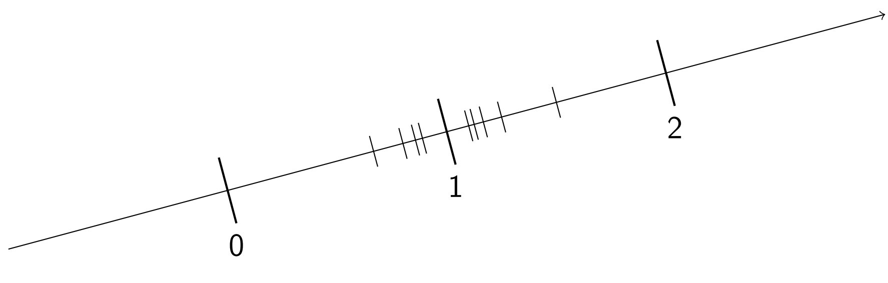
inoffizielles Vorlesungsskript

# Analysis I/II

## Wintersemester 2009 / 2010
## Sommersemester 2009

**15. Juli 2010**


**Prof. Dr. Jürgen Prestin**
Institut für Mathematik
Universität zu Lübeck

# Inhaltsverzeichnis

**Verzeichnis der Sätze und Definitionen** 4

**1 Zahlenfolgen, Grenzwerte, Reihen** 7
1.1 Zahlenfolgen . . . 7
1.1.1 Was ist eine Folge . . . 7
1.1.2 Konvergenz von Folgen . . . 7
1.1.3 Beschränktheit von Folgen . . . 11
1.1.4 Monotonie von Folgen . . . 14
1.1.5 Rechnen mit konvergenten Folgen . . . 14
1.1.6 Konvergenz-Kriterien . . . 20
1.1.7 Beispiele von Folgen-Konvergenz . . . 24
1.1.8 Mengen von Häufungswerten . . . 28
1.2 Undendliche Reihen . . . 29
1.2.1 Definition . . . 29
1.2.2 Beispiele für unendliche Reihen . . . 30
1.2.3 Eulersche Zahl . . . 34
1.2.4 Konvergenz-Kriterien und absolute Konvergenz . . . 36

**2 Stetigkeit und Differenzierbarkeit** 41
2.1 Stetigkeit . . . 41
2.2 Differenzierbarkeit . . . 41
2.2.1 Umkehrfunktionen . . . 41
2.2.2 Beispiele für Umkehrfunktionen . . . 42
2.2.3 schwierige Grenzwerte mit De L'Hospital . . . 47
2.2.4 Abschätzung von Funktionswachstum . . . 51
2.2.5 Approximation durch Polynome . . . 52
2.3 Konvexität . . . 60

**3 Differenzialrechnung in $\mathbb{R}^n$** 68
3.1 Funktionalanalytische Grundlagen . . . 68
3.1.1 Motivation . . . 68
3.1.2 metrische Räume . . . 69
3.1.3 normierte Räume . . . 73
3.1.4 Konvergenz von Vektoren . . . 75
3.1.5 Stetigkeit von Funktionen . . . 76
3.2 Reellwertige Funktionen mehrerer reeller Vektorräume . . . 78
3.2.1 partielle Ableitung . . . 83

3.2.2 totale Ableitung . . . . . 88
3.2.3 Übertragungen ins Multivariable . . . . . 94
3.3 Extremwerte . . . . . 101

**4 Integralrechnung 108**
4.1 Unbestimmte Integrale . . . . . 108
4.1.1 Rechenregeln für unbestimmte Integrale . . . . . 108
4.2 Integraion rationaler Funktionen . . . . . 113
4.2.1 Integration weiterer Funktionenklassen . . . . . 118
4.3 Das bestimmte Integral . . . . . 122
4.3.1 Eigenschaften des Riemann-Integrals . . . . . 123
4.4 Hauptsätze der Integralrechnung . . . . . 125
4.4.1 Substitution und partielle Integration bei bestimmten Integralen . 128
4.5 Uneigentliche Integrale . . . . . 132

**5 Funktionenreihen, Funktionenfolgen 138**
5.1 Gleichmäßige Konvergenz . . . . . 139
5.2 Beweistechnik für punktweise und gleichmäßige Konvergenz . . . . . 142
5.3 Differentiation und Integration . . . . . 143
5.4 Potenzreihen . . . . . 146
5.4.1 Rechnen mit Potenzreihen . . . . . 150
5.5 Fourier-Reihen . . . . . 157
5.5.1 »Periodische Fortsetzung« . . . . . 160
5.5.2 Trigonometrische Polynome . . . . . 160
5.5.3 Fourier-Reihen . . . . . 170

Verzeichnis der Sätze und Definitionen


# Verzeichnis der Sätze und Definitionen


Definition 1.1 . . . . . 7
Definition 1.2 . . . . . 8
Definition 1.3 . . . . . 9
Lemma 1.4 . . . . . 11
Definition 1.5 . . . . . 12
Korollar 1.6 . . . . . 13
Definition 1.7 . . . . . 14
Lemma 1.8 . . . . . 14
Satz 1.9: »Rechenregeln« . . . . . 16
Satz 1.10: »Konvergenzkriterien« . . . . . 20
Definition 1.11: »Cauchy-Folge« . . . . . 21
Satz 1.12: »Konvergenzkriterium von Cauchy« . . . . . 21
Lemma 1.13 . . . . . 27
Definition 1.14 . . . . . 28
Definition 1.15 . . . . . 29
Lemma 1.16 . . . . . 33
Lemma 1.17 . . . . . 34
Lemma 1.18: »Fehlerabschätzung« . . . . . 35
Lemma 1.19 . . . . . 35
Satz 1.20: »Cauchy-Kriterium für Reihen« . . . . . 36
Lemma 1.21 . . . . . 37
Definition 1.22 . . . . . 37
Lemma 1.23 . . . . . 37
Korollar 1.24: »Dreieicksungleichung für *unendlich* viele Summanden« . . . . . 38
Satz 1.25: »Vergleichskriterium« . . . . . 38
Satz 1.26: »Quotientenkriterium« . . . . . 40

Satz 2.21 . . . . . 41
Satz 2.22: »Verallgemeinerung des Mittelwertsatzes« . . . . . 47
Satz 2.23: »Regel von De L'Hospital (1661–1704)« . . . . . 48
Definition 2.24: »Landau-Symbole (1877–1938)« . . . . . 51
Satz 2.25: »Satz von Taylor« . . . . . 52
Korollar 2.26 . . . . . 56
Definition 2.27 . . . . . 60
Lemma 2.28 . . . . . 61
Satz 2.29: »Konvexitätskriterium« . . . . . 61
Korollar 2.30 . . . . . 62

Verzeichnis der Sätze und Definitionen


Korollar 2.31 . . . . . 63
Definition 2.32 . . . . . 64
Lemma 2.33 . . . . . 64
Satz 2.34: »Ungleichung von Jensen (1859–1925)« . . . . . 65

Definition 3.1: »metrische Räume« . . . . . 69
Lemma 3.2 . . . . . 70
Definition 3.3 . . . . . 70
Lemma 3.4 . . . . . 72
Definition 3.5: »Konvergenz in einem metrischen Raum« . . . . . 72
Definition 3.6 . . . . . 73
Lemma 3.7 . . . . . 73
Satz 3.8: »Komponentenweise Konvergenz« . . . . . 75
Definition 3.9 . . . . . 76
Satz 3.10: »$\varepsilon$-$\delta$-Kriterium der Stetigkeit« . . . . . 77
Definition 3.11 . . . . . 78
Satz 3.12 . . . . . 82
Definition 3.13: »Differenzierbarkeit« . . . . . 83
Definition 3.14: »Richtungsableitung« . . . . . 84
Lemma 3.15 . . . . . 85
Definition 3.16: »gradient« . . . . . 85
Satz 3.17: »Satz von Schwarz« . . . . . 86
Definition 3.18: »totale Ableitung« . . . . . 89
Satz 3.19 . . . . . 89
Definition 3.20: »allgemeine Richtungsableitung« . . . . . 90
Satz 3.21 . . . . . 90
Satz 3.22: »Kettenregel« . . . . . 95
Satz 3.23: »Mittelwertsatz für Funktion $f : \mathbb{R}^n \supseteq \Omega \to \mathbb{R}$« . . . . . 97
Satz 3.24: »Taylor-Formel 2. Ordnung« . . . . . 98
Definition 3.25 . . . . . 101
Lemma 3.26: »notwendiges Kriterium für Extremwerte« . . . . . 101
Definition 3.27 . . . . . 102
Definition 3.28 . . . . . 103
Lemma 3.29 . . . . . 104
Satz 3.30: »Hinreichendes Kriterium für Extremwerte« . . . . . 104

Definition 4.1 . . . . . 108
Satz 4.2: »Rechenregeln für unbestimmte Integrale« . . . . . 108
Satz 4.3: »Partialbruchzerlegung« . . . . . 113
Definition 4.4 . . . . . 122
Satz 4.5 . . . . . 123
Satz 4.6: »Mittelwertsatz der Integralrechnung« . . . . . 124
Korollar 4.7 . . . . . 125
Satz 4.8 . . . . . 125
Korollar 4.9 . . . . . 126

Satz 4.10: »Substitution im bestimmten Integral« . . . . . . . . . . . . . . . . . 128
Satz 4.11: »partielle Integration von bestimmten Integralen« . . . . . . . . . . 131
Definition 4.12 . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 132
Definition 4.13 . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 133

Definition 5.1 . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 139
Satz 5.2 . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 141
Satz 5.3 . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 143
Satz 5.4 . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 144
Definition 5.5 . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 146
Satz 5.6 . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 146
Satz 5.7: »Cauchy(1789–1857)-Hadamard(1865–1963)« . . . . . . . . . . . . 148
Lemma 5.8 . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 150
Satz 5.9 . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 151
Satz 5.10 . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 155
Satz 5.11: »Identitätssatz für Potenzreihen« . . . . . . . . . . . . . . . . . . . 156
Satz 5.12 . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 161
Satz 5.13 . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 167
Satz 5.14 . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 168
Satz 5.15 . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 170

1 Zahlenfolgen, Grenzwerte, Reihen


# 1 Zahlenfolgen, Grenzwerte, Reihen

## 1.1 Zahlenfolgen

### 1.1.1 Was ist eine Folge

**Vorlesung I**

**Definition 1.1**

Jede Abbildung (Funktion) $f : \mathbb{N} \to \mathbb{R}$ (oder $\mathbb{C}$) heißt Zahlenfolge oder kurz Folge.

Für $n \in \mathbb{N}$ heißt $a_n = f(n)$ n-tes *Glied* der Folge.

Schreibweise: $a_n$ statt $f(n)$ für einzelne Folgenglieder und $(a_n)$, $(a_n)_{n\in\mathbb{N}}$, $(a_n)_{n=1}^{\infty}$ statt $f$ für die gesamte Folge.

Varianten: Die Abbildung muss aus $\mathbb{N}$ abbilden. Es genügt eine beliebige diskrete Menge. Also zum Beispiel auch $N_0$.

**Beispiel**

$$a_n = \frac{1}{n}, n \in \mathbb{N}, \qquad (e_n) = 1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \ldots$$
$$b_n = 1, n \in \mathbb{N}, \qquad (e_n) = 1, 1, 1, 1, \ldots$$
$$c_n = (-1)^n, n \in \mathbb{N}, (e_n) = -1, 1, -1, 1, \ldots$$
$$d_n = a + (n-1)d, n \in \mathbb{N}, \qquad (e_n) = a, a+d, a+2d, a+3d, \ldots$$
$$e_n = aq^n, n \in \mathbb{N}_0, \qquad (e_n) = a, aq, aq^2, aq^3, \ldots$$

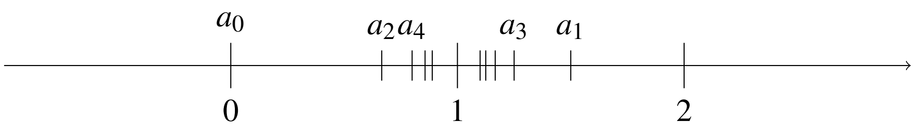
$d_n$ wird *arithmetrische Reihe* und $e_n$ *geometrische Reihe* genannt. [Warum? Verweis auf die Abschnitte]

### 1.1.2 Konvergenz von Folgen

Betrachten wir die Folge $a_n = \frac{(-1)^n}{n}$.

Wir gehen im Folgenden immer von $n \in \mathbb{N}$ aus, wenn diese Angabe weggelassen wird. [Der Hinweis muss in die Definition; Was ist, wenn $\infty$ weggelassen wird?]

Als Folgenglieder ergeben sich $(a_n) = 0, \frac{3}{2}, \frac{2}{3}, \frac{5}{4}, \ldots$, die sich auf dem Zahlenstrahl einzeichnen lassen:




Wir sehen, dass fast alle Folgenglieder »in unmittelbarer Nähe von 1 liegen«. Diese Aussage ist leider etwas schwammig. Wir definieren also die *Konvergenz* als grundelegenden Begriff der Analysis:

### Definition 1.2

Eine Zahlenfolge $(a_n)$ konvergiert gegen den Grenzwert $a \in \mathbb{R}$ (oder $\mathbb{C}$) genau dann wenn

$$\forall \varepsilon > 0 \exists N \in \mathbb{N} \forall n \geq N |a_n - a| < \varepsilon$$

Bezeichnung: $\lim_{n \to \infty} a_n = a$ oder kürzer $a_n n \to \infty \longrightarrow a$

### Beispiel

Betrachten wir mit dieser Definition nun die Folge $a_n = 1 + \frac{(-1)^n}{n}$ können wir beweisen, dass

$$\lim_{n \to \infty} a_n = 1$$

gilt. Für den Beweis benötigen wir die Dreiecksungleichung. Wir erinnern uns an die Weisheit: »Ein Umweg verlängert stets die Reise«. Etwas mathematischer notiert gilt

$$\forall a, b \in \mathbb{R} : |a + b| \leq |a| + |b|$$

Wir sagen auch $|a + b|$ ist der *Abstand* von $a$ und $b$.

**Beweis** Wir geben $\varepsilon > 0$ vor und wählen $N = N(\varepsilon) > \frac{1}{\varepsilon}$. Dann gilt

$$\forall n \geq N(\varepsilon) : |a_n - a| = \left| 1 + \frac{(-1)^n}{n} - 1 \right| = \left| \frac{(-1)^n}{n} \right| = \frac{1}{n}$$

Wir wissen $n > N$ und $N > \frac{1}{\varepsilon}$, also gilt $\frac{1}{n} < \varepsilon$. □

Wir konnten in diesem Beweis also für ein beliebges $\varepsilon$ ein $N$ angeben, sodass

$$\forall n \geq N |a - a_n| < \varepsilon$$

gilt. Da $N$ von $\varepsilon$ abhängt – für jedes $\varepsilon$ muss ein neues $N$ her – schreibt man auch $N(\varepsilon)$. Wir haben jetzt also für jedes vorgegebene $\varepsilon$, dass man uns vorwirft, ein passendes $N$ zur Hand, sodass etwas für alle $n > N$ gilt. Wir betrachten also alle Folgeglieder ab dem Startglied $a_N$. Und für alle diese Glieder, die nach einem bestimmten ersten Glied kommen, ist der Abstand von $a$ (also unserem Grenzwert) und $a_n$ (also dem Wert der Folge an der Stelle $n$) kleiner als $\varepsilon$.

In diesem Beispiel haben wir also gezeigt, dass für jedes gegebene Intervall auf dem Zahlenstrahl um 1 sich immer ein Startindex finden lässt, sodass *alle* Folgenglieder sich nur noch in diesem Intervall befinden. Insbesondere ist hier in »jedes Intervall« auch enthalten, dass wir dieses Intervall beliebig klein wählen können. Und je kleiner wir das Intervall wählen, desto mehr rücken wir der 1 auf die Pelle, ohne sie zu berühren.

Bekommen wir also zum Beispiel $\varepsilon = \frac{1}{100}$ gegeben, berechnen wir wir ein $N > \frac{1}{\varepsilon} = \frac{1}{\frac{1}{100}} = 100$, also zum Beispiel 101. Denn ab $n = 101$ befinden sich alle Folgenglieder im offenen Intervall $(\frac{99}{100}, \frac{101}{100})$.


Und was ist jetzt mit Zahlenfolgen, die nicht konvergent sind? Wir nennen sie divergent. Allerdings gibt es zwei verschiedenen Arten der Divergenz: Entweder die Folge verhält sich so wirr, dass wir auch nicht wissen, was los ist, oder sie wächst über alle Grenzen. Im zweiten Fall nennen wir die Folge bestimmt divergent (denn wir sind nicht mehr völllig desorientiert) und wir können nun analog zum $\varepsilon$ von eben ein $K$ finden, sodass jedes Folgenglied ab einem Startindex $N$ größer ist, als unser $K$. Die Folge wächst also über jede Grenze $K$ hinaus.

### Definition 1.3

a) Zahlenfolgen, die nicht konvergieren, heißen divergent.

b) Divergente Zahlenfolgen heißen *bestimmt divergent* genau dann wenn

$$\forall K > 0 \exists N \forall n \geq N : a_n > K (a_n < K)$$

Bezeichnung: $\lim_{n\to\infty} a_n = \infty(-\infty)$ oder $a_n \underset{n\to\infty}{\longrightarrow} \infty(-\infty)$

c) Jede divergente, nicht bestimmt divergente Teilfolge heißt *unbestimmt divergent*.

d) Konvergente Teilfolgen mit Grenzwert 0 heißen *Nullfolgen*.

e) Sei $n_1 < n_2 < n_3 < \ldots$ eine aufsteigende Indexfolge mit $n_k \in \mathbb{N}$, dann heißt die Folge $(a_{n_k}) = a_{n_1} = a_{n_2} = a_{n_3} = \ldots$ eine *Teilfolge* von $(a_n)$.

Keiner hat gesagt, dass die Folgenglieder der Index**folge** direkt aufeinander folgen müssen. Jede Folge kann eine Indexfolge sein. Und somit müssen die Folgeglieder einer Teilfolge auch nicht direkt aufeinander folgende Elemente einer Folge sein, sondern können wild ausgewählt werden. Nur die Reihenfolge bleibt.

f) Sei $\varepsilon > 0$ und $a \in \mathbb{R}$, dann heißt

$$U_\varepsilon(a) = \{x \mid |x - a| < \varepsilon\}$$

eine Epsilonumgebung von $a$.

Damit haben wir jetzt allen Zahlen im offenen Interval $(a - \varepsilon, a + \varepsilon)$ einen Namen gegeben. Dass schafft keine neue Erkenntnis, sondern spart nur schwammiges Beschreiben. Benutzen wir diese tolle Neuerung in der Definition von konvergenten Folgen, dann konvergergiert eine Folge genau dann wenn

$$\forall \varepsilon > 0 \exists N \forall n \geq N : a_n \in U_\varepsilon(a)$$

Immer noch das Gleiche – drückt nur besser aus, dass sich alle Folgenglieder ab dem Startindex im durch $\varepsilon$ vorgegebenen Intervall um $a$ befinden sollen.

g) Sei $(a_n)$ eine Zahlenfolge und $a \in \mathbb{R}$, dann heißt $a$ ein *Häufungswert* von $(a_n)$ genau dann wenn

$$\forall \varepsilon > 0, N \exists n \geq n : |a_n - a| < \varepsilon$$

Vergleichen wir mit der Definition von Konvergenz, fällt auf, dass jetzt nicht mehr für ein $N$ alle $n > N$ die Bedingung $|a_n - a| < \varepsilon$ gelten soll, sondern es nur noch für alle $N$ ein einziges $n$ geben muss (es drüfen natürlich auch mehr sein), für das die Bedingung erfüllt ist. Es müssen nicht mehr ab einem Startindex alle weiteren Folgenglieder in der $\varepsilon$-Umgebung liegeb, sondern es muss sich für jedes $N$ ein $n$ finden lassen, dass in der Umgebung liegt. Allerdings kann das $N$ beliebig groß gewählt werden und auf diese Weise verlangt die Formel, dass sich unendlich viele Folgenglieder in der Umgebung befinden. Denn für jedes beliebige $N$ findet sich noch ein weiteres $n$. Und danach auch noch für dieses $n$, u.s.w. Es kann also jetzt mehrere Häufungswerte geben, denen sich die Folge annähert – diese heißen dann aber nicht mehr Grenzwert der Folge, da nicht alle Glieder gegen einen Wert streben. [Gibt es einen Satz, dass Häufungswerte und konvergente Teilfogen zusammen hängen?]

h) $(a_n)$ heißt *nach oben (unten) beschränkt*, genau dann wenn

$$\exists C \in \mathbb{R} \forall n \in \mathbb{N} : a_n \leq C (\geq C)$$

$(a_n)$ heißt beschränkt, genau dann wenn

$$\exists C_1, C_2 \in \mathbb{R} \forall n \in \mathbb{N} C_2 \leq a_n \leq C_1$$

Also wenn $(a_n)$ nach oben und nach unten beschränkt ist.

analog dazu reicht es auch, wenn gilt

$$\exists C \in \mathbb{R} \forall n \in \mathbb{N} |a_n| \leq C$$

Denn wir definieren hier nur beschränkt und es ist noch nirgendwo von der kleinsten Schranke die Rede. Wenn es also eine obere und eine untere Schranke gibt, gibt es auch einen Wert, denn der Betrag nie überschreitet.

$C_1$ heißt obere Schranke, $C_2$ heißt untere Schranke.

**Beispiel**

1) $a_n = \frac{1}{n}$ ist Nullfolge mit oberen Schranken $1$, $\Pi$, $e$, $12$ und unteren Schranken $-1$, $0$, $-\pi$, $-e$, $-12$.

   Nullfolge ist $(a_n)$ immer. Die Schranken gelten hingegen nur, weil wir von einem Startindex $n = 1$ bzw. $n \in \mathbb{N}$ ausgehen. Bei anderen diskreten Mengen müssen die Schranken evtl. anders gewählt werden, sie existieren aber auf jeden Fall.

2) Betrachten wir die Folge $a_n = (-1)^n + \frac{1}{n}$, dann gibt es die Teilfolgen $(a_{2n}) = \left(1 + \frac{1}{2n}\right)$ und $(a_{2n+1}) = \left(-1 + \frac{1}{2n+1}\right)$. Es gilt $\lim_{n\to\infty} a_{2n} = 1$ und $\lim_{n\to\infty} a_{2n+1} = -1$. Die Folge $(a_n)$ hat also die konvergenten Teilfogen $(a_{2n})$ und $(a_{2n+1})$ und damit als Häufunsgwerte die Grenzwerte der konvergenten Teilfolgen, also $-1$ und $1$.

### 1.1.3 Beschränktheit von Folgen

**Vorlesung II**

Nachdem wir am Ende der letzten Gewalt-Definition noch eben schnell die Beschränktheit von Folgen mit abgehandelt haben, gehen wir die Sache jetzt etwas genauer an. Denn schnell stellen wir fest, dass irgendeine Schranke in der Praxis wenig hilfreich ist. Wenn ich $\Pi$ oder $e$ oder 2 als obere Schranke nehmen kann, verwirrt das reichlich und ist auch für Klausuren nicht zu gebrauchen...

**Lemma 1.4**

a) Jede nach oben beschränkte Zahlenfolge besitzt eine kleinste obere Schranke.

b) Jede nach unten beschränkte Zahlenfolge besitzt eine kleinste untere Schranke.

**Beweis** ergibt sich unmittelbar aus der Definition von oberer und unterer Schranke. □

Mit dieser Erkenntnis können wir das Kind jetzt benennen. Außerdem lösen wir dabei gleich das Problem, dass die kleinste Schranke nicht immer ein Element der Folge sein muss, dies aber durchaus sein kann. Auf der anderen Seite kann es ein kleinste / größtes Element der Folge geben, zum Beispiel für $n = 1$, wenn die Folge monoton fällt.

## Definition 1.5

a) $(a_n)$ sei nach oben beschränkt. Die kleinste obere Schranke heißt *Supremum von* $(a_n)$. Schreibweise: $\sup\limits_{n\in\mathbb{N}} a_n$

   Wollen wir nur eine Teilfolge von $(a_n)$ betrachten, nehmen wir uns eine Indexmenge $M \subset \mathbb{N}$ her, die alle Indizes der Teilfolge enthält, und schreiben $\sup\limits_{n\in M} a_n$

b) $(a_n)$ sei nach oben beschränkt. Die größte obere Schranke heißt *Infimum von* $(a_n)$. Schreibweise: $\inf\limits_{n\in\mathbb{N}} a_n$

c) Falls ein $n_0 \in \mathbb{N}$ existiert, sodass

$$\forall n \in \mathbb{N} : a_n \leq a_{n_0}$$

   gilt, dann hat $(a_n)$ die *Maximalstelle* $n_0$ mit dem *Maximum* $a_{n_0}$. Schreibweise: $\max\limits_{n\in\mathbb{N}} a_n$

   In der Bedingung muss unbedingt $\leq$ stehen, denn die Bedingung soll für alle $n \in \mathbb{N}$ und somit insbesondere auch für $n_0$ gelten. Und für $n_0$ kann nur $a_{n_0} \leq a_{n_0}$ gelten.

   $n_0$ wird in der Mathematik gerne für eine bestimmte Stelle in der Folge (oder später in der Funktion) verwendet. Der Index 0 gibt nur an, dass es sich um eine feste Stelle mit einer besonderen Eigenschaft handelt, die man im Folgenden meistens näher betrachten will. Nicht mit Nullstellen verwechseln!

   [Wo wird eigentlich der Unterschied von $(a_n)$ und $a_n$ explizit erklärt?]

d) Falles ein $n_0 \in \mathbb{N}$ existiert, sodass

$$\forall n \in \mathbb{N} : a_n \geq a_{n_0}$$

   gilt, dann hat $(a_n)$ die Maximalstelle $n_0$ mit dem Maximum $a_{n_0}$. Schreibweise: $\min\limits_{n\in\mathbb{N}} a_n$

Die Defintionen schließen alle den Fall ein, dass die beschriebene Eigenschaft nicht existiert. a) und b) lösen dass Problem durch die Einleitung »Sei $(a_n)$ beschränkt …«, womit Supremum und Infimum nur für beschränkte Folgen definiert sind, und c) und) schreiben explizit »Falls …«. Also auch Maximum und Maximalstelle und Minimum und Minimalstelle sind nur definiert, falls sie exitsieren.

### Beispiel

Wir betrachten die Folge $(a_n) = \left(\frac{1}{n}\right)$. Es gilt

$$\sup_{n\in\mathbb{N}} a_n = \max_{n\in\mathbb{N}} a_n = 1$$

$$\inf_{n\in\mathbb{N}} a_n = 0$$

Die Folge kann im schlimmsten Fall den Wert 1 annehmen, wenn wir für $n = 1$ einsetzen. In allen anderen Fällen $n > 1$, sind die Werte der Folge kleiner als 1. Also ist 1 das Maximum der Folge. Damit ist 1 auch automatisch die kleinste obere Schranke, denn kein Element der Folge kann größer sein als das Maximum. Wir betrachten diesen Zusammenhang im Folgenden Korollar näher. Die Folge strebt gegen 0, ist dabei aber immer positiv. Damit ist die kleinste untere Schranke 0. Diesmal wird die 0 selber aber von der Folge nie erreicht, da 0 der Grenzwert der Folge ist. Aus diesem Grunde kann kein Minimum existieren, da es für jedes Folgenglied $a_n$ ein kleineres Folgenglied gibt, zum Beispiel $a_{n+1}$. Das Minimum der Folge $(a_n)$ ist also nicht definiert!

### Korollar 1.6

a) Besitzt $(a_n)$ ein Maximum, so gilt

$$\sup_{n\in\mathbb{N}} a_n = \max_{n\in\mathbb{N}} a_n$$

b) Besitzt $(a_n)$ ein Minimum, so gilt

$$\inf_{n\in\mathbb{N}} a_n = \min_{n\in\mathbb{N}} a_n$$

**Beweis** Der Beweis folgt unmittelbar aus der Definition. Existiert ein Maximum bzw. Minimum, so erfüllt dieses auch unmittelbar die Bedingungen für das Suppremum bzw. Infimum. □

Wir sehen also, dass das Maximum, wenn es existiert, immer auch die kleinste obere Schranke darstellt. Gleichzeitig ist das Minimum immer auch die größte untere Schranke. Implizit steckt in dieser Folgerung auch die Erkenntnis, dass Infimum bzw. Supremum immer existieren, wenn Minimum bzw. Maximum existieren.

### 1.1.4 Monotonie von Folgen

**Definition 1.7**

Sei $M = [m_1, m_2]$ ein offenes Intervall und $(a_n)$ eine Folge.

a) $(a_n)$ heißt *(streng) monoton wachsend* auf $M$, falls

$$\forall p, q \in M \cap \mathbb{N} : p < q \Rightarrow a_p \leq a_q (a_p < a_q)$$

b) $(a_n)$ heißt *(streng) monoton fallend* auf $M$, falls

$$\forall p, q \in M \cap \mathbb{N} : p < q \Rightarrow a_p \geq a_q (q_p > a_q)$$

Die tolle Konstruktion $M \cap \mathbb{N}$ verdanken wir der Tatsache, dass $M$ ein Intervall ist und Intervalle immer auf den reellen Zahlen [oder den komplexen?] definiert sind. Unsere Folgen bilden aber immer von $\mathbb{N}$ auf eine Zhal ab. Daher müssen wir das Intervall mit $\mathbb{N}$ schneiden, um nur die ganzen Zahlen aus dem Intervall zu fischen.

Eine Folge ist also genau dann Monoton, wenn ich mir zwei beliebige Zahlen aus der Folge greifen kann, wobei beliebig an dieser Stelle insbesondere belibig dicht zusammen heißt, und die Zahl, die in der Folge weiter hinten steht immer kleiner bzw. immer größer ist. Die Zahlen werden also entsprechend ihrer Position entweder immer größer, oder immer kleiner.

Sprechen wir von einer streng monotonen Folge, muss tatsächlich jedes Folgenglied anders sein. In einer monotonen Folge dürfen aufeinader folgenden Glieder in der Folge auch den gleichen Wert haben.

**Beispiel**

- $\left(\frac{1}{n}\right)$ ist streng monoton fallend auf $\mathbb{N}$
- $\left(-\frac{1}{n}\right)$ ist streng monoton wachsend auf $\mathbb{N}$

  Wir sehen also, wenn $(a_n)$ monoton wächst, fällt $(-a_n)$ monoton.
- $(1)$ ist monoton wachsend und monoton fallend.

  Aber streng ist hier nichts! Schließlich sind die Werte aller Folgenglieder gleich.

### 1.1.5 Rechnen mit konvergenten Folgen

Nach diesem kurzen Abstecher in die Monotonie betrachten wir nun noch einmal konvergente Folgen und halten einige einfache Erkenntnisse fest, die für alle konvergenten Folgen gelten.

### Lemma 1.8

Gegeben sei eine konvergente Zahlenfolge $(a_n)$. Dann gilt

a) $\lim\limits_{n\to\infty} a_n = a \wedge \lim\limits_{n\to\infty} a_n = b \Rightarrow a = b$

Es kann nur einen geben I: Wenn $(a_n)$ eine konvergente Folge ist, dann gibt es nur einen Grenzwert.

b) Es existiert ein $C \in \mathbb{R}$, sodass

$$\forall n \in \mathbb{N} : |a_n| < C$$

Jede konvergente Zahlenfolge ist beschränkt.

c) $\lim\limits_{n\to\infty} a_n = a \Rightarrow \forall (n_k) : \lim\limits_{n\to\infty} a_{n_k} = 0$

Es kann nur einen geben II: Wenn eine Folge konvergiert, dann konvergieren auch alle Teilfolgen (hier realisiert durch die Indexfolge $(n_k)$) dieser Folge gegen dengleichen Grenzwert.

d) Wenn $\lim\limits_{n\to\infty} a_n = a$ gilt, dann ist $a$ der einzige Häufungswert der Folge.

Wenn $\lim\limits_{n\to\infty} a_n = a$ gilt, ist $a$ der Grenzwert der Folge $(a_n)$.

Es kann nur einen geben III: Wenn einen Folge konvergiert, dann hat sie nur einen Häufungswert und dieser ist der Grenzwert.

**Beweis** a) Zum Zwecke des indirekten Beweises nehmen wir $a \neq b$ an. Wir wählen nun $\varepsilon = \left|\frac{b-a}{2}\right|$. Entsprechend der Definition der Konvergenz (siehe Definition 1.2 auf Seite 8) gilt dann

$$\exists N_a(\varepsilon) : \forall n \geq N_a(\varepsilon) : |a_n - a| < \varepsilon$$
$$\exists N_b(\varepsilon) : \forall n \geq N_b(\varepsilon) : |a_n - b| < \varepsilon$$

Für alle $n \geq \max\{N_a(\varepsilon), N_b(\varepsilon)\}$ gilt nun

$$2\varepsilon = 2\left|\frac{b-a}{2}\right| = |b-a| = |b - a_n + a_n - a| \overset{*}{\leq} \underbrace{|a_n - b|}_{<\varepsilon} + \underbrace{|a_n - a|}_{<\varepsilon} < 2\varepsilon \quad ↯$$

Wenn $n$ größer ist als das Maximum von $N_a(\varepsilon)$ und $N_b(\varepsilon)$ gelten *beide* obigen Aussagen und können zu einer neuen zusammengesetz werden. Wir addieren quasi beide Gleichungen.

Mit der Aussage $2\varepsilon < 2\varepsilon$ haben wir den gewünschten Widerspruch gefunden und damit die Aussage gezeigt. Wie in jedem guten Beweis mit Beträgen wurde an der Stelle $*$ die Dreiecksungleichung verwendet.

b) Es sei $\varepsilon = 1$. Dann gibt es aufgrund der Konvergenz der Folge ein $N(\varepsilon = 1)$, sodass für alle $n \geq N(1)$ die Aussage $|a_n - a| < 1$ gilt. Mit der Dreiecksungleichung folgt daraus $\forall n > N(1) : |a_n| \leq |a| + 1$ und damit $\forall n \in \mathbb{N} : |a_n| \leq \max\{|a_1|, |a_2|, \ldots, |a_{N(1)-1}|, |a| + 1\} = C$. Somit ist $C$ die Schranke. [Was macht der hier? Was soll die Menge sein?]

□

**Satz 1.9** Rechenregeln

Es seien $(a_n)$ und $(b_n)$ zwei Folgen mit $\lim\limits_{n\to\infty} a_n = a$ und $\lim\limits_{n\to\infty} b_n = b$. Dann gilt

a) $\lim\limits_{n\to\infty} (a_n \pm b_n) = a \pm b$

   Als Vorraussetzung müssen hier nur die Folgen $(a_n)$ und $(b_n)$ konvergieren, also gegen den Grenzwert $a$ bzw. $b$ streben. Dann ist auch die Folge $(a_n \pm b_n)$ konvergent. Wir müssen also die Konvergenz der Folge $(a_n \pm b_n)$ nicht erst seperat nachweisen, sondern sie ergibt sich ebenfalls aus diesem Satz.

b) $\lim\limits_{n\to\infty} |a_n| = |a|$

   Auch hier erhalten wir die Aussage, dass eine Betragsfolge konvergiert, wenn die Folge ohne Betragsstriche ebenfalls konvergiert. Umgekehrt gilt das auf keinen Fall. Man betrachte die Folge $(|(-1)^n|) = (1)$ und $((-1)^n)$.

c) $\lim\limits_{n\to\infty} a_n b_n = ab$

   Im Spezialfall $\forall n \in \mathbb{N} : a_n = c$ gilt damit auch für ein festes $c \in \mathbb{R}$ und eine beliebige Folge $(x_n)$

$$\lim_{n\to\infty} c x_n = c \lim_{n\to\infty} x_n$$

d) Ist der Grenzwert von $(a_n)$ nicht 0, dann gilt

$$\exists N \in \mathbb{N} \forall n \geq N : a_n \neq 0$$

   und

$$\lim_{n\to\infty} \frac{1}{a_n} = \frac{1}{a}$$

e) Gilt $\forall n \in \mathbb{N} : a_n \geq 0$, hat die Folge $(a_n)$ also keine negativen Glieder, dann gilt

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

   Ich mache noch einmal darauf aufmerksam: Auch hier erhalten wir die Aussage, dass eine Wurzelfolge konvergiert, wenn alle Folgenglieder der Folge unter der Wurzel positiv sind und die Folge unter der Wurzel konvergiert.

**Vorlesung III**

**Beweis** a) Es sei ein $\varepsilon > 0$ gegeben. Wir definieren dann $N_1$ durch

$$\forall n \geq N_1 : |a_n - a| < \frac{\varepsilon}{2}$$

$N_2$ durch

$$\forall n \geq N_2 : |b_n - b| < \frac{\varepsilon}{2}$$

und es sei schließlich $N = \max\{N_1, N_2\}$.

$N$ ist jetzt also so gewählt, dass sich ab Folgenindex $N$ alle Folgenglieder von $(a_n)$ in einer $\varepsilon$-Umgebung von $a$ befinden und alle Folgenglieder von $(b_n)$ in einer $\varepsilon$-Umgebung von $b$ befinden. Nun addieren wir die Aussagen, die nach der Definition der Konvergenzen beider Folgen gelten und erhalten für unsere neue Folge ebenfalls eine Konvergenz.

Dann gilt

$$\forall n \geq N : |(a_n \pm b_n) - (a \pm b)| \leq \underbrace{|a_n - a|}_{<\frac{\varepsilon}{2}} + \underbrace{|b_n - b|}_{<\frac{\varepsilon}{2}} < \varepsilon$$

b) Wir beweisen die Aussage mit der Dreiecksungleichung

$$0 \leq ||a_n| - |a|| \leq |a_n - a| < \varepsilon$$

Wenn der Abstand von $a_n$ und $\varepsilon$ kleiner ist als $a$ (Teil der Bedingung für die Konvergenz von $(a_n)$ gegen $a$), dann ist auch auf jeden Fall immer dann der Abstand von $|a_n|$ und $|a|$ kleiner als $\varepsilon$.

c) Es sei ein $\varepsilon > 0$ gegeben. Dann sei $N_1$ definiert durch

$$\forall n \geq N_1 : |a_n - a| < \frac{\varepsilon}{2(|b|+1)}$$

und $k$ definiert durch

$$\forall n \in \mathbb{N} : |a_n| < k$$

Diese Definition von $k$ ist möglich denn jede konvergente Zahlenfolge beschränkt (siehe Lemma 1.8 auf Seite 15). $k$ stellt also gerade eine untere und obere Schranke der Folge da.

Weiter sei $N_2$ definiert durch

$$\forall n \geq N_2 : |b_n - b| < \frac{\varepsilon}{2k}$$

und es sei schließlich $N = \max\{N_1, N_2\}$.

Dann gilt

$$\begin{aligned}\forall n \geq N : &|a_n a_b - ab| = |a_n b_n - a_n b + a_n b - ab| \leq |a_n b_n - a_n b| + |a_n b - ab| \\ &\leq |a_n||b_n - b| + |b||a_n - a| \overset{\dagger}{<} \underbrace{k\frac{\varepsilon}{2k}}_{=\frac{\varepsilon}{2}} + \underbrace{|b|\frac{\varepsilon}{2|b|+1}}_{\leq\frac{\varepsilon}{2}} \leq \varepsilon\end{aligned}$$

An der mit † gekennzeichneten Stelle wurden alle oben definierten Vorraussetzungen verwendet. Die beiden Ungleichungen vorher ergeben sich nur aus der Dreiecksungleichung und dienen hier nur der Vorbereitung.

Der eigentlich Trick dieses Beweises liegt – wie bereits im Beweis von a) – in der cleveren Definition von $N_1$ und $N_2$. Denn da sich nach der Definition für jedes $\varepsilon$ eine $N$ finden lässt, können wir auch $\varepsilon = \frac{\varepsilon}{2(|b|+1)}$ setzen, wobei $\varepsilon$ jetzt eine beliebige, aber feste Zahl ist, die noch keine nähere Bedeutung hat, und dann $N_1$ als das $N$ defininieren, was sich durch diese Wahl von $\varepsilon$ ergibt. Setzen wir nun an der markierten Stelle in den Ausdruck ein, für den wir eigentlich etwas zeigen wollen, dann fällt alles das weg, was uns irgendwie behindern könnte, und wir erhalten die gewünschte Aussage: Für ein beliebiges $\varepsilon$ haben wir ein $N$ gebastelt, sodass für alle $n \geq N$ die Aussage $|a_n a_b - ab| < \varepsilon$ gilt.

d) Es sei $N_1$ definiert durch

$$\forall n \geq N_1 : \frac{1}{2}|a| < |a_n| < \frac{3}{2}|a|$$

Da die Folge $(a_n)$ gegen $a$ konvergiert nähert sich $a_n$ irgendwann soweit $a$ an, dass es sich in dieser Umgebung befinden muss.

Für diese $a_n$ gilt $|aa_n| > \frac{a^2}{2}$

Weiter sei ein $\varepsilon > 0$ gegeben und $N_2$ definiert durch

$$\forall n \geq N_2 : |a_n - a| < \frac{a^2}{2}\varepsilon$$

und schließlich sei $N = \max\{N_1, N_2\}$. Dann gilt

$$\left|\forall n \geq N : \frac{1}{a_n} - \frac{1}{a}\right| = \left|\frac{a - a_n}{aa_n}\right| < \frac{a^2}{2}\varepsilon\frac{2}{a^2} = \varepsilon$$

e) Wenn für alle $n$ $a_n \geq 0$ gilt, dann muss auch $a \geq 0$ gelten.

Betrachten wir den Fall $a = 0$. Es sei ein $\varepsilon > 0$ gegeben und $N$ definiert durch

$$\forall n \geq N : a_n - 0 < \varepsilon^2$$

dann folgt daraus

$$\forall n \geq N : \sqrt{a_n - 0} < \sqrt{\varepsilon^2} = \varepsilon$$

Im zweiten Fall gelte $a > 0$ und sei auch hier ein $\varepsilon > 0$ gegeben und ein $N$ definiert durch

$$\forall n \geq N : |a_n - a| < \varepsilon\sqrt{a}$$

Dann gilt

$$\begin{aligned}\forall n \geq N : &|\sqrt{a_n} - \sqrt{a}| = |\sqrt{a_n} - \sqrt{a}| \left|\frac{\sqrt{a_n} + \sqrt{a}}{\sqrt{a_n} - \sqrt{a}}\right| \\ &\leq \frac{|a_n - a|}{\sqrt{a}} < \frac{\varepsilon\sqrt{a}}{\sqrt{a}} = \varepsilon\end{aligned}$$

[Ist der zweite Fall so ganz stimmig? Ich weiß nicht so recht. . . ]

□

**Bemerkung** 1. Umkehrungen zu a) und c) gelten nicht!

Zu b) übrigens auch nicht, aber das liegt daran, dass es keine Umkehrfunktion (siehe Bemerkung 4) zu $f(x) = |x|$ gibt. Denn zum Beispiel $|x| = 2$ gilt für $x = 2$ oder $x = -2$. d) enthält in sich selbst auch seine Umkehung und auch e) funktioniert in der Umkehrung, denn für $f(x) = \frac{1}{x}$ und $f(x) = \sqrt{x}$ lassen sich Umkehfunktionen finden. (Man beachte für $\sqrt{x}$ die Beschränkung auf nichtnegative Zahlen!)

2. Aus a) und c) ergibt sich die Linearität der Grenzwertbildung, denn es gilt

$$\forall \alpha, \beta \in \mathbb{R} : \lim_{n\to\infty} (\alpha a_n + \beta b_n) = \alpha \lim_{n\to\infty} a_n + \beta \lim_{n\to\infty} b_n$$

3. Analog zum Satz lassen sich für $a_n \xrightarrow[n\to\infty]{} \infty$, $b_n \xrightarrow[n\to\infty]{} \infty$ und $c_n \xrightarrow[n\to\infty]{} c$ zum Beispiel beweisen

   - $a_n + b_n \xrightarrow[n\to\infty]{} \infty$
   - $a_n + c_n \xrightarrow[n\to\infty]{} \infty$
   - $a_n c_n \xrightarrow[n\to\infty]{} \infty$ für $c > 0$

4. b) mit $f(x) = |x|$, d) mit $f(x) = \frac{1}{x}$ und e) mit $f(x) = \sqrt{x}$ sind Spezialfälle der Aussage [Wieso steht im Script »Frage«? Gilt das nicht?]

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

[Fass mal schlau zusammen, was man jetzt alles darf!]

### 1.1.6 Konvergenz-Kriterien

Bisher haben wir die Konvergenz einer Folge immer direkt mit der Definition der Konvergenz gezeigt. Und wir haben gesehen, dass das meistens ein sehr unhandlicher Beweis ist. Daher werden wir uns im Folgenden einige Techniken aneignen, mit denen man einfacher über die Konvergenz (allerdings nur ja oder nein – es ergibt sich nicht mehr direkt auch der Grenzwert) entscheiden kann.

**Satz 1.10** Konvergenzkriterien

a) *Vergleichskriterium*: Es seien drei Zahlenfolgen $(a_n)$, $(b_n)$ und $(c_n)$ gegeben, sodass gelte

$$\exists N \in \mathbb{N} \forall n \geq N : b_n \leq a_n \leq c_n$$

und es gelte

$$\lim_{n\to\infty} b_n = \lim_{n\to\infty} c_n = a$$

dann gilt auch

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

Sehr anschaulich ist hier das Bild des Einschnürens: Die Folgen $(b_n)$ und $(c_n)$ kommen der Folge $(a_n)$ immer näher, sodass man sich am Ende nur beim gleichen Grenzwert treffen muss. Allerdings gilt wie immer bei der Betrachtung von Grenzwerten von Folgen: Was interessieren mich die ersten Folgenglieder – es muss nur irgendwann klappen! Es kann also durchaus sein, dass das Einschnüren erst ab einem Startindex funktioniert und die Folgenglieder vorher alle wiöld durcheinander tanzen.

b) *Monotoniekriterium*: Jede reele, monotone und beschränkte Zahlenfolge ist konvergent.

Das ist quasi die Umkehrung von »Jede konvergente Zahlenfolge ist beschränkt« (siehe Lemma 1.8 auf Seite 15). Allerdings reicht es nicht aus, dass die Zahlenfolge beschränkt ist. Sie muss auch noch monoton sein, denn sonst muss die Zahlenfolge nicht auf *einen* Häufungswert zusammenlaufen. Man betrachte $((-1)^n)$ – die Zahlenfolge ist beschränkt, denn die Glieder werden nie größer als $|1|$.

**Beweis** a) Für ein beliebiges $\varepsilon > 0$ lässt sich ein $N_1 > N$ finden, sodass gilt

$$\forall n \geq N_1 : a - \varepsilon < b_n \leq a_n \leq c_n < a + \varepsilon$$

Nach Vorraussetzung konvergieren die Folgen $(b_n)$ und $(c_n)$ gegen den Grenzwert $a$. Also befinden sich ab einem Startindex $N_1$ alle Folgenglieder von $(b_n)$ und $(c_n)$ innerhalb einer $\varepsilon$-Umgebung von $a$. Außerdem gilt nach Vorraussetzung, dass die Folgen $(b_n)$ und $(c_n)$ die Folge $(a_n)$ ab dem Startindex $N$ einschnüren – also auch ab $N_1 > N$. Damit muss sich also auch $c_n$ innerhalb der $\varepsilon$-Umgebung befinden. Und wenn sich zu jedem $\varepsilon$ ein $N$ finden lässt, sodass sich $c_n$ ab Startindex $N$ nur noch in der $\varepsilon$-Umgebung befindet, dann ist die Definition von Konvergenz erfüllt.

b) Ohne Beschränkung der Allgemeinheit sei $(a_n)$ monoton wachsend.

Wir können also im Folgenden auch jeweils monoton fallend einsetzen, müssen den Beweis dann nur entsprechend anpassen, kommen aber zum glichen Ziel. Daher die tolle Formulierung »Ohne Beschränkung der Allgemeinheit…«.

Sei weiter $a$ definiert durch $a = \sup_{n\in\mathbb{N}} a_n$ und ein $\varepsilon > 0$ gegeben. Dann ist $a - \varepsilon$ keine obere Schranke von $(a_n)$, denn es existiert ein $N$, sodass gilt $a - \varepsilon < a_N \leq a$.

Da die Folge $(a_n)$ monoton wachsend ist, gilt

$$\forall n \geq N : a - \varepsilon < a_N \leq a$$

und damit gilt die Konvergenz gegen $a$.

Wir nehmen uns also die kleinste obere Schranke von $(a_n)$ her und nennen sie $a$. Für jedes $\varepsilon$ ist $a - \varepsilon$ nun keine obere Schranke mehr, denn sonst wäre $a$ nicht die kleinste obere Schranke. Außerdem ist die Folge monoton wachsend, also sie muss sie irgendwann auch über $a - \varepsilon$ hinauzswachsen, denn $a - \varepsilon$ ist ja keine obere Schranke. Da diese Überlegung für jedes $\varepsilon$ – also auch jedes noch so kleine $\varepsilon$ – funktioniert und die Folge trotzdem nie größer werden kann als $a$, denn $a$ ist eine obere Schranke, muss die Folge gegen $a$ konvergieren.

□

**Definition 1.11** Cauchy-Folge

Eine Zahlenfolge $(a_n)$ heißt *Cauchy-Folge (CF)*, wenn für jedes $\varepsilon > 0$ ein $N$ existiert, sodass gilt

$$\forall m,n \geq N : |a_m - a_n| < \varepsilon$$

**Satz 1.12** Konvergenzkriterium von Cauchy

Eine Zahlenfolge ist genau dann konvergent, wenn sie eine Cauchy-Folge ist.

Die eigentliche Macht dieses Konvergenz-Kriteriums verbirgt sich in der Definition der Cauchy-Folge. Nimmt man Definition und Satz zusammen, kommt man zu der Aussage, dass für Konvergenz »nur noch«

$$\forall \varepsilon > 0 \exists N \in \mathbb{N} \forall m,n \geq N : |a_m - a_n| < \varepsilon$$

zu zeigen ist und nicht mehr

$$\forall \varepsilon > 0 \exists N \in \mathbb{N} \forall n \geq N : |a_n - a| < \varepsilon$$

Es reicht also für die Konvergenz aus, wenn sich alle Folgenglieder ab einem Startindex gegenseitig immer weiter annähern. Es muss nicht mehr der Abstand zu einem festen Grenzwert betrachtet werden. Besonders nützlich wird dieser Unterschied bei der Betrachtung von Reihen werden (siehe Satz 1.20 auf Seite 36).

**Vorlesung IV**

**Beweis** Wir zeigen zwei Richtungen und beginnen mit der Hinrichtung (Jede konvergente Zahlenfolge ist eine Cauchy-Folge). Da die Folge $(a_n)$ konvergent ist, gibt es zu jedem $\varepsilon > 0$ ein $N$, sodass gilt

$$\forall n, m \geq N : |a_n - a| < \frac{\varepsilon}{2} \wedge |a_m - a| < \frac{\varepsilon}{2}$$

Daraus folgt

$$\forall m, n > N : |a_m - a_n| \leq |a_m - a| + |a_n - a| < \varepsilon$$

und damit ist $(a_n)$ eine Cauchy-Folge.

Für die Rückrichtung (Jede Cauchy-Folge ist konvergent) sind zunächst einige Vorüberlegungen nötig.

1. Jede Cauchy-Folge ist beschränkt.

   Sei dazu $\varepsilon > 0$ gegeben und $N$ definiert durch

   $$\forall n, m \geq N : |a_n - a_m| < \varepsilon$$

   Es bleibt zu zeigen, dass eine beschränkte Cauchy-Folge konvergiert.

2. *Satz von Bolzano-Weierstraß*

   Jede beschränkte Folge (allgemeiner jede unendliche beschränkte Menge $M \subset \mathbb{R}$) besitzt einen Häufungspunkt.

   **Beweis** Wir verwenden die Methode der Intervallschachtelung.

   Betrachten wir die folgende Intervallkonstruktion

   $$I_1 \supset I_2 \supset I_3 \supset I_4 \supset \dots$$

   Eine unendliche Menge von abgeschlossenen Intervallen jeweils mit der halben Länge vom vorigen Intervall. Jedes Intervall besitzt unendlich viele Glieder $a_j \in I_n$. Genau ein $a \in \mathbb{R}$ liegt in allen Intervallen (Axiom!) □

Da wir eine *unendliche* Abfolge von Intervallen betrachten, deren Länge sich jeweils halbiert, befindet sich am Grenzwert zum unendlichsten Intervall nur noch der Wert $a$ im Intervall. Aus dieser axiomatischen Aussage – es ist eine Annahme auf der unsere Mathematik basiert – folgt, dass jede unendliche Zahlenmenge, die beschränkt ist, mindestens einen Häufungspunkt besitzen muss. [Sinn?]

Es bleibt zu zeigen, dass eine Cauchy-Folge mit einem Häufungspunkt konvergiert.

3. Eine Zahlenfolge mit einem Häufungspunkt $a$ besitzt einen gegen $a$ konvergierende Teilfolge. [Steht das nicht so schon in der Definition von Häufungspunkt?]

   Es existiert eine Indexfolge $(n_1)$, sodass gilt

   $$|a_{n_1} - a| < 1$$

   Entsprechend existiert auch eine Indexfolge $(n_2)$, sodass gilt

   $$|a_{n_2} - a| < \frac{1}{2}$$

   Setzt man das entsprechend fort, dann gilt allgemein für eine Teilfolge $(a_{n_j})$

   $$|a_{n_j} - a| < 2^{-j}$$

   anders ausgedrückt konvergiert die Folge $(a_{n_j})_j$ gegen $a$.

   Es bleibt zu zeigen, dass eine Cauchy-Folge mit einer konvergenten Teilfolge konvergiert.

4. Besitzt eine Cauchy-Folge eine konvergente Teilfolge mit einem Grenzwert $a$, erwzingt das die Konvergenz der Cauchy-Folge. Denn sei ein $e\varepsilon > 0$ gegeben, dann gilt

   $$\exists N \in \mathbb{N} \forall n,m \geq N : |a_n - a_m| < \frac{\varepsilon}{2}$$

   Wählen wir $m = n_j$, dann ergibt sich

   $$\exists N \in \mathbb{N} \forall n, n_j \geq N : \quad \underbrace{|a_n - a|}_{\text{unabhängig von } j} \overset{\ddagger}{\leq} |a_n - a_{n_j}| + |a_{n_j} - a| \overset{\S}{<} \frac{\varepsilon}{2} + \qquad \underbrace{|a_{n_j} - a|}_{\text{wird für } n \to \infty \text{ kleiner als } \frac{\varepsilon}{2}}$$

   An der Stelle ‡ addieren und subtrahieren wir $a_{n_j}$, was erstmal nichts verändert, und wenden dann die Dreiecksungleichung an. Also der gleiche Trick wie jedes mal...

   An der Stelle § schätzen wir $|a_n - a_{n_j}|$ mit $\frac{\varepsilon}{2}$ ab – was erlaubt ist, da wir es mit einer Cauchy-Folge zu tun haben und die entsprechende Ungleichung aus der Zeile darüber übernehmen können. $|a_{n_j} - a|$ kann stehen bleiben, denn für $j$ gegen $\infty$ strebt der Ausdruck gegen 0 – denn nach 3. besitzt unsere Cauchy-Folge die gegen $a$ konvergierende Teilfolge $(a_{n_j})_j$.

□

### 1.1.7 Beispiele von Folgen-Konvergenz

**Beispiel**

1. Betrachten wir die Folge $\sqrt{\frac{n^2+1}{2n^2-1}}$ und wenden die Rechenregeln an:

$$\lim_{n\to\infty}\sqrt{\frac{n^2+1}{2n^2-1}} \overset{¶}{=} \lim_{n\to\infty}\sqrt{\frac{1+\frac{1}{n^2}}{2-\frac{1}{n^2}}} = \sqrt{\frac{\lim_{n\to\infty}\left(1+\frac{1}{n^2}\right)}{\lim_{n\to\infty}\left(2-\frac{1}{n^2}\right)}} = \sqrt{\frac{1}{2}} = \frac{\sqrt{2}}{2}$$

An der Stelle ¶ haben wir den Bruch mit $\frac{1}{n^2}$ erweitert. Diese Methode den Bruch durch das $n$ mit dem größten Exponenten zu teilen ist auch ohne Limes gültig und hilft bei der Betrachtung von Grenzwerten oft weiter. Denn anschließend kennen wir den Grenzwert von $\frac{1}{n^2}$, vorher kannten wir den Grenzwert von $n^2$ nicht.

2. Wir betrachten die Folge $(a^n)_{n\in\mathbb{N}}$ mit $a\in\mathbb{R}$. Es gilt

- für $a>1$: Die Folge ist bestimmt divergent gegen $\infty$.

  Denn um von einem Folgenglied zum nächsten zu kommen, multiplizieren wir den Wert der Folge an der alten Stelle mit $a$. Und da $a>1$ gilt, sind alle Werte der Folge positiv und der neuen Wert der Folge ist stets größer als der alte. Die Folge ist also monoton steigend.

- für $a\leq -1$: Die Folge ist unbestimmt divergent.

  Die Argumentation ist hier sehr ähnlich. Allerdings wechselt die Folge dabei jeweils das Vorzeichen, denn wegen $a<0$ gilt $a^n>0$ für gerade $n$ und $a^n<0$ für ungerade $n$. Man spricht dann von einer alternierenden Folge.

- für $a=1$: Die Folge konvergiert gegen 1.

  Wer hätte das erwartet.

$$\forall n\in\mathbb{N}: 1^n = 1$$

- für $|a|<1$: Die Folge ist eine Nullfolge.

  Wir multiplizieren den Wert der Folge immer wieder mit einem a, für das $|a<1|$ gilt. Also wird der Wert der Folge immer kleiner. Gilt $a>0$ ist die Folge monoton fallend. Gilt $a<0$, dann wechselt die Folge immer das Vorzeichen. Wir nähern uns der Null also von beiden Seiten.

3. Wir betrachten den Grenzwert

$$\lim_{n\to\infty}\sqrt[n]{n} = 1$$

**Beweis** Sei $a_n = \sqrt[n]{n} - 1$, dann gilt $b_n > 0$, denn es gilt

$$\forall a \in \mathbb{R} : a > 1 \Rightarrow \sqrt[n]{a} > 0.$$

Weiter gilt

$$n = (\sqrt[n]{n} - 1 + 1)^n = (a_n + 1)^n = \sum_{k=0}^{n} \binom{n}{k} a_n^k \geq 11 + \frac{n(n-1)}{2} a_n^2$$

womit folgt

$$n - 1 \geq \frac{n(n-1)}{2} a_n^2$$

worraus schließlich durch Multiplikation mit $\frac{2}{n(n-1)}$ folgt

$$\frac{2}{n} \geq a_n^2 \geq 0$$

und damit ist nach dem Vergleichskriterium (siehe Satz 1.10 auf Seite 20) $(a_n^2)$ eine Nullfolge, womit auch $(a_n)$ eine Nullfolge ist.

Das Vergleichskriterium können wir hier verwenden, da

$$0 \underset{n\to\infty}{\longrightarrow} 0$$

und

$$\frac{2}{n} \underset{n\to\infty}{\longrightarrow} 0$$

gelten.

$\square$

Für die Folge $\left(\sqrt[n]{k}\right)$ können wir für ebenfalls mit dem Vergleichskriterium zeigen

$$\lim_{n\to\infty} \sqrt[n]{k} = 1$$

denn für alle $n > k$ gilt

$$\sqrt[n]{1} < \sqrt[n]{k} < \sqrt[n]{n}$$

Auch hier dürfen wir das Vergleichskriterium verwenden, da

$$\sqrt[n]{1} \underset{n\to\infty}{\longrightarrow} 1$$

und

$$\sqrt[n]{n} \underset{n\to\infty}{\longrightarrow} 1$$

gelten.

**Beispiel** rekursiv definierte Folgen

4. Wir bertrachten die Folge $(a_n)$ gegeben durch $a_1 = 1$ und $an+1 = \frac{6(1+a_n)}{7+a_n}$ für $n \geq 1$.

   Hey, hier haben wir mal was ganz neues! Die Folge ist nicht durch eine Funktionsvorschrift
   $$a_n = f(n)$$
   gegeben, sondern durch die rekursive Vorschrift
   $$a_{n+1} = f(a_n).$$
   Wie bei jeder rekursiven Vorschrift muss ein Anfang vorgegeben sein. Hier durch $a_1 = 1$.

   a) Die Folge $(a_n)$ ist beschränkt, denn es gilt
   $$\forall n \in \mathbb{N} : 0 < a_n < 2$$

   **Beweis** Wir zeigen $a_n > 0$ durch vollständige Induktion.

   *Induktionsanfang*:
   $$a_1 = 1 > 0$$
   *Induktionsvorraussetzung*:
   $$a_n > 0$$
   *Induktionsschritt*:
   $$a_{n+1} = \frac{6(1+a_n)}{7+a_n} > 0$$
   denn nach Induktionsvorraussetzung gilt $a_n > 0$.

   Weiter zeigen wir $a_n < 2$ ebenfalls durch vollständige Induktion.

   *Induktionsanfang*:
   $$a_1 = 1 < 2$$
   *Induktionsvorraussetzung*:
   $$a_n < 2$$
   *Induktionsschritt*: Aus der Induktionsvorraussetzung folgt
   $$4a_n < 8 \Rightarrow 6 + 6a_n < 14 + 2a_n = 2(\underbrace{7 + a_n}_{>0})$$
   und damit gilt
   $$a_{n+1} = \frac{6(1+a_n)}{7+a_n} < 2$$
   □

b) $a_n$ ist monoton wachsend, denn es gilt

$$a_{n+1} - a_n = \frac{6+6a_n}{7+a_n} - a_n = \frac{6+6a_n-7a_n-a_n^2}{7+a_n} = \frac{6-a_n-a_n^2}{7+a_n} \overset{\|}{>} 0$$

An der Stelle $\|$ benutzen wir die Aussage $a_n < 2$ aus a), denn damit gilt ebenfalls $a_n^2 < 4$ und damit gilt $a_n + a_n^2 < 6$.

c) Aus a) und b) folgt nach dem Monotoniekriterium (siehe Satz 1.10 auf Seite 20) die Konvergenz der Folge.

Für den Grenzwert der Folge ergibt sich

$$a = \lim_{n\to\infty} a_n = \lim_{n\to\infty} a_{n+1} = \lim_{n\to\infty} \frac{6+6a_n}{7+a_n} \overset{**}{=} \frac{6+6a}{7+a}$$

An der Stelle $**$ wurde zunächst mit den Rechenregeln (siehe Satz 1.9 auf Seite 16) die Grenzwertberechnung in den Bruch hineingezogen

$$\lim_{n\to\infty} \frac{6+6a_n}{7+a_n} = \frac{\lim_{n\to\infty} 6 + \lim_{n\to\infty} a_n}{\lim_{n\to\infty} 7 + \lim_{n\to\infty} a_n}$$

und anschließend die Gleichheit

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

angewendet.

Damit ist für den Grenzwert der Folge die Gleichung

$$a = \frac{6+6a}{7+a}$$

zu lösen. Wir bestimmen also mit der p-q-Formel die Lösung der quadratischen Gleichung

$$a^2 + a - 6 = 0$$

und erhalten

$$a = 2 \vee a = -3.$$

Da für alle Glieder der Folge $a_n > 0$ gilt, muss auch für den Grenzwert $a > 0$ gelten. Wir erhalten also

$$a = 2.$$

**Lemma 1.13**

Seien $(a_n)$ und $(b_n)$ reelle und konvergente Folgen für die gilt

$$\exists N \in \mathbb{N} \forall n \geq N : a_n \leq b_n$$

dann gilt

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

**Beweis** [Folgt vermutlich direkt aus dem Vergleichskriterium] □

**Bemerkung** In diesem Satz kann »$\leq$« nicht durch »$<$« ersetzt werden, denn der Grenzwert kann trotzdem gleich sein. Man betrachte $(a_n)$ monoton wachsend mit dem Grenzwert 2 und $(b_n) = (2)$. Dann gilt $\forall n \in \mathbb{N} : a_n < b_n$ aber $\lim_{n\to\infty} a_n = \lim_{n\to\infty} b_n = 2$.

**Vorlesung V**

### 1.1.8 Mengen von Häufungswerten

Existieren mehr als ein Häufungswert, so kann man immerhin noch den größten bzw. den kleinsten Häufungswert betrachten und in vielen Fällen mit diesen Aussagen treffen. Im Fall nur eines einzigen Häufungswertes, also eines existierenden Grenzwertes, fallen Grenzwert, größter und kleinster Häufungswert natürlich wieder zusammen, sodass die Folgenden Definition auch dann noch einen Limes definiert, wenn nicht nur ein Häufungswert existiert, und somit eine Erweiterung der Limes-Definition (siehe Definition 1.2 auf Seite 8) darstellt.

**Definition 1.14**

Es sei eine Folge $(a_n)$ mit einer Menge $M \subseteq \mathbb{R}$ von Häufungswerten gegeben. Dann ist der *Limes Superior* gegeben durch

$$\limsup_{n\to\infty} a_n = \begin{cases} \infty & \text{falls } \text{ nach oben unbeschränkt} \\ \sup M & \text{falls } M \text{ nach oben beschränkt und } M \neq \emptyset \\ -\infty & \text{falls } a_n \xrightarrow[n\to\infty]{} -\infty \end{cases}$$

Analog dazu ist der *Limes Inferior* gegeben durch

$$\liminf_{n\to\infty} a_n = \begin{cases} -\infty & \text{falls } (a_n) \text{ nach unten unbeschränkt} \\ \inf M & \text{falls } M \text{ nach unten beschränkt und } M \neq \emptyset \\ \infty & \text{falls } a_n \xrightarrow[n\to\infty]{} \infty \end{cases}$$

# 1.2 Undendliche Reihen

## 1.2.1 Definition

**Beispiel**

$$
\begin{aligned}
1+1-1+1-1+1-1+\ldots &= ?\\
1+\frac{1}{2}+\frac{1}{4}+\frac{1}{8}+\frac{1}{16}+\frac{1}{32}+\ldots &= ?\\
1+\frac{1}{2}+\frac{1}{3}+\frac{1}{4}+\frac{1}{5}+\frac{1}{6}+\frac{1}{7}+\ldots &= ?
\end{aligned}
$$

Was ist das Spannende an diesen Ausdrücken? Sie sind regelmäßig aufgebaut, setzen sich aber bis ins unendliche fort. Wir werden gleich sehen, dass man solche Summationsketten auch als undendliche Summe von Folgengliedern auffassen kann, denn die einzelnen Summanden einer solchen unendlichen Summationskette müssen systematisch aufgebaut sein, damit man dem Summationsausdruck einen Sinn geben kann.

Wir wollen uns in diesem Kapitel damit beschäftigen, solchen Ausdrücken eine Bedeutung zu geben. Oder etwas mathematischer $\sum_{k=1}^{\infty} a_k = a_1 + a_2 + a_3 + a_4 + \ldots$ sinnvoll zu definieren.

Das Problem an solchen Ausdrücken liegt in der Unendlichkeit. Alle Rechenregeln, die für endliche Summen gelten, können wir zunächst vergessen und drüfen sie alle einzeln auf Tauglichkeit prüfen, denn wir werden sehen, dass es einzelne gibt, die *nicht* funktionieren. So kann der Wert eines solchen Ausdruckes zum Beispiel bei alternierenden Summationsfolgen von der Reihenfolge der Summanden abhängen. Das kennen wir von endlichen Summen in der Form nicht!

**Definition 1.15**

Es sei $(a_k)_{k=1}^{\infty}$ eine Zahlenfolge.

a) Die Folge $(s_n)_{n=1}^{\infty}$ mit

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

heißt *Partialsummenfolge* von $(a_k)$ oder *(unendliche) Reihe*.

Bezeichnung:

$$\sum_k a_k := \sum_{k=1}^{\infty} a_k := (s_n)$$

b) Die Reihe $\sum_{k=1}^{\infty} a_k$ konvergiert genau dann gegen $a$, wenn die Folge $(s_n)$ gegen $a$ konvergiert.

Da wir vorher die Äquivalenz von der Reihe $\sum_{k=1}^{\infty} a_n$ und der Summe $(s_n)$ definiert haben, ergibt sich daraus eigentlich keine neue Erkenntnis.

Bezeichnung:

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

Achtung: Der Ausdruck $\sum_{k=1}^{\infty}$ bezeichnet also sowohl die Folge der Partialsummen $(s_n)$ selber, als auch den Grenzwert der Folge $(a_n)$, falls diese konvergiert.

Wir sprechen vom *Wert*, dem *Limes* oder der *Summe* der Reihe.

Jede nicht konvergente Reihe heißt divrgent. Eine reelle divergente Reihe heißt bestimmt divergenz, falls $s_n \underset{n\to\infty}{\longrightarrow} \infty$ oder $s_n \underset{n\to\infty}{\longrightarrow} -\infty$ gilt. Sonst heißt sie unbestimmt divergent.

Die $a_k$ bezeichnen wir als *Glieder* der Reihe und nicht als ~~Summanden~~.

### 1.2.2 Beispiele für unendliche Reihen

**Beispiel**

1. *alternierende Reihe*

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

Für die Partialsummenfolge $(s_n)$ gilt damit

$$s_n = \begin{cases} -1 & \text{für } n \text{ ungerade} \\ 0 & \text{für } n \text{ gerade} \end{cases}$$

Damit ergibt sich: $(s_n)$ ist unbestimmt divergent.

Die Partialsummenfolge (also die Summe für eine feste obere Grenze $n$ und nicht $\infty$) wechselt also immer zwischen $0$ und $-1$ hin und her. Untersuchen wir die Konvergenz der Partualsummenfolge $(s_n)$ (betrachten also jetzt die Summe tatsächlich mit oberer Grenze $\infty$) ergeben sich zwei Häufungswerte bei $-1$ und $0$ und damit die unbestimmte Divergenz der Reihe. Die Reihe hat also keinen Wert.

2. *geometrische Reihe*

$$\sum_{k=0}^{\infty} \qquad \text{mit } q \in \mathbb{R} \text{ und } k \in \mathbb{N}_0$$

Damit ergibt sich für die Partialsummemfolge

$$s_n = \sum_{k=0}^{n} q^k = \begin{cases} \frac{q^{n+1}-1}{q-1} & \text{falls } q \neq 1 \\ n+1 & \text{falls } q = 1 \end{cases}$$

Denn allgemein gilt für eine endliche Folge

$$(q-1)\sum_{k=0}^{n} q^k = \sum k = 0^n q^{k+1} - \sum_{k=0}^{n} q^k = \sum_{k=1}^{n+1} q^k - \sum_{k=0}^{n} q^k = q^{n+1} - 1$$

Damit ergibt sich für die Konvergenz der Reihe: $(s_n)$ konvergiert genau dann, wenn $|q| < 1$ ist.

Im Falle der Konvergenz gilt für den Wert der Reihe

$$\lim_{n\to\infty} s_n = \sum_{k=0}^{\infty} q^k = \frac{1}{1-q}$$

Wenn man die geometrische Reihe noch etwas weiter verallgemeinert, kommt man auf

$$\sum_{k=0}^{\infty} aq^k = \begin{cases} \frac{a}{1-q} & \text{für } |q| < 1, \text{ also für } -1 < q < 1 \\ \infty & \text{für } q \geq 1 \text{ und } a > 0 \\ 0 & \text{für } q \geq 1 \text{ und } a = 0 \\ -\infty & \text{für } q \geq 1 \text{ und } a < 0 \\ \text{nicht definiert} & \text{für } q \leq 1, \text{ denn dann unbestimmt divergent} \end{cases}$$

3. *harmonische Reihe*

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

Betrachten wir zunächst einige Werte für diese Reihe

$$s_{10^3} \approx 7 \quad s_{10^6} \approx 14 \quad s_{10^9} \approx 21$$

Diese Werte scheinen eine Konvergenz nahe zu legen, da für immer größere $n$ die Folge nur sehr langsam wächst. Betrachten wir aber

$$s_{2^{n+1}} = 1 + \frac{1}{2} + \left(\frac{1}{3} + \frac{1}{4}\right) + \left(\frac{1}{5} + \frac{1}{6} + \frac{1}{7} + \frac{1}{8}\right) + \left(\frac{1}{9} + \ldots + \frac{1}{16}\right) + \ldots$$
$$+ \left(\frac{1}{2^n+1} + \ldots + \frac{1}{2^{n+1}}\right) < 1 + \frac{1}{2} + \underbrace{2\frac{1}{4}}_{=\frac{1}{2}} + \underbrace{4\frac{1}{8}}_{=\frac{1}{2}} + \underbrace{8\frac{1}{16}}_{=\frac{1}{2}} + \ldots + \underbrace{2^n\frac{1}{2^{n+1}}}_{=\frac{1}{2}} = \frac{n+3}{2}$$

dann sehen wir, dass $(s_{2^{n+1}})$ keine Nullfolge bildet und damit auch $(s_n)$ keine Nullfolge bilden kann. [Ist das die Begründung?]

Weiter ist $(s_n)$ streng monoton wachsend, denn es gilt

$$s_{n+1} - s_n = \frac{1}{n+1} > 0$$

Damit ergibt sich

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

4. Betrachten wir die Fole

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

dann ergibt sich für die Partialsummenfolge

$$\begin{aligned} s_n &= \sum_{k=1}^{n} \frac{1}{k(k+1)} = \sum_{k=1}^{n} \left(\frac{1}{k} - \frac{1}{k+1}\right) \\ &= \sum_{k=1}^{n} \frac{1}{k} - \sum_{k=1}^{n} \frac{1}{k+1} = \sum_{k=1}^{n} \frac{1}{k} - \sum_{k=2}^{n} \frac{1}{k} \\ &= 1 - \frac{1}{n+1} \underset{n\to\infty}{\longrightarrow} 1 - 0 = 1 \end{aligned}$$

Damit gilt also

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

Das eigentlich interessante an diesem Beispiel ergibt die Betrachtung folgender Zerlegung

$$\sum_{k=1}^{\infty} \left(\frac{1}{k} - \frac{1}{k+1}\right) \neq \sum_{k=1}^{\infty} \frac{1}{k} - \sum_{k=1}^{\infty} \frac{1}{k+1} = \underbrace{\text{»}\infty - \infty\text{«}}_{\text{unbestimmter Ausdruck!}} \neq 1$$

Wir können also eine unendliche Summe auf keinen Fall aufteilen in zwei Summen. Nur weil diese Operation mit endlichen Summen problemlos möglich ist, ist sie mit unendliche Summen (=Reihen) noch lange nicht erlaubt.

Weiter folgt

$$1 = \sum_{k=1}^{\infty} \frac{1}{k(k+1)} < \sum_{k=1}^{\infty} \frac{1}{k^2} < 1 + \sum_{k=2}^{\infty} \frac{1}{k-1} = 1 + \sum_{k=1}^{\infty} \frac{1}{k(k+1)} = 2$$

Wir verwenden als obere Einschnürgrenze $1+\sum_{k=2}^{\infty}\frac{1}{k(k-1)}$, da $k(k-1)$ für $k=1$ undefiniert ist, sodass das erste Folgenglied ausgelagert werden muss. Entsprechend können wir dann die Ungleichung verwenden

$$\frac{1}{k(k+1)} < \frac{1}{kk} < \frac{1}{k(k-1)}$$

In Analysis II werden wir dann auch den Wert dieser Reihe bestimmen können und kommen dann zu

$$\sum_{k=1}^{\infty}\frac{1}{k^2} = \frac{\Pi^2}{6} \approx 1.644$$

5. Die Folge

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

konvergiert genau dann, wenn $\alpha < 1$. [Gibt es da nicht einen Beweis in einer Übungsaufgabe?]

### Lemma 1.16

Es sei $(a_k)$ eine Folge für die $a_k \geq 0$ für alle $k \in \mathbb{N}$ gilt. Dann gilt

$$s_n = \sum_{k=1}^{\infty} a_k \text{ ist konvergent } \iff (s_n) \text{ ist beschränkt}$$

Eine Reihe mit nur positiven Gliedern ist genau dann konvergent, wenn sie beschränkt ist. Denn wenn wir nur positive Glieder in einer Summation haben, dann bedeutet eine Beschränkung automatisch auch eine obere Schranke und die obere Schranke muss dann auch der Grenzwert der Reihe sein.

**Beweis** Die Hinrichtung (Jede konvergente Folge ist beschränkt) folgt direkt aus Lemma 1.8 auf Seite 15. Für die Rückrichtung (Ist die Reihe beschränkt und die Glieder nur positiv, dann ist sie konvergent) betrachten wir die Differenz der Reihenglieder

$$s_{n+1} \geq s_n \Rightarrow s_{n+1} - s_n \geq 0$$

Die Reihe hat nach Vorraussetzung nur positive Glieder und wir summieren diese Glieder in der Partualsummenfolge auf. Also kann ein Folgenglied nur größer oder zumindestens genausogroß wie sein Vorgänger sein, denn sie unterschieden sich nur durch die Addition eines nichtnegativen $a_k$.

Also ist die Reihe monoton wachsend und mit dem Monotoniekriterium (siehe Satz 1.10 auf Seite 20) ist die Reihe auch konvergent. □

**Vorlesung VI**

### 1.2.3 Eulersche Zahl

**Lemma 1.17**

$\sum_{k=0}^{\infty} \frac{1}{k!}$ konvergiert.

Man beachte an dieser Stelle die Definition der Fakultät:

$$0! := 1 \qquad n! := n(n-1)!$$

Die Summe dieser Reihe wird mit $e$ bezeichnet und *Eulersche Zahl* genannt. Es gilt

$$e \approx 2.7182818284590$$

**Beweis** Die Partialsummenfolge

$$s_n = \sum_{k=0}^{n} \frac{1}{k!}$$

ist monoton wachsend und beschränkt.

Denn $k$ ist positiv und es gilt $k! < (k+1)!$, denn $k! := k(k-1)!$. Damit gilt $\frac{1}{k!} > \frac{1}{(k+1)!}$ und damit die Monotonie.

Damit haben wir gezeigt, dass die Folge streng monoton wächst und hauen jetzt noch schnell einen Beleg für die Beschränkung der Folge nach oben raus.

Für alle $n \in \mathbb{N}$ gilt

$$\begin{aligned} s_n = \sum_{k=0}^{n} \frac{1}{k!} &= 2 + \sum_{k=2}^{n} \frac{1}{k!} = 2 + \sum_{k=2}^{n} \frac{1}{2 \cdot 3 \cdot 4 \cdot \ldots \cdot k} \\ &\leq 2 + \sum_{k=2}^{n} \frac{1}{2^{k-1}} \leq 2 \sum_{k=2}^{\infty} \frac{1}{2^{k-1}} = 2 + \sum_{k=1}^{\infty} \left(\frac{1}{2}\right)^k = 3 \end{aligned}$$

Mit Lemma 1.16 auf der vorherigen Seite folgt die Konvergenz der Reihe und damit die Existenz von $e$.

Wir haben in diesem Beweis zum einen die Abschätzung $2^{k-1} = \underbrace{2 \cdot 2 \cdot \ldots \cdot 2}_{k-1 \text{ mal}} \leq 2 \cdot 3 \cdot 4 \cdot \ldots \cdot k$ und zum anderen die geometrische Reihe benutzt um die Reihe nach oben abzuschätzen bzw. zu begrenzen. Denn für geometrische Reihen (siehe Abschnitt »geometrische Reihen« in Kapitel 1.2.2 auf Seite 30) gilt

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

□

**Lemma 1.18** Fehlerabschätzung

Es sei

$$s_n = \sum_{k=0}^{n} \frac{1}{k!}$$

dann gilt für alle $n \in \mathbb{N}$

$$0 < e - s_n = \sum_{k=n+1}^{\infty} < \frac{1}{nn!}$$

**Beweis**

$$\begin{aligned}
\sum_{k=n+1}^{\infty} \frac{1}{k!} &= \frac{1}{(n+1)!} \left(1 + \frac{1}{n+2} + \frac{1}{(n+2)(n+3)} + \ldots\right) \\
&< \frac{1}{(n+1)!} \left(1 + \frac{1}{n+2} + {\frac{1}{n+2}}^2 + \ldots\right) = \frac{1}{(n+1)!} \sum_{k=0}^{\infty} \left(\frac{1}{n+2}\right)^k \\
&\overset{\dagger\dagger}{=} \frac{1}{(n+1)!} \frac{1}{1-\frac{1}{n+2}} = \frac{1}{(n+1)!} \frac{n+2}{n+1} < \frac{1}{(n+1)!} \frac{n+1}{n} = \frac{1}{nn!}
\end{aligned}$$

Für alle die es noch nicht gemerkt haben: An der Stelle †† rechnen wir den Wert einer geometrischen Reihe aus (siehe Abschnitt »geometrische Reihen« in Kapitel 1.2.2 auf Seite 30).

□

**Lemma 1.19**

$e$ ist irrational.

**Beweis** Wir nehmen zum Zecke eines Widerspruches an, es gibt eine Zerlegung $e = \frac{p}{q}$ mit $p, q \in \mathbb{N}$. Nun verwenden wir die Fehlerabschätzung mit $n = q$ und erhaltem

$$0 < e - s_Q = \sum_{k=q+1}^{\infty} \frac{1}{k!} < \frac{1}{qq!}$$

Durch Multiplikatio mit $q!$ erhalten wir

$$0 < q!e - q!s_n = \underbrace{q!e}_{\in \mathbb{N}} - \underbrace{q!\left(1 + \frac{1}{1!} + \frac{1}{2!} + \frac{1}{3!} + \ldots + \frac{1}{q!}\right)}_{\in \mathbb{N}} < \frac{1}{q}$$

$q!e = q!\frac{p}{q} = (q-1)!p$ ist eine natürliche Zahl, da die Multiplikation zweier natürlicher Zahlen einen natürliche Zahl ergibt.

$q!\left(1 + \frac{1}{1!} + \frac{1}{2!} + \ldots + \frac{1}{q!}\right) = \frac{q!}{1} + \frac{q!}{1!} + \frac{q!}{2!} + \ldots + \frac{q!}{q!}$ ist eine natürliche Zahl, da die Addition von natürlichen Zahlen eine natürliche Zahl ergibt.

Subtrahiert man eine natürliche Zahl von einer natürlichen Zahl erhält man – so das Ergebnis wie in diesem Fall positiv ist – wieder einer natürliche Zahl.

Es ergibt sich also

$$0 < x < \frac{1}{q} \leq 1 \wedge x \in \mathbb{N} \quad \lightning$$

$\square$

### 1.2.4 Konvergenz-Kriterien und absolute Konvergenz

**Satz 1.20** Cauchy-Kriterium für Reihen

Die Reihe $\sum_{k=1}^{\infty}$ konvergiert genau dann, wenn

$$\exists \varepsilon > 0 \exists N \in \mathbb{N} \forall m \geq n \geq N : \left| \sum_{k=n}^{m} a_k \right| < \varepsilon$$

gilt.

Dieser Satz ist sehr mächtig, denn durch das Cauchy-Kriterium können wir nun eine unendliche Reihe mit unseren Mitteln für endliche Summen untersuchen.

**Beweis** Die Behauptung ist äquivalent zum Cauchy-Kriterium für die Folge der Partialsummen (siehe Satz 1.12 auf Seite 21).

$$\forall \varepsilon > 0 \exists N \in \mathbb{N} \forall m,n \geq N : |s_n - s_{n-1}| \overset{\ddagger\ddagger}{=} \left| \sum_{k=n}^{m} a_k \right| < \varepsilon$$

□

Der eigentliche Beweis steckt nur in der Umformung an der Stelle ‡‡. Der Rest *ist* die Definition einer Cauchy-Folge.

### Lemma 1.21

Ist $\sum_{k=1}^{\infty} a_k$ konvergent, dann gilt

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

**Beweis** Setze im Cauchy-Kriterium $m = n$. □

Notwendig (aber nicht hinreichend) für die Konvergenz einer Reihe ist, dass die Glieder eine Nullfolge bilden.

Die Folge $\sum_{k=0}^{\infty} 1$ konvergiert nicht, weil $(1)_n$ keine Nullfolge ist.

Die Folge $(\frac{1}{k})$ ist eine Nullfolge, aber $\sum_{k=1}^{\infty} \frac{1}{k} = \infty$ konvergiert nicht.

### Definition 1.22

$\sum_{k=1} a_k$ heißt absolut konvergent, falls $\sum_{k=1}^{\infty} |a_k|$ konvergent ist.

### Lemma 1.23

$\sum_{k=1}^{\infty} a_k$ ist konvergent, wenn $\sum_{k=1}^{\infty} a_k$ absolut konvergent ist.

Absolut konvergent ist also eine stärkere Aussage als konvergent, da absolut konvergent die Konvergent mit einschließt.

Dass die Konveregenz schwächer ist als die absolute Konvergenz können wir auch an der Dreiecksungleichung an der Stelle $*$ im folgenden Beweis sehen.

**Beweis** Nach dem Cauchy-Kriterium ergibt sich aus der Konvergenz der Reihe $\sum_{k=1}^{\infty} |a_k|$

$$\forall \varepsilon > 0 \exists N \in \mathbb{N} \forall m, n \geq N : \varepsilon > \sum_{k=n}^{n} |a_k| \overset{*}{\geq} \left| \sum_{k=n}^{m} a_k \right| .$$

Wieder mit dem Cauchy-Kriterium ergibt sich daraus die Konvergenz der Reihe $\sum_{k=1}^{\infty} a_k$. □

Die Dreiecksungleichung an der markierten Stelle kann nur verwendet werden, da wir nach der Anwendung des Cauchy-Kriteriums nur noch eine endliche Summe betrachten und keine unendliche Summation in Form einer Reihe. Auf eine Reihen dürften wir die Dreieicksungleichung nicht anwenden.

OK, wir dürfen doch. Aber erst nach dem nächsten Korollar.

**Korollar 1.24** Dreieicksungleichung für *unendlich* viele Summanden

Konvergiert die Reihe $\sum_{k=1}^{\infty} |a_k|$, dann gilt

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

**Beweis** Nach der Dreiecksungleichung gilt

$$s_n := \left| \sum_{k=1}^{n} a_k \right| \leq \sum_{k=1}^{n} |a_k| =: S_n.$$

Nach Vorraussetzung sind $(s_n)$ und $(S_n)$ konvergent, sodass mit Lemma 1.13 auf Seite 28 gilt

$$\lim_{n \to \infty} s_n \leq \lim_{n \to \infty} S_n$$

□

Dieser Korollar ist keine direkte Folgerung aus dem vorhergehenden Satz, sondern hätte auch schon deutlich eher im Skript gezeigt werden können. Schließlich verwendet der Beweis nicht das Cauchy-Kriterium, was – wie wir im obigen Satz gesehen haben – auch gegangen wäre, sondern beruft sich auf Lemma 1.13.

**Satz 1.25** Vergleichskriterium

a) Es gelte für alle $k \geq k_0$, dass $|a_k| \leq c_k$ und die Konvergenz von $\sum_{k=1}^{\infty} c_k$, dann gilt

$$\sum_{k=1}^{\infty} a_k \text{ ist absolut konvergennt.}$$

Wir nennen $\sum_{k=1}^{\infty} c_k$ konvergente Majorante.

b) Es gelte für alle $k \geq k_0$, dass $a_k \geq d_k \geq 0$ und die Divergenz von $\sum_{k=1}^{\infty} d_k$, dann gilt

$$\sum_{k=1}^{\infty} a_k \text{ ist divergent.}$$

Wir nennen $\sum_{k=1}^{\infty} d_k$ divergente Minorante.

Kennen wir eine konvergente Reihe, deren Glieder stets größer sind als die einer unbekannten Reihe, dann konvergiert auch die unbekannte Reihe.

Kennen wir eine divergente Reihe, deren Glieder stets kleiner sind als die einer unbekannten Reihe, dann divergiert auch die unbekannte Reihe.

Wobei »stets« jeweils erst ab einem Startindex $k_0$ gelten muss. Also erst irgendwo kurz vor der Unendlichkeit. Die endliche Menge der ersten Glieder ist egal.

**Beweis** a) Für ein gegebenes $\varepsilon > 0$ gilt nach dem Cauchy-Kriterium

$$\exists N \in \mathbb{N} \forall m \geq n \geq N \geq k_0 : \left| \sum_{k=n}^{m} c_k \right| < \varepsilon$$

und daraus folgt mit der Vorraussetzung $|a_k| < c_k$

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

Wieder mit dem Cauchy-Kriterium folgt damit die Konvergenz von $\sum_{k=1}^{\infty} a_k$.

b) Wir nehmen zum Zecke des Widerspruches an, $\sum_{k=1}^{\infty} a_k$ wäre konvergent. Dann gilt mit der Aussage aus Teil a)

$$\sum_{k=1}^{\infty} d_k \text{ ist konvergent} \quad \lightning \text{ zur Vorraussetzung}$$

□

**Satz 1.26** Quotientenkriterium

[—]

2 Stetigkeit und Differenzierbarkeit


# 2 Stetigkeit und Differenzierbarkeit

## 2.1 Stetigkeit

## 2.2 Differenzierbarkeit

### 2.2.1 Umkehrfunktionen

**Vorlesung XVII**

Eher so gegen Ende...

Jetzt wissen wir, was eine Umkehrfunktion ist, aber blicken wir zurück auf den Namen dieses Kapitels, dann erkennen wir schnelle, dass wir doch eigentlich ableiten wollten. Also werden wir jetzt versuchen, eine Umkehrfunktion abzuleiten.

**Satz 2.21**

Sei $g : I \to \mathbb{R}$ in einer Umgebung von $x_0 \in I$ stetig und streng monoton und sei weiter $g$ in $x_0$ differenzierbar, sowie $g'(x_0) \neq 0$. Dann ist $g^{-1}$ in $y = g(x_0)$ differenzierbar und hat den Wert

$$(g^{-1})'(y) = \frac{1}{g'(g^{-1}(y))}.$$

Mit diesem Satz wissen wir jetzt, dass eine Umkehrfunktion in einer Stelle $x_0$ genau dann dieffenzierbar ist, wenn die eigentliche Funktion folgende Eigenschaften erfüllt:

- um $x_0$ herum stetig
- streng monoton (hier ist entweder steigend oder fallend gemeint – hauptsache streng, denn sonst wäre die Umkehrfunktion nicht mehr eindeutig, also keine Funktion)
- in $x_0$ differenzierbar (wenig überraschend, denn warum soll es sonst besser werden, nur weil man die Achsen vertauscht.)
- die Ableitung an der Stelle $x_0$ nicht 0 ist (denn sonst ist die Steigung unendlich, wenn man die Achsen vertauscht)

Und dankenswerter Weise kennnen wir auch noch den Wert. Aber wie immer fällt der Wert eher so nebenbei vom Himmel, während die die Differenzierbarkeit an sich die eigentlich tolle Erkenntnis ist.

### Beweis

$$
\begin{aligned}
(g^{-1})'(y_0) &= \lim_{y \to y_0} \left( \frac{g^{-1}(y) - g^{-1}(y_0)}{y - y_0} \right) \overset{\dagger}{=} \lim_{x \to x_0} \left( \frac{x - x_0}{g(x) - g(x_0)} \right) \\
&= \lim_{x \to x_0} \frac{1}{\frac{g(x)-g(x_0)}{x-x_0}} = \frac{1}{g'(x)} = \frac{1}{g'(g(y_0))}
\end{aligned}
$$

□

An der Stelle † gilt mit der Stetigkeit von $g$

$$g(x) \to g(x_0) \iff x \to x_0.$$

Der Beweis funktioniert eigentlich genau umgekehrt, denn aus der Existenz von $\frac{1}{g'(x)}$ leitet sich die Existenz von $(g^{-1})'(y_0)$ ab.

Wenn die Existenz der Ableitung der Umkehrfunktion bereits gesichert ist, kann man der Wert auch mit der Kettenregel bestimmen:

$$
\begin{aligned}
1 = \frac{\mathrm{d}}{dy} y = \frac{\mathrm{d}}{dy} \left( g(g^{-1}(y)) \right) &= g'(g^{-1}(y)) \cdot (g^{-1})'(y) \\
\implies (g^{-1})'(y) &= \frac{1}{g'(g^{-1}(y))}
\end{aligned}
$$

Geometrisch stellt die Umkehrfunktion eine Spiegelung an der Geraden $x = y$ dar. Bei dieser Betrachtung verwundert es auch nicht weiter, dass man eine Tangente an eine bestimmte Stelle des gespiegelten Graphen legen kann, wenn dies mit dem Ausgangsgraphen gelingt. Und wenn man eine Tangente an den Graphen legen kann, ist die Ableitung gerade die Steigung der Tangenten.

## 2.2.2 Beispiele für Umkehrfunktionen

### die Exponentialfunktion

Die Exponentialfunktion ist definiert durch eine unendliche Reihe

$$f(x) = e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!}.$$

Es gilt für alle $x \in \mathbb{R} : f(x) > 0$, denn für positive $x$ sieht man schnell ein, dass die Reihe positiv ist und für negative $x$ gilt nach weiter oben bewiesenen Gesetzen

$$e^x = \frac{1}{e^{-x}}$$

was also gerade der Kehrwert aus der Reihe für ein positives $x$ ist und damit auch positiv. Weiter ist $f$ streng monoton wachsend, denn die Ableitung der Exponentialfunktion ist die Exponentialfunktion selber und diese ist stets größer als Null. Also existiert die Umkehrfunktion.

Geben wir ihr noch schnell einen Namen: *logarithmus naturalis*:

$$\ln : \{x \in \mathbb{R} \mid x > 0\} \to \mathbb{R}.$$

Für unsere neue Funktion gilt

$$\forall x \in \mathbb{R} : \ln e^x = x$$
$$\forall x > 0 : e^{\ln x} = x$$

Bei der Betrachtung obiger Identitätsfunktionen ist die Berücksichtigung der jeweiligen Funktions- und Wertebereiche enorm wichtig. Die Logarithmus-Funktion bilden von positiven Zahlen auf alle Zahlen ab. Die Exponentialfunktion von allen Zahlen auf positive Zahlen. Je nach dem, wie man diese Funktionen nun verkette, erhält man eine Identität auf allen Zahlen, oder nur auf den positiven Zahlen.

Schließlich gilt für die Ableitung der Umkehrfunktion

$$\forall x > 0 : \frac{\mathrm{d}}{\mathrm{d}x} \ln x = \frac{1}{e^{\ln x}} = \frac{1}{x}$$

### die allgemeine Exponentialfunktion

Die allgemeine Exponentialfunktion kann nur definiert werden für $a > 0$ und $a \neq 1$.

$$f : \mathbb{R} \to \{x \in \mathbb{R} \mid x > 0\} \text{ mit } f(x) = a^x := e^{x \ln a}.$$

Man beachte: Jetzt ist auch die allgemeine Exponentialfunktion über Exponentialfunktion und damit über eine unendliche Reihe definiert.

Für die Ableitung gilt

$$f'(x) = e^{x \ln a} \cdot \ln a = a^x \cdot \ln a \begin{cases} > 0 & \text{für } a > 1 \\ < 0 & \text{für } a < 1. \end{cases}$$

Wir sehen also, dass f streng monoton ist. Also existiert eine Umkehrfunktion und auch ihr geben wir einen Namen. Wir definieren die *allgemeine Logarithmusfunktion*

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

**Vorlesung XVIII**

### Potenzfunktion

$$f : \{x \in \mathbb{R} \mid x > 0\} \to \{x \in \mathbb{R} \mid x > 0\} \text{ mit } f(x) = x^\alpha := e^{\alpha \ln x}$$

Für die Ableitung gilt

$$f'(x) = \frac{\mathrm{d}}{\mathrm{d}x} e^{\alpha \ln x} \frac{\alpha}{x} = \alpha \frac{x^\alpha}{x^1} = \alpha x^{\alpha-1}.$$

Für $\alpha \in \mathbb{N}$ schafft das keine neue Erkenntnis, aber die Erweiterung von$\alpha$ auf $\mathbb{R}$ ist an dieser Stelle neu. Man beachte, dass diese Erweiterung nur für den Bereich der *positiven* reellen Zahlen definiert ist!

Für $\alpha \neq 0$ existiert die Umkehrfunktion

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

Achtung: Für $\alpha \in \mathbb{R}$ ist die Beschränkung auf positive relle Zahlen zwingend notwendig, denn sonst könnte zum Beispiel gelten

$$-1 = (-1)^{\frac{2}{2}} = \left((-1)^2\right)^{\frac{1}{2}} = 1^{\frac{1}{2}} = \sqrt{1} = 1$$

Für den Spezialfall $\alpha \in \mathbb{N}_0$ ist $x \in \mathbb{R}$ möglich und für $\alpha \in \mathbb{Z}$ ist $x \in \mathbb{R}\backslash\{0\}$ erlaubt.

Weitere Probleme ergeben sich bei Wurzeln aus negativen Zahlen. Im Reellen gilt

$$x = -1 \Leftrightarrow x^3 = -1 \Leftrightarrow x^3 + 1 = 0 \Leftrightarrow (x+1)(x^2 - x + 1) = 0.$$

Ist also $\sqrt[3]{-1} = 1$? Ist möglich, aber nur, solange es sich um eine Spezial-Definition der Wurzel-Funktion handelt. Denn sonst könnte auch hier gelten

$$\sqrt[3]{x} = x^{\frac{1}{3}} = x^{\frac{2}{6}} \Rightarrow -1 = 1.$$

Wir beschränken die Definition in dieser Vorlesung also auf $x > 0$.

### Logarithmisches Differenzieren

Für die Ableitung der Logarithmus-Funktion von einer beliebigen Funktion $f$ gilt mit der Ketten-Regel

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

Daraus folgt unmittelbar

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

Die Ableitung einer Funktion lässt sich also bestimmen, indem lediglich der Logarihtmus der Funktion abgelitten wird.

Betrachten wir zum Beispiel die Funktion $f(x) = x^x$ für $x > 0$, dann gilt ganz klassisch berechnet für die Ableitung

$$\frac{\mathrm{d}}{\mathrm{d}x} f(x) = \frac{\mathrm{d}}{\mathrm{d}x} e^{x \ln x} = e^{x \ln x}(\ln x + 1) = x^x(1 + \ln x).$$

Mit Hilfe des logarithmischen Differenzierens ergibt sich

$$\frac{\mathrm{d}}{\mathrm{d}x} \ln f(x) = \frac{\mathrm{d}}{\mathrm{d}x} x \ln x = 1 \ln x + x \frac{1}{x} = 1 + \ln x.$$

Damit gilt für die Ableitung also

$$f'(x) = f(x) \frac{\mathrm{d}}{\mathrm{d}x} \ln f(x) = x^x(1 + \ln x).$$

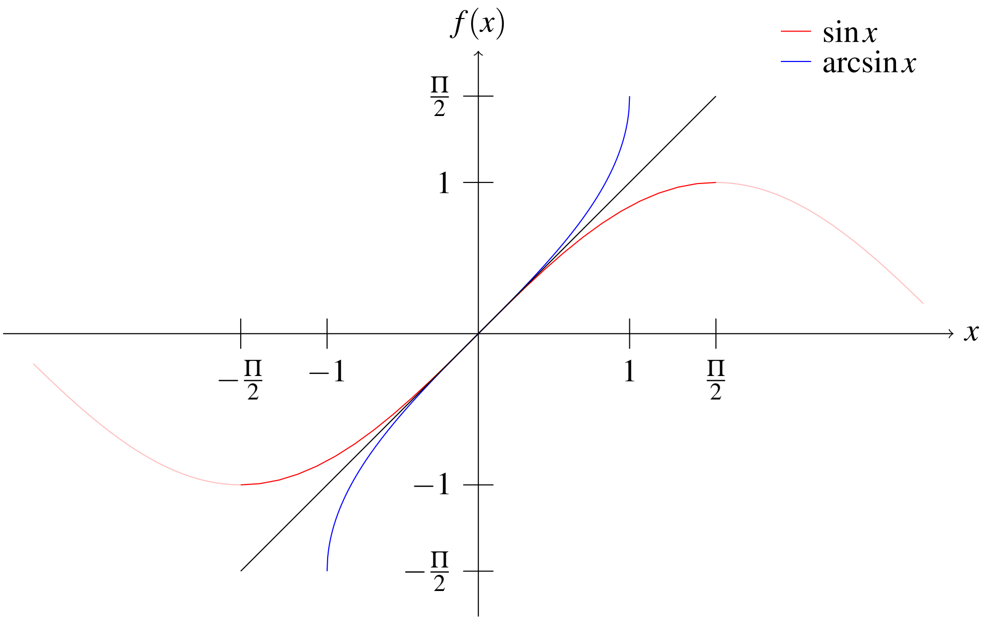
### Zyklometrische oder Arcus-Funktionen

Die Funktion $f(x) = \sin(x)$ ist auf dem Intervall $\left[-\frac{\Pi}{2}, \frac{\Pi}{2}\right]$ streng monoton wachsend und besitzt daher eine Umkehrfunktion. Geben wir auch dieser Umkehrfunktion einen Namen und definieren den *arcus sinus*

$$x = \arcsin y :\Leftrightarrow y = \sin x.$$

$$f^{-1} : [-1, 1] \to \left[-\frac{\Pi}{2}, \frac{\Pi}{2}\right], x \mapsto \arcsin x$$

Geometrisch handelt es sich bei der Umkehrfunktion noch immer um eine Spiegelung an der $y = x$ Achse.



Für die Ableitung ergibt sich

$$\frac{\mathrm{d}}{\mathrm{d}x}\arcsin x = \frac{1}{cos(arcsin(x))} \overset{\ddagger}{=} \frac{1}{\sqrt{1-(\sin(\arcsin x))^2}} = \frac{1}{\sqrt{1-x^2}} \text{ für } x \in (-1,1)$$

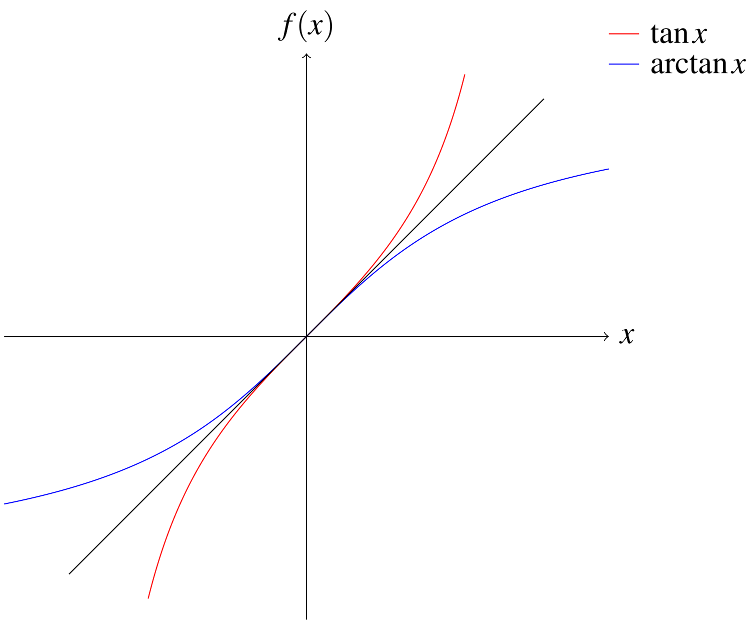
An der Stelle $\ddagger$ dürfen wir quadrieren und die Wurzel ziehen, denn $\arcsin$ bildet auf das Intervall $\left[-\frac{\Pi}{2},\frac{\Pi}{2}\right]$ ab und der $\cos$ ist in diesem Intervall immer echt größer Null.

Der arccos wird völlig analog zum arcsin definiert. Allerdings natürlich über einem entsprechenden Intervall auf dem arccos streng monoton wachsend ist.

Betrachten wir zum Abschluss die Tangens-Funktion. Die Funktion $f(x) = \tan x$ ist auf dem Intervall $(-\frac{\Pi}{2},\frac{\Pi}{2})$ streng monoton wachsend. Es existiert also auch hier eine Umkehrfunktion, die wir analog *arcus tangens* nennen

$$\arctan : \mathbb{R} \to (-\frac{\Pi}{2},\frac{\Pi}{2}), x \mapsto f^{-1}(x).$$

Geometrisch ergibt sich auch hier eine Spiegelung an der $y = x$ Achse.




Für die Ableitung gilt

$$\frac{\mathrm{d}}{\mathrm{d}x}\arctan x = \frac{1}{1+(\tan(\arctan x))^2} = \frac{1}{1+x^2}.$$

### 2.2.3 schwierige Grenzwerte mit De L'Hospital

In diesem Abschnitt werden wir zunächst den Mittelwertsatz in eine etwas allgemeinere Form bringen und den allgemeinen Mittelwertsatz dann für die Regel von De L'Hospital verwenden. Diese vielen aus der Schule bekannte Regel hilft uns weiter, wenn bei der Grenzwertbildung »$\frac{0}{0}$« oder »$\frac{\infty}{\infty}$« entsteht. Haben wir diese Probleme bisher immer gelöst, indem die Funktion solange umgeformt wurde (zum Beispiel mit Polinomdivision) bis ein »echter« Grenzwert entstand, werden wir jetzt sehen, wie uns die Ableitung der Funktion schneller zum Ziel führen kann.

**Satz 2.22** Verallgemeinerung des Mittelwertsatzes

Es seien $f,g : [a,b] \to \mathbb{R}$ zwei auf dem Intervall $(a,b)$ stetige und differenizierbare Funktionen, wobei gelte

$$\forall x \in (a,b) : g'(x) \neq 0.$$

Dann gilt

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

**Bemerkung**

1. Die Nenner beider Brüche werden nicht 0. Links trivial, rechts gegeben durch Anwendung des Mittelwertsatzes auf $g$

$$\exists \xi \frac{g(b)-g(a)}{b-a} = g'(\xi) \neq 0.$$

2. Der Mittelwertsatz ist ein Spezialfall mit $g(x) = x$.
3. Die Verallgemeinerung ist nicht trivial durch Anwendung des Mittelwertsatzes auf $f$ unf $g$ zu erhalten.

$$\exists \xi, \zeta : \frac{f'(\xi)}{g'(\zeta)} = \frac{\frac{f(b)-f(a)}{b-a}}{\frac{g(b)-g(a)}{b-a}} = \frac{f(b)-f(a)}{g(b)-g(a)}.$$

Die Verallgemeinerung besagt, dass $\xi = \zeta$ gilt.

**Beweis** Wir betrachten zunächst die Hildfunktion $F : [a,b] \to \mathbb{R}$ mit

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

Es gilt

$$F(b) - F(a) = f(b) - f(a) - \frac{f(b)-f(a)}{g(b)-g(a)}(g(b)-g(a)) = 0.$$

Mit dem Satz von Rolle [verweis] (»Mittelwertsatz für gleiche Randwerte erzwingt Existenz einer Nullstelle«) folgt

$$\exists \xi 0 = F(\xi) = f'(\xi) - \frac{f(b)-f(a)}{g(b)-g(a)} g'(\xi).$$

□

**Satz 2.23** Regel von De L'Hospital (1661–1704)

Es seien $f, g : (a,b) \to \mathbb{R}$ zwei differenzierbare Funktionen und es gelte

$$\forall x \in (a,b) : g'(x) \neq 0$$

und

$$\lim_{x \to b-} f(x) = \lim_{x \to b-} g(x) = 0 \text{ (oder } \infty).$$

Dann gilt (wobei $c \in \mathbb{R} \cup \{\infty, -\infty\}$)

$$\lim_{x \to b-} \frac{f'(x)}{g'(x)} = c \Longrightarrow \lim_{x \to b-} \frac{f(x)}{g(x)} = c.$$

Der Satz ist analog für $x \to a+$ möglich.

**Beweis** Der Beweis wird hier nur für »$\frac{0}{0}$« durchgeführt. Der Beweis für »$\frac{\infty}{\infty}$« verläuft analog, ist nur technisch etwas aufwändiger.

Wir setzen zunächst $f$ und $g$ auf $[a,b]$ stetig fort, indem wir definieren

$$f(b) := 0 \text{ und } g(b) := 0.$$

Nun betrachten wir ein beliebiges $x \in (a,b)$. Nach dem allgemeinen Mittelwertsatz existiert ein $\xi \in (a,b)$ mit

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

Bilden wir nun den Grenzwert $x \to b-$, so erhalten wir die Behauptung

$$c = \lim_{x \to b-} \frac{f'(\xi)}{g'(\xi)} = \lim_{x \to b-} \frac{f(x)-f(b)}{g(x)-g(b)} = \lim_{x \to b-} \frac{f(x)}{g(x)}.$$

□

Am Ende des Beweises haben wir auf der linken Seite durch den allgemeinen Mittelwertsatz die Existenz eines Wertes

$$c = \lim_{x \to b-} \frac{f'(\xi)}{g'(\xi)}$$

gegeben, der auf rechten Seite durch unsere stetige Fortsetzung von $f$ und $g$ genau dem Ziel der Berechnung

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

entspricht.

**Beispiel**

1. $\lim\limits_{x\to 1}\dfrac{x^3+x^2+x-3}{3x^2+x-4} = »\dfrac{0}{0}« = ?$

$$\lim_{x\to 1}\frac{3x^2+2x+1}{6x+1} = \frac{6}{7} \Rightarrow \frac{x^3+x^2+x-3}{3x^2+x-4} = \frac{6}{7}$$

Diese Berechnung ist auch ohne die Regel von De L'Hospital sehr einfach durchzuführen. Durch Anwendung der Polinomdivision kann man den Bruch um $x-1$ kürzen und erhält

$$\lim_{x\to 1}\frac{x^3+x^2+x-3}{3x^2+x-4} = \lim_{x\to 1}\frac{x^2+2x+3}{3x+4} = \frac{1^2+2\cdot 1+3}{3\cdot 1+4} = \frac{6}{7}$$

Die Regel von De L'Hospital darf *nur verwendet werden*, wenn die Vorraussetzungen erfüllt sind. An diesem Beispiel sehen wir

$$\lim_{x\to 1}\frac{6x+2}{6} = \frac{8}{6} = \frac{4}{3} \neq \frac{6}{7},\ \text{denn}\ \frac{4}{3} \neq »\frac{0}{0}«$$

2. $\lim\limits_{x\to 0}\dfrac{\sin x}{x} \overset{\S}{=} \lim\limits_{\to 0}\dfrac{\cos x}{1} = 1$

Die Gleichheit an der Stelle § gilt nur, da der Grenzwert auf der rechten Seite tatsächlich existiert. Wenn der Grenzwert der Ableitungen nicht dfiniert ist, dann hat die Regel von De L'Hospital keine Aussage. In diesem Fall kann der eigentliche Grenzwert durchaus definiert sein.

**Vorlesung XIX**

Weitere Anwendungsmöglichkeiten der Regel von De L'Hospital

1. Typ »$0\cdot\infty$«: $f(x)\cdot g(x)$

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

Damit ist der nicht direkt definierte Fall wieder auf die Fälle »$\frac{0}{0}$« bzw. »$\frac{\infty}{\infty}$« zurückgeführt.

2. Typ »$\infty - \infty$«: $f(x) - g(x)$

$$\begin{aligned} f(x) - g(x) &= f(x)g(x)\left(\frac{1}{g(x)} - \frac{1}{f(x)}\right) \\ &= \frac{\frac{1}{g(x)} - \frac{1}{f(x)}}{\frac{1}{f(x)g(x)}} \end{aligned}$$

3. Typ »$1^\infty$«: $f(x)^{g()x}$

$$f(x)^{g(x)} = e^{\ln(f(x)^{g(x)})} = e^{g(x)\ln(f(x))}$$

Das ist jetzt wieder der Fall »$0 \cdot \infty$«.

Die e-Funktion ist stetig!

**Beispiel**

$\lim\limits_{x\to\infty}\left(1+\frac{1}{x}\right)^x$

Das Problem: Die Basis geht gegen 1 und der Exponent gegen $\infty$. Es kann also alles mögliche passieren.

$$\begin{aligned} \lim_{x\to\infty}\left(1+\frac{1}{x}\right)^x &= \lim_{x\to\infty} e^{x\ln(1+\frac{1}{x})} \\ &\overset{¶}{=} e^{\lim_{x\to\infty} x\ln(1+\frac{1}{x})} \\ &= e^{\lim_{x\to\infty}\frac{x}{\frac{1}{\ln(1+\frac{1}{x})}}} = e^{\lim_{x\to\infty}\frac{\ln(1+\frac{1}{x})}{\frac{1}{x}}} \end{aligned}$$

An der Stelle ¶ haben wir die Stetigkeit der Exponentialfunktion verwendet! Denn nur dann können wir den Limes in den Exponenten hineinziehen.

Wir haben jetzt beide Möglichkeiten aufgeschrieben, die nach obigen Umformungsregeln mit De L'Hospital gelöst werden können. Wir entscheiden und für die zweite Variante, da hier einfacher Zähler und Nenner getrennt voneinander abgelitten werden können. Wir verwenden also den Fall, wo Zähler und Nenner gegen Null gehen.

$$= e^{\lim_{x\to\infty}\frac{-\frac{1}{x^2}}{\left(1+\frac{1}{x}\right)\left(-\frac{1}{x^2}\right)}} = e^{\lim_{x\to\infty}\frac{1}{1+\frac{1}{x}}} = e^1 = e$$

Da am Ende der Grenzwert des Quotienten der Ableitungen von Zähler und Nenner existiert, gilt die Gleichheit und wir erhalten einen Wert für den gesuchten Grenzwert.

Liefert Beweis von Satz 1.33

$$\begin{aligned}\lim_{n\to\infty}\left(1+\frac{Y_n}{n}\right) &= e^{\lim_{n\to\infty}\frac{n}{Y_n}\ln\left(1+\frac{1}{\frac{n}{Y_n}}\right)} \\ &= e^{\lim_{n\to\infty}Y_n} \text{ falls } \frac{n}{Y_n}\to\infty\end{aligned}$$

### 2.2.4 Abschätzung von Funktionswachstum

**Definition 2.24** Landau-Symbole (1877–1938)

Sei $a\in\mathbb{R}\cup\{\infty,-\infty\}$ und $f$ und $g$ Funktionen in Umgebung von $a$ (ggf. einseitig).

Dann gilt $f(x)=\mathcal{O}(g(x))$ für $x\to a$ per Definition genau dann, wenn

$$\forall x \text{ aus der Umgebung von } a\ \exists C>0:\ |f(x)|\le C|g(x)|.$$

Die schwächere Aussage

$$\lim_{x\to a}\left|\frac{f(x)}{g(x)}\le C\right|$$

ist oft anschaulicher, als obige Definition. In dieser Aussage muss aber ein Grenzwert existieren. In der Definition muss die Funktion $f$ nur immer nach oben beschränkt sein.

Weiter gilt $f(x)=\mathrm{o}(g(x))$ für $x\to a$ per Definition genau dann, wenn

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

**Beispiel**

Sei $n$ für die folgenden Beispiele fest, aber sehr groß.

1. $$\lim_{x\to\infty}\frac{x^n}{e^x}=\lim_{x\to\infty}\frac{nx^{n-1}}{e^x}=\lim_{x\to\infty}=\ldots=\lim_{x\to\infty}\frac{n!}{e^x}=0$$

   De L'Hospital wurde $n$-mal angewendet, bis der Quotient aus einer Konstante und $\infty$ besteht. Dann darf De L'Hospital nicht nochmal verwendet werden. Aber dann wissen wir, dass der Quotient gegen 0 strebt.

2. $\lim\limits_{x\to\infty} \frac{\ln x}{\sqrt[n]{x}} = \lim\limits_{x\to\infty} \frac{\frac{1}{x}}{\frac{1}{n}x^{\frac{1}{n}-1}} = \lim\limits_{x\to\infty} \frac{n}{x^{\frac{1}{n}}} = 0$

Dass heißt also

$$\forall n \in \mathbb{N}: \; x^n = \mathrm{o}(e^x) \text{ für } x \to \infty, \quad \ln x = \mathrm{o}(x^{\frac{1}{n}}) \text{ für } x \to \infty.$$

### 2.2.5 Approximation durch Polynome

Funktionen lassen sich approximieren in Umgebung eines Punktes $x_0$ durch eine konstante Funktion (Polynom nullten Gerades).

$$f(x) \approx p_0(x) := f(x_0)$$

Sehr schnell ist die Approximation ungenau. Man kann die Umgebung, in der die Approximation gut funktioniert erweitern, indem man die Tangente im Punkt $x_0$ verwendet (Polynom ersten Gerades).

$$f(x) \approx p_1(x) := f(x_0) + f'(x_0)(x - x_0)$$

Nun stimmen im Punkt $x_0$ der Wert und die Ableitung von Funktion und Näherung überein.

Das war jetzt alles noch nicht so spannend. Der geneigte Leser sollte was wiedererkannt haben. Also versuchen wir nun, eine Näherung zu finden, die sich noch besser an unsere Funktion anschmiegt. Versuchen wir es mal mit einer Parabel, also einem Polynom zweiten Gerades.

$$f(x) \approx p_2(x)$$

Fesetzsetzung der Freiheistgerade des Polynoms zweiten Gerades.

- $f(x_0) = p_2(x_0)$
- $f'(x_0) = p_2'(x_0)$
- $f''(x_0) = p_2''(x_0)$

Wir sehen also, je mehr Freiheitsgerade des Polynoms, desto besser wird die Approximation. Der Satz von Taylor hilft uns bei der Approximation durch ein Polybom $n$-ten Gerades.

**Satz 2.25** Satz von Taylor

Es sei eine Funktion $f: I \to \mathbb{R}$ gegeben, die $(n+1)$-mal differenzierbar in $I$ sei. Weiter sei ein Entwicklungspunkt $x_o \in I$ gegeben.

Dann gilt für alle $x \in I$

$$\begin{aligned} f(x) &= f(x_0) + \frac{f'(x_0)}{1!}(x-x_0) + \frac{f''(x)}{2!}(x-x_0)^2 + \ldots + \frac{f^{(n)}(x_0)}{n!}(x-x_0)^n + R_n(x) \\ &= \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x-x_0)^k + R_n(x) \end{aligned}$$

Wir bezeichnen das Polynom als *Taylor-Polynom* $n$-ten Gerades $T_n(x,x_0,f) = T_n(x)$ und $R_n(x)$ als Restglied.

Dabei hat das (Lagrange'sche) Restglied mit einem von $x$ abhängigen $\delta \in (0,1)$ die Form

$$R_n(x) = \frac{f^{(n+1)}(x_0 + \delta(x-x_0))}{(n+1)!}(x-x_0)^{n+1}$$

Das Taylor-Polynom ist so gestylt, dass es im Punkt $x_0$ mit der Funktion übereinstimmt. Für das Restglied im Punkte $x_0$ gilt also

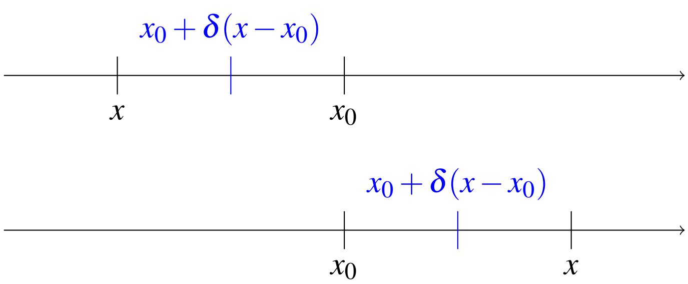
$$R_n(x_0) = 0.$$

Beispielhaft $n = 0$:

$$\begin{aligned} f(x) &= T_0(x) + R_n(x) \\ &= f(x_0) + \frac{f'(x_0 + \delta(x-x_0))}{1!}(x-x_0) \end{aligned}$$

$$\overset{x \neq x_0}{\Leftrightarrow} \frac{f(x) - f(x_0)}{x - x_0} = f'(x_0 + \delta(x-x_0)) = f'(\xi)$$

Das ist genau der Mittelwertsatz!




$x_0 + \delta(x-x_0)$ bezeichnet irgendeinen Punkt zwischen $x$ und $x_0$. Dabei ist es gleich, ob $x$ links oder rechts von $x_0$ liegt.

### Andere Darstellung der Taylor-Formel

$$f(x+h) = \sum_{h=0}^{n} \frac{f^{(k)}(x)}{k!} h^k + \frac{f^(n+1)(\xi)}{(n+1)!} h^{n+1}$$

Wobei $\xi$ zwischen $x$ und $x+h$. (Beide Fälle $h$ positiv oder negativ.)

**Vorlesung XX**

**Beweis** Die Funktion $f$ lebt auf dem Intervall $I$ nach Definition im obigen Satz.

Wir bterachten in $I$ die $(n+1)$-mal differenzierbare Hilfsfunktion

$$\begin{aligned}\varphi(x) :=& f(x) - T_n(x) = f(x) - f(x_0) - \frac{f'(x_0)}{1!}(x-x_0) \\ & - \frac{f''(x_0)}{2!}(x-x_0)^2 - \ldots - \frac{f^{(n)}}{n!}(x-x_0)^n.\end{aligned}$$

Es gilt $0 = \varphi(x_0) = \varphi'(x_0) = \varphi''(x_0) = \ldots = \varphi^{(n)}(x_0)$.

Das Taylor-Polynom stimmt also in allen ersten $n$ Ableitungen mit der Funktion überein.

$$\begin{aligned}\varphi(x_0) =& f(x_0) - T_n(x_0) \\ =& f(x_0) - f(x_0) - \frac{f'(x_0)}{1!}(x_0 - x_0) \\ & - \frac{f''(x_0)}{2!}(x_0-x_0)^2 - \ldots - \frac{f^{(n)}(x_0)}{n!}(x_0-x_0)^n = 0 \\ \varphi'(x_0) =& f'(x_0) - T_n'(x_0) \\ =& f'(x_0) - \frac{f'(x_0)}{1!} - \frac{2f''(x_0)}{2!}(x_0-x_0) - \ldots - \frac{nf^{(n)}(x_0)}{n!}(x_0-x_0)^{n-1} = 0 \\ \varphi''(x_0) =& f''(x_0) - T_n''(x_0) \\ =& f''(x_0) - \frac{2f''(x_0)}{2!} - \frac{6f^{(3)}(x_0)}{6!}(x_0-x_0) - \ldots - \frac{n(n-1)f^{(n)}(x_0)}{(}x_0-x_0)^{n-2} = 0\end{aligned}$$

Weiter gilt für die $(n+1)$-te Ableitung $\varphi^{(n+1)}(x) = f^{(n+1)}$.

Für $\psi(x) = (x-x_0)^{n+1}$ gilt

$$0 = \psi(x_0) = \psi'(x_0) = \ldots = \psi^{(n)}(x_0), \quad \psi^{(n+1)}(x) = (n+1)!$$

Jetzt wollen wir den verallgemeinerten Mittelwertsatz auf diese wunderschönen, gerade vom Himmel gefallenen Funktionen loslassen.

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

$(n+1)$-malige Anwendung des verallgemeinerten Mittelwersatzes. Es gelte $0 < \delta_1, \delta_2, \ldots, \delta_n < 1$.

$$\begin{aligned}\frac{\varphi(x)}{\psi(x)} = \frac{\varphi(x)-\varphi(x_0)}{\psi(x)-\psi(x_0)} &= \frac{\varphi'(x+\delta_1(x-x_0))-\varphi'(x)}{\psi'(x_0+\delta(x-x_0))-\psi'(x_0)}\\ &= \frac{\varphi''(x_0+\delta 1\delta 2(x-x_0))-\varphi''(x)}{\psi''(x_0+\delta_1\delta_2(x-x_0))-\psi''(x_0)}\\ &= \ldots = \frac{\varphi^{(n+1)(x_0+\delta_1\delta_2\ldots\delta_n(x-x_0))}}{(n+1)!}\end{aligned}$$

In jedem Schritt wird jeweils im Zähler und im Nenner $-0$ ergänzt. Denn wir haben ja gesehen, dass die Werte der ersten $n$ Ableitungen von $\varphi$ und $\psi$ von $x_0$ genau 0 sind.

Wir lösen jetzt die Gleichung nach $\varphi(x)$ auf.

$$\varphi(x) = \frac{f^{(n+1)}(\xi)}{(n+1)!}\psi(x) = R_n(x)$$

Am Ende verkürzen wir einfach ein wenig und nennen die Stelle zwischen $x$ und $x_0$ einfach mal anders.

$$\xi = x_0 + \delta_1\delta_2\ldots\delta_n(x-x_0)$$

□

Nur zur Information: Das Restglied $R_n(x)$ lässt sich auch in Integral-Form darstellen.

$$R_n(x) = \frac{1}{n!}\int_{x_0}^{x}(x-t)^n f^{(n+1)}(t)dt$$

**Beispiel**

1. Schießen wir zunächst mit Kanonen auf Spatzen und bilden ein Taylor-Polynom von einem Polynom.

   $f(x) = x^3 + x^2$ mit $x_0 = 1$
   $f(1) = 2,\ f'(1) = 5,\ f''(1) = 8,\ f'''(1) = 6,\ \forall k \geq 4 : f^{(k)}(1) = 0$

   $$\begin{aligned}T_0(x) &= 2\\ T_1(x) &= 2+5(x-1) = 5x-3\\ T_2(x) &= 2+5(x-1)+4(x-1)^2\\ T_3(x) &= 2+5(x-1)+4(x-1)^2+(x-1)^3\\ &= T_4(x) = T_5(x) = \ldots = T_n(x) = f(x)\end{aligned}$$

Wenn ich also ein Taylorpolynom für ein Polynom entwickle, dann bricht die Entwicklung gerade beim Polynom selber ab.

2. Betrachten wir nun eine interessantere Funktion: $f(x) = e^x$, $x_x = 0$, $f^{(k)}(x) = e^x \Rightarrow f^{(k)}(0) = 1 (\forall k \in \mathbb{N})$.

$$\begin{aligned} f(x) &= T_n(x) & &+R_n(x) \\ &= \sum_{k=0}^{n} \frac{1}{k!} x^k & &+\frac{e^\xi}{(n+1)!} x^{n+1} \end{aligned}$$

$$\frac{e^\xi}{(n+1)!} x^{n+1} \xrightarrow[n\to\infty]{} \infty \quad \text{(für festes } x\text{)}$$

## Korollar 2.26

Sei $f : I \to \mathbb{R}$ beliebig oft diefferenzierbar und $x_0 \in I$. Für $x \in I$ gelte $R_n(x) \xrightarrow[n\to\infty]{} 0$. Dann gilt

$$f(x) = \sum_{k=0}^{\infty} \frac{f^{(k)}(x_0)}{k!} (x - x_0)^k.$$

Die Taylor-Reihe stellt also die Ausgangsfunktion selber dar.

## Beispiel

3. $f(x) = \sin x, x_0 = 0$

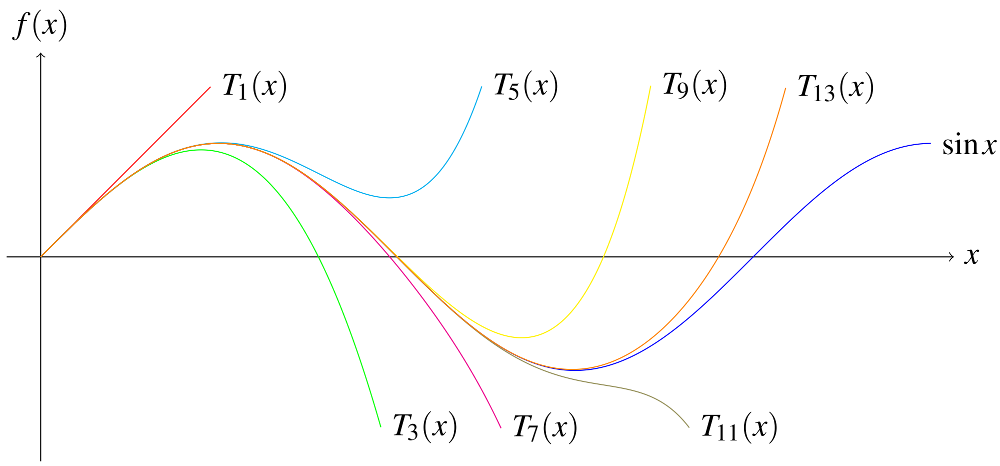
$$R_n(x) = \underbrace{\frac{d^{n+1} \sin}{dx^{n+1}}}_{|\cdot| \leq 1}(\xi) \frac{x^{n+1}}{(n+1)!} \xrightarrow[n\to\infty]{} 0$$

$$\forall x \in \mathbb{R}: \quad \sin x = \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k+1}}{(2k+1)!} = x - \frac{x^3}{6} + \frac{x^5}{120} \pm \ldots$$

An dieser Stelle sehen wir die exakte Definition der Sinus-Funktion über eine unendliche Reihe. In dieser Vorlesung wurde der Sinus aber etwas anschaulicher über seine geometrische Funktion eingeführt.

4. Analog für den Kosinus.

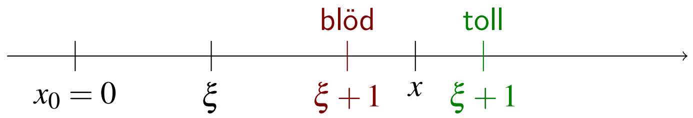
$$\forall x \in \mathbb{R}: \quad \cos x = \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k}}{(2k)!} = 1 - \frac{x^2}{2} + \frac{x^4}{24} \pm \ldots$$




5. $f(x) = \ln(x+1), x_0 = 0$

$$f'(x) = \frac{1}{x+1}$$
$$f^{(k)}(x) = \frac{(-1)^{k+1}(k-1)!}{(x+1)^k} \quad \text{für } k \geq 1$$

$$\begin{aligned} R_n(x) &= \frac{d^{n+1}\ln(x+1)}{\mathrm{d}x^{n+1}} \frac{x^{n+1}}{(n+1)!} \\ &= \frac{(-1)^{n+2}n!}{(\xi+1)^{n+1}} \frac{x^{n+1}}{(n+1)!} \\ &= \frac{(-1)^n x^{n+1}}{(n+1)(\xi+1)^{n+1}} \xrightarrow[n\to\infty]{} 0 \quad \text{falls } 0 \leq x \leq 1 \end{aligned}$$




Für positive $x$ sehen wir direkt an der Skizze, dass wir für $x \leq 1$ immer im tollen Bereich landen. Wir lassen ähnliches für negative $x$ einfach analog vom Himmel fallsen und kommen zu folgender Aussage.

Für $-1 \leq x \leq 1$ gilt

$$\ln(x+1) = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \frac{x^5}{5} \pm \ldots$$
$$\ln 2 = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} \pm \ldots$$

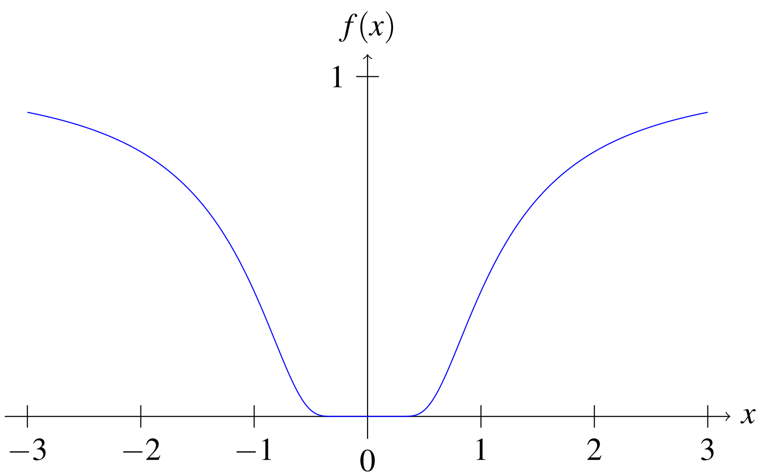
Bis jetzt galt immer $R_n(x) \xrightarrow[n\to\infty]{} 0$, sodass wir die Funktion durch die unendliche Taylor-Reihe ausdrücken konnten. Kommen wir jetzt zu einem Beispiel, bei dem das nicht funktioniniert. Es handelt sich also wirklich um eine notwendige Bedingung.

6. $f(x) = \begin{cases} e^{-\frac{1}{x^2}} & \text{für } x \neq 0 \\ 0 & \text{für } x = 0 \end{cases}$

Eigentlich ist diese Funktion langweilig, denn für Entwicklungen in allen Punkten außer im Nullpunkt geht das Restglied gegen 0. Also betrachten wir die Entwicklung für $x_0 = 0$.

Für alle $x$ außer 0 gilt für die Ableitungen

$$
\begin{aligned}
f'(x) &= \frac{2e^{-\frac{1}{x^2}}}{x^3} \\
f''(x) &= \frac{4e^{-\frac{1}{x^2}}}{x^6} - \frac{6e^{-\frac{1}{x^2}}}{x^4} \\
f^{(3)}(x) &= \frac{8e^{-\frac{1}{x^2}}}{x^9} - \frac{36e^{-\frac{1}{x^2}}}{x^7} + \frac{24e^{-\frac{1}{x^2}}}{x^5} \\
f^{(4)}(x) &= \frac{16e^{-\frac{1}{x^2}}}{x^{12}} - \frac{144e^{-\frac{1}{x^2}}}{x^{10}} + \frac{300e^{-\frac{1}{x^2}}}{x^8} - \frac{120e^{-\frac{1}{x^2}}}{x^6} \\
f^{(5)}(x) &= \frac{32e^{-\frac{1}{x^2}}}{x^{15}} - \frac{480e^{-\frac{1}{x^2}}}{x^{13}} + \frac{2040e^{-\frac{1}{x^2}}}{x^{11}} - \frac{2640e^{-\frac{1}{x^2}}}{x^9} + \frac{720e^{-\frac{1}{x^2}}}{x^7}
\end{aligned}
$$




In der Skizze können wir sehen, dass sich der Graph um 0 herum ganz wunderbar an die X-Achse anschmiegt. (Achtung: Es handelt sich um eine Exponentialfunktion, daher sind selbstverständlich alle Werte $> 0$!) Daher schmiegen sich auch alle Ableitungen um 0 herum an die X-Achse an.

**Vorlesung XXI**

Allgemein gilt für die erste Ableitung

$$f'(x) = \begin{cases} e^{-\frac{1}{x^2}} \frac{2}{x^3} & \text{für } x \neq 0 \\ \lim_{h\to 0} \frac{1}{h}\left(e^{-\frac{1}{h^2}} - 0\right) = \lim_{y\to\infty} \frac{y}{e^{y^2}} = 0 & \text{für } x = 0 \end{cases}$$

Die Substitution von $\frac{1}{h}$ durch $y$ betrachtet nur den Fall, dass $h$ von rechts gegen 0 geht. Für den anderen Fall müsste $y \to -\infty$ betrachtet werden, aber das Ergebnis ist analog 0.

$$f''(x) = \begin{cases} \left(\frac{4}{x^6} - \frac{6}{x^4}\right) e^{-\frac{1}{x^2}} & \text{für } x \neq 0 \\ \lim_{h\to 0} \frac{1}{h}\left(\frac{2e^{-\frac{1}{h^2}}}{h^3} - 0\right) = 0 & \text{für } x = 0 \end{cases}$$

Für einzelne Ableitungen könnte man das jetzt beliebig fortsetzen, aber um allgemeinere Aussagen zu treffen, braucht es einen allgemeinen Beweis, dass die Ableitungen an der Stelle 0 immer 0 sind.

Allgemein zeigt man mit vollständiger Induktion:

$$f^{(k)}(x) = \begin{cases} x^{-3k} e^{-\frac{1}{x^2}} \sum_{j=0}^{2k-2} a_j x^{2j} & \text{für } x \neq 0 \text{ und gewisse } a_j \\ 0 & \text{für } x = 0 \end{cases}$$

Der Beweis ist eine langweilige vollständige Induktion. Mit sowas halten wir uns hier jetzt nicht auf...

$f$ ist beliebig oft differenzierbar und alle Ableitungen $f^{(k)}$ im Punkte 0 sind 0. Für das Taylorpolynom $T_n(x)$ folgert

$$\forall x \in \mathbb{R} \forall n \in \mathbb{N}_0 : \quad T_n(x) = 0.$$

Hier geht also alles schief. Die Taylor-Reihe ist immer 0 und die Funktion ist nur im Punkt 0 genau 0. Woran liegts? Der Rest $R_n(x)$ geht nicht gegen 0! Im Gegenteil, er besteht aus der kompletten Funktion.

Die Funktion ist zwar beliebig oft differenzierbar, aber sie ist nicht analytisch (auch polymorph genannt). Darauf gehen wir aber nicht näher ein. Wichtig ist nur, sie macht alles kaputt, da der Rest nicht gegen 0 geht, da die Ableitungen sehr groß werden.

## 2.3 Konvexität

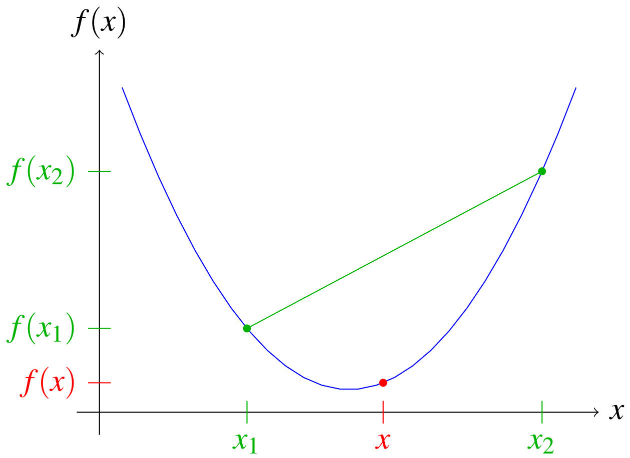
### Definition 2.27

Sei $I$ ein Intervall; die Funktion $f : I \to \mathbb{R}$ heißt *konvex* auf $I$, falls $\forall x, x_1, x_2 \in I$ und $x_1 < x < x_2$ gilt

$$f(x) \leq \frac{x_2 - x}{x_2 - x_1} f(x_1) + \frac{x - x_1}{x_2 - x_1} f(x_2). \qquad (*)$$

- Gilt in $(*)$ »$\leq$« heißt $f$ *konvex*.
- Gilt in $(*)$ »$<$« heißt $f$ *streng konvex*.
- Gilt in $(*)$ »$\geq$« heißt $f$ *konkarv*.
- Gilt in $(*)$ »$>$« heißt $f$ *streng konkarv*.

Im Gegensatz zur punktwiesen Definition von Differenzierbarkeit und Stetigkeit ist die Konvexität auf einem Intervall und nicht in einem Punkt definiert.




In dem Bereich zwischen $x_1$ und $x_2$ muss der Funktionswert immer kleiner sein als der Wert der Geradengleichung $(*)$ der Sekanten.

Für $x_1 < x_2$ gilt $\lambda x_1 + (1 - \lambda) x_2$ ist

- $x_1$ für $\lambda = 1$
- $x_2$ für $\lambda = 0$
- Punkt zwischen $x_1$ und $x_2$ für $0 < \lambda < 1$
- Punkt rechts von $x_2$ für $\lambda < 0$
- Punkt links von $x_1$ für $\lambda > 1$

Auch allgemein für Vektoren $\mathbf{x}_1, \mathbf{x}_2$ und $0 < \lambda < 1$ heißt $\lambda \mathbf{x}_1 + (1-\lambda)\mathbf{x}_2$ konvexe Linearkombination von $\mathbf{x}_1$ und $\mathbf{x}_2$. Dass heißt für die Ungleichung $(*)$

$$(*) \iff \forall x_1, x_2 \in I \text{ und } \forall 0 < \lambda < 1: \quad f(\lambda x_1 + (1-\lambda)x_2) \leq \lambda f(x_1) + (1-\lambda) f(x_2)$$

Weitere Schreibweisen der gleichen Aussage:

$$\forall x_1, x_2 \in I \text{ und } \forall 0 < \alpha, \beta < 1 \text{ und } \alpha + \beta = 1: \quad f(\alpha x_1 + \beta x_2) \leq \alpha f(x_1) + \beta f(x_2)$$

$$\forall x_1, x_2 \in I \text{ und } \forall 0 < \alpha, \beta < 1: \quad f\left(\frac{\alpha x_1 + \beta x_2}{\alpha + \beta}\right) \leq \frac{\alpha}{\alpha + \beta} f(x_1) + \frac{\beta}{\alpha + \beta} f(x_2)$$

### Lemma 2.28

$f$ ist auf $I$ konvex genau dann, wenn gilt

$$\forall x, x_1, x_2 \in I \text{ mit } x_1 < x < x_2: \quad \frac{f(x) - f(x_1)}{x - x_1} \leq \frac{f(x_2) - f(x)}{x_2 - x} \qquad (**)$$

genau dann, wenn gilt

$$\forall x, x_1, x_2 \in I \text{ mit } x_1 < x < x_2: \quad \frac{f(x) - f(x_1)}{x - x_1} \leq \frac{f(x_2) - f(x_1)}{x_2 - x_1} \leq \frac{f(x_2) - f(x)}{x_2 - x}. \qquad (***)$$

**Beweis** Folgt jeweils sofort mit $(*)$, wenn man zusammenfasst zu $f(x) \leq \ldots$ □

### Satz 2.29 Konvexitätskriterium

Die Funktion $f$ sei in dem abgeschlossenen Intervall $[a,b]$ stetig und im offenen Intervall $(a,b)$ differenzierbar. Dann gilt $f$ ist konvex in $[a,b]$ genau dann, wenn $f'$ monoton wachsend in $(a,b)$.

**Beweis** Wir zeigen zunächst die Hinrichtung (»Aus der Konvexität folgt die Monotonie der Ableitung«). Seien dazu $a < x_1 < x_2 < b$. Für alle $x$ mit $x_1 < x < x_2$ gilt dann $(***)$, also auch

$$f'(x_1) \overset{||}{=} \lim_{x \to x_1+} \frac{f(x) - f(x_1)}{x - x_1} \leq \frac{f(x_2) - f(x_1)}{x_2 - x_1} \leq \lim_{x \to x_2-} \frac{f(x_2) - f(x)}{x_2 - x} \overset{**}{=} f'(x_2)$$

An den Stellen $||$ und $**$ benutzen wir, dass die Ableitung $f'$ für $x_1$ und $x_2$ existiert!

Auf der rechten Seite der Ungleichung existiert kein $x$. Daher ändert sich auf der Seite nichrs, wenn $x \to x_1$ geht.

Es gilt also $f'(x_1) \leq f'(x_2)$. Damit ist $f'$ monoton wachsend.

Für die Rückrichtung (»Aus der Monotonie der Ableitung folgt die Konvexität«) seien $x_1 < x < x_2$ beliebig in $[a,b]$.

Nach dem Mittelwertsatz [Verweis] existiert ein $\xi_1 \in (x_1, x)$ mit $\frac{f(x)-f(x_1)}{x-x_1} = f'(\xi_1)$ und ein $\xi_2 \in (x, x_2)$ mit $\frac{f(x_2)-f(x)}{x_2-x} = f'(\xi_2)$. Dass heißt $\xi_1 < \xi_2$ und daraus folgert auch $f'(\xi_1) \leq f!(\xi_2)$, denn die Ableitung ist monoton wachsend. Das ist direkt $(**)$ aus dem Lemma.
□

Auch wenn wir jetzt teilweise nur konvex betrachtet haben, geht das alles genau so durch für streng konvex, konkarv und streng konkarv. Im folgenden Korollar taucht das wieder auf.

### Korollar 2.30

Sei $f$ stetig in abgeschlossenen Intervall $[a,b]$ und im offenen Intervall $(a,b)$ zweimal differenzierbar. Dann gilt

a) $f$ ist konvex genau dann, wenn $f''(x) \geq 0 \forall x \in (a,b)$

b) $f$ ist streng konvex, wenn $f''(x) > 0 \forall x \in (a,b)$

Achtung: In b) gilt nur ein »wenn« und kein »genau dann, wenn«. Wir dürfen also nur von der zweiten Ableitung auf die strenge Konvexität schließen.

**Beweis** a) folgt direkt aus Satz 2.29 auf der vorherigen Seite und Korollar 2.18 e) [verweis]

b) nach a) ist $f$ konvex.

Angenommen es gelte $\forall x \in (a,b): \ f''(x) > 0$, aber $f$ wäre nicht streng konvex, dann existieren $x, x_1, x_2$ mit Gleichheit in $(*)$. Nach Beweis von Satz 2.29 auf der vorherigen Seite gilt $f'(\xi_1) = f'(\xi_2)$. Mit dem Satz von Rolle (»Mittelwertsatz für gleiche Randwerte erzwingt Existenz einer Nullstelle«) [Verweis] folgt daraus, dass es $\xi_1 < \eta < \xi_2$ mit $f''(\eta) = 0$ gibt. Das ist ein Wiederspruch zur Annahme.

□

Wir wissen implizit, dass es umgekehrt auch mit der (strengen) Konkarvität funktioniert.

### Beispiel

1. $f(x) = e^x$ ist streng konvex auf $\mathbb{R}$, denn $f''(x) = e^x > 0$.
2. $f(x) = \ln x$ für $x > 0$ ist streng konkarv für alle $x$, denn $f''(x) = -\frac{1}{x^2} < 0$.

3. Betrachten wir die Ableitungen zu $f(x) = x^P$.

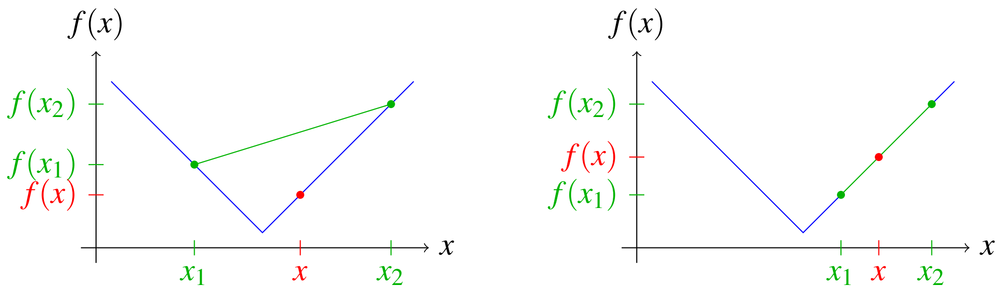
   Wir betrachten $x^P$ für ein reelles $P$, daher kann $x$ nur positive Werte annehmen, damit die Funktion definiert ist. Wir betrachten die Funktion also nur auf dem Definitionsbereich $[0, \infty)$.

$$f'(x) = \begin{cases} Px^{P-1} & \text{für } p \neq 0 \\ 0 & \text{für } p = 0 \end{cases}$$

$$f''(x) = \begin{cases} P(P-1)x^{P-2} & \text{für } P \neq 0 \text{ und } P \neq 1 \\ 0 & \text{sonst} \end{cases}$$

   Also ist $f$ streng konkarv auf $(0, \infty)$ für $0 < P < 1$ und streng konvex für $P < 0$ oder $P > 1$. Im Fall $P = 0$ $(f(x) = 1)$ und $P = 1$ $(f(x) = x)$ betrachten wir eine Gerade und die ist konvex *und* konkarv, aber nicht streng.

4. Die Betragsfunktion $f(x) = |x|$ ist konvex, denn »nach oben offen«, aber nicht streng, denn wenn $x_1, x_2 > 0$ oder $x_1, x_2 < 0$ liegt die Sekante genau auf dem Graphen. Allerdings ist die Betragsfunktion nicht differenzierbar, also funktioniert die Argumentation über die Ableitungen nicht.




**Vorlesung XXII**

## Korollar 2.31

Die Funktion $f$ sei differenzierbar auf dem Intervall $I$. $f$ ist konvex genau dann, wenn $f$ oberhalb jeder Tangente liegt. In Formeln heißt das

$$\forall a, x \in I: \quad f(x) \geq f(a) + f'(a)(x - a)$$

Wir sehen auch hier wieder: Die Tangentengleichung der Funktion $f$ im Punkt $a$ entspricht genau dem Taylor-Polynom erster Ordnung der Funktion $f$ entwickelt im Punkt $a$.

**Beweis** Wir beginnen mit der Hinrichtung (»Wenn $f$ konkarv ist, dann liegt $f$ oberhalb jeder Tangenten«): Für alle $a,x \in I$ gilt nach dem Mittelwertsatz (bzw. der Taylor-Formel für $n = 0$)

$$\exists \xi \text{ zwischen } a \text{ und } x \text{ mit } f(x) = f(a) + f'(\xi)(x-a).$$

Mit Satz 2.29 auf Seite 61 gilt

$$\begin{aligned} f'(\xi) &\geq f'(a) \text{ falls } x > a \\ f'(\xi) &\leq f'(a) \text{ falls } x < a \end{aligned}$$

und damit folgt die Behauptung.

Für die Rückrichtung (»Wenn $f$ oberhalb jeder Tangenten liegt, dann ist $f$ konkarv.«) seien $x_1 < a < x_2$ ($x_1, x_2, a \in I$), dann gilt nach der Tangentenbedingung

$$\frac{f(x_1) - f(a)}{x_1 - a} \leq f'(a) \leq \frac{f(x_1) - f(a)}{x_1 - a}$$

und das ist gerade $(**)$ aus Lemma 2.28 auf Seite 61 und damit folgt die Behauptung. □

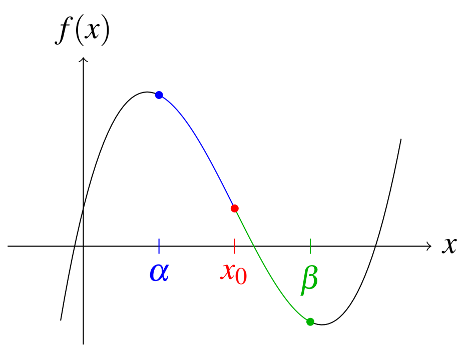
## Definition 2.32

Sei die Funktion $f : (a,b) \to \mathbb{R}$ stetig. Ein Punkt $(x_0, f(x_0))$ heißt Wendepunkt von $f$, wenn Intervalle $(\alpha, x_0), (x_0, \beta) \subset (a,b)$ existieren, sodass

$f$ konvex in $(\alpha, x_0)$ und konkarv in $(x_0, \beta)$

oder

$f$ konkarv in $(\alpha, x_0)$ und konvex in $(x_0, \beta)$.




Die Definition des Wendepunktes erfolgte nur über konvex und konkarv, aber nicht streng konvex oder streng konkarv. Also ist auf einer Geraden jeder Punkt Wendepunkt. Das ist eventuell manchmal etwas unpraktisch. Ersetzt man das »oder« durch ein »entweder … oder«, dann ist der Fall ausgeschlossen.

### Lemma 2.33

Die Funktion $f$ sei zweimal stetig differenzierbar in $(a,b)$. Dann hat $f$ in $(x_0, f(x_0))$ einen Wendepunkt genau dann, wenn (entweder) i) oder ii) gilt.

i) $f''(x) \geq 0$ in $(\alpha, x_0)$,
$f''(x) \leq 0$ in $(x_0, \beta)$

ii) $f''(x) \leq 0$ in $(\alpha, x_0)$,
$f''(x) \geq 0$ in $(x_0, \beta)$

Hinreichende Bedingung: Wenn die Funktion $f$ in der Umgebung von $x_0$ dreimal stetif differenzierbar ist und $f''(x_0) = 0$ und $f'''(x_0) \neq 0$ gilt, dann hat $f$ einen Wendepunkt in $(x_0, f(x_0))$.

Bei der hinreichenden Bedingung gilt nur die Implikation. Es gibt durchaus Wendepunkte, bei denen die Bedingung nicht gilt, aber immer wenn die Bedingung gilt, dannn liegt ein Wendepunkt vor.

**Beweis** Die Aussage folgt direkt aus Korollar 2.30 auf Seite 62. $\square$

### Satz 2.34 Ungleichung von Jensen (1859–1925)

Sei die Funktion $f : I \to \mathbb{R}$ konvex und es seien $\lambda_1, \lambda_2, \ldots, \lambda_n > 0$ mit

$$\sum_{k=1}^{n} \lambda_k = 1$$

dann gilt für alle $x_1, x_2, \ldots, x_n \in I$

$$f\left(\sum_{k=1}^{n} \lambda_k x_k\right) \leq \sum_{k=1}^{n} \lambda_k f(x_k).$$

Ist $f$ eine streng konvexe Funktion, dann gilt die Gleichheit nur, wenn $x_1 = x_2 = \ldots = x_n$.

Der Satz gilt analog für »$\geq$« und »konkarv«.

Ist $f$ linear im Sinne von $f(x) = ax$, dann gilt als wichtigste Eigenschaft der linearen Funktionen $f(x+y) = f(x) + f(y)$. Für lineare Funktionen gilt im Satz immer die Gleichheit.

**Beweis** Wir zeigen die Aussage durch vollständige Induktion.

**Induktionsanfang** Für $n = 1$ erhalten wir die Identität.

Zur reinen Illustration betrachten wir $n = 2$. Wir erhalten genau die Definition der Konvexität. Die Ungleichung von Jensen überträgt diese Ungleichung nun auf mehr als zwei Punkte.

**Induktionsvorraussetzung**

$$f\left(\sum_{k=1}^{n} \lambda_k x_k\right) \leq \sum_{k=1}^{n} \lambda_k f(x_k)$$

**Induktionsschritt**

$$\begin{aligned}
f\left(\sum_{k=1}^{n+1} \lambda_k x_k\right) &\leq f\left(\lambda_{n+1}x_{n+1} + (1-\lambda_{n+1}) \sum_{k=1}^{n} \frac{\lambda_k}{1-\lambda_{n+1}} x_k\right) \\
&\leq \lambda_{n+1} f(x_{n+1}) + (1-\lambda_{n+1}) f\left(\sum_{k=1}^{n} \frac{\lambda_k}{1-\lambda_{n+1}} x_k\right) \\
&\overset{\dagger\dagger}{\leq} \lambda_{n+1} f(x_{n+1}) + (1-\lambda_{n+1}) \sum_{k=1}^{n} \frac{\lambda_k}{1-\lambda_{n+1}} f(x_k)
\end{aligned}$$

□

An der Stelle †† wurde die Induktionsvorraussetzung verwendet. Das ist erlaubt, denn es gilt

$$\sum_{k=1}^{n} \frac{\lambda_k}{1-\lambda_{n+1}} = \frac{1}{1-\lambda_{n+1}}(1-\lambda_{n+1}) = 1$$

### Beispiel

Als typisches Beispiel für die Anwendung der Jensen-Ungleichung betrachten wir die Funktion $f(x) = \ln x$ für $x > 0$. Dann gilt für alle $x_k > 0$ und $\lambda_k > 0$ mit

$$\sum_{k=1}^{n} \lambda_k = 1$$

gilt

$$\ln\left(\sum_{k=1}^{n} \lambda_k x_k\right) \overset{\ddagger\ddagger}{\geq} \sum_{k=1}^{n} \lambda_k \ln x_k \overset{*}{=} \ln\left(\prod_{k=1}^{n} x_k^{\lambda_k}\right).$$

Mit der Monotonie der Logarithmus-Funktion folgt

$$\sum_{k=1}^{n} \lambda_k x_k \geq \prod_{k=1}^{n} x_k^{\lambda_k}.$$

Betrachten wir nun den Spezialfall $\lambda_1 = \lambda_2 = \ldots = \lambda_n = \frac{1}{n}$, dann gilt

$$\underbrace{\frac{1}{n}\sum_{k=1}^{n} xk}_{\text{arithmetisches Mittel}} \geq \underbrace{\sqrt[n]{x_1 x_2 \ldots x_n}}_{\text{geometrisches Mittel}}.$$

Die Gleichheit in obiger Ungleichung gilt genau dann, wenn $x_1 = x_2 = \ldots = x_n$.

An der Stelle ‡‡ wurde die Jensen-Ungleichung angewendet. An der Stelle $*$ hingegen wurden spezielle Eigenschaften der Logarithmus-Funktion verwendet.

# 3 Differenzialrechnung in $\mathbb{R}^n$

## 3.1 Funktionalanalytische Grundlagen

### 3.1.1 Motivation

Wir definieren zunächst Vektoren als Spaltenvektoren, einfach weil das in dieser Vorlesung praktischer ist.

$$\mathbf{x} = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n$$

Und jetzt kommt das neue: Auch Funktionen können ein Vektor sein.

$$\mathbf{f} : \Omega \to \mathbb{R}^m \text{ mit } \Omega \subseteq \mathbb{R}^m$$

Wie funktioniert das jetzt? Naja, wir haben sehr viele Funktionen in einer einzigen.

$$\mathbf{f}(\mathbf{x}) = \mathbf{f}(x_1, x_2, \dots, x_n) = (f_1(x_1, \dots, x_n), \dots, f_m(x_1, \dots, x_n))$$

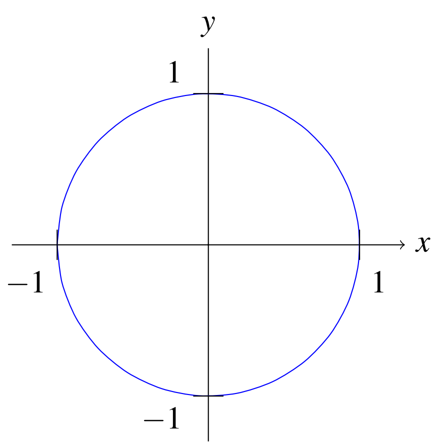
Spannender sind aber einige Spezialfälle von diesem allgemeinem Fall.

1. $n = 1$: $f : \Omega \to \mathbb{R}^m, \Omega \subseteq \mathbb{R}^1$.

   Beispiele:

   $$\mathbf{f}(t) = (\cos t, \sin t) \quad (m = 2) \quad 0 \leq t < 2\Pi$$

   Die Funktionen ist eine Kurve im $\mathbb{R}^2$



$$\mathbf{f}(t) = (\cos t, \sin t, t) \quad (m = 3) \quad t \in \mathbb{R}$$

Diese Funktion ist ebenfalls eine Kurve, aber im $\mathbb{R}^3$. Da zwei Teile der Funktion identisch mit obiger Funktion sind, sieht die Kurve ähnlich aus, steigt aber in der dritten Dimension auf und es entsteht eine Spirale.

2. $m = 1$: $f : \Omega \to \mathbb{R}^1, \Omega \subseteq \mathbb{R}^n$

   Beispiele
   $$f(x_1, x_2) = x_1^2 + x_2^2 \quad \text{»Skalarenfeld«}$$
   Wir schreiben im dem Fall auch $z = f(x,y)$. Dabei sind $x$ und $y$ beide aus dem Definitionsbereich und $z$ aus dem Wertebereich.

### 3.1.2 metrische Räume

**Vorlesung XXIII**

Wir betrachten jetzt Funktionen $f : \Omega \to \mathbb{R}^1, \Omega \subseteq \mathbb{R}^2$ mit zwei Parametern der Bauart $z = f(x,y)$.

Wir brauchen zunächst neue Definitionen für Konvergenz, Stetigkeit und Differenzierbarkeit.

**Definition 3.1** metrische Räume

Sei $E \neq \emptyset$ eine Menge. Eine Abbildung $\rho : E \times E \rho \mathbb{R}$ heißt *Metrik* oder *Abstand* auf $E$, falls für alle $x, y, z \in E$ gilt

a) $\rho(x,y) = 0 \iff x = y$

   Der Abstand zu sich selbst ist Null.

b) $\rho(x,y) = \rho(y,x)$

   Der Abstand ist symmetrisch.

c) $\rho(x,y) \leq \rho(x,z) + \rho(z,y)$

   Der Abstand erfüllt die Dreiecksungleichung.

Das Paar $(E, \rho)$ heißt *metrischer Raum*.

**Beispiel**

a) Es sei $E$ die Menge der Bushaltestellen in Lübeck. Weiter sei $\rho(x,y)$ die kürzeste Fahrstrecke zwischen $x$ und $y$.

   Die Symmetrie ist nur erfüllt, wenn wir davon ausgehen, dass der Bus für den Rückweg den gleichen Weg fährt, wie für den Hinweg. Einbahnstraßen machen diese Annahme in der Regel kaputt.

b) $E$ sei beleibig und $\rho$ die »triviale« Metrik

$$\rho(x,y) = \begin{cases} 1 & \text{für } x \neq y \\ 0 & \text{für } x = y \end{cases}$$

   Die triviale Metrik taugt natürlich für sehr wenige Anwendungen. Aber mit dieser Metrik kann man jede Menge in einen metrischen Raum verwandeln.

c) Es sei $E = \mathbb{R}$ oder $E = \mathbb{C}$ und $\rho(x,y) = |x-y|$.

   Für diese Metrik müssen jetzt tatsächlich Elemente der Menge $E$ her, die man addieren kann. Für die bisherigen Beispiele waren die Elemente der Menge nicht mal Zahlen. Wir werden sehen, dass man viele Dinge definieren kann nur mit Hilfe eines metrischen Raumes. Wir befreien uns also von dem Ballast der ganzen Zahlentheorie und bauen nur auf der Definition der Metrik auf.

### Lemma 3.2

Es sei $(E,\rho)$ ein metrischer Raum, dann gilt für alle $x,y \in E$

$$\rho(x,y) \leq 0.$$

**Beweis**

$$0 \overset{a)}{=} \rho(x,y) \overset{c)}{\leq} \rho(x,y) + \rho(y,x) \overset{b)}{=} 2\rho(x,y)$$

□

### Definition 3.3

Es sei ein metrischer Raum $(E,\rho)$ gegeben.

a) $M \subseteq E$ heißt offen, falls

$$\forall x \in M\ \exists \varepsilon > 0 :\ U_\varepsilon(x) := \{y \in E \mid \rho(y,x) < \varepsilon\} \subseteq M$$

Wir haben hier noch mal eben schnell die $\varepsilon$-Umgebung $U_\varepsilon(x)$ für allgemeine metrischen Räume übertragen. Das ist genauso, wie wir das mit Zahlen schon kennen, aber jetzt braucht es nur noch einen metrischen Raum.

Die Definition der offenen Menge $M$ funktioniert sehr analog zum offenen Intervall auf einer Zahlengeraden. Ich finde zu jedem $x$ – und liege es noch so Dicht an der Grenze der Menge – noch eine Menge mit einem Abstand $\xi$ drumherum, sodass die ganze Menge noch in $M$ liegt. Auf der Zahlengeraden finde ich zu jeder Zahl – und liege sie noch so dicht an der Intervallgrenze – noch eine Zahl, die dichter an der Kante liegt, also eine $\varepsilon$-Umgebung drumherum.

b) $M \subseteq E$ heißt abgeschlossen, falls $E \setminus M$ offen ist.

c) Es sei eine Menge $M \subseteq E$ gegeben. Dann heißt $x \in M$ *innerer Punkt* von $M$, falls

$$\exists \varepsilon > 0: \; U_\varepsilon(x) \subseteq M.$$

*Offene Mengen bestehen nur aus inneren Punkten.* Wir hätten die Definition also anders herum aufeinander Aufbauen können.

d) Ein Punkt $y \in M$ heißt *Randpunkt* von $M \subseteq M$, falls

$$\forall \varepsilon > 0 \; \exists a \in M, b \in E \setminus M: \; a, b \in U_\varepsilon(y)$$

Ein Randpunkt hat also in seiner $\varepsilon$-Umgebung – und sei diese noch so klein – immer Elemente aus der Menge und Elemente, die nicht in der Menge sind.

Die Menge aller Randpunkte $\partial M$ heißt Rand von $M$.

e) $x \in M$ heißt *isolierter Punkt* von $M$, falls

$$\exists \varepsilon > 0: \; U_\varepsilon(x) \cap M = \{x\}.$$

f) $y \in E$ heißt *Berührungspunkt* von $M$, dalls

$$\forall \varepsilon > 0 \; \exists a \in M: \; a \in U_\varepsilon(y).$$

Die Menge $M$ selber besteht nur aus Berührungspunkten von $M$, denn alle Punkte in der Menge berühren natürlich die Menge.

Die Menge aller Berührungspunkte $\overline{M}$ heißt *abgeschlossene Hülle* von $M$.

Die abgeschlossenen Hülle von $M$ sind also alle Punkten aus $M$ und alle Punkten, die $M$ berühren.

g) Ein Punkt $x \in E$ heißt *Häufungspunkt* von $M$, falls

$$\forall \varepsilon > 0\ \exists \text{ unendlich viele } y \in M \cap U_\varepsilon(x).$$

Im Gegensatz zum Häufungspunkt einer Folge müssen es hier unendlich viele verschiedene $y \in M$ sein.

### Lemma 3.4

Es sei ein metrischer Raum $(E, \rho)$ gegeben. Dann gilt

a) $E$ ist offen und abgeschlossen.

Es gibt nur Punkte aus $E$, also sind alle Punkte auch in $E$.

$\emptyset$ ist offen und abgeschlossen.

Also ist die Gesamtemenge der reellen Zahlen eine offene und abgeschlossene Menge.

b) Der Durchschnitt endlich vieler offener Mengen ist offen. (Gilt im allgemeinen nicht für unendlich viele Mengen.)

c) Die Vereinigung beliebig vieler offener Mengen ist offen.

d) offen $\longleftrightarrow$ abgeschlossen
Vereinigung $\longleftrightarrow$ Durchschnitt

Die Vereinigung endlich vieler abgeschlossener Mengen ist abgeschlossen. (Gilt im allgemeinen nicht für unendlich viele Mengen.)

Der Durchscnitt beliebig vieler abgeschlossener Mengen ist abgeschlossen.

**Beweis** a) folgt sofort aus der Definition.

b) Übungsaufgabe

c) Übungsaufgabe

d) Übungsaufgabe

**Gegenbeispiel für d)** $\left[\frac{1}{n}, 1\right]$ für $n \in \mathbb{N}, n \geq 1$

$$\bigcup_{n=2}^{\infty} = (0, 1]$$

$\square$

**Definition 3.5** Konvergenz in einem metrischen Raum

Es sei $(E,\rho)$ ein metrischer Raum und die Folge $(x_k)_k \subset E$ heißt konvergent gegen $a \in E$, falls

$$\forall \varepsilon > 0 \; \exists N = N(\varepsilon) \; \forall k \geq N: \; \rho(x_k, a) < \varepsilon.$$

Equivalent kann man natürlich auch sagen

$$\forall \varepsilon > 0 \; \exists N = N(\varepsilon) \; \forall k \geq N: \; x_k \in U_\varepsilon(a).$$

### 3.1.3 normierte Räume

**Definition 3.6**

Es sei $E$ ein reeller oder komplexer Vektorraum.

$\|\cdot\| : E \to \mathbb{R}$ heißt Norm, falls für alle $\mathbf{x}, \mathbf{y} \in E$ und alle $\lambda \in R(\lambda \in C)$ gilt

a) $\|x\| = 0 \Rightarrow \mathbf{x} = \mathbf{0}$ (Definitheit)

b) $\|\lambda \mathbf{x}\| = |\lambda| \|\mathbf{x}\|$

c) $\|\mathbf{x}+\mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$

Das Paar $(E, \|\cdot\|)$ heißt normierter Raum.

**Lemma 3.7**

Sei $(E, \|\cdot\|)$ ein normierter Raum. Dann gilt

1) $\mathbf{x} = \mathbf{0} \to \|\mathbf{x}\| = 0$

2) $\forall x \in E: \; \|x\| \geq 0$

3) $\forall x, y \in E: \; \rho(x,y) := \|\mathbf{x} - \mathbf{y}\|$ ist eine Metrik auf E.

**Beweis** 1) Setze in Definition 3.6 b) $\lambda = 0$.

2) $0 = \|\mathbf{x} - \mathbf{x}\| \leq \|\mathbf{x}\| + \|-\mathbf{x}\| = 2\|\mathbf{x}\|$

3) a) $\|\mathbf{x} - \mathbf{y}\| = 0 \Leftrightarrow \mathbf{x} = \mathbf{y}$

b) $\|\mathbf{x} - \mathbf{y}\| = \|\mathbf{y} - \mathbf{x}\|$

c) $\|\mathbf{x} - \mathbf{y}\| \leq \|\mathbf{y} - \mathbf{z}\| + \|\mathbf{z} - \mathbf{y}\|$

$\square$

**Beispiel**

1) $E = \mathbb{R}^n, \mathbb{C}^n$

$$\|\mathbf{x}\|_1 = \sum_{j=1}^{n} |x_j| \quad \text{(Manhatten-Norm)}$$

$$\|\mathbf{x}\|_2 = \sqrt{\sum_{j=1}^{n} |x_j|^2} \quad \text{(euklidische Norm)}$$

$$\forall 1 \leq p < \infty : \ \|\mathbf{x}\|_p = \left( \sum_{j=1}^{n} |x_j|^p \right)^{\frac{1}{p}} \quad (p\text{-Norm})$$

$$\|x\|_\infty = \max_{j=1,\dots,n} |x_j| \quad \infty\text{-Norm}$$

Minkowski-Ungleichung für $1 \leq p \leq \infty$:

$$\|\mathbf{x}+\mathbf{y}\|_p \leq \|\mathbf{x}\|_p + \|\mathbf{y}\|_p$$

2) $C([a,b])$ ist der Vektorraum der auf $[a,b]$ stetigen Funktionen.

- Die Nullfunktion ist offensichtlich immer enthalten.
- Man kann Funktionen addieren und erhält neue Funktionen.
- Man kann Funktionen mit Skalaren multiplizieren und erhält neue Funktionen.

$$\|f\| = \max_{x\in[a,b]} |f(x)|$$

$$\|f+g\| = \max_{x\in[a,b]} |f(x)+g(x)| \leq \max_{x\in[a,b]} |f(x)| + \max_{x\in[a,b]} |g(x)|$$

**Vorlesung XXIV**

Nochmal eben zur Erinnerung: Wie funktionieren Durchschnitt und Vereinigung?

$$x \in \bigcup_{j\in J} M_j \Longleftrightarrow \exists j \in J \text{ mit } x \in M_j$$

$$x \in \bigcap_{j\in J} M_j \Longleftrightarrow \forall j \in J \text{ mit } x \in M_j$$

### 3.1.4 Konvergenz von Vektoren

Nachdem wir für metrische Räume die Konvergenz bereits definiert haben (s. Definition 3.5 auf Seite 73), bleibt die Frage, wie Konvergenzen von Vektoren (also im normierten Raum) funktionieren. Die Definition steht, aber es fehlt noch der Griff zum Anfassen...

**Satz 3.8** Komponentenweise Konvergenz

Die Folge $(\mathbf{x}^k)_k \subset \mathbb{R}^n$ mit Normabstand $\|\mathbf{x}-\mathbf{y}\|$ konvergiert gegen $\mathbf{a} \in \mathbb{R}^n$ genau dann, wenn

$$\forall j = 1, \ldots, n: \lim_{k\to\infty} x_j^k = a_j.$$

$x_j^k$ ist die $j$-te Komponente des $k-$ten Vektors der Folge. Hier steht keine Potenz!

**Bemerkung** Die Aussage des Satzes hängt von der Wahl der Norm ab.

Man kann aber zeigen, dass sie davon unabhängig ist: Denn für alle Normen $\|\cdot\|_\alpha$, $\|\cdot\|_\beta$ eines $n$-dimensionalen normierten Raumes $E$ gilt

$$\exists C_1, C_2 > 0\ \forall \mathbf{x} \in E: C_1\|\mathbf{x}\|_\alpha \le \|\mathbf{x}\|_\beta \le C_2\|\mathbf{x}\|_\alpha.$$

In einem endlich dimensionalen Vektorraum sind alle Normen irgendwie gleich. Man kann immer nur von der Norm abhängige Konstanten finden, sodass man die Normen ineinander »umrechnen« kann.

**Beispiel** Die Maximum-Norm und die Summen-Norm (1-Norm, Manhatten-Norm), denn

$$1\|\mathbf{x}\|_\infty \le \|\mathbf{x}\|_1 \le n\|\mathbf{x}\|_\infty.$$

Im dreidimensionalen Raum unterscheiden sich die 1-Norm und die $\infty$-Norm also nur maximal um den Faktor 3. Aber wir sehen, dass der Faktor von $n$ abhängt. Im unendlich dimensionalen Vektorraum funktioniert diese Äquivalenz also nicht mehr.

**Beweis** Da wir gerade gesehen (wenn auch nicht bewiesen) haben, dass die Normen alle äauivalent sind, zeigen wir den Satz nur mit der 1-Norm.

Für die Hinrichtung (»Wenn die Folge der Vektoren konvergiert, dann konvergieren auch die Folgen der Komponenten.«) gilt zunächst

$$\forall \varepsilon > 0\ \exists N\ \forall k \ge N: \|\mathbf{x}^k - \mathbf{a}\|_1 \le \varepsilon.$$

Daraus folgt

$$\forall j = 1,\dots,n \; \forall k \geq N: \; |x_j^k - a_j| \leq \|\mathbf{x}^k - \mathbf{a}\|_1 < \varepsilon$$

dass heißt $\forall j = 1,\dots,n: \; \lim_{k\to\infty} x_j^k = a_j$.

Für die Rückrichtung (»Wenn die Folgen der Komponenten konvergieren, dann konvergiert auch die Folge der Vektoren.«) gilt zunächst

$$\forall j: \; \lim_{k\to\infty} x_j^k = a_j.$$

Daraus folgt

$$\forall \varepsilon > 0 \; \exists N_j \; \forall k \geq N_j |x_j^k - a_j| < \frac{\varepsilon}{n}$$

Wir haben eine *endliche* Anzahl an $N_j$. Genau genommen sind es genau $n$ Stück, denn wir betrachten $n$-dimensionale Vektorräume.

Mit $N := \max\{N_j \mid j = 1,\dots,n\}$ gilt für alle $k \geq N$:

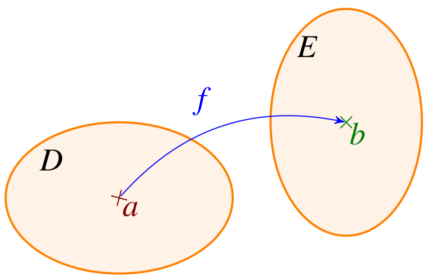
$$\|\mathbf{x}^k - \mathbf{a}\|_1 = \sum_{j=1}^{n} |x_j^k - a_j| < \sum_{j=1}^{n} \frac{\varepsilon}{n} = \varepsilon.$$

□

### 3.1.5 Stetigkeit von Funktionen

**Definition 3.9**

Seien $(D,d)$ und $(E,\rho)$ metrische Räume. Weiter sei die Funktion $f: D \to E$ gegeben und $a \in D$, sowie $b \in E$.




a) Wir sagen $\lim_{x\to a} f(x) = b$ genau dann, wenn

$$\forall (x_k) \subset D \text{ mit } \lim_{x\to\infty} x_k = a \text{ gilt } \lim_{x\to\infty} \lim_{k\to\infty} f(x_k) = b.$$

b) $f$ heißt in $a$ stetig, falls

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

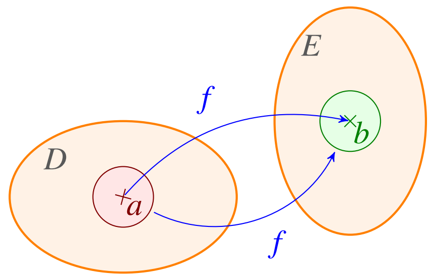
Wir haben also jetzt Grenzwerte und Stetigkeit von Funktionen auf beliebigen metrischen Räumen definiert. Die Funktion $f$ kann also aus einer Menge Äpfel in eine Mengen Birnen abbilden, solange zwischen den Äpfen und zwischen den Birnen ein Abstand definiert ist.

Auch hier haben wir jetzt eine Definition, aber es fehlt irgendwie noch ein Satz, um das ganze in den Griff zu bekommen...

**Satz 3.10** $\varepsilon$-$\delta$-Kriterium der Stetigkeit

Es seien $(D,d)$ und $(E,\rho)$ metrische Räume und $f: D \to E$ eine Funktion. Dann ist $f$ in $a \in D$ stetig genau dann, wenn

$$\forall \varepsilon > 0 \; \exists \delta > 0: \; (\forall x \in D \text{ mit } d(x,a) < \delta \Rightarrow \rho(f(x), f(a)) < \varepsilon)$$




Für jede $\varepsilon$-Umgebung muss eine $\delta$-Umgebung existieren, sodass alles aus der $\delta$-Umgebung nach transformation durch $f$ in der $\varepsilon$-Umgebung liegt. Die Zeichnung impliziert einen geometrischen Abstand. Der Satz isz natürlich nicht darauf beschränkt. Es könnte sich auch um den Preisabstand von Birnen handeln. Das zeichnet sich nur schlechter.

**Beweis** Wir beginnen mit der Hinrichtung (»Aus der Stetigkeit folgt das $\varepsilon$-$\delta$-Kriterium.«). Angenommen das $\varepsilon$-$\delta$-Kriterium ist nicht erfüllt. Dass heißt

$\exists \varepsilon > 0 \forall \delta > 0 \exists z \in D$ mit der Aussage

$$d(z,a) < \delta \text{ und } \rho(f(z), f(a)) \geq \varepsilon.$$

Da dass für alle $\delta$ funktioniert, nehmen wir einfach mal $\delta = \frac{1}{n}$.

Für $\delta := \frac{1}{n} \exists Z_n \in D$ mit

$$d(z_n, a) < \frac{1}{n} \text{ und } \rho(f(z_n), f(a)) \geq \varepsilon.$$

Damit ergibt sich

$$\lim_{n\to\infty} z_n = a \text{ und } \lim_{n\to\infty} f(z) \neq f(a)$$

Damit haben wir ein Widerspruch erzeugt, denn wir gehen davon aus, dass $f$ stetig ist. Jetzt haben wir aber gezeigt, dass $f$ nicht stetig ist.

Für die Rückrichtung (»Aus dem $\varepsilon$-$\delta$-Kriterium folgt die Konvergenz«) sei die Folge $(x_n) \subset D$ so gewählt, dass

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

Wir müssen jetzt zeigen, dass $f(x_n)$ für $n \to \infty$ gegen $f(a)$ konvergiert.

Wegen $x_n \to a$ existiert ein $N$ mit

$$\forall n \geq N \text{ gilt } d(x_n, a) < \delta.$$

Nach dem $\varepsilon$-$\delta$-Kriterium gilt $\rho(f(x_n), f(x)) < \varepsilon$ für alle $n \geq N$. Damit gilt

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

□

## 3.2 Reellwertige Funktionen mehrerer reeller Vektorräume

Würden hier komplexe Funktionswerte stehen, würde das wenig ändern. Betrachtet man aber komplexe Vektorräume, dann wird alles anders und vor allem komplizierter.

- $\mathbb{R}^n$ sei ein normierter Raum mit Norm $\|\cdot\|$
- $\Omega \subset \mathbb{R}^n$ heißt Gebiet, falls
  - $\Omega$ ist eine offene Menge.
  - Für alle $x, y \in \Omega$ existiert eine stetige Funktion $g : [0,1] \to \Omega$ mit $g(0) = x$ und $g(1) = y$

  Für ein Gebiet müssen wir uns nur vorstellen, dass die Menge nur aus einem »Klumpen« besteht und nicht aus mehreren nicht zusammenhängeneden Teilen.
- Sei $f : \Omega \to \mathbb{R}$, wobei $\Omega$ meistens ein Gebiet.

## Definition 3.11

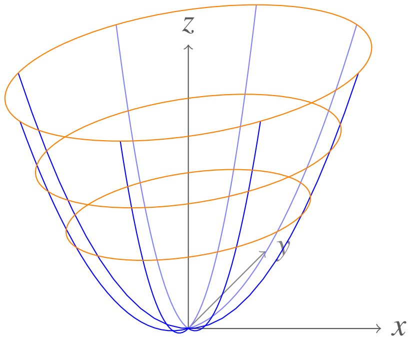
Sei $c \in \mathbb{R}$. Dann heißt

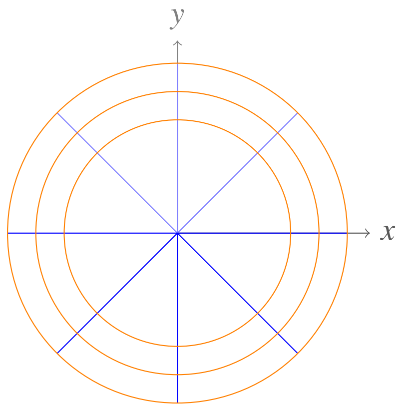
$$N_c = \{\mathbf{x} \in \Omega \mid f(x) = c\}$$

*Niveau-Menge* von $f$ zum Niveau $c$.

## Beispiel

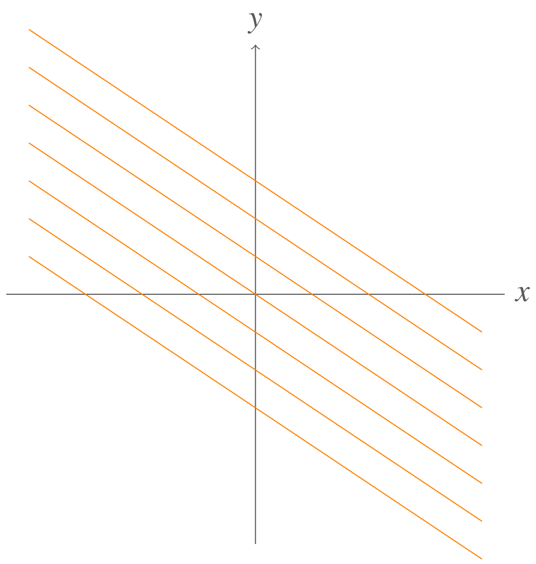
a) $z = f(x,y) = x^2 + y^2$ mit $\Omega = \mathbb{R}^2$




Niveau-Mengen auf den Niveaus 1, $f\left(\frac{5}{4}\right)$ und $f\left(\frac{3}{2}\right)$.

b) Ebene $z = f(x,y) = 2x + 3y$ mit $\Omega = \mathbb{R}^2$




Niveau-Mengen auf den Niveaus $-1$, $-\frac{2}{3}$, $-\frac{1}{3}$, $0$, $\frac{1}{2}$, $\frac{2}{3}$ und 3

$$3y + 2x = c$$
$$y = -\frac{2}{3}x + \frac{c}{3}$$

**Beispiele für stetige Funktionen**

1. Projektionen $\forall x \in \mathbb{R}^n : p : \mathbb{R}^n \to \mathbb{R}^1, p_j(\mathbf{x}) = x_j$ sind stetig, denn
$$\lim_{\mathbf{x}\to\mathbf{a}} p_j(\mathbf{x}) = \lim_{\mathbf{x}\to\mathbf{a}} x_j = a_j = p_j(\mathbf{a}).$$
2. Summen und Produkte stetiger Funktionen sind stetig. (Beweis wie im Eindimensionalen)

   Eigentlich soll da stehen: Wenn zwei Funktionen in einem Punkt stetig ist, dann ist auch die Summe der Funktionen bzw. das Produkt der Funktionen *in diesem Punkt* stetig.

   Lineare Funktionen
$$l(\mathbf{x}) = \sum_{j=1}^{n} b_j x_j = \mathbf{x} \cdot \mathbf{b}^T = \mathbf{b} \cdot \mathbf{x}^T$$
   sind stetig.
$$\mathbf{x} \cdot \mathbf{b}^T = (x_1, x_2, \dots, x_n) \cdot \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}$$
3. Polynome in $n$ Variablen vom Grad $m$
$$p(\mathbf{x}) = p(x_1, x_2, \dots, x_n) = \sum_{d_1=0}^{m} \sum_{d_2=0}^{m} \cdots \sum_{d_n=0}^{m} a_{d_1,d_2,\dots,d_n} x_1^{d_1} x_2^{d_2} \dots x_n^{d_n}$$
   sind stetige Funktionen.

   Polynome sind Linearkombinationen der simpelsten Porjektionen und damit auch stetig.
$$p(\mathbf{x}) = a_{\mathbf{d}} \mathbf{x}^{\mathbf{d}} \text{ mit } \|\mathbf{d}\|_\infty \le m$$

Erstmal ist ein Vektor hoch ein Vektor nicht erklärt und damit mathematischer Unsinn. An dieser Stelle ist eine Abkürzung der obigen Schreibweise gemeit.

Falls $\|\mathbf{d}\|_1 \leq m$, dann hat das Polynom nicht nur den Grad $m$, sondern auch den *Total-Grad m*.

**Beispiel** Polynome mit $n = 2$

Das Polynom
$$p(x,y) = 5 + 7x^1 + 3y^1$$
hat den Grad und den Total-Grad 1.

Das Polynom
$$p(x,y) = 5 + 7x^1 + 3y^1 + 6x^1y^1$$
hat den Grad 1, aber den Total-Grad 2.

Das Polynom
$$p(x,y) = 5 + 7x^1 + 3y^1 + 6x^1y^1 + 4x^2 + 3y^2$$
hat den Grad 2 und den Total-Grad 2.

Das Polynom
$$p(x,y) = 5 + 7x^1 + 3y^1 + 6x^1y^1 + 4x^2 + 3y^2 + 11x^2y - 3xy^2$$
hat den Grad 2 und den Total-Grad 3.

**Beispiel** Stetigkeit eines multivariablen Polynoms

Betrachten wir die Funktion
$$f : \mathbb{R}^2 \to \mathbb{R}$$
mit
$$f(x,y) = \begin{cases} \frac{2xy^2}{x^2+y^4} & \text{falls } x^2 + y^2 > 0 \\ 0 & \text{falls } x^2 + y^2 = 0 \end{cases}$$

Die Stelle $x^2 + y^2 = 0$ ist der Gerade der Ursprung $x = y = 0$. An dieser Stelle passiert irgendwie was spannendes. Wir werden also jetzt versuchen, was passiert, wenn $x$ *und* $y$ gegen 0 gehen.

Wir betrachten zunächst $g_1(x) = f(x,0)$.

Wir setzen also einfach $y = 0$ und können dann eine Funktion in Abhängigkeit von einer Variablen $x$ betrachten.

$$\forall x \in \mathbb{R} : \; g_1(x) = 0$$

Betrachten wir nun $g_2(y) = f(0,y)$.

$$\forall y \in \mathbb{R}: \ g_2(y) = 0$$

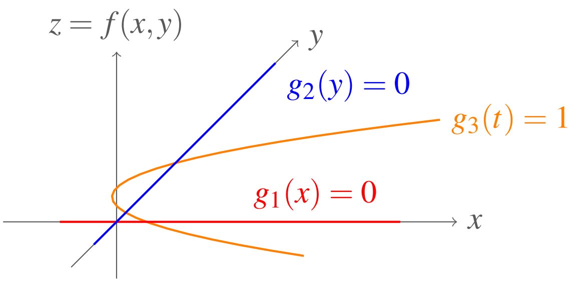
Man könnte jetzt meinen, damit sei der Grenzwert bestimmt.

Betrachten wir jetzt

$$\forall t \in \mathbb{R}: \ g_3(t) = f(t^2,t)$$

$$g_3(t) = \begin{cases} \frac{2t^2}{2t^4} = 1 & \text{falls } t \neq 0 \\ 0 & \text{falls } t = 0 \end{cases}$$

Der Graph einer multivariablen Funktion ist im Prinzip ein »fliegender Teppich«, der über der $x,y$-Ebene schwebt. Je nach Funktionswert hat der Teppich einen anderen Abstand zur Ebene.




Betrachten wir die Lage der Graphen von $g_1, g_2$ und $g_3$ auf dem Teppich, so sehen wir, dass alle Graphen über dem Urpsrung der Ebene vorbeikommen, aber unterschiedliche Funktionswerte haben.

Wir sehen also, dass der »fliegende Teppich« im Nullpuntk ein Loch hat, denn überall auf der Parabel hat er die Höhe 1 und im Urpsrung die Höhe 0. Die Funktion $f$ kann daher nicht stetig sein.

**Vorlesung XXV**

## Satz 3.12

Sei $\Omega \subset \mathbb{R}^n$ beschränkt und abgeschlossen.

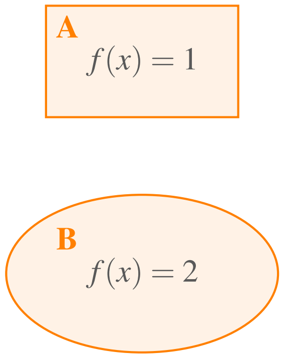
Also quasi die Entsprechung eines abgeschlossenen Intervalls im Eindimensionalen.

Weiter sei $f: \Omega \to \mathbb{R}$ stetig. Dann nimmt $f$ auf $\Omega$ Minimum und Maximum an. In Formeln heißt das

$$\exists \mathbf{a}, \mathbf{b} \in \Omega \ \forall \mathbf{x} \in \Omega: \ f(\mathbf{a}) \leq f(\mathbf{x}) \leq f(\mathbf{b}).$$

Wir haben also ein Minimum (und nicht nur ein Infimum) und ein Maximum (und nicht nur ein Supremum).

Betrachten wir zwei disjunkte Mengen $A$ und $B$, sodass für alle $x \in A$ stets $f(x) = 1$ und für alle $x \in B$ stets $f(x) = 2$ gilt. Sei nun $\Omega = A \cup B$ die Vereinigung beider Mengen. $f$ nimmt nun tatsächlich das Minimum 1 und das Maximum 2 an. Allerdings werden keine Werte zwischen 1 und 2 angenommen, denn die Vereinigung ist kein Gebiet, sondern besteht aus zwei disjunkten Mengen $A$ und $B$.




Falls $f : \Omega \to \mathbb{R}$ stetig ist auf einem Gebiet $\Omega$ und $\mathbf{a}, \mathbf{b} \in \Omega$ sind, dann gilt für alle $c$ mit $f(\mathbf{a}) \leq c \leq f(\mathbf{b})$ mit dem Ziwschenwertsatz die Aussage

$$\exists \mathbf{x} \in \Omega : \; f(\mathbf{c}) = c.$$

Das ist kein Satz und wird nicht bewiesen! Das steht hier einfach nur mal so.

### 3.2.1 partielle Ableitung

Jetzt kommen wir endlich zur Differenzierbarkeit von mutlivariablen Funktionen. Dabei stellt sich als erstes mal, wie man denn nun eigentlich eine Tangente an einen »fliegenden Teppich« legt. Solange man sich nicht auf eine eindimensionale Teilmenge des Teppichs beschränkt, kann die Tangente in jede Richtung zeigen.

**Definition 3.13** Differenzierbarkeit

Sei $\Omega$ ein Gebiet , $f : \Omega \to \mathbb{R}$ eine Funktion und $\mathbf{a} \in \Omega$ fest. Dann heißt

$$x_j \mapsto f(a_1, a_2, \ldots, a_{j-1}, x_j, a_{j+1}, \ldots, a_n)$$

*partielle Funktion* von $f$.

Diese Funktion ist eine gewöhnliche eindimensionale Funktion. Die können wir auch ableiten.

Hat diese Funktion in $x_j = a_j$ eine Ableitung, so nennen wir das die partielle Ableitung von $f$ nach $x_j$ in $\mathbf{a}$.

Wir schreiben

$$\left.\frac{\partial f(\mathbf{x})}{\partial x_j}\right|_{\mathbf{x}=\mathbf{a}} = \frac{\partial f}{\partial x_j}(\mathbf{a}) = f_{x_j}(\mathbf{a}) = \partial_j f(\mathbf{a}).$$

Analog schreiben wir für höhere Ableitungen

$$f_{x_1,x_2} = \frac{\partial^2 f}{\partial x_1 \partial x_2} = \partial_1 \partial_2 f.$$

Äquivalente Definition:

$$\frac{\partial f}{\partial x_j}(\mathbf{a}) = \lim_{h\to 0} \frac{1}{h} \left( f(a_1, \ldots, a_{j-1}, a_j + h, a_{j+1}, \ldots, a_n) - f(\mathbf{a}) \right).$$

Wir schreiben kürzer

$$\partial_j f(\mathbf{a}) = \lim_{h\to 0} \frac{1}{h} \left( f(\mathbf{a} + h\mathbf{e}_j) - f(\mathbf{a}) \right)$$

wobei $\mathbf{e}_j$ der $j$-te Einheitsvektor mit 1 genau an der Stelle $j$ und sonst 0 sei.

$f$ heißt partiell differenzierbar, wenn alle $n$ partiellen Ableitungen existieren.

Jetzt können wir zumindestens entlang der Achsen im Koordinaten-System Tangenten anlegen und entsprechend nach den verschiedenen Variablen ableiten. Es bleibt aber die Frage, was denn mit Tangenten ist, die nicht entlang der Achsen verlaufen.

Wir können also schon mal fragem wie schräg der »fliegende Teppich« entlang der $x$- und $y$-Achsen steht, aber nicht, wie schräg der Teppich zum Beispiel entlang der Winkelhalbierenden von $x$- und $y$-Achse steht.

**Beispiel**

Die Funktion

$$f(x,y) = xe^{x+y}$$

ist partiell differenzierbar.

$$\begin{aligned}
f_x(x,y) &= \partial_1 f(x,y) = (1+x)e^{x+y} \\
f_y(x,y) &= \partial_2 f(x,y) = xe^{x+y} \\
f_{x,y}(x,y) &= \partial_2 \partial_1 f(x,y) = (1+x)e^{x+y} \\
f_{y,x}(x,y) &= \partial_1 \partial_2 f(x,y) = (1+x)e^{x+y} \\
f_{x,x}(x,y) &= (2+x)e^{x+y} \\
f_{y,y}(x,y) &= xe^{x+y}
\end{aligned}$$

Wir verallgemeinern das Konzept der partiellen Ableitungen nun zu einer Richtungsableitung. Wir werden sehen, dass die partiellen Ableitungen Richtungsableitungen in Richtung der Einheitsvektoren sind.

**Definition 3.14** Richtungsableitung

Sei $\Omega$ ein Gebiet, $f : \Omega \to \mathbb{R}$m, $\mathbf{a} \in \Omega$ fest und $\mathbf{v} \neq \mathbf{0}$ beliebiger Vektor in $\mathbb{R}^n$ mit $\|\mathbf{v}\|_2 = 1$. Existiert der Grenzwert

$$\partial_{\mathbf{x}} f(\mathbf{a}) = \lim_{h \to 0} \frac{1}{h} (f(\mathbf{a} + h\mathbf{v}) - f(\mathbf{a}))$$

so heißt er »Richtungsableitung« von $f$ in $\mathbf{a}$ in Richtung $\mathbf{v}$.

Wir sagen auch »$f$ ist in $\mathbf{a}$ in Richtung $\mathbf{v}$ differenzierbar«.

Wichtig an dieser Definition ist die Normierung der Länge des Vektors auf die Länge 1. Sonst erhält man für unterschiedlich lange Vektor natürlich ganz andere Differenzen.

Anschaulich: $\varphi(t) := \mathbf{a} + f\mathbf{v}$ mit $t \in \mathbb{R}$ stellt geometrisch eine Gerade im $\mathbb{R}^n$ dar.

Die Funktion $f_{\mathbf{a},\mathbf{v}}(t) = f \circ \varphi(t) = f(\mathbf{a} + t\mathbf{v})$ hat eine Veränderliche $t \in \mathbb{R}$. Für die Richtungsableitung ergibt sich

$$\partial_{\mathbf{v}} f(\mathbf{a}) = f'_{\mathbf{a},\mathbf{v}}(0) \qquad (*)$$

**Lemma 3.15**

a) $f \longrightarrow \partial_{\mathbf{v}} f(\mathbf{a})$ ist linear in $f$.

b) Produktregel: $\partial_{\mathbf{v}}(f \cdot g)(\mathbf{a}) = f(\mathbf{a})\partial_{\mathbf{v}} g(\mathbf{a}) + g(\mathbf{a})\partial_{\mathbf{v}} f(\mathbf{a})$

**Beweis** a) Folgt sofort aus $(*)$.

b) $\partial_{\mathbf{v}}(f \cdot g)(\mathbf{a}) = (f \cdot g)'_{\mathbf{a},\mathbf{v}}(0) = (f_{\mathbf{a},\mathbf{v}} \cdot g_{\mathbf{a},\mathbf{v}})'(0) = f_{\mathbf{a},\mathbf{v}}(0)g'_{\mathbf{a},\mathbf{v}}(0) + f'_{\mathbf{a},\mathbf{v}}(0)g_{\mathbf{a},\mathbf{v}}(0)$

$\square$

Wir wollen zwar verstehe, dass man die Ableitung in jede Richtung bestimmen kann. Wichtig sind aber vor allem die partiellen Ableitungen. Die Richtungsableitungen werden wir als Linearkombination der partuellen Ableitungen verstehen.

**Definition 3.16** gradient

Sei $\Omega \subset \mathbb{R}^n$ ein Gebiet, $f : \Omega \to \mathbb{R}$ partiell differenzierbar. Der Vektor

$$\operatorname{grad} f(\mathbf{x}) = (\partial_1 f(\mathbf{x}), \partial_2 f(\mathbf{x}), \ldots, \partial_n f(\mathbf{x}))$$

heißt Gradient von $f$ im Punkt $\mathbf{x} \in \Omega$.

$$\nabla := (\partial_1, \partial_2, \ldots, \partial_n)$$

heißt *Nabla-Operator*. Dass heißt

$$\nabla f := (\partial_1 f, \partial_2 f, \ldots, \partial_n f)\,.$$

**Vorlesung XXVI**

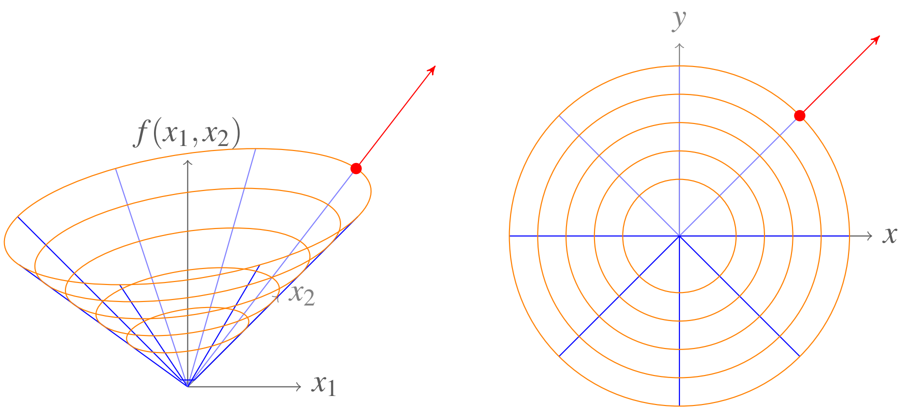
### Beispiel

$$
\begin{aligned}
f(x_1,x_2) &= \sqrt{x_1^2+x_2^2} \\
\forall k \in \{1,2\}: \ \partial_k f(x_1,x_2) &= \frac{x_k}{\sqrt{x_1^2+x_2^2}} \\
\operatorname{grad} f(x_1,x_2) &= \left(\frac{x_1}{\sqrt{x_1^2+x_2^2}}, \frac{x_2}{\sqrt{x_1^2+x_2^2}}\right) = \frac{1}{\sqrt{x_1^2+x_2^2}}(x_1,x_2)
\end{aligned}
$$

Vektoriell aufgeschrieben gilt analog

$$
\begin{aligned}
\mathbf{x} &= (x_1,x_2) \\
f(\mathbf{x}) &= \|\mathbf{x}\|_2 \\
\nabla f(\mathbf{x}) &= \frac{1}{\|\mathbf{x}\|_2}\mathbf{x}
\end{aligned}
$$

Der Gradient der 2-Norm im zweidimentionalen ist also der Vektor selber normiert auf die Länge 1.




Wir sehen die Niveau-Mengen (»Höhenlinien«) auf den Niveaus $1/2, 3/4, 1, 5/4$ und $3/2$ und den Gradienten im Punkt $(\sqrt{9/8}, \sqrt{9/8}, 3/2)$ als Vektor der Länge 1 in Richtung des steilsten Anstiegs[1], also genau konzentrisch von der Mitte weg.

**Satz 3.17** Satz von Schwarz

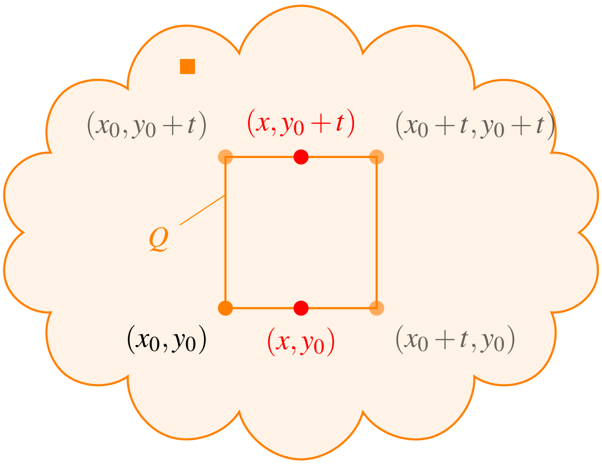
Hermann Aramandus Schwarz, 1843–1921

Sei ein Gebiet $\Omega \subseteq \mathbb{R}^n$ und eine Funktion $f : \Omega \to \mathbb{R}$ gegeben. Weiter sei $f$ zweimal stetig partiell differenzierbar, d. h. für alle $1 \leq p, q \leq n$ existiert eine partielle Ableitung $\partial_p \partial_q f$ und sind stetig in $\mathbf{x}_0$. Dann gilt

$$\partial_p \partial_q f(\mathbf{x}_0) = \partial_q \partial_p f(\mathbf{x}_0)$$

**Beweis** Um den Beweis technisch zu vereinfachen erfolgt der Beweis ohne Beeinträchtigung der Allgemeinheit für $n = 2$.

Sei $(x_0, y_0) \in \Omega$ und $t$ so klein, dass das Quadrat $Q$ mit den Ecken $(x_0, y_0), (x_0, y_0 + t), (x_0 + t, y_0)$ und $(x_0 + t, y_0 + t)$ ganz in $\Omega$ liege.




Für den Beweis bilden wir jetzt zwei Hilfsfunktionen

$$\begin{aligned}
g(x) &= f(x, y_0 + t) - f(x, y_0) \\
g'(x) &= f_x(x, y_0 + t) - f_x(x, y_0) \qquad (*) \\
h(t) &= f(x_0 + t, y_0 + t) - f(x_0 + t, y_0) - f(x_0, y_0 + t) + f(x_0, y_0) \\
&= g(x_0 + t) - g(x_0) \overset{\text{MWS}}{=} g'(\xi_1)t
\end{aligned}$$

Denn nach dem Mittelwertsatz [Verweis] existiert ein $\xi_1$ mit $x_0 < \xi_1 < x_0 + t$.

Wir haben hier den Mittelwertsatz in $X$-Richtung angewandt.

$$\begin{aligned}
&\overset{(*)}{=} (f_k(\xi_1, y_0 + t) - f_k(\xi_1, y_0))t \\
&\overset{\text{MWS}}{=} (f_{x,y}(\xi_1, \eta_1))t^2
\end{aligned}$$

Denn nach Mittelwertsatz [Verweis] existiert ebenfalls ein $\eta_1$ mit $y_0 < \eta_1 < y_0 + t$.

Jetzt haben wir den Mittelwertsatz zum zweiten Mal, aber jetzt in $Y$-Richtung angewandt.

Symmetrisch analog kann die Rolle von $x$ und $y$ vertauscht werden:

$$h(t) = f_{y,x}(\xi_2, \eta_2)t^2, \quad (\xi_2, \eta_2) \in Q$$

Wir betrachten jetzt den Grenzübergang $t \to 0$.

Was apssiert, wenn dieses Quadrat nun auf einen einzigen Punkt zusammenschrumpft. Dann müssen noch immer die Punkte $(\xi_1, \eta_1)$ und $(\xi_2, \eta_2)$ im Quadrat liegen, nur dass das Quadrat jetzt ein Punkt ist.

$$\lim_{t \to 0} h(t) = f_{x,y}(x_0, y_0) = f_{y,x}(x_0, y_0)$$

□

In den meisten Fällen sind die zweiten partiellen Ableitungen auch stetig, sodass man davon ausgehen kann, dass die Reihenfolge der Ableitungen egal ist. Man kann aber auch recht einfach kompliziertere Funktionen aufstellen, bei denen das nicht funktioniert.

### 3.2.2 totale Ableitung

Mit den partiellen Ableitungen kann man die eindimensionale Ableitung nicht so richtig ins multivariable Übertragen, da man immer nur die Ableitung in eine Richtung erhält. Wir werden dieses Problem jetzt umgehen, indem wir sehr kanonisch die eindimensionale Ableitung ins multivairable übetragen.

Zur Erinnerung: für $\Omega \subseteq \mathbb{R}^1, f : \Omega \to \mathbb{R}$ ist differenzierbar in $x_0 \in \Omega$, falls

$$\lim_{h \to 0} \frac{1}{h}(f(x_0 + h) - f(x_0))$$

exsitiert.

Wir schreiben diese Definition jetzt ein wenig um, sodass eine Form entsteht, die sich deutlich besser ins Multivariable überführen lässt.

Äquivalent ist die Aussage

$$\frac{1}{h}(f(x_0 + h) - f(x_0)) = a + r_1(x_0, h) \text{ mit } r_1(x_0, h) = \mathrm{o}(1) \text{ für } h \to 0$$

Wenn wir also so ein $a$ finden, dass diese Aussage gilt, dann ist $a = f'(x_0)$. Wenn $h$ gegen 0 geht, soll der Rest möglichst klein werden, denn genau dass verlangt $\mathrm{o}(1)$.

Weiter umgebaut ist die Aussage äquivalent zu

$$f(x_0+h) = f(x_0) + ha + r(x_0,h) \text{ mit } r(x_0,h) = o(h) \text{ für } h \to 0$$

In dieser Form bedeutet die Aussage, $f$ kann in Umgebung von $x_0$ durch lineare Funktion $f(x_0)+ha$ approximiert werden und es ist $a = f'(x_0)$. Diese Form ist in das Mehrdimensionale übertragbar.

**Definition 3.18** totale Ableitung

Es sei $\Omega \subset \mathbb{R}^n$ und $\mathbf{f}: \Omega \to \mathbb{R}^m$ und $x_o \in \Omega$ sei ein innerer Punkt.

$\mathbf{f}$ heißt *(total) differenzierbar* in $x_0 \in \Omega$, falls eine $m \times n$-Matrix $A \in \mathbb{R}^{m\times n}$ existiert, sodass in einer Umgebung von $\mathbf{x}_0$ gilt

$$\mathbf{f}(\mathbf{x}_0+\mathbf{h}) = \mathbf{f}(x_0) + \mathbf{h}A^T + \mathbf{r}(\mathbf{x}_0,\mathbf{h}) \text{ mit } \mathbf{r}(\mathbf{x}_0,\mathbf{h}) = o(\|\mathbf{h}\|) \text{ für } \|\mathbf{h}\| \to 0$$

Die Norm muss irgendeine Norm im $\mathbb{R}^n$ sein. Da in endlich dimensionalen Vektorräumen eh alle äquivalent sind, ist die genaue Wahl einer Norm egal.

$\mathbf{f}'(\mathbf{x}_0) = A$ heißt *(totale) Ableitung*.

Wir sehen also, es hat geklappt. Wir haben jetzt eine kanonische Verallgemeinerung der Ableitung im Eindimensionalen für Mehrdimensionale Funktionen.

**Bemerkung**

1. Wenn $\mathbf{f}'(\mathbf{x}_0)$ existiert, ist es eindeutig bestimmt.

   Angemommen es existiert ein $B$, dann kann man die Eindeutigkeit zeigen über

   $$\mathbf{0} = \mathbf{0} + \mathbf{h}(A^T - B^T) + \mathbf{r}_A - \mathbf{r}_B$$

2. Abbildung $\mathbf{f} \to \mathbf{f}'$ ist linear.
3. Aus der Existenz der *partiellen* Ableitung folgt *nicht* die Stetigkeit. Aus der *totalen* Ableitung folgt die Stetigkeit.

**Satz 3.19**

Ist $\Omega \subset \mathbb{R}^n$ und $\mathbf{f}: \Omega \to \mathbb{R}^m$ in $\mathbf{x}_0 \in \Omega$ diferenzierbar, dann ist $\mathbf{f}$ in $\mathbf{x}_0$ stetig.

**Beweis** Vorbemerkung zu Normen in $\mathbb{R}^n$ und $\mathbb{R}^m$ und lineare Operatoren $A: \mathbb{R}^n \to \mathbb{R}^m$ mit $\mathbf{x} \mapsto \mathbf{x}A^T$:

$$\|A\|_{\mathbb{R}^n\to\mathbb{R}^m} := \sup_{x\in\mathbb{R}^n, \mathbf{x}\neq\mathbf{0}} \frac{\|\mathbf{x}A^T\|_{\mathbb{R}^m}}{\|\mathbf{x}\|_{\mathbb{R}^n}} = \sup_{\|\mathbf{x}\|_{\mathbb{R}^m}=1} \|\mathbf{x}A^T\|_{\mathbb{R}^m} < \infty$$

Daraus folgt

$$\|\mathbf{x}A^T\|_{\mathbb{R}^m} \leq \|A\|_{\mathbb{R}^n \to \mathbb{R}^m} \cdot \|\mathbf{x}\|_{\mathbb{R}^n} \quad \forall \mathbf{x} \in \mathbb{R}^n$$

Denn $\|A\|_{\mathbb{R}^n \to \mathbb{R}^m}$ ist als das Supremum definiert!

Es sei $\mathbf{x}$ aus Umgebung von $\mathbf{x}_0$. Dann gilt

$$\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{x}_0) = (\mathbf{x} - \mathbf{x}_0)(\mathbf{f}'(\mathbf{x}_0))^T + \mathbf{r}(\mathbf{x}_0, \mathbf{h}).$$

Die eindimensionale Überlegegung war

- $x - x_0 \to 0$
- $r(x_0, h) \to 0$
- Also geht die Summe gegen 0 und damit auch die Differenz der Funktionswerte.

Also gilt

$$\begin{aligned}
\|\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{x}_0)\|_{\mathbb{R}^m} &\leq \|(\mathbf{x} - \mathbf{x}_0)(\mathbf{f}'(\mathbf{x}_0))\|_{\mathbb{R}^m} + \|\mathbf{r}(\mathbf{x}_0, \mathbf{h})\|_{\mathbb{R}^m} \\
&\leq \|\mathbf{x} - \mathbf{x}_0\|_{\mathbb{R}^n} \cdot \|\mathbf{f}'(\mathbf{x}_0)^T\|_{\mathbb{R}^n \to \mathbb{R}^m} + \|\mathbf{r}(\mathbf{x}_0, \mathbf{h})\|_{\mathbb{R}^m}
\end{aligned}$$

$\square$

Jetzt haben wir eine Matrix, die wir Ableitung nennen. Allerdings brauchen wir jetzt noch einige Rechenregeln, wie wir auf diese Matrix kommen, und was wir mit dieser Matrix anfangen können.

**Vorlesung XXVII**

Die Richtungsableitung für den Fall $m = 1$ hat schon die Definition 3.14 auf Seite 85 behandelt. Aber auch das können wir weiter verallgemeinern.

**Definition 3.20** allgemeine Richtungsableitung

Sei $\mathbf{f} : \mathbb{R}^n \supseteq \Omega \to \mathbb{R}^m$ und $x_0 \in \Omega$ ein innerer Punkt. Ferner sei $\mathbf{v} \in \mathbb{R}^n$ mit $\|\mathbf{v}\| = 1$ eine Richtung.

$\mathbf{f}$ besitzt eine *Richtungsableitung* in Richtung $\mathbf{v}$ im Punkt $\mathbf{x}_0$, falls

$$\frac{\partial \mathbf{f} f}{\partial \mathbf{v}}(\mathbf{x}_0) := \lim_{h \to 0} \frac{1}{h} (\mathbf{f}(\mathbf{x}_0 + h\mathbf{v}) - \mathbf{f}(\mathbf{x}_0))$$

existiert. $\dfrac{\partial \mathbf{f}}{\partial \mathbf{v}}$ heißt Richtungsableitung.

Zur Berechnung der Richtungsableitung und Ableitung kommen wir in folgendem Satz.

### Satz 3.21

Sei $\mathbf{f}: \mathbb{R}^n \supseteq \Omega \to \mathbb{R}^m$ und $\mathbf{x}_0 \in \Omega$ ein innerer Punkt.

a) Ist $\mathbf{f}$ differenzierbar in $\mathbf{x}_0$, so existiert die Richtungsableitung in jede Richtung $\mathbf{v} \in \mathbb{R}^n$. Ferner ist

$$\frac{\partial \mathbf{f}}{\partial \mathbf{v}}(\mathbf{x}_0) = \mathbf{v}\mathbf{f}'(\mathbf{x}_0)^T.$$

b) Ist $\mathbf{f}$ in einer Umgebung von $\mathbf{x}_0$ stetig partiell differenzierbar, dann ist $\mathbf{f}$ in $\mathbf{x}_0$ auch diferenzierbar und es gilt

$$\mathbf{f}'(\mathbf{x}_0) = J_{\mathbf{f}}(\mathbf{x}_0)$$

mit

$$J_{\mathbf{f}}(\mathbf{x}_0) = \left(\frac{\partial f_i}{\partial x_j}(\mathbf{x}_0)\right)_{i=1,\dots,m;j=1,\dots,n} = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(\mathbf{x}_0) & \cdots & \frac{\partial f_1}{\partial x_n}(\mathbf{x}_0) \\ \vdots & & \vdots \\ \frac{\partial f_m}{\partial x_1}(\mathbf{x}_0) & \cdots & \frac{\partial f_m}{\partial x_n}(\mathbf{x}_0) \end{pmatrix} \in \mathbb{R}^{m\times n}.$$

$J_{\mathbf{f}}$ heißt *Jacobi-Matrix* oder *Funktionalmatrix*.

Carl Gustav Jacobi lebte von 1804–1851.

Das ist jetzt endlich eine Bildungsvorschrift für die Matrix $A$ aus der Definition 3.18 auf Seite 89

Wir haben hier eine Verallgemeinerung des Gradienten auf mehrere Dimensionen. Denn eine Zeile der Matrix entspricht immer genau einem Gradienten.

Kurz: Stetig partiell differenzierbar $\overset{b)}{\Rightarrow}$ differenzierbar $\overset{a)}{\Rightarrow}$ partiell differenzierbar und in alle Richtungen differenzierbar.

**Beweis** a) Für $\mathbf{v} \in \mathbb{R}^n, \|\mathbf{v}\| = 1$ folgt

$$\mathbf{f}(\mathbf{x}_0 + h\mathbf{v}) - \mathbf{f}(\mathbf{x}_0) = h\mathbf{v}\mathbf{f}'(\mathbf{x}_0)^T + \mathbf{r}(\mathbf{x}_0, h\mathbf{v})$$

mit

$$\lim_{h\to 0} \frac{1}{h}\mathbf{r}(\mathbf{x}_0, h\mathbf{v}) = 0$$

folgt die Behauptung.

b) Wir beweisen zunächst die Aussage für $m = 1$. Dass heißt die Funktion $f$ liefert wieder einen reellen Wert und keinen Vektor. Weiter schreiben wir $\mathbf{a} := \mathbf{x}_0$.

$$\begin{aligned} f(\mathbf{x}) - f(\mathbf{a}) = & \left(f(x_1, x_2, \dots x_n) - f(x_1, \dots, x_{n-1}, a_n)\right) \\ & + \left(f(x_1, \dots, x_{n-1}, a_n) - f(x_1, \dots, x_{n-2}, a_{n-1}, a_n)\right) \\ & + \dots \\ & + \left(f(x_1, a_2, \dots, a_n) - f(a_1, \dots, a_n)\right) \end{aligned}$$

Jetzt sieht der Ausdruck nicht nur viel komplizierter aus, sondern wir können auch in jeder Klammer den *eindimensionalen* Mittelwertsatz anwenden.

$$f(\mathbf{x})-f(\mathbf{a})=\frac{\partial f}{\partial x_n}(\varepsilon_n)(x_n-a_n)+\frac{\partial f}{\partial x_{n-1}}(\varepsilon_{n-1})(x_{n-1}-a_{n-1})+\ldots+\frac{\partial f}{\partial x_1}(\varepsilon_1)(x_1-a_1)$$

wobei $\varepsilon_k=(x_1,\ldots,x_{k-1},a_k+\delta_k(x_k-a_k),a_{n+1},\ldots,a_n)$ mit $0<\delta_k<1$.

Es folgt

$$\begin{aligned}
&\frac{f(\mathbf{x})-f(\mathbf{a})-(\mathbf{x}-\mathbf{a})(\operatorname{grad} f(a))^T}{\|\mathbf{x}-\mathbf{a}\|_1}\\
&=\frac{1}{\|\mathbf{x}-\mathbf{a}\|_1}\sum_{k=1}^{n}\left(\frac{\partial f}{\partial x_k}(\varepsilon_k)-\frac{\partial f}{\partial x_k}(\mathbf{a})\right)(x_k-a_k)\\
&\cdot\left|\frac{1}{\|\mathbf{x}-\mathbf{a}\|_1}\sum_{k=1}^{n}\left(\frac{\partial f}{\partial x_k}(\varepsilon_k)-\frac{\partial f}{\partial x_k}(\mathbf{a})\right)(x_k-a_k)\right|\\
&\leq\underbrace{\max_{1\leq k\leq n}\left|\frac{\partial f}{\partial x_k}(\varepsilon_k)-\frac{\partial f}{\partial x_k}(\mathbf{a})\right|}_{\to 0 \text{ für } x\to a}\cdot\underbrace{\frac{1}{\|\mathbf{x}-\mathbf{a}\|_1}\sum_{k=1}^{n}|x_k-a_k|}_{1}
\end{aligned}$$

Die 1-Norm wird hier nur verwendet, weil diese besonders einfach ist. Da aber eh alle Normen gleich sind (im endlich dimensionalen Raum), können wir irgendeine Norm verwenden.

Also ist $f$ differenzierbar in $\mathbf{a}$ und $f'(\mathbf{a})=\operatorname{grad} f(a)$.

Für $m\geq 2$ ist der Beweis die komponentenweise Betrachtung analog zum Fall $m=1$.

□

**Beispiel**

1. $f:\mathbb{R}^2\to\mathbb{R}^3$ mit $f(x,y)=\big(\cos(x+2y),\sin(x+3y),x^2+y^2\big)$.

   Diese Funktion ordnet jedem Punkt in der Eben – also zum Beispiel auf einer Landkarte – einen dreidimensionalen Vektor zu – also zum Beispiel eine Blickrichtung.

$$J_{\mathbf{f}}=\begin{pmatrix}\frac{\partial f_1}{\partial x} & \frac{\partial f_1}{\partial y}\\ \frac{\partial f_2}{\partial x} & \frac{\partial f_2}{\partial y}\\ \frac{\partial f_3}{\partial x} & \frac{\partial f_3}{\partial y}\end{pmatrix}=\begin{pmatrix}-\sin(x+2y) & -2\sin(x+2y)\\ \cos(x+3y) & 3\cos(x+3y)\\ 2x & 2y\end{pmatrix}$$

2. a) Richtungsableitung von $f(\mathbf{x}) = \|\mathbf{x}\|_2$ in Richtung $\mathbf{v} = \frac{1}{\sqrt{n}}(1,\dots,1)$

Wir halten zunächst fest:

$$J_\mathbf{f}(\mathbf{x}) = \operatorname{grad} f(\mathbf{x}) = \nabla f(\mathbf{x}) = \left(\frac{x_1}{\|\mathbf{x}\|_2},\dots,\frac{x_n}{\|\mathbf{x}\|_2}\right)$$

Diese umfassende Gleichheit von all diesen Dingen funktioniert nur, da diese Funktion in den eindimensionalen Raum abbildet. Dann entspricht der Gradient gerade der Jacobi-Matrix.

$$\begin{aligned}\frac{\partial f}{\partial \mathbf{v}} &= \mathbf{v}(J_\mathbf{f}(\mathbf{x}))^T = \frac{1}{\sqrt{n}}(1,\dots,1)\left(\frac{x_1}{\|\mathbf{x}\|_2},\dots,\frac{x_n}{\|\mathbf{x}\|_2}\right)^T \\ &= \frac{1}{\sqrt{n}}(1,\dots,1)\frac{\mathbf{x}^T}{\|\mathbf{x}\|_2} = \frac{1}{\sqrt{n}\|\mathbf{x}\|_2}\sum_{k=1}^n x_k\end{aligned}$$

b) Richtungsableitung gleicher Funktion $f$ in Richtung $\mathbf{v} = \mathbf{e}_j$

$$\frac{\partial f}{\partial \mathbf{v}} = \mathbf{e}_j(J_\mathbf{f}(\mathbf{x}))^T = \mathbf{e}_j(\operatorname{grad} f(\mathbf{x}))^T = \frac{x_j}{\|\mathbf{x}\|_2}$$

An dieser Stelle dürfen wir uns schwer bestätigt fühlen, denn bereits im Beispiel nach Definition 3.16 auf Seite 85 haben wir den Gradienten der 2-Norm als den Vektor selber auf die Länge 1 normiert erkannt. Bei der parteillen Ableitung oder wie hier analog bei der Richtungsableitung entlang der Koordinatenachsten (also in Richtung eines Einheitsvektors, was eben gerade einer partiellen Ableitung entspricht) entsteht also eine Komponente des auf die Länge 1 normierten Vektors.

c) Die Richtung des steilsten Anstiegs für $f : \mathbb{R}^n \to \mathbb{R}^1$ im Punkt $\mathbf{x}$ ist $\operatorname{grad} f(\mathbf{x})$.

Wir betrachten also zum Beispiel für $n = 2$ wieder den fliegenden Teppich über der $x,y$-Ebene und fragen uns, in welche Richtung es an einem Punkt $\mathbf{x}$ am steilsten nach oben geht.

Begründung:

$$\begin{aligned}\frac{\partial f}{\partial \mathbf{v}}(\mathbf{x}) &= \mathbf{v}\operatorname{grad} f(\mathbf{x})^T = \langle \mathbf{v}, \operatorname{grad} f(\mathbf{x})\rangle_2 \\ &= \|\mathbf{v}\|_2 \|\operatorname{grad} f(\mathbf{x})\|_2 \cos\theta \\ &= \|\operatorname{grad} f(\mathbf{x})\|_2 \cos\theta \\ &\le \|\operatorname{grad} f(\mathbf{x})\|_2\end{aligned}$$

Denn die Ableitung von $f$ in Richtung $\mathbf{v}$ ist ein Maß für den Anstieg des Funktionsgraphen in Richtung $\mathbf{v}$. Der Wert dieser Ableitung ist nun genau dann am größten, denn der Winkel $\theta$ zwischen dem Gradientenvektor und dem Richtungsvektor $v$ 0 ist (also $\cos\theta = 1$). Die Richtungsableitung in einem bestimmen Punkt ist also in Richtung des Gradienten maximal. Der Anstieg ist genau dann am größten, wenn man in Richtung des Gradienten geht.

**Vorlesung XXVIII**

### 3.2.3 Übertragungen ins Multivariable

In diesem Abschnitt werden wir einige Konzepte aus dem Eindimensionalen ($f : \mathbb{R} \to \mathbb{R}$) ins Mutltivariable (hier $f : \mathbb{R}^n \to \mathbb{R}$) übertragen. Dies erleichtert uns die Handhabung von multivariablen Funktionen.

Analog zur Tangenten bei einer Funktion $\mathbb{R} \to \mathbb{R}$ entsteht bei einer multivariablen Funktion eine Tangentialebene. Dieser Ebene berührt den fliegenden Teppich in genau einem Punkt und entspricht in diesem Punkt der »Steigung« (die es in der Form im multivariablen natürlich nicht gibt) der Funktion.

Wir betrachten $f : \mathbb{R}^2 \supseteq \Omega \to \mathbb{R}$, $f$ sei differenzierbar in $(x_0, y_0)$. Dann gilt

$$f(x,y) = f(x_0,y_0) + (x-x_0, y-y_0)\left(\frac{\partial f}{\partial x}(x_0,y_0), \frac{\partial f}{\partial y}(x_0,y_0)\right)^T + r((x_0,y_0),(x-x_0,y-y_0)).$$

Die Tangentialebene an $f$ in $(x_0, y_0)$

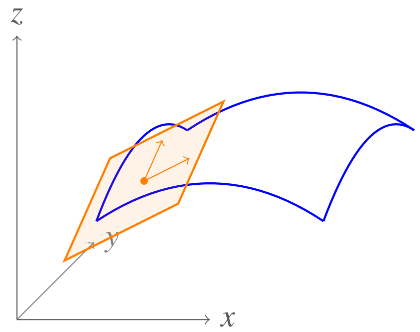
$$\begin{aligned} z = g(x,y) &= f(x_0,y_0) + (x-x_0, y-y_0)\left(\frac{\partial f}{\partial x}(x_0,y_0), \frac{\partial f}{\partial y}(x_0,y_0)\right)^T \\ &= f(x_0,y_0) + (x-x_0) f_x(x_0,y_0) + (y-y_0) f_y(x_0,y_0) \end{aligned}$$

Wir sehen uns also in einem Punkt $(x_0, y_0)$ an, wie stark der fliegende Teppich in die verschiedenen Koordinatenrichtungen ansteigt. Wir berühren also einmal den Teppich und schweben ansonsten darüber.

Alle Richtungsableitungen von $f$ und $g$ stimmen im Punkt $(x_0, y_0)$ überein.

$$\begin{aligned} E &= \left\{(x,y,g(x,y)) \mid (x,y) \in \mathbb{R}^2\right\} \\ &= \left\{(1,0,f_x(x_0,y_0))(x-y_0) + (0,1,f_y(x_0,y_0))(y-y_0) + (x_0,y_0,f(x_0,y_0)) \mid (x,y) \in \mathbb{R}^2\right\} \end{aligned}$$

ist die Gleichung der Tangentialebene an $f$ im Punkt $(x_0, y_0) \in \mathbb{R}^2$.




Wir sehen in der Skizze in Blau den Graphen der Funktion als fliegenden Teppich über der $x,y$-Ebene, sowie in Orange den Punkt $(x_0, y_0, f(x_0, y_0))$ und die partiellen Steigungen an diesem Punkt in $x$- und $y$-Richtung. Mit dem Punkt und diesen beiden Richtungsvektoren ist die orange Ebene vollständig definiert. Die Ebene berüht nur genau in dem orangen Punkt den Graphen der Funktion.

Wie zu jeder Ebene im $\mathbb{R}^2$ kann man über das Kreuzprodukt der Richtungsvektoren den Normalenvektor bestimmen.

Der Normalenvektor dieser Ebene lautet $\mathbf{n} = (f_x(x_0, y_0), f_y(x_0, y_0), -1)$.

**Satz 3.22** Kettenregel

Es seien $\mathbf{f} : \mathbb{R}^n \supseteq \Omega \to \mathbb{R}^m$ differenzierbar in $\mathbf{x}_0 \in \Omega$ und $\mathbf{g} : \mathbb{R}^m \supseteq D \to \mathbb{R}^k$. Es gelte $f(\Omega) \subseteq D$ und $\mathbf{g}$ sei differenzierbar in $\mathbf{y}_0 = f(\mathbf{x}_0) \in D$.

Dann ist

$$\mathbf{h} = \mathbf{g} \circ \mathbf{f} : \Omega \to \mathbb{R}^k \text{ ist differenzierbar in } \mathbf{x}_0 \in \Omega$$

und es gilt

$$(\mathbf{g} \circ \mathbf{f})'(\mathbf{x}_0) = \underbrace{g'(\mathbf{f}(\mathbf{x}_0))}_{\in \mathbb{R}^{n \times m}} \cdot \underbrace{\mathbf{f}'(\mathbf{x}_0)}_{\in \mathbb{R}^{m \times n}}.$$

**Beweis** verläuft wie im Eindimensionalen.

Im Eindimensionalen wird ein Grenzwert betrachtet: Hier $n$-dimensioanler Vektor bestehend aus einzelnen Folgen. Dieser Vektor konvergiert gegen den Nullpunkt.

Mit dieser Mehrdimensionalen Folge kann dann die gleiche Argumentation wie im Eindmimensionalen ausgeführt werden.

□

**Bemerkung** $\mathbf{g}'(\mathbf{f}(\mathbf{x}_0)) \cdot \mathbf{f}'(\mathbf{x}_0)$ ist ein Matrixprodukt und daher nicht vertauschbar (kommutativ).

**Beispiel** Umrechnen in Polarkoordinaten

Wir wollen also einen Punkt $(x,y)$ umrechnen in einen Winkel und einen Abstand vom Nullpunkt.

Sei $g:\mathbb{R}^2\to\mathbb{R}$ eine differenzierbare Abbildung, etwa $g(x,y)=e^{x+y}$, und $x=r\cos\varphi$ und $y=r\sin\varphi$. Dass heißt

$$g(r\cos\varphi, r\sin\varphi)=e^{r(\cos\varphi+\sin\varphi)}$$

Frage: $\dfrac{\partial}{\partial r}=?\quad \dfrac{\partial}{\partial\varphi}=?$

Wir wollen jetzt nicht einfach die Funktion $g$ nach $r$ und $\varphi$ partiell Ableiten. Denn auch wenn wir das hier noch könnten, kann das schnell sehr unangenehm werden. Die Kettenregel vereinfacht hier einiges.

$$f:\mathbb{R}^2\to\mathbb{R}^2 \text{ mit } f(r,\varphi)=(r\cos\varphi, r\sin\varphi)=(x,y)$$

$$J_f=\begin{pmatrix}\cos\varphi & r\sin\varphi\\ \sin\varphi & r\cos\varphi\end{pmatrix},\qquad J_g=\operatorname{grad} g=(e^{x+y},e^{x+y})$$

Das ging also sehr einfach. Die gesamte Ableitung zu bilden wäre deutlich mehr ausgeartet.

$$\begin{aligned}J_{g\circ f}(r,\varphi)&=\left(\frac{\partial}{\partial x}g(f(r,\varphi)),\frac{\partial}{\partial y}g(f(r,\varphi))\right)\cdot\begin{pmatrix}\cos\varphi & r\sin\varphi\\ \sin\varphi & r\cos\varphi\end{pmatrix}\\&=\left(\frac{\partial}{\partial x}g(f(r,\varphi))\cos\varphi+\frac{\partial}{\partial y}g(f(r,\varphi))\sin\varphi\,,\right.\\&\qquad\left.-\frac{\partial}{\partial x}g(f(r,\varphi))r\sin\varphi+\frac{\partial}{\partial y}g(f(r,\varphi))r\cos\varphi\right)\end{aligned}$$

Kurz:

$$\begin{aligned}\frac{\partial}{\partial r}&=\cos\varphi\frac{\partial}{\partial x}+\sin\varphi\frac{\partial}{\partial y}\\ \frac{\partial}{\partial\varphi}&=-r\sin\varphi\frac{\partial}{\partial x}+r\cos\varphi\frac{\partial}{\partial y}\end{aligned}$$

Wir können also jetzt die Kettenregel im Mehrdimensionalen. Versuchen wir uns im folgenden Mal an einer anderen nützlichen Eigenschaft aus dem Eindimensionalen: dem Mittelwertsatz.

Wir betrachten den Mittelwertsatz nur für Funktionen, die in die reellen Zahlen abbilden. Für Funktionen, die ins Mehrdimensionale Abbilden wird das ganze einfach komponentenweise nochmal betrachtet.

**Satz 3.23** Mittelwertsatz für Funktion $f : \mathbb{R}^n \supseteq \Omega \to \mathbb{R}$

Es sei $\Omega \subseteq \mathbb{R}^n$ offen und $f : \Omega \to \mathbb{R}$ differenzierbar. Sind $\mathbf{x}, \mathbf{y} \in \Omega$ mit $\{\mathbf{x} + t(\mathbf{y} - \mathbf{x}) \mid t \in [0,1]\} \subseteq \Omega$ (hinreichend: Konvexität von $\Omega$), dann existiert $\xi \in (0,1)$ mit

$$f(\mathbf{y}) - f(\mathbf{x}) = f'(\mathbf{x} + \xi(\mathbf{y} - \mathbf{x}))(\mathbf{y} - \mathbf{x})^T = (\mathbf{y} - \mathbf{x}) \operatorname{grad} f(\mathbf{x} + \xi(\mathbf{y} - \mathbf{x}))^T$$

Wenn ein Mathematiker in einem differenzierbaren Gebirge auf einer Geraden im Definitionsbereich (Mathematiker laufen immer auf Geraden) wandert, dann weiß er, dass er in kleinen Schritten vor und zurück gehen kann, und auch nur in kleinen Schritten hoch oder runter gehen muss. Er wird nie abrupt runter fallen oder hoch springen.

**Beweis** Sei $h(t) = f(\mathbf{x} + t(\mathbf{yx}))$, dann ist $h : [0,1] \to \mathbb{R}$. Nach dem Mittelwertsatz im Eindimensionalen gilt

$$f(\mathbf{y}) - f(\mathbf{x}) = h(1) - h(0) \overset{MWS}{=} h'(\xi) \cdot 1$$

Mit der Kettenregel folgt

$$h'(\xi) = f'(\mathbf{x} + \xi(\mathbf{y} - \mathbf{x}))(\mathbf{y} - \mathbf{x})^T$$

$\square$

**Taylor-Formel** (entwickelt in $(x_0, y_0), f : \mathbb{R}^2 \to \mathbb{R}$)

$$\begin{aligned}
&0.\ \text{Ordnung:} f(x,y) \approx f(x_0,y_0)\\
&1.\ \text{Ordnung:} f(x,y) \approx f(x_0,y_0) + (x-x_0, y-y_0)(\operatorname{grad} f(x_0,y_0))^T\\
&2.\ \text{Ordnung:} f(x,y) \approx f(x_0,y_0) + (x-x_0, y-y_0)(\operatorname{grad} f(x_0,y_0))^T\\
&+ \underbrace{\frac{1}{2}\left(f_{x,x}(x_0,y_0)(x-x_0)^2 + f_{y,y}(x_0,y_0)(y-y_0)^2 + 2f_{x,y}(x_0,y_0)(x-x_0)(y-y_0)\right)}_{\frac{1}{2}(x-x_0, y-y_0)H_f(x_0,y_0)\begin{pmatrix} x-x_0 \\ y-y_0 \end{pmatrix}}
\end{aligned}$$

wobei

$$H_f(x_0,y_0) = \begin{pmatrix} f_{x,x}(x_0,y_0) & f_{x,y}(x_0,y_0) \\ f_{y,x}(x_0,y_0) & f_{y,y}(x_0,y_0) \end{pmatrix} = \nabla^T \nabla f(x_0,y_0)$$

ist die Hesse-Matrix (Ludwig Otto Hesse, 1811–1874)

Für die Taylor-Entwicklung zweiter Ordnung entsteht jetzt ein Parabel-Kugel-Gebilde, dass in eine Richtung anders verzerrt ist, als in die andere.

**Vorlesung XXIX**

**Satz 3.24** Taylor-Formel 2. Ordnung

Sei $B = B_R(\mathbf{x}_0) \subseteq \mathbb{R}^n$ eine offene Kugel um den Punkt $\mathbf{x}_0$ mit Radius $R$. Weiter sei $f : B \to \mathbb{R}^1$ eine zweimal stetig partiell differenzierbare Funktion. Dann existiert eine Funktion $r : B_R(\mathbf{0}) \to \mathbb{R}^1$ mit

$$\lim_{\mathbf{h}\to\mathbf{h}} \frac{r(\mathbf{h})}{||\mathbf{h}||^2} = 0$$

und für alle $||\mathbf{h}|| < R$ gilt

$$f(\mathbf{x}_0 + \mathbf{h}) = f(\mathbf{x}_0) + \mathbf{h}(\underbrace{\nabla f(\mathbf{x}_0)}_{\operatorname{grad} f(\mathbf{x}_0)})^T + \frac{1}{2}\mathbf{h}H_f(\mathbf{x}_0)\mathbf{h}^T + r(\mathbf{h}).$$

Das Skalarprodukt $\mathbf{h}(\nabla f(\mathbf{x}_0))^T$ kann natürlich auch geschrieben werden als

$$\nabla f(\mathbf{x}_0)\mathbf{h}^T.$$

Das ist tatsächlich eine *Verallgemeinerung* ins Multivariable, d. h. das Eindimensionale steckt in der Formel drin.

Die eigentliche Aussage der Formel steckt in der Vorraussetzung verlangten Grenzwert

$$\lim_{\mathbf{h}\to\mathbf{h}} \frac{r(\mathbf{h})}{||\mathbf{h}||^2} = 0.$$

Denn wir teilen durch ein sehr kleines $||\mathbf{h}||$ und gehen noch immer gegen 0. Der Rest muss also sehr stark gegen 0 gehen. Wäre das nicht verlangt, könnte die gesamte Funktion in der Taylor-Formel durch den Rest $r(\mathbf{h})$ übernommen werden. Dann wäre die Näherung wertlos.

Die Norm ist nur eine feste Norm im $\mathbb{R}^n$. Wie immer sind auch hier alle Normen in endlich dimensionalen Vektorräumen äquivalent.

Dabei ist

$$H_f(\mathbf{x}_0) = \left(\frac{\partial^2 f}{\partial x_i \partial x_j}(\mathbf{x}_0)\right)_{i,j=1,\dots,n}$$

die *Hesse-Matrix*.

$H_f$ ist symmetrisch, d. h.

$$H_f = h_f^T = \nabla^T \nabla f(\mathbf{x}_0) = \begin{pmatrix} \partial_1 \\ \partial_2 \\ \vdots \\ \partial_n \end{pmatrix} (\partial_1, \partial_2, \dots, \partial_n).$$

**Beweis** Wir betrachten die Gerade $\lambda_{\mathbf{h}}(t) = \mathbf{x}_0 + t\mathbf{h}$ für $t \in \mathbb{R}$. Dieser Gerade ist für $\mathbf{h} \in B_R(\mathbf{0})$ und $t \in [-1,1]$ eine Gerade in $B_R(\mathbf{x}_0)$.

Damit sei

$$g(t) := (f \circ \lambda_{\mathbf{h}})(t) = f(\mathbf{x}_0 + t\mathbf{h}) \in C^2([-1,1]).$$

Dann folgt mit der Kettenregel

$$\begin{aligned} g''(t) &= \left( \sum_{i=1}^n \left( \frac{\partial f}{\partial x_i} \circ \lambda_{\mathbf{h}} \right)(t) \cdot h_i \right)' \\ &= \sum_{i=1}^n \left( \frac{\partial f}{\partial x_i} \circ \lambda_{\mathbf{h}} \right)'(t) \cdot h_i \\ &= \sum_{i=1}^n \sum_{j=1}^n \left( \frac{\partial^2}{\partial x_i \partial x_j} \circ \lambda_{\mathbf{h}} \right)(t) \cdot h_i h_j \\ &= \sum_{i=1}^n \sum_{j=1}^n \frac{\partial^2 f}{\partial x_i \partial x_j} (\mathbf{x}_0 + t\mathbf{h}) h_i h_j \end{aligned}$$

Mit dem eindimensionalen »Taylor« folgt parallel

$$\exists \xi(0,t): \; g(t) = g(0) + g'(0)t + \frac{1}{2} g''(0)t^2 + \eta(t)t^2$$

mit

$$\eta(t) = \frac{1}{2} \left( g''(\xi) - g''(0) \right)$$

also

$$\lim_{t \to 0} \eta(t) = 0.$$

Für $t = 1$ gilt

$$\begin{aligned} f(\mathbf{x}_0 + \mathbf{h}) = g(1) &= g(0) + g'(0) \cdot 1 + \frac{1}{2} g''(0) \cdot 1 + \eta(1) \\ &= f(\mathbf{x}_0) + \mathbf{h}(\operatorname{grad} f(\mathbf{x}_0))^T + \frac{1}{2} \mathbf{h} H_f(\mathbf{x}_0) \mathbf{h}^T + \eta(1) \end{aligned}$$

Wir sehen also, wie sich die gewaltige Doppelsumme in eine Vektor-Matrix-Vektor-Multiplikation verwandelt hat.

Nun müssen wir noch das Abklingen des Restgliedes zeigen!

$$r(\mathbf{h}) = \eta(1) = \frac{1}{2}\mathbf{h}\left(H_f(\mathbf{x}_0+\xi\mathbf{h}) - H_f(\mathbf{x}_0)\right)\mathbf{h}^T$$

$$\frac{1}{2}\sum_{i=1}^{n}\sum_{i=1}^{n}\left(f_{x_i,x_j}(\mathbf{x}_0+\xi\mathbf{h}) - f_{x_i,x_j}(\mathbf{x}_0)\right)h_ih_j$$

Um den Grenzwert 0 für $\mathbf{h}\to\mathbf{0}$ zu zeigen, betrachten wir zunächst den Betrag.

$$|r(\mathbf{h})| \leq \underbrace{\max_{1\leq i,j\leq n}\left|f_{x_i,x_j}(\mathbf{x}_0+\xi\mathbf{h}) - f_{x_i,x_j}(\mathbf{x}_0)\right|}_{=\mathrm{o}(1)\text{ für }\|\mathbf{h}\|\to 0}\cdot \underbrace{\sum_{i=1}^{n}\sum_{j=1}^{n}}_{\left(\underbrace{\sum_{i=1}^{n}|h_i|}_{\leq\sqrt{n}\|\mathbf{h}\|_2}\right)\left(\underbrace{\sum_{j=1}^{n}|h_j|}_{\leq\sqrt{n}\|\mathbf{h}\|_2}\right)} |h_ih_j|$$

Damit folgt

$$\frac{|r(\mathbf{h})|}{n\|\mathbf{h}\|_2^2} = \mathrm{o}(1) \text{ für } \|\mathbf{h}\|\to 0$$

$\square$

**Beispiel** $n=2$

Wir betrachten $f(x,y) = x^3+y^3-3xy$ und die Taylor-Entwicklugn 2. Ordnung im Entwicklungspunkt $(-1,1)$.

$$\begin{aligned}\operatorname{grad} f(x,y) &= (3x^2-3y, 3y^2-3x) \Rightarrow \operatorname{grad} f(-1,1) = (0,6)\\ H_f(x,y) &= \begin{pmatrix}6x & -3\\ -3 & 6y\end{pmatrix} \Rightarrow h_f(-1,1) = \begin{pmatrix}-6 & -3\\ -3 & 6\end{pmatrix}\\ &\qquad f(-1,1) = 3\end{aligned}$$

Für den Vektor $\mathbf{h}$ gilt in diesem Fall

$$\mathbf{h} := (x,y)-(-1,1).$$

$$f(x,y) \approx 3 + ((x,y)-(-1,1))(\operatorname{grad} f(-1,1))^T + \frac{1}{2}((x,y)-(-1,1))H_f(-1,1)((x,y)-(-1,1))^T$$

$$\begin{aligned}
&= 3 + (x+1, y-1)\begin{pmatrix}0\\6\end{pmatrix} + \frac{1}{2}(x+1,y-1)\begin{pmatrix}-6 & -3\\-3 & 6\end{pmatrix}\begin{pmatrix}x+1\\y-1\end{pmatrix}\\
&= 3 + (y-1)6 + (x+1,y-1)\begin{pmatrix}-6x-6-3y+3\\-3x-3+6y-6\end{pmatrix}\\
&= 6y-3+\frac{1}{2}(x+1)(-6x-3y-3)+\frac{1}{2}(y-1)(6y-3x-9)\\
&= -3x-3y-3x^2+3y^2-3xy
\end{aligned}$$

$[\frac{1}{2}(x+1)^2(-6)+\frac{1}{2}(y-1)^2 6+(-3)(x+1)(y-1)]$

In diesem Spezialfall ergibt sich aus dem Eindimensionalen

$$\begin{aligned}
x^3 &\approx -1+3(x+1)-3(x+1)^2\\
y^3 &\approx 1+3(y-1)+3(y-1)^2\\
-3xy &= -3(x+1)(y-1)+3y-3y\\
&= -3(x+1)(y-1)+3(y-1)-3(x+1)+3+3
\end{aligned}$$

**Vorlesung C**

## 3.3 Extremwerte

Wir werden jetzt die Erkenntnisse über Extremwerte und deren Bestimmung aus dem Eindimensionalen ins Mehrdimensionale übertragen.

**Definition 3.25**

Sei $\Omega \subseteq \mathbb{R}^n$ offen, $f : \Omega \to \mathbb{R}^1$ eine Funktion und $\mathbf{a} \in \Omega$. Dann hat $f$ in $\mathbf{a}$ ein *lokales Maximum (Minimum)*, falls eine offene Umgtebung $U_\varepsilon(\mathbf{a}) \subset \Omega$, so dass für alle $\mathbf{x} \in U_\varepsilon(\mathbf{a})$ gilt:

$$f(\mathbf{x}) \leq f(\mathbf{a}) \qquad (f(\mathbf{x}) \geq f(\mathbf{a}))$$

Gilt diese Ungleichung für alle $\mathbf{x} \in \Omega$, dann handelt es sich um ein *globales Maximum (Minimum)*.

**Lemma 3.26** notwendiges Kriterium für Extremwerte

Sei $\Omega \subseteq \mathbb{R}^n$ offen und $f : \Omega \to \mathbb{R}$ eine in $\mathbf{a}$ differenzierbare Funktion.

Wenn $f$ ein ein lokales Extremwert in $\mathbf{a}$ besitzt, dann folgt daraus

$$\operatorname{grad} f(\mathbf{a}) = \mathbf{0}$$

Analog zum Eindimensionalen liefert dieses Lemma nur eine Implikation. Die Aussage gilt also *nicht* in der Richtung, dass aus $\operatorname{grad} f(\mathbf{a}) = \mathbf{0}$ unbedingt eine Extremstelle folgt!

Wenn der Gradient der Nullvektor ist, dann müssen alle partiellen Ableitungen (also die Richtungsableitungen entlang aller Koordinatenachsen) null sein.

[Landmasse in der x-y-Ebene. Zwei Geraden durch die Masse parallel zur x- und y-Achse. Oben drüber Teppich und Hochpunkt. Beide Geraden auch oben. Die sind gerade eindimensionale Funktionen, die beide einen Hochpunkt haben müssen, damit der Teppich dort auch einen Hochpunkt hat.]

**Beweis** Für alle $i = 1, \dots, n$ sei $g_i(t) = f(\mathbf{a} + t\mathbf{e}_i)$ mit $t \in \mathbb{R}$ und $\mathbf{a} + t\mathbf{e}_i \in \Omega$.

Wenn die Funktion $g_i$ in $t = 0$ ein lokales Extremum besitzt, dann gilt nach Satz 2.16 [Verweis] für alle $i \in \{1, \dots, n\}$

$$0 \overset{\dagger}{=} g_i'(0) \overset{\ddagger}{=} \operatorname{grad} f(\mathbf{a} + 0\mathbf{e}_i)\mathbf{e}_i^T = \frac{\partial f}{\partial x_i}(\mathbf{a})$$

□

An der Stelle † wurde Satz 2.16 [Verweis] angewendet. An der Stelle ‡ wurde die Kettenregel verwendet

Wie im Eindimensionalen liefert dieses Lemma nur ein notwendige Bedingung. Es gibt aber auch hier fälle wie $f(x) = x^3$ im Eindimensionalen, wo an der Stelle $x_0 = 0$ durchaus die Bedingung erfüllt ist, aber keine Extremstelle vorliegt.

**Definition 3.27**

Sei $f : \Omega \to \mathbb{R}$ differenzierbar und gilt $\operatorname{grad} f(\mathbf{a}) = \mathbf{0}$, so heißt $\mathbf{a}$ ein *stationärer Punkt* von $f$.

Ein stationärer Punkt $\mathbf{a}$ heißt *Sattelpunkt* von $f$, falls in jeder Umgebung von $\mathbf{a}$ Vektoren $\mathbf{b}$ und $\mathbf{c}$ existieren, sodass

$$f(\mathbf{b}) < f(\mathbf{a}) < f(\mathbf{c})$$

»Jeder der sich einen Pferdesattel vorstellt weiß, wenn Sie dass Pferd quer durchschneiden, gibt es in Richtung Maximum ein Minimum.«

### Definition 3.28

Sei $A \in \mathbb{R}^{n\times n}$ eine symmetrische $n \times n$ Matrix, d. h. $A^T = A$.

$$q(\mathbf{h}) = \mathbf{h}A\mathbf{h}^T, \quad \mathbf{h} \in \mathbb{R}^n$$

heißt die *quadratische Form*.

$q$ bzw. $A$ heißt *positiv semidefinit* genau dann, wenn für alle $h \in \mathbb{R}^n$ gilt $q(\mathbf{h}) \geq 0$.

$q$ bzw. $A$ heißt *positiv definit* genau dann, wenn für alle $h \in \mathbb{R}^n \backslash \{\mathbf{0}\}$ gilt $q(\mathbf{h}) \geq 0$.

Analog für *negativ (semi)definit*.

$q$ bzw. $A$ heißt *indefinit* genau dann, wenn zwei Vektoren $\mathbf{h}_1$ und $\mathbf{h}_2$ existieren, sodass $q(\mathbf{h}_1) < 0 < q(\mathbf{h}_2)$.

Da wir diese Erkenntnisse auf die Hesse-Matrix anwenden wollen, ist die Einschränkung auf symmetrische Matrizen durchaus annehmbar. Schließlich ist die Hesse-Matrix immer symmaterisch, wenn sich der Satz von Schwarz [Verweis] anwenden lässt.

### Beispiel

$$A = \begin{pmatrix} a & b \\ b & d \end{pmatrix}, \quad \mathbf{h} = (h_1, h_2)$$
$$q(\mathbf{h}) = \mathbf{h}A\mathbf{h}^T = (h_1, h_2) \begin{pmatrix} a & b \\ b & d \end{pmatrix} \begin{pmatrix} h_1 \\ h_2 \end{pmatrix} = ah_1^2 + 2bh_1h_2 + dh_2^2$$

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \pm 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \mathbf{h} = (h_1, h_2, h_3)$$
$$q(\mathbf{h}) = \mathbf{h}A\mathbf{h}^T = (h_1, h_2, h_3) \begin{pmatrix} 1 & 0 & 0 \\ 0 & \pm 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} h_1 \\ h_2 \\ h_3 \end{pmatrix} = h_1^2 \pm h_2^2 + h_3^2$$

**Bemerkung** 1. $q(t\mathbf{h}) = t^2 q(\mathbf{h}), \quad t \in \mathbb{R}$

2. $q(\mathbf{0}) = 0$
3. Die Definitheit von $q$ lässt sich durch die Eigenwerte von $A$ beschreiben.

**Lemma 3.29**

$$A = \begin{pmatrix} a & b \\ b & d \end{pmatrix}$$

Gilt $\det A = ad - b^2 < 0$, dann gilt $q$ ist indefinit.

Gilt $\det A > 0$ dann ist $q$ positiv definit, wenn $a > 0$ und $q$ negativ definit, wenn $a < 0$.

**Beweis**

1. $a > 0$

$$q(h_1, h_2) = a(\left(h_1^2 + \frac{2h}{a} h_1 h_2 + \frac{d}{a} h_2^2\right))$$

Die quadratische Gleichung in $h_1$ hat die Diskriminante

$$h_2^2 \left(\frac{2b}{2a}\right)^2 - \frac{d}{a} h_2^2 = \left(\frac{h_2}{a}\right)^2 \cdot (b^2 - ad)$$

Falls die Diskriminante negativ ist, dann liegen keine Nullstellen vor. Falls die Diskriminante positiv ist, dann existieren positive und negative Funktionswerte zwischen den Nullstellen.

2. $a < 0$ analog
3. $a = 0, b \neq 0$

$$q(h_1, h_2) = 2bh_1 h_2 + dh_2^2$$

Wir betrachten nur $q(h1_, 1) = 2bh_1 + d$, dann sehen wir, dass $h_1$ existieren mit $q(h_1, 1) > 0$ und $q(h_1, 1) < 0$.

$\square$

Aus der Sicht der moderne Algebra ist das kein sinnvoller Ansatz. Der Weg über die Eigenwerte der Matrix $A$ führt schneller zum Ziel.

**Satz 3.30** Hinreichendes Kriterium für Extremwerte

Sei $f : \Omega \to \mathbb{R}$ mit $\Omega \subseteq \mathbb{R}^n$ offen und $f \in C^2(\Omega)$ ($f$ ist zweimal stetig partiell differenzierbar). $\mathbf{a} \in \Omega$ sei ein stationärer Punkt. Dann gilt

1. $H_f(\mathbf{a})$ positiv definit $\Longrightarrow f$ hat in $\mathbf{a}$ lokales Minimum
2. $H_f(\mathbf{a})$ negativ definit $\Longrightarrow f$ hat in $\mathbf{a}$ lokales Maximum
3. $H_f(\mathbf{a})$ indefinit $\Longrightarrow f$ hat in $\mathbf{a}$ einen Sattelpunkt

**Beweis**

1. $H_f(\mathbf{a})$ positiv definit.

$$q(\mathbf{h}) = \mathbf{h}H_f(\mathbf{a})\mathbf{h}^T$$

Mit der Taylor-Formel gilt

$$f(\vee a+\mathbf{h}) = f(\mathbf{a}) + \mathbf{h}\,\mathrm{grad}\,f(\mathbf{a})^T + \frac{1}{2}q(\mathbf{g}) + r(\mathbf{h})$$

Es gilt $\mathbf{h}\,\mathrm{grad}\,f(\mathbf{a})^T = 0$, da $\mathrm{grad}\,f(\mathbf{a}) = \mathbf{0}$. Es bleibt also zu zeigen, dass $\frac{1}{2}q(\mathbf{g}) + r(\mathbf{h}) \geq 0$.

Stetige Funktionen nehmen auf beschränkten und abgeschlossen Gebieten ihr Minimum und Maximum an.

$q$ ist stetig und positiv (außer in $\mathbf{0}$), also nimmt $q$ auf der Sphäre $\{\mathbf{x} \mid \|\mathbf{x}\| = 1\}$ (beschränkt und abgeschlossen!) ein Minimum $m > 0$ an. Damit gilt

$$\forall \mathbf{h} \in \mathbb{R}^n \backslash \{\mathbf{0}\}: \quad q(\mathbf{h}) = \|\mathbf{h}\|^2 q\left(\frac{h}{\|\mathbf{h}\|}\right) > \|\mathbf{h}\|^2 m$$

$$\left\|\frac{1}{\|\mathbf{h}\|}\mathbf{h}\right\| = 1$$

Außerdem gilt für alle $0 < \varepsilon < \frac{m}{2}$ gibt es ein $R = R(\varepsilon)$ mit

$$\forall \mathbf{h} \in U_R(\mathbf{0}): \quad |r(\mathbf{h})| \leq \varepsilon \|\mathbf{h}\|^2$$

In der Taylor-Formel steht eigentlich $r(h) = o(\|\mathbf{h}\|^2)$, aber die Landau-Symbolik meint genau die oben stehende Aussage.

Also gilt

$$f(\mathbf{a}+\mathbf{h}) - f(\mathbf{a}) = \frac{1}{2}q(\mathbf{h}) + r(\mathbf{h}) \geq \left(\frac{m}{2} - \varepsilon\right)\|\mathbf{h}\|^2 \geq 0$$

2. $H_f(\mathbf{a})$ negativ deinit analog mit anderem Vorzeichen.

3. $H_f(\mathbf{a})$ indefinit, d. h. es existieren Vektoren $\mathbf{h}_1, \mathbf{h}_2$ mit $q(\mathbf{h}_1) < 0 < q(\mathbf{h}_2)$.

$$\begin{aligned} f_1(t) &= f(\mathbf{a} + t\mathbf{h}_1) \\ f_2(t) &= f(\mathbf{a} + t\mathbf{h}_2) \\ f_1'(0) &= f_2'(0) = 0 \\ f_1''(0) &= q(\mathbf{h}_1) < 0 \\ f_2''(0) &= q(\mathbf{h}_2) > 0 \end{aligned}$$

Also ist $t = 0$ ein lokales Maximum für $f_1$ und ein Minimum für $f_2$.

Also ist $\mathbf{q}$ ein Sattel.

□

**Vorlesung CI**

### Beispiel

Wir untersuchen folgende Funktion auf Extrema:

$$f : \mathbb{R}^2 \to \mathbb{R}, \quad f(x,y) = x^3 + y^3 - 3xy$$

Dazu betrachten wir zunächst die notwendige Bedingung

$$\nabla f(x) = (3x^2 - 3y, 3y^2 - 3x) = \mathbf{0}$$

$$\Rightarrow \begin{cases} y = x^2 \\ x = y^2 \end{cases} \quad \Rightarrow x = x^4 \Rightarrow x^4 - x = (x-0)(x-1)(x^2+x+1) = 0$$

$$\begin{aligned} \Rightarrow x_1 = 0 &\Rightarrow y_1 = 0 \\ x_2 = 0 &\Rightarrow y_2 = 1 \end{aligned}$$

Nun betrachten wir die hinreichende Bedinung

$$\begin{aligned} H_f(x,y) &= \begin{pmatrix} 6x & -3 \\ -3 & 6y \end{pmatrix} \\ H_f(0,0) &= \begin{pmatrix} 0 & -3 \\ -3 & 0 \end{pmatrix} \\ H_f(1,1) &= \begin{pmatrix} 6 & -3 \\ -3 & 6 \end{pmatrix} \end{aligned}$$

$\det H_f(0,0) = -9 < 0 \Rightarrow$ indefinite Matrix, also Sattelstelle

$\det H_f(1,1) = 27 > 0 \wedge f_{x,x}(1,1) = 6 > 0 \Rightarrow$ positiv definite Matrix, also Minimalstelle

In diesem Fall kann man die Aufgabe auch ohne höhere Mathematik lösen. Dazu formt man die Funktion um zu

$$f(x,y) = (x-1)^2(x+\frac{1}{2}) + (y-1)^2(y-\frac{1}{2}) + \frac{3}{2}(x-y)^2 - 1 \geq -1$$

und kann jetzt das Ergebnis direkt ablesen.

Aber wir haben ja zum Glück die höhere Mathematik und müssen da nichts direkt ablesen können...

### Beispiel

$$[f : \mathbb{R}^2 \to \mathbb{R}, \quad f(x,y) = 2x^4 + y^4 - 2x^2 - 2y^2$$

Wir betrachten zunächst die notwendige Bedingung

$$\begin{aligned} \nabla f(x,y) = (8x^3 - 4x, 4y^3 - 4y) = \mathbf{0} \\ 0 = 2x^3 - x = x(2x^2-1) = x(\sqrt{2}x - 1)(\sqrt{2}x+1) \\ 0 = y^3 - y = y(y-1)(y+1) \\ x_{1,2,3} = 0; \pm\frac{\sqrt{2}}{2} \\ y_{1,2,3} = 0; \pm 1 \end{aligned}$$

Für die hinreichende Bedingung betrachten wie die Hesse-Matrix

$$H_f(x,y) = \begin{pmatrix} 24x^2 - 4 & 0 \\ 0 & 12y^2 - 4 \end{pmatrix}$$

| $(x,y)$ | $\det H_f(x,y)$ | $f_{x,x}$ | |
|---|---|---|---|
| $(0,0)$ | 16 | -4 | Maximum |
| $(0,1)$ | -32 | | Sattel |
| $(0,-1)$ | -32 | | Sattel |
| $(\frac{\sqrt{2}}{2},0)$ | -32 | | Sattel |
| $(-\frac{\sqrt{2}}{2},0)$ | -32 | | Sattel |
| $(\frac{\sqrt{2}}{2},1)$ | 64, 8 | Minimum | |
| $(\frac{\sqrt{2}}{2},-1)$ | 64 | 8 | Minimum |
| $(-\frac{\sqrt{2}}{2},1)$ | 64 | 8 | Minimum |
| $(-\frac{\sqrt{2}}{2},1)$ | 64 | 8 | Minimum |

[Tabelle schön machen!]

4 Integralrechnung


# 4 Integralrechnung

**Aufgaben** a) Umkehrung der Differentiation, d. h. Berechnung der Stammfunktion $\longrightarrow$ »unbestimmte Integration« (nicht eindeutig)
b) Flächenberechnung (Volumina, Arbeit) $\longrightarrow$ »bestimme Integraion« (eindeutig)

**Probleme** a) Berechnung von Integralen
b) Zusammengang von bestimmten und unbestimmten Integralen

## 4.1 Unbestimmte Integrale

**Definition 4.1**

Sei $I \subseteq \mathbb{R}$ ein Intervall. Die Funktion $F : I \to \mathbb{R}$ heißt eine *Stammfunktion* von $f : I \to \mathbb{R}$, wenn gilt

$$\forall x \in I : F'(x) = f(x)$$

Bekannt aus Folgerung 2.18 [Verweis]: Jede andere Stammfunktion von $f$ auf $I$ hat die From $F(x) + C$ mit einer beliebigen Konstante $C$.

**Beispiel**

- $f(x) = x, F(x) = \frac{x^2}{2}, \frac{x^2}{2} + 1, \frac{x^2}{2} - \Pi, \ldots$
- $f(x) = e^x, F(x) = e^x, e^x + 5, e^x - 1, \ldots$

Ist $F$ eine Stammfunktion von $f$, so heißt $F(x) + C$ (wobei $C$ eine unbestimmte Konstante ist) das *unbestimmte Integral* von $f$. Schreibweise

$$\int f(x)\,\mathrm{d}x = F(x) + C$$

### 4.1.1 Rechenregeln für unbestimmte Integrale

**Satz 4.2** Rechenregeln für unbestimmte Integrale

1. Seien $f, g : I \to \mathbb{R}$ zwei Funktion, die Stammfunktionen bestitzen. Dann gilt für beliebige $\lambda, \mu \in \mathbb{R}$
$$\int \lambda f(x) + \mu g(x)\,\mathrm{d}x = \lambda \int f(x) + \mu \int g(x)$$
2. **Substitutionsregel** $f : I \to \mathbb{R}$ besitze eine Stammfunktion. Weiter sei $\varphi : D \to I$ stetif differenzierbar und $x = \varphi(t)$.
   a) Sei $\int f(x)\,\mathrm{d}x = F(x) + C$ ,d. h. $\forall x \in I : \; F'(x) = f(x)$. Dass heißt
   $$\frac{dF(\varphi(t))}{\mathrm{d}t} = F'(\varphi(t)) \cdot \varphi'(t) = f(\varphi(t))\varphi'(t)$$
   und damit ergibt sich
   $$\int f(\varphi(t))\varphi'(t)\,\mathrm{d}t = F(\varphi(t)) + C$$
   b) Sei außerdem $\varphi$ bijektiv, d. h. $\varphi'(t) \neq 0$ $(\varphi^{-1}(x) = t)$

   Sei $\int f(\varphi(t))\varphi'(t)\,\mathrm{d}t = G(t) + C$, also $G'(t) = f(\varphi(t))\varphi'(t)$, also $\frac{dG(\varphi^{-1}(x)))}{\mathrm{d}x} = G'(\varphi^{-1}(x))(\varphi^{-1})'(x)$, also $\int f(x)\,\mathrm{d}x = G(\varphi^{-1}(x)) * C$.

   [HILFE!!!]

   c) **Partielle Intergration**

   Seien $u, v$ stetig differenzierbar in $I$, dann gilt
   $$\int u(x)v'(x)\,\mathrm{d}x = u(x)v(x) - \int u'(x)v(x)\,\mathrm{d}x$$

**Beweis** 1. Folgt aus der Linearität der Ableitung.:
$$(\lambda F(x) + \mu G(x) + C)' = \lambda F'(x) + \mu G'(x)$$

2. Beweis in Formulierung enthalten.
3. Mit der Produktregel $\frac{\mathrm{d}}{\mathrm{d}x}(u(x)v(x)) = u'(x)v(x) + u(x)v'(x)$ ergibt sich
$$\int u'(x)v(x) + u(x)v'(x)\,\mathrm{d}x = u(x)v(x) + C$$

□

**Beispiel**

$f$ besitze Stammfunktion $F$. Dann gilt für $a \neq 0$ und $\varphi(t) = at + b = x (\Leftrightarrow t = \frac{x-b}{a})$

$$\int f(\phi(t))\,\mathrm{d}t = \int f(at+b)\,\mathrm{d}t = \int f(x)\frac{1}{a}\,\mathrm{d}x$$

Eselsbrücke: Für $x = \varphi(t)$, d. h.

$$\frac{\mathrm{d}x}{\mathrm{d}t} = \varphi'(t)$$

schreiben wir einfach mathematisch falsch

$$\varphi'(t)\,\mathrm{d}t = \mathrm{d}x$$

Damit ergibt sich

$$\int f(at+b)\,\mathrm{d}t = \int f(x)\frac{1}{a}\,\mathrm{d}x.$$

**Vorlesung CII**

## Liste der Grundintegrale

1. $\displaystyle\int x^\alpha\,\mathrm{d}x = \begin{cases} \frac{x^{\alpha+1}}{\alpha+1} + C & \text{für } \alpha \in \mathbb{R}\backslash\{-1\} \\ \ln|x| + C & \text{für } \alpha = -1 \end{cases}$
2. $\displaystyle\int \frac{1}{x}\,\mathrm{d}x = \ln|x| + C \text{ für } x > 0 \text{ oder } x < 0$
3. $\displaystyle\int \sin x\,\mathrm{d}x = -\cos x + C$
4. $\displaystyle\int \cos x\,\mathrm{d}x = \sin x + C$
5. $\displaystyle\int \frac{\mathrm{d}x}{\cos^2 x} = \tan x + C \text{ für } (x + \frac{\Pi}{2} + k\Pi)$
6. $\displaystyle\int \frac{\mathrm{d}x}{\sin^2 x} = -\cot x + C \text{ für } x \neq k\Pi$
7. $\displaystyle\int \frac{\mathrm{d}x}{\sqrt{1-x^2}} = \arcsin x \text{ für } |x| < 1$
8. $\displaystyle\int \frac{\mathrm{d}x}{\sqrt{1+x^2}} = \operatorname{arsinh} x + C = \ln(x + \sqrt{1+x^2}) + C$

9. $\displaystyle\int \frac{\mathrm{d}x}{\sqrt{x^2-1}} = \operatorname{arcosh} x + C = \ln(x+\sqrt{x^2-1}) + C$ für $|x| > 1$

10. $\displaystyle\int \frac{\mathrm{d}x}{1+x^2} = \arctan x + C$

11. $\displaystyle\int e^x\,\mathrm{d}x = e^x + C$

12. $\displaystyle\int a^x\,\mathrm{d}x = \frac{a^x}{\ln a} + C$ für $a > 0$ und $a \neq 1$

**Beispiel** Substitution

Wir betrachten das Integral

$$\int \sqrt{5x+2}\,\mathrm{d}x.$$

Die innere Funktion lautet

$$y = 5x+2 \Rightarrow x = \frac{1}{5}(y-2).$$

Für die Ableitung der inneren Funktion gilt

$$\frac{dy}{\mathrm{d}x} = 5 \Rightarrow dy = 5\,\mathrm{d}x \Rightarrow \mathrm{d}x = \frac{1}{5}dy.$$

Damit ergibt sich für das Integral

$$\begin{aligned}\int \sqrt{5x+2}\,\mathrm{d}x &= \int \sqrt{y}\frac{1}{5}dy \\ &= \frac{2}{3}y^{\frac{3}{2}}\frac{1}{5} + C \\ &= \frac{2}{15}(5x+2)^{\frac{3}{2}} + C\end{aligned}$$

**Beispiel**

Sei $\varphi(t) \neq 0$ und stetig differenzierbar.

$$\int \frac{\varphi'(t)}{\varphi(t)}\,\mathrm{d}t = \ln|\varphi(t)| + C$$

Betrachten wir dazu das angewandte Beispiel für $x \neq k\Pi$

$$\int \cot x\,\mathrm{d}x = \int \frac{\cos x}{\sin x}\,\mathrm{d}x = \ln|\sin x| + C$$

**Beispiel** partielle Integration

Betrachten wir das Integral

$$\int xe^x \,\mathrm{d}x$$

dann bietet sich folgende Wahl an

$$\begin{aligned} u(x) = x &\Rightarrow u'(x) = 1 \\ v'(x) = e^x &\Rightarrow v(x) = e^x \end{aligned}$$

Damit ergibt sich

$$\begin{aligned} \int xe^x \,\mathrm{d}x &= xe^x - \int e^x \,\mathrm{d}x \\ &= (x-1)e^x + C \end{aligned}$$

**Beispiel** weiteres Grundintegral

Wir verwenden die partielle Integration für folgendes Integral

$$\int \ln x \,\mathrm{d}x = \int 1 \ln x \,\mathrm{d}x$$

indem wir wählen

$$\begin{aligned} u'(x) = 1 &\Rightarrow u(x) = x \\ v(x) = \ln x &\Rightarrow v'(x) = \frac{1}{x} \end{aligned}$$

Damit ergibt sich

$$\begin{aligned} \int 1 \ln x \,\mathrm{d}x &= x \ln x - \int x \frac{1}{x} \,\mathrm{d}x \\ &= x \ln x - x + C \qquad \text{für } x > 0 \end{aligned}$$

**Bemerkung** Differentiation ist »nur« ein Handwerk. Wir können jede Funktion mithilfe der uns bekannten Mitteln ableiten. Mit algorithmischen Methoden erhalten wir zu jeder Funktion eine Ableitung.

Integration ist hingegen eine Kunst. Nicht jede Funktion lässt sich integrieren.

Wir eine elementare Funktion differenziert, erhalten wir wieder eine elementare Funktion. Aber was passiert, wenn eine elementare Funktion integriert wird? Erhalten wir auch dann in jedem Fall eine elementare Funktion. Die erstaunliche Antwort lautet *Nein*.

**Beispiel** nicht elementarer Stammfunktionen

- Die Funktion zur Gaußschen Glockenkurve ist stetig differenzierbar (also wunderbar glatt) und es existiert auf jeden Fall eine Stammfunktion.

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

Diese Stammfunktion $F$ lässt sich aber *nicht* als elementare Funktion schreiben!

- $\frac{\sin x}{x}\,\mathrm{d}x$
- $\sin x^2\,\mathrm{d}x$

Fazit: Es gibt nur wenige Klassen von Funktionen, deren Integration in geschlossener Form möglich ist, d. h. deren Integration wieder zu einer Darstellung mit endlich vielen elementaren Funktionen führt.

Das wichtigste Beispiel für diese Klassen sind die rationalen Funktionen.

## 4.2 Integraion rationaler Funktionen

Rationale Funktionen bestehen im Zähler und Nenner aus Polynomen. Damit sind sie von folgender Bauart:

$$R(x) = \frac{p(x)}{q(x)} = \frac{a_0 + a_1x + \ldots a_nx^n}{b_0 + b_1x + \ldots b_mx^m}$$

Wenn das Nenner-Polynom $q$ Nullstellen hat, dann hat die Funktion $R$ Lücken. Wir können natürlich nicht über diese Lücken hinweg integrieren. Also betrachten wir nur Intervalle zwischen den Nullstellen von $q$.

Wir suchen $\int R(x)\,\mathrm{d}x$ für $a_n, b_m \neq 0$.

Wir nehmen ohne Beschränkung der Allgemeinheit $n < m$ an, denn andernfalls können wir die Funktion zerlegen als

$$R(x) = \prod_{n-m} + R_1(x)$$

mit $R_1$ echt gebrochen, d. h. $n < m$.

**Satz 4.3** Partialbruchzerlegung

Jede echt gebrochene rationale Funktion $R(x) = \frac{p(x)}{q(x)}$ lässt sich in eine Summe von *reellen Partialbrüchen* zerlegen.

Sei $q(x) = c(x-b_1)^{k_1}(x-b_2)^{k_2}\dots(x-b_\alpha)^{k_\alpha}q_1(x)^{l_1}q_2(x)^{l_2}\dots q_\beta(x)^{l_\beta}$ mit paarweise verschiedenen reellen Nullstellen $b_j$ und paarweise verschiedenen quadratischen Polynomen $q_j$ ohne reelle Nullstellen.

Für den Grad des Polynoms $R$ ergibt sich damit

$$m = \sum_{j=1}^{\alpha} k_j + 2\sum_{j=1}^{\beta} l_j.$$

Für das Polynom $R$ ergibt sich

$$\begin{aligned}
R(x) = \frac{p(x)}{q(x)} &= \sum_{j=1}^{\alpha}\left(\frac{A_{j_1}}{x-b_j} + \frac{A_{j_2}}{(x-b_j)^2} + \dots + \frac{A_{j_{k_j}}}{(x-b_j)^{k_j}}\right) \\
&\quad + \sum_{j=1}^{B}\left(\frac{B_{j_1}x + C_{j_1}}{q_j(x)} + \frac{B_{j_2}x + C_{j_2}}{(q_j(x))^2} + \dots + \frac{B_{j_{l_1}} + C_{j_{l_j}}}{(q_j(x))^{l_j}}\right) \\
&= \sum_{j=1}^{\alpha}\left(\sum_{\mu=1}^{k_j}\frac{A_{j_\mu}}{(x-b_j)^\mu}\right) + \sum_{j=1}^{B}\left(\sum_{\mu=1}^{l_j}\frac{B_{j_\mu}x + C_{j_\mu}}{(q_j(x))^\mu}\right)
\end{aligned}$$

mit reellen Koeffizienten $A_{j_\mu}, B_{j_\mu}, C_{j_\mu}$.

**Beweis** ohne Beweis. □

**Beispiel** abgefahrenes Beispiel ohne Lösung

$$\frac{1}{(x+3)^4(x-1)(x^2-2x-2)^3(x^2+2)}$$

$$\begin{aligned}
x^2 + 2 &= (x-\sqrt{2}i)(x+\sqrt{2}i) \\
x^2 - 2x - 2 &= (x-1)^2 + 1 > 0
\end{aligned}$$

$$\begin{gathered}
\frac{1}{(x+3)^4(x-1)(x^2-2x-2)^3(x^2+2)} \\
= \frac{A_{11}}{x+3} + \frac{A_{12}}{(x+3)^2} + \frac{A_{13}}{(x+3)^3} + \frac{A_{14}}{(x+3)^4} + \frac{A_{21}}{x-1} \\
\frac{B_{11}x + C_{11}}{x^2-2+2} + \frac{B_{12}x + C_{12}}{(x^2-2x+2)^2} + \frac{B_{13}x + C_{13}}{(x^2-2x+2)^3} + \frac{B_{21}x + C_{21}}{x^2+2}
\end{gathered}$$

Bestimmung der Koeffizienten $A, B, C$ durch

a) Koeffizientenvergleich
b) Einsetzen spezieller Werte

**Beispiel** realistisches Beispiel

$$\frac{x^4+x^3+2x}{x^3+x^2+x+1} = x\frac{x-x^2}{(x+1)(x^2+1)}$$
$$\frac{x-x^2}{(x+1)(x^2+1)} = \frac{A}{x+1}+\frac{Bx+C}{x^2+1}$$
$$\Rightarrow x-x^2 = A(x^2+1)+(Bx+C)(x+1)$$
$$0 = x^2(\underbrace{A+B+1}_{=0})+x(\underbrace{B+C-1}_{=0})+\underbrace{A+C}_{=0}$$

$$A+C=0$$
$$B+C=1$$
$$A+B=-1$$

$$B-A=1$$

$$2B=0$$
$$B=0$$
$$C=1$$
$$A=-1$$

Wir erhalten also

$$\frac{A}{x+1}+\frac{Bx+C}{x^2+1} = \frac{-1}{x+1}+\frac{1}{x^2+1}$$

Alternativ hätte auch das Einsetzen von $x=-1$ sofort zu $A=-1$ geführt. Das funktioniert nicht immer, aber manchmal geht es schneller.

**Vorlesung CIII**

**Beispiel** noch ein realistisches Beispiel

$$R(x) = \frac{x^2+3x+4}{(x^2-1)^2}$$

Zählerpolynom hat einen geringeren Grad als das Nennerpolynom. Also ist keine Polynomdivision nötig um einen nicht gebrochenen Teil abzuspalten.

Wir wissen aus dem Satz, dass wir das Polynom folgendermaßen aufteilen können.

$$\begin{aligned} R(x) &= \frac{x^2+3x+4}{(x-1)^2(x+1)^2} \\ &= \frac{A}{x-1} + \frac{B}{(x-1)^2} + \frac{C}{x+1} + \frac{D}{(x+1)^2} \end{aligned}$$

Durch Umstellen erhalten wir nun eine Gleichung, in die wir geschickt einsetzen können, um die Gleichung zu lösen.

$$x^2+3x+4 = A(x-1)(x+1)^2 + B(x+1)^2 + C(x-1)^2(x+1) + D(x-1)^2$$

$$\begin{aligned} x = 1 &: 8 = 4B \Rightarrow B = 2 \\ x = -1 &: 2 = 4D \Rightarrow D = \frac{1}{2} \\ x = 0 &: 4 = A + 2 + C + \frac{1}{2} \\ x = 2 &: 14 = 9A + 18 + 3C + \frac{1}{2} \\ &\Rightarrow A = -\frac{3}{4}, C = \frac{3}{4} \end{aligned}$$

### Zur Integration der Partialbrüche

- $\displaystyle\int \frac{\mathrm{d}x}{x-a} = \ln|x-a| + C$ für $x \neq a$

  Basierend auf dem Integral von $1/x$. Denn es gilt

  $$\int \frac{1}{x}\,\mathrm{d}x = \ln|x| + C.$$

  Die Betragstriche sind hier sehr wichtig, denn die Logarithmus-Funktion ist nur für positive Zahlen definiert. Für den Fall $x < 0$ gilt

  $$\frac{\mathrm{d}}{\mathrm{d}x}\ln|x| = \frac{1}{|x|} \cdot \frac{\mathrm{d}}{\mathrm{d}x}|x| = \frac{1}{|x|} \cdot (-1) = \frac{1}{x}.$$

- $\displaystyle\int \frac{\mathrm{d}x}{(x-a)^k} = \frac{-1}{(k-1)(x-a)^{k-1}} + C$ für $x \neq a$ und $k = 2, 3, \dots$
- $\displaystyle\int \frac{\mathrm{d}x}{x^2+px+q} = \frac{2}{\sqrt{4q-p^2}} \arctan\left(\frac{2x+p}{\sqrt{4q-p^2}}\right) + C$

  falls Diskriminante $D = p^2 - 4q < 0$, also keine reelle Nullstelle.

  Denn:

$$\begin{aligned} x^2+px+q &= \left(x+\frac{p}{2}\right)^2 + q - \frac{p^2}{4} \\ &= \left(q - \frac{p^2}{4}\right) \cdot \left(\left(\frac{2x+p}{\sqrt{4q-p^2}}\right)^2 + 1\right) \end{aligned}$$

- $$\begin{aligned} \int \frac{ax+b}{x^2+px+q} &= \frac{a}{2}\ln(x^2+px+q) + \left(b - \frac{ap}{2}\right)\int \frac{\mathrm{d}x}{x^2+px+q} \\ &= \frac{a}{2}\ln(x^2+px+q) + \frac{2b-ap}{\sqrt{4q-p^2}} \arctan\left(\frac{2x+p}{\sqrt{4q-p^2}}\right) + C \end{aligned}$$

Basierend auf dem logarithmischen Integrieren:

$$\int \frac{f'(x)}{f(x)}\,\mathrm{d}x = \ln f(x) + C$$

Hat ein Polynom zweiter Ordnung (der Form $x^2 + px + q$) keine reelle Nullstelle, so ist (der Realteil vom) Funktionswert stets positiv. Das logarithmische Integrieren kann also für positive Zahlen problemlos (und ohne Betragstriche) verwendet werden.

Falls Terme mit zwei nicht reellen Nullstellen in höheren Potenzen vorkommen, kann man dieses Problem rekursiv lösen. Das ist allerdings von Hand mehr als unangenehm.

$$\begin{aligned} \int \frac{\mathrm{d}x}{(x^2+px+q)^k} =& \frac{2x+p}{(k-1)(4q-p^2)(x^2+px+q)^{k-1}} \\ &+ \frac{2(2k-3)}{(k-1)(4q-p^2)} \int \frac{\mathrm{d}x}{(x^2+px+q)^{k-1}} \end{aligned}$$

$$\begin{aligned} \int \frac{ax+b}{(x^2+px+q)^k}\,\mathrm{d}x =& \frac{bp-2aq+(2b-ap)x}{(k-1)(4q-p^2)(x^2+px+q)^{k-1}} \\ &+ \frac{(2k-3)(2b-ap)}{(k-1)(4q-p^2)} \int \frac{\mathrm{d}x}{(x^2+px+q)^{k-1}} \end{aligned}$$

**Beispiel** zurück zum ersten Beispiel zur Partialbruchzerlegung

$$\frac{x^4+x^3+2x}{x^3+x^2+x+1} = x - \frac{1}{x+1} + \frac{1}{x^2+1}$$

$$\begin{aligned}\int \frac{x^4+x^3+2x}{x^3+x^2+x+1}\,\mathrm{d}x &= \int x\,\mathrm{d}x - \int \frac{1}{x+1}\,\mathrm{d}x + \int \frac{1}{x^2+1}\,\mathrm{d}x \\ &= \frac{x^2}{2} - \ln|x+1| + \arctan x + C\end{aligned}$$

Wir haben zunächst mit der Polynomdivision begonnen und anschließend »nur« die oben gefundenen Regeln angewendet um die einzelnen Integrale aufzulösen.

**Beispiel** zuruück zum Beispiel von eben

$$\begin{aligned}\int \frac{x^2+3x+4}{(x-1)^2(x+1)^2} &= A(x+1)(x+1)^2 + B(x+1)^2 + C(x-1)^2(x+1) + D(x-1)^2 \\ &= -\frac{3}{4}(x+1)(x+1)^2 + 2(x+1)^2 + \frac{3}{4}(x-1)^2(x+1) + \frac{1}{2}(x-1)^2 \\ &= -\frac{3}{4}\ln|x-1| - \frac{2}{x-1} + \frac{3}{4}\ln|x-1| - \frac{1}{2}\frac{1}{x+1} + C\end{aligned}$$

Jetzt können wir also alle rationale Funktionen problemlos entlang dieses Schemas integrieren. Das ist jetzt ohne Magie und streng nach Algorithmus machbar.

### 4.2.1 Integration weiterer Funktionenklassen

*Idee: Zurückführung auf rationale Funktionen mit Substitution*

a) $A = \int r(x, \sqrt[n]{ax+b})\,\mathrm{d}x$ mit $a \neq 0$ und $r$ rationaler Ausdruck.

$$\begin{aligned}t = \sqrt[n]{ax+b} \Leftrightarrow x &= \frac{t^n - b}{a} \\ \frac{\mathrm{d}x}{\mathrm{d}t} &= \frac{nt^{n-1}}{a}\end{aligned}$$

$$A = \int r(\frac{t^n-b}{a}, t)\frac{nt^{n-1}}{a}\,\mathrm{d}t$$

**Beispiel**

$$\int \frac{x+\sqrt{x}}{x-\sqrt{x}}\,\mathrm{d}x$$

$$t=\sqrt{x} \Rightarrow x=t^2$$
$$\frac{\mathrm{d}x}{\mathrm{d}t}=2t$$

$$\begin{aligned}\int \frac{x+\sqrt{x}}{x-\sqrt{x}}\,\mathrm{d}x &= \int \frac{t^2+t}{t^2-t}2t\,\mathrm{d}t \\ &= 2\int (t+2)\,\mathrm{d}t + 4\int \frac{\mathrm{d}t}{t-1} = t^2+4t+4\ln|t-1|+C \\ &= \sqrt{x}^2+4\sqrt{x}+4\ln|\sqrt{x}-1|+C \\ &= x+4\sqrt{x}+4\ln|\sqrt{x}-1|+C \quad \text{für} 0<x<1 \text{ oder } 1<x\end{aligned}$$

b) $\int r(e^x)\,\mathrm{d}x$ mit $r$ rationaler Ausdruck

$$t=e^x \Rightarrow x=\ln t$$
$$\frac{\mathrm{d}x}{\mathrm{d}t}=\frac{1}{t}$$

$$\int r(e^x)\,\mathrm{d}x = \int r(t)\frac{\mathrm{d}t}{t}$$

**Beispiel**

$$\int \frac{e^{2x}+1}{e^x-1}\,\mathrm{d}x$$

$$t=e^x \Rightarrow x=\ln t$$
$$\frac{\mathrm{d}x}{\mathrm{d}t}=\frac{1}{t}$$

$$\begin{aligned}\int \frac{e^{2x}+1}{e^x-1}\,\mathrm{d}x &= \int \frac{t^2+1}{t-1}\frac{\mathrm{d}t}{t}\\ &= \int \frac{t^2+1}{t^2-t}\,\mathrm{d}t = \int 1\,\mathrm{d}t + 2\int \frac{\mathrm{d}t}{t-1} - \int \frac{\mathrm{d}t}{t}\\ &= t + 2\ln|t-1| - \ln|t| + C\\ &= e^x + 2\ln|e^x-1| - \ln|e^x| + C\\ &= e^x + 2\ln|e^x-1| - x + C \qquad \text{für } x<0, x>0\end{aligned}$$

c) $\int r(\sin x, \cos x)\,\mathrm{d}x$ mit $r$ rationaler Ausdruck

$$\begin{aligned} t = \tan\frac{x}{2} &\Rightarrow x = 2\arctan t\\ \frac{\mathrm{d}x}{\mathrm{d}t} &= \frac{2}{1+t^2}\end{aligned}$$

$$\begin{aligned}\sin 2x &= 2\sin x\cos x\\ \sin x &= \frac{2\sin\frac{x}{2}\cos\frac{x}{2}}{\underbrace{\sin^2\frac{x}{2}+\cos^2\frac{x}{2}}_{=1}}\\ &= \frac{2}{\frac{\sin\frac{x}{2}}{\cos\frac{x}{2}} + \frac{\cos\frac{x}{2}}{\sin\frac{x}{2}}}\\ &= \frac{2}{t+\frac{1}{t}} = \frac{2t}{t^2+1}\end{aligned}$$

$$\begin{aligned}\cos x &= \frac{\cos^2\frac{x}{2} - \sin^2\frac{x}{2}}{\cos^2\frac{x}{2}+\sin^2\frac{x}{2}}\\ &= \frac{1-\frac{\sin^2\frac{x}{2}}{\cos^2\frac{x}{2}}}{1+\frac{\sin^2\frac{x}{2}}{\cos^2\frac{x}{2}}} = \frac{1-t^2}{1+t^2}\end{aligned}$$

$$\int r(\sin x, \cos x)\,\mathrm{d}x = \int r(\frac{2}{t^2+1}, \frac{1-t^2}{1+t^2})\frac{2\,\mathrm{d}t}{1+t^2}$$

**Beispiel**

$$\begin{aligned}\int \frac{\mathrm{d}x}{\sin x} &= \int \frac{t^2+1}{2t}\frac{2\,\mathrm{d}t}{t^2+1}\\ &= \int \frac{\mathrm{d}t}{t} = \ln|t|+C\\ &= \ln|\tan\frac{x}{2}|+C\end{aligned}$$

d) $\int r(x, \sqrt{ax^2+bx+c})\,\mathrm{d}x$ mit $a>0$, $b^2-4ac\neq 0$ und $r$ rationaler Ausdruck

$$\begin{aligned} t = \sqrt{ax^2+bx+c} - x\sqrt{a} &\Rightarrow x\sqrt{a}t = \sqrt{ax^2+bx+c}\\ \Rightarrow ax^2+bx+c = t^2+ax^2+2tx\sqrt{a} &\Rightarrow x = \frac{t^2-c}{b-2t\sqrt{a}}\\ \frac{\mathrm{d}x}{\mathrm{d}t} &= \frac{2t^2\sqrt{a}+2bt-2\sqrt{ac}}{(b-2t\sqrt{a})^2}\end{aligned}$$

**Beispiel**

$$\int \frac{x-\sqrt{x^2-5x+1}}{x^2}\,\mathrm{d}x$$

$$t = \sqrt{x^2-5x+1}-x$$

$$\begin{aligned}\int \frac{x-\sqrt{x^2-5x+1}}{x^2}\,\mathrm{d}x &= \int \frac{2t^3+10t^2+2t}{(1-t^2)^2}\,\mathrm{d}t\\ &= \dots\end{aligned}$$

weiteres Vorgehen wie im Eingangsbeispiel

**Vorlesung CIV**

Wir haben jetzt an vielen Beispielen gesehen, dass die Integration kein Handwerk sondern vielmehr eine tiefe Magie (oder eine Kunst, ganz wie man will) ist. Hier noch einige Beispiele dafür, dass man einer Funktion alles ansehen kann, aber nicht ihr Integral.

$$\int \frac{\mathrm{d}x}{(x-1)^2} = -\frac{1}{x-1} + C$$
$$\int \frac{\mathrm{d}x}{x^2-1} = \frac{1}{2}\ln\left|\frac{x-1}{x+1}\right| + C$$
$$\int \frac{\mathrm{d}x}{x^2+1} = \arctan x + C$$

## 4.3 Das bestimmte Integral

[Sinnlose Skizze mit Fläche von $a$ bis $b$ unter kryptischer Kurve.]

[$a = x_0$, $b = x_n$, dazwischen $x_1, \ldots, x_{n-1}$]

[Zwischen den $x_i$ jeweils ein $\xi_i$]

[einen Balken einzeichnen]

$$\xi_i \in [x_{i-1}, x_i] \text{ für } i = 1, 2, \ldots, n$$
$$\Delta x_i = x_i - x_{i-1}$$

### Definition 4.4

Wir betrachten eine Funktion $f : [a,b] \to \mathbb{R}$. Konvergiert

$$\sum_{i=1}^{n} f(\xi_i)\Delta x_i$$

für $\max_i \Delta x_i \to 0$ und beliebige Wahl von $\xi_i \in [x_{i-1}, x_i]$ gegen den gleichen Grenzwert, so nennen wir diesen Grenzwert das *Riemann-Integral* von $f$ über $[a,b]$.

Wir definieren die Schreibweise

$$\sum_{i=1}^{n} f(\xi_i)\Delta x_i \underset{\to}{\max \Delta x_i} \to 0 \int_a^b f(x)\,\mathrm{d}x$$

und schreiben weiter $f \in \mathscr{R}([a,b])$.

Dieses Integral hat bisher *nichts* mit dem unbestimmten Integral zu tun. Es wird erstmal nur die gleiche Symbolik verwendet.

Ferner gill:

$$\int_a^a f(x)\,\mathrm{d}x = 0, \quad \int_b^a f(x)\,\mathrm{d}x = -\int_a^b f(x)\,\mathrm{d}x$$

Wir können das Riemann-Integral als Flächeninhalt unter dem Graphen betrachten. Im Allgemeinen wird dieses Konzept aber schwierig, wenn man Vorzeichen behaftete Flächeninhalte – also Graphen unterhalb der $x$-Achse betrachtet:

$$\int_0^{2\Pi} \sin x \,\mathrm{d}x = 0$$

**Beispiel**

1. $f(x) = 1$ in $[a,b]$

$$\sum_{i=1}^{n} \Delta x_i = x_n - x_0 = b - a \longrightarrow b - a$$

Die konstante Funktion 1 ist also Riemann integrierbar, denn der Grenzwert existiert immer und ist immer genau $b-a$. (Schließlich betrachtet wir ein Rechteck über der $x$-Achse mit der Höhe 1.)

2. Dirichlet-Funktion

$$f(x) = \begin{cases} 1 & \text{für } x \in \mathbb{Q} \\ 0 & \text{sonst} \end{cases}$$

$$\sum_{i=1}^{n} f(\xi_i)\Delta x_i = \begin{cases} b-a & \text{für } \xi_i \in \mathbb{Q} \forall i \\ 0 & \text{für } \xi_i \notin \mathbb{Q} \forall i \end{cases}$$

Wir erhalten hier verschiedene Grenzwerte, je nach dem, wie man die $x_i$ wählt. $f$ ist also nicht Riemann integrierbar.

$$f \notin \mathscr{R}([a,b])$$

### 4.3.1 Eigenschaften des Riemann-Integrals

**Satz 4.5**

Seien $f, g \in \mathscr{R}([a,b])$.

a) $\int_a^b (\alpha f(x) + \beta g(x))\,\mathrm{d}x = \alpha \int_a^b f(x)\,\mathrm{d}x + \beta \int_a^b g(x)\,\mathrm{d}x \forall \alpha, \beta \in \mathbb{R}$

b) $\int_a^b f(x)\,\mathrm{d}x = \int_a^c f(x)\,\mathrm{d}x + \int_c^b f(x)\,\mathrm{d}x$

c) Sei $f(x) \le g(x) \forall x \in [a,b]$, dann gilt

$$\int_a^b f(x)\,\mathrm{d}x \le \int_a^b g(x)\,\mathrm{d}x.$$

d) $f \in C([a,b]) \Rightarrow f \in \mathscr{R}([a,b])$.

Ist $f$ eine stetige Funktion auf dem Intervall $[a,b]$, dann ist $f$ auch Riemann integrierbar. Die Umkehrung gilt *nicht*. Es gibt durchaus unstetige Funktionen, die sich integrieren lassen.

e) (Folgerung aus c))

[Bild mit krytoschem Grapgen über $x$-Acchse und unterer sowie oberer Schranke.]

Gilt $m \le f(x) \le M \quad \forall x \in [a,b]$, dann wissen wir

$$m(b-a) \le \int_a^b f(x)\,\mathrm{d}x \le M(b-a).$$

f) Dreiecksungleichung:

$$\forall a \le b : \left| \int_a^b f(x)\,\mathrm{d}x \right| \le \int_a^b |f(x)|\,\mathrm{d}x$$

Denn es gilt für alle $x$: $-|f(x)| \le f(x) \le |f(x)|$. Also gilt mit Aussage c) auch

$$-\int_a^b |f(x)|\,\mathrm{d}x \le \int_a^b f(x)\,\mathrm{d}x \le \int_a^b |f(x)|\,\mathrm{d}x.$$

**Beweis** ohne Beweis, bis auf den Beweis im Satz – argh! □

**Satz 4.6** Mittelwertsatz der Integralrechnung

Sei $f \in C([a,b])$, $g \in \mathscr{R}([a,b])$ und $g(x) \ge 0$ für alle $x \in [a,b]$, dann gilt

$$\exists \xi \in [a,b] : \quad \int_a^b f(x)g(x)\,\mathrm{d}x = f(\xi)\int_a^b g(x)\,\mathrm{d}x$$

**Beispiel**

$$g(x) = 1$$

$$\int_a^b f(x)\,\mathrm{d}x = f(\xi)(b-a)$$

[Skizze einer stetigen Funktion über $[a,b]$ und $g$ parallel zur $x$-Achse schneidet $f$, sodass Fläche unter $g$ gerade Fläche unter $f$]

**Beweis** Seien $m = \min_{a \leq x \leq b} f(x)$ und $M \max_{a \leq x \leq b} f(x)$, dann gilt mit $g \geq 0$

$$\Rightarrow mg(x) \leq f(x)g(x) \leq Mg(x)$$
$$\Rightarrow m \int_a^b g(x)\,\mathrm{d}x \leq \int_a^b f(x)g(x)\,\mathrm{d}x \leq M \int_a^b g(x)\,\mathrm{d}x$$

Sei $c$ definiert durch

$$c \int_a^b g(x)\,\mathrm{d}x := \int_a^b f(x)g(x) \text{ mit } m \leq c \leq M.$$

Mit dem Zwischenwertsatz (Satz 2.8 c) [Verweis]) gilt

$$\exists \xi \in [a,b] \text{ mit } c = f(\xi)$$

Also wissen wir

$$\exists \xi \in [a,b] \text{ mir } f(\xi) \int_a^b g(x)\,\mathrm{d}x = \int_a^b f(x)g(x)\,\mathrm{d}x.$$

Wir müssen noch einen Sonderfalls berachten:

Falls $\int_a^b g(x)\,\mathrm{d}x = 0$ kann man alle $c$ mit $m \leq c \leq M$ wählen. □

**Korollar 4.7**

Sei $f \in C([a,b])$ dann gilt

$$\exists \xi \in [a,b] : \int_a^b f(x)\,\mathrm{d}x = f(\xi)(b-a)$$

**Beweis** siehe Beispiel weiter oben □

## 4.4 Hauptsätze der Integralrechnung

Wir suchen jetzt den Zusammenhang zwischen der Stammfunktion (also der der unbestimmten) und der bestimmten Integration.

## Satz 4.8

Sei $f \in C([a,b])$, dann besitzt $f$ eine Stammfunktion $F : [a,b] \to \mathbb{R}$ mit

$$F(x) \int_a^x f(t)\,\mathrm{d}t$$

Dass heißt

$$F'(x) = \frac{\mathrm{d}}{\mathrm{d}x} \int_a^x f(t)\,\mathrm{d}t = f(x)$$

Eigentlich zu zeigen ist die Gleichheit am Ende des Satzes:

$$F'(x) = \frac{\mathrm{d}}{\mathrm{d}x} \int_a^x f(t)\,\mathrm{d}t = f(x)$$

**Beweis** Wir bilden

$$\begin{aligned}\frac{F(x+h)-F(x)}{h} &= \frac{1}{h}\left(\int_a^{x+h} f(t)\,\mathrm{d}t - \int_a^x f(t)\,\mathrm{d}t\right)\\ &= \frac{1}{h}\int_x^{x+h} f(t)\,\mathrm{d}t\\ &\overset{\S}{=} \frac{1}{h} f(\xi)h = f(\xi) \xrightarrow[|h|\to 0]{} f(x)\end{aligned}$$

denn $|x-\xi| \leq |h|$. □

An der Stelle § wird Korollar 4.7 [Verweis] verwendet.

Jetzt passierts – wir führen alles zusammen

## Korollar 4.9

Sei $f \in C([a,b])$ mit der Stammfunktion $F$, dann gilt

$$\int_a^b f(x)\,\mathrm{d}x = F(b) - F(a) =: F(x)|_x = a^x = b =: F(x)|_a^b$$

**Beweis** Nach Korollar 4.9 existiert eine Stammfunktion $F_1$ mit

$$F_1(x) = \int_a^x f(t) + C\,\mathrm{d}t$$

das heißt für alle $F$ gilt

$$F(x) = \int_a^x f(t)\,\mathrm{d}t + C$$

Setzen wir $x = a$ ein, ergibt sich

$$F(a) = C \quad (\text{denn } \int_a^b f(x)\,\mathrm{d}x = 0)$$

Setzen wir $x = b$ ein, ergibt sich

$$F(b) = \int_a^b f(t)\,\mathrm{d}t + C$$

Zusammen gilt

$$F(b) - F(a) = \int_a^b f(t)\,\mathrm{d}t$$

□

**Vorlesung CV**

**Beispiel** einfach

$$\begin{aligned}\int_0^\pi \sin x \,\mathrm{d}x &= -\cos x|_0^\pi \\ &= -\cos\pi - (-\cos 0) = 2\end{aligned}$$

**Beispiel** aus der Physik

Welche Arbeit $A$ ist erforderlich um eine Rakete von der Erdoberfläche senkrecht nach oben auf die Höhe $h$ zu bringen?

Arbeit $=$ Kraft $\cdot$ Weg

Erdanziehungskraft:

$$k(x) = \frac{mM}{x^2}$$

mit

$x$: Entfernung zum Erdmittelpunkt
$M$: Erdmasse
$m$: Raktenmasse
$R$: Erdradius

$$
\begin{aligned}
A &= \int_R^{R+h} \frac{\gamma M m}{x^2}\,\mathrm{d}x \\
&= \left.\frac{-\gamma m M}{x}\right|_R^{R+h} \\
&= \gamma m M \left(\frac{1}{R} - \frac{1}{R+h}\right) \xrightarrow[h\to\infty]{} \frac{\gamma m M}{R}
\end{aligned}
$$

### 4.4.1 Substitution und partielle Integration bei bestimmten Integralen

a) **Substitution**

**1. Weg** Betrachten das unbestimmte Integral und bilden die Stammfunktion $F$ und bilden danach $F(b) - F(a)$.

**Beispiel** Substitution

$$\int_0^{\pi/2} \sin^2 x \cos x \,\mathrm{d}x$$

Wir setzen für die Substitution $u = \sin x$ und erhalten

$$\mathrm{d}u = \cos x \,\mathrm{d}x.$$

Damit erhalten wir das unbestimmte Ingeral

$$
\begin{aligned}
\int \sin^2 x \cos x \,\mathrm{d}x &= \int u^2 \,\mathrm{d}u \\
&= \frac{u^3}{3} + C \\
&= \frac{\sin^3 x}{3} + C
\end{aligned}
$$

Also ergibt sich für das bestimmte Integral

$$\int_0^{\pi/2} \sin^2 x \cos x \,\mathrm{d}x = \left.\frac{\sin^3 x}{3}\right|_0^{\pi/2} = \frac{1}{3}$$

**2. Weg** Substitution im bestimmten Integral *mit* Substitution der Grenzen

**Satz 4.10** Substitution im bestimmten Integral

Gegeben sei $f \in \mathscr{R}([a,b]), f \in C^1([\alpha,\beta])$ mit $\varphi : [\alpha,\beta] \to [a,b]$ und $\varphi$ streng monoton wachsend, also $\varphi(\alpha) = a\varphi(\beta) = b$. Dann gilt

$$\int_a^b f(x)\,\mathrm{d}x = \int_\alpha^\beta f(\varphi(t))\varphi'(t)\,\mathrm{d}t.$$

**Beweis** Sei $F$ Stammfunktion von $f$, das heißt nach Korollar 4.9 auf Seite 126

$$\int_a^b f(x)\,\mathrm{d}x = F(b) - F(a)$$

dann gilt

$$\begin{aligned}\int_\alpha^\beta f(\varphi(t))\varphi'(t)\,\mathrm{d}t &= F(\varphi(t))|_\alpha^\beta \\ &= F(\varphi(\beta)) - F(\varphi(\alpha)) \\ &= F(b) - F(a)\end{aligned}$$

□

**Beispiel**

Wir betrachten das Integral

$$\int_0^{\pi/2} \sin^2 \cos x\,\mathrm{d}x$$

mit der Substitution $u = \sin x$, also $\mathrm{d}u = \cos x\,\mathrm{d}x$.

$\sin x$ auf dem abgeschlossenen Intervall $[0,\pi/2]$ streng monoton. Weiter gilt $\varphi(0) = 0$ und $\varphi(\pi/2) = 1$

$$\int_0^{\pi/2} \sin^2 \cos x\,\mathrm{d}x = \left.\frac{u^3}{3}\right|_0^1 = \frac{1}{3}$$

**Beispiel**

$$\int_1^e \frac{\ln^4 x}{x}\,\mathrm{d}x$$

**1. Weg** Wir setzen $u = \ln x$ und erhalten $\mathrm{d}u = \frac{1}{x}\,\mathrm{d}x$ und erhalten

$$\begin{aligned}\int \frac{\ln^4 x}{x}\,\mathrm{d}x &= \int u^4\,\mathrm{d}u \\ &= \frac{u^5}{5} + C = \frac{\ln^5 x}{5} + C\end{aligned}$$

Also ergibt sich

$$\int_1^e \frac{\ln^4 x}{x}\,\mathrm{d}x = \left.\frac{\ln^5 x}{5}\right|_1^e = \frac{1}{5}$$

**2. Weg** Wir setzen $u = \ln x$ und erhalten $\mathrm{d}u = \frac{1}{x}\,\mathrm{d}x$, $\ln 1 = 0$ und $\ln e = 1$. Weiter ist $\ln$ streng monoton wachsend. Es ergibt sich

$$\int_1^e \frac{\ln^4 x}{x}\,\mathrm{d}x = \int_0^1 u^4\,\mathrm{d}u = \left.\frac{u^5}{5}\right|_0^1 = \frac{1}{5}$$

**Beispiel** Berechnung der Halbkreisfläche

$$\begin{aligned}y &\geq 0 \\ y^2 + x^2 &= r^2 \\ y^2 &= r^2 - x^2 \\ y &= \sqrt{r^2 - x^2}\end{aligned}$$

$$A = \int_{-r}^{r} \sqrt{r^2 - x^2}\,\mathrm{d}x$$

Wir substituieren $x = r\sin t$ und erhalten $\mathrm{d}x = \cos t\,\mathrm{d}t$. Weiter ist $\sin t$ monoton auf dem Intervall $[-r, r]$.

$$\begin{aligned}
\int_{-\pi/2}^{\pi/2} \sqrt{r^2 - r^2 \sin^2 t}\, r \cos t \,\mathrm{d}t &= r^2 \int_{-\pi/2}^{\pi/2} \sqrt{1-\sin^2 t} \cos t \,\mathrm{d}t \\
&= r^2 \int_{-\pi/2}^{\pi/2} \cos^2 t \,\mathrm{d}t \\
&= \frac{r^2}{2} \int_{-\pi/2}^{\pi/2} (1+\cos 2t)\,\mathrm{d}t \\
&= \frac{r^2}{2} \left( t + \frac{\sin 2t}{2} \right) \Big|_{-\pi/2}^{\pi/2} \\
&= \frac{\pi r^2}{2}
\end{aligned}$$

Wir können uns dieses Integral auch durch folgenden Trick überlegen.

$$\begin{aligned}
\int_{-\pi/2}^{\pi/2} \sin^2 x \,\mathrm{d}x &= \int_{-\pi/2}^{\pi/2} \cos^2 x \,\mathrm{d}x \\
&= \int_{-\pi/2}^{\pi/2} \frac{\sin^2 x + \cos^2 x}{2} \,\mathrm{d}x \\
&= \int_{-\pi/2}^{\pi/2} \frac{1}{2} \,\mathrm{d}x \\
&= \frac{\pi}{2}
\end{aligned}$$

b) **partielle Integration**

**Satz 4.11** partielle Integration von bestimmten Integralen

Seien $f,g \in C^1([a,b])$, dann

$$\int_a^b f(x)g'(x)\,\mathrm{d}x = \underbrace{f(x)g(x)|_a^b}_{f(b)g(b)-f(a)g(a)} - \int_a^b f'(x)g(x)\,\mathrm{d}x$$

**Beweis** Nach Korollar 4.9 auf Seite 126 gilt

$$\begin{aligned}
\int_a^b f'(x)g(x)\,\mathrm{d}x + \int_a^b f(x)g'(x)\,\mathrm{d}x &\overset{\P}{=} \int_a^b (f(x)g(x))'\,\mathrm{d}x \\
&= f(x)g(x)|_a^b
\end{aligned}$$

An der Stelle ¶ wurde die Produktregel innerhalb der Integrale angewendet.

□

## 4.5 Uneigentliche Integrale

bisher $\int_a^b f(x)$ nur erklärt falls

- $[a,b]$ endlich
- $f(x)$ beschränkt ist

a) uneigentliche Integrale mit unendlichen Grenzen

**Definition 4.12**

Sei $a$ fest vorgegeben und sei $f \in \mathscr{R}([a,b])$ für alle $b > a$.

$$\int_a^\infty f(x)\,\mathrm{d}x := \lim_{b\to\infty} \int_a^b f(x)\,\mathrm{d}x$$

Existiert der Grenzwert sagen wir »das uneigentliche Integral konvergiert« (ansonsten »divergiert«).

Analog

$$\int_{-\infty}^b f(x)\,\mathrm{d}x := \lim_{a\to-\infty} \int_a^b f(x)\,\mathrm{d}x$$

$$\int_{-\infty}^\infty f(x)\,\mathrm{d}x := \int_{-\infty}^c f(x)\,\mathrm{d}x + \int_c^\infty f(x)\,\mathrm{d}x$$

Unabhängig von der konkreten Wahl von $c$.

**Beispiel** Standardbeispiel

$$\begin{aligned}\int_{-\infty}^\infty \frac{1}{x^2+1}\,\mathrm{d}x &= \int_{-\infty}^c \frac{1}{x^2+1}\,\mathrm{d}x + \int_c^\infty \frac{1}{x^2+1}\,\mathrm{d}x \\ &= \lim_{a\to-\infty}(\arctan c - \arctan a) + \lim_{b\to\infty}(\arctan b - \arctan c) = \pi\end{aligned}$$

[Plot von $f(x)$ und $\arctan x$ und Asymptoten $\pi/2$ und $-\pi/2$]

**Beispiel** Negativbeispiel

$$\int_0^\infty 1\,\mathrm{d}x = \lim_{b\to\infty} x|_0^b = \infty$$

Das uneigentliche Integral konvergiert also nicht. Es divergiert gegen $\infty$.

**Beispiel** Signumfunktion

Sei die Funktion $\operatorname{sgn} : \mathbb{R} \to \{-1, 0, 1\}$ gegeben durch

$$\operatorname{sgn}(x) = \begin{cases} -1 & \text{für } x < 0 \\ 0 & \text{für } x = 0 \\ 1 & \text{für } x > 0 \end{cases}$$

$$\int_{-\infty}^{\infty} \operatorname{sgn} x \,\mathrm{d}x = \int\limits_{\mathbb{R}} \operatorname{sgn} x \,\mathrm{d}x$$

ist divergent, da von zwei Summanden bereits der positive Bereich divergent ist (siehe vorheriges Beispiel).

Aber mit dem Cauchy'scher Hauptwert

$$CH \int_{-\infty}^{\infty} f(x) \,\mathrm{d}x = \lim_{c \to \infty} \int_{-c}^{c} f(x) \,\mathrm{d}x$$

ergibt sich

$$\mathrm{CH} \int_{-\infty}^{\infty} \operatorname{sgn} x \,\mathrm{d}x = 0$$

Achtung: Der Cauchy'sche Hauptwert ist *nicht* unsere Definition von einem uneigentlichen Integral. Wir betrachten stets *zwei* Grenzwerte und in diesem Fall kann der Grenzwert nicht angegeben werden, da beide Grenzwerte divergent sind.

b) Uneigentliche Integrale mit unbeschränkten Integranden

Es sei $[a, b]$ endliches Intervall mit Polstelle $b$ von $f$.

**Definition 4.13**

Sei $f \in \mathscr{R}([a, \beta])$ für alle $\beta$ mit $a < \beta < b$.

$$\int_a^b f(x) \,\mathrm{d}x = \lim_{\beta \to b-} \int_a^{\beta} f(x) \,\mathrm{d}x$$

Existiert der Grenzwert sagen wir »das uneigentliche Integral konvergiert« (ansonsten »divergiert«).

Analog für Polstellen in $a$ und innerhalb von $[a, b]$.

**Beispiel**

$$\begin{aligned}\int_0^1 \frac{1}{\sqrt{1-x^2}}\,\mathrm{d}x &= \lim_{\beta\to -1}(\arcsin\beta - \arcsin 0)\\ &= \frac{\pi}{2} - 0 = \frac{\pi}{2}\end{aligned}$$

Die Funktion passt genau in das Schema, denn sie hat bei 1 eine Polstelle.

[Skizze von Sinus und Arcussinus von $-\pi/2$ bis $\pi/2$]

**Beispiel**

$$\begin{aligned}\int_0^1 \ln x\,\mathrm{d}x &= \lim_{\alpha\to 0+}\int_\alpha^1 \ln x\,\mathrm{d}x\\ &= \lim_{\alpha\to 0+} x\ln x - x|_\alpha^1\\ &= -1\end{aligned}$$

Denn mit L'Hospital gilt

$$\lim_{\alpha\to 0+}\frac{\ln\alpha}{\frac{1}{\alpha}} = \lim_{\alpha\to 0+}\frac{\frac{1}{\alpha}}{-\frac{1}{\alpha^2}} = 0$$

**Beispiel** Verallgemeinerung des vorherigen Beispiels

Wir betrachten drei Fälle. Für $\gamma = 1$ ergibt sich

$$\int_0^1 \frac{1}{x^\gamma} = \ln x|_\alpha^1 = \text{ divergent}$$

Für $gamma \neq 1$ ergibt sich

$$\int_0^1 \frac{1}{x^\gamma} = \lim_{\alpha\to 0+}\left.\frac{x^{-\gamma+1}}{1-\gamma}\right|_\alpha^1 = \begin{cases}\frac{1}{1-\gamma} & \text{für } \gamma < 1\\ \text{divergent} & \text{für } \gamma > 1\end{cases}$$

**Beispiel** und noch eins...

$$\int_{-3}^{3} \frac{1}{x^2-1} = \int_{-3}^{-1} + \int_{-1}^{0} + \int_{0}^{1} + \int_{1}^{3} \frac{1}{x^2+1}\,\mathrm{d}x$$
$$= \lim_{\beta\to -1-} \frac{1}{2} \ln \left| \frac{x-1}{x+1} \right| \Bigg|_{-3}^{\beta} + \ldots \quad \text{divergent!}$$

Denn mit der Partialbruchzerlegung gilt

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

Aber Cauchy'scher Hauptwert existiert:

$$\mathrm{CH} \int_{-3}^{3} \frac{1}{x^2-1}\,\mathrm{d}x = \lim_{\alpha\to 0+} \Bigg( \int_{-3}^{-1-\alpha} \frac{1}{x^2-1}$$
$$+ \int_{-1+\alpha}^{1-\alpha} \frac{1}{x^2-1}\,\mathrm{d}x + \int_{1+\alpha}^{3} \frac{1}{x^2-1}\,\mathrm{d}x = -\ln 2 \Bigg)$$

[Skizze]

c) Vergleichssätze (nur am Beispiel)

$$\int_{1}^{\infty} \frac{\sqrt{x+1}}{x^3+x^2+5}$$

Für hinreichend große $x$ gilt

$$\frac{\sqrt{x+1}}{x^3+x^2+5} < \frac{1}{x^2}$$

also folgt

$$\int_{1}^{\infty} \frac{\sqrt{x+1}}{x^3+x^2+5}\,\mathrm{d}x < c + \int_{1}^{\infty} \frac{1}{x^2}\,\mathrm{d}x < \infty$$

**Beispiel** Die Gamma-Funktion

$$n! = \prod_{k=1}^{n} k, \quad 0! = 1$$

Die Fakultät ist nur diskret definiert. Gibt es eine stetige Funktion, die die Punkte der Fakultät enthält? Ja!

Wir definieren die Gamma-Funktion als uneigentliches Integral für $x > 0$

$$\Gamma(x) = \int_0^\infty t^{x-1} e^{-t}\,\mathrm{d}t = \int_0^1 t^{x-1} e^{-t}\,\mathrm{d}t + \int_1^\infty t^{x-1} e^{-1}\,\mathrm{d}t$$

Beide Teilintegrale konvergieren, denn

$$\int_0^1 t^{x-1} e^{-t}\,\mathrm{d}t < \int_0^1 t^{x-1}\,\mathrm{d}t < \infty$$

$$\int_1^\infty t^{x-1} e^{-t}\,\mathrm{d}t < \infty \text{ denn } \lim_{t\to\infty} \frac{t^{x-1}}{e^t} = 0$$

Damit folgt

$$\exists t_0 \forall t > t_0 \frac{t^{x-1}}{e^t} < 1 \Rightarrow \frac{t^{x-1}}{e^t} < \frac{1}{t^2}$$

$$\text{aber } \int_1^\infty \frac{1}{t^2}\,\mathrm{d}t < \infty$$

Es gilt

$$\Gamma(1) = \int_0^\infty e^{-t}\,\mathrm{d}t = -e^{-t}|_0^\infty = 1 = 1!$$

Und für $0 < a < b < \infty$ gilt

$$\begin{aligned}\int_a^b t^x e^{-t}\,\mathrm{d}t &= -t^x e^{-t}|_a^b + \int_a^b x t^{x-1} e^{-t}\,\mathrm{d}t \\ &= a^x e^{-a} - b^x e^{-b} + \int_a^b x t^{x-1} e^{-t}\,\mathrm{d}t\end{aligned}$$

Betrachten wir den Grenzwert $a \to 0+$ und $b \to \infty$, dann ergibt sich

$$\Gamma(x+1) = x\Gamma(x) \quad \forall x > 0$$

Damit gilt

$$\begin{aligned}\Gamma(2) &= 1 \\ \Gamma(3) &= 2 \\ \Gamma(4) &= 6 \\ \Gamma(5) &= 24\end{aligned}$$

Mit vollständiger Induktion kann man leicht zeigen, dass für $n \in \mathbb{N}$ gilt

$$\Gamma(n) = (n-1)!$$

**Vorlesung CVI**

Wir betrachten nun einen speziellen Zwischenwert. Wer hat sich nicht schon immer gefragt: »Was ist die Fakultät von $\frac{1}{2}$?«

$$\Gamma(\frac{1}{2}) = \int_0^\infty \frac{e^{-t}}{\sqrt{t}}$$

Mit der Substitution $\sqrt{t} = x, \frac{\mathrm{d}t}{\mathrm{d}x} = 2x$ ergibt sich

$$\begin{aligned}\Gamma(\frac{1}{2}) &= \int_0^\infty \frac{e^{-x^2}}{2x}\,\mathrm{d}x \\ &= \underset{\int}{\mathbb{R}} e^{-x^2}\,\mathrm{d}x = \sqrt{\pi}\end{aligned}$$

**Beispiel** Integralsinus

$$\mathrm{Si} := \int_0^x \frac{\sin t}{t}\,\mathrm{d}t$$

Der Integralsinus lässt sich mit uns bekannten Funktionen nicht mit endlichen Schritten angeben. Wir führen also einfach ein neues Symbol ein statt eine Stammfunktion anzugeben.

Wir betrachten

$$\lim_{x \to a} \mathrm{Si}(x)$$

[Plot]

$$\begin{aligned}\int_0^\infty \frac{\sin t}{t}\,\mathrm{d}t &= \int_0^1 \frac{\sin t}{t}\,\mathrm{d}t + \lim_{x\to\infty}\left(\left.\frac{-\cos t}{t}\right|_1^x - \int_1^x \frac{\cos t}{t^2}\,\mathrm{d}t\right) \\ &= \int_0^1 \frac{\sin t}{t}\,\mathrm{d}t + \cos 1 - \underbrace{\int_1^\infty \frac{\cos t}{t^2}\,\mathrm{d}t}_{<\infty} < \infty \\ &\overset{\|}{=} \frac{\Pi}{2}\end{aligned}$$

Die Gleichheit an der Stelle $\|$ ergibt sich an dieser Stelle nicht. Dafür werden später folgende Sätze benötigt. Das steht hier also einfach nur so zum Spaß.

**aber**

$$\int_0^\infty \left|\frac{\sin t}{t}\right|\,\mathrm{d}t = \infty, \quad \int_0^\infty \left|\frac{\cos t}{t^2}\right|\,\mathrm{d}t < \infty$$

# 5 Funktionenreihen, Funktionenfolgen

### Wiederholung

Zahlenreihen: $s_n = \sum_{k=0}^{n} a_k$

Wir haben definiert $\sum_{n=0}^{\infty} := \lim_{n\to\infty} s_n$

Jetzt betrachten wir mit der gleichen Idee Reihen von Funktionen

Gegeben sei eine Folge $(f_k(x))_{k=0}^{\infty}$ $\qquad f_k : D \to \mathbb{R}(\mathbb{C})$ mit $D \subseteq \mathbb{R}(\mathbb{C})$.

Für jedes feste $x$ erhalten wir eine Zahlenfolge. Konvergiert

$$s_n(x) = \sum_{k=0}^{n} f_k(x)$$

für alle $x \in D$ gegen $s(x)$ so heißt die Funktionenfolge bzw. -reihe *punktweise konvergent* gegen die Summe

$$s(x) = \sum_{k=0}^{\infty} f_k(x)$$

**Beispiel**

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

**Beispiel**

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

**Beispiel**

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

Frage: Wie übertragen sich Eigenschaften der Summanden auf die Summe?

**Beispiel**

Wir betrachten eine unendliche Summe und bilden die Ableitung der Summanden

$$\begin{aligned}\sum_{k=0}^{\infty} \frac{\mathrm{d}}{\mathrm{d}x} \frac{x^k}{k!} &= \sum_{k=1}^{\infty} \frac{x^{k-1}}{(k-1)!} \\ &= \sum_{k=0}^{\infty} \frac{x^k}{k!} \\ &= e^x \overset{**}{=} \frac{\mathrm{d}}{\mathrm{d}x} \sum_{k=0}^{\infty} \frac{x^k}{k!}\end{aligned}$$

Die Gleichheit an der Stelle $**$ gilt auf Grund komplizierter Beweise in Kapitel 2.

Es ergibt sich also die Frage, wann man aus der Ableitung der Summanden die Ableitung der unendlichen Summe folgern kann.

## 5.1 Gleichmäßige Konvergenz

**Definition 5.1**

Für alle $k = 0, 1, 2, \ldots$ sei eine beschränkte Funktion $f_k : [a,b] \to \mathbb{R}$. Die Folge $(s_n(x))_{n=0}^{\infty}$ bzw. die Reihe

$$\sum_{k=0}^{\infty} f_k(x)$$

heißt gegen $x(x)$ *gleichmäßig Konvergent* in $[a,b]$, wenn gil

$$\forall \varepsilon > 0 \exists N = N(\varepsilon) \forall x \in [a,b] \forall n \geq N : |s(x) - s_n(x)| < \varepsilon$$

In dieser Definition ist $N$ unabhängig von $x$!

Anders dargestellt gilt

$$\begin{aligned}\text{punktweise} \forall x \in [a,b] \exists \text{ Nullfolge } (a_n(x))_{n \in \mathbb{N}_0} \text{ mit } |s_n(x) - s(x)| < a_n(x) \\ \text{gleichmäßig} \exists \text{ Nullfolge } (a_n(x))_{n \in \mathbb{N}_0} \forall x \in [a,b] \text{ mit } |s_n(x) - s(x)| < a_n(x)\end{aligned}$$

Noch anders dargestellt gilt

$$\text{punktweise} \forall x \in [a,b] \lim_{n \to \infty} |s_n(x) - s(x)| = 0$$

$$\text{gleichmäßig} \forall (x_1) < [a,b] \lim_{n \to \infty} |s_n(x_n) - s(x_n)| = 0$$

### Beispiel

$$
\begin{aligned}
f_n(x) &= x + \frac{1}{n} \quad x \in [a,1] \\
\lim_{n\to\infty} f_n(x) &= x = f(x) \\
|f(x) - f_n(x)| &= \frac{1}{n}
\end{aligned}
$$

$1/n$ ist unabhängig von $x$ eine Nullfolge.

### Beispiel

$$
\begin{aligned}
f_n(x) &= x^n \quad x \in [a,1] \\
f(x) := \lim_{n\to\infty} f_n(x) &= \begin{cases} 0 & \text{für } 0 \le x < 1 \\ 1 & \text{für } x = 1 \end{cases} \\
|f_n(x) - f(x)| &= \begin{cases} x^n & \text{für } 0 \le x < 1 \\ 0 & \text{für } x = 1 \end{cases}
\end{aligned}
$$

Angenommen $\forall \varepsilon > 0 \forall n \ge N(\varepsilon) \forall x \in [a,1] |f_n(x) - f(x)| < \varepsilon$

Wähle $\varepsilon < 1$ und $x = \sqrt[N]{\varepsilon} < 1$, dann gilt $|f_n(x) - f(x)| = \varepsilon$

Also gilt keine gleichmäßige Konvergenz.

### Beispiel

$$
\begin{aligned}
f_k(x) &= \frac{x^2}{(1+x^2)^k} \qquad \text{für } |x| \le 1 \\
s_n(x) &= \sum_{k=0}^{n} \frac{x^2}{(1+x^2)^k} \\
&= x^2 \frac{1 - (1+x^2)^{-n-1}}{1 - \frac{1}{1+x^2}} \\
&= x^2 \frac{1 + x^2 - (1+x^2)^{-n}}{1 + x^2 - 1} \\
&= 1 + x^2 - \frac{1}{(1+x^2)n} \xrightarrow[n\to\infty]{} s(x) = \begin{cases} 1 + x^2 & \text{für } x \neq 0 \\ 0 & \text{für } x = 0 \end{cases}
\end{aligned}
$$

Also ergibt sich

$$|s(x)-s_n(x)| = \begin{cases} \int(1+x^2)^{-n} & \text{für } x \neq 0 \\ 0 & \text{für } x = 0 \end{cases}$$

Es gilt $\forall x \in [-1,1] \forall n \geq N |s(x)-s_n(x)| < \varepsilon$ ist äquivalent zu $\forall x \in [-1,1] |s(x) - s_N(x)| < \varepsilon$ ist äquivalent zu $\forall x \in [-1,1] \backslash \{0\} : (1+x^2) {-} N < \varepsilon$ ist äauivalent zu $\forall x \in [-1,1] \backslash \{0\} N > \frac{-\ln \varepsilon}{\ln(1+x^2)}$. Also kann ich nicht nur in Abhängigkeit von $\varepsilon$ ein $x$ finden. Fazit $N = N(\varepsilon, x)$, also liegt keine gleichmäßige Konvergenz vor.

**Vorlesung CVII**

### Beispiel

$$s_n = 1+x^2-(1+x^2)^{-n}$$
$$s(x)-s_n(x) = (1+x^2)^{-n}$$

Hier ist keine gleichmäßige Konvergenz vorhanden.

[Ich glaube das war die Wiederholung vom letzten Mal?]

### Satz 5.2

[Ist das Satz 5.2?]

Gegeben $f_x \in C([a,b]), k \in \mathbb{N}_0$ und

$$\sum_{k=0}^{\infty} f_k(x)$$

gleichmäßig konvergent in $[a,b]$.

Dann gilt

$$s(x) = \sum_{k=0}^{\infty} f_k(x) \in C([a,b])$$

Wenn die Summe gleichmäßig konvergent und die Funktion stetig ist, dann ist auch die Summe stetig.

**Beweis** Sei $s(x) = s_n(x) + r_n(x)$. Wir betrachten

$$\begin{aligned} |s(x)-s(y)| &\leq |s_n(x)-s_n(y)| + |r_n(x)-r_n(y)| \\ &\leq |s_n(x)-s_n(y)| + |r_n(x)| + |r_n(y)| \quad (*) \end{aligned}$$

Sei jetzt $\varepsilon > 0$ vorgegeben.

a) $\exists N(\varepsilon)$ mit $\forall x \in [a,b]$ gilt $\forall n \geq N : r_n(x) < \varepsilon$ (gleichmäßige Konvergenz)

b) $\exists \delta(\varepsilon)$ mit $\forall |x-y| < \delta$ gilt $|s_N(x) - s_N(y)| < \varepsilon$ (gleichmäßige Stetigkeit in $[a,b]$, Satz 2.8 d))

Setze $n = N$ in $(*)$, dann ergibt sich $\forall x, y \in [a,b]$ mit $|x-y| < \delta$ gilt $|s(x) - s(y)| < 3\varepsilon$ $\square$

**Bemerkung** Gegenbeispiele bereits behandelt.

**Umformulierung für Funktionenfolgen**

$$\lim_{x \to x_0} \lim_{n \to \infty} s_n(x) = \lim_{n \to \infty} \lim_{x \to x_0} s_n(x)$$

Obige Aussage (»Wenn die gleichmäßige Konvergent erfüllt ist, dann kann die Grenzwertbildung vertauscht werden«) ist direkt in Satz 5.2 enthalten.

Als Gegenbeispiel für die Vertauschbarkeit betrachte man die Funktion $f(x) = x^n$. Hier gilt

$$\lim_{x \to 1} \lim_{n \to \infty} x^n = 0$$
$$\lim_{n \to \infty} \lim_{x \to 1} x^n = 1$$

denn die Konvergenz ist nicht gleichmäßig.

## 5.2 Beweistechnik für punktweise und gleichmäßige Konvergenz

Aus der Übung mit Rony Bergmann.

Bei der punktweisen Konvergenz konvergiert der allgemeine Grenzwert (also der Grenzwert für alle $x$) gegen eine Funktion $s$.

$$\forall x \forall \varepsilon > 0 \exists N \forall n \geq N : |s_n(x) - s(x)| < \varepsilon$$

Bei der gleichmäßigen Konvergenz kann die ganze Funktionenfolge ab einem $n$ in einen $\varepsilon$-Schlauch gepresst werden. Dieser Schlauch muss unabhängig von $x$ beliebig klein werden können. (Dies schließt durchaus keine Sprünge in der Grenzwertfunktion aus.)

Teste zunächst, ob für alle $x \in I$ ein

$$\lim_{n \to \infty} s_n(x) \in \mathbb{R}$$

existiert.

Nein: Die Funktionenfolge ist nicht punktweise konvergent und damit auch nicht gleichmäßig konvergent.

Ja: Die Funktion ist punktweise konvergent. Prüfe nun auf gleichmäßige Konvergenz:

Stelle eine gute Vermutung an, ob die Funktionenfolge konvergiert.

Konvergiert die Funktionenfolge gleichmäßig, so existiert eine nicht von $x$ abhängige Nullfolge $(a_n)_{n\in\mathbb{N}}$, die den $\varepsilon$-Schlauch erzeugen kann. Es gilt

$$\forall x \in I : |s_n(x) - s(x)| < a_n$$

Konvergiert die Funktionenfolge nicht gleichmäßig, dann existiert eine von $x$ abhängige Folge $(x_n)_{n\in\mathbb{N}}$, die als Argument für $s_n$ dafür sorgt, dass der Abstand von $s_n(x_n)$ und $s(x_n)$ nicht 0 wird. Es bleibt also immer ein Rest, der verhindert, dass der $\varepsilon$-Schlauch beliebig klein werden kann. Es gilt

$$(s_n(x_n) - s(x_n))_{n\in\mathbb{N}}$$

konvergiert nicht gegen 0. (Es liegt also Konvergenz gegen $u \neq 0$ oder Divergenz vor.)

## 5.3 Differentiation und Integration

### Satz 5.3

Sei $f_k \in C([a,b]) \forall k \in \mathbb{N}_0$ und

$$\sum_{k=0}^{\infty} f_k(x)$$

konvergiere gleichmäßige in $[a,b]$. Dann gilt

$$\int_a^b \sum_{k=0}^{\infty} f_k(x)\,\mathrm{d}x = \sum_{k=0}^{\infty} \int_a^b f_k(x)\,\mathrm{d}x$$

**Beweis**

1. Nach Satz 5.2 ist

$$\sum_{k=0}^{\infty} f_k(x) \in C([a,b])$$

d. h. alle Integrale existieren, da alle Funktionen stetig sind.

2. $\forall n \in \mathbb{N}_0$ gilt

$$\begin{aligned}\int_a^b \sum_{k=0}^{\infty} f_k(x)\,\mathrm{d}x &= \int_a^b \sum_{k=0}^{n} f_k(x)\,\mathrm{d}x + \int_a^b \sum_{k=n+1}^{\infty} f_k(x)\,\mathrm{d}x \\ &= \sum_{k=0}^{n} \int_a^b f_k(x)\,\mathrm{d}x + \int_a^b \sum_{k=n+1}^{\infty} f_k(x)\,\mathrm{d}x \quad \forall n \in \mathbb{N}\end{aligned}$$

   Wir sind fertig, falls

$$\lim_{n\to\infty} \int_a^b \sum_{k=n+1}^{\infty} f_k(x)\,\mathrm{d}x = 0$$

3. Wir nutzen die gleichmäßige Konvergenz

$$\forall \varepsilon > 0 \exists N(\varepsilon) \forall x \in [a,b] \forall n \geq N \left| \sum_{k=n+1}^{\infty} f_k(x) \right| < \varepsilon$$

   Damit folgt

$$\int_a^b \left| \sum_{k=0}^{\infty} f_k(x) \right| \mathrm{d}x \leq (b-a)\varepsilon \forall n \geq N$$

   Damit ist die offene Aussage am Ende von von 2. gezeigt.

□

**Beispiel**

$$\begin{aligned}\int_a^b \sin x\,\mathrm{d}x &= \int_a^b \underbrace{\sum_{k=0}^{\infty} \frac{x^{2k+1}}{(2k+1)!}(-1)^k}_{\text{gleichmäßig konvergent für alle } x \in \mathbb{R} \text{ (Beweis später)}} \mathrm{d}x \\ &= \sum_{k=0}^{\infty} \int_a^b \frac{x^{2k+1}}{(2k+1)!}(-1)^k\,\mathrm{d}x \\ &= \sum_{k=0}^{\infty} \frac{x^{2k+2}}{(2k+2)!}(-1)^{k+1}\Big|_a^b \\ &= \sum_{k=0}^{\infty} -\frac{x^{2k+2}}{(2k+2)!}(-1)^k\Big|_a^b \\ &= \cos a - \cos b\end{aligned}$$

### Satz 5.4

Sei $f_k' \in C([a,b]) \forall k \in \mathbb{N}_0$,

$$\sum_{k=0}^{\infty} f_k(a)$$

konvergiere und

$$\sum_{k=0}^{\infty} f_k'(x)$$

konvergiere gleichmäßig in $[a,b]$.

Dann gilt

$$\sum_{k=0}^{\infty} f_k(x)$$

ist für alle $x \in [a,b]$ differenzierbar und

$$\frac{\mathrm{d}}{\mathrm{d}x} \sum_{k=0}^{\infty} f_k(x) = \sum_{k=0}^{\infty} \frac{\mathrm{d}}{\mathrm{d}x} f_k(x)$$

**Bemerkung** Die Vorraussetzung ist komplizierter als in Satz 5.3.

### Beispiel

$$\sum_{k=1}^{\infty} \frac{\sin k^2 x}{k^2}$$

ist gleichmäßig konvergent und

$$\frac{\sin k^2 x}{k^2} \in C([a,b])$$

$$\sum_{k=0}^{\infty} \frac{\mathrm{d}}{\mathrm{d}x} \frac{\sin k^2 x}{k^2} = \sum_{k=1}^{\infty} \cos k^2 x$$

ist nicht konvergent, zumindestens für $x = 0$.

**Beweis** Sei

$$s*(x) = \sum_{k=0}^{\infty} f_k'(x),$$

dann gilt nach Satz 5.2 $s* \in C([a,b])$. Also gilt nach Satz 5.3

$$\begin{aligned}\int_a^x s*(t)\,\mathrm{d}t &= \sum_{k=0}^{\infty}\int_a^x f_k'(t)\,\mathrm{d}t\\ &= \sum_{k=0}^{\infty}(f_k(x)-f_k(a)) =: s(x)-s(a)\end{aligned}$$

Mit dem Hauptsatz der Integralrechnung ergibt sich

$$s*(x) = s'(x)$$

Also wissen wir, dass die Summe der Ableitungen der Ableitung der Summe entspricht.

□

## 5.4 Potenzreihen

### Definition 5.5

Seien $x_0 \in \mathbb{R}$ (oder $\mathbb{C}$) und $a_k \in \mathbb{R}$ (oder $\mathbb{C}$), $k \in \mathbb{N}_0$

$$\sum_{k=0}^{\infty} a_k(x-x_0)^k$$

heißt *Potenzreihe* mit Enwicklungspunkt $x_0$.

### Beispiel

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

Nur für den Fall, dass es noch keiner gemerkt hat: An dieser Stelle fügt sich alles ganz wunderbar zusammen, da alle bekannten Funktionen als Potenzreihen sauber definiert werden können. Im Folgenden werden wir jetzt die Frage der Konvergenz für Potenzreihen untersuchen.

Jede Potenzreihe konvergiert im Entwicklungspunkt $x_0$. Wie sieht der Konvergenzbereich in $\mathbb{R}$ (oder $\mathbb{C}$) aus?

### Satz 5.6

a) $\sum_{k=0}^{\infty} a_k(x-x_0)^k$ konvergiere für $x = x_1 \neq x_0$. Dann folgt damit, dass die Konvergenzreihe absolut konvergiert für alle $x$ mit $|x-x_0| < |x_1-x_0|$

b) $\sum_{k=0}^{\infty} a_k(x-x_0)^k$ divergiere für $x = x_2$. Dann folgt damit, dass die Konvergenzreihe divergiert für alle $x$ mit $|x-x_0| > |x_1-x_0|$

### Geometrische Veranschaulichung

Wenn es bei $x_1$ auf dem Zahlenstrahl konvergiert, dann konvergiert es auch für alle $x$, die dichter dran sind an $x_0$, als $x_1$. Divergiert es für $x_2$, dann divergiert es auch für alle $x$, die weiter weg von $x_0$ sind als $x_2$.

In der Gaußschen Zahlenebene funktioniert das ganze genauso mit dem euklidischen Abstand. Es entstehen also Kreise in der Ebene statt Bereiche auf der Zahlenebene.

**Vorlesung CVIII**

**Beweis** a) Aus der Konvergenz der Reihe

$$\sum_{k=0}^{\infty} a_k(x_1-x_0)^k$$

folgt

$$\exists C > 0 : \forall k : |a_k(x_1-x_0)^k| < C$$

Damit gilt

$$|x-x_0| = q|x_1-x_0| \text{ mit } 0 \leq q < 1$$

Also ergibt sich

$$|a_k(x-x_0)^k| = |a_k||x_1-x_0|^k q^k < Cq^k$$

und wir erhalten die konvergente Majorante

$$C\sum_{k=0}^{\infty} q^k$$

b) Wir führen einen indirekten Beweis und nehmen zum Zwecke des Widerspruchs an, die Reihe

$$\sum_{k=0}^{\infty} a_k(x-x_0)^k$$

sei konvergent. Dann ergibt sich mit a) die Konvergenz für $x_2$ und wir erhalten einen Widerspruch zur Annahme.

$\square$

Zusammenfassung des Satzes: Es gibt ein $r$ (wobei $r = \infty$ möglich) mit

$$\begin{gathered}\forall x \text{ mit } |x - x_0| < r \text{ gilt Konvergenz (absolut!)}\\ \forall x \text{ mit } |x - x_0| > r \text{ gilt Divergenz}\\ \forall x \text{ mit } |x - x_0| = r \text{ keine Aussage}\end{gathered}$$

Dabei ist $r$ der Konvergenzradius.

### Beispiel

Zur Vereinfachung setzen wir $x_0 := 0$.

a) Für die Folg

$$\sum_{k=0}^{\infty} x^k, \sum_{k=0}^{\infty} \frac{x^k}{k}, \sum_{k=0}^{\infty} \frac{x^k}{k^2}$$

gilt $r = 1$. Auf dem Rand liegt immer Divergenz vor, da die Folge $x^1$ keine Nullfolge ist.

Für die Folge

$$\sum_{k=0}^{\infty} x^k, \sum_{k=0}^{\infty} \frac{x^k}{k}, \sum_{k=0}^{\infty} \frac{x^k}{k^2}$$

gilt $r = 1$. Die harmonische und die alternierende harmonische Reihe (relle Zahlen -1, 1) sind bekannt. Für komplexe Zahlen ist der Rand verschieden.

Für die Folge

$$\sum_{k=0}^{\infty} x^k, \sum_{k=0}^{\infty} \frac{x^k}{k}, \sum_{k=0}^{\infty} \frac{x^k}{k^2}$$

gilt $r = 1$. Auf dem Rand liegt immer Konvergenz vor, da die unendliche Summe von $1/k^2$ immer konvergent ist.

b) Wir betrachten die Folge

$$\sum_{k=0}^{\infty} \frac{x^k}{k!} =: e^x.$$

Die Folge konvergiert für alle $x \in \mathbb{C}$. Also gilt $r = \infty$.

c) Die Folge

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

ist immer divergent für $x \neq 0$. Also gilt $r = 0$.

Nachdem wir jetzt wissen, dass uns der Konvergenzradius schlaue Dinge liefern kann, bleibt noch eine wesentliche Frage: Wie erhalten wir den Konvergenzradius?

**Satz 5.7** Cauchy(1789–1857)-Hadamard(1865–1963)

Gegeben sei eine Potenzreihe mit Koeffizientem $a_k$. Wir definieren

$$\limsup_{k\to\infty} \sqrt[k]{|a_k|} =: A$$

Der Konvergenzradius von

$$\sum_{k=0}^{\infty} a_k(x-x_0)^k$$

ist

$$r = \begin{cases} \infty & \text{falls } A = 0 \\ \frac{1}{A} & \text{falls } 0 < A < \infty \\ 0 & \text{falls } A = \infty \end{cases}$$

**Erinnerung: Wurzelkriterium für Zahlenreihen (Satz 1.26 b))**

$$\sum_{k=0}^{\infty} a_k(x-x_0)^k$$

konvergiert, falls

$$\limsup_{k\to\infty} \sqrt[k]{|a_k||x-x_0|^k} < 1$$

und divergiert, falls

$$\limsup_{k\to\infty} \sqrt[k]{|a_k||x-x_0|^k} > 1$$

**Beweis**

$$\begin{aligned} \limsup_{k\to\infty} \sqrt[k]{|a_k||x-x_0|} &= \limsup_{k\to\infty} \sqrt[k]{|a_k|}|x-x_0| \\ &= A|x-x_0| \\ \Rightarrow r &= \frac{1}{A}, r = \infty (A=0), r = 0 (A=\infty) \end{aligned}$$

$\square$

**Bemerkung** Es gilt auch

$$\exists \lim_{k\to\infty} \left|\frac{a_k}{a_{k+1}}\right| = r \Rightarrow \sum_{k=0}^{\infty} a_k(x-x_0)^k$$

hat Konvergenzradius $r$.

**Beispiel**

$$\sum_{k=0}^{\infty} k(x+1)^k$$
$$\limsup_{k\to\infty} \sqrt[k]{|k|} = 1 = A$$
$$\Rightarrow r = 1$$

**Beispiel**

$$\sum_{k=0}^{\infty} 2^k x^{2k} = \sum_{k=0}^{\infty} 2^k (x^2)^k = \sum_{k=0}^{\infty} 2^k y^k$$

konvergiert genau dann, wenn $|2y| < 1$, also r $= 1/2$

$$a_k = \begin{cases} 2^{k/2} & \text{falls } k \text{ gerade} \\ 0 & \text{falls } k \text{ ungerade} \end{cases}$$
$$\limsup_{k\to\infty} \sqrt[k]{|a_k|} = \sqrt{2} = A$$
$$r = \frac{1}{\sqrt{2}} = \frac{\sqrt{2}}{2}$$

### 5.4.1 Rechnen mit Potenzreihen

Wir betrachten ohne Beschränkung der Allgemeinheit Potenzreihen mit $x_0 = 0$.

$$\sum_{k=0}^{\infty} a_k x^k \text{ habe Konvergenzradius } r > 0$$

Wir definieren für $|x| < r$ die Funktion $f$ mit

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

### Lemma 5.8

Für alle $b$ mit $0 < b < r$ konvergiert

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

gleichmäßig in $|x| \leq b$.

**Beweis** Sei $\varepsilon > 0$ gegeben. Wir wählen dann $N$ so, dass

$$\sum_{k=N}^{\infty} |a_k b^k| < \varepsilon.$$

Damit ergibt sich

$$\forall |x| \leq b : \left| \sum_{k=N}^{\infty} a_k x^n \right| \leq \sum_{k=N}^{\infty} |a_k b^k| < \varepsilon$$

□

### Satz 5.9

Die Funktion

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

habe einen Konvergenzradius $r > 0$ (also |x| < r). Dann gilt

1. $f : (-r, r) \to \mathbb{R}(\mathbb{C})$ ist stetig.
2. Für $-r < x < r$ gilt

$$\begin{aligned} \int_0^x f(t)\,\mathrm{d}t &= \int_0^x \sum_{k=0}^{\infty} a_k t^k \,\mathrm{d}t \\ &= \sum_{k=0}^{\infty} a_k \int t^k \,\mathrm{d}t \\ &= \sum_{k=0}^{\infty} \frac{a_k}{k+1} x^{k+1} \\ &= \sum_{k=1}^{\infty} \frac{a_{k+1}}{k} x^k \end{aligned}$$

Wir erhalten wieder eine Potenzreihe mit Potenzradius $r$.

3. $f$ ist beliebig differenzierbar in $(-r,r)$ mit

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

Rekursiv ergibt sich für alle $p \in \mathbb{N}$

$$f^{(p)} = \sum_{k=p}^{\infty} a_k x^{k-p} \frac{k!}{(k-p)!} = \sum_{k=0}^{\infty} a_{k+p} \frac{(k+p)!}{k!} x^k$$

**Beweis** 1. Lemma 5.8 und Satz 5.2

2. Lemma 5.8 und Satz 5.3

3.
$$\sum_k a_k k x^{k-1}$$
hat den Konvergenzradius $r$. Nach Lemma 5.8 gilt für alle $\varepsilon > 0$
$$\sum_{k=1}^{\infty} a_k k x^{k-1}$$
gleichmäßig konvergent in $|x| < r - \varepsilon$.

Außerdem ist
$$\sum_{k=0}^{\infty} a_k x^k$$
konvergent. Damit liefert Satz 5.4 die Behauptung.

$\square$

**Vorlesung CIX**

**Bemerkung** • **Abelsche Stetigkeit** : Konvergiert
$$\sum_{k=0}^{\infty} a_k r^k,$$
dann konvergiert
$$\sum_{k=0}^{\infty} a_k x^k$$
gleichmäßig in $[0,r]$. Analoges gilt für $x = -r$.

- **Weierstraß'sches Maiorantenkriterium** : Gegeben eine Funktion $f_a : D \to \mathbb{R}(\mathbb{C})$. Existiert eine von $x \in D$ unabhängige Folge $(c_n)$, so dass

$$\forall n \geq N \forall x \in D |f_n(x)| \leq c_n$$

und

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

ist konvergent. Dann ist

$$\sum_{k=0}^{\infty} f_n(x)$$

gleichmäßg konvergent gegen $s : D \to \mathbb{R}(\mathbb{C})$.

### Beispiel

$$\frac{1}{1-x} = \sum_{k=0}^{\infty} x^k \quad \text{für } -1 < x < 1$$

$$\begin{aligned} -\ln(1-x) &= \int_0^x \frac{1}{1-t} \, \mathrm{d}t \\ &= \sum_{k=0}^{\infty} \int_0^x t^k \, \mathrm{d}t \\ &= \sum_{k=0}^{\infty} \frac{x^{k+1}}{k+1} \\ &= \sum_{k=1}^{\infty} \frac{x^k}{k} \qquad \text{für } -1 < x < 1 \end{aligned}$$

Für $x = -1$ ergibt sich

$$\begin{aligned} -\ln 2 &= \sum_{k=0}^{\infty} \frac{(-1)^k}{k} \\ &= 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} - \frac{1}{6} + \dots \end{aligned}$$

Ersetzen wir $x$ durch $-x$, so ergibt sich

$$-\ln(1+x) = \sum_{k=0}^{\infty} \frac{x^{k+1}}{k} x^k \quad \text{für } -1 < x < 1$$

**Beispiel**

$$
\begin{aligned}
\frac{1}{1+x^2} &= \sum_{k=0}^{\infty} (-1)^k x^{2k} \qquad \text{für } |x| < 1 \\
\arctan x &= \int_0^x \frac{\mathrm{d}t}{1+t^2} \\
&= \sum_{k=0}^{\infty} \frac{(-1)^k}{2k-1} x^{2k+1} \qquad \text{für } |x| < 1
\end{aligned}
$$

$$
\begin{aligned}
\arctan 1 = \frac{\pi}{4} &= \sum_{k=0}^{\infty} \frac{(-1)^k}{2k+1} \\
&= 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \ldots
\end{aligned}
$$

**Beispiel**

$$
\begin{aligned}
e^{-x^2} &= \sum_{k=0}^{\infty} \frac{1}{k!} (-x^2)^k \\
&= \sum_{k=0}^{\infty} \frac{1}{k!} (-1)^k x^{2k}
\end{aligned}
$$

**Beispiel**

$$
\begin{aligned}
\operatorname{erf}(x) &= \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2}\,\mathrm{d}t \\
&= \frac{2}{\pi} \sum_{k=0}^{\infty} \frac{(-1)^k}{k!(2k+1)} x^{2k+1} \qquad \forall x \in \mathbb{C} \\
\lim_{x\to\infty} \operatorname{erf}(x) &= 1
\end{aligned}
$$

**Beispiel**

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

**Beispiel** Integralsinus

$$\begin{aligned} \mathrm{Si}(x) &= \int_0^x \frac{\sin t}{t} \,\mathrm{d}t \\ &= \sum_{k=0}^{\infty} \frac{(-1)^k x^{2k+1}}{(2k+1)!(2k+1)} \end{aligned}$$

### Satz 5.10

Gegeben seien die Funktion

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

mit Konvergenzradius $r_1 > 0$ und die Funktion

$$g(x) = \sum_{k=0}^{\infty} b_k x^k$$

mit Konvergenzradius $r_2 > 0$. Dann gilt für alle $x$ mit $|x| < \min(r_1, r_2)$

$$\alpha f(x) + \beta g(x) = \sum_{k=0}^{\infty} (\alpha a_k + \beta b_k) x^k$$

Insbesondere gilt

$$\begin{aligned} f(x)g(x) &= \sum_{n=0}^{\infty} x^n \sum_{n=0}^{\infty} a_k b_{n-k} \\ &= \sum_{n=0}^{\infty} x^n (a_0 b_n + a_1 b_{n-1} + a_2 b_{n-2} + \ldots + a_n b_0) \end{aligned}$$

**Beweis** nach Kapitel 2 $\square$

**Beispiel**

$$\begin{aligned} e^x &= \sum_{k=0}^{\infty} \frac{x^k}{k!} \forall x \in \mathbb{C} \\ \frac{1}{1-x} &= \sum_{k=0}^{\infty} x^k \forall |x| < 1 \\ \frac{e^x}{1-x} &= \sum_{k=0}^{\infty} \left( \sum_{k=0}^{n} \frac{1}{k!} \cdot 1 \right) x^k \forall |x| < 1 \end{aligned}$$

**Satz 5.11** Identitätssatz für Potenzreihen

Gegeben seien die Potenzreihen

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

und

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

mit positiven Konvergenzradien $r_1$ und $r_2$. Weiter sei ein $\varepsilon$ derart gegeben, dass $0 < \varepsilon < \min(r_1, r_2)$ mit

$$\sum_{k=0}^{\infty} a_k x^k = \sum_{k=0}^{\infty} b_k x^k \forall |x| < \varepsilon.$$

Dann gilt für alle $k \in \mathbb{N}_0$

$$a_k = b_k.$$

**Beweis** Für $x = 0$ ergibt sich $a_0 = b_0$. Durch Subtraktion von $a_0, b_0$ und Division durch $x$ ergibt sich

$$\sum_{k=1}^{\infty} a_k x^{k-1} = \sum_{k=1}^{\infty} b_k x^{k-1} \forall |x| < \varepsilon, x \neq 0.$$

Jetzt ergibt sich für $x = 0$ die Aussage $a_1 = b_1$. Verfahre rekursiv weiter ... □

**Bemerkung**

1. Der Identitätssatz für Polynome heißt auch Methode des Koeffizientenvergleichs für Polynome.
2. Die Rückrichtung ist trivial.
3. Eine Funktion kann *bei festem* $x_0$ auf höchstens eine Weise in eine Potenzreihe entwickelt werden. Aber allgemein gilt zum Beispiel

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

4. Der Beweisschritt der Division durch $x$ kann durch die Differntiation ersetzt werden.

   Sei dazu

$$\begin{aligned} f(x) &= \sum_{k=0}^{\infty} a_k (x-x_0)^k \\ a_k &= f(x_0) \\ a_1 &= f'(x_0) \\ s_2 &= \frac{f''(x_0)}{2!} \\ a_k &= \frac{f^{(k)}(x_0)}{k!} \end{aligned}$$

   Dass heißt, jede Potenzmenge ist seine eigene Taylor-Reihe

$$f(x) = \sum_{k=0}^{\infty} \frac{f^{(k)}(x_0)}{k!}(x-x_0)^k$$

5. $I$ sei ein offenes Intervall auf $\mathbb{R}$. Dann heißt die Funktion $f : i \to \mathbb{R}$ in $I$ *reell analytisch*, falls $f$ in jedem Punkt von $I$ in eine Potenzreihe mit positivem Konvergenzradius entwickelbar ist.

## 5.5 Fourier-Reihen

Wärmeausbreitung in endlichem Staal. Frz. Zeitung 19. Jahrhundert; Wettbewerb.

Signalanalyse von akustischen und optischen Wellen.

**Definition** Periodische Funktionen

Sei $T > 0$, $f : \mathbb{R} \supset D \to \mathbb{C}$ heißt $T - Periodisch$, falls

$$\forall x \in D f(x \pm T) = f(x)$$

**Beispiel**

$\sin x, \cos x$ sind $2\pi$-periodisch

Die Peridizität ergibt sich sofort, wenn man die trigonometrischen Funktionen geometrisch erklärt. Aber wir wissen ja inzwischen, dass es viel toller ist, alles Funktionen als Reihen zu definieren. Also tun wir das:

$$
\begin{aligned}
e^{ix} &= \sum_{k=0}^{\infty} \frac{1}{k!}(ix)^k \quad \forall x \in \mathbb{C} \\
&= \underbrace{\sum_{k=0}^{\infty} \frac{1}{(2k)!}(-1)^k x^{2k}}_{\cos x} + i \underbrace{\sum_{k=0} \frac{1}{(2k+1)!}(-1)^k x^{2k+1}}_{\sin x} \\
&= \cos x + i \sin x
\end{aligned}
$$

Nach dem Leibnitz-Kriterium ergibt sich folgende Fehlerabschätzung

$$
\begin{aligned}
\cos 2 &\leq 1 - \frac{2^2}{2!} + \frac{2^4}{4!} = -\frac{1}{3} \\
\sin x &\geq x - \frac{x^3}{3!} \geq x(1 - \frac{x^2}{3!}) \geq x(1 - \frac{2^2}{3!}) = \frac{x}{3} \forall x \in [0,2]
\end{aligned}
$$

$$
\frac{\mathrm{d}}{\mathrm{d}x} \cos x = -\sin x < 0 \quad \forall 0 \leq x \leq 2
$$

Also ist auf dem betrachteten Intervall $\cos x$ steng monoton fallend.

$$
\begin{aligned}
\cos 2 &\leq -\frac{1}{3} \\
\cos 0 &= 1
\end{aligned}
$$

Mit dem Zwischenwertsatz erhalten wir: Es existiert *genau eine* Nullstelle. Wir nennen diese $\pi/2$ und haben damit $\pi$ rein analytisch definiert.

$$
\begin{aligned}
e^{i\pi/2} &= i \sin \pi/2 \\
e^{-i\pi/2} &= -i \sin \pi/2 \\
e^{i\pi/2 + (-i\pi/2)} = 1 &= \sin^2 \pi/2 \\
&\Rightarrow \sin \pi/2 = 1 (\text{da } \sin \pi/2 > 0)
\end{aligned}
$$

Mit dieser Definition von $\pi$ erhalten wir nun

$$
e^{i\pi/2} = i
$$

$$
\boxed{e^{i\pi} = -1}
$$

Diese Formel ist nicht aufgrund ihrer besonderen Bedeutung, sondern aufgrund ihrer besonderen Esthetik hervorgehoben: Alle zentralen Symbole der Analysis in einer Formel verein: $\pi, e, i, -, 1, =$.

$$
e^{2\pi i} = 1
$$

$$\forall x \in \mathbb{C} : e^{i(x+\pi/2)} = e^{ix}$$

Also ist $e^{i\circ}$ ist $2\pi$-periodisch.

$f(x) = \text{const}$ ist $T$-periodisch für alle $T > 0$. Es gibt also keine kleinste Periode.

Diriklet-Funktion

$$\Xi_{\mathbb{Q}}(x) = \begin{cases} 1 & \text{wenn } x \in \mathbb{Q} \\ 0 & \text{sonst} \end{cases}$$

hat genau jedes rationale $T > 0$ als Periode.

Aus der $T$-Periodizität von $f$ folgt die $nT$-Periodizität von $f$ für alle $n \in \mathbb{N}$.

Sinf $f$ und $g$ $T$-periodisch, dann ist die Funktion $\alpha f + \beta g$ $T$-periodisch.

Es sei eine $T$-periodische Funktion $f$ auf $\mathbb{R}$ gegeben. Dann gilt

$$\begin{aligned} \int_a^{a+T} f(x)\,\mathrm{d}x &= \int_0^T f(x)\,\mathrm{d}x \\ &= \int_{-T/2}^{T/2} f(x)\,\mathrm{d}x \end{aligned}$$

Ist $f$ $T$-periodisch, dann gilt

$$F(x) = f\left(\frac{xT}{2\pi}\right)$$

ist $2\pi$-periodisch.

$F$ ist keine Stammfunktion. Aus irgendwelchen Gründen sind uns aber schon wieder die Buchstaben ausgegangen.

Wir können also jetzt nur noch $2\pi$-periodische Funktionen betrachten und können die Ergebnisse durch diese einfache Transformationen auf $T$-periodische Funktionen übertragen.

**Beweis** Es gilt für alle $x \in \mathbb{C}$

$$\begin{aligned} F(x+2\pi) &= f\left(\frac{(x+2\pi)T}{2\pi}\right) \\ &= f\left(\frac{xT}{2\pi} + T\right) \\ &= f\left(\frac{xT}{2\pi}\right) = F(x) \end{aligned}$$

□

### 5.5.1 »Periodische Fortsetzung«

Gegeben sei eine Funktion $f : [0,T] \to \mathbb{C}$. Diese kann fortgestzt werden

- $T$-periodisch (nur für $[0,T) \to \mathbb{C}$) [Skizze von Fleischerhaken immer wieder]
- gerade $2T$-periodisch

  Eine Funktion ist gerade, falls $f(x) = f(-x)$

  [Skizze von Möwen immer wieder]
- ungerade $2T$-periodisch (nur für $f : (0,T) \to \mathbb{C}$) und $f(0) = f(T) = f(-T) = 0$

  Eine Funktion ist ungerade, falls $f(x) = -f(-x)$

  [Skizze von Fleischerhaken am Nullpunkt gespiegelt immer wieder.] Meist wird $f$ und seine periodische Fortsetzung gleich bezeichnet.

### 5.5.2 Trigonometrische Polynome

Grundbausteine der Fourier-Analyse

Polynom $\sum_{k=0}^{n} c_k x^k$ liefert die Potenzreihe $\sum_{k=0}^{\infty} c_k x^k$ und die Taylor-Reihe $\sum_{k=0}^{\infty} \frac{f^{(k)}}{k!} x^k$

Trigonometrische Polynome sind Polynome in $z = e^{ix}$. Dabei gilt für alle $x \in \mathbb{R} : e^{ix} = \cos x + i \sin x$ und bildet den Einheitskreis.

Für alle $x$ gilt

$$
\begin{aligned}
T_n(x) &= \sum_{k=-n}^{n} c_k e^{ikx} \\
&= c_0 + c_1 e^{ix} + c_{-1} e^{-ix} + \ldots + c_n e^{inx} + c_{-n} e^{-inx} \\
&= c_0 + (c_1 + c_{-1}) \cos x + i(c_1 - c_{-1}) \cos x + \ldots + (c_n + c_{-n}) \cos x + i(c_n + c_{-n}) i \sin x \\
\\
&= \frac{a_0}{2} + \sum_{k=1}^{n} (a_k \cos kx + b_k \sin kx) \\
&= \sum_{k=-n}^{n} c_k e^{ikx}
\end{aligned}
$$

mit den Ersetzungsformelm von Euler-Moivre

$$
\begin{gathered}
a_k = c_k + c_{-k} \\
b_k = i(c_k - c_{-k}) \\
(a_0 = 2c_0)
\end{gathered}
$$

für alle $k = 1, 2, 3, \ldots$

Umgekehrt ergibt sich

$$c_k = \frac{1}{2}(a_k - ib_k)$$
$$c_{-k} = \frac{1}{2}(a_k + ib_k)$$
$$c_0 = \frac{a_0}{2}$$

für alle $k = 1, 2, 3, \ldots$

### Satz 5.12

Sei $t_n$ ein trigonometrische Polynom vom Grad $n$, dann gilt

a) $T_n$ ist $2\pi$-periodisch

b) $T_n$ hat in $[0, 2\pi)$ höchstens $2n$ Nullstellen

c) $\forall x \in \mathbb{C} \quad T_n(x) = 0 \iff \forall k \in \{-n, \ldots, n\} \quad c_k = 0 \iff \forall k \in \{1, \ldots, n\} \quad a_k = b_k = 0$ und $a_0 = 0$

d) $T_n$ ist genau dann reell (d. h. $T_n(x) \in \mathbb{R} \forall x \in \mathbb{R}$), wenn $c_k = \overline{-c_{-k}}$ für alle $-n \leq k \leq n$. (Was genau dann der Fall ist, wenn alle $a_k$ und $b_k$ reell sind.)

e) Euler-Moivre: Es gilt für alle $k \in \mathbb{Z}$

$$c_k = \frac{1}{2\pi} \int_0^{2\pi} T_n(x) e^{ikx} \,\mathrm{d}x$$
$$a_k = \frac{1}{\pi} \int_0^{2\pi} T_n(x) \cos kx \,\mathrm{d}x$$
$$b_k = \frac{1}{\pi} \int_0^{2\pi} T_n(x) \sin kx \,\mathrm{d}x$$

Betrachten wir das Polynom

$$\sum_{k=0}^{n} c_k x^k p_n(x),$$

dann gilt

$$c_k = \frac{p_n^{(k)}(0)}{k!}$$

$$e^{ix} = \cos x + i \sin x$$
$$T_n(x) = \sum_{k=0}^{n} c_k e^{ikx}$$
$$T_n(x) = \frac{a_0}{2} + \sum_{k=0}^{n} a_k \cos kx + b_k \sin kx$$

**Beweis** 1. ist trivial

2. Wir schreiben das Polynom ein wenig um und definieren ein $z$, sodass gilt:

$$T_n(x) = \underbrace{\sum_{k=0}^{n} c_k e^{ikx}}_{x\in[0,2\pi]} = \underbrace{\sum_{k=-n}^{n} c_k z^k}_{z \text{ auf dem Einheitskreis in } \mathbb{C}}$$

Der Übergang von $x \in [0, 2\pi)$ zu $z$ auf dem Einheitskreis in $\mathbb{C}$ ist eine Bijektion.

Also fragen wir nach den Nullstellen von

$$z^{-n} \sum_{k=0}^{2n} c_{k-n} z^k$$

auf dem Einheitskreis.

Nach dem Fundamentalsatz der Algebra (»Ein Polynom $n$-ten Gerades hat höchstens genau $n$ Nullstellen.«) hat das Polynom höchstens $2n$ Nullstellen.

3. Hat $T_n$ mehr als $2n$ Nullstellen auf dem Einheitskreis, dann gilt nach b): Das Polynom ist identisch 0, also ist $c_k = 0$ für alle $k \in \mathbb{Z}$.

4. $T_n$ ist reellwertig genau dann, wenn alle $a_k, b_k \in \mathbb{R}$.

   Nach dem Umrechungsformeln von Euler-Moivre gilt

$$a_k = c_k + c_k$$

   Also sind genau dann alle $a_k$ reell, wenn gilt

$$\forall k = 0, \dots, n \quad \operatorname{Im} c_k = -\operatorname{Im} c_{-k}$$

   Weiter gilt nach Euler-Moivre

$$b_k = i(c_k - c_{-k})$$

Also sind genau dann alle $b_k$ reell, wenn gilt

$$\forall k = 0, \ldots, n \quad \operatorname{Re} c_k = \operatorname{Re} c_{-k}$$

5. Es gilt für alle $l = -n, \ldots, n$

$$\begin{aligned}
\int_0^{2\pi} T_n(x) e^{-ilx}\,\mathrm{d}x &= \sum_{k=-n}^{n} c_k \int_0^{2\pi} e^{i(k-l)x}\,\mathrm{d}x \\
&= \sum_{k=-n}^{n} c_k \begin{cases} 2\pi & \text{für } k = l \\ \frac{e^{i(k-l)x}}{i(k-l)}\Big|_{k=0}^{k=2\pi} & \text{für } k \neq l \end{cases} \\
&= 2\pi c_l
\end{aligned}$$

$$T_n(x) = \sum_{k=-n}^{n} \underbrace{\frac{1}{2\pi} \int_0^{2\pi} T_n(x) e^{-ikx}\,\mathrm{d}x}_{c_k} e^{ikx}$$

Im reellen ergibt sich für alle $l = 0, \ldots, n$

$$\begin{aligned}
\int_0^{2\pi} T_n(x) \cos lx\,\mathrm{d}x =& \frac{a_0}{2} \int_0^{2\pi} \cos lx\,\mathrm{d}x \\
&+ \sum_{k=0}^{n} a_k \int_0^{2\pi} \cos kx \cos lx\,\mathrm{d}x \\
&+ \sum_{k=1}^{n} b_k \underbrace{\int_0^{2\pi} \sin kx \cos lx\,\mathrm{d}x}_{\int_{-\pi}^{\pi} \text{ungerade} \cdot \text{gerade} = 0} \\
&= \begin{cases} \pi a_0 & \text{für } l = 0 \\ \pi a_l & \text{für } l > 0 \end{cases} \\
&= \pi a_l
\end{aligned}$$

$a_0$ für $l = 0$ ergibt sich aus dem ersten Integral. $a_l$ für $l > 0$ ergibt sich aus dem zweiten Integral für den Fall $k \neq l$, denn es gilt

$$\cos kx \cos lx = \frac{\cos(k+l) + \cos(k-l)x}{2}.$$

$$
\begin{aligned}
T_n(x) = & \frac{1}{\pi}\int_0^{2\pi} T_n(t)\,\mathrm{d}t \\
& + \sum_{k=1}^{n} \underbrace{\frac{1}{\pi}\int_0^{2\pi} T_n(t)\cos kt\,\mathrm{d}t}_{a_k=a_k(T_n)} \cos kx \\
& + \sum_{k=1}^{n} \frac{1}{\pi}\underbrace{\int_0^{2\pi} T_n(x)\sin kt\,\mathrm{d}t}_{b_k=b_k(T_n)} \sin kx
\end{aligned}
$$

$c_k(T_n), a_k(T_n)$ und $b_k(T_n)$ heißen *Fourier-Koeffizienten*.

□

**Bemerkung** Es gelten die Orthogonalitätsrelationen:

$$
\frac{1}{2\pi}\int_0^{2\pi} e^{ikt}\,\underbrace{\overline{e^{ilt}}}_{e^{-ilt}}
$$

$$
= \delta_{kl} = \begin{cases} 1 & \text{für } k = l \\ 0 & \text{für } k \neq l \end{cases}
$$

$$
\frac{1}{\pi}\int_0^{2\pi} \cos lt \cos kt\,\mathrm{d}t = \begin{cases} 2 & \text{für } l = k = 0 \\ 1 & \text{für } l = k \neq 0 \\ 0 & \text{für } l \neq k \end{cases}
$$

$$
\frac{1}{\pi}\int_0^{\pi} \sin kt + \sin lt\,\mathrm{d}t = \begin{cases} 1 & \text{für } k = l > 0 \\ 0 & \text{sonst} \end{cases}
$$

$$
\frac{1}{\pi}\sin kt \cos lt\,\mathrm{d}t = 0
$$

**Beispiel** »Dirichlet-Kern«

$$
\begin{aligned}
D_n(x) &= \sum_{k=-n}^{n} e^{ikx} = \sum_{k=-n}^{n} (\cos kx + i\sin kx) \\
&= 1 + 2\sum_{k=1}^{n} \cos kx
\end{aligned}
$$

Mit dem immer gegenwärtigen $z := e^{ix}$ ergibt sich

$$\begin{aligned} D_n(x) &= z^{-n}\sum_{k=0}^{2n} z^k \\ &= \begin{cases} \frac{z^{2n+1}-z}{z^n(z-1)} & \text{für } z \neq 1 \\ 2n+1 & \text{für } z = 1 \end{cases} \end{aligned}$$

Im Falle $z \neq 1$ ergibt sich weiter

$$\begin{aligned} D_n(x) &= \frac{z^{n+1/2} - z^{-n-1/2}}{z^{1/2} - z^{-1/2}} \\ &= \frac{e^{i(n+1/2)x} - e^{i(n+1/2)x}}{e^{ix/2} - e^{ix/2}} \\ &= \frac{\sin(n+1/2)x}{\sin x/2} \end{aligned}$$

Wir haben jetzt den Fall $z \neq 1$, also $x \neq 2\pi l$ für alle $l \in \mathbb{Z}$ betrachtet. Im anderen Fall $z = 1$, also $x = 2\pi l$ für alle $l \in \mathbb{Z}$ gilt weiterhin $2n+1$.

**andere Beweisidee**

$$2\sin x/2\cos kx = \sin(k+1/2) - \sin(x-1/2)x$$

Den Dozenten juckt es an dieser Stelle in den Fingern endlich wieder einen Grenzwert zu berachten, also widmen wir uns jetzt dieser Fragestellung.

$$\begin{aligned} \lim_{n\to\infty} D_n(x) &= 1 + \sum_{k=1}^{\infty} \cos kx \\ &= \begin{cases} \infty \text{ (bestimmt diveregnt)} & \text{für } x = 2l\pi \\ \text{unbestimmt divergent} & \text{sonst} \end{cases} \end{aligned}$$

Die ganze Gleichheit ist also irgendwie Mist, da der Grenzwert garnicht existiert. Also sowas besser nicht hinschreiben. (Außer Du bist Dozent einer Analysis-Vorlesung.)

Trigonometrische Reihen sind Reihen der Form

$$
\begin{aligned}
\sum_{k=-\infty}^{\infty} c_k e^{ikx} &:= \lim_{k\to\infty} \sum_{k=-n}^{n} c_k e^{ikx} \\
&= \frac{a_0}{2} + \lim_{n\to\infty} \sum_{k=1}^{n} (a_k \cos kx + b_k \sin kx) \\
&= \frac{a_0}{2} + \sum_{k=1}^{\infty} a_k \cos kx + b_k \sin kx
\end{aligned}
$$

Wir gelangen zu der grundsätzlichen Frage:

Existiert der Grenzwert?

**Beispiel** gutartig

Da $|e^{ix}| = 1$ konvergiert folgende Reihe genau für $|r| < 1$:

$$
\begin{aligned}
\sum_{n=0}^{\infty} \left(re^{ix}\right)^n &= \sum_{n=0}^{\infty} r^n e^{inx} \\
&= \sum_{n=0}^{\infty} r^n \cos nx + i \sum_{n=0}^{\infty} r^n \sin nx \\
&= \frac{1}{1 - re^{ix}} \\
&= \frac{1 - r\cos x + ir\sin x}{(1 - r\cos x)^2 + (r\sin x)^2} + i\frac{r\sin x}{(1-\cos x)^2 + (r\sin x)^2}
\end{aligned}
$$

Also erhalten wir durch diese Umformung

$$
\begin{aligned}
\sum_{n=0}^{\infty} r^n \cos nx &= \frac{1 - r\cos x + ir\sin x}{(1 - r\cos x)^2 + (r\sin x)^2} \\
&= \frac{1 - r\cos x}{1 + r^2 - 2r\cos x}
\end{aligned}
$$

$$
\begin{aligned}
\sum_{n=0}^{\infty} r^n \sin nx &= \frac{r\sin x}{(1-\cos x)^2 + (r\sin x)^2} \\
&= \frac{r\sin x}{1 + r^2 - 2r\cos x}
\end{aligned}
$$

**Beispiel** für obiges Beispiel

Wir betrachten den Fall $r = 1/2$:

$$\sum_{n=0}^{\infty} \frac{\cos nx}{2^n} = \frac{1 - 1/2 \cos x}{5/4 - \cos x} = \frac{4 - 2\cos x}{5 - 4\cos x}$$

**Vorlesung CXI**

$$\sum_{k=1}^{\infty} r^k \sin kx = \frac{r \sin x}{1 - 2r\cos x + r^2}$$

mit $x \in \mathbb{R}$ und $-1 < r < 1$. Die Reihe konvergiert gleichmäßig in $x$.

### Satz 5.13

Die Reihe

$$\sum_{k=-\infty}^{\infty} c_k e^{ikx}$$

konvergiere gleichmäßig auf $[0, 2\pi]$ gegen eine Funktion $f$ in Abhängigkeit von $x$. Dann gilt

a) $f$ ist stetig auf $\mathbb{R}$ und $2\pi$-periodisch. Diese Funktionenklasse bezeichnen mir mit $C_{2\pi}$.

b)
$$c_k = c_k(f) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t) e^{-ikt} \, \mathrm{d}t$$

**Beweis** a) Satz 5.2:
Eine gleichmäßig konvergente Reihe *stetiger* Funktionen konvergiert stets gegen eine stetige Funktion.

b) Nach Satz 5.3 (?) ist die gliedweise Integration erlaubt.

$$f(x) = \sum_{k=-\infty}^{\infty} c_k e^{ikx}$$

$$\frac{1}{2\pi} \int_0^{2\pi} e^{ilx} f(x) = \sum_{k=-\infty}^{\infty} \frac{1}{2\pi} c_k e^{ikx} e^{-ilx} \, \mathrm{d}x$$

$$= \sum_{k=-\infty}^{\infty} c_k \frac{1}{2\pi} \int_0^{2\pi} e^{ix(k-l)} \, \mathrm{d}x = c_l$$

$$\frac{1}{2\pi}\int_0^{2\pi} e^{ix(k-l)}\,\mathrm{d}x = \begin{cases} \left.\frac{e^{ix(k-l)}}{i(k-l)}\right|_0^{2\pi} & \text{falls } k-l\neq 0 \\ 1 & \text{falls } k=l \end{cases}$$

□

**Bemerkung** Theorie der *nicht* gleichmäßig konvergenten trigonometrischen Reihen hat Analysis der letzten 200 Jahre wesentlich geprägt.

[prestin-antritt.pdf / Fourier-Analyse]

### das Fundamentalbeispiel

Im folgenden sei

$$\begin{aligned} s(x) &= \sum_{k=1}^{\infty} \frac{1}{k}\sin kx \\ &= \sum_{k\in\mathbb{Z}\setminus\{0\}} \frac{-i}{2k} e^{ikx} \end{aligned}$$

### Satz 5.14

a) Es gilt für alle $x\in\mathbb{R}$ (punktweise Konveregenz)

$$s(x) = \begin{cases} \frac{1}{2}(\pi - x) & \text{für } 0<x<2\pi \\ 0 & \text{für } x = 2n\pi, n\in\mathbb{Z} \\ 2\pi\text{-periodisch forgesetzt} & \text{sonst} \end{cases}$$

b) $s$ konvergiert auf jedem Intervall $[a,b]$, das kein $2n\pi$ enthält.

c) $s(x) = \dfrac{s(x+)+s(x-)}{2}$

d) Gibbs-Phänomen (1839–1909)

[Skizze von $\sum_{k=1}^{\infty}\frac{1}{k}\sin kx$ und $\sum_{k=0}^{30}\frac{1}{k}\sin kx$]

Für hinreichend große $n$ überschwingen die Partialsummen

$$s_n(x) = \sum_{k=1}^{n}\frac{1}{k}\sin kx$$

den Sprung von $s$ um $\approx 17.89\%$.

### Beispiel

Wir verwenden die Aussage b) aus obigem Satz: Im Interval $0 < x < 2\pi$ liegt gleichmäßige Konvergenz vor; also dürfen wir gliedweise integrieren.

$$\sum_{k=0}^{\infty} \int_{\pi}^{x} \frac{\sin kt}{k} \,\mathrm{d}t = \int_{\pi}^{x} \frac{1}{2}(\pi - t)\,\mathrm{d}t$$

$$-\sum_{k=1}^{\infty} \frac{\cos kx}{k^2} = \frac{\pi}{2}x - \frac{x^2}{4} + C$$

Also konvergiert

$$\sum_{k=1}^{\infty} \frac{\cos kx}{k^2}$$

gleichmäßig auf $\mathbb{R}$. Somit gilt die Gleichung für $0 \leq x \leq 2\pi$.

Bestimmung von $C$:

$$\begin{aligned} 0 &= \int_{0}^{2\pi} \sum_{k=1}^{\infty} \frac{\cos kt}{k^2} \,\mathrm{d}t \\ &= \int_{0}^{2\pi} \left( \frac{\pi}{2}t - \frac{t^2}{4} + C \right) \,\mathrm{d}t \\ &= \pi^3 - \frac{2\pi^3}{3} + 2\pi C \Rightarrow C = -\frac{\pi^2}{6} \end{aligned}$$

$$\sum_{k=1}^{\infty} \frac{\cos kx}{k^2} = \begin{cases} -\left( \frac{\pi}{2}x - \frac{x^2}{4} - \frac{\pi^3}{6} = \frac{(x-\pi)^2}{4} - \frac{\pi^2}{12} \right) & \text{falls } 0 \leq x \leq 2\pi \\ 2\pi\text{-periodisch forgesesetzt} & \text{sonst} \end{cases}$$

[Skizze]

Für $x = 0$ erhalten wir

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

Für $x = \pi$ erhalten wir

$$\sum_{k=1}^{\infty} \frac{(-1)^k}{k^2} = -\frac{\pi^2}{12}$$

Zusammen erhalten wir

$$\sum_{k=1}^{\infty} \frac{1}{(2k)^2} = \frac{\pi^2}{24}$$

$$\sum_{k=1} \infty \frac{1}{(2k+1)^2} = \frac{\pi^2}{8}$$

### 5.5.3 Fourier-Reihen

Gegeben ein $2\pi$-periodisches $f$. Üblicherweise muss vorrausgesetzt werden, dass $f$ Riemann integrierbar über $[0,2\pi]$ ist. Angenehmer wird es, wenn $f \in C_{2\pi}$. Weiter Vorraussetzungen ermöglichen oft weitere Betrachtungen.

Die trigonometrische Reihe

$$\sum_{k=-\infty}^{\infty} c_k e^{ikx}$$

heißt *Fourier-Reihe* einer Funktion $f$, falls

$$\forall k \in \mathbb{Z}: \quad c_k = c_k(f) = \frac{1}{2\pi} \int_0^{2\pi} f(t) \overline{e^{ikt}} \, \mathrm{d}t$$

Die $c_k$ heißen *Fourier-Koeffizienten*

Wir schreiben

$$f(x) \tilde{\sum}_{k=-\infty}^{\infty} c_k(f) e^{ikx}$$

Approximative Variante

$$S_n f(x) = \sum_{k=-n}^{n} c_k(f) e^{ikx}$$

**Satz 5.15**

Sei $f$ zweimal stetig differenzierbar und $2\pi$-periodisch auf $\mathbb{R}$. Dann gilt

$$\forall x \in \mathbb{R}: \quad f(x) = \sum_{k=-\infty}^{\infty} c_k(f) e^{ikx}$$

$$\begin{aligned}
S_n f(x) &= \sum_{k=-n}^{n} \frac{1}{2\pi} \int_0^{2\pi} f(t) e^{-ikt} \, \mathrm{d}t e^{ikt} \\
&= \frac{1}{2\pi} \int_0^{2\pi} f(t) \underbrace{\sum_{k=-n}^{n} e^{ik(x-t)}}_{\text{Dirichlet-Kern } D_n(x-t)} \mathrm{d}t \\
&= \frac{1}{2\pi} \int_0^{2\pi} f(x-t) D_n(t) \, \mathrm{d}t \\
&= \frac{1}{2\pi} \int_0^{2\pi} f(t) D_n(x-t) \, \mathrm{d}t
\end{aligned}$$

## FFT

$$\sum_{k=-n}^{n} c_k e^{ikx_l}$$

für $l = 1, \dots, N$ bei gegebenen Koeffizientenn $c_k$.

$2n+1$ Operationen für ein $x_l$.

$(2n+1)^2$ Operationen für $N = 2n+1$.

FFT benötigt $n \log n$ Operationen.