Analysis II für Lehrämtler

Sommersemester 2022

Prof. Dr. Ivan Veselić

Inhaltsverzeichnis

Vorwort 1

Literatur 3

1 Reihen 4

2 Absolute Konvergenz von Reihen 11

3 Summen und Produkte von Reihen 21

4 Dezimalbruchentwicklung 25

5 Funktionenfolgen 28

6 Funktionenreihen 41

7 Potenzreihen 45

8 Taylor-Entwicklung 52

9 Funktionen im $\mathbb{R}^n$ 65

10 Normen und Metriken 69

11 Konvergenz und Stetigkeit in metrischen Räumen 79

12 Partielle Differenzierbarkeit von Funktionen mehrerer Variablen 99

13 Parameterabhängige Integrale 103

14 Partielle Ableitungen höherer Ordnung 109

15 Differenzierbare Funktionen 113

16 Der Mittelwertsatz und die Taylor-Formel 126

17 Lokale Extrema 131

18 Implizite Funktionen 141

19 Extrema mit Nebenbedingungen 148

Mindmap 151

Vorwort

Wie dieses Skript entstanden ist

Im Sommersemester 2022 habe ich an der TU Dortmund die Vorlesung „Analysis II für Lehrämtler am Gymnasium und Berufskolleg“ gehalten. Die Analysis I in Wintersemester davor wurde wegen der damals aufgrund von Covid geltenden Einschränkungen mit Kamera aufgenommen und die entsprechenden Videos unter Moodle zu Verfügung gestellt. Da es bei Lehramtsstudenten häufiger zu Überschneidungen im Stundenplan kommt, habe ich mich entschieden, bei der Analysis II ebenfalls Videoaufnahmen zu erstellen.

Das vorliegende Skript orientiert sich zum einen an meinen handschriftlichen Notizen, anhand derer ich die Vorlesung gehalten habe, und zum anderen an der Vorlesung, wie sie an der Tafel stattgefunden hat und in den Videos festgehalten ist. Manchmal bin ich in der Vorlesung selbst von den Notizen abgewichen, im Großen und Ganzen hielt ich mich aber recht eng daran. An manchen Stellen habe ich die handschriftlichen Notizen nach der Vorlesung nochmal nachbearbeitet.

Die Videos — genauer: deren URLs in Moodle — sind an den Stellen im Skript, die dem entsprechenden Vorlesungdatum entsprechen, verlinkt.
An manchen Stellen ist die fortlaufende Zählung im vorliegendem Skript anders als die Nummerierung an der Tafel und den Videos. Dies ist jeweils am Rand vermerkt.

Es gibt sicherlich eine Reihe von kleineren Fehlern und ich arbeite — zusammen mit freundlichen Menschen, die mich auf Fehler hinweisen — daran, diese auszumerzen.

Ein genereller Hinweis zum Sinn und Zweck von Skripten: Das Lesen von Vorlesungsnotizen und Skripten kann das eigene Denken und Nachvollziehen der Sachverhalte, Schritte und Argumente nicht ersetzen. Das eigentliche Ziel ist die Verinnerlichung der mathematischen Wahrheiten, so dass man sie wie etwas Eigenes beherrscht und anderen mitteilen kann. Falls Ihnen an einzelnen Stellen ein Skript zu knapp formuliert ist, um die dargestellten Sachverhalte zu verstehen, konsultieren Sie die angegebene Literatur.

Quellen und Bezüge

In der Vorlesung bezog ich mich insbesondere auf die Skripte [Bru21, Kab18, Len17], die mir die Kollegen Rainer Brück, Winfried Kaballo und Daniel Lenz freundlicherweise zu Verfügung gestellt haben, sowie die Bücher [For17a, Kab97, Koe04b, Tao16b].
Als weitere Literatur können die anderen im Literaturverzeichnis aufgeführten Bücher nützlich sein.

Danksagung

Ich bedanke mich bei den Assistenten der Vorlesung, Herrn M. Kämper, Herrn L. Schroeder und Herrn Dr. A. Seelmann. Herrn Seelmann danke ich in besonderer Weise für die Organisation der Übungen und die Vertretung in einzelnen Vorlesungsterminen. Herrn D. Kreutz danke ich für das Aufnehmen, Nachbearbeiten und Bereitstellen der Vorlesungsvideos. Den Hörerinnen und Hörern meiner Vorlesung danke ich für die Hilfe beim Tafelwischen, das Interesse, die Fragen und die aktive Teilnahme.

Auch diese Vorlesung hat mir viel Freude bereitet!

Herrn A. Nuss danke ich für die kompetente und zügige Erstellung des LaTeX-Satzes dieses Skripts und Herrn Seelmann für Korrekturhinweise.

Dortmund, den 29. März 2024 Ivan Veselić

Literatur

[Bru21] Rainer Brück. Analysis I–III für Lehramt Gymnasium, 2021. TU Dortmund, unpublished manuscript.

[For16] Otto Forster. *Analysis 1. Differential- und Integralrechnung einer Veränderlichen.* Heidelberg: Springer Spektrum, 2016.

[For17a] Otto Forster. *Analysis 2. Differentialrechnung im* $\mathbb{R}^n$*, gewöhnliche Differentialgleichungen.* Wiesbaden: Springer Spektrum, 2017.

[For17b] Otto Forster. *Analysis 3. Maß- und Integrationstheorie, Integralsätze im* $\mathbb{R}^n$ *und Anwendungen.* Wiesbaden: Springer Spektrum, 2017.

[FS13] Otto Forster and Thomas Szymczak. *Übungsbuch zur Analysis 2. Aufgaben und Lösungen.* Heidelberg: Springer Spektrum, 2013.

[FW17] Otto Forster and Rüdiger Wessoly. *Übungsbuch zur Analysis 1. Aufgaben und Lösungen.* Heidelberg: Springer Spektrum, 2017.

[Heu08] Harro Heuser. *Lehrbuch der Analysis. Teil 2.* Wiesbaden: Vieweg+Teubner, 2008.

[Heu09] Harro Heuser. *Lehrbuch der Analysis. Teil 1.* Wiesbaden: Vieweg+Teubner, 2009.

[Kab97] Winfried Kaballo. *Einführung in die Analysis II.* Heidelberg: Spektrum Akademischer Verlag, 1997.

[Kab99] Winfried Kaballo. *Einführung in die Analysis III.* Heidelberg: Spektrum Akademischer Verlag, 1999.

[Kab00] Winfried Kaballo. *Einführung in die Analysis I.* Heidelberg: Spektrum Akademischer Verlag, 2000.

[Kab18] Winfried Kaballo. Vorlesung Analysis I / Lehramt, 2018. TU Dortmund, unpublished manuscript.

[Koe04a] Konrad Königsberger. *Analysis 1.* Berlin: Springer, 2004.

[Koe04b] Konrad Königsberger. *Analysis 2.* Berlin: Springer, 2004.

[Len17] Daniel Lenz. Analysis I–II Notizen, 2017. FSU Jena, unpublished notes.

[Tao16a] Terence Tao. *Analysis I*, volume 37. Singapore: Springer; New Delhi: Hindustan Book Agency, 2016.

[Tao16b] Terence Tao. *Analysis II*, volume 38. Singapore: Springer; New Delhi: Hindustan Book Agency, 2016.

Link: Teil 1 der 1. Vorlesung vom 13.04.2022

1 Reihen

In der Analysis I haben wir uns mit Folgen und insbesondere ihrer Konvergenz beschäftigt. Manche von ihnen waren als Summen gegeben; so tauchte beispielsweise im Kontext der Exponentialfunktion die Folge $(E_n)_{n\in\mathbb{N}}$ auf mit

$$E_n := \sum_{k=0}^{n} \frac{1}{k!}.$$

Wir wollen nun systematisch die n-te Partialsummen

$$s_n := \sum_{k=0}^{n} a_k,$$

die einer $\mathbb{C}$-Folge $(a_k)_{k\in\mathbb{N}_0}$ zugeordnet sind, untersuchen.

Definition 1.1. Die Folge der Partialsummen $(s_n)_{n\in\mathbb{N}_0}$ wird (*unendliche*) *Reihe* genannt und wird gekennzeichnet durch

$$\sum_{k=0}^{\infty} a_k = a_0 + a_1 + a_2 + \dots \quad \text{oder einfach} \quad \sum a_k.$$

Die a_k heißen *Glieder* der Reihe.

Definition 1.2. Eine Reihe $\sum a_k$, $a_k \in \mathbb{C}$, heißt *konvergent* zum Wert $s \in \mathbb{C}$, falls

$$\lim_{n\to\infty} s_n = s$$

gilt. Dies ist äquivalent zu der Aussage

$$\lim_{n\to\infty} \mathrm{Re}(s_n) = \lim_{n\to\infty} \sum_{k=0}^{n} \mathrm{Re}(a_k) = \mathrm{Re}(s) \quad \text{und} \quad \lim_{n\to\infty} \mathrm{Im}(s_n) = \mathrm{Im}(s).$$

In diesem Fall schreiben wir

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

und nennen s den *(Grenz-)Wert* der Reihe oder die *Summe* der Reihe.

Konvergiert die Reihe $\sum a_k$ nicht, so heißt sie *divergent*.

Bemerkung 1.3.

(a) Jede Folge (b_n) kann als Reihe (d.h. Folge von Partialsummen) aufgefasst werden, indem man

$$b_{-1} = 0 \quad \text{und} \quad a_k = b_k - b_{k-1}, \quad k \in \mathbb{N},$$

setzt, da dann $b_n = \sum_{k=0}^{n} a_k = s_n$ gilt.

Streng logisch ist die Untersuchung von Reihen und Folgen äquivalent, es stehen jedoch andere Aspekte im Vordergrund.

(b) Konvergiert die Reihe $\sum_{k=0}^{\infty} a_k$, so tut dies auch die Reihe

$$\sum_{k=n_0}^{\infty} a_k$$

für jedes $n_0 \in \mathbb{N}_0$, wobei jedoch der Wert der Reihe von n_0 abhängt. Insbesondere ist bei einer konvergenten Reihe $\sum a_k$ die Folge

$$n \mapsto r_n := \sum_{k=n+1}^{\infty} a_k$$

der Reihenreste wohldefiniert.

Satz 1.4. *Es sei $a_k \in \mathbb{C}$ für $k \in \mathbb{N}_0$. Dann gilt:*

$$\sum a_k \textit{ ist konvergent} \quad \Rightarrow \quad \lim_{n\to\infty} r_n = 0 \quad \Rightarrow \quad \lim_{k\to\infty} a_k = 0.$$

Beweis. Wir setzen zunächst

$$s_n = \sum_{k=0}^{n} a_k.$$

Dann gilt

$$\begin{aligned} & \sum a_k \text{ ist konvergent} \\ \Leftrightarrow \quad & s := \lim_{n\to\infty} s_n \text{ existiert} \\ \Leftrightarrow \quad & 0 = \lim_{n\to\infty}(s - s_n) = \lim_{n\to\infty} r_n. \end{aligned}$$

Da $a_n = r_{n-1} - r_n$ gilt, folgt dann

$$\lim_{n\to\infty} a_n = \lim_{n\to\infty} r_{n-1} - \lim_{n\to\infty} r_n = 0.$$ □

Satz 1.5 (Cauchy-Kriterium für Reihen)**.** *Es sei $\sum a_k$ eine unendliche Reihe. Dann sind folgende Aussagen äquivalent:*

(i) *$\sum a_k$ ist konvergent;*

(ii) *Für alle $\varepsilon > 0$ existieren $n_0 \in \mathbb{N}$ derart, dass*

$$\forall\, m, n \in \mathbb{N} \textit{ mit } m > n \geq n_0 : \quad \left| \sum_{k=n+1}^{m} a_k \right| < \varepsilon.$$

Beweis. Man beachte, dass $s_m - s_n = \sum_{k=n+1}^{m} a_k$, es gilt also

$$\left| \sum_{k=n+1}^{m} a_k \right| = |s_m - s_n|.$$

Anwendung des Cauchy-Kriteriums für Folgen liefert dann direkt die Behauptung. □

Im Folgenden betrachten wir nun zwei Anwendungen des Cauchy-Kriteriums, wo bei beiden auch Monotonie eine Rolle spielt.

Korollar 1.6. *Es sei $(a_k)_{k\in\mathbb{N}_0}$ eine antitone $\mathbb{R}$-Folge und $\sum a_k < \infty$. Dann gilt*

Im Video: **Korollar 1.5a.**

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

Insbesondere geht a_k „schneller gegen Null“ als $\frac{1}{k}$.

Beweis. Da (a_k) antiton ist, gilt nach dem Cauchy-Kriterium

$$2j\cdot a_{2j} \leq 2\Big(a_{j+1} + a_{j+2} + \ldots + a_{2j}\Big) \leq 2\varepsilon,$$

falls j groß genug ist. Analog gilt

$$(2j-1)\cdot a_{2j-1} \leq 2\Big(a_j + a_{j+1} + \ldots + a_{2j-1}\Big) \leq 2\varepsilon.$$

Also ist $k\cdot a_k$ eine Nullfolge. □

LINK: TEIL 2 DER 1. VORLESUNG VOM 13.04.2022

Satz 1.7 (Leibnitz-Kriterium). *Es sei $(a_k)_{k\in\mathbb{N}_0}$ eine antitone Nullfolge in $\mathbb{R}_+$. Dann ist die alternierende Reihe*

Im Video: **Satz 1.5b.**

$$\sum(-1)^k\cdot a_k$$

konvergent.

Beweis. Für gerade n und $m > n$ gilt

$$\sum_{k=n}^{m}(-1)^k a_k = a_n - \underbrace{(a_{n+1} - a_{n+2})}_{\underset{\text{antiton}}{\geq}\ 0} - \underbrace{(a_{n+3} - a_{n+4})}_{\geq 0} - \ldots \leq a_n$$

und

$$\sum_{k=n}^{m} = \underbrace{(a_n - a_{n+1})}_{\geq 0} + \underbrace{(a_{n+2} - a_{n+3})}_{\geq 0} + \ldots \geq 0.$$

Ganz analog für ungerade n und $m > n$ gilt auch

$$-a_n \leq \sum_{k=n}^{m}(-1)^k a_k \leq 0.$$

In beiden Fällen für $n \leq m \in \mathbb{N}$ folgt schließlich

$$\left|\sum_{k=n}^{m}(-1)^k a_k\right| \leq |a_k| \xrightarrow[n\to\infty]{} 0.$$

Damit ist das Cauchy-Kriterium anwendbar und es folgt direkt die Behauptung. □

Fehlerabschätzung für die Konvergenz: Es gilt

$$\left|\sum_{k=1}^{\infty}(-1)^k a_k - \sum_{k=1}^{n-1}(-1)^k a_k\right| \le \sup_{m\in\mathbb{N}}\left|\sum_{k=n}^{m}(-1)^k a_k\right| \le |a_n|.$$

Beispiel 1.8 (Geometrische Reihe). Es sei $z \in \mathbb{C}$, $a_k = z^k$, $k \in \mathbb{N}_0$. Für $z \neq 1$ gilt dann bekannterweise

Im Video: **Beispiel 1.6**.

$$s_n = \sum_{k=0}^{n} z^k = \frac{1-z^{n+1}}{1-z}.$$

Falls $|z| < 1$ gilt, so gilt $z^{n+1} \to 0$ für $n \to \infty$ und damit

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

Falls $|z| > 1$ gilt, dann ist $a_k = z^k$ keine Nullfolge, so dass die Reihe $\sum a_k$ nach Satz 1.4 divergiert.

Für $z = \frac{1}{2}$ gilt beispielsweise

$$s = \sum_{k=0}^{\infty} \frac{1}{2^k} = \frac{1}{1-\frac{1}{2}} = 2.$$

Beispiel 1.9 (Harmonische Reihe). Die Reihe $\sum a_k$ mit $a_k = \frac{1}{k}$, $a_0 = 0$, ist divergent.

Im Video: **Beispiel 1.7**.

Behauptung. Für alle $m \in \mathbb{N}_0$ gilt

$$s_{2^m} = \sum_{k=1}^{2^m} \frac{1}{k} \ge 1 + \frac{m}{2}.$$

Beweis der Behauptung. Durch vollständige Induktion über $m \in \mathbb{N}$:

(IA) $\underline{m = 0:}$
Es gilt offensichtlich $s_1 = a_1 + a_0 = 1 \ge 1 + 0$.

(IV) Es gilt

$$s_{2^m} = \sum_{k=1}^{2^m} \frac{1}{k} \ge 1 + \frac{m}{2}$$

für ein $m \in \mathbb{N}$.

(IS) $\underline{m \mapsto m+1:}$
Es gilt

$$s_{2^{m+1}} = \sum_{k=1}^{2^m} \frac{1}{k} + \sum_{k=2^m+1}^{2^{m+1}} \underbrace{\frac{1}{k}}_{\ge \frac{1}{2^{m+1}}}$$

$$\overset{(IV)}{\geq} \left(1+\frac{m}{2}\right) + \underbrace{\sum_{k=2^m+1}^{2^{m+1}} \frac{1}{2^{m+1}}}_{=\frac{2^{m+1}-2^m}{2^{m+1}}=\frac{1-\frac{1}{2}}{1}=\frac{1}{2}}$$

$$\geq 1+\frac{m}{2}+\frac{1}{2} = 1+\frac{m+1}{2}.$$

Also gilt die Behauptung für alle $m \in \mathbb{N}_0$.

Damit wächst die Folge der Partialsummen unbeschränkt, denn man kann zu $n \in \mathbb{N}$ ein $m \in \mathbb{N}$ finden mit $n \geq 2^m$. Damit folgt

$$s_n \overset{\frac{1}{k}>0}{\geq} s_{2^m} \geq 1+\frac{m}{2} \quad \Rightarrow \quad (s_k)_k \text{ ist isoton.}$$

Insbesondere ist $\sum \frac{1}{k}$ – wie oben erwähnt – divergent.

Die obige Erkenntnis vergleichen wir nun mit Satz 1.4: $\left(\frac{1}{k}\right)_k$ ist zwar eine Nullfolge, aber $\sum \frac{1}{k}$ ist divergent. Die Reihenglieder streben hier „zu langsam gegen Null“; an dieser Stelle kann man diesen Zusammenhang mit Korollar 1.6 vergleichen.

Beispiel 1.10. Aus den obigen Erkenntnissen ergibt sich die Divergenz des uneigentlichen Integrals

Im Video: **Beispiel 1.8.**

$$\int_0^\infty \frac{|\sin(x)|}{x}\,\mathrm{d}x.$$

Beweis. Für alle $n \in \mathbb{N}$ gilt

$$\int_0^{n\cdot\pi} \frac{|\sin(x)|}{x}\,\mathrm{d}x$$

$$= \sum_{k=1}^{n} \int_{(k-1)\pi}^{k\pi} \frac{|\sin(x)|}{\underbrace{x}_{\leq k\pi}}$$

$$\geq \sum_{k=1}^{n} \frac{1}{k\pi} \underbrace{\int_{(k-1)\pi}^{k\pi} |\sin(x)|\,\mathrm{d}x}_{\overset{\text{period.}}{=} \int_0^\pi |\sin(x)|\,\mathrm{d}x}$$

$$= \sum_{k=1}^{n} \frac{1}{k\pi} \cdot \underbrace{\Big[-\cos(x)\Big]_0^\pi}_{=2}.$$

Es gilt also

$$\int_0^{n\pi} \frac{|\sin(x)|}{x}\,\mathrm{d}x \geq \frac{2}{\pi}\sum_{k=1}^{n}\frac{1}{k} \xrightarrow[n\to\infty]{} \infty,$$

so dass das obige uneigentliche Integral divergent ist. □

Später werden wir sehen, dass es einen systematischen Zusammenhang zwischen der Konvergenz von Reihen und der Konvergenz von uneigentlichen Integralen gibt.

LINK: TEIL 1 DER 2. VORLESUNG VOM 14.04.2022

Beispiel 1.11 (Alternierende harmonische Reihe)**.** Die Folge $(a_n)_{n\in\mathbb{N}}$ mit $a_n = \frac{1}{n}$ ist eine antitone Nullfolge. Nach dem Leibnitz-Kriterium 1.7 ist dann die alternierende Reihe

Im Video: **Beispiel 1.5***c*.

$$\sum_{k\in\mathbb{N}} (-1)^{k+1}\frac{1}{k} = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + - \ldots$$

konvergent. Aufwändigere Methoden liefern insbesondere

$$-\sum_{k\in\mathbb{N}} (-1)^k \frac{1}{k} = \ln(2).$$

Dass die Reihenglieder eine Nullfolge bilden müssen, ist zwar eine notwendige Bedingung, aber sie ist nicht hinreichend!

Bekannterweise kennzeichnen wir das Leibnitz-Kriterium durch die Voraussetzungen

$$a_n \searrow 0 \quad (n\to\infty) \quad \Rightarrow \quad \sum_{k\in\mathbb{N}_0} (-1)^k a_k \text{ ist konvergent}$$

und $|s - s_{n-1}| \leq a_n$.

Lemma 1.12.

Im Video: **Lemma 1.9**.

(i) *Es sei $a_k \geq 0$ für alle $k\in\mathbb{N}$. Dann gilt:*

$$\sum a_k \quad \textit{ist konvergent} \quad \Leftrightarrow \quad \textit{Die Folge } (s_n)_{n\in\mathbb{N}} \textit{ ist beschränkt.} \quad :\Leftrightarrow \quad \sum a_k < \infty.$$

(ii) *Es seien $a_k, b_k \geq 0$ für alle $k\in\mathbb{N}$. Gibt es nun ein $C > 0$ und $n_0\in\mathbb{N}$ mit*

$$a_k \leq C\cdot b_k \quad \forall\, k \geq n_0,$$

so gelten:

- *$\sum b_k$ ist konvergent $\Rightarrow$ $\sum a_k$ ist konvergent;*
- *$\sum a_k$ ist divergent $\Rightarrow$ $\sum b_k$ ist divergent.*

Beweis. Übung! □

Satz 1.13. *Es sei $f \in C\big([1,\infty)\big)$ antiton und $f \geq 0$ auf $[1,\infty)$. Dann gilt:*

Im Video: **Satz 1.10**.

$$\text{Die Reihe } \sum_{k=1}^{\infty} f(k) \text{ konvergiert} \quad \Leftrightarrow \quad \text{Das uneigentliche Integral } \int_1^{\infty} f(x)\,\mathrm{d}x \text{ existiert.}$$

Beweisskizze. Da $f \geq 0$ auf $[1,\infty)$ ist, gilt

$$\limsup_{m\to\infty} \sum_{k=1}^{m} f(k) < \infty.$$

Nach Lemma 1.12 (i) ist dann $\sum_{k\in\mathbb{N}} f(k)$ konvergent. Außerdem gilt

$$\limsup_{m\to\infty} \int_1^m f(x)\,\mathrm{d}x < \infty \quad \Rightarrow \quad \int_1^{\infty} f(x)\,\mathrm{d}x \text{ existiert.}$$

Da f antiton ist, gilt weiter

$$f(k) \leq f(x) \quad \text{für } x \in [k-1,k].$$

Damit folgt

$$\sum_{k=2}^{m} \int_{k-1}^{k} \mathrm{d}x\; f(k) \leq \sum_{k=2}^{m} \int_{k-1}^{k} f(x)\,\mathrm{d}x = \int_1^m f(x)\,\mathrm{d}x \leq \underbrace{\ldots}_{\text{analog}} \leq \sum_{k=2}^{m} f(k-1).$$

Ist (a_k) eine isotone und (b_k) eine konvergente Folge mit $a_k \leq b_k$ für alle $k \in \mathbb{N}$, so ist (b_k) beschränkt und damit (a_k) konvergent.

Den Argumentationsvorgang kann man in beiden Richtungen anwenden, weshalb wir insgesamt eine Äquivalenz vorliegen haben. □

Beispiel 1.14. Die Reihe

Im Video: **Beispiel 1.11**.

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

ist konvergent.

Beweis. Es reicht aus, zu zeigen, dass die Folge (s_n) der Partialsummen beschränkt ist, da $\frac{1}{k^2} \geq 0$ für alle $k \in \mathbb{N}$ gilt. Für alle $n \geq 2$ gilt

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

$$\leq 1 + \sum_{k=2}^{n} \underbrace{\frac{1}{k(k-1)}}_{=\frac{1}{k-1}-\frac{1}{k}}$$

$$= 1 + \left(1 - \frac{1}{2}\right) + \left(\frac{1}{2} - \frac{1}{3}\right) + \ldots + \left(\frac{1}{n-1} - \frac{1}{n}\right)$$

$$= 1 + 1 - \frac{1}{n} < 2.$$

Also ist (s_n) beschränkt, weshalb $\sum \frac{1}{k^2}$ konvergent ist. □

$\frac{1}{k^2}$ fällt schnell genug ab, so dass $\sum \frac{1}{k^2} < \infty$ gilt. $\frac{1}{k}$ fällt hingegen ‚zu langsam' ab, so dass $\sum \frac{1}{k}$ divergiert. Mit aufwändigeren Methoden zeigt man sogar:

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

2 Absolute Konvergenz von Reihen

Ist $\sum_{k \in \mathbb{N}_0} a_k = a_0 + a_1 + a_2 + \ldots$ konvergent gegen ein $s \in \mathbb{C}$, so können wir beliebig zusätzliche Klammern setzen, d.h. für jede streng isotone Folge $(k_n)_n \subset \mathbb{N}$ mit $k_0 = 0$ gilt

$$s = \sum_{j=0}^{\infty} \underbrace{\left(\sum_{k=k_j+1}^{k_{j+1}} a_k \right)}_{=:b_j} = \lim_{j \to \infty} \underbrace{\sum_{k=0}^{k_{j+1}} a_k}_{=:\tilde{S}_j(b)},$$

da die Folge $\tilde{S}_j(b)$ der Partialsummen zu der Folge (b_j) eine Teilfolge der Folge $n \mapsto s_n = \sum_{k=0}^{n} a_k$ ist.

Dagegen dürfen Klammern in einer konvergenten in einer konvergenten unendlichen Reihe $\sum a_k$ <u>nicht</u> weggelassen werden! So ist zum Beispiel

$$0 = \sum_{k \in \mathbb{N}_0} 0 = (1-1) + (1-1) + (1-1) + \ldots$$

sicherlich konvergent. Lässt man jedoch die Klammern weg, so erhalten wir

$$1 - 1 + 1 - 1 + 1 - 1 + \ldots = \sum_{k \in \mathbb{N}_0} (-1)^k$$

mit divergenter n-ter Partialsumme

$$s_n = \begin{cases} 0, & \text{falls } n \text{ ungerade,} \\ 1, & \text{falls } n \text{ gerade.} \end{cases}$$

Dies motiviert:

Definition 2.1. Für $a_k \in \mathbb{C}$, $k \in \mathbb{N}_0$, heißt die Reihe $\sum a_k$ *absolut konvergent*, falls die Reihe $\sum |a_k|$ konvergent ist.

LINK: TEIL 2 DER 2. VORLESUNG VOM 14.04.2022

Lemma 2.2. *Es gilt:*

$$\sum a_k \text{ ist absolut konvergent} \quad \Rightarrow \quad \sum a_k \text{ ist konvergent.}$$

Beweis. Für alle $m, n \in \mathbb{N}$ mit $m > n$ gilt

$$\left| \sum_{k=n+1}^{m} a_k \right| \leq \sum_{k=n+1}^{m} |a_k| \underset{\text{für } n \text{ groß}}{\overset{\text{abs. konv.}}{<}} \varepsilon.$$

Damit liefert das Cauchy-Kriterium 1.5 die Behauptung. □

Das Beispiel 1.11 über die harmonische Reihe zeigt wiederum, dass die Umkehrung der Aussage im Allgemeinen nicht gilt.

Bemerkung 2.3. Es sei $a_k \in \mathbb{R}$ für alle $k \in \mathbb{N}$ und $a_k^+ := \max\{0, a_k\}$ bzw. $a_k^- := \max\{0, -a_k\}$. Dann gelten:

(α) $\sum a_k$ ist absolut konvergent $\quad\Leftrightarrow\quad$ $\sum a_k^+$ und $\sum a_k^-$ sind konvergent.

Beweisidee. Es gilt

$$\sum_{k=0}^{n} a_k^{\pm} \leq \sum_{k=0}^{n} \underbrace{|a_k|}_{\geq 0} \leq \sum_{k\in\mathbb{N}_0} |a_k| \leq \sum_{k\in\mathbb{N}_0} \left(a_k^+ + a_k^-\right) = \lim_{m\to\infty} \sum_{k=0}^{m} a_k^+ + \lim_{m\to\infty} \sum_{k=0}^{m} a_k^-.$$

Das Monotoniekriterium für Folgen liefert dann die Behauptung.

(β) Es sei $\sum a_k$ konvergent, aber nicht absolut konvergent. Dann sind $\sum a_k^+$ und $\sum a_k^-$ divergent.

Beweisidee. Widerspruchsbeweis mit Anwendung von Teil (α).

Definition 2.4. Es seien $(a_k), (b_k) \subset \mathbb{C}$. Falls für fast alle $k \in \mathbb{N}$

$$|a_k| \leq b_k \quad \text{bzw.} \quad 0 \leq b_k \leq a_k$$

gilt, so bezeichnen wir $\sum b_k$ als *Majorante bzw. Minorante* von $\sum a_k$.

Aus Lemma 1.12 folgt direkt:

Satz 2.5 (Majoranten- & Minoranten-Kriterium)**.** *Es seien $(a_k), (b_k) \subset \mathbb{C}$. Dann gelten:*

(i) *Falls $\sum b_k$ eine Majorante von $\sum a_k$ ist und zusätzlich konvergent ist, so ist $\sum a_k$ absolut konvergent.*

(ii) *Falls $\sum b_k$ eine Minorante von $\sum a_k$ ist und zusätzlich divergent ist, so ist auch $\sum a_k$ divergent.*

Beispiel 2.6. Es sei $p \in \mathbb{R}$, $p > 0$, und $a_k := \frac{1}{k^p}$.

Für $p \geq 2$ gilt
$$a_k = \frac{1}{k^p} \leq \frac{1}{k^2} =: b_k.$$
Nach dem Majoranten-Kriterium ist dann $\sum a_k$ (absolut) konvergent.

Für $p \leq 1$ gilt
$$b_k := \frac{1}{k} \leq a_k.$$
Nach dem Minoranten-Kriterium ist dann $\sum a_k$ divergent.

Um den Bereich $p \in (1, 2)$ zu klären, benutzen wir das Integral-Vergleich-Kriterium und Beispiel 15.3 (β) aus Analysis I (Lehramt): Ist $p > 0$ und

$$f \colon [1, \infty) \to \mathbb{R}, \quad f(x) := \frac{1}{x^p},$$

so existiert das uneigentliche Integral $\int_1^\infty \frac{1}{x^p}$ genau dann, wenn $p > 1$ ist. Also gilt:

$$\sum_k \frac{1}{k^p} = \sum_k f(k) \text{ ist konvergent} \quad \overset{\text{Int.-Vgl.}}{\Leftrightarrow} \quad \int_1^\infty x^{-p} \text{ ist konvergent} \quad \Leftrightarrow \quad p > 1.$$

Erinnerung: Ist eine Folge (a_k) bestimmt divergent gegen $-\infty$, so schreibt man

$$\liminf_{k\to\infty} a_k = -\infty = \limsup_{k\to\infty} a_k.$$

Satz 2.7 (Wurzelkriterium)**.**

(i) *Gilt*
$$w := \limsup_{k\to\infty} \sqrt[k]{|a_k|} < 1,$$
so ist $\sum a_k$ absolut konvergent;

(ii) *Die Reihe $\sum a_k$ ist divergent, falls*
$$\sqrt[k]{|a_k|} \geq 1 \quad \textit{für unendlich viele Indizes } k$$
gilt – insbesondere falls $w > 1$.

Beweis.

(i) Wähle ein $q \in (w, 1)$. Da $w < q$ gilt, existiert ein $k_0 \in \mathbb{N}$ mit $\sqrt[k]{|a_k|}$ (d.h. $|a_k| \leq q^k$) für alle $k \geq k_0$. Also besitzt $\sum a_k$ eine Majorante $\sum q^k$, die konvergent ist. Nach dem Majoranten-Kriterium aus Satz 2.5 ist dann $\sum a_k$ absolut konvergent.

(ii) Es existiert eine unendliche Teilmenge $K \subset N$ mit

$$\forall\, k \in K: \quad \sqrt[k]{|a_k|} \geq 1 \quad (\text{also } |a_k| \geq 1).$$

Insbesondere ist (a_k) keine Nullfolge, weshalb $\sum a_k$ nach Satz 1.4 divergiert. □

Beispiel 2.8.

(a) Für $a_k = \frac{k^4}{2^k}$ ist die Reihe $\sum a_k$ konvergent, da

$$\lim_{k\to\infty} \sqrt[k]{a_k} = \lim_{k\to\infty} \left(\frac{k^4}{2^k}\right)^{\frac{1}{k}} = \lim_{k\to\infty} \frac{k^{\frac{4}{k}}}{2} = \frac{1}{2} \underbrace{\lim_{k\to\infty} k^{\frac{4}{k}}}_{=1} = \frac{1}{2} < 1$$

und damit das Wurzelkriterium aus Satz 2.7 anwendbar ist.

(b) Analog folgt: Für alle $p \in \mathbb{N}_0$ und $z \in \mathbb{C}$, $|z| < 1$, ist

$$\sum_{k\in\mathbb{N}} k^p z^k$$

absolut konvergent.

(c) Für $a_k = \left(\frac{2+(-1)^k}{4}\right)^k$ gilt

$$\sqrt[k]{|a_k|} = \frac{1}{4}\big(2 + (-1)^k\big),$$

also

$$\limsup_{k\to\infty} \sqrt[k]{|a_k|} = \frac{3}{4} \quad \left(\text{und } \liminf_{k\to\infty} \sqrt[k]{|k|} = \frac{1}{4}\right).$$

Nach dem Wurzelkriterium ist $\sum a_k$ dann (absolut) konvergent.

Satz 2.9 (Quotientenkriterium)**.**

(i) *Gilt $a_k \neq 0$ für alle k und*

$$v := \limsup_{k\to\infty} \left|\frac{a_{k+1}}{a_k}\right| < 1,$$

so ist $\sum a_k$ absolut konvergent;

(ii) *$\sum a_k$ divergiert, falls ein $k_0 \in \mathbb{N}$ existiert mit*

$$\left|\frac{a_{k+1}}{a_k}\right| \geq 1 \quad \forall\, k \geq k_0,$$

insbesondere falls

$$\liminf_{k\to\infty} \left|\frac{a_{k+1}}{a_k}\right| > 1.$$

Beweis.

(i) Wähle ein $q \in (v, 1)$. Insbesondere bedeutet das, dass

$$\limsup_{k\to\infty} \left|\frac{a_{k+1}}{a_k}\right| \leq q.$$

Also existiert ein $k_0 \in \mathbb{N}$ mit

$$\left|\frac{a_{k+1}}{a_k}\right| \leq q \quad \forall\, k \geq k_0.$$

Es folgt

$$\left|\frac{a_k}{a_{k_0}}\right| = \left|\frac{a_k}{a_{k-1}} \cdot \frac{a_{k-1}}{a_{k-2}} \cdots \frac{a_{k_0+1}}{a_{k_0}}\right| \leq q^{k-k_0}.$$

Damit gilt

$$|a_k| \leq q^{k-k_0}|a_{k_0}|.$$

Wegen $q < 1$ folgt, dass $\sum q^k$ konvergent ist und zusammen mit Lemma 1.12 und dem Majoranten-Kriterium die absolute Konvergenz von $\sum a_k$.

(ii) Für alle $k \geq k_0$ gilt

$$\left|\frac{a_{k+1}}{a_k}\right| > 1 \quad \Rightarrow \quad |a_{k+1}| > |a_k|.$$

Damit ist a_k keine Nullfolge und $\sum a_k$ nach Satz 1.4 divergent. □

LINK: TEIL 1 DER 3. VORLESUNG VOM 20.04.2022

Beispiel 2.10. Wir untersuchen verschiedene Reihen $\sum a_k$ auf (absoluter) Konvergenz.

(a) Es sei $a_k := \frac{z^k}{k!}$, $z \neq 0 \in \mathbb{C}$. Dann gilt

$$\lim_{k\to\infty} \left|\frac{a_{k+1}}{a_k}\right| = \lim_{k\to\infty} \left|\frac{z^{k+1}}{(k+1)!} \cdot \frac{k!}{z^k}\right| = \lim_{k\to\infty} \frac{|z|}{k+1} = 0.$$

Also ist $\sum a_k$ absolut konvergent (auch für $z = 0$).

(b) Es sei $a_k := \frac{k!}{k^k} \cdot z^k$, $z \neq 0 \in \mathbb{C}$. Dann gilt

$$\frac{a_{k+1}}{a_k} = \frac{(k+1)!}{(k+1)^{k+1}} \cdot z \cdot \frac{k^k}{k!} = z \cdot \frac{(k+1)k^k}{(k+1)(k+1)^k} = z \cdot \left(\frac{k}{k+1}\right)^k = \frac{z}{\left(1+\frac{1}{k}\right)^k} \xrightarrow[k\to\infty]{} \frac{z}{\mathrm{e}}.$$

Also gilt:

- Falls $|z| < \mathrm{e}$, dann ist $\sum a_k$ absolut konvergent;
- Falls $|z| > \mathrm{e}$, dann ist $\sum a_k$ divergent.

Das Beispiel (b) ist auch mit dem Wurzelkriterium lösbar, was jedoch aufwändiger ist.

(c) Es sei $a_k := k^{-p}$, $p > 0$, $k \in \mathbb{N}$. Wir wissen außerdem, dass

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

Da die Exponentialfunktion $\exp(\cdot)$ stetig für $p > 0$ ist, folgt

$$1 > \sqrt[k]{\frac{1}{k^p}} = k^{-\frac{p}{k}} = \exp\left(-p \cdot \frac{\ln(k)}{k}\right) \xrightarrow[k\to\infty]{} \mathrm{e}^{-p\cdot 0} = 1.$$

Also ist das Wurzelkriterium nicht auf $\sum \frac{1}{k^p}$ anwendbar. Wie sieht es mit dem Quotientenkriterium aus? Es gilt

$$\frac{a_{k+1}}{a_k} = \frac{k^p}{(k+1)^p} = \frac{1}{\left(1+\frac{1}{k}\right)^p} \xrightarrow[k\to\infty]{} \frac{1}{(1+0)^p} = 1$$

und

$$\frac{a_{k+1}}{a_k} < 1 \quad \forall\, k \in \mathbb{N}.$$

Damit ist auch das Quotientenkriterium nicht anwendbar!

(d) Für ein $k \in \mathbb{N}$ sei

$$a_k := \begin{cases} 2^{-k}, & \text{falls } k \text{ gerade,} \\ 3^{-k}, & \text{falls } k \text{ ungerade.} \end{cases}$$

Ist das Quotientenkriterium anwendbar? Es gilt:

$$k \text{ gerade} \quad \Rightarrow \quad \frac{a_{k+1}}{a_k} = \frac{2^k}{3^k} \cdot \frac{1}{3} = \frac{1}{3} \cdot \left(\frac{2}{3}\right)^k \xrightarrow[k\to\infty]{} 0;$$

$$k \text{ ungerade} \quad \Rightarrow \quad \frac{a_{k+1}}{a_k} = \frac{1}{2} \cdot \left(\frac{3}{2}\right)^k \xrightarrow[k\to\infty]{} \infty.$$

Das Quotientenkriterium ist also nicht anwendbar. Wie sieht es mit dem Wurzelkriterium aus? Es gilt

$$\sqrt[k]{a_k} \leq \sqrt[k]{2^{-k}} = \frac{1}{2} < 1.$$

Das Wurzelkriterium ist also anwendbar und liefert sogar die absolute Konvergenz von $\sum a_k$.

Man erkennt, dass das Quotientenkriterium ineffizient bei der Vermischung von zwei schnell (aber unterschiedlich schnell) fallenden Folgen ist.

Bemerkung. Die Reihenglieder $a_k = 0$ tragen nicht zur Summe bei, weshalb solche weggelassen werden können. Damit ist die Annahme $a_n \neq 0$ im Quotientenkriterium keine wirkliche Einschränkung.

Man bemerkt außerdem: Wann immer das Quotientenkriterium eine Aussage zur Konvergenz einer Reihe liefert, dann auch das Wurzelkriterium. Letztere ist aber häufig rechnerisch aufwändiger. Diese Aussage ist im folgenden Satz enthalten:

Satz 2.11. *Es sei $C \in \mathbb{R}$, $a_k > 0$ und $\frac{a_{k+1}}{a_k} \le C$ für alle $k \in \mathbb{N}$. Dann gilt*

$$\liminf_{k\to\infty} \frac{a_{k+1}}{a_k} \le \liminf_{k\to\infty} \sqrt[k]{a_k} \le \limsup_{k\to\infty} \sqrt[k]{a_k} \le \limsup_{k\to\infty} \frac{a_{k+1}}{a_k}.$$

Beweis. Übung!

Definition 2.12.

(a) Eine konvergente, aber nicht absolut konvergente Reihe $\sum a_k$ nennen wir *bedingt konvergent*.

(b) Eine Reihe $\sum b_k$ heißt *Umordnung* der Reihe $\sum a_k$, falls eine Bijektion

$$\varphi\colon \mathbb{N} \to \mathbb{N}$$

existiert, so dass

$$\forall\, k \in \mathbb{N}: \quad b_k = a_{\varphi(k)}.$$

Die Definition ist durch Folgendes motiviert:

Ist $b_1 + \ldots + b_N$ eine Umordnung der endlichen Reihe $a_1 + \ldots + a_N$, so haben beide denselben Wert wegen des Kommutativgesetzes in $\mathbb{C}$. (Man kann beide Reihen auch als unendliche Reihen auffassen, indem man $a_{N+1} = b_{N+1} = a_{N+2} = b_{N+2} = \ldots = 0$ setzt. Dann sind beide Reihen offensichtlich absolut konvergent.)

Im Gegensatz dazu steht:

Beispiel 2.13. Nach dem Leibnitz-Kriterium ist die alternierende harmonische Reihe

$$s := -\sum_{k\in\mathbb{N}} \frac{(-1)^k}{k} = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} - \ldots$$

konvergent, aber nicht absolut konvergent, da $\sum \frac{1}{k}$ divergiert.

Gelte für diese Reihe das „unendliche Kommutativgesetz“, so muss die folgende Umordnung denselben Wert haben:

$$s = \underbrace{\overbrace{\underbrace{\overbrace{\underbrace{\overbrace{\underbrace{1}_{=:t_1} - \frac{1}{2}}^{=:t_2} - \frac{1}{4}}_{=:t_3} + \frac{1}{3}}^{=:t_4} - \frac{1}{6}}_{=:t_5} - \frac{1}{8}}^{=:t_6} + \frac{1}{5}}_{=:t_7} - \frac{1}{10} - \frac{1}{12} + \frac{1}{7} - \frac{1}{14} - \ldots$$

$$
\begin{aligned}
&= \left(1 - \frac{1}{2}\right) - \frac{1}{4} + \left(\frac{1}{3} - \frac{1}{6}\right) - \frac{1}{8} + \left(\frac{1}{5} - \frac{1}{10}\right) - \frac{1}{12} + \left(\frac{1}{7} - \frac{1}{14}\right) - \dots \\
&= \frac{1}{2} - \frac{1}{4} + \frac{1}{6} - \frac{1}{8} + \frac{1}{10} - \frac{1}{12} + \frac{1}{14} - \dots \\
&= \frac{1}{2}\left(1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} - \frac{1}{6} + \frac{1}{7} - \dots\right) \\
&= \frac{1}{2} \cdot s.
\end{aligned}
$$

Die Aussage wäre war, wenn $s = 0$ wäre, aber es ist bekannt, dass $s = \ln(2) \neq 0$ gilt. Damit sind wir bei der Umordnung dieser unendlichen Reihe auf einen anderen Wert gestoßen, weshalb das „unendliche Kommutativgesetz“ für diese Reihe nicht gilt.
Noch zu klären: Konvergiert die Reihe auch ohne Klammersetzung gegen $\frac{1}{2}s$? Die Reihe (ohne Klammern) hat dabei Partialsummen $t_1, t_2, t_3, t_4, \dots$, wobei die Partialsummen

$$t_2 \quad t_3 \quad t_5 \quad t_6 \quad t_8 \quad t_9 \quad t_{11} \quad \dots \to \frac{s}{2}$$

der Reihen mit Klammern eine Teilfolge der Folge der vorigen Partialsummen bilden. Wir kennen also den Grenzwert der Teilfolge und es stellt sich die Frage, ob der Grenzwert auch für die Gesamtfolge gleich bleibt.

Beobachtung.

* Vom Übergang der Folge zur Teilfolge wird maximal ein Folgeglied zwischen zwei aufeinanderfolgenden Folgegliedern weggelassen;
* Der Abstand zum nächsten Folgeglied hat die Form
$$|t_k - t_{k+1}| = \frac{1}{m} \quad \text{mit } m > k.$$
So gilt beispielsweise
$$|t_{10} - t_{11}| = \frac{1}{14};$$
* Wegen $\frac{1}{m} \leq \frac{1}{k} \underset{k\to\infty}{\longrightarrow} 0$ konvergiert $(t_k)_k$ zum selben Limes.

Für die obige Reihe gilt also das „unendliche Kommutativgesetz“ generell nicht.

LINK: TEIL 2 DER 3. VORLESUNG VOM 20.04.2022

Gute Nachricht: Für absolut konvergente Reihen gilt das „unendliche Kommutativgesetz“!

Satz 2.14 (Umordnungssatz). *Es sei $a_k \in \mathbb{C}$ für $k \in \mathbb{N}$. Dann gelten die folgenden Aussagen:*

(i) *Es gilt*

$$\sum_{k\in\mathbb{N}} a_k \text{ ist absolut konvergent}$$

$$\Leftrightarrow \quad \text{Für alle Bijektionen } \varphi\colon \mathbb{N}\to\mathbb{N} \text{ ist } \sum a_{\varphi(k)} \text{ konvergent}$$

$$\Rightarrow \quad \text{Für alle Bijektionen } \varphi\colon \mathbb{N}\to\mathbb{N} \text{ gilt: } \sum_{k\in\mathbb{N}} a_k = \sum_{k\in\mathbb{N}} a_{\varphi(k)};$$

(ii) *Falls $\sum a_k$ bedingt konvergent ist, so existiert eine Bijektion $\varphi\colon \mathbb{N}\to\mathbb{N}$ derart, dass $\sum a_{\varphi(k)}$ divergent ist;*

(iii) *Sind alle $a_k \in \mathbb{R}$, so gilt sogar:*
Falls $\sum a_k$ bedingt konvergent ist, so existiert für alle $\alpha\in\mathbb{R}$ eine Bijektion $\varphi = \varphi_\alpha\colon \mathbb{N}\to\mathbb{N}$ mit

$$\sum_{k\in\mathbb{N}} a_{\varphi(k)} = \alpha.$$

Beweis.

(iii) Da $\sum a_k$ bedingt konvergent ist, sind $\sum a_k^+$ und $\sum a_k^-$ nach Bemerkung 2.3 (β) bestimmt divergent gegen ∞. Wir haben also einen „unerschöpflichen Vorrat an sowohl positiven als auch negativen Reihengliedern“. Wähle dazu ein $\alpha\in\mathbb{R}$. Summiere nun $N\in\mathbb{N}$ positive Glieder, so dass für eine Bijektion $\varphi\colon\mathbb{N}\to\mathbb{N}$ gilt:

$$\sum_{k=1}^{N} a_{\varphi(k)} \geq \alpha.$$

Addiere nun $M\in\mathbb{N}$ negative Reihenglieder, so dass für die Summe gilt:

$$\sum_{k=1}^{N+M} a_{\varphi(k)} \leq \alpha.$$

Wir iterieren dann diese Prozedur.

Da $\sum a_k$ bedingt konvergent ist, gilt $|a_k| \xrightarrow[k\to\infty]{} 0$. Daraus folgt, dass man mit den Approximationen $\sum_{k=1}^{\tilde{N}} a_{\varphi(k)}$ immer näher an α herankommt.

(ii) Wir betrachten zunächst den Fall $a_k\in\mathbb{R}$ für alle $k\in\mathbb{N}$: Da $\sum a_k$ konvergent ist, ist $(|a_k|)_k$ eine Nullfolge; insbesondere ist dann $(a_k^-)_k$ eine Nullfolge. Es existiert also ein $C\in\mathbb{R}$ mit

$$|a_k^-| \leq C \quad \forall\, k\in\mathbb{N}.$$

Da $\sum a_k^+$ divergent ist, existiert ein $N_1 \in \mathbb{N}$ mit

$$\sum_{k=1}^{N_1} a_k^+ > 2C.$$

Das addieren des ersten negativen Reihenglieds $a_1^- \in (a_k^-)_k$ liefert dann

$$\left(\sum_{k=1}^{N_1} a_k^+\right) + \alpha_1 > C.$$

Also existiert ein $N_2 \in \mathbb{N}$ mit

$$\sum_{k=1}^{N_1+N_2} a_k^+ > 4C.$$

Daraus folgt dann

$$\left(\sum_{k=1}^{N_1+N_2} a_k^+\right) + a_1^- + a_2^- = \sum_{k=1}^{M} a_{\varphi(k)} > 2C$$

für ein $M \leq N_1 + N_2 + 2$ und eine Bijektion $\varphi\colon \mathbb{N} \to \mathbb{N}$. Man erkennt dann, dass diese Umordnung divergiert.

Wir betrachten nun den Fall $a_k \in \mathbb{C}$ für alle $k \in \mathbb{N}$: Dabei gilt bekannterweise

$$|a_k| \leq |\mathrm{Re}(a_k)| + |\mathrm{Im}(a_k)|.$$

Wenn nun $\sum |a_k|$ divergent ist, so ist nach der Dreiecksungleichung mindestens einer der beiden Reihen $\sum |\mathrm{Re}(a_k)|$ bzw. $\sum |\mathrm{Im}(a_k)|$ divergent. Wir nehmen nun o.B.d.A. an, dass $\sum |\mathrm{Re}(a_k)|$ divergent ist. Dann existiert (nach der Argumentation im reellen Fall) eine Bijektion $\varphi\colon \mathbb{N} \to \mathbb{N}$ derart, dass $\sum \mathrm{Re}(a_{\varphi(k)})$ divergent ist. Dann ist $\sum a_{\varphi(k)}$ divergent, unabhängig davon was $\sum \mathrm{Im}(a_{\varphi(k)})$ macht.

(i) ($\Rightarrow$) Es sei $\varphi\colon \mathbb{N} \to \mathbb{N}$ eine beliebige Bijektion. Setze nun

$$T_n = \sum_{k=1}^{n} |a_{\varphi(k)}| \leq \sum_{k=1}^{\infty} |a_k| =: A < \infty.$$

Es gilt also

$$\lim_{n\to\infty} T_N \leq A,$$

weshalb $\sum a_{\varphi(k)}$ absolut konvergent und damit insbesondere konvergent ist.

($\Leftarrow$) Siehe Beweis von Teil (ii).

Bei Teil (i) ist noch zu zeigen, dass die Werte der Reihen $\sum a_k$ und $\sum a_{\varphi(k)}$ gleich sind. Dabei sehen wir, dass $\sum a_{\varphi(k)} \leq \sum a_k$, falls $a_k \geq 0$ für alle $k \in \mathbb{N}$. Ein Rollentausch (mit Anwendung der Umkehrabbildung φ^{-1}) liefert dann $\sum a_k \leq \sum a_{\varphi(k)}$ und damit die Gleichheit. Zu zeigen ist noch die Aussage für den allgemeinen Fall $a_k \in \mathbb{C}$ (Übung!). $\square$

3 Summen und Produkte von Reihen

Lemma 3.1. *Sind $a_k, b_k, \lambda \in \mathbb{C}$ für $k \in \mathbb{N}_0$ und $\sum a_k$ und $\sum b_k$ konvergente Reihen, so sind auch*

$$\sum (a_k + b_k) \quad \textit{sowie} \quad \sum \lambda a_k$$

konvergent und es gelten:

$$\sum_{k\in\mathbb{N}_0} (a_k + b_k) = \sum_{k\in\mathbb{N}_0} a_k + \sum_{k\in\mathbb{N}_0} b_k \quad \textit{und} \quad \sum_{k\in\mathbb{N}_0} \lambda a_k = \lambda \sum_{k\in\mathbb{N}_0} a_k.$$

Beweis. Anwendung der Grenzwertsätze für Folgen auf die Folgen der Partialsummen liefert direkt die Behauptung. □

Nun beschäftigen wir uns mit einer wesentlich schwierigeren Frage: Gilt für konvergente Reihen $s = \sum a_k$ und $t = \sum b_k$ auch

$$\sum a_k \cdot \sum b_k = s \cdot t?$$

Wie ist $\sum a_k \cdot \sum b_k$ überhaupt definiert? Vielleicht als $\sum_{k,l\in\mathbb{N}_0} a_k b_l$? In welcher Reihenfolge summiere ich über $(k, l) \in \mathbb{N}_0 \times \mathbb{N}_0$?

LINK: TEIL 1 DER 4. VORLESUNG VOM 21.04.2022

Wir setzen

$$s_m = \sum_{k=0}^{m} a_k \quad \text{und} \quad t_m = \sum_{k=0}^{m} b_k.$$

Offenbar gilt

$$\Big| s \cdot t - s_m t + s_m t - s_m \cdot t_m \Big| \le \underbrace{|s - s_m| \cdot |t|}_{\xrightarrow[m\to\infty]{} 0} + \underbrace{|s_m| \cdot \underbrace{|t - t_m| \cdot |t|}_{\xrightarrow[m\to\infty]{} 0}}_{\xrightarrow[m\to\infty]{} 0} \xrightarrow[m\to\infty]{} 0,$$

da (s_m) konvergent und damit beschränkt ist. Also ist

$$s \cdot t = \lim_{m\to\infty} \Big(s_m \cdot t_m \Big).$$

Außerdem schreiben wir

$$s_m \cdot t_m = \sum_{k=0}^{m} a_k \cdot \sum_{l=0}^{m} b_l = \sum_{k,l=0}^{m} a_k b_l =: \sum_{n=0}^{m} d_n,$$

wobei

$$d_n := \sum_{(k,l)\in\Gamma_n} a_k b_l \quad \text{mit} \quad \Gamma_n := \Big\{ (k, l) \in \mathbb{N}_0^2 \ : \ \max\{k, l\} = n \Big\}.$$

Es gilt also

$$s \cdot t = \sum_{n\in\mathbb{N}_0} d_n.$$

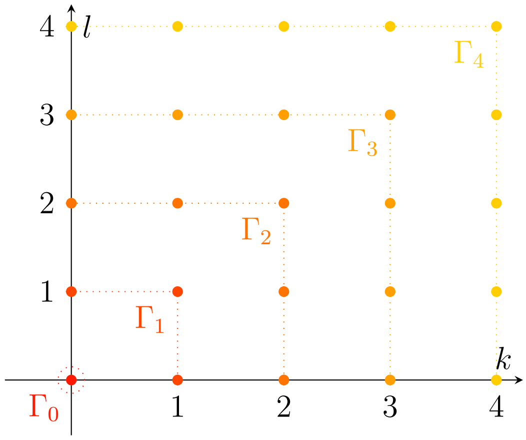

Abbildung 1: Visualisierung der Menge Γ_n für $n = 0, \ldots, 4$.

Alternativ kann man die Doppelsumme auch als unendliche Summe über die Diagonalen

$$\Delta_n := \Big\{(k,l) \in \mathbb{N}_0^2 \;:\; k+l = n\Big\} = \Big\{(j, n-j) \in \mathbb{N}_0^2 \;:\; 0 \le j \le n\Big\}.$$

schreiben.

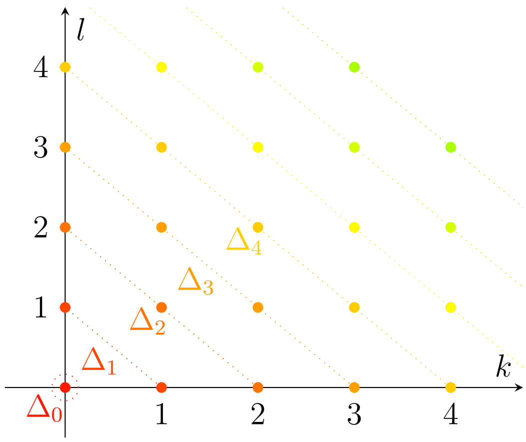

Abbildung 2: Visualisierung der Menge Δ_n für $n = 0, \ldots, 4$.

Damit nun für die Diagonalsummen

$$c_n := \sum_{k+l=n} a_k \cdot b_l = \sum_{j=0}^{n} a_j \cdot b_{n-j}, \quad n \in \mathbb{N},$$

die Gleichheit

$$s \cdot t = \sum_{n \in \mathbb{N}_0} c_n$$

erfüllt ist, muss

$$\sum_{n \in \mathbb{N}_0} c_n = \sum_{n \in \mathbb{N}_0} d_n$$

gelten. Dies gilt genau dann, wenn die Differenzen

$$r_m := \sum_{n=0}^{m} c_n - s_m \cdot t_m = \sum_{(k,l)\in D_m} a_k b_l \tag{Rst}$$

gegen 0 streben, wobei

$$D_m := \left\{(k,l) \in \mathbb{N}_0^2 \;:\; 0 \le k,l \le m, \quad k+l > m\right\} \quad \text{und} \quad D_m^+ := \bigcup_{n=m+1}^{2m} \Delta_n \supset D_m.$$

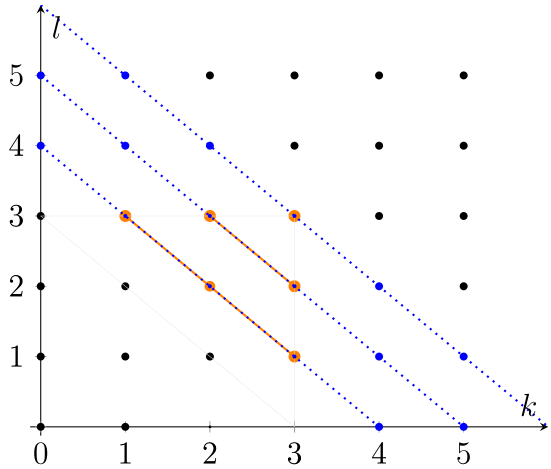

Abbildung 3: Visualisierung der Menge D_m (in orange) und D_m^+ (in blau) für $m = 3$. Man bemerke, dass tatsächlich $D_3 \subset D_3^+$ gilt.

Inklusion ergibt dann die Fehlerabschätzung

$$|r_m| \le \sum_{(k,l)\in D_m} |a_k b_l| \le \sum_{n=m+1}^{2m} c_n^*$$

mit

$$c_n^* := \sum_{(k,l)\in \Delta_n} |a_k b_l| = \sum_{j=0}^{n} |a_j \cdot b_{n-j}|.$$

Dies motiviert den folgenden Satz:

Satz 3.2 (Cauchyprodukt von Reihen)**.** *Es seien $a_k, b_k \in \mathbb{C}$ für alle $k \in \mathbb{N}_0$ und die Reihen $\sum a_k$ sowie $\sum b_k$ absolut konvergent mit Werten $s \in \mathbb{C}$ bzw. $t \in \mathbb{C}$. Dann sind auch die Reihen*

$$\sum c_n \quad \textit{und} \quad \sum c_n^*$$

absolut konvergent und es gilt

$$\sum_{n\in\mathbb{N}_0} c_n = \sum_{n\in\mathbb{N}_0} \sum_{j=0}^{n} (a_j \cdot b_{n-j}) = \left(\sum_{k\in\mathbb{N}_0} a_k\right) \cdot \left(\sum_{k\in\mathbb{N}_0} b_k\right) = s \cdot t. \tag{CP}$$

Damit haben wir eine Analogie zum Distributivgesetz für abzählbar unendliche Summen.

Beweis. Für alle $m \in \mathbb{N}_0$ gilt

$$\sum_{n=0}^{m} |c_n| = \sum_{n=0}^{m} \left| \sum_{j=0}^{n} a_j b_{n-j} \right| \leq \sum_{n=0}^{m} \underbrace{\sum_{j=0}^{n} |a_j b_{n-j}|}_{=c_n^*}.$$

Also ist $\sum c_n^*$ eine Majorante von $\sum |c_n|$. Nun gilt außerdem

$$\sum_{n=0}^{m} c_n^* \leq \sum_{k,l=0}^{m} |a_k b_l| = \left(\sum_{k=0}^{m} |a_k| \right) \cdot \left(\sum_{l=0}^{m} |b_l| \right) \leq \underbrace{\sum_{k \in \mathbb{N}_0} |a_k| \cdot \sum_{l \in \mathbb{N}_0} |b_l|}_{\text{absolut konvergent nach Annahme}}.$$

Also ist $\sum c_n^*$ absolut konvergent und nach dem Majorantenkriterium aus Satz 2.5 dann auch $\sum c_n$. Daraus folgt

$$|r_m| \overset{\text{(Rst)}}{\leq} \sum_{n=m+1}^{2m} c_n^* \xrightarrow[m \to \infty]{\text{Cauchy-Krit.}} 0.$$

Zusammen mit der obigen Vorüberlegung erhält man dann

$$\sum_{n \in \mathbb{N}_0} c_n = \sum_{n \in \mathbb{N}_0} \sum_{j=0}^{n} (a_j b_{n-j}) = s \cdot t.$$ □

Eine abgeschwächte Version des vorigen Satzes liefert der folgende Satz:

Satz 3.3 (von Martens)**.** *Ist $\sum a_k$ absolut konvergent und $\sum b_k$ konvergent, so konvergiert auch $\sum c_k$ und es gilt* (CP)*.*

(Ohne Beweis)

Beispiel 3.4.

(a) Für $z \in \mathbb{C}$, $|z| < 1$, konvergiert die Reihe

$$\sum_{k \in \mathbb{N}_0} z^k = \frac{1}{1-z}$$

absolut. Daraus folgt

$$\left(\sum_{k \in \mathbb{N}_0} z^k \right)^2 = \sum_{n \in \mathbb{N}_0} c_n = \left(\frac{1}{1-z} \right)^2$$

mit

$$c_n := \sum_{j=0}^{n} \underbrace{z^j \cdot z^{n-j}}_{=z^n} = (n+1) \cdot z^n,$$

also

$$\sum_{n \in \mathbb{N}_0} (n+1) z^n = \frac{1}{(1-z)^2} \quad \text{für } |z| < 1.$$

(b) Ist $a_k = b_k = \frac{(-1)^k}{\sqrt{k+1}}$, so sind die Reihen $\sum a_k = \sum b_k$:

- konvergent nach dem Leibnitz-Kriterium 1.7, aber
- nicht absolut konvergent nach dem Minorantenkriterium aus Satz 2.5, da

$$\frac{1}{\sqrt{k+1}} \geq \frac{1}{k+1} \quad \text{und} \quad \sum \frac{1}{k+1} \text{ divergiert.}$$

Wir berechnen nun

$$c_n = \sum_{j=0}^{n} a_j b_{n-j} = \sum_{j=0}^{n} \frac{(-1)^j}{\sqrt{j+1}} \cdot \frac{(-1)^{n-j}}{\sqrt{n-j+1}} = (-1)^n \sum_{j=0}^{n} \Big(\underbrace{(j+1)(n-j+1)}_{=\left(\frac{n}{2}+1\right)^2 - \left(\frac{n}{2}-j\right)^2 \leq \left(\frac{n}{2}+1\right)^2} \Big)^{-\frac{1}{2}}.$$

Es gilt also

$$c_n \leq (-1)^n \sum_{j=0}^{n} \left(\left(\frac{n}{2} + 1 \right)^2 \right)^{-\frac{1}{2}} = (-1)^n \sum_{j=0}^{n} \left(\frac{n}{2} + 1 \right)$$

und es folgt

$$|c_n| \geq \sum_{j=0}^{n} \frac{1}{\frac{n}{2}+1} = 2 \sum_{j=0}^{n} \frac{1}{n+2} = 2 \cdot \frac{n+1}{n+2} \xrightarrow[n\to\infty]{} 2.$$

Damit ist c_n keine Nullfolge, weshalb $\sum c_n$ divergiert. Für bedingt konvergente Reihen $\sum a_k$ und $\sum b_k$ gilt daher (CP) im Allgemeinen nicht!

LINK: TEIL 2 DER 4. VORLESUNG VOM 21.04.2022

4 Dezimalbruchentwicklung

Es sei $a_0 \in \mathbb{Z}$ und $a_k \in \{1, 2, \ldots, 9\}$ für $k \in \mathbb{N}$.

Behauptung 4.1. Die Reihe

$$\sum_{k\in\mathbb{N}_0} \frac{a_k}{10^k} = a_0 + \frac{a_1}{10} + \frac{a_2}{100} + \ldots$$

konvergiert absolut gegen ein $s \in \mathbb{R}$.

Beweis. Nach Voraussetzung gilt $\frac{a_k}{10^k} \geq 0$ für alle $k \in \mathbb{N}_0$, weshalb die Folge $(s_n)_{n\in\mathbb{N}_0}$,

$$s_n = \sum_{k=0}^{n} \frac{a_k}{10^k} \quad \forall n \in \mathbb{N}_0,$$

der n-ten Partialsummen der Reihe isoton ist – es gilt also $s_n \geq 0$ für alle $n \in \mathbb{N}_0$. Nun gilt außerdem

$$s_n \leq a_0 + 0 \cdot \underbrace{\sum_{k=1}^{\infty} \frac{1}{10^k}}_{=-1+\sum_{k\in\mathbb{N}_0} 10^{-k}} = a_0 + \frac{9}{1-\frac{1}{10}} - 9 = a_0 + 10 - 9 = a_0 + 1.$$

(s_n) ist also von oben beschränkt und damit nach Satz 6.3 aus der Vorlesung zur Analysis I (Lehramt) konvergent. □

Den Wert der Reihe notieren wir als

$$s = a_0, a_1 a_2 a_3 \dots$$

und nennen jeden solchen Ausdruck *Dezimalbruch*.

Damit erhalten wir eine Abbildung

$$\Phi \colon \Big\{(a_0, a_1, a_2, \dots) \;:\; a_0 \in \mathbb{Z}, a_k \in \{1, \dots, 9\} \text{ für } k \in \mathbb{N}\Big\} \to \mathbb{R}.$$

Man bemerke, dass Φ <u>nicht</u> injektiv ist, denn es gilt

$$\Phi(1,0,0,0,\dots) = 1,0000\dots = 1 = 10 - 9 = \sum_{k \in \mathbb{N}_0} \frac{9}{10^k} - 9 = \Phi(0,9,9,9,\dots),$$

obwohl $(1,0,0,0\dots) \neq (0,9,9,9,\dots)$.

Definition 4.2. Ein Dezimalbruch heißt *periodisch*, falls es ein $m \in \mathbb{N}_0$ und ein $p \in \mathbb{N}$ gibt mit

$$a_{m+k+p} = a_{m+k} \quad \forall\, k \in \mathbb{N}_0.$$

Dann schreibt man

$$a_0, a_1 a_2 \dots a_m \overline{a_{m+1} a_{m+2} \dots a_{m+p}}.$$

Ist sogar $p = 1$ und $a_m = 9$, also

$$s = a_0, a_1 a_2 \dots a_{m-1} \overline{9},$$

so heißt der Dezimalbruch *uneigentlich*. Ansonsten heißt der Dezimalbruch *eigentlich*: Für alle $m \in \mathbb{N}$ existiert dann ein $k > m$ mit $a_k \neq 9$.

Behauptung 4.3.

(i) Φ ist surjektiv;

(ii) Φ ist injektiv, falls wir es auf die eigentlichen Dezimalbrüche einschränken.

Beweis.

(i) Wir müssen zu einem gegebenen $s \in \mathbb{R}$ eine Dezimalbruchentwicklung finden!

Für $s = 0$ können wir direkt $a_k = 0$ für alle $k \in \mathbb{N}_0$ setzen. Für $s \neq 0$ verläuft der Beweis für $s > 0$ analog zu dem für $s < 0$. Es gelte also o.B.d.A $s > 0$. Wir setzen zunächst

$$a_0 := \lfloor s \rfloor := \max\Big\{k \in \mathbb{Z} \;:\; k \leq s\Big\}.$$

Wir erkennen sofort, dass $a_0 \leq s < a_0 + 1$ gilt. Wir setzen dann

$$a_1 := \lfloor 10(s - a_0) \rfloor$$

und sehen, dass
$$a_1 \leq 10(s - a_0) < a_1 + 1.$$
Es gilt also
$$0 \leq a_1 < 10 \quad \Rightarrow a_1 \in \{1, \dots, 9\}$$
sowie
$$a_0 + \frac{a_1}{10} \leq s < a_0 + \frac{a_1}{10} + \frac{1}{10}.$$
Durch induktive Fortsetzung können wir dann $a_0, a_1, \dots a_n$ bestimmen. Setze nun
$$a_{n+1} := \left\lfloor 10^{n+1}\left(s - \sum_{k=0}^{n} \frac{a_k}{10^k}\right)\right\rfloor.$$
Dann gilt
$$a_{n+1} \in \{1, \dots, 9\} \quad \text{und} \quad \sum_{k=0}^{n} \frac{a_k}{10^k} \leq s < \sum_{k=0}^{n} \frac{a_k}{10^k} + \frac{1}{10^n}$$
sowie
$$0 \leq s - \sum_{k=0}^{n} \frac{a_k}{10^k} < \frac{1}{10^n} \xrightarrow[n\to\infty]{} 0,$$
also $\sum_{k\in\mathbb{N}_0} \frac{a_k}{10^k} = s$. Damit existiert das Urbild, weshalb Φ surjektiv ist.

(ii) Widerspruchsannahme: Wir nehmen an, dass zwei unterschiedliche eigentliche Dezimalbrüche existieren, die den denselben Wert $s \in \mathbb{R}$ besitzen. Es gilt also
$$a_0, a_1 a_2 a_3 \dots = s = b_0, b_1 b_2 b_3 \dots$$
mit einem kleinsten $m \in \mathbb{N}_0$ mit $a_m \neq b_m$. O.B.d.A sei dann
$$a_m < b_m \quad \big(\Leftrightarrow \quad 1 + a_m \leq b_m\big).$$
Dann gilt
$$\begin{aligned} s &= \sum_{k=0}^{m} \frac{a_k}{10^k} + \sum_{k=m+1}^{\infty} \frac{a_k}{10^k} \\ \texttt{[eigentlich]} &< \sum_{k=0}^{m} \frac{a_k}{10^k} + \sum_{k=m+1}^{\infty} \frac{9}{10^k} \\ \texttt{[geometrische Reihe]} &= \sum_{k=0}^{m} \frac{a_k}{10^k} + \frac{1}{10^m} \\ &= \sum_{k=0}^{m-1} \frac{a_k}{10^k} + \frac{a_m + 1}{10^m} \end{aligned}$$

$$\overset{[\texttt{m minimal}]}{\leq} \sum_{k=0}^{m-1} \frac{b_k}{10^k} + \frac{b_m}{10^m}$$

$$\leq \sum_{k=0}^{\infty} \frac{b_k}{10^k} = s.$$

Es gilt also $s < s$, was ein Widerspruch ist. Also muss Φ injektiv sein, wenn wir eigentliche Dezimalbrüche betrachten. □

Damit haben wir eine weitere Beschreibung für rationale Zahlen erhalten: Es gilt $s \in \mathbb{Q}$ genau dann, wenn die eigentliche Dezimalbruchdarstellung von s abbricht – es existiert also ein $N \in \mathbb{N}$, so dass $a_n = 0$ für $n > N$,– oder s periodisch gemäß Definition 4.2 ist.

Wir können analog reelle Zahlen zu jeder Basis $b \in \mathbb{N}$, $b \geq 2$, entwickeln. Die dann auftauchenden Ziffern sind liegen dann in der Menge $\{0, 1, \ldots, b-1\}$. So existiert beispielsweise die Binärdarstellung von $s \in [0, 1]$:

$$s = \sum_{k \in \mathbb{N}_0} \frac{a_k}{2^k} \quad a_k \in \{0, 1\}$$

$$\overset{\text{z.B.}}{=} 0,11010100010\ldots$$

Dabei erfolgt die Unterscheidung, ob eine Entwicklung eigentlich oder uneigentlich ist, analog zum Fall $b = 10$.

Tritt die Ziffer $b - 1$ mit der Periode Eins auf?

LINK: TEIL 1 DER 5. VORLESUNG VOM 27.04.2022

5 Funktionenfolgen

Definition 5.1.

(a) Es sei M eine nichtleere Menge und

$$\mathcal{F}(M, \mathbb{C}) = \mathcal{F}(M) := \left\{ f \colon M \to \mathbb{C} \right\}$$
$$\mathcal{F}(M, \mathbb{R}) = \mathcal{F}_{\mathbb{R}}(M) := \left\{ f \colon M \to \mathbb{R} \right\}$$

die Menge der komplexwertigen / reellwertigen Funktionen auf M – insbesondere gilt $\mathcal{F}_{\mathbb{R}}(M) \subset \mathcal{F}(M)$.

Eine *Funktionenfolge* ist eine Abbildung

$$\varphi \colon \mathbb{N}_0 \to \mathcal{F}_{\mathbb{R}}(M),$$

d.h. für jedes $n \in \mathbb{N}_0$ ist $f_n := \varphi(n)$ eine Funktion $M \to \begin{matrix}\mathbb{C}\\ \mathbb{R}\end{matrix}$ und für jedes $x \in M$ ist $n \mapsto f_n(x)$ eine Zahlenfolge in $\begin{matrix}\mathbb{C}\\ \mathbb{R}\end{matrix}$. Häufig ist $M \subset \mathbb{C}$ oder $M \subset \mathbb{R}$, z.B. ein Intervall.

(b) Es sei $(f_n)_n$ eine Funktionenfolge in $\mathcal{F}(M)$. Wir sagen, dass $(f_n)_N$ *punktweise gegen eine Funktion* $f\colon M \to \mathbb{C}$ *konvergiert*, falls

$$\lim_{n\to\infty} \left| f_n(x) - f(x) \right| = 0 \quad \forall\, x \in M.$$

Die Folge $(f_n)_n$ heißt dann *punktweise konvergent gegen* f und f heißt *Grenzfunktion* oder *Limes* der Folge $(f_n)_n$. Wir schreiben dann

$$f := \lim_{n\to\infty} f_n \quad \text{oder} \quad f_n \xrightarrow[n\to\infty]{\text{pktw}} f.$$

Die Folge $(f_n)_n$ *konvergiert uniform bzw. gleichmäßig gegen* $f\colon M \to \mathbb{C}$, falls

$$\lim_{n\to\infty} \underbrace{\sup_{x\in M} \left| f_n(x) - f(x) \right|}_{=\|f_n - f\|_M = \|f_n - f\|} = 0$$

gilt. In diesem Fall schreiben wir

$$f_n \xrightarrow[n\to\infty]{\text{glm}} f \quad \text{und} \quad \lim_{n\to\infty} \|f_n - f\| = 0.$$

Gilt sogar $\sum_n \underbrace{\|f_n\|}_{\in\mathbb{R}} < \infty$, so heißt $\sum_n f_n$ *normal konvergent.*

Beispiel 5.2. Es sei $M := \overline{B_1} = \left\{ z \in \mathbb{C} \;:\; |z| \le 1 \right\}$ und

$$f_n\colon M \to \mathbb{C}, \quad z \mapsto z^n.$$

Für alle $z \in B_1 = \left\{ z \in \mathbb{C} \;:\; |z| < 1 \right\}$ gilt dann

$$f_n(z) = z^n \xrightarrow[n\to\infty]{} 0 =: f(z),$$

da $\varphi^n := |z|^n \xrightarrow[n\to\infty]{} 0$ wegen $\varphi < 1$. Dagegen gilt $f_n(1) = 1^n = 1 =: f(1)$. Also konvergiert $(f_n)_n$ auf der Menge $B_1 \cup \{z = 1\}$ punktweise gegen die Grenzfunktion

$$f = \begin{cases} 0 & \text{auf } B_1, \\ 1 & \text{auf } \{z = 1\}. \end{cases}$$

Was passiert denn in $z \in \mathbb{C}$, $|z| = 1$, aber $z \neq 1$? Für $\psi \in (0, 2\pi)$ gilt dann $z = \mathrm{e}^{\mathrm{i}\psi}$ und

$$f_n(z) = \left(\mathrm{e}^{\mathrm{i}\psi}\right)^n = \mathrm{e}^{\mathrm{i}n\psi}.$$

Man erkennt, dass $f_n(z)$ die Position am Rand des Einheitskreises darstellt zum Winkel $n \cdot \psi$ darstellt. Mit steigendem n wird dann jedes Mal der Winkel ψ hinzugefügt und die Position am Einheitskreis entsprechend angepasst. Aufgrund der Periodizität des Einheitskreises wird dann deutlich, dass f_n nicht für $n \to \infty$ konvergiert. So gilt für $\psi = \pi$ beispielsweise $z = \mathrm{e}^{\mathrm{i}\pi} = -1$ und damit

$$f_n(-1) = (-1)^n.$$

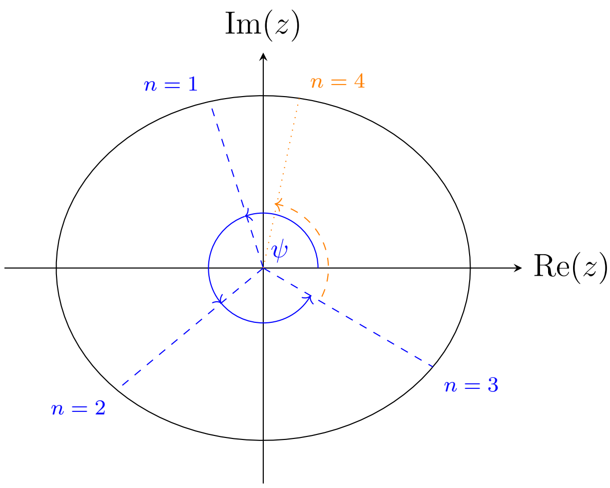

Abbildung 4: Illustration der Funktionenfolge $f_n(z) = \mathrm{e}^{\mathrm{i}n\psi}$ für $\psi = \frac{11}{18}\pi$ und $n = 1, 2, 3$ in blau sowie $n = 4$ in orange. Man bemerke, dass für steigende $n \in \mathbb{N}$ stets der Winkel ψ hinzu addiert und dann der Funktionswert ausgewertet wird. Für $n = 4$ gilt insbesondere $\mathrm{e}^{\mathrm{i}\cdot 4\psi} = \mathrm{e}^{\mathrm{i}\cdot\left(2+\frac{4}{9}\right)\pi} = \mathrm{e}^{\mathrm{i}\cdot\frac{4}{9}\pi}$ aufgrund der 2π-Periodizität. Es wird deutlich, dass $(f_n)_n$ nicht für $n \to \infty$ konvergiert.

Wir untersuchen nun, wann wir sogar gleichmäßige Konvergenz vorliegen haben. Für jedes $\rho \in (0, 1)$ gilt dabei

$$\sup_{|z| \leq \rho} |f_n(z) - f(z)| = \sup_{|z| \leq \rho} |z^n| = \rho^n \xrightarrow[n\to\infty]{} 0,$$

d.h. auf $\overline{B_\rho} = \left\{z \in \mathbb{C} \ : \ |z| \leq \rho\right\}$ konvergiert $(f_n)_n$ sogar gleichmäßig gegen f.

Wir betrachten nun die Einschränkungen $g_n := f_n\big|_I$, $g := f\big|_I$ für $I = [-1, 1] \subset \mathbb{R} \subset \mathbb{C}$. Dann gilt

$$\overbrace{g_n}^{\text{stetig}} \xrightarrow[n\to\infty]{\text{pktw}} \overbrace{g = \mathbb{1}_{\{1\}}}^{\text{unstetig}} \quad \text{auf } (-1, 1] \quad \text{und} \quad g_n \xrightarrow[n\to\infty]{\text{glm}} g \quad \text{auf } [-\rho, \rho], \rho \in (0, 1).$$

$g_n(-1) = (-1)^n$ konvergiert dagegen nicht.

Lemma 5.3 (Äquivalente Formulierungen). *Es seien $f_n \colon M \to \mathbb{C}$ Funktionen für $n \in \mathbb{N}$. Dann sind folgende Aussagen äquivalent:*

(i) *Es gilt*
$$\|f - f_n\| \xrightarrow[n\to\infty]{} 0$$
für ein $f \colon M \to \mathbb{C}$;

(ii) *Es existiert ein $f \colon M \to \mathbb{C}$, so dass für alle $\varepsilon > 0$ ein $n_0 = n_0(\varepsilon) \in \mathbb{N}$ (unabhängig von x) existiert mit*
$$|f_n(x) - f(x)| < \varepsilon \quad \forall\, x \in M, n \geq n_0.$$
Falls ein $n_0 = n_0(\varepsilon, x)$ existiert, das von x abhängig ist und die obige Eigenschaft erfüllt, so haben wir nur punktweise Konvergenz vorliegen.

(iii) *Cauchy-Kriterium:*
Für alle $\varepsilon > 0$ existiert ein $n_0 = n_0(\varepsilon) \in \mathbb{N}$, so dass
$$|f_m(x) - f_n(x)| < \varepsilon \quad \forall\, n, m \geq n_0, x \in M.$$

Beweis.

(i) $\Rightarrow$ (ii) Es sei $\varepsilon > 0$ beliebig. (i) impliziert, dass für n_0 groß genug gilt:
$$\varepsilon > \|f - f_n\|_M = \underbrace{\sup_{x\in M} |f(x) - f_n(x)|}_{n_0 \text{ unabhängig von } x, \text{ da uniform}} \geq |f(y) - f_n(y)|$$
für jedes $y \in M$.

(ii) $\Rightarrow$ (iii) Es sei $\varepsilon > 0$ beliebig. Dann existiert nach (ii) ein $n_0(\varepsilon)$, so dass
$$|f(x) - f_n(x)| < \frac{\varepsilon}{2} \quad \forall x \in M, n \geq n_0.$$
Insbesondere gilt dann
$$|f_m(x) - f(x) + f(x) - f_n(x)| \leq |f_m(x) - f(x)| + |f(x) - f_n(x)| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

LINK: TEIL 2 DER 5. VORLESUNG VOM 27.04.2022

(iii) $\Rightarrow$ (i) Für jedes $x \in M$ ist $\big(f_n(x)\big)_n$ eine $\mathbb{C}$-Cauchyfolge, es existiert also der Grenzwert
$$\lim_{n\to\infty} f_n(x) =: f(x).$$
Sei nun $\varepsilon > 0$ beliebig. Dann existiert nach Teil (iii) ein $n_0(\varepsilon)$, so dass für alle $m, n > n_0$ und $x \in M$ gilt:
$$\varepsilon > |f_m(x) - f_n(x)|,$$

also

$$\varepsilon \geq \lim_{m\to\infty} |f_m(x) - f_n(x)|$$

und insbesondere

$$\varepsilon \geq \sup_{x\in M} \lim_{m\to\infty} |f_m(x) - f_n(x)| = \sup_{x\in M} |f(x) - f_n(x)| = \|f - f_n\|_M. \qquad \square$$

Insbesondere haben wir Folgendes festgestellt:

$$f_n \xrightarrow[n\to\infty]{\text{glm}} f \quad \Rightarrow \quad f_n \xrightarrow[n\to\infty]{\text{ptkw}} f.$$

Hier kann man nämlich $n_0(\varepsilon, x) = n_0(\varepsilon)$ wählen. Für die Umkehrung gilt dagegen

$$f_n \xrightarrow[n\to\infty]{\text{pktw}} f \quad \not\Rightarrow \quad f_n \xrightarrow[n\to\infty]{\text{glm}} f.$$

Hier kann man nicht „$n_0(\varepsilon) = n_0(\varepsilon, x)$" wählen, da die Abhängigkeit von x stört.

Beispiel 5.4.

(a) Es sei g_n wie in Beispiel 5.2. Bekannterweise ist

$$g_n \xrightarrow[n\to\infty]{\text{pktw}} g = \mathbb{1}_{\{1\}},$$

aber es gilt außerdem

$$\sup_{0\in[-1,1]} |g_n(x) - g(x)| \geq \sup_{x\in(-1,1)} |x^n - 0| \geq |x_n^n|$$

für eine beliebige Folge $(x_n)_n$ in $(-1,1)$. Wähle $x_n = \sqrt[n]{\frac{1}{2}}$ und es folgt

$$|x_n^n| = \left|\frac{1}{2}\right| > 0,$$

wir haben also keine gleichmäßige Konvergenz.

(b) Es sei

$$f_n \colon [0,1] \to \mathbb{R}, \quad x \mapsto f_n(x) := \begin{cases} 2nx & \text{für } 0 \leq x \leq \frac{1}{2n}, \\ 2 - 2nx & \text{für } \frac{1}{2n} < x < \frac{1}{n}, \\ 0 & \text{für } \frac{1}{n} \leq x \leq 1. \end{cases}$$

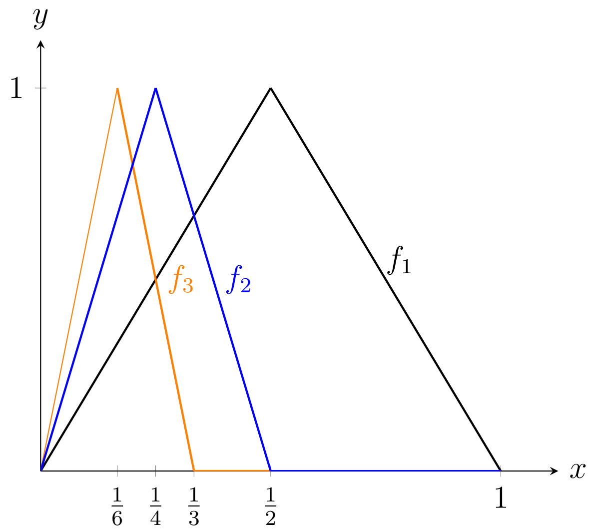

Abbildung 5: Der Graph von f_n für $n = 1$ (in schwarz), $n = 2$ (in blau) und $n = 3$ (in orange).

Ein Kandidat für die Grenzfunktion wäre $f \equiv 0$. Es gilt jedoch $f_n\left(\frac{1}{2n}\right) = 1$, also

$$\|f_n - f\|_{[0,1]} \geq \left| f_n\left(\frac{1}{2n}\right) - 0 \right| = 1 \not\to 0 \quad \text{für } n \to \infty.$$

Dagegen ist $f(0) = f_n(0)$ für alle $n \in \mathbb{N}$ und für jedes $\delta \in (0, 1]$ gilt:

$$\text{Es gibt ein } n_0 \in \mathbb{N} \text{ mit } \frac{1}{n_0} < \delta.$$

Für $n \geq n_0$ folgt dann

$$f_n(\delta) = 0 = f(\delta) \quad \Rightarrow \quad f_n(\delta) \xrightarrow[n\to\infty]{\text{pktw}} f(\delta).$$

Tatsächlich ist sogar

$$\sup_{\delta \leq x \leq 1} |f_n(x) - f(x)| = |0 - 0| = 0, \quad \text{falls } n \geq n_0 \text{ wie oben},$$

es gilt also $f_n \xrightarrow[n\to\infty]{\text{glm}} f$ auf einem Teilintervall $[\delta, 1]$.

(c) Für $n \in \mathbb{N}$ sei

$$f_n \colon \mathbb{R} \to \mathbb{R}, \quad x \mapsto f_n(x) := \frac{x + x^2 + n^2x^5}{1 + n^2x^4}.$$

Dann gilt

$$f(x) = \frac{\frac{x}{n^2} + \frac{x^2}{n^2} + x^5}{\frac{1}{n^2} + x^4} \xrightarrow[n\to\infty]{x \text{ fest}, x\neq 0} \frac{x^5}{x^4} = x.$$

Des Weiteren ist $f_n(0) = \frac{0}{1} = 0$, es gilt also

$$f_n(x) \xrightarrow[n\to\infty]{\text{pktw}} f(x) := x \quad \forall\, x \in \mathbb{R}.$$

Gilt sogar gleichmäßige Konvergenz? Dazu betrachten wir

$$g_n(x) := f_n(x) - f(x) = \frac{x + x^2 + n^2x^5}{1 + n^2x^4} - x = \frac{\overbrace{x^2}^{\geq 0}}{\underbrace{1 + n^2x^4}_{\geq 0}} \geq 0.$$

Für n fixiert ist dann $\lim_{|x|\to\infty} g_n(x) = 0$. Aus der Stetigkeit von g_n folgt dann, dass

$$c_n := \max_{x\in\mathbb{R}} g_n(x)$$

existiert. Um Extremalpunkte zu bestimmen, setzen wir

$$0 \overset{!}{=} g_n'(x) = \frac{2x(1 + n^2 + x^4) - x^2(4n^2x^3)}{(1 + n^2x^4)^2} = \frac{2x - 2n^2x^5}{\underbrace{(1 + n^2x^4)^2}_{>0}}$$

$$\Leftrightarrow 2x(1 - n^2x^4) = 0$$

$$\Leftrightarrow x = 0 \quad \text{oder} \quad x = \pm\sqrt{\frac{1}{n}}.$$

Einsetzen liefert dann

$$g_n(0) = 0 \quad \text{und} \quad g_n\left(\pm\frac{1}{\sqrt{n}}\right) = \frac{1}{2n}.$$

Damit ist $c_n = \frac{1}{2n}$, also $c_n \xrightarrow[n\to\infty]{} 0$. Das zeigt:

$$\|f_n - f\| \xrightarrow[n\to\infty]{} 0, \text{ also } \quad f_n \xrightarrow[n\to\infty]{\text{glm}} f \quad \text{auf } \mathbb{R}.$$

(d) Für $n \in \mathbb{N}$ sei

$$f_n\colon [0,\infty) \to \mathbb{R}, \quad x \mapsto f_n(x) := nx\mathrm{e}^{-nx}.$$

Dann ist $f(0) = 0$ und

$$f_n(x) \xrightarrow[n\to\infty]{} 0 \quad \text{für } x > 0.$$

Es gilt also

$$f_n \xrightarrow[n\to\infty]{\text{pktw}} f \equiv 0 \quad \text{auf } [0,\infty).$$

Gilt sogar gleichmäßige Konvergenz? Dazu betrachten wir

$$g_n(x) := f_n(x) - f(x) = f_n(x) = nx\mathrm{e}^{-nx} \geq 0.$$

Da $g_n(x) \underset{x\to\infty}{0}$ für $n \in \mathbb{N}$ fest gilt und g_n stetig ist, existiert

$$c_n := \max_{x\in[0,\infty)} g_n(x).$$

Zur Bestimmung setzen wir

$$0 \overset{!}{=} g_n'(x) = n\mathrm{e}^{-nx} - n^2 x\mathrm{e}^{-nx} = (1-nx)\underbrace{n\mathrm{e}^{-nx}}_{>0}$$

$$\Leftrightarrow 1 = nx$$

$$\Leftrightarrow x = \frac{1}{n} \in [0,\infty).$$

Wegen $g_n\left(\frac{1}{n}\right) = 1 \cdot \mathrm{e}^{-1} = \frac{1}{\mathrm{e}}$ gilt dann

$$\|f_n - f\|_{[0,\infty)} \geq g_n\left(\frac{1}{n}\right) = \frac{1}{\mathrm{e}} > 0,$$

wir haben also keine gleichmäßige Konvergenz vorliegen.

LINK: TEIL 1 DER 6. VORLESUNG VOM 28.04.2022

Satz 5.5. *Es sei $M \subset \mathbb{C}$ und $y \in M$. Sei weiter $(f_n)_{n\in\mathbb{N}} \subset \mathcal{F}(M)$ mit $f_n \xrightarrow[n\to\infty]{\text{glm.}} f$. Ist dann f_n stetig im Punkt y für alle $n \in \mathbb{N}$, so ist auch f stetig in y.*

Im Video: **Satz 5.6.**

Beweis. Es sei $\varepsilon > 0$ beliebig. Aufgrund der gleichmäßigen Konvergenz von f_n existiert ein $n_0 \in \mathbb{N}$, so dass

$$\sup_{x\in M} |f_n(x) - f(x)| < \frac{\varepsilon}{3} \quad \forall\, n \geq n_0.$$

Da insbesondere f_{n_0} stetig ist im Punkt y, existiert ein $\delta > 0$, so dass

$$|f_{n_0}(x) - f_{n_0}(y)| < \frac{\varepsilon}{3} \quad \forall\, x \in M \cap B_\delta(y).$$

Für diese x gilt dann

$$|f(x) - f(y)| = |f(x) - f_{n_0}(x) + f_{n_0}(x) - f_{n_0}(y) + f_{n_0}(y) - f(y)|$$

$$\leq |f(x) - f_{n_0}(x)| + |f_{n_0}(x) - f_{n_0}(y)| + |f_{n_0}(y) - f(y)|$$

$$< \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon.$$

Also ist f stetig im Punkt y. $\square$

Aber warum musste man für den Satz annehmen, dass alle f_n stetig im Punkt $y \in M$ sind? Reicht diese Annahme nicht für f_{n_0} aus?

Beispiele zeigen, dass die gleichmäßige Konvergenz nicht *notwendig* für die Stetigkeit der Grenzfunktion ist:

Der gleichmäßige Limes stetiger Funktionen ist ebenfalls stetig.

Der Satz besagt, wann man zwei Grenzübergänge vertauschen darf – solange mindestens einer gleichmäßig ist –, es gilt nämlich

$$\lim_{x\to y} f(x) = \lim_{x\to y}\Big(\overbrace{\lim_{n\to\infty}}^{\text{glm}} f_n(x)\Big) = \overbrace{\lim_{n\to\infty}}^{\text{glm}}\Big(\lim_{x\to y} f_n(x)\Big) = \lim_{n\to\infty} f_n(y) = f(y).$$

Insbesondere gilt: Falls f_n stetig ist auf M für alle $n \in \mathbb{N}$ und $f_n \xrightarrow[n\to\infty]{\text{glm.}} f$, so ist auch f stetig auf M.

Nun stellt sich eine weitere Frage: Überträgt sich bei gleichmäßiger Konvergenz einer Funktionenfolge auch die Integrierbarkeit und Differenzierbarkeit von f_n auf ihre Grenzfunktion?

Satz 5.6. *Es sei eine Folge $f_n\colon [a,b] \to \mathbb{R}$ gleichmäßig konvergent gegen eine Funktion $f\colon [a,b] \to \mathbb{R}$. Für jedes $n \in \mathbb{N}$ sei weiter f_n integrierbar auf $[a,b]$. Dann ist auch f integrierbar auf $[a,b]$ und es gilt*

Im Video: **Satz 5.7**.

$$\lim_{n\to\infty} \int_a^b f_n(x)\,\mathrm{d}x = \int_a^b \underbrace{\lim_{n\to\infty}}_{\text{glm}} f_n(x)\,\mathrm{d}x = \int_a^b f(x)\,\mathrm{d}x;$$

das Integralsymbol $\int$ sowie der Limes $\lim_{n\to\infty}$ sind also vertauschbar. Man beachte dabei, dass auch $\int$ für einen Grenzübergang steht.

Beweis. Es sei $\varepsilon > 0$ beliebig. Wegen der gleichmäßigen Konvergenz von (f_n) existiert ein $n_0 \in \mathbb{N}$, so dass für alle $n \geq n_0$ gilt:

$$\|f_n - f\| < \frac{\varepsilon}{b-a} \quad \Leftrightarrow \quad \forall\, x \in [a,b]: \quad f_n(x) - \frac{\varepsilon}{b-a} \leq f(x) \leq f_n(x) + \frac{\varepsilon}{b-a}.$$

Insbesondere ist dann f beschränkt. Es sei nun $Z = \big\{x_0 = a, x_1, \dots, x_k = b\big\}$ eine Zerlegung von $[a,b]$. Dann gilt

$$U_Z(f_n) - \varepsilon = \sum_{j=1}^{k} m_j(f_n)(x_j - x_{j-1}) - \varepsilon \frac{\sum_{j=1}^{k}(x_j - x_{j-1})}{b-a}$$

$$= \sum_{j=1}^{k} \left(m_j(f_n) - \frac{\varepsilon}{b-a} \right) (x_j - x_{j-1})$$

$$\leq \sum_{j=1}^{k} m_j(f)(x_j - x_{j-1}) = U_Z(f).$$

Die Ungleichung gilt für sämtliche Zerlegungen; bilde daher Suprema über alle Zerlegungen Z. Wegen der Integrierbarkeit von f_n folgt dann

$$\underline{\int_a^b} f_n(x)\,\mathrm{d}x - \varepsilon = \sup_Z U_Z(f_n) - \varepsilon \leq \sup_Z U_Z(f) = \underline{\int_a^b} f(x)\,\mathrm{d}x.$$

Ein analoges Argument für Obersummen liefert entsprechend

$$\int_a^b f_n(x)\,\mathrm{d}x + \varepsilon = \overline{\int_a^b} f_n(x)\,\mathrm{d}x + \varepsilon \geq \overline{\int_a^b} f(x)\,\mathrm{d}x.$$

Insgesamt gilt dann für alle $n \geq n_0$:

$$\int_a^b f_n(x)\,\mathrm{d}x - \varepsilon \leq \underline{\int_a^b} f(x)\,\mathrm{d}x \leq \overline{\int_a^b} f(x)\,\mathrm{d}x \leq \int_a^b f_n(x)\,\mathrm{d}x + \varepsilon.$$

Mit einem Grenzübergang erhält man nun

$$\limsup_{n\to\infty} \int_a^b f_n(x)\,\mathrm{d}x \leq \underline{\int_a^b} f(x)\,\mathrm{d}x \leq \overline{\int_a^b} f(x)\,\mathrm{d}x \leq \liminf_{n\to\infty} \int_a^b f_n(x)\,\mathrm{d}x$$

$$\leq \limsup_{n\to\infty} \int_a^b f_n(x)\,\mathrm{d}x =: I(f).$$

Also ist

$$\underline{\int_a^b} f(x)\,\mathrm{d}x = \overline{\int_a^b} f(x)\,\mathrm{d}x = I(f) = \lim_{n\to\infty} \int_a^b f_n(x)\,\mathrm{d}x.$$

Damit ist f integrierbar und die behauptete Formel gilt. □

Wenn wir auf das Beispiel 5.4 (b) zurückblicken, wissen wir zwar, dass keine gleichmäßige Konvergenz vorliegt, aber es gilt dennoch:

$$\int_0^1 f_n(x)\,\mathrm{d}x = \frac{1}{2} \cdot \overbrace{1}^{\text{Höhe}} \cdot \underbrace{\frac{1}{n}}_{\text{Breite}} \xrightarrow[n\to\infty]{} 0,$$

also $\int_0^1 f_n \xrightarrow[n\to\infty]{} \int_0^1 f = 0$.

LINK: TEIL 2 DER 6. VORLESUNG VOM 28.04.2022

Der Satz 5.6 ist aber für $f_n \xrightarrow[n\to\infty]{\text{pkt}} f$ im Allgemeinen <u>nicht</u> wahr.

Beispiel 5.7. Wir betrachten für $n \in \mathbb{N}$ die Funktion

Im Video: **Beispiel 5.9.**

$$g_n \colon [0,1] \to \mathbb{R}, \quad x \mapsto \begin{cases} 2n^2 x & \text{für } 0 \le x \le \frac{1}{2n}, \\ 2n - 2n^2 x & \text{für } \frac{1}{2n} < x < \frac{1}{n}, \\ 0 & \text{für } \frac{1}{n} \le x \le 1. \end{cases}$$

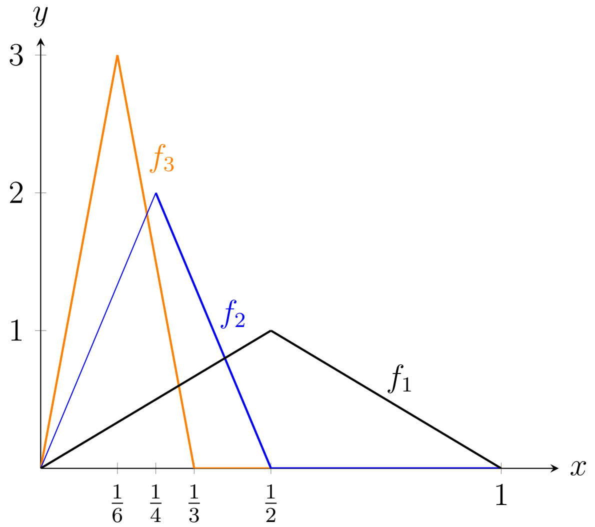

Abbildung 6: Der Graph von g_n für $n = 1$ (in schwarz), $n = 2$ (in blau) und $n = 3$ (in orange).

Wir sehen, dass

$$\int_0^1 g_n(x)\, \mathrm{d}x = \int_0^{\frac{1}{n}} g_n(x)\, \mathrm{d}x = \frac{1}{2} \cdot n \cdot \frac{1}{n} = \frac{1}{2} \neq 0.$$

Aber für alle $\delta \in (0,1]$ existiert ein $n_0 \in \mathbb{N}$ mit $\frac{1}{n_0} < \delta$, so dass

$$g_n(x) = g_n(\delta) = 0 \quad \forall\, n \ge n_0;$$

sogar falls $x \in [\delta, 1]$. Für die Funktionenfolge $(g_n)_n$ gilt also

- $g_n \xrightarrow[n\to\infty]{\text{pkt}} g \equiv 0$ auf $[0,1]$, sogar $g_n \xrightarrow[n\to\infty]{\text{glm}} g \equiv 0$ auf $[\delta, 1]$, aber
- $\int_0^1 g_n(x)\, \mathrm{d}x$ ist konstant gleich $\frac{1}{2}$.

Satz 5.8. *Es sei $f_n \colon [a,b] \to \mathbb{R}$ stetig differenzierbar auf $[a,b]$ für alle $n \in \mathbb{N}$. Es gelte weiter:*

Im Video: **Satz 5.10.**

- *f_n' konvergiere gleichmäßig auf $[a,b]$ und*
- *$\exists\, x_0 \in [a,b]: \quad (f_n(x_0))_n$ ist konvergent.*

Dann existiert eine stetig differenzierbare Funktion $f\colon [a,b] \to \mathbb{R}$ *mit*

$$f_n \xrightarrow[n\to\infty]{\text{glm}} f \text{ auf } [a,b]$$

und

$$f_n' \xrightarrow[n\to\infty]{\text{pktw}} f' \text{ auf } [a,b].$$

Beweis. Wir setzen zunächst

$$a := \lim_{n\to\infty} f_n(x_0) \quad \text{und} \quad g(x) := \lim_{n\to\infty} f_n'(x) \quad \forall\, x \in [a,b].$$

Da nach Voraussetzung f_n' stetig für alle $n \in \mathbb{N}$ und $f_n' \xrightarrow[n\to\infty]{\text{glm}} g$, folgt aus Satz 5.6, dass g stetig – und insbesondere integrierbar – ist auf $[a,b]$. Wir definieren nun

$$f\colon [a,b] \to \mathbb{R}, \quad f(x) := a + \int_{x_0}^{x} g(t)\,\mathrm{d}t, \quad \text{für } x \in [a,b]. \tag{$*$}$$

Nach dem Hauptsatz der Differential- und Integralrechnung ist f differenzierbar und es gilt $f' = g$ auf $[a,b]$. Also ist f sogar stetig differenzierbar, da g stetig auf $[a,b]$ ist.

Behauptung. Es gilt

$$\|f_n - f\|_{[a,b]} \xrightarrow[n\to\infty]{} 0.$$

Beweis der Behauptung. Es sei $\varepsilon > 0$. Dann existiert ein $n_0 \in \mathbb{N}$ mit

$$|f_n(x_0) - a| < \frac{\varepsilon}{2} \quad \forall\, n \geq n_0$$

nach der Definition von a und es folgt

$$\forall\, n \geq n_0: \quad \|f_n - f\|_{[a,b]} = \sup_{x\in[a,b]} |f_n(x) - g(x)| < \frac{\varepsilon}{2(b-a)}.$$

Der Hauptsatz der Differential- und Integralrechnung liefert dann

$$\int_{x_0}^{x} f_n'(t)\,\mathrm{d}t = f_n(x) - f_n(x_0) \quad \forall\, x \in [a,b]. \tag{$**$}$$

Zusammen mit der Definition von f folgt dann für alle $x \in [a,b]$ und $n \geq n_0$:

$$\begin{aligned}
|f_n(x) - f(x)| &\overset{(**)}{=} \left| f_n(x_0) + \int_{x_0}^{x} f_n'(t)\,\mathrm{d}t - f(x) \right| \\
&\overset{(*)}{=} \left| f_n(x_0) - a + \int_{x_0}^{x} \big(f_n'(t) - g(t)\big)\,\mathrm{d}t \right| \\
&\overset{\Delta}{\leq} |f_n(x_0) - a| + \int_{x_0}^{x} \left| f_n'(t) - g(t) \right| \mathrm{d}t
\end{aligned}$$

$$< \frac{\varepsilon}{2} + \underbrace{\int_{x_0}^{x} \mathrm{d}t}_{\leq b-a} \cdot \frac{\varepsilon}{2(b-a)}$$

$$\leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Es gilt also $f_n \xrightarrow[n\to\infty]{\text{glm}} f$ auf $[a,b]$. □

Unter den gegebenen Voraussetzungen darf man also $\lim_{n\to\infty}$ und $\frac{\mathrm{d}}{\mathrm{d}x}$ (was ebenfalls ein Grenzübergang ist) vertauschen:

$$f'(x) = \frac{\mathrm{d}}{\mathrm{d}x} \overbrace{\lim_{n\to\infty}}^{\text{glm}} f_n(x) = \lim_{n\to\infty} f_n'(x).$$

Aber: $C^1(I) \ni f_n \xrightarrow[n\to\infty]{\text{glm}} f$ alleine impliziert weder die Konvergenz der Folge $(f_n')_n$ noch die Differenzierbarkeit von f.

Beispiel 5.9.

Im Video: **Beispiel 5.11.**

(a) Es sei $I = [0,1]$ und

$$f_n \colon I \to \mathbb{R}, \quad f_n(x) := \frac{\sin(nx)}{\sqrt{n}}, \quad n \in \mathbb{N}.$$

Dann gilt für alle $n \in \mathbb{N}$ und $x \in I$:

$$|f_n(x)| \leq \frac{1}{\sqrt{n}} \xrightarrow[n\to\infty]{} 0,$$

also $f_n \xrightarrow[n\to\infty]{\text{glm}} f \equiv 0$.

Behauptung. Die Ableitung

$$f_n'(x) = \sqrt{n}\cos(nx)$$

konvergiert nicht.

Beweis der Behauptung. Angenommen, f_n' konvergiere, es gelte also

$$\lim_{n\to\infty} f_n'(x) = a \in \mathbb{R}.$$

Da $(\sqrt{n})_n$ unbeschränkt ist, muss $(\cos(nx))_n$ eine Nullfolge sein. Insbesondere muss dann die Teilfolge $(\cos(2nx))_n$ ebenfalls eine Nullfolge sein. Die Additionstheoreme der Trigonometrie implizieren dann

$$\underbrace{\cos(2nx)}_{\xrightarrow[n\to\infty]{\text{nach Ann.}} 0} = \cos^2(nx) - \underbrace{\sin^2(nx)}_{=1-\cos^2(nx)} = \underbrace{2\cos^2(nx)}_{\xrightarrow[n\to\infty]{\text{nach Ann.}} 0} - 1,$$

was ein Widerspruch ist.

Für $x = 0$ gilt sogar $f_n'(0) = \sqrt{n} \xrightarrow[n\to\infty]{} \infty$.

(b) Für $n \in \mathbb{N}$ betrachten wir

$$f_n \colon \mathbb{R} \to \mathbb{R}, \quad x \mapsto (-1)^n + \frac{x}{n}.$$

Dann ist $(f_n(x))_n$ für alle $x \in \mathbb{R}$ eine divergente Folge. Für alle $x \in \mathbb{R}$ ist jedoch $\left(f_n'(x) - \frac{1}{n}\right)_n$ eine Nullfolge, es gilt sogar

$$\|f_n'\|_{\mathbb{R}} \xrightarrow[n\to\infty]{} 0, \text{ also } \quad f_n' \xrightarrow[n\to\infty]{\text{glm}} 0 \text{ auf } \mathbb{R}.$$

6 Funktionenreihen

Definition 6.1.

(a) Für $n \in \mathbb{N}_0$ sei $f_n \in \mathcal{F}(M)$ und

$$s_n(x) := \sum_{k=0}^{n} f_k(x) = f_0(x) + f_1(x) + \ldots + f_n(x)$$

die n-te Partialsumme der Funktionenfolge. Wir nennen dann $(s_n)_n \subset \mathcal{F}(M)$ eine *Funktionenreihe* und bezeichnen sie formal mit

$$\sum_{k\in\mathbb{N}_0} f_k = \sum_{k=0}^{\infty} f_k = f_0 + f_1 + f_2 + \ldots,$$

ebenso wie die partiell existierende Grenzfunktion der Folge $(s_n)_n$.

(b) Es sei $s \colon M \to \mathbb{C}$ gegeben. Die Funktionenreihe $\sum_{k\in\mathbb{N}_0} f_k$ heißt *punktweise konvergent gegen* s, falls

$$\forall\, x \in M : \quad s_n \xrightarrow[n\to\infty]{\text{pktw}} s.$$

Wir schreiben dann $s = \sum_{k\in\mathbb{N}_0} f_k$.

$\sum_{k\in\mathbb{N}_0} f_k$ heißt *punktweise konvergent*, falls ein $s \colon M \to \mathbb{C}$ existiert, so dass das Obige gilt.

(c) Die Reihe $\sum_{k\in\mathbb{N}_0} f_k$ heißt *gleichmäßig / uniform konvergent gegen* $s \colon M \to \mathbb{C}$, falls

$$s_n \xrightarrow[n\to\infty]{\text{glm}} s.$$

(d) Die Funktionenreihe $\sum_{k\in\mathbb{N}_0} f_k$ heißt *normal konvergent*, falls die Reihe $\sum_{k\in\mathbb{N}_0} \|f_k\|$ in $\mathbb{R}$ konvergiert.

LINK: TEIL 1 DER 7. VORLESUNG VOM 04.05.2022

(e) Es sei $(f_k)_{k\in\mathbb{N}_0} \subset \mathcal{F}(M)$ und $A \subset M$. Dann heißt $\sum_k f_k$ *normal konvergent auf* A, falls

$$\sum \|f_k\|_A < \infty.$$

Aus dem Cauchy-Kriterium für Funktionenfolgen (vgl. Lemma 5.3 (iii)) ergibt sich:

Lemma 6.2. *Es sei $(f_n)_n \subset \mathcal{F}(M)$. Dann sind die folgenden Aussagen äquivalent:*

(a) *Die Reihe $\sum_{k\in\mathbb{N}_0} f_k$ ist uniform konvergent;*

(b) *Für alle $\varepsilon > 0$ existiert ein $n_0 = n_0(\varepsilon) \in \mathbb{N}$, so dass*

$$\forall\, n, m \geq n_0, \quad \forall\, x \in M : \quad \left| \sum_{k=n+1}^{m} f_k(x) \right| < \varepsilon.$$

Satz 6.3 (Weierstraß-Majoranten-Kriterium)**.** *Für $k \in \mathbb{N}_0$ sei $(f_k)_k \subset \mathcal{F}(M)$. Dann gelten:*

(a)

$$\sum f_k \textit{ ist gleichmäßig konvergent.}$$

$$\Leftrightarrow \textit{Es gilt } r_n \xrightarrow[n\to\infty]{\text{glm}} f \equiv 0, \quad \textit{wobei} \quad r_n := \sum_{k=n+1}^{\infty} f_k.$$

$$\Rightarrow \textit{Es gilt } f_k \xrightarrow[k\to\infty]{\text{glm}} f \equiv 0.$$

(b) *Falls $\sum f_k$ normal konvergent ist, so konvergieren $\sum f_k$ und $\sum |f_k|$ gleichmäßig.*

Beweis. Übung!

Beispiel 6.4. Es sei $z \in \mathbb{C}$ mit $|z| < 1$. Dann gilt

$$\sum_{k\in\mathbb{N}_0} z^k = \frac{1}{1-z}.$$

Wegen der Stetigkeit von $z \mapsto z^k$ gilt außerdem

$$\|z^k\|_{(-1,1)} = \sup_{|z|<1} |z^k| = 1,$$

sodass $f_k(z) = z^k$ keine gleichmäßige Nullfolge ist. Damit ist $\sum x^k$ nicht gleichmäßig konvergent auf $(-1, 1) \subset \mathbb{R}$, insbesondere ist $\sum z^k$ nicht gleichmäßig konvergent auf $B_1 \subset \mathbb{C}$.

Für $\delta \in (0,1)$ ist aber
$$\sum_{k\in\mathbb{N}_0} \|z^k\|_{\overline{B_\delta}} \leq \sum_{k\in\mathbb{N}_0} \delta^k < \infty$$
konvergent. Nach dem Weierstraß-Majoranten-Kriterium 6.3 ist dann $\sum z^k$ gleichmäßig und absolut konvergent auf $\overline{B_\delta} = \left\{z \in \mathbb{C} \;:\; |z| \leq \delta\right\}$.

Lemma 6.5. *Für $k \in \mathbb{N}_0$ sei $(g_k)_k \subset \mathcal{F}(M)$. Existieren $q \in (0,1)$, $C \in \mathbb{R}$ und $l \in \mathbb{N}$ derart, dass*
$$\|g_{n+1} - g_n\| < C \cdot q^n \quad \forall\, n \geq l,$$
so ist $(g_n)_n$ normal konvergent.

Beweis. Wir definieren zunächst $f_n(x) := g_n(x) - g_{n-1}(x)$ für $x \in M$ und $g_{-1} \equiv 0$. Dann sind die Partialsummen $g_n = \sum_{k=0}^n f_k = s_n$ normal konvergent, da
$$\sum_{k=l}^n \|f_k\| \leq C \cdot \sum_{k=l}^n q^k$$
und $\sum_{k=0}^l \|f_k\|$ endlich sowie fest ist. Das Weierstraß-Majoranten-Kriterium 6.3 liefert dann die Behauptung. □

Beispiel 6.6.

(a) Die Funktionenfolge $(E_n)_{n\in\mathbb{N}_0}$ definiert durch
$$E_n \colon \mathbb{C} \to \mathbb{C}, \quad z \mapsto \sum_{k=0}^n \frac{z^k}{k!}$$
konvergiert auf jedem abgeschlossenen Ball $\overline{B_R} = \left\{z \in \mathbb{C} \;:\; |z| \leq R, \quad R > 0\right\}$ uniform, denn es gilt
$$\|f_k\|_{\overline{B_R}} = \left\|\frac{z^k}{k!}\right\|_{\overline{B_R}} \overset{\|z\|\leq R}{\leq} \frac{R^k}{k!}$$
$$\leq \left(\frac{3R}{k}\right)^k, \quad \text{da } \forall\, k \in \mathbb{N}: \quad \left(\frac{k}{3}\right)^3 \leq k! \leq \left(\frac{k}{2}\right)^2 \text{ (Übung!)}$$
$$\overset{6R<k}{\leq} \left(\frac{1}{2}\right)^k.$$
Damit hat $\sum_{k=\lfloor 6R\rfloor+1}^n \|f_k\|_{\overline{B_R}}$ die geometrische Reihe als Majorante, während $\sum_{k=0}^{\lfloor 6R\rfloor} \|f_k\|_{\overline{B_R}}$ einen endlichen, festen Beitrag liefert. Damit ist das Weierstraß-Majoranten-Kriterium 6.3 anwendbar. Dagegen ist
$$\sup_{x\in\mathbb{R}} \left|\frac{x^k}{k!}\right| = \infty, \text{ insbesondere } \quad \left\|\frac{z^k}{k!}\right\|_{\mathbb{C}} = \infty.$$

(b) Für $k \in \mathbb{N}$ betrachten wir $f_k(z) = \frac{z^k}{k}$ und $\delta \in (0,1)$. Dann gilt

$$\|f_k\|_{\overline{B_\delta}} \leq \frac{\delta^k}{k} \leq \delta^k.$$

Nach Lemma 6.5 ist dann $\sum_{k\in\mathbb{N}} f_k$ normal konvergent auf $\overline{B_\delta}$. Für $z = 1$ ergibt sich jedoch die (divergente) harmonische Reihe. Also konvergiert $\sum f_k$ nicht gleichmäßig auf $(0,1) \subset \mathbb{R}$, da

$$\sup_{z\in(0,1)} \left|\frac{z^k}{k}\right| = \frac{1}{k}.$$

Aus den entsprechenden Sätzen für Funktionenfolgen ergeben sich folgende Vertauschungsregeln für Funktionenreihen:

Satz 6.7. *Es seien $\varnothing \neq I = [a,b] \subset \mathbb{R}$, $y \in [a,b]$ und $s, f_k \in \mathcal{F}(I,\mathbb{R})$, $k \in \mathbb{N}_0$. Dann gelten*

(i) *Falls $\sum_{k\in\mathbb{N}_0} f_k$ gleichmäßig konvergent ist mit $\sum_{k\in\mathbb{N}_0} f_k = s$ und f_k stetig in $y \in [a,b]$ für alle $k \in \mathbb{N}_0$ ist, so ist s stetig in y.*

(ii) *Ist $s = \sum f_k$ gleichmäßig konvergent und f_k integrierbar für alle $k \in \mathbb{N}_0$, so ist auch s integrierbar und es gilt*

$$\int_a^b s \,\mathrm{d}x = \sum_{k\in\mathbb{N}_0} \int_a^b f_k \,\mathrm{d}x.$$

(iii) *Falls $f_k \in C^1(I)$ für alle $k \in \mathbb{N}_0$, $\sum f_k'$ gleichmäßig konvergent auf I und $\sum f_k(y)$ konvergent ist, so ist $\sum f_k$ gleichmäßig konvergent gegen eine Grenzfunktion $g \in C^1(I)$ und es gilt*

$$g'(x) = \sum_{k\in\mathbb{N}_0} f_k'(x) \quad \forall\, x \in I.$$

Beispiel 6.8.

(a) Für $k \in \mathbb{N}_0$ betrachten wir die Funktion

$$f_k\colon \mathbb{R} \to \mathbb{R}, \quad f_k(x) := \frac{x^2}{(1+x^2)^k}.$$

Es gilt $f_k(0) = 0$ und damit auch $\sum_{k\in\mathbb{N}_0} f_k(0) = 0$. Für $x \neq 0$ gilt $1 + x^2 > 1$ und damit

$$\sum_{k\in\mathbb{N}_0} f_k(x) = x^2 \cdot \sum_{k\in\mathbb{N}_0} (1+x^2)^{-k} = \frac{x^2}{1 - \frac{1}{1+x^2}} = 1 + x^2.$$

Damit ist $\sum f_k$ auf $\mathbb{R}$ punktweise konvergent gegen

$$s\colon \mathbb{R} \to \mathbb{R}, \quad x \mapsto \begin{cases} 0 & \text{für } x = 0, \\ 1 + x^2 & \text{für } x \neq 0. \end{cases}$$

Link: Teil 2 der 7. Vorlesung vom 04.05.2022

f_k ist stetig auf $\mathbb{R}$ für alle $k \in \mathbb{N}_0$, wohingegen das für s im Punkt $x = 0$ nicht zutrifft. Nach Satz 6.7 wird jedoch deutlich, dass die Konvergenz von $\sum f_k$ auf $\mathbb{R}$ nicht gleichmäßig ist. Dagegen konvergiert $\sum f_k$ gleichmäßig gegen s auf $(-\infty, -\delta) \cup (\delta, \infty)$ für alle $\delta > 0$.

(b) Für $k \in \mathbb{N}_0$ betrachten wir die Funktion

$$f_k \colon \mathbb{R} \to \mathbb{R}, \quad f_k(x) := \frac{\cos(kx)}{k^{1+\alpha}}, \quad \alpha > 0.$$

Wegen $|\cos(kx)| \leq 1$ gilt

$$\sum \|f_k\| \leq \sum \frac{1}{k^{1+\alpha}} < \infty,$$

weshalb $\sum f_k$ normal konvergent ist. Bezeichnet s die Grenzfunktion von $\sum f_k$, so besitzt s eine Stammfunktion. Nach Satz 6.7 ist

$$F(x) = \int_0^x s \,\mathrm{d}t \overset{6.7}{=} \sum_{k \in \mathbb{N}_0} \int_0^x \frac{\cos(kt)}{k^{1+\alpha}} \,\mathrm{d}t = \sum_{k \in \mathbb{N}_0} \frac{\sin(kx)}{k^{2+\alpha}}$$

eine Stammfunktion von s.

7 Potenzreihen

Die zugrundeliegende Frage, mit der wir uns in diesem Kapitel beschäftigen wollen, ist die Folgende: Für welche Werte $z \in \mathbb{C}$ sind beispielsweise die Funktionenreihen

$$\sum_{k \in \mathbb{N}_0} z^k, \quad \sum_{k \in \mathbb{N}_0} \frac{z^k}{k!} \quad \text{sowie} \quad \sum_{k \in \mathbb{N}} \frac{z^k}{k}$$

konvergent oder sogar gleichmäßig konvergent?

Definition 7.1.

(a) Eine Funktionenreihe der Form

$$\sum_{k \in \mathbb{N}_0} f_k, \quad f_k(z) = a_k \cdot (z - z_0)^k, \quad z \in \mathbb{C},$$

mit *Koeffizienten* $a_k \in \mathbb{C}$ und *Entwicklungspunkt* $z_0 \in \mathbb{C}$ heißt *Potenzreihe*. Falls wir diese Funktionenreihe auf $\mathbb{R}$ einschränken, so schreiben wir oft

$$\sum_{k \in \mathbb{N}_0} a_k (x - x_0)^k, \quad x \in \mathbb{R},$$

wobei dies besonders im Fall $x_0, a_k \in \mathbb{R}$, $k \in \mathbb{N}_0$, relevant ist.

(b) Zu einer gegebenen Funktionenreihe mit Koeffizienten $(a_k)_k$ definieren wir den *Konvergenzradius* als

$$\rho = \left(\limsup_{k\to\infty} \sqrt[k]{|a_k|} \right)^{-1} \in [0, \infty]. \tag{$*$}$$

Offensichtlich ist der lim sup nicht-negativ, kann aber die Werte 0 und ∞ annehmen. Als Konvention setzen wir dabei

$$\frac{1}{0} := \infty \quad \text{und} \quad \frac{1}{\infty} = 0.$$

Frage 7.2 (Wo konvergieren die drei oberen Beispiel-Reihen?)**.**

(i) Für $|z| < 1$ ist die Reihe $\sum f_k = \sum z^k$ offensichtlich konvergent. Für $|z| \geq 1$ ist z^k hingegen keine Nullfolge, weshalb dann $\sum f_k$ divergiert. Für den Konvergenzradius gilt entsprechend

$$\rho = \left(\limsup_{k\to\infty} \sqrt[k]{1} \right)^{-1} = 1^{-1} = 1.$$

(ii) Die Reihe $\sum_{k\in\mathbb{N}_0} \frac{z^k}{k!}$ ist konvergent für alle $z \in \mathbb{C}$. Für den Konvergenzradius gilt

$$\rho = \left(\limsup_{k\to\infty} \underbrace{(k!)^{-\frac{1}{k}}}_{\leq \frac{3}{k}} \right)^{-1} \geq \limsup_{k\to\infty} \frac{k}{3} = \infty.$$

(iii) Für $|z| < 1$ konvergiert die Reihe $\sum_{k\in\mathbb{N}} \frac{z^k}{k}$ absolut nach dem Majoranten-Kriterium. Für $|z| > 1$ ist die Reihe $\sum_{k\in\mathbb{N}} \frac{z^k}{k}$ nicht absolut konvergent nach dem Minoranten-Kriterium. Außerdem gilt

$$\sqrt[k]{\frac{1}{k}} = \frac{1}{\sqrt[k]{k}} \xrightarrow[k\to\infty]{} 1 \quad \Rightarrow \quad \rho = 1.$$

Die obigen Beispiele legen den folgenden Satz nahe:

Satz 7.3 (Konvergenzbereich)**.** *Die Potenzreihe $\sum a_k(z - z_0)^k$ aus Definition 7.1*

- *konvergiert absolut für alle $z \in B_\rho(z_0)$,*
- *divergiert für alle $z \in \overline{B_\rho(z_0)}^c = \left\{ z \in \mathbb{C} \ : \ |z - z_0| > \rho \right\}$ und*
- *ist für alle $\delta \in (0, \rho)$ auf $\overline{B_\delta(z_0)} = \left\{ z \in \mathbb{C} \ : \ |z - z_0| \leq \delta \right\}$ sogar normal konvergent.*

Man beachte, dass für $\rho = 0$ der Satz eigentlich nichts aussagt. Aber offensichtlich gilt $\sum a_k(z_0 - z_0)^k = \sum 0 = 0$; die Potenzreihe konvergiert also immer in $z = z_0$. Falls $\rho = \infty$, so gilt

$$B_\rho(z_0) = \left\{ z \in \mathbb{C} \ : \ |z - z_0| < \infty \right\} = C.$$

Also ist die Potenzreihe dann konvergent auf ganz $\mathbb{C}$.

Beweis von Satz 7.3. Es gilt

$$\limsup_{k\to\infty} \sqrt[k]{|a_k(z-z_0)^k|}$$

$$= \limsup_{k\to\infty} \sqrt[k]{|a_k|} \cdot |z-z_0|$$

$$= \frac{1}{\rho}|z-z_0| \begin{cases} < 1, & \text{falls } |z-z_0| < \rho \quad \Rightarrow \quad \text{Reihe konvergiert in } z, \\ > 1, & \text{falls } |z-z_0| > 0 \quad \overset{2.7}{\Rightarrow} \quad \text{Reihe divergiert in } z. \end{cases}$$

Insbesondere ist $\sum_{k\in\mathbb{N}_0} a_k\delta^k$ absolut konvergent für alle $\delta \in (0,\rho)$. Außerdem gilt

$$\sup_{|z-z_0|\leq\delta} \left|a_k(z-z_0)^k\right| \leq |a_k|\delta^k, \text{ ist also eine summierbare Majorante.}$$

Nach dem Weierstraß-Majoranten-Kriterium 6.3 ist dann die Potenzreihe normal und gleichmäßig konvergent auf $\overline{B_\delta(z_0)}$. □

Bemerkung 7.4. Die Menge $I = \mathbb{R} \cap B_\rho(z_0)$ nennt man das *Konvergenzintervall* der Potenzreihe $\sum a_k(z-z_0)^k$. Die Gleichung $(*)$ heißt *Cauchy-Hadamard-Formel* und ist besonders handlich, falls $\lim_{k\to\infty} \sqrt[k]{|a_k|}$ existiert.

LINK: TEIL 1 DER 8. VORLESUNG VOM 05.05.2022

Beispiel 7.5. Über $z \in \partial B_\rho(z_0) = \left\{z \in \mathbb{C} \ : \ |z-z_0| = \rho\right\}$ wird im Satz 7.3 keine Konvergenzaussage gemacht.

(a) Wir betrachten die (alternierende) harmonische Reihe: Die Reihe $\sum_{k\in\mathbb{N}} \frac{z^k}{k}$ ist divergent für $z = 1$ und konvergent für $z = -1$.

(b) Die Reihe $\sum_{k\in\mathbb{N}} z^k$ divergiert für alle $z \in \partial B_1$, da dort $|z^k| = |z|^k = 1$ gilt, $(|z^k|)_k$ also keine Nullfolge ist.

(c) Die Reihe $\sum_{k\in\mathbb{N}} \frac{z^k}{k^2}$ ist normal konvergent auf $\overline{B_1} = \left\{z \in \mathbb{C} \ : \ |z| \leq 1\right\}$, da

$$\left|\frac{z^k}{k^2}\right| \leq \frac{1}{k^2} \quad \forall z \in \overline{B_1}$$

und $\sum \frac{1}{k^2}$ eine konvergente Reihe reeller Zahlen ist.

(d) Jedes Polynom $P(z) = \sum k = 0^n a_k z^k$ ist als Potenzreihe darstellbar mit $a_{n+1} = a_{n+2} = \ldots = 0$. Insbesondere existiert der Grenzwert

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

Insbesondere gilt $\rho = \infty$, weshalb die Reihe auf ganz $\mathbb{C}$ konvergiert.

Bei skalaren Reihen das Wurzelkriterium allgemeiner, aber das Quotientenkriterium einfacher anzuwenden. In manchen Situationen ist das folgende Kriterium anwendbar (und dann auch häufig einfacher als die Verwendung der Cauchy-Hadamard-Formel ($*$)):

Satz 7.6. *Es seien $a_k \in \mathbb{C}$, $a_k \neq 0$ für fast alle $k \in \mathbb{N}$ und der Grenzwert*

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

existierte in $[0,\infty]$. Dann hat die Potenzreihe $\sum a_k(z-z_0)^k$ den Konvergenzradius $\rho = \beta$.

Beweis. Die Aussage folgt direkt aus dem Quotientenkriterium aus Satz 2.9. □

Beispiel 7.7. Es sei $\sum_{k\in\mathbb{N}_0} a_k z^k$ eine Potenzreihe mit Entwicklungspunkt $z_0 = 0$.

(a) Für $a_k = k^k$ erhalten wir

$$\limsup_{k\to\infty} \sqrt[k]{|a_k|} = \lim_{k\to\infty} k = \infty \quad \Rightarrow \quad \rho = 0.$$

(b) Für $a_k = \left(3 + (-1)^k\right)^k$ gilt

$$\limsup_{k\to\infty} \sqrt[k]{|a_k|} = \limsup_{k\to\infty} \left(3 + (-1)^l\right) = 4,$$

also $\rho = \frac{1}{4}$.

Dagegen gilt

$$\liminf_{k\to\infty} \left(3 + (-1)^k\right) = 2 \neq 4,$$

der Grenzwert $\lim_{k\to\infty} |a_k|^{\frac{1}{k}}$ existiert also nicht.

(c) Es sei nun $a_k = \frac{k^k}{(k!)^2}$. Dann gilt

$$\left|\frac{a_k}{a_{k+1}}\right| = \frac{k^k\left((k+1)!\right)^2}{(k!)^2(k+1)^{k+1}} = \underbrace{\frac{(k+1)^2}{(k+1)}}_{=k+1} \cdot \overbrace{\left(\frac{k}{k+1}\right)^k}^{=\left(\left(1+\frac{1}{k}\right)^k\right)^{-1} \to \frac{1}{e}} \xrightarrow[k\to\infty]{} \infty,$$

also $\rho = \beta = \infty$.

(d) Betrachten wir nun die Potenzreihe $\sum_{k\in\mathbb{N}} (-1)^k(2^k+k)(z+1)^k$, also den Entwicklungspunkt $z_0 = -1$ und Koeffizienten $a_k = (-1)^k(2^k+k)$. Dann gilt

$$\sqrt[k]{2^k + k} = 2\left(1 + \frac{k}{2^k}\right)^{\frac{1}{k}} \xrightarrow[k\to\infty]{} 2 \quad \text{(Übung!)},$$

also $\rho = \frac{1}{2}$.

Satz 7.8. *Es sei* $z_0 \in \mathbb{C}$ *und* $(a_k)_{k\in\mathbb{N}_0} \subset \mathbb{C}$ *eine Folge mit*

$$\rho = \left(\limsup_{k\to\infty} \sqrt[k]{|a_k|}\right)^{-1} > 0.$$

So ist durch

$$z \mapsto f(z) := \sum_{k\in\mathbb{N}_0} a_k(z-z_0)^k$$

eine stetige Funktion auf $B_\rho(z_0)$ *definiert.*

Beweis. Die Funktion $z \mapsto f_k(z) := a_k(z-z_0)^k$ ist stetig für alle $k \in \mathbb{N}_0$, während $\sum_{k\in\mathbb{N}_0} f_k$ gleichmäßig auf $\overline{B_\delta(z_0)}$ für alle $\delta \in (0,\rho)$ konvergiert. Also ist f stetig auf $\bigcup_{\delta\in(0,\rho)} \overline{B_\delta(z_0)} \overset{\text{Übg.}}{=} B_\rho(z_0)$. □

Wie es bei Reihen und Funktionen der Fall ist, lassen sich Potenzreihen auch durch Addition und Multiplikation zu neuen Potenzreihen kombinieren.

Satz 7.9. *Es seien zwei Funktionen*

$$f(z) = \sum_{k\in\mathbb{N}_0} a_k(z-z_0)^k \quad \textit{und} \quad g(z) = \sum_{k\in\mathbb{N}_0} b_k(z-z_0)^k$$

durch Potenzreihen um denselben Entwicklungspunkt $z_0 \in \mathbb{C}$ *und Konvergenzradien* $\rho_f > 0$ *bzw.* $\rho_g > 0$ *gegeben. Sei weiter* $c \in \mathbb{C}$ *und setze* $\rho = \min\left\{\rho_f, \rho_g\right\}$. *Folgende Potenzreihen stellen dann Funktionen in* $C\left(B_\rho(z_0)\right)$ *dar:*

- $z \mapsto (f+cg)(z) = \sum_{k\in\mathbb{N}_0} (a_k + cb_k)(z-z_0)^k$;
- $z \mapsto (f\cdot g)(z) = \sum_{k\in\mathbb{N}_0} \Bigg(\underbrace{\sum_{j=0}^{k} a_j b_{k-j}}_{\text{vgl. } c_n \text{ aus (CP)}}\Bigg)(z-z_0)^k.$

Beweis. Die Aussage folgt direkt über die Verwendung von analogen Aussagen über (absolut) konvergente Reihen und insbesondere dem Cauchy-Produkt. □

Satz 7.10. *Für alle* $l \in \mathbb{Z}$ *besitzen Potenzreihen der Form*

$$\sum_{k\in\mathbb{N}} k^l a_k(z-z_0)^k \tag{7.1}$$

denselben Konvergenzradius.

Beispiele für solche Potenzreihen sind

$$\sum_{k\in\mathbb{N}} a_k(z-z_0)^k, \quad \sum_{k\in\mathbb{N}} ka_k(z-z_0)^k \quad \textit{und} \quad \sum_{k\in\mathbb{N}} \frac{a_k}{k}(z-z_0)^k.$$

Beweis. Es gilt

$$\sqrt[k]{|k^l a_k|} = k^{\frac{l}{k}} \cdot \sqrt[k]{|a_k|} = \underbrace{\left(\sqrt[k]{k}\right)^l}_{\xrightarrow[k\to\infty]{} 1} \cdot \sqrt[k]{|a_k|}.$$

Für den invertierten Konvergenzradius von (7.1) folgt dann

$$\limsup_{k\to\infty} \sqrt[k]{k^l |a_k|} = \limsup_{k\to\infty} \sqrt[k]{|a_k|} = \frac{1}{\rho}. \qquad \square$$

LINK: TEIL 2 DER 8. VORLESUNG VOM 05.05.2022

Satz 7.11. *Es sei*

$$f(x) = \sum_{k\in\mathbb{N}_0} a_k (z - z_0)^k$$

die Summenfunktion einer Potenzreihe mit Konvergenzradius ρ und der Schnitt $I := B_\rho(z_0) \cap \mathbb{R}$ ein echtes Intervall. Dann ist $f \in C^1(I)$ und es gilt

$$f'(x) = \sum_{k\in\mathbb{N}} k a_k (x - z_0)^{k-1} \quad \forall\, x \in I. \tag{$\times\times$}$$

Beweis. Die Potenzreihe

$$g(z) = \sum_{k\in\mathbb{N}_0} k a_k (z - z_0)^k$$

konvergiert ebenfalls auf $B_\rho(z_0)$ – nach Satz 7.10 sogar gleichmäßig auf $B_\delta(z_0)$ für alle $\delta \in (0, \rho)$. Da $z \mapsto k a_k (z - z_0)^k$ stetig ist, ist auch g stetig auf $B_delta(z_0)$ für alle $\delta \in (0, \rho)$, insbesondere auf $I_\delta = B_\delta(z_0) \cap \mathbb{R}$, also auch auf $I = \bigcup_{\delta\in(0,\rho)} I_\delta$. Nun impliziert Satz 6.7 (iii) die Formel

$$f'(x) = g(x) \quad \forall\, x \in I,$$

also die Formel ($\times\times$). $\square$

Iteration dieses Satzes liefert dann:

Korollar 7.12. *Unter den Voraussetzungen von Satz 7.11 gelten:*

(a) *Es gilt $f \in C^\infty(I)$;*

(b) *Für alle $m \in \mathbb{N}$ und $x \in I$ gilt*

$$f^{(m)}(x) = \sum_{k=m}^{\infty} k \cdot (k-1) \cdot \ldots \cdot (k - m + 1) a_k (x - z_0)^{k-m}$$

$$= \sum_{k=0}^{\infty} (k+m) \cdot (k + m - 1) \cdot \ldots \cdot (k+1) a_{k+m} (x - z_0)^k;$$

Man bemerke, dass $(x - z_0)^k = 0$ gilt, falls $x = z_0$ und $k \geq 1$. Für $k = 0$ gilt nämlich immer $(x - z_0)^k = 1$.

(c) *Es gilt*

$$f^{(m)}(z_0) = m! \cdot a_m,$$

falls $z_0 \in \mathbb{R}$. *Die Koeffizienten* $a_m = \frac{f^{(m)}(z_0)}{m!}$ *der Potenzreihe von* f *sind also um einen gegebenen Entwicklungspunkt* z_0 *durch die Funktion* f *selbst eindeutig bestimmt – dies gilt sogar für allgemeine* $z_0 \in \mathbb{C}$ *–;*

(d) *Die Funktion*

$$F\colon I \to \mathbb{C}, \quad x \mapsto \sum_{k \in \mathbb{N}_0} \frac{a_k}{k+1}(x - z_0)^{k+1}$$

ist eine Stammfunktion von f *auf* I*;*

(e) *Für jedes* $[a, b] \subset I$ *gilt*

$$\int_a^b f(x)\,\mathrm{d}x = \sum_{k \in \mathbb{N}_0} \int_a^b a_k(x - z_0)^k\,\mathrm{d}x = \sum_{k \in \mathbb{N}_0} \frac{a_k}{k+1}\Big((b - z_0)^{k+1} - (a - z_0)^{k+1}\Big).$$

„Potenzreihen darf man gliedweise Ableiten und Integrieren."

Beweis. Die Teile (a) bis (d) zeigt man direkt über vollständige Induktion über $m \in \mathbb{N}$ und der Anwendung von Satz 7.11.

(c) Einsetzen von $x = z_0 \in \mathbb{R}$ liefert direkt die Aussage.

(d) Diese Potenzreihe hat nach Satz 7.10 denselben Konvergenzbereich. Anwendung von Satz 7.11 liefert dann die Behauptung.

(e) Da $I = B_\rho(z_0) \cap \mathbb{R}$ offen ist, existiert ein $\delta \in (0, \rho)$ derart, dass $[a, b] \subset B_\delta(z_0)$. Daraus folgt die gleichmäßige Konvergenz der Potenzreihe auf $B_\delta(z_0) \supset [a, b]$. Wegen Satz 6.7 ist dann das gliedweise Integrieren erlaubt. □

Beispiel 7.13. Die Funktion

$$h\colon \mathbb{R} \to \mathbb{R}, \quad x \mapsto \begin{cases} \ln\left(\frac{a}{b}\right) & \text{für } x = 0, \\ \frac{a^x - b^x}{x} & \text{für } x \neq 0, \end{cases}$$

ist nach den Regeln von L'Hospital (s. Satz 12.5 aus der Vorlesung zu Analysis I (Lehramt)) stetig für alle $a, b > 0$. Diese wollen wir nun als Potenzreihe darstellen. Es gilt

$$\begin{aligned} \frac{a^x - b^x}{x} &= \frac{1}{x}\left(\mathrm{e}^{x\ln(a)} - \mathrm{e}^{x\ln(b)}\right) \\ &= \frac{1}{x}\left(\sum_{k \in \mathbb{N}_0} \frac{\big(x\ln(a)\big)^k}{k!} - \sum_{k \in \mathbb{N}_0} \frac{\big(x\ln(b)\big)^k}{k!}\right) \end{aligned}$$

$$\overset{7.9}{=} \frac{1}{x}\sum_{k\in\mathbb{N}_0}\frac{x^k}{k!}\Big(\overbrace{\underbrace{\ln^k(a)-\ln^k(b)}_{=1-1=0 \text{ für } k=0}}^{=(\ln(a))^k-(ln(b))^k}\Big)$$

$$= \frac{1}{x}\sum_{k\in\mathbb{N}}\frac{x^k}{k!}\Big(\ln^k(a)-\ln^k(b)\Big)$$

$$\texttt{[Indexverschiebung: j=k-1]} = \sum_{j\in\mathbb{N}_0} x^j\frac{\ln^{j+1}(a)-\ln^{j+1}(b)}{(j+1)!}.$$

Nun gilt außerdem

$$\left|\frac{\ln^{j+1}(a)-\ln^{j+1}(b)}{(j+1)!}\right|^{\frac{1}{j}} \le \left(\frac{\left|\ln^{j+1}(a)\right|+\left|\ln^{j+1}(b)\right|}{(j+1)!}\right)^{\frac{1}{j}}$$

$$\le \frac{\left|\ln^{j+1}(a)\right|^{\frac{1}{j}}+\left|\ln^{j+1}(b)\right|^{\frac{1}{j}}}{\Big((j+1)!\Big)^{\frac{1}{j}}}$$

$$\le \frac{\left|\ln(a)\right|^{1+\frac{1}{j}}+\left|\ln(b)\right|^{1+\frac{1}{j}}}{\left(\frac{j}{3}\right)^{\frac{j}{j}}}$$

$$= \underbrace{\frac{3}{j}}_{\xrightarrow[j\to\infty]{}0}\cdot\Bigg(\underbrace{\left|\ln(a)\right|^{1+\frac{1}{j}}+\left|\ln(b)\right|^{1+\frac{1}{j}}}_{\xrightarrow[j\to\infty]{}|\ln(a)|+|\ln(b)|\in\mathbb{R}}\Bigg)\xrightarrow[j\to\infty]{}0.$$

Für den Konvergenzradius gilt also $\rho=\infty$. Damit konvergiert die Potenzreihe auf jedem beschränkten Intervall $I\subset\mathbb{R}$ gleichmäßig. Nach Satz 7.11 und Korollar 7.12 gilt dann $h\in C^\infty(I)$ und damit insbesondere $h\in C^\infty(\mathbb{R})$ mit

$$h^{(m)}(0)=\frac{\Big(\ln(a)\Big)^{m+1}-\Big(\ln(b)\Big)^{m+1}}{m+1}.$$

8 Taylor-Entwicklung

Als Einstieg betrachten wir die folgende Vorüberlegung: Es sei $I=(a,b)\subset\mathbb{R}$ ein echtes Teilintervall, $x\in I$ und $f\colon I\to\mathbb{R}$ eine auf dem Punkt x_0 differenzierbare Funktion. Wir haben in der Vorlesung zur Analysis I (Lehramt) gezeigt, dass dann eine Funktion $\varepsilon\colon I\to\mathbb{R}$ existiert, so dass

$$\forall\, x\in I:\quad f(x)+f(x_0)+f'(x_0)(x-x_0)+\varepsilon(x)(x-x_0)\quad\text{und}\quad\lim_{x\to x_0}\varepsilon(x)=0.$$

LINK: TEIL 1 DER 9. VORLESUNG VOM 11.05.2022

Nachtrag 7.14. Es sei $f: I \to \mathbb{C}$ eine komplexwertige Funktion, $I \subset \mathbb{R}$ ein echtes Intervall und $x_0 \in I$.

(a) f heißt *differenzierbar im Punkt* x_0, falls ein $b \in \mathbb{C}$ existiert mit

$$\lim_{x \to x_0} \left| b - \frac{f(x_0) - f(x)}{x_0 - x} \right| = 0.$$

Wir setzen dann $f'(x_0) := b$;

(b) f heißt *differenzierbar auf* I, falls f für alle $x_0 \in I$ differenzierbar im Punkt x_0 ist;

(c) Es seien $g := \mathrm{Re}(f)$ und $h := \mathrm{Im}(f)$, also $f = g + \mathrm{i} \cdot h$. Dann ist f genau dann differenzierbar im Punkt x_0, wenn g und h differenzierbar im Punkt x_0 sind. Dann gilt

$$f'(x_0) = g'(x_0) + \mathrm{i} \cdot h'(x_0);$$

(d) Auch für solche Funktionen f gelten die Summen- und Produktregel sowie der Hauptsatz der Differential- und Integralrechnung, wohingegen dies für den Zwischenwert- und Mittelwertsatz nicht der Fall ist;

(e) Die folgenden Aussagen sind äquivalent:

$$\begin{aligned} & f \text{ ist stetig im Punkt } x_0 \\ \Leftrightarrow & g \text{ und } h \text{ sind stetig im Punkt } x_0 \\ \Leftrightarrow & \text{Für alle Folgen } (x_n)_n \subset I \text{ mit } \lim_{n\to\infty} x_n = x_0 \text{ gilt: } \quad \lim_{n\to\infty} f(x_n) = f(x_0); \end{aligned}$$

(f) Es sei $I = [a, b]$. Dann heißt f *Riemann-integrierbar*, falls g und h Riemann-integrierbar sind. Wir setzen dann

$$\int_a^b f(x)\,\mathrm{d}x = \int_a^b g(x)\,\mathrm{d}x + \mathrm{i} \cdot \int_a^b h(x)\,\mathrm{d}x;$$

(g) Wir setzen

$$\|f\| = \|f\|_I = \sup_{x \in I} |f(x)| \in [0, \infty].$$

Auf der Menge $\mathcal{B}(I, \mathbb{C})$ der beschränkten Funktionen $f: I \to \mathbb{C}$ definiert dies eine Norm. Insbesondere ist $\mathcal{B}(I, \mathbb{C})$ ein Vektorraum.

Also ist für $x \approx x_0$ die Funktion

$$P(x) := f(x_0) + f'(x_0) \cdot (x - x_0)$$

eine gute Approximation für $f(x)$. Dabei ist

- P ein Polynom von Grad 1,
- $\varepsilon(x)(x-x_0)$ der Näherungsfehler und
- es gilt
$$P(x_0) = f(x_0) \quad \text{sowie} \quad P'(x_0) = f'(x_0),$$
d.h. der Funktionswert und die Ableitung von P im Punkt x_0 werden auch von f übernommen.

Nun stellen sich dabei mehrere Fragen:

- Kann man den „abstrakten“ Fehler $\varepsilon(x)(x-x_0)$ berechnen oder abschätzen?
- Kann man sogar eine bessere Approximation (genauer gesagt eine Näherung mit kleinerem Fehler) erreichen, indem man ein Polynom mit höherem Grad wählt?

Dazu betrachten wir eine Funktion $f\colon I \to \mathbb{R}$, die n-mal differenzierbar ist. Wir suchen dann ein Polynom derart, dass sogar

$$P(x_0) = f(x_0), \quad P'(x_0) = f'(x_0), \quad \ldots, \quad P^{(n)}(x_0) = f^{(n)}(x_0) \tag{$\times$}$$

gilt. Dazu machen wir den Ansatz

$$P(x) = \sum_{k=0}^{n} a_k (x-x_0)^k = a_0 + a_1(x-x_0) + \ldots + a_n(x-x_0)^n.$$

Dies ist nach Definition eine (abbrechende) Potenzreihe, nach Satz 7.11 und Korollar 7.12 gilt also

$$a_m = \frac{P^{(m)}(x_0)}{m!} = \frac{f^{(m)}(x_0)}{m!}, \quad m \in \{0, \ldots, n\}.$$

Damit haben wir ein eindeutiges Polynom vom Grad n mit den geforderten Eigenschaften bestimmt

Definition 8.1. Es sei $f\colon I = (a,b) \to \mathbb{R}$ eine n-mal differenzierbare Funktion und $x_0 \in (a,b)$. Dann heißt

$$T_{n,x_0}\colon C^n(I) \to \Big\{p\colon I \to \mathbb{R} \;:\; p \text{ ist ein Polynom}\Big\}, \quad f \mapsto T_{n,x_0} f$$

mit

$$\Big(T_{n,x_0} f\Big)(x) = \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!} \cdot (x-x_0)^k$$

das *n-te Taylorpolynom von f um den Entwicklungspunkt x_0*. Die Abbildung T_{n,x_0} ist außerdem linear für jedes $n \in \mathbb{N}_0$ und $x_0 \in I$.

Beispiel 8.2.

(a) Wir betrachten die Funktion

$$f\colon \mathbb{R}\setminus\{-1\} \to \mathbb{R}, \quad x \mapsto \frac{1}{1+x},$$

sowie den Entwicklungspunkt $x_0 = 0$. Für $x \in \mathbb{R}\setminus\{-1\}$ und $k \in \mathbb{N}_0$ gilt (mit vollständiger Induktion)

$$f^{(k)}(x) = \frac{(-1)^k \cdot k!}{(1+x)^{k+1}} \quad \text{und insbesondere} \quad f^{(k)}(0) = (-1)^k \cdot k!,$$

also

$$\big(T_{n,0}f\big)(x) = \sum_{k=0}^{n} (-1)^k \cdot \frac{k!}{k!} \cdot x^k = \sum_{k=0}^{n} (-1)^k \cdot x^k.$$

Dies ist für $x \in (-1,1)$ eine konvergente Potenzreihe, die auf $(-1,1)$ mit f übereinstimmt: Es gilt

$$\lim_{n\to\infty} \big(T_{n,0}f\big)(x) = \sum_{k\in\mathbb{N}_0} (-x)^k = \frac{1}{1+x} = f(x).$$

(b) Wir betrachten die Funktion

$$g\colon (-1,\infty) \to \mathbb{R}, \quad x \mapsto \ln(1+x),$$

sowie den Entwicklungspunkt $x_0 = 0$. Wegen

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

verwenden wir die Formel aus Teil (a). Für alle $x \in (-1,\infty)$ und $k \in \mathbb{N}_0$ erhalten wir dann

$$g^{(k)}(x) = f^{(k-1)}(x) = \frac{(-1)^{k-1} \cdot (k-1)!}{(1+x)^k}$$

und insbesondere

$$g^{(k)}(0) = (-1)^{k-1} \cdot (k-1)!,$$

also

$$\big(T_{n,0}g\big)(x) = \sum_{k=0}^{n} (-1)^{k-1} \cdot \frac{(k-1)!}{k!} \cdot x^k = \sum_{k=0}^{n} (-1)^k \cdot \frac{x^k}{k}.$$

Es folgt

$$\lim_{n\to\infty} \big(T_{n,0}g\big)(x) = -\sum_{k\in\mathbb{N}_0} \frac{(-x)^k}{k} \tag{$*$}$$

und wir erhalten die Abbildung $x \mapsto -s(-x)$, wobei $s(\cdot)$ die in Beispiel 6.6 (b) betrachtete Potenzreihe, die auf $B_1(0) \supset (-1,1)$ konvergiert.

(c) Wir betrachten die Funktion

$$h\colon \mathbb{R} \to \mathbb{R}, \quad x \mapsto \mathrm{e}^x,$$

sowie den Entwicklungspunkt $x_0 = 0$. Wegen $\exp^{(k)}(x) = \exp(x)$ für alle $k \in \mathbb{N}_0$ folgt

$$\big(T_{n,0}h\big)(x) = \sum_{k=0}^{n} \frac{x^k}{k!} = E_n(x) \xrightarrow[n\to\infty]{} \sum_{k\in\mathbb{N}_0} \frac{x^k}{k!} = \mathrm{e}^x = h(x) \quad \forall\, x \in \mathbb{R}.$$

LINK: TEIL 2 DER 9. VORLESUNG VOM 11.05.2022

Wir sehen, dass das Taylorpolynom $T_{n,x_0}f$ nur dann eine gute Approximation für die Funktion f ist, falls der Restterm bzw. Fehler

$$\big(R_{n,x_0}f\big)(x) := f(x) - \big(T_{n,x_0}f\big)(x)$$

gut abgeschätzt werden kann.

Satz 8.3 (Taylor-Formel)**.** *Es sei $f \in C^{n+1}(I)$, $I \subset \mathbb{R}$ ein offenes Intervall und $x, x_0 \in I$. Dann gelten die folgenden Aussagen:*

(i) *Es gilt*

$$f(x) = \big(T_{n,x_0}f\big)(x) + \big(R_{n,x_0}f\big)(x)$$

mit

$$\big(R_{n,x_0}f\big)(x) = \frac{1}{n!} \cdot \int_{x_0}^{x} (x-t)^n f^{(n+1)}(t)\, \mathrm{d}t.$$

Dies wird auch „Integraldarstellung des Restglieds $R_{n,x_0}f$“ genannt;

(ii) *Es existiert ein ξ zwischen x_0 und x, so dass*

$$\big(R_{n,x_0}f\big)(x) = \frac{f^{(n+1)}(\xi)}{(n+1)!}(x-x_0)^{n+1}.$$

Es wird auch „Restglieddarstellung von Lagrange“ genannt. Es sieht fast so aus wie die Summanden von $T_{n,x_0}f$, nur dass die Ableitung an unbekannter Stelle angenommen wird;

(iii) *Es existiert ein ζ zwischen x_0 und x, so dass*

$$\big(R_{n,x_0}f\big)(x) = \frac{f^{(n+1)}(\zeta)}{n!} \cdot (x-\zeta)^n \cdot (x-x_0).$$

Es wird auch „Cauchy-Restglieddarstellung“ genannt.

Beweis.

(i) Beweis durch vollständige Induktion über $n \in \mathbb{N}_0$.

(IA) $\underline{n=0:}$
Nach dem Hauptsatz der Differential- und Integralrechnung gilt

$$\Big(R_{0,x_0}f\Big)(x) = f(x) - f(x_0) = \int_{x_0}^{x} f'(t)\cdot(x-t)^0\cdot\frac{1}{1!}\,\mathrm{d}t,$$

da $\Big(T_{0,x_0}f\Big)(x) = f(x)$ ein konstanten Polynom ist.

(IV) Die Aussage gilt für ein $n \in \mathbb{N}_0$.

(IS) $\underline{n \mapsto n+1:}$
Für das Restglied der Ordnung n gilt

$$\begin{aligned}\Big(R_{n,x_0}f\Big)(x) &\overset{\text{(IV)}}{=} \frac{1}{n!}\int_{x_0}^{x}(x-t)^n f^{(n+1)}(t)\,\mathrm{d}t\\ &\underset{\text{Int.}}{\overset{\text{part.}}{=}} \frac{1}{n!}\left(-\frac{(x-t)^{n+1}}{n+1}f^{(n+1)}(t)\Big|_{x_0}^{x} + \int_{x_0}^{x}\frac{(x-t)^{n+1}}{n+1}f^{(n+2)}(t)\,\mathrm{d}t\right)\\ &= \frac{(x-t)^{n+1}}{(n+1)!}f^{(n+1)}(x_0) + \underbrace{\frac{1}{(n+1)!}\int_{x_0}^{x}(x-t)^{n+1}f^{(n+2)}(t)\,\mathrm{d}t}_{=\left(R_{n+1,x_0}f\right)(x)},\end{aligned}$$

also

$$\begin{aligned}&f(x) - \Big(T_{n+1,x_0}f\Big)(x)\\ =&\left(f(x) - \Big(T_{n,x_0}f\Big)(x)\right) - \frac{f^{(n+1)}(x_0)}{(n+1)!}\cdot(x-x_0)\\ =&\Big(R_{n,x_0}f\Big)(x) - \frac{f^{(n+1)}(x_0)}{(n+1)!}(x-x_0)^{n+1} \overset{\text{s.o.}}{=} \Big(R_{n-1,x_0}f\Big)(x).\end{aligned}$$

Die Aussage gilt also für alle $n \in \mathbb{N}_0$.

(ii) Wir nehmen o.B.d.A an, dass $x > x_0$ (der Beweis ist $x = x_0$ trivial, während dieser für $x < x_0$ bis auf das Vorzeichen analog verläuft). Man wende nun den zweiten Mittelwertsatz der Integralrechnung (s. Satz 14.21 des Skripts aus der Vorlesung zur Analysis I (Lehramt)) mit $g(t) = (x-t)^n > 0$ für $t > x > x_0$ an. Dann gilt

$$\Big(R_{n,x_0}f\Big)(x) = \frac{1}{n!}\int_{x_0}^{x} g(t)f^{(n+1)}(t)\,\mathrm{d}t$$

$$= \frac{1}{n!} f^{(n+1)}(\underbrace{\xi}_{\in(x_0,x)}) \underbrace{\int_{x_0}^{x} \overbrace{(x-t)^n}^{=g(t)} \, \mathrm{d}t}_{=-\frac{(x-t)^{n+1}}{n+1}\Big|_{x_0}^{x} = \frac{(x-x_0)^{n+1}}{n+1}}$$

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

(iii) Ohne Beweis. (Übung!) □

Die Integral-Restglieddarstellung gilt auch für komplexwertige Funktionen $f\colon I \to \mathbb{C}$, $f \in C^{n+1}(I)$ – für die beiden anderen Darstellungen gilt dies wiederum nicht. Analog zu den Aussagen über differenzierbaren Funktionen hat man:

Satz 8.4. *Es sei $I \subset \mathbb{R}$ ein offenes und echtes Intervall, $f \in C^n(I)$ und $x_0 \in I$. Dann gelten die folgenden Aussagen:*

(i) *Es gilt*

$$f(x) = \Big(T_{n,x_0} f\Big)(x) + \varepsilon(x-x_0),$$

wobei

$$\lim_{x \to x_0} \frac{\varepsilon(x-x_0)}{|x-x_0|^n} = 0.$$

„$\varepsilon(x-x_0)$ geht schneller gegen Null als $|x-x_0|^n$“; man schreibt auch

$$\varepsilon(x-x_0) \ll |x-x_0|^n.$$

Man beachte hierbei, dass wir für diese qualitative, nicht explizite Darstellung des Fehlers eine Differenzierbarkeitsordnung weniger fordern.

(ii) *Es seien $f'(x_0) = \ldots = f^{(n-1)} = 0$, aber $f^{(n)}(x_0) \neq 0$. Dann gelten*

(a) *Falls n gerade ist und $f^{(n)}(x_0) > 0$, so hat f ein lokales Minimum im Punkt x_0;*

(b) *Ist n gerade und $f^{(n)}(x_0) < 0$, so hat f ein lokales Maximum im Punkt x_0;*

(c) *Ist n ungerade, so hat f kein lokales Extremum im Punkt x_0.*

(Man vergleiche die Aussagen mit den notwendigen und hinreichenden Bedingungen für lokale Minima und Maxima bei der Kurvendiskussion.)

Beispiel 8.5. Es sei $I = \mathbb{R}$ und $f := \cos\colon I \to \mathbb{R}$ sowie $x_0 = 0$. Dann gilt

$$\Big(T_{2,0} f\Big)(x) = 1 - \frac{x^2}{2}.$$

Zur Bestimmung des Restglieds ist zu beachten, dass ein ξ zwischen 0 und x existiert, so dass

$$\Big(R_{2,0}f\Big)(x) = f'''(\xi)\frac{x^3}{3!} = \sin(\xi)\cdot\frac{x^3}{3!},$$

also

$$\Big|\Big(R_{2,0}f\Big)(x)\Big| \leq \frac{|x|^3}{3!}, \quad \text{da } |\sin(\xi)| \leq 1.$$

Es gilt aber auch $|\sin(\xi)| \leq |\xi| \leq |x|$, so dass

$$\Big|\Big(R_{2,0}f\Big)(x)\Big| \leq \frac{|x|^4}{6}.$$

Die Abschätzung lässt sich weiter verbessern, falls man $f'''(0) = \sin(0) = 0$ beachtet. Damit gilt

$$T_{3,0}f = T_{2,0}f, \quad \text{also} \quad R_{3,0}f = R_{2,0}f.$$

Daraus folgt

$$\Big|\Big(R_{2,0}f\Big)(x)\Big| = \Big|\Big(R_{3,0}f\Big)(x)\Big| \leq |\cos(\zeta)|\cdot\frac{|x|^4}{4!} \leq \frac{|x|^4}{24},$$

wobei ζ eine Stelle zwischen 0 und x ist.

LINK: TEIL 1 DER 10. VORLESUNG VOM 12.05.2022

Taylor-Reihen

Im Rest dieses Kapitels nehmen wir an, dass $f\colon I \to \mathbb{R}$ eine Funktion mit $f \in C^\infty(I)$ für ein offenes Intervall $I \subset \mathbb{R}$ und dass $x_0 \in I$ fixiert ist. Offensichtlich bilden die n-ten Taylorpolynome von f eine Folge

$$\Big(T_{n,x_0}f\Big)_{n\in\mathbb{N}_0} \tag{8.1}$$

von Partialsummen. Dazu können wir uns die folgenden Fragen stellen:

(i) Konvergiert (8.1) gegen eine Potenzreihe?

(ii) Falls ja, ist die Reihe dann eine Darstellung von f?

Definition 8.6. Die Potenzreihe

$$\sum_{k\in\mathbb{N}_0} \frac{f^{(k)}(x_0)}{k!}(x-x_0)^k$$

heißt *Taylor-Reihe von f um den Entwicklungspunkt* x_0. Existiert ein $\delta > 0$, so dass diese Reihe für $x \in B_\delta(x_0) \subset \mathbb{R}$ konvergiert, so schreibt man für den Summenwert

$$\Big(T_{x_0}f\Big)(x) := \sum_{k\in\mathbb{N}_0} \frac{f^{(k)}(x_0)}{k!}(x-x_0)^k, \quad x \in B_\delta(x_0). \tag{$\square$}$$

Gilt

$$\Big(T_{x_0}f\Big)(x) = f(x) \quad \forall x \in B_\delta(x_0),$$

so sagen wir, dass sich f um x_0 *in eine Taylor-Reihe entwickeln* oder *als Taylor-Reihe darstellen lässt.*

Wir sind insbesondere daran interessiert, auf welchen Intervallen die Konvergenz in (□) uniform ist; denn da dürfen wir gliedweise differenzieren und integrieren.

Die Untersuchung sowohl der punktweisen als auch der gleichmäßigen Konvergenz in (□) erfolgt über Abschätzungen des Restglieds $R_{n,x_0}f$.

Lemma 8.7. *Sind $x_0 \in \mathbb{R}$ und $a_k \in \mathbb{R}$, $k \in \mathbb{N}$, derart, dass die Potenzreihe*

$$\sum_{k\in\mathbb{N}} a_k(x-x_0)^k \tag{PR}$$

einen positiven Konvergenzradius ρ besitzt, so sind die a_k gemäß Korollar 7.12 gegeben durch

$$a_k = \frac{f^(k)(x_0)}{k!} \quad \forall\, k \in \mathbb{N},$$

wobei f die unendlich oft differenzierbare Funktion

$$B_\delta(x_0) \ni x \mapsto f(x) = \sum_{k\in\mathbb{N}} a_k(x-x_0)^k$$

ist. Damit ist (PR) die Taylor-Reihe von f und die Taylorpolynome die zugehörigen Partialsummen, es gilt also $s_n = T_{n,x_0}f$.

Beweis. Die Aussage folgt direkt aus Satz 7.8 und Korollar 7.12. □

Dies kann man nun in zwei Richtungen anwenden:

1. Falls man a priori weiß, dass f eine Potenzreihe besitzt, muss man die Ableitungen $f^{(k)}(x_0)$ von f berechnen und zeigen, dass $\rho > 0$ gilt. Dadurch erhält man eine Taylorentwicklung und damit die Potenzreihendarstellung von f.

2. Falls eine Potenzreihendarstellung von f mit positivem Konvergenzradius $\rho > 0$ gegeben ist, kann man daraus die Ableitungen $f^{(k)}(x_0)$ ablesen.

Beispiel 8.8.

(a) Für $f = \exp\colon \mathbb{R} \to \mathbb{R}$ ist bekannterweise $f^{(n)} = f$ auf $\mathbb{R}$ für alle $n \in \mathbb{N}$. Für $x_0 = 0$ und alle $x \in \mathbb{R}$ sowie ξ zwischen 0 und x gilt daher

$$\left|\left(R_{n,0}f\right)(x)\right| = \frac{\mathrm{e}^\xi}{(n+1)!}|x|^{n+1} \le \mathrm{e}^{|x|}\underbrace{\frac{|x|^{n+1}}{(n+1)!}}_{\substack{\longrightarrow 0\\ n\to\infty}} \le \mathrm{e}^r\frac{r^{n+1}}{(n+1)!} \underset{n\to\infty}{\longrightarrow} 0,$$

sofern $x \in \overline{B_r(0)} = [-r, r]$. Wir haben folglich sogar uniforme Konvergenz auf jedem beschränkten Intervall $I \subset \mathbb{R}$ der Reihe:

$$\mathrm{e}^x = \sum_{k\in\mathbb{N}_0} \frac{f^{(k)}(0)}{k!}x^k = \sum_{k\in\mathbb{N}_0} \frac{x^k}{k!}.$$

(b) Für $f = \sin\colon \mathbb{R} \to [-1,1]$ und $x_0 = 0$ gilt

$$f^{(2k)}(x) = (-1)^k \sin(x) \quad \Rightarrow \quad f^{(2k)}(0) = 0$$

und

$$f^{(2k+1)}(x) = (-1)^k \cos(x) \quad \Rightarrow \quad f^{(2k+1)}(0) = (-1)^k$$

für alle $k \in \mathbb{N}_0$. Damit erhalten wir

$$\big(T_0 f\big)(x) = \sum_{j\in\mathbb{N}_0} \frac{(-1)^j}{(2j+1)!} x^{2j+1} \tag{$T\sin$}$$

als formale Potenzreihe, wobei noch der Konvergenzbereich zu klären ist. Um den Restterm abzuschätzen, beachten wir zunächst, dass

$$|\cos(x)| \leq 1 \quad \text{und} \quad |\sin(x)| \leq 1 \quad \forall\, x \in \mathbb{R}.$$

Daraus folgt

$$\Big|\big(R_{2k+1,0} f\big)(x)\Big| = \Big|\big(R_{2k+2,0} f\big)(x)\Big| \leq \frac{|x|^{2k+3}}{(2k+3)!} \xrightarrow[k\to\infty]{} 0.$$

Es gilt also $\rho = \infty$ – der Konvergenzbereich ist also ganz $\mathbb{R}$. Damit ist der Ausdruck in ($T\sin$) tatsächlich überall konvergent und stimmt gemäß Lemma 8.7 mit der Sinus-Funktion überein. Betrachten wir eine konkrete Beispiel-Abschätzung: Für $|x| \leq 1$ und $k \geq 4$ gilt

$$\Big|\big(R_{2k+1,0} f\big)(x)\Big| \leq \frac{1}{(2k+3)!} \leq \frac{1}{11!} < 10^{-7}.$$

(c) Ganz analog kann man zeigen, dass die Taylorentwicklung von

$$f := \cos\colon \mathbb{R} \to \mathbb{R}$$

um den Entwicklungspunkt $x_0 = 0$, also

$$\big(T_0 f\big)(x) = \sum_{k\in\mathbb{N}_0} \frac{(-1)^k}{(2k)!} x^{2k} = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + - \dots$$

auf ganz $\mathbb{R}$ konvergiert. Da bereits bekannt ist, dass die Kosinus-Funktion eine konvergente Potenzreihenentwicklung um $x_0 = 0$ besitzt, folgt mit Lemma 8.7, dass $T_0 f$ mit f übereinstimmt.

Definition 8.9. Für $\alpha \in \mathbb{R}\backslash\mathbb{N}_0$ definieren wir den *verallgemeinerten Binomialkoeffizienten* durch

$$\binom{\alpha}{0} := 1 \quad \text{und} \quad \binom{\alpha}{k} := \frac{\alpha\cdot(\alpha-1)\cdot\ldots\cdot(\alpha-k+1)}{k!} \quad \text{für } k \in \mathbb{N}.$$

LINK: TEIL 2 DER 10. VORLESUNG VOM 12.05.2022

Es gilt dann die Verallgemeinerung der binomischen Formel:

Satz 8.10 (Binomische Reihe)**.** *Für $x \in (-1,1)$ und $\alpha \in \mathbb{R} \setminus \mathbb{N}_0$ gilt*

$$\sum_{k \in \mathbb{N}_0} \binom{\alpha}{k} \cdot x^k = (1+x)^\alpha.$$

Beweis und Diskussion. Übung!

Beispiel 8.11. In der Übung zeigen Sie, dass für $x \in (-1,1)$ die die Arkustangens-Funktion die folgende Potenzreihendarstellung besitzt:

$$\arctan(x) = \sum_{k \in \mathbb{N}_0} \frac{(-1)^k}{2k+1} x^{2k+1}. \tag{ATR}$$

Mit Lemma 8.7 bzw. Satz 7.11 und Korollar 7.12 folgern wir daraus

$$\frac{(-1)^k}{2k+1} = a_{2k+1} = \frac{\arctan^{(2k+1)}(0)}{(2k+1)!} \quad \text{und} \quad 0 = a_{2k} = \frac{\arctan^{(2k)}(0)}{(2k)!},$$

also

$$\arctan^{(n)}(0) = \begin{cases} 0 & \text{für } n \text{ gerade,} \\ (-1)^k (n-1)! & \text{für } n = 2k+1 \text{ ungerade.} \end{cases}$$

Falls schon a priori bekannt ist, dass die Taylor-Reihe von arctan auf $(-1,1)$ konvergiert, könnte man – umgekehrt – die Identität in (ATR) durch die Berechnung aller Ableitungen im Punkt $x_0 = 0$ beweisen. Uns interessiert nun die folgende Fehlerabschätzung: Für $x \in [-1,1]$ ist nämlich $\left(\frac{x^{2k+1}}{2k+1}\right)_{k \in \mathbb{N}_0}$ eine monotone Nullfolge. Also ergibt die Fehlerschranke zum Leibnitz-Kriterium

$$\left| \arctan(x) - \sum_{k=0}^{n} \frac{(-1)^k}{2k+1} x^{2k+1} \right| \le \frac{|x|^{2n+3}}{2n+3} \overset{|x| \le 1}{\le} \frac{1}{2n+3} \xrightarrow[n \to \infty]{} 0 \quad \text{uniform in } [-1,1].$$

Setze nun $x = 1$ ein. Wegen $\sin\left(\frac{\pi}{4}\right) = \cos\left(\frac{\pi}{4}\right)$ gilt $\tan\left(\frac{\pi}{4}\right) = 1$, also $\arctan(1) = \frac{\pi}{4}$. Daraus folgt

$$\left| \frac{\pi}{4} - \sum_{k=0}^{n} \frac{(-1)^k}{2k+1} \right| \le \frac{1}{2n+3} \quad \text{und} \quad \pi = 4 \sum_{k \in \mathbb{N}_0} \frac{(-1)^k}{2k+1}.$$

Für die Bestimmung des Konvergenzradius der Reihe in (ATR) beachte man, dass

$$\sqrt[2k+1]{2k+1} = \mathrm{e}^{\frac{\ln(2k+1)}{2k+1}} \xrightarrow[k \to \infty]{} 1.$$

Also konvergiert die Reihe für alle $x \in (-1,1)$ – sie konvergiert sogar für alle $z \in B_1 \subset \mathbb{C}$. Dagegen konvergiert sie nicht für $|x| > 1$, obwohl $\arctan \in C^\infty(\mathbb{R})$ gilt! Damit stellt die Potenzreihe die Arkustangens-Funktion nur <u>lokal</u> dar.

Was läuft denn beim Konvergenzradius $\rho = 1$ schief? Man betrachte hierzu die Randpunkte $\pm\mathrm{i} \in \partial B_1$. Man erkennt nämlich, dass der Ausdruck

$$\sum_{k \in \mathbb{N}_0} \frac{(-1)^k (\pm\mathrm{i})^{2k+1}}{2k+1} = \pm\mathrm{i} \sum_{k \in \mathbb{N}_0} \frac{(-1)^k \Big(\overbrace{(\pm\mathrm{i})^2}^{=-1} \Big)^k}{2k+1} = \pm\mathrm{i} \sum_{k \in \mathbb{N}_0} \frac{1}{2k+1}$$

divergiert.

Bemerkung 8.12 (Vorsicht: Pathologien)**.**

(a) Es gibt Funktionen $f \in C^\infty(\mathbb{R})$, deren Taylor-Reihe ausschließlich in $x = x_0$ konvergiert.

(b) Es existiert ein Satz von Borel, der besagt, dass zu jeder Folge $(a_k)_k \subset \mathbb{R}$ eine Funktion $f \in C^\infty(\mathbb{R})$ existiert mit

$$f^{(k)}(0) = a_k \quad \forall\, k \in \mathbb{N}.$$

(a_k kann jedoch so schnell wachsen, dass $\rho = 0$ gilt.)

(c) Es gibt Funktionen $f \in C^\infty(\mathbb{R})$, deren Taylor-Reihe auf ganz $\mathbb{R}$ konvergiert – jedoch gegen eine <u>andere</u> Funktion g! Betrachen wir beispielsweise die Funktion

$$h\colon \mathbb{R} \to [0,1), \quad h(x) := \begin{cases} 0 & \text{für } x < 0, \\ \mathrm{e}^{-\frac{1}{x^2}} & \text{für } x \geq 0. \end{cases}$$

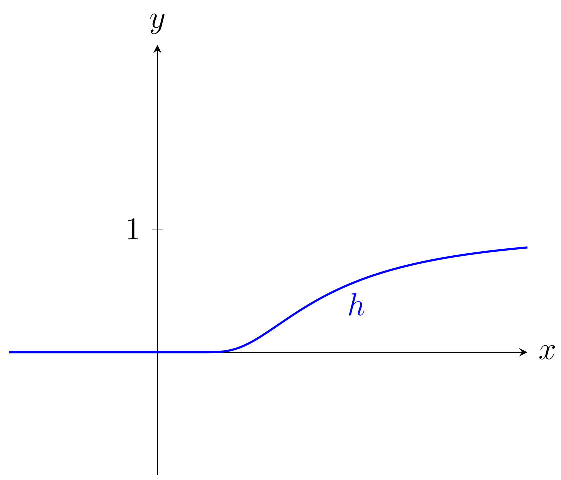

Abbildung 7: Der Graph von h in blau.

Offensichtlich gilt $h \in C^\infty(\mathbb{R} \setminus \{0\})$. Induktiv zeigt man außerdem, dass

$$h \in C^k(\mathbb{R}) \quad \text{und} \quad h^{(k)}(0) = 0 \quad \forall\, k \in \mathbb{N}_0.$$

Damit ist die Taylor-Reihe von h zum Entwicklungspunkt $x_0 = 0$ gegeben durch

$$\big(T_0 h\big)(x) = \sum_{k \in \mathbb{N}_0} 0 = 0.$$

Die Taylor-Reihe konvergiert also gegen die Funktion $g \equiv 0$ – sogar auf ganz $\mathbb{R}$.

Beispiel 8.13 (Fehlerfunktion)**.** Wir betrachten die Funktion

$$\Phi\colon \mathbb{R} \to \mathbb{R}, \quad \Phi(x) = \int_0^x \frac{2}{\sqrt{\pi}} \mathrm{e}^{-t^2}\, \mathrm{d}t.$$

Wegen der gleichmäßigen Konvergenz der Exponentialreihe auf $[0, x]$ gilt dann

$$\Phi(x) = \frac{2}{\sqrt{\pi}} \int_0^x \sum_{k \in \mathbb{N}_0} \frac{(-1)^k}{k!} t^{2k}\, \mathrm{d}t = \frac{2}{\sqrt{\pi}} \sum_{k \in \mathbb{N}_0} \int_0^x \frac{(-1)^k}{k!} t^{2k}\, \mathrm{d}t = \frac{2}{\sqrt{\pi}} \sum_{k \in \mathbb{N}_0} \frac{(-1)^k}{k!} \frac{t^{2k+1}}{2k+1}.$$

Betrachten wir beispielsweise den Punkt $x = 1$, so erhalten wir mit der Fehlerabschätzung aus dem Leibnitz-Kriterium die folgende Identität:

$$\begin{aligned} \Phi(1) &= \frac{2}{\sqrt{\pi}} \sum_{k \in \mathbb{N}_0} \frac{(-1)^k}{k!(2k+1)} \\ &= \frac{2}{\sqrt{\pi}} \cdot \left(1 - \frac{1}{3} + \frac{1}{10} - \frac{1}{216} + \frac{1}{1320} - \frac{1}{3360} + - \dots\right) \\ &= \frac{2}{\sqrt{\pi}} \left(0,762497 \pm 1,4 \cdot 10^{-5}\right) \\ &= 0,860386 \pm 2 \cdot 10^{-5}. \end{aligned}$$

LINK: TEIL 1 DER 11. VORLESUNG VOM 18.05.2022

Bemerkung (Normalverteilung und Fehlerfunktion)**.** Die Funktion

$$f\colon \mathbb{R} \to [0, \infty), \quad x \mapsto f(x) := \frac{1}{\sqrt{2\pi}} \cdot \mathrm{e}^{-\frac{x^2}{2}}$$

ist die Dichte der Standardnormalverteilung (s. Stochastik).

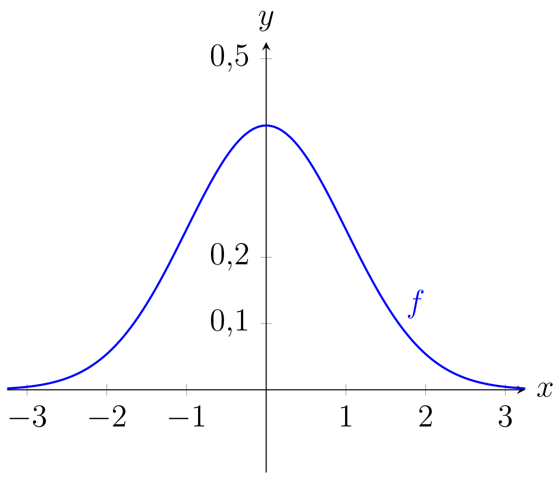

Abbildung 8: Der Graph der Dichte der Standardnormalverteilung f in blau. Sie wird auch *gaußsche Glockenkurve* genannt.

Die Stammfunktion F (als uneigentliches Integral) gegeben durch

$$F(y) := \int_{-\infty}^{y} f(x)\, \mathrm{d}x = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{y} \mathrm{e}^{-\frac{x^2}{2}}\, \mathrm{d}x$$

heißt *Verteilungsfunktion* der Standardnormalverteilung. Im Kontext der Stochastik interpretiert man dies als Wahrscheinlichkeit

$$\mathbb{P}(X \leq y) = F(y),$$

wobei X eine standardnormalverteile Zufallsgröße ist.

Die Funktion Φ aus Bemerkung 8.13 wird auch *Gaußsche Fehlerfunktion* genannt und mit $\Phi = \operatorname{erf}$ gekennzeichnet. Sie ist eine Stammfunktion von $t \mapsto \frac{2}{\sqrt{\pi}}\mathrm{e}^{-t^2}$ – und zwar einzige ungerade.

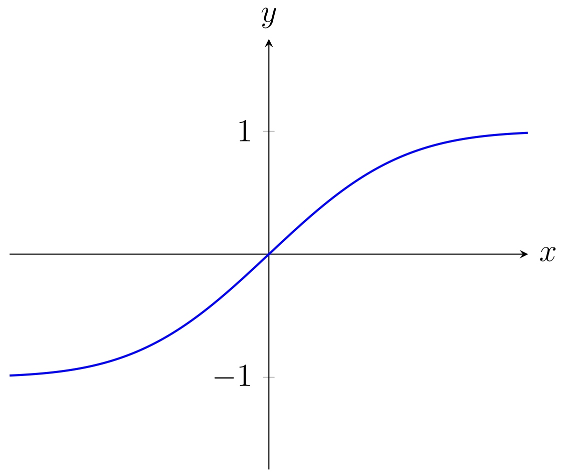

Abbildung 9: Der Graph der Fehlerfunktion $y \mapsto \Phi(y) = \operatorname{erf}(y)$.

Sie besitzt die folgende Beziehung zu der Stammfunktion der Dichte der Standardnormalverteilung:

$$F(y) = \frac{1}{2}\left(1 + \operatorname{erf}\left(\frac{y}{\sqrt{2}}\right)\right).$$

Dabei ist $\operatorname{erf}\left(\frac{y}{\sqrt{2}}\right)$ die Wahrscheinlichkeit, dass der Wert einer standardnormalverteilten Zufallsgröße im Intervall $(-y, y)$ liegt.

9 Funktionen im $\mathbb{R}^n$

Die Vektorräume $\mathbb{R}^n$ und $\mathbb{C}^n$, $n \in \mathbb{N}$, sind bereits aus der Linearen Algebra bekannt. Dabei ist

$$x = (x_1, \ldots, x_n) = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n$$

ein n-Tupel reeller Zahlen. Dabei gibt es zwei geometrische Interpretationen:

- Geometrische Interpretation als *Ortskoordinaten* eines Punktes $p \in \mathbb{R}^n$, beispielsweise innerhalb einer Ebene für den Fall $n = 2$, ansonsten im n-dimensionalen Raum $\mathbb{R}^n$;

- Geometrische Interpretation als *Translation / Parallelverschiebung* des Raumes $\mathbb{R}^n$: Dies entspricht einer (affinen) bijektiven Abbildung

  $$v\colon \mathbb{R}^n \to \mathbb{R}^n, \quad (p_1, \ldots, p_n) \mapsto (p_1 + v_1, \ldots, p_n + v_n).$$

  Der Vektor v ist hier durch seine Wirkung auf einen beliebigen Punkt $p \in \mathbb{R}^n$ festgelegt, insbesondere kann v als gerichtete Strecke $\vec{pq}$ dargestellt werden, falls $q = v(p)$. Ist $p \in \mathbb{R}^n$ fixiert, so liefert dann

  $$v \mapsto v(p)$$

  eine Bijektion zwischen Vektoren und Punkten im $\mathbb{R}^n$.

Für $X \mapsto \mathbb{R}^n$ betrachten wir Funktionen

$$f\colon X \to \mathbb{R}, \quad x = (x_1, \ldots, x_n) \mapsto f(x) = f(x_1, \ldots, x_n)$$

mehrerer (reeller) Veränderlichen. Für $n = 1$ ist schon der *Funktionsgraph*

$$G_f = \Big\{ (\underbrace{x}_{\in \mathbb{R}^n}, f(x)) \in \mathbb{R}^{d+1} \ : \ x \in X \Big\}$$

von f bekannt. Für $n = 2$ würde G_f einer (gekrümmten) Fläche gleichen, falls f „schön“ genug ist.

Beispiel. Wir betrachten verschiedene Funktionen $f\colon \mathbb{R}^2 \to \mathbb{R}$. Dazu seien

$$P(x, y) = x^2 + 2y^2 \quad \text{und} \quad Q(x_1, x_2) = x_1 \cdot x_2.$$

Wir bezeichnen außerdem die Menge

$$\mathcal{N}_0(f) := \Big\{ x \in X \ : \ f(x) = 0 \Big\}$$

als *Nullstellenmenge von f* oder allgemeiner

$$\mathcal{N}_\alpha(f) := \Big\{ x \in X \ : \ f(x) = \alpha \Big\}$$

die *α-Niveaumenge von f* zu $\alpha \in \mathbb{R}$. So werden beispielsweise auf Landkarten

- Höhenlinien – wobei $f\colon \mathbb{R}^2 \to \mathbb{R}$ die Höhe über N.N. kennzeichnet –,

- Isobare mit $f\colon \mathbb{R}^2 \to \mathbb{R}$ als Luftdruck und

- Isotherme mit $f\colon \mathbb{R}^2 \to \mathbb{R}$ als Temperatur dargestellt.

Zu $\alpha > 0$ ist $\mathcal{N}_\alpha(P)$ ellipsenförmig, während $\mathcal{N}_\alpha(Q)$ hyperbelförmig ist.

Was gilt für $\alpha = 0$ und $\alpha < 0$?

Fügt man mehrere Funktionen $f_1, \ldots, f_m \colon \mathbb{R}^n \to \mathbb{R}$ zu einem Vektor

$$f := (f_1, \ldots, f_m)$$

zusammen, so erhält man eine Abbildung

$$f \colon \mathbb{R}^n \to \mathbb{R}^m, \quad (x_1, \ldots, x_n) \mapsto f(x) = \begin{pmatrix} f_1(x_1, \ldots, x_n) \\ \vdots \\ f_m(x_1, \ldots, x_n) \end{pmatrix}.$$

Beispiel 9.1.

(a) Aus der Linearen Algebra sind lineare Abbildungen $L \colon \mathbb{R}^n \to \mathbb{R}^m$ bekannt, die bezüglich einer Basis als Matrix $A \in \mathbb{R}^{m \times n}$ dargestellt werden kann. Ist insbesondere L bijektiv, kann man dies als Koordinatentransformation des Raumes (Basiswechsel) auffassen.

(b) Die Abbildung

$$P \colon \mathbb{R}^3 \to \mathbb{R}^2, \quad (x_1, x_2, x_3) \mapsto (x_1, x_2)$$

ist eine Projektion von $\mathbb{R}^3$ auf $\mathbb{R}^2$. Auf diese Weise können beispielsweise Längen- und Breitengrade einer Stadt auf einer Landkarte dargestellt werden, wobei die Höhe x_3 weggelassen wird.

Betrachten wir nun die (lineare) Abbildung

$$P \colon \mathbb{R}^2 \to \mathbb{R}^2, \quad (x, y) \mapsto \frac{1}{2}(x + y, x + y).$$

Diese projiziert einen beliebigen Punkt $(x, y) \in \mathbb{R}^2$ auf die Diagonale $\frac{1}{2}(x+y, x+y) \in \mathbb{R}^2$. Insbesondere gilt $f(a, a) = (a, a)$ und $f(a, -a) = (0, 0)$ für alle $a \in \mathbb{R}$.

(c) Eine Abbildung $f \colon \mathbb{C} \to \mathbb{C}$ können wir mittels der Identifikation

$$\mathbb{C} \sim \mathbb{R}^2, \quad z = x + \mathrm{i}y \sim (x, y)$$

auch als Abbildung $\mathbb{R}^2 \to \mathbb{R}^2$ aufgefasst werden.

(d) Die Abbildung $z \mapsto z^2$ entspricht der Abbildung

$$x + \mathrm{i}y \mapsto (x + \mathrm{i}y)^2 = (x^2 - y^2) + \mathrm{i} \cdot 2xy.$$

Dies liefert dann die Abbildung

$$f \colon \mathbb{R}^2 \to \mathbb{R}^2, \quad f(x, y) = \begin{pmatrix} x^2 - y^2 \\ 2xy \end{pmatrix}.$$

(e) Die Abbildung $z \mapsto \mathrm{e}^z$ liefert die Abbildung

$$g\colon \mathbb{R}^2 \to \mathbb{R}^2, \quad g(x,y) = \begin{pmatrix} \mathrm{e}^x \cdot \cos(y) \\ \mathrm{e}^x \cdot \sin(y) \end{pmatrix},$$

da ja

$$\mathrm{e}^z = \mathrm{e}^{x+\mathrm{i}y} = \mathrm{e}^x \cdot (\cos(y) + \mathrm{i}\sin(y)).$$

(f) Für $\alpha \in [0, 2\pi]$ sei

$$D_\alpha\colon \mathbb{C} \to \mathbb{C}, \quad z \mapsto \mathrm{e}^{\mathrm{i}\alpha} z.$$

Diese Abbildung beschreibt die Drehung um den Winkel α. Wegen

$$\Big(\cos(\alpha) + \mathrm{i}\sin(\alpha)\Big)(u + \mathrm{i}v) = u\cos(\alpha) - v\sin(\alpha) + \mathrm{i}\Big(v\cos(\alpha) + u\sin(\alpha)\Big)$$

entspricht dies der Abbildung

$$D_\alpha\colon \mathbb{R}^2 \to \mathbb{R}^2, \quad (u,v) \mapsto D_\alpha(u,v) = \begin{pmatrix} u\cos(\alpha) - v\sin(\alpha) \\ v\cos(\alpha) + u\sin(\alpha) \end{pmatrix}.$$

Speziell für $\alpha = \frac{\pi}{4}$ ist $\sin(\alpha) = \cos(\alpha) = \frac{\sqrt{2}}{2}$ und damit

$$D_{\frac{\pi}{4}}(u,v) = \frac{1}{\sqrt{2}} \begin{pmatrix} u - v \\ v + u \end{pmatrix}.$$

LINK: TEIL 2 DER 11. VORLESUNG VOM 18.05.2022

Unter Einbezug der Abbildung Q (s. oben) gilt außerdem

$$\Big(Q \circ D_{\frac{\pi}{4}}\Big)(u,v) = Q\left(\frac{1}{\sqrt{2}}(u - v, v + u)\right) = \left(\frac{1}{\sqrt{2}}\right)^2 (u^2 - v^2) = \frac{u^2 - v^2}{2}.$$

In den „neuen Koordinaten“ wird der gemischte Term des Polynoms $Q(x,y) = x \cdot y$ transformiert in eine Summe von zwei Polynomen jeweils einer Variablen. Da D_α bijektiv ist, darf man sie außerdem als Koordinatentransformation

$$\Psi\colon U \to X \quad X, U \subset \mathbb{R}^n,$$

auffassen. (vgl. Teil (a))

(g) *Polarkoordinaten der Ebene*:
Wir betrachten zunächst die folgende Vorüberlegung im Komplexen: Bekannterweise existiert für alle $z \in \mathbb{R}$ ein $r \geq 0$ und $\vartheta \in \mathbb{R}$ mit

$$z = r\Big(\cos(\vartheta = +\mathrm{i}\sin(\vartheta)\Big) = r \cdot \mathrm{e}^{\mathrm{i}\vartheta},$$

wobei $z \mapsto r = |z|$ eindeutig ist. Ist $z \neq 0$ und $\vartheta \in (-\pi, \pi]$, dann ist auch $z \mapsto \vartheta$ eindeutig (ansonsten nicht). Für $X = \mathbb{R}^2 \setminus \{(0,0)\}$, $U := (0,\infty) \times (-\pi, \pi] \ni (r, \varphi)$ betrachte man die Abbildung

$$\Psi(r,\varphi) = \begin{pmatrix} r\cos(\varphi) \\ r\sin(\varphi) \end{pmatrix}.$$

Die inverse Abbildung ist

$$\Psi^{-1}(x,y) = \begin{pmatrix} \sqrt{x^2+y^2} \\ \operatorname{sgn}(y) \cdot \arccos\left(\frac{x}{\sqrt{x^2+y^2}}\right) \end{pmatrix} = \begin{cases} \begin{pmatrix} \sqrt{x^2+y^2} \\ \arccos\left(\frac{x}{\sqrt{x^2+y^2}}\right) \end{pmatrix} & \text{für } y \geq 0, \\ \begin{pmatrix} \sqrt{x^2+y^2} \\ -\arccos\left(\frac{x}{\sqrt{x^2+y^2}}\right) \end{pmatrix} & \text{für } y < 0. \end{cases}$$

Test: Für $x \neq 0$, $y = 0$ erkennt man dann, dass

$$\begin{aligned} \Psi^{-1}(x,0) = \begin{pmatrix} \sqrt{x^2} \\ 1 \cdot \arccos\left(\frac{x}{\sqrt{x^2}}\right) \end{pmatrix} &= \begin{pmatrix} |x| \\ \arccos\left(\frac{x}{|x|}\right) \end{pmatrix} \\ &= \begin{pmatrix} r \\ \arccos\left(\operatorname{sgn}(x)\right) \end{pmatrix} \\ &= \begin{cases} \left(r, \arccos(1)\right) = (r,0) & \text{für } x > 0, \\ \left(r, \arccos(-1)\right) = (r,\pi) & \text{für } x < 0. \end{cases} \end{aligned}$$

Betrachten wir nun weitere Spezialfälle bzw. ihre Terminologie:

- Funktionen $f\colon \mathbb{R} \to \mathbb{R}^m$ nennt man *Kurven im* $\mathbb{R}^m$, für $m = 3$ heißen solche Funktionen $f\colon \mathbb{R} \to \mathbb{R}^3$ *Kurven im Raum*,
- Funktionen $f\colon \mathbb{R}^2 \to \mathbb{R}^3$ nennt man *Flächen im Raum* und
- Funktionen $f\colon \mathbb{R}^n \to \mathbb{R}^n$ nennt man *Vektorfelder*, z.B. Eisenspäne im Feld eines Magnets.

10 Normen und Metriken

Definition 10.1. Es sei V ein Vektorraum über $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$. Eine Abbildung

$$\|\cdot\|\colon V \to \mathbb{K},$$

die die Eigenschaften

(N1) *(Definitheit)*
Es gilt
$$\|x\| \geq 0 \quad \forall\, x \in V, \quad \|x\| = 0 \quad \Leftrightarrow \quad x = \mathbf{0} \in V.$$

(N2) *(Homogenität)*
Es gilt
$$\|\alpha x\| = |\alpha| \cdot \|x\| \quad \forall\, x \in V, \alpha \in \mathbb{K}.$$

(N3) *(Dreiecksungleichung)*
Es gilt
$$\|x + y\| \leq \|x\| + \|y\| \quad \forall\, x, y \in V.$$

erfüllt, bezeichnet man als *Norm auf* V. Das Tupel $(V, \|\cdot\|)$ heißt dann *normierter Raum*.

Es folgt direkt
$$\Big|\|x\| - \|y\|\Big| \leq \|x - y\| \quad \forall\, x, y \in V.$$

Beispiel 10.2.

(a) Ist V ein euklidischer Vektorraum, d.h. ein Vektorraum über $\mathbb{R}$ mit dem Skalarprodukt
$$\langle \cdot, \cdot \rangle \colon V \times V \to \mathbb{R},$$
so definiert
$$x \mapsto \sqrt{\langle x, x \rangle}$$
eine Norm auf V. (Übung!)

Es gilt insbesondere die *Cauchy-Schwarz-Ungleichung*:
$$\forall\, x, y \in V: \quad |\langle x, y \rangle| \leq \|x\| \cdot \|y\| \leq \frac{\|x\|^2 + \|y\|^2}{2}.$$

(b) Wir betrachten den Spezialfall $V = \mathbb{R}^n$, $n \in \mathbb{N}$: Dann induziert das Skalarprodukt
$$\langle x, y \rangle = \sum j = 1^n x_j \cdot y_j \quad x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \in \mathbb{R}^n$$
die Norm
$$\|x\| = \|x\|_2 = \left(\sum_{j=1}^{n} x_j^2 \right)^{\frac{1}{2}}.$$

(c) Auf $V = \mathbb{K}^n$, $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$, gibt es auch andere Normen wie beispielsweise die 1-Norm bzw. Betragssummennorm
$$\|x\|_1 = \left(\sum_{j=1}^{n} |x_j| \right)^1 = \sum_{j=1}^{n} |x_j|$$

oder allgemeiner die p-Norm

$$\|x\|_p = \left(\sum_{j=1}^{n} |x_j|^p\right)^{\frac{1}{p}} \qquad \text{für } p \in [1, \infty).$$

$\|x\|_p$ verallgemeinert also die beiden Normen $\|x\|_1$ und $\|x\|_2$.

(d) Ergänzend definiert man auf $V = \mathbb{K}^n$ die Norm

$$\|x\|_\infty := \max\Big\{|x_1|, \dots, |x_n|\Big\}$$

und bezeichnet sie als *Maximumsnorm*. Insbesondere gilt

$$\|x\|_\infty \leq \|x\|_2 \leq \sqrt{n} \cdot \|x\|_\infty,$$

$$\|x\|_\infty \leq \|x\|_1 \leq n \cdot \|x\|_\infty,$$

$$\|x\|_2 \leq \|x\|_1 \leq \sqrt{n} \cdot \|x\|_2.$$

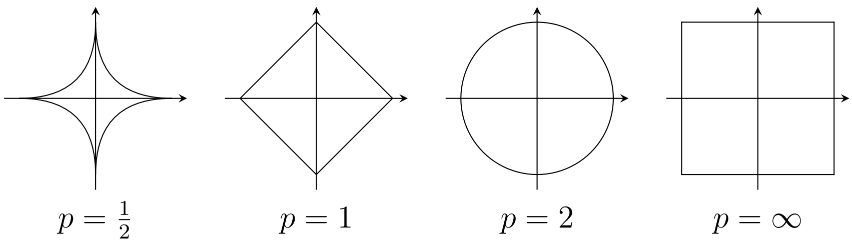

Abbildung 10: Illustrationen der Einheitskreislinie bezüglich der p-Norm für $p = \frac{1}{2}$, $p = 1$, $p = 2$ sowie $p = \infty$ in $\mathbb{R}^2$; beschrieben durch $\Big\{x \in \mathbb{R}^2 \ : \ \|x\|_p = 1\Big\}$.

Diese Eigenschaft nennt man auch *Äquivalenz zwischen Normen*. Allgemein schreibt man: Es sei V ein endlich-dimensionaler Vektorraum, auf dem zwei Normen $\|\cdot\|_1$ und $\|\cdot\|_2$ definiert sind. Die Normen $\|\cdot\|_1$ und $\|\cdot\|_2$ heißen dann *äquivalent*, falls $C_1 > 0$ und $C_2 > 0$ existieren, so dass

$$C_1\|x\|_2 \leq \|x\|_1 \leq C_2\|x\|_2 \quad \forall\, x \in V.$$

LINK: TEIL 1 DER 12. VORLESUNG VOM 19.05.2022

In den Teilen (c) und (d) ist $|x_j| \in [0, \infty)$ für alle $j \in \mathbb{N}$, egal ob $x \in \mathbb{R}^n$ oder $x \in \mathbb{C}^n$ gilt.

(e) Definiere für $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ den *Folgenraum*

$$\mathbb{K}^{\mathbb{N}} := \Big\{ a = (a_i)_{i \in \mathbb{N}} = (a_1, a_2, \ldots) \;:\; a_i \in \mathbb{K} \Big\}.$$

Für zwei Mengen A und B definiert man ganz allgemein

$$A^B := \Big\{ f \colon B \to A \Big\}.$$

Setze nun für $p \in [1, \infty)$ zusätzlich

$$\| \cdot \|_p \colon \mathbb{K}^{\mathbb{N}} \to [0, \infty], \quad a \mapsto \left(\sum_{i \in \mathbb{N}} |a_i|^p \right)^{\frac{1}{p}}$$

und

$$\| \cdot \|_\infty \colon \mathbb{K}^{\mathbb{N}} \to [0, \infty], \quad a \mapsto \sup_{i \in \mathbb{N}} |a_i|.$$

Dann gilt

$$\forall\, p \in [1, \infty], \forall\, a \in \mathbb{K}^{\mathbb{N}} : \quad \|a\|_p = \lim_{n \to \infty} \overbrace{\underbrace{\|(a_1, \ldots, a_n)\|_p}_{\text{Norm auf } \mathbb{K}^n}}^{\text{isoton in } n \in \mathbb{N}}.$$

Für $p \in [1, \infty]$ setzen wir nun

$$l^p(\mathbb{N}, \mathbb{K}) := \Big\{ a \in \mathbb{K}^{\mathbb{N}} \;:\; \|a\|_p < \infty \Big\}.$$

Satz 10.3. *Für alle $p \in [1, \infty]$ und $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ ist*

$$\Big(l^p(\mathbb{N}, \mathbb{K}), \| \cdot \|_p \Big)$$

ein normierter Raum.

Beweis. Übung!

Beispiel 10.4.

(a) Es sei $a_k := (-1)^k \cdot \frac{1}{k}$ für alle $k \in \mathbb{N}$. Bekannterweise ist

$$\sum_{k \in \mathbb{N}} |a_k| = \sum_{k \in \mathbb{N}} \frac{1}{k}) = \infty$$

divergent. Damit ist $a \notin l^1$. Dagegen gilt

$$\sum_{k \in \mathbb{N}} |a_k|^2 = \sum_{k \in \mathbb{N}} \frac{1}{k^2} < \infty$$

und damit $a \in l^2$. Des Weiteren ist $\sup_{k \in \mathbb{N}} |a_k| \leq 1$, also $a_k \in l^\infty$.

(b) Es sei $b_k = 1$ für alle $k \in \mathbb{N}$. Dann ist $\sum b_k = \infty = \sum b_k^2$, aber $\sup |b_k| = 1$. Also ist $b \notin l^1$, $b \notin l^2$ und $b \in l^\infty$.

(c) Es sei $c_k = \frac{1}{k^2}$ für alle $k \in \mathbb{N}$. Dann ist

$$\sum_{k\in\mathbb{N}} |c_k| = \sum_{k\in\mathbb{N}} \frac{1}{k^2} < \infty \quad \Rightarrow \quad c \in l^1$$

und

$$\sum_{k\in\mathbb{N}} |c_k|^2 = \sum_{k\in\mathbb{N}} \frac{1}{k^4} < \infty \quad \Rightarrow \quad c \in l^2.$$

Außerdem ist c_k eine Nullfolge und damit beschränkt, insbesondere gilt $\sup_{k\in\mathbb{N}} |c_k| < \infty$, also $c \in l^\infty$.

Analog zum Bild in Beispiel 10.2 (d) gilt

$$\forall\, a \in \mathbb{K}^{\mathbb{N}} : \quad \|a\|_1 \geq \|a\|_2 \geq \|a\|_\infty \quad \Rightarrow \quad l^1 \subset l^2 \subset l^\infty.$$

„Abfall schneller als $\frac{1}{k}$ $\Rightarrow$ Abfall schneller als $\frac{1}{\sqrt{k}}$ $\Rightarrow$ beschränkte Folge.“

Beispiel 10.5. Es sei $I = [a,b] \subset \mathbb{R}$. Dann sind $V_1 = C\big([a,b],\mathbb{R}\big)$ und $V_2 = C\big([a,b],\mathbb{C}\big)$ Vektorräume von Funktionen. Wir definieren dann drei Normen auf $V \in \{V_1, V_2\}$:

- Die *Supremumsnorm*

$$\|f\|_\infty := \sup_{x\in I} |f(x)| = \max_{x\in I} |f(x)|;$$

- Die L^1-*Norm*

$$\|f\|_1 := \int_a^b |f(x)| \,\mathrm{d}x;$$

- Die L^2-*Norm*

$$\|f\|_2 := \left(\int_a^b |f(x)|^2 \,\mathrm{d}x\right)^{\frac{1}{2}} =: (\langle f, f\rangle)^{\frac{1}{2}}$$

  für das Skalarprodukt

$$\langle f, g\rangle := \int_a^b f(x)\overline{g(x)} \,\mathrm{d}x,$$

  wobei $\overline{g(x)} = g(x)$ für alle $x \in [a,b]$, falls $g \in C([a,b],\mathbb{R})$. Die Cauchy-Schwarz-Ungleichung lautet hier

$$\left|\int_a^b f(x)\overline{g(x)} \,\mathrm{d}x\right| \leq \int_a^b |f(x)||g(x)| \,\mathrm{d}x \leq \|f\|_2\|g\|_2 = \left(\int_a^b |f|^2 \,\mathrm{d}x \cdot \int_a^b |g|^2 \,\mathrm{d}x\right)^{\frac{1}{2}}.$$

Definition 10.6 (Metrik)**.** Es sei $X \neq \varnothing$ eine Menge. Eine Abbildung

$$d\colon X \times X \to \mathbb{R}, \quad (x_1, x_2) \mapsto d(x_1, x_2)$$

heißt dann *Metrik* oder *Abstand auf* X, falls für alle $x, y, z \in X$ die folgenden Eigenschaften erfüllt sind:

(M1) *(Definitheit)*:
Es gilt

$$d(x, y) \geq 0 \quad \text{und} \quad d(x, y) = 0 \quad \Leftrightarrow \quad x = y.$$

(M2) *(Symmetrie)*:
Es gilt

$$d(x, y) = d(y, x).$$

(M3) *(Dreiecksungleichung)*:
Es gilt

$$d(x, y) \leq d(x, z) + d(z, y).$$

Das Tupel (X, d) nennen wir dann *metrischen Raum*.

Beispiel 10.7.

(a) Es sei $(V, \|\cdot\|)$ ein normierter Raum und

$$d(x, y) := \|x - y\| = \| - (y - x)\| = | - 1|\|y - x\|.$$

Dann ist (V, d) ein metrischer Raum.

(b) Ist $Y \subset X$ und (X, d) ein metrischer Raum, dann ist (Y, d_Y) ebenfalls ein metrischer Raum, wobei

$$d_Y(x, y) = d(x, y) \quad \text{für } x, y \in Y$$

die *Einschränkung von d auf Y* bezeichnet. Als Beispiel betrachte man $X = \mathbb{R}^2$, $d(x, y) = \|x - y\|_2$ und $Y = \mathbb{R} \times [0, \infty)$.

Allgemein: Ist $(V, \|\cdot\|)$ ein normierter Raum und $X \subset V$ beliebig, so ist $d(x, y) = \|x - y\|$ für $x, y \in X$ eine Metrik auf X; man nennt sie auch *von der Norm $\|\cdot\|$ induzierte Metrik*. So ist beispielsweise $\mathbb{Q}^n$, $n \in \mathbb{N}$, ein metrischer Raum mit $d(x, y) = \|x - y\|_p$.

LINK: TEIL 2 DER 12. VORLESUNG VOM 19.05.2022

(c) Für eine Menge $X \neq \varnothing$ bezeichnet man die Abbildung

$$d\colon X \times X \to \{0, 1\}, \quad (x, y) \mapsto d(x, y) := \begin{cases} 1, & \text{falls } x \neq y, \\ 0, & \text{falls } x = y, \end{cases}$$

als *diskrete Metrik*. Man bemerke, dass alle Punkte $x, y \in X$ mit $x \neq y$ denselben Abstand voneinander besitzen.

Übung: Zeigen Sie:

- d ist eine Metrik;
- d ist keine von einer Norm $\|\cdot\|$ induzierte Metrik.

(d) Es sei $(V, \|\cdot\|)$ ein normierter Raum und $x_0 \in V$. Dann ist die Abbildung

$$d\colon V \times V \to \mathbb{R}, \quad (x,y) \mapsto \begin{cases} 0, & \text{falls } x = y, \\ \|x - x_0\| + \|x_0 - y\|, & \text{falls } x \neq y, \end{cases}$$

eine von der Norm $\|\cdot\|$ induzierte Metrik – man bezeichnet sie auch als *französische Eisenbahnmetrik*. Damit ist das Tupel (V, d) ein metrischer Raum.

„Alle Wege führen über Paris."

Definition 10.8. Es sei (X, d) ein metrischer Raum, $x_0 \in X$ und $A \subset X$. Setze außerdem $A^c := X \setminus A$.

(a) Für $\varepsilon > 0$ bezeichnet man die Menge

$$B_\varepsilon(x_0) := \left\{ x \in X \ : \ d(x, x_0) < \varepsilon \right\}$$

als *(offenen) ε-Ball um x_0* und

$$\overline{B_\varepsilon(x_0)} = \left\{ x \in X \ : \ d(x, x_0) \leq \varepsilon \right\}$$

als *abgeschlossenen ε-Ball um x_0*.

(b) x_0 heißt *Randpunkt* von A, falls für alle $\varepsilon > 0$ gilt:

$$B_\varepsilon(x_0) \cap A \neq \varnothing \quad \text{und} \quad B_\varepsilon(x_0) \cap A^c \neq \varnothing.$$

Die Menge aller Randpunkte von A kennzeichnet man durch

$$\partial A := \left\{ x \in X \ : \ x \text{ ist ein Randpunkt von } A \right\}.$$

(c) Die Menge $\overline{A} := A \cup \partial A$ heißt *Abschluss von A*.

(d) Die Menge $A^\circ := A \setminus \partial A$ heißt *offenes Inneres von A*. Ein Punkt $x_0 \in A^\circ$ heißt *innerer Punkt von A*.

(e) Die Menge A heißt *offen*, falls $A^\circ = A$, und *abgeschlossen*, falls $\overline{A} = A$.

(f) Ein Punkt x_0 heißt *Häufungspunkt von A*, falls

$$B_\varepsilon(x_0) \setminus \{x_0\} \cap A \neq \varnothing \quad \forall \varepsilon > 0.$$

x_0 heißt *isolierter Punkt von A*, falls ein $\varepsilon > 0$ existiert mit

$$B_\varepsilon(x_0) \cap A = \{x_0\}.$$

(g) Die Menge A heißt *dicht in X*, falls $\overline{A} = X$.

Bemerkung 10.9.

- Es sei X eine Menge mit $\partial X = \emptyset$. Dann ist X offen und abgeschlossen, es gilt also $X^\circ = X = \overline{X}$. Insbesondere ist $\varnothing$ offen und abgeschlossen, es gilt also $\partial\varnothing = \varnothing = \overline{\varnothing}$.
- Es sei x_0 ein innerer Punkt von A. Dann ist x_0 auch ein Häufungspunkt von A.
- Es sei d die diskrete Metrik auf der Menge X und $x_0 \in X$. Dann ist

$$\overline{B_1(x_0)} = \Big\{x \in X \ : \ d(x, x_0) \leq 1\Big\} = X = B_R(x_0) \quad \text{für } R > 1$$

und

$$B_{\frac{1}{2}}(x_0) = \left\{x \in X \ : \ d(d, x_0) < \frac{1}{2}\right\} = \{x_0\} = B_\varepsilon(x_0) \quad \text{für } \varepsilon \in (0,1).$$

- Es sei $A = \left(\frac{1}{n}\right)_{n\in\mathbb{N}} \subset \mathbb{R}$. Dann ist $0 \notin A$ ein Häufungspunkt von A. Außerdem ist jeder Punkt $x_0 \in A$ ein isolierter Punkt von A und

$$A^\circ = \varnothing \quad \text{sowie} \quad \overline{A} = A \cup \{0\} = \partial A.$$

Diese Aussagen gelten für den euklidischen Abstand, aber auch für jede Metrik $d(x,y) = \|x - y\|$ bezüglich einer Norm auf $\mathbb{R}$.

Insbesondere liegt $\mathbb{Q}$ dicht in $\mathbb{R}$.

Satz 10.10. *Es sei (X,d) ein metrischer Raum und $A \subset X$. Dann gelten die folgenden Aussagen:*

(a) *Es gilt*

$$x_0 \in A^\circ \quad \Leftrightarrow \quad \exists\, \varepsilon > 0 : \quad B_\varepsilon(x_0) \subset A;$$

(b) *A° ist offen, es gilt also $(A^\circ)^\circ = A^\circ$;*

(c) *$\overline{A}$ ist abgeschlossen, es gilt also $\overline{(\overline{A})} = \overline{A}$;*

(d) *A ist genau dann offen, wenn A^c abgeschlossen ist.*

Beweis.

(a) Es sei zunächst $x \in A^\circ = A \setminus \partial A$. Dann existiert ein $\varepsilon > 0$, so dass

$$B_\varepsilon(x_0) \cap A^c = \varnothing,$$

also $B_\varepsilon(x_0) \subset A$. Es gelte nun $B_\varepsilon(x) \subset A$. Dann ist direkt $B_\varepsilon(x_0) \cap A^c = \varnothing$.

(b) Es gilt

$$(A^\circ)^\circ = A^\circ \setminus \partial(A^\circ) \subset A^\circ,$$

also $(A^\circ)^\circ \subset A^\circ$. Es ist also nur noch $(A^\circ)^\circ \supset A^\circ$ zu zeigen: Dazu sei $x_0 \in A^\circ$. Nach Teil (a) existiert dann ein $\varepsilon > 0$, so dass

$$B_\varepsilon(x_0) \subset A.$$

Behauptung. Es gilt

$$B_{\frac{\varepsilon}{2}}(x_0) \subset A^\circ.$$

Beweis der Behauptung. Wir nehmen an, dass die Aussage nicht gilt. In diesem Fall existiert dann ein $y \in B_{\frac{\varepsilon}{2}}(x_0) \cap \partial A$. Dann ist

$$B_{\frac{\varepsilon}{2}}(y) \cap A^c \neq \varnothing,$$

es existiert also ein Punkt $z \in B_{\frac{\varepsilon}{2}}(y) \cap A^c$. Nach (M3) gilt jedoch

$$d(x_0, z) \leq d(x_0, y) + d(y, z) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon,$$

also $z \in B_\varepsilon(x_0) \subset A$. Dies ist aber ein Widerspruch wegen $z \in B_{\frac{\varepsilon}{2}}(y) \cap A^c$ bzw. $z \in A^c$. Also muss

$$B_{\frac{\varepsilon}{2}}(x_0) \subset A^\circ$$

gelten.

Die Behauptung zusammen mit Teil (a) liefert dann $x_0 \in (A^\circ)^\circ$. Insgesamt gilt also $(A^\circ)^\circ = A^\circ$.

Teile (c) und (d): Übung! □

Beispiel 10.11.

(a) Es sei $X = \mathbb{R}$ mit einer beliebigen Norm $\|\cdot\|$ und $d(x, y) = \|x - y\|$. Für $a, b \in \mathbb{R}$ mit $a < b$ gilt dann

$$\partial[a, b) = \partial(a, b] = \partial(a, b) = \partial[a, b] = \{a, b\}.$$

Außerdem ist $\overline{(a, b)} = [a, b]$ abgeschlossen und $[a, b]^\circ = (a, b) = [a, b)^\circ$ offen. Hingegen ist $[a, b)$ bzw. $(a, b]$ weder abgeschlossen noch offen.

Außerdem gilt

$$\mathbb{Q}^\circ = \varnothing \quad \text{und} \quad \overline{\mathbb{Q}} = \partial\mathbb{Q} = \mathbb{R}.$$

Analog ist $\overline{\mathbb{Q}^n} = \overline{\mathbb{R}^n}$ usw.

(b) Es sei (X, d) ein metrischer Raum, $x_0 \in X$ und $\delta > 0$. Dann ist die Menge $A = B_\delta(x_0)$ offen, denn zu $x \in B_\delta(x_0)$ kann man $\varepsilon = \delta - d(x, x_0) > 0$ wählen, so dass

$$\forall\, y \in B_\varepsilon(x) d(y, y_0) \leq d(y, x) + d(x, x_0) < \varepsilon + d(x, x_0) = \delta,$$

also $y \in A$ und damit $B_\varepsilon(x) \subset A$. Mit Satz 10.10 (a) folgt dann $x \in A^\circ$, also $A = A^\circ$. Analog ist $\overline{B_\delta(x_0)}$ eine abgeschlossene Menge.

Ist $(X, \|\cdot\|)$ ein normierter Raum und $A = \Big\{x \in X \ : \ d(x, x_0) = \delta\Big\} \subset X$, so gilt

$$\partial B_\delta(x) = A.$$

Dies ist in allgemeinen metrischen Räumen jedoch nicht richtig. So gilt beispielsweise für X mit der diskreten Metrik d und dem Punkt $x_0 \in X$, dass

$$A = \Big\{x \in X \ : \ d(x, x_0) = 1\Big\} = A \setminus \{x_0\}.$$

Für alle $D \subset X$ und $x \in X$ ist aber

$$B_{\frac{1}{2}}(x) \cap D^c = \{x\} \cap D^c = \begin{cases} \{x\}, & \text{falls } x \in D^c, \\ \varnothing, & \text{falls } x \notin D^c. \end{cases}$$

Es folgt $\partial D = \varnothing$ für alle $D \subset X$, also $\varnothing \neq X \setminus \{x_0\}$.

LINK: TEIL 1 DER 13. VORLESUNG VOM 25.05.2022

(c) Es sei $X = \mathbb{R}^2$, $d(x, y) := \|x - y\|$ eine durch eine beliebige Norm induzierte Metrik auf X und

$$A = \Big\{(x, y) \in \mathbb{R}^2 \ : \ 1 \leq x^2 + y^2 < 4\Big\} \cup \Big\{(0, 0)\Big\}.$$

Dann gilt

$$A^\circ = \Big\{(x, y) \in \mathbb{R}^2 \ : \ 1 < x^2 + y^2 < 4\Big\},$$

$$\overline{A} = \Big\{(x, y) \in \mathbb{R}^2 \ : \ 1 \leq x^2 + y^2 \leq 4\Big\} \cup \Big\{(0, 0)\Big\},$$

$$\partial A = \Big\{(x, y) \in \mathbb{R}^2 \ : \ x^2 + y^2 \in \{1, 4\}\Big\} \cup \Big\{(0, 0)\Big\}.$$

Außerdem ist $(0, 0)$ der einzige isolierte Punkt von A.

Satz 10.12. *In einem metrischen Raum (X, d) ist*

(a) *die Vereinigung*

$$\bigcup_{j \in J} U_j$$

von beliebig vielen offenen Mengen $U_j \subset X$, $j \in J$, und der Durchschnitt

$$\bigcap_{k=1}^{n} U_k, \quad n \in \mathbb{N},$$

endlich vieler offener Mengen ebenfalls offen,

(b) *der Durchschnitt*

$$\bigcap_{j \in J} V_j$$

von beliebig vielen abgeschlossenen Mengen $V_j \subset X$, $j \in J$, *und die Vereinigung*

$$\bigcap_{k=1}^{n} V_k, \quad n \in \mathbb{N},$$

endlich vieler abgeschlossener Mengen ebenfalls abgeschlossen.

Beweis.

(a) Es sei J eine beliebige Index menge und für alle $j \in J$ sei $U_j \subset X$ offen. Setze außerdem $A := \bigcup_{j \in J} U_j$. Für ein $x \in A$ gilt, dass ein $j \in J$ existiert mit $x \in U_j$. Da U_j offen ist, existiert nach Satz 10.10 (a) ein $\varepsilon > 0$, so dass

$$B_\varepsilon(x) \subset U_j \subset A,$$

weshalb A ebenfalls offen sein muss.

Seien nun $U_1, \ldots, U_n \subset X$ offen $n \in \mathbb{N}$, und $B := \bigcap_{k=1}^{n} U_k$. Für ein $x \in B$ gilt dann, dass $x \in U_k$ für alle $k \in \{1, \ldots, n\}$. Da U_k offen ist, existiert ein $\varepsilon_k > 0$ mit

$$B_{\varepsilon_k}(x) \subset U_k.$$

Setze nun $\varepsilon := \min\{\varepsilon_1, \ldots, \varepsilon_n\} > 0$. Dann gilt

$$B_\varepsilon(x) \subset U_k \quad \forall\, k \in \{1, \ldots, n\} \quad \Rightarrow \quad B_\varepsilon(x) \subset \bigcap_{k=1}^{n} U_k = B.$$

Damit ist B offen.

(b) Für $j \in J$ seien $K_j \subset X$ jeweils abgeschlossen. Nach Satz 10.10 (d) ist dann $K_j^c = X \setminus K_j$ offen für alle $j \in J$. Nach Teil (a) und den de-morganschen Regeln ist dann

$$\bigcup_{j \in J} K_j^c = \left(\bigcap_{j \in J} K_j \right)^c \quad \text{offen} \quad \overset{10.10\ (d)}{\Rightarrow} \quad \bigcap_{j \in J} K_j \quad \text{abgeschlossen.}$$

Die zweite Aussage folgt analog mit den de-morganschen Regeln. □

11 Konvergenz und Stetigkeit in metrischen Räumen

Definition 11.1. Es sei (X, d) ein metrischer Raum, $(x_n)_{n \in \mathbb{N}} \subset X$ eine Folge und $a \in X$. Dann *konvergiert* $(x_n)_n$ *gegen* a *(bezüglich* d*)*, falls für alle $\varepsilon > 0$ ein $n_0 = n_0(\varepsilon) \in \mathbb{N}$ existiert, so dass

$$\forall\, n \geq n_0 : \quad d(x_n, a) < \varepsilon.$$

In diesem Fall schreiben wir

$$x_n \to a \quad \text{für } n \to \infty \quad \text{oder} \quad x_n \underset{n \to \infty}{\longrightarrow} a \quad \text{oder} \quad \lim_{n \to \infty} x_n = a.$$

Lemma 11.2. *Es sei (X,d) ein metrischer Raum, $(x_n)_{n\in\mathbb{N}} \subset X$ und $a \in X$. Dann sind die folgenden Aussagen äquivalent:*

(a) $a = \lim\limits_{n\to\infty} x_n$;

(b) $\forall\, \varepsilon > 0 \quad \exists\, n_0 = n_0(\varepsilon): \quad x_n \in B_\varepsilon(a) \quad \forall\, n \geq n_0$;

(c) $\forall\, \varepsilon > 0: \quad x_n \in B_\varepsilon(a)$ *für fast alle* $n \in \mathbb{N}$;

(d) $\big(d(a,x_n)\big)_{n\in\mathbb{N}}$ *ist eine Nullfolge in* $\mathbb{R}$.

Für den Fall $X = \mathbb{K}^d$, $\mathbb{K} \in \{\mathbb{R},\mathbb{C}\}$, und $d(x,y) = \|x-y\|$ für eine beliebige Norm $\|\cdot\|$ sind zusätzlich äquivalent:

(e) *Für jedes* $j \in \{1,\dots,d\}$ *gilt*

$$\lim_{n\to\infty} \underbrace{(\overbrace{x_n}^{\substack{n-\text{tes}\\ \text{Glied}}})_j}_{\substack{j-\text{te}\\ \text{Komponente}}} = a_j;$$

(f) $\lim\limits_{n\to\infty} \|x_n - a\|_\infty = 0$.

Beweis. Die Äquivalenz der Aussagen (a) bis (d) erhält man direkt durch Umformung. Wir zeigen nun die weitere Äquivalenz für (e) und (f). Wir erhalten

$$\begin{aligned} & \lim_{n\to\infty} (x_n)_j = a_j \quad \forall\, j \in \{1,\dots,d\} \\ \Leftrightarrow \quad & \max_{j=1}^{n} \left|(x_n)_j - a_j\right| \xrightarrow[n\to\infty]{} 0 \\ \Leftrightarrow \quad & \|x_n - a\|_\infty \xrightarrow[n\to\infty]{} 0. \end{aligned} \tag{$\times\times$}$$

Es sei nun $\|\cdot\|$ eine beliebige Norm auf $\mathbb{K}^n$ und $e_1,\dots,e_n$ die kanonische Basis von $\mathbb{K}^n$. Setze nun

$$C := \sum_{j=1}^{n} \|e_j\|.$$

Für $x = \sum_{j=1}^n \alpha_j e_j \in \mathbb{K}^n$ gilt dann

$$\|x\| = \left\|\sum_{j=1}^{n} \alpha_j e_j\right\| \leq \sum_{j=1}^{n} \|\alpha_j e_j\| = \sum_{j=1}^{n} |\alpha_j|\|e_j\| \leq \max_{j=1}^{n} |a_j| \sum_{j=1}^{n} \|e_j\| = \|x\|_\infty C.$$

Damit impliziert $(\times\times)$ auch $\|x - a\| \xrightarrow[n\to\infty]{} 0$. Die Umkehrrichtung gilt auch, aber mit anderen Methoden. □

LINK: TEIL 2 DER 13. VORLESUNG VOM 25.05.2022

Analog wie im eindimensionalen Fall gelten:

Lemma 11.3. *Es sei (X, d) ein metrischer Raum, $(x_n)_{n\in\mathbb{N}} \subset X$ eine Folge und $a, b \in X$. Dann gilt*

$$x_n \underset{n\to\infty}{\longrightarrow} a \quad \textit{und} \quad x_n \underset{n\to\infty}{\longrightarrow} b \quad \Rightarrow \quad a = b.$$

Dies rechtfertigt erst die Schreibweise $a = \lim_{n\to\infty} x_n$. In allgemeineren Situationen – in topologischen statt metrischen Räumen – muss diese Aussage hingegen nicht gelten.

Beweis von Lemma 11.3. Widerspruchsannahme: Wir nehmen an, dass $a \neq b$ gilt, insbesondere ist dann $d(a, b) > 0$. Setze nun $\varepsilon := \frac{1}{2}d(a, b)$. Dann existiert nach Voraussetzung ein $n_0(\varepsilon) \in \mathbb{N}$, so dass

$$d(x_n, a) < \varepsilon \quad \text{und} \quad d(x_n, b) < \varepsilon \quad \forall\, n \geq n_0(\varepsilon).$$

Daraus folgt

$$d(a, b) \leq d(a, x_n) + d(x_n, b) < \varepsilon + \varepsilon = 2\varepsilon = d(a, b),$$

was ein Widerspruch ist. Also muss $a = b$ gelten. □

Definition 11.4. Es sei V ein Vektorraum über $\mathbb{K}$, $\|\cdot\|$ eine Norm auf V und $M \subset V$. M heißt dann *beschränkt*, falls ein $K \in \mathbb{R}$ existiert mit

$$\|x\| \leq K \quad \forall\, x \in M.$$

Lemma 11.5. *Es sei V ein Vektorraum über $\mathbb{K}$ mit einer Norm $\|\cdot\|$. Dann gelten:*

(i) *Jede konvergente Folge in $(V, \|\cdot\|)$ ist beschränkt;*

(ii) *Enthält V mehr als ein Element, so ist V auch unbeschränkt.*

Beweis.

(i) Der Beweis erfolgt analog zum Beweis für $V = \mathbb{R}$.

(ii) Es sei V ein Vektorraum mit mehr als einem Element – es existiere also ein Punkt $x \in V$ mit $x \neq 0_V$. Sei außerdem

$$M = \operatorname{span}\{x\} = \Big\{\alpha \cdot x \ : \ \alpha \in \mathbb{R}\Big\}.$$

Dann ist

$$\mathbb{K} \ni \alpha \mapsto \|\alpha \cdot x\| = |\alpha|\|x\|$$

eine unbeschränkte Funktion. □

In allgemeinen metrischen Räumen gilt Lemma 11.5 (ii) im Allgemeinen nicht: So kann man beispielsweise auf $\mathbb{K}^n$ eine Metrik definieren durch

$$d(x, y) := \frac{\|x - y\|_2}{1 + \|x - y\|_2} \in [0, 1).$$

Bezüglich dieser Metrik gilt dann

$$B_1(x_0) = \mathbb{K}^n \quad \forall\, x_0 \in \mathbb{K}^n,$$

aber auch

$$d(x_n, x_0) \underset{n\to\infty}{\longrightarrow} 0 \quad \Leftrightarrow \quad \|x_n - x_0\|_2 \underset{n\to\infty}{\longrightarrow} 0.$$

Also ist d in manchen Aspekten unterschiedlich zu $\|\cdot\|$, in anderen Aspekten aber nicht.

Lemma 11.6. *Es sei V ein normierter Vektorraum über $\mathbb{K}$, $(x_n)_{n\in\mathbb{N}}, (y_n)_{n\in\mathbb{N}}$ Folgen in V mit*

$$\lim_{n\to\infty} x_n = a \quad \textit{und} \quad \lim_{n\to\infty} y_n = b$$

und $(\lambda_n)_n$ eine Folge in $\mathbb{K}$ mit

$$\lim_{n\to\infty} \lambda_n = \lambda.$$

Dann gelten

(i) $\lim\limits_{n\to\infty} (x_n + y_n) = a + b$,

(ii) $\lim\limits_{n\to\infty} (\lambda_n \cdot x_n) = \lambda \cdot a$ *und insbesondere*

(iii) $\lim\limits_{n\to\infty} \|x_n\| = \|a\|$.

Beweis. Beweis von (i) und (ii): Übung!

(iii) Aus Teil (ii) folgt direkt

$$\lim_{n\to\infty} \|\lambda_n \cdot x_n - \lambda \cdot a\| = 0.$$

Für $\lambda_n = 1 = \lambda$, $n \in \mathbb{N}$ gilt also

$$\lim_{n\to\infty} \|x_n - a\| = 0.$$

Daraus folgt

$$\Big|\|x_n\| - \|a\|\Big| < \|x_n - a\| \underset{n\to\infty}{\longrightarrow} 0 \quad \Rightarrow \quad \lim_{n\to\infty} \|x_n\| = \|a\|.$$

□

Definition 11.7. Es sei (X, d) ein metrischer Raum und $(x_n)_{n\in\mathbb{N}} \subset X$ eine Folge.

(i) $(x_n)_n$ heißt *Cauchy-Folge (bezüglich d)*, falls

$$\forall\, \varepsilon > 0 \quad \exists\, n_0 = n_0(\varepsilon) \in \mathbb{N} : \quad d(x_m, x_n) < \varepsilon \quad \forall\, n, m \geq n_0.$$

(ii) Existiert zu jeder Cauchy-Folge $(x_n)_n$ in (X, d) ein $a \in X$ mit

$$d(x_n, a) \underset{n\to\infty}{\longrightarrow} 0,$$

so heißt (X, d) *vollständig*.

(iii) Ein vollständiger normierter Raum heißt *Banachraum*.

(iv) Ein vollständiger euklidischer oder unitärer Raum heißt *Hilbertraum*.

In der Vorlesung zu Analysis I haben wir gezeigt, dass $\mathbb{R}$ und $\mathbb{C}$ jeweils mit $d(x,y) = |x-y|$ ein vollständige (metrische) Räume sind.

Beispiel 11.8 (für nicht vollständige Räume)**.**

Im Video: **Beispiel 11.5**.

(a) $\mathbb{Q}$ ist mit $d(x,y) = |x-y|$ ein metrischer Raum, aber nicht vollständig. Wir haben bereits in der Vorlesung zu Analysis I mit dem Heron-Verfahren gezeigt, dass eine Folge $(x_n) \subset \mathbb{Q}$ mit $|x_n - \sqrt{2}| \xrightarrow[n\to\infty]{} 0$ existiert. Dies ist insbesondere eine Cauchy-Folge, aber wegen $\sqrt{2} \in \mathbb{I} = \mathbb{R} \setminus \mathbb{Q}$ konvergiert diese nicht in $\mathbb{Q}$.

(b) Es sei $X = (0,1) \subset \mathbb{R}$ mit $d(x,y) = |x-y|$. Dann ist (X,d) ein metrischer Raum, der aber nicht vollständig ist: So konvergiert $(x_n)_{n\in\mathbb{N}}$ definiert durch $x_n = \frac{1}{n} \in X$ gegen $0 \in \mathbb{R} \setminus X$. Sie ist also eine Cauchy-Folge, besitzt aber keinen Grenzwert in X. Dagegen ist $X = [0,1] \subset \mathbb{R}$ ein vollständiger metrischer Raum.

(c) Auf $V = C([-1,1])$ ist durch

$$\|f\|_1 = \int_{-1}^{1} |f| \,\mathrm{d}x$$

eine Norm gegeben und durch

$$d(f,g) = \|f-g\|_1$$

ein Abstand. $(V, \|\cdot\|_1)$ ist aber nicht vollständig, denn für

$$f_n\colon [-1,1] \to \mathbb{R}, \quad x \mapsto \begin{cases} -1 & \text{für } -1 \le x < -\frac{1}{n}, \\ n \cdot x & \text{für } -\frac{1}{n} \le x < \frac{1}{n}, \\ 1 & \text{für } \frac{1}{n} \le x \le 1, \end{cases}$$

ist $f_n \in C([-1,1])$ für alle $n \in \mathbb{N}$ und

$$\|f_m - f_n\|_1 = \int_{-1}^{1} |f_m(x) - f_n(x)| \,\mathrm{d}x = \left|\frac{1}{m} - \frac{1}{n}\right| \le \frac{1}{m} + \frac{1}{n} \le \frac{2}{n_0} \quad \forall\, n,m \ge n_0.$$

Damit ist $(f_n)_n$ eine Cauchy-Folge, aber es gilt

$$f_n \xrightarrow[n\to\infty]{\text{pktw}} f,$$

wobei

$$f\colon [-1,1] \to \mathbb{R}, \quad x \mapsto f(x) = \operatorname{sgn}(x) = \begin{cases} 1 & \text{für } x \in (0,1], \\ 0 & \text{für } x = 0, \\ -1 & \text{für } x \in [-1,0), \end{cases}$$

also $f \notin C([-1,1])$. Kann es eine andere Funktion $g \in C([-1,1])$ geben, so dass

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

gilt?

LINK: TEIL 1 DER 14. VORLESUNG VOM 01.06.2022

Angenommen es existiere ein $g \in C([-1,1])$ mit

$$\lim_{n\to\infty} \|f_n - g\|_1 = 0.$$

Für alle $\delta \in (0,1)$ gilt dann

$$\int_\delta^1 |1-g(x)|\,\mathrm{d}x \overset{n\geq\frac{1}{\delta}}{=} \int_\delta^1 |f_n(x)-g(x)| \leq \|f_n - g\|_1.$$

Da g stetig ist, ist insbesondere $|1-g(\cdot)|$ stetig, und zusammen mit dem Satz 14.16 (f) des Skripts aus der Vorlesung zur Analysis I (Lehramt) folgt direkt, dass

$$|1-g| \equiv 0 \quad \text{auf } (\delta,1], \text{ also } \quad g \equiv 1.$$

Da $\delta \in (0,1)$ beliebig ist, gilt also $g \equiv 1$ auf $(0,1]$. Analog zeigt man, dass

$$g \equiv -1 \quad \text{auf } [-1,0).$$

Damit ist aber g unstetig im Punkt $x_0 = 0 \in [-1,1]$, was im Widerspruch zu der Annahme $g \in C([-1,1])$ steht. Damit haben wir bestätigt, dass $C([-1,1])$ nicht vollständig bezüglich der Norm $\|\cdot\|_1$ ist.

Auch wenn $C([-1,1])$ nicht vollständig bezüglich $\|\cdot\|_1$ ist, muss dies nicht für alle Normen gelten.

Lemma 11.9. *Es seien $a,b \in \mathbb{R}$ mit $a<b$. Dann ist $C([a,b])$ vollständig bezüglich der Supremumsnorm $\|\cdot\|_\infty$.* Im Video: **Lemma 11.5$\frac{1}{2}$.**

Beweis. Es sei $(f_n)_{n\in\mathbb{N}}$ eine Folge in $C([a,b])$. Dann besagt das Cauchy-Kriterium aus Lemma 5.3 (iii) für $M=[a,b]$, dass

$$(f_n)_n \text{ ist eine Cauchy-Folge bezüglich } \|\cdot\|_\infty$$

$$\Leftrightarrow \quad \exists f\colon M\to\mathbb{C}: \quad \|f-f_n\|_\infty \xrightarrow[n\to\infty]{} 0$$

$$\overset{5.1}{\Leftrightarrow} \quad \exists f\colon M\to\mathbb{C}: \quad f_n \xrightarrow[n\to\infty]{\text{glm}} f.$$

Da f_n nach Voraussetzung stetig in $[a,b]$ für alle $n\in\mathbb{N}$ ist, ist f nach Satz 5.5 ebenfalls stetig in $[a,b]$ – es gilt also $f \in C([a,b])$. Damit ist $C([a,b])$ vollständig bezüglich $\|\cdot\|_\infty$. □

Folgendes ist ebenfalls eine Verallgemeinerung von Sachverhalten, die wir bereits in $\mathbb{R}$ kennengelernt haben:

Definition 11.10. Es sei (X,d) ein metrischer Raum und $(x_n)_{n\in\mathbb{N}} \subset X$ eine Folge. Einen Punkt $a \in X$ nennen wir dann *Häufungspunkt von* $(x_n)_n$, falls Im Video: **Def. 11.9.**

$$\forall \varepsilon > 0: \quad \left|\left\{n\in\mathbb{N} \;:\; x_n \in B_\varepsilon(a)\right\}\right| = \infty.$$

In anderen Worten: Für alle $\varepsilon>0$ existieren unendlich viele Indizes n, so dass $x_n \in B_\varepsilon(a)$. (vgl. Definition 7.1 des Skripts aus der Vorlesung zur Analysis I (Lehramt))

Dies stimmt auch mit der früheren Definition von Häufungspunkten überein: Setze dazu

$$A := \Big\{x \in X \ : \ \exists\, n \in \mathbb{N} : \quad x_n = x\Big\}.$$

Dann ist a genau dann ein Häufungspunkt von A, wenn a ein Häufungspunkt von $(x_n)_n$ ist.

Beweisskizze.

($\Rightarrow$) Es sei a ein Häufungspunkt von A. Für alle $n \in \mathbb{N}$ existiert dann ein $y_n \in B_{\frac{1}{n}}(a) \cap A \setminus \{a\}$. Zu einem beliebigen $\varepsilon > 0$ und $k_0 > \frac{1}{\varepsilon}$ gilt dann $(y_k)_{k \geq k_0} \subset B_\varepsilon(a)$. Damit existieren unendlich viele Indizes, so dass $x_n \in B_\varepsilon(a)$ – a ist also ein Häufungspunkt von (x_n).

($\Leftarrow$) Für alle $\varepsilon > 0$ gilt nach Voraussetzung

$$\Big|B_\varepsilon(a) \cap (x_n)_n\Big| = \Big|B_\varepsilon(a) \cap A\Big| = \infty.$$

Damit gilt insbesondere

$$\Big|(B_\varepsilon(a) \cap A) \setminus \{a\}\Big| = \infty,$$

es gilt also $B_\varepsilon(a) \setminus \{a\} \cap A \neq \varnothing$.

Satz 11.11. *Es sei (X, d) ein metrischer Raum, $(x_n)_{n\in\mathbb{N}} \subset X$ eine Folge und $a \in X$. Dann sind die folgenden Aussagen äquivalent:*

Im Video: **Satz 11.10**.

(i) *a ist ein Häufungspunkt von $(x_n)_n$;*

(ii) *Es existiert eine Teilfolge $(x_{n_k})_{k\in\mathbb{N}}$ von $(x_n)_n$ mit $\lim_{k\to\infty} x_{n_k} = a$.*

Beweis. Der Beweis erfolgt analog wie in der Menge $\mathbb{R}$, nur unter der Verwendung der Metrik d. (vgl. Satz 7.2 (b) des Skripts aus der Vorlesung zur Analysis I (Lehramt)) □

Satz 11.12 (von Bolzano-Weierstraß in $\mathbb{R}^n$). *Eine beschränkte Folge in $\mathbb{R}^n$ besitzt mindestens einen Häufungspunkt.*

Im Video: **Satz 11.11**.

Beweis. Wir bezeichnen die Folgenglieder mit

$$x^k = \left(x_1^k, \ldots, x_n^k\right) \in \mathbb{R}^n,$$

wobei $k \in \mathbb{N}$ den Folgenindex und $n \in \mathbb{N}$ den Koordinatenindex bezeichnet. Dann ist $\left(x_1^k\right)_{k\in\mathbb{N}}$ eine Folge in $\mathbb{R}$, die außerdem beschränkt ist: Es gilt nämlich

$$\sup_{k\in\mathbb{N}} |x_1^k| \leq \sup_{k\in\mathbb{N}} \|x^k\| < \infty,$$

da $(x^k)_k$ beschränkt ist. Der Satz von Bolzano-Weierstraß in $\mathbb{R}$ (vgl. Satz 7.5 aus dem Skript der Vorlesung zur Analysis I (Lehramt)) besagt dann, dass eine Teilfolge $(k_j)_{j\in\mathbb{N}}$ und ein Punkt $a_1 \in \mathbb{R}$ existieren mit

$$\lim_{j\to\infty} x_1^{k_j} = a_1.$$

Die Teilfolge $\left(x^{k_j}\right)_{j\in\mathbb{N}}$ ist in $\mathbb{R}^n$ weiterhiin beschränkt, und damit auch die Folge $\left(x_2^{k_j}\right)_j$ in $\mathbb{R}$. Wie oben besitzt diese eine Teilteilfolge $\left(x_2^{k_{j_m}}\right)_{m\in\mathbb{N}}$, die gegen einen Punkt $a_2 \in \mathbb{R}$ konvergiert. Durch diesen zweiten Übergang zur Teilfolge haben wir erreicht, dass sogar die ersten beiden Koordinaten der Teilfolge konvergieren. Indem wir nun $(n-2)$ weitere Male eine Teilfolge auswählen, finden wir dann einen Punkt $(a_1, a_2, \ldots, a_n) \in \mathbb{R}^n$ und eine Teilfolge $(y^d)_d \subset (x^k)_k$, so dass

$$\lim_{d\to\infty} y_j^d = a_j \quad \forall\, j \in \{1, \ldots, n\}$$

$$\overset{11.2}{\Leftrightarrow} \quad \lim_{d\to\infty} \|y^d - a\| = 0. \qquad \square$$

Wir haben nun den Fall gesehen, dass der Satz von Bolzano-Weierstraß in einem endlich-dimensionalen Vektorraum gilt. Dies ist für unendlichdimensionale Vektorräume jedoch nicht der Fall!

Beobachtung. Ist $M \neq \varnothing$ eine beliebige Menge und $f, g, f_n \colon M \to \mathbb{C}$, $n \in \mathbb{N}$, Funktionen mit

$$f_n \xrightarrow[n\to\infty]{\text{pktw}} f \quad \text{und} \quad f_n \xrightarrow[n\to\infty]{\text{glm}} g,$$

so folgt $f = g$ auf ganz M.

Beweis.

1. Gleichmäßige Konvergenz impliziert auch punktweise Konvergenz, es gilt also $f_n \xrightarrow[n\to\infty]{\text{pktw}} g$.

2. Die Limesfunktion ist eindeutig:

$$\forall\, x \in M : \quad f(x) \xleftarrow[n\to\infty]{} f_n(x) \xrightarrow[n\to\infty]{} g(x).$$

   Insbesondere ist der Limes eindeutig in $\mathbb{C}$, es gilt also $f(x) = g(x)$, also $f = g$ auf ganz M.

Beispiel 11.13. Wir betrachten in $C([0,1])$ die Folge $(f_n)_{n\in\mathbb{N}}$ gegeben durch (s. Beispiel 5.4 (b))

Im Video: **Beispiel 11.12.**

$$f_n(x) := \begin{cases} 2nx & \text{für } 0 \le x \le \frac{1}{2n}, \\ 2 - 2nx & \text{für } \frac{1}{2n} < x < \frac{1}{n}, \\ 0 & \text{für } \frac{1}{n} \le 1 \le 1. \end{cases}$$

f_n ist dann stetig und es gilt $\|f_n\|_\infty = 1$ für alle $n \in \mathbb{N}$. Also ist die folge $(f_n)_n$ beschränkt bezüglich $\|\cdot\|_\infty$. Bekannterweise ist

$$f_n \xrightarrow[n\to\infty]{\text{pktw}} f \equiv 0.$$

Nehmen wir an, dass ein $g \in C([0,1])$ mit

$$\|g - f_{n_k}\|_\infty \xrightarrow[k\to\infty]{} 0,$$

so sogar $g \equiv f \equiv 0$ gelten. Nach Lemma 11.6 folgt dann $\|g\|_\infty = \lim_{k\to\infty} \|f_{n_k}\|_\infty = 1$. Wegen $\|f\| = 0$ wäre dies aber ein Widerspruch.

Link: Teil 2 der 14. Vorlesung vom 01.06.2022

Man beachte den Unterschied:

- Eine Cauchy-Folge in $C([a,b])$ besitzt eine Grenzfunktion $f \in C([a,b])$;
- Eine beschränkte Folge in $C([a,b])$ besitzt im Allgemeinen keine Teilfolge mit einer Grenzfunktion in $C([a,b])$

Hier handelt es sich nicht um einen Widerspruch: Die Annahme, dass man eine Cauchy-Folge vorliegen hat, ist bloß stärker als die Annahme, dass man eine beschränkte Folge vorliegen hat.

Wir setzen nun bereits bekannte Bezüge zwischen Folgen und Geometrie in den Kontext von metrischen Räumen.

Lemma 11.14. *Es sei (X,d) ein metrischer Raum, $A \subset X$ und $x \in X$. Dann gelten:*

Im Video: **Lemma 11.13**.

(i) *Es gilt genau dann $x \in \partial A$, wenn Folgen $(x_n)_{n\in\mathbb{N}} \subset A$ und $(y_n)_n \subset A^c$ existieren mit*

$$\lim_{n\to\infty} x_n = x = \lim_{n\to\infty} y_n;$$

(ii) *Es gilt*

$$\begin{aligned} x \in \overline{A} \quad &\Leftrightarrow \exists\,(x_n)_{n\in\mathbb{N}} \subset A: \quad \lim_{n\to\infty} x_n = x \\ &\Leftrightarrow \forall\,\varepsilon > 0: \quad B_\varepsilon(x) \cap A \neq \varnothing. \end{aligned}$$

(iii) *A ist genau dann abgeschlossen, wenn gilt:*

$$(x_n)_{n\in\mathbb{N}} \subset A \text{ ist eine Folge mit } \lim_{n\to\infty} x_n = x \in X \quad \Rightarrow \quad x \in A;$$

(iv) *∂A ist abgeschlossen.*

Beweis.

(i) ($\Rightarrow$) Es sei $x \in \partial A$. Dann gilt

$$\forall \varepsilon > 0: \quad B_\varepsilon(x) \cap A \neq \varnothing \quad \text{und} \quad B_\varepsilon(x) \cap A^c \neq \varnothing. \tag{$*$}$$

Insbesondere existieren $x_n \in B_{\frac{1}{n}}(x) \cap A$ und $y_n \in B_{\frac{1}{n}}(x) \cap A^c$ für alle $n \in \mathbb{N}$. Also sind $(y_n) \subset A^c$ und $(x_n) \subset A$ Folgen mit

$$|x_n - x| < \frac{1}{n} \xrightarrow[n\to\infty]{} 0 \quad \text{und} \quad |y_n - x| < \frac{1}{n} \xrightarrow[n\to\infty]{} 0; \tag{$\triangle$}$$

insbesondere gilt $\lim_{n\to\infty} x_n = \lim_{n\to\infty} y_n = x$.

($\Leftarrow$) Angenommen es existieren Folgen, die ($\triangle$) (und damit die Annahme) erfüllen. Dann ist offensichtlich auch ($*$) erfüllt.

Teile (ii) bis (iv): Übung! □

Korollar 11.15. *Es sei (X, d) ein metrischer Raum und $A \subset B \subset X$. Dann gelten:*

Im Video: **Korollar 11.14.**

(i) *A ist offen $\Rightarrow$ $A \subset B^\circ$;*

(ii) *B ist abgeschlossen $\Rightarrow$ $\overline{A} \subset B$.*

„A° ist die größte offene Menge, die in A nethalten ist und $\overline{A}$ die kleinste abgeschlossene Menge, die A enthält.“

Beweis.

(i) Es sei A offen und $x \in A$. Nach Satz 10.10 (a) existiert dann ein $\varepsilon > 0$, so dass

$$B_\varepsilon(x) \subset A \overset{\text{Vor.}}{\subset} B.$$

Damit ist $B_\varepsilon(x) \cap B^c = \varnothing$; insbesondere ist x kein Randpunkt von B. Es gilt also $x \in B \notin \partial B = B^\circ$. Damit ist $A \subset B^\circ$.

(ii) Es sei $x \in \overline{A}$ und B abgeschlossen. Dann existiert nach Lemma 11.14 (ii) eine Folge $(x_n)_{n\in\mathbb{N}} \subset A \overset{\text{Vor.}}{\subset} B$ mit $\lim_{n\to\infty} x_n = x$. Da B abgeschlossen ist, gilt wiederum nach Lemma 11.14 (ii), dass $x \in \overline{B} = B$. Es gilt also $\overline{A} \subset B$. □

Beispiel 11.16.

Im Video: **Beispiel 11.15.**

(a) Der Durchschnitt von beliebig vielen offenen Mengen ist im Allgemeinen nicht offen: Dazu betrachten wir für $n \in \mathbb{N}$ die Menge

$$U_n := \left(-\frac{1}{n}, \frac{1}{n}\right) \subset \mathbb{R}.$$

Offensichtlich ist U_n offen für alle $n \in \mathbb{N}$ (vgl. Beispiel 10.11 (a)). Dagegen ist jedoch die Menge

$$\bigcap_{n\in\mathbb{N}} U_n = \{0\}$$

abgeschlossen und nicht offen.

(b) Die Vereinigung von beliebig vielen abgeschlossenen Mengen ist ebenfalls im Allgemeinen nicht abgeschlossen: Dazu betrachten wir für $n \in \mathbb{N}$ die Menge

$$A_n := \left[\frac{1}{n}, 1\right] \subset \mathbb{R}.$$

Offensichtlich ist A_n abgeschlossen. Die Vereinigung

$$\bigcup_{n\in\mathbb{N}} A_n = (0,1]$$

ist jedoch nicht abgeschlossen (und auch nicht offen).

Definition 11.17. Es sei (X,d) ein metrischer Raum und $A \subset X$. Die Menge A heißt *kompakt*, falls jede Folge $(x_n)_{n\in\mathbb{N}} \subset A$ eine Teilfolge besitzt, die gegen einen Punkt $x \in A$ konvergiert.

Im Video: **Def. 11.16**.

Dies motiviert die folgende Umformulierung des Satzes von Bolzano-Weierstraß 11.12:

Satz 11.18 (von Bolzano-Weierstraß)**.** *Jede beschränkte und abgeschlossene Menge $M \subset \mathbb{R}^n$ ist kompakt.*

Im Video: **Satz 11.17**.

Beweis. Es sei $M \subset \mathbb{R}^n$ beschränkt und abgeschlossen sowie $(x_n)_{n\in\mathbb{N}} \subset M$ eine Folge. Nach dem Satz von Bolzano-Weierstraß in $\mathbb{R}^n$ (s. Satz 11.12) existieren Teilfolgen $(y_k)_{k\in\mathbb{N}} \subset (x_n)$ mit $\lim_{k\to\infty} y_k = a \in \mathbb{R}^n$. Wegen $(x_n) \subset M$ gilt insbesondere $(y_k) \subset M$ und aus der Abgeschlossenheit von M folgt direkt $a \in M$. Damit ist M kompakt. □

Im Allgemeinen gilt auch die Umkehrung:

Satz 11.19. *Es sei $(V, \|\cdot\|)$ ein normierter Raum und $A \subset V$ eine kompakte Teilmenge. Dann ist A abgeschlossen und beschränkt.*

Im Video: **Satz 11.18**.

Beweis. Es sei $(x_n)_{n\in\mathbb{N}} \subset A$ eine Folge, die gegen einen Punkt $x \in V$ konvergiert. Da A kompakt ist, existiert eine Teilfolge $(y_k)_{k\in\mathbb{N}} \subset (x_n)$ mit $\lim_{k\to\infty} y_k = y \in A$. Aufgrund der Eindeutigkeit des Limes folgt dann $x = y \in A$, weshalb A nach Lemma 11.14 (iii) abgeschlossen ist.

Nehmen wir nun an, dass A unbeschränkt ist, so gäbe es für alle $n \in \mathbb{N}$ ein $x_n \in A$ mit $\|x_n\| > n$. Nach der Kompaktheitsvoraussetzung von A existiert jedoch eine Teilfolge $(y_k)_k \subset (x_n)$, die gegen $y \in A$ konvergiert. Also existiert ein $N \in \mathbb{N}$ mit

$$\underbrace{\|y_k - y\| < 1}_{\Rightarrow \|y_k\| < \|y\|+1} \quad \forall\, k \geq N.$$

Dies ist jedoch ein Widerspruch zu $\|y_k\| \xrightarrow[k\to\infty]{} \infty$. Also muss A beschränkt sein. □

Beispiel 11.20.

Im Video: **Beispiel 11.19**.

(i) Die folgenden Mengen sind kompakt:

- Abgeschlossene Intervalle $[a, b] \subset \mathbb{R}$;
- Abgeschlossene Quader $[a_1, b_1] \times \ldots \times [a_n, b_n] \subset \mathbb{R}^d$;
- Abgeschlossene Kugeln $\overline{B_R(x_0)} \subset \mathbb{K}^n$ für $R > 0$ und $x_0 \in \mathbb{K}^d$, $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$.

(ii) In beliebigen normierten Räumen ist die Aussage

$$\text{Abgeschlossenheit und Beschränktheit} \quad \Rightarrow \quad \text{Kompaktheit}$$

falsch. Das Beispiel 11.13 liefert eine Folge $(f_n)_{n\in\mathbb{N}} \subset \overline{B_1(0)} \subset C([0,1])$, die keine konvergente Teilfolge besitzt. Offensichtlich ist $\overline{B_1(0)}$ abgeschlossen und beschränkt, wir haben aber keine Kompaktheit vorliegen.

Allgemein gilt: Ist V ein Vektorraum mit $\overline{B_1(0)}$ kompakt, so ist $\dim(V) < \infty$.

LINK: TEIL 1 DER 15. VORLESUNG VOM 02.06.2022

Definition 11.21. Es seien (X, d) und (Y, ρ) metrische Räume, $f\colon X \to Y$ eine Abbildung und $x_0 \in X$. Dann heißt

Im Video: **Def. 11.20**.

- f *stetig in* x_0, falls für alle $\varepsilon > 0$ ein $\delta = \delta(\varepsilon, x_0) > 0$ existiert mit
$$\rho\Big(f(x), f(x_0)\Big) < \varepsilon \quad \forall\, x \in X \text{ mit } d(x, x_0) < \delta;$$
- f *stetig auf* X, falls f stetig in x_0 für alle $x_0 \in X$ ist;
- f *gleichmäßig / uniform stetig auf* X, falls für alle $\varepsilon > 0$ ein $\delta = \delta(\varepsilon) > 0$ existiert, so dass
$$\forall\, x_1, x_2 \in X: \quad d(x_1, x_2) < \delta \quad \Rightarrow \quad \rho\Big(f(x_1), f(x_2)\Big) < \varepsilon;$$
- f *Lipschitz-stetig auf* X, falls ein $L \in \mathbb{R}$ existiert, so dass
$$\rho\Big(f(x_1), f(x_2)\Big) \le L \cdot d(x_1, x_2) \quad \forall\, x_1, x_2 \in X.$$

Offensichtlich gilt

$$x_0 \text{ ist ein isolierter Punkt von } X$$

$$\Rightarrow \quad \text{Für jedes } (Y, \rho) \text{ und } f\colon X \to Y \text{ ist } f \text{ stetig in } x_0.$$

Betrachten wir nun äquivalente Umformulierungen der Stetigkeit (wie in $\mathbb{R}$):

Lemma 11.22. *Es seien (X, d) und (Y, ρ) metrische Räume, $f\colon X \to Y$ eine Abbildung und $x_0 \in X$. Dann sind die folgenden Aussagen äquivalent:*

Im Video: **Lemma 11.21**.

(i) *f ist stetig in x_0;*

(ii) *Für alle* $\varepsilon > 0$ *existiert ein* $\delta > 0$ *mit*

$$f\Big(B_\delta(x_0)\Big) \subset B_\varepsilon\Big(f(x_0)\Big);$$

(iii) *Für jede Folge* $(x_n)_{n\in\mathbb{N}} \subset X$ *mit* $x_0 = \lim_{n\to\infty} x_n$ *gilt*

$$\lim_{n\to\infty} f(x_n) = f(x_0).$$

(Ohne Beweis)

Ähnlich wie in $\mathbb{R}$ / $\mathbb{C}$ kann man aus stetigen Funktionen auch neue konstruieren.

Satz 11.23.

Im Video: **Satz 11.22.**

(a) *Es sei* (X, d) *ein metrischer Raum,* $\lambda \in \mathbb{C}$, $x_0 \in X$ *und* $f, g\colon X \to C$ *Abbildungen, die stetig in* x_0 *sind. Dann sind auch*

$$f \pm g, \quad \lambda \cdot f, \quad f \cdot g \quad \textit{und} \quad \frac{f}{g} \quad (\textit{für } g(x_0) \neq 0)$$

stetig in x_0.

(b) *Es seien* (X, d_X), (Y, d_Y) *und* (Z, d_Z) *metrische Räume,* $x_0 \in X$, $f\colon X \to Y$ *stetig in* x_0, $y_0 := f(x_0)$ *und* $g\colon Y \to Z$ *stetig in* y_0. *Dann ist auch* $(g \circ f)\colon X \to Z$ *stetig in* x_0.

(Ohne Beweis)

Satz 11.24. *Es sei* (X, d) *ein metrischer Raum,* $f\colon X \to \mathbb{K}^n$, $n \in \mathbb{N}$, *eine Abbildung mit Koordinatenfunktionen* $f_1, \ldots, f_n\colon X \to \mathbb{K}$ *und* $x_0 \in X$. *Dann gilt*

Im Video: **Satz 11.23.**

$$f \textit{ ist stetig in } x_0 \quad \Leftrightarrow \quad f_j \textit{ ist stetig in } x_0 \quad \forall\, j \in \{1, \ldots, n\}.$$

Beweis. Die Aussage folgt direkt aus Lemma 11.2 und Lemma (iii). □

Beispiel 11.25.

Im Video: **Beispiel 11.24.**

(a) Es sei $L\colon \mathbb{R}^n \to \mathbb{R}^m$ eine lineare Abbildung. Dann gibt es bezüglich der Standardbasen in $\mathbb{R}^n$ und $\mathbb{R}^m$ eine eindeutige Matrix $A = (a_{jk})_{j=1,\ldots,m;k=1,\ldots,n} \in \mathbb{R}^{m\times n}$ mit

$$L(x) = A \cdot x \quad \text{für } x \in \mathbb{R}^n.$$

Mit der euklidischen Norm erhält man dann

$$\|Ax\|_2^2 = \sum_{j=1}^m \Big|(Ax)_j\Big|^2$$

$$= \sum_{j=1}^m \Big|\sum_{k=1}^n a_{jk} x_k\Big|^2$$

$$\overset{\text{CS}}{\leq} \left(\sum_{j=1}^{m} \sum_{k=1}^{n} |a_{jk}|^2 \right) \cdot \underbrace{\left(\sum_{k=1}^{n} |x_k|^2 \right)}_{=\|x\|^2},$$

also

$$\|Ax\|_2 \leq \left(\sum_{j=1} m \sum_{k=1}^{n} |a_{jk}|^2 \right)^{\frac{1}{2}} \cdot \|x\|_2.$$

Aufgrund der Linearität von L gilt $A \cdot (x-y) = Ax - Ay$ und damit

$$\|Ax - Ay\|_2 = \|A(x-y)\|_2 \leq \left(\sum_{j=1} m \sum_{k=1}^{n} |a_{jk}|^2 \right)^{\frac{1}{2}} \cdot \|x-y\|_2.$$

Also ist $L\colon \mathbb{R}^n \to \mathbb{R}^m$ sogar Lipschitz-stetig, damit insbesondere gleichmäßig stetig und auch stetig auf $\mathbb{R}^n$.

Spezialfall: Für $j \in \{1, \ldots, n\}$ ist die Abbildung

$$\pi_j\colon \mathbb{R}^n \to \mathbb{R}, \quad \pi_j(x_1, \ldots, x_n) = x_j$$

linear und damit gleichmäßig stetig auf $\mathbb{R}^n$.

(b) Es sei

$$P\colon \mathbb{R}^n \to \mathbb{R}, \quad P(x_1, \ldots, x_n) = \underbrace{\sum_{k_1=1}^{m} \cdots \sum_{k_n=1}^{m}}_{n \text{ Summen}} a_{k_1 \cdots k_n} x_1^{k_1} \cdot \ldots \cdot x_n^{k_n}$$

ein Polynom in n Variablen. Diese Funktion ist stetig auf $\mathbb{R}^n$.

Wir betrachten dazu den Spezialfall $n = 2$ und

$$(x, y) \mapsto \sum_{j,k=1}^{m} a_{jk} x^k \cdot y^j.$$

Diese Funktion ist stetig, da sie aus elementar-stetigen Funktionen zusammengebaut werden kann: Für die (stetigen) Abbildungen

$$f\colon \mathbb{R} \to \mathbb{R}, \quad x \mapsto x^k$$

$$g\colon \mathbb{R} \to \mathbb{R}, \quad y \mapsto y^j$$

ist die Zusammensetzung

$$F\colon \mathbb{R}^2 \to \mathbb{R}^2, \quad \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} f(x) \\ g(y) \end{pmatrix}$$

auch stetig (vgl. Satz 11.24). Außerdem sind auch

$$M\colon \mathbb{R}^2 \to \mathbb{R}^2, \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} x \cdot y \\ 0 \end{pmatrix} \quad \text{bzw.} \quad \begin{pmatrix} f(x) \\ g(y) \end{pmatrix} \mapsto \begin{pmatrix} f(x) \cdot g(y) = x^k \cdot y^j \\ 0 \end{pmatrix}$$

und die Projektion

$$\pi_1\colon \mathbb{R}^2 \to \mathbb{R}, \quad \begin{pmatrix} u \\ v \end{pmatrix} \mapsto u$$

alle stetig. Damit ist P als Linearkombination dieser (stetigen) Funktionen ebenfalls stetig.

LINK: TEIL 2 DER 15. VORLESUNG VOM 02.06.2022

(c) Die Abbildung

$$f\colon \mathbb{R}^3 \to \mathbb{R}, \quad (x, y, z) \mapsto \sin\left(\mathrm{e}^{xy^2} + \sqrt{y^2 + z^4} - \arctan\left(\cos\left(\frac{xyz}{1 + z^2}\right)\right)\right)$$

ist stetig, da die folgenden Abbildungen stetig sind:

$$\varphi\colon \mathbb{R}^3 \to \mathbb{R}^3, \quad \begin{pmatrix} x \\ y \\ z \end{pmatrix} \mapsto \begin{pmatrix} \mathrm{e}^{xy^2} \\ \sqrt{y^2 + z^4} \\ \arctan\left(\cos\left(\frac{xyz}{1+z^2}\right)\right) \end{pmatrix},$$

$$\psi\colon \mathbb{R}^3 \to \mathbb{R}^3, \quad \begin{pmatrix} u \\ v \\ w \end{pmatrix} \mapsto \begin{pmatrix} u + v + w \\ 0 \\ 0 \end{pmatrix},$$

$$\pi_1\colon \mathbb{R}^3 \to \mathbb{R}, \quad \begin{pmatrix} u \\ v \\ w \end{pmatrix} \mapsto u,$$

$$\sin\colon \mathbb{R} \to \mathbb{R}.$$

Dabei gilt $f = \sin \circ \pi_1 \circ \psi \circ \rho$.

(d) Ähnlich kann man zeigen, dass die Abbildung

$$f\colon \mathbb{R}^2 \to \mathbb{R}, \quad (x, y) \mapsto \begin{cases} \frac{xy}{x^2+y^2} & \text{für } (x, y) \neq (0, 0), \\ 0 & \text{sonst}, \end{cases}$$

stetig ist auf $X = \mathbb{R}^2 \setminus \{(0, 0)\}$. Wie verhält sich aber f in (der Nähe von) $(0, 0)$? Betrachten wir dazu Nullfolgen $(x_n)_{n\in\mathbb{N}}, (y_n)_n \subset \mathbb{R} \setminus \{0\}$, so gilt

$$\lim_{n\to\infty} f(x_n, 0) = \lim_{n\to\infty} \frac{0}{x_n^2} = 0 = \lim_{n\to\infty} f(0, y_n) = f(0, 0).$$

Also sind die Einschränkungen von f auf die jeweiligen Koordinatenachsen stetig. Solche Funktionen nennt man *partiell stetig*. Allerdings gilt

$$f(x_n, x_n) = \frac{x_n^2}{x_n^2 + x_n^2} = \frac{1}{2} \neq 0,$$

weshalb f nicht stetig in $(0,0)$ ist. Übrigens gilt auch

$$f(x_n, -x_n) = \frac{-x_n^2}{x_n^2 + x_n^2} = -\frac{1}{2} \neq 0.$$

(e) Betrachten nun wir die Abwandlung

$$g\colon \mathbb{R}^2 \to \mathbb{R}, \quad (x,y) \mapsto g(x) := x \cdot f(x) = \begin{cases} \frac{x^2 y}{x^2+y^2} & \text{für } (x,y) \neq (0,0), \\ 0 & \text{sonst.} \end{cases}$$

Wie oben ist g stetig auf $X = \mathbb{R}^2 \setminus \{(0,0)\}$. Wie sieht es aber mit der Stetigkeit in $(0,0)$ aus? Für $(x,y) \neq (0,0)$ ist $|g(x,y)| \leq |y|$, denn

$$|x^2| \leq |x^2 + y^2| \quad \Rightarrow \quad |x^2 y| \leq |x^2 + y^2| \cdot |y|.$$

Für jede Nullfolge $\big((x_n, y_n)\big)_{n\in\mathbb{N}} \subset X$ gilt dann

$$0 \leq \lim_{n\to\infty} |g(x_n, y_n)| \leq \lim_{n\to\infty} |y_n| = 0.$$

Also ist g stetig in $(0,0)$.

(f) Es sei $(V, \|\cdot\|)$ ein normierter Raum und $f(x) := \|x\|$. Dann ist f Lipschitz-stetig, insbesondere gleichmäßig / uniform stetig, da

$$|f(x) - f(y)| = \Big|\|x\| - \|y\|\Big| \leq \|x - y\|,$$

wobei die letzte Abschätzung nach der zweiten Dreiecksungleichung gilt (vgl. (N3)).

Satz 11.26. *Es seien (X, d_X) und (Y, d_Y) metrische Räume und $f\colon X \to Y$ eine Abbildung. Dann sind die folgenden Aussagen äquivalent:*

Im Video: **Satz 11.25**.

(a) *f ist stetig auf X;*

(b) *Es gilt $f\left(\overline{E}\right) \subset \overline{f(E)}$ für alle $E \subset X$;*

(c) *Falls $B = \overline{B} \subset Y$, so ist $A := f^{-1}(B)$ abgeschlossen in X;*

(d) *Falls $V = V^\circ \subset Y$, so ist $U := f^{-1}(V)$ offen in X.*

(Ohne Beweis)

Korollar 11.27. *Es sei (X, d) ein metrischer Raum und $f\colon X \to \mathbb{R}$ eine Abbildung. Dann sind die folgenden Aussagen äquivalent:*

Im Video: **Korollar 11.26**.

(a) *f ist stetig auf X;*

(b) *Die Menge $\{x \in X \ : \ f(x) < c\} = f^{-1}\big((-\infty, c)\big)$ ist für alle $c \in \mathbb{R}$ offen in X;*

(c) *Die Menge $\{x \in X \ : \ f(x) > c\} = f^{-1}\big((c, \infty)\big)$ ist für alle $c \in \mathbb{R}$ offen in X;*

(d) *Die Menge $\{x \in X \ : \ f(x) \leq c\} = f^{-1}\big((-\infty, c]\big)$ ist für alle $c \in \mathbb{R}$ abgeschlossen in X;*

(e) *Die Menge $\{x \in X \ : \ f(x) \geq c\} = f^{-1}\big([c, \infty)\big)$ ist für alle $c \in \mathbb{R}$ abgeschlossen in X.*

Insbesondere gilt:

$$\begin{aligned} &f \text{ ist stetig auf } X \\ \Rightarrow \quad &\text{Alle Niveaumengen } \{x \in X \ : \ f(x) = c\}, \ c \in \mathbb{R}, \ \text{sind abgeschlossen in } X. \end{aligned}$$

Beweisskizze. Es sei $c \in \mathbb{R}$. Die Mengen $(-\infty, c)$ und (c, ∞) sind offensichtlich offen in $\mathbb{R}$, während $(-\infty, c]$ und $[c, \infty)$ in $\mathbb{R}$ abgeschlossen sind. Wenn nun f stetig ist – es gilt also (a) –, so sind für alle $c \in \mathbb{R}$ nach Satz 11.26

- $f^{-1}\big((-\infty, c)\big)$ und $f^{-1}\big((c, \infty)\big)$ offen in X sowie
- $f^{-1}\big((-\infty, c]\big)$ und $f^{-1}\big([c, \infty)\big)$ abgeschlossen in X.

Die andere Richtung der Äquivalenz ist aufwändiger zu zeigen. □

Bisher haben wir uns Aussagen angeschaut, die sich auf die Urbilder $f^{-1}(B)$ einer Funktion $f\colon X \to Y$, $B \subset Y$, bezogen haben. Nun folgt eine Aussage, die sich auf Bilder $f(A)$, $A \subset X$, der Funktion bezieht.

Satz 11.28. *Es seien (X, d_X) und (Y, d_Y) metrischer Räume, $f\colon X \to Y$ eine stetige Funktion und $A \subset X$ kompakt in X. Dann ist das Bild $f(A)$ kompakt in Y.*

Im Video: **Satz 11.27**.

Beweis. Es sei $(y_n)_{n\in\mathbb{N}} \subset f(A)$ eine Folge. Dann existiert eine Folge $(a_n)_n \subset A$ mit $f(a_n) = y_n$ für alle $n \in \mathbb{N}$. Da A nach Voraussetzung kompakt ist, existiert eine Teilfolge $(a_{n_k})_{k\in\mathbb{N}} \subset (a_n)$ und ein $a \in A$ mit

$$\lim_{k\to\infty} a_{n_k} = a.$$

Da f ebenfalls nach Voraussetzung stetig ist, gilt zusätzlich

$$\lim_{k\to\infty} f\left(a_{n_k}\right) = f(a) \in f(A).$$

Also besitzt die Folge $(y_n)_n$ eine Teilfolge $(y_{n_k})_k$, die gegen einen Punkt $f(a) \in f(A)$ konvergiert – $f(A)$ ist damit kompakt. □

Nun folgt ein wichtiger Spezialfall:

Korollar 11.29. *Es sei (X, d) ein metrischer Raum, $A \subset X$ kompakt und $f\colon X \to \mathbb{R}$ eine stetige Funktion. Dann existieren $x_-, x_+ \in A$, so dass*

Im Video: **Korollar 11.28**.

$$\forall\, x \in A: \quad f(x_-) \leq f(x) \leq f(x_+),$$

genauer gesagt

$$\max_{x\in A} f(x) = f(x_+) \quad \text{und} \quad \min_{x\in A} f(x) = f(x_-).$$

Die Funktion f nimmt auf jeder kompakten Teilmenge sein Minimum und Maximum an.

Beweis. Es sei $A \subset X$ kompakt. Nach Satz 11.28 ist dann $f(A)$ kompakt in $\mathbb{R}$. Nach Satz 11.19 ist $f(A)$ also abgeschlossen und beschränkt. Nach Definition 3.1 aus der Vorlesung zur Analysis I (Lehramt) besitzt dann $f(A)$ ein Minimum und ein Maximum. $\square$

Definition 11.30. Es seien (X, d_X) und (Y, d_Y) metrische Räume, $x_0 \in X$ ein Häufungspunkt von X und $f\colon X \setminus \{x_0\} \to Y$ eine Abbildung. Existiert ein $c \in Y$, so dass

Im Video: **Def. 11.29**.

$$\forall\, (x_n)_{n\in\mathbb{N}} \subset X \setminus \{x_0\} \text{ mit } \lim_{n\to\infty} x_n = x_0 \text{ gilt } \lim_{n\to\infty} f(x_n) = c,$$

so *besitzt f in x_0 den Grenzwert c.* Man schreibt auch

$$\lim_{x\to x_0} f(x) = c \quad \text{oder} \quad f(x) \xrightarrow[x\to x_0]{} c.$$

LINK: TEIL 1 DER 16. VORLESUNG VOM 08.06.2022

Eine äquivalente Umformulierung lautet: Ist die *Fortsetzung*

$$F\colon X \to Y, \quad x \mapsto F(x) := \begin{cases} c & \text{für } x = x_0, \\ f(x) & \text{sonst,} \end{cases}$$

der Funktion f von $X \setminus \{x_0\}$ auf X stetig, so hat f den Limes c in x_0.

Ist x_0 kein Häufungspunkt von X, so ist jede Fortsetzung F von f stetig in x_0.

Ähnlich wie in $\mathbb{R}$ gilt:

Satz 11.31. *Es seien (X, d_X) und (Y, d_Y) metrische Räume, X kompakt und $f\colon X \to Y$ eine stetige Funktion. Dann gelten:*

Im Video: **Satz 11.30**.

(a) *f ist sogar gleichmäßig stetig;*

(b) *Falls f bijektiv ist, so ist $f^{-1}\colon Y \to X$ stetig auf Y.*

Beweis.

(a) Der Beweis erfolgt analog wie in $\mathbb{R}$, siehe dazu den Beweis von Lemma 14.14 des Skripts aus der Vorlesung zur Analysis I (Lehramt).

(b) Es sei $y_0 \in Y$ beliebig.

Behauptung. Die Umkehrabbildung f^{-1} ist stetig in y_0.

Beweis der Behauptung. Es sei $(y_n)_{n\in\mathbb{N}} \subset Y$ eine Folge mit $y_n \underset{n\to\infty}{\longrightarrow} y_0$. Setze nun

$$x_n := f^{-1}(y_n) \quad \text{und} \quad x_0 := f^{-1}(y_0).$$

Wegen der Bijektivität von f sind die Ausdrücke wohldefiniert, es gilt also $x_0 \in X$ sowie $x_n \in X$ für alle $n \in \mathbb{N}$. Da nach Voraussetzung X kompakt ist, existiert eine Teilfolge $(x_{n_k})_{k\in\mathbb{N}} \subset (x_n) \subset X$ mit

$$\lim_{k\to\infty} x_{n_k} = \tilde{x}.$$

Aufgrund der Stetigkeit von f und wegen $y_n \underset{n\to\infty}{\longrightarrow} y_0$ folgt dann

$$f(\tilde{x}) = \lim_{k\to\infty} f\left(x_{n_k}\right) = \lim_{k\to\infty} y_{n_k} = y_0 = f(x_0),$$

also $\tilde{x} = x_0$, da f insbesondere injektiv ist.

Konvergiert aber die ganze Folge (x_n) gegen den Grenzwert x_0? Falls dies nicht der Fall ist, so existiert ein $\varepsilon > 0$ und eine (andere) Teilfolge $\left(x_{n_j}\right)_{j\in\mathbb{N}} \subset (x_n)$ mit

$$d\left(x_{n_j}, x_0\right) > \varepsilon \quad \text{für } j \in \mathbb{N}.$$

Wegen der Kompaktheit von X existiert dann eine Teilteilfolge $\left(x_{n_{j_l}}\right)_{l\in\mathbb{N}} \subset \left(x_{n_j}\right)_j$, die gegen einen Punkt $x' \in X$ konvergiert. Wie oben folgt dann $x' = x_0$, im Widerspruch zu

$$d\left(x_{n_{j_l}}, x_0\right) > \varepsilon \quad \forall\, l \in \mathbb{N}.$$

Also gilt

$$\lim_{n\to\infty} x_n = x_0$$

und f^{-1} ist stetig in y_0.

Da $y_0 \in Y$ beliebig gewählt ist, ist f^{-1} stetig auf ganz Y. □

Definition 11.32.

Im Video: **Def. 11.31**.

(a) Es sei (X, d) ein metrischer Raum und $I \subset \mathbb{R}$ ein Intervall. Eine stetige Abbildung $\gamma\colon I \to X$ heißt dann *Weg* und die Bildmenge $\gamma(I) \subset X$ bezeichnet wir als *Spur* bzw. *Bahn von* γ.

(b) Ein metrischer Raum (X, d) heißt *wegzusammenhängend*, falls es zu beliebigen Punkten $x, y \in X$ einen Weg $\gamma\colon [0,1] \to X$ mit

$$\gamma(0) = x \quad \text{und} \quad \gamma(1) = y$$

gibt. Dann heißen x *Anfangspunkt von* γ und y *Endpunkt von* γ.

(c) Eine Teilmenge $U \subset X$ heißt *Gebiet*, falls U offen und wegzusammenhängend ist.

(d) Es sei V ein Vektorraum über $\mathbb{R}$ und $M \subset V$. Dann heißt M *konvex*, falls

$$\forall\, x, y \in M, \quad \forall\, t \in [0,1] : \quad t \cdot x + (1-t) \cdot y \in M.$$

Beispiel 11.33.

Im Video: **Beispiel 11.32**.

(a) Jede konvexe Teilmenge $M \subset V = X$ ist wegzusammenhängend.

(b) In einem metrischen Raum (V, d) ist nicht jeder offene Ball $B_\varepsilon(x_0) \subset V$, $\varepsilon > 0$, wegzusammenhängend: Betrachte dazu die Menge $V = \mathbb{R}$ mit dem Abstand $d(x,y) = |x - y|$. Dann ist $(\mathbb{R}, d)$ ein metrischer Raum, ebenso (X, d) für $X = (-\infty, 0) \cup (0, \infty) \subset \mathbb{R}$. Aber der „Ball“

$$B_2(1) = \Big\{ x \in X \ : \ |x - 1| < 2 \Big\} = (-1, 0) \cup (0, 3)$$

ist nicht wegzusammenhängend.

Um den Wegzusammenhang sicherzustellen ist ein normierter Raum notwendig, genauer: Jeder offene Ball $B_\varepsilon(x_0) \subset V$, $\varepsilon > 0$ in einem normierten Raum $(V, \|\cdot\|)$ ist wegzusammenhängend, da $B_\varepsilon(x_0)$ dann konvex ist.

(c) Jeder Untervektorraum von V ist konvex.

(d) Für $M \subset \mathbb{R}$ gilt:

$$M \text{ ist wegzusammenhängend} \quad \Leftrightarrow \quad M \text{ ist ein Intervall} \quad \Leftrightarrow \quad M \text{ ist konvex.}$$

(e) Die Menge

$$A := \Big\{ (x,y) \in \mathbb{R}^2 \ : \ |y| < |x| \Big\}$$

ist nicht wegzusammenhängend, da $(0,0) \notin A$ ist. Für $x = (-1, 0) \in A$ und $y = (1, 0) \in A$ existiert beispielsweise kein Weg $\gamma \colon [0,1] \to A$.

Satz 11.34. *Es seien (X, d_X) und (Y, d_Y) metrische Räume, $f \colon X \to Y$ eine stetige Funktion und $M \subset X$ wegzusammenhängend. Dann gelten:*

Im Video: **Satz 11.33**.

(a) *Die Bildmenge $f(M) \subset Y$ ist wegzusammenhängend;*

(b) (Zwischenwertsatz)
Gilt insbesondere $Y = \mathbb{R}$, so ist $f(M) \subset \mathbb{R}$ ein Intervall.

Eine wegzusammenhängende Menge kann durch eine stetige Funktion f nicht auseinandergerissen werden.

Beweis. Es seien $y, y' \in f(M)$ beliebig. Dann existieren $x, x' \in M$ mit

$$y = f(x) \quad \text{und} \quad y' = f(x').$$

Da M nach Voraussetzung wegzusammenhängend ist, existiert ein Weg $\gamma\colon [0,1] \to M$ mit

$$\gamma(0) = x \quad \text{und} \quad \gamma(1) = x'.$$

Die Verkettung $\sigma = f \circ \gamma\colon [0,1] \to f(X) \subset Y$ ist dann ebenfalls ein Weg und erfüllt

$$\sigma(0) = f\big(\gamma(0)\big) = f(x) = y \quad \text{und} \quad \sigma(1) = f\big(\gamma(1)\big) = f(x') = y'.$$

Also ist das Bild $f(M)$ wegzusammenhängend. □

12 Partielle Differenzierbarkeit von Funktionen mehrerer Variablen

LINK: TEIL 2 DER 16. VORLESUNG VOM 08.06.2022

Definition 12.1. Es sei $D \subset \mathbb{R}^n$ offen, $f\colon D \to \mathbb{R}^m$ eine Funktion und $a = (a_1, \dots, a_n) \in D$. Für $k \in \{1, \dots, n\}$ heißt f dann

- *partiell differenzierbar in a nach den Variablen x_k*, falls der Limes

$$\begin{aligned}\big(\mathrm{D}_k f\big)(a) &:= \lim_{h\to 0} \frac{f(a_1, \dots, a_{k-1}, a_k + h, a_{k+1}, \dots, a_n) - f(a)}{h} \\ &= \lim_{h\to 0} \frac{f(a + h \cdot e_k) - f(a)}{h}\end{aligned}$$

  existiert. Dann heißt $\big(\mathrm{D}_k f\big)(a) \in \mathbb{R}^m$ *partielle Ableitung von f in a nach x_k* und wird auch mit

$$\mathrm{D}_k f(a), \quad \frac{\partial}{\partial x_k} f(a), \quad \big(\partial_k f\big)(a) \quad \text{und} \quad \big(\partial_{x_k} f\big)(a)$$

  bezeichnet;

- *partiell differenzierbar nach x_k (in D)*, falls f für alle $a \in D$ partiell differenzierbar in a nach x_k ist. In diesem Fall definiert $a \mapsto \big(\mathrm{D}_k f\big)(a)$ eine Abbildung

$$f_{x_k} := \big(\mathrm{D}_k f\big)\colon D \to \mathbb{R}^m$$

  und wird *partielle Ableitung von f nach x_k* genannt.

Die Funktion f heißt dann *partiell differenzierbar (in D)*, falls f für alle $k \in \{1, \dots, n\}$ und alle $a \in D$ partiell differenzierbar in a nach x_k ist.

Beachte: Der Begriff „Partielle Ableitung“ kann einen Vektor $\big(\mathrm{D}_k f\big)(a)$ oder eine Abbildung $\big(\mathrm{D}_k f\big)$ meinen.

Beispiel. Für $n = 2$ und $m = 1$ haben wir

$$\partial_x f(x_0, y_0) := \big(\mathrm{D}_1 f\big)(x_0, y_0) = \frac{\partial f}{\partial x}(x_0, y_0) = \lim_{h \to 0} \frac{f(x_0 + h, y_0) - f(x_0, y_0)}{h}$$

und

$$\partial_y f(x_0, y_0) := \big(\mathrm{D}_2 f\big)(x_0, y_0) = \frac{\partial f}{\partial y}(x_0, y_0) = \lim_{h \to 0} \frac{f(x_0, y_0 + h) - f(x_0, y_0)}{h}.$$

Man erkennt daran die Idee hinter dem partiellen Ableiten: Man betrachtet die x_k-Koordinate als variabel, alle anderen $(n-1)$ Koordinaten als „eingefroren“ bzw. fixiert und wendet dann die Definition aus der Analysis I (Lehramt) an.

Beispiel 12.2.

(a) Gegeben sei die Funktion

$$f\colon \mathbb{R}^3 \to \mathbb{R}, \qquad \begin{pmatrix} x \\ y \\ z \end{pmatrix} \mapsto x^2 \cdot \sin(yz).$$

Dann gilt

$$\begin{aligned} \big(\mathrm{D}_1 f\big)(x, y, z) &= 2x \cdot \sin(yz), \\ \big(\partial_y f\big)(x, y, z) &= x^2 \cdot \cos(yz) \cdot z = \big(\mathrm{D}_2 f\big)(x, y, z), \\ \frac{\partial f}{\partial z}(x, y, z) &= x^2 \cdot \cos(yz) \cdot y. \end{aligned}$$

(b) Wir betrachten die Abbildung

$$r\colon \mathbb{R}^n \to \mathbb{R}, \quad x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \mapsto \|x\|_2 = \left(\sum_{i=1}^n x_i^2\right)^{\frac{1}{2}}.$$

Für $x \neq \mathbf{0}$ und $k \in \{1, \ldots, n\}$ gilt

$$\frac{\partial r}{\partial x_k}(x) = \frac{1}{2}\left(\sum_{i=1}^n x_i^2\right)^{-\frac{1}{2}} \cdot 2x_k = \frac{x_k}{\|x\|_2} = \frac{x_k}{r(x)}.$$

Man bemerke, dass x_j für alle $j \neq k$ eine Konstante bezüglich x_k ist.

(c) Es sei $f\colon [0,\infty) \to \mathbb{R}$ differenzierbar und r wie in Teil (b). Dann heißt die Funktion

$$g = f \circ r\colon \mathbb{R}^n \to \mathbb{R}$$

rotationssymmetrisch. Für $x \neq \mathbf{0}$ und $k \in \{1, \dots, n\}$ gilt dann

$$\big(\mathrm{D}_k g\big)(x) = \frac{\partial g}{\partial x_k}(x) = f'\big(r(x)\big) \cdot \frac{\partial r}{\partial x_k}(x) = f'\big(r(x)\big) \cdot \frac{x_k}{r(x)} = x_k \cdot \frac{f'\big(r(x)\big)}{r(x)}.$$

Man bemerke, dass wir für das zweite Gleichheitszeichen die Kettenregel angewandt haben (Beweis als Übung). Die formale Argumentation lautet:

$$\begin{aligned}\frac{\partial g}{\partial x_k}(x) =& \frac{\partial (f \circ r)}{\partial r(x)} \cdot \frac{\partial r(x)}{\partial x_k} \\ =& \frac{\partial f}{\partial r}\bigg|_{r=\|x\|_2} \cdot \frac{x_k}{r(x)} \\ =& f'\big(r(x)\big) \cdot \frac{x_k}{r(x)}.\end{aligned}$$

Den Beweis in einem allgemeineren Kontext behandeln wir später.

Anzumerken ist, dass die partielle Differenzierbarkeit keine starke Eigenschaft ist. Es gilt insbesondere

$$\text{Partielle Differenzierbarkeit} \quad \not\Rightarrow \quad \text{Stetigkeit.}$$

Beispiel 12.3. Wir betrachten die Funktion

$$f\colon \mathbb{R}^2 \to \mathbb{R}, \qquad \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{cases} \frac{xy}{x^2+y^2}, & \text{für } (x,y) \neq (0,0), \\ 0 & \text{sonst.} \end{cases}$$

In Beispiel 11.25 (d) haben wir bereits gezeigt, dass f nicht stetig in $(0,0)$ ist. Aber tatsächlich ist f partiell differenzierbar auf $\mathbb{R}^2$: Für $(x,y) \neq (0,0)$ ergibt sich

$$\big(\partial_x f\big)(x,y) = \frac{y \cdot (x^2+y^2) - xy \cdot (2x)}{(x^2+y^2)^2} = \frac{y \cdot (y^2 - x^2)}{(x^2+y^2)^2}$$

und

$$\big(\partial_y f\big)(x,y) = \frac{x \cdot (x^2 - y^2)}{(x^2+y^2)^2}.$$

Da

$$f(x,0) = 0 = f(0,y) \quad \forall\, x, y \in \mathbb{R}$$

gilt, haben wir entlang der beiden Koordinatenachsen eine konstante Funktion. Damit sind in $(0,0)$ beide partielle Ableitungen gleich Null:

$$\Big(\partial_x f\Big)(0,0) = 0 = \Big(\partial_y f\Big)(0,0).$$

Also ist f partiell differenzierbar auf $\mathbb{R}^2$.

Beachte: Für $x \neq 0 \neq y$ sind

$$f_x(0,y) = \frac{y^3}{y^4} = \frac{1}{y} \quad \text{und} \quad f_y(x,0) = \frac{x^3}{x^4} = \frac{1}{x}$$

beides unbeschränkte Funktionen auf jedem $B_\varepsilon\Big((0,0)\Big) \setminus \{(0,0)\}$, $\varepsilon > 0$. Tatsächlich ist dies das einzige, was schiefgehen kann.

Satz 12.4. *Es sei $D \subset \mathbb{R}^n$ offen, $f\colon D \to \mathbb{R}^m$ partiell differenzierbar in D und für jedes $k \in \{1,\dots,n\}$ die partielle Ableitung*

$$\partial_k f := \mathrm{D}_k f\colon D \to \mathbb{R}^m$$

beschränkt, d.h. die Bildmenge $f_k(D) \subset \mathbb{R}^m$ ist beschränkt. Dann ist f stetig auf D.

Beweis. Nach Satz 11.24 ist

$$f = (f_1,\dots,f_m)\colon D \to \mathbb{R}^m$$

genau dann stetig, wenn für $j \in \{1,\dots m\}$ jede Koordinatenfunktion

$$f_j\colon D \to \mathbb{R}$$

stetig ist. Also reicht es aus, im Folgenden nur den Fall $m = 1$ zu betrachten. Da D nach Voraussetzung offen ist, existiert zu jedem $x \in D$ ein $\delta > 0$ mit

$$B_{\sqrt{n}\cdot\delta}(x) \subset D.$$

Wegen $\|z\|_2 \leq \sqrt{n}\|z\|_\infty$ (vgl. Beispiel 10.2 (d)) gilt dann

$$\Big\{y \in \mathbb{R}^n \ : \ \|x-y\|_\infty < \delta\Big\} \subset \Big\{y \in \mathbb{R}^n \ : \ \|x-y\|_2 < \sqrt{n}\cdot\delta\Big\} \subset D.$$

Es sei nun $h = (h_1,\dots,h_n) \in \mathbb{R}^n$ mit $\|h\|_\infty < \delta$. Für $l \in \{0,1,2,\dots,n\}$ setzen wir dann

$$z^{(l)} := x + \sum_{i=1}^{l} h_i \cdot e_i.$$

Dabei gilt

$$z^{(0)} = x \quad \text{und} \quad z^{(n)} = x + h.$$

Dabei unterscheiden sich $z^{(l)}$ und $z^{(l-1)}$ nur in der l-ten Koordinate, aber liegen alle in $B_{\sqrt{n}\delta}(x)$. Insbesondere gilt

$$f(x+h) - f(x) = \sum_{l=1}^{n} \left(f\left(z^{(l)}\right) - f\left(z^{(l-1)}\right) \right).$$

Nach dem Mittelwertsatz der Differentialrechnung in $\mathbb{R}$ (vgl. Satz 12.4 aus dem Skript der Vorlesung zur Analysis I (Lehramt)) existiert dann ein $\xi^{(l)} \in \left[z^{(l-1)}, z^{(l)}\right]$ für $l \in \{1, \dots, n\}$, so dass

$$f(x+h) - f(x) = \sum_{l=1}^{n} \left(f\left(z^{(l)}\right) - f\left(z^{(l-1)}\right) \right) = \sum_{l=1}^{n} \left(\mathrm{D}_l f\right)(\xi^{(l)}) \cdot h_l.$$

Für $c := \sum_{l=1}^{n} \|\mathrm{D}_l f\|_\infty < \infty$ und wegen

$$h_l \leq \max_{l=1}^{n} |h_l| = \|h\|_\infty$$

folgt dann

$$|f(x+h) - f(x)| \leq c \cdot \|h\|_\infty \leq c\|h\|_2.$$

Damit ist f sogar Lipschitz-stetig auf D, insbesondere also stetig. □

LINK: TEIL 1 DER 17. VORLESUNG VOM 09.06.2022

13 Parameterabhängige Integrale

Notation. Es seien (X, d_X) und (Y, d_Y) metrische Räume sowie $f\colon X \times Y \to \mathbb{R}^m$ eine Funktion. Für ein beliebiges, aber festes $x \in X$ setzen wir $f_x(y) = f(x, y)$. Damit haben wir eine Funktion

$$f_x\colon Y \to \mathbb{R}^m$$

definiert. Für ein beliebiges, aber festes $y \in Y$ setzen wir ebenso $f^y(x) = f(x, y)$ und definieren damit eine Abbildung

$$f^y\colon X \to \mathbb{R}^m.$$

Häufig sind $X \subset \mathbb{R}^n$ und $Y \subset \mathbb{R}^d$ für $n, d \in \mathbb{N}$. Dies begegnete uns schon im Kontext der

(a) partiellen Stetigkeit, in der wir für die Funktion

$$f\colon \mathbb{R}^2 \to \mathbb{R}, \quad \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{cases} \frac{xy}{x^2+y^2} & \text{für } (x,y) \neq (0,0), \\ 0 & \text{sonst,} \end{cases}$$

gesehen haben, dass $f_x\colon \mathbb{R} \to \mathbb{R}$ und $f^y\colon \mathbb{R} \to \mathbb{R}$ stetig für alle $x, y \in \mathbb{R}$ sind, aber $f\colon \mathbb{R}^2 \to \mathbb{R}$ nicht;

(b) partiellen Ableitung: Ist $f\colon \mathbb{R}^n \to \mathbb{R}^m$ gegeben und eine Koordinate $k \in \{1, 2, \dots, n\}$ ausgewählt, so kann man

$$\mathbb{R}^n = \overbrace{\mathbb{R} \times \mathbb{R} \times \ldots \times \underbrace{\mathbb{R}}_{\ni x_k} \times \ldots \mathbb{R}}^{n\text{-mal}} =: X \times Y$$

schreiben, wobei $X = \mathbb{R}^{n-1}$ und $Y = \mathbb{R}$. Bezeichne nun

$$\hat{x}_k = \Big(x_1, \ldots, x_k, x_{k+1}, \ldots, x_n\Big) \in \mathbb{R}^{n-1}$$

als einen Vektor, bei dem die k-te Koordinate des Vektors $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$ entfernt wurde. Dann ist

$$f_{\hat{x}_k}(x_k) = f(x_1, \ldots, x_n)$$

eine Abbildung $f_{\hat{x}_k} \colon \mathbb{R} \to \mathbb{R}^m$ gemäß der obigen Definition. Die k-te partielle Ableitung $\Big(\mathrm{D}_k f\Big)(x)$, auch gekennzeichnet durch

$$f'_{\hat{x}_k}(x_k) = \frac{\partial}{\partial x_k} f_{\hat{x}_k}(x_k) = \frac{\partial}{\partial x_k} f(x_1, \ldots, x_n),$$

ist dann gerade eine eindimensionale Ableitung (nach x_k) bei eingefrorenem $\hat{x}_k$.

(c) Was nun neu ist: Parameterabhängige Integrale
Betrachten wir dazu die folgende Idee: Wir wollen über eine ausgezeichnete Variable integrieren, während die anderen eingefroren und als Parameter betrachtet werden. Für $a, b \in \mathbb{R}$, $a < b$ sei dazu $Y = [a, b]$ und (X, d) ein metrischer Raum. Sei außerdem $f \colon X \times [a, b] \to \mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ eine Funktion derart, dass für jedes $x \in X$ die *partielle Abbildung* $f_x \colon [a, b] \to \mathbb{R}$ Riemann-integrierbar (z.B. stetig) ist. Dann ist die Funktion

$$F \colon X \to \mathbb{K}, \quad F(x) := \int_a^b f_x(y)\,\mathrm{d}y = \int_a^b f(x, y)\,\mathrm{d}y \tag{PAI}$$

wohldefiniert. Um diese zu studieren, benötigen wir etwas Vorbereitung.

Lemma 13.1. *Es sei (Y, d) ein kompakter metrischer Raum. Dann ist*

$$C(Y, \mathbb{K}) := \Big\{ f \colon Y \to \mathbb{K} \;:\; f \text{ ist stetig auf } Y \Big\}$$

ein normierter Raum bezüglich der Norm

$$\|f\|_\infty = \sup_{y \in Y} |f(y)|.$$

Beweis. Es ist bereits bekannt, dass die Menge $C(Y)$ ein Vektorraum ist bezüglich der Operationen

$$(f + y)(y) = f(y) + g(y) \quad \text{und} \quad (\lambda f)(y) = \lambda \cdot (f(y)), \quad y \in Y,$$

ist. Ist außerdem $\mathbb{K} = \mathbb{R}$, so gilt nach Korollar 11.29 die Identität

$$\sup_{y \in Y} |f(y)| = \max_{y \in Y} |f(y)| \in \mathbb{R} \quad \forall\, f \in C(Y),$$

da Y kompakt ist. Ist $\mathbb{K} = \mathbb{C}$, so muss man

$$f_1 \colon Y \to \mathbb{R}, \quad f_1 = \mathrm{Re}(f),$$

$$f_2\colon Y \to \mathbb{R}, \quad f_2 = \operatorname{Im}(f),$$

betrachten. Man wende nun das Obige auf f_1 und f_2 an und es folgt

$$\sup_{y\in Y} |f(y)| \leq \sup_{y\in Y} |f_1(y)| + \sup_{y\in Y} |f_2(y)| < \infty.$$

Also ist $\|f\|_\infty < \infty$ für alle $f \in C(Y,\mathbb{K})$. Dass die Norm $\|\cdot\|_\infty$ die Normeigenschaften (N1) bis (N3) erfüllt, kann man einfach nachrechnen. □

Satz 13.2. *Es seien (X, d_X) und (Y, d_Y) metrische Räume, Y kompakt, $f\colon X \to Y \to \mathbb{K}$ eine (nicht nur partiell, sondern simultan) stetige Funktion und $x_0 \in X$. Dann gibt es zu jedem $\varepsilon > 0$ ein $\delta = \delta(\varepsilon, x_0) > 0$, so dass*

$$\forall\, x \in B_\delta(x_0), y \in Y: \quad |f(x,y) - f(x_0,y)| < \varepsilon.$$

Dabei sind d_X und d_Y häufig die von der Norm $\|\cdot\|_2$ induzierten Abstände.

Beweis. Widerspruchsannahme: Angenommen man findet ein $\varepsilon_0 > 0$ und zu jedem $n \in \mathbb{N}$ Punkte $x_n \in X$ und $y_n \in Y$, so dass

$$x_n \in B_{\frac{1}{n}}(x_0), \text{ aber } \Big|f(x_n,y_n) - f(x_0,y_n)\Big| \geq \varepsilon_0. \tag{$*$}$$

Offensichtlich impliziert ($*$) dann, dass $x_n \xrightarrow[n\to\infty]{} x_0$. Da Y kompakt ist, existiert dann eine Teilfolge $(y_{n_k})_{k\in\mathbb{N}} \subset (y_n)_n$ mit $y_{n_k} \xrightarrow[k\to\infty]{} y_0$ für ein $y_0 \in Y$. Da f (insbesondere) in (x_0,y_0) stetig ist, folgen partielle Stetigkeit

$$f(x_0, y_{n_k}) \xrightarrow[k\to\infty]{} f(x_0,y_0)$$

und echte Stetigkeit

$$f(x_{n_k}, y_{n_k}) \xrightarrow[k\to\infty]{} f(x_0,y_0).$$

Es folgt

$$\Big|f(x_{n_k}, y_{n_k}) - f(x_0, y_{n_k})\Big| \leq \Big|f(x_{n_k}, y_{n_k}) - f(x_0,y_0)\Big| + \Big|f(x_0,y_0) - f(x_0, y_{n_k})\Big| \xrightarrow[k\to\infty]{} 0,$$

da beide Summanden Nullfolgen sind. Dies steht jedoch im Widerspruch zur Annahme. □

Link: Teil 2 der 17. Vorlesung vom 09.06.2022

Mit fast demselben Beweis folgt diese Umformulierung:

Satz 13.3. *Es seien (X, d_X) und (Y, d_Y) metrische Räume, Y kompakt und $f\colon X \times Y \to \mathbb{K}$ stetig. Dann ist die Abbildung*

$$\Phi\colon X \to C(Y), \quad x \mapsto f_x,$$

stetig bezüglich der Norm $\|\cdot\|_\infty$ auf $C(Y)$.

Wenden wir nun Satz 13.2 direkt an:

Korollar 13.4. *Es sei (X, d) ein metrischer Raum, $a, b \in \mathbb{R}$ mit $a < b$ und $f\colon X \times [a, b] \to \mathbb{K}$ stetig. Dann ist die Funktion $F\colon X \to \mathbb{K}$ aus (PAI) stetig.*

Beweis. Es sei $x_0 \in X$ und $\varepsilon > 0$. Nach Satz 13.2 existiert ein $\delta > 0$, so dass

$$|f(x,t) - f(x_0,t)| < \frac{\varepsilon}{b-a} \quad \forall\, x \in B_\delta(x_0), y \in [a,b].$$

Für alle $x \in B_\delta(x_0)$ folgt dann

$$\begin{aligned}
|F(x) - F(x_0)| &= \left| \int_a^b \Big(f(x,y) - f(x_0,y) \Big) \,\mathrm{d}y \right| \\
&\leq \int_a^b \underbrace{|f(x,y) - f(x_0,y)|}_{<\frac{\varepsilon}{b-a}} \,\mathrm{d}y \\
&< \frac{\varepsilon}{b-a} \int_a^b 1 \,\mathrm{d}y = \varepsilon.
\end{aligned}$$

Also ist F stetig in x_0 für alle $x_0 \in X$ und damit stetig auf ganz X. □

Bei Potenzreihen (Grenzwerte von Summen) untersuchten wir, wann man die Ableitung in die Reihe hineinziehen darf. Das das Integral ebenfalls ein Grenzwert von Summen ist, ergibt sich eine Analogie dazu:

Satz 13.5. *Es sei $\Omega \subset \mathbb{R}^n$ offen, $a, b \in \mathbb{R}$ mit $a < b$ und $f\colon \Omega \times [a,b] \to \mathbb{K}$ stetig. Für jedes $k \in \{1, \dots, n\}$ möge die partielle Ableitung*

$$\frac{\partial}{\partial x_k} f\colon \Omega \times [a,b] \to \mathbb{K}$$

existieren und auf ganz $\Omega \times [a,b]$ stetig sein. Dann ist die Funktion $F\colon \Omega \to \mathbb{K}$ wie in (PAI) partiell differenzierbar auf Ω und für $x \in \Omega$ und $k \in \{1, \dots, n\}$ gilt

$$\frac{\partial}{\partial x_k} F(x) = \int_a^b \frac{\partial}{\partial x_k} f(x,y) \,\mathrm{d}y.$$

Beweis. Wir zeigen den Beweis nur für den Fall $n = 1$, es sei also $\Omega \subset \mathbb{R}$ ein Intervall. Seien weiter $x_0, x \in \Omega$ mit $x_0 \neq x$. Nach der Definition von $F(x)$ in (PAI) gilt dann

$$\frac{F(x) - F(x_0)}{x - x_0} - \int_a^b \partial_x f(x_0, y) \,\mathrm{d}y = \int_a^b \left(\frac{f(x,y) - f(x_0,y)}{x - x_0} - \partial_x f(x_0,y) \right) \mathrm{d}y.$$

Nach dem Mittelwertsatz der Differentialrechnung in $\mathbb{R}$ (vgl. Satz 12.4 aus dem Skript der Vorlesung zur Analysis I (Lehramt)) existiert dann ein $\xi = \xi_{x,y}$ zwischen x_0 und x derart, dass

$$\frac{f(x,y) - f(x_0,y)}{x - x_0} = \partial_x f(\xi, y).$$

Nach Voraussetzung ist $\partial_x f$ stetig auf $\Omega \times [a,b]$, während $[a,b]$ kompakt ist. Somit existiert nach Satz 13.2 ein $\delta > 0$, so dass für alle $x \in B_\delta(x_0) \cap \Omega$ und $y \in [a,b]$ gilt:

$$\left|(\partial_x f)(x,y) - (\partial_x f)(x_0,y)\right| < \frac{\varepsilon}{b-a}.$$

Da ξ zwischen x und x_0 liegt, gilt insbesondere $\xi \in B_\delta(x_0) \cap \Omega$. Für $x \in B_\delta(x_0) \cap \Omega$ folgt dann

$$\left|\frac{F(x)-F(x_0)}{x-x_0} - \int_a^b \partial_x f(x_0,y)\,\mathrm{d}y\right| \le \int_a^b \underbrace{\left|\partial_x f(\xi,y) - \partial_x f(x_0,y)\right|}_{<\frac{\varepsilon}{b-a}}\,\mathrm{d}y < \varepsilon$$

und damit die Behauptung. $\square$

Beispiel 13.6. Für das parameterabhängige Integral

$$\mathbb{R} \ni x \mapsto \underbrace{F(x)}_{\ge 0} := \int_0^1 \frac{\mathrm{e}^{-(1+y^2)x^2}}{1+y^2}\,\mathrm{d}y = \mathrm{e}^{-x^2} \int_0^1 \frac{\mathrm{e}^{-y^2x^2}}{\underbrace{1+y^2}_{\le 1}}\,\mathrm{d}y \le \mathrm{e}^{-x^2}$$

gilt nach Satz 13.5 und der Substitutionsregel der Integration für $u := \varphi(y) := xy$, also $\frac{\mathrm{d}u}{\mathrm{d}y} = x$:

$$\begin{aligned}
\partial_x F(x) = F'(x) &= \int_0^1 \partial_x \left(\frac{\mathrm{e}^{-(1+y^2)x^2}}{1+y^2}\right)\mathrm{d}y \\
&= \int_0^1 -2x \cdot (1+y^2)\frac{\mathrm{e}^{-(1+y^2)x^2}}{1+y^2}\,\mathrm{d}y \\
&= \int_0^1 -2x \cdot \mathrm{e}^{-x^2} \cdot \mathrm{e}^{-x^2y^2}\,\mathrm{d}y \\
&\overset{\text{Sub}}{=} -2x \cdot \mathrm{e}^{-x^2} \int_{\varphi(0)}^{\varphi(1)} \mathrm{e}^{-u^2} \cdot \frac{1}{x}\,\mathrm{d}u \\
&= -2 \cdot \mathrm{e}^{-x^2} \int_0^x \mathrm{e}^{-u^2}\,\mathrm{d}u.
\end{aligned}$$

Leite nun auch

$$\mathbb{R} \ni x \mapsto G(x) := \left(\int_0^x \mathrm{e}^{-u^2}\,\mathrm{d}u\right)^2$$

ab: Wir erhalten mit der Kettenregel

$$G'(x) = \mathrm{e}^{-x^2} \cdot 2\left(\int_0^x \mathrm{e}^{-u^2}\,\mathrm{d}u\right),$$

also
$$G'(x) = -F'(x) \quad \text{bzw.} \quad (F+G)'(x) = F'(x) + G'(x) = 0.$$
Also existiert eine Konstante $c \in \mathbb{R}$ mit
$$F(x) + G(x) = c \quad \forall\, x \in \mathbb{R}.$$
Um c zu bestimmen, setzen wir $x = 0$ ein und berechnen das resultierende Integral mithilfe der Substitutionsregel der Integration. Für $y := \varphi(t) := \tan(t)$ beachten wir dabei, dass $\frac{\mathrm{d}y}{\mathrm{d}t} = \varphi'(t) = 1 + \tan^2(t)$ gilt. Außerdem gilt
$$0 = \tan(0), \quad 1 = \tan\left(\frac{\pi}{4}\right) \quad \Rightarrow \quad \arctan(0) = 0, \quad \arctan(1) = \frac{\pi}{4}.$$
Damit erhalten wir
$$\begin{aligned} c = F(0) + G(0) &= \int_0^1 \frac{1}{1+y^2}\,\mathrm{d}y + 0 \\ &\overset{\text{Sub}}{=} \int_{\arctan(0)}^{\arctan(1)} \underbrace{\frac{1}{1+\tan^2(t)} \cdot (1+\tan^2(t))}_{=1}\,\mathrm{d}t = \arctan(1) = \frac{\pi}{4}. \end{aligned}$$
Es gilt also
$$\left(\int_0^x \mathrm{e}^{-y^2}\,\mathrm{d}y\right)^2 = G(x) = \frac{\pi}{4} - F(x).$$
Wegen $0 \le F(x) \le \mathrm{e}^{-x^2}$ folgt dann im Limes $x \to \infty$:
$$\left(\int_0^\infty \mathrm{e}^{-y^2}\,\mathrm{d}y\right)^2 = \frac{\pi}{4} - 0 \quad \Rightarrow \quad \int_0^\infty \mathrm{e}^{-y^2}\,\mathrm{d}y = \frac{\sqrt{\pi}}{2}.$$
Dies impliziert für die Gauß-Fehlerfunktion:
$$\lim_{x\to\infty} \operatorname{erf}(x) = 1.$$

LINK: TEIL 1 DER 18. VORLESUNG VOM 15.06.2022

Für $v := \varphi(u) := -u$, also $\frac{\mathrm{d}v}{\mathrm{d}u} = -1$, gilt außerdem nach der Substitutionsregel der Integralrechnung, dass
$$\int_0^\infty \mathrm{e}^{-u^2}\,\mathrm{d}u = \int_0^{-\infty} \mathrm{e}^{-v^2} \cdot (-1)\,\mathrm{d}v = \int_{-\infty}^0 \mathrm{e}^{-v^2}\,\mathrm{d}v = \frac{\sqrt{\pi}}{2}.$$
Es gilt also
$$\int_0^\infty \mathrm{e}^{-u^2}\,\mathrm{d}u + \int_{-\infty}^0 \mathrm{e}^{-u^2}\,\mathrm{d}u = \int_{-\infty}^\infty \mathrm{e}^{-u^2}\,\mathrm{d}u = \frac{\sqrt{\pi}}{2} + \frac{\sqrt{\pi}}{2} = \sqrt{\pi}.$$

14 Partielle Ableitungen höherer Ordnung

Definition 14.1. Es sei $\Omega \subset \mathbb{R}^n$ offen.

(a) Eine Funktion $f\colon \Omega \to \mathbb{R}^m$ heißt *stetig partiell differenzierbar*, falls alle partiellen Ableitungen $\mathrm{D}_1 f, \dots, \mathrm{D}_n f$ existieren und stetig sind auf Ω. Der Raum aller solcher Funktionen wird mit $C^1(\Omega, \mathbb{R}^m)$ bezeichnet; für $m = 1$ schreibt man schlichtweg $C^1(\Omega, \mathbb{R}) = C^1(\Omega)$.

Da $g := \mathrm{D}_k f$, $k \in \{1, \dots, n\}$, wieder eine Funktion $g\colon \Omega \to \mathbb{R}^m$ ist, macht es Sinn, Teil (a) wieder anzuwenden.

(b) Falls $g_k \in C^1(\Omega, \mathbb{R}^m)$ für alle $k \in \{1, \dots, n\}$ gilt, wobei wir $g_k = \mathrm{D}_k f$ setzen, so nennen wir f *zweimal stetig differenzierbar* und schreiben $f \in C^2(\Omega, \mathbb{R}^m)$. Dabei heißt

$$\mathrm{D}_j g_k = \mathrm{D}_j \mathrm{D}_k f$$

partielle Ableitung von f zweiter Ordnung (bezüglich x_k und x_j).

(c) Für $d \in \mathbb{N}$ definieren wir rekursiv:

$$C^d(\Omega, \mathbb{R}^m) := \Big\{ f \in C^1(\Omega, \mathbb{R}^m) \ : \ \mathrm{D}_k f \in C^{d-1}(\Omega, \mathbb{R}^m), k \in \{1, \dots, n\} \Big\}$$

und

$$C^\infty(\Omega, \mathbb{R}^m) := \bigcap_{d \in \mathbb{N}} C^d(\Omega, \mathbb{R}^m) \quad \text{sowie} \quad C^d(\Omega) = C^d(\Omega, \mathbb{R}^1).$$

Eine Funktion $f \in C^d(\Omega, \mathbb{R}^m)$ besitzt also alle *partiellen Ableitungen*

$$\mathrm{D}_{k_1} \mathrm{D}_{k_2} \dots \mathrm{D}_{k_l} f, \quad k_1, \dots, k_l \in \{1, \dots, n\},$$

der Ordnung $l \in \{1, \dots, d\}$.

Beispiel 14.2.

(a) Wir betrachten die Funktion

$$g\colon \mathbb{R}^3 \to \mathbb{R}, \quad g(x, y, z) := x \cdot \sin(yz).$$

Dann gilt:

$$\begin{aligned}
\mathrm{D}_1 g(x, y, z) &= \sin(yz), \\
\mathrm{D}_2 g(x, y, z) &= x \cdot z \cdot \cos(yz), \\
\mathrm{D}_3 g(x, y, z) &= x \cdot y \cdot \cos(yz), \\
\mathrm{D}_1 \mathrm{D}_1 g(x, y, z) &= 0,
\end{aligned}$$

$$\begin{aligned}
\mathrm{D}_2\mathrm{D}_2 g(x,y,z) &= -\,x\cdot z^2\cdot\sin(yz),\\
\mathrm{D}_3\mathrm{D}_3 g(x,y,z) &= -\,x\cdot y^2\cdot\sin(yz),\\
\mathrm{D}_1\mathrm{D}_3 g(x,y,z) &= y\cdot\cos(yz),\\
\mathrm{D}_3\mathrm{D}_1 g(x,y,z) &= y\cdot\cos(yz),\\
\mathrm{D}_2\mathrm{D}_3 g(x,y,z) &= x\cdot\cos(yz) - x\cdot yz\cdot\sin(yz),\\
\mathrm{D}_3\mathrm{D}_2 g(x,y,z) &= x\cdot\cos(yz) - x\cdot yz\cdot\sin(yz).
\end{aligned}$$

Man beobachtet, dass

$$\forall\, k,j\in\{1,2,3\}:\quad \mathrm{D}_j\mathrm{D}_k g = \mathrm{D}_k\mathrm{D}_j g \quad \text{auf } \Omega.$$

Aber ist das immer so? Dazu betrachten wir das nächste Beispiel:

(b) Gegeben sei die Funktion

$$f\colon \mathbb{R}^2\to\mathbb{R},\quad (x,y)\mapsto \begin{cases} xy\cdot\frac{x^2-y^2}{x^2+y^2} & \text{für } (x,y)\neq(0,0),\\ 0 & \text{sonst.}\end{cases}$$

Für $(x,y)\neq(0,0)$ gilt dann

$$\begin{aligned}
\mathrm{D}_1 f(x,y) &= y\cdot\frac{x^2-y^2}{x^2+y^2} + xy\cdot\frac{2x(x^2+y^2)-(x^2-y^2)2x}{(x^2+y^2)^2}\\
&= \frac{y}{(x^2+y^2)^2}\Big(\underbrace{(x^2-y^2)(x^2+y^2)}_{=x^4-y^4} + 2x^2(x^2+y^2-x^2+y^2)\Big)\\
&= y\cdot\frac{x^4+y^4+4x^2y^2}{(x^2+y^2)^2}\\
&\overset{x=0}{=} -\frac{y^5}{y^4} = -y.
\end{aligned}$$

Analog erhält man

$$\mathrm{D}_2 f(x,y) = x\cdot\frac{x^4-4x^2y^2-y^4}{(x^2+y^2)^2} \overset{y=0}{=} \frac{x^5}{x^4} = x.$$

Daraus folgt

$$\mathrm{D}_2\mathrm{D}_1 f(x,y) = \frac{x^6 + 9x^4y^2 - 9x^2y^4 - y^6}{(x^2+y^2)^3} = \mathrm{D}_1\mathrm{D}_2 f(x,y).$$

Für $x = y$ gilt insbesondere $\mathrm{D}_2\mathrm{D}_1 f(x,x) = \mathrm{D}_1\mathrm{D}_2 f(x,x) = 0$. Wie sieht es entlang der Koordinatenachsen aus? Für $x, y \in \mathbb{R}$ gilt $f(x,0) = 0 = f(0,y)$, insbesondere

$$\mathrm{D}_1(x,0) = 0 = \mathrm{D}_2 f(0,y) \quad \text{und} \quad \mathrm{D}_1(0,0) = 0 = \mathrm{D}_2 f(0,0).$$

Daraus folgt

$$\mathrm{D}_1(\mathrm{D}_2 f)(0, \overbrace{0}^{\text{fix}}) = \lim_{h\to 0} \frac{\mathrm{D}_2 f(0+h,0) - \mathrm{D}_2 f(0,0)}{h} = \lim_{h\to 0} \frac{1}{h} \cdot \big(h - 0\big) = 1$$

und

$$\mathrm{D}_2(\mathrm{D}_1 f)(\overbrace{0}^{\text{fix}}, 0) = \lim_{h\to 0} \frac{\mathrm{D}_1 f(0,h) - \mathrm{D}_1 f(0,0)}{h} = \lim_{h\to 0} \frac{-h-0}{h} = -1 \neq \mathrm{D}_1\mathrm{D}_2 f(0,0).$$

Fazit.

- Im Punkt $(0,0)$ stimmen $\mathrm{D}_1\mathrm{D}_2 f$ und $\mathrm{D}_2\mathrm{D}_1 f$ nicht überein;
- Außerdem sind $\mathrm{D}_1\mathrm{D}_2 f$ und $\mathrm{D}_2\mathrm{D}_1 f$ nicht stetig in $(0,0)$.

Satz 14.3 (von Schwarz)**.** *Es sei $\Omega \subset \mathbb{R}^n$ offen, $f \in C^1(\Omega, \mathbb{R}^m)$ und $k, j \in \{1, \dots, n\}$ fix. Es sei außerdem $\mathrm{D}_j\mathrm{D}_k f$ eine auf Ω stetige Funktion. Dann ist auch $\mathrm{D}_k\mathrm{D}_j f$ wohldefiniert sowie stetig und es gilt*

$$\mathrm{D}_k\mathrm{D}_j f(x) = \mathrm{D}_j\mathrm{D}_k f(x) \quad \forall\, x \in \Omega. \tag{$*$}$$

Insbesondere gilt $(*)$ für alle $f \in C^2(\Omega, \mathbb{R}^m)$.

LINK: TEIL 2 DER 18. VORLESUNG VOM 15.06.2022

Ein nützliches Werkzeug für den Beweis, aber auch allgemein hilfreich, ist die folgende

Beobachtung 14.4. Es sei $\Omega \subset \mathbb{R}^n$ offen, $f = (f_1, \dots, f_m)\colon \Omega \to \mathbb{R}^m$ eine Funktion, $a \in \Omega$ und $d \in \mathbb{N}_0$ sowie $k \in \{1, \dots, n\}$. Dann gelten die folgenden Aussagen:

(a) f ist genau dann partiell differenzierbar in a bezüglich x_k, wenn f_l für alle $l \in \{1, \dots, m\}$ partiell differenzierbar ist in a bezüglich x_k;

(b) Es gilt

$$\mathrm{D}_k f \in C^n(\Omega, \mathbb{R}^m) \quad \Leftrightarrow \quad \mathrm{D}_k(f_l) \in C^n(\Omega, \mathbb{R}^m) \quad \forall\, l \in \{1, \dots, m\}.$$

Beweis.

(a) f ist nach Definition partiell differenzierbar in $a \in \Omega$ bezüglich x_k, falls ein Vektor $b \in \mathbb{R}^m$ existiert mit
$$\lim_{h\to 0} \frac{1}{h}\Big(f(a+he_k) - f(a)\Big) = b.$$
In diesem Fall schreiben wir $b =: (\mathrm{D}_k f)(a)$. Nach Lemma 11.2 (e) ist dies genau dann der Fall, wenn $b_1, \ldots, b_m \in \mathbb{R}$ existieren mit $b = (b_1, \ldots, b_m) \in \mathbb{R}^m$ und
$$\forall\, l \in \{1, \ldots, m\}: \quad \lim_{h\to 0} \frac{1}{h}\Big(f_l(a+he_k) - f_l(a)\Big) - b_l = 0.$$
Und dies ist äquivalent dazu, dass f_l für alle $l \in \{1, \ldots, m\}$ partiell differenzierbar ist in a bezüglich x_k.

(b) Wende nun ganz analog wie in Teil (a) das Lemma 11.2 (e) an, um die Stetigkeit der vektorwertigen Funktion $\mathrm{D}_k f\colon \Omega \to \mathbb{R}^m$ durch die Stetigkeit aller Koordinatenfunktionen $\mathrm{D}_k f_1, \ldots, \mathrm{D}_k f_m\colon \Omega \to \mathbb{R}$ zu charakterisieren. □

Beweis von Satz 14.3. Nach der Beobachtung 14.4 reicht es aus, nur den Fall $m = 1$ zu betrachten. Da in der Aussage nur zwei Variablen x_k und x_j, $k, j \in \{1, \ldots, n\}$, involviert sind, kann man die restlichen Variablen einfrieren und die Funktion
$$g\colon \mathbb{R}^2 \to \mathbb{R}, \quad g(x_k, x_j) = f(x_1, x_2, \ldots, x_n)$$
betrachten. Es reicht also aus, den Satz nur für den Fall $n = 2$ zu zeigen. Der Fall $j = k$ ist trivial, denn es gilt offensichtlich
$$\mathrm{D}_k\mathrm{D}_j f(x) = \mathrm{D}_k\mathrm{D}_k f(x) = \mathrm{D}_j\mathrm{D}_k f(x) \quad \forall\, x \in \Omega.$$
Wir betrachten nun den Fall $j \neq k$. Dazu beweisen wir $(*)$ unter der stärkeren Annahme, dass $f \in C^2(\Omega, \mathbb{R})$. Es bleibt also für einen beliebigen Punkt $(x_0, y_0) \in \Omega \subset \mathbb{R}^2$ zu zeigen, dass
$$\mathrm{D}_1\mathrm{D}_2 f(x_0, y_0) = \mathrm{D}_2\mathrm{D}_1 f(x_0, y_0).$$
Durch eine globale Verschiebung des Koordinatensystems dürfen wir sogar annehmen, dass der Punkt $(x_0, y_0) = (0, 0)$ ist.

<u>Idee:</u> Stelle den relevanten Differenzenquotienten zweiter Ordnung auf zwei Weisen dar und bilde anschließend den Grenzwert.

Für beliebige $s, t, r > 0$ definieren wir
$$g(r) := f(r, t) - f(r, 0) \quad \Rightarrow \quad g'(r) = \mathrm{D}_1 f(r, t) - \mathrm{D}_1 f(r, 0),$$
und
$$\Delta(s, t) = \Big(f(s, t) - f(0, t)\Big) - \Big(f(s, 0) - f(0, 0)\Big).$$
Wir erhalten
$$\Delta(s, t) = \Big(f(s, t) - f(s, 0)\Big) - \Big(f(0, t) - f(0, 0)\Big) = g(s) - g(0).$$

Nach dem Mittelwertsatz der Differentialrechnung (vgl. Satz 12.4 aus dem Skript der Vorlesung zur Analysis I (Lehramt)) existiert dann ein $\xi \in (0,s)$ mit

$$\Delta(s,t) = g'(\xi) \cdot (s-0) = s\Big(\mathrm{D}_1 f(\xi,t) - \mathrm{D}_1 f(\xi,0)\Big).$$

Für $h(r) := \mathrm{D}_1 f(\xi,r)$, also $h'(r) = \mathrm{D}_2\mathrm{D}_1 f(\xi,r)$, erhalten wir

$$\Delta(s,t) = s \cdot \Big(h(t) - h(0)\Big).$$

Nach erneuter Anwendung des Mittelwertsatzes existiert dann ein $\eta \in (0,t)$, so dass

$$\Delta(s,t) = s \cdot \Big(h'(\eta) \cdot (t-0)\Big) = st \cdot \mathrm{D}_2\mathrm{D}_1 f(\xi,\eta).$$

Dasselbe Vorgehen können wir auch auf die anderen Darstellungen von Δ anwenden und erhalten analog:

$$\Delta(s,t) = \Big(f(s,t) - f(0,t)\Big) - \Big(f(s,0) - f(0,0)\Big) = \ldots = st \cdot \mathrm{D}_1\mathrm{D}_2 f(\tilde{\xi},\tilde{\eta}) \quad \text{mit} \quad \tilde{\xi} \in (0,s), \tilde{\eta} \in (0,t).$$

Nun dividieren wir die Ausdrücke durch $s > 0$ und $t > 0$ und erhalten

$$\mathrm{D}_2\mathrm{D}_1 f(\xi,\eta) = \mathrm{D}_1\mathrm{D}_2 f(\tilde{\xi},\tilde{\eta}).$$

Der Grenzübergang $s \to 0, t \to 0$ impliziert außerdem $\xi, \eta, \tilde{\xi}, \tilde{\eta} \to 0$. Wegen $f \in C^2(\Omega,\mathbb{R})$ folgt dann

$$\mathrm{D}_2\mathrm{D}_1 f(0,0) = \mathrm{D}_1\mathrm{D}_2(0,0).$$ □

15 Differenzierbare Funktionen

Zur Motivation formulieren wir die Differenzierbarkeit einer Funktion einer Veränderlichen um:

Lemma 15.1. *Es seien $a,b \in \mathbb{R}$ mit $a < b$. Eine Funktion $f\colon (a,b) \to \mathbb{R}$ ist genau dann differenzierbar im Punkt $x_0 \in (a,b)$, wenn ein $l \in \mathbb{R}$, ein $\varepsilon > 0$ und eine Funktion $\rho\colon (x_0-\varepsilon, x_0+\varepsilon) \to \mathbb{R}$ existieren, so dass für $h \in (x_0-\varepsilon, x_0+\varepsilon)$ Folgendes gilt:*

$$f(x_0+h) = f(x_0) + l \cdot h + \rho(h)$$

und

$$\rho(h) = o(h), \quad \text{d.h.} \quad \lim_{h\to 0} \frac{\rho(h)}{h} = 0.$$

„$\rho(h)$ geht schneller gegen Null als h.“

Wie bei der Taylorformel sehen wir, dass in diesem Fall der Zuwachs

$$f(x_0+h) - f(x_0)$$

bis auf einen Fehlerterm $\rho(h)$ gut durch die lineare Funktion $h \mapsto l \cdot h$ approximiert wird. Die Forderung der Existenz einer solchen linearen Approximation ist im mehrdimensionalen Fall der richtige Differenzierbarkeitsbegriff.

Definition 15.2. Es sei $\Omega \subset \mathbb{R}^n$ offen und $a \in \Omega$. Eine Funktion $f\colon \Omega \to \mathbb{R}^m$ heißt *(total) differenzierbar im Punkt* a, falls es eine lineare Abbildung $L\colon \mathbb{R}^n \to \mathbb{R}^m$ und ein $\varepsilon > 0$ gibt, so dass für

$$\rho(h) := f(a+h) - f(a) - L \cdot h, \quad h \in B_\varepsilon(a),$$

die folgende Eigenschaft erfüllt ist:

$$\|\rho(h)\| = o(\|h\|) \quad \text{für } h \to \mathbf{0}, \quad \text{d.h.} \quad \lim_{h \to \mathbf{0}} \frac{1}{\|h\|}\Big(\underbrace{f(a+h) - f(a) - L \cdot h}_{\in \mathbb{R}^m}\Big) = \mathbf{0}. \qquad (*)$$

Dann bezeichnen wir L als Darstellungsmatrix von $(\mathrm{D}f)(a) = \mathrm{D}f(a)$ und nennen es *Ableitung von* f *in* a. Wir schreiben insbesondere $\big(\mathrm{D}f(a)\big)(h) = \mathrm{D}f(a) \cdot h$.

Es ist anzumerken, dass wir hier nur Existenz verlangen und keine Eindeutigkeit!

LINK: TEIL 1 DER 19. VORLESUNG VOM 22.06.2022

Erinnerung. Jede lineare Abbildung $L\colon \mathbb{R}^n \to \mathbb{R}^m$ ist Lipschitz-stetig und besitzt eine Darstellungsmatrix $A \in \mathbb{R}^{m \times n}$ bezüglich der Standardbasen.

Bemerkung 15.3.

(a) Mit $x = a + h$ ist $(*)$ äquivalent zu

$$\|f(x) - f(a) - L \cdot (x - a)\| = o(\|x - a\|) \quad \text{für } x \to a.$$

(b) Mit $f = (f_1, \ldots, f_m)\colon \mathbb{R}^n \to \mathbb{R}^m$ und einer reellen Matrix $C = (c_{jk}) \in \mathbb{R}^{m \times n}$ ist

$$f_j(a+h) - f_j(a) - \underbrace{\sum_{k=1}^{n} c_{jk} h_k}_{=(C \cdot h)_j}, \quad j \in \{1, \ldots, m\},$$

gerade die j-te Komponentenfunktion von

$$f(a+h) - f(a) - C \cdot h.$$

Dabei bezeichnet $(C \cdot h)_j$ die j-te Komponente von $Ch \in \mathbb{R}^m$.

Fazit. Die Funktion $f\colon \mathbb{R}^n \to \mathbb{R}^m$ ist genau dann total differenzierbar in a, wenn f_j für alle $j \in \{1, \ldots, m\}$ total differenzierbar in a ist. Dabei entspricht $\mathrm{D}f_j(a)$ gerade der j-ten Komponentenfunktion von $\mathrm{D}f(a)$.

(c) Es sei $m = 1$, $I \subset \mathbb{R}$ ein offenes Intervall und $f\colon I \to \mathbb{R}$ eine Funktion. f ist genau dann total differenzierbar in $a \in I$, wenn $f'(a)$ existiert. In diesem Fall ist $\mathrm{D}f(a)$ gegeben durch

$$\big(\mathrm{D}f(a)\big)(h) = f'(a) \cdot h \quad \forall, h \in \mathbb{R}.$$

Beispiel 15.4.

(a) Es sei $T\colon \mathbb{R}^n \to \mathbb{R}^m$ linear und $c \in \mathbb{R}^m$. Wir definieren $f\colon \mathbb{R}^n \to \mathbb{R}^m$ durch

$$f(x) := c + T(x),$$

f ist also eine affin-lineare Abbildung. Für $a, h \in \mathbb{R}^n$ gilt dann

$$f(a+h) = c + \underbrace{T(a+h)}_{\overset{\text{lin}}{=} T(a)+T(h)} = f(a) + T(h),$$

also

$$\rho(h) := f(a+h) - f(a) - T(h) = 0.$$

Damit ist f (total) differenzierbar in a mit $\mathrm{D}f(a) = T$.

(b) Es sei $A \in \mathbb{R}^{n\times n}$ eine quadratische Matrix und $f\colon \mathbb{R}^n \to \mathbb{R}$ definiert durch

$$f(x) := \langle x, Ax\rangle, \quad x \in \mathbb{R}^n.$$

Für $a, h \in \mathbb{R}^n$ gilt dann

$$\begin{aligned} f(a+h) &= \langle a+h, A\cdot(a+h)\rangle \\ &= \langle a, A\cdot a\rangle + \overbrace{\langle h, Aa\rangle}^{=\langle a, A^{\intercal} h\rangle} + \langle a, Ah\rangle + \langle h, Ah\rangle \\ &= f(a) + \langle a, (A^{\intercal} + A)h\rangle + \rho(h), \quad \rho(h) := \langle h, Ah\rangle. \end{aligned}$$

Mit Cauchy-Schwarz folgt dabei, dass

$$\|\rho(h)\| \leq \|h\| \cdot \|Ah\|,$$

und insbesondere

$$\frac{\|\rho(h)\|}{\|h\|} \leq \|Ah\| \xrightarrow[h\to\mathbf{0}]{} 0,$$

da A die Darstellungsmatrix einer linearen, und damit Lipschitz-stetigen, Funkion ist. Es gilt also

$$\|\rho(h)\| = o(\|h\|).$$

Damit ist f (total) differenzierbar in a mit

$$\big(\mathrm{D}f(a)\big)(h) = \langle a, (A^{\intercal} + A)h\rangle \quad \forall\, h \in \mathbb{R}^n, \quad \text{also} \quad \mathrm{D}f(a) = \langle a(A^{\intercal} + A)\cdot\rangle.$$

Ist speziell $A = I_n$ die Einheitsmatrix, so ist $f(x) = \langle x, x\rangle = \|x\|^2$ und $\mathrm{D}f(a) = 2\langle a, \cdot\rangle$.

Satz 15.5. *Es seien $\Omega \subset \mathbb{R}^n$ offen und $f = (f_1, \dots, f_m)\colon \Omega \to \mathbb{R}^m$ (total) differenzierbar in $a \in \Omega$. Dann gelten die folgenden Aussagen:*

(a) *f ist stetig in a;*

(b) *Für jedes $v \in \mathbb{R}^n$ gilt*

$$\big(\mathrm{D}f(a)\big)(v) = \lim_{\substack{t \to 0 \\ t \in \mathbb{R}}} \frac{1}{t} \cdot \big(f(a + t \cdot v) - f(a)\big).$$

Insbesondere ist f partiell differenzierbar in a bezüglich x_k, wobei

$$\mathrm{D}_k f(a) = \big(\mathrm{D}f(a)\big)(e_k), \quad k \in \{1, \dots, n\}.$$

Außerdem ist $\mathrm{D}f(a)$ eindeutig bestimmt.

Beweis.

(a) Nach Definition der Differenzierbarkeit gilt $f(a+h) = f(a) + \big(\mathrm{D}f(a)\big)(h) + \rho(h)$ mit

$$\frac{\|\rho(h)\|}{\|h\|} \xrightarrow[h \to \mathbf{0}]{} 0.$$

Insbesondere gilt $\|\rho(h)\| = \|h\| \cdot \frac{\|\rho(h)\|}{\|h\|} \xrightarrow[h \to \mathbf{0}]{} 0$. Da auch $\left\|\big(\mathrm{D}f(a)\big)(h)\right\| \xrightarrow[h \to \mathbf{0}]{} 0$ gilt, folgt

$$f(a+h) \xrightarrow[h \to \mathbf{0}]{} f(a),$$

d.h. f ist stetig in a.

(b) Für $v = \mathbf{0} \in \mathbb{R}^n$ ist die Aussage klar. Es sei also $v \neq \mathbf{0}$. Mit $h = t \cdot v, t \in \mathbb{R}$, folgt aus der Differenzierbarkeit von f, dass

$$\lim_{t \to 0} \left\| \frac{f(a+tv) - f(a)}{t} - \big(\mathrm{D}f(a)\big)(v) \right\| = \lim_{t \to 0} \overbrace{\frac{\left\| f(a+tv) - f(a) - \big(\mathrm{D}f(a)\big)(tv) \right\|}{\|tv\|}}^{\xrightarrow[t \to 0]{} 0} \cdot \|v\|$$

$$= 0.$$

Mit der speziellen Wahl $v = e_k$, $k \in \{1, \dots, n\}$, gilt also nach der Definition der partiellen Differenzierbarkeit:

$$\big(\mathrm{D}f(a)\big)(e_k) = \lim_{t \to 0} \frac{f(a + te_k) - f(a)}{t} \overset{\text{Def.}}{=} \mathrm{D}_k f(a).$$

Damit ist $\mathrm{D}f(a)$ eindeutig durch die Wirkung auf den Basisvektoren e_k festgelegt. $\square$

LINK: TEIL 2 DER 19. VORLESUNG VOM 22.06.2022

Wir haben gesehen, dass $\big(\mathrm{D}_k f(a)\big)(e_k) = \mathrm{D}_k f(a)$. Komponentenweise ausgelesen heißt das dann (vgl. Bemerkung 15.3 (b)):

$$\mathrm{D}_k f_j(a) = \big(\mathrm{D} f_j(a)\big)(e_k), \quad \forall\, k \in \{1,\dots,n\}, d \in \{1,\dots,m\}.$$

Ist f also (total) differenzierbar in a, so wird $\mathrm{D}f(a)$ bezüglich der Standardbasen durch die $(m \times n)$-Matrix

$$\begin{pmatrix} \mathrm{D}_1 f_1(a) & \mathrm{D}_2 f_1(a) & \cdots & \mathrm{D}_n f_1(a) \\ \mathrm{D}_1 f_2(a) & \mathrm{D}_2 f_2(a) & \cdots & \mathrm{D}_n f_2(a) \\ \vdots & \vdots & & \vdots \\ \mathrm{D}_1 f_m(a) & \mathrm{D}_2 f_m(a) & \cdots & \mathrm{D}_n f_m(a) \end{pmatrix}$$

dargestellt. Diese Matrix heißt *Fundamentalmatrix* oder *Jacobi-Matrix von f in a* und wird mit $J_f(a)$ bezeichnet. Es gilt dann $\big(\mathrm{D}f(a)\big)(h) = J_f(a) \cdot h$.

Beachte: Für $m = 1$ ist $J_f(a)$ eine $(1 \times n)$-Matrix (Zeilenvektor) und für $n = 1$ eine $(m \times 1)$-Matrix (Spaltenvektor). Dazu kann man die folgende Merkregel mitnehmen: *Für jede Komponentenfunktion von f je eine Zeile.*

Achtung. Aus der Existenz der partiellen Ableitungen von f kann man nicht auf die (totale) Differenzierbarkeit von f schließen, obwohl man die Jacobi-Matrix bilden kann! Dazu betrachten wir die Funktion f wie in Beispiel 12.3, definiert durch

$$f\colon \mathbb{R}^2 \to \mathbb{R}, \quad f(x,y) = \begin{cases} \frac{xy}{x^2+y^2} & \text{für } (x,y) \neq (0,0), \\ 0 & \text{für } (x,y) = (0,0). \end{cases}$$

Hier existieren die partiellen Ableitungen auf $\mathbb{R}^2$, aber f ist in $(0,0)$ sogar unstetig. Hier sind die partiellen Ableitungen in keiner Umgebung von $(0,0)$ beschränkt.

Aber auch die Beschränktheit der partiellen Ableitungen genügt nicht für (totale) Differenzierbarkeit:

Beispiel 15.6. Wir betrachten die Funktion

$$g\colon \mathbb{R}^2 \to \mathbb{R}, \quad g(x,y) := \begin{cases} \frac{x^2 y}{x^2+y^2} & \text{für } (x,y) \neq (0,0), \\ 0 & \text{für } (x,y) = (0,0). \end{cases}$$

Hier gilt $\mathrm{D}_1 g(0,0) = 0 = \mathrm{D}_1 g(0,0)$ und

$$\mathrm{D}_1 g(x,y) = \frac{2xy^3}{(x^2+y^2)^2} \quad \text{und} \quad \mathrm{D}_2 g(x,y) = \frac{x^2(x^2-y^2)}{(x^2+y^2)^2}, \quad (x,y) \neq (0,0).$$

Dabei ist

$$|\mathrm{D}_1 g(x,y)| = \frac{2|x||y|}{x^2+y^2}, \quad \frac{y^2}{x^2+y^2} \le 1 \quad \text{und} \quad |\mathrm{D}_2 g(x,y)| \le \frac{x^2}{x^2+y^2} \cdot \frac{x^2+y^2}{x^2+y^2} \le 1.$$

Insbesondere ist g stetig. Ferner erhalten wir $J_g(0,0) = (0,0)$. Wäre g in $(0,0)$ differenzierbar, so müsste also $\mathrm{D}g(0,0) \equiv 0 \colon \mathbb{R}^2 \to \mathbb{R}$ die Nullabbildung sein, also

$$\frac{g(x,y) - \overbrace{g(0,0)}^{=0} - \overbrace{\big(\mathrm{D}g(0,0)\big)(x,y)}^{=0}}{\|(x,y)\|} = \frac{g(x,y)}{\|(x,y)\|} = \frac{x^2 y}{(x^2+y^2)^{\frac{3}{2}}} \xrightarrow[(x,y)\to(0,0)]{} 0.$$

Für $y = x > 0$ haben wir aber

$$\frac{g(x,x)}{\|(x,x)\|} = \frac{1}{2^{\frac{3}{2}}} \xrightarrow[x\to 0]{} \frac{1}{2^{\frac{3}{2}}} \neq 0,$$

was ein Widerspruch ist. Also ist f nicht differenzierbar in $(0,0)$. Dennoch existiert für alle $v = (v_1, v_2) \neq (0,0)$ der Grenzwert

$$\lim_{t\to 0} \frac{g(tv) - g(0,0)}{t} = \lim_{t\to 0} \frac{t^3 v_1^2 v_2}{t^3(v_1^2 + v_2^2)} = \frac{v_1^2 v_2}{v_1^2 + v_2^2}.$$

Definition 15.7. Es sei $\Omega \subset \mathbb{R}^n$ offen, $f \colon \Omega \to \mathbb{R}^m$ eine Abbildung, $a \in \Omega$ und $v \in \mathbb{R}^n$. Existiert der Grenzwert

$$\mathrm{D}_v f(a) : 0 \lim_{\substack{t\to 0\\ t\in\mathbb{R}}} \frac{f(a+tv) - f(a)}{t} \in \mathbb{R}^m,$$

so heißt $\mathrm{D}_v f(a) \in \mathbb{R}^m$ die *Richtungsableitung von f in a in Richtung v*. Man schreibt auch

$$\mathrm{D}_v f(a) = \frac{\partial f}{\partial v}(a) = \partial_v f(a).$$

Speziell für $v = e_k$, $k \in \{1, \dots, n\}$, erhalten wir wieder die k-te partielle Ableitung von f in a: $\mathrm{D}_{e_k} f(a) = \mathrm{D}_k f(a)$.

Satz 15.5 besagt insbesondere: Falls f differenzierbar ist in $a \in \Omega$, so existiert die Richtungsableitung von f in *jede* Richtung $v \in \mathbb{R}^n$, und es gilt

$$\mathrm{D}_v f(a) = \big(\mathrm{D}f(a)\big)(v) = J_f(a) \cdot v.$$

In Beispiel 15.6 haben wir gesehen, dass aus der Existenz aller Richtungsableitungen noch keine Differenzierbarkeit von f folgt. In dem Beispiel ist auch die Eigenschaft $\mathrm{D}_v f(0,0) = J_f(0,0) \cdot v$ nicht erfüllt: Es gilt nämlich

$$J_f(0,0) = (0,0), \quad \text{aber} \quad \mathrm{D}_v f(0,0) = \frac{v_1^2 v_2}{v_1^2 + v_2^2} \neq 0 \quad \text{für } v_1 \neq 0, v_2 \neq 0.$$

LINK: TEIL 1 DER 20. VORLESUNG VOM 29.06.2022

Aber selbst die Gültigkeit von $\mathrm{D}_v f(a) = J_f(a) \cdot v$ für alle $v \in \mathbb{R}^n$ garantiert noch keine (totale) Differenzierbarkeit:

Beispiel 15.8. Gegeben sei die Funktion

$$g\colon \mathbb{R}^2 \to \mathbb{R}, \quad (x,y) \mapsto \begin{cases} \frac{xy^3}{x^2+y^6} & \text{für } (x,y) \neq (0,0), \\ 0 & \text{sonst.} \end{cases}$$

Für $(x,y) \neq (0,0)$ gilt dann

$$\mathrm{D}_1 g(x,y) = y^3 \cdot \frac{y^6 - x^2}{(y^6+x^2)^2} \quad \text{und} \quad \mathrm{D}_2 g(x,y) = 3xy^2 \cdot \frac{x^2-y^6}{(x^2+y^6)^2}.$$

Außerdem gilt $g(x,0) = 0 = g(0,y)$ für alle $x, y \in \mathbb{R} \setminus \{0\}$ und es folgt

$$\mathrm{D}_1 g(0,0) = 0 = \mathrm{D}_2 g(0,0).$$

Also ist g partiell differenzierbar auf $\mathbb{R}^2$ und es gilt

$$J_g(0,0) = \begin{pmatrix} \mathrm{D}_1 g(0,0) & \mathrm{D}_2 g(0,0) \end{pmatrix} = \begin{pmatrix} 0 & 0 \end{pmatrix}.$$

Bei der Definition von J_g gehen nur die partiellen Ableitungen ein.

Wir berechnen nun nur die Richtungsableitung in $(0,0)$. Für $v = (v_1, v_2) \in \mathbb{R}^2$ gilt dann

$$\begin{aligned} \partial_v g(0,0) &= \lim_{t \to 0} \frac{1}{t}\Big(g(tv) - g(0,0)\Big) \\ &= \lim_{t \to 0} \frac{1}{t} \cdot \frac{t^4 \cdot v_1 v_2^3}{t^2 \cdot v_1 + t^6 \cdot v_2} && (15.1) \\ &= \lim_{t \to 0} \frac{t \cdot v_1 v_2^3}{v_1} = 0 && (15.2) \\ &= J_g(0,0) \cdot v. && (\times) \end{aligned}$$

Aber für $y > 0$ ist $g(y^3, y) = \frac{y^3 y^3}{y^6 y^6} = \frac{1}{2}$, insbesondere

$$\lim_{y \to 0} g(y^3, y) = \frac{1}{2}.$$

Also ist g nicht stetig in $(0,0)$ und nach Satz 15.5 auch nicht differenzierbar in $(0,0)$. Es reicht nicht, sich den Punkt $(0,0)$ nur entlang von den Koordinatenachsen anzunähern, auch nicht entlang von beliebigen Geraden. Man muss auch Kurven betrachten.

Übung: Aber auch Stetigkeit in $(0,0)$ zusammen mit $(\times)$ garantiert nicht (totale) Differenzierbarkeit.

Definition 15.9. Es sei $\Omega \subset \mathbb{R}^n$ offen, $a \in \Omega$ und $f\colon \Omega \to \mathbb{R}$ partiell differenzierbar in a. Dann wird durch

$$\operatorname{grad} f(a) := \big(J_f(a)\big)^{\mathsf{T}} = \begin{pmatrix} \mathrm{D}_1 f(a) \\ \vdots \\ \mathrm{D}_n f(a) \end{pmatrix}$$

der *Gradient von f in a* gekennzeichnet. Ist f sogar auf ganz Ω partiell differenzierbar, so definiert der Gradient eine Abbildung

$$\operatorname{grad}\colon \Omega \to \mathbb{R}^n,$$

die auch als *Gradient(enfeld)* bezeichnet wird. Oft schreibt man auch

$$\nabla f = \operatorname{grad} f$$

mit dem Nabla-Symbol „∇“.

Ist f in $a \in \Omega$ differenzierbar, so ist $w = \operatorname{grad} f(a)$ der eindeutige Vektor in $\mathbb{R}^n$, der

$$\langle w, v\rangle = J_f(a) \cdot v = \big(\mathrm{D} f(a)\big)(v) = \partial_v f(a)$$

für alle $v \in \mathbb{R}^n$ erfüllt. Die Cauchy-Schwarz-Ungleichung impliziert außerdem

$$-\|\operatorname{grad} f(a)\| \cdot \|v\| \leq \langle \operatorname{grad} f(a), v\rangle \leq \|\operatorname{grad} f(a)\| \cdot \|v\|. \tag{$\times\times$}$$

Eine Gleichheit tritt nur auf, falls $\operatorname{grad} f(a)$ und v linear abhängig sind. Im Fall $J_f(a) \neq \mathbf{0} \in \mathbb{R}^{1\times n}$ gibt es dann genau einen normierten Vektor $v \in \mathbb{R}^n$, für den

$$\partial_v f(a) = \langle \operatorname{grad} f(a), v\rangle = \|\operatorname{grad} f(a)\| \cdot \underbrace{\|v\|}_{=1} = \|\operatorname{grad} f(a)\|$$

gilt, d.h. für den die Richtungsableitung den maximal möglichen Wert in ($\times\times$) annimmt. Dies ist der Vektor

$$v = v_{\max} := \frac{1}{\|\operatorname{grad} f(a)\|} \cdot \operatorname{grad} f(a).$$

Der Ausdruck ist wohldefiniert, da aufgrund $J_f(a) \neq 0$ die obige Division erlaubt ist.

Der Gradient von f zeigt in die Richtung des stärksten Anstiegs und seine Norm beschreibt die Stärke des Anstiegs.

Nach all den Beispielen, wo (totale) Differenzierbarkeit nicht sichergestellt werden konnte, gibt es nun einen Hoffnungsschimmer: Der folgende Satz beschreibt nämlich, in welchem Fall (totale) Differenzierbarkeit garantiert ist.

Satz 15.10. *Es sei $\Omega \subset \mathbb{R}^n$ offen, $f\colon \Omega \to \mathbb{R}^m$ partiell differenzierbar auf Ω und $a \in \Omega$. Die partielle Ableitung $\mathrm{D}_l f\colon \Omega \to \mathbb{R}^m$ sei außerdem stetig in a für alle $l \in \{1, \ldots, n\}$. Dann ist f auch (total) differenzierbar in a.*

Beweis. Für den Beweis gehen wir ähnlich vor wie in Satz 12.4. Wie dort reicht es hier aus, den Fall $m = 1$ zu betrachten. Zu $h \in \mathbb{R}^n$ und $l \in \{1, \dots, n\}$ setze nun

$$z^{(l)} = a + \sum_{j=1}^{l} h_j e_j, \quad \text{also} \quad z^{(0)} = a \quad \text{und} \quad z^{(n)} = a + h.$$

Die Menge $\left[z^{(l-1)}, z^{(l)}\right]$ ist ein Intervall, das parallel zur l-ten Koordinatenachse liegt.

LINK: Teil 2 der 20. Vorlesung vom 29.06.2022

Indem wir jeweils $(n-1)$ Koordinaten eingefroren und auf die einzige variable Koordinate den Mittelwertsatz 12.4 aus der Analysis I (Lehramt) anwenden, erhalten wir die Existenz von $\xi^{(1)} \in \left[z^{(0)}, z^{(1)}\right], \dots, \xi^{(n)} \in \left[z^{(l-1)}, z^{(l)}\right]$ mit

$$f(a+h) - f(a) = \sum_{l=1}^{n} \left(f(z^{(l)}) - f(z^{(l-1)})\right) = \sum_{l=1}^{n} \mathrm{D}_l f(\xi^{(l)}) \cdot h_l.$$

Wir setzen

$$\rho(h) := \sum_{l=1}^{n} \left(\mathrm{D}_l f(\xi^{(l)}) - \mathrm{D}_l f(a)\right) \cdot h_l$$

und erhalten

$$f(a+h) - f(a) = \sum_{l=1}^{n} \mathrm{D}_l f(a) \cdot h_l + \rho(h) = \left(\mathrm{D} f(a)\right)(h) + \rho(h).$$

Für $h \to \mathbf{0}$ schrumpfen alle Intervalle $\left[z^{(l-1)}, z^{(l)}\right]$ zu einem Punkt zusammen, „nämlich a“. Insbesondere gilt

$$\lim_{h \to \mathbf{0}} \xi^{(l)} = a_l, \quad l \in \{1, \dots, n\}.$$

Da die partielle Ableitung $\mathrm{D}_l f$ für alle $l \in \{1, \dots, n\}$ stetig in a ist, folgt

$$\lim_{h \to \mathbf{0}} \left(\mathrm{D}_l f(\xi^{(l)}) - \mathrm{D}_l f(a)\right) = 0 \quad \forall\, l \in \{1, \dots, n\},$$

und für ρ folgt

$$\frac{|\rho(h)|}{\|h\|} \leq \sum_{l=1}^{n} \left|\mathrm{D}_l f(\xi^{(l)}) - \mathrm{D}_l f(a)\right| \underbrace{\frac{h_l}{\|h\|}}_{\leq 1} \xrightarrow[h \to \mathbf{0}]{} 0.$$

Also ist f differenzierbar in a. $\square$

Umgekehrt kann man aus der (totalen) Differenzierbarkeit jedoch nicht die Stetigkeit der partiellen Ableitung folgern, wie das folgende Beispiel zeigt.

Beispiel 15.11. Wir betrachten die Funktion

$$f \colon \mathbb{R}^2 \to \mathbb{R}, \quad (x, y) \mapsto \begin{cases} \frac{(x^2+y^2)^{\frac{3}{2}}}{\sin(x^2+y^2)} & \text{für } (x,y) \neq (0,0), \\ 0 & \text{sonst.} \end{cases}$$

Für $(x,y) \neq (0,0)$ und $r^2 := x^2 + y^2$ gilt dann

$$\frac{|f(x,y)|}{\|(x,y)\|} = \frac{1}{r} \cdot r^3 \cdot \underbrace{|\sin(r^{-2})|}_{\leq 1} \leq r^2 \underset{(x,y)\to\mathbf{0}}{\longrightarrow} 0.$$

Wir behauptet nun, dass $\mathrm{D}f(0,0) = \mathbf{0}$, denn damit gilt

$$\rho(h_1,h_2) := f(h_1,h_2) \underbrace{-f(0,0) - 0 \cdot h}_{=0} = f(h_1,h_2),$$

also

$$\frac{|\rho(h)|}{\|h\|} \leq \|h\|^2 \underset{h\to\mathbf{0}}{\longrightarrow} 0.$$

Aber für $x > 0$ gilt

$$\mathrm{D}_1 f(x,0) = \overbrace{3x^2 \cdot \sin(x^{-2})}^{|\ldots|\leq 3x^2} + x^2 \cos(x^{-2}) \cdot (-2x^{-3}) = o(x) - 2\cos\left(\frac{1}{x^2}\right).$$

$\mathrm{D}_1(x,0)$ oszilliert dann für $x \to 0$ und hat dementsprechend keinen Limes, ist also unstetig in $(0,0)$. Aus Symmetriegründen (durch Vertauschung von x und y) folgt dann, dass auch $\mathrm{D}_2 f$ unstetig ist in $(0,0)$.

Analog wie in Analysis I (Lehramt) gelten:

Satz 15.12 (Summenregel)**.** *Es sei $\Omega \subset \mathbb{R}^n$ offen, $a \in \Omega$, $\lambda \in \mathbb{R}$ und $f, g \colon \Omega \to \mathbb{R}^m$ differenzierbar in a. Dann ist auch die Abbildung*

$$(\lambda \cdot f + g) \colon \Omega \to \mathbb{R}^m, \quad x \mapsto (\lambda f + g)(x) = \lambda \cdot f(x) + g(x)$$

differenzierbar in a und es gilt

$$\Big(\mathrm{D}(\lambda \cdot f + g)\Big)(a) = \lambda \cdot \mathrm{D}f(a) + \mathrm{D}g(a).$$

Für die entsprechenden Jacobi-Matrizen gilt also

$$J_{\lambda f + g}(a) = \lambda \cdot J_f(a) + J_g(a).$$

Beweis. Übung!

Satz 15.13 (Kettenregel)**.** *Es seien $\Omega \subset \mathbb{R}^n$, $\Sigma \subset \mathbb{R}^m$ offen, $a \in \Omega$, $b = f(a) \in \Sigma$, $f \colon \Omega \to \Sigma$ differenzierbar in a und $g \colon \Sigma \to \mathbb{R}^l$ differenzierbar in b. Dann ist auch die Verkettung*

$$g \circ f \colon \Omega \to \mathbb{R}^l, \quad (g \circ f)(a) = g\big(f(a)\big)$$

differenzierbar in a und es gilt

$$\Big(\mathrm{D}(g \circ f)\Big)(a) = \mathrm{D}g(b) \circ \mathrm{D}f(a) = \mathrm{D}g\big(f(a)\big) \circ \mathrm{D}f(a).$$

Für die Funktionalmatrizen gilt

$$J_{g\circ f}(a) = J_g(b) \cdot J_f(a) = J_g\big(f(a)\big) \cdot J_f(a)$$

Beweis. Nach Voraussetzung existieren lineare Abbildungen $L = \mathrm{D}f(a)$ und $Q = \mathrm{D}g(b)$ sowie die Funktionen $\rho\colon \mathbb{R}^n \to \mathbb{R}^m$, $r\colon \mathbb{R}^m \to \mathbb{R}^l$ derart, dass

$$\|\rho(h)\| = o(\|h\|) \quad \text{und} \quad \|r(k)\| = o(\|k\|)$$

Es existiert also eine Abbildung $\varepsilon\colon \mathbb{R}^m \to \mathbb{R}^l$ mit $\lim_{k\to\mathbf{0}} \varepsilon(k) = 0$ und $r(k) = \|k\|\varepsilon(k)$.

und

$$f(a+h) = f(a) + L(h) + \rho(h) \quad \text{bzw.} \quad g(b+k) = g(b) + Q(k) + r(k).$$

LINK: TEIL 1 DER 21. VORLESUNG VOM 30.06.2022

Wähle speziell $k := f(a+h) - b$, d.h. es gilt $f(a+h) = b + k$. Man erhält dann

$$\begin{aligned}\Big(g\circ f\Big)(a+h) = g(b+k) &= g(b) + Q(k) + r(k)\\ &= g(b) + Q\Big(\underbrace{f(a+h)-f(a)}_{=L(h)+\rho(h)}\Big) + r(k)\\ &= \Big(g\circ f\Big)(a) + QL(h) + Q\Big(\rho(h)\Big) + r(k).\end{aligned}$$

Behauptung. Es gilt $QL = \mathrm{D}(g\circ f)$, d.h. für

$$R(h) := Q\Big(\rho(h)\Big) + r\Big(f(a+h) - b\Big)$$

muss $R(h) = o(h)$ gelten.

Beweis der Behauptung. Da Q eine lineare Abbildung ist, ist Q nach Beispiel 11.25 (a) Lipschitz-stetig mit Lipschitz-Konstante $\|Q\|$, es gilt also

$$\Big\|Q\Big(\rho(h)\Big)\Big\| \le \|Q\| \cdot \|\rho(h)\|.$$

Damit ist

$$\frac{\Big\|Q\Big(\rho(h)\Big)\Big\|}{\|h\|} \le \|Q\| \cdot \frac{\|\rho(h)\|}{\|h\|} \xrightarrow[h\to 0]{} 0$$

und

$$\frac{r\Big(f(a+h)-b\Big)}{\|h\|} = \underbrace{\frac{\Big\|\overbrace{f(a+h)-f(a)}^{=L(h)+\rho(h)}\Big\|}{\|h\|}}_{\substack{\le \frac{\|L\|}{\|h\|} + \frac{\|\rho(h)\|}{\|h\|}\\ \le \|L\| + \text{const.}}} \cdot \overbrace{\varepsilon\Big(\underbrace{f(a+h)-f(a)}_{\overset{f\text{ stetig}}{\underset{h\to 0}{\longrightarrow}} 0}\Big)}^{\underset{h\to 0}{\longrightarrow} 0},$$

da

- L Lipschitz-stetig mit Lipschitz-Konstante $\|L\|$ ist und
- $\|\rho(h)\|$ als Null„folge“ für kleines h beschränkt ist.

Es gilt also

$$\frac{r\Big(f(a+h)-b\Big)}{\|h\|} \leq \underbrace{\left(\|L\| + \frac{\text{const.}}{\|h\|}\right)}_{\text{beschränkt}} \cdot \underbrace{\varepsilon(h)}_{\xrightarrow[h\to 0]{} 0} \xrightarrow[h\to 0]{} 0$$

und damit $R(h) = o(h)$.

Damit ist der Satz auch bewiesen. $\square$

Beispiel 15.14. Es sei $\Omega \subset \mathbb{R}^n$ und $x \in \Omega$. Setze außerdem $h := g \circ f\colon \Omega \to \mathbb{R}^l = \mathbb{R}$, wobei $f\colon \Omega \to \mathbb{R}^m$ und $g\colon \mathbb{R}^m \to \mathbb{R}^l$. Dann gilt

$$\begin{aligned}
J_h(x) &= \begin{pmatrix} \frac{\partial h}{\partial x_1}(x) & \cdots & \frac{\partial h}{\partial x_n}(x) \end{pmatrix} \in \mathbb{R}^{1\times n} \\
&= \begin{pmatrix} \frac{\partial g}{\partial y_1}\Big(f(x)\Big) & \cdots & \frac{\partial g}{\partial y_m}\Big(f(x)\Big) \end{pmatrix} \cdot \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(x) & \cdots & \frac{\partial f_1}{\partial x_n}(x) \\ \vdots & & \vdots \\ \frac{\partial f_m}{\partial x_1}(x) & \cdots & \frac{\partial f_m}{\partial x_n}(x) \end{pmatrix} \\
&= J_g\Big(f(x)\Big) \cdot J_f(x).
\end{aligned}$$

(a) Speziell für $n = 1$, also $\Omega \subset \mathbb{R}$, und $h\colon \Omega \to \mathbb{R}$ gilt

$$h'(x) = \sum_{k=1}^{m} \frac{\partial g}{\partial y_k}\Big(f(x)\Big) \cdot \frac{\partial f_k}{\partial x}(x).$$

(b) Wir betrachten nun speziell für $n = 2$ die Anwendung auf Polarkoordinaten, es sei also $\Omega = (0,\infty) \times (0, 2\pi)$ und f gegeben durch

$$f\colon \Omega \to \mathbb{R}^2, \quad \begin{pmatrix} r \\ \varphi \end{pmatrix} \mapsto \begin{pmatrix} r\cdot\cos(\varphi) \\ r\cdot\sin(\varphi) \end{pmatrix} =: \begin{pmatrix} x \\ y \end{pmatrix}.$$

Gegeben sei nun beispielsweise die Funktion

$$g\colon \mathbb{R}^2 \to \mathbb{R}, \quad (x,y) \mapsto x \cdot y,$$

für die Verkettung $h = g \circ f\colon \Omega \to \mathbb{R}$ gelte also

$$h(r,\varphi) = r^2 \cdot \cos(\varphi) \cdot \sin(\varphi).$$

Dann gilt

$$\frac{\partial h}{\partial r}(r,\varphi) = \frac{\partial g}{\partial x}(x,y) \cdot \frac{\partial f_1}{\partial r}(r,\varphi) + \frac{\partial g}{\partial y}(x,y) \cdot \frac{\partial f_2}{\partial r}(r,\varphi)$$

$$
\begin{aligned}
&= y \cdot \frac{\partial}{\partial r}\Big(r \cdot \cos(\varphi)\Big) + x \cdot \frac{\partial}{\partial r}\Big(r \cdot \sin(\varphi)\Big) \\
&= \overbrace{y}^{=r\sin(\varphi)} \cdot \cos(\varphi) + \overbrace{x}^{=r\cos(\varphi)} \cdot \sin(\varphi) \\
&= 2r \cdot \sin(\varphi) \cdot \cos(\varphi) \\
&\overset{\text{AT}}{=} r \cdot \sin(2\varphi)
\end{aligned}
$$

und

$$
\begin{aligned}
\frac{\partial h}{\partial \varphi}(r,\varphi) &= \frac{\partial g}{\partial x}(x,y) \cdot \frac{\partial f_1}{\partial \varphi}(r,\varphi) + \frac{\partial g}{\partial y}(x,y) \cdot \frac{\partial f_2}{\partial \varphi}(r,\varphi) \\
&= -y \cdot r \cdot \sin(\varphi) + x \cdot r \cdot \cos(\varphi) \\
&= r \cdot \Big(-r \cdot \sin^2(\varphi) + r \cdot \cos^2(\varphi)\Big) \\
&= r^2 \cdot \cos(2\varphi).
\end{aligned}
$$

(c) Gegeben sei die Funktion

$$
f\colon \mathbb{R}^n \to [0,\infty), \quad (x_1,\dots,x_n) \mapsto \sqrt{x_1^2 + \ldots + x_n^2} = \|x\|_2.
$$

Betrachten wir nun eine Funktion $g\colon [0,\infty) \to \mathbb{R}$, die differenzierbar ist auf $(0,\infty)$. Dann ist die Verkettung $h = g \circ f\colon \mathbb{R}^n \to \mathbb{R}$ rotationsinvariant (vgl. Beispiel 12.2 (c)). Für $x \neq \mathbf{0} \in \mathbb{R}^n$ gilt dann

$$
J_h(x) = \Big(\tfrac{g'(f(x))}{f(x)} \cdot x_1 \quad \cdots \quad \tfrac{g'(f(x))}{f(x)} \cdot x_n \Big) = \frac{g'\big(f(x)\big)}{f(x)} \cdot x^\mathsf{T} = \frac{g'\big(\|x\|_2\big)}{\|x\|_2} \cdot x^\mathsf{T}.
$$

Für $v = \frac{x}{\|x\|_2}$ folgt

$$
\partial_v h(x) = \langle \operatorname{grad} h(x), v \rangle = g'\big(\|x\|_2\big).
$$

Für $g'\big(\|x\|_2\big) > 0$ ist v die Richtung des maximalen Anstiegs von h. Für $g'\big(\|x\|_2\big) < 0$ ist dann analog $w = -v$ die Richtung des maximalen Anstiegs von h. Betrachten wir beispielsweise die differenzierbare Funktion $g(r) := \mathrm{e}^{-r^2}$, so gilt $h(x) = \mathrm{e}^{-\|x\|_2^2}$.

Lemma 15.15. *Es sei $I \subset \mathbb{R}$ ein offenes Intervall, $f \in C^1\big(I \times (a,b)\big)$ und $\varphi, \psi \in C^1(I)$ mit*

$$
a < \varphi(x) < \psi(x) < b \quad \forall\, x \in I.
$$

Es sei außerdem

$$F\colon I \to \mathbb{R}, \quad x \mapsto \int_{\varphi(x)}^{\psi(x)} f(x,y)\,\mathrm{d}y.$$

Dann ist F differenzierbar auf I mit

$$F'(x) = f\big(x,\psi(x)\big)\cdot\psi'(x) - f\big(x,\varphi(x)\big)\cdot\varphi'(x) + \left(\int_{\varphi(x)}^{\psi(x)} \mathrm{D}_1 f(x,y)\,\mathrm{d}y\right), \qquad x \in I.$$

Beweis. Wir definieren zunächst die Funktionen $H\colon (a,b)\times(a,b)\times I \to \mathbb{R}$ und $g\colon I \to \mathbb{R}^3$ durch

$$H(u,v,w) := \int_u^v f(w,y)\,\mathrm{d}y \quad \text{und} \quad g(x) = \begin{pmatrix} \varphi(x) \\ \psi(x) \\ x \end{pmatrix}.$$

Damit erhält man

$$\big(H\circ g\big)(x) = \int_{\varphi(x)}^{\psi(x)} f(x,y)\,\mathrm{d}y = F(x),$$

nach Beispiel 15.14 (a) gilt also

$$F'(x) = \Big(\mathrm{D}_1 H\big(g(x)\big) \quad \mathrm{D}_2 H\big(g(x)\big) \quad \mathrm{D}_3 H\big(g(x)\big)\Big)\cdot\begin{pmatrix} \varphi'(x) \\ \psi'(x) \\ 1 \end{pmatrix}.$$

Der Hauptsatz der Differential- und Integralrechnung liefert dann

$$\mathrm{D}_1 H(u,v,w) = -f(w,u) \quad \text{und} \quad \mathrm{D}_2 H(u,v,w) = f(w,v).$$

Zur Bestimmung von $\mathrm{D}_3 H$ verwenden wir den Satz 13.5 über die Ableitung von parameterabhängigen Integralen und erhalten

$$\mathrm{D}_3 H(u,v,w) = \int_u^v \frac{\partial}{\partial w} f(w,y)\,\mathrm{d}y.$$

Setzt man nun $u = \varphi(x)$, $v = \psi(x)$ und $w = x$ ein, so folgt direkt die Behauptung. □

LINK: TEIL 2 DER 21. VORLESUNG VOM 30.06.2022

16 Der Mittelwertsatz und die Taylor-Formel

Satz 16.1 (Mittelwertsatz)**.** *Es sei $\Omega \subset \mathbb{R}^n$ offen, $f\colon \Omega \to \mathbb{R}$ differenzierbar auf Ω und $a, b \in \Omega$ Punkte derart, dass*

$$[a,b] := \Big\{t\cdot a + (1-t)\cdot b \ : \ t \in [0,1] \subset \mathbb{R} \subset \Omega.$$

Dann existiert ein Punkt

$$\xi \in (a,b) := \Big\{t\cdot a + (1-t)\cdot b \ : \ t \in (0,1) \subset \mathbb{R}\Big\}$$

mit

$$f(b) - f(a) = J_f(\xi)\cdot\overbrace{(b-a)}^{\in\mathbb{R}^n} = \sum_{k=1}^n \mathrm{D}_k f(\xi)\cdot(b_k - a_k).$$

Beweis. Definiere zunächst $\gamma\colon [0,1] \to \Omega$ und $h\colon [0,1] \to \mathbb{R}$ durch

$$\gamma(t) := a + t \cdot (b-a) \quad \text{und} \quad h := f \circ \gamma.$$

Da γ nach Definition stetig auf $[0,1]$ und f nach Voraussetzung differenzierbar, also auch stetig auf Ω ist, ist h nach Satz 11.23 (b) stetig auf $[0,1]$. Da γ auf $(0,1)$ differenziebar ist, ist insbesondere h nach der Kettenregel (s. Satz 15.13) differenzierbar auf $(0,1)$. Nach dem Mittelwertsatz der Differentialrechnung in $\mathbb{R}$ (vgl. Satz 12.4, Analysis I (Lehramt)) existiert dann ein $\tau \in (0,1)$ mit

$$1 \cdot h'(t) = h(1) - h(0).$$

Anwendung der Kettenregel wie in Beispiel 15.14 (a) liefert dann

$$h'(\tau) = \sum_{k=1}^{n} \frac{\partial f}{\partial x_k}\big(\gamma(\tau)\big) \cdot \underbrace{\frac{\partial \gamma_k}{\partial t}(\tau)}_{=(b-a)_k} = J_f(a + \tau \cdot (b-a)) \cdot (b-a).$$

Setze nun $\xi := a + \tau \cdot (b-a)$ und es folgt direkt

$$\mathrm{D}f(\xi) \cdot (b-a) = h(1) - h(0) = f(b) - f(a). \qquad \square$$

Der Mittelwertsatz gilt im Allgemeinen nicht für vektorwertige (insbesondere nicht mal $\mathbb{C}$-wertige) Funktionen. Betrachten wir dazu die Funktion

$$f\colon [0,2\pi] \to \mathbb{R}^2, \quad \varphi \mapsto \begin{pmatrix} \cos(\varphi) \\ \sin(\varphi) \end{pmatrix} \quad \text{bzw.} \quad f\colon [0,2\pi] \to \mathbb{C}, \quad \varphi \mapsto \mathrm{e}^{\mathrm{i}\varphi}.$$

Würde der Mittelwertsatz gelten, gäbe es ein $\xi \in (0, 2\pi)$, so dass

$$\begin{pmatrix} 0 \\ 0 \end{pmatrix} = f(2\pi) - f(0) = 2\pi \cdot f'(\xi) = 2\pi \cdot \begin{pmatrix} -\sin(\varphi) \\ \cos(\varphi) \end{pmatrix}.$$

Dies impliziert wiederum $\sin(\xi) = 0 = \cos(\xi)$, was jedoch für kein $\xi \in \mathbb{R}$ eintritt.

Korollar 16.2. *Es sei $\Omega \subset \mathbb{R}^n$ ein Gebiet – Ω ist also offen und wegzusammenhängend. Ist $f\colon \Omega \to \mathbb{R}$ dann differenzierbar auf Ω mit*

$$\mathrm{D}f(x) = 0 \quad \forall\, x \in \Omega,$$

so ist f konstant auf Ω.

Der Beweis der Aussage in der „Einführung in die Analysis II“ von Winfried Kaballo (1997) ist einfacher und allgemeiner. Es empfiehlt sich also, einen Blick in das Buch zu werfen.

Beweis. Für beliebige $x, y \in \Omega$ müssen wir die Identität

$$f(x) = f(y)$$

zeigen.

(I) Wir betrachten zunächst den Spezialfall $\Omega = B$, wobei B einen Ball bezeichnet – B ist dann insbesondere konvex. Aufgrund der Konvexität folgt direkt $[x, y] \subset B$ und nach dem Mittelwertsatz 16.1 existiert dann ein $\xi \in [x, y]$ mit

$$f(x) - f(y) = \underbrace{J_f(\xi}_{=0} \cdot (x - y) = 0.$$

Die Aussage gilt hier also.

(II) Es sei nun $\Omega \subset \mathbb{R}^n$ ein Gebiet, es existiere also ein Weg $\gamma\colon [0, 1] \to \Omega$ mit $\gamma(0) = x$ und $\gamma(1) = y$. Setze nun $\Gamma := \gamma\big([0, 1]\big) \subset \Omega$.

Behauptung. Es gilt

$$D := \operatorname{dist}(\Gamma, \Omega^c) := \inf_{\substack{b\in\Gamma,\\ a\in\Omega^c}} |a - b| > 0.$$

Beweis der Behauptung. Es existiert ein Ball $B \subset \mathbb{R}^n$ mit $x, y \in B$ und mit einem γ derart, dass $\Gamma \subset B \cap \Omega$ (Ohne Beweis), so dass dann

$$D = \inf_{\substack{b\in\Gamma\\ a\in A}} = \inf |b - a|, \quad A := \Omega^c \cap \overline{B},$$

erfüllt ist. Nun ist aber die Funktion

$$f\colon A \times \Gamma \to \mathbb{R}, \quad (a, b) \mapsto |b - a|,$$

stetig, Γ beschränkt und abgeschlossen in $\mathbb{R}^n$ und A beschränkt und abgeschlossen in $\mathbb{R}^n$. Damit ist $A \times \Gamma$ beschränkt und abgeschlossen in $\mathbb{R}^{2n}$, also kompakt. Also existieren $a_0 \in A$ und $b_0 \in \Gamma$ mit

$$D = |b_0 - a_0| > 0, \quad \text{da } A, \Gamma \text{ disjunkt.}$$

Da γ als weg insbesondere stetig und $[0, 1]$ kompakt ist, ist γ sogar gleichmäßig stetig. Also existiert ein $m \in \mathbb{N}$ mit

$$t, s \in [0, 1], |t - s| \leq \frac{1}{m} \quad \Rightarrow \quad \|\gamma(t) - \gamma(s)\| < D.$$

Für $j \in \{1, \ldots, m\}$ setzen wir nun

$$t_j := \frac{j}{m}, \quad x_j := \gamma(t_j) \quad \text{und} \quad B^j := B_D(x_j).$$

Dann erhält man $x_j \in \Gamma \quad \Rightarrow \quad B^j \subset \Omega$, die Distanz zum Rand wird also eingehalten. Wegen $|t_j - t_{j-1}| \leq \frac{1}{m}$ für alle j, folgt direkt

$$\underbrace{\|x_j - x_{j-1}\|}_{\geq 0} < D \quad \Rightarrow \quad D > 0.$$

Damit ist $B^j \cap B^{j-1} \neq \varnothing$, es existiert also ein Punkt $y_j \in B^j \cap B^{j-1}$. Da B^j für alle $j \in \{1, \dots, m\}$ ein Ball ist, ist f nach Teil (I) konstant: Es gilt

$$f(y_{j+1}) \overset{\text{auf } B_j}{=} f(y_j) \overset{\text{auf } B_{j-1}}{=} f(y_{j-1}). \qquad \square$$

Analog wie für Funktionen mit einer Variablen gilt die Taylorformel auch für Funktionen $f \in C^{k+1}(\mathbb{R}^n, \mathbb{R})$, $k \in \mathbb{N}_0$. Im Folgenden betrachten wir jedoch nur die zweite Ordnung, also den Fall $k = 1$.

Definition 16.3 (Hesse-Matrix)**.** Es sei $\Omega \subset \mathbb{R}^n$ offen, $f \in C^2(\mathbb{R}^n)$ und $x \in \Omega$. Dann wird durch

$$Hf(x) := \begin{pmatrix} \mathrm{D}_1\mathrm{D}_1 f(x) & \cdots & \mathrm{D}_1\mathrm{D}_n f(x) \\ \vdots & & \vdots \\ \mathrm{D}_n\mathrm{D}_1 f(x) & \cdots & \mathrm{D}_n\mathrm{D}_n f(x) \end{pmatrix} \in \mathbb{R}^{n \times n}$$

die *Hesse-Matrix von f in x* bezeichnet. Aufgrund der Annahme $f \in C^2(\mathbb{R}^n)$ und dem Satz 14.3 von Schwarz ist $Df(x)$ sogar eine symmetrische Matrix: Es gilt

$$\mathrm{D}_j\mathrm{D}_k f(x) = \frac{\partial^2 f}{\partial x_j \partial x_k}(x) = \frac{\partial^2 f}{\partial x_k \partial x_j}(x) = \mathrm{D}_k\mathrm{D}_j f(x), \quad j, k \in \{1, \dots, n\}.$$

Satz 16.4 (Taylor-Formel in $\mathbb{R}^n$)**.** *Es sei $\Omega \subset \mathbb{R}^n$ offen, $f \in C^2(\Omega)$, $a \in \Omega$ und $\varepsilon > 0$ derart, dass $B_\varepsilon(a) \subset \Omega$ gilt, wobei $B_\varepsilon(a)$ ebenfalls offen ist. Für alle $h \in B_\varepsilon(0) \subset \mathbb{R}^n$ gilt dann*

$$\begin{aligned} f(a+h) =& f(a) + \big(\mathrm{D}f(a)\big)(h) + \frac{1}{2} \cdot h^\intercal \cdot Hf(a) \cdot h + r(h) \\ \\ =& f(a) + \langle \operatorname{grad} f(a), h \rangle + \frac{1}{2} \cdot h^\intercal \cdot Hf(a) \cdot h + r(h), \end{aligned}$$

mit Restterm $r(h) = o\left(\|h\|^2\right)$.

LINK: TEIL 1 DER 22. VORLESUNG VOM 06.07.2022

Beweis. Es sei $f \in C^2(\Omega)$. Setze nun

$$\gamma\colon (-1, 1) \mapsto \Omega \subset \mathbb{R}^n, \quad t \mapsto a + t \cdot h, \quad \text{sowie} \quad g := f \circ \gamma\colon (-1 - 1) \to \mathbb{R}.$$

Anwendung der Kettenregel 15.13 liefert dann

$$\frac{\mathrm{d}}{\mathrm{d}t} g(t) = \sum_{j=1}^{n} \mathrm{D}_j f\big(\gamma(t)\big) \cdot \overbrace{\frac{\mathrm{d}\gamma_j}{\mathrm{d}t}(t)}^{=h_j} = \mathrm{D}f\big(a + t \cdot h\big) \cdot h.$$

Mit der erneuten Anwendung der Kettenregel erhält man dann

$$g''(t) = \frac{\mathrm{d}}{\mathrm{d}t}\big(g'(t)\big)$$

$$
\begin{aligned}
&= \frac{\mathrm{d}}{\mathrm{d}t}\left(\sum_{j=1}^{n} h_j \cdot \mathrm{D}_j f\big(\gamma(t)\big)\right) \\
&= \sum_{j=1}^{n} h_j \cdot \frac{\mathrm{d}}{\mathrm{d}t}\Big(\mathrm{D}_j f\big(\gamma(t)\big)\Big) \\
&= \sum_{j=1}^{n} h_j \sum_{k=1}^{n} \Big(\mathrm{D}_k \mathrm{D}_j f\Big)\Big(\gamma(t)\Big) \cdot h_k \\
&= h^\intercal \cdot \underbrace{\Big(Hf\Big)\Big(a + t \cdot h\Big)}_{\text{Hesse-Matrix}} \cdot h.
\end{aligned}
$$

Wende nun die Taylor-Formel zweiter Ordnung auf $g\colon (-1,1) \to \mathbb{R}$ an und erhalte

$$
\underbrace{g(1)}_{=f(a+h)} = \underbrace{g(0)}_{=f(a)} + g'(0) + \frac{1}{2} g''(\xi), \quad \xi \in (0,1).
$$

Einsetzen ergibt dann

$$
f(a+h) = f(a) + \mathrm{D}f(a)h + \frac{1}{2}h^\intercal Hf(a)h \underbrace{- \frac{1}{2}h^\intercal Hf(a)h + \frac{1}{2}h^\intercal Hf(a+\xi\cdot h)h}_{=:r(h)=\frac{1}{2}h^\intercal\cdot(-Hf(a)+Hf(a+\xi h))\cdot h}.
$$

Mit der Anwendung der Cauchy-Schwarz-Ungleichung folgt dann

$$
\begin{aligned}
|r(h)| &\leq \frac{1}{2}\Big\| \overbrace{\Big(Hf(a) - Hf(a+\xi h)\Big)h}^{\text{Lipschitz}} \Big\| \cdot \|h\| \\
&\overset{11.25}{\leq} \frac{1}{2}\|Hf(a) - Hf(a+\xi h)\| \cdot \|h\| \cdot \|h\| \\
&= \underbrace{\frac{1}{2}\left(\sum_{j,k=1}^{n} \Big|\overbrace{\mathrm{D}_j\mathrm{D}_k f(a) - \mathrm{D}_j\mathrm{D}_k f(a+\xi h)}^{\xrightarrow[h\to 0]{f\in C^2(\Omega)} 0}\Big|^2\right)^{\frac{1}{2}}}_{\xrightarrow[h\to 0]{} 0} \cdot \|h\|^2 \xrightarrow[h\to 0]{} 0. \qquad \square
\end{aligned}
$$

Die erarbeitete Taylor-Formel ist zentral für die Untersuchung lokaler Extremstellen von Funktionen mehrerer Variablen, die im folgenden Kapitel behandelt wird.

17 Lokale Extrema

Definition 17.1. Es sei (X, d) ein metrischer Raum (z.B. eine Teilmenge des $\mathbb{R}^n, n \in \mathbb{N}$), $a \in X$ und $f\colon X \to \mathbb{R}$. Wir sagen:

- f hat ein *globales Maximum in a*, falls
$$\forall\, x \in X : \quad f(x) \leq f(a);$$
- f hat ein *strikes globales Maximum in a*, falls
$$\forall\, x \in X \setminus \{a\} : \quad f(x) < f(a).$$

Dann heißt $f(a)$ *globales Maximum von f* und a *(strikte) globale Maximalstelle von f.*

Wir sagen:

- f hat ein *lokales Maximum in a*, falls ein $\varepsilon > 0$ existiert, so dass
$$B_\varepsilon(a) \subset X \quad \text{und} \quad f(x) \leq f(a) \quad \forall\, x \in B_\varepsilon(a);$$
- f hat ein *striktes lokales Maximum in a*, falls ein $\varepsilon > 0$ existiert, so dass
$$B_\varepsilon(a) \subset X \quad \text{und} \quad f(x) < f(a) \quad \forall\, x \in B_\varepsilon(a) \setminus \{a\}.$$

In diesem Fall heißt $f(a)$ *lokales Maximum in a* und a *(strikte) lokale Maximalstelle von f.*

Analog sind das *(strikte) globale bzw. lokale Minimum* $f(a)$ von f in a und die *globale bzw. lokale Minimalstelle* a von f definiert.

f nimmt in a ein *Extremum* an, falls es dort ein Maximum oder ein Minimum annimmt und a heißt dann *Extremalstelle von f.*

Satz 17.2 (Notwendiges Kriterium für eine Extremalstelle im Inneren)**.** *Es sei $\Omega \subset \mathbb{R}^n$ offen, $f\colon \Omega \to \mathbb{R}$ differenzierbar in $a \in \Omega$ und a eine lokale Extremalstelle von f. Dann gilt*
$$(\mathrm{D}f)(a) = \underbrace{\mathbf{0}^{\intercal}}_{\in \mathbb{R}^n}.$$

Beweis. Es sei $a \in \Omega$. Da Ω offen ist, existiert ein $\varepsilon > 0$ derart, dass $B_\varepsilon(a) \subset \Omega$. Es sei weiter $h \in \mathbb{R}^n$ mit $\|h\| = 1$ und φ definiert durch
$$\varphi\colon (-\varepsilon, \varepsilon) \to \mathbb{R}, \quad t \mapsto f(a + t \cdot h).$$

Dann ist $\varphi(0) = f(a)$ ein lokales Extremum von φ, da es eines von f ist. Nach dem notwendigen Kriterium für Extremalstellen für Funktionen einer Variablen (vgl. Lemma 12.2, Analysis I (Lehramt)) gilt dann $\varphi'(0) = 0$. Nach der Kettenregel 15.13 gilt außerdem
$$\varphi'(t) = \mathrm{D}f(a + t \cdot h) \cdot h = \Big\langle \operatorname{grad} f(a + t \cdot h), h \Big\rangle.$$

Gilt also $\varphi'(0) = 0$, so ist $\operatorname{grad} f(a) \perp h$. Dieser Schluss gilt insbesondere für den Basisvektor $h = e_j$, $j \in \{1, \dots, n\}$, und damit ist

$$\Big(\operatorname{grad} f(a)\Big)_j = \mathrm{D}_j f(a) = 0.$$

Da dies für alle j gilt, folgt $\mathrm{D}f(a) = 0$. □

Definition 17.3. Es sei $\Omega \subset \mathbb{R}^n$ offen. Ist die Funktion $f\colon \Omega \to \mathbb{R}$ differenzierbar in $a \in \Omega$ mit

$$(\mathrm{D}f)(a) = \mathbf{0}^\intercal,$$

so heißt a *kritische Stelle von* f.

Insbesondere ist jede Extremalstelle von f eine kritische Stelle von f. Die Umkehrung gilt im Allgemeinen jedoch nicht, wie das folgende Beispiel zeigt:

Beispiel 17.4. Gegeben sei die Funktion

$$g\colon \mathbb{R}^2 \to \mathbb{R}, \quad (x, y) \mapsto x \cdot y.$$

Es gilt

$$\mathrm{D}_1 g(x, y) = y \quad \text{und} \quad \mathrm{D}_2 g(x, y) = x, \quad \text{also} \quad \mathrm{D}g(0, 0) = \begin{pmatrix} 0 & 0 \end{pmatrix},$$

wir haben also eine kritische Stelle vorliegen. Dagegen ist $(0, 0)$ jedoch keine Extremalstelle: Dazu beachten wir eine Umgebung von $(0, 0)$. Für $\varepsilon > 0$ gilt nämlich

$$g(\varepsilon, \varepsilon) = \varepsilon^2 > 0 \quad \text{und} \quad g(\varepsilon, -\varepsilon) = -\varepsilon^2 < 0.$$

In dieser Umgebung nimmt g also sowohl negative als auch positive Werte an.

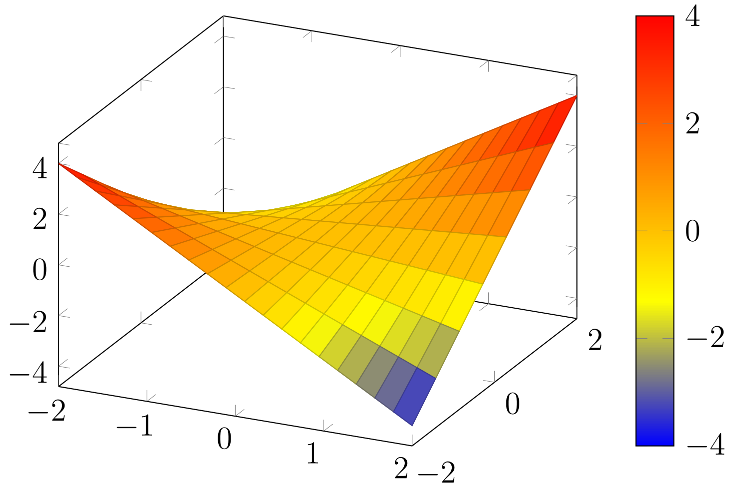

Abbildung 11: Graph der Funktion g für den Definitionsbereich $(-2, 2)^2$.

Der Graph von g sieht aus wie ein Sattel. Tatsächlich wird der Punkt $(0, 0)$ auch *Sattelpunkt von* g genannt.

LINK: TEIL 2 DER 22. VORLESUNG VOM 06.07.2022

Wie in der Analysis I (Lehramt) ist es möglich, hinreichende Kriterien für ein Maximum bzw. für ein Minimum von f zu formulieren. Dies erfolgt unter Bezugnahme auf der zweiten Ableitung (also der Hessematrix $(Hf)(a)$) von f. Dazu definieren wir zunächst bestimmte Eigenschaften von Matrizen.

Definition 17.5. Es sei $A = (a_{jk}) \in \mathbb{R}^{n\times n}$ eine symmetrische Matrix, es gilt also

$$a_{jk} = a_{kj} \quad \forall\, j,k \in \{1,\dots,n\}.$$

Es sei außerdem

$$q = q_A\colon \mathbb{R}^n \to \mathbb{R}, \quad q(x) := x^\intercal A x = \langle x, Ax\rangle$$

die zugehörige *quadratische Form der Matrix A*. Dann heißt A

- *positiv definit*, falls $q(x) > 0$ für alle $x \in \mathbb{R}^n \setminus \{\mathbf{0}\}$,
- *positiv semidefinit*, falls $q(x) \geq 0$ für alle $x \in \mathbb{R}^n$,
- *negativ definit*, falls $q(x) < 0$ für alle $x \in \mathbb{R}^n \setminus \{\mathbf{0}\}$,
- *negativ semidefinit*, falls $q(x) \leq 0$ für alle $x \in \mathbb{R}^n$ und
- *indefinit*, falls $x, y \in \mathbb{R}^n$ existieren mit

$$q(x) > 0 \quad \text{und} \quad q(y) < 0.$$

Insbesondere ist q (quadratisch) homogen: Es gilt

$$\forall\, t \in \mathbb{R}: \quad q(t\cdot x) = \langle t\cdot x, A(t\cdot x)\rangle = t^2 \cdot x^\intercal A x = t^2 \cdot q(x).$$

Satz 17.6. *Es sei $\Omega \subset \mathbb{R}^n$ offen, $f \in C^2(\Omega)$ und $a \in \Omega$ eine kritische Stelle von f. Dann gelten die folgenden Aussagen:*

(a) *Falls $(Hf)(a)$ positiv definit ist, so hat f in a ein striktes lokales Minimum;*

(b) *Falls $(Hf)(a)$ negativ definit ist, so hat f in a ein striktes lokales Maximum;*

(c) *Falls $(Hf)(a)$ indefinit ist, so hat f in a kein Extremum vorliegen. In diesem Fall nennen wir a einen* ***Sattelpunkt von f****.*

Dabei handelt es sich nicht um Äquivalenzen!

Beweis. Es sei $a \in \Omega$. Da Ω offen ist, es existiert ein $\varepsilon > 0$ derart, dass $B_\varepsilon(a) \subset \Omega$. Wegen $f \in C^2(\Omega)$ kann man dann die Taylor-Formel aus Satz 16.4 für $h \in B_\varepsilon(0)$ anwenden: Es gilt

$$f(a+h) - f(a) = \overbrace{\underbrace{\langle \operatorname{grad} f(a), h\rangle}_{\substack{0,\\ \text{da } a \text{ krit. St.}}}}^{=0} + \frac{1}{2}\langle h, (Hf(a))h\rangle + r(h) = \frac{1}{2}\cdot q_{Hf(a)}(h) + r(h)$$

mit $r(h) = o(\|h\|^2)$ für $h \to 0$.

(a) Es sei nun $Hf(a)$ positiv definit. Nach Satz 17.7 (siehe unten) existiert dann eine Konstante $m > 0$ mit

$$q_{Hf(a)}(h) \geq 4m \cdot \|h\|^2, \quad x \in \mathbb{R}^n.$$

Wegen $r(h) = o(\|h\|^2)$ existiert außerdem ein $\delta \in (0, \varepsilon)$ derart, dass

$$|r(h)| \leq m\|h\|^2 \quad \forall h \in B_\delta(0).$$

Daraus folgt

$$\begin{aligned} \forall\, h \in B_\delta(0) \setminus \{\mathbf{0}\} : \quad f(a+h) - f(a) \geq& \frac{1}{2} q_{Hf(a)}(h) - |r(h)| \\ \geq& 2m\|h\|^2 - m\|h\|^2 \\ =& m\|h\|^2 > 0, \end{aligned}$$

f hat also ein lokales Minimum in a.

(b) Die Aussage folgt direkt unter Anwendung von Teil (a) auf $-f$.

(c) Es sei nun $Hf(a)$ indefinit. Dann existieren $h, k \in \mathbb{R}^n$ derart, dass

$$\lambda := q_{Hf(a)}(h) > 0 \quad \text{und} \quad \mu := q_{Hf(a)}(k) < 0.$$

Wähle nun $\delta \in \left(0, \frac{\varepsilon}{\max\{\|h\|,\|k\|\}}\right)$ klein genug, so dass

$$\begin{aligned} \left|r(th)\right| \leq& \frac{\lambda}{4} t^2, \\ \left|r(tk)\right| \leq& \frac{|\mu|}{4} t^2 \quad \text{und} \end{aligned}$$

$$f(a+th) - f(a) = \frac{1}{2} q_{=q_{Hf(a)}}(th) + r(th)$$

für $|t| \leq \delta$. Wegen der (quadratischen) Homogenität von q gilt dann

$$f(a+th) - f(a) \geq \frac{1}{2} \cdot t^2 \cdot \overbrace{q(h)}^{=\lambda} - \frac{\lambda}{4} \cdot t^2 = \frac{\lambda}{2} \cdot t^2 > 0$$

sowie

$$f(a+tk) - f(a) = \frac{1}{2} q(tk) + r(tk) \leq \frac{1}{2} \overbrace{q(k)}^{=\mu<0} + \frac{|\mu|}{4} t^2 = \frac{\mu}{4} t^2 < 0.$$

Dabei haben wir Folgendes benutzt:

$$r(tx) = \|tx\|^2 \cdot \varepsilon(tx) = t^2 \|x\|^2 \cdot \varepsilon(tx),$$

wobei $\lim_{y\to 0}\varepsilon(y) = 0$. Daraus folgt

$$\frac{|r(tx)|}{t^2} = \|x\|^2 \cdot (tx) \xrightarrow[t\to 0]{} 0.$$

Das heißt, dass $r(x) = o(\|x\|^2) \quad \Rightarrow \quad r(tx) = o(t^2)$. Also nimmt f in jeder Umgebung von a sowohl kleinere als auch größere Werte als $f(a)$ an. Damit hat f in a kein Extremum vorliegen. $\square$

Nun folgt ein Exkurs über (in)definite Matrizen aus der Linearen Algebra, damit wir das obige hinreichende Kriterium besser verstehen und anwenden können. Dazu beginnen wir mit einem Satz, der im obigen Beweis des Satzes 17.6 verwendet worden ist.

Satz 17.7. *Es sei $A \in \mathbb{R}^{n\times n}$ eine symmetrische Matrix. Dann sind die folgenden Aussagen äquivalent:*

(i) *A ist positiv definit;*

(ii) *Es existiert ein $\lambda_1 > 0$ derart, dass*

$$\forall\, x \in \mathbb{R}^n : \quad q_A(x) \geq \lambda_1 \cdot \|x\|^2 = q_{\lambda_1 \cdot I_n}(x).$$

Beweis.

(ii) $\Rightarrow$ (i) Es gelte $q_A(x) \geq \lambda_1 \|x\|^2$. Wegen 10.1 (N1) folgt sofort

$$\forall x \in \mathbb{R}^n \setminus \{\mathbf{0}\} : \quad q_A(x) \geq \overbrace{\lambda_1}^{>0} \underbrace{\|x\|}_{>0} > 0.$$

Also ist A positiv definit.

(i) $\Rightarrow$ (ii) Nach Beispiel 15.4 (b) ist $x \mapsto q_A(x)$ (total) differenzierbar und damit insbesondere stetig. Außerdem ist

$$\mathbb{S} := \Big\{x \in \mathbb{R}^n \ : \ \|x\| = 1\Big\} = \partial B_1(\mathbf{0}) \subset \mathbb{R}^n$$

beschränkt und abgeschlossen, also kompakt. Damit nimmt q_A insbesondere ein Minimum auf $\mathbb{S}$ an, d.h. es existiert ein $p \in \mathbb{S}$ mit

$$\inf_{x\in\mathbb{S}} q_A(x) = q_A(p) > 0, \tag{$*$}$$

da A positiv definit ist. Setze nun $\lambda_1 := q_A(p)$. Für alle $y \in \mathbb{R}^n$ folgt dann

$$\begin{aligned} q_A(y) &= \langle A \cdot y, y\rangle \\ &= \|y\|^2 \cdot \left\langle A \cdot \Bigg(\underbrace{\frac{1}{\|y\|} y}_{=:x\in\mathbb{S}}\Bigg), \frac{1}{\|y\|} y\right\rangle \\ &= \|y\|^2 \cdot q_A(x) \\ &\overset{(*)}{\geq} \|y\|^2 \cdot \lambda_1. \end{aligned}$$

$\square$

Genauer gilt sogar:

Satz 17.8. *Es sei $A \in \mathbb{R}^{n\times n}$ eine symmetrische Matrix. Dann gelten die folgenden Aussagen:*

(a) *A ist genau dann positiv definit, wenn alle Eigenwerte von A positiv sind;*

(b) *A ist genau dann positiv semidefinit, wenn alle Eigenwerte von A nicht-negativ sind;*

(c) *A ist genau dann negativ definit, wenn alle Eigenwerte von A negativ sind;*

(d) *A ist genau dann negativ semidefinit, wenn alle Eigenwerte von A nicht-positiv sind;*

(e) *A ist genau dann indefinit, wenn A sowohl positive als auch negative Eigenwerte besitzt.*

(Ohne Beweis)

Link: Teil 1 der 23. Vorlesung vom 07.07.2022

Da es auch vorkommt, dass man die Eigenwerte einer Matrix nicht so einfach einsehen kann, folgt nun ein Kriterium, das einfacher zu prüfen ist:

Satz 17.9 (Hurwitz-Kriterium)**.** *Es sei $A = (a_{jk}) \in \mathbb{R}^{n\times n}$ eine symmetrische Matrix und für $m \in \{1, \dots, n\}$ sei*

$$A_m \in \mathbb{R}^{m\times m} \quad \textit{mit} \quad \big(A_m\big)_{j,k} = a_{jk} \quad \textit{für } j,k \in \{1,\dots,m\},$$

also

$$A = \begin{pmatrix} \mathbf{a_{11}} & \cdots & \mathbf{a_{1m}} & a_{1(m+1)} & \cdots & a_{1n} \\ \vdots & & \vdots & \vdots & & \vdots \\ \mathbf{a_{m1}} & \cdots & \mathbf{a_{mm}} & a_{m(m+1)} & \cdots & a_{mn} \\ a_{(m+1)1} & \cdots & a_{(m+1)m} & a_{(m+1)(m+1)} & \cdots & a_{(m+1)n} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{n1} & \cdots & a_{nm} & a_{n(m+1)} & \cdots & a_{nn} \end{pmatrix}, \quad A_m = \begin{pmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mm} \end{pmatrix}.$$

Dann gelten die folgenden Aussagen:

(i) *A ist genau dann positiv definit, wenn für alle $m \in \{1,\dots,n\}$ der **führende Hauptminor** m**-ter Ordnung** positiv ist, es gilt also*

$$\det(A_m) > 0 \quad \forall\, m \in \{1,\dots,n\};$$

(ii) *A ist genau dann negativ definit, wenn*

$$(-1)^m \det(A_m) > 0 \quad \forall\, m \in \{1,\dots,n\}.$$

Für Hauptminoren gerader Ordnung, also $m = 2l, l \in \mathbb{N}$ gilt insbesondere

$$(-1)^{2l} \det(A_{2l}) = \det(A_{2l}) > 0.$$

(iii) *Existiert ein* $l \in \mathbb{N}$ *mit*
$$l \leq \frac{n}{2} \quad \textit{und} \quad \det(A_{2l}) < 0,$$
so ist A *indefinit.*

(Ohne Beweis)

Korollar 17.10 (Kriterien für $f\colon \mathbb{R}^2 \to \mathbb{R}$). *Es sei* $\Omega \subset \mathbb{R}^2$ *offen,* $f \in C^2(\Omega)$ *und* $a \in \Omega$ *eine kritische Stelle von* f*. Dann gelten die folgenden Aussagen:*

(a) *Es gilt*
$$\det\big(H(f(a)\big) > 0 \quad \textit{und} \quad \mathrm{D}_1\mathrm{D}_1 f(a) > 0$$
$$\Leftrightarrow \quad \det\big(H(f(a)\big) > 0 \quad \textit{und} \quad \mathrm{D}_2\mathrm{D}_2 f(a) > 0$$
$$\Rightarrow \quad f \textit{ hat in } a \textit{ ein striktes lokales Minimum.}$$

(b) *Es gilt*
$$\det\big(H(f(a)\big) > 0 \quad \textit{und} \quad \mathrm{D}_1\mathrm{D}_1 f(a) < 0$$
$$\Leftrightarrow \quad \det\big(H(f(a)\big) > 0 \quad \textit{und} \quad \mathrm{D}_2\mathrm{D}_2 f(a) < 0$$
$$\Rightarrow \quad f \textit{ hat in } a \textit{ ein striktes lokales Maximum.}$$

(c) *Es gilt*
$$\det\big(H(f(a)\big) < 0$$
$$\Rightarrow \quad f \textit{ hat in } a \textit{ einen Sattelpunkt.}$$

(Ohne Beweis)

Beispiel 17.11 (Semidefinite Hesse-Matrix). Gegeben sei die Funktion

Im Video: **Beispiel 17.10**.

$$f\colon \mathbb{R}^2 \to \mathbb{R}, \quad (x,y) \mapsto x^2 + ay^2 + by^4, \quad a, b \in \mathbb{R}.$$

Für die erste Ableitung gilt:
$$\mathrm{D}_1 f(x,y) = 2x, \quad \mathrm{D}_2 f(x,y) = 2ay + 4by^3.$$

Für die zweite Ableitung (wobei $\mathrm{D}_{kj} f := \mathrm{D}_k \mathrm{D}_j f$, $k \in \{1,2\}$) gilt:
$$\mathrm{D}_{11} f(x,y) = 2, \quad \mathrm{D}_{12} f(x,y) = \mathrm{D}_{21} f(x,y) = 0, \quad \mathrm{D}_{22} f(x,y) = 2a + 12by^2.$$

Damit ist die Hesse-Matrix gegeben durch

$$Hf(x,y) = \begin{pmatrix} 2 & 0 \\ 0 & 2a + 12by^2 \end{pmatrix}.$$

In $(0,0)$ gilt dabei

$$\mathrm{D}f(0,0) = \begin{pmatrix} 0 & 0 \end{pmatrix} \quad \text{und} \quad Hf(0,0) = \begin{pmatrix} 2 & 0 \\ 0 & 2a \end{pmatrix}.$$

Betrachten wir nun die folgende Fallunterscheidung:

Fall 1: Es gilt $a > 0$. Dann besitzt f nach Satz 17.6 in $(0,0)$ ein striktes lokales Minimum. Ist dies jedoch auch ein globales Minimum? Dies hängt von b ab! Betrachte dazu weitere Unterfälle:

Fall 1a: Es gilt $b \geq 0$. Dann folgt

$$f(0,0) = 0 < x^2 + ay^2 \leq f(x,y) \quad \text{für } (x,y) \neq (0,0).$$

Wir haben also ein globales Minimum.

Fall 1b: Es gilt $b < 0$. Dann ist

$$ay^2 + by^4 = -by^2\left(-\frac{a}{b} - y^2\right) = |by^2|\left(\left|\frac{a}{b}\right| - y^2\right) < 0, \quad \text{sobald } |y| > \sqrt{\left|\frac{a}{b}\right|}.$$

Es folgt also

$$f(0,y) < 0 = f(0,0), \quad \text{sobald } |y| > \sqrt{\left|\frac{a}{b}\right|},$$

weshalb wir kein globales Minimum vorliegen haben.

Fall 2: Es gilt $a = 0$. Dann ist

$$Hf(0,0) = \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix}$$

positiv semidefinit. Damit ist Satz 17.6 nicht anwendbar. Betrachte nun die folgenden Unterfälle:

Fall 2a: Es gilt $b < 0$. Für $\varepsilon \neq 0$ ist dann

$$f(0,\varepsilon) = b\varepsilon^4 < 0 = f(0,0).$$

Also ist $f(0,0)$ kein Minimum.

Fall 2b: Es gilt $b > 0$. Für $(x,y) \neq (0,0)$ ist dann

$$f(0,0) = 0 < x^2 + by^4 = f(x,y),$$

haben also ein striktes globales Minimum vorliegen.

Fall 2c: Es gilt $b = 0$. Dann ist $f(0,0) = 0$ ein globales, aber kein striktes Minimum, da

$$f(x,y) = x^2 \geq 0, \quad \text{aber} \quad f(0,y) = 0 = f(0,0).$$

Fall 3: Es gilt $a < 0$. Dann ist

$$Hf(0,0) = \begin{pmatrix} 2 & 0 \\ 0 & 2a \end{pmatrix}$$

indefinit. Hier ist $(0,0)$ keine Extremalstelle.

Für $g\colon \mathbb{R}^2 \to \mathbb{R}$, $\quad g(x,y) = x \cdot y$ (s. Beispiel 17.4) gilt ähnlich:

$$(\mathrm{D}g)(0,0) = (y,x) = (0,0) \quad \text{in } (0,0)$$

und

$$(Hg)(0,0) = \begin{pmatrix} \mathrm{D}_{11}g & \mathrm{D}_{21}g \\ \mathrm{D}_{12}g & \mathrm{D}_{22}g \end{pmatrix}(0,0) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Es gilt $\det\big(Hg(0,0)\big) = 0 - 1 < 0$. Also ist $(Hg)(0,0)$ indefinit nach dem Hurwitz-Kriterium 17.9 (iii). Damit haben wir keine Extremalstelle vorliegen, wie schon explizit berechnet wurde.

Beispiel 17.12. Gegeben sei die Funktion

Im Video: **Beispiel 17.11.**

$$f\colon \mathbb{R}^2 \to \mathbb{R}, \quad (x,y) \mapsto x^4 + x^3 + y^2.$$

Dann ist

$$\operatorname{grad} f(x,y) = \begin{pmatrix} 4x^3 + 3x^2 \\ 2y \end{pmatrix} = \begin{pmatrix} x^2 \cdot (4x+3) \\ 2y \end{pmatrix} \overset{!}{=} \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad \Rightarrow \quad y = 0, \quad x \in \left\{-\frac{3}{4}, 0\right\}.$$

Berechne nun die Hesse-Matrix

$$Hf(x,y) = \begin{pmatrix} 12x^2 + 6x & 0 \\ 0 & 2 \end{pmatrix}$$

und werte sie an den kritischen Stellen aus: Es gilt

$$Hf(0,0) = \begin{pmatrix} 0 & 0 \\ 0 & 2 \end{pmatrix} \quad \text{positiv semidefinit}$$

und

$$Hf\left(-\frac{3}{4}, 0\right) = \begin{pmatrix} 12 \cdot \frac{9}{16} - 6 \cdot \frac{3}{4} & 0 \\ 0 & 2 \end{pmatrix} = \begin{pmatrix} \frac{9}{4} & 0 \\ 0 & 2 \end{pmatrix} \quad \text{positiv definit.}$$

Also hat f nach Satz 17.6 in $\left(-\frac{3}{4}, 0\right)$ ein striktes lokales Minimum. Was passiert aber in $(0,0)$? Wegen $\det\big(Hf(0,0)\big) = 0$ helfen uns der Satz 17.9 und das Korollar 17.10 leider nicht weiter. Für $t \in (-1,0)$ gilt aber

$$f(t,0) = t^4 + t^3 = \underbrace{t^3}_{<0} \cdot \underbrace{(t+1)}_{>0} < 0 = f(0,0) < f(x,0) \quad \text{für } x > 0.$$

Also ist $(0,0)$ keine Extremalstelle von f.

LINK: TEIL 2 DER 23. VORLESUNG VOM 07.07.2022

Beispiel 17.13. Gegeben sei die Funktion

$$f\colon \mathbb{R}^3 \to \mathbb{R}, \quad (x, y, z) \mapsto 6xy - 3y^2 - 2x^3 - yz^2.$$

Dann ist

$$\mathrm{D}f(x, y, z) = \begin{pmatrix} 6y - 6x^2 & 6x - 6y - z^2 & -2yz \end{pmatrix}.$$

Aus $\mathrm{D}f(x, y, z) = \mathbf{0}^\intercal$ folgt daher sofort $yz = 0$. Falls $y = 0$ gilt, so ist auch $x = 0$ und damit $z = 0$. Falls $z = 0$ ist, so gilt

$$x = y \quad \text{und} \quad 6y - 6x^2 = 6x \cdot (1 - x), \quad \text{d.h.} \quad x \in \{0, 1\}.$$

Damit sind nur $(0, 0, 0)$ und $(1, 1, 0)$ kritische Stellen von f. Berechne nun die Hesse-Matrix

$$Hf(x, y, z) = \begin{pmatrix} -12x & 6 & 0 \\ 6 & -6 & -2z \\ 0 & -2z & -2y \end{pmatrix}$$

und werte sie an den kritischen Stellen aus: Es gilt

$$A := Hf(0, 0, 0) = \begin{pmatrix} 0 & 6 & 0 \\ 6 & -6 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad \Rightarrow \quad \det\left(Hf(0, 0, 0)\right) = 0.$$

Also ist $Hf(0, 0, 0)$ weder positiv noch negativ (semi)definit. Es gilt jedoch

$$\det(A_2) = \det \begin{pmatrix} 0 & 6 \\ 6 & -6 \end{pmatrix} = 0 - 36 < 0.$$

Also besagt das Hurwitz-Kriterium 17.9 (iii), dass A_2 indefinit ist, d.h. es existiert ein $h_\pm \in \mathbb{R}^2$ mit

$$\langle h_\pm, Ah_\pm \rangle = \pm 1.$$

Für $x_+ := (h_+, 0), x_- := (h_-, 0) \in \mathbb{R}^3$ gilt dann aber

$$\langle x_\pm, (Hf)(0, 0, 0)x_\pm \rangle = \pm 1,$$

weshalb $Hf(0, 0, 0)$ ebenfalls indefinit ist. Also hat f einen Sattelpunkt in $(0, 0, 0)$. In der Stelle $(1, 1, 0)$, also für $B := Hf(1, 1, 0)$, gilt:

$$\det(B_1) = \det(-12) = -12 < 0,$$

$$\det(B_2) = \det \begin{pmatrix} -12 & 6 \\ 6 & -6 \end{pmatrix} = 12 \cdot 6 - 6 \cdot 6 = 36 > 0,$$

$$\det(B_3) = \det \begin{pmatrix} -12 & 6 & 0 \\ 6 & -6 & 0 \\ 0 & 0 & -2 \end{pmatrix} = -2 \cdot \det(B_2) = -72 < 0.$$

Nach dem Hurwitz-Kriterium 17.9 (ii) ist dann $Hf(1, 1, 0)$ negativ definit, weshalb f in $(1, 1, 0)$ ein striktes lokales Maximum besitzt.

18 Implizite Funktionen

Es dürften bereits (mindestens) zwei Arten bekannt sein, Teilmengen des $\mathbb{R}^n, n \in \mathbb{N}$, zu beschreiben:

- Zum einen über **Nullstellengebilde**: So ist zum Beispiel bei einer linearen Abbildung $L\colon \mathbb{R}^n \to \mathbb{R}^m$ die Menge

$$\operatorname{Kern}(L) = \Big\{x \in \mathbb{R}^n \ : \ L(x) = 0\Big\}$$

ein Nullstellengebilde. Es sei konkret eine lineare Abbildung gegeben durch

$$L\colon \mathbb{R}^3 \to \mathbb{R}^3, \quad L\begin{pmatrix} x \\ y \\ z \end{pmatrix} := \underbrace{\begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}}_{=:A} \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x+y \\ 0 \\ z \end{pmatrix}.$$

Dann ist der *Kern von L* gegeben durch

$$\operatorname{Kern}(L) = \Big\{(x,y,z) \in \mathbb{R}^3 \ : \ A \cdot (x,y,z) = \mathbf{0}\Big\} = \Big\{(x,-x,0) \ : \ x \in \mathbb{R}\Big\}.$$

- Zum anderen über **parametrisierte Mengen**: So ist zum Beispiel kann zum Beispiel die Kreislinie $S \subset \mathbb{R}^2$

$$S := \Big\{(x,y) \in \mathbb{R}^2 \ : \ x^2 + y^2 = 1\Big\}$$

parametrisiert werden: Es gilt

$$S = \left\{\begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 \ : \ x^2 + y^2 = 1\right\} = \left\{\begin{pmatrix} \cos(\varphi) \\ \sin(\varphi) \end{pmatrix} \ : \ \varphi \in [0, 2\pi]\right\}.$$

Wir betrachten nun die folgende Fragestellung: Es sei $f\colon \mathbb{R}^2 \to \mathbb{R}$ eine stetige Funktion und für $c \in \mathbb{R}$ sei weiter

$$\mathcal{N}_c(f) = \Big\{(x,y) \in \mathbb{R}^2 \ : \ f(x,y) = c\Big\}$$

die Niveaumenge. Unter welchen Voraussetzungen kann man dann $\mathcal{N}_c(f)$ (zumindest lokal) als Wege bzw. Kurven, d.h. als Graphen einer stetigen Funktion einer reellen Veränderlichen beschreiben? Genauer: Kann man die Gleichung

$$g(x,y) := f(x,y) - c = 0$$

auflösen?

Beispiel 18.1. Gegeben sei die Funktion

$$g\colon \mathbb{R}^2 \to \mathbb{R}, \quad g(x,y) := x^2 + y^2 - 1.$$

Für $(a,b) \in M_+$, wobei

$$M_+ := \Big\{(x,y) \in \mathbb{R}^2 \ : \ |x| < 1, \quad y > 0, \quad g(x,y) = 0\Big\},$$

gibt es eine eindeutige Funktion $\varphi\colon (-1,1) \to \mathbb{R}$ mit

$$\varphi(a) = b \quad \text{und insbesondere} \quad g\Big(a, \varphi(a)\Big) = 0 \quad \text{für } a \in (-1,1),$$

und zwar

$$\varphi\colon (-1,1) \to \mathbb{R}, \quad \varphi(x) := \sqrt{1-x^2}.$$

Es gilt nämlich

$$g(x, \sqrt{1-x^2}) = x^2 + (1-x^2) - 1 = 0.$$

Analog ist für

$$M_- := \Big\{(x,y) \in \mathbb{R}^2 \ : \ |x| < 1, \quad y < 0, \quad g(x,y) = 0\Big\}$$

die Funktion

$$\psi\colon (-1,1) \to \mathbb{R}, \quad \psi(x) := -\sqrt{1-x^2}$$

die eindeutige Parametrisierung von M_- durch $x \in (-1,1)$, denn es gilt

$$g(\underbrace{x}_{\in(-1,1)}, \underbrace{\psi(x)}_{<0}) = x^2 + (-\sqrt{1-x^2})^2 - 1 = x^2 + (1-x^2) - 1 = 0.$$

Der Punkt $(a,b) = (1,0)$ erfüllt ebenfalls $g(a,b) = 1^2 + 0 - 1 = 0$, aber ist weder in M_+ noch in M_- enthalten. In dieser Stelle können wir $g(x,y) = 0$ nicht nach y auflösen, dafür aber nach x: Es gibt eine eindeutig bestimmte Funktion $\chi\colon (-1,1) \to \mathbb{R}$, so dass

$$\chi(0) = 1 \quad \text{und} \quad g\Big(\chi(y), y\Big) = 0 \quad \text{für } y \in (-1,1),$$

und zwar

$$\chi\colon (-1,1) \to \mathbb{R}, \quad \chi(y) := \sqrt{1-y^2}.$$

Es gilt nämlich

$$\chi(0) = 1, \quad g\Big(\chi(y), y\Big) = 1 - y^2 + y^2 - 1 = 0 \quad \text{und} \quad x = \chi(y) > 0 \quad \text{für } y \in (-1,1).$$

LINK: TEIL 1 DER 24. VORLESUNG VOM 13.07.2022

Allgemeiner gilt:

Satz 18.2 (über implizite Funktionen). *Es sei $\Omega \subset \mathbb{R}^{n+1}$ offen, $f \in C^1(\Omega, \mathbb{R})$ und*

$$(a,b) = (\underbrace{a_1, \dots, a_n}_{\in\mathbb{R}^n}, \overbrace{b}^{\in\mathbb{R}}) \in \Omega \quad \textit{mit} \quad f(a,b) = 0$$

sowie $\mathrm{D}_{n+1}f(a,b) \neq 0$*. Dann gibt es einen offenen Würfel* $W \subset \mathbb{R}^n$ *mit Mittelpunkt in* a *und ein offenes Intervall mit Mittelpunkt in* b*, so dass* $W \times I \subset \Omega$ *und die Gleichung*

$$f(x,y) = 0$$

für alle $x \in W$ *genau eine Lösung* y *in* I *besitzt. Genauer gesagt gibt es eine Funktion* $g\colon W \to I$ *mit*

$$f\Big(x, g(x)\Big) = 0 \quad \text{für } x \in W, \quad g(x) \in I \quad \text{und} \quad g(a) = b.$$

Dieses g *erfüllt*

$$g \in C^1(W) \quad \text{und} \quad \mathrm{D}_j g(x) = -\frac{(\mathrm{D}_j f)\Big(x, g(x)\Big)}{(\mathrm{D}_{n+1} f)\Big(x, g(x)\Big)} \quad \forall\, x \in W, j \in \{1, \ldots, n\}. \qquad (*)$$

Wir haben also die Form

$$W = \Big(a_1 - \beta_1, a_1 + \beta_1\Big) \times \ldots \times \Big(a_n - \beta_n, a_n + \beta_n\Big),$$

$$I = \Big(b - \beta_{n+1}, b + \beta_{n+1}\Big) =: (b_-, b_+),$$

mit $\beta_1, \ldots, b_{n+1} > 0$ *und* $b_j = b_k$ *für alle* $j, k \in \{1, \ldots, n\}$*.*

Beweis.

1. Beweis der (lokalen) Existenz und der Eindeutigkeit von g:
Nach Annahme ist o.B.d.A

$$2p := (\mathrm{D}_{n+1} f)(a,b) = \frac{\partial}{\partial y} f(a,y)\Big|_{y=b} > 0$$

und $\mathrm{D}_{n+1} f\colon \Omega \to \mathbb{R}$ stetig. Also existieren $\delta > 0$ und $\beta_{n+1} > 0$ derart, dass

$$(\mathrm{D}_{n+1} f) \geq p \quad \text{auf } Z := \overline{B_\delta(a)} \times [b_-, b_+], \quad \text{und} \quad Z \text{ offen.}$$

Also ist für jedes $x \in \overline{B_\delta(a)}$ die partielle Funktion

$$f_x\colon [b_-, b_+] \to \mathbb{R}, \quad y \mapsto f_x(y) = f(x,y)$$

streng isoton, insbesondere gilt

$$f(x, b_-) < f(x, b) < f(x, b_+) \quad \text{für } x \in \overline{B_\delta(a)}$$

und speziell

$$f(a, b_-) < \underbrace{f(a,b)}_{=0} < f(a, b_+).$$

Da $f \in C^1(\Omega)$ ist f insbesondere stetig, es existiert also ein kleiner Würfel $W \subset B_\delta(a)$ wie oben, so dass

$$f_+ := \min_{x \in \overline{W}}(x, b_+) > 0 \quad \text{und} \quad f_- := \max_{x \in \overline{W}}(x, b_-) < 0.$$

Insbesondere gilt

$$f(x, b_-) < 0 < f(x, b_+) \quad \forall\, x \in \overline{W}.$$

Aufgrund der Stetigkeit von f ist der Zwischenwertsatz (s. Satz 9.17, Analysis I (Lehramt)) anwendbar. Damit hat dann die Gleichung

$$f(x, y) = 0$$

für alle $x \in W$ eine Lösung $y \in I = (b_-, b_+)$. Aufgrund der strengen Isotonie von $f(x, \cdot)$ auf (b_-, b_+) ist die Lösung sogar eindeutig. Insbesondere impliziert $f(a, y) = 0$ und $y \in I$ schon die Identität $y = b$. Aufgrund der Existenz und Eindeutigkeit ist damit eine Funktion $g\colon W \to I$ wohldefiniert, so dass nach Konstruktion gilt:

$$f\Big(x, g(x)\Big) = 0 \quad \text{für } x \in W, \text{ insbesondere} \quad f\Big(a, g(a)\Big) = 0.$$

2. Beweis der Stetigkeit von g:
Es sei $x \in W$ beliebig und $y := g(x) \in I = (b_-, b_+)$. Da I offen ist, existiert ein $\varepsilon > 0$ mit $[y - \varepsilon, y + \varepsilon] \subset I$. Nun wiederholen wir die Argumentation von oben für den Punkt (x, y) statt für (a, b). Damit existiert ein $\alpha > 0$, so dass $f(\xi, \eta) = 0$ für $\xi \in B_\alpha(x)$ genau eine Lösung $\eta \in (y - \varepsilon, y + \varepsilon)$ besitzt. Wegen der (lokalen) Eindeutigkeit muss dann $\eta = g(\xi)$ gelten. Insbesondere folgt

$$|g(\xi) - g(x)| = |\eta - y| < \varepsilon \quad \forall\, \xi \in B_\alpha(x)$$

$$\Leftrightarrow \quad |g(\xi) - g(x)| = |\eta - y| < \varepsilon \quad \forall\, \xi : |\xi - x| < \alpha.$$

Dieses Argument lässt sich für jedes $\tilde{\varepsilon} \in (0, \varepsilon)$ nachahmen. Da $\tilde{\varepsilon} > 0$ dann beliebig klein gewählt werden kann, ist die Stetigkeit von g nachgewiesen.

3. Beweis der Eigenschaft aus $(*)$:
Es sei $j \in \{1, \dots, n\}$, $x \in W$ und $t \in \mathbb{R} \setminus \{0\}$ klein genug, so dass $x + t \cdot e_j \in W$. Setze nun

$$h := g(x + t \cdot e_j) - g(x) \in \mathbb{R}.$$

Wir wenden nun den Mittelwertsatz der Differentialrechnung (Satz 12.4, Analysis I (Lehramt)) **zweimal** an: Es existieren $\tau, \vartheta \in (0, 1)$ derart, dass

$$f(x + te_j, y) - f(x, y) = t \cdot \mathrm{D}_j f(x + \tau \cdot te_j, y),$$

$$f(x + te_j, y) - f(x + te_j, y + h) = -\, h \cdot \mathrm{D}_{n+1} f(x + te_j, y + \vartheta \cdot h).$$

Setze nun $y = g(x)$ ein und erhalte somit

$$y + h = g(x) + h = g(x + te_j).$$

Damit bekommt man

$$t \cdot \mathrm{D}_j f\big(x + \tau \cdot te_j, g(x)\big) = f\big(x + te_j, g(x)\big) - \underbrace{f\big(x, g(x)\big)}_{\substack{=0,\\ \text{nach Def. v. } g}},$$

$$-h \cdot \mathrm{D}_{n+1} f\big(x + te_j, g(x) + \vartheta \cdot h\big) = f\big(x + te_j, g(x)\big) - \underbrace{f\big(x + te_j, g(x + te_j)\big)}_{=0}.$$

Nach Definition gilt dann

$$\frac{g(x + te_j) - g(x)}{t} = \frac{h}{t} = -\frac{1}{t} \cdot t \cdot \frac{\mathrm{D}_j f\big(x + \tau \cdot te_j, g(x)\big)}{\mathrm{D}_{n+1} f\big(x + te_j, g(x) + \vartheta \cdot h\big)}$$

$$= \frac{\mathrm{D}_j f\big(x + \tau \cdot te_j, g(x)\big)}{\mathrm{D}_{n+1} f\big(x + te_j, g(x) + \vartheta \cdot h\big)}.$$

Nun betrachten wir den Grenzübergang $t \to 0$. Wegen der (bereits bewiesenen) Stetigkeit von g impliziert dies

$$h = h(t) = g(x + te_j) - g(x) \xrightarrow[t \to 0]{} 0.$$

Für $j \in \{1, \ldots, n\}$ gilt also

$$(\mathrm{D}_j g)(x) = \lim_{t \to 0} \frac{g(x + te_j) - g(x)}{t} = -\lim_{t \to 0} \frac{\mathrm{D}_j f\big(x + \overbrace{\tau \cdot te_j}^{\to 0}, g(x)\big)}{\mathrm{D}_{n+1} f\big(x + \underbrace{te_j}_{\to o}, g(x) + \underbrace{\vartheta \cdot h}_{\to 0}\big)}$$

$$= -\frac{\mathrm{D}_j f\big(x, g(x)\big)}{\mathrm{D}_{n+1} f\big(x, g(x)\big)} \neq 0.$$

LINK: TEIL 2 DER 24. VORLESUNG VOM 13.07.2022

Damit haben wir die Differenzierbarkeit von g und die Formel aus $(*)$ bewiesen. Da die rechte Seite der Formel stetig von x abhängt, ist sogar $g \in C^1(W)$. □

Bemerkung 18.3 („Implizites Differenzieren“)**.**

(a) Setzen wir bereits voraus, dass g existiert und stetig differenzierbar ist, kann man $(*)$ einfacher herleiten: Die Eigenschaft

$$\forall\, x \in W : \quad f\big(x, g(x)\big) = 0$$

impliziert für $h(x) = f\big(x, g(x)\big)$ nämlich:

$$0 = \frac{\partial}{\partial x_j} h(x) = (\mathrm{D}_j f)\big(x, g(x)\big) + (\mathrm{D}_{n+1} f)\big(x, g(x)\big) \cdot \frac{\partial g(x)}{\partial x_j},$$

wobei $g(x)$ nur eingesetzt, und nicht abgeleitet, wird.

(b) Setzen wir in Satz 18.2 sogar $f \in C^k(\Omega)$, $k \in \mathbb{N}$, voraus, so folgt mit $(*)$, dass $g \in C^k(W)$. Ähnlich wie in Teil (a) kann man diese Information nutzen, um mittels implizitem Differenzieren höhere Ableitungen von g zu berechnen.

(c) Ist $n = 1$, d.h. $\Omega \subset \mathbb{R} \times \mathbb{R}$, und $x, y \in \mathbb{R}$ skalare Variablen, so lautet $(*)$ umformuliert:

$$g'(x) = -\frac{\mathrm{D}_1 f\big(x, g(x)\big)}{\mathrm{D}_2 f\big(x, g(x)\big)}.$$

(d) Der Satz 18.2 über implizite Funktionen besagt, dass

$$\left\{\begin{pmatrix} x \\ y \end{pmatrix} \in W \times I \;:\; f(x,y) = 0\right\} = \left\{\begin{pmatrix} x \\ y \end{pmatrix} \in W \times I \;:\; y = g(x)\right\} = \left\{\begin{pmatrix} x \\ g(x) \end{pmatrix} \;:\; x \in W\right\}.$$

Das heißt, dass die Niveaumenge $\mathcal{N}_0(f)$ in $W \times I \subset \Omega$ durch den Graphen einer stetig differenzierbaren Funktion g gegeben ist:

$$\mathcal{N}_0(f) \cap (W \times I) = G(g).$$

Beispiel 18.4. Gegeben sei die Funktion

$$f \colon \mathbb{R}^2 \to \mathbb{R}, \quad (x, y) \mapsto y^5 + (x^2 + 1)y + (x^3 - 4).$$

Dann ist $\mathrm{D}_1 f(x, y) = 2xy + 3x^2$ und

$$\mathrm{D}_2 f(x, y) = 5y^4 + (x^2 + 1) > 0 \quad \forall\, (x, y) \in \mathbb{R}^2.$$

Damit sind die Ableitungsbedingungen in Satz 18.2 sogar global erfüllt, und haben damit globale Isotonie vorliegen. Also existiert eine Funktion $g \colon \mathbb{R} \to \mathbb{R}$ mit

$$f\big(x, g(x)\big) = 0 \quad \forall\, x \in \mathbb{R}.$$

Die Formel $(*)$ liefert dann

$$g'(x) = -\frac{\mathrm{D}_1 f\big(x, g(x)\big)}{\mathrm{D}_2 f\big(x, g(x)\big)} = -\frac{2x \cdot g(x) + 3x^2}{5 \cdot (g(x))^4 + x^2 + 1}.$$

Falls x ein kritischer Punkt von g ist, es gilt also $0 = g'(x)$, so muss

$$0 = 2xg(x) + 3x^2 = x \cdot \Big(2g(x) + 3x\Big) \quad \Rightarrow \quad x = 0 \quad \text{oder} \quad y := g(x) = -\frac{3}{2}x$$

gelten. Damit haben wir zwei Kandidaten für Minimal- bzw. Maximalstellen. Dazu untersuchen wir den zweiten Kandidaten y näher: Dazu sei

$$h(x) := f\left(x, -\frac{3}{2}x\right) = -\left(\frac{3}{2}\right)^5 x^5 - (x^2+1)\cdot\frac{3}{2}x + x^3 - 4 = -\left(\frac{3}{2}\right)^5 x^5 - \frac{1}{2}x^3 - \frac{3}{2}x - 4.$$

Wegen

$$h'(x) = \underbrace{-5\cdot\left(\frac{3}{2}\right)^5 x^4}_{leq0} \underbrace{-\frac{3}{2}x^2}_{\leq 0} \underbrace{-\frac{3}{2}}_{<0} < 0 \quad \forall\, x \in \mathbb{R},$$

ist h strikt antiton und $h(x) = 0$ hat genau eine Lösung $x^* \in \mathbb{R}$, da h stetig ist und $\lim_{h\to\infty} h(x) = -\infty$ und $\lim_{h\to-\infty} h(x) = \infty$. Also gilt

$$0 = h(x^*) = f\left(x^*, -\frac{3}{2}x^*\right) \quad \Rightarrow \quad g(x^*) = -\frac{3}{2}x^3 \quad \Rightarrow \quad \text{kritischer Punkt.}$$

Da

$$h(-1) = \left(\frac{3}{2}\right)^5 - 2 = \frac{179}{32} > 0 \quad \text{und} \quad h(0) = -4 < 0,$$

ist $x^* \in (-1, 0)$. Über Näherungsverfahren erhält man dann $x^* \approx -0,80307$.

Eine weitere Diskussion ergibt

$$g(0) \in (1,2), \quad g'(0) = 0, \quad g''(0) < 0,$$

$$\lim_{x\to-\infty} g(x) = \infty, \quad g \text{ besitzt in } x^* \text{ ein lokales Maximum.}$$

Daraus folgt, dass g in 0 ein isoliertes lokales Maximum besitzt.

Beispiel 18.5 (Parametrisierte Polynome)**.** es seien $a_0, \dots, a_{m-1} \in C^k(\Omega, \mathbb{R})$, $\Omega \subset \mathbb{R}^n$, $x \in \Omega$ und

$$P\colon \Omega \times \mathbb{R} \to \mathbb{R}, \quad P(x,y) := y^m + \sum_{j=0}^{m-1} a_j(x)\cdot y^j.$$

Ist $y_0 \in \mathbb{R}$ eine einfache Nullstelle von $P(x_0, \cdot)$, d.h. es gilt

$$P(x_0, y_0) = 0 \quad \text{und} \quad \mathrm{D}_y P(x_0, y_0) \neq 0,$$

so ist der Satz 18.2 anwendbar und es existieren offene $W \subset \mathbb{R}^n$ und $I \subset \mathbb{R}$, mit $x_0 \in W$ und $y_0 \in W$, und $g \in C^k(W, I)$ mit $y = g(x)$ als einzige Lösung von

$$P(x,y) = 0 \quad \text{in} \quad \Big\{(x,y) \in \mathbb{R}^{n+1} \;:\; x \in W, y \in I\Big\} = W \times I.$$

LINK: TEIL 1 DER 25. VORLESUNG VOM 14.07.2022

19 Extrema mit Nebenbedingungen

Beispiel 19.1. Man betrachte die stetige Funktion

$$\varphi\colon \mathbb{R}^2 \to \mathbb{R}, \quad \varphi(x,y) = x \cdot y.$$

Nach Korollar 11.29 nimmt φ auf der beschränkten und abgeschlossenen (also kompakten) Menge

$$K := \Big\{(x,y) \in \mathbb{R}^2 \ : \ x^2 + y^2 \leq 2\Big\}$$

sein Minimum und sein Maximum an. Wie berechnen wir aber diese Werte? Liegt in (x_0, y_0) ein Extremum von *phi* im offenen Inneren

$$K^\circ = \Big\{(x,y) \in \mathbb{R}^2 \ : \ x^2 + y^2 < 2\Big\},$$

so muss dort eine kritische Stelle vorliegen, d.h. es gilt

$$\begin{pmatrix} 0 & 0 \end{pmatrix} = (\mathrm{D}\varphi)(x_0, y_0) = \begin{pmatrix} y_0 & x_0 \end{pmatrix}.$$

Aber offensichtlich ist $(0,0)$ keine Extremalstelle von φ, wie wir in Beispiel 17.4 gezeigt haben. Also liegen alle Minimal- und Maximalstellen von $\varphi|_K\colon K \to \mathbb{R}$ auf der Teilmenge

$$M := \partial K = \Big\{(x,y) \in \mathbb{R}^2 \ : \ x^2 + y^2 - 2 = 0\Big\} \subset K,$$

d.h. auf dem Rand von K. Es sind also die Minima und Maxima der Funktion

$$\varphi|_M\colon M \to \mathbb{R}$$

zu bestimmen, d.h. die Extrema von φ unter der *Nebenbedingung*

$$f(x,y) := x^2 + y^2 - 2 = 0.$$

Hier sind wir mit zwei Variablen x und y gestartet. Die Nebenbedingung reduziert die Freiheitsgrade um Eins, so dass

$$\partial K = \left\{ \begin{pmatrix} 2\cos(\varphi) \\ 2\sin(\varphi) \end{pmatrix} \ : \ \varphi \in [0, 2\pi) \right\}$$

ein eindimensionales Gebilde (genauer: eine Mannigfaltigkeit) ist.

Bei Funktionen $\varphi\colon \mathbb{R}^n \to \mathbb{R}$ kann man m Nebenbedingungen stellen, wobei $m \leq n-1$ sein soll, damit die resultierende Menge $n - m \geq 1$ Dimensionen hat und man über diese(n) freie(n) Parameter noch optimieren kann. Betrachten wir dazu wieder $\varphi\colon \mathbb{R}^2 \to \mathbb{R}, \quad \varphi(x,y) := xy$ wie oben und fügen eine weitere Nebenbedingung hinzu, haben also insgesamt die Nebenbedingungen

$$f(x,y) = x^2 + y^2 - 2 = 0 \quad \text{und} \quad F(x,y) = x - y = 0. \tag{$\times$}$$

Hier würde nur die Punkte $(x,y) = (x,x) = (1,1)$ und $(x,y) = (-1,-1)$ den Nebenbedingungen genügen. Um dann Maxima und Minima unter der Nebenbedingung ($\times$) zu finden, müsste man φ nur an diesen Stellen auswerten: Es gilt

$$\varphi(1,1) = 1 \quad \text{und} \quad \varphi(-1,-1) = 1.$$

Also ist φ auf den „erlaubten“ Punkten konstant.

Satz 19.2 (Lagrange-Multiplikatoren)**.** *Es sei* $\Omega \subset \mathbb{R}^n$ *offen,* $\varphi \in C^1(\Omega, \mathbb{R})$ *und* $f = (f_1, \dots, f_m) \in C^1(\Omega, \mathbb{R}^m)$ *mit* $m < n$*. Es sei außerdem*

$$M := \Big\{x \in \Omega \ : \ f(x) = 0\Big\},$$

$x_0 \in M$*,* $\mathrm{Rang}\Big(J_f(x_0)\Big) = m$ *und* $\varphi|_M \colon M \to \mathbb{R}$ *haben ein lokales Extremum in* x_0*. Dann existieren* $\lambda_1, \dots, \lambda_m \in \mathbb{R}$ *mit*

$$\mathrm{D}\varphi(x_0) = \sum_{j=1}^{m} \lambda_j \cdot \mathrm{D}f_j(x_0).$$

Diese $\lambda_1, \dots, \lambda_m$ *heißen* ***Lagrange-Multiplikatoren****.*

Beweis. Wir beschränken uns hier auf den Fall einer einzelnen Nebenbedingung, d.h. $m = 1$. Dann bedeutet $\mathrm{Rang}\Big(J_f(x_0)\Big) = m = 1$ schlichtweg, dass

$$\mathrm{grad}\, f(x_0) \neq \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix},$$

d.h. es gibt mindestens einen Koeffizienten $\mathrm{D}_j f(x_0) \neq 0$. Nach Umnummerierung können wir annehmen, dass $\mathrm{D}_n f(x_0) \neq 0$ gilt, insbesondere ist

$$\lambda := \frac{\mathrm{D}_n \varphi(x_0)}{\mathrm{D}_n f(x_0)} \in \mathbb{R} \tag{$\times\times$}$$

wohldefiniert. Da $x_0 =: (\overbrace{a_1, \dots, a_{n-1}}^{=a \in \mathbb{R}^{n-1}}, b) \in M$, gilt

$$f(a, b) = f(x_0) = 0.$$

Damit ist der Satz 18.2 über implizite Funktionen anwendbar und man kann die Gleichung

$$f(x_1, \dots, x_{n-1}, y) = 0$$

lokal nach y auflösen. Es gibt also ein $r > 0$ und eine eindeutige Funktion

$$g \colon V := \underbrace{B_r(a)}_{\subset \mathbb{R}^{n-1}} \to I := (b - r, b + r),$$

so dass Folgendes gilt:

$$g(a) = b, \quad f\Big(x_1, \dots, x_{n-1}, g(x_1, \dots, x_{n-1})\Big) = 0 \quad \forall\, (x_1, \dots, x_{n-1}) \in V$$

und für $j \in \{1, 2, \dots, n-1\}$ ist

$$\mathrm{D}_j f(a, b) + \mathrm{D}_n f(a, b) \cdot \mathrm{D}_j g(a) = 0. \tag{$\times\times\times$}$$

Vgl. Formel $(*)$ in Satz 18.2 und speziell Bemerkung 18.3 (a).

Achtung: Dort ist die letzte Koordinate $n+1$, wobei es hier nur n ist!

LINK: TEIL 2 DER 25. VORLESUNG VOM 14.07.2022

Nun definieren wir eine Funktion

$$h\colon V \to \mathbb{R}, \quad h(x_1,\ldots,x_{n-1}) := \varphi\Big(x_1,\ldots,x_{n-1},g(x_1,\ldots,x_{n-1})\Big).$$

Da φ in $x_0 := (a,b)$ ein lokales Extremum besitzt (z.B. $\varphi(x_0) \geq \varphi(x)$ für $\|x_0 - x\|$ klein), hat h in $a \in V$ ein lokales Extremum. Dann ist nämlich

$$\begin{aligned} h(a_1,\ldots,a_{n-1}) = h(a) = \varphi\Big(a,g(a)\Big) &= \varphi(x_0) \\ &\geq \varphi\Big(x_1,\ldots,x_{n-1},g(x_1,\ldots,x_{n-1})\Big) \\ &= h(x_1,\ldots,x_{n-1}). \end{aligned}$$

Also gilt für jedes $j \in \{1,\ldots,n-1\}$ für die partielle Ableitung:

$$\begin{aligned} 0 = \mathrm{D}_j h(a) &= \mathrm{D}_j\varphi\Big(a,g(a)\Big) + \mathrm{D}_n\varphi\Big(a,g(a)\Big) \cdot \mathrm{D}_j g(a) \\ &= \mathrm{D}_j\varphi(x_0) + \mathrm{D}_n\varphi(x_0) \cdot (\mathrm{D}_j g)(a) \\ &\overset{(\times\times)}{=} \mathrm{D}_j\varphi(x_0) + \lambda \cdot \mathrm{D}_n f(x_0) \cdot \mathrm{D}_j g(a) \\ &\overset{(\times\times\times)}{=} \mathrm{D}_j\varphi(x_0) - \lambda \cdot \mathrm{D}_j f(x_0). \end{aligned}$$

Also gilt

$$\mathrm{D}_j\varphi(x_0) = \lambda \cdot \mathrm{D}_j f(x_0) \quad \text{für } j \in \{1,\ldots,n-1\}.$$

Für $j = n$ folgt diese Gleichung aus $(\times\times)$. $\square$

Beispiel 19.3. Wir greifen auf die Funktion

$$\varphi\colon \mathbb{R}^2 \to \mathbb{R}, \quad (x,y) \mapsto x \cdot y,$$

zurück. Dazu suchen wir das Maximum und Minimum von $\varphi\|_M\colon M \to \mathbb{R}$, wobei

$$M := \Big\{(x,y) \in \mathbb{R}^2 \;:\; x^2 + y^2 - 2 = 0\Big\},$$

wir haben also die Nebenbedingung

$$f(x, y) = x^2 + y^2 - 2 = 0$$

vorliegen. Für $(x, y) \in M$ gilt

$$\mathrm{D}f(x, y) = \begin{pmatrix} 2x & 2y \end{pmatrix} \neq \begin{pmatrix} 0 & 0 \end{pmatrix}.$$

Also ist Rang $\left(\mathrm{D}f(x, y)\right) = 1 = m$. Ist $(x, y) \in M$ eine Extremalstelle, besagt der Satz 19.2, dass ein $\lambda \in \mathbb{R}$ existiert mit

$$\mathrm{D}\varphi(x, y) = \lambda \cdot \mathrm{D}f(x, y),$$

das heißt

$$\begin{pmatrix} y & x \end{pmatrix} = \lambda \cdot \begin{pmatrix} 2x & 2y \end{pmatrix}. \qquad (\blacksquare)$$

Ist $x = 0$, so ist auch $y = 0$, was der Gleichung $x^2 + y^2 = 2$, also $(x, y) \in M$ widerspricht. Also muss $x \neq 0$ und damit auch $y \neq 0$ gelten. Division in der Gleichung ($\blacksquare$) liefert dann

$$\frac{y}{x} = \frac{x}{y}$$

$$\Leftrightarrow \quad y^2 = x^2$$

$$\Leftrightarrow \quad y = \pm x.$$

Also gilt

$$2 = x^2 + y^2 = x^2 + x^2 = 2x^2 \quad \Rightarrow \quad x^2 = 1.$$

Wir haben also vier mögliche Extremalstellen

$$(1, 1), \quad (1, -1), \quad (-1, 1) \quad \text{und} \quad (-1, -1).$$

Auswertung in φ liefert dann:

$$\varphi(1, 1) = 1, \quad \varphi(1, -1) = -1, \quad \varphi(-1, 1) = -1 \quad \text{und} \quad \varphi(-1, -1) = 1.$$

Also sind $(1, 1)$ und $(-1, -1)$ globale Maximalstellen von $\varphi|_M$ und $(1, -1)$ sowie $(-1, 1)$ globale Minimalstellen von $\varphi|_M$.

LINK: MINDMAP ANALYSIS II

In der letzten Vorlesungsstunde habe ich an der Tafel ein ‚Mindmap‘ mit wichtigen Inhalten der Analysis II-Vorlesung erstellt. Sie sollte den Hörern als Wiederholung und ‚Initialzündung‘ für ihre Lernphase dienen.

Mind map Ana II

Um formulierung von Folgen konverg.

⟺ formal

(absolute) konvergenz von Reihen insbes. " kriterien

⟺ intuitiv

Analogie zu Integralen

Folgen von Fkt

↓

konvergenz begriffe

Achtung:

Im $\mathbb{R}^n$ sind alle Normen äquivalent hier nicht

Funktionen reihen. insb.
Potenzreihen
Taylorreihen
T-Formel, Rest-Darst.

⟷

Extremalwerte.

viele bekannte Fkt von dieser Form:
Polynome, Exp-Reihe
Geom. Reihe.
sinus, cosinus

asymptotisches Verhalten leichter zu untersuchen als z.B. mit L'Hospital.

abstrakte ↕

Abstands d

konvergenz begriffe

Metrik, Norm,

geometrische Begri.

Konv. v. Folgen.

Wegzushang

wichtige Spezialfälle:

im $\mathbb{R}^n$, $\mathbb{C}^n$

... $\Leftrightarrow$ konvergenz d. koeffizienten.

$$\frac{r(h)}{\|h\|} = \left\| \frac{f(a+h)-f(a)}{\|h\|} + L\left(\frac{h}{\|h\|}\right)\right\|$$

(schöne) koordinatentransform.

↕

part und totale Abl in $\mathbb{R}^d$ → parameterabh. Int. (bei uns vor A. zum Bew. / ist aber vom unabh. Interesse)

↕

höhere Abl., z.B. S.v. Schwarz, Hesse-M.

$$\frac{r(h)}{\|h\|} = \left\| \frac{f(a+h)-f(a)+L(h)}{\|h\|}\right\| \to 0$$

$\downarrow$ $Df(a)(h)$

$$\boxed{\begin{array}{l}\longrightarrow \text{Dfn. der Stetigk.} \\ \text{auf metr. Räumen.}\end{array}}$$

dieser Block schräg nach oben!

Analog zu Ana I untersuchen / bw. wir mit diesen Hilfsmitteln:

MWS, Taylorformel (2. Ord. qualit. Restterm)

(Lokale) Extrema

mit Nebenbed.

implizite Fkt ⟷

Anwendung z.B.:
in der Realität habe ich manchmal viele Infos über Fkt aber keine Formel → Untersuchen!