Prof. Dr. Jens Wirth
Institut für Analysis, Dynamik und Modellierung
Universität Stuttgart

Analysis 2

©2022/23 Jens Wirth

Inhaltsverzeichnis

5 Integralrechnung in einer Variablen **7**
5.1 Das Riemannintegral 7
5.2 Hauptsätze der Integralrechnung 16
5.3 Berechnung von Integralen 20
5.4 Die Taylorsche Formel mit Integralrestglied 28
5.5 Uneigentliche Riemannintegrale 30

6 Funktionenfolgen und Reihen **37**
6.1 Doppelfolgen und -Reihen 37
6.2 Gleichmäßige Konvergenz und Stetigkeit 40
6.3 Integration und Differentiation 42
6.4 Potenzreihen 45
6.5 Fourierreihen 50

7 Differentialrechnung in mehreren Veränderlichen **57**
7.1 Vektorräume und lineare Abbildungen 57
7.2 Zur topologischen Struktur metrischer und normierter Räume 61
7.3 Differenzierbarkeit 68
7.4 Hauptsätze über differenzierbare Funktionen 76
7.5 Kurven im Raum 89
7.6 Skalare Funktionen, Differentiale und Integrale 98
7.7 Höhere Ableitungen und Taylorsche Formel 107

A Anhang **119**
A.1 Koordinatensysteme 119

Literatur (-empfehlungen)

[1a] Konrad Königsberger, *Analysis 1*, Springer-Verlag, 2004
https://doi.org/10.1007/978-3-642-18490-1

[1b] Konrad Königsberger, *Analysis 2*, Springer-Verlag, 2004
https://doi.org/10.1007/3-540-35077-2

[2a] Vladimir A. Zorich, *Analysis I*, Springer-Verlag, 2015
https://doi.org/10.1007/978-3-662-48792-1

[2b] Vladimir A. Zorich, *Analysis II*, Springer-Verlag, 2016
https://doi.org/10.1007/978-3-662-48993-2

[3a] Harro Heuser, *Lehrbuch der Analysis*, Band 1, Springer-Verlag, 2003
https://doi.org/10.1007/978-3-322-96828-9

[3b] Harro Heuser, *Lehrbuch der Analysis*, Band 2, Springer-Verlag, 2003
https://doi.org/10.1007/978-3-322-96826-5

[4a] Grigori M. Fichtenholz, *Differential- und Integralrechnung*, Band 1, Verlag Harri Deutsch, 2006

[4b] Grigori M. Fichtenholz, *Differential- und Integralrechnung*, Band 2, Verlag Harri Deutsch, 2009

[4c] Grigori M. Fichtenholz, *Differential- und Integralrechnung*, Band 3, Verlag Harri Deutsch, 2011

[5] Tom Apostol, *Mathematical Analysis*, Addison-Wesley, 1974

[6] Wolfgang Rudin, *Analysis*, Oldenbourg, 2009

[7] Pólya, Szegő, *Aufgaben und Lehrsätze aus der Analysis I*, Springer-Verlag, 1998
https://doi.org/10.1007/978-3-642-61983-0

[8a] Hans von Mangoldt, Konrad Knopp, *Einführung in die Höhere Mathematik*, Band 1, S. Hirzel Verlag, 1963

[8b] Hans von Mangoldt, Konrad Knopp, *Einführung in die Höhere Mathematik*, Band 2, S. Hirzel Verlag, 1965

[9a] Richard Courant, *Vorlesungen über Differential- und Integralrechnung*, Band 1, Springer-Verlag, 1971
https://doi.org/10.1007/978-3-642-61988-5

[9b] Richard Courant, *Vorlesungen über Differential- und Integralrechnung*, Band 2, Springer-Verlag, 1955
https://doi.org/10.1007/978-3-642-61973-1

[10] Henri Cartan, *Differentialrechnung*, BI Wissenschaftsverlag, 1974

Zeitplan Analysis 2

Woche	Themen	
1	Riemannintegral	5.1
2	Hauptsätze, Integralberechnung	5.2, 5.3
3	Integralrestglied, uneigentliche Riemannintegrale	5.4, 5.5
4	Vertauschen von Grenzwerten, gleichmäßige Konvergenz	6.1, 6.2
5	gliedweises Differenzieren, Potenzreihen	6.3, 6.4
6	Fourierreihen	6.5
7	Topologische Struktur des $\mathbb{R}^n$	7.2
8	Differenzierbarkeit	7.3
9,10	Hauptsätze über differenzierbare Funktionen	7.4
11	Kurven und Rektifizierbarkeit	7.5
12	Differentiale und Kurvenintegrale	7.6
13	Hessematrix, Taylorapproximation	7.7
14	Extremwertaufgaben	7.7

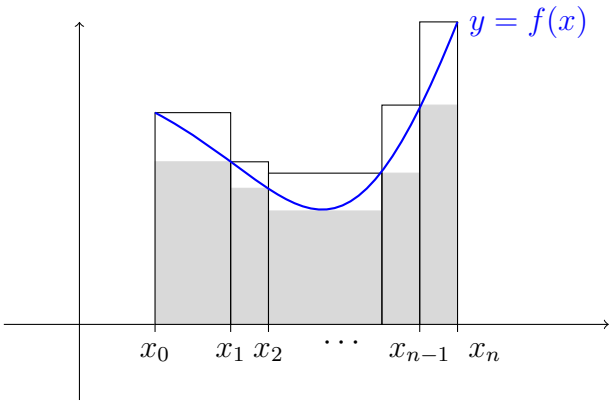
5 Integralrechnung in einer Variablen

5.1 Das Riemannintegral

5.1.1. Unser erstes Ziel ist es, der Menge

$$\{(x, y) \in [a, b] \times \mathbb{R} \mid 0 \leq y \leq f(x)\} \tag{5.1.1-A}$$

unter dem Graphen einer nichtnegativen Funktion $f : [a, b] \to \mathbb{R}_{\geq 0}$ einen Flächeninhalt zuzuordnen. Dies wird nicht für jede Funktion möglich sein, wir benötigen also insbesondere Kriterien für die Zulässigkeit von Funktionen.



Wir skizzieren zuerst die Grundidee. Wir zerlegen das Intervall $[a, b]$ in Teilintervalle und approximieren den Flächeninhalt unter der Funktion von Unten durch die größten Rechtecke, die wir passend zur Zerlegung des Intervalls unter dem Graphen der Funktion unterbringen, und nach oben entsprechend durch kleinsten Rechtecke oberhalb des Graphen. Strebt bei feiner werdenden Zerlegungen jede solche untere und jede obere Approximation gegen denselben Wert, so werden wir diesen als Riemann-Integral bezeichnen.

Es gibt keinen Grund die Konstruktion auf nichtnegative Funktionen zu beschränken, wir zählen Rechtecke unterhalb der reellen Achse im folgenden negativ.

5.1.2 (Zerlegungen)**.** Eine *Zerlegung* $\mathfrak{Z}$ des Intervalls $[a, b]$ ist ein Tupel

$$a = x_0 < x_1 < \ldots < x_{n-1} < x_n = b \tag{5.1.2-A}$$

von geordneten Zwischenstellen $x_1, \ldots, x_{n-1}$. Dabei bezeichnet

$$\delta(\mathfrak{Z}) = \max_{k=1,\ldots,n} |x_k - x_{k-1}| \tag{5.1.2-B}$$

die *Feinheit* der Zerlegung $\mathfrak{Z}$.

Für zwei Zerlegungen $\mathfrak{Z}_1$ und $\mathfrak{Z}_2$ sagen wir, $\mathfrak{Z}_1$ ist eine *Verfeinerung* von $\mathfrak{Z}_2$, wenn $\mathfrak{Z}_1$ alle Zwischenstellen von $\mathfrak{Z}_2$ enthält, in Zeichen[1] $\mathfrak{Z}_2 \subset \mathfrak{Z}_1$. Vereinigt man die Menge der Zwischenstellen zweier Zerlegungen $\mathfrak{Z}_1$ und $\mathfrak{Z}_2$, so nennt man die sich ergebende Zerlegung die *gemeinsame Verfeinerung* $\mathfrak{Z}_1 \cup \mathfrak{Z}_2$. Dabei gilt offenbar $\delta(\mathfrak{Z}_1 \cup \mathfrak{Z}_2) \leq \min\{\delta(\mathfrak{Z}_1), \delta(\mathfrak{Z}_2)\}$.

5.1.3 (Darbouxsche Unter- und Obersummen)**.** Sei nun $f : [a,b] \to \mathbb{R}$ eine beschränkte Funktion und $\mathfrak{Z} = (x_0, \ldots, x_n)$ eine Zerlegung des Intervalls $[a,b]$. Dann definieren wir die *Darbouxsche Untersumme*[2]

$$\mathcal{U}(f;\mathfrak{Z}) = \sum_{k=1}^{n} (x_k - x_{k-1}) \inf_{\xi \in [x_{k-1}, x_k]} f(\xi) \tag{5.1.3-A}$$

und analog die *Darbouxsche Obersumme*

$$\mathcal{O}(f;\mathfrak{Z}) = \sum_{k=1}^{n} (x_k - x_{k-1}) \sup_{\xi \in [x_{k-1}, x_k]} f(\xi). \tag{5.1.3-B}$$

Wir zeigen zuerst einige elementare Eigenschaften dieser Summen.

☞ **5.1.4 Lemma.** **(i)** *Für jede Zerlegung* $\mathfrak{Z}$ *gilt*

$$\mathcal{U}(f;\mathfrak{Z}) \leq \mathcal{O}(f;\mathfrak{Z}). \tag{5.1.4-A}$$

(ii) *Ist* $\mathfrak{Z}_1$ *Verfeinerung von* $\mathfrak{Z}_2$*, so gilt*

$$\mathcal{U}(f;\mathfrak{Z}_1) \geq \mathcal{U}(f;\mathfrak{Z}_2) \qquad \textit{und} \qquad \mathcal{O}(f;\mathfrak{Z}_1) \leq \mathcal{O}(f;\mathfrak{Z}_2). \tag{5.1.4-B}$$

(iii) *Für zwei Zerlegungen* $\mathfrak{Z}_1$ *und* $\mathfrak{Z}_2$ *gilt stets*

$$\mathcal{U}(f;\mathfrak{Z}_1) \leq \mathcal{O}(f;\mathfrak{Z}_2). \tag{5.1.4-C}$$

Beweis. **(i)** und **(ii)** folgen direkt aus der Definition. Für **(iii)** nutzen wir die gemeinsame Verfeinerung $\mathfrak{Z} = \mathfrak{Z}_1 \cup \mathfrak{Z}_2$ und wenden **(ii)** an, weswegen

$$\mathcal{U}(f;\mathfrak{Z}_1) \leq \mathcal{U}(f;\mathfrak{Z}) \leq \mathcal{O}(f;\mathfrak{Z}) \leq \mathcal{O}(f;\mathfrak{Z}_2) \tag{5.1.4-D}$$

gilt. □

✍ **5.1.5 Definition** (Integrierbarkeit)**.** **(i)** *Eine beschränkte Funktion* $f : [a,b] \to \mathbb{R}$ *bezeichnen wir als* R-integrierbar*, falls*

$$\sup_{\mathfrak{Z}} \mathcal{U}(f;\mathfrak{Z}) = \inf_{\mathfrak{Z}} \mathcal{O}(f;\mathfrak{Z}) \tag{5.1.5-A}$$

gilt. Bezeichne weiter $\mathfrak{R}[a,b]$ *die Menge der auf dem Intervall* $[a,b]$ *R-integrierbaren Funktionen.*

[1] Wir verwenden Mengenschreibweise. Auch wenn wir von Tupeln sprechen ist die Reihenfolge der Zwischenstellen durch Ordnung auf $\mathbb{R}$ gegeben und die wirkliche Information in $\mathfrak{Z}$ ist die zugrundeliegende Menge der Zwischenstellen.

[2] Gaston Darboux, 1842–1917

(ii) *Für $f \in \mathfrak{R}[a,b]$ sei*

$$\int_a^b f(x)\,\mathrm{d}x := \sup_{\mathfrak{Z}} \mathcal{U}(f;\mathfrak{Z}) = \inf_{\mathfrak{Z}} \mathcal{O}(f;\mathfrak{Z}) \tag{5.1.5-B}$$

das Riemannintegral *von f über $[a,b]$.*

⚒ **5.1.6 Beispiele.** **(i)** Die Funktion $f(x) = x$ ist auf dem Intervall $[0,1]$ R-integrierbar. Dazu schreiben wir Unter- und Obersummen einer Zerlegung $\mathfrak{Z} = (x_0,\dots,x_n)$ explizit auf. Es gilt

$$\mathcal{O}(f;\mathfrak{Z}) = \sum_{k=1}^{n} (x_k - x_{k-1}) x_k \tag{5.1.6-A}$$

und

$$\mathcal{U}(f;\mathfrak{Z}) = \sum_{k=1}^{n} (x_k - x_{k-1}) x_{k-1} \tag{5.1.6-B}$$

und damit

$$\begin{aligned} \mathcal{O}(f,\mathfrak{Z}) - \mathcal{U}(f,\mathfrak{Z}) &= \sum_{k=1}^{n} (x_k - x_{k-1})(x_k - x_{k-1}) \\ &\le \delta(\mathfrak{Z}) \sum_{k=1}^{n} (x_k - x_{k-1}) = \delta(\mathfrak{Z}), \end{aligned} \tag{5.1.6-C}$$

auf Grund von $\sum_{k=1}^{n}(x_k - x_{k-1}) = x_n - x_0 = 1$ gilt. Für $\delta(\mathfrak{Z}) < \varepsilon$ gilt also $\mathcal{O}(f,\mathfrak{Z}) - \mathcal{U}(f,\mathfrak{Z}) < \varepsilon$ und Integrierbarkeit folgt.

(ii) Weitere Beispiele R-integrierbarer Funktionen werden wir später noch liefern, vorerst ein Nicht-Beispiel. Die Dirichletfunktion $f : [0,1] \to \mathbb{R}$ mit

$$f(x) = \begin{cases} 1, & x \in \mathbb{Q}, \\ 0, & x \notin \mathbb{Q}, \end{cases} \tag{5.1.6-D}$$

ist nicht R-integrierbar. Jede Obersumme liefert den Wert $\mathcal{O}(f;\mathfrak{Z}) = 1$, während jede Untersumme $\mathcal{U}(f;\mathfrak{Z}) = 0$ liefert.

☞ **5.1.7 Satz** (Eigenschaften des Riemannintegrals). **(i)** *(Beschränktheit) Sei $f \in \mathfrak{R}[a,b]$. Dann gilt*

$$\inf_{\xi\in[a,b]} f(\xi) \le \frac{1}{b-a}\int_a^b f(x)\,\mathrm{d}x \le \sup_{\xi\in[a,b]} f(\xi). \tag{5.1.7-A}$$

(ii) *(Linearität) Seien $f, g \in \mathfrak{R}[a,b]$ und $\alpha \in \mathbb{R}$. Dann gilt $f + \alpha g \in \mathfrak{R}[a,b]$ und*

$$\int_a^b f(x) + \alpha g(x)\,\mathrm{d}x = \int_a^b f(x)\,\mathrm{d}x + \alpha \int_a^b g(x)\,\mathrm{d}x. \tag{5.1.7-B}$$

(iii) *(Monotonie) Seien $f, g \in \mathfrak{R}[a,b]$ mit $f(x) \le g(x)$ für $x \in [a,b]$. Dann gilt*

$$\int_a^b f(x)\,\mathrm{d}x \le \int_a^b g(x)\,\mathrm{d}x \tag{5.1.7-C}$$

(iv) *(Zerlegung des Integrals) Sei $f \in \mathfrak{R}[a,b]$ und sei $c \in (a,b)$. Dann gilt für die Einschränkung von f auf Teilintervalle $f \in \mathfrak{R}[a,c]$ und $f \in \mathfrak{R}[c,b]$ sowie*

$$\int_a^b f(x)\,\mathrm{d}x = \int_a^c f(x)\,\mathrm{d}x + \int_c^b f(x)\,\mathrm{d}x. \tag{5.1.7-D}$$

Gilt umgekehrt $f \in \mathfrak{R}[a,c]$ und $f \in \mathfrak{R}[c,b]$, so folgt $f \in \mathfrak{R}[a,b]$.

Beweis. **(i)** Es gilt mit der trivialen Zerlegung $\mathfrak{Z} = (a,b)$

$$(b-a)\inf_{\xi\in[a,b]} f(\xi) = \mathcal{U}(f;\mathfrak{Z}) \le \int_a^b f(x)\,\mathrm{d}x \le \mathcal{O}(f;\mathfrak{Z}) = (b-a)\sup_{\xi\in[a,b]} f(\xi) \tag{5.1.7-E}$$

und damit die Behauptung.

(ii) Wir zeigen Summen und Vielfache getrennt. Es gilt für die Obersummen einer Zerlegung $\mathfrak{Z} = (x_0, x_1, \ldots, x_n)$

$$\begin{aligned}\mathcal{O}(f+g;\mathfrak{Z}) &= \sum_{k=1}^n (x_k - x_{k-1}) \sup_{\xi\in[x_{k-1},x_k]} (f(\xi)+g(\xi)) \\ &\le \sum_{k=1}^n (x_k - x_{k-1}) \left(\sup_{\xi\in[x_{k-1},x_k]} f(\xi) + \sup_{\eta\in[x_{k-1},x_k]} g(\eta) \right) \\ &= \mathcal{O}(f;\mathfrak{Z}) + \mathcal{O}(g;\mathfrak{Z}).\end{aligned} \tag{5.1.7-F}$$

und entsprechend $\mathcal{U}(f+g;\mathfrak{Z}) \ge \mathcal{U}(f;\mathfrak{Z}) + \mathcal{U}(g;\mathfrak{Z})$ für die Untersummen. Also gilt

$$\mathcal{O}(f+g;\mathfrak{Z}) - \mathcal{U}(f+g;\mathfrak{Z}) \le \underbrace{\mathcal{O}(f;\mathfrak{Z}) - \mathcal{U}(f;\mathfrak{Z})}_{\longrightarrow 0} + \underbrace{\mathcal{O}(g;\mathfrak{Z}) - \mathcal{U}(g;\mathfrak{Z})}_{\longrightarrow 0} \tag{5.1.7-G}$$

und aus $f, g \in \mathfrak{R}[a,b]$ folgt $f+g \in \mathfrak{R}[a,b]$ und die Addititvität des Riemannintegrals gilt. Ist weiter $\alpha \ge 0$, so gilt

$$\mathcal{O}(\alpha f;\mathfrak{Z}) = \sum_{k=1}^n (x_k - x_{k-1}) \sup_{\xi\in[x_{k-1},x_k]} (\alpha f(\xi)) = \alpha\,\mathcal{O}(f;\mathfrak{Z}) \tag{5.1.7-H}$$

und entsprechend $\mathcal{U}(\alpha f;\mathfrak{Z}) = \alpha\mathcal{U}(f;\mathfrak{Z})$, während sich bei $\alpha < 0$ die Rollen von $\mathcal{U}$ und $\mathcal{O}$ vertauschen. Damit folgt die Homogenität des Riemannintegrals.

(iii) Aus $f(x) \le g(x)$ auf $[a,b]$ folgt die entsprechend Beziehung auf jedem Teilintervall einer Zerlegung $\mathfrak{Z}$ und damit

$$\mathcal{O}(f;\mathfrak{Z}) \le \mathcal{O}(g;\mathfrak{Z}), \qquad \mathcal{U}(f;\mathfrak{Z}) \le \mathcal{U}(g;\mathfrak{Z}) \tag{5.1.7-I}$$

und da beide Riemannintegrale existieren folgt nach Grenzwertbildung die Ungleichung zwischen den Integralen.

(iv) Jede Zerlegung $\mathfrak{Z}$ von $[a,b]$ kann durch Hinzunahme des Punktes c verfeinert werden. Sei $\mathfrak{Z}'$ diese Verfeinerung und bezeichne $\mathfrak{Z}'_1$ und $\mathfrak{Z}'_2$ die darin enthaltenen

Zerlegungen von $[a,c]$ und $[c,b]$. Jede Zerlegung von $[a,c]$ beziehungsweise von $[c,b]$ kann auf diese Weise erhalten werden.

Dann gilt für jedes $f \in \mathfrak{R}[a,b]$

$$\mathcal{U}(f;\mathfrak{Z}) \le \mathcal{U}(f;\mathfrak{Z}') = \mathcal{U}(f;\mathfrak{Z}'_1) + \mathcal{U}(f;\mathfrak{Z}'_2) \tag{5.1.7-J}$$

und

$$\mathcal{O}(f;\mathfrak{Z}'_1) + \mathcal{O}(f;\mathfrak{Z}'_2) = \mathcal{O}(f;\mathfrak{Z}') \le \mathcal{O}(f;\mathfrak{Z}) \tag{5.1.7-K}$$

und damit

$$\mathcal{O}(f;\mathfrak{Z}'_1) - \mathcal{U}(f;\mathfrak{Z}'_1) \le \mathcal{O}(f;\mathfrak{Z}) - \mathcal{U}(f;\mathfrak{Z}) \longrightarrow 0 \tag{5.1.7-L}$$

und analog für $\mathfrak{Z}_2$. Also folgt $f \in \mathfrak{R}[a,c]$ und $f \in \mathfrak{R}[c,b]$ und

$$\int_a^b f(x)\,\mathrm{d}x = \int_a^c f(x)\,\mathrm{d}x + \int_c^b f(x)\,\mathrm{d}x. \tag{5.1.7-M}$$

Ist umgekehrt $f \in \mathfrak{R}[a,c]$ und $\mathfrak{R}[c,b]$, so folgt aus

$$\mathcal{O}(f;\mathfrak{Z}') - \mathcal{U}(f;\mathfrak{Z}') = \underbrace{\mathcal{O}(f;\mathfrak{Z}'_1) - \mathcal{U}(f;\mathfrak{Z}'_1)}_{\longrightarrow 0} + \underbrace{\mathcal{O}(f;\mathfrak{Z}'_2) - \mathcal{U}(f;\mathfrak{Z}'_2)}_{\longrightarrow 0} \tag{5.1.7-N}$$

schon $f \in \mathfrak{R}[a,b]$.

□

Gerade wegen der letzten Eigenschaft ist es sehr sinnvoll, das Riemannintegral auch für Intervallgrenzen in falscher Reihenfolge vermittels

$$\int_b^a f(x)\,\mathrm{d}x = -\int_a^b f(x)\,\mathrm{d}x, \qquad f \in \mathfrak{R}[a,b], \tag{5.1.7-O}$$

zu definieren.

☞ **5.1.8 Satz** (Integrierbarkeitskriterium)**.** *Sei $f : [a,b] \to \mathbb{R}$. Dann sind die folgenden Aussagen äquivalent:*

- **(i)** *$f \in \mathfrak{R}[a,b]$;*
- **(ii)** *zu jedem $\varepsilon > 0$ existiert eine Zerlegung $\mathfrak{Z}$ mit $0 \le \mathcal{O}(f;\mathfrak{Z}) - \mathcal{U}(f;\mathfrak{Z}) < \varepsilon$;*
- **(iii)** *zu jedem $\varepsilon > 0$ existiert ein δ mit*

$$0 \le \mathcal{O}(f;\mathfrak{Z}) - \mathcal{U}(f;\mathfrak{Z}) < \varepsilon \tag{5.1.8-A}$$

für alle $\mathfrak{Z}$ mit $\delta(\mathfrak{Z}) < \delta$.

Beweis. Nach obiger Definition gilt **(i)** ⇔ **(ii)**. Zu zeigen bleibt die Äquivalenz der letzten beiden Aussagen und damit insbesondere **(ii)** ⇒ **(iii)**. Sei dazu $\mathfrak{Z}_1$ eine Zerlegung mit $\mathcal{O}(f;\mathfrak{Z}_1) - \mathcal{U}(f;\mathfrak{Z}_1) < \frac{\varepsilon}{3}$. Bezeichne weiter n die Anzahl der Teilintervalle der Zerlegung $\mathfrak{Z}_1$ und

$$\Omega = \sup_{\xi\in[a,b]} f(\xi) - \inf_{\xi\in[a,b]} f(\xi) \tag{5.1.8-B}$$

die *Oszillation* von f über $[a,b]$.

Sei nun $\mathfrak{Z}$ eine Zerlegung mit $\delta(\mathfrak{Z}) < \delta := \frac{\varepsilon}{3n\Omega}$. Mit $\mathfrak{Z}_2 = \mathfrak{Z} \cup \mathfrak{Z}_1$ der gemeinsamen Verfeinerung von $\mathfrak{Z}$ und $\mathfrak{Z}_1$ gilt

$$\begin{aligned}\mathcal{O}(f;\mathfrak{Z}) - \mathcal{U}(f;\mathfrak{Z}) \le{}& \big(\mathcal{O}(f;\mathfrak{Z}) - \mathcal{O}(f;\mathfrak{Z}_2)\big)\\ &+ \big(\mathcal{O}(f;\mathfrak{Z}_2) - \mathcal{U}(f;\mathfrak{Z}_2)\big)\\ &+ \big(\mathcal{U}(f;\mathfrak{Z}_2) - \mathcal{U}(f;\mathfrak{Z})\big)\end{aligned} \tag{5.1.8-C}$$

und nach Voraussetzung (da $\mathfrak{Z}_1 \subset \mathfrak{Z}_2$)

$$\mathcal{O}(f;\mathfrak{Z}_2) - \mathcal{U}(f;\mathfrak{Z}_2) \le \mathcal{O}(f;\mathfrak{Z}_1) - \mathcal{U}(f;\mathfrak{Z}_1) < \frac{\varepsilon}{3}. \tag{5.1.8-D}$$

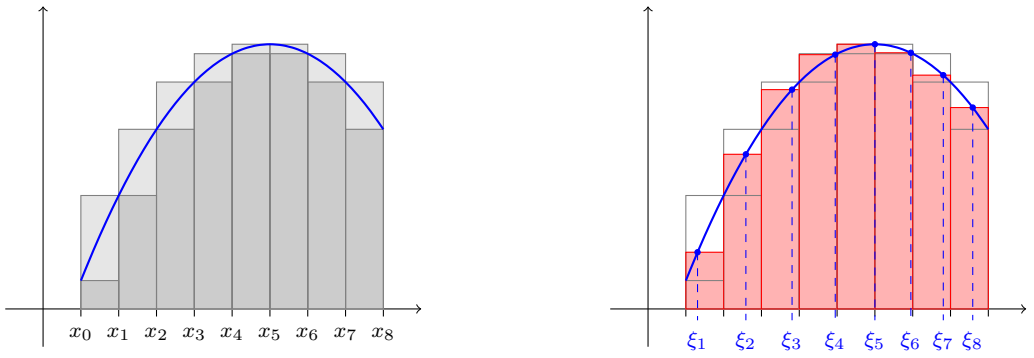
Weiterhin werden höchstens n der Teilintervalle von $\mathfrak{Z}$ beim Übergang zu $\mathfrak{Z}_2$ weiter zerteilt. Da diese die einzigen in der Differenz auftretenden Terme liefern, folgt

$$\mathcal{O}(f;\mathfrak{Z}) - \mathcal{O}(f;\mathfrak{Z}_2) \le n\Omega\delta = \frac{\varepsilon}{3} \tag{5.1.8-E}$$

und entsprechend

$$\mathcal{U}(f;\mathfrak{Z}_2) - \mathcal{U}(f;\mathfrak{Z}) \le n\Omega\delta = \frac{\varepsilon}{3}. \tag{5.1.8-F}$$

Nach Addition folgt $\mathcal{O}(f;\mathfrak{Z}) - \mathcal{U}(f;\mathfrak{Z}) < \varepsilon$ und damit die Behauptung. □



☞ **5.1.9 Korollar** (Riemannsummen). *Angenommen, $f \in \mathfrak{R}[a,b]$. Dann gilt für jede Folge von Zerlegungen $\mathfrak{Z}_m = (x_{0,m}, \dots, x_{n_m,m})$ mit $\delta(\mathfrak{Z}_m) \longrightarrow 0$ und jede Wahl von Stützstellen $\xi_{k,m} \in [x_{k,m}, x_{k-1,m}]$, $k = 1, \dots, n_m$, in den Teilintervallen der Zerlegung $\mathfrak{Z}_m$ für die zugeordneten* Riemannsummen

$$\mathcal{S}(f;\mathfrak{Z}_m,(\xi_{k,m})) = \sum_{k=1}^{n_m} f(\xi_{k,m})(x_{k,m} - x_{k-1,m}) \tag{5.1.9-A}$$

die Grenzwertbeziehung

$$\lim_{m\to\infty} \mathcal{S}(f;\mathfrak{Z}_m,(\xi_{k,m})) = \int_a^b f(x)\,\mathrm{d}x. \tag{5.1.9-B}$$

Beweis. Es gilt

$$\inf_{\xi\in[x_{k-1,m},x_{k,m}]} f(\xi) \le f(\xi_{k,m}) \le \sup_{\xi\in[x_{k-1,m},x_{k,m}]} f(\xi) \tag{5.1.9-C}$$

und die Riemannsumme liegt stets zwischen Untersumme und Obersumme der entsprechenden Zerlegung,

$$\mathcal{U}(f;\mathfrak{Z}) \le \mathcal{S}(f;\mathfrak{Z}_m,(\xi_{k,m})) \le \mathcal{O}(f;\mathfrak{Z}). \tag{5.1.9-D}$$

Wegen $\delta(\mathfrak{Z}_m) \longrightarrow 0$ folgt die Behauptung aus Satz 5.1.8. □

✂ **Ergänzung.** Speziell für die Wahl der $\xi_{k,m}$ als Intervallmitten einer äquidistanten Zerlegung ergibt sich

$$\lim_{n\to\infty} \frac{b-a}{n} \sum_{k=1}^{n} f\left(\frac{2n-2k-1}{2n}a + \frac{2k+1}{2n}b\right) = \int_a^b f(x)\,\mathrm{d}x, \tag{5.1.9-E}$$

mit den rechten Intervallenden entsprechend

$$\lim_{n\to\infty} \frac{b-a}{n} \sum_{k=1}^{n} f\left(\frac{n-k}{n}a + \frac{k}{n}b\right) = \int_a^b f(x)\,\mathrm{d}x. \tag{5.1.9-F}$$

Dies kann man nutzen, um gewisse Grenzwerte als Integrale darzustellen. Wir betrachten zwei Beispiele. Es gilt

$$\lim_{n\to\infty} \frac{1}{n^3} \sum_{k=1}^{n-1} k(n-k) = \lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n-1} \frac{k}{n}\left(1-\frac{k}{n}\right) = \int_0^1 x(1-x)\,\mathrm{d}x = \frac{1}{6} \tag{5.1.9-G}$$

und ebenso

$$\lim_{n\to\infty} \frac{1}{n^2} \sum_{k=1}^{n-1} \sqrt{k(n-k)} = \int_0^1 \sqrt{x(1-x)}\,\mathrm{d}x = \frac{\pi}{8}. \tag{5.1.9-H}$$

Die Berechnung des Integrale diskutieren wir in Abschnitt 5.2. Man kann Riemannsummen aber auch direkt zur Berechnung einfacher Integrale nutzen. So gilt

$$\int_0^1 x\,\mathrm{d}x = \lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} \frac{k}{n} = \lim_{n\to\infty} \frac{1}{n^2} \sum_{k=1}^{n} k = \lim_{n\to\infty} \frac{1}{n^2} \frac{n(n+1)}{2} = \frac{1}{2} \tag{5.1.9-I}$$

unter Ausnutzung der Summenformel der ersten n natürlichen Zahlen.

☞ **5.1.10 Satz.** *Angenommen, $f : [a,b] \to \mathbb{R}$ ist monoton. Dann gilt $f \in \mathfrak{R}[a,b]$.*

Beweis. Wir nehmen ohne Einschränkung der Allgemeinheit an, dass f monoton wachsend ist. Ist nun $\mathfrak{Z} = (x_0, x_1, \ldots, x_n)$ eine Zerlegung von $[a,b]$, so erhalten wir für die zugeordneten Ober- beziehungsweise Untersummen

$$\mathcal{O}(f;\mathfrak{Z}) = \sum_{k=1}^{n} (x_k - x_{k-1}) f(x_k), \qquad \mathcal{U}(f;\mathfrak{Z}) = \sum_{k=1}^{n} (x_k - x_{k-1}) f(x_{k-1}) \tag{5.1.10-A}$$

und damit für die Differenz beider die Abschätzung

$$\begin{aligned} \mathcal{O}(f;\mathfrak{Z}) - \mathcal{U}(f;\mathfrak{Z}) &= \sum_{k=1}^{n} (x_k - x_{k-1})\big(f(x_k) - f(x_{k-1})\big) \\ &\le \delta \sum_{k=1}^{n} \big(f(x_k) - f(x_{k-1})\big) = \delta\big(f(b) - f(a)\big) \end{aligned} \tag{5.1.10-B}$$

durch die Feinheit $\delta = \delta(\mathfrak{Z})$. Also gilt

$$\lim_{\delta(\mathfrak{Z})\to 0} \mathcal{O}(f;\mathfrak{Z}) - \mathcal{U}(f;\mathfrak{Z}) = 0 \tag{5.1.10-C}$$

und die R-Integrierbarkeit folgt. □

☞ **5.1.11 Satz.** *Angenommen, $f : [a,b] \to \mathbb{R}$ ist stetig. Dann gilt $f \in \mathfrak{R}[a,b]$.*

Beweis. Da $[a,b]$ kompakt ist, ist die Funktion f insbesondere gleichmäßig stetig. Es gibt also zu jedem $\varepsilon > 0$ ein $\delta > 0$ mit

$$|x_1 - x_2| < \delta \implies |f(x_1) - f(x_2)| < \varepsilon. \tag{5.1.11-A}$$

Wählt man nun eine Zerlegung $\mathfrak{Z}$ mit Feinheit $\delta(\mathfrak{Z}) < \delta$, so folgt mit

$$f(\xi_k) = \max_{\xi\in[x_{k-1},x_k]} f(\xi) \qquad \text{und} \qquad f(\eta_k) = \min_{\xi\in[x_{k-1},x_k]} f(\xi) \tag{5.1.11-B}$$

und damit $|f(\xi_k) - f(\eta_k)| < \varepsilon$ für die Differenz aus Ober- und Untersummen

$$\mathcal{O}(f;\mathfrak{Z}) - \mathcal{U}(f;\mathfrak{Z}) = \sum_{k=1}^{n} (x_k - x_{k-1})(f(\xi_k) - f(\eta_k)) < \varepsilon(b-a) \tag{5.1.11-C}$$

und die R-Integrierbarkeit folgt. □

☞ **5.1.12 Satz** (Substitutionsregel)**.** *Angenommen, $\varphi : [a,b] \to [c,d]$ ist bijektiv, stetig differenzierbar und es gilt $0 < m \le \varphi'(x) \le M$ für $x \in [a,b]$. Dann gilt*

$$g \in \mathfrak{R}[c,d] \qquad \Longleftrightarrow \qquad f \in \mathfrak{R}[a,b] \tag{5.1.12-A}$$

mit $f(x) = g(\varphi(x))\varphi'(x)$ und darüberhinaus

$$\int_a^b f(x)\,\mathrm{d}x = \int_a^b g(\varphi(x))\varphi'(x)\,\mathrm{d}x = \int_c^d g(y)\,\mathrm{d}y. \tag{5.1.12-B}$$

Beweis. Wir können ohne Beschränkung der Allgemeinheit annehmen, dass $g \ge 0$ und damit auch $f \ge 0$ gilt (sonst betrachten wir statt g die Funktion $g + \mathrm{C}$ für hinreichend großes C).

Nach Voraussetzung ist $\varphi : [a,b] \to [c,d]$ streng monoton wachsend. Jede Zerlegung $\mathfrak{Z} = (x_0, x_1, \ldots, x_n)$ von $[a,b]$ wird durch φ auf eine Zerlegung $\varphi(\mathfrak{Z}) = (y_0, \ldots, y_n)$, $y_k = \varphi(x_k)$, von $[c,d]$ abgebildet. Dabei gilt auf Grund des ersten Mittelwertsatzes der Differentialrechnung angewandt auf φ' und φ

$$m(x_k - x_{k-1}) \le \varphi(x_k) - \varphi(x_{k-1}) \le M(x_k - x_{k-1}) \tag{5.1.12-C}$$

und somit $m\delta(\mathfrak{Z}) \le \delta(\varphi(\mathfrak{Z})) \le M\delta(\mathfrak{Z})$. Darüberhinaus ist φ' gleichmäßig stetig, es gibt also zu jedem $\varepsilon > 0$ ein $\delta > 0$ mit

$$|x - x'| < \delta \qquad \Longrightarrow \qquad |\varphi(x) - \varphi(x')| \le \varepsilon. \tag{5.1.12-D}$$

Wir nehmen im weiteren an, dass $\delta(\mathfrak{Z}) < \delta$ gilt und nutzen

$$y_k - y_{k-1} = \varphi(x_k) - \varphi(x_{k-1}) = \varphi'(\xi_k)(x_k - x_{k-1}) \qquad \text{für ein } \xi_k \in [x_{k-1}, x_k]. \quad (5.1.\textbf{12-E})$$

Sei nun $g \in \mathfrak{R}[c,d]$ und $f(x) = g(\varphi(x))\varphi'(x)$. Dann gilt für Obersummen von f unter Ausnutzung von (5.1.**12-D**) für die erste Abschätzung

$$\begin{aligned}
\mathcal{O}(f;\mathfrak{Z}) &= \sum_{k=1}^{n}(x_k - x_{k-1}) \sup_{\xi\in[x_{k-1},x_k]} g(\varphi(\xi))\varphi'(\xi) \\
&= \sum_{k=1}^{n}(y_k - y_{k-1}) \frac{1}{\varphi'(\xi_k)} \sup_{\xi\in[x_{k-1},x_k]} g(\varphi(\xi))\varphi'(\xi) \\
&\le \sum_{k=1}^{n}(y_k - y_{k-1}) \frac{\varphi'(\xi_k)+\varepsilon}{\varphi'(\xi_k)} \sup_{\xi\in[x_{k-1},x_k]} g(\varphi(\xi)) \\
&\le \Big(1+\frac{\varepsilon}{m}\Big) \sum_{k=1}^{n}(y_k - y_{k-1}) \sup_{\eta\in[y_{k-1},y_k]} g(\eta) = \Big(1+\frac{\varepsilon}{m}\Big)\,\mathcal{O}(g;\varphi(\mathfrak{Z}))
\end{aligned} \qquad (5.1.\textbf{12-F})$$

und entsprechend für Untersummen

$$\begin{aligned}
\mathcal{U}(f;\mathfrak{Z}) &= \sum_{k=1}^{n}(x_k - x_{k-1}) \inf_{\xi\in[x_{k-1},x_k]} g(\varphi(\xi))\varphi'(\xi) \\
&= \sum_{k=1}^{n}(y_k - y_{k-1}) \frac{1}{\varphi'(\xi_k)} \inf_{\xi\in[x_{k-1},x_k]} g(\varphi(\xi))\varphi'(\xi) \\
&\ge \sum_{k=1}^{n}(y_k - y_{k-1}) \frac{\varphi'(\xi_k)-\varepsilon}{\varphi'(\xi_k)} \sup_{\xi\in[x_{k-1},x_k]} g(\varphi(\xi)) \\
&\ge \Big(1-\frac{\varepsilon}{m}\Big) \sum_{k=1}^{n}(y_k - y_{k-1}) \inf_{\eta\in[y_{k-1},y_k]} g(\eta) = \Big(1-\frac{\varepsilon}{m}\Big)\,\mathcal{U}(g;\varphi(\mathfrak{Z}))
\end{aligned} \qquad (5.1.\textbf{12-G})$$

und die R-integrierbarkeit von g über $[c,d]$ impliziert die R-integrierbarkeit von f über $[a,b]$ und die Gleichheit der Integrale. □

✂ **Ergänzung.** Kennt man die Integrierbarkeit beider Funktionen, gilt also $f \in \mathfrak{R}[a,b]$ und $g \in \mathfrak{R}[c,d]$, so folgt die Gleichheit der Integrale allein aus Monotonie von φ und der Bijektivität. Dazu genügt

$$\begin{aligned}
\mathcal{O}(g;\varphi(\mathfrak{Z})) &= \sum_{k=1}^{n}(y_k - y_{k-1}) \sup_{\eta\in[y_{k-1},y_k]} g(\eta) \\
&= \sum_{k=1}^{n}(\varphi(x_k) - \varphi(x_{k-1})) \sup_{\xi\in[x_{k-1},x_k]} g(\varphi(\xi)) \\
&= \sum_{k=1}^{n}(x_k - x_{k-1})\varphi'(\xi_k) \sup_{\xi\in[x_{k-1},x_k]} g(\varphi(\xi)) \\
&\ge \sum_{k=1}^{n}(x_k - x_{k-1})\varphi'(\xi_k) g(\varphi(\xi_k)) = \mathcal{S}(f;\mathfrak{Z},(\xi_k))
\end{aligned} \qquad (5.1.\textbf{12-H})$$

und entsprechend

$$\begin{aligned}\mathcal{U}(g;\varphi(\mathfrak{Z})) &= \sum_{k=1}^{n}(y_k - y_{k-1}) \inf_{\eta\in[y_{k-1},y_k]} g(\eta)\\ &= \sum_{k=1}^{n}(x_k - x_{k-1})\varphi'(\xi_k) \inf_{\xi\in[x_{k-1},x_k]} g(\varphi(\xi)) \qquad (5.1.\textbf{12-I})\\ &\le \sum_{k=1}^{n}(x_k - x_{k-1})\varphi'(\xi_k) g(\varphi(\xi_k)) = \mathcal{S}(f;\mathfrak{Z},(\xi_k))\end{aligned}$$

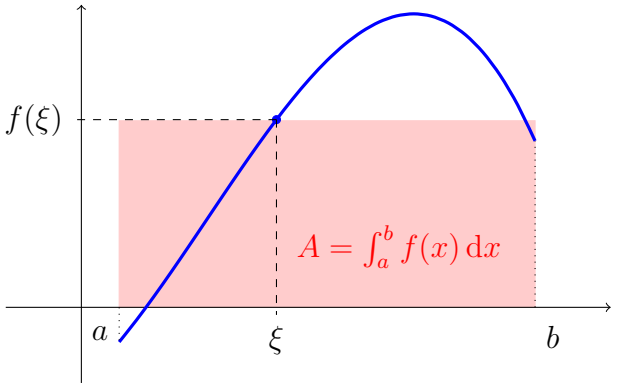
für die Riemannsummen $\mathcal{S}(f;\mathfrak{Z},(\xi_k))$ bezüglich der durch Anwendung des ersten Mittelwertsatzes erhaltenen Wahl von Stützstellen (ξ_k).

Insbesondere benötigt die Substitutionsregel damit (wenn die Integrierbarkeit der Funktionen anderweitig bekannt ist) nicht die starke Voraussetzung $\varphi'(x) \ge m > 0$ des Satzes 5.1.12. Die Monotonie (und damit Bijektivität) der differenzierbaren Funktion φ genügt.

5.2 Hauptsätze der Integralrechnung

☞ **5.2.1 Satz** (1. Mittelwertsatz der Integralrechnung). *Sei $f : [a,b] \to \mathbb{R}$ stetig. Dann existiert ein $\xi \in (a,b)$ mit*

$$\int_a^b f(x)\,\mathrm{d}x = f(\xi)(b-a). \qquad (5.2.\textbf{1-A})$$



Beweis. Da f stetig und $[a,b]$ kompakt ist, nimmt f auf $[a,b]$ sowohl Minimum

$$f(x_\bullet) = \min\{f(x) \mid x \in [a,b]\} \qquad (5.2.\textbf{1-B})$$

als auch Maximum

$$f(x^\bullet) = \max\{f(x) \mid x \in [a,b]\} \qquad (5.2.\textbf{1-C})$$

an. Damit ist $f(x_\bullet)(b-a)$ eine Unter- und $f(x^\bullet)(b-a)$ eine Obersumme des Integrals, es gilt also

$$f(x_\bullet) \le \frac{1}{b-a}\int_a^b f(x)\,\mathrm{d}x \le f(x^\bullet). \qquad (5.2.\textbf{1-D})$$

Gleichheit ist dabei für nichtkonstante Funktionen ausgeschlossen. Gibt es ein $\eta \in (a,b)$ mit $f(\eta) > f(x_\bullet)$, so existiert auch ein Intervall $[\eta-\varepsilon,\eta+\varepsilon]$ mit $f(x) > \frac{1}{2}(f(\eta)+f(x_\bullet)$ für

alle $x \in [\eta-\varepsilon, \eta+\varepsilon]$ und damit auch eine Untersumme mit Wert größer als $(b-a)f(x_\bullet)$. Analoges gilt für $f(x^\bullet)$ und Obersummen. Nach dem Zwischenwertsatz existiert also ein $\xi \in (a,b)$ mit

$$f(\xi) = \frac{1}{b-a}\int_a^b f(x)\,\mathrm{d}x \tag{5.2.1-E}$$

und die Aussage ist gezeigt. □

☞ **5.2.2 Satz** (Hauptsatz der Differential- und Integralrechnung)**.** *Angenommen, $f : [a,b] \to \mathbb{R}$ ist stetig. Dann ist die Funktion $F : [a,b] \to \mathbb{R}$ definiert als*

$$F(x) = \int_a^x f(\xi)\,\mathrm{d}\xi, \qquad x \in [a,b], \tag{5.2.2-A}$$

ist auf (a,b) differenzierbar und es gilt

$$F'(x) = f(x), \qquad x \in (a,b). \tag{5.2.2-B}$$

Damit ist F insbesondere eine Stammfunktion von f.

Beweis. Da f stetig ist, ist f auf jedem Intervall $[a,x]$ R-integrierbar und $F(x)$ wohldefiniert. Wir betrachten den Differenzenquotienten

$$\frac{F(x+h)-F(x)}{h} = \frac{1}{h}\int_x^{x+h} f(\xi)\,\mathrm{d}\xi \tag{5.2.2-C}$$

für $h > 0$. Nach dem 1. Mittelwertsatz der Integralrechnung finden wir damit ein ξ zwischen x und $x+h$ mit

$$\frac{F(x+h)-F(x)}{h} = f(\xi_h) \tag{5.2.2-D}$$

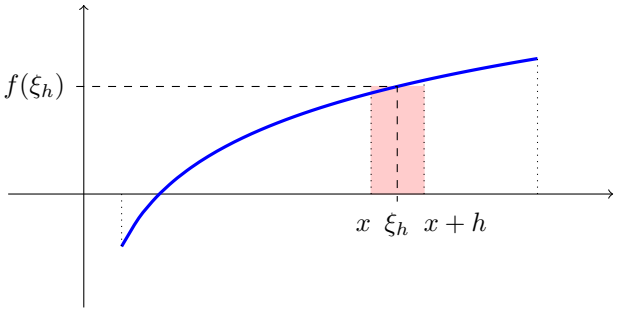
und da f stetig ist gilt wegen $\xi_h \to x$ für $h \searrow 0$

$$\lim_{h\searrow 0}\frac{F(x+h)-F(x)}{h} = \lim_{h\searrow 0} f(\xi_h) = f(x). \tag{5.2.2-E}$$

Entsprechend gilt für den linksseitigen Differenzenquotienten

$$\frac{F(x)-F(x-h)}{h} = \frac{1}{h}\int_{x-h}^{x} f(\xi)\,\mathrm{d}\xi = f(\xi_h) \longrightarrow f(x), \qquad h \searrow 0, \tag{5.2.2-F}$$

mit $\xi_h \in (x-h, x)$ und die Behauptung folgt. □



☞ **5.2.3 Korollar** (Formel von Newton–Leibniz). *Sei $f : [a, b] \to \mathbb{R}$ stetig und F eine Stammfunktion von f. Dann gilt*

$$\int_a^b f(x)\,\mathrm{d}x = F(b) - F(a). \tag{5.2.3-A}$$

Beweis. Da sich beliebige Stammfunktionen nur um Konstanten unterscheiden gilt also

$$F(x) = \int_a^x f(\xi)\,\mathrm{d}\xi + \mathrm{C} \tag{5.2.3-B}$$

mit einer Konstanten C. Also folgt $\mathrm{C} = F(a)$ und somit speziell für $x = b$

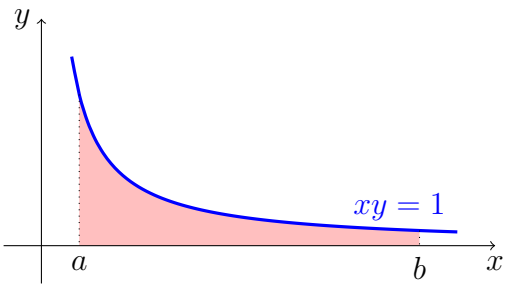
$$F(b) = \int_a^b f(\xi)\,\mathrm{d}\xi + F(a) \tag{5.2.3-C}$$

und damit die Behauptung. □

⚒ **5.2.4 Beispiel.** Mit der Formel von Newton–Leibniz kann man Riemannintegrale berechnen; dazu einige Beispiele:

(i) Die Funktion $f(x) = x^{-1}$ hat als Stammfunktion den Logarithmus. Also gilt

$$\int_a^b \frac{\mathrm{d}x}{x} = \ln x\Big|_a^b = \ln b - \ln a = \ln\frac{b}{a}. \tag{5.2.4-A}$$



(ii) Es gilt

$$\int_a^b \mathrm{e}^x\,\mathrm{d}x = \mathrm{e}^x\Big|_a^b = \mathrm{e}^b - \mathrm{e}^a \tag{5.2.4-B}$$

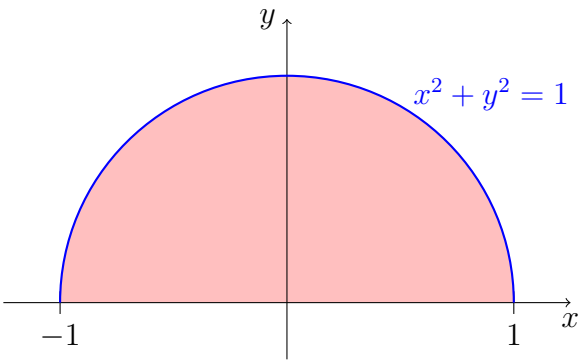
und entsprechend

$$\int_0^\pi \sin x\,\mathrm{d}x = -\cos x\Big|_0^\pi = 2. \tag{5.2.4-C}$$

(iii) Die Berechnung der Integrale entspricht einer Flächenberechnung. So ergibt sich die Fläche eines halben Einheitskreises zu

$$\begin{aligned}\int_{-1}^1 \sqrt{1-x^2}\,\mathrm{d}x &= \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \sqrt{1-\sin^2\theta}\cos\theta\,\mathrm{d}\theta = \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \cos^2\theta\,\mathrm{d}\theta \\ &= \frac{\sin\theta\cos\theta+\theta}{2}\Big|_{-\frac{\pi}{2}}^{\frac{\pi}{2}} = \frac{\pi}{2},\end{aligned} \tag{5.2.4-D}$$

wie zu erwarten war.



(iv) Wir liefern die Rechnung zu den Integralen in (5.1.**9-G**) und (5.1.**9-H**) nach. Es gilt

$$\int_0^1 x(1-x)\,\mathrm{d}x = \left[\frac{1}{2}x^2 - \frac{1}{3}x^3\right]_0^1 = \frac{1}{2} - \frac{1}{3} = \frac{1}{6} \tag{5.2.**4-E**}$$

und

$$\int_0^1 \sqrt{x(1-x)}\,\mathrm{d}x = \frac{1}{2}\int_0^1 \sqrt{1-(2x-1)^2}\,\mathrm{d}x = \frac{1}{4}\int_{-1}^1 \sqrt{1-y^2}\,\mathrm{d}y = \frac{\pi}{8} \tag{5.2.**4-F**}$$

mit $y = 2x - 1$.

☞ **5.2.5 Satz** (2. Mittelwertsatz der Integralrechnung, Mittelwertsatz von Cauchy)**.** *Seien* $f : [a,b] \to \mathbb{R}$ *stetig und* $g \in \mathfrak{R}[a,b]$ *integrierbar mit* $g \geq 0$*. Dann existiert ein* $\xi \in [a,b]$ *mit*

$$\int_a^b f(x)g(x)\,\mathrm{d}x = f(\xi)\int_a^b g(x)\,\mathrm{d}x. \tag{5.2.**5-A**}$$

Beweis. Der Beweis beruht (wie bei dem oben angegebenen ersten Mittelwertsatz) auf dem Zwischenwertsatz. Wir geben ihn der Vollständigkeit halber an und nehmen darüberhinaus an, dass g nicht die Nullfunktion ist.

Bezeichne wiederum

$$y_\bullet = f(x_\bullet) = \min\{f(x) \mid x \in [a,b]\} \tag{5.2.**5-B**}$$

und

$$y^\bullet = f(x^\bullet) = \max\{f(x) \mid x \in [a,b]\}. \tag{5.2.**5-C**}$$

Dann gilt auf Grund von $g \geq 0$

$$y_\bullet g(x) \leq f(x)g(x) \leq y^\bullet g(x) \tag{5.2.**5-D**}$$

und damit auf Grund der Monotonie des Integrals

$$y_\bullet \int_a^b g(x)\,\mathrm{d}x \leq \int_a^b f(x)g(x)\,\mathrm{d}x \leq y^\bullet \int_a^b g(x)\,\mathrm{d}x, \tag{5.2.**5-E**}$$

der Quotient der beiden Integrale liegt also zwischen den Funktionswerten $y_\bullet$ und $y^\bullet$. Nach dem Zwischenwertsatz existiert also $\xi \in [a,b]$ mit

$$\left(\int_a^b g(x)\,\mathrm{d}x\right)^{-1} \int_a^b f(x)g(x)\,\mathrm{d}x = f(\xi) \tag{5.2.**5-F**}$$

und damit die Behauptung. □

✂ **Ergänzung.** Mitunter ist eine alternative Formulierung des gerade gezeigten nützlich, um den Zusammenhang zum zweiten Mittelwertsatz der Differentialrechnung deutlich zu machen. Angenommen beide Funktionen f und g sind stetig. Schreibt man dann $h(x) = f(x)g(x)$, so ergibt sich aus dem gerade gezeigten

$$\frac{H(b)-H(a)}{G(b)-G(a)} = \frac{\int_a^b h(x)\,\mathrm{d}x}{\int_a^b g(x)\,\mathrm{d}x} = \frac{h(\xi)}{g(\xi)} \tag{5.2.5-G}$$

für ein $\xi \in [a,b]$ und mit Stammfunktionen H und G zu h und g. Der zweite Mittelwertsatz der Differentialrechnung ist dabei nun allgemeiner, da nicht jede Ableitung einer differenzierbaren Funktion stetig sein muss. Umgekehrt ist obige Formulierung des Satzes allgemeiner, da diese die Stetigkeit der Funktion g nicht voraussetzt.

5.3 Berechnung von Integralen

5.3.1. Mit der Formel von Newton–Leibniz haben wir die Berechnung von Riemannintegralen stetiger Funktionen auf die Bestimmung von Stammfunktionen, also von unbestimmten Integralen zurückgeführt. Neben den schon im letzten Kapitel diskutierten Grundintegralen spielen dabei wichtige Grundtechniken eine Rolle, die wir jetzt noch systematisch diskutieren wollen.

Wir fassen die *Grundintegrale* noch einmal zusammen. Es gilt jeweils auf allen Intervallen im Definitionsbereich des Integranden

$$\int x^n\,\mathrm{d}x = \frac{x^{n+1}}{n+1} + \mathrm{C}, \qquad n \neq -1, \tag{5.3.1-A}$$

$$\int \frac{\mathrm{d}x}{x} = \ln|x| + \mathrm{C}, \tag{5.3.1-B}$$

$$\int \frac{\mathrm{d}x}{1+x^2} = \arctan x + \mathrm{C}, \tag{5.3.1-C}$$

$$\int \mathrm{e}^x\,\mathrm{d}x = \mathrm{e}^x + \mathrm{C}, \tag{5.3.1-D}$$

$$\int \cos x\,\mathrm{d}x = \sin x + \mathrm{C}, \tag{5.3.1-E}$$

$$\int \sin x\,\mathrm{d}x = -\cos x + \mathrm{C}, \tag{5.3.1-F}$$

$$\int \frac{\mathrm{d}x}{\cos^2 x} = \tan x + \mathrm{C}, \qquad |x| < \frac{\pi}{2}, \tag{5.3.1-G}$$

$$\int \frac{\mathrm{d}x}{\sqrt{1-x^2}} = \arcsin x + \mathrm{C} = \mathrm{C} - \arccos x, \tag{5.3.1-H}$$

$$\int \frac{\mathrm{d}x}{\sqrt{x^2-1}} = \operatorname{arcosh} x + \mathrm{C}, \qquad x > 1, \tag{5.3.1-I}$$

$$\int \frac{\mathrm{d}x}{\sqrt{1+x^2}} = \operatorname{arsinh} x + \mathrm{C}, \tag{5.3.1-J}$$

zum Beweis genügt es, die jeweilige rechte Seite abzuleiten. Darüberhinaus ist es oft hilfreich, sich die Identitäten

$$\arcsin x + \arccos x = \frac{\pi}{2} = \arctan x + \operatorname{arccot} x \tag{5.3.1-K}$$

einzuprägen und sich bewusst zu machen, dass verschiedene Rechenwege zur Berechnung desselben unbestimmten Integrals zu verschiedenen Integrationskonstanten führen können.

Integration rationaler Funktionen

⚒ **5.3.2 Beispiel.** *Rationale Funktionen* sind Quotienten von Polynomfunktionen. Diese können durch Polynomdivision vereinfacht und danach durch Partialbruchzerlegung integriert werden. Wir beginnen mit einem Beispiel. Gegeben sei

$$f(x) = \frac{x^3 + x^2 + x}{x^2 - 1}, \qquad x \neq \pm 1. \tag{5.3.2-A}$$

Polynomdivision erlaubt es, den Grad des Zählers zu reduzieren. Es gilt

$$\frac{x^3 + x^2 + x}{x^2 - 1} = x + \frac{x^2 + 2x}{x^2 - 1} = x + 1 + \frac{2x + 1}{x^2 - 1} \tag{5.3.2-B}$$

und wir können den verbleibenden Bruch weiter als Summe elementar integrierbarer Terme

$$\frac{2x + 1}{x^2 - 1} = \frac{2x}{x^2 - 1} + \frac{1}{x^2 - 1} = \left(\frac{1}{x - 1} + \frac{1}{x + 1}\right) + \frac{1}{2}\left(\frac{1}{x - 1} - \frac{1}{x + 1}\right) \tag{5.3.2-C}$$

schreiben. Also gilt

$$\begin{aligned} \int f(x)\,\mathrm{d}x &= \int \left(x + 1 + \frac{3}{2}\frac{1}{x - 1} + \frac{1}{2}\frac{1}{x + 1}\right)\,\mathrm{d}x \\ &= \frac{1}{2}x^2 + x + \frac{3}{2}\ln|x - 1| + \frac{1}{2}\ln|x + 1| + \mathrm{C} \end{aligned} \tag{5.3.2-D}$$

auf jedem der Teilintervalle $(-\infty, -1)$, $(-1, 1)$ und $(1, +\infty)$.

Es ist kein Zufall, dass am Ende nur Grundintegrale übriggeblieben sind. Wir wollen als nächstes zeigen, dass jede rationale Funktion auf diese Weise auf elementar integrierbare Terme reduziert werden kann.

☞ **5.3.3 Lemma.** *Sei $p(x)$ eine reelle Polynomfunktion mit führendem Koeffizienten 1. Dann existieren eindeutig bestimmte Zahlen $\alpha_i \in \mathbb{R}$, $i = 1, \dots m$, und $\beta_j, \gamma_j \in \mathbb{R}$, $j = 1, \dots n$, mit*

$$\gamma_j > \beta_j^2 \tag{5.3.3-A}$$

und

$$p(x) = \prod_{i=1}^{m}(x - \alpha_i) \prod_{j=1}^{n}(x^2 - 2\beta_j x + \gamma_j). \tag{5.3.3-B}$$

Beweis. Nach dem Fundamentalsatz der Algebra zerfällt p über $\mathbb{C}$ in Linearfaktoren. Es genügt zu zeigen, dass diese geeignet paarweise kombiniert werden können. Da p reell ist, gilt (jetzt für $x \in \mathbb{C}$) auch $p(\overline{x}) = \overline{p(x)}$ und zu jeder nicht reellen Nullstelle

$\lambda \in \mathbb{C} \setminus \mathbb{R}$ mit $p(\lambda) = 0$ ist auch $\overline{\lambda}$ eine Nullstellen. Multipliziert man die zugehörigen Linearfaktoren, so ergibt sich

$$(x-\lambda)(x-\overline{\lambda}) = x^2 - 2x \operatorname{Re}\lambda + |\lambda|^2 = x^2 - 2\beta x + \gamma \tag{5.3.3-C}$$

mit $-|\lambda| < \beta = \operatorname{Re}\lambda < |\lambda|$ und $\gamma = |\lambda|^2$. Auf diese Weise kann man zu jeder komplexen Nullstelle einen quadratischen Term aus $p(x)$ dividieren, das verbleibende Polynom ist wieder reell und nach endlich vielen Schritten verbleiben nur lineare reelle Faktoren. □

Wir werden im weiteren die gleichen Faktoren zusammenfassen und stattdessen

$$p(x) = \prod_{i=1}^{m} (x-\alpha_i)^{\mu_i} \prod_{j=1}^{n} (x^2 - 2\beta_j x + \gamma_j)^{\nu_j} \tag{5.3.3-D}$$

mit paarweise verschiedenen α_i und paarweise verschiedenen Paaren (β_j, γ_j) schreiben. Dann gilt insbesondere

$$\sum_{i=1}^{m} \mu_i + 2\sum_{j=1}^{n} \nu_j = \deg p. \tag{5.3.3-E}$$

Die Partialbruchzerlegung echt gebrochener rationaler Funktionen ist eine direkte Konsequenz dieser Faktorisierung.

☞ **5.3.4 Satz** (reelle Partialbruchzerlegung). *Sei*

$$f(x) = \frac{r(x)}{p(x)} \tag{5.3.4-A}$$

eine rationale Funktion für reelle Polynome p und r mit $1 \le \deg r < \deg p$ und p mit führendem Koeffizienten 1 wie in (5.3.3-D) *faktorisiert. Dann existieren eindeutig bestimmte Zahlen $A_{i,\kappa} \in \mathbb{R}$, $\kappa = 1, \dots, \mu_i$, $i = 1, \dots, m$ und $B_{j,\kappa}, C_{j,\kappa} \in \mathbb{R}$, $\kappa = 1, \dots, \nu_j$, $j = 1, \dots, n$, so dass*

$$f(x) = \sum_{i=1}^{m} \sum_{\kappa=1}^{\mu_i} \frac{A_{i,\kappa}}{(x-\alpha_i)^\kappa} + \sum_{j=1}^{n} \sum_{\kappa=1}^{\nu_j} \frac{B_{j,\kappa}x + C_{j,\kappa}}{(x^2 - 2\beta_j x + \gamma_j)^\kappa} \tag{5.3.4-B}$$

gilt.

Beweis. Wir zeigen nur die Existenz einer solchen Darstellung und überlassen die Eindeutigkeit als Übungsaufgabe.

Sei $\alpha \in \mathbb{R}$ eine beliebige reelle Nullstelle des Polynoms $p(x)$ mit Vielfachheit μ und $q(x)$ so gewählt, dass $p(x) = (x-\alpha)^\mu q(x)$. Dann gilt insbesondere $q(\alpha) \neq 0$ und mit $A = r(\alpha)/q(\alpha)$ gibt es ein Polynom $r_1(x)$ mit $r(x) - Aq(x) = (x-\alpha)r_1(x)$. Also folgt mit $p_1(x) = (x-\alpha)^{\mu-1}q(x)$

$$\frac{r(x)}{p(x)} - \frac{A}{(x-\alpha)^\mu} = \frac{r(x) - Aq(x)}{p(x)} = \frac{r_1(x)}{p_1(x)}, \tag{5.3.4-C}$$

mit $\deg p_1 = \deg p - 1$. Nach endlich vielen Schritten erhalten wir eine echt gebrochene rationale Funktion ohne reelle Nullstellen des Nenners.

Sei nun $x^2 - 2\beta x + \gamma$ ein zu einer komplexen Nullstelle λ der Vielfachheit ν gehörender quadratischer Faktor von p und $p(x) = (x^2 - 2\beta x + \gamma)^\nu q(x)$ mit $q(\lambda) \neq 0$. Dann gibt es eindeutig bestimmte Zahlen $B, C \in \mathbb{R}$ mit $B\lambda + C = r(\lambda)/q(\lambda)$. Damit verschwindet $r(x) - (Bx + C)q(x)$ für $x = \lambda$ und $x = \overline{\lambda}$ und es gibt ein Polynom $r_1(x)$ mit

$$r(x) - (Bx + C)q(x) = (x^2 - 2\beta x + \gamma)r_1(x). \tag{5.3.4-D}$$

Zusammen mit $p_1(x) = (x^2 - 2\beta x + \gamma)^{\nu-1} q(x)$ folgt

$$\frac{r(x)}{p(x)} - \frac{Bx + C}{(x^2 - 2\beta x + \gamma)^\nu} = \frac{r(x) - (Bx + C)q(x)}{p(x)} = \frac{r_1(x)}{p_1(x)} \tag{5.3.4-E}$$

mit $\deg p_1 = \deg p - 2$. Nach endlich vielen Schritten haben wir alle Nullstellen des Nenners abgetrennt, die verbleibende Funktion ist die Nullfunktion. □

⚒ **5.3.5 Beispiel.** Es verbleibt, die einzelnen Summanden zu integrieren. Für die Linearfaktoren nutzen wir

$$\int \frac{\mathrm{d}x}{x - \alpha} = \ln|x - \alpha| + \mathrm{C} \tag{5.3.5-A}$$

und

$$\int \frac{\mathrm{d}x}{(x - \alpha)^{1+\kappa}} = \frac{-1}{\kappa(x - \alpha)^\kappa} + \mathrm{C}, \qquad \kappa \in \mathbb{N}. \tag{5.3.5-B}$$

Weiterhin gilt

$$\int \frac{x - \beta}{x^2 - 2\beta x + \gamma}\,\mathrm{d}x = \frac{1}{2}\ln|x^2 - 2\beta x + \gamma| + \mathrm{C}, \tag{5.3.5-C}$$

sowie (unter Ausnutzung von $\gamma > \beta^2$)

$$\begin{aligned}\int \frac{\mathrm{d}x}{x^2 - 2\beta x + \gamma} &= \int \frac{\mathrm{d}x}{(x - \beta)^2 + \gamma - \beta^2} = \frac{1}{\gamma - \beta^2}\int \frac{\mathrm{d}x}{1 + \left(\frac{x-\beta}{\sqrt{\gamma-\beta^2}}\right)^2} \\ &= \frac{1}{\sqrt{\gamma - \beta^2}} \arctan\frac{x - \beta}{\sqrt{\gamma - \beta^2}} + \mathrm{C}.\end{aligned} \tag{5.3.5-D}$$

Die weiteren aus den quadratischen Faktoren ergeben sich aus

$$\int \frac{x - \beta}{(x^2 - 2\beta x + \gamma)^{1+\kappa}}\,\mathrm{d}x = \frac{-1}{2\kappa(x^2 - 2\beta x + \gamma)^\kappa} + \mathrm{C}, \tag{5.3.5-E}$$

sowie der sich aus partieller Integration

$$\begin{aligned}\int \frac{\mathrm{d}x}{(x^2 - 2\beta x + \gamma)^\kappa} &= \frac{x}{(x^2 - 2\beta x + \gamma)^\kappa} + 2\kappa\int \frac{x(x - \beta)}{(x^2 - 2\beta x + \gamma)^{1+\kappa}}\,\mathrm{d}x \\ &= \frac{x}{(x^2 - 2\beta x + \gamma)^\kappa} + 2\kappa\int \frac{\mathrm{d}x}{(x^2 - 2\beta x + \gamma)^\kappa} \\ &\qquad + 2\kappa\int \frac{\beta x - \gamma}{(x^2 - 2\beta x + \gamma)^{1+\kappa}}\,\mathrm{d}x\end{aligned} \tag{5.3.5-F}$$

zusammen mit dem gerade berechneten Integral

$$\begin{aligned}\int \frac{\beta x-\gamma}{(x^2-2\beta x+\gamma)^{1+\kappa}}\,\mathrm{d}x &= \int \frac{\beta(x-\beta)}{(x^2-2\beta x+\gamma)^{1+\kappa}} + \frac{\beta^2-\gamma}{(x^2-2\beta x+\gamma)^{1+\kappa}}\,\mathrm{d}x \\ &= \frac{-\beta}{2\kappa(x^2-2\beta x+\gamma)^{\kappa}} - \int \frac{\gamma-\beta^2}{(x^2-2\beta x+\gamma)^{1+\kappa}}\,\mathrm{d}x\end{aligned} \tag{5.3.5-G}$$

ergebenden Rekursionsformel

$$\begin{aligned}\int \frac{\gamma-\beta^2}{(x^2-2\beta x+\gamma)^{1+\kappa}}\,\mathrm{d}x = \frac{x-\beta}{2\kappa(x^2-2\beta x+\gamma)^{\kappa}} \\ + \frac{2\kappa-1}{2\kappa}\int \frac{\mathrm{d}x}{(x^2-2\beta x+\gamma)^{\kappa}}.\end{aligned} \tag{5.3.5-H}$$

5.3.6. Für die Bestimmung der Koeffizienten in der Partialbruchzerlegung (5.3.4-B) gibt es verschiedene Methoden. Einerseits liefert Multiplikation mit dem Hauptnenner und Koeffizientenvergleich ein stets eindeutig lösbares Gleichungssystem für die gesuchten Koeffizienten. Koeffizienten zu Linearfaktoren mit maximalem Exponenten lassen sich aber auch direkt ablesen. Es gilt

$$A_{i,\mu_i} = \lim_{x\to\alpha_i} (x-\alpha_i)^{\mu_i} f(x) \tag{5.3.6-A}$$

da nach Multiplikation mit $(x-\alpha_i)^{\mu_i}$ alle anderen Terme auf der rechten Seite von (5.3.4-B) nach Kürzen mindestens einen verbleibenden Faktor $(x-\alpha_i)$ besitzen und somit im Grenzwert mit Null multipliziert werden. Für Koeffizienten zu Termen mit niedrigeren Koeffizienten muss man hier noch entsprechende Terme abziehen,

$$A_{i,\kappa} = \lim_{x\to\alpha_i} (x-\alpha_i)^{\kappa}\left(f(x) - \sum_{\rho=\kappa+1}^{\mu_i} \frac{A_{i,\rho}}{(x-\alpha_i)^{\rho}}\right), \qquad \kappa<\mu_i, \tag{5.3.6-B}$$

oder differenzieren,

$$A_{i,\kappa} = \lim_{x\to\alpha_i} \frac{1}{(\mu_i-\kappa)!}\left(\frac{\mathrm{d}}{\mathrm{d}x}\right)^{\mu_i-\kappa} (x-\alpha_i)^{\mu_i} f(x). \tag{5.3.6-C}$$

Die Lösung eines linearen Gleichungssystems ist in diesen Fällen oft einfacher.

✂ **Ergänzung.** Für die quadratischen Terme kann man analog vorgehen und die komplexen Nullstellen

$$z_{j,\pm} = \beta_j \pm \mathrm{i}\sqrt{\gamma_j-\beta_j^2} \in \mathbb{C}, \qquad j=1,\dots,n, \tag{5.3.6-D}$$

des Nenners nutzen. Es gilt

$$B_{j,\nu_j} z_{j,\pm} + C_{j,\nu_j} = \lim_{x\to z_{j,\pm}} (x^2-2\beta x+\gamma)^{\nu_j} f(x) \tag{5.3.6-E}$$

und die Zahlen $B_{j,\nu_j}, C_{j,\nu_j} \in \mathbb{R}$ sind durch die beiden Grenzwerte eindeutig bestimmt.

⚒ **5.3.7 Beispiel.** Wir betrachten die rationale Funktion

$$f(x) = \frac{x^3}{(x-1)^2(x+1)}, \qquad x\neq\pm1, \tag{5.3.7-A}$$

in der Zähler und Nenner gleichen Grad haben. Polynomdivision führt offenbar auf den ganzen Anteil 1 und damit liefert die Partialbruchzerlegung die Form

$$\frac{x^3}{(x-1)^2(x+1)} = 1 + \frac{A_{1,1}}{x-1} + \frac{A_{1,2}}{(x-1)^2} + \frac{A_2}{x+1} \tag{5.3.7-B}$$

mit noch zu bestimmenden Koeffizienten $A_{1,1}$, $A_{1,2}$ und A_2. Multiplikation beider Seiten mit $(x-1)^2(x+1) = (x-1)(x^2-1) = x^3 - x^2 - x + 1$ liefert

$$x^3 = x^3 - x^2 - x + 1 + A_{1,1}x^2 - A_{1,1} + A_{1,2}x + A_{1,2} + A_2x^2 - 2A_2x + A_2 \tag{5.3.7-C}$$

und Koeffizientenvergleich damit das lineare Gleichungssystem

$$A_{1,1} + A_2 = 1, \qquad A_{1,2} - 2A_2 = 1, \qquad A_{1,1} - A_{1,2} - A_2 = 1 \tag{5.3.7-D}$$

zur Bestimmung der Koeffizienten.

Alternativ können wir auch Grenzwerte nutzen, es gilt

$$\frac{x^3}{(x-1)^2} = (x+1)f(x) = x + 1 + A_{1,1}\frac{x+1}{x-1} + A_{1,2}\frac{x+1}{(x-1)^2} + A_2 \tag{5.3.7-E}$$

und damit

$$A_2 = \lim_{x\to -1}(x+1)f(x) = \lim_{x\to -1}\frac{x^3}{(x-1)^2} = \frac{-1}{4}. \tag{5.3.7-F}$$

Weiterhin gilt

$$\frac{x^3}{x+1} = (x-1)^2 f(x) = (x-1)^2 + A_{1,1}(x-1) + A_{1,2} + A_2\frac{(x-1)^2}{x+1} \tag{5.3.7-G}$$

und somit

$$A_{1,2} = \lim_{x\to 1}(x-1)^2 f(x) = \lim_{x\to 1}\frac{x^3}{x+1} = \frac{1}{2}. \tag{5.3.7-H}$$

Nach Differenzieren von (5.3.7-G) folgt

$$\frac{3x^2(x+1) - x^3}{(x+1)^2} = 2(x-1) + A_{1,1} + A_2\frac{2(x-1)(x+1) - (x-1)^2}{x+1} \tag{5.3.7-I}$$

und damit

$$A_{1,1} = \lim_{x\to 1}\frac{\mathrm{d}}{\mathrm{d}x}(x-1)^2 f(x) = \lim_{x\to 1}\frac{\mathrm{d}}{\mathrm{d}x}\frac{x^3}{x+1} = \lim_{x\to 1}\frac{3x^2(x+1) - x^3}{(x+1)^2} = \frac{5}{4}. \tag{5.3.7-J}$$

Die Formeln (5.3.7-E), (5.3.7-G) und (5.3.7-I) dienen hier nur der Veranschaulichung, dass die Grenzwerte wirklich die Koeffizienten liefern.

Nach Bestimmung der Koeffizienten erhalten wir damit die gewünschte Stammfunktion. Es gilt

$$\int \frac{x^3}{(x-1)^2(x+1)}\,\mathrm{d}x = x + \frac{5}{4}\ln|x-1| - \frac{1}{2}\frac{1}{x-1} - \frac{1}{4}\ln|x+1| + \mathrm{C}. \tag{5.3.7-K}$$

5.3.8. Die Integration vieler Funktionen kann auf die Integration rationaler Funktionen zurückgeführt werden. Dazu vorerst etwas Notation. Ein (reeller) *rationaler Ausdruck* $R(x_1, x_2, x_3, \ldots, x_k)$ in endlich vielen Argumenten ist jeder Ausdruck der sich aus den Variablen x_j, den Rechenoperationen in $\mathbb{R}$ und Konstanten aus $\mathbb{R}$ zusammensetzt.

Im Folgenden sei R stets ein solcher rationaler Ausdruck.

Wurzelausdrücke

⚒ **5.3.9 Beispiel.** Mit der Substitution $x = t^6$ gilt

$$\int \frac{\sqrt{x}}{\sqrt[3]{x} - \sqrt{x} - 2}\,\mathrm{d}x = \int \frac{t^3}{t^2 - t^3 - 2} 6t^5\,\mathrm{d}t = 6\int \frac{t^8}{t^2 - t^3 - 2}\,\mathrm{d}t \tag{5.3.9-A}$$

und das verbleibende Integral kann mittels Partialbruchzerlegung berechnet werden.

✂ **Ergänzung.** Ein Ausdruck der Form $R(x^{r_1}, \ldots, x^{r_k})$ mit Exponenten $r_i \in \mathbb{Q}$ und m dem Hauptnenner der r_i kann stets durch die Substitution

$$x = t^m, \qquad \mathrm{d}x = mt^{m-1}\,\mathrm{d}t \tag{5.3.9-B}$$

auf eine rationale Funktion zurückgeführt werden.

✂ **Ergänzung** (Eulersche Substitution)**.** Ein Ausdruck der Form $R(x, \sqrt{ax^2 + bx + c})$ ist entweder rational (falls $ax^2 + bx + c$ eine doppelte Nullstelle besitzt) oder er kann durch Substitution rational gemacht werden. Wir unterscheiden zwei Fälle.

Fall 1. Es gilt $a > 0$. Dann substituiert man

$$\sqrt{ax^2 + bx + c} = t - \sqrt{a}x \tag{5.3.9-C}$$

also $bx + c = t^2 - 2\sqrt{a}tx$ und damit

$$x = \frac{t^2 - c}{b + 2\sqrt{a}t}, \qquad \mathrm{d}x = \frac{2\sqrt{a}t^2 + 2tb + 2\sqrt{a}c}{(b + 2\sqrt{a}t)^2}\,\mathrm{d}t \tag{5.3.9-D}$$

was direkt auf eine rationale Funktion führt.

Fall 2. Es gilt $a < 0$ und $ax^2 + bx + c$ hat zwei reelle Nullstellen. Mit $ax^2 + bx + c = a(x - \lambda)(x - \mu)$ substituieren wir

$$\sqrt{ax^2 + bx + c} = t(x - \lambda) \tag{5.3.9-E}$$

also $a(x - \lambda)(x - \mu) = t^2(x - \lambda)^2$ und damit $a(x - \mu) = t^2(x - \lambda)$. Das führt auf

$$x = \frac{\lambda t^2 - a\mu}{t^2 - a}, \qquad \mathrm{d}x = \frac{2a(\mu - \lambda)t}{(t^2 - a)^2}\,\mathrm{d}t \tag{5.3.9-F}$$

und nach Substitution erhalten wir eine rationale Funktion.

✂ **Ergänzung.** Integranden der Form $R(x, \sqrt{ax^3 + bx^2 + cx + d})$ oder $R(x, \sqrt{ax^4 + bx^3 + cx^2 + dx + e})$ mit Quadratwurzeln aus Polynomen dritten oder vierten Grades lassen sich im allgemeinen nicht mehr elementar integrieren. Integrale dieser Form werden allgemein als *elliptische Integrale* bezeichnet. Von besonderer Bedeutung sind dabei die drei elliptischen Grundintegrale

$$\begin{aligned} &\int \frac{\mathrm{d}x}{\sqrt{(1 - x^2)(1 - k^2x^2)}}, \\ &\int \frac{x^2}{\sqrt{(1 - x^2)(1 - k^2x^2)}}, \\ &\int \frac{\mathrm{d}x}{(1 + hx^2)\sqrt{(1 - x^2)(1 - k^2x^2)}}, \end{aligned} \tag{5.3.9-G}$$

auf die die Berechnung aller anderen zurückgeführt werden kann. Liouville[3] konnte zeigen, dass diese für $0 < k < 1$ nicht elementar integrierbar sind.

[3] Joseph Liouville, 1809–1882

Trigonometrische Substitutionen

⚒ **5.3.10 Beispiel.** Die trigonometrische Substitutionen

$$t = \sin x, \qquad \mathrm{d}t = \cos x \,\mathrm{d}x, \qquad \mathrm{d}x = \frac{\mathrm{d}t}{\sqrt{1-t^2}} \tag{5.3.10-A}$$

oder

$$t = \tan x, \qquad \mathrm{d}t = (1+\tan^2 x)\,\mathrm{d}x, \qquad \mathrm{d}x = \frac{\mathrm{d}t}{1+t^2} \tag{5.3.10-B}$$

sind mitunter nützlich um Winkelfunktionen aus einem Integranden zu eliminieren.

✂ **Ergänzung** (Weierstraß-Substition)**.** Ein Ausdruck der Form $R(\sin x, \cos x)$ kann auf dem Intervall $[-\pi, \pi]$ durch die Substitution

$$t = \tan\frac{x}{2}, \qquad \mathrm{d}t = \frac{1}{2}\left(1+\tan^2\frac{x}{2}\right)\mathrm{d}x \tag{5.3.10-C}$$

auf eine rationale Funktion zurückgeführt werden. Dazu drücken wir zuerst die Winkelfunktionen durch t aus. Es gilt wegen $(\cos^2 x)(1+\tan^2 x) = 1$ und der Doppelwinkelformeln

$$\sin x = 2\sin\frac{x}{2}\cos\frac{x}{2} = \frac{2\tan\frac{x}{2}}{1+\tan^2\frac{x}{2}} = \frac{2t}{1+t^2}, \tag{5.3.10-D}$$

sowie

$$\cos x = \cos^2\frac{x}{2} - \sin^2\frac{x}{2} = \frac{1-\tan^2\frac{x}{2}}{1+\tan^2\frac{x}{2}} = \frac{1-t^2}{1+t^2}. \tag{5.3.10-E}$$

Ebenso gilt

$$\mathrm{d}x = \frac{2}{1+t^2}\,\mathrm{d}t \tag{5.3.10-F}$$

und jeder Ausdruck der Form $R(\sin x, \cos x)$ führt nach Substitution auf eine rationale Funktion.

⚒ **5.3.11 Beispiel.** Auch wenn die gerade genannte Substitution immer funktioniert, gibt es mitunter kürzere Wege zum Berechnen der Integrale. So gilt mit $t = \sin x$

$$\int \sin^2 x \cos^3 x \,\mathrm{d}x = \int t^2(1-t^2)\,\mathrm{d}t = \frac{1}{3}t^3 - \frac{1}{5}t^5 + \mathrm{C} = \frac{1}{3}\sin^3 x - \frac{1}{5}\sin^5 x + \mathrm{C} \tag{5.3.11-A}$$

oder mit $t = \cos x$

$$\begin{aligned}\int \frac{\mathrm{d}x}{\sin x \cos 2x} &= \int \frac{\mathrm{d}x}{\sin x(2\cos^2 x - 1)} = \int \frac{\mathrm{d}t}{(1-t^2)(1-2t^2)} \\ &= \frac{1}{\sqrt{2}}\ln\left|\frac{1+t\sqrt{2}}{1-t\sqrt{2}}\right| + \frac{1}{2}\ln\frac{1-t}{1+t} + \mathrm{C} \\ &= \frac{1}{\sqrt{2}}\ln\left|\frac{1+\sqrt{2}\cos x}{1-\sqrt{2}\cos x}\right| + \ln\left|\tan\frac{x}{2}\right| + \mathrm{C}.\end{aligned} \tag{5.3.11-B}$$

⚒ **5.3.12 Beispiel.** Wir geben noch eine Beispiel für die Substitution nach Weierstraß. Dazu berechnen wir für $a > |b|$ das Integral

$$\int \frac{\mathrm{d}x}{a + b\cos x}. \tag{5.3.12-A}$$

Die Substitution

$$t = \tan\frac{x}{2}, \qquad \cos x = \frac{1-t^2}{1+t^2}, \qquad \mathrm{d}x = \frac{2}{1+t^2}\,\mathrm{d}t. \tag{5.3.12-B}$$

führt auf

$$\begin{aligned}\int \frac{\mathrm{d}x}{a+b\cos x} &= \int \frac{1}{a+b\frac{1-t^2}{1+t^2}}\frac{2}{1+t^2}\,\mathrm{d}t = \int \frac{2}{a(1+t^2)+b(1-t^2)}\,\mathrm{d}t \\ &= \int \frac{2}{(a+b)+(a-b)t^2}\,\mathrm{d}t = \frac{2}{a+b}\int \frac{\mathrm{d}t}{1+\left(\frac{\sqrt{a-b}}{\sqrt{a+b}}\,t\right)^2} \\ &= \frac{2}{\sqrt{a^2-b^2}}\arctan\left(\frac{\sqrt{a-b}}{\sqrt{a+b}}\,\tan\frac{x}{2}\right)+\mathrm{C}.\end{aligned} \tag{5.3.12-C}$$

5.4 Die Taylorsche Formel mit Integralrestglied

5.4.1. Sei $f : (a,b) \to \mathbb{R}$ eine n-fach differenzierbare Funktion und $x_0 \in (a,b)$. Dann ist das n-te Taylorpolynom definiert als

$$T_{n,x_0}[f](x) = f(x_0) + \sum_{k=1}^{n} \frac{f^{(k)}(x_0)}{k!}(x-x_0)^k \tag{5.4.1-A}$$

und wir haben im letzten Kapitel gesehen, dass das Restglied

$$R_n(x,x_0) = f(x) - T_{n,x_0}[f](x-x_0) \tag{5.4.1-B}$$

nach Peano

$$\lim_{x\to x_0} = \frac{R_n(x,x_0)}{(x-x_0)^n} = 0 \tag{5.4.1-C}$$

und für $n+1$-fach differenzierbare Funktionen nach Lagrange

$$R_n(x,x_0) = \frac{f^{(n+1)}(\xi)}{(n+1)!}(x-x_0)^{n+1} \tag{5.4.1-D}$$

mit einer Zwischenstelle ξ zwischen x und x_0 erfüllt. Integralrechnung erlaubt eine weitere Darstellung des Restgliedes. Diese ist explizit, mitunter für Abschätzungen besser und für Beweise unter Einbeziehung der Taylorschen Formel zu bevorzugen.

☞ **5.4.2 Satz** (Integralrestglied). *Sei $f : (a,b) \to \mathbb{R}$ eine $(n+1)$-fach stetig differenzierbare Funktion und $x_0 \in (a,b)$. Dann gilt für das Restglied $R_n(x,x_0)$ der Taylorschen Formel die Darstellung*

$$\begin{aligned}R_n(x,x_0) &= \frac{1}{n!}\int_{x_0}^{x}(x-t)^n f^{(n+1)}(t)\,\mathrm{d}t \\ &= \frac{(x-x_0)^{n+1}}{n!}\int_0^1 \theta^n f^{(n+1)}(\theta x_0 + (1-\theta)x)\,\mathrm{d}\theta.\end{aligned} \tag{5.4.2-A}$$

Beweis. Wir zeigen dies durch Induktion über n. Für $n = 0$ ergibt die Formel von Newton–Leibniz

$$f(x) - f(x_0) = \int_{x_0}^{x} f'(t)\,\mathrm{d}t = (x - x_0)\int_0^1 f'(\theta x_0 + (1-\theta)x)\,\mathrm{d}\theta \tag{5.4.2-B}$$

und der Induktionsanfang ist gezeigt. Für den Integrationsschritt nutzen wir partielle Integration. Wir zeigen zuerst den Schritt zu $n = 1$ und dann den allgemeinen Schritt. Es gilt unter Nutzung der stetigen Differenzierbarkeit[4]

$$\begin{aligned} f(x) - f(x_0) &= \int_{x_0}^{x} f'(t)\,\mathrm{d}t = tf'(t)\Big|_{t=x_0}^{x} - \int_{x_0}^{x} tf''(t)\,\mathrm{d}t \\ &= (x-x_0)f'(x_0) + x(f'(x) - f'(x_0)) - \int_{x_0}^{x} tf''(t)\,\mathrm{d}t \\ &= (x-x_0)f'(x_0) + \int_{x_0}^{x}(x-t)f''(t)\,\mathrm{d}t. \end{aligned} \tag{5.4.2-C}$$

Für den allgemeinen Schritt nutzen wir die äquivalente zweite Schreibweise. In dieser gilt für jedes $k \le n$ für eine $(n+1)$-fach stetig differenzierbare Funktion f

$$\begin{aligned} R_{k-1}(x, x_0) &= \frac{(x-x_0)^k}{(k-1)!}\int_0^1 \theta^{k-1} f^{(k)}(\theta x_0 + (1-\theta)x)\,\mathrm{d}\theta \\ &= \frac{(x-x_0)^k}{k!}\theta^k f^{(k)}(\theta x_0 + (1-\theta)x)\Big|_{\theta=0}^{1} \\ &\qquad + \frac{(x-x_0)^{k+1}}{k!}\int_0^1 \theta^k f^{(k+1)}(\theta x_0 + (1-\theta)x)\,\mathrm{d}\theta \\ &= \frac{(x-x_0)^k}{k!} f^{(k)}(x_0) + R_k(x, x_0) \end{aligned} \tag{5.4.2-D}$$

und die Behauptung folgt mittels

$$T_{k,x_0}[f](x) = \frac{(x-x_0)^k}{k!} f^{(k)}(x_0) + T_{k-1,x_0}[f](x). \tag{5.4.2-E}$$

□

✂ **Ergänzung.** Aus dem Integralrestglied folgen die anderen Restgliedformeln. So ergibt sich direkt als Folgerung der Stetigkeit des Integranden mit dem 1. Mittelwertsatz der Integralrechnung

$$\begin{aligned} R_n(x, x_0) &= \frac{(x-x_0)^{n+1}}{n!}\int_0^1 (1-\theta)^n f^{(n+1)}(\theta x + (1-\theta)x_0)\,\mathrm{d}\theta \\ &= \frac{(x-x_0)^{n+1}}{n!}(1-\theta_*)^n f^{(n+1)}(x - (1-\theta_*)(x-x_0)) \end{aligned} \tag{5.4.2-F}$$

für ein $\theta_* \in (0,1)$ und damit die Restgliedformel von Cauchy mit $\xi = x - (1-\theta_*)(x-x_0)$. Mit dem 2. Mittelwertsatz folgt entsprechend

$$\begin{aligned} R_n(x, x_0) &= \frac{(x-x_0)^{n+1}}{n!} f^{(n+1)}(\theta_* x + (1-\theta_*)x_0)\int_0^1 (1-\theta)^n\,\mathrm{d}\theta \\ &= \frac{(x-x_0)^{n+1}}{(n+1)!} f^{(n+1)}(\xi) \end{aligned} \tag{5.4.2-G}$$

mit einem (anderen) $\theta_* \in (0,1)$ und damit die Restgliedformel von Lagrange.

[4] Wir benötigen die R-integrierbarkeit der Ableitung, diese folgt hier aus der Stetigkeit. Da nicht jede unstetige Ableitung auch integrierbar ist, bedarf es hier also einer Voraussetzung!

5.5 Uneigentliche Riemannintegrale

5.5.1. Mitunter will man unbeschränkte Funktionen integrieren oder Integrale über nicht endliche Intervalle definieren. Eine naheliegende Idee dafür sind als Grenzwerte definierte uneigentliche Integrale.

Sei dazu $f : (a, b] \to \mathbb{R}$ auf jedem Teilintervall $[a + \varepsilon, b]$ R-integrierbar. Existiert dann der Grenzwert

$$\lim_{\varepsilon \searrow 0} \int_{a+\varepsilon}^{b} f(x)\,\mathrm{d}x \tag{5.5.1-A}$$

so sagen wir, f ist *uneigentlich R-integrierbar* auf $(a, b]$ und bezeichnen den Grenzwert

$$\int_{\to a}^{b} f(x)\,\mathrm{d}x := \lim_{\varepsilon \searrow 0} \int_{a+\varepsilon}^{b} f(x)\,\mathrm{d}x \tag{5.5.1-B}$$

als *uneigentliches Integral*. Wir nutzen die Notation $\to a$ in der Integralgrenze um auf den zusätzlichen Grenzwert hinzuweisen. Entsprechend definieren wir am oberen Intervallende uneigentlichen Integrale

$$\int_{a}^{\to b} f(x)\,\mathrm{d}x := \lim_{\varepsilon \searrow 0} \int_{a}^{b-\varepsilon} f(x)\,\mathrm{d}x, \tag{5.5.1-C}$$

Integrale die an beiden Intervallgrenzen uneigentlich sind oder uneigentliche Integrale mit Intervallgrenzen $\pm\infty$

$$\int_{-\infty}^{+\infty} f(x)\,\mathrm{d}x := \lim_{L \to -\infty} \lim_{R \to +\infty} \int_{L}^{R} f(x)\,\mathrm{d}x \tag{5.5.1-D}$$

bei denen wir das $\to \pm\infty$ auch kurz $\pm\infty$ schreiben, da diese Integrale offensichtlich uneigentlich gemeint sind. Bei zweiseitig uneigentlichen Integralen ist die Reihenfolge der Grenzwertbildung unerheblich, es ergibt sich

$$\int_{\to a}^{\to b} f(x)\,\mathrm{d}x = \int_{\to a}^{c} f(x)\,\mathrm{d}x + \int_{c}^{\to b} f(x)\,\mathrm{d}x. \tag{5.5.1-E}$$

für jedes $c \in (a, b)$.

⚒ **5.5.2 Beispiele.** Als erste Beispiele betrachten wir

(i) die Fläche unter der Exponentialfunktion

$$\int_{0}^{+\infty} \mathrm{e}^{-x}\,\mathrm{d}x = \lim_{R \to \infty} \int_{0}^{R} \mathrm{e}^{-x}\,\mathrm{d}x = \lim_{R \to +\infty} \left(1 - \mathrm{e}^{-R}\right) = 1 \tag{5.5.2-A}$$

(ii) und entsprechend für die Fläche links / oberhalb der Logarithmusfunktion

$$\int_{\to 0}^{1} \ln x\,\mathrm{d}x = \lim_{\varepsilon \searrow 0} \int_{\varepsilon}^{1} \ln x\,\mathrm{d}x = \lim_{\varepsilon \searrow 0} \big[x \ln x - x\big]_{\varepsilon}^{1} = -1. \tag{5.5.2-B}$$

(iii) Wichtig sind uneigentliche Integrale von Potenzfunktionen. Für $\gamma < 1$ gilt

$$\int_{\to 0}^{1} x^{-\gamma}\,\mathrm{d}x = \left[\frac{1}{1-\gamma}x^{1-\gamma}\right]_{\to 0}^{1} = \frac{1}{1-\gamma} \tag{5.5.2-C}$$

während für alle $\gamma \geq 1$ das uneigentliche Integral divergiert.

(iv) Entsprechend gilt für $\gamma > 1$

$$\int_{1}^{+\infty} x^{-\gamma}\,\mathrm{d}x = \left[\frac{1}{1-\gamma}x^{1-\gamma}\right]_{1}^{\to+\infty} = \frac{1}{\gamma-1} \tag{5.5.2-D}$$

mit Divergenz für $\gamma \leq 1$.

5.5.3. Die Konvergenz uneigentlicher Integrale kann wie in obigen Beispielen durch direktes Ausrechnen gezeigt werden. Wie bei Reihen ist es aber mitunter interessanter Funktionen oder Werte durch Integrale zu definieren. Dann hilft als Folgerung des Cauchykriteriums

$$\int_{0}^{\to+\infty} f(x)\,\mathrm{d}x \quad \text{konvergiert} \qquad \Longleftrightarrow \qquad \int_{a}^{b} f(x)\,\mathrm{d}x \longrightarrow 0, \quad a,b \to +\infty. \tag{5.5.3-A}$$

Als Konsequenz ergibt sich wiederum ein Majorantenkriterium für uneigentliche Integrale.

☞ **5.5.4 Satz** (Majorantenkriterium). *Angenommen, f und g sind R-integrierbar auf jedem endlichen Teilintervall von $[0,+\infty)$ und es gilt $|f(x)| \leq g(x)$. Dann gilt*

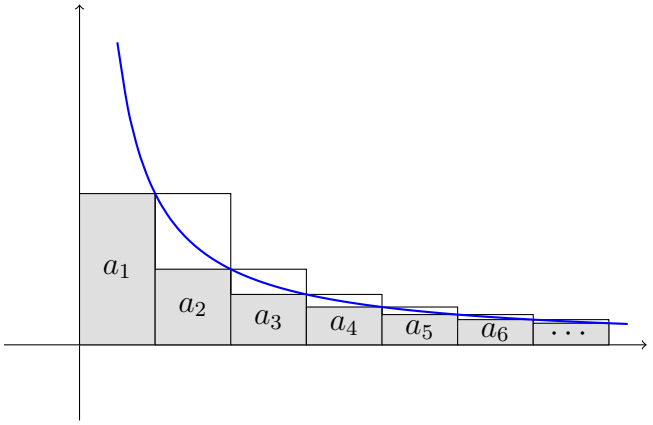
$$\int_{0}^{+\infty} g(x)\,\mathrm{d}x < \infty \qquad \Longrightarrow \qquad \int_{0}^{+\infty} f(x)\,\mathrm{d}x \quad \textit{konvergiert.} \tag{5.5.4-A}$$

Beweis. Die Aussage folgt aus dem Cauchykriterium. Es gilt für $a < b$

$$\left|\int_{a}^{b} f(x)\,\mathrm{d}x\right| \leq \int_{a}^{b} |f(x)|\,\mathrm{d}x \leq \int_{a}^{b} g(x)\,\mathrm{d}x \longrightarrow 0 \tag{5.5.4-B}$$

für $a,b \to \infty$ auf Grund der Konvergenz des Integrals $\int_0^\infty g(x)\,\mathrm{d}x$. □

Eine Anwendung uneigentlicher Integrale ist das Integralvergleichskriterium. Es verknüpft die Konvergenz von Reihen mit der Konvergenz von uneigentlichen Integralen und liefert damit insbesondere neue Vergleichsreihen für Anwendungen des Majorantenkriteriums.



☞ **5.5.5 Satz** (Integralvergleichskriterium). *Sei $f : [1, +\infty) \to \mathbb{R}$ monoton fallend, nichtnegativ und bezeichne $a_k = f(k)$. Dann gilt*

$$\sum_{k=1}^{\infty} a_k < \infty \qquad \Longleftrightarrow \qquad \int_1^{+\infty} f(x)\,\mathrm{d}x < \infty. \tag{5.5.5-A}$$

Beweis. Der Beweis folgt durch Nutzung der Partialsummen der Reihe als geeignete Ober- und Untersummen des Integrals. Auf Grund der Monotonie der Funktion f gilt

$$\int_1^N f(x)\,\mathrm{d}x \le \sum_{k=1}^{N-1} a_k \le a_1 + \int_1^{N-1} f(x)\,\mathrm{d}x \tag{5.5.5-B}$$

und die Konvergenz des Integrals impliziert die Beschränktheit der Partialsummen und damit die Konvergenz der Reihe. Umgekehrt liefert die Konvergenz der Reihe eine konvergente Majorante für das Integral (nämlich die oberhalb liegende Treppenfunktion) und die Konvergenz des Integrals folgt aus dem Majorantenkriterium 5.5.4. □

⚒ **5.5.6 Beispiel.** Die Reihe

$$\sum_{k=2}^{\infty} \frac{1}{k(\ln k)^{\alpha}} \tag{5.5.6-A}$$

konvergiert genau dann, wenn $\alpha > 1$ gilt. Zum Nachweis verwenden wir das Integralvergleichskriterium. Es gilt für $\alpha \neq 1$

$$\int_2^{+\infty} \frac{\mathrm{d}x}{x(\ln x)^{\alpha}} = \left[\frac{1}{1-\alpha}(\ln x)^{1-\alpha}\right]_2^{\to+\infty} = \begin{cases} +\infty, & \alpha < 1 \\ \frac{(\ln 2)^{1-\alpha}}{\alpha-1}, & \alpha > 1, \end{cases} \tag{5.5.6-B}$$

und für $\alpha = 1$

$$\int_2^{+\infty} \frac{\mathrm{d}x}{x \ln x} = \big[\ln \ln x\big]_2^{\to+\infty} = +\infty. \tag{5.5.6-C}$$

Es gibt uneigentliche Integrale deren Konvergenz auf Vorzeichenwechseln beruht. Konvergenzbeweise erinnern an den Beweis des Leibniz-Kriteriums für Reihen und wir geben ein wichtiges Beispiel direkt an:

⚒ **5.5.7 Beispiel.** Das uneigentliche Integral

$$\int_0^{\infty} \frac{\sin x}{x}\,\mathrm{d}x \tag{5.5.7-A}$$

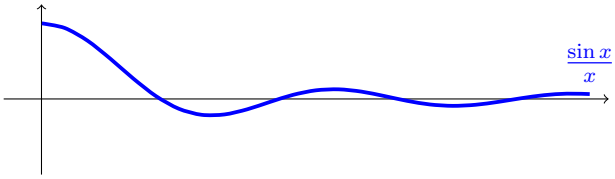
ist konvergent. Sei dazu $\pi < a < b$ und k und l so gewählt, dass $|a - 2k\pi| \le \pi$ und $|b - (2l+1)\pi| \le \pi$. Dann gilt

$$\int_a^b \frac{\sin x}{x}\,\mathrm{d}x \le \int_{2k\pi}^{(2l+1)\pi} \frac{\sin x}{x}\,\mathrm{d}x \le \int_{2k\pi}^{(2k+1)\pi} \frac{\sin x}{x}\,\mathrm{d}x \le \frac{1}{k\pi} \longrightarrow 0 \tag{5.5.7-B}$$

für $a, b \to \infty$ und ebenso mit k' und l' so gewählt, dass $|a - (2k'-1)\pi| \le \pi$ und $|b - 2l'\pi| \le \pi$ als untere Abschätzung

$$\int_a^b \frac{\sin x}{x}\,\mathrm{d}x \ge \int_{(2k'-1)\pi}^{2l'\pi} \frac{\sin x}{x}\,\mathrm{d}x \ge \int_{(2k'-1)\pi}^{2k'\pi} \frac{\sin x}{x}\,\mathrm{d}x \ge \frac{-2}{(2k'-1)\pi} \longrightarrow 0 \tag{5.5.7-C}$$

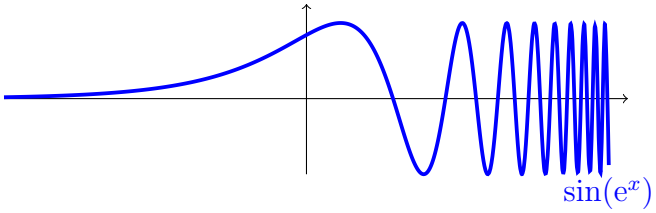
für $a, b \to \infty$. Den Zahlenwert des uneigentlichen Integrals müssen wir vorerst offenlassen.



Es gilt

$$\int_{-\infty}^{+\infty} \sin(\mathrm{e}^x)\,\mathrm{d}x = \int_{0}^{+\infty} \frac{\sin y}{y}\,\mathrm{d}y, \tag{5.5.7-D}$$

ersteres ist also auch konvergent. An diesem Beispiel sieht man, dass für die Konvergenz uneigentlicher Integrale über unbeschränkten Intervallen der Integrand nicht gegen Null streben muss, ein hinreichend schnelles Oszillieren des Integranden reicht ebenso.



Die Gammafunktion

5.5.8. Uneigentliche Integrale kann man zur Definition neuer Funktionen nutzen. Prominentestes Beispiel dazu ist die *Eulersche Gammafunktion* definiert für $x > 0$ als

$$\Gamma(x) := \int_{\to 0}^{+\infty} t^{x-1}\mathrm{e}^{-t}\,\mathrm{d}t. \tag{5.5.8-A}$$

Die Konvergenz des Integrals ergibt sich aus dem Majorantenkriterium mit den Majoranten t^{x-1} auf $[0,1]$ und $c\mathrm{e}^{-t/2}$ auf $[1,+\infty)$. Darüberhinaus gilt

$$\Gamma(1) = \int_{0}^{+\infty} \mathrm{e}^{-t}\,\mathrm{d}t = 1 \tag{5.5.8-B}$$

und mittels partieller Integration

$$\Gamma(x+1) = \int_{\to 0}^{+\infty} t^x\mathrm{e}^{-t}\,\mathrm{d}t = \Big[-t^x\mathrm{e}^{-t}\Big]_{\to 0}^{+\infty} + x\int_{\to 0}^{+\infty} t^{x-1}\mathrm{e}^{-t}\,\mathrm{d}t = x\Gamma(x). \tag{5.5.8-C}$$

Also folgt $\Gamma(2) = \Gamma(1) = 1$ und $\Gamma(3) = 2\Gamma(2) = 2$ und damit per Induktion

$$\Gamma(n+1) = n!. \tag{5.5.8-D}$$

Die Gammafunktion verallgemeinert in diesem Sinne die Fakultät von natürlichen Zahlen auf reelle Zahlen.

✂ **Ergänzung.** Weitere spezielle Werte können durchaus interessant sein. So gilt

$$\Gamma\left(\frac{1}{2}\right) = \int_{\to 0}^{+\infty} t^{-\frac{1}{2}} \mathrm{e}^{-t}\,\mathrm{d}t = 2\int_0^{+\infty} \mathrm{e}^{-s^2}\,\mathrm{d}s = \int_{-\infty}^{+\infty} \mathrm{e}^{-s^2}\,\mathrm{d}s = \sqrt{\pi} \tag{5.5.8-E}$$

mit der Substitution $s = \sqrt{t}$ und damit $\mathrm{d}s = \frac{1}{2\sqrt{t}}\,\mathrm{d}t$. Die Berechnung des letzten Integrals erfordert die Nutzung von Satz 6.3.3 über die Vertauschbarkeit mehrfacher Integrale in Kombination mit Lemma 6.1.2. Ohne diesem direkt vorzugreifen, eine mögliche Rechnung:

$$\begin{aligned}\left(\int_0^{+\infty} \mathrm{e}^{-t^2}\,\mathrm{d}t\right)\left(\int_0^{+\infty} \mathrm{e}^{-s^2}\,\mathrm{d}s\right) &= \int_0^{+\infty}\int_0^{+\infty} \mathrm{e}^{-t^2-s^2}\,\mathrm{d}t\,\mathrm{d}s\\ &= \int_0^{+\infty}\int_0^{+\infty} \mathrm{e}^{-s^2(\theta^2+1)}\,s\,\mathrm{d}\theta\,\mathrm{d}s\\ &\overset{?}{=} \int_0^{+\infty}\int_0^{+\infty} \mathrm{e}^{-s^2(\theta^2+1)}\,s\,\mathrm{d}s\,\mathrm{d}\theta\\ &= \int_0^{+\infty} \frac{\mathrm{d}\theta}{2(\theta^2+1)} = \frac{1}{2}\arctan\theta\Big|_{\theta=0}^{+\infty} = \frac{\pi}{4}\end{aligned} \tag{5.5.8-F}$$

unter Ausnutzung der Substitution $t = s\theta$ mit $\mathrm{d}t = s\,\mathrm{d}\theta$.

✂ **Ergänzung.** Als zweites ebenso, auf Euler zurückgehendes Beispiel, betrachten wir die *Eulersche Betafunktion*

$$\mathrm{B}(a,b) = \int_{\to 0}^{\to 1} t^{a-1}(1-t)^{b-1}\,\mathrm{d}t \tag{5.5.8-G}$$

für Zahlen $a, b > 0$. Das Integral ist symmetrisch in a und b, es gilt also

$$\mathrm{B}(a,b) = \mathrm{B}(b,a), \tag{5.5.8-H}$$

erfüllt eine an das Pascalsche Dreieck erinnernde Rekursionsformel

$$\begin{aligned}\mathrm{B}(a+1,b) + \mathrm{B}(a,b+1) &= \int_{\to 0}^{\to 1} \big(t^a(1-t)^{b-1} + t^{a-1}(1-t)^b\big)\,\mathrm{d}t\\ &= \int_{\to 0}^{\to 1} t^{a-1}(1-t)^{b-1}\big(t+(1-t)\big)\,\mathrm{d}t = \mathrm{B}(a,b),\end{aligned} \tag{5.5.8-I}$$

und partielle Integration liefert

$$\begin{aligned}a\mathrm{B}(a,b+1) &= a\int_{\to 0}^{\to 1} t^{a-1}(1-t)^b\,\mathrm{d}t\\ &= \Big[t^a(1-t)^b\Big]_{t\to 0}^{\to 1} + b\int_{\to 0}^{\to 1} t^a(1-t)^{b-1}\,\mathrm{d}t = b\mathrm{B}(a+1,b)\end{aligned} \tag{5.5.8-J}$$

für $a, b > 0$. Damit ergibt sich die Rekursion

$$\mathrm{B}(a,b) = \mathrm{B}(a+1,b) + \mathrm{B}(a,b+1) = \frac{a+b}{a}\mathrm{B}(a+1,b) \tag{5.5.8-K}$$

und zusammen mit $\mathrm{B}(a,1) = \frac{1}{a}$ die leicht per Induktion zu beweisende Darstellung für natürliche Zahlen als

$$\mathrm{B}(a,n) = \frac{(n-1)!}{(a+n-1)\cdots(a+1)a}. \tag{5.5.8-L}$$

Das legt eine enge Beziehung zur Gammafunktion nahe. Es gilt

$$\begin{aligned}
\Gamma(a+b)\mathrm{B}(a,b) &= \int_{\to 0}^{+\infty} \mathrm{e}^{-s} s^{a+b-1}\,\mathrm{d}s \int_{\to 0}^{\to 1} t^{a-1}(1-t)^{b-1}\,\mathrm{d}t \\
&= \int_{\to 0}^{+\infty} \mathrm{e}^{-s} s^{a+b-1} \int_{\to 0}^{\to 1} t^{a-1}(1-t)^{b-1}\,\mathrm{d}t\,\mathrm{d}s \\
&= \int_{\to 0}^{+\infty} \mathrm{e}^{-s} \int_{\to 0}^{\to 1} (st)^{a-1}(s-st)^{b-1} s\,\mathrm{d}t\,\mathrm{d}s \\
&= \int_{\to 0}^{+\infty} \int_{\to 0}^{\to s} \mathrm{e}^{-\theta}\theta^{a-1}\mathrm{e}^{-(s-\theta)}(s-\theta)^{b-1}\,\mathrm{d}\theta\,\mathrm{d}s \\
&\overset{?}{=} \int_{\to 0}^{+\infty} \int_{\to \theta}^{+\infty} \mathrm{e}^{-\theta}\theta^{a-1}\mathrm{e}^{-(s-\theta)}(s-\theta)^{b-1}\,\mathrm{d}s\,\mathrm{d}\theta \\
&= \int_{\to 0}^{+\infty} \mathrm{e}^{-\theta}\theta^{a-1} \int_{\to \theta}^{+\infty} \mathrm{e}^{-(s-\theta)}(s-\theta)^{b-1}\,\mathrm{d}s\,\mathrm{d}\theta \\
&= \Gamma(a)\Gamma(b),
\end{aligned} \tag{5.5.8-M}$$

wobei zum Vertauschen der Reihenfolge der Integrale wiederum Satz 6.3.3 im Kombination mit Lemma 6.1.2 benötigt wird.

Einige weitere Werte sind elementar ausrechenbar, so gilt

$$\mathrm{B}\left(\frac{1}{2},\frac{1}{2}\right) = \int_{\to 0}^{\to 1} \frac{\mathrm{d}t}{\sqrt{t}\sqrt{1-t}} = 2\int_{\to 0}^{\to 1} \frac{\mathrm{d}s}{\sqrt{1-s^2}} = 2\arcsin s\Big|_{s\to 0}^{\to 1} = \pi \tag{5.5.8-N}$$

und wir erhalten wiederum

$$\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi} = \int_{-\infty}^{+\infty} \mathrm{e}^{-s^2}\,\mathrm{d}s. \tag{5.5.8-O}$$

6 Funktionenfolgen und Reihen

6.1 Doppelfolgen und -Reihen

6.1.1. Für eine Doppelfolge $(x_{k,l})_{k,l\in\mathbb{N}} \in \mathbb{K}^{\mathbb{N}\times\mathbb{N}}$ kann man die beiden iterierten Grenzwerte

$$\lim_{k\to\infty}\lim_{l\to\infty} x_{k,l} \qquad \text{und} \qquad \lim_{l\to\infty}\lim_{k\to\infty} x_{k,l} \tag{6.1.1-A}$$

betrachten. Im Allgemeinen sind diese nicht gleich, wie folgendes einfache Beispiel zeigt: Es gilt

$$\lim_{k\to\infty}\left(\lim_{l\to\infty}\frac{k}{k+l}\right) = \lim_{k\to\infty} 0 = 0 \tag{6.1.1-B}$$

und ebenso

$$\lim_{l\to\infty}\left(\lim_{k\to\infty}\frac{k}{k+l}\right) = \lim_{l\to\infty} 1 = 1. \tag{6.1.1-C}$$

In diesem Abschnitt wollen wir ein einfaches Kriterium dafür angeben, dass die beiden iterierten Grenzwerte einer Doppefolge übereinstimmen. Wir formulieren das Kriterium für Konvergenz in allgemeinen metrischen Räumen, die meisten Anwendungen im Rahmen dieser Vorlesung werden allerdings Konvergenz in $\mathbb{R}$ oder $\mathbb{C}$ betreffen.

☞ **6.1.2 Lemma** (1. Doppelfolgensatz, Weierstraß). *Sei (M,d) metrischer Raum und sei $(x_{k,l})_{k,l\in\mathbb{N}} \in M^{\mathbb{N}\times\mathbb{N}}$ eine Doppelfolge. Angenommen der Grenzwert*

$$y_k = \lim_{l\to\infty} x_{k,l} \qquad \text{existiert für jedes } k\in\mathbb{N} \tag{6.1.2-A}$$

und der Grenzwert

$$z_l = \lim_{k\to\infty} x_{k,l} \qquad \text{existiert gleichmäßig in } l\in\mathbb{N}, \tag{6.1.2-B}$$

es gilt also

$$\forall_{\varepsilon>0}\exists_{K_\varepsilon}\forall_{k>K_\varepsilon}\forall_{l\in\mathbb{N}} \quad d(x_{k,l},z_l)<\varepsilon. \tag{6.1.2-C}$$

Konvergiert dann die Folge $(z_l)_{l\in\mathbb{N}}$, so auch $(y_k)_{k\in\mathbb{N}}$ und es gilt

$$\lim_{k\to\infty} y_k = \lim_{l\to\infty} z_l. \tag{6.1.2-D}$$

Beweis. Sei $\varepsilon>0$. Da der Grenzwert z_l gleichmäßig angenommen wird, gilt für alle $l\in\mathbb{N}$ und $k>K_\varepsilon$

$$d(x_{k,l},z_l)<\varepsilon. \tag{6.1.2-E}$$

Damit folgt für $l\to\infty$ unter Nutzung der Stetigkeit der Metrik d und mit dem Grenzwert $a=\lim_{l\to\infty} z_l$

$$d(y_k,a)\le\varepsilon, \qquad k>K_\varepsilon. \tag{6.1.2-F}$$

Die Folge $(y_k)_{k\in\mathbb{N}}$ damit ebenso konvergent mit Grenzwert a. □

☞ **6.1.3 Satz** (2. Doppelfolgensatz, Weierstraß). *Sei (M,d) vollständig. Existieren dann beide Grenzwerte*

$$y_k = \lim_{l\to\infty} x_{k,l} \qquad \text{und} \qquad z_l = \lim_{k\to\infty} x_{k,l} \tag{6.1.3-A}$$

gleichmäßig bezüglich k beziehungsweise l, so konvergieren beide Folgen $(y_k)_{k\in\mathbb{N}}$ und $(z_l)_{l\in\mathbb{N}}$ und die Grenzwerte stimmen überein.

Beweis. Es genügt, die Konvergenz der Folge $(z_l)_{l\in\mathbb{N}}$ nachzuweisen. Dazu nutzen wir die Dreiecksungleichung

$$d(z_n, z_m) \le d(z_n, x_{k,n}) + d(x_{k,n}, x_{k,m}) + d(x_{k,m}, z_m). \tag{6.1.3-B}$$

Da $z_l = \lim_{k\to\infty} x_{k,l}$ gleichmäßig konvergiert, existiert zu $\varepsilon > 0$ ein K_ε mit $d(z_l, x_{k,l}) < \varepsilon$ für alle $k > K_\varepsilon$ und $l \in \mathbb{N}$. Da weiterhin $x_{k,l}$ gleichmäßig bezüglich l konvergiert, existiert ebenso ein L_ε mit $d(x_{k,n}, x_{k,m}) \le d(x_{k,n}, y_k) + d(y_k, x_{k,m}) < 2\varepsilon$ für $m, n > L_\varepsilon$ und alle k. Also folgt

$$d(z_n, z_m) < 4\varepsilon \tag{6.1.3-C}$$

und die Folge $(z_l)_{l\in\mathbb{N}}$ ist Cauchy. Da (M,d) vollständig ist, folgt Konvergenz und aus Lemma 6.1.2 die Behauptung. □

☞ **6.1.4 Korollar** (Doppelreihensatz von Cauchy). *Sei $(a_{k,l})_{k,l\in\mathbb{N}} \in \mathbb{K}^{\mathbb{N}\times\mathbb{N}}$ eine Doppelfolge in $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$. Ist dann*

$$\sup_{m,n\in\mathbb{N}} \sum_{k=1}^{m}\sum_{l=1}^{n} |a_{k,l}| < \infty \tag{6.1.4-A}$$

beschränkt, so konvergieren die iterierten Doppelreihen

$$\sum_{k=1}^{\infty}\left(\sum_{l=1}^{\infty} a_{k,l}\right) \qquad \text{und} \qquad \sum_{l=1}^{\infty}\left(\sum_{k=1}^{\infty} a_{k,l}\right) \tag{6.1.4-B}$$

absolut und liefern denselben Wert.

Beweis. Nach Voraussetzung konvergiert die Doppelreihe

$$\sum_{k=1}^{\infty}\sum_{l=1}^{\infty} |a_{k,l}| = \sup_{m,n\in\mathbb{N}} \sum_{k=1}^{m}\sum_{l=1}^{n} |a_{k,l}| = \lim_{m\to\infty} \sum_{k=1}^{m}\sum_{l=1}^{m} |a_{k,l}| \tag{6.1.4-C}$$

und es existiert zu jedem $\varepsilon > 0$ ein N_ε, so dass für $m > N_\varepsilon$

$$\sum_{k=1}^{\infty}\sum_{l=1}^{\infty} |a_{k,l}| - \sum_{k=1}^{m}\sum_{l=1}^{m} |a_{k,l}| = \sum_{k=1}^{m-1}\sum_{l=m}^{\infty} |a_{k,l}| + \sum_{k=m}^{\infty}\sum_{l=1}^{m-1} |a_{k,l}| + \sum_{k=m}^{\infty}\sum_{l=m}^{\infty} |a_{k,l}| < \varepsilon. \tag{6.1.4-D}$$

gilt. Dies werden wir nutzen, um den 2. Doppelfolgensatz auf die Doppelfolge der Partialsummen

$$x_{m,n} = \sum_{k=1}^{m}\sum_{l=1}^{n} a_{k,l} \tag{6.1.4-E}$$

anzuwenden. Nach Voraussetzung besitzt diese die Grenzwerte

$$y_n = \lim_{m\to\infty} x_{m,n} = \sum_{l=1}^{n}\sum_{k=1}^{\infty} a_{k,l} \qquad \text{und} \qquad z_m = \lim_{n\to\infty} x_{m,n} = \sum_{k=1}^{m}\sum_{l=1}^{\infty} a_{k,l} \tag{6.1.4-F}$$

da die beiden auftretenden Reihen nach Voraussetzung absolut konvergieren. Beide Grenzwerte werden gleichmäßig angenommen, wir begnügen uns mit dem Nachweis des Ersten: Für $m_2 > m_1 > N_\varepsilon$ gilt

$$|x_{m_2,n} - x_{m_1,n}| = \left|\sum_{k=m_1+1}^{m_2}\sum_{l=1}^{n} a_{k,l}\right| \le \sum_{k=m_1+1}^{m_2}\sum_{l=1}^{n}|a_{k,l}| \le \sum_{k=m_1+1}^{m_2}\sum_{l=1}^{\infty}|a_{k,l}| < \varepsilon \tag{6.1.4-G}$$

gleichmäßig in n. □

✂ **Ergänzung.** Eng verwandt mit dem gerade gezeigten Doppelreihensatz ist der *große Umordnungssatz* für absolut konvergente Reihen. Wir haben im letzten Semester gesehen, dass jede absolut konvergente Reihe $\sum_{k=1}^{\infty} a_k$ durch eine Bijektion $\sigma : \mathbb{N} \to \mathbb{N}$ umgeordnet werden kann und dann auch

$$\sum_{k=1}^{\infty} a_{\sigma(k)} = \sum_{k=1}^{\infty} a_k =: \sum_{k\in\mathbb{N}} a_k \tag{6.1.4-H}$$

gilt. Oft wird diese Aussage als *kleiner Umordnungssatz* bezeichnet, wir können deutlich gravierender Umordnen. Es gilt

✂ **Satz** (Großer Umordnungssatz). *Sei $\sum_{k=1}^{\infty} a_k$ absolut konvergent. Dann konvergiert für jede disjunkte Zerlegung von $\mathbb{N}$ in abzählbar viele Teilmengen*

$$\mathbb{N} = \bigcup_{j\in\mathbb{N}} M_j, \qquad M_i \cap M_j = \varnothing \ \text{ für } i \neq j \tag{6.1.4-I}$$

die Doppelreihe

$$\sum_{j=1}^{\infty}\sum_{k\in M_j} a_k = \sum_{k=1}^{\infty} a_k \tag{6.1.4-J}$$

und liefert den Wert der Ausgangsreihe.

Beweis. Wir beweisen dies direkt, ohne uns auf die Sätze zur Grenzwertvertauschung zu beziehen. Da die Reihe $\sum_{k=1}^{\infty}|a_k|$ konvergiert, existiert zu jedem $\varepsilon > 0$ ein N_ε mit

$$\sum_{k=N_\varepsilon+1}^{\infty} |a_k| < \varepsilon. \tag{6.1.4-K}$$

Wir setzen $K = \{1, \dots, N_\varepsilon\}$. Dann sind alle Mengen $M_j \cap K$ und die Menge $J = \{j \mid M_j \cap K \neq \varnothing\}$ endlich. Weiter konvergieren alle Reihen $\sum_{k\in M_j} a_k$ absolut und

$$\sum_{j>\max J}\left|\sum_{M_j} a_k\right| \le \sum_{k\notin K}|a_k| < \varepsilon \tag{6.1.4-L}$$

liefert die Konvergenz der Doppelreihe. Da auf Grund der Endlichkeit der Summen

$$\sum_{k\in K} a_k = \sum_{j\in J}\sum_{k\in M_j\cap K} a_k \tag{6.1.4-M}$$

gilt, folgt nach Streichen dieser endlich vielen Summanden auf beiden Seiten

$$\left|\sum_{j=1}^{\infty}\sum_{k\in M_j} a_k - \sum_{k=1}^{\infty} a_k\right| = \left|\sum_{j=1}^{\infty}\sum_{k\in M_j\setminus K} a_k - \sum_{k\notin K} a_k\right| \le \sum_{j=1}^{\infty}\sum_{k\in M_j\setminus K}|a_k| + \sum_{k\notin K}|a_k| < 2\varepsilon \tag{6.1.4-N}$$

und somit die Behauptung. □

6.2 Gleichmäßige Konvergenz und Stetigkeit

6.2.1. Seien (M_1, d_1) und (M_2, d_2) metrische Räume und seien $f_k : M_1 \to M_2$ gegebene stetige Funktionen für $k \in \mathbb{N}$. Angenommen, für jedes $x \in M_1$ konvergiert

$$f(x) := \lim_{k\to\infty} f_k(x) \tag{6.2.1-A}$$

punktweise in M_2. Dann ist die sich ergebende Funktion im Allgemeinen nicht stetig. Allerdings liefert der Doppelfolgensatz in Form von Lemma 6.1.2 ein einfaches Kriterium für die Stetigkeit der Grenzfunktion f. Sei dazu $(x_l)_{l\in\mathbb{N}}$ eine konvergente Folge aus M_1. Da f_k stetig ist, existiert damit für jedes k der Grenzwert

$$\lim_{l\to\infty} f_k(x_l) = f_k(\lim_{l\to\infty} x_l) = f_k(x), \qquad x = \lim_{l\to\infty} x_l, \tag{6.2.1-B}$$

und es folgt

$$\lim_{k\to\infty}\lim_{l\to\infty} f_k(x_l) = \lim_{k\to\infty} f_k(x) = f(x). \tag{6.2.1-C}$$

Vertauschbarkeit der beiden Grenzwerte würde jetzt auf

$$\lim_{l\to\infty} f(x_l) = \lim_{l\to\infty}\lim_{k\to\infty} f_k(x_l) = \lim_{k\to\infty}\lim_{l\to\infty} f_k(x_l) = f(x) \tag{6.2.1-D}$$

und damit die Stetigkeit von f führen. Dazu genügt in der Doppelfolge $f_k(x_l)$ gleichmäßige Konvergenz bezüglich l zu fordern. Dies motiviert zu folgender Definition.

✍ **6.2.2 Definition.** *Eine Folge von Funktionen $f_k : M_1 \to M_2$ konvergiert* gleichmäßig *gegen eine Grenzfunktion $f : M_1 \to M_2$, falls*

$$\forall_{\varepsilon>0}\exists_{N_\varepsilon}\forall_{k>N_\varepsilon}\forall_{x\in M_1} \quad d_2(f_k(x), f(x)) < \varepsilon \tag{6.2.2-A}$$

gilt.

☞ **6.2.3 Satz.** *Seien (M_1, d_1) und (M_2, d_2) metrische Räume und $f_k : M_1 \to M_2$ stetige Funktionen für alle $k \in \mathbb{N}$. Konvergiert dann f_k gleichmäßig gegen eine Funktion $f : M_1 \to M_2$, so ist f stetig.*

Beweis. Wir zeigen dies direkt. Sei dazu $\varepsilon > 0$ und $x_0 \in M_1$. Dann existiert auf Grund der gleichmäßigen Konvergenz der f_k ein N_ε, so dass für jedes $k > N_\varepsilon$

$$\sup_{x\in M_1} d_2(f_k(x), f(x)) < \frac{\varepsilon}{3} \tag{6.2.3-A}$$

gilt. Wir wählen ein solches k. Da f_k stetig ist, gibt es weiter ein δ, so dass

$$d_1(x, x_0) < \delta \qquad \Longrightarrow \qquad d_2(f_k(x), f_k(x_0)) < \frac{\varepsilon}{3} \tag{6.2.3-B}$$

gilt. Also folgt mit der Dreiecksungleichung

$$\begin{aligned} d_2(f(x), f(x_0)) &\le d_2(f(x), f_k(x)) + d_2(f_k(x), f_k(x_0)) + d_2(f_k(x_0), f(x_0)) \\ &< \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon \end{aligned} \tag{6.2.3-C}$$

und f ist in x_0 stetig. Da x_0 beliebig war, ist f in jedem $x_0 \in M_1$ stetig und die Behauptung folgt. □

6.2.4. Wir formulieren das gerade Gezeigte für reell- oder komplexwertige Funktionen auf einem kompakten Intervall $K = [a, b]$ (oder allgemeiner einer kompakten Menge $K \subset M$ eines metrischen Raumes) etwas um. Dazu bezeichnen wir mit

$$\mathrm{C}(K) := \{f : K \to \mathbb{C} \mid f \text{ stetig auf } K\} \tag{6.2.4-A}$$

die Menge der stetigen Funktionen auf K und führen auf $\mathrm{C}(K)$ durch

$$d_\infty(f, g) = \max_{x \in K} |f(x) - g(x)| \tag{6.2.4-B}$$

eine Metrik ein. Diese ist wohldefiniert, da nach dem Satz von Weierstraß stetige Funktionen auf Kompakta beschränkt sind und ihre Extremwerte auch annehmen.

Dann entspricht gleichmäßige Konvergenz gerade der Konvergenz in diesem metrischen Raum. Es gilt, als direkte Konsequenz von Satz 6.2.3

Satz. *Der metrische Raum* $(\mathrm{C}(K), d_\infty)$ *ist vollständig.*

Beweis. Jede Cauchyfolge (f_k) bezüglich der Metrik d_∞ ist punktweise für jedes $x \in K$ Cauchyfolge in $\mathbb{C}$ und damit konvergent. Sei $f(x) = \lim_{k\to\infty} f_k(x)$. Dann impliziert Satz 6.2.3 die Stetigkeit von f, also $f \in \mathrm{C}(K)$, und die Cauchyfolgeneigenschaft impliziert Konvergenz bezüglich d_∞; für $k, l > K_\varepsilon$ gilt für jedes $y \in K$ und mit $l \to \infty$

$$|f_l(y) - f_k(y)| \le \max_{x \in K} |f_l(x) - f_k(x)| = d_\infty(f, g) < \varepsilon, \tag{6.2.4-C}$$

und für $l \to \infty$ folgt $|f(y) - f_k(y)| \le \varepsilon$, also $d_\infty(f, f_k) < \varepsilon$. □

Um die gleichmäßige Konvergenz von Reihen nachzuweisen helfen gleichmäßige Majoranten. Dies liefert damit natürlich ebenso absolute Konvergenz.

☞ **6.2.5 Satz** (Weierstraß'scher M-Test)**.** *Seien* $f_k : M \to \mathbb{C}$ *Funktionen. Angenommen, es gilt* $|f_k(x)| \le a_k$*. Dann impliziert*

$$\sum_{k=1}^{\infty} a_k < \infty \qquad \Longrightarrow \qquad \sum_{k=1}^{\infty} f_k(x) \quad \textit{konvergiert absolut und gleichmäßig.} \tag{6.2.5-A}$$

Beweis. Wir zeigen, dass die Folge der Partialsummen gleichmäßig Cauchy ist. Dazu nutzen wir für $m > n$

$$\left| \sum_{k=n}^{m} f_k(x) \right| \le \sum_{k=n}^{m} |f_k(x)| \le \sum_{k=n}^{m} a_k \longrightarrow 0 \tag{6.2.5-B}$$

für $m, n \to \infty$ auf Grund der Konvergenz der Reihe $\sum_{k=1}^{\infty} a_k$. □

6.3 Integration und Differentiation

☞ **6.3.1 Satz.** *Seien $f_k \in \mathfrak{R}[a,b]$ R-integrierbar und sei $f_k(x)$ gleichmäßig konvergent gegen $f(x)$. Dann gilt $f \in \mathfrak{R}[a,b]$ und*

$$\lim_{k\to\infty} \int_a^b f_k(x)\,\mathrm{d}x = \int_a^b f(x)\,\mathrm{d}x. \tag{6.3.1-A}$$

Beweis. Sei $\varepsilon > 0$ gegeben. Dann existiert ein K, so dass für $k > K$

$$\sup_{x\in[a,b]} |f(x) - f_k(x)| < \varepsilon \tag{6.3.1-B}$$

gilt. Da f_k R-integrierbar ist, existiert damit weiterhin eine Zerlegung $\mathfrak{Z}$ mit

$$\mathcal{O}(f_k;\mathfrak{Z}) - \mathcal{U}(f_k;\mathfrak{Z}) < \varepsilon. \tag{6.3.1-C}$$

Da sich aber auf jedem Teilintervall der Zerlegung Suprema und Infima von f höchstens um ε von denen von f_k unterscheiden, folgt

$$\mathcal{O}(f;\mathfrak{Z}) \le \mathcal{O}(f_k,\mathfrak{Z}) + \varepsilon(b-a) \tag{6.3.1-D}$$

und ebenso

$$\mathcal{U}(f;\mathfrak{Z}) \ge \mathcal{U}(f_k,\mathfrak{Z}) - \varepsilon(b-a) \tag{6.3.1-E}$$

und es folgt

$$\mathcal{O}(f;\mathfrak{Z}) - \mathcal{U}(f;\mathfrak{Z}) < \varepsilon + 2\varepsilon(b-a). \tag{6.3.1-F}$$

Da ε beliebig war, folgt die R-integrierbarkeit von f und die Gleichheit der Integrale. □

6.3.2. Ist $f : [a,b]\times[c,d] \to \mathbb{R}$ stetig, so impliziert die Kompaktheit von $[a,b]\times[c,d]$ die gleichmäßige Stetigkeit und für jedes $\varepsilon > 0$ gibt es ein $\delta > 0$ mit

$$|x_1 - x_2| + |y_1 - y_2| < \delta \qquad \Longrightarrow \qquad |f(x_1,y_1) - f(x_2,y_2)| < \varepsilon. \tag{6.3.2-A}$$

Damit konvergiert aber für jedes $x_* \in [a,b]$

$$\lim_{x\to x_*} f(x,y) = f(x_*,y) \tag{6.3.2-B}$$

gleichmäßig bezüglich y und nach Satz 6.3.1 ist

$$x \mapsto \int_c^d f(x,y)\,\mathrm{d}y \tag{6.3.2-C}$$

stetig auf $[a,b]$ und somit insbesondere wiederum R-integrierbar. Wir können also das iterierte Integral

$$\int_a^b \int_c^d f(x,y)\,\mathrm{d}y\,\mathrm{d}x \tag{6.3.2-D}$$

bilden. Der nachfolgende Satz garantiert, dass die Reihenfolge der Integration unter der Stetigkeitsvoraussetzung an f keine Rolle spielt.

☞ **6.3.3 Satz** (Vertauschen von Integralen). *Angenommen, $f : [a,b] \times [c,d] \to \mathbb{R}$ ist stetig. Dann gilt*

$$\int_a^b \int_c^d f(x,y)\,\mathrm{d}y\,\mathrm{d}x = \int_c^d \int_a^b f(x,y)\,\mathrm{d}x\,\mathrm{d}y. \tag{6.3.3-A}$$

Beweis. Wir betrachten allgemeiner die obere Grenze d als variabel und untersuchen die Funktion

$$H(z) = \int_a^b \int_c^z f(x,y)\,\mathrm{d}y\,\mathrm{d}x, \qquad z \in [c,d]. \tag{6.3.3-B}$$

Diese erfüllt $H(c) = 0$ und ist, da wegen (6.3.2-A) für $h \le \delta$

$$\begin{aligned} \frac{H(z+h)-H(z)}{h} - \int_a^b f(x,z)\,\mathrm{d}x &= \int_a^b \left(\frac{1}{h}\int_z^{z+h} f(x,y)\,\mathrm{d}y\right) - f(x,z)\,\mathrm{d}x \\ &\le \int_a^b \left(\frac{1}{h}\int_z^{z+h} (f(x,z)+\varepsilon)\;\mathrm{d}y\right) - f(x,z)\,\mathrm{d}x \\ &\le (b-a)\varepsilon \end{aligned} \tag{6.3.3-C}$$

und entsprechend

$$\frac{H(z+h)-H(z)}{h} - \int_a^b f(x,z)\,\mathrm{d}x \ge -(b-a)\varepsilon \tag{6.3.3-D}$$

gilt, differenzierbar mit

$$\frac{\mathrm{d}}{\mathrm{d}z}H(z) = \int_a^b f(x,z)\,\mathrm{d}x. \tag{6.3.3-E}$$

Weiter gilt nach dem Hauptsatz 5.2.2 der Differential- und Integralrechnung

$$\int_a^b f(x,z)\,\mathrm{d}x = \frac{\mathrm{d}}{\mathrm{d}z}\int_c^z \int_a^b f(x,y)\,\mathrm{d}x\,\mathrm{d}y \tag{6.3.3-F}$$

und damit

$$\int_a^b \int_c^z f(x,y)\,\mathrm{d}y\,\mathrm{d}x = H(z) = \int_c^z \int_a^b f(x,y)\,\mathrm{d}x\,\mathrm{d}y + \mathrm{C}. \tag{6.3.3-G}$$

Speziell mit $z = c$ folgt $\mathrm{C} = 0$ und damit die Behauptung. □

☞ **6.3.4 Korollar** (Leibnizsche Regel). *Angenommen $g : [a,b] \times [c,d] \to \mathbb{R}$ ist stetig, für jedes feste x nach der Variablen y differenzierbar und die (partielle) Ableitung*

$$g_y(x,y) := \frac{\partial g(x,y)}{\partial y} := \lim_{h\to 0} \frac{g(x,y+h)-g(x,y)}{h} \tag{6.3.4-A}$$

auf $[a,b] \times [c,d]$ stetig. Dann gilt

$$\frac{\mathrm{d}}{\mathrm{d}y}\int_a^b g(x,y)\,\mathrm{d}x = \int_a^b g_y(x,y)\,\mathrm{d}x. \tag{6.3.4-B}$$

Beweis. Folgt aus obigem Beweis mit $g(x,y) = \int_c^y g_y(x,\eta)\,\mathrm{d}\eta$. □

☞ **6.3.5 Satz.** *Sei $b-a<\infty$. Angenommen, die Funktionen $f_k : (a,b) \to \mathbb{R}$ sind differenzierbar, es gibt ein $y \in (a,b)$ für welches die Reihe $\sum_{k=1}^{\infty} f_k(y)$ konvergiert und die Reihe der Ableitungen $\sum_{k=1}^{\infty} f_k'(x)$ konvergiert gleichmäßig gegen eine Funktion G,*

$$\sum_{k=1}^{\infty} f_k'(x) = G(x). \tag{6.3.5-A}$$

Dann ist auch $\sum_{k=1}^{\infty} f_k(x)$ gleichmäßig konvergent, die Grenzfunktion F

$$F(x) = \sum_{k=1}^{\infty} f_k(x) \tag{6.3.5-B}$$

differenzierbar und es gilt

$$F'(x) = G(x) \qquad \text{auf } (a,b). \tag{6.3.5-C}$$

Beweis. Wir zeigen, dass die Reihe

$$\sum_{k=1}^{\infty} \frac{f_k(x) - f_k(x_0)}{x - x_0} \tag{6.3.5-D}$$

für jedes x_0 gleichmäßig bezüglich x konvergiert. Speziell mit $x_0 = y$ folgt daraus die gleichmäßige Konvergenz der Reihe

$$\sum_{k=1}^{\infty} (f_k(x) - f_k(y)) = \sum_{k=1}^{\infty} f_k(x) - \sum_{k=1}^{\infty} f_k(y) \tag{6.3.5-E}$$

und damit die gleichmäßige Konvergenz der Reihe $F(x) = \sum_{k=1}^{\infty} f_k(x)$. Weiter folgt, dass es sich bei (6.3.5-D) für $x \neq x_0$ um den Differenzenquotienten

$$\frac{F(x) - F(x_0)}{x - x_0} \tag{6.3.5-F}$$

handelt und, da die Reihe gleichmäßig konvergiert, somit auch

$$\lim_{x\to x_0} \frac{F(x) - F(x_0)}{x - x_0} = \sum_{k=1}^{\infty} \left(\lim_{x\to x_0} \frac{f_k(x) - f_k(x_0)}{x - x_0} \right) = \sum_{k=1}^{\infty} f_k'(x_0) = G(x_0) \tag{6.3.5-G}$$

und damit die Behauptung.

Es bleibt die gleichmäßige Konvergenz der Reihe (6.3.5-D) zu zeigen. Da die Reihe $\sum_k f_k'(x)$ gleichmäßig konvergiert, existiert zu gegebenem $\varepsilon > 0$ ein N, so dass für alle $n > m > N$

$$\left| \sum_{k=m}^{n} f_k'(x) \right| < \varepsilon \tag{6.3.5-H}$$

gilt. Wählt man solche m und n und betrachtet die Hilfsfunktion

$$H_{m,n}(x) = \sum_{k=m}^{n} f_k(x) \qquad \text{mit} \qquad H'_{m,n}(x) = \sum_{k=m}^{n} f'_k(x), \tag{6.3.5-I}$$

so folgt mit dem ersten Mittelwertsatz der Differentialrechnung

$$\sum_{k=m}^{n} \frac{f_k(x) - f_k(x_0)}{x - x_0} = \frac{H_{m,n}(x) - H_{m,n}(x_0)}{x - x_0} = H'_{m,n}(\xi) \tag{6.3.5-J}$$

für ein ξ zwischen x_0 und x und damit

$$\left| \sum_{k=m}^{n} \frac{f_k(x) - f_k(x_0)}{x - x_0} \right| = |H'_{m,n}(\xi)| < \varepsilon. \tag{6.3.5-K}$$

Also folgt die gleichmäßige Konvergenz von (6.3.5-D). □

Statt über Reihen kann man auch direkt über die Folgen der Partialsummen sprechen und erhält

☞ **6.3.6 Korollar.** *Angenommen, die Funktionen $F_k : (a, b) \to \mathbb{R}$ sind differenzierbar, es gibt ein $y \in (a, b)$ für welches die Folge $\lim_{k\to\infty} F_k(y)$ konvergiert und die Folge der Ableitungen konvergiert gleichmäßig gegen eine Funktion G,*

$$\lim_{k\to\infty} F'_k(x) = G(x). \tag{6.3.6-A}$$

Dann konvergiert $F_k(x)$ gleichmäßig bezüglich x und die Grenzfunktion

$$F(x) = \lim_{k\to\infty} F_k(x) \tag{6.3.6-B}$$

ist differenzierbar und es gilt

$$F'(x) = G(x) \qquad \text{auf } (a, b). \tag{6.3.6-C}$$

6.4 Potenzreihen

6.4.1. Eine (komplexe) *Potenzreihe* in einer Variablen $z \in \mathbb{C}$ ist eine Reihe der Form

$$\sum_{k=1}^{\infty} a_k (z - z_0)^k \tag{6.4.1-A}$$

mit Koeffizienten $a_k \in \mathbb{C}$ und mit Entwicklungspunkt $z_0 \in \mathbb{C}$. Reihen dieser Bauart mit reellen Variablen sind uns schon in Form von Taylorreihen begegnet, jetzt können wir einige allgemeine Sätze dafür nachliefern. Zur Untersuchung der Konvergenz bietet sich das Wurzelkriterium an.

☞ **6.4.2 Satz** (Cauchy–Hadamard). *Gegeben sei die Potenzreihe*

$$\sum_{k=1}^{\infty} a_k (z - z_0)^k. \tag{6.4.2-A}$$

Sei weiter $\rho \in [0, +\infty]$ *definiert durch*

$$\frac{1}{\rho} = \limsup_{k\to\infty} \sqrt[k]{|a_k|}. \tag{6.4.2-B}$$

Dann gilt

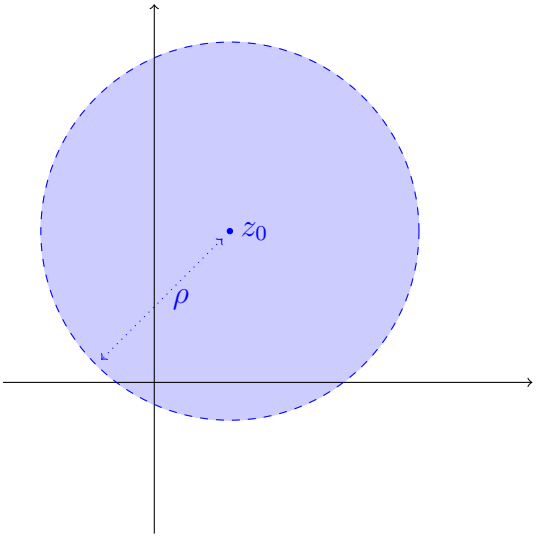
(i) *für* $|z - z_0| < \rho$ *konvergiert die Reihe* (6.4.2-A) *absolut;*

(ii) *für* $0 < \rho' < \rho$ *konvergiert die Reihe* (6.4.2-A) *gleichmäßig auf* $|z - z_0| \le \rho'$ *und bestimmt damit auf der Kreisscheibe* $|z - z_0| < \rho$ *eine stetige Funktion*

$$f(z) = \sum_{k=1}^{\infty} a_k (z - z_0)^k; \tag{6.4.2-C}$$

(iii) *die Reihe ist für* $|z - z_0| > \rho$ *divergent.*

Die Zahl ρ *wird als* Konvergenzradius *der Potenzreihe bezeichnet.*



Beweis. Aussagen **(i)** und **(iii)** folgen direkt aus dem Wurzelkriterium. Dieses liefert absolute Konvergenz für

$$\limsup_{k\to\infty} \sqrt[k]{|a_k||z - z_0|^k} = |z - z_0| \limsup_{k\to\infty} \sqrt[k]{|a_k|} = \frac{|z - z_0|}{\rho} < 1 \tag{6.4.2-D}$$

und damit für $|z - z_0| < \rho$ und entsprechend Divergenz für $|z - z_0| > \rho$. Für **(ii)** nutzen wir den Weierstraß' M-Test. Nach Definition des lim sup gibt es zu gegebenem $\varepsilon > 0$ mit $\rho' + \varepsilon < \rho$ ein K, so dass für alle $k > K$

$$|a_k| < (\rho' + \varepsilon)^{-k} \tag{6.4.2-E}$$

gilt. Also ist für $|z - z_0| < \rho'$

$$\sum_{k>K} \left(\frac{\rho'}{\rho' + \varepsilon} \right)^k \tag{6.4.2-F}$$

eine konvergente Majorante. Da alle Partialsummen Polynome und damit stetig sind, ist die Funktion f als gleichmäßiger Grenzwert stetig auf jeder Menge $|z - z_0| \leq \rho'$ und damit auf $|z - z_0| < \rho$. □

⚒ **6.4.3 Beispiel.** Für Punkte z auf der Kreislinie $|z - z_0| = \rho$ kann im Allgemeinen keine Aussage getroffen werden. Dazu betrachte man die drei Reihen

$$\sum_{k=0}^{\infty} z^k, \qquad \sum_{k=1}^{\infty} \frac{z^k}{k}, \qquad \sum_{k=1}^{\infty} \frac{z^k}{k^2}. \tag{6.4.3-A}$$

Für alle drei Reihen ergibt der Konvergenzradius $\rho = 1$ und wir erhalten Konvergenz im Inneren des Einheitskreises $|z| < 1$, Divergenz auf $|z| > 1$, und ein sehr unterschriedliches Verhalten auf der Linie $|z| = 1$. Während für $|z| = 1$ die Glieder der ersten Reihe keine Nullfolge bilden und die Reihe damit auf der gesamten Kreislinie divergiert, ist die zweite zumindest für $z = 1$ als harmonische Reihe divergent und für $z = -1$ nach dem Leibnizkriterium konvergent. Die dritte Reihe besitzt $\sum \frac{1}{k^2}$ als konvergente Majorante und konvergiert damit für jedes z mit $|z| = 1$.

✂ **Ergänzung.** Wir wollen das zweite Beispiel noch einmal aufgreifen und etwas verallgemeinern. Angenommen a_k ist monotone Nullfolge und $\sum_k a_k$ divergent. Dann hat die Potenzreihe

$$\sum_{k=0}^{\infty} a_k z^k \tag{6.4.3-B}$$

(auf Grund der Beschränktheit der Nullfolge $(a_k)_{k\in\mathbb{N}}$) Konvergenzradius 1. Auf dem Kreis $|z| = 1$ konvergiert diese Reihe genau dann, wenn $z \neq 1$ gilt. Dazu nutzen wir, dass die Partialsummen der geometrischen Reihe

$$\sum_{k=0}^{n} z^k = \frac{z^{n+1} - 1}{z - 1} \tag{6.4.3-C}$$

für alle $z \neq 1$ beschränkt sind. Es gilt

$$\left| \frac{z^{n+1} - 1}{z - 1} \right| \leq \frac{2}{|z - 1|}. \tag{6.4.3-D}$$

Weiter folgt mittels partieller Summation

$$\begin{aligned} \sum_{k=0}^{n} a_k z^k &= a_n \sum_{k=0}^{n} z^k + (a_{n-1} - a_n) \sum_{k=0}^{n-1} z^k + \cdots + (a_0 - a_1) z^0 \\ &= a_n \sum_{k=0}^{n} z^k + \sum_{k=0}^{n-1} (a_k - a_{k+1}) \sum_{l=0}^{k} z^l = a_n \frac{z^{n+1} - 1}{z - 1} + \sum_{k=0}^{n-1} (a_k - a_{k+1}) \frac{z^{k+1} - 1}{z - 1}. \end{aligned} \tag{6.4.3-E}$$

Der erste Summand strebt für $n \to \infty$ gegen Null, für den zweiten ist Konvergenz zu zeigen. Auf Grund der Monotonie der Folge $(a_k)_{k\in\mathbb{N}}$ ist $a_k - a_{k+1} > 0$ und damit

$$\sum_{k=0}^{\infty} (a_k - a_{k+1}) = a_0 \tag{6.4.3-F}$$

absolut konvergent. Damit ist $\frac{2(a_k-a_{k+1})}{|z-1|}$ für $z \neq 1$ eine konvergente Majorante für die Reihe

$$\sum_{k=0}^{\infty}(a_k - a_{k+1})\frac{z^{k+1}-1}{z-1} \tag{6.4.3-G}$$

und Konvergenz folgt.

Zum Berechnen der Reihenwerte auf dem Rand des Konvergenzkreises hilft mitunter der folgende Satz. Wir formulieren ihn nur für Konvergenzradius 1 und den Randpunkt 1, da dies den Beweis kürzer und eleganter macht. Analog gilt die Aussage für jeden anderen Randpunkt des Konvergenzkreises einer Potenzreihe in dem die Reihe konvergiert.

☞ **6.4.4 Satz** (Abelscher Grenzwertsatz). *Angenommen, die Reihe $\sum_{k=0}^{\infty} a_k$ konvergiert. Dann besitzt die Potenzreihe*

$$f(z) = \sum_{k=0}^{\infty} a_k z^k \tag{6.4.4-A}$$

mindestens den Konvergenzradius 1 *und es gilt*

$$\lim_{z\to 1} f(z) = \sum_{k=0}^{\infty} a_k. \tag{6.4.4-B}$$

Beweis. Da die a_k eine Nullfolge bilden, sind sie insbesondere beschränkt und die Konvergenzradiusformel liefert $\rho \geq 1$. Durch Addition einer Konstanten zu a_0 kann darüberhinaus erreicht werden, dass $\sum_{k=0}^{\infty} a_k = 0$ gilt. Wir nehmen dies deshalb im Beweis an.

Es gilt

$$f(z) = \sum_{k=0}^{\infty} a_k z^k = \sum_{k=0}^{\infty}(A_k - A_{k-1})z^k = (1-z)\sum_{k=0}^{\infty} A_k z^k \tag{6.4.4-C}$$

mit den Partialsummen $A_n = \sum_{k=0}^{n} a_k$ und $A_{-1} = 0$. Da $A_n \longrightarrow 0$ gilt, gibt es zu jedem $\varepsilon > 0$ ein N mit $|A_n| < \frac{1}{2}\varepsilon$ für $n > N$. Damit folgt für reelle $0 < z < 1$

$$\left|(1-z)\sum_{k=0}^{\infty} A_k z^k\right| \leq (1-z)\sum_{k=0}^{N-1}|A_k| + \frac{\varepsilon}{2}(1-z)\sum_{k=N}^{\infty} z^k \leq (1-z)\sum_{k=0}^{N-1}|A_k| + \frac{\varepsilon}{2} \tag{6.4.4-D}$$

und für z so nah bei 1, dass $|1-z| < \frac{\varepsilon}{2}\left(\sum_{k=0}^{N-1}|A_k|\right)^{-1}$ gilt folgt

$$|f(z)| = \left|(1-z)\sum_{k=0}^{\infty} A_k z^k\right| < \varepsilon. \tag{6.4.4-E}$$

Also gilt die Behauptung $f(z) \longrightarrow 0$. □

⚒ **6.4.5 Beispiele.** Wir erhalten damit einige interessante Reihenwerte.

(i) Die Logarithmusreihe kann durch

$$\ln(1-x) = -\int_0^x \frac{\mathrm{d}t}{1-t} = -\int_0^x \sum_{k=0}^{\infty} t^k \,\mathrm{d}t = -\sum_{k=0}^{\infty} \frac{x^{k+1}}{k+1} \tag{6.4.5-A}$$

für $|x| < 1$ bestimmt werden und besitzt den Konvergenzradius 1. Sie konvergiert nach Leibniz-Kriterium für $x = -1$, weshalb der Abelsche Grenzwertsatz

$$-\sum_{k=1}^{\infty} \frac{(-1)^k}{k} = \ln 2 \tag{6.4.5-B}$$

impliziert.

(ii) Die Reihendarstellung des Arcustangens ergibt sich ebenso durch Integration einer geometrischen Reihe

$$\arctan x = \int_0^x \frac{\mathrm{d}t}{1+t^2} = \int_0^x \sum_{k=0}^{\infty} (-t^2)^k \,\mathrm{d}t = \sum_{k=0}^{\infty} \frac{(-1)^k}{2k+1} x^{2k+1} \tag{6.4.5-C}$$

für $|x| < 1$. Sie besitzt ebenso den Konvergenzradius 1 und konvergiert für $x = 1$ nach dem Leibniz-Kriterium. Damit erhalten wir

$$\sum_{k=0}^{\infty} \frac{(-1)^k}{2k+1} = \arctan 1 = \frac{\pi}{4}. \tag{6.4.5-D}$$

6.4.6. Speziell im Reellen konvergieren Potenzreihen

$$f(x) = \sum_{k=0}^{\infty} a_k (x - x_0)^k, \qquad a_k \in \mathbb{R}, \tag{6.4.6-A}$$

damit auf Intervallen $(x_0 - \rho, x_0 + \rho)$ und definieren dort stetige Funktionen. Da Potenzreihen auf jedem kompakten Teilintervall gleichmäßig konvergieren und die formal abgeleitete Reihe

$$\sum_{k=1}^{\infty} k a_k (x - x_0)^{k-1} \tag{6.4.6-B}$$

wegen $\sqrt[k]{k} \longrightarrow 1$ und damit

$$\limsup_{k\to\infty} \sqrt[k]{k|a_k|} = \lim_{k\to\infty} \sqrt[k]{k} \limsup_{k\to\infty} \sqrt[k]{|a_k|} = \limsup_{k\to\infty} \sqrt[k]{|a_k|} \tag{6.4.6-C}$$

denselben Konvergenzradius besitzt, impliziert Satz 6.3.5 die Differenzierbarkeit von f und es gilt

$$f'(x) = \sum_{k=1}^{\infty} k a_k (x - x_0)^{k-1}. \tag{6.4.6-D}$$

Das Argument kann man beliebig iterieren, jede durch eine Potenzreihe definierte Funktion ist in ihrem Konvergenzintervall beliebig oft differenzierbar und es gilt

$$a_k = \frac{f^{(k)}(x_0)}{k!}, \tag{6.4.6-E}$$

die Potenzreihe ist also stets die Taylorreihe der durch sie dargestellten Funktion. Insbesondere sind die Koeffizienten einer konvergenten Potenzreihe eindeutig durch die durch die Reihe dargestellte Funktion bestimmt.

6.5 Fourierreihen

6.5.1. Sei $a_0 \in \mathbb{R}$ und seien $a_k, b_k \in \mathbb{R}$ so gewählt, dass

$$\sum_{k=1}^{\infty} |a_k| < \infty \qquad \text{und} \qquad \sum_{k=1}^{\infty} |b_k| < \infty \tag{6.5.1-A}$$

gilt. Dann sind die beiden Reihen in

$$f(x) = a_0 + \sum_{k=1}^{\infty} a_k \cos(kx) + \sum_{k=1}^{\infty} b_k \sin(kx) \tag{6.5.1-B}$$

auf Grund von Satz 6.2.5 gleichmäßig konvergent und bestimmen damit ein stetige Funktion f. Diese ist weiterhin offenbar periodisch mit Periode 2π,

$$f(x + 2\pi) = f(x). \tag{6.5.1-C}$$

Reihen der Form (6.5.1-B) werden als (reelle) *Fourierreihen*[1] bezeichnet. Sie sind wichtig für Anwendungen und treten bei der Beschreibung von Schwingungsvorgängen natürlich auf.

Eine erste Beobachtung ist, dass die Koeffizienten in einer solchen Reihendarstellung eindeutig bestimmt sind. Wir bezeichnen mit $\mathrm{C}_{\mathrm{per}}[-\pi, \pi]$ die Menge der 2π-periodischen stetigen Funktionen und beschränken uns vorerst auf Reihen die auch gleichmäßig konvergieren.

☞ **6.5.2 Lemma** (Euler–Fouriersche Formeln)**.** *Angenommen, $f \in \mathrm{C}_{\mathrm{per}}[-\pi, \pi]$ besitzt eine gleichmäßig konvergente Reihendarstellung der Form* (6.5.1-B)*. Dann gilt*

$$\begin{aligned} a_0 &= \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x)\,\mathrm{d}x, \\ a_k &= \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos(kx)\,\mathrm{d}x, \\ b_k &= \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin(kx)\,\mathrm{d}x. \end{aligned} \tag{6.5.2-A}$$

Beweis. Der Beweis erfolgt durch Nachrechnen. Da die Reihe (6.5.1-B) gleichmäßig konvergiert, gilt

$$\int_{-\pi}^{\pi} f(x)\,\mathrm{d}x = 2\pi a_0 + \sum_{k=1}^{\infty} a_k \int_{-\pi}^{\pi} \cos(kx)\,\mathrm{d}x + \sum_{k=1}^{\infty} b_k \int_{-\pi}^{\pi} \sin(kx)\,\mathrm{d}x = 2\pi a_0. \tag{6.5.2-B}$$

Weiterhin gilt für $k \neq l$, $k, l \in \mathbb{N}$, mittels partieller Integration

$$\begin{aligned} \int_{-\pi}^{\pi} \cos(kx) \cos(lx)\,\mathrm{d}x &= \frac{k}{l} \int_{-\pi}^{\pi} \sin(kx) \sin(lx)\,\mathrm{d}x \\ &= \frac{k^2}{l^2} \int_{-\pi}^{\pi} \cos(kx) \cos(lx)\,\mathrm{d}x = 0 \end{aligned} \tag{6.5.2-C}$$

[1] Jean-Baptiste Joseph Fourier, 1768–1830

und ebenso als Integral über eine ungerade Funktion

$$\int_{-\pi}^{\pi} \cos(kx)\sin(lx)\,\mathrm{d}x = 0, \tag{6.5.2-D}$$

für alle $k, l \in \mathbb{N}$ sowie direkt auf Grund von $\cos x = \sin(x+\pi)$

$$\int_{-\pi}^{\pi} \cos^2(kx)\,\mathrm{d}x = \int_{-\pi}^{\pi} \sin^2(kx)\,\mathrm{d}x = \frac{1}{2}\int_{-\pi}^{\pi} \big(\cos^2(kx)+\sin^2(kx)\big)\,\mathrm{d}x = \pi. \tag{6.5.2-E}$$

Damit folgt

$$\begin{aligned}\int_{-\pi}^{\pi} f(x)\cos(lx)\,\mathrm{d}x &= a_0\int_{-\pi}^{\pi}\cos(lx)\,\mathrm{d}x + \sum_{k=1}^{\infty} a_k \int_{-\pi}^{\pi}\cos(kx)\cos(lx)\,\mathrm{d}x \\ &\quad + \sum_{k=1}^{\infty} b_k \int_{-\pi}^{\pi}\sin(kx)\cos(lx)\,\mathrm{d}x \\ &= \pi a_l\end{aligned} \tag{6.5.2-F}$$

und entsprechend

$$\int_{-\pi}^{\pi} f(x)\sin(lx)\,\mathrm{d}x = \pi b_l. \tag{6.5.2-G}$$

Damit sind alle Koeffizienten berechnet. □

6.5.3. Es stellt sich die Frage, welche stetigen Funktionen in eine solche gleichmäßig konvergente Fourierreihe entwickelt werden können. Wir werden diese Frage nicht allgemein beantworten können, ein hinreichendes Kriterium muss an dieser Stelle reichen. Als erstes bemerken wir, dass obige Formeln für jede stetige 2π-periodische Funktion ausgewertet werden können. Sei also von nun an $f \in \mathrm{C}_{\mathrm{per}}[-\pi,\pi]$ eine 2π-periodische stetige Funktion und seien die Zahlen a_0, a_k und b_k durch die Euler–Fourierschen Formeln (6.5.2-A) definiert. Dann ist die n-te Partialsumme der formal zugeordneten Fourierreihe

$$\begin{aligned}S_n[f](x) &:= a_0 + \sum_{k=1}^{n} a_k\cos(kx) + \sum_{k=1}^{n} b_k \sin(kx) \\ &= \frac{1}{2\pi}\int_{-\pi}^{\pi}\left(1 + 2\sum_{k=0}^{n}\cos(kx)\cos(kt) + 2\sum_{k=1}^{n}\sin(kx)\sin(kt)\right) f(t)\,\mathrm{d}t \\ &= \frac{1}{2\pi}\int_{-\pi}^{\pi}\left(1 + 2\sum_{k=0}^{n}\cos(k(x-t))\right) f(t)\,\mathrm{d}t \\ &= \frac{1}{2\pi}\int_{-\pi}^{\pi} D_n(x-t) f(t)\,\mathrm{d}t\end{aligned} \tag{6.5.3-A}$$

als Integral darstellbar. Wir berechnen noch den dabei auftretenden *Dirichlet-Kern* D_n. Er ist (nach Konstruktion) periodisch mit Periode 2π. Weiter gilt unter Nutzung der geometrischen Summenformel für alle $t \neq 0$

$$\begin{aligned}D_n(t) &= 1 + 2\sum_{k=1}^{n}\cos(kt) = \mathrm{Re}\sum_{k=-n}^{n}\mathrm{e}^{\mathrm{i}kt} \\ &= \mathrm{Re}\left(\mathrm{e}^{-\mathrm{i}nt}\frac{\mathrm{e}^{\mathrm{i}(2n+1)t}-1}{\mathrm{e}^{\mathrm{i}t}-1}\right) = \mathrm{Re}\,\frac{\mathrm{e}^{\mathrm{i}(n+\frac{1}{2})t}-\mathrm{e}^{-\mathrm{i}(n+\frac{1}{2})t}}{\mathrm{e}^{\mathrm{i}\frac{t}{2}}-\mathrm{e}^{-\mathrm{i}\frac{t}{2}}} = \frac{\sin((n+\frac{1}{2})t)}{\sin(\frac{t}{2})}.\end{aligned} \tag{6.5.3-B}$$

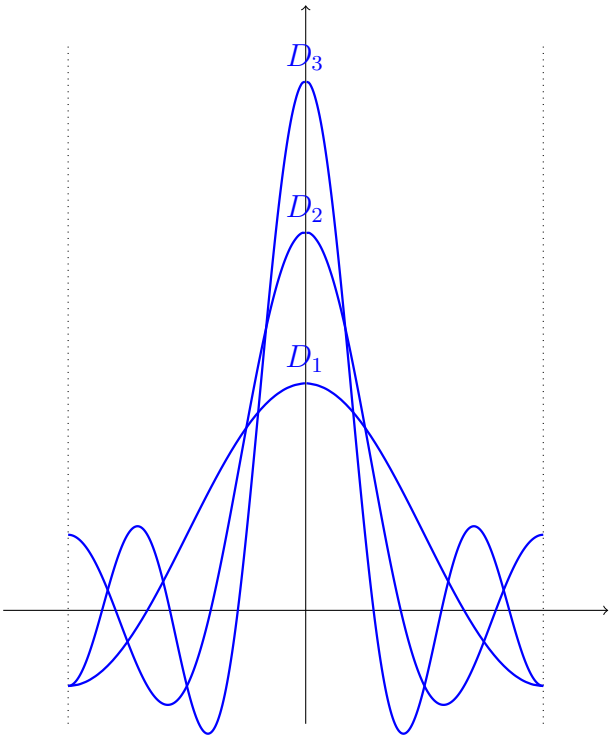
und offenbar $D_n(0) = 2n + 1$ und ebenso

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} D_n(t)\,\mathrm{d}t = 1, \tag{6.5.3-C}$$

sowie die Abschätzung

$$|tD_n(t)| \le 2\frac{\frac{1}{2}t}{\sin(\frac{1}{2}t)} \le \pi \tag{6.5.3-D}$$

für $t \in [-\pi, \pi]$. Das nachfolgende Bild liefert eine grobe Idee, wie die Dirichletkerne aussehen:



Damit können wir nun untersuchen, ob die Partialsummen der einer Funktion zugeordneten Reihe wirklich gegen die Funktion konvergieren. Wir benötigen dafür zuerst eine Hilfsaussage und eine Definition.

☞ **6.5.4 Lemma** (Riemann). *Sei $f \in \mathfrak{R}[-\pi, \pi]$. Dann gilt*

$$\lim_{\lambda\to\infty}\int_{-\pi}^{\pi} f(x)\cos(\lambda x)\,\mathrm{d}x = \lim_{\lambda\to\infty}\int_{-\pi}^{\pi} f(x)\sin(\lambda x)\,\mathrm{d}x = 0. \tag{6.5.4-A}$$

Beweis. Es genügt, den ersten der beiden Grenzwerte zu beweisen. Sei dazu $\varepsilon > 0$ und $\mathfrak{Z} = (-\pi = x_0, x_1, \ldots, x_m = \pi)$ eine Zerlegung des Intervalls $[-\pi, \pi]$, für die die Differenz aus Ober- und Untersumme

$$\mathcal{O}(f;\mathfrak{Z}) - \mathcal{U}(f;\mathfrak{Z}) < \frac{\varepsilon}{2} \tag{6.5.4-B}$$

erfüllt. Bezeichne weiter

$$\mu_k = \inf_{x\in[x_{k-1},x_k]} f(x), \qquad M_k = \sup_{x\in[x_{k-1},x_k]} f(x). \tag{6.5.4-C}$$

Zerlegt man nun

$$\begin{aligned}\int_{-\pi}^{\pi} f(x)\cos(\lambda x)\,\mathrm{d}x = \sum_{k=1}^{m}\int_{x_{k-1}}^{x_k}\big(f(x)-\mu_k\big)\cos(\lambda x)\,\mathrm{d}x \\ + \sum_{k=1}^{m}\mu_k\int_{x_{k-1}}^{x_k}\cos(\lambda x)\,\mathrm{d}x\end{aligned} \tag{6.5.4-D}$$

so können beide Summen einzeln abgeschätzt werden. Es gilt einerseits

$$\begin{aligned}\left|\sum_{k=1}^{m}\int_{x_{k-1}}^{x_k}\big(f(x)-\mu_k\big)\cos(\lambda x)\,\mathrm{d}x\right| &\le \sum_{k=1}^{m}\int_{x_{k-1}}^{x_k}\big|f(x)-\mu_k\big|\,\mathrm{d}x \\ &\le \sum_{k=1}^{m}(M_k-\mu_k)(x_k-x_{k-1}) < \frac{\varepsilon}{2}\end{aligned} \tag{6.5.4-E}$$

und andererseits gilt für $\lambda > 0$

$$\left|\sum_{k=1}^{m}\mu_k\int_{x_{k-1}}^{x_k}\cos(\lambda x)\,\mathrm{d}x\right| \le \sum_{k=1}^{m}|\mu_k|\left|\frac{\sin(\lambda x_k)-\sin(\lambda x_{k-1})}{\lambda}\right| \le \frac{2}{\lambda}\sum_{k=1}^{m}\mu_k < \frac{\varepsilon}{2} \tag{6.5.4-F}$$

für $\lambda > \frac{4}{\varepsilon}\sum_{k=1}^{m}|\mu_k|$. Also folgt

$$\left|\int_{-\pi}^{\pi} f(x)\cos(\lambda x)\,\mathrm{d}x\right| < \varepsilon \tag{6.5.4-G}$$

und damit die Behauptung. □

✍ **6.5.5 Definition.** *Eine stetige Funktion $f:\mathbb{R}\to\mathbb{R}$ heißt im Punkt x lokal* lipschitzstetig[2]*, falls es ein $\delta > 0$ und eine Konstante L_x gibt, so dass*

$$\forall_{y:|x-y|<\delta} \quad |f(x)-f(y)| \le L_x|x-y| \tag{6.5.5-A}$$

gilt.

☞ **6.5.6 Lemma.** *Jede stetig differenzierbare Funktion ist lokal lipschitzstetig.*

Beweis. Hier kann für jedes x die Konstante $\delta = 1$ gewählt werden und der erste Mittelwertsatz der Differentialrechnung liefert wegen

$$f(x)-f(y) = f'(\xi)(x-y) \tag{6.5.6-A}$$

für ein ξ zwischen x und y für $|x-y|<1$ direkt die Abschätzung

$$|f(x)-f(y)| \le L_x|x-y| \qquad \text{für} \quad L_x = \max_{\xi\in[x-1,x+1]}|f'(\xi)| \tag{6.5.6-B}$$

und damit die Behauptung. □

[2] Rudolf Lipschitz, 1832–1903

☞ **6.5.7 Satz** (Lipschitz-Kriterium). *Angenommen $f \in \mathfrak{R}[-\pi,\pi]$ ist 2π-periodisch fortgesetzt und in $x \in [-\pi,\pi]$ lokal lipschitzstetig. Dann gilt*

$$\lim_{n\to\infty} S_n[f](x) = f(x). \tag{6.5.7-A}$$

Beweis. Wir können ohne Beschränkung der Allgemeinheit annehmen, dass $x = 0$ gilt (sonst verschieben wir f entsprechend). Da f in 0 lokal lipschitzstetig ist, gibt es eine Zahl δ und eine Lipschitzkonstante L mit

$$|f(t) - f(0)| \le L|t| \tag{6.5.7-B}$$

für alle t mit $|t| < \delta$. Betrachten wir nun die Differenz

$$\begin{aligned} S_n[f](0) - f(0) &= \frac{1}{2\pi}\int_{-\pi}^{\pi} D_n(t)\,(f(t)-f(0))\,\mathrm{d}t \\ &= \frac{1}{2\pi}\int_{-\delta}^{\delta} D_n(t)\,(f(t)-f(0))\,\mathrm{d}t \\ &\quad + \frac{1}{2\pi}\int_{|t|\ge\delta} D_n(t)\,(f(t)-f(0))\,\mathrm{d}t \end{aligned} \tag{6.5.7-C}$$

und zerlegen diese in Teile t nahe 0 und den Rest, so können wir beide Summanden einzeln abschätzen. Einerseits gilt auf Grund der lokalen Lipschitszstetigkeit von f in 0

$$\left|\frac{1}{2\pi}\int_{-\delta}^{\delta} D_n(t)\,(f(t)-f(0))\,\mathrm{d}t\right| \le \frac{1}{2\pi}\int_{-\delta}^{\delta}\frac{\pi}{|t|}|f(t)-f(0)|\,\mathrm{d}t \le L\delta < \frac{\varepsilon}{2} \tag{6.5.7-D}$$

für δ klein genug und andererseits mit dem Riemann-Lemma

$$\int_{|t|\ge\delta} D_n(t)\,(f(t)-f(0))\,\mathrm{d}t = \int_{|t|\ge\delta}\frac{f(t)-f(0)}{\sin(\frac{t}{2})}\sin((n+1/2)t)\,\mathrm{d}t \longrightarrow 0 \tag{6.5.7-E}$$

für $n \to \infty$ und damit in Summe für hinreichend großes n auch

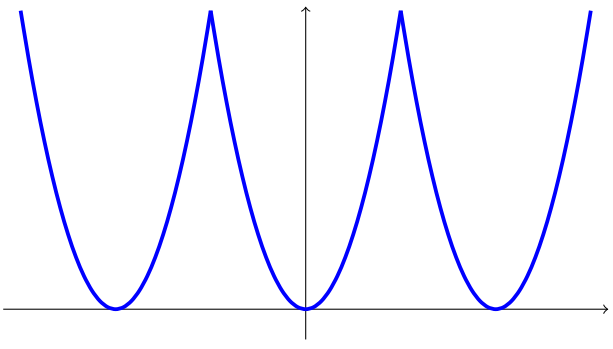
$$|S_n[f](0) - f(0)| < \varepsilon. \tag{6.5.7-F}$$

Damit folgt die Behauptung. □

⚒ **6.5.8 Beispiel.** Wir betrachten ein Beispiel und nehmen die Funktion

$$f(x) = x^2 \tag{6.5.8-A}$$

für $x \in [-\pi,\pi]$. Wir berechnen die Fourier-Koeffizienten der periodischen Fortsetzung.



Es gilt

$$a_0 = \frac{1}{2\pi}\int_{-\pi}^{\pi} x^2\,\mathrm{d}x = \frac{1}{2\pi}\left[\frac{1}{3}x^3\right]_{x=-\pi}^{\pi} = \frac{\pi^2}{3} \tag{6.5.8-B}$$

und mit partieller Integration

$$\begin{aligned} a_k &= \frac{1}{\pi}\int_{-\pi}^{\pi} x^2\cos(kx)\,\mathrm{d}x = \frac{x^2}{k\pi}\sin(kx)\Big|_{x=-\pi}^{\pi} - \frac{2}{k\pi}\int_{-\pi}^{\pi} x\sin(kx)\,\mathrm{d}x \\ &= \frac{2x}{k^2\pi}\cos(kx)\Big|_{x=-\pi}^{\pi} - \frac{2}{k^2\pi}\int_{-\pi}^{\pi}\cos(kx)\,\mathrm{d}x = \frac{4}{k^2}\cos(k\pi) = (-1)^k\frac{4}{k^2}. \end{aligned} \tag{6.5.8-C}$$

Da f gerade ist, gilt weiterhin für alle k

$$b_k = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\sin(kx)\,\mathrm{d}x = 0. \tag{6.5.8-D}$$

Die periodische Fortsetzung ist lokal lipschitzstetig, die zugeordnete Fourierreihe konvergiert damit punktweise gegen f. Es gilt also

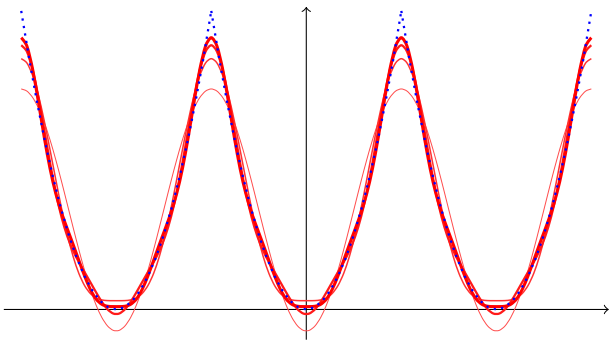
$$f(x) = \frac{\pi^2}{3} + 4\sum_{k=1}^{\infty}(-1)^k\frac{\cos(kx)}{k^2} \tag{6.5.8-E}$$

Speziell mit $x = \pi$ ergibt sich damit

$$\pi^2 = f(\pi) = \frac{\pi^2}{3} + 4\sum_{k=1}^{\infty}\frac{1}{k^2} \tag{6.5.8-F}$$

und damit

$$\sum_{k=1}^{\infty}\frac{1}{k^2} = \frac{\pi^2}{6}. \tag{6.5.8-G}$$



✂ **Ergänzung.** Das Lipschitz-Kriterium in Satz 6.5.7 liefert in der angegebenen Fassung nur punktweise und keine gleichmäßige Konvergenz. Ist eine periodische Funktion $f \in \mathrm{C}_{\mathrm{per}}[-\pi,\pi]$ jedoch global Lipschitz, gilt also

$$|f(x) - f(y)| \le L|x-y| \tag{6.5.8-H}$$

für $x, y \in \mathbb{R}$ mit $|x-y| < 2\pi$, so folgt als Verschärfung des Riemann-Lemmas sogar

$$\left|\frac{1}{\pi}\int_{-\pi}^{\pi} f(t)\sin(\lambda t)\,\mathrm{d}t\right| \le 2L\lambda^{-1} \tag{6.5.8-I}$$

und der oben angegebene Beweis zu Satz 6.5.7 liefert die gleichmäßige Konvergenz. Für lipschitzstetige periodische Funktionen konvergieren Fourierreihen also stets gleichmäßig. Für stetig differenzierbare f folgt (6.5.8-I) direkt durch partielles Integrieren.

⚒ **6.5.9 Beispiel.** Wir geben noch einige interessante Fourierreihen an, für die wir ebenso in allen Stetigkeitspunkten die Konvergenz gezeigt haben.

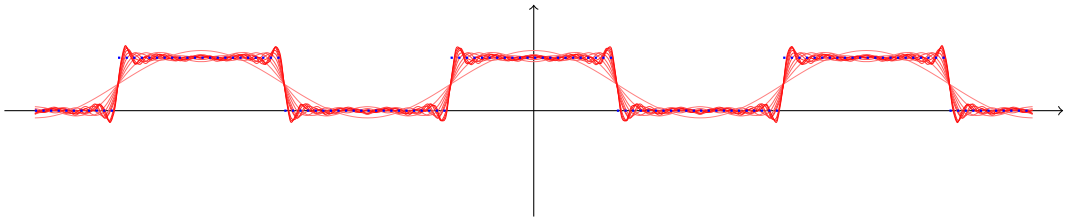
(i) Für die Rechteckkurve

$$f(x) = \begin{cases} 1, & |x - 2k\pi| < \frac{\pi}{2}, \\ 0, & \text{sonst}, \end{cases} \tag{6.5.9-A}$$

sind alle Sinuskoeffizienten Null und es ergibt sich $a_0 = \frac{1}{2}$ und für $l \in \mathbb{N}_0$

$$a_{2l+1} = \frac{1}{\pi} \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \cos((2l+1)x)\,\mathrm{d}x = \frac{1}{(2l+1)\pi} \sin((2l+1)x)\Big|_{-\frac{\pi}{2}}^{\frac{\pi}{2}} = \frac{(-1)^l}{(l+\frac{1}{2})\pi} \tag{6.5.9-B}$$

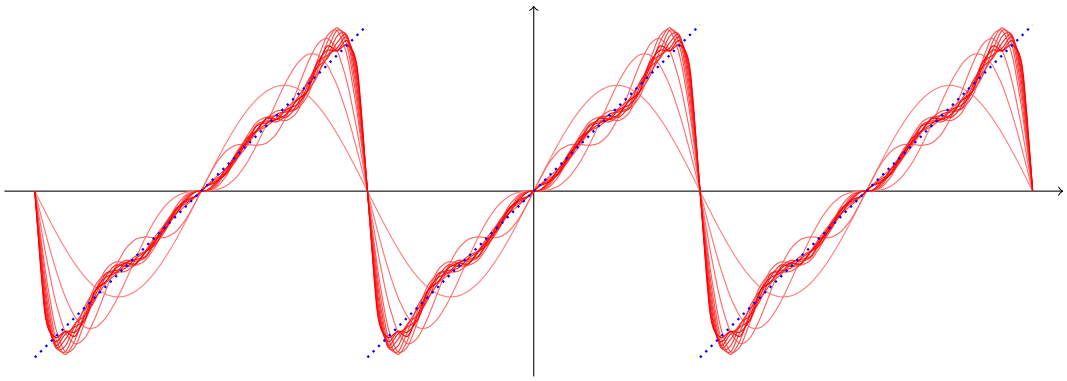
sowie entsprechend $a_{2l+2} = 0$. Die ersten Summanden sind nachfolgend dargestellt:



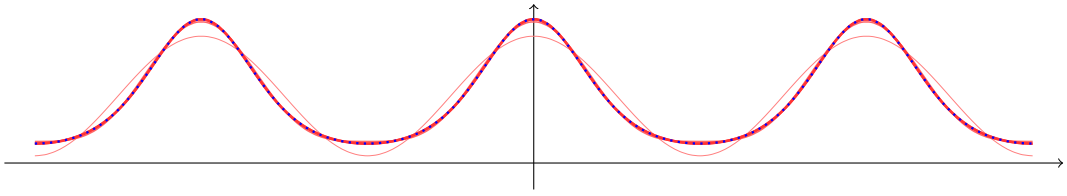
(ii) Für die 2π-periodisch fortgesetzte Sägezahnkurve $f(x) = x$, $x \in [-\pi, \pi]$, sind nur die Sinuskoeffizienten ungleich Null und es gilt

$$\begin{aligned} b_k &= \frac{1}{\pi} \int_{-\pi}^{\pi} x \sin(kx)\,\mathrm{d}x = -\frac{x\cos(kx)}{k\pi}\Big|_{-\pi}^{\pi} + \frac{1}{k\pi} \int_{-\pi}^{\pi} \cos(kx)\,\mathrm{d}x \\ &= (-1)^{k-1} \frac{2}{k}. \end{aligned} \tag{6.5.9-C}$$

Die ersten Summanden sind nachfolgend dargestellt:



(iii) Für mehrfach differenzierbare Funktionen konvergieren Fourierreihen schneller. Ein Beispiel dazu wäre $f(x) = \mathrm{e}^{\cos x}$ mit den nachfolgend dargestellten ersten Summanden:



7 Differentialrechnung in mehreren Veränderlichen

7.1 Vektorräume und lineare Abbildungen

7.1.1. Ab diesem Kapitel werden wir einige Resultate aus der linearen Algebra benötigen, die wir bei Bedarf wiederholen und zusammenfassen. Ziel ist dabei, die bisher für Funktionen $f : (a, b) \to \mathbb{R}$ diskutierte Differential- und Integralrechnung auf Funktionen mehrerer Veränderlicher, also Funktionen $f : \mathbb{R}^n \supset U \to \mathbb{R}^m$ definiert auf Teilmengen $U \subset \mathbb{R}^n$, zu verallgemeinern.

✂ **Ergänzung** (Spaltenvektoren und Matrizen)**.** Wir beginnen mit einem recht allgemeinen Beispiel. Elemente des $\mathbb{R}^n$ sind n-Tupel von Zahlen, die wir als *Spaltenvektoren*

$$\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \tag{7.1.1-A}$$

schreiben wollen. Mit solchen Spaltenvektoren kann man rechnen, wir definieren die beiden Operationen

$$\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} + \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} := \begin{bmatrix} x_1 + y_1 \\ x_2 + y_2 \\ \vdots \\ x_n + y_n \end{bmatrix}, \qquad \lambda \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} := \begin{bmatrix} \lambda x_1 \\ \lambda x_2 \\ \vdots \\ \lambda x_n \end{bmatrix} \tag{7.1.1-B}$$

für $x_j, y_j, \lambda \in \mathbb{R}$. Bezeichnen wir mit $\mathbf{e}_j$ die speziellen Spaltenvektoren mit Eintrag 1 in der j-ten Zeile und sonst den Einträgen 0, so gilt

$$x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \sum_{j=1}^{n} x_j \mathbf{e}_j. \tag{7.1.1-C}$$

Eine Abbildung $f : \mathbb{R}^n \to \mathbb{R}^m$ heißt linear, falls für alle $x, y \in \mathbb{R}^n$ und alle Zahlen $\lambda \in \mathbb{R}$

$$f(x + \lambda y) = f(x) + \lambda f(y) \tag{7.1.1-D}$$

gilt. Solche linearen Abbildungen haben eine besonders einfache Form, es gilt

$$f(x) = f\left(\sum_{j=1}^{n} x_j \mathbf{e}_j\right) = \sum_{j=1}^{n} x_j f(\mathbf{e}_j) \qquad \text{mit} \quad f(\mathbf{e}_j) \in \mathbb{R}^m. \tag{7.1.1-E}$$

Schreibt man

$$f(\mathbf{e}_j) = \begin{bmatrix} a_{1,j} \\ a_{2,j} \\ \vdots \\ a_{m,j} \end{bmatrix}, \qquad A = \begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & & \vdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,n} \end{bmatrix}, \tag{7.1.1-F}$$

so kann man die lineare Abbildung f formal als *Matrix*-Vektor-Multiplikation schreiben,

$$f(x) = Ax = \begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & & \vdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,n} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}. \tag{7.1.1-G}$$

Die Menge der $m \times n$-Matrizen werde als $\mathbb{R}^{m\times n}$ bezeichnet. Verkettung linearer Abbildungen entspricht der Multiplikation von Matrizen. Der Vollständigkeit halber dazu die passende Formel, es gilt

$$\begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,k} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,k} \\ \vdots & \vdots & & \vdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,k} \end{bmatrix} = \begin{bmatrix} b_{1,1} & b_{1,2} & \cdots & b_{1,n} \\ b_{2,1} & b_{2,2} & \cdots & b_{2,n} \\ \vdots & \vdots & & \vdots \\ b_{m,1} & b_{m,2} & \cdots & b_{m,n} \end{bmatrix} \begin{bmatrix} c_{1,1} & c_{1,2} & \cdots & c_{1,k} \\ c_{2,1} & c_{2,2} & \cdots & c_{2,k} \\ \vdots & \vdots & & \vdots \\ c_{n,1} & c_{n,2} & \cdots & c_{n,k} \end{bmatrix} \tag{7.1.1-H}$$

falls

$$a_{i,j} = \sum_{l=1}^{n} b_{i,l} c_{l,j}. \tag{7.1.1-I}$$

Matrixmultiplikation ist assoziativ, da die Verkettung von Abbildung assoziativ ist. Weiter bezeichne $A^\top$ zu einer Matrix A die *transponierte Matrix*, die sich durch Spiegeln an der Diagonalen ergibt,

$$\begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & & \vdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,n} \end{bmatrix}^\top = \begin{bmatrix} a_{1,1} & a_{2,1} & \cdots & a_{m,1} \\ a_{1,2} & a_{2,2} & \cdots & a_{m,2} \\ \vdots & \vdots & & \vdots \\ a_{1,n} & a_{2,n} & \cdots & a_{m,n} \end{bmatrix}. \tag{7.1.1-J}$$

Interpretiert man Spaltenvektoren als $n \times 1$-Matrizen, so liefert $y^\top x = \sum_{j=1}^n y_j x_j$ für $x, y \in \mathbb{R}^n$ das (euklidische) Skalarprodukt von x und y. Weiter gilt $(AB)^\top = B^\top A^\top$ und damit insbesondere $y^\top (Ax) = (A^\top y)^\top x$.

Die zur identischen Abbildung gehörende Matrix wird als *Einheitsmatrix* bezeichnet, diese besitzt die Einträge $a_{j,j} = 1$ und $a_{i,j} = 0$ für $i \neq j$. Eine Matrix A heißt *invertierbar*, wenn die zugeordnete lineare Abbildung invertierbar ist, es also eine Matrix B mit $AB = BA = I$ gibt.

✂ **Ergänzung** (Vektorräume, Basis und Dimension)**.** Sei nun allgemeiner $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$. Ein *Vektorraum* über $\mathbb{K}$ ist eine Menge V versehen mit zwei Operationen $+ : V \times V \to V$ und $\cdot : \mathbb{K} \times V \to V$ und einem speziellen Element $\mathbf{0} \in V$, für die

$$\forall_{x,y\in V} \quad x + y = y + x \tag{7.1.2-A}$$

$$\forall_{x,y,z\in V} \quad x + (y + z) = (x + y) + z \tag{7.1.2-B}$$

$$\forall_{x\in V} \quad 0 \cdot x = \mathbf{0} \quad \wedge \quad 1 \cdot x = x \tag{7.1.2-C}$$

$$\forall_{\lambda\in\mathbb{K}} \forall_{x,y\in V} \quad \lambda \cdot (x + y) = \lambda \cdot x + \lambda \cdot y \tag{7.1.2-D}$$

$$\forall_{\lambda,\mu\in\mathbb{K}} \forall_{x\in V} \quad (\lambda + \mu) \cdot x = \lambda \cdot x + \mu \cdot x \quad \wedge \quad \lambda \cdot (\mu \cdot x) = (\lambda\mu) \cdot x \tag{7.1.2-E}$$

gelten. Elemente von V werden als Vektoren bezeichnet. Wir werden im Weiteren in der Notation nicht mehr zwischen dem Nullvektor $\mathbf{0} \in V$ und der Zahl $0 \in \mathbb{K}$ unterscheiden und bei der skalaren Multiplikation die Punkte weglassen also λx statt $\lambda \cdot x$ schreiben.

Wir sagen, ein Vektorraum V ist *endlich erzeugt*, falls es Vektoren $\mathbf{v}_1, \ldots, \mathbf{v}_n \in V$ mit

$$V = \operatorname{span}\{\mathbf{v}_1, \ldots, \mathbf{v}_n\} := \left\{ \sum_{k=1}^{n} \lambda_k \mathbf{v}_k \mid \lambda_k \in \mathbb{K} \right\} \tag{7.1.2-F}$$

gibt. Wir sagen, das Tupel der Vektoren $(\mathbf{v}_1, \ldots, \mathbf{v}_n)$ bildet eine *Basis* von V, falls sie zusätzlich *linear unabhängig* sind, also

$$\sum_{k=1}^{n} \lambda_k \mathbf{v}_k = 0 \qquad \Longrightarrow \qquad \lambda_1 = \cdots = \lambda_n = 0 \tag{7.1.2-G}$$

gilt. Insbesondere gilt $\mathbf{v}_k \neq 0$. Hat man zwei Basen $\mathbf{v}_1, \ldots, \mathbf{v}_n$ und $\mathbf{w}_1, \ldots, \mathbf{w}_m$ eines endlich erzeugten Vektorraums V gegeben, so haben diese dieselbe Anzahl an Elementen. Dies rechnen wir kurz nach. Angenommen, $n \leq m$. Da $0 \neq \mathbf{w}_1 \in V$ ist, gibt es Zahlen $\lambda_1, \ldots, \lambda_n$ mit $\mathbf{w}_1 = \lambda_1\mathbf{v}_1 + \cdots + \lambda_m\mathbf{v}_m$. Mindestens einer der Koeffizienten λ_k ist ungleich Null, es gilt also

$$v_k = \lambda_k^{-1}\left(\mathbf{w}_1 - \sum_{l\neq k} \lambda_l \mathbf{v}_l\right) \in \operatorname{span}\{\mathbf{w}_1, \mathbf{v}_1, \ldots, \mathbf{v}_{k-1}, \mathbf{v}_{k+1}, \ldots, \mathbf{v}_n\}. \tag{7.1.2-H}$$

Die Menge der Vektoren $\{\mathbf{w}_1, \mathbf{v}_1, \ldots, \mathbf{v}_{k-1}, \mathbf{v}_{k+1}, \ldots, \mathbf{v}_n\}$ erzeugt offenbar immer noch V. Auf diese Weise können wir schrittweise die Elemente $\mathbf{w}_1, \ldots, \mathbf{w}_n$ in das Erzeugendensystem hineintauschen, es kann nie passieren, dass nur vor eingetauschten $\mathbf{w}$-Vektoren von Null verschiedene Koeffizienten stehen (da ja die $\mathbf{w}_k$ linear unabhängig waren) und nach n Schritten folgt $V = \operatorname{span}\{\mathbf{w}_1, \ldots, \mathbf{w}_n\}$. Wäre nun $m > n$, so wäre damit $\mathbf{w}_m \in \operatorname{span}\{\mathbf{w}_1, \ldots, \mathbf{w}_n\}$ im Widerspruch zur linearen Unabhängigkeit.

Eine Abbildung $f : V \to W$ zwischen zwei Vektorräumen V und W heißt *linear*, falls (7.1.**1-D**) gilt. Wir bezeichnen mit

$$\operatorname{Lin}(V, W) = \{f : V \to W \mid f \text{ linear}\} \tag{7.1.2-I}$$

die Menge der linearen Abbildungen von V nach W. Diese ist selbst wiederum ein Vektorraum mit den Operationen $(f + g)(x) := f(x) + g(x)$ und $(\lambda f)(x) := \lambda f(x)$.

Für einen endlich erzeugten Vektorraum V bezeichnen wir die Anzahl der Elemente einer Basis von V als seine *Dimension* $\dim V$. Sei $\mathbf{v}_1, \ldots, \mathbf{v}_n \in V$ eine Basis von V, dann liefert die eindeutige Darstellung der Elemente als Linearkombination

$$x = \sum_{k=1}^{n} x_k \mathbf{v}_k \tag{7.1.2-J}$$

eine wohldefinierte bijektive lineare Abbildung

$$V \ni x = \sum_{k=1}^{n} x_k \mathbf{v}_k \mapsto \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \in \mathbb{K}^n, \tag{7.1.2-K}$$

wir können V also darüber mit $\mathbb{K}^n$ identifizieren. Die Identifikation hängt natürlich von der Wahl der Basis ab. Sind V und W beide endlichdimensional und gilt $\dim V = n$ und $\dim W = m$, so liefern Basen von V und W entsprechend eine Identifikation

$$\operatorname{Lin}(V, W) \longleftrightarrow \mathbb{K}^{m\times n} \tag{7.1.2-L}$$

linearer Abbildungen mit $m \times n$-Matrizen. Insbesondere gilt dann $\dim \operatorname{Lin}(V, W) = mn$.

✂ **Ergänzung** (Unterräume und Dimensionsformel)**.** Eine Teilmenge V_1 eines Vektorraumes V heißt *Unterraum*, wenn sie nichtleer ist und bezüglich der Vektorraumoperationen abgeschlossen ist, also

$$0 \in V_1 \qquad \text{und} \qquad \forall_{x,y\in V_1}\forall_{\alpha\in\mathbb{K}} \quad x + \alpha y \in V_1 \tag{7.1.2-M}$$

gilt. Ist V endlich erzeugt, so auch V_1 und jede Basis von V_1 kann durch Hinzunahme weiterer Vektoren zu einer Basis von V ergänzt werden.

Für zwei Unterräume V_1 und V_2 eines Vektorraumes V bezeichnet

$$V_1 + V_2 := \{\alpha v_1 + \beta v_2 \mid \alpha, \beta \in \mathbb{K},\ v_1 \in V_1,\ v_2 \in V_2\} \tag{7.1.2-N}$$

ihre *Summe*. Diese ist wiederum ein Unterraum. Ebenso ist der Schnitt $V_1 \cap V_2$ zweier Unterräume ein Unterraum. In diesem Fall gilt für die Dimensionen die *Dimensionsformel*

$$\dim(V_1 + V_2) = \dim V_1 + \dim V_2 - \dim(V_1 \cap V_2). \tag{7.1.2-O}$$

Zum Beweis ergänzt man eine Basis von $V_1 \cap V_2$ zu einer von V_1 und ebenso zu einer von V_2.

Gilt zusätzlich, $V_1 \cap V_2 = \{0\}$, so bezeichnet man $V_1 + V_2$ als *direkte Summe*. Für solche direkten Summen gibt es in der Literatur verschiedene Notationen, wir verwenden $V_1 \dot{+} V_2$.

✂ **Ergänzung** (Dimensionssatz, Rang und Defekt)**.** Für $f \in \mathrm{Lin}(V, W)$ sind

$$\ker f = \{x \in V \mid f(x) = 0\} \subseteq V \tag{7.1.3-A}$$

und

$$\operatorname{im} f = \{f(x) \mid x \in V\} \subseteq W \tag{7.1.3-B}$$

beides Unterräume. Auf V liefert

$$x \sim y \qquad :\Longleftrightarrow \qquad x - y \in \ker f \tag{7.1.3-C}$$

eine Äquivalenzrelation. Auf den zugeordneten Äquivalenzklassen ist $[x]_\sim + [y]_\sim := [x + y]_\sim$ und $\lambda[x]_\sim := [\lambda x]_\sim$ wohldefiniert und liefert wiederum die Struktur eines Vektorraumes. Dies ist einfach nachzurechnen und folgt direkt aus der Vektorraumeigenschaft von $\ker f$. Da dann die induzierte Abbildung

$$V/_{\ker f} \ni [x]_\sim \mapsto f([x]_\sim) := f(x) \in \operatorname{im} f \tag{7.1.3-D}$$

wohldefiniert, linear und bijektiv ist, folgt der *Dimensionssatz*

$$\dim \operatorname{im} f + \dim \ker f = \dim V/_{\ker f} + \dim \ker f = \dim V. \tag{7.1.3-E}$$

Für die letzte Gleichheit haben wir genutzt, dass für eine Basis $[\mathbf{w}_1], \ldots, [\mathbf{w}_k]$ von $V/_{\ker f}$ und eine Basis $\mathbf{v}_1, \ldots, \mathbf{v}_l$ von $\ker f$ die Vektoren $\mathbf{v}_1, \ldots, \mathbf{v}_l, \mathbf{w}_1, \ldots, \mathbf{w}_k$ eine Basis von V bilden. Die Zahl

$$\operatorname{rank} f := \dim \operatorname{im} f \tag{7.1.3-F}$$

bezeichnet man als *Rang* der linearen Abbildung f. Die Zahl $\dim \ker f$ wird oft als *Defekt* der linearen Abbildung f bezeichnet.

Gilt $\dim V = \dim W = n$, so gilt für eine lineare Abbildung $f \in \mathrm{Lin}(V, W)$ auf Grund des Dimensionssatzes

$$f \text{ injektiv} \qquad \Longleftrightarrow \qquad \ker f = \{0\} \qquad \Longleftrightarrow \qquad \operatorname{im} f = W \tag{7.1.3-G}$$

und lineare Abbildungen zwischen endlichdimensionalen Vektorräumen gleicher Dimension sind genau dann injektiv wenn sie surjektiv sind.

✂ **Ergänzung** (Determinanten)**.** Oft ist ein alternatives Kriterium für die Invertierbarkeit einer linearen Abbildung nützlich. Dazu definieren wir zu einer Matrix $A \in \mathbb{K}^{n\times n}$ die *Determinante*

$$\det A := \sum_{\sigma \in \mathfrak{S}_n} \operatorname{sign}(\sigma) \prod_{k=1}^{n} a_{k\sigma(k)} \tag{7.1.4-A}$$

als Summe über alle Permutationen $\sigma \in \mathfrak{S}_n$, also bijektiven Abbildungen $\sigma : \{1, \ldots, n\} \to \{1, \ldots, n\}$, und mit

$$\operatorname{sign}(\sigma) = (-1)^{\#\{(k,l) \mid k<l \ \wedge\ \sigma(l)<\sigma(k)\}} \tag{7.1.4-B}$$

der Signatur der Permutation σ. Wir werden später eine alternative Charakterisierung der Determinante geben, vorerst begnügen wir uns mit zwei wesentlichen Eigenschaften. Es gilt $\det I = 1$ für die Einheitsmatrix I und für zwei Matrizen $A, B \in \mathbb{K}^{n\times n}$ gilt

$$\det(AB) = \det A \det B. \tag{7.1.4-C}$$

Damit muss für invertierbare Matrizen $\det A \neq 0$ gelten. Es gilt auch die Umkehrung

$$A \text{ invertierbar} \qquad \Longleftrightarrow \qquad \det A \neq 0. \tag{7.1.4-D}$$

✂ **Ergänzung** (Spur)**.** Für quadratische Matrizen $A = (a_{i,j}) \in \mathbb{K}^{n\times n}$ ist die *Spur* durch die Summe der Diagonaleinträge

$$\operatorname{trace} A := \sum_{i=1}^{n} a_{i,i} \tag{7.1.5-A}$$

gegeben.

7.2 Zur topologischen Struktur metrischer und normierter Räume

7.2.1. Ein metrischer Raum (M, d) ist eine Menge M versehen mit einer Metrik, also einer Abbildung $d : M \times M \to \mathbb{R}$ welche

$$d(x,y) = d(y,x) \geq 0 \quad \wedge \quad (d(x,y) = 0 \iff x = y) \tag{7.2.1-A}$$

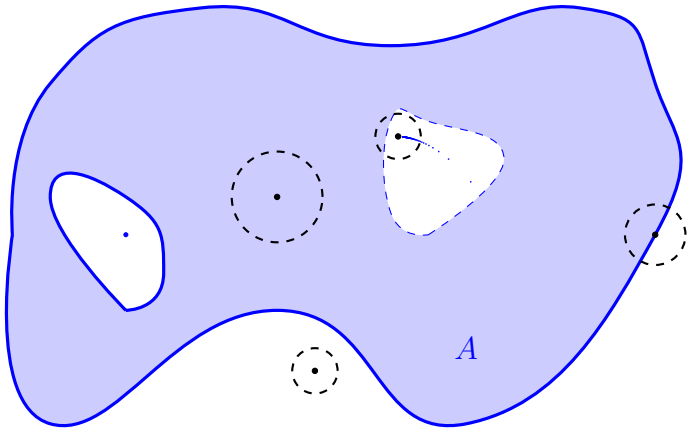
zusammen mit

$$d(x,z) \leq d(x,y) + d(y,z) \tag{7.2.1-B}$$

für alle $x, y, z \in M$ erfüllt. In einem metrischen Raum bezeichne zu $x \in M$ und $\varepsilon > 0$

$$B_\varepsilon(x) = \{y \in M \mid d(x,y) < \varepsilon\} \tag{7.2.1-C}$$

die offene ε-Kugel um x. Diese Kugeln kann man nutzen, um Eigenschaften von Teilmengen des Raumes M zu charakterisieren. Einige davon haben wir schon gesehen.



✍ **7.2.2 Definition.** *Sei (M, d) metrischer Raum und $A \subseteq M$ eine Teilmenge.*

(i) *Ein Element $x \in A$ heißt* innerer Punkt, *falls ein $\varepsilon > 0$ mit*

$$B_\varepsilon(x) \subseteq A \tag{7.2.2-A}$$

existiert.

(ii) *Ein Punkt $x \in M$ heißt* äußerer Punkt *von A, falls ein $\varepsilon > 0$ mit*

$$B_\varepsilon(x) \cap A = \varnothing \tag{7.2.2-B}$$

existiert. Er heißt Randpunkt *von A, falls für jedes $\varepsilon > 0$*

$$B_\varepsilon(x) \cap A \neq \varnothing \qquad \text{und} \qquad B_\varepsilon(x) \setminus A \neq \varnothing \tag{7.2.2-C}$$

gilt. Wir bezeichnen die Menge der Randpunkte von A mit ∂A.

(iii) *Die Menge A heißt* offen*, falls jedes $x \in A$ innerer Punkt ist, also*

$$A \cap \partial A = \varnothing \tag{7.2.2-D}$$

gilt.

(iv) *Die Menge A heißt* abgeschlossen*, falls ihr Komplement $M \setminus A$ offen ist, also*

$$\partial A \subseteq A \tag{7.2.2-E}$$

gilt.

(v) *Ein Punkt $x \in M$ heißt* Häufungspunkt *von A, falls für jedes $\varepsilon > 0$*

$$\big(B_\varepsilon(x) \setminus \{x\}\big) \cap A \neq \varnothing \tag{7.2.2-F}$$

gilt. Die Menge der Häufungspunkte von A wird mit $\operatorname{acc} A$ bezeichnet.

(vi) *Die Menge A heißt* perfekt*, falls $A = \operatorname{acc} A$ gilt.*

⚒ **7.2.3 Beispiel.** Für jedes $r > 0$ und jedes $x \in M$ ist $B_r(x)$ offen. Sei dazu $y \in B_r(x)$. Wählt man nun $\varepsilon < r - d(x, y)$, so folgt mit der Dreiecksungleichung für jedes $z \in B_\varepsilon(y)$

$$d(x, z) \leq d(x, y) + d(y, z) < d(x, y) + \varepsilon < r \tag{7.2.3-A}$$

und damit $z \in B_r(x)$, also $B_\varepsilon(y) \subset B_r(x)$. Analog sieht man, dass

$$\{y \in M \mid d(x, y) > r\} \tag{7.2.3-B}$$

offen und damit auch

$$\{y \in M \mid d(x, y) \leq r\} \tag{7.2.3-C}$$

abgeschlossen ist. Man beachte, dass das nicht $\partial B_r(x) = \{y \mid d(x, y) = r\}$ impliziert. Es ist in allgemeinen metrischen Räumen durchaus möglich, dass kein Punkt genau den Abstand r zu x besitzt.

Wir werden im Weiteren oft Vektorräume mit einer Metrik versehen. Die einfachste Konstruktion dazu ist die Verwendung einer Norm.

✍ **7.2.4 Definition.** *Sei nun V ein Vektorraum über $\mathbb{R}$. Eine* Norm *auf V ist eine Abbildung $\|\cdot\| : V \to \mathbb{R}_{\geq 0}$ mit*

N1

$$\|x\| = 0 \iff x = 0 \tag{7.2.4-A}$$

N2

$$\|\lambda x\| = |\lambda| \, \|x\| \tag{7.2.4-B}$$

N3

$$\|x + y\| \leq \|x\| + \|y\| \tag{7.2.4-C}$$

für $x, y \in V$ und $\lambda \in \mathbb{R}$. Ein Vektorraum versehen mit einer Norm wird als normierter Raum *bezeichnet. Jeder normierte Raum wird durch die induzierte Metrik*

$$d(x, y) = \|x - y\| \tag{7.2.4-D}$$

zu einem metrischen Raum. Er heißt vollständig*, oder* Banachraum*, falls er als metrischer Raum vollständig ist.*

⚒ **7.2.5 Beispiel.** Auf dem $\mathbb{R}^n$ bestimmt $|x| := \sqrt{x^\top x}$ eine Norm. Für das Skalarprodukt $x^\top y$ zweier Vektoren des $\mathbb{R}^n$ gilt die *Cauchy–Schwarz-Ungleichung*[1]

$$|x^\top y| = \left|\sum_{k=1}^{n} x_k y_k\right| \le \left(\sum_{k=1}^{n} x_k^2\right)^{\frac{1}{2}} \left(\sum_{k=1}^{n} y_k^2\right)^{\frac{1}{2}} = |x|\,|y|, \tag{7.2.5-A}$$

die es zuerst zu beweisen gilt. Für $y = 0$ sind beide Seiten Null, wir können also annehmen, dass $y \neq 0$ gilt. Dann gilt für jedes $\lambda \in \mathbb{R}$

$$0 \le |x + \lambda y|^2 = x^\top x + 2\lambda x^\top y + \lambda^2 y^\top y = |x|^2 + 2\lambda x^\top y + \lambda^2 |y|^2. \tag{7.2.5-B}$$

Da die rechte Seite ein quadratisches Polynom in λ ist, welches an der Stelle

$$\lambda = -\frac{x^\top y}{|y|^2} \tag{7.2.5-C}$$

sein Minimum annimmt, folgt nach Einsetzen und Multiplikation mit $|y|^2$

$$0 \le |x|^2 |y|^2 - x^\top y \tag{7.2.5-D}$$

und damit die Behauptung.

Wir weisen nach, dass es sich bei $|\cdot|$ wirklich um eine Norm handelt. Die ersten beiden Normeigenschaften sind direkt aus der Definition ersichtlich, die Dreiecksungleichung folgt aus Cauchy–Schwarz. Es gilt

$$\begin{aligned} |x + y|^2 = (x + y)^\top (x + y) &= x^\top x + 2x^\top y + y^\top y \\ &\le |x|^2 + 2|x|\,|y| + |y|^2 = (|x| + |y|)^2 \end{aligned} \tag{7.2.5-E}$$

und damit $|x + y| \le |x| + |y|$.

Versehen mit der euklidischen Norm $|\cdot|$ ist der $\mathbb{R}^n$ vollständig.

⚒ **7.2.6 Beispiel.** Für Matrizen $A \in \mathbb{R}^{m\times n}$ ist in Analogie zur euklidischen Norm des $\mathbb{R}^n$ durch

$$\|A\|_F = \sqrt{\sum_{i=1}^{m}\sum_{j=1}^{n} |a_{i,j}|^2} = \sqrt{\operatorname{trace} A^\top A} \tag{7.2.6-A}$$

ebenso eine Norm definiert. Diese wird als *Frobeniusnorm*[2] bezeichnet. Es gilt unter Ausnutzung der Ungleichung (7.2.5-A) von Cauchy–Schwarz

$$|Ax|^2 = \sum_{i=1}^{n} \left|\sum_{j=1}^{n} a_{i,j} x_j\right|^2 \le \sum_{i=1}^{n}\sum_{j=1}^{n} |a_{i,j}|^2 \sum_{k=1}^{n} |x_k|^2 = \|A\|_F^2 |x|^2 \tag{7.2.6-B}$$

und damit

$$\|A\| := \sup_{|x|=1} |Ax| \le \|A\|_F. \tag{7.2.6-C}$$

[1] Hermann Amandus Schwarz, 1843–1921
[2] Georg Ferdinand Frobenius, 1849–1917

Das so definierte $\|A\|$ ist ebenso eine Norm auf $\mathbb{R}^{m\times n}$, die (euklidische) *Operatornorm* der Matrix. Sie ist schwerer zu berechnen, hat aber einfachere Eigenschaften und führt oft zu besseren Abschätzungen. Den Nachweis der Normeigenschaften überlassen wir als Übungsaufgabe. Direkt aus der Definition der Operatornorm ergeben sich die Ungleichungen

$$\|AB\| \le \|A\|\|B\| \qquad \text{und} \qquad \|A^\top\| \le \|A\|. \tag{7.2.6-D}$$

Da weiterhin die Ungleichung $|a_{i,j}| = |\mathbf{e}_i^\top A \mathbf{e}_j| \le \|A\|$ gilt, folgt

$$\|A\|_F^2 = \operatorname{trace}(A^\top A) \le n\|A^\top A\| \le n\|A^\top\|\|A\| \le n\|A\|^2 \tag{7.2.6-E}$$

und damit die Abschätzung $\|A\|_F \le \sqrt{n}\|A\|$.

⚒ **7.2.7 Beispiel.** Der Vektorraum $\mathrm{C}[a,b]$ der auf einem kompakten Intervall stetigen Funktionen wird mit der Supremumsnorm $\|f\|_\infty = \sup_{x\in[a,b]} |f(x)|$ zu einem normierten Raum. Gleiches gilt für den Vektorraum $\mathfrak{R}[a,b]$ der auf dem Intervall R-integrierbaren Funktionen versehen mit der Supremumsnorm.

Beide sind als metrische Räume vollständig, dies folgt aus Satz 6.2.3 und Satz 6.3.1.

7.2.8. Als nächstes erinnern wir an die ε-δ-Definition der Stetigkeit in metrischen Räumen. Eine Funktion $f : M_1 \to M_2$ zwischen metrischen Räumen (M_1, d_1) und (M_2, d_2) heißt *stetig* auf M_1, falls

$$\forall_{x\in M_1}\forall_{\varepsilon>0}\exists_{\delta>0} \qquad y \in B_\delta(x) \quad \Longrightarrow \quad f(y) \in B_\varepsilon(f(x)) \tag{7.2.8-A}$$

gilt. Mit den oben eingeführten Bezeichnungen können wir dies umformulieren.

Satz. *Eine Funktion $f : M_1 \to M_2$ ist genau dann stetig, wenn für jede offene Menge $U \subseteq M_2$ das Urbild*

$$f^{-1}[U] := \{x \in M_1 \mid f(x) \in U\} \tag{7.2.8-B}$$

offen in M_1 ist.

Beweis. [⇒] Sei f stetig und $U \subseteq M_2$ offen. Sei weiter $x \in M_1$ mit $f(x) \in U$. Da $f(x)$ innerer Punkt ist, existiert ein $\varepsilon > 0$ mit $B_\varepsilon(f(x)) \subseteq U$. Da f stetig ist, gibt es ein $\delta > 0$ mit

$$f[B_\delta(x)] = \{f(y) \mid y \in B_\delta(x)\} \subseteq B_\varepsilon(f(x)) \tag{7.2.8-C}$$

und x ist innerer Punkt von $f^{-1}[U]$.

[⇐] Da für jedes $x \in M_1$ die Menge $B_\varepsilon(f(x))$ offen in M_2 ist, muss ihr Urbild $f^{-1}[B_\varepsilon(f(x))]$ auch offen in M_1 sein. Da x darin enthalten ist, ist x innerer Punkt davon. Es existiert also ein $\delta > 0$ mit $B_\delta(x) \subseteq f^{-1}[B_\varepsilon(f(x))]$ und damit $f[B_\delta(x)] \subseteq B_\varepsilon(f(x))$ und es folgt die ε-δ-Charakterisierung der Stetigkeit. □

⚒ **7.2.9 Beispiel.** Die Koordinatenabbildungen

$$\mathbb{R}^n \ni x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \mapsto x_k \in \mathbb{R} \tag{7.2.9-A}$$

für $k = 1, \ldots, n$ sind stetig, was direkt aus $|x_k| \leq |x|$ folgt. Verkettungen, Summen und Produkte stetiger Funktionen sind wiederum stetig. Damit sind auf $\mathbb{R}^n$ auch alle Polynomfunktionen

$$\mathbb{R}^n \ni x \mapsto \sum_{|\alpha| \leq n} a_\alpha x^\alpha \tag{7.2.9-B}$$

definiert als Summe über *Multiindizes* $\alpha \in \mathbb{N}_0^n$ mit

$$x^\alpha = \prod_{k=1}^{n} x_k^{\alpha_k}, \qquad |\alpha| = \sum_{k=1}^{n} \alpha_k, \tag{7.2.9-C}$$

stetig. Also sind insbesondere auch die Spurabbildung $\operatorname{trace} : \mathbb{R}^{n\times n} \to \mathbb{R}$ und die Determinantenabbildung $\det : \mathbb{R}^{n\times n} \to \mathbb{R}$ stetig.

⚒ **7.2.10 Beispiel.** Da die Determinantenabbildung $\det : \mathbb{R}^{n\times n} \to \mathbb{R}$ stetig ist, ist die Menge der invertierbaren Matrizen

$$\mathrm{GL}(n) := \{A \in \mathbb{R}^{n\times n} \mid \det A \neq 0\} \subset \mathbb{R}^{n\times n} \tag{7.2.10-A}$$

als Urbild der offenen Menge $\mathbb{R} \setminus \{0\}$ offen.

Wir wollen dies als Alternative und ohne Bezugnahme auf die Stetigkeit der Determinante nachweisen. Dazu gehen wir in zwei Schritten vor. Angenommen, eine Matrix $B \in \mathbb{R}^{n\times n}$ erfüllt $\|B\| = r < 1$. Dann besitzt die Reihe

$$C = I + \sum_{k=1}^{\infty} B^k \tag{7.2.10-B}$$

wegen $\|B^k\| \leq \|B\|^k$ eine konvergente Majorante, die Partialsummen erfüllen für $m \geq n$

$$\left\| \sum_{k=1}^{m} B^k - \sum_{k=1}^{n-1} B^k \right\| = \left\| \sum_{k=n}^{m} B^k \right\| \leq \sum_{k=n}^{m} \|B\|^k \leq \sum_{k=n}^{\infty} r^k = \frac{r^n}{1-r} \longrightarrow 0 \tag{7.2.10-C}$$

für $m, n \to \infty$ und sind damit selbst eine Cauchyfolge in $\mathbb{R}^{n\times n}$. Weiter gilt als Teleskopreihe

$$(I - B)C = I - B + \sum_{k=1}^{\infty} \left(B^k - B^{k+1}\right) = I = C(I - B) \tag{7.2.10-D}$$

und somit ist $I - B$ invertierbar mit Inverser C. Sei nun $A \in \mathbb{R}^{n\times n}$ eine invertierbare Matrix und gelte $\|A^{-1}\|\|A - D\| < 1$. Dann ist die Matrix D ebenso invertierbar. Dazu schreiben wir $A^{-1}D$ als $I - A^{-1}(A - D)$ und nutzen obige Reihe in der Form

$$D^{-1} = (A^{-1}D)^{-1}A^{-1} = A^{-1} + \sum_{k=1}^{\infty} \left(A^{-1}(A - D)\right)^k A^{-1}. \tag{7.2.10-E}$$

Nach Konstruktion gilt $D^{-1}D = I$ aber ebenso liefert direktes Einsetzen in die Reihe $DD^{-1} = I$ und die gewünschte Invertierbarkeit folgt.

⚒ **7.2.11 Beispiel.** Wir wollen das gerade gezeigte Beispiel deutlich verallgemeinern. Dazu zeigen wir, dass für jedes $k \in \mathbb{N}$ die Menge

$$\{A \in \mathbb{R}^{m\times n} \mid \operatorname{rank} A \geq k\} \subseteq \mathbb{R}^{m\times n} \tag{7.2.11-A}$$

offen ist. Sei dazu $A \in \mathbb{R}^{m\times n}$ mit $\operatorname{rank} A = k$ gegeben. Dann finden wir k linear unabhängige Spalten und in diesen weiterhin k linear unabhängige Zeilen, es gibt also eine durch Streichen von Zeilen und Spalten erhaltene $k \times k$-Untermatrix A_0 mit vollem Rang. Für diese gilt damit $\det A_0 \neq 0$.

Bezeichne B_0 die aus einer beliebigen Matrix $B \in \mathbb{R}^{m\times n}$ durch Streichen derselben Zeilen und Spalten entstehende Untermatrix. Streichen von Zeilen und Spalten entspricht der Einschränkung auf einen Teilraum und dem Anwenden einer Orthogonalprojektion, es gilt $\|A_0\| \leq \|A\|$. Wenn wir also eine Matrix $B \in \mathbb{R}^{m\times n}$ mit $\|A-B\| < \varepsilon$ gegeben haben, so folgt $\|A_0 - B_0\| \leq \|A - B\| < \varepsilon$ und da die Determinante stetig ist, folgt $\det B_0 \neq 0$ für ε hinreichend klein. Damit folgt aber $\operatorname{rank} B \geq k$ und die Aussage ist gezeigt.

Wenn normierte Räume nicht endlichdimensional sind, sind nicht mehr alle linearen Abbildungen stetig. Dann hilft die folgende Aussage:

☞ **7.2.12 Lemma.** *Eine lineare Abbildung $f \in \operatorname{Lin}(V, W)$ zwischen normierten Räumen V und W ist genau dann stetig, wenn*

$$\|f\| := \sup_{\|x\|=1} \|f(x)\| < \infty \tag{7.2.12-A}$$

gilt. Man sagt auch, die Abbildung f ist beschränkt. *Die Menge der beschränkten linearen Abbildungen zwischen V und W bezeichnen wir mit $\mathcal{L}(V, W)$.*

Beweisskizze. Wir zeigen nur die Notwendigkeit, also dass beschränkte lineare Abbildungen wirklich stetig sind. Die Nullabbildung ist stetig. Wir nehmen deshalb an, dass $f \neq 0$ und damit auch $\|f\| \neq 0$ gilt.

Zu $\varepsilon > 0$ wählen wir $\delta = \varepsilon/\|f\|$ und erhalten $\|x - y\| < \delta$

$$\|f(x) - f(y)\| = \|f(x - y)\| \leq \|f\|\,\|x - y\| < \|f\|\,\delta = \varepsilon \tag{7.2.12-B}$$

und damit die Behauptung. □

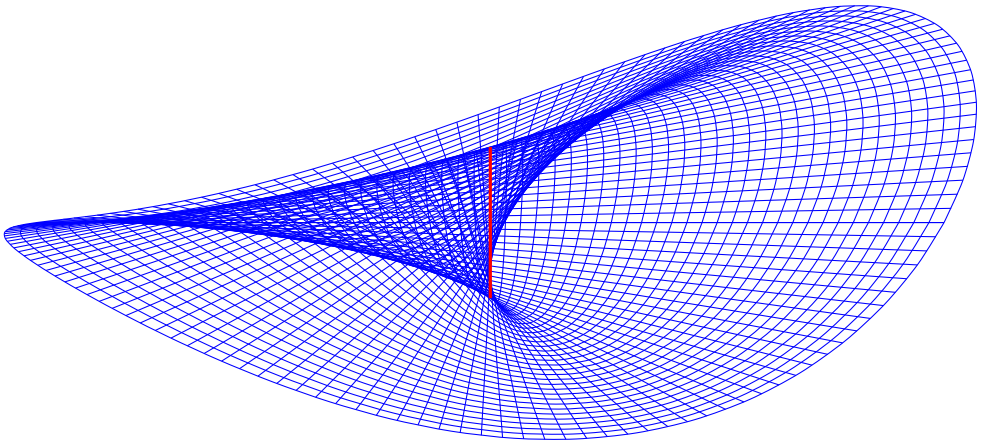
⚒ **7.2.13 Beispiel.** Das Riemannintegral

$$\mathfrak{R}[a, b] \ni f \mapsto \int_a^b f(x)\,\mathrm{d}x \in \mathbb{R} \tag{7.2.13-A}$$

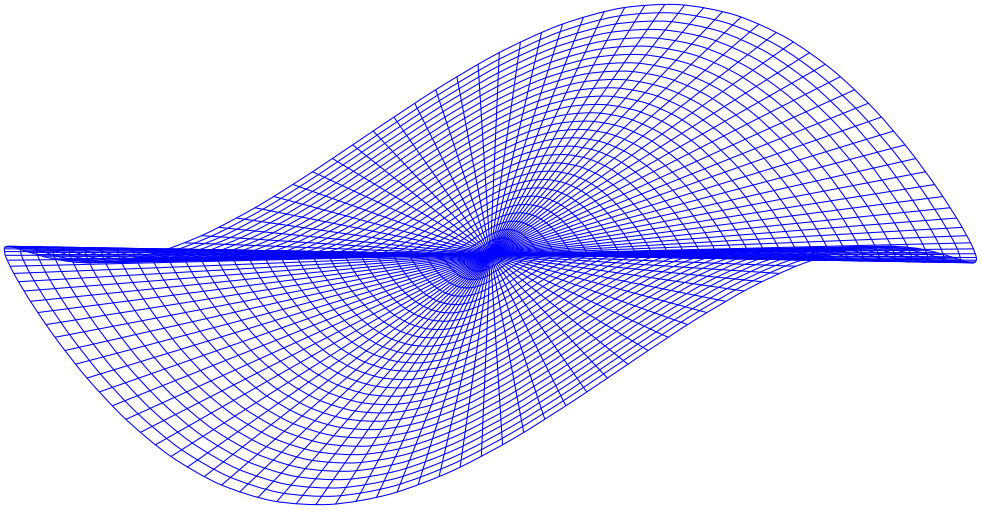
ist stetig. Dies folgt direkt aus der Abschätzung

$$\left|\int_a^b f(x)\,\mathrm{d}x\right| \leq (b - a)\|f\|_\infty. \tag{7.2.13-B}$$

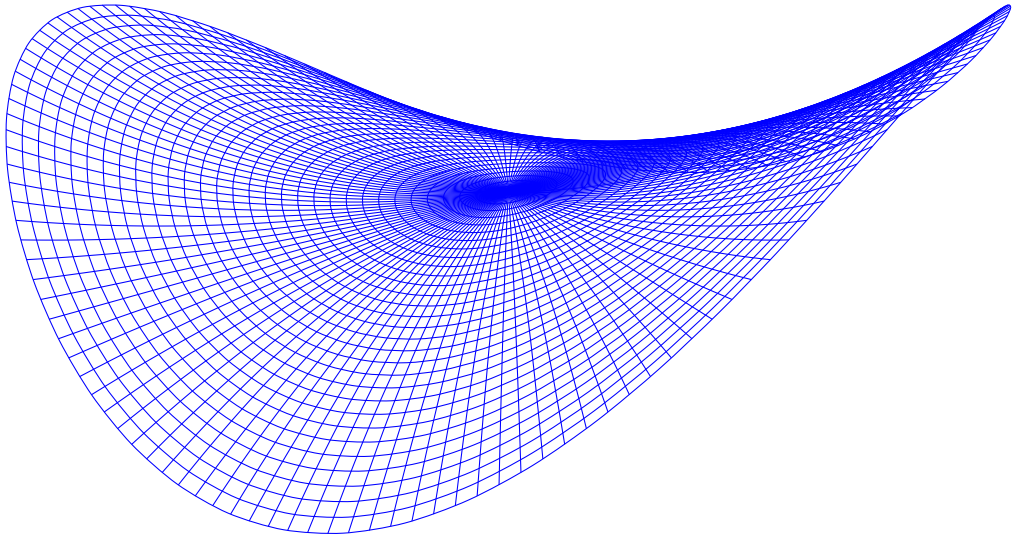
Wir haben dies schon genutzt, das ist genau die Aussage des Vertauschungssatzes 6.3.1.



Graph der Funktion $f(x,y) = \frac{xy}{x^2+y^2}$ mit Singularität über dem Punkt $x = y = 0$.



Graph der Funktion $f(x,y) = \frac{3x^2y-y^3}{(x^2+y^2)^{3/2}}$.



Graph der Funktion $f(x,y) = x^2 - y^2$.

7.3 Differenzierbarkeit

7.3.1. Wir formulieren die Differenzierbarkeit für Abbildungen zwischen normierten Räumen V und W. Das mag sehr allgemein klingen, in Anwendungen werden die Räume oft endlichdimensional und der Einfachheit halber durch $V = \mathbb{R}^n$ und $W = \mathbb{R}^m$ gegeben sein.

✍ **7.3.2 Definition.** *Sei $f : V \supseteq \Omega \to W$, Ω offen und $x_* \in \Omega$.*

(i) *Die Funktion f heißt in x_** Fréchet-differenzierbar[3] *oder kurz* differenzierbar*, falls es eine beschränkte lineare Abbildung $L : V \to W$ mit*

$$f(x) = f(x_*) + L(x - x_*) + \mathbf{o}(\|x - x_*\|), \qquad x \to x_*, \tag{7.3.2-A}$$

gibt. In diesem Fall wird die Abbildung L als Ableitung *der Funktion f in x_* bezeichnet. Wir verwenden die Notation $f'(x_*)$ für die Ableitung von f in x_*.*

(ii) *Die Funktion f heißt auf Ω* Fréchet-differenzierbar *oder kurz* differenzierbar*, falls f in jedem $x_* \in \Omega$ Fréchet-differenzierbar ist. In diesem Fall bezeichnet man*

$$f' : \Omega \to \mathcal{L}(V, W) \tag{7.3.2-B}$$

als Fréchet-Ableitung *oder kurz* Ableitung *der Funktion f.*

Die Bedingung (7.3.2-A) bestimmt die lineare Abbildung L eindeutig, wären L_1 und L_2 beschränkte lineare Abbildungen, die beide (7.3.2-A) erfüllen, so würde (mit $h = x - x_*$)

$$(L_1 - L_2)h = L_1 h - L_2 h = \mathbf{o}(\|h\|) \tag{7.3.2-C}$$

und damit $\|L_1 - L_2\| = 0$ folgen. Dies impliziert aber $L_1 = L_2$. Damit ist die in **(ii)** eingeführte Notation $f' : \Omega \to \mathcal{L}(V, W)$ für die Fréchetableitung wohldefiniert.

✂ **Ergänzung.** Die gewünschte Eindeutigkeit von L aus (7.3.2-A) ist der Grund, warum in obiger Definition $x_0 \in \Omega$ innerer Punkt beziehungsweise Ω offen sein sollte. Während die eindimensionale Ableitung in jedem Häufungspunkt des Definitionsbereichs definierbar ist, benötigt die höherdimensionale Fréchetableitung Grenzwerte aus *genügend vielen Richtungen* für die Charakterisierung der Ableitung.

☞ **7.3.3 Lemma.** *Sei $f : V \supseteq \Omega \to W$ in $x_* \in \Omega$ differenzierbar. Dann ist f in x_* stetig.*

Beweis. Dies folgt direkt aus der Definition. Es gilt

$$f(x) = f(x_*) + f'(x_*)(x - x_*) + \mathbf{o}(\|x - x_*\|), \qquad x \to x_*, \tag{7.3.3-A}$$

und sowohl der zweite als auch der dritte Term auf der rechten Seite strebt für $x \to x_*$ gegen Null. □

⚒ **7.3.4 Beispiel.** Sind $V = \mathbb{R}^n$ und $W = \mathbb{R}^m$ beide endlichdimensional, so entspricht der Raum $\mathcal{L}(V, W)$ gerade dem Raum der Matrizen $\mathbb{R}^{m\times n}$. Eine differenzierbare Funktion $f : \mathbb{R}^n \supseteq \Omega \to \mathbb{R}^m$ besitzt als Ableitung also eine matrixwertige Funktion

$$f' : \Omega \to \mathbb{R}^{m\times n}. \tag{7.3.4-A}$$

[3] René Maurice Fréchet,1878–1973

Die so als Matrix geschriebene Ableitung von f wird als *Jacobimatrix*[4] bezeichnet. Es stellt sich die Frage, wie man diese (Differenzierbarkeit vorausgesetzt) berechnet. Dazu notieren wir f genauer, verwenden die Komponenten der Vektoren $x \in \mathbb{R}^n$ als Variablen und bezeichnen die Einträge des Funktionswertes ebenso mit $f_1, \dots, f_m$. Es gilt also

$$f(x) = f(x_1, \dots, x_n) = \begin{bmatrix} f_1(x_1, \dots, x_n) \\ f_2(x_1, \dots, x_n) \\ \vdots \\ f_m(x_1, \dots, x_n) \end{bmatrix}. \tag{7.3.4-B}$$

Angenommen, diese Funktion ist differenzierbar im Punkt x. Dann gilt für jeden Vektor $h \in \mathbb{R}^n$ und mit der Jacobimatrix $L = f'(x)$

$$f(x+h) - f(x) = Lh + \mathbf{o}(|h|), \qquad h \to 0, \tag{7.3.4-C}$$

damit also speziell mit $h = t\mathbf{e}_j$ und für die i-Komponente

$$\begin{aligned} f_i(x_1, \dots, x_j + t, \dots x_n) - f_i(x_1, \dots, x_j, \dots, x_n) &= \mathbf{e}_i^\top \big(f(x + t\mathbf{e}_j) - f(x)\big) \\ &= \mathbf{e}_i^\top L\mathbf{e}_j\, t + \mathbf{o}(t). \end{aligned} \tag{7.3.4-D}$$

Damit ist der Matrixeintrag $L_{i,j} = \mathbf{e}_i^\top L\mathbf{e}_j$ der Jacobimatrix L als Ableitung der Funktion f_i nach der Variablen x_j (und der Behandlung aller weiteren Variablen als Konstanten) darstellbar. Dafür hat sich eine Notation eingebürgert, wir schreiben

$$\frac{\partial f_i}{\partial x_j} = \lim_{t \to 0} \frac{f_i(x_1, \dots, x_j + t, \dots x_n) - f_i(x_1, \dots, x_j, \dots, x_n)}{t} \tag{7.3.4-E}$$

für die *partielle Ableitung* von f_i nach der j-ten Variablen und erhalten damit für die Jacobimatrix $f'(x)$

$$f'(x) = \begin{bmatrix} \partial f_1/\partial x_1 & \partial f_1/\partial x_2 & \cdots & \partial f_1/\partial x_n \\ \partial f_2/\partial x_1 & \partial f_2/\partial x_2 & \cdots & \partial f_2/\partial x_n \\ \vdots & \vdots & & \vdots \\ \partial f_m/\partial x_1 & \partial f_m/\partial x_2 & \cdots & \partial f_m/\partial x_n \end{bmatrix}. \tag{7.3.4-F}$$

Umgekehrt genügt die Existenz aller partiellen Ableitungen in einem Punkt x, um einer Funktion $f : \mathbb{R}^n \supseteq \Omega \to \mathbb{R}^m$ eine Jacobimatrix zuordnen zu können. In diesem Fall sagen wir, f ist im Punkt $x \in \Omega$ *partiell differenzierbar*.

☞ **7.3.5 Satz.** *Sei $f : \mathbb{R}^n \supseteq \Omega \to \mathbb{R}^m$ auf Ω partiell differenzierbar. Hängt dann die Jacobimatrix* (7.3.4-F) *stetig von x ab, so ist f differenzierbar und die Ableitung durch* (7.3.4-F) *gegeben.*

Beweis. Es genügt, Differenzierbarkeit an jeder Stelle $x_* \in \Omega$ und für den Fall $m = 1$ zu zeigen. Da Ω offen ist, existiert ein $r > 0$ mit $B_{2r}(x_*) \subseteq \Omega$. Die als stetig vorausgesetzten partiellen Ableitungen

$$\Omega \ni x \mapsto \frac{\partial f}{\partial x_j} \in \mathbb{R} \tag{7.3.5-A}$$

[4] Carl Gustav Jacob Jacobi, 1804–1851

sind auf der kompakten Menge $K = \{x \in \mathbb{R}^n \mid |x - x_*| \leq r\} \subseteq B_{2r}(x_*) \subseteq \Omega$ gleichmäßig stetig. Damit gibt es zu jedem $\varepsilon > 0$ ein $\delta > 0$ mit

$$|x - \tilde{x}| < \delta \qquad \Longrightarrow \qquad \left| \frac{\partial f}{\partial x_j}(x) - \frac{\partial f}{\partial x_j}(\tilde{x}) \right| < \varepsilon \tag{7.3.5-B}$$

für $x, \tilde{x} \in K$. Also gilt auf Grund des 1. Mittelwertsatzes der Differentialrechnung für $h = \sum_{j=1}^n h_j \mathbf{e}_j \in \mathbb{R}^n$ mit $|h| < \delta$ und geeigneten $\theta_j \in (0,1)$

$$\begin{aligned} f(x_* + h) - f(x_*) &= \sum_{j=1}^n f\left(x_* + \sum_{k=1}^j h_k \mathbf{e}_k\right) - f\left(x_* + \sum_{k=1}^{j-1} h_k \mathbf{e}_k\right) \\ &= \sum_{j=1}^n \frac{\partial f}{\partial x_j}\left(x_* + \theta_j h_j \mathbf{e}_j + \sum_{k=1}^{j-1} h_k \mathbf{e}_k\right) h_j \\ &= \sum_{j=1}^n \frac{\partial f}{\partial x_j}(x_*) h_j + \varepsilon h_j \end{aligned} \tag{7.3.5-C}$$

und damit

$$f(x_* + h) - f(x_*) = \sum_{j=1}^n \frac{\partial f}{\partial x_j}(x_*) h_j + \mathbf{o}(|h|), \qquad h \to 0. \tag{7.3.5-D}$$

Also ist $f : \Omega \to \mathbb{R}$ differenzierbar.

Für $m > 1$ wendet man obiges Argument auf jede Komponente separat an. □

Eine (partiell) differenzierbare Funktion $f : \mathbb{R}^n \supseteq \Omega \to \mathbb{R}^m$, deren Jacobimatrix stetig ist, werden wir im Weiteren als *stetig differenzierbar* bezeichnen. Nicht jede differenzierbare Funktion ist auch stetig differenzierbar.

⚒ **7.3.6 Beispiel.** Der gerade gezeigte Satz ist eines der wichtigsten Kriterien zum Nachweis der Differenzierbarkeit. Dazu betrachten wir ein Beispiel und untersuchen die Funktion $f : \mathbb{R}^2 \to \mathbb{R}$ mit

$$f(x, y) = \sin(x^2 + y^2) \tag{7.3.6-A}$$

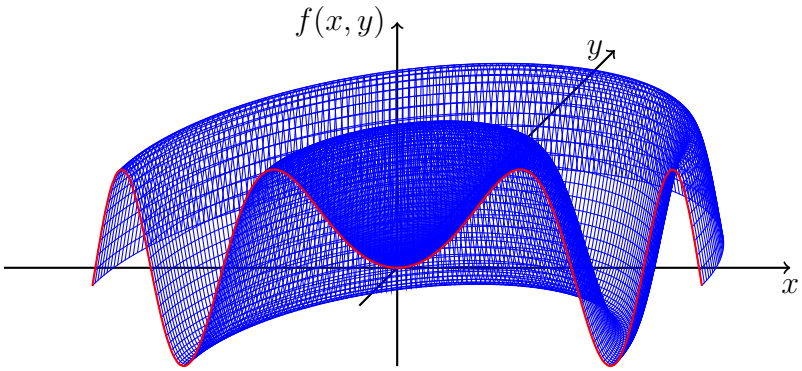
auf Differenzierbarkeit. Die Existenz partieller Ableitungen und damit die partielle Differenzierbarkeit ergibt sich aus den Regeln der eindimensionalen Differentialrechnung, die Funktionen

$$\frac{\partial f}{\partial x}(x, y) = 2x \cos(x^2 + y^2) \qquad \text{und} \qquad \frac{\partial f}{\partial y}(x, y) = 2y \cos(x^2 + y^2) \tag{7.3.6-B}$$

sind als Verkettungen stetiger Funktionen stetig. Damit ist f differenzierbar und die Fréchetableitung durch

$$f'(x, y) = \begin{bmatrix} 2x \cos(x^2 + y^2) & 2y \cos(x^2 + y^2) \end{bmatrix} = 2 \cos(x^2 + y^2) \begin{bmatrix} x & y \end{bmatrix} \tag{7.3.6-C}$$

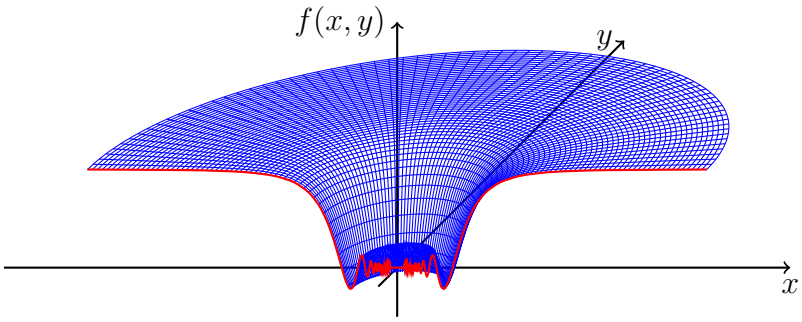
gegeben.



⚒ **7.3.7 Beispiel.** Satz 7.3.5 liefert nur ein hinreichendes Kriterium für die Differenzierbarkeit einer Funktion f. Mitunter hilft nur die direkte Anwendung der Definition. Als Beispiel dazu betrachten wir die Funktion $f : \mathbb{R}^2 \to \mathbb{R}$ mit

$$f(x,y) = \begin{cases} (x^2+y^2)\sin(\frac{1}{x^2+y^2}), & x^2+y^2 > 0, \\ 0, & x = y = 0. \end{cases} \tag{7.3.7-A}$$

Diese ist im Punkt $x = y = 0$ differenzierbar mit Fréchetableitung $f'(0,0) = 0$, allerdings sind die partiellen Ableitungen in der Umgebung des Punktes unbeschränkt und damit auch nicht stetig. Die Funktion f ist also differenzierbar, aber nicht stetig differenzierbar.



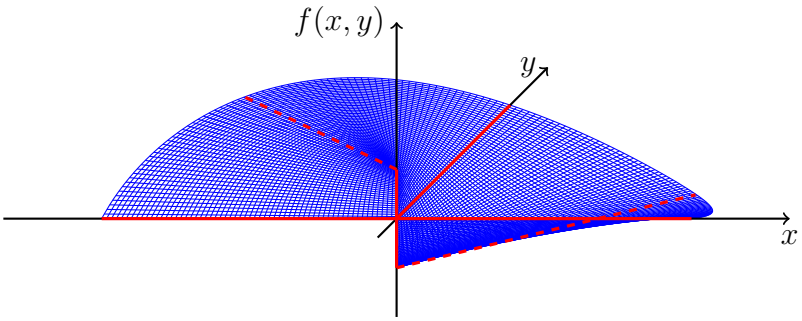
⚒ **7.3.8 Beispiel.** Partielle Differenzierbarkeit impliziert nicht die Stetigkeit der Funktion. Auch dazu ein einfaches Beispiel. Die Funktion $f : \mathbb{R}^2 \to \mathbb{R}$ mit

$$f(x,y) = \begin{cases} \frac{-xy}{x^2+y^2}, & x^2+y^2 > 0, \\ 0, & x = y = 0, \end{cases} \tag{7.3.8-A}$$

ist im Ursprung nicht stetig, obwohl beide partiellen Ableitungen

$$\frac{\partial f}{\partial x}(0,0) = 0 = \frac{\partial f}{\partial y}(0,0) \tag{7.3.8-B}$$

existieren.



⚒ **7.3.9 Beispiel.** Auf den normierten Raum $\mathrm{C}[a,b]$ der auf $[a,b]$ stetigen Funktionen versehen mit Supremumsnorm $\|\cdot\|_\infty$ ist die Abbildung

$$\Psi : \mathrm{C}[a,b] \ni f \mapsto \int_a^b \sin f(x)\,\mathrm{d}x \in \mathbb{R} \tag{7.3.9-A}$$

Fréchet-differenzierbar. Um dies zu zeigen, nutzen wir direkt die Definition der Differenzierbarkeit. Sei dazu $h \in \mathrm{C}[a,b]$. Dann gilt unter Ausnutzung der Additionstheoreme

$$\begin{aligned}\Psi(f+h) &= \int_a^b \sin(f(x)+h(x))\,\mathrm{d}x \\ &= \int_a^b \big(\sin f(x)\cos h(x) + \cos f(x)\sin h(x)\big)\,\mathrm{d}x\end{aligned} \tag{7.3.9-B}$$

und untersuchen die verbleibenden Terme auf der rechten Seite. Unter Nutzung des Lagrange-Restgliedes erhalten wir

$$|\cos\eta - 1| \le \frac{1}{2}\eta^2, \qquad |\sin\eta - \eta| \le \frac{1}{6}|\eta|^3 \tag{7.3.9-C}$$

und damit

$$\begin{aligned}\Psi(f+h) &= \Psi(f) + \int_a^b h(x)\cos f(x)\,\mathrm{d}x + \int_a^b \sin f(x)\big(\cos h(x) - 1\big)\,\mathrm{d}x \\ &\qquad + \int_a^b \cos f(x)\big(\sin h(x) - h(x)\big)\,\mathrm{d}x \\ &= \Psi(f) + \int_a^b h(x)\cos f(x)\,\mathrm{d}x + \mathbf{o}(\|h\|), \qquad h \to 0.\end{aligned} \tag{7.3.9-D}$$

Ebenso gilt

$$\left|\int_a^b h(x)\cos f(x)\,\mathrm{d}x\right| \le \|h\|_\infty (b-a) \tag{7.3.9-E}$$

und der verbleibende lineare Term ist beschränkt. Also ist Ψ differenzierbar und die Fréchet-Ableitung durch

$$\Psi'(f) : \mathrm{C}[a,b] \ni h \mapsto \int_a^b h(x)\cos f(x)\,\mathrm{d}x \in \mathbb{R} \tag{7.3.9-F}$$

gegeben. Die Ableitung Ψ' hängt in diesem Fall stetig von f ab, Ψ ist also insbesondere stetig differenzierbar: Gilt $\|f-g\|_\infty \le \varepsilon$, so folgt wegen $|\cos(\xi)-\cos(\zeta)| \le |\xi-\zeta|$

$$\left|\Psi'(f)[h]-\Psi'(g)[h]\right| \le \|h\|_\infty \int_a^b |\cos f(x)-\cos(g(x))|\,\mathrm{d}x \le \varepsilon\|h\|_\infty\|f-g\|_\infty \quad (7.3.\textbf{9-G})$$

stets $\|\Psi'(f)-\Psi'(g)\| \le \varepsilon(b-a)$.

7.3.10. Mit differenzierbaren Funktionen kann man wiederum rechnen. Ganz analog zum eindimensionalen Fall sieht man, dass die Summe zweier differenzierbarer Funktionen f und g und skalare Vielfache einer Funktion f differenzierbar sind. Es gilt also

$$(f+g)'(x) = f'(x)+g'(x) \quad (7.3.\textbf{10-A})$$

und ebenso

$$(\lambda f)'(x) = \lambda f'(x). \quad (7.3.\textbf{10-B})$$

Beides folgt direkt aus der definierenden Bedingung (7.3.**2-A**), Nachrechnen überlassen wir als Übung. Interessanter ist die Verkettung von Funktionen, die wir als eigenen Satz formulieren.

☞ **7.3.11 Satz** (Kettenregel). *Seien V_1, V_2 und V_3 normierte Räume. Angenommen $f : V_1 \supseteq \Omega_1 \to \Omega_2 \subseteq V_2$ und $g : V_2 \supseteq \Omega_2 \to V_3$ und f ist in x_* differenzierbar und g in $f(x_*)$. Dann ist $g\circ f : \mathbb{R}^n \supseteq \Omega_1 \to \mathbb{R}^k$ in x_* differenzierbar und es gilt*

$$(g\circ f)'(x_*) = g'(f(x_*))f'(x_*). \quad (7.3.\textbf{11-A})$$

Beweis. Wir nutzen die Definition der Ableitung

$$f(x)-f(x_*) = f'(x_*)(x-x_*)+\mathbf{o}(\|x-x_*\|), \qquad x\to x_* \quad (7.3.\textbf{11-B})$$

für f und ebenso für g an der Stelle $f(x_*)$ mit $f(x)$ in der Nähe von $f(x_*)$ (was durch x nahe x_* auf Grund der Stetigkeit von f in x_* garantiert ist) und erhalten

$$\begin{aligned} g(f(x))-g(f(x_*)) &= g'(f(x_*))(f(x)-f(x_*))+\mathbf{o}(\|f(x)-f(x_*)\|) \\ &= g'(f(x_*))f'(x_*)(x-x_*) \\ &\quad + g'(f(x_*))\mathbf{o}(\|x-x_*\|)+\mathbf{o}(\|f(x)-f(x_*)\|). \end{aligned} \quad (7.3.\textbf{11-C})$$

für $x\to x_*$. Da $g'(f(x_*)) \in \mathcal{L}(V_2,V_3)$ beschränkt ist, gilt[5]

$$g'(f(x_*))\mathbf{o}(\|x-x_*\|) \subseteq \mathbf{o}(\|x-x_*\|). \quad (7.3.\textbf{11-D})$$

Da f in x_* differenzierbar ist gilt weiterhin

$$\|f(x)-f(x_*)\| \le \|f'(x_*)\|\|x-x_*\|+\mathbf{o}(\|x-x_*\|) \quad (7.3.\textbf{11-E})$$

und damit $\mathbf{o}(\|f(x)-f(x_*)\|) \subseteq \mathbf{o}(\|x-x_*\|)$ und die Aussage ist bewiesen. □

[5] Da $\mathbf{o}(\|h\|)$ die Menge aller Funktionen f mit $\lim_{h\to} \|f(h)\|/\|h\| = 0$ ist, wäre statt $= \mathbf{o}(...)$ eigentlich immer ein $\in \mathbf{o}(...)$ oder $\subseteq \mathbf{o}(...)$ besser. Wir nutzen diese Notation nur hier.

⚒ **7.3.12 Beispiel.** Im Endlichdimensionalen entspricht die Kettenregel einer Matrixmultiplikation für Jacobimatrizen. Wir zeigen dies an zwei Beispielen.

(i) Die Produktregel für Ableitungen auf $\mathbb{R}$ ist ein Spezialfall der Kettenregel. Seien dazu $f_1 : \mathbb{R} \to \mathbb{R}$ und $f_2 : \mathbb{R} \to \mathbb{R}$ differenzierbar. Betrachtet man dann die Funktionen

$$f(x) = \begin{bmatrix} f_1(x) \\ f_2(x) \end{bmatrix}, \qquad g(u, v) = uv, \tag{7.3.12-A}$$

so sind diese differenzierbar und es gilt $(g \circ f)(x) = f_1(x)f_2(x)$. Weiter gilt mit der Kettenregel und unter Nutzung von $u = f_1(x)$ und $v = f_2(x)$

$$(g \circ f)' = \begin{bmatrix} v & u \end{bmatrix} \begin{bmatrix} f_1'(x) \\ f_2'(x) \end{bmatrix} = f_2(x)f_1'(x) + f_1(x)f_2'(x) \tag{7.3.12-B}$$

und damit die bekannte Produktregel.

(ii) Als zweites Beispiel betrachten wir eine differenzierbare Funktion $f : \mathbb{R}^n \to \mathbb{R}$ und eine weitere differenzierbare Funktion $x : \mathbb{R} \to \mathbb{R}^n$. Dann ist die Verkettung $t \mapsto f(x(t))$ differenzierbar und es gilt mit der Kettenregel

$$\frac{\mathrm{d}}{\mathrm{d}t} f(x(t)) = \begin{bmatrix} \frac{\partial f}{\partial x_1} & \frac{\partial f}{\partial x_2} & \cdots & \frac{\partial f}{\partial x_n} \end{bmatrix} \begin{bmatrix} \dot{x}_1(t) \\ \dot{x}_2(t) \\ \vdots \\ \dot{x}_n(t) \end{bmatrix} = \sum_{j=1}^{n} \frac{\partial f}{\partial x_j} \dot{x}_j(t) \tag{7.3.12-C}$$

unter Ausnutzung der Kurzschreibweise $\dot{x}_j(t) = \frac{\mathrm{d}x_j}{\mathrm{d}t}$ für Ableitungen nach t.

7.3.13. Eine weitere Konsequenz formulieren wir direkt. Eine bijektive und differenzierbare Abbildung $f : \mathbb{R}^n \supseteq \Omega_1 \to \Omega_2 \subseteq \mathbb{R}^n$ heißt *Diffeomorphismus*, falls die Umkehrfunktion f^{-1} ebenfalls differenzierbar ist. Dafür ist es notwendig, dass mit $g = f^{-1}$

$$I = g'(f(x))f'(x) \tag{7.3.13-A}$$

gilt und damit, dass die Jacobimatrix $f'(x)$ für jedes $x \in \Omega_1$ invertierbar ist, also

$$\det f'(x) \neq 0 \tag{7.3.13-B}$$

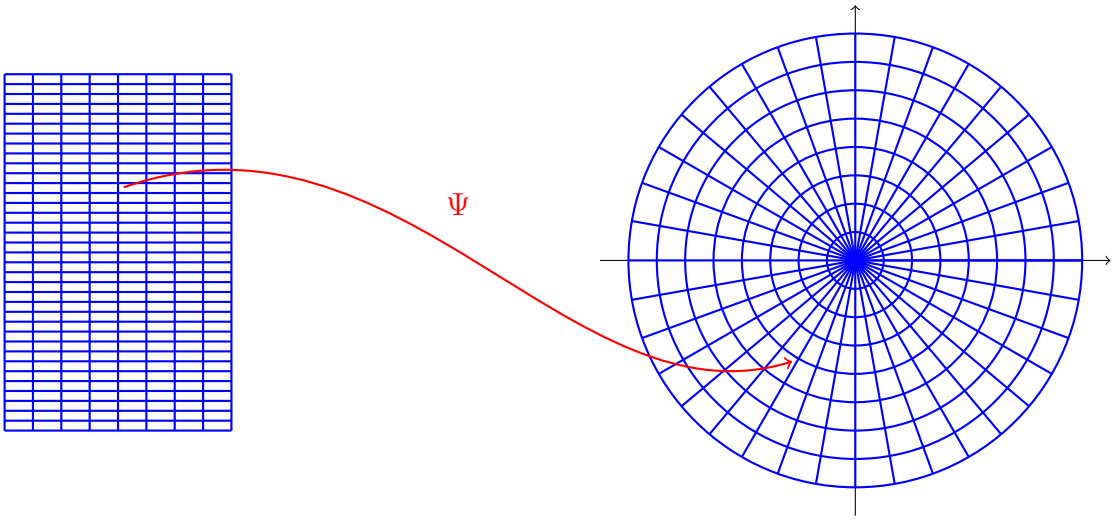
gilt. Wir werden im nächsten Abschnitt zeigen, dass dies (zumindest lokal) auch hinreichend ist. Vorerst betrachten wir ein weiteres Beispiel:

⚒ **7.3.14 Beispiel.** Als Beispiel betrachten wir *Polarkoordinaten in der Ebene*, dazu ordnen wir jedem $r \in \mathbb{R}_{\geq 0}$ und $\theta \in [0, 2\pi)$ den Punkt mit Abstand r zum Ursprung und Winkel θ zur x_1-Achse in $\mathbb{R}^2$ zu. Dem entspricht die Funktion

$$\Psi : \begin{bmatrix} r \\ \theta \end{bmatrix} \mapsto \begin{bmatrix} r\cos\theta \\ r\sin\theta \end{bmatrix} \tag{7.3.14-A}$$

mit Jacobimatrix

$$\Psi'(r, \theta) = \begin{bmatrix} \partial x_1/\partial r & \partial x_1/\partial\theta \\ \partial x_2/\partial r & \partial x_2/\partial\theta \end{bmatrix} = \begin{bmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{bmatrix}. \tag{7.3.14-B}$$



Die Abbildung Ψ ist also differenzierbar. Für $r > 0$ ist die Abbildung ebenso bijektiv und ein Diffeomorphismus. Um dies zu zeigen, kann man die Umkehrabbildung ausrechnen und ableiten (und muss dabei beachten, dass die stückweise definierte Abbildung differenzierbar zusammengesetzt ist). Wir überlassen dies als Übung, geben aber der Vollständigkeit halber die Ableitung der Umkehrfunktion direkt (und in beiden Koordinaten) an. Es gilt

$$\det \Psi'(r,\theta) = r\left(\cos^2\theta + \sin^2\theta\right) = r \tag{7.3.**14-C**}$$

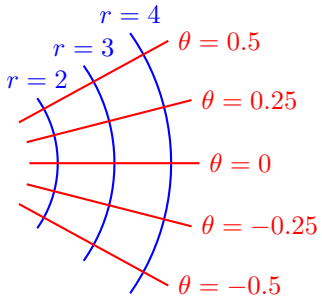
und damit

$$(\Psi^{-1})' \circ \Psi(r,\theta) = \begin{bmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{bmatrix}^{-1} = \frac{1}{r}\begin{bmatrix} r\cos\theta & r\sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}. \tag{7.3.**14-D**}$$

Setzt man die Definition der Polarkoordinaten ein, so erhält man damit

$$(\Psi^{-1})'(x_1,x_2) = \begin{bmatrix} \frac{x_1}{\sqrt{x_1^2+x_2^2}} & \frac{x_2}{\sqrt{x_1^2+x_2^2}} \\ -\frac{x_2}{x_1^2+x_2^2} & \frac{x_1}{x_1^2+x_2^2} \end{bmatrix}. \tag{7.3.**14-E**}$$

✂ **Ergänzung.** Es ist besser, sich *Koordinaten* auf Ω als Abbildungen vorzustellen, die den jeweiligen Punkten $x \in \Omega$ ihre Koordinaten zuordnen (statt wie in dem gerade diskutierten Beispiel den Koordinaten Punkte zuzuordnen). Wenn man in obigem Beispiel statt Ψ zwei *Koordinatenfunktionen* $r : \mathbb{R}^2 \setminus \{0\} \to \mathbb{R}_{>0}$ und $\theta : \mathbb{R}^2 \setminus \{0\} \to [0, 2\pi)$ betrachtet, so sind die eingezeichneten *Koordinatenlinien* gerade Lösungsmengen zu $r(x_1,x_2) = r$ und $\theta(x_1,x_2) = \theta$ für vorgegebene Werte r beziehungsweise θ.



Damit Koordinaten sinnvoll sind, sollten sich Koordinatenlinien stets nichttangential schneiden. Dies entspricht der linearen Unabhängigkeit der Ableitungen der Koordinatenfunktionen und damit der Invertierbarkeit der Jacobimatrix der aus den Koordinatenabbildungen zusammengesetzten Funktion.

7.4 Hauptsätze über differenzierbare Funktionen

☞ **7.4.1 Lemma** (Mittelwertungleichung). *Sei $\Omega \subseteq \mathbb{R}^n$ konvex und $f : \Omega \to \mathbb{R}^m$ stetig differenzierbar. Dann gilt für*

$$|f(x) - f(y)| \leq M|x - y| \tag{7.4.1-A}$$

mit

$$M = \max_{\theta\in[0,1]} \|f'(\theta x + (1-\theta)y)\|. \tag{7.4.1-B}$$

Beweis. Wir betrachten die vektorwertige Hilfsfunktion $h : [0,1] \to \mathbb{R}^m$

$$h(\theta) = f(\theta x + (1-\theta)y) \tag{7.4.1-C}$$

mit

$$h'(\theta) = f'(\theta x + (1-\theta)y)(x-y). \tag{7.4.1-D}$$

Also gilt für jeden Eintrag $h_i'(\theta) = f_i'(\theta x + (1-\theta)y)(x-y)$ und damit

$$h_i(1) - h_i(0) = \int_0^1 h_i'(\theta)\,\mathrm{d}\theta = \int_0^1 f_i'(\theta x + (1-\theta)y)(x-y)\,\mathrm{d}\theta, \tag{7.4.1-E}$$

was wir als

$$f(x) - f(y) = h(1) - h(0) = \int_0^1 h'(\theta)\,\mathrm{d}\theta = \int_0^1 f'(\theta x + (1-\theta)y)(x-y)\,\mathrm{d}\theta \tag{7.4.1-F}$$

notieren. Auch dieses Integral ist als Grenzwert von Riemannsummen darstellbar und wegen

$$|h'(\theta)| \leq \|f'(\theta x + (1-\theta)y)\|\,|x-y| \tag{7.4.1-G}$$

folgt die gesuchte Abschätzung. □

✂ **Ergänzung.** Die Mittelwertungleichung gilt deutlich allgemeiner als gerade bewiesen, nämlich für differenzierbare Abbildungen $f : V \supseteq \Omega \to W$ zwischen normierten Räumen. Wir betrachten dieselbe Hilfsfunktion $h : [0,1] \to W$, diese ist wiederum differenzierbar und erfüllt $\|h'(\theta)\| \leq \tilde{M}$ auf $[0,1]$. Wir wählen $\varepsilon > 0$ und zeigen dafür, dass dann für alle $t \in (0,1)$

$$\|h(t) - h(0)\| \leq (\tilde{M} + \varepsilon)t + \varepsilon \tag{7.4.1-H}$$

gilt. Dafür betrachten wir die Menge U aller $t \in [0,1]$ für die (7.4.1-H) nicht gilt und zeigen, dass diese leer sein muss. Die Menge U ist offen, da

$$U = \{t \mid \|h(t) - h(0)\| > (\tilde{M} + \varepsilon)t + \varepsilon\} = \{t \mid \|h(t) - h(0)\| - (\tilde{M} + \varepsilon)t - \varepsilon > 0\} \subseteq (0,1) \tag{7.4.1-I}$$

und $t \mapsto \|h(t) - h(0)\| - (\tilde{M} + \varepsilon)t - \varepsilon$ stetig ist. Angenommen, $U \neq \varnothing$. Dann erfüllt $c = \inf U$ sowohl $c > 0$ (da (7.4.1-H) für $t = 0$ als echte Ungleichung gilt und damit wegen Stetigkeit auch für einige $t > 0$ gelten muss) als auch $c < 1$ (da $U \subset (0,1)$ nichtleer angenommen wurde). Für dieses c gilt wegen

$$h(c+\theta) - h(c) = h'(c)\theta + \mathbf{o}(\theta) \tag{7.4.1-J}$$

für $0 \leq \theta < \eta$ und mit η hinreichend klein

$$\|h(c+\theta) - h(c)\| \leq \|h'(c)\|\theta + \varepsilon\theta \leq (\tilde{M} + \varepsilon)\theta. \tag{7.4.1-K}$$

Damit folgt aber (da $c \notin U$ und für c somit (7.4.**1-H**) gilt)

$$\|h(c+\theta)-h(0)\| \le \|h(c+\theta)-h(c)\|+\|h(c)-h(0)\| \le (\tilde{M}+\varepsilon)\theta+(\tilde{M}+\varepsilon)c+\varepsilon = (\tilde{M}+\varepsilon)(c+\theta)+\varepsilon \quad (7.4.\textbf{1-L})$$

im Widerspruch zur Wahl von c.

Auf Grund der Stetigkeit von h gilt (7.4.**1-H**) auch für $t = 1$ und $\varepsilon = 0$ und damit folgt die Behauptung

$$\|f(x) - f(y)\| = \|h(1) - h(0)\| \le \tilde{M} \le \|x - y\| \sup_{\theta\in[0,1]} \|f'(\theta x + (1 - \theta)y)\|. \quad (7.4.\textbf{1-M})$$

Es gilt also

✂ **Lemma** (Mittelwertungleichung). *Seien V und W normierte Räume, $\Omega \subseteq V$ konvex und $f : \Omega \to W$ differenzierbar. Dann*

$$\|f(x) - f(y)\| \le \|x - y\| \sup_{\theta\in[0,1]} \|f'(\theta x + (1 - \theta)y)\|. \quad (7.4.\textbf{1-N})$$

☞ **7.4.2 Korollar.** *Angenommen, $f : \mathbb{R}^n \supseteq \Omega \to \mathbb{R}^m$ ist stetig differenzierbar auf einer offenen Menge Ω. Dann ist f lokal lipschitzstetig.*

Eine Abbildung $\Phi : M \to M$ eines metrischen Raumes in sich heißt *kontrahierend*, falls es eine Konstante $K < 1$ mit

$$\forall_{x,y\in M} \quad d(\Phi(x), \Phi(y)) \le Kd(x, y) \quad (7.4.\textbf{2-A})$$

gibt. Kontrahierende Abbildungen sind insbesondere stetig. Darüberhinaus besitzen kontrahierende Abbildungen in vollständigen metrischen Räumen stets Fixpunkte.

☞ **7.4.3 Lemma** (Banachscher[6] Fixpunktsatz). *Sei (M, d) nichtleerer vollständiger metrischer Raum und $\Phi : M \to M$ eine kontrahierende Abbildung. Dann gibt es genau ein $x \in M$ mit $\Phi(x) = x$.*

Beweis. In einem ersten Schritt zeigen wir die Existenz. Sei dazu $x_0 \in M$ beliebig gewählt. Wir betrachten die rekursiv definierte Folge

$$x_{k+1} = \Phi(x_k). \quad (7.4.\textbf{3-A})$$

Es gilt mit dem Kontraktionsfaktor $K < 1$ der Abbildung

$$d(x_2, x_1) = d(\Phi(x_1), \Phi(x_0)) \le Kd(x_1, x_0) \quad (7.4.\textbf{3-B})$$

und per Induktion folgt

$$d(x_{k+1}, x_k) \le K^k d(x_1, x_0). \quad (7.4.\textbf{3-C})$$

Für $l > k$ folgt damit auf Grund der Dreiecksungleichung

$$d(x_l, x_k) \le \sum_{j=k}^{l-1} d(x_{j+1}, x_j) \le d(x_1, x_0) \sum_{j=k}^{l-1} K^j \le d(x_1, x_0)\frac{K^k}{1 - K} \quad (7.4.\textbf{3-D})$$

[6] Stefan Banach, 1892–1945

und die Folge $(x_k)_{k\in\mathbb{N}_0}$ ist Cauchy. Da (M,d) vollständig ist, konvergiert die Folge. Sei

$$x = \lim_{k\to\infty} x_k. \tag{7.4.3-E}$$

Dann folgt aus (7.4.3-A) auf Grund der Stetigkeit von Φ für $k \to \infty$

$$x = \Phi(x) \tag{7.4.3-F}$$

und x ist ein Fixpunkt von Φ. Dabei folgt mit $k = 0$ aus (7.4.3-D)

$$d(x, x_0) \le d(x_1, x_0) \sum_{k=0}^{\infty} K^k = \frac{1}{1-K} d(x_1, x_0). \tag{7.4.3-G}$$

Solche Fixpunkte sind eindeutig, da für jeden Fixpunkt y

$$d(x, y) = d(\Phi(x), \Phi(y)) \le K d(x, y) \tag{7.4.3-H}$$

gilt und damit $d(x, y) = 0$ sein muss. □

Wir formulieren den folgenden Umkehrsatz in vollster Allgemeinheit. Für Anwendungen wird später wieder vor allem der Fall des $\mathbb{R}^n$ von Bedeutung sein.

☞ **7.4.4 Satz** (lokaler Umkehrsatz). *Sei W vollständiger normierter Raum, $\Omega \subseteq W$ offen und $f : \Omega \to W$ stetig differenzierbar. Ist dann für ein $x_* \in \Omega$*

$$f'(x_*) : W \to W \text{ invertierbar mit } (f'(x_*))^{-1} \in \mathcal{L}(W, W) \tag{7.4.4-A}$$

so existieren eine Umgebungen U von x_* und V von $f(x_*)$, so dass $f : U \to V$ bijektiv und die Umkehrabbildung $f^{-1} : V \to U$ differenzierbar ist.

Beweis. Wir stellen zuerst fest, dass $f_1(x) = f'(x_*)^{-1} f(x)$ ebenso differenzierbar ist und zusätzlich $f_1'(x_*) = I$ erfüllt. Weiter ist $f : U \to V$ genau dann bijektiv, wenn auch $f_1 : U \to f'(x_*)^{-1}[V]$ bijektiv ist. Damit genügt es, die Aussage für Funktionen f mit $f'(x_*) = I$ zu beweisen. Wir gehen in mehreren Schritten vor.

Schritt 1: In einem ersten Schritt zeigen wir, dass f auf einer Umgebung stetig invertierbar ist. Dazu betrachten wir die Hilfsfunktionen

$$\Phi(x) = f(x) - x. \tag{7.4.4-B}$$

Die Funktion Φ ist ebenso stetig differenzierbar und erfüllt $\Phi'(x_*) = f'(x_*)^{-1} - I = 0$. Also gibt es eine Umgebung $B_r(x_*)$ auf der $\|\Phi'(x)\| \le \frac{1}{2}$ gilt und damit für $x, y \in B_r(x_*)$ nach der Mittelwertungleichung 7.4.1

$$\|\Phi(x) - \Phi(y)\| \le \frac{1}{2}\|x - y\|. \tag{7.4.4-C}$$

Damit ist Φ kontrahierend und es folgt für $x, y \in B_r(x_*)$ wegen

$$\big(f(x) - f(y)\big) = (x - y) - \big(\Phi(x) - \Phi(y)\big) \tag{7.4.4-D}$$

die Abschätzung

$$\|f(x)-f(y)\| \geq \|x-y\| - \|\Phi(x)-\Phi(y)\| \geq \frac{1}{2}\|x-y\|. \tag{7.4.4-E}$$

Damit ist aber $f : B_r(x_*) \to W$ injektiv. Sei nun $z \in W$ mit $\|z - f(x_*)\| < \frac{1}{2}r$. Dann gilt

$$\|z-\Phi(y)-x_*\| = \|z-f(x_*)+\Phi(x_*)-\Phi(y)\| < \frac{1}{2}r + \frac{1}{2}\|x_*-y\| = r \tag{7.4.4-F}$$

und $y \mapsto z - \Phi(y)$ bildet von $B_r(x_*)$ nach $B_r(x_*)$ ab. Die Abbildung ist kontrahierend. Es gibt also nach dem Banachschen Fixpunktsatz ein y im Abschluss von $B_r(x_*)$ mit

$$y = z - \Phi(y) = z - f(y) + y \tag{7.4.4-G}$$

und damit $f(y) = z$, sowie unter Nutzung von (7.4.3-G)

$$\|y - x_*\| \leq 2\|z - \Phi(x_*) - x_*\| = 2\|z - f(x_*)\| < r \tag{7.4.4-H}$$

und damit $y \in B_r(x_*)$.

Setzen wir nun $V = B_{r/2}(f(x_*))$ und $U = f^{-1}[V] \cap B_r(x_*)$, so ist U als Urbild einer offenen Menge offen und nach dem gerade gezeigten ist $f : U \to V$ bijektiv. Weiter ist $f^{-1} : V \to U$ nach (7.4.4-E) ebenso (Lipschitz) stetig.

Schritt 2: Nun zeigen wir, dass die Funktion $f^{-1} : V \to U$ differenzierbar ist. Für jedes $x \in U$ ist wegen

$$\|f'(x) - I\| = \|\Phi'(x)\| \leq \frac{1}{2} \tag{7.4.4-I}$$

die Abbildung $f'(x) \in \mathcal{L}(W, W)$ invertierbar. Seien nun $x, y \in U$ hinreichend nah. Dann gilt

$$f(x) - f(y) = f'(x)(x-y) + \mathbf{o}(\|x-y\|), \qquad y \to x, \tag{7.4.4-J}$$

und somit nach Anwenden der Inversen $f'(x)^{-1}$

$$(x-y) = f'(x)^{-1}\big(f(x)-f(y)\big) + \mathbf{o}(\|x-y\|), \qquad y \to x \tag{7.4.4-K}$$

Weiter gilt wegen (7.4.4-E) schon $\mathbf{o}(\|x-y\|) \subseteq \mathbf{o}(\|f(x)-f(y)\|)$ und $y \to x$ ist äquivalent zu $f(y) \to f(x)$. Das ist damit aber gerade die Bedingung für die Differenzierbarkeit der Umkehrfunktion f^{-1} im Punkt $f(x)$, es folgt $(f^{-1})'(f(x)) = (f')^{-1}(x)$. □

Beschränken wir uns auf den Fall differenzierbarer Abbildungen $\mathbb{R}^n \to \mathbb{R}^n$, so sind wir wiederum in der Sprache der Jacobimatrizen.

☞ **7.4.5 Korollar.** *Sei $f : \mathbb{R}^n \supseteq \Omega \to \mathbb{R}^n$ stetig differenzierbar und gelte für ein $x_* \in \Omega$*

$$\det f'(x_*) \neq 0. \tag{7.4.5-A}$$

Dann gibt es eine Umgebung U von x_* auf der $f : U \to f[U]$ ein Diffeomorphismus ist.

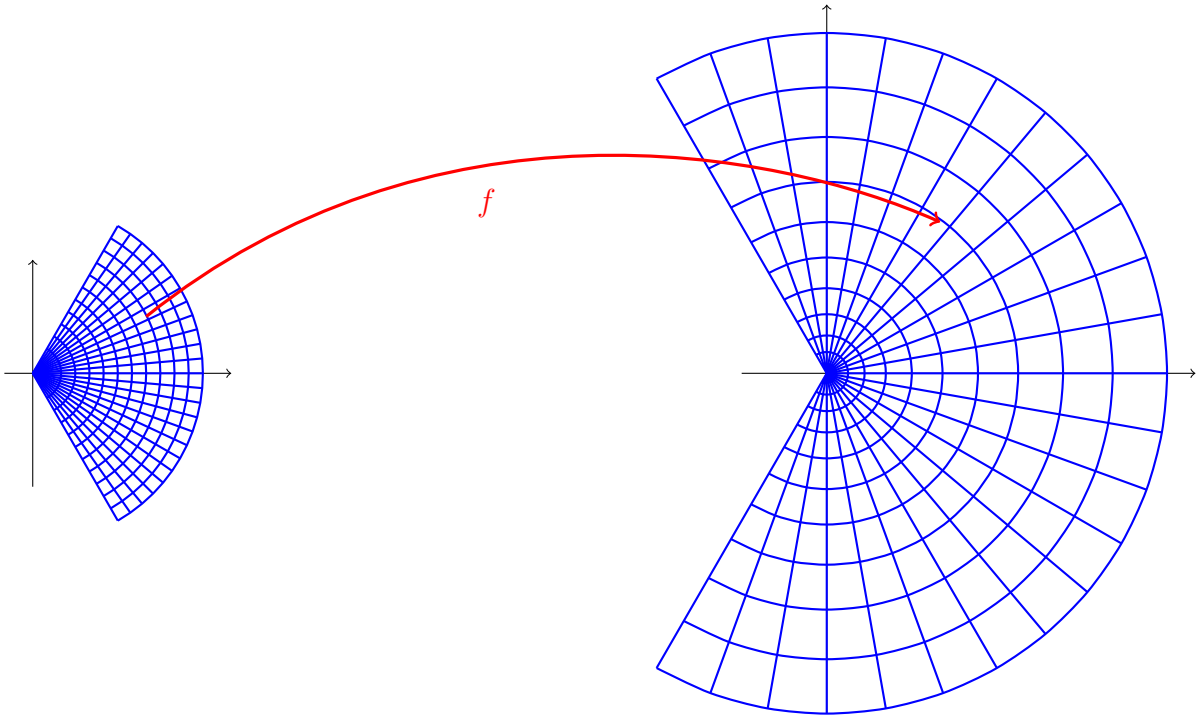
⚒ **7.4.6 Beispiel.** Die Abbildung $\mathbb{C} \ni z \mapsto z^2 \in \mathbb{C}$ ist für $z \neq 0$ (lokal) umkehrbar. Um dies auf obigen Satz zu reduzieren, schreiben wir $z = x + \mathrm{i}y$ mit $z^2 = x^2 - y^2 + 2\mathrm{i}xy$ und betrachten die Abbildung als Funktion

$$f : \mathbb{R}^2 \ni \begin{bmatrix} x \\ y \end{bmatrix} \mapsto \begin{bmatrix} x^2 - y^2 \\ 2xy \end{bmatrix} \in \mathbb{R}^2 \tag{7.4.6-A}$$

mit Jacobimatrix

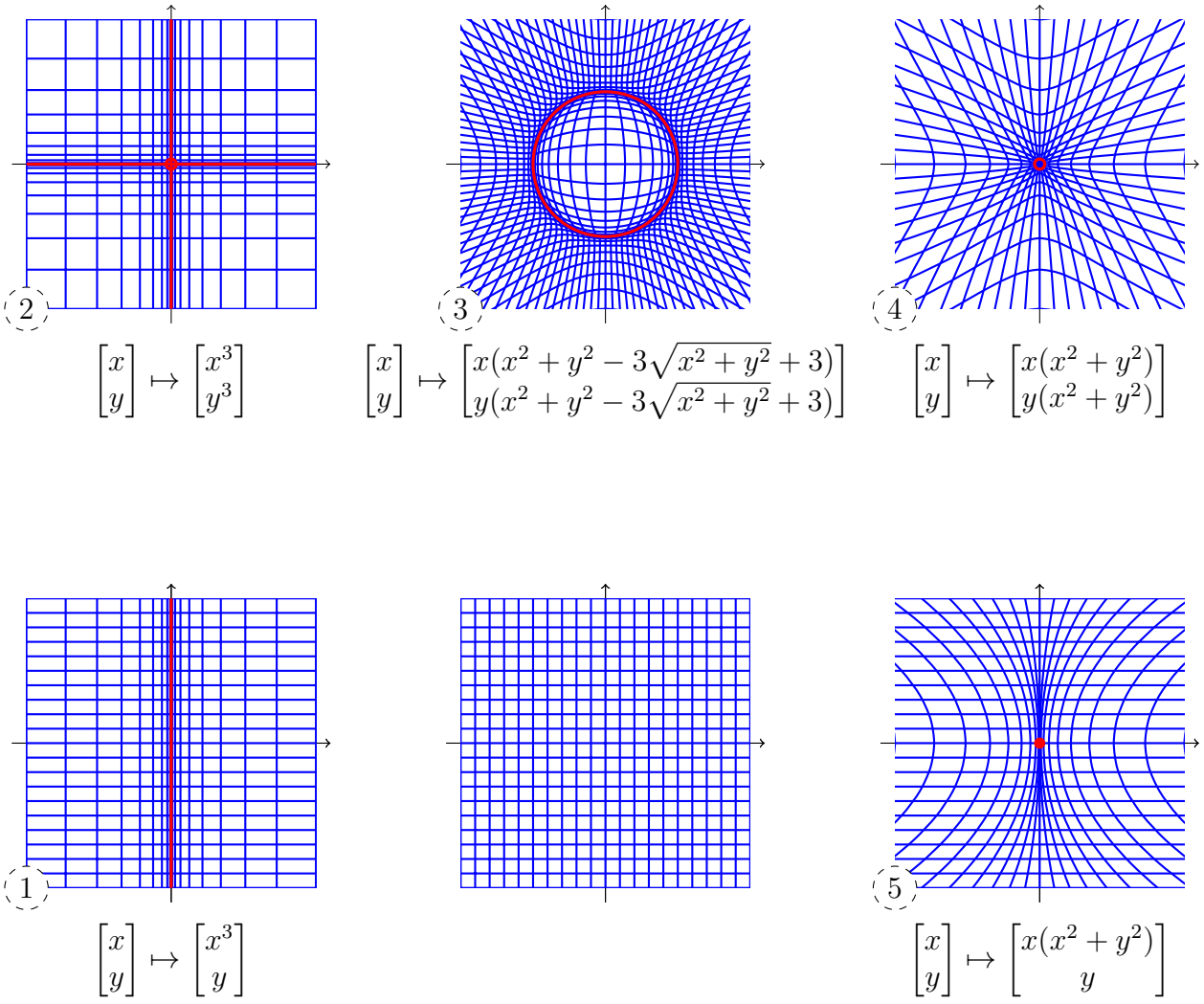
$$f'(x,y) = \begin{bmatrix} 2x & -2y \\ 2y & 2x \end{bmatrix} \qquad \text{mit} \qquad \det f'(x,y) = 4x^2 + 4y^2. \tag{7.4.6-B}$$

Für $z \neq 0$ und damit $x^2 + y^2 > 0$ gilt $\det f'(x,y) \neq 0$.



Um $z = 0$ ist die Funktion nicht lokal bijektiv: es gibt keine Umgebung des Ursprungs auf der $z \mapsto z^2$ injektiv ist, da stets $z^2 = (-z)^2$ gilt und jede Umgebung eine kleinere punktsymmetrische Umgebung enthält.

Es ist schwer, aus der lokalen Aussage des Umkehrsatzes eine globale Umkehrbarkeit der Funktion auf einer größeren Menge zu folgern. Dies ist anders als im Eindimensionalen, wo aus dem Nichtverschwinden der Ableitung auf einem Intervall die Monotonie und damit die Injektivität der Funktion folgt.



Bilder des in der Mitte dargestellten Gitters unter den angegebenen bijektiven differenzierbaren Abbildungen $f : \mathbb{R}^2 \to \mathbb{R}^2$. Rot markiert sind dabei im Bild jeweils die Mengen, auf denen der Rang von f' nicht maximal ist, also die Stellen in denen der lokale Umkehrsatz nicht anwendbar ist. Auf den roten Linien in Bildern 1, 2 und 3 ist der Rang von f' gleich 1, auf dem Schnittpunkt der Linien in Bild 2 und im Mittelpunkt von Bild 4 ist der Rang von f' Null. Im Mittelpunkt von Bild 5 ist der Rang von f' gleich 1.

⚒ **7.4.7 Beispiel.** Die Abbildung $f : \mathbb{R}^3 \to \mathbb{R}^3$ mit

$$f : \begin{bmatrix} x \\ y \\ z \end{bmatrix} \mapsto \begin{bmatrix} xy \\ yz \\ zx \end{bmatrix} \tag{7.4.7-A}$$

erfüllt

$$f'(x,y,z) = \begin{bmatrix} y & x & 0 \\ 0 & z & y \\ z & 0 & x \end{bmatrix}, \qquad \det f'(x,y,z) = 2xyz. \tag{7.4.7-B}$$

Damit besitzt insbesondere der Punkt $(1,1,1)^\top$ eine Umgebung, auf der die Funktion f invertierbar ist. Es gibt also $\delta > 0$ und $\varepsilon > 0$, so dass für jedes $(\xi,\eta,\zeta)^\top \in \mathbb{R}^3$ mit

$$|\xi - 1|^2 + |\eta - 1|^2 + |\zeta - 1|^2 < \delta^2 \tag{7.4.7-C}$$

das Gleichungssystem

$$xy = \xi, \quad yz = \eta, \quad xz = \zeta \tag{7.4.7-D}$$

eindeutig nach x, y und z mit

$$|x - 1|^2 + |y - 1|^2 + |z - 1|^2 < \varepsilon^2 \tag{7.4.7-E}$$

auflösbar ist. Der lokale Umkehrsatz liefert also ein Kriterium für die lokale Auflösbarkeit *expliziter* Gleichungen.

7.4.8. Eine weitere wichtige Anwendung der Differentialrechnung in mehreren Veränderlichen betrifft die (lokale) Auflösbarkeit von *impliziter* Gleichungen

$$f(x,y) = 0 \tag{7.4.8-A}$$

nach einem Teil der Variablen, also dem Finden einer Funktion φ mit $y = \varphi(x)$ und damit $f(x,\varphi(x)) = 0$.

Wir präzisieren die Notation dazu. Für normierte Räume V und W ist $V \times W$ wiederum ein Vektorraum und versehen mit der Norm $\|(v,w)\|^2 = \|v\|^2 + \|w\|^2$ ist dieser auch normiert. Weiter bezeichnen

$$\Pi_V : V \times W \ni (v,w) \mapsto v \in V \quad \text{und} \quad \Pi_W : V \times W \ni (v,w) \mapsto w \in W \tag{7.4.8-B}$$

die kanonischen Projektionen auf die jeweiligen Komponenten und

$$\iota_V : V \ni v \mapsto (v,0) \in V \times W \quad \text{und} \quad \iota_W : W \ni w \mapsto (0,w) \in V \times W \tag{7.4.8-C}$$

die kanonischen Einbettungen und für eine differenzierbare Funktion f auf $V \times W$ definieren wir durch

$$\partial_v f(v,w) = f'(v,w) \circ \iota_V, \qquad \partial_w f(v,w) = f'(v,w) \circ \iota_W \tag{7.4.8-D}$$

die partiellen Ableitungen von f nach den Variablen aus V beziehungsweise W. Diese sind gerade die Ableitungen, wenn man die jeweils andere (Vektor-) Variable als konstant

betrachtet (und damit nur Vektoren aus $V\times\{0\}$ beziehungsweise $\{0\}\times W$ in die Fréchet-Ableitung einsetzt).

Gibt es eine differenzierbare auflösende Funktion $y=\varphi(x)$ zu (7.4.**8-A**), so folgt durch Ableiten der für alle x in einer offenen Menge geltenden Identität $0=g(x):=f(x,\varphi(x))$ mit der Kettenregel

$$0=g'(x)=\partial_x f(x,\varphi(x))+\partial_y f(x,\varphi(x))\varphi'(x). \tag{7.4.8-E}$$

Damit $\varphi'(x)$ eindeutig durch diese Gleichung bestimmt wird, ist es naheliegend die Invertierbarkeit von $\partial_y f(x,y)\in\mathcal{L}(W,W)$ und damit auch die Abbildungseigenschaft $f:V\times W\to W$ zu fordern. Dass diese Forderung ausreicht, besagt der folgende Satz.

☞ **7.4.9 Satz** (Satz über implizite Funktionen)**.** *Seien V und W vollständige normierte Räume, $\Omega\subseteq V\times W$ offen und $f:\Omega\to W$ stetig differenzierbar. Sei weiter für ein $(x_*,y_*)\in\Omega$ mit $f(x_*,y_*)=w$ die partielle Ableitung*

$$\partial_y f(x_*,y_*)\in\mathcal{L}(W,W) \tag{7.4.9-A}$$

invertierbar. Dann gibt es eine Umgebung U von $x_*\in V$ und eine eindeutig bestimmte differenzierbare Funktion $\varphi:U\to W$, so dass

$$f(x,\varphi(x))=w \tag{7.4.9-B}$$

für alle $x\in U$ gilt.

Beweis. Wir führen die Aussage direkt auf den lokalen Umkehrsatz 7.4.4 zurück und nutzen dafür die differenzierbare Hilfsfunktion $F:\Omega\to V\times W$

$$F(x,y)=\begin{bmatrix}x\\ f(x,y)\end{bmatrix} \tag{7.4.9-C}$$

mit Ableitung

$$F'(x,y)=\begin{bmatrix}I & 0\\ \partial_x f(x,y) & \partial_y f(x,y)\end{bmatrix}\in\mathcal{L}(V\times W,V\times W). \tag{7.4.9-D}$$

Dann ist $F'(x_*,y_*)$ invertierbar und es gibt nach Satz 7.4.4 eine Umgebung $\widetilde{U}$ von (x_*,y_*) und eine differenzierbare Umkehrfunktion $F^{-1}:F[\widetilde{U}]\to\widetilde{U}$. Wir setzen[7]

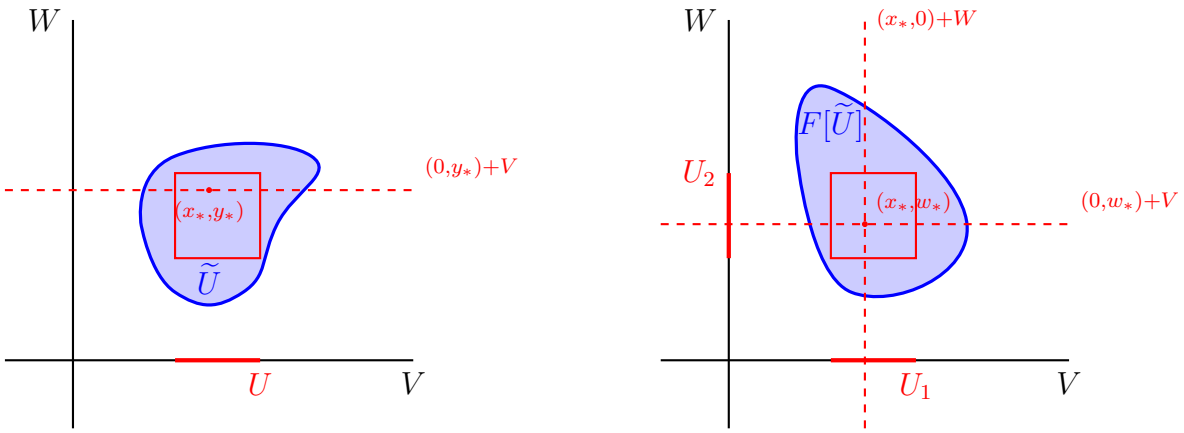
$$U=\Pi_V[\widetilde{U}\cap\big((0,y_*)+\iota_V[V]\big)] \tag{7.4.9-E}$$

und

$$\begin{bmatrix}x\\ \varphi(x)\end{bmatrix}=F^{-1}(x,w),\qquad x\in U. \tag{7.4.9-F}$$

Dann ist nach Konstruktion $\varphi:U\to W$ differenzierbar und erfüllt $f(x,\varphi(x))=w$. □

[7] Wir wählen also U in V so, dass $(x,y_*)\in\widetilde{U}$ für $x\in U$ gilt.



Es gilt leicht mehr. Wählt man Umgebung $U_1 \subseteq \Pi_V[F[\widetilde{U}] \cap \big((0,w_*)+\iota_V[V]\big)]$ und $U_2 \subseteq \Pi_W[F[\widetilde{U}] \cap \big((x_*,0)+\iota_W[W]\big)]$ von (x_*,w_*), $w_* = f(x_*,y_*)$, so klein, dass $U_1 \times U_2 \subseteq F[\widetilde{U}]$ gilt und setzt in Analogie zu (7.4.**9-F**)

$$\begin{bmatrix} x \\ \Phi(x,w) \end{bmatrix} = F^{-1}(x,w), \qquad x \in U_1, \quad w \in U_2, \tag{7.4.9-G}$$

so erhält man eine differenzierbare auflösende Funktion $\Phi : U_1 \times U_2 \to W$ mit $f(x,\Phi(x,w)) = w$. Die Auflösung hängt also differenzierbar von der rechten Seite ab.

Für Abbildungen $f : \mathbb{R}^n \to \mathbb{R}^m$ mit $m \leq n$ kann man obigen Satz wiederum in der Sprache der Jacobimatrizen formulieren und fragen, wann Gleichungen $f(x) = \text{const}$ nach einem Teil der Variablen differenzierbar auflösbar sind. Wir formulieren die Aussage wiederum als Korollar.

☞ **7.4.10 Korollar.** *Sei $f : \mathbb{R}^n \supseteq \Omega \to \mathbb{R}^m$ mit $m \leq n$ stetig differenzierbar und gelte für ein $x_* \in \Omega$*

$$\operatorname{rank} f'(x_*) = m. \tag{7.4.10-A}$$

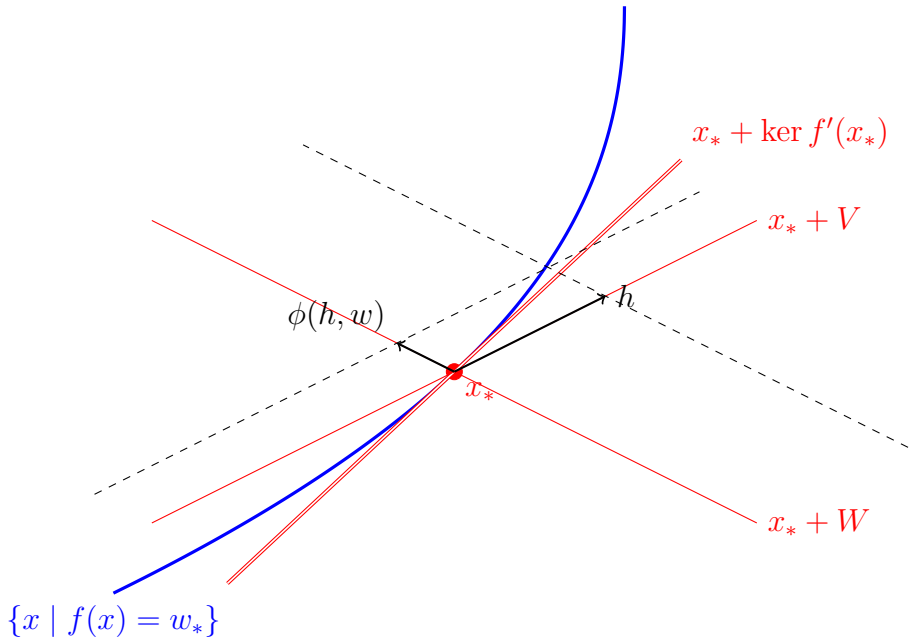
Zerlegt man nun $\mathbb{R}^n = V \dot{+} W$ mit $\dim V = n-m$ und $\dim W = m$ derart, dass

$$\ker f'(x_*) \cap W = \{0\} \tag{7.4.10-B}$$

gilt, dann existiert eine Nullumgebung U_1 in V, eine Umgebung U_2 von $f(x_*)$ und eine differenzierbare Abbildung $\varphi : U_1 \times U_2 \to W$, so dass

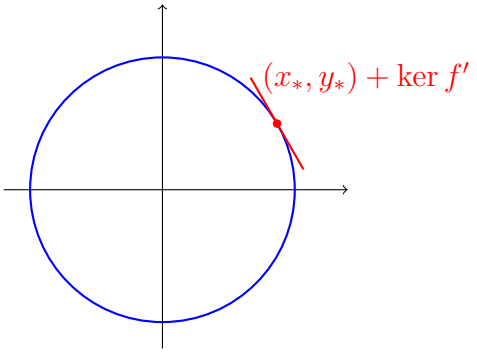
$$f(x_* + h + \varphi(h,w)) = w \tag{7.4.10-C}$$

für alle $h \in U_1$ und alle $w \in U_2$ erfüllt ist.



Insbesondere kann man immer $V = \ker f'(x_*)$ und W einen geeigneten komplementären Unterraum wählen. Eine oft sinnvolle Wahl ist $W = V^\perp = \{y \in \mathbb{R}^n \mid \forall_{x \in V}\ x^\top y = 0\}$.

⚒ **7.4.11 Beispiel.** Wir betrachten die Kreisgleichung $f(x, y) = x^2 + y^2 - 1 = 0$. Diese ist in der Nähe eines jeden Punktes (x_*, y_*) mit $f(x_*, y_*) = 0$ und $\partial_y f(x_*, y_*) = 2y_* \neq 0$ nach y differenzierbar auflösbar. Das sind alle Kreispunkte verschieden von $(\pm 1, 0)^\top$.



Die auflösende Funktion kann man hier direkt angeben, es gilt

$$y = \varphi(x) = \pm\sqrt{1 - x^2} \qquad (7.4.\textbf{11-A})$$

mit dem passend gewählten Vorzeichen, je nachdem ob $y_* > 0$ oder $y_* < 0$ gilt.

⚒ **7.4.12 Beispiel.** Die Lambertsche W-Funktion[8] ist definiert durch die Beziehung

$$f(W(x), x) = W(x)\mathrm{e}^{W(x)} - x = 0 \qquad (7.4.\textbf{12-A})$$

zusammen mit der Bedingung $W(0) = 0$. Die Gleichung ist in der Nähe von 0 differenzierbar auflösbar, da $\partial_W f(W, x) = \mathrm{e}^W(1 + W) \neq 0$ für alle $W > -1$. Darüberhinaus gilt $W = -1$ für $x = -1/\mathrm{e}$ und für die sich ergebende auflösende Funktion W erhalten wir aus der obigen Gleichung

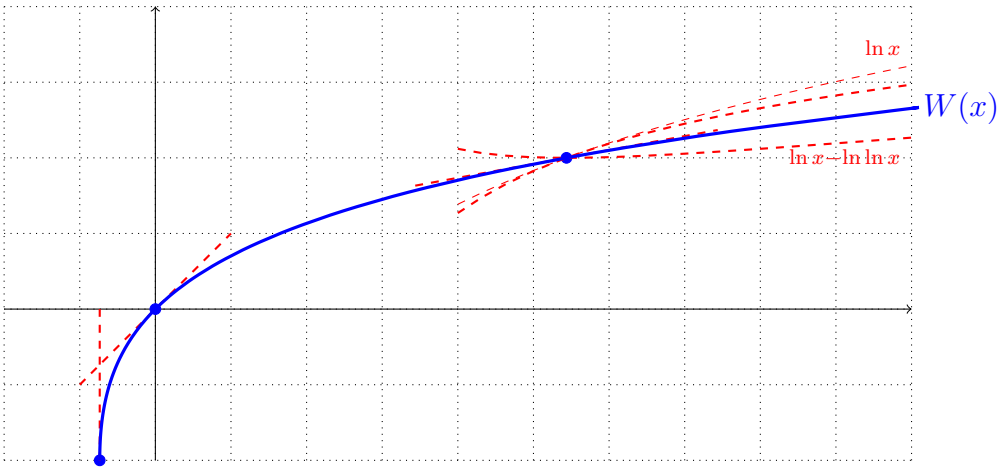
$$W'(x) = \frac{1}{\mathrm{e}^{W(x)}(1 + W(x))} = \frac{1}{x + \mathrm{e}^{W(x)}} \geq 0 \qquad (7.4.\textbf{12-B})$$

[8] Johann Heinrich Lambert, 1728–1777

für alle $W > -1$ (und damit für $x > -\frac{1}{\mathrm{e}}$). Für $W(x) \to \infty$ folgt damit $W'(x) \longrightarrow 0$ und die Lambertsche W-Funktion kann nicht für endliche x gegen Unendlich streben. Damit ist W auf $[-\frac{1}{\mathrm{e}}, +\infty)$ definiert, streng monoton wachsend und auf $(-\frac{1}{\mathrm{e}}, +\infty)$ differenzierbar.

Die Funktion selbst kann nicht explizit angegeben und durch elementare Funktionen ausgedrückt werden. Es existieren Darstellungen als Potenzreihen, die einfachste Möglichkeit sich die Funktion graphisch vorzustellen ist eine Kurvendiskussion mit ausführlicher Diskussion des asymptotischen Verhaltens.

Einige spezielle Werte ergeben sich direkt aus der definierenden Bedingung, so gilt $W(-\frac{1}{\mathrm{e}}) = -1$, $W(0) = 0$ und $W(\mathrm{e}) = 1$. Ebenso gilt $W'(0) = 1$ und $W'(\mathrm{e}) = \frac{1}{2\mathrm{e}}$.



Da $W(t)$ monoton wachsend ist, folgt für $x \geq \mathrm{e}$ dass $\mathrm{e}^{W(x)} = \frac{x}{W(x)} \leq x$ und damit $W(x) \leq \ln x$ gilt. Damit folgt aber weiterhin $\mathrm{e}^{W(x)} = \frac{x}{W(x)} \geq \frac{x}{\ln x}$ und damit

$$\ln x - \ln\ln x \leq W(x) \leq \ln x, \qquad x \geq \mathrm{e}. \tag{7.4.12-C}$$

Dies kann man fortführen und iterativ bessere Ungleichungen erhalten. Interessant ist nur der nächste Schritt: Auf Grund der (elementaren!) Abschätzung

$$-\ln(1-t) \leq 2t\ln 2, \qquad t \in (0, \frac{1}{2}) \tag{7.4.12-D}$$

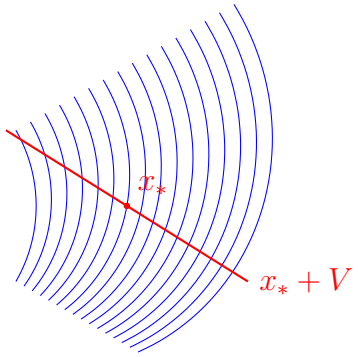
gilt

$$\begin{aligned} W(x) &\leq \ln x - \ln(\ln x - \ln\ln x) = \ln x - \ln\ln x - \ln\left(1 - \frac{\ln\ln x}{\ln x}\right) \\ &\leq \ln x - \ln\ln x + 2\ln 2\frac{\ln\ln x}{\ln x} \qquad \text{für } x \text{ mit } \ln x > 2\ln\ln x > 0, \end{aligned} \tag{7.4.12-E}$$

und da $\frac{\ln\ln x}{\ln x} \longrightarrow 0$ für $x \to \infty$ gilt, folgt

$$W(x) = \ln x - \ln\ln x + \mathbf{o}(1) \qquad \text{für } x \to \infty. \tag{7.4.12-F}$$

7.4.13. Eine differenzierbare Funktion $f : \mathbb{R}^n \supseteq \Omega \to \mathbb{R}^m$ mit $\operatorname{rank} f'(x) = m$ auf Ω wird als *Submersion* bezeichnet. Nach dem Satz über implizite Funktionen wird damit Ω durch die differenzierbar parametrisierbaren Mengen $N_y = \{x \mid f(x) = y\}$ für $y \in \mathbb{R}^m$ zerlegt. Lokal um einen Punkt $x_* \in \Omega$ haben Submersionen also folgende Struktur



mit $V \subseteq \mathbb{R}^n$ einem Unterraum mit $V \dot{+} \ker f'(x_*) = \mathbb{R}^n$. Entlang der gezeichneten Mengen N_y ist die Funktion f konstant, die Werte sind durch die Schnitte mit der Hyperebene $x_* + V$ parametrisiert.

Ein einfaches Beispiel einer Submersion ist offenbar die Projektion

$$\Pi : \begin{bmatrix} x_1 \\ \vdots \\ x_m \\ x_{m+1} \\ \vdots \\ x_n \end{bmatrix} \mapsto \begin{bmatrix} x_1 \\ \vdots \\ x_m \end{bmatrix} \tag{7.4.13-A}$$

auf die ersten m Koordinaten. Zumindest lokal sieht jede Submersion $f : \Omega \to \mathbb{R}^n$ so aus, für jedes $x_* \in \Omega$ existiert ein Diffeomorphismus ψ mit

$$\psi \circ f = \Pi \circ \psi \tag{7.4.13-B}$$

auf einer Umgebung von x_*. Der Diffeomorphismus ψ biegt dabei die Mengen N_y gerade und Π projiziert auf die Hyperebene $x_* + V$. Dies gilt allgemeiner für differenzierbare Abbildung mit Ableitungen von konstantem Rang:

☞ **7.4.14 Satz** (Rangsatz). *Sei $\Omega \subseteq \mathbb{R}^n$ offen und $f : \Omega \to \mathbb{R}^m$ stetig differenzierbar gelte $\operatorname{rank} f'(x) = k$ auf Ω. Dann gibt es zu jedem $x_* \in \Omega$ Umgebungen U_1 von x_* und U_2 von $f(x_*)$ und Diffeomorphismen*

$$\psi : U_1 \to \psi[U_1] \subseteq \mathbb{R}^n \qquad \phi : U_2 \to \phi[U_2] \subseteq \mathbb{R}^m, \tag{7.4.14-A}$$

so dass

$$\phi \circ f \circ \psi^{-1}(x_1, \ldots, x_n) = (x_1, \ldots, x_k, 0, \ldots, 0) \tag{7.4.14-B}$$

die Projektion auf die ersten k Koordinaten ist.

Beweis. Da $\operatorname{rank} f'(x_*) = k$ gilt, gibt es einen k-dimensionalen Unterraum $W \subset \mathbb{R}^m$ und eine Projektion $\Pi_1 : \mathbb{R}^m \to W$ mit $\operatorname{rank} \Pi_1 f'(x_*) = k$ (man nehme zum Beispiel den durch k linear unabhängige Spalten der Jacobimatrix aufgespannten Unterraum W und eine Projektion auf diesen). Dann gilt für die Hilfsfunktion

$$g := \Pi_1 \circ f : \Omega \to W \tag{7.4.**14-C**}$$

nach Konstruktion $g'(x_*) = \Pi_1 f'(x_*)$ und somit $\operatorname{rank} g'(x_*) = k = \dim W$. Also sind die Voraussetzungen des Satzes über die implizite Funktion erfüllt, es bleibt nur die Wahl der passenden Variablen. Dazu wählen wir den Unterraum $V = \ker g'(x_*) \subseteq \mathbb{R}^n$ und finden nach Korollar 7.4.10 eine Nullumgebung $U_1 \subseteq V$ und eine Umgebung $U_2 \subseteq W$ von $g(x_*)$, sowie eine differenzierbare auflösende Funktion $\Phi : U_1 \times U_2 \to \mathbb{R}^n$ mit

$$w = g(x_* + h + \Phi(h, w)) = \Pi_1 \circ f(x_* + h + \Phi(h, w)) \tag{7.4.**14-D**}$$

für alle $h \in U_1$ und $w \in U_2$.

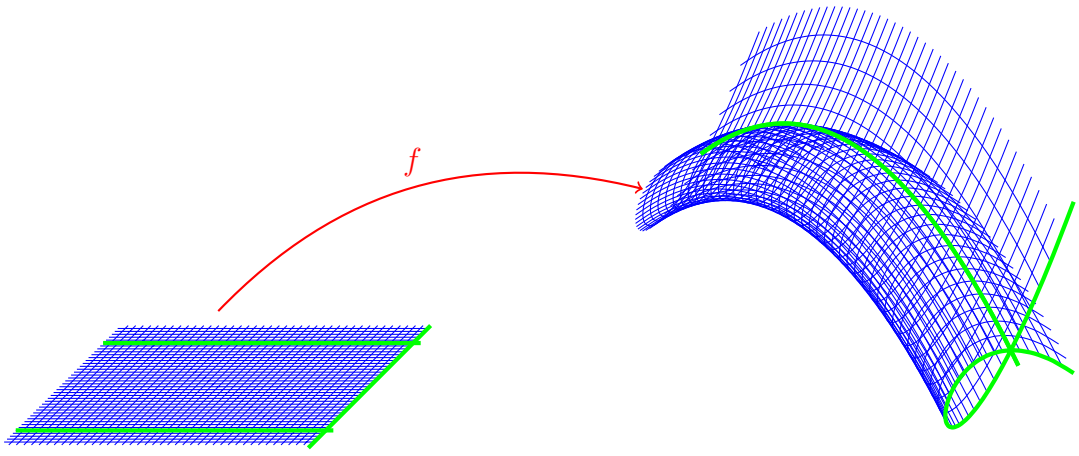
Es verbleibt, die gewünschten Diffeomorphismen abzulesen. Wir setzen $\phi = \Pi_1$ und $\psi(x_* + h + \Phi(h, w)) = (w, h) \in W \times V$ und wegen $\dim(W \times V) = \dim W + \dim V = \dim \operatorname{im} g'(x_*) + \dim \ker g'(x_*) = n$ kann $W \times V$ durch Fortsetzen einer Basis von W mit $\mathbb{R}^n$ identifiziert werden, so dass $(w, h) \mapsto w$ die gewünschte Projektion auf die ersten k Koordinaten ist. □

7.4.15. Eine stetig differenzierbare Funktion $f : \mathbb{R}^n \supseteq \Omega \to \mathbb{R}^m$ mit

$$\operatorname{rank} f'(x) = n \quad \text{auf } \Omega \tag{7.4.**15-A**}$$

wird als *Immersion* bezeichnet. Nach dem gerade gezeigten Rangsatz sehen Immersionen lokal wie (lineare) Einbettungen $\mathbb{R}^n \to \mathbb{R}^m$ für $m \geq n$ verkettet mit einem Diffeomorphismus aus.

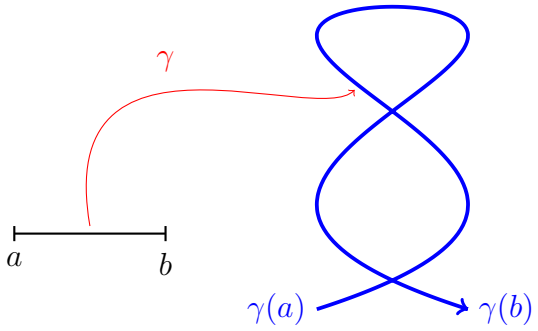
Für $n = 2$ und $m = 3$ beschreiben diese Flächen im Raum (die sich wie in folgendem Bild auch selbst durchdringen dürfen):



Als nächstes wenden wir uns dem Fall $n = 1$ zu. Hier beschreiben Immersionen Kurven. Diese verdienen zuerst eine genauere Betrachtung.

7.5 Kurven im Raum

7.5.1. Unter einer *Parametrisierung* einer *Kurve* im Raum wollen wir eine stetige Abbildung $\gamma : (a, b) \to \mathbb{R}^m$ verstehen. Ist die Abbildung stetig differenzierbar, so nennen wir die Parametrisierung *differenzierbar*.



Gilt darüberhinaus $\dot\gamma(t) = \frac{\mathrm{d}}{\mathrm{d}t}\gamma(t) \neq 0$ für alle $t \in (a, b)$, ist γ also eine Immersion, so heißt die Parametrisierung und die Kurve *regulär*.

Wir haben nicht vorausgesetzt, dass die Parametrisierung γ injektiv sein soll. Das würde wichtige Beispiele ausschließen. So parametrisiert die Abbildung

$$\mathbb{R} \ni t \mapsto \begin{bmatrix} \cos t \\ \sin t \end{bmatrix} \in \mathbb{R}^2 \tag{7.5.1-A}$$

einen Kreis. Um jedoch pathologische Beispiele auszuschliessen, ist es sinnvoll anzunehmen, dass γ *lokal injektiv* ist, es also zu jedem $t \in (a, b)$ eine Umgebung $U_\varepsilon(t)$ mit

$$\gamma\big|_{U_\varepsilon(t)} \quad \text{injektiv} \tag{7.5.1-B}$$

gibt. Dies gilt sicher für alle regulären Parametrisierungen.

Damit können wir nun auch den bisher offen gelassenen Begriff einer *Kurve* definieren. Eine Kurve Γ im $\mathbb{R}^m$ ist für uns eine Teilmenge des $\mathbb{R}^m$, die das Bild einer lokal injektiven stetigen Parametrisierung $\gamma : (a, b) \to \mathbb{R}^m$ ist,

$$\Gamma = \{\gamma(t) \mid t \in (a, b)\}. \tag{7.5.1-C}$$

Punkte die bei der Parametrisierung, mehrfach erreicht werden, heißen *Doppelpunkte*. Wir werden im Folgenden annehmen, dass von uns betrachtete Kurven nur endlich viele Doppelpunkte und auch nur solche mit endlichen Vielfachheiten besitzen.

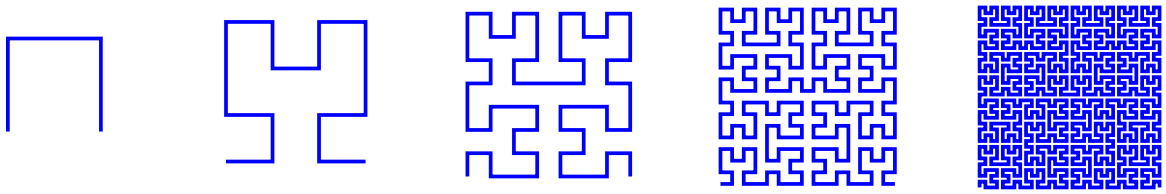
Mitunter lassen sich Parametrisierungen stetig von (a, b) auf $[a, b]$ fortsetzen. Solche Fortsetzungen fügen zu Kurven Anfangs- und Endpunkte hinzu und wir können

$$\gamma : [a, b] \to \mathbb{R}^m \tag{7.5.1-D}$$

als Parametrisierung der Kurve Γ aus (7.5.1-C) mit Endpunkten $\gamma(a)$ und $\gamma(b)$ verstehen. Wir werden im Weiteren nicht zwischen Kurven ohne Endpunkten und Kurven mit Endpunkten unterscheiden.

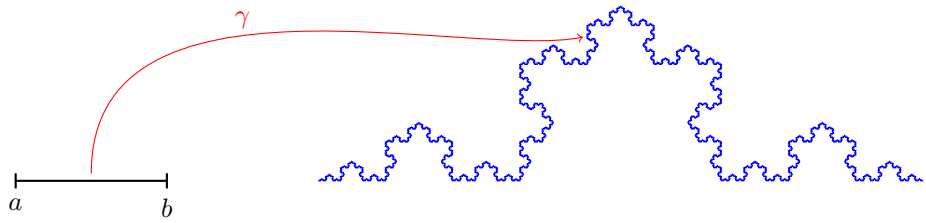
Abgesehen von den Doppelpunkten bestimmt eine Parametrisierung eine *Orientierung* der Kurve, indem wir die Ordnung des Intervalls (a, b) auf die Kurvenpunkte übertragen. In Bildern werden wir deshalb orientierte Kurven oft mit Pfeilen versehen um die Orientierung hervorzuheben. Für zwei Kurvenpunkte $x_1 = \gamma(t_1)$ und $x_2 = \gamma(t_2)$ einer orientierten Kurve Γ, die keine Doppelpunkte sind, schreiben wir $x_1 \prec x_2$ falls x_1 vor x_2 kommt, also $t_1 < t_2$ gilt.

✂ **Ergänzung.** Oft werden auch stetige Funktionen $\gamma : [a, b] \to \mathbb{R}^m$ direkt als Kurven oder Wege bezeichnet. Ohne die hier gemachten Zusatzvoraussetzungen sind diese nicht durch ein *Bild* charakterisierbar, so gibt es stetige und surjektive Abbildungen $\gamma : [0, 1] \to [0, 1] \times [0, 1]$ beziehungsweise $\mathbb{R} \to \mathbb{R}^2$. Ein Beispiel dazu ist der Grenzwert der dargestellten Hilbert-Kurven[9]



Ist die Parametrisierung $\gamma : [a, b] \to \mathbb{R}^m$ injektiv, so spricht man von einer *Jordankurve*[10]. Analog spricht man von einer geschlossenen ebenen Jordankurve, falls die stetige Parametrisierung $\gamma : [0, 1] \to \mathbb{R}^2$ auf $[0, 1)$ injektiv ist und $\gamma(0) = \gamma(1)$ gilt. Für solche Kurven gilt der Jordansche Kurvensatz, sie zerlegen die Ebene in ein inneres und ein äußeres Teilgebiet. Der Beweis dieses Satzes ist (erstaunlicherweise) nichttrivial.

Jordankurven können kompliziert aussehen. Ein bekanntes Beispiel einer ebenen (und selbstähnlichen) Jordankurve ist die Koch-Kurve[11]:



7.5.2. Wir wollen als Nächstes versuchen, orientierten Kurven eine Länge zuzuordnen. Sei dazu Γ eine doppelpunktfreie Kurve im $\mathbb{R}^m$. Eine Zerlegung $\mathfrak{Z} = (x_0, \dots, x_n)$ der Kurve Γ ist eine geordnete Auswahl von Punkten $x_0 \prec x_1 \prec x_2 \prec \cdots \prec x_n$. Jeder solchen Zerlegung können wir eine Länge zuordnen,

$$\ell(\mathfrak{Z}) = \sum_{k=1}^{n} |x_k - x_{k-1}|. \tag{7.5.2-A}$$

Analog zur Konstruktion des Riemannintegrals definieren wir Verfeinerungen $\mathfrak{Z}_2$ einer Zerlegung $\mathfrak{Z}_1$ als Zerlegungen mit mehr Punkten und schreiben dafür $\mathfrak{Z}_1 \subset \mathfrak{Z}_2$. Ebenso

[9] David Hilbert, 1862–1943
[10] Camille Jordan, 1838–1922
[11] Helge von Koch, 1870–1924

bezeichnen wir mit $\mathfrak{Z}_1 \cup \mathfrak{Z}_2$ die gemeinsame Verfeinerung von $\mathfrak{Z}_1$ und $\mathfrak{Z}_2$. Als Konsequenz der Dreiecksungleichung gilt nun

$$\mathfrak{Z}_1 \subset \mathfrak{Z}_2 \qquad \Longrightarrow \qquad \ell(\mathfrak{Z}_1) \le \ell(\mathfrak{Z}_2). \tag{7.5.2-B}$$

Wir sagen die Kurve Γ ist *rektifizierbar*, falls

$$\ell(\Gamma) := \sup_{\mathfrak{Z}} \ell(\mathfrak{Z}) < \infty \tag{7.5.2-C}$$

gilt. In diesem Fall heißt $\ell(\Gamma)$ die *Länge* der Kurve Γ.

Die Voraussetzung der Doppelpunktfreiheit ist unwesentlich, für Kurven mit Doppelpunkten fügen wir diese mehrfach in die orientierte Kurve / in die Zerlegungen ein und gehen analog vor.

7.5.3. Für doppelpunktfreie rektifizierbare Kurven liefert die Länge eine natürliche Parametrisierung. Sei dazu Γ rektifizierbar und $x_0 \in \Gamma$ ein fest gewählter Kurvenpunkt. Dann kann jedem x mit $x_0 \prec x$ durch

$$x \mapsto \ell(\Gamma_{[x_0,x]}), \qquad \Gamma_{[x_0,x]} = \{y \in \Gamma : x_0 \prec y \prec x\} \tag{7.5.3-A}$$

eine Zahl zugeordnet werden. Dabei gilt für $x_0 \prec y \prec x$ gilt

$$\ell(\Gamma_{[x_0,x]}) - \ell(\Gamma_{[x_0,y]}) = \ell(\Gamma_{[y,x]}) \ge |x - y| > 0 \tag{7.5.3-B}$$

und die Zuordnung ist monoton und damit insbesondere injektiv. Um die Zuordnung auch auf $x \prec x_0$ fortzusetzen, nutzen wir entsprechend

$$x \mapsto -\ell(\Gamma_{[x,x_0]}). \tag{7.5.3-C}$$

Als monotone Abbildung besitzt diese eine Umkehrfunktion. Diese ist wegen (7.5.3-B) stetig und liefert damit eine Parametrisierung

$$[\ell_-, \ell_+] \ni t \mapsto x \in \Gamma \qquad \Longleftrightarrow \qquad \begin{cases} t = \ell(\Gamma_{[x_0,x]}), & x_0 \prec x, \\ t = -\ell(\Gamma_{[x,x_0]}), & x \prec x_0, \end{cases} \tag{7.5.3-D}$$

die als *Parametrisierung nach Bogenlänge* bezeichnet wird.

☞ **7.5.4 Satz.** *Sei $\gamma : [a,b] \to \mathbb{R}^m$ stetig differenzierbare lokal injektive Parametrisierung einer Kurve Γ. Dann ist Γ rektifizierbar und es gilt*

$$\ell(\Gamma) = \int_a^b |\dot\gamma(t)| \, \mathrm{d}t. \tag{7.5.4-A}$$

Beweis. Sei $\mathfrak{Z} = (x_0, \dots, x_n)$ eine Zerlegung der Kurve Γ mit $x_0 = \gamma(a)$ und $x_n = \gamma(b)$. Diese bestimmt eine Zerlegung $\widehat{\mathfrak{Z}} = (t_0, t_1, \dots, t_n)$ von $[a,b]$ mit $x_k = \gamma(t_k)$. Da dann nach dem Hauptsatz der Differentialrechnung

$$x_k - x_{k-1} = \gamma(t_k) - \gamma(t_{k-1}) = \int_{t_{k-1}}^{t_k} \dot\gamma(t) \, \mathrm{d}t \tag{7.5.4-B}$$

(mit komponentenweise verstandenem Integral) gilt, folgt

$$|x_k - x_{k-1}| \le \int_{t_{k-1}}^{t_k} |\dot\gamma(t)|\,\mathrm{d}t \tag{7.5.4-C}$$

und damit nach Summation über k

$$\ell(\mathfrak{Z}) = \sum_{k=1}^{n} |x_k - x_{k-1}| \le \sum_{k=1}^{n} \int_{t_{k-1}}^{t_k} |\dot\gamma(t)|\,\mathrm{d}t = \int_a^b |\dot\gamma(t)|\,\mathrm{d}t. \tag{7.5.4-D}$$

Also ist $\ell(\mathfrak{Z})$ nach oben beschränkt und Γ rektifizierbar.

Da $[a,b]$ kompakt ist, ist $\dot\gamma : [a,b] \to \mathbb{R}^m$ gleichmäßig stetig. Es gibt also zu jedem $\varepsilon > 0$ ein $\delta > 0$ mit

$$|s-t| < \delta \qquad \Longrightarrow \qquad |\dot\gamma(s) - \dot\gamma(t)| < \varepsilon. \tag{7.5.4-E}$$

Wählt man nun $\mathfrak{Z}$ so fein, dass die Zerlegung des Intervalls $\delta(\widehat{\mathfrak{Z}}) < \delta$ erfüllt, dann gilt für jedes $\tau \in [t_{k-1}, t_k]$

$$\begin{aligned} \Big| \int_{t_{k-1}}^{t_k} \dot\gamma(t)\,\mathrm{d}t - \dot\gamma(\tau)(t_k - t_{k-1}) \Big| &\le \varepsilon(t_k - t_{k-1}), \\ \left| \int_{t_{k-1}}^{t_k} |\dot\gamma(t)|\,\mathrm{d}t - |\dot\gamma(\tau)|(t_k - t_{k-1}) \right| &\le \varepsilon(t_k - t_{k-1}), \end{aligned} \tag{7.5.4-F}$$

letzteres mit der umgekehrten Dreiecksungleichung. Also folgt

$$\begin{aligned} 0 &\le \int_{t_{k-1}}^{t_k} |\dot\gamma(t)|\,\mathrm{d}t - |x_k - x_{k-1}| \\ &= \underbrace{\int_{t_{k-1}}^{t_k} |\dot\gamma(t)|\,\mathrm{d}t - |\dot\gamma(\tau)|(t_k - t_{k-1})}_{|\dots| \le \varepsilon(t_k - t_{k-1})} + \underbrace{|\dot\gamma(\tau)|(t_k - t_{k-1}) - |x_k - x_{k-1}|}_{|\dots| \le |\dot\gamma(\tau)(t_k - t_{k-1}) - (x_k - x_{k-1})| \le \varepsilon(t_k - t_{k-1})} \\ &\le 2\varepsilon(t_k - t_{k-1}) \end{aligned} \tag{7.5.4-G}$$

und damit nach Summation

$$\ell(\mathfrak{Z}) \ge \int_a^b |\dot\gamma(t)|\,\mathrm{d}t - 2\varepsilon(b-a) \tag{7.5.4-H}$$

und mit $\varepsilon \to 0$ die Behauptung. □

✂ **Ergänzung.** Parametrisiert man die Kurve Γ durch eine auf einem offenen Intervall stetig differenzierbare Funktion $\gamma : (a,b) \to \mathbb{R}^m$, so gilt entsprechend

$$\Gamma \text{ rektifizierbar} \quad \Longleftrightarrow \quad \ell(\Gamma) = \int_{\to a}^{\to b} |\dot\gamma(t)|\,\mathrm{d}t < \infty \tag{7.5.4-I}$$

unter Nutzung eines uneigentlichen Riemannintegrals.

Die durch jedes kompakte Teilintervall $[c,d] \subseteq (a,b)$ parametrisierte Teilkurve ist rektifizierbar, besitzt die durch $\int_c^d |\dot\gamma(t)|\,\mathrm{d}t$ gegebene Kurvenlänge und die Konvergenz des uneigentlichen Riemannintegrals entspricht damit gerade der Rektifizierbarkeit der Ausgangskuve.

⚒ **7.5.5 Beispiele.** Nachdem wir nun eine Formel zur Berechnung der Kurvenlänge einer differenzierbar parametrisierten Kurve haben, können wir dies anwenden und einige Kurvenlängen berechnen.

(i) Ein Kreis mit Radius r kann durch

$$\gamma : [0, 2\pi] \ni t \mapsto \begin{bmatrix} r\cos t \\ r\sin t \end{bmatrix} \tag{7.5.5-A}$$

parametrisiert werden. Für die Länge des Kreisbogens, also den Umfang des Kreises, erhalten wir damit wie zu erwarten

$$\ell = \int_0^{2\pi} \sqrt{r^2\sin^2 t + r^2\cos^2 t}\,\mathrm{d}t = r\int_0^{2\pi} \mathrm{d}t = 2\pi r. \tag{7.5.5-B}$$

(ii) Betrachten wir stattdessen eine *Ellipse* mit Halbachsen a und b, so führt die Parametrisierung

$$\gamma : [0, 2\pi] \ni t \mapsto \begin{bmatrix} a\cos t \\ b\sin t \end{bmatrix} \tag{7.5.5-C}$$

zu dem im Allgemeinen nicht explizit berechenbaren elliptischen Integral

$$\ell = \int_0^{2\pi} \sqrt{a^2\sin^2 t + b^2\cos^2 t}\,\mathrm{d}t = 2\int_{-\pi}^{\pi} \sqrt{(a^2-b^2)\sin^2 t + b^2}\,\mathrm{d}t. \tag{7.5.5-D}$$

(iii) Graphen von Funktionen $f : [a, b] \to \mathbb{R}$ sind parametrisierte Kurven in der Ebene. Dabei ist

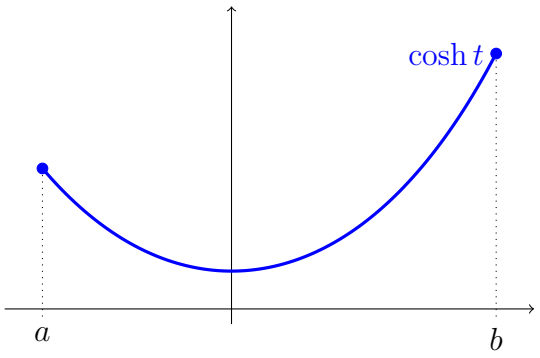
$$\gamma : t \mapsto \begin{bmatrix} t \\ f(t) \end{bmatrix} \tag{7.5.5-E}$$

und für differenzierbares f liefert damit

$$\int_a^b \sqrt{1 + (f'(t))^2}\,\mathrm{d}t \tag{7.5.5-F}$$

die Länge des Graphen über dem Intervall $[a, b]$. Für die *Kettenlinie* $f(t) = \cosh t$ ergibt sich damit zum Beispiel

$$\int_a^b \sqrt{1 + \sinh^2 t}\,\mathrm{d}t = \int_a^b \cosh t\,\mathrm{d}t = \sinh b - \sinh a. \tag{7.5.5-G}$$



7.5.6. Das Integral (7.5.**4-A**) ist unabhängig von der gewählten differenzierbaren Parametrisierung, da es ja die vorher definierte Kurvenlänge berechnet. Wir können die Unabhängigkeit von der Parametrisierung aber auch nachrechnen. Sei

$$\varphi : [c, d] \to [a, b] \tag{7.5.6-A}$$

monoton wachsend und differenzierbar. Dann liefert $\gamma \circ \varphi : [c, d] \to \Gamma$ eine weitere differenzierbare Parametrisierung derselben Kurve Γ und es gilt auf Grund der Kettenregel und unter Ausnutzung von $\varphi'(s) \geq 0$

$$\int_c^d \left|\frac{\mathrm{d}}{\mathrm{d}s}\gamma(\varphi(s))\right| \mathrm{d}s = \int_c^d |\dot\gamma(\varphi(s))|\varphi'(s)\,\mathrm{d}s = \int_a^b |\dot\gamma(t)|\,\mathrm{d}t. \tag{7.5.6-B}$$

Obiger Satz kann genutzt werden, um eine Parametrisierung nach Bogenlänge einer gegebenen Kurve mit regulärer Parametrisierung zu bestimmen. Dazu suchen wir eine Umparametrisierung $\varphi : [0, \ell] \to [a, b]$ mit

$$|\dot\gamma(\varphi)|\frac{\mathrm{d}\varphi}{\mathrm{d}s} = 1 \tag{7.5.6-C}$$

und damit

$$s = \int_0^s \mathrm{d}s = \int_a^{\varphi(s)} |\dot\gamma(\varphi)|\,\mathrm{d}\varphi = F(\varphi(s)) - F(a) \tag{7.5.6-D}$$

mit einer Stammfunktion F zu $t \mapsto |\dot\gamma(t)|$. Da für eine reguläre Parametrisierung $\dot\gamma(t) \neq 0$ gilt, ist F streng monoton wachsend und die Umkehrfunktion φ damit auch differenzierbar. Die Bogenlängenparametrisierung $\gamma \circ \varphi$ also insbesondere wiederum regulär.

✂ **Ergänzung.** Man kann Funktionen entlang rektifizierbarer Kurven bezüglich der Kurvenlänge integrieren. Dies führt zum *Kurvenintergral erster Art*, das wir der Vollständigkeit halber hier erwähnen. Sei dazu Γ eine rektifizierbare Kurve und $f : \Gamma \to \mathbb{R}$ eine auf Γ gleichmäßig stetige Funktion. Dann konvergieren für feiner werdende Zerlegungen $\mathfrak{Z}$ der Kurve die Riemannsummen

$$\mathcal{S}(f; \mathfrak{Z}, (\xi_k)) = \sum_{k=1}^{n} f(\xi_k)|x_k - x_{k-1}|, \qquad \mathfrak{Z} = (x_0, \ldots, x_n), \tag{7.5.6-E}$$

wobei die Zwischenstellen ξ_k gemäß $x_0 \prec \xi_1 \preceq x_1 \prec \cdots \preceq x_{n-1} \prec \xi_n \preceq x_n$ gewählt wurden. Die Konvergenz folgt dabei aus der gleichmäßigen Stetigkeit von f. Wie schon bei der Konstruktion des Riemannintegrals spielt die Wahl der Zwischenstellen keine Rolle, man kann also stets $\xi_k = x_k$ wählen. Wir definieren deswegen

$$\int_\Gamma f(x)\,\mathrm{d}\ell(x) = \lim_{\substack{\delta(\mathfrak{Z})\to 0 \\ \ell(\mathfrak{Z})\to\ell(\Gamma)}} \mathcal{S}(f; \mathfrak{Z}) \tag{7.5.6-F}$$

als Grenzwert in dem Sinne, dass für jedes $\varepsilon > 0$ ein $\delta > 0$ mit

$$\delta(\mathfrak{Z}) < \delta \quad \wedge \quad \ell(\mathfrak{Z}) \geq \ell(\Gamma) - \delta \qquad \Longrightarrow \qquad \left|\int_\Gamma f(x)\,\mathrm{d}\ell(x) - \mathcal{S}(f; \mathfrak{Z})\right| < \varepsilon \tag{7.5.6-G}$$

existiert. Dabei bezeichnet $\delta(\mathfrak{Z}) := \max_k \ell(\Gamma_{[x_{k-1}, x_k]})$ die Feinheit von $\mathfrak{Z}$.

Ist nun $\gamma : [a, b] \to \Gamma$ eine differenzierbare Parametrisierung der Kurve Γ, so kann das gerade definierte Integral über diese Parametrisierung berechnet werden. Es gilt

$$\int_\Gamma f(x)\,\mathrm{d}\ell(x) = \int_a^b f(\gamma(t))|\dot\gamma(t)|\,\mathrm{d}t. \tag{7.5.6-H}$$

Das Integral ist von der gewählten Parametrisierung unabhängig. Dies folgt aus der oben angegebenen Definition und, ebenso wie für die Kurvenlänge gezeigt, aus der Substitutionsregel für Integrale.

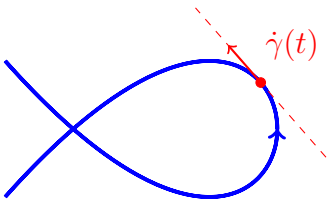
7.5.7. Sei nun $\gamma : (a, b) \to \Gamma$ eine reguläre Parametrisierung und $t \in (a, b)$. Dann kann die Menge

$$\{\gamma(t) + \lambda\dot{\gamma}(t) \mid \lambda \in \mathbb{R}\} \tag{7.5.7-A}$$

als Tangente an die Kurve Γ im Punkt $\gamma(t)$ interpretiert werden. Dies folgt direkt aus der Definition der Ableitung, es gilt

$$\gamma(t + h) = \gamma(t) + h\dot{\gamma}(t) + \mathbf{o}(h). \tag{7.5.7-B}$$

Reguläre Kurven besitzen also in jedem Kurvenpunkt Tangenten (in Doppelpunkten in der Regel mehrere Tangenten).



Ist γ eine Parametrisierung nach Bogenlänge und sogar zweifach differenzierbar, so kann die Gleichung[12]

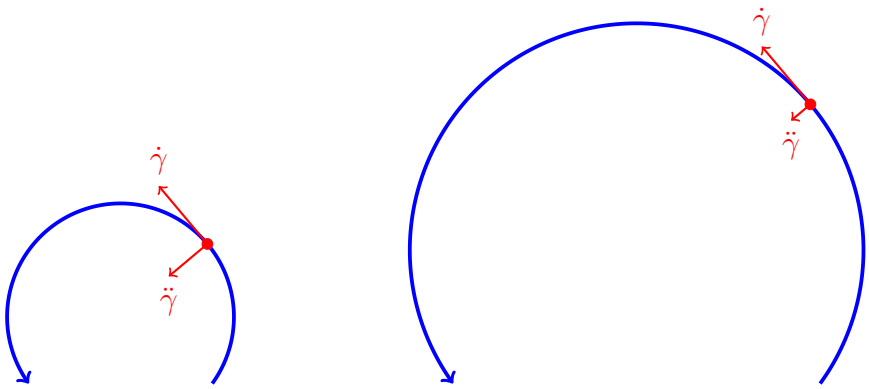
$$1 = |\dot{\gamma}(t)|^2 = \dot{\gamma}(t)^\top \dot{\gamma}(t) \tag{7.5.7-C}$$

nochmals differenziert werden und wir erhalten

$$0 = \ddot{\gamma}(t)^\top \dot{\gamma}(t) + \dot{\gamma}(t)^\top \ddot{\gamma}(t) = 2\dot{\gamma}(t)^\top \ddot{\gamma}(t) \tag{7.5.7-D}$$

und der Vektor $\ddot{\gamma}(t)$ steht in jedem Kurvenpunkt senkrecht auf dem Tangenteneinheitsvektor $\dot{\gamma}(t)$. Dieser Vektor beschreibt die Krümmung der Kurve. Um das zu verdeutlichen, betrachten wir in der Ebene einen Kreis mit Radius $\rho > 0$ in Bogenlängenparametrisierung

$$\gamma(t) = \rho \begin{bmatrix} \cos(\rho^{-1}t) \\ \sin(\rho^{-1}t) \end{bmatrix}, \quad \dot{\gamma}(t) = \begin{bmatrix} -\sin(\rho^{-1}t) \\ \cos(\rho^{-1}t) \end{bmatrix}, \quad \ddot{\gamma}(t) = -\rho^{-1} \begin{bmatrix} \cos(\rho^{-1}t) \\ \sin(\rho^{-1}t) \end{bmatrix}. \tag{7.5.7-E}$$



[12] Dies geht nur mit dem Euklidischen Abstand, der durch ein Skalarprodukt definiert ist.

Allgemein bezeichnet $\varkappa := |\ddot\gamma(t)|$ die Krümmung der in Bogenlänge parametrisierten Kurve im Punkt $\gamma(t)$. Eine in der Ebene

$$\{\gamma(t) + \lambda\dot\gamma(t) + \mu\ddot\gamma(t) \mid \lambda, \mu \in \mathbb{R}\} \tag{7.5.7-F}$$

liegende Kreislinie mit Radius $\varkappa^{-1}$ um den Punkt $\gamma(t) + \varkappa^{-2}\ddot\gamma(t)$ berührt die Kurve zu zweiter Ordnung. Es gilt also mit $\varkappa = |\ddot\gamma(t)|$

$$\gamma(t+h) - \left(\gamma(t) + \varkappa^{-2}\ddot\gamma(t) + \varkappa^{-1}\left(\dot\gamma(t)\sin(\varkappa h) - \varkappa^{-1}\ddot\gamma(t)\cos(\varkappa h)\right)\right) = \mathbf{o}(h^2) \tag{7.5.7-G}$$

für $h \to 0$. Die Aussage folgt direkt aus der taylorschen Formel komponentenweise gelesen: Einsetzen von $h = 0$ liefert

$$\gamma(t) - \left(\gamma(t) + \varkappa^{-2}\ddot\gamma(t) + \varkappa^{-1}\left(-\varkappa^{-1}\ddot\gamma(t)\right)\right) = 0, \tag{7.5.7-H}$$

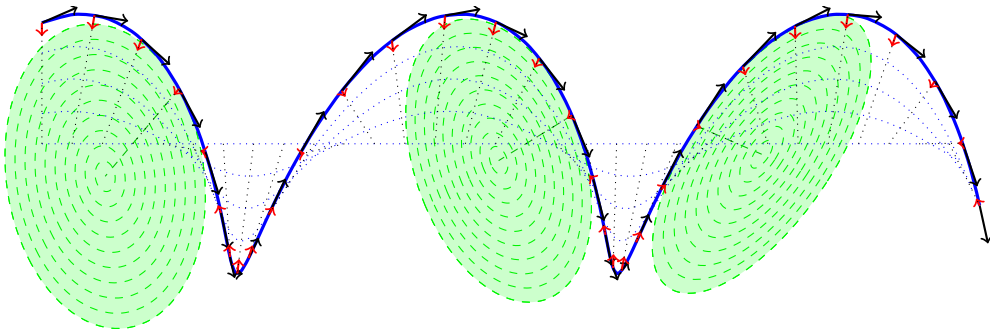
für die erste Ableitung bezüglich h erhält man für $h = 0$

$$\dot\gamma(t) - \left(\varkappa^{-1}\left(\dot\gamma(t)\varkappa\right)\right) = 0 \tag{7.5.7-I}$$

und analog für die zweite Ableitung

$$\ddot\gamma(t) - \left(\varkappa^{-1}\left(\varkappa^{-1}\ddot\gamma(t)\varkappa^2\right)\right) = 0. \tag{7.5.7-J}$$

Funktionen die in einem Punkt zusammen mit den ersten beiden Ableitungen verschwinden sind $\mathbf{o}(h^2)$.



⚒ **7.5.8 Beispiel.** Eine archimedische Spirale wird durch $\gamma : \mathbb{R}_+ \mapsto \mathbb{R}^2$ mit

$$\gamma(t) = \begin{bmatrix} at\cos t \\ at\sin t \end{bmatrix} \tag{7.5.8-A}$$

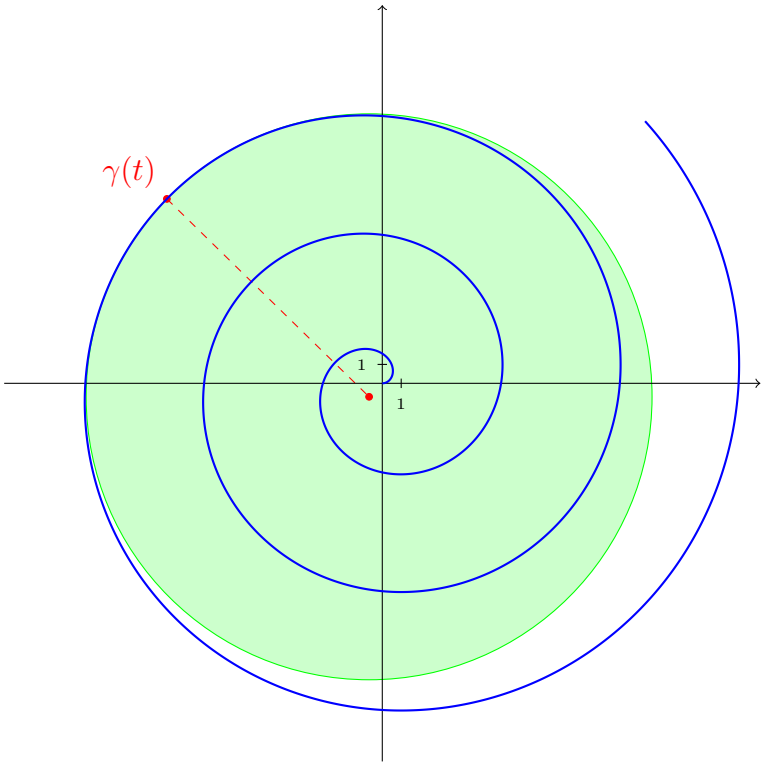
parametrisiert. Wir berechnen Kurvenlänge und Krümmung. Dazu stellen wir zuerst fest, dass γ differenzierbar ist und wegen

$$\dot\gamma(t) = \begin{bmatrix} a\cos t - at\sin t \\ a\sin t + at\cos t \end{bmatrix}, \qquad |\dot\gamma(t)| = a\sqrt{1+t^2} \tag{7.5.8-B}$$

das durch das Intervall $[0, T]$ parametrisierte Kurvenstück die Länge

$$\ell(T) = \int_0^T |\dot\gamma(t)|\,\mathrm{d}t = a\int_0^T \sqrt{1+t^2}\,\mathrm{d}t = \frac{a}{2}\left(T\sqrt{1+T^2} + \operatorname{arsinh} T\right) \tag{7.5.8-C}$$

besitzt.



Für die Krümmung parametrisieren wir die archimedische Spirale nach ihrer Kurvenlänge, verwenden also die Parametrisierung

$$\widetilde{\gamma}(\ell) = \gamma(t(\ell)) \tag{7.5.8-D}$$

mit der Umkehrfunktion $t(\ell)$ zu $\ell(t)$ aus (7.5.8-C). Differenzieren führt unter Nutzung der Kettenregel und der Formel für die Ableitung der (eindimensionalen) Umkehrfunktion auf

$$\widetilde{\gamma}'(\ell) = \dot{\gamma}(t(\ell))t'(\ell) = \frac{1}{\ell'(t)}\dot{\gamma}(t)\Big|_{t=t(\ell)} = \frac{\dot{\gamma}(t)}{|\dot{\gamma}(t)|}\Big|_{t=t(\ell)} \tag{7.5.8-E}$$

sowie wegen $t''(\ell) = -\frac{\ell''(t)}{(\ell'(t))^3}$

$$\widetilde{\gamma}''(\ell) = \ddot{\gamma}(t(\ell))\big(t'(\ell)\big)^2 + \dot{\gamma}(t(\ell))t''(\ell) = \frac{\ddot{\gamma}(t)}{|\dot{\gamma}(t)|^2}\Big|_{t=t(\ell)} - \frac{\dot{\gamma}(t)}{|\dot{\gamma}(t)|^3}\frac{\mathrm{d}}{\mathrm{d}t}|\dot{\gamma}(t)|\Big|_{t=t(\ell)}. \tag{7.5.8-F}$$

Die Krümmung entspricht der Länge dieses Vektors, es bleibt also nur noch die entsprechenden Ableitungen einzusetzen. Es ergibt sich als Vektor

$$\begin{aligned}
&\frac{1}{a(1+t^2)}\begin{bmatrix}-2\sin t - t\cos t\\ 2\cos t - t\sin t\end{bmatrix} - \frac{t}{a(1+t^2)^2}\begin{bmatrix}\cos t - t\sin t\\ \sin t + t\cos t\end{bmatrix}\\
&\qquad = \frac{1}{a(1+t^2)^2}\begin{bmatrix}-(1+t^2)(2\sin t + t\cos t) - t(\cos t - t\sin t)\\ (1+t^2)(2\cos t - t\sin t) - t(\sin t + t\cos t)\end{bmatrix}\\
&\qquad = \frac{2+t^2}{a(1+t^2)^2}\begin{bmatrix}-\sin t - t\cos t\\ \cos t - t\sin t\end{bmatrix}
\end{aligned} \tag{7.5.8-G}$$

und damit als Krümmung im Kurvenpunkt $\gamma(t)$

$$\varkappa(t) = \frac{2+t^2}{a(1+t^2)^2}\sqrt{1+t^2} = \frac{2+t^2}{a(1+t^2)^{3/2}} = \frac{1}{at} + \mathbf{o}(t^{-1}), \qquad t \to \infty. \tag{7.5.8-H}$$

7.6 Skalare Funktionen, Differentiale und Integrale

7.6.1. Differentialrechnung kann auch im Höherdimensionalen zum Berechnen von Extremwerten genutzt werden. Ist $f : \mathbb{R}^n \supseteq \Omega \to \mathbb{R}$ differenzierbar und $x_* \in \Omega$ eine lokale Extremstelle, gilt also entweder $f(x) \leq f(x_*)$ auf einer Umgebung des Punktes x_* oder gilt $f(x) \geq f(x_*)$ auf einer Umgebung von x_*, so folgt

$$\frac{\partial f}{\partial x_1}(x_*) = \cdots = \frac{\partial f}{\partial x_n}(x_*) = 0 \tag{7.6.1-A}$$

und damit $f'(x_*) = 0$. Dies liefert ein *notwendiges Kriterium* für lokale Extremstellen.[13] Punkte $x_* \in \Omega$ mit $f'(x_*) = 0$ bezeichnet man deshalb oft als *kritische Punkte*.

Funktionswerte $y \in f[\Omega]$, für die es keinen kritischen Punkt x_* mit $f(x_*) = y$ gibt, heißen *reguläre Werte*. Ist f stetig differenzierbar und y ein regulärer Wert, so ist nach dem Satz über implizite Funktionen die *Niveaumenge*

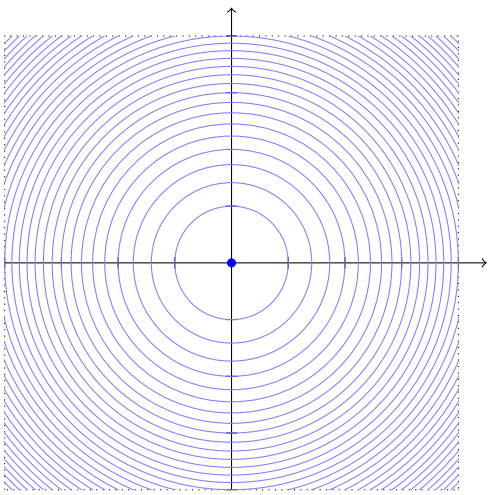
$$f^{-1}[y] = \{x \in \Omega \mid f(x) = y\} \tag{7.6.1-B}$$

lokal um jeden ihrer Punkte als Graph einer stetig differenzierbaren Funktion darstellbar. Gerade für $n = 2$ oder $n = 3$ können Niveaumengen helfen, sich den Verlauf einer Funktion besser vorzustellen. Dazu betrachten wir zwei Beispiele.

⚒ **7.6.2 Beispiel.** Wir geben zwei Beispiele und betrachten Funktionen $f, g : [-4, 4]^2 \to \mathbb{R}$. Für das erste Beispiel nutzen wir

$$f(x, y) = x^2 + y^2 \tag{7.6.2-A}$$

mit kritischem Punkt in $x = 0$ und Kreisen $x^2 + y^2 = r^2 \in \mathbb{N}$ als Niveaulinien. In der Darstellung erkennt man das lokale (und globale) Extremum im Ursprung und (da die Linien zu äquidistanten Niveaus gezeichnet sind) den stärkeren Anstieg der Funktion am Bildrand.

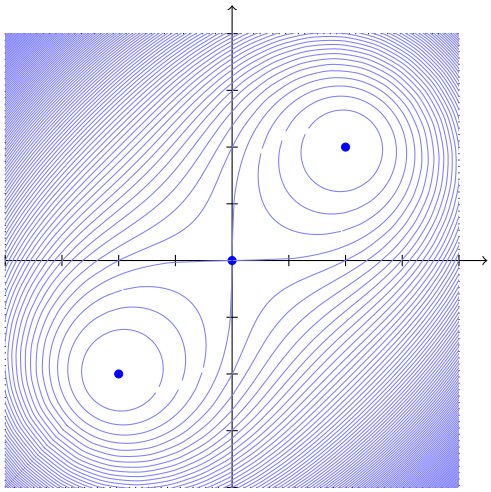


[13] Ein hinreichendes Kriterium werden wir in Kapitel 7.7 diskutieren.

In einem zweiten Beispiel betrachten wir

$$g(x,y) = ((x-2)^2 + (y-2)^2)((x+2)^2 + (y+2)^2) \tag{7.6.2-B}$$

mit kritischen Punkten in $(\pm 2, \pm 2)$ und im Ursprung. Am Bild ist zu sehen, dass $(2,2)$ und $(-2,-2)$ lokale Extrema sind, aber $(0,0)$ nicht.



Bei den in diesem Beispiel erhaltenen Niveaulinien handelt es sich um *Cassinische Kurven*[14], also Kurven bei denen das Produkt der Abstände zu zwei gegebenen Brennpunkten konstant ist.

7.6.3. Im weiteren Verlauf wollen wir uns die Ableitungen skalarer Funktionen genauer anschauen. Dazu ergänzen wir etwas mehr lineare Algebra.

✂ **Ergänzung** (Dualräume)**.** Sei V ein endlichdimensionaler Vektorraum über dem Körper $\mathbb{K}$. Dann bezeichnet $V^* := \mathrm{Lin}(V, \mathbb{K})$ seinen *Dualraum*. Ist $\mathbf{v}_1, \dots, \mathbf{v}_n$ eine Basis von V, dann kann jedes $x \in V$ eindeutig als Linearkombination

$$x = x_1\mathbf{v}_1 + \cdots + x_n\mathbf{v}_n = \sum_{j=1}^{n} x_j \mathbf{v}_j \tag{7.6.3-A}$$

geschrieben werden. Die Zuordnungen $\hat{\mathbf{v}}_j : x \mapsto x_j$ sind dabei linear, also Elemente des Dualraumes V^*. Für Anwendungen eines $\psi \in V^*$ auf ein $x \in V$ schreiben wir $\langle \psi, x \rangle$. Es gilt also für die Koordinatenabbildungen $\hat{\mathbf{v}}_j$

$$\langle \hat{\mathbf{v}}_j, \mathbf{v}_i \rangle = \delta_{i,j} = \begin{cases} 1, & i = j, \\ 0 & \text{sonst.} \end{cases} \tag{7.6.3-B}$$

Man sieht damit direkt, dass die $\hat{\mathbf{v}}_j$ linear unabhängig sind. Darüberhinaus gilt für jedes $\psi \in V^*$

$$\langle \psi, x \rangle = \left\langle \psi, \sum_{j=1}^{n} x_j \mathbf{v}_j \right\rangle = \sum_{j=1}^{n} x_j \langle \psi, \mathbf{v}_j \rangle \tag{7.6.3-C}$$

und damit

$$\psi = \sum_{j=1}^{n} \langle \psi, \mathbf{v}_j \rangle \hat{\mathbf{v}}_j. \tag{7.6.3-D}$$

Also ist $\hat{\mathbf{v}}_1, \dots, \hat{\mathbf{v}}_n$ Basis von V^*, die zu $\mathbf{v}_1, \dots, \mathbf{v}_n$ *duale Basis*. Es gilt $\dim V^* = \dim V$.

[14] Giovanni Domenico Cassini, 1625–1712

✂ **Ergänzung.** Sind V und W zwei Vektorräume und ist $f \in \mathrm{Lin}(V,W)$ eine lineare Abbildung, so induziert diese durch

$$\langle f^\top(\psi), x\rangle := \langle \psi, f(x)\rangle \tag{7.6.3-E}$$

eine lineare Abbildung $f^\top \in \mathrm{Lin}(W^*, V^*)$ zwischen den Dualräumen. Diese wird als die zu f *transponierte Abbildung* bezeichnet.

Im Falle der Vektorräume $\mathbb{K}^m$ und $\mathbb{K}^n$ entspricht Transponieren linearer Abbildungen dem Transponieren von Matrizen, jedoch nur, wenn man die Dualräume auch als Spaltenvektoren schreibt. Es gilt dann mit $\langle x, y\rangle = x^\top y$ für $x, y \in \mathbb{K}^m$ für jedes $A \in \mathbb{K}^{m\times n}$ und $z \in \mathbb{K}^n$

$$\langle x, Az\rangle = x^\top(Az) = (A^\top x)^\top z = \langle A^\top x, z\rangle. \tag{7.6.3-F}$$

7.6.4. Ist $f : \mathbb{R}^n \supseteq \Omega \to \mathbb{R}$ differenzierbar, so bestimmt die Ableitung

$$f' : \Omega \to \mathcal{L}(\mathbb{R}^n, \mathbb{R}) =: (\mathbb{R}^n)^* \tag{7.6.4-A}$$

an jedem Punkt $x \in \Omega$ ein lineares Funktional auf dem $\mathbb{R}^n$. Dafür führen wir nun eine spezielle Notation ein, wir bezeichnen dieses lineare Funktional als $\mathrm{d}f(x)$ und schreiben es als Linearkombination der linearen Funktionale

$$\mathrm{d}x_j : \mathbb{R}^n \to \mathbb{R}, \qquad \mathrm{d}x_j(h) = \langle \mathrm{d}x_j, h\rangle := h_j, \tag{7.6.4-B}$$

also der Ableitungen der Koordinatenfunktionen $x \mapsto x_j$, als

$$\mathrm{d}f = \frac{\partial f}{\partial x_1}\,\mathrm{d}x_1 + \cdots + \frac{\partial f}{\partial x_n}\,\mathrm{d}x_n = \sum_{j=1}^n \frac{\partial f}{\partial x_j}\,\mathrm{d}x_j. \tag{7.6.4-C}$$

Die Abbildung $\mathrm{d}f$ heißt *Differential* der Funktion f.

Nach Konstruktion bilden die Funktionale $\mathrm{d}x_1, \ldots, \mathrm{d}x_n$ die zu $\mathrm{e}_1, \ldots, \mathrm{e}_n$ duale Basis. Es gilt also $\langle \mathrm{d}x_i, \mathrm{e}_j\rangle = \delta_{i,j}$. Die Schreibweise für das Differential entspricht der Basisdarstellung der linearen Abbildung $f'(x) \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}) = (\mathbb{R}^n)^*$ mittels (7.6.3-D)

$$\mathrm{d}f(x) = \sum_{j=1}^n \langle f'(x), \mathrm{e}_j\rangle\,\mathrm{d}x_j = \sum_{j=1}^n \frac{\partial f}{\partial x_j}(x)\,\mathrm{d}x_j, \qquad (\partial_{x_j} f)(x) = \langle f'(x), \mathrm{e}_j\rangle, \tag{7.6.4-D}$$

in Kombination mit der Definition der partiellen Ableitung. Formel (7.6.4-C) ist also nur eine andere Schreibweise der Ableitung (7.3.4-F) im Falle skalarwertiger Funktionen. Der Sinn dieser Notation ergibt sich aus den Eigenschaften solcher Differentiale.

7.6.5. Sei $x : [a,b] \to \Omega$ eine differenzierbare Parametrisierung einer Kurve $\Gamma \subset \Omega$ und $f : \Omega \to \mathbb{R}$ differenzierbar. Dann gilt für die Ableitung der Verkettung $f \circ x : [a,b] \to \mathbb{R}$ nach der Kettenregel

$$\frac{\mathrm{d}f(x(t))}{\mathrm{d}t} = \begin{bmatrix} \frac{\partial f}{\partial x_1} & \cdots & \frac{\partial f}{\partial x_n}\end{bmatrix} \begin{bmatrix} \frac{\mathrm{d}x_1}{\mathrm{d}t} \\ \vdots \\ \frac{\mathrm{d}x_n}{\mathrm{d}t}\end{bmatrix} = \sum_{j=1}^n \frac{\partial f}{\partial x_j}\frac{\mathrm{d}x_j}{\mathrm{d}t}. \tag{7.6.5-A}$$

Wir vergleichen dies mit (7.6.4-C). Da für jedes $x \in \Omega$ das Differential $\mathrm{d}f(x) \in \mathcal{L}(\mathbb{R}^n, \mathbb{R})$ eine lineare Abbildung ist, kann in $\mathrm{d}f(x)$ der Tangentenvektor $\dot{x}(t)$ am Kurvenpunkt t eingesetzt werden. Dies liefert eine Zahl, es gilt

$$\langle\, \mathrm{d}f(x(t)), \dot{x}(t)\rangle = \sum_{j=1}^n \frac{\partial f}{\partial x_j}\frac{\mathrm{d}x_j}{\mathrm{d}t} = \frac{\mathrm{d}f(x(t))}{\mathrm{d}t}. \tag{7.6.5-B}$$

Darüberhinaus ist das Integral

$$\int_\Gamma \mathrm{d}f := \int_a^b \langle\, \mathrm{d}f(x(t)), \dot{x}(t)\rangle \,\mathrm{d}t = \int_a^b \frac{\mathrm{d}f(x(t))}{\mathrm{d}t}\,\mathrm{d}t = f(x(b)) - f(x(a)) \qquad (7.6.\textbf{5-C})$$

von der gewählten Parametrisierung unabhängig. Letzteres gilt allgemeiner, wir bezeichnen eine stetige Abbildung $\omega : \Omega \to (\mathbb{R}^n)^*$ als *Pfaffsche Form*[15] oder *Differentialform* vom Grad 1. Jede solche Differentialform kann als Linearkombination

$$\omega(x) = g_1(x)\,\mathrm{d}x_1 + \cdots + g_n(x)\,\mathrm{d}x_n = \sum_{j=1}^n g_j(x)\,\mathrm{d}x_j \qquad (7.6.\textbf{5-D})$$

mit stetigen Koeffizienten $g_1, \ldots, g_n : \Omega \to \mathbb{R}$ geschrieben werden. Wir definieren das *Kurvenintegral zweiter Art* oder für uns auch kurz *Kurvenintegral* von ω entlang der differenzierbar parametrisierten Kurve Γ analog durch

$$\int_\Gamma \omega := \int_a^b \langle \omega(x(t)), \dot{x}(t)\rangle \,\mathrm{d}t = \int_a^b \sum_{j=1}^n g_j(x(t)) \frac{\mathrm{d}x_j(t)}{\mathrm{d}t}\,\mathrm{d}t. \qquad (7.6.\textbf{5-E})$$

Dieses Integral ist nicht von der gewählten differenzierbaren Parametrisierung der Kurve Γ abhängig. Für jede differenzierbare Umparametrisierung $t = \varphi(s)$, $\varphi : [c,d] \to [a,b]$, ergibt die Substitutionsregel für Integrale

$$\begin{aligned}\int_a^b \left\langle \omega(x(\varphi(s))), \frac{\mathrm{d}x(\varphi(s))}{\mathrm{d}s} \right\rangle \mathrm{d}s &= \int_a^b \langle \omega(x(\varphi(s))), \dot{x}(\varphi(s))\rangle \,\frac{\mathrm{d}\varphi}{\mathrm{d}s}\,\mathrm{d}s \\ &= \int_a^b \langle \omega(x(t)), \dot{x}(t)\rangle \,\mathrm{d}t\end{aligned} \qquad (7.6.\textbf{5-F})$$

und damit die behauptete Unabhängigkeit.

Pfaffsche Formen sind allgemeiner als Differentiale. Für Differentiale besagt (7.6.**5-C**) ebenso, dass das Integral nur von den Endpunkten der Kurve und nicht von deren Verlauf abhängt. Für Formen muss das im Allgemeinen nicht gelten, ein Beispiel dazu ist

$$\int_\Gamma \big(x_1\,\mathrm{d}x_2 - x_2\,\mathrm{d}x_1\big), \qquad (7.6.\textbf{5-G})$$

welches für den Einheitskreis, also $x_1(t) = \cos t$ und $x_2(t) = \sin t$ mit $t \in [0, 2\pi]$,

$$\oint \big(x_1\,\mathrm{d}x_2 - x_2\,\mathrm{d}x_1\big) = \int_0^{2\pi} \cos^2 t + \sin^2 t\,\mathrm{d}t = 2\pi \qquad (7.6.\textbf{5-H})$$

liefert. Es gibt damit also insbesondere auch keine differenzierbare Funktion $f : \mathbb{R}^2 \to \mathbb{R}$ mit $\mathrm{d}f(x) = x_1\,\mathrm{d}x_2 - x_2\,\mathrm{d}x_1$.

✂ **Ergänzung.** Pfaffsche Formen kann man entlang rektifizierbarer Kurven integrieren. Wir geben dazu eine Definition des Kurvenintegrals, die mit obiger im Falle differenzierbar parametrisierbarer Kurven übereinstimmt (und liefern gleichzeitig ein alternatives Argument für die Unabhängigkeit des Integrals von der Parametrisierung).

[15] Johann Friedrich Pfaff, 1765–1825

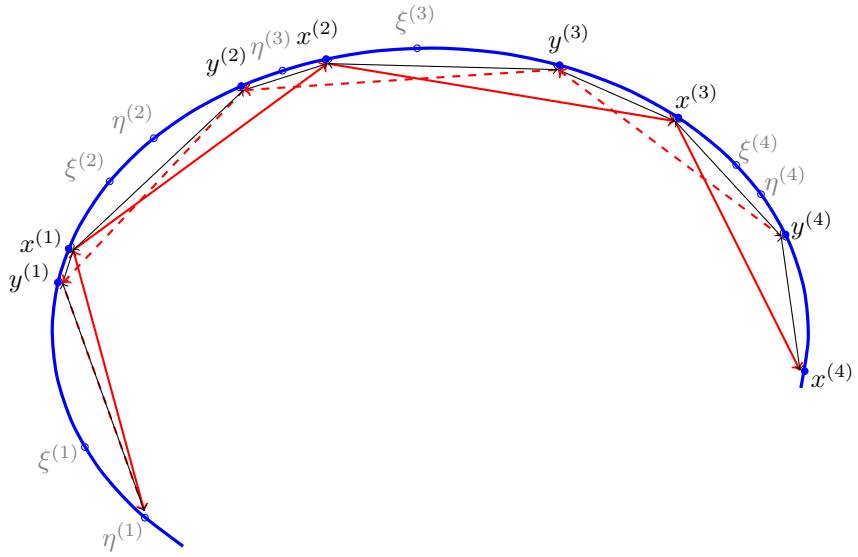
Sei dazu $\Gamma \subset \Omega$ rektifizierbar und $\operatorname{acc}(\Gamma) \cap \partial\Omega = \varnothing$. Für eine Zerlegung $\mathfrak{Z} = (x^{(0)}, \dots, x^{(m)})$ der Kurve Γ und Zwischenstellen $\xi^{(k)}$ mit $x^{(k-1)} \prec \xi^{(k)} \preceq x^{(k)}$ betrachten wir die analog zu einer Riemannsumme gebildete Summe

$$\mathcal{S}(\omega; \mathfrak{Z}, (\xi^{(k)})) = \sum_{k=1}^{m} \langle \omega(\xi^{(k)}), x^{(k)} - x^{(k-1)} \rangle. \tag{7.6.5-I}$$

Wir zeigen, dass diese Summen bei feiner werdenden Zerlegungen konvergieren und ein Integral definieren. Da Γ beschränkt (es ist abgeschlossene Teilmenge einer Kugel vom Radius $\ell(\Gamma) + 1$) und separiert von $\partial\Omega$ ist, ist Γ in einer kompakten Teilmenge von Ω enthalten. Die Koeffizienten g_j der Form ω sind damit auf Γ gleichmäßig stetig. Also existiert zu jedem $\varepsilon > 0$ ein $\delta > 0$, so dass

$$\ell(\Gamma_{[x^{(k-1)}, x^{(k)}]}) < \delta \qquad \Longrightarrow \qquad \sup_{\xi : x^{(k-1)} \prec \xi \prec x^{(k)}} g_j(x) - \inf_{\xi : x^{(k-1)} \prec \xi \prec x^{(k)}} g_j(x) < \varepsilon \tag{7.6.5-J}$$

gilt. Wir wählen nun Zerlegungen $\mathfrak{Z}$ so, dass $\ell(\mathfrak{Z}) \geq \ell(\Gamma) - \varepsilon$ und $\delta(\mathfrak{Z}) = \max_k \ell(\Gamma_{[x^{(k-1)}, x^{(k)}]}) < \delta/2$ gilt.



Dann folgt für zwei solche Zerlegungen $\mathfrak{Z}_1 = (x^{(k)})$ und $\mathfrak{Z}_2 = (y^{(k)})$, beliebige untergeordnete Stützstellensätze $(\xi^{(k)})$ und $(\eta^{(k)})$ und ihre gemeinsame Verfeinerung $\mathfrak{Z}_1 \cup \mathfrak{Z}_2 = (z^{(l)})$

$$\begin{aligned}
\mathcal{S}(\omega; \mathfrak{Z}_1, (\xi^{(k)})) - \mathcal{S}(\omega; \mathfrak{Z}_2, (\eta^{(k)})) &= \sum_{k=1}^{m} \langle \omega(\xi^{(k)}), x^{(k)} - x^{(k-1)} \rangle - \sum_{k=1}^{m'} \langle \omega(\eta^{(k)}), y^{(k)} - y^{(k-1)} \rangle \\
&= \langle \omega(\zeta), z^{(1)} - z^{(0)} \rangle \\
&\quad + \sum_{l=2}^{m+m'-1} \langle \omega(\xi^{(k_l)}) - \omega(\eta^{(\tilde{k}_l)}), z^{(l+j)} - z^{(l+j-1)} \rangle \\
&\quad + \langle \omega(\zeta'), z^{(m+m')} - z^{(m+m'-1)} \rangle
\end{aligned} \tag{7.6.5-K}$$

unter Ausnutzung der Linearität von $\omega(\xi^{(k)})$ beim Ersetzen von

$$x^{(k)} - x^{(k-1)} = \sum_{j=0}^{\nu} z^{(l+j)} - z^{(l+j-1)} \qquad \text{mit } z^{(l-1)} = x^{(k-1)} \text{ und } z^{(l+\nu)} = x^{(k)},\ k = k_l \tag{7.6.5-L}$$

und entsprechend

$$y^{(k)} - y^{(k-1)} = \sum_{j=0}^{\nu} z^{(l+j)} - z^{(l+j-1)} \qquad \text{mit } z^{(l-1)} = y^{(k-1)} \text{ und } z^{(l+\nu)} = y^{(k)},\ k = \tilde{k}_l \tag{7.6.5-M}$$

und mit $\zeta = \min_\prec\{\xi^{(1)}, \eta^{(1)}\}$ und $\zeta' = \max_\prec\{\xi^{(m)}, \eta^{(m')}\}$. Der erste und der letzte Summand sind dabei kleiner als $M\delta$ mit $M = \max_{x\in\Gamma} \|\omega(x)\|$. In der verbleibenden Summe ist jeder Summand wegen $\ell(\Gamma_{[\xi^{(k_l)},\eta^{(\tilde{k}_l)}]}) < \delta$ durch (7.6.**5-J**) mit $m\varepsilon|z^{(l+j)} - z^{(l+j-1)}|$ abschätzbar. Also gilt

$$\left|\mathcal{S}(\omega;\mathfrak{Z}_1,(\xi^{(k)})) - \mathcal{S}(\omega;\mathfrak{Z}_2,(\eta^{(k)}))\right| \le \varepsilon\ell(\mathfrak{Z}_1\cup\mathfrak{Z}_2) + M\delta \le \varepsilon\ell(\Gamma) + M\delta \tag{7.6.5-N}$$

und für feiner werdende Zerlegungen $\mathfrak{Z}$ sind die Riemannsummen $\mathcal{S}(\omega;\mathfrak{Z},(\xi^{(k)}))$ damit Cauchy. Also können wir

$$\int_\Gamma \omega := \lim_{\substack{\delta(\mathfrak{Z})\to 0\\ \ell(\mathfrak{Z})\to\ell(\Gamma)}} \mathcal{S}(\omega;\mathfrak{Z},(\xi^{(k)})) \tag{7.6.5-O}$$

definieren. Die Wahl der Zwischenstellen spielt wiederum keine Rolle, man kann also wie bei Riemannintegralen und Kurvenintegralen erster Art stets $\xi^{(k)} = x^{(k)}$ wählen.

Eine offene und zusammenhängende Teilmenge des $\mathbb{R}^n$ bezeichnet man als ein *Gebiet*. In Gebieten gibt es zu je zwei Punkten also insbesondere eine diese beiden Punkte verbindende Kurve.

☞ **7.6.6 Satz.** *Sei $\Omega \subset \mathbb{R}^n$ ein Gebiet und seien $g_1, \dots, g_n : \Omega \to \mathbb{R}$ stetig. Dann ist die Differentialform*

$$\omega(x) = g_1(x)\,\mathrm{d}x_1 + \cdots + g_n(x)\,\mathrm{d}x_n \tag{7.6.6-A}$$

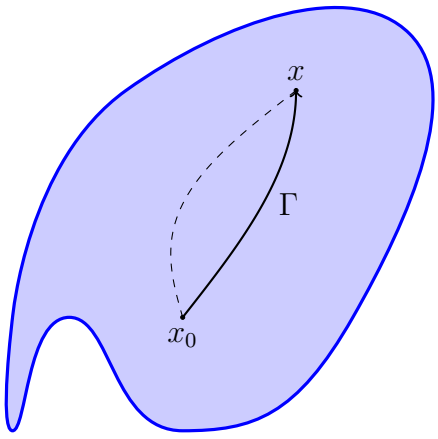
genau dann Differential einer stetig differenzierbaren Funktion $f : \Omega \to \mathbb{R}$, falls für jede geschlossene rektifizierbare[16] Kurve Γ in Ω das Kurvenintegral

$$\oint_\Gamma \omega = \oint_\Gamma \sum_{j=1}^n g_j(x)\,\mathrm{d}x_j = 0 \tag{7.6.6-B}$$

erfüllt. In diesem Fall gilt

$$f(x) - f(x_0) = \int_{x_0}^x \omega = \int_{x_0}^x \sum_{j=1}^n g_j\,\mathrm{d}x_j \tag{7.6.6-C}$$

für eine beliebige x_0 und x verbindende rektifizierbare Kurve als Integrationsweg.



[16] Polygone reichen für den angegebenen Beweis. Die Aussage für alle rektifizierbaren Kurven folgt dann aus der Konstruktion des Integrals.

Beweis. Gilt $\omega = \mathrm{d}f$ für eine Funktion $f : \Omega \to \mathbb{R}$, so folgt für jede x_0 mit x verbindende rektifizierbare Kurve Γ

$$f(x) - f(x_0) = \int_\Gamma \omega \tag{7.6.6-D}$$

und damit für jede geschlossene Kurve (die also einen Punkt x_0 mit sich selbst verbindet) (7.6.6-B). Zu zeigen ist nur ist die Rückrichtung. Angenommen, für jede gechlossene Kurve gilt (7.6.6-B). Dann bestimmt für fest gewähltes x_0 und eine beliebige x_0 mit x verbindende Kurve Γ

$$f(x) = \int_\Gamma \omega = \int_{x_0}^{x} \sum_{j=1}^{n} g_j \,\mathrm{d}x_j \tag{7.6.6-E}$$

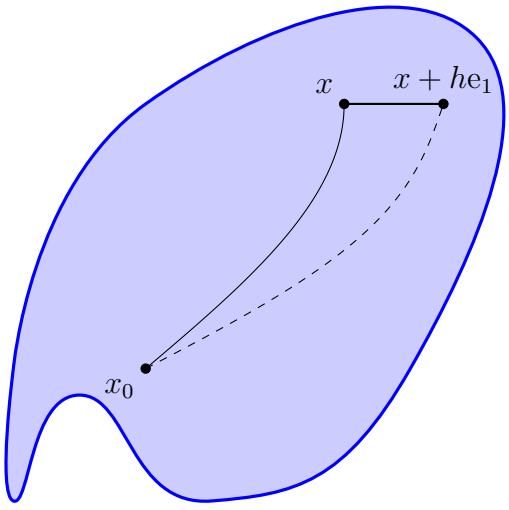
eine (nicht vom gewählten Integrationsweg) abhängende Funktion $f : \Omega \to \mathbb{R}$. Wir zeigen, dass diese stetig partiell differenzierbar ist. Es gilt für h klein genug, so dass die Verbindungsstrecke von x zu $x + h\mathrm{e}_j$ in Ω liegt,

$$\begin{aligned} f(x + h\mathrm{e}_j) - f(x) &= \int_x^{x+h\mathrm{e}_j} \omega = \int_0^h \sum_{i=1}^{n} g_i(x + \theta\mathrm{e}_j)\langle \mathrm{d}x_i, \mathrm{e}_j\rangle \,\mathrm{d}\theta \\ &= \int_0^h g_j(x + \theta\mathrm{e}_j)\,\mathrm{d}\theta = h g_j(x + \theta_h \mathrm{e}_j) \end{aligned} \tag{7.6.6-F}$$

unter Nutzung der Parametrisierung $\gamma(\theta) = x + \theta\mathrm{e}_j$, $\theta \in [0, h]$, mit $\dot\gamma(\theta) = \mathrm{e}_j$, und für ein geeignetes $\theta_h \in [0, h]$. Da die g_j als stetig vorausgesetzt sind, folgt

$$\lim_{h\to 0} \frac{f(x + h\mathrm{e}_j) - f(x)}{h} = g_j(x) \tag{7.6.6-G}$$

und f ist stetig partiell differenzierbar mit $\mathrm{d}f = \omega$.



□

7.6.7. Die Differentialschreibweise (7.6.4-C) ist flexibel genug, auch andere Koordinatensysteme in der Rechnung zuzulassen. Sind auf $\Omega \subset \mathbb{R}^n$ Koordinatenfunktionen

$y_1, \dots, y_n : \Omega \to \mathbb{R}$ gegeben, also stetig differenzierbare Funktionen $y_j = y_j(x)$ mit der Eigenschaft, dass die Differentiale $\mathrm{d}y_1, \dots, \mathrm{d}y_n$ in jedem Punkt $x \in \Omega$ linear unabhängig sind, so kann jede Pfaffsche Form $\omega : \Omega \to (\mathbb{R}^n)^*$ in dieser Basis als

$$\omega(x) = h_1(x)\,\mathrm{d}y_1(x) + \cdots + h_n(x)\,\mathrm{d}y_n(x) \tag{7.6.7-A}$$

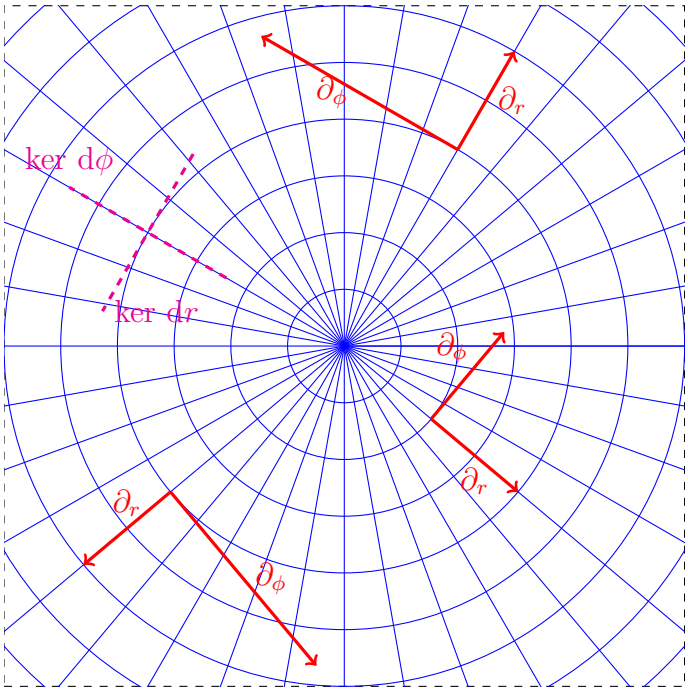
geschrieben werden. Es gilt dabei wiederum $h_j(x) = \langle \omega(x), \mathbf{v}_j(x)\rangle$ mit der zu $\mathrm{d}y_1(x), \dots, \mathrm{d}y_n(x)$ dualen Basis $\mathbf{v}_1(x), \dots, \mathbf{v}_n(x)$ des $\mathbb{R}^n$. Man beachte, dass diese Basen vom Punkt $x \in \Omega$ abhängen müssen.

Betrachtet man speziell Differentiale einer differenzierbaren Funktion $f : \Omega \to \mathbb{R}$ so gilt

$$\mathrm{d}f = \sum_{j=1}^{n} \langle f'(x), \mathbf{v}_j(x)\rangle\,\mathrm{d}y_j \tag{7.6.7-B}$$

und es bleibt, eine sinnvolle Interpretation der Koeffizienten $\langle f'(x), \mathbf{v}_j(x)\rangle$ anzugeben. Es handelt sich um Richtungsableitungen in Richtung des Vektors $\mathbf{v}_j$ und wir vereinbaren dafür die Notation

$$\boldsymbol{\partial}_{y_j} f(x) := \langle f'(x), \mathbf{v}_j(x)\rangle. \tag{7.6.7-C}$$



Nutzt man den lokalen Umkehrsatz und bestimmt zur aus den Koordinatenfunktionen zusammengesetzten Koordinatenabbildung $\Phi : x \mapsto y$ lokal die Inverse Φ^{-1}, so entspricht der gerade definierten partiellen Ableitung $\boldsymbol{\partial}_{y_j}$ die Ableitung

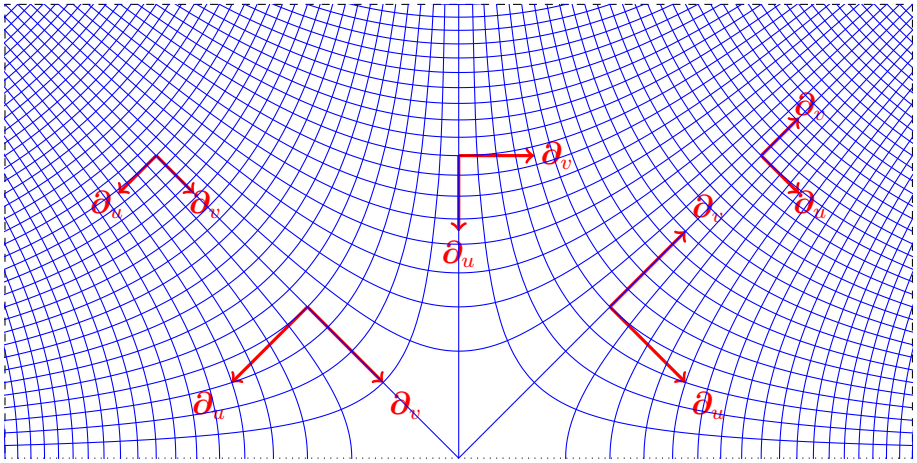
$$(\boldsymbol{\partial}_{y_j} f)(x) = \frac{\partial(f \circ \Phi^{-1})}{\partial y_j}(\Phi(x)). \tag{7.6.7-D}$$

Wir werden in der Notation nicht zwischen diesen Ableitungen unterscheiden und in Zukunft ∂_{y_j} für $\boldsymbol{\partial}_{y_j}$ schreiben. Ebenso ist es sinnvoll, sich die (partiellen) Ableitungen ∂_{y_j} als Vektoren vorzustellen. Dargestellt in obigem Bild sind die partiellen Ableitungen im Fall von Polarkoordinaten als ortsabhängige Basisvektoren.

⚒ **7.6.8 Beispiel.** Wir betrachten ein Beispiel eines Koordinatensystems. Auf der Menge $\Omega : \mathbb{R}\times\mathbb{R}_{>0}$ seien statt der kartesischen Koordinaten x und y die Koordinatenfunktionen

$$u = u(x,y) = \frac{x^2 - y^2}{2} \qquad \text{und} \qquad v = v(x,y) = xy \tag{7.6.8-A}$$

gegeben. Die Koordinatenlinien sind nachfolgend dargestellt:



Wir berechnen zuerst die Differentiale der Koordinatenfunktionen in kartesischen Koordinaten. Es gilt

$$\mathrm{d}u = x\,\mathrm{d}x - y\,\mathrm{d}y \qquad \text{und} \qquad \mathrm{d}v = y\,\mathrm{d}x + x\,\mathrm{d}y \tag{7.6.8-B}$$

Nutzt man diese als Basis von $(\mathbb{R}^2)^*$ zur Darstellung des Differentials im Punkt (x,y), so liefert

$$\mathrm{d}f = (\boldsymbol{\partial}_u f)\,\mathrm{d}u + (\boldsymbol{\partial}_v f)\,\mathrm{d}v = (\partial_x f)\,\mathrm{d}x + (\partial_y f)\,\mathrm{d}y \tag{7.6.8-C}$$

für differenzierbares $f : \Omega \to \mathbb{R}$ eine Darstellung der partiellen Ableitungen ∂_u und ∂_v in kartesischen Koordinaten. Einsetzen der Basisdarstellung liefert

$$(\boldsymbol{\partial}_u f)(x\,\mathrm{d}x - y\,\mathrm{d}y) + (\boldsymbol{\partial}_v f)(y\,\mathrm{d}x + x\,\mathrm{d}y) = (\partial_x f)\,\mathrm{d}x + (\partial_y f)\,\mathrm{d}y \tag{7.6.8-D}$$

und damit Koeffizientenvergleich das lineare Gleichungssystem

$$x\boldsymbol{\partial}_u f + y\boldsymbol{\partial}_v f = \partial_x f \qquad \text{und} \qquad -y\boldsymbol{\partial}_u f + x\boldsymbol{\partial}_v f = \partial_y f \tag{7.6.8-E}$$

und damit nach Lösung Umrechnungsformeln

$$\boldsymbol{\partial}_u f = \frac{x\partial_x f - y\partial_y f}{x^2+y^2}, \qquad \text{und} \qquad \boldsymbol{\partial}_v f = \frac{y\partial_x f + x\partial_y f}{x^2+y^2}. \tag{7.6.8-F}$$

für partielle Ableitungen in die neuen Koordinaten u und v.

7.7 Höhere Ableitungen und Taylorsche Formel

7.7.1. Wir beginnen wieder mit einem Einschub zur linearen Algebra.

✂ **Ergänzung** (Bilinearformen und quadratische Formen)**.** Sei V ein Vektorraum über $\mathbb{K}$. Eine *Bilinearform* auf V ist eine Abbildung

$$\gamma : V \times V \to \mathbb{K} \tag{7.7.1-A}$$

die in jeder Komponente linear ist, also für alle $x, y, z \in V$ und $\alpha \in \mathbb{K}$

$$\gamma(x + \alpha y, z) = \gamma(x, z) + \alpha\gamma(y, z) \qquad \text{und} \qquad \gamma(x, y + \alpha z) = \gamma(x, y) + \alpha\gamma(x, z) \tag{7.7.1-B}$$

erfüllt. Eine Bilinearform heißt *symmetrisch*, falls darüberhinaus

$$\gamma(x, y) = \gamma(y, x) \tag{7.7.1-C}$$

für alle $x, y \in V$ gilt.

Ist V endlich erzeugt und $\mathbf{v}_1, \dots \mathbf{v}_n$ eine Basis von V, so folgt aus der Bilinearität von γ

$$\begin{aligned} \gamma\left(\sum_{i=1}^{n} x_i\mathbf{v}_i, \sum_{j=1}^{n} y_j\mathbf{v}_j\right) &= \sum_{i=1}^{n}\sum_{j=1}^{n} y_i\gamma(\mathbf{v}_i, \mathbf{v}_j)x_j \\ &= \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}^{\top} \begin{bmatrix} \gamma(\mathbf{v}_1, \mathbf{v}_1) & \cdots & \gamma(\mathbf{v}_1, \mathbf{v}_n) \\ \vdots & & \vdots \\ \gamma(\mathbf{v}_n, \mathbf{v}_1) & \cdots & \gamma(\mathbf{v}_n, \mathbf{v}_n) \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \end{aligned} \tag{7.7.1-D}$$

und wir erhalten eine Beschreibung der Bilinearform γ durch die *Grammatrix*[17] mit Einträgen $\gamma(\mathbf{v}_i, \mathbf{v}_j)$. Umgekehrt liefert jede Matrix A durch $(x, y) \mapsto y^{\top}Ax$ eine Bilinearform auf $\mathbb{K}^n$. Die Bilinearform ist symmetrisch genau dann, wenn die zugehörige Grammatrix symmetrisch ist.

Ist $\gamma : V \times V \to \mathbb{K}$ eine Bilinearform, so bezeichnet

$$q_\gamma(v) := \gamma(v, v) \tag{7.7.1-E}$$

die zugeordnete *quadratische Form*. Symmetrische Bilinearformen sind durch die zugeordneten quadratischen Formen eindeutig bestimmt. Es gilt

$$\gamma(x, y) = \frac{1}{2}\left(\gamma(x, y) + \gamma(y, x)\right) = \frac{1}{2}\left(q_\gamma(x + y) - q_\gamma(x) - q_\gamma(y)\right). \tag{7.7.1-F}$$

✂ **Ergänzung.** Jede Bilinearform $\gamma : V \times V \to \mathbb{K}$ wird durch Einsetzen eines Vektors x zu einer linearen Abbildung $\gamma(x, \cdot) : V \to \mathbb{K}$, es gilt also $\gamma(x, \cdot) \in V^*$. Weiterhin ist die Zuordnung $V \ni x \mapsto \gamma(x, \cdot) \in V^*$ linear. Umgekehrt bestimmt jede lineare Abbildung $\varphi : V \to V^*$ eine Bilinearform $\gamma_\varphi(x, y) = \langle\varphi(x), y\rangle$. Damit kann man

$$\mathrm{Bil}(V) = \{\gamma : V \times V \to \mathbb{K} \mid \gamma \text{ bilinear}\} \quad \longleftrightarrow \quad \mathrm{Lin}(V, V^*) = \mathrm{Lin}(V, \mathrm{Lin}(V, \mathbb{R})) \tag{7.7.1-G}$$

identifizieren.

7.7.2. Ist V normierter Raum, so heißt eine bilineare Abbildung $\gamma : V \times V \to \mathbb{R}$ *beschränkt*, falls

$$|\gamma(x, y)| \le c\|x\|\|y\| \tag{7.7.2-A}$$

mit einer Konstanten c unabhängig von x und y gilt. Jeder beschränkten linearen Abbildung $V \mapsto \mathcal{L}(V, \mathbb{R})$ entspricht (durch Einsetzen) eine beschränkte Bilinearform.

Damit ist es sinnvoll, die zweite Ableitung einer Funktion $f : V \to \mathbb{R}$ als Bilinearform aufzufassen.

[17] Jørgen Pedersen Gram, 1850–1916

✍ **7.7.3 Definition** (Hesseform und -matrix). *Eine Funktion $f : V \supseteq \Omega \to \mathbb{R}$ heißt im Punkt $x_* \in \Omega$ zweimal differenzierbar, falls die Ableitung*

$$x \mapsto f'(x) \in \mathcal{L}(V, \mathbb{R}) \tag{7.7.3-A}$$

in einer Umgebung von x_* existiert und in x_* wiederum differenzierbar ist. Dann bezeichnet man die $f''(x_*) \in \mathcal{L}(V, \mathcal{L}(V, \mathbb{R}))$ zugeordnete Bilinearform als Hesseform

$$H_{f,x_*}(h_1, h_2) = \langle f''(x_*)h_1, h_2\rangle \tag{7.7.3-B}$$

der Funktion f im Punkt x_*.

Ist $V = \mathbb{R}^n$, so bezeichnet man die der Standardbasis zugeordnete Grammatrix der Hesseform als die Hessematrix *von f im Punkt x_*. Es gilt also*

$$\mathrm{H}_{f,x_*} = \begin{bmatrix} \partial_{x_1}\partial_{x_1} f & \partial_{x_2}\partial_{x_1} f & \cdots & \partial_{x_n}\partial_{x_1} f \\ \partial_{x_1}\partial_{x_2} f & \partial_{x_2}\partial_{x_2} f & \cdots & \partial_{x_n}\partial_{x_2} f \\ \vdots & \vdots & & \vdots \\ \partial_{x_1}\partial_{x_n} f & \partial_{x_2}\partial_{x_n} f & \cdots & \partial_{x_n}\partial_{x_n} f \end{bmatrix}. \tag{7.7.3-C}$$

☞ **7.7.4 Satz** (Taylorsche Formel). *Sei $f : V \supseteq \Omega \to \mathbb{R}$ stetig differenzierbar und im Punkt $x_* \in \Omega$ zweimal differenzierbar. Dann gilt*

$$f(x_* + h) = f(x_*) + \langle f'(x_*), h\rangle + \frac{1}{2}\langle f''(x_*)h, h\rangle + \mathbf{o}(\|h\|^2), \qquad h \to 0. \tag{7.7.4-A}$$

Beweis. Da f in x_* zweifach differenzierbar ist, gilt

$$f'(x_* + h) - f'(x_*) = f''(x_*)h + \mathbf{o}(\|h\|). \tag{7.7.4-B}$$

Da f' ebenso stetig ist, folgt daraus per Integration

$$\begin{aligned} f(x_* + h) - f(x_*) - \langle f'(x_*), h\rangle &= \Big[f(x_* + th) - \langle f'(x_*), th\rangle\Big]_{t=0}^{1} \\ &= \int_0^1 \langle f'(x_* + th) - f'(x_*), h\rangle \,\mathrm{d}t = \int_0^1 \langle f''(x_*)th, h\rangle \,\mathrm{d}t + \mathbf{o}(\|h\|)\|h\| \\ &= \frac{1}{2}\langle f''(x_*)h, h\rangle + \mathbf{o}(\|h\|^2) \end{aligned} \tag{7.7.4-C}$$

und damit die Behauptung. □

☞ **7.7.5 Satz.** *Sei $f : V \supseteq \Omega \to \mathbb{R}$ stetig differenzierbar und in $x_* \in \Omega$ zweimal differenzierbar. Dann ist die Hesseform H_{f,x_*} symmetrisch.*

Beweis. Wir führen die Hilfsfunktion

$$g(h_1, h_2) = f(x_* + h_1 + h_2) - f(x_* + h_1) - f(x_* + h_2) + f(x_*) \tag{7.7.5-A}$$

ein. Diese ist offenbar symmetrisch und es genügt zu zeigen, dass

$$g(h_1, h_2) - \langle f''(x_*)h_1, h_2\rangle = \mathbf{o}(\|h_1\|^2 + \|h_2\|^2) \tag{7.7.5-B}$$

gilt, da damit

$$\langle f''(x_*)h_1, h_2\rangle - \langle f''(x_*)h_2, h_1\rangle = \mathbf{o}(\|h_1\|^2 + \|h_2\|^2) \tag{7.7.5-C}$$

folgt und die linke Seite auf Grund der Bilinearität damit Null sein muss. Also ist die Hesseform symmetrisch.

Zum Beweis von (7.7.**5-B**). Es gilt

$$\begin{aligned}
&g(h_1,h_2) - \langle f''(x_*)h_1, h_2\rangle \\
&\quad = f(x_* + h_1 + h_2) - f(x_* + h_1) - f(x_* + h_2) + f(x_*) - \langle f''(x_*)h_1, h_2\rangle \\
&\quad = \Big[f(x_* + h_1 + th_2) - f(x_* + th_2) - \langle f'(x_* + h_1) - f'(x_*), th_2\rangle \Big]_{t=0}^{1} \\
&\qquad + \langle f'(x_* + h_1) - f'(x_*) - f''(x_*)h_1, h_2\rangle \\
&\quad = \int_0^1 \langle f'(x_* + h_1 + th_2) - f'(x_* + h_1) - f'(x_* + th_2) + f'(x_*), h_2\rangle \,\mathrm{d}t \\
&\qquad + \langle f'(x_* + h_1) - f'(x_*) - f''(x_*)h_1, h_2\rangle
\end{aligned} \tag{7.7.5-D}$$

und es genügt die einzelnen Terme genauer abzuschätzen. Da f in x_* zweifach differenzierbar ist, gilt

$$\begin{aligned}
f'(x_* + h_1 + th_2) - f'(x_*) &= f''(x_*)(h_1 + th_2) + \mathbf{o}(\|h_1 + th_2\|) \\
f'(x_*) - f'(x_* + h_1) &= -f''(x_*)h_1 + \mathbf{o}(\|h_1\|) \\
f'(x_*) - f'(x_* + th_2) &= -f''(x_*)th_2 + \mathbf{o}(\|th_2\|)
\end{aligned} \tag{7.7.5-E}$$

und in Summe damit für den Integranden

$$\big(\mathbf{o}(\|h_1 + th_2\|) + \mathbf{o}(\|h_1\|) + t\mathbf{o}(\|h_2\|)\big)\|h_2\| = \mathbf{o}(\|h_1\|^2 + \|h_2\|^2). \tag{7.7.5-F}$$

Für den verbleibenden Summanden ergibt die mittlere Zeile

$$\langle f'(x_* + h_1) - f'(x_*) - f''(x_*)h_1, h_2\rangle = \mathbf{o}(\|h_1\|)\|h_2\| = \mathbf{o}(\|h_1\|^2 + \|h_2\|^2). \tag{7.7.5-G}$$

Damit ist (7.7.**5-B**) gezeigt. □

☞ **7.7.6 Korollar** (Schwarz). *Angenommen, $f : \mathbb{R}^n \supseteq \Omega \to \mathbb{R}$ ist in einer Umgebung eines Punktes x zweifach partiell differenzierbar, es existieren also die partiellen Ableitungen*

$$\frac{\partial}{\partial x_i} f \qquad \textit{und} \qquad \frac{\partial}{\partial x_i}\frac{\partial}{\partial x_j} f \tag{7.7.6-A}$$

für $i, j = 1, \ldots, n$. Sind die zweiten partiellen Ableitungen in der Umgebung von x stetig, so gilt

$$\frac{\partial}{\partial x_i}\frac{\partial}{\partial x_j} f(x) = \frac{\partial}{\partial x_j}\frac{\partial}{\partial x_i} f(x). \tag{7.7.6-B}$$

Beweis. Die Stetigkeit der zweiten partiellen Ableitungen impliziert zweifache Differenzierbarkeit und die Aussage folgt. □

Wir geben zwei Beispiele basierend auf Funktionen $f : \mathbb{R}^2 \to \mathbb{R}$.

⚒ **7.7.7 Beispiel.** **(i)** Die Funktion $f(x,y) = x^2 \sin(xy)$ ist zweifach stetig differenzierbar. Weiter gilt

$$\begin{aligned} \partial_x f(x,y) &= 2x\sin(xy) + x^2 y\cos(xy), \\ \partial_y f(x,y) &= x^3\cos(xy) \end{aligned} \tag{7.7.7-A}$$

Wie zu erwarten gilt

$$\begin{aligned} \partial_y\partial_x f(x,y) &= \partial_y\big(2x\sin(xy) + x^2 y\cos(xy)\big) \\ &= 2x^2\cos(xy) + x^2\cos(xy) - x^3 y\sin(xy) \\ &= \partial_x x^3\cos(xy) = \partial_x\partial_y f(x,y). \end{aligned} \tag{7.7.7-B}$$

(ii) Die Voraussetzung der Stetigkeit der zweiten Ableitung ist notwendig. Wir betrachten dazu die Funktion

$$f(x,y) = \begin{cases} xy\frac{x^2-y^2}{x^2+y^2} & x^2+y^2 > 0, \\ 0 & x = y = 0. \end{cases} \tag{7.7.7-C}$$

Dann gilt für die partiellen Ableitungen

$$\partial_x f(0,y) = \begin{cases} -y, & y \neq 0, \\ \lim_{h\to 0}\frac{f(h,0)}{h} = 0, & y = 0, \end{cases} \tag{7.7.7-D}$$

und

$$\partial_y f(x,0) = \begin{cases} x, & x \neq 0, \\ \lim_{h\to 0}\frac{f(0,h)}{h} = 0, & x = 0, \end{cases} \tag{7.7.7-E}$$

und damit

$$\partial_y\partial_x f(0,0) = -1 \neq 1 = \partial_y\partial_x f(0,0). \tag{7.7.7-F}$$

Die Vertauschbarkeit der zweiten Ableitungen gilt für diese Funktion also nicht. Wir schauen uns die Funktion noch etwas genauer an. Die erste Ableitung $f'(x,y)$ ist im Ursprung stetig. Aus obiger Rechnung folgt $\partial_x f(0,0) = \partial_y f(0,0) = 0$. Weiter gilt für $x^2+y^2 > 0$

$$|\partial_x f(x,y)| = \left|\frac{x^4 y + 4x^2y^3 - y^5}{(x^2+y^2)^2}\right| \leq \frac{6(x^2+y^2)^{5/2}}{(x^2+y^2)^2} = 6\sqrt{x^2+y^2} \tag{7.7.7-G}$$

und

$$|\partial_y f(x,y)| = \left|\frac{xy^4 + 4x^3y^2 - x^5}{(x^2+y^2)^2}\right| \leq 6\sqrt{x^2+y^2} \tag{7.7.7-H}$$

und diese Abschätzungen streben für $x^2+y^2 \to 0$ gegen Null.

Im Kapitel 7.6 haben wir ein notwendiges Kriterium für lokale Extremstellen x_* einer differenzierbaren Funktion $f : \mathbb{R}^n \supseteq \Omega \to \mathbb{R}$ angegeben. In diesen kritischen Punkten muss $\mathrm{d}f(x_*) = 0$ gelten. Als nächstes formulieren wir ein *hinreichendes Kriterium* für lokale Extremstellen basierend auf der Hessematrix. Dieses ist analog zu den bekannten Kriterien im Eindimensionalen.

☞ **7.7.8 Satz** (Hinreichendes Kriterium für Extrema). *Sei $f : \mathbb{R}^n \supseteq \Omega \to \mathbb{R}$ zweifach differenzierbar und gelte $\mathrm{d}f(x_*) = 0$ für ein $x_* \in \Omega$. Ist dann die Hesseform positiv definit, gilt also*

$$\mathrm{H}_{f,x_*}(h,h) = \langle f''(x_*)h, h\rangle \geq c_0|h|^2 \qquad \text{für alle } h \in \mathbb{R}^n \tag{7.7.8-A}$$

mit einem $c_0 > 0$, so besitzt f in x_* ein lokales Minimum.

Beweis. Die Behauptung folgt direkt aus der Taylorschen Formel (7.7.4-A),

$$f(x_* + h) - f(x_*) = \frac{1}{2}\langle f''(x_*)h, h\rangle + \mathbf{o}(|h|^2) \geq \frac{c_0}{2}|h|^2 + \mathbf{o}(|h|^2) \tag{7.7.8-B}$$

für $h \to 0$. Für $|h|$ so klein, dass $|\mathbf{o}(|h|^2)| \leq \varepsilon|h|^2$ für ein $\varepsilon < c_0$ gilt, folgt damit

$$f(x_* + h) - f(x_*) \geq \frac{c_0 - \varepsilon}{2}|h|^2 > 0 \tag{7.7.8-C}$$

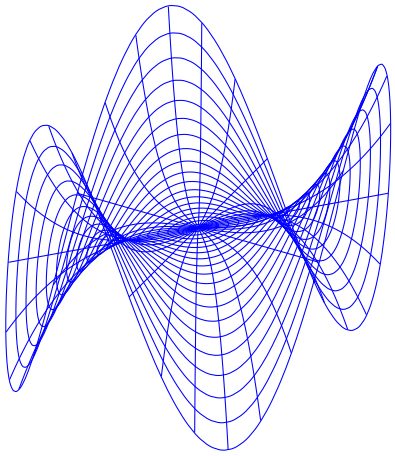
und in x_* liegt ein lokales Minimum vor. □

⚒ **7.7.9 Beispiel.** Das Kriterium ist hinreichend, aber genau wie im Eindimensionalen nicht notwendig. Für die Funktion $f : \mathbb{R}^2 \to \mathbb{R}$ mit

$$f(x,y) = x^4 + y^4 \tag{7.7.9-A}$$

liegt im Ursprung ein Minimum vor. Allerdings ist $\mathrm{H}_0 = 0$ nicht positiv definit.

Betrachtet man die oft als Affensattel bezeichnete Funktion



$$f(x,y) = x^3 - 3xy^2 \tag{7.7.9-B}$$

so liegt im Ursprung keine lokale Extremstelle vor, trotzdem sind Ableitung

$$\mathrm{d}f = (3x^2 - 3y^2)\,\mathrm{d}x - 6xy\,\mathrm{d}y \tag{7.7.9-C}$$

und Hessematrix

$$\mathrm{H}_f = \begin{bmatrix} 6x & -6y \\ -6y & -6x \end{bmatrix} \tag{7.7.9-D}$$

im Ursprung Null.

Wir geben Kriterien für die positive Definitheit der Hesseform / der Hessematrix.

☞ **7.7.10 Lemma.** *Sei $A \in \mathbb{R}^{n\times n}$ symmetrisch und $\gamma_A(x,y) = y^\top Ax$ zugeordnete Bilinearform auf dem $\mathbb{R}^n$. Dann gilt*

$$\gamma_A(x,x) = x^\top Ax \geq c_0|x|^2, \qquad x \neq 0, \tag{7.7.10-A}$$

genau dann, wenn

$$\det(A - \lambda) = 0 \qquad \Longrightarrow \qquad \lambda > 0 \tag{7.7.10-B}$$

gilt.

Beweis. Wir betrachten den Rayleigh[18]-Quotienten

$$g_A(x) = \frac{x^\top Ax}{x^\top x}, \qquad x \neq 0, \tag{7.7.10-C}$$

als Hilfsfunktion auf $\mathbb{R}^n \setminus 0$. Diese ist homogen vom Grad 0, es gilt also für $r > 0$ stets $g_A(rx) = g_A(x)$. Wegen

$$|g_A(x)| \leq \frac{|x|\|A\||x|}{|x|^2} = \|A\| \tag{7.7.10-D}$$

für $x \neq 0$ ist diese beschränkt, nimmt also auf der kompakten Menge

$$\{x \in \mathbb{R}^n \mid \frac{1}{2} \leq |x| \leq 2\} \tag{7.7.10-E}$$

Minimum und Maximum an. Auf Grund der Homogenität ist g_A entlang von Strahlen durch den Ursprung konstant, Maxima und Minima werden also auch im Inneren angenommen. Für solche Stellen x_* gilt das notwendige Kriterium $g'_A(x_*) = 0$ und wegen

$$\begin{aligned}
\frac{1}{(x_*+h)^\top(x_*+h)} &= \frac{1}{x_*^\top x_*}\frac{1}{1+2\frac{h^\top x_*}{x_*^\top x_*}+\frac{h^\top h}{x_*^\top x_*}} \\
&= \frac{1}{x_*^\top x_*}\left(1-2\frac{h^\top x_*}{x_*^\top x_*}+\mathbf{O}(|h|^2)\right) \\
g_A(x_*+h)-g_A(x_*) &= \frac{(x_*+h)^\top A(x_*+h)}{(x_*+h)^\top(x_*+h)} - \frac{x_*^\top Ax_*}{x_*^\top x_*} \\
&= 2\left(\frac{h^\top Ax_*}{x_*^\top x_*} - \frac{x_*^\top Ax_*}{x_*^\top x_*}\frac{h^\top x_*}{x_*^\top x_*}\right)+\mathbf{O}(|h|^2)
\end{aligned} \tag{7.7.10-F}$$

und mit $g_A(x_*) = \lambda$ somit

$$0 = \langle g'_A(x_*), h\rangle = 2\frac{h^\top Ax_*}{x_*^\top x_*} - 2\frac{x_*^\top Ax_*}{x_*^\top x_*}\frac{h^\top x_*}{x_*^\top x_*} = \frac{2}{|x_*|^2}\,h^\top(Ax_* - \lambda x_*) \tag{7.7.10-G}$$

für alle $h \in \mathbb{R}^n$. Damit muss notwendigerweise $Ax_* = \lambda x_*$ gelten. Kritische Punkte müssen also Eigenvektoren zu A und die zugehörigen Werte Eigenwerte sein. Umgekehrt gilt für jedes $x_* \neq 0$ mit $Ax_* = \lambda x_*$ schon $g_A(x_*) = \lambda$. Damit ist (7.7.**10-A**) äquivalent dazu, dass alle (reellen) Eigenwerte der symmetrischen Matrix A positiv sind und das wiederum ist äquivalent zur Forderung (7.7.**10-B**). □

Wir ziehen eine wichtige Folgerung aus dem Beweis.

☞ **7.7.11 Korollar.** *Sei $A \in \mathbb{R}^{n\times n}$ symmetrisch.*

(i) *Dann gibt es $\lambda \in \mathbb{R}$ mit $\ker(A-\lambda) \neq \{0\}$ und damit $\det(A-\lambda) = 0$.*

(ii) *Erfüllt $\mathbf{v}_1 \in \mathbb{R}^n \setminus \{0\}$ die Eigenwertgleichung $A\mathbf{v}_1 = \lambda_1\mathbf{v}_1$ zu einem $\lambda_1 \in \mathbb{R}$, so ist der Unterraum*

$$\{x \in \mathbb{R}^n \mid \mathbf{v}_1^\top x = 0\} =: \{\mathbf{v}_1\}^\perp \tag{7.7.11-A}$$

unter A invariant.

[18] John William Strutt, 3. Baron Rayleigh, 1842–1919

(iii) *Das Polynom* $\chi_A(\lambda) = \det(A - \lambda) = 0$ *zerfällt in reelle Linearfaktoren und es gibt eine Basis* $\mathbf{v}_1, \dots, \mathbf{v}_n$ *bestehend aus Eigenvektoren zu* A *mit* $\mathbf{v}_i^\top \mathbf{v}_j = \delta_{i,j}$.

Beweis. **(i)** folgt aus dem obigen Beweis. **(ii)** folgt, da für jedes $x \in \mathbb{R}^n$ mit $\mathbf{v}_1^\top x = 0$ wegen der Symmetrie $A = A^\top$

$$\mathbf{v}_1^\top (Ax) = (A\mathbf{v}_1)^\top x = \lambda_1 \mathbf{v}_1^\top x = 0 \tag{7.7.11-B}$$

gilt. Damit lässt A den Unterraum $\{\mathbf{v}_1\}^\top$ invariant. **(iii)** folgt rekursiv. Mit (i) finden wir einen ersten Eigenvektor $\mathbf{v}_1$. Da A nach (ii) auch $\{\mathbf{v}_1\}^\perp \to \{\mathbf{v}_1\}^\perp$ abbildet, kann der erste Schritt auch direkt auf $\gamma_A(x, x) = x^\top Ax$ für $x \in \{\mathbf{v}_1\}^\perp$ angewandt werden und wir erhalten als Kriterium für einen kritischen Punkt $x_* \in \{\mathbf{v}_1\}^\perp$ von g_A

$$h^\top (Ax_* - \lambda x_*) = 0, \qquad \text{für alle } h_* \in \{\mathbf{v}_1\}^\perp \tag{7.7.11-C}$$

und da ebenso $Ax_* \in \{\mathbf{v}_1\}^\perp$ gilt, ist dies sogar für alle $h \in \mathbb{R}^n$ erfüllt. Damit finden wir einen weiteren Eigenvektor $\mathbf{v}_2 \in \{\mathbf{v}_1\}^\perp$ zu einem $\lambda_2 \in \mathbb{R}$.

Nach n Schritten haben wir damit eine Basis $\mathbf{v}_1, \dots, \mathbf{v}_n$ mit $A\mathbf{v}_j = \lambda_j \mathbf{v}_j$ und $\mathbf{v}_i^\top \mathbf{v}_j = 0$ für $i \neq j$ konstruiert, jedes auftretende λ_j ist Nullstelle von $\det(A - \lambda)$ gezählt mit der entsprechenden Vielfachheit. Durch Umskalieren kann man $|\mathbf{v}_j| = 1$ erreichen und die Aussage ist gezeigt. □

Insbesondere gilt in der gerade konstruierten Basis $\mathbf{v}_1, \dots, \mathbf{v}_n$

$$\gamma_A(x, y) = \gamma_A\left(\sum_{i=1}^n \xi_i \mathbf{v}_i, \sum_{i=1}^n \eta_j \mathbf{v}_j\right) = \sum_{j=1}^n \lambda_j \xi_j \eta_j. \tag{7.7.11-D}$$

Das Polynom $\chi_A(\lambda) = \det(A - \lambda)$ wird als *charakteristisches Polynom* von A bezeichnet. Nach dem gerade gezeigten, zerfällt es für symmetrische Matrizen $A \in \mathbb{R}^{n\times n}$ in Linearfaktoren zu Eigenwerten λ_j, $j = 1, \dots, n$. Es gilt damit

$$\det A = \prod_{j=1}^n \lambda_j, \qquad \text{und} \qquad \operatorname{trace} A = \sum_{j=1}^n \lambda_j \tag{7.7.11-E}$$

als Konsequenz des Wurzelsatzes von Vieta. Wir geben ein weiteres Kriterium für die positive Definitheit einer symmetrischen Matrix A. Dieses kommt ohne die Bestimmung der Eigenwerte aus.

☞ **7.7.12 Satz** (Sylvester[19]-Kriterium oder Hurwitz[20]-Kriterium). *Sei* $A = (a_{i,j}) \in \mathbb{R}^{n\times n}$ *symmetrisch. Dann sind äquivalent*

(i) *A ist positiv definit, es gilt also* (7.7.**10-A**)

(ii) *die Determinanten der rechten oberen Matrixabschnitte*

$$a_{11} > 0, \quad \det \begin{bmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{bmatrix} > 0, \quad \dots, \quad \det \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{12} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{nn} \end{bmatrix} > 0 \tag{7.7.12-A}$$

[19] James Joseph Sylvester, 1814–1897
[20] Adolf Hurwitz, 1859–1919

sind alle positiv.

Beweis. Die Determinante einer Matrix ist gerade das Produkt der Eigenwerte. Setzt man in (7.7.**10-A**) nur Vektoren aus der linearen Hülle $\operatorname{span}\{\mathrm{e}_1, \ldots, \mathrm{e}_k\}$ ein, so impliziert dies die Positivität der Determinante der oberen $k \times k$-Untermatrix. Damit impliziert **(i)** die Bedingung **(ii)**.

Es bleibt das eigentliche Kriterium, also die Implikation von **(ii)** nach **(i)** zu zeigen. Dazu nutzen wir Induktion über die Dimension n. Für $n = 1$ ist nichts zu zeigen. Angenommen, die Aussage wurde schon für $k \times k$-Matrizen mit $k < n$ gezeigt und sei $A \in \mathbb{R}^{n\times n}$ symmetrisch.

Wir wissen, dass die Matrix A wegen $\det A = \prod_{j=1}^{n} \lambda_j > 0$ nur eine gerade Anzahl negativer Eigenwerte besitzen kann. Angenommen, es gibt zwei linear unabhängige Eigenvektoren $\mathbf{u}$ und $\mathbf{v}$ zu negativen Eigenwerten mit $\mathbf{u}^\top \mathbf{v} = 0$. Dann gilt $\mathbf{u}^\top A\mathbf{u} < 0$ und $\mathbf{v}^\top A\mathbf{v} < 0$. Weiter ist für $\mathbf{w} = v_n\mathbf{u} - u_n\mathbf{v}$ (mit den letzten Einträgen u_n und v_n der Vektoren $\mathbf{u}$ und $\mathbf{v}$) der letzte Eintrag Null und damit nach Induktionsvoraussetzung

$$0 \le \mathbf{w}A\mathbf{w} = (v_n\mathbf{u} - u_n\mathbf{v})^\top A(v_n\mathbf{u} - u_n\mathbf{v}) = v_n^2\mathbf{u}^\top A\mathbf{u} + u_n^2\mathbf{v}^\top A\mathbf{v} \le 0. \qquad (7.7.\textbf{12-B})$$

Also folgt $\mathbf{w} = 0$ und damit $\mathbf{u}, \mathbf{v} \in \mathbb{R}^{n-1} \times \{0\}$. Das ist aber ebenso ein Widerspruch. □

7.7.13 (Sattelpunkte)**.** Ist x_* ein kritischer Punkt einer zweifach differenzierbaren Funktion $f : \mathbb{R}^n \supseteq \Omega \to \mathbb{R}$, für den die Hessematrix H_{f,x_*} sowohl positive als auch negative Eigenwerte besitzt, so gibt es Unterräume $V_-, V_+ \subset \mathbb{R}^n$, so dass

$$f(x_* + h) > f(x_*), \qquad \text{für } h \in V_+ \setminus \{0\} \text{ klein genug}, \qquad (7.7.\textbf{13-A})$$

und

$$f(x_* + h) < f(x_*), \qquad \text{für } h \in V_- \setminus \{0\} \text{ klein genug}. \qquad (7.7.\textbf{13-B})$$

In diesem Fall spricht man von einem *Sattelpunkt*. Sind alle Eigenwerte negativ, so liegt ein lokales Maximum vor. Beides folgt aus dem hinreichenden Kriterium aus Satz 7.7.8 für lokale Minima angewandt auf die Einschränkung $h \mapsto \pm f(x_* + h)$ für $h \in V_\pm$ für Eigenräume V_+ zu positiven Eigenwerten beziehungsweise Eigenräume V_- zu negativen Eigenwerten.

7.7.14. Für höhere partielle Ableitungen hinreichend oft differenzierbarer Funktionen $f : \mathbb{R}^n \supseteq \Omega \to \mathbb{R}$ bietet sich die Multiindexschreibweise an. Da die Reihenfolge partieller Ableitungen nach dem Satz von Schwarz keine Rolle spielt, ist es sinnvoll nur die einzelnen Ableitungsordnungen zu zählen. Wir schreiben also für $\alpha \in \mathbb{N}_0^n$

$$\partial_x^\alpha := \prod_{i=1}^{n} \left(\frac{\partial}{\partial x_i}\right)^{\alpha_i}. \qquad (7.7.\textbf{14-A})$$

Als Folgerung der Produktregel ergibt sich die *Leibnizregel*

$$\partial_x^\alpha(fg) = \sum_{\beta \preceq \alpha} \binom{\alpha}{\beta} \left(\partial_x^\beta f\right) \left(\partial_x^{\alpha-\beta} g\right), \qquad (7.7.\textbf{14-B})$$

wobei wir die Notation

$$\beta \preceq \alpha \qquad :\Longleftrightarrow \qquad \forall_{i=1,\dots,n} \quad \beta_i \leq \alpha_i \tag{7.7.14-C}$$

verwenden und den verallgemeinerten *Binomialkoeffizienten*

$$\binom{\alpha}{\beta} = \frac{\alpha!}{\beta!(\alpha-\beta)!}, \qquad \alpha! = \prod_{i=1}^{n} \alpha_i!, \tag{7.7.14-D}$$

definieren. Den Beweis von (7.7.**14-B**) überlassen wir als Übung.

✂ **Ergänzung.** Formel (7.7.**14-B**) erinnert an den verallgemeinerten binomischen Lehrsatz

$$(x+y)^\alpha = \sum_{\beta \preceq \alpha} \binom{\alpha}{\beta} x^\beta y^{\alpha-\beta} \tag{7.7.14-E}$$

für $x, y \in \mathbb{R}^n$, der zum Rechnen mit Polynomen in mehreren Variablen nützlich ist.

In Satz 7.7.4 haben wir mit genauer Ableitungsordnung die Taylorsche Formel für quadratische Approximation bewiesen. Das Resultat gilt in allgemeinen normierten Räumen. Existieren mehr Ableitungen, so ergeben sich bessere Approximationen. Allerdings wird die Notation in allgemeinen normierten Räumen dabei komplizierter. Wir beschränken uns auf den Fall des $\mathbb{R}^n$, nutzen Multiindexnotation und setzen so viel Glattheit voraus, dass die eindimensionale Taylorsche Formel mit Restglied direkt ein entsprechendes höherdimensionales Resultat liefert.

☞ **7.7.15 Satz** (Taylorsche Formel höherer Ordnung)**.** *Sei $f : \mathbb{R}^n \supseteq \Omega \to \mathbb{R}$ in einer Umgebung von $x_* \in \Omega$ $k+1$-fach stetig differenzierbar. Dann gilt*

$$f(x_*+h) = \sum_{|\alpha| \leq k} \frac{\partial^\alpha f(x_*)}{\alpha!} h^\alpha + \mathbf{O}(|h|^{k+1}), \qquad h \to 0. \tag{7.7.15-A}$$

Beweis. Sei $r > 0$ so klein, dass f in $B_{2r}(x_*)$ $k+1$-fach stetig differenzierbar ist. Dann betrachten wir die Funktion

$$\mathbb{R} \ni t \mapsto f(x_* + th) \in \mathbb{R} \tag{7.7.15-B}$$

für $|h| < r$ und $t \in [0,1]$ und nutzen die eindimensionale Taylorsche Formel. Dazu berechnen wir zuerst die Ableitungen

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t} f(x_*+th) &= \langle f'(x_*+th), h\rangle = \sum_{i=1}^{n} (\partial_{x_i} f)(x_*+th) h_i \\ \frac{\mathrm{d}^2}{\mathrm{d}t^2} f(x_*+th) &= \sum_{j=1}^{n}\sum_{i=1}^{n} (\partial_{x_j}\partial_{x_i} f)(x_*+th) h_i h_j = \sum_{|\alpha|=2} \frac{2}{\alpha_1!\alpha_2!} (\partial^\alpha f)(x_*+th) h^\alpha \end{aligned} \tag{7.7.15-C}$$

und allgemein

$$\frac{\mathrm{d}^l}{\mathrm{d}t^l} f(x_*+th) = \sum_{|\alpha|=l} \frac{l!}{\alpha!} (\partial^\alpha f)(x_*+th) h^\alpha, \tag{7.7.15-D}$$

wobei $\frac{l!}{\alpha!}$ zählt, wie oft die Ableitung $\partial^\alpha f$ auftritt. Mit der Taylorschen Formel und dem Entwicklungspunkt $t = 0$ folgt

$$\begin{aligned} f(x_* + th) &= f(x_*) + \sum_{l=1}^{k} \frac{t^l}{l!} \left(\sum_{|\alpha|=l} \frac{l!}{\alpha!} (\partial^\alpha f)(x_*) h^\alpha \right) \\ &\quad - \frac{t^{k+1}}{k!} \int_0^1 \theta^k \sum_{|\alpha|=k+1} \frac{(k+1)!}{\alpha!} (\partial^\alpha f)(x_* + (1-\theta)th) h^\alpha \, \mathrm{d}\theta \\ &= \sum_{|\alpha|\le k} \frac{\partial^\alpha f(x_*)}{\alpha!} h^\alpha t^{|\alpha|} \\ &\quad - t^{k+1} \sum_{|\alpha|=k+1} \frac{k+1}{\alpha!} h^\alpha \int_0^1 \theta^k (\partial^\alpha f)(x_* + (1-\theta)th) \, \mathrm{d}\theta \end{aligned} \tag{7.7.15-E}$$

unter Nutzung des Integralrestglieds (5.4.2-A). Es genügt, $t = 1$ zu setzen den Integralrest abzuschätzen. Da die Ableitungen $\partial^\alpha f$ in einer Umgebung von x_* stetig und damit insbesondere beschränkt sind, gilt

$$\left| \int_0^1 \theta^k (\partial^\alpha f)(x_* + (1-\theta)h) \, \mathrm{d}\theta \right| \le \max_{x \in B_r(x_*)} \max_{|\alpha|=k+1} |(\partial^\alpha f)(x)| \int_0^1 \theta^k \, \mathrm{d}\theta \tag{7.7.15-F}$$

und die Behauptung folgt. □

☞ **7.7.16 Korollar.** *Unter den Voraussetzungen des vorherigen Satzes und mit*

$$M = \max_{x \in B_r(x_*)} \max_{|\alpha|=k+1} |(\partial^\alpha f)(x)| \tag{7.7.16-A}$$

gilt für $|h| \le r$ die Abschätzung

$$\left| f(x_* + h) - \sum_{|\alpha|\le k} \frac{\partial^\alpha f(x_*)}{\alpha!} h^\alpha \right| \le M \left| \sum_{|\alpha|=k+1} \frac{h^\alpha}{\alpha!} \right| \le M \frac{n^{k+1} r^{k+1}}{(k+1)!}. \tag{7.7.16-B}$$

Beweis. Wegen $|h_j| \le |h| \le r$ gilt stets $|h^\alpha| \le r^{|\alpha|} = r^{k+1}$. Es bleibt die verbleibende Summe über die Multiindizes zu berechnen. Dazu nutzen wir Induktion über n und zeigen

$$\sum_{|\alpha|=m} \frac{m!}{\alpha!} = n^m, \qquad \alpha \in \mathbb{N}_0^n. \tag{7.7.16-C}$$

Für $n = 1$ gilt $\alpha = \alpha_1 = |\alpha| = m$ und damit enthält die Summe genau einen Summanden, nämlich 1. Im Induktionsschritt nutzen wir $\alpha = (\alpha_1, \ldots, \alpha_n, \alpha_{n+1}) = (\alpha', \alpha_{n+1})$ mit $|\alpha'| = m - \alpha_{n+1}$ und erhalten damit

$$\begin{aligned} \sum_{|\alpha|=m} \frac{m!}{\alpha!} &= \sum_{\alpha_{n+1}=0}^{m} \frac{m!}{(m-\alpha_{n+1})!\alpha_{n+1}!} \underbrace{\sum_{|\alpha'|=m-\alpha_{n+1}} \frac{(m-\alpha_{n+1})!}{\alpha'!}}_{=n^{m-\alpha_{n+1}}} \\ &= \sum_{\alpha_{n+1}=0}^{m} \binom{m}{\alpha_{n+1}} 1^{\alpha_{n+1}} n^{m-\alpha_{n+1}} = (n+1)^m. \end{aligned} \tag{7.7.16-D}$$

unter Ausnutzung des binomischen Satzes. □

✂ **Ergänzung.** Die gerade gezeigte Summenformel (7.7.**16-C**) hängt eng mit der Verallgemeinerung des binomischen Lehrsatzes zu einem multinomischen Satz zusammen. Es gilt

$$\left(\sum_{i=1}^{n} x_i\right)^m = \sum_{|\alpha|=m} \frac{m!}{\alpha!} x^\alpha \tag{7.7.16-E}$$

und mit $x_i = 1$ folgt (7.7.**16-C**). Die Quotienten

$$\frac{m!}{\alpha!} = \frac{m!}{\alpha_1! \cdots \alpha_n!} \qquad \text{mit } \alpha_1 + \cdots + \alpha_n = m \tag{7.7.16-F}$$

werden mitunter als *Multinomialkoeffizienten* bezeichnet. Während Binomialkoeffizienten $\binom{m}{k}$ zählen, auf wie viele Weisen k Elemente aus m ausgewählt werden können, zählen Multinomialkoeffizienten auf wie viele Weisen eine m-elementige Menge in Gruppen der Größen $\alpha_1, \alpha_2, \ldots, \alpha_n$ partitioniert werden kann. Genau dies haben wir bereits in Formel (7.7.**15-D**) genutzt.

A Anhang

A.1 Koordinatensysteme

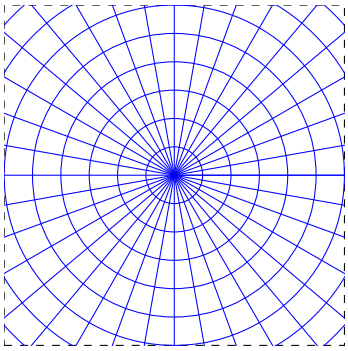
✂ **Ergänzung** (Polarkoordinaten)**.** Punkte der Ebene $\mathbb{R}^2$ kann man in *Polarkoordinaten* als

$$x = r\cos\theta, \qquad y = r\sin\theta \tag{A.1.1-A}$$

beschreiben. Dann gilt für die Differentiale der kartesischen Koordinaten

$$\mathrm{d}x = \cos\theta\,\mathrm{d}r - r\sin\theta\,\mathrm{d}\theta, \qquad \mathrm{d}y = \sin\theta\,\mathrm{d}r + r\cos\theta\,\mathrm{d}\theta. \tag{A.1.1-B}$$

Die Differentiale $\mathrm{d}x$ und $\mathrm{d}y$ haben wir als Basisvektoren des Raumes der linearen Abbildungen $(\mathbb{R}^2)^*$ verwendet. In Polarkoordinaten ist es sinnvoller, $\mathrm{d}r$ und $\mathrm{d}\theta$ zu verwenden. Gleichung (A.1.**1-B**) bestimmt den (vom betrachteten Punkt in der Ebene abhängigen und auch nur außerhalb des Ursprungs definierten) Basiswechsel.



Ist nun $f : \mathbb{R}^2 \to \mathbb{R}$ eine Funktion, so ist deren Differential

$$\mathrm{d}f = \partial_x f\,\mathrm{d}x + \partial_y f\,\mathrm{d}y \tag{A.1.1-C}$$

lokal in der Basis $\mathrm{d}x$, $\mathrm{d}y$ dargestellt. Die partiellen Ableitungen sind also die Koeffizienten in dieser Basisdarstellung. Damit kann die Transformation als Basiswechsel direkt hingeschrieben werden, es gilt

$$\mathrm{d}f = \big((\partial_x f)\cos\theta + (\partial_y f)\sin\theta\big)\,\mathrm{d}r + \big((\partial_y f) r\cos\theta - (\partial_x f) r\sin\theta\big)\,\mathrm{d}\theta \tag{A.1.1-D}$$

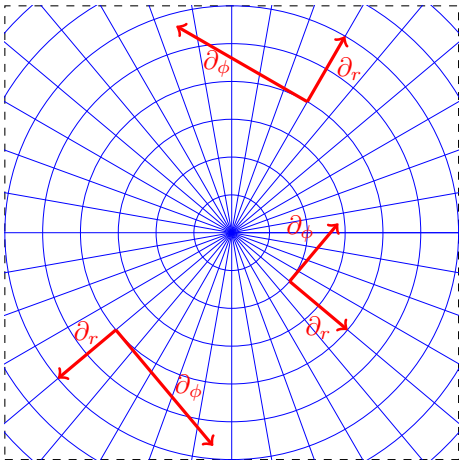
und damit

$$\begin{aligned} \partial_r f &= (\partial_x f)\cos\theta + (\partial_y f)\sin\theta \\ \partial_\theta f &= (\partial_y f)\, r\cos\theta - (\partial_x f)\, r\sin\theta. \end{aligned} \tag{A.1.1-E}$$

Dieses lineare Gleichungssystem besitzt für $r \neq 0$ (da $\mathrm{d}r$ und $\mathrm{d}\theta$ linear unabhängig sind) eine eindeutige Lösung und bestimmt

$$\begin{aligned} \partial_x f &= \cos\theta\,\partial_r f - \frac{\sin\theta}{r}\,\partial_\theta f = \left(\cos\theta\,\partial_r - \frac{\sin\theta}{r}\,\partial_\theta\right) f, \\ \partial_y f &= \sin\theta\,\partial_r f + \frac{\cos\theta}{r}\,\partial_\theta f = \left(\sin\theta\,\partial_r + \frac{\cos\theta}{r}\,\partial_\theta\right) f. \end{aligned} \tag{A.1.1-F}$$

Hier hätte man auch direkt mit der Kettenregel vorgehen und die Koordinatentransformation (7.3.**14-A**) differenzieren können. Die gerade bestimmten Koeffizienten ergeben sich aus der Transponierten der Jacobimatrix (7.3.**14-D**) geschrieben in Polarkoordinaten. Die Ableitungsoperatoren ∂_r und ∂_ϕ kann man sich als zu $\mathrm{d}r$ und $\mathrm{d}\phi$ duale Basis des $\mathbb{R}^2$ vorstellen, das hilft in Bildern und zur Veranschaulichung.[1]



Will man höhere partielle Ableitungen in Polarkoordinaten ausdrücken, so muss man die entsprechenden Formeln für erste partielle Ableitungen nutzen und die Produktregel anwenden. Es gilt

$$\begin{aligned}
\partial_x^2 f &= \left(\cos\theta\,\partial_r - \frac{\sin\theta}{r}\,\partial_\theta\right)\left(\cos\theta\,\partial_r - \frac{\sin\theta}{r}\,\partial_\theta\right) f \\
&= \Bigg(\cos^2\theta\,\partial_r^2 + \frac{\cos\theta\sin\theta}{r^2}\partial_\theta - \frac{\cos\theta\sin\theta}{r}\partial_r\partial_\theta \\
&\qquad + \frac{\sin^2\theta}{r}\partial_r - \frac{\sin\theta\cos\theta}{r}\partial_\theta\partial_r - \frac{\sin\theta\cos\theta}{r^2}\partial_\theta + \frac{\sin^2\theta}{r^2}\partial_\theta^2\Bigg) f \\
&= \left(\cos^2\theta\,\partial_r^2 + \frac{\sin^2\theta}{r}\partial_r - \frac{\sin(2\theta)}{r}\partial_r\partial_\theta + \frac{\sin^2\theta}{r^2}\partial_\theta^2\right) f
\end{aligned} \tag{A.1.**1-H**}$$

und entsprechend

$$\begin{aligned}
\partial_y^2 f &= \left(\sin\theta\,\partial_r + \frac{\cos\theta}{r}\,\partial_\theta\right)\left(\sin\theta\,\partial_r + \frac{\cos\theta}{r}\,\partial_\theta\right) f \\
&= \left(\sin^2\theta\,\partial_r^2 + \frac{\cos^2\theta}{r}\partial_r + \frac{\sin(2\theta)}{r}\partial_r\partial_\theta + \frac{\cos^2\theta}{r^2}\partial_\theta^2\right) f,
\end{aligned} \tag{A.1.**1-I**}$$

sowie

$$\begin{aligned}
2\partial_x\partial_y f &= 2\left(\cos\theta\,\partial_r - \frac{\sin\theta}{r}\,\partial_\theta\right)\left(\sin\theta\,\partial_r + \frac{\cos\theta}{r}\,\partial_\theta\right) f \\
&= 2\Bigg(\cos\theta\sin\theta\,\partial_r^2 - \frac{\cos^2\theta}{r^2}\partial_\theta + \frac{\cos^2\theta}{r}\partial_r\partial_\theta \\
&\qquad - \frac{\sin\theta\cos\theta}{r}\partial_r - \frac{\sin^2\theta}{r}\partial_r\partial_\theta + \frac{\sin^2\theta}{r^2}\partial_\theta - \frac{\sin\theta\cos\theta}{r^2}\partial_\theta^2\Bigg) f \\
&= \left(\sin(2\theta)\,\partial_r^2 - \frac{\sin(2\theta)}{r}\partial_r - \frac{2\cos(2\theta)}{r}\partial_\theta + \frac{2\cos(2\theta)}{r}\partial_r\partial_\theta - \frac{\sin(2\theta)}{r^2}\partial_\theta^2\right) f.
\end{aligned} \tag{A.1.**1-J**}$$

[1] Als Richtungsableitungen sind sie natürlich immer von der Form

$$\lim_{t\to 0}\frac{1}{t}\left(f(x_* + tv) - f(x_*)\right) \tag{A.1.**1-G**}$$

für einen Vektor v. Genau das wird hier genutzt. Die Dualtität der Basis folgt aus der Darstellung des Differentials mit partiellen Ableitungen als Koeffizienten.

Insbesondere gilt für den Laplaceoperator geschrieben in Polarkoordinaten

$$\Delta f := \partial_x^2 f + \partial_y^2 f = \partial_r^2 f + \frac{1}{r}\partial_r f + \frac{1}{r^2}\partial_\theta^2 f. \tag{A.1.1-K}$$

✂ **Ergänzung** (Kugelkoordinaten)**.** Außerhalb der z-Achse lassen sich Punkte im $\mathbb{R}^3$ durch *Kugelkoordinaten*

$$x = r\cos\phi\cos\psi, \qquad y = r\sin\phi\cos\psi, \qquad z = r\sin\psi \tag{A.1.2-A}$$

beschreiben. Dabei ist $r > 0$ und die Winkel $\phi \in [0, 2\pi)$ und $\psi \in (-\frac{\pi}{2}, \frac{\pi}{2})$ entsprechen auf einer Kugeloberfläche gerade Längen- und Breitengraden.

Wir folgen der eben in Polarkoordinaten durchgeführten Rechnung und erhalten entsprechend

$$\begin{aligned} \mathrm{d}x &= \cos\phi\cos\psi\,\mathrm{d}r - r\sin\phi\cos\psi\,\mathrm{d}\phi - r\cos\phi\sin\psi\,\mathrm{d}\psi \\ \mathrm{d}y &= \sin\phi\cos\psi\,\mathrm{d}r + r\cos\phi\cos\psi\,\mathrm{d}\phi - r\sin\phi\sin\psi\,\mathrm{d}\psi \\ \mathrm{d}z &= \sin\psi\,\mathrm{d}r + r\cos\psi\,\mathrm{d}\psi \end{aligned} \tag{A.1.2-B}$$

und damit wegen (Achtung: Koeffizientenmatrix ist transponiert)

$$\begin{aligned} &\begin{bmatrix} \cos\phi\cos\psi & \sin\phi\cos\psi & \sin\psi \\ -r\sin\phi\cos\psi & r\cos\phi\cos\psi & 0 \\ -r\cos\phi\sin\psi & -r\sin\phi\sin\psi & r\cos\psi \end{bmatrix}^{-1} \\ &\quad = \frac{1}{r\cos\psi}\begin{bmatrix} r\cos\phi\cos^2\psi & -\sin\phi & -\cos\phi\sin\psi\cos\psi \\ r\sin\phi\cos^2\psi & \cos\phi & -\sin\phi\sin\psi\cos\psi \\ r\sin\psi\cos\psi & 0 & \cos^2\psi \end{bmatrix} \end{aligned} \tag{A.1.2-C}$$

die Darstellung der kartesischen partiellen Ableitungen in Kugelkoordinaten

$$\begin{aligned} \partial_x &= \cos\phi\cos\psi\,\partial_r - \frac{\sin\phi}{r\cos\psi}\,\partial_\phi - \frac{\cos\phi\sin\psi}{r}\,\partial_\psi, \\ \partial_y &= \sin\phi\cos\psi\,\partial_r + \frac{\cos\phi}{r\cos\psi}\,\partial_\phi - \frac{\sin\phi\sin\psi}{r}\,\partial_\psi, \\ \partial_z &= \sin\psi\,\partial_r + \frac{\cos\psi}{r}\,\partial_\psi. \end{aligned} \tag{A.1.2-D}$$

✂ **Ergänzung** (Zylinderkooordinaten)**.** Außerhalb der z-Achse lassen sich Punkte im $\mathbb{R}^3$ ebenso durch *Zylinderkoordinaten*

$$x = r\cos\phi, \qquad y = r\sin\phi, \qquad z = z \tag{A.1.3-A}$$

beschreiben. Dies sind Polarkoordinaten der Ebene ergänzt um z und wir erhalten entsprechend

$$\mathrm{d}x = \cos\phi\,\mathrm{d}r - r\sin\phi\,\mathrm{d}\phi, \qquad \mathrm{d}y = \sin\phi\,\mathrm{d}r + r\cos\phi\,\mathrm{d}\phi, \qquad \mathrm{d}z \tag{A.1.3-B}$$

als Beschreibung der Differentiale in der lokalen Basis bestehend aus $\mathrm{d}r$, $\mathrm{d}\phi$ und $\mathrm{d}z$. Entsprechend ergibt sich die Darstellung der kartesischen partiellen Ableitungen

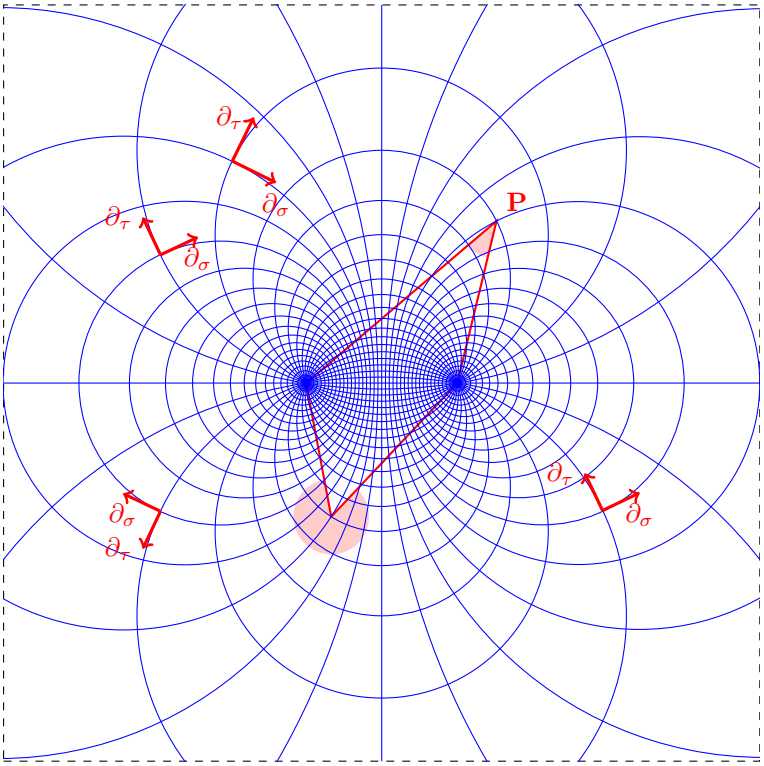
$$\partial_x = \cos\phi\,\partial_r - \frac{\sin\phi}{r}\partial_\phi, \qquad \partial_y = \sin\phi\,\partial_r + \frac{\cos\phi}{r}\partial_\phi, \qquad \partial_z \tag{A.1.3-C}$$

in Zylinderkoordinaten.

✂ **Ergänzung** (Bipolarkoordinaten)**.** Weitere interessante Koordinatensysteme ergeben sich aus physikalischen Anwendungen. So bestimmen Feld- und Äquipotentiallinien zweier Ladungen mit entgegengesetztem Vorzeichen *Bipolarkoordinaten* der Ebene. Der Einfachheit halber wählen wir die beiden Zentren (Ladungen) in den Punkten $(\pm 1, 0)^\top$, Feldlinien sind dabei Kreise durch die beiden Punkte und Äquipotentiallinien sind durch die Bedingung

$$\frac{(x-1)^2 + y^2}{(x+1)^2 + y^2} = \mathrm{e}^{2\tau} = \text{const}, \tag{A.1.4-A}$$

also die Konstanz der Streckenverhältnisse zu den Zentren, charakterisiert. Die Bedingung liefert wiederum Kreise (Kreis des Appollonius [2]), die Koordinatenlinien sehen also wie folgt aus:



Als Koordinaten eines Punktes P verwenden wir den Peripheriewinkel $\sigma \in [0, 2\pi)$ der Kreisbogensegmente zusammen mit dem Logarithmus der Streckenverhältnisse zu den beiden Zentren $\tau \in \mathbb{R}$. Also bestimmt σ den Kreisbogen

$$\text{um} \quad \begin{bmatrix} 0 \\ \cot\sigma \end{bmatrix} \quad \text{mit Radius} \quad \sqrt{1+\cot^2\sigma} = \frac{1}{|\sin\sigma|} \tag{A.1.4-B}$$

für $\sigma \notin \{\pi, 0\}$ und sonst die Linie $y = 0$ zwischen den Punkten beziehungsweise außerhalb. Weiter bestimmt τ einem Kreis mit Mittelpunkt auf der x-Achse durch die Punkte $(\frac{e^\tau-1}{e^\tau+1}, 0)$ und $(\frac{e^\tau+1}{e^\tau-1}, 0)$, also

$$\text{um} \quad \begin{bmatrix} \coth\tau \\ 0 \end{bmatrix} \quad \text{mit Radius} \quad \sqrt{\coth^2\tau - 1} = \frac{1}{|\sinh\tau|} \tag{A.1.4-C}$$

für $\theta \neq \frac{1}{2}$ und sonst der Linie $x = 0$. Schnittpunkte beider Kreise erfüllen also

$$x^2 + y^2 - 2y\cot\sigma - 1 = 0, \qquad x^2 + y^2 - 2x\coth^2\tau + 1 = 0. \tag{A.1.4-D}$$

Die beiden Bedingungen kann man nach x und y auflösen (und die Auflösung ist eindeutig, da man für die ersten Kreise die beiden Halbbögen durch verschiedene σ unterscheidet). Damit folgt

$$x = \frac{\sinh\tau}{\cosh\tau - \cos\sigma}, \qquad y = \frac{\sin\sigma}{\cosh\tau - \cos\sigma}. \tag{A.1.4-E}$$

Also gilt in diesen Koordinaten

$$\begin{aligned} \mathrm{d}x &= \frac{1 - \cosh\tau\cos\sigma}{(\cosh\tau - \cos\sigma)^2}\,\mathrm{d}\tau - \frac{\sinh\tau\sin\sigma}{(\cosh\tau - \cos\sigma)^2}\,\mathrm{d}\sigma, \\ \mathrm{d}y &= -\frac{\sinh\tau\sin\sigma}{(\cosh\tau - \cos\sigma)^2}\,\mathrm{d}\tau + \frac{\cosh\tau\cos\sigma - 1}{(\cosh\tau - \cos\sigma)^2}\,\mathrm{d}\sigma \end{aligned} \tag{A.1.4-F}$$

[2] Apollonios von Perge , ca. 265 v.u.Z. – ca. 190 v.u.Z.

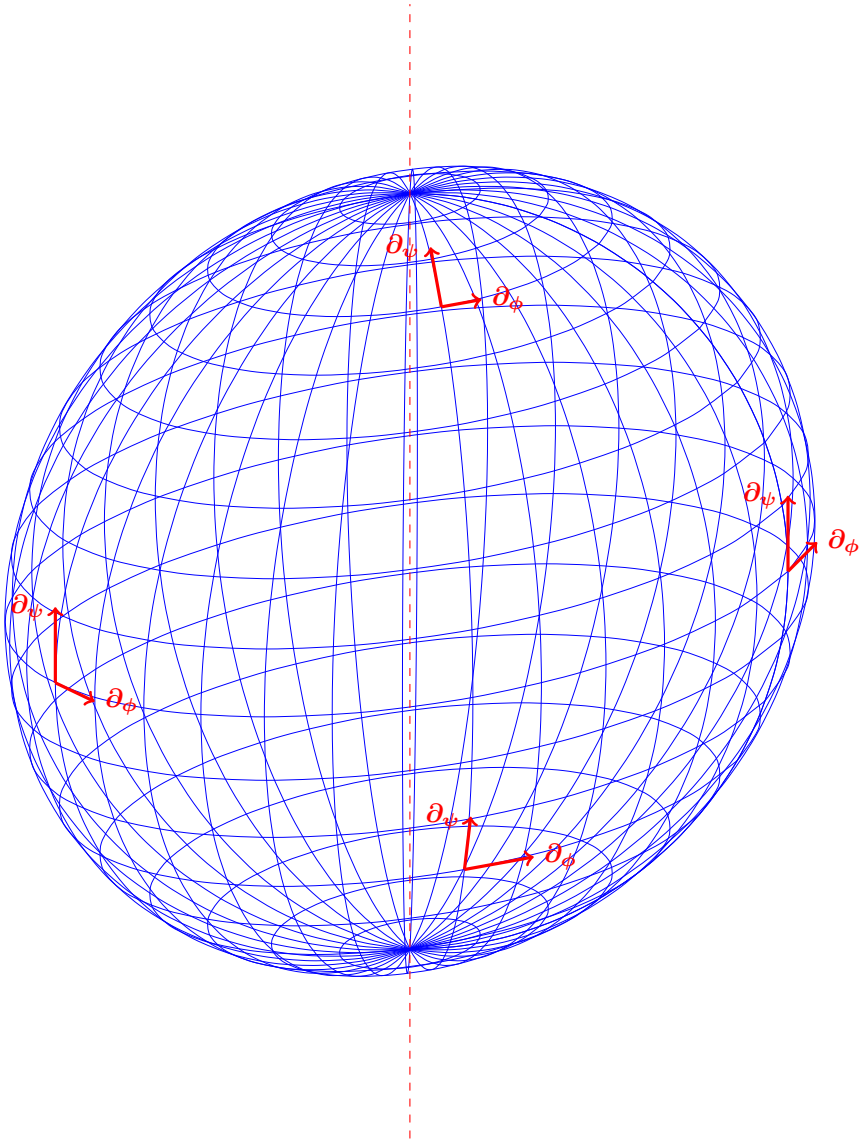
und wegen

$$\begin{bmatrix} 1-\cosh\tau\cos\sigma & -\sinh\tau\sin\sigma \\ -\sinh\tau\sin\sigma & \cosh\tau\cos\sigma-1 \end{bmatrix}^{-1} = \frac{1}{(\cosh\tau-\cos\sigma)^2}\begin{bmatrix} 1-\cosh\tau\cos\sigma & -\sinh\tau\sin\sigma \\ -\sinh\tau\sin\sigma & \cosh\tau\cos\sigma-1 \end{bmatrix} \tag{A.1.4-G}$$

damit auch

$$\begin{aligned} \partial_x &= (1-\cosh\tau\cos\sigma)\,\partial_\tau - \sinh\tau\sin\sigma\,\partial_\sigma, \\ \partial_y &= -\sinh\tau\sin\sigma\,\partial_\tau + (\cosh\tau\cos\sigma-1)\,\partial_\sigma. \end{aligned} \tag{A.1.4-H}$$

✂ **Ergänzung.** Koordinaten sind natürlich nicht nur im gesamten Raum sondern auch auf Flächen sinnvoll. Auf Kugeloberflächen ergeben Kugelkoordinaten mit festem Radius (mit Ausnahme der beiden Pole) sinnvolle Koordinaten.



Es ist nicht sinnvoll, ∂_x, ∂_y und ∂_z in diesen Koordinaten auszudrücken, Funktionen $f : \mathbb{S}^2 \to \mathbb{R}$ definiert auf der Kugeloberfläche $\mathbb{S}^2 = \{(x,y,z)^\top \mid x^2+y^2+z^2 = 1\}$ sind nur tangential zur Fläche sinnvoll linear approximierbar, also ableitbar. Damit genügen stets Linearkombinationen von ∂_ϕ und ∂_ψ zur Beschreibung erster Ableitungen. Dies werden wir uns im kommenden Semester genauer anschauen.

Index

Banachraum, 62

Differentialform, 101
  Pfaffsche Form, 101

Fourierreihen, 50
  Dirichlet-Kern, 51
Funktion
  Ableitung, 68
  Diffeomorphismus, 74
  Differential, 100
  differenzierbar, 68
  Fréchet-Ableitung, 68
  Fréchet-differenzierbar, 68
  Hesseform, 108
  Hessematrix, 108
  Immersion, 88
  Jacobimatrix, 69
  kritische Punkte, 98
    Sattelpunkt, 114
  lipschitzstetig, 53
  Niveaumenge, 98
  Oszillation, 12
  partiell differenzierbar, 69
  partielle Ableitung, 69
  R-integrierbar, 8
  reguläre Werte, 98
  stetig differenzierbar, 70
  Submersion, 87
  uneigentlich R-integrierbar, 30

Gebiet, 103
Gleichungen
  explizite, 82
  implizite, 82
Grundintegrale, 20
  elliptische Integrale, 26
  rationaler Ausdruck, 25

Konvergenz
  gleichmäßig, 40
  punktweise, 40
Koordinaten, 75
  Bipolarkoordinaten, 121
  Koordinatenfunktionen, 75
  Koordinatenlinien, 75
  Kugelkoordinaten, 121
  Polarkoordinaten, 119
  Polarkoordinaten in der Ebene, 74
  Zylinderkoordinaten, 121
Kurve, 89
  Cassinische Kurven, 99
  Doppelpunkte, 89
  Ellipse, 93
  Jordankurve, 90
  Kettenlinie, 93
  Länge, 91
  Orientierung, 90
  Parametrisierung, 89
    differenzierbar, 89
    lokal injektiv, 89
  Parametrisierung nach Bogenlänge, 91
  regulär, 89
  rektifizierbar, 91
Kurvenintegral, 101
  Kurvenintegral zweiter Art, 101
  Kurvenintergral erster Art, 94

Leibnizregel, 114
lokale Extrema
  hinreichendes Kriterium, 110
  notwendiges Kriterium, 98

metrischer Raum
  Abbildung

kontrahierend, 77
stetig, 64
Menge
abgeschlossen, 62
Häufungspunkt, 62
innerer Punkt, 61
offen, 62
perfekt, 62
Randpunkt, 61
äußerer Punkt, 61
Multiindizes, 65
Binomialkoeffizienten, 115
Multinomialkoeffizienten, 117

Potenzreihe, 45
Konvergenzradius, 46

Riemannintegral, 9
Darbouxsche Obersumme, 8
Darbouxsche Untersumme, 8
Riemannsummen, 12
Zerlegung, 7
Feinheit, 8
gemeinsame Verfeinerung, 8
Verfeinerung, 8

spezielle Funktion
Eulersche Betafunktion, 34
Eulersche Gammafunktion, 33
Rationale Funktionen, 21

uneigentliches Integral, 30
Ungleichung
Cauchy–Schwarz-Ungleichung, 63

Vektorraum, 58
Abbildung
beschränkt, 66
Defekt, 60
linear, 59
Rang, 60
Basis, 58
Bilinearform, 107
beschränkt, 107
Grammatrix, 107
symmetrisch, 107
Dimension, 59
Dimensionsformel, 59
Dimensionssatz, 60
Dualraum, 99
duale Basis, 99
transponierte Abbildung, 100
endlich erzeugt, 58
Matrix, 58
charakteristisches Polynom, 113
Determinante, 60
Einheitsmatrix, 58
Frobeniusnorm, 63
invertierbar, 58
Operatornorm, 64
Spur, 60
transponierte Matrix, 58
Norm, 62
normierter Raum, 62
vollständig, 62
quadratische Form, 107
Spaltenvektoren, 57
Unterraum, 59
direkte Summe, 59
Summe, 59
Vektoren
linear unabhängig, 58