Analysis 2

Sommersemester 2020

Ingo Runkel
Fachbereich Mathematik
Universität Hamburg

(Stand: 29. November 2020)

Webseite

www.math.uni-hamburg.de/home/runkel/ss20-ana.html

Bücher

- Rudin, Analysis (Oldenbourg, 4. Auflage)
- Königsberger, Analysis 1, 2 (Springer)
- Forster, Analysis 1, 2 (Vieweg)

und für Mutige

- Dieudonné, Foundations of modern analysis (Academic press)

Denken

Bitte denken Sie beim Durchlesen dieser Notizen mit und sagen mir Bescheid, wenn Sie Fehler finden.

Inhaltsverzeichnis

7 Integration

7.1 Einführung

In Kapitel 6 hatten wir Differentiation kennegelernt. Eine Weise, die Ableitung zu verstehen, war als die beste lineare Nährung an eine Funktion in einem Punkt p. Sei nämlich $f : \mathbb{R} \to \mathbb{R}$, dann

$$f(x) \;=\; f(p) + f'(p)\,(x-p) + r(x)\ ,$$

wobei $r(x)$ für $x \to p$ schneller als linear gegen 0 geht. Falls $f'(p)$ existiert, ist es eindeutig.

Die einfachste Weise, das Integral einzuführen, ist als *Stammfunktion*. Frage:

> Gegeben eine Funktion $f : [a,b] \to \mathbb{R}$, gibt es $F : [a,b] \to \mathbb{R}$ differenzierbar auf $[a,b]$, so dass $F'(x) = f(x)$ für alle $x \in [a,b]$?
> Ist F eindeutig?

Zumindest die zweite Frage können wir gleich verneinen, da für jedes $c \in \mathbb{R}$ die Ableitungen von $F(x)$ und $F(x)+c$ gleich sind. Die erste Frage werden wir im Laufe des Kapitels beantworten mit

$$F(x) \;=\; \int_q^x f(t)\,dt\ ,$$

wobei $q, x \in [a,b]$ und f geeignete Voraussetzungen erfüllt.

Der Nachteil der obigen Definition ist, dass man nicht sofort sieht, warum sie nützlich sein soll. Das war bei der Ableitung besser („Schwierige Funktion in einem Punkt so gut wie möglich linear machen, und linear ist einfach.“).

Aus diesem Blickwinkel ist folgende Definition besser: Für $f : [a,b] \to \mathbb{R}_{\geq 0}$ ist das Integral von f die Fläche zwischen den Graphen von f und der x-Achse. Das beantwortet z.B. die Frage:

> Male den Graphen von $f : [a,b] \to \mathbb{R}$ auf ein Stück dicke Pappe. Schneide entlang f und x-Achse aus. Wieviel wiegt das Stück Pappe (relativ zum Papp-Quadrat $[0,1] \times [0,1]$)?

Etwas formaler: Sei wieder $f : [a,b] \to \mathbb{R}_{\geq 0}$ (bei negativen Werten geht es etwas anders). Setze

$$P = \left\{\, (x,y) \in \mathbb{R}^2 \,\middle|\, x \in [a,b] \text{ und } 0 \leq y \leq f(x) \right\}\ .$$

Dann werden wir in diesem Kapitel aus der Idee „Das Integral von f ist der Flächeninhalt von P“ Sinn machen.

In Analysis 3 lernt man „Lebesgue Integration“ und „Maßtheorie“ kennen. Diese sind eine weitreichende Verallgemeinerung des Riemann-Integrals, das wir hier anschauen und verhalten sich in vielen Anwendungen besser.

7.2 Treppenfunktionen

Siehe [§18 aus Forster 1] und [Kapitel 11.1 aus Königsberger 1].

■ Wir definieren die Teilmenge $\mathcal{T}(a,b) \subset \mathrm{Fun}([a,b],\mathbb{R})$ der *Treppenfunktionen* wie folgt: Sei $f : [a,b] \to \mathbb{R}$. Dann ist $f \in \mathcal{T}(a,b)$ genau dann, wenn es $n \in \mathbb{N}$, sowie

$$a = x_0 < x_1 < \cdots < x_{n-1} < x_n = b$$

und $c_0, \ldots, c_{n-1} \in \mathbb{R}$ gibt, so dass für $i = 0, \ldots, n-1$ und $x \in [a,b]$ gilt:

$$x_i < x < x_{i+1} \quad \Rightarrow \quad f(x) = c_i \ .$$

Die Werte von f an den Sprungstellen x_i dürfen beliebig sein.

■ Wir nennen die Menge $Z = \{a = x_0, x_1, \ldots, x_n = b\}$ eine *Zerlegung von* $[a,b]$.

Lemma 7.2.1. Die Menge der Treppenfunktionen $\mathcal{T}(a,b)$ ist ein Untervektorraum des $\mathbb{R}$-Vektorraumes $\mathrm{Fun}([a,b],\mathbb{R})$.

Beweis. Seien $f, g \in \mathcal{T}(a,b)$ und $\lambda \in \mathbb{R}$. Dass $\lambda f \in \mathcal{T}(a,b)$ ist klar (Warum?). ?
Für $f+g$ betrachte Zerlegungen Z_f und Z_g von $[a,b]$. Dann ist $Z_f \cup Z_g$ eine Zerlegung von $[a,b]$, die zeigt, dass $f+g$ eine Treppenfunktion ist. (Details?) ?
□

■ Wir wollen jeder Treppenfunktion eine Zahl zuordnen (den „Flächeninhalt mit Vorzeichen“ falls die Werte in $\mathbb{R}$ liegen). Dazu suchen wir eine $\mathbb{R}$-lineare Funktion

$$I : \mathcal{T}(a,b) \longrightarrow \mathbb{R}$$

mit folgenden Eigenschaften:

1. („Rechteck gibt Breite mal Höhe“) Für $a \leq l \leq r \leq b$ sei f die Funktion

$$f(x) = \begin{cases} 1 & ;\ l \leq x \leq r \\ 0 & ;\ \text{sonst} \end{cases}$$

   Dann $I(f) = r - l$.

2. ($\mathbb{R}$-linear) Für $f, g \in \mathcal{T}(a,b)$ und $\lambda \in \mathbb{R}$:

$$I(f+g) = I(f) + I(g) \quad \text{und} \quad I(\lambda f) = \lambda I(f) \ .$$

In Eigenschaft 1 ist der Fall $l = r$ explizit zugelassen. Das sagt, dass eine Funktion f, die überall 0 ist, außer an der Stelle $l = r$, immer noch $I(f) = 0$ erfüllt.

Lemma 7.2.2.

(i) Es gibt eine eindeutige Funktion I mit den Eigenschaften 1 und 2.

(ii) Für $f \in \mathcal{T}(a,b)$ mit Zerlegung $x_0, \dots, x_n$ und Werten $c_0, \dots, c_{n-1}$ wie oben gilt

$$I(f) = \sum_{i=0}^{n-1} c_i \, (x_{i+1} - x_i) \ .$$

Beweis.
■ Für die Eindeutigkeit in (i) schreibt man eine Treppenfunktion als Summe von Funktionen wie in 1 oben und benutzt dann 2. (Details? Warum braucht man in 1 hier den Fall $l = r$?) ?
■ Für die Existenz in (i) überlegen wir uns, dass die Vorschrift in (ii) wohldefiniert ist und die Eigenschaften 1 und 2 hat.

Wohldefiniertheit bedeutet hier, dass I in der Tat nur von f abhängt und nicht von der Wahl der Zerlegung.

(Um sich klarzumachen, dass es bei „wohldefiniert“ etwas zu tun gibt, sollte man einmal statt der Vorschrift in (ii) die Formel $I(f) = \sum_{i=0}^{n-1} i \, c_i \, (x_{i+1} - x_i)$ betrachten. Das ist nicht wohldefiniert. Warum?) ?

Sei also $Z = \{x_0, \dots, x_n\}$ eine Zerlegung von $[a,b]$ und $c_0, \dots, c_{n-1} \in \mathbb{R}$, so dass $f(x) = c_i$ für $x_i < x < x_{i+1}$. Sei $Z' = \{x'_0, \dots, x'_m\}$ eine weitere Zerlegung und c'_j die entsprechenden Werte. Wir müssen zeigen

$$\sum_{i=0}^{n-1} c_i \, (x_{i+1} - x_i) \ = \ \sum_{i=0}^{m-1} c'_i \, (x'_{i+1} - x'_i) \ . \qquad (*)$$

Es genügt, sich dies für den Fall zu überlegen, dass Z' eine *Verfeinerung* von Z ist, d.h. falls $Z \subset Z'$. Und hier genügt es wiederum, zu zeigen, dass der Wert der Summe sich nicht ändert, wenn man eine einzelne zusätzliche Unterteilung einbaut, sagen wir t mit $x_i < t < x_{i+1}$. Die neue Zerlegung ist dann $Z \cup \{t\}$. In diesem Fall prüft man nun leicht, dass $(*)$ gilt. (Warum kann man diese beiden Vereinfachungen machen? Warum gilt dann $(*)$?) ?

Es bleibt, zu zeigen, dass $I(f)$ Eigenschaften 1 und 2 hat. Das sieht man duch Nachrechnen. (Details? Insbesondere für $I(f+g) = I(f) + I(g)$, was etwas mehr Nachdenken verlangt.) ? □

- Für $f \in \mathcal{T}(a,b)$ schreiben wir ab jetzt

$$\int_a^b f(x)\,dx \overset{\text{Def.}}{=} I(f)$$

und nennen $\int_a^b f(x)\,dx$ das *Integral der Treppenfunktion* f.

- Für $f, g \in \mathrm{Fun}([a,b],\mathbb{R})$ schreiben wir

$$f \leq g \quad \text{falls} \quad f(x) \leq g(x) \text{ für alle } x \in [a,b]\ .$$

Für $\varphi, \psi \in \mathcal{T}(a,b)$ gilt

$$\varphi \leq \psi \quad \Rightarrow \quad \int_a^b \varphi(x)\,dx \leq \int_a^b \psi(x)\,dx\ .$$

(Warum?) ?

7.3 Das Riemann-Integral

Siehe [§18 aus Forster 1].

Definition 7.3.1. Sei $f \in \mathrm{Fun}([a,b],\mathbb{R})$ beschränkt. Das *Ober– und Unterintegral* von f ist definiert als

$$\text{(Oberintegral)} \qquad \int_a^{b\,*} f(x)\,dx = \inf\left\{ \int_a^b \overline{\varphi}(x)dx \,\middle|\, \overline{\varphi} \in \mathcal{T}(a,b) \text{ und } \overline{\varphi} \geq f \right\}\ ,$$

$$\text{(Unterintegral)} \qquad \int_{a\,*}^{b} f(x)\,dx = \sup\left\{ \int_a^b \underline{\varphi}(x)dx \,\middle|\, \underline{\varphi} \in \mathcal{T}(a,b) \text{ und } \underline{\varphi} \leq f \right\}\ .$$

(Warum muss man für die Definition annehmen, dass f beschränkt ist?) ?

Bemerkung 7.3.2. Sei $\underline{\varphi} \leq f$ und $\overline{\varphi} \geq f$. Dann auch $\underline{\varphi} \leq \overline{\varphi}$ und somit $\int_a^b \underline{\varphi}(x)\,dx \leq \int_a^b \overline{\varphi}(x)\,dx$. Somit ist $\int_a^b \overline{\varphi}(x)\,dx$ eine obere Schranke für die Menge in der Definition des Unterintegrals. Und umgekehrt auch $\int_a^b \underline{\varphi}(x)\,dx$ eine untere Schranke für die Menge in der Defintion des Oberintegrals. Man erhält

$$\int_{a\,*}^{b} f(x)\,dx \;\leq\; \int_a^{b\,*} f(x)\,dx\ .$$

Beispiel 7.3.3.

(1) Für $f \in \mathcal{T}(a,b)$ gilt

$$\int_{a\,*}^{b} f(x)\,dx \;=\; \int_{a}^{b\,*} f(x)\,dx \;=\; \int_{a}^{b} f(x)\,dx \;.$$

(Warum?) ?

(2) Sei $f : [0,1] \to \mathbb{R}$ gegen durch

$$f(x) = \begin{cases} 1 & ; x \in \mathbb{Q} \\ 0 & ; x \notin \mathbb{Q} \end{cases} \quad .$$

Dann gilt

$$\int_{0\,*}^{1} f(x)\,dx = 0 \quad \text{und} \quad \int_{0}^{1\,*} f(x)\,dx = 1 \;.$$

Um dies zu sehen, überlegt man sich zunächst, dass die Menge $\{\varphi \in \mathcal{T}(a,b) \text{ und } \varphi \le f\}$ gleich der Menge

$$\big\{\varphi \in \mathcal{T}(a,b) \;\big|\; \varphi \le 1 \text{ , und } \varphi(x) \le 0 \text{ für alle bis auf endlich viele } x \in [a,b] \cap \mathbb{Q}\,\big\}$$

ist. Daraus folgt dann $\int_{0\,*}^{1} f(x)\,dx = \int_{0\,*}^{1} 0\,dx = 0$.
(Details? Wie geht das Argument für $\int_{0}^{1\,*} f(x)\,dx$?) ?

Definition 7.3.4. Wir nennen eine beschränkte Funktion $f : [a,b] \to \mathbb{R}$ *Riemann-integrierbar*, oder kurz *integrierbar*, falls gilt

$$\int_{a\,*}^{b} f(x)\,dx \;=\; \int_{a}^{b\,*} f(x)\,dx$$

In diesem Fall setzen wir

$$\int_{a}^{b} f(x)\,dx \overset{\text{Def.}}{=} \int_{a\,*}^{b} f(x)\,dx$$

und nennen $\int_a^b f(x)\,dx$ das *(Riemann-)Integral von* f. (Natürlich können wir statt $\int_{a\,*}^{b} f(x)\,dx$ genausogut $\int_{a}^{b\,*} f(x)\,dx$ nehmen.) Die Menge der Riemann-integrierbaren (also insbesondere beschränkten) Funktionen nennen wir

$$\mathcal{R}(a,b) \;\subset\; \mathrm{Fun}([a,b],\mathbb{R}) \;.$$

■ Aus Beispiel (1) oben wissen wir: Treppenfunktionen sind integrierbar,

$$\mathcal{T}(a,b) \ \subset \ \mathcal{R}(a,b) \ .$$

Aus Beispiel (2) oben sehen wir, dass nicht alle Funktionen integrierbar sind:

$$\mathcal{R}(a,b) \ \subsetneq \ \mathrm{Fun}([a,b],\mathbb{R}) \ .$$

Gleich werden wir sehen, dass stetige Funktionen integrierbar sind. Also auch $\mathcal{T}(a,b) \subsetneq \mathcal{R}(a,b)$.

Nebenbemerkung zur Information: Hier ist ein erster Unterschied zwischen dem Riemann-Integral, dass wir hier behandeln, und dem Lebesgue-Integral, dass in Analysis 3 eingeführt wird. Und auch ein Grund, „Riemann-integrierbar“ zu sagen, wenn der Zusammenhang nicht klar ist: Die Funktion aus Beispiel (2) oben ist nicht Riemann-integrierbar. Aber sie ist Lebesgue-integrierbar (mit Wert des Integrals gleich Null). Beim Lebesgue-Integral im R^d, $d \geq 1$, haben abzählbare Punktemengen „Maß Null“, und wenn sich zwei Funktionen auf einer Menge von Maß Null unterscheiden, haben Sie das gleiche Lebesgue-Integral.

Das Integral ist *monoton*:

Lemma 7.3.5. Seien $f, g \in \mathcal{R}(a,b)$ mit $f \leq g$. Dann $\int_a^b f(x)dx \leq \int_a^b g(x)dx$.

Der Beweis ergibt sich sofort aus der Definition des Unterintegrals. (Wie?) ?

Folgendes Kriterium für Integrierbarkeit werden wir häufiger verwenden.

Lemma 7.3.6. Für $f \in \mathrm{Fun}([a,b],\mathbb{R})$ sind äquivalent:

(i) $f \in \mathcal{R}(a,b)$

(ii) Für jedes $\varepsilon > 0$ gibt es $\overline{\varphi}, \underline{\varphi} \in \mathcal{T}(a,b)$ mit $\underline{\varphi} \leq f \leq \overline{\varphi}$ und

$$\int_a^b \overline{\varphi}(x)dx \ - \ \int_a^b \underline{\varphi}(x)dx \ < \ \varepsilon \ .$$

Beweis.
(i)⇒(ii): Klar. (Warum?) ?

(ii)⇒(i): Es gilt

$$0 \overset{\text{Bem. 7.3.2}}{\leq} \int_a^{b*} f(x)dx \ - \ \int_{a*}^b f(x)dx \ \leq \ \int_a^b \overline{\varphi}(x)dx - \int_a^b \underline{\varphi}(x)dx \ < \ \varepsilon \ .$$

Da $\varepsilon > 0$ beliebig gewählt werden kann (mit entsprechenden $\overline{\varphi}, \underline{\varphi}$), müssen Ober- und Unterintegral übereinstimmen. □

Korollar 7.3.7. Falls es für jedes $\varepsilon > 0$ Treppenfunktionen $\overline{\varphi}, \underline{\varphi} \in \mathcal{T}(a,b)$ gibt, so dass $\underline{\varphi} \leq f \leq \overline{\varphi}$ und $\overline{\varphi}(x) - \underline{\varphi}(x) < \varepsilon$ für alle $x \in [a,b]$, dann $f \in \mathcal{R}(a,b)$.

(Warum gilt das Korollar?) ?

Satz 7.3.8. Stetige Funktionen $[a,b] \to \mathbb{R}$ sind (Riemann-)integrierbar.

(Warum sind stetige Funktionen auf $[a,b]$ automatisch beschränkt?) ?

Satz 7.3.9. Monotone Funktionen $[a,b] \to \mathbb{R}$ sind (Riemann-)integrierbar.

(Warum sind monotone Funktionen auf $[a,b]$ automatisch beschränkt?) ?
Die Funktion f braucht hier nicht stetig zu sein.

Beweis. Wir führen den Beweis für monoton steigendes f.

Sei $\delta > 0$ beliebig. Wähle eine Zerlegung $a = x_0, x_1, \ldots, x_n = b$ von $[a,b]$, so dass $x_{i+1} - x_i < \delta$. Definiere $\overline{\varphi}, \underline{\varphi} \in \mathcal{T}(a,b)$ als

$$\underline{\varphi}(x) = f(x_i) \ \text{ für } \ x \in [x_i, x_{i+1}) \ , \quad \overline{\varphi}(x) = f(x_{i+1}) \ \text{ für } \ x \in (x_i, x_{i+1}] \ ,$$

sowie $\overline{\varphi}(a) = f(a)$ und $\underline{\varphi}(b) = f(b)$ (und nach der Definition oben sowieso schon $\underline{\varphi}(a) = f(a)$ und $\overline{\varphi}(b) = f(b)$). Dann gilt für alle $x \in [a,b]$, dass

$$\underline{\varphi}(x) \ \leq \ f(x) \ \leq \ \overline{\varphi}(x) \ .$$

(Warum?) Andererseits ?

$$\begin{aligned}
&\int_a^b \overline{\varphi}(x)dx - \int_a^b \underline{\varphi}(x)dx \\
&= \sum_{i=0}^{n-1} f(x_{i+1})(x_{i+1} - x_i) - \sum_{i=0}^{n-1} f(x_i)(x_{i+1} - x_i) \\
&= \sum_{i=0}^{n-1} \underbrace{\big(f(x_{i+1}) - f(x_i)\big)}_{\geq 0} \underbrace{(x_{i+1} - x_i)}_{<\delta} \\
&\leq \sum_{i=0}^{n-1} \big(f(x_{i+1}) - f(x_i)\big)\,\delta \ = \ \big(f(b) - f(a)\big)\,\delta \ .
\end{aligned}$$

Mit Lemma 7.3.6 folgt $f \in \mathcal{R}(a,b)$. □

(Mit diesem Satz kann man Beispiele für integrierbare Funktionen basteln, die unendlich viele Sprungstellen haben. Fällt Ihnen ein Beispiel ein?) ?

Satz 7.3.10. $\mathcal{R}(a,b)$ ist ein Untervektorraum von $\mathrm{Fun}([a,b],\mathbb{R})$. Die Abbildung

$$\mathcal{R}(a,b) \longrightarrow \mathbb{R} \quad , \quad f \longmapsto \int_a^b f(x)dx$$

ist $\mathbb{R}$-linear.

Beweis. Sei $\lambda \in \mathbb{R}$ und $f, g \in \mathcal{R}(a,b)$. Die folgenden vier Aussagen müssen gezeigt werden:

$$\lambda f \in \mathcal{R}(a,b) \quad , \quad f+g \in \mathcal{R}(a,b) \quad ,$$
$$\int_a^b \lambda f(x)dx = \lambda \int_a^b f(x)dx \quad , \quad \int_a^b (f+g)(x)dx = \int_a^b f(x)dx \; + \; \int_a^b g(x)dx \; .$$

Wir zeigen nur $f+g \in \mathcal{R}(a,b)$, der Rest ist zum selber Überlegen (Details?). ?

Sei $\varepsilon > 0$. Nach Lemma 7.3.6 gibt es $\overline{\varphi}, \underline{\varphi}, \overline{\psi}, \underline{\psi} \in \mathcal{T}(a,b)$ mit $\underline{\varphi} \le f \le \overline{\varphi}$, $\underline{\psi} \le g \le \overline{\psi}$ und

$$\int_a^b \overline{\varphi}(x)dx \; - \; \int_a^b \underline{\varphi}(x)dx < \varepsilon \quad , \quad \int_a^b \overline{\psi}(x)dx \; - \; \int_a^b \underline{\psi}(x)dx < \varepsilon \; .$$

Dann gilt aber auch $\underline{\varphi} + \underline{\psi} \le f + g \le \overline{\varphi} + \overline{\psi}$ und

$$\int_a^b (\overline{\varphi}(x) + \overline{\psi}(x))dx \; - \; \int_a^b (\underline{\varphi}(x) + \underline{\psi}(x))dx < 2\varepsilon \; ,$$

da wir nach Lemma 7.2.2 und Beispiel 7.3.3 (1) schon wissen, dass das Integral auf Treppenfunktionen linear ist. Nach Lemma 7.3.6 ist $f+g \in \mathcal{R}(a,b)$. $\square$

Satz 7.3.11. Sei $f \in \mathcal{R}(a,b)$ mit $m \le f \le M$ für geeignete Schranken $m, M \in \mathbb{R}$. Sei $G : [m,M] \to \mathbb{R}$ stetig. Für die Funktion $h := G \circ f$ gilt

$$h \in \mathcal{R}(a,b) \; .$$

Beweis. Sei $\varepsilon > 0$ beliebig.

■ Wir brauchen Satz 5.3.12 aus Analysis 1 („Stetige Abbildungen von einem kompakten metrischen Raum in einen metrischen Raum sind gleichmäßig stetig.“)

Somit ist G auch gleichmäßig stetig, und es gibt $\delta > 0$, so dass

$$\forall u, v \in [m,M] \; : \; |u - v| \le \delta \; \Rightarrow \; |G(u) - G(v)| \le \varepsilon \; .$$

(Das ist äquivalent zu der Formulierung mit „<“.) Wir brauchen unten die folgende Konsequenz: Sei $[c,d] \subset [m,M]$ mit $d - c \le \delta$. Dann auch

$$\sup\{G(x) | x \in [c,d]\} \; - \; \inf\{G(x) | x \in [c,d]\} \; \le \; \varepsilon \; . \qquad (*)$$

(Warum?) ?

■ Da $f \in \mathcal{R}(a,b)$ können wir eine Zerlegung $a = x_0 < x_1 < \cdots < x_n = b$ finden, zusammen mit Treppenfunktionen $\overline{\varphi}$, $\underline{\varphi}$ *bezüglich dieser Zerlegung*, so dass $\underline{\varphi} \leq f \leq \overline{\varphi}$ und

$$\int_a^b \big(\overline{\varphi}(x) - \underline{\varphi}(x)\big)dx \ < \ \delta^2 \ .$$

Wir können annehmen, dass $\overline{\varphi}(x)$, $\underline{\varphi}(x)$ auf den halboffenen Intervallen $[x_i, x_{i+1})$ konstant sind. (Die Werte an den Sprungstellen der Treppenfunktion können ja beliebig gewählt werden.)

■ Nun definieren wir Treppenfunktionen $\overline{\psi}$, $\underline{\psi}$, die h einsperren. Setze

$$m_i = \inf\{h(x) | x \in [x_i, x_{i+1}]\} \ \ , \ \ M_i = \sup\{h(x) | x \in [x_i, x_{i+1}]\} \ .$$

Für $x \in [x_i, x_{i+1})$ definiere $\underline{\psi}(x) = m_i$, $\overline{\psi}(x) = M_i$. Für $x = b$ setzen wir $\underline{\psi}(b) = \overline{\psi}(b) = h(b)$. Dann per Konstruktion

$$\underline{\psi} \leq h \leq \overline{\psi} \ .$$

■ Wir unterteilen die Indexmenge $\{0, 1, \ldots, n-1\}$ in zwei Gruppen: $i \in A$ falls $\overline{\varphi}(x) - \underline{\varphi}(x) < \delta$ für alle $x \in [x_i, x_{i+1})$, und $i \in B$ sonst.

Beh.: Sei $\Delta_i = x_{i+1} - x_i$. Dann $\sum_{i \in B} \Delta_i < \delta$.

Bew.:

$$\sum_{i \in B} \delta \, \Delta_i \ \leq \ \sum_{i \in B} (\overline{\varphi}(x_i) - \underline{\varphi}(x_i))\Delta_i \ \leq \ \int_a^b \big(\overline{\varphi}(x) - \underline{\varphi}(x)\big)dx \ < \ \delta^2$$

■ Für $i \in A$ und $x \in [x_i, x_{i+1})$ gilt dann wegen $(*)$, dass $M_i - m_i \leq \varepsilon$.

Für $i \in B$ haben wir zumindest $M_i - m_i \leq 2K$, wobei K eine obere Schranke für $|G|$ auf $[m, M]$ ist.

Jetzt können wir endlich die Integrierbarkeit von h zeigen:

$$\begin{aligned}
&\int_a^b \big(\overline{\psi}(x) - \underline{\psi}(x)\big)dx \\
&= \sum_{i=0}^{n-1} \big(\overline{\psi}(x_i) - \underline{\psi}(x_i)\big)\Delta_i \\
&= \sum_{i \in A} (M_i - m_i)\Delta_i + \sum_{i \in B} (M_i - m_i)\Delta_i \\
&\leq \sum_{i \in A} \varepsilon \, \Delta_i + \sum_{i \in B} 2K\Delta_i \ \leq \ \varepsilon(b-a) \ + \ 2K\delta \ .
\end{aligned}$$

Da wir ε und δ beliebig klein wählen können, folgt $h \in \mathcal{R}(a,b)$ aus Lemma 7.3.6 □

Beweis von Satz 7.3.8 („stetige Funktionen sind integrierbar“).
Sei $G : [a,b] \to \mathbb{R}$ stetig. Betrachte $f : [a,b] \to [a,b]$, $f(x) = x$. Nach Satz 7.3.9 ist f integrierbar. Nach Satz 7.3.11 ist $G \circ f = G$ integrierbar. □

■ Für $f \in \mathrm{Fun}([a,b],\mathbb{R})$ setzen wir

$$f_+(x) = \begin{cases} f(x) & ; f(x) > 0 \\ 0 & ; f(x) \leq 0 \end{cases} \quad , \quad f_-(x) = \begin{cases} -f(x) & ; f(x) < 0 \\ 0 & ; f(x) \geq 0 \end{cases} \quad .$$

Damit gilt

$$f = f_+ - f_- \quad \text{und} \quad |f| = f_+ + f_- \ .$$

Satz 7.3.12. Sei $f, g \in \mathcal{R}(a,b)$. Dann gilt:

(i) $f_+, f_-, |f| \in \mathcal{R}(a,b)$. Es gilt

$$\Big| \int_a^b f(x)dx \Big| \ \leq \ \int_a^b |f(x)|dx \ .$$

(ii) Für jedes $p \geq 0$ ist $|f|^p \in \mathcal{R}(a,b)$.

(iii) $f \cdot g \in \mathcal{R}(a,b)$.

(Für $p < 0$ gilt die Aussage in (ii) nicht, selbst wenn $f(x) \neq 0$ für alle $x \in [a,b]$. Fällt Ihnen ein Gegenbeispiel ein?) ?

Beweis.

(i) Beh.: $f_+ \in \mathcal{R}(a,b)$.

⌈ Wende Satz 7.3.11 auf $G = id_+$ an, d.h. auf die Funktion $G(x) = x$ für $x > 0$ und $G(x) = 0$ für $x \leq 0$. Denn dann $f_+ = G \circ f$. ⌋

Für f_- geht das Argument genauso. Dann auch $|f| = f_+ + f_- \in \mathcal{R}(a,b)$ da $\mathcal{R}(a,b)$ ein $\mathbb{R}$-Vektorraum ist. Die Abschätzung folgt aus

$$\pm(f_+(x) - f_-(x)) \ \leq \ f_+(x) + f_-(x)$$

und Lemma 7.3.5. (Wie?) ?

(ii) Wende Satz 7.3.11 auf $G(x) = |x|^p$ an.

(iii) Schreibe $fg = \frac{1}{2}(|f+g|^2 - |f|^2 - |g|^2)$. Nach Teil (ii) ist fg dann auch integrierbar.

$\square$

Nebenbemerkung zur Information: Nach den vorherigen Sätze über die Menge der Riemann-integrierbaren Funktionen kann man sich fragen, ob es ein alternative Beschreibungen dieser Menge gibt. Eine solche Beschreibung ist, dass $\mathcal{R}(a,b)$ genau die Menge der beschränkten und fast überall stetigen Funktionen ist, siehe [Satz 11.33 aus Rudin].

Bemerkung 7.3.13.

(1) Sei $f \in \mathcal{R}(a,b)$, und sei $g : [a,b] \to \mathbb{R}$ eine Funktion, die sich von f nur an endlich vielen Stellen unterscheidet. Dann auch $g \in \mathcal{R}(a,b)$ und

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

(Warum?) ?

(2) Falls $f : [a,b] \to \mathbb{R}$ auf $[a,b]$ integrierbar ist, so ist auch jede Einschränkung auf ein Teilintervall $[c,d] \subset [a,b]$ integrierbar, also $f|_{[c,d]} \in \mathcal{R}(c,d)$.

Umgekehrt, falls $g \in \mathcal{R}(c,d)$, so ist die Funktion

$$f : [a,b] \to \mathbb{R} \quad , \quad f(x) = \begin{cases} g(x) & ; x \in [c,d] \\ 0 & ; \text{sonst} \end{cases}$$

auf $[a,b]$ integrierbar.

(Warum gilt das beides?) ?

(3) Für eine Zerlegung $a = x_0 < x_1 < \cdots < x_n = b$ von $[a,b]$ betrachte Funktionen f_i, die auf $[x_i, x_{i+1}]$ integrierbar sind, und außerhalb gleich Null sind. Dann ist $\sum_{i=0}^{n-1} f_i$ integrierbar auf $[a,b]$. Z.B. ist jede stückweise stetige Funktion integrierbar.

Bemerkung 7.3.14. Eine komplexwertige Funktionen $f : [a,b] \to \mathbb{C}$ nennen wir Riemann-integrierbar, wenn die Funktionen $\mathrm{Re}(f), \mathrm{Im}(f) : [a,b] \to \mathbb{R}$ Riemann-integrierbar sind. Wir werden auch für die Menge der komplexwertigen Riemann-integrierbaren Funktionen die Notation

$$\mathcal{R}_{\mathbb{C}}(a,b) \quad \text{oder auch nur} \quad \mathcal{R}(a,b)$$

verwenden. Für reell-wertige integrierbare Funktionen schreiben wir entsprechend auch $\mathcal{R}_{\mathbb{R}}(a,b)$. Wir setzen

$$\int_a^b f(x)\,dx \stackrel{\text{Def.}}{=} \int_a^b \mathrm{Re}(f(x))\,dx + i\int_a^b \mathrm{Im}(f(x))\,dx\ .$$

Das Integral ist $\mathbb{C}$-linear. Insbesondere gilt für $\lambda \in \mathbb{C}$, dass

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

(Warum? Können Sie folgende Bemerkung in Formeln ausdrücken? „Das obige Integral auf komplexwertigen Funktionen ist die eindeutige $\mathbb{C}$-lineare Fortsetzung des ursprünglichen Integrals auf reell-wertigen Funktion.") ?

Falls $f \in \mathcal{R}_{\mathbb{C}}(a,b)$, so sind auch die komplex konjugierte Funktion $\bar{f}$ und der Absolutbetrag $|f|$ integrierbar. Es gilt auch hier

$$\Big|\int_a^b f(x)dx\Big| \;\leq\; \int_a^b |f(x)|dx\ .$$

Der Beweis geht über die Cauchy-Schwarzsche Ungleichung und wird in Kapitel 7.6 in etwas allgemeinerer Form nachgeholt (siehe Lemma 7.6.3).

7.4 Integration und Differentiation

In diesem Kapitel steht $\mathbb{K}$ für $\mathbb{R}$ oder $\mathbb{C}$, und $\mathcal{R}(a,b)$ steht für $\mathcal{R}_{\mathbb{K}}(a,b)$.

■ Bislang haben wir für Funktionen $f \in \mathcal{R}(a,b)$ das Integral $\int_a^b f(x)dx$ definiert. Unter Benutzung von Bemerkung 7.3.13 (2) erweitern wir die Notation wie folgt: für $c,d \in [a,b]$ setzen wir

$$\int_c^d f(x)\,dx \;=\; \begin{cases} \int_c^d f(x)\,dx & ; c \leq d \\ -\int_d^c f(x)\,dx & ; c > d \end{cases}\qquad .$$

In Worten: Ist die obere Grenze d kleiner als die untere Grenze c, so ist das Integral von c nach d definiert als Minus das Integral von d nach c. Z.B.

$$\int_1^2 1\,dx = 1 \quad \text{und} \quad \int_2^1 1\,dx = -1\ .$$

Diese Konvention wird sich als praktisch erweisen. Z.B. für folgende Beobachtung:

Lemma 7.4.1. Sei $f \in \mathcal{R}(a,b)$. Für $c, d, y \in [a,b]$ gilt

$$\int_c^d f(x)\,dx \;=\; \int_c^y f(x)\,dx \;+\; \int_y^d f(x)\,dx\ .$$

Der Beweis ist zum selber überlegen. (Details?) ?

Satz 7.4.2. Sei $a < b$ und $f \in \mathcal{R}(a,b)$. Für $p \in [a,b]$ definiere die Funktion

$$F : [a,b] \longrightarrow \mathbb{K} \quad , \quad y \longmapsto \int_p^y f(x)\,dx\ .$$

Dann gilt:

(i) Die Funktion F ist auf $[a,b]$ stetig.

(ii) Sei $s \in [a,b]$ ein Punkt, in dem die Funktion f stetig ist. Dann ist F in s differenzierbar und es gilt

$$F'(s) = f(s)\ .$$

Beweis. Sei $s \in [a,b]$.

(i) Für $y \in [a,b]$ gilt

$$F(y) - F(s) \;=\; \int_p^y f(x)dx - \int_p^s f(x)dx \overset{\text{Lem. 7.4.1}}{=} \int_s^y f(x)dx\ .$$

Sei K eine obere Schranke für $|f|$, also $|f(x)| \le K$ für alle $x \in [a,b]$. Dann folgt

$$|F(y) - F(s)| \;=\; \Big|\int_s^y f(x)dx\Big| \;\le\; \Big|\int_s^y |f(x)|dx\Big| \;\le\; K|y-s|\ .$$

(Warum braucht man die Betragsstriche um das Integral, obwohl man über $|f(x)|$ integriert?) Diese Bedingung nennt man „Lipschitz-stetig“. Sie impliziert, dass F stetig ist. (Wie?) ? ?

(ii) Betrachte den Differenzenquotienten

$$\phi(t) = \frac{F(t) - F(s)}{t - s}\ .$$

Wir müssen zeigen, dass $\lim_{t\to s} \phi(t)$ existiert und durch $f(s)$ gegeben ist. Wir können umformen:

$$\begin{aligned} \phi(t) - f(s) &= \frac{1}{t-s}\int_s^t f(x)\,dx - \frac{f(s)}{t-s}\int_s^t 1\,dx \\ &= \frac{1}{t-s}\int_s^t \big(f(x) - f(s)\big)\,dx \end{aligned}$$

Sei $\varepsilon > 0$ beliebig. Da f in s stetig ist, gibt es $\delta > 0$, so dass $|t-s| < \delta \Rightarrow |f(t) - f(s)| < \varepsilon$. Für $|t-s| < \delta$ können wir somit abschätzen

$$|\phi(t) - f(s)| \leq \left|\frac{1}{t-s}\int_s^t |f(x) - f(s)|\,dx\right| < \frac{1}{|t-s|}\,\varepsilon\,|t-s| \; .$$

Dies zeigt, dass $\lim_{t\to 0} \phi(t) = f(s)$.

□

Definition 7.4.3. Sei $f : [a,b] \to \mathbb{K}$ stetig. Wir nennen eine Funktion $F : [a,b] \to \mathbb{K}$ eine *Stammfunktion* zu f, falls F auf $[a,b]$ differenzierbar ist und $F' = f$ erfüllt.

Bemerkung 7.4.4.

(1) *Stammfunktionen existieren:* Satz 7.4.2 zeigt, dass jede stetige Funktion eine Stammfunktion besitzt.

(2) *Stammfunktionen sind nicht eindeutig:* Seien $F_1, F_2 : [a,b] \to \mathbb{K}$ zwei Stammfunktionen zu einer stetigen Funktion $f : [a,b] \to \mathbb{K}$. Dann gibt es ein $\lambda \in \mathbb{K}$, sodass für alle $x \in [a,b]$ gilt: $F_2(x) = F_1(x) + \lambda$. Dies folgt durch Ableiten von $F_2 - F_1$. (Details?) ?

Satz 7.4.5. (Hauptsatz der Differential- und Integralrechnung)
Sei $a < b$, und sei $f : [a,b] \to \mathbb{K}$ stetig. Für jede Stammfunktion $F : [a,b] \to \mathbb{K}$ von f und für alle $c, d \in [a,b]$ gilt, dass

$$\int_c^d f(x)\,dx = F(d) - F(c) \; .$$

Beweis. Nach Satz 7.4.2 ist $\Phi(x) = \int_a^x f(y)dy$ eine Stammfunktion von f. Nach Bemerkung 7.4.4 gilt $F(x) = \Phi(x) + \lambda$ für ein geeignetes (nicht von x abhängiges) $\lambda \in \mathbb{K}$. Dann rechnet man:

$$\int_c^d f(x)\,dx = \int_a^d f(x)\,dx - \int_a^c f(x)\,dx = \Phi(d) - \Phi(c) = F(d) - F(c) \; .$$

□

■ Wegen des Hauptsatzes benutzt man auch folgende Notation für Stammfunktionen F von f:

$$F = \int f(x)\,dx \quad \text{oder kürzer} \quad F = \int f \ .$$

Streng genommen ist dies Notationsmissbrauch, und man sollte sagen, dass $\int f$ die *Menge der Stammfunktionen* bezeichnet. Aber $F = \int f$ ist gebräuchlich (und praktisch) und sollte gelesen werden als „F hat die Eigenschaft, eine Stammfunktion von f zu sein".

Man nennt $\int f(x)dx$ auch ein *unbestimmtes Integral* und $\int_a^b f(x)dx$ ein *bestimmtes Integral*. Ein unbestimmtes Integral ist eine Funktion, die bis auf eine additive Konstante definiert ist. Ein bestimmtes Integral ist eine Zahl.

■ Eine weitere nützliche Notation ist $F(d) - F(c) = F(x)\big|_c^d$. Dann sagt der Hauptsatz:

Für $F = \int f$ gilt $\int_c^d f(x)dx = F(x)|_c^d$.

Endlich können wir also Integrale ausrechnen.

Beispiel 7.4.6.

(1) $\int x\,dx = \frac{1}{2}x^2$ und $\int_0^1 x\,dx = \frac{1}{2}x^2|_0^1 = \frac{1}{2}$

(2) Für $u \in \mathbb{R} \setminus \{-1\}$ gilt $\int x^u = \frac{1}{u+1}x^{u+1}$

Hier muss man allerdings aufpassen, was man meint. Für $u < 0$ ist der Integrand bei $x = 0$ nicht definiert. Für Intervalle $[a, b]$ mit $a < 0 < b$ (und für $u < 0$) ist x^u nicht in $\mathcal{R}(a, b)$ (selbst wenn man bei $x = 0$ von Hand den Wert 0 festlegt), da x^u nicht beschränkt ist. Die Gleichung

$$\int_a^b x^u\,dx = \left(\tfrac{1}{u+1}x^{u+1}\right)|_a^b$$

macht also nur Sinn für $a, b > 0$ oder $a, b < 0$.

(3) $\int x^{-1} = \log(x)$

Für bestimmte Integrale müssen die Grenzen hier beide positiv sein, da $\log(x)$ nur für positive x definiert ist. Z.B. $\int_2^3 x^{-1}dx = \log(x)\big|_2^3 = \log(3) - \log(2) = \log(3/2)$.

Aber $\int_{-3}^{-2} x^{-1}dx$ ist auch ein wohldefiniertes Integral. Um das zu berechnen, muss man eine andere Stammfunktion wählen: $\int x^{-1}dx = \log(-x)$. Das ist nun für negative x definiert, z.B. $\int_{-3}^{-2} x^{-1}dx = \log(2) - \log(3) = -\log(3/2)$.

Man kann diese beiden Fälle in einer Formel kombinieren:

$$\int x^{-1} dx \;=\; \log|x| \;.$$

(4) Für $\lambda \in \mathbb{K}$, $\lambda \neq 0$ gilt $\int e^{\lambda x} = \frac{1}{\lambda} e^{\lambda x}$.

(5) $\int \frac{1}{1+x^2} = \arctan(x)$

Dieses Beispiel, und auch Beispiel (3) illustrieren, warum integrieren „schwerer“ ist, als differenzieren. Bei Differentiation hatten wir Ketten- und Produktregel, und wir konnten die Ableitung wieder durch uns schon bekannte Funktionen ausdrücken. Bei Integrieren ist das typischerweise nicht so.

(6) Betrachte die Funktion $K : [-1,1] \to \mathbb{R}$, $K(x) = \sqrt{1-x^2}$. Dann gilt $x^2 + K(x)^2 = 1$, sodass der Graph $(x, K(x))$ gerade einen Halbkreis in der x-y-Ebene beschreibt. Also ist

$$\int_{-1}^{1} K(x)\, dx$$

der Flächeninhalt einer halben Kreisscheibe. Eine Weise, dieses Integral zu bestimmen, ist, dass jemand einem eine Stammfunktion verrät. Da Ableiten leichter ist als Integrieren, kann man dann prüfen, ob die Stammfunktion wirklich eine ist. Also:

Beh.: $\int K = \frac{1}{2}\big(\arcsin x + x\sqrt{1-x^2}\big)$

⌈ Z01A5. (Details?) ⌋ ?

Da $\arcsin(\pm 1) = \pm\frac{\pi}{2}$ erhalten wir für das bestimmte Integral

$$\int_{-1}^{1} K(x)\, dx \;=\; \frac{1}{2}\big(\arcsin x + x\sqrt{1-x^2}\big)\Big|_{-1}^{1} \;=\; \frac{\pi}{4} - \frac{-\pi}{4} = \frac{\pi}{2}\;.$$

In Analysis 1 haben wir π dadurch definiert, dass die erste positive Nullstelle der Cosinusfunktion bei $\pi/2$ liegt (siehe Z12A2 aus Analysis 1). Die Rechnung oben ist unsere erste geometrische Interpretation der so definierten Zahl π: Es ist die Fläche einer Kreisscheibe mit Radius 1.

(Später behandeln wir noch Kurvenintegrale, dann erhalten wir auch die Interpretation von 2π als Kreisumfang.)

Satz 7.4.7. (Substitutionsregel)
Seien $a < b$, $c < d$ und $\phi : [a,b] \to [c,d]$, $f : [c,d] \to \mathbb{K}$ Funktionen. Ferner sei f stetig und ϕ auf $[a,b]$ stetig differenzierbar. Dann gilt

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

Beweis. Zunächst einmal stellen wir fest, dass nach Satz 7.3.8 die Integrale auf beiden Seiten der Behauptung existieren. Sei $F(x) = \int_c^x f(y)\, dy$ eine Stammfunktion zu f auf $[c,d]$. Nach Satz 7.4.5 gilt

$$\int_{\phi(a)}^{\phi(b)} f(y)\, dy \;=\; F(x)\Big|_{\phi(a)}^{\phi(b)} \;=\; F(\phi(b)) - F(\phi(a)) \;.$$

Setze $G = F \circ \phi$. Dann $G'(x) = F'(\phi(x))\phi'(x) = f(\phi(x))\phi'(x)$, also ist G eine Stammfunktion für das Integral auf der linken Seite der Behauptung. Mit Satz 7.4.5 erhält man

$$\int_a^b f(\phi(x))\, \phi'(x)\, dx \;=\; G(x)\Big|_a^b \;=\; F(\phi(b)) - F(\phi(a)) \;.$$

□

Nebenbemerkung zur Information: Die obige Substitutionsregel und deren einfacher Beweis werden sich als 1-dimensionaler Unfall herausstellen. Wir werden in Analysis 3 eine analoge Regel für mehr-dimensionale Integrale kennenlernen. Dort darf die Funktion f beliebig (Lebesgue-)integrabel sein, die Funktion f muss dann aber (insbesondere) injektiv sein. Dann ist auch die Interpretation der Formel klarer: die Bijektion ϕ verzerrt das Volumen, und dieser Effekt wird durch den Term ϕ' kompensiert.

Beispiel 7.4.8.

(1) Sei $\phi : [a,b] \to [c,d]$ von der Form $\phi(x) = \lambda x + \mu$ für $\lambda, \mu \in \mathbb{R}$ (also ist ϕ eine affin-lineare Abbildung). Für $f : [c,d] \to \mathbb{K}$ sagt die Substitutionsregel:

$$\int_a^b f(\lambda x + \mu)\, \lambda\, dx \;=\; \int_{\lambda a + \mu}^{\lambda b + \mu} f(y)\, dy \;.$$

(2) Für $f(y) = 1/y$

$$\int_a^b \frac{\phi'(x)}{\phi(x)} dx \;=\; \int_{\phi(a)}^{\phi(b)} \frac{1}{y}\, dy \;=\; \log(x)\Big|_{\phi(a)}^{\phi(b)} \;=\; \log\big(\phi(b)/\phi(a)\big) \;.$$

Das sieht man auch direkt durch die Kettenregel: $\log(\phi(x))' = \phi'(x)/\phi(x)$, so dass man als unbestimmtes Integral erhält:

$$\int \frac{\phi'(x)}{\phi(x)} dx \;=\; \log(\phi(x)) \; .$$

Bemerkung 7.4.9. Hier sind zwei weitere Integrationstechniken.

(1) *Partielle Integration:* Die Funktionen $u, v : [a, b] \to \mathbb{K}$ seien stetig differenzierbar. Die Produktregel für Ableitungen gibt uns

$$(uv)' = u'v + v'u$$

Zusammen mit dem Hauptsatz erhalten wir damit

$$\int_a^b u'(x)v(x)\,dx \;=\; -\int_a^b u(x)v'(x)\,dx \;+\; \big(u(x)v(x)\big)\Big|_a^b \; .$$

Zum Beispiel

$$\begin{aligned} \int_a^b \sin(x)\,x\,dx \;&=\; \int_a^b \cos(x)\,dx \;-\; \big(\cos(x)\,x\big)\Big|_a^b \\ &=\; \big(\sin(x) - x\cos(x)\big)\Big|_a^b \; . \end{aligned}$$

Hier war $u(x) = -\cos(x)$, $v(x) = x$.

(2) *Partialbruchzerlegung:* Die Idee ist, einen Ausdruck von der Form

$$\frac{\text{Polynom}}{\text{anderes Polynom}}$$

umzuschreiben als

$$q(x) + \frac{a_1}{(x-p_1)^{m_1}} + \cdots + \frac{a_n}{(x-p_n)^{m_n}} \; ,$$

wobei $q(x)$ ein Polynom ist und $a_i, p_i \in \mathbb{C}$ (hier braucht man im Allgemeinen $\mathbb{C}$, auch wenn man mit $\mathbb{R}$ anfängt). In der Funktionentheorie lernt man, dass das immer möglich ist, hier betrachten wir das als Ansatz.

Jeden einzelnen Term kann man dann mit Beispiel 7.4.6 (2,3) integrieren:

$$\begin{aligned} \int \frac{1}{x-p} \;&=\; \log(x-p) \; , \\ \int \frac{1}{(x-p)^m} \;&=\; \frac{-1}{m-1}\,\frac{1}{(x-p)^{m-1}} \quad \text{für } m \geq 2 \; . \end{aligned}$$

Im ersten Fall muss man aufpassen, was passiert, wenn p nicht reell ist – wir hatten den Logarithmus bislang nicht für komplexe Argumente definiert ($\to$ Funktionentheorie).

Zum Beispiel

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

Also

$$\begin{aligned} \int \frac{1}{x^2-1}dx \;&=\; \frac{1}{2}\Big(-\int \frac{1}{x+1}dx + \int \frac{1}{x-1}dx\Big) \\ &=\; \tfrac{1}{2}\Big(-\log(x+1) + \log(x-1)\Big) \;=\; \log\sqrt{\frac{x-1}{x+1}} \;. \end{aligned}$$

(Das ist auch ein gutes Beispiel, dass man beim Einsetzen von Grenzen in uneigentliche Integrale aufpassen muss. Der Ausdruck oben funktioniert prima für $\int_2^3 (x^2-1)^{-1}dx$, aber was ist mit $\int_0^{1/2}(x^2-1)^{-1}dx$?) ?

Beispiel 7.4.10. Es gilt

$$\int \frac{1}{\cos x}dx \;=\; \int \frac{\cos x}{1-(\sin x)^2}dx \;=\; \ldots \;=\; \frac{1}{2}\log\frac{1+\sin(x)}{1-\sin(x)} \;.$$

(Jetzt kann man die Antwort natürlich durch Ableiten nachprüfen. Aber wie kommt man bloß darauf? *Hinweis:* In „…“ macht man erst eine Substitution und dann eine Partialbruchzerlegung.) ?

(Sowohl $\cos(x)$ als auch $\sin(x)$ sind ja periodisch mit Periode 2π. Damit gibt der Ausdruck oben also $\int_0^{2\pi} 1/(\cos x)dx = 0$. Stimmt das?) ?

Der Hauptsatz (Satz 7.4.5) gilt auch allgemeiner. Hier geben wir eine Variante.

Satz 7.4.11. Sei $f \in \mathcal{R}(a,b)$. Angenommen, f ist stückweise stetig, d.h. es gibt eine Zerlegung $a = x_0 < x_1 < \cdots < x_n = b$, so dass f auf jedem offenen Teilintervall (x_i, x_{i+1}) stetig ist.

Sei $F : [a,b] \to \mathbb{K}$ stetig und an jeder Stelle $s \in (x_i, x_{i+1})$, $i = 0, \ldots, n-1$ differenzierbar mit $F'(s) = f(s)$. Dann gilt

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

(Warum wäre es nicht so schlau, Stetigkeit von f auf den abgeschossenen Intervallen $[x_i, x_{i+1}]$ zu fordern?) ?

Der Beweis ist eine Übung.

Beispiel 7.4.12. Sei sgn : $\mathbb{R} \to \mathbb{R}$ die Vorzeichenfunktion

$$\operatorname{sgn}(x) = \begin{cases} 1 & ; x > 0 \\ \text{egal} & ; x = 0 \\ -1 & ; x < 0 \end{cases} \quad .$$

Diese ist in jedem Intervall $[a,b]$ stückweise stetig und es gilt

$$\int_a^b \operatorname{sgn}(x)\, dx \;=\; |x|\,\Big|_a^b \quad .$$

Nebenbemerkung zur Information: Man kann den Hauptsatz noch weiter verallgemeinern, zum Beispiel auf Funktionen, die in jedem Punkt einen links- und einen rechtsseitigen Grenzwert haben. Solche Funktionen können abzählbar viele Unstetigkeitsstellen haben. Siehe [Kap. 11.4 aus Königsberger 1].

7.5 Uneigentliche Integrale

(Siehe z.B. [Kap. 11.9 aus Königsberger 1] oder [§20 aus Forster 1].)

In diesem Kapitel steht $\mathbb{K}$ wieder für $\mathbb{R}$ oder $\mathbb{C}$, und $\mathcal{R}(a,b)$ steht für $\mathcal{R}_{\mathbb{K}}(a,b)$ (und es ist bei dieser Notation wieder $a < b$ vorausgesetzt).

Definition 7.5.1. Sei $-\infty < A < B \leq \infty$ und $f : [A,B) \to \mathbb{K}$ so, dass $f \in \mathcal{R}(A,b)$ für alle $A < b < B$. Falls $\lim_{b \nearrow B} \int_A^b f(x)dx$ existiert (in $\mathbb{K}$), sagen wir $\int_A^B f(x)dx$ *konvergiert* und setzen

$$\int_A^B f(x)dx \overset{\text{Def.}}{=} \lim_{b \nearrow B} \int_A^b f(x)dx \; .$$

Falls $\int_A^B |f(x)|dx$ konvergiert, sagen wir $\int_A^B f(x)dx$ *konvergiert absolut.*

▪ Genauso definiert man für $-\infty \leq A < B < \infty$ und $f : (A,B] \to \mathbb{K}$, dass

$$\int_A^B f(x)dx \overset{\text{Def.}}{=} \lim_{a \searrow A} \int_a^B f(x)dx \; .$$

(Wie oben muss hier $f \in \mathcal{R}(a,B)$ sein für alle $A < a < B$.)

▪ Für $-\infty \leq A < B \leq \infty$ und $f : (A,B) \to \mathbb{K}$ wählt man einen Punkt c mit $A < c < B$ und definiert

$$\int_A^B f(x)dx \overset{\text{Def.}}{=} \int_A^c f(x)dx + \int_c^B f(x)dx \; .$$

(Hier muss $f \in \mathcal{R}(a,b)$ für alle $A < a < b < B$ gelten.) Diese Definition ist unabhängig von der Wahl von c. (Warum?) ?

■ Riemann-Integrale der drei oben diskutierten Arten nennt man *uneigentliche Integrale*. Das bisherige Riemann-Integral nennt man dann manchmal auch *eigentliches Integral*.

Bemerkung 7.5.2. Sei $-\infty < A < B < \infty$ und $f \in \mathcal{R}(A,B)$. Das Symbol $\int_A^B f(x)\,dx$ kann nun mehrere Bedeutungen haben:

- Die ursprüngliche Bedeutung als eigentliches Integral wie in Definition 7.3.4.
- Als eine der drei obigen Sorten von uneigentlichem Integral. Z.B.

$$\int_A^B f(x)dx \; = \; \lim_{b \nearrow B} \int_A^b f(x)dx \; .$$

Alle vier Interpretationsmöglichkeiten geben den gleichen Wert für $\int_A^B f(x)\,dx$. (Was sind die anderen beiden Möglichkeiten im zweiten Punkt? Warum geben alle vier das gleiche Ergebnis?) ?

Beispiel 7.5.3.

(1) Für $f(x) = x^{-\frac{1}{2}}$ gilt:

$$\int_1^\infty \frac{1}{\sqrt{x}} dx = \lim_{b \nearrow \infty} 2\sqrt{x}\Big|_1^b = \lim_{b \nearrow \infty} \big(2\sqrt{b} - 2\big) \qquad \text{konvergiert nicht,}$$

$$\int_0^1 \frac{1}{\sqrt{x}} dx = \lim_{a \searrow 0} 2\sqrt{x}\Big|_a^1 = \lim_{a \searrow 0} \big(2 - 2\sqrt{a}\big) = 2 \quad \text{konvergiert absolut.}$$

(2) Für $f(x) = x^{-2}$ gilt:

$$\int_1^\infty \frac{1}{x^2} dx = \lim_{b \nearrow \infty} \frac{-1}{x}\Big|_1^b = \lim_{b \nearrow \infty} \Big(- \frac{1}{b} + 1 \Big) = 1 \quad \text{konvergiert absolut,}$$

$$\int_0^1 \frac{1}{x^2} dx = \lim_{a \searrow 0} \Big(- 1 + \frac{1}{a} \Big) \qquad \text{konvergiert nicht.}$$

(3) Ein Beispiel bei dem beide Grenzen unendlich sind:

$$\int_{-\infty}^\infty \frac{1}{1+x^2} dx \; = \; \pi \qquad \text{konvergiert absolut.}$$

(Details?) ?

(4) Hier ein Beispiel für nicht-absolute Konvergenz:

Beh: $\int_0^\infty x^{-1} \sin(x)\, dx$ konvergiert, aber nicht absolut.

Die Funktion $\sin(x)/x$ lässt sich stetig nach 0 fortsetzen (durch 1). Somit müssen wir nur die Grenze ∞ getrennt behandeln. Setze

$$I(b) = \int_0^b \frac{\sin(x)}{x}\, dx \qquad \text{und} \qquad J(b) = \int_0^b \left|\frac{\sin(x)}{x}\right| dx \ .$$

Es ist zu zeigen, dass $\lim_{b\to\infty} I(b)$ existiert, und dass $\lim_{b\to\infty} J(b)$ nicht existiert.

Der Beweis ist eine Übung, für die der nächste Satz nützlich ist.

Nebenbemerkung zur Information: Das Integral $I(b) = \int_0^b x^{-1} \sin(x)\, dx$ hat keine geschlossene Form. Damit meint man, dass $I(b)$ nicht durch „elementare Funktionen" wie sin, log, etc. ausgedrückt werden kann. Wenn man möchte, kann man es aber als sogenannte Hypergeometrische Funktion schreiben:

$$I(b) \ = \ b \cdot {}_1F_2\big(\tfrac{1}{2}\,;\, \tfrac{3}{2}, \tfrac{3}{2}\,;\, -b^2/4\big) \ .$$

Satz 7.5.4. (Majorantenkriterium)
Sei $-\infty < a < B \leq \infty$, seien $f : [a, B) \to \mathbb{K}$ und $g : [a, B) \to \mathbb{R}$ mit

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

auf $[a, B)$ und $f, g \in \mathcal{R}(a, b)$ für alle $a < b < B$. Falls $\int_a^B g(x)dx$ konvergiert, so konvergiert auch $\int_a^B f(x)dx$.

Beweis. Sei $G = \int_a^B g(x)dx$. Sei $\varepsilon > 0$. Nach der Definition des Grenzwertes gibt es $R > a$, so dass für alle $R < u < B$ gilt

$$\Big| \int_a^u g(x)dx - G \Big| < \varepsilon \ .$$

Es folgt, dass für alle u, v mit $R < u < v < B$:

$$\begin{aligned}\Big| \int_u^v g(x)dx \Big| \ &= \ \Big| \int_a^v g(x)dx - G + G - \int_a^u g(x)dx \Big| \\ &\leq \ \Big| \int_a^v g(x)dx - G \Big| \ + \ \Big| \int_a^u g(x)dx - G \Big| \ < \ 2\varepsilon\end{aligned}$$

Also gilt

$$\Big| \int_u^v f(x)dx \Big| \leq \int_u^v |f(x)|dx \leq \int_u^v g(x)dx < 2\varepsilon \ .$$

Setze $F(x) = \int_a^x f(y)dy$, so dass $|F(v) - F(u)| = |\int_u^v f(x)dx|$.

- Falls $B = \infty$ betrachte die Folge $(F_m)_m$ mit $F_m = F(m)$ und $m > a$. Für $m, n > R$ gilt $|F_m - F_n| < 2\varepsilon$.
- Falls $B < \infty$ betrachte die Folge $(F_m)_m$ mit $F_m = F(B - \frac{1}{m})$ und $B - \frac{1}{m} > a$. Für $\frac{1}{m}, \frac{1}{n} < B - R$ gilt $|F_m - F_n| < 2\varepsilon$.

In beiden Fällen ist $(F(m))_m$ eine Cauchy-Folge und konvergiert in $\mathbb{K}$. Ferner gilt $\lim_{x\to B} F(x) = \lim_{m\to\infty} F_m$ (Warum?). Also existiert $\lim_{x\to B} F(x)$, und somit konvergiert $\int_a^B f(x)dx$. ? $\square$

■ Das Majorantenkriterium gibt es genauso für $-\infty \leq A < b < \infty$ und Funktionen $f : (A, b] \to \mathbb{K}$ und $g : (A, b] \to \mathbb{R}$ mit $|f(x)| \leq g(x)$ auf $(A, b]$ und $f, g \in \mathcal{R}(a, b)$ für alle $A < a < b$.

Bemerkung 7.5.5.

(1) Wenn das Integral $\int_a^B f(x)dx$ nach dem Majorantenkriterium konvergiert, dann konvergiert es auch absolut. (Warum?) ?

(2) Folgende negative Formulierung des Majorantenkriteriums ist auch nützlich: Seinen f, g wie im Majorantenkriterium, mit $g \geq 0$, aber jetzt gelte $|f(x)| \geq g(x)$ auf $[a, B)$. Falls $\int_a^B g(x)dx$ divergiert, so divergiert auch $\int_a^B |f(x)|dx$. (Warum?) ?

Insbesondere kann $\int_a^B f(x)dx$ dann nicht absolut konvergieren. (Aber es kann noch konvergieren, siehe Beispiel 7.5.3 (4)).

Satz 7.5.6. (Integralkriterium)
Sei $f : [1, \infty) \to \mathbb{R}$ monoton fallend mit $f(x) \geq 0$ auf $[1, \infty)$. Setze

$$a_n = \sum_{k=1}^{n} f(k) \; - \; \int_1^{n+1} f(x)dx \; .$$

Dann existiert $\lim_{n\to\infty} a_n$ und es gilt

$$0 \leq \lim_{n\to\infty} a_n \leq f(1) \; .$$

Beweis. Da f monoton fallend ist, gilt $f(k) \geq f(x) \geq f(k+1)$ für alle $k \leq x \leq k+1$. Nach Lemma 7.3.5 folgt $f(k) \geq \int_k^{k+1} f(x) \geq f(k+1)$. Also gilt

$$a_n - a_{n-1} = f(n) - \int_n^{n+1} f(x)\, dx \; \geq \; 0 \; ,$$

und somit ist die Folge $(a_n)_{n\in\mathbb{N}}$ monoton steigend. Ferner gilt

$$0 \le f(k) - \int_k^{k+1} f(x)dx \le f(k) - f(k+1)\ .$$

Summiert man dies von $k = 1, \dots, n$, erhält man $0 \le a_n \le f(1) - f(n+1)$. Da $f(x) \ge 0$, folgt, dass a_n nach oben durch $f(1)$ beschränkt ist. Somit konvergiert die Folge $(a_n)_{n\in\mathbb{N}}$ gegen einen Wert a (Satz 4.3.2) und für a gilt $0 \le a \le f(1)$. □

Korollar 7.5.7. Sei f wie in Satz 7.5.6. Die Reihe $\sum_{k=1}^{\infty} f(k)$ konvergiert genau dann, wenn das Integral $\int_1^{\infty} f(x)dx$ konvergiert.

(Wieso gilt das Korollar?) ?

Beispiel 7.5.8. Wir haben folgende Ableitungen:

$$(\log(x))' = \frac{1}{x}\quad,\quad (\log(\log(x)))' = \frac{1}{x\log x}\quad,\quad (-1/\log(x))' = \frac{1}{x\,(\log x)^2}\ .$$

Daraus ergeben sich die folgenden Aussagen über uneigentliche Integrale und Reihen:

(1)
$$\int_1^{\infty} \frac{1}{x}dx = \lim_{b\to\infty} \log x\Big|_1^b \quad \text{konvergiert nicht,}$$
und somit divergiert auch $\sum_{k=1}^{\infty} 1/k$.

(2)
$$\int_2^{\infty} \frac{1}{x\log x}dx = \lim_{b\to\infty} \log(\log x)\Big|_2^b \quad \text{konvergiert nicht,}$$
und somit divergiert auch $\sum_{k=2}^{\infty} 1/(k\log k)$.

(3)
$$\int_2^{\infty} \frac{1}{x(\log x)^2}dx = \lim_{b\to\infty} (-1/\log x)\Big|_2^b \quad \text{konvergiert,}$$
und somit konvergiert auch $\sum_{k=2}^{\infty} 1/(k(\log k)^2)$.

(4) Der Grenzwert

$$\gamma \stackrel{\text{Def.}}{=} \lim_{n\to\infty} \Big(1 + \frac{1}{2} + \cdots + \frac{1}{n} - \log(n)\Big)$$

existiert und erfüllt $0 \le \gamma \le 1$. (Man benutzt $\lim_{n\to\infty} \log\frac{n+1}{n} = 0$ – Wo?). Die Konstante γ heisst *Euler-Konstante* und hat den ungefähren Wert $\gamma = 0.57721566\ldots$?

7.6 Kurvenintegrale

(Siehe z.B. [Kap. 6, Def. 6.26 etc. aus Rudin] oder [Kap. 12 aus Königsberger 1].)

In diesem Kapitel verwenden wir wieder die Konvention, dass $a < b$ gilt, so dass das Intervall $[a, b]$ nicht entartet ist.

Definition 7.6.1. Sei $f = (f_1, \dots, f_d) : [a, b] \to \mathbb{R}^d$. Wir schreiben $f \in \mathcal{R}(a, b)$, falls $f_i \in \mathcal{R}(a, b)$ für $i = 1, \dots, d$ und sagen f *ist (Riemann-) integrierbar*. In diesem Fall definieren wir

$$\int_a^b f(x)dx \overset{\text{Def.}}{=} \Big(\int_a^b f_1(x)dx \ , \ \dots \ , \ \int_a^b f_d(x)dx \Big) \in \mathbb{R}^d \ .$$

■ Mit $\mathbb{C} = \mathbb{R}^2$ ist die Erweiterung des Integrals von $\mathbb{R}$-wertigen Funktionen auf $\mathbb{C}$-wertige Funktionen aus Bemerkung 7.3.14 ein Spezialfall der obigen Definition.

In Kapitel 6 aus Analysis 1 hatten wir die Ableitung von vektorwertigen Funktionen definiert. Der Hauptsatz der Differential- und Integralrechnung gilt auch in diesem Fall:

Satz 7.6.2. Sei $F : [a, b] \to \mathbb{R}^d$ auf $[a, b]$ stetig differenzierbar. Mit $f = F' : [a, b] \to \mathbb{R}^d$ gilt dann

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

Beweis. Seien $F_i : [a, b] \to \mathbb{R}$ die Komponentenfunktionen von F, also $F(x) = (F_1(x), \dots, F_d(x))$. Nach Satz 6.5.2 gilt die erste Gleichheit in

$$F'(x) = \big(F_1'(x), \dots, F_d'(x) \big) = \big(f_1(x), \dots, f_d(x) \big) \ .$$

Die zweite Gleichheit gilt per Annahme. Jetzt braucht man nur noch den Hauptsatz (Satz 7.4.5) komponentenweise anwenden. □

Lemma 7.6.3. Sei $f : [a, b] \to \mathbb{R}^d$ integrierbar. Dann ist auch $|f| : [a, b] \to \mathbb{R}$ integrierbar und es gilt

$$\Big| \int_a^b f(x)dx \Big| \ \leq \ \int_a^b |f(x)|dx \ .$$

Beweis. Schreibe $f = (f_1, \dots, f_d)$. Dann gilt

$$|f(x)| \ = \ \sqrt{f_1(x)^2 + \dots + f_d(x)^2} \ .$$

Da alle $f_i \in \mathcal{R}(a,b)$, ist nach Satz 7.3.10 und Satz 7.3.12 auch $|f| \in \mathcal{R}(a,b)$.

Setze $v = \int_a^b f(x)dx \in \mathbb{R}^d$. Für $v = 0$ ist nichts zu zeigen. Sei also $v \neq 0$. Wir können schreiben

$$|v|^2 = \sum_{i=1}^d v_i^2 = \sum_{i=1}^d \int_a^b v_i f_i(x) dx \ .$$

Jetzt benutzen wir die Cauchy-Schwarzsche Ungleichung (Lemma 6.5.5):

$$\big|\langle v, f(x)\rangle\big| \ = \ \Big|\sum_{i=1}^d v_i\, f_i(x)\Big| \ \leq \ |v|\,|f(x)| \ .$$

Damit erhalten wir insgesamt

$$|v|^2 \ = \ \int_a^b \Big(\sum_{i=1}^d v_i\, f_i(x)\Big)dx \ \overset{(*)}{\leq} \ \int_a^b |v|\,|f(x)|dx \ .$$

Die Ungleichung $(*)$ folgt aus der obigen punktweisen Ungleichung (Warum ist es egal, dass die Betragsstriche fehlen?) und Lemma 7.3.5 für $\mathbb{R}$-wertige Integrale. Die Aussage folgt, indem man durch $|v|$ teilt. ? □

Definition 7.6.4. Eine *parametrisierte Kurve im* $\mathbb{R}^d$ oder kurz *Kurve im* $\mathbb{R}^d$ ist eine stetige Abbildung $\gamma : [a,b] \to \mathbb{R}^d$. Eine Kurve heißt *Bogen*, falls γ injektiv ist, und *geschlossen* falls $\gamma(a) = \gamma(b)$.

Beispiel 7.6.5. Für $L > 0$ betrachte

$$\gamma : [0, L] \to \mathbb{R}^2 \quad , \quad t \mapsto \big(\sin(t)\,,\, \sin(2t)\big)$$

Für alle $L > 0$ ist γ eine Kurve im $\mathbb{R}^2$. Ferner gilt

- $L = \pi, 2\pi, 3\pi, \dots$: γ ist eine geschlossene Kurve (und sonst nicht).
- $L < \pi$: γ ist ein Bogen (und sonst nicht).

(Warum?) ?

Sei $\gamma : [a,b] \to \mathbb{R}^k$ eine Kurve und $P = \{x_0, \dots, x_n\}$ eine Zerlegung (auch genannt *Parition*) von $[a,b]$ (also $a = x_0 < x_1 < \dots < x_n = b$). Wir definieren

$$\Lambda(P,\gamma) \ \overset{\text{Def.}}{=} \ \sum_{i=0}^{n-1} \big|\,\gamma(x_{i+1}) - \gamma(x_i)\,\big| \ .$$

Definition 7.6.6. Eine Kurve $\gamma : [a,b] \to \mathbb{R}^k$ heisst *rektifizierbar*, falls

$$\Lambda(\gamma) \overset{\text{Def.}}{=} \sup\left\{\Lambda(P,\gamma) \,\middle|\, P \text{ Zerlegung von } [a,b]\right\} < \infty .$$

In diesem Fall heisst $\Lambda(\gamma)$ die *Länge* von γ .

Beispiel 7.6.7.

(1) Betrachte $\gamma : [0,1] \to \mathbb{R}^d$, $t \mapsto (t, 0, \ldots, 0)$. Dann gilt für jede Zerlegung P, dass $\Lambda(P,\gamma) = 1$ (Warum?). Also ist γ rektifizierbar mit $\Lambda(\gamma) = 1$. ?

(2) Die Koch-Kurve im $\mathbb{R}^2$ erhält man durch rekursives Anwenden der folgenden Vorschrift. 0) Beginne mit dem Intervall $[0,1]$ auf der x-Achse im $\mathbb{R}^2$. 1) Unterteile das Intervall in drei gleiche Teile. 2) Ersetze das mittlere Interval durch zwei gleich lange Intervalle, die die zwei anderen Seiten eines gleichseiten Dreiecks darstellen. 3) Wiederhole ab Schritt 1 für alle so gewonnen Intervalle.

Siehe z.B. „de.wikipedia.org/wiki/Koch-Kurve“ auf Wikipedia für Bilder.

Achtung: Die obige Vorschrift lässt wichtige Fragen unbeantwortet. Konvergiert die Vorschrift in irgend einem Sinn? Kann man die Koch-Kurve als Abbildung $\gamma : [0,1] \to \mathbb{R}^2$ darstellen? Wenn ja, ist γ dann auch stetig? D.h. ist die Koch-Kurve eine parametrisierte Kurve im Sinne von Definition 7.6.4?

Wir nehmen für dieses Beispiel an, dass es ein solches γ gibt und überlegen informell, welche Eigenschaften wir erwarten. Dass γ existiert (und stetig ist), können wir dann besser mit den Methoden des nächsten Kapitels zeigen.

Informelle Betrachtung: Das Intervall in Schritt 0) hat Länge 1. Bei Schritt 1 und 2 ersetzt man es durch 4 Intervalle der Länge $\frac{1}{3}$. Sei Λ_n die Länge Λ_n nach n Iterationsschritten. Dann gilt $\Lambda_{n+1} = \frac{4}{3}\Lambda_n$. Mit $\Lambda_0 = 1$ erhält man $\Lambda_n = (4/3)^n$. Für $n \to \infty$ divergiert dieser Ausdruck. Wir erwarten also, dass die Koch-Kurve nicht rektifizierbar ist.

Bemerkung 7.6.8.

1. *(Spur einer Kurve)* Die *Spur einer Kurve* $\gamma : [a,b] \to \mathbb{R}^d$ bezeichnet die Bildmenge einer Kurve, also die Teilmenge

$$\left\{\gamma(t) \,\middle|\, t \in [a,b]\right\} \subset \mathbb{R}^d .$$

Die Spur enthält weniger Information als die parametrisierte Kurve. Z.B. kann eine Kurve „mehrmals die gleiche Spur durchlaufen“ wie in Beispiel 7.6.5. Das zeigt auch, dass die Spur im Allgemeinen nicht ausreicht, um die Länge einer Kurve zu bestimmen.

2. *(Umparametrisierung)* Sei $\gamma : [a,b] \to \mathbb{R}^d$ eine rektifizierbare Kurve. Sei $\varphi : [A,B] \to [a,b]$ stetig, bijektiv, und monoton steigend. Setze

$$\delta = \gamma \circ \varphi \ : \ [A,B] \to \mathbb{R}^d \ .$$

Dann haben γ und δ die gleiche Spur (Warum)? Ferner haben wir folgende Mengengleichheit in $\mathbb{R}$: ?

$$\big\{\Lambda(P,\gamma) \,\big|\, P \text{ Zerlegung von } [a,b]\,\big\} \ = \ \big\{\Lambda(P',\delta) \,\big|\, P' \text{ Zerlegung von } [A,B]\,\big\}$$

(Warum sind diese Mengen gleich? *Hinweis:* Eine Zerlegung von $[A,B]$ wird mit der Bijektion φ auf eine Zerlegung von $[a,b]$ abgebildet. Was passiert mit $\Lambda(\dots)$ unter dieser Abbildung?) ?

Also ist auch δ rektifizierbar und γ und δ haben die gleiche Länge: $\Lambda(\gamma) = \Lambda(\delta)$.

Satz 7.6.9. Sei $\gamma : [a,b] \to \mathbb{R}^d$ stetig differenzierbar. Dann ist γ rektifizierbar und es gilt

$$\Lambda(\gamma) \ = \ \int_a^b \big|\gamma'(t)\big| \, dt \ .$$

Beweis. Für $a \le x < y \le b$ gilt

$$\big|\gamma(y) - \gamma(x)\big| \ \overset{\text{Satz 7.6.2}}{=} \ \Big|\int_x^y \gamma'(t)\, dt\Big| \ \overset{\text{Lem. 7.6.3}}{\le} \ \int_x^y \big|\gamma'(t)\big| \, dt \ .$$

(Die Funktion $\gamma' : [a,b] \to \mathbb{R}^d$ ist integrierbar, da sie stetig ist.) Also gilt für jede Zerlegung $P = \{x_0, x_1, \dots, x_n\}$ von $[a,b]$, dass

$$\Lambda(P,\gamma) \ \le \ \int_a^b \big|\gamma'(t)\big| \, dt \ .$$

(Warum?) Insbesondere ist γ rektifizierbar. ?

Also existiert die Länge $\Lambda(\gamma)$. Es bleibt zu zeigen, dass sie durch das Integral in der Behauptung gegeben ist, also dass die Ungleichung oben in der Tat die kleinste obere Schranke für die Werte $\Lambda(P,\gamma)$ darstellt.

Sei $\varepsilon > 0$ beliebig. Die Funktion $\gamma' : [a,b] \to \mathbb{R}^d$ ist gleichmäßig stetig (Warum?). Also gibt es $\delta > 0$, so dass für $s,t \in [a,b]$ gilt ?

$$|s-t| < \delta \quad \Rightarrow \quad |\gamma'(s) - \gamma'(t)| < \varepsilon \ .$$

Sei P eine beliebige Zerlegung von $[a,b]$, so dass $\Delta_i = x_{i+1} - x_i < \delta$. Dann gilt

$$\begin{aligned}
\big|\gamma(x_{i+1}) - \gamma(x_i)\big| &= \Big|\int_{x_i}^{x_{i+1}} \gamma'(t)\,dt\Big| \\
&= \Big|\int_{x_i}^{x_{i+1}} \big(\gamma'(x_i) - \big(\gamma'(x_i) - \gamma'(t)\big)\big)dt\Big| \\
&\geq \underbrace{\Big|\int_{x_i}^{x_{i+1}} \gamma'(x_i)dt\Big|}_{=\,|\gamma'(x_i)|\,\Delta_i} - \underbrace{\Big|\int_{x_i}^{x_{i+1}} \big(\gamma'(t) - \gamma'(x_i)\big)dt\Big|}_{=\,(*)} \ .
\end{aligned}$$

Wir brauchen eine Abschätzung für den Ausdruck $(*)$:

$$(*) \overset{\text{Lem. 7.6.3}}{\leq} \int_{x_i}^{x_{i+1}} \underbrace{\big|\gamma'(t) - \gamma'(x_i)\big|}_{<\varepsilon} dt \ < \ \varepsilon\,\Delta_i \ .$$

Soweit wissen wir also, dass

$$\big|\gamma(x_{i+1}) - \gamma(x_i)\big| \ > \ |\gamma'(x_i)|\,\Delta_i \ - \ \varepsilon\,\Delta_i \ . \qquad (**)$$

Als nächstes vergleichen wir $|\gamma'(x_i)|\,\Delta_i$ und $\int_{x_i}^{x_{i+1}} |\gamma'(t)|dt$. Das geht so ähnlich wie oben:

$$\begin{aligned}
\int_{x_i}^{x_{i+1}} \big|\gamma'(t)\big|\,dt &= \int_{x_i}^{x_{i+1}} \Big|\gamma'(x_i) - \big(\gamma'(x_i) - \gamma'(t)\big)\Big|dt \\
&\leq \underbrace{\int_{x_i}^{x_{i+1}} \big|\gamma'(x_i)\big|dt}_{=\,|\gamma'(x_i)|\,\Delta_i} + \underbrace{\int_{x_i}^{x_{i+1}} \big|\gamma'(t) - \gamma'(x_i)\big|dt}_{<\,\varepsilon\,\Delta_i} \ .
\end{aligned}$$

Somit können wir in $(**)$ weiter abschätzen:

$$\big|\gamma(x_{i+1}) - \gamma(x_i)\big| \ > \ \int_{x_i}^{x_{i+1}} \big|\gamma'(t)\big|\,dt \ - \ 2\varepsilon\,\Delta_i \ .$$

Per Summe über i erhalten wir insgesamt

$$\Lambda(P,\gamma) \ > \ \int_a^b \big|\gamma'(t)\big|\,dt \ - \ 2\varepsilon\,(b-a) \ .$$

Daraus folgt, dass $\int_a^b \big|\gamma'(t)\big|\,dt$ in der Tat die kleinste obere Schranke für die Menge $\{\Lambda(P,\gamma) | P \text{ Zerlegung }\}$ ist. (Wieso folgt das?) $\square$?

Beispiel 7.6.10. Für $z \in \mathbb{C}$, $|z| = 1$ gibt es genau ein $t \in [0, 2\pi)$, so dass $z = e^{it}$ (Analysis 1 Zettel 12 Aufgabe 3, siehe auch [Satz 8.7 aus Rudin]). Also ist $\gamma : [0, \alpha] \to \mathbb{C}$, $\gamma(t) = e^{it}$ injektiv für $0 \leq \alpha < 2\pi$. Ausgedrückt in Koordinaten von $\mathbb{C} = \mathbb{R}^2$ können wir auch schreiben $\gamma(t) = (\cos(t), \sin(t))$. Also ist $\gamma'(t) = (-\sin(t), \cos(t))$ stetig und nach Satz 7.6.9 gilt

$$\Lambda(\gamma) = \int_0^\alpha |\gamma'(t)| dt = \int_0^\alpha 1 \, dt = \alpha \ .$$

Dieses Integral berechnet die Länge des Einheitskreisbogens zwischen den Punkten $(1, 0)$ und $(\cos \alpha, \sin \alpha)$. Dies zeigt, dass unsere Definition von sin und cos über die Exponentialreihe mit der geometrischen Definition übereinstimmt.

8 Folgen und Reihen von Funktionen

Hier betrachten wir Funktionen mit Werten in $\mathbb{C}$. Die Aussagen schließen dann Funktionen mit Werten in $\mathbb{R}$ ein.

8.1 Punktweise Konvergenz

(Siehe [Kap. 7 aus Rudin] und [Kap. 11.1 aus Königsberger 1].)

In diesem Abschnitt schauen wir uns ein paar Beispiele zu punktweiser Konvergenz von Funktionen an (siehe nächste Definition), um zu illustrieren, dass sich dieser Konvergenzbegriff nicht so gut verhält.

Definition 8.1.1. Sei E eine Menge und $(f_n)_{n\in\mathbb{N}}$ mit $f_n : E \to \mathbb{C}$ eine Folge von Funktionen.

(i) Angenommen, für jedes $x \in E$ konvergiert die Folge $(f_n(x))_{n\in\mathbb{N}}$ in $\mathbb{C}$. Dann definieren wir $f : E \to \mathbb{C}$ durch

$$f(x) \stackrel{\text{Def.}}{=} \lim_{n\to\infty} f_n(x) \ .$$

Wir sagen $(f_n)_{n\in\mathbb{N}}$ *konvergiert punktweise auf E gegen f*. Die Funktion f heißt *Grenzwert* oder *Grenzfunktion* von $(f_n)_{n\in\mathbb{N}}$. Wir schreiben $f = \lim_{n\to\infty} f_n$.

(ii) Angenommen, für jedes $x \in E$ konvergiert die Reihe $\sum_{n=1}^{\infty} f_n(x)$ in $\mathbb{C}$. Dann definieren wir $g : E \to \mathbb{C}$ als

$$g(x) \stackrel{\text{Def.}}{=} \sum_{n=1}^{\infty} f_n(x)$$

und nennen g die *Summe der Reihe* $\sum_{n=1}^{\infty} f_n$.

In den nächsten beiden Kapiteln beschäftigen wir uns mit der Frage, was mit Grenzwerten, Stetigkeit, Differenzieren und Integrieren passiert, wenn wir bei Funktionen einen Grenzübergang $\lim_{n\to\infty} f_n$ machen. Um zu sehen, was uns da alles passieren kann, schauen wir uns Beispiele an.

Beispiel 8.1.2.

(1) *Grenzwerte vertauschen nicht:*

Sei $E = [1, \infty)$ und $f_n : E \to \mathbb{R}$, $f_n(x) = \frac{x}{x+n}$. Dann gilt

$$\begin{aligned} &\text{für fixes } n \in \mathbb{N} : \qquad && \lim_{x\to\infty} f_n(x) \;=\; 1 \ , \\ &\text{für fixes } x \in E : && \lim_{n\to\infty} f_n(x) \;=\; 0 \ . \end{aligned}$$

Setze $f = \lim_{n\to\infty} f_n$. Dann $f = 0$. Also:

$$\lim_{x\to\infty}\left(\lim_{n\to\infty} f_n(x)\right) \;=\; 0 \;\neq\; 1 \;=\; \lim_{n\to\infty}\left(\lim_{x\to\infty} f_n(x)\right) .$$

(2) *Alle Funktionen in einer Folge sind stetig, aber die Grenzfunktion nicht:*

Sei $E = [0,1]$, $f_n : E \to \mathbb{R}$, $f_n(x) = x^n$. Dann ist f_n stetig auf E, aber es gilt

$$\lim_{n\to\infty} f_n(x) = \begin{cases} 0 & ; x < 1 \\ 1 & ; x = 1 \end{cases} .$$

(3) *Alle Funktionen in einer Folge sind differenzierbar, und die Grenzfunktion ist differenzierbar, aber die Grenzfunktion der Ableitungen ist nicht die Ableitung der Grenzfunktion:*

Sei $E = [0,1]$, $f_n : E \to \mathbb{R}$, $f_n(x) = \frac{1}{n+1}x^{n+1}$. Dann gilt für $x \in [0,1]$, dass $f(x) := \lim_{n\to\infty} f_n(x) = 0$. Aber $f_n'(x) = x^n$, und somit $\lim_{n\to\infty} f_n'(1) = 1$ (siehe Beispiel 2). Also

$$\lim_{n\to\infty} f_n' \neq (\lim_{n\to\infty} f_n)' .$$

(4) *Alle Funktionen in einer Folge sind differenzierbar, und die Grenzfunktion ist differenzierbar, aber der Grenzwert der Ableitungen existiert in keinem Punkt:*

Sei $E = \mathbb{R}$, $f_n : E \to \mathbb{R}$,

$$f_n(x) = \frac{\sin(nx)}{\sqrt{n}} .$$

Dann gilt für $x \in \mathbb{R}$, dass $f(x) := \lim_{n\to\infty} f_n(x) = 0$. Aber $f_n'(x) = \sqrt{n}\cos(nx)$, und der Grenzwert $\lim_{n\to\infty} f_n'(x)$ existiert für kein x.

(Warum? *Hinweis:* Das ist übrigens gar nicht so ganz einfach. Für $x = 0$ ist es klar. Für allgemeine x betrachtet man die Abbildung $\mathbb{Z} \to \mathbb{R}/(2\pi\mathbb{Z})$, $n \mapsto xn + 2\pi\mathbb{Z}$. Das ist ein Gruppenhomomorphismus. Der Kern ist entweder $k\mathbb{Z}$ für ein $k \in \mathbb{N}$, oder $\{0\}$. Der erste Fall ist wieder leicht, im zweiten Fall muss man sich überlegen, dass das Bild von $\mathbb{Z}$ dann dicht liegt.) ?

(5) *Alle Funktionen in einer Folge sind integrierbar, und die Grenzfunktion ist integrierbar, aber der Grenzwert der Integrale ist nicht das Integral der Grenzfunktion:*

Sei $E = [0,1]$, $f_n : E \to \mathbb{R}$, $f_n(x) = n^2(1-x)x^n$. Für $x = 1$ gilt $f_n(1) = 0$, also $\lim_{n\to\infty} f_n(1) = 0$. Für $0 \le x < 1$ gilt $\lim_{n\to\infty} n^a x^n = 0$ für alle $a \in \mathbb{R}$ (Das folgt aus Beispiel 5.5.3. – Wie?). Also insgesamt ?

$$f(x) := \lim_{n\to\infty} f_n(x) = 0 \quad \text{für alle } x \in [0,1].$$

Andererseits gilt

$$\int_0^1 f_n(x)dx = n^2 \int_0^1 (x^n - x^{n+1})dx = n^2\Big(\frac{1}{n+1} - \frac{1}{n+2}\Big) = \frac{n^2}{(n+1)(n+2)}$$

und somit

$$\lim_{n\to\infty} \int_0^1 f_n(x)dx = 1 \neq 0 = \int_0^1 f(x)dx \ .$$

Also kann es passieren, dass alle Funktionen in einer Funktionenfolge integrierbar sind, die Grenzfunktion ebenfalls integrierbar ist, aber die Folge der Integrale nicht gegen das Integral der Grenzfunktion strebt.

(Wofür ist das Folgende ein Beispiel? Wähle eine bijektive Funktion $u : \mathbb{N} \to \mathbb{Q} \cap [0,1]$. Sei $f_n : [0,1] \to \mathbb{R}$ wie folgt gegeben: $f_n(x) = 1$, falls es ein $m \le n$ gibt mit $u(m) = x$, und $f_n(x) = 0$ sonst. Betrachte $(f_n)_{n\in\mathbb{N}}$.) ?

8.2 Gleichmäßige Konvergenz

Definition 8.2.1. Sei E eine Menge und $(f_n)_{n\in\mathbb{N}}$ eine Funktionenfolge mit $f_n : E \to \mathbb{C}$. Die Folge $(f_n)_{n\in\mathbb{N}}$ konvergiert *gleichmäßig* gegen eine Funktion $f : E \to \mathbb{C}$, falls gilt:

$$\forall \varepsilon > 0 \ \ \exists N \in \mathbb{N} \ \ \forall n \ge N, x \in E \ : \ |f_n(x) - f(x)| < \varepsilon \ .$$

Eine Reihe $\sum f_k$ konvergiert *gleichmäßig*, falls die Folge $s_n = \sum_{k=1}^n f_k$ der Partialsummen gleichmäßig konvergiert.

Bemerkung 8.2.2.

(1) Für punktweise Konvergenz würde die Bedingung so aussehen:

$$\forall \varepsilon > 0, x \in E \ \ \exists N \in \mathbb{N} \ \ \forall n \ge N \ : \ |f_n(x) - f(x)| < \varepsilon \ .$$

Der Unterschied ist also, dass bei punktweiser Konvergenz N von x abhängen darf, und bei gleichmäßiger Konvergenz muss das gleiche N für alle $x \in E$ funktionieren. Insbesondere:

$$\begin{aligned} &(f_n)_{n\in\mathbb{N}} \text{ konv. gleichmäßig gegen } f \\ &\quad \Rightarrow \quad \forall x \in E \ : \ f(x) = \lim_{n\to\infty} f_n(x) \\ &\overset{\text{Def. 8.1.1}}{\Leftrightarrow} \quad (f_n)_{n\in\mathbb{N}} \text{ konv. punktweise gegen } f, \text{ d.h. } f = \lim_{n\to\infty} f_n \end{aligned}$$

(2) Aus punktweiser Konvergenz folgt nicht gleichmäßige Konvergenz. Dazu betrachten wir Beispiel 8.1.2 (2): $f_n : [0,1] \to \mathbb{R}$, $f_n(x) = x^n$ und $f(x) = 0$ für $0 \le x < 1$ und $f(1) = 1$. Es gilt

$$f = \lim_{n\to\infty} f_n \ .$$

(Wenn nichts anderes dabei steht, heißt diese Notation immer punktweise konvergenz.) Aber:

Beh.: $(f_n)_{n\in\mathbb{N}}$ konvergiert nicht gleichmäßig gegen f.

(Warum gilt das? *Hinweis:* Widerspruchsbeweis. Wählen Sie $\varepsilon = \frac{1}{3}$ und nehmen Sie an, es gäbe ein geeignetes N. Verwenden Sie den Zwischenwertsatz, um zu argumentieren, dass $f_N(x) = \frac{1}{2}$ für ein geeignetes x. Erklären Sie, warum das schlecht ist.) ?

(3) Die Supremumsnorm auf Funktionen $f : E \to \mathbb{C}$ ist definiert als

$$\|f - g\|_\infty \ = \ \sup\big\{|f(x) - g(x)| \,\big|\, x \in E\big\} \ \in \ \mathbb{R} \cup \{\infty\} \ .$$

Es gilt

$$\begin{aligned} & (f_n)_{n\in\mathbb{N}} \text{ konv. gleichmäßig gegen } f \\ \Leftrightarrow \quad & \forall\varepsilon > 0 \ \ \exists N \in \mathbb{N} \ \ \forall n \ge N \ : \ \|f_n - f\|_\infty < \varepsilon \\ \Leftrightarrow \quad & \lim_{n\to\infty} \|f_n - f\|_\infty \ = \ 0 \ . \end{aligned}$$

(Warum gelten diese Äquivalenzen?) Man sagt auch „$(f_n)_{n\in\mathbb{N}}$ konvergiert in der sup-Norm gegen f“. ?

(4) Sei $Y = \mathrm{Fun}_b(E, \mathbb{C})$ die Menge der beschränkten Funktionen $E \to \mathbb{C}$. Mit $d(f, g) = \|f - g\|_\infty$ wird Y ein metrischer Raum. Eine Folge beschränkter Funktionen $(f_n)_{n\in\mathbb{N}}$ ist dann einfach eine Folge in Y. Teil 3 oben kann man so umformulieren:

$$\begin{aligned} & f_n \to f \text{ in } Y \text{ für } n \to \infty \\ \Leftrightarrow \quad & (f_n)_{n\in\mathbb{N}} \text{ konv. gleichmäßig gegen } f \end{aligned}$$

(Warum?) ?

Satz 8.2.3. (Cauchy-Kriterium für gleichmäßige Konvergenz)
Sei $(f_n)_{n\in\mathbb{N}}$ eine Funktionenfolge auf E. Es sind äquivalent:

(i) $(f_n)_{n\in\mathbb{N}}$ konvergiert gleichmäßig auf E.

(ii) $\forall \varepsilon > 0 \;\; \exists N \in \mathbb{N} \;\; \forall m, n \geq N, x \in E \;:\; |f_m(x) - f_n(x)| < \varepsilon$.

Beweis.
(i)⇒(ii): Sei $f(x) = \lim_{n\to\infty} f_n(x)$ die Grenzfunktion. Sei $\varepsilon > 0$ beliebig. Dann gibt es N, so dass für alle $n \geq N$ und $x \in E$,

$$|f_n(x) - f(x)| < \tfrac{\varepsilon}{2} \ .$$

Dann gilt aber auch für alle $m, n \geq N$, dass

$$|f_m(x) - f_n(x)| \;\leq\; |f_m(x) - f(x)| + |f(x) - f_n(x)| \;<\; \tfrac{\varepsilon}{2} + \tfrac{\varepsilon}{2} \ .$$

(ii)⇒(i): Für jedes x bildet $(f_n(x))_{n\in\mathbb{N}}$ eine Cauchy-Folge in $\mathbb{C}$. Also existiert für jedes x der Grenzwert

$$f(x) := \lim_{n\to\infty} f_n(x) \ ,$$

d.h. $(f_n)_{n\in\mathbb{N}}$ konvergiert punktweise gegen f.

Beh.: $(f_n)_{n\in\mathbb{N}}$ konvergiert gleichmäßig gegen f.

⌈ Sei $\varepsilon > 0$ beliebig. Per Voraussetzung gibt es $N \in \mathbb{N}$, so dass für alle $m, n \geq N$ und $x \in E$ gilt

$$|f_m(x) - f_n(x)| < \varepsilon \ .$$

Da die Betragsfunktion $|\cdot| : \mathbb{C} \to \mathbb{R}$ stetig ist, folgt (für alle $n \geq N$ und $x \in E$)

$$|f(x) - f_n(x)| \;=\; \lim_{m\to\infty} |f_m(x) - f_n(x)| \;\leq\; \varepsilon \ .$$

(Warum? Und warum „≤“ und nicht „<“? Und warum beweist das trotzdem die Behauptung?) ?
⌋

□

Satz 8.2.4. (Konvergenzkriterium von Weierstraß)
Sei $(h_n)_{n\in\mathbb{N}}$ eine Funktionenfolge auf E. Angenommen, es gibt Konstanten $M_n \in \mathbb{R}$, $n \in \mathbb{N}$, so dass

(i) $|h_n(x)| \leq M_n$ für alle $x \in E$, $n \in \mathbb{N}$, und

(ii) die Reihe $\sum_{n=1}^{\infty} M_n$ konvergiert.

Dann konvergiert die Reihe von Funktionen $\sum h_n$ gleichmäßig auf E.

Beweis. Sei $f_n = \sum_{k=1}^{n} h_k$ die n-te Partialsumme. Wir zeigen, dass $(f_n)_{n\in\mathbb{N}}$ das Cauchy-Kriterium für gleichmäßige Konvergenz erfüllt.

Sei also $\varepsilon > 0$ beliebig. Nach dem Cauchy-Kriterium für Reihen (Satz 4.5.3 (i)) gibt es ein $N \in \mathbb{N}$, so dass für alle $m \geq n \geq N$ gilt:

$$\sum_{k=n}^{m} M_k \; < \; \varepsilon \; .$$

Daraus folgt, dass für alle $x \in E$ und $m \geq n \geq N$ gilt

$$\big|f_m(x) - f_{n-1}(x)\big| \; = \; \Big|\sum_{k=n}^{m} h_k(x)\Big| \; \leq \; \sum_{k=n}^{m} \big|h_k(x)\big| \; \leq \; \sum_{k=n}^{m} M_k \; < \; \varepsilon \; .$$

Nach Satz 8.2.3 konvergiert $(f_n)_{n\in\mathbb{N}}$ somit gleichmäßig. □

Beispiel 8.2.5. Sei $E = \mathbb{R}$ und $h_n : \mathbb{R} \to \mathbb{R}$, $n \in \mathbb{N}$ gegeben durch

$$h_n(x) = \begin{cases} \frac{(-1)^{(n-1)/2}}{n^2} \sin(nx) & ; n \text{ ungerade} \\ 0 & ; n \text{ gerade} \end{cases}$$

Wir können $M_n = n^{-2}$ wählen. Dann konvergiert $\sum M_n$, also konvergiert $\sum_{n=1}^{\infty} h_n$ gleichmäßig gegen eine Funktion $f : \mathbb{R} \to \mathbb{R}$. Über f werden wir später mehr sagen können.

Bislang war E nur eine Menge. Um über Grenzwerte $x \to p$ von Funktionen auf E reden zu können, muss E in einem metrischen Raum enthalten sein.

Satz 8.2.6. Sei X ein metrischer Raum und $E \subset X$. Sei $(f_n)_{n\in\mathbb{N}}$ eine Folge von Funktionen $f_n : E \to \mathbb{C}$, die gleichmäßig gegen $f : E \to \mathbb{C}$ konvergiert. Sei $x \in X$ ein Häufungspunkt von E und sei, für $n \in \mathbb{N}$,

$$F_n := \lim_{t\to x} f_n(t) \; .$$

(Insbesondere nehmen wir an, dass alle diese Grenzwerte existieren.) Dann konvergiert die Folge $(F_n)_{n\in\mathbb{N}}$ in $\mathbb{C}$ und es gilt

$$\lim_{n\to\infty} F_n \; = \; \lim_{t\to x} f(t) \; .$$

■ Die Folgerung im Satz oben kann man auch so formulieren: Es gilt

$$\lim_{n\to\infty}\left(\lim_{t\to x} f_n(t)\right) = \lim_{t\to x}\left(\lim_{n\to\infty} f_n(t)\right),$$

und alle diese Grenzwerte existieren. (Wobei die Existenz von $\lim_{t\to x} f_n(t)$ und $\lim_{n\to\infty} f_n(t)$ Voraussetzungen im Satz sind.)

Beweis.
Beh.: $(F_n)_n$ konvergiert.

⌈ Sei $\varepsilon > 0$ beliebig. Da $(f_n)_n$ gleichmäßig gegen f konvergiert, gibt es Satz 8.2.3 ein N, so dass für alle $m, n \geq N$ und $t \in E$ gilt $|f_m(t) - f_n(t)| < \varepsilon$. Für $t \to x$ folgt $|F_m - F_n| \leq \varepsilon$. Also ist $(F_n)_n$ eine Cauchy-Folge in $\mathbb{C}$. ⌋

Setze $F := \lim_{n\to\infty} F_n$. Wir müssen nun zeigen, dass $\lim_{t\to x} f(t) = F$. Dies ist (per Definition) äquivalent zu folgender Behauptung:

Beh.: $\forall\varepsilon > 0 \;\; \exists\delta > 0 \;\; \forall t \in E : 0 < d(t, x) < \delta \;\Rightarrow\; |f(t) - F| < \varepsilon.$

⌈ Sei $\varepsilon > 0$ beiliebig. Wegen der gleichmäßigen Konvergenz gibt es N_1, sodass für alle $n \geq N_1$ und $t \in E$:

$$|f(t) - f_n(t)| < \tfrac{\varepsilon}{3} .$$

Da F_n gegen F konvergiert, gibt es N_2, sodass für alle $n \geq N_2$:

$$|F - F_n| < \tfrac{\varepsilon}{3} .$$

Fixiere ein n, sodass $n \geq N_1$ und $n \geq N_2$. Da $f_n(t) \to F_n$ für $x \to t$ gibt es ein $\delta > 0$, sodass für alle $t \in E$:

$$0 < d(t, x) < \delta \quad\Rightarrow\quad |f_n(t) - F_n| < \tfrac{\varepsilon}{3} .$$

Insgesamt erhalten wir für $t \in E$ und $0 < d(t, x) < \delta$:

$$|f(t)-F| \leq |f(t)-f_n(t)|+|f_n(t)-F_n|+|F_n-F| < \tfrac{\varepsilon}{3}+\tfrac{\varepsilon}{3}+\tfrac{\varepsilon}{3} .$$ ⌋

□

Der nächste Satz gibt eine der wichtigsten Eigenschaften von gleichmäßig konvergenten Folgen: Sie erhalten die Stetigkeit. Dies folgt relativ direkt aus dem vorhergehenden Satz.

Satz 8.2.7. Sei X ein metrischer Raum und $E \subset X$. Sei $(f_n)_{n\in\mathbb{N}}$ eine Folge von stetigen Funktionen $f_n : E \to \mathbb{C}$, die gleichmäßig gegen $f : E \to \mathbb{C}$ konvergiert. Dann ist f stetig auf E.

Beweis. Wir benutzen die Charakterisierung von Stetigkeit aus Satz 5.2.4. Sei $x \in E$ ein Häufungspunkt von E. Es gilt

$$\begin{aligned}\lim_{t\to x} f(t) \;&\overset{\text{Def. } f}{=}\; \lim_{t\to x}\Big(\lim_{n\to\infty} f_n(t)\Big) \;\overset{\text{Satz 8.2.6}}{=}\; \lim_{n\to\infty}\Big(\lim_{t\to x} f_n(t)\Big)\\ &\overset{f_n \text{ stetig}}{=}\; \lim_{n\to\infty} f_n(x) \;\overset{\text{Def. } f}{=}\; f(x)\ .\end{aligned}$$

(Warum sind die Voraussetzungen von Satz 8.2.6 hier erfüllt?) □ ?

Wenn wir über integrierbare Funktionen $[a,b] \to \mathbb{C}$ reden, nehmen wir wieder an, dass $a < b$, ohne dass wir das jedes Mal dazusagen.

Satz 8.2.8. Sei $(f_n)_{n\in\mathbb{N}}$ eine Folge von Funktionen $f_n \in \mathcal{R}(a,b)$, die gleichmäßig gegen $f : [a,b] \to \mathbb{C}$ konvergiert. Dann ist $f \in \mathcal{R}(a,b)$ und es gilt

$$\int_a^b f(x)\,dx \;=\; \lim_{n\to\infty}\int_a^b f_n(x)\,dx\ .$$

■ Dieser Satz sagt, dass unter den angegebenen Voraussetzungen Integration und Grenzwert $n \to \infty$ vertauschen:

$$\int_a^b \Big(\lim_{n\to\infty} f_n(x)\Big)\,dx \;=\; \lim_{n\to\infty}\Big(\int_a^b f_n(x)\,dx\Big)\ .$$

Beweis. Wir zeigen den Satz für reellwertige Funktionen f_n. Die Aussage für komplexwertige f_n erhält man dann durch Zerlegen in Real- und Imaginärteil. ($\to$Übung)

Sei $\varepsilon > 0$ beliebig. Da $f_n \to f$ gleichmäßig, gibt es N, so dass für $n \geq N$ und alle $x \in [a,b]$ gilt $|f_n(x) - f(x)| < \varepsilon$. Oder, anders gesagt, dass

$$f_n(x) - \varepsilon \;<\; f(x) \;<\; f_n(x) + \varepsilon\ .$$

Fixiere nun ein $n \geq N$. Da $f_n \in \mathcal{R}(a,b)$, gibt es Treppenfunktionen $\underline{\varphi}, \overline{\varphi} \in \mathcal{T}(a,b)$, so dass $\underline{\varphi}(x) \leq f_n(x) \leq \overline{\varphi}(x)$, und

$$\int_a^b \overline{\varphi}(x)\,dx \;-\; \int_a^b f_n(x)\,dx \;<\; \varepsilon \quad , \quad \int_a^b f_n(x)\,dx \;-\; \int_a^b \underline{\varphi}(x)\,dx \;<\; \varepsilon\ .$$

(Warum gibt es solche Treppenfunktionen?) Dann sind auch $\underline{\psi}(x) = \underline{\varphi}(x) - \varepsilon$?

und $\overline{\psi}(x) = \overline{\varphi}(x) + \varepsilon$ Treppenfunktionen, und es gilt

$$\underline{\psi}(x) \;\leq\; f_n(x) - \varepsilon \;<\; f(x) \;<\; f_n(x) + \varepsilon \;\leq\; \overline{\psi}(x) \;.$$

Dass $f \in \mathcal{R}(a,b)$ folgt nun aus Lemma 7.3.6:

$$\begin{aligned}
&\int_a^b \big(\overline{\psi}(x) - \underline{\psi}(x)\big) dx \\
&= \int_a^b \big(\overline{\varphi}(x) + \varepsilon - f_n(x)\big) dx \;+\; \int_a^b \big(f_n(x) - \underline{\varphi}(x) + \varepsilon\big) dx \\
&< \; 2(b-a)\varepsilon + 2\varepsilon \;.
\end{aligned}$$

Da f integrierbar ist, können wir für den Wert von $\int f$ die folgende Abschätzung machen:

$$\Big| \int_a^b f(x)\,dx \;-\; \int_a^b f_n(x)\,dx \Big| \;\leq\; \int_a^b \underbrace{\big|f(x) - f_n(x)\big|}_{<\varepsilon}\,dx \;<\; (b-a)\varepsilon \;.$$

Daraus folgt $\lim_{n\to\infty} \int_a^b f_n(x)dx = \int_a^b f(x)dx$. (Warum?) □ ?

Korollar 8.2.9. Sei $(h_n)_{n\in\mathbb{N}}$ eine Folge von Funktionen $h_n \in \mathcal{R}(a,b)$, sodass die Reihe $\sum h_n$ gleichmäßig konvergiert. Dann ist $\sum h_n$ integrierbar, und es gilt

$$\int_a^b \Big(\sum_{n=1}^{\infty} h_n(x) \Big)\,dx \;=\; \sum_{n=1}^{\infty} \Big(\int_a^b h_n(x)\,dx \Big) \;.$$

(Wie folge das Korollar aus dem vorherigen Satz?) ?

Satz 8.2.10. Sei $(f_n)_{n\in\mathbb{N}}$ eine Funktionenfolge mit $f_n : [a,b] \to \mathbb{C}$. Es gelte:

- Für alle $n \in \mathbb{N}$ ist $f_n : [a,b] \to \mathbb{C}$ differenzierbar auf $[a,b]$,
- $(f_n(x_0))_{n\in\mathbb{N}}$ konvergiert für ein $x_0 \in [a,b]$,
- $(f'_n)_{n\in\mathbb{N}}$ konvergiert gleichmäßig auf $[a,b]$.

Dann folgt:

(i) $(f_n)_{n\in\mathbb{N}}$ konvergiert gleichmäßig gegen ein $f : [a,b] \to \mathbb{C}$, und

(ii) f ist differenzierbar, und für alle $x \in [a,b]$ gilt

$$f'(x) \;=\; \lim_{n\to\infty} f'_n(x) \;.$$

Beweis. Wir beweisen dies für reellwertige f_n. Für komplexwertige f_n erhält man die Aussage durch Zerlegen in Komponentenfunktionen. (Wie?) ?

(i) Wir werden den Mittelwertsatz benutzen, um das Cauchy-Kriterium für gleichmäßige Konvergenz zu zeigen. Sei also $\varepsilon > 0$ beliebig.

Beh.: Es gibt N, sodass für alle $m, n \geq N$ und $x \in [a, b]$ gilt $|f_m(x) - f_n(x)| < \varepsilon$.

⌈ Es gibt N, sodass für alle $m, n \geq N$ gilt

- $|f_m(x_0) - f_n(x_0)| < \frac{\varepsilon}{2}$ (da $(f_n(x_0))_{n\in\mathbb{N}}$ konvergiert),
- $|f'_m(x) - f'_n(x)| < \frac{\varepsilon}{2(b-a)}$ für alle $x \in [a, b]$ (da $(f'_n)_{n\in\mathbb{N}}$ gleichmäßig konvergiert).

Setze $g(x) = f_m(x) - f_n(x)$. Nach dem Mittelwertsatz (Satz 6.2.4) gibt es für jedes $y \in [a, b]$ ein ξ zwischen x und y, sodass $g(x) - g(y) = g'(\xi)\,(x - y)$. Mit der Abschätzung oben für $g' = f'_m - f'_n$ erhalten wir

$$\big|\, g(x) - g(y) \,\big| \;<\; \tfrac{\varepsilon}{2(b-a)}\, |x - y| \;. \tag{$*$}$$

Damit bekommen wir weiter

$$\begin{aligned} \big|\, f_m(x) - f_n(x) \,\big| \;&=\; \big|\, g(x) \,\big| \\ &\leq\; \big|\, g(x) - g(x_0) \,\big| + \big|\, g(x_0) \,\big| \\ &<\; \tfrac{\varepsilon}{2(b-a)}\, |x - x_0| + \tfrac{\varepsilon}{2} \;\leq\; \varepsilon \;. \end{aligned}$$ ⌋

Teil (i) folgt nun aus Satz 8.2.3.

(ii) Für $x \in [a, b]$ und $t \in [a, b] \setminus \{x\}$ setze

$$\phi_n(t) \;=\; \frac{f_n(t) - f_n(x)}{t - x} \quad , \quad \phi(t) \;=\; \frac{f(t) - f(x)}{t - x} \;.$$

Beh.: $(\phi_n)_{n\in\mathbb{N}}$ konvergiert gleichmäßig gegen ϕ.

⌈ Nach Teil (i) konvergiert ϕ_n zumindest punktweise gegen ϕ (Warum?). Es bleibt zu zeigen, dass die Konvergenz gleichmäßig ist. Dazu benutzen wir wieder das Cauchy-Kriterium (Satz 8.2.3). Seien N, m, n und g wie in Teil (i). Dann gilt für alle t in $[a, b] \setminus \{x\}$: ?

$$\begin{aligned} \big|\, \phi_m(t) - \phi_n(t) \,\big| \;&=\; \left|\, \frac{f_m(t) - f_m(x) - f_n(t) + f_n(x)}{t - x} \,\right| \\ &=\; \left|\, \frac{g(t) - g(x)}{t - x} \,\right| \;\overset{(*)}{<}\; \frac{\varepsilon}{2(b-a)} \;. \end{aligned}$$ ⌋

Jetzt können wir den Satz über das Vertauschen der Grenzwerte anwenden (Satz 8.2.6). Da für jedes n der Grenzwert $\lim_{t\to x}\phi(t)$ existiert, folgt

$$\lim_{n\to\infty}\underbrace{\left(\lim_{t\to x}\phi_n(t)\right)}_{=f_n'(x)} = \lim_{t\to x}\underbrace{\left(\lim_{n\to\infty}\phi_n(t)\right)}_{=\phi(t)} .$$

Insbesondere existiert der Grenzwert $f'(x) = \lim_{t\to x}\phi(t)$ (also ist f in x differenzierbar).

□

8.3 Potenzreihen

Eine wichtige Anwendung von gleichmäßiger Konvergenz ist, dass man in Potenzreihen die unendliche Summe mit Differenzieren und Integrieren vertauschen kann. Das behandeln wir in diesem Kapitel.

Definition 8.3.1. Für $\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$ sei $U \subset \mathbb{K}$ offen, $f : U \to \mathbb{K}$ und $a \in U$. Falls es ein $r > 0$ und $c_n \in \mathbb{K}$ $(n = 0, 1, \dots)$ gibt, so dass für alle $x \in U$ mit $|x - a| < r$ gilt

$$f(x) = \sum_{n=0}^{\infty} c_n\,(x-a)^n$$

(insbesondere muss die Reihe für alle $|x - a| < r$ konvergieren), so sagen wir

$$f \text{ ist bei } a \begin{cases} \text{reell analytisch} & , \text{ falls } \mathbb{K} = \mathbb{R} \\ \text{komplex analytisch} & , \text{ falls } \mathbb{K} = \mathbb{C} \end{cases} ,$$

oder auch kurz *f ist in a analytisch.* Wir sagen auch, *f kann um a in eine Potenzreihe entwickelt werden.*

Bemerkung 8.3.2.

(1) Wir werden häufig $a = 0$ setzen, um die Ausdrücke übersichtlicher zu machen. Das ist keine Einschränkung, da wir immer $y = x - a$ ersetzen können.

(2) Das Konvergenzverhalten von Potenzreihen hatten wir in Kapitel 4.7 betrachtet. Der Konvergenzradius R einer Potenzreihe $\sum_{n=0}^{\infty} c_n x^n$ (alles in $\mathbb{C}$) war

$$R^{-1} = \limsup_{n\to\infty} \sqrt[n]{|c_n|} \;\in\; \mathbb{R}_{\geq 0} \cup \{\infty\} ,$$

wobei wir $R = \infty$ setzen, falls $\limsup(\cdots) = 0$, und $R = 0$, falls $\limsup(\cdots) = \infty$. Die Reihe $\sum c_n x^n$ konvergiert für $|x| < R$ und divergiert für $|x| > R$. Für $|x| = R$ hängt das Konvergenzverhalten von x und der genauen Reihe ab.

(3) Hier interessieren wir uns für eine Mischform aus reell und komplex analytischen Funktionen, nämlich Funktionen von $U \subset \mathbb{R}$ nach $\mathbb{C}$ von der Form $f(x) = \sum_n c_n(x-a)^n$ mit $c_n \in \mathbb{C}$, aber $x, a \in \mathbb{R}$.

(4) In der Definition von reell/komplex analytisch wird nur die *Existenz* von $(c_n)_{n\in\mathbb{N}_0}$ verlangt, sodass gilt $f(x) = \sum_{n=0}^{\infty} c_n(x-a)^n$ für $|x-a|$ klein genug. Eine Frage an dieser Stelle ist, ob die c_n auch *eindeutig* durch f festgelegt sind. Das beantworten wir weiter unten mit „Ja."

Beispiel 8.3.3.

(1) Jedes Polynom ist analytisch in jedem Punkt. (Warum? Was sind die c_n falls $a = 1$ bei der Funktion $f(x) = x^3 + x$?) ?

(2) Die Exponentialfunktion $\exp(x) = \sum_{n=0}^{\infty} x^n/n!$ ist schon als Potenzreihe definiert und hat Konvergenzradius $R = \infty$ (Beispiel 4.7.4). Die Definition zeigt, dass f in 0 analytisch ist. Aber es gilt auch:

Beh.: exp ist in jedem Punkt analytisch.

⌈ Es gilt

$$\exp(x) = \exp(a)\exp(x-a) = \sum_{n=0}^{\infty} \tfrac{\exp(a)}{n!}(x-a)^n$$

wobei die Reihe auf der rechten Seite auch Konvergenzradius unendlich hat. ⌋

(3) Die Funktion $|x| : \mathbb{R} \to \mathbb{R}$ ist in jedem Punkt $a \in \mathbb{R} \setminus \{0\}$ analytisch (Warum?) und in $a = 0$ nicht analytisch. Letzteres wird eine einfache Konsequenz des Resultates unten, dass analytische Funktionen beliebig oft differenzierbar sind. ?

(4) Die Funktion $f : \mathbb{R} \to \mathbb{R}$, $f(x) = 0$ für $x \leq 0$ und $f(x) = \exp(-x^{-1})$ für $x > 0$ ist auf ganz $\mathbb{R}$ beliebig oft differenzierbar, aber in 0 nicht analytisch. ($\to$Übung)
(Auf $\mathbb{R} \setminus \{0\}$ ist f schon analytisch, das zeigen wir hier aber nicht.)

Satz 8.3.4. Die Reihe $\sum_{n=0}^{\infty} c_n x^n$ (mit $c_n \in \mathbb{C}$) habe Konvergenzradius $R > 0$. Dann gilt:

(i) $\sum_{n=0}^{\infty} c_n x^n$ konvergiert gleichmäßig auf $[-r, r]$ für jedes $0 < r < R$.

(ii) Die Funktion $f(x) = \sum_{n=0}^{\infty} c_n x^n$ ist stetig und differenzierbar auf $(-R, R)$ mit $f'(x) = \sum_{n=1}^{\infty} n c_n x^{n-1}$.

(Fällt Ihnen ein Gegenbeispiel ein für: „Die Reihe in (i) konvergiert gleichmäßig auf $(-R, R)$“?) ?

Beweis.

(i) Wir verwenden das Weierstraß-Kriterium (Satz 8.2.4). Setze $M_n = |c_n| r^n$ und $h_n(x) = c_n x^n$. Dann gilt für alle $x \in [-r, r]$, dass $|h_n(x)| \leq M_n$. Da $r < R$ folgt ferner, dass $\sum_{n=0}^{\infty} M_n = \sum_{n=0}^{\infty} |c_n| r^n$ konvergiert (da Potenzreihen für $|x| < R$ absolut konvergieren, siehe Beispiel 4.9.3). Nach Satz 8.2.4 konvergiert $\sum h_n$ gleichmäßig auf $[-r, r]$.

(ii) Wir wollen Satz 8.2.10 anwenden. Für $r < R$ setze $f_n : [-r, r] \to \mathbb{C}$, $f_n(x) = \sum_{k=0}^{n} c_k x^k$. Es gilt:

- $(f_n)_{n \in \mathbb{N}}$ ist differenzierbar auf $[-r, r]$,
- $(f_n(0))_{n \in \mathbb{N}}$ konvergiert gegen c_0 (da $f_n(0) = c_0$ für alle n),
- Beh.: $(f_n')_{n \in \mathbb{N}}$ konvergiert gleichmäßig auf $[-r, r]$.

  ⌈ Der Konvergenzradius von $\sum_{n=1}^{\infty} n c_n x^{n-1}$ ist ebenfalls R. (Warum? Hinweis: $\lim_{n\to\infty} \sqrt[n]{n} = 1$.) Also folgt die Behauptung aus Teil (i). ⌋ ?

Nach Satz 8.2.10 ist dann f für alle $x \in [-r, r]$ differenzierbar mit

$$f'(x) \;=\; \lim_{n\to\infty} f_n'(x) \;=\; \lim_{n\to\infty} \sum_{k=1}^{n} k c_k x^{k-1} \;=\; \sum_{k=1}^{\infty} k c_k x^{k-1} \; .$$

Da dies für alle $0 < r < R$ auf $[-r, r]$ gilt, gilt es auch auf $(-R, R)$. Somit ist f auf $(-R, R)$ differenzierbar, und insbesondere stetig.

□

Korollar 8.3.5. Die Reihe $f(x) = \sum_{n=0}^{\infty} c_n x^n$ habe Konvergenzradius $R > 0$. Dann gilt:

(i) f ist auf $(-R, R)$ beliebig oft differenzierbar, und die k'te Ableitung ist gegeben durch

$$f^{(k)}(x) \;=\; \sum_{n=k}^{\infty} n(n-1)\cdots(n-k+1)\, c_n\, x^{n-k} \; .$$

(ii) Die Koeffizienten der Reihe $\sum_n c_n x^n$ sind durch die Ableitungen von f wie folgt bestimmt:

$$c_k = \frac{1}{k!} f^{(k)}(0) .$$

■ Teil (ii) des Korollars zeigt die Eindeutigkeits in Bemerkung 8.3.2 (4).

Beweis.

(i) Im Beweis von Satz 8.3.4 (ii) haben wir gesehen, dass eine Potenzreihe und ihre Ableitung den gleichen Konvergenzradius haben. Teil (i) des Korollars folgt nun durch wiederholtes Anwenden von Satz 8.3.4 (ii).

(ii) Setze $x = 0$ in Teil (i).

□

Korollar 8.3.6. Die Reihe $f(x) = \sum_{n=0}^{\infty} c_n x^n$ habe Konvergenzradius $R > 0$. Dann gilt für $[a, b] \subset (-R, R)$, dass

$$\int_a^b f(x)\,dx = \sum_{n=0}^{\infty} \frac{c_n}{n+1} x^{n+1}\Big|_a^b .$$

Beweis. Nach Satz 8.3.4 konvergiert $\sum_{n=0}^{\infty} c_n x^n$ gleichmäßig auf $[a, b]$. Da $c_n x^n$ stetig ist, gilt $c_n x^n \in \mathcal{R}(a, b)$. Damit ist Satz 8.2.8 anwendbar und ergibt die Behauptung. □

Beispiel 8.3.7. In Beispiel 6.4.5 (2) hatten wir die Taylorentwicklung und deren Fehlerterm benutzt, um zu zeigen, dass für $-\frac{1}{2} < x < 1$ gilt

$$\log(1 + x) = \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} x^n .$$

Hier bekommen wir nun einen anderen Beweis, die obige Reihenentwicklung für alle $|x| < 1$ zeigt.

Es gilt $(\log(1 + x))' = (1 + x)^{-1}$ und, für $|x| < 1$,

$$\frac{1}{1+x} = \sum_{n=0}^{\infty} (-x)^n .$$

(Diese Reihe ist die geometrische Reihe und hat daher Konvergenzradius 1.) Für $[a, b] \subset (-1, 1)$ gilt nach Korollar 8.3.6, dass

$$\begin{aligned}\log(1 + b) - \log(1 + a) &= \int_a^b \frac{1}{1+x} dx = \sum_{n=0}^{\infty} \int_a^b (-x)^n dx \\ &= \sum_{n=0}^{\infty} (-1)^n \frac{1}{n+1} (b^{n+1} - a^{n+1}) .\end{aligned}$$

Für $|x| < 1$ und $a = 0, b = x$ erhält man die Reihenentwicklung von $\log(1+x)$.

Satz 8.3.8. Konvergiert die Reihe $\sum_{n=0}^{\infty} c_n$, dann konvergiert die Reihe $\sum_{n=0}^{\infty} c_n x^n$ gleichmäßig auf $[0,1]$.

Beweis. Setze $s_{-1} = 0$, $s_n = \sum_{k=0}^{n} c_k$, $(n \geq 0)$ und $s = \lim_{n\to\infty} s_n$ (so dass per Definition $s = \sum_{k=0}^{\infty} c_k$). Dann

$$\begin{aligned}
\sum_{k=0}^{m} c_k x^k &= \sum_{k=0}^{m} \big((s_k - s) - (s_{k-1} - s)\big) x^k \\
&= \sum_{k=0}^{m} (s_k - s) x^k - \sum_{k=-1}^{m-1} (s_k - s) x^{k+1} \\
&= (1-x) \sum_{k=0}^{m-1} (s_k - s) x^k + (s_m - s) x^m + s
\end{aligned}$$

Sei $\varepsilon > 0$ beliebig. Wähle $N \in \mathbb{N}$ so dass für alle $n \geq N$ gilt

$$|s_n - s| < \varepsilon \ .$$

Dann gilt für alle $m > n \geq N$

$$\begin{aligned}
\Big| \sum_{k=n+1}^{m} c_k x^k \Big| &= \Big| \sum_{k=0}^{m} c_k x^k - \sum_{k=0}^{n} c_k x^k \Big| \\
&= \Big| (1-x) \sum_{k=n}^{m-1} (s_k - s) x^k + (s_m - s) x^m - (s_n - s) x^n \Big| \\
&\leq (1-x) \sum_{k=n}^{m-1} |s_k - s| x^k + |s_m - s| x^m + |s_n - s| x^n \\
&\leq \varepsilon (1-x) \sum_{k=0}^{\infty} x^k + \varepsilon + \varepsilon = 3\varepsilon \ .
\end{aligned}$$

Somit erfüllt die Reihe $\sum_{n=0}^{\infty} c_k x^k$ auf $[0,1]$ das Cauchy-Kriterium für gleichmäßige Konvergenz (Satz 8.2.3). $\square$

Mit Satz 8.2.7 erhalten wir:

Korollar 8.3.9. Konvergiert die Reihe $\sum_{n=0}^{\infty} c_n r^n$ für ein $0 < r < \infty$, und ist $f(x) = \sum_{n=0}^{\infty} c_n x^n$ für $|x| < r$, so gilt

$$\lim_{x \to r} f(x) = \sum_{n=0}^{\infty} c_n r^n \ .$$

(Wieso funktioniert die Verallgemeinerung von $x \in [0,1]$ zu $[0,r]$ in ?
Satz 8.3.8 wie wir es hier brauchen?)

Beispiel 8.3.10.

(1) In Korollar 8.3.9 folgt die Existenz des Grenzwertes $\lim_{x\to r} f(x)$ aus der Konvergenz der Reihe $\sum c_n r^n$. Man kann fragen, ob man umgekehrt aus der Existenz des Grenzwertes die Konvergenz der Reihe folgern kann. Das ist aber nicht so:

Betrachte $\sum_{n=0}^{\infty}(-x)^n$. Dies konvergiert für $|x| < 1$ gegen $f(x) = \frac{1}{1+x}$. Nun gilt

$$\lim_{x\to 1} f(x) = \frac{1}{2} \ ,$$

insbesondere existiert der Grenzwert. Aber die Reihe $\sum_{n=0}^{\infty}(-1)^n$ konvergiert nicht.

(2) Aus Analysis 1 wissen wir, dass die Reihe $\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n}$ konvergiert. Aus Beispiel 8.3.7 wissen wir, dass für $|x| < 1$ gilt

$$\log(1+x) = \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n}\, x^n \ .$$

Mit dem Korollar erhalten wir

$$\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} \ = \ \lim_{x\to 1} \log(1+x) \ = \ \log(2) \ .$$

Dies beweist eine der Behauptungen am Anfang von Kapitel 4.

8.4 Der Satz von Stone-Weierstraß

In diesem Kapitel sei

- $\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$, und
- (X, d_X) sei ein metrischer Raum.

Wir definieren

$$\mathcal{C}_{\mathbb{K}}(X) = \{f : X \to \mathbb{K} \mid f \text{ stetig und beschränkt}\} \ .$$

Wir schreiben auch $\mathcal{C}(X)$ statt $\mathcal{C}_{\mathbb{K}}(X)$.

■ Wie in Bemerkung 8.2.2: für $f \in \mathcal{C}(X)$ ist die *Supremumsnorm* gegeben durch

$$\|f\|_\infty \;=\; \sup\left\{|f(x)| \;\middle|\; x \in X\right\} .$$

Wir schreiben in diesem Kapitel kurz $\|f\|$ statt $\|f\|_\infty$. Da f beschränkt ist, gilt $\|f\| < \infty$. Durch

$$d(f,g) \;=\; \|f-g\| \qquad \text{für} \;\; f,g \in \mathcal{C}(X)$$

wird eine Metrik auf $\mathcal{C}(X)$ definiert. (Warum?) ?

Wir haben es jetzt also mit zwei metrische Räumen zu tun. Nämlich zum einen mit (X, d_X), dem Definitionsbereich der Funktionen, und zum anderen mit $(\mathcal{C}(X), d)$, dem Raum der Funktionen selber.

Erinnerung:

- Ein metrischer Raum ist *vollständig*, wenn jede Cauchy-Folge konvergiert, siehe Definition 4.2.5.
- Sei $E \subset X$. Die *abgeschlossenen Hülle von E in X*, oder kurz der *Abschluss von E*, ist $\overline{E} = E \cup \{\text{Häufungspunkte von } E \text{ in } X\}$, siehe Definition 3.2.13.

Bemerkung 8.4.1.

(1) In Bemerkung 8.2.2 (4) hatten wir gesehen, dass Konvergenz im metrischen Raum $(\mathcal{C}_{\mathbb{K}}(X), d)$ das gleiche ist, wie gleichmäßige Konvergenz von Funktionenfolgen.

(2) Sei $E \subset \mathcal{C}(X)$. Die abgeschlossene Hülle von E in $\mathcal{C}(X)$ heisst auch *gleichmäßig abgeschlossene Hülle*.

Nach Teil (1) und Satz 4.1.8 liegt $f \in \mathcal{C}(X)$ genau dann in der gleichmäßig abgeschlossenen Hülle von E, wenn es eine Folge $(f_n)_{n\in\mathbb{N}}$ von Funktionen in E gibt, die gleichmäßig gegen f konvergiert.

Satz 8.4.2. $(\mathcal{C}(X), d)$ ist vollständig.

Beweis. Sei $(f_n)_{n\in\mathbb{N}}$ eine Cauchy-Folge in $\mathcal{C}(X)$. Dies ist äquivalent zu:

$$\forall \varepsilon > 0 \;\; \exists N \;\; \forall m,n \geq N, x \in X \;:\; |f_m(x) - f_n(x)| < \varepsilon \; .$$

(Warum?). ?

Nach dem Cauchy-Kriterium für gleichmäßige Konvergenz (Satz 8.2.3) konvergiert f_n auf X gleichmäßig gegen eine Grenzfunktion f. Da alle f_n stetig sind, ist nach Satz 8.2.7 auch f stetig.

Beh.: f ist beschränkt.

⌈ Da $f_n \to f$ in der sup-Norm, gibt es ein N, so dass für alle $n \geq N$ gilt $\|f_n - f\| < 1$. Da f_n beschränkt ist, muss auch f beschränkt sein. ⌋

(Bei punktweiser Konvergenz $f_n \to f$ von beschränkten stetigen Funktionen f_n stimmt die Behauptung im Allgemeinen nicht. Fällt Ihnen ein Gegenbeispiel ein?) ? □

Beispiel 8.4.3. Die Aussage von Satz 8.4.2 ist bemerkenswert, da wir ja noch gar nicht so viele vollständige Mengen kennen. Im wesentlichen kennen wir bislang $\mathbb{R}^n$.

Hier geben wir noch ein paar Beispiele, die illustrieren, dass viele Funktionenräume nicht vollständig sind (immer bezüglich der sup-Norm). Wir betrachten hier jeweils den metrischen Raum $X = [a, b] \subset \mathbb{R}$ für $a < b$.

(1) Die Treppenfunktionen $\mathcal{T}(a, b)$ sind nicht vollständig.

(2) Die Menge der polynomialen Funktionen auf $[a, b]$ ist nicht vollständig.

(3) Die Menge der auf $[a, b]$ differenzierbaren Funktionen ist nicht vollständig.

Andererseits:

(4) Konstante Funktionen sind vollständig.

(5) Die integrierbaren Funktionen $\mathcal{R}(a, b)$ ist vollständig.

(Warum gilt das alles?) ?

Bemerkung 8.4.4. Sei M eine Menge. Eine Untermenge A der Menge der Funktionen $M \to \mathbb{K}$ ist eine (Unter-)Algebra (über $\mathbb{K}$ und mit Eins) falls

(1) für $f, g \in A$ auch $f + g \in A$ und $fg \in A$ gilt,

(2) für $f \in A$ und $c \in \mathbb{K}$ auch $cf \in A$ gilt, und

(3) die Funktion $1 : M \to \mathbb{C}$, $m \mapsto 1$ in A ist.

Dies ergibt sich sofort aus dem Vergleich mit der Definition aus der Linearen Algebra.

■ In [Def. 7.28 aus Rudin] wird nicht $1 \in A$ vorausgesetzt, während wir hier nur Algebren mit Eins betrachten.

Beispiel 8.4.5.

(1) Die Menge der Polynome

$$A_1 = \left\{ f : \mathbb{R} \to \mathbb{R} \,\middle|\, f(x) = \sum_{k=0}^{N} c_k x^k \ (N \in \mathbb{N}\,,\ c_k \in \mathbb{R}) \right\}$$

ist eine Algebra.

(2) Die Menge der ungeraden Polynome

$$A_2 = \left\{ f : \mathbb{R} \to \mathbb{R} \,\middle|\, f(x) = \sum_{k=0}^{N} c_k x^{2k+1} \ (N \in \mathbb{N}\,,\ c_k \in \mathbb{R}) \right\}$$

ist keine Algebra (da $1 \notin A_2$ und $x \in A_2$, aber $x \cdot x \notin A_2$).

(3) Die Menge der geraden Polynome

$$A_3 = \left\{ f : \mathbb{R} \to \mathbb{R} \,\middle|\, f(x) = \sum_{k=0}^{N} c_k x^{2k} \ (N \in \mathbb{N}\,,\ c_k \in \mathbb{R}) \right\}$$

ist eine Algebra.

(4) Die Menge

$$A_4 = \left\{ f : \mathbb{R} \to \mathbb{C} \,\middle|\, f(x) = \sum_{k=0}^{N} c_k e^{ikx} \ (N \in \mathbb{N}\,,\ c_k \in \mathbb{C}) \right\}$$

ist eine Algebra.

Satz 8.4.6. Sei $A \subset \mathcal{C}(X)$ eine Algebra. Dann ist auch der gleichmäßige Abschluss $\overline{A} \subset \mathcal{C}(X)$ eine Algebra.

Der Beweis ist eine $\to$Übung.

Definition 8.4.7. Sei M eine Menge und E eine Teilmenge der Funktionen $M \to \mathbb{K}$. Wir sagen E *separiert Punkte auf* M, falls es für alle $x, y \in M$, $x \neq y$, ein $f \in E$ gibt, sodass $f(x) \neq f(y)$.

Beispiel 8.4.8. Bezeichnungen wie in Beispiel 8.4.5.

(1) A_1 separiert Punkte auf $\mathbb{R}$: Für $u, v \in \mathbb{R}$ mit $u \neq v$ wähle $f(x) = x$. Dann $f \in A_1$, $f(u) = u$ und $f(v) = v$.

(2) A_2 separiert Punkte auf $\mathbb{R}$: dito.

(3) A_3 separiert nicht Punkte auf $\mathbb{R}$: Z.B. gilt $f(1) = f(-1)$ für alle $f \in A_3$. (Allerdings separiert A_3 Punkte auf $\mathbb{R}_{\geq 0}$.)

(4) A_4 separiert nicht Punkte auf $\mathbb{R}$: Z.B. gilt $f(0) = f(2\pi)$ für alle $f \in A_4$. (Allerdings separiert A_4 Punkte auf $[0, 2\pi)$.)

(5) Wenn E eine injektive Funktion $f : M \to \mathbb{K}$ enthält (wie in den ersten beiden Beispielen), dann ist klar, dass E Punkte separiert. Aber das ist keine notwendige Bedingung. Z.B. sei $M = [a, b]$ und $E = \mathcal{T}(a, b)$, die Treppenfunktionen auf $[a, b]$. Dann separiert E Punkte, aber es gibt keine einzelne Funktion $\varphi \in E$ (oder auch nur eine endliche Teilmenge von Funktionen in E), die schon alle Punkte $u, v \in [a, b]$ separiert. (Warum?) ?

Satz 8.4.9. (Satz von Stone-Weierstraß)
Sei K ein kompakter metrischer Raum, und sei $A \subset \mathcal{C}_{\mathbb{R}}(K)$ eine Algebra (über $\mathbb{R}$ und mit Eins), die auf K Punkte separiert. Dann $\overline{A} = \mathcal{C}_{\mathbb{R}}(K)$.

■ In [Satz 7.31 aus Rudin] ist eine etwas andere Aussage formuliert. Dort wird noch „verschwindet an keinem Punkt von K“ gefordert, dafür muss die Algebra nicht unbedingt eine Algebra mit 1 sein. Der Beweis hier folgt [Kap. VII.3 aus Dieudonné].

Wir werden folgendes Lemma über kompakte Mengen benötigen:

Lemma 8.4.10. Seien K_n, $n \in \mathbb{N}$ nicht-leere kompakte Mengen mit $K_{n+1} \subset K_n$. Dann ist $\bigcap_{n\in\mathbb{N}} K_n$ nicht leer.

Beweis. Angenommen, $\bigcap_{n\in\mathbb{N}} K_n = \emptyset$. Dann ist $\{U_n\}_{n\in\mathbb{N}}$ mit $U_n = K_1 \setminus K_n$ eine offene (in K_1) Überdeckung von K_1 (Warum?). Sei $U_{n_1} \cup \cdots \cup U_{n_k} = K_1$? eine endliche Teilüberdeckung. Da $U_{n+1} \supset U_n$ wird dann K_1 schon durch ein einzelnes U_{n_i} überdeckt. Dazu muss aber K_{n_i} leer sein, Widerspruch. □

Beweis von Satz 8.4.9. Wir werden den Satz in sechs Schritten beweisen.

Schritt 1:
Beh.: Es gibt eine Folge von Polynomen $(p_n)_{n\in\mathbb{N}}$, die punktweise auf $[0, 1]$ gegen $\sqrt{x}$ konvergiert.

⌈ Wir konstruieren rekursiv eine Folge von Polynomen $(p_n)_{n\in\mathbb{N}}$. Es sei

$$p_1(x) = 0 \qquad \text{und} \qquad p_{n+1}(x) = p_n(x) + \tfrac{1}{2}(x - p_n(x)^2) \ .$$

1. Es gilt $0 \leq p_n(x) \leq \sqrt{x}$ für $x \in [0,1]$: Dies folgt per Induktion. Für $n = 1$ ist die Aussage wahr. Angenommen, die Aussage sei für alle $m \leq n$ gezeigt. Dann gilt

$$p_n(x) \leq \sqrt{x} \;\Rightarrow\; \tfrac{1}{2}(\sqrt{x} + p_n(x)) \leq \sqrt{x} \leq 1$$

   und

$$\begin{aligned} \sqrt{x} - p_{n+1}(x) &= \sqrt{x} - p_n(x) - \tfrac{1}{2}(x - p_n(x)^2) \\ &= (\sqrt{x} - p_n(x))(1 - \tfrac{1}{2}(\sqrt{x} + p_n(x))) \geq 0 \;, \end{aligned}$$

   da beide Terme des Produktes getrennt ≥ 0 sind. Ferner gilt

$$p_{n+1}(x) = p_n(x) + \tfrac{1}{2}(x - p_n(x)^2) \geq 0 \;,$$

   da per Induktionsannahme beide Summanden ≥ 0 sind. Dies zeigt Aussage 1.

2. $p_{n+1}(x) \geq p_n(x)$ für $x \in [0,1]$: Folgt aus der Rekursionsvorschrift für $p_{n+1}(x)$ und Teil 1.

3. Die Folge $(p_n)_{n\in\mathbb{N}}$ konvergiert punktweise auf $[0,1]$ gegen $\sqrt{x}$: Sei $x \in [0,1]$ fix. Die Folge $(p_n(x))_{n\in\mathbb{N}}$ ist monoton steigend und beschränkt (nach Teil 1 und 2), und ist somit nach Satz 4.3.2 konvergent. Setze

$$p(x) = \lim_{n\to\infty} p_n(x) \;.$$

   Aus der Rekursionsformel für $p_{n+1}(x)$ erhält man im Grenzübergang $n \to \infty$ somit $p(x) = p(x) + \frac{1}{2}(x - p(x)^2)$, also $p(x)^2 = x$. Da $p(x) \geq 0$ folgt $p(x) = \sqrt{x}$. $\rfloor$

Beh.: Die Folge $(p_n)_{n\in\mathbb{N}}$ konvergiert sogar gleichmäßig auf $[0,1]$ gegen $\sqrt{x}$.

$\lceil$ Betrachte $g_n(x) = \sqrt{x} - p_n(x)$. Dann ist g_n stetig, $g_n(x) \geq 0$ und die Folge $(g_n(x))_{n\in\mathbb{N}}$ ist monoton fallend für jedes $x \in [0,1]$. Sei $\varepsilon > 0$ beliebig. Sei $I_n = \{x \in [0,1] | g_n(x) \geq \varepsilon\}$. Dann ist I_n abgeschlossen, also kompakt, und es gilt $I_{n+1} \subset I_n$ (Warum?). ?
Da $g_n(x)$ punktweise gegen 0 konvergiert, folgt $\bigcap_n I_n = \emptyset$. Nach Lemma 8.4.10 ist dies nur möglich, falls es N gibt mit $I_n = \emptyset$ für $n \geq N$ (Warum?). $\rfloor$?

Schritt 2:
Beh.: Sei $f \in \overline{A}$. Dann auch $|f| \in \overline{A}$.

⌈ Sei $M = \|f\|$. Setze $f_n(x) = p_n(f(x)^2/M^2)$. Da $f(x)^2/M^2 \in [0,1]$, konvergiert nach Schritt 1 die Folge $(f_n)_{n\in\mathbb{N}}$ gleichmäßig auf $[0,1]$ gegen $\sqrt{f(x)^2/M^2} = |f(x)|/M$ (Warum?). Nach Satz 8.4.6 ist $\overline{A}$ eine Algebra, und da f_n sich aus (endlichen) Summen und Produkten von Elementen aus $\overline{A}$ ergibt, gilt $f_n \in \overline{A}$. Also liegt $|f|/M$ im gleichmäßigen Abschluss von $\overline{A}$, der aber gleich $\overline{A}$ ist. ⌋ ?

Schritt 3: Setze

$$\max(f,g)(x) = \begin{cases} f(x) & ; f(x) \geq g(x) \\ g(x) & ; f(x) < g(x) \end{cases} , \quad \min(f,g)(x) = \begin{cases} f(x) & ; f(x) \leq g(x) \\ g(x) & ; f(x) > g(x) \end{cases}$$

Beh.: Seien $f, g \in \overline{A}$. Dann auch $\max(f,g) \in \overline{A}$ und $\min(f,g) \in \overline{A}$.

⌈ Es gilt $\max(f,g) = \frac{1}{2}(f + g + |f - g|)$ und $\min(f,g) = \frac{1}{2}(f + g - |f - g|)$. Damit folgt das Ergebnis aus Satz 8.4.6 und Schritt 2. ⌋

Wie setzen $\max(g_1, \ldots, g_n) = \max(g_1, \max(g_2, \ldots, \max(g_{n-1}, g_n) \ldots))$. Die obige Behauptung zeigt auch, dass $\max(g_1, \ldots, g_n) \in \overline{A}$, falls $g_1, \ldots, g_n \in \overline{A}$. Analoges gilt für $\min(g_1, \ldots, g_n)$.

Schritt 4:
Beh.: Für alle $a, b \in K$, $a \neq b$ und $\alpha, \beta \in \mathbb{R}$ gibt es $f \in A$, so dass $f(a) = \alpha$ und $f(b) = \beta$.

⌈ Nach Annahme separiert A Punkte auf K, so dass es ein $g \in A$ gibt mit $g(a) \neq g(b)$. Wähle

$$f(x) = \alpha + (\beta - \alpha)\frac{g(x) - g(a)}{g(b) - g(a)} .$$

Dann $f \in A$ (Warum?) und f hat die Eigenschaften aus der Behauptung. ⌋ ?

Schritt 5:
Beh.: Sei $f \in \mathcal{C}_{\mathbb{R}}(K)$, $a \in K$, $\varepsilon > 0$. Es gibt $g \in \overline{A}$, so dass $g(a) = f(a)$ und $g(x) \leq f(x) + \varepsilon$ für alle $x \in K$.

⌈ Nach Schritt 4 gibt es für jedes $b \in K$ ein $h_b \in A$, so dass $h_b(a) = f(a)$ und $h_b(b) = f(b)$. Da h_b und f stetig sind, gibt es eine offene Umgebung U_b mit $h_b(x) \leq f(x) + \varepsilon$ für alle $x \in U_b$ (Warum?). Somit ist $\{U_b\}_{b\in K}$ eine offene Überdeckung von K. Da K kompakt ist, gibt es eine endliche Teilüberdeckung ?

$$K \subset U_{b_1} \cup \cdots \cup U_{b_n} .$$

Setze $g = \min(h_{b_1}, \ldots, h_{b_n})$. Dann $g \in \overline{A}$ und g erfüllt die geforderte Eigenschaft (Warum?). ⌋ ?

Schritt 6:
Beh.: $\overline{A} = \mathcal{C}_{\mathbb{R}}(K)$.

⌈ Sei $f \in \mathcal{C}_{\mathbb{R}}(K)$ und $\varepsilon > 0$ vorgegeben. Für jedes $a \in K$ gibt es nach Schritt 5 ein $g_a \in \overline{A}$ mit $g_a(a) = f(a)$ und $g_a(x) \leq f(x) + \varepsilon$ für alle $x \in K$. Sei V_a eine offene Umgebung von a mit $g_a(x) \geq f(x) - \varepsilon$ für alle $x \in V_a$. Dann ist $\{V_a\}_{a \in K}$ eine offene Überdeckung von K. Da K kompakt ist, gibt es eine endliche Teilüberdeckung

$$K \subset V_{a_1} \cup \cdots \cup V_{a_n} \ .$$

Setze $h = \max(g_{a_1}, \ldots, g_{a_n})$. Dann ist $h \in \overline{A}$. Sei $x \in K$ beliebig. Es gilt $x \in V_{a_i}$ für ein $i = 1, \ldots, n$. Per Konstruktion gilt $h(x) \geq g_{a_i}(x) \geq f(x) - \varepsilon$. Ferner gilt $h(x) \leq f(x) + \varepsilon$, da dies für alle g_a gilt. Also $\|f - h\| \leq \varepsilon$. Somit folgt, dass $f \in \overline{A}$, oder dass f ein Häufungspunkt von $\overline{A}$ ist. In beiden Fällen ist $f \in \overline{A}$. ⌋

□

Beispiel 8.4.11.

(1) Beh.: Für jede stetige Funktion $f : [a, b] \to \mathbb{R}$ und jedes $\varepsilon > 0$ gibt es ein Polynom $P(x) = \sum_{n=0}^{M} c_n x^n$ mit $|f(x) - P(x)| < \varepsilon$ für alle $x \in [a, b]$.

⌈ Wir haben bereits gesehen, dass die Polynome eine Algebra A bilden, die Punkte separiert. Nach Satz 8.4.9 ist $\overline{A} = \mathcal{C}_{\mathbb{R}}([a, b])$. Also gibt es eine Folge $(p_n)_{n \in \mathbb{N}}$ mit $p_n \to f$ gleichmäßig und $p_n \in A$. Daraus folgt die Behauptung. (Wie?) ⌋ ?

(2) Mit dem vorherigen Beispiel kann man auch illustrieren, dass Satz 8.4.9 nicht mehr gilt, wenn man die Bedingung „kompakt" weglässt.

Betrachte dazu $X = [-1, 1] \setminus \{0\}$ und die Algebra A der Polynome auf $[-1, 1]$, aber jetzt eingeschränkt auf X. Dann separiert A immer noch Punkte (da es die Funktion $p(x) = x$ enthält). Aber jetzt ist die Stufenfunktion $f(x) = 0$ für $x < 0$ und $f(x) = 1$ für $x > 0$ stetig auf X, also $f \in \mathcal{C}_{\mathbb{R}}(X)$. Und jede Funktion in $\overline{A}$ lässt sich stetig auf $[-1, 1]$ fortsetzen (Warum?). Die Funktion f lässt sich aber nicht stetig auf $[-1, 1]$ fortsetzen. ?

(Fällt Ihnen auch jeweils ein Gegenbeispiel ein, wenn man statt „kompakt" in Satz 8.4.9 die Bedingungen „Algebra" oder „separiert Punkte" weglässt?) ?

(3) Beh.: Sei $f : \mathbb{R} \to \mathbb{R}$ eine stetige Funktion mit $f(x) = f(x + 2\pi)$ für alle $x \in \mathbb{R}$, und sei $\varepsilon > 0$. Dann gibt es $Q(x) = \sum_{m=0}^{M}(c_m \sin(mx) + d_m \cos(mx))$ mit $|Q(x) - f(x)| < \varepsilon$ für alle $x \in \mathbb{R}$.

Hier kann man Satz 8.4.9 erstmal nicht anwenden, da $\mathbb{R}$ nicht kompakt ist. Andererseits reicht es ja, die Funktion f auf $[0, 2\pi]$ zu kennen, und das ist kompakt. Allerdings gilt für alle Funktionen Q, dass $Q(0) = Q(2\pi)$, also separiert die entsprechende Algebra nicht die Punkte von $[0, 2\pi]$. Und $[0, 2\pi)$ ist wieder nicht kompakt.

Wir kommen nach dem nächsten Beispiel wieder auf die Behauptung oben zurück.

(4) Sei $S^n = \left\{p \in \mathbb{R}^{n+1} \,\middle|\, |p| = 1\right\}$.

Beh.: Sei $f : S^n \to \mathbb{R}$ stetig und $\varepsilon > 0$. Dann gibt es ein Polynom in $n+1$ Variablen, $P(x_1, x_2, \ldots, x_{n+1})$, so dass $|f(p) - P(p)| < \varepsilon$ für alle $p \in S^n$.

⌈ Betrachte zunächst die Algebra der Polynome in $n + 1$ Variablen $x_1, x_2, \ldots, x_n$ auf $\mathbb{R}^{n+1}$. Diese separiert Punkte auf $\mathbb{R}^{n+1}$ (da sie die Koordinatenfunktionen $(x_1, \ldots, x_{n+1}) \mapsto x_i$ enthält). Wir wählen nun $K = S^n$ (dies ist kompakt) und für A die Einschränkung der Polynome auf die Menge S^n (A separiert Punkte, weil die Funktionen dies schon vor der Einschränkung taten). Wie vorher folgt die Behauptung dann aus Satz 8.4.9. ⌋

Für $n = 3$ heißen die homogenen Polynome in A auch „Kugelflächenfunktionen“.

(3′) Wir wenden uns wieder der Behauptung aus Teil 3 zu. Für jedes f wie in der Behauptung gibt es ein eindeutiges $g : S^1 \to \mathbb{R}$, so dass $f(x) = g(e^{ix})$ (Warum?). Da f beschränkt ist, ist auch g beschränkt. ?

Beh.: g ist stetig.

⌈ Setze $\phi : \mathbb{R} \to S^1$, $\phi(x) = e^{ix}$. Betrachte die Einschränkung von ϕ auf ein Intervall $[a, b]$ mit $0 < b - a < 2\pi$. Dann ist die Funktion $\phi|_{[a,b]}$ injektiv, also bijektiv auf ihr Bild. Nach Satz 3.5.7 ist die Umkehrfunktion ϕ^{-1} stetig, und die Verkettung $g = f \circ \phi^{-1}$ ist auch stetig auf $\phi([a, b])$. Da dies für alle $[a, b]$ wie oben gilt, ist g insgesamt stetig auf S^1. ⌋

Nun verwenden wir Teil 4 für $n = 2$ um ein $P(x_1, x_2)$ zu finden mit $|g(p) - P(p)| < \varepsilon$ für alle $p \in S^1$. Die Funktion $Q(x) = P(\cos(x), \sin(x))$ hat dann die gewünschten Eigenschaften. (Warum?) ?

■ Für $\mathcal{C}_{\mathbb{C}}(K)$ statt $\mathcal{C}_{\mathbb{R}}(K)$ braucht man eine Zusatzannahme um eine analoge Aussage zu Satz 8.4.9 zu bekommen. Diese ist wie folgt. Eine Algebra A von Funktionen $X \to \mathbb{C}$ heisst *selbstadjungiert*, falls

$$f \in A \;\Rightarrow\; \bar{f} \in A \;,$$

wobei $\bar{f}$ für die Funktion $\bar{f}(x) = \overline{f(x)}$ steht, das heisst, x wird der komplex konjugierte Wert von $f(x)$ zugeordnet.

Satz 8.4.12. Sei K ein kompakter metrischer Raum und $A \subset \mathcal{C}_{\mathbb{C}}(K)$ eine selbstadjungierte Algebra (über $\mathbb{C}$ und mit 1), die auf K Punkte separiert. Dann $\overline{A} = \mathcal{C}_{\mathbb{C}}(K)$.

Beweis. Sei $A_0 = \{f \in A | f : K \to \mathbb{R}\}$. Dann ist A_0 eine Algebra in $\mathcal{C}_{\mathbb{R}}(K)$. (Warum?) ?

Beh.: A_0 separiert Punkte auf K.

⌈ Für $x \neq y$ sei $f \in A$ so, dass $f(x) \neq f(y)$. Wir zerlegen $f = u+iv$ mit $u, v : K \to \mathbb{R}$. Es gilt $u(x) \neq u(y)$ oder $v(x) \neq v(y)$. Ferner ist $u = \frac{1}{2}(f + \bar{f}) \in A_0$ und $v = \frac{1}{2i}(f - \bar{f}) \in A_0$ (hier geht „selbstadjungiert" ein, sonst gilt nicht unbedingt, dass $f \pm \bar{f} \in A$). ⌋

Nach Satz 8.4.9 gilt $\overline{A}_0 = \mathcal{C}_{\mathbb{R}}(K)$. Für $u, v \in \overline{A}_0$ gilt $f = u + iv \in \overline{A}$, also $\overline{A} = \mathcal{C}_{\mathbb{C}}(K)$. □

Beispiel 8.4.13. Wenn man die Voraussetzung „selbstadjungiert" weglässt, wird die Aussage falsch. Betrachte z.B. den Einheitskreis $K = S^1 = \{z \in \mathbb{C} | \, |z| = 1\}$ und die Algebra der Polynome $A = \{\sum_{n=0}^{M} c_n z^n \,|\, M \in \mathbb{N}, c_n \in \mathbb{C}\}$. Diese separiert Punkte (da $p(z) = z$ in A ist).

Für $f \in \mathcal{C}_{\mathbb{C}}(S^1)$ definiere

$$I(f) = \int_0^{2\pi} f(e^{i\varphi}) e^{i\varphi} \, d\varphi \;.$$

Da $\varphi \mapsto f(e^{i\varphi})e^{i\varphi}$ stetig ist, ist es integrierbar. Jetzt muss man folgende Behauptung zeigen:

Beh.: Für alle $f \in \overline{A}$ ist $I(f) = 0$.

(Warum gilt das? *Hinweis:* Überlegen Sie sich die Behauptung erst für $f \in A$, und prüfen Sie dann, dass man Satz 8.2.8 anwenden kann.) ?

Jetzt rechnet man noch nach, dass $I(\bar{z}) \neq 0$, woraus dann folgt, dass die Abbildung $z \mapsto \bar{z}$ zwar in $\mathcal{C}_{\mathbb{C}}(S^1)$ liegt, aber nicht in $\overline{A}$ liegt.

8.5 Fourier Reihen

Der Satz von Stone-Weierstraß gibt ein einfaches Kriterium für eine einleuchtende Annäherungseigenschaft („nirgendwo weiter weg als ein vorgegebenes ε“). Aber der Satz hilft nicht bei den Fragen:

- Wie finde ich explizt eine Approximation für eine Funktion f, die überall besser ist als das vorgegebene ε?

- Wenn ich mir nur einen endlich-dimensionalen Unterraum meines Funktionenraumes hernehme (z.B. Polynome bis zu einem maximalen Grad), was ist dann die kleinste Abweichung ε, die ich für ein vorgegebenes f erreichen kann?

In diesem Kapitel behandeln wir einen anderen Approximationsbegriff („am Besten im quadratischen Mittel“) und werden sehen, wie man diesbezüglich die obigen Fragen beantworten kann.

In diesem Kapitel steht $\mathcal{R}(a,b)$ für $\mathcal{R}_{\mathbb{C}}(a,b)$, wobei $a < b$.

Definition 8.5.1. Ein *orthogonales Funktionensystem* ist eine Folge von Funktionen $(\phi_n)_{n\in\mathbb{N}}$ mit $\phi_n \in \mathcal{R}(a,b)$ und $\phi_n \neq 0$, sodass

$$\int_a^b \phi_m(x)\,\overline{\phi_n(x)}\,dx \;=\; 0 \quad \text{für alle} \quad m \neq n\ .$$

Falls überdies für alle $n \in \mathbb{N}$ gilt, dass

$$\int_a^b \left|\phi_n(x)\right|^2 dx \;=\; 1\ ,$$

dann heißt $(\phi_n)_{n\in\mathbb{N}}$ *orthonormal*, oder auch ein *ON-System*.

Beispiel 8.5.2. Die folgenden $(\phi_n)_{n\in\mathbb{N}}$ sind Beispiele für orthogonale Funktionensysteme auf $[a,b]$:

$[a,b]$	ϕ_1	ϕ_2	ϕ_3	$\dots$	ϕ_n	ϕ_{n+1}	$\dots$
$[-\pi,\pi]$	$\frac{1}{\sqrt{2\pi}}$	$\frac{1}{\sqrt{2\pi}}e^{ix}$	$\frac{1}{\sqrt{2\pi}}e^{-ix}$	$\dots$	$\frac{1}{\sqrt{2\pi}}e^{mix}$	$\frac{1}{\sqrt{2\pi}}e^{-mix}$	$\dots$
$[-\pi,\pi]$	$\frac{1}{\sqrt{2\pi}}$	$\frac{1}{\sqrt{\pi}}\cos(x)$	$\frac{1}{\sqrt{\pi}}\sin(x)$	$\dots$	$\frac{1}{\sqrt{\pi}}\cos(mx)$	$\frac{1}{\sqrt{\pi}}\sin(mx)$	$\dots$
$[-1,1]$	$\sqrt{\frac{1}{2}}$	$\sqrt{\frac{3}{2}}\,x$	$\sqrt{\frac{5}{2}}\,\frac{3x^2-1}{2}$	$\dots$	$\sqrt{\frac{2n+1}{2}}\,P_{n-1}(x)$	$\dots$	

In den ersten beiden Zeilen gilt $n = 2m+1$ im allgemeinen Ausdruck (Warum stimmen die ersten beiden Zeilen?). In der letzten Zeile steht P_n für das n'te Legendre Polynom ($\rightarrow$Übung). ?

Definition 8.5.3. Sei $(\phi_n)_{n\in\mathbb{N}}$ ein ON-System auf $[a,b]$ und sei $f \in \mathcal{R}(a,b)$. Dann nennt man

$$c_n \;=\; \int_a^b f(x)\,\overline{\phi_n(x)}\,dx$$

den *n-ten Fourier-Koeffizienten von f bezüglich* $(\phi_n)_{n\in\mathbb{N}}$. Wir nennen

$$\sum_{n=1}^{\infty} c_n\phi_n(x)$$

die *Fourier-Reihe von f bezüglich* $(\phi_n)_{n\in\mathbb{N}}$ und schreiben auch $f \sim \sum_n c_n\phi_n$.

Bemerkung 8.5.4. In der Definition oben wird keine Bedingung an die Konvergenz der Reihe $\sum c_n\phi_n(x)$ gestellt. Daher verwenden wir das Symbol „$\sim$" anstatt „=". Somit ist „$f \sim \sum c_n\,\phi_n$" nur eine abkürzende Schreibweise für den Sachverhalt, dass die $(c_n)_{n\in\mathbb{N}}$ durch das obige Integral definiert sind. (Jedes einzelne Integral dagegen existiert nach Voraussetzung.)

Satz 8.5.5. Sei $(\phi_n)_{n\in\mathbb{N}}$ ein ON-System auf $[a,b]$ und sei $f \in \mathcal{R}(a,b)$. Sei ferner $f \sim \sum_n c_n\,\phi_n$, und seien $\gamma_n \in \mathbb{C}$, $n \in \mathbb{N}$. Schreibe

$$s_N(x) \;=\; \sum_{k=1}^{N} c_k\,\phi_k(x) \qquad \text{und} \qquad t_N(x) \;=\; \sum_{k=1}^{N} \gamma_k\,\phi_k(x)\ .$$

(Also ist s_N die N'te Partialsumme der Reihe $\sum c_n\phi_n$.) Dann gilt

$$\int_a^b \big|f(x) - s_N(x)\big|^2 dx \;\le\; \int_a^b \big|f(x) - t_N(x)\big|^2 dx$$

mit „=" genau dann, wenn $c_k = \gamma_k$, $k = 1,\dots,N$.

Beweis. Wir kürzen ab $\int f\,\overline{\phi_m} = \int_a^b f(x)\overline{\phi_m(x)}dx$, etc.

Beh.: Für $m = 1,\dots,N$ gilt $\int (f - s_N)\,\overline{\phi_m} = 0 = \int (\overline{f} - \overline{s_N})\,\phi_m$.

⌈ Es gilt

$$\int (f - s_N)\,\overline{\phi_m} \;=\; \underbrace{\int f\,\overline{\phi_m}}_{=c_m} \;-\; \underbrace{\int s_N\,\overline{\phi_m}}_{=\sum_{k=1}^N c_k \int \phi_k\,\overline{\phi_m}} \;=\; c_m - c_m = 0$$

Die zweite Gleichheit folgt aus $\overline{\int f} = \int \overline{f}$. (Wie? Und warum gilt das überhaupt?). ?

⌋

Wir berechnen

$$\begin{aligned}\int |f-t_N|^2 &= \int \big|(f-s_N)+(s_N-t_N)\big|^2 \\ &= \int |f-s_N|^2 + \int |s_N-t_N|^2 \\ &\quad + \underbrace{\int (f-s_N)(\overline{s_N}-\overline{t_N})}_{=0\ (*)} + \underbrace{\int (\overline{f}-\overline{s_N})(s_N-t_N)}_{=0\ (*)} \\ &\geq \int |f-s_N|^2\end{aligned}$$

Hier folgt $(*)$ jeweils aus der Behauptung oben (Wie?). In der Ungleichung am Ende gilt „=“ genau dann, wenn ?

$$0 = \int |s_N-t_N|^2 = \sum_{k,l=1}^{N} (c_k-\gamma_k)(\overline{c_l}-\overline{\gamma_l}) \int \phi_k\overline{\phi_l} = \sum_{k=1}^{N} |c_k-\gamma_k|^2 ,$$

also wenn $c_k = \gamma_k$ für alle $k = 1, \dots, N$. □

Bemerkung 8.5.6.

(1) Den Ausdruck $\int_a^b \big|f(x)-s_N(x)\big|^2 dx$ aus Satz 8.5.5 können wir als „quadratisches Mittel“ interpretieren. Damit beantwort Satz 8.5.5 die Fragen vom Anfang des Kapitels: Er beschreibt die beste Näherung an ein gegebenes f im quadratischen Mittel durch eine Linearkombination von Funktionen $\phi_1, \dots, \phi_N$, wobei die Koeffizienten durch explizite Integrale bestimmt sind.

(2) Für $f \in \mathcal{R}(a,b)$ schreibt man auch $\|f\|_2 = \sqrt{\int_a^b |f(x)|^2 dx}$ und nennt dies die *L2-Norm von f* (→Übung). Die Sup-Norm dagegen war gegeben durch $\|f\|_\infty = \sup\{|f(x)| \mid x \in [a,b]\}$. Es gilt $\|f\|_2 \leq \sqrt{b-a}\,\|f\|_\infty$. In die umgekehrte Richtung gibt es keine solche Abschätzung (→Übung). Die L2-Norm verrät einem also im Allgemeinen nichts über die Sup-Norm.

(3) Satz 8.5.5 sagt, dass $\sum_{k=1}^N c_k\phi_k$ die beste Annäherung an f unter allen Funktionen der Form $\sum_{k=1}^N \gamma_k\phi_k$ in der L2-Norm ist. Das ist aber im Allgemeinen nicht die beste Annäherung in der Sup-Norm (sogar typischerweise nicht).

Ein einfaches Beispiel ist wie folgt. Sei $0 < r < 1$ und $f : [0,1] \to \mathbb{C}$, $f(x) = 0$ für $x \leq r$ und $f(x) = 1$ für $x > r$. Betrachte $N = 1$ und $\phi_1(x) = 1$. (Für welches $c_1 \in \mathbb{C}$ ist $c_1\phi_1$ die beste Näherung an f in der L2-Norm? Und in der Sup-Norm?) ?

Satz 8.5.7. (Besselsche Ungleichung)
Sei $f \in \mathcal{R}(a,b)$, sei $(\phi_n)_{n\in\mathbb{N}}$ ein ON-System auf $[a,b]$, und sei $f \sim \sum_n c_n \phi_n$. Dann gilt

$$\sum_{n=1}^{\infty} |c_n|^2 \ \le \ \int_a^b \big|f(x)\big|^2 \, dx \ .$$

Beweis. Im Beweis von Satz 8.5.5 hatten wir gesehen, dass (in der Notation wie dort benutzt)

$$\int |f - t_N|^2 \ = \ \int |f - s_N|^2 \ + \ \int |s_N - t_N|^2 \ .$$

Dies wenden wir jetzt auf $t_N = 0$ an und lassen auf der rechten Seite den ersten Summanden weg (im Beweis von Satz 8.5.5 hatten wir dagegen den zweiten Summanden weggelassen):

$$\int |f|^2 \ \ge \ \int |s_N|^2 \ .$$

Die Aussage folgt nun aus $\int |s_N|^2 = \sum_{k=1}^N |c_k|^2$. (Wieso gilt das? Und wieso folgt die Aussage?) ? □

Bemerkung 8.5.8. Insbesondere folgt aus Satz 8.5.7, dass $\lim_{n\to\infty} c_n = 0$. Das ist überraschend, weil wir außer Integrierbarkeit und ON-System gar keine weiteren Annahmen an f oder die ϕ_n gemacht haben. Z.B. ist die Antwort auf die Frage: „Gibt es ein $f \in \mathcal{R}(a,b)$, so dass $\big|\int_a^b f(x)\sin(mx)dx\big| \ge 1$ für alle $m \in \mathbb{N}$?“ damit „Nein.“ (Warum?) ?

Definition 8.5.9. Ein *trigonometrisches Polynom* ist eine Summe der Form

$$a_0 + \sum_{n=1}^{N} \Big(a_n \cos(nx) + b_n \sin(nx) \Big) \qquad \text{bzw.} \qquad \sum_{n=-N}^{N} d_n e^{inx} \ ,$$

mit $x \in \mathbb{R}$ und $a_i, b_i, d_i \in \mathbb{C}$. Eine *trigonometrisches Reihe* erhält man, wenn man in diesen Ausdrücken N durch ∞ ersetzt. Dies steht in beiden Fällen für den Grenzübergang $N \to \infty$, insbesondere erhält man die Notation

$$\sum_{n=-\infty}^{\infty} d_n e^{inx} \ \overset{\text{Def.}}{=} \ \lim_{N\to\infty} \Big(\sum_{n=-N}^{N} d_n e^{inx} \Big) \ .$$

■ Für den Rest des Kapitels bezeiche f eine Funktion mit den folgenden Eigenschaften:

- $f : \mathbb{R} \to \mathbb{C}$
- f ist 2π-periodisch, also $f(x+2\pi) = f(x)$ für alle $x \in \mathbb{R}$.
- f ist auf $[-\pi, \pi]$ integrierbar, also $f|_{[-\pi,\pi]} \in \mathcal{R}(-\pi, \pi)$.

Es folgt, dass f auf jedem Intervall $[a, b]$ integrierbar ist, und es gilt für jedes $c \in \mathbb{R}$, dass

$$\int_{-\pi}^{\pi} f(x)\,dx \;=\; \int_{c}^{c+2\pi} f(x)\,dx \;.$$

(Warum?) ?

■ Betrachte das erste ON-System aus Beispiel 8.5.2, etwas umnummeriert: $\phi_k(x) = \frac{1}{\sqrt{2\pi}} e^{ikx}$ auf $[-\pi, \pi]$, $k \in \mathbb{Z}$. Dann bedeutet $f(x) \sim \sum c_k \phi_k(x) = \sum c_k \frac{1}{\sqrt{2\pi}} e^{ikx}$ im Sinne von Definition 8.5.3, dass

$$c_k = \int_{-\pi}^{\pi} f(x)\, \tfrac{1}{\sqrt{2\pi}} e^{-ikx}\,dx$$

Um nicht so viele $\sqrt{2\pi}$ schreiben zu müssen, verteilt man die Faktoren anders. Man setzt $d_k = c_k \frac{1}{\sqrt{2\pi}}$, und in diesem Fall benutzt man auch die Notation $\widehat{f}(k) = d_k$. Insgesamt erhält man die trigonometrische Reihe

$$f \sim \sum_{k=-\infty}^{\infty} \widehat{f}(k)\, e^{ikx} \qquad \text{mit} \qquad \widehat{f}(k) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x)\, e^{-ikx}\,dx \;.$$

Die Bedeutung von „$\sim$“ ist hier also leicht anders als in Definition 8.5.3, aber es gibt keine Verwechselungsgefahr, weil die e^{ikx} ohne den Faktor $1/\sqrt{2\pi}$ nicht orthonormal sind.

■ Für $m, n \in \mathbb{Z}$ gilt

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} e^{imx}\, \overline{e^{inx}}\,dx \;=\; \begin{cases} 1 & ; m = n \\ 0 & ; m \neq n \end{cases} \;.$$

Die Besselsche Ungleichung (Satz 8.5.7) lautet hier

$$\sum_{k \in \mathbb{Z}} \left|\widehat{f}(k)\right|^2 \;\leq\; \frac{1}{2\pi} \int_{-\pi}^{\pi} \left|f(x)\right|^2 dx \;.$$

(Warum?) Aus Satz 4.9.6 über die Umordnung von absolut konvergenten Reihen wissen wir, dass die genaue Reihenfolge, in der man die Summe über $\mathbb{Z}$ berechnet, keine Rolle spielt. ?

Beispiel 8.5.10. Betrachte die Funktion $f : \mathbb{R} \to \mathbb{C}$, die auf $[0, 2\pi)$ gegeben ist durch

$$f(x) = \begin{cases} 1 & ;\ x \in (0, \pi) \\ 0 & ;\ x \in (\pi, 2\pi) \\ \text{egal} & ;\ x = 0\,,\ x = \pi \end{cases} .$$

Die Funktion f ist dann auf $\mathbb{R}$ durch die 2π-periodische Fortsetzung gegeben. Es gilt: (Warum?) ?

$$\widehat{f}(k) \ = \ \frac{1}{2\pi}\int_0^{2\pi} f(x)\, e^{-ikx}\, dx \ = \ \frac{1}{2\pi}\int_0^{\pi} e^{-ikx}\, dx \ = \ \begin{cases} \frac{1}{2} & ;\ k = 0 \\ \frac{1}{\pi i k} & ;\ k \text{ ungerade} \\ 0 & ;\ \text{sonst} \end{cases}$$

Wenn wir die Terme für k und $-k$ zusammenfassen, erhalten wir die Fourier-Reihe (wobei wir $k = 2n + 1$ gesetzt haben, damit wir nur ungerade Terme summieren)

$$f(x) \ \sim \ \frac{1}{2} + \sum_{n=0}^{\infty} \frac{2}{\pi(2n+1)} \sin\big((2n+1)x\big) \ .$$

Wir lernen eine ganze Reihe von Dingen:

- Für $x = 0, \pi$ hatten wir den Wert von $f(x)$ nicht vorgegeben, aber die Fourier-Reihe konvergiert in diesem Fall (trivialerweise) gegen $\frac{1}{2}$, was genau die Hälft der Sprunghöhe ist.
  (Zur Info: Dies ist immer so, siehe [Kap. 16.3 aus Königsberger 1].)
- Wir wissen ersteinmal nicht, für welche $x \in \mathbb{R}$ die Fourier-Reihe auf der rechten Seite konvergiert, und gegen was. Aber Satz 8.5.12 unten wird uns sagen, dass die Reihe für $x \notin \pi\mathbb{Z}$ punktweise gegen f konvergiert.

  Da die einzelne Summanden stetig sind, die Grenzfunktion aber nicht, konvergieren Fourier-Reihen also im Allgemeinen nicht gleichmäßig.
- Wenn wir das Konvergenzergebnis aus Satz 8.5.12 voraussetzen und $x = \frac{\pi}{2}$ einsetzen, so erhalten wir: (Wie?) ?

$$\frac{\pi}{4} \ = \ \sum_{n=0}^{\infty} \frac{(-1)^n}{2n+1} \ = \ 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \cdots$$

■ Für die Partialsummen der Fourier-Reihe von f benutzen wir auch folgende Notation:

$$s_N(f; x) \ = \ s_N(x) \ = \ \sum_{k=-N}^{N} \widehat{f}(k)\, e^{ikx} \ .$$

■ Der *Dirichlet-Kern* ist gegeben durch

$$D_N(x) \;=\; \sum_{k=-N}^{N} e^{ikx} \;\overset{(*)}{=}\; \begin{cases} 2N+1 & ; x \in 2\pi\mathbb{Z} \\ \sin((N+\frac{1}{2})x)/\sin(\frac{1}{2}x) & ; x \notin 2\pi\mathbb{Z} \end{cases} .$$

(Warum gilt $(*)$? *Hinweis:* Prüfen Sie zuerst durch Ausmultiplizieren, dass $(e^{ix/2} - e^{-ix/2})D_N(x) = e^{i(N+\frac{1}{2})x} - e^{-i(N+\frac{1}{2})x}$.) ?
(Übrigens gibt es keine Funktion f, so dass $s_N(f;x) = D_N(x)$ für alle N gilt. Warum?) ?

Lemma 8.5.11. Der Dirichlet-Kern erfüllt

$$s_N(f;x) \;=\; \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x-t)\, D_N(t)\, dt \;.$$

Beweis. Es gilt

$$\begin{aligned} s_N(f;x) \;&=\; \sum_{k=-N}^{N}\Big(\frac{1}{2\pi}\int_{-\pi}^{\pi} f(u)e^{-iku}du\Big)e^{ikx} \;=\; \frac{1}{2\pi}\int_{-\pi}^{\pi} f(u)\Big(\sum_{k=-N}^{N} e^{ik(x-u)}\Big)du \\ &=\; \frac{1}{2\pi}\int_{-\pi}^{\pi} f(u)D_N(x-u)\,du \;\overset{(*)}{=}\; \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x-t)D_N(t)\,dt \;. \end{aligned}$$

wobei in $(*)$ die Substitution $t = x - u$ ausgeführt wird. (Aber sollte das nicht die Integralgrenzen ändern? Und ein Minuszeichen geben?) ? □

Satz 8.5.12. Angenommen, für ein $x \in \mathbb{R}$ existieren $\delta > 0$ und M, sodass

$$\forall t \in (-\delta, \delta) \;:\; \big|\, f(x+t) - f(x) \,\big| \;\le\; M\,|t|$$

(d.h. f ist Lipschitz-stetig bei x). Dann gilt

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

Beweis. Definiere $g : [-\pi,\pi] \to \mathbb{C}$ wie folgt. Für $0 < |t| \le \pi$ setze

$$g(t) = \frac{f(x-t) - f(x)}{\sin\frac{t}{2}} \qquad \text{und} \qquad g(0) = 0 \;.$$

Die Voraussetzung an f impliziert, dass g auf $[-\pi,\pi]$ beschränkt ist (Warum?). ?

Beh.: $g \in \mathcal{R}(-\pi,\pi)$.

(Warum gilt die Behauptung? *Hinweis:* Überlegen Sie, warum g für jedes $\varepsilon > 0$ auf $[-\pi,-\varepsilon]$ und $[\varepsilon,\pi]$ integrierbar ist. Und dann, warum dies ?

zusammen mit Beschränktheit ausreicht, um das Kriterium in Lemma 7.3.6 zu erfüllen.)

Wir brauchen auch die Identität

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} D_N(t)\,dt \;=\; 1\ . \qquad (*)$$

(Warum gilt die?) Damit rechnen wir weiter ?

$$\begin{aligned}
s_N(f;x) - f(x) &\overset{\text{Lem. 8.5.11}}{=} \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x-t)\,D_N(t)\,dt \;-\; f(x) \\
&\overset{(*)}{=} \frac{1}{2\pi}\int_{-\pi}^{\pi} \big(f(x-t)-f(x)\big)\,D_N(t)\,dt \\
&= \frac{1}{2\pi}\int_{-\pi}^{\pi} g(t)\,\sin\big((N+\tfrac{1}{2})t\big)\,dt \\
&\overset{\text{Add.}}{=} \frac{1}{2\pi}\int_{-\pi}^{\pi}\Big(g(t)\,\cos\tfrac{t}{2}\Big)\,\sin(Nt)\,dt \;+\; \frac{1}{2\pi}\int_{-\pi}^{\pi}\Big(g(t)\,\sin\tfrac{t}{2}\Big)\,\cos(Nt)\,dt\ .
\end{aligned}$$

In Schritt „Add.“ haben wir eines der Additionstheoreme für sin, cos benutzt:

$$\sin(a+b) = \sin(a)\cos(b) + \cos(a)\sin(b)\ .$$

(Rechnen Sie dies einmal mit der Exponentialdarstellung nach. Vielleicht auch gleich die entsprechende Aussage für Cosinus.) ?

Da $g(t)\cos\frac{t}{2}$ und $g(t)\sin\frac{t}{2}$ in $\mathcal{R}(-\pi,\pi)$ sind, können wir die Besselsche Ungleichung (Satz 8.5.7), und genauer Bemerkung 8.5.8, auf das ON-System $\sin(mt)/\sqrt{2\pi}$, $\cos(mt)/\sqrt{2\pi}$ anwenden. Demnach geht jedes der beiden Integrale oben gegen 0 für $N \to \infty$.

(Wieso darf man die Besselsche Ungleichung überhaupt anwenden? Da steht doch einmal $g(t)\cos\frac{t}{2}$ und einmal $g(t)\sin\frac{t}{2}$.) ? □

Bemerkung 8.5.13.

(1) Wenn zwei Potenzreihen um 0 in einer Umgebung $(-\varepsilon,\varepsilon)$ von 0 übereinstimmen, dann sind sie schon im gesamten Definitionsbereich $(-R,R)$ gleich (R: Konvergenzradius), siehe Korollar 8.3.5 (ii).

Bei Fourier-Reihen ist das anders. Für zwei Funktionen $f \neq g$, die den Bedingungen von Satz 8.5.12 genügen, und die in einer kleinen Umgebung $(-\varepsilon,\varepsilon)$ von 0 übereinstimmen (Fällt Ihnen ein konkretes Beispiel für f, g ein?), haben trotzdem verschiedene Fourier-Reihen. ?

(2) Zur Info: Jedes f, dass die Bedingung in Satz 8.5.12 erfüllt, ist stetig. Umgekehrt reicht es aber nicht aus, nur anzunehmen, dass f stetig ist. Ein (etwas aufwendiges) Gegenbeispiel wird in [Kap. 16.4 aus Königsberger 1] konstruiert. Dort divergiert die Fourier-Reihe einer stetigen Funktion in einem Punkt.

Nebenbemerkung zur Information: Als nächstes könnte man z.B. den Parsevalschen Satz behandeln. Der sagt unter anderem folgendes (siehe [Satz 8.16 aus Rudin] oder [Kap. 16.7 aus Königsberger 1]): Für eine 2π-periodische Funktion f mit $f|_{[-\pi,\pi]} \in \mathcal{R}(-\pi,\pi)$ gilt

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} |f(x)|^2\,dx \;=\; \sum_{k\in\mathbb{Z}} \left|\,\widehat{f}(k)\,\right|^2 .$$

Dies kann man dann z.B. auf das Beispiel $f(x) = x$ für $x \in (-\pi,\pi]$ anwenden (siehe Aufgabe 2 von Zettel 7). Dann $\widehat{f}(0) = 0$ und $|\widehat{f}(k)|^2 = k^{-2}$ ($k \neq 0$), und somit

$$\frac{\pi^2}{3} \;=\; \frac{1}{2\pi}\int_{-\pi}^{\pi} x^2\,dx \;=\; 2\sum_{k=1}^{\infty}\frac{1}{k^2} .$$

9 Funktionen in mehreren Variablen

9.1 Die Operatornorm

■ Notation für lineare Abbildungen: Seien V, W Vektorräume (über einem gegebenen Körper, z.B. über $\mathbb{R}$). Dann schreiben wir

$$L(V,W) \overset{\text{Def.}}{=} \left\{ A : V \to W \,\middle|\, A \text{ linear} \right\} .$$

Statt $L(V,V)$ schreiben wir auch kurz $L(V)$.

Definition 9.1.1. Für $A \in L(\mathbb{R}^m, \mathbb{R}^n)$ setze

$$\|A\|_{\text{op}} \overset{\text{Def.}}{=} \sup \left\{ |Ax| \,\middle|\, x \in \mathbb{R}^m \,,\, |x| \le 1 \right\} ,$$

wobei $|Ax|$ für die Euklidische Norm im $\mathbb{R}^n$ steht, und $|x|$ für die Euklidische Norm im $\mathbb{R}^m$. Wir nennen $\|A\|_{\text{op}}$ die *Operatornorm* auf $L(\mathbb{R}^m, \mathbb{R}^n)$, und schreiben auch kurz $\|A\|$.

Bemerkung 9.1.2.

(1) Die Abbildung $x \mapsto |Ax|$ ist stetig (Warum?), und nimmt daher ihr Maximum auf der kompakten Menge $\left\{ x \in \mathbb{R}^m \,|\, |x| \le 1 \right\}$ an. Somit gilt $\|A\|_{\text{op}} < \infty$. ?

(Nebenbemerkung: Für unendlich dimensionale normierte Vektorräume gilt das nicht mehr. →Übung.)

Es gilt auch (Warum?) ?

$$\|A\|_{\text{op}} = \sup \left\{ |Ax| \,\middle|\, x \in \mathbb{R}^m \,,\, |x| = 1 \right\} = \sup \left\{ \frac{|Ax|}{|x|} \,\middle|\, x \in \mathbb{R}^m \,,\, x \neq 0 \right\} .$$

(2) Die Operatornorm $\| - \|_{\text{op}}$ ist nicht das gleiche wie die Supremumsnorm $\| - \|_\infty$. In der Tat gilt für $A \in L(\mathbb{R}^m, \mathbb{R}^n)$, dass $\|A\|_{\text{op}} < \infty$, aber

$$\|A\|_\infty < \infty \quad \Leftrightarrow \quad A = 0 .$$

(Warum?) Aber es gibt folgende Beziehung. Sei $B^m \subset \mathbb{R}^m$ der abgeschlossene Einheitsball, $B^m = \{x \in \mathbb{R}^m |\, |x| \le 1\}$, und es bezeichne $A|_{B^m}$ die Einschränkung von A auf B^m. Dann gilt: ?

$$\|A\|_{\text{op}} = \left\| \, A|_{B^m} \right\|_\infty$$

(Warum?) ?

(3) Für alle $x \in \mathbb{R}^m$ gilt

$$|Ax| \;\leq\; \|A\|_{\text{op}}\,|x| \;,$$

denn für $x = 0$ ist dies klar und für $x \neq 0$ folgt dies aus dem letzten Ausdruck für $\|A\|_{\text{op}}$ (Details?). Wenn man die Operatornorm nach unten abschätzen will, kann man den Ausdruck auch so umschreiben: ?

$$\text{für } x \neq 0: \quad \|A\|_{\text{op}} \;\geq\; \frac{|Ax|}{|x|} \;.$$

Wenn man die Operatornorm nach oben abschätzen will, kann man versuchen ein $\lambda \in \mathbb{R}$ zu finden, so dass gilt $\forall x \in \mathbb{R}^m \;:\; |Ax| \leq \lambda |x|$. Dann folgt $\|A\|_{\text{op}} \leq \lambda$. (Warum?) ?

(4) Die Operatornorm is *sub-multiplikativ* . D.h. für $A \in L(\mathbb{R}^m, \mathbb{R}^n)$ und $B \in L(\mathbb{R}^k, \mathbb{R}^m)$ gilt

$$\|AB\|_{\text{op}} \;\leq\; \|A\|_{\text{op}}\,\|B\|_{\text{op}} \;.$$

Dies folgt aus der Beobachtung, dass für alle $x \in \mathbb{R}^k$ gilt $|ABx| \leq \|A\|_{\text{op}}|Bx| \leq \|A\|_{\text{op}}\|B\|_{\text{op}}|x|$, und aus Teil 3.

Beispiel 9.1.3.

(1) Für die Identitätsabbildung $id \in L(\mathbb{R}^n)$ gilt: $\|id\|_{\text{op}} = 1$. (Warum?) ?

(2) Sei $A \in L(\mathbb{R}, \mathbb{R}^2)$ die Abbildung mit darstellender Matrix $\left(\begin{smallmatrix}1\\2\end{smallmatrix}\right)$. Also $Ax = (x, 2x)$. Dann gilt für $x \neq 0$, dass $|Ax|/|x| = \sqrt{x^2 + (2x)^2}/|x| = \sqrt{5}$. Also $\|A\|_{\text{op}} = \sqrt{5}$.

(3) Sei $A \in L(\mathbb{R}^2, \mathbb{R}^2)$ die Abbildung mit darstellender Matrix $\left(\begin{smallmatrix}0&1\\0&0\end{smallmatrix}\right)$. Also $A(x_1, x_2) = (x_2, 0)$. Dann gilt $|Ax| = |x_2|$, und das wird unter allen x mit $|x| = 1$ maximiert durch $(0, \pm 1)$. Also $\|A\|_{\text{op}} = 1$.

Lemma 9.1.4. Sei $A \in L(\mathbb{R}^m, \mathbb{R}^n)$. Dann ist A gleichmäßig stetig auf $\mathbb{R}^m$.

Beweis. Wir müssen zeigen:

$$\forall \varepsilon > 0 \;\exists \delta > 0 \;\forall x, y \in \mathbb{R}^m \;:\; |x - y| < \delta \;\Rightarrow\; |Ax - Ay| < \varepsilon \;.$$

Aus Bemerkung 9.1.2 (1) wissen wir, dass $\|A\|_{\text{op}}$ endlich ist, und aus Teil 3, dass $|Ax - Ay| = |A(x - y)| \leq \|A\|_{\text{op}}\,|x - y|$. Wir können also $\delta = \varepsilon/\|A\|_{\text{op}}$ wählen. $\square$

In Aufgabe 6 auf Zettel 7 haben wir den Begriff einer Norm auf einem Vektorraum definiert. Wir wenden dies auf den $\mathbb{R}$-Vektorraum $L(\mathbb{R}^m, \mathbb{R}^n)$ an:

Satz 9.1.5. $(\, L(\mathbb{R}^m, \mathbb{R}^n)\,,\, \| - \|_{\mathrm{op}}\,)$ ist ein normierter $\mathbb{R}$-Vektorraum.

Der Beweis ist eine →Übung. Insbesondere ist $d(A,B) \overset{\text{Def.}}{=} \|A-B\|_{\mathrm{op}}$ eine Metrik auf $L(\mathbb{R}^m, \mathbb{R}^n)$.

Definition 9.1.6. Wir bezeichnen mit $GL(\mathbb{R}^n)$ (oder kurz GL_n) die Menge der invertierbaren linearen Abbildungen $\mathbb{R}^n \to \mathbb{R}^n$:

$$GL_n \;=\; \left\{\, A \in L(\mathbb{R}^n) \,\middle|\, A \text{ invertierbar} \,\right\} \;\subset\; L(\mathbb{R}^n)\ .$$

Satz 9.1.7.

(i) Sei $A \in GL_n$ und $B \in L(\mathbb{R}^n)$. Falls

$$\|A - B\| \;<\; \frac{1}{\|A^{-1}\|}\ ,$$

so folgt, dass $B \in GL_n$.

(ii) GL_n ist eine offene Teilmenge von $L(\mathbb{R}^n)$.

(iii) Einer invertierbaren Matrix ihr Inverses zuordnen ist stetig. D.h. die Abbildung $GL_n \to GL_n$, $A \mapsto A^{-1}$ ist stetig.

(Warum ist $\|A^{-1}\|$ in Teil 1 ungleich 0? Muss man wirklich $(\|A^{-1}\|)^{-1}$ sagen, oder kann man den "doppelten Kehrwert" zu $\|A\|$ vereinfachen? Ist die Abbildung in Teil 3 auch gleichmäßig stetig auf GL_n?) ?

Beweis.

(i) Beh.: B ist injektiv

⌈ Sei $x \in \mathbb{R}^n$ gegeben. Wir werden zeigen, dass $|Bx| \geq \lambda |x|$ für ein $\lambda > 0$. Falls $Bx = 0$, so folgt dann, dass $|x| = 0$, also auch $x = 0$, und somit gilt die Behauptung.
Betrachte die Abschätzungen

$$\begin{aligned} |Ax| \;&=\; |Bx + (A-B)x| \leq |Bx| + |(A-B)x| \\ &\leq\; |Bx| + \|A-B\|\,|x| \end{aligned}$$

und

$$|x| \;=\; |A^{-1}Ax| \;\leq\; \|A^{-1}\|\,|Ax| \quad\Rightarrow\quad \frac{1}{\|A^{-1}\|}|x| \leq |Ax|\ .$$

Zusammen erhält man

$$|Bx| \;\geq\; \underbrace{\left(\frac{1}{\|A^{-1}\|} - \|A-B\|\right)}_{>0 \text{ per Annahme}} |x|\ . \qquad (*)$$

⌋

Die Invertierbarkeit von B folgt nun aus einem Ergebnis der Linearen Algebra:

> Sei $F : V \to W$ eine lineare Abbildung zwischen zwei Vektorräumen. Angenommen, V und W sind endlich-dimensional mit $\dim(V) = \dim(W)$. Dann sind äquivalent: a) F is injektiv, b) F ist surjektiv, und c) F ist bijektiv.
>
> (Braucht man da wirklich, dass V und W endlich-dimensional sind? Und die gleiche Dimension haben?) ?

(ii) Folgt aus Teil (i): Für jedes $A \in GL_n$ enthält GL_n den offenen Ball um A in $L(\mathbb{R}^n)$ von Radius $1/\|A^{-1}\|$.

(iii) Sei $A \in GL_n$. Wir müssen zeigen:

$$\forall \varepsilon > 0\ \exists \delta > 0\ \forall B \in GL_n\ :\ \|A - B\| < \delta\ \Rightarrow\ \|A^{-1} - B^{-1}\| < \varepsilon\ .$$

Zunächst einmal gilt

$$\|A^{-1} - B^{-1}\| \ =\ \|A^{-1}(B - A)B^{-1}\| \ \leq\ \|A^{-1}\|\,\|B^{-1}\|\,\|A - B\|\ . \quad (**)$$

Wir werden jetzt noch $\|B^{-1}\|$ abschätzen, dann können wir ein passendes δ bestimmen.

Setze $x = B^{-1}y$ in $(*)$:

$$|y| \ \geq\ \left(\frac{1}{\|A^{-1}\|} - \|A - B\|\right) |B^{-1}y|\ .$$

Also (Warum?) ?

$$\|B^{-1}\| \ \leq\ \frac{\|A^{-1}\|}{1 - \|A^{-1}\|\,\|A - B\|}$$

Für $\delta < \frac{1}{2\|A^{-1}\|}$ folgt also $\|B^{-1}\| \leq 2\|A^{-1}\|$. Wenn man überdies $\delta < \frac{\varepsilon}{2\|A^{-1}\|^2}$ fordert, so wird aus $(**)$ die gewünschte Ungleichung $\|A^{-1} - B^{-1}\| < \varepsilon$.

$\square$

Beispiel 9.1.8.

(1) Für $n = 1$ sind A, B von der Form $Ax = a \cdot x$ und $Bx = b \cdot x$ für $a, b \in \mathbb{R}$, $a \neq 0$, und es gilt $\|A\| = |a|$, etc. Die einzelnen Aussagen des Satzes sind dann schon bekannt:

(i) Falls $|a-b| < 1/|a^{-1}| = |a|$, so folgt $b \neq 0$.
(ii) $\mathbb{R} \setminus \{0\}$ ist offen in $\mathbb{R}$.
(iii) Die Abbildung $a \mapsto a^{-1}$ ist auf $\mathbb{R} \setminus \{0\}$ stetig.

(2) Die Abschätzung in Teil (i) ist scharf, z.B. würde „$\leq$" nicht ausreichen. (Fällt Ihnen für jedes n ein Beispiel ein?) ?

(3) Zur Info: Wenn man mehr Ergebnisse aus der Linearen Algebra benutzt, kann man Teil (ii) und (iii) des Satzes auch anders zeigen:

(ii) $B \in GL_n$ genau dann, wenn $\det(B) \neq 0$. Aber $\det(B)$ ist ein Polynom in den Einträgen der Matrixdarstellung von B. Also ist $B \mapsto \det(B)$ stetig. Somit ist das Urbild der offenen Menge $\mathbb{R} \setminus \{0\}$ offen. (Hier benutzt man implizit die Äquivalenz von Normen auf endlich dimensionalen Vektorräumen. Sehen Sie, wo?) ?

(iii) Die Matrixdarstellung von A^{-1} ist durch $A^{\#}/\det(A)$ gegeben, wobei $A^{\#}$ die Adjunkte von A ist. Die Einträge von $A^{\#}$ sind polynomial in denen von A, also ist $A \mapsto A^{\#}$ stetig, also auch $A \mapsto A^{-1}$.

■ Notation für Matrixdarstellung einer linearen Abbildung: Seien V, W endlich dimensionale Vektorräume, und sei $\{v_1, \ldots, v_m\}$ eine Basis von V und $\{w_1, \ldots, w_n\}$ eine Basis von W. Dann gibt es eindeutige Konstanten a_{ij}, sodass für alle $j = 1, \ldots, m$ gilt

$$Av_j = \sum_{i=1}^{n} a_{ij}\, w_i$$

(Sind die Indizes da auch richtig rum?). Die *Matrixdarstellung von A bezüglich $\{v_j\}$ und $\{w_i\}$* ist ?

$$[A] = \begin{pmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nm} \end{pmatrix} .$$

Für welche Basiswahlen die Darstellung $[A]$ gemeint ist, muss aus dem Zusammenhang klar sein oder dazu gesagt werden.

Bemerkung 9.1.9. Sei $A \in L(\mathbb{R}^m, \mathbb{R}^n)$ und sei $[A]$ die Matrixdarstellung bezüglich der Standardbasen von $\mathbb{R}^m$ und $\mathbb{R}^n$. Setze

$$\|A\|_F \overset{\text{Def.}}{=} \sqrt{\textstyle\sum_{i,j} (a_{ij})^2} .$$

Dies ist ebenfalls eine Norm auf $L(\mathbb{R}^m, \mathbb{R}^n)$ (die *Frobenius-Norm*). (Warum ist $\|-\|_F$ eine Norm?) Die Norm $\|-\|_F$ ist leichter zu berechnen als $\|-\|_{\text{op}}$, aber es gilt z.B. für $id \in L(\mathbb{R}^n)$, dass $\|id\|_F = \sqrt{n}$ aber $\|id\|_{\text{op}} = 1$. ?

Satz 9.1.10. Sei S ein metrischer Raum und $A : S \to L(\mathbb{R}^m, \mathbb{R}^n)$, $p \mapsto A(p)$ eine Abbildung mit Matrixdarstellung

$$[A(p)] = \begin{pmatrix} a_{11}(p) & \cdots & a_{1m}(p) \\ \vdots & & \vdots \\ a_{n1}(p) & \cdots & a_{nm}(p) \end{pmatrix} \quad .$$

Dann ist A genau dann stetig (bzgl. der Operatornorm auf $L(\mathbb{R}^m, \mathbb{R}^n)$), wenn alle $a_{ij} : S \to \mathbb{R}$ ($i = 1, \dots, n$ und $j = 1, \dots, m$) stetig sind.

Beweis. Bezüglich der Norm $\| - \|_F$ ist dies gerade die Aussage von Analysis 1, Zettel 11, Aufgabe 5. Nach Zettel 7, Aufgabe 6 sind $\| - \|_F$ und $\| - \|_{\text{op}}$ äquivalent. (Wieso hilft es einem hier, dass die beiden Normen äquivalent sind?) ? □

9.2 Differentiation

■ Wir können schon ableiten:

$$\begin{aligned} f : (a,b) \to \mathbb{R} \quad &, \qquad f'(x) = \lim_{h\to 0} \frac{f(x+h) - f(x)}{h} \in \mathbb{R} \ , \\ f : (a,b) \to \mathbb{R}^n \quad &, \qquad f'(x) = \lim_{h\to 0} \frac{f(x+h) - f(x)}{h} \in \mathbb{R}^n \ , \end{aligned}$$

wobei die erste Definition ein Spezialfall der zweiten ist, mit $n = 1$. Als nächstes würden wir gerne auch den Definitionsbereich von f mehr-dimensional wählen dürfen:

$$U \subset \mathbb{R}^m \text{ offen} \quad , \quad f : U \to \mathbb{R}^n \ .$$

Der Differenzenquotient $(f(x+h) - f(x))/h$ von oben macht jetzt aber keinen Sinn mehr, da $h \in \mathbb{R}^m$ sein müsste, aber durch Vektoren können wir nicht teilen.

■ Wir hatten eine weitere Sichtweise auf die Ableitung. Sie beschreibt die „beste lineare Näherung an eine Funktion“:

$$\begin{aligned} f : (a,b) \to \mathbb{R} \ , \quad f(x+h) &= f(x) + f'(x) \cdot h + r_x(h) \ , \quad \lim_{h\to 0} \frac{|r_x(h)|}{|h|} = 0 \ , \\ f : (a,b) \to \mathbb{R}^n \ , \quad f(x+h) &= f(x) + f'(x) \cdot h + r_x(h) \ , \quad \lim_{h\to 0} \frac{|r_x(h)|}{|h|} = 0 \ . \end{aligned}$$

Wir hatten gesehen, dass diese Definition von $f'(x)$ und die im Punkt oben äquivalent sind. In beiden Fällen ist die definierende Eigenschaft von $f'(x)$, dass

$$\lim_{h\to 0} \frac{\big| f(x+h) - f(x) - f'(x) \cdot h \big|}{|h|} = 0$$

Die Sichtweise „lineare Näherung“ verallgemeinern wir jetzt so:

Definition 9.2.1. Sei $U \subset \mathbb{R}^m$ offen, $f : U \to \mathbb{R}^n$, und sei $x \in U$. Existiert ein $A \in L(\mathbb{R}^m, \mathbb{R}^n)$, so dass

$$\lim_{h \to 0} \frac{|f(x+h) - f(x) - Ah|}{|h|} = 0 \ , \tag{$*$}$$

dann ist f *differenzierbar an der Stelle* x (oder *in* x) und man schreibt

$$df(x) = A \ .$$

Ist f in jedem $x \in U$ differenzierbar, so ist f *differenzierbar auf* U.

Wir nennen $df(x)$ das *Differential von* f *in* x (oder auch die *Ableitung von* f, oder *totale Ableitung*, *totales Differential*).

■ Die Notation df für die Ableitung wird auch in [Kap. 2.1 aus Königsberger II] verwendet. In [Def. 9.11 aus Rudin] wird statt dessen weiter f' benutzt.

Lemma 9.2.2. Gibt es $A_1, A_2 \in L(\mathbb{R}^m, \mathbb{R}^n)$, die $(*)$ erfüllen, so gilt $A_1 = A_2$.

Beweis.
Beh.: Es gilt $\lim_{h \to 0} |(A_1 - A_2)h|/|h| = 0$.

⌈ Wir berechnen

$$\begin{aligned} &\frac{|(A_1 - A_2)h|}{|h|} \\ =\ & \frac{|(f(x+h) - f(x) - A_1 h) - (f(x+h) - f(x) - A_2 h)|}{|h|} \\ \leq\ & \frac{|f(x+h) - f(x) - A_1 h|}{|h|} + \frac{|f(x+h) - f(x) - A_2 h|}{|h|} \end{aligned}$$

Die rechte Seite geht per Voraussetzung gegen 0 für $h \to 0$. Also auch die linke Seite. ⌋

Sei nun $x \in \mathbb{R}^m$ gegeben mit $x \neq 0$. Setzt $h = tx$ für ein $t \in \mathbb{R} \setminus \{0\}$. Mit $h = tx$ erhält man aus der obigen Behauptung, dass

$$0 = \lim_{t \to 0} \frac{|(A_1 - A_2)tx|}{|tx|} = \lim_{t \to 0} \frac{|(A_1 - A_2)x|}{|x|} = \frac{|(A_1 - A_2)x|}{|x|}$$

Also $(A_1 - A_2)x = 0$ für alle $x \in \mathbb{R}^m$ und somit $A_1 = A_2$. □

Bemerkung 9.2.3.

(1) Die Eindeutigkeitsaussage in Lemma 9.2.2 rechtfertigt die Notation $df(x)$. Falls $f : U \to \mathbb{R}^n$ auf ganz U differenzierbar ist, so erhalten wir eine Abbildung

$$df \;:\; U \longrightarrow L(\mathbb{R}^m, \mathbb{R}^n) \;.$$

In Worten:

> df ist eine Abbildung, die jedem Punkt x in U eine lineare Abbildung $df(x)$ zuordnet. Die lineare Abbildung $df(x)$ wiederum bildet einen Vektor v aus $\mathbb{R}^m$ auf einen Vektor $\big(df(x)\big)v$ in $\mathbb{R}^n$ ab.

(2) Der Bezug zu $f'(x)$ für eine differenzierbare Abbildung $f : (a,b) \to \mathbb{R}^n$ ist wie folgt. Für jedes $x \in (a,b)$ und $v \in \mathbb{R}$ gilt (Warum?) ?

$$\big(df(x)\big)v \;=\; f'(x) \cdot v \;\in\; \mathbb{R}^n \;.$$

Oder, äquivalent dazu: Die Matrixdarstellung von $df(x)$ bezüglich der Standardbasis von $\mathbb{R}$ und $\mathbb{R}^n$ ist

$$[df(x)] \;=\; \begin{pmatrix} (f'(x))_1 \\ \vdots \\ (f'(x))_n \end{pmatrix} \;\overset{(*)}{=}\; \begin{pmatrix} f_1'(x) \\ \vdots \\ f_n'(x) \end{pmatrix} \qquad ,$$

wobei wir in $(*)$ Satz 6.5.2 benutzt haben: Man kann die Ableitung $f'(x)$ komponentenweise ausrechnen.

(3) Sei U offen in $\mathbb{R}^m$ und $f : U \to \mathbb{R}^n$ differenzierbar in $x \in U$. Definiere die Funktion r_x mit Werten in $\mathbb{R}^n$ durch die Gleichung

$$f(x+h) \;=\; f(x) \;+\; \big(df(x)\big)h \;+\; r_x(h) \;.$$

(In welcher Menge darf h hier Werte annehmen?) Dann ist ?

$$h \;\mapsto\; f(x) \;+\; \big(df(x)\big)h \qquad (*)$$

eine *affine-lineare Abbildung* vom $\mathbb{R}^m$ in den $\mathbb{R}^n$ (d.h. eine lineare Abbildung, die um einen konstanten Vektor verschoben ist). Dies ist die eindeutige affine-lineare Abbildung $\mathbb{R}^m \to \mathbb{R}^n$ mit der Eigenschaft, dass

$$\lim_{h \to 0} \frac{|r_x(h)|}{|h|} \;=\; 0 \;.$$

(Warum gilt dieser Grenzwert? Warum ist die affine-lineare Abbildung $(*)$ eindeutig?) In diesem Sinne ist $(*)$ die beste (affin-)lineare Näherung an die Funktion $h \mapsto f(x+h)$. ?

(4) Sei $A : \mathbb{R}^m \to \mathbb{R}^n$ eine lineare Abbildung. Dann gilt für alle $x \in \mathbb{R}^m$, dass $df(x) = A$ (Warum?). Also ist in diesem Fall die Abbildung $df : \mathbb{R}^m \to L(\mathbb{R}^m, \mathbb{R}^n)$ konstant und nimmt überall den Wert A an. ?

Interessantere Beispiele schauen wir uns im nächsten Kapitel an, wenn wir partielle Ableitungen eingeführt haben.

(5) Seien U, f, x wie in Teil (3). Wie im Fall $m = 1$ gilt auch für allgemeine $m \in \mathbb{N}$: Ist f differenzierbar in x, so ist f auch stetig in x. (Warum gilt das?) ?

Satz 9.2.4. (Kettenregel)
Seien $U \subset \mathbb{R}^m$, $V \subset \mathbb{R}^n$ und

$$f : U \to \mathbb{R}^n \quad , \quad g : V \to \mathbb{R}^k$$

gegeben mit U, V offen und $f(U) \subset V$. Die Funktion f sei in x differenzierbar, und g sei in $f(x)$ differenzierbar. Dann ist $p := g \circ f : U \to \mathbb{R}^k$ in x differenzierbar, und es gilt

$$dp(x) \;=\; dg(f(x)) \circ df(x) \; .$$

Beweis. Wir setzen $y = f(x)$ und definieren Funktionen r_x, s_y als

$$\begin{aligned} f(x+h) &= f(x) \;+\; \big(df(x)\big)h \;+\; r_x(h) \; , \\ g(y+u) &= g(y) \;+\; \big(dg(y)\big)u \;+\; s_y(u) \; . \end{aligned}$$

Für $h, u \neq 0$ definieren wir weiter zwei Funktionen ρ, σ mit Werten in $\mathbb{R}_{\geq 0}$ durch die Gleichungen

$$|r_x(h)| \;=\; \rho(h)\,|h| \quad , \quad |s_y(u)| \;=\; \sigma(u)\,|u| \; .$$

Per Annahme gilt

$$\lim_{h\to 0} \rho(h) \;=\; 0 \quad , \quad \lim_{u \to 0} \sigma(u) \;=\; 0 \; .$$

(Warum?). Wenn wir weiterhin definieren, dass $\rho(0) = 0$ und $\sigma(0) = 0$, dann sind ρ und σ also in Null stetig. ?

Für $p = g \circ f$ setzen wir genauso

$$p(x+h) \;=\; p(x) \;+\; \big(dg(y) \circ df(x)\big)h \;+\; t_x(h) \; .$$

Die Aussage des Satzes ergibt sich aus der folgenden Behauptung.

Beh.: $\lim_{h\to 0} \frac{|t_x(h)|}{|h|} = 0$.

⌈ Wir drücken $p(x+h)$ durch f und g aus:

$$\begin{aligned} p(x+h) &= g(f(x+h)) \\ &= g\big(\, f(x) \;+\; df(x)h \;+\; r_x(h)\,\big) \\ &= g(f(x)) \;+\; dg(y)\big(df(x)h + r_x(h)\big) \;+\; s_y\big(df(x)h + r_x(h)\big) \end{aligned}$$

Daraus können wir ablesen, dass

$$t_x(h) \;=\; dg(y)\big(r_x(h)\big) \;+\; s_y\big(df(x)h + r_x(h)\big)\ .$$

Also gilt für den Betrag die Abschätzung

$$\begin{aligned} |t_x(h)| &\le \|dg(y)\|\,|r_x(h)| \;+\; \big|s_y\big(df(x)h + r_x(h)\big)\big| \\ &= \|dg(y)\|\,\rho(h)|h| \;+\; \sigma\big(df(x)h + r_x(h)\big)\,\underbrace{\big|df(x)h + r_x(h)\big|}_{\le \|df(x)\|\,|h| + \rho(h)\,|h|}\ . \end{aligned}$$

Nach Teilen durch $|h|$ erhalten wir

$$\frac{|t_x(h)|}{|h|} \;\le\; \|dg(y)\|\,\rho(h) \;+\; \sigma\big(df(x)h + r_x(h)\big)\,\big(\|df(x)\| + \rho(h)\big)\ .$$

Da $\rho(h) \to 0$ und $\sigma\big(df(x)h + r_x(h)\big) \to 0$ für $h \to 0$ (Warum?), ⓘ?
folgt die Behauptung. ⌋

□

9.3 Partielle Ableitungen

Definition 9.3.1. Sei $U \subset \mathbb{R}^m$ offen, $f : U \to \mathbb{R}^n$, $p \in U$ und $v \in \mathbb{R}^m$, $v \neq 0$. Falls der Grenzwert

$$\lim_{t\to 0} \frac{f(p+tv) - f(p)}{t}$$

existiert, so sagen wir, die *Ableitung von f in Richtung v* existiert, oder *Richtungsableitung von f bezüglich v* existiert. Den Grenzwert bezeichnen wir dann mit $(D_v f)(p)$.

In der Definition haben wir nicht genau gesagt, für welche t der Ausdruck im Grenzwert überhaupt Sinn macht. Es gibt aber für jedes p ein $\delta > 0$, so dass $p + tv \in U$ für alle $t \in (-\delta, \delta)$. (Warum?) ?

Mit der Notation von Definition 9.3.1 haben wir folgende Aussage.

Lemma 9.3.2. Wenn f an der Stelle p differenzierbar ist, dann existieren alle Richtungsableitungen, und es gilt

$$(D_v f)(p) \;=\; df(p)v \;\in\; \mathbb{R}^n\ .$$

Beweis. Wir benutzen Bemerkung 9.2.3 (3):

$$\frac{f(p+tv)-f(p)}{t} \;=\; \frac{df(p)tv + r_p(tv)}{t} \;=\; df(p)v \;+\; \frac{r_p(tv)}{t}$$

Die Behauptung folgt nun aus $\lim_{t\to 0} r_p(tv)/t \;=\; 0$ (Warum stimmt dieser Grenzwert?) ? □

Bemerkung 9.3.3.

(1) Die Richtungsableitung erfüllt, für $\lambda \in \mathbb{R}$, $\lambda \neq 0$,

$$(D_{\lambda v}f)(p) \;=\; \lambda\,(D_v f)(p)\;.$$

(Warum? (In Ihrer Begründung dürfen Sie nicht annehmen, dass f in p differenzierbar ist, nur die Richtungsableitung in Richtung v muss es geben.)) ?

Deswegen ist es bei Richtungsableitungen üblich, v als Einheitsvektor zu wählen, $|v| = 1$.

(2) Die Umkehrung von Lemma 9.3.2 gilt nicht. Betrachte hierzu die Funktion $f : \mathbb{R}^2 \to \mathbb{R}$,

$$f(x,y) = \begin{cases} \frac{x^2y}{x^2+y^2} & ;(x,y) \neq (0,0) \\ 0 & ;(x,y) = (0,0) \end{cases}\qquad .$$

Beh.: f ist stetig auf $\mathbb{R}^2$.

⌈ Wir müssen nur Stetigkeit in 0 prüfen. Für $v \in \mathbb{R}^2$, $v \neq 0$ gilt $|f(v)| = |v_1^2 v_2|/|v|^2 \leq |v|^3/|v|^2 = |v|$. Also $\lim_{v\to 0} f(v) = 0$. ⌋

Sei $v \in \mathbb{R}^2$ ein Einheitsvektor, $|v| = 1$. Wir berechnen die Richtungsableitung von f in Richtung v an der Stelle $p = 0$:

$$(D_v f)(0) \;=\; \lim_{t\to 0}\frac{f(tv)}{t} \;=\; \lim_{t\to 0}\frac{1}{t}\cdot\frac{t^3v_1^2v_2}{t^2(v_1^2+v_2^2)} \;=\; v_1^2\,v_2$$

Insbesondere existieren alle Richtungsableitungen.

Beh.: f ist in 0 nicht differenzierbar.

⌈ Angenommen, es gäbe die Ableitung $df(0)$. Nach Lemma 9.3.2 gilt dann $(D_v f)(0) = df(0)v$. Wenn wir für v die Basisvektoren $e_1 = (1,0)$ und $e_2 = (0,1)$ wählen, erhalten wir

$$(D_{e_1}f)(0) \;=\; 0 \quad , \quad (D_{e_2}f)(0) \;=\; 0\;.$$

Die lineare Abblidung $df(0)$ liegt durch ihre Wirkung auf eine Basis eindeutig fest, und es folgt also $df(0) = 0$. Aber für $v = (1,1)$ erhalten wir $(D_v f)(0) = 1$, im Widerspruch zu $df(0) = 0$. ⌋

(3) Wie vorher sei $U \subset \mathbb{R}^m$ offen und $f : U \to \mathbb{R}^n$. Es seien $f_1, \dots, f_n : U \to \mathbb{R}$ die Komponentenfunktionen von f, also $f(p) = (f_1(p), \dots, f_n(p))$. Nach Satz 6.5.2 gilt (Wieso?) ?

$$(D_v f)(p) = \begin{pmatrix} (D_v f_1)(p) \\ \vdots \\ (D_v f_n)(p) \end{pmatrix} ,$$

wobei die Richtungsableitung auf der linken Seite genau dann existiert, wenn alle Richtungsableitungen auf der rechten Seite existieren.

Wenn v einer der Standard-Basisvektoren $e_1, \dots, e_m$ von $\mathbb{R}^m$ ist, schreiben wir auch kurz $(D_j f)(p)$ statt $(D_{e_j} f)(p)$. Für die Richtungsableitung bezüglich e_i der Komponentenfunktion f_i schreiben wir auch

$$\frac{\partial f_i}{\partial x_j}(p) \overset{\text{Def.}}{=} (D_j f_i)(p) \overset{\text{Def.}}{=} (D_{e_j} f_i)(p)$$

und nennen $\frac{\partial f_i}{\partial x_j}(p)$ die *partielle Ableitung von f_i nach x_j an der Stelle p.*

(4) Für eine Funktion $g : (a, b) \to \mathbb{R}$, $t \mapsto g(t)$, benutzt man statt der Notation $g'(p)$ für die Ableitung von g an der Stelle t auch

$$\frac{dg}{dt}(t) \overset{\text{Def.}}{=} g'(t) ,$$

oder auch z.B. $\frac{d}{dt} g(t)$.

Wir schreiben in diesem Fall die Definition der partiellen Ableitung noch einmal aus: Setze $g(t) = f_i(p_1, \dots, p_j + t, \dots, p_m)$. Dann

$$\frac{\partial f_i}{\partial x_j}(p) = \frac{dg}{dt}(0) .$$

In Worten: „Man stellt sich alle Variablen von f_i außer x_j als feste Werte vor. Dann hat man noch eine Funktion in einer reellen Variablen mit Werten in $\mathbb{R}$. Diese leitet man wie gewohnt ab.“

Satz 9.3.4. Sei $U \subset \mathbb{R}^m$ offen und $f : U \to \mathbb{R}^n$. Ist f in $p \in U$ differenzierbar, so existieren alle partiellen Ableitungen, und $df(p)$ hat die Matrixdarstellung

$$\big[df(p)\big] = \begin{pmatrix} (D_1 f_1)(p) & \cdots & (D_m f_1)(p) \\ \vdots & & \vdots \\ (D_1 f_n)(p) & \cdots & (D_m f_n)(p) \end{pmatrix} = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(p) & \cdots & \frac{\partial f_1}{\partial x_m}(p) \\ \vdots & & \vdots \\ \frac{\partial f_n}{\partial x_1}(p) & \cdots & \frac{\partial f_n}{\partial x_m}(p) \end{pmatrix} .$$

Beweis. Klar nach Lemma 9.3.2 und Bemerkung 9.3.3 (3). (Details? *Hinweis:* Um die j-te Spalte der Matrix zu bekommen, müssen Sie $df(p)e_j$ in der Standardbasis ausdrücken.) ? □

■ Wir haben gesehen, dass man aus der Existenz der partiellen Ableitungen in p nicht folgern kann, dass $df(p)$ existiert. Im Beispiel von Bemerkung 9.3.3 (2) ist $\frac{\partial f}{\partial x_1}(0) = 0 = \frac{\partial f}{\partial x_2}(0)$, also ist die Matrix auf der rechten Seite der Gleichung in Satz 9.3.4 die Nullmatrix, aber f ist in 0 nicht differenzierbar, und die linke Seite der Gleichung existiert erst gar nicht.

Beispiel 9.3.5. („Höhenprofil einer Bergwanderung (im $\mathbb{R}^m$)")
Sei $U \subset \mathbb{R}^m$ offen ($m = 2$ für einen wirklichen Berg) und $h : U \to \mathbb{R}$, differenzierbar auf U. Wir denken uns $h(p)$ ist die „Höhe vom Berg am Punkt $p \in U$. Für $a < b$ sei $\gamma : (a,b) \to U$ differenzierbar. Wir denken uns γ als den Weg und $\gamma(t)$ als den Ort, wo wir zum Zeitpunkt $t \in (a,b)$ gerade sind. Dann ist

$$h \circ \gamma \ : \ (a,b) \to \mathbb{R}$$

das Höhenprofil unserer Wanderung. Der Höhenunterschied pro Zeit, den wir zurücklegen ist

$$\tfrac{d}{dt}\big(h \circ \gamma\big) \overset{\text{Satz 9.2.4}}{=} dh(\gamma(t))\,\gamma'(t) \quad , \quad \text{wobei } dh(\gamma(t)) \in L(\mathbb{R}^m, \mathbb{R}) \ .$$

Definition 9.3.6. Sei $U \subset \mathbb{R}^m$ offen und $f : U \to \mathbb{R}$ differenzierbar. Dann definieren wir

$$(\nabla f)(p) \overset{\text{Def.}}{=} \begin{pmatrix} (D_1 f)(p) \\ \vdots \\ (D_m f)(p) \end{pmatrix} \quad .$$

Die Funktion $\nabla f : U \to \mathbb{R}^m$ heißt der *Gradient von f*.

Beispiel 9.3.7. Mit der Notation aus Beispiel 9.3.5: Es gilt, mit $p = \gamma(t)$,

$$dh(p)\,\gamma'(t) \ = \ \big((\nabla h)(p)\,,\, \gamma'(t) \big) \ ,$$

wobei die Klammer auf der rechten Seite für das Skalarprodukt im $\mathbb{R}^m$ steht (Warum?). In Worten: „Das Skalarprodukt von Gradient und Richtungsvektor gibt die Höhenänderung pro Zeit." In der Tat gibt der Gradient die Richtung des steilsten Anstieges (→Übung). ?

Definition 9.3.8. Eine Teilmenge $E \subset \mathbb{R}^m$ heißt konvex, falls sie für je zwei Punkte auch deren Verbindungsgerade enthält:

$$\forall p, q \in E \ \forall t \in [0,1] \ : \ p + t(q-p) \ \in \ E \ .$$

Beispiel 9.3.9. Die offene Einheitskreisscheibe $\{p \in \mathbb{R}^2 \mid |v| < 1\}$ ist konvex, der Einheitskreis $\{p \in \mathbb{R}^2 \mid |v| = 1\}$ ist nicht konvex. (Beweis? Ist die Vereinigung von zwei konvexen Mengen konvex? Ihr Durchschnitt? Wie sehen die konvexen Teilmengen von $\mathbb{R}$ aus?). ?

Satz 9.3.10. (Schrankensatz)
Sei $U \subset \mathbb{R}^m$ offen und konvex, und sei $f : U \to \mathbb{R}^n$ differenzierbar auf U. Angenommen, es gibt $M \in \mathbb{R}$, so dass $\|df(p)\|_{\mathrm{op}} \leq M$ für alle $p \in U$. Dann gilt für alle $p, q \in U$:

$$\big| f(p) - f(q) \big| \;\leq\; M \cdot |p - q| \;.$$

(Können Sie ein Gegenbeispiel angeben, wenn man die Bedingung „konvex“ weglässt?) ?

■ Satz 6.5.6 gibt einen Beweis des Satzes im Spezialfall $m = 1$. (Wie?) ?

Beweis von Satz 9.3.10. Wir führen die Aussage auf Satz 6.5.6 zurück. Sei $\gamma : [0,1] \to U$, $t \mapsto p + t(q-p)$. Dann $\gamma([0,1]) \subset U$ per Voraussetzung. Nach der Kettenregel (Satz 9.2.4) ist

$$g \overset{\text{Def.}}{=} f \circ \gamma \;:\; [0,1] \longrightarrow \mathbb{R}^n$$

differenzierbar auf $(0,1)$ mit Ableitung

$$g'(t) \;=\; df(\gamma(t))\,\gamma'(t) \;=\; df(\gamma(t))(q-p) \;.$$

Da $\|df\|$ per Annahme beschränkt ist, können wir abschätzen

$$\big|g'(t)\big| \;\leq\; \|df(\gamma(t))\|\,|q-p| \;\leq\; M \cdot |q-p| \;.$$

Nach Satz 6.5.6 gilt nun $|g(1) - g(0)| \leq M \cdot |q - p|$. Dies ist genau die Behauptung des Satzes. □

Definition 9.3.11. Sei $U \subset \mathbb{R}^m$ offen und $f : U \to \mathbb{R}^n$. Wie sagen, *f ist stetig differenzierbar (auf U)*, falls f auf U differenzierbar ist, und falls die Abbildung

$$df \;:\; U \longrightarrow L(\mathbb{R}^m, \mathbb{R}^n)$$

stetig ist. Wir sagen auch, *f ist eine $\mathcal{C}^1$-Abbildung*, und wir schreiben $\mathcal{C}^1(U, \mathbb{R}^n)$ (oder kurz $\mathcal{C}^1(U)$) für die Menge der stetig differenzierbaren Funktionen $U \to \mathbb{R}^n$.

■ Mit ε-δ ausgeschrieben, heißt die Stetigkeitsbedingung:

$$\forall \varepsilon > 0, p \in U \;\exists \delta > 0 \;\forall x \in U \;:\; |x - p| < \delta \;\Rightarrow\; \|df(x) - df(p)\|_{\mathrm{op}} < \varepsilon \;.$$

Satz 9.3.12. Sei $U \subset \mathbb{R}^m$ offen und $f : U \to \mathbb{R}^n$. Es gilt: f ist genau dann stetig differenzierbar auf U, falls alle partiellen Ableitungen $D_j f_i$ existieren und auf U stetig sind.

■ Erinnerung: Falls f in p differenzierbar ist, so hatten wir in Satz 9.3.4 folgende Matrixdarstellung für $df(p)$:

$$\big[df(p)\big] \;=\; \begin{pmatrix} (D_1 f_1)(p) & \cdots & (D_m f_1)(p) \\ \vdots & & \vdots \\ (D_1 f_n)(p) & \cdots & (D_m f_n)(p) \end{pmatrix} \quad .$$

Beweis von Satz 9.3.12.

$\Rightarrow$: Nach Lemma 9.3.2 existieren die $D_j f_i(p)$ für alle $p \in U$. Nach Satz 9.3.4 geben die $D_j f_i(p)$ die Matrixdarstellung von $df(p)$. Nach Satz 9.1.10 ist df stetig.

$\Leftarrow$: Es genügt, die Aussage für $n = 1$ zu zeigen ($\to$Übung). Sei also $f : U \to \mathbb{R}$ so, dass $D_j f : U \to \mathbb{R}$ existiert ($j = 1, \ldots, m$) und stetig ist.

Sei $x \in U$. Zunächst schreiben wir die Differenz $f(x+h) - f(x)$ so um, dass wir partielle Ableitungen verwenden können. Die Komponenten von h seien $h = h_1 e_1 + \cdots + h_m e_m$ (mit $e_1, \ldots, e_m$ die Standardbasis von $\mathbb{R}^m$). Es gilt

$$\begin{aligned} &f(x + h_1 e_1 + \cdots + h_m e_m) - f(x) \\ &\quad = \; f(x + h_1 e_1 + \cdots + h_m e_m) \;-\; f(x + h_1 e_1 + \cdots + h_{m-1} e_{m-1}) \\ &\quad\;\; + \; f(x + h_1 e_1 + \cdots + h_{m-1} e_{m-1}) \;-\; f(x + h_1 e_1 + \cdots + h_{m-2} e_{m-2}) \\ &\quad\;\; + \; \cdots \\ &\quad\;\; + \; f(x + h_1 e_1) \;-\; f(x) \; . \end{aligned} \qquad (*)$$

Auf die Differenz in jeder Zeile können wir nun den Mittelwertsatz (Satz 6.2.4) anwenden. Es gibt $\xi_i \in (0, h_i)$, so dass

$$\begin{aligned} f(x + h_1 e_1) \;-\; f(x) \;&=\; (D_1 f)(x + \xi_1 e_1)\, h_1 \\ f(x + h_1 e_1 + h_2 e_2) \;-\; f(x + h_1 e_1) \;&=\; (D_2 f)(x + h_1 e_1 + \xi_2 e_2)\, h_2 \\ &\cdots \end{aligned}$$

Sei $\varepsilon > 0$ beliebig. Es gilt $\delta > 0$, so dass für alle $u \in \mathbb{R}^m$ mit $|u| < \delta$ gilt:

a) $x + u \in U$ (möglich, da U offen ist)

b) Für $j = 1, \dots, m$ ist $|D_j f(x+u) - D_j f(x)| < \frac{\varepsilon}{m}$ (möglich, da $D_j f$ in x stetig ist)

Aus $(*)$ bekommen wir dann für $|h| < \delta$ insgesamt die Abschätzung

$$\begin{aligned}
&\Big| f(x+h) - f(x) \; - \; \sum_{j=1}^{m} (D_j f(x)) h_j \Big| \\
&\quad = \Big| \Big((D_m f)(x + h_1 e_1 + \dots + h_{m-1} e_{m-1} + \xi_m e_m) \; - \; (D_m f)(x) \Big) h_m \\
&\qquad + \cdots \\
&\qquad + \Big((D_1 f)(x + \xi_1 e_1) - (D_1 f)(x) \Big) h_1 \Big| \\
&\quad \leq \frac{\varepsilon}{m} \Big(|h_m| + |h_{m-1}| + \dots + |h_1| \Big) \; \leq \; \varepsilon |h| \; .
\end{aligned}$$

Anders ausgedrückt: Für jedes $\varepsilon > 0$ finden wir ein $\delta > 0$, so dass für alle $0 < |h| < \delta$ gilt

$$\frac{\big| f(x+h) - f(x) \; - \; \sum_{j=1}^{m} (D_j f(x)) h_j \big|}{|h|} \; \leq \; \varepsilon \; .$$

Dies zeigt, dass f in x differenzierbar ist (Warum?). Dass $df : U \to L(\mathbb{R}^m, \mathbb{R})$ stetig ist, folgt wieder aus Satz 9.1.10 und Satz 9.3.4. ?

□

Beispiel 9.3.13.

(1) Betrachte $f : \mathbb{R}^2 \to \mathbb{R}^2$, $f(x, y) = (xy, x^2 + y^2)$. Dann ist f auf $\mathbb{R}^2$ differenzierbar, und es gilt

$$[df(x, y)] \; = \; \begin{pmatrix} \frac{\partial f_1}{\partial x}(x, y) & \frac{\partial f_1}{\partial y}(x, y) \\ \frac{\partial f_2}{\partial x}(x, y) & \frac{\partial f_2}{\partial y}(x, y) \end{pmatrix} \; = \; \begin{pmatrix} y & x \\ 2x & 2y \end{pmatrix} \quad ,$$

da alle partiellen Ableitungen existieren und auf ganz $\mathbb{R}^2$ stetig sind.

(2) Wenn man den Zusatz „... und auf U stetig sind“ in Satz 9.3.12 weglässt, dann gilt die Aussage nicht mehr. Betrachte dazu die Funktion $f : \mathbb{R}^2 \to \mathbb{R}$ aus Bemerkung 9.3.3 (2),

$$f(x, y) = \begin{cases} \frac{x^2 y}{x^2 + y^2} & ; (x, y) \neq (0, 0) \\ 0 & ; (x, y) = (0, 0) \end{cases} \qquad .$$

Diese Funktion war in 0 nicht differenzierbar. Wir prüfen nun, dass dennoch alle partiellen Ableitungen auf ganz $\mathbb{R}^2$ existieren, und überall stetig sind, außer in 0.

In Bemerkung 9.3.3 hatten wir bereits gesehen, dass

$$\frac{\partial f}{\partial x}(0,0) \;=\; 0 \;=\; \frac{\partial f}{\partial y}(0,0)\ .$$

Für $(x,y) \neq (0,0)$ berechnet man

$$\frac{\partial f}{\partial x}(x,y) \;=\; \frac{2xy^3}{(x^2+y^2)^2} \quad , \quad \frac{\partial f}{\partial y}(x,y) \;=\; \frac{x^2(x^2-y^2)}{(x^2+y^2)^2}\ .$$

Beide partiellen Ableitungen sind in 0 nicht stetig. (Warum?) ?

9.4 Umkehrabbildung und implizite Funktionen

Kurzer Exkurs: Kontrahierende Abbildungen

Definition 9.4.1. Sei (X,d) ein metrischer Raum und $\varphi : X \to X$. Wir nennen φ eine *Kontraktion* oder *kontrahierende Abbildung*, falls es ein $\lambda \in \mathbb{R}$ mit $0 \leq \lambda < 1$ gibt, so dass

$$\forall x,y \in X \;:\; d(\varphi(x),\varphi(y)) \;\leq\; \lambda \cdot d(x,y)\ .$$

■ Eine Kontraktion ist automatisch stetig (und sogar global Lipschitz stetig). Eine lineare Abbildung $A : V \to V$ für einen normierten $\mathbb{R}$-Vektorraum V ist genau dann eine Kontraktion, wenn $\|A\|_{\mathrm{op}} < 1$. (Warum gilt das beides?). ?

■ Für eine Menge E und eine Abbildung $f : E \to E$ nennen wir ein Element $u \in E$ mit $f(u) = u$ einen *Fixpunkt von* f. Z.B. hat jeder lineare Endomorphismus den Vektor 0 als Fixpunkt.

Satz 9.4.2. (Banachscher Fixpunktsatz)
Sei (X,d) ein vollständiger metrischer Raum und $\varphi : X \to X$ eine Kontraktion. Dann hat φ genau einen Fixpunkt (d.h. es gibt ein eindeutiges $p \in X$ mit $\varphi(p) = p$).

Beweis. Existenz: Sei $x_0 \in X$ beliebig. Setze $x_{n+1} = \varphi(x_n)$.

Beh.: Die Folge $(x_n)_{n\in\mathbb{N}_0}$ konvergiert in X, und $p = \lim_{n\to\infty} x_n$ ist ein Fixpunkt von φ.

⌈ Es gilt $d(x_{n+1}, x_n) = d(\varphi(x_n), \varphi(x_{n-1})) \leq \lambda\, d(x_n, x_{n-1})$. Damit sieht man induktiv, dass $d(x_{n+1}, x_n) \leq \lambda^n\, d(x_1, x_0)$. Für $m > n$ können wir abschätzen (Wie?) ?

$$\begin{aligned} d(x_m, x_n) &\leq \sum_{k=n}^{m-1} d(x_{k+1}, x_k) \leq \sum_{k=n}^{m-1} \lambda^k\, d(x_1, x_0) \\ &\leq \lambda^n\, d(x_1, x_0) \sum_{j=0}^{\infty} \lambda^j = \frac{\lambda^n}{1-\lambda}\, d(x_1, x_0)\ . \end{aligned}$$

Damit ist $(x_n)_{n\in\mathbb{N}_0}$ eine Cauchy-Folge (Warum? *Hinweis:* $\lambda^n \to 0$ für $n \to \infty$.). Da X vollständig ist, konvergiert die Folge. Setze $p = \lim_{n\to\infty} x_n$. Da φ stetig ist, gilt $\varphi(p) = \lim_{n\to\infty} \varphi(x_n) = \lim_{n\to\infty} x_{n+1} = p$. ⌋ ?

Eindeutigkeit: Angenommen $p, q \in X$ erfüllen $\varphi(p) = p$ und $\varphi(q) = q$. Dann

$$d(p, q) = d(\varphi(p), \varphi(q)) \leq \lambda\, d(p, q)\ .$$

Da $\lambda < 1$ folgt $d(p, q) = 0$, also $p = q$. □

Bemerkung 9.4.3.

(1) Sei $U \subset \mathbb{R}^m$ offen und $\varphi : U \to \mathbb{R}^m$ auf U differenzierbar. Angenommen, es gibt $0 \leq M < 1$, so dass $\|d\varphi(p)\| \leq M$ für alle $p \in U$. Nach dem Schrankensatz (Satz 9.3.10) ist φ auf jeder konvexen Teilmenge von U eine Kontraktion. Gibt es ferner eine konvexe, abgeschlossene (in $\mathbb{R}^m$) Teilmenge $A \subset U$ mit $\varphi(A) \subset A$, so hat φ auf A genau einen Fixpunkt.

(2) Lässt man in Satz 9.4.2 die Bedingung „vollständig" weg, so ist die Aussage falsch. (Können Sie ein Gegenbeispiel angeben?) ?

(3) Die Abbildung $\varphi : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$, $\varphi(x) = \sqrt{x^2+1}$ erfüllt

$$\varphi'(x) = \frac{x}{\sqrt{x^2+1}} < 1\ .$$

Nach dem Mittelwertsatz gilt also $|\varphi(x) - \varphi(y)| < |x - y|$ für alle $x, y \in \mathbb{R}_{\geq 0}$ mit $x \neq y$. Aber φ hat keinen Fixpunkt (Warum?). Das ist der Grund, warum man bei der Definition von Kontraktion nicht einfach sagt $d(\varphi(x), \varphi(y)) < d(x, y)$, sondern die Konstante $\lambda < 1$ einführt. ?

(4) Auch Abbildungen, die keine Kontraktionen sind, können einen eindeutigen Fixpunkt haben (Beispiel?). ?

Umkehrabbildung und implizite Funktionen

Satz 9.4.4. (Satz von der lokalen Umkehrbarkeit)
Sei $U \subset \mathbb{R}^n$ offen, $f : U \to \mathbb{R}^n$ eine $\mathcal{C}^1$-Abbildung, und sei $a \in U$, sodass $df(a) \in GL_n(\mathbb{R})$. Dann gibt es offene Mengen $A, B \subset \mathbb{R}^n$ mit $a \in A \subset U$, so dass

(i) $f|_A : A \to B$ ist eine Bijektion, und

(ii) die Umkehrabbildung $g := f^{-1} : B \to A$ ist stetig differenzierbar auf B, und für $b = f(a)$ gilt

$$dg(b) = \big(df(a)\big)^{-1} .$$

Beweis.
Beh.: Es ist genug, den Satz unter der Annahme zu zeigen, dass $a = 0$, $f(0) = 0$ und $df(0) = \text{id}$.

⌈ Sei $b = f(a)$ und $A = df(a)$. Betrachte $F(x) = A^{-1}\big(f(x + a) - b\big)$. Dann hat F die gewünschten Eigenschaften (Warum). ?
Angenommen, der Satz gilt für F mit Umkehrabbildung G. Da $f(x) = AF(x-a)+b$ gilt (i) auch für f. Die Umkehrabbildung ist $g(y) = G(A^{-1}(y-b))+a$ und erfüllt $dg(y) = dG(A^{-1}(y-b))A^{-1}$ (Warum?). Damit ist dg ebenfalls stetig und $dg(b) = A^{-1}$. ⌋ ?

In folgenden gelte also $a = 0$, $f(0) = 0$, $df(0) = id$.

(i) Wir zeigen die Existenz des Inversen mit dem Fixpunktsatz. Sei $r > 0$, sodass der abgeschlossene Ball $\overline{B}_{2r}$ um 0 in U liegt und $\|df(x)-id\| \leq \frac{1}{2}$ für alle $x \in \overline{B}_{2r}$. Letzteres ist möglich, da df stetig ist.

Sei $y \in \mathbb{R}^n$ mit $|y| < r$ gegeben. Betrachte $\varphi : U \to \mathbb{R}^n$,

$$\varphi(x) = y + x - f(x) .$$

Dann ist $f(x) = y$ äquivalent zu $\varphi(x) = x$, und anstatt Urbilder von y unter f zu suchen, können wir Fixpunkte von φ suchen.

Beh.: φ ist eine Kontraktion auf $\overline{B}_{2r}$.

⌈ Dies folgt aus dem Schrankensatz (Satz 9.3.10): $d\varphi(x) = \text{id} - df(x)$, und somit $\|d\varphi(x)\| \leq \frac{1}{2}$. Also gilt für alle $x, x' \in \overline{B}_{2r}$, dass $|\varphi(x) - \varphi(x')| \leq \frac{1}{2}|x - x'|$.
Ferner $|\varphi(x)| = |\varphi(x) - \varphi(0) + \varphi(0)| \leq |\varphi(x) - \varphi(0)| + |y| \leq \frac{1}{2}|x - 0| + |y| < r + r$. Also sogar $\varphi(\overline{B}_{2r}) \subset B_{2r}$, dem offenen Ball um 0 von Radius 2r. ⌋

Da $\overline{B}_{2r}$ vollständig ist, gibt der Banachsche Fixpunktsatz (Satz 9.4.2) ein eindeutiges x_0 in $\overline{B}_{2r}$ mit $f(x_0) = y$. Da $x_0 = \varphi(x_0)$, liegt x_0 sogar in B_{2r}. Wähle

$$B = B_r \quad , \quad A = f^{-1}(B_r) \cap B_{2r} \ .$$

Insgesamt haben wir gezeigt, dass $f|_A : A \to B$ eine Bijektion ist. (Warum kann man nicht direkt $A = f^{-1}(B_r)$ setzen und muss nochmal mit B_{2r} schneiden? Warum sind A und B offen?) ?

(ii) Sei $g : B \to A$ die Umkehrfunktion von $f : A \to B$. Sei $y \in B$ beliebig und sei $x = g(y) \in A$.

Beh.: $df(x) \in GL_n(\mathbb{R})$ für alle $x \in A$.

⌈ Wir benutzen Satz 9.1.7: Für $A \in GL_n(\mathbb{R})$ und $B \in L(\mathbb{R}^n)$ mit $\|A - B\| < 1/\|A^{-1}\|$ ist auch B invertierbar. Setze nun $A = df(0) = id$ and $B = df(x)$. Wegen $A \subset \overline{B}_{2r}$ ist dann $\|id - df(x)\| \leq \frac{1}{2} < 1 = 1/\|id^{-1}\|$. Somit ist $df(x) \in GL_n(\mathbb{R})$. ⌋

Beh.: Für alle $y \in B$ ist g differenzierbar in y mit $dg(y) = df(x)^{-1}$.

⌈ Sei $k \in \mathbb{R}^n$, so dass $y + k \in B$. Dann gibt es ein eindeutiges $h \in \mathbb{R}^n$, so dass $x + h = g(y + k)$. Mit φ wie oben gilt

$$\begin{aligned} \varphi(x+h) - \varphi(x) &= y + x + h - f(x+h) - y - x + f(x) \\ &= -f(x+h) + f(x) + h \ . \end{aligned}$$

Andererseits ist nach der Kontraktionseigenschaft auch

$$\varphi(x+h) - \varphi(x) \ \leq \ \tfrac{1}{2}\,|x + h - x| \ = \ \tfrac{1}{2}\,|h| \ .$$

Zusammen erhält man

$$|h| - |f(x+h) - f(x)| \ \leq \ |h - (f(x+h) - f(x))| \ \leq \ \tfrac{1}{2}\,|h| \ ,$$

also

$$\tfrac{1}{2}\,|h| \ \leq \ |f(x+h) - f(x)| \ = \ |k| \ .$$

(Warum gilt die letzte Gleichheit?) Jetzt haben wir alles zusammen, um die Bedingung für die Ableitung nachzurechnen. ?
Zunächst formen wir um:

$$\begin{aligned} &g(y+k) - g(y) - df(x)^{-1}k \\ &\quad = \ x + h - x - df(x)^{-1}k \\ &\quad = \ df(x)^{-1}\big(df(x)h - k\big) \\ &\quad = \ -df(x)^{-1}\big(f(x+h) - f(x) - df(x)h\big) \ . \end{aligned}$$

Dann

$$\begin{aligned}
&\frac{|g(y+k) - g(y) - df(x)^{-1}k|}{|k|} \\
&\qquad = \frac{\big| - df(x)^{-1}\big(f(x+h) - f(x) - df(x)h\big)\big|}{|k|} \\
&\qquad \leq 2\,\frac{\|df(x)^{-1}\|\,\big|f(x+h) - f(x) - df(x)h\big|}{|h|} \xrightarrow{k\to 0} 0
\end{aligned}$$

(Warum durften wir in der letzten Abschätzung $|k|$ durch $|h|$ ersetzen? Und warum sagen wir $k \to 0$ wenn wir doch eigentlich $h \to 0$ benutzen?) ⌋ ?

Beh.: Die Abbildung $dg : B \to L(\mathbb{R}^n)$ ist stetig.

⌈ Wir wissen bereits, dass $dg(y) = df(g(y))^{-1}$, und g ist stetig (da differenzierbar) und df ist stetig (per Annahme). ⌋

□

Bemerkung 9.4.5.

(1) Im Fall $n = 1$ folgt die Aussage des Satzes aus Satz 6.1.9. (Dort wird aber streng monoton steigend/fallend gefordert, hier aber nicht. Wieso können wir Satz 6.1.9 trotzdem anwenden?) ?

(2) Die Folgerung des Satzes über lokale Umkehrbarkeit kann man auch wie folgt formulieren: Sei $f : A \to B$ stetig differenzierbar und bijektiv, mit Umkehrfunktion $g : B \to A$. Die Komponentenfunktionen $f_j : A \to \mathbb{R}$, $j = 1, \ldots, n$ geben ein Gleichungssystem

$$\begin{aligned}
y_1 &= f_1(x_1, x_2, \ldots, x_n)\ , \\
y_2 &= f_2(x_1, x_2, \ldots, x_n)\ , \\
&\ \vdots \\
y_n &= f_n(x_1, x_2, \ldots, x_n)\ .
\end{aligned}$$

Für jede Wahl von $y \in B$ gibt es genau eine Lösung $x \in A$ des Gleichungssystems, und die Lösung x hängt stetig differenzierbar von y ab.

Korollar 9.4.6. (Offenheitssatz)
Sei $U \subset \mathbb{R}^n$ offen und $f : U \to \mathbb{R}^n$ eine $\mathcal{C}^1$-Abbildung, sodass $df(p) \in GL_n(\mathbb{R})$ für alle $p \in U$. Dann ist $f(U)$ offen in $\mathbb{R}^n$.

Beweis. Sei $q \in f(U)$. Wir müssen zeigen, dass es eine in $\mathbb{R}^n$ offene Umgebung von q gibt, die in $f(U)$ enthalten ist.

Sei $p \in U$ mit $q = f(p)$. Nach Satz 9.4.4 gibt es Umgebungen A von p und B von q, so dass $f|_A : A \to B$ eine Umkehrabbildung $g : B \to A$ hat, die auch $\mathcal{C}^1$ ist. Insbesondere ist g stetig, also sind Urbilder offener Mengen offen. Das Urbild $g^{-1}(V)$ einer Umgebung V von p ist also eine offene Umgebung von q in $\mathbb{R}^n$, und per Konstruktion $g^{-1}(V) = f(V) \subset f(U)$. □

■ Notation: Sei $A \in L(\mathbb{R}^{m+n}, \mathbb{R}^n)$. Wir definieren lineare Abbildungen $A_x \in L(\mathbb{R}^m, \mathbb{R}^n)$ und $A_y \in L(\mathbb{R}^n, \mathbb{R}^n)$ als, für $h \in \mathbb{R}^m$, $k \in \mathbb{R}^n$,

$$A_x h \overset{\text{Def.}}{=} A \begin{pmatrix} h \\ 0 \end{pmatrix} \quad , \quad A_y k \overset{\text{Def.}}{=} A \begin{pmatrix} 0 \\ k \end{pmatrix} \ .$$

In Matrixdarstellung bedeutet dies, dass $[A] = \big([A_x] \,|\, [A_y] \big)$, d.h. dass sich die $n \times (m+n)$-Matrix $[A]$ durch nebeneinanderschreiben der $n \times m$-Matrix $[A_x]$ und der $n \times n$-Matrix $[A_y]$ ergibt.

Das folgende ist die lineare Version des Satzes über implizite Funktionen, den wir danach formulieren und beweisen.

Satz 9.4.7. Sei $A \in L(\mathbb{R}^{m+n}, \mathbb{R}^n)$, sodass A_y invertierbar ist. Dann gibt es zu jedem $h \in \mathbb{R}^m$ ein eindeutiges $k \in \mathbb{R}^n$, sodass

$$A \begin{pmatrix} h \\ k \end{pmatrix} = 0 \ .$$

Es gilt $k = -A_y^{-1} A_x h$.

Beweis. Wir müssen zeigen, dass $A_x h + A_y k = 0$. Dies ergibt sich sofort durch einsetzen. □

Beispiel 9.4.8. Betrachte $f : \mathbb{R}^2 \to \mathbb{R}$, $f(x, y) = x^2(1 - x^2) - y^2$. Die Lösungsmenge von $f(x, y) = 0$ sieht aus wie eine umgekippte 8:

- $x^2(1 - x^2)$ hat Nullstellen $-1, 0, 1$ und ist ≥ 0 für $x \in [-1, 1]$ und < 0 sonst.
- Daher gibt es nur für $x \in [-1, 1]$ überhaupt Lösungen von $y^2 = x^2(1 - x^2)$. Für $x \neq -1, 0, 1$ sind es genau zwei Lösungen, $y = \pm x\sqrt{1 - x^2}$, und für $x = -1, 0, 1$ ist $y = 0$ die einzige Lösung.
- Umgekehrt gesehen: in Abhängigkeit von y hat $f(x, y) = 0$ für x entweder 0, 2, 3, oder 4 Lösungen.

Die Funktion f ist auf $\mathbb{R}^2$ stetig differenzierbar mit

$$[df(x,y)] \;=\; \begin{pmatrix} \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} \end{pmatrix} \;=\; \begin{pmatrix} 2x(1-2x^2) & -2y \end{pmatrix} \;.$$

Wir wollen uns fragen, wie man die Lösungsmenge in der Nähe einer gegeben Lösung parametrisieren kann. Dazu nehmen wir uns drei Punkte heraus:

- $p = (\frac{3}{5}, \frac{12}{25})$: Dann $[df(p)] = \begin{pmatrix} \frac{42}{125} & -\frac{24}{25} \end{pmatrix}$. In der Nähe von p wird $f(x,y) = 0$ durch $(x, g(x))$ gelöst, mit $g(x) = x\sqrt{1-x^2}$. Aber es gibt auch eine Funktion $y \mapsto h(y)$, so dass $f(h(y), y) = 0$ (in der Nähe von p) ($\to$Satz 9.4.9 unten).
- $p = (1, 0)$: Dann $[df(p)] = \begin{pmatrix} -2 & 0 \end{pmatrix}$. Die Lösungen von $f(x,y) = 0$ können in der Nähe von p nicht in der Form $(x, g(x))$ geschrieben werden (Warum?), aber immer noch als $f(h(y), y)$ ($\to$Satz 9.4.9 unten). ?
- $p = (0, 0)$: Dann $[df(p)] = \begin{pmatrix} 0 & 0 \end{pmatrix}$. Die Lösungen von $f(x,y) = 0$ können in der Nähe von p weder als $(x, g(x))$ noch als $(h(y), y)$ geschrieben werden. (Warum?) ?

Satz 9.4.9. (Satz über implizite Funktionen)
Sei $U \subset \mathbb{R}^{m+n}$ und $f : U \to \mathbb{R}^n$ eine $\mathcal{C}^1$-Abbildung. Angenommen,

- $a \in \mathbb{R}^m$, $b \in \mathbb{R}^n$ erfüllen $f(a,b) = 0$, und
- für $A := df(a,b) \in L(\mathbb{R}^{m+n}, \mathbb{R}^n)$ ist $A_y \in L(\mathbb{R}^n)$ invertierbar.

Dann gibt es offene Mengen $V \subset \mathbb{R}^m$ und $W \subset \mathbb{R}^n$ mit $(a,b) \in V \times W \subset U$ und eine eindeutige Abbildung $g : V \to W$, sodass für alle $x \in V$, $y \in W$:

$$f(x,y) = 0 \quad \Leftrightarrow \quad y = g(x) \;.$$

Die Abbildung g ist $\mathcal{C}^1$ (auf V) und erfüllt

$$dg(a) \;=\; -A_y^{-1} A_x \;\in\; L(\mathbb{R}^m, \mathbb{R}^n) \;.$$

■ In Worten heißt dass, dass in der Nähe der gegebenen Lösungen a, b alle Lösungen (x,y) von $f(x,y) = 0$ von der Form $(x, g(x))$ sind, also genau durch den Graphen der Funktion g gegeben sind (wobei $x \in V \subset \mathbb{R}^m$ und $g(x) \in W \subset \mathbb{R}^n$, sodass $(x, g(x)) \in \mathbb{R}^{m+n}$).

Da g stetig differenzierbar ist, hängt die Lösung y der Bedingung $f(x,y) = 0$ stetig differenzierbar von dem Parameter x ab.

Bemerkung 9.4.10. Die lineare Version in Satz 9.4.7 können wir in der Tat als Spezialfall von Satz 9.4.9 erhalten. Aus Satz 9.4.7 haben wir $A \in L(\mathbb{R}^{m+n}, \mathbb{R}^n)$ mit A_y invertierbar.

In Satz 9.4.9 wählen wir $U = \mathbb{R}^{m+n}$, und f ist die lineare Abbildung $f(p) = Ap$ (wobei $p \in \mathbb{R}^{m+n}$). Wir wählen $a = 0$, $b = 0$ (dann $f(a,b) = 0$). Die Notation A verursacht keine Probleme, da hier gilt $df(p) = A$ für alle $p \in \mathbb{R}^{m+n}$ (Warum?). ?

Wir erhalten offene Umgebungen $V \subset \mathbb{R}^m$, $W \subset \mathbb{R}^n$ von 0 (in der Tat können wir in diesem Fall $V = \mathbb{R}^m$ und $W = \mathbb{R}^n$ wählen), und ein $g : V \to W$, so dass $f(x,y) = 0 \Leftrightarrow y = g(x)$. Die Abbildung g kennen wir hier aus Satz 9.4.7 explizit: $g(x) = -A_y^{-1}A_x x$. Dann ist auch klar, dass $dg(x) = -A_y^{-1}A_x$, wieder für alle $x \in \mathbb{R}^m$. Insbesondere gilt $dg(0) = -A_y^{-1}A_x$, wie in Satz 9.4.9 gezeigt.

Beweis von Satz 9.4.9. Wir führen die Aussage auf den Satz von der lokalen Umkehrbarkeit zurück (Satz 9.4.4). Dazu brauchen wir zuerst eine lokal invertierbare Abbildung. Wir setzen

$$F : U \to \mathbb{R}^{m+n} \quad , \quad (x,y) \mapsto (x, f(x,y)) \ ,$$

wobei $x \in \mathbb{R}^m$, $y \in \mathbb{R}^n$. Es gilt $F(a,b) = (a,0)$.

Beh. 1: F ist $\mathcal{C}^1$ und $dF(a,b) \in GL_{m+n}(\mathbb{R})$.

⌈ Es gilt, für $h \in \mathbb{R}^m$, $k \in \mathbb{R}^n$:

$$dF(x,y)\begin{pmatrix} h \\ k \end{pmatrix} \ = \ \begin{pmatrix} h \\ df(x,y)(h,k) \end{pmatrix}$$

(Warum?). Insbesondere ist $dF(a,b)(h,k) = 0$ äquivalent zu $h = 0$ und $df(a,b)(0,k) = A_y k = 0$. Da A_y invertierbar ist, folgt auch $k = 0$. Somit ist $dF(a,b)$ injektiv, und da $dF(a,b)$ eine lineare Abbildung $\mathbb{R}^{m+n} \to \mathbb{R}^{m+n}$ ist, ist es auch bijektiv (→Lineare Algebra). ?

Es bleibt, zu zeigen, dass die Abbildung $U \mapsto L(\mathbb{R}^{m+n})$, $(x,y) \mapsto dF(x,y)$ stetig ist. Es gilt

$$\big(dF(x,y)(h,k) - dF(x',y')\big)(h,k) \ = \ \begin{pmatrix} 0 \\ (df(x,y) - df(x',y'))(h,k) \end{pmatrix} \ ,$$

also $\|dF(x,y)(h,k) - dF(x',y')\|_{\mathrm{op}} = \|df(x,y)(h,k) - df(x',y')\|_{\mathrm{op}}$ (Warum?). Da df stetig ist, ist also auch dF stetig. ⌋ ?

Nach Satz 9.4.4 erhalten wir offene Mengen $A, B \subset \mathbb{R}^{m+n}$ mit $(a, b) \in A$, $(a, 0) \in B$, so dass $F|_A : A \to B$ bijektiv ist. Die Umkehrabbildung $G = (F|_A)^{-1} : B \to A$ ist stetig differenzierbar mit

$$dG(a, b) = dF(a, b)^{-1} \ .$$

Setze

$$V' = \left\{ x \in \mathbb{R}^m \,\middle|\, (x, 0) \in B \right\} \ .$$

Dann ist $V' \subset \mathbb{R}^m$ offen (Warum?) und $a \in V'$. Definiere ferner $g : V' \to \mathbb{R}^n$ durch die Vorschrift ?

$$(x, g(x)) \ = \ G(x, 0) \ .$$

Da G eine $\mathcal{C}^1$-Abbildung ist, ist g ebenfalls $\mathcal{C}^1$. Für $x \in V'$ ist per Konstruktion $(x, g(x)) \in A \subset U$.

Beh. 2: Für $x \in V'$, $y \in \mathbb{R}^n$ mit $(x, y) \in A$ gilt: $f(x, y) = 0 \Leftrightarrow y = g(x)$.

$\lceil$ $\Rightarrow$: Falls $(x, y) \in A$ mit $f(x, y) = 0$, so gilt $F(x, y) = (x, 0)$. Dann auch $G(x, 0) = (x, y)$, also $y = g(x)$.

$\Leftarrow$: Es gilt $(x, 0) = F(G(x, 0)) = F(x, g(x)) = (x, f(x, g(x)))$. Also $f(x, g(x)) = 0$.

Jetzt brauchen wir noch geignete offene Mengen V, W. Seien $X \subset \mathbb{R}^m$ und $W \subset \mathbb{R}^n$ offen, so dass $X \times W \subset A$ und $(a, b) \in X \times W$ (Warum gibt es solche X, W?). Wir setzen ?

$$V \ \stackrel{\text{Def.}}{=} \ g^{-1}(W) \cap X \ .$$

Dann:

- V ist offen (Warum?), ?
- $a \in V$, da $b \in W$ und $g(a) = b$, also $a \in g^{-1}(W)$, und sowieso $a \in X$,
- $V \subset V'$ und $g(V) \subset W$, so dass $g|_V : V \to W$.
- $(x, y) \in V \times W$ erfüllt $f(x, y) = 0$ genau dann, wenn $y = g(x)$. (Das ist Beh. 2 und die Tatsache, dass $V \times W \subset A$.)

Die Ableitung $dg(a)$ erhalten wir nun durch Ableiten von beiden Seiten der Gleichung $f(x, g(x)) = 0$

$$0 \ = \ df(a, b) \circ (id, dg(a)) \ = \ A_x + A_y dg(a) \ ,$$

also $dg(a) = -A_y^{-1} A_x$. $\square$

Beispiel 9.4.11. Betrachte $m = 3$, $n = 2$ und $f : \mathbb{R}^{3+2} \to \mathbb{R}^2$, gegeben durch

$$\begin{aligned} f_1(x_1, x_2, x_3, y_1, y_2) &= 3 - 4x_2 + x_1 y_2 + 2e^{y_1} \ , \\ f_2(x_1, x_2, x_3, y_1, y_2) &= 2x_1 - x_3 - 6y_1 + y_2 \cos y_1 \ . \end{aligned}$$

Es gilt

$$f(3, 2, 7, 0, 1) = 0 \ .$$

Wir setzen also $a = (3, 2, 7)$ und $b = (0, 1)$. Die Matrixdarstellung der Ableitung von f im Punkt (a, b) ist

$$[df(a, b)] = \begin{pmatrix} 1 & -4 & 0 & 2 & 3 \\ 2 & 0 & -1 & -6 & 1 \end{pmatrix} \ .$$

Also

$$[A_x] = \begin{pmatrix} 1 & -4 & 0 \\ 2 & 0 & -1 \end{pmatrix} \ , \quad [A_y] = \begin{pmatrix} 2 & 3 \\ -6 & 1 \end{pmatrix} \ .$$

Die Matrix $[A_y]$ ist invertierbar. Also kann man den Satz über implizite Funktionen anwenden. Es gibt also Umgebungen $V \subset \mathbb{R}^3$ und $W \subset \mathbb{R}^2$ mit $a \in V$ und $b \in W$, und eine $\mathcal{C}^1$-Funktion $g : V \to W$, sodass $g(3, 2, 7) = (0, 1)$ und

$$\forall x \in V \ : \ f(x, g(x)) = 0 \ .$$

Dies sind alle Lösungen von $f(x, y) = 0$, die $x \in V$ und $y \in W$ erfüllen. Die Ableitung von g an der Stelle a ist gegeben durch

$$[dg(a)] = [-A_y^{-1} A_x] = \begin{pmatrix} \frac{1}{4} & \frac{1}{5} & -\frac{3}{20} \\ -\frac{1}{2} & \frac{6}{5} & \frac{1}{10} \end{pmatrix} \ .$$

9.5 Höhere Ableitungen

■ Sei $U \subset \mathbb{R}^m$ offen und $k > 0$. Falls sie existiert, nennen wir die iterierte partielle Ableitung $D_{i_1} \cdots D_{i_k} f_j$ eine *partielle Ableitung k-ter Ordnung*.

Wir sagen, $f : U \to \mathbb{R}^n$ ist eine *$\mathcal{C}^k$-Abbildung* (oder *k-mal stetig differenzierbar*), wenn für alle $i_1, \ldots, i_k \in \{1, \ldots, m\}$ und $j \in \{1, \ldots, n\}$ die partiellen Ableitungen $D_{i_1} \cdots D_{i_k} f_j$ auf U existieren und stetig sind. Wir sagen, f ist eine *$\mathcal{C}^\infty$-Abbildung*, wenn f eine $\mathcal{C}^k$-Abbildung ist für jedes $k \in \mathbb{N}$.

■ Es kann sein, dass $D_i D_j f$ und $D_j D_i f$ auf U existieren, aber verschieden sind ($\to$Z10A4). Die nächsten beiden Aussagen zeigen, dass dies für $\mathcal{C}^2$-Funktionen nicht passiert, also dass in diesem Fall partielle Ableitungen vertauschen.

Satz 9.5.1. Sei $U \subset \mathbb{R}^2$ offen und $f : U \to \mathbb{R}$, sodass die partiellen Ableitungen $D_1 f$, $D_2 f$, $D_2 D_1 f$ auf U existieren. Falls $D_2 D_1 f$ in einem Punkt $p \in U$ stetig ist, so existiert $(D_1 D_2 f)(p)$ und es gilt

$$(D_1 D_2 f)(p) \;=\; (D_2 D_1 f)(p) \ .$$

Beweis. Sei $p = (a, b)$ und seien $h, k \in \mathbb{R}$ mit $h, k \neq 0$. Sei Q das abgeschlossene Rechteck mit Eckpunkten (a, b), $(a + h, b)$, $(a, b + k)$ und $(a + h, b + k)$. Ist z.B. $h, k > 0$, so können wir auch schreiben $Q = [a, a+h] \times [b, b+k]$. Wir fordern, dass h, k klein genug sind, so dass $Q \subset U$.

Wir betrachten den Ausdruck

$$\Delta \;=\; f(a + h, b + k) - f(a + h, b) \;-\; \big(f(a, b + k) - f(a, b)\big) \ .$$

Da $D_2 f$ auf ganz U existiert, gilt

$$\lim_{k \to 0} \frac{\Delta}{k} \;=\; (D_2 f)(a + h, b) - (D_2 f)(a, b) \ . \qquad (*)$$

Wir sind fertig, wenn wir zeigen können, dass

$$\lim_{h \to 0} \Big(\lim_{k \to 0} \frac{\Delta}{hk} \Big) \;=\; (D_2 D_1 f)(a, b)$$

(also insbesondere müssen wir zeigen, dass der Grenzwert existiert). Denn der Grenzwert auf der rechten Seite ist per Definition gleich $(D_1 D_2 f)(a, b)$. Dazu zeigen wir:

Beh.: Es gibt $(x, y) \in Q$ mit $\Delta = h\,k \cdot (D_2 D_1 f)(x, y)$.

⌈ Für t zwischen (oder gleich) a und $a+h$ setzen wir $u(t) = f(t, b+k) - f(t, b)$. Dann $\Delta = u(a + h) - u(a)$. Da $D_1 f$ auf U existiert, haben wir $u'(t) = (D_1 f)(t, b + k) - (D_1 f)(t, b)$ für alle t. Nach dem Mittelwertsatz gibt es ein x (echt) zwischen a und $a + h$, sodass

$$\Delta = h\, u'(x) \ .$$

Aber mit $v(s) = (D_1 f)(x, s)$ gilt $u'(x) = v(b + k) - v(b)$. Per Annahme existiert $D_2 D_1 f$ auf U, also gilt $v'(s) = (D_2 D_1 f)(x, s)$. Erneutes Anwenden des Mittelwertsatzes ergibt ein y zwischen b und $b + k$, sodass

$$\Delta \;=\; h \cdot \big(v(b + k) - v(b)\big) \;=\; h\,k \cdot v'(y) \ .$$

⌋

Sei nun $\varepsilon > 0$ gegeben. Da D_2D_1f in (a,b) stetig ist, gibt es $\delta > 0$, sodass für $0 < |h|, |k| < \delta$ und alle $(x,y) \in Q$ gilt

$$\Big| (D_2D_1f)(x,y) - (D_2D_1f)(a,b) \Big| \; < \; \varepsilon \; .$$

(Warum gibt es so ein δ? Eigentlich müsste man doch den Euklidischen Abstand der Punkte (a,b) und (x,y) betrachten.) ?

Insbesondere ergibt sich für die Wahl (x,y) aus der Behauptung oben, dass

$$\Big| \frac{\Delta}{hk} - (D_2D_1f)(a,b) \Big| \; < \; \varepsilon \; .$$

Im Grenzübergang $k \to 0$ erhält man zusammen mit $(*)$, dass

$$\Big| \frac{(D_2f)(a+h,b) - (D_2f)(a,b)}{h} - (D_2D_1f)(a,b) \Big| \; \leq \; \varepsilon \; .$$

Das ist nun aber genau die Aussage, dass $\lim_{h\to 0} \frac{(D_2f)(a+h,b)-(D_2f)(a,b)}{h}$ existiert und durch $(D_2D_1f)(a,b)$ gegeben ist. (Warum?) □ ?

Korollar 9.5.2. Sei $U \subset \mathbb{R}^m$ offen. Ist f eine $\mathcal{C}^k$-Abbildung, dann gilt für jede Permutation π von $\{1, \ldots, k\}$ die folgende Gleichung auf U:

$$D_{i_1} \cdots D_{i_k} f \; = \; D_{i_{\pi(1)}} \cdots D_{i_{\pi(k)}} f \; .$$

Beweis. Die Permutationsgruppe wird von Transpositionen erzeugt. Also ist es genug, zu zeigen, dass

$$D_{i_1} \cdots D_{i_m} D_a D_b D_{j_1} \cdots D_{j_n} f \; = \; D_{i_1} \cdots D_{i_m} D_b D_a D_{j_1} \cdots D_{j_n} f \; ,$$

wobei $m+n+2 = k$. (Warum ist das genug? Warum gilt das?) □ ?

Eine Anwendung von höheren Ableitungen ist die mehrdimensionale Version des Taylorschen Satzes (Satz 6.4.3). Für $p, q \in \mathbb{R}^m$ nennen wir die Menge $\{tp + (1-t)q \in \mathbb{R}^m | t \in [0,1]\}$ die *Verbindungsstrecke von p und q.*

Satz 9.5.3. (Taylor-Approximation in mehreren Variablen)
Sei $U \subset \mathbb{R}^m$ und sei $f : U \to \mathbb{R}$ eine $\mathcal{C}^{k+1}$ Abbildung. Betrachte $p \in U$ und $h \in \mathbb{R}^m$ sodass die Verbindungsstrecke von p und $p+h$ in U liegt. Definiere $R_{k+1}(h)$ durch

$$f(p+h) \; = \; f(p) \; + \; \sum_{j=1}^{k} \frac{1}{j!} \sum_{i_1,\ldots,i_j=1}^{m} (D_{i_1} \cdots D_{i_j} f)(p)\, h_{i_1} \cdots h_{i_j} \; + \; R_{k+1}(h)$$

Dann gibt es einen Punkt ξ auf der Verbindungsstrecke von p und $p+h$, sodass

$$R_{k+1}(h) \; = \; \frac{1}{(k+1)!} \sum_{i_1,\ldots,i_{k+1}=1}^{m} (D_{i_1} \cdots D_{i_{k+1}} f)(\xi)\, h_{i_1} \cdots h_{i_{k+1}} \; .$$

Beweis. Wir führen die Aussage auf Satz 6.4.3 zurück. Setze dazu $F(t) = f(p+th)$, mit $t \in [0,1]$. Wir berechnen die Ableitungen von F mit der Kettenregel (Satz 9.2.4):

$$\begin{aligned}
F'(t) &= \sum_{i_1=1}^{m} (D_{i_1} f)(p+th)\, h_{i_1} \ , \\
F''(t) &= \sum_{i_1,i_2=1}^{m} (D_{i_1} D_{i_2} f)(p+th)\, h_{i_1} h_{i_2} \ , \\
&\vdots \\
F^{(j)}(t) &= \sum_{i_1,\dots,i_j=1}^{m} (D_{i_1} \cdots D_{i_j} f)(p+th)\, h_{i_1} \cdots h_{i_j} \ .
\end{aligned}$$

(Wie folgen diese Gleichungen aus der Kettenregel?) Insbesondere ist F ebenfalls $\mathcal{C}^{k+1}$. Der (ein-dimensionale) Taylorsche Satz (Satz 6.4.3) ergibt ?

$$F(t) \ = \ \sum_{j=0}^{k} \frac{F^{(j)}(0)}{j!} t^j \ + \ \frac{F^{(k+1)}(\zeta)}{(k+1)!}$$

für ein $\zeta \in (0,t)$. Für $t=1$ ist dies genau die gewünschte Aussage. □

■ Für den Restterm gilt insbesondere, dass

$$\lim_{h \to 0} \frac{|R_{k+1}(h)|}{|h|^k} \ = \ 0 \ .$$

(Warum?) In diesem Sinne ist das Taylorpolynom bis Grad k die beste Näherung an f im Raum der Polynome von Grad $\leq k$ mit Variablen $h_1, \dots, h_m$. Für $k=1$ erhalten wir gerade die beste lineare Näherung, also die totale Ableitung: ?

$$f(p+h) \ = \ f(p) \ + \ \underbrace{\sum_{i=1}^{m} (D_i f)(p)\, h_i}_{=df(p)h} \ + R_2(h) \ .$$

Beispiel 9.5.4. Sei $U \subset \mathbb{R}^2$ die Menge aller (x,y) mit $x+y+1 \neq 0$. Betrachte

$$f \ : \ U \longrightarrow \mathbb{R} \quad , \quad (x,y) \ \longmapsto \ \frac{1}{x+y+1} \quad .$$

Die Taylor-Approximation um $(0,0)$ von Grad 2 ist

$$f(x,y) = 1 - x - y + x^2 + 2xy + y^2 + R_3(x,y) \ .$$

Z.B.

$$(D_1D_2f)(x,y) = \frac{2}{(x+y+1)^3} \qquad \text{also} \qquad (D_1D_2f)(0,0) = 2\ .$$

Genauso sieht man $(D_1D_1f)(0,0) = 2 = (D_2D_2f)(0,0)$. (Wie passt das zu den Koeffizienten oben?) ?

9.6 Extrema unter Nebenbedingungen

Beispiel 9.6.1. Wir wollen ein achsenparalleles Rechteck in eine Ellipse legen, so dass das Rechteck maximalen Flächeninhalt hat. Die Ellipse soll Durchmesser $2a$ und $2b$ haben. Die Ecken des maximalen Rechtecks werden sicher die Ellipse berühren. Wir können die Frage also so formulieren:

Für welche $x, y \in \mathbb{R}_{\geq 0}$ mit $(\frac{x}{a})^2 + (\frac{y}{b})^2 = 1$ ist $F = 4xy$ maximal?

■ In Definition 6.2.1 hatten wir die Begriffe „lokales Maximum / Minimum / Extremum“ eingeführt.

Satz 9.6.2. (Lagrange Multiplikatoren)
Sei $U \subset \mathbb{R}^n$ offen und seien $f : U \to \mathbb{R}$, $\varphi : U \to \mathbb{R}^k$ $\mathcal{C}^1$-Abbildungen. Es gelte

- $M \overset{\text{Def.}}{=} \varphi^{-1}(\{0\}) \neq \emptyset$,
- $d\varphi(x) : \mathbb{R}^n \to \mathbb{R}^k$ ist surjektiv für alle $x \in M$.

Angenommen $x_0 \in M$ ist ein lokales Extremum von f auf M. Dann gibt es $\lambda_1, \ldots, \lambda_k \in \mathbb{R}$, sodass

$$df(x_0) \;=\; \lambda_1 d\varphi_1(x_0) \;+\; \cdots \;+\; \lambda_k d\varphi_k(x_0)$$

wobei φ_i die Komponentenfunktionen vor φ sind.

■ Die Konstanten $\lambda_1, \ldots, \lambda_k$ nennt man auch *Lagrange Multiplikatoren*

Beweis. Schreibe $\alpha_i = d\varphi_i(x_0) \in L(\mathbb{R}^n, \mathbb{R})$, $i = 1, \ldots, k$. Da $d\varphi(x_0) \in L(\mathbb{R}^n, \mathbb{R}^k)$ surjektiv ist, sind $\alpha_1, \ldots, \alpha_k$ linear unabhängig. (Das ist lineare Algebra. Details?) ?

Setze $d = n-k$. Dann $d \geq 0$ (Warum?). Wir können $\beta_1, \ldots, \beta_d \in L(\mathbb{R}^n, \mathbb{R})$ finden, so dass ?

$$\{\alpha_1, \ldots, \alpha_k, \beta_1, \ldots, \beta_d\}$$

eine Basis von $L(\mathbb{R}^n, \mathbb{R})$ bilden. (Das ist lineare Algebra. Details?) Da auch ?

$df(x_0) \in L(\mathbb{R}^n, \mathbb{R})$, gibt es (eindeutige) $\lambda_i, \mu_j \in \mathbb{R}$, so dass

$$df(x_0) \;=\; \sum_{i=1}^{k} \lambda_i \, \alpha_i \;+\; \sum_{j=1}^{d} \mu_j \, \beta_j \;. \tag{$*$}$$

Wir müssen zeigen, dass $\mu_1 = \cdots = \mu_d = 0$. Dazu benutzen wir:

Beh.: Für alle $v \in \mathbb{R}^n$ gilt: $d\varphi(x_0)v = 0 \Rightarrow df(x_0)v = 0$.

⌈ Definiere $F : U \to \mathbb{R}^n$ als

$$F(x) \;=\; \big(\varphi_1(x), \ldots, \varphi_k(x), \beta_1(x), \ldots, \beta_d(x)\big)$$

Dann, für $x \in U$ und $h \in \mathbb{R}^n$,

$$dF(x) \;=\; \big(d\varphi_1(x)h, \ldots, d\varphi_k(x)h, \beta_1(h), \ldots, \beta_d(h)\big) \;.$$

(Warum?) Insbesondere ist F stetig differenzierbar auf U und $dF(x_0) \in GL_n(\mathbb{R})$, da alle Zeilenvektoren in $[dF(x_0)]$ linear unabhängig sind. ?

Der Satz über lokale Umkehrbarkeit (Satz 9.4.4) gibt uns offene Mengen $A, B \subset \mathbb{R}^n$, so dass $x_0 \in A$ und $F|_A : A \to B$ bijektiv. Sei $G : B \to A$ die Umkehrfunktion.

Sei nun $v \in \mathbb{R}^n$ gegeben mit $v \neq 0$ und $d\varphi(x_0)v = 0$. Dann gilt

$$dF(x_0)v \;=\; \big(0, \ldots, 0, \beta_1(v), \ldots, \beta_d(v)\big) \;\stackrel{\text{Def.}}{=}\; a \;.$$

Für $t \in (-\delta, \delta)$ und $\delta > 0$ klein genug setze $\gamma(t) = G(F(x_0) + ta)$. Dann $\gamma(0) = x_0$ und

$$\gamma'(0) \;=\; dG(F(x_0))a \;=\; dG(F(x_0))dF(x_0)v \;=\; v \;.$$

(Warum gilt die letzte Gleichheit?) Ferner gilt $\gamma(t) \in M$ für alle $t \in (-\delta, \delta)$, denn aus ?

$$F(\gamma(t)) \;=\; F(x_0) + ta \;=\; (0, \ldots, 0, \beta_1(x_0) + ta_1, \ldots, \beta_d(x_0) + ta_d)$$

folgt $\varphi(\gamma(t)) = 0$.

Per Annahme hat f in x_0 ein lokales Extremum auf M. Also hat auch $g(t) = f(\gamma(t))$ in $t = 0$ ein lokales Extremum. Da g in 0 differenzierbar ist, gilt nach Satz 6.2.2, dass $g'(0) = 0$. Aber $g'(0) = df(\gamma(0))\gamma'(0) = df(x_0)v$. ⌋

Es gibt $u_1, \ldots, u_d \in \mathbb{R}^n$, sodass $\alpha_i(u_j) = 0$ für alle $i = 1, \ldots, k$ und $j = 1, \ldots, d$, und $\beta_i(u_j) = \delta_{i,j}$ (Das ist lineare Algebra. Details?). Für jedes $j = 1, \ldots, d$ gilt nun: $d\varphi(x_0)u_j = 0$, also nach Behauptung $df(x_0)u_j = 0$, also nach $(*)$ auch $\mu_j = 0$. Insgesamt wird aus $(*)$ die Gleichung ?

$$df(x_0) = \sum_{i=1}^{k} \lambda_i \, \alpha_i \, .$$

□

Beispiel 9.6.3.

(1) Hier fahren wir mit Beispiel 9.6.1 fort.

Wir setzen $f, \varphi : \mathbb{R}^2 \to \mathbb{R}$ mit $f(x, y) = 4xy$ und $\varphi(x, y) = (\frac{x}{a})^2 + (\frac{y}{b})^2 - 1$, sodass $M = \varphi^{-1}(\{0\}) = \{(x, y) | (\frac{x}{a})^2 + (\frac{y}{b})^2 = 1\}$ die gewünschte Ellipse ist. Es gilt

$$[df(x, y)] = \begin{pmatrix} 4y & 4x \end{pmatrix} \quad , \quad [d\varphi(x, y)] = \begin{pmatrix} \frac{2x}{a^2} & \frac{2y}{b^2} \end{pmatrix} \, .$$

Auf M ist $d\varphi(x, y)$ surjektiv. Die Bedingung $df(x, y) = \lambda d\varphi(x, y)$ kann nur gelöst werden, wenn die beiden Vektoren kolinear sind. Dazu muss die folgende Determinante Null sein:

$$0 = \det \begin{pmatrix} 4y & 4x \\ \frac{2x}{a^2} & \frac{2y}{b^2} \end{pmatrix} = 8\Big(\big(\tfrac{y}{b}\big)^2 - \big(\tfrac{x}{a}\big)^2 \Big)$$

Auf M gibt es also höchstens 4 lokale Extrema (der Satz über Lagrange Multiplikatoren ist keine Äquivalenz, nur eine Implikation):

$$\frac{x}{a} = \pm \frac{1}{\sqrt{2}} \quad , \quad \frac{y}{b} = \pm \frac{1}{\sqrt{2}}$$

Da M kompakt ist, und f stetig, nimmt f auf M sein Maximum und Minimum an (Satz 5.3.5). Es gilt

$$f(\tfrac{a}{\sqrt{2}}, \tfrac{b}{\sqrt{2}}) = f(-\tfrac{a}{\sqrt{2}}, -\tfrac{b}{\sqrt{2}}) = 2ab \quad , \quad f(\tfrac{a}{\sqrt{2}}, -\tfrac{b}{\sqrt{2}}) = f(-\tfrac{a}{\sqrt{2}}, \tfrac{b}{\sqrt{2}}) = -2ab$$

Für $x \geq 0$ und $y \geq 0$ ist das eindeutige Maximum also $x = a/\sqrt{2}$ und $y = b/\sqrt{2}$.

(2) Wenn man die Bedingung „$d\varphi(x)$ ist surjektiv“ weglässt, stimmt die Aussage von Satz 9.6.2 nicht mehr. Z.B. $f, \varphi : \mathbb{R}^2 \to \mathbb{R}$, $f(x, y) = (1 + x)\cos(y)$, $\varphi(x, y) = x^3$. Dann $M = \varphi^{-1}(\{0\}) = \{(0, y) | y \in \mathbb{R}\}$.

Es ist $f(0,y) = \cos(y)$, und es gibt lokale Extrema bei $(0, m\pi)$ mit $m \in \mathbb{Z}$. Andererseits gilt

$$[df(0,y)] = \begin{pmatrix} \cos(y) & -\sin(y) \end{pmatrix} \quad , \quad [d\varphi(0,y)] = \begin{pmatrix} 0 & 0 \end{pmatrix} \; , \qquad (*)$$

sodass niemals $df(0,y) = \lambda\, d\varphi(x,y)$ gilt.

Wählt man dagegen $\tilde{\varphi}(x,y) = x$, so ist M wie vorher, aber $[d\tilde{\varphi}(0,y)] = (1\ 0)$, und dies ist surjektiv. Die Gleichung $df(0,y) = \lambda\, d\tilde{\varphi}(0,y)$ hat nun Lösungen genau bei $\sin(y) = 0$, also $y = m\pi$ mit $m \in \mathbb{Z}$.

9.7 Differentiation von Integralen

Sei $f : [a,b] \times [c,d] \to \mathbb{R}$, $(x,t) \mapsto f(x,t)$. In diesem Kapitel untersuchen wir, wann die folgende Gleichheit gilt (und wann beide Seiten Sinn machen)

$$\frac{\partial}{\partial t} \int_a^b f(x,t)\, dx \;=\; \int_a^b \frac{\partial}{\partial t} f(x,t)\, dx \;.$$

Beispiel 9.7.1.

(1) Sei $f : [1,2] \times \mathbb{R}_{>0} \to \mathbb{R}$, $f(x,t) = \log(xt)$. Dann

$$\begin{aligned} \frac{\partial}{\partial t} \int_1^2 \log(xt)\, dx \;&=\; \frac{\partial}{\partial t}\Big(\log(t) + \big(x\log(x) - x\big)\Big|_1^2 \Big) \;=\; \frac{1}{t} \;, \\ \int_1^2 \frac{\partial}{\partial t} \log(xt)\, dx \;&=\; \int_1^2 \frac{1}{t}\, dx \;=\; \frac{1}{t} \;. \end{aligned}$$

In diesem Fall gilt also die Gleichheit $(*)$.

(2) Betrachte $f : [0,1] \times \mathbb{R}_{\geq 0} \to \mathbb{R}$. Wir schreiben auch $f_t(x) = f(x,t)$, wenn wir $f(x,t)$ als Funktion von x für festes t verstehen wollen. Der Graph von f_t setze sich zusammen aus den Verbindungsstrecken $(0,0) \to (\sqrt{t}, \sqrt{t})$, $(\sqrt{t}, \sqrt{t}) \to (2\sqrt{t}, 0)$, und $(2\sqrt{t}, 0) \to (1,0)$. Insbesondere ist $f(x,0) = 0$ für $x \in [0,1]$.

$$\begin{aligned} \frac{\partial}{\partial t} \int_0^1 f(x,t)\, dx \Big|_{t=0} \;&=\; \frac{\partial}{\partial t} t \Big|_{t=0} \;=\; 1 \;, \\ \int_0^1 \frac{\partial}{\partial t} f(x,t) \Big|_{t=0} dx \;&=\; \int_0^1 0\, dx \;=\; 0 \;. \end{aligned}$$

In diesem Fall gilt die Gleichheit $(*)$ nicht. (Warum nehmen wir in dem Beispiel immer $\sqrt{t}$ und nicht einfach t?) ?

Satz 9.7.2. Sei $f : [a,b] \times [c,d] \to \mathbb{R}$ und $s \in [c,d]$. Schreibe $f_t(x) = f(x,t)$. Es gelte:

- $f_t \in \mathcal{R}(a,b)$ für alle $t \in [c,d]$.
- $(D_2 f)(x,t)$ existiert auf $[a,b] \times [c,d]$.
- Für alle $\varepsilon > 0$ gibt es $\delta > 0$, sodass für alle $x \in [a,b]$ und $t \in [c,d]$ mit $|t-s| < \delta$ gilt
$$\big|D_2 f(x,t) - D_2 f(x,s)\big| < \varepsilon \ .$$

Setze $F(t) = \int_a^b f(x,t)\,dx$. Dann existiert $F'(s)$ und $x \mapsto D_2 f(x,s)$ ist in $\mathcal{R}(a,b)$. Ferner gilt
$$F'(s) \ = \ \int_a^b D_2 f(x,s)\,dx \ .$$

Beweis. Schreibe $(D_2 f)_s$ für die Funktion $x \mapsto D_2 f(x,s)$. Für $t \in [c,d]$, $t \neq s$ setze
$$\phi_t(x) \ = \ \frac{f(x,t) - f(x,s)}{t-s} \ .$$
Beh.: Für $t \to s$ konvergiert ϕ_t auf $[a,b]$ gleichmäßig gegen $(D_2 f)_s$

⌈ Sei $\varepsilon > 0$ gegeben. Wir müssen zeigen: Es gibt ein $\delta > 0$, sodass für alle $t \in [c,d]$ mit $0 < |t-s| < \delta$ und alle $x \in [a,b]$ gilt:
$$\big|\phi_t(x) - (D_2 f)_s(x)\big| \ < \ \varepsilon \ .$$

Wähle δ wie in der Voraussetzung des Satzes. Nach dem Mittelwertsatz gibt es für jedes x ein u_x zwischen t und s, sodass
$$f(x,t) - f(x,s) \ = \ (t-s)\,(D_2 f)(x,u_x) \ ,$$
also gilt $\phi_t(x) = (D_2 f)(x,u_x)$. Da auch $|u_x - s| < \delta$, folgt per Annahme, dass
$$\big|D_2 f(x,u_x) - D_2 f(x,s)\big| < \varepsilon \ .$$
⌋

Da $\phi_t \in \mathcal{R}(a,b)$, folgt aus der Behauptung und aus Satz 8.2.8 (Aber der Satz ist doch für Folgen. Macht das was?), dass $(D_2 f)_s \in \mathcal{R}(a,b)$ und ?
$$\lim_{t\to s} \int_a^b \phi_t(x)\,dx \ = \ \int_a^b (D_2 f)_s(x)\,dx \ .$$

Insbesondere existiert der Grenzwert auf der linken Seite. Wie berechnen:

$$\begin{aligned}\lim_{t\to s}\int_a^b \phi_t(x)\,dx &= \lim_{t\to s}\frac{\int_a^b f(x,t)\,dx \;-\; \int_a^b f(x,s)\,dx}{t-s}\\ &= \lim_{t\to s}\frac{F(t)-F(s)}{t-s} \;=\; F'(s)\ .\end{aligned}$$

□

Bemerkung 9.7.3. Die Voraussetzungen von Satz 9.7.2 sind z.B. erfüllt, falls $D_2 f$ auf $[a,b]\times[c,d]$ existiert und falls f und $D_2 f$ auf $[a,b]\times[c,d]$ stetig sind. Denn dann ist jedes f_t auf $[a,b]$ stetig, also integrierbar. Und da $[a,b]\times[c,d]$ kompakt ist, ist nach Satz 5.3.12 $D_2 f$ gleichmäßig stetig auf $[a,b]\times[c,d]$. Dies impliziert die dritte Voraussetzung im Satz. (Warum?) (Welche Voraussetzung von Satz 9.7.2 ist in Beispiel 9.7.1 (2) nicht erfüllt?). ?

?

Beispiel 9.7.4. Sei $a>0$ und $g:[-a,a]\to\mathbb{R}$ eine $\mathcal{C}^\infty$-Funktion mit $g(0)=0$. Definiere die Funktion $q:(-a,a)\to\mathbb{R}$ als

$$q(x) \;=\; \begin{cases} g(x)/x & ; x\neq 0\\ g'(0) & ; x=0\end{cases}\quad .$$

Da g insbesondere differenzierbar ist, existiert $g'(0)$. Da $g(0)=0$ gilt $g'(0)=\lim_{t\to 0}\frac{g(t)-g(0)}{t-0} = \lim_{t\to 0}\frac{g(t)}{t}$. Also ist q auf $(-a,a)$ stetig. Aber ist q auch differenzierbar? In der Tat gilt:

Beh.: q ist $\mathcal{C}^\infty$.

⌈ Man kann dies mit der Taylorentwicklung beweisen. Aber wir wollen Satz 9.7.2 benutzen. Für $t\in(-a,a)$ und $t\neq 0$ gilt

$$g(t) \;=\; \int_0^t g'(y)\,dy \;=\; t\int_0^1 g'(tx)\,dx\ .$$

(Was passiert im 2. Schritt?). Dann gilt für $t\neq 0$, dass ?

$$q(t) \;=\; \frac{g(t)}{t} \;=\; \int_0^1 g'(tx)\,dx$$

Für $t=0$ ist der Integrant auf der rechten Seite konstant gleich $g'(0)$, also auch das Integral, also gilt die Gleichung oben für alle $t\in(-a,a)$. Wir haben es geschafft, $q(t)$ auszudrücken, ohne durch t zu teilen!

Definiere $f : [0,1] \times [-a,a]$ als $f(x,t) = g'(tx)$. Dann existiert $D_2 f$ auf ganz $[0,1] \times [-a,a]$ und f und $D_2 f$ sind auf $[0,1] \times [-a,a]$ stetig (Warum gilt das beides?). Also können wir Satz 9.7.2 anwenden: ?

$$q'(t) \; = \; \int_0^1 D_2 f(x,t)\, dx$$

Nun gilt $D_2 f(x,t) = g''(tx)x$. Dies erfüllt aber wieder die Voraussetzungen von Satz 9.7.2, etc, und wir erhalten

$$q^{(n)}(t) \; = \; \int_0^1 \left(\frac{\partial}{\partial t} \right)^n f(x,t)\, dx \; = \; \int_0^1 x^n g^{(n+1)}(xt)\, dx \; .$$

10 Gewöhnliche Differentialgleichungen

Hier folgen wir im wesentlichen [§12 aus Forster II].

■ Unter einem Intervall $I \subset \mathbb{R}$ verstehen wir ein offenes, halboffenes, oder abgeschlossenes Intervall in $\mathbb{R}$, das nicht leer ist und nicht nur aus einem Punkt besteht (damit enthält I also immer unendlich viele Punkte). Das Intervall I darf nach unten und/oder nach oben unbeschränkt sein (d.h. $\pm\infty$ sind als Grenzen erlaubt).

10.1 Beispiele und Definition

Beispiel 10.1.1. Hier geben wir einige wichtige Beispiele von Differentialgleichungen.

(1) *Radioaktiver Zerfall:* Seien $M, h > 0$ Konstanten. Wir suchen Funktionen $u : [0, \infty) \to \mathbb{R}$ mit $u(0) = M$, die differenzierbar sind und für alle t die Gleichung

$$u'(t) \;=\; -h \cdot u(t)$$

erfüllen. Wir denken uns t als Zeit, M als Masse des radioaktiven Materials und h ist das Inverse der Halbwertzeit (bis auf einen Faktor (Welchen?)). ?

Z.B. ist $u(t) = M \exp(-ht)$ eine Lösung. Gibt es noch andere?

(2) *Gedämpfte Schwingung:* Seien M (Masse), b (Dämpfungskonstante), k (Federsteifheit) in $\mathbb{R}_{>0}$, und x_0 (Auslenkung zur Zeit $t = 0$) in $\mathbb{R}$. Wir suchen Funktionen $u : [0, \infty) \to \mathbb{R}$ mit $u(0) = x_0$, die differenzierbar sind und für alle t (Zeit) die Gleichung

$$mu''(t) + bu'(t) + ku(t) \;=\; 0$$

erfüllen. Z.B. ist für $\omega = \sqrt{k/m - b^2/(2m)^2}$ und $\gamma = b/(2m)$ die Funktion $u(t) = x_0 e^{-\gamma t}\cos(\omega t)$ eine Lösung (Warum?), zumindest für $k/m - b^2/(2m)^2 \geq 0$. Gibt es noch andere Lösungen? Was ist, wenn $k/m - b^2/(2m)^2 < 0$?) ?

(3) *Wärmegleichung:* Sei $\kappa > 0$ (Wärmeleitfähigkeit) und $h : \mathbb{R}^3 \to \mathbb{R}_{\geq 0}$ eine $\mathcal{C}^1$-Funktion (Wärmeverteilung zur Zeit $t = 0$). Wir suchen Funktionen $u : \mathbb{R}_{\geq 0} \times \mathbb{R}^3 \to \mathbb{R}$ (die Wärmeverteilung zum Zeitpunkt t) mit $u(0, x, y, z) = h(x, y, z)$ für alle x, y, z und

$$\tfrac{\partial}{\partial t} u(p) \;=\; \kappa \left(\left(\tfrac{\partial}{\partial x}\right)^2 u(p) + \left(\tfrac{\partial}{\partial y}\right)^2 u(p) + \left(\tfrac{\partial}{\partial z}\right)^2 u(p) \right)$$

für alle $p = (t, x, y, z) \in \mathbb{R}_{\geq 0} \times \mathbb{R}^3$.

In Beispiele 1 und 2 hat die Funktion nur einen Parameter, in Beispiel 3 mehrere Parameter. In Beispielen 1 und 2 spricht man von „gewöhnlichen Differentialgleichungen“ und in Beispiel 3 von „partiellen Differentialgleichungen“. Wir befassen uns hier nur mit gewöhnlichen Differentialgleichungen.

Definition 10.1.2. Sei $G \subset \mathbb{R} \times \mathbb{R}^n$ und $f : G \to \mathbb{R}^n$, $(x, y) \mapsto f(x, y)$. Dann nennt man

$$y' = f(x, y) \tag{$*$}$$

ein *System von n (gewöhnlichen) Differentialgleichungen erster Ordnung*. Eine *Lösung von* $(*)$ ist ein Intervall $I \subset \mathbb{R}$ und eine differenzierbare Funktion

$$\varphi : I \to \mathbb{R}^n ,$$

sodass

- der Graph $\left\{ (x, \varphi(x)) \,\middle|\, x \in I \right\}$ von φ eine Teilmenge von G ist, und
- für alle $x \in I$ gilt $\varphi'(x) = f(x, \varphi(x))$.

Bemerkung 10.1.3.

(1) „Erste Ordnung“ bezieht sich darauf, dass es nur eine Ableitung gibt. „System von n Differentialgleichung“ bezieht sich darauf, dass man $(*)$ auch komponentenweise schreiben kann:

$$\begin{aligned} y_1' &= f_1(x, y_1, \ldots, y_n) , \\ y_2' &= f_2(x, y_1, \ldots, y_n) , \\ &\vdots \\ y_n' &= f_n(x, y_1, \ldots, y_n) . \end{aligned}$$

(2) Beispiel 10.1.1 (1) passt direkt zur Definition: Setze $n = 1$, $G = \mathbb{R} \times \mathbb{R}$, $f(x, y) = -hy$, sodass $y' = -hy$.

(3) Treten Ableitung bis zur k-ten Ordnung auf, spricht man von einer *gewöhnlichen Differentialgleichung k-ter Ordnung*. Beispiel = 10.1.1 (2) ist eine gewöhnliche Differentialgleichung zweiter Ordnung.

(4) Mit einem Trick sind gewöhnliche Differentialgleichungen höherer Ordnung äquivalent zu Gleichungen erster Ordnung. Wähle dazu f wie folgt:

$$\begin{aligned} y_1' &= f_1(x, y_1, \ldots, y_n) \stackrel{\text{Def.}}{=} y_2 \;, \\ y_2' &= f_2(x, y_1, \ldots, y_n) \stackrel{\text{Def.}}{=} y_3 \;, \\ &\vdots \\ y_{n-1}' &= f_{n-1}(x, y_1, \ldots, y_n) \stackrel{\text{Def.}}{=} y_n \;, \\ y_n' &= f_n(x, y_1, \ldots, y_n) \;. \end{aligned}$$

Dann $y_1'' = y_2' = y_3$, etc., und allgemein $y_i = y_1^{(i-1)}$ für $i = 2, \ldots, n$. Also kann man die n Gleichungen oben als eine Gleichung schreiben:

$$y_1^{(n)} = f_n(x, y_1, y_1', \ldots, y_1^{(n-1)}) \;.$$

Für Beispiel = 10.1.1 (2) wählt man $n = 2$, $G = \mathbb{R} \times \mathbb{R}^2$, $f_1(x, y_1, y_2) = y_2$, wie oben, und

$$f_2(x, y_1, y_2) = -\tfrac{k}{m} y_1 - \tfrac{b}{m} y_2 \;.$$

Dann, mit $y_1' = y_2$ und $y_1 = u$,

$$u'' = -\tfrac{k}{m} u - \tfrac{b}{m} u' \;.$$

Für den genauen Vergleich von Lösungsmengen von Gleichungen erster und höherer Ordnung siehe [Ende von §12 aus Forster II], wir behandeln dies hier nicht.

10.2 Existenz und Eindeutigkeit

Wir möchten gerne, dass Lösungen zu gewöhnlichen Differentialgleichungen erster Ordnung existieren und eindeutig sind. (Schließlich wird ja in jedem Punkt vorgegeben, wie sich die Lösungsfunktion zu ändern hat, was soll da passieren?) Die folgenden beiden Beispiele sollen unsere Erwartungen etwas dämpfen.

Beispiel 10.2.1.

(1) Setze $G = \mathbb{R} \times \mathbb{R}$, $f(x, y) = \sqrt[3]{y^2} = |y|^{\frac{2}{3}}$, also betrachten wir die Differentialgleichung $|y|^{\frac{2}{3}}$ und suchen Lösungen $\varphi : \mathbb{R} \to \mathbb{R}$ mit $\varphi(x_0) = 0$ für ein $x_0 \in \mathbb{R}$.

Die Funktion $\varphi(x) = 0$ für alle $x \in \mathbb{R}$ ist eine solche Lösung, aber auch $\varphi(x) = \frac{1}{27}\,(x - x_0)^3$.

Allgemeiner erhalten wir eine ganze Familie von Lösungen mit $\varphi(x_0) = 0$: Wähle $a, b \in \mathbb{R}$ mit $x_0 \in [a, b]$ und betrachte die Funktion

$$\varphi(x) = \begin{cases} \frac{1}{27}(x-a)^3 & ; x > a \\ 0 & x \in [a, b] \\ \frac{1}{27}(x-b)^3 & ; x < b \end{cases}$$

(Ist das überhaupt auf ganz $\mathbb{R}$ differenzierbar? Löst das die Differentialgleichung?) ?

Im Allgemeinen können wir also nicht erwarten, dass eine Differentialgleichung erster Ordnung eine eindeutige Lösung hat, selbst wenn wir einen Anfangswert $\varphi(x_0)$ vorgeben.

(2) Setze $G = \mathbb{R} \times \mathbb{R}$, $f(x, y) = 2xy^2$, sodass $y' = 2xy^2$. Sei $c > 0$ eine Konstante. Wir suchen Lösungen mit $\varphi(0) = c$. Wir werden später sehen (Satz 10.2.8), dass für diese Wahl von f gilt: Wenn es überhaupt eine Lösung $\varphi : I \to \mathbb{R}$ gibt mit $\varphi(0) = c$, dann ist diese eindeutig.

Eine Lösung ist durch $I = (-1/\sqrt{c}, 1/\sqrt{c})$ und $\varphi : I \to \mathbb{R}$,

$$\varphi(x) = \frac{1}{c^{-1} - x^2}.$$

gegeben. Diese Lösung divergiert für $x \to \pm 1/\sqrt{c}$. Sie lässt sich somit nicht stetig (oder gar differenzierbar) auf ein größeres Interval ausdehnen, obwohl f auf ganz $\mathbb{R} \times \mathbb{R}$ definiert ist.

Im Allgemeinen können wir also nicht erwarten, dass Lösungen sich auf immere größere Intervalle ausdehnen lassen, selbst wenn der Definitionsbereich G von f dies erlauben würde.

Die folgende Integraldarstellung der Lösungsbedingung einer Differentialgleichung wird später für die Eindeutigkeit und Existenz von Lösungen gebraucht.

Satz 10.2.2. Sei $G \subset \mathbb{R} \times \mathbb{R}^n$ und $f : G \to \mathbb{R}^n$ stetig. Sei ferner

- $(a, c) \in G$ und I ein Intervall mit $a \in I$,
- $\varphi : I \to \mathbb{R}^n$ stetig mit $\varphi(a) = c$, und der Graph von φ sei Teilmenge von G.

Es sind äquivalent:

(i) φ ist auf I differenzierbar und löst die Differentialgleichung $y' = f(x, y)$.

(ii) Für jedes $x \in I$ gilt

$$\varphi(x) \;=\; c + \int_a^x f(t, \varphi(t))\, dt \;.$$

Beweis. (i)⇒(ii): Es gilt

$$\int_a^x f(t, \varphi(t))\, dt \;=\; \int_a^x \varphi'(t)\, dt \;=\; \varphi(x) - \varphi(a) \;=\; \varphi(x) - c \;.$$

(Warum ist $\varphi'(t)$ überhaupt integrierbar? *Hinweis:* $t \mapsto f(t, \varphi(t))$ ist stetig.) ?

(ii)⇒(i): Da $t \mapsto f(t, \varphi(t))$ stetig ist, ist $\varphi(x)$ nach Satz 7.4.2 differenzierbar mit Ableitung $f(x, \varphi(x))$. □

■ Die Aussage des Satzes ist bemerkenswert, weil in der Reformulierung der Bedingung $y' = f(x, y)$ durch Integrale nicht vorausgesetzt werden muss, dass die Lösung $\varphi(x)$ differenzierbar ist. Stetigkeit reicht. Differenzierbarkeit ist dann eine Folgerung.

Definition 10.2.3. (Lipschitz-Stetigkeit)
Sei $M \subset \mathbb{R}^m$ eine Teilmenge und $g : M \to \mathbb{R}^n$ eine Funktion. Wir sagen, g ist *Lipschitz-stetig* mit *Lipschitz-Konstante* $L \geq 0$, falls gilt

$$\forall x, y \in M \;:\; |g(x) - g(y)| \;\leq\; L\,|x - y| \;.$$

Wir sagen g ist *lokal Lipschitz stetig*, falls es für jedes $p \in M$ eine offene Umgebung $U \subset \mathbb{R}^m$ von p gibt, sodass g auf $M \cap U$ Lipschitz-stetig ist (mit einer Lipschitz-Konstanten L, die von U abhängt).

■ Eine (lokal) Lipschitz-stetige Funktion ist automatisch stetig. (Warum?) ?

Beispiel 10.2.4.

(1) $g : \mathbb{R} \to \mathbb{R}$, $g(x) = \sin(x)$ ist Lipschitz-stetig mit Lipschitz-Konstante $L = 1$. Dies folgt aus dem Schrankensatz und $|g'(x)| \leq 1$ für alle x.

(2) $g : \mathbb{R} \to \mathbb{R}$, $g(x) = x^2$ ist lokal Lipschitz-stetig, aber nicht Lipschitz-stetig. (Warum?) ?

(3) $g : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$, $g(x) = \sqrt{x}$ ist werder Lipschitz-stetig noch lokal Lipschitz-stetig. (Warum? *Hinweis:* Wählen Sie $x = 0$ und y immer näher an 0.) ?

Lemma 10.2.5. Sei $M \subset \mathbb{R}^m$ offen, und $g : M \to \mathbb{R}^n$ sei $\mathcal{C}^1$. Dann ist g lokal Lipschitz-stetig.

Beweis. Sei $p \in M$ beliebig und $\overline{B}_r(p)$ der abgeschlossene Ball von Radius r um p. Da M offen ist, ist $\overline{B}_r(p) \subset M$ für $r > 0$ klein genug. Da $\overline{B}_r(p)$ kompakt ist, ist $\|df\|_{\mathrm{op}}$ dort beschränkt, sagen wir durch $L > 0$. Nach dem Schrankensatz (Satz 9.3.10) gilt auf $U = B_r(p)$ (dem offenen Ball), dass $|g(x) - g(y)| \leq L\,|x - y|$ für alle $x, y \in U$. □

Für Differentialgleichungen brauchen wir folgende Variation des Begriffs der lokalen Lipschitz-Stetigkeit.

Definition 10.2.6. Sei $G \subset \mathbb{R} \times \mathbb{R}^n$ und $f : G \to \mathbb{R}^n$. Wir sagen, dass f *lokal eine Lipschitz-Bedingung erfüllt*, falls es für jedes $p \in G$ eine offene (in G) Umgebung U von p gibt, und eine Konstante L_U (die *Lipschitz-Konstante*), sodass für alle $x \in \mathbb{R}$ und $y_1, y_2 \in \mathbb{R}^n$ mit $(x, y_1), (x, y_2) \in U$ gilt:

$$|f(x, y_1) - f(x, y_2)| \;\leq\; L_U\,|y_1 - y_2|\ .$$

■ Man beachte, dass hier nur Werte von f bei gleichem $x \in \mathbb{R}$ verglichen werden. Dennoch ist die lokale Lipschitz-Bedingung stärker als nur zu sagen „für fixes x ist $f(x, y)$ in y lokal Lipschitz-stetig“. Aber die Lipschitz-Bedingung ist schwächer als „f ist auf G lokal Lipschitz-stetig“. (Können Sie für die stärker/schwächer Behauptungen jeweils ein Beispiel finden?) ?

Lemma 10.2.7. Sei $G \subset \mathbb{R} \times \mathbb{R}^n$ offen und $f : G \to \mathbb{R}^n$. Angenommen, die partiellen Ableitungen $\frac{\partial f_j}{\partial y_i}(x, y)$ existieren für $i, j = 1, \ldots, n$ und sind stetig auf G als Funktionen von (x, y). Dann genügt f lokal einer Lipschitz-Bedingung.

■ Hier wird nichts über Stetigkeit oder partielle Differentierbarkeit von $f(x, y)$ bezüglich x verlangt. Aber die partiellen Ableitungen nach y_i müssen als Funktionen auf G stetig sein, nicht nur als Funktionen von y für fixes x (Was ist da der Unterschied?). Umgekehrt sind die Voraussetzungen des Lemmas aber insbesondere erfüllt, falls f als Funktion auf G eine $\mathcal{C}^1$-Abbildung ist. ?

Beweis von Lemma 10.2.7. Der Beweis ist analog zu dem von Lemma 10.2.5. Sei $p \in G$ beliebig. Für fixes x betrachtet man $g_x : y \mapsto f(x, y)$ und das dazugehörige totale Differential $dg_x(y) \in L(\mathbb{R}^n)$. Nach Voraussetzung ist $(x, y) \mapsto dg_x(y)$ stetig auf G, also auch in einer kompakten Umgebung von p. Dort ist $\|dg_x(y)\|_{\mathrm{op}}$ dann beschränkt, und dann geht es weiter wie im Beweis von Lemma 10.2.5. (Details?) □ ?

Satz 10.2.8. (Eindeutigkeitssatz)
Sei $G \subset \mathbb{R} \times \mathbb{R}^n$, und $f : G \to \mathbb{R}^n$ sei stetig und erfülle lokal eine Lipschitz-Bedingung. Sei $I \subset \mathbb{R}$ ein Intervall und $x_0 \in I$. Angenommen $\varphi, \psi : I \to \mathbb{R}^n$ sind Lösungen von $y' = f(x, y)$ und erfüllen $\varphi(x_0) = \psi(x_0)$. Dann gilt

$$\forall x \in I\ :\ \varphi(x) = \psi(x)\ .$$

Beweis. Sei $J \subset I$ die Teilmenge von I, auf der φ und ψ übereinstimmen:

$$J = \left\{ x \in I \,\middle|\, \varphi(x) = \psi(x) \right\} .$$

Per Voraussetzung ist J nicht leer, da $x_0 \in J$.

Beh. 1: J ist abgeschlossen in I.

> ⌈ Die Abbildung $\eta : I \to \mathbb{R}^n$, $x \mapsto \varphi(x) - \psi(x)$ ist stetig, also ist $J = \eta^{-1}(\{0\})$ abgeschlossen in I. ⌋

Beh. 2: J ist offen in I

> ⌈ Sei $a \in J$. Wir müssen zeigen, dass es ein $\delta > 0$ gibt, sodass aus $x \in I$ und $|x - a| < \delta$ folgt, dass $x \in J$.
>
> Setze $b = \varphi(a) = \psi(a)$ und $p = (a, b)$. Da f lokal eine Lipschitz-Bedingung erfüllt, gibt es $\varepsilon, \delta' > 0$ und $L > 0$, sodass für alle x, y_1, y_2 mit $(x, y_1), (x, y_2) \in G$ gilt:
>
> $$\begin{aligned} &|x - a| < \delta' \text{ und } |y_1 - b| < \varepsilon \,,\ |y_2 - b| < \varepsilon \\ &\quad \Rightarrow \quad |f(x, y_1) - f(x, y_2)| < L\,|y_1 - y_2| \,. \end{aligned} \tag{$*$}$$
>
> Da φ, ψ stetig sind, gibt es $\delta'' > 0$, sodass für alle $x \in I$ mit $|x - a| < \delta''$ gilt, dass $|\varphi(x) - \varphi(a)| < \varepsilon$ und $|\psi(x) - \psi(a)| < \varepsilon$. Indem wir δ'' gegebenenfalls verkleinern, können wir verlangen, dass
>
> - $(a - \delta'', a + \delta'') \subset I$, falls a ein innerer Punkt von I ist,
> - $[a, a + \delta'') \subset I$ oder $(a - \delta'', a] \subset I$, falls a der untere, bzw. obere, Randpunkt von I ist.
>
> Folgende Wahl von δ wird sich als ausreichend klein herausstellen:
>
> $$\delta = \tfrac{1}{2} \min\left(\delta', \delta'', \tfrac{1}{L}\right) .$$
>
> Per Konstruktion ist $A = [a - \delta, a + \delta] \cap I$ eines von $[a - \delta, a + \delta]$, $[a - \delta, a]$ oder $[a, a + \delta]$. In jedem Fall ist es kompakt. Da $|\varphi(x) - \psi(x)|$ stetig ist, nimmt es auf A ein Maximum M an, sagen wir im Punkt $x_{\max}$: $|\varphi(x_{\max}) - \psi(x_{\max})| = M$.
>
> Es folgt, dass für alle $x \in A$ gilt
>
> $$|f(x, \varphi(x)) - f(x, \psi(x))| \overset{(*)}{\leq} L|\varphi(x) - \psi(x)| \leq LM \,. \tag{$**$}$$

Aus der Integraldarstellung (Satz 10.2.2) folgt, dass für alle $x \in I$:

$$\varphi(x) - \psi(x) \;=\; \int_a^x \Big(f(t,\varphi(t)) - f(t,\psi(t))\Big) dt$$

Für $x \in A$ können wir dies abschätzen als

$$\begin{aligned} \big|\varphi(x) - \psi(x)\big| \;&\leq\; \Big| \int_a^x \big|f(t,\varphi(t)) - f(t,\psi(t))\big| dt \Big| \\ &\overset{(**)}{\leq}\; LM\,|x-a| \;\leq\; LM\delta \;\leq\; \tfrac{1}{2}M \;. \end{aligned}$$

(Die äußeren Betragsstriche in der ersten Abschätzung brauchen wir, da auch $x < a$ erlaubt ist.) Da dies insbesondere für $x_{\max} \in A$ gilt, erhalten wir

$$M \;=\; |\varphi(x_{\max}) - \psi(x_{\max})| \leq \tfrac{1}{2}M \;.$$

Dies impliziert $M = 0$. Also gilt $\varphi(x) = \psi(x)$ für alle $x \in I$ mit $|x - a| < \delta$. ⌋

Wenn wir zusammenhängende Menge sorgfältig behandelt hätten, wären wir jetzt fertig: Das Komplement von J in I ist offen nach Beh. 1, also ist I die disjunkte Vereinigung der offenen Mengen J und $I \setminus J$. Aber I ist zusammenhängend, also muss eine von beiden Mengen leer sein, aber J ist nicht leer, also ist $I \setminus J$ leer, also ist $I = J$. Haben wir aber nicht, somit machen wir das von Hand:

Beh. 3: $J = I$.

⌈ Wir behandeln die obere Intervallgrenze von I im Detail, das Argument für die untere Grenze ist analog. Sei

$$D_+ \;=\; \big\{\, a \in \mathbb{R} \,\big|\, a \geq x_0 \text{ und } [x_0, a] \subset J \,\big\} \;.$$

Falls D_+ nach oben unbeschränkt ist, gilt $I \cap [x_0, \infty) = [x_0, \infty) = J \cap [x_0, \infty)$. (Warum?) ?

Falls D_+ nach oben beschränkt ist, setze $S = \sup D_+$.

Falls S die obere Intervallgrenze von I ist, sind wir fertig. (Bis auf eine Fallunterscheidung ob $S \in I$ oder $S \notin I$ – Wie geht die?) ?

Ansonsten gibt es $\delta > 0$, so dass auch $[x_0, S + \delta) \subset I$. Per Konstruktion ist $[x_0, S) \subset J$ (Warum ist S a priori nicht dabei?). Da ?
J abgeschlossen ist in I (Beh. 1), ist auch der Abschluss (in I)

von $[x_0, S)$ wieder in J. Da S im inneren von I liegt, ist dieser Abschluss durch $[x_0, S]$ gegeben. Da J offen ist in I (Beh. 2), und nun auch $S \in J$, muss J eine Umgebung von S enthalten, sagen wir $(S-\varepsilon, S+\varepsilon) \in J$ für $0 < \varepsilon < \delta$. Dann enthält J aber Punkte, die echt größer sind als S, Widerspruch (zu S obere Schranke von J). ⌋

(Wo haben wir im Beweis jetzt eigentlich gebraucht, dass $f(x, y)$ auch in x stetig ist?) ? □

Beispiel 10.2.9. Wir hatten schon folgende Beispiele betrachtet:

	Wo?	$G =$	$f(x,y) =$	lokale Lipschitz-Bedingung?
1	Bsp. 10.1.1 (1)	$\mathbb{R} \times \mathbb{R}$	$-hy$	Ja: f ist $\mathcal{C}^1$ auf G.
2	Bsp. 10.1.1 (2)	$\mathbb{R} \times \mathbb{R}^2$	$\begin{pmatrix} y_2 \\ -\frac{k}{m}y_1 - \frac{b}{m}y_2 \end{pmatrix}$	Ja: f ist $\mathcal{C}^1$ auf G.
3	Bsp. 10.2.1 (1)	$\mathbb{R} \times \mathbb{R}$	$\lvert y \rvert^{\frac{2}{3}}$	Nicht bei $(x, 0)$ für $x \in \mathbb{R}$.
4	Bsp. 10.2.1 (2)	$\mathbb{R} \times \mathbb{R}$	$2xy^2$	Ja: f ist $\mathcal{C}^1$ auf G.

Insbesondere sind die Lösungen, die wir für die Beispiele in Zeilen 1, 2, 4 gefunden haben, auf den dort gegebenen I eindeutig, und das Beispiel in Zeile 3 zeigt, dass man die lokale Lipschitz-Bedingung in Satz 10.2.8 nicht einfach weglassen darf.

Bemerkung 10.2.10. Sei I ein Intervall.

(1) Die Menge

$$\mathcal{C}(I, \mathbb{R}^n) = \{ f : I \to \mathbb{R}^n \text{ stetig und beschränkt} \}$$

mit Metrik $d(f, g) = \| f - g \|_\infty = \sup\{ |f(x) - g(x)| \mid x \in I \}$ ist vollständig (dies folgt durch Reduktion des Problems auf die Komponentenfunktionen von f zusammen mit Satz 8.4.2 (Details?)). ?

Die Teilmenge

$$\mathcal{C}^1(I, \mathbb{R}^n) = \{ f : I \to \mathbb{R}^n \text{ stetig diff.bar und beschränkt} \}$$

von $\mathcal{C}(I, \mathbb{R}^n)$ dagegen ist nicht vollständig. (Warum?) ?

(2) Sei $g \in \mathcal{C}(I, \mathbb{R}^n)$ und $r > 0$. Dann ist der „Ball mit Rand“ um g mit Radius r abgeschlossen. Genauer:

Beh.: Die Teilmenge

$$V_r(g) \;=\; \left\{ f \in \mathcal{C}(I, \mathbb{R}^n) \,\middle|\, |f(x) - g(x)| \le r \;\forall x \in I \right\} \subset \mathcal{C}(I, \mathbb{R}^n)$$

ist abgeschlossen in $\mathcal{C}(I, \mathbb{R}^n)$.

⌈ Betrachte die Abbildung $H : \mathcal{C}(I, \mathbb{R}^n) \to \mathbb{R}$, $H(f) = \|f - g\|_\infty$. Dann ist H stetig (Warum?) und es gilt $V_r(g) = H^{-1}([0, r])$. Also ist $V_r(g)$ Urbild einer abgeschlossenen Menge unter einer stetigen Funktion und damit abgeschlossen. ⌋ ?

Da abgeschlossene Teilmengen von vollständigen Räumen wieder vollständig sind (Beispiel 4.2.6), ist $V_r(g)$ vollständig.

Gleich benötigen wir den Fall, dass für ein $b \in \mathbb{R}^n$ die Funktion g die konstante Funktion $g(x) = b$ ist. Wir schreiben in diesem Fall $V_r(b)$ statt $V_r(g)$.

Satz 10.2.11. (Existenzsatz von Picard-Lindelöf)
Sei $G \subset \mathbb{R} \times \mathbb{R}^n$ offen, und $f : G \to \mathbb{R}^n$ sei stetig und erfülle lokal eine Lipschitz-Bedingung. Für jedes $(a, b) \in G$ gibt es ein $\varepsilon > 0$ und eine differenzierbare Funktion

$$\varphi : [a - \varepsilon, a + \varepsilon] \to \mathbb{R}^n \;,$$

sodass gilt:

- $\varphi(a) = b$, und
- φ löst die Differentialgleichung $y' = f(x, y)$.

Bemerkung 10.2.12.

(1) Die Voraussetzungen hier sind etwas stärker als beim Eindeutigkeitssatz (Satz 10.2.8) (G ist hier offen). Insbesondere können wir den Eindeutigkeitssatz anwenden und es folgt, dass die Lösung $\varphi : [a - \varepsilon, a + \varepsilon] \to \mathbb{R}$ die eindeutige Lösung auf $[a - \varepsilon, a + \varepsilon]$ ist, die $\varphi(a) = b$ erfüllt.

(2) Der Existenzsatz (Satz 10.2.11) sagt allerdings nichts darüber aus, wie groß man das Interval I wählen kann, auf dem die Lösung existiert. Dies behandeln wir hier nicht, aber dazu kann man auch Aussagen machen, siehe z.B. [Kap. 4.2 aus Königsberger II].

Beweis von Satz 10.2.11. Da f lokal eine Lipschitz-Bedingung erfüllt, gibt es eine offene Umgebung U von (a,b) in G und ein $L > 0$, sodass für alle $x \in \mathbb{R}$ und $y_1, y_2 \in \mathbb{R}^n$ mit $(x,y_1),(x,y_2) \in G$ gilt:

$$\big|f(x,y_1) - f(x,y_2)\big| \;\leq\; L\,|y_1 - y_2|\;. \tag{$*$}$$

Da G offen ist, ist U auch offen in $\mathbb{R}\times\mathbb{R}^n$. Daher können wir $\delta > 0$ und $r > 0$ finden, sodass folgende kompakte Menge Q eine Teilmenge von U ist:

$$Q \;=\; \big\{(x,y)\,\big|\,|x-a| \leq \delta \text{ und } |y-b| \leq r\,\big\}\;.$$

(Warum gibt es solche δ, r?) Insbesondere gilt $(*)$ auch auf Q. ?

Da Q kompakt ist und f stetig ist, ist f auf Q beschränkt. Es gibt also $M \geq 0$ mit

$$\forall (x,y) \in Q \;:\; |f(x,y)| \leq M\;. \tag{$**$}$$

Wir wählen nun das folgende ε, dessen Sinnhaftigkeit sich im Laufe des Beweises herausstellen wird:

$$\varepsilon \;=\; \min\big(\delta, \tfrac{r}{M}, \tfrac{1}{2L}\big)\;.$$

Wir setzen $I = [a-\varepsilon, a+\varepsilon]$ und betrachten den metrischen Raum

$$V \;=\; \big\{\,f \in \mathcal{C}(I,\mathbb{R}^n)\,\big|\,\forall x \in I \;:\; |f(x)-b| \leq r\,\big\} \subset \mathcal{C}(I,\mathbb{R}^n)\;.$$

Dies ist der Raum $V = V_r(b)$ aus Bemerkung 10.2.10 (2), und dort haben wir gesehen, dass V ein vollständiger Raum ist bezüglich der Metrik $\|f-g\|_\infty$, die sich aus der Supremums-Norm ergibt.

Wir definieren die Abbildung $T : V \to \mathcal{C}(I,\mathbb{R}^n)$, $\psi \mapsto T(\psi)$ als

$$(T(\psi))(x) \;=\; b + \int_a^x f(t,\psi(t))\,dt \quad \text{für alle } \; x \in I\;.$$

Da $|t-a| \leq \delta$ und $|\psi(t) - b| \leq r$ (denn $\psi \in V$) gilt $(t,\psi(t)) \in Q \subset G$, sodass die Auswertung $f(t,\psi(t))$ Sinn macht. (Warum ist $T(\psi) \in \mathcal{C}(I,\mathbb{R}^n)$? ?
D.h. warum ist $x \mapsto (T(\psi))(x)$ stetig und beschränkt?)

Um Klammern zu sparen, schreiben wir ab jetzt $T\psi$ statt $T(\psi)$ (aber T ist im Allgemeinen nicht linear).

Beh. 1: Für $\psi \in V$ ist auch $T\psi \in V$.

⌈ Sei $\psi \in V$ gegeben. Für alle $x \in I$ gilt

$$\begin{aligned}\big|(T\psi)(x) - b\big| \;&=\; \Big|\int_a^x f(t,\psi(t))\,dt\Big| \;\leq\; \Big|\int_a^x \big|f(t,\psi(t))\big|\,dt\Big| \\ &\overset{(**)}{\leq}\; M\,|x-a| \;\leq\; M\varepsilon \;\leq\; r\end{aligned}$$

Da wir schon wissen, dass $T\psi \in \mathcal{C}(I,\mathbb{R}^n)$, ist dies genau die Bedingung, dass $T\psi \in V$. ⌋

Beh. 2: T ist kontrahierend.

⌈ Seien $\varphi, \psi \in V$. Wie breits oben beobachtet, gilt für alle $t \in I$, dass $(t, \phi(t))$ und $(t, \psi(t))$ in Q liegen. Es gilt für alle $x \in I$:

$$\begin{aligned}
&\big|(T\varphi)(x) - (T\psi)(x)\big| \;=\; \Big| \int_a^x \big(f(t,\varphi(t)) - f(t,\psi(t))\big)\, dt \Big| \\
&\overset{(*)}{\leq} \; \Big| \int_a^x L \underbrace{\big|\varphi(t) - \psi(t)\big|}_{\leq \|\varphi-\psi\|_\infty} dt \Big| \;\leq\; L\,|x-a| \cdot \|\varphi - \psi\|_\infty \\
&\leq\; \tfrac{1}{2}\,\|\varphi - \psi\|_\infty \;.
\end{aligned}$$

In der letzten Abschätzung haben wir benutzt, dass $|x - a| \leq \varepsilon$ und $\varepsilon \leq \frac{1}{2L}$. Insgesamt gilt also, dass $\|T\varphi - T\psi\|_\infty \leq \frac{1}{2}\,\|\varphi - \psi\|_\infty$.⌋

Der Banachsche Fixpunktsatz (Satz 9.4.2) gibt nun ein eindeutiges $\varphi \in V$, welches $\varphi = T\varphi$ erfüllt, also

$$\varphi(x) \;=\; b + \int_a^x f(t,\varphi(t))\, dt \quad \text{für alle } \; x \in I \;.$$

Nach Satz 10.2.2 ist dies gleichbedeutend damit, dass φ auf I differenzierbar ist und die Differentialgleichung $y' = f(x,y)$ mit der Anfangsbedingung $\varphi(a) = b$ löst. □

Bemerkung 10.2.13. (Iterationsverfahren von Picard-Lindelöf)
Seien G und f wie in Satz 10.2.11 und sei $(a,b) \in G$. Der Beweis des Satzes zeigt, dass es ein $\varepsilon > 0$ gibt, sodass die eindeutige Lösung φ der Differentialgleichung $y' = f(x,y)$ auf $I = [a - \varepsilon, a + \varepsilon]$ mit $\varphi(a) = b$ iterativ gefunden werden kann.

Definiere rekursiv Funktionen $\varphi_k : I \to \mathbb{R}^n$, $k = 0, 1, 2, \ldots$ wie folgt. Für alle $x \in I$ setze

$$\begin{aligned}
\varphi_0(x) \;&=\; b \;, \\
\varphi_{k+1}(x) \;&=\; b + \int_a^x f(t, \varphi_k(t))\, dt \;.
\end{aligned}$$

Wenn ε so gewählt ist wie im Satz, dann konvergiert die Funktionenfolge $(\varphi_k)_{k \in \mathbb{N}_0}$ auf I gleichmäßig gegen φ.

Beispiel 10.2.14. Wir betrachten das Iterationsverfahren für Beispiel 10.1.1 (1): $G = \mathbb{R} \times \mathbb{R}$ und $f(x,y) - hy$. Als Startpunkt wählen wir $(a,b) = (0, M)$.

Die Funktion f ist sogar (global) Lipschitz-stetig mit $L = 1$ und wir können Q aus dem Beweis beliebig groß wählen. Wegen $L = 1$ erhalten wir $\varepsilon = \frac{1}{2}$, und wir erhalten eine garantierte gleichmäßige Konvergenz auf $I = [-\frac{1}{2}, \frac{1}{2}]$.

In der Tat konvergiert die Folge φ_k sogar auf jedem kompakten Intervall gleichmäßig. Wir erhalten für alle $x \in \mathbb{R}$

$$
\begin{aligned}
\varphi_0(x) &= M \ , \\
\varphi_1(x) &= M + \int_0^x (-h\varphi_0(t))dt \ = \ M - hMx \ , \\
\varphi_2(x) &= M + \int_0^x (-h) \cdot M\,(1-ht)dt \ = \ M\left(1 - hx + \tfrac{1}{2}(-hx)^2\right) \ , \\
&\vdots \\
\varphi_k(x) &= M \sum_{n=0}^{k} \frac{1}{n!}(-hx)^n \ .
\end{aligned}
$$

Die Funktionenfolge $(\varphi_k)_{k\in\mathbb{N}_0}$ konvergiert auf $\mathbb{R}$ punktweise gegen Me^{-hx} und auf jedem kompakten Intervall sogar gleichmäßig. Insbesondere ist die Wahl von ε im Beweis von Satz 10.2.11 typischerweise nicht optimal.

Index