

Analysis II

Walter Bergweiler

Sommersemester 2000

Fassung vom 14. Juli 2000

Diese Vorlesung ist eine Fortsetzung der Vorlesung Analysis I aus dem Wintersemester 1999/2000. Die Numerierung dieser Vorlesung wird hier fortgesetzt; Verweise wie “nach Satz $x.y.z$ ” mit $x \leq 4$ beziehen sich darauf.

Verweise wie “nach LA I, Satz ...” beziehen sich auf die Vorlesung Lineare Algebra I von Professor Anand Srivastav aus dem dem Wintersemester 1999/2000.

Inhaltsverzeichnis

5	Integration	1
5.1	Das Riemann-Integral	1
5.2	Der Hauptsatz der Differential- und Integralrechnung	6
5.3	Integrationstechniken	9
5.4	Integralform des Taylorrestglieds	13
5.5	Folgen und Reihen integrierbarer Funktionen	16
5.6	Uneigentliche Integrale	16
5.7	Bogenlänge ebener Kurven	19
6	Topologie metrischer Räume	22
6.1	Metrik, Norm und Skalarprodukt	22
6.2	Offene und abgeschlossene Mengen	25
6.3	Stetigkeit	29
6.4	Kompaktheit	33
6.5	Zusammenhang	37
7	Differentialrechnung in mehreren Veränderlichen	39
7.1	Partielle Ableitungen	39
7.2	Differenzierbarkeit	44
7.3	Höhere Ableitungen und Taylorformel	49
7.4	Lokale Extrema	51
7.5	Der Satz über implizite Funktionen	56
7.6	Extrema unter Nebenbedingungen	62
7.7	Vertauschung von Grenzprozessen	64
7.8	Taylorreihen	66

5 Integration

5.1 Das Riemann-Integral

Der Flächeninhalt eines Rechtecks ist das Produkt der Seitenlängen. Die Idee beim Riemann-Integral ist die Approximation allgemeinerer Flächen durch Rechtecke und damit Definition des Flächeninhaltes durch einen Grenzübergang.

Definition 5.1.1 Seien $a, b \in \mathbb{R}$ mit $a < b$. Sei $I := [a, b]$. Weiter sei $n \in \mathbb{N}$ und $a = x_0 < x_1 < x_2 < \dots < x_{n-1} < x_n = b$. Dann heißt $Z := \{x_0, x_1, \dots, x_n\}$ eine *Zerlegung* von I . Für $k \in \{1, \dots, n\}$ heißt $I_k := [x_{k-1}, x_k]$ das k -te Teilintervall der Zerlegung und $|I_k| := x_k - x_{k-1}$ die *Länge* von I_k . Weiter heißt $|Z| := \max_k |I_k|$ die *Feinheit* der Zerlegung Z .

Sei $\xi = (\xi_1, \xi_2, \dots, \xi_n) \in I_1 \times I_2 \times \dots \times I_n$ (also $\xi_k \in I_k$ für alle k). Dann heißen die ξ_k (oder auch das n -tupel ξ) *Stützstellen* zur Zerlegung Z .

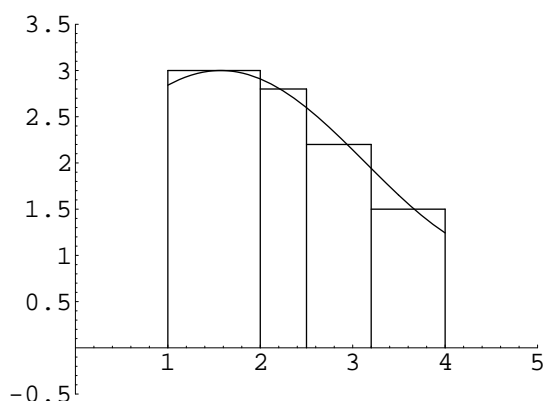
Sei zusätzlich noch $f : I \rightarrow \mathbb{C}$. Dann heißt

$$S(Z, \xi) := S(f, Z, \xi) := \sum_{k=1}^n f(\xi_k) |I_k|$$

Riemannsche Summe von f bzgl. Z und ξ .

Eine Zerlegung eines kompakten Intervalls ist nach Definition also nichts anderes als eine endliche Teilmenge des Intervalls, die die Endpunkte enthält. Schreiben wir im folgenden aber eine Zerlegung Z von $[a, b]$ in der Form $Z = \{x_0, x_1, \dots, x_n\}$, so werden wir immer $a = x_0 < x_1 < x_2 < \dots < x_{n-1} < x_n = b$ annehmen.

Für die Zerlegung $Z = \{1, 2, 2, 5, 3, 2, 4\}$ des Intervalls $[1, 4]$, geeignete Stützstellen ξ und die durch $x \mapsto 2 + \sin x$ gegebene Funktion f sind in folgender Abbildung die Rechtecke dargestellt, deren Flächensumme $S(f, Z, \xi)$ ist.



Die Idee ist nun, einen Grenzübergang $|Z| \rightarrow 0$ durchzuführen.

Definition 5.1.2 Sei $I \subset \mathbb{R}$ kompaktes Intervall und $f : I \rightarrow \mathbb{C}$. Dann heißt f *Riemann-integrierbar* (kurz auch: *integrierbar*) falls gilt:

$$\exists S \in \mathbb{C} \forall \varepsilon \in \mathbb{R}_+ \exists \delta \in \mathbb{R}_+ \forall \text{ Zerlegungen } Z \text{ von } I \forall \text{ Stützstellen } \xi \text{ zu } Z : \\ |Z| < \delta \Rightarrow |S(f, Z, \xi) - S| < \varepsilon.$$

Existiert so ein S , so ist es eindeutig und heißt *Riemann-Integral* (kurz auch: *Integral*) von f über I . Es wird mit $\int_a^b f(x)dx$, $\int_a^b f$ oder $\int_I f$ bezeichnet.

Die Eindeutigkeit von S ist dabei klar; man vgl. den Beweis, daß eine konvergente Zahlenfolge nur einen Grenzwert hat.

Man kann $\int_a^b f$ nicht direkt als Grenzwert von $S(Z, \xi)$ für $|Z| \rightarrow 0$ definieren, da ja Z und ξ durch $|Z|$ nicht eindeutig festgelegt sind. Dennoch übertragen sich viele Resultate über Grenzwerte ohne weiteres. Beispielsweise ist f genau dann integrierbar mit $S = \int_a^b f$ ist, wenn für jede Folge (Z_n) von Zerlegungen mit $|Z_n| \rightarrow 0$ und jede zugehörige Stützstellenfolge (ξ_n) gilt, daß $S(Z_n, \xi_n) \rightarrow S$.

Auch das Cauchy Kriterium gilt analog:

Satz 5.1.1 (Cauchy Kriterium für Integrierbarkeit) Sei $I \subset \mathbb{R}$ kompaktes Intervall und $f : I \rightarrow \mathbb{C}$. Dann ist f integrierbar genau dann, wenn

$$\forall \varepsilon \in \mathbb{R}_+ \exists \delta \in \mathbb{R}_+ \forall \text{ Zerlegungen } Z_1, Z_2 \text{ von } I \forall \text{ Stützstellen } \xi_1, \xi_2 \text{ zu } Z_1, Z_2 : \\ |Z_1| < \delta \wedge |Z_2| < \delta \Rightarrow |S(Z_1, \xi_1) - S(Z_2, \xi_2)| < \varepsilon.$$

Der Beweis ist analog zum Cauchy Kriterium für Funktionengrenzwerte (Satz 3.5.4).

Satz 5.1.2 Sei $I \subset \mathbb{R}$ kompaktes Intervall und $f : I \rightarrow \mathbb{C}$. Ist f stetig, so ist f integrierbar.

Zum Beweis benutzen wir folgenden

Hilfssatz 5.1.1 Sei $I \subset \mathbb{R}$ kompaktes Intervall und $f : I \rightarrow \mathbb{C}$. Seien $\varepsilon, \delta \in \mathbb{R}_+$ und es gelte $|f(x) - f(y)| < \varepsilon$ für alle $x, y \in I$ mit $|x - y| < \delta$. Seien weiter Z, Z' Zerlegungen mit Stützstellen ξ, ξ' . Ist dann $Z \subset Z'$ und $|Z| < \delta$, so gilt $|S(Z, \xi) - S(Z', \xi')| < \varepsilon|I|$.

Beweis. Seien I_1, I_2, \dots, I_m die Teilintervalle von Z und I'_1, I'_2, \dots, I'_n die Teilintervalle von Z' . Dann gilt $m \leq n$ und es gilt $I_1 = I'_1 \cup I'_2 \cup \dots \cup I'_{\ell_1}$, $I_2 = I'_{\ell_1+1} \cup \dots \cup I'_{\ell_2}$, \dots , $I_m = I'_{\ell_{m-1}+1} \cup \dots \cup I'_{\ell_m}$ mit $0 < \ell_1 < \dots < \ell_m := n$. Es folgt mit $\ell_0 := 0$, daß

$$\begin{aligned} |S(Z, \xi) - S(Z', \xi')| &= \left| \sum_{k=1}^m f(\xi_k) |I_k| - \sum_{k=1}^n f(\xi'_k) |I'_k| \right| \\ &= \left| \sum_{k=1}^m f(\xi_k) \sum_{j=\ell_{k-1}+1}^{\ell_k} |I'_j| - \sum_{k=1}^m \sum_{j=\ell_{k-1}+1}^{\ell_k} f(\xi'_j) |I'_j| \right| \\ &\leq \sum_{k=1}^m \sum_{j=\ell_{k-1}+1}^{\ell_k} \underbrace{|f(\xi_k) - f(\xi'_j)|}_{< \varepsilon} |I'_j| \\ &< \varepsilon |I|. \end{aligned}$$

Beweis von Satz 5.1.2. Sei $\varepsilon > 0$. Da f gleichmäßig stetig ist, existiert $\delta > 0$ mit $|f(x) - f(y)| < \varepsilon' := \frac{\varepsilon}{2|I|}$ für alle $x, y \in I$ mit $|x - y| < \delta$.

Es seien jetzt Z_1, Z_2 Zerlegungen von I mit Stützstellen ξ_1, ξ_2 . Es gelte $|Z_1| < \delta$ und $|Z_2| < \delta$. Für $Z := Z_1 \cup Z_2$ und beliebige Stützstellen ξ zu Z folgt mit Hilfssatz 5.1.1 dann

$$|S(Z_1, \xi_1) - S(Z_2, \xi_2)| \leq |S(Z_1, \xi_1) - S(Z, \xi)| + |S(Z, \xi) - S(Z_2, \xi_2)| < \varepsilon'|I| + \varepsilon'|I| = \varepsilon.$$

Die Behauptung folgt mit dem Cauchy Kriterium (Satz 5.1.1). \square

Beispiel. Sei $I = [0, 1]$ und $f := id_I$, also $f(x) = x$. Dann ist f stetig. Es sei $Z_n = \{0, \frac{1}{n}, \frac{2}{n}, \dots, \frac{n-1}{n}, 1\}$ und $\xi_n = (\frac{1}{n}, \frac{2}{n}, \dots, \frac{n-1}{n}, 1)$. Dann gilt $|Z_n| = \frac{1}{n} \rightarrow 0$ und

$$S(Z_n, \xi_n) = \sum_{k=1}^n \frac{k}{n} \cdot \frac{1}{n} = \frac{1}{n^2} \sum_{k=1}^n k = \frac{1}{n^2} \frac{n(n+1)}{2} = \frac{1}{2} \left(1 + \frac{1}{n}\right) \rightarrow \frac{1}{2},$$

also $\int_0^1 x \, dx = \frac{1}{2}$.

Die Umkehrung von Satz 5.1.2 gilt nicht, d. h., integrierbare Funktionen müssen nicht stetig sein. Zum Beispiel zeigt man leicht, daß Funktionen mit endlich (oder sogar abzählbar unendlich) vielen Unstetigkeitsstellen integrierbar sind. Es gilt aber der folgende Satz.

Satz 5.1.3 *Integrierbare Funktionen sind beschränkt.*

Wir skizzieren nur die *Beweisidee*: Sind I, f wie vorher, und ist f unbeschränkt, so existiert eine Folge (x_k) in I mit $|f(x_k)| \rightarrow \infty$. Sei nun Z eine Zerlegung mit Stützstellen $\xi = (\xi_1, \dots, \xi_n)$. Sei I_{j_k} das Teilintervall, welches x_k enthält. (Falls $x_k \in Z$ ist, aber x_k kein Randpunkt von I ist, existieren zwei Teilintervalle mit dieser Eigenschaft. In diesem Falle sei I_{j_k} irgendeines dieser beiden.) Wir betrachten die Stützstellen $\xi^{(k)}$, die wir dadurch erhalten, daß wir ξ_{j_k} durch x_k ersetzen, also $\xi^{(k)} = (\xi_1, \dots, \xi_{j_k-1}, x_k, \xi_{j_k+1}, \dots, \xi_n)$. Dann folgt $|S(Z, \xi^{(k)})| \rightarrow \infty$. \square

Satz 5.1.4 *Sei I kompaktes Intervall und seien $f, g : I \rightarrow \mathbb{C}$ integrierbar. Sei $c \in \mathbb{C}$. Dann sind auch $f+g, c \cdot f$ und \overline{f} integrierbar und es gilt $\int_I (f+g) = \int_I f + \int_I g$, $\int_I c \cdot f = c \cdot \int_I f$ und $\int_I \overline{f} = \overline{\int_I f}$.*

Beweis. Die erste Behauptung folgt aus

$$S(f+g, Z, \xi) = S(f, Z, \xi) + S(g, Z, \xi).$$

Die anderen erhält man analog. \square

Satz 5.1.5 *Sei I kompaktes Intervall und $f : I \rightarrow \mathbb{C}$. Dann ist f genau dann integrierbar, wenn $\operatorname{Re} f$ und $\operatorname{Im} f$ integrierbar sind. In diesem Fall gilt $\int_I f = \int_I \operatorname{Re} f + i \int_I \operatorname{Im} f$.*

Der *Beweis* folgt unmittelbar aus Satz 5.1.4.

Man kann sich also auf die Integration reellwertiger Funktionen beschränken. Für beschränktes $f : I \rightarrow \mathbb{R}$ und eine Zerlegung Z von I nennt man

$$\overline{S}(Z) := \overline{S}(f, Z) := \sup_{\xi} S(Z, \xi)$$

bzw.

$$\underline{S}(Z) := \underline{S}(f, Z) := \inf_{\xi} S(Z, \xi)$$

Riemannsche Ober- bzw. Untersumme von f . Dabei sind Supremum bzw. Infimum über alle Stützstellen ξ zur Zerlegung Z zu nehmen. Für $Z' \subset Z$ gilt dann

$$\underline{S}(Z') \leq \underline{S}(Z) \leq \overline{S}(Z) \leq \overline{S}(Z').$$

Man nennt

$$\overline{\int}_I f := \inf_Z \overline{S}(Z)$$

bzw.

$$\underline{\int}_I f := \sup_Z \underline{S}(Z)$$

oberes bzw. unteres Riemann-Integral von f . Hier sind Supremum bzw. Infimum über alle Zerlegungen Z von I zu nehmen. Man beachte, daß oberes und unteres Riemann-Integral (für beschränktes f) immer existieren.

Satz 5.1.6 (Riemannsches Integrierbarkeitskriterium) Sei I kompaktes Intervall und $f : I \rightarrow \mathbb{R}$ beschränkt. Dann sind die folgenden Aussagen äquivalent:

(i) f ist integrierbar,

(ii) $\underline{\int}_I f = \overline{\int}_I f$,

(iii) Für alle $\varepsilon > 0$ existiert eine Zerlegung Z mit $\overline{S}(Z) - \underline{S}(Z) < \varepsilon$.

Im Falle der Integrierbarkeit gilt $\int_I f = \underline{\int}_I f = \overline{\int}_I f$.

Beweis. Übung.

Hilfssatz 5.1.2 Seien I, J kompakte Intervalle, $f : I \rightarrow J$ integrierbar und $\phi : J \rightarrow \mathbb{R}$ stetig. Dann ist $\phi \circ f$ integrierbar.

Beweis. Sei $\varepsilon > 0$. Da ϕ gleichmäßig stetig nach Satz 3.4.4 existiert $\delta > 0$ mit $|\phi(x) - \phi(y)| < \varepsilon' := \frac{\varepsilon}{2|I|}$ für alle $x, y \in I$ mit $|x - y| < \delta$. Sei $K := \max \phi(J) - \min \phi(J)$. Man kann annehmen, daß $\delta < \varepsilon/(2K + 1)$.

Nun existiert eine Zerlegung $Z = \{x_0, x_1, \dots, x_n\}$ von I mit $\overline{S}(f, Z) - \underline{S}(f, Z) < \delta^2$. Wir definieren nun $M_k := \sup f(I_k)$, $m_k := \inf f(I_k)$, $M_k^* := \sup \phi(f(I_k))$, $m_k^* := \inf \phi(f(I_k))$, $A = \{k : M_k - m_k < \delta\}$ und $B = \{k : M_k - m_k \geq \delta\}$. Für $k \in A$ ist dann $M_k^* - m_k^* < \varepsilon'$. Außerdem gilt für alle k , insbesondere also für $k \in B$, auch $M_k^* - m_k^* \leq K$. Desweiteren ist

$$\sum_{k \in B} |I_k| \leq \frac{1}{\delta} \sum_{k \in B} (M_k - m_k) |I_k| \leq \frac{1}{\delta} \sum_{k=1}^n (M_k - m_k) |I_k| = \frac{1}{\delta} (\overline{S}(f, Z) - \underline{S}(f, Z)) < \delta.$$

Es folgt

$$\begin{aligned} \overline{S}(\phi \circ f, Z) - \underline{S}(\phi \circ f, Z) &= \sum_{k=1}^n (M_k^* - m_k^*) |I_k| \\ &\leq \sum_{k \in A} \varepsilon' |I_k| + \sum_{k \in B} K |I_k| \\ &< \varepsilon' |I| + K \delta \\ &< \varepsilon. \quad \square \end{aligned}$$

Satz 5.1.7 Sei I kompaktes Intervall und seien $f, g : I \rightarrow \mathbb{C}$ integrierbar. Dann sind auch $|f|$ und $f \cdot g$ integrierbar. Außerdem gilt

$$\left| \int_I f \right| \leq \int_I |f| \leq |I| \sup_{x \in I} |f(x)|.$$

Beweis. Für reellwertiges f folgt die Integrierbarkeit von $|f|$ direkt aus Hilfssatz 5.1.2 (mit $\phi(x) = |x|$). Mit $\phi(x) = x^2$ folgt aus diesem Hilfssatz auch die Integrierbarkeit von f^2 . Wegen $fg = \frac{1}{4}((f+g)^2 - (f-g)^2)$ erhält man mit Satz 5.1.4 dann auch die Integrierbarkeit von $f \cdot g$, falls f und g beide reellwertig sind.

Für komplexwertiges f, g erhält man die Integrierbarkeit von $|f|$ und $f \cdot g$ jetzt mit Satz 5.1.5 (und wiederum Hilfssatz 5.1.2).

Die Abschätzung von $\left| \int_I f \right|$ folgt wegen

$$|S(f, Z, \xi)| \leq S(|f|, Z, \xi) \leq |I| \sup_{x \in I} |f(x)|$$

für alle Zerlegungen Z von I mit Stützstellen ξ . \square

Mit dem letzten Argument zeigt man auch leicht folgenden Satz.

Satz 5.1.8 Sei I kompaktes Intervall und seien $f, g : I \rightarrow \mathbb{R}$ integrierbar. Gilt $f(x) \leq g(x)$ für alle $x \in I$, so gilt $\int_I f \leq \int_I g$.

Folgerung 5.1.1 Ist I kompaktes Intervall und $f : I \rightarrow \mathbb{R}$ integrierbar, so ist

$$|I| \inf f(I) \leq \int_I f \leq |I| \sup f(I).$$

Die linke bzw. rechte Seite in obiger Ungleichung sind auch die Riemannsche Unter- bzw. Obersumme zu der Zerlegung von I , die nur aus den beiden Randpunkten besteht.

Satz 5.1.9 (Mittelwertsatz der Integralrechnung) Sei I kompaktes Intervall und $f : I \rightarrow \mathbb{R}$ integrierbar. Dann existiert $\mu \in \mathbb{R}$ mit $\inf f(I) \leq \mu \leq \sup f(I)$ und $\int_I f = \mu|I|$. Ist f stetig, so existiert $\xi \in \text{int}(I)$ mit $\mu = f(\xi)$.

Beweis. Die erste Behauptung folgt aus obiger Folgerung, die zweite aus dem Zwischenwertsatz. \square

Analog beweist man das folgende Ergebnis.

Satz 5.1.10 (Erweiterter Mittelwertsatz der Integralrechnung) Sei I kompaktes Intervall und seien $f, g : I \rightarrow \mathbb{R}$ integrierbar. Sei $g(x) \geq 0$ für alle $x \in I$. Dann existiert $\mu \in \mathbb{R}$ mit $\inf f(I) \leq \mu \leq \sup f(I)$ und $\int_I fg = \mu \int_I g$. Ist f stetig, so existiert $\xi \in \text{int}(I)$ mit $\mu = f(\xi)$.

Satz 5.1.11 Sei I kompaktes Intervall und $f : I \rightarrow \mathbb{R}$ monoton. Dann ist f integrierbar.

Beweis. Ohne Beschränkung der Allgemeinheit sei f monoton steigend und nicht konstant. Dann ist $f(b) > f(a)$. Sei $\varepsilon > 0$ und Z Zerlegung von I mit $|Z| < \delta := \varepsilon/(f(b) - f(a))$. Dann gilt

$$\begin{aligned} \overline{S}(f, Z) - \underline{S}(f, Z) &= \sum_{k=1}^n (f(x_k) - f(x_{k-1})) |I_k| \\ &< \delta \sum_{k=1}^n (f(x_k) - f(x_{k-1})) \\ &= \delta (f(b) - f(a)) \\ &= \varepsilon. \quad \square \end{aligned}$$

Satz 5.1.12 (i) Sind I, J kompakte Intervalle mit $J \subset I$ und ist $f : I \rightarrow \mathbb{C}$ integrierbar (über I), so ist auch $f|_J$ integrierbar (über J).

(ii) Seien $a, b, c \in \mathbb{R}$, $a < b < c$ und sei $f : [a, c] \rightarrow \mathbb{C}$. Sind dann $f|_{[a, b]}$ und $f|_{[b, c]}$ integrierbar, so ist auch f integrierbar und es gilt

$$\int_a^c f = \int_a^b f + \int_b^c f.$$

Beweis. Übung.

Wegen (ii) setzt man für $a < b$ und integrierbares $f : [a, b] \rightarrow \mathbb{C}$ auch $\int_b^a f := -\int_a^b f$. Außerdem setzt man $\int_a^a f = 0$ für $f : \{a\} \rightarrow \mathbb{C}$. Mit diesen Setzungen gilt die Formel in (ii) für alle $a, b, c \in \mathbb{R}$, wenn die auftretenden Integrale definiert sind.

5.2 Der Hauptsatz der Differential- und Integralrechnung

Definition 5.2.1 Sei I Intervall, $F : I \rightarrow \mathbb{C}$ differenzierbar und $f : I \rightarrow \mathbb{C}$. Dann heißt F *Stammfunktion* von f , falls $F' = f$.

Seien F, G Stammfunktionen von f, g und sei $c \in \mathbb{C}$. Unmittelbar sieht man ein, daß dann $F + G$ Stammfunktion von $f + g$ und $c \cdot F$ Stammfunktion von $c \cdot f$ ist.

Mit F ist auch $F + c$ Stammfunktion von f . Umgekehrt gilt nach Satz 4.3.1, daß wenn F_1 und F_2 Stammfunktionen von f sind, so existiert $c \in \mathbb{C}$ mit $F_1 = F_2 + c$.

Beispiel. Sei $I := (-1, 1)$, $f : I \rightarrow \mathbb{R}$, $f(x) = -\frac{1}{\sqrt{1-x^2}}$. Eine Stammfunktion von f ist $\arccos |x|$, eine andere ist $-\arcsin |x|$. Man beachte, daß $\arccos = \frac{\pi}{2} - \arcsin$.

Satz 5.2.1 (Hauptsatz der Differential- und Integralrechnung) Seien $a, b \in \mathbb{R}$, $a < b$ und $f : [a, b] \rightarrow \mathbb{C}$ integrierbar.

(i) Besitzt f eine Stammfunktion F , so gilt $\int_a^b f = F(b) - F(a)$.

(ii) Sei $F : [a, b] \rightarrow \mathbb{C}$ definiert durch $F(x) = \int_a^x f$. Ist f stetig in $y \in [a, b]$, so ist F differenzierbar in y mit $F'(y) = f(y)$. Ist f stetig in $[a, b]$, so ist F Stammfunktion von f .

Beweis. Ohne Beschränkung der Allgemeinheit sei $f : [a, b] \rightarrow \mathbb{R}$.

zu (i): Sei $Z = \{x_0, x_1, \dots, x_n\}$ Zerlegung von $[a, b]$. Nach dem Mittelwertsatz (der Differentialrechnung) existiert zu $k \in \{1, \dots, n\}$ dann $\xi_k \in (x_{k-1}, x_k)$ mit

$$F(x_k) - F(x_{k-1}) = F'(\xi_k)(x_k - x_{k-1}) = f(\xi_k)(x_k - x_{k-1}).$$

Mit $\xi = (\xi_1, \dots, \xi_n)$ erhält man

$$\begin{aligned} F(b) - F(a) &= F(x_n) - F(x_0) \\ &= \sum_{k=1}^n F(x_k) - F(x_{k-1}) \\ &= \sum_{k=1}^n f(\xi_k)(x_k - x_{k-1}) \\ &= S(f, Z, \xi). \end{aligned}$$

Hieraus folgt die Behauptung.

zu (ii): Seien $x, y \in [a, b]$, $x \neq y$. Dann gilt nach Satz 5.1.12

$$F(x) - F(y) = \int_a^x f - \int_a^y f = \int_x^y f$$

und damit

$$\begin{aligned} \left| \frac{F(x) - F(y)}{x - y} - f(y) \right| &= \left| \frac{1}{x - y} \int_y^x f(t) dt - f(y) \frac{1}{x - y} \int_y^x 1 dt \right| \\ &= \frac{1}{|x - y|} \left| \int_y^x f(t) - f(y) dt \right| \\ &\leq \sup_{t \in [y, x]} |f(t) - f(y)|. \end{aligned}$$

Ist f stetig in y , so strebt die rechte Seite für $x \rightarrow y$ gegen 0 und es folgt

$$\lim_{x \rightarrow y} \frac{F(x) - F(y)}{x - y} = f(y). \quad \square$$

Die im Hauptsatz auftretende Differenz $F(b) - F(a)$ schreiben wir im folgenden oft in der Form

$$F \Big|_a^b, \quad F(x) \Big|_a^b \quad \text{oder} \quad F(x) \Big|_{x=a}^{x=b}.$$

Beispiel. Es ist $\arctan' x = \frac{1}{1+x^2}$ für $x \in \mathbb{R}$ und damit

$$\int_0^1 \frac{dx}{1+x^2} = \arctan \Big|_0^1 = \arctan 1 - \arctan 0 = \frac{\pi}{4}.$$

Man beachte, daß aus der Integrierbarkeit nicht die Existenz einer Stammfunktion folgt, und aus der Existenz einer Stammfunktion auch nicht die Integrierbarkeit.

Beispiele. 1. Sei $f : [-1, 1] \rightarrow \mathbb{R}$,

$$f(x) = \begin{cases} -1 & \text{falls } -1 \leq x < 0, \\ 1 & \text{falls } 0 \leq x \leq 1. \end{cases}$$

Dann ist f integrierbar und $\int_{-1}^1 f = \int_{-1}^0 f + \int_0^1 f = -1 + 1 = 0$. Wäre F Stammfunktion, so wäre nach dem Hauptsatz $F(1) - F(-1) = \int_{-1}^1 f = 0$, nach dem Satz von Rolle würde also $\xi \in (-1, 1)$ mit $f(\xi) = F'(\xi) = 0$ existieren, ein Widerspruch. Also hat f keine Stammfunktion.

2. Sei $F : [-1, 1] \rightarrow \mathbb{R}$,

$$F(x) = \begin{cases} x^2 \sin \frac{1}{x^2} & \text{falls } x \neq 0, \\ 0 & \text{falls } x = 0 \end{cases}$$

Dann ist F differenzierbar (vgl. das Beispiel in §4.1) mit

$$F'(x) = \begin{cases} 2x \sin \frac{1}{x^2} - \frac{2}{x} \cos \frac{1}{x^2} & \text{falls } x \neq 0, \\ 0 & \text{falls } x = 0 \end{cases}$$

Setzt man $f := F'$, so ist also F Stammfunktion von f . Es ist aber f nicht integrierbar, da $f\left(\frac{1}{\sqrt{2\pi k}}\right) = 2\sqrt{2\pi k} \rightarrow \infty$ für $k \rightarrow \infty$, $k \in \mathbb{N}$, und damit f unbeschränkt ist.

Der Hauptsatz reduziert also die Berechnung von Integralen auf das Auffinden von Stammfunktionen. Statt “ F ist Stammfunktion von f ” schreibt man auch

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

Da für konstantes C mit F auch $F+C$ Stammfunktion ist, schreibt man manchmal auch

$$\int f(x)dx = F(x) + C.$$

Man nennt Stammfunktionen auch *unbestimmte Integrale*.

Obige Schreibweisen sind mit Vorsicht zu benutzen! So gilt etwa

$$\int \frac{dx}{1+x^2} = \arctan x \quad \text{und} \quad \int \frac{dx}{1+x^2} = -\operatorname{arccot} x,$$

aber es gilt $\arctan x \neq \operatorname{arccot} x$. (Es ist $\arctan x + \operatorname{arccot} x = \frac{\pi}{2}$.)

Falls nichts anderes angegeben ist, gelten die in den folgenden *Beispielen* angegebenen Stammfunktionen in jedem Intervall, in denen der Integrand definiert ist.

- $\int x^\alpha dx = \frac{x^{\alpha+1}}{\alpha+1}$ für $\alpha \in \mathbb{C}$, $\alpha \neq -1$,
- $\int \frac{1}{x} dx = \ln|x|$,
- $\int e^x dx = e^x$,
- $\int \cos x dx = \sin x$,

- $\int \sin x \, dx = -\cos x,$
- $\int \frac{dx}{1+x^2} = \arctan x,$
- $\int \frac{dx}{1-x^2} = \frac{1}{2} \ln \left| \frac{1+x}{1-x} \right|,$
- $\int \frac{dx}{\sqrt{1-x^2}} = \arcsin x,$
- $\int \frac{dx}{\cos^2 x} = \tan x,$
- $\int \frac{dx}{\sin^2 x} = -\cot x,$
- $\int \frac{dx}{\sqrt{1+x^2}} = \ln \left(x + \sqrt{1+x^2} \right),$
- $\int \frac{dx}{\sqrt{x^2-1}} = \ln \left(x + \sqrt{x^2-1} \right)$ für $x \in (1, \infty).$

Die in den letzten beiden Beispielen angegebenen Stammfunktionen sind Umkehrfunktionen der hyperbolischen Funktionen \sinh und \cosh , d. h., die durch $\operatorname{arsinh} : \mathbb{R} \rightarrow \mathbb{R}, x \rightarrow \ln \left(x + \sqrt{1+x^2} \right)$, und $\operatorname{arcosh} : [1, \infty) \rightarrow \mathbb{R}, x \rightarrow \ln \left(x + \sqrt{x^2-1} \right)$, definierten Funktionen erfüllen $\operatorname{arsinh}(\sinh x) = x$ und $\operatorname{arcosh}(\cosh x) = x$ für alle $x \in \mathbb{R}$. Sie heißen *area sinus (bzw. cosinus) hyperbolicus*. (Die geeignet definierten Umkehrfunktionen von \tanh und \coth heißen *area tangens (bzw. cotangens) hyperbolicus*.) Im Intervall $(-\infty, -1)$ gilt

- $\int \frac{dx}{\sqrt{x^2-1}} = -\operatorname{arcosh}(-x).$

5.3 Integrationstechniken

Kennt man die Ableitungen zweier Funktionen f und g , so kann man die von $f \cdot g$, f/g , $f \circ g$, usw. daraus berechnen. Entsprechendes gilt *nicht* für die Berechnung von Stammfunktionen. Dennoch gibt es einige nützliche Regeln.

Satz 5.3.1 (Partielle Integration) Seien $f, g : [a, b] \rightarrow \mathbb{C}$ stetig differenzierbar. Dann gilt

$$\int_a^b f g' = f g \Big|_a^b - \int_a^b f' g.$$

Beweis. Sei $F : [a, b] \rightarrow \mathbb{C}, F(x) = f g \Big|_a^x - \int_a^x f' g$. Nach Produktregel und Hauptsatz, Teil (ii), folgt dann $F' = f' g + f g' - f' g = f g'$. Die Behauptung folgt jetzt mit dem Hauptsatz, Teil (i). \square

Man schreibt die Regel der partiellen Integration auch als

$$\int f(x) g(x) dx = f(x) g(x) - \int f'(x) g(x) dx.$$

Beispiele. 1. Für $a \in \mathbb{R}$ ist

$$\int_0^a x e^x dx = x e^x \Big|_0^a - \int_0^a e^x dx = a e^a - e^x \Big|_0^a = a e^a - e^a + 1.$$

Es gilt also

$$\int x e^x dx = x e^x - e^x = (x - 1)e^x.$$

(Die additive Konstante 1 ist hier weggelassen.)

2. Für $x > 0$ ist

$$\int_1^x \ln t dt = \int_1^x \underbrace{1}_{g'} \cdot \underbrace{\ln t}_f dt = t \cdot \ln t \Big|_1^x - \int_1^x t \cdot \frac{1}{t} dt = x \ln x - x + 1.$$

Wiederum schreibt man auch $\int \ln x = x \ln x - x = x(\ln x - 1)$.

Neben der partiellen Integration ist die wichtigste Integrationsmethode die Substitution.

Satz 5.3.2 (Substitutionsregel) Sei $g : [a, b] \rightarrow [c, d]$ stetig differenzierbar und $f : [a, b] \rightarrow \mathbb{C}$ stetig. Dann gilt

$$\int_a^b f(g(t))g'(t)dt = \int_{g(a)}^{g(b)} f(x)dx.$$

Beweis. Nach dem Hauptsatz hat f eine Stammfunktion F . Nach Kettenregel ist $(F \circ g)'(t) = F'(g(t))g'(t) = f(g(t))g'(t)$ und damit ist $F \circ g$ Stammfunktion von $(f \circ g) \cdot g'$. Mit dem Hauptsatz folgt

$$\int_a^b (f \circ g) \cdot g' = F \circ g \Big|_a^b = F \Big|_{g(a)}^{g(b)} = \int_{g(a)}^{g(b)} f. \quad \square$$

Die Substitutionsregel kann man sich merken, indem man $x = g(t)$ setzt. Dann ist $\frac{dx}{dt} = g'(t)$.

Beispiele. 1. Es ist

$$\int_0^2 \frac{t dt}{1+t^4} \stackrel{x=t^2}{=} \frac{1}{2} \int_0^4 \frac{dx}{1+x^2} = \frac{1}{2} \arctan 4.$$

2. Es ist

$$\int_1^e \frac{\sqrt{1+\ln t}}{t} dt \stackrel{x=\ln t}{=} \int_0^1 \sqrt{1+x} dx \stackrel{y=1+x}{=} \int_1^2 \sqrt{y} dy = \frac{2}{3} y^{\frac{3}{2}} \Big|_1^2 = \frac{2}{3} (2\sqrt{2} - 1).$$

3. Wir betrachten für $x \in [-1, 1]$ das Integral

$$\int_0^x \sqrt{1-y^2} dy.$$

Mit $y = \sin t$, $t = \arcsin y$, wobei $y \in [-1, 1]$, $t \in [-\frac{\pi}{2}, \frac{\pi}{2}]$, ist dann $\sqrt{1-y^2} = \sqrt{1-\sin^2 t} = \sqrt{\cos^2 t} = |\cos t| = \cos t$. Die Substitutionsregel (mit $g(t) = \sin t$, $g'(t) = \cos t$) liefert also

$$\int_0^s \cos^2 t dt = \int_0^{\sin s} \sqrt{1-y^2} dy$$

für $s \in [-1, 1]$. Es folgt also

$$\begin{aligned}
\int_0^x \sqrt{1-y^2} dy &= \int_0^{\arcsin x} \cos^2 t dt \\
&= \int_0^{\arcsin x} \frac{1}{2}(1 + \cos 2t) dt \\
&= \frac{1}{2} \left(t + \frac{\sin 2t}{2} \right) \Big|_0^{\arcsin x} \\
&= \frac{1}{2} (t + \sin t \cos t) \Big|_0^{\arcsin x} \\
&= \frac{1}{2} \left(\arcsin x + x\sqrt{1-x^2} \right).
\end{aligned}$$

Wir beschreiben jetzt die *Integration rationaler Funktionen*, d. h., Funktionen f der Form $f = \frac{P}{Q}$ mit teilerfremden Polynomen P, Q . Wir dürfen $\text{grad}(P) < q := \text{grad}(Q)$ annehmen. (Andernfalls führen wir zunächst eine Polynomdivision durch.) Nach dem Fundamentalsatz der Algebra (Satz 1.9.2) existieren $c \in \mathbb{C}$, $k \in \mathbb{N}$, $a_1, \dots, a_k \in \mathbb{C}$ und $m_1, \dots, m_k \in \mathbb{N}$ mit

$$Q(z) = \prod_{j=1}^k (z - a_j)^{m_j}$$

für $z \in \mathbb{C}$. Es gilt dann $\sum_{j=1}^k m_j = q$. Wir dürfen $c = 1$ annehmen. Wir setzen nun die Koeffizienten von Q als reell voraus. Dann gilt $\overline{Q(z)} = Q(\bar{z})$ und damit ist mit a auch \bar{a} Nullstelle von Q . Wir fassen in dem Produkt für Q nun konjugiert komplexe Nullstellen zusammen und erhalten wegen

$$(z - a)(z - \bar{a}) = z^2 - (a + \bar{a})z + a\bar{a} = z^2 - 2 \operatorname{Re} a \cdot z + |a|^2$$

für Q dann eine Darstellung

$$Q(x) = \prod_{j=1}^M (x - a_j)^{m_j} \prod_{j=1}^N (x^2 + b_j x + c_j)^{n_j}$$

mit $a_j, b_j, c_j \in \mathbb{R}$, $c_j > b_j^2/4$, $m_j, n_j \in \mathbb{N}$ und $\sum_{j=1}^M m_j + 2 \sum_{j=1}^N n_j = q$.

Der *Satz von der Partialbruchzerlegung* besagt nun, daß $A_{i,j} \in \mathbb{R}$, $i = 1, \dots, M$, $j = 1, \dots, m_i$, und $B_{i,j}, C_{i,j} \in \mathbb{R}$, $i = 1, \dots, N$, $j = 1, \dots, n_i$, existieren, so daß

$$\begin{aligned}
\frac{P(x)}{Q(x)} &= \frac{A_{1,1}}{x - a_1} + \frac{A_{1,2}}{(x - a_1)^2} + \dots + \frac{A_{1,m_1}}{(x - a_1)^{m_1}} \\
&+ \frac{A_{2,1}}{x - a_2} + \frac{A_{2,2}}{(x - a_2)^2} + \dots + \frac{A_{2,m_2}}{(x - a_2)^{m_2}} \\
&+ \dots \\
&+ \frac{A_{M,1}}{x - a_M} + \frac{A_{M,2}}{(x - a_M)^2} + \dots + \frac{A_{M,m_M}}{(x - a_M)^{m_M}} \\
&+ \frac{B_{1,1}x + C_{1,1}}{x^2 + b_1x + c_1} + \frac{B_{1,2}x + C_{1,2}}{(x^2 + b_1x + c_1)^2} + \dots + \frac{B_{1,n_1}x + C_{1,n_1}}{(x^2 + b_1x + c_1)^{n_1}} \\
&+ \dots \\
&+ \frac{B_{N,1}x + C_{N,1}}{x^2 + b_Nx + c_N} + \frac{B_{N,2}x + C_{N,2}}{(x^2 + b_Nx + c_N)^2} + \dots + \frac{B_{N,n_N}x + C_{N,n_N}}{(x^2 + b_Nx + c_N)^{n_N}}.
\end{aligned}$$

Der *Beweis* läßt sich durch vollständige Induktion über q führen; hier nicht. Die Koeffizienten $A_{i,j}, B_{i,j}, C_{i,j}$ kann man durch Multiplikation mit $Q(x)$ und Koeffizientenvergleich bestimmen.

Die in der Partialbruchzerlegung rechts stehenden Terme lassen sich nun integrieren. Es ist zunächst

$$\int \frac{A}{x-a} dx = A \ln|x-a|$$

und

$$\int \frac{A}{(x-a)^m} dx = -\frac{A}{(m-1)(x-a)^{m-1}}$$

für $m \geq 2$. Wegen $x^2 + bx + c = \left(x + \frac{b}{2}\right)^2 + c - \frac{b^2}{4}$ lassen sich die Terme mit $(x^2 + bx + c)^n$ im Nenner durch eine Substitution auf die Form

$$\frac{\beta x + \gamma}{(x^2 + \alpha^2)^n}$$

bringen. Nun ist

$$\int \frac{\beta x}{x^2 + \alpha^2} dx = \frac{\beta}{2} \ln(x^2 + \alpha^2)$$

und

$$\int \frac{\beta x}{(x^2 + \alpha^2)^n} dx = -\frac{\beta}{2} \frac{1}{(n-1)(x^2 + \alpha^2)^{n-1}}$$

für $n \geq 2$. Schließlich ist noch

$$\int \frac{dx}{x^2 + \alpha^2} = \frac{1}{\alpha} \arctan \frac{x}{\alpha}$$

und für $n \geq 2$ zeigt partielle Integration, daß

$$\begin{aligned} \int \frac{dx}{(x^2 + \alpha^2)^n} &= \frac{1}{\alpha^2} \int \frac{\alpha^2 + x^2 - x^2}{(x^2 + \alpha^2)^n} dx \\ &= \frac{1}{\alpha^2} \int \frac{dx}{(x^2 + \alpha^2)^{n-1}} - \frac{1}{\alpha^2} \int \underbrace{x}_f \cdot \underbrace{\frac{x}{(x^2 + \alpha^2)^n}}_{g'} dx \\ &= \frac{1}{\alpha^2} \int \frac{dx}{(x^2 + \alpha^2)^{n-1}} - \frac{1}{\alpha^2} \left[x \left(-\frac{1}{2(n-1)(x^2 + \alpha^2)^{n-1}} \right) \right. \\ &\quad \left. + \int \frac{dx}{2(n-1)(x^2 + \alpha^2)^{n-1}} \right]. \end{aligned}$$

Damit können die Integrale

$$\int \frac{dx}{(x^2 + \alpha^2)^n}$$

rekursiv berechnet werden.

Beispiel. Man berechne

$$\int_1^2 \frac{x^4 + 1}{x^3 + x} dx.$$

Zunächst ist

$$f(x) := \frac{x^4 + 1}{x^3 + x} = x - \frac{x^2 - 1}{x^3 + x} = x - \frac{x^2 - 1}{x(x^2 + 1)}.$$

Es existieren nun $A, B, C \in \mathbb{R}$ mit

$$\frac{x^2 - 1}{x(x^2 + 1)} = \frac{A}{x} + \frac{Bx + C}{x^2 + 1}.$$

Zur Bestimmung der Konstanten A, B, C kann man natürlich mit dem Nenner durchmultiplizieren und anschließend Koeffizientenvergleich machen, was auf ein lineares Gleichungssystem für A, B, C führt. Einfacher ist, obige Gleichung mit x zu multiplizieren und anschließend $x = 0$ zu setzen. Dies führt sofort auf $A = -1$ und damit auf

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

Dies ergibt

$$\begin{aligned} \int_1^2 f(x) dx &= \int_1^2 \left(x + \frac{1}{x} - \frac{2x}{x^2 + 1} \right) dx \\ &= \left. \frac{x^2}{2} + \ln x - \ln(x^2 + 1) \right|_1^2 \\ &= 2 + \ln 2 - \ln 5 - \frac{1}{2} + \ln 2 \\ &= \frac{3}{2} + 2 \ln 2 - \ln 5 \quad \left(= \frac{3}{2} + \ln \frac{4}{5} = 1,276 \dots \right) \end{aligned}$$

Viele Integrale lassen sich durch geeignete Substitutionen auf Integrale über rationale Funktionen zurückführen. Sei etwa R eine rationale Funktionen von zwei Veränderlichen, d. h., $R(x, y) = P(x, y)/Q(x, y)$, wobei P, Q bei festem x Polynome in y und bei festem y Polynome in x sind. Dann kann

$$\int R(\sin x, \cos x) dx$$

durch die Substitution $t = \tan \frac{x}{2}$, $x = 2 \arctan t$, in ein Integral über eine rationale Funktion überführt werden. Die Rechnungen sind allerdings aufwendig.

Auch Integrale der Form

$$\int R(x, \sqrt{ax^2 + bx + c}) dx$$

lassen sich durch geeignete Substitutionen in Integrale über rationale Funktionen umformen.

5.4 Integralform des Taylorrestglieds

Mit partieller Integration läßt sich auch eine Formel für das Restglied der Taylorentwicklung gewinnen. Für eine n -mal differenzierbare Funktion $f : I \rightarrow \mathbb{R}$ und $\xi, x \in I$ hatten wir (siehe §4.4)

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

als Taylorpolynom (vom Grad n) bezeichnet. Für das *Restglied*

$$R_n(x) := f(x) - T_n(x)$$

hatten wir für $(n + 1)$ -mal differenzierbares f die Darstellung

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

mit η zwischen ξ und x bewiesen (Satz 4.4.3). Man nennt dies auch die *Lagrangesche Form des Restglieds*. Wir leiten noch eine *Integralform des Restglieds* her.

Wir bezeichnen für ein Intervall I und $n \in \mathbb{N}$ die Menge der n -mal stetig differenzierbaren Funktionen von I nach \mathbb{C} mit $C^n(I)$. Mit $C(I)$ bezeichnen wir die Menge der stetigen (englisch: continuous) Funktionen von I nach \mathbb{C} . Statt $C^n([a, b])$ bzw. $C([a, b])$ schreiben wir $C^n[a, b]$ bzw. $C[a, b]$.

Satz 5.4.1 Für $f \in C^{n+1}(I)$ und $\xi, x \in I$ gilt

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

Beweis. Es ist

$$\begin{aligned} f(x) &= f(\xi) + \int_{\xi}^x f'(t) dt \\ &= f(\xi) + f'(t)(t - x) \Big|_{t=\xi}^{t=x} - \int_{\xi}^x f''(t)(t - x) dt \\ &= f(\xi) + f'(\xi)(x - \xi) + \int_{\xi}^x f''(t)(x - t) dt \\ &= f(\xi) + f'(\xi)(x - \xi) + \frac{1}{2} f''(\xi)(x - \xi)^2 + \frac{1}{2} \int_{\xi}^x f'''(t)(x - t)^2 dt \\ &= \dots, \end{aligned}$$

und die Behauptung folgt durch vollständige Induktion. \square

Ist f reellwertig, so erhält man mit dem erweiterten Mittelwertsatz der Integralrechnung (Satz 5.1.10) aus der Integralform die Lagrangesche Form des Restglieds. (In Satz 4.4.3 wurde allerdings nicht die Stetigkeit von $f^{(n+1)}$ vorausgesetzt.)

In §4.4 hatten wir das Beispiel $f(x) = \ln x$, $\xi = 1$, untersucht und gezeigt, daß $\lim_{n \rightarrow \infty} R_n(x) = 0$ für $\frac{1}{2} \leq x \leq 2$ gilt. Mit Satz 5.4.1 können wir zeigen, daß dies sogar für $0 < x \leq 2$ gilt, denn für $0 < x < 1$ ist

$$\begin{aligned} |R_n(x)| &= \left| \frac{1}{n!} \int_x^1 (x - t)^n f^{(n+1)}(t) dt \right| \\ &= \int_x^1 \frac{(t - x)^n}{t^{n+1}} dt \\ &= \int_x^1 \left(1 - \frac{x}{t}\right)^n \frac{dt}{t} \\ &\leq (1 - x)^n \int_x^1 \frac{dt}{t} \\ &= (1 - x)^n \ln \frac{1}{x}. \end{aligned}$$

Für $f \in C^{n+1}(I)$ folgt aus Satz 5.4.1, und für reellwertiges f auch aus Satz 4.4.3, daß eine Umgebung U von ξ und eine Konstante $K \in \mathbb{R}_+$ mit

$$|R_n(x)| \leq K|x - \xi|^{n+1}$$

für $x \in U$ existiert.

Nun kann T_n und R_n bereits für n -mal differenzierbares f definiert werden. Dann erhalten wir aber nur schlechtere Abschätzungen.

Satz 5.4.2 Für $f \in C^n(I)$ und ξ, x in I gilt

$$\lim_{x \rightarrow \xi} \frac{R_n(x)}{(x - \xi)^n} = 0.$$

Beweis. Es gilt

$$\begin{aligned} |R_n(x)| &= |f(x) - T_n(x)| \\ &= \left| f(x) - T_{n-1}(x) - \frac{1}{n!} f^{(n)}(\xi)(x - \xi)^n \right| \\ &= \left| \frac{1}{(n-1)!} \int_{\xi}^x f^{(n)}(t)(x-t)^{n-1} dt - \frac{1}{n!} f^{(n)}(\xi)(x - \xi)^n \right| \\ &= \left| \frac{1}{(n-1)!} \int_{\xi}^x (f^{(n)}(t) - f^{(n)}(\xi))(x-t)^{n-1} dt \right| \\ &\leq |x - \xi|^n \frac{1}{(n-1)!} \sup_t |f^{(n)}(t) - f^{(n)}(\xi)|, \end{aligned}$$

wobei das Supremum über alle t zwischen ξ und x zu nehmen ist. Die Behauptung folgt. \square

Für reellwertiges f erhalten wir die Behauptung von Satz 5.4.2 auch direkt aus der Lagrangeschen Form des Restglieds:

$$R_n(x) = R_{n-1}(x) - \frac{1}{n!} f^{(n)}(\xi)(x - \xi)^n = \frac{1}{n!} (f^{(n)}(\eta) - f^{(n)}(\xi))(x - \xi)^n.$$

Die obigen Restgliedabschätzungen lassen sich gut mit den sogenannten *Landau-Symbolen* beschreiben. Ist $M \subset \mathbb{C}$ und ξ Häufungspunkt von M , wobei für $M \subset \mathbb{R}$ auch $\xi = \pm\infty$ zugelassen ist, und ist $\alpha : M \rightarrow \mathbb{C}$ und $\beta : M \rightarrow \mathbb{R}_+$, so schreibt man

$$\alpha(x) = o(\beta(x)) \text{ für } x \rightarrow \xi, \text{ falls } \lim_{x \rightarrow \xi} \frac{\alpha(x)}{\beta(x)} = 0,$$

und

$$\alpha(x) = O(\beta(x)) \text{ für } x \rightarrow \xi, \text{ falls } \limsup_{x \rightarrow \xi} \frac{\alpha(x)}{\beta(x)} < \infty.$$

Dabei läßt man den Zusatz “für $x \rightarrow \xi$ ” oft weg, wenn dieser aus dem Zusammenhang hervorgeht. Offensichtlich gilt $\alpha(x) = O(\beta(x))$ für $x \rightarrow \xi$ genau dann, wenn eine Konstante $K \in \mathbb{R}_+$ und eine Umgebung U von ξ existieren, so daß $\frac{|f(x)|}{g(x)} \leq K$ für alle $x \in U$ gilt.

Für $\xi, x \in I$ und $f : I \rightarrow \mathbb{C}$ (mit einem Intervall I) erhalten wir damit:

$$\begin{aligned} f \in C^{n+1}(I) &\Rightarrow R_n(x) = O(|x - \xi|^{n+1}) \text{ für } x \rightarrow \xi, \\ f \in C^n(I) &\Rightarrow R_n(x) = o(|x - \xi|^n) \text{ für } x \rightarrow \xi. \end{aligned}$$

5.5 Folgen und Reihen integrierbarer Funktionen

In §3.6 bzw. §4.4 hatten wir untersucht, wann sich bei einer konvergenten Funktionenfolge aus der Stetigkeit bzw. Differenzierbarkeit der Folgenglieder die der Grenzfunktion folgern läßt. Wir betrachten hier das entsprechende Problem für Integrierbarkeit.

Satz 5.5.1 *Sei I kompaktes Intervall und (f_n) eine Folge über I integrierbarer Funktionen, die gleichmäßig gegen eine Funktion f konvergiert. Dann ist f über I integrierbar und es gilt $\lim_{n \rightarrow \infty} \int_I f_n = \int_I f$.*

Beweis. Ist bereits bekannt, daß f integrierbar ist, so folgt $\lim_{n \rightarrow \infty} \int_I f_n = \int_I f$ sehr leicht. Denn ist $\varepsilon > 0$, so existiert $N \in \mathbb{N}$ mit $\sup_{x \in I} |f_n(x) - f(x)| < \frac{\varepsilon}{|I|}$ für alle $n \geq N$. Es folgt $|\int_I f_n - \int_I f| = |\int_I (f_n - f)| \leq |I| \sup_{x \in I} |f_n(x) - f(x)| < \varepsilon$ und damit die Behauptung.

Zu zeigen ist also lediglich die Integrierbarkeit von f . Diese zeigen wir mit dem Cauchy Kriterium (Satz 5.1.1). Sei dazu $\varepsilon > 0$. Dann existiert $N \in \mathbb{N}$ mit $\sup_{x \in I} |f_n(x) - f(x)| < \frac{\varepsilon}{3|I|}$ für $n \geq N$. Für jede Zerlegung Z von I mit Stützstellen ξ folgt dann

$$|S(f_N, Z, \xi) - S(f, Z, \xi)| = |S(f_N - f, Z, \xi)| < \frac{\varepsilon}{3|I|}|I|.$$

Da f_N integrierbar existiert $\delta > 0$, so daß für alle Zerlegungen Z_1, Z_2 mit Stützstellen ξ_1, ξ_2 gilt: ist $|Z_1| < \delta$ und $|Z_2| < \delta$, so ist

$$|S(f_N, Z_1, \xi_1) - S(f_N, Z_2, \xi_2)| < \frac{\varepsilon}{3}.$$

Für $|Z_1| < \delta$ und $|Z_2| < \delta$ folgt damit

$$\begin{aligned} & |S(f, Z_1, \xi_1) - S(f, Z_2, \xi_2)| \\ & \leq |S(f, Z_1, \xi_1) - S(f_N, Z_1, \xi_1)| + |S(f_N, Z_1, \xi_1) - S(f_N, Z_2, \xi_2)| \\ & \quad + |S(f_N, Z_2, \xi_2) - S(f, Z_2, \xi_2)| \\ & < \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} \\ & = \varepsilon. \quad \square \end{aligned}$$

Natürlich gilt auch wieder eine Analogon von Satz 5.5.1 für Reihen.

5.6 Uneigentliche Integrale

Bei der Definition von $\int_I f$ haben wir vorausgesetzt, daß I kompakt ist, und aus der Existenz des Integrals konnte die Beschränktheit von f gefolgert werden (Satz 5.1.3). Wir werden jetzt $\int_I f$ in einigen Fällen definieren, in denen I nicht kompakt oder f nicht beschränkt ist.

Definition 5.6.1 Sei $-\infty < a < b \leq \infty$ und $f : [a, b) \rightarrow \mathbb{C}$. Für alle $c \in [a, b)$ sei f integrierbar über $[a, c]$. Existiert dann $\lim_{c \rightarrow b} \int_a^c f$, so heißt f *uneigentlich integrierbar* über $[a, b)$. Der Grenzwert heißt *uneigentliches Integral* von f über $[a, b)$ und wird mit $\int_a^b f$ bezeichnet.

Man sagt auch, daß das uneigentliche Integral *konvergiert* (bzw. *divergiert*). Analog definiert man für $-\infty \leq a < b < \infty$ und $f : (a, b] \rightarrow \mathbb{C}$ das uneigentliche Integral $\int_a^b f$. Im Falle $f : (a, b) \rightarrow \mathbb{C}$ sagt man, daß das uneigentliche Integral $\int_a^b f$ existiert, wenn für $c \in (a, b)$ die beiden uneigentlichen Integrale $\int_a^c f$ und $\int_c^b f$ existieren. Man setzt dann $\int_a^b f = \int_a^c f + \int_c^b f$. (Dies hängt nicht von c ab.)

Beispiele. 1. Für $\alpha \in \mathbb{R}$ und $r \in \mathbb{R}_+$ gilt

$$\int_1^r \frac{dx}{x^\alpha} = \begin{cases} \frac{1}{1-\alpha} (r^{1-\alpha} - 1) & \text{falls } \alpha \neq 1, \\ \ln r & \text{falls } \alpha = 1. \end{cases}$$

Es folgt, daß $\int_1^\infty \frac{dx}{x^\alpha}$ für $\alpha > 1$ konvergiert, mit $\int_1^\infty \frac{dx}{x^\alpha} = \frac{1}{1-\alpha}$, während $\int_1^\infty \frac{dx}{x^\alpha}$ für $\alpha \leq 1$ divergiert.

2. Wir können in obigem Beispiel auch den Grenzwert für $r \rightarrow 0$ betrachten und erhalten, daß $\int_0^1 \frac{dx}{x^\alpha}$ für $\alpha < 1$ konvergiert und den Wert $\frac{1}{\alpha-1}$ hat, während das Integral für $\alpha \geq 1$ divergiert.

3. Es ist

$$\begin{aligned} \int_{-\infty}^{\infty} \frac{dx}{1+x^2} &= \lim_{a \rightarrow -\infty} \int_a^0 \frac{dx}{1+x^2} + \lim_{b \rightarrow -\infty} \int_0^b \frac{dx}{1+x^2} \\ &= \lim_{a \rightarrow -\infty} (-\arctan a) + \lim_{b \rightarrow -\infty} \arctan b \\ &= \frac{\pi}{2} + \frac{\pi}{2} \\ &= \pi. \end{aligned}$$

4. Das uneigentliche Integral

$$\int_{-\infty}^{\infty} \frac{x dx}{1+x^2}$$

existiert nicht, denn

$$\lim_{r \rightarrow \infty} \int_0^r \frac{x dx}{1+x^2} = \lim_{r \rightarrow \infty} \frac{1}{2} \ln(1+r^2) = \infty,$$

womit

$$\int_0^{\infty} \frac{x dx}{1+x^2}$$

nicht existiert. (Auch

$$\int_{-\infty}^0 \frac{x dx}{1+x^2}$$

existiert nicht. Man beachte aber, daß

$$\lim_{r \rightarrow \infty} \int_{-r}^r \frac{x dx}{1+x^2} = 0$$

gilt.)

Satz 5.6.1 (Cauchy Kriterium für uneigentliche Integrale) Sei $-\infty < a < b < \infty$ und $f : [a, b) \rightarrow \mathbb{C}$. Für alle $c \in [a, b)$ sei f integrierbar über $[a, c]$. Dann konvergiert $\int_a^b f$ genau dann, wenn

$$\forall \varepsilon \in \mathbb{R}_+ \exists \delta \in \mathbb{R}_+ \forall x, y \in (b - \delta, b) : \left| \int_x^y f \right| < \varepsilon.$$

Beweis. Man wende das Cauchy Kriterium für Funktionengrenzwerte (Satz 3.5.4) auf $F(x) := \int_a^x f$ und beachte, daß $F(y) - F(x) = \int_x^y f$. \square

Ist $b = \infty$ in Satz 5.6.1, so lautet das Cauchy Kriterium: $\int_a^\infty f$ konvergiert genau dann, wenn

$$\forall \varepsilon \in \mathbb{R}_+ \exists r \in \mathbb{R}_+ \forall x, y \in (r, \infty) : \left| \int_x^y f \right| < \varepsilon.$$

Satz 5.6.2 (Vergleichskriterium für uneigentliche Integrale) Seien $-\infty < a < b \leq \infty$, $f : [a, b) \rightarrow \mathbb{C}$ und $g : [a, b) \rightarrow \mathbb{R}$. Für alle $c \in [a, b)$ seien f und g integrierbar über $[a, c]$ und es gelte $|f(x)| \leq g(x)$ für alle $x \in \mathbb{R}$. Konvergiert dann $\int_a^b g$, so konvergieren auch $\int_a^b |f|$ und $\int_a^b f$.

Beweis. Wegen

$$\left| \int_x^y f \right| \leq \int_x^y |f| \leq \int_x^y g$$

für alle $x, y \in (a, b)$ folgt die Behauptung direkt aus dem Cauchy Kriterium. \square

Durch Kontraposition kann man mit dem Vergleichskriterium natürlich auch Divergenz nachweisen.

Die Sätze 5.6.1 und 5.6.2 und die Bemerkungen dazu gelten analog für $f : (a, b] \rightarrow \mathbb{C}$, $-\infty \leq a < b < \infty$.

Beispiel. Wegen $|\sin t| \leq 1$ für $t \in \mathbb{R}$ konvergiert

$$\int_1^\infty \frac{\sin t}{t^\alpha} dt$$

nach Vergleichskriterium und Beispiel 1 nach Definition 5.6.1 für $\alpha > 1$.

Tatsächlich konvergiert das Integral sogar für $\alpha > 0$, denn für $y > x > (2\varepsilon)^{-1/\alpha} > 1$ ist

$$\begin{aligned} \left| \int_x^y \frac{\sin t}{t^\alpha} dt \right| &\stackrel{\text{part. Int.}}{=} \left| -\frac{\cos t}{t^\alpha} \Big|_x^y - \alpha \int_x^y \frac{\cos t}{t^{\alpha+1}} dt \right| \\ &\leq \frac{|\cos y|}{y^\alpha} + \frac{|\cos x|}{x^\alpha} + \alpha \int_x^y \frac{|\cos t|}{t^{\alpha+1}} dt \\ &\leq \frac{1}{y^\alpha} + \frac{1}{x^\alpha} + \alpha \int_x^y \frac{dt}{t^{\alpha+1}} \\ &= \frac{2}{x^\alpha} \\ &< \varepsilon, \end{aligned}$$

und die Konvergenz folgt mit dem Cauchy Kriterium. Man kann hieraus folgern, daß

$$\int_0^\infty \sin t^2 dt$$

konvergiert, denn für $0 < x < y < \infty$ folgt mit der Substitutionsregel

$$\int_x^y \sin t^2 dt = \frac{1}{2} \int_{\sqrt{x}}^{\sqrt{y}} \frac{\sin s}{\sqrt{s}} ds.$$

Satz 5.6.3 (Integralkriterium für Reihen) Sei $f : [1, \infty) \rightarrow \mathbb{R}_+$ monoton fallend. Dann konvergiert $\int_1^\infty f(x) dx$ genau dann, wenn $\sum_{k=1}^\infty f(k)$ konvergiert.

Beweis. Nach Satz 5.1.11 ist f über kompakte Teilintervalle von $[1, \infty)$ integrierbar. Für $k \in \mathbb{N}$ gilt

$$f(k+1) \leq \int_k^{k+1} f(x) dx \leq f(k)$$

und damit folgt

$$\sum_{k=m+1}^{n+1} f(k) \leq \int_m^n f(x) dx \leq \sum_{k=m}^n f(k)$$

für $m, n \in \mathbb{N}$, $n > m$. Die Behauptung folgt aus dem Cauchy Kriterium. \square

Beispiel. 1. Für $\alpha > 1$ konvergiert

$$\sum_{k=2}^\infty \frac{1}{k(\ln k)^\alpha},$$

da

$$\int_2^r \frac{dx}{x(\ln x)^\alpha} = -\frac{1}{(\alpha+1)(\ln x)^{\alpha+1}} \Big|_2^r.$$

2. Die Reihe

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

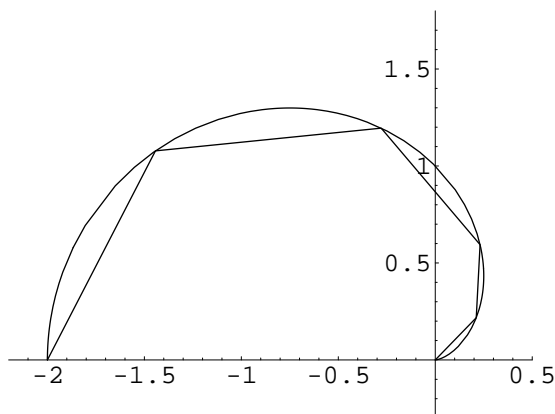
divergiert, da

$$\int_2^r \frac{dx}{x \ln x} = \ln \ln x \Big|_2^r.$$

5.7 Bogenlänge ebener Kurven

Bisher haben wir Integrale als Flächeninhalt interpretiert. Wir können Integrale auch zur Berechnung von Kurvenlängen benutzen. Dabei ist eine (ebene) Kurve eine stetige Abbildung eines Intervalls nach $\mathbb{R}^2 = \mathbb{C}$. Die Idee ist, Kurven durch Streckenzüge zu approximieren.

Als *Beispiel* betrachten wir die durch $f : [0, \pi] \rightarrow \mathbb{C}$, $t \mapsto (1 - \cos t)e^{it}$ gegebene Kurve:



Betrachtet man obige Funktion im Intervall $[0, 2\pi]$, erhält man die sogenannte *Kardioide*.

Definition 5.7.1 Sei I kompaktes Intervall und $f : I \rightarrow \mathbb{C}$ stetig. Für eine Zerlegung $Z = \{x_0, x_1, \dots, x_n\}$ von I sei

$$L(f, Z) := \sum_{k=1}^n |f(x_k) - f(x_{k-1})|$$

und es sei

$$L(f) := \sup L(f, Z),$$

wobei das Supremum über alle Zerlegungen Z von I zu nehmen ist. Ist $L(f) < \infty$, so heißt f *rektifizierbar* und $L(f)$ heißt *Länge* von f .

Satz 5.7.1 Sei I kompaktes Intervall und $f : I \rightarrow \mathbb{C}$ stetig differenzierbar. Dann ist f rektifizierbar und $L(f) = \int_I |f'|$.

Beweis. Sei $Z = \{x_0, x_1, \dots, x_n\}$ Zerlegung von I . Nach Satz 4.2.5 existieren $\xi_k \in (x_{k-1}, x_k)$ mit $|f(x_k) - f(x_{k-1})| \leq |f'(\xi_k)|(x_k - x_{k-1})$. Mit $\xi := (\xi_1, \dots, \xi_n)$ folgt $L(f, Z) \leq S(|f'|, Z, \xi)$ und damit $L(f) \leq \int_I |f'| < \infty$.

Damit ist f rektifizierbar und für $x, y \in I$ mit $x < y$ auch $f|[x, y]$ rektifizierbar mit $L(f|[x, y]) \leq \int_x^y |f'|$. Wir definieren $\ell : I \rightarrow \mathbb{R}$ durch $\ell(x) = L(f|[x_0, x])$. Dann gilt

$$|f(y) - f(x)| = L(f, \{x, y\}) \leq \ell(y) - \ell(x) = L(f|[x, y]) \leq \int_x^y |f'|$$

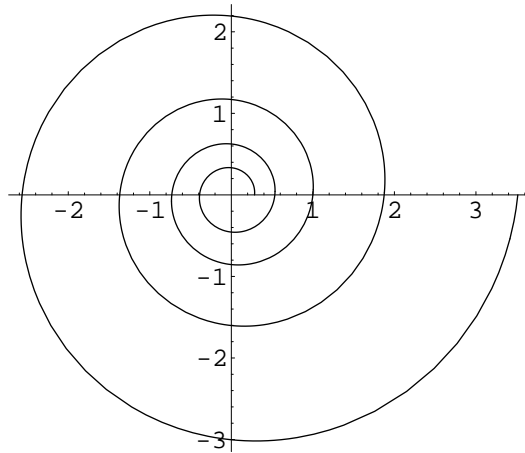
für $x, y \in I$ mit $x < y$. Division durch $y - x$ und Grenzübergang $x \rightarrow y$ bzw. $y \rightarrow x$ liefert, daß ℓ differenzierbar ist, mit $\ell' = |f'|$. Die Behauptung folgt aus dem Hauptsatz. \square

Beispiele. 1. Sei $c > 0$ und $f : [0, 2\pi] \rightarrow \mathbb{C}$, $f(t) = e^{ct}e^{it} = e^{(c+i)t}$. Dann gilt $f'(t) = (c+i)e^{(c+i)t}$ und $|f'(t)| = |c+i|e^{ct}|e^{it}| = \sqrt{c^2+1}e^{ct}$. Es folgt

$$L(f) = \int_0^{2\pi} \sqrt{c^2+1}e^{ct}dt = \sqrt{c^2+1} \frac{e^{ct}}{c} \Big|_0^{2\pi} = \frac{\sqrt{c^2+1}}{c} (e^{2\pi c} - 1).$$

Die durch f gegebene Kurve heißt *logarithmische Spirale*.

Für $c = 1/10$ sieht der durch $-2\pi \leq t \leq 4\pi$ gegebene Teil wie folgt aus:

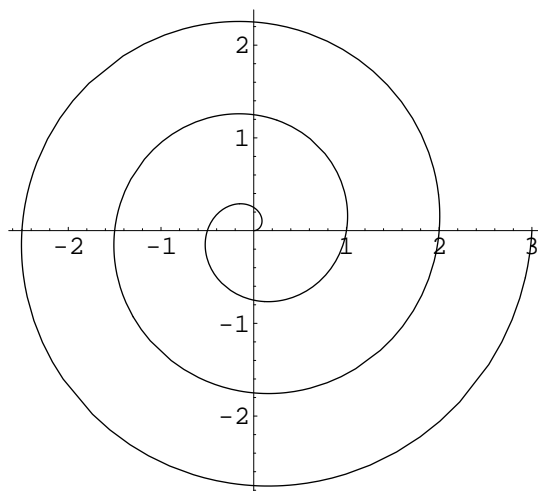


2. Sei $c > 0$ und $f : [0, 2\pi] \rightarrow \mathbb{C}$, $f(t) = cte^{it}$. Dann gilt $f'(t) = c(1 + it)e^{it}$ und $|f'(t)| = c|1 + it| = c\sqrt{1 + t^2}$. Es folgt

$$\begin{aligned} L(f) &= c \int_0^{2\pi} \sqrt{1 + t^2} dt \\ &= \frac{c}{2} \left(t\sqrt{1 + t^2} + \ln \left(t + \sqrt{1 + t^2} \right) \right) \Big|_0^{2\pi} \\ &= \frac{c}{2} \left(2\pi\sqrt{1 + 4\pi^2} + \ln \left(2\pi + \sqrt{1 + 4\pi^2} \right) \right). \end{aligned}$$

Die durch f gegebene Kurve heißt *Archimedische Spirale*.

Für $c = 1/2\pi$ sieht der durch $0 \leq t \leq 3$ gegebene Teil wie folgt aus:



3. Sei I kompaktes Intervall und $f : I \rightarrow \mathbb{R}$. Die Länge des Graphen von f ist die Länge der durch $t \mapsto (t, f(t)) = t + if(t)$ gegebenen Kurve. Für stetig differenzierbares f ist diese durch $\int_I \sqrt{1 + f'(t)^2} dt$ gegeben.

Für die Länge L des Graphen des durch $f : [0, 1] \rightarrow \mathbb{R}$, $f(x) = x^2/2$ Parabelstücks gilt also

$$L = \int_0^1 \sqrt{1+x^2} dx = \frac{1}{2} \left(\sqrt{2} + \ln(1 + \sqrt{2}) \right).$$

6 Topologie metrischer Räume

6.1 Metrik, Norm und Skalarprodukt

Bisher haben wir Funktionen $f : M \rightarrow N$ mit $M, N \subset \mathbb{R}, \mathbb{C}$ betrachtet. Jetzt werden wir allgemeinere Mengen M, N zulassen, etwa \mathbb{R}^n oder \mathbb{C}^n mit $n \in \mathbb{N}$, oder auch Mengen von Funktionen. Wichtige Beispiele sind hier $C^n(I)$ und $C(I)$ mit einem Intervall I . Die genannten Mengen sind Vektorräume. Wir benötigen aber noch weitere "Struktur" auf diesen Mengen.

Definition 6.1.1 Sei M nichtleere Menge. Eine Funktion $d : M \times M \rightarrow [0, \infty)$ heißt *Metrik*, wenn für alle $x, y, z \in M$ die folgenden drei Bedingungen erfüllt sind:

- (i) $d(x, y) = 0 \Leftrightarrow x = y$,
- (ii) $d(x, y) = d(y, x)$ (*Symmetrie*)
- (iii) $d(x, y) + d(y, z) \leq d(x, z)$ (*Dreiecksungleichung*)

Das Paar (M, d) heißt dann *metrischer Raum*.

Beispiele. 1. $M = \mathbb{R}$ oder $M = \mathbb{C}$, $d(x, y) = |x - y|$. Wir nennen d die *euklidische Metrik* und $d(x, y)$ den *euklidischen Abstand* von x und y .

2. $M \neq \emptyset$ beliebig,

$$d(x, y) = \begin{cases} 1 & \text{falls } x \neq y, \\ 0 & \text{falls } x = y. \end{cases}$$

Dies nennt man die *diskrete Metrik* auf M .

3. $M = \mathbb{C}$,

$$d(x, y) = \begin{cases} |x - y| & \text{falls } \lambda, \mu \in \mathbb{R} \text{ mit } \lambda x + \mu y = 0 \text{ existieren,} \\ |x| + |y| & \text{sonst.} \end{cases}$$

Liegen die Punkte x, y auf einer Geraden durch 0, so ist $d(x, y)$ also der euklidische Abstand. Andernfalls ist $d(x, y)$ die Summe der Abstände zu 0. (Aus naheliegende Gründen nennt man das auch "französische Eisenbahnmetrik".)

Weitere wichtige Beispiele metrischer Räume sind normierte Räume; vgl. LA I, Definition 1.46. Der Vollständigkeit halber wiederholen wir die Definition. Hier und im folgenden ist dabei immer $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$. (In LA I wurde nur der Fall $\mathbb{K} = \mathbb{R}$ betrachtet, aber ansonsten ist die Definition wörtlich die gleiche.)

Definition 6.1.2 Sei V ein Vektorraum über \mathbb{K} . Eine Funktion $\|\cdot\| : V \rightarrow [0, \infty)$ (also $x \mapsto \|x\|$) heißt *Norm*, wenn für alle $x, y \in V$ und $\lambda \in \mathbb{K}$ die folgenden drei Bedingungen erfüllt sind:

- (i) $\|x\| = 0 \Leftrightarrow x = 0$,
- (ii) $\|\lambda \cdot x\| = |\lambda| \cdot \|x\|$,
- (iii) $\|x + y\| \leq \|x\| + \|y\|$ (*Dreiecksungleichung*)

Das Paar $(V, \|\cdot\|)$ heißt dann *normierter Raum*.

Satz 6.1.1 Sei $(V, \|\cdot\|)$ normierter Raum. Dann ist die durch $d(x, y) := \|x - y\|$ definierte Funktion $d : V \times V \rightarrow [0, \infty)$ eine Metrik auf V , d. h., (V, d) ist metrischer Raum.

Beweis. Seien $x, y, z \in V$. Dann gilt

$$d(x, y) = 0 \Leftrightarrow \|x - y\| = 0 \Leftrightarrow x - y = 0 \Leftrightarrow x = y,$$

$$d(x, y) = \|x - y\| = \|(-1)(y - x)\| = |(-1)| \cdot \|(y - x)\| = \|(y - x)\| = d(y, x)$$

und

$$d(x, z) = \|x - z\| = \|x - y + y - z\| \leq \|x - y\| + \|y - z\| = d(x, y) + d(y, z). \quad \square$$

Beispiele. 1. $V = \mathbb{K}$, $\|x\| = |x|$ für $x \in V$.

2. $p \in \mathbb{N}$, $V = \mathbb{K}^p$, $\|x\|_\infty := \max_j |x_j|$ für $x = (x_1, \dots, x_p)^T \in V$. (Wie in LA I wollen wir Vektoren im K^p als Spaltenvektoren schreiben. Es steht "T" für "transponiert".) Es ist $(V, \|\cdot\|_\infty)$ normierter Raum.

3. $p \in \mathbb{N}$, $V = \mathbb{K}^p$, $\|x\|_1 := \sum_{j=1}^p |x_j|$ für $x = (x_1, \dots, x_p)^T \in V$. Dann ist $(V, \|\cdot\|_1)$ normierter Raum.

4. $V = C[a, b]$, $\|f\|_\infty := \max_{x \in [a, b]} |f(x)|$ für $f \in V$. Dann ist $(V, \|\cdot\|_\infty)$ normierter Raum. Die zugehörige Metrik d ist durch $d(f, g) := \max_{x \in [a, b]} |f(x) - g(x)|$ gegeben.

5. $V = C[a, b]$, $\|f\|_1 := \int_a^b |f|$ für $f \in V$. Dann ist $(V, \|\cdot\|_1)$ normierter Raum. Die zugehörige Metrik d_1 ist durch $d_1(f, g) := \int_a^b |f - g|$ gegeben.

Die einzige Bedingung in Definition 6.1.2, die vielleicht nicht offensichtlich ist, ist " \Leftarrow " in (i). Wir zeigen dies durch Kontraposition. Sei dazu $f \in C[a, b]$, $f \neq 0$. Dann existiert $\xi \in [a, b]$ mit $f(\xi) \neq 0$. Da f stetig, existiert $\delta > 0$ mit $|f(x) - f(\xi)| < \varepsilon := \frac{|f(\xi)|}{2}$ für $|x - \xi| < \delta$, $x \in [a, b]$. Mit $I := [a, b] \cap [\xi - \frac{\delta}{2}, \xi + \frac{\delta}{2}]$ folgt $|f(x)| \geq |f(\xi)| - |f(x) - f(\xi)| > |f(\xi)| - \varepsilon = \varepsilon$ für $x \in I$ und damit $\|f\|_1 = \int_a^b |f'| \geq \int_I |f'| \geq \varepsilon |I| > 0$.

6. Sei V der Vektorraum der Riemann-integrierbaren Funktionen von $[a, b]$ nach \mathbb{C} und sei $\|f\|_1 := \int_a^b |f|$ für $f \in V$. Dann ist $(V, \|\cdot\|_1)$ kein normierter Raum, denn Bedingung (i) aus Definition 6.1.2 gilt nicht, wie das Beispiel

$$f(x) = \begin{cases} 1 & \text{falls } x = a, \\ 0 & \text{falls } a < x \leq b, \end{cases}$$

zeigt.

Definition 6.1.3 Sei V ein Vektorraum über \mathbb{K} . Eine Funktion $\langle \cdot, \cdot \rangle : V \times V \rightarrow K$ (also $(x, y) \mapsto \langle x, y \rangle$) heißt *Skalarprodukt* (oder *inneres Produkt*), wenn für alle $x, y, z \in V$ und $\lambda \in \mathbb{K}$ die folgenden fünf Bedingungen erfüllt sind:

- (i) $\langle x, y \rangle = \overline{\langle y, x \rangle}$, wobei der Querstrich für komplexe Konjugation steht,
- (ii) $\langle x, x \rangle \geq 0$,
- (iii) $\langle x, x \rangle = 0 \Leftrightarrow x = 0$,
- (iv) $\langle \lambda x, y \rangle = \lambda \langle x, y \rangle$,
- (v) $\langle x + y, z \rangle = \langle x, z \rangle + \langle y, z \rangle$.

Das Paar $(V, \langle \cdot, \cdot \rangle)$ heißt dann *Prähilbertraum*.

Statt $\langle x, y \rangle$ sind auch die Schreibweisen $x \cdot y$ und (x, y) üblich.

Satz 6.1.2 Sei $(V, \langle \cdot, \cdot \rangle)$ Prähilbertraum. Dann ist durch $\|x\| := \sqrt{\langle x, x \rangle}$ eine Norm auf V definiert, d. h., $(V, \|\cdot\|)$ ist normierter Raum.

Wir stellen den Beweis von Satz 6.1.2 noch etwas zurück und ziehen zunächst einige einfache Folgerungen aus (i)-(v). Dabei seien wieder $x, y, z \in V$ und $\lambda \in \mathbb{K}$:

- $\langle x, \lambda y \rangle = \bar{\lambda} \langle x, y \rangle$, denn $\langle x, \lambda y \rangle = \overline{\langle \lambda y, x \rangle} = \overline{\lambda \langle y, x \rangle} = \bar{\lambda} \overline{\langle y, x \rangle} = \bar{\lambda} \langle x, y \rangle$,
- $\langle x, y + z \rangle = \langle x, y \rangle + \langle x, z \rangle$,
- $\langle 0, x \rangle = \langle x, 0 \rangle = 0$, denn $\langle 0, x \rangle = \langle 0 - 0, x \rangle = \langle 0, x \rangle - \langle 0, x \rangle = 0$.

In LA I, §1.8, wurde gezeigt, daß für $V = \mathbb{R}^p$ durch $\langle x, y \rangle := \sum_{j=1}^p x_j y_j$ ein Skalarprodukt gegeben ist. (Dabei ist $x = (x_1, \dots, x_p)^T, y = (y_1, \dots, y_p)^T \in V$.) Analog zeigt man, daß für $V = \mathbb{C}^p$ durch $\langle x, y \rangle := \sum_{j=1}^p x_j \bar{y}_j$ ein Skalarprodukt gegeben ist.

Satz 6.1.3 Sei $(V, \langle \cdot, \cdot \rangle)$ Prähilbertraum, x, y in V . Dann gilt (mit $\|\cdot\|$ wie in Satz 6.1.2)

- $\|x + y\|^2 = \|x\|^2 + \|y\|^2 + 2 \operatorname{Re} \langle x, y \rangle$ (Satz des Pythagoras).
- $\|x + y\|^2 + \|x - y\|^2 = 2\|x\|^2 + 2\|y\|^2$ (Parallelogrammidentität).
- $|\langle x, y \rangle|^2 \leq \|x\| \cdot \|y\|$ (Cauchy-Schwarz-Ungleichung).

Für $V = \mathbb{R}^p$ sind dies die Sätze 1.43 und 1.44 aus LA I. Für Vektorräume über \mathbb{R} kann man den dortigen Beweis wörtlich übernehmen. Für Vektorräume über \mathbb{C} sind nur geringfügige Modifikationen nötig.

Beweis von Satz 6.1.2. Seien $x, y \in V$ und $\lambda \in \mathbb{K}$. Dann gilt

$$\|x\| = 0 \Leftrightarrow \langle x, x \rangle = 0 \Leftrightarrow x = 0$$

und

$$\|\lambda x\| = \sqrt{\langle \lambda x, \lambda x \rangle} = \sqrt{\lambda \bar{\lambda} \langle x, x \rangle} = \sqrt{|\lambda|^2 \langle x, x \rangle} = |\lambda| \cdot \|x\|.$$

Die Dreiecksungleichung folgt (vgl. LA I, Korollar 1.45) unmittelbar aus der Cauchy-Schwarz-Ungleichung und dem Satz des Pythagoras. \square

Eine Analyse des Beweises zeigt, daß für $x, y \neq 0$ in der Cauchy-Schwarz-Ungleichung genau dann Gleichheit gilt, wenn $\lambda \in \mathbb{K}$ mit $x = \lambda y$ existiert. (Es gilt dann $\lambda = \langle x, y \rangle / \langle y, y \rangle$.) Gleichheit in der Dreiecksungleichung gilt genau dann, wenn $\lambda \in \mathbb{R}_+$ mit $x = \lambda y$ existiert.

Beispiele. 1. Das Beispiel $V = \mathbb{K}^p$ mit $\langle x, y \rangle := \sum_{j=1}^p x_j \overline{y_j}$ für $x = (x_1, \dots, x_n)^T, y = (y_1, \dots, y_n)^T \in V$ wurde bereits diskutiert. Die zugehörige Norm heißt *euklidische Norm* und wird mit $\|\cdot\|_2$ bezeichnet, also $\|x\|_2 = \sqrt{\sum_{j=1}^p x_j \overline{x_j}} = \sqrt{\sum_{j=1}^p |x_j|^2}$. Die zugehörige Metrik heißt *euklidische Metrik*.

2. Sei $V = C[a, b]$ und $\langle f, g \rangle := \int_a^b f \overline{g}$. Die Skalarprodukteigenschaften sind einfach nachzurechnen; für (iii) vgl. Beispiel 5 nach Satz 6.1.1.

Satz 6.1.4 *Seien (M, d_M) und (N, d_N) metrische Räume. Dann ist durch*

$$d((x_1, y_1), (x_2, y_2)) := d_M(x_1, x_2) + d_N(y_1, y_2)$$

eine Metrik auf $M \times N$ gegeben.

Beweis. Übung.

Es gibt auch viele andere Metriken auf $M \times N$. So ist z. B. auch durch

$$d((x_1, y_1), (x_2, y_2)) := \max\{d_M(x_1, x_2), d_N(y_1, y_2)\}$$

eine Metrik auf $M \times N$ gegeben.

Ein Analogon gilt für normierte Räume: sind $(V, \|\cdot\|_V)$ und $(W, \|\cdot\|_W)$ normierte Räume, so sind durch $\|(x, y)\| := \|x\|_V + \|y\|_W$ und $\|(x, y)\| := \max\{\|x\|_V, \|y\|_W\}$ Normen auf $V \times W$ gegeben.

Entsprechend kann man Metriken bzw. Normen auf dem kartesischen Produkt von endlich vielen metrischen bzw. normierten Räumen definieren.

6.2 Offene und abgeschlossene Mengen

Wir übertragen die für \mathbb{R} und \mathbb{C} in §3.3 eingeführten Konzepte auf metrische Räume. (Dabei sind normierte Räume und Prähilberträume gemäß Satz 6.1.1 und 6.1.2 als metrische Räume zu betrachten.)

Definition 6.2.1 Sei (M, d) metrischer Raum und $x \in M$. Für $\varepsilon \in \mathbb{R}_+$ heißt dann $U(x, \varepsilon) := \{y \in M : d(x, y) < \varepsilon\}$ die ε -Umgebung von x . Sei weiter $A \subset M$. Existiert $\varepsilon \in \mathbb{R}_+$ mit $U(x, \varepsilon) \subset A$, so heißt A *Umgebung* von x . Weiter heißt A *offen*, falls A Umgebung jedes Punktes von A ist, und A heißt *abgeschlossen*, falls $M \setminus A$ offen ist.

Vieles, was wir in Analysis I über offene und abgeschlossene Teilmengen von \mathbb{R} und \mathbb{C} gesagt haben, gilt (mit den gleichen Argumenten) in allgemeinen metrischen Räumen. Beispielsweise überträgt sich Satz 3.3.5 ohne weiteres: *Vereinigungen offener Mengen und endliche Durchschnitte offener Mengen sind wieder offen. Durchschnitte abgeschlossener Mengen und endliche Vereinigungen abgeschlossener Mengen sind wieder abgeschlossen.*

Sei (M, d) metrischer Raum und $N \subset M$. Dann ist auch $(N, d|(N \times N))$ metrischer Raum. Für $A \subset N$ können wir Offenheit und Abgeschlossenheit außer in (M, d) jetzt auch in $(N, d|(N \times N))$ betrachten. Wir sagen dann, daß A *relativ offen* (bzw. *abgeschlossen*) in N ist. Den Zusatz "relativ" lassen wir dabei in der Regel weg. (Für $M = \mathbb{C}$ stimmen diese Begriffe mit den in Definition 3.3.1 definierten Begriffen überein.)

Aus der Offenheit von A in M folgt die Offenheit von A in N . Allgemeiner gilt:

Satz 6.2.1 Sei (M, d) metrischer Raum und $A \subset N \subset M$. Dann ist A offen in N genau dann, wenn ein (in M) offenes $B \subset M$ mit $A = B \cap N$ existiert.

Beweis. Wir bezeichnen die ε -Umgebungen in M bzw. N mit $U_M(\cdot, \varepsilon)$ bzw. $U_N(\cdot, \varepsilon)$. Dann gilt $U_N(x, \varepsilon) = U_M(x, \varepsilon) \cap N$ für alle $x \in N$. Aus $U_M(x, \varepsilon) \subset B$ folgt also $U_N(x, \varepsilon) \subset B \cap N$ und damit " \Leftarrow ".

Um " \Rightarrow " zu zeigen, sei A offen in N . Für jedes $x \in A$ existiert dann $\varepsilon_x > 0$ mit $U_N(\varepsilon_x, x) \subset A$. Hieraus erhält man $A = \bigcup_{x \in A} U_N(\varepsilon_x, x)$ und damit leistet $B := \bigcup_{x \in A} U_M(\varepsilon_x, x)$ das Verlangte.

Definition 6.2.2 Sei (M, d) metrischer Raum und (a_n) Folge in M . Dann heißt (a_n) *konvergent*, falls $a \in M$ mit folgender Eigenschaft existiert:

Zu jedem $\varepsilon \in \mathbb{R}_+$ existiert ein $N \in \mathbb{N}$, so daß $d(a_n, a) < \varepsilon$ für alle $n \in \mathbb{N}$ mit $n \geq N$.

Dieses a heißt dann *Grenzwert* der Folge (a_n) und wir schreiben $a = \lim_{n \rightarrow \infty} a_n$ oder $a_n \rightarrow a$ (für $n \rightarrow \infty$).

Weiter heißt $b \in M$ *Häufungswert* von (a_n) , falls für jedes $\varepsilon \in \mathbb{R}_+$ unendlich viele $n \in \mathbb{N}$ mit $d(a_n, b) < \varepsilon$ existieren.

Diese Definition ist völlig analog zu den Definitionen 2.1.2 und 2.3.1 in Analysis I. Viele der dort gemachten Aussagen über Grenzwerte und Häufungswerte übertragen sich leicht. So gilt wieder die Aussage von Satz 2.1.1: *eine konvergente Folge hat genau einen Grenzwert.*

In Quantorenschreibweise lautet die Aussage, daß a Grenzwert von (a_n) ist, wie folgt:

- $\forall \varepsilon \in \mathbb{R}_+ \exists N \in \mathbb{N} \forall n \in \mathbb{N} : n \geq N \Rightarrow d(a_n, a) < \varepsilon,$

Leicht sieht man wieder, daß die folgende Aussage äquivalent dazu ist:

- \forall Umgebungen U von $a \exists N \in \mathbb{N} \forall n \in \mathbb{N} : n \geq N \Rightarrow a_n \in U.$

Beispiel. Sei $V := C[a, b]$ und d die von der Norm $\|\cdot\|_\infty$ erzeugte Metrik, also $d(f, g) = \|f - g\|_\infty = \max_{x \in [a, b]} |f(x) - g(x)|$ für $f, g \in V$. Sei (f_n) Folge in V und $f \in V$. Dann gilt

$$\begin{aligned} f_n &\rightarrow f \text{ (bzgl. der Metrik in } V) \\ \Leftrightarrow \forall \varepsilon \in \mathbb{R}_+ \exists N \in \mathbb{N} \forall n \in \mathbb{N} : n \geq N &\Rightarrow \|f_n - f\|_\infty < \varepsilon \\ \Leftrightarrow \forall \varepsilon \in \mathbb{R}_+ \exists N \in \mathbb{N} \forall n \in \mathbb{N} : n \geq N &\Rightarrow \forall x \in [a, b] : |f_n(x) - f(x)| < \varepsilon \\ \Leftrightarrow \forall \varepsilon \in \mathbb{R}_+ \exists N \in \mathbb{N} \forall n \in \mathbb{N} \forall x \in [a, b] : n \geq N &\Rightarrow |f_n(x) - f(x)| < \varepsilon \\ \Leftrightarrow f_n &\rightarrow f \text{ gleichmäßig} \end{aligned}$$

Äquivalent dazu, daß a Häufungswert von (a_n) ist, sind:

- $\forall \varepsilon \in \mathbb{R}_+ \forall N \in \mathbb{N} \exists n \in \mathbb{N} : n \geq N \wedge d(a_n, a) < \varepsilon,$
- \forall Umgebungen U von $a \forall N \in \mathbb{N} \exists n \in \mathbb{N} : n \geq N \wedge a_n \in U.$

Es gilt wieder die Aussage von Satz 2.3.2:

- (i) a ist Häufungswert von $(a_n) \Leftrightarrow$ Es existiert eine Teilfolge von (a_n) , die gegen a konvergiert.
- (ii) a ist Grenzwert von $(a_n) \Leftrightarrow$ Alle Teilfolgen von (a_n) konvergieren gegen $a.$

Auch die Aussage von Satz 3.3.3 gilt in metrischen Räumen:

Satz 6.2.2 Eine Teilmenge A eines metrischen Raumes ist genau dann abgeschlossen, wenn für jede konvergente Folge in A auch ihr Grenzwert in A ist.

Die folgende Definition ist analog zu Definition 3.5.1.

Definition 6.2.3 Sei (M, d) metrischer Raum, $A \subset M$ und $\xi \in M$. Dann heißt ξ Häufungspunkt von A , falls eine Folge (a_n) in $A \setminus \{\xi\}$ mit $a_n \rightarrow \xi$ existiert.

Satz 6.2.3 Sei (M, d) metrischer Raum und $A \subset M$. Dann ist A genau dann abgeschlossen, wenn A alle Häufungspunkte von A enthält.

Beweis. “ \Rightarrow ”. Sei A abgeschlossen und ξ Häufungspunkt von A . Dann existiert eine Folge (a_n) in $A \setminus \{\xi\}$ mit $a_n \rightarrow \xi$. Nach Satz 6.2.2 folgt $\xi \in A$.

“ \Leftarrow ”. A enthalte alle seine Häufungspunkte. Wir benutzen wieder Satz 6.2.2. Sei dazu (a_n) Folge in A mit $a_n \rightarrow \xi$. Zu zeigen ist, daß $\xi \in A$ gilt. Existiert $N \in \mathbb{N}$ mit $a_n = \xi$, so folgt $\xi \in A$. Andernfalls ist (a_n) Folge in $A \setminus \{\xi\}$ und damit ist ξ Häufungspunkt von A , also $\xi \in A$ nach Voraussetzung. \square

Definition 6.2.4 Sei (M, d) metrischer Raum und $A \subset M$. Sei X die Menge aller abgeschlossenen Teilmengen von M , die A enthalten. Dann heißt

$$\bar{A} := \bigcap_{S \in X} S$$

der Abschluß von A .

Dies ist genau die Definition, die für $M = \mathbb{C}$ in Analysis I, Aufgabe 42, gegeben wurde. Die dort gemachten Aussagen gelten auch in dem jetzigen Kontext. Außerdem gilt

$$\bar{A} = A \cup \{x \in M : x \text{ ist Häufungspunkt von } A\}.$$

Definition 6.2.5 Sei (M, d) metrischer Raum und (x_n) Folge in M . Dann heißt (x_n) Cauchyfolge, wenn zu jedem $\varepsilon \in \mathbb{R}_+$ ein $N \in \mathbb{N}$ existiert, so daß für alle $m, n \in \mathbb{N}$ mit $m \geq N$ und $n \geq N$ gilt, daß $d(x_m, x_n) < \varepsilon$ ist.

In Quantorenschreibweise lautet die in der Definition der Cauchyfolge angegebene Eigenschaft:

$$\forall \varepsilon \in \mathbb{R}_+ \exists N \in \mathbb{N} \forall m, n \in \mathbb{N} : m \geq N \wedge n \geq N \Rightarrow d(a_m, a_n) < \varepsilon.$$

Satz 6.2.4 *Konvergente Folgen sind Cauchyfolgen.*

Beweis (vgl. Beweis zu Satz 2.3.4). Sei (M, d) metrischer Raum und sei (x_n) konvergente Folge in M , etwa $x_n \rightarrow \xi$. Sei $\varepsilon > 0$. Dann existiert $N \in \mathbb{N}$ mit $d(x_n, \xi) < \frac{\varepsilon}{2}$ für $n \geq N$. Für $m, n \in \mathbb{N}$ mit $m, n \geq N$ folgt dann $d(x_m, x_n) \leq d(x_m, \xi) + d(\xi, x_n) < \varepsilon$. \square

Die Umkehrung gilt nicht, wie etwa das Beispiel $M = \mathbb{Q}$ (mit der euklidischen Metrik) zeigt, und wie wir später an weiteren Beispielen sehen werden. Vom Cauchy Kriterium für Folgenkonvergenz (Satz 2.3.4) gilt also nur “eine Hälfte”.

Definition 6.2.6 Ein metrischer Raum heißt *vollständig*, wenn jede Cauchyfolge in ihm konvergiert. Ein vollständiger normierter Raum heißt *Banachraum* und ein vollständiger Prähilbertraum heißt *Hilbertraum*.

Satz 6.2.5 *Sei (M, d) vollständiger metrischer Raum und $A \subset M$ abgeschlossen. Dann ist A vollständig (d. h., $(A, d|(A \times A))$ ist vollständiger metrischer Raum).*

Beweis. Ist (x_n) Cauchyfolge in A , so ist (x_n) Cauchyfolge in M , also gilt $x_n \rightarrow \xi$ für ein $\xi \in M$. Mit Satz 6.2.2 folgt $\xi \in A$. \square

Beispiele. 1. $(C[a, b], \|\cdot\|_\infty)$ ist vollständig. Denn sei (f_n) Cauchyfolge in $C[a, b]$, d. h., (vgl. das Beispiel vor Definition 3.5.1)

$$\forall \varepsilon \in \mathbb{R}_+ \exists N \in \mathbb{N} \forall x \in [a, b] \forall m, n \in \mathbb{N} : m \geq N \wedge n \geq N \Rightarrow |f_m(x) - f_n(x)| < \varepsilon.$$

Nach dem Cauchy Kriterium für Funktionenfolgen (Satz 3.6.1) folgt, daß (f_n) gleichmäßig gegen eine Funktion $f : [a, b] \rightarrow \mathbb{C}$ konvergiert. Nach Satz 3.6.2 ist $f \in C[a, b]$. Es folgt, daß (f_n) in $(C[a, b], \|\cdot\|_\infty)$ gegen f konvergiert.

2. $(C[a, b], \|\cdot\|_1)$ ist nicht vollständig. Ohne Beschränkung der Allgemeinheit sei $[a, b] = [-1, 1]$. Wir betrachten die durch

$$f(x) = \begin{cases} 0 & \text{falls } -1 \leq x \leq 0, \\ nx & \text{falls } 0 < x \leq \frac{1}{n}, \\ 1 & \text{falls } \frac{1}{n} \leq x \leq 1 \end{cases}$$

definierte Folge (f_n) . Für $m > n$ gilt

$$\|f_m - f_n\|_1 = \int_{-1}^1 |f_m - f_n| \leq \int_0^{\frac{1}{n}} |f_m - f_n| \leq \frac{1}{n}$$

und damit ist (f_n) Cauchyfolge. Die Folge (f_n) ist aber nicht konvergent. Dies scheint intuitiv klar, da der der einzige “Kandidat” f für eine Grenzfunktion $f(x) = 0$ für $x < 0$ und $f(x) = 1$ für $x > 0$ erfüllen sollte, und damit in 0 nicht stetig sein kann.

Für einen formalen Beweis nehmen wir an, daß (f_n) konvergent ist, also $f_n \rightarrow f \in C[-1, 1]$. Es existiert $\delta \in (0, 1]$ mit $|f(x) - f(0)| \leq \frac{1}{4}$ für $|x| < \delta$. Es folgt

$$\|f - f_n\|_1 \geq \int_{\frac{1}{n}}^\delta |f(x) - 1| dx \geq \int_{\frac{1}{n}}^\delta 1 - |f(0)| - |f(x) - f(0)| dx \geq \int_{\frac{1}{n}}^\delta \frac{1}{4} dx = \frac{\delta - \frac{1}{n}}{4}$$

für $|f(0)| \leq \frac{1}{2}$, $n \geq \frac{1}{\delta}$, und

$$\|f - f_n\|_1 \geq \int_{-\delta}^0 |f(x)| dx \geq \int_{-\delta}^0 \frac{1}{4} dx = \frac{\delta}{4}$$

für $|f(0)| > \frac{1}{2}$. In beiden Fällen erhält man einen Widerspruch.

3. $(\mathbb{K}^p, \|\cdot\|_\infty)$ ist vollständig (für $p \in \mathbb{N}$). Denn sei (x^k) Cauchyfolge in \mathbb{K}^p , mit $x^k = (x_1^k, \dots, x_p^k)$. Wegen $|x_j^k - x_j^\ell| \leq \|x^k - x^\ell\|_\infty$ für $k, \ell \in \mathbb{N}$ und $j \in \{1, \dots, p\}$ folgt, daß $(x_j^k)_{k \in \mathbb{N}}$ für jedes $j \in \{1, \dots, p\}$ eine Cauchyfolge ist, und damit also konvergent, etwa $x_j^k \rightarrow \xi_j$. Mit $\xi = (\xi_1, \dots, \xi_p)$ folgt dann

$$\|x^k - \xi\|_\infty = \max_j |x_j^k - \xi_j| \rightarrow 0,$$

also $x^k \rightarrow \xi$ für $k \rightarrow \infty$.

Genauso kann man zeigen, daß $(\mathbb{K}^p, \|\cdot\|_1)$ und $(\mathbb{K}^p, \|\cdot\|_2)$ vollständig sind. Dies folgt aber auch unmittelbar aus der Tatsache, daß die Normen $\|\cdot\|_1$ und $\|\cdot\|_2$ äquivalent zur Norm $\|\cdot\|_\infty$ sind; vgl. LA I, Proposition 1.49. Denn sind die Normen $\|\cdot\|'$ und $\|\cdot\|''$ auf dem normierten Raum V äquivalent, so ist eine Folge (x_n) in V genau dann konvergent (bzw. Cauchyfolge) in $(V, \|\cdot\|')$, wenn sie konvergent (bzw. Cauchyfolge) in $(V, \|\cdot\|'')$ ist; vgl. LA I, Satz 1.51.

Ebenso gilt, daß $A \subset V$ genau dann offen (bzw. abgeschlossen) in $(V, \|\cdot\|')$ ist, wenn A offen (bzw. abgeschlossen) in $(V, \|\cdot\|'')$ ist.

In LA I, Definition 1.48, wurde die Äquivalenz von Normen nur für \mathbb{R} -Vektorräume definiert. Die Definition für \mathbb{C} -Vektorräume ist aber genau dieselbe. Der Vollständigkeit halber sei die Definition hier noch aufgeführt.

Definition 6.2.7 Sei V Vektorraum über \mathbb{K} und seien $\|\cdot\|'$ und $\|\cdot\|''$ Normen auf V . Dann heißen $\|\cdot\|'$ und $\|\cdot\|''$ äquivalent, falls $\alpha, \beta \in \mathbb{R}_+$ existieren, so daß $\alpha\|x\|' \leq \|x\|'' \leq \beta\|x\|'$ für alle $x \in V$ gilt.

6.3 Stetigkeit

Definition 6.3.1 Seien $(M, d_M), (N, d_N)$ metrische Räume und sei $f : M \rightarrow N$ eine Funktion. Sei $\xi \in M$. Dann heißt f stetig in ξ , falls für jede Folge (x_n) in M mit $x_n \rightarrow \xi$ gilt, daß $f(x_n) \rightarrow f(\xi)$. Für $A \subset M$ heißt f stetig in A , falls f stetig in jedem Punkt von A ist. Schließlich heißt f stetig, wenn f stetig in M ist.

Für $M, N \subset \mathbb{C}$, mit der euklidischen Metrik versehen, ist dies genau Definition 3.1.1. Ansonsten ist obige Definition wörtlich dieselbe. Die Sätze 3.1.1 und 3.1.2 übertragen sich unmittelbar:

Satz 6.3.1 Seien $(M, d_M), (N, d_N), (P, d_P)$ metrische Räume und seien $g : M \rightarrow N$ stetig (in $\xi \in M$) und $f : N \rightarrow P$ stetig (in $g(\xi) \in g(M)$). Dann ist $f \circ g$ stetig (in ξ).

Satz 6.3.2 Seien $(M, d_M), (N, d_N)$ metrische Räume, $\xi \in M$ und $f : M \rightarrow N$. Dann ist f stetig in ξ genau dann, wenn zu jedem $\varepsilon \in \mathbb{R}_+$ ein $\delta \in \mathbb{R}_+$ existiert, so daß für jedes $x \in M$ mit $d_M(x, \xi) < \delta$ gilt, daß $d_N(f(x), f(\xi)) < \varepsilon$ ist.

Beispiel. Sei (M, d) metrischer Raum, $A \subset M$ und $f : A \rightarrow \mathbb{R}$ definiert durch $f(x) := \inf_{a \in A} d(a, x)$. Es ist also $f(x)$ der Abstand von x zur Menge A .

Behauptung. f ist stetig.

Beweis. Sei $x, y \in M$ und $a \in A$. Dann gilt $d(x, a) \leq d(x, y) + d(y, a)$ und damit $f(x) \leq d(x, y) + f(y)$. Es folgt $f(x) - f(y) \leq d(x, y)$ und aus Symmetriegründen auch $f(y) - f(x) \leq d(x, y)$, also $|f(y) - f(x)| \leq d(x, y)$. Die Stetigkeit von d folgt jetzt aus Satz 6.3.2. (Man kann dort $\delta := \varepsilon$ wählen.) \square

Auch Satz 3.3.2 überträgt sich (mit Beweis):

Satz 6.3.3 Seien $(M, d_M), (N, d_N)$ metrische Räume und $f : M \rightarrow N$. Dann gilt:

- (i) Sei $\xi \in M$. Dann ist f stetig in ξ genau dann, wenn für jede Umgebung U von $f(\xi)$ das Urbild $f^{-1}(U)$ eine Umgebung von ξ ist.
- (ii) Die Funktion f ist genau dann stetig, wenn für jedes offene $U \subset N$ das Urbild $f^{-1}(U)$ offen ist.

In Satz 2.2.1 wurde für Zahlenfolgen (a_n) und (b_n) mit $a_n \rightarrow a$ und $b_n \rightarrow b$ gezeigt, daß $a_n + b_n \rightarrow a + b$ und $a_n \cdot b_n \rightarrow a \cdot b$ gilt. Hieraus wurde in §3.1 die Stetigkeit von Summe und Produkt stetiger Funktionen gefolgert. Entsprechende Aussagen gelten auch hier wieder (wobei die Beweise analog sind).

Satz 6.3.4 Sei (M, d) metrischer Raum, $(V, \|\cdot\|)$ normierter Raum (über \mathbb{K}), $\xi \in M$, $\lambda \in \mathbb{K}$ und $f, g : M \rightarrow V$ stetig in ξ . Dann sind $f + g$ und $\lambda \cdot f$ stetig in ξ . Ist speziell $V = \mathbb{K}$ und $\|\cdot\| = |\cdot|$, so ist auch $f \cdot g$ stetig in ξ . Im Falle $g(\xi) \neq 0$ ist auch $\frac{f}{g}$ stetig in ξ .

Wir untersuchen Abbildungen $f : \mathbb{K}^p \rightarrow \mathbb{K}$ auf Stetigkeit bzgl. der Norm $\|\cdot\|_\infty$. Für $j \in \{1, \dots, p\}$ heißt die durch $\pi_j(x_1, \dots, x_p) = x_j$ gegebene Abbildung $\pi_j : \mathbb{K}^p \rightarrow \mathbb{K}$ die *Projektion* auf die j -te Koordinate. Offensichtlich ist π_j stetig. (Man wähle $\delta := \varepsilon$ in Satz 6.3.2.) Nach Satz 6.3.4 sind dann auch Funktionen f der Form

$$f(x_1, \dots, x_p) = \sum_{(k_1, \dots, k_p) \in I} a_{k_1, \dots, k_p} x_1^{k_1} \cdot \dots \cdot x_p^{k_p}$$

stetig, wobei $I \subset \mathbb{N}_0^p$ endlich und $a_{k_1, \dots, k_p} \in \mathbb{K}$ für $(k_1, \dots, k_p) \in I$. Solche Funktionen heißen *Polynome* (in p Veränderlichen). Quotienten von Polynomen heißen *rationale Funktionen*. Sie sind dort definiert, wo der Nenner nicht verschwindet.

Wir untersuchen jetzt Abbildungen $f : \mathbb{K}^p \rightarrow \mathbb{K}^q$ auf Stetigkeit, wieder bzgl. der Norm $\|\cdot\|_\infty$ (sowohl im Definitions- wie im Zielbereich). Für $j \in \{1, \dots, q\}$ heißt $f_j := \pi_j \circ f$ die j -te *Koordinatenfunktion*. Es ist also $f(x) = (f_1(x), \dots, f_q(x))^T$ für $x \in \mathbb{K}^p$.

Satz 6.3.5 Eine Funktion von \mathbb{K}^p nach \mathbb{K}^q ist genau dann stetig (in $\xi \in \mathbb{K}^p$), wenn alle ihre Koordinatenfunktionen stetig (in ξ) sind.

Beweis. Sei $f = (f_1, \dots, f_q)^T : \mathbb{K}^p \rightarrow \mathbb{K}^q$.

“ \Rightarrow ”. Dies folgt aus Satz 6.3.1, da $f_j = \pi_j \circ f$ und π_j stetig.

“ \Leftarrow ”. Es seien alle f_j stetig in ξ . Sei $\varepsilon > 0$. Dann existiert zu $j \in \{1, \dots, q\}$ ein $\delta_j > 0$ mit $|f_j(x) - f_j(\xi)| < \varepsilon$ für $\|x - \xi\|_\infty < \delta_j$. Mit $\delta := \min_j \delta_j$ folgt $\|f(x) - f(\xi)\|_\infty = \max_j |f_j(x) - f_j(\xi)| < \varepsilon$ für $\|x - \xi\|_\infty < \delta$. \square

Man könnte nun vermuten, daß eine Funktion $f = (f_1, \dots, f_q)^T : \mathbb{K}^p \rightarrow \mathbb{K}^q$ stetig ist, wenn sie für jedes j bei festgehaltenem $x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_p$ stetig in x_j ist. Dies ist aber nicht der Fall.

Beispiel. Sei $f : \mathbb{R}^2 \rightarrow \mathbb{R}$,

$$f(x, y) = \begin{cases} \frac{2xy}{x^2+y^2} & \text{falls } (x, y) \neq (0, 0), \\ 0 & \text{falls } (x, y) = (0, 0). \end{cases}$$

Dann ist f unstetig in $(0, 0)$, da $f(x, x) = 1$ für alle $x \in \mathbb{R} \setminus \{0\}$, aber $f(0, 0) = 0$. Wird eine Variable festgehalten, ist f aber stetig in der anderen.

Es folgt aus Satz 6.3.5, daß lineare Abbildungen von \mathbb{K}^p nach \mathbb{K}^q stetig sind, da jede Koordinatenfunktion ein Polynom ist. Denn ist $f = (f_1, \dots, f_q) : \mathbb{K}^p \rightarrow \mathbb{K}^q$ linear, so existieren $a_{ij} \in \mathbb{K}$, $1 \leq i \leq q$, $1 \leq j \leq p$, mit $f_i(x_1, \dots, x_p) = \sum_{j=1}^p a_{ij}x_j$. Entscheidend dabei ist, daß \mathbb{K}^p und \mathbb{K}^q endlichdimensional sind. Lineare Abbildungen zwischen unendlichdimensionalen normierten Räumen brauchen nicht stetig zu sein.

Beispiel. Es ist $(C^1[-1, 1], \|\cdot\|_\infty)$ normierter Raum und $L : C^1[-1, 1] \rightarrow \mathbb{C}$, $f \mapsto f'(0)$ ist linear. Mit $f_n(x) = \frac{1}{n} \sin nx$ gilt $\|f_n\|_\infty \leq \frac{1}{n}$, also $f_n \rightarrow 0$, aber $L(f_n) = f'_n(0) = \cos n0 = 1 \neq 0 = L(0)$. Damit ist L nicht stetig.

Definition 6.3.2 (vgl. Definition 3.4.2) Seien $(M, d_M), (N, d_N)$ metrische Räume und $f : M \rightarrow N$. Dann heißt f *gleichmäßig stetig*, falls zu jedem $\varepsilon \in \mathbb{R}_+$ ein $\delta \in \mathbb{R}_+$ existiert, so daß für alle $x, y \in M$ mit $d_M(x, y) < \delta$ gilt, daß $d_N(f(x), f(y)) < \varepsilon$ ist.

Satz 6.3.6 Seien $(V, \|\cdot\|_V), (W, \|\cdot\|_W)$ normierte Räume und sei $f : V \rightarrow W$ linear. Dann sind äquivalent:

- (i) f ist gleichmäßig stetig,
- (ii) f ist stetig,
- (iii) f ist stetig in 0,
- (iv) $\exists C \in \mathbb{R}_+ \forall x \in V : \|f(x)\|_W \leq C\|x\|_V$.

Beweis. Wir zeigen, daß (i) \Rightarrow (ii) \Rightarrow (iii) \Rightarrow (iv) \Rightarrow (i). Dabei ist (i) \Rightarrow (ii) \Rightarrow (iii) trivial.

zu (iii) \Rightarrow (iv): Sei f stetig in 0. Zu $\varepsilon := 1$ existiert dann $\delta > 0$ mit $\|f(x)\|_W \leq 1$ für $\|x\|_V \leq \delta$. Für $x \in V \setminus \{0\}$ folgt dann

$$\|f(x)\|_W = \frac{\|x\|_V}{\delta} \left\| f\left(\frac{\delta}{\|x\|_V}x\right) \right\|_W \leq \frac{\|x\|_V}{\delta}.$$

Mit $C := \frac{1}{\delta}$ folgt also $\|f(x)\|_W \leq C\|x\|_V$, und das gilt auch für $x = 0$.

zu (iv) \Rightarrow (i): Sei $\varepsilon > 0$ und $\delta := \frac{\varepsilon}{C}$. Für $x, y \in V$ mit $\|x - y\|_V < \delta$ gilt dann $\|f(x) - f(y)\|_W = \|f(x - y)\|_W \leq C\|x - y\|_V < \varepsilon$. \square

Seien $(V, \|\cdot\|_V)$, $(W, \|\cdot\|_W)$ normierte Räume. Den Vektorraum der stetigen linearen Abbildungen von V nach W bezeichnen wir mit $L(V, W)$. Nach Satz 6.3.6, (iv), ist für $f \in L(V, W)$

$$\|f\|_{L(V,W)} := \sup_{x \in V \setminus \{0\}} \frac{\|f(x)\|_W}{\|x\|_V} = \sup_{\substack{x \in V \\ \|x\|_V=1}} \|f(x)\|_W < \infty.$$

Man nennt $\|f\|_{L(V,W)}$ *Operatornorm* von f . Es ist nicht schwer zu zeigen, daß $(L(V, W), \|\cdot\|_{L(V,W)})$ normierter Raum ist. Darüberhinaus ist $(L(V, W), \|\cdot\|_{L(V,W)})$ Banachraum, falls $(W, \|\cdot\|_W)$ Banachraum ist.

Beispiel. Sei $(V, \|\cdot\|_V) = (\mathbb{R}^p, \|\cdot\|_\infty)$, $(W, \|\cdot\|_W) = (\mathbb{R}^q, \|\cdot\|_\infty)$, $f = (f_1, \dots, f_q)^T : \mathbb{R}^p \rightarrow \mathbb{R}^q$. Dann existieren $a_{ij} \in \mathbb{K}$, $1 \leq i \leq q$, $1 \leq j \leq p$, mit $f_i(x_1, \dots, x_p) = \sum_{j=1}^p a_{ij}x_j$ für alle i . In Matrixschreibweise haben wir also

$$\begin{pmatrix} f_1(x) \\ f_2(x) \\ \vdots \\ f_q(x) \end{pmatrix} = \begin{pmatrix} a_{11} & \dots & a_{1p} \\ a_{21} & \dots & a_{2p} \\ \vdots & & \vdots \\ a_{q1} & \dots & a_{qp} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_p \end{pmatrix}.$$

Nun ist für $x = (x_1, \dots, x_p)^T \in \mathbb{R}^p$

$$|f_i(x_1, \dots, x_p)| \leq \sum_{j=1}^p |a_{ij}| \cdot |x_j| \leq \|x\|_\infty \sum_{j=1}^p |a_{ij}|$$

für alle i und damit

$$\|f(x)\|_\infty \leq \|x\|_\infty \max_i \sum_{j=1}^p |a_{ij}|.$$

Wählt man, bei festem i , nun $x_j = \pm 1$, so daß x_j und a_{ij} das gleiche Vorzeichen haben, so hat man in obiger Abschätzung Gleichheit. Es folgt

$$\|f\|_{L(V,W)} = \max_i \sum_{j=1}^p |a_{ij}|.$$

Man nennt dies auch die *Zeilensummennorm* der Matrix $(a_{ij})_{\substack{1 \leq i \leq q \\ 1 \leq j \leq p}}$.

Definition 6.3.3 (vgl. Definition 3.5.2) Seien (M, d_M) , (N, d_N) metrische Räume, ξ Häufungspunkt von M und $f : M \rightarrow N$. Wir sagen, daß $f(x)$ für $x \rightarrow \xi$ konvergiert, falls $\eta \in N$ existiert, so daß für jede Folge (x_n) in $M \setminus \{\xi\}$ mit $x_n \rightarrow \xi$ auch $f(x_n) \rightarrow \eta$ gilt. Dieses η heißt dann *Grenzwert* von f für $x \rightarrow \xi$ und wird mit $\lim_{x \rightarrow \xi} f(x)$ bezeichnet. Wir schreiben auch $f(x) \rightarrow \eta$ für $x \rightarrow \xi$.

Wie in Satz 3.5.1 sehen wir, daß f genau dann stetig in ξ ist, wenn $\lim_{x \rightarrow \xi} f(x) = f(\xi)$ gilt.

Auch Satz 3.5.3 gilt entsprechend: *Es gilt $\lim_{x \rightarrow \xi} f(x) = \eta$ genau dann, wenn*

$$\forall \varepsilon \in \mathbb{R}_+ \exists \delta \in \mathbb{R}_+ \forall x \in M \setminus \{\xi\} : d_M(x, \xi) < \delta \Rightarrow d_N(f(x), \eta) < \varepsilon.$$

6.4 Kompaktheit

Definition 6.4.1 Sei (M, d) metrischer Raum.

- (i) (M, d) heißt *kompakt*, falls jede Folge in M eine konvergente Teilfolge besitzt.
- (ii) (M, d) heißt *totalbeschränkt* (oder *präkompakt*), falls für alle $\varepsilon \in \mathbb{R}_+$ eine endliche Teilmenge E von M existiert, so daß $M \subset \bigcup_{x \in E} U(x, \varepsilon)$.
- (iii) (M, d) heißt *beschränkt*, falls $\sup_{x, y \in M} d(x, y) < \infty$.

Kurz sagen wir auch, daß M (statt (M, d)) kompakt bzw. (total)beschränkt ist. Für $K \subset M$ sagen wir, daß K kompakt bzw. (total)beschränkt ist, wenn der metrische Raum $(K, d|(K \times K))$ diese Eigenschaft hat.

Teil (i) obiger Definition entspricht damit Definition 3.4.1. Statt die Existenz einer konvergenten Teilfolge zu fordern, kann man natürlich auch die Existenz eines Häufungswertes fordern; vgl. Satz 2.3.2, (i).

Aus der Totalbeschränktheit folgt die Beschränktheit, denn aus

$$M \subset \bigcup_{x \in E} U(x, \varepsilon)$$

folgt $d(a, b) \leq \max_{x, y \in E} d(x, y) + 2\varepsilon$ für $a, b \in M$.

Sind I, M nichtleere Mengen, so nennt man eine durch $i \mapsto x_i$ gegebene Funktion von I nach M auch *Familie* und bezeichnet sie mit $(x_i)_{i \in I}$. (Für $I = \mathbb{N}$ sind dies Folgen.)

Definition 6.4.2 Sei (M, d) metrischer Raum, $K \subset M$ und $(U_i)_{i \in I}$ eine Familie offener Teilmengen von M mit $K \subset \bigcup_{i \in I} U_i$. Dann heißt $(U_i)_{i \in I}$ eine *offene Überdeckung* von K . Existiert eine endliche Teilmenge $E \subset I$ mit $K \subset \bigcup_{i \in E} U_i$, so sagt man, daß $(U_i)_{i \in I}$ eine *endliche Teilüberdeckung* besitzt.

Die Totalbeschränktheit besagt also, daß für jedes $\varepsilon > 0$ die offene Überdeckung $(U(x, \varepsilon))_{x \in M}$ eine endliche Teilüberdeckung besitzt.

Satz 6.4.1 Sei (M, d) metrischer Raum und $K \subset M$. Dann sind äquivalent:

- (i) jede offene Überdeckung von K besitzt eine endliche Teilüberdeckung,
- (ii) K ist kompakt,
- (iii) K ist vollständig und totalbeschränkt.

Die Äquivalenz von (i) und (ii) heißt auch *Satz von Heine-Borel*.

In sogenannten *topologischen Räumen* – die eine Verallgemeinerung der metrischen Räume sind – gilt Satz 6.4.1 nicht. Dort wird die in (i) formulierte Eigenschaft als Kompaktheit bezeichnet und die in Definition 6.4.1, (i), angegebene als Folgenkompaktheit.

Beweis von Satz 6.4.1. zu (i) \Rightarrow (ii): Sei (x_n) Folge in K , die keine konvergente Teilfolge (und damit also auch keinen Häufungswert) besitzt. Dann ist $A := \{x_n : n \in \mathbb{N}\}$ unendlich.

Sei nun $y \in K$. Da y kein Häufungswert von (x_n) ist, existiert $\varepsilon_y \in \mathbb{R}_+$ und $N \in \mathbb{N}$, so daß $d(x_n, y) \geq \varepsilon_y$ für alle $n \geq N$ gilt. Folglich ist $A \cap U(y, \varepsilon_y)$ endlich.

Nun ist $(U(y, \varepsilon_y))_{y \in K}$ offene Überdeckung von K . Nach (i) existiert nun eine endliche Teilüberdeckung, d. h., es existiert eine endliche Teilmenge E von K mit $K \subset \bigcup_{y \in E} U(y, \varepsilon_y)$. Es folgt $A = A \cap K \subset \bigcup_{y \in E} A \cap U(y, \varepsilon_y)$. Damit ist A endlich, ein Widerspruch.

zu (ii) \Rightarrow (iii): Sei K kompakt. Wir zeigen zuerst, daß K vollständig ist. Sei dazu (x_n) Cauchyfolge in K . Dann hat (x_n) eine konvergente Teilfolge (x_{n_k}) , etwa $x_{n_k} \rightarrow \xi \in K$. Wir zeigen, daß $x_n \rightarrow \xi$. Sei dazu $\varepsilon > 0$. Da (x_n) Cauchyfolge ist, existiert $N \in \mathbb{N}$ mit $d(x_n, x_m) < \frac{\varepsilon}{2}$ für $m, n \geq N$. Da $x_{n_k} \rightarrow \xi$ existiert $K \in \mathbb{N}$ mit $d(x_{n_k}, \xi) < \frac{\varepsilon}{2}$ für $k \geq K$. Sei nun $k \geq K$ mit $n_k \geq N$. Für $n \geq N$ folgt dann $d(x_n, \xi) \leq d(x_n, x_{n_k}) + d(x_{n_k}, \xi) < \varepsilon$. Es gilt also $x_n \rightarrow \xi$.

Wir zeigen jetzt, daß K totalbeschränkt ist. Dazu nehmen wir an, daß dies nicht der Fall ist. Dann existiert $\varepsilon > 0$ mit $K \not\subset \bigcup_{x \in E} U(x, \varepsilon)$ für jede endliche Teilmenge E von K . Es sei nun $x_1 \in K$ beliebig. Dann gilt $K \not\subset U(x_1, \varepsilon)$ und damit existiert $x_2 \in K \setminus U(x_1, \varepsilon)$. Wegen $K \not\subset U(x_1, \varepsilon) \cup U(x_2, \varepsilon)$ existiert $x_3 \in K \setminus (U(x_1, \varepsilon) \cup U(x_2, \varepsilon))$. Induktiv findet man so eine Folge (x_n) in K mit $x_{n+1} \in K \setminus \bigcup_{j=1}^n U(x_j, \varepsilon)$. Es folgt $d(x_m, x_n) \geq \varepsilon$ für $m \neq n$. Dies impliziert, daß keine Teilfolge von (x_n) eine Cauchyfolge ist. Damit konvergiert aber auch keine Teilfolge von (x_n) , ein Widerspruch.

zu (iii) \Rightarrow (i): Sei $(U_i)_{i \in I}$ offene Überdeckung von K . Wir nehmen an, daß keine endliche Teilüberdeckung existiert.

Wegen der Totalbeschränktheit existiert eine endliche Teilmenge E_1 von K mit $K \subset \bigcup_{x \in E_1} U(x, \frac{1}{2})$. Nun ist $(U_i)_{i \in I}$ auch offene Überdeckung von $K \cap U(x, \frac{1}{2})$ für alle $x \in K$. Es folgt, daß $x_1 \in E_1$ existiert, so daß keine endliche Teilüberdeckung von $K \cap U(x_1, \frac{1}{2})$ existiert.

Wegen der Totalbeschränktheit existiert weiter eine endliche Teilmenge E_2 von K mit $K \cap U(x_1, \frac{1}{2}) \subset \bigcup_{x \in E_2} U(x, \frac{1}{4})$. Dabei kann man $U(x_1, \frac{1}{2}) \cap U(x, \frac{1}{4}) \neq \emptyset$, also $d(x_1, x) < \frac{1}{2} + \frac{1}{4}$, für alle $x \in E_2$ annehmen. Es existiert nun $x_2 \in E_2$, so daß keine endliche Teilüberdeckung von $K \cap U(x_2, \frac{1}{4})$ existiert.

Induktiv findet man so eine Folge (x_n) , so daß keine endliche Teilüberdeckung von $K \cap U(x_n, \frac{1}{2^n})$ existiert und außerdem

$$d(x_n, x_{n+1}) < \frac{1}{2^n} + \frac{1}{2^{n+1}} = \frac{3}{2^{n+1}}$$

gilt. Für $m > n$ folgt

$$d(x_n, x_m) < \sum_{k=n+1}^m \frac{3}{2^k} = \frac{3}{2^{n+1}} \sum_{j=0}^{\infty} \frac{1}{2^j} = \frac{3}{2^n}.$$

Damit ist (x_n) Cauchyfolge und wegen der Vollständigkeit von K also konvergent, etwa $x_n \rightarrow \xi \in K$. Da $(U_i)_{i \in I}$ offene Überdeckung von K ist, existiert $i_0 \in I$ mit $\xi \in U_{i_0}$.

Nun existiert $\varepsilon > 0$ mit $U(\xi, \varepsilon) \subset U_{i_0}$. Für genügend großes $n \in \mathbb{N}$ ist $d(x_n, \xi) < \frac{\varepsilon}{2}$ und $\frac{1}{2^n} < \frac{\varepsilon}{2}$. Es folgt

$$K \cap U\left(x_n, \frac{1}{2^n}\right) \subset U\left(x_n, \frac{1}{2^n}\right) \subset U\left(\xi, d(x_n, \xi) + \frac{1}{2^n}\right) \subset U(\xi, \varepsilon) \subset U_{i_0}.$$

Damit existiert eine endliche Teilüberdeckung von $K \cap U(x_n, \frac{1}{2^n})$, ein Widerspruch. \square

Satz 6.4.2 (i) *Kompakte Mengen sind beschränkt und abgeschlossen.*

(ii) *Abgeschlossene Teilmengen kompakter Mengen sind kompakt.*

Beweis. Sei K kompakt (im metrischen Raum (M, d)).

zu (i): Zunächst ist K wegen Satz 6.4.1 totalbeschränkt und damit auch beschränkt.

Um die Abgeschlossenheit zu zeigen, benutzen wir Satz 6.2.4. Sei dazu (x_n) Folge in K mit $x_n \rightarrow \xi \in M$. Zu zeigen ist, daß $\xi \in K$. Dies folgt aber, da (x_n) Cauchyfolge und K nach Satz 6.4.1 vollständig ist.

zu (ii): Sei $A \subset K$ abgeschlossen. Wir geben zwei Beweise, daß A kompakt ist.

Beweisvariante 1. Sei (x_n) Folge in A . Dann besitzt (x_n) eine Teilfolge, die in K konvergiert, etwa $x_{n_k} \rightarrow \xi \in K$. Da A abgeschlossen, folgt mit Satz 6.2.4, daß $\xi \in A$. Also ist A kompakt.

Beweisvariante 2. Sei $(U_i)_{i \in I}$ offene Überdeckung von A . Da $M \setminus A$ offen ist, ist dann durch $K \subset \bigcup_{i \in I} U_i \cup M \setminus A$ eine offene Überdeckung von K gegeben. Damit existiert eine endliche Teilmenge $E \subset I$ mit $K \subset \bigcup_{i \in E} U_i \cup M \setminus A$, also $A \subset \bigcup_{i \in E} U_i$. \square

Nach Satz 3.4.1 gilt für $K \subset \mathbb{C}$ auch die Umkehrung von Satz 6.4.2, (i). Wir werden gleich sehen, daß dies in allgemeinen metrischen Räumen nicht gilt. Es gilt aber der folgende Satz.

Satz 6.4.3 *Sei $p \in \mathbb{N}$ und $K \subset \mathbb{R}^p$. Bezüglich der durch die Norm $\|\cdot\|_\infty$ erzeugten Metrik ist K genau dann kompakt, wenn K beschränkt und abgeschlossen ist.*

Auch dieser Satz wird nach Heine und Borel benannt. Er gilt auch für andere Normen in \mathbb{R}^p ; siehe Satz 6.4.6 später.

Beweis von Satz 6.4.3. “ \Rightarrow ”. Satz 6.4.2, (i).

“ \Leftarrow ”. Nach Satz 6.4.1 reicht es zu zeigen, daß K vollständig und totalbeschränkt ist. Die Vollständigkeit folgt dabei wegen Satz 6.2.5 (und der Vollständigkeit von \mathbb{R}^p) unmittelbar aus der Abgeschlossenheit.

Wegen der Beschränktheit von K existiert $r \in \mathbb{R}_+$ mit $\|x\|_\infty \leq r$ für alle $x \in K$. Ist dann $\varepsilon > 0$ und $N \in \mathbb{N}$ mit $N\varepsilon \geq r$, so folgt

$$K \subset \bigcup_{j_1, \dots, j_n = -N}^N U((j_1\varepsilon, \dots, j_n\varepsilon), \varepsilon)$$

und damit die Totalbeschränktheit von K . \square

Beispiel. Wir betrachten den Prähilbertraum $(C[0, 2\pi], \langle \cdot, \cdot \rangle)$ und bezeichnen die zugehörige Norm mit $\|\cdot\|_2$. Für $n \in \mathbb{Z}$ sei $e_n \in C[0, 2\pi]$ definiert durch $e_n(x) = e^{inx}$. Dann gilt

$$\langle e_n, e_m \rangle = \int_0^{2\pi} e_n \overline{e_m} = \int_0^{2\pi} e^{i(n-m)x} dx = \frac{e^{i(n-m)x}}{i(n-m)} \Big|_0^{2\pi} = 0$$

für $n \neq m$ und

$$\|e_n\|_2 = \sqrt{\langle e_n, e_m \rangle} = \sqrt{\int_0^{2\pi} 1} = \sqrt{2\pi}$$

für $n \in \mathbb{Z}$. Nach dem Satz von Pythagoras ist damit $\|e_n - e_m\|_2^2 = \|e_n\|_2^2 + \|e_m\|_2^2 = 4\pi$, also $\|e_n - e_m\|_2 = \sqrt{4\pi}$.

Wir betrachten jetzt $K := \{e_n : n \in \mathbb{N}\}$. Dann ist K offensichtlich beschränkt. Außerdem ist K abgeschlossen, denn ist (x_k) eine Folge in K , die in $(C[0, 2\pi], \langle \cdot, \cdot \rangle)$ konvergiert, so existieren $M, N \in \mathbb{N}$ mit $x_k = e_N$ für $k \geq M$.

Andererseits ist K nicht kompakt, da $(e_n)_{n \in \mathbb{N}}$ keine konvergente Teilfolge besitzt. Es folgt, daß die Umkehrung von Satz 6.4.2, (i), im allgemeinen gilt.

Satz 6.4.4 *Seien (M, d_M) und (N, d_N) metrische Räume, $K \subset M$ kompakt und $f : K \rightarrow N$ stetig. Dann ist $f(K)$ kompakt.*

Für $M = N = \mathbb{C}$ (mit euklidischer Metrik) ist dies Satz 3.4.2. Der dortige Beweis kann wörtlich übernommen werden.

Ebenso überträgt sich Satz 3.4.3:

Satz 6.4.5 *Sei (M, d) metrischer Raum, $K \subset M$ kompakt und $f : K \rightarrow \mathbb{R}$ stetig. Dann nimmt f auf K sein Minimum und Maximum an.*

Satz 6.4.6 *Alle Normen auf dem \mathbb{R} -Vektorraum \mathbb{R}^p (mit $p \in \mathbb{N}$) sind äquivalent.*

Beweis. Es reicht zu zeigen, daß jede Norm äquivalent zu $\|\cdot\|_\infty$ ist. Sei also $\|\cdot\|$ Norm auf \mathbb{R}^p .

Sei $\{e_1, \dots, e_p\}$ die Standardbasis des \mathbb{R}^p . Mit $\beta := \sum_{j=1}^p \|e_j\|$ ist dann

$$\|x\| = \left\| \sum_{j=1}^p x_j e_j \right\| \leq \sum_{j=1}^p |x_j| \|e_j\| \leq \beta \|x\|_\infty$$

für $x = (x_1, \dots, x_p)^T \in \mathbb{R}^p$. Insbesondere folgt, daß die durch $x \mapsto \|x\|$ gegebene Abbildung von $(\mathbb{R}^p, \|\cdot\|_\infty)$ nach $(\mathbb{R}, |\cdot|)$ stetig ist; vgl. auch Satz 6.3.6. Da $K := \{x \in \mathbb{R}^p : \|x\|_\infty = 1\}$ beschränkt und abgeschlossen und damit kompakt in $(\mathbb{R}^p, \|\cdot\|_\infty)$ ist, existiert deshalb nach Satz 6.3.5 ein $\xi \in K$ mit $\alpha := \|\xi\| = \min_{x \in K} \|x\|$. Für $x \in \mathbb{R}^p \setminus \{0\}$ folgt

$$\|x\| = \left\| \frac{x}{\|x\|_\infty} \right\| \cdot \|x\|_\infty \geq \alpha \|x\|_\infty.$$

Damit erhalten wir $\alpha \|x\|_\infty \leq \|x\| \leq \beta \|x\|_\infty$ für alle $x \in \mathbb{R}^p$. \square

Analog zu Satz 3.4.4 ist das folgende Resultat (und sein Beweis).

Satz 6.4.7 *Seien (M, d_M) und (N, d_N) metrische Räume, M kompakt und $f : M \rightarrow N$ stetig. Dann ist f gleichmäßig stetig.*

6.5 Zusammenhang

Definition 6.5.1 Ein metrischer Raum (M, d) heißt *unzusammenhängend*, falls nichtleere, offene Teilmengen U, V von M existieren, so daß $M = U \cup V$ und $U \cap V = \emptyset$ gilt. Andernfalls heißt (M, d) *zusammenhängend*.

Bemerkungen. 1. Aus $U, V \neq \emptyset$ und $U \cap V = \emptyset$ folgt $U, V \neq M$.

2. Statt zu fordern, daß U, V beide offen sind, kann man auch fordern, daß U, V beide abgeschlossen sind. Denn sind U, V wie in Definition 6.5.1, so sind $A := M \setminus U$ und $B := M \setminus V$ abgeschlossen, nichtleer und es gilt $A \cup B = M$ und $A \cap B = \emptyset$.

3. Eine Teilmenge N von M heißt (un)zusammenhängend, falls dies für den metrischen Raum $(N, d|(N \times N))$ gilt. Es gilt dann, daß N genau dann unzusammenhängend ist, wenn (in M) offene Mengen $U', V' \subset M$ existieren, so daß $U' \cap N \neq \emptyset$, $V' \cap N \neq \emptyset$, $N \subset U' \cup V'$ und $U' \cap V' \cap N = \emptyset$; vgl. Satz 6.2.1.

Satz 6.5.1 Eine nichtleere Teilmenge von \mathbb{R} ist genau dann zusammenhängend, wenn sie ein Intervall ist oder wenn sie aus einem Punkt besteht.

Beweis. Einpunktige Mengen sind offensichtlich zusammenhängend. Sei $I \subset \mathbb{R}$ und I enthalte mindestens zwei Punkte.

“ \Rightarrow ”. Sei I zusammenhängend. Seien $\alpha, \beta \in I$ und $\gamma \in (\alpha, \beta)$. Dann folgt $\gamma \in I$, da sonst mit $U' := (-\infty, \gamma)$ und $V' = (\gamma, \infty)$ gilt, daß $I \subset U' \cup V'$, $U' \cap I \neq \emptyset$, $V' \cap I \neq \emptyset$ und $U' \cap V' = \emptyset$. Mit $a := \inf I$ und $b := \sup I$ folgt hieraus, daß I eines der Intervalle $[a, b]$, $[a, b)$, $(a, b]$ oder (a, b) ist.

“ \Leftarrow ”. Sei I Intervall. Wir nehmen an, daß I nicht zusammenhängend ist, etwa $I = A \cup B$ mit nichtleeren, disjunkten, in I abgeschlossenen Mengen $A, B \subset I$. Sei $a \in A, b \in B$. Ohne Beschränkung der Allgemeinheit sei $a < b$. Wegen $a, b \in I$ folgt dann $[a, b] \subset I$. Sei $c := \sup(A \cap [a, b])$. Dann gilt $c \in A$ (und damit insbesondere $c < b$), da A abgeschlossen. Andererseits ist $(c, b] \subset B$ und wegen der Abgeschlossenheit von B also auch $c \in B$. Insgesamt folgt $c \in A \cap B$, ein Widerspruch. \square

Satz 6.5.2 Seien (M, d_M) und (N, d_N) metrische Räume, M zusammenhängend und $f : M \rightarrow N$ stetig. Dann ist $f(M)$ zusammenhängend.

Beweis. Ohne Beschränkung der Allgemeinheit sei $f(M) = N$. Sei N unzusammenhängend, etwa $N = U \cup V$ mit nichtleeren, disjunkten, offenen Teilmengen U, V von N . Mit $\tilde{U} := f^{-1}(U)$ und $\tilde{V} := f^{-1}(V)$ gilt dann $M \subset \tilde{U} \cup \tilde{V}$, und es sind \tilde{U} und \tilde{V} offen (vgl. Satz 6.3.3, (ii)), nichtleer und disjunkt. Damit ist M unzusammenhängend. \square

Definition 6.5.2 Sei (M, d) metrischer Raum.

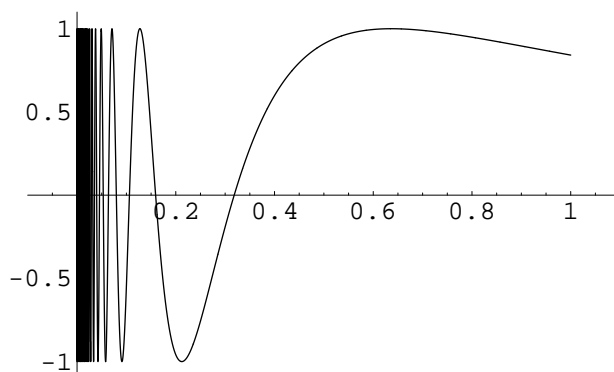
- (i) Sei $a, b \in \mathbb{R}$, $a < b$ und $\gamma : [a, b] \rightarrow M$ stetig. Dann heißt γ *Weg* von $\gamma(a)$ nach $\gamma(b)$. Die Punkte $\gamma(a)$ bzw. $\gamma(b)$ heißen *Anfangs-* bzw. *Endpunkt* des Weges γ .
- (ii) (M, d) heißt *wegzusammenhängend*, wenn für alle $x, y \in M$ ein Weg von x nach y existiert.

Satz 6.5.3 Ein wegzusammenhängender metrischer Raum ist zusammenhängend.

Beweis. Sei (M, d) wegzusammenhängender metrischer Raum. Wir nehmen an, daß M unzusammenhängend ist, etwa $M = U \cup V$ mit nichtleeren, disjunkten, offenen Teilmengen U, V von M . Sei $u \in U, v \in V$. Dann existiert ein Weg $\gamma : [a, b] \rightarrow M$ von u nach v . Es ist dann $\gamma([a, b])$ unzusammenhängend, nach Satz 6.5.2 also $[a, b]$ unzusammenhängend, im Widerspruch zu Satz 6.5.1. \square

Die Umkehrung von Satz 6.5.3 gilt nicht, wie das folgende *Beispiel* zeigt. Sei $M := \{(x, \sin \frac{1}{x}) : x \in (0, 1]\} \cup \{(0, y) : y \in [-1, 1]\}$ und sei d die euklidische Metrik, eingeschränkt auf M . Dann ist (M, d) zusammenhängend, aber nicht wegzusammenhängend.

Die Menge M sieht wie folgt aus:



Definition 6.5.3 Sei V Vektorraum (über \mathbb{K}).

- (i) Für $x, y \in V$ heißt $[x, y] := \{(1-t)x + ty : t \in [0, 1]\}$ *Strecke* (von x nach y).
- (ii) Für $x_1, x_2, \dots, x_n \in V$ heißt $[x_1, x_2, \dots, x_n] := [x_1, x_2] \cup [x_2, x_3] \cup \dots \cup [x_{n-1}, x_n]$ *Streckenzug* (von x_1 nach x_n).

In einem normierten Raum $(V, \|\cdot\|)$ ist für $x, y \in V$ die durch $t \mapsto (1-t)x + ty$ gegebene Funktion $\gamma : [0, 1] \rightarrow V$ stetig, also ein Weg von x nach y . Damit ist die Strecke $[x, y] = \gamma([0, 1])$ das Bild eines Weges. Auch Streckenzüge sind Bilder von Wegen.

Satz 6.5.4 Sei $(V, \|\cdot\|)$ normierter Raum und $M \subset V$ offen und zusammenhängend. Dann existiert für alle $x, y \in M$ ein in M verlaufender Streckenzug von x nach y , d. h., es existieren $x_1, x_2, \dots, x_n \in M$ mit $x_1 = x, x_n = y$ und $[x_1, x_2, \dots, x_n] \subset M$. Insbesondere ist M wegzusammenhängend.

Beweis. Sei $x \in M$. Wir betrachten die Menge U aller $y \in M$, für die ein in M verlaufender Streckenzug von x nach y existiert. Zu zeigen ist, daß $U = M$ gilt.

Wir zeigen zunächst, daß U offen ist. Sei dazu $y \in U$ und $[x_1, x_2, \dots, x_n]$ ein in M verlaufender Streckenzug von x nach y . Da M offen und $y \in U \subset M$, existiert $\varepsilon > 0$ mit $U(y, \varepsilon) \subset M$. Für $z \in U(y, \varepsilon)$ ist dann $[x_1, x_2, \dots, x_n, z]$ ein in M verlaufender Streckenzug von x nach z . Es folgt $U(y, \varepsilon) \subset U$ und damit ist U offen.

Ein analoges Argument zeigt, daß auch $V := M \setminus U$ offen ist. Wegen $U \neq \emptyset, U \cup V = M$ und $U \cap V = \emptyset$ folgt nun aus dem Zusammenhang von M , daß $V = \emptyset$, also $U = M$. \square

7 Differentialrechnung in mehreren Veränderlichen

7.1 Partielle Ableitungen

Definition 7.1.1 Sei $p \in \mathbb{N}$, $U \subset \mathbb{R}^p$ offen, $\xi \in U$ und $f : U \rightarrow \mathbb{R}$. Für $j \in \{1, \dots, p\}$ heißt dann f *partiell differenzierbar* in ξ nach der j -ten Variablen (oder nach x_j), falls

$$\begin{aligned} & \partial_j f(\xi) \\ := & \lim_{x_j \rightarrow \xi_j} \frac{f(\xi_1, \dots, \xi_{j-1}, x_j, \xi_{j+1}, \dots, \xi_p) - f(\xi_1, \dots, \xi_{j-1}, \xi_j, \xi_{j+1}, \dots, \xi_p)}{x_j - \xi_j} \\ = & \lim_{h \rightarrow 0} \frac{f(\xi_1, \dots, \xi_{j-1}, \xi_j + h, \xi_{j+1}, \dots, \xi_p) - f(\xi_1, \dots, \xi_{j-1}, \xi_j, \xi_{j+1}, \dots, \xi_p)}{h} \end{aligned}$$

existiert. Der Grenzwert heißt dann *partielle Ableitung* (von f nach der j -ten Variablen an der Stelle ξ) und wird auch mit $\frac{\partial f}{\partial x_j}(\xi)$ oder $f_{x_j}(\xi)$ bezeichnet.

Definiert man für ein geeignetes Intervall I , welches ξ_j als inneren Punkt enthält, $g : I \rightarrow \mathbb{R}$ durch $g(x) = f(\xi_1, \dots, \xi_{j-1}, x, \xi_{j+1}, \dots, \xi_p)$, so existiert also $\partial_j f(\xi)$ genau dann, wenn $g'(\xi_j)$ existiert, und es gilt dann $\partial_j f(\xi) = g'(\xi_j)$. Man berechnet also $\partial_j f$ durch Differenzieren nach x_j bei festgehaltenem $x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_p$.

Ist $p = 2$ bzw. $p = 3$, so bezeichnet man die Variablen statt mit x_1, x_2 (und x_3) in der Regel mit x, y bzw. x, y, z . Dementsprechend schreibt man auch zum Beispiel $\frac{\partial f}{\partial y}$ und f_y statt $\frac{\partial f}{\partial x_2}$ und f_{x_2} . Analog verfährt man bei anderen Bezeichnungen für die Variablen.

Üblich (aber gelegentlich problematisch) sind auch Schreibweisen wie $\frac{\partial f(x,y)}{\partial x}$ oder $\frac{\partial f}{\partial x}(x, y)$ für $\partial_1 f(x, y)$.

Beispiele. 1. Es sei $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, $f(x, y) = xe^y + \sin(xy)$. Dann ist $\partial_1 f(x, y) = e^y + y \cos(xy)$, $\partial_2 f(x, y) = xe^y + x \cos(xy)$.

2. Es sei $f : \mathbb{R}^p \setminus \{0\} \rightarrow \mathbb{R}$, $f(x) = \frac{1}{\|x\|_2^2}$, also $f(x_1, \dots, x_p) = \frac{1}{x_1^2 + \dots + x_p^2}$. Dann gilt $\partial_j f(x_1, \dots, x_p) = -\frac{2x_j}{(x_1^2 + \dots + x_p^2)^2}$.

Ist $U \subset \mathbb{R}^p$ offen, $k \in \{1, \dots, p\}$, $f : U \rightarrow \mathbb{R}$ und existiert $\partial_k f(x)$ für alle $x \in U$, so ist durch $x \mapsto \partial_k f(x)$ eine Funktion $\partial_k f : U \rightarrow \mathbb{R}$ gegeben. Man kann auch diese wieder auf partielle Differenzierbarkeit untersuchen. Für $\partial_j \partial_k f$ schreiben wir dabei kurz ∂_{jk} (oder $\partial_{j,k}$), bzw. $\frac{\partial^2 f}{\partial x_j \partial x_k}$ oder $f_{x_j x_k}$.

Entsprechend definiert man höhere partielle Ableitungen, etwa

$$\partial_1 \partial_3^2 \partial_2 f = \frac{\partial^4 f}{\partial x_1 \partial x_3^2 \partial x_2} = f_{x_2 x_3 x_3 x_1}.$$

Die Anzahl der auftretenden Ableitungen wird also *Ordnung* der Ableitung bezeichnet, im angegebenen Beispiel handelt es sich also um eine partielle Ableitung 4. Ordnung.

Beispiele. 1. Es sei $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, $f(x, y) = xe^y + \sin(xy)$ wie oben, mit $\partial_1 f(x, y) = e^y + y \cos(xy)$, $\partial_2 f(x, y) = xe^y + x \cos(xy)$. Es ist

$$\begin{aligned}\partial_{11} f(x, y) &= -y^2 \sin(xy), \\ \partial_{21} f(x, y) &= e^y + \cos(xy) - xy \sin(xy), \\ \partial_{12} f(x, y) &= e^y + \cos(xy) - xy \sin(xy) = \partial_{21} f(x, y), \\ \partial_{22} f(x, y) &= xe^y - x^2 \sin(xy).\end{aligned}$$

2. Sei $f : \mathbb{R}^2 \rightarrow \mathbb{R}$,

$$f(x, y) = \begin{cases} xy \frac{x^2 - y^2}{x^2 + y^2} & \text{falls } (x, y) \neq (0, 0), \\ 0 & \text{falls } (x, y) = (0, 0), \end{cases}$$

Dann gilt

$$\partial_1 f(0, y) = \lim_{x \rightarrow 0} \frac{f(x, y)}{x} = \begin{cases} 0 & \text{falls } y = 0, \\ \lim_{x \rightarrow 0} y \frac{x^2 - y^2}{x^2 + y^2} & \text{falls } y \neq 0 \end{cases} = -y$$

und analog $\partial_2 f(x, 0) = x$. Es folgt $\partial_{12} f(0, 0) = 1 \neq -1 = \partial_{21} f(0, 0)$.

Satz 7.1.1 (Schwarz) Sei $p \in \mathbb{N}$, $U \subset \mathbb{R}^p$ offen, $\xi \in U$, $f : U \rightarrow \mathbb{R}$ und $j, k \in \{1, \dots, p\}$. Es sei weiter vorausgesetzt, daß die zweiten partiellen Ableitungen ∂_{jk} und ∂_{kj} in U existieren und in ξ stetig sind. Dann gilt $\partial_{jk}(\xi) = \partial_{kj}(\xi)$.

Beweis. Ohne Beschränkung der Allgemeinheit sei $p = 2$, $j = 1$, $k = 2$. Zu zeigen ist also $\partial_{12}(\xi) = \partial_{21}(\xi)$.

Sei $\varepsilon > 0$ mit $U(\xi, \varepsilon) \subset U$ (wobei $U(\xi, \varepsilon) \subset U$ mit der Norm $\|\cdot\|_\infty$ gebildet sei). Es sei $\xi = (a, b)$ und

$$F : U(\xi, \varepsilon) \rightarrow \mathbb{R}, \quad F(x, y) = f(x, y) - f(x, b) - f(a, y) + f(a, b).$$

Für festes $y \in (b - \varepsilon, b + \varepsilon)$, $y \neq b$, betrachten wir

$$\phi : (a - \varepsilon, a + \varepsilon) \rightarrow \mathbb{R}, \quad \phi(x) = f(x, y) - f(x, b).$$

Es ist dann $F(x, y) = \phi(x) - \phi(a)$. Dann ist ϕ differenzierbar und nach Mittelwertsatz existiert für $x \neq a$ dann α_1 zwischen a und x mit

$$F(x, y) = \phi(x) - \phi(a) = (x - a)\phi'(\alpha_1) = (x - a)(\partial_1 f(\alpha_1, y) - \partial_1 f(\alpha_1, b)).$$

Erneute Anwendung des Mittelwertsatzes liefert β_1 zwischen b und y mit

$$F(x, y) = (x - a)(y - b)\partial_{21} f(\alpha_1, \beta_1).$$

Jetzt betrachten wir für festes $x \in (a - \varepsilon, a + \varepsilon)$, $x \neq a$, die Funktion

$$\psi : (b - \varepsilon, b + \varepsilon) \rightarrow \mathbb{R}, \quad \psi(y) = f(x, y) - f(x, b).$$

Man findet für $y \neq b$ dann β_2 zwischen b und y mit

$$F(x, y) = \psi(y) - \psi(b) = (y - b)\psi'(\beta_2) = (y - b)(\partial_2 f(x, \beta_2) - \partial_2 f(x, b))$$

und anschließend α_2 zwischen a und x mit

$$F(x, y) = (x - a)(y - b)\partial_{12}f(\alpha_2, \beta_2).$$

Es folgt

$$\partial_{21}f(\alpha_1, \beta_1) = \partial_{12}f(\alpha_2, \beta_2).$$

Da für $(x, y) \rightarrow (a, b)$ auch $(\alpha_1, \beta_1) \rightarrow (a, b)$ und $(\alpha_2, \beta_2) \rightarrow (a, b)$ gilt, folgt die Behauptung jetzt aus der Stetigkeit von $\partial_{12}f$ und $\partial_{21}f$ in $\xi = (a, b)$. \square

Satz 7.1.2 Sei $p \in \mathbb{N}$, $U \subset \mathbb{R}^p$ offen, $\xi \in U$ und $f : U \rightarrow \mathbb{R}$. Die partiellen Ableitungen von f mögen existieren und sie seien in $\xi = (\xi_1, \dots, \xi_p)^T$ stetig. Dann gilt

$$\frac{f(x) - f(\xi) - \sum_{j=1}^p \partial_j f(\xi)(x_j - \xi_j)}{\|x - \xi\|} \rightarrow 0$$

für $x = (x_1, \dots, x_p)^T \rightarrow \xi$. Desweiteren ist f stetig in ξ .

Bemerkungen. 1. Wegen der Äquivalenz der Normen in \mathbb{R}^p (Satz 6.4.6) ist es irrelevant, welche Norm in Satz 7.1.2 betrachtet wird.

2. Aus der Existenz der partiellen Ableitungen allein folgt noch nicht die Stetigkeit, man vgl. etwa das Beispiel nach Satz 6.3.5 oder betrachte eine beliebige Funktion $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, die $f(0, t) = f(t, 0) = 0$ für $t \in \mathbb{R}$ erfüllt. Dann gilt $\partial_1 f(0, 0) = \partial_2 f(0, 0) = 0$, aber f braucht nicht stetig sein.

3. Mit den Landau-Symbolen läßt sich die Aussage von Satz 7.1.2 auch in der Form

$$f(x) - f(\xi) = \sum_{j=1}^p \partial_j f(\xi)(x_j - \xi_j) + o(\|x - \xi\|)$$

für $x \rightarrow \xi$ schreiben. Sie läßt sich wie folgt interpretieren: die durch $(h_1, \dots, h_n) \mapsto \sum_{j=1}^p \partial_j f(\xi)h_j$ definierte lineare Funktion $T : \mathbb{R}^p \rightarrow \mathbb{R}$ ist (in der Nähe von 0) eine gute Approximation der durch $h \mapsto f(\xi + h) - f(\xi)$ definierten Funktion, denn

$$f(\xi + h) - f(\xi) = T(h) + o(\|h\|)$$

für $h \rightarrow 0$. Wir können dies auch als

$$f(x) - f(\xi) = T(x - \xi) + o(\|x - \xi\|)$$

für $x \rightarrow \xi$ schreiben. Es ist T die einzige lineare Funktion mit dieser Eigenschaft, da für $L \neq 0$ linear $L(h) \neq o(\|h\|)$, d. h.,

$$\frac{L(h)}{\|h\|} \not\rightarrow 0,$$

für $h \rightarrow 0$ gilt.

Beweis von Satz 7.1.2. Sei $\varepsilon > 0$ mit $U(\xi, \varepsilon) \subset U$ (wobei $U(\xi, \varepsilon) \subset U$ wieder mit der Norm $\|\cdot\|_\infty$ gebildet sei). Für $0 < \|x - \xi\|_\infty < \varepsilon$ gilt dann

$$\begin{aligned} f(x) - f(\xi) &= f(x_1, \dots, x_p) - f(\xi_1, \dots, \xi_p) \\ &= f(x_1, \dots, x_p) - f(\xi_1, \dots, \xi_{p-1}, x_p) \end{aligned}$$

$$\begin{aligned}
& + f(\xi_1, \dots, \xi_{p-1}, x_p) - f(\xi_1, \dots, \xi_{p-2}, x_{p-1}, x_p) \\
& + \dots \\
& + f(\xi_1, x_2, \dots, x_p) - f(x_1, \dots, x_p) \\
= & \sum_{j=1}^p f(\xi_1, \dots, \xi_j, x_{j+1}, \dots, x_p) - f(\xi_1, \dots, \xi_{j-1}, x_j, \dots, x_p) \\
= & \sum_{j=1}^p (x_j - \xi_j) \partial_j f(\xi_1, \dots, \xi_{j-1}, y_j, x_{j+1}, \dots, x_p) \\
= & \sum_{j=1}^p (x_j - \xi_j) \partial_j f(\eta^j)
\end{aligned}$$

mit y_j zwischen ξ_j und x_j , also $\eta^j := (\xi_1, \dots, \xi_{j-1}, y_j, x_{j+1}, \dots, x_p)^T \in U(\xi, \|x - \xi\|_\infty)$ Mit L wie in obiger Bemerkung folgt also

$$f(x) - f(\xi) - L(x - \xi) = \sum_{j=1}^p (x_j - \xi_j) (\partial_j f(\eta^j) - \partial_j f(\xi))$$

und damit

$$\frac{|f(x) - f(\xi) - L(x - \xi)|}{\|x - \xi\|_\infty} \leq \sum_{j=1}^p |\partial_j f(\eta^j) - \partial_j f(\xi)|.$$

Die erste Behauptung folgt jetzt aus der Stetigkeit der $\partial_j f$. Die Stetigkeit von f folgt da $L(x - \xi) \rightarrow 0$ für $x \rightarrow \xi$. \square

Die partiellen Ableitungen beschreiben das Änderungsverhalten einer Funktion in Richtung der Koordinatenachsen. Analoges läßt sich auch für andere Richtungen machen.

Definition 7.1.2 Sei $p \in \mathbb{N}$, $U \subset \mathbb{R}^p$ offen, $\xi \in U$, $f : U \rightarrow \mathbb{R}$ und $v \in \mathbb{R}^p$ mit $\|v\|_2 = 1$. Dann heißt

$$\partial_v f(\xi) := \frac{\partial f}{\partial v}(\xi) := \lim_{t \rightarrow 0} \frac{f(\xi + tv) - f(\xi)}{t}$$

die *Richtungsableitung* von f an der Stelle ξ in Richtung v , falls der Grenzwert existiert.

Ist $\{e_1, \dots, e_p\}$ die Standardbasis des \mathbb{R}^p , so gilt also $\partial_{e_j} = \partial_j$ für alle j .

Satz 7.1.3 Sei $p \in \mathbb{N}$, $U \subset \mathbb{R}^p$ offen, $\xi \in U$, $f : U \rightarrow \mathbb{R}$ und $v \in \mathbb{R}^p$ mit $\|v\|_2 = 1$. Die partiellen Ableitungen von f mögen existieren und sie seien in $\xi = (\xi_1, \dots, \xi_p)^T$ stetig. Dann existiert $\partial_v f(\xi)$ und es gilt

$$\partial_v f(\xi) = \sum_{j=1}^p \partial_j f(\xi) v_j.$$

Beweis. Nach Satz 7.1.2 gilt

$$\lim_{t \rightarrow 0} \frac{f(\xi + tv) - f(\xi) - \sum_{j=1}^p \partial_j f(\xi) v_j t}{t} = 0,$$

und daraus folgt die Behauptung. \square

Definition 7.1.3 Sei $p \in \mathbb{N}$, $U \subset \mathbb{R}^p$ offen, $\xi \in U$ und $f : U \rightarrow \mathbb{R}$. Dann heißt $\text{grad}f(\xi) := (\partial_1 f(\xi), \dots, \partial_p f(\xi))$ der *Gradient* von f an der Stelle ξ , falls die partiellen Ableitungen existieren.

Unter den Voraussetzungen von Satz 7.1.3 gilt dann $\partial_v f(\xi) = \text{grad}f(\xi) \cdot v$. (Hierbei steht “ \cdot ” für das Matrixprodukt der $(1 \times p)$ -Matrix $\text{grad}f(\xi)$ mit der $(p \times 1)$ -Matrix $(v_1, \dots, v_p)^T$. Wir können dies auch als Skalarprodukt auffassen: $\partial_v f(\xi) = \langle \text{grad}f(\xi), v \rangle$.)

In Satz 7.1.2 erhalten wir

$$\frac{f(x) - f(\xi) - \text{grad}f(\xi)(x - \xi)}{\|x - \xi\|} \rightarrow 0$$

für $x \rightarrow \xi$.

Definition 7.1.4 Seien $p, q \in \mathbb{N}$, $U \subset \mathbb{R}^p$ offen, $\xi \in U$ und $f = (f_1, \dots, f_q)^T : U \rightarrow \mathbb{R}^q$. Dann heißt die $(q \times p)$ -Matrix

$$J_f(\xi) := \begin{pmatrix} \partial_1 f_1(\xi) & \dots & \partial_p f_1(\xi) \\ \partial_1 f_2(\xi) & \dots & \partial_p f_2(\xi) \\ \vdots & & \vdots \\ \partial_1 f_q(\xi) & \dots & \partial_p f_q(\xi) \end{pmatrix}$$

die *Funktionalmatrix* (oder *Jacobimatrix*) von f an der Stelle ξ , falls die partiellen Ableitungen existieren.

Die k -te Zeile von $J_f(\xi)$ ist also durch $\text{grad}f_k(\xi)$ gegeben.

Satz 7.1.4 Seien $p, q \in \mathbb{N}$, $U \subset \mathbb{R}^p$ offen, $\xi \in U$ und $f = (f_1, \dots, f_q)^T : U \rightarrow \mathbb{R}^q$. Die partiellen Ableitungen der f_k mögen existieren und sie seien in $\xi = (\xi_1, \dots, \xi_p)^T$ stetig. Dann gilt

$$\frac{f(x) - f(\xi) - J_f(\xi)(x - \xi)}{\|x - \xi\|} \rightarrow 0$$

für $x \rightarrow \xi$.

Beweis. Die Behauptung folgt unmittelbar aus Satz 7.1.2 (bzw. der auf Definition 7.1.3 folgenden Bemerkung), da die k -te Koordinate des Zählers durch $f_k(x) - f_k(\xi) - \text{grad}f_k(\xi)(x - \xi)$ gegeben ist. \square

Für $U \subset \mathbb{R}^p$ offen und $n \in \mathbb{N}$ bezeichnet man die Menge aller Funktionen von U nach \mathbb{R}^q , für die alle partiellen Ableitungen bis zur Ordnung n existieren und stetig sind, mit $C^n(U, \mathbb{R}^q)$ oder, falls der Zielraum aus dem Zusammenhang klar ist, auch mit $C^n(U)$. Für solche Funktionen sind nach dem Satz von Schwarz die partiellen Ableitungen bis zur Ordnung n unabhängig von der Reihenfolge der Differentiation. Weiter setzt man $C^\infty(U, \mathbb{R}^q) := \bigcap_{n=0}^{\infty} C^n(U, \mathbb{R}^q)$.

7.2 Differenzierbarkeit

Aus LA I, §2.2, ist bekannt, daß $(q \times p)$ -Matrizen zu linearen Abbildungen von \mathbb{R}^p nach \mathbb{R}^q korrespondieren. Insbesondere entspricht die Funktionalmatrix einer linearen Abbildung. Beachtet man noch, daß lineare Abbildungen von \mathbb{R}^p nach \mathbb{R}^q stetig sind (siehe §6.3), so führt dies auf folgende Definition.

Definition 7.2.1 Seien $(V, \|\cdot\|_V)$, $(W, \|\cdot\|_W)$ Banachräume, $U \subset V$ offen, $\xi \in U$ und $f : U \rightarrow W$. Dann heißt f differenzierbar in ξ , falls $T \in L(V, W)$ existiert, so daß

$$\lim_{x \rightarrow \xi} \frac{f(x) - f(\xi) - T(x - \xi)}{\|x - \xi\|_V} = \lim_{h \rightarrow 0} \frac{f(\xi + h) - f(\xi) - T(h)}{\|h\|_V} = 0.$$

Die Abbildung T ist dann eindeutig bestimmt (siehe die 2. Bemerkung unten) und heißt (*totale*) *Ableitung* (oder *Fréchet-Ableitung*) von f an der Stelle ξ . Sie wird mit $Df(\xi)$ bezeichnet.

Die Funktion f heißt (*total*) *differenzierbar*, falls f in jedem Punkt von U differenzierbar ist. Ist dann die durch $x \mapsto Df(x)$ gegebene Funktion $Df : U \rightarrow L(V, W)$ stetig, heißt f *stetig differenzierbar*.

Bemerkungen. 1. Mit $L(V, W)$ hatten wir den Vektorraum der stetigen linearen Abbildungen von V nach W bezeichnet; vgl. §6.3.

2. Wir zeigen, daß T in obiger Definition eindeutig bestimmt ist. Dazu nehmen wir an, daß $T_1, T_2 \in L(V, W)$ mit $T_1 \neq T_2$ existieren, die die in der Definition angegebene Eigenschaft haben. Dann gilt

$$\lim_{h \rightarrow 0} \frac{T_1(h) - T_2(h)}{\|h\|_V} = \lim_{h \rightarrow 0} \frac{(T_1 - T_2)(h)}{\|h\|_V} = 0.$$

Andererseits existiert $y \in V$ mit $T_1(y) \neq T_2(y)$. Es folgt

$$\frac{(T_1 - T_2)(ty)}{\|ty\|_V} = \frac{t(T_1 - T_2)(y)}{t\|y\|_V} = \frac{(T_1 - T_2)(y)}{\|y\|_V}$$

für $t \in \mathbb{R}_+$ und damit

$$\lim_{t \rightarrow 0} \frac{(T_1 - T_2)(ty)}{\|ty\|_V} = \frac{(T_1 - T_2)(y)}{\|y\|_V} \neq 0,$$

ein Widerspruch.

3. Die Interpretation der in obiger Definition gegebenen Eigenschaft ist analog zu der von Satz 7.1.2; vgl. Bemerkung 3 dort. Die lineare Abbildung $Df(\xi)$ ist (in der Nähe von 0) eine gute Approximation der durch $h \mapsto f(\xi + h) - f(\xi)$ gegebenen Funktion. Denn ist $R(h) := f(\xi + h) - f(\xi) - Df(\xi)(h)$, so gilt $R(h) = o(\|h\|_V)$ für $h \rightarrow 0$. Es ist $Df(\xi)$ die einzige lineare Abbildung mit dieser Eigenschaft; vgl. die vorige Bemerkung.

4. Im allgemeinen hängt Df von den Normen $\|\cdot\|_V, \|\cdot\|_W$ ab. Dies gilt aber nicht, falls $V = \mathbb{R}^p, W = \mathbb{R}^q$, da dort alle Normen äquivalent sind (Satz 6.4.6).

Beispiel. Sei $f \in L(V, W)$. Dann ist f differenzierbar und es gilt $Df(\xi) = f$ für alle $\xi \in V$.

Satz 7.2.1 *Differenzierbare Funktionen sind stetig.*

Beweis. Seien U, V, W, f wie in Definition 7.2.1 und sei f differenzierbar in $\xi \in U$. Für $x \in U$ sei $R(x) := f(x) - f(\xi) - Df(\xi)(x - \xi)$. Nach Definition der Differenzierbarkeit gilt dann $R(x)/\|x - \xi\|_V \rightarrow 0$ für $x \rightarrow \xi$. Es folgt $R(x) \rightarrow 0$ für $x \rightarrow \xi$. Da $Df(\xi)$ stetig ist, folgt auch $Df(\xi)(x - \xi) \rightarrow 0$ für $x \rightarrow \xi$. Insgesamt folgt $f(x) \rightarrow f(\xi)$ für $x \rightarrow \xi$. Damit ist f stetig in ξ . \square

Man beachte, daß im Beweis die Stetigkeit von $Df(\xi)$ benutzt wird. Dies ist ein Grund dafür, in Definition 7.2.1 T als linear *und stetig* zu fordern.

Satz 7.2.2 *Seien $p, q \in \mathbb{N}$, $U \subset \mathbb{R}^p$ offen, $\xi \in U$ und $f = (f_1, \dots, f_q)^T : U \rightarrow \mathbb{R}^q$. Dann sind äquivalent:*

- (i) f ist differenzierbar in ξ .
- (ii) Alle partiellen Ableitungen $\partial_k f_j(\xi)$ existieren und es gilt

$$\frac{f(x) - f(\xi) - J_f(\xi)(x - \xi)}{\|x - \xi\|} \rightarrow 0$$

für $x \rightarrow \xi$.

Es gilt dann $Df(\xi)(h) = J_f(\xi)(h)$ für alle $h \in \mathbb{R}^p$, d. h., die Matrix $J_f(\xi)$ entspricht der linearen Abbildung $Df(\xi)$ bzgl. der Standardbasis.

Beweis. (ii) \Rightarrow (i). Diese Richtung ist trivial, denn offensichtlich hat die durch $T(h) = J_f(\xi)h$ definierte Funktion $T \in L(V, W)$ die in Definition 7.2.1 verlangte Eigenschaft.

(i) \Rightarrow (ii). Sei $A = (a_{ij})_{\substack{1 \leq i \leq q \\ 1 \leq j \leq p}}$ die zu $Df(\xi)$ gehörige Matrix (bzgl. der Standardbasen $\{e_1, e_2, \dots\}$). Dann folgt

$$\frac{f(\xi + h) - f(\xi) - Ah}{\|h\|} \rightarrow 0$$

für $h \rightarrow 0$ und damit

$$\frac{f_j(\xi + te_k) - f_j(\xi) - a_{jk}t}{t} \rightarrow 0$$

für $t \rightarrow 0$. Hieraus folgt, daß die partiellen Ableitungen $\partial_k f_j(\xi)$ existieren und daß $a_{jk} = \partial_k f_j(\xi)$ und $A = J_f(\xi)$ gilt. \square

Aus der totalen Differenzierbarkeit in ξ folgt also mit Satz 7.2.2 die partielle Differenzierbarkeit in ξ , aber nicht umgekehrt (vgl. die Beispiele nach Satz 7.1.2, wo die partiellen Ableitungen existieren, aber die Funktion sogar unstetig ist). Existieren aber die partiellen Ableitungen in einer Umgebung von ξ und sind sie in ξ stetig, so ist f in ξ total differenzierbar; vgl Satz 7.1.4 und 7.2.2.

Für stetige Differenzierbarkeit sind die Verhältnisse einfacher.

Satz 7.2.3 *Seien $p, q \in \mathbb{N}$, $U \subset \mathbb{R}^p$ offen und $f = (f_1, \dots, f_q)^T : U \rightarrow \mathbb{R}^q$. Dann ist f genau dann stetig differenzierbar, wenn $f \in C^1(U, \mathbb{R}^q)$, d. h., wenn alle partiellen Ableitungen der Koordinatenfunktionen von f existieren und stetig sind.*

Beweis. Da aus der Stetigkeit der partiellen Ableitungen die Differenzierbarkeit folgt, können wir f als differenzierbar annehmen. Für $h \in \mathbb{R}^p$ und $x, \xi \in U$ ist dann $(Df(x) - Df(\xi))(h) = (J_f(x) - J_f(\xi))(h)$. Legen wir in $\mathbb{R}^p, \mathbb{R}^q$ die Norm $\|\cdot\|_\infty$ zugrunde, so erhalten wir (vgl. das Beispiel nach Satz 6.3.6)

$$\|Df(x) - Df(\xi)\|_{L(\mathbb{R}^p, \mathbb{R}^q)} = \max_j \sum_{k=1}^p |\partial_k f_j(x) - \partial_k f_j(\xi)|.$$

Hieraus folgt leicht die Behauptung. \square

Die üblichen Regeln für die Differenzierbarkeit von $f + g$, und $\lambda \cdot f$, wobei $f, g : U \rightarrow W$ und $\lambda \in \mathbb{K}$, übertragen sich unmittelbar. Entsprechendes gilt auch für $f \cdot g$ und f/g , falls $W = \mathbb{R}$ ist; vgl. auch Satz 6.3.4. Wir formulieren die Kettenregel:

Satz 7.2.4 *Seien $(V, \|\cdot\|_V)$, $(W, \|\cdot\|_W)$, $(X, \|\cdot\|_X)$ Banachräume, $U \subset V$ offen, $U' \subset W$ offen, $\xi \in U$, $g : U \rightarrow W$ differenzierbar in ξ , $g(U) \subset U'$ und $f : U' \rightarrow X$ differenzierbar in $g(\xi)$. Dann ist $f \circ g$ differenzierbar in ξ und*

$$D(f \circ g)(\xi) = Df(g(\xi)) \circ Dg(\xi).$$

Beweis. Es ist

$$R_g(h) := g(\xi + h) - g(\xi) - Dg(\xi)(h) = o(\|h\|_V)$$

für $h \rightarrow 0$ und

$$R_f(k) := f(g(\xi) + k) - f(g(\xi)) - Df(g(\xi))(k) = o(\|k\|_W)$$

für $k \rightarrow 0$. Mit $\Delta(h) := g(\xi + h) - g(\xi) = Dg(\xi)(h) + R_g(h)$ folgt

$$\begin{aligned} & f(g(\xi + h)) - f(g(\xi)) \\ &= f(g(\xi) + \Delta(h)) - f(g(\xi)) \\ &= Df(g(\xi))(\Delta(h)) + R_f(\Delta(h)) \\ &= (Df(g(\xi)) \circ Dg(\xi))(h) + \underbrace{Df(g(\xi))(R_g(h)) + R_f(\Delta(h))}_{=: R(h)}. \end{aligned}$$

Zu zeigen ist, daß $R(h) = o(\|h\|_V)$ für $h \rightarrow 0$. Wegen

$$\|Df(g(\xi))(R_g(h))\|_X \leq \|Df(g(\xi))\|_{L(W, X)} \|R_g(h)\|_W = o(\|h\|_V)$$

für $h \rightarrow 0$ müssen wir nur noch $\|R_f(\Delta(h))\|_X = o(\|h\|_V)$ für $h \rightarrow 0$ zeigen. Dies folgt aber, da $\Delta(h) = O(\|h\|_V)$ und damit $\Delta(h) \rightarrow 0$ für $h \rightarrow 0$, und da

$$\frac{\|R_f(\Delta(h))\|_X}{\|h\|_V} = \underbrace{\frac{\|R_f(\Delta(h))\|_X}{\|\Delta(h)\|_W}}_{=o(1)} \cdot \underbrace{\frac{\|\Delta(h)\|_W}{\|h\|_V}}_{=O(1)}$$

falls $\Delta(h) \neq 0$. \square

Wir betrachten jetzt Abbildungen $f : I \rightarrow V$ wobei I Intervall und V Banachraum. Hier können wir die Ableitung wieder – wie in Analysis I – über den Differenzenquotienten bilden.

Definition 7.2.2 Sei I Intervall, $(V, \|\cdot\|)$ Banachraum, $f : I \rightarrow V$ und $\xi \in I$. Dann heißt

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

Ableitung von f an der Stelle ξ , falls der Grenzwert existiert.

Andererseits können wir ein offenes Intervall (oder das Innere eines beliebigen Intervalls) als offene Teilmenge des Banachraums \mathbb{R} auffassen und die Fréchet-Ableitung betrachten.

Satz 7.2.5 Seien I, V, f wie in Definition 7.2.2 und $\xi \in \text{int}(I)$. Dann existiert $f'(\xi)$ genau dann, wenn $Df(\xi)$ existiert, und es gilt dann $f'(\xi) = Df(\xi)(1)$.

Beweis. Existiert $f'(\xi)$, so gilt

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

und damit folgt, daß $Df(\xi)$ existiert und durch $h \mapsto hf'(\xi)$ gegeben ist. Es folgt also $f'(\xi) = Df(\xi)(1)$. Analog folgt aus der Existenz von $Df(\xi)$ die von $f'(\xi)$. \square

Mit den Bezeichnungen von Satz 7.2.5 gilt $\|f'(\xi)\| = \|Df(\xi)\|_{L(\mathbb{R}, V)}$, falls $f'(\xi)$ (und damit $Df(\xi)$) existiert.

Der Mittelwertsatz der Differentialrechnung (Satz 4.2.2) gilt nur für reellwertige Funktionen. Für komplexwertige Funktionen hatten wir nur eine Ungleichung erhalten (Satz 4.2.5). Diese können wir auf die Banachraumsituation übertragen.

Satz 7.2.6 Sei $(V, \|\cdot\|)$ Banachraum und seien $f : [a, b] \rightarrow V$ und $\varphi : [a, b] \rightarrow \mathbb{R}$ stetig und in (a, b) differenzierbar. Es gelte $\|f'(t)\| \leq \varphi'(t)$ für $t \in (a, b)$. Dann gilt $\|f(b) - f(a)\| \leq \varphi(b) - \varphi(a)$.

Gilt

$$M := \sup_{t \in (a, b)} \|f'(t)\| < \infty,$$

so kann man hier $\varphi(t) = Mt$ wählen und erhält

$$\|f(b) - f(a)\| \leq M(b - a).$$

Diese Abschätzung heißt auch *Mittelwertungleichung*.

Beweis von Satz 7.2.6. Für $\varepsilon > 0$ betrachten wir

$$A := \{t \in [a, b] : \|f(t) - f(a)\| \leq \varphi(t) - \varphi(a) + \varepsilon(t - a) + \varepsilon\}.$$

Wir werden zeigen, daß $A = [a, b]$ gilt. Insbesondere gilt dann $b \in A$, und mit $\varepsilon \rightarrow 0$ folgt dann die Behauptung.

Um zu zeigen, daß $A = [a, b]$ gilt, beachten wir zunächst, daß A abgeschlossen ist. Dies folgt unmittelbar aus der Stetigkeit von f und φ . Außerdem folgt aus der Stetigkeit, daß eine Umgebung von a in A enthalten ist.

Sei nun $c := \sup\{t \in [a, b] : [a, t] \subset A\}$. Dann gilt $c \in A$ und $c \in (a, b]$. Wir nehmen an, daß $c < b$ gilt. Nach Definition der Differenzierbarkeit existiert $\delta > 0$ mit

$$\|f(t) - f(c) - f'(c)(t - c)\| < \frac{\varepsilon}{2}(t - c)$$

und

$$|\varphi(t) - \varphi(c) - \varphi'(c)(t - c)| < \frac{\varepsilon}{2}(t - c)$$

für $c \leq t \leq c + \delta$ und damit

$$\begin{aligned} \|f(t) - f(c)\| &\leq \|f'(c)\|(t - c) + \frac{\varepsilon}{2}(t - c) \\ &\leq |\varphi'(c)|(t - c) + \frac{\varepsilon}{2}(t - c) \\ &\leq \varphi(t) - \varphi(c) + \frac{\varepsilon}{2}(t - c) + \frac{\varepsilon}{2}(t - c) \\ &= \varphi(t) - \varphi(c) + \varepsilon(t - c). \end{aligned}$$

Außerdem gilt

$$\|f(c) - f(a)\| \leq \varphi(c) - \varphi(a) + \varepsilon(c - a) + \varepsilon$$

wegen $c \in A$. Es folgt

$$\|f(t) - f(a)\| \leq \|f(t) - f(c)\| + \|f(c) - f(a)\| \leq \varphi(t) - \varphi(a) + \varepsilon(t - a) + \varepsilon$$

für $c \leq t \leq c + \delta$ und damit $[a, t + \delta] \subset A$, ein Widerspruch. \square

Satz 7.2.7 Seien $(V, \|\cdot\|_V)$, $(W, \|\cdot\|_W)$ Banachräume, $U \subset V$ offen, $a, b \in U$ mit $[a, b] \in U$ und $f : U \rightarrow W$ differenzierbar. Ist $M := \sup_{x \in [a, b]} \|Df(x)\|_{L(V, W)} < \infty$, so gilt $\|f(b) - f(a)\|_W \leq M\|b - a\|_V$.

Beweis. Sei $\gamma : [0, 1] \rightarrow [a, b]$, $t \mapsto a + t(b - a)$, und $g := f \circ \gamma$. Dann ist g differenzierbar in $(0, 1)$ mit $Dg(t) = Df(\gamma(t)) \circ D\gamma(t)$ und damit

$$g'(t) = Dg(t)(1) = Df(\gamma(t))(D\gamma(t)(1)) = Df(\gamma(t))(\gamma'(t)) = Df(\gamma(t))(b - a).$$

Es folgt

$$\|g'(t)\|_W \leq \|Df(\gamma(t))\|_{L(V, W)} \|b - a\|_V \leq M\|b - a\|_V.$$

Mit Satz 7.2.6 folgt

$$\|f(b) - f(a)\|_W = \|g(1) - g(0)\|_W \leq M\|b - a\|_V. \quad \square$$

Auch die Ungleichung in Satz 7.2.7 nennt man Mittelwertungleichung. Eine typische Anwendung ist der folgende Satz (vgl. Satz 4.3.1).

Satz 7.2.8 Seien $(V, \|\cdot\|_V)$, $(W, \|\cdot\|_W)$ Banachräume, $U \subset V$ offen und zusammenhängend, $f : U \rightarrow W$ differenzierbar und $Df(x) = 0$ für alle $x \in U$. Dann ist f konstant, d. h., es existiert $c \in W$ mit $f(x) = c$ für alle $x \in U$.

Beweis. Sei $x \in U$. Nach Satz 7.2.7 gilt dann $f(y) = f(x)$ für alle $y \in U$ mit $[x, y] \subset U$. Daraus folgt, daß $f(z) = f(x)$ für alle $z \in U$ gilt, die mit x durch einen in U verlaufenden Streckenzug verbunden werden können. Dies sind nach Satz 6.5.4 aber alle $z \in U$. \square

Auch der Begriff des Riemann-Integrals läßt sich ohne weiteres auf Funktionen $f : [a, b] \rightarrow V$, mit $(V, \|\cdot\|)$ Banachraum, übertragen. Für eine Zerlegung $Z = \{x_0, x_1, \dots, x_n\}$ von $[a, b]$ mit Stützstellen $\xi = (\xi_1, \dots, \xi_n)$ heißt $S(f, Z, \xi) = \sum_{k=1}^n f(\xi_k) |I_k|$ wieder Riemannsche Summe und f heißt Riemann-integrierbar, falls gilt:

$$\exists S \in \mathbb{C} \forall \varepsilon \in \mathbb{R}_+ \exists \delta \in \mathbb{R}_+ \forall \text{ Zerlegungen } Z \text{ von } I \forall \text{ Stützstellen } \xi \text{ zu } Z : \\ |Z| < \delta \Rightarrow \|S(f, Z, \xi) - S\| < \varepsilon.$$

Dieses S ist dann wieder eindeutig und wird mit $\int_a^b f$ oder $\int_a^b f(x) dx$ bezeichnet. Wie in §5.1 gilt wieder das Cauchyriterium und wie dort zeigt man, daß stetige Funktionen integrierbar sind. Man erhält für stetiges f auch wieder

$$\left\| \int_a^b f \right\| \leq \int_a^b \|f\| \leq (b-a) \max_{x \in [a, b]} \|f(x)\|.$$

Für $(V, \|\cdot\|_V)$, $(W, \|\cdot\|_W)$ Banachräume, $U \subset V$ offen, $a, b \in U$ mit $[a, b] \in U$, $f : U \rightarrow W$ stetig differenzierbar, und $\gamma : [0, 1] \rightarrow [a, b]$, $t \mapsto a + t(b-a)$, gilt dann

$$f(b) - f(a) = \int_0^1 (f \circ \gamma)'(t) dt = \int_0^1 Df(\gamma(t))(b-a) dt.$$

Hieraus kann man (für *stetig* differenzierbares f) wieder die Aussage von Satz 7.2.7 gewinnen.

7.3 Höhere Ableitungen und Taylorformel

Seien $(V, \|\cdot\|_V)$, $(W, \|\cdot\|_W)$ Banachräume, $U \subset V$ offen und $f : U \rightarrow W$ differenzierbar. Ist dann $Df : U \rightarrow L(V, W)$ wieder differenzierbar, so heißt $D(Df)$ *zweite Ableitung* und wird mit D^2f bezeichnet. Entsprechend definiert man die höheren Ableitungen $D^n f = D(D^{n-1}f)$, $n \geq 3$, falls sie existieren. Wir nennen dann f *n-mal differenzierbar* und, falls $D^n f$ auch stetig ist, *n-mal stetig differenzierbar*.

Es ist $D^2f : U \rightarrow L(V, L(V, W))$, für $\xi \in U$ also $D^2f(\xi) \in L(V, L(V, W))$. Für $x, y \in V$ gilt also $D^2f(\xi)(x) \in L(V, W)$ und $(D^2f(\xi)(x))(y) \in W$.

Der Vektorraum $L(V, L(V, W))$ kann in offensichtlicher Weise mit dem Vektorraum der stetigen, bilinearen Abbildungen von $V \times V$ nach W identifiziert werden. Aus diesem Grunde werden wir im folgenden $D^2f(\xi)$ als stetige, bilineare Abbildung von $V \times V$ nach W betrachten und schreiben daher $D^2f(\xi)(x, y)$ statt $(D^2f(\xi)(x))(y)$.

Analog betrachten wir die höheren Ableitungen als stetige, multilineare Abbildungen und schreiben daher $D^n f(\xi)(x_1, \dots, x_n)$ statt $(\dots (D^n f(\xi)(x_1)) \dots)(x_n)$.

Wir wollen jetzt die Resultate aus §4.4 und §5.4 über die Taylorentwicklung übertragen. Für $\xi \in U$ und $x \in V$ setzen wir $h := x - \xi$ und definieren das *n-te*

Taylorpolynom durch

$$T_n(x) := f(\xi) + Df(\xi)(h) + \frac{1}{2!}D^2f(\xi)(h, h) + \dots + \frac{1}{n!}D^n f(\xi) \underbrace{(h, h, \dots, h)}_{n\text{-mal}}$$

und das Restglied $R_n(x)$ wieder durch

$$R_n(x) := f(x) - T_n(x).$$

Wir übertragen Satz 4.4.3. (Alternativ könnten wir auch Satz 5.4.1 entsprechend verallgemeinern.)

Satz 7.3.1 Sei $(V, \|\cdot\|_V)$ Banachraum, $U \subset V$ offen, $n \in \mathbb{N}$, $\xi, x \in U$ mit $[\xi, x] \subset U$ und $h := x - \xi$. Die Funktion $f : U \rightarrow \mathbb{R}$ sei $(n+1)$ -mal differenzierbar. Dann existiert $\eta \in [\xi, x]$ mit

$$R_n(x) = \frac{1}{(n+1)!} D^{n+1}f(\eta) \underbrace{(h, h, \dots, h)}_{(n+1)\text{-mal}}.$$

Hilfssatz 7.3.1 Seien $(V, \|\cdot\|_V)$, $(W, \|\cdot\|_W)$ Banachräume, $h \in V$ und $A : L(V, W) \rightarrow W$, $A(T) = T(h)$. Dann ist A differenzierbar und $DA(T) = A$ für alle $T \in L(V, W)$.

Beweis. A ist linear und stetig, und nach Beispiel zu Definition 7.2.1 folgt die Behauptung. \square

Beweis von Satz 7.3.1 Da $[\xi, x] \subset U$ und U offen existiert $\delta > 0$ mit $\xi + th = (1-t)\xi + tx \in U$ für $t \in I := (-\delta, 1 + \delta)$. Wir betrachten $\gamma : I \rightarrow U$, $\gamma(t) = \xi + th = (1-t)\xi + tx$, und $g := f \circ \gamma$. Dann ist g nach Kettenregel $(n+1)$ -mal differenzierbar und es gilt nach Satz 4.4.3

$$g(1) = g(0) + g'(0) + \frac{1}{2!}g''(0) + \dots + \frac{1}{n!}g^{(n)}(0) + \frac{1}{(n+1)!}g^{(n+1)}(s)$$

mit $s \in (0, 1)$. Nach Kettenregel und Satz 7.2.5 gilt

$$g'(t) = Dg(t)(1) = (Df(\gamma(t)) \circ D\gamma(t))(1) = Df(\gamma(t))(\gamma'(t)) = Df(\gamma(t))(h)$$

für $t \in I$. Mit A wie in Hilfssatz 7.3.1 gilt also

$$g'(t) = A(Df(\gamma(t))) = (A \circ Df \circ \gamma)(t).$$

Es folgt

$$\begin{aligned} g''(t) &= Dg'(t)(1) \\ &= D(A \circ Df \circ \gamma)(t)(1) \\ &= \left(DA \underbrace{(Df \circ \gamma)(t)}_T \circ (D^2f(\gamma(t))) \circ D\gamma(t) \right) (1) \\ &= (A \circ D^2f(\gamma(t)) \circ D\gamma(t))(1) \\ &= (A \circ D^2f(\gamma(t)))(\gamma'(t)) \\ &= (D^2f(\gamma(t))(\gamma'(t)))(h) \\ &= (D^2f(\gamma(t))(h))(h) \\ &= D^2f(\gamma(t))(h, h) \end{aligned}$$

Ebenso zeigt man, daß

$$g^{(k)}(t) = D^k f(\gamma(t))(h, h, \dots, h)$$

für $k \leq n + 1$ gilt. Hieraus folgt sofort die Behauptung. \square

Aus Satz 7.3.1 folgt, daß für $(n + 1)$ -mal stetig differenzierbares f

$$R_n(x) = O(\|x - \xi\|^{n+1})$$

für $x \rightarrow \xi$ gilt. Wie in §5.4, insbesondere Satz 5.4.2, erhält man auch eine schwächere Abschätzung, falls f nur n -mal stetig differenzierbar ist.

Satz 7.3.2 Sei $(V, \|\cdot\|_V)$ Banachraum, $U \subset V$ offen, $n \in \mathbb{N}$, $\xi \in U$, $x \in V$ mit $[\xi, x] \subset U$ und $h := x - \xi$. Die Funktion $f : U \rightarrow \mathbb{R}$ sei n -mal stetig differenzierbar. Dann gilt

$$R_n(x) = o(\|x - \xi\|^n)$$

für $x \rightarrow \xi$.

Wir betrachten den Spezialfall $V = \mathbb{R}^p$, also $U \subset \mathbb{R}^p$ offen, $f : U \rightarrow \mathbb{R}$ differenzierbar und $\xi \in U$. Mit $h = (h_1, \dots, h_p)^T \in \mathbb{R}^p$ ist dann

$$Df(\xi)(h) = \text{grad} f(\xi)h = \sum_{j=1}^p \partial_j f(\xi)h_j.$$

Die Funktion $Df : U \rightarrow L(V, \mathbb{R})$ ist genau dann stetig differenzierbar, wenn dies für die Funktion $\text{grad} f : U \rightarrow \mathbb{R}^p$ gilt. Letzteres ist genau dann der Fall, wenn dies für alle Koordinatenfunktionen von $\text{grad} f$, also alle partiellen Ableitungen von f , gilt. Dies ist (nach Satz 7.2.3) genau dann der Fall, wenn alle zweiten partiellen Ableitungen existieren und stetig sind, d. h., wenn $f \in C^2(U, \mathbb{R})$.

Definition 7.3.1 Sei $U \subset \mathbb{R}^p$ offen, $f \in C^2(U, \mathbb{R})$ und $\xi \in U$. Dann heißt

$$H_f(\xi) := (\partial_{ij} f(\xi))_{\substack{1 \leq i \leq p \\ 1 \leq j \leq p}}$$

Hesse-Matrix von f an der Stelle ξ .

Für $f \in C^2(U, \mathbb{R})$ gilt dann

$$D^2 f(\xi)(h, h) = h^T H_f(\xi)h = \sum_{i,j=1}^n \partial_{ij} f(\xi)h_i h_j.$$

Man beachte, daß die Hesse-Matrix nach dem Satz von Schwarz symmetrisch ist.

7.4 Lokale Extrema

In Definition 4.2.1 haben wir für Funktionen $f : M \rightarrow \mathbb{R}$ mit $M \subset \mathbb{C}$ definiert, wann f ein lokales bzw. globales Extremum hat. Diese Begriffe können wörtlich auf den Fall, daß M metrischer Raum ist, übertragen werden.

Für $f \in C^2(I, \mathbb{R})$ mit einem Intervall I und $\xi \in \text{int}(I)$ hatten wir gezeigt:

$$\begin{array}{l} f'(\xi) = 0 \text{ und } f''(\xi) > 0 \text{ (bzw. } < 0) \\ \xrightarrow{\text{Satz 4.3.3}} f \text{ hat lokales Minimum (bzw. Maximum) in } \xi \\ \xrightarrow{\text{Satz 4.2.1}} f'(\xi) = 0 \end{array}$$

Wir wollen diese Aussagen jetzt auf Funktionen $f \in C^2(U, \mathbb{R})$ mit $U \subset \mathbb{R}^p$ übertragen. Dabei ersetzen wir f' durch den Gradienten und f'' durch die Hesse-Matrix.

Satz 7.4.1 Sei $U \subset \mathbb{R}^p$ offen und $f \in C^1(U, \mathbb{R})$ habe in $\xi \in U$ ein lokales Extremum. Dann gilt $\text{grad}f(\xi) = 0$.

Beweis. Für $j \in \{1, \dots, p\}$ hat die durch $t \mapsto f(\xi + te_j)$ definierte Funktion g_j ein lokales Extremum in 0. Nach Satz 4.3.3 folgt $\partial_j f(\xi) = g_j'(0) = 0$. \square

Es sei angemerkt, daß ein entsprechender Satz auch für Funktionen gilt, die auf offenen Teilmengen von Banachräumen definiert sind. (Beweis selbst).

Beispiel. Sei $K := \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \leq 4, y \geq 0\}$ und $f : K \rightarrow \mathbb{R}$, $f(x, y) = x^2 + y^2 - \sqrt{3}x - y$. Man bestimme $\max_{(x,y) \in K} f(x, y)$ und $\min_{(x,y) \in K} f(x, y)$. (Die Existenz von Maximum und Minimum folgt, da K kompakt und f stetig.)

Für ein im Inneren von K gelegenes Extremum (x, y) folgt nach Satz 7.4.1, daß $(0, 0) = \text{grad}f(x, y) = (2x - \sqrt{3}, 2y - 1)$, also $(x, y) = (\frac{1}{2}\sqrt{3}, \frac{1}{2})$.

Wir untersuchen jetzt den Rand von K auf Extrema. Mit $\Gamma_1 := \{(x, 0) : -2 \leq x \leq 2\}$ und $\Gamma_2 := \{(2 \cos t, 2 \sin t) : 0 \leq t \leq \pi\}$ gilt $\partial K = \Gamma_1 \cup \Gamma_2$. Desweiteren ist $\max_{(x,y) \in \Gamma_1} f(x, y) = \max_{-2 \leq x \leq 2} f(x, 0) = \max_{-2 \leq x \leq 2} g(x)$ mit $g(x) := x^2 - \sqrt{3}x$ und $\max_{(x,y) \in \Gamma_2} f(x, y) = \max_{0 \leq t \leq \pi} f(2 \cos t, 2 \sin t) = \max_{0 \leq t \leq \pi} h(t)$ mit $h(t) := 4 - 2\sqrt{3} \cos t - 2 \sin t$. Analoges gilt für die Minima.

Wegen $g'(x) = 2x - \sqrt{3}$ hat g' nur die Nullstelle $\frac{1}{2}\sqrt{3}$ und damit sind Extremwerte von g nur in $-2, \frac{1}{2}\sqrt{3}$ und 2 möglich. Wegen $h'(t) = 2\sqrt{3} \sin t - 2 \cos t$ hat h' in $(0, \pi)$ nur die Nullstelle $\frac{\pi}{6}$ und damit sind Extremwerte von h nur in $0, \frac{\pi}{6}$ und π möglich.

Insgesamt erhält man 5 Kandidaten für mögliche Extremstellen: $\xi_1 = (\frac{1}{2}\sqrt{3}, \frac{1}{2})$, $\xi_2 = (\frac{1}{2}\sqrt{3}, 0)$, $\xi_3 = (2 \cos \frac{\pi}{6}, 2 \sin \frac{\pi}{6}) = (\sqrt{3}, 1)$, $\xi_4 = (-2, 0)$ und $\xi_5 = (2, 0)$.

Wegen $f(\xi_1) = -1$, $f(\xi_2) = g(\frac{1}{2}\sqrt{3}) = -\frac{3}{4}$, $f(\xi_3) = h(\frac{\pi}{6}) = 0$, $f(\xi_4) = g(-2) = h(\pi) = 4 + 2\sqrt{3}$ und $f(\xi_5) = g(2) = h(0) = 4 - 2\sqrt{3}$ folgt, daß f sein globales Maximum $4 + 2\sqrt{3}$ in $\xi_4 = (-2, 0)$ und sein globales Minimum -1 in $\xi_1 = (\frac{1}{2}\sqrt{3}, \frac{1}{2})$ annimmt.

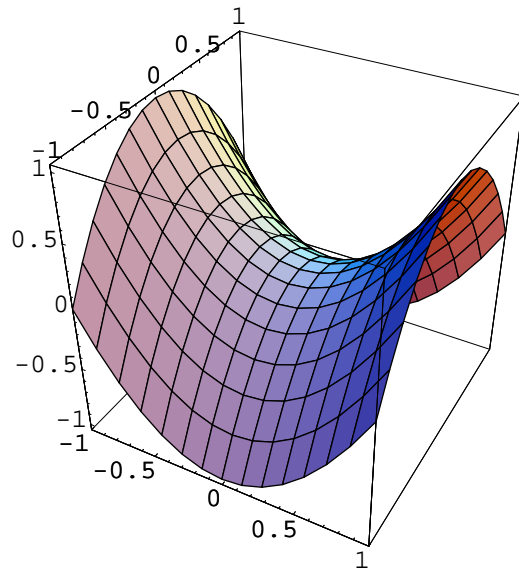
Wir haben hier die Extrema zur Übung mit Hilfe der Differentialrechnung bestimmt. Im konkreten Beispiel hätte man das leichter tun können, in dem man $f(x, y)$ in der Form $f(x, y) = (x - \frac{1}{2}\sqrt{3})^2 + (y - \frac{1}{2})^2 - 1$ schreibt.

Satz 7.4.2 Sei $U \subset \mathbb{R}^p$ offen, $f \in C^2(U, \mathbb{R})$, $\xi \in U$ und $\text{grad}f(\xi) = 0$. Dann gilt

- (i) Ist $H_f(\xi)$ positiv definit, so hat f ein lokales Minimum in ξ .
- (ii) Ist $H_f(\xi)$ negativ definit, so hat f ein lokales Maximum in ξ .
- (iii) Ist $H_f(\xi)$ indefinit, so hat f kein lokales Extremum in ξ .

Im semidefiniten Fall ist keine allgemeine Aussage möglich. (Dies entspricht dem Fall $f''(\xi) = 0$ im Falle einer Veränderlichen.)

Für $p = 2$ sieht im Falle (iii) der Graph von f (also $\{(x, y, f(x, y)) : (x, y) \in U\} \subset \mathbb{R}^3$) lokal wie ein "Sattel" aus, im Beispiel $f(x, y) = x^2 - y^2$, $\xi = (0, 0)$, etwa wie folgt:



Man spricht daher auch von einem *Sattelpunkt*.

Beweis von Satz 7.4.2. Sei $h \in \mathbb{R}^p$ mit $\xi + h \in U$. Aufgrund der Taylorformel gilt

$$f(\xi + h) - f(\xi) = \underbrace{Df(\xi)(h)}_{=0} + \frac{1}{2}D^2f(\xi)(h, h) + R(h) = \frac{1}{2}h^T H_f(\xi)h + R(h)$$

mit $R(h) = o(\|h\|^2)$ für $h \rightarrow 0$.

zu (i). Ist $H_f(\xi)$ positiv definit, so existiert $c > 0$ mit $h^T H_f(\xi)h \geq c\|h\|^2$ für alle $h \in \mathbb{R}^p$. Weiter existiert $\delta > 0$ mit $|R(h)| < \frac{c}{2}\|h\|^2$ für $0 < \|h\| \leq \delta$. Es folgt

$$f(\xi + h) - f(\xi) \geq \frac{1}{2}h^T H_f(\xi)h - |R(h)| > 0$$

für $0 < \|h\| \leq \delta$. Also ist ξ lokales Minimum.

zu (ii). Dies folgt analog zu (i).

zu (iii). Ist $H_f(\xi)$ indefinit, so existieren $h_1, h_2 \in \mathbb{R}^p$ mit $c_1 := h_1^T H_f(\xi)h_1 < 0$ und $c_2 := h_2^T H_f(\xi)h_2 > 0$. Für $t \in \mathbb{R}$ und $j \in \{1, 2\}$ folgt $(th_j)^T H_f(\xi)(th_j) = t^2 c_j$ und $R(th_j) = o(\|th_j\|^2) = o(t^2)$ für $t \rightarrow 0$. Für genügend kleines t ist dann $|R(th_j)| < \frac{1}{2}(th_j)^T H_f(\xi)(th_j)$ und damit

$$f(\xi + th_1) - f(\xi) \leq \frac{1}{2}h^T H_f(\xi)h + |R(h)| < 0$$

und

$$f(\xi + th_2) - f(\xi) \geq \frac{1}{2}h^T H_f(\xi)h - |R(h)| > 0.$$

Es folgt, daß ξ kein lokales Extremum ist. \square

Wir betrachten den Spezialfall $p = 2$ genauer. Für eine (2×2) -Matrix

$$A = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$$

ist die Determinante $\det(A)$ durch $\det(A) = ac - b^2$ gegeben und es gilt:

- (i) A ist positiv definit $\Leftrightarrow a > 0$ und $\det(A) > 0$.
- (ii) A ist negativ definit $\Leftrightarrow a < 0$ und $\det(A) > 0$.
- (iii) A ist indefinit $\Leftrightarrow \det(A) < 0$.

Damit erhalten wir das folgende Resultat direkt aus Satz 7.4.2.

Satz 7.4.3 Sei $U \subset \mathbb{R}^2$ offen, $f \in C^2(U, \mathbb{R})$, $\xi \in U$ und $\text{grad}f(\xi) = 0$. Sei $\Delta(\xi) := \partial_{11}f(\xi)\partial_{22}f(\xi) - (\partial_{12}f(\xi))^2$. Dann gilt

- (i) Ist $\Delta(\xi) > 0$ und $\partial_{11}f(\xi) > 0$, so hat f ein lokales Minimum in ξ .
- (ii) Ist $\Delta(\xi) > 0$ und $\partial_{11}f(\xi) < 0$, so hat f ein lokales Maximum in ξ .
- (iii) Ist $\Delta(\xi) < 0$, so hat f kein lokales Extremum (sondern einen Sattelpunkt) in ξ .

Beispiel 1. Sei $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, $f(x, y) = xy - x^3 - y^2$. Dann ist $f \in C^2(\mathbb{R}^2, \mathbb{R})$ und $\partial_1 f(x, y) = y - 3x^2$ und $\partial_2 f(x, y) = x - 2y$. Aus $\text{grad}f(x, y) = 0$ folgt $y = 3x^2$ und $x = 2y$, also $y = 3(2y)^2 = 12y^2$ und damit $y = 0$ oder $y = \frac{1}{12}$. Zusammen mit $x = 2y$ folgt $(x, y) = \xi_1 := (0, 0)$ oder $(x, y) = \xi_2 := (\frac{1}{6}, \frac{1}{12})$. Tatsächlich gilt auch $\text{grad}f(\xi_j) = 0$ für $j \in \{1, 2\}$.

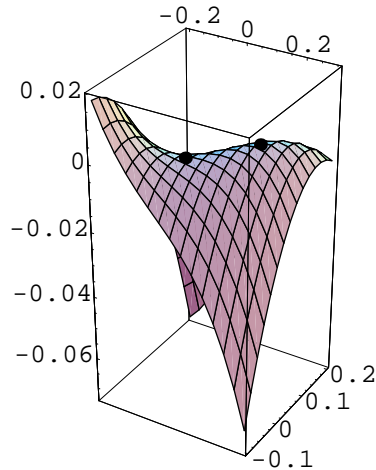
Wegen $\partial_{11}f(x, y) = -6x$, $\partial_{12}f(x, y) = \partial_{21}f(x, y) = 1$ und $\partial_{22}f(x, y) = -2$ folgt

$$H_f(x, y) = \begin{pmatrix} -6x & 1 \\ 1 & -2 \end{pmatrix}$$

und damit $\Delta(x, y) = \det H_f(x, y) = 12x - 1$.

Somit ist $\Delta(\xi_1) = \Delta(0, 0) = -1 < 0$ und folglich ist ξ_1 Sattelpunkt. Weiter ist $\Delta(\xi_2) = \Delta(\frac{1}{6}, \frac{1}{12}) = 1 > 0$ und wegen $\partial_{11}f(\xi_2) = -1 < 0$ hat f in ξ_2 ein lokales Maximum.

Der Graph von f – mit den Punkten ξ_1, ξ_2 markiert – sieht wie folgt aus:



Beispiel 2. Sei $f : \mathbb{R}^3 \rightarrow \mathbb{R}$, $f(x, y, z) = 2x^4 + y^4 + z^4 - 4xyz - 2x^2$. Dann ist $f \in C^2(\mathbb{R}^3, \mathbb{R})$ und $\partial_1 f(x, y, z) = 8x^3 - 4yz - 4x$, $\partial_2 f(x, y, z) = 4y^3 - 4xz$, $\partial_3 f(x, y, z) = 4z^3 - 4xy$. Eine längere Rechnung zeigt, daß $\text{grad} f(x, y, z) = 0$ für genau 7 Punkte (x, y, z) gilt, nämlich für $(x, y, z) = \xi_j$, $j \in \{1, \dots, 7\}$, mit $\xi_1 := (0, 0, 0)$, $\xi_2 := (\frac{1}{\sqrt{2}}, 0, 0)$, $\xi_3 := (-\frac{1}{\sqrt{2}}, 0, 0)$, $\xi_4 := (1, 1, 1)$, $\xi_5 := (1, -1, -1)$, $\xi_6 := (-1, 1, -1)$ und $\xi_7 := (-1, -1, 1)$. Es ist

$$H_f(x, y, z) = \begin{pmatrix} 24x^2 - 4 & -4z & -4y \\ -4z & 12y^2 & -4x \\ -4y & -4x & 12z^2 \end{pmatrix}$$

und damit

$$H_f(\xi_1) = \begin{pmatrix} -4 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Diese Matrix ist negativ semidefinit, Satz 7.4.2 also nicht anwendbar. Es ist aber $f(0, 0, z) > 0$ für alle $z \in \mathbb{R}$ und $f(x, 0, 0) < 0$ für $0 < x < 1$, und daher hat f in ξ_1 kein lokales Extremum. Weiter gilt

$$H_f(\xi_{2,3}) = \begin{pmatrix} 8 & 0 & 0 \\ 0 & 0 & \pm 2\sqrt{2} \\ 0 & \pm 2\sqrt{2} & 0 \end{pmatrix}.$$

Das charakteristische Polynom dieser Matrix ist durch $(8 - \lambda)(\lambda^2 - 8)$ gegeben. Damit hat $H_f(\xi_{2,3})$ sowohl positive wie negative Eigenwerte, ist also indefinit, und folglich ist in $\xi_{2,3}$ kein lokales Extremum. Weiter ist

$$H_f(\xi_4) = \begin{pmatrix} 20 & -4 & -4 \\ -4 & 12 & -4 \\ -4 & -4 & 12 \end{pmatrix}.$$

Diese Matrix ist positiv definit, wie man etwa durch Betrachtung der Abschnittsdeterminanten feststellt. Damit liegt in ξ_4 ein lokales Minimum vor.

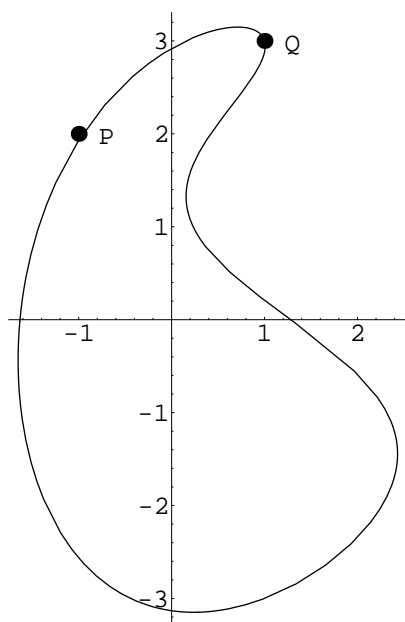
Analog – oder durch Symmetrieüberlegungen – zeigt man, daß f auch in ξ_5, ξ_6, ξ_7 lokale Minima hat.

7.5 Der Satz über implizite Funktionen

Wir betrachten die Frage, wann sich eine Gleichung der Form $F(x, y) = 0$ (mit einer differenzierbaren, auf einer Teilmenge von \mathbb{R}^2 definierten Funktion F) nach y auflösen läßt, d. h., wann eine Funktion f existiert, so daß die Gleichung $F(x, y) = 0$ äquivalent zur Gleichung $y = f(x)$ ist. Es gilt dann $F(x, f(x)) = 0$. Wir nennen dann $F(x, y) = 0$ eine *implizite* Form der *explizit* gegebenen Gleichung $y = f(x)$.

Etwas genauer werden wir folgende Fragestellung untersuchen: gegeben $(\xi, \eta) \in \mathbb{R}^2$ mit $F(\xi, \eta) = 0$. Existieren dann Umgebungen U und V von ξ und η und eine (differenzierbare) Funktion $f : U \rightarrow V$ mit $f(\xi) = \eta$ und $F(x, f(x)) = 0$ für alle $x \in U$? Inwieweit ist f eindeutig bestimmt?

Für eine konkrete Funktion F könnte $\{(x, y) \in \mathbb{R}^2 : F(x, y) = 0\}$ etwa so aussehen:



Hier scheint es so zu sein, daß für $(\xi, \eta) = P$ Umgebungen U, V und eine Funktion f mit den verlangten Eigenschaften existieren, während das für $(\xi, \eta) = Q$ nicht der Fall ist.

Man erkennt nun, daß für $(\xi, \eta) = Q$ die durch $y \mapsto F(\xi, y)$ definierte Funktion in einer Umgebung von η immer ≤ 0 oder immer ≥ 0 ist, also in η ein lokales Extremum hat. Es folgt, daß $\partial_2 F(Q) = 0$ gilt.

Wir werden zeigen, daß (für stetig differenzierbares F) Umgebungen U, V und eine Funktion f mit den verlangten Eigenschaften existieren, wenn $\partial_2 F(\xi, \eta) \neq 0$ gilt. Die Beweisidee dabei ist folgende: Wir nehmen an, daß f existiert und das f_0

eine Funktion “nahe” bei f ist. Dann gilt

$$\begin{aligned} F(x, f_0(x)) &= F(x, f_0(x)) - \underbrace{F(x, f(x))}_{=0} \\ &= \partial_2 F(x, y')(f_0(x) - f(x)) \\ &\approx \partial_2 F(\xi, \eta)(f_0(x) - f(x)) \end{aligned}$$

mit einer Zwischenstelle y' . Es folgt

$$f(x) \approx f_0(x) - \frac{F(x, f_0(x))}{\partial_2 F(\xi, \eta)} =: Af_0(x).$$

Dies legt nahe, daß $f_1 := A(f_0)$ eine bessere Approximation für f ist als f_0 , und daß die rekursiv durch $f_{n+1} := A(f_n)$ definierte Folge (f_n) gegen f konvergiert.

Zur Umsetzung dieser Beweisidee benötigen wir den folgenden Satz.

Satz 7.5.1 (Banachscher Fixpunktsatz) *Sei (M, d) metrischer Raum und $f : M \rightarrow M$. Es existiere $\alpha \in [0, 1)$ mit $d(f(x), f(y)) \leq \alpha d(x, y)$ für alle $x, y \in M$. Dann hat f genau einen Fixpunkt, d. h., es existiert genau ein $\xi \in M$ mit $f(\xi) = \xi$. Darüberhinaus gilt $\lim_{n \rightarrow \infty} f^n(x) = \xi$ für alle $x \in M$. (Dabei ist f^n die n -te Iterierte von f . Diese sind rekursiv durch $f^1 := f$ und $f^{n+1} := f \circ f^n$ definiert.)*

Beweis. Übung.

Wir werden die oben skizzierten Aussagen über implizite Funktionen in allgemeinerer Form beweisen. Seien dazu $G \subset \mathbb{R}^p$ offen, $H \subset \mathbb{R}^q$ offen, $\xi \in G$, $\eta \in H$ und $F \in C^1(G \times H, \mathbb{R}^q)$ mit $F(\xi, \eta) = 0$. Wir fragen nach der Existenz von Umgebungen U und V von ξ und η und einer differenzierbaren Funktion $f : U \rightarrow V$ mit $f(\xi) = \eta$ und $F(x, f(x)) = 0$ für alle $x \in U$.

Um die oben (für $p = q = 1$) diskutierte Bedingung $\partial_2 F(\xi, \eta) \neq 0$ zu verallgemeinern, betrachten wir für $x \in G$ die Funktion $F_2^x : H \rightarrow \mathbb{R}^q$, $y \mapsto F(x, y)$ und setzen

$$D_2 F(x, y) := DF_2^x(y)$$

und

$$\partial_2^* F(x, y) := J_{F_2^x}(y) = \begin{pmatrix} \partial_{p+1} F_1(x, y) & \cdots & \partial_{p+q} F_1(x, y) \\ \partial_{p+1} F_2(x, y) & \cdots & \partial_{p+q} F_2(x, y) \\ \vdots & & \vdots \\ \partial_{p+1} F_q(x, y) & \cdots & \partial_{p+q} F_q(x, y) \end{pmatrix}.$$

Man schreibt auch $\frac{\partial F}{\partial y}$ statt $\partial_2^* F$. Analog definiert man $D_1 F$ und $\partial_1^* F$, in dem man $F_1^y : G \rightarrow \mathbb{R}^p$, $x \mapsto F(x, y)$, betrachtet.

An Stelle der für $p = q = 1$ genannte Bedingung $\partial_2 F(\xi, \eta) \neq 0$ benötigt man jetzt, daß $D_2 F(\xi, \eta)$ invertierbar ist. Dies ist nach Linearer Algebra äquivalent zu $\det(\partial_2^* F(\xi, \eta)) \neq 0$.

Satz 7.5.2 (Satz über implizite Funktionen) *Sei $G \subset \mathbb{R}^p$ offen, $H \subset \mathbb{R}^q$ offen, $\xi \in G$, $\eta \in H$ und $F \in C^1(G \times H, \mathbb{R}^q)$ mit $F(\xi, \eta) = 0$. Es sei $D_2 F(\xi, \eta)$ invertierbar. Dann existieren offene Umgebungen $U \subset G$ und $V \subset H$ von ξ und η*

und eine Funktion $f \in C^1(U, V)$ mit $f(\xi) = \eta$ und $F(x, f(x)) = 0$ für $x \in U$, und umgekehrt folgt für $(x, y) \in U \times V$ mit $F(x, y) = 0$, daß $y = f(x)$ gilt.

Desweiteren gilt $Df(x) = -D_2F(x, f(x))^{-1} \circ D_1F(x, f(x))$ für $x \in U$, insbesondere also $Df(\xi) = -D_2F(\xi, \eta)^{-1} \circ D_1F(\xi, \eta)$.

Beweis. Wir werden nur die Existenz von $f \in C(U, V)$ mit den verlangten Eigenschaften beweisen, und den Beweis der stetigen Differenzierbarkeit von f nur skizzieren. Ist aber die Differenzierbarkeit von f bekannt, so folgt die Formel für Df aus der Kettenregel.

Zur Abkürzung sei $D := D_2F(\xi, \eta) \in L(\mathbb{R}^q, \mathbb{R}^q)$. Nach Voraussetzung existiert dann D^{-1} . Seien $\varepsilon, \delta > 0$, $U := U(\xi, \delta)$ und $V := U(\eta, \varepsilon)$. Dabei können wir ε, δ so klein wählen, daß die folgenden drei Bedingungen gelten:

(i) $\bar{U} \subset G$ und $\bar{V} \subset H$.

(ii) Für $x \in U$ und $y \in V$ gilt

$$\|id_{\mathbb{R}^q} - D^{-1}(D_2F(x, y))\|_{L(\mathbb{R}^q, \mathbb{R}^q)} \leq \frac{1}{2}.$$

(iii) Für $x \in U$ gilt

$$\|D^{-1}(F(x, \eta))\| \leq \frac{\varepsilon}{4}.$$

Dabei können (ii) und (iii) erreicht werden, weil die linken Seiten stetig in $G \times H$ bzw. G sind und für $(x, y) = (\xi, \eta)$ bzw. $x = \xi$ verschwinden. Wir setzen nun $X := C(\bar{U}, \mathbb{R}^q) = \{g : \bar{U} \rightarrow \mathbb{R}^q : g \text{ stetig}\}$ und $\|g\|_X := \max_{x \in \bar{U}} \|g(x)\|$ für $g \in X$. Dann ist $(X, \|\cdot\|_X)$ Banachraum; vgl. Beispiel 1 zu Satz 6.2.5. Weiter setzen wir

$$M := \left\{ g \in X : g(\xi) = \eta, \forall x \in \bar{U} : \|g(x) - \eta\|_X \leq \frac{\varepsilon}{2} \right\}.$$

Dann ist $M \neq \emptyset$, da die konstante Funktion $g = \eta$ in M ist. Außerdem ist M abgeschlossen. Nach Satz 6.2.5 ist M also vollständig.

Wir betrachten jetzt die Funktion $A : M \rightarrow X$,

$$A(g)(x) : g(x) - D^{-1}(F(x, g(x))).$$

Wir werden zeigen:

(iv) Für $g, h \in M$ gilt $\|A(g) - A(h)\|_X \leq \frac{1}{2}\|g - h\|_X$.

(v) $A(M) \subset M$.

Der Banachsche Fixpunktsatz liefert nun $f \in M \subset C\left(\bar{U}, \overline{U(\eta, \frac{\varepsilon}{2})}\right)$ und damit $f \in C(U, V)$ mit $A(f) = f$. Daraus folgt $D^{-1}(F(x, f(x))) = 0$, also $F(x, f(x)) = 0$ für $x \in U$.

zu (iv): Für $x \in \bar{U}$ betrachten wir $\phi^x : V \rightarrow \mathbb{R}^q$, $y \mapsto y - D^{-1}(F(x, y))$, also $\phi^x = id_{\mathbb{R}^q}|_V - D^{-1} \circ F_2^x$. Es ist dann $\phi^x(g(x)) = A(g)(x)$ für $g \in M$.

Nach Kettenregel ist ϕ^x differenzierbar mit

$$D\phi^x(y) = id_{\mathbb{R}^q} - D^{-1}(D_2F(x, y)).$$

Nach (ii) folgt $\|D\phi^x(y)\|_{L(\mathbb{R}^q, \mathbb{R}^q)} \leq \frac{1}{2}$ für $y \in V$. Nach Mittelwertungleichung (Satz 7.2.7) folgt

$$\|\phi^x(y_1) - \phi^x(y_2)\| \leq \sup_{y \in [y_1, y_2]} \|D\phi^x(y)\|_{L(\mathbb{R}^q, \mathbb{R}^q)} \|y_1 - y_2\|$$

für $y_1, y_2 \in V$, und damit

$$\|\phi^x(y_1) - \phi^x(y_2)\| \leq \frac{1}{2} \|y_1 - y_2\|,$$

und das gilt aus Stetigkeitsgründen sogar für $y_1, y_2 \in \bar{V}$. Es folgt

$$\|A(g)(x) - A(h)(x)\| = \|\phi^x(g(x)) - \phi^x(h(x))\| \leq \frac{1}{2} \|g(x) - h(x)\|$$

für $g, h \in M$ und $x \in \bar{U}$ und damit (iv).

zu (v): Für $g \in M$ und $x \in \bar{U}$ gilt

$$\begin{aligned} \|A(g)(x) - \eta\| &\leq \|A(g)(x) - A(\eta)(x)\| + \|A(\eta)(x) - \eta\| \\ &\stackrel{\text{(iv)}}{\leq} \frac{1}{2} \|g(x) - \eta\| + \|D^{-1}(F(x, \eta))\| \\ &\stackrel{\text{(iii)}}{\leq} \frac{1}{2} \cdot \frac{\varepsilon}{2} + \frac{\varepsilon}{4}, \end{aligned}$$

und daraus folgt (v).

Wir haben also die Existenz einer Funktion $f \in C(U, V)$ mit $F(x, f(x)) = 0$ für $x \in U$ nachgewiesen. Seien nun $x \in U$ und $y \in V$ mit $F(x, y) = 0$. Dann ist $\phi^x(y) = y$ und damit $\|f(x) - y\| = \|\phi^x(f(x)) - \phi^x(y)\| \leq \frac{1}{2} \|f(x) - y\|$ und damit $\|f(x) - y\| = 0$, also $f(x) = y$.

Um die Differenzierbarkeit von f in ξ zu zeigen, betrachtet man die Entwicklung

$$\begin{aligned} 0 &= \underbrace{F(x, f(x))}_{=0} - \underbrace{F(\xi, \eta)}_{=0} \\ &= D_1 F(\xi, \eta)(x - \xi) + D_2 F(\xi, \eta)(f(x) - \eta) + \dots, \end{aligned}$$

die wegen $f(\xi) = \eta$

$$f(x) - f(\xi) = -(D_2 F(\xi, \eta))^{-1} \circ D_1 F(\xi, \eta)(x - \xi) + \dots$$

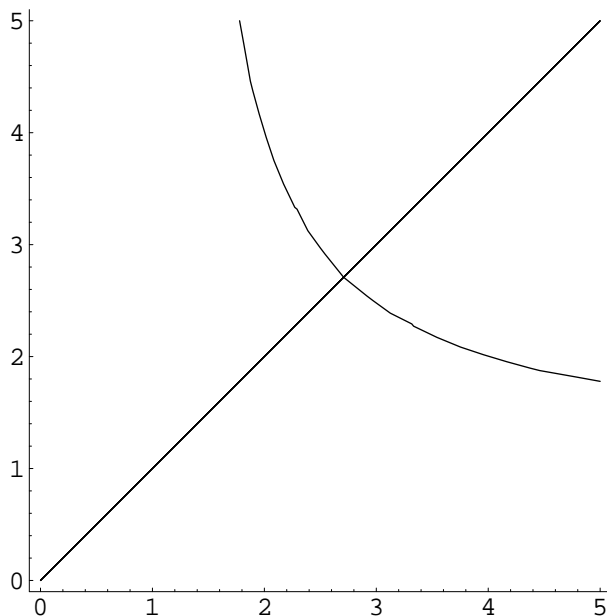
liefert. Die Differenzierbarkeit von f in ξ erhält man nun durch eine Analyse der hier nur durch ... angedeuteten Restglieder. Die Differenzierbarkeit von f in einer Umgebung von ξ erhält man analog. (Dabei muß das im ersten Teil des Beweises gewählte δ eventuell noch verkleinert werden.) \square

Beispiel 1. Sei $F : \mathbb{R}_+ \times \mathbb{R}_+ \rightarrow \mathbb{R}$, $F(x, y) = x^y - y^x$. Es gilt $F(2, 4) = 0$ und $\partial_2 F(x, y) = x^y \ln x - xy^{x-1}$, also $\partial_2 F(2, 4) = 2^4 \ln 2 - 2 \cdot 4 = 8(2 \ln 2 - 1) > 0$. Nach dem Satz über implizite Funktionen existieren $\varepsilon, \delta > 0$ und stetig differenzierbares $f : (2 - \delta, 2 + \delta) \rightarrow (4 - \varepsilon, 4 + \varepsilon)$ mit $F(x, f(x)) = 0$, d. h., $x^{f(x)} = f(x)^x$.

Es ist $\partial_1 F(x, y) = yx^{y-1} - y^x \ln y$, also $\partial_1 F(2, 4) = 4 \cdot 2^3 - 4^2 \ln 4 = 32(1 - \ln 2)$. Es folgt

$$f'(2) = -\frac{\partial_1 F(2, 4)}{\partial_2 F(2, 4)} = -4 \frac{2 \ln 2 - 1}{1 - \ln 2} = -3,177\dots$$

Die durch $F(x, y) = 0$ gegebene Teilmenge des \mathbb{R}^2 sieht wie folgt aus:



Eine Lösung der Gleichung $F(x, y) = 0$ ist natürlich auch durch $y = x$ gegeben. Wegen $\partial_2 F(x, x) = x^x(\ln x - 1) \neq 0$ für $x \neq e$ folgt aus dem Satz über implizite Funktionen, daß jeder Punkt $x \neq e$ eine Umgebung U besitzt, so daß $F(x, y) \neq 0$ für $x, y \in U$ mit $x \neq y$. Dies ist für $x = e$ nicht der Fall.

Beispiel 2. Sei $G = \mathbb{R}$, $H = \mathbb{R}_+ \times \mathbb{R}_+$ und $F = (F_1, F_2)^T : G \times H \rightarrow \mathbb{R}^2$,

$$F(x, y_1, y_2) = \begin{pmatrix} F_1(x, y_1, y_2) \\ F_2(x, y_1, y_2) \end{pmatrix} = \begin{pmatrix} x \ln y_1 + y_1 e^{xy_2} - 1 \\ x e^{y_1} + \sqrt{y_1 y_2} - y_2 e^x \end{pmatrix}.$$

Es ist dann $F \in C^1(G \times H, \mathbb{R}^2)$ und $F(0, 1, 1) = 0$. Es gilt

$$\begin{aligned} \partial_2^* F(x, y_1, y_2) &= \begin{pmatrix} \partial_2 F_1(x, y_1, y_2) & \partial_3 F_1(x, y_1, y_2) \\ \partial_2 F_2(x, y_1, y_2) & \partial_3 F_2(x, y_1, y_2) \end{pmatrix} \\ &= \begin{pmatrix} \frac{x}{y_1} + e^{xy_2} & y_1 x e^{xy_2} \\ x e^{y_1} + \frac{1}{2} \sqrt{\frac{y_2}{y_1}} & \frac{1}{2} \sqrt{\frac{y_1}{y_2}} - e^x \end{pmatrix}, \end{aligned}$$

also

$$A := \partial_2^* F(0, 1, 1) = \begin{pmatrix} 1 & 0 \\ \frac{1}{2} & -\frac{1}{2} \end{pmatrix}.$$

Wegen $\det A = -\frac{1}{2} \neq 0$ ist A invertierbar. Es gilt

$$A^{-1} = \begin{pmatrix} 1 & 0 \\ 1 & -2 \end{pmatrix}.$$

Nach dem Satz über implizite Funktionen existiert also eine Umgebung $U \subset \mathbb{R}$ von 0 und eine Umgebung $V \subset \mathbb{R}_+ \times \mathbb{R}_+$ von $(1, 1)$ sowie eine Funktion $f = (f_1, f_2)^T :$

$U \rightarrow V$ mit $F(x, f_1(x), f_2(x)) = 0$ für alle $x \in U$. Wegen

$$\partial_1^* F(x, y_1, y_2) = \begin{pmatrix} \partial_1 F_1(x, y_1, y_2) \\ \partial_1 F_2(x, y_1, y_2) \end{pmatrix} = \begin{pmatrix} \ln y_1 + y_1 y_2 e^{xy_2} \\ e^{y_1} - y_2 e^x \end{pmatrix}$$

ist

$$\partial_1^* F(0, 1, 1) = \begin{pmatrix} 1 \\ e - 1 \end{pmatrix}$$

und damit

$$f'(0) = -\partial_2^* F(0, 1, 1)^{-1} \partial_1^* F(0, 1, 1) = - \begin{pmatrix} 1 & 0 \\ 1 & -2 \end{pmatrix} \begin{pmatrix} 1 \\ e - 1 \end{pmatrix} = \begin{pmatrix} -1 \\ 2e - 3 \end{pmatrix}.$$

Satz 7.5.3 (Umkehrsatz) Sei $G \subset \mathbb{R}^p$ offen, $f \in C^1(G, \mathbb{R}^p)$, $\alpha \in G$ und $Df(\alpha)$ invertierbar. Dann existieren offene Umgebungen $U \subset G$ von α und $V \subset \mathbb{R}^p$ von $\beta := f(\alpha)$, so daß $f : U \rightarrow V$ bijektiv ist (d. h., $f|_U$ ist injektiv und $V := f(U)$ ist eine offene Umgebung von β).

Die Umkehrfunktion $g : V \rightarrow U$ von $f : U \rightarrow V$ ist stetig differenzierbar und es gilt $Dg(x) = Df(g(x))^{-1}$ für $x \in V$.

Beweis. Wir wenden den Satz über implizite Funktionen auf die Funktion $F : \mathbb{R}^p \times \mathbb{R}^p \rightarrow \mathbb{R}^p$, $F(x, y) = f(y) - x$, an. Es ist dann $F(\beta, \alpha) = 0$ und $D_2 F(x, y) = Df(y)$, also $D_2 F(\beta, \alpha) = Df(\alpha)$.

Nach dem Satz über implizite Funktionen existieren offene Umgebungen W von α und V von β sowie eine Funktion $g \in C^1(V, W)$ mit $0 = F(x, g(x)) = f(g(x)) - x$ für $x \in V$, und $F(x, y) \neq 0$ falls $x \in V$, $y \in W$, $y \neq g(x)$.

Wegen $f(g(x)) - x = 0$ für $x \in V$ folgt, daß g injektiv ist. Mit $U := g(V)$ folgt also, daß $f : U \rightarrow V$ bijektiv mit Umkehrfunktion $g : V \rightarrow U$ ist.

Noch zu zeigen ist, daß U offen ist. Dazu zeigen wir zunächst, daß $U = f^{-1}(V) \cap W$ gilt. Sei dazu $y \in U$. Dann existiert $x \in V$ mit $y = g(x)$. Es folgt $f(y) = f(g(x)) = x \in V$, also $y \in f^{-1}(V)$, und wegen $U \subset W$ also $U \subset f^{-1}(V) \cap W$. Ist umgekehrt $y \in f^{-1}(V) \cap W$, so folgt $x := f(y) \in V$. Wegen $F(x, y) = f(y) - x = 0$ folgt $y = g(x) \in U$. Insgesamt erhalten wir $U = f^{-1}(V) \cap W$. Da f stetig und V offen ist, ist $f^{-1}(V)$ offen und damit auch $f^{-1}(V) \cap W$ offen.

Die Formel $Dg(x) = Df(g(x))^{-1}$ folgt aus dem Satz über implizite Funktionen – oder direkt aus der Kettenregel. \square

Der Satz ist nur lokal, d. h., auch wenn $Df(x)$ für alle $x \in G$ invertierbar ist, muß $f : G \rightarrow f(G)$ nicht bijektiv sein. Man betrachte etwa die Funktion $f : \mathbb{C} \rightarrow \mathbb{C} \setminus \{0\}$, $f(z) = e^z$. In reeller Schreibweise haben wir

$$f(x, y) = \begin{pmatrix} f_1(x, y) \\ f_2(x, y) \end{pmatrix} = \begin{pmatrix} e^x \cos y \\ e^x \sin y \end{pmatrix}$$

und damit

$$\det J_f(x, y) = \det \begin{pmatrix} e^x \cos y & -e^x \sin y \\ e^x \sin y & e^x \cos y \end{pmatrix} = e^{2x} \neq 0,$$

womit $Df(x)$ für alle $x \in \mathbb{R}^2$ invertierbar ist. Andererseits gilt aber $f(0, 2\pi) = f(0, 0)$, womit f nicht injektiv ist.

Definition 7.5.1 Seien (M, d_M) , (N, d_N) metrische Räume. Eine Abbildung $f : M \rightarrow N$ heißt *offen*, wenn $f(U)$ für jedes offene $U \subset M$ offen ist.

Man vergleiche diese Definition mit der durch Satz 6.3.3 gegebenen Charakterisierung der Stetigkeit. (“Urbilder offener Mengen sind offen.”)

Eine Folgerung aus dem Umkehrsatz ist das folgende Resultat.

Satz 7.5.4 Sei $G \subset \mathbb{R}^p$ offen, $f \in C^1(G, \mathbb{R}^p)$ und $Df(x)$ invertierbar für alle $x \in G$. Dann ist f offen.

7.6 Extrema unter Nebenbedingungen

In §4.3.3 haben wir für $U \subset \mathbb{R}^p$ Funktionen $f : U \rightarrow \mathbb{R}$ auf lokale Extrema untersucht. Jetzt betrachten wir den Fall, daß zusätzlich noch eine Funktion $g : U \rightarrow \mathbb{R}^q$ gegeben ist, und wollen lokale Extrema von f unter der Nebenbedingung $g = 0$ bestimmen, d. h., wir untersuchen $f|_{\{x \in U : g(x) = 0\}} = f|_{g^{-1}(0)}$ auf lokale Extrema.

Beispiel. Man maximiere (oder minimiere) $f(x, y) = x + y$ unter der Nebenbedingung $g(x, y) = x^2 + y^2 - 1 = 0$.

Eine Methode besteht darin, die Gleichung $g(x, y) = 0$ nach y aufzulösen und das Resultat $y = h(x)$ in f einzusetzen, also das Extremum von $\tilde{f}(x) = f(x, h(x))$ zu bestimmen.

Im obigen *Beispiel* ist $h(x) = \pm\sqrt{1-x^2}$ und $\tilde{f}(x) = x \pm \sqrt{1-x^2}$, $x \in [-1, 1]$.

Der Nachteil dieser Methode ist, daß eine Auflösung der Gleichung $g(x, y)$ nach y (oder x) nicht immer möglich ist, oder die Auflösung oft kompliziert ist. Stattdessen werden wir jetzt den Satz über implizite Funktionen benutzen.

Satz 7.6.1 (Lagrangesche Multiplikatorenregel) Sei $q < p$, $G \subset \mathbb{R}^p$ offen, $f \in C^1(G, \mathbb{R})$, $g \in C^1(G, \mathbb{R}^q)$, $\xi \in G$ und $g(\xi) = 0$. Die Funktion f habe in ξ ein lokales Extremum unter der Nebenbedingung $g = 0$. Hat dann $Dg(\xi)$ Rang q , so existiert $\Lambda \in L(\mathbb{R}^q, \mathbb{R})$ mit $\Lambda \circ Dg(\xi) = Df(\xi)$.

Ist $\lambda = (\lambda_1, \dots, \lambda_q)$ die Λ darstellende $(1 \times q)$ -Matrix bezüglich der Standardbasis, so erhält die Gleichung $\Lambda \circ Dg(\xi) = Df(\xi)$ die Form

$$\text{grad} f(\xi) = \lambda \cdot J_g(\xi),$$

mit $g = (g_1, \dots, g_q)^T$ also

$$\sum_{j=1}^q \lambda_j \partial_k g_j(\xi) = \partial_k f(\xi).$$

Die λ_j nennt man *Lagrangesche Multiplikatoren*.

Beweis von Satz 7.6.1. Nach der Voraussetzung über den Rang von $Dg(\xi)$ existieren q linear unabhängige Spalten in $J_g(\xi)$. Ohne Beschränkung der Allgemeinheit seien die letzten q Spalten linear unabhängig. Wir setzen nun $r := p - q$. Dann ist $\mathbb{R}^p = \mathbb{R}^r \times \mathbb{R}^q$ und f und g können als Abbildungen von $\mathbb{R}^r \times \mathbb{R}^q$ nach \mathbb{R} aufgefaßt

werden. Dementsprechend schreiben wir $x \in \mathbb{R}^p$ in der Form $x = (z, y)$ mit $z \in \mathbb{R}^r$ und $y \in \mathbb{R}^q$ und damit $f(z, y)$ bzw. $g(z, y)$ statt $f(x)$ bzw. $g(x)$. Wir definieren D_1f, D_2f, D_1g und D_2g wie zuvor.

Nach obiger Annahme über die Spalten von $J_g(\xi)$ ist dann $D_2g(\xi)$ invertierbar. Setzen wir noch $\xi = (\zeta, \eta)$, so folgt mit dem Satz über implizite Funktionen, daß Umgebungen $U \subset \mathbb{R}^r$ von ζ und $V \subset \mathbb{R}^q$ von η und eine Funktion $h \in C^1(U, V)$ existieren, so daß $h(\zeta) = \eta$ und $g(z, h(z)) = 0$ für $z \in U$. Damit ist in ζ ein lokales Extremum der durch $z \mapsto f(z, h(z))$ definierten Funktion $\varphi : U \rightarrow \mathbb{R}$. Es folgt $D\varphi(\zeta) = 0$.

Nun ist (nach Kettenregel)

$$D\varphi(\zeta) = D_1f(\zeta, \eta) + D_2f(\zeta, \eta) \circ Dh(\zeta)$$

und (nach Satz über implizite Funktionen)

$$Dh(\zeta) = -D_2g(\zeta, \eta)^{-1} \circ D_1g(\zeta, \eta).$$

Es folgt

$$0 = D\varphi(\zeta) = D_1f(\xi) - D_2f(\xi) \circ D_2g(\xi)^{-1} \circ D_1g(\xi).$$

Mit $\Lambda := D_2f(\xi) \circ D_2g(\xi)^{-1}$ folgt also $D_1f(\xi) = \Lambda \circ D_1g(\xi)$. Trivialerweise ist auch $D_2f(\xi) = \Lambda \circ D_2g(\xi)$. Es folgt $Df(\xi) = \Lambda \circ Dg(\xi)$. \square

Beispiel 1. Es seien $f, g : \mathbb{R}^2 \rightarrow \mathbb{R}$, $f(x, y) = x + y$, $g(x, y) = x^2 + y^2 - 1 = 0$. Man bestimme das Maximum und Minimum von f unter der Nebenbedingung $g = 0$. (Die Existenz von Maximum und Minimum folgt, da f stetig und $g^{-1}(0) = \{(x, y) \in \mathbb{R}^2 : g(x, y) = 0\}$ kompakt.)

Es sind $f, g \in C^1(\mathbb{R}^2, \mathbb{R})$ mit $\text{grad}f(x, y) = (1, 1)$ und $\text{grad}g(x, y) = (2x, 2y)$. Aus $\text{grad}g(x, y) = (0, 0)$ folgt $(x, y) = (0, 0)$. Da aber $g(0, 0) = -1 \neq 0$, ist die Voraussetzung über den Rang von Dg in allen Punkte von $g^{-1}(0)$ erfüllt.

Wir betrachten die Gleichung

$$\text{grad}f(x, y) = \lambda \text{grad}g(x, y),$$

also $(1, 1) = \lambda(2x, 2y)$. Es folgt $\lambda \neq 0$ und damit $x = y$. Zusammen mit $g(x, y) = x^2 + y^2 - 1 = 0$ erhalten wir $(x, y) = \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right) =: \xi_1$ oder $(x, y) = \left(-\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right) =: \xi_2$. Es folgt, daß f sein Maximum $\sqrt{2}$ in ξ_1 und sein Minimum $-\sqrt{2}$ in ξ_2 annimmt.

Beispiel 2. Man maximiere das Volumen einer Konservendose bei konstanter Oberfläche, d. h., man maximiere die durch $V(r, h) = \pi r^2 h$ gegebene Funktion $V : \mathbb{R}_+^2 \rightarrow \mathbb{R}$ unter der Nebenbedingung $g(r, h) := 2\pi r^2 + 2\pi r h - F = 0$, wobei $F \in \mathbb{R}_+$ eine gegebene Konstante ist.

Zunächst ist wegen $2\pi r^2 < F$ für $g(r, h) = 0$ die Menge $\{r \in \mathbb{R}_+ : \exists h \in \mathbb{R}_+ \text{ mit } g(r, h) = 0\}$ beschränkt und damit gilt $V(r, h) \rightarrow 0$ für $h \rightarrow 0$ (und $g(r, h) = 0$). Außerdem gilt für $g(r, h) = 0$ auch $2\pi r h < F$ und damit $r < F/2\pi h$, also $V(r, h) < \pi h F^2 / (4\pi^2 h^2) = F^2 / 4\pi h$. Es folgt $V(r, h) \rightarrow 0$ für $h \rightarrow \infty$ (und $g(r, h) = 0$). Zusammen mit der Stetigkeit von V folgt aus diesen Überlegungen, daß das gesuchte Maximum existiert.

Es gilt nun $V, g \in C^1(\mathbb{R}^2, \mathbb{R})$ mit $\text{grad}V(r, h) = (2\pi r h, \pi r^2)$ und $\text{grad}g(r, h) = (4\pi r + 2\pi h, 2\pi r)$. Aus $\text{grad}g(r, h) = (0, 0)$ folgt $(r, h) = (0, 0)$. Da aber $g(0, 0) =$

$-F \neq 0$, ist die Voraussetzung über den Rang von Dg wieder in allen Punkte von $g^{-1}(0)$ erfüllt. Die Gleichung

$$\text{grad}V(r, h) = \lambda \text{grad}g(r, h)$$

hat jetzt die Form

$$\begin{aligned} 2\pi r h &= \lambda(4\pi r + 2\pi h), \\ \pi r^2 &= \lambda 2\pi r. \end{aligned}$$

Aus der zweiten Gleichung folgt $\lambda = \frac{r}{2}$ und Einsetzen in die erste liefert $2\pi r h = \frac{r}{2}(4\pi r + 2\pi h) = 2\pi r^2 + \pi r h$ und damit $r h = 2r^2$, also $h = 2r$.

Bemerkung. In Beispiel 2 kann man natürlich auch die Nebenbedingung $g(r, h) = 0$ nach h auflösen und erhält $h = \alpha(r) := \frac{F - 2\pi r^2}{2\pi r}$. Man hat dann

$$\beta(r) := V(r, \alpha(r)) = \pi r^2 \frac{F - 2\pi r^2}{2\pi r} = \frac{F}{2} r - \pi r^3$$

zu maximieren. Aus $\beta'(r) = \frac{F}{2} - 3\pi r^2$ folgt, daß β sein Maximum bei $r = r_0 := \sqrt{\frac{F}{6\pi}}$ annimmt. Der zugehörige Wert von h ist durch

$$h_0 := \alpha(r_0) = \frac{F - 2\pi r_0^2}{2\pi r_0} = \dots = 2\sqrt{\frac{F}{6\pi}}$$

gegeben.

Die interessante Aussage in diesem Beispiel ist, daß das Maximum für $h = 2r$ angenommen wird. Diese erhält man mit Hilfe der Lagrangeschen Multiplikatoren direkt. Durch das Auflösen der Nebenbedingung nach einer Variablen erhält man zunächst die eher uninteressante Aussage, wie sich an der Stelle, wo das Maximum angenommen wird, r und h aus F berechnen lassen, und man kann die Bedingung $r = 2h$ dann erst durch Elimination von F daraus gewinnen.

7.7 Vertauschung von Grenzprozessen

Gegeben sei eine Funktion f von zwei Veränderlichen. Der Satz von Schwarz besagt, daß unter geeigneten Voraussetzungen die Differentiation von $f(x, y)$ nach x mit der nach y vertauscht werden kann.

Wir untersuchen die Frage, wann die Differentiation nach x mit der Integration über y vertauscht werden kann und wann Integration über x mit Integration über y vertauscht werden kann.

Satz 7.7.1 *Seien I, J kompakte Intervalle und sei $f : I \times J \rightarrow \mathbb{R}$ stetig. Dann ist auch $F : I \rightarrow \mathbb{R}$,*

$$F(x) = \int_J f(x, y) dy$$

stetig.

Existiert $\partial_1 f$ und ist $\partial_1 f$ stetig (in $I \times J$), so ist F stetig differenzierbar und es gilt

$$F'(x) = \int_J \partial_1 f(x, y) dy.$$

In den Randpunkten von $I \times J$ ist dabei $\partial_1 f$ als "einseitige" Ableitung zu verstehen. (Wir haben partielle Ableitungen nur für auf offenen Mengen definierte Funktionen eingeführt, aber die Erweiterung ist hier offensichtlich.)

Man kann die Formel in obigem Satz auch einprägsam als

$$\frac{d}{dx} \int_J f(x, y) dy = \int_J \frac{\partial}{\partial x} f(x, y) dy$$

schreiben.

Die Erweiterung des Satzes auf Funktionen von $p \geq 3$ Veränderlichen ist offensichtlich. Es folgt

$$\frac{d}{dx_j} \int_{I_k} f(x_1, \dots, x_p) dx_k = \int_{I_k} \frac{\partial}{\partial x_j} f(x_1, \dots, x_p) dx_k$$

für stetiges $f, \frac{\partial f}{\partial x_j} : I_1 \times \dots \times I_p \rightarrow \mathbb{R}$ (und $k \neq j$).

Beweis von Satz 7.7.1. Sei $\xi \in I$. Da $I \times J$ kompakt ist, ist f gleichmäßig stetig und damit existiert zu $\varepsilon > 0$ ein $\delta > 0$ mit

$$|f(x, y) - f(\xi, y)| < \frac{\varepsilon}{|J|}$$

für $|x - \xi| < \delta$. Es folgt

$$|F(x) - F(\xi)| = \left| \int_I f(x, y) - f(\xi, y) dy \right| \leq \int_I |f(x, y) - f(\xi, y)| dy \leq \varepsilon.$$

Damit ist F stetig in ξ .

Sei nun $\partial_1 f$ stetig und damit gleichmäßig stetig. Nach Mittelwertsatz existiert dann zu $x \in I$ und $y \in J$ ein z zwischen ξ und x mit

$$\frac{f(x, y) - f(\xi, y)}{x - \xi} = \partial_1 f(z, y).$$

Es folgt, daß zu $\varepsilon > 0$ ein $\delta > 0$ existiert mit

$$\left| \frac{f(x, y) - f(\xi, y)}{x - \xi} - \partial_1 f(\xi, y) \right| < \frac{\varepsilon}{|J|}$$

für $|x - \xi| < \delta$. Durch Integration über y erhält man

$$\left| \frac{F(x) - F(\xi)}{x - \xi} - \int_J \partial_1 f(\xi, y) dy \right| \leq \varepsilon$$

für $|x - \xi| < \delta$. Dies liefert die Differenzierbarkeit von F und die angegebene Darstellung von F' . Diese Darstellung, zusammen mit dem zuerst bewiesenen, liefert wiederum die Stetigkeit von F' . \square

Beispiel. Sei $F : \mathbb{R} \rightarrow \mathbb{R}$,

$$F(x) = \int_1^2 \frac{e^{xy}}{y} dy.$$

Dann ist F differenzierbar und es gilt

$$F'(x) = \int_1^2 e^{xy} dy = \frac{e^{xy}}{x} \Big|_{y=1}^{y=2} = \frac{e^{2x} - e^x}{x}.$$

Der folgende Satz – und Verallgemeinerungen davon – werden nach Fubini benannt.

Satz 7.7.2 *Seien I, J kompakte Intervalle und sei $f : I \times J \rightarrow \mathbb{R}$ stetig. Dann gilt*

$$\int_I \left(\int_J f(x, y) dy \right) dx = \int_J \left(\int_I f(x, y) dx \right) dy.$$

Beweis. Sei $I = [a, b]$. Wir betrachten $\phi : I \rightarrow \mathbb{R}$,

$$\phi(t) = \int_a^t \left(\int_J f(x, y) dy \right) dx,$$

und $\psi : I \rightarrow \mathbb{R}$,

$$\psi(t) = \int_J \left(\int_a^t f(x, y) dx \right) dy.$$

Nach dem Hauptsatz ist ϕ differenzierbar mit

$$\phi'(x) = \int_J f(x, y) dy.$$

Nach Satz 7.7.1 und dem Hauptsatz ist aber auch ψ differenzierbar mit

$$\psi'(t) = \int_J f(x, y) dy.$$

Es folgt $\phi' = \psi'$ und wegen $\phi(a) = 0 = \psi(a)$ damit $\phi(b) = \psi(b)$. \square

7.8 Taylorreihen

Es sei $U \subset \mathbb{R}^p$ offen und $f : U \rightarrow \mathbb{R}$. Wie in §7.3 sehen wir, daß f genau dann n -mal stetig differenzierbar ist, wenn $f \in C^n(U, \mathbb{R})$, d. h., wenn die partiellen Ableitungen der Ordnung n existieren und stetig sind. Es gilt dann

$$D^n f(\xi)(h, \dots, h) = \sum_{j_1=1}^p \sum_{j_2=1}^p \dots \sum_{j_n=1}^p \partial_{j_1 j_2 \dots j_n} f(\xi) h_{j_1} h_{j_2} \dots h_{j_n}.$$

Nach dem Satz von Schwarz kommt es auf die Reihenfolge der partiellen Ableitungen nicht an. Wir wollen daher gleiche Ableitungen zusammenfassen.

Für $\alpha = (\alpha_1, \dots, \alpha_p) \in \mathbb{N}_0^p$ und $x = (x_1, \dots, x_p) \in \mathbb{R}^p$ setzen wir

$$\begin{aligned} |\alpha| &:= \alpha_1 + \alpha_2 + \dots + \alpha_p, \\ \alpha! &:= \alpha_1! \cdot \alpha_2! \cdot \dots \cdot \alpha_p!, \\ \partial^\alpha f &:= \partial_1^{\alpha_1} \partial_2^{\alpha_2} \dots \partial_p^{\alpha_p} f, \\ x^\alpha &:= x_1^{\alpha_1} \cdot x_2^{\alpha_2} \cdot \dots \cdot x_p^{\alpha_p}. \end{aligned}$$

Wir nennen $\alpha = (\alpha_1, \dots, \alpha_p)$ auch *Multiindex*.

Eine kombinatorische Überlegung (die wir hier nicht weiter ausführen) liefert folgendes Ergebnis.

Hilfssatz 7.8.1 Die Anzahl der n -Tupel $(j_1, \dots, j_n) \in \mathbb{N}_0^n$, für welche für alle $\ell \in \{1, \dots, n\}$ die Menge $\{k : j_k = \ell\}$ genau α_ℓ Elemente hat, ist durch $\frac{n!}{\alpha!}$ gegeben.

Der Hilfssatz besagt, daß die partielle Ableitung $\partial^\alpha f$ auf $\frac{n!}{\alpha!}$ verschiedene Arten in der Form $\partial_{j_1 j_2 \dots j_n} f$ geschrieben werden kann. Wir erhalten damit

$$D^n f(\xi)(h, \dots, h) = \sum_{j_1=1}^p \sum_{j_2=1}^p \dots \sum_{j_n=1}^p \partial_{j_1 j_2 \dots j_n} f(\xi) h_{j_1} h_{j_2} \dots h_{j_n} = \sum_{\substack{\alpha \in \mathbb{N}_0^n \\ |\alpha|=n}} \frac{n!}{\alpha!} \partial^\alpha f(\xi) h^\alpha.$$

Das Taylorpolynom erhält damit die Gestalt

$$T_n(x) = f(\xi) + \sum_{k=1}^n \sum_{\substack{\alpha \in \mathbb{N}_0^n \\ |\alpha|=k}} \frac{1}{\alpha!} \partial^\alpha f(\xi) h^\alpha.$$

Beispiel. Sei $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, $f(x, y) = e^{x^2+y}$. Es ist dann $\partial_2 f = f$ und damit $\partial_2^k f = f$ für alle $k \in \mathbb{N}$. Es folgt, daß $\partial^{(j,k)} f = \partial_1^j \partial_2^k f = \partial_1^j f$ für alle $j, k \in \mathbb{N}_0$. Nun ist

$$\begin{aligned} \partial_1 f(x, y) &= 2xe^{x^2+y} \\ \partial_1^2 f(x, y) &= (2 + 4x^2)e^{x^2+y} \\ \partial_1^3 f(x, y) &= (12x + 8x^3)e^{x^2+y} \end{aligned}$$

Das zweite Taylorpolynom zum Entwicklungspunkt $\xi = (0, 0)$ ist damit gegeben durch

$$\begin{aligned} T_2(x, y) &= f(0, 0) + \partial_1 f(0, 0)x + \partial_2 f(0, 0)y \\ &\quad + \frac{1}{2} \partial_1^2 f(0, 0)x^2 + \partial_{12} f(0, 0)xy + \frac{1}{2} \partial_2^2 f(0, 0)y^2 \\ &= 1 + y + x^2 + \frac{1}{2}y^2 \end{aligned}$$

Nach Satz 7.3.1 existiert $(\sigma, \tau) \in [(0, 0), (x, y)]$ mit

$$\begin{aligned} R_2(x, y) &= f(x, y) - T_2(x, y) \\ &= \frac{1}{3!} \partial_1^3 f(\sigma, \tau)x^3 + \frac{1}{2!} \partial_1^2 \partial_2 f(\sigma, \tau)x^2 y + \frac{1}{2!} \partial_1 \partial_2^2 f(\sigma, \tau)xy^2 + \frac{1}{3!} \partial_2^3 f(\sigma, \tau)y^3. \end{aligned}$$

Ist etwa $t := \|(x, y)\|_\infty < 1$, so folgt $\|(\sigma, \tau)\|_\infty < 1$ und damit

$$|R_2(x, y)| \leq \left(\frac{1}{3}20 + \frac{1}{2}6 + \frac{1}{2}2 + \frac{1}{3} \right) e^2 t^3 = \frac{15e^2}{2} t^3.$$

Man kann nun auch Potenzreihen und Taylorreihen in mehreren Veränderlichen betrachten. Eine *Potenzreihe* in p Variablen ist dabei eine Reihe der Form

$$\sum_{\alpha \in \mathbb{N}_0^p} c_\alpha (x - \xi)^\alpha,$$

wobei $p \in \mathbb{N}$, $c_\alpha \in \mathbb{R}$ und $x, \xi \in \mathbb{R}^p$. (Auch $c_\alpha \in \mathbb{C}$ und $x, \xi \in \mathbb{C}^p$ wäre hier möglich.)
Für $f \in C^\infty(U, \mathbb{R})$ und $\xi \in U$ heißt

$$\sum_{\alpha \in \mathbb{N}_0^p} \frac{1}{\alpha!} \partial^\alpha f(\xi) (x - \xi)^\alpha$$

Taylorreihe von f um ξ . Eine Taylorreihe ist also eine Potenzreihe, deren Koeffizienten c_α die Form $c_\alpha = \partial^\alpha f(\xi) (x - \xi)^\alpha$ haben. Da für $p \geq 2$ keine Summationsreihenfolge besonders ausgezeichnet ist, betrachten wir hier nur absolut konvergente Potenz- und Taylorreihen, denn diese sind ja auch unbedingt konvergent (Satz 2.7.2).

Satz 7.8.1 *Es seien $\xi, \eta \in \mathbb{R}^p$ und die Potenzreihe*

$$\sum_{\alpha \in \mathbb{N}_0^p} c_\alpha (x - \xi)^\alpha$$

konvergiere absolut für $x = \eta$. Dann konvergiert die Potenzreihe absolut und gleichmäßig in

$$\{x \in \mathbb{R}^p : \forall j \in \{1, \dots, p\} : |x_j - \xi_j| \leq |\eta_j - \xi_j|\}.$$

Beweis. Gilt $|x_j - \xi_j| \leq |\eta_j - \xi_j|$ für alle j , so ist

$$|c_\alpha (x - \xi)^\alpha| = |c_\alpha| \cdot |x_1 - \xi_1|^{\alpha_1} \cdot \dots \cdot |x_p - \xi_p|^{\alpha_p} \leq |c_\alpha| \cdot |\eta_1 - \xi_1|^{\alpha_1} \cdot \dots \cdot |\eta_p - \xi_p|^{\alpha_p},$$

und die Behauptung folgt aus dem Majorantenkriterium. \square

Den folgenden Satz erhält man nun wie in einer Veränderlichen (vgl. Satz 4.4.2). Für $\alpha = (\alpha_1, \dots, \alpha_p), \beta = (\beta_1, \dots, \beta_p) \in \mathbb{N}_0^p$ schreiben wir dabei $\alpha \geq \beta$, falls $\alpha_j \geq \beta_j$ für alle j gilt.

Satz 7.8.2 *Es seien die Voraussetzungen von Satz 7.8.1 gegeben und es sei*

$$U := \{x \in \mathbb{R}^p : \forall j \in \{1, \dots, p\} : |x_j - \xi_j| < |\eta_j - \xi_j|\}$$

und $f : U \rightarrow \mathbb{R}$,

$$f(x) := \sum_{\alpha \in \mathbb{N}_0^p} c_\alpha (x - \xi)^\alpha.$$

Dann gilt $f \in C^\infty(U, \mathbb{R})$ und die partiellen Ableitungen dürfen gliedweise gebildet werden, d. h., für $\beta \in \mathbb{N}_0^p$ gilt

$$\partial^\beta f(x) = \sum_{\alpha \geq \beta} c_\alpha \frac{\alpha!}{(\alpha - \beta)!} (x - \xi)^{\alpha - \beta}.$$

Insbesondere gilt $\partial^\beta f(\xi) = \beta! c_\beta$.

Wie bei Funktionen einer Veränderlichen muß die Taylorreihe einer C^∞ -Funktion nicht konvergieren, und wenn sie konvergiert, muß ihre Summe nicht mit dem Funktionswert übereinstimmen. Es gilt aber z. B. der folgende Satz.

Satz 7.8.3 Sei $\xi \in \mathbb{R}^p$, $r \in \mathbb{R}_+$, $U := \{x \in \mathbb{R}^p : \forall j \in \{1, \dots, p\} : |x_j - \xi_j| < r_j\}$ und $f \in C^\infty(U, \mathbb{R})$. Es existiere $C \in \mathbb{R}_+$ mit

$$|\partial^\alpha f(x)| \leq C \frac{\alpha!}{r^\alpha}$$

für alle $x \in U$ und alle $\alpha \in \mathbb{N}_0^p$. Dann konvergiert die Taylorreihe von f um ξ auf jeder kompakten Teilmenge von U absolut und gleichmäßig, und es gilt

$$f(x) = \sum_{\alpha \in \mathbb{N}_0^p} \frac{1}{\alpha!} \partial^\alpha f(\xi) (x - \xi)^\alpha.$$

für alle $x \in U$.

Beweis. Für $\rho \in \mathbb{N}_0^p$ mit $0 < \rho_j < r_j$ für alle j sei

$$Q_\rho := \{x \in \mathbb{R}^p : \forall j \in \{1, \dots, p\} : |x_j - \xi_j| < \rho_j\}.$$

Dann ist $Q_\rho \subset U$ und für $x \in Q_\rho$ und $\alpha \in \mathbb{N}_0^p$ gilt

$$\frac{1}{\alpha!} |\partial^\alpha f(\xi) (x - \xi)^\alpha| \leq C \frac{1}{r^\alpha} \rho^\alpha = C \left(\frac{\rho_1}{r_1}\right)^{\alpha_1} \cdots \left(\frac{\rho_p}{r_p}\right)^{\alpha_p}.$$

Wegen

$$\sum_{\alpha \in \mathbb{N}_0^p} \left(\frac{\rho_1}{r_1}\right)^{\alpha_1} \cdots \left(\frac{\rho_p}{r_p}\right)^{\alpha_p} = \frac{1}{1 - \frac{\rho_1}{r_1}} \cdots \frac{1}{1 - \frac{\rho_p}{r_p}} < \infty$$

folgt die absolute und in Q_ρ gleichmäßige Konvergenz der Taylorreihe nach Vergleichskriterium. Da jede kompakte Teilmenge von U in einer der Mengen Q_ρ enthalten ist, folgt hieraus die erste Behauptung.

Sei nun $T(x)$ die Summe der Taylorreihe. Um zu zeigen, daß $f(x) = T(x)$ gilt, reicht es zu zeigen, daß $R_n(x) := f(x) - T_n(x) \rightarrow 0$ für $n \rightarrow \infty$ gilt. Nach Satz 7.3.1 existiert aber zu $x \in Q_\rho$ ein $\eta \in [\xi, x] \subset Q_\rho$ mit

$$\begin{aligned} |R_n(x)| &= \left| \sum_{\substack{\alpha \in \mathbb{N}_0^p \\ |\alpha|=n+1}} \frac{1}{\alpha!} \partial^\alpha f(\eta) (x - \xi)^\alpha \right| \\ &\leq C \sum_{\substack{\alpha \in \mathbb{N}_0^p \\ |\alpha|=n+1}} \frac{1}{r^\alpha} |(x - \xi)^\alpha| \\ &\leq C \sum_{\substack{\alpha \in \mathbb{N}_0^p \\ |\alpha|=n+1}} \left(\frac{\rho_1}{r_1}\right)^{\alpha_1} \cdots \left(\frac{\rho_p}{r_p}\right)^{\alpha_p}, \end{aligned}$$

und $R_n(x) \rightarrow 0$ folgt ähnlich wie oben. \square

Man beachte in obigem Beweis, daß aus $R_n(x) \rightarrow 0$ noch nicht die Konvergenz der Reihe folgen muß.

Beispiel. Sei $U := \{(x, y) \in \mathbb{R}^2 : x < 1 \text{ und } y < 1\}$ und $f : U \rightarrow \mathbb{R}$, $f(x, y) = \frac{1}{1-x} - \frac{1}{1-y}$. Dann gilt $f \in C^\infty(U, \mathbb{R})$ mit

$$\partial^{(m,n)} f(x, y) = \begin{cases} \frac{m!}{(1-x)^{m+1}} & \text{falls } n = 0, m \geq 1, \\ \frac{n!}{(1-y)^{n+1}} & \text{falls } m = 0, n \geq 1, \\ 0 & \text{sonst.} \end{cases}$$

Die Taylorreihe $T(x, y)$ um $\xi = (0, 0)$ ist also durch

$$T(x, y) = \sum_{m,n=0}^{\infty} c_{mn} x^m y^n = \sum_{m=0}^{\infty} x^m - \sum_{n=0}^{\infty} y^n,$$

gegeben, mit

$$c_{mn} = \begin{cases} 1 & \text{falls } n = 0, m \geq 1, \\ -1 & \text{falls } m = 0, n \geq 1, \\ 0 & \text{sonst.} \end{cases}$$

Sie konvergiert genau dann, wenn $\|(x, y)\|_\infty < 1$, also wenn $|x| < 1$ und $|y| < 1$. Weiter gilt $f(x, x) = T_n(x, x) = 0$ für alle $x \in (-\infty, 1)$. Für $x \in (-\infty, -1)$ gilt also $R_n(x, x) = 0$, obwohl die Taylorreihe divergiert.

In konkreten Fällen berechnet man – wie auch in obigem Beispiel – die Taylorreihe am einfachsten oft nicht durch sukzessives Berechnen der partiellen Ableitungen, sondern durch Rückgriff auf bekannte Potenzreihenentwicklungen in einer Veränderlichen.

Beispiel. Sei $f : \{(x, y) \in \mathbb{R}^2 : x < 1\} \rightarrow \mathbb{R}$,

$$f(x, y) = \frac{e^{xy}}{1-x}.$$

Für $|x| < 1$ und $y \in \mathbb{R}$ ist dann

$$\begin{aligned} f(x, y) &= e^{xy} \frac{1}{1-x} \\ &= \left(\sum_{k=0}^{\infty} \frac{1}{k!} (xy)^k \right) \left(\sum_{\ell=0}^{\infty} x^\ell \right) \\ &= \sum_{k=0}^{\infty} \sum_{\ell=0}^{\infty} \frac{1}{k!} x^{k+\ell} y^k \\ &= \sum_{k=0}^{\infty} \sum_{j=k}^{\infty} \frac{1}{k!} x^j y^k \\ &= \sum_{(j,k) \in \mathbb{N}_0^2} c_{jk} x^j y^k \end{aligned}$$

mit

$$c_{jk} = \begin{cases} \frac{1}{k!} & \text{falls } j \geq k, \\ 0 & \text{sonst.} \end{cases}$$