Vorlesungsmanuskript zu

Höhere Mathematik für Elektrotechniker II

Werner Balser
Institut für Angewandte Analysis

Sommersemester 2009

Inhaltsverzeichnis

Kapitel 1

Integralrechnung

Die folgenden Abschnitte sind zum Teil nur eine Wiederholung von Stoff des ersten Semesters! Dabei sind viele Einzelheiten für einen Anwender der Mathematik nicht so interessant – wichtig ist aber ein gewisses Verständnis für die Definition eines Integrals, da wir im folgenden Semester auch mehrdimensionale Integrale betrachten wollen, deren Definition sehr ähnlich ist.

Beachte, dass in dieser Vorlesung 0 *keine* natürliche Zahl ist, obwohl es eine DIN-Norm gibt, die das Gegenteil sagt! Die Menge der ganzen bzw. natürlichen Zahlen wird auch mit $\mathbb{Z}$ bzw. $\mathbb{N}$ bezeichnet. Weiter steht $\mathbb{Q}$ für die Menge aller rationalen Zahlen, also die Menge aller Brüche der Form p/q, wobei p eine ganze und q eine natürliche Zahl ist. Schließlich schreiben wir $\mathbb{R}$ bzw. $\mathbb{C}$ für die Menge aller reellen bzw. komplexen Zahlen.

1.1 Riemann-Summen und Riemann-Integral

Wenn nichts anderes gesagt wird, betrachten wir im Folgenden immer Funktionen $f : [a,b] \longrightarrow \mathbb{R}$ mit festen reellen Zahlen $a < b$.

Definition 1.1.1 *Eine Menge* $\mathcal{Z} = \{x_0, \ldots, x_N\}$*,* $N \in \mathbb{N}$*, heißt* Zerlegung *von* $[a,b]$*, wenn gilt*

$$a = x_0 \; < \; x_1 \; < \ldots < \; x_{N-1} \; < \; x_N = b.$$

Die Zahl $|\mathcal{Z}| = \max\{x_k - x_{k-1} : 1 \le k \le N\}$ *heißt die* Feinheit *der Zerlegung, während* N die Zahl der Teilintervalle *genannt werden soll. Die Zahlen* x_k *heißen auch* die Teilpunkte der Zerlegung. *Ein Vektor* $\xi = (\xi_1, \ldots, \xi_N)^T$ *heißt ein* Zwischenpunktvektor *zu* $\mathcal{Z}$*, falls*

$$x_{k-1} \; \le \; \xi_k \; \le \; x_k \qquad \forall\; k = 1, \ldots, N.$$

Für jede Zerlegung $\mathcal{Z}$ *und jeden Zwischenpunktvektor* ξ *zu* $\mathcal{Z}$ *heißt die Zahl*

$$S(\mathcal{Z}, \xi) \;=\; \sum_{k=1}^{N} f(\xi_k)\,(x_k \;-\; x_{k-1})$$

die zu $\mathcal{Z}$ und ξ gehörige Riemannsumme *von* f*. Eine Folge* $(\mathcal{Z}_n)$ *von Zerlegungen heißt* zulässig*, falls* $\lim\limits_{n\to\infty} |\mathcal{Z}_n| = 0$ *gilt. Falls für jede zulässige Zerlegungsfolge und jede Wahl von Zwischenpunktvektoren die zugehörige Folge von Riemannsummen konvergiert, dann heißt* f über das Intervall $[a,b]$ Riemann-integrierbar*. In der nächsten Behauptung wird gezeigt, dass der Grenzwert der Riemannsummen in diesem*

Fall nicht von der Wahl der Zerlegungsfolge oder der Zwischenpunktvektoren abhängt. Diesen Wert nennen wir dann das (Riemann-)Integral *von f über $[a,b]$ und schreiben*

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

Für die geometrische Bedeutung von Riemannsumme und Integral, siehe die untenstehende Bemerkung.

Behauptung 1.1.2 *Wenn für jede zulässige Zerlegungsfolge und jede Wahl von Zwischenpunktvektoren die zugehörige Folge von Riemannsummen konvergiert, dann ist der Grenzwert immer der gleiche, also unabhängig von der Wahl der Zerlegungsfolge oder der Zwischenpunktvektoren.*

Beweis: Seien $(\mathcal{Z}_n)$ und $(\tilde{\mathcal{Z}}_n)$ zwei zulässige Zerlegungsfolgen, und seien (ξ_n) bzw. $(\tilde{\xi}_n)$ zugehörige Folgen von Zwischenpunktvektoren. Durch *Mischen* der Folgen bilden wir zwei weitere Folgen $(\hat{\mathcal{Z}}_n)$ und $(\hat{\xi}_n)$ mit

$$\hat{\mathcal{Z}}_n \;=\; \begin{cases} \mathcal{Z}_k & (n=2k) \\ \tilde{\mathcal{Z}}_k & (n=2k+1) \end{cases}\,, \qquad \hat{\xi}_n \;=\; \begin{cases} \xi_k & (n=2k) \\ \tilde{\xi}_k & (n=2k+1) \end{cases}\,.$$

Dabei entsteht wieder eine zulässige Zerlegungsfolge mit zugehörigen Zwischenpunktvektoren, welche beide Ausgangsfolgen als Teilfolgen beisitzt. Da die zu $(\hat{\mathcal{Z}}_n)$ und $(\hat{\xi}_n)$ gehörigen Riemannsummen konvergieren müssen, müssen also die beiden Teilfolgen den gleichen Grenzwert besitzen, was zu beweisen war. □

Aufgabe 1.1.3 *Zeige: Die sogenannte* Dirichletsche Sprungfunktion

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

ist nicht über $[a,b]$ integrierbar. Es darf bei der Lösung dieser Aufgabe benutzt werden, dass zwischen zwei reellen Zahlen immer eine rationale, aber auch eine irrationale Zahl liegt.

Aufgabe 1.1.4 *Zeige: Eine konstante Funktion ist über jedes abgeschlossene Intervall integrierbar. Finde den Wert des Integrals!*

Bemerkung 1.1.5 *Wenn $f(x) \geq 0$ ist für alle $x \in [a,b]$, dann ist eine Riemannsumme die Summe der Flächeninhalte von Rechtecken mit Breite $x_k - x_{k-1}$ und Höhe $f(\xi_k)$. Daraus ergibt sich, dass das Integral von f über $[a,b]$ den Flächeninhalt der Figur in einer (x,y)-Ebene unterhalb des Graphen von f bzw. oberhalb der x-Achse und mit seitlichen Begrenzungslinien $x=a$ bzw. $x=b$ definiert.*

Die folgende Proposition erlaubt, dass wir im nächsten Abschnitt zur weiteren Untersuchung der Integrierbarkeit nur beschränkte Funktionen betrachten:

Proposition 1.1.6 *Für Integrale gelten die folgenden Aussagen:*

(a) **(Fundamentalabschätzung)** *Jede über $[a,b]$ integrierbare Funktion f ist dort beschränkt, und es gilt*

$$\left|\int_a^b f(x)\,dx\right| \;\leq\; (b\;-\;a)\;\sup\{\,|f(x)|\,:\,a\leq x\leq b\,\}\,.$$

(b) **(Linearität des Integrals)** *Sind f und g über $[a,b]$ integrierbar, und sind $\alpha,\beta\in\mathbb{R}$, so ist auch $\alpha\,f+\beta\,g$ über $[a,b]$ integrierbar, und es gilt*

$$\int_a^b (\alpha\,f(x)+\beta\,g(x))\,dx \;=\; \alpha\int_a^b f(x)\,dx \;+\; \beta\int_a^b g(x)\,dx\,.$$

Beweis: Die Fundamentalabschätzung folgt, da der Betrag jeder Riemannsumme höchstens gleich der rechten Seite ist (auch falls f unbeschränkt sein sollte). Sei jetzt f auf $[a,b]$ nach oben unbeschränkt, und sei $\mathcal{Z}_n$ eine zulässige Zerlegungsfolge. Für jedes n ist dann f in mindestens einem Teilintervall von $\mathcal{Z}_n$ nach oben unbeschränkt, und deshalb kann man in einem solchen Teilintervall den Zwischenpunkt so wählen, dass die zugehörige Riemannsumme $\geq n$ ausfällt. Deshalb kann f nicht integrierbar sein. Durch Übergang zu $-f$ folgt, dass auch nach unten unbeschränkte Funktionen nicht integrierbar sind. Also gilt (a). Teil (b) ist aber unmittelbar klar wegen der Definition des Integrals. $\square$

1.2 Ober- und Untersummen

Definition 1.2.1 *Sei f auf $[a,b]$ beschränkt, und sei $\mathcal{Z} = \{x_0, \ldots, x_N\}$ eine Zerlegung von $[a,b]$. Für*

$$m_k = \inf\{f(x) : x_{k-1} \leq x \leq x_k\}, \quad M_k = \sup\{f(x) : x_{k-1} \leq x \leq x_k\}$$

heißen

$$U(\mathcal{Z}) \;=\; \sum_{k=1}^{N} m_k\,(x_k \;-\; x_{k-1}), \quad O(\mathcal{Z}) \;=\; \sum_{k=1}^{N} M_k\,(x_k \;-\; x_{k-1})$$

die zu $\mathcal{Z}$ gehörige Unter- und Obersumme *von f. Offenbar gilt für alle Zwischenpunktvektoren immer*

$$U(\mathcal{Z}) \;\leq\; S(\mathcal{Z},\xi) \;\leq\; O(\mathcal{Z}).$$

Seien zwei Zerlegungen $\mathcal{Z}_1, \mathcal{Z}_2$ von $[a,b]$ gegeben. Wir nennen $\mathcal{Z}_2$ Verfeinerung von $\mathcal{Z}_1$*, wenn $\mathcal{Z}_2$ alle Teilpunkte von $\mathcal{Z}_1$ enthält. Wir schreiben $\mathcal{Z}_1 + \mathcal{Z}_2$ für diejenige Zerlegung, welche genau aus allen Teilpunkten von $\mathcal{Z}_1$ und $\mathcal{Z}_2$ besteht, und nennen $\mathcal{Z}_1 + \mathcal{Z}_2$* Überlagerung von $\mathcal{Z}_1$ *und $\mathcal{Z}_2$.*

Aufgabe 1.2.2 *Finde den Zusammenhang zwischen den Obersummen von f und den Untersummen der Funktion $-f$.*

Lemma 1.2.3 *Sei $|f(x)| \leq K$ für alle $x \in [a,b]$, und sei $\mathcal{Z}_0$ eine Zerlegung von $[a,b]$ mit N Teilintervallen. Dann gilt für jede Zerlegung $\mathcal{Z}$*

$$\begin{aligned} U(\mathcal{Z}) \;&\leq\; U(\mathcal{Z}+\mathcal{Z}_0) \;\leq\; U(\mathcal{Z}) \;+\; 2\,N\,K\,|\mathcal{Z}|, \\ O(\mathcal{Z}) \;&\geq\; O(\mathcal{Z}+\mathcal{Z}_0) \;\geq\; O(\mathcal{Z}) \;-\; 2\,N\,K\,|\mathcal{Z}|. \end{aligned}$$

(Ohne Beweis)

Korollar zu Lemma 1.2.3 *Sei f auf $[a,b]$ beschränkt. Für beliebige Zerlegungen $\mathcal{Z}_j$ von $[a,b]$ gilt dann stets*

$$U(\mathcal{Z}_1) \;\leq\; O(\mathcal{Z}_2)\,.$$

Beweis: Für $\mathcal{Z} = \mathcal{Z}_1$, $\mathcal{Z}_0 = \mathcal{Z}_2$ folgt aus dem Lemma $U(\mathcal{Z}_1) \leq U(\mathcal{Z}_1 + \mathcal{Z}_2)$. Trivialerweise gilt $U(\mathcal{Z}_1 + \mathcal{Z}_2) \leq O(\mathcal{Z}_1 + \mathcal{Z}_2)$, und durch Vertauschen von $\mathcal{Z}_1, \mathcal{Z}_2$ folgt aus dem Lemma $O(\mathcal{Z}_1 + \mathcal{Z}_2) \leq O(\mathcal{Z}_2)$. $\square$

Definition 1.2.4 *Sei f auf $[a,b]$ beschränkt. Nach obigem Korollar sind die Obersummen (Untersummen) von f stets nach unten (oben) beschränkt. Wir definieren deshalb*

$$I_* \;=\; \underline{\int_a^b} f(x)\,dx \;=\sup\{U(\mathcal{Z})\}, \qquad I^* \;=\; \overline{\int_a^b} f(x)\,dx \;=\; \inf\{O(\mathcal{Z})\},$$

wobei das Supremum bzw. Infimum jeweils über alle Zerlegungen von $[a,b]$ gebildet wird. Die so definierten Werte I_*, I^* heißen das Unter- bzw. Oberintegral von f *über $[a,b]$. Offenbar ist $I_* \leq I^*$, und wir werden sehen, dass f genau dann über $[a,b]$ integrierbar ist, wenn $I_* = I^*$ gilt.*

Wir charakterisieren jetzt das Riemann-Integral durch Grenzwerte von Ober- und Untersummen:

Satz 1.2.5 *Für jede auf $[a, b]$ beschränkte Funktion f gilt:*

(a) *Für jede zulässige Zerlegungsfolge $(\mathcal{Z}_n)$ konvergieren die Folgen $(U(\mathcal{Z}_n))$ und $(O(\mathcal{Z}_n))$, und es gilt*

$$\underline{\int}_a^b f(x)\,dx = \lim_{n\to\infty} U(\mathcal{Z}_n), \qquad \overline{\int}_a^b f(x)\,dx = \lim_{n\to\infty} O(\mathcal{Z}_n).$$

(b) *Die Funktion f ist genau dann integrierbar über $[a, b]$, wenn ihr Unter- und Oberintegral übereinstimmen, und dann ist*

$$\int_a^b f(x)\,dx = \underline{\int}_a^b f(x)\,dx = \overline{\int}_a^b f(x)\,dx\ .$$

(Ohne Beweis)

Proposition 1.2.6 (Riemannsches Integrabilitätskriterium)
Eine beschränkte Funktion $f : [a, b] \longrightarrow \mathbb{R}$ ist genau dann über $[a, b]$ integrierbar, wenn es zu jedem $\varepsilon > 0$ eine Zerlegung $\mathcal{Z}$ von $[a, b]$ gibt mit $O(\mathcal{Z}) - U(\mathcal{Z}) < \varepsilon$.

Beweis: Sei f integrierbar, dann folgt aus dem vorausgegangenen Satz für jede zulässige Zerlegungsfolge $(\mathcal{Z}_n)$, dass $\lim_{n\to\infty}(O(\mathcal{Z}_n) - U(\mathcal{Z}_n)) = 0$, und daraus folgt die eine Richtung der Behauptung. Umgekehrt folgt aus der Ungleichung $O(\mathcal{Z}) - U(\mathcal{Z}) < \varepsilon$, dass $I^* - I_* < \varepsilon$ ist, und da ε beliebig klein sein kann, folgt die Gleichheit von Ober- und Unterintegral, also die Integrierbarkeit. □

Mit diesem Kriterium zeigt man die folgenden Resultate:

Satz 1.2.7 *Jede auf $[a, b]$ stetige, aber auch jede monotone Funktion ist integrierbar.*

(Ohne Beweis)

Satz 1.2.8 *Sind f und g über $[a, b]$ integrierbar, so gilt dasselbe auch für $f\,g$.*

(Ohne Beweis)

Lemma 1.2.9 *Sei $f : [a, b] \longrightarrow A \subset \mathbb{R}$ über $[a, b]$ integrierbar, und gelte für $g : A \to \mathbb{R}$ eine Lipschitzbedingung auf A. Dann ist $g \circ f$ über $[a, b]$ integrierbar.*

(Ohne Beweis)

Proposition 1.2.10 (Dreiecksungleichung für Integrale)
Ist f über $[a, b]$ integrierbar, so auch $|f|$, und es gilt

$$\left|\ \int_a^b f(x)\,dx\ \right| \le \int_a^b |f(x)|\,dx.$$

Beweis: Für $g(x) = |x|$ folgt aus dem obigen Lemma die Integrierbarkeit von $|f|$, und die Ungleichung folgt aus der Dreiecksungleichung für die Riemannsummen. □

Aufgabe 1.2.11 *Sei f über $[a,b]$ integrierbar, und sei $p > 1$. Zeige mit* Lemma 1.2.9 *die Integrierbarkeit von $|f|^p$ über $[a,b]$.*

Satz 1.2.12 *Für $a < c < b$ ist eine Funktion f genau dann über $[a,b]$ integrierbar, wenn sie sowohl über $[a,c]$ als auch über $[c,b]$ integrierbar ist, und dann gilt*

$$\int_a^b f(x)\,dx \;=\; \int_a^c f(x)\,dx \;+\; \int_c^b f(x)\,dx\,.$$

Beweis: Aus Lemma 1.2.3 folgt, dass bei Verfeinerung einer Zerlegung die Untersummen bzw. Obersummen nicht abnehmen bzw. nicht zunehmen können. Daraus ergibt sich, dass wir o. B. d. A. nur Zerlegungen betrachten können, die c als einen Teilpunkt enthalten. Damit folgt die Behauptung mit dem Riemannschen Integrabilitätskriterium. □

Aufgabe 1.2.13 *Zeige, dass stückweise stetige Funktionen immer integrierbar sind.*

Bemerkung 1.2.14 *Sei f über $[a,b]$ integrierbar, und seien $x_0, x_1 \in [a,b]$. Ist $x_0 < x_1$, so folgt aus dem vorstehenden Satz die Integrierbarkeit von f über $[x_0, x_1]$. Es ist üblich,*

$$\int_{x_1}^{x_0} f(x)\,dx \;=\; -\int_{x_0}^{x_1} f(x)\,dx\,, \qquad \int_{x_0}^{x_0} f(x)\,dx \;=\; 0$$

zu setzen. Mit diesen Vereinbarungen folgt dann für beliebige $x_0, x_1, x_2 \in [a,b]$ aus obigem Satz

$$\int_{x_0}^{x_1} f(x)\,dx \;+\; \int_{x_1}^{x_2} f(x)\,dx \;=\; \int_{x_0}^{x_2} f(x)\,dx\,.$$

Lemma 1.2.15 *Sei f über $[a,b]$ integrierbar, und sei $x_0 \in [a,b]$, sowie*

$$F(x) \;=\; \int_{x_0}^{x} f(t)\,dt \qquad \forall\; x \in [a,b]\,.$$

Dann erfüllt F eine Lipschitzbedingung auf $[a,b]$ und ist insbesondere dort stetig.

Beweis: Für $x_1, x_2 \in [a,b]$ folgt aus der obigen Bemerkung, dass $F(x_1) - F(x_2) = \int_{x_2}^{x_1} f(t)\,dt$ ist. Aus der Fundamentalabschätzung folgt dann die Behauptung mit der Lipschitzkonstanten $L = \sup_{[a,b]} |f(t)|$. □

Aufgabe 1.2.16 *Gib ein Beispiel einer über $[a,b]$ integrierbaren Funktion f, für welche die in* Lemma 1.2.15 *definierte Funktion F* nicht *auf $[a,b]$ differenzierbar ist.*

1.3 Die Hauptsätze der Differenzial- und Integralrechnung

Definition 1.3.1 *Sei I ein beliebiges Intervall, und sei $f : I \longrightarrow \mathbb{R}$. Falls eine auf I differenzierbare Funktion F existiert mit der Eigenschaft $F'(x) = f(x)$ für alle $x \in I$, dann nennen wir F* eine Stammfunktion *zu f. Wir schreiben in diesem Fall auch*

$$F(x) \;=\; \int f(x)\,dx$$

und nennen das rechtsstehende Symbol auch unbestimmtes Integral *von f.*

Aufgabe 1.3.2 *Finde Stammfunktionen zu den Funktionen*

$$\cos x\,, \qquad x^2+1\,, \qquad \frac{1}{\cos^2 x}\,, \qquad \frac{1}{x},$$

jeweils für x auf dem natürlichen Definitionsbereich.

Behauptung 1.3.3 *Ist F auf einem Intervall I Stammfunktion zu f, und ist $c \in \mathbb{R}$, so ist auch $F + c$ eine Stammfunktion zu f. Umgekehrt, sind F_1 und F_2 zwei Stammfunktionen zu f auf I, so ist $F_2 - F_1$ konstant.*

Beweis: Folgt aus dem ersten Mittelwertsatz der Differenzialrechnung. □

Satz 1.3.4 (Erster Hauptsatz der Analysis) *Falls f über $[a, b]$ integrierbar ist und eine Stammfunktion F besitzt, dann gilt*

$$\int_a^b f(x)\,dx \;=\; F(b) \;-\; F(a) \quad \Big(=:\; F(x)\big|_a^b \Big)\,.$$

Beweis: Für eine beliebige Zerlegung $\mathcal{Z} = \{x_0, \dots, x_N\}$ folgt mit dem ersten Mittelwertsatz der Differenzialrechnung

$$F(b) - F(a) = \sum_{k=1}^{N} (F(x_k) - F(x_{k-1})) = \sum_{k=1}^{N} f(\xi_k)\,(x_k - x_{k-1})\,,$$

mit geeigneten Zwischenwerten $\xi_k \in (x_{k-1}, x_k)$. Mit der Definition des Integrals folgt dann die Behauptung. □

Aufgabe 1.3.5 *Berechne die Integrale $\int_0^1 (x^2 + 1)\,dx$ und $\int_0^{\pi/2} \cos x\,dx$.*

Satz 1.3.6 (Zweiter Hauptsatz der Analysis) *Sei f auf $[a, b]$ stetig, und sei $x_0 \in [a, b]$. Dann ist die Funktion*

$$F(x) \;=\; \int_{x_0}^{x} f(t)\,dt \quad \forall\; x \in [a, b]$$

eine Stammfunktion zu f.

Beweis: Bemerkung 1.2.14 stellt zunächst sicher, dass F immer definiert ist, und dass

$$\frac{F(x + h) - F(x)}{h} \;=\; \frac{1}{h} \int_x^{x+h} f(t)\,dt\,.$$

Weiter ist $\int_x^{x+h} f(x)\,dt = h\,f(x)$. Also folgt mit der Fundamentalabschätzung

$$\left| \frac{F(x + h) - F(x)}{h} \;-\; f(x) \right| \;=\; \left| \frac{1}{h} \int_x^{x+h} (f(t) \;-\; f(x))\,dt \right| \;\le\; m_x(h)\,,$$

wobei $m_x(h)$ das Maximum von $|f(t) - f(x)|$ auf dem von x und $x + h$ begrenzten Intervall bezeichnet. Da f stetig ist, gilt $m_x(h) \to 0$ für $h \to 0$. Daraus folgt die Behauptung. □

1.4 Weitere Ergebnisse

Satz 1.4.1 (Partielle Integration) *Seien f, g über $[a, b]$ integrierbar, und seien F bzw. G Stammfunktionen zu f bzw. g. Dann gilt*

$$\int_a^b f(x)\, G(x)\, dx \;=\; F(x)\, G(x)\big|_a^b \;-\; \int_a^b F(x)\, g(x)\, dx.$$

Beweis: Es ist $(F(x)\, G(x))' = f(x)\, G(x) + F(x)\, g(x)$, und da F, G differenzierbar (also insbesondere stetig) sind, folgt die Integrierbarkeit von $f\, G$ und $F\, g$ (also auch die von $(F\, G)'$). Mit dem ersten Hauptsatz folgt dann die Behauptung. □

Aufgabe 1.4.2 *Berechne die Integrale*

$$\int_0^\pi x\, \cos x\, dx\,, \qquad\qquad \int_0^{2\pi} \cos x\, \sin x\; dx$$

mit Hilfe von partieller Integration.

Satz 1.4.3 (Substitutionsregel) *Seien f stetig auf $[a, b]$ und g stetig differenzierbar auf $[\alpha, \beta]$, und gelte $g([\alpha, \beta]) \subset [a, b]$ sowie*

$$g(\alpha) = a, \quad g(\beta) = b, \quad \text{oder} \quad g(\alpha) = b, \quad g(\beta) = a.$$

Dann gilt

$$\int_{g(\alpha)}^{g(\beta)} f(x)\, dx \;=\; \int_\alpha^\beta f(g(t))\, g'(t)\, dt.$$

Beweis: Sei F Stammfunktion zu f. Mit der Kettenregel folgt $(F(g(t)))' = f(g(t))\, g'(t)$. Also gilt nach dem ersten Hauptsatz

$$\int_{g(\alpha)}^{g(\beta)} f(x)\, dx \;=\; F(g(\beta)) \;-\; F(g(\alpha)) \;=\; \int_\alpha^\beta f(g(t))\, g'(t)\, dt.$$

Das ist die Behauptung. □

Aufgabe 1.4.4 *Berechne* $\int_0^1 x\, e^{x^2}\, dx$ *und* $\int_0^1 \sqrt{1 - x^2}\, dx$ *mit der Substitutionsregel.*

Aufgabe 1.4.5 *Zeige für $t = \tan(x/2)$ die Gleichungen*

$$\cos^2(x/2) \;=\; \frac{1}{1 + t^2}\;, \quad \sin^2(x/2) \;=\; \frac{t^2}{1 + t^2}\;,$$

und schließe daraus mit Hilfe der Additionstheoreme für die trigonometrischen Funktionen, dass

$$\cos x \;=\; \frac{1 - t^2}{1 + t^2}\;, \quad \sin x \;=\; \frac{2\, t}{1 + t^2}\;.$$

Zeige damit, dass Integrale über eine rationale Funktion von $\sin x$ und $\cos x$ durch die Substitution $x = 2 \arctan t$ in ein Integral verwandelt werden, dessen Integrand eine rationale Funktion von t ist.

Aufgabe 1.4.6 *Zeige mit der Substitutionsregel* $\int_0^a f(x)\,dx = \int_{-a}^0 f(-x)\,dx$*. Benutze dies, um zu zeigen, dass das Integral einer ungeraden Funktion über ein Intervall* $[-a,a]$ *immer* 0 *ergibt.*

Lemma 1.4.7 (Integration von Ungleichungen)

(a) *Seien* f, g *über* $[a,b]$ *integrierbar, und gelte* $f(x) \le g(x)$ *für alle* $x \in [a,b]$*. Dann folgt* $\int_a^b f(x)\,dx \le \int_a^b g(x)\,dx$*.*

(b) *Sei* f *stetig und nicht-negativ auf* $[a,b]$*, und gelte* $\int_a^b f(x)\,dx = 0$*. Dann folgt* $f(x) = 0$ *für alle* $x \in [a,b]$*.*

Beweis: Zu (a): Für jede Zerlegung von $[a,b]$ und jede Wahl von Zwischenpunkten ist die Riemannsumme von f nicht größer als die von g. Daher folgt die Behauptung mit der Definition des Integrals. Zu (b): Die Funktion $F(x) = \int_0^x f(t)\,dt$ ist monoton wachsend, da $F'(x) = f(x) \ge 0$ ist. Wegen $F(a) = F(b) = 0$ folgt also, dass F die Nullfunktion ist, und deshalb gilt dasselbe auch für f. □

Definition 1.4.8 *Sei* $a < b$*, und sei* f *über* $[a,b]$ *integrierbar. Die Zahl*

$$\mu \;=\; \frac{1}{b-a}\int_a^b f(x)\,dx$$

heißt der Mittelwert von f über $[a,b]$.

Satz 1.4.9 (Erster Mittelwertsatz der Integralrechnung)
Sei $a < b$*, sei* f *über* $[a,b]$ *integrierbar, und sei* μ *der Mittelwert von* f *über* $[a,b]$*. Dann ist*

$$m = \inf\{f(x) : a \le x \le b\} \;\le\; \mu \;\le\; \sup\{f(x) : a \le x \le b\} \;=\; M.$$

Falls f *auf* $[a,b]$ *stetig ist, existiert ein* $\xi \in [a,b]$ *mit* $\mu = f(\xi)$*.*

Beweis: Wegen $m \le f(x) \le M$ folgt die Ungleichung aus Lemma 1.4.7 (a), und der Zusatz ergibt sich aus dem Zwischenwertsatz. □

Satz 1.4.10 (Erweiterter Mittelwertsatz der Integralrechnung)
Seien g *und* $f\,g$ *über* $[a,b]$ *integrierbar, und sei* $g(x) \ge 0$ *für alle* $x \in [a,b]$*. Dann existiert ein* $\mu \in [m,M]$*, mit* m*,* M *wie im vorigen Satz, so dass*

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

Ist f *stetig auf* $[a,b]$*, so gibt es ein* $\xi \in [a,b]$ *mit* $f(\xi) = \mu$*.*

Beweis: Es gilt

$$m\int_a^b g(x)\,dx \;\le\; \int_a^b f(x)\,g(x)\,dx \;\le M \int_a^b g(x)\,dx\,,$$

und daraus folgt die Behauptung. □

Satz 1.4.11 (Zweiter Mittelwertsatz der Integralrechnung)
Seien f *stetig differenzierbar und monoton, sowie* g *stetig auf* $[a,b]$*. Dann existiert ein* $\xi \in [a,b]$ *mit*

$$\int_a^b f(x)\,g(x)\,dx \;=\; f(a)\int_a^\xi g(x)\,dx \;+\; f(b)\int_\xi^b g(x)\,dx\,.$$

Beweis: O. B. d. A. sei f wachsend, also $f'(x) \geq 0$. Sei G Stammfunktion von g, dann folgt mit partieller Integration

$$\int_a^b f(x)\, g(x)\, dx \;=\; f(x)\, G(x)|_a^b \;-\; \int_a^b f'(x)\, G(x)\, dx\,.$$

Aus dem vorherigen Satz und dem Zwischenwertsatz folgt die Existenz von ξ mit

$$\int_a^b f'(x)\, G(x)\, dx \;=\; G(\xi) \int_a^b f'(x)\, dx \;=\; G(\xi)\; f(x)|_a^b\,.$$

Daraus ergibt sich die Behauptung. □

Satz 1.4.12 (Gliedweise Integration) *Gegeben sei ein abgeschlossenes Intervall $[a,b] \subset \mathbb{R}$, sowie Funktionen $f_n, g_k : [a,b] \longrightarrow \mathbb{K}$ für alle $n, k \in \mathbb{N}$. Dann gilt:*

(a) *Sind alle f_n über $[a,b]$ integrierbar, und ist die Funktionenfolge (f_n) auf $[a,b]$ gleichmäßig konvergent, so ist auch die Grenzfunktion f über $[a,b]$ integrierbar, und es gilt*

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

(b) *Sind alle g_k über $[a,b]$ integrierbar, und ist die Funktionenreihe $\sum_{k=1}^{\infty} g_k$ auf $[a,b]$ gleichmäßig konvergent, so ist auch die Grenzfunktion f über $[a,b]$ integrierbar, und es gilt*

$$\int_a^b f(x)\, dx \;=\; \sum_{k=1}^{\infty} \int_a^b g_k(x)\, dx.$$

(Ohne Beweis)

Aufgabe 1.4.13 *Finde eine Potenzreihenentwicklung für* $\arctan x$ *und Entwicklungspunkt* 0.

Lösung: Aus dem ersten Hauptsatz folgt wegen $\arctan 0 = 0$, dass

$$\arctan x \;=\; \int_0^x \frac{1}{1+t^2}\; dt \qquad \forall\; x \in \mathbb{R}.$$

Aus der geometrischen Reihe folgt

$$\frac{1}{1+t^2} \;=\; \sum_{k=0}^{\infty} (-1)^k\, t^{2k} \qquad \forall\; |t| < 1,$$

und die Reihe ist sogar gleichmäßig konvergent für $|t| \leq r$, mit beliebigem $r < 1$. Deshalb folgt aus Satz 1.4.12, dass

$$\arctan x \;=\; \sum_{k=0}^{\infty} (-1)^k\, \frac{x^{2k+1}}{2k+1}\;,$$

zunächst nur für $|x| \leq r$, aber da ja r beliebig dicht bei 1 sein kann, ist dies sogar richtig für alle x mit $|x| < 1$. □

1.5 Die Partialbruchzerlegung

Auf der einen Seite garantiert der zweite Hauptsatz zu jeder stetigen Funktion f die Existenz einer Stammfunktion F, andererseits gibt er keinen Hinweis darauf, wie man F explizit berechnen kann. Tatsächlich ist es so, dass man z. B. für $f(x) = \exp(-x^2)$ keine Stammfunktion durch die uns bekannten Funktionen ausdrücken kann. In diesem Abschnitt wollen wir allerdings zeigen, dass man beliebige rationale Funktionen in einfache Ausdrücke zerlegen kann, für welche man dann einzeln Stammfunktionen findet. Dazu benötigen wir einige Hilfsmittel:

Satz 1.5.1 (Fundamentalsatz der Algebra)
Jedes nicht konstante Polynom $q \in \mathbb{C}[z]$ besitzt mindestens eine Nullstelle in $\mathbb{C}$.

(Ohne Beweis)

Korollar zu Satz 1.5.1 *Zu jedem $q \in \mathbb{C}[z]$ vom Grade $n \geq 1$ existieren verschiedene $z_1, \ldots, z_m \in \mathbb{C}$, natürliche Zahlen $\nu_1, \ldots, \nu_m$ mit $\nu_1 + \ldots + \nu_m = n$, und ein $a \in \mathbb{C} \setminus \{0\}$ derart, dass*

$$q(z) \;=\; a\,(z - z_1)^{\nu_1} \cdot \ldots \cdot (z - z_m)^{\nu_m} \qquad \forall\, z \in \mathbb{C}. \tag{1.5.1}$$

(Ohne Beweis)

Satz 1.5.2 *Zu jedem reellen Polynom $p \in \mathbb{R}[x]$ vom Grade $n \geq 1$ existieren verschiedene $x_1, \ldots, x_\mu \in \mathbb{R}$, verschiedene $(a_1, b_1), \ldots, (a_\nu, b_\nu) \in \mathbb{R}^2$, wobei auch $\mu = 0$ oder $\nu = 0$ eintreten kann, eine reelle Zahl $a \neq 0$, sowie natürliche Zahlen $\nu_1, \ldots, \nu_\mu, \sigma_1, \ldots, \sigma_\nu$ derart, dass folgendes gilt:*

(a) *Es ist $\sum_{k=1}^{\mu} \nu_k + 2 \sum_{k=1}^{\nu} \sigma_k = n$.*

(b) *Es ist $a_k^2 < 4\, b_k$, d. h., $x^2 - a_k\, x + b_k$ hat keine reellen Nullstellen, für alle $k = 1, \ldots, \nu$.*

(c) *Es gilt die Darstellung*

$$\begin{aligned} p(x) \;&=\; a\,(x - x_1)^{\nu_1} \cdot \ldots \cdot (x - x_\mu)^{\nu_\mu}\; (x^2 - a_1\, x + b_1)^{\sigma_1} \\ &\quad\; \cdot \ldots \cdot (x^2 - a_\nu\, x + b_\nu)^{\sigma_\nu} \qquad \forall\, x \in \mathbb{R}. \end{aligned} \tag{1.5.2}$$

(Ohne Beweis)

Satz 1.5.3 (Partialbruchzerlegung im Komplexen) *Seien $p, q \in \mathbb{C}[z]$ mit $0 \leq \deg p < \deg q$, und gelte* (1.5.1). *Dann gibt es eindeutig bestimmte Zahlen $a_{jk} \in \mathbb{C}$ mit*

$$r(z) \;=\; \frac{p(z)}{q(z)} \;=\; \sum_{k=1}^{m} \sum_{j=1}^{\nu_k} \frac{a_{jk}}{(z - z_k)^j} \qquad \forall\, z \in \mathbb{C} \setminus \{z_1, \ldots, z_m\}. \tag{1.5.3}$$

(Ohne Beweis)

Satz 1.5.4 (Partialbruchzerlegung im Reellen) *Seien $p, q \in \mathbb{R}[x]$ mit $0 \leq \deg p < \deg q$, und gelte* (1.5.2). *Dann gibt es eindeutig bestimmte Zahlen a_{jk}, α_{jk}, $\beta_{jk} \in \mathbb{R}$ so, dass*

$$r(x) \;=\; \sum_{k=1}^{\mu} \sum_{j=1}^{\nu_k} \frac{a_{jk}}{(x - x_k)^j} \;+\; \sum_{k=1}^{\nu} \sum_{j=1}^{\sigma_k} \frac{\alpha_{jk}\, x + \beta_{jk}}{(x^2 - a_k\, x + b_k)^j} \qquad \forall\, x \in D, \tag{1.5.4}$$

mit $D = \mathbb{R} \setminus \{x_1, \ldots, x_\mu\}$.

(Ohne Beweis)

Bemerkung 1.5.5 (Durchführung der Partialbruchzerlegung)
Hat man den entsprechenden Ansatz für die Partialbruchzerlegung gemacht, so multipliziert man am besten beide Seiten mit dem Nennerpolynom $q(x)$ und erhält so eine äquivalente Polynomgleichung. Danach bieten sich folgende Möglichkeiten an:

1. Durch Sammeln aller Terme mit gleichen Potenzen der Variablen, also durch Koeffizientenvergleich, erhält man ein lineares Gleichungssystem, welches in jedem Fall eindeutig gelöst werden kann.
2. Durch Einsetzen geeigneter Werte, vor allem der Nullstellen des Nennerpolynoms, erhält man Gleichungen, welche in der Regel sehr einfach nach einer der Unbekannten auflösbar sind. In jedem Fall folgt aus dem Nullstellensatz für Polynome, dass sich durch Einsetzen von n verschiedenen Werten ein Gleichungssystem ergibt, welches eindeutig lösbar ist.
3. Durch Differenzieren der Polynomgleichung kann man eine neue, einfachere Gleichung erhalten, aus der man einige Unbekannte bestimmen kann.

Man kann auch Schritte der drei obigen Typen mischen, wenn man die bereits berechneten Werte für einige der Unbekannten in die Polynomgleichung einsetzt. In jedem Fall muss man ein Gleichungssystem in n Unbekannten lösen, und der obige Satz besagt gerade, dass diese Unbekannten eindeutig bestimmbar sind. Siehe dazu auch die folgenden Beispiele und Aufgaben.

Beispiel 1.5.6 *Sei $r(x) = 1/(x^2 - x)$. Da der Nenner die Nullstellen 0 und 1 hat, welche beide die Vielfachheit 1 haben, folgt aus obigem Satz, dass es eindeutig bestimmte Zahlen a und b aus $\mathbb{R}$ geben muss, für welche*

$$r(x) \;=\; \frac{1}{x^2 - x} \;=\; \frac{a}{x} \;+\; \frac{b}{x-1} \qquad \forall\, x \in \mathbb{R} \setminus \{0, 1\}.$$

Um sie zu finden, multiplizieren wir die Gleichung mit dem Nennerpolynom und erhalten die äquivalente Gleichung

$$1 \;=\; a\,(x-1) \;+\; b\,x \qquad \forall\, x \in \mathbb{R}.$$

Einsetzen von $x = 0$ ergibt $a = -1$, einsetzen von $x = 1$ ergibt $b = 1$.

Sei jetzt $r(x) = 1/(x^3 + x)^2$. Jetzt sind sechs Zahlen $a, b, c, d, e, f \in \mathbb{R}$ zu finden mit

$$r(x) \;=\; \frac{1}{(x^3+x)^2} \;=\; \frac{a}{x} \;+\; \frac{b}{x^2} \;+\; \frac{c\,x + d}{x^2+1} \;+\; \frac{e\,x + f}{(x^2+1)^2} \qquad \forall\, x \in \mathbb{R} \setminus \{0\}.$$

Dies ist äquivalent zu

$$1 \;=\; a\,x\,(x^2+1)^2 \;+\; b\,(x^2+1)^2 \;+\; (c\,x+d)\,x^2\,(x^2+1) \;+\; (e\,x+f)\,x^2 \quad \forall\, x \in \mathbb{R}.$$

Für $x = 0$ folgt $b = 1$. Da die Gleichung auch für komplexe x richtig sein muss, erhält man durch Einsetzen von $\pm i$ die Gleichungen

$$1 \;=\; -(e\,i + f) \;=\; -(-e\,i + f).$$

Daraus folgt $e = 0$, $f = -1$. Bringt man die schon gefundenen Terme auf die linke Seite, so erhält man

$$-x^2\,(x^2+1) \;=\; a\,x\,(x^2+1)^2 \;+\; (c\,x+d)\,x^2\,(x^2+1) \quad \forall\, x \in \mathbb{R}.$$

Jetzt kann man offenbar durch $x\,(x^2+1)$ teilen und erhält so

$$-x \;=\; a\,(x^2+1) \;+\; (c\,x+d)\,x \quad \forall\, x \in \mathbb{R}.$$

Einsetzen von $x = 0$ liefert $a = 0$. Einsetzen von $\pm i$ ergibt die Gleichungen

$$-1 \;=\; i\,c + d \;=\; -i\,c + d,$$

woraus $d = -1$ und $c = 0$ folgen. Eine Probe ist hier sicher angebracht!

Als Anwendung der Partialbruchzerlegung wollen wir nun das Problem angehen, zu einer beliebigen rationalen Funktion (mit reellen Koeffizienten) eine Stammfunktion zu finden:

Aufgabe 1.5.7 *Gegeben sei eine rationale Funktion $r \in \mathbb{R}(x)$. Finde eine Stammfunktion.*

Lösung: Sei $r = p/q$. Falls $\deg p \geq \deg q$ ist, kann man Polynome $p_1, p_2 \in \mathbb{R}[x]$ finden mit

$$r \; = \; \frac{p}{q} \; = \; p_1 \; + \; \frac{p_2}{q},$$

und $\deg p_2 < \deg q$. Da man zu p_1 sofort eine Stammfunktion angeben kann, soll im Weiteren angenommen sein, dass $\deg p < \deg q$ ist. Dann kann man nach dem Satz über die Partialbruchzerlegung im Reellen die Funktion r in eine Summe von einfachen Termen zerlegen. Es bleibt also zu zeigen, dass man zu jedem einzelnen Term eine Stammfunktion angeben kann. Dies wollen wir jetzt tun, wobei in jedem Fall durch Differenzieren nachgeprüft werden kann, dass die angegebene Funktion F tatsächlich Stammfunktion zu f ist:

1. $f(x) = (x - x_0)^{-1}$ hat die Stammfunktion $F(x) = \log |x - x_0|$.
2. $f(x) = (x - x_0)^{-j}$ hat die Stammfunktion $F(x) = (1 - j)^{-1} \, (x - x_0)^{1-j}$ für alle natürlichen Zahlen $j \geq 2$.
3. Sei jetzt $f(x) = (\alpha \, x + \beta)/(x^2 - a \, x + b)$, wobei $a^2 < 4 \, b$ sei, damit der Nenner keine reelle Nullstelle mehr hat. Wir zerlegen $f = g + h$ mit
$$g(x) \; = \; \frac{\alpha}{2} \, \frac{2 \, x - a}{x^2 - a \, x + b} \; , \quad h(x) \; = \; \frac{\beta + a \, \alpha/2}{x^2 - a \, x + b} \; .$$
Mit der Kettenregel zeigt man, dass $G(x) = (\alpha/2) \, \log |x^2 - a \, x + b|$ Stammfunktion zu g ist, und
$$H(x) \; = \; \frac{2 \, \beta + \alpha \, a}{\sqrt{4 \, b - a^2}} \; \arctan \frac{2 \, x - a}{\sqrt{4 \, b - a^2}}$$
ist Stammfunktion zu h.
4. Als letztes sei $f(x) = (\alpha \, x + \beta)/(x^2 - a \, x + b)^j$ für eine natürliche Zahl $j \geq 2$ und $a^2 < 4 \, b$. Wir zerlegen $f = f_1 + f_2$, mit
$$f_1(x) \; = \; \frac{\alpha}{2} \, \frac{2 \, x - a}{(x^2 - a \, x + b)^j} \; , \quad f_2(x) \; = \; \frac{a \, \alpha/2 + \beta}{(x^2 - a \, x + b)^j} \; .$$
Eine Stammfunktion zu f_1 ist
$$F_1(x) = \frac{-\alpha}{2 \, (j - 1)} \; \frac{1}{(x^2 - a \, x + b)^{j-1}} \; .$$
Es reicht deshalb aus, eine Stammfunktion zu $g_j(x) = (x^2 - a \, x + b)^{-j}$ zu finden. Man rechnet nach, dass gilt
$$(4 \, b - a^2) \, (j - 1) \, g_j(x) \; = \; \frac{2 \, (2 \, j - 3)}{(x^2 - a \, x + b)^{j-1}} + \frac{d}{dx} \, \frac{2 \, x - a}{(x^2 - a \, x + b)^{j-1}} \; .$$
Der zweite Term hat natürlich eine Stammfunktion, während der erste Term, bis auf eine Konstante, gleich g_{j-1} ist. Daher kann man aus dieser Formel rekursiv eine Stammfunktion zu g_j berechnen.

Es sei noch erwähnt, dass man die oben angegebenen Stammfunktionen sowie viele weitere in zahlreichen Formelsammlungen nachschlagen kann. □

Aufgabe 1.5.8 *Führe eine Partialbruchzerlegung im Reellen für folgende rationale Funktionen durch:*

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

Aufgabe 1.5.9 *Finde Stammfunktionen zu den drei rationalen Funktionen aus der vorigen Aufgabe.*

1.6 Der Taylorsche Satz

Im Folgenden betrachten wir ein abgeschlossenes Intervall $[a,b] \subset \mathbb{R}$ und ein $f : [a,b] \longrightarrow \mathbb{R}$.

Lemma 1.6.1 (Mehrfache Stammfunktion) *Sei f stetig auf $[a,b]$, und sei für ein $x_0 \in [a,b]$ und $n \in \mathbb{N}$*

$$F_n(x) \;=\; \int_{x_0}^{x} \frac{(x-t)^{n-1}}{(n-1)!}\, f(t)\, dt \qquad \forall\, x \in [a,b].$$

Dann ist F_n auf $[a,b]$ n-mal stetig differenzierbar. Weiter gilt $F_n^{(k)}(x_0) = 0$ für $0 \leq k \leq n-1$, und die n-te Ableitung von F_n ist f.

(Ohne Beweis)

Satz 1.6.2 (Satz von Taylor) *Sei f auf $[a,b]$ $(n+1)$-mal stetig differenzierbar. Dann gilt für alle $x, x_0 \in [a,b]$*

$$f(x) \;=\; \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}\, (x-x_0)^k \;+\; \frac{1}{n!} \int_{x_0}^{x} (x-t)^n\, f^{(n+1)}(t)\, dt.$$

(Ohne Beweis)

Korollar zu Satz 1.6.2 *Sei f mindestens $2n$-mal stetig differenzierbar auf einem offenen Intervall (a,b), für ein $n \in \mathbb{N}$, und gelte für ein $x_0 \in (a,b)$:*

$$f'(x_0) = f''(x_0) = \ldots = f^{(2n-1)}(x_0) = 0, \quad f^{(2n)}(x_0) \neq 0.$$

Dann hat f im Punkt x_0 ein lokales Extremum, und zwar ein lokales Minimum für $f^{(2n)}(x_0) > 0$, und ein lokales Maximum für $f^{(2n)}(x_0) < 0$.

Beweis: Aus dem Taylorschen Satz folgt

$$f(x) - f(x_0) \;=\; \frac{1}{(2n-1)!} \int_{x_0}^{x} (x-t)^{2n-1}\, f^{(2n)}(t)\, dt \quad \forall\, x \in (a,b).$$

Da $f^{(2n)}$ stetig ist, gibt es ein $\varepsilon > 0$ für welches $f^{(2n)}(x)$ auf $(x_0 - \varepsilon, x_0 + \varepsilon)$ stets das gleiche Vorzeichen hat wie $f^{(2n)}(x_0)$. Daraus folgt für $x > x_0$ sofort, dass auch $f(x) - f(x_0)$ gleiches Vorzeichen wie $f^{(2n)}(x_0)$ hat. Beachtet man, dass per Definition gilt

$$\int_{x}^{x_0} (x-t)^{2n-1}\, f^{(2n)}(t)\, dt \;=\; -\int_{x_0}^{x} (x-t)^{2n-1}\, f^{(2n)}(t)\, dt$$

so folgt dasselbe aber auch für $x < x_0$. $\square$

Definition 1.6.3 *Sei f auf $[a,b]$ wenigstens n-mal differenzierbar, und sei $x_0 \in [a,b]$. Dann heißt*

$$p_n(x) \;=\; \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}\, (x-x_0)^k$$

das n-te Taylorpolynom von f an der Stelle x_0. *Die Differenz $r_n = f - p_n$ heißt* das Taylorsche Restglied n-ter Ordnung.

Ist f sogar beliebig oft differenzierbar auf $[a,b]$, so heißt die Potenzreihe

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

die Taylorreihe von f im Punkt x_0.

Bemerkung 1.6.4 *Wir haben eine Reihe von Funktionen über Potenzreihen definiert. Es ist leicht zu sehen, dass diese Reihen gerade die Taylorreihen der entsprechenden Funktionen im Punkt $x_0 = 0$ sind.*

Per Definition konvergiert die Talyorreihe von f genau dann für ein x gegen $f(x)$, wenn gilt $r_n(x) \longrightarrow 0$ für $n \to \infty$. Dies muss aber durchaus nicht gelten: Es kann vorkommen, dass die Taylorreihe für alle $x \neq x_0$ divergiert, oder dass sie konvergiert, aber sozusagen gegen den falschen Wert*, soll heißen, nicht gegen $f(x)$.*

Der Satz von Taylor kann kurz so ausgedrückt werden:

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

Man nennt diese Gleichung auch die Darstellung des Restgliedes in Integralform*. Aus dem erweiterten Mittelwertsatz folgt, wenn man dort f durch $f^{(n+1)}$ und g durch $(x-t)^n/n!$ ersetzt und noch den Zwischenwertsatz benutzt, dass für einen geeigneten Wert ξ zwischen x und x_0 gilt*

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

Man nennt diese Gleichung auch die Darstellung des Restgliedes in differenzieller Form*.*

Behauptung 1.6.5 *Für alle $\alpha \in \mathbb{R}$ gilt*

$$(1+x)^\alpha \;=\; \sum_{k=0}^{\infty} \binom{\alpha}{k} x^k \qquad \forall\, x \in (-1,1)\,.$$

Die Reihe heißt auch die Binomialreihe*.*

Beweis: Die Behauptung ist klar wenn $\alpha \in \mathbb{N}_0$ ist. Im anderen Fall folgt, dass die Binomialreihe den Konvergenzradius $R = 1$ hat. Außerdem ist die Reihe gerade die Taylorreihe von $f(x) = (1+x)^\alpha$ (mit $x_0 = 0$), und das Restglied ist in diesem Fall gleich

$$r_n(x) \;=\; \binom{\alpha}{n} (\alpha - n) \int_0^x (x-t)^n \; (1+t)^{\alpha-n-1}\, dt.$$

Zu zeigen ist also, dass das Restglied für $|x| < 1$ gegen 0 geht wenn $n \to \infty$. Dies folgt allerdings für x nahe bei -1 nicht direkt aus den üblichen Abschätzungen. Deshalb gehen wir anders vor:

Nach einer der unten stehenden Aufgaben folgt, dass es genau eine auf $(-1,1)$ stetig differenzierbare Funktion $y(x)$ gibt, die die Bedingungen

$$y'(x) \;=\; \frac{\alpha}{(1+x)}\, y(x) \quad \forall\, x \in (-1,1)\,, \quad y(0) \;=\; 1$$

erfüllt. Offensichtlich sind diese Gleichungen richtig für $y(x) = (1+x)^\alpha$. Man rechnet aber leicht nach, dass auch die durch die Binomialreihe definierte Funktion eine Lösung dieses Anfangswertproblems ist, und daher muss die Behauptung gelten. □

Aufgabe 1.6.6 *Zeige durch Induktion über n: Ist f auf einem Intervall I n-mal differenzierbar, und ist die n-te Ableitung von f die Nullfunktion, so ist f ein Polynom vom Grade höchstens gleich $n-1$.*

Aufgabe 1.6.7 *Sei f mindestens $(2n+1)$-mal stetig differenzierbar auf einem offenen Intervall (a,b), für ein $n \in \mathbb{N}$, und gelte für ein $x_0 \in (a,b)$:*

$$f'(x_0) = f''(x_0) = \ldots = f^{(2n)}(x_0) = 0, \quad f^{(2n+1)}(x_0) \neq 0.$$

Zeige: Dann hat f im Punkt x_0 kein lokales Extremum.

Aufgabe 1.6.8 *Finde die Taylorreihe von $f(x) = \log x$ im Punkt $x_0 = 1$ und untersuche ihre Konvergenz.*

Aufgabe 1.6.9 (Lineare homogene Differenzialgleichung erster Ordnung) *Sei $I \subset \mathbb{R}$ ein Intervall, sei $f : I \longrightarrow \mathbb{R}$ stetig, und sei $F(x) = \int_{x_0}^{x} f(t)\,dt$, für ein $x_0 \in I$. Sei schließlich $y_0 \in \mathbb{R}$.*

1. Zeige: Die Funktion $y(x) = y_0\,\mathrm{e}^{F(x)}$ erfüllt die beiden Bedingungen
$$y'(x) \;=\; f(x)\,y(x)\,, \quad \forall\, x \in I, \quad y(x_0) \;=\; y_0.$$
Man nennt die erste der beiden Gleichungen eine *lineare homogene Differenzialgleichung erster Ordnung* (und y eine Lösung derselben), während die zweite auch *Anfangsbedingung* genannt wird. Beide zusammen stellen *ein Anfangswertproblem* dar.

2. Zeige weiter: Ist y irgend eine Funktion auf I, die beiden Bedingungen genügt, so ist $y(x)\,\mathrm{e}^{-F(x)}$ konstant, und durch Einsetzen von $x = x_0$ folgt dann $y(x) = y_0\,\mathrm{e}^{F(x)}$.

Schließe hieraus dass das obige Anfangswertproblem genau eine Lösung y besitzt.

Aufgabe 1.6.10 (Differenzialgleichung mit getrennten Veränderlichen) *Seien $I_1, I_2 \subset \mathbb{R}$ zwei Intervalle, seien $f : I_1 \longrightarrow \mathbb{R}$ und $g : I_2 \longrightarrow \mathbb{R}$ stetig, und sei $g(x) > 0$ auf I_2. Seien weiter F und G Stammfunktionen zu f bzw. $1/g$ (also G streng monoton wachsend auf I_2). Seien schließlich $x_0 \in I_1$, $y_0 \in I_2$.*

1. Zeige: Ist $I \subset I_1$ ein Intervall, und ist $y : I \longrightarrow I_2$ stetig differenzierbar, mit
$$y'(x) \;=\; f(x)\,g(y(x)) \quad \forall\, x \in I, \qquad y(x_0) \;=\; y_0, \tag{1.6.1}$$
so folgt $F(x) - F(x_0) = \int_{x_0}^{x} f(t)\,dt = \int_{x_0}^{x} y'(t)/g(y(t))\,dt = G(y(x)) - G(y_0)$ für $x \in I$. Da G injektiv ist, kann diese Gleichung nach $y(x)$ aufgelöst werden, und wir erhalten
$$y(x) \;=\; G^{-1}\left(G(y_0) + F(x) - F(x_0)\right) \qquad \forall x \in I. \tag{1.6.2}$$

2. Zeige: Ist $I \subset I_1$ ein Intervall und so, dass für $x \in I$ immer gilt $G(y_0) + F(x) - F(x_0) \in G(I_2)$, und definiert man y durch (1.6.2), so erfüllt y die Bedingungen (1.6.1).

Schließe hieraus dass das Anfangswertproblem (1.6.1) *auf dem in* 2. *angegebenen Intervall genau eine Lösung y besitzt.*

1.7 Uneigentliche Integrale

Im Folgenden betrachten wir ein Intervall der Form $[a, b)$ mit $a \in \mathbb{R}$ und $b \in \overline{\mathbb{R}}$, $a < b$.

Definition 1.7.1 *Sei $f : [a, b) \longrightarrow \mathbb{R}$ über jedes Intervall der Form $[a, c]$, mit $a \leq c < b$, integrierbar, und existiere $\lim_{c \to b-} \int_a^c f(x)\,dx$. Dann heißt f* über $[a, b)$ uneigentlich integrierbar*, und wir nennen*
$$\int_a^b f(x)\,dx \;=\; \lim_{c \to b-} \int_a^c f(x)\,dx$$
das uneigentliche Integral von f über $[a, b)$. *Wir sagen manchmal:* Das uneigentliche Integral ist konvergent*, falls f uneigentlich integrierbar ist. Falls $|f|$ über $[a, b)$ uneigentlich integrierbar ist, sagen wir auch:*

Das uneigentliche Integral von f ist absolut konvergent. *Beachte: Falls $b < \infty$ ist, und falls f auf $[a, b]$ integrierbar ist, gilt für alle $c \in [a, b)$:*

$$\int_a^b f(x)\,dx \;=\; \int_a^c f(x)\,dx \;+\; \int_c^b f(x)\,dx.$$

Die Fundamentalabschätzung impliziert mit $M = \sup |f(x)|$, dass

$$\left|\int_c^b f(x)\,dx\right| \;\le M\,(b-c) \to 0 \quad (c \to b-).$$

Deshalb ist f auch uneigentlich integrierbar über $[a, b)$, und das uneigentliche Integral ist gleich dem Integral. In analoger Weise kann man auch uneigentliche Integrale über Intervalle der Form $(a, b]$ definieren, wobei a auch gleich $-\infty$ sein kann.

Beispiel 1.7.2 *Sei $\alpha > 1$. Aus $\int_1^a x^{-\alpha}dx = (\alpha-1)^{-1}\,(1 - a^{1-\alpha}) \to (\alpha-1)^{-1}$ für $a \to \infty$ folgt die Konvergenz des uneigentlichen Integrals $\int_1^\infty x^{-\alpha}dx$, und*

$$\int_1^\infty \frac{dx}{x^\alpha} \;=\; \frac{1}{\alpha-1}.$$

Beispiel 1.7.3 *Sei $0 < \alpha < 1$. Aus $\int_\varepsilon^1 x^{-\alpha}dx = (1-\alpha)^{-1}\,(1 - \varepsilon^{1-\alpha}) \to (1-\alpha)^{-1}$ für $\varepsilon \to 0+$ folgt die Konvergenz des uneigentlichen Integrals $\int_0^1 x^{-\alpha}dx$, und*

$$\int_0^1 \frac{dx}{x^\alpha} \;=\; \frac{1}{1-\alpha}.$$

Satz 1.7.4 (Cauchy-Kriterium für uneigentliche Integrale)
Das uneigentliche Integral $\int_a^b f(x)\,dx$ ist genau dann konvergent, wenn gilt

$$\forall\ \varepsilon > 0\ \exists\ c \in [a, b)\ \forall\ \xi \in [c, b),\ \forall\ \eta \in [\xi, b): \quad \left|\int_\xi^\eta f(x)\,dx\right| < \varepsilon. \tag{1.7.1}$$

(Ohne Beweis)

Korollar zu Satz 1.7.4 *Aus der absoluten Konvergenz des uneigentlichen Integrals $\int_a^b f(x)\,dx$ folgt die Konvergenz.*

Beweis: Folgt aus $|\int_\xi^\eta f(x)\,dx| \le \int_\xi^\eta |f(x)|\,dx$ und dem Cauchy-Kriterium. □

Beispiel 1.7.5 *Sei $\alpha > 0$, und gelte $1 \le \xi < \eta$. Dann ist*

$$\int_\xi^\eta \frac{\sin x}{x^\alpha}\,dx \;=\; -\left.\frac{\cos x}{x^\alpha}\right|_\xi^\eta \;-\; \alpha \int_\xi^\eta \frac{\cos x}{x^{\alpha+1}}\,dx\ .$$

Daraus folgt

$$\left|\int_\xi^\eta \frac{\sin x}{x^\alpha}\,dx\right| \;\le \frac{1}{\xi^\alpha} \;+\; \frac{1}{\eta^\alpha} \;+\; \alpha \int_\xi^\eta \frac{1}{x^{\alpha+1}}\,dx\ .$$

Mit Beispiel 1.7.2 *und dem Cauchy-Kriterium folgt hieraus die Konvergenz des uneigentlichen Integrals $\int_1^\infty x^{-\alpha}\,\sin x\,dx$.*

Satz 1.7.6 (Majorantenkriterium für uneigentliche Integrale)
Gilt $|f(x)| \leq g(x)$ *für* $x \in [a, b)$, *und ist das uneigentliche Integral* $\int_a^b g(x)\,dx$ *konvergent, so ist* $\int_a^b f(x)\,dx$ *absolut konvergent.*

Beweis: Folgt aus dem Cauchy-Kriterium, denn $\int_\xi^\eta |f(x)|\,dx \leq \int_\xi^\eta g(x)\,dx$. □

Satz 1.7.7 (Integralkriterium)
Sei $p \in \mathbb{Z}$, *und sei* f *auf* $[p, \infty)$ *positiv und monoton fallend. Dann gilt: Genau dann ist das uneigentliche Integral* $\int_p^\infty f(x)\,dx$ *konvergent, wenn die Reihe* $\sum_{n=p}^\infty f(n)$ *konvergiert, und*

$$\sum_{n=p+1}^{\infty} f(n) \;\leq\; \int_p^\infty f(x)\,dx \;\leq\; \sum_{n=p}^{\infty} f(n) \; .$$

Beweis: Aus der Monotonie von f folgt die Ungleichung $f(n+1) \leq f(x) \leq f(n)$ für alle $x \in [n, n+1]$, und hieraus folgt $f(n+1) \leq \int_n^{n+1} f(x)\,dx \leq f(n)$, was die Behauptung ergibt. □

Aufgabe 1.7.8 *Zeige, dass das uneigentliche Integral* $\int_1^\infty x^{-\alpha} dx$ *für* $0 \leq \alpha \leq 1$ *nicht konvergiert.*

Aufgabe 1.7.9 *Zeige, dass das uneigentliche Integral* $\int_0^1 x^{-\alpha} dx$ *für* $\alpha \geq 1$ *nicht konvergiert.*

Aufgabe 1.7.10 *Zeige, dass das uneigentliche Integral* $\int_1^\infty x^{-\alpha} \sin x\,dx$ *für* $0 \leq \alpha \leq 1$ *nicht absolut konvergiert.*

1.8 Die Gamma-Funktion

Definition 1.8.1 *Sei* $f : [a, b] \longrightarrow \mathbb{C}$. *Dann ist* $f(x) = u(x) + i\,v(x)$ *für alle* $x \in [a, b]$. *Wir nennen* f über $[a, b]$ integrierbar, *wenn* u *und* v *über* $[a, b]$ *integrierbar sind, und setzen*

$$\int_a^b f(x)\,dx = \int_a^b u(x)\,dx \;+\; i \int_a^b v(x)\,dx.$$

Beispiel 1.8.2 *Für* $t \in \mathbb{R}_+$ *und* $z = x + i\,y \in \mathbb{C}$ *ist*

$$t^z \;=\; \mathrm{e}^{(x+i\,y)\log t} \;=\; t^x\,(\cos(y \log t) \;+\; i\,\sin(y \log t)) \; .$$

Also ist per Definition für $0 < \varepsilon < a$:

$$\int_\varepsilon^a t^{z-1}\,\mathrm{e}^{-t}\,dt \;=\; \int_\varepsilon^a t^{x-1}\,\mathrm{e}^{-t}\,\cos(y \log t)\,dt \;+\; i \int_\varepsilon^a t^{x-1}\,\mathrm{e}^{-t}\,\sin(y \log t)\,dt \; .$$

Beide Integrale auf der rechten Seite haben für $a \to \infty$ *immer einen Grenzwert. Dagegen muss* $x > 0$ *sein, damit auch für* $\varepsilon \to 0$ *ein Grenzwert existiert. Wir sagen deshalb:* Das uneigentliche Integral $\int_0^\infty t^{z-1}\,\mathrm{e}^{-t}\,dt$ konvergiert für $x = \mathrm{Re}\ z > 0$.

Definition 1.8.3 *Für* $\operatorname{Re} z > 0$ *heißt die Abbildung* $z \mapsto \Gamma(z)$ *mit*

$$\Gamma(z) \;=\; \int_0^\infty t^{z-1}\,\mathrm{e}^{-t}\,dt$$

die Gamma-Funktion.

Proposition 1.8.4 *Für die Gamma-Funktion gilt die Darstellung*

$$\Gamma(z) \;=\; \int_1^\infty t^{z-1}\,\mathrm{e}^{-t}\,dt \;+\; \sum_{n=0}^{\infty} \frac{(-1)^n}{n!\,(n+z)}\,, \qquad \operatorname{Re} z \;>\; 0\,.$$

Das uneigentliche Integral konvergiert für alle $z \in \mathbb{C}$. *Die unendliche Reihe konvergiert absolut in der Menge* $G = \mathbb{C} \setminus \{0,\ -1,\ -2,\ \ldots\}$. *Wir sagen deshalb, dass sich die Gamma-Funktion in die Menge* G *fortsetzen läßt.*

Beweis: Aus der Exponentialreihe folgt die Entwicklung

$$t^{z-1}\,\mathrm{e}^{-t} \;=\; \sum_{n=0}^{\infty} \frac{(-1)^n\,t^{n+z-1}}{n!}\,,$$

wobei die Reihe auf jedem Intervall der Form $[\varepsilon, 1]$, mit $\varepsilon > 0$, gleichmäßig konvergiert. Deshalb gilt wegen des Satzes über gliedweise Integration

$$\int_\varepsilon^1 t^{z-1}\,\mathrm{e}^{-t}\,dt \;=\; \sum_{n=0}^{\infty} \frac{(-1)^n\,t^{n+z}}{n!\,(n+z)}\Bigg|_\varepsilon^1\,,$$

und für $\varepsilon \to 0$ folgt die Behauptung. □

Aufgabe 1.8.5 *Zeige durch partielle Integration* $\Gamma(z+1) = z\,\Gamma(z)$ *für alle* $z \in \mathbb{C} \setminus \{0,\ -1,\ -2,\ \ldots\}$.

Aufgabe 1.8.6 *Berechne* $\Gamma(1)$, *und zeige mit Hilfe der vorigen Aufgabe, dass* $\Gamma(n+1) = n!$ *ist, für alle* $n \in \mathbb{N}_0$.

Satz 1.8.7 (Stirlingsche Formel) *Für reelle Werte von* x *gilt*

$$\lim_{x\to\infty} \frac{\Gamma(x)\ \mathrm{e}^x\ \sqrt{x}}{x^x} \;=\; \sqrt{2\,\pi}\,.$$

(Ohne Beweis)

1.9 Numerische Integration

Unter *numerischer Integration* versteht man die zahlenmäßige Berechnung eines Integrals $\int_a^b f(x)\,dx$, meist mit Hilfe eines Rechners. Die Idee besteht dabei in der Regel darin, das große Intervall $[a,b]$ in kleinere Intervalle zu zerlegen, und auf diesen kleinen Intervallen die Funktion f durch eine einfach zu integrierende Funktion, meist ein Polynom kleinen Grades, zu ersetzen. Genauer: Ist $\mathcal{Z} = \{x_0, \ldots, x_N\}$ eine Zerlegung von $[a,b]$, so wählt man für jedes $k = 1, \ldots, N$ eine Funktion p_k und berechnet statt $\int_a^b f(x)\,dx = \sum_{k=1}^N \int_{x_{k-1}}^{x_k} f(x)\,dx$ die Summe $S = \sum_{k=1}^N \int_{x_{k-1}}^{x_k} p_k(x)\,dx$. Die Frage ist, wie man die Zerlegung $\mathcal{Z}$ und die Funktionen p_k am günstigsten wählt, damit die berechnete Zahl mit dem richtigen Wert des Integrals möglichst genau übereinstimmt! Meistens wählt man die Punkte der Zerlegung *äquidistant*, also

$$x_k = a + k\,h\,, \quad k = 0, \ldots, N\,, \qquad h = (b - a)/N\,.$$

Für die Wahl der Funktionen p_k diskutieren wir die drei folgenden Beispiele:

1. p_k sei ein konstantes Polynom, etwa $p_k(x) = f(\xi_k)$, mit einem Zwischenpunkt $\xi_k \in [x_{k-1}, x_k]$. In diesem Fall berechnet man an Stelle des Integrals gerade eine Riemann-Summe.

2. **(Trapezregel)** p_k sei ein lineares Polynom (eine Gerade), welche in den Punkten x_{k-1} und x_k mit f übereinstimmt; also

$$p_k(x) = f(x_{k-1}) + (x - x_{k-1})\,\frac{f(x_k) - f(x_{k-1})}{x_k - x_{k-1}}\,.$$

   Dies ergibt bei äquidistanter Zerlegung den Wert

$$S = \frac{h}{2}\,\big(f(x_0) + 2\,f(x_1) + \ldots + 2\,f(x_{N-1}) + f(x_N)\big)\,.$$

   Man sagt: die beiden Randpunkte $x_0 = a$ und $x_N = b$ haben bei der Trapezregel das Gewicht 1, alle übrigen dagegen das Gewicht 2. Den Vorfaktor $h/2$ kann man sich so erklären, dass man für konstante Funktionen den richtigen Wert erhält.

3. **(Simpson-Regel, oder Keplersche Fassregel)** Hier ist p_k ein quadratisches Polynom, welches in den Punkten x_{k-1} und x_k sowie im Mittelpunkt $(x_k + x_{k-1})/2$ mit f übereinstimmt. Also stellt man sich am besten eine äquidistante Zerlegung mit einer geraden Anzahl von Teilpunkten vor. Wählt man die neuen Bezeichnungen

$$x_k = a + k\,h\,, \quad k = 0, \ldots, 2\,N\,, \qquad h = \frac{b - a}{2\,N}$$

   ergibt dies die Formel

$$S = \frac{h}{3}\,\big(f(x_0) + 4\,f(x_1) + 2\,f(x_2) + 4\,f(x_3) + \ldots \\ + 2\,f(x_{2\,N-2}) + 4\,f(x_{2\,N-1}) + + f(x_{2\,N})\big)\,.$$

   Hier haben also die Randpunkte wieder Gewicht 1, die übrigen abwechselnd Gewichte 4 bzw. 2.

Man kann die folgenden *Fehlerabschätzungen* zeigen:

Für die Trapezregel:

$$\Big|\int_a^b f(x)\,dx - S\Big| \le \frac{b-a}{12}\,h^2 \max_{a \le x \le b} \big|f''(x)\big|\,,$$

für die Simpson-Regel:

$$\Big|\int_a^b f(x)\,dx - S\Big| \le \frac{b-a}{180}\,h^4 \max_{a \le x \le b} \big|f''''(x)\big|\,.$$

Aufgabe 1.9.1 *Zeige: Die Simpson-Regel liefert für Polynome vom Grad ≤ 3 den exakten Integralwert.*

Kapitel 2

Ungleichungen und Transformationen

2.1 Ungleichungen

Die folgenden Ungleichungen spielen in der Mathematik eine zentrale Rolle:

Aufgabe 2.1.1 *Seien $p, q > 1$ mit $1/p + 1/q = 1$, und seien $a, b \in \mathbb{R}_+$. Zeige: Es ist stets*

$$a^{1/p}\, b^{1/q} \;\le\; \frac{a}{p} + \frac{b}{q}\,,$$

und Gleichheit gilt genau dann, wenn $a = b$ ist. Anleitung: *Finde das Maximum der Funktion $f(x) = x^{1/p}\, b^{1/q}\, (x/p + b/q)^{-1}$, für $x \ge 0$.*

Lemma 2.1.2 (Höldersche Ungleichung) *Für $p, q > 1$ mit $1/p + 1/q = 1$ sowie $a = (a_1, \ldots, a_n)^T, b = (b_1, \ldots, b_n)^T \in \mathbb{K}^n$ gilt*

$$\sum_{k=1}^n |a_k\, b_k| \;\le\; \Big(\sum_{k=1}^n |a_k|^p \Big)^{1/p} \Big(\sum_{k=1}^n |b_k|^q \Big)^{1/q}\,.$$

Das Gleichheitszeichen gilt genau dann, wenn $\lambda, \mu \in \mathbb{R}$ existieren, welche nicht beide gleich 0 sind, mit $\lambda\, |a_k|^p = \mu\, |b_k|^q$ für alle k.

Beweis: O. B. d. A. seien $A = \sum_{k=1}^n |a_k|^p > 0$ und $B = \sum_{k=1}^n |b_k|^q > 0$. Aus Aufgabe 2.1.1 folgt dann

$$\frac{|a_k|}{A^{1/p}}\, \frac{|b_k|}{B^{1/q}} \;\le\; \frac{|a_k|^p}{p\,A} + \frac{|b_k|^q}{q\,B} \qquad \forall\; k = 1, \ldots, n,$$

mit Gleichheit genau für $|a_k|^p/A = |b_k|^q/B$. Durch Summation über k folgt die Behauptung. □

Aufgabe 2.1.3 *Zeige mit Hilfe der Hölderschen Ungleichung für $a_1, \ldots, a_n \in \mathbb{C}$ und $p, q > 1$ mit $1/p + 1/q = 1$*

$$\sum_{k=1}^n |a_k| \;\le n^{1/q} \left(\sum_{k=1}^n |a_k|^p \right)^{1/p}$$

Aufgabe 2.1.4 (Höldersche Ungleichung für Integrale) *Sei $a < b$, und seien $f, g : [a,b] \longrightarrow \mathbb{K}$ integrierbar. Sei $\mathcal{Z} = \{x_0, \ldots, x_N\}$ eine Zerlegung von $[a,b]$, und sei $\xi = (\xi_1, \ldots, \xi_n)^T$ ein zugehöriger Zwischenpunktvektor. Zeige mit Hilfe der Hölderschen Ungleichung*

$$\sum_{k=1}^{N} |f(\xi_k)\, g(\xi_k)|\,(x_k - x_{k-1}) \quad \leq$$
$$\left(\sum_{k=1}^{N} |f(\xi_k)|^p (x_k - x_{k-1})\right)^{1/p} \left(\sum_{k=1}^{N} |g(\xi_k)|^q (x_k - x_{k-1})\right)^{1/q}$$

und schließe hieraus mit der Definition des Integrals auf die Gültigkeit der Ungleichung

$$\int_a^b |f(x)g(x)|\,dx \leq \left(\int_a^b |f(x)|^p\,dx\right)^{1/p} \left(\int_a^b |g(x)|^q\,dx\right)^{1/q}$$

mit $p, q > 1$ und $1/p + 1/q = 1$.

Lemma 2.1.5 (Minkowskische Ungleichung) *Für $a = (a_1, \ldots, a_n)^T, b = (b_1, \ldots, b_n)^T \in \mathbb{K}^n$ und $p \geq 1$ gilt*

$$\left(\sum_{k=1}^{n} |a_k + b_k|^p \right)^{1/p} \quad \leq \quad \left(\sum_{k=1}^{n} |a_k|^p \right)^{1/p} \quad + \quad \left(\sum_{k=1}^{n} |b_k|^p \right)^{1/p} .$$

Beweis: Für $p = 1$ ist die Behauptung klar wegen der Dreiecksungleichung. Sei jetzt $p > 1$, und sei o. B. d. A. angenommen, dass $A = \sum_{k=1}^{n} |a_k + b_k|^p > 0$. Dann folgt mit der Hölderschen Ungleichung und $1/q = 1 - 1/p$, also $(p-1)\,q = p$:

$$\begin{aligned}\sum_{k=1}^{n} |a_k|\,|a_k + b_k|^{p-1} \quad &\leq \quad \left(\sum_{k=1}^{n} |a_k|^p \right)^{1/p} \left(\sum_{k=1}^{n} |a_k + b_k|^{(p-1)q} \right)^{1/q} \\ &= \quad A^{1/q} \left(\sum_{k=1}^{n} |a_k|^p \right)^{1/p} ,\end{aligned}$$

und dasselbe gilt, wenn wir die a_k und b_k vertauschen. Daraus folgt wegen

$$A \;\leq \sum_{k=1}^{n} |a_k|\,|a_k + b_k|^{p-1} \;+\; \sum_{k=1}^{n} |b_k|\,|a_k + b_k|^{p-1}$$

die Ungleichung

$$A \;\leq A^{1/q} \left(\left(\sum_{k=1}^{n} |a_k|^p \right)^{1/p} \;+\; \left(\sum_{k=1}^{n} |b_k|^p \right)^{1/p} \right) ,$$

und Division durch $A^{1/q}$ ergibt die Behauptung. □

Aufgabe 2.1.6 (Minkowskische Ungleichung für Integrale)
Zeige für $p \geq 1$ und f, g so, dass die rechtsstehenden Integrale existieren:

$$\left(\int_a^b |f(x) + g(x)|^p dx\right)^{1/p} \leq \left(\int_a^b |f(x)|^p dx\right)^{1/p} + \left(\int_a^b |g(x)|^p dx\right)^{1/p} .$$

2.2 Äquivalenzrelationen

Definition 2.2.1 *Sei X eine beliebige nicht-leere Menge. Eine Relation auf X, also eine Teilmenge $R \subset X \times X$, heißt* eine Äquivalenzrelation *auf X, falls für beliebige $x, x_1, x_2, x_3 \in X$ gilt:*

(R) $(x, x) \in R$ (Reflexivität).

(S) $(x_1, x_2) \in R \implies (x_2, x_1) \in R$ (Symmetrie).

(T) $(x_1, x_2) \in R$ und $(x_2, x_3) \in R \implies (x_1, x_3) \in R$ (Transitivität).

Statt $(x_1, x_2) \in R$ schreiben wir auch $x_1 \sim x_2$ und sagen in Worten: x_1 ist äquivalent zu x_2.

Für $x \in X$ sei $A_x = \{\tilde{x} \in X : \quad x \sim \tilde{x}\ \}$. Wir nennen ein solches A_x eine Äquivalenzklasse. *Ein beliebiges Element einer Äquivalenzklasse heißt auch* ein Repräsentant *dieser Äquivalenzklasse.*

Proposition 2.2.2 *Sei X eine beliebige nicht-leere Menge mit einer Äquivalenzrelation auf X. Seien A_{x_1} und A_{x_2} zwei Äquivalenzklassen mit $A_{x_1} \cap A_{x_2} \neq \emptyset$. Dann gilt bereits $A_{x_1} = A_{x_2}$.*

Beweis: Sei $x \in A_{x_1} \cap A_{x_2}$. Dann ist $x \sim x_1$ und $x \sim x_2$, und aus der Symmetrie und Transitivität folgt dann $x_1 \sim x_2$, also $x_2 \in A_{x_1}$. Sei jetzt $x \in A_{x_2}$, also $x_2 \sim x$. Dann folgt aber aus der Transitivität, dass $x_1 \sim x$, also $x \in A_{x_1}$ ist. Somit ist $A_{x_2} \subset A_{x_1}$. Die Umkehrung gilt aber ebenso. □

Korollar zu Proposition 2.2.2 (Zerlegung in Äquivalenzklassen) *Sei X eine beliebige nicht-leere Menge mit einer Äquivalenzrelation auf X. Dann bilden die Äquivalenzklassen eine Zerlegung von X, d. h., X ist die Vereinigung aller Äquivalenzklassen, und je zwei verschiedene Äquivalenzklassen sind disjunkt.*

Beweis: Da jedes $x \in X$ in der Äquivalenzklasse A_x liegt, folgt $X = \cup_{x \in X} A_x$. Dass verschiedene Äquivalenzklassen disjunkt sind, ist äquivalent zur Proposition. □

Aufgabe 2.2.3 *Sei $n \in \mathbb{N}$. Zeige: Auf $\mathbb{Z}$ wird durch*

$$p \sim q \iff \exists\, m \in \mathbb{Z} : \quad p - q = n\, m$$

eine Äquivalenzrelation definiert. Bestimme die Anzahl der Äquivalenzklassen!

Aufgabe 2.2.4 *Seien $a < b$, und sei $p \geq 1$. Zeige: Auf der Menge $R[a, b]$ aller über $[a, b]$ Riemann-integrierbaren Funktionen wird durch*

$$f \sim g \iff \int_a^b |f(x) - g(x)|^p \, dx = 0$$

eine Äquivalenzrelation definiert. Zeige weiter, dass die Anzahl der Äquivalenzklassen unendlich ist, und dass jede Äquivalenzklasse überabzählbar ist.

Aufgabe 2.2.5 *Zeige mit den Bezeichnungen und Voraussetzungen der vorigen Aufgabe:*

$$f_1 \sim g_1\,, \quad f_2 \sim g_2 \implies f_1 + f_2 \sim g_1 + g_2\,, \quad f_1\, f_2 \sim g_1\, g_2\,.$$

Schließe hieraus, dass es sinnvoll ist, von der Summe und dem Produkt zweier Äquivalenzklassen zu sprechen.

2.3 Periodische Funktionen

Definition 2.3.1 *Eine Funktion $f : \mathbb{R} \longrightarrow \mathbb{C}$ heißt* periodisch, *falls es ein $T > 0$ gibt mit*

$$f(x+T) \;=\; f(x) \qquad \forall\, x \in \mathbb{R}\,.$$

Jedes solche T heißt dann auch eine Periode *von f, und wir sagen auch, dass f dann T-periodisch ist. Beachte, dass mit T auch $n\,T$ für $n \in \mathbb{N}$ eine Periode von f ist, und dass für konstante Funktionen* jedes *$T > 0$ eine Periode ist.*

Beispiel 2.3.2 *Für $\omega > 0$ sind die Funktionen*

$$\sin(n\,\omega\,x)\,, \qquad \cos(n\,\omega\,x) \qquad n \in \mathbb{N}$$

alle periodisch mit Periode $T = 2\,\pi/\omega$.

Die Dirichletsche Sprungfunktion

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

ist periodisch, und jedes rationale $T > 0$ ist eine Periode, aber irrationale T sind keine Perioden. Insbesondere gibt es keine kleinste *Periode.*

Aufgabe 2.3.3 *Zeige: Ist f auf einem halboffenen Intervall der Länge $T > 0$ definiert, so gibt es eine eindeutig bestimmte Fortsetzung von f, welche T-periodisch ist. Falls f etwa auf $[0,T)$ stetig ist, ist dann die T-periodische Fortsetzung auf $\mathbb{R}$ stetig?*

2.4 Die Laplace-Transformation

Definition 2.4.1 *Sei $f : [0,\infty) \longrightarrow \mathbb{C}$ stetig. Wir setzen*

$$g(z) \;=\; (\mathcal{L}f)(z) \;=\; \int_0^\infty f(t)\,\mathrm{e}^{-zt}\,dt$$

Man kann zeigen: Wenn es überhaupt ein $z \in \mathbb{C}$ gibt, für welches dieses uneigentliche Integral konvergiert, dann konvergiert es in einer Halbebene, d. h. für alle z mit $\mathrm{Re}\ z > c$, für ein $c \in \mathbb{R}$. Wir nennen dann die dort definierte Funktion g die Laplace-Transformierte *der Funktion f. In der Sprache der linearen Algebra kann man sagen, dass die Abbildung $f \longmapsto \mathcal{L}f$ eine lineare Abbildung eines geeigneten Vektorraums von Funktionen in einen anderen ist. Diese Abbildung wird auch* Laplace-Transformation *genannt.*

Aufgabe 2.4.2 *Zeige: Für $f(t) = t^{\alpha-1}\,\mathrm{e}^{\beta\,t}$ ist die Laplace-Transformierte gleich $g(z) = \Gamma(\alpha)\,(z-\beta)^{-\alpha}$; dabei muss $\mathrm{Re}\ \alpha > 0$ sein, damit das Integral überhaupt konvergiert, und dann konvergiert es für alle z mit $\mathrm{Re}\ z > \mathrm{Re}\ \beta$.*

Aufgabe 2.4.3 *Finde die Laplace-Transformierte von $\cos t$ bzw. $\sin t$.*

Bemerkung 2.4.4 *Die Laplace-Transformation hat folgende Eigenschaften, die hier nicht bewiesen werden sollen. Dabei sollen die auftretenden Funktionen immer so sein, dass ihre Laplace-Transformation für genügend große Werte von $\mathrm{Re}\ z$ existiert.*

1. $(\mathcal{L}f)(z) \longrightarrow 0$ wenn $\operatorname{Re} z \to \infty$.

2. $$(\mathcal{L}f^{(n)})(z) \;=\; z^n\,(\mathcal{L}f)(z) \;-\; \sum_{j=1}^{n} z^{n-j}\,f^{(j-1)}(0) \qquad \forall\, n \in \mathbb{N}.$$

3. $$\frac{d^n}{dz^n}(\mathcal{L}f)(z) \;=\; (-1)^n\,(\mathcal{L}\,t^n f(t))(z) \qquad \forall\, n \in \mathbb{N}.$$

4. $(\mathcal{L}\,f(t-\alpha))(z) \;=\; \mathrm{e}^{-\alpha\,z}\,(\mathcal{L}\,f)(z) \qquad \forall\,\alpha > 0,$ falls wir $f(t) = 0$ für $t < 0$ setzen.

5. $$(\mathcal{L}\int_0^t f(\tau)\,d\tau)(z) \;=\; \frac{1}{z}\,(\mathcal{L}\,f)(z)\,.$$

Satz 2.4.5 (Faltungssatz) *Für $h(t) = \int_0^t f(t-\tau)\,g(\tau)\,d\tau$ gilt*

$$\mathcal{L}\,h \;=\; (\mathcal{L}\,f)\;(\mathcal{L}\,g)\,.$$

(Ohne Beweis)

Aufgabe 2.4.6 (Schwingungsgleichung) *Benutze die Laplace-Transformation zur Lösung der Differenzialgleichung*

$$m\,y'' \;-r\,y' \;+\; k\,y \;=f(t)\,,$$

mit positiven reellen Zahlen m, r, k, einer auf $[0,\infty)$ stetigen Funktion f, sowie Anfangsbdingungen $y(0) = y_0$, $y'(0) = y_1$.

2.5 Fourier- und Z-Transformation

Definition 2.5.1 *Sei $f : (-\infty,\infty) \longrightarrow \mathbb{C}$ stetig. Wir setzen*

$$g(\xi) \;=\; (\mathcal{F}f)(\xi) \;=\; \int_{-\infty}^{\infty} f(x)\,\mathrm{e}^{i\,\xi\,x}\,dx\,,$$

falls das Integral für gewisse $\xi \in \mathbb{R}$ existiert. Wir nennen dann die dort definierte Funktion g die Fourier-Transformierte *der Funktion f. Die Abbildung $f \longmapsto \mathcal{F}f$ ist offenbar wieder linear und wird auch* Fourier-Transformation *genannt.*

Die Fourier-Transformation hängt eng mit der Laplace-Transformation zusammen und hat auch ähnliche Eigenschaften, auf die wir hier nicht näher eingehen.

Definition 2.5.2 *Sei $(f_n)_{n=0}^{\infty}$ eine Folge reeller oder komplexer Zahlen, sodass gilt*

$$\frac{1}{R} \;=\; \limsup_{n\to\infty} \sqrt[n]{|f_n|} \;<\; \infty\,.$$

Dann konvergiert bekanntlich die Potenzreihe

$$F(z) \;=\; \big(\mathcal{Z}(f_n)\big)(z) \;=\; \sum_{n=0}^{\infty} f_n\,z^{-n}$$

für alle $z \in C$ mit $|z| > 1/R$ absolut. Die Funktion F heißt die Z-Transformierte *der Folge (f_n). Die entsprechende lineare Abbildung heißt* die Z-Transformation*.*

Bemerkung 2.5.3 *Für die Z-Transformation gelten folgende Regeln:*

1. **(Faltungssatz)** Für Folgen (f_n), (g_n) gilt
$$\mathcal{Z}\Big(\sum_{k=0}^{n} f_k\, g_{n-k}\Big) \;=\; \big(\mathcal{Z}(f_n)\big)\;\big(\mathcal{Z}(g_n)\big)\,.$$

2. **(Summationssatz)**
$$\mathcal{Z}\Big(\sum_{k=0}^{n-1} f_k\Big) \;=\; \frac{1}{z-1}\big(\mathcal{Z}(f_n)\big)$$

3. **(Differenzensatz)** Für $\Delta^0 f_n = f_n$,
$$\Delta f_n \;=\; \Delta^1 f_n \;=\; f_{n+1} - f_n\,, \qquad \Delta^{k+1} f_n \;=\; \Delta^k f_{n+1} - \Delta^k f_n$$
für alle $n \in \mathbb{N}_0$ und $k \in \mathbb{N}$ gilt
$$\mathcal{Z}\big(\Delta^k f_n\big) \;=\; (z-1)^k\; \mathcal{Z}(f_n) \qquad \forall\, k \in \mathbb{N}_0\,.$$

4. **(Verschiebungssatz)** Setzt man $f_n = 0$ für alle $n < 0$, so gilt
$$\mathcal{Z}(f_{n-k}) \;=\; z^{-k}\; \mathcal{Z}(f_n) \qquad \forall\, k \in \mathbb{N}_0\,.$$

5. **(Ähnlichkeitssatz)** Für beliebiges $\lambda \in \mathbb{C} \setminus \{0\}$ gilt
$$\mathcal{Z}(\lambda^n\, f_n)(z) \;=\; \mathcal{Z}(f_n)(z/\lambda)\,.$$

Aufgabe 2.5.4 *Seien $m \in \mathbb{N}$ und $a_0, \ldots, a_m \in \mathbb{C}$ mit $a_m \neq 0$ beliebig gegeben. Sind Werte $f_0, \ldots, f_{m-1}$ bekannt, so legt die Gleichung*
$$\sum_{k=0}^{m} a_k\, f_{n+k} \;=\; 0 \qquad \forall\, n \;\geq\; 0$$
eine Folge (f_n) eindeutig fest. Versuche, mit Hilfe der Z-Transformation ein Bildungsgesetz für diese Folge zu finden.

Kapitel 3

Topologie in normierten Räumen

3.1 Der n-dimensionale Raum

Die reellen, aber auch die komplexen Zahlen wurden bereits im letzten Semester eingeführt, und in dieser Vorlesung kann das Symbol $\mathbb{K}$ sowohl die Menge $\mathbb{R}$ aller reellen, aber auch die Menge $\mathbb{C}$ aller komplexen Zahlen bedeuten. Soweit es um die Operationen der Addition und Multiplikation geht, gelten für die beiden Mengen die gleichen Rechenregeln – allerdings kann man für komplexe Zahlen keine Ungleichungen betrachten und insbesondere nicht davon sprechen, welche von zwei komplexen Zahlen die größere ist.

Definition 3.1.1 *Sei n eine natürliche Zahl. Wir schreiben $\mathbb{K}^n$ für die Menge aller* Spaltenvektoren der Länge n *mit Elementen in $\mathbb{K}$, d. h. genauer, für die Menge aller*

$$x \;=\; \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \;=\; (x_1, \dots, x_n)^T\,,$$

mit Zahlen $x_j \in \mathbb{K}$. Dabei ist für x wie oben und einen zweiten Spaltenvektor

$$y \;=\; \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} \;=\; (y_1, \dots, y_n)^T$$

sowie eine Zahl $\lambda \in \mathbb{K}$ eine Addition *definiert durch*

$$x \,+\, y \;=\; \begin{pmatrix} x_1 + y_1 \\ x_2 + y_2 \\ \vdots \\ x_n + y_n \end{pmatrix}\,,$$

und außerdem definieren wir eine Multiplikation *von x mit einer Zahl $\lambda \in \mathbb{K}$ durch*

$$\lambda\, x \;=\; \begin{pmatrix} \lambda\, x_1 \\ \lambda\, x_2 \\ \vdots \\ \lambda\, x_n \end{pmatrix}\,.$$

Das Skalarprodukt *von x und y ist die Zahl*

$$\langle x, y\rangle \;=\; \overline{x}^T y \;:=\; \sum_{j=1}^{n} \overline{x}_j \, y_j \,,$$

wobei $\overline{x}_j$ die zu x_j konjugiert komplexe Zahl bezeichnet.

Die Menge $\mathbb{K}^n$ ist *ein Beispiel* für das, was man in der Mathematik einen *Vektorraum* nennt. Die genaue Definition dieses Begriffes wird im nächsten Semester gegeben – für den Moment ist nur wichtig, dass man zwei Elemente eines allgemeinen Vektorraumes addieren und jedes Element mit einer Zahl multiplizieren kann, und dass für diese *Operationen* dieselben Rechenregeln wie in der elementaren Vektorrechnung aus dem ersten Semester gelten. Allerdings ist zu beachten, dass in einem allgemeinen Vektorraum *kein Skalarprodukt* definiert zu sein braucht. In den folgenden Abschnitten kann man sich an Stelle eines solchen allgemeinen Vektorraumes aber immer die Menge $\mathbb{K}^n$ vorstellen. Besonders gut eignet sich die Menge $\mathbb{R}^2$, d. i. die Ebene, oder $\mathbb{R}^3$, also der dreidimensionale Raum, zur Veranschaulichung der in den nächsten Abschnitten betrachteten Begriffe.

3.2 Normierte Räume

Definition 3.2.1 *Sei X ein beliebiger Vektorraum über $\mathbb{K}$. Eine Abbildung*

$$\|\cdot\| : \quad X \longrightarrow \mathbb{R}\,, \qquad x \longmapsto \|x\|$$

heißt eine Norm *auf X, wenn folgendes gilt:*

(N1) $\forall\, x \in X : \qquad \|x\| \geq 0\,; \quad \|x\| = 0 \iff x = 0.$

(N2) $\forall\, x \in X,\ \forall\, \alpha \in \mathbb{K} : \quad \|\alpha\, x\| \;=\; |\alpha|\, \|x\|.$

(N3) $\forall\, x_1, x_2 \in X : \quad \|x_1 + x_2\| \;\leq\; \|x_1\| + \|x_2\|$ (Dreiecksungleichung).

Der Vektorraum X, zusammen mit einer Norm auf X, heißt dann ein normierter Raum. *Sind auf X zwei Normen $\|\cdot\|_1, \|\cdot\|_2$ gegeben, so nennen wir diese* äquivalent*, wenn es Konstanten $C, K \in \mathbb{R}_+$ gibt, für die*

$$\forall\, x \in X : \quad C\,\|x\|_1 \;\leq\; \|x\|_2 \;\leq\; K\,\|x\|_1\,.$$

Dies ist eine Äquivalenzrelation auf der Menge aller Normen auf X.

Aufgabe 3.2.2 (Dreiecksungleichung nach unten) *Zeige: In einem normierten Raum X gilt*

$$\Big|\, \|x_1\| \;-\; \|x_2\| \,\Big| \;\leq \|x_1 \;-\; x_2\| \qquad \forall\, x_1, x_2 \in X\,.$$

Beispiel 3.2.3 *Für $x = (x_1, \ldots, x_n)^T \in \mathbb{K}^n$ und $1 \leq p \leq \infty$ sei*

$$\|x\|_p \;=\; \begin{cases} \Big(|x_1|^p + \ldots + |x_n|^p\Big)^{1/p} & (1 \leq p < \infty), \\[2ex] \sup\{|x_1|, \ldots, |x_n|\} & (p = \infty). \end{cases}$$

Dadurch ist für jedes feste p eine Norm auf $\mathbb{K}^n$ definiert; für $p < \infty$ ist die Dreiecksungleichung äquivalent zur Minkowskischen Ungleichung. Wir nennen $\|\cdot\|_p$ die p-Norm auf $\mathbb{K}^n$*, und sprechen für $p = 2$*

auch von der Euklidischen Norm*. Beachte, dass die Euklidische Norm eines Vektors $x \in \mathbb{K}^n$ gerade die Quadratwurzel aus dem Skalarprodukt von x mit sich selber ist. Offenbar gilt $|x_k| \leq \|x\|_p$ für alle p und alle $k = 1, \dots, n$, und daraus ergeben sich folgende Abschätzungen:*

$$\|x\|_\infty \leq \|x\|_p \leq n^{1/p}\, \|x\|_\infty \qquad \forall\, p \in [1, \infty), \quad \forall x \in \mathbb{K}^n.$$

Also sind die Normen $\|\cdot\|_p$ und $\|\cdot\|_\infty$ äquivalent. Daraus ergibt sich aber leicht, dass zwei beliebige p-Normen auf $\mathbb{K}^n$ immer äquivalent sind.

Beispiel 3.2.4 *Sei $[a, b]$ ein abgeschlossenes Intervall, und sei $C[a, b]$ die Menge aller dort stetigen Funktionen mit Werten in $\mathbb{K}$. Diese Funktionenmenge ist ein Vektorraum über $\mathbb{K}$. Für $f \in C[a, b]$ sei*

$$\|f\|_p = \begin{cases} \Big(\displaystyle\int_a^b |f(x)|^p dx\Big)^{1/p} & (1 \leq p < \infty), \\ \sup\{|f(x)| : \ a \leq x \leq b\} & (p = \infty). \end{cases}$$

Dies sind Normen auf $C[a, b]$, für jedes solche p. Diese Normen sind allerdings nicht äquivalent zueinander!

Definition 3.2.5 *Die Norm $\|f\|_\infty = \sup\{|f(x)| : \ a \leq x \leq b\}$ heißt* die Supremumsnorm der Funktion f*. Manchmal nennen wir $\|x\|_\infty = \sup\{|x_1|, \dots, |x_n|\}$ auch* die Supremumsnorm des Vektors $x \in \mathbb{K}^n$.

Aufgabe 3.2.6 *Zeige, dass die oben definierte Äquivalenz von Normen tatsächlich die drei Eigenschaften einer Äquivalenzrelation besitzt.*

Aufgabe 3.2.7 *Zeige: Für alle $x \in \mathbb{K}^n$ gilt $\lim\limits_{p\to\infty} \|x\|_p = \|x\|_\infty$.*

Aufgabe 3.2.8 *Finde eine Funktionenfolge (f_n) in $C[a, b]$, für die gilt*

$$\|f_n\|_\infty = 1 \ \forall\ n \geq 1, \qquad \lim_{n\to\infty} \|f_n\|_p = 0 \ \forall\ p \in [1, \infty).$$

Schließe daraus, dass die p-Norm und die Supremumsnorm auf $C[a, b]$ nicht äquivalent sind.

3.3 Innere Punkte und offene Mengen

Definition 3.3.1 *Sei X ein normierter Raum. Für ein $r \in \mathbb{R}_+$ und $x_0 \in X$ heißt*

$$\mathcal{U}_r(x_0) = K(x_0, r) := \{x \in X : \ \|x - x_0\| < r\}$$

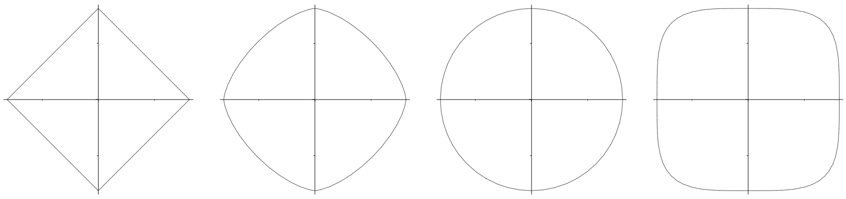
die r-Umgebung von x_0 *oder* die Kugel um x_0 mit Radius r*. Ein x_0 heißt* innerer Punkt *einer Teilmenge $A \subset X$, wenn es ein $r > 0$ gibt mit $\mathcal{U}_r(x_0) \subset A$. Die Teilmenge A heißt* offen*, wenn jeder ihrer Punkte ein innerer Punkt ist.*

Aufgabe 3.3.2 *Skizziere in $\mathbb{R}^2$ die Kugel vom Radius $r = 1$ um den Nullpunkt für eine der p-Normen mit $p = 1$, $p = 3/2$, $p = 2$ und $p = 4$.*

Lösung: Eine elegante Lösung der Aufgabe mit MAPLE geschieht mit dem Kommando

```
> plot([signum(cos(t))*(abs(cos(t)))^(2/p),
  signum(sin(t))*(abs(sin(t)))^(2/p),t=0..2*Pi],
  scaling=constrained,tickmarks=[1,1]);
```

und vorherige Zuweisung der einzelnen Werte für p. Tut man dies, ergeben sich die folgenden vier Bilder:



Man kann ganz allgemein zeigen, dass die Kontur der Kugel für wachsendes p gegen die entsprechende Figur für den Wert $p = \infty$ strebt, und diese ist ein achsenparalleles Quadrat der Seitenlänge 2 mit dem Ursprung als Mittelpunkt. $\square$

Aufgabe 3.3.3 *Begründe mit Hilfe einer Skizze, warum in $\mathbb{R}^2$ mit der euklidischen Norm die Menge $O = \{(x, y)^T : \quad x < 2\, y^2\}$ offen ist.*

Behauptung 3.3.4 *In jedem normierten Raum sind Kugeln um beliebige Punkte mit beliebigen Radien immer offen.*

Beweis: Seien $x_0 \in X$, $r \in \mathbb{R}_+$, und sei $x \in K(x_0, r)$, also $\|x - x_0\| < r$. Sei $\rho = r - \|x - x_0\|$. Dann ist $\rho > 0$, und für $\tilde{x} \in \mathcal{U}_\rho(x)$ folgt mit der Dreiecksungleichung

$$\|\tilde{x} - x_0\| \;\le\; \|\tilde{x} - x\| \;+\; \|x - x_0\| \;<\; \rho \;+\; \|x - x_0\| \;=\; r.$$

Also ist $\mathcal{U}_\rho(x) \subset K(x_0, r)$, und somit ist x innerer Punkt von $K(x_0, r)$. Daher folgt die Behauptung. $\square$

Proposition 3.3.5 (Rechenregeln für offene Mengen) *In jedem normierten Raum X gilt:*

(a) *Der Gesamtraum X und die leere Menge sind offen.*

(b) *Der Durchschnitt endlich vieler offener Mengen ist offen.*

(c) *Die Vereinigung beliebig vieler offener Mengen ist offen.*

Beweis: Aussage (a) folgt sofort mit der Definition offener Mengen. Zu (b): Seien $O_j \subset X$ alle offen, für $j = 1, \ldots, n$. Sei O ihr Durchschnitt, und sei $x \in O$. Dann gibt es Radien $r_j > 0$ mit $\mathcal{U}_{r_j}(x) \subset O_j$ für alle $j = 1, \ldots, n$. Sei $r = \min\{r_1, \ldots, r_n\}$. Dann folgt $\mathcal{U}_r(x) \subset \mathcal{U}_{r_j}(x) \subset O_j$ für alle $j = 1, \ldots, n$, und deshalb ist $\mathcal{U}_r(x) \subset O$. Daher ist x innerer Punkt von O, und da $x \in O$ beliebig war, folgt dass O offen ist. Zu (c): Seien $O_j \subset X$ offen, für alle $j \in J$. Sei O die Vereinigung aller O_j, und sei $x \in O$. Dann gibt es ein $j = j_x \in J$ mit $x \in O_j$, und somit existiert ein $r = r_j > 0$ mit $\mathcal{U}_r(x) \subset O_j$. Dann gilt aber erst recht $\mathcal{U}_r(x) \subset O$, und deshalb ist O offen. $\square$

Aufgabe 3.3.6 *Sei X ein normierter Raum, und seien $O_n = K(0, 1/n) \subset X$ für $n \in \mathbb{N}$. Zeige: Alle O_n sind offen, aber ihr Durchschnitt ist es nicht.*

Definition 3.3.7 *Sei X ein normierter Raum. Die Familie aller offenen Teilmengen von X heißt auch* eine Topologie auf *X.*

Proposition 3.3.8 *Äquivalente Normen auf einem Vektorraum X über $\mathbb{K}$ definieren die gleiche Topologie auf X.*

(Ohne Beweis)

Nach obiger Proposition ist es im Folgenden unerheblich, welche der p-Normen auf $\mathbb{K}^n$ wir betrachten. Wir wollen aber vereinbaren, dass wir immer die euklidische Norm zugrundelegen, wenn nicht ausdrücklich etwas anderes gesagt ist!

Aufgabe 3.3.9 *Zeige: Jedes offene Intervall ist eine offene Teilmenge von $\mathbb{R}$, und jede offene Teilmenge von $\mathbb{R}$ ist die Vereinigung von höchstens abzählbar vielen offenen Intervallen.*

3.4 Häufungspunkte, isolierte Punkte, abgeschlossene Mengen

Definition 3.4.1 *Sei X ein normierter Raum, und sei $B \subset X$ beliebig. Ein Punkt $x_0 \in X$ heißt* Häufungspunkt *von B, wenn*

$$\forall\, \varepsilon > 0 \quad \exists\, x \in B \setminus \{x_0\} : \quad \|x - x_0\| < \varepsilon\,.$$

Die Menge aller Häufungspunkte von B sei mit B' bezeichnet. Ein Punkt $x_0 \in B$ heißt ein isolierter Punkt *von B, wenn er kein Häufungspunkt von B ist. Eine Teilmenge $A \subset X$ heißt* abgeschlossen*, falls ihr Komplement $O = X \setminus A$ offen ist.*

Lemma 3.4.2 *Sei X ein normierter Raum, und sei $B \subset X$ beliebig. Genau dann ist B abgeschlossen, wenn es alle seine Häufungspunkte enthält.*

Beweis: Sei $x_0 \notin B$ ein Häufungspunkt von B. Dann gehört x_0 zum Komplement von B, kann aber kein innerer Punkt des Komplementes sein. Also kann B nicht abgeschlossen sein. Sei jetzt B nicht abgeschlossen. Dann gibt es ein $x_0 \notin B$, welches kein innerer Punkt des Komplementes von B ist. Also muss es in jeder Kugel um x_0 einen Punkt von B geben, und deshalb ist x_0 ein Häufungspunkt von B. □

Proposition 3.4.3 (Rechenregeln für abgeschlossene Mengen) *Gegeben sei ein normierter Raum X. Dann gelten die folgenden Regeln:*

(a) *Der Gesamtraum X und die leere Menge sind abgeschlossen.*

(b) *Die Vereinigung endlich vieler abgeschlossener Mengen ist abgeschlossen.*

(c) *Der Durchschnitt beliebig vieler abgeschlossener Mengen ist abgeschlossen.*

Beweis: Folgt mit Hilfe der de Morganschen Regeln aus den ensprechenden Regeln für offene Mengen. □

Satz 3.4.4 *Sei X ein normierter Raum, und sei eine Teilmenge $A \subset X$ offen und abgeschlossen zugleich. Dann ist $A = X$ oder $A = \emptyset$.*

Beweis: Sei $B = X \setminus A$, und seien A und B beide nicht leer. Für $x_0 \in A$ und $x_1 \in B$ bilden die Vektoren

$$x(t) \;=\; t\,x_1 \;+\; (1-t)\,x_0\,, \qquad 0 \;\le t \;\le 1,$$

die *Verbindungsstrecke* von x_0 nach x_1. Sei $t_0 = \sup\{t \in [0,1] : \; x(t) \in A\}$. Dann kann $x = x(t_0)$ kein innerer Punkt von A, aber auch keiner von B sein. Falls x zu A gehört, ist deshalb A nicht offen, und sonst muss $x \in B$ sein, so dass dann B nicht offen sein kann. □

Aufgabe 3.4.5 *Finde abgeschlossene Mengen in einem normierten Raum X, deren Vereinigung nicht abgeschlossen ist.*

3.5 Abgeschlossene Hülle, offener Kern, Rand

Definition 3.5.1 *Sei X ein normierter Raum. Für ein $B \subset X$ sei*

$$\overset{\circ}{B} \;=\; \bigcup_{\substack{O \subset B \\ O \text{ offen}}} O.$$

Nach den Rechenregeln für offene Mengen ist $\overset{\circ}{B}$ offen, und wenn O eine offene Teilmenge von B ist, dann folgt $O \subset \overset{\circ}{B}$. Also ist $\overset{\circ}{B}$ die größte offene Teilmenge von B *und wird als* der offene Kern von B *bezeichnet. Weiter sei*

$$\overline{B} \;=\; \bigcap_{\substack{A \supset B \\ A \text{ abgeschl.}}} A.$$

Nach den Rechenregeln für abgeschlossene Mengen ist $\overline{B}$ abgeschlossen, und wenn A eine abgeschlossene Obermenge von B ist, dann folgt $A \supset \overline{B}$. Also ist $\overline{B}$ die kleinste abgeschlossene Obermenge von B *und wird als* die abgeschlossene Hülle von B *bezeichnet.*

Die Menge $\partial B = \overline{B} \setminus \overset{\circ}{B}$ heißt der Rand *von B, und jedes $x \in \partial B$ heißt* ein Randpunkt *von B.*

Aufgabe 3.5.2 *Sei X ein normierter Raum, und sei $B \subset X$. Zeige:*

$$\overline{B} \;=\; B \iff B \text{ abgeschlossen}, \qquad \overset{\circ}{B} \;=\; B \iff B \text{ offen}.$$

Proposition 3.5.3 *Sei X normierter Raum, und seien $B \subset X$, $C = X \setminus B$. Dann gilt:*

(a) $\overset{\circ}{B}$ *ist die Menge aller inneren Punkte von B.*

(b) $\overline{C} \;=\; X \setminus \overset{\circ}{B}, \qquad \overset{\circ}{C} \;=\; X \setminus \overline{B}.$

(c) $\partial C \;=\; \partial B \;=\; X \setminus (\overset{\circ}{B} \cup \overset{\circ}{C}).$

(d) $\overline{B} \;=\; B \,\cup\, B' \;=\; B \,\cup\, \partial B.$

(Ohne Beweis)

Aufgabe 3.5.4 *Sei X ein normierter Raum. Zeige: Ist $B \subset X$ mit $\partial B = \emptyset$, so folgt $B = X$ oder $B = \emptyset$. Überlege, ob auch die Umkehrung gilt.*

Aufgabe 3.5.5 *Sei X normierter Raum, und seien $x_0 \in X$ und $r \in \mathbb{R}_+$. Zeige:*

$$\overline{K(x_0, r)} = \{x \in X : \; \|x - x_0\| \le r\}\,.$$

Wir nennen diese Menge auch die abgeschlossene Kugel um x_0 vom Radius r.

3.6 Beschränkte und kompakte Mengen

Definition 3.6.1 *Sei X ein normierter Raum, und sei $B \subset X$. Wir nennen B* beschränkt*, falls ein $r \in \mathbb{R}_+$ existiert, für welches $B \subset K(0, r)$ ist.*

Wir sagen, dass die Mengen $\{O_\alpha : \alpha \in I\}$ *eine* offene Überdeckung *von* B *bilden, wenn alle* O_α *offen sind, und wenn*

$$B \subset \bigcup_{\alpha \in I} O_\alpha .$$

Wir sagen weiter, dass eine solche offene Überdeckung von B eine endliche Teilüberdeckung besitzt*, wenn es* $\alpha_1, \ldots, \alpha_n \in I$ *gibt mit*

$$B \subset \bigcup_{j=1}^{n} O_{\alpha_j} .$$

Wir nennen die Menge B kompakt*, wenn* jede *offene Überdeckung von* B *eine endliche Teilüberdeckung besitzt.*

Aufgabe 3.6.2 (Durchmesser einer Menge) *Sei* X *ein normierter Raum, und sei* $B \subset X$*. Wir nennen* $d = \sup\{\|x_1 - x_2\| : x_1, x_2 \in B\}$ den Durchmesser *von* B*. Zeige: Genau dann ist* B *beschränkt, wenn sein Durchmesser endlich ist.*

Aufgabe 3.6.3 *Sei* $X \neq \{0\}$ *ein normierter Raum, und seien* $x_0 \in X$ *und* $r \in \mathbb{R}_+$*. Berechne den Durchmesser* d *von* $K(x_0, r)$*. Untersuche ob dieser Durchmesser angenommen wird, d. h., ob es* $x_1, x_2 \in K(x_0, r)$ *mit* $d = \|x_1 - x_2\|$ *gibt.*

Satz 3.6.4 *Sei* X *normierter Raum, und sei* $B \subset X$ *kompakt. Dann ist* B *abgeschlossen und beschränkt.*

Beweis: Die Menge aller Kugeln $K(0, n)$, mit beliebigem $n \in \mathbb{N}$, bilden sicher eine offene Überdeckung von B. Da B kompakt ist, muss es eine endliche Teilüberdeckung geben. Da $K(0, n) \subset K(0, n+1)$ ist, bedeutet das die Existenz eines $n_0 \in \mathbb{N}$ mit $B \subset K(0, n_0)$. Also ist B beschränkt.

Sei $x_0 \notin B$. Die Mengen $O_n = \{x \in X : \|x - x_0\| > 1/n\}$, $n \in \mathbb{N}$, sind offen wegen Aufgabe 3.5.5 und bilden offenbar eine Überdeckung von B. Da B kompakt ist, und da wiederum $O_n \subset O_{n+1}$ ist, gibt es ein $n_0 \in \mathbb{N}$ mit $B \subset O_{n_0}$. Daraus folgt, dass $K(x_0, 1/n_0)$ keinen Punkt von B enthält, und deshalb ist x_0 kein Häufungspunkt von B. Also ist B abgeschlossen. □

Aufgabe 3.6.5 *Zeige: Ist* X *normierter Raum und* $K \subset X$ *kompakt, so ist jede abgeschlossene Teilmenge von* K *ebenfalls kompakt. Insbesondere ist die leere Menge kompakt!*

Aufgabe 3.6.6 *Zeige: In jedem normierten Raum ist der Durchschnitt beliebig vieler kompakter Mengen wieder kompakt.*

3.7 Konvexe und sternförmige Mengen

Definition 3.7.1 *Sei* X *ein beliebiger Vektorraum über* $\mathbb{K}$*. Eine Teilmenge* $B \subset X$ *heißt* sternförmig *bzgl. eines* $x_0 \in B$*, wenn für jedes* $x \in B$ *die ganze Verbindungsstrecke von* x_0 *nach* x *zu* B *gehört. Wir nennen* B konvex*, wenn es sternförmig bzgl. jedes Punktes* $x_0 \in B$ *ist. Also ist* B *genau dann konvex, wenn mit je zwei Punkten auch deren Verbindungsstrecke zu* B *gehört.*

Aufgabe 3.7.2 *Sei* X *normierter Raum. Zeige: Jede Kugel in* X *ist konvex.*

Behauptung 3.7.3 *Durchschnitte beliebig vieler konvexer Mengen sind wieder konvex.*

Beweis: Folgt direkt aus der Definition (auch, falls der Durchschnitt leer sein sollte). □

Definition 3.7.4 *Sei X ein beliebiger Vektorraum über $\mathbb{K}$, und sei $B \subset X$. Die Menge*

$$B_h \;=\; \bigcap_{\substack{C \supset B \\ C \text{ konvex}}} C$$

ist nach obiger Behauptung konvex und enthält B. Also ist B_h die kleinste konvexe Obermenge *von B und wird* die konvexe Hülle *von B genannt.*

Aufgabe 3.7.5 *Sei X ein beliebiger Vektorraum über $\mathbb{K}$, und sei $x_0 \in X$. Zeige: Der Durchschnitt beliebig vieler bzgl. x_0 sternförmiger Mengen ist wieder sternförmig bzgl. x_0. Benutze dies, um eine sternförmige Hülle bzgl. x_0 zu definieren.*

Aufgabe 3.7.6 *Sei X ein beliebiger Vektorraum über $\mathbb{K}$, sei $A \subset X$ konvex, und seien $x_1, \ldots, x_n \in A$ sowie $\alpha_1, \ldots, \alpha_n \in \mathbb{R}$ mit $\alpha_j \geq 0$ für alle j und $\sum_{j=1}^n \alpha_j = 1$. Zeige:*

$$\sum_{j=1}^{n} \alpha_j \, x_j \;\in\; A\,.$$

Man nennt jede solche Summe auch eine Konvexkombination *der Vektoren $x_1, \ldots, x_n$.*

Aufgabe 3.7.7 *Sei X ein beliebiger Vektorraum über $\mathbb{K}$, und seien $x_1, \ldots, x_n \in X$ gegeben. Zeige: Die Menge aller Konvexkombinationen der $x_1, \ldots, x_n$ ist genau gleich der konvexen Hülle der endlichen Menge $\{x_1, \ldots, x_n\}$. Interpretiere diese Menge anschaulich für $n \leq 3$.*

3.8 Zusammenhang

Definition 3.8.1 *Sei X ein normierter Raum, und sei $B \subset X$. Wir nennen B* unzusammenhängend, *falls offene Mengen $O_1, O_2 \subset X$ existieren mit*

$$B \;\subset\; O_1 \cup O_2\,, \quad B \cap O_1 \cap O_2 \;=\; \emptyset\,, \quad B \cap O_j \;\neq\; \emptyset\,, \quad 1 \leq j \leq 2.$$

Falls dies nicht so ist, heißt B zusammenhängend. *Seien $x_0, \ldots, x_m \in X$. Die Menge aller Verbindungsstrecken von x_{j-1} nach x_j, für $j = 1, \ldots, m$, heißt* der Streckenzug *oder* das Polygon *zu den Eckpunkten x_j. Wir sagen auch: Der Streckenzug* verbindet x_0 mit x_m in der Menge B, *wenn alle Verbindungsstrecken in B liegen. Wir nennen B* polygonzusammenhängend, *wenn sich je zwei Punkte von B durch ein Polygon in B verbinden lassen.*

Aufgabe 3.8.2 *Gib ein Beispiel für eine Teilmenge von $\mathbb{R}^2$, welche zusammenhängend, aber nicht polygonzusammenhängend ist.*

Aufgabe 3.8.3 *Zeige: Konvexe und sternförmige Mengen sind immer polygonzusammenhängend.*

Lemma 3.8.4 *Sei X ein normierter Raum, und sei $B \subset X$ polygonzusammenhängend. Dann ist B zusammenhängend.*

Beweis: Seien offene Mengen O_1, O_2 in X mit $B \subset O_1 \cup O_2$, $B \cap O_j \neq \emptyset$, für $1 \leq j \leq 2$, gegeben. Zu zeigen ist dann $B \cap O_1 \cap O_2 \neq \emptyset$. Seien $x \in B \cap O_1$, $\tilde{x} \in B \cap O_2$ gewählt. Dann gibt es einen Streckenzug, der beide Punkte in B verbindet. Seien $x_0(= x), x_1, \ldots, x_m(= \tilde{x})$ seine Eckpunkte, und sei $k \in \{0, \ldots, m\}$ maximal gewählt, so dass $x_k \in B \cap O_1$ gilt. Falls $k = m$ ist, ist nichts mehr zu zeigen. Andernfalls sei $t_0 = \sup\{t \in [0, 1] : \; x(t) = (1-t)\, x_k + t\, x_{k+1} \in O_1\}$. Da O_1 offen ist, folgt $x(t_0) \notin O_1$, also $x(t_0) \in O_2$. Da O_2 ebenfalls offen ist, gilt $x(t) \in O_2$ für $t < t_0$, mit $t_0 - t$ genügend klein. Dann ist aber $x(t) \in B \cap O_1 \cap O_2$, was zu zeigen war. □

Definition 3.8.5 *Sei X ein normierter Raum. Eine offene und zusammenhängende Teilmenge von X heißt* ein Gebiet.

Satz 3.8.6 (Polygonzusammenhang von Gebieten) *In einem normierten Raum ist jedes Gebiet G polygonzusammenhängend.*

(Ohne Beweis)

Aufgabe 3.8.7 *Sei B eine zusammenhängende Teilmenge eines normierten Raumes X, und sei $B \subset C \subset \overline{B}$. Zeige: Dann ist auch C zusammenhängend.*

Aufgabe 3.8.8 *Zeige: Eine Teilmenge von $\mathbb{R}$ ist genau dann zusammenhängend, wenn sie ein Intervall ist.*

Kapitel 4

Konvergenz und Stetigkeit

4.1 Konvergenz, Vollständigkeit

Definition 4.1.1 *Sei X ein normierter Raum, und sei $(x_m)_{m=p}^\infty$ eine Folge in X. Wir nennen (x_m)* beschränkt*, wenn ein $r \in \mathbb{R}_+$ existiert mit $\|x_m\| \le r$ für alle $m \ge p$. Wir nennen (x_m)* Cauchy-Folge*, wenn*

$$\forall\, \varepsilon > 0 \quad \exists\, N \in \mathbb{R}_+ \quad \forall\, m, \mu \in \mathbb{N} : \quad m, \mu \ge N \implies \|x_m - x_\mu\| < \varepsilon\,.$$

Weiter heißt (x_m) konvergent*, wenn ein $x \in X$ existiert, für welches*

$$\forall\, \varepsilon > 0 \quad \exists\, N \in \mathbb{R}_+ \quad \forall\, m \in \mathbb{N} : \quad m \ge N \implies \|x_m - x\| < \varepsilon\,.$$

Wenn dem so ist, heißt x Grenzwert *der Folge. Wir schreiben dann $\lim\limits_{m\to\infty} x_m = x$ oder $x_m \longrightarrow x$ $(m \to \infty)$. Schließlich heißt die Folge (x_m)* divergent*, wenn sie nicht konvergiert.*

Die folgenden Ergebnisse beweist man genauso wie die entsprechenden Aussagen aus dem ersten Semester, wobei man für Vektoren statt $|\cdot|$ jeweils $\|\cdot\|$ schreiben muss. Beachte allerdings, dass nicht behauptet wird, dass jede Cauchy-Folge konvergiert; dies trifft im Allgemeinen nicht zu!

Proposition 4.1.2 *In jedem normierten Raum X gilt:*

(a) *Eine konvergente Folge hat nur einen Grenzwert.*

(b) *Eine konvergente Folge ist Cauchy-Folge.*

(c) *Eine Cauchy-Folge ist beschränkt.*

(d) *Wenn eine Cauchy-Folge eine konvergente Teilfolge besitzt, so ist sie selber konvergent.*

Satz 4.1.3 (Rechenregeln für Grenzwerte) *Sei X normierter Raum, und seien $x_m, y_m \in X$, $\alpha_m \in \mathbb{K}$, für alle $m \in \mathbb{N}$.*

(a) *Falls die Folgen (x_m) und (y_m) beide konvergieren, dann ist auch $(x_m + y_m)$ konvergent, und es gilt*

$$\lim_{m\to\infty} (x_m + y_m) \;=\; \lim_{m\to\infty} x_m \;+\; \lim_{m\to\infty} y_m\,.$$

(b) *Falls die Folgen (x_m) und (α_m) beide konvergieren, dann ist auch $(\alpha_m\, x_m)$ konvergent, und es gilt*

$$\lim_{m\to\infty} \alpha_m\, x_m \;=\; \lim_{m\to\infty} \alpha_m \,\lim_{m\to\infty} x_m\,.$$

Definition 4.1.4 *Ein normierter Raum X heißt* vollständig, *falls jede Cauchy-Folge konvergiert. Ein vollständiger normierter Raum heißt auch* Banach-Raum.

Nicht jeder normierte Raum hat diese wichtige Eigenschaft der Vollständigkeit. Wir werden aber in Abschnitt 4.2 zeigen, dass die Räume $\mathbb{K}^n$ vollständig sind.

Aufgabe 4.1.5 *Beweise* Proposition 4.1.2 *und die anschließenden Rechenregeln.*

Aufgabe 4.1.6 *Zeige: In $C[a,b]$ mit der Supremumsnorm bedeutet Konvergenz einer Folge dasselbe wie die gleichmäßige Konvergenz. Schließe aus einem Satz aus der Erstsemestervorlesung, dass $C[a,b]$ vollständig ist. Zeige weiter, dass der Teilraum der Polynome nicht vollständig ist.*

Aufgabe 4.1.7 *Sei X ein normierter Raum, und sei $B \subset X$. Zeige: Jeder Häufungspunkt von B ist Grenzwert einer Folge mit Gliedern in B.*

Aufgabe 4.1.8 *Sei X ein normierter Raum, und sei $B \subset X$. Zeige: Genau dann ist B abgeschlossen, wenn für jede konvergente Folge (x_n), deren Glieder alle zu B gehören, auch der Grenzwert in B liegt.*

Aufgabe 4.1.9 (Intervallschachtelungsprinzip) *Sei X ein vollständiger normierter Raum. Seien A_n nicht-leere, abgeschlossene und beschränkte Teilmengen von X mit $A_n \supset A_{n+1}$, für alle $n \in \mathbb{N}$, und so dass der Durchmesser der A_n für $n \to \infty$ gegen 0 geht. Zeige: Dann gibt es genau ein $x \in X$ mit $x \in A_n$ für alle $n \in \mathbb{N}$.*

4.2 Konvergenz und Kompaktheit

Die für diese Vorlesung wichtigsten Räume sind $\mathbb{R}^n$ oder $\mathbb{C}^n$. Dabei sei $n \in \mathbb{N}$ immer fest gewählt, und wir stellen uns $n \geq 2$ vor, obwohl die gemachten Aussagen auch alle für $n = 1$ richtig sind. Wenn wir einen Vektor $x \in \mathbb{K}^n$ betrachten, soll immer $x = (x_1, \dots, x_n)^T$ gelten; für Folgen (x_m) in $\mathbb{K}^n$ sei entsprechend $x_m = (x_1^{(m)}, \dots, x_n^{(m)})^T$ gesetzt. Als Norm auf $\mathbb{K}^n$ sei stets die euklidische Norm gewählt, obwohl die folgenden Aussagen auch für jede andere p-Norm gelten. Der Einfachheit halber schreiben wir nur $\|\cdot\|$ an Stelle von $\|\cdot\|_2$.

Satz 4.2.1 (Vollständigkeit von $\mathbb{K}^n$) *In jeder Dimension n gilt:*

(a) *Eine Folge (x_m) aus $\mathbb{K}^n$ konvergiert genau dann gegen ein $x \in \mathbb{K}^n$, wenn gilt*

$$\lim_{m\to\infty} x_k^{(m)} \;=\; x_k \qquad \forall\, k \;=\; 1, \dots, n.$$

(b) *Der Raum $\mathbb{K}^n$ ist vollständig.*

Beweis: Zu (a): Gelte $\lim_{m\to\infty} x_m = x$, also per Definition $\lim_{m\to\infty} \|x_m - x\| = 0$. Wegen $|x_k^{(m)} - x_k| \leq \|x_m - x\|$ folgt dann also $\lim_{m\to\infty} x_k^{(m)} = x_k$, für alle $k = 1, \dots, n$. Umgekehrt folgt aus $\lim_{m\to\infty} x_k^{(m)} = x_k$ und den Rechenregeln für Grenzwerte, dass $\lim_{m\to\infty} \sum_{k=1}^n |x_k^{(m)} - x_k|^2 = 0$ ist, und aus der Stetigkeit der Quadratwurzel folgt dann $\lim_{m\to\infty} \|x_m - x\| = 0$. Zu (b): Wenn (x_m) eine Cauchy-Folge in $\mathbb{K}^n$ ist, so ist für jedes $k = 1, \dots, n$ die Folge $(x_k^{(m)})_{m=1}^\infty$ eine Cauchy-Folge in $\mathbb{K}$. Nach dem Cauchy-Kriterium sind diese Folgen alle konvergent. Wenn wir die Grenzwerte mit x_k bezeichnen und daraus einen Vektor $x \in \mathbb{K}^n$ bilden, folgt mit Teil (a) des Satzes, dass (x_m) gegen x konvergiert. □

In einem allgemeinen normierten Raum ist es nicht leicht zu erkennen, ob eine Teilmenge kompakt ist; in $\mathbb{K}^n$ ist dies anders:

Satz 4.2.2 (Satz von Heine-Borel) *Eine Teilmenge von $\mathbb{K}^n$ ist genau dann kompakt, wenn sie abgeschlossen und beschränkt ist.*

(Ohne Beweis)

Korollar zu Satz 4.2.2 (Satz von Bolzano und Weierstraß) *Eine beschränkte Folge in $\mathbb{K}^n$ besitzt eine konvergente Teilfolge.*

(Ohne Beweis)

4.3 Stetigkeit in normierten Räumen

Im Folgenden seien X_1, X_2 zwei normierte Räume. Für diese Vorlesung ist es ausreichend, sich $X_1 = \mathbb{K}^n$ und $X_2 = \mathbb{K}^m$ vorzustellen (mit $n, m \in \mathbb{N}$), aber die folgenden Begriffe sind auch allgemein sinnvoll.

Definition 4.3.1 *Sei $D \subset X_1$ nicht leer, und sei $f : D \longrightarrow X_2$. Wir nennen f* stetig in einem Punkt *$x_0 \in D$, falls gilt*

$$\forall\, \varepsilon > 0 \quad \exists\, \delta > 0 \quad \forall\, x \in D : \quad \|x - x_0\| < \delta \implies \|f(x) - f(x_0)\| < \varepsilon. \tag{4.3.1}$$

Falls f in jedem Punkt von D stetig ist, sagen wir kurz: f ist auf D stetig. *Wir sagen, dass f* auf D einer Lipschitzbedingung genügt, *falls eine Konstante $L \in \mathbb{R}_+$ existiert, so dass*

$$\|f(x_0) \;-\; f(x_1)\| \;\leq\; L\,\|x_0 \;-\; x_1\| \qquad \forall\, x_0, x_1 \in D.$$

Jedes solche L heißt auch Lipschitzkonstante für f (auf D).

Beispiel 4.3.2 *Überraschenderweise ist in unendlich-dimensionalen normierten Räumen nicht jede lineare Abbildung von X_1 nach X_2 stetig. Auf jedem normierten Raum X ist aber die Abbildung $x \longmapsto \|x\|$ stetig, ja erfüllt sogar eine Lipschitzbedingung mit Konstante $L = 1$ auf dem ganzen Raum X; dies ist nämlich genau die Dreiecksungleichung nach unten. Ausserdem ist die identische Abbildung stetig auf X.*

Bemerkung 4.3.3 *Die folgenden Aussagen sind zum Teil im ersten Semester bewiesen worden:*

(a) *Falls x_0 ein isolierter Punkt von D ist, ist jede auf D definierte Funktion in x_0 stetig.*

(b) *Folgenstetigkeit ist äquivalent zur Stetigkeit; d. h., f ist genau dann stetig in x_0, wenn für jede Folge (x_n) aus D gilt*
$$\lim_{n\to\infty} x_n \;=\; x_0 \quad \implies \quad \lim_{n\to\infty} f(x_n) \;=\; f(x_0)\,.$$

(c) *Erfüllt f eine Lipschitzbedingung auf D, so ist f dort stetig.*

(d) *Sind $f : D \longrightarrow X_2$ und $g : D \longrightarrow X_2$ stetig in $x_0 \in D$, so ist auch $g + f$ dort stetig.*

(e) *Sind $f : D \longrightarrow X_2$ und $g : D \longrightarrow \mathbb{K}$ stetig in $x_0 \in D$, so ist auch $g\,f$ dort stetig.*

(f) *Die Hintereinanderausführung stetiger Funktionen ist stetig. Genauer: Ist auch X_3 ein normierter Raum, ist $D_2 \subset X_2$, und sind $f : D \longrightarrow D_2$ stetig in $x_0 \in D$ sowie $g : D_2 \longrightarrow X_3$ stetig in $f(x_0)$, so ist $g \circ f$ stetig in x_0.*

(g) *Im Fall $X_2 = \mathbb{K}$ ist der Kehrwert einer Funktion f dort stetig, wo f selber stetig und von 0 verschieden ist.*

Proposition 4.3.4 *Seien X_1, X_2 normierte Räume, sei $D \subset X_1$ nicht leer, und sei $f : D \longrightarrow X_2$. Genau dann ist f stetig auf D, wenn zu jeder offenen Teilmenge $O_2 \subset X_2$ eine offene Teilmenge $O_1 \subset X_1$ existiert, für welche $f^{-1}(O_2) = D \cap O_1$ ist.*

(Ohne Beweis)

Aufgabe 4.3.5 *Sei $A = [\alpha_{jk}]$ eine (n, m)-Matrix mit Elementen aus $\mathbb{K}$, und sei*

$$\|A\| = \Big(\sum_{j=1}^{n} \sum_{k=1}^{m} |\alpha_{jk}|^2 \Big)^{1/2}.$$

Zeige: $\|A\,x\| \leq \|A\|\,\|x\|$ für alle $x \in \mathbb{K}^m$, wobei links die euklidische Norm in $\mathbb{K}^n$, und rechts für x die euklidische Norm in $\mathbb{K}^m$ zu nehmen ist. Leite daraus die Stetigkeit von $x \mapsto A\,x$ auf $\mathbb{K}^m$ ab.

Aufgabe 4.3.6 *Zeige: Ist $[a, b]$ eine abgeschlossenes Intervall in $\mathbb{R}$, und ist $X = C[a, b]$ der Raum der auf $[a, b]$ stetigen Funktionen mit der Supremumsnorm, so erfüllt die Abbildung $f \longmapsto \int_a^b f(x)\,dx$ auf $C[a, b]$ eine Lipschitzbedingung und ist deshalb dort stetig.*

4.4 Stetigkeit und Zusammenhang

Der folgende Satz entspricht genau dem Zwischenwertsatz aus dem ersten Semester:

Satz 4.4.1 *Seien X_1, X_2 normierte Räume, sei $D \subset X_1$ nicht leer, und sei $f : D \longrightarrow X_2$ stetig auf D. Ist B eine zusammenhängende Teilmenge von D, so ist $f(B)$ ebenfalls zusammenhängend.*

(Ohne Beweis)

Aufgabe 4.4.2 (Zwischenwertsatz) *Sei X ein normierter Raum, sei $D \subset X$ nicht leer, und sei $f : D \longrightarrow \mathbb{R}$ stetig auf D. Zeige mit Hilfe von* Satz 4.4.1*: Ist B eine zusammenhängende Teilmenge von D, und sind $x_1, x_2 \in B$ mit $f(x_1) = y_1 < f(x_2) = y_2$, so gibt es zu jedem $y \in (y_1, y_2)$ ein $x \in B$ mit $f(x) = y$.*

4.5 Stetigkeit und Kompaktheit

Definition 4.5.1 *Seien X_1, X_2 normierte Räume, sei $D \subset X_1$ nicht leer, und sei $f : D \longrightarrow X_2$. Wir nennen f* auf D gleichmäßig stetig, *falls folgendes gilt:*

$$\forall\ \varepsilon > 0\ \exists\ \delta > 0\ \forall\ x_1, x_2 \in D : \ \|x_1 - x_2\| < \delta \implies \|f(x_1) - f(x_2)\| < \varepsilon.$$

Falls $X_2 = \mathbb{R}$ ist, definieren wir die Begriffe Maximum *und* Minimum *genau wie im ersten Semester.*

Satz 4.5.2 *Seien X_1, X_2 normierte Räume, sei $D \subset X_1$ nicht leer und kompakt, und sei $f : D \longrightarrow X_2$ stetig auf D. Dann ist f sogar gleichmäßig stetig auf D, und $f(D)$ ist kompakt. Ist $X_2 = \mathbb{R}$, so nimmt f auf D ein Maximum und ein Minimum an.*

(Ohne Beweis)

Satz 4.5.3 (Stetigkeit der Umkehrabbildung) *Seien X_1, X_2 zwei normierte Räume, sei $D \subset X_1$ nicht leer und kompakt, und sei $f : D \longrightarrow X_2$ stetig und injektiv auf D. Dann ist die Umkehrfunktion auf $\tilde{D} = f(D)$ stetig.*

Beweis: Sei (y_m) eine Folge aus $\tilde{D}$, welche gegen ein $y \in \tilde{D}$ konvergiert. Seien x_m und x die (eindeutig bestimmten) Urbilder dieser Punkte. Zu zeigen ist $\lim_{m\to\infty} x_m = x$. Wenn wir das Gegenteil annehmen, dann existiert ein $\varepsilon > 0$, für welches $\|x_m - x\| \geq \varepsilon$ für unendlich viele m, d. h. für eine Teilfolge von (x_m), gilt. Wegen der Folgenkompaktheit von D enthält diese Teilfolge selbst wieder eine Teilfolge, welche gegen ein $\tilde{x} \in D$ konvergiert. Wegen der Stetigkeit von f konvergieren die Bilder dieser Teilfolge, also die entsprechende Teilfolge von (y_m), gegen $f(\tilde{x})$, und daher gilt $f(\tilde{x}) = y$, also $\tilde{x} = x$. Das widerspricht aber der Wahl der Teilfolge. □

Aufgabe 4.5.4 *Finde ein Beispiel für eine Funktion mit nicht-kompaktem Definitionsbereich aber kompakter Wertemenge.*

Aufgabe 4.5.5 *Finde ein Beispiel für eine Funktion mit beschränktem Definitionsbereich aber unbeschränkter Wertemenge.*

4.6 Stetigkeit und gleichmäßige Konvergenz

Definition 4.6.1 *Seien X_1, X_2 normierte Räume, sei $D \subset X_1$ nicht leer, und seien $f_m, g_k : D \longrightarrow X_2$. Dann nennen wir (f_m) eine* Funktionenfolge *auf D, und $\sum_{k=1}^\infty g_k$ eine* Funktionenreihe *auf D. Die Funktionenfolge (f_m) heißt* auf D punktweise konvergent*, falls für jedes $x \in D$ die Folge $(f_m(x))$ konvergent ist. Ist dies der Fall, so heißt $f : D \longrightarrow X_2$, mit $f(x) = \lim_{m\to\infty} f_m(x)$ für alle $x \in D$,* die Grenzfunktion *der Folge. Analog heißt die Funktionenreihe $\sum_{k=1}^\infty g_k$* auf D punktweise konvergent*, falls für jedes $x \in D$ die Reihe $\sum_{k=1}^\infty g_k(x)$ konvergent ist, und die Grenzfunktion f ist dann durch $f(x) = \sum_{k=1}^\infty g_k(x)$ für alle $x \in D$ gegeben. Die Funktionenfolge (f_m) heißt* auf D gleichmäßig konvergent*, falls sie punktweise gegen die Grenzfunktion f konvergiert, und falls weiter gilt:*

$$\forall\, \varepsilon > 0\ \exists\, N \in \mathbb{R}_+\ \forall\, m \in \mathbb{N}\ \forall\, x \in D : m \geq N \implies \|f_m(x) - f(x)\| < \varepsilon.$$

Analog heißt die Funktionenreihe $\sum_{k=1}^\infty g_k$ auf D gleichmäßig konvergent*, falls sie punktweise gegen die Grenzfunktion f konvergiert, und falls weiter gilt:*

$$\forall\, \varepsilon > 0\ \exists\, N \in \mathbb{R}_+\ \forall\, m \in \mathbb{N}\ \forall\, x \in D : m \geq N \implies \Big\| \sum_{k=1}^{m} g_k(x) - f(x) \Big\| < \varepsilon.$$

Die gleichmäßige Konvergenz der Reihe $\sum_{k=1}^\infty g_k$ ist also äquivalent mit der gleichmäßigen Konvergenz der Folge ihrer Partialsummen.

Genau wie im ersten Semester zeigt man die folgenden Resultate in jedem vollständigen normierten Raum X:

Satz 4.6.2 (Cauchy-Kriterium für gleichmäßige Konvergenz) *Mit den Bezeichnungen wie in obiger Definition gilt: Die Funktionenfolge (f_n) ist genau dann auf D gleichmäßig konvergent, wenn*

$$\forall\, \varepsilon > 0\ \exists\, N \in \mathbb{R}_+\ \forall\, m, \mu \in \mathbb{N}\ \forall x \in D : m, \mu \geq N \implies \|f_m(x) - f_\mu(x)\| < \varepsilon.$$

Die Funktionenreihe $\sum_{k=1}^\infty g_k$ ist genau dann auf D gleichmäßig konvergent, wenn

$$\forall\, \varepsilon > 0\ \exists\, N \in \mathbb{R}_+\ \forall\, m, \mu \in \mathbb{N}\ \forall x \in D : \mu \geq m \geq N \implies \Big\| \sum_{k=m+1}^{\mu} g_k(x) \Big\| < \varepsilon.$$

Satz 4.6.3 (Majorantenkriterium für gleichmäßige Konv.) *Mit den Bezeichnungen wie in obiger Definition gilt: Falls Zahlen $a_k \in \mathbb{R}_+$ existieren, für welche gilt*

$$\|g_k(x)\| \;\le\; a_k \;\forall\, x \in D, \qquad \sum_{k=1}^{\infty} a_k \;<\; \infty,$$

dann ist die Funktionenreihe $\sum_{k=1}^{\infty} g_k$ auf D gleichmäßig konvergent.

Satz 4.6.4 *Mit den Bezeichnungen wie in obiger Definition gilt:*

(a) *Sind alle f_n in einem Punkt $x_0 \in D$ stetig, und ist die Funktionenfolge (f_n) auf D gleichmäßig konvergent, so ist die Grenzfunktion f ebenfalls stetig in x_0.*

(b) *Sind alle g_k in einem Punkt $x_0 \in D$ stetig, und ist die Funktionenreihe $\sum_{k=1}^{\infty} g_k$ auf D gleichmäßig konvergent, so ist die Grenzfunktion f ebenfalls stetig in x_0.*

Aufgabe 4.6.5 *Seien $g_k(x_1, x_2) = (x_1^k + x_2)/k^2$, für $k \in \mathbb{N}$ und $x_1, x_2 \in \mathbb{R}$. Finde möglichst große Bereiche D, auf denen die Funktionenreihe $\sum_{k=1}^{\infty} g_k$ gleichmäßig konvergiert, und untersuche dort die Stetigkeit der Grenzfunktion.*

4.7 Stetigkeit von vektorwertigen Abbildungen

Definition 4.7.1 *Sei X ein normierter Raum, sei $D \subset X$ nicht leer, und sei $f : D \longrightarrow \mathbb{K}^m$. Dann ist jeder Funktionswert $f(x)$ ein Vektor*

$$f(x) \;=\; (f_1(x), \ldots, f_m(x))^T \qquad \forall\, x \in D.$$

Wir nennen dann die Funktionen $f_k : D \longrightarrow \mathbb{K}$ auch die Koordinatenfunktionen *der vektorwertigen Funktion f und schreiben kurz $f = (f_1, \ldots, f_n)^T$. Ist $X = \mathbb{K}^n$, so nennen wir f, aber auch die f_k,* Funktionen der n Variablen $x_1, \ldots, x_n$. *Wenn $n \ge 2$ ist, sprechen wir auch von Funktionen* mehrerer Variabler *und schreiben statt $f(x)$ auch $f(x_1, \ldots, x_n)$.*

Bemerkung 4.7.2 *Wenn eine Funktion f wie oben gegeben ist, zeigt man leicht, dass die Stetigkeit von f an einer Stelle x_0 äquivalent zur Stetigkeit jeder der Koordinatenfunktionen an dieser Stelle ist. Beachte allerdings, dass bei Funktionen mehrerer Variabler die Stetigkeit nicht schon dann gesichert ist, wenn die Funktion stetig in jeder einzelnen Variablen ist! Dies ist gerade Inhalt der folgenden Aufgaben.*

Aufgabe 4.7.3 *Benutze die Äquivalenz von Stetigkeit und Folgenstetigkeit zusammen mit* Satz 1*, um zu zeigen, dass die Stetigkeit von f äquivalent zur Stetigkeit der Koordinatenfunktionen ist.*

Aufgabe 4.7.4 *Für $x_1, x_2 \in \mathbb{R}$ sei*

$$f(x_1, x_2) \;=\; \begin{cases} \dfrac{x_1^2\, x_2}{x_1^4 + x_2^2} & \text{für } (x_1, x_2) \neq (0,0), \\ 0 & \text{für } (x_1, x_2) = (0,0). \end{cases}$$

Zeige: Wenn man für x_1 oder x_2 den Wert 0 einsetzt, wird dieses f eine stetige Funktion der anderen Variablen auf ganz $\mathbb{R}$. Als Funktion von zwei Variablen ist f aber nicht stetig im Nullpunkt.

4.8 Polynome, rationale Funktionen und Potenzreihen

Definition 4.8.1 *Für $x = (x_1, \dots, x_n)^T \in \mathbb{K}^n$ und $p = (p_1, \dots, p_n) \in \mathbb{N}_0^n$ sei*

$$x^p \;=\; x_1^{p_1} \cdot \ldots \cdot x_n^{p_n} \,.$$

Jedes solche p heißt ein Multi-Index*, und die Zahl $|p| = p_1 + \ldots + p_n$ heißt* Betrag *oder* Länge *von p.*

Eine Funktion $P : \mathbb{K}^n \longrightarrow \mathbb{K}$ der Form

$$P(x) \;=\; P(x_1, \dots, x_n) \;=\; \sum_{0 \le |p| \le m} a_p\, x^p \qquad \forall\; x \in \mathbb{K}^n$$

mit $a_p \in \mathbb{K}$, heißt ein Polynom *in n Variablen. Die Zahlen a_p heißen* die Koeffizienten *des Polynoms. Für die Menge aller Polynome mit Koeffizienten in $\mathbb{K}$ schreiben wir wieder $\mathbb{K}[x]$; beachte aber, dass hier $x \in \mathbb{K}^n$ ist! Falls alle Koeffizienten $a_p = 0$ sind, ist $P(x) = 0$ für alle $x \in \mathbb{K}^n$, und wir nennen dieses Polynom wieder* das Nullpolynom *oder* die Nullfunktion *Im anderen Fall können wir annehmen, dass $a_p \neq 0$ ist für mindestens ein p mit $|p| = m$, und dann nennen wir m auch* den Grad *des Polynoms. Den Grad des Nullpolynoms definieren wir wieder als $-\infty$.*

Für zwei Polynome $p, q \in \mathbb{K}[x]$, wobei q nicht das Nullpolynom ist, heißt $r = p/q$ eine rationale Funktion *in n Variablen. Sie ist überall da definiert, wo das Nennerpolynom nicht verschwindet. Die Menge aller rationalen Funktionen wird wieder mit $\mathbb{K}(x)$ bezeichnet.*

Beispiel 4.8.2 *Sei A eine quadratische n-reihige Matrix. Dann ist die Abbildung $p(x) = x^T A\, x$, $x \in \mathbb{K}^n$, ein Polynom. Der Grad dieses Polynoms ist gleich zwei, außer wenn A die Nullmatrix ist. Man spricht dann auch von der* quadratischen Form mit der Koeffizientenmatrix A. *Für $a \in \mathbb{K}^n$ ist $p(x) = a^T\, x$ ein Polynom vom Grad eins, außer wenn a der Nullvektor ist.*

Definition 4.8.3 *Ein Ausdruck der Form*

$$\sum_p a_p\; (z - z_0)^p\,, \qquad a_p \in \mathbb{C}, \quad z_0 \;=\; (z_1^{(0)}, \dots, z_n^{(0)})^T \in \mathbb{C}^n \tag{4.8.1}$$

mit Summation über alle Multi-Indizes p, heißt eine Potenzreihe in n Variablen $z = (z_1, \dots, z_n)^T$. *Die Zahlen a_p heißen* die Koeffizienten*, und z_0 heißt* der Entwicklungspunkt *der Potenzreihe.*

Beispiel 4.8.4 *Für $z_0 = 0$ und $a_p = 1$ für alle p erhalten wir die* geometrische Reihe in n Variablen. *Es gilt*

$$\sum_p z^p \;=\; \sum_{p_1=0}^{\infty} \cdots \sum_{p_1=0}^{\infty} z_1^{p_1} \cdot \ldots \cdot z_n^{p_n} \;=\; \frac{1}{1 - z_1} \cdot \ldots \cdot \frac{1}{1 - z_n}\,,$$

wobei die Reihe für alle z mit $\|z\|_\infty = \max\{|z_1|, \dots, |z_n|\} < 1$ absolut konvergiert.

Bemerkung 4.8.5 *Die Menge der Multi-Indizes ist abzählbar unendlich, also kann man durch Anordnung der Terme eine Potenzreihe in mehreren Variablen als normale Reihe schreiben. Allerdings ist klar, dass bei Konvergenzuntersuchungen der richtige Konvergenzbegriff der der* absoluten Konvergenz *sein muss, da sonst der Wert der Reihe von der gewählten Anordnung der Terme abhängen kann.*

Für Potenzreihen in mehreren Variablen gibt es nichts, was dem Konvergenzradius bei bewöhnlichen Potenzreihen entspricht. Das Konvergenzverhalten einer solchen Reihe ist nicht leicht zu beschreiben. Es gilt aber folgendes Ergebnis:

Proposition 4.8.6 (Konvergenzgebiet einer Potenzreihe in mehreren Variablen) *Für Potenzreihen in mehreren Variablen gelten folgende Resultate:*

(a) *Sei die Reihe* (4.8.1) *für ein* $z = z_1 = (z_1^{(1)}, \ldots, z_n^{(1)})^T$ *mit* $z_k^{(1)} \neq z_k^{(0)}$, $1 \leq k \leq n$, *absolut konvergent. Dann konvergiert sie absolut für alle* $z = (z_1, \ldots, z_n)^T$ *mit* $|z_k - z_k^{(0)}| < |z_k^{(1)} - z_k^{(0)}|$ *für alle* $k = 1, \ldots, n$. *Die Konvergenz ist gleichmäßig für*

$$|z_k - z_k^{(0)}| \leq r_k \, |z_k^{(1)} - z_k^{(0)}| \qquad \forall k = 1, \ldots, n\,, \tag{4.8.2}$$

für beliebige Zahlen $0 < r_k < 1$, *für* $k = 1, \ldots, n$.

(b) *Ist* $\alpha_k = \max\{|a_p| : \ |p| = k\}$, *für* $k \in \mathbb{N}_0$, *und hat die Potenzreihe* $\sum_{k=0}^{\infty} \alpha_k \, t^k$ *einen Konvergenzradius* $R > 0$, *so konvergiert* (4.8.1) *absolut für*

$$\|z - z_0\|_\infty = \max\{|z_1 - z_1^{(0)}|, \ldots, |z_n - z_n^{(0)}|\} \ < \ R.$$

Beweis: Zu (a): Aus der Konvergenz für $z = z_1$ folgt die Existenz von $K > 0$ mit $|a_p \, (z_1 - z_0)^p| \leq K$ für alle Multi-Indizes p. Daher gilt für alle z wie in (4.8.2)

$$|a_p \, (z - z_0)^p| \ \leq \ K \, r_1^{p_1} \cdot \ldots \cdot r_n^{p_n} \qquad \forall \, p.$$

Mit dem Majorantenkriterium und der Konvergenz der geometrischen Reihe folgt die Behauptung.

Zu (b): Sei $r = \|z - z_0\|_\infty < R$. Dann folgt $|(z - z_o)^p| \leq r^{|p|}$, also

$$\sum_p |a_p| \ |(z - z_0)^p| \ \leq \ \sum_{k=0}^{\infty} r^k \sum_{|p|=k} |a_p|.$$

Die Anzahl der Multi-Indizes vom Betrag k ist nicht größer als $(k+1)^n$, und deshalb ist die Reihe $\sum_k (k+1)^n \alpha_k \, r^k$ eine konvergente Majorante. □

Aufgabe 4.8.7 *Zeige: Ist* $p = e_k = (0, \ldots, 0, 1, 0, \ldots, 0)^T$ *der* k*-te Basisvektor der kanonischen Basis von* $\mathbb{K}^n$, *mit der* 1 *an der* k*-ten Stelle, so ist* $x^p = x_k$.

Aufgabe 4.8.8 *Zeige: Sind* p *und* q *zwei Multi-Indizes, so ist* $x^{p+q} = x^p \, x^q$.

Aufgabe 4.8.9 *Zeige: Jedes Polynom* $p \in \mathbb{K}[x]$ *vom Grade höchstens gleich* 2 *kann geschrieben werden als* $p(x) = x^T A \, x + a^T \, x + b$, *mit einer quadratischen* n*-reihigen Matrix* A, *einem Vektor* $a \in \mathbb{K}^n$ *und einer Zahl* $b \in \mathbb{K}$.

Aufgabe 4.8.10 *Zeige die Abzählbarkeit der Menge aller Multi-Indizes.*

Kapitel 5

Differenzialrechnung mehrerer Variabler

Im Folgenden sei immer ein festes nicht-leeres Gebiet $G \subset \mathbb{R}^n$ und eine Funktion $f : G \longrightarrow \mathbb{R}$ betrachtet. Wir schreiben dann auch $f(x) = f(x_1, \ldots, x_n)$.

5.1 Richtungsableitungen, partielle Ableitungen, Gradient

Definition 5.1.1 *Sei e ein beliebiger* Einheitsvektor, *d. h., $e \in \mathbb{R}^n$ und $\|e\| = 1$. Falls für ein $x \in G$ der Grenzwert*

$$f_e(x) \;=\; \lim_{t \to 0} \frac{f(x + t\,e) \;-\; f(x)}{t}$$

existiert, heißt $f_e(x)$ die Ableitung von f an der Stelle x in Richtung von e. *Jede solche Ableitung heißt auch kurz* eine Richtungsableitung von f im Punkt x. *Ist e der k-te Basisvektor e_k der kanonischen Basis von $\mathbb{R}^n$, so heißt diese Richtungsableitung auch* k-te partielle Ableitung *von f im Punkt x, oder* die Ableitung von f nach der Variablen x_k im Punkt x, *und wir schreiben auch*

$$f_{x_k}(x) \;=\; \frac{\partial f(x)}{\partial x_k} \;=\; \lim_{t \to 0} \frac{f(x_1, \ldots, x_{k-1}, x_k + t, x_{k+1}, \ldots, x_n) - f(x_1, \ldots, x_n)}{t}$$

für diese partielle Ableitung. Wenn an einer Stelle $x \in G$ alle partiellen Ableitungen existieren, so heißt f im Punkt x partiell differenzierbar, *und der Zeilenvektor*

$$\operatorname{grad} f(x) \;=\; (f_{x_1}(x), \ldots, f_{x_n}(x))$$

heißt der Gradient *von f im Punkt x.*

Bemerkung 5.1.2 *Die Berechnung einer partiellen Ableitung geschieht wie folgt: Alle Variablen werden wie Konstante behandelt mit Ausnahme einer, und nach dieser einen Variablen wird nach den Regeln aus dem ersten Semester differenziert. Wie man andere Richtungsableitungen, außer über die Definition, berechnet, wird später behandelt. Wichtig ist,* dass aus der partiellen Differenzierbarkeit, oder sogar aus der Existenz aller Richtungsableitungen, i. A. nicht die Stetigkeit folgt. *Dazu siehe das folgende Beispiel:*

Beispiel 5.1.3 *Wir definieren f wie in* Aufgabe 4.7.4. *Sei $e = (\cos\phi, \sin\phi)^T$, für irgend ein $\phi \in \mathbb{R}$. Dann ist $\|e\| = 1$, und jeder Einheitsvektor in $\mathbb{R}^2$ kann so geschrieben werden. Es gilt offenbar*

$$\lim_{t \to 0} \frac{f(t\,\cos\phi, t\,\sin\phi) \;-\; f(0,0)}{t} \;=\; \cos^2\phi \,/\, \sin\phi \qquad \forall\ \sin\phi \neq 0\,,$$

und der Grenzwert ist gleich 0 *falls* $\sin\phi = 0$ *ist. Somit existiert die Ableitung in Richtung von* e *und hat den angegebenen Wert. Also existieren alle Richtungsableitungen im Punkt* $x = 0$*. Insbesondere ist* f *im Nullpunkt partiell differenzierbar. In* Aufgabe 4.7.4 *wurde aber gezeigt, dass* f *im Nullpunkt nicht stetig ist!*

Aufgabe 5.1.4 *Sei* $p = (p_1, \ldots, p_n)$ *ein Multi-Index, und sei* $f(x) = x^p$*. Zeige, dass* f *in jedem Punkt* $x \in \mathbb{R}^n$ *partiell differenzierbar ist, und dass* $f_{x_k}(x) = p_k\, x^{p-e_k}$ *ist – auch für* $p_k = 0$*, wenn wir davon absehen, dass dann* $p - e_k$ *kein Multi-Index mehr ist.*

Aufgabe 5.1.5 *Sei* $A = [\alpha_{jk}]$ *eine reelle, nicht notwendigerweise symmetrische* (n,n)*-Matrix, und sei* $f(x) = x^T A\, x$*. Zeige, dass* f *in jedem Punkt* $x \in \mathbb{R}^n$ *partiell differenzierbar ist, und dass* $f_{x_k}(x) = \sum_{j=1}^n (\alpha_{jk} + \alpha_{kj})\, x_j$ *ist. Versuche, den Gradienten von* f *auf möglichst einfache Weise zu schreiben.*

Definition 5.1.6 *Falls* f *in jedem Punkt von* G *partiell differenzierbar ist, und falls alle partiellen Ableitungen auf* G *stetig sind, sagen wir kurz, dass* f auf G stetig partiell differenzierbar ist*.*

Falls es zu jedem $x_0 \in G$ *einen Radius* $r > 0$ *mit* $K(x_0, r) \subset G$ *so gibt, dass* f *auf* $K(x_0, r)$ *einer Lipschitzbedingung genügt, so sagen wir dass* f *auf* G lokal eine Lipschitzbedingung *erfüllt.*

Lemma 5.1.7 *Falls* f *auf* G *stetig partiell differenzierbar ist, dann erfüllt* f *auf* G *lokal eine Lipschitzbedingung und ist somit dort stetig.*

(Ohne Beweis)

Lemma 5.1.8 (Vertauschen von Integral und Ableitung) *Sei* $n = 2$*, sei* f *auf* G *nach* x_2 *partiell differenzierbar, und seien* f *und* f_{x_2} *auf* G *stetig. Seien weiter* $a < b$ *und* $c < d$ *so, dass* $[a,b] \times [c,d] \subset G$ *ist, und sei*

$$g(x) \;=\; \int_a^b f(t,x)\,dt \qquad \forall x \in [c,d].$$

Dann ist g *auf* $[c,d]$ *stetig differenzierbar, und es gilt*

$$g'(x) \;=\; \int_a^b f_{x_2}(t,x)\,dt \qquad \forall x \in [c,d].$$

(Ohne Beweis)

Bemerkung 5.1.9 *Beachte, dass obiges Lemma sinngemäß auch dann gilt, wenn die Variable* x *ein Vektor ist; in diesem Fall kann man ja alle* x_j *bis auf eines festhalten und nach der verbleibenden Variablen partiell differenzieren.*

Definition 5.1.10 *Sei* $\alpha \in \mathbb{N}_0$*. Die durch das Integral*

$$J_\alpha(x) \;=\; \frac{1}{\pi}\int_0^\pi \cos(x\,\sin t - \alpha\, t)\,dt \qquad \forall\; x \in \mathbb{R}$$

definierte Funktion heißt Besselsche Funktion *vom Index* α*. Mit dem obigen Lemma folgt, dass diese Funktion beliebig oft differenzierbar ist.*

Aufgabe 5.1.11 *Zeige, dass die Besselfunktion vom Index* $\alpha \in \mathbb{N}_0$ *eine Lösung der Differenzialgleichung*

$$x^2\, y'' \;+\; x\, y' \;+\; (x^2 - \alpha^2)\, y \;=\; 0$$

ist. Diese Gleichung heißt auch Besselsche Differenzialgleichung.

5.2 Vertauschen der Differenziationsreihenfolge

Definition 5.2.1 *Wenn f auf G nach x_k partiell differenzierbar ist, und wenn f_{x_k} in einem Punkt x nach x_j partiell differenzierbar ist, dann heißt diese Ableitung auch* die zweite partielle Ableitung von f nach x_k und x_j im Punkt x, *und wir schreiben auch hierfür*

$$f_{x_k x_j}(x) \;=\; \frac{\partial^2 f(x)}{\partial x_j\, \partial x_k} \;=\; \frac{\partial^2}{\partial x_j\, \partial x_k}\, f(x)\,.$$

Entsprechend werden höhere als zweite partielle Ableitungen definiert. Im Allgemeinen muss man hier auf die Reihenfolge der Differenziationen *achten!*

Beispiel 5.2.2 *Für $x_1, x_2 \in \mathbb{R}$ sei*

$$f(x_1,x_2) \;=\; \begin{cases} \dfrac{x_1\, x_2^3}{x_1^2+x_2^2} & \text{für } (x_1,x_2) \neq (0,0),\\[2mm] 0 & \text{für } (x_1,x_2) = (0,0). \end{cases}$$

Dann sind

$$f_{x_1}(x_1,x_2) \;=\; \begin{cases} \dfrac{(x_2^2-x_1^2)\, x_2^3}{(x_1^2+x_2^2)^2} & \text{für } (x_1,x_2) \neq (0,0),\\[2mm] 0 & \text{für } (x_1,x_2) = (0,0), \end{cases}$$

$$f_{x_2}(x_1,x_2) \;=\; \begin{cases} \dfrac{(3\, x_1^2+x_2^2)\, x_1\, x_2^2}{(x_1^2+x_2^2)^2} & \text{für } (x_1,x_2) \neq (0,0),\\[2mm] 0 & \text{für } (x_1,x_2) = (0,0). \end{cases}$$

Daraus ergibt sich $f_{x_1 x_2}(0,0) = 1$, $f_{x_2 x_1}(0,0) = 0$.

Satz 5.2.3 (Satz von Schwarz) *Seien $j, k \in \{1,\dots,n\}$, mit $j \neq k$, gegeben. Sei f in G nach x_j und x_k partiell differenzierbar, und seien f und diese beiden partiellen Ableitungen in G stetig. Ferner existiere $f_{x_j x_k}$ auf ganz G und sei stetig in einem Punkt $x_0 \in G$. Dann existiert auch $f_{x_k x_j}$ im Punkt x_0, und es gilt*

$$f_{x_j x_k}(x_0) \;=\; f_{x_k x_j}(x_0)\,.$$

(Ohne Beweis)

Aufgabe 5.2.4 *Überprüfe die Behauptungen in* Beispiel 5.2.2.

Aufgabe 5.2.5 *Berechne alle zweiten partiellen Ableitungen von*

$$f(x_1,x_2) = x_1^3 \cos(x_1^2 + x_2).$$

Aufgabe 5.2.6 *Berechne alle zweiten partiellen Ableitungen von f aus* Aufgabe 5.1.5.

5.3 Totale Differenzierbarkeit

Definition 5.3.1 *Wir nennen f* im Punkt $x_0 \in G$ total differenzierbar, *wenn es eine lineare Abbildung $\ell : \mathbb{R}^n \longrightarrow \mathbb{R}$ gibt, sodass für alle genügend kleinen $h \in \mathbb{R}^n$, für welche dann insbesondere $x_0 + h \in G$ liegt, gilt*

$$f(x_0+h) \;=\; f(x_0) \;+\; \ell(h) \;+\; \|h\|\, r(h), \qquad r(h) \longrightarrow 0 \quad (h \to 0).$$

Falls dies für alle $x_0 \in G$ gilt, heißt f kurz auf G total differenzierbar*, oder einfach* auf G differenzierbar. *Jede lineare Abbildung $\ell : \mathbb{R}^n \longrightarrow \mathbb{R}$ ist nach Resultaten aus der linearen Algebra von der Form $\ell(h) = a\,h$ mit einem Zeilenvektor $a = (a_1, \ldots, a_n)$, d. h., einer $(1, n)$-Matrix. Diese Matrix heißt dann* die Ableitung *oder* das totale Differenzial *von f im Punkt x_0, und wir schreiben $f'(x_0)$ an Stelle von a. Der Graph von $g(x) = f'(x_0)\,(x - x_0) + f(x_0)$ ist eine lineare Mannigfaltigkeit in $\mathbb{R}^{n+1}$ und heißt, etwas ungenau,* die Tangentialhyperebene *zu f im Punkt x_0.*

Bemerkung 5.3.2 *Man sieht leicht mit obiger Definition, dass die Summe zweier total differenzierbarer Funktionen selbst wieder total differenzierbar ist.*

Beispiel 5.3.3 *Ein Polynom $p \in \mathbb{R}[x]$ höchstens ersten Grades kann immer geschrieben werden als $p(x) = a\,x + b$, mit einem Zeilenvektor a und einer reellen Zahl b. Beachtet man, dass $p(x_0 + h) = p(x_0) + a\,h$ gilt, so folgt direkt aus der Definition, mit $r(h)$ gleich der Nullfunktion, dass p an jeder Stelle von $\mathbb{R}^n$ total differenzierbar ist, und dass $f'(x_0) = a$ ist, unabhängig von x_0.*

Satz 5.3.4 (Berechnung der Ableitung) *Sei f in $x_0 \in G$ total differenzierbar. Dann ist f dort auch stetig und partiell differenzierbar, und es gilt*

$$f'(x_0) \;=\; \operatorname{grad} f(x_0)\,.$$

Ferner existiert auch für jeden Einheitsvektor e die Ableitung von f in Richtung von e, und es ist

$$f_e(x_0) \;=\; \operatorname{grad} f(x_0)\,e.$$

Beweis: Die Stetigkeit folgt direkt aus der Definition der totalen Differenzierbarkeit. Setzt man dort $h = t\,e$ mit kleinem $t \in \mathbb{R}$ ein, so folgt wegen $\|h\| = |t|$

$$\frac{f(x_0 + t\,e) \;-\; f(x_0)}{t} \;=\; f'(x_0)\,e \pm r(t\,e) \;\longrightarrow\; f'(x_0)\,e \qquad (t \to 0),$$

und deshalb existiert die Ableitung von f in Richtung von e und ist gleich $f'(x_0)\,e$. Für $e = e_k$ folgt, dass f im Punkt x_0 nach x_k partiell differenzierbar ist, und $f_{x_k}(x_0) = f'(x_0)\,e_k$. Also ist $\operatorname{grad} f(x_0) = f'(x_0)$. □

Bemerkung 5.3.5 *Wir haben gezeigt, dass die Richtungsableitung einer total differenzierbaren Funktion das innere Produkt aus dem Gradienten und dem Einheitsvektor e ist. Dieses innere Produkt ist am größten, wenn e die gleiche Richtung wie der Gradient von f hat. Daraus ergibt sich folgende Interpretation des Gradienten:*

Der Gradient einer total differenzierbaren Funktion f zeigt in Richtung des stärksten Anstieges der Funktionswerte von f.

Die Definition der totalen Differenzierbarkeit ist nicht leicht zu überprüfen. Der folgende Satz gibt aber ein bequemes Kriterium für die Differenzierbarkeit auf dem Gebiet G:

Satz 5.3.6 *Sei f auf G stetig partiell differenzierbar. Dann ist f dort auch total differenzierbar, und die Ableitung f' ist dann stetig auf G.*

(Ohne Beweis)

Definition 5.3.7 *Sei jetzt $f = (f_1, \dots, f_m) : G \longrightarrow \mathbb{R}^m$ eine vektorwertige Funktion auf G. Dann nennen wir f in einem Punkt $x_0 \in G$ oder auf G total oder nach x_k partiell differenzierbar, wenn dasselbe für jede der Koordinatenfunktionen f_j zutrifft. Wir setzen dann*

$$f_{x_k}(x_0) \;=\; \frac{\partial f}{\partial x_k}(x_0) \;=\; \left(\frac{\partial f_1}{\partial x_k}(x_0), \dots, \frac{\partial f_m}{\partial x_k}(x_0)\right)^T .$$

Eine partielle Ableitung ist also jetzt ein Spaltenvektor, und entsprechend ist die Ableitung von f gleich der Matrix

$$f'(x_0) \;=\; \frac{\partial f}{\partial x}(x_0) \;=\; \begin{bmatrix} \frac{\partial f_1}{\partial x_1}(x_0) & \frac{\partial f_1}{\partial x_2}(x_0) & \dots & \frac{\partial f_1}{\partial x_n}(x_0) \\ \frac{\partial f_2}{\partial x_1}(x_0) & \frac{\partial f_2}{\partial x_2}(x_0) & \dots & \frac{\partial f_2}{\partial x_n}(x_0) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1}(x_0) & \frac{\partial f_m}{\partial x_2}(x_0) & \dots & \frac{\partial f_m}{\partial x_n}(x_0) \end{bmatrix}$$

Man nennt diese Matrix auch Funktionalmatrix *oder* Jacobi-Matrix *von f. Für $m = n$ ist die Funktionalmatrix quadratisch, und ihre Determinante heißt* Funktionaldeterminante. *Außer für $m = 1$ ist es nicht üblich, vom Gradienten von f zu sprechen.*

Aufgabe 5.3.8 *Für eine n-reihige quadratische Matrix A sei $f(x) = x^T A\, x$ die zugehörige quadratische Form. Zeige, dass dann für $x_0, h \in \mathbb{R}^n$ gilt*

$$f(x_0 + h) \;=\; f(x_0) \;+\; x_0^T (A + A^T)\, h \;+\; \|h\|\, e^T(h)\, A\, h,$$

mit einem Einheitsvektor $e(h)$, nämlich $e(h) = \|h\|^{-1} h$. Leite daraus ohne Benutzung von Satz 5.3.6 *direkt aus der Definition her, dass f im Punkt x_0 total differenzierbar ist, und dass $f'(x_0) = x_0^T (A + A^T)$ ist.*

Aufgabe 5.3.9 *Zeige, dass Polynome und rationale Funktionen auf ihrem natürlichen Definitionsbereich überall total differenzierbar sind.*

Aufgabe 5.3.10 *Zeige, dass* affine Abbildungen *$f(x) = A\, x + b$, mit einer reellen (m, n)-Matrix A und einem $b \in \mathbb{R}^m$, überall total differenzierbar sind, und dass $f'(x) = A$ ist, für alle $x \in \mathbb{R}^n$.*

5.4 Die Kettenregel

Satz 5.4.1 (Kettenregel) *Seien $f : G \longrightarrow G_1 \subset \mathbb{R}^m$ und $g : G_1 \longrightarrow \mathbb{R}^p$ gegeben, und seien f in $x_0 \in G$ und g in $y_0 = f(x_0) \in G_1$ total differenzierbar. Dann ist $\ell = g \circ f$ in x_0 total differenzierbar, und es gilt*

$$\ell'(x_0) \;=\; g'(f(x_0))\; f'(x_0)\,.$$

(Ohne Beweis)

Korollar zu Satz 5.4.1 *Sei (a, b) ein offenes Intervall in $\mathbb{R}$ sowie $G \subset \mathbb{R}^n$ ein Gebiet, seien $f_k : (a, b) \to \mathbb{R}$ für $k = 1, \dots, n$, mit $f(t) = (f_1(t), \dots, f_n(t))^T \in G$ für alle $t \in (a, b)$, und sei $g : G \longrightarrow \mathbb{R}$. Falls dann alle f_k an einer Stelle $t_0 \in (a, b)$ differenzierbar sind, und falls g im Punkt $x_0 = (f_1(t_0), \dots, f_n(t_0))^T$ total differenzierbar ist, dann ist $g \circ f$ in t_0 differenzierbar, und es gilt*

$$(g \circ f)'(t_0) \;=\; \operatorname{grad}\, g(f(t_0))\, f'(t_0) \;=\; \sum_{k=1}^{n} g_{x_k}(f_1(t_0), \dots, f_n(t_0))\, f_k'(t_0)\,.$$

Beweis: Folgt direkt aus der Kettenregel! □

Als eine Anwendung zeigen wir:

Behauptung 5.4.2 *Seien $a < b$ und $c < d$, sei $f(x_1, x_2)$ auf $(a,b) \times (c,d)$ nach x_2 partiell differenzierbar, und seien f und f_{x_2} auf $(a,b) \times (c,d)$ stetig. Seien weiter $g_1, g_2 : (c,d) \longrightarrow (a,b)$ stetig differenzierbar, und sei schließlich*

$$g(x) \;=\; \int_{g_1(x)}^{g_2(x)} f(t,x)\,dt \qquad \forall\; x \in (c,d).$$

Dann ist g auf (c,d) differenzierbar, und

$$g'(x) \;=\; f(g_2(x),x)\,g_2'(x) \;-\; f(g_1(x),x)\,g_1'(x) \;+\; \int_{g_1(x)}^{g_2(x)} f_{x_2}(t,x)\,dt$$

für alle $x \in (c,d)$.

Beweis: Sei $F(\xi,\eta,\zeta) = \int_\xi^\eta f(t,\zeta)\,dt$, für $\xi,\eta \in (a,b)$, $\zeta \in (c,d)$. Dann ist F nach allen Variablen partiell differenzierbar, und

$$F_\xi(\xi,\eta,\zeta) \;=\; -f(\xi,\zeta), \quad F_\eta(\xi,\eta,\zeta) \;=\; f(\eta,\zeta),$$

bzw. (wegen Lemma 5.1.8)

$$F_\zeta(\xi,\eta,\zeta) \;=\; \int_\xi^\eta f_{x_2}(t,\zeta)\,dt.$$

Da alle partiellen Ableitungen stetig sind, ist F auch total differenzierbar. Durch Anwendung der Kettenregel auf $g(x) = F(g_1(x), g_2(x), x)$ folgt dann die Behauptung. □

Aufgabe 5.4.3 *Sei $G \subset \mathbb{R}^2$ ein Gebiet, und sei $F = F(x,y)$ auf G total differenzierbar. Sei weiter $f(x)$ eine auf (a,b) differenzierbare Funktion, für die $(x, f(x))^T \in G$ und $F(x, f(x))$ konstant ist für alle $x \in (a,b)$. Zeige: Ist $F_y(x,y) \neq 0$ auf G, so ist*

$$f'(x) \;=\; -\,\frac{F_x(x,f(x))}{F_y(x,f(x))} \quad \forall\; x \in (a,b)\,.$$

Aufgabe 5.4.4 *Verallgemeinere obige Aufgabe auf den Fall $x \in \mathbb{R}^n$, $y \in \mathbb{R}^m$.*

5.5 Der Mittelwertsatz und der Satz von Taylor

Satz 5.5.1 (Mittelwertsatz) *Sei $f : G \longrightarrow \mathbb{R}$ total differenzierbar auf G, und seien $a, b \in G$ so, dass die Verbindungsstrecke von a nach b ganz zu G gehört. Dann gibt es ein ξ auf dieser Verbindungsstrecke mit*

$$f(b) \;-\; f(a) \;=\; f'(\xi)\;(b \;-\; a)\,.$$

Beweis: Setze $g(t) = f(t\,b + (1-t)\,a)$, $0 \le t \le 1$, und wende den ersten Mittelwertsatz der Differentialrechnung sowie die Kettenregel an. □

Bei vektorwertigen Funktionen gilt im Allgemeinen kein Mittelwertsatz. Als Ersatz zeigen wir aber, dass stetig differenzierbare Funktionen mit beschränkter Ableitung eine Lipschitzbedingung erfüllen:

Satz 5.5.2 (Lipschitzbedingung für stetig differenzierbare Funktionen)
Sei G konvex, sei $f : G \longrightarrow \mathbb{R}^m$ stetig partiell differenzierbar auf G, und sei $L = \sup\{\|f'(x)\| : x \in G\} < \infty$. Dann gilt

$$\|f(x_1) \; - \; f(x_2)\| \;\le\; L\,\|x_1 \; - \; x_2\| \qquad \forall\, x_1, x_2 \in G\,.$$

Beweis: Seien $x_1, x_2 \in G$, und sei $g(t) = \|f(t\,x_2 + (1-t)\,x_1) - f(x_1)\|$, für $0 \le t \le 1$. Dieses g ist stetig auf $[0,1]$ und differenzierbar in jedem Punkt t mit $g(t) > 0$. Falls $g(1) = 0$ ist, ist nichts zu zeigen. Im anderen Fall gibt es ein $t_0 \in [0,1)$ mit $g(t_0) = 0$ und $g(t) > 0$ für alle $t \in (t_0, 1]$. Dann ist der erste Mittelwertsatz der Differentialrechnung auf g und das Intervall $[t_0, 1]$ anwendbar, und wir erhalten die Existenz eines $\xi \in (t_0, 1)$ mit

$$g(1) \;=\; g(1) \;-\; g(t_0) \;=\; g'(\xi)\,(1 \;-\; t_0) \;\le g'(\xi)\,.$$

Man zeigt aber mit der Kettenregel, dass

$$g'(t) \;=\; \frac{1}{g(t)}\,(f^T(x(t)) \;-\; f^T(x_1))\,f'(x(t))\,(x_2 - x_1)\,,$$

mit $x(t) = t\,x_2 + (1-t)\,x_1$, falls nur $g(t) > 0$ ist. Daraus folgt aber

$$\|g'(t)\| \;\le\; \|f'(x(t))\|\,\|x_2 - x_1\| \;\le\; L\,\|x_2 - x_1\|\,,$$

und hieraus ergibt sich die Behauptung. □

Im nächsten Satz benutzen wir folgende Bezeichnungen:

$$\nabla \;=\; \left(\frac{\partial}{\partial x_1}, \ldots, \frac{\partial}{\partial x_n}\right)$$

bezeichnet *den Gradientenoperator*, d. h. $\nabla f = \operatorname{grad} f$. Für $h = (h_1, \ldots, h_n)^T \in \mathbb{R}^n$ sei

$$\nabla\, h \;=\; \sum_{k=1}^{n} h_k\,\frac{\partial}{\partial x_k},$$

d. h. $\nabla\, h$ ist ein *Differentialoperator*, den man auf jede partiell differenzierbare Funktion f anwenden kann, und $(\nabla\, h)\,f = \sum_{k=1}^{n} h_k\,\frac{\partial f}{\partial x_k}$. Weiter nennen wir eine Funktion *$(m+1)$-mal stetig partiell differenzierbar* auf G, wenn auf G alle ihre partiellen Ableitungen bis zur $(m+1)$-ten Ordnung existieren und stetig sind. Insbesondere kann man dann den Operator $\nabla\, h$ auf jede solche Funktion $(m+1)$-mal anwenden, und wir schreiben dann auch $(\nabla\, h)^k\, f$ für die Funktion, die man durch k-maliges Anwenden dieses Operators auf f erhält. Wir definieren noch $(\nabla\, h)^0\, f = f$.

Satz 5.5.3 (Der Satz von Taylor) *Sei $f : G \longrightarrow \mathbb{R}$ $(m+1)$-mal stetig partiell differenzierbar, und seien x und h so, dass die Verbindungsstrecke von x nach $x+h$ ganz in G liegt. Dann gilt*

$$f(x+h) \;=\; \sum_{k=0}^{m} \frac{(\nabla\, h)^k f(x)}{k!} \;+\; \int_0^1 \frac{(1-t)^m}{m!}\;(\nabla\, h)^{m+1} f(x + t\,h)\,dt.$$

Beweis: Sei $g(t) = f(x + t\,h)$, $0 \le t \le 1$. Mit dem Taylorschen Satz (Satz 1.6.2) für eine Dimension folgt dann

$$g(1) \;=\; \sum_{k=0}^{m} \frac{g^{(k)}(0)}{k!} \;+\; \int_0^1 \frac{(1-t)^m}{m!}\; g^{(m+1)}(t)\,dt.$$

Mit der Kettenregel und vollständiger Induktion zeigt man, dass $g^{(k)}(t) = (\nabla\, h)^k f(x + t\,h)$ gilt, für $k = 0, \ldots, m+1$, und daraus folgt die Behauptung. □

Definition 5.5.4 *Wir nennen die Summe* $\sum_{k=0}^{m} \frac{(\nabla h)^k f(x)}{k!}$ das Taylorpolynom *der Ordnung* m*; dieser Ausdruck ist ein Polynom vom Grade höchstens gleich* m *in den Variablen* $h_1, \ldots, h_n$*. Die Differenz* $r_m(x,h) = f(x+h) - \sum_{k=0}^{m} \frac{(\nabla h)^k f(x)}{k!}$ *heißt dann auch* das Restglied *der Ordnung* m*.*

Aufgabe 5.5.5 *Sei, zusätzlich zu den übrigen Voraussetzungen des Mittelwertsatzes, noch angenommen, dass* grad f *auf* G *stetig ist. Zeige, dass dann* f *auf* G *lokal einer Lipschitzbedingung genügt.*

Aufgabe 5.5.6 *Sei* $p = (p_1, \ldots, p_n)$ *ein Multi-Index, und sei*

$$D^p \;=\; \frac{\partial^{p_1}}{\partial x_1^{p_1}} \cdots \frac{\partial^{p_n}}{\partial x_n^{p_n}}\,,$$

d. h., $D^p f$ *erhält man aus* f *durch* p_1*-maliges partielles Differenzieren nach* x_1*,* p_2*-maliges Differenzieren nach* x_2*, u. s. w., wobei stillschweigend vorausgesetzt ist, dass die Reihenfolge des Differenzierens ohne Bedeutung ist, was nach dem Satz von Schwarz sicher gilt, wenn* f *hinreichend oft stetig partiell differenzierbar ist. Zeige:*

$$\frac{(\nabla h)^k f(x)}{k!} \;=\; \sum_{|p|=k} \frac{h^p\, D^p f(x)}{p!}\,,$$

wenn man $p! = (p_1!) \cdots (p_n!)$ *setzt.*

Kapitel 6

Implizite Funktionen, lokale Extrema

6.1 Der Banachsche Fixpunktsatz

Definition 6.1.1 *Sei X ein normierter Raum, sei $D \subset X$ nicht leer, und sei $\phi : D \longrightarrow X$. Wir nennen ϕ* eine Kontraktion *auf D, falls ϕ einer Lipschitzbedingung mit einer Lipschitzkonstanten $L < 1$ genügt. Ein solches L heißt dann auch* Kontraktionsparameter. *Ein $x \in D$ heißt* ein Fixpunkt von ϕ, *falls $\phi(x) = x$ ist.*

Satz 6.1.2 (Banachscher Fixpunktsatz) *Sei X ein normierter Raum, sei $D \subset X$ vollständig, und sei $\phi : D \longrightarrow D$ eine Kontraktion auf D. Dann hat ϕ genau einen Fixpunkt $\xi \in D$. Weiter gilt: Ist $L\,(< 1)$ ein Kontraktionsparameter von ϕ, ist $x_0 \in D$, und setzt man $x_{m+1} = \phi(x_m)$ für $m \in \mathbb{N}_0$, so folgt*

$$\|x_m \; - \; \xi\| \;\leq \frac{1}{1-L}\;\|x_{m+1} \; - \; x_m\| \;\leq\; \frac{L^m}{1-L}\;\|x_1 \; - \; x_0\|\;,$$

und daher gilt $\lim x_m = \xi$.

Beweis: Es ist auf Grund der Definition der (x_m)

$$\|x_{m+1} \; - \; x_m\| \;=\; \|\phi(x_m) \; - \; \phi(x_{m-1})\| \;\leq\; L\,\|x_m \; - \; x_{m-1}\|\;,$$

und daher folgt mit Induktion über m:

$$\|x_{m+1} \; - \; x_m\| \;\leq\; L^m\,\|x_1 \; - \; x_0\| \qquad \forall\, m \in \mathbb{N}_0\,.$$

Für $p \in \mathbb{N}_0$ ist $x_{m+p+1} - x_m = \sum_{k=m}^{m+p} (x_{k+1} - x_k)$, und daraus folgt

$$\|x_{m+p+1} - x_m\| \leq \sum_{k=m}^{m+p} \|x_{k+1} - x_k\| \leq \|x_1 - x_0\| \sum_{k=m}^{\infty} L^k \;=\frac{\|x_1 - x_0\|\,L^m}{1-L}\;.$$

Da $L < 1$ ist, folgt hieraus, dass (x_m) eine Cauchy-Folge ist, und wegen der Vollständigkeit von D existiert $\xi = \lim_{m\to\infty} x_m \in D$. Da jede Kontraktion auch stetig ist, folgt hieraus $\phi(\xi) = \lim_{m\to\infty} \phi(x_m) = \lim_{m\to\infty} x_{m+1} = \xi$. Also ist ξ ein Fixpunkt von ϕ. Ist $\tilde{\xi}$ ebenfalls ein Fixpunkt, so gilt $\|\tilde{\xi} - \xi\| = \|\phi(\tilde{\xi}) - \phi(\xi)\| \leq L\,\|\tilde{\xi} - \xi\|$, was nur für $\|\tilde{\xi} - \xi\| = 0$ richtig ist. Daher ist also ξ der einzige Fixpunkt von f. Weiter folgt

$$\begin{aligned} \|x_m - \xi\| &\leq \|x_m - x_{m+1}\| \;+\; \|x_{m+1} - \xi\| \\ &= \|x_m - x_{m+1}\| \;+\; \|\phi(x_m) - \phi(\xi)\| \\ &\leq \|x_m - x_{m+1}\| \;+\; L\,\|x_m - \xi\|\;, \end{aligned}$$

woraus folgt

$$(1-L)\,\|x_m-\xi\| \;\le\; \|x_m-x_{m+1}\| \;\le\; L^m\,\|x_1\;-\;x_0\| \qquad \forall\, m\in\mathbb{N}_0\,.$$

Das ist alles, was noch zu zeigen war. □

Bemerkung 6.1.3 *Der für uns interessanteste Fall des Banachschen Fixpunktsatzes betrifft den Raum $X=\mathbb{R}^n$. Satz 5.5.2 zeigt, dass eine Funktion ϕ dann eine Kontraktion auf einer konvexen offenen Menge G ist, wenn sie dort stetig partiell differenzierbar und das Supremum ihrer Ableitung echt kleiner als 1 ist. Ist $D=\overline{G}$, und ist ϕ auf D stetig, so sieht man leicht, dass ϕ auch auf D eine Kontraktion (mit dem gleichen Kontraktionsparameter wie auf G) ist. Ausserdem ist D als abgeschlossene Teilmenge eines vollständigen Raumes selber vollständig. Es ist aber oft nicht leicht zu zeigen, dass ϕ die Menge D in sich abbildet. Deshalb halten wir fest:*

Ist $D=\overline{K(a,r)}$, mit einem $a\in\mathbb{R}^n$ und $r>0$, so ist D abgeschlossen und somit auch vollständig. Ist $\phi: D\longrightarrow X$ eine Kontraktion mit Kontraktionsparameter L, und gilt $\|\phi(a)-a\|\le(1-L)\,r$, so folgt für $x\in D$ (also $\|x-a\|\le r$), dass $\|\phi(x)-a\|\le\|\phi(x)-\phi(a)\|+\|\phi(a)-a\|\le L\,\|x-a\|+(1-L)\,r\le r$. Deshalb folgt $\phi: D\longrightarrow D$, so dass der Banachsche Fixpunktsatz anwendbar ist.

In dieser Form lassen sich die Voraussetzungen des Banachschen Fixpunktsatzes meist am leichtesten nachprüfen.

Aufgabe 6.1.4 (Newton-Interation) *Zeige: Ist $f:(a,b)\longrightarrow\mathbb{R}$ zweimal stetig differenzierbar auf dem offenen Intervall (a,b), und ist $f(x_0)=0$ für ein $x_0\in(a,b)$, aber $f'(x_0)\ne 0$, so ist die Funktion $\phi(x)=x-f(x)/f'(x)$ auf einem kleinen Intervall um x_0 herum kontrahierend.*

Aufgabe 6.1.5 *Seien f, x_0 und ϕ wie oben. Zeige die Existenz eines $\varepsilon>0$ derart, dass ϕ das Intervall $(x_0-\varepsilon,x_0+\varepsilon)$ in sich abbildet.*

6.2 Das Newton-Verfahren

Wir wollen nun die mehrdimensionale Version des Newton-Verfahrens zur Berechnung von Lösungen nicht-linearer Gleichungssysteme kennenlernen: Dazu sei $D\subset\mathbb{R}^n$ nicht leer und $f=(f_1,\ldots,f_n)^T: D\longrightarrow\mathbb{R}^n$. Die Gleichung $f(x)=0$ ist offenbar eine bequeme Schreibweise für das System von n (im Allgemeinen nicht linearen) Gleichungen

$$f_j(x_1,\ldots,x_n) \;=\; 0\,, \qquad 1\le j\le n, \tag{6.2.1}$$

in den *Unbekannten* $x_1,,\ldots,x_n$. Jede Lösung der Gleichung $f(x)=0$ heißt dann auch *eine Nullstelle von* f (in D).

Um den Banachschen Fixpunktsatz anwenden zu können, muss man eine andere Funktion ϕ finden, deren Fixpunkte gerade die Nullstellen von f sind. Die beim Newton-Verfahren verwendete *Schrittfunktion* ist

$$\phi(x) \;=\; x \;-\; f'(x)^{-1}f(x) \qquad (x\in D). \tag{6.2.2}$$

Dabei muss vorausgesetzt werden, dass D eine offene Menge ist, auf welcher f stetig partiell differenzierbar ist, und dass die Ableitung $f'(x)$ (welche eine n-reihige quadratische Matrix ist) immer invertierbar ist. Wenn f sogar zweimal stetig partiell differenzierbar auf D ist, dann ist ϕ dort wenigstens einmal stetig partiell differenzierbar, und aus $f(\xi)=0$ folgt $\phi'(\xi)=0$. Daher ist ϕ auf einer genügend kleinen Kugel um ξ kontrahierend und bildet diese Kugel in sich ab. Aus dem Banachschen Fixpunktsatz folgt dann die *lokale Konvergenz* des Newton-Verfahrens, d. h. genauer: *Ist x_0 nahe genug bei ξ, und ist $x_{n+1}=\phi(x_n)=x_n-f'(x_n)^{-1}f(x_n)$ für $n\in\mathbb{N}_0$, so konvergiert (x_n) gegen ξ.* Allerdings ist die Berechnung von $f'(x_n)^{-1}$ in jedem Iterationsschritt recht aufwendig, und deshalb benutzt man oft das sogenannte *vereinfachte Newton-Verfahren* mit der Schrittfunktion

$$\phi(x) \;=\; x \;-\; A\,f(x) \qquad (x\in D),$$

wobei A eine feste quadratische Matrix ist. Natürlich wird man die Matrix A geeignet wählen müssen, damit ϕ kontrahierend ist. Im Idealfall wäre $A = f'(\xi)^{-1}$ zu setzen, denn dann ist $\phi'(\xi) = 0$. Allerdings ist uns ξ ja nicht bekannt, sondern soll ja gerade erst berechnet werden.

Aufgabe 6.2.1 *Sei D eine offene Menge, und sei $f : D \longrightarrow \mathbb{R}^n$ zweimal stetig partiell differenzierbar. Weiter sei $\xi \in D$ so, dass $f(\xi) = 0$, aber $\det f'(\xi) \neq 0$ ist. Sei schließlich ϕ wie in* (6.2.2)*. Zeige, dass dann $\phi'(\xi) = 0$ ist, und folgere hieraus die Existenz eines $r > 0$, für welches ϕ die abgeschlossene Kugel um ξ mit Radius r in sich abbildet und dort kontrahierend ist.*

Aufgabe 6.2.2 *Sei D eine offene Menge, und sei $f : D \longrightarrow \mathbb{R}^n$ stetig partiell differenzierbar. Weiter sei $\xi \in D$ so, dass $f(\xi) = 0$, aber $\det f'(\xi) \neq 0$ ist, und sei A eine Matrix mit $\|I - A\, f'(\xi)\| < 1$. Sei schließlich ϕ die Schrittfunktion des vereinfachten Newtonverfahrens mit obiger Matrix A. Zeige, dass dann $\|\phi'(\xi)\| < 1$ ist, und folgere hieraus die Existenz eines $r > 0$, für welches ϕ die abgeschlossene Kugel um ξ mit Radius r in sich abbildet und dort kontrahierend ist.*

6.3 Implizite Funktionen

Man kann das Newton-Verfahren verwenden, um die Lösbarkeit eines allgemeinen nichtlinearen Gleichungssystems wie (6.2.1), aber mit weniger Gleichungen als Unbekannten, zu untersuchen. Dazu verwenden wir folgende Bezeichnungen: Für $n, m \in \mathbb{N}$ sei $G \subset \mathbb{R}^{n+m}$ ein Gebiet, und $f = (f_1, \ldots, f_m)^T : G \longrightarrow \mathbb{R}^m$ eine dort definierte Funktion. Dann ist eine Nullstelle von f genau die Lösung eines Systems von m Gleichungen in $n + m$ Unbekannten. Also liegt die Vermutung nahe, dass man für gewisse der Unbekannten (n an der Zahl) beliebige Werte einsetzen kann und die Gleichungen nach den übrigen m Unbekannten auflösen kann. Allerdings können die Gleichungen auch abhängig sein; um dies auszuschließen, definieren wir:

Definition 6.3.1 *Mit obigen Bezeichnungen nennen wir die Gleichung $f(x) = 0$* wohlgestellt*, wenn die Matrix $f'(x)$ für jedes $x \in G$ den Rang m hat.*

Nach welchen der Unbekannten man auflösen kann, ist im Allgemeinen nicht klar. Durch eine geeignete Umnumerierung der Unbekannten können wir uns aber auf den Fall beschränken, dass wir nach den letzten m Unbekannten auflösen wollen. Um diese besser von den anderen zu unterscheiden, fassen wir einen Vektor aus $\mathbb{R}^{n+m}$ hier als ein Paar (x, y) mit $x = (x_1, \ldots, x_n)^T \in \mathbb{R}^n$, $y = (y_1, \ldots, y_m)^T \in \mathbb{R}^m$ auf. Die Aufgabe lautet also, die Gleichungen

$$f_j(x_1, \ldots, x_n, y_1, \ldots, y_m) \;=\; 0\,, \qquad 1 \le j \le m, \tag{6.3.1}$$

bei gegebenen Werten für $x_1, \ldots, x_n$, nach den Unbekannten $y_1, \ldots, y_m$ aufzulösen.

Falls f partiell differenzierbar ist, schreiben wir

$$\frac{\partial f}{\partial x}(x,y) \;=\; f_x(x,y) \;=\; \begin{bmatrix} \frac{\partial f_1}{\partial x_1}(x,y) & \ldots & \frac{\partial f_1}{\partial x_n}(x,y) \\ \vdots & & \vdots \\ \frac{\partial f_m}{\partial x_1}(x,y) & \ldots & \frac{\partial f_m}{\partial x_n}(x,y) \end{bmatrix},$$

und analog

$$\frac{\partial f}{\partial y}(x,y) \;=\; f_y(x,y) \;=\; \begin{bmatrix} \frac{\partial f_1}{\partial y_1}(x,y) & \ldots & \frac{\partial f_1}{\partial y_m}(x,y) \\ \vdots & & \vdots \\ \frac{\partial f_m}{\partial y_1}(x,y) & \ldots & \frac{\partial f_m}{\partial y_m}(x,y) \end{bmatrix}.$$

Beachte, dass f_y eine quadratische Matrix ist!

In Vektorschreibweise ist (6.3.1) äquivalent zur Gleichung $f(x, y) = 0$. Es kann vorkommen, dass diese Gleichung für jedes x unlösbar ist. Es gilt aber nach dem nächsten Satz: Wenn ein x_0 existiert, für welches die Gleichung lösbar ist, und wenn eine zusätzliche Bedingung an f erfüllt ist, dann ist die Gleichung für alle x in der Nähe von x_0 eindeutig nach y auflösbar. Die Auflösung ergibt dann eine Funktion $y = g(x)$, welche wir eine *implizite Funktion* nennen wollen.

Satz 6.3.2 (Hauptsatz über implizite Funktionen) *Für $n, m \in \mathbb{N}$ sei G ein Gebiet in $\mathbb{R}^{n+m}$, und $f : G \longrightarrow \mathbb{R}^m$ sei stetig partiell differenzierbar auf G. Für ein $(\xi, \eta)^T \in G$ gelte $f(\xi, \eta) = 0$, und die Matrix*

$$\frac{\partial f}{\partial y}(x, y) = \begin{bmatrix} \frac{\partial f_1}{\partial y_1}(x, y) & \ldots & \frac{\partial f_1}{\partial y_m}(x, y) \\ \vdots & & \vdots \\ \frac{\partial f_m}{\partial y_1}(x, y) & \ldots & \frac{\partial f_m}{\partial y_m}(x, y) \end{bmatrix}$$

sei für $(x, y) = (\xi, \eta)$ invertierbar. Dann gibt es $r_1, r_2 > 0$ derart, dass eine eindeutig bestimmte Funktion $g : K(\xi, r_1) \longrightarrow K(\eta, r_2)$ existiert mit

$$\forall\ x \in K(\xi, r_1)\,,\ y \in K(\eta, r_2) : \qquad f(x, y) = 0 \quad \Longleftrightarrow \quad y = g(x)\ .$$

Weiter gilt: Die Funktion g ist stetig partiell differenzierbar auf $K(\xi, r_1)$, und

$$g'(x) = -\left(\frac{\partial f}{\partial y}(x, g(x))\right)^{-1} \frac{\partial f}{\partial x}(x, g(x)) \quad \forall\ x \in K(\xi, r_1)\,. \tag{6.3.2}$$

(Ohne Beweis)

6.4 Differenzierbarkeit der Umkehrfunktion

Satz 6.4.1 (Existenz und Ableitung der Umkehrfunktion) *Sei $G \subset \mathbb{R}^n$ ein Gebiet, und sei $f = (f_1, \ldots, f_n)^T : G \longrightarrow \mathbb{R}^n$ stetig partiell differenzierbar auf G. Sei weiter für ein $\eta \in G$*

$$\det f'(\eta) = \det \begin{bmatrix} \frac{\partial f_1}{\partial x_1}(\eta) & \ldots & \frac{\partial f_1}{\partial x_n}(\eta) \\ \vdots & & \vdots \\ \frac{\partial f_n}{\partial x_1}(\eta) & \ldots & \frac{\partial f_n}{\partial x_n}(\eta) \end{bmatrix} \neq 0\,.$$

Dann gibt es Gebiete G_1, G_2 mit $\eta \in G_1 \subset G$, $f(\eta) \in G_2 \subset \mathbb{R}^n$ derart, dass $f : G_1 \longrightarrow G_2$ bijektiv und $\det f'(x) \neq 0$ für alle $x \in G$ ist. Die Umkehrfunktion ist dann auf G_2 differenzierbar, und es gilt

$$\frac{\partial f^{-1}}{\partial x}(x) = \big(f'(f^{-1}(x))\big)^{-1} \quad \forall\ x \in G_2\,.$$

Beweis: Eine Anwendung des Hauptsatzes über implizite Funktionen auf $f(x, y) = f(y) - x$ und $\xi = f(\eta)$ liefert die Existenz von $r_1, r_2 > 0$, so dass

$$\forall\ x \in K(\xi, r_1) \quad \exists_1\ y \in K(\eta, r_2) : \quad x = f(y)\,.$$

Aus Proposition 4.3.4 folgt, dass die Urbildmenge von $K(\xi, r_1)$ eine offene Teilmenge $O \subset G$ ist, und wir setzen $G_1 = O \cap K(\eta, r_2)$. Dann ist $f : G_1 \longrightarrow G_2 = K(\xi, r_1)$ bijektiv. Aus dem oben schon angewandten Hauptsatz folgt aber auch die stetige partielle Differenzierbarkeit der Umkehrfunktion (also auch deren Stetigkeit), und da G_2 zusammenhängend ist, folgt mit Satz 4.4.1 der Zusammenhang von G_1. Die behauptete Aussage über die Ableitung von f^{-1} folgt dann aus (6.3.2). □

Definition 6.4.2 *Sei $G \subset \mathbb{R}^n$ ein Gebiet, und sei $f : G \longrightarrow \mathbb{R}^n$. Wir nennen f* lokal umkehrbar *oder* lokal injektiv*, wenn es zu jedem $x \in G$ ein $r > 0$ gibt, für welches die Restriktion von f auf $K(x, r)$ injektiv ist. Falls f auf G injektiv ist, nennen wir f manchmal auch* global umkehrbar*.*

Der obige Satz zeigt, dass eine stetig partiell differenzierbare Funktion, deren Funktionaldeterminante nie verschwindet, lokal umkehrbar ist. In Aufgabe 6.4.4 wird ein Beispiel dafür gegeben, dass ein solches f nicht global umkehrbar sein muss.

Aufgabe 6.4.3 *Finde eine einfache Funktion $f : \mathbb{R}^{n+m} \longrightarrow \mathbb{R}^m$, für welche die Gleichung $f(x, y) = 0$ für jedes $x \in \mathbb{R}^n$ unlösbar ist.*

Aufgabe 6.4.4 *Sei $(r, \phi)^T \mapsto f(r, \phi) = (r \cos\phi, r \sin\phi)^T$, für $r \in \mathbb{R}_+$, $\phi \in \mathbb{R}$. Zeige: Es ist $\det f'(r, \phi) \neq 0$ für alle diese r, ϕ. Also ist die Abbildung f lokal umkehrbar. Ist f auch global umkehrbar?*

6.5 Lokale Extrema

Definition 6.5.1 *Sei $G \subset \mathbb{R}^n$ ein Gebiet, und sei $f : G \longrightarrow \mathbb{R}$. Ein $\xi \in G$ heißt* lokales Maximum (Minimum) *von f, falls ein $r > 0$ existiert, für welches*

$$\forall\, x \in G \cap K(\xi, r) \; : \qquad f(x) \;\le\; f(\xi) \quad \big(f(x) \;\ge\; f(\xi)\big)\,.$$

In beiden Fällen nennen wir ξ dann auch lokales Extremum *von f.*

Satz 6.5.2 *Sei $G \subset \mathbb{R}^n$ ein Gebiet, und habe $f : G \longrightarrow \mathbb{R}$ in einem Punkt $\xi \in G$ ein lokales Extremum. Falls dann für einen Einheitsvektor e die Richtungsableitung $f_e(\xi)$ existiert, folgt $f_e(\xi) = 0$. Insbesondere folgt $\operatorname{grad} f(\xi) = 0$, falls f im Punkt ξ partiell differenzierbar ist.*

Beweis: Die Funktion $g(t) = f(x + t\,e)$ ist für kleine $|t|$ definiert und hat im Nullpunkt ein lokales Extremum. Also folgt $g'(0) = f_e(x) = 0$. □

Definition 6.5.3 *Sei $G \subset \mathbb{R}^n$ ein Gebiet, und sei $f : G \longrightarrow \mathbb{R}$. Falls f in $\xi \in G$ partiell differenzierbar ist, und falls $\operatorname{grad} f(\xi) = 0$ ist, heißt ξ ein* kritischer *oder* stationärer *Punkt von f. Falls f in ξ zweimal partiell differenzierbar ist, heißt die Matrix*

$$H_f(\xi) \;=\; \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2}(\xi) & \frac{\partial^2 f}{\partial x_1 \partial x_2}(\xi) & \dots & \frac{\partial^2 f}{\partial x_1 \partial x_n}(\xi) \\ \vdots & \vdots & & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1}(\xi) & \frac{\partial^2 f}{\partial x_n \partial x_2}(\xi) & \dots & \frac{\partial^2 f}{\partial x_n^2}(\xi) \end{bmatrix}$$

die Hesse-Matrix *von f im Punkt ξ. Falls f auf G zweimal stetig partiell differenzierbar ist, ist H_f nach dem Satz von Schwarz symmetrisch.*

Satz 6.5.4 *Sei $G \subset \mathbb{R}^n$ ein Gebiet, sei $f : G \longrightarrow \mathbb{R}$ zweimal stetig partiell differenzierbar auf G, und sei $\xi \in G$ ein kritischer Punkt von f. Dann gilt:*

(a) *Falls $H_f(\xi)$ positiv definit ist, ist ξ ein lokales Minimum von f.*

(b) *Falls $H_f(\xi)$ negativ definit ist, ist ξ ein lokales Maximum von f.*

(c) *Falls $H_f(\xi)$ indefinit ist, ist ξ kein lokales Extremum von f.*

(Ohne Beweis)

Aufgabe 6.5.5 *Sei $G \subset \mathbb{R}^n$ ein Gebiet, und seien $g, h : G \longrightarrow \mathbb{R}$ partiell differenzierbar sowie $h(x) \neq 0$ für alle $x \in G$. Zeige: Genau dann ist $x \in G$ kritischer Punkt von $f = g/h$, wenn* $\operatorname{grad} g(x) = f(x) \operatorname{grad} h(x)$ *gilt.*

Aufgabe 6.5.6 *Sei A eine n-reihige symmetrische Matrix, und sei $f(x) = \|x\|^{-2}\, x^T A\, x$. Zeige: Die stationären Punkte von f sind genau die Eigenvektoren von A. Man nennt dieses f auch den* Raleigh-Quotienten.

6.6 Extrema unter Nebenbedingungen

Definition 6.6.1 *Sei $G \subset \mathbb{R}^n$ ein Gebiet, und seien $f : G \longrightarrow \mathbb{R}$, $g : G \longrightarrow \mathbb{R}^m$, mit $1 \leq m \leq n-1$, und sei $\xi \in G$ mit $g(\xi) = 0$ gegeben. Dann heißt ξ* ein lokales Maximum (Minimum) *von f* unter der Nebenbedingung $g(x) = 0$, *falls ein $r > 0$ existiert mit*

$$\forall\ x \in G \cap K(\xi, r) \cap \{x \in G : g(x) = 0\} \ :\quad f(x) \ \leq\ f(\xi) \quad \big(f(x) \ \geq\ f(\xi)\big)\,.$$

Wieder sprechen wir in beiden Fällen auch von einem lokalen Extremum, *aber jetzt unter der Nebenbedingung $g(x) = 0$.*

Satz 6.6.2 (Lagrangesche Multiplikatorregel) *Sei G ein beliebiges Gebiet in $\mathbb{R}^n$, und seien $f : G \longrightarrow \mathbb{R}$, $g = (g_1, \ldots, g_m)^T : G \longrightarrow \mathbb{R}^m$, mit $1 \leq m \leq n-1$, beide stetig partiell differenzierbar auf G. Sei weiter ξ ein lokales Extremum von f unter der Nebenbedingung $g(x) = 0$, und habe $g'(\xi)$ den Rang m. Dann gibt es ein $\lambda_0 = (\lambda_1^{(0)}, \ldots, \lambda_m^{(0)})^T \in \mathbb{R}^m$ derart, dass die Funktion h, definiert durch*

$$h(x, \lambda) \ =\ f(x) \ +\ \lambda^T g(x) \qquad \forall\ \lambda \in \mathbb{R}^n,\ x \in G,$$

einen kritischen Punkt bei (ξ, λ_0) hat; d. h. also, dass die Gleichungen

$$\frac{\partial f}{\partial x_j}(\xi) \ + \sum_{k=1}^m \lambda_k^{(0)} \frac{\partial g_k}{\partial x_j}(\xi) \ =\ 0, \quad 1 \leq j \leq n,$$

$$g_k(\xi) \ =\ 0, \qquad 1 \ \leq\ k \leq\ n,$$

gelten müssen.

(Ohne Beweis)

Aufgabe 6.6.3 *Zeige dass die Funktion $f(x, y) = x + y$ auf der Menge $\{(x, y)^T \in \mathbb{R}^2 : x^2 + y^2 = 1\}$ ein Maximum und ein Minimum annimmt, und bestimme diese Punkte mit der Lagrangeschen Multiplikatorregel.*

Aufgabe 6.6.4 *Finde dasjenige Rechteck vom Umfang 1, dessen Flächeninhalt maximal ist. Gibt es auch eines mit minimalem Flächeninhalt?*

Literaturverzeichnis

[1] **T. Apostol**, *Mathematical Analysis*, Addison & Wesley, Reading, 1979.

[2] **I. N. Bronstein, K. A. Semendjajew, G. Musiol, und H. Mï¿$\frac{1}{2}$hlig**, *Taschenbuch der Mathematik*, Harri Deutsch, Frankfurt am Main, 1995.

[3] **K. Endl und W. Luh**, *Analysis II*, Aula Verlag, Wiesbaden, 1989.

[4] **F. Erwe**, *Differential- und Integralrechnung II*, BI Hochschultaschenbücher, Bibliographisches Institut, Mannheim, 1973.

[5] **H. Grauert und I. Lieb**, *Differential- und Integralrechnung II*, Heidelberger Taschenbücher, Springer, Berlin, 196.

[6] **H. Heuser**, *Lehrbuch der Analysis 2*, Teubner, Stuttgart, 1992.

[7] **W. Luh und M. W. 2**, *Aufgabensammlung Analysis*, Aula Verlag, Wiesbaden, 1992.

[8] **K. Meyberg und P. Vachenauer**, *Höhere Mathematik 1*, Springer, Berlin, 1999.

[9] ——, *Höhere Mathematik 2*, Springer, Berlin, 1999.

[10] **W. Walter**, *Analysis 2*, Springer, Berlin, 1992.

Index