

Skript zur Vorlesung Analysis 2

Sommersemester 2013

Prof. Dr. Benjamin Schlein

Inhaltsverzeichnis

1	Fourier-Reihen	2
2	Gewöhnliche Differentialgleichungen	10
2.1	Differentialgleichungen erster Ordnung, elementare Lösungsmethoden . .	11
2.2	Existenz und Eindeutigkeit	15
2.3	Differentialgleichungen höherer Ordnung	21
2.4	Lineare Differentialgleichungen	22
2.5	Lineare Differentialgleichungen mit konstanten Koeffizienten	26
2.6	Grundlagen der Stabilitätstheorie	34
3	Differentialrechnung in mehreren Veränderlichen	38
3.1	Definition der Ableitung für Funktionen auf \mathbb{R}^n	38
3.2	Mittelwertsatz	49
3.3	Höhere Ableitungen, Taylor Entwicklung, lokale Extrema	51
3.4	Umkehrabbildung und Satz über implizite Funktionen	59
3.5	Mannigfaltigkeiten in \mathbb{R}^n	66
3.6	Extrema mit Nebenbedingungen	72
3.7	Integrale, die von einem Parameter abhängen.	77
3.8	Konservative Vektorfelder	81
3.9	Holomorphe Funktionen	91

1 Fourier-Reihen

Wir betrachten in diesem Kapitel periodische Funktionen. Eine Funktion $f : \mathbb{R} \rightarrow \mathbb{C}$ heisst periodisch, mit Periode $T > 0$, falls $f(t + T) = f(t)$ für alle $t \in [0; T)$. Jede auf $[0; T)$ definierte Funktion f kann durch die Definition $f(t + kT) := f(t)$ für alle $k \in \mathbb{Z}$, und alle $t \in [0; T)$ periodisch fortgesetzt werden.

Ein wichtiges Beispiel einer periodischen Funktion ist die Exponentialfunktion $f(t) = e^{it}$. f hat die Periode $T = 2\pi$, weil

$$e^{i(t+2\pi)} = e^{it} e^{2\pi i} = e^{it}$$

für alle $t \in [0; 2\pi)$. $T = 2\pi$ ist die Fundamentalperiode der Funktion f , d.h. es existiert keine Periode $\tilde{T} > 0$ mit $\tilde{T} < T$. Für $j \in \mathbb{Z}$ ist die Funktion $f(t) = e^{ijt}$ auch periodisch. Die Fundamentalperiode von $f(t) = e^{ijt}$ ist $2\pi/|j|$. $T = 2\pi$ ist auch eine Periode von $f(t) = e^{ijt}$, für alle $j \in \mathbb{Z}$. In der Tat

$$e^{ij(t+2\pi)} = e^{ijt} e^{2\pi ij} = e^{ijt}$$

für alle $j \in \mathbb{Z}$. $\{e^{ijt}\}_{j \in \mathbb{Z}}$ ist damit eine unendliche Familie von 2π -periodischen Funktionen. Analog, für ein beliebiges $L > 0$, ist $\{e^{ijt/L}\}_{j \in \mathbb{Z}}$ eine unendliche Familie von Funktionen mit Periode $T = 2\pi/L$.

Lemma 1.1. *Seien $j, k \in \mathbb{Z}$. Dann gilt*

$$\int_0^{2\pi} e^{ikx} e^{-ijx} dx = \begin{cases} 2\pi & \text{falls } j = k \\ 0 & \text{sonst} \end{cases}$$

Bemerkung (aus Analysis 1): eine \mathbb{C} -wertige Funktion $f : [a; b] \rightarrow \mathbb{C}$ ist integrierbar, falls $\operatorname{Re} f$ und $\operatorname{Im} f$ integrierbar sind. In diesem Fall definieren wir

$$\int_a^b f dx = \int_a^b \operatorname{Re} f(x) dx + i \int_a^b \operatorname{Im} f(x) dx$$

Beweis: Wir haben $e^{ikx} e^{-ijx} = e^{i(k-j)x} = \cos((k-j)x) + i \sin((k-j)x)$. Für $k \neq j$ gilt

$$\int_0^{2\pi} \cos((k-j)x) dx = \frac{1}{k-j} (\sin(2\pi(k-j)) - \sin(0)) = 0$$

und analog,

$$\int_0^{2\pi} \sin((k-j)x) dx = 0$$

Dagegen, für $k = j$ ist $e^{i(k-j)x} = 1$ und

$$\int_0^{2\pi} e^{i(k-j)x} dx = 2\pi$$

□

Definition 1.2. Sei $f : \mathbb{R} \rightarrow \mathbb{C}$ eine 2π -periodische Funktion, auf $[0; 2\pi]$ integrierbar. Für ein beliebiges $j \in \mathbb{Z}$ ist dann die Funktion $e^{-ijx} f(x)$ auch 2π -periodisch und auf $[0; 2\pi]$ integrierbar. Wir definieren den j -ten Fourierkoeffizienten von f durch

$$\widehat{f}(j) = \frac{1}{2\pi} \int_0^{2\pi} f(x) e^{-ijx} dx$$

Weiter, für $N \in \mathbb{N}$, definieren wir die N -te Fourier Partialsumme

$$(\mathcal{F}_N f)(x) = \sum_{j=-N}^N \widehat{f}(j) e^{ijx}$$

Konvergiert die Folge $(\mathcal{F}_N f)(x)$ für $N \rightarrow \infty$, dann wird der Grenzwert durch

$$(\mathcal{F}f)(x) = \sum_{j=-\infty}^{\infty} \widehat{f}(j) e^{ijx} = \lim_{N \rightarrow \infty} (\mathcal{F}_N f)(x)$$

bezeichnet. $(\mathcal{F}f)$ wird die Fourier-Reihe von f genannt.

Wir werden sehen, dass unter geeigneter Annahme der Funktion f , die Fourier-Reihe von f mit f übereinstimmt; die Fourier-Reihe gibt also eine nützliche Darstellung von periodischen Funktionen als Limes von Linearkombinationen von den Funktionen e^{ijx} (ähnlich wie die Taylorreihe eine nützliche Darstellung von analytischen Funktionen gibt). Um zu zeigen, dass $\mathcal{F}f = f$ gilt, brauchen wir das folgende Lemma.

Lemma 1.3 (Lemma von Riemann-Lebesgue). Sei $f : [a; b] \rightarrow \mathbb{C}$ integrierbar. Dann gilt

$$\lim_{k \rightarrow \infty} \int_a^b e^{\pm ikx} f(x) dx = 0$$

Bemerkung: Die Idee hinter dem Lemma von Riemann-Lebesgue ist die folgende: Die Funktion e^{ikx} hat (Fundamental-) Periode $2\pi/k$. Deswegen gilt $\int_I e^{ikx} dx = 0$ für jedes Intervall der Länge $2\pi/k$. Wenn wir annehmen können, dass f auf diesen kleinen Intervallen näherungsweise konstant ist, dann muss auch $\int_I f(x) e^{ikx} dx$ näherungsweise verschwinden. Da wir $[a; b]$ in kleine Intervalle der Länge $(2\pi)/k$ zerlegen können, muss auch das Integral $\int_a^b f(x) e^{ikx} dx$ klein sein. Wenn man eine reguläre Funktion f betrachtet, dann kann man die Aussage des Riemann-Lebesgue Lemmas verfeinern. Ist f m -Mal differenzierbar, dann existiert eine Konstante C_m mit

$$\left| \int_a^b f(x) e^{ikx} dx \right| \leq C_m |k|^{-m}$$

Wir werden zurück zur Beziehung zwischen Regularität und Abfall von oszillierenden Integralen in den Übungen kommen.

Beweis: O.B.d.A. betrachten wir eine reelwertige Funktion $f : [a; b] \rightarrow \mathbb{R}$. Sei $\varepsilon > 0$ fest gewählt. Aus Analysis 1 (Proposition 9.3) existiert eine Teilung $T = \{a = x_0 < x_1 < \dots < x_n = b\}$ von $[a; b]$, mit

$$\overline{S}(T) - \frac{\varepsilon}{2} \leq \int_a^b f(x) dx \leq \overline{S}(T)$$

D.h. es existieren $h_j = \sup_{x \in [x_{j-1}; x_j]} f(x)$, mit

$$\sum_{j=1}^n h_j(x_j - x_{j-1}) - \frac{\varepsilon}{2} \leq \int_a^b f(x) dx \leq \sum_{j=1}^n h_j(x_j - x_{j-1})$$

Mit anderen Worten, für beliebige $\varepsilon > 0$ existiert eine Treppenfunktion

$$g(x) := \sum_{j=1}^n h_j \mathbf{1}_{[x_{j-1}; x_j]}(x)$$

auf $[a; b]$ mit $g(x) \geq f(x)$ für alle $x \in [a; b]$ und

$$\int_a^b |g(x) - f(x)| dx = \int_a^b (g(x) - f(x)) \leq \varepsilon/2 \quad (1)$$

Hier benutzen wir die Notation $\mathbf{1}_I(x)$ für die charakteristische Funktion des Intervalls I , definiert durch $\mathbf{1}_I(x) = 1$, falls $x \in I$ und $\mathbf{1}_I(x) = 0$, falls $x \notin I$.

Nun bemerken wir, dass

$$\int_a^b g(x) e^{ikx} dx = \sum_{j=1}^n h_j \int_a^b \mathbf{1}_{[x_{j-1}; x_j]}(x) e^{ikx} dx = \sum_{j=1}^n h_j \int_{x_{j-1}}^{x_j} e^{ikx} dx = \sum_{j=1}^n h_j \frac{e^{ikx_j} - e^{ikx_{j-1}}}{ik}$$

und damit

$$\left| \int_a^b g(x) e^{ikx} dx \right| \leq \frac{2}{|k|} \sum_{j=1}^n |h_j| \rightarrow 0$$

für $k \rightarrow \infty$ (oder $k \rightarrow -\infty$). Also existiert $K > 0$ gross genug, mit

$$\left| \int_a^b g(x) e^{ikx} dx \right| \leq \frac{\varepsilon}{2}$$

für alle $k > K$. Aus (1) bekommen wir

$$\begin{aligned} \left| \int_a^b f(x) e^{ikx} dx \right| &\leq \left| \int_a^b (f(x) - g(x)) e^{ikx} dx \right| + \left| \int_a^b g(x) e^{ikx} dx \right| \\ &\leq \int_a^b |f(x) - g(x)| dx + \left| \int_a^b g(x) e^{ikx} dx \right| \\ &\leq \varepsilon \end{aligned}$$

□

Satz 1.4. Sei $f : \mathbb{R} \rightarrow \mathbb{C}$ (2π) -periodisch und differenzierbar. Dann gilt

$$\lim_{N \rightarrow \infty} (\mathcal{F}_N f)(x) = f(x)$$

für alle $x \in [0; 2\pi]$.

Beweis: Es gilt

$$\mathcal{F}_N f(x) = \sum_{k=-N}^N \widehat{f}(k) e^{ikx} = \frac{1}{2\pi} \int_0^{2\pi} dt f(t) \sum_{k=-N}^N e^{ik(x-t)}$$

Aus Lemma 1.1 gilt

$$f(x) = \frac{1}{2\pi} \int_0^{2\pi} dt f(x) \sum_{k=-N}^N e^{ik(x-t)}$$

Damit

$$\begin{aligned} \mathcal{F}_N f(x) - f(x) &= \frac{1}{2\pi} \int_0^{2\pi} dt (f(t) - f(x)) \sum_{k=-N}^N e^{-ik(t-x)} \\ &= \frac{1}{2\pi} \int_{-x}^{2\pi-x} ds (f(x+s) - f(x)) \sum_{k=-N}^N e^{-iks} \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} ds (f(x+s) - f(x)) \sum_{k=-N}^N e^{-iks} \end{aligned}$$

wobei wir die Periodizität von f und e^{iks} im letzten Schritt benutzt haben. Wir berechnen nun

$$\sum_{k=-N}^N e^{-iks} = e^{-iNs} (1 + e^{is} + \dots + e^{2iNs}) = e^{-iNs} \frac{e^{is(2N+1)} - 1}{e^{is} - 1} = \frac{\sin(s(N+1/2))}{\sin s/2}$$

Damit gilt

$$\mathcal{F}_N f(x) - f(x) = \frac{1}{2\pi} \int_{-\pi}^{\pi} ds \frac{f(x+s) - f(x)}{\sin s/2} \sin((N+1/2)s)$$

Wir definieren

$$g(s) := \begin{cases} \frac{f(x+s) - f(x)}{\sin(s/2)} & \text{falls } s \neq 0 \\ 2f'(x) & \text{falls } s = 0 \end{cases}$$

Da f differenzierbar ist, ist g stetig bei $s = 0$, weil

$$\lim_{s \rightarrow 0} \frac{f(x+s) - f(x)}{\sin(s/2)} = 2 \lim_{s \rightarrow 0} \frac{f(x+s) - f(x)}{s} \cdot \frac{s/2}{\sin(s/2)} = 2f'(x)$$

f differenzierbar impliziert insbesondere, dass f stetig ist. Damit ist g auf $[-\pi; \pi]$ stetig, und deswegen sicher integrierbar. Lemma 1.3 impliziert also, dass

$$\int_{-\pi}^{\pi} ds g(s) \sin((N+1/2)s) = \frac{1}{2i} \left[\int_{-\pi}^{\pi} ds g(s) e^{is(N+1/2)} - \int_{-\pi}^{\pi} ds g(s) e^{-is(N+1/2)} \right] \rightarrow 0$$

für $N \rightarrow \infty$. □

Wir haben in Satz 1.4 die punktweise Konvergenz der Fourier-Reihe gegen f . Unter der Annahme, dass f differenzierbar ist, ist die Konvergenz eigentlich gleichmäßig. Um das zu zeigen, werden wir die zwei folgenden Lemmata brauchen.

Lemma 1.5. Sei $f \in C^1(\mathbb{R}; \mathbb{C})$ 2π -periodisch. Wie üblich bezeichnen wir mit $\widehat{f}(j)$ die Fourier Koeffizienten von f . Weiter bezeichnen wir mit $\widehat{f}'(j)$ die Fourier Koeffizienten von f' . Es gilt

$$\widehat{f}'(j) = ij\widehat{f}(j)$$

Proof. Durch partielle Integration bekommen wir

$$\widehat{f}'(j) = \frac{1}{2\pi} \int_0^{2\pi} f'(x)e^{-ijx} dx = \frac{f(2\pi) - f(0)}{2\pi} + \frac{ij}{2\pi} \int_0^{2\pi} dx f(x)e^{-ijx} = ij\widehat{f}(j)$$

□

Lemma 1.6. Sei $f : \mathbb{R} \rightarrow \mathbb{C}$ 2π -periodisch. Dann gilt

$$\int_0^{2\pi} |f(x) - \mathcal{F}_N f(x)|^2 dx = \int_0^{2\pi} |f(x)|^2 dx - 2\pi \sum_{k=-N}^N |\widehat{f}(k)|^2 \quad (2)$$

und

$$\int_0^{2\pi} |\mathcal{F}_N f(x)|^2 dx = 2\pi \sum_{k=-N}^N |\widehat{f}(k)|^2 \leq \int_0^{2\pi} |f(x)|^2 dx \quad (3)$$

Beweis: Wir berechnen

$$\begin{aligned} \int_0^{2\pi} |f(x) - \mathcal{F}_N f(x)|^2 dx &= \int_0^{2\pi} dx \left(\overline{f(x)} - \sum_{k=-N}^N \overline{\widehat{f}(k)} e^{-ikx} \right) \left(f(x) - \sum_{j=-N}^N \widehat{f}(j) e^{ijx} \right) \\ &= \int_0^{2\pi} |f(x)|^2 dx + \sum_{j,k=-N}^N \widehat{f}(j) \overline{\widehat{f}(k)} \int_0^{2\pi} dx e^{ix(j-k)} \\ &\quad - \sum_{j=-N}^N \widehat{f}(j) \int_0^{2\pi} f(x) e^{-ijx} dx - \sum_{k=-N}^N \overline{\widehat{f}(k)} \int_0^{2\pi} f(x) e^{-ikx} dx \\ &= \int_0^{2\pi} |f(x)|^2 dx - 2\pi \sum_{j=-N}^N |\widehat{f}(j)|^2 \end{aligned}$$

Das zeigt (2). Analog finden wir

$$\int_0^{2\pi} |\mathcal{F}_N f(x)|^2 dx = \sum_{j=-N}^N \overline{\widehat{f}(k)} \widehat{f}(j) \int_0^{2\pi} dx e^{ix(j-k)} = 2\pi \sum_{j=-N}^N |\widehat{f}(j)|^2$$

Aus (2) folgt nun (3). □

Satz 1.7. Sei $f \in C^1(\mathbb{R}; \mathbb{C})$ 2π -periodisch. Dann konvergiert die Funktionenfolge $\mathcal{F}_N f \rightarrow f$ gleichmässig, für $N \rightarrow \infty$.

Beweis: Wir haben schon punktweise Konvergenz gezeigt, d.h.

$$f(x) = \lim_{N \rightarrow \infty} \mathcal{F}_N f(x) = \lim_{N \rightarrow \infty} \sum_{j=-N}^N \widehat{f}(j) e^{ijx} = \sum_{j=-\infty}^{\infty} \widehat{f}(j) e^{ijx}.$$

Es gilt

$$\sum_{k=-N}^N |k\widehat{f}(k)|^2 = \sum_{k=-N}^N |\widehat{f}'(k)|^2 \leq \frac{1}{2\pi} \int_0^{2\pi} |f'(x)|^2 dx.$$

Damit konvergiert die Summe $\sum_{k=-\infty}^{\infty} |k|^2 |\widehat{f}(k)|^2$ absolut. Wir erhalten

$$\begin{aligned} |\mathcal{F}_N f(x) - f(x)| &= \left| \sum_{k=N+1}^{\infty} (\widehat{f}(k)e^{ikx} + \widehat{f}(-k)e^{-ikx}) \right| \\ &\leq \sum_{k=N+1}^{\infty} (|\widehat{f}(k)| + |\widehat{f}(-k)|) \\ &\leq \sum_{k=N+1}^{\infty} |k|^2 (|\widehat{f}(k)|^2 + |\widehat{f}(-k)|^2) + \sum_{k=N+1}^{\infty} \frac{1}{|k|^2} \rightarrow 0 \end{aligned}$$

für $N \rightarrow \infty$. Hier haben wir die Ungleichung $2|\widehat{f}(k)| \leq |k|^{-2} + |k|^2|\widehat{f}(k)|^2$ benutzt. \square

Was können wir nun sagen über die Fourier-Reihe, falls die periodische Funktion f nicht differenzierbar ist? Im nächsten Satz zeigen wir, dass wir immer noch Konvergenz von $\mathcal{F}_N f$ gegen f haben, aber in einem schwächeren Sinn; wir erhalten nämlich Konvergenz im Sinn von quadratischem Mittel.

Satz 1.8. Sei $f : \mathbb{R} \rightarrow \mathbb{C}$ 2π -periodisch und über $[0; 2\pi]$ integrierbar. Dann gilt

$$\lim_{N \rightarrow \infty} \int_0^{2\pi} dx |f(x) - \mathcal{F}_N f(x)|^2 = 0 \quad (4)$$

und die Parsevalsche Identität

$$\lim_{N \rightarrow \infty} \sum_{k=-N}^N |\widehat{f}(k)|^2 = \frac{1}{2\pi} \int_0^{2\pi} |f(x)|^2 dx$$

Ist f differenzierbar, so folgt (4) aus der gleichmässigen Konvergenz $\mathcal{F}_N f \rightarrow f$. Im Allgemeinen zeigen wir (4) indem wir zunächst f durch eine differenzierbare Funktion approximieren. Dazu benutzen wir das folgende Lemma.

Lemma 1.9. Sei $f : \mathbb{R} \rightarrow \mathbb{C}$ 2π -periodisch und auf $[0; 2\pi]$ integrierbar. Sei $\varepsilon > 0$. Dann existiert $g \in C^1(\mathbb{R}; \mathbb{C})$, 2π -periodisch, so dass

$$\int_0^{2\pi} |f - g|^2 dx \leq \varepsilon$$

Beweis: Sei $K := \sup_{x \in [0; 2\pi)} |f(x)| < \infty$. Wie im Beweis von Lemma 1.3, finden wir zunächst eine Teilung $0 = x_0 < x_1 < \dots < x_n = 2\pi$, und eine Treppenfunktion $h(x) = \sum_{j=1}^n h_j \mathbf{1}_{[x_j; x_{j-1})}(x)$ mit $|h_j| \leq K$ für alle $j = 1, \dots, n$, so dass

$$\int_0^{2\pi} |f(x) - h(x)| dx \leq \frac{\varepsilon}{8K}.$$

Dann gilt $|f(x) - h(x)| \leq |f(x)| + |h(x)| \leq 2K$ und damit

$$\int_0^{2\pi} |f(x) - h(x)|^2 dx \leq 2K \int_0^{2\pi} |f(x) - h(x)| dx \leq \frac{\varepsilon}{4}.$$

Nun approximieren wir für ein beliebiges $j \in \{1, \dots, n\}$ die charakteristische Funktion $\mathbf{1}_{[x_{j-1}; x_j]}$ durch eine differenzierbare Funktion. Wir setzen $\delta = \varepsilon/4 \sum_{j=1}^n h_j^2$. Ist $|x_j - x_{j-1}| < \delta$ dann setzen wir einfach $\theta_j(x) = 0$. Sonst setzen wir

$$\theta_j(x) = \begin{cases} 0 & \text{falls } x \in [0; x_{j-1}] \\ \sin^2((x - x_{j-1})\pi/\delta) & \text{falls } x \in [x_{j-1}; x_{j-1} + \delta/2] \\ 1 & \text{falls } x \in [x_{j-1} + \delta/2; x_j - \delta/2] \\ \sin^2((x - x_j)\pi/\delta) & \text{falls } x \in [x_j - \delta/2; x_j] \\ 0 & \text{falls } x \in [x_j, 2\pi] \end{cases}$$

Wir setzen θ_j durch Periodizität auf \mathbb{R} fort. Nach einfacher Rechnungen ist $\theta_j \in C^1(\mathbb{R})$, 2π -periodisch, mit

$$\int_0^{2\pi} |\theta_j(x) - \mathbf{1}_{[x_{j-1}; x_j]}(x)|^2 dx \leq \delta = \frac{\varepsilon}{4 \sum_{j=1}^n h_j^2}$$

Wir definieren nun $g(x) = \sum_{j=1}^n h_j \theta_j(x)$. Offenbar gilt $g \in C^1(\mathbb{R})$, 2π -periodisch. Weiter

$$h(x) - g(x) = \sum_{j=1}^n h_j \left(\mathbf{1}_{[x_{j-1}; x_j]}(x) - \theta_j(x) \right).$$

Da die verschiedenen Summanden auf disjunkten Intervallen getragen werden, gilt auch

$$|h(x) - g(x)|^2 = \sum_{j=1}^n h_j^2 \left| \mathbf{1}_{[x_{j-1}; x_j]}(x) - \theta_j(x) \right|^2$$

und damit

$$\int_0^{2\pi} |h(x) - g(x)|^2 dx = \sum_{j=1}^n h_j^2 \int_0^{2\pi} \left| \mathbf{1}_{[x_{j-1}; x_j]}(x) - \theta_j(x) \right|^2 dx \leq \delta \sum_{j=1}^n h_j^2 \leq \frac{\varepsilon}{4}$$

Das gibt

$$\int_0^{2\pi} |f(x) - g(x)|^2 dx \leq 2 \int_0^{2\pi} |f(x) - h(x)|^2 dx + 2 \int_0^{2\pi} |h(x) - g(x)|^2 dx \leq \varepsilon$$

Hier haben wir benutzt, dass $|f(x) - g(x)| \leq |f(x) - h(x)| + |h(x) - g(x)|$ und also, dass

$$|f(x) - g(x)|^2 \leq (|f(x) - h(x)| + |h(x) - g(x)|)^2 \leq 2|f(x) - h(x)|^2 + 2|h(x) - g(x)|^2.$$

□

Beweis von Satz 1.8: Sei $\varepsilon > 0$ festgewählt. Dann finden wir $g \in C^1(\mathbb{R}; \mathbb{C})$ mit

$$\int |f(x) - g(x)|^2 dx < \frac{\varepsilon}{9}$$

Das impliziert auch, dass

$$\int |\mathcal{F}_N f(x) - \mathcal{F}_N g(x)|^2 = \int |\mathcal{F}_N(f - g)(x)|^2 \leq \int |f(x) - g(x)|^2 \leq \frac{\varepsilon}{9}$$

Weiter, da g differenzierbar ist, gilt $\mathcal{F}_N g \rightarrow g$ gleichmässig. Aus Analysis 1 (Satz 9.13) folgt, dass

$$\int |\mathcal{F}_N g(x) - g(x)|^2 dx \rightarrow 0$$

für $N \rightarrow \infty$. Für N gross genug ist also

$$\int |\mathcal{F}_N g(x) - g(x)|^2 dx \leq \frac{\varepsilon}{9}$$

Insgesamt,

$$\begin{aligned} & \int |\mathcal{F}_N f(x) - f(x)|^2 dx \\ & \leq \int (|f(x) - g(x)| + |g(x) - \mathcal{F}_N g(x)| + |\mathcal{F}_N g(x) - \mathcal{F}_N f(x)|)^2 dx \\ & \leq 3 \int |f(x) - g(x)|^2 dx + 3 \int |g(x) - \mathcal{F}_N g(x)|^2 dx + 3 \int |\mathcal{F}_N g(x) - \mathcal{F}_N f(x)|^2 dx \\ & \leq \varepsilon \end{aligned}$$

falls N gross genug ist. □

Bemerkungen:

- 2π -periodische Funktionen können als Funktionen auf dem Einheitskreis $S^1 = \{e^{i\varphi} : \varphi \in \mathbb{R}\}$ gedacht werden.
- Ganz ähnlich kann man auch periodische Funktionen mit einer beliebigen Periode $L > 0$ betrachten (solche Funktionen werden mit Funktionen auf dem Kreis von Radius $L/2\pi$ identifiziert werden). In diesem Fall wird die Fourier-Reihe durch die Funktionen $\{e^{2\pi i j x/L}\}_{j \in \mathbb{Z}}$ definiert.
- Sei

$$V := \{f : \mathbb{R} \rightarrow \mathbb{C} : f \text{ } 2\pi\text{-periodisch und auf } [0; 2\pi] \text{ stetig}\}$$

Es ist einfach zu sehen, dass V ein unendlich dimensionaler Vektorraum ist. Für $f, g \in V$ definieren wir das Skalarprodukt

$$\langle f, g \rangle = \int_0^{2\pi} \overline{f(x)} g(x)$$

Die Funktionen $\{e^{ijx}\}_{j \in \mathbb{Z}}$ sind wegen Lemma 1.1 ein Orthonormalsystem auf V bezüglich dem Skalarprodukt $\langle \cdot, \cdot \rangle$. Satz 1.8 besagt dann, dass $\{e^{ijx}\}_{j \in \mathbb{Z}}$ eine Orthonormalbasis von V ist, d.h., dass jedes Element von V beliebig gut durch endliche lineare Kombinationen von den orthonormal Funktionen $\{e^{ijx}\}_{j \in \mathbb{Z}}$ approximiert werden kann. Die Fourier-Reihe $f(x) = \sum_{j=-\infty}^{\infty} \hat{f}(j)e^{ijx}$ gibt dann einfach die Darstellung von f als Grenzwert von endlichen linearen Kombinationen der Basis-Funktionen. Die Fourierkoeffizienten werden deswegen durch die Produkte $\hat{f}(j) = \langle e^{ijx}, f \rangle$ gegeben. Bemerke, dass V bezüglich der vom Skalarprodukt induzierten Metrik nicht vollständig ist. Um dieses Problem zu lösen, kann man die Vervollständigung \tilde{V} von V betrachten (jeder Skalarproduktraum kann vervollständigt werden). \tilde{V} ist ein Vektorraum, mit einem Skalarprodukt $[\cdot, \cdot]$ so, dass: 1) \tilde{V} vollständig, bezüglich der von $[\cdot, \cdot]$ induzierten Metrik ist, 2) V kann mit einem dichten Unterraum von \tilde{V} identifiziert werden, 3) Für $f, g \in V$ gilt $[f; g] = \langle f; g \rangle$. \tilde{V} ist ein sogenannter Hilbertraum (ein Skalarproduktraum, der vollständig ist, bezüglich der aus dem Skalarprodukt induzierten Metrik); es wird mit $L^2([0; 2\pi])$ bezeichnet. Mehr zu diesem Thema in der Vorlesung Funktionalanalysis.

- Ein Grund, warum Fourier-Reihen sehr nützlich sind, ist die Tatsache, dass Ableitungen auf Fourierkoeffizienten sehr einfach wirken. Aus Lemma 1.5 folgt, dass die Fourierkoeffizienten von $f^{(m)}(x)$ einfach durch $(ij)^m \hat{f}(j)$ gegeben sind. Differentialoperatoren sind, in diesem Sinn, diagonal im Fourierraum (wo die Funktion f durch ihre Fourierkoeffizienten $\{\hat{f}(j)\}_{j \in \mathbb{Z}}$ parametrisiert wird).

2 Gewöhnliche Differentialgleichungen

Differentialgleichungen sind Gleichungen, bei denen die Unbekannten Funktionen sind. Die Differentialgleichung definiert eine Beziehung zwischen den gesuchten Funktionen und ihren Ableitungen. Gewöhnliche Differentialgleichungen (auf Englisch “ordinary differential equations” oder einfach ODEs) sind Differentialgleichungen, wo die unbekannt Funktionen einer einzelnen reellen Variablen sind. Bei partiellen Differentialgleichungen sind dagegen die unbekannt Funktionen von mehreren Variablen. Hier werden wir nur gewöhnliche Differentialgleichungen betrachten (partielle Differentialgleichungen werden erst im vierten Semester untersucht).

Differentialgleichungen haben sehr viele Anwendungen. Die ganze Physik wird z.B. durch Differentialgleichungen formuliert: Die Newtonsche Gleichung der klassischen Mechanik, die Maxwell Gleichungen der Elektrodynamik, die Schrödingergleichung der Quantenmechanik, die Einsteingleichung der allgemeinen Relativitätstheorie sind alle Beispiele von Differentialgleichungen. Dabei ist nur die Newtonsche Gleichung eine gewöhnliche Differentialgleichung, die anderen sind partielle Differentialgleichungen. Die Newtonsche Gleichung beschreibt die Bewegung von Teilchen und Körpern unter der Wirkung von Kräften. Seien $x(t) = (x_1(t), x_2(t), x_3(t)) \in \mathbb{R}^3$ die Koordinaten eines Teilchens mit Masse m zur Zeit t . Sei $F(x) = (F_1(x), F_2(x), F_3(x))$ ein Kraftfeld. D.h. $F(x)$ ist die Kraft, die im Punkt x auf das Teilchen wirkt. Dann besagt die Newtonsche Gleichung, dass die Beschleunigung des Teilchens, die aus der zweiten Ableitung $x''(t)$ gegeben ist, proportional zur wirkenden Kraft ist. Genauer,

$$mx''(t) = F(x(t)) \tag{5}$$

Die Ableitung der vektorwertigen Funktion $x(t)$ ist komponentenweise zu verstehen; d.h. $x''(t) = (x_1''(t), x_2''(t), x_3''(t))$. Um die Trajektorie der Teilchen zu bestimmen, muss man also eine Funktion $x(t)$ finden, so dass für alle t erfüllt ist. Z.B., die Erde bewegt sich unter der Wirkung des Gravitationsfelds der Sonne. In einem Koordinatensystem, wo die Sonne an der Stelle $x = 0$ liegt, ist die Gravitationskraft, die die Sonne auf einem Körper der Masse m ausübt aus

$$F(x) = -Gm \frac{x}{|x|^3}$$

gegeben, für eine geeignete Konstante G . Bezeichnet also $x(t)$ die Position der Erde zur Zeit t , so muss $x(t)$ die Gleichung

$$mx''(t) = -Gm \frac{x}{|x|^3} \Rightarrow x''(t) = -G \frac{x}{|x|^3} \quad (6)$$

erfüllen. Diese Differentialgleichung hat mehrere Lösungen. Die Lösung kann eindeutig festgestellt werden, falls man geeignete Anfangsbedingungen spezifiziert. Schon Kepler hat herausgefunden, dass Lösungen von (6) immer auf einer Ebene bleiben und Ellipsen, Hyperbeln oder Parabeln beschreiben (für die Erde ist die Lösung eine Ellipse).

Gewöhnliche Differentialgleichungen werden nach ihrer Ordnung klassifiziert; die Ordnung der Differentialgleichung ist die Ordnung der höchsten Ableitung in der Gleichung. Eine Differentialgleichung erster Ordnung ist eine Differentialgleichung der Form $y'(x) = f(x, y(x))$ für die n unbekannt Funktionen $y(x) = (y_1(x), \dots, y_n(x))$ einer reellen Variable $x \in \mathbb{R}$. Eine Differentialgleichung m -ter Ordnung hat die Form $y^{(m)}(x) = f(x, y(x), y'(x), \dots, y^{(m-1)}(x))$. Die Lösung einer Differentialgleichung ist normalerweise nicht eindeutig. Sie wird aber oft eindeutig durch Spezifizierung von geeigneten Anfangsbedingungen. Z.B. eine Gleichung erster Ordnung für die n unbekannt Funktionen $y(x) = (y_1(x), \dots, y_n(x))$ wird oft eindeutig, falls wir die Bedingung $y(x_0) = (y_1^{(0)}, \dots, y_n^{(0)})$ für ein $x_0 \in \mathbb{R}$ und für einen Vektor $(y_1^{(0)}, \dots, y_n^{(0)}) \in \mathbb{R}^n$ verlangen. Gleichungen höherer Ordnung brauchen natürlich mehr Anfangsbedingungen. Eine Gleichung m -ter Ordnung wird oft eindeutig, falls wir Anfangsbedingungen für $y, y', \dots, y^{(m-1)}$ verlangen. Eine Differentialgleichung mit Anfangsbedingungen wird als ein Anfangswertproblem oder ein Cauchy-Problem bezeichnet.

Bei der Untersuchung von gewöhnlichen Differentialgleichungen werden für uns die folgenden Fragen eine wichtige Rolle spielen: Existiert eine Lösung der Differentialgleichung? Ist die Lösung unter Berücksichtigung von geeigneten Anfangsbedingungen eindeutig (d.h. ist die Lösung des Anfangswertproblems eindeutig)? Ist es möglich die Lösung explizit zu finden? Welche Methoden können verwendet werden, um die Lösung einer Differentialgleichung zu finden? Wie hängt die Lösung von den Anfangsbedingungen ab (Stabilitätstheorie für Differentialgleichungen)? Wir werden sehen, es ist nur selten möglich die Lösung einer Differentialgleichung explizit zu schreiben. Dagegen können Existenz und Eindeutigkeit der Lösungen unter allgemeinen Voraussetzungen gezeigt werden.

2.1 Differentialgleichungen erster Ordnung, elementare Lösungsmethoden

Wir betrachten hier gewöhnliche Differentialgleichungen erster Ordnung.

Definition 2.1. Sei $n \geq 1$, $U \subset \mathbb{R}^{n+1}$, $f \in C(U; \mathbb{R}^n)$. Dann ist

$$y'(x) = f(x, y(x)) \quad (7)$$

eine gewöhnliche Differentialgleichung erster Ordnung. Eine Lösung dieser Differentialgleichung auf einem Intervall $I \subset \mathbb{R}$ ist eine Funktion $y \in C^1(I; \mathbb{R}^n)$ so, dass $(x, y(x)) \in U$ und (7) erfüllt für alle $x \in I$ ist. Für $x_0 \in \mathbb{R}$, $y_0 \in \mathbb{R}^n$ mit $(x_0, y_0) \in U$ heisst

$$\begin{cases} y'(x) &= f(x, y(x)) \\ y(x_0) &= y_0 \end{cases} \quad (8)$$

ein Anfangswertproblem oder ein Cauchy-Problem. Eine Lösung des Anfangswertproblems (8) ist eine Lösung der Differentialgleichung (7), die auch die Anfangsbedingung $y(x_0) = y_0$ erfüllt (insbesondere muss $x_0 \in I$ sein). Ist $n = 1$, so heisst die Differentialgleichung skalar (die gesuchte Funktion hat Werten in \mathbb{R}). Ist dagegen $n > 1$, so heisst die Differentialgleichung vektoriell (man spricht in diesem Fall von einem System von Differentialgleichungen).

Wir betrachten ein paar Beispiele von Differentialgleichungen, wo die Lösungen explizit berechnet werden können (der Einfachheit halber betrachten wir hier Beispiele von skalaren Gleichungen; wir werden einige Beispiele von vektoriellen Gleichungen später betrachten, wenn wir lineare Differentialgleichungen untersuchen werden).

Beispiele:

- Sei $n = 1$, $I \subset \mathbb{R}$ ein offenes Intervall, $U = I \times \mathbb{R}$, und $f(x, y) = g(x)$ (unabhängig von y), für ein $g \in C(I)$. Wir betrachten die Differentialgleichung

$$\varphi'(x) = g(x)$$

Sei $G \in C^1(I)$ eine Stammfunktion von g , mit $G' = g$. Dann ist G eine Lösung der Differentialgleichung. Sei φ eine andere Lösung der Differentialgleichung. Dann gilt $(\varphi - G)'(x) = 0$ für alle $x \in I$. Das zeigt, dass jede Lösung die Form $\varphi(x) = G(x) + c$ hat, für eine Konstante $c \in \mathbb{R}$. Betrachten wir nun das Anfangswertproblem

$$\begin{cases} \varphi'(x) &= g(x) \\ \varphi(x_0) &= y_0 \end{cases}$$

für ein $x_0 \in I$ und ein $y_0 \in \mathbb{R}$. Die Lösung des Anfangswertproblems ist insbesondere die Lösung der Differentialgleichung und hat deswegen die Form

$$\varphi(x) = G(x) + c$$

Die Bedingung

$$y_0 = \varphi(x_0) = G(x_0) + c \quad \Rightarrow \quad c = y_0 - G(x_0)$$

bestimmt die Konstante c eindeutig. Die einzige Lösung des Anfangswertproblems ist aus

$$\varphi(x) = G(x) - G(x_0) + y_0$$

gegeben. Bemerke, dass die eindeutige Lösung auch als

$$\varphi(x) = y_0 + \int_{x_0}^x g(t) dt$$

geschrieben werden kann.

- Sei wieder $n = 1$, $U = \mathbb{R}^2$, und $f(x, y) = -y$. Die Differentialgleichung (7) nimmt dann die Form

$$\varphi'(x) = -\varphi(x) \tag{9}$$

Die Funktion $\varphi(x) = ce^{-x}$ erfüllt diese Differentialgleichung auf \mathbb{R} , für beliebige $c \in \mathbb{R}$. Wir behaupten jede Lösung auf \mathbb{R} hat diese Form. Sei in der Tat φ eine Lösung von (9) auf \mathbb{R} . Dann gilt

$$\frac{d}{dx} (e^x \varphi(x)) = e^x (\varphi(x) + \varphi'(x)) = 0$$

für alle $x \in \mathbb{R}$. Es existiert also eine Konstante $c \in \mathbb{R}$ mit $e^x \varphi(x) = c$ für alle $x \in \mathbb{R}$, d.h. mit $\varphi(x) = ce^{-x}$ für alle $x \in \mathbb{R}$. Betrachten wir nun das Anfangswertproblem

$$\begin{cases} \varphi'(x) &= -\varphi(x) \\ \varphi(x_0) &= y_0 \end{cases}$$

für $x_0, y_0 \in \mathbb{R}$. Die Lösung des Anfangswertproblem hat die Form $y(x) = ce^{-x}$. Die Anfangsbedingung $y(x_0) = y_0$ bestimmt die Konstante $c \in \mathbb{R}$ durch

$$y_0 = y(x_0) = ce^{-x_0} \quad \Rightarrow \quad c = y_0 e^{x_0}$$

Die eindeutige Lösung des Anfangswertproblems ist also $y(x) = y_0 \exp(-(x - x_0))$.

- Wir betrachten das Anfangswertproblem

$$\begin{cases} \varphi'(x) &= a(\varphi(x) - b\varphi^2(x)) \\ \varphi(0) &= y_0 \end{cases} \tag{10}$$

für $a, b, y_0 > 0$. Die Differentialgleichung in (10) heisst die logistische Gleichung oder die Differentialgleichung des beschränkten exponentiellen Wachstums, und hat z.B. Anwendungen in der Biologie (die Lösung beschreibt das Wachstum einer idealen Bakterienpopulation). Um die Gleichung zu lösen bemerken wir, dass

$$\frac{1}{\varphi(x) - b\varphi^2(x)} \varphi'(x) = a$$

Integration über x gibt

$$\int_0^x \frac{1}{\varphi(t) - b\varphi^2(t)} \varphi'(t) dt = a \int_0^x dt = ax$$

Wir substituieren $y = \varphi(t)$ und bekommen

$$\int_{\varphi(0)}^{\varphi(x)} \frac{1}{y - by^2} dy = ax$$

Aus

$$\frac{1}{y - by^2} = \frac{1}{y(1 - by)} = \frac{1}{y} + \frac{b}{1 - by}$$

finden wir

$$\log \frac{\varphi(x)(1 - by_0)}{y_0(1 - b\varphi(x))} = ax$$

Nach leichter algebraischer Manipulationen bekommen wir die eindeutige Lösung des Anfangswertproblems

$$\varphi(x) = \frac{y_0 e^{ax}}{1 + by_0(e^{ax} - 1)}$$

Im letzten Beispiel haben wir die Methode der Trennung der Variablen benutzt. Wir zeigen im nächsten Satz, dass diese Methode immer angewandt werden kann, falls die Funktion $f(x, y)$ auf der rechten Seite von (7) das Produkt einer Funktion von x mit einer Funktion von y ist.

Satz 2.2. Seien $I, J \subset \mathbb{R}$ offene Intervalle, $g \in C(I)$, $h \in C(J)$, mit $0 \notin h(J)$. Sei $(x_0, y_0) \in I \times J$. Seien

$$G(x) = \int_{x_0}^x g(t)dt, \quad \text{und} \quad H(y) = \int_{y_0}^y \frac{1}{h(t)}dt$$

Weiter, sei $I' \subset I$ ein offenes Intervall mit $G(I') \subset H(J)$ und $x_0 \in I'$. Dann existiert genau eine Lösung $\varphi \in C^1(I')$ des Anfangswertproblems

$$\begin{cases} \varphi'(x) = g(x)h(\varphi(x)) \\ \varphi(x_0) = y_0 \end{cases} \quad (11)$$

Ferner ist $\varphi : I' \rightarrow J$ die einzige Funktion mit

$$H(\varphi(x)) = G(x) \quad \text{für alle } x \in I'. \quad (12)$$

Bemerkung: Die Aussage impliziert, dass Differentialgleichungen der Form (11) durch Trennung der Variablen gelöst werden können. Das bedeutet, dass (11) zunächst als

$$\frac{1}{h(\varphi(x))} \varphi'(x) = g(x)$$

umgeschrieben werden kann. Integration über x ergibt dann

$$\int_{x_0}^x \frac{1}{h(\varphi(t))} \varphi'(t)dt = \int_{x_0}^x g(t)dt$$

und damit

$$\int_{\varphi(x_0)}^{\varphi(x)} \frac{1}{h(y)} dy = \int_{x_0}^x g(t)dt$$

und

$$H(\varphi(x)) = G(x)$$

Die eindeutige Lösung des Anfangswertproblems kann dann durch Umkehrung der Funktion H bestimmt werden.

Beweis: Da $H \in C^1(J)$ mit $H'(y) = 1/h(y) \neq 0$ für alle $y \in J$ ist H injektiv. Damit ist $H : J \rightarrow H(J)$ bijektiv und also invertierbar. Sei $T : H(J) \rightarrow J$ die Umkehrfunktion. Dann ist $T \in C^1(H(J))$ mit $T'(z) = 1/H'(T(z)) = h(T(z))$, für alle $z \in H(J)$. Die Gleichung (12) definiert eindeutig eine Funktion $\varphi = T \circ G \in C^1(I')$. Diese Funktion erfüllt $\varphi'(x) = h(T \circ G(x))G'(x) = h(\varphi(x))g(x)$ und $\varphi(x_0) = T \circ G(x_0) = T(0) = y_0$. D.h. φ ist eine Lösung des Anfangswertproblems. Das zeigt die Existenz der Lösung. Es bleibt die Eindeutigkeit zu zeigen. Sei dazu $\tilde{\varphi} \in C^1(I')$ eine andere Lösung des Anfangswertproblems. Es folgt, dass $\tilde{\varphi}(I') \subset J$. Sei $\psi = H \circ \tilde{\varphi} - G$. Dann gilt

$$\psi' = (H' \circ \tilde{\varphi})\tilde{\varphi}' - G' = \frac{1}{h \circ \tilde{\varphi}}\tilde{\varphi}' - g = 0$$

auf I' , Damit muss ψ konstant auf I' sein. Da aber $\psi(x_0) = H(\tilde{\varphi}(x_0)) - G(x_0) = 0$, muss $\psi(x) = 0$ für alle $x \in I'$. D.h. $H \circ \tilde{\varphi} = G$ auf I' , und deswegen, $\tilde{\varphi} = \varphi$. Das zeigt die Eindeutigkeit der Lösung. \square

2.2 Existenz und Eindeutigkeit

In diesem Abschnitt möchten wir zeigen, dass unter geeigneten Voraussetzungen an der Funktion f , das Anfangswertproblem (8) eine eindeutige Lösung besitzt. Dazu werden wir den Banachschen Fixpunktsatz anwenden. Erinnerung aus Analysis 1, dass ein metrischer Raum vollständig heißt, wenn jede Cauchy-Folge in M konvergiert. Wir haben in Analysis 1 gezeigt, dass \mathbb{R}^n , versehen mit der Standardmetrik vollständig für alle $n \in \mathbb{N}$ ist.

Satz 2.3 (Banachscher Fixpunktsatz). *Sei M , versehen mit der Metrik d , ein vollständiger metrischer Raum. $T : M \rightarrow M$ eine Abbildung mit der Eigenschaft, dass es eine Konstante $0 < c < 1$ existiert, mit*

$$d(T(x_1), T(x_2)) \leq c d(x_1, x_2)$$

für alle $x_1, x_2 \in M$ (eine solche Abbildung heißt eine Kontraktion; Kontraktionen sind insbesondere stetig). Dann gibt es genau ein $x \in M$ mit $T(x) = x$ (ein solches x heißt ein Fixpunkt der Abbildung T ; der Satz besagt, dass jede Kontraktion auf einem vollständigen metrischen Raum genau einen Fixpunkt besitzt).

Beweis: Wir zeigen zunächst die Eindeutigkeit. Nehme an, dass x_1, x_2 zwei Fixpunkte der Abbildung T sind. Dann gilt

$$d(x_1, x_2) = d(T(x_1), T(x_2)) \leq c d(x_1, x_2)$$

Da aber $c < 1$ ist diese Ungleichung nur möglich, falls $d(x_1, x_2) = 0$. Also $x_1 = x_2$. Nun zeigen wir die Existenz eines Fixpunktes. Sei $x_0 \in M$ beliebig. Dann definieren wir rekursiv eine Folge x_n in M durch $x_1 = T(x_0)$ und $x_{n+1} = T(x_n)$. Für $n \geq 1$ gilt dann

$$d(x_{n+1}, x_n) = d(T(x_n), T(x_{n-1})) \leq c d(x_n, x_{n-1}) \leq \dots \leq c^n d(x_1, x_0)$$

Es folgt, dass, für beliebige $n > m$,

$$\begin{aligned} d(x_n, x_m) &\leq d(x_n, x_{n-1}) + d(x_{n-1}, x_{n-2}) + \cdots + d(x_{m+1}, x_m) \\ &= \sum_{j=m+1}^n d(x_j, x_{j-1}) \\ &\leq d(x_1, x_0) \sum_{j=m+1}^n c^j \\ &\leq d(x_1, x_0) \sum_{j=m+1}^{\infty} c^j = \frac{d(x_1, x_0)}{1-c} c^{m+1} \rightarrow 0 \end{aligned}$$

für $m \rightarrow \infty$. D.h. x_n ist eine Cauchy-Folge auf M . Da M vollständig ist, muss x_n konvergieren. Sei $x = \lim_{n \rightarrow \infty} x_n$. Da aber T stetig ist, muss

$$T(x) = \lim_{n \rightarrow \infty} T(x_n) = \lim_{n \rightarrow \infty} x_{n+1} = x$$

Also, x ist ein Fixpunkt von T . □

Bemerkung: Der Beweis besagt, dass für jede $x_0 \in M$, die Folge $T \circ T \circ \cdots \circ T(x_0)$ gegen den Fixpunkt konvergiert. In praktischen Situationen, ergibt dies ein Verfahren, um den Fixpunkt von T zu approximieren.

Um die Existenz und Eindeutigkeit der Lösung von Anfangswertprobleme zu beweisen, werden wir den Banach'schen Fixpunktsatz auf dem Raum

$$C(I, \mathbb{R}^n) = \{f : I \rightarrow \mathbb{R}^n \text{ stetig} \}$$

anwenden. Hier ist $I \subset \mathbb{R}$ ein kompaktes Intervall. Dieser Raum ist mit der Metrik

$$d(f, g) = \sup_{x \in I} |f(x) - g(x)|$$

versehen. Wir haben in Analysis 1 gezeigt, dass $C(I, \mathbb{R}^n)$ vollständig ist (siehe Proposition 6.25 und die Diskussion danach; bemerke, dass die Kompaktheit von I impliziert, wegen dem Satz von Maximum, dass jede stetige Funktion auf I auch beschränkt ist. Deswegen ist $C(I; \mathbb{R}^n) = C_b(I; \mathbb{R}^n)$). Das nächsten Lemma impliziert dann, dass jede abgeschlossene Teilmenge von $C(I; \mathbb{R}^n)$ vollständig ist.

Lemma 2.4. *Sei M ein vollständiger metrischer Raum, und $A \subset M$ abgeschlossen. Dann ist A vollständig (bezüglich der von M induzierten Metrik).*

Beweis: Sei x_n eine Cauchy-Folge in A . Dann ist x_n auch eine Cauchy-Folge in M . Die Vollständigkeit von M impliziert, dass x_n in M konvergent. Sei $x \in M$ der Grenzwert von x_n , als Folge in M . Da A abgeschlossen ist und $x_n \in A$ für alle n ist, muss dann aber $x \in A$ sein. Damit ist x_n auch in A konvergent. □

Ein anderer Begriff spielt bei der Untersuchung der Existenz und Eindeutigkeit von Lösungen von Anfangswertproblemen eine wichtige Rolle. Das ist der Begriff der Lipschitz-Stetigkeit.

Definition 2.5. Seien (M_1, d_1) , (M_2, d_2) zwei metrische Räume, $A \subset M_1$. Eine Funktion $f : A \rightarrow M_2$ heisst Lipschitz-stetig falls eine Konstante $L > 0$ existiert, mit

$$d_{M_2}(f(x), f(y)) \leq L d_{M_1}(x, y)$$

für alle $x, y \in A$.

Bemerkungen:

- Jede Lipschitz-stetige Funktion ist gleichmässig stetig und damit auch stetig.
- Nicht alle gleichmässig stetigen Funktionen sind Lipschitz-stetig. Z.B. $f(x) = \sqrt{x}$ auf $[0; 1]$ ist gleichmässig stetig, aber nicht Lipschitz-stetig. Tatsache (Übung): Eine differenzierbare Funktion $f : (a; b) \rightarrow \mathbb{R}$ ist genau dann Lipschitz-stetig, falls die Ableitung beschränkt ist.

Wir sind nun bereit, um Existenz und Eindeutigkeit von Lösungen von Anfangswertprobleme der Form (8) zu zeigen. Wir bezeichnen im Folgenden mit $\|\cdot\|$ die euklidische Norm auf \mathbb{R}^n , die durch $\|a\|^2 = \sum_{j=1}^n |a_j|^2$ für $a = (a_1, \dots, a_n) \in \mathbb{R}^n$ definiert ist.

Satz 2.6 (Picard-Lindelöf). Sei $\Omega \subset \mathbb{R} \times \mathbb{R}^n$ offen, $(x_0, y_0) \in \Omega$, $f \in C(\Omega; \mathbb{R}^n)$ Lipschitz-stetig in der zweiten Variablen. Das bedeutet, dass $L > 0$ mit

$$\|f(x, y) - f(x, y')\| \leq L \|y - y'\|$$

für alle $x \in \mathbb{R}, y \in \mathbb{R}^n, y' \in \mathbb{R}^n$ mit $(x, y), (x, y') \in \Omega$ existiert. Dann gibt es ein $\varepsilon > 0$, so dass das Anfangswertproblem

$$\begin{cases} \varphi'(x) = f(x, \varphi(x)) \\ \varphi(x_0) = y_0 \end{cases} \quad (13)$$

eine eindeutige Lösung $\varphi \in C^1([x_0 - \varepsilon; x_0 + \varepsilon]; \mathbb{R}^n)$ besitzt.

Bemerkung: Satz 2.6 zeigt nur die Existenz und Eindeutigkeit einer lokalen Lösung, in der Nähe vom Punkt x_0 , wo die Anfangsbedingung gegeben ist. Wir werden später sehen, unter welchen Bedingungen die Existenz und Eindeutigkeit einer globalen Lösung gezeigt werden kann.

Der Beweis benutzt die Tatsache, dass eine Funktion $\varphi \in C([x_0 - \varepsilon; x_0 + \varepsilon], \mathbb{R}^n)$ genau dann eine Lösung des Anfangswertproblem (13) ist, wenn

$$\varphi(x) = y_0 + \int_{x_0}^x f(t, \varphi(t)) dt \quad (14)$$

Aus (14) folgt in der Tat sofort, dass $\varphi(x_0) = y_0$ ist. Ferner, aus der Stetigkeit von φ und von f , und aus dem Hauptsatz der Integralrechnung folgt auch, dass $\varphi \in C^1([x_0 - \varepsilon; x_0 + \varepsilon]; \mathbb{R}^n)$ mit $\varphi'(x) = f(x, \varphi(x))$ gilt. Andererseits, falls $\varphi \in C^1([x_0 - \varepsilon; x_0 + \varepsilon], \mathbb{R}^n)$ eine Lösung von (13) ist, dann folgt

$$\varphi(x) = \varphi(x_0) + \int_{x_0}^x \varphi'(t) dt = y_0 + \int_{x_0}^x f(t, \varphi(t)) dt.$$

Beweis: Sei $\delta > 0$ so klein, dass $K = [x_0 - \delta; x_0 + \delta] \times \overline{B_\delta(y_0)} \subset \Omega$. Da $K \in \mathbb{R}^{n+1}$ kompakt und f stetig ist, folgt, dass

$$M := \sup\{\|f(x, y)\| : (x, y) \in K\} < \infty$$

Wir wählen nun

$$0 < \varepsilon \leq \min\left\{\delta, \frac{\delta}{2L}, \frac{\delta}{2M}\right\} \quad (15)$$

und wir setzen $I = [x_0 - \varepsilon; x_0 + \varepsilon]$. Wir definieren

$$A = \{g \in C(I; \mathbb{R}^n) : \|g(x) - y_0\| \leq \delta \text{ für alle } x \in I\}$$

A ist dann eine abgeschlossene Teilmenge von $C(I; \mathbb{R}^n)$, versehen mit der Metrik $d(f, g) = \sup_{x \in I} \|f(x) - g(x)\|$ (Beweis: Übung). Es folgt aus Lemma 2.4, dass A ein vollständiger metrischer Raum ist. Wir definieren nun die Abbildung $T : A \rightarrow C(I; \mathbb{R}^n)$ durch

$$(T\phi)(x) = y_0 + \int_{x_0}^x f(t, \phi(t)) dt.$$

Offenbar ist $T\phi \in C(I; \mathbb{R}^n)$, für alle $\phi \in A$ (d.h. die Abbildung ist wohldefiniert). Weiter gilt, für alle $x \in I$ und alle $\phi \in A$,

$$\begin{aligned} \|(T\phi)(x) - y_0\| &= \left\| \int_{x_0}^x f(t, \phi(t)) dt \right\| \leq \varepsilon \sup\{\|f(t, \phi(t))\| : t \in I\} \\ &\leq \varepsilon \sup\{\|f(x, y)\| : x \in I, y \in \overline{B_\delta(y_0)}\} \leq \varepsilon \sup\{\|f(x, y)\| : (x, y) \in K\} = \varepsilon M \leq \delta/2 \end{aligned}$$

aus der Wahl (15). Damit gilt $T\phi \in A$, für alle $\phi \in A$. Weiter, für $\phi, \psi \in A$, finden wir

$$\begin{aligned} d(T\phi, T\psi) &= \sup_{x \in I} \|(T\phi)(x) - (T\psi)(x)\| = \sup_{x \in I} \left\| \int_{x_0}^x (f(t, \phi(t)) - f(t, \psi(t))) dt \right\| \\ &\leq \varepsilon \sup_{t \in I} \|f(t, \phi(t)) - f(t, \psi(t))\| \leq L\varepsilon \sup_{t \in I} \|\phi(t) - \psi(t)\| = \varepsilon L d(\phi, \psi) \leq \frac{1}{2} d(\phi, \psi) \end{aligned}$$

Damit ist $T : A \rightarrow A$ eine Kontraktion. Es folgt aus Satz 2.3, dass $\varphi \in A$ mit $T(\varphi) = \varphi$. Da $\varphi \in C(I; \mathbb{R}^n)$ ist $t \rightarrow f(t, \varphi(t))$ stetig, und damit $\varphi = T\varphi \in C^1([x_0 - \varepsilon; x_0 + \varepsilon]; \mathbb{R}^n)$ existiert. Ferner gilt $\varphi(x_0) = (T\varphi)(x_0) = y_0$ und, aus dem Hauptsatz der Integralrechnung,

$$\varphi'(x) = f(x, \varphi(x))$$

Damit ist φ eine Lösung des Anfangswertproblems (13) auf I (wir haben hier das Argument unten (14) wiederholt). Das zeigt die Existenz einer Lösung. Wir zeigen nun die Eindeutigkeit. Sei dazu $\psi \in C^1([x_0 - \varepsilon; x_0 + \varepsilon]; \mathbb{R}^n)$ eine andere Lösung von (13). Ist $\psi \in A$, so muss $T\psi = \psi$, weil ψ eine Lösung von (13) ist. Dann muss aber $\psi = \varphi$, weil φ der einzelne Fixpunkt von T ist. Ist $\psi \notin A$, dann muss es ein $x \in I$ geben, mit $\|\psi(x) - y_0\| > \delta$. O.B.d.A. nehmen wir an, es existiert $x \in I$, $x > x_0$ mit $\|\psi(x) - y_0\| > \delta$. Wir setzen dann

$$x_1 = \inf\{x \in I, x > x_0 : \|\psi(x) - y_0\| > \delta\}$$

Aus Stetigkeit von ψ muss dann $\|\psi(x_1) - y_0\| = \delta$ sein. Also

$$\delta = \|\psi(x_1) - y_0\| = \left\| \int_{x_0}^{x_1} f(t, \psi(t)) dt \right\| \leq \varepsilon \sup\{\|f(x, y)\| : (x, y) \in K\} \leq \delta/2$$

was ein Widerspruch ist. □

Bemerkungen:

- Die Lipschitz-Bedingung ist tatsächlich für die Existenz der Lösung nicht notwendig (Stetigkeit von f ist für die Existenz hinreichend). Dagegen ist die Lipschitz-Bedingung für die Eindeutigkeit der Lösung wichtig. Betrachte in der Tat das Anfangswertproblem

$$\begin{cases} \varphi'(x) &= \sqrt{|\varphi(x)|} \\ \varphi(0) &= 0 \end{cases}$$

In diesem Fall ist $f(x, y) = \sqrt{|y|}$ stetig, aber nicht Lipschitz-stetig in der Nähe von $y = 0$. Für ein beliebiges $a \geq 0$ ist dann die Funktion

$$\varphi(x) = \begin{cases} 0 & \text{falls } x < a \\ \frac{1}{4}(x - a)^2 & \text{falls } x \geq a \end{cases}$$

eine Lösung. Ferner ist auch $\varphi(x) = 0$ eine Lösung. Es existieren also unendlich viele Lösungen dieses Anfangswertproblems.

- Satz 2.6 besagt die Existenz und Eindeutigkeit einer Lösung auf einem genügend kleinen Intervall um x_0 . Im Allgemeinen existieren keine globalen Lösungen. Betrachte in der Tat das Anfangswertproblem

$$\begin{cases} \varphi'(x) &= 2x\varphi^2(x) \\ \varphi(0) &= 1 \end{cases} \quad (16)$$

Durch Trennung der Variablen finden wir die eindeutige Lösung

$$\varphi(x) = \frac{1}{1 - x^2}$$

auf dem Intervall $(-1; 1)$. Auf dem Intervall $[a; b]$ existiert also keine Lösung, falls $a \leq -1$ oder $b \geq 1$ (insbesondere existiert keine Lösung auf \mathbb{R}).

Aus der letzten Bemerkung stellt sich die Frage, ob es möglich ist, unter stärkeren Annahmen an f , die Existenz und Eindeutigkeit einer globalen Lösung zu zeigen. Die Antwort ist ja: Eine Lösung auf einem vorgegebenen Intervall $[a; b]$ existiert immer (und ist eindeutig), falls die Funktion $f(x, y)$ in der Variablen y auf ganz \mathbb{R}^n die Lipschitz-Bedingung erfüllt (die Funktion $f(x, y) = xy^2$, die in (16) vorkommt, ist nur für y in einem kompakten Intervall Lipschitz-stetig). Das ist der Inhalt des nächsten Satzes.

Satz 2.7 (Picard-Lindelöf, globale Version). *Sei $I = [a; b] \subset \mathbb{R}$ ein nicht-leeres kompaktes Intervall, $x_0 \in I$, $f \in C(I \times \mathbb{R}^n, \mathbb{R}^n)$ Lipschitz-stetig in der zweiten Variablen. D.h. es existiere $L > 0$ mit*

$$\|f(x, y) - f(x, y')\| \leq L\|y - y'\|$$

für alle $x \in I, y, y' \in \mathbb{R}^n$. Dann hat für jede $y_0 \in \mathbb{R}^n$ das Anfangswertproblem

$$\begin{cases} \varphi'(x) &= f(x, \varphi(x)) \\ \varphi(x_0) &= y_0 \end{cases} \quad (17)$$

eine eindeutige Lösung $\varphi \in C^1(I; \mathbb{R}^n)$.

Bemerkung: Satz 2.7 kann auch benutzt werden, um die Existenz und Eindeutigkeit von Lösungen auf \mathbb{R} zu zeigen. In der Tat eine Lösung auf \mathbb{R} existiert und genau dann eindeutig ist, wenn sie auf dem Intervall $[-m; m]$ existiert und ist eindeutig, für alle $m \in \mathbb{N}$.

Beweis: Für $\varphi \in C(I; \mathbb{R}^n)$, definieren wir

$$\|\varphi\|_L = \sup_{x \in [a; b]} e^{-2L|x-x_0|} \|\varphi(x)\|$$

Es ist einfach zu überprüfen, dass $\|\cdot\|_L$ eine Norm ist. Ferner, es gilt

$$e^{-2L(b-a)} \|\varphi\|_\infty \leq \|\varphi\|_L \leq \|\varphi\|_\infty \quad (18)$$

wobei $\|\varphi\|_\infty = \sup_{x \in [a; b]} \|\varphi(x)\|$. Sei φ_n eine Folge in $C(I; \mathbb{R}^n)$. Dann impliziert (18):

$$\begin{aligned} \varphi_n \text{ konvergiert bezüglich } \|\cdot\|_L &\iff \varphi_n \text{ konvergiert bezüglich } \|\cdot\|_\infty, \text{ und} \\ \varphi_n \text{ ist Cauchy-Folge bezüglich } \|\cdot\|_L &\iff \varphi_n \text{ ist Cauchy-Folge bezüglich } \|\cdot\|_\infty. \end{aligned}$$

(Man sagt, die zwei Normen $\|\cdot\|_\infty$ und $\|\cdot\|_L$ sind äquivalent). Es folgt insbesondere, dass $(C(I; \mathbb{R}^n), \|\cdot\|_L)$ ein vollständiger metrischer Raum ist. Auf $C(I; \mathbb{R}^n)$ definieren wir nun die Abbildung

$$(T\phi)(x) = y_0 + \int_{x_0}^x f(t, \phi(t)) dt$$

Dann gilt, für beliebige $x \in [a; b], x > x_0$,

$$\begin{aligned} \|(T\phi)(x) - (T\psi)(x)\| &= \left\| \int_{x_0}^x (f(t, \phi(t)) - f(t, \psi(t))) dt \right\| \leq \int_{x_0}^x \|f(t, \phi(t)) - f(t, \psi(t))\| dt \\ &\leq L \int_{x_0}^x \|\phi(t) - \psi(t)\| dt = L \int_{x_0}^x e^{2L|t-x_0|} e^{-2L|t-x_0|} \|\phi(t) - \psi(t)\| dt \\ &\leq L \|\phi - \psi\|_L \int_{x_0}^x e^{2L(t-x_0)} dt \leq \frac{1}{2} e^{2L|x-x_0|} \|\phi - \psi\|_L \end{aligned}$$

Analog gilt auch für $x \in [a; b]$ mit $x < x_0$,

$$\|(T\phi)(x) - (T\psi)(x)\| \leq \frac{1}{2} e^{2L|x-x_0|} \|\phi - \psi\|_L$$

Damit gilt

$$e^{-2L|x-x_0|} \|(T\phi)(x) - (T\psi)(x)\| \leq \frac{1}{2} \|\phi - \psi\|_L$$

für alle $x \in [a; b]$ und also

$$\|T\phi - T\psi\|_L \leq \frac{1}{2} \|\phi - \psi\|_L$$

Es folgt, dass T eine Kontraktion ist. Das impliziert, dass es einen eindeutigen Fixpunkt $\varphi \in C(I; \mathbb{R}^n)$, mit $T\varphi = \varphi$ gibt. Es ist dann einfach zu sehen, dass $\varphi \in C^1(I; \mathbb{R}^n)$ eine Lösung von (17) ist. Zur Eindeutigkeit: Ist $\psi \in C^1(I; \mathbb{R}^n)$ eine Lösung von (17), so ist insbesondere $\psi \in C(I; \mathbb{R}^n)$. Damit kann man T auf ψ anwenden. Da ψ eine Lösung des Anfangswertproblem ist, muss aber $T\psi = \psi$. Damit ist $\psi = \varphi$, weil T nur einen Fixpunkt haben kann. \square

2.3 Differentialgleichungen höherer Ordnung

Differentialgleichungen höherer Ordnung hängen auch von den höheren Ableitungen der gesuchten Funktion $y(x)$ ab.

Definition 2.8. Seien $n, k \in \mathbb{N} \setminus \{0\}$ fest, $\Omega \subset \mathbb{R} \times \mathbb{R}^{n \times k}$ offen, $f \in C(\Omega, \mathbb{R}^n)$. Dann ist

$$y^{(k)}(x) = f(x, y(x), y'(x), \dots, y^{(k-1)}(x)) \quad (19)$$

eine Differentialgleichung k -ter Ordnung. Eine Lösung von (19) auf einem Intervall $I \subset \mathbb{R}$ ist eine Funktion $y \in C^k(I; \mathbb{R}^n)$ so, dass

$$\left(x, y(x), y'(x), \dots, y^{(k)}(x) \right) \in \Omega$$

und (19) gilt, für alle $x \in I$. Für gegebene $(x_0, y_0, y_1, \dots, y_k) \in \Omega$ ist

$$\begin{cases} y^{(k)} = f(x, y(x), \dots, y^{(k-1)}(x)) \\ y^{(j)}(x_0) = y_j, \quad \text{für } j = 1, 2, \dots, (k-1) \end{cases} \quad (20)$$

ein Anfangswertproblem oder ein Cauchy-Problem k -ter Ordnung.

Man kann Resultate über die Existenz und Eindeutigkeit der Lösung von Anfangswertproblemen k -ter Ordnung aus den entsprechenden Resultaten für Gleichungen erster Ordnung herleiten, indem man bemerkt, dass eine Gleichung k -ter Ordnung zu einer Gleichung erster Ordnung in mehreren Variablen äquivalent ist. In der Tat, das Anfangswertproblem (20) kann wie folgt umgeschrieben werden. Wir definieren die neue Funktion $\psi(x) = (y(x), y'(x), \dots, y^{(k-1)}(x))$. Dann ist ψ eine Funktion mit Werten in $\mathbb{R}^{n \times k}$. Wir definieren ferner

$$\tilde{f}(x, z_0, z_1, \dots, z_{k-1}) := (z_1, z_2, \dots, z_{k-1}, f(x, z_0, z_1, \dots, z_{k-1}))$$

für alle $(x, z_0, z_1, \dots, z_{k-1}) \in \Omega$. Auch \tilde{f} hat Werte in $\mathbb{R}^{n \times k}$. Es ist dann leicht zu sehen, dass (20) zu dem Anfangswertproblem

$$\psi'(x) = \tilde{f}(x, \psi(x))$$

mit der Anfangsbedingung $\psi(x_0) = (y_0, y_1, \dots, y_{k-1}) \in \mathbb{R}^{n \times k}$ äquivalent ist. Damit haben wir ein Problem k -ter Ordnung in Dimension n in einem Problem erster Ordnung in Dimension nk umgeschrieben. Wir erhalten deswegen das folgende Existenz- und Eindeutigkeitsresultat.

Satz 2.9. Seien $n, k \in \mathbb{N} \setminus \{0\}$ fest, $\Omega \subset \mathbb{R} \times \mathbb{R}^{n \times k}$ offen, $f \in C(\Omega, \mathbb{R}^n)$ Lipschitz-stetig in alle Argumenten nach dem ersten. D.h. es existiere $L > 0$ mit

$$\|f(x, z_0, z_1, \dots, z_{k-1}) - f(x, z'_0, z'_1, \dots, z'_{k-1})\| \leq L \|z - z'\|$$

für alle $x \in \mathbb{R}$, $z = (z_0, \dots, z_{k-1})$, $z' = (z'_0, \dots, z'_{k-1}) \in \mathbb{R}^{n \times k}$ mit $(x, z), (x, z') \in \Omega$. Sei $(x_0, y_0, \dots, y_{k-1}) \in \Omega$. Dann existiert ein $\varepsilon > 0$ so, dass das Anfangswertproblem (20) eine eindeutige Lösung $\varphi \in C^k([x_0 - \varepsilon; x_0 + \varepsilon], \mathbb{R}^n)$ hat.

Beweis: Es genügt zu zeigen, dass die Funktion

$$\tilde{f}(x, z_0, z_1, \dots, z_{k-1}) = (z_1, z_2, \dots, z_{k-1}, f(x, z_0, z_1, \dots, z_{k-1}))$$

Lipschitz-stetig in $z = (z_0, z_1, \dots, z_{k-1})$ ist. Dazu bemerken wir, dass

$$\begin{aligned} \|\tilde{f}(x, z) - \tilde{f}(x, z')\| &= \|(z_1 - z'_1, z_2 - z'_2, \dots, z_{k-1} - z'_{k-1}, f(x, z) - f(x, z'))\| \\ &\leq \|z - z'\| + \|f(x, z) - f(x, z')\| \leq (L + 1) \|z - z'\|. \end{aligned}$$

□

2.4 Lineare Differentialgleichungen

Die Differentialgleichung erster Ordnung

$$y'(x) = f(x, y(x)) \tag{21}$$

heisst linear, falls die Funktion $f(x, y)$ affin in der Variable $y \in \mathbb{R}^n$ ist, d.h. falls eine offene Teilmenge $A \subset \mathbb{R}$, eine matrixwertige Funktion $a \in C(A; \mathbb{R}^{n \times n})$ und eine vektorwertige Funktion $b \in C(A; \mathbb{R}^n)$ existieren mit

$$f(x, y) = a(x)y + b(x) \tag{22}$$

Für ein beliebiges $x \in A$ bezeichnet hier $a(x)y$ die Anwendung der $n \times n$ Matrix $a(x)$ auf dem Vektor $y \in \mathbb{R}^n$. Die Differentialgleichung (21) heisst linear und homogen, falls $f(x, y)$ linear in y ist, d.h. falls f die Form (22) hat, mit $b = 0$.

Skalare lineare Differentialgleichungen: Wir betrachten zunächst den skalaren Fall, mit $n = 1$. Sei $I \subset \mathbb{R}$ ein Intervall und $a \in C(I)$. Für $x_0 \in I$ und $y_0 \in \mathbb{R}$ beliebig, untersuchen wir das skalare, lineare und homogene Anfangswertproblem

$$\begin{cases} y'(x) = a(x)y(x) \\ y(x_0) = y_0 \end{cases}$$

Das Anfangswertproblem hat eine eindeutige Lösung (aus Satz 2.7). Durch Trennung der Variablen finden wir, dass die eindeutige Lösung aus

$$\varphi(x) = y_0 \exp\left(\int_{x_0}^x a(t) dt\right)$$

gegeben ist.

Sei nun, wie vorher, $I \subset \mathbb{R}$ ein Intervall und $a \in C(I)$. Weiter, sei $b \in C(I)$. Für beliebige $x_0 \in I$ und $y_0 \in \mathbb{R}$, untersuchen wir das skalare, lineare (aber inhomogene) Anfangswertproblem

$$\begin{cases} y'(x) = a(x)y(x) + b(x) \\ y(x_0) = y_0 \end{cases}$$

Aus Satz 2.7, hat dieses Anfangswertproblem eine eindeutige Lösung. Die Lösung kann durch die Methode der Variation der Konstante gefunden werden. Man findet zunächst die allgemeine Lösung der homogenen Differentialgleichung $y'(x) = a(x)y(x)$, die aus

$$y(x) = c \exp\left(\int_{x_0}^x a(t)dt\right)$$

für eine beliebige Konstante $c \in \mathbb{R}$ gegeben ist. Um die inhomogene Gleichung zu lösen, betrachtet man den Ansatz

$$y(x) = c(x) \exp\left(\int_{x_0}^x a(t)dt\right)$$

bei welchem die Konstante c aus der Lösung der homogenen Gleichung nun von x abhängt. Dann ist

$$\begin{aligned} y'(x) &= c'(x) \exp\left(\int_{x_0}^x a(t)dt\right) + c(x)a(x) \exp\left(\int_{x_0}^x a(t)dt\right) \\ &= c'(x) \exp\left(\int_{x_0}^x a(t)dt\right) + a(x)y(x) \end{aligned}$$

Wir sehen also, dass $y(x)$ eine Lösung des inhomogenen Anfangswertproblems ist, g.d.w.

$$c'(x) \exp\left(\int_{x_0}^x a(t)dt\right) = b(x) \quad \iff \quad c'(x) = b(x) \exp\left(-\int_{x_0}^x a(t)dt\right)$$

Wir finden also, dass die eindeutige Lösung des inhomogenen Anfangswertproblems aus

$$\varphi(x) = \left(y_0 + \int_{x_0}^x b(t) \exp\left(-\int_{x_0}^t a(s)ds\right) dt\right) \exp\left(\int_{x_0}^x a(t)dt\right)$$

gegeben ist.

Vektorielle lineare Differentialgleichungen: Wir kommen nun zum allgemeinen Fall $n \geq 1$. Sei $I \subset \mathbb{R}$ ein Intervall, $x_0 \in I$, $a \in C(I, \mathbb{R}^{n \times n})$ eine matrix-wertige stetige Funktion auf I . Wir untersuchen das lineare, homogene Anfangswertproblem

$$\begin{cases} y'(x) = a(x)y(x) \\ y(x_0) = y_0 \end{cases} \quad (23)$$

für ein beliebiges $y_0 \in \mathbb{R}^n$.

Es lohnt sich in diesem Fall zunächst eine matrix-wertige Differentialgleichung zu lösen. Aus Satz 2.7 folgt nämlich, dass eine eindeutige Lösung $\varphi \in C^1(I, \mathbb{R}^{n \times n})$ des Anfangswertproblems

$$\begin{cases} \varphi'(x) = a(x)\varphi(x) \\ \varphi(x_0) = \mathbf{1} \end{cases} \quad (24)$$

existiert, wobei $\mathbf{1}$ die Identitätsmatrix auf \mathbb{R}^n ist. Bemerke, dass, für alle $x \in I$, $\varphi(x)$ hier eine $n \times n$ Matrix bezeichnet. Die Ableitung $\varphi'(x)$ ist wieder eine Matrix, mit Einträgen $(\varphi'(x))_{ij} = \varphi'_{ij}(x)$, wobei $\varphi_{ij}(x)$ die Einträge von $\varphi(x)$ sind (d.h. die Matrix wird Einträge-weise differenziert). Das Produkt $a(x)\varphi(x)$ soll dann als Produkt von zwei Matrizen verstanden werden. Die Matrixgleichung (24) ist einfach ein System von n^2 Differentialgleichungen, oder äquivalent, eine vektorielle Differentialgleichung für eine Unbekannte $\varphi(x)$ mit n^2 Komponenten (deswegen kann man Satz 2.7 anwenden). Analog existiert eine eindeutige Lösung $\psi \in C^1(I; \mathbb{R}^{n \times n})$ des Anfangswertproblems

$$\begin{cases} \psi'(x) = -\psi(x)a(x) \\ \psi(x_0) = \mathbf{1} \end{cases} \quad (25)$$

Wir behaupten nun, dass

$$\psi(x)\varphi(x) = \mathbf{1}$$

für alle $x \in I$. In der Tat, die Anfangsbedingung impliziert, dass $\psi(x_0)\varphi(x_0) = \mathbf{1}$. Andererseits

$$\frac{d}{dx} [\psi(x)\varphi(x)] = \psi'(x)\varphi(x) + \psi(x)\varphi'(x) = -\psi(x)a(x)\varphi(x) + \psi(x)a(x)\varphi(x) = 0$$

D.h. $\psi(x)\varphi(x)$ ist konstant auf I und deswegen $\psi(x)\varphi(x) = \mathbf{1}$, für alle $x \in I$. Das impliziert insbesondere, dass die Lösungen $\varphi(x)$ und $\psi(x)$ invertierbar sind, für alle $x \in I$.

Wir können nun die Lösung $\varphi(x)$ des Anfangswertproblems (24) benutzen, um die Lösung von (23) zu konstruieren. Aus Satz 2.7 wissen wir nämlich schon, dass (23) eine eindeutige Lösung besitzt. Wir behaupten nun, dass die eindeutige Lösung von (23) aus $y(x) = \varphi(x)y_0$ gegeben ist. In der Tat, $y(x_0) = \varphi(x_0)y_0 = \mathbf{1}y_0 = y_0$ und

$$y'(x) = \varphi'(x)y_0 = a(x)\varphi(x)y_0 = a(x)y(x) \quad (26)$$

Mit anderen Worten, die eindeutige Lösung von (23) bekommt man einfach durch Anwendung der Matrix $\varphi(x)$ auf die Anfangsbedingung $y_0 \in \mathbb{R}^n$.

Aus der Darstellung der Lösung von (23) als $y(x) = \varphi(x)y_0$ folgt einfach, dass der Lösungsraum der linearen Differentialgleichung in (23) eine lineare Struktur hat. Für gegebene $a \in C(I; \mathbb{R}^{n \times n})$ definieren wir nämlich den Lösungsraum der Differentialgleichung $y'(x) = a(x)y(x)$ als

$$L_h := \{y \in C^1(I; \mathbb{R}^n) : y'(x) = a(x)y(x)\}$$

(d.h. L_h ist die Menge aller Lösungen der Differentialgleichung, unabhängig von der Anfangsbedingung). Wir haben schon bewiesen, dass ein beliebiges $y \in L$ die Form $y(x) = \varphi(x)y(x_0)$ hat. Das impliziert offenbar, dass L_h ein Vektorraum ist. Da die Matrix $\varphi(x)$ invertierbar ist, folgt auch, dass $y^{(1)}, \dots, y^{(m)} \in L_h$ genau dann linear unabhängig sind, wenn $y^{(1)}(x_0), \dots, y^{(m)}(x_0) \in \mathbb{R}^n$ linear unabhängig sind. Das impliziert, dass $\dim L_h = \dim \mathbb{R}^n = n$ (mit anderen Worten, die Formel $y(x) = \varphi(x)y(x_0)$ erlaubt uns L_h mit \mathbb{R}^n zu identifizieren).

Die Lösung der Matrix-Gleichung (24) erlaubt uns auch inhomogene lineare Differentialgleichungen zu berechnen. Sei nämlich $I \subset \mathbb{R}$ ein Intervall, $x_0 \in I$, $y_0 \in \mathbb{R}^n$, $a \in C(I; \mathbb{R}^{n \times n})$ und $b \in C(I; \mathbb{R}^n)$. Dann hat das Anfangswertproblem

$$\begin{cases} y'(x) = a(x)y(x) + b(x) \\ y(x_0) = y_0 \end{cases} \quad (27)$$

die eindeutige Lösung

$$y(x) = \varphi(x) \left(y_0 + \int_{x_0}^x \varphi^{-1}(t)b(t)dt \right) \quad (28)$$

wobei $\varphi \in C^1(I; \mathbb{R}^{n \times n})$ die eindeutige Lösung von (24) ist. In der Tat, aus $\varphi(x_0) = \mathbf{1}$ folgt sofort, dass (28) die Bedingung $y(x_0) = y_0$ erfüllt. Weiter gilt

$$y'(x) = \varphi'(x) \left(y_0 + \int_{x_0}^x \varphi^{-1}(t)b(t) \right) + \varphi(x)\varphi^{-1}(x)b(x) = a(x)y(x) + b(x)$$

Wir haben in (28) benutzt, dass $\varphi(x)$ für alle $x \in I$ invertierbar ist. Es folgt aus diesem Ausdruck für die Lösung des Anfangswertproblems (27), dass der Lösungsraum der inhomogenen linearen Differentialgleichung $y'(x) = a(x)y(x) + b(x)$, definiert durch

$$L_i = \{y \in C^1(I; \mathbb{R}^n) : y'(x) = a(x)y(x) + b(x), \text{ für alle } x \in I\}$$

aus

$$L_i = L_h + \varphi(x) \int_{x_0}^x \varphi^{-1}(t)b(t) = \left\{ y(x) = y_h(x) + \varphi(x) \int_{x_0}^x \varphi^{-1}(t)b(t) : y_h \in L_h \right\}$$

gegeben ist. Allgemeiner, für eine beliebige Lösung $z \in L_i$, gilt $L_i = z + L_h$. D.h. L_i ist ein affiner Raum.

Lineare Differentialgleichungen höherer Ordnung: Lineare homogene und inhomogene Differentialgleichungen höherer Ordnung können als lineare homogene und inhomogene Differentialgleichungen erster Ordnung mit höherer Dimension geschrieben werden, ähnlich wie in Sektion 2.3 erklärt wird. Sei zum Beispiel $I \subset \mathbb{R}$ ein Intervall, $x_0 \in I$, $a_0, \dots, a_{n-1} \in C(I)$ und $b \in C(I)$ reelwertig. Der Lösungsraum L_h der linearen homogenen Differentialgleichung

$$y^{(n)}(x) + a_{n-1}(x)y^{(n-1)}(x) + \dots + a_0y(x) = 0$$

der Ordnung n ist ein linearer Vektorraum mit Dimension n . Zu jeder $(y_0, y_1, \dots, y_{n-1})$ gibt es genau eine Lösung $y \in L_h$, mit $y(x_0) = y_0$, $y'(x) = y_1, \dots, y^{(n-1)}(x_0) = y_{n-1}$. Der Lösungsraum L_i der linearen inhomogenen Gleichung

$$y^{(n)}(x) + a_{n-1}(x)y^{(n-1)}(x) + \dots + a_0y(x) + b(x) = 0$$

ist so, dass $L_i = z + L_h$, für ein beliebiges $z \in L_i$.

Bemerke, dass es im Gegensatz zum skalaren Fall $n = 1$, wo die Lösung von homogenen und inhomogenen Problemen mit Trennung der Variablen und Variationen der

Konstanten immer gefunden werden kann, bei vektoriellen linearen Problemen (und also bei Probleme höherer Ordnung) kein allgemeines Rezept gibt, um Lösungen zu finden. Wir haben nur gezeigt, dass die Lösung jeder vektoriellen linearen Gleichung zur Berechnung der Lösung $\varphi \in C^1(I; \mathbb{R}^{n \times n})$ der Matrix-Gleichung (24) reduziert werden kann (und wir haben die Darstellung der Lösung durch die Matrix $\varphi(x)$ benutzt, um allgemeine Eigenschaften der Lösungen zu diskutieren). Im Allgemeinen kann man aber φ nicht explizit berechnen. Eine Ausnahme, wo die Berechnung von φ auf die Lösung von linearen Gleichungssystemen reduziert werden kann, ist der Fall von linearen Gleichungen mit konstanten Koeffizienten, die durch Konstanten $a \in \mathbb{R}^{n \times n}$ und $b \in \mathbb{R}^n$ charakterisiert ist.

2.5 Lineare Differentialgleichungen mit konstanten Koeffizienten

Sei $n \in \mathbb{N}$ und $A \in \mathbb{R}^{n \times n}$ eine festgewählte $n \times n$ Matrix. In (23) setzen wir dann $a(x) = A$ für alle $x \in \mathbb{R}^n$. Wir bekommen die lineare homogene Differentialgleichung mit konstanten Koeffizienten

$$y'(x) = Ay(x) \quad (29)$$

für eine unbekannte Funktion $y \in C^1(\mathbb{R}; \mathbb{R}^n)$. Sei

$$L_h = \{y \in C^1(\mathbb{R}; \mathbb{R}^n) : y'(x) = Ay(x)\}$$

der Lösungsraum der Differentialgleichung (29). Wir wissen schon, dass L_h ein Vektorraum der Dimension n ist. Ist eine Basis $y^{(1)}, \dots, y^{(n)}$ des Lösungsraums L_h gegeben, so kann man die eindeutige Lösung des Anfangswertproblems

$$\begin{cases} y'(x) = Ay(x) \\ y(x_0) = y_0 \end{cases} \quad (30)$$

bestimmen, indem man den Vektor y_0 als eine lineare Kombination der Basisvektoren ausdrückt:

$$y_0 = \sum_{j=1}^n c_j y^{(j)}(x_0)$$

Das ist möglich, weil $\{y^{(j)}(x_0)\}_{j=1}^n$ eine Basis von \mathbb{R}^n ist. Dann ist die eindeutige Lösung von (30) aus

$$y(x) = \sum_{j=1}^n c_j y^{(j)}(x)$$

gegeben.

Wie können wir nun eine Basis von L_h finden? Sei $v \in \mathbb{R}^n$ ein Eigenvektor von A mit Eigenwert λ , d.h. $Av = \lambda v$. Dann ist $y(x) = ve^{\lambda x} \in L_h$, weil

$$y'(x) = \lambda ve^{\lambda x} = Ave^{\lambda x} = Ay(x).$$

Nehmen wir nun an, dass die Matrix A n linear unabhängige Eigenvektoren $v_1, \dots, v_n \in \mathbb{R}^n$, mit Eigenwerten $\lambda_1, \dots, \lambda_n \in \mathbb{R}$ (nicht notwendigerweise verschiedenen), besitzt. Dann sind die Funktionen $y_j(x) = v_j e^{\lambda_j x}$, für $j = 1, \dots, n$, linear unabhängig und damit eine Basis von L_h .

Es passiert oft, dass eine Matrix $A \in \mathbb{R}^{n \times n}$ auf \mathbb{C} , aber nicht auf \mathbb{R} diagonalisierbar ist. Sei $\lambda = \gamma + i\omega \in \mathbb{C} \setminus \mathbb{R}$ ein komplexer Eigenwert von A , mit Eigenvektor $v \in \mathbb{C}^n \setminus \{0\}$; wir zerlegen $v = u + iw$, mit $u, w \in \mathbb{R}^n$. Da A reelle Einträge hat, ist auch $\bar{\lambda} = \gamma - i\omega$ ein Eigenwert von A , mit Eigenvektor $\bar{v} = u - iw$. Die zwei Funktionen

$$\begin{aligned}\tilde{y}_1(x) &= ve^{\lambda x} = (u + iw)e^{i\omega x}e^{\gamma x} = [(u \cos \omega x - w \sin \omega x) + i(u \sin \omega x + w \cos \omega x)]e^{\gamma x} \\ \tilde{y}_2(x) &= (u - iw)e^{-i\omega x}e^{\gamma x} = [(u \cos \omega x - w \sin \omega x) - i(u \sin \omega x + w \cos \omega x)]e^{\gamma x}\end{aligned}$$

sind dann komplexe Lösungen der Differentialgleichung (29). Weil wir uns vor allem für reelle Lösungen interessieren, möchten wir \tilde{y}_1 und \tilde{y}_2 durch die reellen linearen Kombinationen

$$\begin{aligned}y_1(x) &= \frac{\tilde{y}_1(x) + \tilde{y}_2(x)}{2} = (u \cos \omega x - w \sin \omega x)e^{\gamma x} \\ y_2(x) &= \frac{\tilde{y}_1(x) - \tilde{y}_2(x)}{2i} = (u \sin \omega x + w \cos \omega x)e^{\gamma x}\end{aligned}\tag{31}$$

ersetzen.

Seien also $\tilde{v}_1, \dots, \tilde{v}_n \in \mathbb{C}^n$ linear unabhängige Eigenvektoren von A , zu den Eigenwerten $\lambda_1, \dots, \lambda_n \in \mathbb{C}$. Für jedes $j = 1, \dots, n$ unterscheiden wir zwei Fälle. Sei zunächst $\lambda_j \in \mathbb{R}$. Dann ist mit \tilde{v}_j auch $\overline{\tilde{v}_j}$ ein Eigenvektor von A zum Eigenwert λ . Mindestens einer der zwei Vektoren $\operatorname{Re} \tilde{v}_j = (\tilde{v}_j + \overline{\tilde{v}_j})/2$ und $\operatorname{Im} \tilde{v}_j = (\tilde{v}_j - \overline{\tilde{v}_j})/2i$ ist nicht Null und deswegen ein reeller Eigenvektor $v \in \mathbb{R}^n$ von A zum Eigenwert λ . Wir setzen, dann

$$y_j(x) = ve^{\lambda x}$$

Sei nun $\lambda_j \in \mathbb{C} \setminus \mathbb{R}$. Dann ist mit λ_j auch $\bar{\lambda}_j$ ein Eigenwert von A . D.h. es existiert $i \neq j$ mit $\lambda_i = \bar{\lambda}_j$. Dann setzen wir, gemäss (31),

$$\begin{aligned}y_j(x) &= (\operatorname{Re} v_j \cos(\operatorname{Im} \lambda_j) - \operatorname{Im} v_j \sin(\operatorname{Im} \lambda_j))e^{\operatorname{Re} \lambda_j x} \\ y_i(x) &= (\operatorname{Re} v_j \sin(\operatorname{Im} \lambda_j) - \operatorname{Im} v_j \cos(\operatorname{Im} \lambda_j))e^{\operatorname{Re} \lambda_j x}\end{aligned}$$

Damit konstruieren wir n linear unabhängige und reelle Lösungen y_j der Differentialgleichung $y'(x) = Ay(x)$; das gibt eine Basis vom Lösungsraum L_h .

Es gibt natürlich auch den Fall, dass die Matrix A nicht diagonalisierbar ist. D.h., dass keine n linearen unabhängigen Eigenvektoren von a existieren. In diesem Fall ist die Suche nach einer Basis des Lösungsraums L_h komplizierter. Es hilft, die Lösung der Differentialgleichung $y'(x) = Ay(x)$ durch Exponenzierung von A zu konstruieren.

Lösung durch Exponentialabbildung: Die Matrix $A = (a_{ij})$ ist ein Element von $\mathbb{R}^{n \times n}$. Auf diesem Raum ist die euklidische Norm aus

$$\|A\|^2 = \sum_{i,j=1}^n |a_{ij}|^2 = \operatorname{Tr} A^* A$$

gegeben. Es gibt eine andere natürliche Norm für Matrixen, nämlich die Operator-Norm. Wir definieren die Operatornorm von A durch

$$\|A\|_{\text{op}} = \sup_{v \in \mathbb{R}^n} \frac{\|Av\|}{\|v\|}$$

wobei $\|Av\|$ und $\|v\|$ die euklidischen Normen von Av und v , als Elemente von \mathbb{R}^n , sind. Es ist einfach zu zeigen, dass $\|\cdot\|_{\text{op}}$ wirklich eine Norm ist. Die Operatornorm hat die Eigenschaft, dass $\|Av\| \leq \|A\|_{\text{op}}\|v\|$, für einen beliebigen Vektor $v \in \mathbb{R}^n$ (das folgt direkt aus der Definition). Sind also $A, B \in \mathbb{R}^{n \times n}$ zwei Matrizen, dann gilt (AB bezeichnet die Multiplikation der zwei Matrizen A und B ; das entspricht der Verknüpfung der zwei Abbildungen)

$$\|ABv\| \leq \|A\|\|Bv\| \leq \|A\|\|B\|\|v\|$$

für alle $v \in \mathbb{R}^n$. Es folgt, dass

$$\|AB\|_{\text{op}} \leq \|A\|_{\text{op}}\|B\|_{\text{op}}$$

Es gilt

$$\frac{1}{\sqrt{n}}\|A\| \leq \|A\|_{\text{op}} \leq \|A\| \quad (32)$$

für jede $A \in \mathbb{R}^{n \times n}$. D.h. die zwei Normen $\|\cdot\|$ und $\|\cdot\|_{\text{op}}$ auf $\mathbb{R}^{n \times n}$ sind äquivalent (das gilt übrigens für jede zwei Normen auf einem beliebigen endlich dimensionalen Vektorraum). Man kann (32) wie folgt beweisen. Es gilt

$$\begin{aligned} \|Av\|^2 &= \sum_{j=1}^n \left| \sum_{i=1}^n a_{ji}v_i \right|^2 = \sum_{j=1}^n \sum_{i_1=1}^n \sum_{i_2=1}^n a_{ji_1} \bar{a}_{ji_2} v_{i_1} \bar{v}_{i_2} \\ &\leq \frac{1}{2} \sum_{j=1}^n \sum_{i_1=1}^n \sum_{i_2=1}^n (|a_{ji_1}|^2 |v_{i_2}|^2 + |a_{ji_2}|^2 |v_{i_1}|^2) = \|v\|^2 \|A\|^2 \end{aligned}$$

Das impliziert, dass $\|Av\|/\|v\| \leq \|A\|$ für alle $v \in \mathbb{R}^n$, und damit, dass $\|A\|_{\text{op}} \leq \|A\|$. Andererseits,

$$\|A\|^2 = \text{Tr } A^*A = \sum_{j=1}^n \langle e_j, A^*Ae_j \rangle = \sum_{j=1}^n \|Ae_j\|^2 \leq \sum_{j=1}^n \|A\|_{\text{op}}^2 \|e_j\|^2 \leq n \|A\|_{\text{op}}^2$$

wobei die Vektoren $e_j = (0, \dots, 0, 1, 0, \dots, 0)$ die Standardbasis von \mathbb{R}^n sind.

Wir betrachten nun für $A \in \mathbb{R}^{n \times n}$ die Folge

$$B_N := \sum_{j=0}^N \frac{A^j}{j!}$$

Wir bemerken, dass, für $N > M$,

$$\|B_N - B_M\|_{\text{op}} = \left\| \sum_{j=M+1}^N \frac{A^j}{j!} \right\|_{\text{op}} \leq \sum_{j=M+1}^N \frac{\|A^j\|_{\text{op}}}{j!} \leq \sum_{j=M+1}^N \frac{\|A\|^j}{j!}$$

Aus der Konvergenz der Reihe $\sum_{j=0}^{\infty} \|A\|^j/j!$ auf \mathbb{R} folgt, dass $\|B_N - B_M\|_{\text{op}} \rightarrow 0$, da $N, M \rightarrow \infty$. Das impliziert auch, dass $\|B_N - B_M\| \rightarrow 0$ bezüglich der euklidischen Norm

auf $\mathbb{R}^{n \times n}$. Also ist B_N eine Cauchy-Folge auf $\mathbb{R}^{n \times n}$, bzg. der euklidischen Norm, und damit konvergiert B_N . Wir definieren

$$\exp(A) := \lim_{N \rightarrow \infty} \sum_{j=0}^N \frac{A^j}{j!} \equiv \sum_{j=0}^{\infty} \frac{A^j}{j!}$$

Wir bemerken nun, dass die matrixwertige Funktion

$$\varphi(x) = \exp(A(x - x_0))$$

definiert für beliebige $x \in \mathbb{R}$, das Anfangswertproblem

$$\begin{cases} \varphi'(x) = A\varphi(x) \\ \varphi(x_0) = \mathbf{1} \end{cases} \quad (33)$$

löst. Die Anfangsbedingung $\varphi(x_0) = \mathbf{1}$ ist offenbar erfüllt. Wir zeigen nun, dass $\varphi(x)$ die Differentialgleichung erfüllt. Sei dazu $R > 0$ fest. Auf $x \in [x_0 - R; x_0 + R]$ konvergiert die matrixwertige Potenzreihe

$$B_N(x) = \sum_{j=0}^N \frac{A^j}{j!} (x - x_0)^j$$

gleichmässig gegen $\exp(A(x - x_0))$. Die Ableitung

$$B'_N(x) = \sum_{j=0}^N \frac{A^j}{j!} j (x - x_0)^{j-1} = A \sum_{j=1}^N \frac{A^{j-1}}{(j-1)!} (x - x_0)^{j-1} = A \sum_{j=0}^{N-1} \frac{A^j}{j!} (x - x_0)^j$$

konvergiert auch gleichmässig gegen $A \exp(A(x - x_0))$. Aus Analysis 1 (Proposition 8.29) folgt, dass $\varphi \in C^1([-R; R]; \mathbb{R}^{n \times n})$, mit $\varphi'(x) = A\varphi(x)$ für alle $x \in [-R; R]$ (bemerke, dass Proposition 8.29 in Analysis 1 nur für \mathbb{R} -wertige Funktionenfolgen formuliert ist. Das Resultat lässt sich aber trivial auf matrixwertige Funktionen erweitern, indem man die n^2 Komponenten der Matrix separat untersucht). Da $R > 0$ beliebig ist, folgt dass $\varphi \in C^1(\mathbb{R}; \mathbb{R}^{n \times n})$ das Anfangswertproblem (33) auf ganz \mathbb{R} löst.

Wie in (26), ist nun die eindeutige Lösung des Anfangswertproblems (30) durch Anwendung der Matrix $\varphi(x)$ auf die Anfangsbedingung, d.h.

$$y(x) = \varphi(x)y_0 = \exp(A(x - x_0))y_0$$

Ferner, das inhomogene Anfangswertproblem mit konstanten Koeffizienten $y'(x) = ay(x) + b$, mit der Anfangsbedingung $y(x_0) = y_0$, hat, gemäss (28) die eindeutige Lösung

$$y(x) = e^{a(x-x_0)} \left(y_0 + \int_{x_0}^x e^{-a(x-x_0)} b dt \right).$$

Aus einem praktischen Sichtpunkt ist die Berechnung der Exponentialabbildung $\exp(a(x - x_0))$ durch die Diagonalisierung von A möglich. Ist nämlich $A = U^{-1}DU$,

für eine diagonale Matrix $D = \text{diag}(d_1, \dots, d_n)$, so gilt $A^m = (U^{-1}DU)^m = U^{-1}D^mU$. Es folgt

$$\begin{aligned} \exp(A(x-x_0)) &= \sum_{m=0}^{\infty} A^m \frac{(x-x_0)^m}{m!} = \sum_{m=0}^{\infty} U^{-1}D^mU \frac{(x-x_0)^m}{m!} \\ &= U^{-1} \sum_{m=0}^{\infty} D^m \frac{(x-x_0)^m}{m!} U \\ &= U^{-1} \text{diag} \left(\sum_{m=0}^{\infty} \frac{d_1^m (x-x_0)^m}{m!}, \dots, \sum_{m=0}^{\infty} \frac{d_n^m (x-x_0)^m}{m!} \right) U \\ &= U^{-1} \text{diag} \left(e^{(x-x_0)d_1}, \dots, e^{(x-x_0)d_n} \right) U. \end{aligned}$$

Ist die Matrix a nicht diagonalisierbar, so ist die Berechnung von $\exp((x-x_0)a)$ schwieriger. In diesem Fall kann die jordanische Normalform der Matrix A verwendet werden. Für beliebige $A \in \mathbb{R}^{n \times n}$ kann man nämlich eine invertierbare Matrix U und eine blockdiagonale Matrix

$$J = \begin{pmatrix} J_1 & 0 & 0 & \dots & 0 \\ 0 & J_2 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & J_k \end{pmatrix}$$

finden, so dass $A = U^{-1}JU$. J heisst die Jordan Normalform von A . Die Blöcke J_i haben die Form

$$J_i = \begin{pmatrix} \lambda_i & 1 & 0 & \dots & 0 \\ 0 & \lambda_i & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & 0 & \lambda_i & 1 \\ 0 & \dots & 0 & 0 & \lambda_i \end{pmatrix} \quad (34)$$

Die Einträge λ_i auf der Diagonalen der Matrixen J_i sind die Eigenwerte von A . Die Anzahl der Blöcke mit Eigenwert λ_i ist aus der geometrischen Vielfachheit von λ_i gegeben (d.h. die Dimension des Eigenraumes mit Eigenwert λ_j). Die Gesamtdimension der Jordanblöcke mit Eigenwert λ_i ist dagegen die algebraische Vielfachheit von λ_i (die Vielfachheit von λ_i als Nullstelle des charakteristischen Polynoms). Sind algebraische und geometrische Vielfachheit gleich, so ist jeder Jordanblock mit Eigenwert λ_i eine 1×1 Matrix mit Eintrag λ_i (ist das der Fall für alle Eigenwerte von A , dann ist J diagonal und A diagonalisierbar).

Der Ausdruck $A = U^{-1}JU$ erlaubt uns, die Exponentialabbildung $\exp(tA)$ zu berechnen, für ein beliebiges $t \in \mathbb{R}$ ($t = x - x_0$ in unserer Anwendung). In der Tat

$$\exp(tA) = \exp(U^{-1}tJU) = U^{-1} \exp(tJ)U$$

Die Anwendung der Exponentialabbildung an der Blockdiagonale Matrix tJ ist wieder blockdiagonal, mit Blöcken $\exp(tJ_i)$, wobei die J_i die Form (34) haben. Sei J_i eine $\ell \times \ell$

Matrix. Dann liefert die Berechnung von $\exp(tJ_i)$ (Beweis: Übung)

$$\exp \begin{pmatrix} t\lambda_i & t & 0 & \dots & 0 \\ 0 & t\lambda_i & t & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & 0 & t\lambda_i & t \\ 0 & \dots & 0 & 0 & t\lambda_i \end{pmatrix} = e^{t\lambda_i} \begin{pmatrix} 1 & t & \frac{t^2}{2!} & \dots & \frac{t^{\ell-1}}{(\ell-1)!} \\ 0 & 1 & t & \dots & \frac{t^{\ell-2}}{(\ell-2)!} \\ \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & 0 & 1 & t \\ 0 & \dots & 0 & 0 & 1 \end{pmatrix}$$

Damit kann man in Prinzip die Exponentialabbildung $\varphi(x) = \exp(A(x - x_0))$ für jede Matrix $A \in \mathbb{R}^{n \times n}$ berechnen. Mit dieser Methode kann man also immer die eindeutige Lösung des Anfangswertproblems (30) finden.

Skalare, lineare, homogene Differentialgleichungen höherer Ordnung mit konstanten Koeffizienten: Eine skalare, lineare, homogene Differentialgleichung der Ordnung $n \in \mathbb{N}$ mit konstanten Koeffizienten hat die Form

$$a_n \varphi^{(n)}(x) + a_{n-1} \varphi^{(n-1)}(x) + \dots + a_1 \varphi'(x) + a_0 \varphi(x) = 0 \quad (35)$$

für eine Funktion $\varphi \in C^n(\mathbb{R})$, und für Konstanten $a_0, \dots, a_n \in \mathbb{R}$. Definieren wir $y = (\varphi, \varphi', \dots, \varphi^{(n-1)}) \in C^1(\mathbb{R}; \mathbb{R}^n)$, dann nimmt (35) die Form

$$y'(x) = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & 0 & 1 \\ -\frac{a_0}{a_n} & -\frac{a_1}{a_n} & -\frac{a_2}{a_n} & \dots & -\frac{a_{n-1}}{a_n} \end{pmatrix} y(x) =: Ay(x)$$

Es ist einfach zu sehen, dass die Matrix A genau dann (in \mathbb{C}) diagonalisierbar ist, wenn sie n verschiedene Eigenwerte hat (Übung: Gilt $Av_1 = \lambda v_1$ und $Av_2 = \lambda v_2$ für ein $\lambda \in \mathbb{C}$, dann existiert $\kappa \in \mathbb{C}$ mit $v_1 = \kappa v_2$).

Obwohl die Matrix A nicht immer diagonalisierbar ist, ist es hier trotzdem einfach, eine Basis für den n dimensionalen Lösungsraum

$$L_h = \left\{ y \in C^n(\mathbb{R}) : \sum_{j=0}^n a_j \varphi^{(j)}(x) = 0 \right\}$$

zu bestimmen, ohne explizit die jordsche Normalform von A zu berechnen. Wir definieren dazu das Polynom

$$p(s) = a_n s^n + a_{n-1} s^{n-1} + \dots + a_1 s + a_0$$

Dann nimmt (35) die Form

$$p \left(\frac{d}{dx} \right) \varphi = 0 \quad (36)$$

Aus dem Fundamentalsatz der Algebra (Satz 2.33 in Analysis 1) existieren $k \in \mathbb{N}$, $\lambda_1, \dots, \lambda_k \in \mathbb{C}$ voneinander unterschiedliche, $n_1, \dots, n_k \in \mathbb{N} \setminus \{0\}$ mit $\sum_{j=1}^k n_j = n$ so, dass

$$p(s) = a_n \prod_{j=1}^k (s - \lambda_j)^{n_j} \quad (37)$$

Wir können also (36) als

$$p\left(\frac{d}{dx}\right)\varphi = a_n \prod_{j=1}^k \left(\frac{d}{dx} - \lambda_j\right)^{n_j} \varphi = 0 \quad (38)$$

schreiben. Bemerke hier, dass die Ordnung der Operatoren $(d/dx - \lambda_j)^{n_j}$ keine Rolle spielt (die verschiedenen Monomen kommutieren miteinander). Wir bemerken ferner, dass

$$\left(\frac{d}{dx} - \lambda\right)^k [f(x)e^{\lambda x}] = f^{(k)}(x)e^{\lambda x} \quad (39)$$

In der Tat

$$\left(\frac{d}{dx} - \lambda\right) [f(x)e^{\lambda x}] = f'(x)e^{\lambda x} + \lambda f(x)e^{\lambda x} - \lambda f(x)e^{\lambda x} = f'(x)e^{\lambda x}$$

Wenden wir diese Formel k Mal an, so finden wir (39). Aus (36) erhalten wir also

$$\begin{aligned} p\left(\frac{d}{dx}\right)[x^\ell e^{\lambda_m x}] &= a_n \prod_{j \neq m} \left(\frac{d}{dx} - \lambda_j\right)^{n_j} \left(\frac{d}{dx} - \lambda_m\right)^{n_m} [x^\ell e^{\lambda_m x}] \\ &= a_n \prod_{j \neq m} \left(\frac{d}{dx} - \lambda_j\right)^{n_j} \left[\frac{d^{n_m}}{dx^{n_m}} x^\ell\right] e^{\lambda_m x} = 0 \end{aligned}$$

für alle $\ell = 0, 1, 2, \dots, n_m - 1$. Die n Funktionen

$$\{y_{m,\ell}(x) = x^\ell e^{\lambda_m x} : m = 1, \dots, k, \text{ und } \ell = 0, 1, \dots, n_m - 1\}$$

sind also Lösungen der Differentialgleichung $p(d/dx)\varphi = 0$. Diese Funktionen sind linear unabhängig (Beweis: Übung), und definieren also eine Basis des Lösungsraums L_h , wenn wir L_h als einen Vektorraum über \mathbb{C} betrachten. Mit anderen Worten, jede Lösung von (35) kann als eine endliche lineare Kombination der Funktionen $y_{m,\ell}$ geschrieben werden. Für gegebene Anfangsbedingungen $y(x_0) = y_0, y'(x_0) = y_1, \dots, y^{(n-1)}(x_0) = y_{n-1}$ kann man also die eindeutige Lösung des Anfangswertproblems bestimmen, indem man Konstanten $\alpha_{m,\ell}$ findet, so dass

$$y(x) = \sum_{m=1}^k \sum_{\ell=1}^{n_m-1} \alpha_{m,\ell} x^\ell e^{\lambda_m x} \quad (40)$$

alle Anfangsbedingungen an der Stelle $x = x_0$ erfüllt. Die resultierende Lösung (40) ist, für reelle Anfangsbedingungen $y_0, \dots, y_{n-1} \in \mathbb{R}$ automatisch reell. Falls man aber eine Basis für L_h betrachtet als \mathbb{R} -Vektorraum, sucht, so muss man wie oben die Lösungen $x^\ell e^{\lambda_j x}, x^\ell e^{\bar{\lambda}_j x}$ mit $\lambda_j \in \mathbb{C} \setminus \mathbb{R}$ durch die reellen linearen Kombinationen

$$x^\ell \cos((\operatorname{Im} \lambda_j)x) e^{(\operatorname{Re} \lambda_j)x} \quad \text{und} \quad x^\ell \sin((\operatorname{Im} \lambda_j)x) e^{(\operatorname{Re} \lambda_j)x}$$

ersetzen (das ist immer möglich, wenn die Koeffizienten $a_0, \dots, a_n \in \mathbb{R}$, weil für jede Nullstelle λ von p auch $\bar{\lambda}$ eine Nullstelle ist).

Skalare, lineare, inhomogene Differentialgleichungen höherer Ordnung mit konstanten Koeffizienten: Auch inhomogene, skalare, lineare Differentialgleichungen von höherer Ordnung mit konstanten Koeffizienten kann man mit dieser Methode lösen. Aus der Untersuchung von allgemeinen linearen Differentialgleichungen wissen wir schon, dass der Lösungsraum der inhomogenen Gleichung

$$a_n y^{(n)}(x) + a_{n-1} y^{(n-1)}(x) + \dots + a_0 y(x) + b = 0 \quad (41)$$

aus $L_i = z + L_h$ gegeben ist, wobei z eine beliebige Lösung der inhomogenen Gleichung ist. Also braucht man, um den Lösungsraum von (41) zu finden, einfach eine einzige Lösung von (41) zu finden. Das ist aber einfach. Ist z.B. $a_0 \neq 0$, dann kann man einfach z als die konstante Funktion $z(x) = -b/a_0$ wählen. Allgemeiner, sei $j \in \{0, 1, \dots, n\}$ der kleinste Index mit $a_j \neq 0$. Dann ist $z(x) = -(b/a_j)x^j/j!$ eine Lösung von (41). Der Lösungsraum von (41) ist also aus

$$L_i = \{-(b/a_j)x^j/j!\} + L_h$$

gegeben. Die eindeutige Lösung von (41), unter den Anfangsbedingungen $y(x_0) = y_0$, $y'(x_0) = y_1, \dots, y^{(n-1)}(x_0) = y_{n-1}$, kann man also bestimmen, indem man Konstanten $\alpha_{m,\ell}$ findet, so dass

$$y(x_0) = -\frac{b}{a_j} \frac{x^j}{j!} + \sum_{m=1}^k \sum_{\ell=0}^{n_m-1} \alpha_{m,\ell} x^\ell e^{\lambda_m x}$$

die Anfangsbedingungen erfüllt.

Beispiel: Sei

$$y'' + 2\gamma y + \omega_0^2 y = 0$$

Die Gleichung kann als $p(d/dx)y = 0$ geschrieben werden, mit dem Polynom

$$p(s) = s^2 + 2\gamma s + \omega_0^2$$

Das Polynom p hat die Nullstellen

$$s = -\gamma \pm \sqrt{\gamma^2 - \omega_0^2}$$

Ist $|\gamma| > |\omega|$, dann sind die zwei Nullstellen verschiedenen und reell. Damit sind

$$y_1(x) = e^{(-\gamma + \sqrt{\gamma^2 - \omega_0^2})x} \quad \text{und} \quad y_2(x) = e^{(-\gamma - \sqrt{\gamma^2 - \omega_0^2})x}$$

eine Basis für den Lösungsraum. Ist dagegen $|\gamma| = |\omega_0|$, dann hat p die einzige reelle Nullstelle $s_0 = -\gamma$. Damit sind

$$y_1(x) = e^{-\gamma x} \quad \text{und} \quad y_2(x) = x e^{-\gamma x}$$

eine Basis für den Lösungsraum. Ist dagegen $|\gamma| < |\omega_0|$, dann sind die zwei Nullstellen von p komplex. In diesem Fall sind

$$y_1(x) = \cos(\sqrt{\omega_0^2 - \gamma^2} x) e^{-\gamma x} \quad \text{und} \quad y_2(x) = \sin(\sqrt{\omega_0^2 - \gamma^2} x) e^{-\gamma x}$$

eine Basis des Lösungsraums.

2.6 Grundlagen der Stabilitätstheorie

In diesem Abschnitt untersuchen wir die Abhängigkeit der Lösung einer Differentialgleichung von der Anfangsbedingung und der Form der Gleichung. In vielen Situationen erwartet man, dass eine kleine Änderung der Anfangsbedingungen nur eine kleine Änderung der Lösung erzeugen kann. Analog, falls wir das Anfangswertproblem $y'(x) = f(x, y(x)), y(x_0) = y_0$ untersuchen, dann erwarten wir, dass eine kleine Änderung der Funktion f nur eine kleine Änderung der Lösung produziert. Mit anderen Worten, wir untersuchen die Stabilität der Lösung von Anfangswertproblemen. Da Differentialgleichungen nur selten explizit gelöst werden können, ist ihre Stabilität sehr wichtig in praktischen Anwendungen. In der Physik passiert es zum Beispiel sehr oft, dass man eine komplizierte Differentialgleichung $y'(x) = f(x, y(x))$ durch eine einfachere Differentialgleichung $y'(x) = f_0(x, y(x))$ ersetzt, wobei die Differenz $f - f_0$ in geeignetem Sinne klein ist. Damit diese Approximation nützlich sein kann, muss aber die Differenz der zwei Lösungen $y(x)$ und $y_0(x)$ (unter geeigneten Anfangsbedingungen) klein sein. Wir brauchen also die Stabilität der Gleichung ohne, dass wir die Gleichung explizit lösen können.

Ein wichtiges Hilfsmittel um Stabilität zu beweisen ist das Lemma von Gronwall.

Lemma 2.10 (Gronwall Lemma). *Sei $I = [x_0; x_1]$ mit $x_1 > x_0$, $a, b \in \mathbb{R}$ mit $b \geq 0$, $y \in C(I)$, so dass*

$$y(x) \leq a + b \int_{x_0}^x y(t) dt$$

für alle $x \in I$. Dann gilt

$$y(x) \leq a \exp(b(x - x_0))$$

für alle $x \in I$.

Bemerkung:

- Das Lemma von Gronwall zeigt insbesondere, dass jede Lösung der Differentialungleichung $y'(x) \leq by(x)$, mit der Anfangsbedingung $y(x_0) = y_0$, aus der Lösung der Differentialgleichung $y'(x) = by(x)$, mit der selben Anfangsbedingung, nach oben beschränkt wird. Sei in der Tat $y \in C^1([x_0; x_1])$ so, dass $y'(x) \leq by(x)$ und $y(x_0) = y_0$. Dann gilt

$$y(x) = y(x_0) + \int_{x_0}^x y'(t) dt \leq y_0 + b \int_{x_0}^x y(t) dt$$

Lemma 2.10 zeigt dann, dass

$$y(x) \leq y_0 e^{b(x-x_0)}$$

was genau die eindeutige Lösung der Gleichung $y'(x) = by(x)$ mit der Anfangsbedingung $y(x_0) = y_0$ ist.

- Die differentielle Form von dem Gronwall-Lemma, die wir in der ersten Bemerkung diskutiert haben, gilt (im Gegensatz zur Integralform im Lemma 2.10) auch für $b < 0$. In diesem Fall zeigt das Lemma von Gronwall, dass jede $y \in C^1([x_0; x_1])$, mit $y'(x) \leq by(x)$ und $y(x_0) = y_0$ exponentiell abfällt, für $x > x_0$.

Beweis: Sei $\varphi : I \rightarrow \mathbb{R}$ durch $\varphi(x) = y(x)e^{-b(x-x_0)}$ definiert. Dann muss

$$\varphi(x) \leq ae^{-b(x-x_0)} + be^{-b(x-x_0)} \int_{x_0}^x y(t) dt =: \psi(x)$$

Es gilt

$$\psi'(x) = -bae^{-b(x-x_0)} - b^2 \int_{x_0}^x \varphi(t)e^{b(t-x)} + b\varphi(x) = -b[\psi(x) - \varphi(x)] \leq 0$$

für alle $x \in I$. Damit gilt $\varphi(x) \leq \psi(x) \leq \psi(x_0) = a$ für alle $x \in I$. \square

Im nächsten Satz benutzen wir Gronwall-Lemma, um die Stabilität von gewöhnlichen Differentialgleichungen bzgl. Variationen der Anfangsbedingungen zu beweisen.

Satz 2.11. Sei $\Omega \subset \mathbb{R} \times \mathbb{R}^n$ offen, $f \in C(\Omega; \mathbb{R}^n)$ Lipschitz-stetig im zweiten Argument. D.h. es existiere $L > 0$ mit

$$\|f(x, y_1) - f(x, y_2)\| \leq L\|y_1 - y_2\|$$

für alle $x \in \mathbb{R}$, $y_1, y_2 \in \mathbb{R}^n$ mit $(x, y_1), (x, y_2) \in \Omega$. Seien $\varphi, \psi \in C^1(I; \mathbb{R}^n)$ zwei Lösungen der Differentialgleichung

$$y'(x) = f(x, y(x))$$

auf einem Intervall $I \subset \mathbb{R}$. Dann gilt, für beliebige $x_0, x_1 \in I$,

$$\|\varphi(x_1) - \psi(x_1)\| \leq \|\varphi(x_0) - \psi(x_0)\| e^{L|x_1-x_0|}$$

Bemerkung: Insbesondere folgt aus Satz 2.11, dass die Lösung $\varphi(x)$ des Anfangswertproblems

$$\begin{cases} \varphi'(x) = f(x, \varphi(x)) \\ \varphi(x_0) = y_0 \end{cases}$$

stetig (sogar Lipschitz-stetig) von der Anfangsbedingung y_0 abhängt.

Beweis: Sei, o.B.d.A, $x > x_0$. Wir setzen $y(x) = \varphi(x) - \psi(x)$. Aus

$$y'(x) = \varphi'(x) - \psi'(x) = f(x, \varphi(x)) - f(x, \psi(x))$$

folgt

$$y(x) = y(x_0) + \int_{x_0}^x (f(t, \varphi(t)) - f(t, \psi(t))) dt$$

und damit

$$\begin{aligned} \|y(x)\| &\leq \|y(x_0)\| + \int_{x_0}^x \|f(t, \varphi(t)) - f(t, \psi(t))\| dt \\ &\leq \|y(x_0)\| + L \int_{x_0}^x \|y(t)\| dt \end{aligned}$$

Anwendung von Lemma 2.10 auf die Funktion $g(x) = \|y(x)\|$ impliziert, dass

$$\|y(x)\| \leq \|y(x_0)\| e^{L(x-x_0)}$$

\square

Analog kann man auch die Stabilität von Differentialgleichungen bzgl. Änderungen (Störungen) der Gleichung zeigen. Wir betrachten die Differentialgleichung $y'(x) = f(x, \varphi(x), z)$, die von einem zusätzlichen Parameter $z \in \mathbb{R}^m$ abhängt. Für jede $z \in \mathbb{R}^m$ haben wir eine andere Differentialgleichung. Die Frage, die wir im nächsten Satz untersuchen, ist, wie ändert sich die Lösung der Differentialgleichung, falls wir den Parameter z (und eventuell auch die Anfangsbedingung) ein bisschen variieren.

Satz 2.12. Sei $I = [x_0; x_1] \subset \mathbb{R}$, $f \in C(I \times \mathbb{R}^n \times \mathbb{R}^m; \mathbb{R}^n)$. Es existieren $M, L > 0$ mit

$$\|f(x, y_1, z_1) - f(x, y_2, z_2)\| \leq L\|y_1 - y_2\| + M\|z_1 - z_2\|$$

für alle $x \in I$, $y_1, y_2 \in \mathbb{R}^n$, $z_1, z_2 \in \mathbb{R}^m$. Seien $y_1, y_2 \in \mathbb{R}^n$ und $z_1, z_2 \in \mathbb{R}^m$ beliebig gewählt und seien $\varphi_1, \varphi_2 \in C^1(I; \mathbb{R}^n)$ die eindeutigen Lösungen von den Anfangswertproblemen

$$\begin{cases} \varphi_1'(x) = f(x, \varphi_1(x), z_1) \\ \varphi_1(x_0) = y_1 \end{cases} \quad \text{und} \quad \begin{cases} \varphi_2'(x) = f(x, \varphi_2(x), z_2) \\ \varphi_2(x_0) = y_2 \end{cases}$$

Dann gilt

$$\|\varphi_1(x) - \varphi_2(x)\| \leq (\|y_1 - y_2\| + M|x_1 - x_0|\|z_1 - z_2\|) e^{L|x-x_0|}$$

für alle $x \in I$.

Bemerkung: Wählen wir $z_1 = z_2$, dann sind wir zurück bei Satz 2.11.

Beweis: Sei $y(x) = \varphi_1(x) - \varphi_2(x)$. Aus

$$\begin{aligned} \varphi_1(x) &= y_1 + \int_{x_0}^x f(t, \varphi_1(t), z_1) dt \\ \varphi_2(x) &= y_2 + \int_{x_0}^x f(t, \varphi_2(t), z_2) dt \end{aligned}$$

finden wir

$$y(x) = (y_1 - y_2) + \int_{x_0}^x (f(t, \varphi_1(t), z_1) - f(t, \varphi_2(t), z_2))$$

und damit

$$\begin{aligned} \|y(x)\| &\leq \|y_1 - y_2\| + \int_{x_0}^x \|f(t, \varphi_1(t), z_1) - f(t, \varphi_2(t), z_2)\| \\ &\leq \|y_1 - y_2\| + \int_{x_0}^x (L\|y(t)\| + M\|z_1 - z_2\|) \\ &\leq (\|y_1 - y_2\| + M\|z_1 - z_2\|\|x_1 - x_0\|) + L \int_{x_0}^x \|y(t)\| \end{aligned}$$

für alle $x \in I$. Lemma 2.10 impliziert also, dass

$$\|y(x)\| \leq (\|y_1 - y_2\| + \|z_1 - z_2\|\|x_1 - x_0\|) e^{L(x-x_0)}$$

für alle $x > x_0$. □

Satz 2.11 zeigt die orbitale Stabilität der Differentialgleichung $y'(x) = f(x, y(x))$, unter geeigneter Annahme an f . Orbitale Stabilität bedeutet, dass falls wir zwei Anfangsbedingungen y_1, y_2 betrachten, mit $\|y_1 - y_2\|$ klein, dann bleibt der Abstand $\|y_1(x) - y_2(x)\|$ klein, für alle festen $x \in \mathbb{R}$. Der Fehler kann aber im Limes $x \rightarrow \infty$ gross werden. Eine stärkere Form von Stabilität ist die sogenannte asymptotische Stabilität. In diesem Fall bleibt der Fehler klein, gleichmässig in x und verschwindet, da $x \rightarrow \infty$. Ein besonderes Beispiel von asymptotischer Stabilität hat man bei Attraktoren.

Definition 2.13. Sei $\Omega \subset \mathbb{R}^n$ offen, $f \in C(\Omega; \mathbb{R}^n)$. Ein Punkt $y^* \in \Omega$ heisst ein Attraktor, falls eine offene Umgebung V von y^* existiert, so dass für alle $y_0 \in V$, das Anfangswertproblem

$$\begin{cases} y'(x) = f(y(x)) \\ y(x_0) = y_0 \end{cases}$$

eine eindeutige Lösung $y \in C^1([x_0; \infty); \mathbb{R}^n)$ hat, mit der Eigenschaft

$$\lim_{x \rightarrow \infty} y(x) = y^* .$$

Ist $y^* \in \mathbb{R}^n$ ein Attraktor, dann kann man sich einfach überzeugen, dass $f(y^*) = 0$ (sonst würde $y'(x) \rightarrow f(y^*) \neq 0$ und $y(x)$ könnte nicht konvergieren. Beweis: Übung). Andererseits, jede $y^* \in \mathbb{R}^n$ mit $f(y^*) = 0$ definiert eine stationäre (d.h. konstante) Lösung $y(x) = y^*$ von der Gleichung $y'(x) = f(y)$. Die Frage, ob ein solcher Punkt ein Attraktor ist, ist gerade die Frage, ob die stationäre Lösung $y(x) = y^*$ asymptotisch stabil ist. Man findet, dass die Stabilität von stationären Lösungen mit der Ableitung von f an der Stelle y^* zu tun hat. Das diskutieren wir im nächsten Satz, für den skalaren Fall $n = 1$ (eine analoge Aussage gilt auch für $n > 1$; in diesem Fall braucht man aber Kenntnisse aus der Differentialrechnung in mehreren Veränderlichen, die wir noch nicht haben).

Satz 2.14. Sei $f \in C^1(\mathbb{R})$, mit $\sup_{x \in \mathbb{R}} |f'(x)| < \infty$, und $y^* \in \mathbb{R}$ mit $f(y^*) = 0$ und $f'(y^*) < 0$. Dann ist y^* ein Attraktor für die Differentialgleichung $y'(x) = f(y(x))$.

Beweis: O.B.d.A. können wir annehmen, dass $y^* = 0$. Sei $f'(0) = -\lambda$, für ein $\lambda > 0$. Wir schreiben (da $f(0) = 0$)

$$\begin{aligned} f(y) &= f(0) + \int_0^1 \frac{d}{dt} f(ty) dt = y \int_0^1 f'(ty) dt \\ &= f'(0)y + y \int_0^1 (f'(ty) - f'(0)) dt = f'(0)y + yR(y) \end{aligned}$$

mit

$$R(y) = \int_0^1 (f'(ty) - f'(0)) dt$$

Die Stetigkeit von f' impliziert, dass $|R(y)| \rightarrow 0$ für $y \rightarrow 0$. Wir finden also ein $\varepsilon > 0$ mit

$$\sup_{y \in [-\varepsilon; \varepsilon]} |R(y)| \leq \lambda/2.$$

Für ein beliebiges $x_0 \in \mathbb{R}$, sei nun $y \in C^1([x_0; \infty))$ die Lösung der Differentialgleichung $y'(x) = f(y(x))$ mit der Anfangsbedingung $y(x_0) = y_0$ für ein $y_0 \in [-\varepsilon/2; \varepsilon/2]$ (bemerke, dass Satz 2.7 die Existenz einer globalen Lösung für dieses Anfangswertproblem garantiert).

Wir behaupten, dass $|y(x)| \leq \varepsilon$ für alle $x \in [x_0; \infty)$. Ist das nicht der Fall, so setzen wir

$$x_1 = \inf\{x \in [x_0; \infty) : |y(x)| > \varepsilon\}$$

Aus der Stetigkeit von $y(x)$ gilt dann $|y(x)| \leq \varepsilon$ für alle $x \in [x_0; x_1]$. Wir setzen nun $\varphi(x) = y(x) \exp(-f'(0)(x - x_0))$, für alle $x > x_0$. Es gilt dann $\varphi(x_0) = y_0$ und

$$\begin{aligned} \varphi'(x) &= (y'(x) - f'(0)y(x)) \exp(-f'(0)(x - x_0)) \\ &= (f(y(x)) - f'(0)y(x)) \exp(-f'(0)(x - x_0)) \\ &= y(x)R(y(x)) \exp(-f'(0)(x - x_0)) = R(y(x))\varphi(x) \end{aligned} \quad (42)$$

Da $|y(x)| \leq \varepsilon$ für alle $x \in [x_0; x_1]$, ist $|R(y(x))| \leq \lambda/2$ für alle $x \in [x_0; x_1]$. Damit ist $|\varphi'(x)| \leq (\lambda/2)|\varphi(x)|$ für alle $x \in [x_0; x_1]$. Die Identität

$$\varphi(x) = \varphi(x_0) + \int_{x_0}^x \varphi'(t) dt$$

zeigt also, dass

$$|\varphi(x)| \leq |y_0| + \int_{x_0}^x |\varphi'(t)| dt \leq |y_0| + (\lambda/2) \int_{x_0}^x |\varphi(t)| dt$$

Aus Lemma 2.10 folgt, dass

$$|\varphi(x)| \leq |y_0| e^{(\lambda/2)(x-x_0)}$$

Das ergibt

$$|y(x)| \leq |y_0| e^{-(\lambda/2)(x-x_0)}$$

für alle $x \in [x_0; x_1]$. Damit ist $|y(x_1)| \leq |y_0| \leq \varepsilon/2$. Aus Stetigkeit von y existiert also ein $\delta > 0$ mit $|y(x)| \leq \varepsilon$ für alle $x \in [x_1 - \delta; x_1 + \delta]$, im Widerspruch zur Definition von x_1 . Das zeigt, dass $|y(x)| \leq \varepsilon$ für alle $x \in \mathbb{R}$.

Aus (42) folgt also, dass $|\varphi'(x)| \leq (\lambda/2)|\varphi(x)|$ für alle $x > x_0$. Das ergibt $|\varphi(x)| \leq |y_0| \exp((\lambda/2)(x - x_0))$ für alle $x > x_0$ und

$$|y(x)| \leq |y_0| e^{-(\lambda/2)(x-x_0)}$$

für alle $x > x_0$. Das zeigt, dass $y(x) \rightarrow 0$, für $x \rightarrow \infty$, für alle Anfangsbedingungen y_0 mit $|y_0| \leq \varepsilon/2$. \square

3 Differentialrechnung in mehreren Veränderlichen

3.1 Definition der Ableitung für Funktionen auf \mathbb{R}^n

Wiederholung von Begriffen aus der linearen Algebra und Analysis 1. In Analysis 1 (siehe Definition 2.34) haben wir den Begriff vom Vektorraum definiert. Ein Vektorraum über \mathbb{R} ist nämlich eine Menge V , versehen mit einer Addition $+$: $V \times V \rightarrow V$ und einer skalaren Multiplikation \cdot : $\mathbb{R} \times V \rightarrow V$, die eine Reihe von Axiomen erfüllen. Ein normierter Vektorraum ist ein Vektorraum V , auf dem eine Abbildung $\|\cdot\|$: $V \rightarrow \mathbb{R}$ definiert ist, mit den Eigenschaften: i) $\|x\| \geq 0$ für alle $x \in V$, $\|x\| = 0$ genau dann,

wenn $x = 0$; ii) $\|\alpha x\| = |\alpha|\|x\|$ für alle $x \in V$ und $\alpha \in \mathbb{R}$; iii) $\|x + y\| \leq \|x\| + \|y\|$. Eine Norm $\|\cdot\|$ auf einem Vektorraum V erzeugt immer eine Metrik auf V , die durch $d(x, y) = \|x - y\|$ definiert wird. Also ist jeder normierte Vektorraum ein metrischer Raum. Wir sagen der normierte Vektorraum V ist vollständig, falls V , versehen mit der aus der Norm induzierten Metrik, ein vollständiger metrischer Raum ist.

Seien nun V und W zwei Vektorräume über \mathbb{R} . Eine Abbildung $L : V \rightarrow W$ heisst linear, falls $L(x + \lambda y) = L(x) + \lambda L(y)$ für alle $x, y \in V$ und $\lambda \in \mathbb{R}$. Wir bezeichnen die Menge aller linearen Abbildungen $L : V \rightarrow W$ mit $\tilde{\mathcal{L}}(V, W)$. Seien $L, M \in \tilde{\mathcal{L}}(V, W)$ und $\lambda \in \mathbb{R}$. Wir definieren dann die Abbildungen $L + M, \lambda L : V \rightarrow W$ durch

$$(L + M)(x) = L(x) + M(x), \quad \text{und} \quad (\lambda L)(x) = \lambda L(x)$$

für alle $x \in V$. Offenbar gilt $L + M, \lambda L \in \tilde{\mathcal{L}}(V, W)$. Damit hat $\tilde{\mathcal{L}}(V, W)$ die Struktur eines Vektorraumes über \mathbb{R} (es ist leicht zu überprüfen, dass Summe und skalare Multiplikation alle notwendigen Axiome erfüllen). Nehmen wir nun an, V, W seien normierte Vektorräume. Für $L \in \tilde{\mathcal{L}}(V, W)$ definieren wir dann

$$\|L\|_{\text{op}} := \sup_{v \in V \setminus \{0\}} \frac{\|Lv\|}{\|v\|} = \sup_{v \in V, \|v\| \leq 1} \|Lv\| = \sup_{v \in V, \|v\|=1} \|Lv\| \quad (43)$$

Eine lineare Abbildung $L \in \tilde{\mathcal{L}}(V, W)$ heisst beschränkt, falls $\|L\|_{\text{op}} < \infty$. Das ist leicht zu zeigen: Eine Abbildung $L \in \tilde{\mathcal{L}}(V, W)$ ist genau dann beschränkt, falls sie auf V stetig ist. Ferner, eine Abbildung $L \in \tilde{\mathcal{L}}(V, W)$ ist genau dann auf V stetig, falls sie in $v = 0$ stetig ist. Wir bezeichnen

$$\mathcal{L}(V, W) := \{L \in \tilde{\mathcal{L}}(V, W) : L \text{ beschränkt ist}\} = \{L \in \tilde{\mathcal{L}}(V, W) : L \text{ stetig ist}\}$$

die Teilmenge von $\tilde{\mathcal{L}}(V, W)$ die aus beschränkten Abbildungen besteht. Es ist einfach zu zeigen, dass $\mathcal{L}(V, W)$ ein linearer Unterraum von $\tilde{\mathcal{L}}(V, W)$ ist. Also ist $\mathcal{L}(V, W)$ selbst ein Vektorraum. Man kann dann leicht beweisen, dass (43) eine Norm auf $\mathcal{L}(V, W)$ definiert. Man nennt $\|L\|_{\text{op}}$ die Operatornorm von L . $\mathcal{L}(V, W)$, versehen mit der Norm $\|\cdot\|_{\text{op}}$ ist also ein normierter Vektorraum. Tatsache: Ist W vollständig, so ist auch $\mathcal{L}(V, W)$ ein vollständig normierter Vektorraum, unabhängig davon, ob V vollständig ist oder nicht (ein vollständiger normierter Vektorraum heisst ein Banach-Raum).

Sind V, W, X drei Vektorräume, und $L \in \tilde{\mathcal{L}}(V, W)$ und $M \in \tilde{\mathcal{L}}(W, X)$ zwei lineare Abbildungen, so können wir die Verknüpfung $L \circ M : V \rightarrow X$ durch $(L \circ M)(v) := L(M(v))$ definieren. Wir bezeichnen oft die Verknüpfung $L \circ M$ als $L \cdot M$ oder einfach als LM . Es ist leicht zu sehen, dass $LM \in \tilde{\mathcal{L}}(V, X)$ eine lineare Abbildung ist. Sind ferner V, W, X drei normierte Vektorräume und $L \in \tilde{\mathcal{L}}(V, W)$ und $M \in \tilde{\mathcal{L}}(W, X)$ beschränkt, dann ist auch LM beschränkt und

$$\|LM\|_{\text{op}} \leq \|L\|_{\text{op}} \|M\|_{\text{op}}$$

Also, für jede $L \in \mathcal{L}(V, W)$ und $M \in \mathcal{L}(W, X)$, ist $LM \in \mathcal{L}(V, X)$. Insbesondere, für jede $L, M \in \mathcal{L}(V, V)$ ist $LM \in \mathcal{L}(V, V)$. Das definiert ein Produkt auf dem Vektorraum $\mathcal{L}(V, V)$.

Sei V ein Vektorraum. Eine lineare Abbildung $L \in \tilde{\mathcal{L}}(V, \mathbb{R})$ heisst ein lineares Funktional auf V . Sei V ein normierter Vektorraum. Der Raum $\mathcal{L}(V, \mathbb{R})$ aller stetigen linearen Funktionalen auf V heisst der Dualraum von V und wird oft mit V^* bezeichnet. Versehen mit der Operatornorm

$$\|Lv\| = \sup_{v \in V, \|v\| \leq 1} |Lv|$$

ist V^* ein normierter Vektorraum. Da \mathbb{R} vollständig ist, ist V^* immer vollständig.

Wir werden in dieser Vorlesung nur endlich dimensionale Vektorräume betrachten. Jeder Vektorraum V mit $\dim V = n < \infty$ ist isomorph zu \mathbb{R}^n (ein Isomorphismus ist nach Wahl einer Basis von V gegeben). Auf \mathbb{R}^n ist die standard euklidische Norm durch

$$\|(x_1, \dots, x_n)\|^2 = \sum_{j=1}^n |x_j|^2$$

definiert. Auf \mathbb{R}^n ist auch ein Skalarprodukt definiert. Für $x = (x_1, \dots, x_n)$ und $y = (y_1, \dots, y_n)$ setzen wir

$$x \cdot y = \sum_{j=1}^n x_j y_j$$

Dann gilt $\|x\|^2 = x \cdot x$. Wir haben in Analysis 1 bewiesen, dass \mathbb{R}^n , versehen mit der euklidischen Norm $\|\cdot\|$, ein vollständiger normierter Vektorraum ist. Tatsache: Auf einem endlich dimensionalen Vektorraum sind alle zwei Normen äquivalent. D.h., falls $\|\cdot\|_1, \|\cdot\|_2$ zwei Normen auf \mathbb{R}^n sind, dann es gibt eine Konstante $c > 0$ mit

$$\frac{1}{c} \|x\|_1 \leq \|x\|_2 \leq c \|x\|_1$$

für alle $x \in \mathbb{R}^n$. Die Äquivalenz der zwei Normen impliziert, dass eine Folge x_n auf \mathbb{R}^n genau dann bzg. $\|\cdot\|_1$ konvergiert, wenn sie bzg. $\|\cdot\|_2$ konvergiert und dass eine Folge x_n auf \mathbb{R}^n genau dann bzg. $\|\cdot\|_1$ eine Cauchy-Folge ist, wenn sie bzg. $\|\cdot\|_2$ eine Cauchy-Folge ist. Da \mathbb{R}^n versehen mit der Standardnorm vollständig ist, ist \mathbb{R}^n bzg. einer beliebigen Norm ein vollständiger Vektorraum (es ist natürlich möglich, auf \mathbb{R}^n eine Metrik d zu finden, so dass (\mathbb{R}^n, d) nicht vollständig ist; eine solche Metrik d kann dann aber nicht von einer Norm induziert werden).

Sei $L : \mathbb{R}^n \rightarrow \mathbb{R}^m$ eine lineare Abbildung. Dann existiert eine $m \times n$ Matrix $\tilde{L} = (\ell_{ij})$, so dass $L(x) = \tilde{L}x$, wobei das Produkt $\tilde{L}x$ durch

$$(\tilde{L}x)_i = \sum_{j=1}^n \ell_{ij} x_j$$

definiert ist. Wir identifizieren deswegen die lineare Abbildung L mit der entsprechenden Matrix \tilde{L} . Jede lineare Abbildung $L : \mathbb{R}^n \rightarrow \mathbb{R}^m$ ist beschränkt. In der Tat, falls wir auch mit $L = (\ell_{ij})$ die Matrix bezeichnen, die der Abbildung L zugeordnet ist, so gilt $\|L\|_{\text{op}} \leq \|L\|$, wobei

$$\|L\| = \text{Tr } L^* L = \sum_{i=1}^m \sum_{j=1}^n |\ell_{ij}|^2 < \infty$$

die Standardnorm von L ist, falls wir L als ein Element von $\mathbb{R}^{m \times n}$ betrachten. Es folgt, dass jede lineare Abbildung zwischen \mathbb{R}^n und \mathbb{R}^m automatisch stetig ist, d.h. $\tilde{\mathcal{L}}(\mathbb{R}^n, \mathbb{R}^m) = \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$.

Der Dualraum zu \mathbb{R}^n ist der Vektorraum $(\mathbb{R}^n)^* = \mathcal{L}(\mathbb{R}^n, \mathbb{R})$ aller linearen Funktionalen auf \mathbb{R}^n (weil jedes lineare Funktional auf \mathbb{R}^n stetig ist). Ein beliebiges lineares Funktional L auf \mathbb{R}^n kann also mit einer $1 \times n$ Matrix identifiziert werden. Mit anderen Worten, jedes lineare Funktional L auf \mathbb{R}^n kann mit einem Vektor $a \in \mathbb{R}^n$ identifiziert werden, so dass

$$L(x) = a \cdot x = \sum_{j=1}^n a_j x_j$$

wobei $a \cdot x$ das Skalarprodukt auf \mathbb{R}^n bezeichnet. Man kann sich leicht davon überzeugen, dass für ein solches lineares Funktional,

$$\|L\|_{\text{op}} = \sup_{x \in \mathbb{R}^n, \|x\| \leq 1} |L(x)| = \sup_{x \in \mathbb{R}^n, \|x\| \leq 1} |a \cdot x| = \|a\|$$

wobei $\|a\|$ die Standardnorm auf \mathbb{R}^n bezeichnet. Es folgt, dass $(\mathbb{R}^n)^* \simeq \mathbb{R}^n$ als normierte Vektorräume identifiziert werden können.

Partielle Ableitungen. Nach dieser kurzen Wiederholung aus der linearen Algebra, sind wir bereit, den Begriff von Ableitung auf mehrdimensionale Vektorräume zu definieren. Wir werden der Einfachheit halber Funktionen betrachten, die auf einer Teilmenge von \mathbb{R}^n definiert sind, mit Werten auf \mathbb{R}^m . Die Definitionen können aber einfach auf Funktionen zwischen zwei beliebigen (endlich dimensionalen) Vektorräumen verallgemeinert werden (weil jeder endlich dimensionale Vektorraum isomorph zu \mathbb{R}^n ist). Auf \mathbb{R}^n werden wir immer die euklidische Standardnorm betrachten. Es ist aber einfach, die Definitionen auf beliebige andere Normen zu erweitern (weil jede Norm zur Standardnorm äquivalent ist). Also, obwohl wir nur Funktionen $f : \mathbb{R}^n \supset U \rightarrow \mathbb{R}^m$ betrachten werden, kann man den Begriff von Ableitung für beliebige Funktionen zwischen zwei normierten endlich dimensional Vektorräumen definieren.

Definition 3.1. Sei $U \subset \mathbb{R}^n$ offen, $x_0 \in U$, $i \in \{1, \dots, n\}$. Eine Funktion $f : U \rightarrow \mathbb{R}$ heisst im Punkt x_0 partiell differenzierbar in der i -ten Koordinate, falls der Limes

$$\frac{\partial f}{\partial x_i}(x_0) := \lim_{h \rightarrow 0} \frac{f(x_0 + h e_i) - f(x_0)}{h}$$

existiert. Hier sind die Vektoren e_1, \dots, e_n die Standardbasis von \mathbb{R}^n , d.h. wir haben $e_i = (0, \dots, 0, 1, 0, \dots, 0)$. In diesem Fall heisst die Zahl $\partial f / \partial x_i(x_0)$ die i -te partielle Ableitung von f an der Stelle x_0 . Die Funktion f heisst an der Stelle $x_0 \in U$ partiell differenzierbar, falls f in jeder Koordinate an der Stelle x partiell differenzierbar ist. Die Funktion f heisst auf U partiell differenzierbar, falls f an der Stelle x partiell differenzierbar ist, für jede $x \in U$.

Bemerkung: Die partielle Ableitung in der i -ten Koordinate ist die gewöhnliche Ableitung bezüglich der i -ten Variablen von f , wenn die anderen $(n-1)$ Koordinaten konstant gehalten werden. D.h. die i -te partielle Ableitung von f an der Stelle $x^* = (x_1^*, \dots, x_n^*) \in$

\mathbb{R}^n ist die Ableitung der Funktion einer Variablen $t \rightarrow f(x_1^*, x_2^*, \dots, x_i^* + t, x_{i+1}^*, \dots, x_n^*)$ an der Stelle $t = 0$.

Bemerkung: Analog kann man die partielle Ableitungen einer vektorwertigen Funktion definieren. Sei wie oben $U \subset \mathbb{R}^n$ offen und $f : U \rightarrow \mathbb{R}^m$. Für $j = 1, \dots, m$, sei $f_m : U \rightarrow \mathbb{R}$ die m -te Komponente von f ; d.h. es gelte $f(x) = (f_1(x), \dots, f_m(x))$. Dann sagen wir, dass f im Punkt x_0 in der i -ten Koordinate partiell differenzierbar ist, falls f_j an der Stelle x_0 in der i -ten Koordinate partiell differenzierbar ist, für alle $j = 1, \dots, m$. In diesem Fall ist die i -te partielle Ableitung von f aus dem Vektor

$$\frac{\partial f}{\partial x_i}(x_0) = \left(\frac{\partial f_1}{\partial x_i}(x_0), \dots, \frac{\partial f_m}{\partial x_i}(x_0) \right) \in \mathbb{R}^m$$

gegeben.

Es stellt sich heraus, dass der Begriff von partieller Differenzierbarkeit ein bisschen zu schwach ist. Viele Resultate, die wir für differenzierbare Funktionen auf \mathbb{R} kennen, gelten für eine auf einer offenen Teilmenge $U \subset \mathbb{R}^n$ definierte, partiell differenzierbare Funktion f nicht. Z.B. zeigt das folgende Beispiel, dass partielle Differenzierbarkeit einer Funktion nicht ihre Stetigkeit impliziert.

Beispiel: Auf \mathbb{R}^2 definieren wir die Funktion

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

Wir behaupten, dass f auf \mathbb{R}^2 partiell differenzierbar ist. In der Tat ist f offenbar an der Stelle (x, y) in der ersten Koordinate partiell differenzierbar für alle $(x, y) \neq (0, 0)$. Wir behaupten, f auch an der Stelle $(0, 0)$ in der ersten Koordinate partiell differenzierbar ist. In der Tat

$$\lim_{h \rightarrow 0} \frac{f(h, 0) - f(0, 0)}{h} = \lim_{h \rightarrow 0} \frac{0}{h} = 0$$

Analog kann man zeigen, dass f überall in der zweiten Koordinate partiell differenzierbar ist. Also ist f auf \mathbb{R}^2 partiell differenzierbar. Wir behaupten nun, dass f an der Stelle $(0, 0)$ nicht stetig ist. In der Tat haben wir

$$\lim_{k \rightarrow \infty} f(1/k, 1/k) = \lim_{k \rightarrow \infty} \frac{1/k^2}{1/k^2 + 1/k^2} = \frac{1}{2} \neq 0$$

Also, obwohl die Folge $(1/k, 1/k) \rightarrow (0, 0)$ konvergiert, ist $f(1/k, 1/k) \not\rightarrow f(0, 0)$.

Ableitung auf \mathbb{R}^n . Wir brauchen also einen stärkeren Begriff von Differenzierbarkeit für Funktionen, die auf Teilmengen von \mathbb{R}^n definiert sind. Um den richtigen Begriff zu finden, möchten wir zunächst den Begriff von Differenzierbarkeit für Funktionen einer Variablen umschreiben. Sei $U \subset \mathbb{R}$ offen, und $x_0 \in U$. Eine Funktion $f : U \rightarrow \mathbb{R}$ ist an der Stelle x_0 falls der Grenzwert

$$\lim_{h \rightarrow 0} \frac{f(x_0 + h) - f(x_0)}{h}$$

existiert. Es folgt: f ist in x_0 differenzierbar, falls ein Zahl $L \in \mathbb{R}$ existiert s.d.

$$|f(x_0 + h) - f(x_0) - Lh| = o(|h|)$$

im Limes $h \rightarrow 0$. Das bedeutet, f ist an der Stelle x_0 differenzierbar, falls sich f in der Nähe von x_0 durch eine lineare Funktion approximieren lässt. Ist das der Fall, so ist L eindeutig bestimmt und aus $L = f'(x_0)$ gegeben. Dieser Begriff lässt sich nun auf Funktionen verallgemeinern, die auf einer offenen Teilmenge von \mathbb{R}^n definiert sind.

Definition 3.2. Sei $U \subset \mathbb{R}^n$ offen, $f : U \rightarrow \mathbb{R}^m$ und $x_0 \in U$. Die Funktion f heisst an der Stelle x_0 differenzierbar, wenn eine lineare Abbildung $L : \mathbb{R}^n \rightarrow \mathbb{R}^m$ existiert, so dass

$$\lim_{h \rightarrow 0} \frac{f(x_0 + h) - f(x_0) - L(h)}{\|h\|} = 0$$

In diesem Fall heisst die Abbildung L die Ableitung oder das Differential von f an der Stelle x_0 und wird mit $L = Df(x_0)$ bezeichnet. f heisst auf U differenzierbar, falls f an der Stelle x differenzierbar ist, für alle $x \in U$.

Bemerkung: Ist f an der Stelle x_0 differenzierbar, so ist die Ableitung $Df(x_0)$ eindeutig bestimmt. Gilt in der Tat

$$\lim_{h \rightarrow 0} \frac{f(x_0 + h) - f(x_0) - L(h)}{\|h\|} = 0, \quad \text{und} \quad \lim_{h \rightarrow 0} \frac{f(x_0 + h) - f(x_0) - M(h)}{\|h\|} = 0$$

so muss

$$\lim_{h \rightarrow 0} \frac{\|L(h) - M(h)\|}{\|h\|} = 0$$

Da

$$\frac{\|Lx - Mx\|}{\|x\|} = \frac{\|L(x/K) - M(x/K)\|}{\|x/K\|}$$

für alle $K > 0$, erhalten wir

$$\frac{\|Lx - Mx\|}{\|x\|} = \lim_{K \rightarrow \infty} \frac{\|L(x/K) - M(x/K)\|}{\|x/K\|} = \lim_{h \rightarrow 0} \frac{\|L(h) - M(h)\|}{\|h\|} = 0$$

für alle $x \in \mathbb{R}^n$. Das bedeutet

$$\|L - M\|_{\text{op}} = \sup_{x \in \mathbb{R}^n \setminus \{0\}} \frac{\|Lx - Mx\|}{\|x\|} = 0$$

und deswegen $L = M$.

Bemerkung: Ist f an der Stelle x_0 differenzierbar, so ist die Ableitung $Df(x_0) : \mathbb{R}^n \rightarrow \mathbb{R}^m$ eine lineare Abbildung. Man kann also $Df(x_0)$ mit einer $n \times m$ Matrix identifizieren. Wie für jede lineare Abbildung zwischen endlich dimensionale Vektorräume, ist $Df(x_0)$ beschränkt, $\|Df(x_0)\|_{\text{op}} < \infty$, und damit auch stetig.

Bemerkung: Es folgt, dass eine Funktion f ist an der Stelle x_0 differenzierbar, falls sie sich lokal durch eine lineare Abbildung approximieren lässt, d.h. falls $L \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ existiert, s.d.

$$\|f(x_0 + h) - f(x_0) - Lh\| = o(\|h\|)$$

Bemerkung: Sei $U \subset \mathbb{R}^n$ offen. Sei $f : U \rightarrow \mathbb{R}^m$ mit Komponenten $f_1, \dots, f_m : U \rightarrow \mathbb{R}$. D.h. es gelte $f(x_1, \dots, x_n) = (f_1(x_1, \dots, x_n), \dots, f_m(x_1, \dots, x_n))$ für alle $(x_1, \dots, x_n) \in U$.

U . Dann ist f an der Stelle $x_0 \in U$ genau dann differenzierbar, wenn f_j an der Stelle x_0 differenzierbar ist, für alle $j = 1, \dots, m$. Das folgt aus der Tatsache, dass eine Folge $x^{(n)} = (x_1^{(n)}, \dots, x_m^{(n)}) \in \mathbb{R}^m$ genau dann gegen $x = (x_1, \dots, x_m) \in \mathbb{R}^m$ konvergiert, wenn $x_j^{(n)} \rightarrow x_j$ für alle $j = 1, \dots, m$.

Satz 3.3. Sei $U \subset \mathbb{R}^n$ offen, $x_0 \in U$, und $f : U \rightarrow \mathbb{R}^m$ an der Stelle x_0 differenzierbar. Dann ist f an der Stelle x_0 stetig.

Beweis: Sei L die Ableitung von f an der Stelle x_0 . Wir schreiben

$$f(x_0 + h) - f(x_0) = [f(x_0 + h) - f(x_0) - L(h)] + L(h)$$

Da f differenzierbar an der Stelle x_0 ist, gilt $\|f(x_0 + h) - f(x_0) - L(h)\| \rightarrow 0$ für $h \rightarrow 0$. Andererseits, $\|L(h)\| \leq \|L\|_{\text{op}} \|h\| \rightarrow 0$ für $h \rightarrow 0$. Also

$$\|f(x_0 + h) - f(x_0)\| \leq \|f(x_0 + h) - f(x_0) - L(h)\| + \|L(h)\| \rightarrow 0$$

für $h \rightarrow 0$. Das zeigt, dass f an der Stelle x_0 stetig ist. □

Richtungsableitungen. Ist f an der Stelle x_0 differenzierbar, so existieren alle partiellen Ableitungen von f an der Stelle x_0 . Ferner existieren alle Richtungsableitungen.

Proposition 3.4. Sei $U \subset \mathbb{R}^n$ offen, $x_0 \in U$ und $f : U \rightarrow \mathbb{R}^m$ an der Stelle x_0 differenzierbar. Dann

$$\lim_{t \rightarrow 0} \frac{f(x_0 + tv) - f(x_0)}{t} = Df(x_0)(v)$$

für alle $v \in \mathbb{R}^n$ (hier ist $t \in \mathbb{R}$) und insbesondere existiert der Grenzwert auf der linken Seite. Man nennt den Grenzwert auf der linken Seite die Richtungsableitung von f in der Richtung v .

Beweis: Aus Differenzierbarkeit folgt, dass

$$\lim_{h \rightarrow 0} \frac{\|f(x_0 + h) - f(x_0) - L(h)\|}{\|h\|} = 0$$

Insbesondere, falls $h = tv$ für ein festes $v \in \mathbb{R}^n$ und $t \in \mathbb{R}$, gilt (weil, wegen Linearität, $L(tv) = tL(v)$)

$$0 = \lim_{t \rightarrow 0} \frac{\|f(x_0 + tv) - f(x_0) - L(tv)\|}{\|tv\|} = \|v\|^{-1} \lim_{t \rightarrow 0} \left\| \frac{f(x_0 + tv) - f(x_0)}{t} - L(v) \right\|$$

und damit

$$\lim_{t \rightarrow 0} \frac{f(x_0 + tv) - f(x_0)}{t} = L(v)$$

□

Wählen wir $v = e_j$, dann impliziert Proposition 3.4, dass alle partiellen Ableitungen $(\partial f / \partial x_j)(x_0)$ für $j = 1, \dots, n$ existieren. Es folgt auch, dass die partielle Ableitung $(\partial f / \partial x_j)(x_0)$ die j -te Kolumne der Matrix $Df(x_0)$ ist. Mit anderen Worten, sei $f : U \rightarrow \mathbb{R}^m$, mit Komponenten $f_1, \dots, f_m : U \rightarrow \mathbb{R}$, so dass $f(x) = (f_1(x), f_2(x), \dots, f_m(x))$ für alle $x \in U$. Sei f an der Stelle x_0 differenzierbar. Dann kann die lineare Abbildung $Df(x_0) : \mathbb{R}^n \rightarrow \mathbb{R}^m$ durch die $m \times n$ Matrix mit Einträge

$$(Df(x_0))_{i,j} = \frac{\partial f_i}{\partial x_j}(x_0) \quad \text{für } i = 1, \dots, m \text{ und } j = 1, \dots, n \quad (44)$$

dargestellt werden. Diese Matrix heisst die Funktionalmatrix, oder die Jacobi-Matrix von f an der Stelle x_0 .

Wie wir schon gemerkt haben, impliziert die Existenz der Jacobi-Matrix nicht, dass f an der Stelle x_0 differenzierbar ist. Man findet aber, dass Existenz und Stetigkeit der partiellen Ableitungen die Differenzierbarkeit von f implizieren. Dieses Kriterium ist wichtig, weil es uns erlaubt, die Differenzierbarkeit von Funktionen, einfach durch Untersuchung der partiellen Ableitungen, zu beweisen.

Proposition 3.5. *Sei $U \subset \mathbb{R}^n$ offen, $f : U \rightarrow \mathbb{R}^m$. Ferner, nehmen wir an, dass die partiellen Ableitungen $\partial f / \partial x_j(x)$ auf U existieren und stetig sind. Dann ist f auf U differenzierbar.*

Beweis: O.B.d.A. betrachten wir den Fall $m = 1$ (die Differenzierbarkeit von $f(x) = (f_1(x), \dots, f_m(x))$ ist mit der Differenzierbarkeit von f_1, \dots, f_m äquivalent). Der Einfachheit halber untersuchen wir zunächst den Fall $n = 2$. Wir nehmen an $0 = (0, 0) \in U$, und wir zeigen die Differenzierbarkeit in diesem Punkt. Sei $h = (h_1, h_2)$ so klein, dass $B_{\|h\|}(0) \subset U$. Wir schreiben

$$f(h_1, h_2) - f(0, 0) = f(h_1, h_2) - f(h_1, 0) + f(h_1, 0) - f(0, 0)$$

Da die Abbildung $y \rightarrow f(h_1, y)$ stetig differenzierbar ist (aus Existenz und Stetigkeit der partiellen Ableitung in der y -Richtung), können wir schreiben

$$\begin{aligned} f(h_1, h_2) - f(h_1, 0) &= \int_0^{h_2} dy \frac{\partial f}{\partial y}(h_1, y) \\ &= \frac{\partial f}{\partial y}(0, 0)h_2 + \int_0^{h_2} dy \left(\frac{\partial f}{\partial y}(h_1, y) - \frac{\partial f}{\partial y}(0, 0) \right) \end{aligned}$$

Analog ist $x \rightarrow f(x, 0)$ stetig differenzierbar. Deswegen

$$\begin{aligned} f(h_1, 0) - f(0, 0) &= \int_0^{h_1} dx \frac{\partial f}{\partial x}(x, 0) \\ &= \frac{\partial f}{\partial x}(0, 0) + \int_0^{h_1} dx \left(\frac{\partial f}{\partial x}(x, 0) - \frac{\partial f}{\partial x}(0, 0) \right) \end{aligned}$$

Also haben wir

$$\begin{aligned} f(h_1, h_2) - f(0, 0) &= \frac{\partial f}{\partial x}(0, 0)h_1 + \frac{\partial f}{\partial y}(0, 0)h_2 \\ &= \int_0^{h_1} dx \left(\frac{\partial f}{\partial x}(x, 0) - \frac{\partial f}{\partial x}(0, 0) \right) + \int_0^{h_2} dy \left(\frac{\partial f}{\partial y}(h_1, y) - \frac{\partial f}{\partial y}(0, 0) \right) \end{aligned}$$

Die Differenzierbarkeit von f an der Stelle $(0, 0)$ folgt, falls wir zeigen können, dass

$$\int_0^{h_1} dx \left(\frac{\partial f}{\partial x}(x, 0) - \frac{\partial f}{\partial x}(0, 0) \right) + \int_0^{h_2} dy \left(\frac{\partial f}{\partial y}(h_1, y) - \frac{\partial f}{\partial y}(h_1, 0) \right) = o(\|(h_1, h_2)\|)$$

da $(h_1, h_2) \rightarrow 0$. Sei also $\varepsilon > 0$ fest. Da die partiellen Ableitungen stetig sind, finden wir $\delta > 0$ so dass

$$\left| \frac{\partial f}{\partial x}(x, y) - \frac{\partial f}{\partial x}(0, 0) \right| \leq \frac{\varepsilon}{2} \quad \text{und} \\ \left| \frac{\partial f}{\partial y}(x, y) - \frac{\partial f}{\partial y}(0, 0) \right| \leq \frac{\varepsilon}{2}$$

für alle $(x, y) \in \mathbb{R}^2$ mit $\|(x, y)\| \leq \delta$. Sei nun $\|(h_1, h_2)\| \leq \delta$. Dann gilt auch $\|(h_1, y)\| \leq \delta$, für alle $0 \leq y \leq h_2$ (angenommen $h_2 > 0$, sonst ist die Aussage war für alle $h_2 \leq y \leq 0$). Damit gilt

$$\left| \frac{\partial f}{\partial y}(h_1, y) - \frac{\partial f}{\partial y}(0, 0) \right| \leq \frac{\varepsilon}{2}$$

für alle $0 \leq y \leq h_2$ und also

$$\left| \int_0^{h_2} dy \left(\frac{\partial f}{\partial y}(h_1, y) - \frac{\partial f}{\partial y}(0, 0) \right) \right| \leq \frac{\varepsilon|h_2|}{2} \leq \frac{\varepsilon\|h\|}{2}$$

Ähnlich gilt $\|(x, 0)\| \leq \delta$ für alle $0 \leq x \leq h_1$ (oder $h_1 \leq x \leq 0$, falls $h_1 < 0$). Deswegen

$$\left| \frac{\partial f}{\partial x}(x, 0) - \frac{\partial f}{\partial x}(0, 0) \right| \leq \frac{\varepsilon}{2}$$

für alle $0 \leq x \leq h_1$ und also

$$\left| \int_0^{h_1} dx \left(\frac{\partial f}{\partial x}(x, 0) - \frac{\partial f}{\partial x}(0, 0) \right) \right| \leq \frac{\varepsilon|h_1|}{2} \leq \frac{\varepsilon\|h\|}{2}$$

Es folgt, dass für alle $\varepsilon > 0$ ein $\delta > 0$ existiert, so dass

$$\frac{1}{\|(h_1, h_2)\|} \left| \int_0^{h_1} dx \left(\frac{\partial f}{\partial x}(x, 0) - \frac{\partial f}{\partial x}(0, 0) \right) + \int_0^{h_2} dy \left(\frac{\partial f}{\partial y}(h_1, y) - \frac{\partial f}{\partial y}(0, 0) \right) \right| \leq \varepsilon$$

für alle $\|(h_1, h_2)\| \leq \delta$. Das zeigt die Behauptung. Die Verallgemeinerung zu $n \geq 3$ lassen wir als Übung. \square

Der Gradient. Sei $U \subset \mathbb{R}^n$, offen und $f : U \rightarrow \mathbb{R}$ eine reel-wertige Abbildung, differenzierbar an der Stelle $a \in U$. Das Differential von f an der Stelle a ist dann eine lineare Abbildung $Df(a) : \mathbb{R}^n \rightarrow \mathbb{R}$ und kann mit einer $1 \times n$ Matrix identifiziert werden. Mit anderen Worten, $Df(a)$ ist ein lineares Funktional auf \mathbb{R}^n . Wie jedes lineare Funktional auf \mathbb{R}^n kann $Df(a)$ mit einem Vektor $v = (v_1, \dots, v_n) \in \mathbb{R}^n$, mit der Eigenschaft, dass

$$Df(a)(y) = v \cdot y = \sum_{j=1}^n v_j y_j$$

für alle $y = (y_1, \dots, y_n) \in \mathbb{R}^n$ identifiziert werden. Man nennt den Vektor v den Gradienten von f an der Stelle a und man benutzt die Notation $v = \nabla f(a)$. Nach (44) sind die Komponenten vom Gradient aus

$$\nabla f(a) = \left(\frac{\partial f}{\partial x_1}(a), \dots, \frac{\partial f}{\partial x_n}(a) \right)$$

gegeben. Für einen beliebigen Einheitsvektor $e \in \mathbb{R}^n$ gilt, nach Proposition 3.4,

$$\frac{d}{dt}f(a + te)|_{t=0} = Df(a)(e) = \nabla f(a) \cdot e$$

D.h. die Zuwachsrate der Funktion f in der Richtung e ist aus dem Skalarprodukt $\nabla f(a) \cdot e$ gegeben. Nehmen wir an $\nabla f(a) \neq 0$. Das Skalarprodukt $\nabla f(a) \cdot e$ ist dann maximal über allen möglichen Einheitsvektoren $e \in \mathbb{R}^n$, mit $\|e\| = 1$, falls $e = \nabla f(a)/\|\nabla f(a)\|$. Für $e = \nabla f(a)/\|\nabla f(a)\|$ gilt dann

$$\frac{d}{dt}f(a + te) = \nabla f(a) \cdot \frac{\nabla f(a)}{\|\nabla f(a)\|} = \|\nabla f(a)\|$$

Wir haben bewiesen, dass der Vektor $\nabla f(a)$ in die Richtung der grössten Zuwachsrate der Funktion f an der Stelle a zeigt. Die Länge von $\nabla f(a)$ ist dann genau die grösste Zuwachsrate von f an der Stelle a (das gilt auch, falls $\nabla f(a) = 0$).

Stetige Differenzierbarkeit. Sei $U \subset \mathbb{R}^n$ offen. Eine Funktion $f : U \rightarrow \mathbb{R}^m$ heisst auf U stetig differenzierbar, falls die Ableitung $Df(x)$ existiert, für alle $x \in U$, und falls die Abbildung $Df : U \rightarrow \mathcal{L}(\mathbb{R}^n; \mathbb{R}^m)$ stetig ist. Aus Proposition 3.5 folgt, dass f auf U genau dann stetig differenzierbar ist, wenn die partielle Ableitung $\partial f_i / \partial x_j(x)$ für alle $i = 1, \dots, m$ und alle $j = 1, \dots, n$ auf U existiert und stetig ist. Wir setzen

$$C^1(U; \mathbb{R}^m) := \{f : U \rightarrow \mathbb{R}^m : f \text{ auf } U \text{ stetig differenzierbar ist}\}.$$

Rechenregeln. Wir sammeln in der nächsten Proposition ein paar nützliche elementare Regeln für die Berechnung von Ableitungen von Funktionen mit mehreren Veränderlichen.

Proposition 3.6. *Sei $U \subset \mathbb{R}^n$ offen, $a \in U$.*

- a) *Seien $f, g : U \rightarrow \mathbb{R}^m$ differenzierbar an der Stelle a und $\lambda \in \mathbb{R}$. Dann ist auch $f + \lambda g : U \rightarrow \mathbb{R}^m$ differenzierbar an der Stelle a und $D(f + \lambda g)(a) = Df(a) + \lambda Dg(a)$.*
- b) *Ist f konstant auf U , so gilt $Df(x) = 0$ für alle $x \in U$.*
- c) *Ist $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ linear, dann gilt $Df(x) = f$, für alle $x \in \mathbb{R}^n$.*

Beweis: Teil (a) und (b) sind offenbar, nach Definition der Ableitung und Linearität des Limes. Zu Teil (c) bemerken wir, dass $f(x + h) = f(x) + f(h)$. Das impliziert, dass

$$\|f(x + h) - f(a) - f(h)\| = 0 = o(\|h\|)$$

Damit ist die lineare Abbildung $L = f$ die Ableitung von f an der Stelle x . □

Kettenregel. Die Ableitung der Verknüpfung zweier Funktionen kann durch die Kettenregel berechnet werden.

Satz 3.7. Sei $U \subset \mathbb{R}^n$ und $G \subset \mathbb{R}^p$ offen. $f : U \rightarrow \mathbb{R}^p$, $g : G \rightarrow \mathbb{R}^m$, mit $f(U) \subset G$. Sei $a \in U$, f differenzierbar an der Stelle a , g differenzierbar an der Stelle $f(a)$. Dann ist die Funktion $g \circ f : U \rightarrow \mathbb{R}^m$ differenzierbar an der Stelle $a \in U$, und

$$D(g \circ f)(a) = Dg(f(a)) \cdot Df(a)$$

wobei das Produkt auf der rechten Seite die Komposition der zwei linearen Abbildungen $Df(a) : \mathbb{R}^n \rightarrow \mathbb{R}^p$ und $Dg(f(a)) : \mathbb{R}^p \rightarrow \mathbb{R}^m$ ist. Mit anderen Worten, die $m \times n$ Matrix $D(g \circ f)(a)$ ist aus dem Produkt der $m \times p$ Matrix $Dg(f(a))$ mit der $p \times n$ Matrix $Df(a)$ gegeben.

Beweis: Sei $b = f(a)$, $L = Df(a)$, $M = Dg(b)$. Für $h \in \mathbb{R}^n$, $\tilde{h} \in \mathbb{R}^p$ setzen wir

$$\eta_1(h) = f(a+h) - f(a) - L(h), \quad \text{und} \quad \eta_2(\tilde{h}) = g(b+\tilde{h}) - g(b) - M(\tilde{h})$$

Nach Differenzierbarkeit von f an der Stelle a und von g an der Stelle b , gilt

$$\|\eta_1(h)\| = o(\|h\|), \quad \text{und} \quad \|\eta_2(\tilde{h})\| = o(\|\tilde{h}\|)$$

für $h, \tilde{h} \rightarrow 0$. Sei nun $h \in \mathbb{R}^n$ beliebig und $\tilde{h} = L(h) + \eta_1(h)$. Dann gilt

$$b + \tilde{h} = f(a) + L(h) + \eta_1(h) = f(a+h)$$

Also

$$\begin{aligned} (g \circ f)(a+h) &= g(f(a+h)) = g(b+\tilde{h}) = g(b) + M(\tilde{h}) + \eta_2(\tilde{h}) \\ &= g(f(a)) + M(L(h)) + M(\eta_1(h)) + \eta_2(\tilde{h}) \end{aligned}$$

Die Behauptung folgt, falls wir zeigen können, dass i) $\|M(\eta_1(h))\| = o(\|h\|)$ und ii) $\|\eta_2(\tilde{h})\| = o(\|h\|)$ für $h \rightarrow 0$. Um i) zu zeigen, bemerken wir einfach, dass

$$\frac{\|M(\eta_1(h))\|}{\|h\|} \leq \|M\|_{\text{op}} \frac{\|\eta_1(h)\|}{\|h\|} \rightarrow 0$$

für $h \rightarrow 0$, weil $\eta_1(h) = o(\|h\|)$. Andererseits, um ii) zu beweisen, benutzen wir, dass $\|\eta_1(h)\| \leq \|h\|$ für $\|h\|$ klein genug (weil $\eta_1(h) = o(\|h\|)$). Deswegen gilt

$$\|\tilde{h}\| = \|L(h) + \eta_1(h)\| \leq \|L(h)\| + \|\eta_1(h)\| \leq (\|L\| + 1)\|h\|$$

für $\|h\|$ klein genug. Da $\eta_2(\tilde{h}) = o(\|\tilde{h}\|)$ existiert, für ein beliebiges $\varepsilon > 0$ ein $\delta > 0$ mit

$$\|\eta_2(\tilde{h})\| \leq \varepsilon \|\tilde{h}\| \leq (\|L\| + 1)\varepsilon \|h\|$$

für alle $h \in \mathbb{R}^n$ mit $\|h\| \leq \delta$. Das bedeutet, dass $\eta_2(\tilde{h}) = o(\|h\|)$ und zeigt ii). \square

Beispiel. Sei $q : \mathbb{R} \rightarrow \mathbb{R}^3$ die Bahn eines Teilchens als Funktion der Zeit. Sei $T : \mathbb{R} \times \mathbb{R}^3 \rightarrow \mathbb{R}$ die Temperatur als Funktion von der Zeit und von der Position im Raum. Die vom Teilchen zur Zeit t gespürte Temperatur ist aus der Funktion $t \rightarrow T(t, q(t))$ gegeben. Sind q und T differenzierbar, so ist auch $t \rightarrow T(t, q(t))$ differenzierbar, mit

$$\frac{d}{dt}T(t, q(t)) = \frac{\partial T}{\partial t}(t, q(t)) + \sum_{j=1}^3 \frac{\partial T}{\partial x_j}(t, q(t))q'_j(t)$$

Das Resultat folgt mit der Definition $f : \mathbb{R} \rightarrow \mathbb{R}^4$ durch $f(t) = (t, q_1(t), q_2(t), q_3(t))$. Nach Differenzierbarkeit von q ist auch f differenzierbar, mit

$$f'(t) = (1, q'_1(t), q'_2(t), q'_3(t))$$

Da $T(t, q(t)) = (T \circ f)(t)$ folgt, dass

$$(T \circ f)'(t) = DT(f(t)) \cdot f'(t) = (\nabla T)(f(t)) \cdot f'(t) = \frac{\partial T}{\partial t}(t, q(t)) + \sum_{j=1}^3 \frac{\partial T}{\partial x_j}(t, q(t))q'_j(t)$$

3.2 Mittelwertsatz

Für eine auf $[a; b]$ stetige und auf $(a; b)$ differenzierbare Funktion $f : [a; b] \rightarrow \mathbb{R}$ besagt der Mittelwertsatz, dass ein $\xi \in (a; b)$ existiert, mit $f(b) - f(a) = f'(\xi)(b - a)$. Wir zeigen hier eine analoge Aussage für Funktionen mit mehreren Veränderlichen.

Satz 3.8 (Mittelwertsatz). *Sei $U \subset \mathbb{R}^n$ offen, $f : U \rightarrow \mathbb{R}$ differenzierbar. Seien $a, b \in U$ mit*

$$[a; b] = \{(1 - \lambda)a + \lambda b : \lambda \in [0; 1]\} \subset U$$

Dann gibt es ein $\xi \in [a; b]$ ($\xi \neq a, b$) mit

$$f(b) - f(a) = \nabla f(\xi) \cdot (b - a)$$

Beweis: Sei $\phi : [0, 1] \rightarrow \mathbb{R}^n$ durch $\phi(t) = (1 - t)a + tb$ definiert. Sei $\psi : [0; 1] \rightarrow \mathbb{R}$ durch $\psi(t) = f(\phi(t))$ definiert. Nach der Kettenregel ist ψ ist dann auf $[0; 1]$ stetig und auf $(0; 1)$ differenzierbar, mit

$$\psi'(t) = \nabla f(\phi(t)) \cdot \phi'(t) = \nabla f((1 - t)a + tb) \cdot (b - a)$$

Aus dem Mittelwertsatz für Funktionen einer Variablen existiert $t_0 \in (0; 1)$ mit

$$f(b) - f(a) = \psi(1) - \psi(0) = \psi'(t_0)(1 - 0) = \nabla f((1 - t_0)a + t_0b) \cdot (b - a)$$

Die Behauptung folgt, mit $\xi = (1 - t_0)a + t_0b$. □

Für Funktionen mit Werten auf \mathbb{R}^m , $m > 1$, gilt i.A. der Mittelwertsatz nicht (unabhängig davon, ob die Funktion eine oder mehrere Veränderliche hat; siehe Bemerkung unter Satz 8.9 in Analysis 1). Man kann aber eine Mittelwertabschätzung zeigen (siehe Proposition 8.10 in Analysis 1 für die Mittelwertabschätzung für Funktionen einer Variablen).

Satz 3.9 (Mittelwertabschätzung). Sei $U \subset \mathbb{R}^n$ offen, $f : U \rightarrow \mathbb{R}^m$ differenzierbar, $a, b \in U$ mit

$$[a; b] = \{(1 - \lambda)a + \lambda b : \lambda \in [0; 1]\} \subset U$$

Sei $\|Df(x)\|_{\text{op}} \leq M$ für alle $x \in [a; b]$. Dann gilt

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

Beweis: O.B.d.A. können wir annehmen, dass $f(b) \neq f(a)$. Wir setzen

$$e = \frac{f(b) - f(a)}{\|f(b) - f(a)\|} \in \mathbb{R}^m.$$

Wir definieren die lineare Funktion $\phi : \mathbb{R}^m \rightarrow \mathbb{R}$ durch $\phi(x) = x \cdot e$. Wir bemerken, dass, wegen Linearität, ϕ auf \mathbb{R}^m differenzierbar ist, mit $D\phi = \phi$. Ferner, da $\|e\| = 1$, gilt $\|\phi\|_{\text{op}} \leq 1$ (eigentlich $\|\phi\|_{\text{op}} = 1$). Wir definieren auch die Funktion $g : \mathbb{R}^n \rightarrow \mathbb{R}$ durch

$$g(x) = \phi(f(x) - f(a))$$

Es gilt $g(a) = 0$ und $g(b) = \|f(b) - f(a)\|$. Aus der Kettenregel ist g auf U differenzierbar, mit

$$Dg(x) = D\phi(f(x) - f(a)) \cdot Df(x) = \phi \cdot Df(x)$$

Für $x \in [a; b]$ gilt also

$$\|Dg(x)\|_{\text{op}} \leq \|\phi\|_{\text{op}} \|Df(x)\|_{\text{op}} \leq M$$

Der Mittelwertsatz 3.8 für die Funktion g impliziert, dass ein $\xi \in [a; b]$ mit

$$\|f(b) - f(a)\| = g(b) - g(a) = Dg(\xi) \cdot (b - a) \leq \|Dg(\xi)\|_{\text{op}} \|b - a\| \leq M\|b - a\|$$

existiert. □

Eine Anwendung der Mittelwertabschätzung ist der Beweis der Tatsache, dass eine auf einem offenen und zusammenhängenden Gebiet $U \subset \mathbb{R}^n$ definierte Funktion f mit $Df = 0$ auf U konstant sein muss.

Definition 3.10. Ein Streckenzug auf \mathbb{R}^n ist eine Menge der Form

$$[a_1; a_2] \cup [a_2; a_3] \cup \dots \cup [a_{p-1}; a_p]$$

für ein $p \in \mathbb{N}$, und für Punkten $a_1, \dots, a_n \in \mathbb{R}^n$ (hier bezeichnet $[a; b] = \{(1 - \lambda)a + \lambda b : \lambda \in [0; 1]\}$ der Segment zwischen a und b). a_0 heisst Anfangspunkt und a_n heisst Endpunkt des Streckenzuges. Wir sagen dann der Streckenzug verbindet die Punkten a_0 und a_p . Eine offene Teilmenge $U \subset \mathbb{R}^n$ heisst zusammenhängend falls je zwei Punkten in U durch einen Streckenzug in U verbinden werden können.

Proposition 3.11. Sei $U \subset \mathbb{R}^n$ offen und zusammenhängend, $f : U \rightarrow \mathbb{R}^m$ differenzierbar, mit $Df(x) = 0$ für alle $x \in U$. Dann ist f konstant auf U .

Beweis: Ist $[a; b] \subset U$ dann gilt, aus Satz 3.9 mit $M = 0$, $f(b) = f(a)$. Seien nun $x, y \in U$ beliebig. Da U zusammenhängend ist gibt es ein Streckenzug $[a_0; a_1] \cup \dots \cup [a_{p-1}; a_p]$ in U , mit $a_0 = x$ und $a_p = y$. Also

$$f(x) = f(a_1) = f(a_2) = \dots = f(a_{p-1}) = f(y).$$

□

Bemerkung: die Annahme, dass U zusammenhängend ist, ist notwendig. Sei

$$U = \{x \in \mathbb{R}^2 : |x| < 1 \quad \text{oder} \quad |x - 3| < 1\}$$

und $f : U \rightarrow \mathbb{R}$ durch $f(x) = 0$ falls $|x| < 1$ und $f(x) = 1$ falls $|x - 3| < 1$ definiert. Dann ist U offen, und $Df(x) = 0$ für alle $x \in U$, aber f ist auf U nicht konstant.

3.3 Höhere Ableitungen, Taylor Entwicklung, lokale Extrema

Sei $U \subset \mathbb{R}^n$ offen, und $f : U \rightarrow \mathbb{R}^m$ differenzierbar. Die Ableitung von f ist dann eine Funktion $Df : U \rightarrow \mathcal{L}(\mathbb{R}^n; \mathbb{R}^m)$ mit Werten in den linearen Abbildungen zwischen \mathbb{R}^n und \mathbb{R}^m . Man kann $\mathcal{L}(\mathbb{R}^n; \mathbb{R}^m)$ mit dem Vektorraum \mathbb{R}^{mn} identifizieren (weil jede lineare Abbildung in $\mathcal{L}(\mathbb{R}^n; \mathbb{R}^m)$ mit einer $m \times n$ Matrix identifiziert werden kann). Man kann sich also fragen, ob die Abbildung Df differenzierbar ist. Ist Df an der Stelle $a \in U$ differenzierbar, dann heisst f an der Stelle a zweimal differenzierbar. Die zweite Ableitung ist eine lineare Abbildung $D^2f(a) : \mathbb{R}^n \rightarrow \mathbb{R}^{mn}$, d.h. $D^2f(a) \in \mathcal{L}(\mathbb{R}^n; \mathbb{R}^{n \times m}) \simeq \mathbb{R}^{mn^2}$. Iterativ kann man höhere Ableitungen definieren. Ist f auf U k -mal differenzierbar, und ist die k -te Ableitung $D^k f : U \rightarrow \mathbb{R}^{mn^k}$ an der Stelle a differenzierbar, dann sagt man, dass f an der Stelle a $(k + 1)$ -mal differenzierbar ist, und man bezeichnet die $(k + 1)$ -te Ableitung mit $D^{k+1} f(a) \in \mathcal{L}(\mathbb{R}^n; \mathbb{R}^{mn^k}) \simeq \mathbb{R}^{mn^{k+1}}$.

Man bemerke, dass die Abbildung $D^2f(a) : \mathbb{R}^n \rightarrow \mathcal{L}(\mathbb{R}^n; \mathbb{R}^m) \simeq \mathbb{R}^{n \times m}$ mit der bilinearen Abbildung $D^2f(a) : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}^m$, definiert durch

$$(D^2f(a))(v, v') = (D^2f(a)(v))(v'),$$

identifiziert werden kann (wir benutzen die selbe Notation $D^2f(a)$ für die lineare Abbildung $\mathbb{R}^n \rightarrow \mathcal{L}(\mathbb{R}^n; \mathbb{R}^m)$ und für die bilineare Abbildung $\mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}^m$). Wir erinnern hier, dass eine Abbildung $b : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}^m$ bilinear heisst, falls die zwei Bedingungen

$$\begin{aligned} b(v_1 + \lambda v_2, v) &= b(v_1, v) + \lambda b(v_2, v) && \text{(Linearität im ersten Argument)} \\ b(v, v_1 + \lambda v_2) &= b(v, v_1) + \lambda b(v, v_2) && \text{(Linearität im zweiten Argument)} \end{aligned} \tag{45}$$

für alle $v, v_1, v_2 \in \mathbb{R}^n$, $\lambda \in \mathbb{R}$ erfüllt sind. Analog kann $D^3f(a)$ mit einer trilinearen Form identifiziert werden und $D^k f(a)$ mit einer k -linearen Abbildung auf \mathbb{R}^n , mit Werten in \mathbb{R}^m .

Höhere partielle Ableitungen. Sei nun $f : U \rightarrow \mathbb{R}$ partiell differenzierbar. Dann ist $\partial f / \partial x_j$ wieder eine Funktion auf U mit Werten in \mathbb{R} . Ist diese neue Funktion partiell differenzierbar, so können wir die partiellen Ableitungen zweiter Ordnung

$$\frac{\partial^2 f}{\partial x_j \partial x_i} = \frac{\partial}{\partial x_j} \left(\frac{\partial f}{\partial x_i} \right)$$

definieren. Iterativ kann man partielle Ableitungen höherer Ordnung definieren (die Ordnung einer partiellen Ableitung ist die gesamte Anzahl von partiellen Ableitungen). Um die Notation ein bisschen zu vereinfachen, schreiben wir, für eine partielle Ableitung der Ordnung k ,

$$\partial_{i_1} \partial_{i_2} \dots \partial_{i_k} f = \frac{\partial^k f}{\partial x_{i_1} \partial x_{i_2} \dots \partial x_{i_k}} = \frac{\partial}{\partial x_{i_1}} \left(\frac{\partial}{\partial x_{i_2}} \left(\dots \frac{\partial}{\partial x_{i_{k-1}}} \left(\frac{\partial f}{\partial x_{i_k}} \right) \right) \right).$$

Wir sagen, die Funktion $f : U \rightarrow \mathbb{R}$ ist k -mal partiell differenzierbar, falls alle partiellen Ableitungen der Ordnung kleiner oder gleich k existieren. Wir sagen, dass eine Funktion $f : U \rightarrow \mathbb{R}^m$ k -mal partiell differenzierbar ist, falls $f(x) = (f_1(x), \dots, f_m(x))$ und jede Komponente $f_1, \dots, f_m : U \rightarrow \mathbb{R}$ k -mal partiell differenzierbar ist.

Für $k \in \mathbb{N}$, $k \geq 1$, bezeichnen wir mit $C^k(U; \mathbb{R}^m)$ die Menge der Funktionen $f : U \rightarrow \mathbb{R}^m$, die auf U k -mal differenzierbar sind, so dass $D^k f$ stetig ist. Nach Proposition 3.5 ist $C^k(U; \mathbb{R}^m)$ genau die Menge der Funktionen $f : U \rightarrow \mathbb{R}^m$, für die alle partiellen Ableitungen der Ordnung kleiner oder gleich k existieren und stetig sind.

Zur Berechnung von höheren partiellen Ableitungen ist es sehr nützlich zu bemerken, dass (unter geeigneten Annahmen an f) sich partielle Ableitungen miteinander vertauschen, d.h. $\partial_i \partial_j f = \partial_j \partial_i f$. Das wird in dem nächsten Satz bewiesen.

Satz 3.12 (Schwarz). *Sei $U \subset \mathbb{R}^2$ offen, $(x_0, y_0) \in U$, und $f : U \rightarrow \mathbb{R}$ in U partiell differenzierbar. Falls $\partial_y \partial_x f$ auf U existiert und an der Stelle (x_0, y_0) stetig ist, dann existiert auch $\partial_x \partial_y f$ an der Stelle (x_0, y_0) und*

$$\partial_x \partial_y f(x_0, y_0) = \partial_y \partial_x f(x_0, y_0).$$

Bemerkung. O.B.d.A. können wir $(x_0, y_0) = (0, 0)$ betrachten. Da

$$\partial_y f(x, 0) = \lim_{y \rightarrow 0} \frac{f(x, y) - f(x, 0)}{y}$$

erhalten wir

$$\begin{aligned} \partial_x \partial_y f(0; 0) &= \lim_{x \rightarrow 0} \frac{\partial_y f(x, 0) - \partial_y f(0, 0)}{x} = \lim_{x \rightarrow 0} \lim_{y \rightarrow 0} \frac{\frac{f(x, y) - f(x, 0)}{y} - \frac{f(0, y) - f(0, 0)}{y}}{x} \\ &= \lim_{x \rightarrow 0} \lim_{y \rightarrow 0} \frac{f(x, y) - f(x, 0) - f(0, y) + f(0, 0)}{xy} \end{aligned} \quad (46)$$

Analog gilt

$$\partial_y \partial_x f(0; 0) = \lim_{y \rightarrow 0} \lim_{x \rightarrow 0} \frac{f(x, y) - f(0, y) - f(x, 0) + f(0, 0)}{xy}$$

Das Problem ist also zu zeigen, dass die zwei Grenzwerte vertauscht werden können.

Beweis: O.B.d.A. nehmen wir an, dass $(x_0, y_0) = (0, 0)$ und dass $\partial_y \partial_x f(0, 0) = 0$ (sonst ersetze $f(x, y)$ durch $f(x, y) - \partial_y \partial_x f(0, 0)xy$). Wir definieren die Funktion $\phi(x, y) = f(x, y) - f(x, 0)$. Gemäss (46) sind wir an

$$\frac{f(x, y) - f(0, y) - f(x, 0) + f(0, 0)}{xy} = \frac{\phi(x, y) - \phi(0, y)}{xy}$$

interessiert. Für festgehaltene y ist ϕ differenzierbar nach x und es gilt

$$\partial_x \phi(x, y) = \partial_x f(x, y) - \partial_x f(x, 0)$$

Der Mittelwertsatz (für Funktionen einer Variablen) zeigt, es existiert $0 < \theta < 1$ mit

$$\phi(x, y) - \phi(0, y) = x \partial_x \phi(\theta x, y) = x [\partial_x f(\theta x, y) - \partial_x f(\theta x, 0)] \quad (47)$$

Nun ist die Funktion $y \rightarrow \partial_x f(\theta x, y)$, für festgehaltene x und θ nach y differenzierbar, mit Ableitung $\partial_y \partial_x f(\theta x, y)$ (wir benutzen hier die Existenz der zweiten partiellen Ableitung $\partial_y \partial_x f$ auf U). Der Mittelwertsatz (für eine Variable) impliziert also, dass ein $0 < \theta' < 1$ existiert, mit

$$\partial_x f(\theta x, y) - \partial_x f(\theta x, 0) = y \partial_y \partial_x f(\theta x, \theta' y)$$

Aus (47) folgt, dass

$$\phi(x, y) - \phi(0, y) = xy \partial_y \partial_x f(\theta x, \theta' y)$$

und damit

$$\frac{f(x, y) - f(0, y) - f(x, 0) + f(0, 0)}{xy} = \partial_y \partial_x f(\theta x, \theta' y)$$

Sei nun $\varepsilon > 0$ beliebig fest gewählt. Da $\partial_y \partial_x f$ an der Stelle $(0, 0)$ stetig ist, und da $\partial_y \partial_x f(0, 0) = 0$, existiert ein $\delta > 0$ mit $|\partial_y \partial_x f(w, z)| \leq \varepsilon$ für alle $(w, z) \in \mathbb{R}^2$ mit $\|(w, z)\| \leq \delta$. Seien also $(x, y) \in \mathbb{R}^2$, mit $\|(x, y)\| \leq \delta$. Dann gilt auch $\|(\theta x, \theta' y)\| \leq \delta$, für alle $\theta, \theta' \in (0, 1)$. Deswegen gilt

$$|\partial_y \partial_x f(\theta x, \theta' y)| \leq \varepsilon$$

und

$$\left| \frac{\frac{f(x, y) - f(x, 0)}{y} - \frac{f(0, y) - f(0, 0)}{y}}{x} \right| = \left| \frac{f(x, y) - f(0, y) - f(x, 0) + f(0, 0)}{xy} \right| \leq \varepsilon$$

Das gilt für alle $\|(x, y)\| \leq \delta$, und also insbesondere für feste $x \in (-\delta, \delta)$ und $y \rightarrow 0$. Da

$$\lim_{y \rightarrow 0} \frac{f(x, y) - f(x, 0)}{y} = \partial_y f(x, 0), \quad \text{und} \quad \lim_{y \rightarrow 0} \frac{f(0, y) - f(0, 0)}{y} = \partial_y f(0, 0)$$

erhalten wir

$$\left| \frac{\partial_y f(x, 0) - \partial_y f(0, 0)}{x} \right| \leq \varepsilon$$

für alle $x \in (-\delta, \delta)$. Da $\varepsilon > 0$ beliebig ist, es folgt, dass

$$\lim_{x \rightarrow 0} \frac{\partial_y f(x, 0) - \partial_y f(0, 0)}{x} = 0$$

(und insbesondere, dass der Grenzwert existiert). Das zeigt, dass $\partial_x \partial_y f(0, 0) = 0$. \square

Durch wiederholte Anwendung von Satz 3.12 bekommen wir das folgende Korollar für partielle Ableitungen beliebiger Ordnung.

Korollar 3.13. Sei $U \subset \mathbb{R}^n$ offen. Sei $k \in \mathbb{N}$ und $f \in C^k(U)$. Dann gilt für alle $i_1, \dots, i_k \in \{1, \dots, n\}$ und alle Permutationen π der Zahlen $\{1, \dots, k\}$,

$$\partial_{i_1} \dots \partial_{i_k} f = \partial_{i_{\pi 1}} \dots \partial_{i_{\pi k}} f.$$

Beispiel: Sei $f \in C^4(U)$. Dann $\partial_{x_1} \partial_{x_1} \partial_{x_2} \partial_{x_2} f = \partial_{x_1} \partial_{x_2} \partial_{x_1} \partial_{x_2} f = \partial_{x_1} \partial_{x_2} \partial_{x_2} \partial_{x_1} f = \dots$

Bemerkung: Nicht nur die verschiedenen partiellen Ableitungen, sondern auch beliebige Richtungsableitungen vertauschen sich miteinander.

Differentialoperatoren. Man kann partielle Ableitungen als Operatoren interpretieren, die auf differenzierbare Funktionen wirken. Sei

$$p(\xi_1, \dots, \xi_n) = \sum_{i_1, \dots, i_n \geq 0: i_1 + \dots + i_n \leq k} p_{i_1, \dots, i_n} \xi_1^{i_1} \dots \xi_n^{i_n}$$

ein Polynom in den n Variablen ξ_1, \dots, ξ_n . Dann definieren wir den entsprechenden Differentialoperator

$$p(\partial_1, \dots, \partial_n) = \sum_{i_1, \dots, i_n \geq 0: i_1 + \dots + i_n \leq k} p_{i_1, \dots, i_n} \partial_1^{i_1} \dots \partial_n^{i_n}$$

Der Operator $p(\partial_1, \dots, \partial_n)$ ist linear und bildet Funktionen in $C^k(U)$ nach Funktionen in $C(U)$. Der Operator ist wohldefiniert aus Satz 3.12, weil die verschiedenen partiellen Ableitungen sich miteinander vertauschen (wäre das nicht der Fall, so würden zwei verschiedene Operatoren dem selben Polynom entsprechen). Eine weitere Folgerung von Satz 3.12 ist die folgende Bemerkung: Seien p_1, p_2 zwei Polynome in n Variablen der Ordnung k_1 und k_2 , sei $p_1 \cdot p_2$ das Produkt der zwei Polynome (ein Polynom in n Variablen der Ordnung $k_1 + k_2$). Dann gilt

$$p_1(\partial_1, \dots, \partial_n) \cdot p_2(\partial_1, \dots, \partial_n) = (p_1 \cdot p_2)(\partial_1, \dots, \partial_n)$$

als Identität zweier Operatoren auf $C^{(k_1+k_2)}(U)$. Das Produkt auf der linken Seite ist die Komposition von zwei (linearen) Abbildungen.

Beispiel: Sei $p(\xi_1, \dots, \xi_n) = \sum_{j=1}^n \xi_j^2$. Der Laplace-Operator auf \mathbb{R}^n ist aus

$$\Delta := p(\partial_1, \dots, \partial_n) = \sum_{j=1}^n \partial_j^2 = \sum_{j=1}^n \frac{\partial^2}{\partial x_j^2}$$

gegeben. Der Laplace-Operator wirkt auf $C^2(\mathbb{R}^n)$.

Taylor Entwicklung. Wir erinnern uns an den Begriff der Taylor-Entwicklung für Funktionen einer Variablen (siehe Kapitel 8.6 im Skript zu Analysis 1). Sei $f \in C^{m+1}([a; x])$. Dann existiert $\xi \in (a; x)$, so dass

$$f(x) = f(a) + f'(a)(x-a) + \dots + \frac{f^{(m)}(a)}{m!} (x-a)^m + \frac{f^{(m+1)}(\xi)}{(m+1)!} (x-a)^{m+1}.$$

Sei nun $U \subset \mathbb{R}^n$ offen, $a \in U$ und $h \in \mathbb{R}^n$ mit $a + h \in U$. Sei weiter $f \in C^{m+1}(U)$ \mathbb{R} -wertig. Wir setzen $\phi(t) = f(a + th)$. Dann ist $\phi \in C^{m+1}([0, 1])$, mit $\phi(1) = f(a + h)$ und $\phi(0) = f(a)$. Das impliziert, dass

$$f(a + h) = f(a) + \sum_{j=1}^m \frac{\phi^{(j)}(0)}{j!} + \frac{\phi^{(m+1)}(\theta)}{(m+1)!}$$

für ein $\theta \in (0; 1)$. Wir müssen die Ableitungen von ϕ berechnen. Es gilt

$$\phi'(t) = Df(a + th)(h) = h \cdot \nabla f(a + th)$$

und deswegen $\phi'(0) = h \cdot \nabla f(a)$. Induktiv bekommen wir

$$\phi^{(j)}(t) = (h \cdot \nabla)^j f(a + th)$$

Der Operator $(h \cdot \nabla)^j$ ist ein Differentialoperator der Ordnung j . Man findet

$$\frac{(h \cdot \nabla)^j}{j!} = \left(\sum_{i=1}^n h_i \partial_i \right)^j = \sum_{i_1, \dots, i_n \geq 0: i_1 + \dots + i_n = j} \frac{h_1^{i_1} \dots h_n^{i_n}}{i_1! i_2! \dots i_n!} \partial_1^{i_1} \dots \partial_n^{i_n}$$

und also die Taylor-Entwicklung

$$\begin{aligned} f(a + h) &= f(a) + \sum_{j=1}^m \sum_{i_1, \dots, i_n \geq 0: i_1 + \dots + i_n = j} \frac{1}{i_1! i_2! \dots i_n!} \left(\partial_1^{i_1} \dots \partial_n^{i_n} f \right) (a) h_1^{i_1} \dots h_n^{i_n} \\ &+ \sum_{i_1, \dots, i_n: i_1 + \dots + i_n = m+1} \frac{1}{i_1! i_2! \dots i_n!} \left(\partial_1^{i_1} \dots \partial_n^{i_n} f \right) (a + \theta h) h_1^{i_1} \dots h_n^{i_n} \end{aligned} \quad (48)$$

Es ist nützlich, eine kompaktere Notation für die höheren partiellen Ableitungen einzuführen. Ein Multiindex ist eine n -Tupel $i = (i_1, \dots, i_n)$ mit $i_j \in \mathbb{N}$ für alle $j = 1, \dots, n$. Der Betrag des Multiindex $i = (i_1, \dots, i_n)$ wird als

$$|i| = i_1 + i_2 + \dots + i_n$$

definiert. Für den Multiindex i definieren wir weiter die partielle Ableitung der Ordnung $|i|$

$$\partial^i := \partial_1^{i_1} \partial_2^{i_2} \dots \partial_n^{i_n}$$

und $h^i := h_1^{i_1} \dots h_n^{i_n}$ für alle $h = (h_1, \dots, h_n) \in \mathbb{R}^n$. Wir setzen auch $i! := i_1! i_2! \dots i_n!$. Dann lässt sich (48) als

$$f(a + h) = \sum_{j=0}^m \sum_{i: |i|=j} \frac{(\partial^i f)(a)}{i!} h^i + \sum_{i: |i|=m+1} \frac{(\partial^i f)(a + \theta h)}{i!} h^i \quad (49)$$

schreiben, für ein beliebiges $f \in C^{m+1}(U)$, $U \subset \mathbb{R}^n$ offen, $a \in U$ und h klein genug. In (49) bilden die ersten m Termen das m -te Taylor-Polynom von f an der Stelle a :

$$p_m(h) = \sum_{j=0}^m \sum_{i: |i|=j} \frac{(\partial^i f)(a)}{i!} h^i$$

Der letzte Term auf der rechten Seite von (49) heisst das Restglied. Das Restglied ist offenbar $O(\|h\|^{m+1})$, für $h \rightarrow 0$. In den Übungen wird ferner bewiesen, dass

$$\lim_{h \rightarrow 0} \frac{1}{\|h\|^{m+1}} \left[f(a+h) - \sum_{j=0}^{m+1} \sum_{i:|i|=j} \frac{(\partial^i f)(a)}{i!} h^i \right] = 0.$$

Wir haben in (45) bemerkt, dass die r -te Ableitung $D^r f(a)$ als die r -lineare Form

$$\begin{aligned} D^r f(a) : \mathbb{R}^n \times \dots \times \mathbb{R}^n &\rightarrow \mathbb{R} \\ (v_1, v_2, \dots, v_r) &\rightarrow D^r f(a)(v_1, \dots, v_r) = (((D^r f(a)(v_1))(v_2)) \dots)(v_r) \end{aligned}$$

interpretiert werden kann. Man kann dann überprüfen, dass

$$\frac{1}{r!} D^r f(a)(h, h, \dots, h) = \sum_{i:|i|=r} \frac{(\partial^i f)(a)}{i!} h^i = \sum_{i:|i|=r} \frac{1}{i_1! i_2! \dots i_n!} (\partial_1^{i_1} \dots \partial_n^{i_n} f)(a) h_1^{i_1} \dots h_n^{i_n}$$

Damit können wir die Taylor-Entwicklung (49) als

$$f(a+h) = \sum_{r=0}^m \frac{D^r f(a)(h, \dots, h)}{r!} + \frac{D^{m+1} f(a+\theta h)(h, \dots, h)}{(m+1)!} \quad (50)$$

umschreiben. Man bemerke, dass, für feste a , $D^r f(a)(h, \dots, h)$ ein homogenes Polynom von Grad r in h ist. D.h. $g(h) := D^r f(a)(h, \dots, h)$ ist ein Polynom in h , mit der Eigenschaft $g(th) = t^r g(h)$ für alle $t \in \mathbb{R}$.

Lokale Extrema und kritische Punkte. Sei $U \subset \mathbb{R}^n$ offen und $f : U \rightarrow \mathbb{R}$. Ein Punkt $a \in U$ heisst ein lokales Minimum von f , falls eine offene Umgebung $A \subset U$ von a existiert, mit $f(a) = \min\{f(x) : x \in A\}$. a heisst ein lokales Maximum von f , falls eine offene Umgebung $A \subset U$ von a existiert, so dass $f(a) = \max\{f(x) : x \in A\}$. $a \in U$ heisst ein lokales Extremum, falls a entweder ein lokales Minimum oder ein lokales Maximum ist. Für eine Funktion ϕ einer Variablen haben wir in Analysis 1 bewiesen, dass, falls ϕ an der Stelle $t \in \mathbb{R}$ differenzierbar ist, mit $\phi'(t) \neq 0$, t kein Extremum sein kann. Im nächsten Satz zeigen wir die analoge Aussage für Funktionen mehrerer Veränderlichen.

Satz 3.14. *Sei $U \subset \mathbb{R}^n$ offen, $a \in U$ und $f : U \rightarrow \mathbb{R}$ differenzierbar an der Stelle a . Es gelte $\nabla f(a) \neq 0$. Dann ist a kein Extremum von f .*

Beweis: Sei $e \in \mathbb{R}^n$ ein Einheitsvektor mit $e \cdot \nabla f(a) \neq 0$. Wir setzen $\phi(t) = f(a+te)$ für $t \in \mathbb{R}$, mit $|t|$ klein genug (damit $a+te \in U$). ϕ ist an der Stelle $t = 0$ differenzierbar, mit $\phi'(0) = e \cdot \nabla f(a) \neq 0$. Also ist $0 \in \mathbb{R}$ keine Extremalstelle von ϕ . D.h. ϕ nimmt in jeder Umgebung von 0 Werten grösser als $\phi(0) = f(a)$ und Werten kleiner als $\phi(0) = f(a)$ an. Das zeigt, dass a keine Extremalstelle von f ist. \square

Seien U, f wie oben. Wir sagen $a \in U$ ist ein kritischer Punkt von f , falls f in a differenzierbar ist und $\nabla f(a) = 0$. Ist $a \in U$ ein Extremum von f , so muss entweder f an der Stelle a nicht differenzierbar sein, oder a muss ein kritischer Punkt sein.

Sei nun $f \in C^p(U)$ und a eine kritische Stelle von f . Es existiere $1 < r < p$ mit $D^r f(a) \neq 0$. Sei r die kleinste ganze Zahl mit dieser Eigenschaft. Dann gilt, aus (50),

$$f(a+h) = f(a) + \frac{1}{r!} D^r f(a)(h, \dots, h) + O(\|h\|^{r+1}) \quad (51)$$

für $h \rightarrow 0$. Die Frage, ob a ein Maximum, ein Minimum oder keine Extremalstelle ist, wird vom Verhalten von $D^r f(a)(h, \dots, h)$ bestimmt.

Definition 3.15. Sei $p : \mathbb{R}^n \rightarrow \mathbb{R}$ ein homogenes Polynom. Wir sagen

- p ist positiv definit, wenn $p(h) > 0$ für alle $h \neq 0$
- p ist positiv semidefinit, wenn $p(h) \geq 0$ für alle h
- p ist negativ definit, wenn $p(h) < 0$ für alle $h \neq 0$
- p ist negativ semidefinit, wenn $p(h) \leq 0$ für alle h
- p ist indefinit, wenn p weder positiv noch negativ semidefinit ist

Ist p indefinit, so nimmt p Werte mit beiden Vorzeichen an.

Bemerkung. Ist $p : \mathbb{R}^n \rightarrow \mathbb{R}$ ein homogenes Polynom von ungeradem Grad r , so ist entweder $p \equiv 0$ oder p indefinit. In der Tat

$$p(-h) = (-1)^r p(h) = -p(h)$$

D.h. entweder ist $p \equiv 0$ oder p nimmt positive und negative Werte an.

Beispiele: Sei $n = 2$, $r = 2$. Dann ist

$$\begin{aligned} p(h_1, h_2) &= h_1^2 + h_2^2 && \text{positiv definit} \\ p(h_1, h_2) &= (h_1 + h_2)^2 && \text{positiv semidefinit} \\ p(h_1, h_2) &= -h_1^2 - h_2^2 && \text{negativ definit} \\ p(h_1, h_2) &= -(h_1 + h_2)^2 && \text{seminegativ definit} \\ p(h_1, h_2) &= h_1 h_2 && \text{indefinit} \end{aligned}$$

Aus (51) folgt einfach, falls a ein lokales Minimum ist, so muss $D^r f(a)$ positiv semidefinit sein, und falls a ein lokales Maximum ist, so muss $D^r f(a)$ negativ semidefinit sein. Die umgekehrten Aussagen gelten i.A. nur, wenn $D^r f(a)$ positiv bzw. negativ definit sind (statt nur semidefinit). Das ist der Inhalt der nächsten Proposition.

Proposition 3.16. Sei $U \subset \mathbb{R}^n$ offen, $f \in C^p(U)$, $a \in U$ eine kritische Stelle von f , und $r < p$ so, dass (wie in (51))

$$f(a+h) = f(a) + \frac{1}{r!} D^r f(a)(h, \dots, h) + O(\|h\|^{r+1})$$

für $h \rightarrow 0$. Dann gilt

- a) Ist $D^r f(a)(h, \dots, h)$ positiv definit, so ist a ein lokales Minimum.
- b) Ist $D^r f(a)(h, \dots, h)$ negativ definit, so ist a ein lokales Maximum.

c) Ist $D^r f(a)(h, \dots, h)$ indefinit, so ist a kein Extremum.

Bemerkung: Ist $D^r f(a)(h, \dots, h)$ positiv semidefinit (aber nicht positiv definit) oder negativ semidefinit (aber nicht negativ definit), so wird in Proposition 3.16 keine Aussage über die kritische Stelle a gemacht. In diesem Fall ist eine tiefere Untersuchung notwendig.

Beweis: a) Das Polynom $h \rightarrow D^r f(a)(h, \dots, h)$ ist stetig und $D^r f(a)(h, \dots, h) > 0$ für alle $h \in S = \{h \in \mathbb{R}^n : \|h\| = 1\}$ (S^{n-1} ist die Einheitskugel in \mathbb{R}^n). Da $S \subset \mathbb{R}^n$ kompakt ist, folgt aus dem Satz vom Maximum, dass $\alpha = \inf_{h \in S} D^r f(a)(h, \dots, h) > 0$. Für ein beliebiges $h \in \mathbb{R}^n$, $h \neq 0$ schreiben wir $h = \|h\|e$ für ein $e \in S$. Dann gilt

$$D^r f(a)(h, \dots, h) = \|h\|^r D^r f(a)(e, \dots, e) \geq \alpha \|h\|^r.$$

Das gibt

$$\begin{aligned} f(a+h) &= f(a) + \frac{1}{r!} D^r f(a)(h, \dots, h) + O(\|h\|^{r+1}) \geq f(a) + \frac{\alpha}{r!} \|h\|^r + O(\|h\|^{r+1}) \\ &= f(a) + \frac{\alpha}{r!} \|h\|^r (1 + O(\|h\|)) \geq f(a) \end{aligned}$$

für alle $h \in \mathbb{R}^n$ klein genug. Das zeigt, dass a ein lokales Minimum ist. Analog zeigt man die Aussage b). Um c) zu zeigen, finden wir $e_1, e_2 \in S$ mit $D^r f(a)(e_1, \dots, e_1) > 0$ und $D^r f(a)(e_2, \dots, e_2) < 0$. Für $\lambda > 0$ beliebig finden wir

$$f(a + \lambda e_1) = f(a) + \frac{\lambda^r}{r!} D^r f(a)(e_1, \dots, e_1) (1 + O(\lambda)) > f(a)$$

und

$$f(a + \lambda e_2) = f(a) + \frac{\lambda^r}{r!} D^r f(a)(e_2, \dots, e_2) (1 + O(\lambda)) < f(a)$$

für alle $\lambda > 0$ klein genug. Damit ist a kein Extremum. \square

Besonders wichtig ist der Fall, dass an einem kritischen Punkt a einer Funktion $f \in C^3(U)$, die zweite Ableitung nicht verschwindet. In diesem Fall hängt die Frage, ob a ein Minimum, ein Maximum oder keine Extremalstelle ist, mit dem Verhalten der quadratischen Form $D^2 f(a)(h, h)$ zusammen. Wir bemerken, dass, falls $h = (h_1, \dots, h_n)$,

$$D^2 f(a)(h, h) = \sum_{i,j=1}^n \frac{\partial^2 f}{\partial x_i \partial x_j}(a) h_i h_j.$$

Wir definieren die $n \times n$ Matrix

$$H_{ij} = \partial_i \partial_j f(a) = \frac{\partial^2 f}{\partial x_i \partial x_j}(a).$$

Die Matrix H_{ij} heisst die Hesse'sche Matrix von f an der Stelle a , die quadratische Form $D^2 f(a)(h, h) = \sum_{i,j=1}^n H_{ij} h_i h_j$ die Hesse'sche Form. Aus Satz 3.12 folgt, dass die Matrix H_{ij} symmetrisch ist (d.h. $H_{ij} = H_{ji}$). Die Hesse'sche Form heisst nicht entartet, falls $\det(H_{ij}) \neq 0$. Tatsache: Ist die Hesse'sche Form nicht entartet, dann ist sie entweder positiv definit, negativ definit oder indefinit (der Fall, dass $D^2 f(a)$ positiv

semidefinit, aber nicht positiv definit, und der Fall, dass $D^2f(a)$ negativ semidefinit, aber nicht negativ definit ist, sind ausgeschlossen). Diese Aussage folgt aus der Bemerkung, dass eine symmetrische $n \times n$ Matrix immer durch eine unitäre Matrix diagonalisierbar ist. Seien $\lambda_1, \dots, \lambda_n$ die (nicht notwendigerweise verschiedenen) Eigenwerte von H_{ij} . Ist $D^2f(a)$ nicht entartet, so gilt $\lambda_i \neq 0$ für alle $i = 1, \dots, n$. Es gibt also nur drei Möglichkeiten: 1) alle Eigenwerte sind positiv, 2) alle Eigenwerte sind negativ, 3) es gibt positive und negative Eigenwerte. Im Fall 1) ist $D^2f(a)$ positiv definit, im Fall 2) ist $D^2f(a)$ negativ definit und im Fall 3) ist $D^2f(a)$ indefinit. Es folgt aus dieser Bemerkung, dass, falls die Hesse'sche Form nicht entartet ist, man immer durch Untersuchung von $D^2f(a)$ entscheiden kann, ob der kritische Punkt a ein Maximum, ein Minimum oder kein Extremum ist.

3.4 Umkehrabbildung und Satz über implizite Funktionen

Für differenzierbare Funktionen einer Variablen $f : \mathbb{R} \supset \Omega \rightarrow \mathbb{R}$ haben wir in Analysis 1 gezeigt, dass, unter der Bedingung $f'(a) \neq 0$, die Umkehrabbildung lokal wohldefiniert und an der Stelle $f(a)$ differenzierbar, mit $(f^{-1})'(f(a)) = 1/f'(a)$, ist. Wir möchten nun eine analoge Aussage für Funktionen mehrerer Variablen beweisen. Die Bedingung $f'(a) \neq 0$ wird hier durch die Bedingung ersetzt, dass $Df(a)$ invertierbar ist.

Satz 3.17 (Satz über die Umkehrabbildung). *Sei $U \subset \mathbb{R}^n$ offen, $f : U \rightarrow \mathbb{R}^n$ stetig differenzierbar, $a \in U$ und $Df(a) \in \mathcal{L}(\mathbb{R}^n; \mathbb{R}^n)$ invertierbar. Dann existieren offene Umgebungen V von a und W von $f(a)$ so, dass $f : V \rightarrow W$ bijektiv und $f^{-1} : W \rightarrow V$ stetig differenzierbar. Ferner gilt*

$$Df^{-1}(f(a)) = (Df(a))^{-1}.$$

Bemerkung: Seien $f_1, \dots, f_n : U \rightarrow \mathbb{R}$ die Komponenten von f , d.h. es gelte $f(x) = (f_1(x), \dots, f_n(x))$ für alle $x \in U$. Die Invertierbarkeit von $Df(a)$ ist dann äquivalent zur Bedingung, dass

$$\det \left(\frac{\partial f_i(a)}{\partial x_j} \right)_{i,j \leq n} \neq 0.$$

Im Beweis des Satzes (genauer gesagt, um die Stetigkeit der Ableitung von f^{-1} zu zeigen) werden wir die folgende Proposition benutzen.

Proposition 3.18. *Sei*

$$GL(\mathbb{R}^n) = \{L \in \mathcal{L}(\mathbb{R}^n; \mathbb{R}^n) : L \text{ ist invertierbar}\}$$

Die Menge $GL(\mathbb{R}^n)$ ist in $\mathcal{L}(\mathbb{R}^n; \mathbb{R}^n)$ offen (bezüglich der aus der Operatornorm (oder aus jeder anderen Norm) induzierten Topologie). Die Abbildung $i : GL(\mathbb{R}^n) \rightarrow GL(\mathbb{R}^n)$, definiert durch $i(L) = L^{-1}$ ist stetig.

Beweis: Sei $L \in \mathcal{L}(\mathbb{R}^n; \mathbb{R}^n)$ invertierbar. Dann gilt

$$\|L^{-1}\|_{\text{op}} = \sup_{y \neq 0} \frac{\|L^{-1}y\|}{\|y\|} = \sup_{x \neq 0} \frac{\|x\|}{\|Lx\|} = \frac{1}{\inf_{x \neq 0} \frac{\|Lx\|}{\|x\|}}$$

und deswegen

$$\inf_{x \neq 0} \frac{\|Lx\|}{\|x\|} = \frac{1}{\|L^{-1}\|_{\text{op}}}$$

Das impliziert, dass

$$\|Lx\| \geq \frac{1}{\|L^{-1}\|_{\text{op}}} \|x\| \quad (52)$$

für alle $x \in \mathbb{R}^n$.

Sei nun $L_0 \in GL(\mathbb{R}^n)$ invertierbar und $L \in \mathcal{L}(\mathbb{R}^n; \mathbb{R}^n)$ mit $\|L - L_0\|_{\text{op}} < \|L_0^{-1}\|^{-1}$. Dann gilt, für ein beliebiges $x \in \mathbb{R}^n$,

$$\|Lx\| \geq \|L_0x\| - \|(L - L_0)x\| \geq (\|L_0^{-1}\|_{\text{op}}^{-1} - \|L - L_0\|_{\text{op}}) \|x\| \geq c\|x\| \quad (53)$$

für ein $c > 0$. Das zeigt, dass L injektiv und deswegen auch invertierbar ist (eine lineare Abbildung $L \in \mathcal{L}(\mathbb{R}^n; \mathbb{R}^n)$ ist genau dann bijektiv, wenn sie injektiv ist). Das impliziert, dass

$$\{L \in \mathcal{L}(\mathbb{R}^n; \mathbb{R}^n) : \|L - L_0\|_{\text{op}} < \|L_0^{-1}\|_{\text{op}}^{-1}\} \subset GL(\mathbb{R}^n)$$

und also, dass $GL(\mathbb{R}^n)$ offen ist.

Aus (53) finden wir auch

$$\|L^{-1}\|_{\text{op}} = \frac{1}{\inf_{x \neq 0} \frac{\|Lx\|}{\|x\|}} \leq \frac{1}{\frac{1}{\|L_0^{-1}\|_{\text{op}}} - \|L - L_0\|_{\text{op}}} = \frac{\|L_0^{-1}\|_{\text{op}}}{1 - \|L_0^{-1}\|_{\text{op}} \|L - L_0\|_{\text{op}}}$$

Da $L^{-1} - L_0^{-1} = L^{-1}(L_0 - L)L_0^{-1}$ erhalten wir

$$\|L^{-1} - L_0^{-1}\|_{\text{op}} \leq \|L^{-1}\|_{\text{op}} \|L_0^{-1}\|_{\text{op}} \|L - L_0\|_{\text{op}} \leq \frac{\|L_0^{-1}\|_{\text{op}}^2}{1 - \|L_0^{-1}\|_{\text{op}} \|L - L_0\|_{\text{op}}} \|L - L_0\|_{\text{op}}$$

Das zeigt, dass die Abbildung $i(L) = L^{-1}$ stetig an der Stelle L_0 ist, für jede $L_0 \in GL(\mathbb{R}^n)$. \square

Beweis von Satz 3.17. O.B.d.A können wir annehmen, dass $a = 0$ und $f(0) = 0$ (sonst ersetzen wir f durch die Funktion $\tilde{f}(x) = f(x + a) - f(a)$). Wir definieren $\phi : U \rightarrow \mathbb{R}^n$ durch $\phi(x) = f(x) - Df(0)(x)$. Dann gilt $f(x) = Df(0)(x) + \phi(x)$ und $Df(x) = Df(0) + D\phi(x)$ für alle $x \in U$. Insbesondere gilt $\phi(0) = f(0) = 0$ und $D\phi(0) = 0$. Da f stetig differenzierbar ist, ist die Abbildung $x \rightarrow D\phi(x)$ stetig. Also existiert $r_0 > 0$ mit

$$\|D\phi(x)\|_{\text{op}} \leq \frac{1}{2\|Df(0)^{-1}\|_{\text{op}}}$$

für alle $x \in \overline{B}_{r_0} := \{x \in \mathbb{R}^n : \|x\| \leq r_0\}$. Aus der Mittelwertsatzung gilt dann

$$\|\phi(x_1) - \phi(x_2)\| \leq \frac{1}{2\|Df(0)^{-1}\|_{\text{op}}} \|x_1 - x_2\| \quad (54)$$

für alle $x_1, x_2 \in \overline{B}_{r_0}$. Insbesondere, mit $x_2 = 0$, finden wir $\|\phi(x)\| \leq (2\|Df(0)^{-1}\|_{\text{op}})^{-1} \|x\|$ für alle $x \in \overline{B}_{r_0}$ (weil $\phi(0) = 0$).

Schritt 1. Für alle $r \leq r_0$ gilt $f(\overline{B}_r) \supset \overline{B}_{r/2\|Df(0)^{-1}\|_{\text{op}}}$. D.h. für jede $y \in \mathbb{R}^n$ mit $\|y\| \leq r/(2\|Df(0)^{-1}\|_{\text{op}})$ existiert ein $x \in \overline{B}_r$ mit $f(x) = y$.

Beweis. Für beliebige $y \in \overline{B}_{r/2\|Df(0)^{-1}\|_{\text{op}}}$ definieren wir $\psi_y : \overline{B}_r \rightarrow \mathbb{R}^n$ durch $\psi_y(x) = Df(0)^{-1}(y - \phi(x))$. Es gilt $f(x) = y$ genau dann, wenn $\psi_y(x) = x$, d.h. wenn x ein Fixpunkt von ψ_y ist. Wir möchten den Banachschen Fixpunktsatz anwenden, um zu zeigen, dass ψ_y einen Fixpunkt in \overline{B}_r besitzt. Dazu bemerken wir zunächst, dass

$$\|\psi_y(x)\| = \|Df(0)^{-1}\|_{\text{op}} \|y - \phi(x)\| \leq \|Df(0)^{-1}\|_{\text{op}} (\|y\| + \|\phi(x)\|) \leq r$$

für alle $y \in \overline{B}_{r/2\|Df(0)^{-1}\|_{\text{op}}}$ und $x \in \overline{B}_r$ (wir haben hier (54) gebraucht). Das zeigt, dass für alle $y \in \overline{B}_{r/2\|Df(0)^{-1}\|_{\text{op}}}$, $\psi_y : \overline{B}_r \rightarrow \overline{B}_r$. Ferner haben wir

$$\psi_y(x_1) - \psi_y(x_2) = Df(0)^{-1}(\phi(x_1) - \phi(x_2))$$

und deswegen

$$\|\psi_y(x_1) - \psi_y(x_2)\| = \|Df(0)^{-1}\|_{\text{op}} \|\phi(x_1) - \phi(x_2)\| \leq \frac{1}{2} \|x_1 - x_2\|$$

für alle $x \in \overline{B}_r$ (wieder wegen (54)). Es folgt, dass für alle $y \in \overline{B}_{r/2\|Df(0)^{-1}\|_{\text{op}}}$, ψ_y eine Kontraktion auf \overline{B}_r ist. Da \overline{B}_r ein vollständiger metrischer Raum ist, folgt aus Satz 2.3, dass für alle $y \in \overline{B}_{r/2\|Df(0)^{-1}\|_{\text{op}}}$ ein $x \in \overline{B}_r$ mit $f(x) = y$ existiert. \square

Schritt 2. Es gilt

$$\frac{1}{2\|Df(0)^{-1}\|_{\text{op}}} \|x_1 - x_2\| \leq \|f(x_1) - f(x_2)\| \leq \frac{3\|Df(0)\|_{\text{op}}}{2} \|x_1 - x_2\|$$

für alle $x_1, x_2 \in \overline{B}_{r_0}$. Insbesondere ist f auf \overline{B}_{r_0} injektiv.

Beweis: Wir haben $f(x) = \phi(x) + Df(0)(x)$. Aus (54) gilt

$$\begin{aligned} \|f(x_1) - f(x_2)\| &\leq \|Df(0)\| \|x_1 - x_2\| + \|\phi(x_1) - \phi(x_2)\| \\ &\leq \left(\|Df(0)\|_{\text{op}} + \frac{1}{2\|Df(0)^{-1}\|_{\text{op}}} \right) \|x_1 - x_2\| \end{aligned}$$

Aus $1 = Df(0)Df(0)^{-1}$ folgt, dass $1 \leq \|Df(0)\|_{\text{op}} \|Df(0)^{-1}\|_{\text{op}}$ und damit

$$\|Df(0)^{-1}\|_{\text{op}}^{-1} \leq \|Df(0)\|_{\text{op}}.$$

Das ergibt

$$\|f(x_1) - f(x_2)\| \leq \frac{3\|Df(0)\|_{\text{op}}}{2} \|x_1 - x_2\|$$

für alle $x_1, x_2 \in \overline{B}_{r_0}$. Andererseits gilt aus (52) und wieder aus (54),

$$\begin{aligned} \|f(x_1) - f(x_2)\| &\geq \|Df(0)(x_1 - x_2)\| - \|\phi(x_1) - \phi(x_2)\| \\ &\geq \left(\frac{1}{\|Df(0)^{-1}\|_{\text{op}}} - \frac{1}{2\|Df(0)^{-1}\|_{\text{op}}} \right) \|x_1 - x_2\| \\ &= \frac{1}{2\|Df(0)^{-1}\|_{\text{op}}} \|x_1 - x_2\|. \end{aligned}$$

□

Sei nun $r < r_0$ fest. Wir setzen $W_0 = B_{r/(2\|Df(0)^{-1}\|_{\text{op}})} = \{x \in \mathbb{R}^n : \|x\| < r/2\|Df(0)^{-1}\|_{\text{op}}\}$, und $V_0 = f^{-1}(W_0) \cap B_{r_0} = \{x \in B_{r_0} : \|f(x)\| < r\}$. W_0 ist offen. Da f stetig, ist auch V_0 offen. Es folgt aus Schritten 1 und 2, dass $f : V_0 \rightarrow W_0$ bijektiv ist. Die Injektivität folgt aus Schritt 2, weil $V_0 \subset B_{r_0}$. Die Surjektivität folgt dagegen aus Schritt 1, weil für jede $y \in B_{r/(2\|Df(0)^{-1}\|_{\text{op}})}$, $x \in \overline{B}_r \subset B_{r_0}$ mit $f(x) = y$ existiert. Wir bezeichnen die Inverse mit $g : W_0 \rightarrow V_0$. Für $y_1, y_2 \in W_0$ setze $x_1 = g(y_1)$ und $x_2 = g(y_2)$. Aus Schritt 2 haben wir

$$\|g(y_1) - g(y_2)\| = \|x_1 - x_2\| \leq 2\|Df(0)^{-1}\|_{\text{op}}\|f(x_1) - f(x_2)\| = 2\|Df(0)^{-1}\|_{\text{op}}\|y_1 - y_2\|$$

Also ist g auf W_0 stetig (sogar Lipschitz-stetig).

Schritt 3. g ist an der Stelle 0 differenzierbar. Es gilt $Dg(0) = Df(0)^{-1}$.

Beweis. Da $g(0) = 0$ müssen wir zeigen, dass

$$g(y) - Df(0)^{-1}(y) = o(\|y\|)$$

für $y \rightarrow 0$. Sei $0 < \varepsilon < 1$ festgewählt. Da $x \rightarrow \phi(x)$ stetig, existiert $\delta > 0$ so, dass $\|D\phi(x)\| \leq \varepsilon/(2\|Df(0)^{-1}\|_{\text{op}})$ für alle $\|x\| \leq \delta$. Nach Definition von $r_0 > 0$ gilt $\delta \leq r_0$. Für $y \in \mathbb{R}^n$ mit $\|y\| < \delta/(2\|Df(0)^{-1}\|_{\text{op}})$, sei $x = g(y)$. Dann gilt $\|x\| \leq \delta$. Ferner

$$f(x) = y = Df(0)(x) + \phi(x)$$

Wir multiplizieren rechts und links mit der Matrix $Df(0)^{-1}$ und erhalten

$$g(y) - Df(0)^{-1}(y) = -Df(0)^{-1}(\phi(x))$$

Aus der Mittelwertabschätzung für ϕ ,

$$\begin{aligned} \|g(y) - Df(0)^{-1}(y)\| &= \|Df(0)^{-1}\phi(x)\| = \|Df(0)^{-1}(\phi(x) - \phi(0))\| \\ &\leq \|Df(0)^{-1}\|_{\text{op}}\|\phi(x) - \phi(0)\| \leq \frac{\varepsilon}{2}\|x\| \\ &\leq \frac{\varepsilon}{2}\|g(y)\| \leq \varepsilon\|Df(0)^{-1}\|_{\text{op}}\|y\| \end{aligned}$$

Da $\varepsilon > 0$ beliebig ist, folgt die Behauptung. □

Wir haben somit folgendes bewiesen: Für jede $a \in U$ mit $Df(a)$ invertierbar, existieren offene Umgebungen V_0 von a und W_0 von $f(a)$, so dass $f : V_0 \rightarrow W_0$ bijektiv ist, und so, dass $f^{-1} : W_0 \rightarrow V_0$ stetig und an der Stelle a differenzierbar ist, mit $Df^{-1}(f(a)) = (Df(a))^{-1}$. Da $x \rightarrow Df(x)$ stetig, und da die Menge der invertierbaren linearen Abbildungen $GL(\mathbb{R}^n)$ in $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^n)$ offen ist, finden wir eine offene Umgebung $V \subset V_0$ von a so, dass $Df(x)$ invertierbar ist, für alle $x \in V$. Wir setzen $W = f(V)$; da f^{-1} stetig ist, ist auch W offen. f^{-1} ist dann in jedem Punkt von W differenzierbar und $Df^{-1}(f(x)) = (Df(x))^{-1} = Df(x)^{-1}$. Da die Abbildung $i : GL(\mathbb{R}^n) \rightarrow GL(\mathbb{R}^n)$ stetig ist, ist $Df(x)^{-1} = i(Df(x))$ als Komposition zweier stetiger Abbildungen wieder stetig. Damit ist $f^{-1} : W \rightarrow V$ stetig differenzierbar. □

Definition 3.19. Seien X, Y zwei metrische Räume. Ein Homöomorphismus von X nach Y ist eine Bijektion $f : X \rightarrow Y$, so dass f und f^{-1} stetig sind. Ist $f : X \rightarrow Y$ ein Homöomorphismus, so ist auch f^{-1} ein Homöomorphismus. Eine stetige Bijektion $f : X \rightarrow Y$ ist genau dann ein Homöomorphismus, wenn $f(U)$ offen in Y für jede U offen in X ist (weil eine Abbildung genau dann stetig ist, wenn das Urbild jeder offenen Menge wieder offen ist; siehe Analysis 1, Prop. 6.18). Zwei metrische Räume X, Y heißen homöomorph, wenn ein Homöomorphismus $f : X \rightarrow Y$ existiert.

Seien nun $U \subset \mathbb{R}^n$ und $V \subset \mathbb{R}^m$ offen. Ein Diffeomorphismus $f : U \rightarrow V$ heißt ein Diffeomorphismus, falls f und f^{-1} stetig differenzierbar sind. Zwei offene Mengen $U \subset \mathbb{R}^n$ und $V \subset \mathbb{R}^m$ heißen diffeomorph, wenn ein Diffeomorphismus $f : U \rightarrow V$ existiert. Eine Bijektion $f : U \rightarrow V$ ist genau dann ein Diffeomorphismus, wenn f^{-1} ein Diffeomorphismus ist.

Bemerkung: Ist $U \subset \mathbb{R}^n$, $V \subset \mathbb{R}^m$ und $f : U \rightarrow V$ ein Diffeomorphismus, dann gilt $f^{-1} \circ f(x) = x$ für alle $x \in U$. Die Kettenregel impliziert, dass

$$Df^{-1}(f(x)) \circ Df(x) = 1,$$

wobei $Df^{-1}(f(x)) \in \mathcal{L}(\mathbb{R}^m; \mathbb{R}^n)$ und $Df(x) \in \mathcal{L}(\mathbb{R}^n; \mathbb{R}^m)$. Das ist nur möglich, falls $m \geq n$. Analog impliziert $f \circ f^{-1}(x) = x$, dass $Df(f^{-1}(x)) \circ Df^{-1}(x) = 1$ und also, dass $n \geq m$. Es folgt, dass $n = m$. Mit anderen Worten können nur Mengen der gleichen Dimension zueinander diffeomorph sein.

Tatsache: Sind $U \subset \mathbb{R}^n$ und $V \subset \mathbb{R}^m$ nicht leer und offen, und $f : U \rightarrow V$ ein Homöomorphismus, dann muss $n = m$ sein.

Bemerkung: Der Satz der Umkehrabbildung besagt, dass falls $U \subset \mathbb{R}^n$ offen ist, $a \in U$, $f : U \rightarrow \mathbb{R}^n$ stetig differenzierbar, mit $Df(a)$ invertierbar, dann ist f lokal in der Nähe von a ein Diffeomorphismus. D.h. es existieren offene Umgebungen V von a und W von $f(a)$, so dass $f : V \rightarrow W$ ein Diffeomorphismus ist.

Eine wichtige Anwendung des Satzes über die Umkehrabbildung ist der Satz über implizite Funktionen. Oft werden Teilmengen von \mathbb{R}^n durch Gleichungen definiert. Z.B. ist $\{(x, y) \in \mathbb{R}^2 : y = x^3\}$ eine Teilmenge von \mathbb{R}^2 . In diesem Fall ist die Teilmenge besonders einfach, weil sie als Graph einer Funktion geschrieben werden kann. Das ist i.A. nicht möglich. Z.B. für den Einheitskreis $S = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 = 1\}$ existiert keine Funktion g , definiert auf einer Teilmenge $U \subset \mathbb{R}$, mit der Eigenschaft, dass $S = \{(x, g(x)) : x \in U\}$. Trotzdem ist es in diesem Fall möglich, S lokal als Graph zu schreiben. Betrachten wir z.B. den Punkt $(0, 1)$ auf S . Es ist dann einfach zu sehen, dass offene Umgebungen $U \subset \mathbb{R}$ von 0 und $V \subset \mathbb{R}$ von 1 und eine differenzierbare Funktion $g : U \rightarrow V$ existiert, mit der Eigenschaft, dass

$$S \cap (U \times V) = \{(x, g(x)) : x \in U\}.$$

In diesem Fall ist es sogar möglich, $g(x) = \sqrt{1 - x^2}$ explizit zu schreiben.

Ein anderes Beispiel ist das sogenannte kartesische Blatt

$$T = \{(x, y) \in \mathbb{R}^2 : x^3 - 2xy + y^3 = 0\}.$$

Wie S , kann auch T nicht global als Graph einer Funktion geschrieben werden. Ist es möglich, T lokal als Graph zu schreiben? Der Punkt $(1, 1)$ ist z.B. in T . Es ist einfach zu sehen (vgl. Bild von T), dass offene Umgebungen $U, V \subset \mathbb{R}$ von 1 und eine differenzierbare Funktion $f : U \rightarrow V$ existieren, so dass $f(1) = 1$ und

$$T \cap (U \times V) = \{(x, f(x)) : x \in U\}.$$

Kann T in der Nähe von jedem seiner Punkte lokal als Graph einer Funktion geschrieben werden? Nein: Es ist einfach zu sehen, dass in der Nähe von $(0, 0)$, T nicht als Graph geschrieben werden kann. Was unterscheidet also die Punkte $(1, 1)$ und $(0, 0)$ auf T ? Sei $f(x, y) = x^3 - 2xy + y^3$, so dass T Menge aller Nullstellen von f ist. Dann gilt

$$\frac{\partial f}{\partial y}(x, y) = -2x + 3y^2$$

Wir zeigen im nächsten Satz, dass der fundamentale Unterschied zwischen $(1, 1)$ und $(0, 0)$ die Tatsache ist, dass $(\partial f / \partial y)(1, 1) = -2 \neq 0$ während $(\partial f / \partial y)(0, 0) = 0$.

Satz 3.20 (Satz über implizite Funktionen). *Seien $m, n \in \mathbb{N} \setminus \{0\}$, $U \subset \mathbb{R}^m \times \mathbb{R}^n$ offen und nicht leer. Sei $f \in C^1(U; \mathbb{R}^n)$ und $(x_0, y_0) \in U$ mit $f(x_0, y_0) = 0$. Es gelte*

$$\det \left(\frac{\partial f_i}{\partial y_j}(x_0, y_0) \right)_{1 \leq i, j \leq n} \neq 0 \quad (55)$$

Dann existieren offene Umgebungen $V \subset \mathbb{R}^m$ von x_0 und $W \subset \mathbb{R}^n$ von y_0 und eine stetig differenzierbare Funktion $g : V \rightarrow W$, so dass

$$\{(x, y) \in V \times W : f(x, y) = 0\} = \{(x, g(x)) : x \in V\}$$

Ferner gilt

$$Dg(x_0) = -(D_y f(x_0, y_0))^{-1} \cdot D_x f(x_0, y_0). \quad (56)$$

Bemerkung: $D_y f(x_0, y_0)$ und $D_x f(x_0, y_0)$ bezeichnen die Ableitung von f als Funktion von y bei festen $x = x_0$, bzw. die Ableitung von f als Funktion von x , bei festen $y = y_0$. $D_y f(x_0, y_0)$ ist eine $n \times n$ Matrix und $D_x f(x_0, y_0)$ eine $n \times m$ Matrix. Die Bedingung (55) bedeutet genau, dass die Matrix $D_y f(x_0, y_0)$ invertierbar ist. In diesem Fall ist die Inverse $D_y f(x_0, y_0)^{-1}$ wieder eine $n \times n$ Matrix und $D_y f(x_0, y_0)^{-1} \cdot D_x f(x_0, y_0)$ eine $n \times m$ Matrix. Damit ist (56) konsistent mit der Tatsache, dass g eine Teilmenge von \mathbb{R}^m auf einer Teilmenge von \mathbb{R}^n abbildet.

Bemerkung: Seien $f_1, \dots, f_n : U \rightarrow \mathbb{R}$ die Komponenten von f , s.d. $f(x, y) = (f_1(x, y), \dots, f_n(x, y))$ für alle $(x, y) \in U$ gelte. Dann ist die vektorielle Gleichung $f(x, y) = 0$ das System von n Gleichungen

$$\begin{cases} f_1(x_1, \dots, x_m, y_1, \dots, y_n) = 0 \\ f_2(x_1, \dots, x_m, y_1, \dots, y_n) = 0 \\ \dots \\ f_n(x_1, \dots, x_m, y_1, \dots, y_n) = 0 \end{cases}$$

Der Satz besagt, dass, falls $D_y f(x_0, y_0)$ invertierbar ist, kann man das Gleichungssystem lokal für (y_1, \dots, y_n) lösen. Sind die Funktionen f_1, \dots, f_n linear oder affin, dann ist die Behauptung schon aus der linearen Algebra bekannt (in diesem Fall ist die Ableitung $D_y f(x_0, y_0)$ unabhängig von (x_0, y_0) und die Behauptung gilt natürlich global).

Beweis: Wir definieren die Hilfsfunktion $F : U \rightarrow \mathbb{R}^m \times \mathbb{R}^n$ durch $F(x, y) = (x, f(x, y))$. Da $f \in C^1(U; \mathbb{R}^n)$, ist F stetig differenzierbar. Die Ableitung von F an der Stelle (x_0, y_0) ist aus der Blockmatrix

$$DF(x_0, y_0) = \begin{pmatrix} 1_{\mathbb{R}^m} & 0 \\ D_x f(x_0, y_0) & D_y f(x_0, y_0) \end{pmatrix}$$

mit der $n \times m$ Matrix $D_x f(x_0, y_0) = ((\partial f_i / \partial x_j)(x_0, y_0))$ und mit der $n \times n$ Matrix $D_y f(x_0, y_0) = ((\partial f_i / \partial y_j)(x_0, y_0))$ gegeben ($1_{\mathbb{R}^m}$ ist die Identität auf \mathbb{R}^m). Nach Annahme ist $D_y f(x_0, y_0)$ invertierbar; sei $D_y f(x_0, y_0)^{-1}$ die Inverse. Wir behaupten, dass auch $DF(x_0, y_0)$ invertierbar ist. In der Tat, explizite Berechnung zeigt, dass

$$\begin{aligned} & \begin{pmatrix} 1_{\mathbb{R}^m} & 0 \\ D_x f(x_0, y_0) & D_y f(x_0, y_0) \end{pmatrix} \cdot \begin{pmatrix} 1_{\mathbb{R}^m} & 0 \\ -D_y f(x_0, y_0)^{-1} D_x f(x_0, y_0) & D_y f(x_0, y_0)^{-1} \end{pmatrix} \\ &= \begin{pmatrix} 1_{\mathbb{R}^m} & 0 \\ 0 & 1_{\mathbb{R}^n} \end{pmatrix}. \end{aligned}$$

Deswegen ist $DF(x_0, y_0)$ invertierbar. Aus dem Satz über die Umkehrabbildung (Satz 3.17) folgt, dass offene Umgebungen $U_1 \subset \mathbb{R}^m \times \mathbb{R}^n$ von (x_0, y_0) und $U_2 \subset \mathbb{R}^m \times \mathbb{R}^n$ von $F(x_0, y_0) = (x_0, 0)$ existieren, so dass $F : U_1 \rightarrow U_2$ ein Diffeomorphismus ist. Da U_1 eine offene Umgebung von (x_0, y_0) ist, kann man offene Umgebungen $\tilde{V} \subset \mathbb{R}^m$ von x_0 und $W \subset \mathbb{R}^n$ von y_0 , mit $\tilde{V} \times W \subset U_1$. Dann ist $F(\tilde{V} \times W) \subset U_2$ eine offene Umgebung von $(x_0, 0)$ (weil F ein Homöomorphismus ist), und $F : \tilde{V} \times W \rightarrow F(\tilde{V} \times W)$ wieder ein Diffeomorphismus. Sei $G : F(\tilde{V} \times W) \rightarrow \tilde{V} \times W$ die Inverse dieses Diffeomorphismus'. Da $F(x, y) = (x, f(x, y))$, existiert eine Funktion $\tilde{G} : F(\tilde{V} \times W) \rightarrow W$ mit $G(x, y) = (x, \tilde{G}(x, y))$ für alle $(x, y) \in F(\tilde{V} \times W)$. Da G differenzierbar ist, ist auch \tilde{G} differenzierbar. Da $F(\tilde{V} \times W)$ eine offene Umgebung von $(x_0, 0)$ ist, finden wir eine offene Umgebung $V \subset \tilde{V}$ von x_0 mit $\{(x, 0) : x \in V\} \subset F(\tilde{V} \times W)$. Dann können wir $\phi : V \rightarrow W$ durch $\phi(x) = \tilde{G}(x, 0)$ definieren (d.h. durch $G(x, 0) = (x, \phi(x))$ für alle $x \in V$). Da \tilde{G} differenzierbar ist, ist auch ϕ differenzierbar. Für $(x, y) \in V \times W \subset \tilde{V} \times W$ gilt dann

$$\begin{aligned} f(x, y) = 0 &\iff F(x, y) = (x, 0) \iff G(x, 0) = (x, y) \\ &\iff y = \tilde{G}(x, 0) \iff y = \phi(x) \end{aligned}$$

D.h.

$$\{(x, y) \in V \times W : f(x, y) = 0\} = \{(x, \phi(x)) : x \in V\}$$

Aus $f(x, \phi(x)) = 0$ für alle $x \in V$ folgt, mit der Kettenregel, dass

$$0 = Df(x_0, \phi(x_0)) \cdot Dg(x_0) = D_x f(x_0, \phi(x_0)) \cdot 1_{\mathbb{R}^m} + D_y f(x_0, \phi(x_0)) \cdot D\phi(x_0)$$

Das gibt (56). □

Beispiel. Sei $f(x, y) = x^3 - 2xy + y^3$ und, wie oben, $T = \{(x, y) \in \mathbb{R}^2 : f(x, y) = 0\}$ das kartesische Blatt. Es gilt

$$\frac{\partial f}{\partial y}(x, y) = -2x + 3y^2, \quad \text{und} \quad \frac{\partial f}{\partial x}(x, y) = 3x^2 - 2y$$

Ist $(x_0, y_0) \in T$ mit $-2x_0 + 3y_0^2 \neq 0$, dann kann man, in der Nähe von (x_0, y_0) , T als Graph einer Funktion von x schreiben. Wir bemerken, $(x_0, y_0) \in T$ mit $-2x_0 + 3y_0^2 = 0$ impliziert, dass

$$\frac{27}{8}y_0^6 - 2y_0^3 = 0 \Rightarrow 2y_0^3 \left(\frac{27}{16}y_0^3 - 1 \right) = 0$$

Das ist nur bei $y_0 = 0$ oder $y_0 = 2\sqrt{2}/3$ möglich. Also existieren für jede $(x_0, y_0) \in T$, mit der Ausnahmen $(x_0, y_0) = (0, 0)$ und $(x_0, y_0) = (4/3, 2\sqrt{2}/3)$, offene Umgebungen $V \subset \mathbb{R}$ von x_0 und $W \subset \mathbb{R}$ von y_0 und eine C^1 -Funktion $\phi : V \rightarrow W$, so dass

$$T \cap (V \times W) = \{(x, y) \in V \times W : f(x, y) = 0\} = \{(x, \phi(x)) : x \in V\}$$

und

$$\phi'(x_0) = -\frac{(\partial f / \partial x)(x_0, y_0)}{(\partial f / \partial y)(x_0, y_0)} = -\frac{3x_0^2 - 2y_0}{3y_0^2 - 2x_0}.$$

Analog finden wir, dass für alle $(x_0, y_0) \in T$ mit den Ausnahmen von $(x_0, y_0) = (0, 0)$ und $(x_0, y_0) = (2\sqrt{2}/3, 4/3)$ offene Umgebungen V von y_0 und W von x_0 und eine C^1 -Funktion $\psi : V \rightarrow W$ existieren, so dass

$$T \cap (W \times V) = \{(x, y) \in W \times V : f(x, y) = 0\} = \{(\psi(y), y) : y \in V\}$$

und

$$\psi'(y_0) = -\frac{3x_0^2 - 2x_0}{3y_0^2 - 2y_0}.$$

3.5 Mannigfaltigkeiten in \mathbb{R}^n

Wir untersuchen in diesem Abschnitt besondere Teilmengen von \mathbb{R}^n , genannt Mannigfaltigkeiten oder Untermannigfaltigkeiten des \mathbb{R}^n , die lokal wie \mathbb{R}^k aussehen, für ein $k \leq n$.

Bevor wir zur genaueren Definition von Mannigfaltigkeit kommen, betrachten wir einige Beispiele von Teilmengen von \mathbb{R}^n , die lokal wie \mathbb{R}^k für $k = 1$ oder $k = 2$ aussehen. Im Fall $k = 1$ spricht man von Kurven. Das Begriff von Kurven kann verschiedene Bedeutungen haben. Eine *parametrisierte Kurve* ist eine Abbildung $\phi : I \rightarrow \mathbb{R}^n$, für ein Intervall $I \subset \mathbb{R}$. Eine parametrisierte Kurve kann zum Beispiel die Bewegung eines Teilchens im Raum beschreiben, als Funktion der Zeit $t \in I$. Eine parametrisierte Kurve ist also nicht nur durch die Bahn des Teilchens charakterisiert, sondern auch von dem Zeitplan (die parametrisierte Kurve bestimmt die Position des Teilchens zu jeder Zeit, nicht nur seine Trajektorie). Sind wir nur an der Bahn interessiert, und nicht am Zeitplan, so können wir die folgende Äquivalenzrelation im Raum der parametrisierten Kurven definieren. Wir sagen zwei parametrisierte Kurve $\varphi_1 : I_1 \rightarrow \mathbb{R}^n$, $\varphi_2 : I_2 \rightarrow \mathbb{R}^n$, für zwei Intervalle $I_1, I_2 \subset \mathbb{R}$ sind äquivalent, falls eine monoton wachsende stetige und surjektive Funktion $\psi : I_1 \rightarrow I_2$ existiert, so dass $\phi_1 = \phi_2 \circ \psi$. In diesem Fall heisst ψ eine Parametertransformation. Man kann sich leicht davon überzeugen, dass das wirklich eine Äquivalenzrelation definiert. Man kann dann eine Kurve als eine Äquivalenzklasse von parametrisierten Kurven definieren. Mit anderen Worten, eine Kurve wird somit als die Bildmenge einer parametrisierten Kurve definiert.

Man könnte auch Kurven als Graph von Funktionen definieren. Das ist aber zu restriktiv; z.B. der Kreis $S^1 = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 = 1\}$ kann nicht als Graph $\{(x, y) \in \mathbb{R}^2 : y = f(x)\}$ einer Funktion von x geschrieben werden. Es ist auch unmöglich, S^1 als Graph $\{(x, y) \in \mathbb{R}^2 : x = f(y)\}$ einer Funktion von y zu schreiben. Immerhin, der Kreis S^1 ist die Vereinigung der Graphen von zwei Funktionen, nämlich $f_1(x) = (1 - x^2)^{1/2}$ und $f_2(x) = -(1 - x^2)^{1/2}$. Wenn wir nur Funktionen auf offene Teilmengen von \mathbb{R} betrachten möchten, so können wir S^1 als die Vereinigung der Graphen von vier Funktionen schreiben, nämlich $y = \pm\sqrt{1 - x^2}$ auf $x \in (-1; 1)$ und $x = \pm\sqrt{1 - y^2}$ auf $y \in (-1; 1)$. Zwei dieser Abbildungen definieren y als Funktion von x , die anderen zwei geben x als Funktion von y . Ausgehend aus diesem Beispiel kann man also berlegen, Kurven als Vereinigungen von Graphen zu definieren. Gemäss dieser Definition kann man sich auch vorstellen, eine Kurve differenzierbar zu nennen, wenn sie als Vereinigung von Graphen von differenzierbaren Funktionen geschrieben werden kann (wir werden von C^1 -Kurven sprechen, unter der Annahme, dass die Funktionen stetig differenzierbar sind). Man muss hier ein bisschen aufpassen. Es gibt einen Unterschied zwischen der gegebenen Definition von differenzierbarer Kurve und differenzierbarer parametrisierte Kurve. Z.B. die parametrisierte Kurve $\phi(t) = (t^3, t^2) \in \mathbb{R}^2$ ist differenzierbar. Die Bildmenge $T = \{\phi(t) : t \in \mathbb{R}\}$ kann aber neben $(0, 0)$ nicht als Graph einer differenzierbaren Funktion geschrieben werden. Deswegen ist T keine differenzierbare Kurve im obigen Sinne. Wir werden sehen, die Bildmenge der parametrisierten Kurve $\phi(t)$ ist keine differenzierbare Kurve, weil $\phi'(0) = 0$.

Analog kann man Teilmengen von \mathbb{R}^n , die lokal wie \mathbb{R}^2 aussehen, betrachten. In diesem Fall spricht man von Flächen. Auch hier muss man zwischen parametrisierten Flächen und Fläche als Bildmenge von parametrisierten Flächen unterscheiden. Wie für Kurven, kann man Flächen als Vereinigung von Graphen beschreiben. Die Sphäre $S^2 = \{(x, y, z) : x^2 + y^2 + z^2 = 1\} \subset \mathbb{R}^3$ ist nicht der Graph einer einzelne Funktion, kann aber als Vereinigung der folgenden sechs Graphen betrachtet werden, die auf offenen Teilmengen von \mathbb{R}^2 definiert sind: $z = \pm\sqrt{1 - x^2 - y^2}$, definiert auf $\|(x, y)\| < 1$, $y = \pm\sqrt{1 - x^2 - z^2}$ auf $\|(x, z)\| < 1$ und $x = \pm\sqrt{1 - y^2 - z^2}$ definiert auf $\|(y, z)\| < 1$. Weil die Funktionen, aus dessen Graphen S^2 besteht, stetig differenzierbar sind, sagt man S^2 ist eine differenzierbare Fläche, oder eine C^1 -Fläche. Wir erweitern diese Definitionen, um differenzierbare Mannigfaltigkeiten $M \subset \mathbb{R}^n$ der Dimension k einzuführen.

Definition 3.21. Seien $n, k \in \mathbb{N}$, mit $k < n$. Ein C^1 -Mannigfaltigkeitstück der Dimension k in \mathbb{R}^n ist eine Teilmenge von \mathbb{R}^n die, nach allfälliger Unnumerierung der Koordinaten, die Form

$$\{(x_1, \dots, x_n) \in \mathbb{R}^n : (x_{k+1}, \dots, x_n) = \phi(x_1, \dots, x_k) \text{ und } (x_1, \dots, x_k) \in G\}$$

hat, wobei $G \subset \mathbb{R}^k$ offen und zusammenhängend ist und $\phi \in C^1(G; \mathbb{R}^{n-k})$. Eine Menge $M \subset \mathbb{R}^n$ heisst eine C^1 -Mannigfaltigkeit der Dimension k , falls für jede $a \in M$ eine offene Umgebung $U \subset \mathbb{R}^n$ von a existiert, so dass $U \cap M$ ein C^1 -Mannigfaltigkeitstück der Dimension k ist. Eine C^1 -Mannigfaltigkeit der Dimension $k = 1$ heisst eine C^1 -Kurve. Eine C^1 -Mannigfaltigkeit der Dimension $k = 2$ heisst eine C^1 -Fläche. Eine C^1 -Mannigfaltigkeit M der Dimension $n - 1$ in \mathbb{R}^n heisst eine Hyperfläche.

Bemerkung: Kurz gesagt, eine C^1 -Mannigfaltigkeit der Dimension k ist eine Teilmenge von \mathbb{R}^n , die lokal aus dem Graph einer stetig differenzierbaren Abbildung $\phi : \mathbb{R}^k \supset G \rightarrow \mathbb{R}^{n-k}$ gegeben ist.

Beispiele: Der Kreis $S^1 = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 = 1\}$ ist eine C^1 -Kurve. Die Sphäre $S^{n-1} = \{(x_1, \dots, x_n) \in \mathbb{R}^n : \|(x_1, \dots, x_n)\| = 1\} \subset \mathbb{R}^n$ ist eine C^1 -Mannigfaltigkeit der Dimension $(n-1)$ (d.h. S^{n-1} ist eine Hyperfläche). Die Bildmenge $\{(t^3, t^2) \in \mathbb{R}^2 : t \in \mathbb{R}\}$ der parametrisierten Kurve $\phi(t) = (t^3, t^2)$ ist keine C^1 -Mannigfaltigkeit, weil sie in der Nähe von $(0, 0)$ nicht als Graph einer stetig differenzierbaren Funktion geschrieben werden kann. Die Bildmenge der parametrisierten Kurve $\phi(t) = (\cos t, \sin(2t))$ ist keine C^1 -Kurve, weil sie in der Nähe von $\phi(\pi/2) = (0, 0)$ nicht als Graph einer Funktion geschrieben werden kann (nach Definition dürfen Mannigfaltigkeiten keine "Selbst-Durchschnitte" haben).

Statt Teilmengen von \mathbb{R}^n durch Vereinigung von Graphen zu definieren, kann man sie als Lösungsmengen von Gleichungen definieren. Z.B. definiert die Gleichung $x^2 + y^2 = 1$ den Kreis S^1 , also eine Mannigfaltigkeit der Dimension eins. Man kann sich analog vorstellen, dass die Gleichung $f(x_1, \dots, x_n) = 0$, für eine Funktion $f : \mathbb{R}^n \supset U \rightarrow \mathbb{R}$, eine Teilmenge von \mathbb{R}^n definiert, die lokal wie \mathbb{R}^{n-1} aussieht. Im nächsten Satz zeigen wir, dass, falls $a \in U$ die Gleichung $f(a) = 0$ erfüllt, und falls $\nabla f(a) \neq 0$, dann ist die Menge $\{x \in U : f(x) = 0\}$ in der Nähe von a ein Mannigfaltigkeitstück.

Proposition 3.22. *Sei $U \subset \mathbb{R}^n$ offen, $f \in C^1(U)$, und $a \in U$ mit $f(a) = 0$ und $\nabla f(a) \neq 0$. Dann es existiert eine offene Umgebung $G \subset \mathbb{R}^n$ von a so, dass $G \cap \{x \in U : f(x) = 0\}$ ein C^1 -Mannigfaltigkeitstück der Dimension $(n-1)$ ist. D.h. die Lösungsmenge der Gleichung $f(x) = 0$ ist, in der Nähe von a eine Mannigfaltigkeit. Gilt ferner $\nabla f(x) \neq 0$ für alle $x \in U$ mit $f(x) = 0$, dann ist $\{x \in U : f(x) = 0\}$ eine C^1 -Mannigfaltigkeit der Dimension $n-1$.*

Beweis: Die Bedingung $\nabla f(a) \neq 0$ impliziert, dass $j \in \{1, \dots, n\}$ mit $\partial f / \partial x_j(a) \neq 0$ existiert. O.B.d.A. nehmen wir an $\partial f / \partial x_n(a) \neq 0$. Nach dem Satz über implizite Funktionen, existieren eine Umgebung $V \subset \mathbb{R}^{n-1}$ von (a_1, \dots, a_{n-1}) , eine Umgebung $W \subset \mathbb{R}$ von a_n und eine Funktion $\phi \in C^1(V)$ mit Werten in W , so dass

$$\begin{aligned} V \times W \cap \{(x_1, \dots, x_n) \in U : f(x_1, \dots, x_{n-1}, x_n) = 0\} \\ = \{(x_1, \dots, x_{n-1}, \phi(x_1, \dots, x_{n-1})) \in \mathbb{R}^n : (x_1, \dots, x_{n-1}) \in V\}. \end{aligned}$$

Die Behauptung folgt, mit $G = V \times W$. □

Allgemeiner, sei $f : \mathbb{R}^n \subset U \rightarrow \mathbb{R}^m$, für ein $m < n$. Die Gleichung $f(x) = 0$ ist dann ein System von m Gleichungen

$$\begin{cases} f_1(x_1, \dots, x_n) = 0 \\ f_2(x_1, \dots, x_n) = 0 \\ \dots \\ f_m(x_1, \dots, x_n) = 0 \end{cases}$$

Sind die m Gleichungen in geeignetem Sinne unabhängig, so kann man sich vorstellen, dass $f(x) = 0$ eine Mannigfaltigkeit der Dimension $k = n - m$ definiert. Wir müssen verstehen, in welchen Sinne die Gleichungen unabhängig sein müssen. Dazu definieren wir den Begriff vom Rang einer linearen Abbildung.

Definition 3.23. Sei $L \in \mathcal{L}(\mathbb{R}^n; \mathbb{R}^m)$. Der Rang der Matrix L ist

$$\operatorname{Rg} L = \dim \operatorname{Ran}(L) = \dim L(\mathbb{R}^n) = n - \dim \ker(L).$$

Mit anderen Worten, der Rang von L ist die maximale Anzahl von linear unabhängigen Spalten in der Matrix L . Aus der linearen Algebra, $\operatorname{Rg} L$ ist auch die maximale Anzahl von linear unabhängigen Zeilen in L .

Definition 3.24. Sei $U \subset \mathbb{R}^n$ offen und $f \in C^1(U; \mathbb{R}^m)$, mit $m \leq n$. Wir sagen, f ist regulär an der Stelle $a \in U$, falls $Df(a)$ Rang m hat. Ist $m = n$, so ist f genau dann an der Stelle a regulär, wenn $Df(a)$ invertierbar ist.

Satz 3.25. Sei $U \subset \mathbb{R}^n$ offen, $f \in C^1(U; \mathbb{R}^m)$. Sei $a \in U$ mit $f(a) = 0$ und so, dass f regulär an der Stelle a ist. Dann existiert eine offene Umgebung $G \subset U$ von a in \mathbb{R}^n , so dass $G \cap \{x \in U : f(x) = 0\}$ ein Mannigfaltigkeitsstück in \mathbb{R}^n der Dimension $k = n - m$ ist. Ist f an der Stelle x regulär, für alle $x \in U$ mit $f(x) = 0$, dann ist $\{x \in U : f(x) = 0\}$ eine Mannigfaltigkeit in \mathbb{R}^n der Dimension $k = n - m$.

Beweis: Nach Unnumerierung der Koordinaten können wir annehmen, dass die m Spalten

$$\frac{\partial f}{\partial x_{k+1}}(a), \frac{\partial f}{\partial x_{k+2}}(a), \dots, \frac{\partial f}{\partial x_n}(a) \quad (57)$$

linear unabhängig sind (wir haben hier $k = n - m$ gesetzt). Wir schreiben $Df(a) = (D_1f(a), D_2f(a))$, wobei $D_1f(a)$ die $m \times k$ Matrix ist, die aus allen partiellen Ableitungen nach x_1, \dots, x_k besteht und $D_2f(a)$ die $m \times m$ Matrix ist, die aus allen partiellen Ableitungen nach x_{k+1}, \dots, x_n besteht. Die Matrix $D_2f(a)$ ist nach (57) invertierbar. Der Satz über implizite Funktionen impliziert, dass eine offene Umgebung V von (a_1, \dots, a_k) in \mathbb{R}^k , eine offene Umgebung W von (a_{k+1}, \dots, a_n) in \mathbb{R}^m und eine stetig differenzierbare Funktion $\phi : V \rightarrow W$ existieren, so dass

$$V \times W \cap \{x \in U : f(x) = 0\} = \{(x_1, \dots, x_k, \phi(x_1, \dots, x_k)) : (x_1, \dots, x_k) \in V\}$$

Die Behauptung folgt, mit $G = V \times W$. □

Bemerkung: Der Satz zeigt, dass die richtige Verallgemeinerung der Bedingung $\nabla f(a) \neq 0$ in Proposition 3.22 aus der Bedingung $\operatorname{Rg} Df(a) = m$ gegeben ist.

Statt C^1 -Mannigfaltigkeiten durch Graphen von stetig differenzierbaren Funktionen zu definieren, kann man auch lokale reguläre Parametrisierungen (genannt Karten) benutzen. Wir geben hier die alternative Definition, obwohl wir im Folgenden weiter mit der ursprünglichen Definition arbeiten.

Alternative Definition von Mannigfaltigkeiten: Seien $n, k \in \mathbb{N}$, mit $1 \leq k < n$. Eine k -dimensionale C^1 -Mannigfaltigkeit in \mathbb{R}^n (oder eine C^1 -Untermannigfaltigkeit des \mathbb{R}^n) ist eine nicht-leere Menge $M \subset \mathbb{R}^n$, so dass für alle $a \in M$ eine offene Umgebung $U \subset \mathbb{R}^n$ von a , eine offene Menge $G \subset \mathbb{R}^k$ und eine reguläre Abbildung $\varphi \in C^1(G; \mathbb{R}^n)$ so, dass $\varphi(G) = M \cap U$ und $\varphi : G \rightarrow M \cap U$ ein Homöomorphismus ist. Das Paar (G, φ) heisst eine Karte von M in der Nähe vom Punkt a . Übung: Zeigen Sie, dass diese Definition mit der ursprünglichen Definition übereinstimmt.

Bemerkung: Die alternative Definition von Mannigfaltigkeiten durch Karten lässt sich auch zum Fall $k = n$ erweitern. Dann gilt: Eine Teilmenge $M \subset \mathbb{R}^n$ ist genau dann eine Mannigfaltigkeit der Dimension n , wenn M offen ist.

Bemerkung: Statt von C^1 -Mannigfaltigkeit in \mathbb{R}^n spricht man in der Literatur oft von C^1 -Untermannigfaltigkeit des \mathbb{R}^n . Man benutzt das Wort Untermannigfaltigkeit, weil die Mengen, die wir betrachten, immer Teilmengen von \mathbb{R}^n sind. Das Wort Mannigfaltigkeit der Dimension k wird dann für allgemeinere Mengen benutzt, die lokal das Bild einer regulären Funktion auf einer offenen Teilmenge von \mathbb{R}^k sind. Ein wichtiges Resultat der Differentialgeometrie besagt dann, dass jede C^1 -Mannigfaltigkeit der Dimension n in \mathbb{R}^{2n} eingebettet werden kann (Einbettungssatz von Whitney). Bemerke, dass die Dimension (mindestens) $2n$ sein muss; die Klein'sche Flasche ist ein berühmtes Beispiel einer zwei dimensionale C^1 -Mannigfaltigkeit, die nicht in \mathbb{R}^3 eingebettet werden kann. In dieser Vorlesung werden wir immer C^1 -Untermannigfaltigkeiten des \mathbb{R}^n betrachten; wir werden aber das Wort C^1 -Mannigfaltigkeiten in \mathbb{R}^n benutzen.

Sei nun $M \subset \mathbb{R}^n$ eine Mannigfaltigkeit der Dimension $k < n$ und $a \in M$ ein Punkt auf der Mannigfaltigkeit. Wir möchten den Begriff von Tangentialraum zu M an $a \in M$ einführen. Dazu betrachten wir parametrisierte Kurven auf M , die durch a gehen. Sei I ein offenes Intervall in \mathbb{R} . Eine differenzierbare parametrisierte Kurve auf M ist eine differenzierbare Abbildung $\varphi : I \rightarrow \mathbb{R}^n$, mit $\varphi(t) \in M$ für alle $t \in I$. Es gelte $\varphi(t_0) = a$. Der Vektor

$$\varphi'(t_0) = (\varphi'_1(t_0), \dots, \varphi'_n(t_0))$$

heißt der Tangentialvektor zu der Kurve φ im Punkt $\varphi(t_0) = a$.

Definition 3.26. Der Tangentialraum $T_a(M)$ zu M an der Stelle $a \in M$ besteht aus allen Tangentialvektoren zu differenzierbaren Kurven auf M , die durch a gehen. Mit anderen Worten, $\xi \in \mathbb{R}^n$ ist genau dann Element von $T_a(M)$, wenn ein Intervall $I \subset \mathbb{R}$, ein $t_0 \in I$ und eine Kurve $\varphi \in C^1(I; \mathbb{R}^n)$ mit $\varphi(t) \in M$ für alle $t \in I$, $\varphi(t_0) = a$ und $\varphi'(t_0) = \xi$ existieren.

Proposition 3.27. Sei M eine Mannigfaltigkeit in \mathbb{R}^n der Dimension $k < n$ und $a \in M$. $T_a(M)$ ist ein linearer Unterraum von \mathbb{R}^n , mit $\dim T_a(M) = k$.

Beweis: Die Definition von $T_a(M)$ hängt nur von M in der Nähe von a ab. D.h. $T_a(M) = T_a(M \cap U)$ für jede, beliebig kleine Umgebung $U \subset \mathbb{R}^n$ von a . Aus diesem Grund können wir annehmen, dass eine offene Menge $V \subset \mathbb{R}^k$ und ein $\psi \in C^1(V; \mathbb{R}^n)$ existieren, so dass

$$M = \{(x, \psi(x)) : x \in V\}$$

Wir führen die Notation $x^{(1)} = (x_1, \dots, x_k)$ und $x^{(2)} = (x_{k+1}, \dots, x_n)$ ein. Dann $a = (a^{(1)}, a^{(2)})$, wobei $a^{(2)} = \psi(a^{(1)})$. Sei nun $t \rightarrow \varphi(t)$ eine differenzierbare Kurve auf M , mit $\varphi(t_0) = a$. Wir bezeichnen $\phi_1(t) = (\varphi_1(t), \dots, \varphi_k(t))$ und $\phi_2(t) = (\varphi_{k+1}(t), \dots, \varphi_n(t))$, wobei $\varphi_1, \dots, \varphi_n$ die Komponenten von φ sind. Da $\varphi(t) \in M$ für alle t , muss gelten $\phi_2(t) = \psi(\phi_1(t))$ für alle t genügend nahe zu t_0 . Also $\varphi(t) = (\phi_1(t), \psi(\phi_1(t)))$ und

$$\varphi'(t_0) = (\phi'_1(t_0), D\psi(a^{(1)})(\phi'_1(t_0)))$$

Wir setzen $v = \phi'_1(t_0) \in \mathbb{R}^k$. Dann ist $\varphi'(t_0) = (v, D\psi(a^{(1)})v) \in \mathbb{R}^n$.

Andererseits, für gegebene $v \in \mathbb{R}^k$, können wir die Kurve $\varphi(t) = (a^{(1)} + tv, \psi(a^{(1)} + tv))$ definieren. Dann ist φ offenbar eine differenzierbare Kurve auf M , mit $\varphi(0) = a$ und $\varphi'(0) = (v, D\psi(a^{(1)})(v))$. Wir haben also gezeigt, dass

$$T_a(M) = \{(v, D\psi(a^{(1)})v) : v \in \mathbb{R}^k\}$$

$T_a(M)$ ist also ein linearer Raum, mit Dimension k (die Vektoren $(e_i, D\psi(a^{(1)})(e_i))$, für $i = 1, \dots, k$ sind eine Basis von $T_a(M)$, falls e_i , $i = 1, \dots, k$, die Standard-Basis von \mathbb{R}^k bezeichnet). \square

Falls die Mannigfaltigkeit M als Lösungsmenge einer Gleichung $f(x) = 0$ gegeben ist, so kann man eine andere Charakterisierung von $T_a(M)$ angeben. Sei zunächst $U \subset \mathbb{R}^n$ und $f \in C^1(U)$ reelwertig. Es gelte $\nabla f(x) \neq 0$ für alle $x \in U$ mit $f(x) = 0$. Dann ist $M = \{x \in U : f(x) = 0\}$ eine Mannigfaltigkeit der Dimension $n - 1$. Wir sind im Tangentialraum $T_a(M)$ interessiert, für ein $a \in M$. Wir wissen schon $T_a(M)$ ist ein linearer Raum mit $\dim T_a(M) = n - 1$. Sei $\varphi(t)$ eine parametrisierten Kurve auf M , mit $\varphi(t_0) = a$. Dann gilt $f(\varphi(t)) = 0$ für alle t . Wir erhalten:

$$0 = \frac{d}{dt} f(\varphi(t))|_{t=t_0} = \nabla f(\varphi(t_0)) \cdot \varphi'(t_0)$$

und deswegen

$$T_a(M) \subset \{\xi \in \mathbb{R}^n : \xi \cdot \nabla f(a) = 0\}$$

Da der Raum $\{\xi \in \mathbb{R}^n : \xi \cdot \nabla f(a) = 0\}$ auch Dimension $n - 1$ hat, gilt

$$T_a(M) = \{\xi \in \mathbb{R}^n : \xi \cdot \nabla f(a) = 0\}$$

Wir haben bewiesen, dass der Gradient $\nabla f(a)$ senkrecht zum Tangentialraum $T_a(M)$ steht.

Sei nun $f \in C^1(U; \mathbb{R}^m)$ regulär an der Stelle x , für alle $x \in U$ mit $f(x) = 0$. Dann ist $M = \{x \in U : f(x) = 0\}$ eine C^1 -Mannigfaltigkeit der Dimension $k = n - m$. Seien f_1, \dots, f_m die Komponenten von f . Ähnlich wie oben, gilt

$$T_a(M) = \{\xi \in \mathbb{R}^n : \nabla f_1(a) \cdot \xi = \dots = \nabla f_m(a) \cdot \xi = 0\}$$

Da die m Zeilen $\nabla f_j(a)$, $j = 1, \dots, m$, von $Df(a)$ linear unabhängig sind (weil f regulär ist), ist es klar, dass der Raum $\{\xi \in \mathbb{R}^n : \nabla f_1(a) \cdot \xi = \dots = \nabla f_m(a) \cdot \xi = 0\}$ die Dimension $k = n - m$ hat.

Es ist manchmal auch nützlich, neben dem Begriff vom Tangentialraum $T_a(M)$ auch den Begriff der Tangentialebene einzuführen. Die Tangentialebene an der Mannigfaltigkeit M im Punkt a ist die Teilmenge von \mathbb{R}^n , die aus allen Tangentialvektoren zu M in a besteht, die aber vom Punkt a ausgehen. Mit anderen Worten,

$$\text{Tangentialebene zu } M \text{ in } a = \{a + \xi \in \mathbb{R}^n : \xi \in T_a(M)\}$$

Falls $M = \{x \in U : f(x) = 0\}$ für eine Funktion $f \in C^1(U; \mathbb{R}^m)$ regulär im Punkt a , dann gilt

$$\begin{aligned} \text{Tangentialebene zu } M \text{ in } a &= \{a + \xi \in \mathbb{R}^n : \xi \cdot \nabla f_j(a) = 0 \text{ für alle } j = 1, \dots, m\} \\ &= \{\xi \in \mathbb{R}^n : (\xi - a) \cdot \nabla f_j(a) = 0 \text{ für alle } j = 1, \dots, m\} \end{aligned}$$

Bemerke, dass im Gegensatz zum Tangentialraum $T_a(M)$, die Tangentialebene kein linearer Raum ist.

3.6 Extrema mit Nebenbedingungen

Wir betrachten eine reelwertige Funktion f , definiert auf einer offenen Teilmenge $\Omega \subset \mathbb{R}^n$. In diesem Abschnitt möchten wir Extrema von $f(x)$ finden, unter der zusätzlichen Nebenbedingung $x \in M$, wobei $M \subset \mathbb{R}^n$ eine Mannigfaltigkeit mit Dimension $k < n$, enthalten in Ω , ist.

Wir sagen $a \in \Omega$ ist ein lokales bedingtes Minimum von f mit der Nebenbedingung $x \in M$, falls eine offene Umgebung $A \subset \Omega$ in \mathbb{R}^n existiert, so dass $f(a) = \min\{f(x) : x \in A \cap M\}$. Analog definiert man den Begriff vom lokalen bedingten Maximum. $a \in \Omega$ heisst ein lokales bedingtes Extremum von f , falls a entweder ein lokales bedingtes Minimum oder ein lokales bedingtes Maximum ist. Nehmen wir an $f \in C^1(\Omega)$. Wegen der Nebenbedingungen können bedingte Extrema $a \in M$ existieren, für die $\nabla f(a) \neq 0$. Wir suchen also andere Kriterien, um Extrema mit Nebenbedingungen zu charakterisieren.

Manchmal kann man dieses Problem einfach lösen, indem man eine Parametrisierung von M benutzt. Sei z.B. $f \in C^1(\mathbb{R}^2)$ und nehmen wir an, wir suchen

$$s := \sup\{f(x_1, x_2) : x_1^2 + x_2^2 = 1\}$$

Dann können wir einen Parameter $t \in [0, 2\pi)$ einführen und $x_1 = \cos t, x_2 = \sin t$ schreiben. Damit ist das gesuchte Supremum s durch

$$s = \sup_{t \in [0, 2\pi)} g(t)$$

, wobei $g(t) = f(\cos t, \sin t)$, gegeben. Das Problem mit Nebenbedingungen wurde damit zu einem Problem (in einer Dimension, statt zwei) ohne Nebenbedingungen reduziert. Extrema sind nun durch $g'(t) = 0$ charakterisiert.

Analog kann man die Nebenbedingung entfernen, falls M explizit als Graph einer differenzierbaren Funktion geschrieben werden kann. Nehmen wir an, wir suchen

$$s := \sup\{x^2yz : x^2 + y^2 + z^2 = 1, x, y, z \geq 0\}$$

Das Maximum wird angenommen, weil die Funktion $f(x, y, z) = x^2yz$ stetig und die Menge $\{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 + z^2 = 1, x, y, z \geq 0\}$ kompakt ist. Das Maximum hat offenbar $x, y, z > 0$. Wir haben

$$\{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 + z^2 = 1, x, y, z > 0\} = \{(\sqrt{1 - y^2 - z^2}, y, z) : (y, z) \in G\}$$

mit der offenen Menge $G = \{(y, z) \in \mathbb{R}^2 : y^2 + z^2 < 1, y, z > 0\}$. Wir haben also

$$s = \sup\{g(y, z) : (y, z) \in G\},$$

wobei $g(y, z) = f(1 - y^2 - z^2, y, z) = (1 - y^2 - z^2)yz$. Um s zu finden, berechnen wir also

$$\nabla g(y, z) = (z - 3y^2z - z^3, y - 3z^2y - y^3)$$

Die Bedingung $\nabla g(y, z) = 0$ gibt, nach kurzer Rechnung, $y = z = 1/2$. Das Maximum der Funktion f wird also im Punkt $(x, y, z) = (1/\sqrt{2}, 1/2, 1/2)$ angenommen und beträgt $f(1/\sqrt{2}, 1/2, 1/2) = 1/8$.

Das letzte Beispiel war einfach, weil wir die Mannigfaltigkeit explizit als Graph schreiben konnten. Das ist natürlich nicht immer möglich. Deswegen ist der folgende Satz nützlich, um Extrema einer differenzierbaren Funktion f unter der Nebenbedingung $g(x) = 0$ zu finden.

Proposition 3.28 (Lagrange-Multiplikatoren, eine Nebenbedingung). *Sei $U \subset \mathbb{R}^n$ offen, $g \in C^1(U)$ reelwertig, und $M = \{x \in U : g(x) = 0\}$. Sei $a \in M$, so dass $\nabla g(a) \neq 0$. Dann ist M in der Nähe von a eine Mannigfaltigkeit der Dimension $n - 1$ (eine Hyperfläche). Sei nun f eine reelwertige Funktion, definiert und differenzierbar in der Nähe von a . Sei a eine bedingte lokale Extremalstelle von f auf M . Dann existiert $\lambda \in \mathbb{R}$ mit*

$$\nabla f(a) = \lambda \nabla g(a)$$

Der Parameter λ heisst ein Lagrange-Multiplikator.

Proof. Nehmen wir an $\nabla f(a)$ ist nicht proportional zu $\nabla g(a)$. Dann existiert $\xi \in \mathbb{R}^n$ mit $\xi \cdot \nabla g(a) = 0$ aber $\xi \cdot \nabla f(a) \neq 0$. Die Bedingung $\xi \cdot \nabla g(a) = 0$ impliziert, dass $\xi \in T_a(M)$. Also existiert eine parametrisierte Kurve $\varphi : I \rightarrow M$, mit $\varphi(t_0) = a$ und $\varphi'(t_0) = \xi$ für ein $t_0 \in I$. Betrachte nun die Funktion $h(t) = f(\varphi(t))$. Es gilt

$$h'(t_0) = \nabla f(\varphi(t_0)) \cdot \varphi'(t_0) = \nabla f(a) \cdot \xi \neq 0$$

Das heisst, in jeder Umgebung von t_0 nimmt $h(t)$ Werte kleiner und grösser als $h(t_0) = f(a)$ an. Das zeigt, dass f auf $U \cap M$ Werte kleiner und grösser als $f(a)$ annimmt, für jede offene Umgebung U von a . Deswegen ist a keine bedingte Extremalstelle von f . \square

Die Proposition gibt uns eine Rezept, um Extrema einer differenzierbaren Funktion f mit der Nebenbedingung $g(x) = 0$ zu finden. Die Gleichungen

$$\begin{cases} \nabla f(x) = \lambda \nabla g(x) \\ g(x) = 0 \end{cases}$$

bilden ein System von $(n + 1)$ -Gleichungen für die $(n + 1)$ Unbekannten x_1, \dots, x_n, λ . Lösungen dieses Gleichungssystemes, zusammen mit Punkten $x \in \mathbb{R}^n$ mit $g(x) = 0$, wo f nicht differenzierbar ist, sind dann die einzigen möglichen Kandidaten für bedingte lokale Extremalstellen.

Beispiel: Betrachten wir noch einmal das Beispiel von oben. Wir suchen

$$s := \sup\{x^2yz : x^2 + y^2 + z^2 = 1, x, y, z \geq 0\}$$

Wir setzen $f(x, y, z) = x^2yz$ und $g(x, y, z) = x^2 + y^2 + z^2 - 1$. Bei bedingten Extrema muss gelten: $\nabla f(x, y, z) = \lambda \nabla g(x, y, z)$ für ein $\lambda \in \mathbb{R}$. Das ergibt die vier Gleichungen

$$\begin{cases} 2xyz & = 2\lambda x \\ x^2z & = 2\lambda y \\ x^2y & = 2\lambda z \\ x^2 + y^2 + z^2 & = 1 \end{cases}$$

Da $x, y, z > 0$ gelten muss, implizieren die zweite und dritte Gleichung, dass $y = z$. Die erste Gleichung gibt dann $\lambda = y^2$, und damit liefert die dritte Gleichung $x^2 = 2y^2$. Einsetzen in die letzte Gleichung ergibt $4y^2 = 1$, und damit $y = z = 1/2$ und $x = 1/\sqrt{2}$.

Bis jetzt haben wir Probleme mit einer Nebenbedingung der Form $g(x) = 0$ untersucht, für eine reellwertige Funktion g . Im Folgenden untersuchen wir den Fall, dass g Werte in \mathbb{R}^m hat, für ein $m < n$. In diesem Fall ist die vektorielle Gleichung $g(x) = 0$ eigentlich ein System mit m Gleichungen. Man spricht dann von m Nebenbedingungen. Auch in diesem Fall kann man Extrema finden, indem man Lagrange-Multiplikatoren einführt.

Proposition 3.29 (Lagrange Multiplikatoren, m Nebenbedingungen). *Sei $U \subset \mathbb{R}^n$ offen, $g \in C^1(U; \mathbb{R}^m)$ für ein $m < n$. Sei $M = \{x \in U : g(x) = 0\}$. Sei $a \in M$, so dass g regulär an der Stelle a ist (d.h. $\text{Rg}(Dg(a)) = m$; in diesem Fall ist M in der Nähe von a ein Mannigfaltigkeitsstück der Dimension $k = n - m$). Sei f eine reellwertige C^1 -Funktion, definiert in einer Umgebung von a in \mathbb{R}^n . Sei a eine lokale bedingte Extremalstelle von f auf M . Dann ist $\nabla f(a)$ eine lineare Kombination von $\nabla g_1(a), \dots, \nabla g_m(a)$. D.h. es existieren $\lambda_1, \dots, \lambda_m \in \mathbb{R}$ mit*

$$\nabla f(a) = \sum_{j=1}^m \lambda_j \nabla g_j(a)$$

wobei $g_1, \dots, g_m : U \rightarrow \mathbb{R}$ die Komponenten von g sind.

Beweis: Nehmen wir an, dass $\nabla f(a)$ nicht eine lineare Kombination von $\nabla g_1(a), \dots, \nabla g_m(a)$ ist. Dann existiert $\xi \in \mathbb{R}^n$, so dass $\xi \cdot \nabla g_i(a) = 0$ für alle $i = 1, \dots, m$ und $\xi \cdot \nabla f(a) \neq 0$. Das impliziert, dass $\xi \in T_a(M)$. Damit existiert eine differenzierbare parametrisierte Kurve $\varphi : I \rightarrow M$, so dass $\varphi(t_0) = a$ und $\varphi'(t_0) = \xi$. Sei nun $h(t) = f(\varphi(t))$. Dann ist

$$h'(t_0) = \nabla f(a) \cdot \xi \neq 0$$

Damit ist t_0 keine Extremalstelle von $f(\varphi(t))$. Das zeigt, dass a keine bedingte lokale Extremalstelle von f sein kann. \square

Um bedingte Extremalstellen einer Funktion f auf \mathbb{R}^n mit m Nebenbedingungen $g(x) = 0$ zu finden, muss man also die $n + m$ Gleichungen

$$\begin{cases} \nabla f(x) = \sum_{j=1}^m \lambda_j \nabla g_j(x) \\ g(x) = 0 \end{cases}$$

für die $(n + m)$ Unbekannten $x_1, \dots, x_n, \lambda_1, \dots, \lambda_m$ lösen. Die einzigen Kandidaten für Extremalstellen sind Lösungen dieser Gleichungen und Punkte auf M , wo f nicht differenzierbar ist.

Beispiel (aus der statistischen Mechanik): Ein Molekül habe n mögliche Zustände, mit Energien E_1, \dots, E_n . In der statistischen Mechanik wird ein System von vielen Molekülen durch die Wahrscheinlichkeiten $p_1, \dots, p_n \in [0; 1]$ beschrieben, dass ein Molekül im j -ten Zustand gefunden wird (die Wahrscheinlichkeit p_j gibt die Fraktion zwischen den

Molekülen im j -ten Zustand und der gesamten Anzahl von Molekülen). Der Zustand des Vielteilchensystems wird durch Maximierung der Entropie

$$H(p_1, \dots, p_n) = - \sum_{j=1}^n p_j \log p_j$$

unter den Nebenbedingungen $p_1 + \dots + p_n = 1$ und

$$\bar{E} = \sum_{j=1}^n p_j E_j$$

bestimmt (wir möchten also die Entropie bei fester mittlerer Energie \bar{E} maximieren).

Sei $p = (p_1, \dots, p_n)$. Wir setzen $g_1(p) = p_1 + \dots + p_n - 1$ und $g_2(p) = \sum_{j=1}^n p_j E_j$. Es gilt $\nabla g_1(p) = (1, 1, \dots, 1)$ und $\nabla g_2(p) = (E_1, \dots, E_n)$. Da

$$\nabla H(p) = (-\log p_1 - 1, \dots, -\log p_n - 1)$$

finden wir aus der Gleichung $\nabla H(p) = \lambda_1 \nabla g_1(p) + \lambda_2 \nabla g_2(p)$, dass

$$-1 - \log p_j = \lambda_1 + \lambda_2 E_j$$

für alle $j = 1, \dots, n$. Das ergibt $\log p_j = -1 - \lambda_1 - \lambda_2 E_j$ und also

$$p_j = e^{-(\lambda_1+1)} \cdot e^{-\lambda_2 E_j} =: k e^{-\lambda_2 E_j}$$

Die Bedingung $g_1(p) = 0$ bestimmt die Konstante k . Wir finden

$$p_j = \frac{e^{-\lambda_2 E_j}}{\sum_{j=1}^n e^{-\lambda_2 E_j}}$$

Die Bedingung $g_2(p) = 0$ ergibt die Gleichung

$$\bar{E} = \sum_{j=1}^n \frac{E_j e^{-\lambda_2 E_j}}{\sum_{j=1}^n e^{-\lambda_2 E_j}}$$

Diese Gleichung erlaubt uns im Prinzip, λ_2 zu bestimmen. In der statistischen Mechanik setzte man $\lambda_2 = k_B/T$, wobei k_B die sogenannte Boltzmann-Konstante ist, und T die absolute Temperatur des Systems ist. Die Temperatur wird von der mittleren Energie bestimmt. Der Zustand vom System zur Temperatur T (d.h. zur mittleren Energie \bar{E}) ist der sogenannte Gibbs-Zustand, charakterisiert durch die Wahrscheinlichkeiten

$$p_j = \frac{e^{-k_B E_j/T}}{\sum_{i=1}^n e^{-k_B E_i/T}}.$$

Lagrange-Multiplikatoren und die oben erklärten Strategien, um bedingte Extrema zu finden, sind auch nützlich, um (globale) Extrema von Funktionen auf Teilmengen von \mathbb{R}^n zu suchen, die einen Rand haben. Wir betrachten ein Beispiel, um die Situation zu erklären.

Beispiel: Wir suchen das globale Maximum und das globale Minimum von $f(x, y, z) = x^2 + y + z^2$ auf $D := \{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 + z^2 \leq 1 \text{ und } x^2 + y^2 \leq 1/2\}$.

Wir zerlegen D in verschiedenen Teile. Wir setzen:

$$\begin{aligned} D_1 &= \{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 + z^2 < 1 \text{ und } x^2 + y^2 < 1/2\}, \\ D_2 &= \{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 + z^2 = 1 \text{ und } z > 1/\sqrt{2}\}, \quad D_3 = -D_2 \\ D_4 &= \{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 = 1/2, z \in (-1/\sqrt{2}; 1/\sqrt{2})\}, \\ D_5 &= \{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 = 1/2, z = 1/\sqrt{2}\}, \quad D_6 = -D_5. \end{aligned}$$

Wir suchen separate lokale Extrema in den Mengen D_1, \dots, D_5 . Das globale Maximum von f in D ist das grösste lokale Maximum, das wir in D_1, \dots, D_6 finden werden, und analog für das globale Minimum von f .

- 1) Extrema auf D_1 werden durch die Bedingung $\nabla f(x) = 0$ charakterisiert. Da $\nabla f(x) = (2x, 1, 2z)$, existiert kein Punkt $x \in D_1$ mit $\nabla f(x) = 0$. Also gibt es kein lokales Extremum in D_1 .
- 2) Lokale Extrema auf D_2 werden durch die Nebenbedingung $g(x, y, z) = x^2 + y^2 + z^2 = 1$ charakterisiert. Ist $x \in D_2$ ein lokales bedingtes Extremum, dann muss $\lambda \in \mathbb{R}$ existieren mit $\nabla f(x) = \lambda \nabla g(x)$. Mit $\nabla g(x) = 2(x, y, z)$ wir finden

$$\begin{cases} 2x & = 2\lambda x \\ 1 & = 2\lambda y \\ 2z & = 2\lambda z \\ x^2 + y^2 + z^2 & = 1 \end{cases}$$

Da $z \neq 0$ in D_2 , folgt $\lambda = 1$. Das ergibt $y = 1/2$ und $x^2 + z^2 = 3/4$. Der Kreis $\{(x, 1/2, z) : x^2 + z^2 = 3/4\}$ schneidet D_2 wenn $z > 1/\sqrt{2}$. Auf diesem Kreis ist f konstant, gegeben aus

$$f(x, 1/2, z) = 5/4, \quad \text{für alle } (x, z) \in \mathbb{R}^2 \text{ mit } x^2 + z^2 = 3/4.$$

- 3) Aus Symmetrie, kann die Funktion f auf D_3 nur auf dem Kreis $\{(x, 1/2, z) : x^2 + z^2 = 3/4\}$ extremal sein, wo $f = 5/4$ ist.
- 4) Wir betrachten nun die Teilmenge D_4 , wo Extrema von f mit der Nebenbedingung $g(x, y, z) = x^2 + y^2 - 1/2 = 0$ gesucht werden sollen. Die Gleichung $\nabla f(x) = \lambda \nabla g(x)$ wird zu

$$\begin{cases} 2x & = 2\lambda x \\ 1 & = 2\lambda y \\ 2z & = 0 \\ x^2 + y^2 & = 1/2 \end{cases}$$

Wir finden die Lösungen $(x, y, z) = (0, \pm 1/\sqrt{2}, 0)$ und $(x, y, z) = (\pm 1/2, 1/2, 0)$. Es gilt

$$\begin{aligned} f(0, \pm 1/\sqrt{2}, 0) &= \pm 1/\sqrt{2} \\ f(\pm 1/2, 1/2, 0) &= 3/4. \end{aligned}$$

- 5) Auf D_5 haben wir die Nebenbedingungen $g_1(x, y, z) = x^2 + y^2 - 1/2 = 0$ und $g_2(x, y, z) = z - 1/\sqrt{2} = 0$. Die Gleichung $\nabla f(x) = \lambda_1 \nabla g_1(x) + \lambda_2 \nabla g_2(x)$ gibt

$$\begin{cases} 2x & = 2\lambda_1 x \\ 1 & = 2\lambda_1 y \\ 2z & = \lambda_2 \\ x^2 + y^2 & = 1/2 \\ z & = 1/\sqrt{2} \end{cases}$$

Es gibt die Lösungen $(\pm 1/2, 1/2, 1/\sqrt{2})$ und $(0, \pm 1/\sqrt{2}, 1/\sqrt{2})$. Es gilt

$$\begin{aligned} f(\pm 1/2, 1/2, 1/\sqrt{2}) &= 5/4 \\ f(0, \pm 1/\sqrt{2}, 1/\sqrt{2}) &= \pm 1/\sqrt{2} + 1/2. \end{aligned}$$

- 6) Aus Symmetrie können Extrema in D_6 nur an den Stellen $(\pm 1/2, 1/2, -1/\sqrt{2})$ und $(0, \pm 1/\sqrt{2}, -1/\sqrt{2})$ gefunden werden. Wie in D_5 finden wir $f(\pm 1/2, 1/2, -1/\sqrt{2}) = 5/4$ und $f(0, \pm 1/\sqrt{2}, -1/\sqrt{2}) = \pm 1/\sqrt{2} + 1/2$.

Durch Vergleich der Werte von f in den gefundenen Kandidaten für Maxima und Minima, finden wir

$$\max_{x \in D} f(x) = 5/4, \quad \text{und} \quad \min_{x \in D} f(x) = -\frac{1}{\sqrt{2}}$$

Das Maximum wird auf dem Kreisbogen $y = 1/2, x^2 + z^2 = 3/4$, mit $|z| \geq 1/\sqrt{2}$ angenommen. Das Minimum wird in $(0, -1/\sqrt{2}, 0)$ angenommen.

3.7 Integrale, die von einem Parameter abhängen.

Sei $f(x, y)$ eine stetige Funktion von zwei Variablen, definiert auf einer Produktmenge $I \times U$, wobei $I \subset \mathbb{R}$ ein Intervall ist und $U \subset \mathbb{R}^n$. Für feste $y \in U$ können wir dann f über x integrieren (weil $f(x, y)$ für feste y als Funktion von x stetig und deswegen auch integrierbar ist). Das Resultat ist eine Funktion auf U . Wir möchten in diesem Abschnitt einige Eigenschaften von dieser Funktion diskutieren.

Satz 3.30. *Seien $a, b \in \mathbb{R}, a < b, U \subset \mathbb{R}^n$ und $f \in C([a; b] \times U)$. Dann ist die Funktion $\varphi : U \rightarrow \mathbb{R}$, definiert durch*

$$\varphi(y) = \int_a^b f(x, y) dx \tag{58}$$

stetig.

Beweis: Sei $y \in U$ und y_n eine Folge in U , mit $y_n \rightarrow y$. O.B.d.A. können wir $r > 0$ finden, mit $y_n \in \overline{B}_r(y) = \{z \in \mathbb{R}^n : \|z - y\| \leq r\} \subset U$ für alle $n \in \mathbb{N}$ (sonst betrachten wir nur y_n für n gross genug). Wir setzen dann $F_n(x) := f(x, y_n)$ und $F(x) := f(x, y)$. Aus Stetigkeit von f gilt offenbar $F_n(x) \rightarrow F(x)$ punktweise. Wir behaupten nun $F_n \rightarrow F$ gleichmässig auf $[a; b]$. Da $[a; b] \times \overline{B}_r(y)$ kompakt ist, ist f auf $[a; b] \times \overline{B}_r(y)$ gleichmässig stetig. Für $\varepsilon > 0$ fest gewählt existiert also $\delta > 0$, so dass $|f(x, y) - f(z, w)| < \varepsilon$ für alle $(x, y), (z, w) \in [a; b] \times \overline{B}_r(y)$ mit $\|(x, y) - (z, w)\| < \delta$. Wir finden nun $N \in \mathbb{N}$ mit

$\|y_n - y\| < \delta$ für alle $n > N$. Dann gilt auch $\|(x, y_n) - (x, y)\| < \delta$ für alle $n > N$.
Deswegen muss

$$|F_n(x) - F(x)| = |f(x, y_n) - f(x, y)| < \varepsilon$$

für alle $n > N$ und alle $x \in [a; b]$. Das zeigt die gleichmässige Konvergenz von $F_n \rightarrow F$.
Aus Analysis 1 (Satz 9.13) folgt, dass

$$\lim_{n \rightarrow \infty} \int_a^b F_n(x) dx = \int_a^b F(x) dx$$

und also, dass $\varphi(y_n) \rightarrow \varphi(y)$. Damit ist φ stetig. □

Ist f nach y differenzierbar, so ist auch φ , definiert wie in (58), differenzierbar. Das ist der Inhalt vom nächsten Satz.

Satz 3.31. Seien $a, b \in \mathbb{R}$, $a < b$, $n, m \in \mathbb{N}$, $U \subset \mathbb{R}^n$ offen und $f : [a; b] \times U \rightarrow \mathbb{R}^m$ stetig.
Wir nehmen an, $\partial f / \partial y_j(x, y)$ existiert und ist stetig auf $[a; b] \times U$ für ein $j \in \{1, \dots, n\}$.
Dann hat auch die Funktion

$$\varphi(y) = \int_a^b f(x, y) dx$$

eine stetige partielle Ableitung $\partial \varphi / \partial y_j(y)$ auf U , gegeben aus

$$\frac{\partial \varphi}{\partial y_j}(y) = \int_a^b dx \frac{\partial f}{\partial y_j}(x, y).$$

Ist ferner f nach y_1, \dots, y_n partiell differenzierbar und sind alle partiellen Ableitungen $\partial f / \partial y_j(x, y)$ stetig auf $[a, b] \times U$, dann ist $\varphi \in C^1(U; \mathbb{R}^m)$ mit

$$D\varphi(y) = \int_a^b dx D_y f(x, y)$$

wobei $D_y f(x, y)$ die $m \times n$ Matrix ist, deren Einträge aus $\partial f_i / \partial y_j(x, y)$ gegeben sind.

Beweis: O.B.d.A. betrachten wir den Fall $m = 1$ (sonst wiederholen wir das Argument für die m Komponenten von $f = (f_1, \dots, f_m)$). Sei $y \in U$ festgewählt. Wir finden $\rho > 0$, so dass $y + te_j \in U$ für alle $t \in [-\rho, \rho]$. Für beliebige $|t| \leq \rho$ haben wir

$$\frac{\varphi(y + te_j) - \varphi(y)}{t} = \int_a^b dx \frac{f(x, y + te_j) - f(x, y)}{t} \tag{59}$$

Aus dem Mittelwertsatz existiert für jede $t \in [-\rho; \rho]$ und $x \in [a; b]$ ein $s(t, x)$ mit $|s(t, x)| \leq |t|$, so dass

$$\frac{f(x, y + te_j) - f(x, y)}{t} = \frac{\partial f}{\partial y_j}(x, y + s(t, x)e_j)$$

Die Abbildung $(x, \lambda) \rightarrow (\partial f / \partial y_j)(x, y + \lambda e_j)$ ist stetig und damit auf der kompakten Menge $[a; b] \times [-\rho, \rho]$ gleichmässig stetig. Für ein beliebiges $\varepsilon > 0$ gibt es also ein $0 < \delta < \rho$, so dass

$$\left| \frac{\partial f}{\partial y_j}(x, y + \lambda e_j) - \frac{\partial f}{\partial y_j}(x, y) \right| \leq \varepsilon$$

für alle $|\lambda| \leq \delta$ und alle $x \in [a; b]$. Damit gilt

$$\left| \frac{f(x, y + te_j) - f(x, y)}{t} - \frac{\partial f}{\partial y_j}(x, y) \right| = \left| \frac{\partial f}{\partial y_j}(x, y + s(t, x)e_j) - \frac{\partial f}{\partial y_j}(x, y) \right| \leq \varepsilon$$

für alle $|t| \leq \delta$ und $x \in [a; b]$ (weil dann $|s(t, x)| \leq |t| \leq \delta$) ist. Damit gilt

$$\frac{f(x, y + te_j) - f(x, y)}{t} \rightarrow \frac{\partial f}{\partial y_j}(x, y)$$

für $t \rightarrow 0$, gleichmässig in x . Aus Analysis 1 (Satz 9.13) folgt, dass

$$\lim_{t \rightarrow 0} \int_a^b dx \frac{f(x, y + te_j) - f(x, y)}{t} \rightarrow \int_a^b dx \frac{\partial f}{\partial y_j}(x, y).$$

Aus (59) folgt, dass φ nach y_j partiell differenzierbar ist, und dass

$$\frac{\partial \varphi}{\partial y_j}(y) = \int_a^b dx \frac{\partial f}{\partial y_j}(x, y).$$

Da $\partial f / \partial y_j$ stetig ist, folgt die Stetigkeit von $\partial \varphi / \partial y_j$ aus Satz 3.30. Die andere Behauptung folgt aus Proposition 3.5. \square

Wir untersuchen nun Funktionen $f(x, y)$ auf der Produktmenge $[a; b] \times [c; d] \subset \mathbb{R}^2$. Wir definieren das Doppelintegral von f .

Definition 3.32. Seien $a < b \in \mathbb{R}$, $c < d \in \mathbb{R}$. Sei $f : [a; b] \times [c; d] \rightarrow \mathbb{R}$. Das Doppelintegral

$$\int_a^b \int_c^d f(x, y) dy dx$$

existiert, wenn folgendes gilt

- i) Für jedes $x \in [a; b]$ ist die Funktion $y \rightarrow f(x, y)$ auf $[c; d]$ integrierbar.
- ii) Die Funktion $F : [a; b] \rightarrow \mathbb{R}$ definiert durch

$$F(x) = \int_c^d f(x, y) dy$$

ist integrierbar.

In diesem Fall setzt man

$$\int_a^b \int_c^d f(x, y) dy dx = \int_a^b F(x) dx = \int_a^b \left(\int_c^d f(x, y) dy \right) dx$$

Das Doppelintegral

$$\int_c^d \int_a^b f(x, y) dx dy$$

wird analog definiert.

Ist $f \in C([a; b] \times [c; d])$ stetig, so spielt es keine Rolle, ob wir zunächst über x oder über y integrieren.

Satz 3.33 (Fubini). *Seien $a, b, c, d \in \mathbb{R}$, mit $a < b$ und $c < d$. Sei $f \in C([a; b] \times [c; d])$. Dann existieren beide Doppelintegrale*

$$\int_a^b \int_c^d f(x, y) dy dx, \quad \int_c^d \int_a^b f(x, y) dx dy$$

und sind gleich.

Beweis: Aus der Stetigkeit von f folgt die Existenz von

$$F(x) = \int_c^d f(x, y) dy$$

für alle $x \in [a; b]$. Aus Satz 3.30 folgt, dass F stetig und deswegen auf $[a; b]$ integrierbar ist. Das zeigt die Existenz vom Doppelintegral

$$\int_a^b \int_c^d f(x, y) dy dx.$$

Analog zeigt man die Existenz vom Doppelintegral

$$\int_c^d \int_a^b f(x, y) dx dy.$$

Um zu zeigen, dass die zwei Integrale gleich sind, definieren wir $A : [a; b] \times [c; d] \rightarrow \mathbb{R}$ durch

$$A(x; y) = \int_c^y f(x, t) dt$$

Aus Analysis 1 folgt, dass A partiell nach y differenzierbar ist, mit

$$\frac{\partial A}{\partial y}(x, y) = f(x, y)$$

Wir behaupten nun, dass A stetig ist. Sei $(x_k; y_k)$ eine Folge in \mathbb{R}^2 mit $(x_k; y_k) \rightarrow (x; y)$. Dann gilt

$$|A(x_k; y_k) - A(x; y)| \leq |A(x_k; y_k) - A(x_k; y)| + |A(x_k; y) - A(x; y)|$$

Wir haben

$$|A(x_k; y_k) - A(x_k; y)| = \left| \int_{y_k}^y f(x_k; t) dt \right| \leq C |y_k - y| \rightarrow 0$$

als $k \rightarrow \infty$, weil $\sup\{f(x; y) : x \in [a; b], y \in [c; d]\} < \infty$ aus der Stetigkeit von f und aus der Kompaktheit von $[a; b] \times [c; d]$. Andererseits $|A(x_k, y) - A(x, y)| \rightarrow 0$, für $k \rightarrow \infty$, weil aus Satz 3.30 die Funktion $A(x, y)$ stetig in x ist, für beliebig $y \in [c; d]$. Wir setzen also

$$\varphi(y) = \int_a^b A(x, y) dx = \int_a^b \int_c^y f(x, t) dt dx$$

Insbesondere $\varphi(d) = \int_a^b \int_c^d f(x, y) dy dx$. Nach Satz 3.31 finden wir, dass φ differenzierbar ist, mit

$$\varphi'(y) = \int_a^b \frac{\partial A}{\partial y}(x, y) dx = \int_a^b f(x, y) dx$$

Das gibt

$$\varphi(d) = \varphi(c) + \int_c^d \varphi'(y) dy = \int_c^d \int_a^b f(x, y) dx dy$$

und zeigt die Behauptung. □

3.8 Konservative Vektorfelder

Wir führen den Begriff vom Vektorfeld ein.

Definition 3.34. *Ein Vektorfeld mit Definitionsbereich $U \subset \mathbb{R}^n$ ist eine Abbildung $K : U \rightarrow \mathbb{R}^n$. Ist $U \subset \mathbb{R}^n$ offen, so sagen wir die Abbildung $K : U \rightarrow \mathbb{R}^n$ ist ein C^k -Vektorfeld, falls $K \in C^k(U; \mathbb{R}^n)$.*

Wir haben schon oft Abbildungen betrachtet, die eine Teilmenge von \mathbb{R}^n auf \mathbb{R}^n abbilden (zB. Diffeomorphismen). Die Interpretation von Vektorfeldern ist aber anders. Ein Vektorfeld wird interpretiert als eine Abbildung, die zu jedem Punkt im Raum einen Vektor in \mathbb{R}^n zuordnet. Typische Beispiele von Vektorfeldern sind Kraftfelder. Eine elektrische Ladung im Punkt $x = 0$ erzeugt an der Stelle $x \in \mathbb{R}^3$ die Kraft

$$K(x) = -c \frac{x}{\|x\|^3}$$

für eine Konstante $c \in \mathbb{R}$. $K(x)$ ist ein Beispiel eines Vektorfelds. Ein anderes Beispiel von Vektorfeldern sind Geschwindigkeitsfelder. Die Strömung einer Flüssigkeit kann durch das Geschwindigkeitsfeld $v(x)$ beschrieben werden, das die momentane Geschwindigkeit der Flüssigkeit an der Stelle $x \in \mathbb{R}^n$ spezifiziert. Aus diesen Beispielen ist klar, dass Vektorfelder eine sehr wichtige Rolle in der Physik spielen (natürlich spielen in der Physik auch skalare Felder, wie zum Beispiel die Temperatur $T(x)$ als Funktion vom Ort, eine wichtige Rolle, und manchmal ist es auch nützlich, Matrix-wertige Felder zu betrachten; hier untersuchen wir aber nur Vektorfelder).

Feldlinien. Sei nun $K : U \rightarrow \mathbb{R}^n$ ein Vektorfeld. Eine parametrisierte Kurve $\gamma : I \rightarrow \mathbb{R}^n$ (wobei $I \subset \mathbb{R}$ ein Intervall ist) heisst eine Feldlinie vom Vektorfeld K , falls der Tangentialvektor $\gamma'(t)$ für alle $t \in I$ proportional zum Vektor $K(\gamma(t))$ ist. Man bemerke, der Begriff von Feldlinie ist von der Parametrisierung der Kurve γ unabhängig. In der Tat, falls $\psi : \tilde{I} \rightarrow I$ eine monotone differenzierbare Funktion ist, so gilt

$$\frac{d}{dt} \gamma(\psi(t)) = \gamma'(\psi(t)) \psi'(t)$$

und deswegen ist $(\gamma \circ \psi)'(t)$ immer proportional zu $\gamma'(\psi(t))$. Eine natürliche Parametrisierung einer Feldlinie ist also durch die Gleichung

$$\gamma'(t) = K(\gamma(t))$$

bestimmt. Diese Differentialgleichung für $\gamma(t)$ gibt uns die Möglichkeit, Feldlinien eines Kraftfelds zu finden. Ist K ein C^1 -Vektorfeld auf U und ist $x_0 \in U$, dann folgt aus der Theorie der gewöhnlichen Differentialgleichungen, siehe z.B. Satz 2.6, dass man immer mindestens ein Stück Feldlinie von K durch x_0 finden kann, und dass diese lokale Feldlinie eindeutig bestimmt ist.

Zentralfelder. Ein Vektorfeld $K : U \rightarrow \mathbb{R}^n$ heisst ein Zentralfeld, falls K die Form $K(x) = f(\|x\|)x$ hat, für eine Funktion $f(\|x\|)$ die nur von der Länge $\|x\|$ von x abhängt. Das elektrische Feld $K(x) = -\text{const}x/\|x\|^3$, erzeugt auf $\mathbb{R}^3 \setminus \{0\}$ aus einer Ladung im Ursprung, ist ein Beispiel eines Zentralfelds.

Gradientenfelder. Ein Vektorfeld $K : U \rightarrow \mathbb{R}^n$ definiert auf $U \subset \mathbb{R}^n$ heisst ein Gradientenfeld, falls eine reel-wertige Funktion $\varphi \in C^1(U)$ existiert, mit $K(x) = \nabla\varphi(x)$.

Die Länge einer parametrisierten Kurve. Sei $I = [a; b] \subset \mathbb{R}$ und $\gamma \in C^1(I; \mathbb{R}^n)$ eine differenzierbare parametrisierte Kurve in \mathbb{R}^n . Wir möchten die Länge von γ definieren. Ist γ eine Gerade, so ist die Länge von γ durch $\|\gamma(b) - \gamma(a)\|$ gegeben. Falls γ keine Gerade ist, so können wir versuchen, die Länge zu definieren, indem wir γ in viele kleine Teile zerlegen. Für $n \in \mathbb{N}$ finden wir $a = t_0 < t_1 < t_2 < \dots < t_n = b$. Eine erste Näherung für die Länge von γ ist aus

$$\sum_{j=1}^n \|\gamma(t_j) - \gamma(t_{j-1})\| \simeq \sum_{j=1}^n (t_j - t_{j-1}) \|\gamma'(t_{j-1})\|$$

gegeben. Wir können nun die Länge von γ berechnen, indem wir den Limes dieses Ausdrucks für $n \rightarrow \infty$ betrachten (angenommen, die Folge der Teilungen ist so, dass $\sup_j |t_j - t_{j-1}| \rightarrow 0$). Wir definieren also die Länge der parametrisierten differenzierbaren Kurve γ durch

$$L(\gamma) = \int_a^b dt \|\gamma'(t)\| \quad (60)$$

Aus der Annahme $\gamma \in C^1([a; b]; \mathbb{R}^n)$ folgt, dass $\|\gamma'(t)\|$ stetig von t abhängt. Deswegen ist die rechte Seite von (60) wohldefiniert und endlich.

Man bemerke auch, dass die Definition (60) unabhängig von der Parametrisierung der Kurve ist. Sei nämlich $\psi : [c; d] \rightarrow [a; b]$ eine monoton steigende injektive und differenzierbare Funktion. Wir definieren die parametrisierte Kurve $\tilde{\gamma} : [c; d] \rightarrow \mathbb{R}^n$ durch $\tilde{\gamma}(t) = \gamma(\psi(t))$. Dann gilt

$$\tilde{\gamma}'(t) = \gamma'(\psi(t))\psi'(t)$$

Deswegen

$$L(\tilde{\gamma}) = \int_c^d dt \|\tilde{\gamma}'(t)\| = \int_c^d dt \psi'(t) \|\gamma'(\psi(t))\|$$

wobei wir benutzt haben, dass $\psi'(t) \geq 0$ (aus der Monotonie). Mit der Variablentransformation $s = \psi(t)$ finden wir $L(\tilde{\gamma}) = L(\gamma)$. D.h., wie behauptet, die Länge ist von der Parametrisierung unabhängig.

Allgemeiner kann man die Länge einer stückweise stetig differenzierbaren parametrisierten Kurve definieren.

Definition 3.35. Eine parametrisierte Kurve $\gamma : [a; b] \rightarrow \mathbb{R}^n$ heisst *stückweise stetig differenzierbar*, falls sie stetig auf $[a; b]$ ist, und falls eine endliche Teilung $a = t_0 < t_1 < \dots < t_{n-1} < t_n = b$ mit der Eigenschaft, dass $\gamma \in C^1([t_{j-1}; t_j]; \mathbb{R}^n)$ für alle $j = 1, \dots, n$ existiert (erinnere, dass $\gamma \in C^1([t_{j-1}; t_j]; \mathbb{R}^n)$ genau dann, wenn $\gamma \in C^1((t_{j-1}; t_j); \mathbb{R}^n)$ und γ und γ' können auf dem abgeschlossenen Intervall $[t_{j-1}; t_j]$ stetig fortgesetzt werden).

Für eine stückweise stetig differenzierbare Kurve $\gamma : [a; b] \rightarrow \mathbb{R}^n$, können wir dann die Länge durch

$$L(\gamma) = \sum_{j=1}^n \int_{t_{j-1}}^{t_j} dt \|\gamma'(t)\| \quad (61)$$

definieren, wobei die Teilung $a = t_0 < t_1 < \dots < t_n = b$ so gewählt wird, dass $\gamma \in C^1([t_{j-1}; t_j]; \mathbb{R}^n)$ für alle $j = 1, \dots, n$.

Linienintegrale. Sei nun $\gamma : [a; b] \rightarrow \mathbb{R}^n$ eine parametrisierte Kurve und $K : U \rightarrow \mathbb{R}^n$ ein Vektorfeld, mit $\gamma(I) \subset U$. Wir interpretieren γ als die Bahn eines Teilchen und $K(x)$ als die auf dem Teilchen im Punkt x wirkende Kraft. Ein wichtiger Begriff in der Physik ist die Arbeit, die das Teilchen leisten muss, um sich durch das Kraftfeld zu bewegen (oder die Arbeit, die das Kraftfeld auf dem Teilchen leistet). Ist γ eine Gerade und $K(x) = K$ konstant auf der Geraden, so ist die Arbeit aus $K \cdot (\gamma(b) - \gamma(a))$ gegeben. Allgemeiner können wir die Arbeit berechnen, indem wir das Intervall $[a; b]$ in kleine Teilintervalle zerlegen. Seien $a = t_0 < t_1 < \dots < t_n = b$. Dann können wir die Arbeit durch

$$\sum_{j=1}^n K(\gamma(t_{j-1})) \cdot (\gamma(t_j) - \gamma(t_{j-1})) \simeq \sum_{j=1}^n K(\gamma(t_{j-1})) \gamma'(t_{j-1})(t_j - t_{j-1})$$

approximieren. Nehmen wir das Limes $n \rightarrow \infty$ (mit $\sup_j(t_j - t_{j-1}) \rightarrow 0$), so konvergiert die linke Seite (angenommen z.B. das Vektorfeld K ist stetig) zum Integral

$$\int_a^b K(\gamma(t)) \cdot \gamma'(t) dt$$

Das motiviert die folgende Definition.

Definition 3.36. Sei $U \subset \mathbb{R}^n$ offen, $K : U \rightarrow \mathbb{R}^n$ ein stetiges Vektorfeld. Sei $\gamma : [a; b] \rightarrow U$ eine stückweise stetig differenzierbare Kurve in U . Wir definieren dann das *Linienintegral (oder Wegintegral) von K entlang γ* durch

$$\int_{\gamma} K \cdot dx := \sum_{j=1}^n \int_{t_{j-1}}^{t_j} K(\gamma(t)) \cdot \gamma'(t) dt$$

wobei die Teilung $a = t_0 < t_1 < \dots < t_n = b$ so gewählt wurde, dass $\gamma \in C^1([t_{j-1}; t_j]; \mathbb{R}^n)$ für alle $j = 1, \dots, n$.

In der folgenden Proposition sammeln wir einige wichtige Eigenschaften von Linienintegralen.

Proposition 3.37. Sei $U \subset \mathbb{R}^n$ offen, K ein stetiges Vektorfeld in U , $\gamma : [a; b] \rightarrow U$ eine stückweise stetig differenzierbare Kurve in U .

i) Sei $M = \sup\{\|K(x)\| : x \in \gamma([a; b])\}$ (bemerke, dass $M < \infty$ wegen Stetigkeit von K und Kompaktheit von $\gamma([a; b])$). Dann gilt

$$\left| \int_{\gamma} K \cdot dx \right| \leq ML(\gamma)$$

wobei $L(\gamma)$ die Länge von γ ist, wie in (61) definiert.

ii) Sei $\psi : [c; d] \rightarrow [a; b]$ monoton steigend, mit $\psi(c) = a$ und $\psi(d) = b$, und stückweise stetig differenzierbar (ψ ist eine Parametertransformation). Wir definieren $\tilde{\gamma} : [c; d] \rightarrow U$ durch $\tilde{\gamma}(t) = \gamma(\psi(t))$. Dann gilt

$$\int_{\tilde{\gamma}} K \cdot dx = \int_{\gamma} K \cdot dx$$

D.h. der Wert des Linienintegrals ist von der Parametrisierung der Kurve unabhängig.

Beweis: i) Sei $a = t_0 < t_1 < \dots < t_n = b$ eine Teilung, so dass $\gamma \in C^1([t_{j-1}, t_j]; U)$ für alle $j = 1, \dots, n$. Dann gilt

$$\begin{aligned} \left| \int_{t_{j-1}}^{t_j} K(\gamma(t)) \cdot \gamma'(t) dt \right| &\leq \int_{t_{j-1}}^{t_j} |K(\gamma(t)) \cdot \gamma'(t)| dt \\ &\leq \int_{t_{j-1}}^{t_j} \|K(\gamma(t))\| \|\gamma'(t)\| dt \leq M \int_{t_{j-1}}^{t_j} \|\gamma'(t)\| dt \end{aligned}$$

Also

$$\left| \int_{\gamma} K \cdot dx \right| = \left| \sum_{j=1}^n \int_{t_{j-1}}^{t_j} K(\gamma(t)) \cdot \gamma'(t) dt \right| \leq M \sum_{j=1}^n \int_{t_{j-1}}^{t_j} \|\gamma'(t)\| dt = ML(\gamma)$$

ii) Sei $c = t_0 < t_1 < \dots < t_n = d$ eine Teilung von $[c; d]$ mit der Eigenschaft, dass $\psi \in C^1([t_{j-1}, t_j])$ und $\gamma \in C^1([\psi(t_{j-1}), \psi(t_j)]; U)$ für alle $j = 1, \dots, n$. Dann gilt

$$\begin{aligned} \int_{t_{j-1}}^{t_j} K(\tilde{\gamma}(t)) \cdot \tilde{\gamma}'(t) dt &= \int_{t_{j-1}}^{t_j} K(\gamma(\psi(t))) \cdot \gamma'(\psi(t)) \psi'(t) dt \\ &= \int_{\psi(t_{j-1})}^{\psi(t_j)} K(\gamma(s)) \cdot \gamma'(s) ds \end{aligned}$$

mit der Variablentransformation $s = \psi(t)$. Also, da ψ monoton wachsend ist, finden wir

$$\begin{aligned} \int_{\tilde{\gamma}} K \cdot dx &= \sum_{j=1}^n \int_{t_{j-1}}^{t_j} K(\tilde{\gamma}(t)) \tilde{\gamma}'(t) dt \\ &= \sum_{j=1}^n \int_{\psi(t_{j-1})}^{\psi(t_j)} K(\gamma(s)) \cdot \gamma'(s) ds = \int_{\gamma} K \cdot dx \end{aligned}$$

□

Operationen mit Kurven. Für eine stückweise stetig differenzierbare Kurve $\gamma : [a; b] \rightarrow \mathbb{R}^n$, kann man die stückweise stetig differenzierbare Kurve $-\gamma : [a; b] \rightarrow \mathbb{R}^n$ durch $-\gamma(t) = \gamma(b + a - t)$ definieren. Es ist einfach zu sehen, dass $-\gamma$ die selbe Kurve wie γ beschreibt, aber in umgekehrte Richtung parametrisiert.

Sind $\gamma_1 : [a; b] \rightarrow \mathbb{R}^n$ und $\gamma_2 : [c; d] \rightarrow \mathbb{R}^n$ zwei stückweise stetig differenzierbare Kurven, so dass $\gamma_1(b) = \gamma_2(c)$, so kann man die Kurve $\gamma_3 = \gamma_1 + \gamma_2 : [a; b + d - c]$ durch

$$\gamma_3(t) = \begin{cases} \gamma_1(t) & \text{falls } t \in [a; b] \\ \gamma_2(t + c - b) & \text{falls } t \in [b; b + d - c] \end{cases}$$

definieren. Anschaulich, γ_3 ist die "Vereinigung" der zwei Kurven γ_1 und γ_2 . Es ist einfach zu sehen, dass $L(-\gamma) = L(\gamma)$, und $L(\gamma_1 + \gamma_2) = L(\gamma_1) + L(\gamma_2)$. Für Linienintegrale finden wir

$$\int_{-\gamma} K \cdot dx = - \int_{\gamma} K \cdot dx \quad (62)$$

und

$$\int_{\gamma_1 + \gamma_2} K \cdot dx = \int_{\gamma_1} K \cdot dx + \int_{\gamma_2} K \cdot dx. \quad (63)$$

Um Gleichung 62 zu zeigen, sei $\gamma : [a; b] \rightarrow \mathbb{R}^n$ und $a = t_0 < t_1 < \dots < t_n = b$ eine Teilung mit der Eigenschaft, dass $\gamma \in C^1([t_{j-1}; t_j]; U)$ für alle $j = 1, \dots, n$. Wir setzen $\tilde{t}_j = a + b - t_{n-j}$. Dann ist $a = \tilde{t}_0 < \tilde{t}_1 < \dots < \tilde{t}_n = b$ eine Teilung mit der Eigenschaft, dass $-\gamma \in C^1([\tilde{t}_{j-1}; \tilde{t}_j]; U)$ für alle $j = 1, \dots, n$. Erinnerung hier, dass $-\gamma(t) = \gamma(a + b - t)$. Wir finden

$$\begin{aligned} \int_{-\gamma} K \cdot dx &= \sum_{j=1}^n \int_{\tilde{t}_{j-1}}^{\tilde{t}_j} K(-\gamma(t)) \cdot (-\gamma)'(t) dt \\ &= - \sum_{j=1}^n \int_{a+b-t_{n-j+1}}^{a+b-t_{n-j}} K(\gamma(a+b-t)) \cdot \gamma'(a+b-t) dt \\ &= \sum_{j=1}^n \int_{t_{n-j+1}}^{t_{n-j}} K(\gamma(s)) \cdot \gamma'(s) ds \\ &= - \sum_{j=1}^n \int_{t_{n-j}}^{t_{n-j+1}} K(\gamma(s)) \cdot \gamma'(s) ds = - \int_{\gamma} K \cdot dx \end{aligned}$$

Die Gleichung (63) kann ähnlich bewiesen werden.

Konservative Vektorfelder. Falls das Vektorfeld K ein Gradientenfeld ist, dann ist die Berechnung seiner Linienintegrale besonders einfach.

Satz 3.38. Sei $U \subset \mathbb{R}^n$ offen, $\varphi \in C^1(U)$ und $K = \nabla\varphi$ (dann ist K ein stetiges Vektorfeld auf U). Sei $\gamma \in C^1([a; b]; U)$ eine stetig differenzierbare Kurve auf U . Dann gilt

$$\int_{\gamma} K \cdot dx = \varphi(\gamma(a)) - \varphi(\gamma(b))$$

$\gamma(a)$ heisst der Anfangspunkt der Kurve γ und $\gamma(b)$ der Endpunkt.

Beweis: Wir bemerken, dass

$$\frac{d}{dt}\varphi(\gamma(t)) = \nabla\varphi(\gamma(t)) \cdot \gamma'(t)$$

Deswegen finden wir

$$\begin{aligned} \int_{\gamma} K \cdot dx &= \int_a^b \nabla\varphi(\gamma(t)) \cdot \gamma'(t) dt \\ &= \int_a^b \frac{d}{dt} \varphi(\gamma(t)) dt = \varphi(\gamma(a)) - \varphi(\gamma(b)) \end{aligned}$$

□

Bemerkung: Die Aussage von Satz 3.38 gilt auch, falls die Kurve γ stückweise stetig differenzierbar ist. In diesem Fall finden wir eine Teilung $a = t_0 < t_1 < \dots < t_n = b$ mit der Eigenschaft, dass $\gamma \in C^1([t_{j-1}; t_j]; U)$ für alle $j = 1, \dots, n$. Dann gilt, ähnlich wie im Beweis des Satzes,

$$\int_{t_{j-1}}^{t_j} \nabla\varphi(\gamma(t)) \cdot \gamma'(t) dt = \varphi(\gamma(t_j)) - \varphi(\gamma(t_{j-1}))$$

für alle $j = 1, \dots, n$. Damit

$$\int_{\gamma} K \cdot dx = \sum_{j=1}^n (\varphi(\gamma(t_j)) - \varphi(\gamma(t_{j-1}))) = \varphi(\gamma(b)) - \varphi(\gamma(a))$$

Bemerkung: Es folgt aus dem Satz, dass Linienintegrale von Gradientenfeldern entlang einer Kurve γ nur vom Anfangspunkt und Endpunkt von γ abhängen, nicht von dem Weg dazwischen.

Bemerkung: Eine parametrisierte Kurve $\gamma : [a; b] \rightarrow \mathbb{R}^n$ heisst geschlossen, wenn $\gamma(a) = \gamma(b)$, d.h. falls Anfangspunkt und Endpunkt der Kurve übereinstimmen. Es folgt aus Satz 3.38, dass das Linienintegral von einem Gradientenfeld entlang einer geschlossenen Kurve verschwindet.

Definition 3.39. Sei $U \subset \mathbb{R}^n$. Ein Vektorfeld $K : U \rightarrow \mathbb{R}^n$ heisst konservativ (oder manchmal exakt), falls das Linienintegral

$$\int_{\gamma} K \cdot dx$$

nur vom Anfangs- und Endpunkt von γ abhängt, für jede Kurve γ in U . Mit anderen Worten, K ist konservativ, falls für alle zwei stückweise stetig differenzierbaren Kurven γ_1, γ_2 mit übereinstimmenden Anfangs- und Endpunkten, gilt

$$\int_{\gamma_1} K \cdot dx = \int_{\gamma_2} K \cdot dx$$

Es folgt aus Satz 3.38, dass Gradientenfelder konservativ sind. In der Physik spielen konservative Kraftfelder eine besonders wichtige Rolle. Ist K konservativ, dann ist die Arbeit, die ein Teilchen leistet, wenn es sich auf einer Bahn γ bewegt, nur vom Anfangs- und Endpunkt der Bahn abhängig. Das impliziert, wie wir bald sehen werden, dass man für konservative Kraftfelder ein Potential einführen kann, so dass die geleistete Arbeit einfach die Differenz vom Potential im End- und im Anfangspunkt ist. D.h. man kann ein Potential einführen, so dass Energieerhaltung gilt.

Proposition 3.40. *Sei $U \subset \mathbb{R}^n$ offen und $K : U \rightarrow \mathbb{R}^n$ ein stetiges Vektorfeld. Dann ist K genau dann konservativ, wenn*

$$\int_{\gamma} K \cdot dx = 0$$

für alle geschlossenen stückweise stetig differenzierbaren Kurve γ in U .

Beweis: Sei zunächst K konservativ und $\gamma : [a; b] \rightarrow U$ eine geschlossene stückweise stetig differenzierbare Kurve mit $\gamma(a) = \gamma(b) =: x_0$. Es bezeichne $\tilde{\gamma} : [a; b] \rightarrow \mathbb{R}^n$ die konstante Kurve $\tilde{\gamma}(t) = x_0$ für alle $t \in [a; b]$. Weil γ und $\tilde{\gamma}$ die selben Anfangs- und Endpunkte haben, finden wir (siehe Prop. 3.37)

$$\int_{\gamma} K \cdot dx = \int_{\tilde{\gamma}} K \cdot dx = 0.$$

Nehmen wir nun an, dass

$$\int_{\gamma} K \cdot dx = 0$$

für alle geschlossenen stückweise stetig differenzierbaren Kurven γ auf U . Seien γ_1 und γ_2 zwei beliebige stückweise stetig differenzierbare Kurven auf U , mit übereinstimmenden Anfangs- und Endpunkten. Wie oben bezeichnen wir mit $-\gamma_2$ die Kurve γ_2 , mit umgekehrter Richtung. Der Anfangspunkt von $-\gamma_2$ ist dann der Endpunkt von γ_1 und analog der Anfangspunkt von γ_1 ist der Endpunkt von $-\gamma_2$. Wir definieren auch die Kurve $\gamma = \gamma_1 + (-\gamma_2)$, die parametrisierte Kurve gegeben aus der "Vereinigung" von γ_1 und $-\gamma_2$. Genauer gesagt, falls $\gamma_1 : [a; b] \rightarrow U$ und $-\gamma_2 : [c; d] \rightarrow U$, so definieren wir $\gamma : [a; b + d - c] \rightarrow U$ durch $\gamma(t) = \gamma_1(t)$ falls $t \in [a; b]$ und $\gamma(t) = -\gamma_2(t - b + c)$ falls $t \in [b; b + d - c]$. Die Kurve γ ist dann offenbar eine geschlossene stückweise stetig differenzierbare Kurve, und deswegen

$$0 = \int_{\gamma} K \cdot dx = \int_{\gamma_1} K \cdot dx + \int_{-\gamma_2} K \cdot dx = \int_{\gamma_1} K \cdot dx - \int_{\gamma_2} K \cdot dx$$

Das zeigt, dass

$$\int_{\gamma_1} K \cdot dx = \int_{\gamma_2} K \cdot dx.$$

Also ist K konservativ. □

Wir haben schon bemerkt, dass jedes Gradientenfeld konservativ ist. Wir zeigen nun die Umkehrung dieser Aussage: Jedes konservative Vektorfeld ist ein Gradientenfeld. Dazu werden wir das folgende Hilfslemma anwenden.

Lemma 3.41. *Jede offene Teilmenge $U \subset \mathbb{R}^n$ lässt sich als Vereinigung einer disjunkten Familie offener zusammenhängender Mengen darstellen (die Vereinigung braucht nicht endlich zu sein).*

Bemerkung. Diese Zerlegung ist eigentlich eindeutig; die offenen zusammenhängenden Teilmengen werden als Zusammenhangskomponenten bezeichnet.

Beweis: Wir definieren eine Relation zwischen Punkten in U . Für $x, y \in U$ schreiben wir $x \sim y$ genau dann, wenn x und y sich durch einen Streckenzug in U verbinden lassen. Offenbar definiert \sim eine Äquivalenzrelation. Die Äquivalenzklassen sind offen. Sei nämlich $x \in U$ beliebig. Wir zeigen die Äquivalenzklasse $[x]$ ist offen. Dazu finden wir $r > 0$, so dass die offene Kugel $B_r(x)$ von Radius r um x in U enthalten ist. Dann ist $x \sim y$, für alle $y \in B_r(x)$. Damit ist $B_r(x) \subset [x]$, und $[x]$ ist offen. Die Äquivalenzklassen sind offenbar zusammenhängend und paarweise disjunkt. \square

Wir können nun zeigen, dass jedes konservative Feld ein Gradientenfeld ist.

Satz 3.42. *Sei $U \subset \mathbb{R}^n$ offen, K ein konservatives stetiges Vektorfeld auf U . Dann existiert $\varphi : U \rightarrow \mathbb{R}$ stetig differenzierbar, mit $K = \nabla\varphi$.*

Beweis: Wir betrachten zunächst den Fall, dass U zusammenhängend ist. Wir wählen $x_0 \in U$ fest. Für ein beliebiges $x \in U$ finden wir eine stückweise stetig differenzierbare Kurve γ_x in U mit Anfangspunkt x_0 und Endpunkt x . Wir setzen

$$\varphi(x) = \int_{\gamma_x} K \cdot dx$$

Da K konservativ ist, ist die Definition unabhängig von der Wahl der Kurve γ_x , natürlich unter der Annahme, dass der Endpunkt von γ_x gerade x ist. Wir bemerken, dass $\varphi(x_0) = 0$. Wir behaupten, dass $\nabla\varphi(x) = K(x)$. Das würde den Satz für den Fall U zusammenhängend beweisen, weil K aus Annahme stetig ist. Um die Behauptung zu zeigen, bemerken wir, dass

$$\frac{\partial\varphi}{\partial x_i}(x) = \frac{d}{dt}\varphi(x + te_i)|_{t=0}$$

und dass

$$\begin{aligned} \varphi(x + te_i) &= \int_{\gamma_{x+te_i}} K \cdot dx = \int_{\gamma_x + [x; x+te_i]} K \cdot dx \\ &= \int_{\gamma_x} K \cdot dx + \int_{[x; x+te_i]} K \cdot dx \end{aligned}$$

wobei $[x; x + te_i]$ das Segment zwischen x und $x + te_i$ bezeichnet. Hier wählen wir t so klein, dass das Segment $[x; x + te_i]$ in U enthalten ist (das ist natürlich möglich, weil U offen ist). Wir parametrisieren das Segment $[x; x + te_i]$ durch die Kurve $\gamma : [0; t] \rightarrow U$, definiert durch $\gamma(s) = x + se_i$. Dann ist $\gamma'(s) = e_i$, und damit

$$\int_{[x; x+te_i]} K \cdot dx = \int_0^t K(x + se_i) \cdot e_i ds = \int_0^t K_i(x + se_i) ds$$

wobei K_i die i -te Komponente vom Vektorfeld K bezeichnet. Nach dem Hauptsatz der Integralrechnung, finden wir

$$\frac{d}{dt}\varphi(x + te_i)|_{t=0} = \frac{d}{dt} \int_0^t K_i(x + se_i) ds \Big|_{t=0} = K_i(x + te_i)|_{t=0} = K_i(x)$$

Damit ist $\nabla\varphi(x) = K(x)$, wie behauptet.

Sei nun U nicht zusammenhängend. Aus Lemma 3.41 finden wir eine Familie U_i von disjunkten zusammenhängenden offenen Mengen in \mathbb{R}^n , mit $U = \cup_i U_i$. Für jede i können wir dann wie oben eine stetig differenzierbare Funktion $\varphi_i : U_i \rightarrow \mathbb{R}$ konstruieren, mit $\nabla\varphi_i(x) = K(x)$ für alle $x \in U_i$. Da die Teilmengen disjunkt sind, können wir $\varphi : U \rightarrow \mathbb{R}$ durch $\varphi(x) := \varphi_i(x)$ für alle $x \in U_i$ definieren. Dann ist φ wohldefiniert, stetig differenzierbar, mit $\nabla\varphi(x) = K(x)$ für alle $x \in U$. \square

Satz 3.42 zeigt, zusammen mit Satz 3.38, dass für jedes konservative Vektorfeld $K : U \rightarrow \mathbb{R}^n$, eine Potentialfunktion $\varphi : U \rightarrow \mathbb{R}$ gefunden werden kann, mit $K = \nabla\varphi$ und deswegen, mit

$$\int_{\gamma} K \cdot dx = \varphi(\text{Endpunkt}) - \varphi(\text{Anfangspunkt})$$

Das bedeutet, für jedes konservative Vektorfeld kann man ein Potential einführen, so dass geleistete Arbeit = Unterschied im Potential (die Energie ist erhalten). Bemerke, dass die Potentialfunktion immer nur bis auf eine additive Konstante bestimmt ist.

Charakterisierung von konservativen Vektorfeldern. Wir haben bis jetzt bewiesen, dass jedes konservative Vektorfeld ein Gradientenfeld ist. Die nächste natürliche Frage ist, wie können wir entscheiden, ob ein gegebenes Vektorfeld ein Gradientenfeld ist oder nicht. Es ist sehr einfach, notwendige Bedingungen zu finden. Ist $K = \nabla\varphi$, so muss gelten

$$\frac{\partial K_i}{\partial x_j}(x) = \frac{\partial^2 \varphi}{\partial x_j \partial x_i}(x) = \frac{\partial^2 \varphi}{\partial x_i \partial x_j}(x) = \frac{\partial K_j}{\partial x_i}(x) \quad (64)$$

für alle $1 \leq i < j \leq n$ (die Bedingungen sind symmetrisch bzg. Änderung $i \rightarrow j, j \rightarrow i$). Hier bezeichnet K_i die i -te Komponente vom Vektorfeld K . Ein Vektorfeld K mit der Eigenschaft (64) heisst rotationfrei (manchmal geschlossen). Im nächsten Satz beweisen wir, dass diese Bedingungen nicht nur notwendig, sondern auch hinreichend sind, falls das Vektorfeld auf einer konvexen offenen Menge definiert und differenzierbar ist.

Satz 3.43. Sei $U \subset \mathbb{R}^n$ offen und konvex (d.h. es gelte $\lambda x + (1 - \lambda)y \in U$, für alle $x, y \in U$ und $\lambda \in [0; 1]$). Sei $K \in C^1(U; \mathbb{R}^n)$. Dann existiert $\varphi \in C^1(U)$ mit $K = \nabla\varphi$ genau dann, wenn

$$\frac{\partial K_i}{\partial x_j}(x) = \frac{\partial K_j}{\partial x_i}(x) \quad (65)$$

für alle $1 \leq i < j \leq n$ und alle $x \in U$. Das impliziert, K ist genau dann konservativ, wenn (65) erfüllt ist.

Beweis: O.B.d.A. nehmen wir an, $0 \in U$. Für jede $x \in U$ ist dann das Segment $[0; x] = \{tx : t \in [0; 1]\}$ in U enthalten (weil U konvex ist). Deswegen können wir eine Funktion $\varphi : U \rightarrow \mathbb{R}$ durch

$$\varphi(x) = \int_0^1 K(tx) \cdot x dt$$

definieren. Da $K \in C^1(U; \mathbb{R}^n)$, folgt aus Satz 3.31, dass $\varphi \in C^1(U; \mathbb{R})$, mit

$$\frac{\partial \varphi}{\partial x_j}(x) = \int_0^1 \frac{\partial}{\partial x_j} \left(\sum_{i=1}^n K_i(tx) x_i \right) dt = \int_0^1 \left(K_j(tx) + \sum_{i=1}^n \frac{\partial K_i}{\partial x_j} tx_i \right) dt \quad (66)$$

Andererseits, bemerken wir, dass

$$\frac{d}{dt}(tK_j(tx)) = K_j(tx) + \sum_{i=1}^n \frac{\partial K_j}{\partial x_i}(tx) tx_i = K_j(tx) + \sum_{i=1}^n \frac{\partial K_i}{\partial x_j}(tx) tx_i$$

In der letzten Gleichung haben wir die Bedingung $\partial K_j / \partial x_i = \partial K_i / \partial x_j$ benutzt. Vergleich mit (66) gibt

$$\frac{\partial \varphi}{\partial x_j}(x) = \int_0^1 \frac{d}{dt} (tK_j(tx)) dt = K_j(x).$$

□

Beispiel: Sei $K(x, y) = (y^2, 2xy + y^2)$. Wir suchen eine Funktion $\varphi : \mathbb{R}^2 \rightarrow \mathbb{R}$ mit $\nabla \varphi = K$. Wir bemerken zunächst, dass

$$\frac{\partial K_1}{\partial y} = 2y = \frac{\partial K_2}{\partial x}$$

Das heisst, es existiert eine Potentialfunktion φ mit $\nabla \varphi = K$. Sie muss erfüllen, dass

$$\frac{\partial \varphi}{\partial x} = y^2, \quad \text{und} \quad \frac{\partial \varphi}{\partial y} = 2xy + y^2$$

Die erste Gleichung impliziert, dass $\varphi(x, y) - xy^2$ unabhängig von x sein soll. Damit können wir $\varphi(x, y) = xy^2 + \psi(y)$ schreiben, für eine geeignete Funktion ψ . Durch Einsetzen in die Gleichung für $\partial \varphi / \partial y$ finden wir

$$2xy + \psi'(y) = 2xy + y^2$$

Das ergibt $\psi'(y) = y^2$ und damit $\psi(y) = y^3/3 + c$ für eine Konstante $c \in \mathbb{R}$. Potentialfunktionen von K haben die Form

$$\varphi(x, y) = xy^2 + \frac{y^3}{3} + c.$$

Man bemerke, dass die Bedingung (65) nicht auf beliebigen Gebieten hinreichend ist (sie ist immer notwendig). Das zeigen wir mit dem folgenden Beispiel.

Beispiel: Sei

$$K(x_1, x_2) = \left(\frac{-x_2}{x_1^2 + x_2^2}; \frac{x_1}{x_1^2 + x_2^2} \right)$$

definiert auf der offenen Menge $\mathbb{R}^2 \setminus \{0\}$. Dann gilt

$$\frac{\partial K_1}{\partial x_2}(x) = \frac{-x_1^2 + x_2^2}{(x_1^2 + x_2^2)^2} = \frac{\partial K_2}{\partial x_1}(x)$$

für alle $x \in \mathbb{R}^2 \setminus \{0\}$. Sei aber $\gamma : [0, 2\pi] \rightarrow \mathbb{R}^2$ definiert durch $\gamma(t) = (\cos t, \sin t)$ der Einheitskreis (eine geschlossene Kurve auf $\mathbb{R}^2 \setminus \{0\}$). Dann ist

$$\int_{\gamma} K \cdot dx = \int_0^{2\pi} K(\cos t, \sin t) \cdot (-\sin t, \cos t) dt = 2\pi \neq 0$$

Also, K ist sicher nicht konservativ.

Tatsächlich gilt die Äquivalenz

$$K \text{ Gradientfeld} \iff \frac{\partial K_i}{\partial x_j} = \frac{\partial K_j}{\partial x_i}$$

nicht nur auf konvexen, sondern allgemeiner auf sogenannten einfach zusammenhängenden Gebieten. Ein Gebiet $G \subset \mathbb{R}^n$ heisst einfach zusammenhängend, wenn jede geschlossene Kurve in G stetig zu einem Punkt deformiert werden kann (wir verzichten hier auf die genaue Definition dieses Begriffes). Das Gebiet $\mathbb{R}^2 \setminus \{0\}$ ist nicht einfach zusammenhängend, weil jede Kurve um den Ursprung nicht stetig innerhalb $\mathbb{R}^2 \setminus \{0\}$ zu einem Punkt deformiert werden kann (bemerke dagegen, dass $\mathbb{R}^3 \setminus \{0\}$ einfach zusammenhängend ist).

3.9 Holomorphe Funktionen

In diesem Abschnitt betrachten wir Funktionen einer komplexen Variablen, mit Werten in \mathbb{C} . Bemerke, dass diese Funktionen das Thema der Vorlesung ‘‘Einführung in der komplexen Analysis’’ sind (oft wird diese Vorlesung auch als ‘‘Funktionentheorie’’ genannt); hier geben wir nur eine kurze Einführung in dieses wichtige Gebiet der Analysis.

Wir wissen, dass \mathbb{C} mit \mathbb{R}^2 identifiziert werden kann. Eine Funktion $f : \mathbb{C} \rightarrow \mathbb{C}$ kann deswegen mit einer Funktion $\tilde{f} : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ identifiziert werden, die durch $\tilde{f}(x, y) = (\operatorname{Re} f(x + iy), \operatorname{Im} f(x + iy))$ definiert wird. Wir werden aber sehen, der Begriff von Differenzierbarkeit einer Funktion auf \mathbb{C} mit Werten auf \mathbb{C} , ist nicht mit dem Begriff von Differenzierbarkeit von Funktionen auf \mathbb{R}^2 , mit Werten auf \mathbb{R}^2 äquivalent.

Definition 3.44. Sei $\Omega \subset \mathbb{C}$ offen. Eine Funktion $f : \Omega \rightarrow \mathbb{C}$ heisst komplex differenzierbar an der Stelle $z_0 \in \Omega$, wenn der Grenzwert

$$f'(z_0) = \lim_{z \rightarrow z_0} \frac{f(z) - f(z_0)}{z - z_0}$$

existiert. In diesem Fall heisst die Zahl $f'(z_0) \in \mathbb{C}$ die Ableitung von f an der Stelle z_0 . Die Funktion f heisst auf Ω komplex differenzierbar, falls f an der Stelle z differenzierbar ist, für alle $z \in \Omega$. Die Funktion f heisst holomorph in Ω , wenn sie auf Ω differenzierbar ist, und falls die Ableitung $f'(z)$ auf Ω stetig ist. Die Menge der holomorphen Funktionen auf Ω wird mit $\mathcal{H}(\Omega)$ bezeichnet.

Bemerkung: Die Funktion $f : \Omega \rightarrow \mathbb{C}$ ist genau dann an der Stelle $z_0 \in \Omega$ komplex differenzierbar, wenn ein $a \in \mathbb{C}$ existiert, so dass

$$\lim_{z \rightarrow z_0} \frac{f(z) - f(z_0) - a(z - z_0)}{|z - z_0|} = 0$$

Mit anderen Worten, f ist an der Stelle z_0 komplex differenzierbar, falls

$$f(z_0 + h) - f(z_0) - ah = o(|h|)$$

für $h \rightarrow 0$ (hier ist $h \in \mathbb{C}$). Die Abbildung $L : \mathbb{C} \rightarrow \mathbb{C}$ definiert durch $L(h) = ah$ ist ein Beispiel einer komplex linearen Abbildung auf \mathbb{C} (weil $L(h_1 + h_2) = L(h_1) + L(h_2)$ für alle $h_1, h_2 \in \mathbb{C}$ und $L(\alpha h) = \alpha L(h)$ für alle $\alpha, h \in \mathbb{C}$). Es ist eigentlich einfach zu sehen, dass jede komplex lineare Abbildung auf \mathbb{C} die Form $L(h) = ah$ für ein $a \in \mathbb{C}$ hat. D.h. die Funktion f ist genau dann differenzierbar, wenn eine komplex lineare Abbildung $L : \mathbb{C} \rightarrow \mathbb{C}$ existiert, mit der Eigenschaft, dass

$$f(z_0 + h) - f(z_0) - L(h) = o(|h|)$$

für $h \rightarrow 0$. Das erklärt den Unterschied zum Begriff von Differenzierbarkeit in \mathbb{R}^2 . Versehen wir f als die Abbildung $\tilde{f}(x, y) = (\operatorname{Re} f(x + iy), \operatorname{Im} f(x + iy))$, definiert auf einer offenen Teilmenge von \mathbb{R}^2 mit Werten in \mathbb{R}^2 , so ist \tilde{f} an der Stelle $z_0 = (x_0, y_0)$ genau dann differenzierbar, wenn eine reell lineare Abbildung $L : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ existiert, so dass

$$\tilde{f}(z_0 + h) - \tilde{f}(z_0) - L(h) = o(\|h\|)$$

für $h \rightarrow 0$ (hier ist $h \in \mathbb{R}^2$). Während jede komplex lineare Abbildung reell linear ist, ist nicht jede reell lineare Abbildung komplex linear. Deswegen impliziert die komplexe Differenzierbarkeit einer Funktion $f : \Omega \rightarrow \mathbb{C}$ immer die reelle Differenzierbarkeit von $\tilde{f}(x, y) = (\operatorname{Re} f(x + iy), \operatorname{Im} f(x + iy))$, aber die Umkehrung ist nicht wahr.

Satz 3.45. *Sei $\Omega \subset \mathbb{C}$ offen. Die Funktion $f : \Omega \rightarrow \mathbb{C}$ ist genau dann holomorph, wenn $\operatorname{Re} f, \operatorname{Im} f$ als Funktionen zweier reeller Variablen auf Ω stetig differenzierbar sind, und*

$$\frac{\partial}{\partial x} \operatorname{Re} f(x + iy) = \frac{\partial}{\partial y} \operatorname{Im} f(x + iy) \quad \text{und} \quad \frac{\partial}{\partial y} \operatorname{Re} f(x + iy) = -\frac{\partial}{\partial x} \operatorname{Im} f(x + iy) \quad (67)$$

In diesem Fall gilt

$$f'(z) = \partial_x \operatorname{Re} f(z) + i \partial_x \operatorname{Im} f(z) = \partial_y \operatorname{Im} f(z) - i \partial_y \operatorname{Re} f(z)$$

für alle $z \in \Omega$. Diese zwei Bedingungen werden Cauchy-Riemann-Gleichungen genannt.

Beweis: Die stetige Differenzierbarkeit von $\operatorname{Re} f, \operatorname{Im} f$ impliziert, dass für alle $z \in \Omega$,

$$f(z + h) - f(z) - L(h) = o(\|h\|)$$

für $h \rightarrow 0$, mit der reell linearen Abbildung

$$L(h_1 + ih_2) = (\partial_x \operatorname{Re} f(z)h_1 + \partial_y \operatorname{Re} f(z)h_2) + i(\partial_x \operatorname{Im} f(z)h_1 + \partial_y \operatorname{Im} f(z)h_2)$$

Damit f komplex differenzierbar an der Stelle z_0 ist, muss L komplex linear sein, d.h. es muss ein $a = (a_1 + ia_2) \in \mathbb{C}$ existieren, mit

$$L(h_1 + ih_2) = (a_1 + ia_2)(h_1 + ih_2) = (a_1h_1 - a_2h_2) + i(a_1h_2 + a_2h_1)$$

Koeffizientenvergleich ergibt

$$\begin{aligned} a_1 &= \partial_x \operatorname{Re} f(z) = \partial_y \operatorname{Im} f(z) \\ a_2 &= -\partial_y \operatorname{Re} f(z) = \partial_x \operatorname{Im} f(z) \end{aligned}$$

Sind die Cauchy-Riemann-Gleichungen erfüllt, so ist f an der Stelle z komplex differenzierbar, mit Ableitung

$$f'(z) = \partial_x \operatorname{Re} f(z) + i\partial_x \operatorname{Im} f(z) = \partial_y \operatorname{Im} f(z) - i\partial_y \operatorname{Re} f(z)$$

Die Stetigkeit von $\operatorname{Re} f(z)$ und $\operatorname{Im} f(z)$ impliziert dann, dass $f \in \mathcal{H}(\Omega)$. Die umgekehrte Implikation folgt einfach aus der Bemerkung, dass jede komplex lineare Abbildung auch reell linear ist. Damit ist jede komplex differenzierbare Funktion auch reell differenzierbar. \square

Beispiel: Jedes Polynom $f(z) = a_n z^n + \dots + a_1 z + a_0$ in einer komplexen Variablen ist auf \mathbb{C} holomorph. Wie bei Polynomen einer reellen Variablen findet man

$$f'(z) = na_n z^{n-1} + (n-1)a_{n-1} z^{n-2} + \dots + a_1$$

Jede rationale Funktion $P(z)/Q(z)$, wobei P, Q Polynome sind, ist holomorph auf $\mathbb{C} \setminus \{z \in \mathbb{C} : Q(z) = 0\}$. Die Exponentialfunktion $f(z) = \exp(z)$ ist auf \mathbb{C} holomorph, mit $f'(z) = e^z$. Jede Potenzreihe $f(z) = \sum_{n=0}^{\infty} a_n (z - z_0)^n$ ist innerhalb ihres Konvergenzradius komplex differenzierbar. Das folgt, weil die Potenzreihe gleichmässig innerhalb des Konvergenzradius konvergiert. Das impliziert, aus Prop. 8.29 in Analysis 1, dass man die Ableitung und den Grenzwert vertauschen kann. Genauer gesagt, gilt Prop. 8.29 nur für die Ableitung nach einer reellen Variablen. Trotzdem, kann man Prop. 8.29 benutzen, um die partiellen Ableitungen von $\operatorname{Re} f, \operatorname{Im} f$ nach $x = \operatorname{Re} z$ und $y = \operatorname{Im} z$ zu berechnen. Dann kann man leicht sehen, dass die partiellen Ableitungen stetig sind und dass sie die Cauchy-Riemann-Gleichungen erfüllen. Damit folgt auch, dass f komplex differenzierbar ist.

Die Funktion $f(z) = |z|^2$ ist a.d.S. $z_0 \in \mathbb{C}$ nicht differenzierbar, für alle $z_0 \in \mathbb{C} \setminus \{0\}$. In der Tat $f(x + iy) = x^2 + y^2 = \operatorname{Re} f(x + iy)$. Deswegen ist $\partial_x \operatorname{Im} f = \partial_y \operatorname{Im} f \equiv 0$ während

$$\partial_x \operatorname{Re} f(x + iy) = 2x, \quad \text{und} \quad \partial_y \operatorname{Im} f(x + iy) = 2y$$

Die Cauchy-Riemann-Gleichungen sind nur an der Stelle $x = y = 0$ erfüllt.

Die Eigenschaften der komplexen Ableitung sind ähnlich denen der entsprechenden Eigenschaften der Ableitung auf \mathbb{R} .

Proposition 3.46. *Sei $\Omega \subset \mathbb{C}$ offen, $f, g : \Omega \rightarrow \mathbb{C}$ an der Stelle $z \in \Omega$ komplex differenzierbar. Dann*

i) $f + g$ und fg sind ebenfalls an der Stelle z komplex differenzierbar, mit

$$(f + g)'(z) = f'(z) + g'(z), \quad \text{und} \quad (fg)'(z) = f'(z)g(z) + f(z)g'(z)$$

ii) Ist $g(z) \neq 0$, so ist auch f/g an der Stelle z komplex differenzierbar, mit

$$(f/g)'(z) = \frac{f'(z)g(z) - f(z)g'(z)}{g^2(z)}$$

iii) Sei $U \subset \mathbb{C}$ offen und $h : U \rightarrow \mathbb{C}$ an der Stelle $f(z) \in U$ komplex differenzierbar. Dann ist $h \circ f$ an der Stelle z komplex differenzierbar und

$$(h \circ f)'(z) = h'(f(z))f'(z)$$

Die Beweise sind den entsprechenden Beweisen in Analysis 1 sehr ähnlich.

Wir definieren nun den Begriff von Linienintegralen einer Funktion $f : \mathbb{C} \supset \Omega \rightarrow \mathbb{C}$ entlang einer stückweise stetig differenzierbaren Kurve γ .

Definition 3.47. Sei $\Omega \subset \mathbb{C}$ offen, $f : \Omega \rightarrow \mathbb{C}$ stetig und γ eine stückweise stetig differenzierbare Kurve auf Ω (d.h. $\gamma : [a; b] \rightarrow \Omega$ ist stetig und es existiert eine endliche Teilung $a = t_0 < t_1 < \dots < t_n = b$ mit $\gamma \in C^1([t_{j-1}; t_j]; \Omega)$ für alle $j = 1, \dots, n$). Dann definieren wir das Linien- oder Wegintegral von f entlang γ durch

$$\int_{\gamma} f(z) dz = \sum_{j=1}^n \int_{t_{j-1}}^{t_j} f(\gamma(t)) \gamma'(t) dt$$

Bemerkung: Das Linienintegral einer stetigen komplexen Funktion $f : \mathbb{C} \supset \Omega \rightarrow \mathbb{C}$ entlang einer stückweise stetig differenzierbaren Kurve $\gamma : [a; b] \rightarrow \Omega$ ist unabhängig von der Parametrisierung von γ . Sei nämlich $\psi : [c; d] \rightarrow [a; b]$ stückweise stetig differenzierbar und $\tilde{\gamma} : [c; d] \rightarrow \Omega$ definiert durch $\tilde{\gamma}(t) = \gamma(\psi(t))$. Dann gilt

$$\int_{\tilde{\gamma}} f(z) dz = \int_{\gamma} f(z) dz$$

Das kann man ähnlich wie in Proposition 3.37 zeigen, weil

$$\int_{t_{j-1}}^{t_j} f(\gamma(\psi(t))) \gamma'(\psi(t)) \psi'(t) dt = \int_{\psi(t_{j-1})}^{\psi(t_j)} f(\gamma(s)) \gamma'(s) ds$$

mit der Variablentransformation $s = \psi(t)$.

Das Linienintegral von f entlang γ ist eine komplexe Zahl. Ausgedrückt durch Real- und Imaginärteil von f und γ , ist es durch

$$\int_a^b [(\operatorname{Re} f(\gamma(t)) \operatorname{Re} \gamma'(t) - \operatorname{Im} f(\gamma(t)) \operatorname{Im} \gamma'(t)) + i (\operatorname{Re} f(\gamma(t)) \operatorname{Im} \gamma'(t) + \operatorname{Im} f(\gamma(t)) \operatorname{Re} \gamma'(t))] dt$$

gegeben (unter der Annahme, dass $\gamma \in C^1([a; b]; \Omega)$, sonst muss man das Intervall in eine geeignete Teilung zerlegen). Wir können $f : \mathbb{C} \supset \Omega \rightarrow \mathbb{C}$ auch als Funktion $\tilde{f}(x, y) = (\operatorname{Re} f(x + iy), \operatorname{Im} f(x + iy))$ interpretieren. \tilde{f} ist ein Vektorfeld, definiert auf

einer Teilmenge von \mathbb{R}^2 . Wir haben das Linienintegral eines Vektorfeldes schon in Sektion 3.8 definiert. Das Linienintegral von \tilde{f} , versehen als Vektorfeld auf \mathbb{R}^2 , entlang γ ist aus

$$\int_{\gamma} \tilde{f} \cdot dx = \int_a^b \tilde{f}(\gamma(t)) \cdot \gamma'(t) dt = \int_a^b \operatorname{Re} f(\gamma(t)) \operatorname{Re} \gamma'(t) + \operatorname{Im} f(\gamma(t)) \operatorname{Im} \gamma'(t) dt$$

unter der Annahme, dass $\gamma \in C^1([a; b]; \Omega)$ ist, gegeben. Das Linienintegral vom Vektorfeld \tilde{f} ist also nicht dasselbe, wie das Linienintegral der komplexen Funktion f . Dagegen gilt

$$\begin{aligned} \operatorname{Re} \int_{\gamma} f(z) dz &= \int_{\gamma} (\operatorname{Re} f, -\operatorname{Im} f) \cdot dx \\ \operatorname{Im} \int_{\gamma} f(z) dz &= \int_{\gamma} (\operatorname{Im} f, \operatorname{Re} f) \cdot dx \end{aligned}$$

Man bemerke, dass die Cauchy-Riemann-Gleichungen äquivalent mit der Tatsache sind, dass die zwei Vektorfelder $(\operatorname{Re} f, -\operatorname{Im} f)$ und $(\operatorname{Im} f, \operatorname{Re} f)$ rotationsfrei sind. Aus dieser Bemerkung folgt: Ist f holomorph auf einem konvexen $\Omega \subset \mathbb{C}$, und ist γ eine geschlossene stückweise stetig differenzierbare Kurve auf Ω , dann gilt $\int_{\gamma} f(z) dz = 0$.

Satz 3.48. *Sei $\Omega \subset \mathbb{C}$ offen, $f \in \mathcal{H}(\Omega)$, $a < b$, $\gamma, \tilde{\gamma} : [a; b] \rightarrow \Omega$ geschlossene und stückweise stetig differenzierbare Kurven auf Ω , so dass $s\tilde{\gamma}(t) + (1-s)\gamma(t) \in \Omega$ für alle $s \in [0; 1]$ und $t \in [a; b]$ (das ist sicher der Fall, wenn Ω konvex ist). Dann gilt*

$$\int_{\gamma} f(z) dz = \int_{\tilde{\gamma}} f(z) dz$$

Beweis: Um den Beweis zu vereinfachen nehmen wir an $\gamma, \tilde{\gamma} \in C^1([a; b])$ (sonst muss man das Intervall $[a; b]$ mit einer geeigneten Teilung zerlegen). Sei $\phi : [0, 1] \times [a; b] \rightarrow \Omega$, definiert durch

$$\phi(s; t) = s\tilde{\gamma}(t) + (1-s)\gamma(t)$$

Wir setzen

$$g(s) := \int_{\phi(s, \cdot)} f(z) dz = \int_a^b f(\phi(s, t)) \partial_t \phi(s, t) dt = \int_a^b \psi(s, t) dt$$

mit $\psi : [0, 1] \times [a; b] \rightarrow \mathbb{C}$ definiert durch

$$\psi(s, t) = f(\phi(s, t)) \partial_t \phi(s, t)$$

Dann gilt

$$g(0) = \int_{\gamma} f(z) dz, \quad \text{und} \quad g(1) = \int_{\tilde{\gamma}} f(z) dz$$

Wir möchten zeigen, dass $g(0) = g(1)$. Wir bemerken, dass ψ stetig ist und dass

$$\begin{aligned} \partial_s \psi(s, t) &= f'(\phi(s, t)) (\tilde{\gamma}(t) - \gamma(t)) \partial_t \phi(s, t) + f(\phi(s, t)) \partial_t (\tilde{\gamma}(t) - \gamma(t)) \\ &= \partial_t [f(\phi(s, t)) (\tilde{\gamma}(t) - \gamma(t))] \end{aligned}$$

auch stetig ist. Deswegen können wir Satz 3.31 anwenden. Wir erhalten

$$\begin{aligned} g'(s) &= \int_a^b \partial_s \psi(s, t) dt = \int_a^b \partial_t [f(\phi(s, t))(\tilde{\gamma}(t) - \gamma(t))] \\ &= f(\phi(s, b))(\tilde{\gamma}(b) - \gamma(b)) - f(\phi(s, a))(\tilde{\gamma}(a) - \gamma(a)) = 0 \end{aligned}$$

weil $\gamma(a) = \gamma(b)$ und $\tilde{\gamma}(a) = \tilde{\gamma}(b)$ und also auch $\phi(s, a) = \phi(s, b)$ für alle $s \in [0, 1]$. \square

Falls im letzten Satz $\Omega \subset \mathbb{C}$ auch konvex ist, so wissen wir schon, dass

$$\int_{\gamma} f(z) dz = \int_{\tilde{\gamma}} f(z) dz = 0$$

Satz 3.48 ist aber wichtig, weil er auch für nicht konvexe Ω gilt. Als Anwendung dieses Satzes sei $\Omega \subset \mathbb{C}$ eine offene konvexe Menge und $w \in \Omega$. Sei weiter f holomorph auf der nicht konvexen (und nicht einfach zusammenhängenden) Menge $\Omega \setminus \{w\}$. f kann aber in w eine Singularität haben. Deswegen ist das Linienintegral von f entlang einem geschlossenen Kreis um w im Allgemeinen nicht Null. Der Satz besagt aber, dass das Linienintegral entlang jedem Kreis (oder jede einfache geschlossene Kurve) um w (mit genügend kleinem Radius) immer denselben Wert hat. Diese Bemerkung benutzen wir im nächsten wichtigen Satz.

Satz 3.49. Sei $\Omega \subset \mathbb{C}$ offen, $f \in \mathcal{H}(\Omega)$, $x \in \Omega$ und $r \in (0, \infty)$ so klein, dass $\overline{B}_r(x) = \{z \in \mathbb{C} : |z - x| \leq r\} \subset \Omega$. Dann gilt, für alle $w \in B_r(x) = \{z \in \mathbb{C} : |z - x| < r\}$,

$$f(w) = \frac{1}{2\pi i} \int_{\gamma_{x,r}} \frac{f(z)}{z - w} dz.$$

Hier bezeichnet $\gamma_{x,r} : [0, 2\pi] \rightarrow \mathbb{C}$ die parametrisierte Kurve $\gamma_{x,r}(t) = x + re^{it}$ (das ist der Kreis mit Radius r um x , parametrisiert im Gegenuhrzeigersinn).

Beweis: Sei $0 < \rho < r - |w - x|$. Dann gilt $\overline{B}_\rho(w) \subset B_r(x)$. Ähnlich wie für $\gamma_{x,r}$, betrachten wir den Kreis $\gamma_{w,\rho}(t) = w + \rho e^{it}$ mit Radius ρ um w . Dann gilt, für alle $s \in [0; 1]$ und alle $t \in [0; 2\pi]$,

$$\phi(s, t) = s\gamma_{x,r}(t) + (1 - s)\gamma_{w,\rho}(t) \in \overline{B}_{x,r} \setminus B_{w,\rho} \subset \Omega \setminus \{w\}$$

Das folgt, weil $B_{x,r}$ konvex ist und weil

$$\begin{aligned} |\phi(s, t) - w| &= |s(x + re^{it}) + (1 - s)(w + \rho e^{it}) - w| \\ &= |s(x - w + (r - \rho)e^{it}) + \rho e^{it}| \\ &= |\rho + s((x - w)e^{-it} + (r - \rho))| \\ &\geq \rho + s(\operatorname{Re}(x - w)e^{-it} + (r - \rho)) > \rho \end{aligned}$$

Satz 3.48 impliziert also, dass

$$\begin{aligned} \int_{\gamma_{x,r}} \frac{f(z)}{z - w} dz &= \int_{\gamma_{w,\rho}} \frac{f(z)}{z - w} dz \\ &= \int_0^{2\pi} \frac{f(w + \rho e^{it})}{\rho e^{it}} i \rho e^{it} dt = i \int_0^{2\pi} f(w + \rho e^{it}) \end{aligned}$$

Deswegen

$$\frac{1}{2\pi i} \int_{\gamma_{x,t}} \frac{f(z)}{z-w} dz = \frac{1}{2\pi} \int_0^{2\pi} f(w + \rho e^{it}) dt$$

Da diese Formel für beliebige $\rho > 0$ klein genug gilt, gilt sie auch im Limes $\rho \rightarrow 0$. Wir finden

$$\frac{1}{2\pi i} \int_{\gamma_{x,t}} \frac{f(z)}{z-w} dz = \lim_{\rho \rightarrow 0} \frac{1}{2\pi} \int_0^{2\pi} f(w + \rho e^{it}) dt = f(w)$$

weil $f(w + \rho e^{it}) \rightarrow f(w)$ für $\rho \rightarrow 0$, gleichmässig in $t \in [0; 2\pi]$ (das kann z.B. durch Verwendung der Mittelwertabschätzung gezeigt werden). \square

Satz 3.49 gibt eine Darstellung von $f(w)$ durch ein Integral einer Funktion, die auf dem Kreis $\gamma_{x,r}$ stetig und differenzierbar ist. Das erlaubt uns, die Ableitung von f an der Stelle w zu berechnen, indem wir das Integrand nach w differenzieren.

Satz 3.50. Sei $\Omega \subset \mathbb{C}$ offen, $f \in \mathcal{H}(\Omega)$. Dann ist f auf Ω beliebig oft komplex differenzierbar. Falls $w \in B_r(x)$ und $\overline{B}_r(x) \subset \Omega$ gilt

$$f^{(n)}(w) = \frac{n!}{2\pi i} \int_{\gamma_{x,r}} \frac{f(z)}{(z-w)^{n+1}} dz$$

wobei $\gamma_{x,r}(t) = x + re^{it}$ der Kreis mit Radius r um x , parametrisiert im Gegenuhrzeigersinn, ist.

Beweis: Aus Satz 3.49 finden wir

$$f(w) = \frac{1}{2\pi i} \int_{\gamma_{x,r}} \frac{f(z)}{z-w} dz = \frac{1}{2\pi i} \int_0^{2\pi} \frac{f(x + re^{it})}{x - w + re^{it}} dt$$

Das Integrand ist nach w differenzierbar. Aus Satz 3.31 können wir Ableitung und Integral vertauschen. Wir bekommen

$$f'(w) = \frac{1}{2\pi i} \int_{\gamma_{x,r}} \frac{f(z)}{(z-w)^2} dz \tag{68}$$

Genauer gesagt, kann man Satz 3.31 nur für reelle Ableitungen anwenden. Man kann aber Satz 3.31 anwenden, um die partiellen Ableitungen $\partial_{w_1} \operatorname{Re} f(w_1 + iw_2)$, $\partial_{w_2} \operatorname{Re} f(w_1 + iw_2)$, $\partial_{w_1} \operatorname{Im} f(w_1 + iw_2)$ und $\partial_{w_2} \operatorname{Im} f(w_1 + iw_2)$ zu berechnen. Es ist dann einfach zu sehen, dass die Cauchy-Riemann-Gleichungen erfüllt sind, und dass $f'(w)$ in der Tat aus (68) gegeben ist (einfach weil die Funktion $(w-z)^{-1}$ auf $\gamma_{x,r}$ komplex differenzierbar ist, und Cauchy-Riemann-Gleichungen erfüllt). Aus (68) kann man dann analog die zweite Ableitung $f''(z)$ berechnen. Induktiv, findet man

$$f^{(n)}(w) = \frac{n!}{2\pi i} \int_{\gamma_{x,r}} \frac{f(z)}{(z-w)^n} dz.$$

\square

Es folgt aus dem letzten Satz, dass eine holomorphe Funktion automatisch beliebig oft komplex differenzierbar ist. Wir sehen, dass komplexe Differenzierbarkeit ein viel stärkerer Begriff, als reelle Differenzierbarkeit ist. Das hat viele wichtige Folgerungen und Anwendungen.

Proposition 3.51. Sei $\Omega \subset \mathbb{C}$ offen und $f : \Omega \rightarrow \mathbb{C}$ holomorph. Dann sind $\operatorname{Re} f$, $\operatorname{Im} f$ harmonische Funktionen auf Ω , d.h.

$$\Delta \operatorname{Re} f(z) = 0, \quad \Delta \operatorname{Im} f(z) = 0$$

für alle $z \in \Omega$.

Beweis: Es folgt aus Satz 3.50, dass $\operatorname{Re} f$, $\operatorname{Im} f$ beliebig oft differenzierbar sind. Aus den Cauchy-Riemann-Gleichungen (67) folgt, dass

$$\partial_x^2 \operatorname{Re} f(z) = \partial_x \partial_y \operatorname{Im} f(z) = \partial_y \partial_x \operatorname{Im} f(z) = -\partial_y^2 \operatorname{Re} f(z)$$

Deswegen gilt

$$\Delta \operatorname{Re} f(z) = \partial_x^2 \operatorname{Re} f(z) + \partial_y^2 \operatorname{Re} f(z) = 0$$

Analog zeigt man, dass $\Delta \operatorname{Im} f = 0$. □

Die letzte Proposition zeigt, dass Real- und Imaginärteil einer holomorphen Funktion immer harmonisch sind. Andererseits kann man zeigen, dass jede harmonische Funktion auf einer konvexen Teilmenge von \mathbb{R}^2 der Realteil (oder der Imaginärteil) einer holomorphen Funktion ist.

Proposition 3.52. Sei $\Omega \subset \mathbb{C}$ offen und konvex, $u \in C^2(\Omega)$ (hier wird Ω als Teilmenge von \mathbb{R}^2 versehen) mit $\Delta u = 0$. Dann existiert $v \in C^2(\Omega)$, so dass $u + iv \in \mathcal{H}(\Omega)$.

Beweis: Wir definieren das Vektorfeld $g : \Omega \rightarrow \mathbb{R}^2$ durch $g(x, y) = (-\partial_y u(x, y), \partial_x u(x, y))$. Dann gilt

$$\partial_y g_1(x, y) = -\partial_y^2 u(x, y) = \partial_x^2 u(x, y) = \partial_x g_2(x, y)$$

weil $\Delta u = 0$. Damit ist g auf Ω rotationsfrei. Es folgt aus Satz 3.43, dass eine Potentialfunktion für g existiert. D.h. es existiert $v \in C^1(\Omega)$ mit $g(x, y) = (\partial_x v(x, y), \partial_y v(x, y))$. Dann ist es leicht zu sehen, dass $f = u + iv$ die Cauchy-Riemann-Gleichungen erfüllt. Damit ist f holomorph und v harmonisch. □

Eine wichtige Folgerung dieser Charakterisierung von harmonischen Funktionen ist das folgende Korollar.

Korollar 3.53. Sei $\Omega \subset \mathbb{R}^2$ offen und konvex (einfach zusammenhängend ist genug) und $u \in C^2(\Omega)$ harmonisch (d.h. $\Delta u = 0$). Dann ist u beliebig oft differenzierbar.

Eine andere wichtige Anwendung von Satz 3.50 ist der Satz von Liouville

Satz 3.54 (Liouville). Sei $f \in \mathcal{H}(\mathbb{C})$ beschränkt. Dann ist f konstant.

Beweis: Sei $M = \sup_{z \in \mathbb{C}} |f(z)|$. Aus Satz 3.49 folgt, dass

$$f'(w) = \frac{1}{2\pi i} \int_{\gamma_{0,r}} \frac{f(z)}{(z-w)^2} dz = \frac{1}{2\pi} \int_0^{2\pi} \frac{f(re^{it})}{re^{it} - w} re^{it} dt$$

für alle $r > |w|$. Deswegen

$$|f'(w)| \leq \frac{M}{2\pi} \int_0^{2\pi} \frac{r}{|re^{it} - w|^2} dt \leq \frac{Mr}{(r - |w|)^2}$$

Da $r > |w|$ beliebig ist, finden wir

$$|f'(w)| \leq \lim_{r \rightarrow \infty} \frac{Mr}{(r - |w|)^2} = 0$$

Also $f'(w) = 0$ für alle $w \in \mathbb{C}$. Das impliziert, dass f konstant ist. \square

Der Fundamentalsatz der Algebra folgt einfach aus dem Satz von Liouville.

Satz 3.55 (Fundamentalsatz der Algebra). *Jedes Polynom von Grad $p \geq 1$ besitzt mindestens eine Nullstelle in \mathbb{C} .*

Beweis: Sei P ein Polynom von Grad $p \geq 1$, mit $P(z) \neq 0$ für alle $z \in \mathbb{C}$. Dann ist $1/P(z)$ eine holomorphe Funktion auf \mathbb{C} (man spricht von einer ganzen Funktion). Ferner, da für jede $M > 0$, R mit $|P(z)| > M$ für alle $z \in \mathbb{C}$ mit $|z| > R$, existiert, schliessen wir, dass $1/P(z)$ beschränkt auf \mathbb{C} ist. Aus Satz 3.54 folgt, dass $1/P(z)$ eine konstante Funktion ist. Das impliziert auch, dass $P(z)$ konstant ist, in Widerspruch zu der Annahme, dass $P(z)$ ein Polynom von Grad $p \geq 1$ ist. \square

Tatsächlich sind holomorphe Funktionen auf einem offenen Gebiet $\Omega \subset \mathbb{C}$ nicht nur beliebig oft komplex differenzierbar, sondern sogar analytisch auf Ω . Der Begriff von analytischer Funktion ist hier genau wie für Funktionen auf \mathbb{R} definiert.

Definition 3.56. *Sei $\Omega \subset \mathbb{C}$ offen, $f : \Omega \rightarrow \mathbb{C}$. Die Funktion f ist an der Stelle $z_0 \in \Omega$ analytisch, falls ein $r > 0$ und eine Folge a_n existieren, so dass*

$$f(z) = \sum_{n=0}^{\infty} a_n (z - z_0)^n$$

für alle $z \in B_r(z_0) = \{w \in \mathbb{C} : |z_0 - w| < r\}$. f heisst analytisch auf Ω , wenn f analytisch an der Stelle z_0 ist, für alle $z_0 \in \Omega$.

Wie für reelle Funktionen zeigt man, dass analytische Funktionen beliebig oft differenzierbar sind. Das folgt aus der Tatsache, dass eine Potenzreihe $\sum_{n \geq 0} a_n (z - z_0)^n$ mit Konvergenzradius r , für alle $r_0 < r$, auf $\overline{B}_{r_0}(z_0) = \{z \in \mathbb{C} : |z - z_0| \leq r_0\}$ gleichmässig konvergent ist. Deswegen kann man Summe und Ableitung vertauschen; siehe Analysis 1, Proposition 8.29 (siehe auch das Argument in den Beispielen unter Satz 3.45 um Prop. 8.29 aus Analysis 1 für komplexe Ableitungen zu benutzen). Induktiv zeigt man, dass f beliebig oft differenzierbar ist. Die Umkehrung gilt für reelwertige Funktionen auf $U \subset \mathbb{R}$ i.A. nicht. D.h. es existieren Funktionen $f : \mathbb{R} \rightarrow \mathbb{R}$, die unendlich oft differenzierbar sind, aber nicht analytisch. Das ist bei komplex differenzierbaren Funktionen nicht möglich. Jede holomorphe Funktion ist automatisch analytisch.

Satz 3.57. *Sei $\Omega \subset \mathbb{C}$, $f \in \mathcal{H}(\Omega)$, $z_0 \in \Omega$. Sei $r > 0$ so klein, dass $\overline{B}_r(z_0) = \{z \in \mathbb{C} : |z - z_0| \leq r\} \subset \Omega$. Dann gilt, für alle $z \in B_r(z_0) = \{z \in \mathbb{C} : |z - z_0| < r\}$,*

$$f(z) = \sum_{n=0}^{\infty} \frac{f^{(n)}(z_0)}{n!} (z - z_0)^n$$

Insbesondere ist f analytisch auf Ω .

Beweis: Für $z \in B_r(z_0)$ schreiben wir

$$f(z) = \frac{1}{2\pi i} \int_{\gamma_{z_0,r}} \frac{f(w)}{w-z} dw$$

wobei $\gamma_{z_0,r}(t) = z_0 + re^{it}$, $t \in [0; 2\pi]$, der Kreis von Radius r um z_0 ist, parametrisiert im Gegenuhrzeigersinne. Weiter, für $w \in \gamma_{z_0,r}([0; 2\pi])$ gilt

$$\frac{1}{w-z} = \frac{1}{w-z_0+z_0-z} = \frac{1}{w-z_0} \frac{1}{1-\frac{z-z_0}{w-z_0}}$$

Da $|z-z_0| < |w-z_0| = r$, finden wir

$$\frac{1}{w-z} = \frac{1}{w-z_0} \sum_{n=0}^{\infty} \frac{(z-z_0)^n}{(w-z_0)^n} = \sum_{n=0}^{\infty} \frac{(z-z_0)^n}{(w-z_0)^{n+1}}$$

wobei die Summe für $w \in \gamma_{z_0,r}$ gleichmäßig konvergiert. Deswegen können wir Integral und Summe vertauschen und bekommen

$$f(z) = \frac{1}{2\pi i} \int_{\gamma_{z_0,r}} \frac{f(w)}{w-z} dw = \sum_{n=0}^{\infty} \frac{1}{2\pi i} \int_{\gamma_{z_0,r}} \frac{f(w)}{(w-z_0)^{n+1}} (z-z_0)^n$$

Aus Satz 3.50 erhalten wir

$$f(z) = \sum_{n=0}^{\infty} \frac{f^{(n)}(z_0)}{n!} (z-z_0)^n.$$

□

Wie wir schon bei der Untersuchung von analytischen Funktionen auf \mathbb{R} diskutiert haben, hat die Analytizität wichtige Folgerungen. Z.B. gilt der folgende Identitätssatz.

Satz 3.58 (Identitätssatz). *Sei $U \subset \mathbb{C}$ eine nichtleere offene und zusammenhängende Teilmenge von \mathbb{C} und seien f, g holomorph auf U . Dann sind die folgenden Aussagen äquivalent.*

- i) $f(z) = g(z)$ für alle $z \in U$.
- ii) Die Menge $\{z \in U : f(z) = g(z)\}$ enthält unendlich viele Punkte und besitzt einen Häufungspunkt in U .
- iii) Es gibt ein $z_0 \in U$, so dass $f^{(n)}(z_0) = g^{(n)}(z_0)$ für alle $n \geq 0$.

Beweis: Die Implikationen i) \Rightarrow ii) und i) \Rightarrow iii) sind trivial. Wir zeigen ii) \Rightarrow iii) und iii) \Rightarrow i).

ii) \Rightarrow iii): Sei $h = f - g$ und $z_0 \in U$ ein Häufungspunkt der Menge $M = \{z \in U : h(z) = 0\}$. Wir behaupten, dass $h^{(n)}(z_0) = 0$ für alle $n \in \mathbb{N}$. Nehmen wir an, es existiert $m \in \mathbb{N}$ mit $h^{(m)}(z_0) \neq 0$. Sei m der kleinste Index mit dieser Eigenschaft. Dann, aus Satz 3.57, gibt es eine Umgebung G von z_0 und eine auf G holomorphe Funktion φ mit $\varphi(z_0) \neq 0$ und

$$h(z) = (z-z_0)^m \varphi(z)$$

für alle $z \in G$. Aus Stetigkeit von φ gilt also $\varphi(z) \neq 0$ in einer Umgebung von z_0 . Damit ist auch $h(z) \neq 0$ für alle $z \neq z_0$ in einer Umgebung von z_0 . D.h. z_0 ist kein Häufungspunkt von M , in Widerspruch zur Annahme.

iii) \Rightarrow i): Sei $h = f - g$ und $S_k = \{z \in U : h^{(k)}(z) = 0\}$. Da $h^{(k)}$ stetig ist, ist S_k geschlossen, für alle $k \in \mathbb{N}$. Damit ist auch $S := \bigcap_{k \geq 0} S_k$ abgeschlossen. Wir behaupten nun, S ist auch offen, als Teilmenge von U . Für $z_0 \in S$ beliebig, können wir die holomorphe Funktion h in einer Potenzreihe um z_0 entwickeln. Das zeigt, dass $h(z) = 0$ in einer offenen Umgebung von z_0 . Für jede $z_0 \in S$ existiert also $\varepsilon > 0$, so dass $B_\varepsilon(z_0) \subset S$. Da U zusammenhängend ist, muss entweder $S = U$ oder $S = \emptyset$ gelten. Die Annahme iii) impliziert, dass $S = U$. \square

Bemerkung: Sei $U \subset \mathbb{C}$ und $I \subset \mathbb{R}$ ein Intervall mit $I \subset U$. Sei f eine beliebige Funktion auf I . Dann gibt es höchstens eine auf U holomorphe Funktion, die auf I mit f übereinstimmt.

Die komplexe Analysis ist manchmal nützlich, um gewisse bestimmte Integrale von Funktionen auf \mathbb{R} zu berechnen. Wir betrachten zwei Beispiele.

Beispiel: wir möchten das uneigentliche Integral

$$\int_0^\infty \frac{\sin x}{x} dx$$

berechnen. Wir wissen aus Analysis 1, dass das Integral konvergiert. Um den Wert des Integrales zu berechnen, betrachten wir die Funktion $f(z) = e^{iz}/z$, die auf $\mathbb{C} \setminus \{0\}$ holomorph ist. Wir definieren weiter die stückweise stetig differenzierbare Kurve $\gamma : [0; 4] \rightarrow \mathbb{C}$ durch

$$\gamma(t) = \begin{cases} r + (R-r)t, & \text{falls } t \in [0; 1] \\ Re^{i\pi(t-1)}, & \text{falls } t \in [1; 2] \\ -R + (R-r)(t-2), & \text{falls } t \in [2; 3] \\ re^{i\pi(4-t)}, & \text{falls } t \in [3; 4] \end{cases}$$

Dann gilt, mit Satz 3.48,

$$\int_\gamma f(z) dz = 0$$

weil wir $\tilde{\gamma}(t) = i$ für alle t können, und dann $\phi(s, t) = s\gamma(t) + (1-s)\tilde{\gamma}(t) \in \mathbb{C} \setminus \{0\}$, wo f holomorph ist (und natürlich ist $\int_{\tilde{\gamma}} f(z) dz = 0$). Es folgt, dass

$$\int_r^R f(x) dx + \int_{-R}^{-r} f(x) dx = - \int_0^\pi f(Re^{it}) i Re^{it} dt + \int_0^\pi f(re^{it}) i e^{it} dt \quad (69)$$

Auf der linken Seite, haben wir

$$\int_r^R \frac{e^{ix} - e^{-ix}}{x} dx = 2i \int_r^R \frac{\sin x}{x} dx \rightarrow 2i \int_0^\infty \frac{\sin x}{x} dx$$

für $r \rightarrow 0$ und $R \rightarrow \infty$. Auf der rechten Seite von (69) haben wir dagegen

$$\int_0^\pi f(re^{it}) i re^{it} dt = i \int_0^\pi e^{ire^{it}} dt \rightarrow i\pi$$

für $r \rightarrow 0$, weil $\exp(ir \exp(it)) \rightarrow 1$ gleichmässig, da $r \rightarrow 0$ (und deswegen dürfen wir Limes und Integral vertauschen). Andererseits betrachten wir

$$\int_0^\pi f(Re^{it})iRe^{it} dt = i \int_0^\pi e^{iRe^{it}} dt$$

Also

$$\left| \int_0^\pi f(Re^{it})iRe^{it} dt \right| \leq \int_0^\pi e^{-R \sin t} dt$$

Für ein beliebiges $\varepsilon > 0$, gilt

$$\left| \int_0^\pi f(Re^{it})iRe^{it} dt \right| \leq \int_0^\varepsilon e^{-R \sin t} dt + \int_\varepsilon^{\pi-\varepsilon} e^{-R \sin t} dt + \int_{\pi-\varepsilon}^\pi e^{-R \sin t} dt \leq 2\varepsilon + \pi e^{-R \sin \varepsilon}$$

Deswegen

$$\lim_{R \rightarrow \infty} \left| \int_0^\pi f(Re^{it})iRe^{it} dt \right| \leq 2\varepsilon$$

für beliebiges $\varepsilon > 0$. Es folgt, dass

$$\lim_{R \rightarrow \infty} \left| \int_0^\pi f(Re^{it})iRe^{it} dt \right| = 0$$

und deswegen, aus (69), bekommen wir im Limes $R \rightarrow \infty, r \rightarrow 0$,

$$2i \int_0^\infty \frac{\sin x}{x} dx = i\pi$$

Das ergibt

$$\int_0^\infty \frac{\sin x}{x} dx = \frac{\pi}{2}.$$

Beispiel: das Gausssche Integral. Wir möchten das Integral

$$\int_{-\infty}^\infty e^{-x^2} dx$$

berechnen. Wir setzen $a = \sqrt{\pi/2}(1+i)$ und betrachten das komplexe Linienintegral

$$\int_{\gamma_R} \frac{e^{-z^2}}{1+e^{-2az}} dz$$

wobei, für $R > 0$, γ_R das Parallelogramm ist, mit den Eckpunkten $-R, R, R+a, -R+a$. Wir zerlegen $\gamma_R = \gamma_{1,R} + \gamma_{2,R} + \gamma_{3,R} + \gamma_{4,R}$, wobei $\gamma_{1,R} = [-R; R]$, $\gamma_{2,R} = [R; R+a]$, $\gamma_{3,R} = [R+a; -R+a]$ und $\gamma_{4,R} = [-R+a; -R]$ ist. Sei $\gamma_{2,R}(t) = R+ta$, $t \in [0; 1]$ eine Parametrisierung des Segmentes $\gamma_{2,R}$. Dann gilt

$$\int_{\gamma_{2,R}} \frac{e^{-z^2}}{1+e^{-2az}} dz = \int_0^1 \frac{e^{-(R+ta)^2}}{1+e^{-2a(R+ta)}} a dt \quad (70)$$

Bemerke, dass

$$\operatorname{Re}(R+ta)^2 = R^2 + \sqrt{2\pi}tR \geq R^2 - \sqrt{2\pi}R$$

weil $a^2 = i\pi$, und $t \in [0; 1]$. Deswegen ist

$$|e^{-(R+ta)^2}| = e^{-\operatorname{Re}(R+ta)^2} \leq e^{-(R^2 - \sqrt{2\pi}R)} \leq e^{-R^2/2}$$

falls R gross genug ist. Andererseits

$$1 + e^{-2a(R+ta)} = 1 + e^{-2aR}e^{-2\pi it} = 1 + e^{-\sqrt{2\pi}R}e^{-i(2\pi t + \sqrt{2\pi}R)}$$

und damit

$$\begin{aligned} |1 + e^{-2a(R+ta)}|^2 &= (1 + e^{-\sqrt{2\pi}R} \cos(2\pi t - \sqrt{2\pi}R))^2 + e^{-2\sqrt{2\pi}R} \sin^2(2\pi t - \sqrt{2\pi}R) \\ &= 1 + e^{-2\sqrt{2\pi}R} + 2e^{-\sqrt{2\pi}R} \cos(2\pi t - \sqrt{2\pi}R) \\ &\geq (1 - e^{-\sqrt{2\pi}R})^2 \geq 1/2 \end{aligned}$$

für R gross genug. Aus (70) finden wir also, dass

$$\left| \int_{\gamma_{2,R}} \frac{e^{-z^2}}{1 + e^{-2az}} dz \right| \leq C e^{-R^2/2} \rightarrow 0$$

für $R \rightarrow \infty$. Analog kann man zeigen, dass

$$\left| \int_{\gamma_{4,R}} \frac{e^{-z^2}}{1 + e^{-2az}} dz \right| \rightarrow 0$$

für $R \rightarrow \infty$. Andererseits,

$$\begin{aligned} \int_{\gamma_{1,R}} \frac{e^{-z^2}}{1 + e^{-2az}} dz + \int_{\gamma_{3,R}} \frac{e^{-z^2}}{1 + e^{-2az}} dz &= \int_{-R}^R \frac{e^{-t^2}}{1 + e^{-2at}} dt - \int_{-R}^R \frac{e^{-(t+a)^2}}{1 + e^{-2a(t+a)}} dt \\ &= \int_{-R}^R \left[\frac{e^{-t^2}}{1 + e^{-2at}} + \frac{e^{-t^2-2at}}{1 + e^{-2at}} \right] dt \\ &= \int_{-R}^R e^{-t^2} dt \end{aligned}$$

Wir erhalten, dass

$$\int_{-\infty}^{\infty} e^{-t^2} dt = \lim_{R \rightarrow \infty} \int_{\gamma_R} \frac{e^{-z^2}}{1 + e^{-2az}} dz$$

Die Funktion $f(z) = e^{-z^2}/(1 + e^{-2az})$ ist überall holomorph, ausser in den Punkten $z \in \mathbb{C}$ mit $e^{-2az} = -1$. Man findet, f ist auf $\mathbb{C} \setminus \{a(n + 1/2) : n \in \mathbb{Z}\}$ holomorph. Die einzige Singularität von f innerhalb der von γ_R berandeten Menge ist im Punkt $z_0 = a/2$. Deswegen, für $r > 0$ klein genug,

$$\int_{\gamma_R} \frac{e^{-z^2}}{1 + e^{-2az}} dz = \int_{\gamma_{a/2,r}} \frac{e^{-z^2}}{1 + e^{-2az}} dz$$

wobei $\gamma_{a/2,r}$ der Kreis von Radius r um $a/2$, parametrisiert im Gegenuhreigersinn, ist. Für r klein genug, können wir den Nenner in einer Taylorreihe um den Punkt $a/2$ entwickeln. Wir finden

$$\begin{aligned} 1 + e^{-2az} &= 1 - e^{-2a(z-a/2)} \\ &= \sum_{n \geq 1} \frac{(-2a)^n}{n!} (z - a/2)^n \\ &= -2a(z - a/2) \left[1 + \sum_{n \geq 2} \frac{(-2a)^{n-1}}{n!} (z - a/2)^{n-1} \right] \\ &=: -2a(z - a/2)g(z) \end{aligned}$$

für eine analytische Funktion g , definiert in einer Umgebung von $a/2$, mit $g(a/2) = 1$. Da $g(a/2) \neq 0$ ist, ist auch $h(z) = e^{-z^2}/g(z)$ eine analytische Funktion, mit $h(a/2) = e^{-i\pi/4} = (1 - i)/\sqrt{2}$. Also, aus Satz 3.49, finden wir, nach einer kleinen Rechnung,

$$\begin{aligned} \int_{\gamma_R} \frac{e^{-z^2}}{1 + e^{-2az}} dz &= \int_{\gamma_{a/2,r}} \frac{e^{-z^2}}{1 + e^{-2az}} dz \\ &= -\frac{1}{2a} \int_{\gamma_{a/2,r}} \frac{h(z)}{z - a/2} dz \\ &= -\frac{(2\pi i)h(a/2)}{2a} = \sqrt{\pi} \end{aligned}$$

Wir haben damit bewiesen, dass

$$\int_{-\infty}^{\infty} e^{-x^2} dx = \sqrt{\pi}.$$