X. Differentialrechnung mehrerer Veränderlicher

In diesem Kapitel wenden wir uns der Differentialrechnung von vektorwertigen Funktionen zu, die zudem von mehreren Argumenten abhängen, d.h., wir werden Funktionen $f\colon U \to \mathbb{R}^m$ betrachten, wobei $U \subseteq \mathbb{R}^n$ in der Regel eine offene Teilmenge sein wird. In Abschnitt X.1 diskutieren wir zunächst differenzierbare Kurven, d.h. den Fall $n = 1$, und in Abschnitt X.2 wenden wir uns dem allgemeinen Fall zu.

X.1. Kurven im $\mathbb{R}^n$

Nach den eher abstrakten Überlegungen des vorangegangenen Kapitels wenden wir uns nun konkreten geometrischen Objekten zu, nämlich Kurven im $\mathbb{R}^n$. Wir definieren Tangenten an eine Kurve und die Bogenlänge einer Kurve.

Definition X.1.1. Seien $a < b$ reelle Zahlen und

$$\gamma : [a, b] \to \mathbb{R}^n, \ t \mapsto \gamma(t) = \big(\gamma_1(t), \ldots, \gamma_n(t)\big)$$

eine Abbildung.
(a) Ist γ stetig, so heißt γ eine *stetige Kurve* oder ein *Weg* in $\mathbb{R}^n$.
(b) Der Weg γ heißt (*stetig*) *differenzierbar*, wenn alle Komponenten γ_j, $j = 1, \ldots, n$, (stetig) differenzierbar sind. Er heißt *stückweise stetig differenzierbar*, wenn γ stetig ist und eine Zerlegung

$$a = t_0 < t_1 < \ldots < t_k = b$$

existiert, so dass die Kurven $\gamma|_{[t_i, t_{i+1}]}$ stetig differenzierbar sind.
(c) Ist γ in t differenzierbar, so heißt

$$\dot\gamma(t) := \gamma'(t) := \big(\gamma_1'(t), \ldots, \gamma_n'(t)\big)$$

die *Ableitung* oder der *Geschwindigkeitsvektor* von γ bei t. Die Zahl

$$\|\dot\gamma(t)\|_2 := \Big(\sum_{j=1}^n |\gamma_j'(t)|^2\Big)^{\frac{1}{2}}$$

heißt *Geschwindigkeit von* γ in t. (Wir lassen den Index bei $\|\dot\gamma(t)\|_2$ meistens weg, wenn dadurch keine Verwechslungen möglich sind.)
(d) Ist $D = [a,b]$, so heißt $\gamma(a)$ *Anfangspunkt* und $\gamma(b)$ *Endpunkt* der Kurve.
(e) Ist γ komponentenweise integrierbar auf $[a,b]$, so setzen wir

$$\int_a^b \gamma(t)\,dt := \Big(\int_a^b \gamma_1(t)\,dt, \ldots, \int_a^b \gamma_n(t)\,dt\Big) \in \mathbb{R}^n. \qquad \blacksquare$$

Beispiel X.1.2. (a) Die Abbildung

$$\gamma : [0, 2\pi] \to \mathbb{R}^2,\ t \mapsto p + r(\cos t, \sin t)$$

beschreibt eine Kreiskurve vom Radius r um den Punkt $p \in \mathbb{R}^2$.
(b) Eine Ellipse um den Ursprung mit den Halbachsen a und b läßt sich durch die Kurve

$$\gamma : [0, 2\pi] \to \mathbb{R}^2,\ t \mapsto (a\cos t, b\sin t)$$

beschreiben.
(c) Die einfachsten Kurven sind (affine) Geraden:

$$\gamma : \mathbb{R} \to \mathbb{R}^n,\ t \mapsto p + tv,$$

wobei $p, v \in \mathbb{R}^n$ sind. Dann ist $\dot\gamma(t) = v$ konstant.
(d) Die *Neilsche Parabel*

$$\gamma : \mathbb{R} \to \mathbb{R}^2,\ t \mapsto (t^2, t^3)$$

ist überall differenzierbar, auch wenn ihr Bild im Nullpunkt eine Spitze besitzt. Es gilt $\dot\gamma(0) = 0$.
(e) Eine Schraubenlinie im $\mathbb{R}^3$ läßt sich durch die Kurve

$$\gamma : \mathbb{R} \to \mathbb{R}^3,\ \gamma(t) = (\cos t, \sin t, t)$$

beschreiben. $\blacksquare$

Definition X.1.3. Die *Gesamtbogenlänge* einer stückweise stetig differenzierbaren Kurve $\gamma : [a,b] \to \mathbb{R}$ ist

$$s(\gamma) := \sum_{j=0}^{k-1} \int_{t_j}^{t_{j+1}} \|\dot\gamma(t)\|_2\,dt,$$

wenn $a = t_0 < t_1 < \ldots t_k = b$ eine Unterteilung ist, für die $\gamma\,|_{[t_j,t_{j+1}]}$ für alle $j = 0, \ldots, k-1$, stetig differenzierbar ist. Dieses Integral existiert, weil die Integranden jeweils stetig sind. Die Funktion

$$s : [a,b] \to \mathbb{R},\ t \mapsto s(\gamma\,|_{[a,t]})$$

wird *Bogenlängenfunktion* genannt. Sie ist eine stetige, monoton wachsende Funktion (Nachweis!). Ist γ stetig differenzierbar, so auch s, und es gilt (nach dem Hauptsatz)

$$s'(t) = \|\dot\gamma(t)\|_2. \qquad \blacksquare$$

Beispiel X.1.4. (a) Wir betrachten das Geradenstück

$$\gamma : [0,1] \to \mathbb{R}^n, \quad t \mapsto a + t(b-a).$$

Dann ist

$$s(\gamma) = \int_0^1 \|\dot\gamma(t)\| \, dt = \int_0^1 \|b-a\| \, dt = \|b-a\|.$$

(b) Für den Kreisbogen $\gamma : [0, 2\pi] \to \mathbb{R}, t \mapsto (r\cos t, r\sin t)$ gilt

$$\|\dot\gamma(t)\| = \|(-r\sin t, r\cos t)\| = r,$$

also

$$s(\gamma) = \int_0^{2\pi} r \, dt = 2\pi r.$$

(c) Wir können für stückweise stetig differenzierbare Funktionen $f : [a,b] \to \mathbb{R}$ die Bogenlänge des Funktionsgraphen $\gamma : [a,b] \to \mathbb{R}^2$, $t \mapsto (t, f(t))$ berechnen. Mit $\dot\gamma(t) = (1, f'(t))$ erhalten wir

$$s(\gamma) = \int_a^b \sqrt{1 + f'(t)^2} \, dt.$$

Für den Viertel-Einheitskreis erhalten wir so aus $f(t) = \sqrt{1-t^2}$ für $0 \le t \le 1$ und $f'(t) = \frac{-t}{\sqrt{1-t^2}}$:

$$\sqrt{1+f'(t)^2} = \sqrt{1 + \frac{t^2}{1-t^2}} = \frac{1}{\sqrt{1-t^2}} = \arcsin'(t),$$

und hieraus

$$\int_0^1 \sqrt{1+f'(t)^2} \, dt = \int_0^1 \arcsin'(t) \, dt = \arcsin(1) - \arcsin(0) = \frac{\pi}{2}.$$

Man beachte, dass dieses Integral an der Stelle 1 uneigenlich ist, da der Integrand dort unbeschränkt ist. Nachdem wir die Zahl π in der Analysis I durch die Nullstellen der Cosinusfunktion definiert haben, zeigt uns obige Rechnung, dass die geometrische Interpretation der Zahl π als die Bogenlänge der halben Einheitskreislinie mit unserer Definition konsistent ist. ■

Definition X.1.5. Ist $\gamma\colon [a,b] \to \mathbb{R}^n$ eine stückweise stetig differenzierbare Kurve und $\varphi : [c,d] \to [a,b]$ stückweise stetig differenzierbar, so ist die Komposition $\gamma \circ \varphi : [c,d] \to \mathbb{R}^n$ wieder stückweise stetig differenzierbar. Ist φ bijektiv mit $\varphi' \ge 0$, so heißt φ eine *Umparametrisierung*. Insbesondere ist dann $\varphi(c) = a$ und $\varphi(d) = b$. ■

Satz X.1.6. *Ist φ eine Umparametrisierung der stückweise stetig differenzierbaren Kurve $\gamma : [a,b] \to \mathbb{R}^n$, so gilt $s(\gamma) = s(\gamma \circ \varphi)$, d.h. Umparametrisieren erhält die Bogenlänge.*

Beweis. Sei $t_0 = a < t_1 < \ldots < t_n = b$ eine Zerlegung, für die alle Wege $\gamma|_{[t_i,t_{i+1}]}$ stetig differenzierbar sind. Da $\varphi : [c,d] \to [a,b]$ bijektiv und monoton wachsend ist, gilt $c = \varphi^{-1}(t_0) < \varphi^{-1}(t_1) < \ldots < \varphi^{-1}(t_n) = d$. Wenden wir die Kettenregel komponentenweise an, so folgt $(\gamma \circ \varphi)'(t) = \dot\gamma(\varphi(t)) \cdot \varphi'(t)$ und daher wegen $\varphi'(t) \geq 0$:

$$\begin{aligned}\int_{\varphi^{-1}(t_i)}^{\varphi^{-1}(t_{i+1})} \|(\gamma \circ \varphi)'(t)\|\, dt &= \int_{\varphi^{-1}(t_i)}^{\varphi^{-1}(t_{i+1})} \|\dot\gamma(\varphi(t))\| \cdot \varphi'(t)\, dt \\ &= \int_{t_i}^{t_{i+1}} \|\dot\gamma(\tau)\|\ d\tau = s(\gamma|_{[t_i,t_{i+1}]}).\end{aligned}$$

Durch Zusammensetzen der Stücke erhält man die Behauptung. ■

Bemerkung X.1.7. Oft ist es bequem, eine Kurve auf ihre Bogenlänge zu parametrisieren. Ist $\gamma : [a,b] \to \mathbb{R}^n$ stetig differenzierbar mit $\dot\gamma(t) \neq 0$ für alle $t \in [a,b]$, so ist die Bogenlänge $s : [a,b] \to \mathbb{R}$ stetig differenzierbar mit $s'(t) = \|\dot\gamma(t)\| > 0$, also eine Bijektion $s : [a,b] \to [0, s(\gamma)]$ mit einer stetig differenzierbaren Umkehrfunktion s^{-1} (vgl. den Satz über die Differenzierbarkeit der Umkehrfunktion V.1.11). Die Kurve $\widetilde\gamma := \gamma \circ s^{-1} : [0, s(\gamma)] \to \mathbb{R}^n$ ist daher stetig differenzierbar mit

$$\dot{\widetilde\gamma}(t) = \dot\gamma(s^{-1}(t)) \cdot (s^{-1})'(t) = \dot\gamma(s^{-1}(t)) \frac{1}{s'(s^{-1}(t))} = \frac{\dot\gamma(s^{-1}(t))}{\|\dot\gamma(s^{-1}(t))\|}.$$

Dies ist offensichtlich ein Einheitsvektor. Es gilt also $\|\dot{\widetilde\gamma}\| = 1$ für alle $t \in [0, s(\gamma)]$. Daher heißt $\widetilde\gamma$ *über die Bogenlänge parametrisiert.* ■

Definition X.1.8. (*Kurvenintegral*) Sei $\gamma : [a,b] \to X \subseteq \mathbb{R}^n$ stückweise stetig differenzierbar und $f : X \to \mathbb{R}^k$ eine Funktion, für die die Komposition $f \circ \gamma$ integrabel ist. Dann heißt

$$\int_\gamma f := \int_a^b f(\gamma(t)) \cdot \|\dot\gamma(t)\|\, dt \in \mathbb{R}^k$$

das *Integral von f längs γ*. Hierbei beachten wir, dass das Produkt der beiden integrablen Funktionen $f \circ \gamma$ und $\|\dot\gamma\|$ ebenfalls integrabel ist (Lemma VI.1.13).■

Bemerkung X.1.9. (a) Man beachte, dass $\int_\gamma f$ ein Punkt im $\mathbb{R}^k$ ist.

(b) Die Bogenlänge der Kurve γ läßt sich mit dieser Definition als $s(\gamma) = \int_\gamma 1$ schreiben.

(c) Wir zeigen in den Übungen, dass das Integral $\int_\gamma f$ von der Parametrisierung von γ unabhängig ist.

(d) Ist speziell $\gamma = \mathrm{id}_{[a,b]}$ (als Kurve in $\mathbb{R}$), so ist $\int_\gamma f = \int_a^b f(t)\, dt$. ■

$\boxed{\text{Integralabschätzung}}$

Satz X.1.10. *Sei $X \subseteq \mathbb{R}^n$ eine Teilmenge, $\gamma : [a,b] \to X$ stückweise stetig differenzierbar und $f : X \to \mathbb{R}^k$ eine stetige Funktion. Dann ist*

$$\Big\| \int_\gamma f \Big\|_2 \leq \int_\gamma \|f\|_2 \leq s(\gamma) \cdot \sup_{a \leq t \leq b} \|f(\gamma(t))\|_2.$$

Für $\gamma = \mathrm{id}_{[a,b]}$ und $X = [a,b]$ erhalten wir insbesondere für jede Kurve $f\colon [a,b] \to \mathbb{R}^k$ die Abschätzung

$$\Big\| \int_a^b f(t)\, dt \Big\|_2 \leq \int_a^b \|f(t)\|_2 \, dt \leq M(b-a).$$

Beweis. Sei $\langle u, v\rangle = \sum_{j=1}^n u_j v_j$ das euklidische Skalarprodukt von $u, v \in \mathbb{R}^k$. Die Cauchy–Schwarzsche Ungleichung (Bemerkung IX.1.6) besagt dann

$$|\langle u, v\rangle| \leq \|u\|_2 \|v\|_2.$$

Für $v \in \mathbb{R}^k$ und integrable Kurven $\varphi : [a,b] \to \mathbb{R}^k$ gilt dann wegen der Linearität des Integrals

$$\int_a^b \langle \varphi(t), v\rangle \, dt = \int_a^b \sum_{j=1}^n \varphi_j(t) v_j \, dt = \sum_{j=1}^n \int_a^b \varphi_j(t)\, dt \cdot v_j = \langle \int_a^b \varphi(t)\, dt, v\rangle.$$

Für $v := \int_\gamma f$ ergibt sich damit aus der Cauchy–Schwarzschen Ungleichung

$$\begin{aligned}\|v\|_2^2 = \langle \int_\gamma f, v\rangle &= \int_a^b \langle f \circ \gamma(t), v\rangle \|\dot\gamma(t)\|_2 \, dt \\ &\leq \|v\|_2 \int_a^b \|f(\gamma(t))\|_2 \cdot \|\dot\gamma(t)\|_2 \, dt = \|v\|_2 \int_\gamma \|f\|_2,\end{aligned}$$

also

$$\|v\|_2 \leq \int_\gamma \|f\|_2 = \int_a^b \|f(\gamma(t))\|_2 \cdot \|\dot\gamma(t)\|_2 \, dt \leq \int_a^b \|\dot\gamma(t)\|_2 \, dt \cdot \sup_{a \leq t \leq b} \|f(\gamma(t))\|.$$

■

Folgerung X.1.11. *Ist $\gamma\colon [a,b] \to \mathbb{R}^n$ eine stückweise stetig differenzierbare Kurve, so gilt*

$$s(\gamma) \geq \|\gamma(b) - \gamma(a)\|.$$

Die Geraden sind also die kürzesten Verbindungen zweier Punkte.

Beweis. Wir wenden den zweiten Teil von Satz X.1.10 an und erhalten

$$\|\gamma(b) - \gamma(a)\| = \Big\| \int_a^b \dot\gamma(t)\, dt \Big\| \leq \int_a^b \|\dot\gamma(t)\| \, dt = s(\gamma).$$

■

Satz X.1.12. (Rechenregeln für Ableitungen von Skalarprodukten) *Sei* $D \subseteq \mathbb{R}$ *ein Intervall und* $\langle x, y \rangle := \sum_{j=1}^n x_j y_j$ *das Skalarprodukt auf dem* $\mathbb{R}^n$.

(1) *Sind* $\gamma, \varphi : D \to \mathbb{R}^n$ *in einem Punkt* $t \in D$ *differenzierbar, so ist auch* $\langle \gamma, \varphi \rangle : D \to \mathbb{R},\ t \mapsto \langle \gamma(t), \varphi(t) \rangle$ *in* t *differenzierbar, und es gilt*

$$\langle \gamma, \varphi \rangle'(t) = \langle \dot\gamma(t), \varphi(t) \rangle + \langle \gamma(t), \dot\varphi(t) \rangle.$$

(2) *Sind* $\gamma : D \to \mathbb{R}^n$ *und* $\varphi : D \to \mathbb{R}$ *wie oben, so ist*

$$\varphi \cdot \gamma : D \to \mathbb{R}^n,\ t \mapsto \varphi(t)\gamma(t)$$

in t *differenzierbar mit*

$$(\varphi \cdot \gamma)'(t) = \dot\varphi(t)\gamma(t) + \varphi(t)\dot\gamma(t).$$

Beweis. Wir wenden die Produktregel komponentenweise an (Übung). ■

X.2. Differenzierbare Abbildungen

In diesem Abschnitt werden wir sehen, wie sich das Konzept der Differenzierbarkeit in geeigneter Weise auf Funktionen in mehreren Veränderlichen übertragen lässt. Der begriffliche Aufwand wird hier dadurch etwas höher als im Eindimensionalen, dass die Ableitung einer Funktion $f: U \to \mathbb{R}^m$, $U \subseteq \mathbb{R}^n$ offen, in einem Punkt p jetzt eine lineare Abbildung $\mathtt{d}f(p): \mathbb{R}^n \to \mathbb{R}^m$ ist. Im Eindimensionalen kann man lineare Abbildungen $\mathbb{R} \to \mathbb{R}$ mit Zahlen identifizieren, so dass die Ableitung wieder eine Funktion $f': U \to \mathbb{R}$ wird, aber in der allgemeinen Situation erhalten wir eine Funktion

$$\mathtt{d}f: U \to \mathrm{Hom}(\mathbb{R}^n, \mathbb{R}^m),$$

und der Raum $\mathrm{Hom}(\mathbb{R}^n, \mathbb{R}^m)$, den wir mit dem Raum $M_{m,n}(\mathbb{R})$ der $(m \times n)$-Matrizen identifizieren können, hat die Dimension nm.

Im folgenden betrachten wir nur die euklidische Norm $\|\cdot\| := \|\cdot\|_2$ auf $\mathbb{R}^n$. Da alle Normen auf dem $\mathbb{R}^n$ äquivalent sind, spielt es keine Rolle, welche Norm wir hier verwenden.

Definition X.2.1. Sei $U \subseteq \mathbb{R}^n$ offen und $f: U \to \mathbb{R}^m$ eine Funktion.
(a) Die Funktion f heißt in $x \in U$ *differenzierbar*, wenn eine lineare Abbildung $A \in \mathrm{Hom}(\mathbb{R}^n, \mathbb{R}^m)$ existiert, so dass

$$f(x+h) = f(x) + A(h) + \varphi(h) \quad \text{mit } \lim_{h \to 0} \frac{\varphi(h)}{\|h\|} = 0 \tag{2.1}$$

gilt. Die Abbildung $h \mapsto f(x) + A(h)$ ist eine *affine Abbildung*, die f im Sinne von (2.1) in x von *erster Ordnung* approximiert.

(b) Ist f in x differenzierbar, so möchte man auch von *der* Ableitung von f in x reden. Hierzu hat man die Eindeutigkeit der linearen Abbildung A zu verifizieren. Sei dazu (2.1) erfüllt und $v \in \mathbb{R}^n$. Für ausreichend kleine $t \in \mathbb{R}$ ist dann $x + tv \in U$, und wir erhalten

$$\lim_{t\to 0} \frac{f(x+tv) - f(x)}{t} = \lim_{t\to 0} \frac{A(tv) + \varphi(tv)}{t} = Av + \lim_{t\to 0} \frac{\varphi(tv)}{t} = Av.$$

Also ist die lineare Abbildung A eindeutig durch f bestimmt. Die lineare Abbildung

$$\mathtt{d}f(x) := A \in \mathrm{Hom}(\mathbb{R}^n, \mathbb{R}^m)$$

heißt *Ableitung* oder *Differential von f im Punkt x*. Für $v \in \mathbb{R}^n$ heißt

$$\mathtt{d}f(x)(v) = \lim_{t\to 0} \frac{f(x+tv) - f(x)}{t} \tag{2.2}$$

die *Richtungsableitung von f in x in Richtung v*.

(c) Die Funktion f heißt *in U differenzierbar*, wenn sie in allen Punkten $x \in U$ differenzierbar ist.

(d) Sie heißt *in U stetig differenzierbar*, wenn f in U differenzierbar und die Funktion

$$\mathtt{d}f : U \to \mathrm{Hom}(\mathbb{R}^n, \mathbb{R}^m) \cong M_{m,n}(\mathbb{R}) \cong \mathbb{R}^{nm}$$

stetig ist. Die Menge der stetig differenzierbaren Funktionen $f : U \to \mathbb{R}^m$ bezeichnet man mit $C^1(U, \mathbb{R}^m)$. ■

Bemerkung X.2.2. (a) Wir müssen den Begriff der Differenzierbarkeit im Mehrdimensionalen anders definieren als im Eindimensionalen, da der Ausdruck

$$\frac{f(x+h) - f(x)}{h}$$

für Vektoren $h \in \mathbb{R}^n$ keinen Sinn ergibt.

(b) Die Definition X.2.1 passt jedoch gut mit der eindimensionalen Situation zusammen: Für $n = 1$ ist f eine Kurve und

$$\dot{f}(x) = \lim_{h\to 0} \frac{f(x+h) - f(x)}{h} = \mathtt{d}f(x)(1).$$

Beachte dabei, dass $\mathtt{d}f(x) : \mathbb{R} \to \mathbb{R}^n$ eine lineare Abbildung ist, die Größe $\dot{f}(x) = \mathtt{d}f(x)(1) \in \mathbb{R}^m$ also ein Vektor. ■

Lemma X.2.3. *Ist $U \subseteq \mathbb{R}^n$ offen und $f : U \to \mathbb{R}^m$ eine Funktion, so ist f in $x \in U$ genau dann differenzierbar, wenn dies für alle Komponentenfunktionen $f_j : U \to \mathbb{R}$, $j = 1, \ldots, m$, gilt.*

Beweis. Sei zunächst f in x differenzierbar und $A \in \mathrm{Hom}(\mathbb{R}^n, \mathbb{R}^m)$ mit

$$f(x+h) = f(x) + A(h) + \varphi(h) \quad \text{mit } \lim_{h\to 0} \frac{\varphi(h)}{\|h\|} = 0.$$

Sind $A_j \in \mathrm{Hom}(\mathbb{R}^n, \mathbb{R})$ die Komponentenfunktionen der linearen Abbildung A, so erhalten wir für die Komponentenfunktionen von f

$$f_j(x+h) = f_j(x) + A_j(h) + \varphi_j(h) \quad \text{mit } \lim_{h\to 0} \frac{\varphi_j(h)}{\|h\|} = 0, \tag{2.3}$$

denn da Konvergenz mit komponentenweiser Konvergenz gleichbedeutend ist (Satz IX.1.9), ist insbesondere $\lim_{h\to 0} \frac{\varphi(h)}{\|h\|} = 0$ äquivalent zu $\lim_{h\to 0} \frac{\varphi_j(h)}{\|h\|} = 0$ für alle $j \in \{1, \ldots, m\}$. Also ist jede Komponentenfunktion f_j in x differenzierbar.

Ist dies umgekehrt der Fall und gilt (2.3) für alle j, so betrachten wir die lineare Abbildung $A = (A_1, \ldots, A_m)\colon \mathbb{R}^n \to \mathbb{R}^m$ und erhalten mit $\varphi = (\varphi_1, \ldots, \varphi_m)$ die Beziehung

$$f(x+h) = f(x) + A(h) + \varphi(h) \quad \text{mit } \lim_{h\to 0} \frac{\varphi(h)}{\|h\|} = 0. \qquad \blacksquare$$

Das obige Lemma macht deutlich, dass die höhere Komplexität von Funktionen $f\colon U \to \mathbb{R}^m$ bzgl. Differenzierbarkeitseigenschaften weniger von der Anzahl m der Komponenten im Bildbereich kommt, als vielmehr von der Anzahl n der Komponenten im Urbildbereich.

Lemma X.2.4. *Ist $U \subseteq \mathbb{R}^n$ offen und $f\colon U \to \mathbb{R}^m$ in $x \in U$ differenzierbar, so ist f in x stetig.*

Beweis. Gemäß (2.1) haben wir

$$f(x+h) = f(x) + \mathtt{d}f(x)(h) + \varphi(h) \quad \text{mit } \lim_{h\to 0} \frac{\varphi(h)}{\|h\|} = 0.$$

Aus der Stetigkeit der linearen Abbildung $\mathtt{d}f(x)$ (Theorem IX.4.15) folgt

$$\lim_{h\to 0} \mathtt{d}f(x)(h) = 0,$$

und weiter ist $\lim_{h\to 0} \varphi(h) = \lim_{h\to 0} \frac{\varphi(h)}{\|h\|}\|h\| = 0$. Also erhalten wir

$$\lim_{h\to 0} f(x+h) - f(x) = 0,$$

d.h., f ist in x stetig. $\blacksquare$

Wir kommen nun zu einer Charakterisierung der Differenzierbarkeit, die im folgenden einige Beweise vereinfacht.

Satz X.2.5. *Die Abbildung $f : U \to \mathbb{R}^m$ ist genau dann in $x \in U$ differenzierbar, wenn eine Abbildung*

$$\Phi : U \to \mathrm{Hom}(\mathbb{R}^n, \mathbb{R}^m)$$

so existiert, dass Φ *im Punkt* x *stetig ist und die Beziehung*

$$f(x+h) = f(x) + \Phi(x+h)(h)$$

für $x+h \in U$ *gilt. In diesem Fall ist* $\mathtt{d}f(x) = \Phi(x)$.

Beweis. Sei zunächst

$$f(x+h) = f(x) + \Phi(x+h)(h),$$

wobei Φ in x stetig ist. Für

$$A := \Phi(x) \quad \text{und} \quad \varphi(h) := \big(\Phi(x+h) - \Phi(x)\big)(h) = \Phi(x+h)(h) - A(h)$$

gilt dann $f(x+h) = f(x) + A(h) + \varphi(h)$ sowie

$$\lim_{h\to 0} \frac{\varphi(h)}{\|h\|} = \lim_{h\to 0} \big(\Phi(x+h) - \Phi(x)\big)\Big(\frac{h}{\|h\|}\Big) = 0,$$

da für $h \neq 0$ die Beziehung

$$\begin{aligned} \Big\|\big(\Phi(x+h) - \Phi(x)\big)\big(\frac{h}{\|h\|}\big)\Big\| &\leq \big\|\Phi(x+h) - \Phi(x)\big\| \cdot \Big\|\frac{h}{\|h\|}\Big\| \\ &= \|\Phi(x+h) - \Phi(x)\| \underset{h\to 0}{\longrightarrow} 0 \end{aligned}$$

gilt.

Sei jetzt f in x differenzierbar und (2.1) erfüllt. Wir definieren $\Phi(x+h) \in \mathrm{Hom}(\mathbb{R}^n, \mathbb{R}^m)$ durch

$$\Phi(x+h)(v) := \begin{cases} A(v) & \text{für } h = 0 \\ A(v) + \langle h, v\rangle \frac{\varphi(h)}{\|h\|^2} & \text{für } h \neq 0. \end{cases}$$

Dann gilt für $h \neq 0$:

$$\begin{aligned} f(x+h) &= f(x) + A(h) + \varphi(h) = f(x) + A(h) + \langle h, h\rangle \frac{\varphi(h)}{\|h\|^2} \\ &= f(x) + \Phi(x+h)(h). \end{aligned}$$

Die Stetigkeit von Φ in x erhalten wir mit der Cauchy–Schwarzschen Ungleichung: Zunächst ist

$$\|\Phi(x+h)(v) - \Phi(x)(v)\| = |\langle h, v\rangle| \frac{\|\varphi(h)\|}{\|h\|^2} \leq \|h\| \cdot \|v\| \frac{\|\varphi(h)\|}{\|h\|^2} = \|v\| \frac{\|\varphi(h)\|}{\|h\|}$$

für alle $v \in \mathbb{R}^n$. Für $\|v\| \leq 1$ erhalten wir also

$$\|\Phi(x+h) - \Phi(x)\| \leq \frac{\|\varphi(h)\|}{\|h\|} \underset{h\to 0}{\longrightarrow} 0.$$

Damit ist Φ in x stetig. ∎

Satz X.2.6. (Rechenregeln für Ableitungen)

(a) *Sind* $f, g : U \to \mathbb{R}^m$ *in einem Punkt* $x \in U \subseteq \mathbb{R}^n$ *differenzierbare Funktionen, so ist die Funktion* $\lambda f + \mu g$ *für alle* $\lambda, \mu \in \mathbb{R}$ *in* x *differenzierbar, und es gilt*

$$\mathtt{d}(\lambda f + \mu g)(x) = \lambda \cdot \mathtt{d}f(x) + \mu \cdot \mathtt{d}g(x) \qquad \textit{(Linearität).}$$

(b) *Seien* $U \subseteq \mathbb{R}^n$ *und* $V \subseteq \mathbb{R}^m$ *offene Mengen,* $f : U \to V$ *im Punkte* $x \in U$ *differenzierbar und* $g : V \to \mathbb{R}^k$ *im Punkte* $f(x) \in V$ *differenzierbar. Dann ist die Funktion* $g \circ f : U \to \mathbb{R}^k$ *in* x *differenzierbar, und es gilt die*

$\boxed{\text{Kettenregel}}$

$$\mathtt{d}(g \circ f)(x) = \mathtt{d}g\big(f(x)\big) \circ \mathtt{d}f(x).$$

Beweis. (a) Mit Satz X.2.5 erhalten wir Funktionen $\Phi, \Psi \colon U \to \mathrm{Hom}(\mathbb{R}^n, \mathbb{R}^m)$, die in x stetig sind, so dass folgende Beziehungen für $x + u \in U$ gelten:

$$f(x+h) = f(x) + \Phi(x+h)(h) \quad \text{und} \quad g(x+h) = g(x) + \Psi(x+h)(h).$$

Dann ist

$$(\lambda f + \mu g)(x+h) = (\lambda f + \mu g)(x) + (\lambda\Phi + \mu\Psi)(x+h)(h),$$

und die Funktion $\lambda\Phi + \mu\Psi \colon U \to \mathrm{Hom}(\mathbb{R}^n, \mathbb{R}^m)$ ist in x stetig. Hieraus folgt die Differenzierbarkeit von f in x und

$$\mathtt{d}(\lambda f + \mu g)(x) = \lambda\Phi(x) + \mu\Psi(x) = \lambda\, \mathtt{d}f(x) + \mu\, \mathtt{d}g(x).$$

(b) Mit Satz X.2.5 erhalten wir eine Funktion $\Phi \colon U \to \mathrm{Hom}(\mathbb{R}^n, \mathbb{R}^m)$, die in x stetig ist, und eine Funktion $\Psi \colon V \to \mathrm{Hom}(\mathbb{R}^m, \mathbb{R}^k)$, die in $f(x)$ stetig ist, so dass folgende Beziehungen gelten:

$$f(x+h) = f(x) + \Phi(x+h)(h), \qquad \text{und} \qquad g(f(x)+k) = g\big(f(x)\big) + \Psi(f(x)+k)(k).$$

Dann ist $f(x+h) = f(x) + k$ mit $k = \Phi(x+h)(h)$ und daher

$$(g \circ f)(x+h) = g\big(f(x)\big) + \Psi\big(f(x+h)\big)\big(\Phi(x+h)(h)\big).$$

Wir haben $\Psi\big(f(x+h)\big) \circ \Phi(x+h) \in \mathrm{Hom}(\mathbb{R}^n, \mathbb{R}^k)$ und

$$\lim_{h \to 0} \Psi\big(f(x+h)\big) \circ \Phi(h+x) = \Psi\big(f(x)\big) \circ \Phi(x),$$

da f in x stetig ist (Lemma X.2.4) und die Komposition

$$\mathrm{Hom}(\mathbb{R}^m, \mathbb{R}^k) \times \mathrm{Hom}(\mathbb{R}^n, \mathbb{R}^m) \to \mathrm{Hom}(\mathbb{R}^n, \mathbb{R}^k), \quad (A, B) \mapsto A \circ B$$

wegen $\|A \circ B\| \leq \|A\| \cdot \|B\|$ stetig ist, denn sie ist eine bilineare Abbildung (Satz IX.4.18). Also ist $g \circ f$ in x differenzierbar, und das Differential ist gegeben durch

$$\mathtt{d}(g \circ f)(x) = \Psi\big(f(x)\big) \circ \Phi(x) = \mathtt{d}g\big(f(x)\big) \circ \mathtt{d}f(x). \qquad \blacksquare$$

Beispiel X.2.7. (a) Für eine affine Abbildung

$$f : \mathbb{R}^n \to \mathbb{R}^m, \ x \mapsto A(x) + b$$

ist $\mathtt{d}f(x) = A$ für alle $x \in \mathbb{R}^n$ (vgl. Def. X.2.1).

$$\boxed{\text{Allgemeine Produktregel}}$$

(b) Ist $f : \mathbb{R}^n \times \mathbb{R}^m \cong \mathbb{R}^{n+m} \to \mathbb{R}^k$ bilinear, so ist f überall differenzierbar mit

$$\mathtt{d}f(x,y)(v,w) = f(x,w) + f(v,y).$$

Hierbei schreiben wir Elemente aus $\mathbb{R}^n \times \mathbb{R}^m \cong \mathbb{R}^{n+m}$ jeweils als Paare (x,y) bzw. (v,w) mit $x,v \in \mathbb{R}^n$ und $y,w \in \mathbb{R}^m$. Für den Beweis schreiben wir

$$f(x+h, y+k) = f(x,y) + f(x,k) + f(h,y) + f(h,k).$$

Da f stetig ist, existiert wegen Satz IX.4.17 ein $C > 0$ mit

$$\|f(h,k)\| \leq C\|h\| \cdot \|k\|$$

für $(h,k) \in \mathbb{R}^n \times \mathbb{R}^m$. Für das quadratische Restglied $f(h,k)$ ergibt sich daher

$$\frac{\|f(h,k)\|}{\|(h,k)\|} \leq C\frac{\|h\| \cdot \|k\|}{\|(h,k)\|} \leq C\|k\|,$$

also $\lim_{(h,k)\to(0,0)} \frac{\|f(h,k)\|}{\|(h,k)\|} = 0$. Da die Abbildung

$$\mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^k, \quad (h,k) \mapsto f(x,k) + f(h,y)$$

linear ist, ergibt sich hieraus die Differenzierbarkeit von f in (x,y) sowie die Formel für $\mathtt{d}f(x,y)$. ■

Definition X.2.8. Ist $f : U \to \mathbb{R}^m$ im Punkt $x \in U \subseteq \mathbb{R}^n$ differenzierbar, so ist $\mathtt{d}f(x) \in \mathrm{Hom}(\mathbb{R}^n, \mathbb{R}^m)$ durch eine Matrix darstellbar. Wir wollen ihre Komponenten berechnen. Sei dazu

$$f(x) = \begin{pmatrix} f_1(x) \\ \cdot \\ \cdot \\ \cdot \\ f_m(x) \end{pmatrix}$$

mit $f_i : U \to \mathbb{R}$, $i = 1, \ldots, m$. Ist $e_j \in \mathbb{R}^n$, $j = 1, \ldots, n$, der j-te kanonische Basisvektor des $\mathbb{R}^n$ (die einzige Komponente ungleich 0 ist eine 1 an der i-ten Stelle), so heißt

$$\big(D_j f\big)(x) := \frac{\partial f}{\partial x_j}(x) := \lim_{t\to 0} \frac{1}{t}\big(f(x+te_j) - f(x)\big) = \mathtt{d}f(x)(e_j)$$

die *j-te partielle Ableitung von f in* $x \in U$. Entsprechend definiert man $\dfrac{\partial f_i}{\partial x_j}(x)$. Die Matrix

$$J_x(f) := \left(\frac{\partial f_i}{\partial x_j}(x)\right)_{\substack{i=1,\ldots,m \\ j=1,\ldots,n}} = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(x) & \cdots & \frac{\partial f_1}{\partial x_n}(x) \\ \frac{\partial f_2}{\partial x_1}(x) & \cdots & \frac{\partial f_2}{\partial x_n}(x) \\ \vdots & & \vdots \\ \frac{\partial f_m}{\partial x_1}(x) & \cdots & \frac{\partial f_m}{\partial x_n}(x) \end{pmatrix}$$

heißt *Jacobimatrix* von f in x. ■

Satz X.2.9. *Ist $U \subseteq \mathbb{R}^n$ offen und die Funktion $f : U \to \mathbb{R}^m$ in $p \in U$ differenzierbar, so existieren alle partiellen Ableitungen $\frac{\partial f_i}{\partial x_j}$ in p, und die lineare Abbildung $\mathtt{d}f(p)$ wird bzgl. der kanonischen Basen in $\mathbb{R}^n$ und $\mathbb{R}^m$ durch die Jacobimatrix $J_p(f)$ dargestellt.*

Beweis. Die Existenz der partiellen Ableitungen wurde in Definition X.2.1 gezeigt. Ist

$$f(x) = \begin{pmatrix} f_1(x) \\ \cdot \\ \cdot \\ \cdot \\ f_m(x) \end{pmatrix}, \quad \text{so ist} \quad \mathtt{d}f(p)(v) = \begin{pmatrix} \mathtt{d}f_1(p)(v) \\ \cdot \\ \cdot \\ \cdot \\ \mathtt{d}f_m(p)(v) \end{pmatrix} \in \mathbb{R}^m,$$

wenn $v \in \mathbb{R}^n$ ist (siehe Lemma X.2.3). Für $v = \sum_{j=1}^n v_j e_j$ folgt hieraus

$$\mathtt{d}f_i(p)(v) = \sum_{j=1}^n \mathtt{d}f_i(p)(e_j) \cdot v_j = \sum_{j=1}^n \frac{\partial f_i}{\partial x_j}(p) v_j,$$

also

$$\mathtt{d}f(p)(v) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(x) & \cdots & \frac{\partial f_1}{\partial x_n}(x) \\ \frac{\partial f_2}{\partial x_1}(x) & \cdots & \frac{\partial f_2}{\partial x_n}(x) \\ \vdots & & \vdots \\ \frac{\partial f_m}{\partial x_1}(x) & \cdots & \frac{\partial f_m}{\partial x_n}(x) \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} = J_p(f) \cdot v,$$

d.h., das Differential $\mathtt{d}f(p)$ wird durch die Jacobimatrix $J_p(f)$ dargestellt. ■

In diesem Sinne können wir $f(x+h) - f(x) = \mathtt{d}f(x)(h) + \varphi(h)$ wie folgt durch Matrizen und Vektoren beschreiben:

$$\begin{pmatrix} f_1(x+h) - f_1(x) \\ f_2(x+h) - f_2(x) \\ \vdots \\ f_m(x+h) - f_m(x) \end{pmatrix} = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(x) & \cdots & \frac{\partial f_1}{\partial x_n}(x) \\ \frac{\partial f_2}{\partial x_1}(x) & \cdots & \frac{\partial f_2}{\partial x_n}(x) \\ \vdots & & \vdots \\ \frac{\partial f_m}{\partial x_1}(x) & \cdots & \frac{\partial f_m}{\partial x_n}(x) \end{pmatrix} \begin{pmatrix} h_1 \\ h_2 \\ \vdots \\ h_n \end{pmatrix} + \begin{pmatrix} \varphi_1(h) \\ \varphi_2(h) \\ \vdots \\ \varphi_m(h) \end{pmatrix}.$$

Beispiel X.2.10. Wir betrachten die Funktion

$$f : \mathbb{R}^3 \to \mathbb{R}^2, \quad (x, y, z) \mapsto \begin{pmatrix} x \cdot y \\ \sin x + \cos y \end{pmatrix}.$$

Dann ist

$$f_1(x, y, z) = xy \quad \text{und} \quad f_2(x, y, z) = \sin x + \cos y.$$

Die partiellen Ableitungen sind

$$\frac{\partial f_1}{\partial x}(x, y, z) = y, \quad \frac{\partial f_1}{\partial y}(x, y, z) = x, \quad \frac{\partial f_2}{\partial x}(x, y, z) = \cos x,$$

$$\frac{\partial f_2}{\partial y}(x,y,z) = -\sin y, \qquad \text{und} \qquad \frac{\partial f_1}{\partial z}(x,y,z) = \frac{\partial f_2}{\partial z}(x,y,z) = 0.$$

Damit wird das Differential von f in (x,y,z) durch die Matrix

$$J_{(x,y,z)}(f) = \begin{pmatrix} \frac{\partial f_1}{\partial x}(x,y,z) & \frac{\partial f_1}{\partial y}(x,y,z) & \frac{\partial f_1}{\partial z}(x,y,z) \\ \frac{\partial f_2}{\partial x}(x,y,z) & \frac{\partial f_2}{\partial y}(x,y,z) & \frac{\partial f_2}{\partial z}(x,y,z) \end{pmatrix} = \begin{pmatrix} y & x & 0 \\ \cos x & -\sin y & 0 \end{pmatrix}$$

dargestellt. ■

Bemerkung X.2.10b. Ist $g \circ f$ eine Komposition differenzierbarer Abbildungen $f\colon U \to \mathbb{R}^m$ und $g\colon V \to \mathbb{R}^k$, wobei $U \subseteq \mathbb{R}^n$ offen ist, so erhalten wir aus der Kettenregel die Beziehung

$$\mathtt{d}(g \circ f)(p) = \mathtt{d}g(f(p)) \circ \mathtt{d}f(p)$$

für die Ableitungen. Auf der Ebene der zugehörigen Jacobi-Matrizen wird hieraus die Produktformel

$$J_p(g \circ f) = J_{f(p)}(g) \cdot J_p(f),$$

wobei $\cdot$ für das Produkt einer $(k \times m)$-Matrix mit einer $(m \times n)$-Matrix steht. Für die partiellen Ableitungen der Komposition $g \circ f$ ergibt sich damit insbesondere

$$\frac{\partial (g \circ f)}{\partial x_j}(p) = \sum_{\ell=1}^m \frac{\partial g}{\partial x_\ell}(f(p)) \frac{\partial f_\ell}{\partial x_j}(p),$$

wenn man sich überlegt, wie die Einträge der Produktmatrix aussehen.

Ein wichtiger Spezialfall hiervon ergibt sich für $n = 1$. Dann sind $f\colon U \to \mathbb{R}^m$ und $g \circ f\colon U \to \mathbb{R}^k$ Kurven und wir erhalten

$$(g \circ f)'(t) = \mathtt{d}(g \circ f)(t)(1) = \mathtt{d}g(f(t))\mathtt{d}f(t)(1) = \mathtt{d}g(f(t))(f'(t))$$

bzw.

$$(g \circ f)'(t) = \sum_{\ell=1}^m \frac{\partial g}{\partial x_\ell}(f(p)) f_\ell'(t).$$ ■

Definition X.2.11. Ist $f : U \to \mathbb{R}$, $U \subseteq \mathbb{R}^n$ offen, im Punkt $x \in U$ differenzierbar, so heißt der (Zeilen-)Vektor

$$\operatorname{grad} f(x) := \nabla f(x) = \big(D_1 f(x), \cdots, D_n f(x)\big) = J_x(f) = \Big(\frac{\partial f}{\partial x_1}(x), \cdots, \frac{\partial f}{\partial x_n}(x)\Big)$$

der *Gradient von f in x.* ■

Für jedes $v \in \mathbb{R}^n$ ist die Ableitung in Richtung v dann durch

$$\mathtt{d}f(x)(v) = \langle \nabla f(x), v \rangle = \sum_{j=1}^{n} D_j f(x) \cdot v_j$$

gegeben. Für alle Vektoren $v \in \mathbb{R}^n$ mit $\|v\| = 1$ gilt mit der Cauchy–Schwarzschen Ungleichung $|\mathtt{d}f(x)(v)| = |\langle \nabla f(x), v \rangle| \leq \|\nabla f(x)\|$. Für den speziellen Einheitsvektor $v = \frac{\nabla f(x)}{\|\nabla f(x)\|}$ gilt sogar ohne $\|\cdot\|$ Gleichheit:

$$\langle \nabla f(x), v \rangle = \frac{\|\nabla f(x)\|^2}{\|\nabla f(x)\|} = \|\nabla f(x)\|$$

(falls $\nabla f(x) \neq 0$). Der Gradient zeigt also in die Richtung des steilsten Anstiegs der Funktion f im Punkt x.

Für den Fall $n = 2$ nennt man die Teilmengen $H_c := \{x \in U \colon f(x) = c\}$ *Höhenlinien der Funktion* f. Man kann diese Linien verwenden, um sich das Verhalten der Funktion f zu veranschaulichen (Man denke zum Beispiel an eine Landkarte, die den Bereich U beschreibt, auf der man die Höhe des jeweiligen Punktes durch Höhenlinien einträgt). Ist nun $D \subseteq \mathbb{R}$ ein Intervall und $\gamma \colon D \to U$ eine Kurve, die in einer Höhenlinie verläuft, d.h. $\gamma(D) \subseteq H_c$ bzw. $f(\gamma(t)) = c$ für alle $t \in D$, so erhalten wir durch Ableiten mit der Kettenregel

$$0 = \frac{d}{dt} f(\gamma(t)) = \langle \nabla f(\gamma(t)), \dot\gamma(t) \rangle.$$

Geometrisch interpretiert man dies so, dass die Geschwindigkeit der Kurve γ senkrecht zum Gradienten $\nabla f(\gamma(t))$ in dem Punkt $\gamma(t)$ ist. Die Höhenlinien verlaufen also in jedem Punkt senkrecht zum Gradienten. Beschreibt f die Höhenfunktion einer Landkarte und ist $\gamma \colon D \to U$ ein Weg, den ein Wanderer durchläuft, so bedeutet $f(\gamma(t)) = c$, dass der Wanderer auf einem Höhenweg entlangläuft. Das ist zwar nicht anstrengend, er legt dabei aber auch keinen Höhenunterschied zurück. Ein Bergsteiger würde eher einen Weg mit

$$\dot\gamma(t) = \nabla f(\gamma(t)),$$

einen sogenannten *Gradientenweg*, vorziehen. ■

Beispiel X.2.12. Die Existenz der partiellen Ableitungen $\dfrac{\partial f_i(x)}{\partial x_j}$ ist zwar notwendig, aber nicht hinreichend für die Differenzierbarkeit (Stetigkeit) im Punkt x. Als Beispiel betrachten wir die Funktion

$$f : \mathbb{R}^2 \to \mathbb{R}, \ (x, y) \mapsto \begin{cases} 0, & \text{falls } (x, y) = 0 \\ \frac{xy}{x^2+y^2}, & \text{sonst.} \end{cases}$$

Im Punkt $(x, y) = (0, 0)$ ist dann

$$\frac{\partial f}{\partial x}(0,0) = \lim_{t \to 0} \frac{1}{t}\big(f(t,0) - f(0,0)\big) = 0 \quad \text{und} \quad \frac{\partial f}{\partial y}(0,0) = 0.$$

Aber die Funktion f ist im Nullpunkt *unstetig*, da für alle $t \neq 0$ gilt

$$f(t,t) = \frac{t^2}{t^2+t^2} = \frac{1}{2}.$$

Setzt man etwas stärkere Regularität der partiellen Ableitungen voraus, so lässt sich die Differenzierbarkeit allerdings doch durch die partiellen Ableitungen nachprüfen.

Satz X.2.13. *Die Funktion $f: U \to \mathbb{R}^m$ sei überall partiell differenzierbar, und die partiellen Ableitungen seien in $x \in U$ stetig. Dann ist f in x differenzierbar.*

Beweis. Wir dürfen o.B.d.A. $m = 1$ annehmen (vgl. Lemma X.2.3). Wir schreiben $h \in \mathbb{R}^n$ als $h = \sum_{j=1}^n h_j e_j$. Dann ist

$$\begin{aligned} f(x+h) - f(x) = \sum_{k=1}^n & f(x_1+h_1, \ldots, x_{k-1}+h_{k-1}, x_k+h_k, x_{k+1}, \ldots, x_n) \\ & - f(x_1+h_1, \ldots, x_{k-1}+h_{k-1}, x_k, x_{k+1}, \ldots, x_n). \end{aligned}$$

Wenden wir den Mittelwertsatz der Differentialrechnung auf jeden Summanden (als Funktion von h_k) an, so finden wir Zahlen $\vartheta_j \in]0,1[$, $j = 1, \ldots, n$, mit

$$f(x+h)-f(x) = \sum_{k=1}^n \big(D_k f\big)(x_1+h_1, \ldots, x_{k-1}+h_{k-1}, x_k+\vartheta_k h_k, x_{k+1}, \ldots, x_n) \cdot h_k.$$

Wir definieren nun $\Phi(x+h) \in \mathrm{Hom}(\mathbb{R}^n, \mathbb{R}^m)$ durch

$$\Phi(x+h)(v) := \sum_{k=1}^n \big(D_k f\big)(x_1+h_1, \ldots, x_{k-1}+h_{k-1}, x_k+\vartheta_k h_k, x_{k+1}, \ldots, x_n) \cdot v_k.$$

Dann erhalten wir

$$f(x+h) - f(x) = \Phi(x+h)(h).$$

Da

$$\begin{aligned} \lim_{h \to 0} \Phi_k(x+h) &= \lim_{h \to 0} (D_k f)(x_1+h_1, \ldots, x_{k-1}+h_{k-1}, x_k+\vartheta_k h_k, x_{k+1}, \ldots, x_n) \\ &= D_k f(x) = \Phi_k(x) \end{aligned}$$

nach Voraussetzung gilt, ist Φ in x stetig und daher f in x differenzierbar. ■

Mittelwertsatz

Satz X.2.14. *Sei $U \subseteq \mathbb{R}^n$ offen und $f: U \to \mathbb{R}$ differenzierbar. Sei $x+th \in U$ für alle $t \in [0,1]$. Dann existiert ein $\vartheta \in]0,1[$ mit*

$$f(x+h) - f(x) = \mathtt{d}f(x+\vartheta h)(h).$$

Beweis. Wir betrachten die differenzierbare Funktion

$$g: [0,1] \to \mathbb{R}, \quad t \mapsto f(x+th) \quad \text{mit} \quad g'(t) = \mathtt{d}f(x+th)(h).$$

Nach dem Mittelwertsatz der Differentialrechnung V.2.2 existiert ein $\vartheta \in]0,1[$ mit

$$f(x+h) - f(x) = g(1) - g(0) = g'(\vartheta) = \mathtt{d}f(x+\vartheta h)(h),$$

wobei die letzte Gleichung aus der Kettenregel folgt. ■

Bemerkung X.2.15. Für Funktionen $: U \to \mathbb{R}^m$ mit $m \geq 2$ gilt der Mittelwertsatz im allgemeinen nicht. Als Beispiel hierzu betrachten wir die Spiralkurve:

$$\gamma : \mathbb{R} \to \mathbb{R}^3,\ t \mapsto (\cos t, \sin t, t).$$

Dann ist $\gamma(0) = (1, 0, 0)$, $\gamma(2\pi) = (1, 0, 2\pi)$ und

$$\dot\gamma(t) = \mathtt{d}\gamma(t)(1) = (-\sin t, \cos t, 1).$$

Für $0 < t < 2\pi$ zeigt $\dot\gamma(t)$ nie in Richtung von $\gamma(2\pi) - \gamma(0)$, also ist

$$\gamma(2\pi) - \gamma(0) \neq \dot\gamma(t) \cdot 2\pi$$

für alle $t \in\ [0, 2\pi]$. ■

Für $m \geq 2$ ist die folgende Version des Mittelwertsatzes die nächstbeste und sehr nützlich:

$\boxed{\text{Satz vom endlichen Zuwachs}}$

Satz X.2.16. *Sei $U \subseteq \mathbb{R}^n$ offen und $f : U \to \mathbb{R}^m$ stetig differenzierbar sowie $x + th \in U$ für alle $t \in [0, 1]$. Dann ist*

$$f(x + h) - f(x) = \int_0^1 \mathtt{d}f(x + th)(h)\, dt.$$

Ist $\|\mathtt{d}f(x + th)\| \leq M$ für alle $t \in [0, 1]$, so gilt

$$\|f(x + h) - f(x)\| \leq M \cdot \|h\|.$$

Beweis. Für $g(t) := f(x+th)$ gilt $g'(t) = \mathtt{d}f(x+th)(h)$ (Bemerkung X.2.10b) und daher

$$f(x + h) - f(x) = g(1) - g(0) = \int_0^1 g'(t)\, dt = \int_0^1 \mathtt{d}f(x + th)(h)\, dt,$$

womit die erste Aussage schon gezeigt wäre. Für die zweite wenden wir die Integralabschätzung aus Satz X.1.9 auf $\gamma = \mathrm{id}_{[0,1]}$ an und erhalten so:

$$\begin{aligned}\|f(x + h) - f(x)\| &= \Big\| \int_0^1 \mathtt{d}f(x + th)(h)\, dt\Big\| \leq \int_0^1 \|\mathtt{d}f(x + h)(h)\|\, dt \\ &= \int_0^1 M \cdot \|h\|\, dt = M \cdot \|h\|.\end{aligned}$$

■

X.3. Höhere partielle Ableitungen und Taylorentwicklung

In diesem Abschnitt werden wir den Taylorschen Satz für differenzierbare Funktionen von mehreren Veränderlichen kennenlernen. Nachdem wir uns überlegt haben, wie wir die vielen Glieder, die in der Taylorentwicklung auftreten, geschickt bezeichnen, werden wir sehen, dass man im Prinzip genauso wie im Eindimensionalen vorgehen kann.
Es seien $U \subseteq \mathbb{R}^n$ eine offene Menge und $f : U \to \mathbb{R}^m$ eine Funktion. Wir nehmen an, dass die partiellen Ableitungen $D_j f(x) = \frac{\partial f}{\partial x_j}(x)$ für alle $x \in U$ und $j \in \{1, \ldots, n\}$ existieren. Für jedes $j \in \{1, \ldots, n\}$ ist dann $D_j f$ eine Funktion $D_j f : U \to \mathbb{R}^m$.

Definition X.3.1. (a) Sind die Funktionen $D_j f : U \to \mathbb{R}^m$, $j \in \{1, \ldots, n\}$ wieder partiell differenzierbar, so können wir für alle $x \in U$ die *höheren partiellen Ableitungen*

$$D_i D_j f(x) := D_i\big(D_j f\big)(x) =: \frac{\partial^2 f}{\partial x_i \partial x_j}(x)$$

definieren. Die Funktion f heißt dann *zweimal partiell differenzierbar*.
(b) Die Funktion f heißt *k-mal partiell differenzierbar* ($k \geq 2$), wenn sie $(k-1)$-mal partiell differenzierbar ist und alle partiellen Ableitungen

$$D_{i_{k-1}} D_{i_{k-2}} \ldots D_{i_1} f := D_{i_{k-1}}\big(D_{i_{k-2}} \ldots (D_{i_1} f) \cdots\big)$$

wieder partiell differenzierbar sind. Für jedes k-Tupel $(i_1, \ldots, i_k) \in \{1, \ldots, n\}^k$ erhalten wir dann wieder Funktionen

$$D_{i_k} D_{i_{k-1}} \ldots D_{i_1} f = \frac{\partial^k f}{\partial x_{i_k} \cdots \partial_{x_{i_1}}} : U \to \mathbb{R}^m,$$

die *k-ten partiellen Ableitungen von f*.
(c) Die Funktion f heißt *k-mal stetig partiell differenzierbar*, falls sie k-mal partiell differenzierbar ist und alle partiellen Ableitungen k-ter Ordnung stetig sind. ■

Als Beispiel betrachten wir die Funktion $f : \mathbb{R}^2 \to \mathbb{R},\ f(x, y) = 3x^2 y + y^3$. Dann ist f zweimal stetig partiell differenzierbar mit den partiellen Ableitungen:

$$\begin{aligned} &D_1 f(x, y) = 6xy, \quad D_1 D_1 f(x, y) = 6y, \quad D_2 f(x, y) = 3x^2 + 3y^2, \\ &D_2 D_2 f(x, y) = 6y, \quad D_1 D_2 f(x, y) = 6x = D_2 D_1 f(x, y). \end{aligned}$$

dass die „gemischten“ Ableitungen in der letzten Zeile übereinstimmen, ist kein Zufall:

Satz von Schwarz

Satz X.3.2. [1] *Sei $U \subseteq \mathbb{R}^n$ offen und $f : U \to \mathbb{R}^m$ zweimal stetig partiell differenzierbar. Dann gilt für alle $i, j \in \{1, \dots, n\}$*

$$D_i D_j f = D_j D_i f.$$

Beweis. Wir dürfen o.B.d.A. $m = 1$ annehmen, da wir die Komponenten von f getrennt behandeln können. Sei $u \in U$. Da U offen ist, existiert ein $\varepsilon > 0$, so dass $u + se_i + te_j \in U$ für alle Zahlen s und t mit $|s|, |t| < \varepsilon$ gilt. Sei $U' :=]-\varepsilon, \varepsilon[\times]-\varepsilon, \varepsilon[\subseteq \mathbb{R}^2$ und

$$\varphi \colon U' \to \mathbb{R}, \quad \varphi(s,t) := f(u + se_i + te_j).$$

Dann besagt die Voraussetzung insbesondere, dass $D_1 D_2 \varphi$ existiert und stetig ist. Zu zeigen ist nun

$$(D_i D_j f)(u) = (D_1 D_2 \varphi)(0,0) = (D_2 D_1 \varphi)(0,0) = (D_j D_i f)(u).$$

Nach Definition ist

$$\begin{aligned} D_2 D_1 \varphi(0,0) &= \frac{d}{dt}\Big|_{t=0} \lim_{s \to 0} \frac{\varphi(s,t) - \varphi(0,t)}{s} \\ &= \lim_{t \to 0} \lim_{s \to 0} \frac{\frac{1}{s}\big(\varphi(s,t) - \varphi(0,t)\big) - \frac{1}{s}\big(\varphi(s,0) - \varphi(0,0)\big)}{t}. \end{aligned}$$

Wir wenden den Mittelwertsatz der Differentialrechnung auf die zweite Variable dieses Ausdrucks an und erhalten so

$$D_2 D_1 \varphi(0,0) = \lim_{t \to 0} \lim_{s \to 0} \frac{1}{s}\big((D_2\varphi)(s, \vartheta_{s,t} t) - (D_2\varphi)(0, \vartheta_{s,t} t)\big)$$

für ein $\vartheta_{s,t} \in]0,1[$, das von t und s abhängt. Auf den so entstandenen Ausdruck wenden wir den Mittelwertsatz noch einmal an, diesmal für die erste Variable, und erhalten

$$D_2 D_1 \varphi(0,0) = \lim_{t \to 0} \lim_{s \to 0} D_1 D_2 \varphi(\widetilde{\vartheta}_{s,t} s, \vartheta_{s,t} t)$$

mit $0 < \vartheta_{s,t}, \widetilde{\vartheta}_{s,t} < 1$. Da $D_1 D_2 \varphi$ nach Voraussetzung stetig ist, folgt somit

$$D_2 D_1 \varphi(0,0) = D_1 D_2 \varphi(0,0). \qquad \blacksquare$$

[1] Hermann Amandus Schwarz (1843-1921), deutscher Mathematiker. Schüler von Kummer und Weierstraß in Berlin. Er war Professor in Halle, Zürich, Göttingen und der Berliner Akademie der Wissenschaften. Schwarz beschäftigte sich insbesondere mit der Funktionentheorie und zeigte vielfache Anwendungsmöglichkeiten auf. Nach ihm benannt sind die Cauchy-Schwarz-Ungleichung und der Satz von Schwarz.

Beispiel X.3.3. Der Satz von Schwarz hat eine interessante Konsequenz für die Existenz von „Stammfunktionen" von Funktionen in mehreren Veränderlichen. Gegeben sei eine stetig partiell differenzierbare Funktion $v : \mathbb{R}^2 \to \mathbb{R}^2$ (ein sogenanntes *Vektorfeld*) und gesucht sei eine *zweimal stetig partiell* differenzierbare Funktion $f : \mathbb{R}^2 \to \mathbb{R}$ mit

$$v = \operatorname{grad} f = (D_1 f, D_2 f),$$

d.h., f ist eine Lösung der *partiellen Differentialgleichung*

$$\frac{\partial f}{\partial x_1} = v_1, \quad \frac{\partial f}{\partial x_2} = v_2.$$

Der Satz von Schwarz liefert eine notwendige Bedingung für die Funktion v. Ist obige Gleichung erfüllt, so erhalten wir

$$D_2 v_1 = D_2 D_1 f = D_1 D_2 f = D_1 v_2.$$

Wir betrachten hierzu ein konkretes Beispiel: Für die Funktion

$$v : \mathbb{R}^2 \to \mathbb{R}^2, \ (x, y) \mapsto (-y, x)$$

ist $D_2 v_1(x, y) = -1 \neq 1 = D_1 v_2(x, y)$, also existiert keine Funktion f mit $v = \operatorname{grad} f$. In diesem Sinn hat das Vektorfeld v keine Potentialfunktion (es ist kein Gradientenfeld). Dies ist gleichbedeutend zu der Tatsache, dass es keine stetig differenzierbare Funktion $f : \mathbb{R}^2 \to \mathbb{R}$ gibt, die die beiden Gleichungen

$$\frac{\partial f}{\partial x}(x, y) = -y \quad \text{und} \quad \frac{\partial f}{\partial y}(x, y) = x$$

erfüllt. ■

Wir führen einige Bezeichnungen ein, die uns in diesem Abschnitt sehr viel Schreibarbeit ersparen werden.

Definition X.3.4. (a)

- Ein n-Tupel $\alpha = (\alpha_1, \ldots, \alpha_n) \in \mathbb{N}_0^n$ heißt *Multi-Index.*
- Die Zahl $|\alpha| := \alpha_1 + \ldots + \alpha_n$ heißt *Ordnung von* α.
- Die Zahl $\alpha! := \alpha_1! \cdot \ldots \cdot \alpha_n!$ heißt α*-Fakultät.*
- Die Zahl $\binom{\beta}{\alpha} := \prod_{j=1}^n \binom{\beta_j}{\alpha_j}$ heißt *Binomialkoeffizient.*
- Die Funktion $\mathbb{R}^n \to \mathbb{R}, x \mapsto x^\alpha := x_1^{\alpha_1} \cdot \ldots \cdot x_n^{\alpha_n}$ heißt *Monom vom Exponenten* α.
- Ist $P(x) = \sum_\alpha c_\alpha x^\alpha$ ein Polynom (die Summe sei endlich), so definieren wir seinen *Grad* durch $\deg P := \max\{|\alpha| : c_\alpha \neq 0\}$.

Auf der Menge der Multiindizes definiert man Addition und Subtraktion sowie eine partielle Ordnung:

- Wir definieren $\beta \leq \alpha :\iff \beta_i \leq \alpha_i$ für alle $i \in \{1, \ldots, n\}$, und
- $\alpha \pm \beta := (\alpha_1 \pm \beta_1, \ldots, \alpha_n \pm \beta_n)$, wobei $\alpha - \beta$ nur für $\beta \leq \alpha$ definiert ist.

(b) Für $k \in \mathbb{N}$ sei $D_i^k f := D_i(D_i^{k-1} f) =: \frac{\partial^k}{\partial x_i^k} f$ und $D_i^0 f := f$. Nun definieren wir eine Kurzschreibweise für höhere gemischte Ableitungen. Für $\alpha \in \mathbb{N}_0^n$ sei

$$D^\alpha f := D_1^{\alpha_1} D_2^{\alpha_2} \cdots D_n^{\alpha_n} f =: \frac{\partial^{|\alpha|}}{\partial x_1^{\alpha_1} \cdots \partial x_n^{\alpha_n}} f, \quad D^0 f := f.$$

(c) Eine Funktion $f : U \to \mathbb{R}^m$ heißt C^k*-Funktion* oder k*-mal stetig differenzierbar*, kurz: $f \in C^k(U, \mathbb{R}^m)$, falls sie k-mal stetig partiell differenzierbar ist (vgl. Satz X.2.13). Weiter sei

$$C^\infty(U, \mathbb{R}^m) := \bigcap_{k=0}^{\infty} C^k(U, \mathbb{R}^m).$$

Man beachte, dass sich für $f \in C^k(U, \mathbb{R}^m)$ jede partielle Ableitung der Ordnung $\leq k$ als ein $D^\alpha f$ schreiben lässt (Satz von Schwarz). Schließlich setzen wir noch

$$C^k(U) := C^k(U, \mathbb{R}).$$

■

Beispiel X.3.5. Ist $\beta \in \mathbb{N}_0^n$ und $f_\beta : \mathbb{R}^n \to \mathbb{R}$, $f_\beta(x) = x^\beta = x_1^{\beta_1} \cdot \ldots \cdot x_n^{\beta_n}$, so gilt

$$D^\alpha f_\beta(x) = \begin{cases} \frac{\beta!}{(\beta-\alpha)!} x^{\beta-\alpha}, & \text{falls } \alpha \leq \beta \\ 0, & \text{sonst.} \end{cases}$$

Um dies einzusehen, wendet man mehrfach die Produktregel an und erhält

$$\begin{aligned} D^\alpha x^\beta &= (D_1^{\alpha_1} x_1^{\beta_1}) \cdot \ldots \cdot (D_n^{\alpha_n} x_n^{\beta_n}) \\ &= \begin{cases} \frac{\beta_1!}{(\beta_1-\alpha_1)!} x_1^{\beta_1-\alpha_1} \cdot \ldots \cdot \frac{\beta_n!}{(\beta_n-\alpha_n)!} x_n^{\beta_n-\alpha_n} & \text{für } \alpha \leq \beta \\ 0 & \text{sonst} \end{cases} \end{aligned}$$

für alle $\alpha \leq \beta$. An der Stelle $x = 0$ erhalten wir insbesondere

$$(D^\alpha x^\beta)(0) = \begin{cases} 0, & \text{falls } \alpha \neq \beta \\ \alpha!, & \text{sonst.} \end{cases}$$

■

Definition X.3.6. Sei $f : U \to \mathbb{R}^m$ eine k-mal stetig differenzierbare Funktion. Das *k-te Taylorpolynom* von f bei $u \in U$ ist das Polynom

$$\begin{aligned} T_u^k(f)(x) &= \sum_{|\alpha| \leq k} \frac{(D^\alpha f)(u)}{\alpha!} x^\alpha \\ &= f(u) + (D_1 f)(u) x_1 + (D_2 f)(u) x_2 \\ &\quad + \frac{1}{2}(D_1^2 f)(u) x_1^2 + \frac{1}{2}(D_2^2 f)(u) x_2^2 + (D_1 D_2 f)(u) x_1 x_2 + \ldots. \end{aligned}$$

Man beachte dabei, dass $(D^\alpha f)(u)$ jeweils ein Vektor in $\mathbb{R}^m$ ist. ■

Wir erhalten die gleiche Charakterisierung des Taylorpolynoms wie in Kapitel VIII:

Bemerkung X.3.7. (a) Das Polynom $T_u^k(f)$ hat in 0 bis zur Ordnung k die gleichen Ableitungen wie f in u, d.h., für alle α mit $|\alpha| \leq k$ gilt

$$D^\alpha(T_u^k(f))(0) = (D^\alpha f)(u).$$

Dies folgt aus Beispiel X.3.5:

$$\begin{aligned} D^\alpha(T_n^k(f))(0) &= \Big(\sum_{|\beta|\leq k} D^\alpha \frac{\big(D^\beta f\big)(u)}{\beta!} x^\beta \Big)(0) \\ &= \sum_{|\beta|\leq k} \frac{\big(D^\beta f\big)(u)}{\beta!} \big(D^\alpha x^\beta\big)(0) = \frac{\big(D^\alpha f\big)(u)}{\alpha!}\alpha! = \big(D^\alpha f\big)(u). \end{aligned}$$

(b) Diese Eigenschaft bestimmt das k-te Taylorpolynom eindeutig, denn ist $p(x) = \sum_{|\alpha|\leq k} a_\alpha \cdot x^\alpha$ mit $a_\alpha \in \mathbb{R}^m$ ein Polynom mit $\big(D^\alpha p\big)(0) = \big(D^\alpha f\big)(u)$ für alle α mit $|\alpha| \leq k$, so ist

$$\frac{\big(D^\alpha f\big)(u)}{\alpha!} = \frac{\big(D^\alpha p\big)(0)}{\alpha!} = a_\alpha. \qquad \blacksquare$$

$\boxed{\text{Rechenregeln für Taylorpolynome}}$

Analog zum Fall $n = m = 1$ leiten wir die folgenden Rechenregeln ab.

Satz X.3.8. *Sei $U \subseteq \mathbb{R}^n$ offen und $u \in U$.*

(a) *Sind $f, g \in C^k(U, \mathbb{R}^m)$, $\lambda, \mu \in \mathbb{R}$ und $u \in U$, so gilt*

$$T_u^k(\lambda f + \mu g) = \lambda T_u^k(f) + \mu T_u^k(g).$$

(b) *Ist $f \in C^k(U)$ und $g \in C^k(U, \mathbb{R}^m)$, so gilt die Produktregel:*

$$T_u^k(f \cdot g) = T_0^k\big(T_u^k(f) \cdot T_u^k(g)\big).$$

(c) *Ist $V \subseteq \mathbb{R}^m$ offen und $g \in C^k(U, \mathbb{R}^m)$ mit $g(U) \subseteq V$ sowie $f \in C^k(V, \mathbb{R}^\ell)$, so gilt für $g(u) = v$ die allgemeine Kettenregel*

$$T_u^k(f \circ g) = T_0^k\big(T_v^k(f) \circ (T_u^k(g) - v)\big).$$

Beweis. (a) Dies folgt aus der Linearität der Abbildungen

$$D^\alpha : C^k(U, \mathbb{R}^m) \to C^{k-|\alpha|}(U, \mathbb{R}^m).$$

(b) Nach Verschieben um u dürfen wir o.B.d.A. $u = 0$ annehmen. Wir setzen $\varphi(x) := f(x) - T_0^k(f)(x)$ und $\psi(x) := g(x) - T_0^k(g)(x)$. Dann ist $\big(D^\alpha\varphi\big)(0) = \big(D^\alpha\psi\big)(0) = 0$ für alle α mit $|\alpha| \leq k$ und

$$(f \cdot g)(x) = T_0^k(f)(x) \cdot T_0^k(g)(x) + \varphi(x) \cdot T_0^k(g)(x) + f(x) \cdot \psi(x). \tag{3.1}$$

Wir erinnern uns nun an die Leibnizformel

$$(h_1 \cdot h_2)^{[n]} = \sum_{k=0}^{n} \binom{n}{k} h_1^{[k]} \cdot h_2^{[n-k]}$$

für die Ableitung von Produkten von Funktionen einer Veränderlichen (Satz VII.2.1). Diese Formel lässt sich leicht auf den Fall von mehreren Veränderlichen verallgemeinern:

$$\begin{aligned} D^\alpha(h_1 \cdot h_2) &= \sum_{\beta_1=0}^{\alpha_1} \cdots \sum_{\beta_n=0}^{\alpha_n} \binom{\alpha_1}{\beta_1} \cdots \binom{\alpha_n}{\beta_n} D^\beta(h_1) D^{\alpha-\beta}(h_2) \\ &= \sum_{\beta \leq \alpha} \binom{\alpha}{\beta} D^\beta(h_1) D^{\alpha-\beta}(h_2), \end{aligned} \tag{3.2}$$

indem man die Faktoren $D_j^{\alpha_j}$ von D^α nacheinander anwendet. An (3.2) lesen wir nun unmittelbar ab, dass

$$D^\alpha(\varphi \cdot T_0^k(g))(0) = D^\alpha(f \cdot \psi)(0) = 0$$

für alle α mit $|\alpha| \leq k$ gilt, d.h., die Restglieder in (3.1) liefern keinen Beitrag zum k-ten Taylorpolynom. Also gilt (b).

(c) Nach Ersetzen von g durch die Funktion $\widetilde{g}(x) := g(u+x) - v$ und f durch die Funktion $\widetilde{f}(x) := f(v+x)$, dürfen wir o.B.d.A. annehmen, dass $u = v = 0$ ist, denn die verschobenen Funktionen haben, bis auf den konstanten Term, die gleichen Taylorpolynome. Insbesondere ist dann $g(0) = 0$. Nun vereinfacht sich die Behauptung zu

$$T_0^k(f \circ g) = T_0^k\big(T_0^k(f) \circ T_0^k(g)\big).$$

Fall 1: Wir zeigen zuerst durch Induktion nach k, dass aus $T_0^k(f) = 0$ schon $T_0^k(f \circ g) = 0$ folgt.

Für $k = 0$ folgt dies aus $(f \circ g)(0) = f(g(0)) = f(0) = 0$.

Wir nehmen nun an, dass die Behauptung für $k-1$ gilt, d.h., $T_0^{k-1}(\widetilde{f}) = 0$ impliziert $T_0^{k-1}(\widetilde{f} \circ g) = 0$ für C^{k-1}-Funktionen $\widetilde{f} \in C^{k-1}(V, \mathbb{R}^\ell)$. Ist nun $|\alpha| = k$ und $\alpha_j > 0$, so erhalten wir mit der Kettenregel (Bemerkung X.2.10b) und der Leibnizformel:

$$\begin{aligned} D^\alpha(f \circ g)(0) &= D^{\alpha - e_j} D_j(f \circ g)(0) = \sum_{i=1}^{n} D^{\alpha - e_j}\big((D_i(f) \circ g) \cdot D_j(g_i)\big)(0) \\ &= \sum_{i=1}^{n} \sum_{\beta \leq \alpha - e_j} \binom{\alpha - e_j}{\beta} \underbrace{D^\beta(D_i(f) \circ g)(0)}_{=0} \cdot D^{\alpha-\beta}(g_i)(0), \end{aligned}$$

denn wir können die Induktionsvoraussetzung auf die Funktionen $D_i(f)$ anwenden, deren partielle Ableitungen bis zur Ordnung $k-1$ in 0 verschwinden. Damit ist $D^\alpha(f\circ g)(0)=0$. Für $|\alpha|<k$ folgt $D^\alpha(f\circ g)(0)=0$ ohnehin aus der Induktionsvoraussetzung. Daher ist $T_0^k(f\circ g)=0$. Wir haben also

$$0=T_0^k(f\circ g)=T_0^k\big(\underbrace{T_0^k(f)}_{=0}\circ T_0^k(g)\big)$$

gezeigt.
Fall 2: Allgemein setzen wir $\varphi := f-T_0^k(f)$ und beachten $T_0^k(\varphi)=0$. Dann können wir (1) anwenden und erhalten mit Fall 1:

$$T_0^k(f\circ g)=T_0^k\big(T_0^k(f)\circ g+\varphi\circ g\big)=T_0^k\big(T_0^k(f)\circ g\big).$$

Da $T_0^k(f)$ ein Polynom ist, erhalten wir durch mehrmaliges Anwenden von (a) und (b):

$$T_0^k(f\circ g)=T_0^k\big(T_0^k(f)\circ T_0^k(g)\big).$$ ■

Beispiel X.3.9. Gesucht sei das Taylorpolynom $T_0^2(f)$ der Funktion

$$f\colon\mathbb{R}^2\to\mathbb{R},\qquad f(x_1,x_2)=e^{x_1^2+\cos x_2}.$$

Wir wollen die allgemeine Kettenregel Satz X.3.8(c) anwenden und schreiben dazu $f=g\circ h$ für

$$g\colon\mathbb{R}\to\mathbb{R},\quad g(x)=e^x\quad\text{ und }\quad h\colon\mathbb{R}^2\to\mathbb{R},\quad h(x_1,x_2)=x_1^2+\cos x_2.$$

In unserem Fall ist $u=(0,0)$ und $v=h(u)=1$. Wir haben also

$$T_0^2(f)=T_0^2\big(T_1^2(g)\circ(T_0^2(h)-1)\big).$$

Über die Reihenentwicklung der Kosinusfunktion erhalten wir direkt

$$T_0^2(h)(x)=x_1^2+1-\frac{x_2^2}{2},$$

denn alle Terme höherer Ordnung tragen nichts zu den Ableitungen bis zur Ordnung 2 in 0 bei. Weiter ist

$$T_1^2(g)(y)=g(1)+g'(1)y+\frac{1}{2}g''(1)=e+ey+\frac{1}{2}ey^2.$$

Wir erhalten also

$$T_1^2(g)\circ(T_0^2(h)-1)=e+e(x_1^2-\frac{1}{2}x_2^2)+e^2(x_1^2-\frac{1}{2}x_2^2)^2.$$

Das Taylorpolynom der Ordnung 2 von diesem Polynom an der Stelle $u=0$ erhalten wir durch Weglassen der Terme höherer Ordnung:

$$T_0^2(f)(x)=T_0^2\big(T_0^2(g)\circ(T_1^2(h)-1)\big)(x)=e+e(x_1^2-\frac{1}{2}x_2^2).$$ ■

Bemerkung X.3.10. Sei $U \subseteq \mathbb{R}^n$ offen und $f \in C^k(U)$. Weiter sei $x + sh \in U$ für $s \in [0,1]$. Wir betrachten die Funktion $g\colon [0,1] \to \mathbb{R}^n, s \mapsto x + sh$. Für $k \geq 1$ und $\tau \in [0,1]$ gilt dann $T_\tau^k(g)(t) = g(\tau) + t \cdot h$, also $T_\tau^k(g)(t) - g(\tau) = t \cdot h$. Mit Satz X.3.8(c) und der Tatsache, dass $T_{g(\tau)}^k(f)(th)$ schon ein Polynom der Ordnung $\leq k$ in t ist, erhalten wir

$$T_\tau^k(f \circ g)(t) = T_0^k\big(T_{g(\tau)}^k(f)(th)\big) = T_{g(\tau)}^k(f)(th) = \sum_{|\alpha| \leq k} \frac{(D^\alpha f)(g(\tau))}{\alpha!} h^\alpha \cdot t^{|\alpha|}.$$

Aus

$$T_\tau^k(f \circ g)(t) = \sum_{m=0}^{k} \frac{(f \circ g)^{[m]}(\tau)}{m!} t^m$$

(der Formel für das Taylorpolynom von $f \circ g\colon D \to \mathbb{R}$) folgt daher durch Koeffizientenvergleich

$$\frac{(f \circ g)^{[m]}(\tau)}{m!} = \frac{1}{m!} \frac{d^m}{ds^m}\Big|_{s=\tau} f(x + sh) = \sum_{|\alpha| = m} \frac{\big(D^\alpha f\big)(x + \tau h)}{\alpha!} h^\alpha. \tag{3.3}$$

■

Nach diesen Vorbereitungen wenden wir uns der Taylorschen Formel zu, die angibt, wie gut eine Funktion durch ihr Taylorpolynom der Ordnung k approximiert wird. Die wesentliche Idee ist, dass wir die Taylorformel für eine Veränderliche auf die Verbindungsstrecke von x und $x + h$ anwenden.

Satz von Taylor

Satz X.3.11. *Sei $U \subseteq \mathbb{R}^n$ offen, $f \in C^{k+1}(U)$, und die Verbindungsstrecke $\{x + sh \mid 0 \leq s \leq 1\}$ sei in U enthalten. Dann existiert ein $\theta \in]0,1[$ mit*

$$f(x + h) = T_x^k(f)(h) + \sum_{|\alpha| = k+1} \frac{\big(D^\alpha f\big)(x + \theta h)}{\alpha!} h^\alpha.$$

Beweis. Sei $\varphi(s) := f(x + sh)$ für $0 \leq s \leq 1$. Dann besagt die Taylorformel mit der Restglieddarstellung nach Lagrange (VII.1.6)

$$\varphi(1) = \sum_{j=0}^{k} \frac{1}{j!} \varphi^{[j]}(0) + \frac{1}{(k+1)!} \varphi^{[k+1]}(\theta)$$

für ein $\theta \in]0,1[$. Setzen wir (3.3) hier ein, so ergibt sich

$$f(x + h) = \varphi(1) = \sum_{|\alpha| \leq k} \frac{\big(D^\alpha f\big)(x)}{\alpha!} h^\alpha + \sum_{|\alpha| = k+1} \frac{\big(D^\alpha f\big)(x + \theta h)}{\alpha!} h^\alpha.$$ ■

Satz X.3.12. (Restgliedabschätzung) *Für* $f \in C^{k+1}(U)$ *gilt*

$$f(x+h) = T_x^{k+1}(f)(h) + \varphi(h) \quad \text{mit} \quad \lim_{h\to 0} \frac{\varphi(h)}{\|h\|^{k+1}} = 0.$$

Beweis. Da U offen ist, dürfen wir annehmen, daß $U_\varepsilon(x) \subseteq U$ ist und $\|h\| < \varepsilon$. Nach Satz X.3.11 gilt für ein $\tau \in]0,1[$:

$$\varphi(h) := f(x+h) - T_x^{k+1}(f)(h) = f(x+h) - T_x^k(f)(h) - \sum_{|\alpha|=k+1} \frac{(D^\alpha f)(x)}{\alpha!} h^\alpha$$
$$= \sum_{|\alpha|=k+1} \frac{\big(D^\alpha f\big)(x+\tau h) - \big(D^\alpha f\big)(x)}{\alpha!} h^\alpha.$$

Wegen $|h_j| \leq \|h\| = \|h\|_2$ erhalten wir

$$|h^\alpha| = |h_1|^{\alpha_1} \cdot \ldots \cdot |h_n|^{\alpha_n} \leq \|h\|^{\alpha_1+\ldots+\alpha_n} = \|h\|^{|\alpha|} = \|h\|^{k+1}$$

und daher $\frac{|h^\alpha|}{\|h\|^{k+1}} \leq 1$. Der Ausdruck

$$\frac{\big(D^\alpha f\big)(x+\tau h) - \big(D^\alpha f\big)(x)}{\alpha!}$$

in der obigen Summe geht für $h \to 0$ gegen Null, da $f \in C^{k+1}(U)$ ist. Hiermit erhalten wir $\lim_{h\to 0} \frac{\varphi(h)}{\|h\|^{k+1}} = 0$. ■

Man beachte, dass diese Restgliedabschätzung die Approximation

$$f(x+h) = f(x) + \mathtt{d}f(x)(h) + \varphi(x) \quad \text{mit} \quad \frac{\varphi(h)}{\|h\|} \longrightarrow 0$$

verallgemeinert, die zur Differenzierbarkeit äquivalent ist. In der Tat erhalten wir für $k = 0$ das Taylorpolynom

$$T_x^1(f)(h) = f(x) + \mathtt{d}f(x)(h) = f(x) + \sum_{j=1}^n (D_i f)(x) h_i.$$

X.4. Das lokale Verhalten von Funktionen

In diesem Abschnitt werden wir Extrema von differenzierbaren reellwertigen Funktionen studieren, die auf offenen Teilmengen des $\mathbb{R}^n$ definiert sind. Wir werden hierbei sehen, dass das Verhalten der Funktion in Bezug auf Extrema im wesentlichen durch das Taylorpolynom zweiter Ordnung bestimmt wird.

Sei $U \subseteq \mathbb{R}^n$ eine offene Teilmenge und $f \in C^2(U,\mathbb{R})$. Dann ist das Taylorpolynom zweiter Ordnung in $u \in U$ gegeben durch

$$\begin{aligned} T_u^2(f)(x) &= f(u) + \sum_{i=1}^n \big(D_i f\big)(u)\cdot x_i + \frac{1}{2}\sum_{i,j=1}^n \big(D_i D_j f\big)(u) x_i x_j \\ &= f(u) + \sum_{i=1}^n \frac{\partial f}{\partial x_i}(u)\cdot x_i + \frac{1}{2}\sum_{i,j=1}^n \frac{\partial^2 f}{\partial x_i \partial x_j}(u) x_i x_j. \end{aligned}$$

Hierbei beachten wir, dass jeder Multiindex α vom Grad 1 die Gestalt

$$(0,\ldots,0,1,0,\ldots,0)$$

(eine 1 an der i-ten Stelle) besitzt. Multiindizes vom Grad haben die Gestalt

$$(0,\ldots,0,2,0,\ldots,0) \quad \text{oder} \quad (0,\ldots,0,1,0,\ldots,0,1,0,\ldots,0).$$

In diesem Abschnitt werden wir das Taylorpolynom der Ordnung 2 einer Funktion verwenden, um ihr lokales Verhalten zu beschreiben.

Definition X.4.1. Ist $f \in C^2(U,\mathbb{R})$ und $u \in U$, so heißt

$$H_u(f) := \left(\frac{\partial^2 f}{\partial x_i \partial x_j}(u)\right)_{i,j=1,\ldots,n}$$

die *Hessematrix* von f in u. Nach dem Satz von Schwarz ist $H_u(f)$ eine *symmetrische* Matrix. Die Abbildung

$$\widetilde{H}_u(f) : \mathbb{R}^n \to \mathbb{R},\ x \mapsto \langle H_u(f)x, x\rangle = \sum_{i,j=1}^n \frac{\partial^2 f}{\partial x_i \partial x_j}(u) x_i x_j$$

heißt *Hesseform* von f in u. Sie ist eine *quadratische Form* auf $\mathbb{R}^n$. ∎

Mit der obigen Definition gilt für das Taylorpolynom zweiter Ordnung einer Funktion $f \in C^2(U,\mathbb{R})$

$$T_u^2(f)(h) = f(u) + \mathtt{d}f(u)(h) + \tfrac{1}{2}\widetilde{H}_u(f)(h) = f(u) + J_u(f)h + \frac{1}{2}h^\top H_u(f)h$$

und $f(u+h) = T_u^2(f)(h) + \varphi(h)$ mit $\frac{|\varphi(h)|}{\|h\|^2} \longrightarrow 0$ für $h \to 0$ (vgl. X.3.11).

Definition X.4.2. Sei $U \subseteq \mathbb{R}^n$ offen und $f \in C^1(U)$. Ein Punkt $u \in U$ heißt *kritischer Punkt*, wenn $\mathtt{d}f(u) = 0$ ist. In diesem Fall heißt $f(u)$ *kritischer Wert* von f.

Wir beachten, dass u genau dann kritischer Punkt ist, wenn alle partiellen Ableitungen von f in u verschwinden, d.h.

$$D_1(f)(u) = \ldots = D_n(f)(u) = 0$$

gilt. ∎

Definition X.4.3. Sei $f\colon U \to \mathbb{R}$ differenzierbar.
(a) Ein Punkt $u \in U$ heißt ein (*isoliertes*) *lokales Maximum* von f, wenn ein $\varepsilon > 0$ so existiert, dass $U_\varepsilon(u) \subseteq U$ und $f(u+h) \leq f(u)$ $(f(u+h) < f(u))$ für alle h mit $0 \neq \|h\| < \varepsilon$ gilt. Der Begriff des (*isolierten*) *lokalen Minimums* wird analog definiert. Die Funktion f hat in u ein *lokales Extremum*, wenn sie in u ein lokales Maximum oder ein lokales Minimum hat.
(b) Der Punkt u heißt *globales Maximum* bzw. *globales Minimum*, falls für alle $v \in U$ gilt $f(v) \leq f(u)$ (bzw. $f(v) \geq f(u)$). ■

Notwendige Bedingung für Extrema

Lemma X.4.4. *Hat $f \in C^1(U)$ in u ein lokales Extremum, so ist u ein kritischer Punkt.*

Beweis. Sei $u \in U$ ein lokales Extremum von f und $v \in \mathbb{R}^n$. Wir haben zu zeigen, dass $\mathtt{d}f(u)(v) = 0$ gilt. Hierzu wählen wir $\delta > 0$ so klein, dass $u+tv \in U$ für $|t| \leq \delta$ gilt (die Existenz folgt aus der Offenheit von U). Wir betrachten nun die Funktion

$$\varphi_v : [-\delta, \delta] \to \mathbb{R}, \quad t \mapsto f(u+tv).$$

Diese Funktion hat im Nullpunkt ein lokales Extremum und daher ist

$$0 = \varphi_v'(0) = \mathtt{d}f(u)(v).$$

■

Beispiel X.4.5. (a) Sei $f : \mathbb{R}^2 \to \mathbb{R}$, $f(x,y) = x^2 + y^2$. Dann ist

$$\nabla f(x,y) = (2x, 2y) = 0$$

genau dann, wenn $(x,y) = (0,0)$ ist. Daher ist der Nullpunkt der einzige kritische Punkt, und es liegt dort ein globales Minimum vor.
(b) Genauso hat die Funktion $f(x,y) = -x^2 - y^2$ im Nullpunkt ein globales Maximum.
(c) Die Funktion $f : \mathbb{R}^2 \to \mathbb{R}$, $(x,y) \mapsto x^2 - y^2$ hat zwar in $(0,0)$ einen kritischen Punkt, aber trotzdem kein Extremum. Der Nullpunkt ist ein sogenannter *Sattelpunkt*. ■

Definition X.4.6. (Wiederholung aus der linearen Algebra) Sei A eine symmetrische $(n \times n)$-Matrix.

(a) A heißt *positiv definit*, wenn $\langle Ax, x\rangle > 0$ für alle $x \neq 0$ gilt.
(b) A heißt *positiv semidefinit*, wenn $\langle Ax, x\rangle \geq 0$ für alle $x \in \mathbb{R}^n$ gilt.
(c) A heißt *negativ* (*semi-*)*definit*, wenn $-A$ *positiv* (*semi-*)*definit* ist.
(d) A heißt *indefinit*, wenn es $x, y \in \mathbb{R}^n$ gibt, so dass $\langle Ax, x\rangle > 0$ und $\langle Ay, y\rangle < 0$ gelten. ■

HAUPTACHSENTRANSFORMATION

Bemerkung X.4.7. (a) Zu jeder symmetrischen Matrix A existiert eine Orthonormalbasis $v_1, \ldots, v_n$ von $\mathbb{R}^n$ aus Eigenvektoren von A, d.h., dass für jedes $j = 1, \ldots, n$ ein $\lambda_j \in \mathbb{R}$ mit $Av_j = \lambda_j v_j$ existiert. Für $x = \sum_{j=1}^n x_j v_j$ gilt dann

$$\langle Ax, x \rangle = \sum_{j=1}^{n} x_j^2 \lambda_j.$$

Daraus folgt, dass A genau dann positiv (semi-)definit ist, wenn alle Eigenwerte $\lambda_j > 0$ $(\lambda_j \geq 0)$ sind. Die Matrix ist genau dann indefinit, wenn es sowohl positive als auch negative Eigenwerte gibt.
(b) Ist A eine positiv definite symmetrische Matrix und $B \in \mathrm{GL}_n(\mathbb{R})$ eine invertierbare Matrix, so ist A genau dann positiv definit, wenn $B^\top AB$ positiv definit ist. Für $v \in \mathbb{R}^n$ haben wir nämlich

$$\langle B^\top ABv, v \rangle = \langle ABv, Bv \rangle,$$

und da Multiplikation mit B bijektiv ist, ist dieser Ausdruck genau dann für all $0 \neq v \in \mathbb{R}^n$ positiv, wenn $\langle A.w, w \rangle > 0$ für alle $0 \neq w \in \mathbb{R}^n$ gilt. Also ist A genau dann positiv definit, wenn dies für $B^\top AB$ der Fall ist. ■

HURWITZKRITERIUM

Satz X.4.8. [1] *Eine symmetrische $n \times n$-Matrix A ist genau dann positiv definit, wenn alle Hauptminoren positiv sind, d.h. wenn für alle $k \in \{1, \ldots, n\}$ gilt*

$$\det \begin{pmatrix} a_{11} & \ldots & a_{1k} \\ \vdots & & \vdots \\ a_{k1} & \ldots & a_{kk} \end{pmatrix} > 0.$$

Beweis. Für $k \leq n$ setzen wir

$$A_k := \begin{pmatrix} a_{11} & \ldots & a_{1k} \\ \vdots & & \vdots \\ a_{k1} & \ldots & a_{kk} \end{pmatrix}.$$

Notwendigkeit der Bedingung: Ist A positiv definit, so sind auch alle Matrizen A_k positiv definit. In der Tat, für $0 \neq x \in \mathbb{R}^k$ sei $\widetilde{x} = (x_1, \ldots, x_k, 0, \ldots, 0)$ der zugehörige Vektor im $\mathbb{R}^n$. Dann ist

$$\langle A_k x, x \rangle = \langle A\widetilde{x}, \widetilde{x} \rangle > 0.$$

[1] Adolf Hurwitz (1859–1919), deutscher Mathematiker. Er studierte bei Felix Klein in München und bei Kummer, Kronecker und Weierstraß in Berlin. Professor in Königsberg und Zürich. Er beschäftigte sich vor allem mit Zahlentheorie, aber auch mit Funktionentheorie, wo er das Geschlecht von Riemannschen Flächen untersuchte. Nach ihm sind z.B. das Hurwitzpolynom und das Hurwitzkriterium aus der Stabilitätstheorie dynamischer Systeme benannt; Satz X.4.8 ist eine Variation davon.

Insbesondere sind dann alle Eigenwerte von A_k positiv, also auch $\det A_k > 0$.
Die Bedingung ist hinreichend: Wir zeigen dies durch Induktion nach n. Für $n = 1$ ist die Behauptung trivial. Sei nun $n > 1$. Nach Induktionsvoraussetzung ist die Matrix A_{n-1} dann positiv definit. Also existiert eine orthogonale $(n-1) \times (n-1)$-Matrix S (d.h. $SS^\top = S^\top S = \mathbf{1}$) mit

$$S^\top A_{n-1} S = \begin{pmatrix} \alpha_1 & \dots & 0 \\ \vdots & & \vdots \\ 0 & \dots & \alpha_{n-1} \end{pmatrix}, \quad \alpha_1, \dots, \alpha_{n-1} > 0.$$

Sei $\widetilde{S}$ die $n \times n$-Matrix, die durch

$$\widetilde{S} = \begin{pmatrix} S & 0 \\ 0 & 1 \end{pmatrix}$$

gegeben ist. Dann ist $\widetilde{S}$ ebenfalls orthogonal und wir erhalten

$$B := \widetilde{S}^\top A \widetilde{S} = \begin{pmatrix} \alpha_1 & \dots & 0 & b_1 \\ \vdots & & \vdots & \vdots \\ 0 & \dots & \alpha_{n-1} & b_{n-1} \\ b_1 & \dots & b_{n-1} & b_n \end{pmatrix}.$$

Nach Voraussetzung ist $\det B = \det A > 0$. Wir setzen

$$T := \begin{pmatrix} 1 & \dots & 0 & c_1 \\ \vdots & & \vdots & \vdots \\ 0 & \dots & 1 & c_{n-1} \\ 0 & \dots & 0 & 1 \end{pmatrix} \quad \text{mit} \quad c_j := -\frac{b_j}{\alpha_j}$$

und erhalten

$$C := T^\top B T = \begin{pmatrix} \alpha_1 & \dots & 0 & 0 \\ \vdots & & \vdots & \vdots \\ 0 & \dots & \alpha_{n-1} & 0 \\ 0 & \dots & 0 & \alpha_n \end{pmatrix} \quad \text{mit} \quad \alpha_n = b_n - \frac{b_1^2}{\alpha_1} - \dots - \frac{b_{n-1}^2}{\alpha_{n-1}}.$$

Wegen $\det T = 1$ ist die Determinante dieser Matrix $\alpha_1 \cdots \alpha_n$ positiv und somit $\alpha_n > 0$. Daher ist die Matrix $C = (\widetilde{S}T)^\top A \widetilde{S} T$ positiv definit und somit auch A (Bemerkung X.4.7(b)). ∎

Das Hurwitzkriterium lässt sich nicht analog zu einem Kriterium für die positive Semidefinitheit verallgemeinern. Für die Matrix

$$A = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

gilt

$$\det(0) \geq 0, \quad \det \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \geq 0 \quad \text{und} \quad \det A \geq 0,$$

aber A ist nicht positiv semidefinit.

Der Vorteil des Hurwitzkriteriums ist, dass man die Eigenwerte nicht kennen muss, um auszurechnen, ob eine Matrix positiv oder negativ definit ist.

Folgerung X.4.9. *Sei $A = (a_{ij})$ eine symmetrische (2×2)-Matrix.*

(1) *Ist* $\det A < 0$*, so ist A indefinit.*

(2) *Ist* $\det A > 0$ *und* $a_{11} > 0$*, so ist A positiv definit.*

(3) *Ist* $\det A > 0$ *und* $a_{11} < 0$*, so ist A negativ definit.*

Beweis. (1) Ist $\det(A) < 0$, so haben die beiden Eigenwerte von A verschiedene Vorzeichen. Also ist A indefinit.

(2), (3) folgen direkt aus dem Hurwitzkriterium. ∎

Hinreichende Bedingung für Extrema

Satz X.4.10. *Sei $U \subseteq \mathbb{R}^n$ eine offene Menge, $f \in C^2(U, \mathbb{R})$ und $x \in U$ ein kritischer Punkt von f. Dann gilt:*

(a) *Ist die Hessematrix $H_x(f)$ positiv definit, so ist x ein isoliertes lokales Minimum.*

(b) *Ist $H_x(f)$ negativ definit, so ist x ein isoliertes lokales Maximum.*

(c) *Ist $H_x(f)$ indefinit, so handelt es sich bei x nicht um einen Extremalpunkt.*

Kritische Punkte, die keine lokalen Extrema sind, nennt man *Sattelpunkte.*

Beweis. Nach Definition X.4.1 gilt

$$\begin{aligned} f(x+h) &= f(x) + \mathtt{d}f(x)(h) + \frac{1}{2}\langle H_x(f)(h), h\rangle + \varphi(h) \\ &= f(x) + \frac{1}{2}\langle H_x(f)(h), h\rangle + \varphi(h) \end{aligned}$$

mit $\lim_{h\to 0} \frac{\varphi(h)}{\|h\|^2} = 0$. Zu jedem $\varepsilon > 0$ existiert also ein $\delta > 0$ mit $U_\delta(x) \subseteq U$ und $|\varphi(h)| \leq \varepsilon \cdot \|h\|^2$ für alle h mit $\|h\| < \delta$.

(a) Sei $A := \frac{1}{2}H_x(f)$ positiv definit und $\lambda_1 > 0$ der kleinste Eigenwert von A. Ist $v_1, \ldots, v_n$ eine Orthonormalbasis aus Eigenvektoren und $h = \sum_{j=1}^n h_j v_j$, so ist

$$\langle Ah, h\rangle = \sum_{j=1}^n |h_j|^2 \lambda_j \geq \lambda_1 \|h\|^2$$

für alle $h \in \mathbb{R}^n$. Sei nun $\varepsilon := \frac{\lambda_1}{2}$ und δ wie oben. Für $\|h\| < \delta$ gilt dann

$$\begin{aligned} f(x+h) &= f(x) + \langle Ah, h\rangle + \varphi(h) \\ &\geq f(x) + \lambda_1\|h\|^2 - \frac{\lambda_1}{2}\|h\|^2 = f(x) + \frac{\lambda_1}{2}\|h\|^2. \end{aligned}$$

Also ist x ein isoliertes lokales Minimum.

(b) Wende (a) auf $-f$ an.

(c) Ist $v \in \mathbb{R}^n$ ein Vektor mit $\langle Av, v\rangle > 0$, so ist

$$f(x+tv) = f(x) + t^2\langle Av, v\rangle + \varphi(tv) \quad \text{mit} \quad \frac{\varphi(tv)}{t^2} \xrightarrow[t\to 0]{} 0.$$

Ist t so klein, dass $\frac{\varphi(tv)}{t^2} > -\langle Av, v\rangle$ gilt, so ist $f(x+tv) > f(x)$. Analog zeigt man für einen Vektor w mit $\langle Aw, w\rangle < 0$ die Existenz eines $\delta > 0$ mit $f(x+tw) < f(x)$ für $|t| < \delta$. Folglich ist x ein Sattelpunkt. ∎

Aus dem Beweis von (c) erhalten wir eine wichtige Folgerung, die wir als notwendige Bedingung für Extrema verstehen können:

Folgerung X.4.11. *Sei $U \subseteq \mathbb{R}^n$ eine offene Menge, $f \in C^2(U, \mathbb{R})$ und $x \in U$ ein kritischer Punkt von f. Dann gilt:*

(a) *Ist x ein lokales Minimum, so ist $H_x(f)$ positiv semidefinit.*

(b) *Ist x ein lokales Maximum, so ist $H_x(f)$ negativ semidefinit.*

Beweis. (a) Ist $H_x(f)$ nicht positiv semidefinit und $\langle H_x(f)v, v\rangle < 0$, so folgt aus dem Beweis von Satz X.4.10(c) die Existenz eines $\delta > 0$, so dass $f(x + tv) < f(x)$ für alle t mit $|t| < \delta$ gilt. Also kann x kein lokales Minimum sein.
(b) Wir wenden (a) auf $-f$ an. ∎

Bemerkung X.4.12. Ist die Hessematrix semidefinit, so lassen sich keine allgemeinen Aussagen machen. Die Funktionen $f_j\colon \mathbb{R}^2 \to \mathbb{R}$:

$$f_1(x, y) = x^2 + y^4, \quad f_2(x, y) = x^2 \quad \text{und} \quad f_3(x, y) = x^2 + y^3$$

besitzen alle im Nullpunkt die Hessematrix

$$H_0(f_1) = H_0(f_2) = H_0(f_3) = \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix}.$$

Die Funktion f_1 hat im Nullpunkt ein isoliertes Minimum; f_2 hat dort ein (nicht-isoliertes) Minimum, und f_3 besitzt kein Extremum. ∎

Beispiel X.4.13. Wir wollen für die Funktion

$$f : \mathbb{R}^2 \to \mathbb{R}, \ f(x, y) = \big(3x + 4y + \sin(xy)\big)\big(2x + y - (\cos x)(1 - \cos y)\big)$$

das lokale Verhalten im Nullpunkt bestimmen. Dazu berechnen wir ihr Taylorpolynom der Ordnung 2 (siehe die Bemerkung unten):

$$\begin{aligned} T_0^2(f)(x, y) &= T_0^2\Big(\big(3x + 4y + xy\big)\big(2x + y - (1 - \frac{x^2}{2})(1 - (1 - \frac{y^2}{2}))\big)\Big) \\ &= (3x + 4y)(2x + y) = 6x^2 + 8xy + 3xy + 4y^2 = 6x^2 + 11xy + 4y^2, \end{aligned}$$

d.h., die Hessematrix ergibt sich zu

$$H_0(f) = \begin{pmatrix} 12 & 11 \\ 11 & 8 \end{pmatrix}.$$

Daher ist der Nullpunkt wegen $12 > 0$ und $\det H_0(f) = 96 - 121 < 0$ ein Sattelpunkt.

Bemerkung zur Berechnung der Taylorpolynome: Zunächst wissen wir, dass die Funktion sin sich auf $\mathbb{R}$ durch eine Potenzreihe darstellen lässt:

$$\sin x = \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k+1)!} x^{2k+1}.$$

Da sich konvergente Potenzreihen gliedweise differenzieren lassen, stimmt diese Reihe mit der Taylorreihe $T_0(\sin)$ überein.

Betrachten wir nun die Funktion $f\colon \mathbb{R}^2 \to \mathbb{R}, f(x,y) = \sin(xy)$, so ist f beliebig oft partiell differenzierbar und hat die Reihenentwicklung

$$(\dagger) \qquad f(x,y) = \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k+1)!} x^{2k+1} y^{2k+1}.$$

Halten wir jeweils x oder y fest, so ergibt sich eine überall konvergente Potenzreihe in einer Variablen, die wir gemäß Satz VI.5.3 gliedweise differenzieren dürfen. Wir erhalten daher sukzessive für die partiellen Ableitungen:

$$D_1^m D_2^k f(x,y) = \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k+1)!} D_1^m(x^{2k+1}) D_2^k(y^{2k+1}).$$

Hieraus erkennen wir insbesondere, dass ($\dagger$) mit der Taylorreihe der Funktion f übereinstimmt. Insbesondere ist

$$T_0^2(f)(x,y) = xy,$$

was wir oben verwendet haben. Alternativ kann man mit der Kettenregel für Taylorpolynome argumentieren (Satz X.3.8(c)). ■

X.5. Vertauschbarkeit von Ableitung und Integral

In diesem kurzen Abschnitt lernen wir eine wichtige Rechenmethode kennen, die das Vertauschen von Ableiten und Integration betrifft, sofern verschiedene Variablen betroffen sind.

Satz X.5.1. *Sei $U \subseteq \mathbb{R}^n$ eine offene Menge und $D = [a,b] \subseteq \mathbb{R}$ ein Intervall. Die Funktion $f : D \times U \to \mathbb{R}$ sei stetig. Dann ist die Funktion*

$$F : U \to \mathbb{R}, \ x \mapsto \int_a^b f(t,x)\,dt$$

stetig. Hat f stetige partielle Ableitungen $\frac{\partial f}{\partial x_i} : D \times U \to \mathbb{R}$, $i = 1, \ldots, n$, so ist auch F stetig nach x_i differenzierbar, und es gilt

$$\frac{\partial}{\partial x_i} \int_a^b f(t,x)\,dt = \int_a^b \frac{\partial f}{\partial x_i}(t,x)\,dt.$$

Beweis. (a) Sei $x \in U$. Da U offen ist, existiert ein $r > 0$ mit $U_{2r}(x) \subseteq U$. Die Menge

$$D \times \overline{U_r(x)} = D \times \{y \in \mathbb{R}^n : \|y - x\| \leq r\} \subseteq \mathbb{R}^{n+1}$$

ist nach dem Satz von Heine–Borel kompakt, denn sie ist abgeschlossen und beschränkt, also ist f auf $D \times \overline{U_r(x)}$ gleichmäßig stetig (Satz IX.3.19). Zu jedem $\varepsilon > 0$ existiert also ein $\delta \in]0, r[$ mit

$$|f(t, x + h) - f(t, x)| < \frac{\varepsilon}{b - a} \quad \text{für alle } h \text{ mit } \|h\| < \delta.$$

Es folgt

$$|F(x + h) - F(x)| \leq \int_a^b |f(t, x + h) - f(t, x)| \, dt \leq \frac{\varepsilon}{b - a}(b - a) = \varepsilon$$

für alle h mit $\|h\| < \delta$, d.h., F ist stetig.
(b) Für $x \in U$ und $h \in \mathbb{R}$ mit $x + he_i \in U$ definieren wir

$$g(t, x, h) := \begin{cases} \frac{f(t,x+he_i)-f(t,x)}{h}, & \text{falls } h \neq 0 \\ \frac{\partial f}{\partial x_i}(t, x), & \text{falls } h = 0. \end{cases}$$

Wir behaupten, dass diese Funktion auf der Menge

$$\widetilde{U} := \{(t, x, h) \in D \times U \times \mathbb{R} \colon x + he_i \subseteq U\}$$

stetig ist. In allen Punkten (x, y, h) mit $h \neq 0$ ist dies klar. Wir müssen die Stetigkeit also nur in den Punkten der Gestalt $(x, y, 0)$ nachweisen. Nach dem Mittelwertsatz der Differentialrechnung ist

$$\frac{f(t, x + he_i) - f(t, x)}{h} = \frac{\partial f}{\partial x_i}(t, x + \vartheta_h he_i)$$

für ein $\vartheta_h \in]0, 1[$, falls $x + [0, 1]he_i \subseteq U$ ist. Gilt $(t_n, x_n, h_n) \to (t_0, x_0, 0)$, so auch $\vartheta_{h_n} h_n \longrightarrow 0$. Für ausreichend große n ist $x_n + [0, 1]h_n \subseteq U$. Für solche n erhalten wir daher

$$\begin{aligned} g(t_n, x_n, h_n) &= \frac{f(t_n, x_n + h_n e_i) - f(t_n, x_n)}{h} = \frac{\partial f}{\partial x_i}(t_n, x_n + \vartheta_n h_n e_i) \\ &\longrightarrow \frac{\partial f}{\partial x_i}(t_0, x_0) = g(t_0, x_0, 0), \end{aligned}$$

d.h., die Funktion g ist stetig. Es gilt also wegen (a)

$$\frac{\partial F}{\partial x_i}(x) = \lim_{\substack{h \to 0 \\ h \neq 0}} \int_a^b g(t, x, h) \, dt \overset{(a)}{=} \int_a^b g(t, x, 0) \, dt = \int_a^b \frac{\partial f}{\partial x_i}(t, x) \, dt,$$

und diese Funktion hängt nach dem ersten Teil stetig von x ab. ■

Bemerkung X.5.2. Hat f stetige partielle Ableitungen der Ordnung $\leq k$, so sieht man induktiv

$$\big(D^\alpha F\big)(x) = \int_a^b \big(D^{(0,\alpha)} f\big)(t, x) \, dt$$

für alle $|\alpha| \leq k$, wobei $(0, \alpha) = (0, \alpha_1, \ldots, \alpha_n) \in \mathbb{N}_0^{n+1}$ ist. ■

X.6. Kurvenintegrale und Pfaffsche Formen

In diesem Abschnitt lernen wir eine neue Variation zu dem Thema Kurvenintegrale kennen. Die Kurvenintegrale, die wir in Abschnitt X.1 behandelt haben, waren von dem Typ

$$\int_\gamma f = \int_a^b f(\gamma(t))\|\dot\gamma(t)\|\,dt,$$

wobei $\gamma\colon [a,b] \to \mathbb{R}^n$ eine Kurve war. Der Nachteil dieser Kurvenintegrale ist, dass hier der Geschwindigkeitsvektor der Kurve nur als Skalar eingeht und nicht als vektorielle Größe. Will man zum Beispiel die Arbeit modellieren, die man entlang eines Weges in einem Kraftfeld verrichtet, so sollte das Ergebnis eine skalare Größe sein, auch wenn das Kraftfeld eine vektorielle Funktion ist. Den angemessenen Rahmen für solche Integrale bildet der Kalkül der Differentialformen. Die Differentialformen, die bei den Kurvenintegralen auftreten, nennt man Pfaffsche Formen bzw. Differentialformen erster Ordnung oder 1-Formen. Mit ihnen lässt sich zum Beispiel die Frage danach, ob ein gegebenes Vektorfeld ein Gradient einer Funktion ist, in der ihr angemessenen Allgemeinheit diskutieren. In der Sprache der Pfaffschen Formen ist es die Frage nach der Existenz einer Stammfunktion einer gegebenen Pfaffschen Form. Diese wiederum lässt sich durch das Verschwinden von Integralen der Pfaffschen Form entlang geschlossener Wege charakterisieren.

Definition X.6.1. Sei $U \subseteq \mathbb{R}^n$ offen.
(a) Eine *Pfaffsche Form auf U* ist eine Funktion

$$\omega\colon U \to \mathrm{Hom}(\mathbb{R}^n, \mathbb{R}),$$

d.h., für jedes $x \in U$ ist $\omega(x)\colon \mathbb{R}^n \to \mathbb{R}$ eine lineare Abbildung. Wir schreiben $\Omega^1(U)$ für den Raum der Pfaffschen Formen auf U.

Um die $(1 \times n)$-Matrix zu erhalten, die die lineare Abbildung $\omega(p)$ bzgl. der kanonischen Basis des $\mathbb{R}^n$ darstellt, betrachten wir die Funktionen

$$f_j\colon U \to \mathbb{R}, \qquad f_j(p) := \omega(p)(e_j)$$

und erhalten die darstellende Zeilenmatrix $\begin{pmatrix} f_1(p), & \cdots, & f_n(p) \end{pmatrix}$.
(b) Jeder differenzierbaren Funktion $F\colon U \to \mathbb{R}$ können wir eine Pfaffsche Form zuordnen, denn die Ableitung

$$\mathtt{d}F\colon U \to \mathrm{Hom}(\mathbb{R}^n, \mathbb{R})$$

ordnet jedem Punkt $x \in U$ eine lineare Abbildung $\mathtt{d}F(x)\colon \mathbb{R}^n \to \mathbb{R}$ zu. Die Pfaffsche Form $\mathtt{d}F$ nennen wir das *totale Differential* von F.

(c) Ist $\gamma\colon [a,b] \to U$ eine stetig differenzierbare Kurve und ω auf U eine stetige Pfaffsche Form mit den Komponentenfunktionen $f_j(x) := \omega(x)(e_j)$, so definieren wir das *Integral von ω über γ* wir folgt:

$$\int_\gamma \omega := \int_a^b \omega\big(\gamma(t)\big(\dot\gamma(t)\big)\, dt = \int_a^b \sum_{j=1}^n f_j\big(\gamma(t)\big)\gamma_j'(t)\, dt.$$

Ist γ lediglich stückweise stetig differenzierbar und

$$a = x_0 < x_1 < \ldots < x_m = b$$

eine Unterteilung, so dass die Wege

$$\gamma_j := \gamma|_{[x_j,x_{j+1}]}\colon [x_j, x_{j+1}] \to U$$

stetig differenzierbar sind, so definieren wir

$$\int_\gamma \omega := \sum_{j=0}^{m-1} \int_{\gamma_j} \omega.$$ ■

Möchte man konkret mit Pfaffschen Formen rechnen, so erweist es sich als praktisch, sie in den natürlichen Koordinaten des $\mathbb{R}^n$ zu beschreiben. Für jedes $j \in \{1,\ldots,n\}$ betrachten wir dazu die Funktion

$$x_j\colon \mathbb{R}^n \to \mathbb{R}, \quad p = (p_1,\ldots,p_n) \mapsto p_j.$$

Da x_j stetig differenzierbar ist, erhalten wir Pfaffsche Formen $\mathtt{d}x_j \in \Omega^1(\mathbb{R}^n)$, $j = 1,\ldots,n$. Für jedes $p \in \mathbb{R}^n$ ist dann

$$\mathtt{d}x_j(p)(h) = h_j,$$

d.h., $\mathtt{d}x_j\colon \mathbb{R}^n \to \mathrm{Hom}(\mathbb{R}^n,\mathbb{R})$ ist die konstante Abbildung, die jedem p die lineare Abbildung x_j zuordnet. In diesem Sinn haben wir $\mathtt{d}x_j(p) = x_j$ für alle $p \in \mathbb{R}^n$. Es ist klar, dass die linearen Abbildungen $x_1,\ldots,x_n$ eine Basis des n-dimensionalen Vektorraums $\mathrm{Hom}(\mathbb{R}^n,\mathbb{R})$ bilden.

Ist nun $\omega \in \Omega^1(U)$ eine Pfaffsche Form und definieren wir die Funktionen $f_j\colon U \to \mathbb{R}$ durch $f_j(p) := \omega(p)(e_j)$, so erhalten wir für jedes $p \in U$:

$$\omega(p) = \sum_{j=1}^n f_j(p) \cdot \mathtt{d}x_j(p),$$

also

$$\omega = \sum_{j=1}^n f_j \mathtt{d}x_j. \tag{6.1}$$

Wir können daher jede Pfaffsche Form wie in (6.1) darstellen, und diese Darstellung ist eindeutig.

Ist $\omega = \mathtt{d}F$ das totale Differential einer stetig differenzierbaren Funktion, so erhalten wir insbesondere

$$\mathtt{d}F(x) = \sum_{j=1}^n \frac{\partial F}{\partial x_j}\ \mathtt{d}x_j \tag{6.2}$$

(man vergleiche dies mit der Kettenregel).

Verhalten bei Parametertransformation

Natürlich müssen wir uns überlegen, in welcher Form das Integral einer Pfaffschen Form über eine Kurve von ihrer Parametrisierung abhängt.

Sei $\omega = \sum_j f_j \mathtt{d}x_j$ eine stetige Pfaffsche Form in der offenen Menge $U \subseteq \mathbb{R}^n$ und $\gamma: [a,b] \to U$ eine stetig differenzierbare Kurve. Weiter sei $\varphi: [c,d] \to [a,b]$ stetig differenzierbar mit $\varphi(c) = a$ und $\varphi(d) = b$. Dann ist $\gamma \circ \varphi: [c,d] \to U$ ebenfalls eine stetig differenzierbare Kurve mit dem gleichen Bild sowie dem gleichen Anfangs- und Endpunkt wie γ.

Lemma X.6.2. $\int_\gamma \omega = \int_{\gamma\circ\varphi} \omega$ (Parametrisierungsinvarianz).

Beweis. Nach der Kettenregel gilt $(\gamma \circ \varphi)'(t) = \varphi'(t)\gamma'\big(\varphi(t)\big)$. Aus der Transformationsformel für eindimensionale Integrale (Substitutionsregel) ergibt sich daher

$$\begin{aligned}\int_{\gamma\circ\varphi} \omega &= \int_c^d \omega\big(\gamma \circ \varphi(t)\big)\big((\gamma \circ \varphi)'(t)\big)\,dt \\ &= \int_c^d \omega\big(\gamma\big(\varphi(t)\big)\big)\big(\gamma'\big(\varphi(t)\big)\big)\varphi'(t)\,dt \\ &= \int_a^b \omega\big(\gamma(s)\big)\big(\gamma'(s)\big)\,ds = \int_\gamma \omega.\end{aligned}$$

■

Bemerkung X.6.3. Analog zeigt man, dass für eine stetig differenzierbare Bijektion $\varphi: [c,d] \to [a,b]$ mit $\varphi(c) = b$ und $\varphi(d) = a$ die Beziehung

$$\int_{\gamma\circ\varphi} \omega = -\int_\gamma \omega$$

gilt.

Wir stellen hier ein interessantes Phänomen fest: durch die Umkehrung der Orientierung ändert das Kurvenintegral sein Vorzeichen. Anschaulich bedeutet dies, dass beim Durchlaufen der Kurve in umgekehrter Richtung das Kurvenintegral sein Vorzeichen umkehrt. ■

Wir berechnen nun das Kurvenintegral eines totalen Differentials.

Satz X.6.4. *Sei $U \subseteq \mathbb{R}^n$ offen und $F: U \to \mathbb{R}$ stetig differenzierbar sowie $\gamma: [a,b] \to U$ eine stückweise stetig differenzierbare Kurve. Dann gilt*

$$\int_\gamma \mathtt{d}F = F(\gamma(b)) - F(\gamma(a)).$$

Beweis. Wir nehmen zuerst an, dass γ stetig differenzierbar ist. Aus der Kettenregel und dem Hauptsatz der Differential- und Integralrechnung folgt

$$\int_\gamma \mathtt{d}F = \int_a^b \mathtt{d}F\big(\gamma(t)\big)\big(\gamma'(t)\big)\,dt = \int_a^b (F\circ\gamma)'(t)\,dt$$
$$= (F\circ\gamma)(b) - (F\circ\gamma)(a) = F(\gamma(b)) - F(\gamma(a)).$$

Der allgemeine Fall eines stückweise stetig differenzierbaren Weges ergibt sich nun durch Zusammensetzen der einzelnen Integrale, was zu einer Teleskopsumme führt. ■

Satz X.6.4 besagt insbesondere, dass das Integral eines totalen Differentials NUR von Anfangs- und Endpunkt des Weges abhängt. Ist γ ein geschlossener Weg, d.h. gilt $\gamma(a) = \gamma(b)$, so erhalten wir insbesondere

$$\int_\gamma \mathtt{d}F = 0.$$

Beispiel X.6.5. In $U := \mathbb{R}^2 \setminus \{0\}$ betrachten wir die Pfaffsche Form

$$\omega(x,y) = -\frac{y}{x^2+y^2}\mathtt{d}x + \frac{x}{x^2+y^2}\mathtt{d}y$$

und den geschlossenen Weg

$$\gamma\colon [0,2\pi] \to U, \quad t \mapsto (\cos t, \sin t).$$

Dann ist $\gamma'(t) = (-\sin t, \cos t)$ und daher

$$\int_\gamma \omega = \int_0^{2\pi} \frac{(\sin t)^2 + (\cos t)^2}{(\sin t)^2 + (\cos t)^2}\,dt = \int_0^{2\pi} 1\,dt = 2\pi.$$

Wir sehen insbesondere, dass ω kein totales Differential sein kann, da das Integral über die geschlossene Kurve γ nicht verschwindet. ■

Ein Einschub über Zusammenhang

In diesem Unterabschnitt werden wir kurz einige Aspekte des Zusammenhangsbegriffs diskutieren.

Definition X.6.6. (a) Ein metrischer Raum (U,d) heißt *zusammenhängend*, wenn er nicht disjunkte Vereinigung von zwei offenen nichtleeren Teilmengen ist. D.h., sind $U_1, U_2 \subseteq U$ offen mit

$$U = U_1 \cup U_2 \quad \text{und} \quad U_1 \cap U_2 = \emptyset, \tag{6.3}$$

so ist eine der Mengen U_j leer.

Ist eine Zerlegung wie in (6.3) gegeben, so ist ist $U_1^c = U \setminus U_1 = U_2$ offen, also ist U_1 auch abgeschlossen. Man überlegt sich leicht, dass die Zerlegungen wie in (6.3) genau denjenigen Mengen entsprechen, die gleichzeitig offen und abgeschlossen sind. Triviale Fälle erhalten wir für $U_1 = \emptyset$ oder $U_1 = U$. Der Raum U ist also genau dann zusammenhängend, wenn $\emptyset$ und X die einzigen Teilmengen sind, die gleichzeitig offen und abgeschlossen sind.

(b) Man nennt einen metrischen Raum (U, d) *bogenzusammenhängend*, wenn für jedes Paar $x, y \in U$ eine stetige Kurve $\gamma\colon [a, b] \to U$ mit $\gamma(a) = x$ und $\gamma(b) = y$ existiert, d.h., x und y lassen sich durch die Kurve γ verbinden. ∎

Das folgende Lemma ist von grundlegender Bedeutung für die Analysis. Es liefert eine topologische Charakterisierung der Intervalle in $\mathbb{R}$.

Lemma X.6.7. *Eine Teilmenge $I \subseteq \mathbb{R}$ ist genau dann zusammenhängend, wenn sie ein Intervall ist.*

Beweis. Ist I kein Intervall, so existieren $a < b$ in I und ein $c \in]a, b[\setminus I$. Dann sind die Mengen

$$U_1 := \{x \in I \colon x < c\} \quad \text{und} \quad U_2 := \{x \in I \colon x > c\}$$

nichtleer, offen und disjunkt mit $U_1 \cup U_2 = I$. Also ist I nicht zusammenhängend.

Wir haben noch zu zeigen, dass jedes Intervall zusammenhängend ist. Dazu argumentieren wir indirekt. Wir nehmen an, dass $I = U_1 \cup U_2$ eine Zerlegung in zwei nichtleere disjunkte Mengen ist, die in dem metrischen Raum I jeweils offen sind. Dann existieren $a \in U_1$ und $b \in U_2$ und wir dürfen o.B.d.A. annehmen, dass $a < b$ gilt.

Wir betrachten die disjunkten Teilmengen

$$J_1 := U_1 \cap [a, b] \quad \text{und} \quad J_2 := U_2 \cap [a, b]$$

des Intervalls $[a, b]$. Wegen $[a, b] \subseteq I$ ist $J_1 \cup J_2 = [a, b]$. Wegen $U_1 = I \setminus U_2$ und $U_2 = I \setminus U_1$ sind die beiden Teilmengen U_i auch abgeschlossen in I, so dass auch die Teilmengen J_i von $[a, b]$ jeweils offen und abgeschlossen sind.

Sei $c := \sup J_1$. Da J_2 in $[a, b]$ offen ist und b enthält, ist $c < b$. Andrerseits ist $c \in J_1$, da jede Umgebung von c die Menge J_1 schneidet (vgl. Lemma IX.3.13). Da J_1 auch in $[a, b]$ offen ist, existiert ein $\varepsilon > 0$ mit $]c-\varepsilon, c+\varepsilon[\subseteq J_1$, im Widerspruch zur Definition von c. Damit ist unser indirekter Beweis abgeschlossen. ∎

Lemma X.6.8. *Ist $f\colon X \to Y$ eine stetige Abbildunge zwischen metrischen Räumen und X (bogen-)zusammenhängend, so ist auch die Teilmenge $f(X) \subseteq Y$ (bogen-)zusammenhängend.*

Beweis. Bogenzusammenhang: Sei X bogenzusammenhängend und $p, q \in f(X)$. Dann existieren Punkte $x, y \in X$ mit $f(x) = p$ und $f(y) = q$. Aus dem Bogenzusammenhang von X folgt die Existenz einer stetigen Kurve $\gamma\colon [a, b] \to X$

mit $\gamma(a) = x$ und $\gamma(b) = y$. Dann ist $f \circ \gamma\colon [a,b] \to f(X)$ eine stetige Kurve, die p und q verbindet. Also ist $f(X)$ bogenzusammenhängend.

Zusammenhang: Sei X zusammenhängend und seien $U_1, U_2 \subseteq f(X)$ offene Teilmengen mit

$$f(X) = U_1 \cup U_2 \quad \text{und} \quad U_1 \cap U_2 = \emptyset.$$

Dann sind $O_i := f^{-1}(U_i)$ offene Teilmengen von X mit

$$X = O_1 \cup O_2 \quad \text{und} \quad O_1 \cap O_2 = \emptyset.$$

Da X zusammenhängend ist, ist eine der Mengen O_i leer. Dann ist aber auch $U_i = f(O_i)$ leer. Hieraus folgt, dass $f(X)$ zusammenhängend ist. ∎

Lemma X.6.9. *Ist (U,d) ein bogenzusammenhängender metrischer Raum, so ist (U,d) auch zusammenhängend.*

Beweis. Sei $U = U_1 \cup U_2$ eine disjunkte Zerlegung in nichtleere offene Teilmengen. Dann existieren $x \in U_1$ und $y \in U_2$, und wir finden eine stetige Kurve $\gamma\colon [a,b] \to U$ mit $\gamma(a) = x$ und $\gamma(b) = y$. Dann ist

$$\gamma([a,b]) = \big(\gamma([a,b]) \cap U_1\big) \cup (\gamma([a,b]) \cap U_2)$$

eine Zerlegung von $\gamma([a,b])$ in zwei paarweise disjunkte offene Teilmengen, was im Widerspruch zum Zusammenhang von $\gamma([a,b])$ steht (Lemma X.6.7 und Lemma X.6.8). ∎

Definition X.6.10. (a) Ist U eine offene Teilmenge von $\mathbb{R}^n$, so ist jede offene Teilmenge von U auch eine offene Teilmenge von $\mathbb{R}^n$ (Nachweis!). D.h., U ist genau dann zusammenhängend, wenn keine zwei nichtleeren offenen Teilmengen $U_1, U_2 \subseteq \mathbb{R}^n$ so existieren, dass

$$U = U_1 \cup U_2 \quad \text{und} \quad U_1 \cap U_2 = \emptyset.$$

Eine nichtleere offene zusammenhängende Teilmenge $U \subseteq \mathbb{R}^n$ nennen wir ein *Gebiet.*

(b) Wir nennen eine Teilmenge $U \subseteq \mathbb{R}^n$ *sternförmig bzgl.* $p \in U$, wenn für jeden Punkt $q \in U$ die ganze Verbindungsstrecke

$$[p,q] := \{\lambda p + (1-\lambda)q \colon \lambda \in [0,1]\}$$

in U liegt.

Ist U sternförmig bzgl. dem Punkt p, so ist U auch wegzusammenhängend, denn sind x und y Punkte in U, so kann man sie durch eine Kurve verbinden, die zuerst die Strecke $[x,p]$ und dann die Strecke $[p,y]$ durchläuft (Übung).

(c) Ist $U \subseteq \mathbb{R}^n$ *konvex*, d.h., liegt mit $x, y \in U$ auch deren Verbindungsstrecke $[x,y]$ in U, so ist U sternförmig bzgl. jedem Punkt $p \in U$, insbesondere also auch zusammenhängend.

(d) Die Menge $U :=]0,1[\cup]1,2[\subseteq \mathbb{R}$ ist nicht zusammenhängend. ∎

Lemma X.6.11. *Ist $U \subseteq \mathbb{R}^n$ ein Gebiet, d.h. offen und zusammenhängend, sowie $x, y \in U$, so existiert eine stückweise stetig differenzierbare Kurve $\gamma\colon [0,1] \to U$ mit $\gamma(0) = x$ und $\gamma(1) = y$.*

Beweis. Wir betrachten die Teilmenge U_x aller Punkte $p \in U$, für die eine stückweise stetig differenzierbare Kurve

$$\eta\colon [0,1] \to U$$

mit $\eta(0) = x$ und $\eta(1) = p$ existiert.

U_x **ist offen:** Ist $p \in U_x$, so folgt aus der Offenheit von U die Existenz eines $\varepsilon > 0$, so dass die Kugel $U_\varepsilon(p)$ ganz in U enthalten ist. Sei $q \in U_\varepsilon(p)$ und $\eta\colon [0,1] \to U$ eine stückweise stetig differenzierbare Kurve mit $\eta(0) = x$ und $\eta(1) = p$. Wir verlängern nun diese Kurve wie folgt:

$$\widetilde{\eta}\colon [0,1] \to U, \quad t \mapsto \begin{cases} \eta(2t) & \text{für } t \in [0, \frac{1}{2}], \\ p + (2t-1)(q-p) & \text{für } t \in]\frac{1}{2}, 1]. \end{cases}$$

Dann ist $\widetilde{\eta}$ eine stückweise stetig differenzierbare Kurve in U, die x mit q verbindet. Also ist $U_\varepsilon(p) \subseteq U_x$, folglich ist U_x offen.

$U \setminus U_x$ **ist offen:** Ist $p \notin U_x$, so finden wir wie oben ein $\varepsilon > 0$ mit $U_\varepsilon(p) \subseteq U$. Wäre ein Punkt $q \in U_\varepsilon(p)$ in der Menge U_x enthalten, so könnten wir die Kurve von x nach q durch ein Streckenstück, das q in $U_\varepsilon(p)$ mit p verbindet, verlängern und würden so wie oben eine stückweise stetig differenzierbare Kurve von x nach p erhalten, ein Widerspruch. Also ist $U_\varepsilon(p)$ ganz in $U \setminus U_x$ enthalten und $U \setminus U_x$ somit offen.

Nun ist $U = U_x \cup (U \setminus U_x)$ eine Zerlegung von U in zwei offene disjunkte Teilmengen. Da U_x nicht leer ist und U zusammenhängend, muss $U \setminus U_x$ leer sein. Also ist $U = U_x$. Insbesondere ist $y \in U_x$, und dies war zu zeigen. ■

Lokal konstante Funktionen

Definition X.6.12. Eine Funktion $f\colon X \to \mathbb{R}$ auf einem metrischen Raum (X, d) heißt *lokal konstant*, wenn für jedes $p \in X$ ein $\varepsilon > 0$ existiert, so dass f auf der Kugel $U_\varepsilon(p)$ konstant ist. ■

Bemerkung X.6.13. Für eine lokal konstante Funktion ist jede Niveaumenge $f^{-1}(c)$, $c \in \mathbb{R}$, offen und wegen

$$f^{-1}(c) = X \setminus \bigcup_{d \neq c} f^{-1}(d)$$

auch abgeschlossen.

Hieraus folgt insbesondere, dass auf einem zusammenhängenden metrischen Raum alle lokal konstanten Funktionen schon konstant sind. Ist umgekehrt

X nicht zusammenhängend und $X = U_1 \dot\cup U_2$ eine disjunkte Zerlegung in zwei nichtleere offene Teilmengen, so wird durch

$$f\colon X \to \mathbb{R}, \quad f(x) := \begin{cases} 1 & \text{für } x \in U_1 \\ 2 & \text{für } x \in U_2 \end{cases}$$

eine lokal konstante Funktion auf X definiert, die nicht konstant ist. ∎

Satz X.6.14. *Sei $U \subseteq \mathbb{R}^n$ ein Gebiet. Eine stetig differenzierbare Funktion $F\colon U \to \mathbb{R}$ ist genau dann konstant, wenn $\mathtt{d}F = 0$ gilt.*

Beweis. Ist F konstant, so gilt trivialerweise $\mathtt{d}F = 0$.

Wir nehmen nun an, dass $\mathtt{d}F = 0$ gilt, und fixieren einen Punkt $x \in U$. Nach Lemma X.6.11 existiert zu jedem $y \in U$ eine stückweise stetig differenzierbare Kurve $\gamma\colon [a,b] \to U$ mit $\gamma(a) = x$ und $\gamma(b) = y$. Mit Satz X.6.4 ergibt sich nun

$$F(y) - F(x) = \int_\gamma \mathtt{d}F = \int_\gamma 0 = 0.$$

Also ist F konstant. ∎

Folgerung X.6.15. *Ist $U \subseteq \mathbb{R}^n$ offen und $F\colon U \to \mathbb{R}$ differenzierbar, so ist F genau dann lokal konstant, wenn $\mathtt{d}F = 0$ ist.*

Beweis. Ist F lokal konstant, so gilt trivialerweise $\mathtt{d}F = 0$. Ist andererseits $\mathtt{d}F = 0$ und $p \in U$, so existiert ein $\varepsilon > 0$ mit $U_\varepsilon(p) \subseteq U$. Da die offene Kugel $U_\varepsilon(p)$ zusammenhängend ist, also ein Gebiet, folgt aus Satz X.6.14, dass F auf $U_\varepsilon(p)$ konstant ist. Folglich ist F lokal konstant. ∎

Beispiel X.6.16. Die Teilmenge $U :=]0,1[\cup]1,2[\subseteq \mathbb{R}$ ist offen, aber kein Gebiet. Die Funktion

$$F\colon U \to \mathbb{R}, \quad F(x) := \begin{cases} 0 & \text{für } x \in]0,1[\\ 1 & \text{für } x \in]1,2[\end{cases}$$

ist beliebig oft differenzierbar, und es gilt $\mathtt{d}F = 0$, aber sie ist nicht konstant. ∎

Stammfunktionen Pfaffscher Formen

Definition X.6.17. Sei ω eine stetige Pfaffsche Form auf der offenen Teilmenge $U \subseteq \mathbb{R}^n$. Eine stetig differenzierbare Funktion $F\colon U \to \mathbb{R}$ heißt *Stammfunktion von* ω, wenn $\mathtt{d}F = \omega$ gilt. ∎

Bemerkung X.6.18. (a) Ist F eine Stammfunktion von ω und $c\colon U \to \mathbb{R}, p \mapsto c$ die konstante Funktion mit Wert c, so ist auch $F + c$ eine Stammfunktion von ω, da $\mathtt{d}(F+c) = \mathtt{d}F + \mathtt{d}c = \mathtt{d}F = \omega$ gilt.

Sind umgekehrt F_1 und F_2 zwei Stammfunktionen von ω, so gilt die Beziehung $\mathtt{d}(F_1 - F_2) = 0$. Ist U ein Gebiet, so schließen wir mit Satz X.6.14, dass $F_1 - F_2$ konstant ist.

Auf einem Gebiet U sind Stammfunktionen also bis auf eine additive Konstante eindeutig bestimmt, sofern sie existieren.

(b) Sei $n = 1$ und U ein offenes Intervall. Ist ω eine stetige Pfaffsche Form auf U, so können wir ω schreiben als

$$\omega = f \cdot \mathtt{d}x.$$

Ist $F\colon U \to \mathbb{R}$ eine stetig differenzierbare Funktion, so ist

$$\mathtt{d}F = F' \cdot \mathtt{d}x.$$

Also ist F genau dann eine Stammfunktion der Pfaffschen Form ω, wenn F eine Stammfunktion der Funktion f im Sinne der Differentialrechnung einer Veränderlichen ist (Satz VI.2.2).

Insbesondere folgt aus dem Hauptsatz der Differential- und Integralrechnung die Existenz einer Stammfunktion für jede stetige Pfaffsche Form ω auf einem Intervall. Man bekommt sie zum Beispiel durch

$$F(x) := \int_{x_0}^{x} f(t)\, dt,$$

wobei $x_0 \in U$ ein fester Punkt ist .

(c) Im Gegensatz zum eindimensionalen Fall existiert für $n = 2$ nicht für jede stetige Pfaffsche Form eine Stammfunktion. Ein Gegenbeispiel hierzu liefert für $U := \mathbb{R}^2 \setminus \{0\}$ die Pfaffsche Form

$$\omega = -\frac{y}{x^2 + y^2}\mathtt{d}x + \frac{x}{x^2 + y^2}\mathtt{d}y$$

aus Beispiel X.6.5. Wir haben gesehen, dass eine geschlossene Kurve existiert, für die das Integral $\int_\gamma \omega$ nicht verschwindet. Also hat ω keine Stammfunktion.■

Kriterium für die Existenz von Stammfunktionen

Satz X.6.19. *Genau dann besitzt die stetige Pfaffsche Form ω auf dem Gebiet U eine Stammfunktion, wenn für jeden geschlossenen Weg γ in U das Integral von ω über γ verschwindet.*

Beweis. Die Notwenigkeit der Bedingung, dass Integrale über geschlossene Kurven verschwinden, folgt aus Satz X.6.4.

Wir nehmen nun an, dass diese Bedingung erfüllt ist. Wir wählen einen festen Punkt $p \in U$. Zu jedem Punkt $q \in U$ existiert dann eine stückweise stetig differenzierbare Kurve $\gamma\colon [0,1] \to U$ mit $\gamma(0) = p$ und $\gamma(1) = q$ (Lemma X.6.11). Wir definieren

$$F_\gamma(q) := \int_\gamma \omega.$$

Um so eine Funktion auf U definieren zu können, müssen wir einsehen, dass $F_\gamma(p)$ nicht von der Wahl der Kurve γ abhängt. Sei also $\eta\colon[0,1]\to U$ eine weitere stückweise stetig differenzierbare Kurve mit $\eta(0)=p$ und $\eta(1)=q$. Wir betrachten die Kurve

$$\alpha\colon[0,2]\to U,\quad t\mapsto\begin{cases}\gamma(t), & \text{für } t\in[0,1]\\ \eta(2-t), & \text{für } t\in]1,2].\end{cases}$$

Dann ist α eine stückweise stetig differenzierbare Kurve mit $\alpha(0)=\gamma(0)=p=\eta(0)=\alpha(2)$, d.h., α ist geschlossen. Gemäß unserer Voraussetzung ist daher

$$0=\int_\alpha\omega=\int_\gamma\omega-\int_\eta\omega,$$

da im zweiten Teilstück von α die Kurve η rückwärts durchlaufen wird (siehe Bemerkung X.6.3). Hieraus ergibt sich

$$F_\gamma(q)=\int_\gamma\omega=\int_\eta\omega=F_\eta(q).$$

Wir können daher durch

$$F\colon U\to\mathbb{R},\quad q\mapsto F_\gamma(q)$$

eine Funktion definieren. Da es nicht auf die Kurve ankommt, die p und q verbindet, schreiben wir

$$F(q)=\int_p^q\omega:=\int_\gamma\omega,$$

wobei γ irgendeine Kurve von p nach q ist. Wir zeigen jetzt, dass F eine Stammfunktion von ω ist. Dazu schreiben wir $\omega=\sum_{j=1}^n f_j\,\mathtt{d}x_j$ mit stetigen Funktionen $f_j\colon U\to\mathbb{R}$. Wir haben zu zeigen, dass F stetig differenzierbar ist mit

$$\frac{\partial F}{\partial x_j}=f_j,\quad j=1,\ldots,n.$$

Sei dazu $x\in U$ und $\varepsilon>0$ mit $U_\varepsilon(x)\subseteq U$. Für $\|h\|<\varepsilon$ haben wir dann

$$F(x+h)=\int_p^{x+h}\omega=\int_p^x\omega+\int_x^{x+h}\omega=F(x)+\int_x^{x+h}\omega.$$

Da die Verbindungsstrecke der Punkte x und $x+h$ ganz in $U_\varepsilon(x)$ und damit auch in U liegt, können wir den Weg

$$\gamma_h\colon[0,1]\to U,\quad t\mapsto x+th$$

betrachten, der x und $x+h$ verbindet. Damit erhalten wir mit $\dot\gamma_h(t) = h$ für alle t:

$$\begin{aligned} F(x+h) - F(x) &= \int_x^{x+h} \omega = \int_{\gamma_h} \omega = \int_0^1 \omega(x+th)(h)\,dt \\ &= \sum_{j=1}^n \int_0^1 f_j(x+th)h_j\,dt = \sum_{j=1}^n \Big(\int_0^1 f_j(x+th)\,dt\Big)\cdot h_j. \end{aligned}$$

Für die Ableitungen ergibt sich hieraus mit Satz X.5.1:

$$\begin{aligned} \frac{\partial F(x)}{\partial x_j} &= \lim_{h_j\to 0} \frac{1}{h_j}\big(F(x+h_je_j) - F(x)\big) \\ &= \lim_{h_j\to 0} \frac{1}{h_j}\int_0^1 f_j(x+th_je_j)\,dt\cdot h_j \\ &= \lim_{h_j\to 0} \int_0^1 f_j(x+th_je_j)\,dt = \int_0^1 f_j(x)\,dt = f_j(x). \end{aligned}$$

■

Beispiel X.6.20. In einem Gebiet $U \subseteq \mathbb{R}^3$ sei das *Vektorfeld* $F\colon U \to \mathbb{R}^3$ gegeben. Wir denken uns F als ein zeitlich konstantes Kraftfeld, z.B. ein elektrisches Feld. Sind F_1, F_2, F_3 die Komponenten dieses Feldes, so können wir diese auch als eine Pfaffsche Form interpretieren:

$$\omega := F_1\cdot \mathtt{d}x_1 + F_2\cdot \mathtt{d}x_2 + F_3\cdot \mathtt{d}x_3.$$

Ist nun $\gamma\colon [a,b] \to U$ eine stückweise stetig differenzierbare Kurve, so interpretieren wir das Integral

$$\int_\gamma \omega = \int_a^b \sum_{j=1} F_j\big(\gamma(t)\big)\gamma_j'(t)\,dt = \int_a^b \langle F\big(\gamma(t)\big), \dot\gamma(t)\rangle\ dt$$

als die Arbeit, die man aufwenden muss, um sich von dem Punkt $\gamma(a)$ zum Punkt $\gamma(b)$ entlang des Weges γ zu bewegen. Das ist dadurch gerechtfertigt, dass man für ein kleines Stück des Weges näherungsweise annehmen kann, dass F konstant ist und

$$\gamma(t) = \gamma(t_i) + \frac{t-t_i}{t_{i+1}-t_i}\big(\gamma(t_{i+1}) - \gamma(t_i)\big)$$

gilt. Dann ist

$$\langle F\big(\gamma(t)\big), \dot\gamma(t)\rangle\cdot(t_{i+1}-t_i) = \langle F\big(\gamma(t_i)\big), \gamma(t_{i+1}) - \gamma(t_i)\rangle.$$

Dieser Ausdruck ist proportional zu der Länge des Weges, die hier durch die Differenz $\gamma(t_{i+1}) - \gamma(t_i)$ gegeben ist, zu der Größe des Kraftfelds F im Punkt

$\gamma(t_i)$ und zum Kosinus $\cos\alpha$ des Winkels α zwischen diesen beiden Vektoren, da

$$\langle F\big(\gamma(t_i)\big), \gamma(t_{i+1}) - \gamma(t_i)\rangle = \cos\alpha \cdot \|F\big(\gamma(t_i)\big)\| \cdot \|\gamma(t_{i+1}) - \gamma(t_i)\|$$

gilt. Ist das Kraftfeld senkrecht zur Richtung des Weges, so wird keine Arbeit verrichtet; ist es dagegen parallel zum Weg, so kommt es auf seine Richtung an, ob Energie notwendig ist, um dagegen anzukämpfen, oder ob potentielle Energie dadurch frei wird, dass man in Richtung des Feldes gezogen wird.

(a) Ist $F = E$ ein elektrostatisches Feld, so ist die Arbeit, die man zum Verschieben einer Ladung entlang eines Weges γ verrichten muss, aus physikalischen Gründen nur abhängig von Anfangs- und Endpunkt des Weges, denn sonst würde sich Energie dadurch billig gewinnen lassen, dass man eine Ladung auf eine geschlossene Bahn schickt, auf der sie Energie gewinnen kann. Man spricht daher auch von *konservativen Kraftfeldern*.

In diesem Fall hängt das Integral $\int_\gamma \omega$ also nur vom Anfangs- und Endpunkt des Weges ab und verschwindet insbesondere für geschlossene Wege. Nach Satz X.6.19 hat die Pfaffsche Form ω daher auf jedem Gebiet U eine Stammfunktion, die gegeben ist durch

$$\Phi(p) := \int_{p_0}^{p} \omega,$$

wobei $p_0 \in U$ ein fest gewählter Punkt ist. Die Funktion $-\Phi$ nennt man ein *Potential* des Kraftfelds, das durch ω beschrieben wird. Ist γ eine Kurve von p nach q, so ist

$$\int_\gamma \omega = \Phi(q) - \Phi(p),$$

d.h., die Differenz der Werte der Potentialfunktion Φ gibt die Arbeit an, die entlang des Weges verrichtet wurde.

(b) Ein besonders einfaches Feld ist das elektrische Feld einer Punktladung im $\mathbb{R}^3$. Ist eine Ladung der Größe q im Punkt $p_0 \in \mathbb{R}^3$ positioniert, so ist das zugehörige elektrische Feld gegeben durch

$$E\colon \mathbb{R}^3 \setminus \{p_0\} \to \mathbb{R}^3, \qquad E(x) = \frac{q}{4\pi\varepsilon_0} \frac{x - p_0}{\|x - p_0\|^3},$$

wobei ε_0 die elektrische Feldkonstante ist.

Wir behaupten, dass die Funktion

$$\Phi\colon \mathbb{R}^3 \setminus \{p_0\} \to \mathbb{R}, \qquad \Phi(x) = \frac{q}{4\pi\varepsilon_0} \frac{1}{\|x - p_0\|}$$

eine Potentialfunktion dieses Vektorfeldes ist. In der Tat haben wir

$$\frac{\partial}{\partial x_1} \frac{1}{\|x - p_0\|} = -\frac{1}{2} \frac{1}{\|x - p_0\|^3} 2x_1 = -\frac{1}{\|x - p_0\|^3} x_1,$$

$$\frac{\partial}{\partial x_2} \frac{1}{\|x - p_0\|} = -\frac{1}{\|x - p_0\|^3} x_2, \qquad \frac{\partial}{\partial x_3} \frac{1}{\|x - p_0\|} = -\frac{1}{\|x - p_0\|^3} x_3$$

und daher

$$\nabla\Phi = -E.$$

Das Gravitationsfeld einer Masse im $\mathbb{R}^3$ lässt sich vollkommen analog behandeln, da es die gleiche Struktur besitzt. ■

Geschlossene Pfaffsche Formen

Die Bedingung, dass die Integrale einer Pfaffschen Form über alle geschlossenen Kurven verschwinden, kann man in der Regel nicht so leicht nachprüfen. Die folgende Bedingung ist eng damit verwandt, lässt sich aber sehr leicht verifizieren.

Definition X.6.21. Sei $U \subseteq \mathbb{R}^n$ eine offene Teilmenge. Eine stetig differenzierbare Pfaffsche Form $\omega = \sum_j f_j \,\mathtt{d}x_j$ auf U heißt *geschlossen*, wenn für alle $j, k \in \{1, \ldots, n\}$ gilt:

$$\frac{\partial f_k}{\partial x_j} = \frac{\partial f_j}{\partial x_k}. \qquad \blacksquare$$

Besitzt die stetig differenzierbare Pfaffsche Form ω auf U eine Stammfunktion, so ist ω geschlossen, denn aus

$$\omega = \mathtt{d}F = \sum_j \frac{\partial F}{\partial x_j} \,\mathtt{d}x_j = \sum_j f_j \,\mathtt{d}x_j$$

folgt

$$\frac{\partial f_k}{\partial x_j} = \frac{\partial^2 F}{\partial x_j \partial x_k} = \frac{\partial^2 F}{\partial x_k \partial x_j} = \frac{\partial f_j}{\partial x_k}$$

aus dem Satz von Schwarz X.3.2. Die Geschlossenheit einer Pfaffschen Form ist also notwendig für die Existenz einer Stammfunktion.

In der Regel ist sie aber nicht hinreichend. Für $U := \mathbb{R}^2 \setminus \{0\}$ und die Pfaffsche Form

$$\omega = -\frac{y}{x^2 + y^2}\mathtt{d}x + \frac{x}{x^2 + y^2}\mathtt{d}y$$

ist

$$f_1(x, y) = -\frac{y}{x^2 + y^2} \quad \text{und} \quad f_2(x, y) = \frac{x}{x^2 + y^2}.$$

Damit ergibt sich

$$\frac{\partial f_1}{\partial y} = -\frac{1}{x^2 + y^2} + \frac{2y^2}{(x^2 + y^2)^2} = \frac{2y^2 - x^2 - y^2}{(x^2 + y^2)^2} = \frac{y^2 - x^2}{(x^2 + y^2)^2}$$

und

$$\frac{\partial f_2}{\partial x} = \frac{1}{x^2 + y^2} - \frac{2x^2}{(x^2 + y^2)^2} = \frac{x^2 + y^2 - 2x^2}{(x^2 + y^2)^2} = \frac{y^2 - x^2}{(x^2 + y^2)^2}.$$

Also ist ω geschlossen. Andererseits haben wir schon mehrfach gesehen, dass ω keine Stammfunktion besitzt.

Ob die Geschlossenheit einer Pfaffschen Form für die Existenz einer Stammfunktion hinreichend ist, entscheidet sich an der geometrischen bzw. topologischen Struktur des Gebietes U. Hierüber kann man in einer Vorlesung über „Algebraische Topologie“ mehr lernen. Wir diskutieren hier nur eine einfache Bedingung.

Satz X.6.22. *Ist U ein Gebiet, das sternförmig bzgl. dem Punkt $p \in U$ ist, so besitzt jede geschlossene Pfaffsche Form auf U eine Stammfunktion.*

Beweis. Nach Translation des Gebiets dürfen wir o.B.d.A. $p = 0$ annehmen. Sei $\omega = \sum_j f_j \, \mathtt{d}x_j$. Wir definieren eine Funktion $F\colon U \to \mathbb{R}$ durch das Integral

$$F(x) := \int_0^1 \Big(\sum_{j=1}^n f_j(tx)x_j \Big) \, dt = \int_{\gamma_x} \omega,$$

wobei wir $\gamma_x\colon [0,1] \to U, \gamma_x(t) = tx$, setzen. Man beachte, dass wegen der Sternförmigkeit von U bzgl. 0 das Bild der Kurve γ_x in U liegt.

Mit Satz X.5.1 sehen wir zunächst, dass F differenzierbar ist, und dass wir die Ableitungen erhalten durch

$$\begin{aligned}
\frac{\partial F}{\partial x_k}(x) &= \sum_{j=1}^n \int_0^1 \frac{\partial}{\partial x_k} \big(f_j(tx)x_j \big) \, dt \\
&= \sum_{j=1}^n \int_0^1 \frac{\partial f_j(tx)}{\partial x_k} t x_j \, dt + \sum_{j=1}^n \int_0^1 f_j(tx) \frac{\partial x_j}{\partial x_k} \, dt \\
&= \int_0^1 \sum_{j=1}^n \frac{\partial f_k(tx)}{\partial x_j} t x_j \, dt + \int_0^1 f_k(tx) \, dt.
\end{aligned}$$

Setzen wir $L(t) := f_k(tx)$, so folgt aus der Kettenregel

$$L'(t) = \sum_{j=1}^n \frac{\partial f_k(tx)}{\partial x_j} x_j,$$

und wir erhalten

$$\begin{aligned}
\frac{\partial F}{\partial x_k}(x) &= \int_0^1 L'(t) t \, dt + \int_0^1 L(t) \, dt \\
&= \int_0^1 \big(tL(t) \big)' \, dt = 1 \cdot L(1) - 0 \cdot L(0) = L(1) = f_k(x).
\end{aligned}$$

■

Beispiel X.6.23. Wir kommen nochmal auf das Beispiel X.6.5 zurück. Natürlich ist das Gebiet $U = \mathbb{R}^2 \setminus \{0\}$ nicht sternförmig bzgl. einem seiner Punkte. Nehmen wir allerdings einen ganzen Strahl, zum Beispiel $]-\infty, 0] \times \{0\} = -\mathbb{R}^+ e_1$ heraus, so ist das Restgebiet

$$U_1 := \mathbb{R}^2 \setminus \big( ] - \infty, 0] \times \{0\} \big)$$

sternförmig bzgl. dem Punkt $(1, 0)$. Nach Satz X.6.22 hat ω daher auf dem kleineren Gebiet U_1 eine Stammfunktion. ■