

A N A L Y S I S II

Sommersemester 2016

Ernst Kuwert

Mathematisches Institut

Universität Freiburg

Inhaltsverzeichnis

| | | |
|----|---|----|
| 16 | Topologie im \mathbb{R}^n | 1 |
| 17 | Partielle Ableitungen | 9 |
| 18 | Die Ableitung | 13 |
| 19 | Schrankensatz | 23 |
| 20 | Extremwerte und konvexe Funktionen | 27 |
| 21 | Taylorentwicklung | 33 |
| 22 | Parameterabhängige Integrale | 41 |
| 23 | Kurvenintegrale | 47 |
| 24 | Komplexe Analysis | 59 |
| 25 | Diffeomorphismen | 65 |
| 26 | Implizite Funktionen | 73 |
| 27 | Existenz und Eindeutigkeit für das Anfangswertproblem | 79 |
| 28 | Lineare Differentialgleichungen | 87 |
| 29 | Separation der Variablen | 91 |
| 30 | Lineare Differentialgleichungen | 95 |

16 Topologie im \mathbb{R}^n

Das griechische Wort $\tau\acute{o}\pi\omicron\varsigma$ bedeutet soviel wie Ort oder Lage. Mathematisch geht es in der Topologie um Mengen mit einem Konvergenzbegriff, sogenannte topologische Rume, und um die stetigen Abbildungen zwischen diesen Rumen. Das ist ein sehr allgemeiner Ansatz, wir werden uns hier auf metrische Rume konzentrieren. Unser Hauptziel ist dabei die Wiederholung der Konzepte im \mathbb{R}^n . Im Laufe der Vorlesung werden weitere Beispiele von metrischen Rumen eine Rolle spielen.

Definition 16.1 (Metrischer Raum) *Ein metrischer Raum ist eine Menge X mit einer Funktion $d : X \times X \rightarrow [0, \infty)$, die fur alle $x, y, z \in X$ folgende Eigenschaften hat:*

Positivitat: $d(x, y) \geq 0$ mit Gleichheit genau wenn $x = y$,

Symmetrie: $d(y, x) = d(x, y)$

Dreiecksungleichung: $d(x, z) \leq d(x, y) + d(y, z)$.

Wir nennen $d(x, y)$ auch den Abstand von x und y .

In dieser Definition kann X eine beliebige Menge sein, insbesondere muss X kein Vektorraum sein. Betrachten Sie als Beispiel die Menge X aller Bahnhofe in Frankreich und

$$(16.1) \quad d(x, y) = \begin{cases} \text{minimale Fahrzeit von } x \text{ nach } y \text{ uber Paris} & \text{fur } x \neq y, \\ 0 & \text{fur } x = y. \end{cases}$$

Viele interessante metrische Rume sind normierte Vektorrume.

Definition 16.2 (Norm) *Eine Norm auf dem reellen (oder komplexen) Vektorraum X ist eine Funktion $\| \cdot \| : X \rightarrow \mathbb{R}$ mit folgenden Eigenschaften:*

Positivitat: $\|x\| \geq 0$, mit Gleichheit genau wenn $x = 0$.

Halblinearitat: $\|\lambda x\| = |\lambda| \cdot \|x\|$ fur alle $\lambda \in \mathbb{R}$, $x \in X$.

Dreiecksungleichung: $\|x + y\| \leq \|x\| + \|y\|$ fur alle $x, y \in X$.

Das wichtigste Beispiel ist naturlich die Euklidische Norm auf dem \mathbb{R}^n :

$$(16.2) \quad |x| = \left(\sum_{i=1}^n x_i^2 \right)^{\frac{1}{2}} \quad \text{fur } x = (x_1, \dots, x_n) \in \mathbb{R}^n.$$

Wir verwenden oft doppelte Betragstriche, einfache jedoch fur die Euklidische Norm, auf \mathbb{R} oder \mathbb{C} ist diese ja gleich dem Betrag, also keine Verwechslungsgefahr. Positivitat und Halblinearitat sind fur die Euklidische Norm klar, die Dreiecksungleichung folgt aus der Ungleichung von Cauchy-Schwarz, siehe Analysis I, Satz 5.4. Andere Normen auf \mathbb{R}^n sind zum Beispiel die 1-Norm und die Maximumnorm

$$(16.3) \quad \|x\|_1 = \sum_{i=1}^n |x_i| \quad \text{und} \quad \|x\|_\infty = \max_{1 \leq i \leq n} |x_i|.$$

Jeder normierte Vektorraum $(X, \|\cdot\|)$ wird zu einem metrischen Raum, indem wir den Abstand von zwei Punkten x, y erklären durch

$$(16.4) \quad d(x, y) = \|x - y\| \quad \text{für } x, y \in X.$$

Denn offensichtlich gilt $d(x, y) \geq 0$ mit Gleichheit nur für $x = y$, sowie

$$\begin{aligned} d(y, x) &= \|y - x\| = \|(-1)(x - y)\| = |(-1)| \|x - y\| = d(x, y), \\ d(x, z) &= \|x - z\| = \|(x - y) + (y - z)\| \leq \|x - y\| + \|y - z\| = d(x, y) + d(y, z). \end{aligned}$$

Insbesondere ist \mathbb{R}^n ein metrischer Raum mit dem üblichen euklidischen Abstandsbegriff.

Definition 16.3 Sei X ein metrischer Raum. Die offene Kugel um x_0 mit Radius $r > 0$ ist

$$B_r(x_0) = \{x \in X : d(x, x_0) < r\}.$$

Bezüglich der Euklidischen Norm auf \mathbb{R}^n gilt also wie gewohnt

$$B_r(x_0) = \{x \in \mathbb{R}^n : |x - x_0| < r\}.$$

Es ist instruktiv, sich die Kugeln $B_r(x_0)$ für die französische Eisenbahnmetrik aus (16.1) sowie die Kugeln $B_1(0)$ für die Normen $\|\cdot\|_1$ und $\|\cdot\|_\infty$ auf \mathbb{R}^n zu überlegen.

Definition 16.4 (Offene Mengen) Sei X ein metrischer Raum. Eine Menge $\Omega \subset X$ heißt offen, falls zu jedem $x \in \Omega$ ein $\varepsilon > 0$ existiert mit $B_\varepsilon(x) \subset \Omega$.

Beispiel 16.1 Die Kugel $B_r(x_0)$ ist offen in X , vgl. Analysis I, Beispiel 5.3. Sei nämlich $x \in B_r(x_0)$ gegeben. Dann ist $\varepsilon = r - d(x, x_0) > 0$ und für $y \in B_\varepsilon(x)$ folgt

$$d(y, x_0) \leq d(y, x) + d(x, x_0) < \varepsilon + d(x, x_0) = r,$$

also $B_\varepsilon(x) \subset B_r(x_0)$, was zu zeigen war.

Satz 16.1 (Topologie) Das System der offenen Teilmengen eines metrischen Raums X bildet eine Topologie, das heißt es gelten folgende Eigenschaften:

- (a) \emptyset, X sind offen.
- (b) Der Durchschnitt von endlich vielen offenen Mengen ist offen.
- (c) Die Vereinigung von beliebig vielen offenen Mengen ist offen.

BEWEIS: (vgl. Analysis I, Satz 5.8) Aussage (a) ist klar. Für (b) sei $x \in \bigcap_{i=1}^N \Omega_i$, wobei $\Omega_1, \dots, \Omega_N$ endlich viele offene Teilmengen von X sind. Dann gibt es $\varepsilon_i > 0$ mit $B_{\varepsilon_i}(x) \subset \Omega_i$. Es folgt $\varepsilon = \min_{1 \leq i \leq N} \varepsilon_i > 0$ sowie $B_\varepsilon(x) \subset B_{\varepsilon_i}(x) \subset \Omega_i$ für jedes i , das heißt $B_\varepsilon(x) \subset \bigcap_{i=1}^N \Omega_i$.

Für (c) sei nun $x \in \bigcup_{\lambda \in \Lambda} \Omega_\lambda$, wobei Λ eine beliebige Indexmenge ist. Dann ist $x \in \Omega_{\lambda_0}$ für (mindestens) ein $\lambda_0 \in \Lambda$. Da Ω_{λ_0} offen, gibt es ein $\varepsilon > 0$ mit $B_\varepsilon(x) \subset \Omega_{\lambda_0}$, also erst recht $B_\varepsilon(x) \subset \bigcup_{\lambda \in \Lambda} \Omega_\lambda$. \square

Ein abzählbarer Schnitt von offenen Mengen muss nicht offen sein. Zum Beispiel sind die Kugeln $B_{\frac{1}{n}}(0)$, $n \in \mathbb{N}$, offen im \mathbb{R}^n , nicht aber der Schnitt

$$\bigcap_{n=1}^{\infty} B_{\frac{1}{n}}(0) = \{0\}.$$

Eine offene Menge $\Omega \subset X$ mit $x \in \Omega$ nennt man auch offene Umgebung von x . Insbesondere wird die offene Kugel $B_{\varepsilon}(x)$ als ε -Umgebung von x bezeichnet.

Lemma 16.1 (Hausdorff-Trennungseigenschaft) *In einem metrischen Raum X gibt es zu zwei Punkten $x, y \in X$ mit $x \neq y$ ein $\varepsilon > 0$ mit $B_{\varepsilon}(x) \cap B_{\varepsilon}(y) = \emptyset$.*

BEWEIS: Sei $z \in B_{\varepsilon}(x) \cap B_{\varepsilon}(y)$. Dann folgt $d(x, y) \leq d(x, z) + d(z, y) < 2\varepsilon$. Also ist die Behauptung richtig für jedes $\varepsilon \leq \frac{1}{2}d(x, y)$. \square

Definition 16.5 (Konvergenz) *Sei X ein metrischer Raum. Die Folge $(x_k)_{k \in \mathbb{N}}$ von Punkten $x_k \in X$ konvergiert gegen $x \in X$, falls gilt:*

Für alle $\varepsilon > 0$ gibt es ein $K \in \mathbb{R}$ mit $x_k \in B_{\varepsilon}(x)$ für alle $k > K$.

Äquivalent dazu ist $d(x_k, x) \rightarrow 0$ mit $k \rightarrow \infty$.

Der Grenzwert ist eindeutig bestimmt, denn wäre $y \neq x$ ebenfalls Grenzwert von (x_k) , so wählen wir $\varepsilon > 0$ wie in Lemma 16.1 und erhalten für k hinreichend groß den Widerspruch

$$x_k \in B_{\varepsilon}(x) \cap B_{\varepsilon}(y) = \emptyset.$$

Definition 16.6 (abgeschlossene Teilmenge) *Eine Teilmenge A eines metrischen Raums X heißt abgeschlossen, wenn folgende Implikation stets gilt:*

$$x_k \in A, \quad x_k \rightarrow x \quad \Rightarrow \quad x \in A.$$

Die Eigenschaften *offen* und *abgeschlossen* sind nicht Gegensätze. Die leere Menge und der ganze Raum X sind sowohl offen als auch abgeschlossen. Es gilt aber folgende Komplementarität.

Satz 16.2 *In einem metrischen Raum X gilt für jede Menge $M \subset X$:*

$$M \text{ offen} \quad \Leftrightarrow \quad X \setminus M \text{ abgeschlossen.}$$

BEWEIS: Im Fall $X = \mathbb{R}^n$ wurde das in Analysis I, Satz 5.7, gezeigt. Das Argument gilt analog für jeden metrischen Raum X . \square

Folgerung 16.1 *Für die abgeschlossenen Teilmengen eines metrischen Raums X gilt:*

- a) \emptyset, X sind abgeschlossen.
- b) Die Vereinigung von endlich vielen abgeschlossenen Mengen ist abgeschlossen.
- c) Der Durchschnitt von beliebig vielen abgeschlossenen Mengen ist abgeschlossen.

BEWEIS: Folgt aus Satz 16.1 und Satz 16.2. □

Die Vereinigung von unendlich vielen abgeschlossenen Mengen ist nicht notwendig abgeschlossen, zum Beispiel $\bigcup_{n=1}^{\infty} [\frac{1}{n}, 1] = (0, 1] \subset \mathbb{R}$.

Beispiel 16.2 (induzierte Metrik) Ist (X, d) metrischer Raum, so ist jede Teilmenge $M \subset X$ selbst ein metrischer Raum mit der induzierten Abstandsfunktion

$$(16.5) \quad d_M : M \times M \rightarrow [0, \infty), \quad d_M(x, y) = d(x, y).$$

Zum Beispiel ist die Sphäre $\mathbb{S}^{n-1} = \{x \in \mathbb{R}^n : |x| = 1\}$ ein metrischer Raum mit dem Euklidischen Abstand $d_{\mathbb{S}^{n-1}}(x, y) = |x - y|$. Für die Kugeln bezüglich der induzierten Abstandsfunktion gilt allgemein

$$B_r^M(x) = \{y \in M : d_M(y, x) < r\} = \{y \in M : d(y, x) < r\} = M \cap B_r(x).$$

Die offenen Mengen in (M, d_M) sind genau die Mengen $M \cap \tilde{U}$ mit \tilde{U} offen in X . Denn ist $x \in M \cap \tilde{U}$, so gibt es ein $\varepsilon > 0$ mit $B_\varepsilon(x) \subset \tilde{U}$, also $B_\varepsilon^M(x) \subset M \cap \tilde{U}$. Die Mengen des Typs sind also offen. Ist umgekehrt $U \subset (M, d_M)$ eine beliebige offene Menge, so gibt es zu jedem $x \in U$ ein $\varepsilon(x) > 0$ mit $B_{\varepsilon(x)}^M(x) \subset U$. Es folgt

$$U = \bigcup_{x \in U} B_{\varepsilon(x)}^M(x) = \bigcup_{x \in U} M \cap B_{\varepsilon(x)}(x) = M \cap \underbrace{\bigcup_{x \in U} B_{\varepsilon(x)}(x)}_{\text{offen in } X}.$$

Weiter gilt: die abgeschlossenen Mengen in (M, d_M) sind genau von der Form $M \cap \tilde{A}$ mit \tilde{A} abgeschlossen in X . Denn es gilt

$$M \setminus (M \cap \tilde{A}) = M \cap (X \setminus \tilde{A}) = \text{offen in } M.$$

Also $M \cap \tilde{A}$ abgeschlossen in M nach Satz 16.2. Und jede in M abgeschlossene Menge hat diese Form, denn für \tilde{U} offen in X ist

$$A = M \setminus (M \cap \tilde{U}) = M \cap \underbrace{(X \setminus \tilde{U})}_{\text{abg. in } X}.$$

In der eindimensionalen Analysis wurden meist Funktionen auf einem Intervall I mit Randpunkten $a < b$ betrachtet. Im mehrdimensionalen Fall werden wir oft Kugeln $B_r(x)$ oder achsenparallele Quader $I_1 \times \dots \times I_n$ betrachten, bisweilen aber auch kompliziertere Mengen. Dafür sind die folgenden Begriffe nützlich.

Definition 16.7 Sei X ein metrischer Raum und $M \subset X$. Dann definieren wir

$$\begin{aligned} \text{int } M &= \{x \in M : \exists \varepsilon > 0 \text{ mit } B_\varepsilon(x) \subset M\} && \text{(Menge der inneren Punkte von } M), \\ \overline{M} &= \{x \in X : \forall \varepsilon > 0 \text{ ist } B_\varepsilon(x) \cap M \neq \emptyset\} && \text{(Abschluss von } M), \\ \partial M &= \{x \in X : \forall \varepsilon > 0 \text{ sind } B_\varepsilon(x) \cap M, B_\varepsilon(x) \cap (X \setminus M) \neq \emptyset\} && \text{(Rand von } M). \end{aligned}$$

Trivialerweise gilt $\text{int } M \subset M \subset \overline{M}$. Außerdem ist $\text{int } \Omega = \Omega$ für $\Omega \subset X$ offen sowie $\overline{M} = M$ für $M \subset X$ abgeschlossen.

Beispiel 16.3 Auf dem \mathbb{R}^n mit der euklidischen Abstandsfunktion $d(x, y) = |x - y|$ gilt für die Kugel $B_r(x) = \{y \in \mathbb{R}^n : d(y, x) < r\}$:

$$\overline{B_r(x)} = \{y \in \mathbb{R}^n : d(y, x) \leq r\} \quad \text{und} \quad \partial B_r(x) = \{x \in \mathbb{R}^n : d(y, x) = r\}.$$

Beweis Übungsaufgabe.

Satz 16.3 Sei M Teilmenge des metrischen Raums X .

(a) $\text{int } M$ ist offen, und es gilt die Implikation

$$\Omega \text{ offen, } \Omega \subset M \quad \Rightarrow \quad \Omega \subset \text{int } M.$$

(b) \overline{M} ist abgeschlossen, und es gilt die Implikation

$$A \text{ abgeschlossen, } A \supset M \quad \Rightarrow \quad A \supset \overline{M}.$$

(c) ∂M ist abgeschlossen und es gilt $\partial M = \overline{M} \setminus \text{int } M$.

BEWEIS: Für (a) sei $x \in \text{int } M$, also $B_r(x) \subset M$ für ein $r > 0$. Für $y \in B_r(x)$ gilt dann $B_\varepsilon(y) \subset B_r(x) \subset M$ mit $\varepsilon = r - d(y, x) > 0$, vgl. Beispiel 16.1. Es folgt $B_r(x) \subset \text{int } M$, damit ist $\text{int } M$ offen. Sei nun Ω offen und $\Omega \subset M$. Zu $x \in \Omega$ gibt es ein $\varepsilon > 0$ mit $B_\varepsilon(x) \subset \Omega$, also auch $B_\varepsilon(x) \subset M$, das heißt $x \in \text{int } M$.

Für (b) verwenden wir (a) und Satz 16.2. Nach Definition ist $X \setminus \overline{M} = \text{int}(X \setminus M)$, also ist $\overline{M} = X \setminus \text{int}(X \setminus M)$ abgeschlossen. Ist nun $A \subset X$ eine beliebige abgeschlossene Menge mit $A \supset M$, so ist $X \setminus A$ offen sowie $X \setminus A \subset X \setminus M$, also $X \setminus A \subset \text{int}(X \setminus M)$ nach (a), und somit $A \supset \overline{M}$. Dies beweist (b).

Nach Definition gilt weiter $\partial M = \overline{M} \cap \overline{(X \setminus M)}$, also ist ∂M abgeschlossen nach (b) und Folgerung 16.1. Ferner ist ebenfalls nach Definition $X \setminus \text{int } M = \overline{X \setminus M}$, folglich

$$\partial M = \overline{M} \cap (X \setminus \text{int } M) = \overline{M} \setminus \text{int } M.$$

□

Im Abschluss von M können noch zwei Sorten Punkte unterschieden werden, die Häufungspunkte und die isolierten Punkte.

Definition 16.8 Ein Punkt $x \in X$ heißt

Häufungspunkt von $M \Leftrightarrow$ für jedes $\varepsilon > 0$ ist $M \cap B_\varepsilon(x) \setminus \{x\}$ nichtleer,

isolierter Punkt von $M \Leftrightarrow$ es gibt ein $\varepsilon > 0$ mit $M \cap B_\varepsilon(x) = \{x\}$.

Ist $x \in X$ Häufungspunkt von M , so enthält $B_\varepsilon(x) \cap M \setminus \{x\}$ sogar unendlich viele Punkte. Denn würde die Menge nur aus endlich vielen Punkten y_1, \dots, y_N bestehen, so ist $\delta = \min_{1 \leq i \leq N} d(y_i, x) > 0$ und dann $B_\delta(x) \cap M \setminus \{x\} = \emptyset$, ein Widerspruch. Insbesondere können wir eine Folge $x_k \in M \setminus \{x\}$ bestimmen mit $x_k \rightarrow x$.

Definition 16.9 Eine Teilmenge M eines metrischen Raums X heißt dicht, falls $\overline{M} = X$.

Bekanntes Beispiel sind die rationalen Zahlen \mathbb{Q} im metrischen Raum \mathbb{R} , beziehungsweise die rationalen Punkte \mathbb{Q}^n im \mathbb{R}^n .

Definition 16.10 (Stetigkeit) Eine Abbildung $f : X \rightarrow Y$ zwischen metrischen Räumen heißt stetig im Punkt $x_0 \in X$, wenn es zu jedem $\varepsilon > 0$ ein $\delta > 0$ gibt mit

$$d(f(x), f(x_0)) < \varepsilon \quad \text{für alle } x \in X \text{ mit } d(x, x_0) < \delta,$$

oder äquivalent mit $f(B_\delta(x_0)) \subset B_\varepsilon(f(x_0))$. Die Funktion f heißt stetig, wenn sie in jedem Punkt $x_0 \in X$ stetig ist.

Wir müssten hier eigentlich $d_X(\cdot, \cdot)$ und $d_Y(\cdot, \cdot)$ schreiben, denn im allgemeinen sind X und Y verschiedene metrische Räume, jedoch führt die einfachere Notation nicht zu Missverständnissen.

Definition 16.11 (Lipschitzstetigkeit) Eine Abbildung $f : X \rightarrow Y$ heißt Lipschitzstetig mit Konstante $L \geq 0$, falls

$$d(f(x), f(x')) \leq L d(x, x') \quad \text{für alle } x, x' \in X.$$

Beispiel 16.4 Die Abstandsfunktion von einem Punkt $x_0 \in X$, das heißt

$$f : X \rightarrow \mathbb{R}, f(x) = d(x, x_0),$$

ist Lipschitzstetig mit Konstante $L = 1$, denn aus der Dreiecksungleichung folgt

$$f(x) = d(x, x_0) \leq d(x, x') + d(x', x_0) = d(x, x') + f(x').$$

Durch Vertauschen von x und x' folgt $|f(x) - f(x')| \leq d(x, x')$ wie gewünscht.

Die folgende Umformulierung der Stetigkeit ist unverzichtbar in allgemeinen topologischen Räumen.

Satz 16.4 (Charakterisierung stetiger Abbildungen) Eine Abbildung $f : X \rightarrow Y$ zwischen metrischen Räumen ist genau dann stetig, wenn für jede offene Menge $V \subset Y$ das Urbild $f^{-1}(V)$ offen in X ist.

BEWEIS: Sei f stetig, $V \subset Y$ offen und $x_0 \in f^{-1}(V)$. Dann ist $y_0 = f(x_0) \in V$, also gilt $B_\varepsilon(y_0) \subset V$ für geeignetes $\varepsilon > 0$. Es gibt dann ein $\delta > 0$ mit $f(B_\delta(x_0)) \subset B_\varepsilon(y_0) \subset V$, also $B_\delta(x_0) \subset f^{-1}(V)$ wie verlangt.

Umgekehrt sei $y_0 = f(x_0)$ und $\varepsilon > 0$ gegeben. Nach Voraussetzung ist dann $f^{-1}(B_\varepsilon(y_0))$ offen, das heißt es gibt ein $\delta > 0$ mit $B_\delta(x_0) \subset f^{-1}(B_\varepsilon(y_0))$ beziehungsweise $f(B_\delta(x_0)) \subset B_\varepsilon(y_0)$. \square

Im Gegensatz zur Definition 16.10 der Stetigkeit, die in jedem Punkt des Definitionsbereichs einzeln überprüft werden kann, bezieht sich Satz 16.4 auf die Funktion als Ganzes, die Charakterisierung ist nicht lokal.

Aus Analysis 1 wissen wir, dass es neben *offen* und *abgeschlossenen* Mengen eine dritte wichtige Klasse gibt.

Definition 16.12 (Folgenkompaktheit) Ein metrischer Raum X heißt folgenkompakt, wenn jede Folge $(x_k)_{k \in \mathbb{N}}$ in X eine Teilfolge $(x_{k_p})_{p \in \mathbb{N}}$ hat, die gegen ein $x \in X$ konvergiert.

Eine alternative Charakterisierung der Kompaktheit mittels Überdeckungen kommt später. Zunächst beziehen wir uns stets auf Definition 16.12. Jede Teilmenge des \mathbb{R}^n ist metrischer Raum mit der induzierten, also Euklidischen Metrik. Das folgende nützliche Kriterium wurde in Analysis I, Satz 10.2, gezeigt.

Satz 16.5 (Kompaktheit im \mathbb{R}^n) *Eine Menge $M \subset \mathbb{R}^n$ ist genau dann kompakt, wenn sie abgeschlossen und beschränkt ist.*

Diese Aussage ist in vielen metrischen Räumen falsch, das heißt es kann abgeschlossene und beschränkte Mengen geben, die nicht kompakt sind. Folgende Aussagen über Stetigkeit und kompakte Mengen sind oft nützlich.

Satz 16.6 (Bilder kompakter Mengen) *Sei $f : X \rightarrow Y$ stetige Abbildung zwischen metrischen Räumen, und X sei kompakt. Dann gilt:*

- (1) $f(X)$ ist kompakte Teilmenge von Y .
- (2) Ist f injektiv, so ist $f^{-1} : f(X) \rightarrow X$ stetig.

BEWEIS: Sei (y_k) eine Folge in $M = f(X)$, also $y_k = f(x_k)$. Da X kompakt, gibt es eine Teilfolge mit $x_{k_j} \rightarrow x \in X$. Aus der Stetigkeit von f folgt $y_{k_j} = f(x_{k_j}) \rightarrow f(x) \in M$.

Sei nun f injektiv. Angenommen f^{-1} ist in $y = f(x)$ unstetig. In metrischen Räumen kann Stetigkeit über Folgen definiert werden, vgl. Satz 7.1 in Analysis I. Also gibt es eine Folge $y_k = f(x_k)$ mit $y_k \rightarrow y$, aber $d_X(x_k, x) \geq \varepsilon$ für alle k . Da X kompakt, gibt es eine Teilfolge $x_{k_j} \rightarrow x' \in X$ und es gilt $d(x', x) \geq \varepsilon$. Aber f ist stetig in x' , also $f(x') = \lim_{j \rightarrow \infty} f(x_{k_j}) = y$, im Widerspruch zur Injektivität von f . \square

Satz 16.7 (Extrema) *Eine stetige Funktion auf einem kompakten metrischen Raum X ist beschränkt und nimmt ihr Infimum und Supremum an.*

BEWEIS: vgl. Analysis I, Satz 10.1. \square

Beispiel 16.5 Sei X metrischer Raum und $K \subset X$ kompakt. Dann gibt es zu jedem $x_0 \in X$ einen nächsten Punkt $x \in K$, das heißt

$$d(x, x_0) = \inf_{y \in K} d(y, x_0) = \text{dist}(x_0, K).$$

Der Punkt x ist nicht notwendig eindeutig, betrachte etwa $K = \{1, -1\} \subset \mathbb{R}$ und $x_0 = 0$.

Satz 16.8 (Gleichmäßige Stetigkeit) *Sei X kompakter metrischer Raum. Dann ist jede stetige Abbildung $f : X \rightarrow Y$ sogar gleichmäßig stetig.*

BEWEIS: vgl. Satz 14.4 in Analysis I. Wäre f nicht gleichmäßig stetig, so gibt es ein $\varepsilon > 0$ und $x_n, x'_n \in X$ mit $d(x_n, x'_n) \rightarrow 0$, aber $d(f(x_n), f(x'_n)) \geq \varepsilon$. Da X kompakt, konvergiert die Folge x_n nach evtl. Auswahl einer Teilfolge gegen ein $x \in X$. Wegen $d(x'_n, x) \leq d(x'_n, x_n) + d(x_n, x)$ konvergiert dann auch die Folge x'_n gegen x , und es gilt aufgrund der Stetigkeit

$$\varepsilon \leq d(f(x_n), f(x'_n)) \leq d(f(x_n), f(x)) + d(f(x), f(x'_n)) \rightarrow 0,$$

ein Widerspruch. \square

17 Partielle Ableitungen

Für Funktionen auf dem \mathbb{R}^n gibt es mehrere Ableitungskonzepte. Die partiellen Ableitungen sind am einfachsten, es sind die eindimensionalen Ableitungen in Richtung der Koordinatenachsen. Im Folgenden bezeichnet e_1, \dots, e_n die Standardbasis des \mathbb{R}^n , also $e_j = (0, \dots, 1, \dots, 0)$ mit der 1 an der j -ten Stelle.

Definition 17.1 Sei $\Omega \subset \mathbb{R}^n$ offen und $f : \Omega \rightarrow \mathbb{R}^m$. Die partielle Ableitung von f nach x_j an der Stelle $x \in \Omega$ ist der Grenzwert (falls existent)

$$\partial_j f(x) = \lim_{t \rightarrow 0} \frac{f(x + te_j) - f(x)}{t} = \frac{d}{dt} f(x + te_j)|_{t=0}.$$

Andere Bezeichnungen: $\frac{\partial f}{\partial x_j}(x)$ oder $f_{x_j}(x)$.

Sei lokal f auf einem Quader $\Omega = I_1 \times \dots \times I_n$ gegeben. Halten wir alle $x_k \in I_k$ fest mit Ausnahme von $k = j$, so ergibt sich die Funktion einer reellen Variablen

$$f_{x,j} : I_j \rightarrow \mathbb{R}, f_{x,j}(x_j) = f(x_1, \dots, x_{j-1}, x_j, x_{j+1}, \dots, x_n).$$

Die gewöhnliche eindimensionale Ableitung von $f_{x,j}$ ist

$$f'_{x,j}(x_j) = \lim_{h \rightarrow 0} \frac{f_{x,j}(x_j + h) - f_{x,j}(x_j)}{h} = \lim_{h \rightarrow 0} \frac{f(x + he_j) - f(x)}{h} = \partial_j f(x).$$

Bei der Berechnung der partiellen Ableitung $\partial_j f$ können wir also die gewohnte eindimensionale Ableitung nach x_j bilden und dabei die anderen Variablen als Konstanten behandeln.

Die wohlbekanntenen Differentiationsregeln für Funktionen einer Variablen ergeben in diesem Kontext direkt folgende Aussagen.

Satz 17.1 (Ableitungsregeln) Sei $\Omega \subset \mathbb{R}^n$ offen und $x \in \Omega$. Die Existenz der partiellen Ableitungen $\partial_j f(x)$ und $\partial_j g(x)$ sei hier stets vorausgesetzt. Dann gelten folgende Aussagen:

(a) Linearität: für $f, g : \Omega \rightarrow \mathbb{R}^m$ und $\alpha, \beta \in \mathbb{R}$ gilt

$$\partial_j(\alpha f + \beta g)(x) = \alpha \partial_j f(x) + \beta \partial_j g(x).$$

(b) Komponentenweise Differentiation: für $f : \Omega \rightarrow \mathbb{R}^m$ gilt, wenn eine der Seiten existiert,

$$\partial_j f(x) = \sum_{i=1}^m \partial_j f_i(x) e_i.$$

(c) Produktregel: für $f : \Omega \rightarrow \mathbb{R}^m$ und $g : \Omega \rightarrow \mathbb{R}$ gilt

$$\partial_j(fg)(x) = (\partial_j f)(x)g(x) + f(x)(\partial_j g)(x).$$

(d) Quotientenregel: für $f : \Omega \rightarrow \mathbb{R}^m$ und $g : \Omega \rightarrow \mathbb{R}$ mit $g(x) \neq 0$ gilt

$$\partial_j \left(\frac{f}{g} \right) (x) = \frac{(\partial_j f)(x)g(x) - f(x)(\partial_j g)(x)}{g(x)^2}.$$

(e) Kettenregel: sei f reellwertig. Ist I offenes Intervall mit $f(\Omega) \subset I$ und $\varphi : I \rightarrow \mathbb{R}$ differenzierbar, so gilt

$$\partial_j(\varphi \circ f)(x) = \varphi'(f(x))\partial_j f(x).$$

Beispiel 17.1 Wir betrachten die Euklidische Abstandsfunktion vom Nullpunkt

$$r : \mathbb{R}^n \rightarrow \mathbb{R}, r(x) = |x| = \sqrt{x_1^2 + \dots + x_n^2}.$$

In $x \neq 0$ existieren die partiellen Ableitungen, und zwar gilt mit der Kettenregel

$$\partial_j r(x) = \frac{2x_j}{2\sqrt{x_1^2 + \dots + x_n^2}} = \frac{x_j}{r} \quad \text{für } r = r(x) = |x|.$$

Im Nullpunkt sind die partiellen Ableitungen dagegen nicht definiert, denn $r(0 + te_i) = |t|$ ist in $t = 0$ nicht differenzierbar. Die Funktion $\partial_j r$ ist in $x \neq 0$ wieder partiell differenzierbar, und wir erhalten mit der Quotientenregel die zweiten partiellen Ableitungen

$$\partial_i(\partial_j r)(x) = \frac{(\partial_i x_j)r - x_j \partial_i r}{r^2} = \frac{1}{r} \left(\delta_{ij} - \frac{x_i x_j}{r^2} \right).$$

Ist $\varphi : (0, \infty) \rightarrow \mathbb{R}$ zweimal differenzierbar, so berechnen wir weiter für $f = \varphi \circ r : \mathbb{R}^n \rightarrow \mathbb{R}$

$$\partial_j f(x) = \varphi'(r)\partial_j r = \varphi'(r)\frac{x_j}{r},$$

$$\partial_i(\partial_j f)(x) = \varphi''(r)\partial_i r \partial_j r + \varphi'(r)\partial_i(\partial_j r) = \varphi''(r)\frac{x_i x_j}{r^2} + \frac{\varphi'(r)}{r} \left(\delta_{ij} - \frac{x_i x_j}{r^2} \right).$$

Wir betrachten nun den Laplaceoperator $\Delta f = \sum_{i=1}^n \partial_i^2 f$. Die Gleichung $\Delta f = 0$ spielt in der komplexen Analysis, der Theorie der Minimalflächen und der Elektrostatik eine zentrale Rolle, ihre Lösungen heißen harmonische Funktionen. Wir rechnen jetzt die rotationssymmetrischen harmonischen Funktionen aus, und zwar erhalten wir die Gleichung

$$0 \stackrel{!}{=} \Delta f(x) = \varphi''(r) + \frac{n-1}{r}\varphi'(r) = r^{1-n}(r^{n-1}\varphi'(r))'.$$

Diese Gleichung hat die Lösungen, mit Integrationskonstanten $a, b \in \mathbb{R}$,

$$\varphi(r) = \begin{cases} a \frac{r^{2-n}}{2-n} + b & \text{für } n \geq 3 \\ a \log r + b & \text{für } n = 2. \end{cases}$$

Für $n = 3$ ist $f(x) = -\frac{1}{r}$ das Coulombpotential einer Punktladung.

Wir wollen zweite und höhere partielle Ableitung nochmal allgemein einführen. Sei für $f : \Omega \rightarrow \mathbb{R}^m$ die Ableitungsfunktion $\partial_j f : \Omega \rightarrow \mathbb{R}^m$ erklärt. Ist diese im Punkte $x \in \Omega$ nach x_i partiell differenzierbar, so setzen wir

$$(17.1) \quad \partial_{i_j}^2 f(x) := \partial_i(\partial_j f)(x) \quad (\text{alternative Notation } \frac{\partial^2 f}{\partial x_i \partial x_j}(x) \text{ oder } f_{x_i x_j}(x)).$$

Entsprechend für Ableitungen beliebiger Ordnung: ist für $i_1, \dots, i_k \in \{1, \dots, n\}$ die Ableitungsfunktion $\partial_{i_1 \dots i_k}^k f : \Omega \rightarrow \mathbb{R}^n$ definiert und ist diese in $x \in \Omega$ nach x_i partiell differenzierbar, so setzen wir

$$(17.2) \quad \partial_{i_1 \dots i_k}^{k+1} f(x) = \partial_i(\partial_{i_1 \dots i_k}^k f)(x).$$

Die folgenden Klassen von Funktionen spielen in der Analysis eine wichtige Rolle. Die Tatsache, dass die partiellen Ableitungen nicht nur existieren sondern auch stetig sind, ist in nahezu allen Anwendungen wesentlich.

Definition 17.2 (C^k -Räume) Sei $\Omega \subset \mathbb{R}^n$ offen und $k \in \mathbb{N}_0 \cup \{\infty\}$. Wir bezeichnen mit $C^k(\Omega, \mathbb{R}^m)$ die Menge aller k -mal stetig differenzierbaren Funktionen auf Ω mit Werten im \mathbb{R}^m , das heißt alle Ableitungen $\partial_{i_1 \dots i_j}^j f$ der Ordnung $j \leq k$ (bzw. $j < \infty$ im Fall $k = \infty$) sind definiert und stetig auf Ω . Im reellwertigen Fall, also $m = 1$, setzen wir $C^k(\Omega, \mathbb{R}) = C^k(\Omega)$.

Wir wollen nun zeigen, dass die Operatoren ∂_i und ∂_j auf C^2 -Funktionen vertauschen. Aus der Existenz der partiellen Ableitungen $\partial_i \partial_j f$ und $\partial_j \partial_i f$ allein folgt das nicht:

$$(17.3) \quad f(x, y) = \begin{cases} xy \frac{x^2 - y^2}{x^2 + y^2} & \text{für } (x, y) \neq (0, 0) \\ 0 & \text{für } (x, y) = (0, 0) \end{cases}$$

Für diese Funktion ist $\partial_1 \partial_2 f(0, 0) = 1$, aber $\partial_2 \partial_1 f(0, 0) = -1$. Beide Ableitungen existieren, aber sie sind nicht gleich.

Satz 17.2 (von Schwarz) Sei $\Omega \subset \mathbb{R}^n$ offen. Ist $f \in C^2(\Omega)$, so vertauschen für $1 \leq i, j \leq n$ die Ableitungen nach x_i und x_j :

$$\partial_i \partial_j f = \partial_j \partial_i f \quad \text{auf } \Omega.$$

BEWEIS: Für $g : \Omega \rightarrow \mathbb{R}$ ist die Ableitung $\partial_i g(x)$ Grenzwert der Differenzenquotienten $\Delta_i^s g(x) = \frac{1}{s}(g(x + se_i) - g(x))$ für $s \rightarrow 0$. Wir betrachten den zweifachen Differenzenquotienten

$$\begin{aligned} \Delta_i^s(\Delta_j^t f)(x) &= \frac{1}{s}(\Delta_j^t f(x + se_i) - \Delta_j^t f(x)) \\ &= \frac{f(x + se_i + te_j) - f(x + se_i) - f(x + te_j) + f(x)}{st}. \end{aligned}$$

Offenbar gilt $\Delta_i^s(\Delta_j^t f)(x) = \Delta_j^t(\Delta_i^s f)(x)$. Es reicht also für $s, t > 0$ zu zeigen:

$$\lim_{(s,t) \rightarrow (0,0)} \Delta_i^s(\Delta_j^t f)(x) = \partial_j(\partial_i f)(x).$$

Der Mittelwertsatz der Differentialrechnung liefert allgemein die Darstellung

$$(\Delta_i^s g)(x) = \frac{d}{ds} g(x + se_i)|_{s=\sigma} = \partial_i g(x + \sigma e_i) \quad \text{für ein } \sigma \in [0, s].$$

Um das auf $g = \Delta_j^t f$ anzuwenden, berechnen wir die Ableitung

$$\partial_i(\Delta_j^t f)(x) = \frac{d}{ds} \frac{f(x + se_i + te_j) - f(x + se_i)}{t} \Big|_{s=0} = \frac{\partial_i f(x + te_j) - \partial_i f(x)}{t} = \Delta_j^t(\partial_i f)(x).$$

Wir wenden den Mittelwertsatz zweimal an. Es gibt dann $\sigma \in [0, s]$, $\tau \in [0, t]$ mit

$$\Delta_i^s(\Delta_j^t f)(x) = \partial_i(\Delta_j^t f)(x + \sigma e_i) = \Delta_j^t(\partial_i f)(x + \sigma e_i) = \partial_j(\partial_i f)(x + \sigma e_i + \tau e_j).$$

Da $\partial_j(\partial_i f)$ stetig im Punkt x ist, folgt $\Delta_i^s(\Delta_j^t f)(x) \rightarrow \partial_j(\partial_i f)(x)$ wie behauptet. \square

Folgerung 17.1 Für eine Funktion $f \in C^k(\Omega, \mathbb{R}^m)$ vertauschen die partiellen Ableitungen bis zur Ordnung k , das heißt für jede Permutation $\sigma \in S_k$ gilt

$$\partial_{i_{\sigma(1)}} \dots \partial_{i_{\sigma(k)}} f = \partial_{i_1} \dots \partial_{i_k} f.$$

BEWEIS: Nach Satz 17.2 können benachbarte Operatoren ∂_i, ∂_j vertauscht werden. Die symmetrische Gruppe wird durch Vertauschungen erzeugt (siehe Lineare Algebra). \square

Der Begriff der partiellen Ableitung allein ist nicht geeignet, um die mehrdimensionale Differentialrechnung zu entwickeln. Entscheidendes Manko ist, dass aus der Existenz der partiellen Ableitungen $\partial_1 f, \dots, \partial_n f$ in $x \in \Omega$ nicht die Stetigkeit von f im Punkt x folgt.

Beispiel 17.2 Sei $\Omega = \mathbb{R}^2$ und

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

Dann gilt $f(x, 0) = 0 = f(0, y)$, insbesondere $\partial_1 f(0, 0) = 0 = \partial_2 f(0, 0)$. Aber für $c(t) = (t, t)$ gilt $f(c(t)) = 1/2$ für alle $t \neq 0$, das heißt f ist nicht stetig im Nullpunkt.

Also kann die Verkettung $f \circ c$ mit einer Kurve unstetig sein. Aber dann ist $f \circ c$ auch nicht differenzierbar, siehe Analysis I, Satz 9.1, und eine Kettenregel kann es nicht geben. Die Definition der partiellen Ableitungen macht explizit von den Koordinaten auf \mathbb{R}^n Gebrauch. Es wäre denkbar, dass sich ein besserer Ableitungsbegriff ergibt, wenn alle Richtungen gleichberechtigt betrachtet werden. Dies führt auf den Begriff der Richtungsableitung.

Definition 17.3 (Richtungsableitung) Sei $\Omega \subset \mathbb{R}^n$ offen und $f : \Omega \rightarrow \mathbb{R}^m$. Die Richtungsableitung von f an der Stelle $x \in \Omega$ in Richtung $v \in \mathbb{R}^n$ ist der Grenzwert (falls existent)

$$\partial_v f(x) = \lim_{t \rightarrow 0} \frac{f(x + tv) - f(x)}{t} = \frac{d}{dt} f(x + tv)|_{t=0}.$$

Beispiel 17.3 Die Richtungsableitung von $r(x) = |x|$ in $x \in \mathbb{R}^n \setminus \{0\}$ in Richtung $v \in \mathbb{R}^n$ ist

$$\partial_v r(x) = \frac{d}{dt} \sqrt{|x|^2 + 2t\langle x, v \rangle + t^2|v|^2} |_{t=0} = \left\langle \frac{x}{|x|}, v \right\rangle.$$

Es gibt aber wieder schlechte Nachrichten: selbst wenn in $x \in \Omega$ alle Richtungsableitungen existieren, kann die Funktion im Punkt x trotzdem unstetig sein.

Beispiel 17.4 Betrachte jetzt auf $\Omega = \mathbb{R}^2$ die Funktion

$$f(x, y) = \begin{cases} \frac{2xy^2}{x^2 + y^4} & (x, y) \neq (0, 0) \\ 0 & (x, y) = (0, 0). \end{cases}$$

Dann existieren im Punkt $(0, 0)$ alle Richtungsableitungen, denn für $v = (a, b) \neq (0, 0)$ ist

$$\partial_v f(0, 0) = \lim_{t \rightarrow 0} \frac{2ab^2}{a^2 + t^2b^4} = \begin{cases} 2b^2/a & \text{für } a \neq 0 \\ 0 & \text{für } a = 0. \end{cases}$$

Dennoch ist f im Nullpunkt unstetig, denn für $c(t) = (t^2, t)$ gilt $f(c(t)) = 1$ für alle $t \neq 0$.

18 Die Ableitung

Das Konzept der mehrdimensionalen Ableitung beruht auf dem Ansatz, dass eine differenzierbare Funktion mit einer affin-linearen Funktion lokal in erster Ordnung übereinstimmt, siehe Analysis I, Lemma 9.2. Zur Abgrenzung von den partiellen Ableitungen verwendet man auch den Begriff der totalen Ableitung. Wir betrachten hier Abbildungen zwischen \mathbb{R}^n und \mathbb{R}^m , den Raum der linearen Abbildungen bezeichnen wir mit $L(\mathbb{R}^n, \mathbb{R}^m)$. Die Definition ist aber auch für beliebige normierte Räumen X, Y sinnvoll, nur muss dann $L(X, Y)$ als Raum der stetigen linearen Abbildungen erklärt werden. Man spricht von Differenzierbarkeit im Sinne von Fréchet.

Definition 18.1 (Ableitung) Sei $\Omega \subset \mathbb{R}^n$ offen. $f : \Omega \rightarrow \mathbb{R}^m$ heißt differenzierbar in $x_0 \in \Omega$, falls es ein $A \in L(\mathbb{R}^n, \mathbb{R}^m)$ gibt, so dass gilt:

$$(18.1) \quad \lim_{x \rightarrow x_0} \frac{f(x) - (f(x_0) + A(x - x_0))}{|x - x_0|} = 0.$$

Mit der Substitution $h = x - x_0$ erhalten wir die äquivalente Fassung

$$(18.2) \quad \lim_{h \rightarrow 0} \frac{f(x_0 + h) - (f(x_0) + Ah)}{|h|} = 0.$$

Die Abbildung $A \in L(\mathbb{R}^n, \mathbb{R}^m)$ ist durch (18.1) eindeutig bestimmt, siehe Satz 18.1, und heißt Ableitung von f in x_0 . Notation: $Df(x_0) = A$.

Eine Basis kommt in der Definition nicht explizit vor. Zum Rechnen werden aber in aller Regel die Standardbasen benutzt. Eine lineare Abbildung $A \in L(\mathbb{R}^n, \mathbb{R}^m)$ hat dann eine zugehörige Matrix $a \in \mathbb{R}^{m \times n}$, und zwar gilt

$$Ax = \sum_{i=1}^m \sum_{j=1}^n a_{ij} x_j e_i.$$

Umgekehrt entspricht jeder Matrix $a \in \mathbb{R}^{m \times n}$ durch diese Formel eine lineare Abbildung $A \in L(\mathbb{R}^n, \mathbb{R}^m)$. Oft wird zwischen linearer Abbildung und Matrix gar nicht unterschieden.

Satz 18.1 (Berechnung und Eindeutigkeit der Ableitung) Die Funktion $f : \Omega \rightarrow \mathbb{R}^m$ sei in $x_0 \in \Omega$ differenzierbar. Dann hat f in x_0 die Richtungsableitungen

$$(18.3) \quad \partial_v f(x_0) = Df(x_0)v \quad \text{für alle } v \in \mathbb{R}^n,$$

und $Df(x_0)$ hat bezüglich der Standardbasen die Matrixdarstellung (Jacobimatrix)

$$(18.4) \quad (\partial_j f_i(x_0)) = \begin{pmatrix} \partial_1 f_1(x_0) & \dots & \dots & \partial_n f_1(x_0) \\ \vdots & & & \vdots \\ \vdots & & & \vdots \\ \partial_1 f_m(x_0) & \dots & \dots & \partial_n f_m(x_0) \end{pmatrix} \in \mathbb{R}^{m \times n}.$$

Insbesondere ist die Ableitung durch (18.1) eindeutig bestimmt.

BEWEIS: Für $v = 0$ sind beide Seiten von (18.3) nach Definition gleich Null. Für $v \neq 0$ berechnen wir mit $A = Df(x_0)$,

$$\begin{aligned} \left| \frac{f(x_0 + tv) - f(x_0)}{t} - Av \right| &= \frac{|f(x_0 + tv) - (f(x_0) + A(tv))|}{|t|} \\ &= \frac{|f(x_0 + tv) - (f(x_0) + A(tv))|}{|tv|} |v|. \end{aligned}$$

Für $t \rightarrow 0$ geht die rechte Seite gegen Null nach (18.1), also folgt $\partial_v f(x_0) = DF(x_0)v$. Setzen wir $v = e_j$ ein und berechnen die Ableitung komponentenweise, siehe Satz 17.1, so folgt weiter

$$Df(x_0)e_j = \partial_j f(x_0) = \sum_{i=1}^m \partial_j f_i(x_0)e_i.$$

□

Um die Differenzierbarkeit einer Funktion $f : \Omega \rightarrow \mathbb{R}^m$ im Punkt $x \in \Omega \subset \mathbb{R}^n$ zu zeigen, kann man in zwei Schritten vorgehen. Erstens berechnet man die Jacobimatrix, also die partiellen Ableitungen im Punkt x . Zweitens prüft man, ob die Entwicklung (18.1) gilt, wenn A die Jacobimatrix ist. Nach Satz 18.1 ist das die einzig mögliche Wahl.

Beispiel 18.1 Die komplexe Funktion $f(z) = z^2$ lautet in reellen Koordinaten

$$f : \mathbb{R}^2 \rightarrow \mathbb{R}^2, f(z) = \begin{pmatrix} x^2 - y^2 \\ 2xy \end{pmatrix} \quad \text{wobei } z = \begin{pmatrix} x \\ y \end{pmatrix}.$$

Wir schreiben die Punkte im \mathbb{R}^2 hier als Spaltenvektoren, zwecks Konsistenz mit der Notation der Jacobimatrix. Diese ist

$$A = \begin{pmatrix} 2x & -2y \\ 2y & 2x \end{pmatrix}.$$

Damit berechnen wir für $\zeta = \begin{pmatrix} \xi \\ \eta \end{pmatrix}$ nach (18.1) den Approximationsfehler

$$\begin{aligned} f(z + \zeta) - (f(z) + A\zeta) &= \begin{pmatrix} (x + \xi)^2 - (y + \eta)^2 \\ 2(x + \xi)(y + \eta) \end{pmatrix} \\ &\quad - \left(\begin{pmatrix} x^2 - y^2 \\ 2xy \end{pmatrix} + \begin{pmatrix} 2x & -2y \\ 2y & 2x \end{pmatrix} \begin{pmatrix} \xi \\ \eta \end{pmatrix} \right) \\ &= \begin{pmatrix} \xi^2 - \eta^2 \\ 2\xi\eta \end{pmatrix}. \end{aligned}$$

Mit $|\zeta| = (\xi^2 + \eta^2)^{1/2}$ ist die Norm rechts abgeschätzt durch $C|\zeta|^2$, also folgt

$$\frac{f(z + \zeta) - (f(z) + A\zeta)}{|\zeta|} \rightarrow 0 \quad \text{mit } \zeta \rightarrow 0.$$

Beispiel 18.2 (Lineare Abbildungen) Sei $\Omega \subset \mathbb{R}^n$ offen und $A \in L(\mathbb{R}^n, \mathbb{R}^m)$. Dann ist

$$f : \Omega \rightarrow \mathbb{R}^m, f(x) = Ax \quad \text{für alle } x \in \Omega,$$

in allen $x_0 \in \Omega$ differenzierbar mit Ableitung $Df(x_0) = A$. Dies folgt sofort wegen $f(x_0 + h) = A(x_0 + h) = Ax_0 + Ah = f(x_0) + Ah$.

Beispiel 18.3 (Quadratische Formen) Sei $b : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$ symmetrische Bilinearform. Wir betrachten die quadratische Form

$$f : \mathbb{R}^n \rightarrow \mathbb{R}, f(x) = \frac{1}{2} b(x, x).$$

Um die Ableitung im Punkt $x \in \mathbb{R}^n$ zu bestimmen, entwickeln wir

$$f(x+h) = \frac{1}{2} b(x+h, x+h) = \underbrace{f(x) + b(x, h)}_{\text{affinlinear in } h} + \underbrace{\frac{1}{2} b(h, h)}_{\text{quadratisch in } h}.$$

Es folgt $Df(x)h = b(x, h)$, denn der Restterm hat die Abschätzung

$$|b(h, h)| \leq \sum_{i,j=1}^n |b(e_i, e_j)| |h_i| |h_j| \leq C|h|^2 \quad \text{mit } C = \sum_{i,j=1}^n |b(e_i, e_j)|.$$

Beispiel 18.4 (Funktionen einer Variablen) Natürlich muss das Konzept auch in diesem Fall Sinn machen. Die Funktion $f : I = (a, b) \rightarrow \mathbb{R}^m$ habe in $x \in I$ die Ableitung $f'(x) \in \mathbb{R}^m$ im Sinne von Analysis 1. Dann ist f differenzierbar in x im Sinne von Definition 18.1 mit

$$Df(x) : \mathbb{R} \rightarrow \mathbb{R}^m, \quad Df(x)h = f'(x)h.$$

Denn es gilt für $h \neq 0$, siehe auch Lemma 9.2 in Analysis I,

$$\frac{|f(x+h) - (f(x) + f'(x)h)|}{|h|} = \left| \frac{f(x+h) - f(x)}{h} - f'(x) \right| \rightarrow 0 \quad \text{mit } h \rightarrow 0.$$

Für reelle Funktionen $f : \Omega \rightarrow \mathbb{R}$ ist $Df(x) \in L(\mathbb{R}^n, \mathbb{R})$, also Element des Dualraums von \mathbb{R}^n . Es ist anschaulicher, den zugehörigen Vektor im \mathbb{R}^n zu betrachten.

Definition 18.2 (Gradient) Sei $\Omega \subset \mathbb{R}^n$ offen und $f : \Omega \rightarrow \mathbb{R}$ differenzierbar in $x \in \Omega$. Der Gradient von f im Punkt x ist der Vektor

$$\text{grad } f(x) = \sum_{j=1}^n \partial_j f(x) e_j = \begin{pmatrix} \partial_1 f(x) \\ \vdots \\ \partial_n f(x) \end{pmatrix} \in \mathbb{R}^n.$$

Formal gehen wir vom Zeilenvektor $Df(x)$ zum Spaltenvektor $\text{grad } f(x)$ mit denselben Einträgen über. Eine Charakterisierung ohne Koordinaten ist wie folgt: der Gradient ist der eindeutig bestimmte Vektor im \mathbb{R}^n mit der Eigenschaft

$$(18.5) \quad \langle \text{grad } f(x), v \rangle = Df(x)v \quad \text{für alle } v \in \mathbb{R}^n.$$

Dabei ist $\langle \cdot, \cdot \rangle$ das Standardskalarprodukt. Ist $\text{grad } f(x) = 0$, so heißt x kritischer Punkt von f . Ist x nicht kritisch, so ist die Richtung von $\text{grad } f(x)$ diejenige, in der f am stärksten ansteigt. Denn für $v \in \mathbb{R}^n$ mit $|v| = 1$ folgt aus der Ungleichung von Cauchy-Schwarz

$$(18.6) \quad \partial_v f(x) = \langle \text{grad } f(x), v \rangle \leq |\text{grad } f(x)|, \quad \text{Gleichheit genau wenn } v = \frac{\text{grad } f(x)}{|\text{grad } f(x)|}.$$

Beispiel 18.5 Der Gradient der Funktion $f(x) = \varphi(r)$ mit $r(x) = |x|$ ist nach Beispiel 17.1

$$\text{grad } f(x) = \varphi'(r) \frac{x}{r} \quad \text{für } x \neq 0.$$

Beispiel 18.6 Sei $b : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$ symmetrische Bilinearform und $B \in \mathbb{R}^{n \times n}$ die zugehörige Matrix, also $B_{ij} = b(e_i, e_j)$. Es gilt dann, da B symmetrisch,

$$b(x, y) = \langle Bx, y \rangle \quad \text{für alle } x, y \in \mathbb{R}^n.$$

Wir betrachten wieder die quadratische Form

$$f : \mathbb{R}^n \rightarrow \mathbb{R}, \quad f(x) = \frac{1}{2} b(x, x).$$

Nach Beispiel 18.3 gilt für alle $v \in \mathbb{R}^n$

$$\langle \text{grad } f(x), v \rangle = Df(x)v = b(x, v) = \langle Bx, v \rangle.$$

Also ist $\text{grad } f(x) = Bx$.

In Analysis I haben wir die Ableitung mit der Existenz der Tangente an den Graphen der Funktion motiviert. Im n -dimensionalen erwarten wir analog die Existenz einer n -dimensionalen Tangentialebene. Eine reellwertige Funktion f auf $\Omega \subset \mathbb{R}^n$ kann immer als Höhenfunktion einer Landschaft über der Grundfläche Ω interpretiert werden. Betrachte dazu den Graph der Funktion

$$G = \{(y, f(y)) : y \in \Omega\} \subset \Omega \times \mathbb{R} \subset \mathbb{R}^{n+1}.$$

Wir wollen zeigen, dass der Graph im Punkt $p = (x, f(x))$ eine Tangentialebene hat, wenn f im Punkt x differenzierbar ist. Betrachte dazu für $\lambda > 0$ die Mengen

$$G_{p,\lambda} = \frac{1}{\lambda}(G - p) = \left\{ \left(\frac{y-x}{\lambda}, \frac{f(y)-f(x)}{\lambda} \right) : y \in \Omega \right\}.$$

Der Graph G wird um $-p$ verschoben, wobei $p = (x, f(x))$ im Nullpunkt landet, dann wird mit dem Faktor $\frac{1}{\lambda}$ gestreckt. Wir wollen die $G_{p,\lambda}$ wieder als Graphen schreiben. Substituieren wir $y = x + \lambda z$, so folgt mit $\Omega_{x,\lambda} = \{z : x + \lambda z \in \Omega\}$

$$G_{p,\lambda} = \{(z, f_{x,\lambda}(z)) : z \in \Omega_{x,\lambda}\} \quad \text{für} \quad f_{x,\lambda}(z) = \frac{f(x + \lambda z) - f(x)}{\lambda}.$$

Da Ω offen, gibt es ein $\varepsilon > 0$ mit $B_\varepsilon(x) \subset \Omega$. Es folgt $B_R(0) \subset \Omega_{x,\lambda}$ für $\lambda < \frac{\varepsilon}{R}$. Für $\lambda > 0$ hinreichend klein ist $f_{x,\lambda}(z)$ also definiert, und es gilt

$$\lim_{\lambda \searrow 0} f_{x,\lambda}(z) = Df(x)z \quad \text{für alle } z \in \mathbb{R}^n.$$

In diesem Sinn konvergieren die reskalierten Graphen $G_{p,\lambda}$ gegen die Menge

$$T_p G = \{(z, Df(x)z) : z \in \mathbb{R}^n\} \subset \mathbb{R}^{n+1}.$$

$T_p G$ ist das Bild der linearen Abbildung $z \mapsto (z, Df(x)z)$, also linearer Unterraum von \mathbb{R}^{n+1} mit Basis $(e_1, \partial_1 f(x)), \dots, (e_n, \partial_n f(x))$. Einheitsnormale von $T_p G$ ist

$$\nu(p) = \frac{(-\text{grad } f(x), 1)}{\sqrt{1 + |\text{grad } f(x)|^2}} \quad \text{für } p = (x, f(x)).$$

Im Beweis der Differentiationsregeln brauchen wir eine Abschätzung für lineare Abbildungen $A \in L(\mathbb{R}^n, \mathbb{R}^m)$ aus Analysis I, Beispiel 7.10. Und zwar hatten wir mit Cauchy-Schwarz

$$(18.7) \quad |Ax| = \left| \sum_{j=1}^n x_j A e_j \right| \leq \left(\sum_{j=1}^n |A e_j|^2 \right)^{1/2} \left(\sum_{j=1}^n x_j^2 \right)^{1/2} = |A| |x|.$$

Dabei bezeichnet $|A| = \left(\sum_{j=1}^n |A e_j|^2 \right)^{1/2}$ die Euklidische Norm von $A \in L(\mathbb{R}^n, \mathbb{R}^m)$. Es folgt, dass jede lineare Abbildung $A : \mathbb{R}^n \rightarrow \mathbb{R}^m$ Lipschitzstetig ist mit Konstante $|A|$, vgl. Analysis I, Beispiel 7.10:

$$|Ax - Ay| = |A(x - y)| \leq |A| |x - y|.$$

Die optimale, also kleinstmögliche Norm $\|A\|$ mit einer Abschätzung (18.7) heißt Operatornorm. Für uns reicht die Euklidische Norm aus, die Optimalität spielt keine Rolle. Wird \mathbb{R}^n durch einen unendlichdimensionalen Raum ersetzt, so gilt (18.7) im allgemeinen nicht und lineare Abbildungen sind dann nicht automatisch stetig.

Satz 18.2 (Differenzierbarkeit \Rightarrow Stetigkeit) Sei $\Omega \subset \mathbb{R}^n$ offen. Ist $f : \Omega \rightarrow \mathbb{R}^m$ differenzierbar in x_0 , so ist f stetig in x_0 .

BEWEIS: Wie soeben besprochen, sind affinere Funktionen stetig auf \mathbb{R}^n . Es reicht daher zu zeigen, dass die Funktion $\varphi(x) = f(x) - (f(x_0) + Df(x_0)(x - x_0))$ stetig in x_0 ist. Aber $\varphi(x_0) = 0$, und nach Definition der Differenzierbarkeit gilt

$$\varphi(x) = |x - x_0| \frac{\varphi(x)}{|x - x_0|} \rightarrow 0 \quad \text{mit } x \rightarrow x_0.$$

□

Wir müssen jetzt die Differentiationsregeln erarbeiten.

Satz 18.3 (Kettenregel) Seien $f : U \rightarrow \mathbb{R}^m$ und $g : V \rightarrow \mathbb{R}^p$ mit $U \subset \mathbb{R}^n$, $V \subset \mathbb{R}^m$ offen und $f(U) \subset V$. Sind f in x_0 und g in $f(x_0)$ differenzierbar, so ist auch $g \circ f$ in x_0 differenzierbar und es gilt die Kettenregel

$$D(g \circ f)(x_0) = Dg(f(x_0)) Df(x_0).$$

Für die zugehörigen Jacobimatrizen bedeutet das mit $y_0 = f(x_0)$

$$\frac{\partial (g \circ f)_i}{\partial x_k}(x_0) = \sum_{j=1}^m \frac{\partial g_i}{\partial y_j}(y_0) \frac{\partial f_j}{\partial x_k}(x_0) \quad \text{für } 1 \leq i \leq p, 1 \leq k \leq n.$$

BEWEIS: Sei $y_0 = f(x_0)$, $Df(x_0) = A$, $Dg(y_0) = B$. Wir definieren für hinreichend kleine $\xi \in \mathbb{R}^n \setminus \{0\}$, $\eta \in \mathbb{R}^m \setminus \{0\}$ die Funktionen

$$\varepsilon_f(\xi) = \frac{f(x_0 + \xi) - (f(x_0) + A\xi)}{|\xi|} \quad \text{und} \quad \varepsilon_g(\eta) = \frac{g(y_0 + \eta) - (g(y_0) + B\eta)}{|\eta|}.$$

Mit $\varepsilon_f(0) = 0$ und $\varepsilon_g(0) = 0$ sind beide Funktionen nach Voraussetzung im Nullpunkt stetig. Offensichtliche Kandidatin für die Ableitung von $g \circ f$ in x_0 ist BA , also berechnen wir

$$\begin{aligned} & \frac{(g \circ f)(x_0 + \xi) - ((g \circ f)(x_0) + BA\xi)}{|\xi|} \\ = & \frac{g(y_0 + A\xi + |\xi|\varepsilon_f(\xi)) - (g(y_0) + BA\xi)}{|\xi|} \\ = & \frac{g(y_0) + B\eta + |\eta|\varepsilon_g(\eta) - (g(y_0) + BA\xi)}{|\xi|} \quad \text{wobei } \eta = A\xi + |\xi|\varepsilon_f(\xi) \\ = & B\varepsilon_f(\xi) + \frac{|\eta|}{|\xi|}\varepsilon_g(\eta). \end{aligned}$$

Wegen $|B\varepsilon_f(\xi)| \leq |B||\varepsilon_f(\xi)|$ und $|\eta| \leq (|A| + |\varepsilon_f(\xi)|)|\xi| \leq C|\xi|$ konvergiert die rechte Seite wie gewünscht gegen Null. \square

Beispiel 18.7 Spezialfall ist die Verkettung $f \circ c$ einer Kurve $c : (a, b) \rightarrow \Omega \subset \mathbb{R}^n$ und einer Funktion $f : \Omega \rightarrow \mathbb{R}$. Ist c differenzierbar in $t \in (a, b)$ und f differenzierbar in $c(t)$, so folgt

$$\frac{d(f \circ c)}{dt}(t) = \sum_{j=1}^n \frac{\partial f}{\partial x_j}(c(t)) \frac{dc_j}{dt}(t),$$

beziehungsweise in vektorieller Form

$$(f \circ c)'(t) = Df(c(t))c'(t) = \langle \text{grad } f(c(t)), c'(t) \rangle.$$

Ist $f \circ c$ konstant, so folgt $\text{grad } f(c(t)) \perp c'(t)$. Anschaulich: der Gradient von f steht senkrecht auf Kurven in der Niveaumenge $\{x \in \Omega : f(x) = \text{const.}\}$, also auf die ganze Niveaumenge. Im Fall $n = 2$ kann man sich die Niveaumenge als Höhenlinie vorstellen.

Wie bei Funktionen einer Variablen kann die Ableitung vektorwertiger Funktionen auf die einzelnen Komponenten zurückgeführt werden.

Satz 18.4 (komponentenweise Differentiation) $f : \Omega \rightarrow \mathbb{R}^m$ ist genau dann in $x_0 \in \Omega$ differenzierbar, wenn alle Komponenten $f_i : \Omega \rightarrow \mathbb{R}$, $i = 1, \dots, m$, in x_0 differenzierbar sind. Ist $P_i : \mathbb{R}^m \rightarrow \mathbb{R}$ Projektion auf die i -te Koordinate, so gilt $Df_i(x_0) = P_i Df(x_0)$.

BEWEIS: Es gilt nach Definition

$$Df(x_0) = A \quad \Leftrightarrow \quad \lim_{x \rightarrow x_0} \frac{f(x) - (f(x_0) + A(x - x_0))}{|x - x_0|} = 0.$$

Die Konvergenz im \mathbb{R}^n ist gleichbedeutend mit der Konvergenz aller Komponenten. Durch Anwendung von P_i ergibt sich daher weiter die äquivalente Formulierung

$$(18.8) \quad \lim_{x \rightarrow x_0} \frac{f_i(x) - (f_i(x_0) + P_i A(x - x_0))}{|x - x_0|} = 0 \quad \text{für alle } i = 1, \dots, m.$$

Aus $Df(x_0) = A$ folgt somit $Df_i(x_0) = P_i A$. Sei umgekehrt $Df_i(x_0) = A_i$ für $i = 1, \dots, m$. Wir definieren $A \in L(\mathbb{R}^n, \mathbb{R}^m)$ durch $Av = \sum_{i=1}^m (A_i v) e_i$. Dann ist $P_i A = A_i$, also gilt (18.8) und somit $Df(x_0) = A$. \square

Wir zeigen schließlich die weiteren klassischen Ableitungsregeln.

Satz 18.5 (Ableitungsregeln) Sei $\Omega \subset \mathbb{R}^n$ offen, und $f, g : \Omega \rightarrow \mathbb{R}$ seien differenzierbar im Punkt $x \in \Omega$. Dann gelten folgende Aussagen:

(a) Linearität: für $\alpha, \beta \in \mathbb{R}$ ist $\alpha f + \beta g$ in x differenzierbar mit Ableitung

$$D(\alpha f + \beta g)(x) = \alpha Df(x) + \beta Dg(x).$$

(b) Produktregel: fg ist in x differenzierbar mit Ableitung

$$D(fg)(x) = Df(x)g(x) + f(x)Dg(x).$$

(c) Quotientenregel: ist $g(x) \neq 0$, so ist f/g auf einer Umgebung von x definiert und

$$D\left(\frac{f}{g}\right)(x) = \frac{Df(x)g(x) - f(x)Dg(x)}{g(x)^2}.$$

BEWEIS: Wir setzen $Df(x) = A$, $Dg(x) = B$, und für $h \neq 0$

$$\varepsilon_f(h) = \frac{f(x+h) - (f(x) + Ah)}{|h|}, \quad \varepsilon_g(h) = \frac{g(x+h) - (g(x) + Bh)}{|h|}.$$

Nach Voraussetzung gilt $\varepsilon_f(h) \rightarrow 0$, $\varepsilon_g(h) \rightarrow 0$ mit $h \rightarrow 0$. Mit der jeweils behaupteten Ableitung ist nun für $h \rightarrow 0$ der Grenzwert in (18.2) nachzuprüfen. Für (a) gilt

$$\frac{(\alpha f + \beta g)(x+h) - ((\alpha f + \beta g)(x) + (\alpha A + \beta B)h)}{|h|} = \alpha \varepsilon_f(h) + \beta \varepsilon_g(h) \rightarrow 0.$$

Für (b) berechnen wir mit etwas mehr Mühe

$$\begin{aligned} & \frac{(fg)(x+h) - ((fg)(x) + (Ag(x) + f(x)B)h)}{|h|} \\ &= \frac{(f(x) + Ah + \varepsilon_f(h)|h|)(g(x) + Bh + \varepsilon_g(h)|h|) - (f(x)g(x) + g(x)Ah + f(x)Bh)}{|h|} \\ &= \frac{1}{|h|} (Ah)(Bh) + \varepsilon_f(h)(g(x) + Bh + \varepsilon_g(h)|h|) + \varepsilon_g(h)(f(x) + Ah). \end{aligned}$$

Wie in (18.7) bemerkt gilt $|Ah| \leq |A||h|$ sowie $|Bh| \leq |B||h|$, also geht die rechte Seite mit $h \rightarrow 0$ gegen Null. In (c) können wir $m = 1$ und $f \equiv 1$ annehmen, denn sonst schreiben wir $f/g = f(1/g)$ und verwenden (b). Es gilt

$$\begin{aligned} & \frac{1}{|h|} \left(\frac{1}{g(x+h)} - \left(\frac{1}{g(x)} - \frac{Bh}{g(x)^2} \right) \right) \\ &= \frac{1}{|h|} \frac{1}{g(x)g(x+h)} \left(g(x) - (g(x) + Bh + \varepsilon_g(h)|h|) + \frac{g(x+h)}{g(x)} Bh \right) \\ &= \frac{1}{g(x)g(x+h)} \left(\left(\frac{g(x+h)}{g(x)} - 1 \right) \frac{Bh}{|h|} - \varepsilon_g(h) \right). \end{aligned}$$

Wegen $g(x) \neq 0$ und $g(x+h) \rightarrow g(x)$ mit $h \rightarrow 0$ nach Satz 18.2 geht die rechte Seite wieder gegen Null mit $h \rightarrow 0$. \square

Die Quotientenregel kann auch eleganter mit der Kettenregel gezeigt werden: man verwendet $1/g = h \circ g$ mit $h(y) = \frac{1}{y}$. Für die Produktregel gibt es ein ähnliches Argument: es ist $fg = h \circ \phi$ mit $\phi(x) = (f(x), g(x)) \in \mathbb{R}^2$ und $h(y_1, y_2) = y_1 y_2$. Nach Satz 18.4 ist ϕ differenzierbar, und h nach Beispiel 18.3.

Wie besprochen kann aus der Existenz der partiellen Ableitungen nicht auf die Differenzierbarkeit geschlossen werden, ja nicht einmal auf die Stetigkeit. Das ist schade, denn die partiellen Ableitungen sind so schön einfach auszurechnen, während die Definition 18.1 eventuell schwierig zu verifizieren ist. Zum Glück können wir aber doch die partiellen Ableitungen einsetzen.

Satz 18.6 (stetig partiell differenzierbar \Rightarrow differenzierbar) Sei $\Omega \subset \mathbb{R}^n$ offen. Die Funktion $f : \Omega \rightarrow \mathbb{R}^m$ sei in Ω nach x_1, \dots, x_n partiell differenzierbar. Sind die Funktionen $\partial_j f : \Omega \rightarrow \mathbb{R}^m$ in $x \in \Omega$ stetig, so ist f in x differenzierbar.

BEWEIS: Wegen Satz 18.4 können wir $m = 1$ annehmen. Mit Satz 18.1 kennen wir bereits die einzig mögliche Kandidatin für die Ableitung, nämlich

$$A : \mathbb{R}^n \rightarrow \mathbb{R}, \quad Ah = \sum_{k=1}^n \partial_k f(x) h_k.$$

Für $h \in \mathbb{R}^n$ hinreichend klein ist $f(x+h)$ mit $f(x) + Ah$ zu vergleichen, dazu wollen wir den Mittelwertsatz verwenden. Da wir nur in Achsenrichtungen differenzieren können, laufen wir längs der Kanten des Quaders, das heißt wir betrachten die Punkte $p_k = x + \sum_{i=1}^k h_i e_i$ mit $k = 0, \dots, n$. Es gilt für geeignete $s_k \in [0, 1]$

$$f(p_k) - f(p_{k-1}) = f(p_{k-1} + h_k e_k) - f(p_{k-1}) = \partial_k f(p_{k-1} + s_k h_k e_k) h_k.$$

Es folgt nun

$$\begin{aligned} \frac{|f(x+h) - (f(x) + Ah)|}{|h|} &= \frac{1}{|h|} \left| \sum_{k=1}^n (f(p_k) - f(p_{k-1}) - \partial_k f(x) h_k) \right| \\ &= \frac{1}{|h|} \left| \sum_{k=1}^n \left(\partial_k f(p_{k-1} + s_k h_k e_k) - \partial_k f(x) \right) h_k \right| \\ &\leq \sum_{k=1}^n \left| \partial_k f \left(x + \sum_{i=1}^{k-1} h_i e_i + s_k h_k e_k \right) - \partial_k f(x) \right|. \end{aligned}$$

Die rechte Seite geht mit $h \rightarrow 0$ gegen Null, da $\partial_k f$ im Punkt x stetig ist. \square

Es gibt differenzierbare Funktionen, die nicht stetig differenzierbar sind. In Analysis I, Serie 13, Aufgabe 4 hatten wir das Beispiel

$$f : \mathbb{R} \rightarrow \mathbb{R}, \quad f(x) = \begin{cases} x^2 \cos \frac{1}{x} & \text{für } x \neq 0, \\ 0 & \text{für } x = 0. \end{cases}$$

Aber hier gilt: Ausnahmen bestätigen die Regel, in den meisten Fällen ist Satz 18.6 das Mittel der Wahl, um die Differenzierbarkeit einer Funktion zu begründen. Dabei ist hilfreich, dass die Ableitungsregeln auch in der Klasse der C^k -Funktionen gelten.

Folgerung 18.1 Sei $k \in \mathbb{N}_0 \cup \{\infty\}$.

- (a) Mit $f, g \in C^k(\Omega, \mathbb{R}^m)$ gilt $\alpha f + \beta g \in C^k(\Omega, \mathbb{R}^m)$ für $\alpha, \beta \in \mathbb{R}$.
- (b) Aus $f, g \in C^k(\Omega)$ folgt $fg \in C^k(\Omega)$, sowie $f/g \in C^k(\Omega)$ falls $g \neq 0$ auf Ω .
- (c) Sind $f \in C^k(U, \mathbb{R}^m)$, $g \in C^k(V, \mathbb{R}^p)$ mit $U \subset \mathbb{R}^n$, $V \subset \mathbb{R}^m$ offen und $f(U) \subset V$, so ist $g \circ f \in C^k(U, \mathbb{R}^p)$.

BEWEIS: Im Fall $k = 0$ sind die Aussagen wohlbekannt. Die Behauptungen (a) und (b) folgen nun aus den Rechenregeln für die partielle Ableitung, siehe Satz 17.1, mit Induktion über k . Sind zum Beispiel $f, g \in C^k(\Omega)$ für ein $k \geq 1$, so gilt induktiv $\partial_j(fg) = (\partial_j f)g + f(\partial_j g) \in C^{k-1}(\Omega)$, also $fg \in C^k(\Omega)$.

Für $k \geq 1$ sind die Abbildungen f und g aus (c) differenzierbar nach Satz 18.6. Dann ist $g \circ f$ ebenfalls differenzierbar wegen der Kettenregel, Satz 18.3, mit partiellen Ableitungen

$$\frac{\partial(g \circ f)_i}{\partial x_k} = \sum_{j=1}^m \frac{\partial g_i}{\partial y_j} \circ f \frac{\partial f_j}{\partial x_k}.$$

Nun sind $\partial_k f_j \in C^{k-1}(U)$, $\partial_j g_i \in C^{k-1}(V)$ nach Voraussetzung, also $\partial_j g_i \circ f \in C^{k-1}(U)$ nach Induktion. Es folgt $\partial_k(g \circ f)_i \in C^{k-1}(U)$ mit der Produktregel aus (b), also ist $g \circ f$ von der Klasse C^k . □

19 Schrankensatz

Ein Grundproblem in der Analysis ist es, Informationen über die Ableitung in Eigenschaften der Funktion zu übersetzen. Für Funktionen einer Variablen, also $f : [a, b] \rightarrow \mathbb{R}$, stehen dazu zwei Argumente zur Verfügung:

a) der Mittelwertsatz (Analysis I, Kapitel 10):

$$f(b) - f(a) = f'(\tau)(b - a) \quad \text{für ein } \tau \in (a, b);$$

b) der Hauptsatz der Differential- und Integralrechnung (Analysis I, Kapitel 15):

$$f(b) - f(a) = \int_a^b f'(t) dt.$$

Der Mittelwertsatz hat den Nachteil, dass er keine Kontrolle der Zwischenstelle liefert, zum Beispiel bei Abhängigkeit von weiteren Parametern. Auch gilt er so nicht für vektorwertige Funktionen, wie das Beispiel $f(t) = (\cos t, \sin t)$ auf $[0, 2\pi]$ zeigt. Deshalb verwenden wir im Folgenden meistens den Hauptsatz, allerdings muss $f(t)$ dazu eine C^1 -Funktion sein. Genauer reicht es wenn $f(t)$ stetig auf $[a, b]$ und stückweise C^1 ist. Denn sei $a = t_0 < \dots < t_N = b$ eine Unterteilung, so dass f' auf den offenen Teilintervallen stetig ist und einseitige Grenzwerte hat, nicht notwendig gleich. Dann gilt

$$f(b) - f(a) = \sum_{k=1}^N (f(t_k) - f(t_{k-1})) = \sum_{k=1}^N \int_{t_{k-1}}^{t_k} f'(t) dt = \int_a^b f'(t) dt.$$

Wie kann dieses eindimensionale Argument nun für Funktionen mehrerer Variabler $f : \Omega \rightarrow \mathbb{R}$ mit $\Omega \subset \mathbb{R}^n$ eingesetzt werden? Die einfache Antwort: indem f längs Kurven $\gamma : [a, b] \rightarrow \Omega$, $\gamma = \gamma(t)$, ausgewertet wird.

Lemma 19.1 Sei $\gamma : I = [a, b] \rightarrow \Omega$ stetig und stückweise C^1 . Dann gilt für $f \in C^1(\Omega, \mathbb{R}^m)$

$$(19.1) \quad f(\gamma(b)) - f(\gamma(a)) = \int_a^b Df(\gamma(t))\gamma'(t) dt.$$

BEWEIS: Nach Folgerung 18.1 ist die Funktion $f \circ \gamma$ stückweise C^1 , und mit dem Hauptsatz der Differential- und Integralrechnung und der Kettenregel gilt

$$f(\gamma(b)) - f(\gamma(a)) = \int_a^b \frac{d}{dt} f(\gamma(t)) dt = \int_a^b Df(\gamma(t))\gamma'(t) dt.$$

□

Im Beweis trat das Integral einer vektorwertigen Funktion auf. Dieses kann komponentenweise erklärt werden, das heißt für $v \in C^0([a, b], \mathbb{R}^m)$ ist

$$\int_a^b v(t) dt = \int_a^b \left(\sum_{i=1}^m v_i(t)e_i \right) dt = \sum_{i=1}^m \left(\int_a^b v_i(t) dt \right) e_i.$$

Alternativ kann man prüfen, dass die Definition des Integrals in Analysis I, Kapitel 14, mittels Riemannscher Summen ohne Änderung auch für Funktionen mit Werten im \mathbb{R}^m funktioniert. Man kann sich so oder so davon überzeugen, dass der Hauptsatz der Differential- und Integralrechnung ganz analog für vektorwertige Funktionen gilt.

Definition 19.1 Ein metrischer Raum X heißt *wegweise zusammenhängend*, falls es zu je zwei Punkten $x_0, x_1 \in X$ eine stetige Abbildung $c : [0, 1] \rightarrow X$ gibt mit $c(0) = x_0, c(1) = x_1$.

Ein nur stetiger Weg kann unanschaulich kompliziert sein, zum Beispiel kann er die Fläche eines Quadrats überdecken (Peano 1890). Um den Hauptsatz wie in Lemma 19.1 anzuwenden, brauchen wir einen stückweise C^1 Weg, dazu das folgende Lemma.

Lemma 19.2 Sei $\Omega \subset \mathbb{R}^n$ offen und *wegweise zusammenhängend*. Dann gibt es einen Polygonzug, also einen stetigen, stückweise linearen Weg $\gamma : [0, 1] \rightarrow \Omega$ mit $\gamma(0) = x_0, \gamma(1) = x_1$.

BEWEIS: Sei $c \in C^0([0, 1], \Omega)$ mit $c(0) = x_0$ und $c(1) = x_1$ gegeben. Betrachte die Zerlegung $t_k = k/N$ für $k = 0, \dots, N$ des Intervalls $[0, 1]$, und definiere einen Polygonzug $\gamma : [0, 1] \rightarrow \mathbb{R}^n$ mit $\gamma(t_k) = c(t_k)$ für alle k durch

$$\gamma(t) = \frac{t_k - t}{t_k - t_{k-1}} c(t_{k-1}) + \frac{t - t_{k-1}}{t_k - t_{k-1}} c(t_k) \quad \text{für } t \in [t_{k-1}, t_k].$$

Wir wollen zeigen, dass dieser Polygonzug für N hinreichend groß ganz in Ω verläuft. Da $c([0, 1])$ kompakt ist und $\mathbb{R}^n \setminus \Omega$ abgeschlossen, gibt es ein $\varrho > 0$ mit

$$|x - c(t)| \geq \varrho \quad \text{für alle } t \in [0, 1], x \in \mathbb{R}^n \setminus \Omega.$$

Denn sonst gibt es Folgen $x_j \in c([0, 1]), y_j \in \mathbb{R}^n \setminus \Omega$ mit $|x_j - y_j| \rightarrow 0$. Nach Übergang zu einer Teilfolge gilt $x_j \rightarrow x \in c([0, 1])$, und weiter $y_j \rightarrow x$, also $x \in \mathbb{R}^n \setminus \Omega$, ein Widerspruch. Da c gleichmäßig stetig auf $[0, 1]$ ist, gilt

$$\text{osc}(c, \delta) = \sup\{|c(t) - c(s)| : 0 \leq s, t \leq 1, |t - s| \leq \delta\} \rightarrow 0 \quad \text{mit } \delta \rightarrow 0.$$

Daraus folgt, dass γ gleichmäßig gegen c konvergiert mit $N \rightarrow \infty$: für $t \in [0, 1]$ sei $k \in \{1, \dots, N\}$ mit $t \in [t_{k-1}, t_k]$, also

$$\begin{aligned} |\gamma(t) - c(t)| &\leq \frac{t_k - t}{t_k - t_{k-1}} |c(t_{k-1}) - c(t)| + \frac{t - t_{k-1}}{t_k - t_{k-1}} |c(t_k) - c(t)| \\ &\leq \text{osc}(c, \frac{1}{N}) \rightarrow 0 \quad \text{mit } N \rightarrow \infty. \end{aligned}$$

Für N groß ist also $|\gamma(t) - c(t)| < \varrho$ für alle $t \in [0, 1]$, und damit $\gamma([0, 1]) \subset \Omega$. □

Satz 19.1 (Konstanzsatz) Sei $\Omega \subset \mathbb{R}^n$ offen und *wegweise zusammenhängend*. Dann gilt:

$$Df(x) = 0 \quad \text{für alle } x \in \Omega \quad \Rightarrow \quad f \text{ ist konstant.}$$

BEWEIS: Zu $x_0, x_1 \in \Omega$ gibt es nach Lemma 19.2 einen Polygonzug $\gamma : [0, 1] \rightarrow \Omega$ mit $\gamma(0) = x_0$ und $\gamma(1) = x_1$. Da γ stetig und stückweise C^1 , folgt aus Lemma 19.1

$$f(x_1) - f(x_0) = \int_0^1 Df(\gamma(t))\gamma'(t) dt = 0.$$

□

Definition 19.2 Eine Menge $M \subset \mathbb{R}^n$ heißt *konvex*, falls folgende Implikation gilt:

$$x_0, x_1 \in M \quad \Rightarrow \quad (1 - t)x_0 + tx_1 \in M \quad \text{für alle } t \in [0, 1].$$

Satz 19.2 (Schrankensatz) Sei $\Omega \subset \mathbb{R}^n$ offen und konvex, und $f \in C^1(\Omega, \mathbb{R}^m)$. Es gebe ein $L < \infty$ mit $|Df(x)| \leq L$ für alle $x \in \Omega$. Dann folgt

$$|f(x_1) - f(x_0)| \leq L|x_1 - x_0| \quad \text{für alle } x_0, x_1 \in \Omega.$$

BEWEIS: Für jede stetige Funktion $\varphi : I = [a, b] \rightarrow \mathbb{R}^m$ gilt die Ungleichung

$$(19.2) \quad \left| \int_a^b \varphi \right| \leq \int_a^b |\varphi|.$$

Dies folgt durch Anwendung der Dreiecksungleichung auf die Riemannschen Summen. Sei nun $\gamma(t) = (1-t)x_0 + tx_1$ für $0 \leq t \leq 1$. Aus (19.1) und (18.7) folgt, da $\gamma'(t) = x_1 - x_0$,

$$|f(x_1) - f(x_0)| = \left| \int_0^1 Df(\gamma(t))(x_1 - x_0) dt \right| \leq \int_0^1 |Df(\gamma(t))(x_1 - x_0)| dt \leq L|x_1 - x_0|.$$

□

Die folgende lokale Variante des Schrankensatzes ist ebenfalls oft nützlich.

Folgerung 19.1 Sei $\Omega \subset \mathbb{R}^n$ offen und $f \in C^1(\Omega, \mathbb{R}^m)$. Dann gibt es zu jeder kompakten Menge $K \subset \Omega$ eine Konstante $L < \infty$ mit

$$|f(x) - f(y)| \leq L|x - y| \quad \text{für alle } x, y \in K.$$

BEWEIS: Angenommen nicht, dann gibt es zu jedem $k \in \mathbb{N}$ Punkte $x_k, y_k \in K$ mit

$$|f(x_k) - f(y_k)| > k|x_k - y_k| \quad \text{für } k = 1, 2, \dots$$

Da f stetig auf der kompakten Menge K ist, gibt es ein $M < \infty$ mit $|f(x)| \leq M$ für alle $x \in K$ nach Satz 16.7. Weiter können wir nach Wahl einer Teilfolge und Umnummerierung annehmen, dass $x_k \rightarrow x \in K$ mit $k \rightarrow \infty$. Aber

$$|x_k - y_k| < \frac{1}{k} |f(x_k) - f(y_k)| \leq \frac{2M}{k} \rightarrow 0 \quad \text{mit } k \rightarrow \infty,$$

also folgt $y_k \rightarrow x$ mit $k \rightarrow \infty$. Wähle nun ein $r > 0$ mit $\overline{B_r(x)} \subset \Omega$. Da Df stetig ist, gibt es wieder nach Satz 16.7 ein $L < \infty$ mit

$$|Df(y)| \leq L \quad \text{für alle } y \in \overline{B_r(x)}.$$

Für hinreichend große k gilt $x_k, y_k \in B_r(x)$, also liefert Satz 19.2

$$k|x_k - y_k| < |f(x_k) - f(y_k)| \leq L|x_k - y_k|,$$

ein Widerspruch für k hinreichend groß. □

20 Extremwerte und konvexe Funktionen

In diesem Abschnitt diskutieren wir lokale Extrema von Funktionen mehrerer Variabler, und verallgemeinern die notwendigen und hinreichenden Kriterien aus Analysis 1. Dabei spielt die zweite Ableitung eine entscheidende Rolle. Wir behandeln im Anschluss Grundtatsachen über konvexe Funktionen. Als bekannt setzen wir voraus: auf einer kompakten Teilmenge des \mathbb{R}^n nimmt eine stetige Funktion ihre Extremwerte an.

Definition 20.1 Die Funktion $f : M \rightarrow \mathbb{R}$, $M \subset \mathbb{R}^n$, hat in $x \in M$ ein lokales Minimum, falls es ein $\delta > 0$ gibt mit

$$f(y) \geq f(x) \quad \text{für alle } y \in B_\delta(x) \cap M.$$

Ist sogar $f(y) > f(x)$ für $y \in B_\delta(x) \setminus \{x\}$, so heißt das Minimum isoliert. Ein (isoliertes) lokales Maximum ist entsprechend definiert.

Satz 20.1 (notwendige Bedingung für Extrema) Sei $\Omega \subset \mathbb{R}^n$ offen, und $f : \Omega \rightarrow \mathbb{R}$ habe in $x \in \Omega$ ein lokales Extremum. Ist f differenzierbar in x , so folgt $Df(x) = 0$.

BEWEIS: Für $v \in \mathbb{R}^n$ hat die Funktion $t \mapsto f(x + tv)$ ein lokales Extremum bei $t = 0$, also folgt aus der eindimensionalen Version und Satz 18.1

$$0 = \frac{d}{dt} f(x + tv)|_{t=0} = Df(x)v \quad \text{für alle } v \in \mathbb{R}^n.$$

□

Definition 20.2 Ein Punkt $x \in \Omega$ mit $Df(x) = 0$ heißt kritischer Punkt von f .

Kritische Punkte sind also Kandidaten für Extremalstellen. Es gibt aber auch andere kritische Punkte, das zeigt schon das eindimensionale Beispiel $f(x) = x^3$ im Punkt $x = 0$. Um die Situation genauer zu analysieren brauchen wir die zweite Ableitung.

Definition 20.3 Sei $f \in C^2(\Omega)$ mit $\Omega \subset \mathbb{R}^n$ offen. Die zweite Ableitung von f im Punkt $x \in \Omega$ ist die Bilinearform

$$D^2 f(x) : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}, \quad D^2 f(x)(v, w) = \sum_{i,j=1}^n \partial_{ij}^2 f(x) v_i w_j.$$

Die Matrix $(\partial_{ij}^2 f(x)) \in \mathbb{R}^{n \times n}$ heißt Hessematrix von f an der Stelle x , und als Hesseform bezeichnet man die zugehörige quadratische Form

$$v \mapsto D^2 f(x)(v, v) = \sum_{i,j=1}^n \partial_{ij}^2 f(x) v_i v_j.$$

Die Hessematrix ist symmetrisch, und $D^2 f(x)$ ist symmetrische Bilinearform. Denn nach Satz 17.2 gilt $\partial_{ij}^2 f = \partial_{ji}^2 f$ für $f \in C^2(\Omega)$, und daraus folgt

$$D^2 f(x)(v, w) = \sum_{i,j=1}^n \partial_{ij}^2 f(x) v_i w_j = \sum_{i,j=1}^n \partial_{ji}^2 f(x) w_j v_i = D^2 f(x)(w, v).$$

Als erstes wollen wir die Formel für die zweite Ableitung längs Kurven herleiten.

Lemma 20.1 Sei $\Omega \subset \mathbb{R}^n$ offen, $f \in C^2(\Omega)$ und $\gamma \in C^2(I, \Omega)$. Dann gilt

$$(20.1) \quad (f \circ \gamma)''(t) = D^2 f(\gamma(t))(\gamma'(t), \gamma'(t)) + Df(\gamma(t))\gamma''(t).$$

BEWEIS: Nach Kettenregel ist $(f \circ \gamma)'(t) = \sum_{j=1}^n \partial_j f(\gamma(t))\gamma'_j(t)$, und weiter

$$(f \circ \gamma)''(t) = \sum_{i,j=1}^n \partial_{ij}^2 f(\gamma(t))\gamma'_i(t)\gamma'_j(t) + \sum_{j=1}^n \partial_j f(\gamma(t))\gamma''_j(t).$$

□

Wir benötigen nun eine lokale Entwicklung, die die zweite Ableitung mit einbezieht.

Lemma 20.2 Sei $\Omega \subset \mathbb{R}^n$ offen und $f \in C^2(\Omega)$. Dann gilt

$$\frac{f(x+h) - (f(x) + Df(x)h + \frac{1}{2}D^2f(x)(h,h))}{|h|^2} \rightarrow 0 \quad \text{mit } h \rightarrow 0.$$

BEWEIS: Setze $\gamma(t) = x + th$, das heißt nach Lemma 20.1 gilt

$$(f \circ \gamma)'(t) = Df(x+th)h \quad \text{und} \quad (f \circ \gamma)''(t) = D^2f(x+th)(h,h).$$

Wir berechnen mit dem Hauptsatz und partieller Integration

$$\begin{aligned} (f \circ \gamma)(1) &= (f \circ \gamma)(0) + \int_0^1 (f \circ \gamma)'(t) dt \\ &= (f \circ \gamma)(0) + (f \circ \gamma)'(0) + \int_0^1 (1-t)(f \circ \gamma)''(t) dt. \end{aligned}$$

Einsetzen von $\gamma(t) = x + th$ liefert

$$(20.2) \quad f(x+h) = f(x) + Df(x)h + \frac{1}{2}D^2f(x)(h,h) + \int_0^1 (1-t)(D^2f(x+th) - D^2f(x))(h,h) dt.$$

Wir schätzen den Integranden ab. Nach Cauchy-Schwarz gilt für $Q \in \mathbb{R}^{n \times n}$

$$\left| \sum_{i,j=1}^n Q_{ij}h_ih_j \right| = \left| \sum_{i=1}^n (Qh)_i h_i \right| = |\langle Qh, h \rangle| \leq |Qh| |h| \leq |Q| |h|^2.$$

Zu $\varepsilon > 0$ gibt es ein $\delta > 0$ mit $|D^2f(y) - D^2f(x)| < \varepsilon$ für $|y - x| < \delta$. Es folgt

$$|(D^2f(x+th) - D^2f(x))(h,h)| \leq \varepsilon |h|^2 \quad \text{für } |h| < \delta.$$

Damit ist das Lemma bewiesen. □

Als zweites Hilfsmittel brauchen wir folgende Tatsache über quadratische Formen.

Lemma 20.3 Sei $b : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$ eine symmetrische Bilinearform, und

$$\lambda = \inf\{b(x,x) : x \in \mathbb{R}^n, |x| = 1\}.$$

Dann gibt es ein $v \in \mathbb{R}^n$ mit $|v| = 1$ und $b(v,v) = \lambda$.

BEWEIS: Die Funktion $f : \mathbb{R}^n \rightarrow \mathbb{R}$, $f(x) = b(x, x)$, ist stetig auf \mathbb{R}^n , denn es gilt

$$f(x) = \sum_{i,j=1}^n b_{ij}x_i x_j \quad \text{wobei } b_{ij} = b(e_i, e_j).$$

Da $\{x \in \mathbb{R}^n : |x| = 1\}$ kompakt ist, existiert ein Minimierer v nach Satz 16.7. \square

Das Lemma gilt allgemeiner auf jedem Euklidischen Vektorraum V mit $n := \dim V < \infty$, und zwar wird dies auf den Fall \mathbb{R}^n reduziert durch Wahl einer Orthonormalbasis $\mathcal{B} = \{v_1, \dots, v_n\}$, also $\langle v_i, v_j \rangle = \delta_{ij}$. Eine Orthonormalbasis lässt sich aus einer beliebigen Basis mit dem Verfahren von Gram-Schmidt explizit konstruieren (Übungsaufgabe). Für gegebenes $b : V \times V \rightarrow \mathbb{R}$ haben wir dann die induzierte Bilinearform

$$b_{\mathcal{B}} : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}, \quad b_{\mathcal{B}}(x, y) = b(x_{\mathcal{B}}, y_{\mathcal{B}}) \quad \text{wobei } x_{\mathcal{B}} = \sum_{i=1}^n x_i v_i.$$

Nach Wahl von \mathcal{B} gilt $\|x_{\mathcal{B}}\| = |x|$. Durch Substitution $v = x_{\mathcal{B}}$ folgt

$$\inf\{b(v, v) : \|v\| = 1\} = \inf\{b_{\mathcal{B}}(x, x) : |x| = 1\}.$$

Das Infimum wird rechts in einem Punkt x angenommen, also links im Punkt $v = x_{\mathcal{B}}$.

Definition 20.4 Eine symmetrische Bilinearform $b : \mathbb{R}^n \rightarrow \mathbb{R}^n \rightarrow \mathbb{R}$ heißt *positiv definit* (bzw. *positiv semidefinit*), falls gilt:

$$b(v, v) > 0 \quad (\text{bzw. } b(v, v) \geq 0) \quad \text{für alle } v \in V \setminus \{0\}.$$

Notation: $b > 0$ bzw. $b \geq 0$. Entsprechend für negativ (semi-)definit.

Beachten Sie, dass *definit* die strikte Ungleichung bedeutet, anders als zum Beispiel bei der Monotonie von Funktionen, wo wir zum Ausschluss der Gleichheit den Begriff *streng monoton* verwenden. Wir bemerken auch, dass es sich nur um eine teilweise Ordnung handelt, es muss nicht einer der Fälle $b \geq 0$ oder $b \leq 0$ gelten. Für $b(x, y) = x_1 y_1 - x_2 y_2$ auf \mathbb{R}^2 gilt zum Beispiel $b(e_1, e_1) > 0$, aber $b(e_2, e_2) < 0$.

Satz 20.2 (Lokale Extrema) Sei $f \in C^2(\Omega)$, $\Omega \subset \mathbb{R}^n$ offen und $x \in \Omega$.

- (a) Wenn f in x ein lokales Minimum hat, so ist $D^2 f(x)$ positiv semidefinit.
- (b) Ist $Df(x) = 0$ und $D^2 f(x)$ positiv definit, so hat f in x ein isoliertes lokales Minimum.

BEWEIS: In (a) gilt $Df(x) = 0$ nach Satz 20.1. Für $v \in \mathbb{R}^n$ beliebig hat $t \mapsto f(x + tv)$ bei $t = 0$ ein lokales Minimum, also folgt aus dem eindimensionalen Fall und (20.1)

$$0 \leq \frac{d^2}{dt^2} f(x + tv)|_{t=0} = D^2 f(x)(v, v).$$

Für (b) verwende (20.2): zu $\varepsilon > 0$ gibt es ein $\delta > 0$ mit

$$f(x + h) - f(x) = \frac{1}{2} D^2 f(x)(h, h) + R(h) \quad \text{mit } |R(h)| < \varepsilon |h|^2 \text{ für } |h| < \delta.$$

Nach Voraussetzung ist $D^2f(x)(h, h) > 0$ für $h \neq 0$, wir brauchen aber hier eine quantitative Version. Nach Lemma 20.3 gibt es ein $v \in \mathbb{R}^n$, $|v| = 1$, mit

$$\lambda := \inf_{|w|=1} D^2f(x)(w, w) = D^2f(x)(v, v) > 0.$$

Wir wählen $\varepsilon < \frac{\lambda}{2}$. Mit dem zugehörigen $\delta > 0$ gilt

$$f(x+h) - f(x) \geq \frac{\lambda}{2}|h|^2 - \varepsilon|h|^2 = \left(\frac{\lambda}{2} - \varepsilon\right)|h|^2 > 0 \text{ für } 0 < |h| < \delta.$$

□

Um die Funktion in der Nähe eines kritischen Punkts zu verstehen, ist der folgende Satz aus der Linearen Algebra nützlich.

Satz 20.3 (Hauptachsentransformation) Sei $b : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$ symmetrische Bilinearform. Dann gibt es eine Orthonormalbasis v_1, \dots, v_n und $\lambda_1 \leq \dots \leq \lambda_n$, so dass gilt:

$$b(v_i, v_j) = \lambda_i \delta_{ij} \quad \text{für alle } i, j = 1, \dots, n.$$

BEWEIS: Setze $\lambda = \inf\{b(x, x) : |x| = 1\}$ und wähle $v \in V$ mit $|v| = 1$ und $b(v, v) = \lambda$, siehe Lemma 20.3. Wir behaupten

$$(20.3) \quad b(v, w) = \lambda \langle v, w \rangle \quad \text{für alle } w \in \mathbb{R}^n.$$

Die Gleichung stimmt für $w = v$, und die Menge der $w \in \mathbb{R}^n$ mit (20.3) ist ein Unterraum. Es reicht daher, die Gleichung für $w \in \mathbb{R}^n$ mit $\langle v, w \rangle = 0$ und $|w| = 1$ zu zeigen. Dann ist $|(\cos t)v + (\sin t)w|^2 = 1$ für $t \in \mathbb{R}$, und aus der Minimumeigenschaft folgt wie behauptet

$$b(v, w) = \frac{1}{2} \frac{d}{dt} b((\cos t)v + (\sin t)w, (\cos t)v + (\sin t)w)|_{t=0} = 0 = \lambda \langle v, w \rangle.$$

Jetzt konstruiere induktiv $\lambda_1 \leq \dots \leq \lambda_n$ und orthonormale v_1, \dots, v_n mit

$$b(v_i, w) = \lambda_i \langle v_i, w \rangle \quad \text{für alle } w \in V.$$

Mit $w = v_j$ ist das die Behauptung des Satzes. Für $k = 1$ nehmen wir $\lambda_1 = \lambda$ und $v_1 = v$ wie oben. Seien nun v_1, \dots, v_k und $\lambda_1 \leq \dots \leq \lambda_k$ schon bestimmt für $1 \leq k \leq n - 1$. Setze

$$\lambda_{k+1} = \inf\{b(x, x) : x \in V_k, |x| = 1\} \quad \text{mit } V_k = \{v_1, \dots, v_k\}^\perp.$$

Es gilt $V_{k+1} \subset V_k$ und damit $\lambda_{k+1} \geq \lambda_k$. Nach Lemma 20.3, angewandt im Raum V_k statt \mathbb{R}^n , gibt es ein $v_{k+1} \in V_k$ mit $|v_{k+1}| = 1$ und $b(v_{k+1}, v_{k+1}) = \lambda_{k+1}$. Aus (20.3) folgt

$$b(v_{k+1}, w) = \lambda_{k+1} \langle v_{k+1}, w \rangle \quad \text{für alle } w \in V_k.$$

Aber da b symmetrisch, gilt induktiv für $1 \leq i \leq k$

$$b(v_{k+1}, v_i) = b(v_i, v_{k+1}) = \lambda_i \langle v_i, v_{k+1} \rangle = 0 = \lambda_{k+1} \langle v_{k+1}, v_i \rangle.$$

Es folgt $b(v_{k+1}, w) = \lambda_{k+1} \langle v_{k+1}, w \rangle$ für alle $w \in V$, der Induktionsschluss. □

Jede lineare Abbildung $B : \mathbb{R}^n \rightarrow \mathbb{R}^n$ induziert die Bilinearform

$$b(v, w) = \langle Bv, w \rangle \quad \text{für } v, w \in \mathbb{R}^n.$$

B heißt symmetrisch, wenn b symmetrisch ist. Wegen $b(v, w) - \lambda \langle v, w \rangle = \langle Bv - \lambda v, w \rangle$ ist Gleichung (20.3) gleichbedeutend mit

$$Bv = \lambda v,$$

das heißt v ist Eigenvektor von B zum Eigenwert $\lambda \in \mathbb{R}$. Satz 20.3 besagt: symmetrische Endomorphismen des \mathbb{R}^n sind diagonalisierbar, genauer gibt es eine Orthonormalbasis aus Eigenvektoren. Diese Konzepte haben eine unendlichdimensionale Verallgemeinerung, den Spektralsatz, der zum Beispiel in der Quantenmechanik von Bedeutung ist. Der Eigenvektor zum kleinsten Eigenwert heißt dort Grundzustand, die weiteren Eigenvektoren sind die angeregten Zustände. Ein anderes Beispiel sind Schwingungen mit Grund- und Obertönen.

In den Koordinaten $x = x_1 v_1 + \dots + x_n v_n$ bezüglich der Eigenvektorbasis von b hat die quadratische Form die Darstellung

$$b(x, x) = \sum_{i=1}^n \lambda_i x_i^2 \quad \text{mit } \lambda_1 \leq \dots \leq \lambda_n.$$

Für $n = 2$ wollen wir die Mengen $M_c = \{x \in \mathbb{R}^2 : b(x, x) = c\}$ beschreiben; dabei können wir nach Übergang zu $-b$ annehmen, dass $\lambda_2 > 0$ ist. Es ergeben sich drei Fälle:

$\lambda_1 > 0$: Im Nullpunkt hat $b(x, x)$ ein globales, isoliertes Minimum und für $c > 0$ ist M_c eine achsensymmetrische Ellipse mit Scheiteln in $(\pm \sqrt{c/\lambda_1}, 0)$ und $(0, \pm \sqrt{c/\lambda_2})$.

$\lambda_1 = 0$: Auch hier ist im Nullpunkt ein globales Minimum, allerdings ist M_0 die gesamte x_1 -Achse; für $c > 0$ besteht M_c aus den parallelen Geraden $x_2 = \pm \sqrt{c/\lambda_2}$.

$\lambda_1 < 0$: M_0 ist Vereinigung der beiden Ursprungsgeraden $x_2 = \pm \sqrt{-\lambda_1/\lambda_2} x_1$. Für $c > 0$ ist M_c eine nach oben und unten geöffnete Hyperbel mit Scheiteln $(0, \pm \sqrt{c/\lambda_2})$ und M_0 als Asymptotenlinien. Für $c < 0$ erhalten wir ebenfalls eine Hyperbel mit Asymptoten M_0 , die aber nach links und rechts geöffnet ist und die Scheitel $(\pm \sqrt{c/\lambda_1}, 0)$ hat.

Betrachten wir in den drei Fällen die zugehörigen Graphen im \mathbb{R}^3 , so haben wir für $\lambda_1 > 0$ anschaulich eine Mulde, für $\lambda_1 = 0$ einen Hohlweg und für $\lambda_1 < 0$ einen Sattel. Aus der Mulde wird für $-b$ eine Kuppe. Ein kritischer Punkt x einer Funktion f heißt nicht degeneriert, wenn die Eigenwerte λ_i von $D^2 f(x)$ alle ungleich Null sind. In diesem Fall bezeichnet man die Anzahl der negativen Eigenwerte als den Index des kritischen Punkts. Für $f(x) = b(x, x)$ sind die drei Fälle oben der Reihe nach Index Null, degeneriert und Index Eins, die Kuppe ist Index zwei.

Definition 20.5 (konvexe Funktion) Sei $K \subset \mathbb{R}^n$ konvex. Dann heißt $f : K \rightarrow \mathbb{R}$ konvex, falls für alle $x, y \in K$ gilt:

$$f((1-t)x + ty) \leq (1-t)f(x) + tf(y) \quad \text{für alle } t \in [0, 1].$$

f heißt strikt konvex, falls die strikte Ungleichung gilt für $x \neq y$ und $t \in (0, 1)$.

Wie man leicht sieht, ist Konvexität von f äquivalent dazu, dass der Epigraph

$$G^+(f) = \{(x, z) \in K \times \mathbb{R} : z \geq f(x)\}$$

eine konvexe Menge im $\mathbb{R}^{n+1} = \mathbb{R}^n \times \mathbb{R}$ ist.

Satz 20.4 (Konvexitätskriterien) Sei $\Omega \subset \mathbb{R}^n$ offen und konvex, und $f \in C^1(\Omega)$. Die folgenden Aussagen sind äquivalent:

- (a) f ist konvex.
- (b) $f(y) \geq f(x) + Df(x)(y - x)$ für alle $x, y \in \Omega$.
- (c) $(Df(y) - Df(x))(y - x) \geq 0$ für alle $x, y \in \Omega$.

Ist sogar $f \in C^2(\Omega)$, so ist außerdem äquivalent:

- (d) $D^2f(x) \geq 0$ für alle $x \in \Omega$.

BEWEIS: Die Aussage wird jeweils auf den eindimensionalen Fall reduziert, indem wir für $x, y \in \Omega$ die Funktion $\varphi(t) = (1 - t)f(x) + tf(y) - f((1 - t)x + ty)$ betrachten. Unter Voraussetzung (a) hat φ in $t = 0$ ein Minimum, daraus folgt (b):

$$0 \leq \varphi'(0) = f(y) - f(x) - Df(x)(y - x).$$

Aussage (c) folgt aus (b) durch Vertauschen von x und y und Addition. Die Implikation (c) \Rightarrow (a) zeigen wir durch Widerspruch. Angenommen $\varphi(t)$ hat in $\tau \in (0, 1)$ ein Minimum $\varphi(\tau) < 0$. Für $t_1 < t_2$ gilt nach (c) mit $x(t) = (1 - t)x + ty$

$$\varphi'(t_1) - \varphi'(t_2) = \frac{1}{t_2 - t_1} (Df(x(t_2)) - Df(x(t_1)))(x(t_2) - x(t_1)) \geq 0.$$

Für $t < \tau$ folgt $\varphi'(t) \geq \varphi'(\tau) = 0$, und hieraus $\varphi(0) \leq \varphi(\tau) < 0$, ein Widerspruch.

Sei nun $f \in C^2(\Omega)$. Nach (20.2) wissen wir

$$f(x + h) = f(x) + Df(x)h + \int_0^1 (1 - t)D^2f(x + th)(h, h) dt.$$

Mit $h = y - x$ folgt die die Implikation (d) \Rightarrow (b). Umgekehrt folgt (d) aus (b) mit Satz 20.2, denn die Funktion $g(y) = f(y) - (f(x) + Df(x)(y - x))$ hat in x ein Minimum. \square

Eine Funktion f mit $f((1 - t)x + ty) \geq (1 - t)f(x) + tf(y)$ für alle $x, y \in \Omega$, $t \in [0, 1]$, heißt konkav und es gelten entsprechende Aussagen mit umgekehrten Ungleichungen.

21 Taylorentwicklung

Wir wollen nun die Taylorentwicklung von Funktionen zunächst einer und dann mehrerer Variabler herleiten. Die Idee der Taylorentwicklung ist es, eine gegebene Funktion f mit einem Polynom zu vergleichen, das mit f an einer festen Stelle x_0 „von höherer Ordnung“ übereinstimmt, das heißt einschließlich einer Reihe von Ableitungen. Dieses Polynom sollte dann auch nahe bei x_0 die Funktion gut approximieren, und das will quantifiziert werden. Für lineare sowie quadratische Polynome haben wir das in den vorigen Abschnitten schon behandelt.

Zur Erinnerung: eine Funktion $P : \mathbb{R} \rightarrow \mathbb{R}$ heißt Polynom vom Grad $k \in \mathbb{N}_0$, wenn es $a_0, \dots, a_k \in \mathbb{R}$ gibt mit $a_k \neq 0$, so dass gilt:

$$P(x) = \sum_{j=0}^k a_j x^j \quad \text{für alle } x \in \mathbb{R}.$$

Im Raum aller Funktionen $f : \mathbb{R} \rightarrow \mathbb{R}$ ist die Menge \mathbb{P}_k der Polynome vom Grad $\leq k$ der durch $1, x, \dots, x^k$ erzeugte Unterraum. Es gilt: für jedes $x_0 \in \mathbb{R}$ bilden die Funktionen $1, x - x_0, \dots, (x - x_0)^k$ eine Basis von \mathbb{P}_k . Wegen $\dim \mathbb{P}_k \leq k + 1$ müssen wir nur die lineare Unabhängigkeit zeigen. Dazu verwenden wir die Ableitungsregel

$$(21.1) \quad P(x) = \sum_{j=0}^k a_j (x - x_0)^j \quad \Rightarrow \quad \left(\frac{d}{dx}\right)^i P(x)|_{x=x_0} = i! a_i.$$

Ist $P(x)$ die Nullfunktion, so folgt $a_i = 0$ für $i = 0, \dots, k$ wie behauptet.

Lemma 21.1 Sei $I = (a, b) \subset \mathbb{R}$, $x_0 \in I$ und $k \in \mathbb{N}_0$. Zu $f \in C^k(I)$ gibt es genau ein Polynom $P \in \mathbb{P}_k$ mit $P^{(i)}(x_0) = f^{(i)}(x_0)$ für $i = 0, 1, \dots, k$, und zwar

$$(21.2) \quad P_k(x) = \sum_{j=0}^k \frac{f^{(j)}(x_0)}{j!} (x - x_0)^j.$$

P_k heißt Taylorpolynom der Ordnung k von f mit Entwicklungspunkt x_0 .

BEWEIS: Für das in (21.2) definierte Polynom gilt $P^{(i)}(x_0) = f^{(i)}(x_0)$ für $i = 0, \dots, k$, wie man mit (21.1) sieht. Zur Eindeutigkeit sei $P \in \mathbb{P}_k$ mit $P^{(i)}(x_0) = 0$ für alle $i = 0, \dots, k$. Wie oben gezeigt gilt eine Darstellung $P(x) = \sum_{j=0}^k a_j (x - x_0)^j$ für alle $x \in \mathbb{R}$. Mit (21.1) folgt $a_i = 0$ für alle $i = 0, \dots, k$. \square

Folgerung 21.1 Das k -te Taylorpolynom mit Entwicklungspunkt x_0 eines Polynoms f vom Grad höchstens k ist f selbst.

In der Situation von Lemma 21.1 heißt die Funktion

$$(21.3) \quad R_k : (a, b) \rightarrow \mathbb{R}, \quad R_k(x) = f(x) - P_k(x)$$

das Restglied k -ter Ordnung der Taylorentwicklung in x_0 . Knackpunkt bei der Taylorentwicklung ist die Abschätzung dieses Restglieds und damit eine Aussage darüber, wie gut die Funktion durch das Taylorpolynom approximiert wird. Hierfür gibt es verschiedene mögliche Darstellungen von R_k .

Satz 21.1 (Integraldarstellung des Restglieds) Sei $f \in C^{k+1}(I)$ für ein $k \in \mathbb{N}_0$, und $P_k(x) = \sum_{j=0}^k \frac{f^{(j)}(x_0)}{j!} (x - x_0)^j$ das k -te Taylorpolynom im Punkt $x_0 \in I$. Dann gilt

$$f(x) = P_k(x) + R_k(x) \quad \text{mit} \quad R_k(x) = \frac{1}{k!} \int_{x_0}^x (x - y)^k f^{(k+1)}(y) dy.$$

BEWEIS: Durch Induktion über $k \in \mathbb{N}_0$. Für $k = 0$ folgt aus dem Hauptsatz

$$f(x) = f(x_0) + \int_{x_0}^x f'(y) dy.$$

Für $k \geq 1$ folgt induktiv mit partieller Integration, vgl. Lemma 20.2 für den Fall $k = 1$,

$$\begin{aligned} f(x) &= P_{k-1}(x) + \frac{1}{(k-1)!} \int_{x_0}^x (x-y)^{k-1} f^{(k)}(y) dy \\ &= P_{k-1}(x) + \frac{1}{(k-1)!} \left(\left[-\frac{(x-y)^k}{k} f^{(k)}(y) \right]_{y=x_0}^{y=x} + \int_{x_0}^x \frac{(x-y)^k}{k} f^{(k+1)}(y) dy \right) \\ &= P_k(x) + \frac{1}{k!} \int_{x_0}^x (x-y)^k f^{(k+1)}(y) dy. \end{aligned}$$

□

Die zweite Darstellung des Restglieds ist vielleicht etwas populärer.

Satz 21.2 (Lagrange-Darstellung des Restglieds) Sei $f \in C^{k+1}(I)$ für ein $k \in \mathbb{N}_0$. Dann gibt es zu $x_0, x \in I$ ein ξ zwischen x_0 und x , so dass gilt:

$$(21.4) \quad f(x) = \sum_{j=0}^k \frac{f^{(j)}(x_0)}{j!} (x - x_0)^j + R_k(x) \quad \text{mit} \quad R_k(x) = \frac{f^{(k+1)}(\xi)}{(k+1)!} (x - x_0)^{k+1}.$$

BEWEIS: Wir verwenden den Mittelwertsatz der Integralrechnung mit Gewicht, siehe Folgerung 14.2, Analysis I: ist $\varphi \in C^0(I)$ mit $\varphi \geq 0$, so gibt es zu $f \in C^0(I)$ ein $\xi \in I$ mit

$$\int_I f \varphi = f(\xi) \int_I \varphi.$$

Sei nun $x > x_0$. Dann können wir $I = [x_0, x]$ und $\varphi(y) = (x - y)^k$ wählen. Es folgt

$$\frac{1}{k!} \int_{x_0}^x (x - y)^k f^{(k+1)}(y) dy = \frac{1}{k!} f^{(k+1)}(\xi) \int_{x_0}^x (x - y)^k dy = \frac{f^{(k+1)}(\xi)}{(k+1)!} (x - x_0)^{k+1},$$

für ein $\xi \in [x_0, x]$. Der Fall $x < x_0$ ist analog, der Satz ist bewiesen. □

Beispiel 21.1 Betrachte für $x \in (-1, 1)$ die Funktion $f(x) = (1 - x)^{-1/2}$, mit Ableitungen

$$f'(x) = \frac{1}{2}(1 - x)^{-3/2} \quad \text{und} \quad f''(x) = \frac{3}{4}(1 - x)^{-5/2}.$$

Es gilt $f(0) = 1$ und $f'(0) = 1/2$, also lautet das Taylorpolynom der Ordnung Eins in $x_0 = 0$

$$P_1(x) = f(0) + f'(0)x = 1 + \frac{1}{2}x,$$

mit der Lagrange-Restglieddarstellung

$$R_1(x) = \frac{f''(\xi)}{2} x^2 = \frac{3}{8} (1 - \xi)^{-5/2} x^2 \quad \text{für ein } \xi \in [0, x].$$

Als Anwendung erhalten wir für die relativistische Energie eines Teilchens mit Ruhemasse m_0 und Geschwindigkeit v , wenn wir $\beta = v/c$ setzen,

$$E = \frac{m_0 c^2}{\sqrt{1 - \beta^2}} = m_0 c^2 \left(1 + \frac{1}{2} \beta^2 + \frac{f''(\xi)}{2} \beta^4 \right) = m_0 c^2 + \frac{1}{2} m_0 v^2 + \Delta E.$$

Dabei ist der erste Term die Ruheenergie und der zweite die klassische kinetische Energie. Für den relativistischen Korrekturterm ergibt sich aus der Restglieddarstellung die Abschätzung

$$\frac{\Delta E}{E_{kin}} = f''(\xi) \beta^2 \leq f''(\beta^2) \beta^2 < 0,008 \quad \text{für } \beta \leq 0,1.$$

Bei Geschwindigkeiten $v \leq \frac{1}{10} c$ beträgt die relativistische Korrektur weniger als ein Prozent der klassischen kinetischen Energie.

Allgemein gilt folgende Approximationseigenschaft des Taylorpolynoms.

Satz 21.3 (Approximation durch das Taylorpolynom) Sei $f \in C^k(I)$ für $k \in \mathbb{N}_0$, und P_k das k -te Taylorpolynom von f mit Entwicklungspunkt $x_0 \in I$. Dann ist P_k das eindeutig bestimmte Polynom vom Grad höchstens k mit

$$\lim_{x \rightarrow x_0} \frac{f(x) - P_k(x)}{(x - x_0)^k} = 0.$$

BEWEIS: Nach Satz 21.2 gibt es zu $x \in I$ ein ξ zwischen x_0 und x mit

$$\frac{f(x) - P_k(x)}{(x - x_0)^k} = \frac{f(x) - P_{k-1}(x)}{(x - x_0)^k} - \frac{1}{k!} f^{(k)}(x_0) = \frac{1}{k!} (f^{(k)}(\xi) - f^{(k)}(x_0)).$$

Da $f^{(k)}$ stetig, ist $|f^{(k)}(\xi) - f^{(k)}(x_0)| < \varepsilon$ für $|x - x_0| < \delta$, womit die Konvergenz gegen Null bewiesen ist. Für die Eindeutigkeit ist zu zeigen, dass für $P(x) = \sum_{j=0}^k a_j (x - x_0)^j$ gilt:

$$\lim_{x \rightarrow x_0} \frac{P(x)}{(x - x_0)^k} = 0 \quad \Rightarrow \quad a_0 = \dots = a_k = 0.$$

Sei induktiv schon $a_0 = \dots = a_{j-1} = 0$ gezeigt mit $0 \leq j \leq k$. Dann folgt

$$a_j = \lim_{x \rightarrow x_0} (x - x_0)^{-j} P(x) = \lim_{x \rightarrow x_0} (x - x_0)^{k-j} (x - x_0)^{-k} P(x) = 0.$$

□

Wir haben bis jetzt die Differenz $f(x) - P_k(x)$ für festes k und $x \rightarrow x_0$ untersucht. Jetzt nehmen wir einen anderen Standpunkt ein und fragen uns, ob die Folge $P_k(x)$ die Funktion $f(x)$ für $k \rightarrow \infty$ approximiert.

Definition 21.1 Für $f \in C^\infty(I)$ und $x_0 \in I$ heißt die Reihe

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

Taylorreihe von f mit Entwicklungspunkt x_0 .

Die Taylorreihe ist eine Potenzreihe, mit Entwicklungspunkt $x_0 \in \mathbb{R}$. Nach dem Satz vom Konvergenzradius gibt ein $R \in [0, \infty]$, so dass die Reihe für $|x - x_0| < R$ absolut konvergiert, für $|x - x_0| > R$ dagegen divergiert, siehe Analysis I, Satz 13.1. Selbst wenn die Reihe einen positiven Konvergenzradius hat, muss sie aber keineswegs gegen die gegebene Funktion konvergieren.

Beispiel 21.2 Betrachte $f : \mathbb{R} \rightarrow \mathbb{R}$ mit

$$f(x) = \begin{cases} e^{-\frac{1}{x}} & \text{für } x > 0 \\ 0 & \text{für } x \leq 0 \end{cases}$$

Es gilt $f \in C^\infty(\mathbb{R})$ nach Folgerung 11.2, Analysis I, insbesondere $f^{(j)}(0) = 0$ für alle $j \in \mathbb{N}_0$. Also sind die Koeffizienten der Taylorreihe alle Null und damit auch alle Partialsummen, die Reihe konvergiert somit gegen die Nullfunktion, nicht gegen f . Übrigens hätten wir dies auch aus dem Identitätssatz für Potenzreihen, Satz 13.7, Analysis I, schließen können, denn die Menge $f^{-1}\{0\}$ hat einen Häufungspunkt in $x = 0$.

Definition 21.2 (analytische Funktion) $f : I \rightarrow \mathbb{R}$, I offen, heißt *analytisch*, wenn jedes $x_0 \in I$ eine Umgebung $(x_0 - \delta, x_0 + \delta)$ hat, auf der f durch eine konvergente Potenzreihe mit Entwicklungspunkt x_0 dargestellt wird:

$$f(x) = \sum_{k=0}^{\infty} a_k (x - x_0)^k \quad \text{für alle } x \in (x_0 - \delta, x_0 + \delta).$$

Die darstellende Potenzreihe kann nur die Taylorreihe sein. Denn f ist C^∞ auf $(x_0 - \delta, x_0 + \delta)$ nach Satz 13.5, Analysis I, und gliedweise Differentiation der Reihe ergibt

$$f^{(i)}(x_0) = \sum_{k=0}^{\infty} \delta_{ik} k! a_k = i! a_i.$$

Eine Funktion ist also genau dann analytisch, wenn sie C^∞ ist und für jedes $x_0 \in I$ gilt: die Taylorreihe mit Entwicklungspunkt x_0 konvergiert punktweise gegen f nahe bei x_0 .

Die mehrdimensionale Taylorentwicklung orientiert sich am Fall $n = 1$, nur ist der Notationsaufwand größer. Sei $\Omega \subset \mathbb{R}^n$ offen und konvex. Für $f \in C^k(\Omega)$ definieren wir die k -te Ableitung $D^k f(x)$ im Punkt $x \in \Omega$ als k -Linearform $D^k f(x) : \mathbb{R}^n \times \dots \times \mathbb{R}^n \rightarrow \mathbb{R}$, wobei

$$(21.5) \quad D^k f(x)(v_1, \dots, v_k) = \sum_{i_1, \dots, i_k=1}^n (\partial_{i_1 \dots i_k}^k f)(x)(v_1)_{i_1} \dots (v_k)_{i_k}.$$

Betrachte jetzt für $x_0, x \in \Omega$ die C^k -Funktion, vgl. Folgerung 18.1,

$$\varphi : [0, 1] \rightarrow \mathbb{R}, \varphi(t) = f(x_0 + th) \quad \text{mit } h = x - x_0.$$

Wir zeigen durch Induktion die Formel

$$(21.6) \quad \varphi^{(k)}(t) = D^k f(x_0 + th)(h, \dots, h).$$

Für $k = 1$ gilt das nach Kettenregel und Satz 18.1, denn

$$\varphi'(t) = Df(x_0 + th)h = \sum_{i=1}^n \partial_i f(x_0 + th)h_i.$$

Für $k \geq 2$ ergibt sich induktiv mit Satz 17.2 von Schwarz

$$\begin{aligned} \varphi^{(k)}(t) &= \frac{d}{dt} \sum_{i_1, \dots, i_{k-1}=1}^n (\partial_{i_1 \dots i_{k-1}}^{k-1} f)(x_0 + th) h_{i_1} \dots h_{i_{k-1}} \\ &= \sum_{i_1, \dots, i_{k-1}=1}^n \sum_{i=1}^n (\partial_{i_1 \dots i_{k-1} i}^k f)(x_0 + th) h_{i_1} \dots h_{i_{k-1}} h_i \\ &= D^k f(x_0 + th)(h, \dots, h). \end{aligned}$$

Satz 21.2, angewandt auf die Funktion φ , liefert sofort eine erste Fassung der mehrdimensionalen Taylorentwicklung.

Lemma 21.2 Sei $\Omega \subset \mathbb{R}^n$ offen und konvex, und sei $f \in C^{k+1}(\Omega)$ für ein $k \in \mathbb{N}_0$. Dann gibt es zu $x_0, x \in \Omega$ ein $\xi = (1 - \tau)x_0 + \tau x$, $\tau \in [0, 1]$, so dass mit $h = x - x_0$ gilt:

$$f(x) = \sum_{j=0}^k \frac{D^j f(x_0)(h, \dots, h)}{j!} + \frac{D^{k+1} f(\xi)(h, \dots, h)}{(k+1)!}.$$

BEWEIS: Wir wenden auf die C^{k+1} -Funktion $\varphi(t) = f(x_0 + th)$ die eindimensionale Taylorsche Formel an, mit Entwicklungspunkt $t_0 = 0$. Nach Satz 21.2 gibt es ein $\tau \in [0, 1]$ mit

$$\varphi(1) = \sum_{j=0}^k \frac{\varphi^{(j)}(0)}{j!} + \frac{\varphi^{(k+1)}(\tau)}{(k+1)!}.$$

Einsetzen von (21.6) liefert die Behauptung. □

Die k -te Ableitung $D^k f(x)(h, \dots, h)$ ist eine Summe von n^k Termen, von denen viele aber gleich sind wegen der Vertauschbarkeit der partiellen Ableitungen. Es ist ökonomischer, die Summe danach zu ordnen, wie oft nach den einzelnen Variablen x_1, \dots, x_n differenziert wird. Gleichzeitig führt das wie im Eindimensionalen auf eine Taylordarstellung mit Basispolynomen. Für einen Multiindex $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{N}_0^n$ setzen wir

$$\begin{aligned} |\alpha| &= \alpha_1 + \dots + \alpha_n && \text{Ordnung von } \alpha, \\ \alpha! &= (\alpha_1)! \dots (\alpha_n)! && \alpha\text{-Fakultät}, \\ x^\alpha &= x_1^{\alpha_1} \dots x_n^{\alpha_n} && \text{Monom mit Exponent } \alpha, \\ D^\alpha &= \partial_1^{\alpha_1} \dots \partial_n^{\alpha_n} && (D^0 = \text{Id}). \end{aligned}$$

Im Operator D^α wird also α_i mal nach x_i differenziert.

Satz 21.4 (Taylorentwicklung im \mathbb{R}^n) Sei $\Omega \subset \mathbb{R}^n$ offen und konvex, und sei $f \in C^{k+1}(\Omega)$ für ein $k \in \mathbb{N}_0$. Dann gibt es zu $x_0, x \in \Omega$ ein $\xi = (1 - \tau)x_0 + \tau x$, $\tau \in [0, 1]$, so dass gilt:

$$f(x) = \sum_{|\alpha| \leq k} \frac{D^\alpha f(x_0)}{\alpha!} (x - x_0)^\alpha + \sum_{|\alpha|=k+1} \frac{D^\alpha f(\xi)}{\alpha!} (x - x_0)^\alpha.$$

BEWEIS: Sei α Multiindex der Ordnung $|\alpha| = k$. Wieviele Tupel (i_1, \dots, i_k) gibt es, in denen jedes $i \in \{1, \dots, n\}$ genau α_i mal vorkommt? Wähle α_1 Stellen für $i = 1$, aus den übrigen α_2 Stellen für $i = 2$, etc. Das ergibt die Zahl

$$\binom{k}{\alpha_1} \cdot \binom{k - \alpha_1}{\alpha_2} \cdot \dots \cdot \binom{k - (\alpha_1 + \dots + \alpha_{n-1})}{\alpha_n} = \frac{k!}{(\alpha_1)! \dots (\alpha_n)!} = \frac{k!}{\alpha!}.$$

Die behauptete Entwicklung folgt nun aus Lemma 21.2 und der Vertauschbarkeit der partiellen Ableitungen, siehe Folgerung 17.1. \square

Sei $I(n, k)$ die Menge aller n -Multiindizes mit $|\alpha| = k$. Es gilt dann

$$I(n + 1, k) = \bigcup_{\ell=0}^k \{(\alpha, \ell) : \alpha \in I(n, k - \ell)\}.$$

Also gilt $\#I(n + 1, k) = \sum_{\ell=0}^k \#I(n, k - \ell)$, und induktiv folgt leicht $\#I(n, k) \leq (k + 1)^{n-1}$. Für große k ist das viel kleiner als die Zahl n^k aus der vorigen Darstellung der Taylorformel.

Eine Funktion $P : \mathbb{R}^n \rightarrow \mathbb{R}$ heißt Polynom vom Grad $k \geq 0$, wenn es $a_\alpha \in \mathbb{R}$, $|\alpha| \leq k$, gibt mit $a_\alpha \neq 0$ für mindestens ein $|\alpha| = k$, so dass gilt:

$$P(x) = \sum_{|\alpha| \leq k} a_\alpha x^\alpha \quad \text{für alle } x \in \mathbb{R}^n.$$

Für $x_0 \in \mathbb{R}^n$ beliebig bilden die Monome $(x - x_0)^\alpha$ mit $0 \leq |\alpha| \leq k$ eine Basis des Raums \mathbb{P}_k der Polynome vom Grad $\leq k$. Dies folgt wie für $n = 1$ aus der Ableitungsregel

$$P(x) = \sum_{|\beta| \leq k} a_\beta (x - x_0)^\beta \quad \Rightarrow \quad D^\alpha P(x_0) = \alpha! a_\alpha \quad \text{für } |\alpha| \leq k.$$

Es folgt analog Lemma 21.1): das k -te Taylorpolynom

$$(21.7) \quad P_k(x) = \sum_{|\alpha| \leq k} \frac{D^\alpha f(x_0)}{\alpha!} (x - x_0)^\alpha$$

ist das eindeutige Polynom vom Grad höchstens k mit $D^\alpha P(x_0) = D^\alpha f(x_0)$ für $|\alpha| \leq k$.

Folgerung 21.2 *Das k -te Taylorpolynom mit Entwicklungspunkt x_0 eines Polynoms f vom Grad höchstens k ist f selbst.*

Beispiel 21.3 (Polynomialformel) Die Funktion $f(x) = (x_1 + \dots + x_n)^k$ ist ein Polynom vom Grad k , und es gilt

$$D^\alpha f(0) = \begin{cases} k! & \text{falls } |\alpha| = k, \\ 0 & \text{sonst.} \end{cases}$$

Mit Folgerung 21.2 (oder direkt durch Abzählen) ergibt sich

$$(x_1 + \dots + x_n)^k = \sum_{|\alpha|=k} \frac{k!}{\alpha!} x^\alpha.$$

Auch im Mehrdimensionalen approximiert das k -te Taylorpolynom bis auf Terme höherer Ordnung, nur müssen wir jetzt im Nenner Beträge setzen, da bekanntlich durch Vektoren nicht dividiert werden kann.

Satz 21.5 (Approximation durch das Taylorpolynom im \mathbb{R}^n) Sei $f \in C^k(\Omega)$ für $k \in \mathbb{N}_0$, und P_k das k -te Taylorpolynom von f mit Entwicklungspunkt $x_0 \in \Omega$. Dann ist P_k das eindeutig bestimmte Polynom vom Grad höchstens k mit

$$\lim_{x \rightarrow x_0} \frac{f(x) - P_k(x)}{|x - x_0|^k} = 0.$$

BEWEIS: Nach Satz 21.4, mit k statt $k + 1$, gibt es zu $x \in \Omega$ ein ξ zwischen x_0 und x mit

$$f(x) - P_k(x) = \sum_{|\alpha|=k} \frac{D^\alpha f(\xi) - D^\alpha f(x_0)}{\alpha!} (x - x_0)^\alpha.$$

Da $D^\alpha f$ stetig und $|(x - x_0)^\alpha| \leq |x - x_0|^k$, folgt die Konvergenz gegen Null. Für die Eindeutigkeit zeigen wir für ein beliebiges Polynom $P(x)$ vom Grad $\leq k$ folgende Implikation:

$$\lim_{x \rightarrow x_0} \frac{P(x)}{|x - x_0|^k} = 0 \quad \Rightarrow \quad P(x) = 0 \text{ für alle } x \in \mathbb{R}^n.$$

Angenommen es gibt $x_1 \in \mathbb{R}^n$ mit $P(x_1) \neq 0$, oBdA $x_1 \neq x_0$. Mit $x_t = x_0 + t(x_1 - x_0)$, $t \in \mathbb{R}$, ist $\varphi(t) = P(x_t)$ eindimensionales Polynom vom Grad $\leq k$, und wegen $\lim_{t \rightarrow 0} x_t = x_0$ folgt

$$\frac{\varphi(t)}{|t|^k} = |x_1 - x_0|^k \frac{P(x_t)}{|t(x_1 - x_0)|^k} = |x_1 - x_0|^k \frac{P(x_t)}{|x_t - x_0|^k} \rightarrow 0 \quad \text{mit } t \rightarrow 0.$$

Nach Beweis von Satz 21.3 ist $\varphi(t) \equiv 0$, im Widerspruch zu $\varphi(1) = P(x_1) \neq 0$. □

Um das relative Verhalten von Funktionen bei Grenzprozessen zu beschreiben, werden oft die Landauschen Symbole \mathcal{O} und o benutzt. Seien f, g zwei Funktionen, die auf $B_\delta(x_0)$ definiert sind, und es gelte $g(x) \neq 0$ für x nahe bei x_0 . Dann schreibt man

$$\begin{aligned} f(x) = o(g(x)) \text{ für } x \rightarrow x_0 &\Leftrightarrow \lim_{x \rightarrow x_0} \frac{|f(x)|}{|g(x)|} = 0, \\ f(x) = \mathcal{O}(g(x)) \text{ für } x \rightarrow x_0 &\Leftrightarrow \limsup_{x \rightarrow x_0} \frac{|f(x)|}{|g(x)|} < \infty. \end{aligned}$$

In Worten: die Funktion $f(x)$ ist klein- o von $g(x)$ beziehungsweise groß- \mathcal{O} von $g(x)$ für $x \rightarrow x_0$. Diese Begriffe sind analog für Grenzwerte $|x| \rightarrow \infty$ usw. erklärt. Im obigen Approximationssatz gilt $f(x) - P_k(x) = o(|x - x_0|^k)$ für $x \rightarrow x_0$, aber häufig wird das in Form einer Entwicklung geschrieben:

$$f(x) = P_k(x) + o(|x - x_0|^k) \quad \text{für } x \rightarrow x_0.$$

Beispiel 21.4 Wir berechnen hier mit der Multiindexnotation die Taylorentwicklung erster Ordnung im Punkt $(1, 1)$ für

$$f(x, y) = \frac{x - y}{x + y}.$$

Es ist $f(1, 1) = 0$, und die partiellen Ableitungen der Funktion lauten

$$\begin{aligned} D^{(1,0)}f(x, y) &= \frac{2y}{(x+y)^2} & D^{(0,1)}f(x, y) &= -\frac{2x}{(x+y)^2} \\ D^{(2,0)}f(x, y) &= -\frac{4y}{(x+y)^3} & D^{(1,1)}f(x, y) &= \frac{2(x-y)}{(x+y)^3} & D^{(0,2)}f(x, y) &= \frac{4x}{(x+y)^3}. \end{aligned}$$

Das Taylorpolynom erster Ordnung ist somit

$$\begin{aligned} P_1(x, y) &= f(1, 1) + D^{(1,0)}f(1, 1)((x, y) - (1, 1))^{(1,0)} + D^{(0,1)}f(1, 1)((x, y) - (1, 1))^{(0,1)} \\ &= \frac{1}{2}(x-1) - \frac{1}{2}(y-1) = \frac{1}{2}(x-y). \end{aligned}$$

Das Restglied lautet in Lagrangedarstellung mit Zwischenpunkt (ξ, η)

$$\begin{aligned} R_1(x, y) &= \frac{D^{(2,0)}f(\xi, \eta)}{2!0!}((x, y) - (1, 1))^{(2,0)} + \frac{D^{(1,1)}f(\xi, \eta)}{1!1!}((x, y) - (1, 1))^{(1,1)} \\ &\quad + \frac{D^{(0,2)}f(\xi, \eta)}{0!2!}((x, y) - (1, 1))^{(0,2)} \\ &= \frac{2}{(\xi + \eta)^3}(-\eta(x-1)^2 + (\xi - \eta)(x-1)(y-1) + \xi(y-1)^2). \end{aligned}$$

Eine Funktion $f : \Omega \rightarrow \mathbb{R}$, die in der Nähe jedes Punkts $x_0 \in \Omega$ durch eine Potenzreihe

$$P(x) = \sum_{k=0}^{\infty} \sum_{|\alpha|=k} a_{\alpha}(x - x_0)^{\alpha} \quad \text{mit } a_{\alpha} \in \mathbb{R}$$

dargestellt werden können, heißt reell-analytisch. Unsere eindimensionalen Überlegungen lassen sich auch in diesem Punkt verallgemeinern, worauf wir jedoch aus Zeitgründen verzichten.

22 Parameterabhängige Integrale

In diesem Abschnitt behandeln wir Integrale, deren Integranden von zusätzlichen Parametern abhängen. Dies ist eine typische Problematik in zahlreichen Anwendungen. Sei $\Omega \subset \mathbb{R}^n$ offen und $I = [a, b]$ ein kompaktes Intervall. Für eine gegebene Funktion $f : \Omega \times I \rightarrow \mathbb{R}$, $f = f(x, y)$, betrachten wir die neue Funktion

$$(22.8) \quad \phi : \Omega \rightarrow \mathbb{R}, \quad \phi(x) = \int_I f(x, y) dy.$$

Diese Funktion wird als parameterabhängiges Integral bezeichnet, wobei die Parameter hier die Punkte $x = (x_1, \dots, x_n) \in \Omega$ sind. Damit ϕ wohldefiniert ist, müssen die Integrale existieren, also sollte für jedes $x \in \Omega$ die Funktion $f(x, \cdot) : I \rightarrow \mathbb{R}$, $y \mapsto f(x, y)$, Riemann-integrierbar sein. Wir interessieren uns für die Stetigkeit und Ableitung der Funktion $\phi(x)$. Ein nützlicher Begriff ist dabei die Oszillationsfunktion einer Funktion $f : D \rightarrow \mathbb{R}$:

$$\text{osc}(f, \delta) = \sup\{|f(x) - f(x')| : x, x' \in D, |x - x'| < \delta\}.$$

Die Funktion f ist genau dann gleichmäßig stetig, wenn $\lim_{\delta \rightarrow 0} \text{osc}(f, \delta) = 0$.

Satz 22.1 (Stetigkeit von Parameterintegralen) Sei $f : \Omega \times I \rightarrow \mathbb{R}$, $f = f(x, y)$, wobei $\Omega \subset \mathbb{R}^n$ offen und $I = [a, b]$ kompakt. Ist $f \in C^0(\Omega \times I)$, so ist die Funktion

$$\phi : \Omega \rightarrow \mathbb{R}, \quad \phi(x) = \int_I f(x, y) dy,$$

wohldefiniert und stetig.

BEWEIS: Die Funktion $\phi(x)$ ist wohldefiniert, denn für $x \in \Omega$ ist $f(x, \cdot) \in C^0(I)$ und damit Riemann-integrierbar (siehe Satz 14.5). Wir berechnen für $x, x' \in \Omega$

$$|\phi(x') - \phi(x)| \leq \int_I |f(x', y) - f(x, y)| dy \leq |I| \sup_{y \in I} |f(x', y) - f(x, y)|.$$

Wähle $R > 0$ mit $\overline{B_R(x)} \subset \Omega$. Dann ist $f(x, y)$ gleichmäßig stetig auf $\overline{B_R(x)} \times I$, vgl. Satz 14.4, insbesondere folgt für $|x' - x| < \delta \leq R$

$$\sup_{y \in I} |f(x', y) - f(x, y)| \leq \text{osc}(f|_{\overline{B_R(x)} \times I}, \delta) \rightarrow 0 \quad \text{mit } \delta \rightarrow 0.$$

□

Wir gehen direkt weiter zur Differenzierbarkeit und Berechnung der Ableitung.

Satz 22.2 (Differentiation unter dem Integral) Sei $f : \Omega \times I \rightarrow \mathbb{R}$, $f = f(x, y)$, wobei $\Omega \subset \mathbb{R}^n$ offen und $I = [a, b]$ kompakt. Es gelte:

- (a) $f(x, \cdot)$ ist Riemann-integrierbar für jedes $x \in \Omega$.
- (b) Die partielle Ableitung $\frac{\partial f}{\partial x_j}$ existiert und ist stetig auf $\Omega \times I$.

Dann ist $\phi : \Omega \rightarrow \mathbb{R}$, $\phi(x) = \int_I f(x, y) dy$, nach x_j partiell differenzierbar, und zwar gilt

$$\frac{\partial \phi}{\partial x_j}(x) = \int_I \frac{\partial f}{\partial x_j}(x, y) dy \quad \text{für alle } x \in \Omega.$$

Sind f und $\frac{\partial f}{\partial x_1}, \dots, \frac{\partial f}{\partial x_n}$ in $C^0(\Omega \times I)$, so ist $\phi \in C^1(\Omega)$.

BEWEIS: Nach dem Hauptsatz der Differential- und Integralrechnung gilt

$$\frac{f(x + he_j, y) - f(x, y)}{h} = \frac{1}{h} \int_0^1 \frac{d}{ds} f(x + she_j, y) ds = \int_0^1 \frac{\partial f}{\partial x_j}(x + she_j, y) ds.$$

Sei wieder $\overline{B_R(x)} \subset \Omega$ und $|h| < \delta \leq R$, dann schätzen wir wie folgt ab:

$$\begin{aligned} \left| \frac{\phi(x + he_j) - \phi(x)}{h} - \int_I \frac{\partial f}{\partial x_j}(x, y) dy \right| &= \left| \int_I \left(\frac{f(x + he_j, y) - f(x, y)}{h} - \frac{\partial f}{\partial x_j}(x, y) \right) dy \right| \\ &= \left| \int_I \int_0^1 \left(\frac{\partial f}{\partial x_j}(x + she_j, y) - \frac{\partial f}{\partial x_j}(x, y) \right) ds dy \right| \\ &\leq \int_I \int_0^1 \left| \frac{\partial f}{\partial x_j}(x + she_j, y) - \frac{\partial f}{\partial x_j}(x, y) \right| ds dy \\ &\leq |I| \operatorname{osc} \left(\frac{\partial f}{\partial x_j} \Big|_{\overline{B_R(x)} \times I}, \delta \right) \rightarrow 0 \quad \text{mit } \delta \rightarrow 0. \end{aligned}$$

Damit ist die Vertauschung der Ableitung mit dem Integral gerechtfertigt. Satz 22.1 impliziert weiter $\frac{\partial \phi}{\partial x_j} \in C^0(\Omega)$. Sind nun f und alle $\frac{\partial f}{\partial x_j}$ in $C^0(\Omega \times I)$, so folgt $\phi, \frac{\partial \phi}{\partial x_j} \in C^0(\Omega)$ aus Satz 22.1 bzw. wie gerade gezeigt. \square

Beispiel 22.1 Wir berechnen hier das Integral der Gaußschen Dichtefunktion (das früher auf 10-Mark-Scheinen zu finden war)

$$\int_{-\infty}^{\infty} e^{-x^2} dx = \sqrt{\pi}.$$

Der Beweis ist trickreich, ich wäre wohl selbst nicht drauf gekommen. Setze

$$F : [0, \infty) \rightarrow \mathbb{R}, F(x) = \left(\int_0^x e^{-\xi^2} d\xi \right)^2,$$

und berechne mit Hauptsatz und anschließender Substitution $\xi = xy$, also $d\xi = xdy$,

$$F'(x) = 2e^{-x^2} \int_0^x e^{-\xi^2} d\xi = \int_0^1 2xe^{-(1+y^2)x^2} dy = \int_0^1 \frac{\partial f}{\partial x}(x, y) dy,$$

wobei $f(x, y) = -e^{-(1+y^2)x^2} / (1+y^2)$. Da f auf $(0, \infty) \times [0, 1]$ glatt ist, können wir nach Satz 22.2 den Operator $\frac{\partial}{\partial x}$ herausziehen, und mit $\phi(x) = \int_0^1 f(x, y) dy$ folgt

$$\phi'(x) = \int_0^1 \frac{\partial f}{\partial x}(x, y) dy = F'(x).$$

Nun gilt $F(0) - \phi(0) = \int_0^1 (1+y^2)^{-1} dy = \arctan 1 = \pi/4$, also $F(x) = \phi(x) + \pi/4$ für alle $x \in [0, \infty)$. Aber $|\phi(x)| \leq e^{-x^2} \rightarrow 0$ mit $x \rightarrow \infty$, und so

$$\int_0^{\infty} e^{-x^2} dx = \lim_{x \rightarrow \infty} \sqrt{F(x)} = \frac{\sqrt{\pi}}{2}.$$

In dieser Vorlesung werden wir aus Zeitgründen kein mehrdimensionales Integral behandeln, dies soll in Analysis 3 ausführliches Thema sein. Immerhin können wir als nützliche Anwendung hier die Vertauschbarkeit der Integrationsreihenfolge in Mehrfachintegralen folgern.

Satz 22.3 (Kleiner Fubini) Seien $I = [a, b]$, $J = [\alpha, \beta]$ kompakte Intervalle. Dann gilt

$$\int_{\alpha}^{\beta} \left(\int_a^b f(x, y) dx \right) dy = \int_a^b \left(\int_{\alpha}^{\beta} f(x, y) dy \right) dx \quad \text{für } f \in C^0(I \times J).$$

BEWEIS: Wir betrachten die Funktionen $\phi, \psi : [a, b] \rightarrow \mathbb{R}$ mit

$$\phi(x) = \int_{\alpha}^{\beta} \left(\int_a^x f(\xi, y) d\xi \right) dy \quad \text{und} \quad \psi(x) = \int_a^x \left(\int_{\alpha}^{\beta} f(\xi, y) dy \right) d\xi.$$

Nach Satz 22.1 sind $y \mapsto \int_a^x f(\xi, y) d\xi$ sowie $\xi \mapsto \int_{\alpha}^{\beta} f(\xi, y) dy$ stetig, und damit beide Seiten wohldefiniert mit $\phi(a) = \psi(a) = 0$. Wir zeigen $\phi'(x) = \psi'(x)$ für alle $x \in I$, woraus die Behauptung $\phi(b) = \psi(b)$ folgt. Der Hauptsatz der Differential- und Integralrechnung liefert

$$\psi'(x) = \int_{\alpha}^{\beta} f(x, y) dy.$$

Weiter hat die Funktion $F(x, y) = \int_a^x f(\xi, y) d\xi$ die partielle Ableitung $\frac{\partial F}{\partial x} = f \in C^0(I \times J)$, und aus Satz 22.2 folgt

$$\phi'(x) = \int_{\alpha}^{\beta} \frac{\partial F}{\partial x}(x, y) dy = \int_{\alpha}^{\beta} f(x, y) dy.$$

□

Alternativ kann der kleine Fubini auch durch Approximation mit Riemannschen Summen in beiden Variablen bewiesen werden.

Wir kommen jetzt zu einer Anwendung in der Variationsrechnung, und zwar betrachten wir Integrale des folgenden Typs:

$$\mathcal{F} : C^1(I, \mathbb{R}^n) \rightarrow \mathbb{R}, \quad \mathcal{F}(u) = \int_a^b f(t, u(t), u'(t)) dt.$$

Abstrakt ist \mathcal{F} eine reelle Funktion auf dem Raum $C^1(I, \mathbb{R}^n)$. Da es sich aber nicht um eine Funktion von endlich vielen reellen Variablen handelt, wie wir sie bisher hatten, wird meistens die Bezeichnung *Funktional* oder *Variationsintegral* benutzt. Das Funktional \mathcal{F} ist dabei definiert durch die *Lagrangefunktion*

$$f : I \times \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}, \quad f = f(t, x, v).$$

Hier ein paar Beispiele.

Beispiel 22.2 Die Bogenlänge von $u \in C^1(I, \mathbb{R}^n)$ ist das Funktional

$$\mathcal{F}(u) = \int_a^b |u'(t)| dt, \quad \text{also } f(t, x, v) = |v|.$$

Die Formel kann wie folgt motiviert werden: für eine Zerlegung $a = t_0 < \dots < t_N$ ergibt sich als Näherung der Bogenlänge

$$\sum_{i=1}^N |u(t_i) - u(t_{i-1})| \approx \sum_{i=1}^N |u'(t_i)| \Delta t_i.$$

Rechts steht aber die Riemannsche Summe für die Funktion $|u'(t)|$.

Beispiel 22.3 Soll der Kalorienbedarf beim Querfeldeinlauf ermittelt werden, so spielt nicht nur die Länge der Strecke eine Rolle, sondern auch die wechselnde Qualität des Bodens. Dies könnte durch eine Gewichtsfunktion als Faktor beschrieben werden:

$$\mathcal{F}(u) = \int_a^b \omega(u(t)) |u'(t)| dt, \quad \text{also } f(t, x, v) = \omega(x)|v|.$$

Beispiel 22.4 Hier beschreibt $u \in C^1(I, \mathbb{R}^n)$ die Bahn eines Teilchens der Masse m in einem Kraftfeld mit Potential $V : \mathbb{R}^n \rightarrow \mathbb{R}$. Es ist dann $\frac{m}{2}|u'(t)|^2$ die kinetische und $V(u(t))$ die potentielle Energie des Teilchens zur Zeit t . Das Wirkungsintegral bildet die Differenz aus kinetischer und potentieller Energie, integriert auf I :

$$\mathcal{F}(u) = \int_a^b \left(\frac{m}{2}|u'(t)|^2 - V(u(t)) \right) dt, \quad \text{also } f(t, x, v) = \frac{m}{2}|v|^2 - V(x).$$

Funktionen, für die ein Funktional einen kleinsten oder größten Wert annimmt, sind natürlich von zentralem Interesse. Wir werden zeigen, dass eine extremale Funktion eine gewisse Differentialgleichung erfüllt, die Euler-Lagrange Gleichung. Unser Ansatz besteht darin, Variationen $u(\varepsilon, t)$ der extremalen Funktion $u(t)$ zu betrachten, die von einem Parameter abhängen:

$$u : (-\varepsilon_0, \varepsilon_0) \times I \rightarrow \mathbb{R}^n, \quad u = u(\varepsilon, t), \quad \text{wobei } u(0, \cdot) = u.$$

Ableitung der Variation nach dem Parameter ergibt das zugehörige Vektorfeld

$$\varphi : I \rightarrow \mathbb{R}^n, \quad \varphi(t) = \frac{\partial u}{\partial \varepsilon}(0, t).$$

Lemma 22.1 (Erste Variation) Sei $f = f(t, x, v)$ Lagrangefunktion mit f und $D_v f$ stetig differenzierbar auf $I \times \mathbb{R}^n \times \mathbb{R}^n$. Für $u \in C^2((-\varepsilon_0, \varepsilon_0) \times I, \mathbb{R}^n)$ betrachte

$$\phi(\varepsilon) = \mathcal{F}(u(\varepsilon, \cdot)) = \int_a^b f(t, u(\varepsilon, t), \frac{\partial u}{\partial t}(\varepsilon, t)) dt.$$

Dann gilt mit $\varphi = \frac{\partial u}{\partial \varepsilon}(0, \cdot) : I \rightarrow \mathbb{R}^n$ die Formel

$$(22.9) \quad \frac{d\phi}{d\varepsilon}(0) = \int_a^b \langle L_f(u), \varphi \rangle dt + \left[\langle D_v f(t, u, u'), \varphi \rangle \right]_{t=a}^{t=b}.$$

Dabei ist $\langle \cdot, \cdot \rangle$ das Standardskalarprodukt im \mathbb{R}^n und $L_f(u) : I \rightarrow \mathbb{R}^n$ ist gegeben durch

$$L_f(u) = D_x f(t, u, u') - \frac{d}{dt} [D_v f(t, u, u')].$$

BEWEIS: Die Verkettung $(\varepsilon, t) \mapsto f(t, u(\varepsilon, t), \frac{\partial u}{\partial t}(\varepsilon, t))$ ist von der Klasse C^1 . Deshalb kann nach Satz 22.2 unter dem Integralzeichen differenziert werden. Es folgt mit der Kettenregel

$$\phi'(0) = \int_a^b \sum_{i=1}^n \left(\frac{\partial f}{\partial x_i}(t, u, u') \frac{\partial u_i}{\partial \varepsilon}(0, t) + \frac{\partial f}{\partial v_i}(t, u, u') \frac{\partial^2 u_i}{\partial \varepsilon \partial t}(0, t) \right) dt.$$

Mit $\varphi(t) = \frac{\partial u}{\partial \varepsilon}(0, t)$ folgt, indem wir hinten die Ableitungen vertauschen,

$$(22.10) \quad \phi'(0) = \int_a^b \sum_{i=1}^n \left(\frac{\partial f}{\partial x_i}(t, u, u') \varphi_i(t) + \frac{\partial f}{\partial v_i}(t, u, u') \varphi_i'(t) \right) dt.$$

Schließlich mit partieller Integration im hinteren Term

$$\phi'(0) = \int_a^b \sum_{i=1}^n \underbrace{\left(\frac{\partial f}{\partial x_i}(t, u, u') - \frac{d}{dt} \left[\frac{\partial f}{\partial v_i}(t, u, u') \right] \right)}_{= L_f(u)_i} \varphi_i dt + \left[\sum_{i=1}^n \frac{\partial f}{\partial v_i}(t, u, u') \varphi_i \right]_{t=a}^{t=b}.$$

Die Behauptung des Lemmas folgt, indem wir die Summen als Skalarprodukte schreiben. \square

Um eine optimale Funktion zu charakterisieren, müssen wir sie mit hinreichend vielen Variationen vergleichen. Das folgende Lemma gibt an, wieviele wir tatsächlich brauchen.

Lemma 22.2 Sei $I = (a, b)$. Für $f \in C^0(I, \mathbb{R}^n)$ gelte

$$(22.11) \quad \int_a^b \langle f, \varphi \rangle = 0 \quad \text{für alle } \varphi \in C_c^\infty(I, \mathbb{R}^n).$$

Dann ist f die Nullfunktion.

BEWEIS: Wir zeigen das Lemma erst für $n = 1$. Es gibt eine Funktion $\eta \in C^\infty(\mathbb{R})$ mit

$$\eta(s) = 0 \quad \text{für } |s| \geq 1, \quad \eta \geq 0 \quad \text{und} \quad \int_{\mathbb{R}} \eta = 1.$$

Ein konkretes Beispiel ist, bei passender Wahl von $a > 0$, die Funktion

$$\eta(s) = \begin{cases} a \exp \frac{1}{s^2-1} & \text{für } |s| < 1 \\ 0 & \text{sonst.} \end{cases}$$

Angenommen es ist $\varepsilon := f(t_0) > 0$ für ein $t_0 \in I$. Dann gibt es ein $\delta > 0$ mit $f(t) \geq \varepsilon/2$ für $t \in [t_0 - \delta, t_0 + \delta] \subset I$. Betrachte die reskalierte Funktion

$$\eta_{t_0, \delta}(t) = \frac{1}{\delta} \eta \left(\frac{t - t_0}{\delta} \right).$$

Dann gilt $\eta_{t_0, \delta}(t) = 0$ für $|t - t_0| \geq \delta$, sowie $\int_{\mathbb{R}} \eta_{t_0, \delta} = 1$. Nach Voraussetzung

$$0 = \int_a^b f(t) \eta_{t_0, \delta}(t) dt = \int_{t_0 - \delta}^{t_0 + \delta} f(t) \eta_{t_0, \delta}(t) dt \geq \frac{\varepsilon}{2} \int_a^b \eta_{t_0, \delta}(t) dt = \frac{\varepsilon}{2},$$

ein Widerspruch. Im Fall $f : I \rightarrow \mathbb{R}^n$ liefert die Voraussetzung, für alle $\varphi \in C_c^\infty(I)$,

$$0 = \int_a^b \langle f, \varphi e_i \rangle = \int_a^b f_i \varphi.$$

Aus obigem folgt $f_i = 0$ für $i = 1, \dots, n$, das Lemma ist bewiesen. \square

Satz 22.4 (Euler-Lagrange-Gleichungen) Sei $\mathcal{F}(u) = \int_a^b f(t, u, u')$ ein Variationsintegral mit $f \in C^2(I \times \mathbb{R}^n \times \mathbb{R}^n)$, $f = f(t, x, v)$. Sei $u \in C^2(I, \mathbb{R}^n)$ stationärer Punkt, d. h.

$$\frac{d}{d\varepsilon} \mathcal{F}(u + \varepsilon\varphi)|_{\varepsilon=0} = 0 \quad \text{für alle } \varphi \in C_c^\infty(I, \mathbb{R}^n).$$

Dann gelten die Euler-Lagrange-Gleichungen

$$L_f(u) = 0 \quad \Leftrightarrow \quad \frac{\partial f}{\partial x_i}(t, u, u') - \frac{d}{dt} \frac{\partial f}{\partial v_i}(t, u, u') = 0 \quad \text{für } i = 1, \dots, n.$$

Bemerkung. Es handelt sich um ein System von n Differentialgleichungen zweiter Ordnung, wie man durch Ausdifferenzieren des zweiten Terms sieht.

BEWEIS: Die Randterme in (22.1) verschwinden, da $\varphi(a) = \varphi(b) = 0$ nach Voraussetzung. Die Aussage folgt dann aus den Lemmas 22.1 und 22.2. \square

Beispiel 22.5 (Bogenlänge) Wir betrachten die Bogenlänge aus Beispiel 22.2

$$\mathcal{F}(u) = \int_a^b |u'(t)| dt \quad \text{für } u \in C^2([a, b], \mathbb{R}^n).$$

Die Lagrangefunktion und ihre Ableitung sind

$$f(t, x, v) = |v|, \quad D_v f(t, x, v) = \frac{v}{|v|} \quad \text{falls } v \neq 0.$$

Die Euler-Lagrange-Gleichungen lauten somit

$$L_f(u) = -\frac{d}{dt} \frac{u'}{|u'|} = 0.$$

Die Gleichung sagt aus, dass der Einheitstangentenvektor $\frac{u'}{|u'|}$ konstant ist. Es ist nicht schwer zu sehen, dass $u(t)$ dann die Strecke von $u(a)$ nach $u(b)$ durchläuft. Allerdings wird zur Herleitung der Euler-Lagrange Gleichungen gebraucht, dass $u'(t) \neq 0$ für alle $t \in [a, b]$.

Beispiel 22.6 Bewegung eines Massenpunkts in einem konservativen Kraftfeld:

$$\mathcal{F}(u) = \int_a^b \left(\frac{m}{2} |u'|^2 - V(c(t)) \right) dt, \quad f(t, x, v) = \frac{m}{2} |v|^2 - V(x).$$

Das zugehörige Kraftfeld ist gegeben durch $F(x) = -\text{grad } V(x)$; das Minuszeichen ist in der Physik üblich. Dann ergibt sich

$$\frac{\partial f}{\partial x_i}(t, x, v) = F_i(x), \quad \frac{\partial f}{\partial v_i}(t, x, v) = mv_i.$$

Die Euler-Lagrange-Gleichungen lauten somit $F(u) - mu'' = 0$, es sind die Newtonschen Bewegungsgleichungen.

Viele interessante Parameterintegrale sind uneigentliche Integrale, zum Beispiel bei der Definition der Gammafunktion oder der Fouriertransformation. Aus Zeitgründen können wir darauf jetzt nicht eingehen, werden aber Parameterintegrale nochmals innerhalb der Theorie des Lebesgue-Integrals im dritten Semester aufgreifen.

23 Kurvenintegrale

Wir kehren hier zur eindimensionalen Analysis zurück und interessieren uns zunächst für Kurven im \mathbb{R}^n . Man bezeichnet jede stetige Abbildung eines Intervalls in den \mathbb{R}^n als Kurve. Dies hört sich zunächst vernünftig an, allerdings lässt die Forderung der Stetigkeit noch Abbildungen zu, die anschaulich weit entfernt davon sind, eine Kurve zu sein. So hat G. Peano 1905 stetige Kurven konstruiert, die das Intervall $[0, 1]$ surjektiv auf die Fläche eines Dreiecks abbilden. Wir sollten also lieber etwas mehr verlangen als nur Stetigkeit.

Definition 23.1 (C^1 -Kurve) Ist $\gamma \in C^1(I, \mathbb{R}^n)$, I Intervall, so heißt γ C^1 -Kurve im \mathbb{R}^n .

Beispiel 23.1 Für $p, v \in \mathbb{R}^n$ mit $v \neq 0$ ist $\gamma(t) = p + tv$ eine parametrisierte Gerade. Der Fall $v = 0$ ist jedoch durch Definition 23.1 nicht ausgeschlossen. Die Kurve

$$\gamma : \mathbb{R} \rightarrow \mathbb{R}^2, \gamma(t) = (r \cos t, r \sin t) \quad \text{mit } r > 0,$$

durchläuft den Kreis vom Radius r um den Nullpunkt unendlich oft. Durch Hinzufügen einer dritten Komponente entsteht die Schraubenlinie

$$\gamma : \mathbb{R} \rightarrow \mathbb{R}^3, \gamma(t) = (r \cos t, r \sin t, at) \quad \text{für } a > 0.$$

Bei einem Umlauf wächst die dritte Komponente um die Höhe $2\pi a$. Das Bild der Kurve

$$\gamma : \mathbb{R} \rightarrow \mathbb{R}^2, \gamma(t) = (\cos t, \sin 2t)$$

sieht aus wie eine etwas deformierte Acht.

Definition 23.2 (Bogenlänge) Die Bogenlänge einer Kurve $\gamma \in C^1(I, \mathbb{R}^n)$, $I = [a, b]$, ist

$$L(\gamma) = \int_I |\gamma'|.$$

Eigentlich müsste diese Formel durch Approximation von γ mit Polygonzügen, deren Länge elementar definiert ist, begründet werden. Für eine Zerlegung $a = t_0 < \dots < t_N = b$ sollte näherungsweise gelten:

$$\sum_{i=1}^N |\gamma(t_i) - \gamma(t_{i-1})| = \sum_{i=1}^N \left| \int_{t_{i-1}}^{t_i} \gamma'(t) dt \right| \approx \sum_{i=1}^N |\gamma'(t_i)| \Delta t_i \approx \int_I |\gamma'(t)| dt.$$

Wir verzichten aber hier darauf, dieses Argument rigoros zu machen.

Beispiel 23.2 Die Länge von $\gamma(t) = (r \cos t, r \sin t, at)$, $0 \leq t \leq 2\pi$, ist

$$L(\gamma) = \int_0^{2\pi} |(-r \sin t, r \cos t, a)| dt = 2\pi \sqrt{r^2 + a^2}.$$

Die Schraubenlinie verläuft auf dem Mantel des Zylinders $x^2 + y^2 = r^2$, und man kann das Ergebnis durch Abrollen des Zylinders mit dem Satz des Pythagoras bestätigen.

Anschaulich besteht eine Kurve aus ihrem Bild im \mathbb{R}^n und einem Fahrplan, wie dieses Bild durchlaufen werden soll. In der Physik spielt der Fahrplan eine wesentliche Rolle, zum Beispiel für unsere Jahreszeiten. Die Bogenlänge ist jedoch eine geometrische Größe, und sollte nicht vom Fahrplan abhängen. Das soll nun präzisiert werden.

Definition 23.3 (Umparametrisierung) Seien $\gamma_{1,2} : I_{1,2} \rightarrow \mathbb{R}^n$ parametrisierte Kurven. Dann heißt γ_2 Umparametrisierung von γ_1 , falls eine Bijektion $\varphi \in C^1(I_2, I_1)$ mit $\varphi' \neq 0$ existiert, so dass $\gamma_2 = \gamma_1 \circ \varphi$.

Die Bijektion φ nennt man auch Parametertransformation. An dieser Stelle ergibt sich die Frage, warum wir die Bedingung $\varphi' \neq 0$ verlangen. Eine Antwort ist, dass wir jedenfalls die Differenzierbarkeit der Parametertransformationen verlangen wollen, außerdem sollte die Sache symmetrisch bezüglich γ_1 und γ_2 sein. Die Bedingung $\varphi' \neq 0$ ist aber äquivalent dazu, dass die inverse Parametertransformation φ^{-1} differenzierbar ist. Tatsächlich ist die Relation $\gamma_1 \sim \gamma_2$, falls γ_2 Umparametrisierung von γ_1 , eine Äquivalenzrelation (Übungsaufgabe).

Lemma 23.1 (Invarianz der Bogenlänge) Sind $\gamma_{1,2} \in C^1(I_{1,2}, \mathbb{R}^n)$ und ist γ_2 eine Umparametrisierung von γ_1 , so folgt $L(\gamma_2) = L(\gamma_1)$.

BEWEIS: Sei $\gamma_2 = \gamma_1 \circ \varphi$ für $\varphi \in C^1(I_2, I_1)$ mit $\varphi' \neq 0$. Dann folgt je nach Vorzeichen von φ'

$$|\gamma_2'(t)| = |\gamma_1'(\varphi(t))| |\varphi'(t)| = \pm |\gamma_1'(\varphi(t))| |\varphi'(t)|,$$

also liefert die Substitution $s = \varphi(t)$ für $I_1 = [a_1, b_1]$ sowie $I_2 = [a_2, b_2]$

$$L(\gamma_2) = \int_{a_2}^{b_2} |\gamma_2'(t)| dt = \pm \int_{a_2}^{b_2} |\gamma_1'(\varphi(t))| |\varphi'(t)| dt = \pm \int_{\varphi(a_2)}^{\varphi(b_2)} |\gamma_1'(s)| ds = L(\gamma_1).$$

□

Es ist eine naheliegende Frage, ob für eine gegebene Kurve eine besonders schöne Umparametrisierung existiert. Was dabei besonders schön sein soll, sagt folgende Definition.

Definition 23.4 (Parametrisierung nach der Bogenlänge) Eine Kurve $c \in C^1(I, \mathbb{R}^n)$ heißt nach der Bogenlänge parametrisiert, wenn gilt:

$$|c'(s)| = 1 \quad \text{für alle } s \in I.$$

Physikalisch betrachtet wird eine nach der Bogenlänge parametrisierte Kurve mit Absolutgeschwindigkeit Eins durchlaufen. Geometrisch folgt für jedes Teilintervall $[a, b] \subset I$

$$L(c|_{[a,b]}) = \int_a^b |c'(s)| ds = (b - a),$$

das heißt das Intervall wird durch c längentreu abgebildet.

Satz 23.1 (Parametrisierung nach der Bogenlänge) Sei $\gamma \in C^1([a, b], \mathbb{R}^n)$ eine Kurve mit Länge $L = L(\gamma)$. Es gelte $\gamma'(t) \neq 0$ für alle $t \in I$. Dann gibt es eine C^1 -Bijektion $\varphi : [0, L] \rightarrow [a, b]$ mit $\varphi' > 0$, so dass $c = \gamma \circ \varphi$ nach der Bogenlänge parametrisiert ist.

BEWEIS: Wir betrachten die Bogenlängenfunktion

$$(23.12) \quad \sigma : [a, b] \rightarrow [0, L], \quad \sigma(t) = \int_a^t |\gamma'(\tau)| d\tau.$$

Es gilt $\sigma'(t) = |\gamma'(t)| > 0$ nach Voraussetzung, also ist σ eine Bijektion der Klasse C^1 . Bezeichnet $\varphi : [0, L] \rightarrow [a, b]$ die Umkehrfunktion, so folgt

$$|(\gamma \circ \varphi)'(s)| = |\gamma'(\varphi(s))| |\varphi'(s)| = |\gamma'(\varphi(s))| \frac{1}{\sigma'(\varphi(s))} = 1.$$

□

Eine Kurve γ mit $\gamma'(t) \neq 0$ für alle t nennt man regulär (oder immergiert). Verzichtet man auf die Bedingung der Regularität, so kann das Bild sogar einer C^∞ -Kurve Singularitäten aufweisen. Ein Beispiel ist die Neilsche Parabel $\gamma : \mathbb{R} \rightarrow \mathbb{R}^2$, $\gamma(t) = (t^2, t^3)$. Das Bild $\gamma(\mathbb{R}) = \{(x, \pm x^{3/2}) : x \in \mathbb{R}\}$ hat eine Spitze im Nullpunkt. Die Regularität einer Kurve bleibt unter Umparametrisierungen erhalten, denn es gilt $(\gamma \circ \varphi)' = (\gamma' \circ \varphi)\varphi'$ mit $\varphi' \neq 0$ nach Definition. Insbesondere kann auf die Voraussetzung $\gamma' \neq 0$ in Satz 23.1 nicht verzichtet werden.

Ist der Endpunkt einer Kurve der Anfangspunkt einer zweiten Kurve, so ist es nahelegend, diese zu einer Kurve zusammensetzen. Im allgemeinen wird das Ergebnis dann keine C^1 -Kurve mehr sein. Auch stückweise lineare Kurven sind in der Regel nicht C^1 . Es ist daher praktisch, als Verallgemeinerung stückweise C^1 -Kurven einzuführen.

Definition 23.5 (stückweise C^1) Eine Kurve $\gamma \in C^0([a, b], \mathbb{R}^n)$ heißt stückweise C^1 , wenn es eine Zerlegung $a = t_0 < \dots < t_N = b$ gibt, so dass mit $I_k = [t_{k-1}, t_k]$ gilt:

$$\gamma|_{I_k} \in C^1(I_k, \mathbb{R}^n) \quad \text{für alle } k = 1, \dots, N.$$

Notation: $\gamma \in PC^1(I, \mathbb{R}^n)$ (piecewise continuously differentiable).

In den Unterteilungspunkten existieren die links- und rechtsseitigen Grenzwerte $\gamma'_\pm(t_i)$, die jedoch nicht übereinstimmen müssen. Setzen wir willkürlich $\gamma'(t_i) = 0$, so ist die Funktion $\gamma' : [a, b] \rightarrow \mathbb{R}^n$ stückweise stetig, insbesondere ist $|\gamma'| : [a, b] \rightarrow \mathbb{R}$ Riemann-integrierbar, siehe Kapitel 5, Abschnitt 1. Die Bogenlänge ist also auch für $\gamma \in PC^1([a, b], \mathbb{R}^n)$ durch Definition 23.2 erklärt. Insbesondere hängt die Länge nicht von der Wahl der Unterteilung ab.

Im Folgenden bezeichne $\langle \cdot, \cdot \rangle$ stets das Standardskalarprodukt im \mathbb{R}^n .

Definition 23.6 Ist $F \in C^0(\Omega, \mathbb{R}^n)$, $\Omega \subset \mathbb{R}^n$ offen, und $\gamma \in PC^1([a, b], \Omega)$, so heißt

$$\int_\gamma F \cdot d\vec{x} := \int_a^b \langle F(\gamma(t)), \gamma'(t) \rangle dt$$

das Kurvenintegral von F längs c .

In der Physik ist zum Beispiel F das Gravitationsfeld, und das Kurvenintegral ergibt die Arbeit, die beim Transport einer Masse längs einer Kurve innerhalb des Feldes verrichtet wird. Dabei wird $d\vec{x}$ als vektorielles Längenelement interpretiert, und der Punkt \cdot bedeutet das Skalarprodukt im \mathbb{R}^n . Wir haben diese Notation übernommen, obwohl sie für uns rein symbolisch ist, das heißt $d\vec{x}$ hat keine eigene mathematische Bedeutung, ähnlich wie beim Riemannintegral. Das Kurvenintegral ist allein durch die rechte Seite in Definition 23.6 erklärt. Allerdings ist die Merkregel, dass $d\vec{x}$ durch $\gamma'(t) dt$ zu ersetzen ist, hilfreich.

Beispiel 23.3 (Gravitationsfeld) Das Gravitationsfeld der Sonne ist gegeben durch

$$F : \mathbb{R}^3 \setminus \{0\} \rightarrow \mathbb{R}^3, F(x) = -C \frac{x}{|x|^3} \quad \text{mit } C > 0.$$

Beispiel 23.4 (Winkelvektorfeld) Wir betrachten hier das Vektorfeld

$$W : \mathbb{R}^2 \setminus \{0\} \rightarrow \mathbb{R}^2, W(x, y) = \left(-\frac{y}{x^2 + y^2}, \frac{x}{x^2 + y^2} \right).$$

Ist $\gamma : [a, b] \rightarrow \mathbb{R}^2 \setminus \{0\}$, $\gamma(t) = r(t)(\cos \theta(t), \sin \theta(t))$ mit $r, \theta \in C^1(I)$, so folgt

$$(23.13) \quad \int_{\gamma} W \cdot d\vec{x} = \int_a^b \left\langle \frac{1}{r} \begin{pmatrix} -\sin \theta \\ \cos \theta \end{pmatrix}, r' \begin{pmatrix} \cos \theta \\ \sin \theta \end{pmatrix} + r\theta' \begin{pmatrix} -\sin \theta \\ \cos \theta \end{pmatrix} \right\rangle dt = \theta(b) - \theta(a).$$

Also liefert das Kurvenintegral von W die Differenz der Polarwinkel von Endpunkt und Anfangspunkt der Kurve.

Eine Parametertransformation $\varphi : I_2 \rightarrow I_1$ heißt orientierungserhaltend (bzw. orientierungsumkehrend), wenn $\varphi' > 0$ (bzw. $\varphi' < 0$) ist. Anschaulich wird bei Umparametrisierung mit einer orientierungsumkehrenden Parametertransformation die Kurve umgekehrt durchlaufen, Anfangs- und Endpunkt tauschen ihre Rollen. Während die Bogenlänge unter allen Umparametrisierungen invariant ist, ist das Kurvenintegral nur bei orientierungserhaltenden Umparametrisierungen invariant, bei orientierungsumkehrenden Umparametrisierungen wechselt es sein Vorzeichen. Dies wird sogleich gezeigt.

Lemma 23.2 *Das Kurvenintegral hat folgende Eigenschaften:*

(a) *Linearität:* sind $F_{1,2} \in C^0(\Omega, \mathbb{R}^2)$ und $\lambda_{1,2} \in \mathbb{R}$, so gilt für $\gamma \in PC^1([a, b], \Omega)$

$$\int_{\gamma} (\lambda_1 F_1 + \lambda_2 F_2) \cdot d\vec{x} = \lambda_1 \int_{\gamma} F_1 \cdot d\vec{x} + \lambda_2 \int_{\gamma} F_2 \cdot d\vec{x}.$$

(b) *Additivität bei Zerlegungen:* ist $\gamma \in PC^1([a, b], \mathbb{R}^n)$ und $a = t_0 < \dots < t_N = b$ eine beliebige Zerlegung von $[a, b]$, so folgt mit $\gamma_i = \gamma|_{[t_{i-1}, t_i]}$

$$\int_{\gamma} F \cdot d\vec{x} = \sum_{i=1}^N \int_{\gamma_i} F \cdot d\vec{x}.$$

(c) *Invarianz bei Umparametrisierungen:* sei $\gamma \in PC^1(I_1, \mathbb{R}^2)$ und $\varphi \in C^1(I_2, I_1)$ eine Parametertransformation. Dann gilt, je nach Vorzeichen von φ' ,

$$\int_{\gamma \circ \varphi} F \cdot d\vec{x} = \pm \int_{\gamma} F \cdot d\vec{x}.$$

BEWEIS: (a) und (b) folgen aus der Definition und den Eigenschaften des Riemannintegrals. Für (c) sei $I_1 = [a_1, b_1]$ und $I_2 = [a_2, b_2]$. Mit der Substitution $\varphi(t) = s$ ergibt sich

$$\begin{aligned} \int_{\gamma \circ \varphi} F \cdot d\vec{x} &= \int_{a_2}^{b_2} \langle (F \circ \gamma \circ \varphi)(t), (\gamma \circ \varphi)'(t) \rangle dt \\ &= \int_{a_2}^{b_2} \langle F \circ \gamma(\varphi(t)), \gamma'(\varphi(t)) \rangle \varphi'(t) dt \\ &= \int_{\varphi(a_2)}^{\varphi(b_2)} \langle (F \circ \gamma)(s), \gamma'(s) \rangle ds. \end{aligned}$$

Ist φ orientierungserhaltend, so gilt $\varphi(a_2) = a_1$ und $\varphi(b_2) = b_1$ und wir bekommen das Kurvenintegral. Ist φ orientierungsumkehrend, so sind die Grenzen vertauscht und wir bekommen das negative Kurvenintegral. \square

Wir benötigen wie beim Riemann-Integral eine Standardabschätzung des Kurvenintegrals.

Lemma 23.3 Sei $F \in C^0(\Omega, \mathbb{R}^n)$, $\Omega \subset \mathbb{R}^n$ offen, und $\gamma \in PC^1(I, \Omega)$ mit $I = [a, b]$. Dann gilt mit $\|F \circ \gamma\|_I = \sup_{t \in I} |F(\gamma(t))|$

$$\left| \int_{\gamma} F \cdot d\vec{x} \right| \leq \|F \circ \gamma\|_I L(\gamma).$$

BEWEIS: Aus der Ungleichung von Cauchy-Schwarz und der Standardabschätzung des Riemann-Integrals folgt

$$\begin{aligned} \left| \int_a^b \langle F(\gamma(t)), \gamma'(t) \rangle dt \right| &\leq \int_a^b |\langle F(\gamma(t)), \gamma'(t) \rangle| dt \\ &\leq \int_a^b |F(\gamma(t))| |\gamma'(t)| dt \\ &\leq \|F \circ \gamma\|_I \int_a^b |\gamma'(t)| dt \\ &= \|F \circ \gamma\|_I L(\gamma). \end{aligned}$$

\square

Die Physiker nennen das Gravitationsfeld konservativ, weil der Energieerhaltungssatz gilt. Die verrichtete Arbeit zum Beispiel bei Transport einer Masse vom Mathematischen Institut zum Kandel entspricht genau der zugewonnenen Lageenergie, und diese kommt beim Herunterrollen auch wieder heraus, theoretisch jedenfalls. Der Begriff des konservativen Feldes ist auch in der Mathematik interessant.

Definition 23.7 Sei $\Omega \subset \mathbb{R}^n$ offen. Ein Vektorfeld $F \in C^0(\Omega, \mathbb{R}^n)$ heißt Gradientenfeld (bzw. konservativ), wenn es eine Funktion $\varphi \in C^1(\Omega)$ gibt mit $\text{grad } \varphi = F$. Die Funktion φ heißt Stammfunktion (bzw. Potential) von F .

Lemma 23.4 (Eindeutigkeit der Stammfunktion) Ist $\Omega \subset \mathbb{R}^n$ wegweise zusammenhängend, so ist eine Stammfunktion von $F \in C^0(\Omega, \mathbb{R}^n)$ eindeutig bestimmt, bis auf eine additive Konstante.

BEWEIS: Sind $\varphi_1, \varphi_2 \in C^1(\Omega)$ Stammfunktionen von F , so folgt

$$\text{grad}(\varphi_2 - \varphi_1) = \text{grad } \varphi_2 - \text{grad } \varphi_1 = F - F = 0,$$

also ist $\varphi_2 - \varphi_1$ konstant nach Satz 19.1, das heißt $\varphi_2 = \varphi_1 + c$. \square

Wir werden jetzt sehen, dass die Existenz einer Stammfunktion gleichbedeutend damit ist, dass das Kurvenintegral für Kurven mit gleichem Anfangs- und Endpunkt stets denselben Wert hat. Zuvor eine Definition.

Definition 23.8 Eine Kurve $\gamma : [a, b] \rightarrow \mathbb{R}^n$ heißt geschlossen, wenn $\gamma(a) = \gamma(b)$.

Aus $\gamma(a) = \gamma(b)$ folgt nicht notwendig $\gamma'(a) = \gamma'(b)$, zum Beispiel ist im Fall der Acht aus Beispiel 23.1 $\gamma(\pi/2) = \gamma(3\pi/2) = (0, 0)$, während $\gamma'(\pi/2) = (-1, -2) \neq (1, -2) = \gamma'(3\pi/2)$. Anschaulich schneidet sich die Kurve hier mit einem Winkel.

Satz 23.2 (Wegunabhängigkeit des Kurvenintegrals) Sei $\Omega \subset \mathbb{R}^n$ offen und wegweise zusammenhängend. Für ein Vektorfeld $F \in C^0(\Omega, \mathbb{R}^n)$ sind folgende Aussagen äquivalent:

- (a) F ist ein Gradientenfeld.
- (b) Für jede geschlossene Kurve $\gamma \in PC^1([a, b], \Omega)$ ist $\int_{\gamma} F \cdot d\vec{x} = 0$.
- (c) Für je zwei Kurven $\gamma_0, \gamma_1 \in PC^1([a, b], \Omega)$ mit $\gamma_0(a) = \gamma_1(a)$, $\gamma_0(b) = \gamma_1(b)$ gilt

$$\int_{\gamma_0} F \cdot d\vec{x} = \int_{\gamma_1} F \cdot d\vec{x}.$$

BEWEIS: Ist $F = \text{grad } \varphi$ mit $\varphi \in C^1(\Omega)$, so folgt für $\gamma \in PC^1([a, b], \Omega)$ geschlossen

$$\int_{\gamma} F \cdot d\vec{x} = \int_a^b \langle \text{grad } \varphi(\gamma(t)), \gamma'(t) \rangle dt = \int_a^b (\varphi \circ \gamma)'(t) dt = \varphi(\gamma(b)) - \varphi(\gamma(a)) = 0.$$

Für $\gamma_{0,1} \in PC^1([a, b], \Omega)$ mit gleichem Anfangs- und Endpunkt ist die Kurve

$$\gamma(t) = \begin{cases} \gamma_0(t) & a \leq t \leq b \\ \gamma_1(2b - t) & b \leq t \leq 2b - a \end{cases}$$

geschlossen und stückweise C^1 , und aus (b) ergibt sich mit Lemma 23.2

$$0 = \int_{\gamma} F \cdot d\vec{x} = \int_{\gamma_0} F \cdot d\vec{x} - \int_{\gamma_1} F \cdot d\vec{x}.$$

Für (c) \Rightarrow (a) sei $x_0 \in \Omega$ fest. Zu $x \in \Omega$ wählen wir eine Kurve $\gamma_x \in PC^1([0, 1], \Omega)$ mit $\gamma_x(0) = x_0$ und $\gamma_x(1) = x$, und setzen

$$\varphi(x) = \int_{\gamma_x} F \cdot d\vec{x}.$$

Die Existenz von γ_x ist gesichert nach Lemma 19.2 in Kapitel 7, genauer können wir γ_x stückweise linear wählen. Nach Voraussetzung (c) hängt das Kurvenintegral nicht von der Wahl von γ_x ab. Daher ist die Funktion $\varphi : \Omega \rightarrow \mathbb{R}$ wohldefiniert. Sei nun $x \in \Omega$. Für $|h|$ klein erhalten wir eine Kurve von x_0 nach $x + he_j$, indem wir γ_x mit der Kurve $c : [0, 1] \rightarrow \mathbb{R}^n$, $c(t) = x + the_j$, zusammensetzen. Es folgt

$$\frac{\varphi(x + he_j) - \varphi(x)}{h} = \int_c F \cdot d\vec{x} = \int_0^1 \langle F(x + the_j), e_j \rangle dt \rightarrow F_j(x) \quad \text{mit } h \rightarrow 0.$$

Also gilt $\partial_j \varphi = F_j$ für $j = 1, \dots, n$. □

Die zentrale Frage dieses Kapitels ist: wie können wir entscheiden, ob ein gegebenes Vektorfeld ein Gradientenfeld ist? Für C^1 -Vektorfelder gibt es eine notwendige Bedingung, die offensichtlich ist.

Satz 23.3 (Rotationsfreiheit von Gradientenfeldern) Sei $\Omega \subset \mathbb{R}^n$ offen. Ist $F \in C^1(\Omega, \mathbb{R}^n)$ ein Gradientenfeld, so gilt für alle $i, j = 1, \dots, n$

$$\partial_i F_j = \partial_j F_i \quad \text{in } \Omega.$$

BEWEIS: Ist $F = \text{grad } \varphi$, so folgt $\varphi \in C^2(\Omega)$ und wegen der Vertauschbarkeit der partiellen Ableitungen, Satz 17.2 in Kapitel 6, gilt

$$\partial_i F_j = \partial_i \partial_j \varphi = \partial_j \partial_i \varphi = \partial_j F_i.$$

□

Für $n = 3$ lässt sich die Bedingung schreiben als $\text{rot } F = 0$, wobei

$$\text{rot } F = (\partial_2 F_3 - \partial_3 F_2, \partial_3 F_1 - \partial_1 F_3, \partial_1 F_2 - \partial_2 F_1).$$

Beispiel 23.5 $F : \mathbb{R}^2 \rightarrow \mathbb{R}^2$, $F(x, y) = (-y, x)$, hat keine Stammfunktion auf \mathbb{R}^2 , denn

$$\partial_1 F_2 = 1, \text{ aber } \partial_2 F_1 = -1.$$

Beispiel 23.6 Für das Winkelvektorfeld

$$W : \mathbb{R}^2 \setminus \{0\} \rightarrow \mathbb{R}^2, W(x, y) = \left(-\frac{y}{x^2 + y^2}, \frac{x}{x^2 + y^2} \right)$$

vgl. Beispiel 23.4, berechnen wir

$$\partial_1 W_2 = \frac{y^2 - x^2}{(x^2 + y^2)^2} = \partial_2 W_1,$$

das heißt die notwendige Bedingung aus Satz 23.3 ist erfüllt. Dennoch ist das Kurvenintegral nicht wegunabhängig: für $k \in \mathbb{Z}$ und $r > 0$ ist die Kurve $\gamma_k : [0, 2\pi] \rightarrow \mathbb{R}^2 \setminus \{0\}$, $\gamma_k(t) = r(\cos kt, \sin kt)$, geschlossen und es gilt nach Beispiel 23.4

$$\int_{\gamma_k} W \cdot d\vec{x} = 2\pi k \quad (\neq 0 \text{ für } k \in \mathbb{Z} \setminus \{0\}).$$

Im Unterschied zum vorangegangenen Beispiel besitzt das Vektorfeld W allerdings zumindest lokal Stammfunktionen. Zum Beispiel ist

$$\varphi : \{(x, y) \in \mathbb{R}^2 : y > 0\} \rightarrow \mathbb{R}, \varphi(x, y) = \arccos \frac{x}{\sqrt{x^2 + y^2}}$$

eine Stammfunktion von W auf der oberen Halbebene; analog hat man Stammfunktionen auch auf der unteren, rechten und linken Halbebene.

Wir wollen jetzt untersuchen, wie sich das Kurvenintegral längs einer Schar von Kurven ändert. Statt Schar verwenden wir den moderneren Ausdruck Homotopie. Dies ist ein fundamentales Konzept der Analysis.

Definition 23.9 (Homotopie) Eine Homotopie in Ω zwischen Kurven $\gamma_0, \gamma_1 \in C^0([a, b], \Omega)$ ist eine Abbildung $\gamma \in C^0([a, b] \times [0, 1], \Omega)$ mit

$$\gamma(\cdot, 0) = \gamma_0 \quad \text{und} \quad \gamma(\cdot, 1) = \gamma_1.$$

Gilt $\gamma_0(a) = \gamma_1(a) = p$, $\gamma_0(b) = \gamma_1(b) = q$, und gibt es eine Homotopie mit $\gamma(a, t) = p$, $\gamma(b, t) = q$ für alle $t \in [0, 1]$, so heißen γ_0, γ_1 homotop in Ω mit festen Endpunkten. Gilt $\gamma_0(a) = \gamma_0(b)$, $\gamma_1(a) = \gamma_1(b)$, und gibt es eine Homotopie mit $\gamma(a, t) = \gamma(b, t)$ für alle $t \in [0, 1]$, so heißen γ_0, γ_1 in Ω geschlossen homotop.

Es gilt folgende allgemeine Formel für die Änderung des Kurvenintegrals unter (hinreichend glatten) Homotopien.

Lemma 23.5 (Homotopieformel) Sei $\Omega \subset \mathbb{R}^n$ offen und $F \in C^1(\Omega, \mathbb{R}^n)$. Dann gilt für $\gamma \in C^1([a, b] \times [0, 1], \Omega)$, falls $\partial_s \partial_t \gamma \in C^0([a, b] \times [0, 1], \mathbb{R}^2)$,

$$\begin{aligned} \int_{\gamma(\cdot, 1)} F \cdot d\vec{x} - \int_{\gamma(\cdot, 0)} F \cdot d\vec{x} &= \int_{\gamma(b, \cdot)} F \cdot d\vec{x} - \int_{\gamma(a, \cdot)} F \cdot d\vec{x} \\ &+ \int_0^1 \int_a^b \left(\langle DF \circ \gamma \frac{\partial \gamma}{\partial t}, \frac{\partial \gamma}{\partial s} \rangle - \langle DF \circ \gamma \frac{\partial \gamma}{\partial s}, \frac{\partial \gamma}{\partial t} \rangle \right) ds dt. \end{aligned}$$

Gilt $\partial_i F_j = \partial_j F_i$ für $i, j = 1, \dots, n$, und hat die Homotopie feste Endpunkte oder ist geschlossen, so ist die rechte Seite Null.

BEWEIS: Nach Zusatz zum Satz von Schwarz, Satz 17.2 in Kapitel 6, gilt $\partial_t \partial_s \gamma = \partial_s \partial_t \gamma$. Wir berechnen mit Satz 22.2 und partieller Integration bezüglich $s \in [a, b]$

$$\begin{aligned} \frac{\partial}{\partial t} \int_{\gamma(\cdot, t)} F \cdot d\vec{x} &= \frac{\partial}{\partial t} \int_a^b \langle F(\gamma(s, t)), \frac{\partial \gamma}{\partial s}(s, t) \rangle ds \\ &= \int_a^b \langle DF \circ \gamma \frac{\partial \gamma}{\partial t}, \frac{\partial \gamma}{\partial s} \rangle ds + \int_a^b \langle F \circ \gamma, \frac{\partial^2 \gamma}{\partial t \partial s} \rangle ds \\ &= \int_a^b \langle DF \circ \gamma \frac{\partial \gamma}{\partial t}, \frac{\partial \gamma}{\partial s} \rangle ds + \langle F \circ \gamma, \frac{\partial \gamma}{\partial t} \rangle \Big|_{s=a}^{s=b} - \int_a^b \langle DF \circ \gamma \frac{\partial \gamma}{\partial s}, \frac{\partial \gamma}{\partial t} \rangle ds. \end{aligned}$$

Integration bezüglich t ergibt die Formel. Ist $\partial_i F_j = \partial_j F_i$ oder mit anderen Worten DF symmetrisch, so verschwindet das Doppelintegral. Bei festen Endpunkten sind $\gamma(\cdot, a)$ und $\gamma(\cdot, b)$ konstant, also die zugehörigen Kurvenintegrale Null. Ist die Homotopie geschlossen, so gilt $\gamma(\cdot, a) = \gamma(\cdot, b)$ und die Kurvenintegrale rechts heben sich gegenseitig weg. \square

Wir können an dieser Stelle als Anwendung den Fundamentalsatz der Algebra (Kapitel 2, Satz 3.10) beweisen. Es gibt vielleicht einfachere Beweise, aber dieses Argument ist jedenfalls sehr anschaulich.

Satz 23.4 (Fundamentalsatz der Algebra) Jedes komplexe Polynom vom Grad $n \geq 1$ hat mindestens eine Nullstelle $z_0 \in \mathbb{C}$.

BEWEIS: Wir verwenden das Winkelvektorfeld $W : \mathbb{R}^2 \setminus \{0\} \rightarrow \mathbb{R}^2$. Nach Beispiel 23.6 gilt $\partial_1 W_2 = \partial_2 W_1$. Sei $p(z) = z^n + a_{n-1} z^{n-1} + \dots + a_0$ mit $a_i \in \mathbb{C}$ und $n \geq 1$. Schreibe $p(z) = p_n(z) + q(z)$ mit $p_n(z) = z^n$. Da q höchstens den Grad $n-1$ hat, gilt $z^{-n} q(z) \rightarrow 0$ mit $|z| \rightarrow \infty$, also für $R > 0$ hinreichend groß

$$|q(Re^{i\theta})| \leq \frac{1}{2} R^n \quad \text{für } \theta \in [0, 2\pi].$$

Betrachte nun die geschlossene Kurve $\gamma_0 : [0, 2\pi] \rightarrow \mathbb{R}^2$, $\gamma_0(\theta) = p(Re^{i\theta})$, und die Homotopie

$$\gamma : [0, 2\pi] \times [0, 1] \rightarrow \mathbb{R}^2, \quad \gamma(\theta, t) = p_n(Re^{i\theta}) + (1-t)q(Re^{i\theta}).$$

Wir berechnen

$$|\gamma(\theta, t)| \geq |p_n(Re^{i\theta})| - |q(Re^{i\theta})| \geq R^n - \frac{1}{2} R^n > 0.$$

Da $\gamma(\theta, 1) = p_n(Re^{i\theta}) = R^n e^{in\theta}$, folgt aus Lemma 23.5 und Beispiel 23.6

$$\int_{\gamma_0} W \cdot d\vec{x} = \int_{\gamma(\cdot, 1)} W \cdot d\vec{x} = 2\pi n.$$

Wir betrachten jetzt eine zweite, ebenfalls glatte Homotopie:

$$\tilde{\gamma} : [0, 2\pi] \times [0, R] \rightarrow \mathbb{R}^2, \tilde{\gamma}(\theta, \varrho) = p(\varrho e^{i\theta}).$$

Es ist $\tilde{\gamma}(\theta, 0) = p(0) = a_0$ eine konstante Abbildung. Hätte p keine Nullstelle in \mathbb{C} , so ist auch dies eine geschlossene Homotopie in $\mathbb{R}^2 \setminus \{0\}$, und es folgt wieder mit Lemma 23.5

$$\int_{\gamma_0} W \cdot d\vec{x} = \int_{\tilde{\gamma}(\cdot, 0)} W \cdot d\vec{x} = 0,$$

das heißt $2\pi n = 0$, ein Widerspruch. □

Lemma 23.6 (affine Homotopie) Sei $\Omega \subset \mathbb{R}^n$ offen und $F \in C^1(\Omega, \mathbb{R}^n)$ mit $\partial_i F_j = \partial_j F_i$ für $i, j = 1, \dots, n$. Für Kurven $\gamma_0, \gamma_1 \in PC^1([a, b], \Omega)$ betrachte die affine Homotopie

$$\gamma : [a, b] \times [0, 1] \rightarrow \mathbb{R}^n, \gamma(s, t) = (1-t)\gamma_0(s) + t\gamma_1(s).$$

Haben γ_0, γ_1 gleiche Endpunkte oder sind geschlossen, und gilt $\gamma([a, b] \times [0, 1]) \subset \Omega$, so folgt

$$\int_{\gamma_0} F \cdot d\vec{x} = \int_{\gamma_1} F \cdot d\vec{x}.$$

BEWEIS: Sind $\gamma_0, \gamma_1 \in C^1([a, b], \Omega)$, so folgt die Aussage direkt aus Lemma 23.5. Für γ_0, γ_1 stückweise C^1 zerlegen wir $[a, b]$ in Teilintervalle, auf denen beide Kurven C^1 sind, und wenden Lemma 23.5 auf den Teilintervallen an. Die Randintegrale heben sich bei Addition heraus. □

Satz 23.5 (Homotopieinvarianz des Kurvenintegrals) Sei $\Omega \subset \mathbb{R}^n$ offen und $F \in C^1(\Omega, \mathbb{R}^n)$ mit $\partial_i F_j = \partial_j F_i$ auf Ω für $1 \leq i, j \leq n$. Sind dann $\gamma_0, \gamma_1 \in PC^1([a, b], \Omega)$ homotop in Ω mit festen Endpunkten (oder geschlossen homotop), so gilt

$$\int_{\gamma_0} F \cdot d\vec{x} = \int_{\gamma_1} F \cdot d\vec{x}.$$

BEWEIS: Ist die Homotopie hinreichend glatt, so folgt die Aussage aus Lemma 23.5. Es geht also um das technische Problem, dass die gegebene Homotopie $\gamma \in C^0([a, b] \times [0, 1], \Omega)$ eventuell nur stetig ist. Die Lektüre des Beweises könnte zurückgestellt werden.

Aus Kompaktheitsgründen gibt es ein $\varepsilon > 0$ mit $B_{2\varepsilon}(p) \subset \Omega$ für alle $p \in \gamma([a, b] \times [0, 1])$, vgl. Lemma 19.1 in Kapitel 7. Da γ auf der kompakten Menge $[a, b] \times [0, 1]$ gleichmäßig stetig ist, gibt es weiter ein $\delta > 0$ mit

$$|\gamma(s_0, t) - \gamma(s_1, t)|, |\gamma(s, t_0) - \gamma(s, t_1)| < \varepsilon \quad \text{für } |s_0 - s_1|, |t_0 - t_1| < \delta.$$

Wir ersetzen jetzt $\gamma(\cdot, t)$ durch stückweise lineare Kurven. Für $N \in \mathbb{N}$ mit $(b-a)/N < \delta$ und $s_k = a + k(b-a)/N$, $k = 0, 1, \dots, N$, definieren wir $\tilde{\gamma} : [a, b] \times [0, 1] \rightarrow \mathbb{R}^n$ durch

$$\tilde{\gamma}(s, t) = \frac{s_k - s}{s_k - s_{k-1}} \gamma(s_{k-1}, t) + \frac{s - s_{k-1}}{s_k - s_{k-1}} \gamma(s_k, t) \quad \text{für } s \in [s_{k-1}, s_k].$$

Es gilt $\tilde{\gamma}(a, t) = \gamma(a, t)$, $\tilde{\gamma}(b, t) = \gamma(b, t)$ für alle $t \in [0, 1]$. Für $s \in [s_{k-1}, s_k]$ haben wir

$$|\tilde{\gamma}(s, t) - \gamma(s, t)| = \left| \frac{s_k - s}{s_k - s_{k-1}} (\gamma(s_{k-1}, t) - \gamma(s, t)) + \frac{s - s_{k-1}}{s_k - s_{k-1}} (\gamma(s_k, t) - \gamma(s, t)) \right| < \varepsilon.$$

Für $\lambda \in [0, 1]$ folgt $|(1 - \lambda)\gamma(s, t) + \lambda\tilde{\gamma}(s, t) - \gamma(s, t)| \leq |\tilde{\gamma}(s, t) - \gamma(s, t)| < \varepsilon$, das heißt die affine Homotopie zwischen $\gamma(\cdot, t)$ und $\tilde{\gamma}(\cdot, t)$ liegt in Ω . Insbesondere folgt aus Lemma 23.6

$$(23.14) \quad \int_{\gamma_0} F \cdot \vec{dx} = \int_{\tilde{\gamma}(\cdot, 0)} F \cdot \vec{dx} \quad \text{und} \quad \int_{\gamma_1} F \cdot \vec{dx} = \int_{\tilde{\gamma}(\cdot, 1)} F \cdot \vec{dx}.$$

Weiter gilt für $|t_0 - t_1| < \delta$ und $s \in [s_{k-1}, s_k]$

$$|\tilde{\gamma}(s, t_0) - \tilde{\gamma}(s, t_1)| = \left| \frac{s_k - s}{s_k - s_{k-1}} (\gamma(s_{k-1}, t_0) - \gamma(s_{k-1}, t_1)) + \frac{s - s_{k-1}}{s_k - s_{k-1}} (\gamma(s_k, t_0) - \gamma(s_k, t_1)) \right| < \varepsilon,$$

und es folgt für alle $\lambda \in [0, 1]$

$$|((1 - \lambda)\tilde{\gamma}(s, t_0) + \lambda\tilde{\gamma}(s, t_1)) - \gamma(s, t_0)| \leq |\tilde{\gamma}(s, t_1) - \tilde{\gamma}(s, t_0)| + |\tilde{\gamma}(s, t_0) - \gamma(s, t_0)| < 2\varepsilon.$$

Für $1/N < \delta$ folgt mit $t_l = l/N$ für $l = 0, 1, \dots, N$ aus Lemma 23.6

$$\int_{\tilde{\gamma}(\cdot, t_l)} F \cdot \vec{dx} = \int_{\tilde{\gamma}(\cdot, t_{l-1})} F \cdot \vec{dx} \quad \text{für } l = 1, \dots, N,$$

und Kombination mit (23.14) beweist den Satz. \square

Definition 23.10 Eine Menge $\Omega \subset \mathbb{R}^n$ heißt einfach zusammenhängend, wenn jede geschlossene Kurve $\gamma \in C^0([a, b], \Omega)$ in Ω geschlossen homotop zu einer konstanten Kurve ist.

Beispiel 23.7 Eine Menge $\Omega \subset \mathbb{R}^n$ heißt sternförmig, wenn es ein $x_0 \in \Omega$ gibt mit

$$(1 - t)x + tx_0 \in \Omega \quad \text{für alle } x \in \Omega, t \in [0, 1].$$

Eine sternförmige Menge ist einfach zusammenhängend, denn jede geschlossene Kurve $\gamma_0 \in C^0([a, b], \Omega)$ ist homotop zur konstanten Kurve in x_0 , nämlich durch die Homotopie

$$\gamma : [a, b] \times [0, 1] \rightarrow \Omega, \quad \gamma(s, t) = (1 - t)\gamma_0(s) + tx_0.$$

Satz 23.6 Sei $\Omega \subset \mathbb{R}^n$ offen und einfach zusammenhängend. Dann sind für ein Vektorfeld $F \in C^1(\Omega, \mathbb{R}^n)$ folgende Aussagen äquivalent:

- (a) $\partial_i F_j = \partial_j F_i$ für $i, j = 1, \dots, n$.
- (b) F hat eine Stammfunktion.

BEWEIS: Aus (a) folgt mit Satz 23.5 für jeden geschlossenen Weg $\gamma \in PC^1([a, b], \Omega)$

$$\int_{\gamma} F \cdot \vec{dx} = 0.$$

Hieraus ergibt sich mit Satz 23.2 die Existenz einer Stammfunktion. Die umgekehrte Implikation wurde schon in Satz 23.3 festgestellt. \square

Beispiel 23.8 Ein Spezialfall von Satz 23.6 ist das Lemma von Poincaré: Ist $\Omega \subset \mathbb{R}^n$ Stammfunktion φ . Tatsächlich kann diese explizit angegeben werden: Integration längs $\gamma_x(t) = (1-t)x_0 + tx$, $t \in [0, 1]$, ergibt

$$\varphi(x) = \int_{\gamma_x} F \cdot d\vec{x} = \int_0^1 \langle F((1-t)x_0 + tx, x - x_0) \rangle dt.$$

Wir fassen unsere Resultate über Kurvenintegrale in einer kleinen Tabelle zusammen:

| | | |
|-----------------------------------|--|--|
| F Gradientenfeld | Satz 23.2 \Leftrightarrow | $\int F \cdot d\vec{x}$ wegunabhängig |
| \Downarrow Satz 23.3 | \Uparrow 1-fach zshg. Satz 23.6 \Uparrow | \Downarrow |
| $\partial_i F_j = \partial_j F_i$ | Satz 23.5 \Leftrightarrow | $\int F \cdot d\vec{x}$ homotopieinvariant |

Zu begründen ist noch die Implikation von rechts nach links in der unteren Zeile. Ist das Kurvenintegral homotopieinvariant, so ist es auf einer Umgebung $B_\rho(x) \subset \Omega$ aber wegunabhängig. Also hat das Vektorfeld auf $B_\rho(x)$ eine Stammfunktion, und es folgt $\partial_i F_j = \partial_j F_i$ auf $B_\rho(x)$.

Das Standardskalarprodukt erlaubt es, jedem Vektorfeld eine 1-Form bijektiv zuzuordnen, und zwar definiert man für das Vektorfeld $A : \Omega \rightarrow \mathbb{R}^n$ die 1-Form $\alpha : \Omega \rightarrow (\mathbb{R}^n)^*$ durch

$$\alpha(x)v = \langle A(x), v \rangle \quad \text{für } x \in \Omega, v \in \mathbb{R}^n.$$

Dann haben A und α die gleichen Koordinatenfunktionen, denn es ist

$$\alpha_i(x) = \alpha(x)e_i = \langle A(x), e_i \rangle = A_i(x);$$

Insbesondere ist die Gleichung $A = \text{grad } \varphi$ äquivalent zu $\alpha = d\varphi$. Etwas abstrakter ergibt sich das auch aus der Charakterisierung des Gradienten in Gleichung (18.5), Kapitel 6:

$$A = \text{grad } \varphi \quad \Leftrightarrow \quad \langle A(x), v \rangle = d\varphi(x)v \quad \text{für alle } x \in \Omega, v \in \mathbb{R}^n \quad \Leftrightarrow \quad \alpha = d\varphi.$$

Nach Definition von α gilt weiter $\alpha(\gamma(t))\gamma'(t) = \langle A(\gamma(t)), \gamma'(t) \rangle$, und damit

$$\int_\gamma \alpha = \int_\gamma A \cdot d\vec{x}.$$

Es wird folgende Terminologie eingeführt.

Definition 23.11 Eine 1-Form α auf der offenen Menge $\Omega \subset \mathbb{R}^n$ heißt

- (a) *exakt*, wenn es eine Funktion $\varphi \in C^1(\Omega)$ gibt mit $\alpha = d\varphi$,
- (b) *geschlossen*, wenn $\alpha \in C^1$ und $\partial_i \alpha_j = \partial_j \alpha_i$ auf Ω für $i, j = 1, \dots, n$.

Ist α dem Vektorfeld A zugeordnet, so ist demnach α genau dann exakt, wenn A ein Gradientenfeld ist, und genau dann geschlossen, wenn A rotationsfrei ist. Wir können somit alle unsere Resultate in der Sprache der 1-Formen neu formulieren:

- genau dann ist α exakt, wenn das Kurvenintegral wegunabhängig ist;
- ist α exakt, so auch geschlossen;
- ist α geschlossen, so ist das Kurvenintegral homotopieinvariant;
- auf einem einfach zusammenhängenden Gebiet ist jede geschlossene 1-Form exakt.

Der Vorteil der 1-Formen gegenüber den anschaulicheren Vektorfeldern liegt nun im Transformationsverhalten. Seien dazu $U \subset \mathbb{R}^n$ und $V \subset \mathbb{R}^m$ offene Mengen, und $\phi \in C^1(U, V)$. Ist ω eine 1-Form auf V , so erhalten wir eine 1-Form $\phi^*\omega$ auf U , den pullback von ω unter ϕ , durch die Formel

$$(\phi^*\omega)(x)v = \omega(\phi(x))D\phi(x)v \quad \text{für } x \in U, v \in \mathbb{R}^n.$$

Sind dy_i die Koordinatendifferentiale auf \mathbb{R}^m , so berechnen wir

$$(\phi^*dy_i)(x)v = dy_i \cdot D\phi(x)v = d\phi_i(x)v$$

beziehungsweise kurz $\phi^*dy_i = d\phi_i$ für $i = 1, \dots, m$, und allgemeiner

$$\phi^*\omega = \sum_{i=1}^m (\omega_i \circ \phi) d\phi_i \quad \text{mit} \quad d\phi_i = \sum_{j=1}^n \frac{\partial \phi_i}{\partial x_j} dx_j.$$

Ist $\phi \in C^{k+1}$ und $\omega \in C^k$ für $k \in \mathbb{N}_0$, so ist demnach $\phi^*\omega \in C^k$.

Satz 23.7 (Transformation des Kurvenintegrals) *Seien $U \subset \mathbb{R}^n$, $V \subset \mathbb{R}^m$ offen und $\phi \in C^1(U, V)$. Für eine stetige 1-Form ω auf V und $\gamma \in PC^1([a, b], U)$ gilt*

$$\int_{\phi \circ \gamma} \omega = \int_{\gamma} \phi^*\omega.$$

BEWEIS: Aus den Definitionen ergibt sich

$$\int_{\phi \circ \gamma} \omega = \int_a^b \omega(\phi(\gamma(t))) D\phi(\gamma(t)) \gamma'(t) dt = \int_a^b (\phi^*\omega)(\gamma(t)) \gamma'(t) dt = \int_{\gamma} \phi^*\omega.$$

□

Das Kurvenintegral von Vektorfeldern benutzt wesentlich das Skalarprodukt, was bei der Analyse des Transformationsverhaltens mit berücksichtigt werden müsste. Bei der Umrechnung des Laplaceoperators auf krummlinige Koordinaten wird uns so etwas noch begegnen.

24 Komplexe Analysis

In diesem Abschnitt wollen wir einen kurzen Ausflug in die komplexe Analysis – die sogenannte Funktionentheorie – unternehmen, und zwar wollen wir jetzt komplexe Kurvenintegrale betrachten. Im folgenden sei stets Ω eine offene Teilmenge von $\mathbb{C} \cong \mathbb{R}^2$.

Definition 24.1 Die Funktion $f : \Omega \rightarrow \mathbb{C}$ heißt komplex differenzierbar in $z \in \Omega$ mit Ableitung $f'(z) = c \in \mathbb{C}$, falls gilt:

$$\lim_{w \rightarrow z} \frac{f(w) - f(z)}{w - z} = c.$$

f heißt komplex differenzierbar oder holomorph auf Ω , wenn f in allen $z \in \Omega$ komplex differenzierbar ist.

Formal ist diese Definition völlig analog zur Definition der Differenzierbarkeit einer Funktion einer reellen Variablen, nur dass eben jetzt der Differenzenquotient in \mathbb{C} statt in \mathbb{R} gebildet wird. Demzufolge gelten auch alle üblichen Differentiationsregeln:

- Differenzierbarkeit in $z \in \Omega$ impliziert Stetigkeit in $z \in \Omega$.
- Linearität der Ableitung: für $f, g : \Omega \rightarrow \mathbb{C}$ und $\alpha, \beta \in \mathbb{C}$ gilt $(\alpha f + \beta g)' = \alpha f' + \beta g'$.
- Produktregel: für $f, g : \Omega \rightarrow \mathbb{C}$ gilt $(fg)' = f'g + fg'$.
- Quotientenregel: für $f, g : \Omega \rightarrow \mathbb{C}$ mit $g \neq 0$ gilt $\left(\frac{f}{g}\right)' = \frac{f'g - fg'}{g^2}$.
- Kettenregel: für offene $U, V \subset \mathbb{C}$ und $f : U \rightarrow V$, $g : V \rightarrow \mathbb{C}$ gilt $(g \circ f)' = (g' \circ f)f'$.

Die Beweise sind weitgehend analog zu den reellen Beweisen und wir wollen aus Zeitgründen darauf verzichten. Wir halten nur fest, dass zum Beispiel Polynome $P(z) = a_0 + a_1 z + \dots + a_n z^n$ differenzierbar sind mit Ableitung $P'(z) = a_1 + \dots + n a_n z^{n-1}$. Soweit verläuft die Diskussion der komplexen Differenzierbarkeit also ganz parallel zum reellen Fall.

Vergessen wir die komplexe Multiplikation, so ist \mathbb{C} nichts anderes als der \mathbb{R}^2 mit Standardbasis $1 = (1, 0)$ und $i = (0, 1)$. Eine Funktion $f : \Omega \rightarrow \mathbb{C}$ ist dann eine vektorwertige Funktion $f : \Omega \rightarrow \mathbb{R}^2$, $f(x, y) = (u(x, y), v(x, y))$, wobei

$$u(x, y) = \operatorname{Re} f(x + iy) \quad \text{und} \quad v(x, y) = \operatorname{Im} f(x + iy).$$

Es stellt sich nun die Frage: in welcher Beziehung stehen die komplexe Differenzierbarkeit von f und die Differenzierbarkeit als reelle, vektorwertige Funktion?

Satz 24.1 (Cauchy-Riemannsches Differentialgleichungen) Für $f : \Omega \rightarrow \mathbb{C}$, $f = u + iv$, sind folgende Aussagen äquivalent:

- f ist auf Ω komplex differenzierbar.
- $f : \Omega \rightarrow \mathbb{R}^2$ ist (reell) differenzierbar mit $u_x = v_y$ und $u_y = -v_x$.

BEWEIS: Die Multiplikation mit einer komplexen Zahl c ist eine lineare Abbildung von \mathbb{R}^2 nach \mathbb{R}^2 , genauer gilt für $c = a + ib$ mit einer einfachen Rechnung

$$cz = Cz \quad \text{mit } C = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}.$$

Auf der linken Seite steht dabei das Produkt der komplexen Zahlen c und z , rechts die Anwendung der Matrix $C \in \mathbb{R}^{2 \times 2}$ auf den Vektor $z \in \mathbb{R}^2$. Ist f in $z \in \Omega$ komplex differenzierbar mit $f'(z) = c$, so folgt mit dieser Wahl von C

$$\frac{f(w) - f(z) - C(w - z)}{|w - z|} = \frac{f(w) - f(z) - c(w - z)}{w - z} \frac{w - z}{|w - z|} \rightarrow 0,$$

also gilt $Df(z) = C$. Außerdem erhalten wir mit $f = u + iv$ im Punkt z die Gleichungen

$$(24.15) \quad f' = u_x - iu_y = v_y + iv_x = \frac{1}{2}(f_x - if_y).$$

Sei jetzt umgekehrt f reell differenzierbar in $z = (x, y)$, und die Cauchy-Riemann Gleichungen seien im Punkt z erfüllt. Dann gilt für $a = u_x$ und $b = -u_y$

$$Df(z) = \begin{pmatrix} u_x & u_y \\ v_x & v_y \end{pmatrix} = \begin{pmatrix} a & -b \\ b & a \end{pmatrix},$$

also folgt mit $c = a + ib$ für $w \rightarrow z$

$$\frac{f(w) - f(z) - c(w - z)}{w - z} = \frac{f(w) - f(z) - Df(z)(w - z)}{|w - z|} \frac{|w - z|}{w - z} \rightarrow 0.$$

□

Die komplexe Differenzierbarkeit ist also stärker als die reelle Differenzierbarkeit der vektorwertigen Funktion, und zwar müssen die partiellen Ableitungen zusätzlich die Gleichungen $u_x = v_y$ sowie $u_y = -v_x$ erfüllen. Wir wollen als nächstes das komplexe Kurvenintegral definieren. Das Riemann-Integral einer \mathbb{C} -wertigen Funktion $f = u + iv : I \rightarrow \mathbb{C}$ ist wie üblich komponentenweise definiert, das heißt

$$\int_I f = \int_I u + i \int_I v \in \mathbb{C}.$$

Definition 24.2 (komplexes Kurvenintegral) Sei $\Omega \subset \mathbb{C}$ offen und $f \in C^0(\Omega, \mathbb{C})$. Für $\gamma \in PC^1([a, b], \Omega)$ definieren wir

$$\int_\gamma f dz = \int_a^b f(\gamma(t))\gamma'(t) dt.$$

Im Integral rechts steht das Produkt der komplexen Zahlen $f(\gamma(t))$ und $\gamma'(t)$. Um das Integral in reelle Kurvenintegrale umzuformen, schreiben wir $f = u + iv$ mit $u, v \in C^0(\Omega)$ sowie $\gamma = x + iy$ mit $x, y \in PC^1(I)$ für $I = [a, b]$, und berechnen

$$\int_I (f \circ \gamma)\gamma' = \int_I ((u + iv) \circ \gamma) (x' + iy') = \int_I ((u \circ \gamma)x' - (v \circ \gamma)y') + i((u \circ \gamma)y' + (v \circ \gamma)x'),$$

das heißt es gilt

$$(24.16) \quad \int_{\gamma} f dz = \int_{\gamma} (u dx - v dy) + i \int_{\gamma} (u dy + v dx).$$

Als Merkregel kann man hier $dz = dx + idy$ benutzen, die rechte Seite der Formel ergibt sich dann durch formales Ausmultiplizieren.

Beispiel 24.1 Sei γ der gegen den Uhrzeigersinn durchlaufene Kreis mit Radius $r > 0$ um den Punkt $z_0 \in \mathbb{C}$, das heißt genauer

$$\gamma : [0, 2\pi] \rightarrow \mathbb{C} \setminus \{0\}, \quad \gamma(\theta) = z_0 + re^{i\theta}.$$

Es gilt $\gamma'(\theta) = ire^{i\theta}$. Wir berechnen nun für $k \in \mathbb{Z}$

$$\int_{\gamma} (z - z_0)^k dz = \int_0^{2\pi} (re^{i\theta})^k ire^{i\theta} d\theta = \begin{cases} 0 & \text{falls } k \in \mathbb{Z} \setminus \{-1\}, \\ 2\pi i & \text{falls } k = -1. \end{cases}$$

Satz 24.2 (Cauchys Integralsatz) Die Funktion $f : \Omega \rightarrow \mathbb{C}$, $\Omega \subset \mathbb{C}$ offen, sei stetig partiell differenzierbar und holomorph. Ist dann $\gamma \in PC^1(I, \Omega)$ geschlossen homotop zu einer konstanten Kurve, so folgt

$$\int_{\gamma} f dz = 0.$$

BEWEIS: Nach Satz 23.5 in diesem Kapitel müssen wir nur prüfen, ob die Differentialformen $u dx - v dy$ und $u dy + v dx$ geschlossen sind, das heißt ob gilt

$$u_y = -v_x \quad \text{und} \quad u_x = v_y.$$

Das sind aber genau die Cauchy-Riemannschen Differentialgleichungen. □

Etwas allgemeiner besagt Satz 23.5, dass das Kurvenintegral von $f(z) dz$, f holomorph, den gleichen Wert ergibt für zwei Kurven, die homotop mit festen Endpunkten oder geschlossen homotop sind. Tatsächlich kann auf die Stetigkeit der partiellen Ableitungen verzichtet werden, es reicht als Voraussetzung dass f holomorph ist. Diese Verschärfung ist für uns aber nicht relevant.

Definition 24.3 (komplexe Stammfunktion) Sei $\Omega \subset \mathbb{C}$ offen und $f : \Omega \rightarrow \mathbb{C}$. Dann heißt $F : \Omega \rightarrow \mathbb{C}$ (komplexe) Stammfunktion von f auf Ω , falls gilt:

$$F'(z) = f(z) \quad \text{für alle } z \in \Omega.$$

Im Gegensatz zur Situation bei Funktionen einer reellen Variablen, wo bekanntlich jede stetige Funktion eine Stammfunktion besitzt, müssen für die Existenz einer komplexen Stammfunktion wieder Integrabilitätsbedingungen erfüllt sein.

Folgerung 24.1 (Existenz komplexer Stammfunktionen) Sei $\Omega \subset \mathbb{C}$ offen und einfach zusammenhängend. Für $f \in C^1(\Omega, \mathbb{C})$, $f = u + iv$, sind folgende Aussagen äquivalent:

- (a) f besitzt auf Ω eine komplexe Stammfunktion.
- (b) f ist komplex differenzierbar auf Ω .

BEWEIS: Sei $F = U + iV$ Stammfunktion von f , also $u + iv = U_x - iU_y = V_y + iV_x$ nach (24.15). Es folgt $U, V \in C^2(\Omega)$ und

$$u_x - v_y = V_{yx} - V_{xy} = 0 \quad \text{und} \quad u_y + v_x = U_{xy} - U_{yx} = 0.$$

Gelten umgekehrt die Cauchy-Riemann Gleichungen für f , so sind die 1-Formen $u dx - v dy$ sowie $u dy + v dx$ geschlossen, besitzen also Stammfunktionen $U, V : \Omega \rightarrow \mathbb{R}$ nach Satz 23.6. Für U, V verifiziert man leicht die Cauchy-Riemann Gleichungen, also ist $F = U + iV$ komplex differenzierbar und mit (24.15) erhalten wir $F' = f$. \square

Im folgenden schreiben wir $D_r(z_0) = \{z \in \mathbb{C} : |z - z_0| < r\}$, und setzen

$$\int_{\partial D_r(z_0)} f(z) dz = \int_{\gamma} f(z) dz, \quad \text{für } \gamma : [0, 2\pi] \rightarrow \partial D_r(z_0), \gamma(t) = z_0 + re^{it}.$$

Bekanntlich hängt das Integral nicht von der Wahl der Parametrisierung von $\partial D_r(z_0)$ ab, solange der Kreis im positiven Sinn (also gegen den Uhrzeigersinn) durchlaufen wird. Der folgende Satz besagt unter anderem, dass eine holomorphe Funktion auf einer Kreisscheibe schon durch ihre Werte auf dem Rand der Kreisscheibe bestimmt ist. Für reell differenzierbare Funktionen ist das keineswegs so, zum Beispiel gibt es (viele) Funktionen $\eta \in C^\infty(\mathbb{R})$ mit $\eta(x) = 0$ für $|x| \geq 1$.

Satz 24.3 (Cauchy Integralformel) Sei $f \in C^1(\Omega, \mathbb{C})$ holomorph und $\overline{D_r(z_0)} \subset \Omega$. Dann gilt für alle $z \in D_r(z_0)$

$$f(z) = \frac{1}{2\pi i} \int_{\partial D_r(z_0)} \frac{f(\zeta)}{\zeta - z} d\zeta.$$

BEWEIS: In $\Omega \setminus \{z\}$ ist der Kreis $\partial D_r(z_0)$ homotop zu $\partial D_\varepsilon(z)$, jeweils mit der mathematisch positiven Orientierung, falls $0 < \varepsilon < r - |z - z_0|$. Die Konstruktion der Homotopie sei den LeserInnen überlassen. Nun ist die Funktion $\zeta \mapsto f(\zeta)/(\zeta - z)$ auf $\Omega \setminus z$ komplex differenzierbar nach der Quotientenregel, und damit das Kurvenintegral homotopieinvariant nach (der nachfolgenden Bemerkung zu) Satz 24.2 Es ergibt sich

$$\begin{aligned} \frac{1}{2\pi i} \int_{\partial D_r(z_0)} \frac{f(\zeta)}{\zeta - z} d\zeta &= \frac{1}{2\pi i} \int_{\partial D_\varepsilon(z)} \frac{f(\zeta)}{\zeta - z} d\zeta \\ (\zeta = z + \varepsilon e^{it}) &= \frac{1}{2\pi i} \int_0^{2\pi} \frac{f(z + \varepsilon e^{it})}{\varepsilon e^{it}} i\varepsilon e^{it} dt \\ &= \frac{1}{2\pi} \int_0^{2\pi} f(z + \varepsilon e^{it}) dt. \end{aligned}$$

Da f im Punkt z stetig ist, geht die rechte Seite gegen $f(z)$ mit $\varepsilon \rightarrow 0$. \square

Satz 24.4 (holomorph = analytisch) Sei $f \in C^1(\Omega, \mathbb{C})$ komplex differenzierbar und $\overline{D_r(z_0)} \subset \Omega$. Dann gilt für alle $z \in D_r(z_0)$ die Potenzreihenentwicklung

$$f(z) = \sum_{k=0}^{\infty} a_k (z - z_0)^k \quad \text{mit} \quad a_k = \frac{1}{2\pi i} \int_{\partial D_r(z_0)} \frac{f(\zeta)}{(\zeta - z_0)^{k+1}} d\zeta.$$

Insbesondere ist f auf Ω unendlich oft komplex differenzierbar.

BEWEIS: Aus der Cauchyschen Integralformel, Satz 24.3, erhalten wir durch Entwicklung des Integranden in eine geometrische Reihe für $z \in D_r(z_0)$ die Darstellung

$$f(z) = \frac{1}{2\pi i} \int_{\partial D_r(z_0)} \frac{f(\zeta)}{\zeta - z_0} \frac{1}{1 - \frac{z-z_0}{\zeta-z_0}} d\zeta = \frac{1}{2\pi i} \int_{\partial D_r(z_0)} \frac{f(\zeta)}{\zeta - z_0} \sum_{k=0}^{\infty} \left(\frac{z-z_0}{\zeta-z_0}\right)^k d\zeta.$$

Aus der Linearität des Kurvenintegrals folgt mit der Definition der a_k

$$\begin{aligned} \frac{1}{2\pi i} \int_{\partial D_r(z_0)} \frac{f(\zeta)}{\zeta - z_0} \sum_{k=0}^n \left(\frac{z-z_0}{\zeta-z_0}\right)^k d\zeta &= \frac{1}{2\pi i} \sum_{k=0}^n (z-z_0)^k \int_{\partial D_r(z_0)} \frac{f(\zeta)}{(\zeta-z_0)^{k+1}} d\zeta \\ &= \sum_{k=0}^n a_k (z-z_0)^k. \end{aligned}$$

Das komplexe Kurvenintegral ist wie in Lemma 23.3 abgeschätzt durch Länge der Kurve mal Supremum des Integranden. Es folgt mit $M = \max_{|\zeta-z_0|=r} |f(\zeta)|$ für $|z-z_0| < r$

$$\begin{aligned} \left| f(z) - \sum_{k=0}^n a_k (z-z_0)^k \right| &= \left| \frac{1}{2\pi i} \int_{\partial D_r(z_0)} \frac{f(\zeta)}{\zeta - z_0} \sum_{k=n+1}^{\infty} \left(\frac{z-z_0}{\zeta-z_0}\right)^k d\zeta \right| \\ &\leq \frac{1}{2\pi} 2\pi r \frac{M}{r} \sum_{k=n+1}^{\infty} \left(\frac{|z-z_0|}{r}\right)^k \\ &= \frac{M}{1 - \frac{|z-z_0|}{r}} \left(\frac{|z-z_0|}{r}\right)^{n+1} \rightarrow 0 \quad \text{mit } n \rightarrow \infty. \end{aligned}$$

Damit ist die Potenzreihendarstellung bewiesen.

Sei nun allgemein $f(z) = \sum_{k=0}^{\infty} a_k (z-z_0)^k$ eine auf $D_r(z_0)$ konvergente Potenzreihe. Die Polynome $f_n(z) = \sum_{k=0}^n a_k (z-z_0)^k$ sind komplex differenzierbar, also gilt auf $D_r(z_0)$ nach Satz 24.1

$$Df_n = \begin{pmatrix} c_n & -d_n \\ d_n & c_n \end{pmatrix} \quad \text{wobei } c_n + id_n = f'_n = \sum_{k=1}^n k a_k (z-z_0)^{k-1}.$$

Die gliedweise differenzierte Potenzreihe $g(z) = \sum_{k=1}^{\infty} a_k (z-z_0)^{k-1}$ konvergiert aber ebenfalls lokal gleichmäßig auf $D_r(z_0)$, siehe Lemma 3.1 in Kapitel 5. Da partielle Ableitungen eindimensionale Ableitungen sind, können wir die bekannte Vertauschbarkeit von Konvergenz und Ableitung, Satz 3.2 in Kapitel 5, hier anwenden und erhalten mit $n \rightarrow \infty$ auf $D_r(z_0)$

$$Df = \begin{pmatrix} c & -d \\ d & a \end{pmatrix} \quad \text{wobei } c + id = g.$$

Nach Satz 24.1 ist f also komplex differenzierbar mit $f' = g$. Durch Induktion ergibt sich, dass f auf $D_r(z_0)$ unendlich oft komplex differenzierbar ist. \square

Wir sehen jetzt noch deutlicher, dass komplexe Differenzierbarkeit eine viel stärkere Bedingung darstellt als reelle Differenzierbarkeit. Die Existenz von Potenzreihendarstellungen ist sehr einschneidend, zum Beispiel ist eine holomorphe Funktion $f : \Omega \rightarrow \mathbb{C}$, deren Nullstellenmenge einen Häufungspunkt in Ω hat, automatisch die Nullfunktion. Dies folgt jetzt aus dem Identitätssatz für Potenzreihen, siehe Satz 4.11 in Kapitel 2. Für reell differenzierbare Funktionen ist eine solche Aussage keineswegs wahr. Aus Zeitgründen müssen wir den Ausflug in die Funktionentheorie nun leider beenden.

25 Diffeomorphismen

Thema dieses und des folgenden Abschnitts ist die lokale Lösbarkeit nichtlinearer Gleichungen. Sei $\Omega \subset \mathbb{R}^n$ offen und $f \in C^1(\Omega, \mathbb{R}^m)$, und es sei schon eine Lösung von $f(x_0) = y_0$ gegeben. Dann stellen wir uns folgende Fragen:

- (1) Hat die Gleichung $f(x) = y$ zu jedem y nahe bei y_0 eine Lösung x nahe bei x_0 ?
- (2) Ist x_0 die einzige Lösung von $f(x) = y_0$ in einer Umgebung von x_0 ?
- (3) Falls nicht, wie sieht die Lösungsmenge $f^{-1}\{y_0\}$ nahe bei x_0 aus?

Betrachten wir erst den Fall, dass f affin-linear ist. Wegen $f(x_0) = y_0$ hat f dann die Form $f(x) = y_0 + A(x - x_0)$ mit $A \in L(\mathbb{R}^n, \mathbb{R}^m)$ und es gilt

$$f(x) = y \quad \Leftrightarrow \quad A(x - x_0) = y - y_0.$$

Die Lineare Algebra liefert folgende, globale Aussagen:

- (1) Es gibt eine Lösung für alle $y \in \mathbb{R}^m \quad \Leftrightarrow \quad \text{rang } A = m.$
- (2) Es gibt höchstens eine Lösung $x \in \mathbb{R}^n \quad \Leftrightarrow \quad \ker A = \{0\} \quad \Leftrightarrow \quad \text{rang } A = n.$
- (3) $f^{-1}\{y_0\}$ ist affiner Unterraum der Dimension $n - \text{rang } A.$

Sei nun $f \in C^1(\Omega, \mathbb{R}^m)$ beliebig, und $R_f(\xi) = f(x_0 + \xi) - (f(x_0) + Df(x_0)\xi)$ für $\xi \in \mathbb{R}^n$ hinreichend klein. Ist $f(x_0) = y_0$, so folgt mit $A = Df(x_0)$

$$f(x) = y \quad \Leftrightarrow \quad A(x - x_0) + R_f(x - x_0) = y - y_0.$$

Wir hoffen, dass sich dies als Störung der linearen Gleichung behandeln lässt, so dass sich die Aussagen (1), (2) und (3) lokal geeignet übertragen. In diesem Abschnitt geht es um den Fall $n = m$, das heißt es gibt genausoviele Unbekannte wie Gleichungen.

Definition 25.1 Eine Abbildung $f : U \rightarrow V$ zwischen offenen Mengen $U, V \subset \mathbb{R}^n$ heißt *Diffeomorphismus der Klasse C^r* , wobei $r \in \mathbb{N} \cup \{\infty\}$, falls f bijektiv ist und sowohl f als auch f^{-1} sind r -mal stetig differenzierbar.

Beispiel 25.1 Sei $I \subset \mathbb{R}$ ein offenes Intervall. Ist $f \in C^1(I)$ mit $f' > 0$ auf ganz I (bzw. $f' < 0$ auf ganz I), so ist $J := f(I)$ ein offenes Intervall und $f : I \rightarrow J$ ein C^1 -Diffeomorphismus. Dies ergibt sich aus Folgendem:

- $f : I \rightarrow J$ ist bijektiv, da streng monoton wachsend. Nach dem Zwischenwertsatz, genauer Satz 2.2 in Kapitel 3, ist J wieder ein offenes Intervall.
- Die Umkehrabbildung $g = f^{-1}$ ist differenzierbar mit $g' = 1/(f' \circ g)$, insbesondere ist g von der Klasse C^1 , vgl. Kapitel 4, Satz 1.6.

Umgekehrt: ist $f : I \rightarrow f(I)$ ein C^1 -Diffeomorphismus mit Umkehrabbildung $g : f(I) \rightarrow I$, so ergibt die Kettenregel

$$g(f(x)) = x \quad \Rightarrow \quad g'(f(x))f'(x) = 1 \quad \Rightarrow \quad f'(x) \neq 0.$$

Nach dem Zwischenwertsatz gilt entweder $f' > 0$ auf ganz I oder $f' < 0$ auf ganz I . Insbesondere ist f streng monoton.

Zum Beispiel ist die Abbildung $f : (-1, 1) \rightarrow (-1, 1)$, $f(x) = x^3$, zwar bijektiv, genauer streng monoton wachsend, und von der Klasse C^1 , aber sie ist kein C^1 -Diffeomorphismus, denn es gilt $f'(0) = 0$. Die Umkehrabbildung

$$g : (-1, 1) \rightarrow (-1, 1), g(y) = \begin{cases} \sqrt[3]{y} & \text{für } y \geq 0 \\ -\sqrt[3]{-y} & \text{für } y < 0 \end{cases}$$

ist im Punkt $y = 0$ nicht differenzierbar.

Beispiel 25.2 Sei $U = \{(r, \theta) \in \mathbb{R}^2 : r > 0, 0 < \theta < 2\pi\}$ und $V = \mathbb{R}^2 \setminus \{(x, 0) : x \geq 0\}$. Die Polarkoordinatenabbildung

$$f : U \rightarrow V, f(r, \theta) = (r \cos \theta, r \sin \theta),$$

ist ein Diffeomorphismus der Klasse C^∞ . Prüfen Sie nach, dass die Umkehrabbildung $g : V \rightarrow U$ wie folgt gegeben ist:

$$g(x, y) = \begin{cases} (\sqrt{x^2 + y^2}, \arccos \frac{x}{\sqrt{x^2 + y^2}}) & \text{für } y \geq 0 \\ (\sqrt{x^2 + y^2}, 2\pi - \arccos \frac{x}{\sqrt{x^2 + y^2}}) & \text{für } y < 0. \end{cases}$$

Die Funktion \arccos ist unendlich oft differenzierbar auf dem Intervall $(-1, 1)$, also ist g unendlich oft differenzierbar für $y \in V$, $y \neq 0$. Aber für $x < 0$ gilt alternativ die Formel

$$g(x, y) = (\sqrt{x^2 + y^2}, \frac{\pi}{2} + \arccos \frac{y}{\sqrt{x^2 + y^2}}).$$

Somit ist in der Tat $g \in C^\infty(V, U)$.

Beispiel 25.3 (Inversion) Die Inversion an der Standardsphäre $\mathbb{S}^{n-1} = \{x \in \mathbb{R}^n : |x| = 1\}$ ist der Diffeomorphismus

$$f : \mathbb{R}^n \setminus \{0\} \rightarrow \mathbb{R}^n \setminus \{0\}, f(x) = \frac{x}{|x|^2}.$$

Die Abbildung $f = f^{-1}$ ist von der Klasse C^∞ . Die beschränkte Menge $U = \{x \in \mathbb{R}^n : 0 < |x| < 1\}$ wird auf die unbeschränkte Menge $V = \{y \in \mathbb{R}^n : |y| > 1\}$ abgebildet.

Lemma 25.1 (Ableitung der Umkehrfunktion) Seien $U \subset \mathbb{R}^n$, $V \subset \mathbb{R}^m$ offene Mengen und $f : U \rightarrow V$ sei bijektiv mit Umkehrabbildung $g : V \rightarrow U$. Sind f in x_0 und g in $y_0 = f(x_0)$ differenzierbar, so ist $Df(x_0) \in L(\mathbb{R}^n, \mathbb{R}^m)$ invertierbar, insbesondere $m = n$, und es gilt $Dg(y_0) = Df(x_0)^{-1}$.

BEWEIS: Aus $g(f(x)) = x$ und $f(g(y)) = y$ folgt mit der Kettenregel

$$Dg(y_0)Df(x_0) = \text{Id}_{\mathbb{R}^n} \quad Df(x_0)Dg(y_0) = \text{Id}_{\mathbb{R}^m}.$$

Also ist $Df(x_0)$ injektiv und surjektiv, das heißt invertierbar, und die Dimensionsformel impliziert $m = n$. \square

Seien $U \subset \mathbb{R}^n$ und $V \subset \mathbb{R}^m$ offene Mengen. Theoretisch hätten wir in unserer Definition 25.1 eines Diffeomorphismus $f : U \rightarrow V$ die Möglichkeit $m \neq n$ zulassen können. Lemma 25.1 zeigt, dass dies nichts gebracht hätte, denn es gilt automatisch $n = m$; man spricht von der Invarianz der Dimension unter Diffeomorphismen. Eine Bijektion $f : U \rightarrow V$ heißt Homeomorphismus, wenn f und f^{-1} beide stetig sind. Nach einem Satz von Brouwer (1910) gilt die Invarianz der Dimension, also $m = n$, schon für Homeomorphismen. Dies ist auf dem Hintergrund eines Beispiels von Peano (1890) zu sehen, der surjektive, stetige Abbildungen von einem Intervall auf die Fläche eines Quadrats konstruiert hat. Der Satz von Brouwer wird mit dem Konzept des Abbildungsgrads bewiesen, das in der algebraischen Topologie oder der nichtlinearen Funktionalanalysis eingeführt wird.

Bekanntlich heißt $\det Df(x_0)$ Jacobideterminante von f in x_0 . In der Situation von Lemma 25.1 folgt aus dem Determinantenmultiplikationssatz

$$(25.1) \quad \det Dg(y_0) \det Df(x_0) = 1 \quad \text{für } y_0 = f(x_0).$$

Lemma 25.2 (Höhere Ableitungen der Umkehrfunktion) *Seien $U, V \subset \mathbb{R}^n$ offen und $f : U \rightarrow V$ bijektiv. Ist $f \in C^r(U, V)$ für ein $r \in \mathbb{N} \cup \{\infty\}$ und ist die Umkehrabbildung $g : V \rightarrow U$ differenzierbar, so ist auch $g \in C^r(V, U)$.*

BEWEIS: Nach Lemma 25.1 gilt $Dg = (Df)^{-1} \circ g$. Die Cramersche Regel für die Berechnung der inversen Matrix impliziert

$$(25.2) \quad \frac{\partial g_j}{\partial y_i} = (-1)^{i+j} \frac{M_{ij}(Df)}{\det(Df)} \circ g.$$

Dabei bezeichnet $M_{ij}(Df)$ die Determinante der Matrix, die aus Df durch Streichen der i -ten Zeile und j -ten Spalte entsteht. Wir zeigen die Behauptung durch Induktion über $r \in \mathbb{N}$. Da g nach Voraussetzung differenzierbar und somit stetig ist, vgl. Satz 18.2 in Kapitel 6, ist für $f \in C^1$ die rechte Seite in (25.2) stetig als Produkt, Quotient und Verkettung stetiger Funktionen, und damit $g \in C^1$. Ist $f \in C^r$ und induktiv schon $g \in C^{r-1}$, so ist die rechte Seite von der Klasse C^{r-1} als Produkt, Quotient und Verkettung von C^{r-1} -Funktionen, siehe Folgerung 18.1 in Kapitel 6, und damit $g \in C^r$, was zu zeigen war. \square

Nach diesen Vorüberlegungen wollen wir nun die Frage der Existenz einer Lösung angehen. Dazu die folgenden Definitionen.

Definition 25.2 *Eine Folge $x_k, k \in \mathbb{N}$, in einem metrischen Raum (X, d) heißt Cauchyfolge, wenn zu jedem $\varepsilon > 0$ ein $N \in \mathbb{R}$ gibt mit*

$$d(x_k, x_l) < \varepsilon \quad \text{für alle } k, l > N.$$

Ein metrischer Raum heißt vollständig, wenn jede Cauchyfolge x_k in X konvergiert, das heißt es gibt ein $x \in X$ mit $d(x, x_k) \rightarrow 0$ mit $k \rightarrow \infty$.

Natürlich ist \mathbb{R}^n mit der Euklidischen Abstandsfunktion ein vollständiger metrischer Raum. Aber jede abgeschlossene Teilmenge $A \subset \mathbb{R}^n$ ist mit dem euklidischen Abstand auch ein vollständiger metrischer Raum, denn eine Cauchyfolge $x_k \in A$ ist auch Cauchyfolge in \mathbb{R}^n und konvergiert damit gegen ein $x \in \mathbb{R}^n$, und es gilt $x \in A$ wegen A abgeschlossen.

Satz 25.1 (Fixpunktsatz von Banach) Sei (X, d) ein vollständiger metrischer Raum, und $F : X \rightarrow X$ eine Kontraktion, das heißt es gibt ein $\theta \in [0, 1)$ mit

$$(25.3) \quad d(F(x), F(y)) \leq \theta d(x, y) \quad \text{für alle } x, y \in X.$$

Dann gibt es genau ein $x \in X$ mit $F(x) = x$.

BEWEIS: Die Eindeutigkeit ist klar, denn aus $F(x) = x$ und $F(y) = y$ folgt

$$d(x, y) = d(F(x), F(y)) \leq \theta d(x, y) \quad \Rightarrow \quad d(x, y) = 0, \text{ also } x = y.$$

Um den Fixpunkt zu konstruieren, betrachten wir die rekursiv definierte Folge $x_{n+1} = F(x_n)$ mit beliebigem Startwert $x_0 \in X$. Es folgt aus (25.3) für $n \geq 1$

$$(25.4) \quad d(x_{n+1}, x_n) = d(F(x_n), F(x_{n-1})) \leq \theta d(x_n, x_{n-1}).$$

Wir können uns einen müder werdenden Frosch vorstellen, dessen Sprünge jedes Mal um ein Faktor $\theta \in [0, 1)$ kürzer werden. Wie weit kann der Frosch insgesamt kommen? Es folgt per Induktion aus (25.4)

$$(25.5) \quad d(x_{n+1}, x_n) \leq \theta^n d(x_1, x_0) \quad \text{für } n \in \mathbb{N}_0,$$

und hieraus weiter mit der Dreiecksungleichung

$$d(x_n, x_0) \leq \sum_{j=0}^{n-1} d(x_{j+1}, x_j) \leq \sum_{j=0}^{n-1} \theta^j d(x_1, x_0) \leq \frac{1}{1-\theta} d(x_1, x_0).$$

Indem wir x_n statt x_0 als Startwert auffassen, haben wir für $m > n$

$$(25.6) \quad d(x_m, x_n) \leq \frac{1}{1-\theta} d(x_{n+1}, x_n) \leq \frac{\theta^n}{1-\theta} d(x_1, x_0).$$

Also ist $(x_n)_{n \in \mathbb{N}_0}$ eine Cauchyfolge, und konvergiert nach Voraussetzung gegen ein $x \in X$. Da F nach Voraussetzung Lipschitzstetig ist (mit Konstante θ), folgt

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

und die Existenz des Fixpunkts ist gezeigt. □

Wir bemerken, dass wir a priori abschätzen können, wie weit die Iteration im n -ten Schritt noch vom gesuchten Fixpunkt entfernt ist, und zwar folgt mit $m \rightarrow \infty$ aus (25.6)

$$d(x, x_n) \leq \frac{\theta^n}{1-\theta} d(x_1, x_0).$$

Satz 25.2 (über inverse Funktionen) Sei $\Omega \subset \mathbb{R}^n$ offen und $f \in C^1(\Omega, \mathbb{R}^n)$. Ist $Df(x_0) \in L(\mathbb{R}^n, \mathbb{R}^n)$ invertierbar, so gibt es eine offene Umgebung U von x_0 , so dass gilt:

- (a) $V = f(U)$ ist offene Umgebung von $y_0 = f(x_0)$
- (b) $f : U \rightarrow V$ ist Diffeomorphismus der Klasse C^1 .

Zusatz. Ist $f \in C^r(\Omega, \mathbb{R}^n)$ für ein $r \in \mathbb{N} \cup \{\infty\}$, so ist $g = (f|_U)^{-1} \in C^r(V, \mathbb{R}^n)$.

BEWEIS: Wir können $x_0 = 0$ und $y_0 = 0$ annehmen, andernfalls betrachten wir die Abbildung

$$\tilde{f} : \tilde{\Omega} = \{\xi \in \mathbb{R}^n : x_0 + \xi \in \Omega\} \rightarrow \mathbb{R}^n, \tilde{f}(\xi) = f(x_0 + \xi) - f(x_0).$$

Schritt 1 *Formulierung als Fixpunktproblem*

Mit $A := Df(0)$ und $R_f(x) := f(x) - Ax$ können wir die Gleichung wie folgt umformen:

$$f(x) = y \Leftrightarrow Ax + R_f(x) = y \Leftrightarrow x = A^{-1}(y - R_f(x)).$$

Für $y \in \mathbb{R}^n$ definieren wir also $\phi_y : \Omega \rightarrow \mathbb{R}^n$, $\phi_y(x) = A^{-1}(y - R_f(x))$, und erhalten

$$(25.7) \quad f(x) = y \Leftrightarrow \phi_y(x) = x.$$

Schritt 2 *Konstruktion der Lösung*

Wir bestimmen $\varepsilon > 0$ und $\delta > 0$, so dass für jedes $y \in B_\varepsilon(0)$ die Abbildung $\phi_y : \overline{B_\delta(0)} \rightarrow \overline{B_\delta(0)}$ eine Kontraktion ist. Setze $\Lambda = \|A^{-1}\| \in (0, \infty)$. Nach Voraussetzung ist $DR_f(x) = Df(x) - A$ stetig mit $DR_f(0) = 0$, folglich gibt es ein $\delta_0 > 0$ mit

$$\overline{B_{\delta_0}(0)} \subset \Omega \quad \text{und} \quad \|DR_f(x)\| \leq \frac{1}{2\Lambda} \quad \text{für } |x| \leq \delta_0.$$

Aus dem Schrankensatz, siehe Satz 19.2 in Kapitel 7, folgt

$$(25.8) \quad |x_1|, |x_2| \leq \delta_0 \Rightarrow |R_f(x_1) - R_f(x_2)| \leq \frac{1}{2\Lambda} |x_1 - x_2|.$$

Wir berechnen nun

$$\begin{aligned} |\phi_y(x_1) - \phi_y(x_2)| &= |A^{-1}(y - R_f(x_1)) - A^{-1}(y - R_f(x_2))| \\ &= |A^{-1}(R_f(x_1) - R_f(x_2))| \\ &\leq \Lambda |R_f(x_1) - R_f(x_2)|. \end{aligned}$$

Also folgt aus (25.8) die Kontraktionseigenschaft

$$(25.9) \quad |x_1|, |x_2| \leq \delta_0 \Rightarrow |\phi_y(x_1) - \phi_y(x_2)| \leq \frac{1}{2} |x_1 - x_2|.$$

Bisher ist $y \in \mathbb{R}^n$ noch beliebig. Wir schätzen weiter ab

$$\begin{aligned} |\phi_y(x)| &= |A^{-1}(y - R_f(x))| \\ &\leq \|A^{-1}\| (|y| + |R_f(x)|) \\ &= \Lambda (|y| + |R_f(x) - R_f(0)|) \quad (\text{da } R_f(0) = 0) \\ &\leq \Lambda |y| + \frac{1}{2} |x| \quad \text{für } |x| \leq \delta_0 \text{ nach (25.8)}. \end{aligned}$$

Also folgt für $\delta \in (0, \delta_0]$, wenn wir $\varepsilon = \delta/(2\Lambda) > 0$ wählen,

$$(25.10) \quad |x| \leq \delta, |y| < \varepsilon \Rightarrow |\phi_y(x)| < \Lambda\varepsilon + \frac{1}{2}\delta = \delta.$$

Wegen (25.10) und (25.9) ist $\phi_y : \overline{B_\delta(0)} \rightarrow \overline{B_\delta(0)}$ eine wohldefinierte Kontraktion mit Konstante $\theta = 1/2$. Nach dem Banachschen Fixpunktsatz gibt es zu jedem $y \in B_\varepsilon(0)$ genau ein $x \in \overline{B_\delta(0)}$ mit $\phi_y(x) = x$, das heißt $f(x) = y$ nach (25.7). Tatsächlich ist $x \in B_\delta(0)$, denn nach (25.10) gilt $|x| = |\phi_y(x)| < \delta$. Die Mengen $V = B_\varepsilon(0)$ und $U = f^{-1}(V) \cap B_\delta(0)$ sind offene Umgebungen des Nullpunkts, siehe Satz 16.4 in Kapitel 6 für die Offenheit von U , und wie gezeigt ist $f : U \rightarrow V$ bijektiv. Insbesondere ist Behauptung (a) bewiesen.

Schritt 3 Differenzierbarkeit der inversen Abbildung

Sei $g : V \rightarrow U$ die Umkehrabbildung von $f : U \rightarrow V$. Dann gilt

$$(25.11) \quad |g(y)| = |\phi_y(g(y))| \leq \Lambda|y| + \frac{1}{2}|g(y)| \quad \Rightarrow \quad |g(y)| \leq 2\Lambda|y|.$$

Insbesondere ist g stetig im Nullpunkt mit $g(0) = 0$. Wir zeigen nun $Dg(0) = A^{-1}$. Für $y \neq 0$ ist $g(y) \neq 0$ und es gilt die Abschätzung

$$\frac{|g(y) - A^{-1}y|}{|y|} = \frac{|\phi_y(g(y)) - A^{-1}y|}{|y|} = \frac{|A^{-1}(R_f(g(y)))|}{|y|} \leq \Lambda \frac{|R_f(g(y))|}{|g(y)|} \frac{|g(y)|}{|y|}.$$

Mit $y \rightarrow 0$ geht die rechte Seite aber gegen Null, denn $|g(y)|/|y| \leq 2\Lambda$ nach (25.11) und $|R_f(x)|/|x| \rightarrow 0$ mit $x = g(y) \rightarrow 0$. Damit ist $Dg(0) = A^{-1}$ gezeigt. Um die Differenzierbarkeit für alle $y \in V$ zu bekommen, wählen wir $\delta > 0$ so klein, dass $\det Df(x) \neq 0$ für alle $x \in B_\delta(0)$. Ist dann $y \in V$, so sind die Voraussetzungen des Satzes im Punkt $x = g(y)$ erfüllt, und es folgt aus dem Bewiesenen $Dg(y) = Df(x)^{-1}$.

Lemma 25.2 liefert schließlich $g \in C^1(V, U)$. Ist $f \in C^r(U, V)$ für ein $r \in \mathbb{N} \cup \{\infty\}$, so ist $g \in C^r(V, U)$, ebenfalls nach Lemma 25.2. \square

Als unmittelbare Konsequenz des Satzes halten wir fest:

Folgerung 25.1 Sei $\Omega \subset \mathbb{R}^n$ offen und $f \in C^1(\Omega, \mathbb{R}^n)$. Ist $Df(x)$ invertierbar für alle $x \in \Omega$, so ist $f(\Omega) \subset \mathbb{R}^n$ offen.

BEWEIS: Nach Satz 25.2 hat jeder Punkt $y \in f(\Omega)$ eine offene Umgebung $V \subset f(\Omega)$. \square

Beispiel 25.4 Wie wir in Beispiel 25.1 gesehen haben, bildet eine eindimensionale Funktion $f : (a, b) \rightarrow \mathbb{R}$ mit $f' \neq 0$ das gesamte Definitionsintervall diffeomorph auf das Bildintervall ab, das heißt es gilt eine globale Version des Umkehrsatzes. Das folgende Beispiel zeigt, dass eine entsprechende Aussage für Funktionen mehrerer Variabler im allgemeinen nicht wahr ist. In reellen Koordinaten $z = x + iy$ lautet die komplexe Exponentialfunktion

$$\exp : \mathbb{R}^2 \rightarrow \mathbb{R}^2, \quad \exp(x, y) = (e^x \cos y, e^x \sin y).$$

Es gilt $\exp(\mathbb{R}^2) = \mathbb{R}^2 \setminus \{0\}$. Die Jacobideterminante von \exp ist nirgends Null, genauer gilt

$$D \exp(x, y) = \begin{pmatrix} e^x \cos y & -e^x \sin y \\ e^x \sin y & e^x \cos y \end{pmatrix} \quad \Rightarrow \quad \det D \exp(x, y) = e^{2x} \neq 0.$$

Die Abbildung ist jedoch nicht injektiv, denn es ist $\exp(x, y + 2k\pi) = \exp(x, y)$ für alle $k \in \mathbb{Z}$.

Wir diskutieren jetzt ein Beispiel, das unter anderem beim Übergang von der Lagrangefunktion zur Hamiltonfunktion in der klassischen Mechanik auftritt. Für eine differenzierbare Funktion $L : U \rightarrow \mathbb{R}$ definiert man die zugehörige Gradientenabbildung durch

$$f : U \rightarrow V, f(x) = DL(x) \quad \text{wobei } V = f(U) \subset \mathbb{R}^n.$$

Ist $f : U \rightarrow V$ injektiv und $g : V \rightarrow U$ die Umkehrfunktion von f , so können wir die Legendretransformierte oder duale Funktion von L wie folgt erklären:

$$H : V \rightarrow \mathbb{R}, H(y) = \left(\sum_{i=1}^n y_i x_i - L(x) \right) \Big|_{x=g(y)}.$$

Für $L : \mathbb{R}^n \rightarrow \mathbb{R}$, $L(x) = \sqrt{1 + |x|^2}$ berechnet man zum Beispiel $V = \{y \in \mathbb{R}^n : |y| < 1\}$ und

$$f(x) = \frac{x}{\sqrt{1 + |x|^2}} \quad g(y) = \frac{y}{\sqrt{1 - |y|^2}} \quad \text{sowie}$$

$$H(y) = \left(\sum_{i=1}^n y_i x_i - \sqrt{1 + |x|^2} \right) \Big|_{x=g(y)} = \sqrt{1 - |y|^2}.$$

Satz 25.3 (Involutionseigenschaft der Legendretransformation) *Sei für $L \in C^2(U)$ die Gradientenabbildung $f : U \rightarrow V$, $f(x) = DL(x)$, diffeomorph, und sei H die Legendretransformierte von L . Dann folgt $DH = g$ mit $g = f^{-1} \in C^1(V, U)$, insbesondere $H \in C^2(V)$, und die Legendretransformierte von H ist wieder L :*

$$L(x) = \left(\sum_{i=1}^n y_i x_i - H(y) \right) \Big|_{y=f(x)} \quad \text{für alle } x \in U.$$

BEWEIS: Wir berechnen unter Verwendung von $\frac{\partial L}{\partial x_i}(g(y)) = f_i(g(y)) = y_i$

$$\frac{\partial H}{\partial y_j} = \frac{\partial}{\partial y_j} \left(\sum_{i=1}^n y_i g_i(y) - L \circ g \right) = g_j + \sum_{i=1}^n y_i \frac{\partial g_i}{\partial y_j} - \sum_{i=1}^n \frac{\partial L}{\partial x_i} \circ g \frac{\partial g_i}{\partial y_j} = g_j,$$

also gilt $DH = g$. Die Darstellung von L folgt mit $y = f(x)$ in der Definition von H . \square

Geometrisch gelangen wir zu der Legendretransformation, wenn wir für den Graph einer Funktion L die Schar $\{E_x : x \in U\}$ der affinen Tangentialebenen betrachten:

$$E_x = \{(\xi, z) \in \mathbb{R}^n \times \mathbb{R} : z = L(x) + DL(x)(\xi - x)\} \quad \text{für } x \in U.$$

Wir wollen diese Schar durch die Steigungen $y = DL(x)$ parametrisieren; natürlich muss dazu die Gradientenabbildung injektiv sein. Die Legendretransformierte $H : V \rightarrow \mathbb{R}$ von L ist dann definiert, und es folgt

$$E_x = \{(\xi, z) \in \mathbb{R}^n \times \mathbb{R} : z = \sum_{i=1}^n y_i \xi_i - H(y)\} \quad \text{für } y = DL(x).$$

Umgekehrt kann das Problem betrachtet werden, zu einer gegebenen Ebenenschar $E_y = \{(\xi, z) \in \mathbb{R}^n \times \mathbb{R} : z = \sum_{i=1}^n y_i \xi_i - H(y)\}$ einen Graphen $L : U \rightarrow \mathbb{R}$ zu bestimmen, dessen affine Tangentialebenen gerade die E_y sind. Es muss dann H die Legendretransformierte der gesuchten Funktion L sein, und unter den Annahmen von Satz 25.3 kann L wiederum durch Legendretransformation von H bestimmt werden.

Satz 25.4 (Youngsche Ungleichung) Sei $U \subset \mathbb{R}^n$ konvex und $L \in C^2(U)$ mit $D^2L(x) > 0$ für alle $x \in U$. Dann ist die Legendretransformierte $H \in C^2(V)$ definiert, wobei $V = DL(U)$, und es gilt

$$\sum_{i=1}^n x_i y_i \leq L(x) + H(y) \quad \text{für alle } x \in U, y \in V.$$

Gleichheit tritt ein genau für $y = DL(x)$ beziehungsweise äquivalent $x = DH(y)$.

BEWEIS: Die Gradientenabbildung $f \in C^1(U, V)$, $f(x) = DL(x)$, ist injektiv, denn wegen $Df(x) = D^2L(x) > 0$ folgt für $x_0, x_1 \in U$ mit $x_0 \neq x_1$

$$\langle f(x_1) - f(x_0), x_1 - x_0 \rangle = \int_0^1 \langle Df((1-t)x_1 + tx_0)(x_1 - x_0), x_1 - x_0 \rangle dt > 0.$$

Da außerdem $Df(x)$ invertierbar ist für alle $x \in U$, ist f ein C^1 -Diffeomorphismus nach dem Umkehrsatz. Betrachte nun für festes $y \in V$ die Funktion $\varphi \in C^2(U)$ mit

$$\varphi(x) = L(x) + H(y) - \sum_{i=1}^n y_i x_i.$$

Dann gilt $\varphi(g(y)) = 0$ nach Definition von H sowie $D\varphi(g(y)) = DL(g(y)) - y = 0$ nach Definition von $g = f^{-1}$, und $D^2\varphi(x) = D^2L(x) > 0$ für alle $x \in U$. Damit folgt die Behauptung aus Satz 20.4 in Kapitel 7. \square

Aus dem Satz folgt für die Legendretransformierte die Darstellung

$$H(y) = \inf_{x \in U} \left(\sum_{i=1}^n y_i x_i - L(x) \right).$$

Mit dieser Formel kann die Legendretransformierte auch dann definiert werden, wenn L nicht differenzierbar, sondern lediglich konvex ist.

26 Implizite Funktionen

Wir betrachten jetzt den Fall eines unterbestimmten Systems, wenn es also weniger Gleichungen gibt als Unbekannte. Wir können die Funktion dann wie folgt schreiben, indem wir die Variablen in zwei Gruppen einteilen:

$$f : \Omega \rightarrow \mathbb{R}^k, \quad f = f(x, y), \quad \text{wobei } (x, y) \in \Omega \subset \mathbb{R}^m \times \mathbb{R}^k = \mathbb{R}^n.$$

Gegeben sei eine Lösung (x_0, y_0) der Gleichung $f(x, y) = z_0$. Wir interessieren uns dafür, wie die Lösungsmenge dieser Gleichung nahe bei (x_0, y_0) aussieht. Können wir nach y auflösen, d. h. die Lösungsmenge als Graph einer Funktion $y = g(x)$ darstellen?

Beispiel 26.1 Betrachte die Gleichung

$$f(x, y) = x^2 + y^2 = 1 \quad \text{für } (x, y) \in \mathbb{R} \times \mathbb{R} = \mathbb{R}^2$$

Sei eine Lösung $(x_0, y_0) \in \mathbb{R}^2$ gegeben, also ein Punkt auf dem Einheitskreis. Ist $y_0 > 0$, so ist die Lösungsmenge nahe bei (x_0, y_0) Graph der Funktion $y = \sqrt{1 - x^2}$. Analog haben wir im Fall $y_0 < 0$ die lokale Graphendarstellung $y = -\sqrt{1 - x^2}$. Dagegen ist die Lösungsmenge in keiner Umgebung von $(1, 0)$ (und ebenso in keiner Umgebung von $(-1, 0)$) als Graph über der x -Achse darstellbar, denn es gibt die zwei Lösungen $y = \pm\sqrt{1 - x^2}$.

Es kann auch vorkommen, dass (x_0, y_0) der einzige Punkt mit $f(x_0, y_0) = z_0$ ist, z.B. löst nur der Nullpunkt die Gleichung $x^2 + y^2 = 0$ in \mathbb{R}^2 .

Beispiel 26.2 Sei $f : \mathbb{R}^m \times \mathbb{R}^k \rightarrow \mathbb{R}^k$ linear. Wir unterteilen die $k \times (m + k)$ -Matrix von f in eine $k \times m$ -Matrix A und eine $k \times k$ -Matrix B , d. h.

$$f(x, y) = Ax + By \quad \text{mit } A \in L(\mathbb{R}^m, \mathbb{R}^k), \quad B \in L(\mathbb{R}^k, \mathbb{R}^k).$$

Die Gleichung $Ax + By = z_0$ hat zu festem $x \in \mathbb{R}^m$ eine eindeutige Auflösung nach y dann und nur dann, wenn B invertierbar ist. Ist das der Fall, so lautet die Auflösung $y = B^{-1}(z_0 - Ax)$.

Allgemein schreiben wir die Jacobimatrix von $f = f(x, y)$ in der Form

$$Df(x, y) = (D_x f, D_y f) \in (\mathbb{R}^{k \times m}, \mathbb{R}^{k \times k}).$$

Wenn wir nach $y = g(x)$ auflösen wollen, so sollte nach Beispiel 26.2 die Ableitung $D_y f(x_0, y_0) \in \mathbb{R}^{k \times k}$ invertierbar sein. In den Anwendungen ist die Einteilung in die beiden Variablengruppen nicht immer vorgegeben, das heißt es könnte nach verschiedenen Gruppen von je k Variablen aufgelöst werden. So kann der Einheitskreis in einer Umgebung von $(1, 0)$ zwar nicht als Graph $y = g(x)$ geschrieben werden, wohl aber als Graph $x = g(y)$, und außer in den vier Punkten $\pm e_1, \pm e_2$ könnte sowohl nach x als auch nach y aufgelöst werden.

Merkregel. Die Ableitung nach den Variablen, nach denen aufgelöst werden soll, muss invertierbar sein. Im Spezialfall $k = 1$ bedeutet das $\frac{\partial f}{\partial y}(x_0, y_0) \neq 0$.

Satz 26.1 (über implizite Funktionen) Sei $\Omega \subset \mathbb{R}^m \times \mathbb{R}^k$ offen und $f \in C^1(\Omega, \mathbb{R}^k)$. Ist $f(x_0, y_0) = z_0$ und $D_y f(x_0, y_0) \in L(\mathbb{R}^k, \mathbb{R}^k)$ invertierbar, so gibt es offene Umgebungen U von x_0 bzw. V von y_0 sowie eine Funktion $g \in C^1(U, V)$ mit

$$(26.12) \quad \{(x, y) \in U \times V : f(x, y) = z_0\} = \{(x, g(x)) : x \in U\},$$

insbesondere ist $g(x_0) = y_0$. Die Funktion g hat die Ableitung

$$(26.13) \quad Dg(x_0) = -(D_y f(x_0, y_0))^{-1} D_x f(x_0, y_0).$$

Zusatz. Für jedes $r \in \mathbb{N} \cup \{\infty\}$ gilt die Implikation

$$f \in C^r(\Omega, \mathbb{R}^k) \quad \Rightarrow \quad g \in C^r(U, \mathbb{R}^k).$$

BEWEIS: Wir verwenden einen Trick, um den Satz über inverse Funktionen anwenden zu können, und zwar betrachten wir $F : \Omega \rightarrow \mathbb{R}^m \times \mathbb{R}^k$, $F(x, y) = (x, f(x, y))$. Es gilt

$$DF = \begin{pmatrix} E_m & 0 \\ D_x f & D_y f \end{pmatrix} \in \begin{pmatrix} \mathbb{R}^{m \times m} & \mathbb{R}^{m \times k} \\ \mathbb{R}^{k \times m} & \mathbb{R}^{k \times k} \end{pmatrix}.$$

Es folgt $\det DF(x_0, y_0) = \det D_y f(x_0, y_0) \neq 0$ nach Voraussetzung. Nach dem Umkehrsatz existieren offene Umgebungen $U_0 \times V$ von (x_0, y_0) sowie W von (x_0, z_0) , so dass $F : U_0 \times V \rightarrow W$ diffeomorph ist. Wir bezeichnen die zugehörige Umkehrabbildung mit $G \in C^1(W, U_0 \times V)$. Ist $(x, z) \in W$, also $(x, z) = (x, f(x, y))$ mit $(x, y) \in U_0 \times V$ nach Konstruktion, so folgt

$$G(x, z) = G(x, f(x, y)) = G(F(x, y)) = (x, y).$$

Also ist G von der Form $G(x, z) = (x, g_0(x, z))$ mit $g_0 \in C^1(W, \mathbb{R}^k)$. Sei nun $U = \{x \in U_0 : (x, z_0) \in W\}$. Da W offen in $\mathbb{R}^m \times \mathbb{R}^k$, ist U offen in \mathbb{R}^m und für $(x, y) \in U \times V$ gilt

$$\begin{aligned} f(x, y) = z_0 &\Leftrightarrow F(x, y) = (x, z_0) \\ &\Leftrightarrow (x, y) = G(x, z_0) \quad (\text{da } (x, z_0) \in W) \\ &\Leftrightarrow y = g_0(x, z_0). \end{aligned}$$

Also gilt die Aussage des Satzes mit $g(x) = g_0(x, z_0)$. Die Formel für die Ableitung folgt aus der Kettenregel:

$$f(x, g(x)) = z_0 \quad \Rightarrow \quad D_x f(x_0, y_0) + D_y f(x_0, y_0) Dg(x_0) = 0.$$

□

Beispiel 26.3 Wir wollen als triviales Beispiel untersuchen, wie die Nullstelle eines quadratischen Polynoms von seinen Koeffizienten abhängt. Sei

$$f : \mathbb{R}^2 \times \mathbb{R} \rightarrow \mathbb{R}, \quad f(p, q, \lambda) = \lambda^2 + 2p\lambda + q = (\lambda + p)^2 - (p^2 - q).$$

Die Menge $N = \{(p, q, \lambda) \in \mathbb{R}^2 \times \mathbb{R} : f(p, q, \lambda) = 0\}$ ist die Vereinigung der beiden disjunkten Graphen $G^\pm = \{(p, q, \lambda^\pm(p, q)) : p^2 > q\}$, wobei $\lambda^\pm(p, q) = -p \pm \sqrt{p^2 - q}$, mit der Menge $\overline{G}^+ \cap \overline{G}^- = \{(p, q, \lambda) : p^2 = q, \lambda = -p\}$. Im Fall

$$\frac{\partial f}{\partial \lambda}(p_0, q_0, \lambda_0) = 2(\lambda_0 + p_0) \neq 0 \quad \Leftrightarrow \quad p_0^2 - q_0 > 0$$

liegt (p_0, q_0, λ_0) in einem der beiden Graphen G^+ oder G^- , und N ist in einer Umgebung $U \times V$ als Graph von λ^+ oder λ^- darstellbar. Ist dagegen

$$\frac{\partial f}{\partial \lambda}(p_0, q_0, \lambda_0) = 2(\lambda_0 + p_0) = 0 \quad \Leftrightarrow \quad p_0^2 - q_0 = 0,$$

so macht der implizite Funktionensatz keine Aussage. Tatsächlich lässt sich die Menge N in keiner Umgebung von (p_0, q_0, λ_0) als Graph $\lambda = \lambda(p, q)$ darstellen: für $p^2 < q$ hat die Gleichung überhaupt keine Lösung, für $p^2 = q$ genau eine und für $p^2 > q$ die zwei verschiedenen Lösungen $\lambda^\pm(p, q)$.

Beispiel 26.4 Betrachte jetzt $f : \mathbb{R}^n \times \mathbb{R} \rightarrow \mathbb{R}$, $f(b, \lambda) = \lambda^n + b_{n-1}\lambda^{n-1} + \dots + b_0$. Sei λ_0 eine einfache Nullstelle von $f(a, \lambda)$ für $a \in \mathbb{R}^n$ fest, das heißt es gilt

$$f(a, \lambda) = (\lambda - \lambda_0) q(\lambda) \quad \text{für ein Polynom } q(\lambda) \text{ mit } q(\lambda_0) \neq 0.$$

Es folgt $\frac{\partial f}{\partial \lambda}(a, \lambda_0) = q(\lambda_0) \neq 0$. Nach dem Satz über implizite Funktionen existiert eine Umgebung $U \times V$ von (a, λ_0) , so dass zu jedem $b \in U$ genau eine Nullstelle $\lambda(b) \in V$ von $f(b, \cdot)$ existiert. Diese hängt unendlich oft differenzierbar von b ab, und es gilt für $0 \leq i \leq n-1$

$$\frac{\partial \lambda}{\partial b_i}(a) = -\left(\frac{\partial f}{\partial \lambda}(a, \lambda_0)\right)^{-1} \frac{\partial f}{\partial b_i}(a, \lambda_0) = -\frac{\lambda_0^i}{n \lambda_0^{n-1} + (n-1) a_{n-1} \lambda_0^{n-2} + \dots + a_1}.$$

Wir kommen nun zu einer geometrischen Anwendung des Satzes über implizite Funktionen. Ist $f \in C^1(\mathbb{R}^2)$ und $z_0 \in \mathbb{R}$, so kann im allgemeinen die Niveaumenge

$$M = \{(x, y) \in \mathbb{R}^2 : f(x, y) = z_0\}$$

Punkte enthalten, in denen M nicht lokal wie eine Linie aussieht, z.B. Kreuzungspunkte von Linien oder isolierte Punkte. Ist aber $Df(x, y) \neq 0$ für alle $(x, y) \in M$, so ist M nach dem impliziten Funktionensatz in der Nähe jedes Punkts als C^1 -Graph über der x -Achse oder der y -Achse darstellbar und damit wirklich eine Höhenlinie im strengen Sinn des Worts. Teilmengen des \mathbb{R}^n , die lokal aussehen wie ein Unterraum, heißen Untermannigfaltigkeiten; vergleiche hierzu auch die Diskussion des Tangentialraums von differenzierbaren Graphen mittels Blowup, siehe Kapitel 6 Abschnitt 3.

Definition 26.1 Sei $1 \leq m \leq n$. Eine Menge $M \subset \mathbb{R}^n$ heisst m -dimensionale Untermannigfaltigkeit des \mathbb{R}^n der Klasse C^r , wobei $r \in \mathbb{N} \cup \{\infty\}$, falls gilt: zu jedem $p \in M$ gibt es eine offene Umgebung $\Omega \subset \mathbb{R}^n$ und einen C^r -Diffeomorphismus $\phi : \Omega \rightarrow \phi(\Omega)$ mit

$$\phi(M \cap \Omega) = (\mathbb{R}^m \times \{0\}) \cap \phi(\Omega).$$

Wir nennen den Diffeomorphismus ϕ eine (lokale) Plättung von M . Im Einzelfall kann der Nachweis, dass eine gegebene Menge $M \subset \mathbb{R}^n$ eine Untermannigfaltigkeit ist, anhand der Definition mühevoll sein. Für Mengen, die als Niveaumengen einer Funktion gegeben sind, liefert jedoch der Satz über implizite Funktionen folgendes Kriterium.

Satz 26.2 (Untermannigfaltigkeitskriterien) Sei $M \subset \mathbb{R}^n$ und $m + k = n$. Dann sind folgende Aussagen äquivalent:

- (1) M ist eine m -dimensionale Untermannigfaltigkeit der Klasse C^r .
- (2) Niveaumengenkriterium: Zu $p \in M$ gibt es eine offene Umgebung $\Omega \subset \mathbb{R}^n$ und eine Funktion $f \in C^r(\Omega, \mathbb{R}^k)$, so dass $M \cap \Omega = f^{-1}(0)$ und $\text{rang } Df = k$ auf Ω .
- (3) Graphenkriterium: Zu $p \in M$ gibt es eine offene Umgebung $U \times V \subset \mathbb{R}^m \times \mathbb{R}^k$ und $g \in C^r(U, V)$, so dass nach geeigneter Permutation der Koordinaten gilt:

$$M \cap (U \times V) = \{(x, g(x)) : x \in U\}.$$

BEWEIS: Wir zeigen (1) \Rightarrow (2) \Rightarrow (3) \Rightarrow (1). Es gelte (1), das heißt zu jedem $p \in M$ gibt es eine lokale Plättung $\phi : \Omega \rightarrow \phi(\Omega)$ der Klasse C^r mit $p \in \Omega$. Sei $\pi^\perp : \mathbb{R}^m \times \mathbb{R}^k \rightarrow \mathbb{R}^k$ die Projektion auf den zweiten Faktor. Dann folgt mit $f = \pi^\perp \circ \phi$ für $q \in \Omega$

$$f(q) = 0 \quad \Leftrightarrow \quad \phi(q) \in \mathbb{R}^m \times \{0\} \quad \Leftrightarrow \quad q \in \phi^{-1}(\mathbb{R}^m \times \{0\}) = M \cap \Omega.$$

Außerdem gilt $\text{rang } Df = \text{rang } (\pi^\perp \circ D\phi) = k$ und $f \in C^r(\Omega, \mathbb{R}^k)$.

Ist (2) erfüllt, so ist nach evtl. Permutation der Koordinaten $D_y f(p)$ invertierbar, wobei $(x, y) \in \mathbb{R}^n = \mathbb{R}^m \times \mathbb{R}^k$, und (3) folgt aus dem Satz über implizite Funktionen.

Für (3) \Rightarrow (1) können wir annehmen, dass die Graphendarstellung ohne Permutation der Koordinaten gilt. Wir setzen $\Omega = U \times V$ und

$$\phi : \Omega \rightarrow \phi(\Omega), \quad \phi(x, y) = (x, y - g(x)).$$

Dann ist $\phi \in C^r(\Omega, \mathbb{R}^n)$ injektiv und es gilt

$$D\phi(x, y) = \begin{pmatrix} E_m & 0 \\ -Dg(x) & E_k \end{pmatrix} \Rightarrow \det D\phi(x, y) = 1 \quad \text{für alle } (x, y) \in \Omega.$$

Also ist $\phi : \Omega \rightarrow \phi(\Omega)$ ein Diffeomorphismus der Klasse C^r nach dem Umkehrsatz. Da $(x, y) \in M \cap \Omega$ genau wenn $y = g(x)$, also $\phi(x, y) \in \mathbb{R}^k \times \{0\}$, ist die in (1) verlangte lokale Plättung gefunden. \square

Beispiel 26.5 Die Sphäre $\mathbb{S}^m = \{x \in \mathbb{R}^{m+1} : |x| = 1\}$ ist eine m -dimensionale Untermannigfaltigkeit im \mathbb{R}^{m+1} der Klasse C^∞ . Denn es gilt

$$\mathbb{S}^m = f^{-1}(0) \quad \text{für } f : \mathbb{R}^{m+1} \setminus \{0\} \rightarrow \mathbb{R}, \quad f(x) = |x|^2 - 1,$$

und $Df(x) \neq 0$ für alle $x \in \mathbb{R}^{m+1} \setminus \{0\}$. Die Behauptung folgt also aus Satz 26.2.

Definition 26.2 Ein Vektor $v \in \mathbb{R}^n$ heisst Tangentialvektor von $M \subset \mathbb{R}^n$ im Punkt $p \in M$, falls es eine Abbildung $\gamma : (-\varepsilon, \varepsilon) \rightarrow M$ gibt mit $\gamma(0) = p$, $\gamma'(0) = v$. Die Menge der Tangentialvektoren von M im Punkt p wird mit $T_p M$ bezeichnet.

Folgerung 26.1 Sei $M \subset \mathbb{R}^n$ eine m -dimensionale C^1 -Untermannigfaltigkeit, und $n = m + k$. Ist $p \in M \cap \Omega = f^{-1}(0)$ für eine Funktion $f \in C^1(\Omega, \mathbb{R}^k)$ mit $\text{rang } Df = k$ auf Ω , so gilt

$$T_p M = \ker Df(p).$$

Insbesondere ist $T_p M$ ein m -dimensionaler Unterraum des \mathbb{R}^n .

BEWEIS: Für $\gamma : (-\varepsilon, \varepsilon) \rightarrow M$ mit $\gamma(0) = p$ und $\gamma'(0) = w$ gilt

$$0 = \frac{d}{dt} f(\gamma(t))|_{t=0} = Df(\gamma(0))\gamma'(0) = Df(p)w \quad \Rightarrow \quad T_p M \subset \ker Df(p).$$

Nach Satz 26.2 gibt es andererseits, nach eventueller Permutation der Koordinaten, offene Mengen $U \subset \mathbb{R}^m$, $V \subset \mathbb{R}^k$ mit $p \in U \times V$, sowie $g \in C^1(U, V)$ mit

$$M \cap (U \times V) = \{(x, g(x)) : x \in U\}.$$

Die Graphenabbildung $G \in C^1(U, \mathbb{R}^n)$, $G(x) = (x, g(x))$, bildet nach M ab. Schreiben wir $p = (x_0, g(x_0))$ für geeignetes $x_0 \in U$, so folgt für alle $v \in \mathbb{R}^m$

$$DG(x_0)v = \frac{d}{dt}G(x_0 + tv)|_{t=0} \in T_pM \quad \Rightarrow \quad \text{Bild } DG(x_0) \subset T_pM.$$

Zusammenfassend ist $\text{Bild } DG(x_0) \subset T_pM \subset \ker Df(p)$. Aber $DG(x_0)$ ist injektiv, denn $DG(x_0)v = (v, Dg(x_0)v)$, und wegen $\text{rang } Df(p) = k$ folgt mit der Dimensionsformel

$$\dim \text{Bild } DG(x_0) = m = n - \text{rang } Df(p) = \dim \ker Df(p).$$

Also gilt $\text{Bild } DG(x_0) = \ker Df(p) = T_pM$. □

Die Folgerung zeigt, dass der Tangentialraum einer m -dimensionalen Untermannigfaltigkeit tatsächlich ein m -dimensionaler Vektorraum ist. Dies zeigt, dass die Dimension einer C^1 -Untermannigfaltigkeit wohldefiniert ist, das heißt es kann nicht Plättungen mit verschiedenen Dimensionen von M geben. Auf die Idee wäre vermutlich auch kaum jemand gekommen. Wir kommen nun zur sogenannten Lagrange-Multiplikatorenregel.

Folgerung 26.2 (Extrema mit Nebenbedingungen) Sei $\Omega \subset \mathbb{R}^n$ offen, $f \in C^1(\Omega, \mathbb{R}^k)$ und $\varphi \in C^1(\Omega)$. Gilt dann für ein $p \in f^{-1}(0)$

(1) $\varphi(q) \geq \varphi(p)$ für alle $q \in \Omega$ mit $f(q) = 0$,

(2) $\text{rang } Df(p) = k$,

so gibt es $\lambda_1, \dots, \lambda_k \in \mathbb{R}$ mit $\text{grad } \varphi(p) = \sum_{i=1}^k \lambda_i \text{grad } f_i(p)$.

BEWEIS: Nach eventueller Verkleinerung von Ω ist $\text{rang } Df = k$ auf Ω und $M := f^{-1}(0)$ ist eine m -dimensionale Untermannigfaltigkeit, wobei $m = n - k$. Ist $\gamma : (-\varepsilon, \varepsilon) \rightarrow M$ mit $\gamma(0) = p$ und $\gamma'(0) = v$, so hat $\varphi \circ \gamma$ in $t = 0$ ein lokales Minimum und folglich

$$0 = \frac{d}{dt}\varphi(\gamma(t))|_{t=0} = \langle \text{grad } \varphi(p), v \rangle,$$

das heißt $\text{grad } \varphi(p) \in (T_pM)^\perp$. Da $f_j|_M \equiv 0$, gilt analog $\text{grad } f_j(p) \in (T_pM)^\perp$ für $1 \leq j \leq k$. Aber $\dim (T_pM)^\perp = k$ nach Folgerung 26.1, und die Vektoren $\text{grad } f_j(p)$ sind die Zeilenvektoren der Matrix $Df(p)$ mit Rang k . Wegen der Gleichheit von Zeilenrang und Spaltenrang ist $\{\text{grad } f_j(p) : 1 \leq j \leq k\}$ eine Basis von $(T_pM)^\perp$, und die Behauptung folgt. □

Beispiel 26.6 Für eine symmetrische Matrix $B \in \mathbb{R}^{n \times n}$ betrachten wir hier nochmals das Problem, die quadratische Form $\varphi(x) = \langle Bx, x \rangle$ zu minimieren unter der Nebenbedingung $f(x) = |x|^2 - 1 = 0$. Da $\mathbb{S}^{n-1} = f^{-1}(0)$ kompakt und φ stetig, wird das Infimum in einem Punkt $x_0 \in \mathbb{S}^{n-1}$ angenommen. Nach Folgerung 26.2 gibt es dann ein $\lambda \in \mathbb{R}$ mit $\text{grad } \varphi(x_0) = \lambda \text{grad } f(x_0)$, also $Bx_0 = \lambda x_0$. Somit hat jede symmetrische Matrix B mindestens einen Eigenvektor, vergleiche Kapitel 7, Satz 20.3.

Wir haben die Sätze über inverse und implizite Funktionen im Endlichdimensionalen formuliert, damit das Wesentliche nicht durch zuviel Abstraktion verdeckt wird. An der Verallgemeinerung auf Abbildungen zwischen Banachräumen besteht aber großes Interesse: in

den Anwendungen ist die Gleichung $f(x) = y$ zum Beispiel eine Differential- oder Integralgleichung, die durch eine gesuchte Funktion x in einem geeigneten Funktionenraum X gelöst werden soll. Eine Inspektion des Beweises des Umkehrsatzes ergibt, dass die Konstruktion der inversen Abbildung einschließlich ihrer Differenzierbarkeit ohne wesentliche Änderungen auch dann richtig ist, wenn f eine offene Teilmenge des Banachraums X in den Banachraum Y abbildet. Dabei muss die euklidische Norm von A durch die Operatornorm $\|A\| = \sup_{\|x\|=1} \|Ax\|$ ersetzt werden, und der Raum $L(\mathbb{R}^n, \mathbb{R}^m)$ muss ersetzt werden durch

$$L(X, Y) = \{A : X \rightarrow Y \mid A \text{ linear}, \|A\| < \infty\}.$$

Es ist leicht zu sehen, dass die Bedingung $\|A\| < \infty$ äquivalent dazu ist, dass die lineare Abbildung A stetig ist. Dies ist eine *conditio sine qua non*, darum nehmen wir sie in die Definition von $L(X, Y)$ auf. Insbesondere wird in der Definition der Differenzierbarkeit verlangt, dass $Df(x_0) \in L(X, Y)$.

Expliziten Gebrauch von den Koordinaten des \mathbb{R}^n haben wir nur gemacht, um die höhere Differenzierbarkeit der Inversen g zu etablieren. Im Unendlichdimensionalen wird dazu alternativ die sogenannte Neumannsche Reihe benutzt.

27 Existenz und Eindeutigkeit für das Anfangswertproblem

Aus Zeitgründen müssen wir uns hier auf einen zentralen Aspekt aus der Theorie der gewöhnlichen Differentialgleichungen beschränken, nämlich die Existenz und Eindeutigkeit von Lösungen von Anfangswertaufgaben. Als Einstieg betrachten wir das Problem, die zeitliche Entwicklung einer Population (Bakterien, Bevölkerung, Kontostand, Atome, ...) vorherzusagen oder rückwärtig zu bestimmen. Wir interessieren uns also für die Größe $x(t)$ der Population innerhalb eines gewissen Zeitintervalls I . Dabei ist zur Zeit $t_0 \in I$ ein Wert $x(t_0) = x_0$ gegeben. Wir sprechen von einem Anfangswertproblem, auch wenn t_0 nicht der linke Endpunkt von I ist. Je nach Kontext sind viele verschiedene Wachstumsgesetze denkbar:

Beim natürlichen Wachstum (Kontostand, radioaktiver Zerfall) ist die Wachstums- oder Zerfallsgeschwindigkeit proportional zur vorhandenen Menge, mit fester Zuwachs- bzw. Zerfallsrate $\alpha \in \mathbb{R}$:

$$x' = \alpha x.$$

Lösung des zugehörigen Anfangswertproblems ist $x(t) = x_0 e^{\alpha(t-t_0)}$. Dabei ergibt sich die Eindeutigkeit wie folgt: ist $x : I \rightarrow \mathbb{R}$ beliebige Lösung der Differentialgleichung, so gilt

$$\frac{d}{dt} \left(e^{-\alpha t} x(t) \right) = e^{-\alpha t} (x'(t) - \alpha x(t)) = 0,$$

also $x(t) = c e^{\alpha t}$ für ein $c \in \mathbb{R}$. Aus der Anfangsbedingung folgt weiter $x_0 = c e^{\alpha t_0}$ wie behauptet. Das sogenannte logistische Wachstum ist das Vermehrungsgesetz

$$x' = (\alpha - \beta x)x = \alpha x - \beta x^2 \quad \text{mit } \alpha, \beta > 0.$$

Als Motivation für den Zusatzterm $-\beta x$ kann das Beispiel einer Schafherde dienen, für die nur eine feste Weidefläche zur Verfügung steht. Ab einem gewissen Schwellenwert sollte die Zahl der Tiere dann wieder abnehmen. Im Fall $x_0 = x(t_0) > 0$ ist eine Lösung für $t \geq t_0$ durch folgende Formel gegeben, wobei die Eindeutigkeit schon weniger offensichtlich ist:

$$x(t) = \frac{1}{\frac{\beta}{\alpha} + \left(\frac{1}{x_0} - \frac{\beta}{\alpha} \right) e^{-\alpha(t-t_0)}}.$$

Für $t \rightarrow \infty$ konvergiert die Lösung gegen α/β . Das sogenannte Räuber-Beute Modell von Volterra und Lotka betrachtet zwei Populationen $x(t)$ und $y(t)$, zum Beispiel Gänse und Füchse. Bei zuviel Füchsen wird die Vermehrungsrate der Gänse negativ, bei zuwenig Gänsen ist die Vermehrungsrate der Füchse negativ ($\alpha, \beta, \gamma, \delta$ sind positive Konstanten):

$$\begin{aligned} x' &= (\alpha - \beta y)x \\ y' &= (-\gamma + \delta x)y. \end{aligned}$$

Es ergeben sich also zwei gekoppelte Gleichungen, deren Lösung nicht ohne weiteres ermittelt werden kann. Soll das Modell außerdem die jahreszeitliche Änderung der Futtersituation berücksichtigen, so müssen die Koeffizienten durch zeitabhängige Funktionen ersetzt werden. Allgemein interessieren wir uns für folgende Situation.

Definition 27.1 Sei G offen in $\mathbb{R} \times \mathbb{R}^n$ und $f \in C^0(G, \mathbb{R}^n)$, $f = f(t, x)$. Eine Funktion $x \in C^1(I, \mathbb{R}^n)$, $I \subset \mathbb{R}$ Intervall, heißt Lösung der Differentialgleichung $x' = f(\cdot, x)$, falls

$$(27.14) \quad x'(t) = f(t, x(t)) \text{ für alle } t \in I \quad (\text{insbesondere } (t, x(t)) \in G).$$

Gilt außerdem

$$(27.15) \quad x(t_0) = x_0 \text{ für gegebenes } (t_0, x_0) \in G,$$

so heißt x Lösung des zugehörigen Anfangswertproblems.

Die Differentialgleichung heißt autonom, wenn das Vektorfeld f nicht von der Zeit abhängt. Wir können uns $x(t)$ dann als Bahn eines Teilchens vorstellen, dass zur Zeit t_0 in x_0 startet und durch Vorgabe der Momentangeschwindigkeit $x'(t) = f(x(t))$ zur Zeit t gesteuert wird. Im nichtautonomen Fall ist die Steuerung zusätzlich zeitabhängig. Es stellen sich folgende Fragen:

1. Ist eine Lösung des Anfangswertproblems eindeutig bestimmt?
2. Existiert eine Lösung des Anfangswertproblems?
3. Wie hängt die Lösung vom Anfangswert x_0 und dem Vektorfeld f ab?

In der Vorlesung werden aus Zeitgründen nur die ersten beiden Fragen befriedigend beantwortet, wobei wir mit der Eindeutigkeit beginnen.

Beispiel 27.1 Sei $G = \mathbb{R} \times \mathbb{R}$ und $f(t, x) = 2\sqrt{|x|}$. Dann hat das Anfangswertproblem

$$x' = f(\cdot, x) \quad x(0) = 0$$

unendlich viele verschiedene Lösungen in $C^1(\mathbb{R})$, und zwar für $-\infty \leq \alpha \leq 0 \leq \beta \leq \infty$

$$x_{\alpha, \beta}(t) = \begin{cases} -(t - \alpha)^2 & \text{für } t < \alpha \\ 0 & \text{für } \alpha \leq t \leq \beta \\ (t - \beta)^2 & \text{für } t > \beta. \end{cases}$$

Für die Eindeutigkeit ist die Stetigkeit des Vektorfeldes f demnach nicht ausreichend.

Satz 27.1 (Eindeutigkeit der Lösung) Sei $G \subset \mathbb{R} \times \mathbb{R}^n$ offen. Ist $f \in C^0(G)$ mit $D_x f \in C^0(G)$, so hat das Anfangswertproblem

$$x' = f(\cdot, x), \quad x(t_0) = x_0$$

höchstens eine Lösung $x \in C^1(I, \mathbb{R}^n)$.

Für den Beweis brauchen wir die nachstehende Folgerung aus dem Schrankensatz, vgl. Satz 19.2 und Folgerung 19.1 in Kapitel 7.

Lemma 27.1 Sei $f \in C^0(G, \mathbb{R}^n)$, $f = f(t, x)$, mit $D_x f \in C^0(G, \mathbb{R}^{n \times n})$. Dann ist f lokal Lipschitzstetig bzgl. x : zu $(t_0, x_0) \in G$ gibt es eine Umgebung $U_\varepsilon(t_0, x_0) := \{(t, x) : |t - t_0| < \varepsilon, |x - x_0| < \varepsilon\} \subset G$ und eine Konstante $L \in [0, \infty)$ mit

$$|f(t, x_1) - f(t, x_2)| \leq L |x_1 - x_2| \quad \text{für alle } (t, x_1), (t, x_2) \in U_\varepsilon(t_0, x_0).$$

BEWEIS VON SATZ 27.1 Seien $x_1, x_2 \in C^1(I, \mathbb{R}^n)$ zwei Lösungen des Anfangswertproblems. Wir zeigen erst $x_1 = x_2$ auf $(t_0 - \delta, t_0 + \delta)$ für ein geeignetes $\delta > 0$. Seien dazu $\varepsilon > 0$, $L < \infty$ wie in Lemma 27.1. Wähle $\delta > 0$ mit $(t, x_1(t)), (t, x_2(t)) \in U_\varepsilon(t_0, x_0)$ für $|t - t_0| < \delta$. Für $u(t) = |x_1(t) - x_2(t)|^2$ berechnen wir

$$\begin{aligned} |u'| &= 2|\langle x_1 - x_2, x_1' - x_2' \rangle| \\ &= 2|\langle x_1 - x_2, f(\cdot, x_1) - f(\cdot, x_2) \rangle| \\ &\leq 2|x_1 - x_2| |f(\cdot, x_1) - f(\cdot, x_2)| \quad (\text{Cauchy-Schwarz}) \\ &\leq 2Lu. \end{aligned}$$

Daraus folgen für $t \in (t_0 - \delta, t_0 + \delta)$ die Ungleichungen

$$\frac{d}{dt}(e^{-2Lt}u(t)) \leq 0 \quad \text{und} \quad \frac{d}{dt}(e^{2Lt}u(t)) \geq 0.$$

Durch Integration ergibt sich $u(t) \leq u(t_0)e^{2L|t-t_0|}$, und wegen $u(t_0) = |x_1(t_0) - x_2(t_0)|^2 = 0$ nach Voraussetzung ist $u(t) = 0$ bzw. $x_1(t) = x_2(t)$ für alle $t \in (t_0 - \delta, t_0 + \delta)$.

Jetzt zeigen wir $x_1 = x_2$ auf ganz I . Die Menge $M = \{t \in I : x_1(t) = x_2(t)\}$ ist nichtleer, da $t_0 \in M$ nach Voraussetzung, und ist abgeschlossen in I , da x_1, x_2 stetig sind. Nach Schritt 1 gibt es zu $t \in M$ aber ein $\delta > 0$ mit $(t - \delta, t + \delta) \subset M$, d. h. M ist offen. Daraus folgt $M = I$. \square

Wir kommen nun zur Frage der Existenz. Das folgende Beispiel zeigt, dass wir im allgemeinen nur eine zeitlich lokale Lösung erwarten können.

Beispiel 27.2 Betrachte $f \in C^\infty(\mathbb{R} \times \mathbb{R})$, $f(t, x) = x^2$. Das zugehörige Anfangswertproblem

$$x' = x^2, \quad x(0) = 1$$

hat die Lösung $x : (-\infty, 1) \rightarrow \mathbb{R}$, $x(t) = 1/(1-t)$. Die Lösung ist nach rechts nicht fortsetzbar, denn es gilt $\lim_{t \nearrow 1} x(t) = \infty$.

In Satz 1.3, Kapitel 5, wurde gezeigt, dass die Grenzfunktion einer Folge stetiger Funktionen bei gleichmäßiger Konvergenz stetig ist. Die folgende Konsequenz ist für uns jetzt wichtig.

Satz 27.2 (Vollständigkeit von $C^0(I, \mathbb{R}^n)$) Sei $I \subset \mathbb{R}$ ein Intervall. Dann ist $C^0(I, \mathbb{R}^n)$, versehen mit der Supremumsnorm, ein Banachraum.

BEWEIS: Sei $x_k \in C^0(I, \mathbb{R}^n)$ eine Cauchyfolge bezüglich der Supremumsnorm, das heißt $\|x_k - x_l\|_I < \varepsilon$ für $k, l > N$. Dann folgt

$$|x_k(t) - x_l(t)| < \varepsilon \quad \text{für alle } t \in I, k, l > N.$$

Da \mathbb{R}^n vollständig ist, existiert $x(t) = \lim_{k \rightarrow \infty} x_k(t)$. Mit $l \rightarrow \infty$ folgt

$$|x_k(t) - x(t)| \leq \varepsilon \quad \text{für alle } t \in I, k > N,$$

bzw. $\|x_k - x\|_I \leq \varepsilon$ für $k > N$. Somit konvergiert x_k bezüglich der Supremumsnorm, d.h. gleichmäßig, gegen x , und es gilt $x \in C^0(I)$ nach Lemma 1.3 in Kapitel 5, Analysis 1. \square

Eine entscheidende Beobachtung zur Konstruktion der lokalen Lösung ist, dass das Anfangswertproblem als Integralgleichung geschrieben werden kann.

Lemma 27.2 Sei $f \in C^0(G, \mathbb{R}^n)$ und $(t_0, x_0) \in G$. Für $x : I \rightarrow \mathbb{R}^n$ mit $\{(t, x(t)) : t \in I\} \subset G$ sind folgende Aussagen äquivalent:

(a) $x \in C^1(I, \mathbb{R}^n)$ ist Lösung des Anfangswertproblems

$$x'(t) = f(t, x(t)) \text{ für alle } t \in I, \quad x(t_0) = x_0.$$

(b) $x \in C^0(I, \mathbb{R}^n)$ erfüllt die Gleichung

$$x(t) = x_0 + \int_{t_0}^t f(s, x(s)) ds \text{ für alle } t \in I.$$

BEWEIS: (b) folgt aus (a) durch Integration von t_0 bis t , mit dem Hauptsatz der Differential- und Integralrechnung. Umgekehrt folgt aus (b) die Stetigkeit der Funktion $t \mapsto f(t, x(t))$, und hieraus wieder mit dem Hauptsatz $x'(t) = f(t, x(t))$ durch Differentiation. \square

Satz 27.3 (Kurzzeitexistenzsatz von Picard-Lindelöf) Sei $f \in C^0(G, \mathbb{R}^n)$, mit $D_x f \in C^0(G, \mathbb{R}^{n \times n})$, und $(t_0, x_0) \in G$. Dann gibt es ein $\delta > 0$, so dass das Anfangswertproblem

$$x' = f(\cdot, x) \text{ auf } I = [t_0 - \delta, t_0 + \delta], \quad x(t_0) = x_0,$$

eine Lösung $x \in C^1(I, \mathbb{R}^n)$ besitzt.

BEWEIS: Als erstes formulieren wir das Problem um in eine Fixpunktgleichung. Zu $(t_0, x_0) \in G$ seien $\varepsilon > 0$ und $L < \infty$ wie in Lemma 27.1 gewählt, und $\delta \in (0, \varepsilon]$ zunächst beliebig. Betrachte im Banachraum $X = C^0(I, \mathbb{R}^n)$ mit Norm $\|\cdot\|_I$ die abgeschlossene Teilmenge

$$A = \{x \in X : |x(t) - x_0| \leq \varepsilon \text{ für alle } t \in I\},$$

sowie die Abbildung

$$F : A \rightarrow X, \quad [F(x)](t) = x_0 + \int_{t_0}^t f(s, x(s)) ds.$$

Nach Lemma 27.2 ist ein Fixpunkt von F , also $F(x) = x$, eine Lösung des Anfangswertproblems. Wir zeigen nun für $\delta > 0$ hinreichend klein folgende Aussagen:

(1) $F(A) \subset A$ (Selbstabbildung)

(2) $\|F(x) - F(y)\|_I \leq \frac{1}{2} \|x - y\|_I$ für alle $x, y \in A$ (Kontraktion).

Aus dem Banachschen Fixpunktsatz, Satz 25.1 in Kapitel 8, folgt dann die Existenz eines Fixpunkts $x \in A$, also der gewünschten Lösung der Anfangswertproblems. Da $f : G \rightarrow \mathbb{R}^n$ stetig, ist $M = \sup\{|f(t, x)| : |t - t_0| \leq \varepsilon, |x - x_0| \leq \varepsilon\} < \infty$, und für $x \in A$ folgt

$$|[F(x)](t) - x_0| = \left| \int_{t_0}^t f(s, x(s)) ds \right| \leq M |t - t_0| \leq M\delta.$$

Also gilt (1) für $\delta \leq \varepsilon/M$. Weiter erhalten wir für $x, y \in A$ aus der Lipschitzbedingung

$$\begin{aligned} |[F(x)](t) - [F(y)](t)| &= \left| \int_{t_0}^t (f(s, x(s)) - f(s, y(s))) ds \right| \\ &\leq |t - t_0| \sup_{s \in I} |f(s, x(s)) - f(s, y(s))| \\ &\leq L\delta \sup_{s \in I} |x(s) - y(s)|. \end{aligned}$$

Für $\delta = \min(\varepsilon/M, 1/(2L))$ folgt also auch Bedingung (2), und der Satz ist bewiesen. \square

Beispiel 27.3 Im Satz von Banach wird der Fixpunkt bekanntlich durch Iteration der Abbildung F mit einem geeigneten Startwert bestimmt. Wir wollen das für folgendes triviale Beispiel explizit durchrechnen:

$$x' = \alpha x, \quad x(0) = 1.$$

Hier ist $G = \mathbb{R} \times \mathbb{R}$, $f(t, x) = \alpha x$ und $t_0 = 0$. Die Iterationsvorschrift lautet

$$[F(x)](t) = 1 + \int_0^t \alpha x(s) ds.$$

Wählen wir als Startfunktion $x_0(t) \equiv 1$, so sind die ersten Iterationsschritte

$$\begin{aligned} x_1(t) &= 1 + \int_0^t \alpha ds = 1 + \alpha t, \\ x_2(t) &= 1 + \int_0^t \alpha(1 + \alpha s) ds = 1 + \alpha t + \frac{\alpha^2 t^2}{2}, \\ x_3(t) &= 1 + \int_0^t \alpha \left(1 + \alpha s + \frac{\alpha^2 s^2}{2}\right) ds = 1 + \alpha t + \frac{\alpha^2 t^2}{2} + \frac{\alpha^3 t^3}{6}. \end{aligned}$$

Durch Induktion sieht man ohne Mühe

$$x_k(t) = \sum_{j=0}^k \frac{(\alpha t)^j}{j!}.$$

Insbesondere konvergiert das Verfahren für $k \rightarrow \infty$ gegen die Lösung $x(t) = e^{\alpha t}$, und zwar lokal gleichmäßig auf ganz \mathbb{R} .

Für $D \subset \mathbb{R}^n$ offen betrachten wir nun Vektorfelder $f : G = (\alpha, \beta) \times D \rightarrow \mathbb{R}^n$ mit $f, D_x f$ stetig, wobei $-\infty \leq \alpha < 0 < \beta \leq \infty$. Zu gegebenen Anfangsdaten $x(0) = x_0 \in D$ liefert der Satz von Picard-Lindelöf eine C^1 -Lösung $x : (-\delta, \delta) \rightarrow D$, und wir fragen uns: unter welchen Voraussetzungen kann das Anfangswertproblem auf dem ganzen Intervall (α, β) gelöst werden? Oder anders:

Was muss passieren, damit die Lösung vorzeitig ihren Geist aufgibt?

Sei t^+ das Supremum aller $t > 0$, so dass das Anfangswertproblem auf $[0, t)$ lösbar ist. Dann ist $t^+ > 0$, und das Anfangswertproblem ist auch auf $[0, t^+)$ noch lösbar: wähle dazu eine Folge $t_k \nearrow t^+$ mit zugehörigen Lösungen $x_k \in C^1([0, t_k), D)$, und definiere

$$x : [0, t^+) \rightarrow D, \quad x|_{[0, t_k)} = x_k.$$

Für $t_k < t_l$ gilt $x_k = x_l$ auf $[0, t_k)$ wegen der eindeutigen Lösbarkeit, Satz 27.1. Damit ist x wohldefinierte Lösung des Anfangswertproblems. Nach Definition des Supremums ist das Anfangswertproblem auf keinem Intervall $[0, t)$ mit $t > t^+$ lösbar; wir nennen $t^+ \in (0, \infty]$ die maximale Lebensdauer der Lösung (Englisch: *lifespan*). Analog bestimmen wir das maximale Alter $t^- < 0$, und erhalten insgesamt ein maximales Intervall (t^-, t^+) , auf dem das Anfangswertproblem eine Lösung hat.

Lemma 27.3 (Divergenz der Lösung) *Sei $f : (\alpha, \beta) \times D \rightarrow \mathbb{R}^n$ mit $f, D_x f$ stetig, wobei $D \subset \mathbb{R}^n$ offen und $-\infty \leq \alpha < 0 < \beta \leq \infty$. Gilt für die maximale Lösung $x \in C^1((t^-, t^+), D)$ des Anfangswertproblems $t^+ < \beta$, so verlässt $x(t)$ für $t \nearrow t^+$ jedes Kompaktum, das heißt für jede kompakte Menge $K \subset D$ gilt $x(t) \notin K$ für t hinreichend nahe bei t^+ . Entsprechendes gilt im Fall $t^- > \alpha$.*

Bemerkung. Im Fall $D = \mathbb{R}^n$ bedeutet das

$$t^+ < \beta \Rightarrow \lim_{t \nearrow t^+} |x(t)| = \infty, \quad \text{bzw.} \quad t^- > \alpha \Rightarrow \lim_{t \searrow t^-} |x(t)| = \infty.$$

BEWEIS: Sei $t^+ < \beta$ und $K \subset D$ kompakt. Angenommen es gibt ein $\tau < t^+$ mit $x(t) \in K$ für alle $t \in (\tau, t^+)$. Da f stetig, ist $M = \sup\{|f(t, x)| : t \in [0, t^+], x \in K\} < \infty$ und es folgt

$$|x(t_2) - x(t_1)| = \int_{t_1}^{t_2} |f(t, x(t))| dt \leq M |t_2 - t_1| \quad \text{für alle } t_{1,2} \in [0, t^+).$$

Somit existiert $x^+ = \lim_{t \nearrow t^+} x(t) \in K$. Nach Satz 27.3 gibt es nun auf einem Intervall $I = (t^+ - \tau, t^+ + \tau)$ eine Lösung y des Anfangswertproblems

$$y' = f(\cdot, y) \text{ auf } I, \quad y(t^+) = x^+.$$

Nach dem Eindeutigkeitsatz stimmen x und y links von t^+ überein, also ergibt Zusammensetzen eine Lösung \tilde{x} der Differentialgleichung auf $[0, t^+ + \tau)$ mit $\tilde{x}(0) = x_0$, im Widerspruch zur Maximalität von t^+ . Der Satz ist damit nicht ganz gezeigt, denn die obige indirekte Annahme schließt nicht ein oszillierendes Verhalten aus, bei dem die Lösung für $t \nearrow t^+$ immer wieder nach K zurückkommt. Aber angenommen es gibt eine Folge $t_k \nearrow t^+$ mit $x(t_k) \in K$. Aus Kompaktheitsgründen gilt $\tilde{K} = \{x \in \mathbb{R}^n : \text{dist}(x, K) \leq \varrho\} \subset D$ für $\varrho > 0$ klein, sowie $\tilde{M} = \sup\{|f(t, x)| : t \in [0, t^+], x \in \tilde{K}\} < \infty$. Wie oben gezeigt ist $\tilde{t}_k = \sup\{t \geq t_k : x(t) \in \tilde{K}\} < t^+$, und es folgt

$$\varrho \leq |x(\tilde{t}_k) - x(t_k)| = \left| \int_{t_k}^{\tilde{t}_k} f(t, x(t)) dt \right| \leq \tilde{M}(\tilde{t}_k - t_k).$$

Für t_k hinreichend nahe bei t^+ ist dies ein Widerspruch, und der Beweis ist komplett. \square

Wir wollen nun im Fall $f : (\alpha, \beta) \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ Wachstumsabschätzungen für die Lösung herleiten, um unter geeigneten Annahmen die Divergenz auszuschließen. Das folgende Argument verallgemeinert die Abschätzung aus dem Eindeutigkeitsatz.

Lemma 27.4 (Gronwall) *Sei $u \in C^0([t_0, t_1])$, und es gelte*

$$u(t) \leq B(t) + \int_{t_0}^t a(s)u(s) ds \quad \text{für alle } t \in [t_0, t_1],$$

wobei $a \in C^0([t_0, t_1])$, $a \geq 0$, und $B \in C^1([t_0, t_1])$. Dann folgt mit $A(t) = \int_{t_0}^t a(s) ds$

$$u(t) \leq e^{A(t)} \left(B(t_0) + \int_{t_0}^t e^{-A(s)} B'(s) ds \right) \quad \text{für alle } t \in [t_0, t_1].$$

BEWEIS: Wir setzen $g(t) = \int_{t_0}^t a(s)u(s) ds$, also $u(t) \leq B(t) + g(t)$, und berechnen

$$\frac{d}{dt}(e^{-A}g) = e^{-A}a(u - g) \leq e^{-A}aB = -\frac{d}{dt}(e^{-A}B) + e^{-A}B'.$$

Integration von t_0 bis t liefert wegen $g(t_0) = 0$ und $A(t_0) = 0$

$$e^{-A(t)}g(t) \leq B(t_0) - e^{-A(t)}B(t) + \int_{t_0}^t e^{-A(s)}B'(s) ds.$$

Multiplikation mit $e^{A(t)}$ und Einsetzen in $u(t) \leq B(t) + g(t)$ ergibt die Ungleichung. \square

Satz 27.4 (Langzeitexistenz bei linearem Wachstum) Sei $f : (\alpha, \beta) \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ mit $f, D_x f$ stetig und $-\infty \leq \alpha < 0 < \beta \leq \infty$. Es gebe $a, b \in C^0((\alpha, \beta))$ mit $a \geq 0$, so dass gilt:

$$|f(t, x)| \leq a(t)|x| + b(t) \quad \text{für alle } t \in (\alpha, \beta), x \in \mathbb{R}^n.$$

Dann ist das Anfangswertproblem $x' = f(\cdot, x)$, $x(0) = x_0 \in \mathbb{R}^n$, auf ganz (α, β) lösbar, und mit $A(t) = \int_0^t a(s) ds$ gilt die Abschätzung

$$(27.16) \quad |x(t)| \leq e^{A(t)} \left(|x_0| + \int_0^t e^{-A(s)} b(s) ds \right) \quad \text{für alle } t \in (\alpha, \beta).$$

BEWEIS: Für $t \in [0, t^+)$ gilt die Abschätzung

$$|x(t)| \leq |x_0| + \int_0^t |f(s, x(s))| ds \leq |x_0| + \int_0^t (a(s)|x(s)| + b(s)) ds,$$

das heißt es ist $|x(t)| \leq B(t) + \int_0^t a(s) ds$ mit $B(t) = |x_0| + \int_0^t b(s) ds$. Aus dem Lemma von Gronwall folgt die gewünschte Abschätzung auf $[0, t^+)$. Wäre $t^+ < \beta$, so gilt $|x(t)| \rightarrow \infty$ mit $t \nearrow t^+$ nach Lemma 27.3, im Widerspruch zur Abschätzung. \square

28 Lineare Differentialgleichungen

Wir betrachten hier das Anfangswertproblem

$$(28.1) \quad x'(t) = A(t)x(t) + b(t) \text{ für } t \in I = (\alpha, \beta), \quad x(t_0) = x_0.$$

Dabei ist $A : I \rightarrow \mathbb{R}^{n \times n}$, $b : I \rightarrow \mathbb{R}^n$ sowie $x : I \rightarrow \mathbb{R}^n$. Allerdings ist es nützlich, auch komplexwertige Koeffizienten und Lösungen zuzulassen. Durch Aufspaltung in Real- und Imaginärteil ist ein komplexes $n \times n$ System äquivalent zu einem reellen $(2n) \times (2n)$ System, das heißt unsere Existenz- und Eindeigkeitstheorie ist voll anwendbar. Wir betrachten erst das sogenannte homogene Problem, das heißt den Fall $b(t) \equiv 0$.

Satz 28.1 (Anfangswertisomorphismus) Sei $A \in C^0(I, \mathbb{K}^{n \times n})$ für $\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$. Dann ist die Menge $L_A = \{x \in C^1(I, \mathbb{K}^n) : x' = Ax \text{ auf } I\}$ ein n -dimensionaler \mathbb{K} -Vektorraum, genauer ist für jedes $t_0 \in I$ die Abbildung

$$\delta_{t_0} : L_A \rightarrow \mathbb{K}^n, \quad \delta_{t_0}(x) = x(t_0)$$

ein Vektorraumisomorphismus.

BEWEIS: Mit $x, y \in L_A$ ist auch $\lambda x + \mu y \in L_A$ für $\lambda, \mu \in \mathbb{K}$, denn

$$(\lambda x + \mu y)' = \lambda x' + \mu y' = \lambda Ax + \mu Ay = A(\lambda x + \mu y).$$

Also ist L_A ein Untervektorraum von $C^1(I, \mathbb{K}^n)$, und die Abbildung $\delta_{t_0} : L_A \rightarrow \mathbb{K}^n$ ist wohldefiniert und linear. Da $|A(t)x| \leq |A(t)||x|$, gibt es nach Satz 27.4 zu jedem $x_0 \in \mathbb{K}^n$ eine Lösung von $x' = Ax$ auf ganz I mit Anfangswert $x(t_0) = x_0$, das heißt die Abbildung δ_{t_0} ist surjektiv. Aus der Eindeutigkeit der Lösung nach Satz 27.1 folgt, dass δ_{t_0} auch injektiv ist. \square

Eine Basis von L_A bezeichnet man als Fundamentalsystem der Gleichung $x' = Ax$. Explizit erhält man ein solches Fundamentalsystem, indem man die Gleichung zu linear unabhängigen Anfangsdaten löst, etwa $x_j(t_0) = e_j$, wobei e_1, \dots, e_n die Standardbasis ist.

Für beliebige n Lösungen $x_i : I \rightarrow \mathbb{K}^n$ der Differentialgleichung impliziert der Satz folgende Alternative: entweder bilden die Vektoren $x_1(t), \dots, x_n(t)$ eine Basis des \mathbb{K}^n für jedes $t \in I$, oder es gibt Zahlen $\lambda_1, \dots, \lambda_n \in \mathbb{K}$, nicht alle Null, mit $\lambda_1 x_1(t) + \dots + \lambda_n x_n(t) = 0$ für alle $t \in I$. Wir können n Funktionen $x_1, \dots, x_n : I \rightarrow \mathbb{K}^n$ zu einer matrixwertigen Funktion $X : I \rightarrow \mathbb{K}^{n \times n}$ zusammenfassen, so dass x_j die j -te Spalte von X ist. Es gilt dann

$$x'_j = Ax_j \text{ für } j = 1, \dots, n \quad \Leftrightarrow \quad X' = AX.$$

Dabei steht rechts die Matrixmultiplikation. Die Äquivalenz folgt aus der Tatsache, dass die Ableitung von X spaltenweise berechnet werden kann, und dass die j -te Spalte von AX gleich Ax_j ist. Die Tatsache, dass die Vektoren $x_1(t), \dots, x_n(t)$ für Lösungen des Systems $x' = Ax$ entweder für alle $t \in I$ oder für kein $t \in I$ linear abhängig sind, ergibt sich alternativ auch aus folgender Formel.

Satz 28.2 (Formel von Liouville) Sei $X \in C^1(I, \mathbb{K}^{n \times n})$ eine Lösung von $X' = AX$ auf I , wobei $A \in C^0(I, \mathbb{K}^{n \times n})$. Dann gilt für beliebiges $t_0 \in I$

$$\det X(t) = \det X(t_0) \cdot \exp\left(\int_{t_0}^t \operatorname{tr} A(s) ds\right) \quad \text{für alle } t \in I.$$

BEWEIS: Ist $\det X(t_0) = 0$, so gilt $\det X(t) = 0$ für alle $t \in I$ nach Satz 28.1, und die Formel trifft zu. Andernfalls ist $\det X(t) \neq 0$ für alle $t \in I$, und es gilt die allgemeine Formel

$$(\det X)'(t) = \det X(t) \cdot \operatorname{tr}(X(t)^{-1}X'(t)).$$

Setzen wir $X' = A(t)X(t)$ ein und beachten $\operatorname{tr}(X^{-1}AX) = \operatorname{tr}(XX^{-1}A) = \operatorname{tr}(A)$, so folgt

$$(\det X)'(t) = \det X(t) \cdot \operatorname{tr}(A(t)).$$

Damit gilt weiter

$$\frac{d}{dt} \left(e^{-\int_{t_0}^t \operatorname{tr} A(s) ds} \det X(t) \right) = 0,$$

und durch Integration folgt die Behauptung. \square

Für die konkrete Berechnung eines Fundamentalsystems gibt es keine allgemeinen Rezepte. Eine Ausnahme bilden die Gleichungen mit konstanten Koeffizienten, für die ein Fundamentalsystem durch Übergang zu einer Basis bestimmt werden kann, in der die Gleichungen entweder ganz entkoppeln oder auf sehr spezielle Weise gekoppelt sind, genauer ist die Koeffizientenmatrix entweder diagonal oder hat Jordansche Normalform. Dies ist (vielleicht) in der Linearen Algebra behandelt worden, darum will ich hier etwas anders vorgehen: die Matrix-Exponentialfunktion ist definiert durch

$$\exp : \mathbb{K}^{n \times n} \rightarrow \mathbb{K}^{n \times n}, \exp(A) = \sum_{k=0}^{\infty} \frac{A^k}{k!}.$$

Für $|A| \leq R$ gilt $|A^k| \leq |A|^k \leq R^k$. Wegen $\sum_{k=0}^{\infty} \frac{R^k}{k!} = e^R < \infty$ konvergiert die Reihe nach dem Majorantenkriterium, und die Konvergenz ist gleichmäßig für $|A| \leq R$. Insbesondere ist die \exp stetig auf $\mathbb{K}^{n \times n}$.

Satz 28.3 (Homogene Systeme mit konstanten Koeffizienten) Für $A \in \mathbb{K}^{n \times n}$ ist

$$X : \mathbb{R} \rightarrow \mathbb{K}^{n \times n}, X(t) = \exp(tA),$$

die Lösung des Anfangswertproblems $X' = AX$, $X(0) = E_n$. Die Spaltenvektoren $x_j(t)$, $j = 1, \dots, n$, bilden ein Fundamentalsystem für $x' = Ax$ zu den Anfangswerten $x_j(0) = e_j$.

BEWEIS: Für $|t| \leq T$ und $k \geq 1$ gilt

$$\left| \frac{d}{dt} \frac{(tA)^k}{k!} \right| \leq \left| \frac{t^{k-1}A^k}{(k-1)!} \right| \leq |A| \frac{(T|A|)^{k-1}}{(k-1)!}.$$

Wegen $\sum_{k=0}^{\infty} \frac{(T|A|)^k}{k!} = e^{T|A|} < \infty$ konvergiert die formal differenzierte Reihe absolut und gleichmäßig für $|t| \leq T$. Es folgt durch gliedweise Differentiation, vgl. Satz 3.2 in Kapitel 5,

$$X'(t) = \sum_{k=1}^{\infty} \frac{t^{k-1}A^k}{(k-1)!} = A \sum_{k=0}^{\infty} \frac{1}{k!} (tA)^k = AX(t).$$

Hier haben wir benutzt, dass die Multiplikation von links mit A eine stetige Abbildung auf $\mathbb{K}^{n \times n}$ ist. Weiter gilt $X(0) = A^0 = E_n$ per Definition, und somit ist $X(t)$ ein Fundamentalsystem nach Satz 28.1. \square

Folgerung 28.1 Die Matrix-Exponentialabbildung hat folgende Eigenschaften:

- (a) $\exp : \mathbb{K}^{n \times n} \rightarrow \text{Gl}_n(\mathbb{K})$ und $\exp(0) = E_n$.
- (b) $\det(\exp(A)) = \exp(\text{tr}(A))$.
- (c) $\exp(SAS^{-1}) = S \exp(A) S^{-1}$ für alle $A \in \mathbb{K}^{n \times n}$, $S \in \text{Gl}_n(\mathbb{K})$.
- (d) $\exp(A+B) = \exp(A)\exp(B)$, falls $[A, B] = AB - BA = 0$.

BEWEIS: Die Invertierbarkeit von $e^A = e^{tA}|_{t=1}$ folgt aus Satz 28.3 sowie Satz 28.1, und $e^0 = E_n$ gilt nach Definition. Gleichung (b) wurde bereits in Satz 28.2 gezeigt. Für (c) argumentieren wir mit dem Eindeutigkeitssatz: die Funktionen $t \mapsto e^{tSAS^{-1}}$ sowie $t \mapsto Se^{tA}S^{-1}$ lösen beide die Differentialgleichung $X' = (SAS^{-1})X$ zum Anfangswert $X(0) = E_n$ und sind damit gleich; für $t = 1$ folgt (c). Bei (d) gehen wir ähnlich vor, aber in zwei Schritten: zunächst sind sowohl $t \mapsto e^{tA}B$ als auch $t \mapsto Be^{tA}$ Lösungen des Anfangswertproblems $X' = AX$ mit $X(0) = B$, wobei für die zweite Funktion die Voraussetzung $[A, B] = 0$ eingeht:

$$\frac{d}{dt}(Be^{tA}) = BAe^{tA} = A(Be^{tA}).$$

Also gilt $e^A B = B e^A$. Jetzt berechnen wir für $X(t) = e^{tA} e^{tB}$

$$X'(t) = Ae^{tA}e^{tB} + e^{tA}Be^{tA} = (A+B)e^{tA}e^{tB} = (A+B)X(t).$$

Es folgt $X(t) = e^{t(A+B)}$, insbesondere $e^A e^B = e^{A+B}$. □

Die Eigenschaften (c) und (d) können alternativ auch aus der Definition der Matrix-Exponentialfunktion als Reihe hergeleitet werden. Für (d) ist dabei wesentlich, dass $e^A e^B$ mit dem Cauchyprodukt für Reihen berechnet werden kann, wenn $[A, B] = 0$. Nach (c) reicht es zur Berechnung von e^A den Fall zu behandeln, wenn A Jordansche Normalform hat. Wir wollen das nicht allgemein durchführen, auch wenn es nicht schwer ist, sondern uns auf folgendes Beispiel beschränken.

Die Auslenkung eines schwingungsfähigen Systems (Oszillators) mit Eigenfrequenz $\omega_0 \geq 0$ und Reibungskoeffizient $\beta \geq 0$ wird beschrieben durch die Differentialgleichung

$$x'' + 2\beta x' + \omega_0^2 x = 0.$$

Wir schreiben die Gleichung äquivalent in ein System erster Ordnung um:

$$\begin{pmatrix} x \\ y \end{pmatrix}' = \begin{pmatrix} 0 & 1 \\ -\omega_0^2 & -2\beta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}.$$

Das charakteristische Polynom der Koeffizientenmatrix A ist

$$p_A(\lambda) = \lambda^2 + 2\beta\lambda + \omega_0^2 = (\lambda + \beta)^2 + \omega_0^2 - \beta^2.$$

Für $\beta \neq \omega_0$ hat A zwei verschiedene, eventuell komplexe Eigenwerte $\lambda^\pm = -\beta \pm \sqrt{\beta^2 - \omega_0^2} \in \mathbb{C}$, mit zugehörigen Eigenvektoren $v^\pm = (1, \lambda^\pm)$. Es gilt dann

$$\frac{d}{dt} e^{\lambda^\pm t} v^\pm = \lambda^\pm e^{\lambda^\pm t} v^\pm A(e^{\lambda^\pm t} v^\pm).$$

Die Vektoren $e^{\lambda^{\pm}t}v^{\pm}|_{t=0} = v^{\pm}$ sind als Eigenvektoren zu v verschiedenen Eigenvektoren linear unabhängig, also bilden die Funktionen $e^{\lambda^{\pm}t}v^{\pm}$ ein Fundamentalsystem über \mathbb{C} . Ist $\beta > \omega_0$, so folgt $\lambda^{\pm} \in [-2\beta, 0)$ und die resultierenden Lösungen $x^{\pm}(t) = e^{\lambda^{\pm}t}$ fallen streng monoton und exponentiell (Kriechfall). Ist $\beta < \omega_0$, so erhalten wir durch Linearkombination das reelle Fundamentalsystem

$$\begin{aligned} x_1(t) &= e^{-\beta t} \cos \sqrt{\omega_0^2 - \beta^2} t & x_2(t) &= e^{-\beta t} \sin \sqrt{\omega_0^2 - \beta^2} t \\ y_1(t) &= x_1'(t) & y_2(t) &= x_2'(t) \end{aligned}$$

Fall 3 $0 = \beta^2 - \omega_0^2$

In diesem Fall hat das Polynom die eine reelle Nullstelle $-\beta$. Ein Fundamentalsystem lautet

$$\begin{aligned} x_1(t) &= e^{-\beta t} & x_2(t) &= t e^{-\beta t} \\ y_1(t) &= -e^{-\beta t} & y_2(t) &= e^{-\beta t} - \beta t e^{-\beta t} \end{aligned}$$

An der Stelle $t = 0$

$$\begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}.$$

29 Separation der Variablen

Als Ausgangspunkt betrachten wir nochmals die Euler-Lagrange-Gleichungen.

Satz 29.1 (Energieerhaltungssatz) Sei $\Omega \subset \mathbb{R}^n$ offen und $f \in C^2(\Omega \times \mathbb{R}^n)$, $f = f(x, v)$ (also f unabhängig von t). Ist $x \in C^2(I, \Omega)$ eine Lösung der Euler-Lagrange-Gleichungen, so gilt der „Energieerhaltungssatz“

$$\frac{d}{dt}[\langle D_v f(x, x'), x' \rangle - f(x, x')] = 0.$$

BEWEIS:

$$\frac{d}{dt} \left[\sum_{j=1}^n \frac{\partial f}{\partial v_j}(x, x') x'_j - f(x, x') \right] = \sum_{j=1}^n \left(\underbrace{\frac{d}{dt} \frac{\partial f}{\partial v_j}(x, x') - \frac{\partial f}{\partial x_j}(x, x')}_{=0} \right) x'_j = 0$$

□

Beispiel 29.1 Sei $f(x, v) = \frac{m}{2}|v|^2 - V(x)$ die Lagrangefunktion eines Teilchens der Masse $m > 0$ im Potential V . Dann gilt

$$\begin{aligned} \langle D_v f(x, x'), x' \rangle - f(x, x') &= \langle mx', x' \rangle - \left(\frac{m}{2} |x'|^2 - V(x) \right) \\ &= \frac{m}{2} |x'|^2 + V(x). \end{aligned}$$

Es folgt der Energieerhaltungssatz

$$\frac{m}{2} |x'|^2 + V(x) = E \quad (= \text{Konstante}).$$

Bei Systemen mit einem Freiheitsgrad ($n = 1$) lautet der Energieerhaltungssatz

$$x' = \pm \sqrt{\frac{2}{m}(E - V(x))}.$$

Anfangswertprobleme dieses Typs können nun „explizit“ gelöst werden mit dem nachstehendem Verfahren.

Satz 29.2 (Separation der Variablen) Seien $f \in C^0(I)$, $g \in C^0(J)$ für Intervalle $I, J \subset \mathbb{R}$, und $g(x) \neq 0$ für $x \in J$. Betrachte das Anfangswertproblem

$$(29.1) \quad x'(t) = \frac{f(t)}{g(x(t))} \quad \text{für } t \in I, \quad x(t_0) = x_0.$$

Definiere $F \in C^1(I)$, $G \in C^1(J)$ durch

$$F(t) = \int_{t_0}^t f(s) ds, \quad G(x) = \int_{x_0}^x g(y) dy.$$

Ist I so klein gewählt dass $F(I) \subset G(J)$, so folgt:

a) Es gibt genau ein $x \in C^1(I)$ mit $G(x(t)) = F(t)$ für alle $t \in I$.

b) Diese Funktion ist die eindeutig bestimmte Lösung des Anfangswertproblems (29.1).

BEWEIS: Nach Analysis 1 ist $J^* = G(J)$ ein offenes Intervall und es existiert die Umkehrfunktion $H \in C^1(J^*)$ von G . Für $x \in C^1(I)$ mit $x(I) \subset J$ gilt:

$$\begin{aligned} & x \text{ ist Lösung von (29.1)} \\ \Leftrightarrow & \frac{d}{dt} G(x(t)) = \frac{d}{dt} F(t) \quad \text{und } x(t_0) = x_0 \\ \Leftrightarrow & G(x(t)) = F(t) \quad (\text{beachte } G(x_0) = F(t_0) = 0) \\ \Leftrightarrow & x(t) = H(F(t)). \end{aligned}$$

Insbesondere ist eine Lösung von (29.1) eindeutig bestimmt. Definieren wir $x = H \circ F$, so ist $x \in C^1(I)$ mit $x(I) = H(F(I)) \subset H(G(J)) = J$ nach Voraussetzung und x ist Lösung von (29.1). \square

Lösungsrezept.

- $\frac{dx}{dt} = \frac{f(t)}{g(x)} \longrightarrow g(x) dx = f(t) dt$
- Integration von t_0 bis t und x_0 bis x :

$$G(x) = \int_{x_0}^x g(y) dy = \int_{t_0}^t f(s) ds = F(t)$$

- Auflösung nach x ergibt $x = x(t)$ oder Auflösung nach t ergibt $t = t(x)$.

Beispiel 29.2 (Homogene, lineare Gleichung) Betrachte für $a \in C^0(I)$ das Anfangswertproblem

$$x'(t) = a(t)x(t), \quad x(t_0) = x_0.$$

Nach dem Lösungsrezept ergibt sich (unter welchen Voraussetzungen?)

- $\frac{dx}{x} = a(t) dt$
- Integration von t_0 bis t bzw. x_0 bis x :

$$\log \frac{x}{x_0} = \int_{x_0}^x \frac{dy}{y} = \int_{t_0}^t a(s) ds$$

- $x(t) = x_0 \exp \int_{t_0}^t a(s) ds$.

Beispiel 29.3 (Rotationsminimalflächen) Wird der Graph einer Funktion $r : I \rightarrow (0, \infty)$, $r = r(x)$, um die x -Achse rotiert, so hat die entstehende Fläche (wie man zeigen kann) den Flächeninhalt

$$A(r) = 2\pi \int_I r(x) \sqrt{1 + r'(x)^2} dx.$$

Falls r diesen Flächeninhalt relativ zu allen Graphen $u : I \rightarrow (0, \infty)$ mit $u|_{\partial I} = r|_{\partial I}$ minimiert, so ist r Lösung der zugehörigen Euler-Lagrange-Gleichung. Aus Satz 29.1 folgt

$$\begin{aligned} \frac{r r'}{\sqrt{1 + (r')^2}} r' - r \sqrt{1 + (r')^2} &= -a \quad (= \text{Konst.}) \\ \Leftrightarrow (r')^2 &= \left(\frac{r}{a}\right)^2 - 1. \end{aligned}$$

Wir lösen mit Separation unter der Annahme $r' > 0$:

$$\frac{dr}{dt} = \sqrt{\left(\frac{r}{a}\right)^2 - 1} \quad \rightarrow \quad \left(\left(\frac{r}{a}\right)^2 - 1\right)^{-1/2} dr = dx.$$

Integration von r_0 bis r , x_0 bis x liefert $\text{Arcosh } \frac{r}{a} - \text{Arcosh } \frac{r_0}{a} = x - x_0$ beziehungsweise

$$r(x) = a \cosh \frac{x - x_1}{a} \quad \text{mit } x_1 = x_0 - \text{Arcosh } \frac{r_0}{a}.$$

30 Lineare Differentialgleichungen

Für lineare Differentialgleichungen erhält man eine sehr präzise Theorie. Wir betrachten hier allgemein ein System

$$(30.1) \quad x'(t) = A(t)x(t) + b(t), \quad x(t_0) = x_0.$$

Dabei sind $A : I \rightarrow \mathbb{R}^{n \times n} \simeq L(\mathbb{R}^n, \mathbb{R}^n)$ und $b \in I \rightarrow \mathbb{R}^n$ stetig und $x_0 \in \mathbb{R}^n$. Gesucht ist eine Lösung $x \in C^1(I, \mathbb{R}^n)$ von (30.1). Es erweist sich aber als nützlich, im Folgenden auch komplexwertige Funktionen zuzulassen, also

$$A(t) \in \mathbb{C}^{n \times n}, \quad b(t) \in \mathbb{C}^n, \quad x_0 \in \mathbb{C}^n \quad \text{und} \quad x(t) \in \mathbb{C}^n.$$

Tatsache. Das lineare Anfangwertproblem (30.1) ist auf ganz I (und nicht nur lokal) lösbar. Dies ergibt sich durch eine genauere Analyse des Satzes von Picard-Lindelöf, siehe Skript Analysis II, SS '97.

Satz 30.1 Sei $A \in C^0(I, \mathbb{K}^{n \times n})$ mit $\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$. Dann ist die Lösungsmenge des homogenen Systems

$$L_A = \{x \in C^1(I, \mathbb{K}^n) : x' = Ax \text{ auf } I\}$$

ein n -dimensionaler \mathbb{K} -Vektorraum. Und zwar ist für jedes $t_0 \in I$ die Abbildung

$$\delta_{t_0} : L_A \rightarrow \mathbb{K}^n, \quad \delta_{t_0}(x) = x(t_0)$$

ein Vektorraumisomorphismus.

BEWEIS: Mit $x, y \in L_A$ ist auch $\lambda x + \mu y \in L_A$ für $\lambda, \mu \in \mathbb{R}$. Deshalb ist L_A ein Untervektorraum von $C^1(I, \mathbb{K}^n)$. Die Abbildung δ_{t_0} ist surjektiv, weil das Anfangwertproblem zu jedem $x_0 \in \mathbb{K}^n$ eine Lösung auf ganz I hat, siehe obige Tatsache. δ_{t_0} ist injektiv, weil die Lösung des Anfangwertproblems nach Satz 29.1 eindeutig bestimmt ist. \square

Definition 30.1 Eine Basis von L_A heißt (Lösungs-)Fundamentalsystem auf I .

Als Beispiel betrachten wir die Differentialgleichung zweiter Ordnung (Schwingung mit Reibung)

$$x'' + 2\beta x' + \omega_0^2 x = 0 \quad (\beta, \omega_0 \geq 0).$$

Wir können die Gleichung äquivalent in ein System erster Ordnung umschreiben:

$$\begin{aligned} x' &= y \\ y' &= -\omega_0^2 x - 2\beta y \end{aligned}$$

Ein Fundamentalsystem ergibt sich aus dem Ansatz $x(t) = e^{\lambda t}$ mit $\lambda \in \mathbb{C}$, also

$$0 = e^{\lambda t}(\lambda^2 + 2\beta\lambda + \omega_0^2).$$

Fall 1 $0 < \beta^2 - \omega_0^2 =: \omega^2$ (wobei $\omega > 0$)

Dann hat das Polynom die beiden reellen Nullstellen $\lambda_{\pm} = -\beta \pm \omega < 0$. Die beiden Lösungen

$$\begin{pmatrix} x_{\pm}(t) \\ y_{\pm}(t) \end{pmatrix} = \begin{pmatrix} e^{\lambda_{\pm} t} \\ \lambda_{\pm} e^{\lambda_{\pm} t} \end{pmatrix}$$

bilden ein Fundamentalsystem, denn

$$\det \begin{pmatrix} x_+(0) & x_-(0) \\ y_+(0) & y_-(0) \end{pmatrix} = \det \begin{pmatrix} 1 & 1 \\ \lambda_+ & \lambda_- \end{pmatrix} = \lambda_- - \lambda_+ = -2\omega < 0.$$

Fall 2 $0 > \beta^2 - \omega_0^2 =: (i\omega)^2$ (wobei $\omega > 0$). Dann hat das Polynom die beiden nichtreellen Nullstellen $\lambda_{\pm} = -\beta \pm i\omega$. Analog zu oben erhalten wir nun ein komplexes Fundamentalsystem

$$\begin{pmatrix} x_{\pm}(t) \\ y_{\pm}(t) \end{pmatrix} = \begin{pmatrix} e^{\lambda_{\pm}t} \\ \lambda_{\pm} e^{\lambda_{\pm}t} \end{pmatrix}.$$

Über \mathbb{R} erhalten wir durch Linearkombination

$$\begin{aligned} x_1(t) &= e^{-\beta t} \cos \omega t & x_2(t) &= e^{-\beta t} \sin \omega t \\ y_1(t) &= x_1'(t) & y_2(t) &= x_2'(t) \end{aligned}$$

Fall 3 $0 = \beta^2 - \omega_0^2$

In diesem Fall hat das Polynom die eine reelle Nullstelle $-\beta$. Ein Fundamentalsystem lautet

$$\begin{aligned} x_1(t) &= e^{-\beta t} & x_2(t) &= t e^{-\beta t} \\ y_1(t) &= -e^{-\beta t} & y_2(t) &= e^{-\beta t} - \beta t e^{-\beta t} \end{aligned}$$

An der Stelle $t = 0$

$$\begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}.$$