Analysis II für Physiker

Prof. Dr. Andreas Kollross

Florian Rappl

14. März 2009

Inhaltsverzeichnis

1 Differentialrechnung im $\mathbb{R}^n$ **2**

1.1 Der $\mathbb{R}^n$ als metrischer Raum 2
1.2 Kompaktheit 7
1.3 Partielle Ableitungen 10
1.4 Totale Differenzierbarkeit 17
1.5 Taylorformel und lokale Extrema 26
1.6 Banachscher Fixpunktsatz, implitzite Funktionen und Umkehrsatz 34
1.7 Parameterabhängige Integrale und Variationsrechnung 44

2 Gewöhnliche Differentialgleichungen **56**

2.8 Einführung 56
2.9 Existenz und Eindeutigkeitssatz 57
2.10 Elementare Lösungsmethoden 62
2.11 Lineare Differentialgleichungen 68

Kapitel 1

Differentialrechnung im $\mathbb{R}^n$

1.1 Der $\mathbb{R}^n$ als metrischer Raum

Stetigkeit und Konvergenz Unser Ziel im ersten Teil der Vorlesung ist es, Methoden der Differentialrechnung für Abbildung $\mathbb{R}^m \to \mathbb{R}^n$ zu entwickeln. Dazu benötigen wir eine Reihe von topologischen Begriffen wie:
Stetigkeit, Konvergenz, Abstand, Norm, Offene / Abgeschlossene Mengen, Kompaktheit, Umgebungen, die wir im folgenden wiederholen wollen.
Auf dem reelen Vektorraum, $\mathbb{R}^n$, der n-Tupel von reellen Zahlen ist ein euklidisches Skalarprodukt gegeben durch

$$\langle x, y\rangle := x_1y_1 + x_2y_2 + ... + x_ny_n$$

Für $x = (x_1, ..., x_n) \in \mathbb{R}^n, y = (y_1, ..., y_n) \in \mathbb{R}^n$. Wir nennen dies das kanonische Skalarprodukt. Setzt man $\|x\| := \sqrt{\langle x, x\rangle}$, so erhält man die euklidische Norm auf $\mathbb{R}^n$.
Allgemeiner heißt die Abbildung $V \to \mathbb{R}, x \mapsto \|x\|$ auf einem reellen Vektorraum eine Norm, falls folgende drei Axiome gelten:

1. $\|x\| = 0 \Leftrightarrow x = 0$
2. $\|\lambda x\| = |\lambda| \cdot \|x\|$ für alle $\lambda \in \mathbb{R}, x \in V$
3. $\|x + y\| \leq \|x\| + \|y\|$

Auf dem $\mathbb{R}^n$ kann man auch weitere Normen definieren, z.B.:
$\|x\|_1 := |x_1|+|x_2|+...+|x_n|$ oder $\|x\|_\infty := max\{|x_1|, ...|x_n|\}$ Auf jeden normierten Vektorraum (= reeller Vektorraum mit Norm) kann man eine Abstandsfunktion (oder Metrik) definieren durch $d(x, y) := \|x-y\|$. Falls nichts anderes gesagt wird, verwenden wir im folgenden stetig die aus den euklidischen Norm abgeleitete euklidische Abstandsfunktion:

$$d(x, y) := \sqrt{(x_1 - y_1)^2 + ... + (x_n - y_n)^2}$$

Der $\mathbb{R}^n$ wird damit zu einem metrischen Raum.

Erinnerung Eine Menge X mit einer Abbildung $d : X \times X \to \mathbb{R}$ heißt metrischer Raum, falls gilt:

1. $d(x, y) = 0 \Leftrightarrow x = y$
2. $d(x, y) = d(y, x)$ für alle $x, y \in X$
3. $d(x, y) \leq d(x, z) + d(z, y)$ für alle $x, y, z \in X$

Ist die Abstandsfunktion wie oben aus einer Norm abgeleitet, so folgen diese Eigenschaften direkt aus denen der Norm.

Bemerkung Eine Teilmenge $M \subseteq X$ eines metrischen Raumes ist, mit der Abstandsfunktion $d_{|M\times M}$, ebenfalls ein metrischer Raum. Insbesondere ist jede Teilmenge des $\mathbb{R}^n$ mit der Einschränkung der euklidischen Abstandsfunktion wieder ein metrischer Raum. Viele topologische Begriffe können auf den Begriff des Abstandes zurückgeführt werden, d.h. wir können sie für metrisch definieren, z.B. Konvergenz von Folgen, Cauchyfolgen, Stetigkeit.

Satz 1.1.1 *Sei $x_\nu = (x_{1\nu}, ..., x_{n\nu}), \nu \in \mathbb{N}$, eine Folge im $\mathbb{R}^n$. Dann sind äquivalent:*

1. *Die Folge x_ν konvergiert.*
2. *Jede Komponentfolge $x_{i\nu}, i \in \{1, ..., n\}, \nu \in \mathbb{N}$ konvergiert.*
3. *Die Folge x_ν ist eine Cauchyfolge.*

Beweis (3) $\Rightarrow$ (1):
Sei x_ν eine Cauchyfolge. Wegen $|x_{i\nu} - x_{i\mu}| \leq d(x_\nu, x_\mu) = \sqrt{(x_{1\nu} - x_{1\mu})^2 + ... + (x_{n\nu} - x_{n\mu})^2}$ ist jedes $x_{i\nu}$ eine Cauchyfolge, also konvergent.
(2) $\Rightarrow$ (1):
Angenommen, jede Komponentfolge $x_{i\nu}$ konvergiert. Sei $\varepsilon > 0$ vorgegeben. Für jedes $i \in \{1, ..., n\}$ gibt es ein $k_i \in \mathbb{N}$, so dass $|x_{i\nu} - x_i| < \frac{\varepsilon^2}{n}$ für $\nu \geq k_i$, wobei $x_i := \lim_{\nu\to\infty} x_{i\nu}$ gilt. Dann gilt für $\nu \geq max\{k_1, ..., k_n\}$, dass $d((x_{1\nu}, ..., x_{n\nu}), (x_1, ..., x_n)) < \sqrt{n\frac{\varepsilon^2}{n}} = \varepsilon$, was $x_n \to (x_1, ..., x_n)$ zeigt.
(1) $\Rightarrow$ (3):
Ist klar.
Der Satz zeigt, dass $\mathbb{R}^n$ ein vollständiger metrischer Raum bzw. ein vollständiger normierter Vektorraum (oder auch Banachraum genannt) ist. $\square$

Korollar 1.1.2 *Eine Folge* $(x_{1\nu}, ..., x_{n\nu})$ *im* $\mathbb{R}^n$ *konvergiert genau dann gegen einen Punkt* $(x_1, ..., x_n) \in \mathbb{R}^n$*, wenn für* $i = 1, ..., n$ *die reelle Zahlenfolge* $x_{i\nu}$ *gegen* x_i *konvergiert.*

Beweis „$\Leftarrow$" haben wir im Beweis Satz 1 gezeigt. „$\Rightarrow$" folgt daraus, dass Grenzwerte in einem metrischen Raum eindeutig sind. $\square$

Erinnerung Seien X, Y metrische Räume und $f : X \to Y$ eine Abbildung. Die Abbildung f heißt an der Stelle $x \in X$ (folgen)stetig, wenn für jede Folge x_ν in X mit $x_\nu \to X$ auch $f(x_\nu) \to f(x)$ gilt.
Stetigkeit und Konvergenz sind eigentlich verwandte Konzepte. Beides sind topologische Begriffe, d.h. sie lassen sich (auch offene Metrik) anhand eines Systems von offenen Mengen untersuchen.
Eine Teilmenge $U \subseteq X$ eines metrischen Raumes X heißt offen, wenn es für jeden Punkt $x \in U$ ein $\varepsilon > 0$ gibt, so dass der Ball $B_\varepsilon(x) := \{p \in X | d(x,p) < \varepsilon\} \subseteq X$ ganz in U enthalten ist.

Beispiele für offene Mengen Offene Intervalle, X, $\emptyset$, Bälle $B_R(p)$, endliche Schnitte, beliebige Vereinigungen von offenen Mengen.
Eine Teilmenge $M \subseteq X$ eines metrischen Raums heißt Umgebung eines Punktes $x \in X$, wenn es eine offene Teilmenge U von X mit $x \in U \subseteq M$ gibt.

Lemma 1.1.3 *Eine Folge* x_n *in einem metrischen Raum* X *konvergiert genau dann gegen* $x \in X$*, wenn zu jeder Umgebung* U *von* X *ein* $n_0 \in \mathbb{N}$ *existiert, so dass* $x_\nu \in U$ *für alle* $\nu \geq n_0$ *gilt.*

Beweis Lemma 4.30 im Analysis I Skript von Prof. Bunke.

Lemma 1.1.4 *Seien* X, Y *metrische Räume und* $f : X \to Y$ *eine Abbildung. Dann sind äquivalent:*

1. f *ist an der Stelle* $x \in X$ *(folgen)stetig*

2. *Für jede Umgebung* U *von* $f(x)$ *in* Y *ist das Urbild* $f^{-1}(U)$ *eine Umgebung von* x *in* X *(„topologisch" stetig)*

3. *Für jedes* $\varepsilon > 0$ *gibt es ein* $\delta \geq 0$*, so dass für alle* $x' \in X$ *gilt:* $d(x, x') < \delta \Rightarrow d(f(x), f(x')) < \varepsilon$

4. *Für jede Umgebung* U *von* $f(x)$ *in* Y *gibt es ein* $\delta > 0$*, so dass für alle* $x' \in X$ *gilt* $d(x, x') < \delta \Rightarrow f(x') \in U$

Beweis Vergleiche Lemma 4.5 im Analysis I Skript von Prof. Bunke.
Eine Abbildung $f : X \to Y$ zwischen metrischen Räumen heißt stetig, wenn sie an jeder Stelle $x \in X$ stetig ist.

Lemma 1.1.5 *Eine Abbildung $f : X \to Y$ zwischen metrischen Räumen ist genau dann stetig, wenn Urbilder offener Mengen offen sind.*

Beweis Aus Lemma 4.52 im Analysis I Skript (gilt für topologische Räume), folgt eben auch fast sofort aus 1.1.4 2. Dies liefert eine bequeme Möglichkeit nachzuweisen, das gewisse Teilmengen des $\mathbb{R}^n$ offen sind.
Ist $f : \mathbb{R}^n \to \mathbb{R}$ eine stetige Funktion und $c \in \mathbb{R}$, so ist etwa $f^{-1}((-\infty, c)) = \{x \in \mathbb{R}^n | f(x) < c\}$ offen.

Beispiele $\{(x_1, ..., x_n) \in \mathbb{R}^n | x_1 < 0\}$ offen
$\{(x_1, ..., x_8) \in \mathbb{R}^n | x_1^2 + x_8^2 = 5x_7\} = f^{-1}(\{0\})$ abgeschlossen
$B_\varepsilon(x) = \{p \in \mathbb{R}^n | d(x, p) < \varepsilon\} = f^{-1}((-\infty, \varepsilon)), f(p) = d(x, p)$

Satz 1.1.6 *Sei X ein metrischer Raum und seien $f : X \to \mathbb{R}$ für $i : 1, ..., n$ Abbildungen. Die Abbildung $f : X \to \mathbb{R}^n, f(x) := (f_1(x), ..., f_n(x))$ ist genau dann stetig, wenn alle Komponentenfunktionen stetig sind.*

Beweis Sei $x \in X$. Wir zeigen: $f(x)$ ist genau dann an der Stelle $x \in X$ stetig, wenn alle f_i an der Stelle $x \in X$ stetig sind. Sei $x_\nu \to x$ eine konvergente Folge in X. Aus Korollar 1.1.2 folgt, dass $f(x_\nu)$ genau dann gegen $f(x)$ konvergiert, wenn alle $f_i(x_\nu)$ gegen $f_i(x)$ konvergieren. □

Beispiel 1.1.7 *Stetige Abbildungen*

1. *In Analysis I wurde gezeigt, dass die Abbildungen* $add : \mathbb{R} \times \mathbb{R} \to \mathbb{R}, (x, y) \mapsto x + y; mult : \mathbb{R} \times \mathbb{R} \to \mathbb{R}, (x, y) \mapsto xy; div : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$
   $\{0\}, (x, y) \mapsto \frac{x}{y}$; *stetig sind.*

2. *Verkettungen stetiger Abbildungen sind stetig (Lemma 4.54 im Analysis I Skript).*

3. *Sind $f, g : X \to \mathbb{R}$ auf dem metrischen Raum stetige Funktionen, so sind auch die Summe und das Produkt*
   $f + g : X \to \mathbb{R}, x \mapsto f(x) + g(x),$
   $fg : X \to \mathbb{R}, x \mapsto f(x)g(x)$
   stetig. Gilt außerdem $g(x) \neq 0$ für alle $x \in X$ so ist auch
   $\frac{f}{g} : X \to \mathbb{R}, x \mapsto \frac{f(x)}{g(x)}$
   stetig.

Beweis *Nach Satz 1.1.6 ist die Abbildung* $X \to \mathbb{R}^2, x \mapsto (f(x), g(x))$ *stetig. Verkettung dieser Abbildung mit den Funktionen add, mult, div in 1 liefert die obigen Funktionen* $f+g, fg, \frac{f}{g}$*. Nach 2 sind diese stetig.* $\square$

4. *Die Koordinatenfunktionen* $(x_1, ..., x_n) \mapsto x_i, \mathbb{R}^n \to \mathbb{R}$ *sind stetig / folgt aus (1.6), da* $id_{\mathbb{R}} : x \mapsto x$*.*
5. *Konstante Abbildungen* $f(x) = c$ *für alle* $x \in X$ *sind stetig.*
6. *Ein Monom auf* $\mathbb{R}^n$ *vom Grad* r *ist eine Funktion* $\mathbb{R}^n \to \mathbb{R}$ *der Gestalt* $(x_1, ..., x_n) \mapsto x_1^{k_1}...x_n^{k_n}$*, wobei* $k_1, ..., k_n \in \mathbb{N}_0$ *natürliche Zahlen mit* $k_1 + ... + k_n = r$ *sind. Eine Polynomfunktion* $F : \mathbb{R}^n \to \mathbb{R}$ *vom Grad* $\leq r$ *ist eine Linearkombination von Monomen vom Grad* $\leq r$*.* $F(x_1, ..., x_n) = \sum_{k_1+...+k_n \leq r} c_{k_1...k_n} x_1^{k_1}...x_n^{k_n}$*, wobei* $c_{k_1...k_n} \in \mathbb{R}$*. Stetigkeit folgt aus 3, 4, 5.*
7. *Inbesondere sind lineare Abbildungen* $\mathbb{R}^n \to \mathbb{R}^m$ *stetig. Denn die Komponentenfunktionen sind Polynome vom Grad* 1.
   $(f_i(x_1, ..., x_n) = \sum_{j=1}^n a_{ij} x_j$*,* f *durch die mxn-Matrix* (aij) *gegeben)*

Aber Allgemeiner gilt: Seien nun V und W normierte reelle Vektorräume und $A : V \to W$ eine lineare Abbildung.

Satz 1.1.8 *A ist genau dann stetig, wenn es beschränkt ist, das heißt wenn es ein* $C \in \mathbb{R}_+$ *gibt, so dass* $\|A(x)\| \leq C\|x\|$ *für alle* $x \in V$*.*

Beweis *„⇒" Sei A stetig. Dann ist A insbesondere im Nullpunkt stetig und es gibt zu*$\varepsilon = 1$ *ein* $\delta > 0$*, so dass* $\|A(z)\| < 1$ *für ale* $z \in V$ *mit* $\|z\| < \delta$ *gilt. Setze* $C = \frac{2}{\delta}$*. Sei* $x \in V$
$\{0\}$ *beliebig,* $z := \frac{x}{C\|x\|}$*. Dann gilt* $\|z\| = \frac{\delta\|x\|}{2\|x\|} < \delta$*, also* $\|A(z)\| < 1$*. Nun gilt:* $A(z) = \frac{A(x)}{C\|x\|}$*.*
Also folgt $A(x) < C\|x\|$*.*
„⇐" Es gebe $C \geq 0$ *mit* $\|A(x)\| \leq C\|x\|$ *für alle* $x \in V$*. Dann gilt* $\|A(x) - A(x')\| = \|A(x-x')\| \leq C\|x-x'\|$ *was die Stetigkeit mit Hilfe des* $\varepsilon-\delta$*-Kriteriums 1.1.4 3 zeigt.*

Beispiele

1. Sei $C([a;b])$ der reelle Vektorraum der stetigen Funktionen. $f : [a;b] \to \mathbb{R}$. Durch $\|f\| := \sup\{|f(x)||x \in [a;b]\}$ ist auf $C([a;b])$ eine Norm gegeben: Die Supremumsnorm. Vergleiche Analysis I Skript 4.4 (Beispiel 9).
   Nun betrachten wir die lineare Abbildung: $I : C([a;b]) \to \mathbb{R}, I(f) := \int_a^b f(x)dx$

Behauptung I ist stetig.

Beweis Es gilt für alle $f \in ([a;b])$ die Abschätzung $|I(f)| \leq (b-a)\|f\|$. Also ist I nach Satz 1.1.8 stetig. □

2. Sei $C^1([0;1]) \subset C([0;1])$ der Untervektorraum aller stetigen differenzierbaren Funktionen, ebenfalls mit der Supremumsnorm. Sei $D : C^1([0;1]) \to C([0;1])$ die durch die Differentation $D(f) := f'$ gegebene Abbildung.

   Behauptung D ist nicht stetig.

   Beweis Betrachte die Funktionen $f_\nu \in C^1([0;1])$ für $\nu \in \mathbb{N}$ mit $f(x) = x^\nu$. Es gilt $\|f_\nu = 1$ und $\|D(f_\nu)\| = \|\nu x^(\nu - 1)\| = \nu$. Es gibt keine Konstante $C \in \mathbb{R}_+$ mit $\|D(f_\nu)\| \leq C\|f_\nu\|$, also ist D nach Satz (1.8) unstetig. □

1.2 Kompaktheit

Definition 1.2.1 *Sei A eine Teilmenge eines metrischen Raumes X (z.B. $\mathbb{R}^n$). Unter einer offenen Überdeckung von A versteht man eine Familie $(U_i)_{i\in I}$ von offenen Teilmengen $U_i \subseteq X$, so dass $A \subseteq \bigcup_{i\in I} U_i$. Die Indexmenge kann dabei endlich oder unendlich sein.*

Definition 1.2.2 *Eine Teilmenge A eines metrischen Raumes X heißt kompakt, wenn jede offene Überdeckung von A eine endliche Teilüberdeckung besitzt. Genauer: A heißt kompakt, wenn es zu jeder offenen Überdeckung $(U_i)_{i\in I}$ von A endlich viele Indizes $i_1, ..., i_k$ gibt, so dass $A \subseteq U_{i_1} \cap ... \cap U_{i_k}$ gilt.*

Beachte Die Definition sagt nicht: A ist kompakt, wenn es eine endliche Überdeckung mit offenen Teilmengen gibt. (Die gibt es nämlich immer, z.B. X ist offen!)

Satz 1.2.3 *Sei X ein metrischer Raum, a_ν eine konvergente Folge und a der Grenzwert. Dann ist $A := \{a_\nu | \nu \in \mathbb{N}\} \cup \{a\}$ kompakt.*

Beweis Sei $(U_i)_{i\in I}$ eine beliebige Offene Überdeckung von A. Dann gibt es ein $i* \in I$, so dass $a \in U_{i*}$. Dieses U_{i*} ist eine Umgebung von a und da a_ν gegen a konvergiert, gibt es nach Lemma (1.3) eine Schranke $N \in \mathbb{N}$, so dass alle a_ν mit $\nu > N$ in U_{i*} liegen.
Außerdem liegt jedes $a_k, 1 \leq k < N$ in einem U_{i_k} und wir haben somit die endliche Überdeckung $U_{i_1} \cup ... \cup U_{i_N} \cup U_{1*}$ gefunden.

Lässt man den Grenzwert der Folge weg, gilt der Satz im Allgemeinen nicht mehr. Setze z.B.:

$$A = \{\tfrac{1}{1+n} | n \in \mathbb{N}\} = \{\tfrac{1}{2}, \tfrac{1}{3}, \tfrac{1}{4}, ...\}$$

und definiert eine offene Überdeckung von A durch die Vereinigung von offenen Intervallen.

$$(\tfrac{1}{3}; 1) \cup (\tfrac{1}{4}; \tfrac{1}{2}) \cup (\tfrac{1}{5}; \tfrac{1}{3})...$$

Jedes Intervall enthält genau einen Punkt von A und es gibt keine (echte) Teilüberdeckung, also auch keine endliche. □

Einige wichtige Sätze über Kompaktheit in metrischen Räumen und im $\mathbb{R}^n$:

Satz 1.2.4 *Ein metrischer Raum ist genau dann kompakt, wenn er Folgenkompakt ist, d.h. wenn jede Folge eine in diesem Raum konvergente Teilfolge besitzt.*

Beweis *Lemma 4.43 im Analysis I Skript von Prof. Bunke.*

Eine Teilmenge $A \subset \mathbb{R}^n$ ist somit genau dann kompakt, wenn jede Folge (a_ν) mit $a_\nu \in A$ eine Teilfolge a_{ν_k} besitzt, so dass $a_{\nu_k} \longrightarrow a \in A$.
Kompakte metrische Räume sind insbesondere vollständig (denn eine Cauchyfolge mit konvergenter Teilfolge geht gegen den Limes der Teilfolge).
Mit Hilfe von (1.2.4) wurde in Analysis I die folgende wichtige Charakterisierung von kompakten Mengen in $\mathbb{R}^n$ bewiesen:

Satz 1.2.5 (Heine-Borel) *Eine Teilmenge des $\mathbb{R}^n$ mit der euklidischen Metrik ist genau dann kompakt, wenn sie abgeschlossen und beschränkt ist.*

Beweis *Lemma 4.45 im Analysis I Skript von Prof. Bunke.*

Erinnerung Eine Teilmenge eines metrischen Raumes heißt beschränkt, wenn sie in einem Ball enthalten ist. Eine Teilmenge M des $\mathbb{R}^n$ ist genau dann beschränkt, wenn es ein $R > 0$ gibt, so dass $\|x\| \leq R$ für alle $x \in M$. In metrischen Räumen sind Abgeschlossenheit und Folgenabgeschlossenheit äquivalent.

Beispiel 1.2.6 (Kompakte Mengen)

1. *Endliche Teilmengen von metrischen (topologisch Hausdorffschen-) Räumen sind kompakt.*

2. *Die abgeschlossene Kugel im $\mathbb{R}^n$*

$$K_r(x) := B_r(x) = \{p \in \mathbb{R}^n | \|p - x\| \leq r\}$$

und Sphäre

$$S_r^{n-1}(x) := \{p \in \mathbb{R}^n | \|p - x\| = r\}$$

sind kompakt. Denn sie sind offensichtlich beschränkt und Urbilder der abgeschlossenen Teilmenge $[0; r]$ *bzw.* $\{r\} \subset \mathbb{R}$ *unter der stetigen Funktion* $p \mapsto \|p - x\|$.

3. *Seien* $a, b \in \mathbb{R}$ *mit* $a \leq b$. *Dann ist* $[a; b] \subset \mathbb{R}$ *kompakt, da abgeschlossen und beschränkt.*

4. *Ein Quader* $[a_1; b_1] \times [a_2; b_2] \times ... \times [a_n; b_n] \subset \mathbb{R}^n$.

Satz 1.2.7 *Ist* $f : X \to Y$ *eine stetige Abbildung zwischen metrischen Räumen und* X *kompakt, so ist* $f(x) \subseteq Y$ *kompakt.*

Beweis Denn ist $(U_i)_{i \in I}$ eine offene Überdeckung von $f(x)$, so ist $(f^{-1}(U_i))_{i \in I}$ eine von X. □

Korollar 1.2.8 *Sei* X *ein kompakter metrischer Raum und* $f : X \to \mathbb{R}$ *eine stetige Funktion. Dann nimmt* f *auf* X *sein Maximum und Minimum an.*

Beweis Denn nach (1.2) ist $f(x)$ kompakt in $\mathbb{R}$, also beschränkt und abgeschlossen. Also etwa $\sup(f(X)) < \infty$. Sei $t_n \in f(X)$ eine Folge mit $t_n \to \sup(f(X))$. Dann folgt $\sup(f(X)) \in f(X)$. □

Korollar 1.2.9 *Jede beschränkte Folge mit* $\mathbb{R}^n$ *besitzt eine konvergente Teilfolge.*

Beweis Da die Folge beschränkt ist liegen alle Folgenglieder in einer genügend groß gewählten Teilmenge des $\mathbb{R}^n$, etwa in $K_R(0)$ mit $R > 0$ genügend groß. Nun folgt die Behauptung aus (1.2.4).
Dies gilt nicht in jedem metrischem Raum! Betrachten den Raum $C([0; 1])$ mit der Supremumsnorm und die Funktionenfolgen $f_\nu \in C([0; 1]), \nu = 1, 2, ...$ Dann gilt für zwei $\nu, \mu \in \mathbb{N}$: $\|f_\nu - f_\mu\| = 1$ und es gilt $\|f_\nu\| = 1 \forall \nu \in \mathbb{N}$. □

Die Folge kann also keine Teilfolge besitzen. Insbesondere ist die Einheitssphäre in $C([0; 1])$ nicht kompakt.

1.3 Partielle Ableitungen

Unser Ziel ist es nun, Abbildungen $\mathbb{R}^n \to \mathbb{R}^m$ auf Differenzierbarkeit zu untersuchen und das Differential für solche Abbildungen zu erklären. Zunächst wollen wir uns aber auf die Methoden der eindimensionalen Differentialrechnung beschränken und diese auf solche Abbildungen anwenden.
Sei $U \subseteq \mathbb{R}^n$ eine offene Teilmenge und $f : U \to \mathbb{R}$ eine Funktion. Der Graph von f ist die Menge $\Gamma_f := \{(x, y) \in U \times \mathbb{R} | y = f(x)\} \subset \mathbb{R}^{n+1}$ die wir uns im Fall $n = 2$ als eine Fläche im dreidimensionalen Raum vorstellen.
Eine Funktion $f : U \to \mathbb{R}$ ist auch festgelegt durch ihre Niveamengen $N_F(c) = f^{-1}(\{c\}) = \{x \in U | f(x) = c\}$, die wir uns im Fall $n = 2$ als Höhenlinien vorstellen können.

Definition 1.3.1 *Sei $U \subseteq \mathbb{R}^n$ offen, $f : U \to \mathbb{R}$ eine Funktion und $x \in U$. Die Funktion f heißt im Punkt $x \in U$ partiell differenzierbar bezüglich der i-ten Koordinatenrichtung, falls die Funktion $t \mapsto f(x_1, ..., x_{i-1}, t, x_{i+1}, ..., x_n)$ bei $t = x_i$ differenzierbar ist (im gewöhnlichen Sinne - als eindimensionale reelle Funktion). In diesem Fall nennt man den Wert der Ableitung der obigen eindimensionalen Funktion $t = x_1$, die i-te partielle Ableitung von f in x. Bezeichnung:*

$$D_i f(x) \text{ oder } \frac{\partial f}{\partial x_i}(x)$$

Beispiele

- $f_i(t) := f(x_1, ..., t, ..., x_n) \Rightarrow \frac{\partial f}{\partial x_i}(x) = \frac{df_i}{dt}(x_i)$
- Sei $U = \mathbb{R}^3$ und $f(x_1, x_2, x_3) := x_2^5 + x_1 x_2 + 4x_3^3$. Dann ist $D_2 f(x) = D_2 f(x_1, x_2, x_3) = \frac{\partial f}{\partial x_2}(x) = 5x_2^4 + x_1$.

Anschaulich Die Graphen der Funktionen $t \mapsto f(x_1, ..., t, ..., x_n)$ erhält man als Schnitte durch den Graphen von f.

Definition 1.3.2 *Sei f wie in (1.3.1). f heißt partiell differenzierbar, falls für alle $x \in U$ und alle $i = 1, ..., n$ die Funktion f in x partiell differenzierbar bezüglich der i-ten Koordinatenrichtung ist. f heißt stetig partiell differenzierbar, falls zusätzlich alle partiellen Ableitungen $D_i f := \frac{\partial f}{\partial x_i} : U \to \mathbb{R}, x \mapsto D_i f(x)$ stetig sind.*

Beispiel 1.3.3 (Beispiele)

1. Betrachte die Funktion $r : \mathbb{R}^n \to \mathbb{R}$ („Radius") mit $r(x) := \|x\| = \sqrt{x_1^2 + ... + x_n^2}$. Die Niveaumengen $N_r(c)$ sind für $c > 0$ Sphären um den Ursprung, leer für $c < 0$ und $N_r(0) = \{0\}$.

Behauptung Die Funktion r ist auf $\mathbb{R}^n \backslash \{0\}$ partiell differenzierbar mit $\frac{\partial r}{\partial x_i}(x) = \frac{x_i}{r(x)}$ für $x = (x_1, ..., x_n) \in \mathbb{R}^n \backslash \{0\}$.

Beweis Dies gilt, da die Funktion einer Variablen t, $t \mapsto \sqrt{x_1^2 + ... + t^2 + ... + x_n^2}$ differenzierbar ist. Wir berechnen die partielle Ableitung, indem wir $x_2, ..., x_n$ als Konstanten ansehen:

$$\frac{\partial r}{\partial x_1} = \frac{\partial}{\partial x_1}\sqrt{x_1^2 + ... + x_n^2} = \frac{2x_1}{2\sqrt{x_1^2 + ... + x_n^2}} = \frac{x_i}{r(x)}$$

2. **Radialsymmetrische Funktionen:** Sei $f : (0, \infty) \to \mathbb{R}$ eine differenzierbare Funktion. Damit ist die Verkettung $f \circ r : x \mapsto f(r(x))$ (die wir abkürzend als $f(r)$ bezeichnen) auf $\mathbb{R}^n \backslash \{0\}$ definiert und ist partiell differenzierbar. Aus der Kettenregel:

$$\frac{\partial}{\partial x_i} f(r) = f'(r)\frac{\partial r}{\partial x_i} = f'(r) \cdot \frac{x_i}{r}$$

3. **Eine Funktion, die partiell differenzierbar, aber unstetig ist:** Betrachte für $n \geq 2$:

$$F : \mathbb{R}^n \to \mathbb{R}, F(x) = \begin{cases} \frac{x_1 \cdot \dots \cdot x_n}{r(x)^{2n}} & \text{für } x \neq 0 \\ 0 & \text{für } x = 0 \end{cases}$$

Aus Beispiel (2) folgt mit der Produktregel, dass F auf $\mathbb{R}^n \backslash \{0\}$ partiell differenzierbar ist. Wir berechnen für $x \neq 0$:

$$\frac{\partial F}{\partial x_1}(x) = \frac{x_2 \cdot \dots \cdot x_n}{r^{2n}} + x_1 \cdot \dots \cdot \frac{\partial}{\partial x_1}(r^{-2n}) = \frac{x_2 \cdot \dots \cdot x_n}{r^{2n}} - 2n\frac{x_1^2 \cdot x_2 \cdot \dots \cdot x_n}{r^{2n+2}}$$

(Analog für $i = 2, ..., n$)
Da F aber auf den Koordinatenachsen gleich Null ist, ist die Funktion auch in $x = 0$ partiell differenzierbar. Andererseits ist F bei $x = 0$ unstetig: Wähle die Folge $a_\nu = \left(\frac{1}{\nu}, ..., \frac{1}{\nu}\right)$, $\nu \in \mathbb{N}$, dann gilt $r(a_\nu) = \sqrt{(\frac{1}{r})^2 + ... + (\frac{1}{r})^2} = \frac{\sqrt{n}}{r}$, somit

$$F(a_\nu) = \frac{\frac{1}{\nu^n}}{\frac{n^n}{\nu^{2n}}} = (\frac{\nu}{n})^n \to_{\nu \to \infty} \infty$$

Wegen $F(0) = 0$ ist F also in 0 unstetig! Für $n \geq 2$ folgt also aus „partiell differenzierbar” *nicht* „stetig” (Stetig folgt aber weiterhin aus „stetig partiell differenzierbar” bzw. (total) differenzierbar!).

$\square$

Definition 1.3.4 *Sei* $U \subseteq \mathbb{R}^n$ *offen und* $f : U \to \mathbb{R}$ *partiell differenzierbare Funktion. Wir nennen den Vektor*

$$grad(f(x)) := (\frac{\partial f}{\partial x_1}(x), ..., \frac{\partial f}{\partial x_n}(x))$$

den Gradient von f *im Punkt* $(x \in Y)$*. Andere Schreibweise* $\nabla f(x)$ *(„Nabla”).*

Zum Beispiel gilt für die Funktion $r(x) = \|x\|$:

$$\nabla r = \frac{x}{r}, x \neq 0$$

Definition 1.3.5 *Sei* $U \subseteq \mathbb{R}^n$ *offen. Unter einem Vektorfeld versteht man eine Abbildung*

$$v : U \to \mathbb{R}^n \tag{1.1}$$

die also jedem Punkt $x \in U$ *einem Vektor* $v(x)$ *in* $\mathbb{R}^n$ *zuordnet. Ein spezielles Vektorfeld ist der Gradient einer partiell differenzierbaren Funktion*

$$\nabla f : U \to \mathbb{R}^n, x \mapsto \nabla f(x) \tag{1.2}$$

Definition 1.3.6 *Sei* $U \subseteq \mathbb{R}^n$ *offen,* $v = (v_1, ..., v_n) : U \to \mathbb{R}^n$ *ein partiell differenzierbares Vektorfeld (d.h. alle Komponentenfunktionen* $v_i : U \to \mathbb{R}$ *sind partiell differenzierbar). Dann heißt die Funktion*

$$div(v) := \sum_{i=1}^{n} \frac{\partial v_i}{\partial x_i} \tag{1.3}$$

die Divergenz des Vektorfeldes v*.*

Für die Divergenz erhalten wir folgende Rechenregel: Seien $f : U \to \mathbb{R}$ und $v : U \to \mathbb{R}^n$ partiell differenzierbar. Dann gilt (nach der gewöhnlichen eindimensionalen Produktregel):

$$\frac{\partial}{\partial x_i}(f \cdot v_i) = \frac{\partial f}{\partial x_i} v_i + f \frac{\partial v_i}{\partial x_i} \tag{1.4}$$

Summation über i liefert:

$$div(f \cdot f) = \langle \nabla f, v \rangle + f \cdot div(v) \tag{1.5}$$

Beispiel 1.3.7 *Betrachte das Vektorfeld* $F : \mathbb{R}^n \backslash \{0\} \to \mathbb{R}^n$, $F(x) = \frac{x}{r}$, *wobei* $r = \|x\|$, *also folgt wie oben mit 1.3.3 (2):*

$$div(\frac{x}{r}) = \langle \nabla \frac{1}{r}, x \rangle + \frac{1}{r} \cdot div(x)$$

Aus 1.3.3 (2) $\frac{\partial}{\partial x_i} f(r) = f'(r) \frac{\partial r}{\partial x_i} = f'(r) \frac{\partial x_i}{r}$ *hier* $f(r) = \frac{1}{r}$.

$$\Rightarrow div(\frac{x}{r}) = \langle -\frac{1}{r^2} \frac{x}{r}, x \rangle + \frac{n}{r} = \frac{n-1}{r}$$

Wir wollen nun auch höhere partielle Ableitungen betrachten. Sei $U \subseteq \mathbb{R}^n$ offen und $f : U \to \mathbb{R}$ eine partiell differenzierbare Funktion. Sind alle partiellen Ableitungen selbst wieder partiell ableitbar, so heißt f zweimal, partiell differenzierbar. In diesem Fall existieren die zweiten Ableitungen:

$$D_i D_j f = \frac{\partial^2 f}{\partial x_i \partial x_j} \forall i, j = 1, ..., n \tag{1.6}$$

Entsprechend definiert man k-mal partiell differenzierbar. Sind k-ten partiellen Ableitungen

$$D_{i_1} ... D_{i_k} f = \frac{\partial^k f}{\partial x_{i_1} \cdot ... \cdot \partial x_{i_k}}$$

zusätzlich stetig, so nennt man f k-mal stetig partiell differenzierbar.

Satz 1.3.8 *Sei* $U \subseteq \mathbb{R}^n$ *offen und* $f : U \to \mathbb{R}$ *zweimal stetig partiell differenzierbar. Dann gilt für alle* $a \in U$ *und* $i, j = 1, ..., n$

$$D_i D_j f(a) = D_j D_i f(a) \tag{1.7}$$

Beweis Wir können ohne Einschränkung $n = 2$ und $a = (0, 0)$ annehmen. Für (x_1, x_2) schreiben wir (x, y).
Wegen der Offenheit von U gibt es ein $\delta > 0$, so dass $(-\delta, \delta) \times (-\delta, \delta) \subset U$ gilt.
Für $x, y \in (-\delta, \delta)$ betrachten wir den Term:

$$\underbrace{f(x, y) - f(x, 0)}_{:= F_y(x)} - \underbrace{(f(0, y) - f(0, 0))}_{:= F_y(0)} = F_y(x) - F_y(0)$$

Wobei wir für $y \in (-\delta, \delta)$ die neue Funktion $F_y : (-\delta, \delta) \to \mathbb{R}$, $F_y(x) := f(x, y) - f(x, 0)$ eingeführt wurde. Nach dem Mittelwertsatz der Differentialrechnung (Lemma 5.11, Analysis I Skript von Prof. Bunke) gibt es ein $\zeta \in (-|x|, |x|)$ mit $F_y(x) - F_y(0) = F'_y(\zeta) \cdot x = (D_1 f(\zeta, y) - D_1 f(\zeta, 0)) \cdot x$. Wieder nach dem Mittelwertsatz der Differentialrechnung gibt es ein η, $\eta \in (-|y|, |y|)$, so dass $D_1 f(\zeta, y) - D_1 f(\zeta, 0) = D_2 D_1 f(\zeta, \eta) \cdot y$. Insgesamt haben wir also gezeigt, dass $f(x, y) - f(x, 0) - f(0, y) + f(0, 0) = D_2 D_1 f(\zeta, \eta) \cdot yx$ für ein $|\zeta| < |x|$ und $|\eta| < |y|$

gilt.
Nun vertauschen wir die Rollen der beiden Variablen x und y, das heißt wir betrachten die Funktion $\tilde{f}(y,x) := f(x,y)$. Das obige Argument auf die Funktion $\tilde{f}$ angewendet zeigt, dass es $|\tilde{\zeta}| < |x|$ und $|\tilde{\eta}| < |y|$ gibt mit $yxD_2D_1\tilde{f}(\tilde{\eta},\tilde{\zeta}) = \tilde{f}(y,x) - \tilde{f}(y,0) - \tilde{f}(0,x) + \tilde{f}(0,0) = f(x,y) - f(0,y) - f(x,0) + f(0,0) = D_2D_1(\zeta,\eta)\cdot y\cdot x$.
Nun gilt aber $D_2D_1\tilde{f}(\tilde{\eta},\tilde{\zeta}) = D_1D_2f(\zeta,\eta) = D_2D_1(\zeta,\eta)$. Lässt man nun (x,y) gegen $(0,0)$ gehen, so auch $(\tilde{\zeta},\tilde{\eta})$ und (ζ,η) gegen Null wegen der Stetigkeit der zweiten partiellen Ableitung folgt

$$D_1D_2f(0,0) = D_2D_1f(0,0)$$

□

Beispiel 1.3.9 *Sei U eine offene Menge im $\mathbb{R}^3$ und sei $v : U \to \mathbb{R}^3$ ein partiell differenzierbares Vektorfeld. Wir definieren ein neues Vektorfeld $rot(v) : U \to \mathbb{R}^3$*

$$rot(v) := (\frac{\partial v_3}{\partial x_2} - \frac{\partial v_2}{\partial x_3}, \frac{\partial v_1}{\partial x_3} - \frac{\partial v_3}{\partial x_1}, \frac{\partial v_2}{\partial x_1} - \frac{\partial v_1}{\partial x_2})$$

genannt die Rotation (englisch „curl") des Vektorfeldes v. Sei nun $f : U \to \mathbb{R}$ eine zweimal stetig partiell differenzierbare Funktion. Wir berechnen:

$$rot(\nabla f) = (\frac{\partial^2 f}{\partial x_2 \partial x_3} - \frac{\partial^2 f}{\partial x_3 \partial x_2}, \frac{\partial^2 f}{\partial x_3 \partial x_1} - \frac{\partial^2 f}{\partial x_1 \partial x_3}, \frac{\partial^2 f}{\partial x_1 \partial x_2} - \frac{\partial^2 f}{\partial x_2 \partial x_1}) = 0$$

Damit sich ein partiell differenzierbares Vektorfeld als Gradient einer (2-mal stetig partiell differenzierbaren) Funktion darstellen lässt, muss die notwendige Bedingung $rot(v) = 0$ erfüllt sein.

$$\frac{\partial v_i}{\partial x_j} = \frac{\partial v_j}{\partial x_i} \tag{1.8}$$

Definition 1.3.10 *Sei $U \subseteq \mathbb{R}^n$ offen, $f : U \to \mathbb{R}$ eine zweimal stetig partiell differenzierbare Funktion. Setze*

$$\Delta f := div(\nabla f) = \frac{\partial^2 f}{\partial x_1^2} + ... + \frac{\partial^2 f}{\partial x_n^2}$$

wobei Δ der Laplace-Operator ist.

Schreibeweise $\frac{\partial^2 f}{\partial x_i^2}$ ist gleich $\frac{\partial^2 f}{\partial x_i \partial x_i}$. Funktionen f mit $\Delta f = 0$ nennt man harmonische Funktionen.

Bemerkung 1.3.11 $\Delta f = 0$ *ist ein Beispiel für eine partielle Differentialgleichung. Allgemein nennt man eine Gleichung Differentialgleichung, wenn eine Funktion gesucht ist und in der Gleichung neben der Funktion auch (partielle) Ableitungen der Funktion vorkommen.*

Hängt die gesuchte Funktion nur von einer Variablen ab, so handelt es sich um eine gewöhnliche Differentialgleichung. Hängt die gesuchte Funktion dagegen von mehreren Variablen ab (und treten partielle Ableitungen auf), nennt man die Gleichung eine partielle Differentialgleichung.
Die Gleichung $\Delta f = 0$ heißt auch Potentialgleichung.

Beispiel 1.3.12 *Wirkung des Laplace Operators auf radialsymmetrische Funktionen:*
Sei $f : (0, \infty) \to \mathbb{R}$ *zweimal stetig differenzierbar. Wir wollen den Laplace Operator auf die rotationssymmetrische Funktion*

$$\mathbb{R}^n \backslash \{0\} \to \mathbb{R}, x \mapsto f(\|x\|) = f(r)$$

wirken lassen. Nach Beispiel 1.3.3 (2) gilt $\nabla f(r) = f'(r)\frac{x}{r}$, *also*

$$\Delta f(r) = div(\nabla f) = div(f'(r)\frac{x}{r})$$

Wir vereinfachen mit der Produktregel für *div*:

$$\langle \nabla f'(r), \frac{x}{r} \rangle + f'(r) \cdot div(\frac{x}{r}) = \langle f''(r)\frac{x}{r}, xr \rangle + f'(r)\frac{n-1}{r}$$

Also $\Delta f(r) = f''(r) + \frac{n-1}{r} \cdot f'(r)$, das heißt falls $f(r)$ genau dann harmonisch auf $\mathbb{R}^n \backslash \{0\}$ ist, wenn $f : (0, \infty) \to \mathbb{R}$ eine Lösung der gewöhnlichen Differentialgleichung

$$f''(t) + \frac{n-1}{t} f'(t) = 0 \tag{1.9}$$

ist. Wir werden im zweiten Teil der Vorlesung sehen, wie die Menge der Lösungen einer solchen Gleichung bestimmt werden kann. Für den Moment beschränken wir uns darauf, zu bemerken, dass

$$f(r) = \frac{1}{r^{n-2}} \forall n \in \mathbb{N}$$

und im Fall $n = 2$ auch

$$f(r) = \ln r$$

Lösungen von 1.3 sind. Denn für $n > 2$ gilt:

$$\frac{d^2}{dr^2} r^{2-n} = \frac{d}{dr}(2-n)r^{1-n} = (1-n)(2-n)r^{-n} = -\frac{n-1}{r}\frac{d}{dr}r^{2-n}$$

und

$$\frac{d^2}{dr^2}\ln r = \frac{d}{dr}\frac{1}{n} = -r^{-2} = -\frac{1}{r}\frac{d}{dr}\ln r$$

Also gilt $\Delta\frac{1}{r^{n-2}} = 0$ und $\Delta \ln r = 0$ für $n = 2$. Der Laplace Operator kommt in vielen Gleichungen der mathematischen Physik vor (da t meistens $n = 2$ oder $n = 3$ ist), zum Beispiel in der Wärmeleistungsgleichung:

$$\Delta f - \frac{\partial f}{\partial t} = 0 \tag{1.10}$$

Genauer: $\frac{\partial^2 f}{\partial x_1^2} + ... + \frac{\partial^2 f}{\partial x_i^2} - \frac{\partial^2 f}{\partial t} = 0$. Hier ist $f : U \times I \to \mathbb{R}$ eine Funktion, wobei U eine offene Teilmenge des $\mathbb{R}^n$ eines räumlichen Bereichs beschreibt und das Intervall $I \subseteq \mathbb{R}$ als Zeitintervall aufgefasst wird.
Die Funktion f beschreibt eine Temperaturverteilung über den räumlichen Bereich U und ihren zeitlichen Verlauf. Ist für ein festes $t_0 \in I$ eine Funktion $g : U \to \mathbb{R}$ (2-mal stetig partiell Differenzierbar), also eine Temperaturverteilung zu einem festen Startzeitpunkt gegeben, so sagt die Wärmeleitungsgleichung die zeitliche Änderung der Temperatur durch die Wärmeleitung vorraus. Ist etwa g harmonisch, so ist

$$f(x, t) := g(x)$$

eine zeitunabhängige (also stationäre) Lösung. Nimmt g bei $u \in U$ ein isoliertes Maximum an, so besagt die Gleichung für eine Lösung f mit $f(x, t_0) = g(x)$, dass $\frac{\partial f}{\partial t}(x, t_0) < 0$ eine Abkühlung ist.
Eine weitere Wichtige Gleichung ist die Wellengleichung oder auch Schwingungsgleichung:

$$\Delta f - \frac{\partial^2 f}{\partial t^2} = 0 \tag{1.11}$$

die die Ausbreitung von Wellen oder Schwingungen beschreibt, wobei f die Amplitude der Schwingung ist.

Beispiel 1.3.13 *Wir betrachten die eindimensionale Wellengleichung:*

$$\frac{\partial^2 f}{\partial x^2} - \frac{\partial^2 f}{\partial t^2} = 0 \tag{1.12}$$

für eine zweimal partiell differenzierbare Funktion $f : \mathbb{R}^2 \to \mathbb{R}, (x, t) \mapsto f(x, t)$.

Sind $g, h : \mathbb{R} \to \mathbb{R}$ zweimal stetige differenzierbare Funktionen, so ist eine Lösung der eindimensionalen Wellengleichung gegeben durch:

$$f(x, t) := g(x + t) + h(x - t)$$

Denn es gilt für dieses f:

$$\frac{\partial^2 f}{\partial x^2}(x, t) = g''(x + t) + h''(x - t) = \frac{\partial^2}{\partial t^2}$$

Bemerkung Tatsächlich sind alle Lösungen von dieser Gestalt. Zum Beispiel $g(x) = h(x) = \cos(x)$.
$\Rightarrow f(x,t) = \cos(x+t) + \cos(x-t) = \cos(x)\cos(t) + \sin(x)\sin(t) + \cos(x)\cos(t) - \sin(x)\sin(t) = 2\cos(x)\cos(t)$.

1.4 Totale Differenzierbarkeit

Wir wollen nun die sogenannte totale Differenzierbarkeit einer Abbildung $f : \mathbb{R}^n \to \mathbb{R}^m$ als gewisse Approximierbarkeit durch lineare Abbildungen definieren. Das Differential einer solchen Abbildung ist die approximierte lineare Abbildung. Totale Differenzierbarkeit ist ein wesentlich stärkerer Begriff als partielle Differenzierbarkeit. So folgt aus der totalen Differenzierbarkeit insbesondere die Stetigkeit.

Definition 1.4.1 *Sei $U \subseteq \mathbb{R}^n$ eine offene Menge und $f : U \to \mathbb{R}^n$ eine Abbildung, f heißt im Punkt $x \in U$ total differenzierbar (oder Differenzierbar), falls es eine lineare Abbildung $A : \mathbb{R}^n \to \mathbb{R}^m$ gibt, die f in der Nähe des Punktes x in folgenden Sinne approximiert. Es gibt eine Umgebung von x, in der gilt*

$$f(x+\xi) = f(x) + A(\xi) + \varphi(\xi) \tag{1.13}$$

Wobei φ eine in einer Umgebung der Null in $\mathbb{R}^n$ definierten Funktion mit Werten in $\mathbb{R}^n$ ist, so dass

$$\lim_{\xi \to 0} \frac{\varphi(\xi)}{\|\xi\|} = 0$$

Bemerkung 1.4.2

1. *Die Bedingung an φ ist eine Forderung an die Güte der Approximation. Würde man etwas von $\lim_{\xi \to 0} \varphi(\xi) = 0$ verlangen, so wäre aus der Definition lediglich die Stetigkeit von f bei x folgen. Insbesondere folgt aus der Definition sofort, dass eine in x differenzierbare Abbildung auch in x stetig it.*

2. *Für $m = n = 1$ liefert dies die übliche Differenzierbarkeit von Funktionen in einer Variablen und $A = f'(x)$ ist die gewöhnliche Ableitung (Tangentensteigung).*

3. *Ist die lineare Abbildung A durch eine $m \times n$-Matrize (a_{ij}) gegeben und schreiben die Komponenten*

$$f = \begin{pmatrix} f_1 \\ \vdots \\ f_m \end{pmatrix}, \varphi = \begin{pmatrix} \varphi_1 \\ \vdots \\ \varphi_m \end{pmatrix},$$

so lässt (1.13) schreiben als

$$f_i(x+\xi) = f_i(x) + \sum_{j=1}^{n} a_{ij}\xi_j + \varphi_i(\xi), i = 1, ..., m \tag{1.14}$$

Dies zeigt auch, dass f genau dann in $x \in U$ differenzierbar ist, wenn alle Komponentenfunktionen in x differenzierbar sind. (Denn $\lim_{\xi\to 0} \frac{\varphi(\xi)}{\|\xi\|} = 0 \Leftrightarrow \lim_{\xi\to 0} \frac{\varphi_i(\xi)}{\|\xi\|} = 0 \forall i$).

4. *Die Bedingung (1.13) in der Definition schreibt man oft so:*

$$f(x+\xi) = f(x) + A\xi + o(\|\xi\|) \tag{1.15}$$

Wobei o das „Landau-Symbol" bezeichnet. Dies steht für eine in einer Umgebung der Null definierten Funktion φ, für die gilt:

$$\varphi(o) = 0, \lim_{\xi\to 0} \frac{\varphi(\xi)}{\|\xi\|} = 0$$

Dieses Symbol wird oft als Abkürzung für einen Fehlerterm (der eventuell zu vernachlässigen ist) benutzt. Mit dem Argument $\|\xi\|$ gibt man die Güte der Approximation an, z.B: ist für t reell:

$$t^3\cos(t) = o(t^3)$$

Beachte aber, dass $o(\|\xi\|)$ nicht bedeutet, dass der Fehler eine Funktion von $\|\xi\|$ ist.

Satz 1.4.3 *Sei $U \subseteq \mathbb{R}^n$ offen und $f : U \to \mathbb{R}^n$ eine Abbildung, die im Punkt $x \in U$ differenzierbar ist, so dass*

$$f(x+\xi) = f(x) + A\xi + o(\|\xi\|)$$

mit der reellen $m \times n$-Matrix A. Dann sind alle Komponenten von f partiell differenzierbar in x und es gilt:

$$a_{ij} = \frac{\partial f_i}{\partial x_j}(x) \tag{1.16}$$

Insbesondere Ist die Matrix A durch f eindeutig festgelegt.

Beweis Seien $i \in \{1, ..., m\}$ und $j \in \{1, ..., n\}$. Es gilt

$$f_i(x+\xi) = f_i(x) + \sum_{k=1}^{n} a_{ik}\xi_k + \varphi_i(\xi)$$

wobei $\varphi_i(\xi) = o(\|\xi\|)$. Setze nun $\xi := h \cdot e_j$, wobei e_j den j-ten kanonischen Basisvektor des $\mathbb{R}^n$ bezeichnet, $h \in (-\varepsilon, \varepsilon)$.

$$f_i(x+he_j) = f_i(x) + a_{ij}h + \varphi_i(he_j)$$

was für $h \neq 0$ äquivalent ist zu

$$\frac{f_i(x+he_j) - f_i(x)}{h} = a_{ij} + \underbrace{\frac{\varphi_i(he_j)}{h}}_{\longrightarrow_{h\to 0} 0} \tag{1.17}$$

Dies zeigt, dass f_i bei x nach den j-ten Koordinatenrichtungen partiell differenzierbar ist mit

$$\frac{\partial f_i}{\partial x_j}(x) = a_{ij}$$

□

Definition 1.4.4 *Sei $U \subseteq \mathbb{R}^n$ offen und $f : U \to \mathbb{R}^m$ eine Abbildung, die im Punkt $x \in U$ partiell differenzierbar ist. Dann nennt man die Matrix*

$$Df(x) = (\frac{\partial f_i}{\partial x_j})_{1\le i\le m, 1\le j\le n}$$

die aus den partiellen Ableitungen von f bei x gebildet wird die Jacobi-Matrix als Funktionalmatrix von f bei x.

Andere Schreibweisen:

$$Df(x) = J_f(x) = \frac{df}{dx}(x) = \frac{\partial(f_1, ..., f_m)}{\partial(x_1, ..., x_n)}(x)$$

Ist f bei x differenzierbar, so nennt man $Df(x)$ auch das Differential oder die Ableitung von f bei x.

Beispiel Sei $f : \mathbb{R}^3 \to \mathbb{R}^2, f(x_1, x_2, x_3) = (x_1x_2, x_2^2)$. Dann gilt:

$$D_f(x_1, x_2, x_3) = \begin{pmatrix} x_2 & x_1 & 0 \\ 0 & 2x_2 & 0 \end{pmatrix}$$

Um die Differenzierbarkeit einer Abbildung nachzuweisen, steht uns bis jetzt nur die Definition (1.4.1) zur Verfügung. Ein oft einfach anzuwendendes Kriterium ist das folgende:

Satz 1.4.5 *Sei $U \subseteq \mathbb{R}^n$ offen und $f : U \to \mathbb{R}$ eine in U partiell differenzierbare Funktion. Alle partiellen Ableitungen D_if seien im Punkte $x \in U$ stetig. Dann ist f in x (total) differenzierbar.*

Beweis Der Einfachheit halber führen wir den Beweis nur für $n = 2$ Dimensionen durch (Das Argument lässt sich auf n-Dimensionen verallgemeinern). Sei $\delta > 0$, so dass das Achsenparallele Quadrat mit Mittelpunkt x und Seitenlänge 2δ ganz in U liegt.
Sei nun $\xi = \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix}$ mit $|\xi_1|, |\xi_2| < \delta$. Wir gehen nun von Punkt x zum Punkt $x + \xi$, indem wir uns parallel zu den Koordinatenachsen bewegen. Seien $x' = x + \xi_1 e_1, x'' = x + \xi_1 e_1 + \xi_2 e_2 = x + \xi$. Dann gilt, wobei wir 2 mal den Mittelwertsatz der (eindimensionalen) Differentialrechnung anwenden:

$$f(x+\xi)-f(x) = f(x'')-f(x')+f(x')-f(x) = \underbrace{f(x' + \xi_2 e_2) - f(x')}_{D_2 f(x'+\theta_2\xi_2 e_2)\xi_2} + \underbrace{f(x + \xi_1 e_1) - f(x)}_{D_1 f(x+\theta_1\xi_1 e_1)\xi_1}$$

Mit $\theta_1, \theta_2 \in [0, 1]$.
Setze nun $a_i := D_i f(x)$ für $x = 1, 2$, dann können wir schreiben:

$$f(x + \xi) = f(x) + \sum_{i=1}^{2} a_i \xi_i - \underbrace{\sum_{i=1}^{2} (D_i(x^{('\cdot(i-1))} + \theta_i \xi_i e_i) - a_i)\xi_i}_{=:\varphi(\xi)}$$

Aufgrund der Stetigkeit über partielle Ableitungen $D_i f$ bei x gilt:

$$\lim_{\xi \to 0} \frac{\varphi(\xi)}{\|\xi\|} = 0$$

□

Bemerkung 1.4.6 *Es gelten also folgende Implikationen: Stetig partiell differenzierbar $\Rightarrow$ (total) differenzierbar $\Rightarrow$ stetig. Außerdem, folgt aus (total) differenzierbar $\Rightarrow$ partiell differenzierbar.*

Insbesondere sind stetig partiell differenzierbare Funktionen auch stetig. Die Umkehrungen geilten im Allgemeinen nicht, vergleiche Beispiel (1.3.3)(3). Wegen dieser Zusammenhänge nennt man stetig partiell differenzierbare Funktionen kurz stetig differenzierbar.

Beispiel 1.4.7 *Aus Satz (1.4.5) folgt, dass Polynome auf dem $\mathbb{R}^n$ differenzierbar sind (denn sie sind stetig partiell differenzierbar).*

Satz 1.4.8 (Kettenregel) *Seien $U \subseteq \mathbb{R}^n$ und $V \subseteq \mathbb{R}^m$ offene Mengen und $g : U \to \mathbb{R}^m, f : V \to \mathbb{R}^k$ Abbildungen mit $g(U) \subseteq V$. Die Abbildung g sei im Punkt $x \in U$ differenzierbar und die Abbildung f im Punkt $y := g(x)$ differenzierbar. Dann ist die Verkettung*

$$f \circ g : U \to \mathbb{R}^k, x \mapsto f(g(x)) \tag{1.18}$$

im Punkt x differenzierbar und es gilt für ihre Ableitung

$$D(f \circ g)(x) = Df(g(x)) \cdot Dg(x), (f \circ g)'(x) = f'(g(x)) \cdot g'(x) \tag{1.19}$$

Wobei die Multiplikation auf der rechten Seite Matrizenmultiplikation (bzw. Verkettung von linearen Abbildungen) bezeichnet.

Beweis Sei $A := Dg(x), B := Df(y)$. Wir wollen zeigen, dass $f \circ g$ bei x differenzierbar ist mit $D(f \circ g)(x) = BA$. Es gilt

$$\begin{aligned} g(x+\xi) &= g(x) + A\xi + \varphi(\xi) \\ f(y+\eta) &= f(y) + B\eta + \psi(\eta) \\ \varphi(\xi) &= o(\|\xi\|), \psi(\eta) = o(\|\eta\|) \end{aligned}$$

Untersuche nun

$$\begin{aligned} f \circ g(x+\xi) &= f(g(x+\xi)) = f(\underbrace{g(x)}_{y} + \underbrace{A\xi + \varphi(\xi)}_{=:\eta}) = \\ &= f(g(x)) + B(A\xi + \varphi(\xi)) + \psi(A\xi + \varphi(\xi)) = \\ &= f \circ g(x) + BA\xi + \underbrace{B\varphi(\xi) + \psi(A\xi + \psi(\xi))}_{=\chi(\xi)} \end{aligned}$$

Der Satz ist bewiesen, wenn wir $\chi(\xi) = o(\|\xi\|)$ gezeigt haben.

$$\chi(\xi) = o(\|\xi\|) \Leftrightarrow \lim_{\xi \to 0} \frac{\chi(\xi)}{\|\xi\|} = 0$$

Aus $\varphi(\xi) = o(\|\xi\|)$ folgt sofort $B\varphi(\xi) = o(\|\xi\|)$. Außerdem ist wegen $\varphi(\xi) = o(\|\xi\|)$ der Term $\frac{\varphi(\xi)}{\|\xi\|}$ für kleine ξ beschränkt. Es gibt also eine Konstante $k > 0$ mit $\varphi(\xi) < K\|\xi\|$. Da (stetige) lineare Abbildungen $\xi \mapsto A\xi$ beschränkt ist, gibt es also eine Konstante $c > 0$ mit $\|A\xi\| < c \cdot \|\xi\|$.
Wegen $\psi(\eta) = o(\|\eta\|)$ gilt $\psi(\eta) = \|\eta\| \cdot \psi_1(\eta)$ mit $\lim_{\eta \to 0} \psi_1(\eta) = 0$. Daraus folgt $\|\psi(A\xi + \varphi(\xi))\| \leq \|A\xi + \varphi(\xi)\| \cdot \|\psi_1(A\xi + \varphi(\xi))\| < (c+k)\|\xi\| \cdot \|\psi_1(A\xi + \varphi(\xi))\|$.
Also gilt $\lim_{\xi \to 0} \frac{\psi(A\xi + \varphi(\xi))}{\|\xi\|} = 0$, insgesamt haben wir $\chi(\xi) = o(\|\xi\|)$ gezeigt. □

Beispiel 1.4.9 *Seien $U, V \subseteq \mathbb{R}^n$ und $g : U \to V$ ein Diffeomorphismus, d.h. g ist bijektiv, differenzierbar und die Umkehrabbildung $f := g^{-1}$ ist ebenfalls differenzierbar. Dann gilt $\forall x \in U$, dass $Dg(x)$ eine invertierbare Matrix ist und es gilt $(Dg(x))^{-1} = D(g^{-1})(g(x))$.*

Beweis Es gilt $f \circ g = id_U$, also gilt nach Kettenregel

$$Df(g(x))Dg(x) = D(f \circ g)(x) = D(id_U)(x) = 1_n \Rightarrow (Dg(x))^{-1} = D(g^{-1})(g(x)) \tag{1.20}$$

□

Beispiel 1.4.10 *Sei* $U \subseteq \mathbb{R}^n, I \subset \mathbb{R}$ *offen und seien* $f : U \to \mathbb{R}$ *eine differenzierbare Funktion, sowie* $g : I \to \mathbb{R}^n$ *eine differenzierbare Kurve. Dann ist die Verkettung* $f \circ g : I \to \mathbb{R}$ *differenzierbar und es gilt*

$$(f \circ g)'(t) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(g(t)) \frac{dg_i}{dt}(t) = \langle \nabla f(g(t)), \dot{g}(t) \rangle$$

Beweis Dies folgt direkt aus der Kettenregel, denn

$$\begin{aligned} Df(g(t)) &= \left(\frac{\partial f}{\partial x_1}(g(t)), ..., \frac{\partial f}{\partial x_n}(g(t))\right) \\ Dg(t) &= \begin{pmatrix} \frac{dg_1}{dt}(t) \\ \vdots \frac{dg_n}{dt}(t) \end{pmatrix} \end{aligned}$$

Die Behauptung folgt nun aus:

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

Definition 1.4.11 *Sei* $U \subseteq \mathbb{R}^n$ *offen und* $f : U \to \mathbb{R}$ *eine Abbildung. Sei weiter* $x \in U$ *ein Punkt und* $v \in \mathbb{R}^n$ *ein Vektor. Unter der Richtungsableitung von* f *bei* x *in Richtung* v *versteht man (im Fall der Existenz) den Wert:*

$$D_v f(x) := \frac{d}{dt} f(x + tv)|_{t=0} \tag{1.21}$$

Für $v = e_i$ ist also $D_v f(x) = D_{e_i} f(x) = D_i f(x)$ die i-te partielle Ableitung.

Satz 1.4.12 *Sei* $U \subseteq \mathbb{R}^n$ *offen und* $f : U \to \mathbb{R}^n$ *eine stetige differenzierbare Ableitung. Dann gilt für jedes* $x \in U$ *und* $v \in \mathbb{R}^n$

$$\underbrace{D_v f(x)}_{\textit{Richtungsableitung}} \quad = \quad \underbrace{Df(x)v}_{\textit{Matrix mal Vektor}}$$

Ist insbesondere $f : U \to \mathbb{R}$ *eine stetige Funktion, so gilt:*

$$D_v f(x) = \langle v, \nabla f(x) \rangle \tag{1.22}$$

Beweis Sei $g : \mathbb{R} \to \mathbb{R}^n$ definiert durch $g(t) := x + tv = (x_1 + tv_1, ..., x_n + tv_n)$. Für genügend kleines $\varepsilon > 0$ gilt $g((-\varepsilon, \varepsilon)) \subseteq U$, also ist $h := f \circ g : (-\varepsilon, \varepsilon) \to \mathbb{R}^m$ definiert. Nach Definition der Richtungsableitung gilt

$$D_v f(x) = \frac{d}{dt} f(x + tv)|_{t=0} = \frac{dh}{dt}(0)$$

Nun folgt aus der Kettenregel (bzw. Beispiel (1.4.10)) für jedes $j = 1, ..., m$:

$$\frac{dh_j}{dt}(0) = \sum_{i=1}^{n} \frac{\partial f_j}{\partial x_i}(g(0))\frac{dg_i}{dt}(0) = \sum_{i=0}^{n} \frac{\partial f_j}{\partial x_i}(x)v_i$$

worraus die erste Behauptung folgt. Für $m = 1, f_1 = f$, gilt speziell

$$D_v f(x) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(x)v_i = \langle v, \nabla(x)\rangle$$

□

Bemerkung 1.4.13 *Gilt* $\|v\| = 1$ *und* $\nabla f(x) \neq 0$*, ist der Winkel*

$$\cos\theta = \langle v, \frac{\nabla f(x)}{|\nabla f(x)\|}\rangle = \frac{1}{\|\nabla f(x)\|} D_v f(x) \tag{1.23}$$

Daraus sehen wir, dass für $\|v\| = 1$ die Richtungsableitung genau dann maximal wird, wenn v und $\nabla f(x)$ in diesselbe Richtung zeigen, das heißt der Vektor $\nabla f(x)$ gibt also die Richtung des stärksten Anstiegs von f an (und seine Länge die Steigung in diese Richtung).

Beispiel 1.4.14 *Wir betrachten die eindimensionale Wellengleichung*

$$\frac{\partial^2 f}{\partial x^2} - \frac{\partial^2 f}{\partial t^2} = 0$$

für eine zweimal stetig differenzierbare Funktion $f : \mathbb{R}^2 \to \mathbb{R}, (x,t) \mapsto f(x,t)$*. Wir führen eine Koordinatentransformation durch. Dazu betrachten wir die Abbildung* $g : \mathbb{R}^2 \to \mathbb{R}^2, (\xi, \eta) \mapsto (\xi + \eta, \xi - \eta)$ *(sog. Lichtartige Koordinaten) und betrachten nun die Funktion* $F := f \circ g$ *auf* $\mathbb{R}^2$*. Berechne*

$$\begin{aligned} F_\xi : &= \frac{\partial F}{\partial \xi}(\xi,\eta) = \frac{\partial f}{\partial x}(g(\xi,\eta)) \cdot \frac{\partial g_1}{\partial \xi} + \frac{\partial f}{\partial t}(g(\xi,\eta)) \cdot \frac{\partial g_2}{\partial \xi} = \\ &= \frac{\partial f}{\partial x}(g(\xi,\eta)) + \frac{\partial f}{\partial t}(g(\xi,\eta)) \\ F_\nu : &= \frac{\partial F}{\partial \nu}(\xi,\eta) = \frac{\partial f}{\partial x}\frac{\partial g_1}{\partial \eta} + \frac{\partial f}{\partial t}\frac{\partial g_2}{\partial \eta} = \frac{\partial f}{\partial x} - \frac{\partial f}{\partial t} \end{aligned}$$

Es folgt

$$\begin{aligned} (F_\xi)_\eta &= \frac{\partial}{\partial \eta}(\frac{\partial f}{\partial x} + \frac{\partial f}{\partial t})(g(\xi,\eta)) = \\ &= (\frac{\partial^2 f}{\partial x \partial x} - \frac{\partial^2 f}{\partial t \partial x} + \frac{\partial^2 f}{\partial x \partial t} - \frac{\partial^2 f}{\partial t \partial t})(g(\xi,\eta)) = \\ &= \frac{\partial^2 f}{\partial x^2} - \frac{\partial^2 f}{\partial t^2} \end{aligned}$$

f ist also genau dann eine Lösung der Wellengleichung, wenn $F_{\xi\eta} = 0$ gilt. Dies gilt genau dann wenn F_ξ nicht von η abhängt, also genau dann wenn

$$F(\xi, \eta) = \int_{\xi_0}^{\xi} \varphi(s)ds + \underbrace{h(\eta)}_{*}$$

(Konstante kann von η abhängen.) Für eine stetig differenzierbare Funktion $\varphi : \mathbb{R} \to \mathbb{R}$ und zweimal stetig differenzierbar Funktion $h : \mathbb{R} \to \mathbb{R}$. Ingesamt folgt: f ist eine Lösung der Gleichung genau dann wenn:*

$$F(\xi, \eta) = \underbrace{f(\xi + \eta, \xi - \eta)}_{f(x,t)} = g(\xi) + h(\eta) = g(\frac{x+t}{2}) + h(\frac{x-t}{2})$$

Mittelwertsatz im Mehrdimensionalen Im Beweis von Satz (1.3.8) und Satz (1.4.5) haben wir den eindimensionalen Mittelwertsatz in folgender Form verwendet:

Mittelwertsatz Ist f eine differenzierbare Funktion auf dem Intervall I und sind $x + \xi \in I$, so gibt es $\theta \in (0, 1)$, so dass gilt:

$$f(x + \xi) - f(x) = f'(\theta\xi + x)\xi$$

In dieser Form lässt sich der Mittelwertsatz nicht auf Vektorförmige Funktionen $f : I \to \mathbb{R}^n$ übertragen.
Setzt man jedoch vorraus, dass f stetig differenzierbar ist, so folgt aus dem Hauptsatz der Differential- und Integralrechnung:

$$f(x + \xi) - f(x) = \int_{x}^{x+\xi} f'(u)du = (\int_0^1 f'(x + t\xi)dt)\xi$$

Im Gegensatz zur ersten Variante haben wir hier den Wert der Ableitung an einer Zwischenstelle durch den Mittelwertsatz der Ableitung auf dem Intervall zwischen x und $x + \xi$ ersetzt. In dieser Form lässt sich der Satz ins Mehrdimensionale übertragen.
Da das Differential einer Vektorwertigen Funktion mehrere Variablen einer Matrix ist, benötigen wir dazu ein Integral über eine matrixwertige Funktion. Sei $A(t) := (a_{ij}(t))$ eine reelle nxm-Matrix, deren Einträge stetige Funktionen auf einem Intervall $[t_0, t_1]$ seien. Dann definieren wir das Integral komponentenweise:

$$\int_{t_0}^{t_1} A(t)dt = \begin{pmatrix} \int_{t_0}^{t_1} a_{11}(t)dt & \dots & \int_{t_0}^{t_1} a_{1n}(t)dt \\ \vdots & & \vdots \\ \int_{t_0}^{t_1} a_{n1}(t)dt & \dots & \int_{t_0}^{t_1} a_{nn}(t)dt \end{pmatrix}$$

Satz 1.4.15 *Sei* $U \subseteq \mathbb{R}^n$ *offen und* $f : U \to \mathbb{R}^m$ *eine stetige differenzierbare Abbildung. Sei* $x \in U$ *und* $\xi \in \mathbb{R}^n$ *derart, dass die ganze Strecke* $x + t\xi, 0 \leq t \leq 1$ *in* U *liegt. Dann gilt:*

$$f(x+\xi) - f(x) = \int_0^1 Df(x+t\xi)dt\xi$$

Beweis Sei $f = (f_1, ..., f_n)$. Betrachte $g_i(t) := f_i(x+t\xi)$. Damit gilt: $f_i(x+\xi) - f_i(x) = g_i(1) - g_i(0) = \int_0^1 g_i'(t)dt = \int_0^1 \sum_{j=1}^n \frac{\partial f_i}{\partial x_j}(x+t\xi)\xi_j dt$ □

Einige wichtige Folgerungen:

Definition 1.4.16 (Norm einer linearen Abbildung) *Seien* V, W *normierte Vektorräume und* $A : V \to W$ *eine stetige (beschränkte) lineare Abbildung. Setze*

$$\|A\| := \sup\{\|A(v)\| \mid v \in V, \|v\| \leq 1\}$$

Nach Satz (1.1.8) und Beispiel (1.1.7)(7) gilt $\|A\| < \infty$.

Begründung

$$\|A(x)\| = \|(A(\frac{x}{\|x\|})\|x\|)\| \leq \|A\|\|x\|$$

Lemma 1.4.17 *Sei* $v : [a,b] \to \mathbb{R}^m$ *eine stetige Vektorwertige Funktion auf dem Intervall* $[a,b] \subset \mathbb{R}$. *Dann gilt, wobei* $\|\cdot\|$ *hier die euklidische Norm auf* $\mathbb{R}^n$ *bezeichnet:*

$$\|\int_a^b v(t)dt\| \leq \int_a^b \|v(t)\|dt$$

Beweis Setze $u := \int_a^b v(t) \in \mathbb{R}^n$ und $K := \|u\|$. Dann gilt:

$$\begin{aligned} K^2 = \langle u, u\rangle &= \langle \int_a^b v(t)dt, u\rangle = \int_a^b \langle v(t), u\rangle dt \leq \\ &\leq \int_a^b \|v(t)\| \cdot \|u\| dt = K \cdot \int_a^b \|v(t)\|dt \end{aligned}$$

□

Korollar 1.4.18 *Sei $U \subseteq \mathbb{R}$ offen und $f : U \to \mathbb{R}^m$ eine stetig differenzierbare Abbildung. Sei $x \in U$ und $\xi \in \mathbb{R}^n$ derart, dass die ganze Strecke $x + t\xi, 0 \leq t \leq 1$ in U liegt. Dann gilt:*

$$\|f(x+\xi) - f(x)\| \leq M\|\xi\|$$

Wobei

$$M := \sup_{t \in [0,1]} \|Df(x+t\xi)\|$$

die Norm der Matrix darstellt.

Beweis Nach dem Mittelwertsatz (1.4.15) gilt:

$$\|f(x+\xi) - f(x)\| = \| \int_0^1 (Df(x+\xi)\xi)dt\| \leq \int_0^1 \|Df(x+t\xi)\| \cdot \|\xi\| dt \leq M \cdot \|\xi\|$$

□

1.5 Taylorformel und lokale Extrema

Das Differential liefert eine Annäherung durch eine affin-lineare Funktion, oder mit anderen Worten, durch ein Polynom ersten Grades. Dies ist jedoch nicht immer ausreichend um die Eigenschaften einer Abbildung zu untersuchen. Deshalb werden wir jetzt auch Annäherungen durch Polynome beliebig hohen Grades betrachten. Dies führt uns auf die Taylorformel, die Approximationen beliebig hoher Ordnung ermöglicht. Wir werden dies insbesondere anwenden, um Funktionen auf lokale Extrema zu untersuchen.

Bezeichnung 1.5.1 *Sei $\alpha = (\alpha_1, ..., \alpha_n) \in \mathbb{N}^n$ ein n-Tupel natürlicher Zahlen. Wir definieren $|\alpha| := \alpha_1 + ... + \alpha_n, \alpha! = \alpha_1! \cdot \dots \cdot \alpha_n!$. Ist f eine $|\alpha|$-mal stetig partiell differenzierbare Funktion*

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

wobei $D_i^{\alpha_i} = \underbrace{D_i ... D_i}_{a_i \text{ mal}}$ und für $x = (x_1, ..., x_n) \in \mathbb{R}^n$. Sei $x^\alpha := x_1^{\alpha_1} \cdot \dots \cdot x_n^{\alpha_n}$.

Beispiel Sei $\alpha = (1, 4, 2)$ und $f : \mathbb{R}^3 \to \mathbb{R}$ 7-mal stetig partiell differenzierbar. Dann gilt: $\alpha = 1 + 4 + 2 = 7, \alpha! = 1! \cdot 4! \cdot 2! = 48$.

$$D^\alpha f = \frac{\partial^7 f}{\partial x_1^1 \partial x_2^4 \partial x_3^2}$$

Satz 1.5.2 *Sei $U \subseteq \mathbb{R}^n$ offen und $f : U \to \mathbb{R}$ eine k-mal stetig partiell differenzierbare Funktion. Sei $x \in U, \xi \in \mathbb{R}^n$, so dass die Strecke $x + t\xi, 0 \leq t \leq 1$ ganz in U liegt. Dann ist die Funktion*

$$g : [0,1] \to \mathbb{R}, g(t) := f(x + t\xi)$$

stetig differenzierbar und es gilt

$$g(t) = \frac{d^k g}{dt^k}(t) = \sum_{|\alpha|=k} \frac{k!}{\alpha!} D^\alpha f(x + t\xi)\xi^\alpha \tag{1.24}$$

Die Summe läuft hier über alle n-Tupel mit $|\alpha| = k$.

Beweis

1. Wir zeigen zunächst durch Induktion über k, dass

$$\frac{d^k g}{dt^k}(t) = \sum_{i_1,...,i_k=1}^{n} D_{i_k} ....D_{i_1} f(x + t\xi)\xi_{i_1} ..\xi_{i_k}$$

Für $k = 1$ liefert die Kettenregel (1.4.10):

$$\frac{dg}{dt}(t) = \frac{d}{dt} f(x_1 + t\xi_1, ..., x_n + t\xi_n) = \sum_{i=1}^{n} D_i f(x + t\xi)\xi_i$$

$k - 1 \rightsquigarrow k$:

$$\frac{d^{k-1} g}{dt^{k-1}}(t) = \sum_{i_1,...,i_{k-1}=1}^{n} D_{i_{k-1}} ....D_{i_1} f(x + t\xi)\xi_{i_1} ..\xi_{i_{k-1}}$$

$$\Rightarrow \frac{d^k g}{dt^k}(t) = \frac{d}{dt}(\frac{d^{k-1} g}{dt^{k-1}})(t) = \sum_{j=1}^{n} D_j(\sum_{i_1,...,i_{k-1}=1}^{n} D_{ik_1}...D_{i_1} f(x+t\xi)\xi_{i_1}...\xi_{i_{k-1}})\xi_j$$

2. Da f k-mal stetig ist kommt es auf die Reihenfolge der Differentation nicht an, und wir können in (1.24) einige Summanten zusammenfassen. Kommt unter den Indizes $(i_1, ..., i_k)$ die Zahl 1 genau α_1-mal vor, 2 genau α_2-mal, usw., so können wir schreiben:

$$D_{i_k}...D_{i_1} f(x + t\xi)\xi_{i_1}...\xi_{i_n} = D^{\alpha_1}...D^{\alpha_n}(x + t\xi)\xi_1^{\alpha_1}...\xi_n^{\alpha_n}$$

Nun gibt es genau

$$\frac{k!}{\alpha_1!...\alpha_n!} = \frac{k!}{\alpha!}$$

k-Tupel von Indizes $(i_1, ..., i_k)$ bei denen jeweils die Zahl ν genau α_ν-mal vorkommt, wobei $\alpha_1 + ... + \alpha_n = k$, so können wir (1.24) umschreiben als:

$$\frac{d^k g}{dt^k}(t) = \sum_{|\alpha|=k} \frac{k!}{\alpha!} D_1^{\alpha_1}...D_n^{\alpha_n} f(x + t\xi)\xi_1^{\alpha_1}...\xi_n^{\alpha_n} = \sum_{|\alpha|=k} \frac{k!}{\alpha!} D^\alpha f(x + t\xi)\xi^\alpha$$

$\square$

Definition 1.5.3 *Wir setzen für* $m \in \mathbb{N}$*:*

$$P_m(\xi) := \sum_{|\alpha|=m} \frac{D^\alpha f(x)}{\alpha!} \xi^\alpha \tag{1.25}$$

Dann ist P_m *ein (homogenes) Polynom (m-ten Grades) in der Variablen* $\xi = (\xi_1, ..., \xi_n) \in \mathbb{R}^n$*. Wir nennen*

$$T_m(\xi) := \sum_{k=0}^{m} P_k(\xi) = \sum_{|\alpha|\leq m} \frac{D^\alpha f(x)}{\alpha!} \xi^\alpha \tag{1.26}$$

das Taylorpolynom m-ten Grades von f *im Punkt* x*. Für* $n = 1$ *liefert dies das aus Analysis I bekannte Taylorpolynom für Funktionen einer Variablen, so dass dann gilt:*

$$\sum_{k=0}^{m} P_k(t) = \sum_{k=0}^{m} \frac{f^{(k)}(x)}{k!} t^k \tag{1.27}$$

Satz 1.5.4 (Taylorformel) *Sei* $U \subseteq \mathbb{R}^n$ *offen,* $x \in U$ *und* $\xi \in \mathbb{R}^n$*, so dass* $x + t\xi \in U \forall t \in [0; 1)$*. Sei* $f : U \to \mathbb{R}$ *eine* $(k+1)$*-mal stetig differenzierbare Funktion, dann existiert ein* $\theta \in [0, 1]$*, so dass*

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

Beweis Indem wir den Satz auf die eindimensionale Taylorformel zurückführen. Dazu betrachten wir wieder wie in (1.5.2) die Funktion $g : [0, 1] \to \mathbb{R}, g(t) := f(x + t\xi)$. Nach der Taylorformel für Funktionen einer Veränderlichen existiert ein $\theta \in [0, 1]$, so dass

$$g(1) = \sum_{m=0}^{k} \frac{g^{(m)}(\theta)}{m!} + \frac{g^{(k+1)}(\theta)}{(k+1)!}$$

Nun folgt die Behauptung, wenn wir die in Satz (1.5.2) berechneten Terme für die Ableitung von g einsetzen. $\square$

Beispiel Wir berechnen das Taylorpolynom zweiten Grades der Funktion $f(x_1, x_2) = e^{x_1} \cdot x_2^2$ im Punkt $(0, 0)$. Dazu berechnen wir:

$$f(0, 0) = 0, D_1 f(x_1, x_2) = e^{x_1} x_2^2 \Rightarrow 0, ...D_2^2 f(x_1, x_2) = 2e^{x_1} \Rightarrow 2$$

Alle Ableitungen bis zur 2. Ordnung verschwinden außer $D_2^2 f(0, 0)$. Somit folgt:

$$T_2(\xi) = T_2(\xi_1, \xi_2) = \sum_{|\alpha|\leq 2} \frac{D^\alpha f(0, 0)}{\alpha!} \xi^\alpha = \frac{D_2 D_2 f(0, 0)}{0! \cdot 2!} \xi_1^0 \xi_2^2 = \xi_2^2$$

Korollar 1.5.5 *Sei* $U \subseteq \mathbb{R}^n$ *offen,* $x \in U$ *und* $\delta > 0$*, so dass* $B_\delta(x) \in U$*. Sei* $f : U \to \mathbb{R}$ *eine* k*-mal stetig differenzierbare Funktion. Dann gilt* $\forall \xi \in \mathbb{R}^n$ *mit* $\|\xi\| < \delta$*:*

$$f(x+y) = \sum_{m=0}^{k} P_m(\xi) + o(\|\xi\|^k) \tag{1.28}$$

Beweis Nach (1.5.4) gibt es ein von ξ abhängiges $\theta \in [0,1]$. Somit folgt:

$$\begin{aligned} f(x+y) &= \sum_{|\alpha| \leq k-1} \frac{D^\alpha f(x)}{\alpha!}\xi^\alpha + \sum_{|\alpha|=k} \frac{D^\alpha f(x+\theta\xi)}{\alpha!}\xi^\alpha = \\ &= \sum_{|\alpha| \leq k} \frac{D^\alpha f(x)}{\alpha!}\xi^\alpha + \sum_{|\alpha|=k} r_\alpha(\xi)\xi^\alpha \end{aligned}$$

Wobei $r_\alpha(\xi) := \frac{D^\alpha f(x+\theta\xi) - D^\alpha f(x)}{\alpha!}$. Wegen der Stetigkeit von $D^\alpha f$ gilt $\lim_{\xi \to 0} r_\alpha(\xi) = 0$. Setze nun $\varphi(\xi) := \sum_{|\alpha|=k} r_\alpha(\xi) \cdot \xi^\alpha$. Es folgt wegen:

$$\frac{|\xi^\alpha|}{\|\xi\|^k} = \frac{|\xi^{\alpha_1} ... \xi^{\alpha_n}|}{\|\xi\|^{\alpha_1} ... \|\xi\|^{\alpha_n}} \leq 1$$

für $|\alpha| = k$. Daher ist $\lim_{\xi \to 0} \frac{\varphi(\xi)}{\|\xi\|^k} = 0$, das heißt $\varphi(\xi) = o(\|\xi\|^k)$. □

Wir wollen das Ergebnis von Korollar (1.5.5) nun in den Spezialfällen $m = 0, 1, 2$ betrachten.

1. **Beispiel 1.5.6** *Es gibt nur ein* n*-Tupel* $\alpha = (\alpha_1, ..., \alpha_n) \in \mathbb{N}_0^n$ *mit* $|\alpha| = 0$*, nämlich* $(0, ..., 0)$*. Wir erhalten:*

   $$T_0(\xi) = \frac{D^0 f(x)}{0!}\xi_1^0 ... \xi_n^0 = f(x)$$

   Das heißt das 0*-te Taylorpolynom ist die Konstante:*

   $$T_0(\xi) \equiv f(x) \tag{1.29}$$

   Aus (1.5.5) folgt in diesem Fall die Aussage, dass f *bei* x *stetig ist.*

   Approximation durch affin-lineare Abbildung Bei $m = 1$ gilt: Die Tupel $\alpha \in \mathbb{N}_0^n$ mit $|\alpha| = 1$ sind gerade $e_1 = (1, 0, ..., 0), ..., e_n = (0, ..., 0, 1)$. Wir erhalten

   $$T_1(\xi) = f(x) + \sum_{|\alpha|=1} \frac{D^\alpha f(x)}{\alpha!}$$

   Daraus folgt folgende Aussage:

   $$T_1(\xi) = f(x) + \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(x)\xi_i = f(x) + \langle \nabla f(x), \xi \rangle \tag{1.30}$$

3. $m = 2$: Es gibt zwei Arten von n-Tupeln $\alpha \in \mathbb{N}_0^n$, so dass die Summe der Einträge 2 ergibt, nämlich
   - die Vektoren $2e_i = (0, ..., 0, \underbrace{2}_{i-te}, 0, ..., 0)$ für $i \in \{1, ..., n\}$.
   - die Vektoren $e_i + e_j = (0, ..., 0, \underbrace{1}_{i-te}, 0, ..., 0, \underbrace{1}_{j-te}, 0, ..., 0)$ für $i, j \in \{1, ..., n\}, i < j$.

   Also ergibt sich:

$$\begin{aligned} T_2(\xi) &= f(x) + \langle \nabla f(x), \xi \rangle + \sum_{i=1}^{n} \frac{1}{2} \frac{\partial^2 f}{\partial x_i^2}(x)\xi_i^2 + \sum_{1 \leq i < j \leq n} \frac{\partial^2 f}{\partial x_j \partial x_i}(x)\xi_i\xi_j = \\ &= f(x) + \langle \nabla f(x), \xi \rangle + \frac{1}{2} \sum_{i,j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j}(x)\xi_i\xi_j \end{aligned}$$

   Wegen der Vertauschbarkeit der zweiten Ableitungen. Der letzte Term ist eine quadratische Form, so dass man mit Hilfe einer Matrix A mit $(a_{ij}) = \frac{\partial^2 f}{\partial x_i \partial x_j}(x)$ auch schreiben kann:

$$T_2(\xi) = f(x) + \langle \nabla f(x), \xi \rangle + \frac{1}{2}\langle A\xi, \xi \rangle$$

Definition 1.5.7 *Sei $U \subseteq \mathbb{R}^n$ eine offene Teilmenge und $f : U \to \mathbb{R}$ eine zweimal stetig differenzierbare Funktion. Unter der Hesseschen Matrix (oder Hesse-Matrix) versteht man die nxn-Matrix:*

$$((Hessf)(x))_{ij} := \frac{\partial^2 f}{\partial x_i \partial x_j}(x), 1 \leq i, j \leq n$$

Wegen der Vertauschbarkeit der zweiten Ableitungen ist die Hessesche Matrix symmetrisch, das heißt:

$$((Hessf)(x))^T = ((Hessf)(x))$$

Korollar 1.5.8 *Sei $U \subseteq \mathbb{R}^n$ offen, $f : U \to \mathbb{R}$ zweimal stetig differenzierbar und $x \in U$. Dann gilt:*

$$f(x + \xi) = f(x) + \langle \nabla f(x), \xi \rangle + \frac{1}{2}\langle (Hessf)(x)\xi, \xi \rangle + o(\|\xi\|^2)$$

Beweis Wurde in (1.5.5) (3) gezeigt.

Beispiel Für $n = 2$:

$$f(x+\xi) = f(x)+D_1f(x)\xi_1+D_2f(x)\xi_2+\frac{1}{2}D_1D_1f(x)\xi_1^2+D_1D_2f(x)\xi_1\xi_2+\frac{1}{2}D_2D_2f(x)\xi_2^2+o(\|\xi\|^2)$$

Definition 1.5.9 *Sei $U \subseteq \mathbb{R}^n$ offen und $f : U \to \mathbb{R}$. Ein Punkt $x \in U$ heißt* lokales Maximum *(bzw. lokales Minimum) von f, falls es eine Umgebung $V \subseteq U$ von x gibt $f(y) \leq f(x)$ (bzw. $f(y) \geq f(x)$) $\forall y \in V$.*
Ein Punkt $x \in U$ heißt isoliertes lokales Maximum*, falls es eine Umgebung $V \subseteq U$ von x gibt, so dass $f(y) < f(x) \forall y \in V \setminus \{x\}$. Ein* lokales Extremum *ist ein lokales Maximum oder Minimum.*

Unmittelbar aus Lemma 5.9, Analysis I Skript von Prof. Bunke, folgt:

Satz 1.5.10 *Sei $U \subseteq \mathbb{R}^n$ offen und besitze die partiell differenzierbare Funktion $f : U \to \mathbb{R}$ bei x ein lokales Extrema. Dann gilt:*

$$\nabla f(x) = 0$$

Beweis Betrachte $g_i(t) := f(x+te_i)$. Da U offen ist, gibt es ein Intervall $(-\varepsilon, \varepsilon)$ auf dem g_i definiert ist. Die Funktion g_i hat bei $0 \in (-\varepsilon, \varepsilon)$ ein lokales Extremum also gilt: $0 = g_i'(0) = \frac{\partial f}{\partial x_i}(x)$. Also verschwinden alle partiellen Ableitungen von f bei x und es folgt $\nabla f(x) = 0$. □

Definition 1.5.11 *Sei $U \subseteq \mathbb{R}^n$ offen, $f : U \to \mathbb{R}$ partiell differenzierbar. Die Punkte $x \in U$ mit $0 = \nabla f(x)$ nennen wir kritische Punkte der Funktion f.*

Wir wollen uns überlegen, wie man lokale Extrema anhand der Taylorformel bestimmen kann.

Lemma 1.5.12 *Sei $U \subseteq R^n$ offen, $f : U \to \mathbb{R}$ zweimal stetig differenzierbar. Sei $x \in U$ ein kritischer Punkt ($\nabla f(x) = 0$) und $\xi \in \mathbb{R}^n$, so dass $\langle (Hessf)(x)\xi, \xi\rangle > 0$. Dann gibt es ein $\delta > 0$, so dass $f(x + t\xi) > f(x)$ für alle $t \in (-\delta, \delta)\setminus\{0\}$.*

Beweis Es gilt nach (1.5.8):

$$f(x + t\xi) = f(x) + \frac{t^2}{2}\langle (Hessf)(x)\xi, \xi\rangle + \varphi(t)$$

mit $\varphi(t) = o(\|t\|^2)$, das heißt $\lim_{t\to 0} \frac{\varphi(t)}{t^2} = 0$. Also gibt es ein $\delta > 0$, so dass $|\varphi(t)| < \frac{\lambda}{2}t^2$ für $t \in (-\delta, \delta)\setminus\{0\}, \lambda := \langle (Hessf)(x)\xi, \xi\rangle$.

$$f(x + t\xi) = f(x) + \frac{t^2}{2}\lambda + \varphi(t) > f(x)$$

□

Wiederholung Eine symmetrische reelle nxn-Matrix A heißt:

- positiv definit, falls $\langle A\xi, \xi\rangle > 0 \forall \xi \in \mathbb{R}^n \setminus \{0\}$
- negativ definit, falls $\langle A\xi, \xi\rangle < 0 \forall \xi \in \mathbb{R}^n \setminus \{0\}$
- positiv semidefinit, falls $\langle A\xi, \xi\rangle \geq 0 \forall \xi \in \mathbb{R}^n$
- negativ semidefinit, falls $\langle A\xi, \xi\rangle \leq 0 \forall \xi \in \mathbb{R}^n$
- indefinit, falls es $\xi, \eta \in \mathbb{R}^n$ gibt mit $\langle A\xi, \xi\rangle > 0$ und $\langle A\eta, \eta\rangle < 0$.

Aus dem Lemma folgt:

Satz 1.5.13 *Sei $f : U \to \mathbb{R}$ zweimal stetig differenzierbar, $U \subseteq \mathbb{R}^n$ offen und haben f bei $x \in U$ ein lokales Maximum (bzw. Minimum). Dann ist $((Hessf)(x))$ negativ semidefinit (bzw. positiv semidefinit), das heißt es gilt:*

$$\langle (Hessf)(x)\xi, \xi\rangle \leq 0 (\geq 0) \forall \xi \in \mathbb{R}^n$$

Beweis Habe f bei x ein lokales Maximum. Gäbe es ein $\xi \in \mathbb{R}^n$ mit $\langle (Hessf)(x)\xi, \xi\rangle > 0$, so würde aus (1.5.12) ein Widerspruch folgen. □

Dass die Hessematrix an einem kritischen Punkt positiv (bzw. negativ) semidefinit ist, ist jedoch nicht hinreichend dafür, dass die Funktion dort ein Extremum annimmt.

Beispiel 1.5.14

$$\begin{aligned} f_1(x,y) &= x^2 + y^2 \\ f_2(x,y) &= x^2 \\ f_3(x,y) &= x^2 + y^3 \end{aligned}$$

Für alle 3 Funktionen verschwindet der Gradient im Nullpunkt und es gilt:

$$(Hessf_k)(0,0) = \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix}, k = 1,2,3$$

Die Hessematrix ist also positiv semidefinit, denn:

$$\langle (Hessf_k)(0,0), \xi, \xi\rangle = 2\xi_1^2 \geq 0, \forall \xi = (\xi_1, \xi_2) \in \mathbb{R}^2$$

Aber

- f_1 hat bei 0 ein isoliertes lokales Minimum.
- f_2 hat ein lokales Minimum bei $(0, y)$ - ist also nicht isoliert, da f_2 nicht von y abhängt, das heißt Punkte auf y-Achse haben selben Funktionswerte.
- f_3 hat bei 0 weder ein lokales Minimum noch ein lokales Maximum.

Gilt jedoch die stärkere Aussage, dass die Hessematrix positiv (bzw. negativ) definit ist und einen kritischen Punkt hat, dann folgt sogar die Existenz eines isolierten lokalen Extremums.

Satz 1.5.15 *Sei $U \subset \mathbb{R}^n$ offen, $f : U \to \mathbb{R}$ zweimal stetig differenzierbar. Habe f bei x einen kritischen Punkt und sei Hessematrix $(Hessf)(x)$ bei x positiv (bzw. negativ) definit. Dann hat f bei x ein isoliertes lokales Minimum (bzw. Maximum).*

Beweis Sei $S = \{\xi \in \mathbb{R}^n \mid \|\xi\| = 1\}$ die Einheitssphäre um Null in $\mathbb{R}^n$. Da S kompakt ist, nimmt die stetige Funktion $\xi \mapsto \langle A\xi, \xi\rangle$ für jede Matrix $A \in Mat(n, n; \mathbb{R})$ ihr Maximum an. Speziell für $A = (Hessf)(x)$ gilt, dass dieses Minimum μ positiv ist, da $(Hessf)(x)$ als positiv definit vorrausgesetzt wurde. Es gilt:

$$f(x + \xi) = f(x) + \frac{1}{2}\langle A\xi, \xi\rangle + \varphi(\xi)$$

mit $\varphi(\xi) = o(\|\xi\|^2)$. Wähle nun $\delta > 0$, so dass

$$|\varphi(\xi)| \leq \frac{\mu}{4}\|\xi\|^2$$

für $\|\xi\| < \delta$.

$$f(x + \xi) = f(x) + \frac{1}{2}\underbrace{\langle A\frac{\xi}{\|\xi\|}, \frac{\xi}{\|\xi\|}\rangle}_{\geq \mu}\|\xi\|^2 + \varphi(\xi) \geq f(x) + \frac{1}{4}\mu\|\xi\|^2 > f(x)$$

Bemerkung 1.5.16 *Sei $A \in Mat(n, n; \mathbb{R})$ eine symmetrische reelle nxn-Matrix. Dann gibt es eine Orthonormalbasis $v_1, ..., v_n \in \mathbb{R}^n$ aus Eigenvektoren zu reellen Eigenwerten $\lambda_1, ..., \lambda_n \in \mathbb{R}$. Stellen wir einen Vektor $\xi \in \mathbb{R}^n$ bezüglich dieser Basis dar als*

$$\xi = \xi_1 v_1 + ...\xi_n v_n$$

so folgt

$$\langle \xi, A\xi\rangle = \langle \xi_1 v_1 + ... + \xi_n v_n, \lambda_1\xi_1 v_1 + ... + \lambda_n \xi_n v_n\rangle = \sum_{i=1}^{n} \lambda_i \xi_i^2$$

Daran sieht man, dass A genau dann

- *positiv definit ist, wenn* $\lambda_1, ..., \lambda_n > 0$
- *negativ definit ist, wenn* $\lambda_1, ..., \lambda_n < 0$
- *positiv semidefinit ist, wenn* $\lambda_1, ..., \lambda_n \geq 0$
- *negativ semidefinit ist, wenn* $\lambda_1, ..., \lambda_n \leq 0$
- *indefinit ist, falls es einen positiven und negativen Eigenwert gibt*

Für $n > 2$ sind die Eigenwerte aber im Allgemeinen schwierig zu bestimmen. Deshalb wollen wir noch ein einfacheres Kriterium für die Definitheit einer symmetrischen Matrix wiederholen.

Satz 1.5.17 (Hurwitz) *Sei* $A \in Mat(n, n; \mathbb{R})$ *eine reelle symmetrische Matrix. Dann ist A genau dann positiv definit, wenn gilt:*
Die Determinanten $\det \begin{pmatrix} a_{11} & \dots & a_{1k} \\ \vdots & & \vdots \\ a_{k1} & \dots & a_{kk} \end{pmatrix}$ *sind für* $k = 1, ..., n$ *alle positiv. Da A genau dann negativ definit ist, wenn* $-A$ *positiv ist, folgt sofort:*
A ist genau dann negativ definit wenn die Determinanten positiv sind für gerade k und negativ für ungerade.

$$\det \begin{pmatrix} -a_{11} & \dots & -a_{1k} \\ \vdots & & \vdots \\ -a_{k1} & \dots & -a_{kk} \end{pmatrix} = (-1)^k \det \begin{pmatrix} a_{11} & \dots & a_{1k} \\ \vdots & & \vdots \\ a_{k1} & \dots & a_{kk} \end{pmatrix}$$

Warnung Ersetzt man „definit" durch „semidefinit" in (1.5.17), so gilt nur noch die „$\Rightarrow$" Richtung. Denn zum Beispiel für die Matrix

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

sind die drei zu bestimmenden Determinaten alle ≥ 0, aber für $\xi = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$ folgt $\langle A\xi, \xi \rangle = -1$. A ist nicht positiv semidefinit (sondern indefinit!).

1.6 Banachscher Fixpunktsatz, implitzite Funktionen und Umkehrsatz

Definition 1.6.1 *Eine Abbildung* $f : X \to Y$ *zwischen metrischen Räumen heißt* Lipschitz-stetig*, falls es eine Konstante* $C \in \mathbb{R}$ *gibt mit* $d_y(f(x), f(y)) \leq$

$C \cdot d_x(x, y)$ für alle $x, y \in X$, wobei d_x, d_y die Abstandsfunktionen auf X bzw. Y bezeichnet. Eine Abbildung heißt kontrahierend oder Kontraktion, falls die obige Bedingung für ein $C < 1$ erfüllt ist.

Satz 1.6.2 (Banachscher Fixpunktsatz, Kontraktionsprinzip) *Sei X ein vollständiger metrischer Raum und $f : X \to X$ eine Kontraktion. Sei $x_0 \in X$ beliebig als Startwert und die Folge $(x_n)_{n \in \mathbb{N}}$ induktiv definiert durch $x_{n+1} := f(x_n) \forall n \in \mathbb{N}$. Dann hat f genau einen Fixpunkt, das heißt es gibt genau ein $x \in X$ mit $f(x) = x$ und die Folge x_n konvergiert gegen x.*

Beweis Wir zeigen zuerst, dass es höchstens einen Fixpunkt geben kann. Seien $x, x' \in X$ mit $f(x) = x, f(x') = x', x \neq x'$. Dann folgt $d(f(x), f(y)) \leq \underbrace{C}_{<1} d(x, x')$. Dies ist ein Widerspruch!
Wir zeigen nun, dass x_n eine Cauchyfolge ist. Mehrfaches Anwenden der Dreiecksungleichung liefert:

$$\begin{aligned} d(x_n, x_{n+k}) &\leq d(x_n, x_{n+1}) + d(x_{n+1}, x_{n+k}) \leq \\ &\leq d(x_n, x_{n+1}) + d(x_{n+1}, x_{n+2}) + ... + d(x_{n+k-1}, x_{n+k}) \end{aligned}$$

Wir wenden nun die Kontraktionseigenschaft an, um die einzelnen Summanden weiter abzuschätzen:

$$d(x_n, x_{n+1}) = d(f(x_{n-1}), f(x_n)) \leq C \cdot d(x_{n-1}, x_n) \leq ... \leq C^n d(x_0, x_1)$$

Also gilt $d(x_n, x_{n+k}) \leq \sum_{i=0}^{k-1} C^{n+i} d(x_0, x_1)$. Da die geometrische Reihe $\sum_k C^k$ für $C < 1$ konvergiert, folgt die Cauchy Eigenschaft für x_n. Die Folge x_n hat also einen Limes $x \in X$ und wegen

$$x = \lim_{n \to \infty} x_n = \lim_{n \to \infty} f(x_n) = f(\lim_{n \to \infty} x_n) = f(x)$$

ist dieser ein Fixpunkt von f - wir wir gesehen haben, der einzige. □

Bemerkung 1.6.3 *Aus dem Beweis folgt außerdem die Fehlerabschätzung*

$$d(x_n, x) \leq \sum_{k=0}^{\infty} C^k d(x_0, x_1) = (\frac{1}{1-C} - \frac{1-C^n}{1-C}) d(x_0, x_1) = \frac{C^n}{1-C} d(x_0, x_1)$$

Beispiel 1.6.4 *Heronverfahren zur Näherungsweisen Berechnung der Quadratwurzel. Statt $x^2 = 2$ für positive x zu lösen, suchen wir eine Lösung der dazu äquivalenten Fixpunktgleichung $2x = x + \frac{2}{x} \Leftrightarrow x = \frac{x}{2} + \frac{1}{x}$. Also definieren wir die Funktion $f(x) = \frac{x}{2} + \frac{1}{x}$. Nach dem Mittelwertsatz gilt:*

$$d(f(x), f(y)) = |f(x) - f(y)| = |f'(\xi)| \cdot \underbrace{|x - y|}_{d(x,y)}$$

Für ein ξ zwischen x und y. Wir berechnen

$$|f'(x)| = |\frac{1}{2} - \frac{1}{x^2}| \leq \frac{1}{2}, x \geq 1$$

Kontraktion im Intervall $[1, \infty)$ mit $C = \frac{1}{2}$. Setze nun $x_0 = 1$. Berechnung von $x_1 = \frac{3}{2}, x_2 = \frac{17}{12}$ sorgt bei 2. Schritt schon für sehr genaue Annährung an $\sqrt{2}$ (mit Ergebnis $x_2 = 1,416$).

Implizit definierte Funktionen Unter einer *implizit definierten Funktion* versteht man eine Abbildung, die nicht durch eine *explizite* Abbildungsvorschrift, wie zum Beispiel $f(x) = 3x^2 + 5, y(t) = \begin{pmatrix} t^2 \\ \cos t \end{pmatrix}, ...$ gegeben ist, sondern durch Gleichungen, wie zum Beispiel:

$$f(x) \arctan(f(x) + x) = x^2 + x, f(x)^3 + x^5 = 1, ...$$

Wir wollen im Folgenden eine Bedingung kennenlernen, unter der eine eindeutige Lösung für den Funktionswert $f(x)$ existiert.

Beispiel 1.6.5 *Sei $f : \mathbb{R}^2 \to \mathbb{R}$ eine stetig differenzierbare Funktion auf dem $\mathbb{R}^2$. Eine Funktion $g : \mathbb{R} \to \mathbb{R}$ könnte dann implizit durch eine Gleichung*

$$f(t, g(t)) = 0$$

gegeben sein. Anders ausgedrückt: Der Graph von g ist eine Höhenlinie der Funktion f (nämlich zum Wert $f = 0$).

Es geht also um die Frage, unter welchen Bedingungen die Höhenlinie $f^{-1}(\{0\})$ einer Funktion $f : \mathbb{R}^2 \to \mathbb{R}$ als Graph $\{(t, g(t)) \mid t \in \mathbb{R}\}$ einer Funktion $g : \mathbb{R} \to \mathbb{R}$ darstellt werden kann.
Betrachten wir ein Beispiel. Sei $c > 0$ und $f : \mathbb{R}^2 \to \mathbb{R}$ mit

$$f(x, y) = x^2 + y^2 - c$$

Dann ist $f^{-1}(\{0\}) = \{(x, y) \in \mathbb{R}^2, f(x, y) = 0\}$ der Kreis um 0 mit Radius $\sqrt{c}$. Offensichtlich gibt es *keine* Funktion $g : \mathbb{R} \to \mathbb{R}$, so dass $f^{-1}(\{0\})$ der Graph von g wäre, denn für $|x| < \sqrt{x}$ gibt es jeweils 2 Lösungen $y = \pm\sqrt{c - x^2}$ und für $|x| > \sqrt{c}$ keine. Schränken wir die Funktion f aber geeignet ein, zum Beispiel auf die Menge $U := \{(x, y) \in \mathbb{R}^2 \mid |x| < \sqrt{c}, y > 0\}$, dann gibt es eine Funktion $g : (-\sqrt{c}, \sqrt{c}) \to \mathbb{R}_+$ mit $f(x, g(x)) = 0$, nämlich $g(x) = \sqrt{c - x^2}$. Wir wollen ein Kriterium für die lokale Existenz einer solchen Funktion beweisen.

Beispiel 1.6.6 *Sei $U \subseteq \mathbb{R}^2$ offen und $F : U \to \mathbb{R}$ eine differenzierbare Funktion. Sei $g : I \to \mathbb{R}$ eine auf dem Intervall $I \subseteq \mathbb{R}$ definierte differenzierbare Funktion,*

so dass der Graph von g in U enthalten ist und es gilt: $F(x, g(x)) = 0$.
Durch differenzieren dieser Gleichung mit Hilfe der Kettenregel erhalten wir:

$$D_1F(x, g(x)) + D_2F(x, g(x))g'(x) = 0$$

Unter der Vorraussetzung, dass $D_2F(x, g(x)) \neq 0$ ist, folgt:

$$g'(x) = -\frac{D_1F(x, g(x))}{D_2F(x, g(x))}$$

Für die Ableitung der implizit definierten Funktion g. Wie der folgende Satz zeigt, ist die Bedingung $D_2F(x, g(x)) \neq 0$ sogar hinreichend für die lokale Existenz einer impliziten Funktion. Wir formulieren die Aussage für beliebig viele Dimensionen.

Satz 1.6.7 (Über implizite Funktionen) *Seien $U_1 \subseteq \mathbb{R}^k$ und $U_2 \subseteq \mathbb{R}^m$ offene Mengen und*

$$F : U_1 \times U_2 \to \mathbb{R}^m$$
$$(x, y) \mapsto F(x, y)$$

eine stetig differenzierbare Abbildung. Sei $(a, b) \in U_1 \times U_2$ ein Punkt, so dass $F(a, b) = 0$ und die mxm-Matrix $\frac{\partial F}{\partial y}(a, b)$ invertierbar ist. Dann gibt es offene Umgebungen $V_1 \subseteq U_1$ von a und $V_2 \subseteq U_2$ von b und eine stetig differenzierbare Abbildung $g : V_1 \to V_2$, so dass gilt:

1. *$F(x, g(x)) = 0 \forall x \in V_1$*

2. *Ist $(x, y) \in V_1 \times V_2$ ein Punkt mit $F(x, y) = 0$, so folgt $y = g(x)$.*

Beweis

1. Wir können ohne Einschränkung $(a, b) = (0, 0)$ annehmen. Setze $B := \frac{\partial F}{\partial y}(0, 0) \in GL(m, \mathbb{R})$. Definiere nun:

   $$G : U_1 \times U_2 \to \mathbb{R}^m, G(x, y) := y - B^{-1}F(x, y)$$

   Es gilt:
   $$\frac{\partial G}{\partial y}(x, y) = 1_m - B^{-1}\frac{\partial F}{\partial y}(x, y)$$

   Daraus folgt:
   $$\frac{\partial G}{\partial y}(0, 0) = 0$$

   Da alle Komponenten der Matrix $\frac{\partial G}{\partial y}$ stetig von (x, y) abhängen, gibt es eine Umgebung der Null $W_1 \subseteq U_1$ und $W_2 \subseteq U_2$, so dass $\|\frac{\partial G}{\partial y}\| \leq \frac{1}{2}$ für alle $(x, y) \in W_1 \times W_2$. Wähle ein $r > 0$, so dass:

   $$\tilde{V_2} := \{y \in \mathbb{R}^n \mid \|y\| < r\} \subseteq W_2$$

Wegen der Stetigkeit von G und da $G(0,0) = 0$ ist, gibt es eine offene Nullumgebung $\tilde{V}_1 \subseteq W_1$, so dass

$$\varepsilon := \sup_{x \in V_1} \|G(x,0)\| < \frac{r}{2}$$

2. Wir zeigen jetzt, dass es genau eine stetige Abbildung $g : \tilde{V}_1 \to \tilde{V}_2$ gibt, so dass $F(x, g(x)) = 0 \forall x \in \tilde{V}_1 \Leftrightarrow \underbrace{g(x) = G(x, g(x))}_{\text{Fixpunktgleichung}}$ für alle $x \in \tilde{V}_1$ gilt.
   Nun haben wir die Gleichung für das gesuchte $g(x)$ also als Fixpunktgleichung (für die Abbildung $y \mapsto G(x,y)$) formuliert. Aus der obigen Abschätzigen für die Norm von $\frac{\partial G}{\partial y}$ und Korollar (1.4.18) folgt, dass die Abbildung $y \mapsto G(x,y)$ für jedes $x \in W_1$ eingeschränkt auf $\tilde{V}_2$ Lipschitz-Stetig mit Konstante $C = \frac{1}{2}$ ist.
   Wir betrachten un die Folge $g_0(x) := 0, g_{n+1}(x) := G(x, g_n(x)), x \in \tilde{V}_1$. Wir zeigen nun, dass diese Folge den offenen Ball $\tilde{V}_2$ nicht verläßt. Es gilt $\|g_1(x)\| < \varepsilon$. Induktiv zeigt man, dass

$$\|g_{n+1}(x) - g_n(x)\| \leq \frac{1}{2^n}\varepsilon$$

   ist. Mit $g_N(x) = \sum_{i=0}^{N-1}(g_{i+1}(x) - g_i(x))$ folgt:

$$g_N(x)\| \leq \sum_{i=0}^{N-1} \frac{1}{2^2}\varepsilon \leq 2\varepsilon < r$$

   Wie im Beweis des Banachschen Fixpunktsatzes sieht man nun, dass die Folge $g_n(x)$ gegen den einzigen Fixpunkt von $y \mapsto G(x,y)$ auf $\tilde{V}_2$ konvergiert. Aus der Fehlerabschätzung (1.6.3) folgt, dass die Folge sogar gleichmäßig konvergiert und somit nach Lemma 4.67, Analysis I Skript von Prof. Bunke, die Abbildung $x \mapsto g(x)$ stetig ist.

3. Noch zu zeigen: $g : \tilde{V}_1 \to \tilde{V}_2$ stetig differenzierbar. Setze dazu: $A := \frac{\partial F}{\partial x}(0,0) \in Mat(m,m;\mathbb{R})$. Aus der Differenzierbarkeit von F bei $(0,0)$ haben wir

$$F(x,y) = Ax + By + \varphi(x,y)$$

   mit $\varphi(x,y) = o(\|(x,y)\|)$ wegen $F(x,g(x)) = 0 \forall x \in \tilde{V}_1$ erhalten wir daraus

$$g(x) = -B^{-1}Ax - B^{-1}\varphi(x, g(x)) \tag{1.31}$$

   Zeige: Es gibt Konstanten $0 < \delta$ und $0 < K$, so dass $\overline{B_\delta(0)} \subseteq \tilde{V}_1$ (Abschluss des δ-Balls) und $\|g(x)\| \leq K\|x\| \forall \|x\| < \delta$.

Bewies hierzu $C_1 := \|B^{-1}A\|, C_2 := \|B^{-1}\|$ - Wegen $\varphi(x,y) = o(\|(x,y)\|)$ gibt es Konstante $\delta_1, \delta_2 > 0$, so dass $\|\varphi(x,y)\| \leq \frac{1}{2C_2}\|(x,y)\| \leq \frac{1}{2C_2}(\|x\| + \|y\|)$ für $\|x\| \leq \delta_1, \|y\| \leq \delta_2$. Da g stetig ist, gibt es $0 < \delta \leq \delta_1$, so dass

$$\|g(x)\| \leq \delta_2, \|x\| \leq \delta$$

Deshalb gilt:

$$\|\varphi(x, g(x))\| \leq \frac{1}{2C_2}(\|x\| + \|g(x)\|)$$

für $\|x\| \leq \delta$. Aus (1.31) erhalten wir nun

$$\|g(x)\| \leq C_1\|x\| + C_2\|\varphi(x, g(x))\| \leq C_1\|x\| + \frac{1}{2}(\|x\| + \|g(x)\|)$$

$$\Rightarrow \|g(x)\| \leq \underbrace{2(C_1 + \frac{1}{2})}_{=:K}\|x\|$$

Nun können wir (1.31) benutzen, um die Differenzierbarkeit von g bei 0 zu zeigen. Setze $\psi(x) := -B^{-1}\varphi(x, g(x))$. Dann hat (1.31) die Form

$$g(x) = -B^{-1}Ax + \psi(x)$$

Nun folgt aus $\varphi(x,y) = o(\|(x,y)\|)$ und $\|g(x)\| \leq K\|x\|$ für kleine x, dass $\psi(x) = o(\|x\|)$. Dies zeigt, dass g bei 0 differenzierbar ist mit Ableitung:

$$Dg(0) = -(\frac{\partial F}{\partial y}(0,0))^{-1}\frac{\partial F}{\partial x}(0,0) = -B^{-1}A$$

Im letzten Schritt wenden wir die bis jetzt gezeigte Aussage auf alle Punkte $(x, g(x))$ in einer kleinen Umgebung von $(0,0) = (0, g(x))$ an. Die Bedingung, dass $\frac{\partial F}{\partial y}(x,y)$ invertierbar ist, gilt nämlich (wegen der Stetigkeit der Determinante). In einer ganzen Umgebung von $(0,0)$ und es folgt, dass (in eventuell verkleinerten Umgebung $V_1 \subseteq \tilde{V}_1, V_2 \subseteq \tilde{V_2}$) mit $g : V_1 \to V_2$ differenzierbar ist. Da $Dg(x) = -(\frac{\partial F}{\partial x}(x, g(x)))^{-1}(\frac{\partial F}{\partial y}(x, g(x)))$ für alle $x \in V_1$ gilt, g stetig ist und F stetig differenzierbar ist.

$\square$

Beispiel 1.6.8 *Betrachte*

$$F : \mathbb{R}^4 \to \mathbb{R}^2, (x,y,u,v) \mapsto (xy + u\cos(u), (1-y)v + x\sin(u))$$

Diese Abbildung bildet den Punkt $(\underbrace{1,0}_{a}, \underbrace{0,0}_{b})$ *auf* $(0,0)$ *ab. Es gilt*

$$\frac{\partial F}{\partial(u,v)}(x,y,u,v) = \begin{pmatrix} \cos(u) - u\sin(u) & 0 \\ x\cos(u) & 1-y \end{pmatrix}$$

Somit ist $\frac{\partial F}{\partial(u,v)}(1,0,0,0) = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$ *invertierbar.*

Nach Satz (1.6.7) existiert also offene Umgebung V_1 von $(1,0) \in \mathbb{R}^2$, V_2 von $(0,0) \in \mathbb{R}^2$ und eine stetig differenzierbare Abbildung

$$g : V_1 \to V_2, (x,y) \mapsto (g_1(x,y), g_2(x,y))$$

so dass $F(x,y,g_1(x,y),g_2(x,y)) = (0,0)$ für alle $(x,y) \in V_1$. Außerdem gilt für jedes $(x,y) \in V_1$ und $(u,v) \in V_2$ mit $F(x,y,u,v) = 0$, so dass

$$u = g_1(x,y), v = g_2(x,y)$$

Beispiel 1.6.9 *Betrachten wir noch einmal Beispiel (1.6.6), den Kreis mit Radius $\sqrt{c}$ um 0 als Nullstellengebilde der Funktion $f(x,y) = x^2 + y^2 - c, c > 0$ auf $\mathbb{R}^2$. Da $\nabla f(x,y) = (2x, 2y)$ nirgends auf S (Kreisbahn) mit $S = f^{-1}(\{0\}) = \{(x,y) \in \mathbb{R}^2 \mid f(x,y) = 0\}$ verschwindet, können wir S durch eine Umgebung jedes Punktes auf S als Graph einer Funktion darstellen und zwar entweder als Graph einer Funktion von x oder eine Funktion von y.*

$$\begin{aligned} S &= \{(x,y) \in \mathbb{R}^2 \mid -\sqrt{c} < x < \sqrt{c}, y = \sqrt{c - x^2}\} \cup \\ &\cup \{(x,y) \in \mathbb{R}^2 \mid -\sqrt{c} < x < \sqrt{c}, y = -\sqrt{c - x^2}\} \cup \\ &\cup \{(x,y) \in \mathbb{R}^2 \mid -\sqrt{c} < y < \sqrt{c}, x = \sqrt{c - x^2}\} \cup \\ &\cup \{(x,y) \in \mathbb{R}^2 \mid -\sqrt{c} < y < \sqrt{c}, x = -\sqrt{c - x^2}\} \cup \end{aligned}$$

Umkehrabbildungen Seien U_1, U_2 offene Mengen des $\mathbb{R}^n$ und $f : U_1 \to U_2$ eine stetig differenzierbare Abbildung. Wir wollen uns nun mit der frage beschäftigen unter welchen Bedingungen f ein C^1-Diffeomorphismus ist, das heißt wann f eine stetig differenzierbare Umkehrabbildung besitzt. Eine notwendige Bedingung dafür haben wir schon im Beispiel (1.4.9) hergeleitet. Sei $a \in U_1, b = f(a)$. Aus $g \circ f = Id_{U_1}$ mit der Kettenregel $Dg(b)Df(a) = 1$ - die Einheitsmatrix. Also muss die Funktionalmatrix $Df(a)$ invertierbar sein. Der folgende Satz zeigt, dass diese Bedingung auch hinreichend ist, wenn man die offenen Mengen U_1, U_2 eventuell verkleinert.

Satz 1.6.10 (Umkehrsatz) *Sei $U \subset \mathbb{R}^n$ offen und $f : U \to \mathbb{R}^n$ eine stetig differenzierbare Abbildung. Sei $a \in U, b := f(a)$. Die Jacobi-Matrix $Df(a)$ sei invertierbar. Dann gibt es offene Umgebungen U_1 von a und U_2 von b, so dass f die Mengen U_1 bijektiv auf U_2 abbildet und so, dass die Umkehrabbildung*

$$g := (f_{|U_1})^{-1} : U_2 \to U_1$$

stetig differenzierbar ist.

Beweis Da eine Umehrabbildung implizit durch die Gleichung

$$x - f(g(x)) = 0$$

definiert ist können wir den Impliziten Funktionensatz (1.6.7) anwenden. Betrachten $F : \mathbb{R}^n \times U \to \mathbb{R}^n, (x, y) \mapsto x - f(y)$. Es gilt $F(b, a) = b - f(a) = 0$. Satz (1.6.7) ist anwendbar, da $\frac{\partial F}{\partial y}(b, a) = -Df(a)$ nach Vorraussetzung invertierbar ist. Es gibt also offene Umgebungen V_1 von b und V_2 von a und eine stetig differenzierbare Abbildung

$$g : V_1 \to V_2, 0 = F(x, g(x)) = x - f(g(x))$$

Setze $U_2 := V_2 \cap f^{-1}(V_1) = \{y \in V_2 \mid f(y) \in V_1\}$. Dann ist U_1 offen und wird durch f bijektiv auf $U_2 := V_1$ abgebildet. □

Beispiel 1.6.11 *Polarkoordinaten in* $\mathbb{R}^2$. *Wir betrachten die Abbildung*

$$f : (0, \infty) \times \mathbb{R} \to \mathbb{R}^2, (r, \varphi) \mapsto (r \cos \varphi, r \sin \varphi)$$

Wir berechnen die Jacobi-Matrix:

$$Df(r, \varphi) = \begin{pmatrix} \cos \varphi & -r \sin \varphi \\ \sin \varphi & r \cos \varphi \end{pmatrix}$$

Um die Invertierbarkeit zu überprüfen wird die Determinante der Jacobi-Matrix berechnet:

$$\det(Df(r, \varphi)) = r(\cos^2 \varphi + \sin^2 \varphi) = r > 0$$

Die Jacobi-Matrix ist also überall invertierbar. Die Abbildung f ist aber nicht injektiv, denn für $r > 0, \varphi \in \mathbb{R}$ werden die Punkte $\{(r, \varphi + 2\pi k) \mid k \in \mathbb{Z}\}$ alle auf denselben Wert abgebildet.
Man kann f zu einem globalen Diffeomorphismus machen, indem man es auf eine geeignete Menge einschränkt, zum Beispiel:

$$U_1 = (0, \infty) \times (-\frac{\pi}{2}, \frac{\pi}{2}), U_2 = f(U_1)$$

In diesem Fall kann man die Umkehrabbildung sogar explizit angeben:

$$g : U_2 \to U_1, (x, y) \mapsto (\sqrt{x^2 + y^2}, \arctan(\frac{y}{x}))$$

$\Rightarrow f$ ist ein lokaler, aber kein globaler - Diffeomorphismus.

Extrema unter Nebenbedingungen Wir betrachten folgendes Problem: Auf einer offenen Teilmenge des $\mathbb{R}^n$ seien zwei Funktionen f und h gegeben. Wir wollen lokale Extrema der Funktion h, eingeschränkt auf die Nullstellenmenge

$$M := \{x \in U \mid f(x) = 0\}$$

der Funktion f finden. Dabei reicht es die Punkte zu betrachten, an denen der Gradient von h verschwindet, wie folgendes Beispiel zeigt:
Sei $U = \mathbb{R}^2, f : \mathbb{R}^2 \to \mathbb{R}$ gegeben durch $f(x, y) = x^2 + y^2 - c, c > 0$. M ist also der Kreis um 0 mit Radius $\sqrt{c}$, h sei „Höhenlinienfunktion" $h(x, y) = y$. Offensichtich nimmt $h|_M$ sein Maximum bei $(0, \sqrt{c})$ und sein Minimum bei $(0, -\sqrt{c})$ an. Der Gradient von h, $\nabla h(x, y) = (0, 1)$. Der folgende Satz gibt ein Kriterium für das Auffinden von Extrema unter Nebenbedingungen.

Satz 1.6.12 *Sei $U \subseteq \mathbb{R}^n$ offen, $f : U \to \mathbb{R}$ eine stetig differenzierbare Funktion und*

$$M := \{x \in U \mid f(x) = 0\}$$

die Nullstellenmenge von f. Sei $a \in M$ ein Punkt mit $\nabla f(a) \neq 0$. Weiter sei $h : U \to \mathbb{R}$ eine stetig differenzierbare Funktion, die im Punkt a ein lokales Maximum (bzw. Minimum) unter der Nebenbedingung $f = 0$ besitze, das heißt es existiert eine Umgebung $V \subseteq U$ mit $h(a) \geq (\leq) h(x)$ für alle $x \in M \cap V$. Dann existiert ein $\lambda \in \mathbb{R}$, so dass gilt:

$$\nabla h(a) = \lambda \nabla f(a)$$

Bemerkung Man nennt λ einen Langrangeschen-Multiplikator. Im obigen Beispiel gibt es ein solches λ genau in den Punkten $(0, \pm\sqrt{c})$, denn $\nabla f(x, y) = (2x, 2y)$ und $\nabla h(x, y) \equiv (0, 1)$.
Es gilt $\nabla h(0, \pm\sqrt{c}) = \lambda \nabla f(0, \pm\sqrt{c}), \lambda = \pm\frac{1}{2}$.

Beweis Da $\nabla f(a) \neq 0$, können wir nach einer eventuellen Umnummerierung der Koordinaten annehmen, dass $\frac{\partial f}{\partial x_n}(a) \neq 0$ gilt. Nach dem impliziten Funktionensatz (1.6.7) gibt es also eine offene Umgebung $V_1 \subseteq \mathbb{R}^{n-1}$ von $(a_1, ..., a_{n-1})$, eine offene Umgebung $V_2 \subseteq \mathbb{R}$ von a_n mit $V_1 \times V_2 \subseteq U$, sowie einer stetig differenzierbare Funktion $g : V_1 \to V_2$, so dass gilt:

$$M \cap (V_1 \times V_2) = \{(x_1, .., x_n) \in V_1 \times V_2 \mid x_n = g(x_1, ..., x_{n-1})\}$$

Da also $0 = f(x_1, ..., x_{n-1}, g(x_1, ..., x_{n-1}))$ für $(x_1, ..., x_{n-1}) \in V_1$ gilt, folgt aus der Kettenregel:

$$0 = \frac{\partial f}{\partial x_i}(a) + \frac{\partial f}{\partial x_n}(a) \frac{\partial g}{\partial x_i}(a_1, ..., a_{n-1}), i = 1, ..., n-1 \qquad (1.32)$$

Statt h betrachten wir jetzt die Funktion:

$$H : V_1 \to \mathbb{R}, H(x_1, ..x_{n-1}) := h(x_1, ..., x_{n-1}, g(x_1, ..., x_{n-1}))$$

Da h im Punkt a ein lokales Extremum unter der Nebenbedingung $f = 0$ hat, gibt es eine Umgebung $V \subseteq V_1$ von $(a_1, ..., a_{n-1})$, so dass $H(a_1, ..., a_{n-1}) \geq H(x_1, ..., x_{n-1}), \forall x \in V_1$ gilt. Somit besitzt H ein lokales Extremum bei $(a_1, ..., a_{n-1})$ und es folgt:

$$\frac{\partial H}{\partial x_i}(a_1, ..., a_{n-1}) = 0, i = 1, ..., n-1$$

Und mit der Kettenregel:

$$0 = \frac{\partial h}{\partial x_i}(a) + \frac{\partial h}{\partial x_n}(a)\frac{\partial g}{\partial x_i}(a_1, ..., a_{n-1}) \tag{1.33}$$

Setze λ gleich

$$\lambda = \frac{\frac{\partial h}{\partial x_n}(a)}{\frac{\partial f}{\partial x_n}(a)}$$

Aus (1.32) und (1.33) folgt:

$$\frac{\partial h}{\partial x_i}(a) = \lambda \frac{\partial f}{\partial x_i}(a), i = 1, ..., n$$

□

Beispiel 1.6.13 *Sei $U = \mathbb{R}^n, h(x, y) = x^2 - 3y^2$. Wir bestimmen die Extrema von h unter der Nebenbedingung $f = 0$ für $f(x, y) = x^2 + y^2 - 1$. Also die Extrema von h auf dem Einheitskreis um 0.*

Es gilt:

$$\nabla h(x, y) = (2x, -6y), \nabla f(x, y) = (2x, 2y)$$

Man sieht hier sofort, dass ∇h nur dort ein Vielfaches von ∇f sein kann, wo entweder $x = 0$ odr $y = 0$ gilt.
$\Rightarrow (\pm 1, 0), (0, \pm 1)$ - es gilt $h(\pm 1, 0) = 1, h(0, \pm 1) = -3$, dies sind also die globalen Maxima bzw. Minima von h auf dem Einheitskreis um 0. Da der Kreis kompakt ist wissen wir, dass das Minimum und das Maximum tatsächlich angenommen wird.

Beispiel 1.6.14 *Sei $A = (a_{ij}) \in Mat(n, n; \mathbb{R})$ eine symmetrische nxn-Matrix. Damit kann man eine Funktion auf $\mathbb{R}^n$ definieren durch*

$$h(x) = \langle x, Ax \rangle = \sum_{i,j=1}^{n} a_{ij} x_i x_j$$

(die zugehörige quadratische Form). Wir wollen die Extrema von h, eingeschränkt auf die Einheitssphäre um 0 bestimmen, das heißt die Extrema von h unter der Nebenbedingung $f = 0$ mit

$$f(x) := \langle x, x\rangle - 1 = \|x\|^2 - 1 = \sum_{i=1}^{n} x_i^2 - 1$$

Also gilt hier $M := f^{-1}(\{0\}) = \{x \in \mathbb{R}^n \mid \|x\| = 1\}$. Dazu werden wir wieder die Lagrange-Multiplikator-Methode (1.6.12) anwenden. Berechne dazu den Gradient von f.

$$\nabla f(x) = (2x_1, 2x_2, ...) = 2x$$

$$\begin{aligned}
\frac{\partial h}{\partial x_k}(x) &= \frac{\partial}{\partial x_k}\sum_{i,j=1}^{n} a_{i,j}x_i x_j = \sum_{i,j=1}^{n}(a_{ij}x_i\frac{\partial x_j}{\partial x_k} + a_{ij}\frac{\partial x_i}{\partial x_k}x_j) = \\
&= \sum_{j=1}^{n} a_{kj}x_j + \sum_{i=1}^{n} a_{ik}x_i = 2\sum_{i=1}^{n} a_{ki}x_i
\end{aligned}$$

Die k-te Komponente von $\nabla h(x)$ ist also $2\sum_{i=1}^{n} a_{ki}x_i$. Das ist gerade die k-te Komponente des Vektors $2Ax$ also folgt

$$\nabla h(x) = 2Ax$$

damit gilt:

$$\nabla h(x) = \lambda\nabla f(x) \Leftrightarrow 2Ax = 2\lambda x \Leftrightarrow Ax = \lambda x$$

Da M kompakt ist, wird das Maximum (bzw. Minimum) in einem Punkt $v \in M$ angenommen. Dieser Punkt ist also ein Eigenvektor zum Eigenwert λ von A. Wegen

$$h(v) = \langle v, Av\rangle = \langle v, \lambda v\rangle = \lambda\underbrace{\langle v, v\rangle}_{=1} = \lambda$$

folgt, dass ein Eigenvektor zum größten (kleinsten) Eigenwert von A ist. Insbesondere haben wir damit bewiesen, dass eine reelle symmetrische nxn-Matrix mindestens einen reellen Eigenwert besitzt.

1.7 Parameterabhängige Integrale und Variationsrechnung

Wir wollen uns nun zunächst mit folgenden Problem beschäftigen: Sei $(x, y) \mapsto f(x, y)$ eine Funktion mit zwei Variablen. Für feste y betrachten wir das Integral $a \leq x \leq b$ und erhalten damit eine Funktion $\varphi(y) := \int_a^b f(x, y)dx$ des Parameters

y. Wir wollen untersuchen, unter welchen Vorraussetzungen an f die Funktion φ stetig oder differenzierbar von dem Parameter y abhängt (und wann man Limesbildung und Integral vertauschen darf bzw. unter welchem Integral differenziert werden darf). Dies werden wir unter anderem nutzen, um die Eulerschen Differentialgleichungen der Variationsrechnung herzuleiten.

Lemma 1.7.1 *Sei* $[a,b] \subset \mathbb{R}$ *ein kompaktes Intervall und* $U \subseteq \mathbb{R}^m$ *eine beliebige Teilmenge. Die Funktion*

$$f : [a,b] \times U \to \mathbb{R}, (x,y) \mapsto f(x,y)$$

sei stetig. Weiter sei $(y_k)_{k\in\mathbb{N}}$ *eine konvergente Folge von Punkten in* U *mit* $c := \lim_{k\to\infty} y_k \in U$. *Dann konvergieren die Funktionen*

$$F_k : [a,b] \to \mathbb{R}, F_k(x) := f(x, y_k)$$

für $k \to \infty$ *gleichmäßig gegen die Funktion:*

$$F : [a,b] \to \mathbb{R}, F(x) := f(x,c)$$

Beweis Nach Satz (1.2.3) ist die Menge $Q := \{y_k \mid k \in \mathbb{N}\} \cup \{c\}$ kompakt, abgeschlossen und beschränkt. Deshalb ist die Menge $[a,b] \times Q \subset \mathbb{R}^{m+1}$ abgeschlossen und beschränkt - somit kompakt. Nach Lemma 4.59, Analysis I Skript, ist also $f_{|[a,b]\times Q} \to \mathbb{R}$ gleichmäßig stetig. Sei $\varepsilon > 0$ gegeben dann existiert ein $\delta > 0$, so dass für alle $(x,y), (x',y')$ gilt:

$$\|(x,y) - (x',y')\| < \delta \Rightarrow |f(x,y) - f(x',y')| < \varepsilon$$

Wegen $\lim y_k = c$ gibt es ein $N \in \mathbb{N}$ mit $\|c - y_k\| < \delta \forall k \geq N$. Daher gilt für $k \geq N$, $|f(x,y_k) - f(x,c)| < \varepsilon$ für alle $x \in [a,b]$. Das implitziert wiederrum

$$F_k(x) - F(x)| < \varepsilon \forall x \in [a,b], k \geq N$$

Das heißt die Funktionenfolge konvergiert gleichmäßig gegen F. □

Lemma 1.7.2 *Die Folge stetiger Funktionen* $F_k : [a,b] \to \mathbb{R}$ *konvergieren gleichmäßig gegen die (somit stetige) Grenzfunktion* $F : [a,b] \to \mathbb{R}$, $F(x) = \lim F_k(x)$. *Dann darf man Integration und Limesbildung vertauschen, das heißt es gilt*

$$\lim_{k\to\infty} \int_a^b F_k(x)dx = \int_a^b \lim_{k\to\infty} F_k(x)dx = \int_a^b F(x)dx$$

Beweis Sei $\varepsilon > 0$ vorgegeben. Dann gibt es wegen der gleichmäßigen Konvergenz ein $N \in \mathbb{N}$, so dass für alle $x \in [a, b]$ gilt:

$$|F_k(x) - F(x)| < \frac{\varepsilon}{b-a} =: c, k \geq N$$

Somit gilt $F_k(x) - c < F(x) < F_k(x) + c$ für alle $x \in [a, b]$. Daraus folgt:

$$\underbrace{\int_a^b (F_k(x) - c)dx}_{=\int_a^b F_k(x)dx - \varepsilon} \leq \int_a^b F(x)dx \leq \underbrace{\int_a^b (F_k(x) + c)dx}_{=\int_a^b F_k(x)dx + \varepsilon}$$

Also folgt:

$$|\int_a^b F_k(x)dx - \int_a^b F(x)dx| < \varepsilon$$

□

Satz 1.7.3 *Sei* $[a, b] \subset \mathbb{R}$ *ein kompaktes Intervall,* $U \subseteq \mathbb{R}^n$ *eine beliebige Teilmenge und* $f : [a, b] \times U \to \mathbb{R}$ *eine stetige Funktion. Für* $y \in U$ *sei* $\varphi(y) := \int_a^b f(x, y)dx$*. Dann ist die dadurch definierte Funktion* $\varphi : U \to \mathbb{R}$ *stetig.*

Beweis Sei $c \in U$ ein beliebiger punkt und $(y_k)_{k \in \mathbb{N}}$ eine Punktfolge in U mit $\lim y_k = c$. Dann gilt, wenn wir wieder (1.7.1) anwenden:

$$F_k(x) := f(x, y_k), F(x) := f(x, c)$$

$$\begin{aligned} \varphi(y_k) &= \int_a^b F_k(x)dx \wedge \\ \varphi(y) &= \int_a^b F(x)dx \end{aligned}$$

Da nach (1.7.1) F_k gleichmäßig gegen F konvergiert, können wir nach (1.7.2) Integral und Limes vertauschen, das heißt es gilt:

$$\lim_{k \to \infty} \int_a^b F_k(x)dx = \int_a^b F(x)dx$$

Also gilt $\varphi(y_k) \longrightarrow_{k \to \infty} \varphi(c)$, was zu zeigen war. □

Lemma 1.7.4 *Seien* I, J *zwei kompakte Intervalle* $f : I \times J \to \mathbb{R}, (x, y) \mapsto f(x, y)$ *eine stetige Funktion, die nach der zweiten Variable stetig differenzierbar ist. Sei weiter* $c \in J$ *und* $y_k \in J$ *eine Folge, die gegen* c *konvergiert mit* $y_k \neq c$*. Seien auf dem Intervall* I *Funktionen* F_k, F *definiert durch*

$$F_k(x) := \frac{f(x, y_k) - f(x, c)}{y_k - c}, F(x) = \frac{\partial f}{\partial y}(x, c)$$

dann kovergiert die Funktionenfolge F_k *gleichmäßig gegen* F*.*

Beweis Sei $\varepsilon > 0$ vorgegeben. Nach Lemma 5.49, Analysis I Skript von Prof. Bunke, ist die Funktion $D_2 f : I \times J \to \mathbb{R}$ gleichmäßig stetig (da sie stetig ist und $I \times J$ kompakt). Also gibt es ein $\delta > 0$, so dass für alle $y, y' \in J$ gilt:

$$|y - y'| < \delta \Rightarrow |D_2 f(x, y) - D_2 f(x, y')| < \varepsilon$$

Nach dem Mittelwertsatz der Differentialrechnung für Funktionen einer Variable (Lemma 5.11, Analysis I Skript von Prof. Bunke) gibt es zu jedem $k \in \mathbb{N}$ ein η_k zwischen y_k und c mit $F_k(x) = D_2 f(x, \eta_k)$. Wähle nun ein $N \in \mathbb{N}$, so dass $|c - y_k| < \delta \forall k \geq N$. Da η_k zwischen c und y_k liegt, gilt auch $|c - \eta_k| < \delta$ für $k \geq N$ und es folgt:

$$|F(x) - F_k(x)| = |D_2 f(x, c) - D_2 f(x, \eta_k) < \varepsilon$$

für alle $x \in I$ und $k \geq N$. □

Damit können wir eine Bedingung beweisen, unter der wir „Unter dem Integral differenzieren" dürfen.

Satz 1.7.5 *Seien I, J kompakte Intervalle und $f : I \times J \to \mathbb{R}$ eine stetige Funktion, die nach der zweiten Variablen stetig partiell differenzierbar ist. Für $y \in J$ definiere*

$$\varphi(y) := \int_I f(x, y) dx$$

Dann ist die Funktion $\varphi : J \to \mathbb{R}$ stetig differenzierbar und es gilt:

$$\frac{\partial \varphi}{\partial y}(y) = \int_I \frac{\partial f}{\partial y}(x, y) dx$$

Beweis Sei $c \in J$ und $y_k \in J$ eine gegen c konvergente Folge mit $y_k \neq c$. Die Funktionen F_k und F seien wie in (1.7.4) definiert und aus (1.7.4) folgt, dass F_k gleichmäßig gegen F konvergiert. Es folgt:

$$\begin{aligned}
\frac{\varphi(y_k) - \varphi(c)}{y_k - c} &= \frac{\int_I f(x, y_k) dx - \int_I f(x, c) dx}{y_k - c} = \int_I \frac{f(x, y_k) - f(x, c)}{y_k - c} dx = \\
&= \int_I F_k(x) dx, k \to \infty, F_k \to F \text{ gleichmäßig} \\
\int_I F(x) dx &= \int_I \frac{\partial f}{\partial y}(x, y) dx
\end{aligned}$$

Somit existiert $\varphi'(c)$ für jedes $c \in J$ und es gilt:

$$\varphi'(c) = \int_I \frac{\partial f}{\partial y}(x, y) dx$$

□

Beispiel 1.7.6 *Wir wollen (1.7.5) benutzen, um das Integral*

$$\int_0^x xe^x dx$$

zu berechnen. Dies kann man natürlich mit Hilfe der partiellen Integration tun. Statt dessen betrachten wir aber die Funktion:

$$F(y) = \int_0^a e^{xy} dx$$

Zum Beispiel für $y \in [\frac{1}{2}, 2]$. Dann gilt:

$$F(y) = [\frac{1}{y} e^{xy}]_{x=0}^{x=a} = \frac{1}{y}(e^{ay} - 1)$$

Somit gilt:

$$F'(y) = -\frac{1}{y^2}(e^{ay} - 1) + \frac{1}{y} a e^{ay}$$

Nach Satz (1.7.3) gilt aber andererseits:

$$F'(y) = \int_0^a \frac{\partial}{\partial y} e^{xy} dx = \int_0^a xe^{xy} dx$$

Der Wert des gesuchten Integrals ist also:

$$F'(1) = -e^a + 1 + ae^a$$

Als eine erste Anwendung wollen wir uns noch einmal mit der in (1.3) gestellten Frage beschäftigen, wann ein stetig differenzierbarer Vektor der Gradient einer Funktion ist. Sei also $U \subseteq \mathbb{R}^n$ offen und $v = (v_1, ..., v_n) : U \to \mathbb{R}^n$ ein stetig differenzierbares Vektorfeld $f : U \to \mathbb{R}$, so dass $\nabla f(x) = v(x)$. Also:

$$\frac{\partial f}{\partial x_i}(x) = v_i(x), \forall x \in U, i = 1, ..., n$$

Bereits in (1.3) haben wir uns überlegt, dass dann (da f dann zweimal stetig differenzierbar) nach Satz (1.3.8) gilt:

$$\frac{\partial^2 f}{\partial x_i \partial x_j}(x) = \frac{\partial^2 f}{\partial x_j \partial x_i}(x) \Rightarrow \frac{\partial v_j}{\partial x_i}(x) = \frac{\partial v_i}{\partial x_j}(x), \forall x \in U, i, j = 1, ..., n$$

Insbesondere gilt für $n = 3$, dass $rot(v) \equiv 0$. Die Bedingung $D_i v_j = D_j v_i$ ist also notwendig für die Existenz einer Funktion f mit $\nabla f = v$. Wir wollen jetzt auch eine hinreichende Bedingung herleiten. Sei f, v wie oben mit $\nabla f = v$. Seien $a, x \in U$ zwei Punkte so, dass die Strecke zwischen a, x ganz in U liegt, das heißt es gilt $a + t(x - a) \in U$ für alle $t \in [0, 1]$. Dann können wir die Funktion

$g : [0,1] \to \mathbb{R}, t \mapsto f(a + t(x-a))$ betrachten. Für die Ableitung $g'(t)$ gilt nach der Kettenregel (1.4.10):

$$g'(t) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(a + t(x-a))(x_i - a_i) = \sum_{i=1}^{n} v_i(a + t(x-a))(x_i - a_i)$$

Es folgt

$$f(x) = g(1) = g(0) + \int_0^1 g'(t)dt = f(a) + \sum_{i=1}^{n} \int_0^1 v_i(a + t(x-a))dt(x_i - a_i)$$

Satz 1.7.7 *Sei $U \subset \mathbb{R}^n$ eine offene Teilmenge, die sternförmig bezüglich eines Punktes $a \in U$ ist, das heißt für jedes $x \in U$ liegt die Strecke zwischen a, x in U (also $a + t(x-a) \in U$ für alle $t \in [0,1]$). Sei $v = (v_1, ..., v_n) : U \to \mathbb{R}^n$ ein stetig differenzierbares Vektorfeld für das $D_i v_j(x) = D_j v_i(x)$ für alle $x \in U, i, j = 1, ...n$ gilt. Dann ist die durch*

$$f(x) = \sum_{i=1}^{n} \int_0^1 v_i(a + t(x-a))dt(x_i - a_i)$$

definierte Funktion stetig differenzierbar und es gilt:

$$\nabla f(x) = v(x), \forall x \in U$$

Beweis Zur Vereinfachung setzen wir ohne Einschränkung $a = 0$. Wir betrachten also

$$f(x) = \sum_{i=1}^{n} \int_0^1 v_i(tx)dtx_i$$

Nach geeigneter Einschränkung können wir auf die Funktion $(t, x_j) \mapsto v_i(tx_1, ..., tx_n)$ Satz (1.7.5) anwenden. Dann folgt:

$$\begin{aligned}
\frac{\partial}{\partial x_j} f(x) &= \sum_{i=1}^{n} \frac{\partial}{\partial x_j}\left(\int_0^1 v_i(tx)dtx_i\right) = \sum_{i=1}^{n} \int_0^1 \frac{\partial}{\partial x_j} v_i(tx)dtx_i + \int_0^1 v_i(tx)dt\frac{\partial x_i}{\partial x_j} = \\
&= \sum_{i=1}^{n} \int_0^1 \frac{\partial v_i}{\partial x_j}(tx)tdtx_i + \int_0^1 v_j(tx)dt = \sum_{i=1}^{n} \int_0^1 \frac{\partial v_j}{\partial x_i}(tx)tx_i + v_j(tx)dt = \\
&= \int_0^1 \frac{\partial}{\partial t}(tv_j(x))dt = tv_j(tx) \mid_{t=0}^{t=1} = v_j(x)
\end{aligned}$$

Also ist $\nabla f(x) = v(x)$ und da v insbesondere stetig ist, ist f also auch stetig differenzierbar. $\square$

Bemerkung (1.7.6) ist ein Spezialfall des Poincaré-Lemmas.

Beispiel 1.7.8 *Wir betrachten zwei Vektorfelder auf* $\mathbb{R}^2$.

1. $$v(x_1, x_2) = (v_1(x_1, x_2), v_2(x_1, x_2)) = (x_1x_2, x_1x_2)$$

   Es gilt:
   $$D_1v_2(x_1, x_2) = x_2 \neq D_2v_1(x_1, x_2) = x_1$$
   Es gibt also keine stetig differenzierbare Funktion $f : \mathbb{R}^2 \to \mathbb{R}$ mit $\nabla f(x) = v(x)$.

2. $$w(x_1, x_2) = (w_1(x_1, x_2), w_2(x_1, x_2)) = (x_1x_2^2, x_1^2x_2)$$

   Es gilt:
   $$D_1w_1(x_1, x_2) = 2x_1x_2 = D_2w_2(x_1, x_2)$$
   Nachdem die Integrabilitätsbedingung erfüllt ist, gibt es also (da $\mathbb{R}^2$ sternförmig bezüglich 0 und v_2 stetig differenzierbar) nach (1.7.7) eine stetig differenzierbare Funktion $f : \mathbb{R}^2 \to \mathbb{R}$ mit $\nabla f(x) = w(x)$, nämlich:

   $$\begin{aligned} f(x_1, x_2) &= \int_0^1 w_1(tx_1, tx_2)x_1dt + \int_0^1 w_2(tx_1, tx_2)x_2dt = \\ &= \int_0^1 t^3x_1^3x_2^2dt + \int_0^1 t^3x_1^2x_2^2dt = \frac{1}{2}x_1^2x_2^2 \end{aligned}$$

Doppel- und mehrfache Integrale Seien $[a, b] \subset \mathbb{R}, [c, d] \subset \mathbb{R}$ zwei kompakte Intervalle und $f : [a, b] \times [c, d] \to \mathbb{R}$ eine stetige Funktion. Nach Satz (1.7.3) ist die Funktion
$$F(y) := \int_a^b f(x, y)dx$$
stetig auf $[c, d]$. Daher existiert das *Doppelintegral*

$$\int_c^d F(y)dy = \int_c^d (\int_a^b f(x, y)dx)dy$$

Man könnte f aber zuerst bezüglich y und dann bezüglich x integrieren. Der folgende Satz zeigt, dass sich dabei der selbe Wert ergibt.

Satz 1.7.9 *Mit den obigen Bezeichnungen gilt:*

$$\int_c^d (\int_a^b f(x, y)dx)dy = \int_a^b (\int_c^d f(x, y)dy)dx$$

Beweis Wir definieren eine Funktion $\varphi : [c,d] \to \mathbb{R}$ durch:

$$\varphi(y) := \int_a^b (\int_c^y f(x,t)dt)dx$$

Dann gilt $\varphi(c) = 0$ und nach (1.7.5) ist φ differenzierbar und es gilt:

$$\varphi'(y) = \int_a^b \frac{\partial}{\partial y}(\int_c^y f(x,t)dt)dx = \int_a^b f(x,y)dx$$

Daraus folgt:

$$\int_c^d \int_a^b f(x,y)dxdy = \int_c^d \varphi'(y)dy = \varphi(d) - \varphi(c) = \int_a^b \int_c^d f(x,t)dtdx$$

□

Bemerkung Für das Doppelintegral schreibt man auch:

$$\int_{[a,b]\times[c,d]} f(x,y)dxdy := \int_a^b \int_c^d f(x,y)dydx$$

Ein analoger Satz gilt auch für eine stetige Funktion, die auf einem Quader $I_1 \times I_2 \times I_3 \times ... \times I_n = [a_1,b_1] \times ... \times [a_n,b_n] \subset \mathbb{R}^n$ definiert ist.

(Anfänge der) Variationsrechnung In der Variationsrechnung geht es darum, in einem Raum von Funktionen $\mathfrak{F}$ diejenigen zu finden, die ein gewisses Funktional (eine Abbildung $\mathfrak{F} \to \mathbb{R}$) haben. Typisches Problem: Finde unter allen Kurven, die zwei Punkte verbinden, diejenige(n) mit kürzesten Länge. Ein einfaches Variationsproblem könnte, zum Beispiel, so aussehen: Wir betrachten den Raum von Funktionen $\mathfrak{F} := \{\varphi \in C^2([a,b]) \mid \varphi(a) = c_1, \varphi(b) = c_2\}$, also die Menge aller zweimal stetig differenzierbaren Funktionen, die an den Endpunkten a und b vorgegebene Werte c_1, c_2 annehmen. Auf F definieren wir jetzt eine Abbildung $S : \mathfrak{F} \to \mathbb{R}$ die wir minimieren wollen, das heißt wir wollen eine Funktion $\varphi \in \mathfrak{F}$, so dass $S(\varphi) = \inf\{S(\psi) \mid \psi \in \mathfrak{F}\}$. Das Funktional S soll wie folgt definiert sein:

$$S(\varphi) := \int_a^b L(t, \varphi(t), \varphi'(t))dt$$

wobei

$$L : [a,b] \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}, (t,y,p) \mapsto L(t,y,p)$$

eine zweimal stetig differenzierbare Funktion ist.

Beispiel 1.7.10 *Seien* $a, b, c, d \in \mathbb{R}$ *mit* $a < b$ *und sei*

$$\mathfrak{F} := \{\varphi \in C^2([a,b]) \mid \varphi(a) = c, \varphi(b) = d\}$$

als Funktional S *verwenden wir die Bogenlänge, welche in Analysis I Definition 6.25 eingeführt worden ist, der Kurven:*

$$[a,b] \to \mathbb{R}^2, t \mapsto (t, \varphi(t)), S(\varphi) := \int_a^b \sqrt{1 + \varphi'(t)^2} dt$$

Das heißt wir setzen $L(t, y, p) := \sqrt{1 + p^2}$*. Somit hängt* S *hier im Beispiel nicht von* t, y *ab.*

Satz 1.7.11 *Sei* $\mathfrak{F}$ *wie oben,* $L : [a,b] \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ *eine zweimal stetig differenzierbare Funktion und* $S(\varphi) := \int_a^b L(t, \varphi(t), \varphi'(t)) dt$*. Falls* $S(\varphi) = \inf\{S(\psi) \mid \psi \in \mathfrak{F}\}$ *für ein* $\varphi \in \mathfrak{F}$ *gilt, dann erfüllt* φ *die* Eulersche Differentialgleichung*:*

$$\frac{d}{dt} \frac{\partial L}{\partial p}(t, \varphi(t), \varphi'(t)) - \frac{\partial L}{\partial y}(t, \varphi(t), \varphi'(t)) = 0$$

Für alle $t \in [a, b]$*.*

Beweis Sei $\varphi \in \mathfrak{F}$ eine Funktion mit $S(\varphi) \leq S(\psi) \forall \psi \in \mathfrak{F}$. Wir führen nun eine Variation der Funktion durch. Dazu wählen wir eine beliebig zweimal stetig differenzierbare Funktion

$$g : [a,b] \to \mathbb{R}, g(a) = g(b) = 0$$

und betrachten die Funktionenschar mit $t \mapsto \varphi(t) + \varepsilon g(t)$. Für jedes $\varepsilon \in \mathbb{R}$ ist dann $\varphi + \varepsilon g$ eine Funktion in $\mathfrak{F}$ und wir definieren $F : \mathbb{R} \to \mathbb{R}, F(\varepsilon) := S(\varphi + \varepsilon g)$. Die Funktion F nimmt also bei $\varepsilon = 0$ ein Minimum an. Mit Satz (1.7.5) folgt: Da F bei $\varepsilon = 0$ ein Minimum hat, gilt

$$\frac{\partial F}{\partial \varepsilon}(0) = 0 \Rightarrow \frac{\partial F}{\partial \varepsilon}(\varepsilon) = \int_a^b \frac{\partial}{\partial \varepsilon} L(t, \varphi(t) + \varepsilon g(t), \varphi'(t) + \varepsilon g'(t)) dt$$

$$= \int_a^b \frac{\partial L}{\partial y}(...) g(t) + \frac{\partial L}{\partial p} g'(t) dt$$

Den zweiten Summanden behandeln wir nun mit partieller Integration.

$$\int_a^b \frac{\partial L}{\partial p}(...) g'(t) dt = \underbrace{\frac{\partial L}{\partial p}(...) g(t) \mid_{t=a}^{t=b}}_{=0, g(a)=g(b)=0} - \int_a^b \frac{d}{dt} \frac{\partial L}{\partial p}(...) g(t) dt$$

Also erhalten wir

$$0 = \frac{dF}{d\varepsilon}(0) = \int_a^b \underbrace{(\frac{\partial l}{\partial y}(t, \varphi(t), \varphi'(t)) - \frac{d}{dt}\frac{\partial L}{\partial p}(t, \varphi(t), \varphi'(t)))}_{(*)} g(t)dt$$

Der Satz ist gezeigt, da man aus der obigen Gleichung, die für alle zweimal stetig differenzierbaren Funktionen $g : [a, b] \to \mathbb{R}$ mit $g(a) = g(b) = 0$ gilt, folgern kann, dass die Funktion $(*)$ gleich Null ist. Das zeigen wir mit dem nachfolgenden Lemma. $\square$

Lemma 1.7.12 *Seien $a, b \in \mathbb{R}$, $a < b$ mit $f : [a, b] \to \mathbb{R}$ eine stetige Funktion, die folgende Eigenschaft hat: Für jede zweimal stetig differenzierbare Funktion $g : [a, b] \to \mathbb{R}$ mit $g(a) = g(b) = 0$ gilt $\int_a^b f(t)g(t)dt = 0$. Dann ist $f(t) = 0 \forall t \in [a, b]$.*

Beweis Da f stetig ist genügt es zu zeigen, dass $f(t) = 0$ für alle $t \in (a, b)$. Annahme: Es gebe $x \in (a, b)$ mit $f(x) \neq 0$. Ohne Einschränkung ist $\varepsilon := f(x) > 0$. Dann existiert ein $\delta > 0$, so dass $[x - \delta, x + \delta] \subset (a, b)$ und so dass $f(t) \geq \frac{\varepsilon}{2} \forall t \in [x - \delta, x + \delta]$. Man kann nun eine zweimal stetig differenzierbare Funktion $g : [a, b] \to [0, \infty)$ konstruieren mit $g(x) > 0, g(t) = 0$ für $t \notin [x - \delta, x + \delta]$. Nun folgt:

$$0 = \int_a^b f(t)g(t)dt = \int_{x-\delta}^{x+\delta} f(t)g(t)dt \geq \frac{\varepsilon}{2} \int_{x-\delta}^{x+\delta} g(t)dt > 0$$

Somit folgt der gewünschte Widerspricht. Also muss doch $f(t) = 0 \forall t \in (a, b)$ sein. $\square$

Zurück zu Beispiel (1.7.10): Wir wollen diejenige Funktion φ aus $\mathfrak{F}$ bestimmen, deren Graph minimale Länge hat, also:

$$S(\varphi) := \int_a^b \sqrt{1 + \varphi'(t)^2} dt$$

Die Funktion L hat also die einfache Gestalt

$$L(t, y, p) = \sqrt{1 + p^2}$$

und hängt somit nicht von t und y ab. Es gilt

$$\frac{\partial L}{\partial y} = 0, \frac{\partial L}{\partial p}(t, y, p) = \frac{p}{\sqrt{1 + p^2}}$$

Die Eulersche Differentialgleichung (1.7.11) liefert uns:

$$\frac{d}{dt}\frac{\partial L}{\partial p}(t, \varphi(t), \varphi'(t)) - \frac{\partial L}{\partial y}(t, \varphi(t), \varphi'(t)) = 0$$

Also in diesem Fall:

$$\frac{d}{dt}\frac{\varphi'}{\sqrt{1+\varphi'^2}} = 0 \Leftrightarrow \frac{\sqrt{1+\varphi'^2}\varphi'' - \varphi'\frac{\varphi'\varphi''}{\sqrt{1+\varphi'^2}}}{1+\varphi'^2} = 0$$

$$\varphi''(t)\underbrace{(1-\frac{\varphi'^2(t)}{1+\varphi'^2(t)})}_{\neq 0} = 0 \Rightarrow \varphi''(t) = 0$$

Wenn es also überhaupt eine Lösung des Variationsproblems gibt, dann muss φ ein Polynom 1. Grades sein, welches $\varphi(a) = c$ und $\varphi(b) = d$ erfüllt.

$$\Rightarrow \varphi(t) = \frac{d}{b-a}(t-a) + \frac{c}{a-b}(t-b)$$

Das heißt nur das verbindende Geradenstück kann eine Lösung sein. Dass das verbindende Geradenstück tatsächlich minimale Länge hat kann man hier durch geometrische Überlegungen nachweisen. Im Allgemeinen ist es schweirig zu zeigen, dass das Minimum tatsächlich angenommen wird.

Beispiel 1.7.13 (Anwendungen in der Physik) ***Hamiltonsches Prinzip der kleinsten Wirkung.*** *Wir bemerken zunächst, dass sich Satz (1.7.11)leicht auf den Fall verallgemeinern lässt, dass $\varphi : [a,b] \to \mathbb{R}^n$ eine Funktion nach $\mathbb{R}^n$ mit Komponenten $\varphi = (\varphi_1, ..., \varphi_n)$ ist. Dazu wählt man eine Funktion*

$$L : [a,b] \times \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}, (t, y_1, ..., y_n, p_1, ..., p_n) \mapsto L(t, y_1, ..., y_n, p_1, ..., p_n)$$

und betrachtet

$$\mathfrak{F} = \{\varphi : [a,b] \to \mathbb{R}^n \mid \varphi \in C^2([a,b]), \varphi(a) = c, \varphi(b) = d\}$$

für $c, d \in \mathbb{R}^n$ mit:

$$S := \int_a^b L(t, \varphi(t), \varphi'(t))dt = \int_a^b L(t, \varphi_1(t), ..., \varphi_n(t), \varphi_1'(t), ..., \varphi_n'(t))dt$$

Für eine dem Funktional minimierende Funktion $\varphi \in \mathfrak{F}$ gelten dann die Eulerschen Differentialgleichungen:

$$\frac{d}{dt}\frac{\partial L}{\partial p_i}(t, \varphi(t), \varphi'(t)) - \frac{\partial L}{\partial y_i}(t, \varphi(t), \varphi'(t)) = 0, i = 1, ..., n$$

Beim Hamiltonschen Prinzip der kleinsten Wirkung beschreibt die Funktion $\varphi_1, ..., \varphi_n$ den Zustand des physikalischen Systems zur Zeit t, $L(t, \varphi, \varphi')$ ist die Lagrange-Funktion.

Bei reibungsfreien, mechanischen Systemen ist L gleich der Differenz $T - U$ aus kinetischer Energie T und potentieller Energie U.

$$S(\varphi) = \int_a^b L(t, \varphi(t), \varphi'(t))dt$$

ist das sogenannte Wirkungsintegral. Das Hamiltonsche Prinzip sagt: Für den tatsächlich ablaufenden Vorgang φ wird die Wirkung $S(\varphi)$ minimal, wird also durch die Eulerschen Differentialgleichungen beschrieben.

Beispiel 1.7.14 *Bewegung eines Massenpunktes im $\mathbb{R}^3$ unter dem Einfluss eines zeitlichen konstanten Potentials $U : \mathbb{R}^3 \to \mathbb{R}, (x_1, x_2, x_3) \mapsto U(x_1, x_2, x_3)$. Die Bewegung eines Massenpunktes wird beschrieben durch die vektorwertige Funktion $x(t) = (x_1(t), x_2(t), x_3(t)), x : [a, b] \to \mathbb{R}^3$. Seine Geschwindigkeit ist $v(t) = \frac{dx}{dt}(t) = (\dot{x}_1(t), \dot{x}_2(t), \dot{x}_3(t))$. Die kinetische Energie ist somit*

$$T = \frac{m}{2}(\dot{x}_1^2 + \dot{x}_2^2 + \dot{x}_3^2) = \frac{m}{2}\|v\|^2$$

Für die Lagrangefunktion $L = T - U$ gilt:

$$L(x, v) = \frac{m}{2}(v_1^2 + v_2^2 + v_3^2) - U(x_1, x_2, x_3)$$

(sie hängt nicht explizit von der Zeit ab) Es gilt:

$$\frac{\partial L}{\partial x_i} = -\frac{-\partial U}{\partial x_i}, \frac{\partial L}{\partial v_i} = mv_i$$

Die Eulerschen Differentialgleichungen lauten somit:

$$\underbrace{\frac{d}{dt}(m\dot{x}_i(t))}_{=m\ddot{x}_i(t)} + \frac{\partial U}{\partial x_i}(x(t)) = 0, i = 1, 2, 3$$

Dies kann man so zusammenfassen:

$$m\ddot{x}(t) = -\nabla U(x(t))$$

Aus dieser Gleichung folgt insbesondere, dass die Gesamtenergie $E = T + U$ zeitlich konstant bleibt:

$$E(t) = \frac{m}{2}\sum_{i=1}^{3} \dot{x}_i(t)^2 + U(x_1(t), x_2(t), x_3(t))$$

Für die Ableitung nach der Zeit ergibt sich somit:

$$\begin{aligned} \frac{dE}{dt}(t) &= m\sum_{i=1}^{3} \dot{x}_i(t)\ddot{x}_i(t) + \sum_{i=1}^{3} \frac{\partial U}{\partial x_i}(x(t))\dot{x}_i(t) = \\ &= \sum_{i=1}^{3} x_i(m\ddot{x}_i(t) + \frac{\partial U}{\partial x_i}(x(t))) = x_i(-\nabla U(x(t)) + \nabla U(x(t))) = 0 \end{aligned}$$

Kapitel 2

Gewöhnliche Differentialgleichungen

2.8 Einführung

Aufgabe Bestimme alle Funktionen $y = f(x), x \in (0, \infty)$, so dass die Tangente an den Graphen in jedem Punkt $(x, y) \in \Gamma_f$ die y-Achse im Punkt $(0, \frac{1}{2}y)$ schneidet.
Bei dieser Aufgabe ist als Lösung eine Funktion $f : (0, \infty) \to \mathbb{R}$ gesucht. Über den Verlauf der Funktion ist nur bekannt: Verläuft der Graph durch den Punkt (x, y), so ist die Gerade durch $(0, \frac{1}{2}y)$ eine Tangente in (x, y). Also beträgt die Tangentensteigung bei x genau $\frac{y}{2x}$. Mit anderen Worten: Hat f bei x den Wert $f(x)$ (und ist f differenzierbar bei x), so hat die Ableitung den Wert $f'(x) = \frac{f(x)}{2x}$. Es handelt sich also um eine Differentialgleichung (das heißt es kommen die gesuchte Funktion und ihre Ableitungen darin vor).

Gewöhnliche Differentialgleichung Funktionen einer Variablen gesucht.

Partielle Differentialgleichung Funktionen mehrerer Variablen gesucht; es kommen partielle Ableitungen vor.

Hier Gewöhnliche Differentialgleichung, erster Ordnung - explizit:

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

Kurzschreibweise:

$$y' = \frac{y}{2x} \text{ für } f'(x) = \frac{f(x)}{2x}$$

Geometrische Anschauung Die Funktion $F(x, y)$ gibt ein „Richtungs- oder Steigungsfeld" auf $(0, \infty) \times \mathbb{R}$ vor:
Zeichne durch jeden Punkt (x, y) eine Gerade mit Steigung $\frac{y}{2x}$.

Näherungsweise Lösung (durch Polygonzug) An einer beliebigen Stelle x_0 kann ein Anfangswert $y_0 = y(x_0)$ beliebig vorausgesetzt werden. Dadurch erhält man eine Näherungslösung durch einen sog. Polygonzug, der bei x_i die Steigung $F(x_i, y(x_i))$ hat. Wir wollen jetzt eine Lösung der Gleichung herleiten durch *Trennung der Variablen*:
Eine Lösung ist $y \equiv 0$. Angenommen es existiert eine Lösung $y(x)$ auf einem Intervall $I \subseteq (0, \infty)$ mit $y(x) \neq 0$ für alle $x \in I$. dann gilt (für alle $x \in I$):

$$y'(x) = \frac{y(x)}{2x} = y(x)\frac{1}{2x} \Leftrightarrow \underbrace{\frac{y'(x)}{y(x)}}_{(\ln|y(x)|)'} = \frac{1}{2x} \Leftrightarrow \ln|y(x)| = \int_{x_0}^{x} \frac{1}{2x'}dx' = \frac{1}{2}\ln x + C$$

$$|y(x)| = e^{\frac{1}{2}\ln x + C} = \sqrt{x}e^C \Leftrightarrow y(x) = \sqrt{x}B, B \in \mathbb{R}\backslash\{0\}$$

Die Menge der Lösungen ist $\{f : (0, \infty) \to \mathbb{R} \mid f(x) = B\sqrt{x}, B \in \mathbb{R}\}$.

Später Existenz- und Eindeutigkeitssatz.

Weiteres Beispiel

$$y' = \frac{y}{x}, x > 0$$

Wir nehmen an, dass $y : I \to \mathbb{R}$ keine Nullstellen hat. Somit gilt:

$$\frac{y'(x)}{y(x)} = \frac{1}{x} \Rightarrow |y(x)| = e^{\ln x + C} \Rightarrow y(x) = Bx, B \in \mathbb{R}$$

2.9 Existenz und Eindeutigkeitssatz

Definition 2.9.1 *Sei U eine Teilmenge von $\mathbb{R} \times \mathbb{R}^n$ und $f : U \to \mathbb{R}^n, (x, y) \mapsto f(x, y), y = (y_1, ..., y_n)$ eine stetige Abbildung. Dann nennt man*

$$y' = f(x, y) \tag{2.1}$$

ein System von (gewöhnlichen) Differentialgleichungen erster Ordnung, falls $n = 1$. Unter einer Lösung von (2.1) versteht man eine auf einem Intervall $I \subseteq \mathbb{R}$ definierte differenzierbare Funktion $\varphi : I \to \mathbb{R}^n$, so dass der Graph von φ in U enthalten ist, das heißt

$$\Gamma_\varphi = \{(x, y) \in I \times \mathbb{R}^n \mid y = \varphi(x)\}$$

und so dass $\varphi'(x) = f(x, \varphi(x))$ für alle $x \in I$ gilt.

Bemerkung 2.9.2 *1. Das System lautet ausgeschrieben:*

$$\begin{aligned} y_1' &= f_1(x, y_1, ..., y_n) \\ &\vdots \\ y_n' &= f_n(x, y_1, ..., y_n) \end{aligned}$$

2. *Im Spezialfall wo die „rechte Seite" f der Differentialgleichung von der „Zeit" x abhängt nennt man das System autonom.*

3. *Falls f nicht von y abhängt, läßt sich eine Lösung einfach durch Integration*

$$\varphi(x) = c + \int_a^x f(x)dx$$

   bestimmen.

Wir wollen uns jetzt mit der Frage beschäftigen, wann Differentialgleichungen eindeutig lösbar sind.

Definition 2.9.3 *Sei $U \subseteq \mathbb{R} \times \mathbb{R}^n$ und $f : U \to \mathbb{R}^n$ eine Funktion. Man sagt f genüge in U einer Lipschitz-Bedingung mit der Lipschitz-Konstante L, wenn für alle $(x, y), (x, \tilde{y}) \in U$ gilt:*

$$\|f(x, y) - f(x, \tilde{y})\| \leq L\|y - \tilde{y}\|$$

Das heißt wenn bei festgehaltenen x die Abbildung $y \mapsto f(x, y)$ Lipschitz-Stetig mit Konstante L ist.
Wir sagen f genüge lokal einer Lipschitz-Bedingung, wenn jeder Punkt $(x, y) \in U$ eine Umgebung B besitzt, so dass f auf $U \cap B$ einer Lipschitz-Bedingung (mit einer B abhängigen Konstanten $K \geq 0$) genügt.

Satz 2.9.4 (Eindeutigkeitssatz) *Sei $U \subset \mathbb{R} \times \mathbb{R}^n$ und $f : U \to \mathbb{R}^n$ eine stetige Funktion, die lokal einer Lipschitz-Bedingung genügt. Seien $\varphi, \psi : I \to \mathbb{R}^n$ zwei Lösungen der Differentialgleichung*

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

auf dem Intervall $I \subseteq \mathbb{R}$. Gilt dann $\varphi(x_0) = \psi(x_0)$ für ein $x_0 \in I$ so folgt $\varphi(x) = \psi(x)$ für alle $x \in I$.

Beweis Wir zeigen zunächst: Ist $\varphi(a) = \psi(a)$ für ein $a \in I$, so gibt es ein $\varepsilon > 0$, so dass $\varphi(x) = \psi(x)$ für ale $x \in I$ mit $|x - a| \leq \varepsilon$. Durch Integration der beiden Gleichungen

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

erhalten wir wegen $\varphi(a) = \psi(a)$:

$$\varphi(x) - \psi(x) = \int_a^x (f(t, \varphi(t)) - f(t, \psi(t)))dt$$

Da f lokal einer Lipschitz-Bedingung genügt, gibt es Konstanten $L \geq 0$ und $\delta \geq 0$, so dass

$$\|f(t, \varphi(t)) - f(t, \psi(t))\| \leq L\|\varphi(t) - \psi(t)\| \forall t \in I$$

mit $|t-a| \leq \delta$. Daraus folgt:

$$\|\varphi(t)-\psi(t)\| \leq L|\int_a^x \|\varphi(t)-\psi(t)\|dt|$$

Setzt man weiter:

$$M(x) := \sup\{\|\varphi(t)-\psi(t)\| \mid |t-a| \leq |x-a|\}$$

so folgt weiter für alle $\xi \in I$ mit $|\xi - a| \leq |x-a| \leq \delta$.

$$\|\varphi(\xi)-\psi(\xi)\| \leq L|\xi-a|M(x) \leq L|x-a|M(x)$$

Also ergibt sich $M(x) \leq L|x-a|M(x)$. Für $|x-a| \leq \varepsilon := \min(\delta, \frac{L}{2})$ folgt damit:

$$M(x) \leq \frac{M(x)}{2} \Rightarrow M(x) = 0$$

Das heißt aber, dass $\varphi(x) = \psi(x)$ für $|x-a| \leq \varepsilon$ nach Definitio von $M(x)$.
Wir zeigen nun, dass die beiden Lösungen φ und ψ nicht nur auf einer ε-Umgebung von a übereinstimmen, sondern auf dem ganzen Intervall I. Wir zeigen nun, dass $\varphi(x) = \psi(x)$ für alle $x \in I$ mit $x \geq x_0$ (Der Beweis für $x \leq x_0$ verläuft genauso). Sei $x_1 := \sum\{\xi \in I \mid \varphi_{|[x_0,\xi]} = \psi_{|[x_0,\xi]}\}$. Falls $x_1 = \infty$ oder gleich der rechten Intervallgrenze sind wir fertig. Andernfalls existiert ein $\delta > 0$, so dass $[x_0, x_0+\delta] \subseteq I$. Da φ und ψ stetig sind gilt $\varphi(x_1) = \psi(x_1)$ und wir können den ersten Teil des Beweises anwenden. Es gibt ein $\varepsilon > 0$, so dass $\varphi(x) = \psi(x)$ für alle $x \in I$ in einer ε-Umgebung von x_1. Das widerspricht aber der Definition von x_1. Also stimmen φ, ψ auf ganz I überein. □

Um den Eindeutigkeitssatz anzuwenden, muss man meist nicht die Lipschitz-Bedingung direkt nachweisen.

Lemma 2.9.5 *Sei $U \subseteq \mathbb{R} \times \mathbb{R}^n$ offen und $f : U \to \mathbb{R}^n, (x,y) \mapsto f(x,y)$ bezüglich der Variablen $y = (y_1, ..., y_n)$ stetig partiell differenzierbar. Dann genügt f eine lokale Lipschitz-Bedingung.*

Beweis Sei (a,b) ein Punkt aus U. Da U offen ist, gibt es ein $r > 0$, so dass die kompakte Menge

$$V := \{(x,y) \in \mathbb{R} \times \mathbb{R}^n \mid |x-a| \leq r, |y-b| \leq r\}$$

ganz in U liegt. Die Norm der Matrix $\frac{\partial f}{\partial y}$ nimmt auf V ihr Maximum L an. Außerdem gilt, dass V mit zwei Punkten (x,y) und $(x,\tilde{y})$ auch das verbindende Geradenstück enthält. Aus Korollar (1.4.18) folgt nun, dass für alle (x,y) und $(x,\tilde{y}) \in V$ gilt:

$$\|f(x,y) - f(x,\tilde{y})\| \leq L\|y-\tilde{y}\|$$

□

Das man auf die Lipschitz-Bedingung in (2.9.4) nicht verzichten kann, zeigt folgendes Beispiel einer Differentialgleichung in der der Eindeutigkeitssatz nicht gilt:

Beispiel 2.9.6 *Wir betrachten die Differentialgleichung*

$$y' = y^{\frac{2}{3}}$$

also $f(x, y) = y^{2/3}$ *definiert auf* $\mathbb{R} \times \mathbb{R}$. *Eine Lösung ist damit* $\varphi_0 \equiv 0$. *Weitere Lösungen sind gegeben durch:*

$$\psi_0(x) = \frac{1}{27}(x - a)^3$$

Für ein beliebiges $a \in \mathbb{R}$, *denn es gilt:*

$$\psi_a(x) = \frac{1}{9}(x - a)^2 = \psi_0(x)^{2/3}$$

Da $\varphi_0(a) = 0$ und $\psi_a(a) = 0$ ist der Eindeutigkeitssatz verletzt. Durch „Ausstückelung" können wir aber noch mehr Lösungen ψ der Gleichung bekommen, die $\psi(a) = 0$ erfüllen.

Satz 2.9.7 *[Existenzsatz von Picard-Lindelöf] Sei* $U \subseteq \mathbb{R} \times \mathbb{R}^n$ *offen und* $f : U \to \mathbb{R}^n$ *eine stetige Funktion, die lokal einer Lipschitz-Bedingung genügt. Dann gibt es zu jeden* $(a, c) \in U$ *ein* $\varepsilon > 0$ *und eine Lösung*

$$\varphi : [a - \varepsilon, a + \varepsilon] \to \mathbb{R}^n$$

der Differentialgleichung $y' = f(x, y)$ *mit der Anfangsbedingung* $\varphi(a) = c$.

Beweisskizze Man verwendet das Picard-Lindelöfsche Iterationsverfahren

$$T(\varphi)(x) = c + \int_a^x f(t, \varphi(t))dt, \varphi_{n+1} = T(\varphi_n)$$

Für ε genügend klein ist T, eingeschränkt auf eine geeignete Teilmenge des Raumes der stetigen Funktion $f : [a-\varepsilon, a+\varepsilon] \to \mathbb{R}^n$ (versehen mit der Supremumsnorm) eine Kontraktion und die Folge φ_n konvergiert gegen die Lösung der Differentialgleichung. □

Beispiel Man kann das Picard-Lindelöfsche Verfahren benutzen, um Näherungslösungen einer Differentialgleichung zu erhalten. Betrachte z.B. die Differentialgleichung $y' = y$ auf $\mathbb{R} \times \mathbb{R}$.

$$\varphi_0(x) = 0, \varphi_{n+1}(x) = 1 + \int_0^x \varphi_n(t)dt$$

$$\Rightarrow \varphi_1(x) = 1, \varphi_2(x) = 1 + x, \varphi_3(x) = 1 + x + \frac{x^2}{2}$$

$\Rightarrow$ Dies konvergiert gegen die Lösung $\varphi_\infty = e^x = \varphi(x)$ mit Anfangsbedingung $\varphi(0) = 1$. Unter der Schwachen Vorraussetzung, dass die rechte Seite stetig ist, liefert der Satz von Peano die lokale Existenz von Lösungen (diese sind nicht notwendig eindeutig durch eine Anfangsbedingung bestimmt).

Differentialgleichungen höherer Ordnung

Definition 2.9.8 *Sei $U \subseteq \mathbb{R} \times \mathbb{R}^n$ und $f : U \to \mathbb{R}$ eine stetige Funktion. Dann heißt*

$$y^{(n)} = f(x, y, y', ..., y^{(n-1)}) \tag{2.2}$$

eine (gewöhnliche) Differentialgleichung n-ter Ordnung. Unter einer Lösung von (2.2) versteht man eine n-mal differenzierbare Funktion $\varphi : I \to \mathbb{R}$ definiert auf einem reellen Intervall I, so dass der Graph von φ in U enthalten ist und

$$\varphi^{(n)} = f(x, \varphi(x), \varphi'(x), ..., \varphi^{n-1}(x))$$

für alle $x \in I$ gilt.

Man kann eine Differentialgleichung n-ter Ordnung auf ein System von Differentialgleichungen 1. Ordnung zurückführen. Dann definiere:

$$\begin{cases} y_0' & = & y_1 \\ y_1' & = & y_2 \\ & \vdots & \\ y_{n-2}' & = & y_{n-1} \\ y_{n-1}' & = & f(x, y_0, ..., y_{n-1}) \end{cases} \tag{2.3}$$

Man sieht leicht: Ist φ eine Lösung von (2.2) so ist

$$\Phi = (\varphi, \varphi', \varphi'', ...\varphi^{(n-1)})$$

eine Lösung von (2.3) und umgekehrt, ist $\Phi : I \to \mathbb{R}^n$ eine Lösung von (2.3) mit

$$\Phi = (\varphi_0, \varphi_1, ..., \varphi_{n-1})$$

dann ist φ_0 n-mal differenzierbar und eine Lösung von (2.2).

Beispiel 2.9.9 (Aus der Physik - vergleiche (1.7.14))

$$m\ddot{x} = -\nabla U$$

Diese 3 Gleichungen 2. Ordnung sind äquivalent zu:

$$\begin{cases} \dot{x} & = & v \\ \dot{v} & = & -\frac{1}{m}\nabla U \end{cases}$$

was 6 Gleichungen 1. Ordnung entspricht. Dies entspricht dem Übergang zum Phasenraum.

Damit sind die Sätze (2.9.3) und (2.9.7) (Eindeutigkeit und Existenz von Lösungen) auch auf Differentialgleichungen höherer Ordnung anwendbar. Um die Lösung einer Differentialgleichung n-ter Ordnung eindeutig festzulegen, muss man also nicht nur den Wert einer Funktion an einer Stelle a des Definitionsbereiches vorschreiben, sondern den Wert aller Ableitungen der Ordnung $\leq n-1$ im Punkt a.

Beispiel 2.9.10 *Die Differentialgleichung 2. Ordnung*

$$y'' + y = 0$$

besitzt offensichtlich die Lösungen $y(x) = \cos x, y(x) = \sin x, y(x) = e^{ix}$ *aber auch die Linearkombinationen*

$$\varphi(x) = c_0 \cos(x) + c_1 \sin(x)$$

mit $c_0, c_1 \in \mathbb{R}$.

Da $\varphi(0) = c_0, \varphi'(0) = c_1$, ist φ die einzige Lösung mit $\varphi(0) = c_0, \varphi'(0) = c_1$ und man erhält alle Lösungen $y'' + y = 0$, wenn man c_0, c_1 alle Werte in $\mathbb{R}$ durchlaufen lässt.

2.10 Elementare Lösungsmethoden

Differentialgleichungen mit getrennten Variablen Seien I, J offene Intervalle und $f : I \to \mathbb{R}, g : J \to \mathbb{R}$ zwei stetige Funktionen, wobei wir zusätzlich vorraussetzen, dass $g(y) \neq 0 \forall y \in J$. Die Differentialgleichung

$$y' = f(x)g(y) \tag{2.4}$$

auf $I \times J$ heißt dann Differentialgleichung mit getrennten Variablen.
Angenommen $\varphi : I' \to \mathbb{R}$ sei eine Lösung von (2.4) auf dem Teilintervall $I' \subseteq I$ mit $\varphi(x_0) = y_0$. Somit folgt:

$$\frac{\varphi'(x)}{g(\varphi(x))} = f(x) \tag{2.5}$$

Definiere $F : I \to \mathbb{R}, G : J \to \mathbb{R}$ durch

$$F(x) := \int_{x_0}^{x} f(t)dt, G(y) := \int_{y_0}^{y} \frac{1}{g(t)}dt$$

Da der Integrand bei G keine Nullstelle hat, ist G entweder streng monoton wachsend oder streng monoton fallend. Insbesondere hat G eine Umkehrfunktion

$G^{-1} : G(J) \to J$.
Damit folgt aus (2.5) nun

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

Es folgt $G(\varphi(x)) = F(x)$. Also folgt $\varphi(x) = G^{-1}(F(x)) \forall x \in I'$. Dies zeigt die eindeutige Lösbarkeit (und liefert einen expliziten Ausdruck für die Lösung). Lokale Existenz folgt aus dem Satz von Peano, folgt aber auch direkt aus obiger Formel, wenn man das Intervall I' geeignet wählt.

Informell

$$\frac{dy}{dx} = f(x)g(y) \Rightarrow \frac{dy}{g(y)} = f(x)dx \rightsquigarrow \int \frac{dy}{g(y)} = \int f(x)dx + C$$

Beispiel 2.10.1 *Betrachten $y' = y^2$. Die rechte Seite ist in y stetig partiell differenzierbar, also gilt der Eindeutigkeitssatz.*

Über Seperation der Variablen erhält man folgende Ergebnisse:

$$f(x) = 1, g(y) = y^2 \Rightarrow F(x) = x - x_0, G(y) = \int_{y_0}^{y} \frac{1}{t^2} dt = \frac{1}{y_0} - \frac{1}{y}$$

$$y(x) = -\frac{1}{x - x_0 - \frac{1}{y_0}} = \frac{1}{c - x}, c = x_0 + \frac{1}{y_0}$$

Eine weitere spezielle Lösung ist $y \equiv 0$. Es genügt also im Folgenden, Lösungen ohne Nullstellen zu betrachten, da mit $t \mapsto \varphi(t)$ auch $t \mapsto -\varphi(-t)$ eine Lösung ist, können wir uns auf die positive Lösung beschränken.
Insbesondere gibt es außer $y \equiv 0$ bei dieser Gleichung keine auf ganz $\mathbb{R}$ definierte Lösung.

Lineare Differentialgleichungen Sei I ein Intervall und seien $a, b : I \to \mathbb{R}$ stetige Funktionen. Dann nennt man

$$y' = a(x)y + b(x)$$

eine lineare Differentialgleichung und zwar homogen, falls $b \equiv 0$, sonst inhomogen. Wir betrachten zunächst die homogene lineare Differentialgleichung

$$y' = a(x)y \tag{2.6}$$

Eine Lösung ist $\varphi \equiv 0$, wir können die Gleichung durch Trennung der Variablen lösen (wegen Eindeutigkeit können wir ohne Einschränkung $y(x) \neq 0$ annehmen):

$$\frac{y'(x)}{y(x)} = a(x) \Rightarrow \ln |y(x)|' = a(x) \Rightarrow |y(x)| = e^{\int_{x_0}^{x} a(t)dt + \tilde{c}}$$

Das zeigt:

Satz 2.10.2 *Mit den obigen Bezeichnungen gilt: Sei $x_0 \in I$ und $c \in \mathbb{R}$. Dann gibt es genau eine Lösung φ der Linearen homogenen Differentialgleichung (2.6) mit $\varphi(x_0) = c$, nämlich*

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

Beispiel 2.10.3 *Die Lösungen der Differentialgleichung*

$$y' = ky, k \in \mathbb{R}$$

und der Anfangsbedingung $\varphi(x_0) = c$ haben die Gestalt:

$$\varphi(x) = ce^{k(x-x_0)}$$

Wir wollen nun die inhomgene lineare Differentialgleichung

$$y' = a(x)y + b(x)$$

betrachten. Diese löst man mit der sogenannten Variation der Konstante.
Sei $\varphi : I \to \mathbb{R}$ eine Lösung der homogenen Differentialgleichung, das heißt

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

Dann ist mit φ jedes $c \in \mathbb{R}$ auch $x \mapsto c\varphi(x)$ eine Lösung der homogenen linearen Differentialgleichung. Wir machen nun einen Ansatz für die inhomogene Gleichung, indem wir die Konstante c durch eine Funktion $u : I \to \mathbb{R}, x \mapsto u(x)$ ersetzen („Variation der Konstante "), also annehmen, dass eine Lösung ψ der inhomogenen Gleichung gegeben ist durch

$$\psi : x \mapsto \psi(x) = u(x)\varphi(x)$$

wobei φ eine Lösung der homogenen Differentialgleichung und $u : U \to \mathbb{R}$ differenzierbar ist.

$$\begin{aligned}\psi'(x) &= \frac{d}{dx}(u(x)\varphi(x)) = u'(x)\varphi(x) + \underbrace{u(x)\varphi'(x)}_{u(x)a(x)\varphi(x)} = a(x)\psi(x) + b(x) = \\ &= a(x)u(x)\varphi(x) + b(x)\end{aligned}$$

Die Funktion $\psi = u\varphi$ löst also inhomogene lineare Differentialgleichungen genau dann wenn

$$u'(x)\varphi(x) = b(x)$$

Nach dem Eindeutigkeitssatz können wir annehmen, dass φ keine Nullstellen besitzt. Wir erhalten:

$$u(x) = \int_{x_0}^{x} \frac{b(t)}{\varphi(t)}dt + const$$

Durch geeignete Wahl der Integrationskonstanten erreichen wir, dass ψ eine Anfangsbedingung $\psi(x_0) = c$ erfüllt.

Satz 2.10.4 *Sei $I \subseteq \mathbb{R}$ ein Intervall und seien $a, b : I \to \mathbb{R}$ stetige Funktionen. Dann gibt es zu einem beliebigen $x_0 \in I$ und $c \in \mathbb{R}$ genau eine Lösung $\psi : I \to \mathbb{R}$ der Differentialgleichung:*

$$y' = a(x)y + b(x)$$

die der Anfangsbedingung $\psi(x_0) = c$ genügt, nämlich

$$\psi(x) = \varphi(x)(c + \int_{x_0}^{x} \frac{b(t)}{\varphi(t)} dt)$$

Wobei $\varphi(x) = \exp(\int_{x_0}^{x} a(t)dt)$ ist.

Beispiel 2.10.5 *Gesucht sei eine Lösung der Differentialgleichung*

$$y' = 4xy + (1-x)e^{4x}$$

für die $y(1) = 0$ gilt.

Berechnung der homogenen Gleichung ergibt:

$$\varphi(x) = \exp(\int_1^x 4tdt) = e^{2x^2-2}$$

Die Lösung der inhomogenen Differentialgleichung ergibt sich aus (2.10.4):

$$\begin{aligned}
\psi(x) &= \varphi(x)(c + \int_1^x \frac{b(t)}{\varphi(t)} dt) = e^{2x^2-2}(c + \int_1^x (1-t)e^{4t-2t^2+2} dt) = \\
&= e^{ex^2-2}(c + \frac{1}{4}\int_1^x e^{4t-2t^2+2}(4-4t)dt) = e^{2x^2-2}(c + \frac{1}{4}[e^{4t-2t^2+2}]_{t=1}^{t=x}) = \\
&= e^{2x^2-2}(c + \frac{1}{4}e^{4x-2x^2+2} - \frac{1}{4}e^4) \underbrace{=}_{\psi(1)=0 \Rightarrow c=0} \frac{1}{4}(e^{4x} - e^{2x^2+2})
\end{aligned}$$

Exakte Differentialgleichungen und integrierende Faktoren Bei den bis jetzt vorkommenden Beispielen haben wir stets explizit gegebene Lösungen gefunden, also $y(x) = ...$ Bei vielen Differentialgleichungen ist es aber gar nicht möglich, solch eine explizite Lösung anzugeben. Manchmal kann man jedoch wenigstens implizit definierte Funktionen als Lösungen finden (was fast genauso gut ist), also Lösungen etwa einer Differentialgleichung auf $\mathbb{R} \times \mathbb{R}$ deren Graphen auf Höhenlinien einer Funktion

$$h : \mathbb{R}^2 \to \mathbb{R}$$

verlaufen. Sei also $y : \mathbb{R} \to \mathbb{R}$ eine differenzierbare Funktion und $h : \mathbb{R}^2 \to \mathbb{R}$ ebenfalls differenzierbar. Der Graph von y verläuft genau dann in einer Niveaumenge von h, wenn $h(x, y(x)) = \mathit{const.}$ gilt. Nach Kettenregel (1.4.10) ist dies äquivalent zu

$$\frac{d}{dx}h(x, y(x)) = \frac{d}{dx}(h \circ [h \mapsto (x, y(x))]) = \frac{\partial h}{\partial x}(x, y(x)) + \frac{\partial h}{\partial y}(x, y(x))y'(x) = 0$$

Kurz:

$$h_x + h_y y' = 0$$

Definition 2.10.6 *Sei $U \subseteq \mathbb{R} \times \mathbb{R}$ und seien $f, g : U \to \mathbb{R}$ Funktionen. Die Differentialgleichung $f(x,y) + g(x,y)y' = 0$ heißt exakt, falls es eine partiell differenzierbare Funktion $h : U \to \mathbb{R}, (x,y) \mapsto h(x,y)$ gibt, so dass $f(x,y) = \frac{\partial h}{\partial x}(x,y)$ und $g(x,y) = \frac{\partial h}{\partial y}(x,y)$ für alle $(x,y) \in U$ gilt.*

Bemerkung Die obige Differentialgleichung $f + gy' = 0$ wird oft

$$f dx + g dy = 0$$

geschrieben.

Korollar 2.10.7 *Mit Bezeichnungen wie in (2.10.6) gilt: Ist $U \subseteq \mathbb{R} \times \mathbb{R}$ sternförmig und sind f, g stetig differenzierbar, so ist die Differentialgleichung*

$$f + gy' = 0$$

exakt genau dann, wenn $\frac{\partial f}{\partial y}(x,y) = \frac{\partial g}{\partial x}(x,y)$ für alle $(x,y) \in U$ kurz $f_y = g_x$ gilt.

Beweis Folgt direkt aus (1.7.7). □

Beispiel 2.10.8 *Die Differentialgleichung*

$$e^y + x + xe^y y' = 0$$

auf $\mathbb{R} \times \mathbb{R}$ ist exakt, denn:

$$\frac{\partial(e^y + x)}{\partial y} = e^y = \frac{\partial(xe^y)}{\partial y}$$

Eine Funktion $h : \mathbb{R}^2 \to \mathbb{R}, (x,y) \mapsto h(x,y)$ mit $h_x = f$ und $h_y = g$ ist gegeben durch:

$$h(x,y) = xe^y + \frac{1}{2}x^2$$

Eine Lösung der Differentialgleichung erhalten wir aus

$$ye^{y(x)} + \frac{1}{2}x^2 = c, c \in \mathbb{R}$$

Zum Beispiel mit der Anfangsbedingung $\varphi(1) = 0 \Rightarrow c = \frac{3}{2}$. Also für x nahe 1 gilt:

$$e^{y(x)} = \frac{c - \frac{1}{2}x^2}{x} = \frac{1}{2}(\frac{3}{x} - x) \Rightarrow y(x) = \ln(\frac{1}{2}(\frac{3}{x} - x))$$

In diesem Fall kann man sogar explizit nach $y(x)$ auflösen. Allgemein nennt man eine solche Funktion h oft (etwas altmodisch) ein Integral der Differentialgleichung (früher: „Integral" ist eine Lösung einer Differentialgleichung, Flächenbestimmung ist „Quadratur").
Falls $f + gy' = 0$ nicht exakt ist, muss sie mit Hilfe eines sog. integrierenden Faktors exakt gemacht werden, was in der Praxis oft möglich ist.

Definition 2.10.9 *Mit den Bezeichnungen wie in (2.10.6). Eine Überall auf U von Null verschiedenen Funktion $\mu : U \to \mathbb{R}$ heißt integrierender Faktor (auch Eulerscher Multiplikator) für Differentialgleichungen $f + gy' = 0$, wenn die dann äquivalente Differentialgleichungen $\mu f + \mu g y' = 0$ exakt ist, wenn also*

$$(\mu f)_y = \mu_y f + \mu f_y = \mu_x g + \mu g_x = (\mu g)_x$$

gilt.

Das Problem dabei ist, dass wir, um so einen integrierenden Faktor zu finden die partiellen Differentialgleichungen $(\mu f)_y = (\mu g)_x$ lösen müssen. In der Praxis findet man oft Lösungen, indem man für μ einen speziellen Ansatz macht (z.B. μ hängt nur von x ab, oder nur von y, oder $\mu(x, y) = f(x, y)$ oder $\mu(x, y) = f(x)g(y)$ oder $\mu(x, y) = x^k y^l$... etc.). Es genügt dabei, eine Lösung zu finden.

Beispiel 2.10.10 *Die Gleichung $(4xy + 3y^2 - x) + x(x + 2y)y' = 0$ auf $\mathbb{R} \times \mathbb{R}$. Hier gilt $f(x, y) = 4xy + 3y^2 - x$ und $g(x, y) = x(x+2y)$. Die Gleichung ist nicht exakt, denn*

$$f_y = 4x + 6y \neq 2x + 2y = g_x$$

Wir wollen versuchen, einen integrierenden Faktor μ zu finden. Die Funktion μ ist genau dann ein integrierender Faktor, wenn

$$\mu_y f + \mu f_y = \mu_x g + \mu g_x$$

gilt oder

$$\mu(f_y - g_x) = \mu_x g - \mu_y f$$

Wir versuchen es jetzt mit dem Ansatz, dass μ nur von x abhängen soll. Dann lautet die Gleichung für μ:

$$\frac{\mu_x}{\mu} = \frac{f_y - g_x}{g} = \frac{2x + 4y}{x(x + 2y)} = \frac{2}{x}$$

Folglich lautet eine Lösung

$$\mu = \exp(\int \frac{2}{x} dx) = \exp(2 \ln |x|) = x^2$$

Mit dem integrierenden Faktor multipliziert, lautet die Gleichung nun

$$(4x^3 y + 3x^2 y^2 - x^3) + (x^4 + 2x^3 y)y' = 0$$

Eine Funktion $h : \mathbb{R}^2 \to \mathbb{R}$ mit $h_x = f, h_y = g$ ist gegeben durch

$$h(x,y) = x^4 y + x^3 y^2 - \frac{1}{4} x^4$$

Lösungen der ursprünglichen Differentialgleichung sind also implizit gegeben durch

$$x^4 y + x^3 y^2 - \frac{1}{4} x^4 = c, c \in \mathbb{R}$$

Substitution Wenn sich eine Differentialgleichung mit den besprochenen Methoden nicht lösen lässt, kann man sie oft durch eine Substitution (Koordinatentransformation) in eine lösbare Form bringen, das heißt durch Einfürung neuer Variablen. Die „Kunst" dabei ist es, die richtige Substitution durchzuführen, also eine die die Gleichung vereinfacht. Manchmal kann man an der Form der Gleichung erkennen, welche Substitutionen sinnvoll sein könnte. Tritt etwa ein Term mehrfach auf, lohnt es sich zu überlegen, ob man diesen Term als neue Variable einführen will.

Beispiel 2.10.11 *In der Gleichung*

$$3(2x+y) + (2x+y-1)y' = 0$$

tritt der Term $(2x+y)$ *zweifach auf. Setze* $v := 2x + y \Leftrightarrow y = v - 2x$ *mit* $y' = v' - 2$*. Die Gleichung lässt sich nun schreiben als*

$$\begin{aligned} 3v + (v-1)(v'-2) &= 0 \Leftrightarrow v + vv' - v' + 2 = 0 \Leftrightarrow (v-1)v' + v + 2 = 0 \Leftrightarrow \\ &\Leftrightarrow \frac{dv}{dx} = -\frac{(v+2)}{(v-1)} = \frac{v+2}{1-v} \end{aligned}$$

Die Gleichung liegt nun in getrennten Variablen vor

$$\frac{v-1}{v+2} v' = -1 \Leftrightarrow v - 3\ln|v+2| = -x + c$$

Jetzt machen wir die Substitution wieder Rückgängig und erhalten

$$2x + y - 3\ln|2x+y+2| = -x + c \Leftrightarrow 3x + y = 3\ln|2x+y+2| + c$$

2.11 Lineare Differentialgleichungen

Definition 2.11.1 *Sei* $I \subseteq \mathbb{R}$ *ein Intervall und sei* $A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix}, I \to Mat(n,n;\mathbb{R})$ *eine stetige Abbildung, das heißt alle Funktionen* a_{ij} *stetig sind.*

Dann heißt $y' = A(x)y$ ein homogenes lineares Differentialgleichungssystem. Sei weiter $b = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} : I \to \mathbb{R}^n$ eine stetige Abbildung. Dann heißt

$$y' = A(x) + b(x)$$

ein inhomogenes lineares Differentialgleichungssystem.

Wir wollen nun auch komplexe Differentialgleichungen zulassen. Ein komplex lineares Differentialgleichungssystem hat die Form

$$y' = A(x)y + b(x)$$

wobei $A : I \to Mat(n, n; \mathbb{C})$ und $b : I \to \mathbb{C}^n$ stetige Abbildungen sind. Eine Lösung ist eine differenzierbare Funktion $\varphi : I \to \mathbb{C}^n$, so dass $\varphi'(x) = A(x)\varphi(x) + b(x)$ gilt, wobei Differenzierbarkeit für Abbildungen nach $\mathbb{C}$ Komponentenweise zu verstehen ist, das heißt Real- und Imaginärteil sollen beide differenzierbar sein. Wir schreiben ab jetzt $\mathbb{K}$ für einen der Körper $\mathbb{R}$ oder $\mathbb{C}$.

Satz 2.11.2 *Sei $I \subseteq \mathbb{R}$ ein offenes Intervall und seien $A : I \to Mat(n, n; \mathbb{K}), b : I \to \mathbb{K}^n$ stetige Abbildungen. Dann gibt es zu jedem $x_0 \in I$ und $c \in \mathbb{K}^n$ genau eine Lösung $\varphi : I \to \mathbb{K}^n, y' = A(x)y + b(x)$ mit der Anfangsbedingung $\varphi(x_0) = c$.*

Bemerkung Wir haben schon in Beispiel (2.10.1) gesehen, dass im Allgemeinen eine Differentialgleichung $y' = f(x, y)$ keine Lösung auf dem ganzen Intervall I zu haben braucht, auch wenn f auf ganz $I \times \mathbb{R}^n$ definiert ist (und einer Lipschitz-Bedingung genügt).

Beweis Wir zeigen nun die Eindeutig, diese gilt, da die rechte Seite einer lokalen Lipschitz-Bedingung genügt auf einem kompakten Teilintervall $J \subseteq I$ nimmt $\|A(x)\|$ sein Maximum L an und es folgt

$$\|f(x, y) - f(x, \tilde{y})\| = \|A(x)(y - \tilde{y})\| \leq L\|y - \tilde{y}\|$$

$\square$

Satz 2.11.3 *Sei $I \subseteq \mathbb{R}$ ein Intervall und $A : I \to Mat(n, n; \mathbb{K})$ stetig. Wir bezeichnen mit L_H die Menge der Lösungen $\varphi : I \to \mathbb{K}^n$ der homogenen lineare Differentialgleichung:*

$$y' = A(x)y$$

Dann ist L_H ein Vektorraum der Dimension n über $\mathbb{K}$. Für ein k-Tupel $\varphi_1, ..., \varphi_k$, $k \leq n$ von Lösungen sind folgende Aussagen äquivalent:

1. *$\varphi_1, ..., \varphi_k$ sind als Funktionen linear unabhängig über $\mathbb{K}$.*

2. *Es existiert ein $x_0 \in I$, so dass die Vektoren $\varphi_1(x_0), ..., \varphi_k(x_0) \in \mathbb{K}^n$ linear unabhängig über $\mathbb{K}$ sind.*

3. *Für jedes $x_0 \in I$ sind die Vektoren $\varphi_1(x_0), ...\varphi_k(x_0) \in \mathbb{K}^n$ linear unabhängig.*

Beweis (a) Zeige: L_H ist Vektorraum. Wir zeigen, dass L_H ein Untervektorraum des Raumes aller Abbildung $I \to \mathbb{K}^n$ ist. Dies gilt, da

1. $0 \in L_H$

2. $\varphi, \psi \in L_H \Leftrightarrow (\varphi + \psi) \in L_H \Rightarrow (\varphi + \psi)' = \varphi' + \psi' = A(\varphi + \psi)$

3. $\lambda \in \mathbb{K}, \varphi \in L_H \Rightarrow (\lambda\varphi)' = \lambda\varphi' = A\lambda\varphi \Rightarrow \lambda\varphi \in L_H$

(b) Äquivalenz von (1),(2),(3). Zeige nur (1)$\Rightarrow$(3), da die anderen direkt daraus folgen (trivial sind).
Seien also $\varphi_1, ..., \varphi_k$ linear unabhängig. Angenommen es gäbe $x_0 \in I$, mit $\varphi_1(x_0), ..., \varphi_k(x_0)$ linear abhängig, das heißt es gäbe $\lambda_1, ..., \lambda_k \in \mathbb{K}, \lambda_1\varphi_1(x_0) + ... + \lambda_k\varphi_k(x_0) = 0$ wobei nicht alle λ_i gleich Null sind. Betrachte nun die Lösung

$$\varphi := \lambda_1\varphi_1 + ... + \lambda_k\varphi_k$$

für die gilt: $\varphi(x_0) = 0$. Wegen der Eindeutigkeit muss φ also die Nulllösung $\varphi \equiv 0$ sein, das heißt $\varphi_1, ..., \varphi_k$ sind linear abhängig über $\mathbb{K}$. Somit folgt ein Widerspruch!
(c) Zeige: $\dim L_H = n$. Seien $e_1, ..., e_n \in \mathbb{K}^n$ die kanonischen Basisvektoren. Nach (2.10.2) gibt es Lösungen $\varphi_1, ..., \varphi_n$ mit $\varphi_i(x_0) = e_i$. Diese Lösungen sind linear unabhängig, also $\dim L_H \geq n$. Andererseits ist $\dim L_H \leq n$, denn gäbe es $n + 1$ linear unabhängige Lösungen $\psi_1, ..., \psi_{n+1}$ so wären (wegen (1)$\Leftrightarrow$(3)) die Vektoren $\psi_1(x_0), ..., \psi_{n+1}(x_0) \in \mathbb{K}^n$ linear unabhängig. Daraus folgt ein Widerspruch! $\square$