Script
zur
Vorlesung

Mathematik für Physiker III
– Differentialrechnung –

Prof. Dr. Frank Loose WS 2004/2005

Eberhard-Karls-Universität Tübingen

Florian Jessen

Jessen@pit.physik.uni-tuebingen.de

14. März 2009

Made with $\LaTeX 2\varepsilon$

Vorwort

Dieses Script entsteht während der Vorlesung „Mathematik für Physiker III“ des laufenden Semesters (WS2004/2005). Es wird dabei nicht garantiert, dass sich keine Fehler, sowohl in Formeln, als auch in den Formulierungen von Definitionen, Sätzen oder auch Beweisen einschleichen. Der Leser sollte daher entsprechend kritisch damit umgehen und gefundene Fehler weitergeben, damit diese korrigiert werden können.

Wer die schönen Bilder aus der Vorlesung vermissen sollte, ist hiermit aufgerufen diese in digitaler Form (png, ps, o.ä.) zu erstellen, damit diese integriert werden können. Gleiches gilt für anderweitige Mitarbeit am Gesamtwerk. Kenntnisse im Umgang mit LaTeX2ε sind dabei hilfreich.

In der Zwischenzeit wurde das Script von Herrn Prof. Loose korrekturgelesen. Die dabei entdeckten Fehler sind korrigert – womit allerdings weitere Fehler nicht ausgeschlossen sind. Auch wurden einige Stellen mit dem Hinweis ✎ *Übung* ✐ versehen. Hier sind jeweils kurze, beinahe triviale oder aber auch besonders lange Beweise ausgelassen, die im Selbststudium nachvollzogen werden müssen. Teilweise werden diese Aufgaben auch in den Übungen behandelt. Verweise auf Sätze und Definitionen aus anderen Scripten der Serie, die hier ebenfalls für die Beweise verwendet werden, sind mit MfPh... markiert.

Inhaltsverzeichnis

Kapitel 1

Metrische Räume

19.10.2004

Motivation

Will man Begriffe wie Stetigkeit, Differenzierbarkeit, Integrierbarkeit, ... für Funktionen $f(x)$ mehrerer reeller Veränderlicher $x = (x_1, \dots, x_n) \in \mathbb{R}^n (n \in \mathbb{N} = 1, 2, 3, ...)$ einführen, so ist zuerst nicht klar, was die richtige Verallgemeinerung der offenen Intervalle $(a, b) \subseteq \mathbb{R}$ und der abgeschlossenen Intervalle $[a, b] \subseteq \mathbb{R}$ ist. In diesem Paragraphen werden daher Begriffe wie *offen*, *abgeschlossen*, *Randpunkt* u.ä. allgemein für *metrische Räume* eingeführt und später zunächst auf Teilmengen des $\mathbb{R}^n (n \in \mathbb{N})$, oder auch auf Teilmengen von *Funktionenräumen* angewendet.

1.1 Metriken

Definition 1.1.
Sei X eine Menge und $d : X \times X \to [0, \infty), (x, y) \mapsto d(x, y)$ eine Abbildung. Dann heißt d eine Metrik auf X, wenn für alle $x, y, z \in X$ gilt:

(a) $d(x, y) = 0$ genau dann wenn $x = y$

(b) $d(x, y) = d(y, x)$ *Symmetrie*

(c) $d(x, y) \leq d(x, z) + d(z, y)$ *Dreiecksungleichung*

Ein Paar (X, d) heißt *metrischer Raum*, wenn X eine Menge und d eine Metrik auf X ist.

1.1.1 Beispiele

(I) Sei $X = \mathbb{R}^n (n \in \mathbb{N})$ und für $x, y \in \mathbb{R}^n$ sei

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

Es ist offenbar $d(x, y) \geq 0$, $d(x, y) = 0$ genau dann wenn $x = y$ und $d(x, y) = d(y, x)$.
Die Dreiecksungleichung ist nicht sofort offensichtlich, folgt jedoch aus 1.2.4.
Damit ist d eine Metrik auf $\mathbb{R}^n$ und heißt *Euklidische Metrik* (auf $\mathbb{R}^n$). Es wird $(\mathbb{R}^n, d)$ als der *Euklidische Raum* bezeichnet.

(II) Speziell für $n = 1$ erhält man für $x, y \in \mathbb{R}$:

$$d(x, y) = \sqrt{(x - y)^2} = |x - y| \ .$$

Hier ist die Dreiecksungleichung klar.

(III) Ist (X, d) ein metrischer Raum und $A \subseteq X$ eine Teilmenge, so ist

$$d_A : A \times A \to [0, \infty) \ , \quad d_A := d|_{A \times A}$$

eine Metrik auf A. Sie heißt die *von d induzierte Metrik*. So wird beispielsweise jede Teilmenge $A \subseteq \mathbb{R}^n$ mit der von der Euklidischen Metrik induzierten Metrik zu einem metrischen Raum.

(IV) Sei X eine beliebige Menge. Setzt man

$$d : X \times X \to [0, \infty) \ ,$$

$$d(x, y) := \begin{cases} 0 & \text{für } x = y \\ 1 & \text{für } x \neq y \end{cases} \ ,$$

so ist d eine Metrik auf X. ✎ *Übung* ✐

1.2 Normen

Definition 1.2.
Sei V ein reeller Vektorraum. Eine Abbildung $\|.\| : V \to \mathbb{R}\,, \quad x \mapsto \|x\|$ heißt eine *Norm auf V*, wenn für alle $x, y \in V$ und $\lambda \in \mathbb{R}$ gilt:

(a) $\|x\| \geq 0$ und $\|x\| = 0$ genau wenn $x = 0$

(b) $\|\lambda x\| = |\lambda| \cdot \|x\|$

(c) $\|x + y\| \leq \|x\| + \|y\|$ *Dreiecksungleichung*

Ein Paar $(V, \|.\|)$ bestehend aus einem Vektorraum V mit einer Norm $\|.\|$ heißt ein *normierter (Vektor-)Raum*.

1.2.1 Bemerkungen

Sei $(V, \|.\|)$ ein normierter Raum. Dann wird durch

$$d : V \times V \to [0, \infty) \ , \quad d(x, y) := \|x - y\|$$

eine Metrik auf V definiert. (Sie heißt die *von der Norm induzierte Metrik auf V*.)

¡? Beweis ¿!

(a) $d(x, y) = 0 \overset{\text{Def!}}{\iff} \|x - y\| = 0 \overset{\text{Def 1.2 (a)}}{\iff} x - y = 0 \iff x = y$

(b) $d(y, x) \overset{\text{Def!}}{=} \|y - x\| = \|(-1)(x - y)\| \overset{\text{Def 1.2 (b)}}{=} |-1| \, \|x - y\| = \|x - y\| = d(x, y)$

(c) $d(x, z) = \|x - z\| = \|x - y + y - z\| \overset{\text{Def 1.2 (c)}}{\leq} \|x - y\| + \|y - z\| = d(x, y) + d(y, z)$

QED.

1.2.2 Beispiele

(I) Für $V = \mathbb{R}^n (n \in \mathbb{N})$ sei für $x \in \mathbb{R}^n$ definiert:

$$\|x\| := \sqrt{x_1^2 + \dots + x_n^2}\ .$$

Es ist offenbar $\|x\| \geq 0$ und $\|x\| = 0$ genau wenn $x = 0$,

$$\|\lambda x\| = \sqrt{(\lambda x_1)^2 + \dots + (\lambda x_n)^2} = \sqrt{\lambda^2 (x_1^2 + \dots + x_n^2)} = \sqrt{\lambda^2}\sqrt{x_1 + \dots + x_n} = |\lambda|\,\|x\|$$

und die Dreiecksungleichung folgt aus 1.2.4. Damit ist $\|.\|$ eine Norm auf $\mathbb{R}^n$ und heißt *Euklidische Norm* (auf $\mathbb{R}^n$). Die von ihr induzierte Metrik ist die *Euklidische Metrik*.

(II) Für $V = \mathbb{R}^n$ $(n \in \mathbb{N})$ setze man

$$\|x\| := \max\{|x_1|, |x_2|, \dots, |x_n|\}\ .$$

Es ist dann auch $\|.\|$ eine Norm auf $\mathbb{R}^n$ und heißt *Maximumsnorm* (auf $\mathbb{R}^n$).

(III) Sei X eine beliebige Menge und V der Vektorraum der beschränkten, reellwertigen Funktionen auf X,

$$V = \left\{ f : X \to \mathbb{R} \;\middle|\; \sup_{x \in X} |f(x)| < \infty \right\}\ .$$

Dann ist mit

$$\|f\| := \sup |f(x)|$$

$(V, \|.\|)$ ein normierter Raum.

✎ *Übung* ✐

21.10.2004

1.2.3 Erinnere

Ist V ein reeller Vektorraum, so heißt eine Abbildung $\langle\,\cdot\,,\,\cdot\,\rangle : V \times V \to \mathbb{R}$ ein *Skalarprodukt auf V*, wenn für alle $x, y, z \in V$ und $\lambda, \mu \in \mathbb{R}$ gilt:

(I) $\langle \lambda x + \mu y, z\rangle = \lambda \langle x, z\rangle + \mu \langle y, z\rangle$
$\langle x, \lambda y + \mu z\rangle = \lambda \langle x, y\rangle + \mu \langle x, z\rangle$ *Bilinearität*

(II) $\langle x, y\rangle = \langle y, x\rangle$ *Symmetrie*

(III) $\langle x, x\rangle \geq 0 \;\; \forall x \in V$ und $\langle x, x\rangle = 0$ genau wenn $x = 0$ ist *positive Definitheit*

Ein Paar $(V, \langle\,\cdot\,,\,\cdot\,\rangle)$ bestehend aus einem Vektorraum V und einem Skalarprodukt $\langle\,\cdot, \cdot\rangle$ heißt ein *Euklidischer Vektorraum.*
Ist nun $(V, \langle\,\cdot\,,\,\cdot\,\rangle$ ein Euklidischer Vektorraum, so setzt man für jedes $x \in V$

$$\|x\| := \sqrt{\langle x, x\rangle}\ .$$

Es gilt dann für alle $x, y \in V$ die *Ungleichung von Cauchy-Schwarz*

$$|\langle x, y\rangle| \leq \|x\| \cdot \|y\|\ \ .$$

1.2.4 Bemerkungen

Ist $(V, \langle\,\cdot\,,\,\cdot\,\rangle)$ ein Euklidischer Vektorraum, so wird durch $\|x\| := \sqrt{\langle x, x\rangle}$ eine Norm auf V definiert.

¡? Beweis ¿!

- $\|x\| = 0 \Longleftrightarrow \sqrt{\langle x, x\rangle} = 0 \Longleftrightarrow \langle x, x\rangle = 0 \overset{1.2.3\,(\mathrm{III})}{\Longleftrightarrow} x = 0$
- $\|\lambda x\| = \sqrt{\langle \lambda x, \lambda y\rangle} = \sqrt{\lambda^2 \langle x, x\rangle} = \sqrt{\lambda^2}\sqrt{\langle x, x\rangle} = |\lambda|\,\|x\|$
- $\|x+y\|^2 = \langle x+y, x+y\rangle = \langle x, x\rangle + \langle x, y\rangle + \langle y, x\rangle + \langle y, y\rangle$
  $= \|x\|^2 + 2\cdot\langle x, y\rangle + \|y\|^2 \leq \|x\|^2 + 2\,\|x\|\,\|y\| + \|y\|^2 = (\|x\| + \|y\|)^2$

QED.

1.2.5 Beispiele

(I) Auf $V = \mathbb{R}^n$ wird durch

$$\langle x, y\rangle = x_1 y_1 + x_2 y_2 + \cdots + x_n y_n$$

ein Skalarprodukt gegeben, das so genannte *kanonische Skalarprodukt*. Die von ihm induzierte Norm 1.2.3 ist die Euklidische Norm. (Sie ist also nach 1.2.4 eine Norm und damit nach Definition 1.2. die Euklidische Metrik auch tatsächlich eine Metrik.)

(II) Sei V der $\mathbb{R}$-Vektorraum der quadrat-summierbaren, reellen Folgen, d.h.

$$V = \left\{ (a_n)_{n\in\mathbb{N}} \;\middle|\; \sum_{n=1}^{\infty} a_n^2 < \infty \right\} .$$

Es konvergiert dann für jedes $(a_n), (b_n) \in V$ auch die folgende Reihe $\sum\limits_{n=1}^{\infty} a_n b_n$ absolut, denn wegen $(a_n \pm b_n)^2 \geq 0$ (also $|a_n b_n| \leq \frac{1}{2}a_n^2 + \frac{1}{2}b_n^2$) ist:

$$\sum_{n=1}^{\infty} |a_n b_n| \leq \sum \frac{1}{2}a_n^2 + \frac{1}{2}b_n^2 = \frac{1}{2}\sum a_n^2 + \frac{1}{2}\sum b_n^2 < \infty$$

Es ist daher $\langle\,\cdot\,,\,\cdot\,\rangle : V \times V \to \mathbb{R}$,

$$\langle (a_n), (b_n)\rangle := \sum_{n=1}^{\infty} a_n b_n$$

ein Skalarprodukt auf V und macht $l^2 := (V, \langle\,\cdot\,,\,\cdot\,\rangle)$ zu einem Euklidischen Vektorraum.

1.2.6 Kommentar

Ist V ein reeller Vektorraum, so induziert

$$\text{Skalarprodukt } \langle\,\cdot\,,\,\cdot\,\rangle \quad \hookrightarrow \quad \text{Norm } \|.\| \quad \hookrightarrow \quad \text{Metrik } d$$

1.3 Offene und abgeschlossene Bälle und Mengen

Definition 1.3.
Sei (X, d) ein metrischer Raum.

(a) Für $x_0 \in X$ und $r > 0$ heißt dann

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

der *(offene) Ball um* x_0 *mit Radius* r.

(b) Eine Teilmenge $U \subseteq X$ heißt *offen*, wenn es für jeden Punkt $x \in U$ ein $r > 0$ gibt, so dass $B_r(x) \subseteq U$ ist.

1.3.1 Beispiele

(I) Der (offene) Ball $B_r(x_0) \subseteq X$ ist selbst eine offene Menge, denn ist $x \in B_r(x_0)$ beliebig, so setze $\varepsilon := r - d(x, x_0) > 0$. Es ist dann wegen der Dreiecksungleichung $B_\varepsilon(x) \subseteq B_r(x_0)$, weil für $y \in B_\varepsilon(x)$ gilt:

$$\begin{aligned} d(y, x_0) &\leq d(y, x) + d(x, x_0) \\ &< \varepsilon + d(x, x_0) \\ &= r - d(x, x_0) + d(x, x_0) \\ &= r \end{aligned}$$

(II) Ein offenes Intervall $(a, b) \subseteq \mathbb{R}$ ist bezüglich $(\mathbb{R}, |.|)$ eine offene Menge.

(III) Ist X eine beliebige Menge versehen mit der *diskreten Metrik* 1.1.1 (IV), so ist dort <u>jede</u> Teilmenge von X offen! ✎ *Übung* ✐

Definition 1.4.
Sei (X, d) ein metrischer Raum und $x_0 \in X$. Eine Teilmenge $S \subseteq X$ heißt eine *Umgebung von* x_0, wenn es eine offene Menge $U \subseteq X$ gibt mit

$$x_0 \in U \subseteq S \ .$$

1.3.2 Bemerkungen

Eine Teilmenge $U \subseteq X$ ist genau dann offen, wenn sie Umgebung aller ihrer Punkte ist.

<u>¡? Beweis ¿!</u>

$\Rightarrow$ Sei $U \subseteq X$ offen und $x \in U$ beliebig. Dann ist $x \in U \subseteq U$, also ist U Umgebung von x.

$\Leftarrow$ Sei U Umgebung aller ihrer Punkte und $x \in U$ beliebig. $\Longrightarrow \exists V \subseteq X$ offen mit $x \in V \subseteq U \overset{V \text{ offen}}{\Longrightarrow} \exists r > 0 : B_r(x) \subseteq V \subseteq U$, also ist U offen.

QED.

1.3.3 Bemerkungen

Sei (X, d) ein metrischer Raum. Dann gilt:

(1) $\emptyset$, X sind offen

(2) sind $U, V \subseteq X$ offen, so auch $U \cap V$

(3) sind $U_i \subseteq X$ offen $(i \in I)$, so ist auch $\bigcup_{i\in I} U_i$ offen.

¡? Beweis ¿!

- i.O.
- Seien U, V offen und $x \in U \cap V \Longrightarrow \exists r > 0 : B_r(x) \subseteq U$; $\exists s > 0 : B_s(x) \subseteq V$. Setze dann $\varepsilon := \min(r, s) > 0 \Longrightarrow B_\varepsilon(x) \subseteq U \cap V$, also ist $U \cap V$ offen.
- Seien U_i $(i \in I)$ offen und $x \in \bigcup_{i\in I} U_i =: U \Longrightarrow \exists i_0 \in I : x \in U_{i_0} \Longrightarrow \exists r > 0 : B_r(x) \subseteq U_{i_0} \subseteq U$, also ist U offen.

QED.

1.3.4 Kommentar

Beachte, dass beliebige Durchschnitte offener Mengen i.a. nicht offen zu sein brauchen. Z.B. ist $I_n = \left(-\frac{1}{n} + \frac{1}{n}\right) \subseteq \mathbb{R}$ offen $\forall n \in \mathbb{N}$, aber $\{0\} = \bigcap_{n=1}^{\infty} I_n$ ist es nicht.

Definition 1.5.
Sei X ein metrischer Raum. Eine *Teilmenge* $A \subseteq X$ heißt *abgeschlossen*, wenn ihr Komplement $X \backslash A$ offen ist.

26.10.2004

1.3.5 Kommentar

(1) Man beachte, dass i.a. eine Teilmenge $Y \subseteq X$ eines metrischen Raumes X weder offen noch abgeschlossen ist, z.B. ist $(a, b] \subseteq \mathbb{R}$ weder offen noch abgeschlossen.

(2) Es kann auch sein, dass eine Teilmenge $Y \subseteq X$ sowohl offen, als auch abgeschlossen ist, z.B. ist nach 1.3.3 (1) $\emptyset$, X offen und abgeschlossen.

(3) Durch Komplementbildung (und die de-Morgan-Regeln) erhält man aus 1.3.3:

  (i) Endliche Vereinigungen abgeschlossener Mengen sind abgeschlossen

  (ii) Beliebige Durchschnitte abgeschlossener Mengen sind abgeschlossen

Definition 1.6.
Sei X ein metrischer Raum und $Y \subseteq X$ eine beliebige Teilmenge. Es heißt

(a)
$$\overset{\circ}{Y} = \bigcup_{U \text{ offen}, U \subseteq Y} U$$

der *offene Kern von* Y (das *Innere von* Y)

(b)
$$\bar{Y} = \bigcap_{A \text{ abgeschl.}, A \supseteq Y} A$$

die *abgeschlossene Hülle von* Y (der *Abschluss von* Y)

1.3.6 Kommentar

Es ist $\overset{\circ}{Y} \subseteq Y \subseteq \bar{Y}$. Weiter ist $\overset{\circ}{Y}$ die größte offene Teilmenge, die in Y enthalten ist und $\bar{Y}$ die kleinste abgeschlossene Teilmenge, die Y enthält. Y ist offen, genau wenn $Y = \overset{\circ}{Y}$ ist und Y abgeschlossen, genau wenn $Y = \bar{Y}$ ist.

1.4 Konvergente Folgen

Definition 1.7.
Sei (X, d) ein metrischer Raum. Eine Folge $(x_n)_{n \in \mathbb{N}}$ in X heißt *konvergent gegen* $a \in X$, geschrieben

$$(x_n) \longrightarrow a$$

oder

$$\lim_{n \to \infty} (x_n) = a \ ,$$

wenn gilt: Für jede Umgebung S von a gibt es ein $N \in \mathbb{N}$, so dass für alle $n \geq N$ gilt : $x_n \in S$.

1.4.1 Kommentar

Eine äquivalente Definition ist offenbar durch folgende gegeben:

$$\lim_{n \to \infty} (x_n) = a \Leftrightarrow \forall \varepsilon > 0 \ \ \exists N \in \mathbb{N} \ \ \forall n \geq N : d(x_n, a) < \varepsilon \ .$$

✎ *Übung* ✐

1.4.2 Bemerkung

Sei X ein metrischer Raum. Dann gilt: eine Teilmenge A in X ist genau dann abgeschlossen, wenn für jede Folge (x_n) in A, die einen Grenzwert $a = \lim_{n \to \infty} (x_n)$ in X hat, notwendig $a \in A$ sein muss.

¡? Beweis ¿!

$\Rightarrow$ Sei A abgeschlossen, $x_n \in A \ \ (n \in \mathbb{N})$ und $a = \lim(x_n) \in X$. Angenommen $a \notin A$. Da $X \backslash A$ dann Umgebung von a ist (denn $X \backslash A$ ist offen), folgt: $x_n \in X \backslash A$ für fast alle $n \in \mathbb{N}$.
↯ Widerspruch ↯
Also doch $a \in A$.

$\Leftarrow$ Sei $a \in X \backslash A$. Angenommen $\forall n \in \mathbb{N} : B_{1/n}(a) \cap A \neq \emptyset$. Wähle dann $x_n \in B_{1/n}(a) \cap A \subseteq A$. Es ist dann $\lim(x_n) = a$, also muss $a \in A$ sein.✎ *Übung* ✐
↯ Widerspruch ↯
Also $\exists n \in \mathbb{N} : B_{1/n}(a) \cap A = \emptyset$, also ist $B_{1/n}(a) \subseteq X \backslash A$. Also ist $X \backslash A$ offen, d.h. A ist abgeschlossen.

QED.

Definition 1.8.
Sei X ein metrischer Raum und $Y \subseteq X$ eine Teilmenge. Man nennt $x \in X$ einen *Randpunkt von* Y, geschrieben $x \in \partial Y$, wenn jede Umgebung $S \subseteq X$ von x sowohl Y als auch $X \backslash Y$ schneidet

$$S \cap Y \neq \emptyset \ , \quad S \cap (X \backslash Y) \neq \emptyset$$

1.5 Cauchy-Folgen

Definition 1.9.
Sei X ein metrischer Raum. Eine Folge (x_n) in X heißt eine *Cauchy-Folge*, wenn es zu jedem $\varepsilon > 0$ ein $N \in \mathbb{N}$ gibt, so dass für alle $n, m \geq N$ gilt

$$d(x_n, x_m) < \varepsilon$$

1.5.1 Kommentar

Ist (x_n) eine konvergente Folge und $a = \lim(x_n)$, so ist (x_n) notwendig eine Cauchy-Folge, denn ist $\varepsilon > 0$ beliebig, so gibt es ein $N \in \mathbb{N}$, so dass für alle $n \geq N$ gilt $d(x_n, a) < \frac{\varepsilon}{2}$. Dann gilt für alle $n, m \geq N$ nach der Dreiecksungleichung

$$d(x_n, x_m) \leq d(x_n, a) + d(a, x_m) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon \; .$$

Umgekehrt muss aber eine Cauchy-Folge in einem metrischen Raum nicht notwendigerweise konvergent sein, z.B. ist in $X = \mathbb{R}\backslash\{0\}$ $x_n = \frac{1}{n}$ sicher eine Cauchy-Folge, hat aber keinen Grenzwert in X.

Definition 1.10.
(a) Ein metrischer Raum (X, d) heißt *vollständig*, wenn in ihm jede Cauchy-Folge konvergiert.

(b) Ein normierter Raum $(V, \|.\|)$ heißt ein *Banachraum*, wenn V zusammen mit der induzierten Metrik vollständig ist.

1.5.2 Kommentar

Zwei Normen $\|.\|_1$ und $\|.\|_2$ auf einem Vektorraum V heißen äquivalent, wenn es Konstanten $a, b > 0$ gibt, so dass für alle $x \in V$ gilt:

$$a \left\|x\right\|_1 \leq \left\|x\right\|_2 \leq b \left\|x\right\|_1 \; .$$

$(V, \|.\|_1)$ ist dann ein Banachraum, genau dann wenn $(V, \|.\|_2)$ einer ist. Zwei äquivalente Normen erzeugen (über die induzierten Metriken d_1 und d_2) stets die gleiche Topologie auf X, d.h. die gleichen offenen Mengen.

28.10.2004 ✎ *Übung* ✎

1.5.3 Bemerkung

Der Euklidische Raum $(\mathbb{R}^n, d)$ ist vollständig.

¡? Beweis ¿!

Die Euklidische Norm $\|.\|$ und die Maximumsnorm $\|.\|_\infty$ sind äquivalent, denn für $x \in \mathbb{R}$ gilt:

$$\left\|x\right\|_\infty^2 = \max\{\left|x_1\right|^2, \ldots, \left|x_n\right|^2\} \leq \left|x_1\right|^2 + \cdots + \left|x_n\right|^2 = \left\|x\right\|^2 \;,$$

$$\left\|x\right\|^2 = x_1^2 + \cdots + x_n^2 \leq n \cdot \max\{\left|x_1\right|^2, \ldots, \left|x_n\right|^2\} = n \cdot \max\{\left|x_1\right|, \ldots, \left|x_n\right|\}^2 = n \cdot \left\|x\right\|_\infty^2 \; .$$

Also ist für alle $x \in \mathbb{R}$

$$\left\|x\right\|_\infty \leq \left\|x\right\| \leq \sqrt{n} \left\|x\right\|_\infty$$

Es reicht daher zu zeigen, dass $(\mathbb{R}^n, \|.\|_\infty)$ ein Banachraum ist. Ist $(x^{(k)})_{k \in \mathbb{N}}$ eine Cauchy-Folge im $(\mathbb{R}^n, \|.\|_\infty)$, so ist für jedes $i \in \{1, \ldots, n\}$ $(x_i^{(k)})$ eine Cauchy-Folge in $\mathbb{R}$, denn

$$\left|x_i^{(k)} - x_i^{(l)}\right| \leq \max_{j=1}^{n} \left|x_j^{(k)} - x_j^{(l)}\right| = \left\|x^{(k)} - x^{(l)}\right\|_\infty < \varepsilon$$

wenn $k, l \geq k_0$ ist, für ein $k_0 \in \mathbb{N}$. Also ist $(x_i^{(k)})_{k\in\mathbb{N}}$ eine Cauchy-Folge in $\mathbb{R}$. Da $(\mathbb{R}, |.|)$ vollständig ist, existiert also ein $a_i \in \mathbb{R}$, so dass $(x_i^{(k)})_{k\in\mathbb{N}} \longrightarrow a_i \quad \forall \;\; i = 1, \ldots, n$. Für $a := (a_1, \ldots, a_n) \in \mathbb{R}^n$ gilt dann: Ist $\varepsilon > 0$ und $k_0^i(\varepsilon) > 0$ so, dass für alle $k \geq k_0^i(\varepsilon)$ gilt:

$$\left| x_i^{(k)} - a_i \right| < \varepsilon \ ,$$

so gilt für

$$k_0(\varepsilon) := \max_{i=1}^{n} \{k_0^i(\varepsilon)\} \ \ :$$

$$\left\| x^{(k)} - a \right\|_\infty = \max_{i=1}^{n} \left| x_i^{(k)} - a_i \right| < \varepsilon \quad \forall k \geq k_0(\varepsilon) \ .$$

Also konvergiert $(x^{(k)})$ gegen a.
QED.

Kapitel 2

Stetigkeit

28.10.2004

2.1 Stetigkeit

Definition 2.1.
Seien X, Y metrische Räume und $a \in X$. Eine Abbildung $f : X \to Y$ heißt *stetig in* a, wenn für jede Folge (x_n) in X, die gegen a konvergiert, ihre Bildfolge $\big(f(x_n)\big)$ in Y gegen $f(a)$ konvergiert.

2.1.1 Schreibweise

Wenn für jede Folge (x_n) in X, die gegen a konvergiert, ihre Bildfolge $(f(x_n))$ in Y gegen $b \in Y$ konvergiert, so schreiben wir dafür kurz

$$\lim_{x \to a} f(x) = b\ .$$

Stetigkeit von f in a bedeutet daher in Kurzschreibweise:

$$\lim_{x \to a} f(x) = f(a)\ .$$

2.1.2 Bemerkung

Seien X und Y metrische Räume und $a \in X$. Eine Abbildung $f : X \to Y$ ist genau dann stetig in a, wenn gilt: zu jedem $\varepsilon > 0$ gibt es ein $\delta > 0$, so dass für alle $x \in B_\delta(a)$ gilt, dass $f(x) \in B_\varepsilon\big(f(a)\big)$ ist.

$$f : X \to Y \text{ stetig in } a$$
$$\Updownarrow$$
$$\text{für alle Folgen } (x_n) \text{ mit } (x_n) \longrightarrow a \text{ gilt: } \big(f(x_n)\big) \longrightarrow f(a)$$
$$\Updownarrow$$
$$\forall \varepsilon > 0\ \exists \delta > 0 : f\big(B_\delta(a)\big) \subseteq B_\varepsilon\big(f(a)\big)$$

¡? Beweis ¿!

- $\Downarrow$ Angenommen: Es gibt ein $\varepsilon > 0$ so dass für alle $\delta > 0$ gilt:$f(B_\delta(a)) \not\subseteq B_\varepsilon(f(a))$. Speziell für $\delta = \frac{1}{n}\quad (n \in \mathbb{N})$ ist also dann $B_{\frac{1}{n}}(a) \setminus f^{-1}\big(B_\varepsilon(f(a))\big) \neq \emptyset$; sei $x_n \in B_{\frac{1}{n}}(a) \setminus f^{-1}\big(B_\varepsilon(f(a))\big)$. Aber dann ist $(x_n) \longrightarrow a$, jedoch $f(x_n) \notin B_\varepsilon(f(a))$, also sicher $\big(f(x_n)\big) \not\longrightarrow f(a)$.
- $\Uparrow$ Sei (x_n) in X konvergent gegen a, $(x_n) \longrightarrow a$ und $\varepsilon > 0$ beliebig. Dann existiert ein $\delta > 0$, so dass $f\big(B_\delta(a)\big) \subseteq B_\varepsilon\big(f(a)\big)$. Ist dann $n_0 \in \mathbb{N}$ so groß, dass $x_n \in B_\delta(a)$ für alle $n \geq n_0$, so ist $f(x_n) \in f\big(B_\delta(a)\big) \subseteq B_\varepsilon\big(f(a)\big)\ \ \forall\ \ n \geq n_0$. Also ist $f(x_n) \longrightarrow f(a)$ und damit f stetig in a.

QED.

2.1.3 Bemerkungen

Seien X und Y metrische Räume. Eine Abbildung $f : X \to Y$ ist genau dann stetig, wenn das Urbild jeder offenen Menge (in Y) offen (in X) ist.

¡? Beweis ¿!

- $\Rightarrow$ Sei $V \subseteq Y$ offen und $U := f^{-1}(V) \subseteq X$ sowie $x \in U$. Dann existiert ein $\varepsilon > 0$, so dass $B_\varepsilon\big(f(x)\big) \subseteq V$ (weil $f(x) \in V$ und V offen ist). Da f stetig in x ist, existiert ein $\delta > 0$, so dass $f\big(B_\delta(x)\big) \subseteq B_\varepsilon\big(f(x)\big) \subseteq V$, d.h. $B_\delta(x) \subseteq f^{-1}(V) = U$. Also ist U offen.
- $\Leftarrow$ Sei $x \in X$ und $\varepsilon > 0$ beliebig. Da $B_\varepsilon(f(x)) \subseteq Y$ offen ist, muss also $f^{-1}(B_\varepsilon(f(x)) \subseteq U$ offen sein und $x \in f^{-1}\big(B_\varepsilon(f(x))\big)$. Also existiert $\delta > 0$, so dass $B_\delta(x) \subseteq f^{-1}\big(B_\varepsilon(f(x))\big)$, d.h. $f\big(B_\delta(x)\big) \subseteq B_\varepsilon\big(f(x)\big)$. Daher ist f stetig in x.

QED.

2.1.4 Beispiele

(I) Die Addition

$$\mathrm{add} : \mathbb{R} \times \mathbb{R} \to \mathbb{R}\ , \quad (x, y) \mapsto x + y\ ,$$

die Multiplikation

$$\mathrm{mult} : \mathbb{R} \times \mathbb{R} \to \mathbb{R}\ , \quad (x, y) \mapsto x \cdot y$$

und die Division

$$\mathrm{di} : \mathbb{R} \times \mathbb{R}\backslash\{0\} \to \mathbb{R}\ , \quad (x, y) \mapsto \frac{x}{y}$$

sind stetig.
✎ *Übung* ✐

(II) Die Hintereinanderschaltung $g \circ f : X \to Z$ von zwei stetigen Abbildungen $f : X \to Y$ und $g : Y \to Z$ ist wieder stetig ($W \subseteq Z$ offen $\Longrightarrow (g \circ f)^{-1}(W) = f^{-1}(g^{-1}(W))$ ist offen).

(III) Eine Abbildung $f : X \to \mathbb{R}^n\ ,\quad f = (f_1, \ldots, f_n)\ ,\quad f_i : X \to \mathbb{R} \quad (i = 1, \ldots, n)$ ist genau dann stetig, wenn $f_i : X \to \mathbb{R}$ stetig ist, für alle $i = 1, \ldots, n$.
✎ *Übung* ✐

(IV) Sind $f, g : X \to \mathbb{R}$ stetige Funktionen, so sind auch die Funktionen $f + g : X \to \mathbb{R}\ ,\quad f \cdot g : X \to \mathbb{R}$ (und auch $\frac{f}{g} : X \to \mathbb{R}$, falls $g(x) \neq 0\ \ \forall x \in X$) stetig, denn z.B. ist: $f+g = \mathrm{add}\circ(f, g)$.

2.2 Gleichmäßige Konvergenz

Definition 2.2.
Seien X und Y metrische Räume. Eine Funktionen-Folge (f_n) von Funktionen $f_n : X \to Y$ $(n \in \mathbb{N})$ heißt *gleichmäßig konvergent gegen eine Funktion*

$$f : X \to Y\ , \quad (f_n) \xrightarrow{\text{glm.}} f\ ,$$

wenn gilt:
Zu jedem $\varepsilon > 0$ gibt es ein $n_0 \in \mathbb{N}$, so dass für alle $n \geq n_0$ und für alle $x \in X$ gilt:

$$d(f_n(x), f(x)) < \varepsilon\ .$$

02.11.2004

punktweise Konvergenz: $(f_n) \longrightarrow f$ punktweise

$$\forall x \in X \; \forall \varepsilon > 0 \; \exists n_0(x, \varepsilon)) \in \mathbb{N} \; \forall n \geq n_0 : d(f_n(x), f(x)) < \varepsilon$$

gleichmäßige Konvergenz:

$$\forall \varepsilon > 0 \; \exists n_0(\varepsilon)) \in \mathbb{N} \; \forall n \geq n_0 \forall x \in X : d\big(f_n(x), f(x)\big) < \varepsilon$$

Satz 2.3.
Ist (f_n) gleichmäßig konvergente Folge von Abbildungen $f_n : X \to Y$ gegen eine Abbildung $f : X \to Y$, $(f_n) \stackrel{\text{glm.}}{\longrightarrow} f$, so gilt: Sind alle f_n $(n \in \mathbb{N})$ stetig, so ist auch f stetig.

¡? Beweis ¿!

Sei $a \in X$ und $(x_k)_{k \in \mathbb{N}}$ eine Folge in X mit $(x_k) \longrightarrow a$. Sei $\varepsilon > 0$. Dann existiert zunächst ein $n_1 \in \mathbb{N}$, so dass für alle $x \in X$ gilt:

$$d\big(f_{n_1}(x), f(x)\big) < \frac{\varepsilon}{3} \; .$$

Sei nun $k_1 \in \mathbb{N}$ so, dass für alle $k \geq k_1$ gilt (Stetigkeit von f_{n_1} in a):

$$d\big(f_{n_1}(x_k), f_{n_1}(a)\big) < \frac{\varepsilon}{3}$$

Dann gilt für alle $k \geq k_1$:

$$\begin{aligned} d(f(x_k), f(a)) &\leq d\big(f(x_k), f_{n_1}(x_k)\big) + d\big(f_{n_1}(x_k), f_{n_1}(a)\big) + d\big(f_{n_1}(a), f(a)\big) \\ &< \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} \\ &= \varepsilon \end{aligned}$$

Also konvergiert $\big(f(x_k)\big)$ gegen $f(a)$, d.h. f ist stetig in a.
QED.

2.2.1 Beispiele

(I) Die Folge $f_n : [0,1] \to [0,1]$ gegeben durch $f_n(x) = x^n$ besteht aus stetigen Funktionen und ihre Grenzfolge $f : [0,1] \to [0,1]$ ist

$$f(x) = \begin{cases} 0 & \text{für } 0 \leq x < 1 \\ 1 & \text{für } x = 1 \end{cases}$$

was nicht stetig ist. Die Konvergenz von f_n gegen f ist auch nur *punktweise* aber nicht gleichmäßig.

(II) Sei $[a,b] \in \mathbb{R}$ ein beschränktes, abgeschlossenes Intervall und $V = \mathcal{C}([a,b])$ der $\mathbb{R}$-Vektorraum der stetigen Funktionen auf $[a,b]$,

$$V = \{f : [a,b] \to \mathbb{R} | \; f \text{ stetig}\} \; .$$

Dann ist V zusammen mit der Supremumsnorm

$$\|f\|_\infty := \sup_{x \in [a,b]} |f(x)|$$

ein Banachraum (also $(V, \|.\|)$ vollständig), denn $\|.\|_\infty$ ist eine Norm auf V und ist (f_n) eine Cauchy-Folge in V, so ist wegen

$$|f_n(x) - f_m(x)| \leq \|f_n - f_m\|_\infty \; , \; \forall x \in X \; ,$$

$(f_n(x))$ auch eine Cauchy-Folge (in $\mathbb{R}$). Weil $\mathbb{R}$ vollständig ist, existiert ein $f(x) \in \mathbb{R}$ so dass

$$f(x) = \lim_{n \to \infty} \big(f_n(x)\big) \; .$$

Dann konvergiert (f_n) sogar gleichmäßig gegen f ✎ *Übung* ✐, also ist f (mit Satz 2.3.) sogar stetig, also in V. Daher ist $(f_n) \longrightarrow f$ sogar in $(V, \|.\|_\infty)$, also $(V, \|.\|_\infty)$ vollständig.

Satz 2.4. Banach'scher Fixpunktsatz
Sei X ein vollständiger metrischer Raum und $A \subseteq X$ eine abgeschlossene Teilmenge. Es sei $f : A \to A$ eine Abbildung, so dass es eine Konstante $0 < \Theta < 1$ gibt, so dass für alle $x, y \in A$ gilt

$$d\big(f(x), f(y)\big) \leq \Theta d(x, y)$$

(eine *Kontraktion*). Dann gibt es genau einen *Fixpunkt von* f, d.h. ein $a \in A$ mit $f(a) = a$.

¡? Beweis ¿!

- Eindeutigkeit
  Es kann höchstens einen Fixpunkt von f geben, denn ist $f(a) = a$ und $f(b) = b$, so ist

  $$d(a, b) = d\big(f(a), f(b)\big) \leq \Theta d(a, b) \quad \text{mit } 0 < \Theta < 1$$

  also muss $d(a, b) = 0$ und damit $a = b$ sein.

- Existenz
  Sei $x_0 \in A$. Betrachte die Iteriertenfolge $x_{n+1} := f(x_n)$. Es ist dann zunächst (x_n) beschränkt in X, weil für alle $n \in \mathbb{N}$

  $$\begin{aligned} d(x_n, x_0) &\leq d(x_n, x_{n-1}) + d(x_{n-1}, x_{n-2}) + \cdots + d(x_1, x_0) \\ &\leq \Theta^{n-1} d(x_1, x_0) + \Theta^{n-2} d(x_1, x_0) + \cdots + \Theta^0 d(x_1, x_0) \\ &= \big(1 + \Theta + \cdots + \Theta^{n-1}\big)\, d(x_1, x_0) \\ &= \frac{1 - \Theta^n}{1 - \Theta} d(x_1, x_0) \\ &\leq \frac{1}{1 - \Theta} d(x_1, x_0) =: M\ . \end{aligned}$$

  Damit ist weiter

  $$d(x_n, x_m) \leq d(x_n, x_0) + d(x_0, x_m) \leq 2M \quad \forall n, m \in \mathbb{N}\ .$$

  Es ist (x_n) sogar Cauchy-Folge, denn ist $\varepsilon > 0$ beliebig, $n_0 \in \mathbb{N}$ und $n, m \geq n_0$, so gilt

  $$d(x_n, x_m) \leq \Theta^{n_0} d(x_{n-n_0}, x_{m-n_0}) \leq \Theta^{n_0} \cdot 2M < \varepsilon$$

  wenn $n_0 = n_0(\varepsilon)$ groß genug ist (weil $(\Theta^n) \longrightarrow 0$ eine Nullfolge ist). Da X vollständig ist, konvergiert (x_n) gegen ein $a \in X$, $(x_n) \longrightarrow a$. Da A abgeschlossen ist und $x_n \in A$, ist auch $a = \lim(x_n) \in A$ (1.4.2). Schließlich ist f stetig (f Kontraktion $\Longrightarrow f$ stetig ✎ *Übung* ✐) und daher ist dann:

  $$f(a) = f\big(\lim(x_n)\big) = \lim_{n\to\infty} f(x_n) = \lim_{n\to\infty} (x_{n+1}) = a\ .$$

QED.

Kapitel 3

Kompaktheit

02.11.2004

3.1 Kompakte Mengen

Definition 3.1.
Sei X ein metrischer Raum und $Y \subseteq X$ eine Teilmenge.

(a) Eine Familie $(U_i)_{i \in I}$ von Teilmengen $U_i \subseteq X$ heißt eine *offene Überdeckung von* Y, wenn jedes $U_i \subseteq X$ offen ist und $Y \subseteq \bigcup_{i \in I} U_i$ ist.

(b) Eine Teilmenge $K \subseteq X$ heißt *kompakt*, wenn gilt: Zu jeder offenen Überdeckung $(U_i)_{i \in I}$ von K gibt es eine endliche Teilüberdeckung, d.h. es existieren endlich viele $i_1, i_2, \ldots, i_n \in I$, so dass bereits

$$K \subseteq U_{i_1} \cup U_{i_2} \cup \cdots \cup U_{i_n} \ .$$

3.1.1 Beispiele

(I) Jede endliche Teilmenge $K = \{x_1, \ldots, x_n\}$ eines metrischen Raumes ist kompakt.

(II) Ist (x_n) eine konvergente Folge in X und $a = \lim x_n$, so ist die Teilmenge

$$K = \{x_n | n \in \mathbb{N}\} \cup \{a\}$$

kompakt, denn: ist $(U_i)_{i \in I}$ eine offene Überdeckung von K, so wähle zunächst ein $i_n \in I$ mit $x_n \in U_{i_n}$ und $i_0 \in I$ mit $a \in U_{i_0}$. Da $(x_n) \longrightarrow a$, existiert ein $n_0 \in \mathbb{N}$, so dass $x_n \in U_{i_0}$, für alle $n \geq n_0$. Dann ist K aber bereits in

$$U_{i_1} \cup \cdots \cup U_{i_{n-1}} \cup U_{i_0}$$

enthalten. Also ist K kompakt.

Definition 3.2.
Sei X ein metrischer Raum und $(x_n)_{n \in \mathbb{N}}$ eine Folge in X. Es heißt $a \in X$ ein *Häufungspunkt von* (x_n), wenn in jeder Umgebung von a unendlich viele Folgenglieder liegen.

04.11.2004

3.1.2 Bemerkungen

(1) Eine konvergente Folge hat genau einen Häufungspunkt, nämlich ihren Grenzwert $a = \lim(x_n)$. ✎ *Übung* ✎

(2) Es ist $a \in X$ genau dann ein Häufungspunkt von (x_n), wenn es eine Teilfolge $(x_{n_k})_{k\in\mathbb{N}}$ gibt, die gegen a konvergiert. ✎ *Übung* ✐

Satz 3.3. Bolzano-Weierstraß
Ist X ein metrischer Raum und $K \subseteq X$ kompakt, so besitzt jede Folge in K einen Häufungpunkt in K.

¡? Beweis ¿!

Sei (x_n) eine Folge in K. Angenommen: (x_n) hat keinen Häufungspunkt. Dann gibt es für alle $a \in K$ eine offene Umgebung $U_a \subseteq X$, so dass U_a nur endlich viele Folgenglieder enthält. Aber $(U_a)_{a\in K}$ ist eine offene Überdeckung von K,

$$K = \bigcup_{a\in K} U_a \ .$$

Weil K kompakt ist, existieren $a_1, \ldots, a_n \in K$ so dass bereits

$$K \subseteq U_{a_1} \cup \cdots \cup U_{a_n} \ .$$

Aber dann liegen in K nur endlich viele Folgenglieder.
ϟ Widerspruch ϟ
QED.

3.1.3 Kommentar

Man kann sogar zeigen, dass eine Teilmenge $K \subseteq X$ eines metrischen Raumes genau dann kompakt ist, wenn jede Folge in K einen Häufungspunkt in K hat (vgl. Heuser: `Lehrbuch der Analysis, Teil 2, $111`).

3.2 Beschränkte Mengen

Definition 3.4.
Sei X ein metrischer Raum.

(a) Eine Teilmenge $B \subseteq X$ heißt *beschränkt*, wenn es ein $c > 0$ gibt, so dass für alle $x, y \in B$ gilt:

$$d(x, y) \leq c \ .$$

(b) Für eine Teilmenge $Y \subseteq X$ definiert man den *Durchmesser von* Y durch

$$\operatorname{diam}(Y) := \sup\{d(x, y) | x, y \in Y\} \in [0, \infty] \ .$$

3.2.1 Kommentar

Es ist offenbar $B \subseteq X$ genau dann beschränkt, wenn ihr Durchmesser endlich ist: $\operatorname{diam}(B) < \infty$.

3.2.2 Bemerkungen

Ist $K \subseteq X$ kompakte Teilmenge eines metrischen Raumes X, so ist K abgeschlossen und beschränkt.

¡? Beweis ¿!

Sei $p \in X$ fest. Da für $U_n = B_n(p)$ sogar gilt, dass $(U_n)_{n\in\mathbb{N}}$ ganz X überdeckt, ist (U_n) insbesondere offene Überdeckung von K, also gibt es ein $n_0 \in \mathbb{N}$, so dass

$$K \subseteq U_1 \cup \cdots \cup U_{n_0} = U_{n_0}$$

ist. Für alle $x, y \in K$ ist dann

$$d(x,y) \leq \underbrace{d(x,p)}_{<n_0} + \underbrace{d(y,p)}_{<n_0} < n_0 + n_0 = 2n_0 =: c \ .$$

Also ist K beschränkt.
K ist auch abgeschlossen, denn für eine Folge (x_n) in K mit einem Grenzwert $a \in X \quad , (x_n) \longrightarrow a$ gilt, dass $a \in K$ liegen muss, denn (x_n) hat nur einen Häufungspunkt, nämlich a, und dieser muss nach Satz 3.3. in K liegen. Nach 1.4.2 ist daher K abgeschlossen.
QED.

3.2.3 Bemerkung

Sei X ein metrischer Raum. Ist $K \subseteq X$ eine abgeschlossene Teilmenge und ist $K \subseteq L$ für eine kompakte Teilmenge $L \subseteq X$, so ist K bereits kompakt.
¡? Beweis ¿!

Sei $(U_i)_{i\in I}$ eine offene Überdeckung von K. Setzt man $U = X \backslash K$, also offen, so ist $\{U, U_i | i \in I\}$ Überdeckung von L, also existiert eine endliche Teilmenge der Art, dass[1]

$$L \subseteq U \cup U_{i_1} \cup \cdots \cup U_{i_n}$$

ist. Es ist dann

$$K \subseteq U_{i_1} \cup \cdots \cup U_{i_n} \ ,$$

also ist auch K kompakt.
QED.

Satz 3.5. Heine-Borel
Eine Teilmenge K im euklidischen Raum $\mathbb{R}^n$ ist genau dann kompakt, wenn sie beschränkt und abgeschlossen ist.

¡? Beweis ¿!

$\Rightarrow$ gilt sogar in beliebigen metrischen Räumen (siehe Bemerkung 3.2.2).

$\Leftarrow$ Sei $K \subseteq \mathbb{R}^n$ beschränkt und abgeschlossen. Dann existiert ein $R > 0$, so dass $K \subseteq W_R(0) := [-R, R]^n \subseteq \mathbb{R}^n$. Wir werden in Satz 3.7. zeigen, dass W_R kompakt ist. Damit ist nach 3.2.3 auch K kompakt.

QED.

3.3 Schachtelungsprinzip

Proposition 3.6. Schachtelungsprinzip
Ist X ein vollständiger metrischer Raum und $(A_n)_{n\in\mathbb{N}}$ eine Familie nicht leerer, abgeschlossener Teilmengen von X mit $A_n \supseteq A_{n+1}$ und mit $\big(\operatorname{diam}(A_n)\big)_{n\in\mathbb{N}} \xrightarrow[\infty]{n} 0$. Dann gibt es genau ein $a \in X$, welches in allen A_n liegt.

[1] oder genauer: Sei o.E. $0 \notin I$, setze $U_0 := U$. Dann ist $(U_i)_{i\in I\cup\{0\}}$ offene Überdeckung von L, also existiert $i_1, \ldots, i_n \in I$ mit $L \subseteq \ldots$

¡? Beweis ¿!

- Eindeutigkeit
  Seien $a, b \in \bigcap_{n=1}^{\infty} A_n$. Angenommen: $d := d(a,b) > 0$ Wähle dann $n \in \mathbb{N}$ so groß, dass $\operatorname{diam}(A_n) < d$ ist (immer möglich, weil $\operatorname{diam}(A_n) \longrightarrow 0$). Dann kann a und b nicht in A_n liegen.
  ↯ Widerspruch ↯
  Also ist $d(a,b) = 0$ und damit $a = b$.

- Existenz
  Da $A_n \neq \emptyset$, existiert zunächst $x_n \in A$, $\forall n \in \mathbb{N}$. Wegen $A_n \supseteq A_{n+1}$ gilt dann für $n \geq n_0$, dass $x_n \in A_{n_0}$ ist. Sei nun $\varepsilon > 0$ beliebig und $n_0 \in \mathbb{N}$ so groß, dass $\operatorname{diam}(A_n) < \varepsilon$ ist, $\forall n \geq n_0$. Dann gilt für alle $n, m \geq n_0$
  $$d(x_n, x_m) < \varepsilon \ .$$
  Also ist (x_n) eine Cauchy-Folge. Da X vollständig ist, existiert damit ein $a \in X$ mit$(x_n) \longrightarrow a$. Damit ist auch $(x_k)_{k=n}^{\infty} \longrightarrow a$, aber $(x_k) \in A_n$, $\forall k \geq n$. Da A_n abgeschlossen ist, ist damit auch $a \in A_n$. Also ist $a \in \bigcap_{n=1}^{\infty} A_n$.

QED.

Satz 3.7.
Sei $R > 0$ und $W = \{x \in \mathbb{R}^n | \, \|x\|_\infty \leq R\}$ (Würfel mit Kantenlänge $2R$ um den Nullpunkt). Dann ist W kompakt.

¡? Beweis ¿!

Sei $(U_i)_{i \in I}$ eine offene Überdeckung von W. Angenommen: Es gibt keine endliche Teilüberdeckung. Zerlege nun W in die 2^n Teilwürfel der Kantenlänge R,

$$Q^{(S_1, \ldots, S_n)} := I_{S_1} \times I_{S_2} \times \cdots \times I_{S_n}$$

mit $S_i \in \{0, 1\}$ und $I_0 = [-R, 0]$ und $I_1 = [0, R]$. Für mindestens einen dieser 2^n Teilwürfel gilt, dass er (auch) nicht durch endlich viele Mitglieder von $(U_i)_{i \in I}$ überdeckt wird, diesen nennen wir
09.11.2004 W_2 (und $W =: W_1$). Zerlege nun W_2 wieder in 2^n Teilwürfel (mit Kantenlänge R) und suche aus diesen wieder einen heraus, der nicht von endlich vielen U_i überdeckt wird. Erhalte so $(W_k)_{k \in \mathbb{N}}$, wobei $W_{k+1} \subseteq W_k$, W_k ist abgeschlossen und $\big(\operatorname{diam}(W_k)\big) \longrightarrow 0$. Da $\mathbb{R}^n$ vollständig, gilt wegen des Schachtelungsprinzipes Proposition 3.6., dass es einen $a \in W$ gibt mit $a \in \bigcap_{k=1}^{\infty} W_k$. Da aber $W \subseteq \bigcup_{i \in I} U_i$, gibt es ein $i_0 \in I$, so dass $a \in U_{i_0}$. Da U_{i_0} offen, existiert ein $\varepsilon > 0$, so dass $B_\varepsilon(a) \subseteq U_{i_0}$ ist. Wähle nun ein $k_0 \in \mathbb{N}$ so groß, dass $\operatorname{diam}(W_{k_0}) < \varepsilon$ ist. Damit ist dann (wegen $a \in W_{k_0}$)

$$W_{k_0} \subseteq B_\varepsilon(a) \subseteq U_{i_0} \ .$$

Aber dann wird W_{k_0} bereits von einem einzigen Mitglied der Familie $(U_i)_{i \in I}$ überdeckt, wo es doch nicht mal durch endlich viele Mitglieder überdeckbar sein soll.
↯ Widerspruch ↯
(Also: Heine-Borel: $K \subseteq \mathbb{R}^n$ kompakt $\Leftrightarrow$ K abgeschlossen und beschränkt)

QED.

Satz 3.8.
Seien X und Y metrische Räume, $f : X \to Y$ stetig und $K \subseteq X$ sei kompakt. Dann ist auch $f(K)$ kompakt.

¡? Beweis ¿!

Sei $(V_i)_{i \in I}$ eine offene Überdeckung von $f(K) \subseteq Y$. Setzt man $U_i := f^{-1}(V_i) \subseteq X$, so ist U_i offen, weil f stetig ist und auch eine Überdeckung von K, $K \subseteq \bigcup_{i \in I} U_i$. Also existieren $i_1, \ldots, i_n \in I$, so dass $K \subseteq U_{i_1} \cup \cdots \cup U_{i_n}$. Damit ist

$$f(K) \subseteq f(U_{i_1}) \cup \cdots \cup f(U_{i_n}) \subseteq V_{i_1} \cup \cdots \cup V_{i_n} \ .$$

Also ist $f(K)$ kompakt.
QED.

Satz 3.9. Weierstraß
Sei $K \subseteq \mathbb{R}^n$ kompakt und $f : K \to \mathbb{R}$ eine stetige Funktion. Dann nimmt f ihr Supremum und ihr Infimum an.

3.3.1 Kommentar

Ist $Y \subseteq \mathbb{R}^n$ beliebig und $f : Y \to \mathbb{R}$ eine Funktion, so ist i.a.

$$\sup_{y\in Y} \big(f(y)\big) \in \mathbb{R} \cup \{\infty\} \ , \quad \inf_{y\in Y} \big(f(y)\big) \in \mathbb{R} \cup \{-\infty\} \ .$$

Dass f sein Supremum (bzw. Infimum) annimmt, bedeutet, dass es ein $y_0 \in Y$ gibt, so dass

$$\sup_{y\in Y} \big(f(y)\big) = f(y_0) \quad \text{bzw.} \quad \inf_{y\in Y} \big(f(y)\big) = f(y_0)$$

ist, insbesondere ist dann f nach oben bzw. unten beschränkt, denn $f(y_0) \in \mathbb{R}$

¡? Beweis ¿! (nur Supremum)

Sei

$$c := \sup_{x\in K} f(x) \in \mathbb{R} \cup \{\infty\} \ .$$

Nach Definition des Supremums gibt es dann eine Folge (x_n) in K mit folgender Eigenschaft:

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

Da K kompakt ist, hat (x_n) einen Häufungspunkt $a \in K$. Also gibt es eine Teilfolge $(x_{n_k})_{k\in\mathbb{N}}$ mit $(x_{n_k}) \longrightarrow a$. Dann ist wegen der Stetigkeit von f aber

$$f(a) = f\left(\lim_{k\to\infty} (x_{n_k})\right) = \lim_{k\to\infty} f(x_{n_k}) = c \ .$$

Also nimmt f ihr Supremum in K an.
QED.

3.4 Wegzusammenhang

Definition 3.10.
Sei X ein metrischer Raum. Eine Teilmenge $Y \subseteq X$ heißt *wegzusammenhängend*, wenn es zu jedem $y_0, y_1 \in Y$ einen *Weg von y_0 nach y_1 in Y* gibt, d.h. eine stetige Abbildung $\alpha : [0,1] \to Y$ mit $\alpha(0) = y_0$ und $\alpha(1) = y_1$.

Satz 3.11.
Seien X und Y metrische Räume und $f : X \to Y$ stetig, dann gilt: Ist $A \subseteq X$ wegzusammenhängend, so ist auch $f(A) \subseteq Y$ wegzusammenhängend.

¡? Beweis ¿!

Seien $y_0, y_1 \in f(A)$ beliebig. Dann existieren $x_0, x_1 \in A$, so dass $f(x_0) = y_0$ und $f(x_1) = y_1$. Weil A wegzusammenhängend ist, existiert ein Weg $\alpha : [0,1] \to A$ von x_0 nach x_1. Dann ist $\beta := f \circ \alpha : [0,1] \to f(A)$ ein Weg von y_0 nach y_1.
QED.

Definition 3.12.
Sei X ein metrischer Raum. Eine Teilmenge $G \subseteq X$ heißt ein *Gebiet in* X, wenn G offen und wegzusammenhängend ist.

3.4.1 Kommentar

(1) Das Definitionsgebiet unserer Funktionen mehrerer Veränderlichen $f(x_1, \ldots, x_n)$ wird üblicherweise ein Gebiet $G \subseteq \mathbb{R}^n$ sein, $f : G \to \mathbb{R}$.

(2) Für den Fall $n = 1$ ist $G \subseteq \mathbb{R}$ genau dann ein Gebiet, wenn G ein offenes Intervall $(a, b) \subseteq \mathbb{R}$ ist (mit $a \in [-\infty, \infty)$ und $b \in (-\infty, \infty]$). Setzt man nämlich $a := \inf(G) \in [-\infty, \infty)$ und $b := \sup(G) \in (-\infty, \infty]$, so folgt zunächst, dass $G \subseteq [a, b]$, dann aber sogar in (a, b), weil G offen ist. Es ist aber sogar $G = (a, b)$, denn ist $\xi \in (a, b)$ beliebig, so wähle zunächst ein $x_0 \in (a, \xi) \cap G$ und ein $x_1 \in (\xi, b) \cap G$. Es gibt nun einen Weg α in G von x_0 nach x_1, da G wegzusammenhängend. Der Zwischenwertsatz liefert dann ein $t \in (0, 1)$, so dass $\xi = \alpha(t) \in G$ ist. Also ist $G = (a, b)$.

(Dass (a, b) ein gebiet ist, ist klar.)

Kapitel 4

Differenzierbarkeit

09.11.2004

4.1 Partielle Ableitungen

Definition 4.1.
Sei $n \in \mathbb{N}$ und $G \subseteq \mathbb{R}^n$ ein Gebiet. Für $x \in G$ und $j \in \{1, 2, \ldots, n\}$ sagen wir, dass eine Funktion $f : G \to \mathbb{R}$ *im Punkt x in die j-te Koordinatenrichtung partiell differenzierbar* ist, falls der Grenzwert

$$\lim_{h \to 0} \frac{1}{h} \left(f(x + h\hat{e}_j) - f(x) \right)$$

existiert.

4.1.1 Kommentar

(1) Hierbei ist $\hat{e}_j := (\delta_{jl})_{l \in \{1,\ldots,n\}}$ der j-te kanonische Basisvektor von $\mathbb{R}^n$.

(2) Dass der Limes von $\frac{1}{h} \left(f(x + h\hat{e}_j) - f(x) \right)$ existiert, bedeutet präzise, dass für alle Nullfolgen $(h_k)_{k \in \mathbb{N}}$ in $\mathbb{R}$ mit $h_k \neq 0, \forall k \in \mathbb{N}$ und h_k derart, dass $x + h_k \hat{e}_j \in G$ ist, $\forall k \in \mathbb{N}$, der Grenzwert

$$\lim_{k \to \infty} \frac{1}{h_k} \left(f(x + h_k \hat{e}_j) - f(x) \right)$$

existiert und gleich ist.

(3) Wir schreiben dann

$$D_j f(x) = \lim_{h \to 0} \frac{1}{h} \left(f(x + h\hat{e}_j) - f(x) \right)$$

oder auch

$$\frac{\partial f}{\partial x_j}(x)$$

und nennen diese Zahl die *j-te partielle Ableitung von f in x.* 11.11.2004

(4) Man beachte, dass man in dieser Definition die Differenzierbarkeit von f auf die Situation der Differenzierbarkeit einer Funktion in einer Veränderlichen zurückgeführt hat. Betrachtet man nämlich $x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n$ als konstant (sozusagen als Parameter) und betrachtet die Funktion

$$f_j : (x_j - \varepsilon, x_j + \varepsilon) \to \mathbb{R} \ ,$$
$$f_j(t) := f(x_1, \ldots, x_{j-1}, t, x_{j+1}, \ldots, x_n) \ ,$$

so ist f offenbar genau dann in die j-te Richtung in x differenzierbar, wenn f_j in x_j differenzierbar ist (im Sinne der Differenzierbarkeit für Funktionen einer Veränderlichen) und es gilt dann:

$$D_j f(x) = f_j'(x_j)$$

($\varepsilon > 0$ z.B. so klein, dass $B_\varepsilon(x) \subseteq G$ ist).

4.1.2 Beispiel

Betrachte die Radiusfunktion $r : \mathbb{R}^n \backslash \{0\} \to \mathbb{R}_+$ ($\mathbb{R}_+ = \{x \in \mathbb{R} | x > 0\}$),

$$r(x) = \|x\| = \sqrt{\sum_j x_j^2} \ .$$

Dann ist r für alle $x \in \mathbb{R}^n \backslash \{0\}$ und $j \in \{1, \ldots, n\}$ partiell in die j-te Richtung differenzierbar und es gilt:

$$D_j r(x) = \frac{1}{2r(x)} 2x_j = \frac{x_j}{\|x\|} \ .$$

4.2 Gradienten

Definition 4.2.
Sei $G \subseteq \mathbb{R}^n$ ein Gebiet und f eine reelle Funktion auf G, $f : G \to \mathbb{R}$.

(a) Für $x \in G$ heißt f in x *partiell differenzierbar* wenn f in x in alle n Richtungen partiell differenzierbar ist. Man nennt dann

$$\operatorname{grad}(f)(x) := (D_1 f(x), \ldots, D_n f(x)) =: \nabla f(x)$$

den *Gradienten von f in x.*

(b) Es heißt f *partiell differenzierbar*, wenn f in allen Punkten $x \in G$ partiell differenzierbar ist.

(c) Es heißt f *stetig partiell differenzierbar*, kurz $f \in \mathcal{C}^1(G)$, wenn alle partiellen Ableitungen $D_j f : G \to \mathbb{R}$ existieren und stetig sind ($j = 1, \ldots, n$).

4.2.1 Kommentar

(1) Man beachte, dass die partielle Differenzierbarkeit einer Funktion $f : G \to \mathbb{R}$, $G \subseteq \mathbb{R}^n$ ein Gebiet, $n \geq 2$ in einem Punkt $x \in G$ i.a. keineswegs (wie in einer Dimension) die Stetigkeit von f in x nach sich zieht! Betrachte beispielsweise die Funktion $f : \mathbb{R}^2 \to \mathbb{R}$,

$$f(x) = \begin{cases} 0 & \text{für} \quad xy = 0 \\ 1 & \text{für} \quad xy \neq 0 \end{cases} \ .$$

Dann ist f offenbar in $(x, y) = (0, 0)$ partiell differenzierbar (und $\operatorname{grad}(f)(0,0) = (0,0)$) allerdings keineswegs stetig, denn z.B. für die Folge $\left(\frac{1}{k}, \frac{1}{k}\right)_{k \in \mathbb{N}}$ gilt, dass $\left(\frac{1}{k}, \frac{1}{k}\right) \longrightarrow 0$, aber

$$\lim_{k \to \infty} f\left(\frac{1}{k}, \frac{1}{k}\right) = 1 \neq 0 = f(0,0) \ .$$

(2) Wir werden bald einen stärkeren Differenzierbarkeitsbegriff einer Funktion $f : G \to \mathbb{R}^n$, $G \subseteq \mathbb{R}^n$ ein Gebiet, kennenlernen, der insbesondere die Stetigkeit von f nach sich zieht.

4.2.2 Kommentar

Weil der Begriff der partiellen Differenzierbarkeit auf den einer Funktion in einer Veränderlichen zurück geführt werden kann, gelten auch die bekannten Rechenregeln:

(1) Sind $f, g : G \to \mathbb{R}$ partiell differenzierbar, $\lambda \in \mathbb{R}$, so ist auch $f + g : G \to \mathbb{R}$, $(f+g)(x) = f(x) + g(x)$ und auch $\lambda f : G \to \mathbb{R}$, $(\lambda f)(x) = \lambda \cdot f(x)$ partiell differenzierbar und es gilt für $j = 1, \ldots, n$

$$D_j(f + g) = D_j f + D_j g \ ,$$

$$D_j(\lambda f) = \lambda D_j f \ .$$

Insbesondere ist also $\mathcal{C}^1(G) = \{f : G \to \mathbb{R} | f$ stetig partiell differenzierbar$\}$ ein reeller Vektorraum.

(2) Sind $f, g : G \to \mathbb{R}$ partiell differenzierbar, so ist auch ihr Produkt $f \cdot g : G \to \mathbb{R}$, $(f \cdot g)(x) = f(x) \cdot g(x)$ und ihr Quotient $\frac{f}{g} : G \to \mathbb{R}$, $\left(\frac{f}{g}\right)(x) = \frac{f(x)}{g(x)}$ partiell differenzierbar (wenn $g : G \to \mathbb{R}^* = \mathbb{R}\backslash\{0\}$) und es gilt:

$$D_j(fg) = (D_j f)g + f(D_j g) \ ,$$

$$D_j\left(\frac{f}{g}\right) = \frac{1}{g^2}\big((D_j f)g - f(D_j g)\big) \ .$$

(3) Ist $f : G \to \mathbb{R}$ partiell differenzierbar und $h : \mathbb{R} \to \mathbb{R}$ differenzierbar, so ist auch $h \circ f : G \to \mathbb{R}$, $(h \circ f)(x) = h(f(x))$ partiell differenzierbar und es gilt:

$$D_j(h \circ f) = (h' \circ f)D_j f$$

(D.h. $\forall x \in G : D_j(h \circ f)(x) = h'(f(x) \cdot D_j f(x))$.)

4.2.3 Beispiel

Ist $f : \mathbb{R}^n\backslash\{0\} \to \mathbb{R}$ eine *rotationssymmetrische Funktion* , d.h. $f(x) = f(y)$ wenn $\|x\| = \|y\|$ ist, so gibt es eine Funktion $h : \mathbb{R}_+ \to \mathbb{R}$, so dass $f(x) = h(\|x\|)$ für alle $x \in \mathbb{R}^n\backslash\{0\}$ (nämlich z.B. $h(r) = f(r, 0, 0, \dots)$). Es ist also dann mit $r : \mathbb{R}^n\backslash\{0\} \to \mathbb{R}_+$, $r(x) = \|x\|$,

$$f = h \circ r \ .$$

Ist nun f partiell differenzierbar, so ist auch h differenzierbar ($h'(r) = D_1 f(r, 0, 0, \dots)$) und daher gilt mit 4.2.2 (3):

$$D_j f(x) = D_j(h \circ r)(x) = h'(r(x))D_j r(x) = h'(\|x\|)\frac{x_j}{\|x\|} \ ,$$

also

$$\mathrm{grad}\,(f)(x) = h'(\|x\|)\frac{x}{\|x\|}$$

oder kurz

$$\mathrm{grad}\,(h \circ r) = \frac{h' \circ r}{r}\,\mathrm{id} \ .$$

Z.B. ist für $p \in \mathbb{Z}$ und $h(r) = r^p$:

$$\mathrm{grad}\,(\|x\|^p) = p\,\|x\|^{p-1} \cdot \frac{x}{\|x\|} = p\,\|x\|^{p-2} \cdot x$$

oder kurz:

$$\mathrm{grad}\,(r^p) = pr^{p-2}\,\mathrm{id} \ .$$

Definition 4.3.
Seien $m, n \in \mathbb{N}$, $G \subseteq \mathbb{R}^n$ ein Gebiet und $f : G \to \mathbb{R}^m$ eine Abbildung, $f = (f_1, \dots, f_m)$ mit $f_i : G \to \mathbb{R}$. Wir sagen, dass f stetig (partiell) differenzierbar in $x \in G$ ist, falls f_i dies ist für $i = 1, \dots, m$.

4.3 Vektorfelder und Divergenz

Definition 4.4.
Sei $G \subseteq \mathbb{R}^n$ ein Gebiet.

(a) Eine Abbildung $v : G \to \mathbb{R}^n$ heißt dann ein *Vektorfeld auf* G.

(b) Ist $v : G \to \mathbb{R}^n$ ein partiell differenzierbares Vektorfeld, so nennt man die Funktion $\operatorname{div}(v) : G \to \mathbb{R}$, gegeben durch

$$\operatorname{div}(v)(x) = \sum_{j=1}^{n} \frac{\partial v_j}{\partial x_j}(x) = \sum_{j=i}^{n} D_j v_j(x) \ ,$$

die *Divergenz von* v.

16.11.2004

4.3.1 Beispiele

(I) Die Identität $\operatorname{id} : G \to \mathbb{R}^n \ \ x \mapsto x$ ist ein (partiell differenzierbares) Vektorfeld, $\operatorname{id}(x) = (x_1, \dots, x_n)$. Ihre Divergenz ist

$$\operatorname{div}(\operatorname{id}) = \sum_{j=1}^{n} \frac{\partial x_j}{\partial x_j} = 1 + \cdots + 1 = n \ .$$

(II) Die Abbildung $v : \mathbb{R}^n \backslash \{0\} \to \mathbb{R}^n$, $\ x \mapsto \frac{x}{\|x\|}$ ist ein (partiell differenzierbares) Vektorfeld. Die Divergenz ist

$$\begin{aligned} \operatorname{div}(v)(x) &= \sum_{j=1}^{n} \frac{\partial}{\partial x_j}\left(\frac{x_j}{\|x\|}\right) = \sum_{j=1}^{n} \frac{1 \cdot \|x\| - x_j \frac{x_j}{\|x\|}}{\|x\|^2} \\ &= \sum_{j=1}^{n} \frac{1}{\|x\|} - \sum_{j=1}^{n} \frac{x_j \cdot x_j}{\|x\|^3} = \frac{n}{\|x\|} - \frac{\|x\|^2}{\|x\|^3} \\ &= (n-1)\frac{1}{\|x\|} \end{aligned}$$

(oder kurz $\operatorname{div}(\frac{1}{r}\operatorname{id}) = \frac{n-1}{r}$).

(III) Ist $f : G \to \mathbb{R}$ partiell differenzierbar, so ist ihr Gradient $v = \operatorname{grad}(f) : G \to \mathbb{R}^n$ ein Vektorfeld auf G. Ist auch $\operatorname{grad}(f)$ noch einmal partiell differenzierbar, so nennt man $\Delta f : G \to \mathbb{R}$ gegeben durch

$$\Delta f = \operatorname{div}(v) = \operatorname{div} \circ \operatorname{grad}(f) : G \to \mathbb{R} \ ,$$

also

$$\Delta f(x) = \operatorname{div}(\operatorname{grad}(f))(x) = \operatorname{div}(D_1 f, \dots, D_n f)(x) = \sum_{j=1}^{n} D_j(D_j f)(x) = \sum_{j=1}^{n} \frac{\partial^2 f}{\partial x_j^2}(x)$$

den *Laplace von* f.

(IV) Für die Radiusfunktion $r : \mathbb{R}^n \backslash \{0\} \to \mathbb{R}_+$ ist also z.B. wegen $\nabla r = \frac{\operatorname{id}}{r}$ (vgl. 4.2.3)

$$\Delta r = \frac{n-1}{r} \ .$$

Definition 4.5.
Sei G ein Gebiet und $k \in \mathbb{N}$. Eine Funktion $f : G \to \mathbb{R}$ heißt *k-mal stetig partiell differenzierbar*, wenn für alle $(j_1, \ldots, j_k)$ mit $j_k \in \mathbb{N}$ gilt: f ist in die j_1-te Richtung (stetig) partiell differenzierbar, $D_{j_1} f$ ist partiell (stetig) in die j_2-te Richtung differenzierbar, usw. $D_{j_k} \ldots D_{j_1} f$ existiert und ist stetig.

4.3.2 Kommentar

(1) Ist f k-mal stetig partiell differenzierbar, so schreibt man auch (für jedes $j = (j_1, \ldots, j_k) \in \{1, \ldots, n\}^k$)

$$\frac{\partial^k f}{\partial x_{j_k} \ldots \partial x_{j_1}}(x) := D_{j_k} \ldots D_{j_1} f(x) \ .$$

(2) Wiederholte Anwendung von 4.2.2 (1) liefert, dass

$$\mathcal{C}^k(G) := \{f : G \to \mathbb{R} | f \ k\text{-mal stetig partiell differenzierbar}\}$$

ein reeller Vektorraum ist.

(3) Für $k = 0$ setzt man

$$\mathcal{C}^0(G) := \mathcal{C}(G) := \{f : G \to \mathbb{R} | f \text{ stetig}\} \ .$$

(4) Für $k = 2$, $f \in \mathcal{C}^2(G)$ und $x \in G$ nennt man die Matrix

$$\text{Hess}\,(f)(x) = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1} & \cdots & \frac{\partial^2 f}{\partial x_n^2} \end{pmatrix} (x)$$

die *Hesse-Matrix von f in x.*

Satz 4.6. Satz von H.A. Schwarz
Sei $G \subseteq \mathbb{R}^n$ ein Gebiet, $f : G \to \mathbb{R}$ zweimal stetig differenzierbar und $1 \leq i, j \leq n$. Dann gilt für alle $x \in G$:

$$D_i D_j f(x) = D_j D_i f(x) \ .$$

¡? Beweis ¿!

O.B.d.A. sei $i = 1, j = 1, n = 2, x = (0, 0) \in G$.
Wähle $\delta > 0$ so klein, dass $(-\delta, \delta)^2 \subseteq G$ ist (existiert, da G offen). Betrachte nun für jedes feste $y \in (-\delta, \delta)$ die differenzierbare Funktion $F_y : (-\delta, \delta) \to \mathbb{R}$

$$F_y(x) = f(x, y) - f(x, 0) \ .$$

Mit dem Mittelwertsatz der Differenzialrechnung existiert für jedes x ein ξ zwischen 0 und x so, dass

$$\begin{aligned} F_y(x) - F_y(0) &= F_y'(\xi) \cdot x \\ &= (D_1 f(\xi, y) - D_1 f(\xi, 0)) \cdot x \end{aligned}$$

ist. Betrachte nun die folgende differenzierbare Funktion $(-\delta, \delta) \to \mathbb{R}$ $y \mapsto D_1 f(\xi, y)$. Der Mittelwertsatz liefert: Für alle $y \in (-\delta, \delta)$ existiert ein η zwischen 0 und y so dass

$$D_1 f(\xi, y) - D_1 f(\xi, 0) = D_2 D_1 f(\xi, \eta) \cdot y \ .$$

Zusammen erhält man

$$\begin{aligned} f(x,y) - f(x,0) - f(0,y) + f(0,0) &= (f(x,y) - f(x,0)) - (f(0,y) - f(0,0)) \\ &= F_y(x) - F_y(0) = D_2D_1f(\xi,\eta)xy \ . \end{aligned}$$

Betrachtet man andererseits für jedes feste $x \in (-\delta,\delta)$ die Funktion $G_x : (-\delta,\delta) \to \mathbb{R}$,

$$G_x(y) = f(x,y) - f(0,y) \ ,$$

so erhält man für jedes y ein $\widetilde{\eta}$ zwischen 0 und y, so dass gilt:

$$G_x(y) - G_x(0) = (D_2f(x,\widetilde{\eta}) - D_2f(0,\widetilde{\eta}))y \ .$$

Der Mittelwertsatz liefert schließlich auf die Funktion $x \mapsto D_2f(x,\widetilde{\eta})$ angewandt für jedes x ein $\widetilde{\xi}$ zwischen 0 und x, so dass

$$D_2f(x,\widetilde{\eta}) - D_2(0,\widetilde{\eta}) = D_1D_2(\widetilde{\xi},\widetilde{\eta})x$$

ist. Insgesamt liefert dies

$$\begin{aligned} f(x,y) - f(x,0) - f(0,y) + f(0,0) &= (f(x,y) - f(0,y)) - (f(x,0) - f(0,0)) \\ &= G_x(y) - G_x(0) = D_1D_2f(\widetilde{\xi},\widetilde{\eta})xy \end{aligned}$$

(mit $\xi,\widetilde{\xi}$ zwischen 0 und x, $\eta,\widetilde{\eta}$ zwischen 0 und y).
Für $xy \neq 0$ folgt also

$$D_1D_2f(\widetilde{\xi},\widetilde{\eta}) = D_2D_1f(\xi,\eta) \ .$$

Wähle nun eine Folge (x_k,y_k) mit $(x_k,y_k) \longrightarrow (0,0)$ und $x_ky_k \neq 0$. Dann gilt auch $(\widetilde{\xi}_k,\widetilde{\eta}_k) \longrightarrow 0$ und $(\xi_k,\eta_k) \longrightarrow 0$ und wegen der Stetigkeit von D_1D_2f und D_2D_1f in $(0,0)$ ist dann

$$D_1D_2f(0,0) = \lim_{k\to\infty} D_1D_2f\big(\widetilde{\xi}_k,\widetilde{\eta}_k\big) = \lim_{k\to\infty} D_2D_1f(\xi_k,\eta_k) = D_2D_1f(0,0) \ .$$

QED.

4.3.3 Kommentar

Für eine 2-mal stetig differenzierbare Funktion $f : G \to \mathbb{R}$ ist also die Hesse-Matrix in jedem Punkt *symmetrisch.*

4.3.4 Beispiel

(I) Ist f 2-mal stetig differenzierbar auf $G \subseteq \mathbb{R}^n$, $f : G \to \mathbb{R}$, so gilt für $\Delta f : G \to \mathbb{R}$ gerade

$$\Delta f = \sum_{i=1}^{n} D_iD_if = \operatorname{spur}(\operatorname{Hess}(f)) \ .$$

(II) Seien nun $G \subseteq \mathbb{R}^3$ ein Gebiet und $v : G \to \mathbb{R}^3$ ein partiell differenzierbares Vektorfeld, $v = (v_1,v_2,v_3)$. Man definiert die *Rotation von* v als das Vektorfeld $\operatorname{rot}(v) : G \to \mathbb{R}^3$,

$$\operatorname{rot}(v)(x) = (D_2v_3 - D_3v_2, D_3v_1 - D_1v_3, D_1v_2 - D_2v_1)(x) \ .$$

Führt man formal ∇ als „vektorwertigen Differentialoperator“ ein,

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

so schreiben sich grad und div (und für $n = 3$ auch rot) folgendermaßen:

$$\operatorname{grad}(f) = \nabla f \ ,$$

$$\operatorname{div}(f) = \langle \nabla, v \rangle \ ,$$
$$\operatorname{rot}(v) = \nabla \times v \ .$$

Ist $n \in \mathbb{N}$ beliebig, so setzt man für ein partiell differenzierbares Vektorfeld $v : G \to \mathbb{R}^n$ ($G \subseteq \mathbb{R}^n$ ein Gebiet) manchmal

$$\operatorname{curl}(v) : G \to \mathbb{R}^{\binom{n}{2}}$$

($\binom{n}{2} = \frac{(n-1)n}{2} = \#\{\mathit{Paare} \text{ in } \{1, \ldots, n\}\}$)

$$(\operatorname{curl}(v))_{i,j} = D_i v_j - D_j v_i \ ,$$

z.B. für $n = 2$ ist $\operatorname{curl}(v) = D_1 v_2 - D_2 v_1$ und für $n = 3$ ist $\operatorname{curl}(v) = \operatorname{rot}(v)$. Beachte, dass nur für $n = 3$

$$\binom{n}{2} = n$$

gilt.

18.11.2004

Korollar 4.7.

(a) Sei $G \subseteq \mathbb{R}^n$ ein Gebiet und $f : G \to \mathbb{R}$ eine 2-mal stetig partiell differenzierbare Funktion. Dann gilt

$$\operatorname{curl} \circ \operatorname{grad} f = 0 \ .$$

(b) Sei $v : G \to \mathbb{R}^3$ ein 2-mal stetig partiell differenzierbares Vektorfeld, $G \subseteq \mathbb{R}^3$ ein Gebiet. Dann ist

$$\operatorname{div}(\operatorname{rot} v) = 0 \ .$$

¡? Beweis ¿!

(a) $\operatorname{curl}(v)_{i,j} = \frac{\partial v_i}{\partial x_j} - \frac{\partial v_j}{\partial x_i}$ und $v_i = (\operatorname{grad}\ f)_i = \frac{\partial f}{\partial x_i}$ eingesetzt ergibt mit der 2-fachen stetigen partiellen Differenzierbarkeit von f:

$$\operatorname{curl} \circ \operatorname{grad}(f)_{ij} = \frac{\partial^2 f}{\partial x_j \partial x_i} - \frac{\partial^2 f}{\partial x_i \partial x_j} = 0 \ .$$

(b) Für $v = (v_1, v_2, v_3)$ ist

$$\begin{aligned} \operatorname{div}(\operatorname{rot}\ v) &= \operatorname{div}(D_2 v_3 - D_3 v_2, D_3 v_1 - D_1 v_3, D_1 v_2 - D_2 v_1) \\ &= D_1 D_2 v_3 - D_1 D_3 v_2 + D_2 D_3 v_1 - D_2 D_1 v_3 + D_3 D_1 v_2 - D_3 D_2 v_1 \\ &= 0 \ . \end{aligned}$$

QED.

4.3.5 Beispiel

Sei $F : \mathbb{R}^n \backslash \{0\} \to \mathbb{R}$ eine zweimal stetig partiell differenzierbare Funktion, die rotationssymmetrisch ist, also $F(x) = F(y)$ für $\|x\| = \|y\|$. Es gibt dann ein zweimal stetig differenzierbares $f : \mathbb{R}_+ \to \mathbb{R}$ mit $F(x) = f\big(r(x)\big)$, wobei $r(x) = \|x\|$ sei. Behauptung: Auch $G = \Delta F : \mathbb{R}^n \backslash \{0\} \to \mathbb{R}$ ist dann rotationssymmetrisch, also $G = g \circ r$ für ein zweifach stetig differenzierbares $g : \mathbb{R}_+ \to \mathbb{R}$ und es gilt:

$$g(r) = f''(r) + \frac{n-1}{r} f'(r) \ ,$$

also

$$\Delta F(x) = f''(\|x\|) + \frac{n-1}{\|x\|} f'(\|x\|) \ .$$

¡? Beweis ¿!

Es ist $F = f \circ r$. Damit gilt:

$$
\begin{aligned}
\Delta F(x) &= \operatorname{div}\big(\operatorname{grad}(f \circ r)\big)(x) \\
&\overset{4.2.3}{=} \operatorname{div}\left(f' \circ r \cdot \frac{x}{\|x\|}\right) \\
&\overset{\text{Übung}}{=} \left\langle \operatorname{grad}(f' \circ r), \frac{x}{\|x\|} \right\rangle + f' \circ r \cdot \operatorname{div}\left(\frac{x}{\|x\|}\right) \\
&\overset{4.2.3 \text{ bzw. } 4.3.1}{=} \left\langle f'' \circ r \cdot \frac{x}{\|x\|}, \frac{x}{\|x\|} \right\rangle + f' \circ r \cdot (n-1)\frac{1}{\|x\|} \\
&= \left(f'' \circ r + \frac{n-1}{r} f' \circ r\right)(x) \\
&= \left(f'' + \frac{n-1}{r} f'\right) \circ r(x)
\end{aligned}
$$

Mit $g(r) := f''(r) + \frac{n-1}{r} f'(r)$ folgt die Behauptung.
QED.

4.4 Totale Differenzierbarkeit

Motivation

Eine Funktion $f : I \to \mathbb{R}$, $I \subseteq \mathbb{R}$ ein offenes Intervall, ist differenzierbar in $x \in I$, wenn der Grenzwert

$$\lim_{h \longrightarrow 0} \frac{1}{h}\big(f(x+h) - f(x)\big) = f'(x)$$

existiert. Ist $\delta > 0$ klein genug, so dass $(x-\delta, x+\delta) \subseteq I$ ist und setzt man $\varphi : (-\delta, \delta) \to \mathbb{R}$,

$$\varphi(h) := f(x+h) - f(x) - f'(x)h \ ,$$

so erhält man, dass f auf $(x-\delta, x+\delta)$ folgende Darstellung gestattet:

$$f(x+h) = f(x) + f'(x)h + \varphi(h)$$

und die Differenzierbarkeit von f in x drückt sich so aus:

$$\lim_{h \to 0} \frac{\varphi(h)}{h} = \lim_{h \to 0} \left(\frac{f(x+h) - f(x)}{h}\right) = 0 \ .$$

Man sagt dazu, dass φ für $h \longrightarrow 0$ von *höherer als 1. Ordnung* gegen Null geht.
Äquivalent zur Differenzierbarkeit für f in x ist also folgende Approximationseigenschaft durch eine lineare Abbildung in einer Umgebung von x:
Es existiert ein $\delta > 0$, eine Funktion $\varphi : (-\delta, \delta) \to \mathbb{R}$ und eine lineare Abbildung $A : \mathbb{R} \to \mathbb{R}$, so dass gilt:

$$f(x+h) = f(x) + A(h) + \varphi(h)$$

mit $\lim\limits_{h \to 0} \frac{\varphi(h)}{h} = 0$. Eine lineare Abbildung $A : \mathbb{R} \to \mathbb{R}$ ist nämlich einfach durch eine Zahl $a \in \mathbb{R}$ gegeben: $A(h) = ah$; und dieses $a \in \mathbb{R}$ ist dann gerade die Ableitung von f in x.

Definition 4.8.
Sei $G \subseteq \mathbb{R}^n$ ein Gebiet und $f : G \to \mathbb{R}^m$ eine Abbildung. f heißt (total) differenzierbar in $x \in G$, falls eine lineare Abbildung $A : \mathbb{R}^n \to \mathbb{R}^m$ existiert, ein $\delta > 0$ so dass $B_\delta(x) \subseteq G$ ist und es eine Funktion $\varphi : B_\delta(0) \to \mathbb{R}^n$ gibt, so dass für alle $h \in B_\delta(0)$ gilt:

$$f(x+h) = f(x) + Ah + \varphi(h)$$

mit $\lim\limits_{h\to 0} \frac{\varphi(h)}{\|h\|} = 0$.

4.4.1 Kommentar

(1) Ist $\varphi : B_\delta(0) \to \mathbb{R}^n$ eine Funktion und $k \in \mathbb{N}$, so dass

$$\lim_{h\to 0} \frac{\varphi(h)}{\|h\|^k} = 0$$

ist, so sagt man, dass φ von *höherer als k-ter Ordnung in Null verschwindet* und schreibt für ein solches φ (etwas unpräzise) $o\left(\|h\|^k\right)$ („*klein-oh von Norm h hoch k*“). Mit dieser Sprachregelung ist also dann ein $f : G \to \mathbb{R}$ differenzierbar in x, wenn

$$f(x+h) = f(x) + Ah + o(\|h\|)$$

ist.

(2) Eine lineare Abbildung $A : V \to W$ zwischen zwei Vektorräumen V und W der Dimensionen n bzw. m wird bekanntlich bei einer Basiswahl der Vektorräume durch eine $(m \times n)$-Matrix beschrieben, $(a_{ij})_{\substack{1\le i\le m\\ 1\le j\le n}}$. Bei der Wahl der kanonischen Basen $(\hat{e}_1, \dots, \hat{e}_n)$ von $\mathbb{R}^n$ bzw. $(\hat{e}_1, \dots, \hat{e}_m)$ von $\mathbb{R}^m$ gilt für die zugehörige Matrix (a_{ij}) dann:

$$Ah = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} h_1 \\ \vdots \\ h_n \end{pmatrix}.$$

Wir identifizieren bisweilen die lineare Abbildung $A : \mathbb{R}^n \to \mathbb{R}^m$ mit der Matrix (a_{ij}), die sie beschreibt.

(3) Natürlich sagen wir, dass $f : G \to \mathbb{R}^m$ differenzierbar ist, wenn f in allen Punkten $x \in G$ differenzierbar ist. (Beachte: Die lineare Abbildung $A : \mathbb{R}^n \to \mathbb{R}^m$ hängt im Allgemeinen natürlich von x ab!)

23.11.2004

4.4.2 Bemerkung

(1) Sei $G \subseteq \mathbb{R}^n$ ein Gebiet und $f : G \to \mathbb{R}^m$ eine Abbildung, $f = (f_1, \dots, f_m)$. Dann gilt: Ist f in $x \in G$ (total) differenzierbar, so ist $f_i : G \to \mathbb{R}$, $i = 1, \dots, m$ in x partiell differenzierbar.

(2) Ist f in $x \in G$ total differenzierbar mit approximierender linearer Abbildung $A : \mathbb{R}^n \to \mathbb{R}^m$ und Matrix (a_{ij}), so gilt für die partiellen Ableitungen von f_i in x:

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

4.4.3 Kommentar

(1) Diese Bemerkung zeigt, dass die lineare Abbildung $A : \mathbb{R}^n \to \mathbb{R}^m$, deren Existenz in der Definition der totalen Differenzierbarkeit gefordert wird, eindeutig bestimmt ist. Man nennt diese lineare Abbildung *die Ableitung von f in x* und schreibt

$$Df(x) = A$$

(Manchmal auch einfach $f'(x)$, obwohl hier $f'(x)$ eine lineare Abbildung ist!).

(2) Ist (a_{ij}) die Matrix-Beschreibung von $Df(x)$, so nennt man (a_{ij}) die *Jacobi-Matrix von f in x*. Es ist also:

$$J_f(x) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(x) & \cdots & \frac{\partial f_1}{\partial x_n}(x) \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1}(x) & \cdots & \frac{\partial f_m}{\partial x_n}(x) \end{pmatrix} .$$

¡? Beweis ¿!

(1) Ist f in $x \in G$ total differenzierbar, so gibt es also eine lineare Abbildung $A : \mathbb{R}^n \to \mathbb{R}^m$ mit

$$f(x+h) = f(x) + Ah + o(\|h\|) .$$

Betrachtet man die i-te Komponente in dieser Gleichung, so gilt mit der beschreibenden Matrix (a_{ij}) von A

$$f_i(x+h) = f_i(x) + \sum_{k=1}^{n} a_{ik} h_k + o(\|h\|) .$$

(Diesmal steht $o(\|h\|)$ für eine Funktion $\varphi : B_\delta(0) \to \mathbb{R}^1$ mit $\lim\limits_{h\to 0} \frac{\varphi(h)}{\|h\|} = 0$.) Speziell für $h := t\hat{e}_j$ ($|t| < \delta$, $\hat{e}_j = (0, \ldots, 0, 1, 0, \ldots, 0)$) gilt damit:

$$f_i(x + t\hat{e}_j) = f_i(x) + a_{ij} t + o(|t|) .$$

(Beachte $\|h\| = \|t\hat{e}_j\| = |t| \, \|\hat{e}_j\| = |t|$.) Deshalb existiert nun der Grenzwert des Differenzenquotienten

$$\frac{f_i(x + t\hat{e}_j) - f_i(x)}{t} = a_{ij} + \frac{o(|t|)}{t}$$

und damit ist f_i in x in die j-te Richtung partiell differenzierbar und

(2) wegen $\frac{o(|t|)}{t} \underset{0}{\overset{t}{\longrightarrow}} 0$ sieht man auch unmittelbar, dass

$$D_j f_i(x) = \lim_{t\to 0} \frac{f_i(x + t\hat{e}_j) - f_i(x)}{t} = a_{ij}$$

ist.

QED.

4.4.4 Bemerkung

Sei $G \subseteq \mathbb{R}^n$ ein Gebiet, $x \in G$ und $f : G \to \mathbb{R}^m$ total differenzierbar in x. Dann ist f auch stetig in x.

¡? Beweis ¿!

Zu zeigen ist $\lim_{h\to 0} f(x+h) = f(x)$. Wir wissen aber, dass

$$f(x+h) = f(x) + Df(x)h + o(\|h\|)$$

für $h \in B_\delta(0),\ \delta > 0$ klein. Da

$$\lim_{h\to 0} Df(x)h = 0 \quad \text{und} \quad \lim_{h\to 0} o(\|h\|) = \lim_{h\to 0} \underbrace{\frac{o(\|h\|)}{\|h\|}}_{\to 0} \cdot \underbrace{\|h\|}_{\to 0} = 0$$

ist, folgt die Behauptung.
QED.

4.4.5 Kommentar

Wir hatten die Implikationen

- total differenzierbar $\Rightarrow$ partiell differenzierbar
- total differenzierbar $\Rightarrow$ stetig

Für beide Implikationen gilt im Allgemeinen nicht die Umkehrung und zudem gibt es zwischen partiell differenzierbar & stetig keine Implikation. Es gilt aber:

Satz 4.9.
Sei $G \subseteq \mathbb{R}^n$ ein Gebiet und $f : G \to \mathbb{R}^m$ eine Abbildung. Sind alle Komponenten $f_i : G \to \mathbb{R}\ \ (i = 1, \dots, m)$ von f stetig partiell differenzierbar, so ist f total differenzierbar.

¡? Beweis ¿!

Da f genau dann total differenzierbar ist, wenn die Komponenten $f_i\ \ \forall i = 1, \dots, m$ total differenzierbar sind, sei **o.B.d.A.** $m = 1$. Sei $x \in G$ und $\delta > 0$ so klein, dass $B_\delta(x) \subseteq G$ ist. Sei nun $h \in B_\delta(0)$ beliebig. Betrachte nun weiter

$$\begin{array}{lcl}
y^{(0)} & := & x \\
y^{(1)} & := & x + h_1\hat{e}_1 \\
y^{(2)} & := & x + h_1\hat{e}_1 + h_2\hat{e}_2 \\
\vdots & & \\
y^{(j)} & := & x + \sum_{k=1}^{j} h_k\hat{e}_k \\
\vdots & & \\
y^{(n)} & := & x + h\ .
\end{array}$$

Betrachte nun die Funktion $g_j : I_j \subseteq \mathbb{R} \to \mathbb{R}\ ,\ g_j(t) = f(y^{(j-1)} + t\hat{e}_j)\ \ \ j = 1, \dots, n$ mit einem Intervall $I_j \subseteq \mathbb{R}$ wobei $0, h_j \in I_j$. Weil f partiell differenzierbar ist, folgt: g_j ist differenzierbar und daher gibt es nach dem Mittelwertsatz ein θ_j zwischen 0 und h_j, so dass gilt:

$$f\big(y^{(j)}\big) - f\big(y^{(j-1)}\big) = g_j(h_j) - g_j(0) \overset{\text{MWS}}{=} g_j'(\theta_j)h_j = D_j f\big(y^{(j-1)} + \theta_j\hat{e}_j\big)h_j$$

Mit $z^{(j)} := y^{(j-1)} + \theta_j \hat{e}_j$ erhält man damit:

$$\begin{aligned}
f(x+h) - f(x) &= f\big(y^{(n)}\big) - f\big(y^{(0)}\big) \\
&= \Big(f(y^{(n)}) - f(y^{(n-1)})\Big) + \Big(f(y^{(n-1)}) - f(y^{(n-2)})\Big) + \cdots + \Big(f(y^{(1)}) - f(y^{(0)})\Big) \\
&= \sum_{j=1}^{n} f\big(y^{(j)}\big) - f\big(y^{(j-1)}\big) \\
&= \sum_{j=1}^{n} D_j f\big(z^{(j)}\big) h_j \ .
\end{aligned}$$

Setzt man nun

$$a_j := D_j f(x) \ ,$$

so gilt also

$$f(x+h) - f(x) = \sum_{j=1}^{n} a_j h_j + \varphi(h)$$

mit

$$\varphi(h) = \sum_{j=1}^{n} \Big(D_j f(z^{(j)}) - a_j\Big) h_j$$

und daher

$$\frac{|\varphi(h)|}{\|h\|} \leq \sum_{j=1}^{n} \underbrace{\Big|D_j f(z^{(j)}) - D_j f(x)\Big|}_{\xrightarrow{h \to 0}\ 0} \cdot \underbrace{\frac{|h_j|}{\|h\|}}_{\leq 1} \xrightarrow{h \to 0} 0 \ ,$$

weil $D_j f$ in x stetig ist und $z^{(j)} \xrightarrow{h \to 0} x$. Also ist f in x total differenzierbar.

QED.

4.4.6 Kommentar

(1) Betrachtet man Satz 4.9. zusammen mit 4.4.4, so sieht man, dass eine stetig partiell differenzierbare Funktion insbesondere stetig ist.

Eine Funktion $f : G \to \mathbb{R}$ ($G \subseteq \mathbb{R}^n$ ein Gebiet) ist also genau dann in $\mathcal{C}^k(G)$, wenn alle partiellen Ableitungen $D_{j_k} \cdots D_{j_1} f$ $(j_1, \ldots, j_k \in \{1, \ldots, n\})$ existieren und stetig sind. Die niedrigeren Ableitungen existieren dann sowieso und sind automatisch stetig.

(2) Da eine stetig partiell differenzierbare Funktion (total) differenzierbar ist nennen wir sie nun nur noch kurz *stetig differenzierbar*.

25.11.2004

$$\begin{array}{ccccc}
\text{stetig partiell differenzierbar} & \Rightarrow & \text{total differenzierbar} & \Rightarrow & \text{partiell differenzierbar} \\
 & & \Downarrow & & \\
 & & \text{stetig} & &
\end{array}$$

4.5 Rechenregeln

Motivation

Bisher haben wir die Kettenregel für Funktionen in einer Veränderlichen nur für folgende Situationen verallgemeinern können: Ist $G \subseteq \mathbb{R}^n$ ein Gebiet, $f : G \to \mathbb{R}$ partiell differenzierbar und $h : \mathbb{R} \to \mathbb{R}$ differenzierbar, so ist auch $h \circ f : G \to \mathbb{R}$ partiell differenzierbar und

$D_j(h \circ f) = h' \circ f \cdot D_j f$ (4.2.2 (3)). Noch nicht betrachtet haben wir folgende (wichtige) allgemeinere Version: Falls $G \subseteq \mathbb{R}^n$ und $D \subseteq \mathbb{R}^m$ Gebiete sind, $f : G \to \mathbb{R}^m$ differenzierbar mit $f(G) \subseteq D$ und $g : D \to \mathbb{R}^r$ differenzierbar, ist dann auch $g \circ f : G \to \mathbb{R}^r$ differenzierbar und wie hängt dann die Ableitung $D(g \circ f)(x) : \mathbb{R}^n \to \mathbb{R}^r$ mit den Ableitungen $Df(x) : \mathbb{R}^n \to \mathbb{R}^m$ und $Dg(y) : \mathbb{R}^m \to \mathbb{R}^r$ (mit $y = f(x)$) zusammen?

Satz 4.10. Kettenregel
Seien $G \subseteq \mathbb{R}^n$ und $D \subseteq \mathbb{R}^m$ Gebiete. Sei $f : G \to \mathbb{R}^m$ in $x \in G$ differenzierbar, $f(G) \subseteq D$ und $g : D \to \mathbb{R}^r$ in $y = f(x) \in D$ differenzierbar. Dann ist auch $g \circ f : G \to \mathbb{R}^r$ in $x \in G$ differenzierbar und es gilt:

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

4.5.1 Kommentar

(1) Beachte, dass auf der rechten Seite die Hintereinanderausführung der linearen Abbildungen $Dg(y) : \mathbb{R}^m \to \mathbb{R}^r$ und $Df(x) : \mathbb{R}^n \to \mathbb{R}^m$ gemeint ist. Beschreibt man $D(g \circ f)(x)$, $Dg(y)$, $Df(x)$ durch ihre Jacobi-Matrizen, so wird die Hintereinanderausführung bekanntlich zur Matrizen-Multiplikation und man erhält also die Kettenregel dann in der Form ($y = f(x)$):

$$J_{g \circ f}(x) = J_g(y) \cdot J_f(x) \ ,$$

d.h.

$$\frac{\partial (g \circ f)_k}{\partial x_j}(x) = \sum_{i=1}^{m} \frac{\partial g_k}{\partial y_i}(y) \cdot \frac{\partial f_i}{\partial x_j}(x) \quad \text{für } 1 \leq j \leq n \text{ und } 1 \leq k \leq r \ .$$

(2) Ist f nur partiell differenzierbar in x und ist g nur partiell differenzierbar in $y := f(x)$, so ist i.a. $g \circ f$ nicht mehr partiell differenzierbar in x und selbst wenn, so ist dann i.a.

$$J_{g \circ f}(x) \neq J_g\big(f(x)\big) \cdot J_f(x) \ .$$

Definition 4.11.
Seien $(V, \|.\|)$ und $(W, \|.\|)$ normierte Vektorräume endlicher Dimension. Ist $T : V \to W$ eine lineare Abbildung, so definiert man die *(Operator-) Norm von T* durch

$$\|T\| := \sup \{\|Tx\| \mid \|x\| \leq 1\} \ .$$

4.5.2 Kommentar

(1) Ist z.B. $V = \mathbb{R}^n$ mit der euklidischen Norm $\|.\|$ so ist $B^n = \{x \in \mathbb{R}^n | \, \|x\| \leq 1\}$ sicher abgeschlossen und beschränkt und daher kompakt. Da die Abbildung $f : B \to \mathbb{R}$, $x \mapsto \|Tx\|$ stetig ist, nimmt f ihr Supremum an und daher ist $\|T\| < \infty$.

(2) Ist V abstakt mit $n = \dim V$, so ist nach einer Basiswahl V isomorph zu $\mathbb{R}^n$. Nun sind je zwei Normen auf $\mathbb{R}^n$ äquivalent ✎ *Übung* ✐und daher ist auch die Menge $B = \{x \in V | \, \|x\| \leq 1\}$ mit einer beliebigen Norm kompakt. Das zeigt, dass $\|T\| < \infty$ immer richtig ist.

(3) Die Abbildung

$$\|.\| : \mathrm{Hom}(V, W) \to \mathbb{R}_+ \cup \{0\} = [0, \infty)$$

erfüllt die Eigenschaften einer Norm.✎ *Übung* ✐

4.5.3 Bemerkung

(1) Sei $T : V \to W$ linear zwischen endlich dimensionalen normierten Vektorräumen. Dann gilt für alle $x \in V$:

$$\|Tx\| \leq \|T\| \cdot \|x\| \ .$$

(2) Sind U, V, W normierte Vektorräume endlicher Dimension und $\|.\| : \mathrm{Hom}(V, W) \to [0, \infty)$ bzw. $\|.\| : \mathrm{Hom}(U, V) \to [0, \infty)$ und $\|.\| : \mathrm{Hom}(U, W) \to [0, \infty)$ die induzierten Normen, so gilt für alle $T \in \mathrm{Hom}(V, W)$ und $S \in \mathrm{Hom}(W, U)$:

$$\|S \circ T\| \leq \|S\| \cdot \|T\| \ .$$

¡? Beweis ¿!

(1) **O.B.d.A.** sei $x \neq 0$. Sei $y := \frac{x}{\|x\|}$ (also $\|y\| = 1$). Dann ist

$$\|Tx\| = \big\|T(\|x\| \cdot y)\big\| \overset{T \text{ linear}}{=} \big\| \|x\| \, T(y)\big\| \overset{\|\cdot\| \text{ Norm}}{=} \|x\| \cdot \|Ty\| \leq \|T\| \cdot \|x\| \ .$$

(2) Für $x \in V$ mit $\|x\| \leq 1$ ist

$$\|STx\| \overset{4.5.3\,(1)}{\leq} \|S\| \cdot \|Tx\| \overset{4.5.3\,(1)}{\leq} \|S\| \, \|T\| \, \|x\| \leq \|S\| \, \|T\| \ .$$

Also

$$\|ST\| \leq \|S\| \, \|T\| \ .$$

QED.

¡? Beweis ¿! der Kettenregel

Sei $A := Df(x)$ und $B := Dg(y)$, $y := f(x)$. Für $\delta > 0$ klein genug, ist dann also für $\xi, \eta \in B_\delta(0)$.

$$f(x + \xi) = f(x) + A\xi + \varphi(\xi)$$

mit $\lim\limits_{\xi \to 0} \frac{\varphi(\xi)}{\|\xi\|} = 0$ und

$$g(y + \eta) = g(y) + B\eta + \psi(\eta)$$

mit $\lim\limits_{\eta \to 0} \frac{\psi(\eta)}{\|\eta\|} = 0$.
Es folgt

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

(Wenn man bedenkt, dass für $\|\xi\| < \delta_1 < \delta$ auch $\|\eta\| = \|A\xi + \varphi(\xi)\| < \delta$ ist, für $0 < \delta_1 < \delta$ klein genug.) Setzen wir daher:

$$\chi(\xi) := B\varphi(\xi) + \psi(A\xi + \varphi(\xi)) \ ,$$

so erhalten wir

$$g \circ f(x + \xi) = g \circ f(x) + BA\xi + \chi(\xi) \ .$$

Können wir nun zeigen, dass $\frac{\chi(\xi)}{\|\xi\|} \underset{0}{\overset{\xi}{\longrightarrow}} 0$ ist, so folgt, dass $g \circ f$ in x differenzierbar ist mit

$$D(g \circ f)(x) = BA = Dg(y) \circ Df(x) \ .$$

Wegen

$$\|B\varphi(\xi)\| \leq \|B\| \cdot \|\varphi(\xi)\|$$

ist schon einmal

$$\frac{\|B\varphi(\xi)\|}{\|\xi\|} \leq \|B\| \, \frac{\|\varphi(\xi)\|}{\|\xi\|} \xrightarrow[0]{\xi} 0 \; ,$$

weil $\frac{\|\varphi(\xi)\|}{\|\xi\|} \xrightarrow[0]{\xi} 0$
Noch zu zeigen ist also

$$\frac{\|\psi(A\xi + \varphi(\xi))\|}{\|\xi\|} \xrightarrow[0]{\xi} 0 \; .$$

Zunächst: Da $\frac{\varphi(\xi)}{\|\xi\|} \xrightarrow[0]{\xi} 0$, gibt es ein $C > 0 : \frac{\|\varphi(\xi)\|}{\|\xi\|} \leq C \;\; \forall \xi : \;\|\xi\| < \delta_1$ (und $\delta_1 > 0$ klein genug). 30.11.2004

Setzt man $\psi_1(\eta) := \frac{\psi(\eta)}{\|\eta\|}$ für $0 < |\eta| < \delta$, so gilt $\psi(\eta) = \|\eta\| \cdot \psi_1(\eta)$ mit $\psi_1(\eta) \xrightarrow[0]{\eta} 0$.

Jetzt gilt

$$\begin{aligned}
\frac{\|\psi(A\xi + \varphi(\xi))\|}{\|\xi\|} &= \frac{\|A\xi + \varphi(\xi)\| \cdot \|\psi_1(A\xi + \varphi(\xi)\|}{\|\xi\|} \\
&\leq \frac{(\|A\xi\| + \|\varphi(\xi)\|) \cdot \|\psi_1(A\xi + \varphi(\xi)\|}{\|\xi\|} \\
&\leq \frac{\|A\| \, \|\xi\| + C\,\|\xi\|}{\|\xi\|} \cdot \|\psi_1(A\xi + \varphi(\xi)\| \\
&= (\|A\| + C) \underbrace{\left\| \psi_1(\underbrace{A\xi + \varphi(\xi)}_{\to 0}) \right\|}_{\to 0} \; ,
\end{aligned}$$

weil $A\xi + \varphi(\xi) \longrightarrow 0$ für $\xi \longrightarrow 0$ geht. Insgesamt ist also tatsächlich

$$\frac{\chi(\xi)}{\|\xi\|} \longrightarrow 0 \;\; (\text{für } \xi \longrightarrow 0) \; .$$

QED.

4.5.4 Beispiel: Laplace in Polarkoordinaten

Sei $f : \mathbb{R}_+ \times \mathbb{R} \to \mathbb{R}^2$, $f(r,\varphi) = (r\cos\varphi, r\sin\varphi) = (x(\varphi), y(\varphi))$ und $u : G \to \mathbb{R}$ zweimal stetig differenzierbar, $G \subseteq \mathbb{R}^2$ ein Gebiet. Drückt man u und Δu in Polarkoordinaten aus, d.h. betrachtet man $u \circ f$ und $(\Delta u) \circ f$, so möchte man $(\Delta u) \circ f$ durch Differenzieren an $u \circ f$ ausdrücken (also in $D_r D_r, D_r D_\varphi, D_\varphi D_\varphi$ und ggf. niedrigere Ordnungen). Behauptung:

$$(\Delta u) \circ f = \frac{\partial^2 (u \circ f)}{\partial r^2} + \frac{1}{r^2} \frac{\partial^2 (u \circ f)}{\partial \varphi^2} + \frac{1}{r} \frac{\partial (u \circ f)}{\partial r} \tag{4.1}$$

Denn (schreibe $\frac{\partial u}{\partial r}$ statt $\frac{\partial (u \circ f)}{\partial r}$ usw.) nach der Kettenregel ist:

$$\begin{aligned}
\frac{\partial u}{\partial r} &= \frac{\partial u}{\partial x}\frac{\partial x}{\partial r} + \frac{\partial u}{\partial y}\frac{\partial y}{\partial r} \\
&= \cos(\varphi)\frac{\partial u}{\partial x} + \sin(\varphi)\frac{\partial u}{\partial y} \; .
\end{aligned}$$

Damit

$$
\begin{aligned}
\frac{\partial^2 u}{\partial r^2} &= \cos(\varphi)\frac{\partial}{\partial r}\frac{\partial u}{\partial x}+\sin(\varphi)\frac{\partial}{\partial r}\frac{\partial u}{\partial y}\\
&= \cos(\varphi)\left(\frac{\partial u^2}{\partial x^2}\frac{\partial x}{\partial r}+\frac{\partial u^2}{\partial y\partial x}\frac{\partial y}{\partial r}\right)+\sin(\varphi)\left(\frac{\partial^2 u}{\partial x\partial y}\frac{\partial x}{\partial r}+\frac{\partial^2 u}{\partial y^2}\frac{\partial y}{\partial r}\right)\\
&= \cos^2(\varphi)\frac{\partial^2 u}{\partial x^2}+\sin(\varphi)\cos(\varphi)\frac{\partial^2 u}{\partial y\partial x}+\sin(\varphi)\cos(\varphi)\frac{\partial^2 u}{\partial x\partial y}+\sin^2(\varphi)\frac{\partial^2 u}{\partial y^2}\\
&\overset{\substack{\text{H.A.Schwarz}\\ \text{Satz 4.6.}}}{=} \cos^2(\varphi)\frac{\partial^2 u}{\partial x^2}+2\sin(\varphi)\cos(\varphi)\frac{\partial^2 u}{\partial x\partial y}+\sin^2(\varphi)\frac{\partial^2 u}{\partial y^2}
\end{aligned}
$$

Ähnlich:

$$
\begin{aligned}
\frac{\partial^2 u}{\partial \varphi^2} &= -r\cos(\varphi)\frac{\partial u}{\partial x}+r^2\sin(\varphi)\frac{\partial^2 u}{\partial x^2}-r^2\sin(\varphi)\cos(\varphi)\frac{\partial^2 u}{\partial y\partial x}\\
&\quad -r\sin(\varphi)\frac{\partial u}{\partial y}-r^2\sin(\varphi)\cos(\varphi)\frac{\partial^2 u}{\partial x\partial y}+r^2\cos^2(\varphi)\frac{\partial^2 u}{\partial y^2}\ .
\end{aligned}
$$

Es folgt

$$
\frac{\partial^2(u\circ f)}{\partial r^2}+\frac{1}{r^2}\frac{\partial^2(u\circ f)}{\partial\varphi^2}+\frac{1}{r}\frac{\partial(u\circ f)}{\partial r}=\left(\frac{\partial^2 u}{\partial x^2}+\frac{\partial^2 u}{\partial y^2}\right)\circ f\ .
$$

Frage: Wie kommt man auf den Ausdruck Gleichung (4.1)?

Dazu schränkt man die Abbildung f auf $\mathbb{R}_+\times(-\pi,\pi)=:G$ ein. Dann erhält man, dass f ein *Diffeomorphismus* ist (d.h. f ist differenzierbar, bijektiv und $g=f^{-1}$ ist differenzierbar). Ihre Umkehrung $g:D\to G\ ,\ D=\operatorname{Bild}(f)\ ,\ g(x,y)=f^{-1}(x,y)=(r(x,y),\varphi(x,y))$ ist gegeben durch:

$$
r(x,y)=\sqrt{x^2+y^2}
$$

$$
\varphi(x,y)=\begin{cases}
\arctan\frac{y}{x} & \text{für}\ \ x>0,y>0\\
\frac{\pi}{2} & \text{für}\ \ x=0,y>0\\
\arctan\frac{y}{x}+\pi & \text{für}\ \ x<0,y>0\\
\arctan\frac{y}{x}-\pi & \text{für}\ \ x<0,y<0\\
-\frac{\pi}{2} & \text{für}\ \ x=0,y<0\\
\arctan\frac{y}{x} & \text{für}\ \ x>0,y\leq 0
\end{cases}\ .
$$

Wie bekommt man $(\Delta u)\circ f$?

Es ist (wir unterdrücken wieder f und g):

$$
\frac{\partial u}{\partial x}=\frac{\partial u}{\partial r}\frac{\partial r}{\partial x}+\frac{\partial u}{\partial\varphi}\frac{\partial\varphi}{\partial x}\ ,
$$

$$
\frac{\partial u}{\partial y}=\frac{\partial u}{\partial r}\frac{\partial r}{\partial y}+\frac{\partial u}{\partial\varphi}\frac{\partial\varphi}{\partial y}\ .
$$

Dann folgt mit der Produktregel

$$
\begin{aligned}
\frac{\partial^2 u}{\partial x^2} &= \left[\frac{\partial}{\partial x}\left(\frac{\partial u}{\partial r}\right)\frac{\partial r}{\partial x}+\frac{\partial u}{\partial r}\frac{\partial^2 r}{\partial x^2}\right]+\left[\frac{\partial}{\partial x}\left(\frac{\partial u}{\partial\varphi}\right)\frac{\partial\varphi}{\partial x}+\frac{\partial u}{\partial\varphi}\frac{\partial^2\varphi}{\partial x^2}\right]\\
&= \left[\frac{\partial^2 u}{\partial y^2}\left(\frac{\partial r}{\partial x}\right)^2+\frac{\partial^2 u}{\partial\varphi\partial r}\frac{\partial\varphi}{\partial x}\frac{\partial r}{\partial x}+\frac{\partial u}{\partial r}\frac{\partial^2 r}{\partial x^2}\right]+\left[\frac{\partial^2 u}{\partial r\partial\varphi}\frac{\partial r}{\partial x}\frac{\partial\varphi}{\partial x}+\frac{\partial^2 u}{\partial\varphi^2}\left(\frac{\partial\varphi}{\partial x}\right)^2+\frac{\partial u}{\partial\varphi}\frac{\partial^2\varphi}{\partial x^2}\right]\\
&= \left(\frac{\partial r}{\partial x}\right)^2\frac{\partial^2 u}{\partial r^2}+\left(2\frac{\partial r}{\partial x}\frac{\partial\varphi}{\partial x}\right)\frac{\partial^2 u}{\partial r\partial\varphi}+\left(\frac{\partial\varphi}{\partial x}\right)^2\frac{\partial^2 u}{\partial\varphi^2}+\left(\frac{\partial^2 r}{\partial x^2}\right)\frac{\partial u}{\partial r}+\left(\frac{\partial^2\varphi}{\partial x^2}\right)\frac{\partial u}{\partial\varphi}\ .
\end{aligned}
$$

Ähnlich für $\frac{\partial^2 u}{\partial y^2}$ indem man überall x durch y ersetzt. Damit folgt

$$\begin{aligned}\Delta u &= \left(\left(\frac{\partial r}{\partial x}\right)^2 + \left(\frac{\partial r}{\partial y}\right)^2\right)\frac{\partial^2 u}{\partial r^2} + \left(2\frac{\partial r}{\partial x}\frac{\partial \varphi}{\partial x} + 2\frac{\partial r}{\partial y}\frac{\partial \varphi}{\partial y}\right)\frac{\partial^2 u}{\partial r \partial \varphi} \\ &\quad + \left(\left(\frac{\partial \varphi}{\partial x}\right)^2 + \left(\frac{\partial \varphi}{\partial y}\right)^2\right)\frac{\partial^2 u}{\partial \varphi^2} + \left(\frac{\partial^2 r}{\partial x^2} + \frac{\partial^2 r}{\partial y^2}\right)\frac{\partial u}{\partial r} + \left(\frac{\partial^2 \varphi}{\partial x^2} + \frac{\partial^2 \varphi}{\partial y^2}\right)\frac{\partial u}{\partial \varphi} \ .\end{aligned}$$

Um nun die Koeffizienten $\frac{\partial r}{\partial x}, \frac{\partial r}{\partial y}, \frac{\partial^2 r}{\partial x^2}, \frac{\partial^2 r}{\partial y^2}, ..$ auszurechen, differenziert man nicht etwa $g(x,y) = (r(x,y), \varphi(x,y))$ (bei komplizierteren Koordinatenransformationen $f : G \to D$ kann man $g = f^{-1}$ oft gar nicht explizit angeben), sondern benutzt erneut die Kettenregel: Weil nämlich $g \circ f = \mathrm{id}$ ist, ist nach der Kettenregel $Dg(y) \cdot Df(x) = D(g \circ f)(x) = D(\mathrm{id})(x) = \mathrm{id}$ also $Dg(y) = (Df(x))^{-1}$. (Man sieht schon: Ist $f : G \to D$ ein Diffeomorphismus, so muss $Df(x) \in Gl_n(\mathbb{R})\ \forall x \in G$ sein!) Bei uns bedeutet dies

$$Dg(y) = \begin{pmatrix} \cos\varphi & -r\sin\varphi \\ \sin\varphi & r\cos\varphi \end{pmatrix}^{-1} = \frac{1}{r}\begin{pmatrix} r\cos\varphi & r\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix}$$

nach der Formel

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \frac{1}{ad - bc}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix} .$$

Es ist also:

$$\frac{\partial r}{\partial x} = \cos\varphi \ , \qquad \frac{\partial r}{\partial y} = \sin\varphi \ ,$$
$$\frac{\partial \varphi}{\partial x} = -\frac{1}{r}\sin\varphi \ , \quad \frac{\partial \varphi}{\partial y} = \frac{1}{r}\cos\varphi \ .$$

Für $\frac{\partial^2 r}{\partial x^2}, \frac{\partial^2 r}{\partial y^2}, \frac{\partial^2 \varphi}{\partial x^2}, \frac{\partial^2 \varphi}{\partial y^2}$ benutzt man schließlich erneut die Kettenregel

$$\begin{aligned}\frac{\partial^2 r}{\partial x^2} &= \frac{\partial}{\partial x}\cos\varphi \\ &= \underbrace{\frac{\partial}{\partial r}\cos\varphi}_{=0} \cdot \frac{\partial r}{\partial x} + \underbrace{\frac{\partial}{\partial \varphi}\cos\varphi}_{-\sin\varphi} \cdot \frac{\partial \varphi}{\partial x} \\ &= \frac{1}{r}\sin^2(\varphi) \ .\end{aligned}$$

Ähnlich ist:

$$\begin{aligned}\frac{\partial^2 r}{\partial y^2} &= \frac{\cos^2(\varphi)}{r} \ , \\ \frac{\partial^2 \varphi}{\partial x^2} &= 2\frac{\sin\varphi\cos\varphi}{r^2} \ , \\ \frac{\partial^2 \varphi}{\partial y^2} &= -2\frac{\sin\varphi\cos\varphi}{r^2} \ .\end{aligned}$$

Setzt man nun alles ein, so erhält man tatsächlich

$$\Delta u = \frac{\partial^2 u}{\partial r^2} + \frac{1}{r}\frac{\partial^2 u}{\partial \varphi^2} + \frac{1}{r}\frac{\partial u}{\partial r} \ .$$

(und dann gilt sie auch für $f : \mathbb{R}_+ \times \mathbb{R} \to \mathbb{R}^2$ wie gesehen). 02.12.2004

4.6 Mittelwertsatz

Motivation

(1) Der wichtige Mittelwertsatz der Differentialrechnung in einer Veränderlichen lässt sich für eine stetig differenzierbare Funktion $f : [a,b] \to \mathbb{R}$ aus dem Hauptsatz wie folgt ableiten:

$$\begin{aligned} f(b) - f(a) &= \int_a^b f'(x)\,\mathrm{d}x \overset{\substack{x=a+t(b-a)\\ =:a+th}}{=} \int_0^1 f'(a+th)h\,\mathrm{d}t \\ &= \left(\int_0^1 f'(a+th)\,\mathrm{d}t\right) h \\ &= f'(a+\Theta h)h \end{aligned}$$

mit einer Zahl $\Theta \in [0,1]$ nach dem Mittelwertsatz der Integralrechung.

(2) Für die Verallgemeinerung in mehreren Veränderlichen vereinbaren wir für eine stetig differenzerbare Abbildung $A : \mathbb{R} \supseteq I \to M(m \times n, \mathbb{R})$, $t \to A(t)$, dass $\int_{t_1}^{t_2} A(\tau)\,\mathrm{d}\tau$ die $m \times n$ Matrix ist, die die Einträge $\int_{t_1}^{t_2} a_{ij}(\tau)\,\mathrm{d}\tau$ an der (i,j)-ten Stelle hat $(1 \le i \le m\ ,\ 1 \le j \le n)$.

Satz 4.12. Mittelwertsatz der Differentialrechnung
Sei $G \subseteq \mathbb{R}^n$ ein Gebiet und $f : G \to \mathbb{R}^m$ stetig differenzierbar. Sei $x \in G$ und $h \in \mathbb{R}$ derart, das die ganze Strecke $\{x + th \in \mathbb{R}^n | t \in [0,1]\}$ noch ganz in G liegt. Dann gilt:

$$f(x+th) - f(x) = \left(\int_0^1 Df(x+th)\,\mathrm{d}t\right) \cdot h$$

4.6.1 Kommentar

Man beachte, dass man in höheren Dimensionen das Integral

$$\int_0^1 Df(x+th)\,\mathrm{d}t$$

i.a. nicht durch den Wert des Integranden an einer Zwischenstelle $x + \Theta h$ ersetzen kann.

¡? Beweis ¿!

Setzen wir $\alpha : [0,1] \to G$, $\alpha(t) = x + th$, so ist

$$\begin{aligned} f(x+h) - f(x) &= f(\alpha(1)) - f(\alpha(0)) \\ &= \int_0^1 \frac{\mathrm{d}}{\mathrm{d}t}(f \circ \alpha)(t)\,\mathrm{d}t \\ &\overset{\text{Kettenregel}}{=} \int_0^1 Df\big(\alpha(t)\big)D\alpha(t)\,\mathrm{d}t \end{aligned}$$

nach dem Hauptsatz der Differenzial- und Integralrechung, angewendet auf die Funktion $f_i \circ \alpha : [0,1] \to \mathbb{R}$ $(i = 1, \dots, m)$. Es ist aber $\dot{\alpha}(t) := D\alpha(t) = h$ unabhängig von t, also ist

$$f(x+h) - f(x) = \left(\int_0^1 Df(x+th)\,\mathrm{d}t\right) h\ .$$

QED.

4.6.2 Kommentar

Die wichtige Konsequenz aus dem Mittelwertsatz für Funktionen in einer Veränderlichen, dass nämlich die Differenz der Funktionswerte durch die Differenz der Argumente mal einer *Schranke* auf die 1. Ableitung abgeschätzt werden kann, liefert aber auch diese integrierte Form des Mittelwertsatzes in mehreren Veränderlichen noch:

Korollar 4.13. Schrankensatz
Ist G ein Gebiet und $x, x+h \in G$ wie in den Voraussetzungen des Mittelwertsatzes gegeben und $f : G \to \mathbb{R}^m$ stetig differenzierbar, so gilt mit $M := \sup\{\|Df(x+th)\| \,|t \in [0,1]\} < \infty$:

$$\|f(x+h) - f(x)\| \le M\,\|h\| \ .$$

Lemma 4.14.
Ist $\alpha : [0,1] \to \mathbb{R}^n$ eine stetig Abbildung und $\|.\|$ die euklidische Norm auf $\mathbb{R}^n$, so gilt:

$$\left\|\int_0^1 \alpha(t)\ \mathrm{d}t\right\| \le \int_0^1 \|\alpha(t)\|\ \mathrm{d}t$$

¡? Beweis ¿! des Schrankensatzes

$$\begin{aligned}
\|f(x+h) - f(x)\| &\overset{\text{Satz 4.12.}}{=} \left\|\left(\int_0^1 Df(x+th)\ \mathrm{d}t\right) h\right\| \\
&= \left\|\left(\int_0^1 Df(x+th)h\ \mathrm{d}t\right)\right\| \\
&\overset{\text{Lemma}}{\le} \left(\int_0^1 \|Df(x+th)h\|\ \mathrm{d}t\right) \\
&\overset{4.5.3}{\le} \left(\int_0^1 \|Df(x+th)\|\,\|h\|\ \mathrm{d}t\right) \\
&\le \int_0^1 M\ \mathrm{d}t \cdot \|h\| \\
&= M\,\|h\| \ .
\end{aligned}$$

QED.

¡? Beweis ¿! des Lemmas

Für das kanonische Skalarprodukt $\langle\ \cdot\ ,\ \cdot\ \rangle : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ hat man mit $I := \int_0^1 \alpha(t)\ \mathrm{d}t \in \mathbb{R}^n$ (**o.B.d.A.** $I \neq 0$):

$$\begin{aligned}
\|I\|^2 &= \langle I, I\rangle = \left\langle \int_0^1 \alpha(t)\ \mathrm{d}t, I\right\rangle \\
&\overset{\text{Bilinearität}}{=} \int_0^1 \langle \alpha(t), I\rangle\ \mathrm{d}t \\
&\overset{\substack{\text{Cauchy-Schwarz}\\ 1.2.3}}{\le} \int_0^1 \|\alpha(t)\|\,\|I\|\ \mathrm{d}t \\
&= \|I\| \int_0^1 \|\alpha(t)\|\ \mathrm{d}t \ .
\end{aligned}$$

Damit folgt

$$\left\|\int_0^1 \alpha(t)\ \mathrm{d}t\right\| \le \int_0^1 \|\alpha(t)\|\ \mathrm{d}t \ .$$

QED.

Korollar 4.15.
Seien $G \subseteq \mathbb{R}^n$ ein Gebiet und $f : G \to \mathbb{R}$ stetig differenzierbar, so dass überall $\operatorname{grad} f(x) = 0$ ist. Dann muss f konstant sein.

¡? Beweis ¿!

Sei $x, y \in G$ und $\alpha : [0,1] \to G$ ein (stetiger) Weg von $x = \alpha(0)$ nach $y = \alpha(1)$[1]. Wir setzen

$$t_0 := \sup\{t \in [0,1] | f(\alpha(\tau)) = f(x) \ , \ \forall \tau \in [0,t]\} \ .$$

Wegen der Stetigkeit von f in $\alpha(t_0)$ ist dann auch

$$f\big(\alpha(t_0)\big) = \lim_{\substack{\tau \to t_0 \\ \tau < t_0}} \underbrace{\alpha(\tau)}_{f(x)} = f(x) \ .$$

Angenommen $t_0 < 1$. Dann gibt es ein $\delta > 0$, so dass $B_\delta(\alpha(t_0)) \subseteq G$ und ein h mit $0 < h < \delta$, so dass auch $\alpha(t_0 + h) = B_\delta(\alpha(t_0))$ ist. Da die ganze Strecke $\big\{\alpha(t_0) + s\big(\alpha(t_0 + h) - \alpha(t_0)\big)\big| s \in [0,1]\big\}$ in $B_\delta(\alpha(t_0)) \subseteq G$ ist, gilt nach dem Schrankensatz für alle $h' \in [0,h]$:

$$\|f(\alpha(t_0 + h')) - f(\alpha(t_0)\| \leq M \, \|\alpha(t_0 + h') - \alpha(t_0)\|$$

mit

$$M = \sup\{\big\|\operatorname{grad} f\,\big(\alpha(t_0) + s\big(\alpha(t_0 + h') - \alpha(t_0)\big)\big)\big\| \, | s \in [0,1]\} = 0 \ .$$

Also ist $f(\alpha(t_0 + h')) = f(\alpha(t_0)) = f(x) \quad \forall h' \in [0,h]$ im ↯ Widerspruch ↯ dazu, dass t_0 das Supremum ist. Daraus folgt $t_0 = 1$, also

$$f(y) = f(\alpha(1)) = f(x)$$

und f konstant.

QED.

[1] Oder direkter (wenn α sogar stetig differenzierbar ist):

$$f(y) - f(x) = \int_0^1 \frac{\mathrm{d}}{\mathrm{d}t}(f \circ \alpha)(t)\ \mathrm{d}t = \int_0^1 \Big\langle \underbrace{\operatorname{grad} f\big(\alpha(t)\big)}_{=0}, \dot{\alpha}(t) \Big\rangle\ \mathrm{d}t = 0$$

Kapitel 5

Lokale Extrema

02.12.2004

Motivation

Um nun auch eine „Graphendiskussion“ für Funktionen $f : G \to \mathbb{R}$ in mehreren Veränderlichen durchführen zu können, $G \subseteq \mathbb{R}^n$ ein Gebiet, brauchen wir insbesondere eine „Taylorentwicklung“ von f bis zur zweiten Ordnung. Nun ist aber die erste Ableitung von f in x ein Vektor $\operatorname{grad} f(x)$ (bzw. eine lineare Abbildung $\mathbb{R}^n \to \mathbb{R}$) und die zweite Ableitung $\operatorname{Hess} f(x)$ eine Matrix (bzw., wie sich zeigen wird, eine *quadratische Form* $\mathbb{R}^n \to \mathbb{R}$) und höhere Ableitungen werden noch komplizierter. Wir erwarten etwa in Analogie zum eindimensionalen Fall, wo gilt

$$f(x+h) = f(x) + f'(x) \cdot h + \frac{1}{2} f''(x) h^2 + o\Big(|h|^2\Big) \ ,$$

etwa folgendes:

$$f(x+h) = f(x) + \underbrace{\sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(x) h_i}_{\langle \operatorname{grad} f(x), h\rangle} + \frac{1}{2} \underbrace{\sum_{i,j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j}(x) h_i h_j}_{\langle h, \operatorname{Hess} f(x), h\rangle} + o\Big(\|h\|^2\Big) \ .$$

07.12.2004

5.1 Taylor-Entwicklung

5.1.1 Notation

Um höhere Terme in der Taylor-Entwicklung günstig zu notieren, führt man für $\alpha = (\alpha_1, \cdots, \alpha_n) \in \mathbb{N}_0^n$ ein

$$|\alpha| := \alpha_1 + \cdots + \alpha_n \ ,$$

$$\alpha! := \alpha_1! + \cdots + \alpha_n! \ ,$$

für $f : G \to \mathbb{R}$ eine $|\alpha|$ -mal stetig differenzierbare Funktion:

$$D^\alpha f := D_1^{\alpha_1} \cdot \cdots \cdot D_n^{\alpha_n} f = \frac{\partial^{|\alpha|} f}{\partial x_1^{\alpha_1} \cdots \partial x_n^{\alpha_n}} \ ,$$

wobei $D_j^{\alpha_j} = D_j \cdot \cdots \cdot D_j$ (α_j-mal) ist und schließlich für $h = (h_1, \cdots, h_n) \in \mathbb{R}^n$:

$$h^\alpha := h_1^\alpha \cdots h_n^{\alpha_n} \ .$$

Lemma 5.1.
Sei $G \subseteq \mathbb{R}^n$ ein Gebiet, $f : G \to \mathbb{R}$ sei k-mal stetig differenzierbar, $x \in G$ und $h \in \mathbb{R}^n$ derart, dass die geradlinige Verbindung $[x, x+h] := \{x + t \cdot h \in \mathbb{R}^n | t \in [0,1]\}$ ganz in G liegt. Dann ist die Funktion $\varphi : [0,1] \to \mathbb{R}$, $\varphi(t) = f(x+th)$ auch k-mal stetig differenzierbar und es gilt:

$$\left(\varphi^{(k)}(t)\right) = \frac{d^k\varphi}{dt^k}(t) = \sum_{|\alpha|=k} \frac{k!}{\alpha!} D^\alpha f(x+th) \cdot h^\alpha$$

5.1.2 Kommentar

(1) $\sum\limits_{|\alpha|=k}$ bedeutet, dass sich die Summe über alle n-Tupel $\alpha \in \mathbb{N}_0^n$ erstreckt, für die $|\alpha| = k$ ist. (Wie viele gibt es davon? Antwort: $\binom{n+1+k}{k}$)

(2) Ist $\alpha = (\alpha_1, \cdots, \alpha_n) \in \mathbb{N}_0^n$ mit $|\alpha| = k$, so gibt $\frac{k!}{\alpha!}$ gerade die Anzahl der Möglichkeiten an, eine k-elementige Teilmenge M in n Teilmengen $S_1, \cdots, S_n$ disjunkt zu zerlegen, $M = S_1 \dot\cup \cdots \dot\cup S_n$, so dass S_i gerade α_i Elemente hat $(i = 1, \cdots, n)$(z.B. für $n = 2$: $\frac{k!}{\alpha_1!\alpha_2!} = \frac{k!}{\alpha!_1(k-\alpha_1)!} = \binom{k}{\alpha_1}$).
✎ *Übung* ✐

¡? Beweis ¿!

(1) Nach der Kettenregel ist

$$\dot\varphi(t) = Df(x+th) \cdot h = \sum_{j=1}^n \frac{\partial f}{\partial x_j}(x+th) \cdot h_j \ .$$

Nach der Kettenregel erhält man weiter

$$\ddot\varphi(t) = \sum_{j=1}^n \sum_{i=1}^n \frac{\partial^2 f}{\partial x_i \partial x_j}(x+th) \cdot h_i h_j \ .$$

Schließlich erhält man

$$\varphi^{(k)}(t) = \sum_{j_1,\cdots,j_k=1}^n D_{j_1} \cdot \cdots \cdot D_{j_k} f(x+th) h_{j_1} \cdots h_{j_k} \ .$$

(2) Kommt nun in dem k-Tupel $(j_1, \cdots, j_k)$ die 1 α_1-mal, die 2 α_2-mal, ... vor, so gilt nach dem Satz vom Schwarz (Satz 4.6.), dass

$$D_{j_1} D_{j_2} \cdots D_{j_k} f(x+th) h_{j_1} \cdots h_{j_k} = D_1^{\alpha_1} \cdots D_n^{\alpha_n} f(x+th) h_1^{\alpha_1} \cdots h_n^{\alpha_n} = D^\alpha f(x+th) h^\alpha \ ,$$

wobei $|\alpha| = k$ sein muss. Es gibt nun aber gerade $\frac{k!}{\alpha_1! \cdots \alpha_n!}$ k-Tupel, in denen 1 α_1-mal, 2 α_2-mal, ... auftritt. Man erhält also durch Zusammenfassung der Summanden tatsächlich

$$\varphi^{(k)}(t) = \sum_{i_1,\cdots,i_k=1}^n D_{i_1} \cdot \cdots \cdot D_{i_k} f(x+th) h_{i_1} \cdots h_{i_k} = \sum_{|\alpha|=k} \frac{k!}{\alpha!} D^\alpha f(x+th) h^\alpha \ .$$

QED.

Satz 5.2. Taylor
Sei $G \subseteq \mathbb{R}^n$ ein Gebiet, $k \in \mathbb{N}_0$ und $f : G \to \mathbb{R}$ $(k+1)$-mal stetig differenzierbar. Sei $x \in G$ und $h \in \mathbb{R}^n$ derart, dass die Verbindungsstrecke $[x, x+h]$ ganz in G verläuft. Dann gibt es ein $\Theta \in [0,1]$, so dass gilt:

$$f(x+th) = \sum_{|\alpha| \le k} \frac{D^\alpha f(x)}{\alpha!} h^\alpha + \sum_{|\alpha|=k+1} \frac{D^\alpha f}{\alpha!}(x + \Theta h) h^\alpha \ .$$

5.1.3 Kommentar

(1) Man nennt

$$P^{(k)}_{f,x}(h) = \sum_{|\alpha|\le k} \frac{D^\alpha f(x)}{\alpha!} h^\alpha$$

das *Taylor-Polynom vom Grad k von f im Punkt x*. Schreiben wir

$$P^{(k)}_{f,x} = P_0(h) + P_1(h) + \cdots + P_k(h)$$

wo $P_i(h) \in \mathbb{R}[h]$ der i-te homogene Bestandteil ist, so ist also

$$\begin{aligned}
P_0(h) &= f(x)\ , \\
P_1(h) &= \sum_{j=1}^n \cdot \frac{D_i f(x)}{1} \cdot h_j = \langle \operatorname{grad} f(x), h\rangle\ , \\
P_2(h) &= \sum_{|\alpha|=2} \frac{D^\alpha f(x)}{\alpha!} h^\alpha = \sum_{j=1}^n \frac{D_j^2 f(x)}{2} h_j^2 + \sum_{i<j} \frac{D_i D_j f(x)}{1} h_i h_j \\
&= \frac{1}{2} \sum_{j=1}^n \frac{\partial^2 f}{\partial x_j^2}(x) h_j^2 + \frac{1}{2} \sum_{i\ne j} \frac{\partial^2 f}{\partial x_i \partial x_j}(x) h_i h_j = \frac{1}{2} \sum_{i,j=1}^n \frac{\partial^2 f}{\partial x_i \partial x_j}(x) h_i h_j \\
&= \frac{1}{2} \langle h, \operatorname{Hess} f(x) h\rangle\ .
\end{aligned}$$

(2) Vom „Restterm“

$$R^{(k)}_{f,x}(h) := \sum_{|\alpha|=k+1} \frac{D^\alpha f(x+\Theta h)}{\alpha!} h^\alpha =: R(h)$$

wird sich herausstellen, dass er $o\Big((\|h\|^k)\Big)$ ist, also $\lim_{h\to 0} \frac{R(h)}{\|h\|^k} = 0$, so dass also das Taylor Polynom $P(h) := P^{(k)}_{f,x}(h)$ eine Approximation von f in einer Umgebung $B_\delta(x) \subseteq G$ ($\delta > 0$ klein genug) ist, die besser als von k-ter Ordnung ist.

¡? Beweis ¿! der Taylor-Formel

Betrachte die Kurve $\gamma : [0,1] \to G\ ,\quad t \mapsto x + th$ und setze $\varphi : [0,1] \to \mathbb{R}\ ,\quad \varphi(t) = f(\gamma(t)) = f(x+th)$. Es ist dann φ $(k+1)$-mal stetig differenzierbar, also gibt es nach der Taylor-Formel für Funktionen in einer Veränderlichen ein $\Theta \in [0,1]$ so, dass gilt

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

Nach Lemma 5.1. erhält man

$$\begin{aligned}
f(x+h) &= \varphi(1) = \sum_{m=0}^k \frac{1}{m!} \left(\sum_{|\alpha|=m} \frac{m!}{\alpha!} D^\alpha f(x) h^\alpha \right) + \frac{1}{(k+1)!} \sum_{|\alpha|=k+1} \frac{(k+1)!}{\alpha!} D^\alpha f(x+\Theta h) h^\alpha \\
&= \sum_{|\alpha|\le k} \frac{D^\alpha f(x)}{\alpha!} h^\alpha + \sum_{|\alpha|=k+1} \frac{1}{\alpha!} D^\alpha f(x+\Theta h) h^\alpha\ .
\end{aligned}$$

QED.

09.12.2004

Korollar 5.3.
Sei $G \subseteq \mathbb{R}^n$ ein Gebiet und $f : G \to \mathbb{R}$ k-mal stetig differenzierbar ($k \in \mathbb{N}$). Sei $x \in G$ und $\delta > 0$ so klein, dass $B_\delta(x) \subseteq G$ ist. Es gibt dann eine Funktion $\varphi : B_\delta(0) \to \mathbb{R}$ mit

$$\lim_{h\to 0} \frac{\varphi(h)}{\|h\|^k} = 0 \ ,$$

so dass für alle $h \in B_\delta(0)$ gilt:

$$f(x+h) = \sum_{|\alpha|\le k} \frac{D^\alpha f(x)}{\alpha!} h^\alpha + \varphi(h) \ .$$

Kurz, wenn $P^{(k)}(h) = P^{(k)}_{f,x}(h)$ das k-te Taylor-Polynom von f in x bezeichnet:

$$f(x+h) = P^{(k)}(h) + o\left(\|h\|^k\right) \ .$$

¡? Beweis ¿!

Nach Taylors Satz gibt es für jedes $h \in B_\delta(0)$ ein $\Theta \in [0,1]$, $\Theta = \Theta(h)$, so dass gilt:

$$f(x+h) = \sum_{|\alpha|\le k-1} \frac{D^\alpha f(x)}{\alpha!} h^\alpha + \sum_{|\alpha|=k} \frac{D^\alpha f(x+\Theta h)}{\alpha!} h^\alpha \ .$$

Wir setzen $\varphi : B_\delta(0) \to \mathbb{R}$,

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

Dann gilt offenbar:

$$f(x+h) = \sum_{|\alpha|\le k} \frac{D^\alpha f(x)}{\alpha!} h^\alpha + \varphi(h) \ .$$

Wegen der Stetigkeit von $D^\alpha f$ in x und weil $x + \Theta h \xrightarrow{h \to 0} x$, folgt:

$$\lim_{h\to 0} D^\alpha f(x+\Theta h) = D^\alpha f(x) \quad \forall \alpha : |\alpha| = k \ .$$

Deshalb ist:

$$\begin{aligned} \frac{|\varphi(h)|}{\|h\|^k} &\le \frac{1}{\|h\|^k} \sum_{|\alpha|\le k} \frac{1}{\alpha!} \left|D^\alpha f(x+\Theta h) - D^\alpha f(x)\right| \cdot |h^\alpha| \\ &= \frac{1}{\alpha!} \sum_{|\alpha|\le k} \underbrace{\left|D^\alpha f(x+\Theta h) - D^\alpha f(x)\right|}_{\xrightarrow{h \to 0}\ 0} \cdot \underbrace{\frac{|h_1^{\alpha_1}| \cdots |h_n^{\alpha_n}|}{\|h\|^k}}_{\le 1} \xrightarrow{h\to 0} 0 \ , \end{aligned}$$

weil z.B.

$$\frac{|h_1|}{\|h\|} = \frac{|h_1|}{\sqrt{h_1^2 + \cdots + h_n^2}} \le 1$$

ist. Also ist tatsächlich

$$\varphi(h) = o\left(\|h\|^k\right) \ .$$

QED.

5.1.4 Kommentar

(1) Für $k = 0$ erhält man folgende Aussage: „eine stetige Funktion ist in $x \in G$ stetig“, denn

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

mit $\lim\limits_{h \to 0} \varphi(h) = 0$, also

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

(2) Für $k = 1$ erhält man die Aussage, dass eine stetig differenzierbare Funktion $f : G \to \mathbb{R}$ „in jedem Punkt $x \in G$ differenzierbar ist“, denn:

$$f(x+h) = f(x) + \underbrace{\sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(x) h_i}_{Df(x)h} + \varphi(h)$$

mit $\lim\limits_{h \to 0} \frac{\varphi(h)}{\|h\|} = 0$.

(3) Aber für $k = 2$ erhält man nun die (sehr nützliche) Aussage, dass eine 2-mal stetig differenzierbare Funktion $f : G \to \mathbb{R}$ die folgende Darstellung erlaubt ($x \in G \ , \quad \|h\| < \delta$):

$$f(x+h) = f(x) + \underbrace{\sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(x) h_i}_{=\langle \operatorname{grad} f(x), h \rangle} + \frac{1}{2} \underbrace{\sum_{i,j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j}(x) h_i h_j}_{=\langle h, \operatorname{Hess} f(x) h \rangle} + \varphi(h)$$

mit $\lim\limits_{h \to 0} \frac{\varphi(h)}{\|h\|^2} = 0$.

5.2 Lokale Extrema

Definition 5.4.
Sei $G \subseteq \mathbb{R}^n$ ein Gebiet und $x \in G$. Man sagt, dass eine Funktion $f : G \to \mathbb{R}$ *in x ein lokales Maximum* hat, wenn es ein $\delta > 0$ gibt, so dass $B_\delta(x) \subseteq G$ ist und für alle $y \in B_\delta(x)$ gilt:

$$f(y) \leq f(x) \ .$$

Ist $f(y) \geq f(x) \ , \quad \forall y \in B_\delta(x)$ für ein $\delta > 0$, so heißt x *ein lokales Minimum von f*. Hat f in $x \in G$ ein lokales Maximum oder ein lokales Minimum, so spricht man allgemeiner von einem *lokalen Extremum*.

5.2.1 Bemerkung

Sei $f : G \to \mathbb{R}$ eine Funktion, $G \subseteq \mathbb{R}^n$ ein Gebiet, $x \in G$ und f habe in x ein lokales Extremum. Ist f in x partiell differenzierbar, so muss gelten:

$$\operatorname{grad} f(x) = 0 \ .$$

¡? Beweis ¿!

Definiert man für $\delta > 0$ klein genug (und $i \in \{1, \ldots, n\}$) die Funktion $g_i : (-\delta, \delta) \to \mathbb{R}$ durch $g_i(t) = f(x + t\hat{e}_i)$, so ist g_i differenzierbar in $t = 0$ und $\dot{g}_i(0) = \frac{\partial f}{\partial x_i}(x)$. Es hat aber nun g_i in $t = 0$ ein lokales Extremum, weshalb $\dot{g}_i(0) = 0$ sein muss. Also muss $\operatorname{grad} f(x) = 0$ gelten.
QED.

Motivation

Das bewiesene Kriterium ist ein sogenanntes *notwendiges Kriterium*, um ein lokales Extremum für eine differenzierbare Funktion $f : G \to \mathbb{R}$ zu finden. Die Lösung der Gleichung $\nabla f(x) = 0$ liefert damit die Kandidaten für die lokalen Extrema. Um nun auch ein *hinreichendes Kriterium* zu finden, führt man folgende Begriffe ein.

Definition 5.5.
Sei V ein endlich dimensionaler $\mathbb{R}$ -Vektorraum. Man nennt eine Abbildung $q : V \to \mathbb{R}$ eine *quadratische Form auf* V, wenn es eine symmetrische Bilinearform $s : V \times V \to \mathbb{R}$ gibt, so dass gilt

$$q(v) = s(v, v) \ , \quad \forall v \in V \ .$$

5.2.2 Beispiel

Ist $V = \mathbb{R}^n$ und A eine symmetrische Matrix, $A \in \mathrm{Sym}_n(\mathbb{R})$, so wird (bekanntlich) durch $s : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$,

$$s(x, y) = \langle x, Ay \rangle = x^T Ay \ ,$$

eine symmetrische Bilinearform auf $\mathbb{R}^n$ gegeben. Die zugehörige quadratische Form $q : \mathbb{R}^n \to \mathbb{R}$ ist also gegeben durch

$$q(x) = \langle x, Ax \rangle = \sum_{i,j=1}^{n} a_{ij} x_i x_j \overset{a_{ij}=a_{ji}}{=} \sum_{i=1}^{n} a_{ii}^2 x_i^2 + 2 \sum_{i<j} a_{ij} x_i x_j \ .$$

5.2.3 Bemerkung

Ist V ein reeller Vektorraum und $q : V \to \mathbb{R}$ die quadratische Form einer Bilinearform $s : V \times V \to \mathbb{R}$, so gilt die sogenannte *Polarisationsformel*:

$$s(v, w) = \frac{1}{4}(q(v + w) - q(v - w)) \ ,$$

für alle $v, w \in \mathbb{R}$.
¡? Beweis ¿!
Ist $q(v) = s(v, v)$, $\forall v \in V$, so folgt aus der Bilinearität und der Symmetrie von s

$$\begin{aligned} q(v \pm w) &= s(v \pm w, v \pm w) \\ &= s(v, v) \pm s(v, w) \pm s(w, v) \pm s(w, w) \\ &= q(v) \pm 2s(v, w) + q(w) \end{aligned}$$

Damit ist auch $q(v + w) - q(v - w) = 4s(v, w)$. Es folgt die Behauptung.
QED.

5.3 Definitheit

Definition 5.6.
Sei V ein endlich dimensionaler $\mathbb{R}$ -Vektorraum und $q : V \to \mathbb{R}$ eine quadratische Form auf V. Dann sagt man:

(a) q ist *positiv definit*, wenn für alle $v \in V \backslash \{0\}$ gilt: $q(v) > 0$

(b) q ist *negativ definit*, wenn für alle $v \in V \backslash \{0\}$ gilt: $q(v) < 0$

(c) q ist *indefinit*, wenn es $v, w \in V$ gibt, so dass gilt: $q(v) < 0$ und $q(w) > 0$

14.12.2004

5.3.1 Kommentar

(1) Ist $(v_1, \ldots, v_n) \subseteq V$ eine Basis von V und q eine quadratische Form auf V zur Bilinearform $s : V \times V \to \mathbb{R}$, so gilt mit der beschreibenden Matrix $A \in \mathrm{Sym}_n(\mathbb{R})$ (d.h. $A = (a_{ij})$, $a_{ij} = s(v_i, v_j)$):

$$q(v) = s(v, v) = \sum_{i,j=1}^{n} s(x_i v_i, x_j v_j) = \sum_{i,j=1}^{n} x_i x_j a_{ij} = \langle x, Ax\rangle \ ,$$

wenn $v = \sum_{i=1}^{n} x_i v_i$, mit $x_i \in \mathbb{R}$ ist. Nach einer Basiswahl ist also jede quadratische Form auf V von der Form wie in 5.2.2. Insbesondere muss man der symmetrischen Matrix $A \in \mathrm{Sym}_n(\mathbb{R})$ entnehmen können, ob q positiv definit ist. Wir sagen, dass $A \in \mathrm{Sym}_n(\mathbb{R})$ positiv definit ist, wenn $q : \mathbb{R}^n \to \mathbb{R}$, $q(x) = \langle x, Ax\rangle$ positiv definit ist, also

$$\langle x, Ax\rangle > 0 \ , \ \forall x \in \mathbb{R}^n \backslash \{0\} \ .$$

(2) Ist $A \in \mathrm{Sym}_n(\mathbb{R})$, so besagt ein wichtiger Satz der linearen Algebra (Spektralsatz für selbstadjungierte Endomorphismen), dass es eine Orthonormalbasis $(e_1, \ldots, e_n) \in \mathbb{R}^n$ gibt, so dass gilt

$$Ae_j = \lambda_j e_j \quad (j = 1, \ldots, n)$$

mit $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$. Hierbei sind $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$ offenbar die Eigenwerte von A.

5.3.2 Bemerkung

Sei $A \in \mathrm{Sym}_n(\mathbb{R})$. Dann ist A genau dann positiv definit, wenn alle Eigenwerte von A positiv sind. A ist negativ definit, wenn alle Eigenwerte negativ sind. A ist indefinit, wenn sowohl positive als auch negative Eigenwerte von A existieren.

$\Rightarrow$ Sei $\lambda \in \mathbb{R}$ ein Eigenwert. Also gibt es ein $x \in \mathbb{R}^n$, $x \neq 0$, so dass $Ax = \lambda x$. Dann ist

$$\lambda \left\| x \right\|^2 = \langle x, \lambda x\rangle = \langle x, Ax\rangle > 0 \Rightarrow \lambda > 0 \ .$$

$\Leftarrow$ Ist $(e_1, \ldots, e_n) \in \mathbb{R}^n$ eine Orthonormalbasis von Eigenvektoren von $A : \mathbb{R}^n \to \mathbb{R}^n$, $Ae_j = \lambda_j e_j$ mit $\lambda_j > 0$, $j = 1, \ldots, n$, so gilt für alle $x \in \mathbb{R}\backslash\{0\}$ folgendes: Ist $x = \sum\limits_{j=1}^{n} x_j e_j$ die Darstellung von x bzgl. $(e_1, \ldots, e_n)$, so ist wenigstens ein $x_j \neq 0$ und es gilt

$$\begin{aligned} \langle x, Ax\rangle &= \sum_{i,j=1}^{n} \langle x_i e_i, A(x_j e_j)\rangle = \sum_{i,j=1}^{n} x_i x_j \langle e_i, Ae_j\rangle \\ &= \sum_{i,j=1}^{n} x_i x_j \underbrace{\langle e_i, \lambda_j e_j\rangle}_{=\lambda_j\langle e_i, e_j\rangle = \lambda_j \delta_{ij}} = \sum_{j=1}^{n} \underbrace{x_j^2}_{\geq 0} \underbrace{\lambda_j}_{>0} > 0 \ , \end{aligned}$$

da mindestens einen Summand > 0. Also ist A positiv definit.

5.3.3 Kommentar

(1) Ist eine symmetrische Matrix in Diagonalform, so sieht man ihr unmittelbar an, ob sie positiv definit ist, denn dann müssen ihre Diagonal-Einträge sämtlich positiv sein. Ist sie nicht in Diagonalform, so ist die Sache nicht offensichtlich, z.B. ist

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

indefinit, denn für $v = (1,1)$ und $w = (1,-1)$ ist

$$\left\langle \begin{pmatrix}1\\1\end{pmatrix}, \underbrace{A\begin{pmatrix}1\\1\end{pmatrix}}_{=\begin{pmatrix}1\\1\end{pmatrix}} \right\rangle = \left\langle \begin{pmatrix}1\\1\end{pmatrix}, \begin{pmatrix}1\\1\end{pmatrix} \right\rangle = 2 > 0 \ ,$$

$$\left\langle \begin{pmatrix}1\\-1\end{pmatrix}, \underbrace{A\begin{pmatrix}1\\-1\end{pmatrix}}_{=\begin{pmatrix}-1\\1\end{pmatrix}} \right\rangle = \left\langle \begin{pmatrix}-1\\1\end{pmatrix}, \begin{pmatrix}1\\-1\end{pmatrix} \right\rangle = -2 < 0 \ .$$

(2) Man müsste also eigentlich alle Eigenwerte von A bestimmen, von denen man weiß, dass sie alle reell sind, also die Nullstellen des charakteristischen Polynoms

$$\chi_A(\lambda) = \det(\lambda \mathbb{1} - A) = \lambda^n - \operatorname{spur}(A)\lambda^{n-1} \pm \cdots + (-1)^n \det A \ ,$$

was i.a. schwer ist. Es gibt aber folgenden *Satz von Hurwitz*, der einem das Leben vereinfacht (`vgl.z.B. G.Fischer: Lineare Algebra, Vieweg-Verlag`): Eine symmetrische Matrix $A = (a_{ij})_{1\le i,j\le n}$ ist genau dann positiv definit, wenn die Determinanten ihrer Hauptminoren alle positiv sind, also

$$\det \begin{pmatrix} a_{11} & \cdots & a_{1k} \\ \vdots & \ddots & \vdots \\ a_{k1} & \cdots & a_{kk} \end{pmatrix} > 0 \qquad (k = 1, \dots, n) \ .$$

Die Matrix A ist negativ definit, falls $-A$ positiv definit ist.

Satz 5.7.
Sei $G \subseteq \mathbb{R}^n$ ein Gebiet und $f : G \to \mathbb{R}$ eine zweimal stetig differenzierbare Funktion. Sei $x \in G$ derart, dass $\operatorname{grad} f(x) = 0$.

(a) Ist die Hesse-Matrix $\operatorname{Hess} f(x)$ positiv definit, so hat f in x ein (sogar striktes) lokales Minimum.

(b) Ist die Hesse-Matrix $\operatorname{Hess} f(x)$ negativ definit, so hat f in x ein (sogar striktes) lokales Maximum.

(c) Ist die Hesse-Matrix $\operatorname{Hess} f(x)$ indefinit, so hat f in x kein lokales Extremum.

Ein lokales Minimum von f heißt strikt, falls $\exists \delta > 0 : \forall y \in B_\delta(x)\backslash\{0\} : f(y) > f(x)$. Analog für Maxima.

Lemma 5.8.
Ist $q : \mathbb{R}^n \to \mathbb{R}$ eine positiv definite quadratische Form, so existiert eine Zahl $\alpha > 0$, so dass für alle $x \in \mathbb{R}^n$ gilt:

$$q(x) \ge \alpha \, \|x\|^2 \ .$$

¡? Beweis ¿!

Sei $S = \{x \in \mathbb{R}^n | \, \|x\| = 1\}$ die Einheitssphäre im $\mathbb{R}^n$. Da S beschränkt und abgeschlossen ist, ist S kompakt. Da q stetig ist ✎ *Übung* ✐, nimmt $q|_S : S \to \mathbb{R}$ ihr Minimum an (Weierstraß). Sei

$$\alpha := \min\{q(x) | x \in S\} > 0 \ .$$

Ist nun $x \in \mathbb{R}^n \backslash \{0\}$ beliebig, so gilt:

$$q(x) = q\left(\|x\| \frac{x}{\|x\|}\right) = \|x\|^2 \cdot q\underbrace{\left(\frac{x}{\|x\|}\right)}_{\in S} \geq \|x\|^2 \cdot \alpha \ ,$$

denn für eine quadratische Form ist stets $q(\lambda x) = \lambda^2 q(x)$, also

$$q(x) \geq \alpha \|x\|^2 \ , \quad \forall x \in \mathbb{R}^n \ ,$$

denn für $x = 0$ ist die Ungleichung offensichtlich auch richtig.
QED.

5.3.4 Beispiele

(I) Sei $f : \mathbb{R}^n \to \mathbb{R}$,

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

$c \in \mathbb{R}$ und $(x_0, y_0) = (0, 0)$. Es ist $\operatorname{grad} f(x, y) = (2x, 2y)$, also $\operatorname{grad} f(0, 0) = (0, 0)$. Weiter ist

$$\operatorname{Hess} f(0, 0) = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$$

offensichtlich positiv definit. Also hat f in $(0, 0)$ ein striktes (sogar globales) Minimum. 16.12.2004

(II) $f(x, y) = -x^2 - y^2 + c \Rightarrow (0, 0)$ ist lokales Maximum.

(III) Die Funktion $f : \mathbb{R}^2 \to \mathbb{R}$, $(x, y) \mapsto x^2 - y^2 + c$ erfüllt

$$\operatorname{grad} f(0, 0) = (0, 0)$$

und

$$\operatorname{Hess} f(0, 0) = \begin{pmatrix} 2 & 0 \\ 0 & -2 \end{pmatrix} \ ,$$

also hat f in $(0, 0)$ kein lokales Extremum.

¡? Beweis ¿! des Satz 5.7.

(a) Sei $\delta > 0$ so klein, dass $B_\delta(x) \subseteq G$ ist. Nach Korollar 5.3. gibt es dann eine Funktion $\varphi : B_\delta(0) \to \mathbb{R}$ mit $\lim\limits_{h \to 0} \frac{\varphi(h)}{\|h\|^2} = 0$, so dass gilt:

$$f(x + h) = f(x) + \langle \operatorname{grad} f(x), h\rangle + \frac{1}{2} \langle h, \operatorname{Hess} f(x), h\rangle + \varphi(h) \ .$$

Sei nun $\alpha > 0$ so klein, dass

$$\langle h, \operatorname{Hess} f(x), h\rangle \geq \alpha \|h\|^2$$

ist, für alle $h \in \mathbb{R}^n$ (vgl. Lemma 5.8.), denn $\operatorname{Hess} f(x)$ ist positiv definit. Wähle nun (zu $\varepsilon = \frac{\alpha}{4}$) $0 < \delta' < \delta$ so klein, dass für alle $h \in B_{\delta'}(0) \backslash \{0\}$ gilt:

$$\frac{|\varphi(h)|}{\|h\|^2} < \frac{\alpha}{4} \ ,$$

also $|\varphi(h)| \leq \frac{\alpha}{4} \|h\|^2 \ \ \forall h \in B_{\delta'}(0) \backslash \{0\}_n$. Dann ist für alle $h \in B_{\delta'}(0)$.

$$\begin{aligned} f(x + h) &= f(x) + \underbrace{\langle \operatorname{grad} f(x), h\rangle}_{=0} + \frac{1}{2} \langle h, \operatorname{Hess} f(x), h\rangle + \varphi(h) \\ &\geq f(x) + \frac{1}{2}\alpha \|h\|^2 - \frac{\alpha}{4} \|h\|^2 \\ &= f(x) + \frac{\alpha}{4} \|h\|^2 \\ &> f(x) \ . \end{aligned}$$

Also hat f ein (sogar striktes) lokales Minimum.

(b) Folgt aus obigem, wenn man von f zu $-f$ übergeht.

(c) Ist $h \in \mathbb{R}^n \backslash \{0\}$ so, dass $\alpha := \langle h, \operatorname{Hess} f(x), h\rangle > 0$ gilt und **o.B.d.A.** $\|h\| = 1$, so gilt für alle $t > 0$ mit t klein genug, dass

$$\begin{aligned} f(x+th) &= f(x) + \underbrace{\langle \operatorname{grad} f(x), th\rangle}_{=0} + \frac{1}{2}\langle th, \operatorname{Hess} f(x), th\rangle + \varphi(th) \\ &\geq f(x) + \frac{\alpha}{2}t^2 + \varphi(th) \ . \end{aligned}$$

Für $|t|$ klein genug, ist aber

$$|\varphi(th)| \leq \frac{\alpha}{4}\|th\|^2 = \frac{\alpha}{4}t^2\underbrace{\|h\|^2}_{=1} = \frac{\alpha}{4}t^2 \ .$$

Also ist

$$\begin{aligned} f(x+th) &\geq f(x) + \frac{\alpha}{2}t^2 - \frac{\alpha}{4}t^2 \\ &= f(x) + \frac{\alpha}{4}t^2 \\ &> f(x) \end{aligned}$$

für $0 < t < \delta$ für δ klein genug. Also gibt es in jeder Umgebung U von x einen Punkt $y \in U$ ($y = x + th$), so dass $f(y) > f(x)$ ist. Ähnlich sieht man: Es gibt auch $\widetilde{y} = x + t\widetilde{h}$ (mit t klein und $\widetilde{h} \in \mathbb{R}^n$ so, dass $\left\|\widetilde{h}\right\| = 1$ und $\left\langle \widetilde{h}, \operatorname{Hess} f(x), \widetilde{h}\right\rangle < 0$ ist), so dass $f(\widetilde{y}) < f(x)$ ist. Also hat f in x kein lokales Extremum.

QED.

5.3.5 Kommentar

(1) Ist $\operatorname{Hess} f(x)$ nur positiv semidefinit, d.h. $\langle h, \operatorname{Hess} f(x), h\rangle \geq 0$, für alle $h \in \mathbb{R}^n$ (und $\operatorname{grad} f(x) = 0$), so kann man noch nicht entscheiden, ob in x ein lokales Minimum vorliegt.

(i) Für $f : \mathbb{R}^2 \to \mathbb{R}$, $(x,y) \mapsto x^2 + y^4$: lokales Minimum

(ii) Für $f : \mathbb{R}^2 \to \mathbb{R}$, $(x,y) \mapsto x^2 + y^3$: kein lokales Extremum

(iii) Für $f : \mathbb{R}^2 \to \mathbb{R}$, $(x,y) \mapsto x^2$: lokales Maximum, aber nicht strikt

(2) Aus Satz 5.7. folgt aber folgendes notwendige Kriterium: Hat f in x ein lokales Extremum (Minimum), so ist $\operatorname{Hess} f(x)$ wenigstens (positiv) semidefinit.

Kapitel 6

Implizite Funktionen

16.12.2004

Motivation

Eine Funktion $g : \mathbb{R} \to \mathbb{R}$ ist nicht immer *explizit* gegeben (also $y = g(x)$), sondern häufig nur *implizit* durch eine Gleichung $F(x, g(x)) = 0$. Hierbei ist $F : \mathbb{R}^2 \to \mathbb{R}$ eine gegebene Funktion. Man möchte dann die implizite Gleichung $F(x, y) = 0$ explizit machen: $y = g(x)$, d.h. „nach y auflösen“.

Beispiel

Betrachte die Funktion $F : \mathbb{R}^2 \to \mathbb{R}$, $(x, y) \mapsto x^2+y^2-1$. Dann ist $C = \{(x, y) \in \mathbb{R}^2 : F(x, y) = 0\}$ die Einheitskreislinie in $\mathbb{R}^2$. C ist nun leider nicht der Graph einer Funktion $g : \mathbb{R} \to \mathbb{R}$, denn

- für $x \in \mathbb{R}$ mit $|x| > 1$ gibt es überhaupt keine $y \in \mathbb{R}$ mit $(x, y) \in C$,
- für $x \in \mathbb{R}$ mit $|x| < 1$ gibt es gleich zwei $y \in \mathbb{R}$ mit $(x, y) \in C$ (nämlich $\pm\sqrt{1 - x^2}$).

Ist allerdings $(a, b) \in C$, so kann man hier $F(x, y) = 0$ *lokal um* (a, b) nach x oder y auflösen, d.h. es gibt eine Umgebung $U_1 \subseteq \mathbb{R}$ von a und eine Umgebung $U_2 \subseteq \mathbb{R}$ von b und eine Funktion $g : U_1 \to U_2$, so dass für alle $(x, y) \in U_1 \times U_2$ gilt:

$$F(x, y) = 0 \quad \Leftrightarrow \quad y = g(x) \ .$$

Tatsächlich: Ist etwa $b > 0$, so wähle $\varepsilon > 0$ klein genug, $U_1 = (a - \varepsilon, a + \varepsilon)$, $U_2 = (b - \varepsilon, b + \varepsilon)$ und $g(x) = \sqrt{1 - x^2}$.

Beachte aber: Für $b = 0$ kann man in keiner Umgebung von $(\pm1, 0)$ die Gleichung $F(x, y) = 0$ nach y auflösen, wohl aber nach x (durch $x = \pm\sqrt{1 - y^2}$).

Beachte schließlich: Die Punkte $(\pm1, 0) \in C$ sind genau die Punkte, wo die partielle Ableitung von F nach y verschwindet, $\frac{\partial F}{\partial y} = 2y$, (d.h. die Tangente an C ist vertikal).

6.1 Satz über die impliziten Funktionen

6.1.1 Kommentar

Der Satz über implizite Funktionen gibt nun gerade eine hinreichende Bedingung dafür an, dass man eine implizite Gleichung $F(x, y) = 0$ lokal um einen Punkt (a, b) mit $F(a, b) = 0$ explizit machen kann: $y = g(x)$. Wir formulieren ihn nun gleich für die Situation, wo auch a und b mehrere Komponenten haben kann.

Theorem 6.1. Satz über implizite Funktionen
Sei $G \subseteq \mathbb{R}^{n+m} = \mathbb{R}^n \times \mathbb{R}^m$ ein Gebiet und $F : G \to \mathbb{R}^m$ eine stetig differenzierbare Abbildung. Sei $(a,b) \in G$ so, dass $F(a,b) = 0$ ist und

$$\frac{\partial F}{\partial y}(a,b) = \begin{pmatrix} \frac{\partial F_1}{\partial y_1} & \cdots & \frac{\partial F_1}{\partial y_m} \\ \vdots & \ddots & \vdots \\ \frac{\partial F_m}{\partial y_1} & \cdots & \frac{\partial F_m}{\partial y_m} \end{pmatrix} (a,b)$$

regulär ist (d.h. $\frac{\partial F}{\partial y}(a,b) \in Gl_m(\mathbb{R})$, also det $\left(\frac{\partial F}{\partial y}(a,b)\right) \neq 0$). Dann existieren offene Umgebungen $U_1 \subseteq \mathbb{R}^n$ von a und $U_2 \subseteq \mathbb{R}^m$ von b, so dass $U_1 \times U_2 \subseteq G$ ist und eine stetig differenzierbare Funktion $g : U_1 \to U_2$, so dass für alle $(x,y) \in U_1 \times U_2$ gilt:

$$F(x,y) = 0 \quad \Leftrightarrow \quad y = g(x).$$

21.12.2004

6.1.2 Beweisstrategie

Wie werden nun zunächst (für U_1, U_2 klein genug) mit Hilfe des Banachschen Fixpunktsatzes (Satz 2.4.) eine eindeutig bestimmte Funktion $g : U_1 \to U_2$ mit $F\big(x, g(x)\big) = 0$ konstruieren. In einem zweiten Schritt wird (wieder mit dem Banachschen Fixpunktsatz) gezeigt, dass g stetig sein muss. In einem dritten Schritt schließlich wird gezeigt, dass g sogar stetig differenzierbar ist. Wir ziehen den dritten Schritt vor:

Lemma 6.2.
Seien $U_1 \subseteq \mathbb{R}^n$ und $U_2 \subseteq \mathbb{R}^m$ offen und $F : U_1 \times U_2 \to \mathbb{R}^m$ eine stetig differenzierbare Abbildung. Sei $(a,b) \in U_1 \times U_2$ mit $F(a,b) = 0$ und $\frac{\partial F}{\partial y}(a,b) \in \mathrm{Gl}_m(\mathbb{R})$ und sei $g : U_1 \to U_2$ eine stetige Funktion mit $g(a) = b$, so dass für alle $x \in U_1$ gilt: $F\big(x, g(x)\big) = 0$. Dann ist g im Punkt a sogar (total) differenzierbar und es gilt:

$$\underbrace{\frac{\partial g}{\partial x}(a)}_{\text{ist } m \times n \text{ Matrix}} = - \underbrace{\left(\frac{\partial F}{\partial y}(a,b)\right)^{-1}}_{\text{ist } m \times m \text{ Matrix}} \cdot \underbrace{\frac{\partial F}{\partial x}(a,b)}_{\text{ist } m \times n \text{ Matrix}} . \tag{6.1}$$

6.1.3 Kommentar

(1) Wenn wir schon wüssten, dass g differenzierbar ist, so würde die Formel aus Gleichung (6.1) für die Ableitung unmittelbar aus der Kettenregel folgen:

$$\begin{aligned} & F\big(x, g(x)\big) = 0 \\ \Rightarrow\quad & 0 = D_x\big(F(x, g(x))\big) = D_x F\big(x, g(x)\big) \cdot \mathrm{id} + D_y F\big(x, g(x)\big) \circ D_x g(x) \\ \Rightarrow\quad & D_x g(a) = -\left(D_y F(a,b)\right)^{-1} \cdot D_x F(a,b) \; . \end{aligned}$$

(2) Wendet man das Lemma auch für die Punkte $\big(x, g(x)\big)$ in einer Umgebung von (a,b) an, so sieht man an Gleichung (6.1) und der Stetigkeit von $(x,y) \to DF(x,y) = (D_x F, D_y F)$, dass g sogar stetig differenzierbar auf ganz U_1 ist, wenn man gegebenfalls U_1 noch etwas verkleinert, so dass die Bedingung $D_y F\big(x, g(x)\big) \in \mathrm{Gl}_m(\mathbb{R})$, für alle $x \in U_1$ erfüllt ist. Ist also Lemma 6.2. bewiesen, so reicht es im Beweis des impliziten Funktionensatzes lediglich eine stetige Funktion $g : U_1 \to U_2$ nachzuweisen.

¡? Beweis ¿! von Lemma 6.2.

Sei **o.B.d.A.** $(a,b) = (0,0)$. Wir setzen $A := \frac{\partial F}{\partial x}(0,0) \in \mathrm{Mat}\,(m \times n; \mathbb{R})$ und $B := \frac{\partial F}{\partial y}(0,0) \in \mathrm{Gl}_m(\mathbb{R})$. Wegen der Differenzierbarkeit von F in $(0,0)$ gibt es eine Funktion $\varphi : U_1 \times U_2 \to \mathbb{R}$, so dass

$$\lim_{(x,y)\to(0,0)} \frac{\varphi(x,y)}{\|(x,y)\|} = 0$$

mit

$$F(x,y) = \underbrace{F(0,0)}_{=0} + \underbrace{\left\langle \left(\frac{\partial F}{\partial x}, \frac{\partial F}{\partial y} \right)(0,0), (x,y) \right\rangle}_{=Ax+By} + \varphi(x,y) \ .$$

Wegen $F\big(x, g(x)\big) = 0$ erhält man daraus die Darstellung

$$0 = F\big(x, g(x)\big) = Ax + Bg(x) + \varphi\big(x, g(x)\big) \ ,$$

also wegen $B \in \mathrm{Gl}_m(\mathbb{R})$:

$$g(x) = -B^{-1}Ax - B^{-1}\varphi\big(x, g(x)\big) \ . \qquad (6.2)$$

(Wir wissen schon, dass $-B^{-1}A$ der Kandidat für das Differential von g in $a = 0$ ist, daher behaupten wir:)

Behauptung: $\lim\limits_{x\to 0} \frac{\Psi(x)}{\|x\|} = 0$ für $\Psi(x) := -B^{-1}\varphi\big(x, g(x)\big)$.

Schritt 1: Es wird zunächst eine „verbesserte Stetigkeit“ von g in $a = 0$ gezeigt ($\hookrightarrow$ Lipschitz-Stetigkeit): Behauptung: Es existieren Konstanten $\delta > 0$ und $C > 0$, so dass für alle $x \in B_\delta(0)$ gilt:

$$\|g(x)\| \leq C \cdot \|x\| \ .$$

Dazu: Wegen $\frac{\varphi(x,y)}{\|(x,y)\|} \longrightarrow 0$ für $(x,y) \longrightarrow 0$, existien (zu $\varepsilon \leq \frac{1}{2\|B^{-1}\|}$) $\delta_1, \delta_2 > 0$, so dass für $\|x\| < \delta_1$, $\|y\| < \delta_2$ gilt:

$$\begin{aligned} \|\varphi(x,y)\| &\leq \frac{1}{2\|B^{-1}\|}\|(x,y)\| \leq \frac{1}{2\|B^{-1}\|}\|(x,0)+(0,y)\| && (6.3) \\ &\leq \frac{1}{2\|B^{-1}\|}\left(\|(x,0)\| + \|(0,y)\|\right) \leq \frac{1}{2\|B^{-1}\|}\left(\|x\| + \|y\|\right) \ . && (6.4) \end{aligned}$$

Es folgt

$$\|\Psi(x)\| = \left\|B^{-1}\varphi\big(x, g(x)\big)\right\| \leq \left\|B^{-1}\right\| \cdot \left\|\varphi\big(x, g(x)\big)\right\| \leq \frac{1}{2}\left(\|x\| + \|g(x)\|\right) \ .$$

(Hierbei muss man $\delta_1 < 0$ eventuell nochmal verkleinern, so dass $\|g(x)\| < \delta_2$ ist, für $\|x\| < \delta_1$, was man wegen der Stetigkeit von g in $x = 0$ machen kann.) Insgesamt bekommt man nun aus Gleichung (6.2):

$$\|g(x)\| \leq \left\|B^{-1}A\right\| \cdot \|x\| + \frac{1}{2}\|x\| + \frac{1}{2}\|g(x)\| \ .$$

(Nun kann man den Term $g(x)$ auf der rechten Seite „links verschlucken“ und erhält)

$$\frac{1}{2}\|g(x)\| \leq \left(\left\|B^{-1}A\right\| + \frac{1}{2} \right) \|x\| \ ,$$

also

$$\|g(x)\| \leq C \cdot \|x\| \ ,$$

für alle $x \in B_\delta(0)$ (mit $\delta := \delta_1$) und $C := 2\left\|B^{-1}A\right\| + 1$.

Schritt 2 : Zu $\varepsilon > 0$ ist nun

$$\begin{aligned} \|\Psi(x)\| &= \left\|-B^{-1}\varphi\big(x, g(x)\big)\right\| \\ &\leq \left\|B^{-1}\right\| \cdot \left(\varepsilon \left\|\big(x, g(x)\big)\right\|\right) \\ &\leq \left\|B^{-1}\right\| \cdot \varepsilon \left(\|x\| + \|g(x)\|\right) \\ &\overset{\text{Schritt 1}}{\leq} \left\|B^{-1}\right\| \cdot \varepsilon \left(\|x\| + C\|x\|\right) \\ &\leq \left[\left\|B^{-1}\right\| (1+C)\varepsilon\right] \|x\| \ , \end{aligned}$$

wenn $\|x\|$ klein genug ist. Das zeigt (da $\varepsilon > 0$ beliebig), dass

$$\lim_{x\to 0} \frac{\Psi(x)}{\|x\|} = 0$$

ist (und damit g in $a = 0$ differenzierbar mit $Dg(0) = -B^{-1}A$.)

QED.

¡? Beweis ¿! des impliziten Funktionensatzes (Theorem 6.1.)

Sei **o.B.d.A.** $(a, b) = (0, 0)$.

Schritt 1: Wir zeigen, dass es abgeschlossene Kugeln $\overline{B_1} = \{x \in \mathbb{R}^n \,|\, \|x\| \leq r_1\}$ und $\overline{B_2} = \{y \in \mathbb{R}^m \,|\, \|y\| \leq r_2\}$ gibt mit $\overline{B_1} \times \overline{B_2} \subseteq G$ derart, dass es genau eine Funktion $g : \overline{B_1} \to \overline{B_2}$ gibt, so dass für alle $(x, y) \in \overline{B_1} \times \overline{B_2}$ gilt:

$$F(x, y) = 0 \;\Leftrightarrow\; y = g(x) \; .$$

(Die Stetigkeit von g wird später gezeigt!)

Betrachte dazu die Funktion $\Phi : G \to \mathbb{R}^m$ gegeben durch

$$\Phi(x, y) = y - B^{-1}F(x, y)$$

(wobei wieder $B = \frac{\partial F}{\partial y}(0, 0)$ abgekürzt wird). Es ist dann

$$\begin{aligned} F(x, y) = 0 \quad &\Leftrightarrow \quad B^{-1}F(x, y) = 0 \\ &\Leftrightarrow \quad y - B^{-1}F(x, y) = y \\ &\Leftrightarrow \quad \Phi(x, y) = y \; , \end{aligned}$$

also y ein Fixpunkt der Abbildung $y \mapsto \Phi(x, y)$ (für gegebenes x) ist. Weil $F(0, 0) = 0$ ist und F stetig, kann man für jedes $r_2 > 0$ zunächst $r_1 > 0$ so klein wählen, dass $\|\Phi(x, 0)\| = \left\|B^{-1}F(x, 0)\right\| \leq \frac{1}{2} r_2$ ist, wenn $\|x\| \leq r_1$ ist.

Nach dem Schrankensatz (Korollar 4.13.) ist nun für $y, y' \in \overline{B_2}$ und jedem $x \in \overline{B_1}$:

$$\|\Phi(x, y) - \Phi(x, y')\| \leq \sup\left\{\|D_y\Phi(x, z)\| \,\middle|\, z \in [y, y']\right\} \cdot \|y - y'\| \; .$$

Weil aber

$$D_y\Phi(0, 0) = \text{ id } - B^{-1} \cdot D_yF(0, 0) = \text{ id } - B^{-1}B = 0$$

ist und Φ stetig differenzierbar, kann man $r_1, r_2 > 0$ so klein wählen, dass $\|D_y\Phi(x, y)\| \leq \frac{1}{2}$ ist, für alle $(x, y) \in \overline{B_1} \times \overline{B_2}$. Deshalb ist dann

$$\|\Phi(x, y) - \Phi(x, y')\| \leq \frac{1}{2} \|y - y'\| \tag{6.5}$$

und damit

$$\|\Phi(x, y)\| \leq \|\Phi(x, y) - \Phi(x, 0)\| + \|\Phi(x, 0)\| \leq \frac{1}{2} \underbrace{\|y\|}_{\leq r_2} + \underbrace{\left\|B^{-1}F(x, 0)\right\|}_{\leq \frac{1}{2} r_2} \leq r_2 \tag{6.6}$$

für alle $(x, y) \in \overline{B_1} \times \overline{B_2}$. Damit ist für jedes $x \in \overline{B_1}$ die Abbildung $\Phi(x, -) : \overline{B_2} \to \overline{B_2}$ wegen Gleichung (6.5) eine Kontraktion und weil $\overline{B_2} \subseteq \mathbb{R}^m$ abgeschlossen (und $\mathbb{R}^m$ vollständig) ist, hat entsprechend des Banachschen Fixpunktsatzes (Satz 2.4.) $\Phi(x, -) : \overline{B_2} \to \overline{B_2}$ genau einen Fixpunkt y; wir nennen ihn $g(x)$. Also gibt es zu jedem $x \in \overline{B_1}$ genau ein $y = g(x) \in \overline{B_2}$, so dass $F(x, y) = 0$ ist (insbesondere ist $g(0) = 0$).

23.12.2004

Schritt 2: Es bleibt zu zeigen, dass die Funktion $g : \overline{B_1} \to \overline{B_2}$ stetig ist. Betrachte dazu den Vektorraum $\mathcal{C}(\overline{B_1}, \mathbb{R}^m)$ mit der Supremumsnorm

$$\|h\| = \sup_{x \in \overline{B_1}} \|h(x)\| \ .$$

Wie im Beispiel 2.2.1 sieht man, dass $\mathcal{C}(\overline{B_1}, \mathbb{R}^m)$ vollständig ist, weil gleichmäßige Limiten von stetigen Funktionen wieder stetig sind. In $\mathcal{C}(\overline{B_2}, \mathbb{R}^m)$ betrachten wir die abgeschlossene Teilmenge

$$A := \left\{ h \in \mathcal{C}(\overline{B_1}, \mathbb{R}^m) \middle| \mathrm{Bild}\,(h) \subseteq \overline{B_2}, h(0) = 0 \right\}$$

und auf A die Abbildung $\Psi : A \to A$:

$$\Psi(h)(x) = \Phi\big(x, h(x)\big) \ .$$

Es ist dann

$$\begin{aligned} \Psi(h) = h \quad &\Leftrightarrow \quad \Phi\big(x, h(x)\big) = h(x) \quad \forall \ x \in \overline{B_1} \\ &\Leftrightarrow \quad 0 = F\big(x, h(x)\big) \quad \forall \ x \in \overline{B_1} \\ &\Leftrightarrow \quad h(x) = g(x) \ \ \forall \ x \in \overline{B_1} \ \ \text{nach Schritt 1} \\ &\Leftrightarrow \quad h = g \ . \end{aligned}$$

Aber Ψ ist eine Kontraktion, denn

$$\begin{aligned} \|\Psi(h) - \Psi(h')\| &= \sup_{x \in \overline{B_1}} \left\| \big(\Psi(h)(x) - \Psi(h')(x)\big) \right\| \\ &= \sup_{x \in \overline{B_1}} \left\| \Phi\big(x, h(x)\big) - \Phi\big(x, h'(x)\big) \right\| \\ &\overset{\text{Gleichung(6.5)}}{\leq} \sup_{x \in \overline{B_1}} \left\| \frac{1}{2} \left(h(x) - h'(x)\right) \right\| \\ &= \frac{1}{2} \sup_{x \in \overline{B_1}} \|(h(x) - h'(x)\| \\ &= \frac{1}{2} \|h - h'\| \ . \end{aligned}$$

(Es ist wegen Gleichung (6.6) auch $\Psi(A) \subseteq A$.)

Also hat Ψ nach dem Banachschen Fixpunktsatz genau einen Fixpunkt $h \in A$ (also h stetig) und dieser muss g sein. Also ist g stetig.

Schritt 3: Wegen Lemma 6.2. und seines anschließenden Kommentars ist g sogar stetig differenzierbar und damit das Theorem 6.1. (mit $U_1 := B_1$ und $U_2 := B_2$, r_1 evtl. noch einmal verkleinern und $g|B_1 : B_1 \to B_2$) bewiesen.

QED.

6.1.4 Beispiel

Sei $G \subseteq \mathbb{R}^n$ ein Gebiet mit $f : G \to \mathbb{R}$ stetig differenzierbar. Für $c \in \mathbb{R}$ betrachten wir die *Niveauflächen*

$$N_c = \{x \in G | f(x) = c\} \ .$$

Ist nun $\mathrm{grad}\, f(x) \neq 0$ für alle $x \in N_c$, so behaupten wir, dass N_c überall lokal aussieht, wie der Graph einer stetig differenzierbaren Funktion g in $(n-1)$ Veränderlichen. Denn ist z.B. $\frac{\partial f}{\partial x_n}(a) \neq 0, \ \ a \in N_c$, so gilt mit $a' = (a_1, \ldots, a_{n-1})$: Es gibt Umgebungen $U' \subseteq \mathbb{R}^{n-1}$ von a' und $I \subseteq \mathbb{R}$ von a_n, so dass: $U' \times I \subseteq G$ und es existiert ein stetig differenzierbares $g : U' \to I \subseteq \mathbb{R}$, so dass

$$f(x', x_n) - c = 0 \quad \Leftrightarrow \quad x_n = g(x') \ ,$$

für alle $(x', x_n) \in U \times I$. Für alle $x \in U \times I$ ist also

$$x \in N_c \quad \Leftrightarrow \quad x \in \operatorname{Graph}(g) = \{(x', g(x') | x' \in U'\} \ .$$

Z.B. ist für $f(x) = \|x\|^2$: $N_c = \{x \in \mathbb{R}^n | \, \|x\|^2 = c\}$ (mit $c > 0$).
Wegen $\operatorname{grad} f(x) = 2x \neq 0$, $\forall x \in N_c$, ist damit N_c lokalGraph einer Funktion in $(n-1)$ Veränderlichen.

6.2 Umkehrabbildungen

Motivation

Seien $G, D \subseteq \mathbb{R}^n$ Gebiete und $f : G \to D$ stetig differenzierbar. Häufig möchte man wissen, ob f eine Koordinatentransformation ist, d.h. ob f bijektiv ist (also $g = f^{-1} : D \to G$ existiert) und ob $g = f^{-1} : D \to G$ selbst wieder stetig differenzierbar ist. (Ist dann $y = f(x)$, so ist f ein *Koordinatenwechsel von x nach y.*) Man definiert deshalb:

Definition 6.3.
Seien $G, D \subseteq \mathbb{R}^n$ Gebiete. Eine stetig differenzierbare Abbildung $f : G \to D$ heißt ein *Diffeo(morphismus)*, genau dann wenn f bijektiv und $f^{-1} : D \to G$ stetig differenzierbar ist.

6.2.1 Bemerkungen

Seien $G, D \subseteq \mathbb{R}^n$ Gebiete. Ist $f : G \to D$ ein Diffeomorphismus, so muss für jedes $x \in G$ die Ableitung $Df(x) : \mathbb{R}^n \to \mathbb{R}^n$ ein Isomorphismus (also invertierbar, $J_f(x) \in Gl_n(\mathbb{R})$) sein.

¡? Beweis ¿!

Ist g Inverses zu f, also insbesondere $g \circ f(x) = x$, für alle $x \in G$, so ist nach der Kettenregel

$$Dg(f(x)) \circ Df(x) = D(\operatorname{id})(x) = \operatorname{id}$$

bzw. in Matrizenschreibweise

$$J_g\big(f(x)\big) \cdot J_f(x) = \mathbb{1} \ .$$

$J_g(f(x))$ ist damit Inverses zu $J_f(x)$. (Es muss also $\det J_f(x) \neq 0$ sein!)
QED.

6.2.2 Kommentar

(1) Es kann durchaus vorkommen, dass $f : G \to D$ stetig differenzierbar ist und bijektiv ohne dass die Umkehrung $g = f^{-1} : D \to G$ differenzierbar ist. Ist $f : \mathbb{R} \to \mathbb{R}$, $f(x) = x^3$ so ist f sicher stetig differenzierbar und bijektiv, ihre Umkehrung $g : \mathbb{R} \to \mathbb{R}$, $y \mapsto \sqrt[3]{y}$ ist aber (im Nullpunkt) nicht differenzierbar. Beachte: $f'(0) = 0$ liegt nicht in $Gl_1(\mathbb{R}) = \mathbb{R}^* = \mathbb{R}\backslash\{0\}$.

(2) In Dimension $n = 1$ ist die Bedingung „$f'(x) \neq 0$“ für eine Funktion $f : I \to \mathbb{R}$ (I offen) sogar hinreichend dafür, dass f ein Diffeomorphismus auf ein offenes Intervall $J \subseteq \mathbb{R}$ ist. (Nach 6.2.1 ist diese Bedingung auch notwendig.)

Da f' stetig ist, folgt nämlich mit dem Zwischenwertsatz aus $f'(x) \neq 0 \ \forall x \in I$, dass $f'(x) > 0 \ \ \forall x \in I$ <u>oder</u> $f'(x) < 0 \ \ \forall x \in I$, sagen wir $f' > 0$. Dann ist f monoton wachsend und damit existieren die Grenzwerte (mit $I = (a, b)$, $a \in [-\infty, \infty)$, $b \in (-\infty, \infty]$)

$$\begin{aligned} \lim_{x \to a} f(x) &= c \in [-\infty, \infty) \ , \\ \lim_{x \to b} f(x) &= d \in (-\infty, \infty] \ . \end{aligned}$$

Es muss dann für $J = (c, d)$ gelten: $f(I) = J$ (erneut Zwischenwertsatz), d.h. $f : I \to J$ surjektiv. f ist auch injektiv (Satz von Rolle). Also existiert $g : J \to I$, $g = f^{-1}$ und ist stetig differenzierbar (vgl. Mathematik für Physiker I) mit

$$g'(y) = \frac{1}{f'(g(y))} \ .$$

Daher ist $f : I \to J$ ein Diffeomorphismus.

Wir verallgemeinern nun diesen Sachverhalt auf Funktionen in mehreren Veränderlichen (allerdings nur in einer lokalen Version).

Satz 6.4. Satz über die Umkehrabbildung
Sei $G \subseteq \mathbb{R}^n$ ein Gebiet und $f : G \to \mathbb{R}^n$ eine stetig differenzierbare Abbildung. Ist nun $a \in G$ derart, dass $Df(a)$ invertierbar ist, so existieren Umgebungen $U \subseteq G$ von a und $V \subseteq \mathbb{R}^n$ von $b := f(a)$, so dass $f(U) = V$ und $f|_U : U \to V$ ein Diffeomorphismus ist.

6.2.3 Kommentar

Ist $f : G \to \mathbb{R}^n$ stetig differenzierbar, $Df(a)$ invertierbar, $U, V \subseteq \mathbb{R}^n$ und $f : U \to V$ wie in obigem Satz, so gilt für $g : V \to U$ in $b = f(a)$:

$$Dg(b) = (Df(a))^{-1} \ .$$

Denn weil $g \circ f = \mathrm{id}$ ist, folgt nach der Kettenregel

$$Dg(b)Df(a) = \mathrm{id} \ ,$$

also

$$Dg(b) = Df(a)^{-1} \ .$$

11.01.2005

¡? Beweis ¿!

Man möchte die Gleichung $y = f(x)$ um $x_0 = a$ und $y_0 = b$ nach x auflösen, $x = g(y)$. Deshalb betrachtet man folgende Abbildung $F : G \times \mathbb{R}^n \to \mathbb{R}^n$,

$$F(x, y) = y - f(x) \ .$$

Es ist

$$\frac{\partial F}{\partial x}(a) = -Df(a) \ ,$$

also in $Gl_n(\mathbb{R})$. Deshalb liefert der implizite Funktionensatz Umgebungen $U \subseteq G$ von a, $V \subseteq \mathbb{R}^n$ von b und ein stetig differenzierbares $g : V \to U$, so dass gilt:

$$F(x, y) = 0 \quad \Leftrightarrow \quad x = g(y) \ ,$$

für alle $(x, y) \in U \times V$ und damit gilt für alle $(x, y) \in U \times V$:

$$f(x) = y \quad \Leftrightarrow F(x, y) = 0 \quad \Leftrightarrow \quad x = g(y) \ .$$

Das zeigt, dass $f|_U : U \to V$ bijektiv ist und g stetig differenzierbar, also ist $f|_U$ ein Diffeomorphismus.
QED.

6.2.4 Beispiel

Räumliche Polarkoordinaten sind gegeben durch $f : \mathbb{R}_+ \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}^3$

$$f(r, \theta, \varphi) = (r \sin\theta \cos\varphi, r \sin\theta \sin\varphi, r \cos\theta) \ .$$

f ist (z.B.) injektiv auf

$$G = \mathbb{R}_+ \times (0, \pi) \times (-\pi, \pi)$$

und für ihre *Funktionaldetermiante (Jacobideterminante)* $\det J_f : G \to \mathbb{R}$ gilt:

$$\det(J_f(r, \theta, \varphi)) = r^2 \cdot \sin\theta \neq 0 \ ,$$

für alle $(r, \vartheta, \varphi) \in G$. Deshalb ist f mit dem Umkehrsatz (zunächst nur) lokaler Diffeomorphismus, aber weil f injektiv ist, auch ein (globaler) Diffeomorphismus auf sein Bild $D := f(G)$.

6.3 Nebenbedingungen

Motivation

Sei $G \subseteq \mathbb{R}^n$ ein Gebiet und $f : G \to \mathbb{R}$ stetig differenzierbar. Haufig ist es so, dass man an lokalen Extrema $a \in G$ von f „unter einer Nebenbedingung“ $h = 0$ interessiert ist, z.B. in der Theoretischen Mechanik, wenn Zwangskräfte durch die Nebenbedingung $h = 0$ gegeben sind. Ist $\operatorname{grad} h(a) \neq 0$ an so einer Stelle a mit $h(a) = 0$, so gibt einem der implizite Funktionensatz die Möglichkeit eine notwendige Bedingung zu bekommen, indem man das Problem zurückführt auf ein solches ohne Nebenbedingungen.

Definition 6.5.
Sei $G \subseteq \mathbb{R}^n$ ein Gebiet und seien $f, h : G \to \mathbb{R}$ stetig differenzierbar. Sei

$$M = \{x \in G | h(x) = 0\}$$

und $a \in M$.

(a) f hat ein *lokales Maximum (bzw. Minimum) unter der Nebenbedingung* $h = 0$, wenn es eine offene Umgebung $U \subseteq G$ von a gibt, so dass für alle $x \in M \cap U$ gilt:

$$f(x) \leq f(a) \quad (\text{bzw. } f(x) \geq f(a)\ ) \ .$$

(b) f hat ein *lokales Extremum unter der Nebenbedingung* $h = 0$, wenn f ein lokales Maximum oder lokales Minimum unter der Nebenbedingung hat.

Satz 6.6.
Seien $f, h : G \to \mathbb{R}$ stetig differenzierbar, $G \subseteq \mathbb{R}^n$ ein Gebiet und $a \in M = \{x \in G | h(x) = 0\}$. Es habe f in $a \in M$ ein lokales Extremum unter der Nebenbedingung $h = 0$ und es sei $\operatorname{grad} h(a) \neq 0$. Dann gibt es ein $\lambda \in \mathbb{R}$, so dass gilt:

$$\operatorname{grad} f(a) = \lambda \operatorname{grad} h(a) \ .$$

6.3.1 Kommentar

(1) Der Gradient von h (in a) steht senkrecht auf den Niveauflächen von h ✎ *Übung* ✐. Damit f

ein lokales Extremum in $a \in M$ unter der Nebenbedingung $h = 0$ hat, muss $\operatorname{grad} f(a)$ nicht notwendig verschwinden, sondern nur senkrecht auf $M = \{x|h(x) = 0\}$ stehen.

(2) Man kann Lemma 6.2. auch so ausdrücken, dass es unter den beschriebenen Vorraussetzungen ein $\lambda \in \mathbb{R}$ gibt, so dass für die Funktion $g := f + \lambda h : G \to \mathbb{R}$ gilt

$$\operatorname{grad}(g)(a) = 0 \ .$$

λ wird dann als ein *Lagrange-Multiplikator für das Problem* bezeichnet.

¡? Beweis ¿!

Da $\operatorname{grad} h(a) \neq 0$ ist, ist wenigstens eine partielle Ableitung von h in a von Null verschieden, sagen wir $\frac{\partial h}{\partial x_n}(a)$. Sei $a' = (a_1, \ldots, a_{n-1})$. Der implizite Funktionensatz liefert dann:

- Umgebungen $V \subseteq \mathbb{R}^{n-1}$ von a', $I \subseteq \mathbb{R}$ von a_n, so dass $V \times I \subseteq G$ ist,
- ein stetig differenzierbares $g : V \to I$, so dass für alle $(x', x_n) \in V \times I$ gilt:

$$h(x', x_n) = 0 \quad \Leftrightarrow \quad x_n = g(x') \ .$$

Hat nun $f : G \to \mathbb{R}$ ein lokales Extremum unter der Nebenbedingung $h = 0$, so hat $f \circ \varphi : V \to \mathbb{R}$ ein lokales Extremum in a' (ohne Nebenbedingung!). Hierbei ist $\varphi : V \to M \subseteq \mathbb{R}^n$,

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

Es ist also nach 5.2.1

$$\operatorname{grad}(f \circ \varphi)(a') = 0 \ .$$

Für $i = 1, \ldots, n-1$ rechnen wir nun (mit der Kettenregel):

$$0 = \frac{\partial (f \circ \varphi)}{\partial x'_i}(a') = \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}\big(\underbrace{\varphi(a')}_{=a}\big) \qquad \underbrace{\frac{\partial \varphi_j}{\partial x'_i}(a')}_{substack=\delta_{ij} \text{für } 1 \leq j \leq n-1} \qquad = \frac{\partial f}{\partial x_i}(a) \cdot 1 + \frac{\partial f}{\partial x_n}(a) \frac{\partial g}{\partial x'_i}(a') \ . \tag{6.7}$$

Andererseits ist wegen $h \circ \varphi(x') = 0 \ \forall x' \in V$:

$$0 = \frac{\partial (h \circ \varphi)}{\partial x'_i}(a') = \frac{\partial h}{\partial x_i}(a) + \frac{\partial h}{\partial x_n}(a) \frac{\partial g}{\partial x'_i}(a') \ ,$$

also wegen $\frac{\partial h}{\partial x_n}(a) \neq 0$:

$$\frac{\partial g}{\partial x'_i}(a') = -\frac{\frac{\partial h}{\partial x_i}(a)}{\frac{\partial h}{\partial x_n}(a)} \ . \tag{6.8}$$

Setzen wir nun

$$\lambda := \frac{\frac{\partial f}{\partial x_n}(a)}{\frac{\partial h}{\partial x_n}(a)} \in \mathbb{R} \ ,$$

so folgt natürlich für $i = n$ zunächst:

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

aber dann auch für $i = 1, \ldots, n-1$, wenn man Gleichung (6.8) in Gleichung (6.7) einsetzt:

$$\frac{\partial f}{\partial x_i}(a) = -\frac{\partial f}{\partial x_n}(a) \cdot \left(-\frac{\frac{\partial h}{\partial x_i}(a)}{\frac{\partial h}{\partial x_n}(a)} \right) = \lambda \frac{\partial h}{\partial x_i}(a) \ ,$$

also tatsächlich

$$\operatorname{grad} f(a) = \lambda \operatorname{grad} h(a) \ .$$

QED.

6.3.2 Beispiel

Wir betrachten die Funktion $f : K \to \mathbb{R}$,

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

auf

$$K = \{(x,y) \in \mathbb{R}^2 | x^2 + y^2 \leq 1\} \ .$$

Da K beschränkt und abgeschlossen, also kompakt ist (und f stetig), nimmt f sein Supremum (und sein Infimum) $c = \sup_{x \in K} f(x)$ an! Wir wollen $c \in \mathbb{R}$ berechnen und die Stellen $(x,y) \in K$, an denen f dieses annimmt.

(I) $\operatorname{grad} f(x,y) = \begin{pmatrix} -2x \\ 2y \end{pmatrix} = 0 \Rightarrow (x,y) = (0,0)$,
Hess $f(0,0) = \begin{pmatrix} -2 & 0 \\ 0 & 2 \end{pmatrix}$ indefinit! Also ist $(0,0)$ kein lokales Extremum.
Schluss: $f|_G : G \to \mathbb{R}$, $G = \{(x,y) \in K | x^2 + y^2 < 1\}$ hat kein Maximum. (Beachte, dass 5.2.1 nur eine notwendige Bedingung liefert, wenn der Definitionsbereich ein Gebiet ist, also offen).
Folgerung: Die Maxima (und die Minima) von f müssen auf

$$M = \{(x,y) \in \mathbb{R} | x^2 + y^2 = 1\}$$

liegen.

(II) Betrachte deshalb die Nebenbedingung $h = 0$, wo $h : \mathbb{R}^2 \to \mathbb{R}$ gegeben ist durch

$$h(x,y) = x^2 + y^2 - 1 \ .$$

Hat nun $f|_M : M \to \mathbb{R}$ ein Maximum in $(a,b) \in M$, so muss gelten

$$\operatorname{grad} f(a,b) = \lambda \operatorname{grad} h(a,b)$$

für ein $\lambda \in \mathbb{R}$. (Es ist nämlich $\operatorname{grad} h(x,y) = (2x, 2y) \neq 0$ für alle $(x,y) \in M$.) Wir erhalten deshalb die folgenden Gleichungen

$$-2x = \lambda \cdot 2x \ , \tag{6.9}$$
$$2y = \lambda \cdot 2y \ , \tag{6.10}$$
$$x^2 + y^2 = 1 \ . \tag{6.11}$$

Ist $x \neq 0$, so folgt aus Gleichung (6.9), dass $\lambda = -1$ und daher aus Gleichung (6.10) $y = 0$. Wegen Gleichung (6.11) folgt $x = \pm 1$. Ist dagegen $x = 0$, so ist Gleichung (6.9) offenbar erfüllt, aus Gleichung (6.11) folgt $y = \pm 1$ (und dann aus Gleichung (6.10), dass $\lambda = +1$). Man erhält also die vier Kandidaten:

$$(a,b) \in \{(\pm 1, 0), (0, \pm 1)\} \ .$$

Da schließlich $f(\pm 1, 0) = -1$ und $f(0, \pm 1) = +1$ ist, ist $c = +1$ und wird genau in den Punkten $(a,b) = (0, \pm 1)$ angenommen.

Kapitel 7

Gewöhnliche Differenzialgleichungen

13.01.2005

Motivation

Ein deterministisches System auf einem Phasenraum G ist dadurch gegeben, dass jedem Punkt $x \in G$ seine „Dynamik“ $\varphi(x) : I(x) \to G$ zugeordnet wird. $I(x) \subseteq \mathbb{R}$ ist ein offenes Intervall mit $0 \in I(x)$. Wir schreiben dann

$$\varphi^t(x) := \varphi(x)(t)$$

und verlangen

$$\varphi^0(x) = x \ .$$

Weiterhin muss die *Verträglichkeitsbedingung* gelten:

$$\varphi^s(\varphi^t(x)) = \varphi^{s+t}(x)$$

für alle $t \in I(x), s \in I(\varphi^t(x))$. Insbesondere gilt für alle $t \in I(x)$:

$$s \in I(\varphi^t(x)) \ \Leftrightarrow \ s + t \in I(x) \ .$$

7.1 Dynamische Systeme

Definition 7.1.
Ein *dynamisches System* auf einem Gebiet $G \subseteq \mathbb{R}^n$ ist gegeben durch eine stetig differenzierbare Abbildung $\varphi : \Omega \to G$, $(t, x) \mapsto \varphi^t(x) (:= \varphi(t, x))$ mit folgenden Eigenschaften:

(a) $\Omega \subseteq \mathbb{R} \times G$ ist ein Gebiet, so dass $\{0\} \times G \subseteq \Omega$ ist und für jedes $x \in G$

$$I(x) := \left\{ t \in \mathbb{R} \middle| (t, x) \in \Omega \right\}$$

wegzusammenhängend ist.

(b) (i) für alle $x \in G$ gilt

$$\varphi^0(x) = x$$

(ii) für alle $x \in G$ und $t \in I(x)$ gilt: Es ist $s \in I(\varphi^t(x))$, genau dann wenn $s + t \in I(x)$ und es gilt dann:

$$\varphi^s(\varphi^t(x)) = \varphi^{s+t}(x)$$

7.1.1 Kommentar

(1) Wir nennen $G \subseteq \mathbb{R}^n$ den *Phasenraum* des dynamischen Systems $\varphi : \Omega \to G$

(2) Weil $\Omega \subseteq \mathbb{R} \times G$ offen ist, ist auch $I(x) \subseteq \mathbb{R}$ offen und wegen Definition 7.1. (a) auch wegzusammenhängend, also ein offenes Intervall

$$I(x) = (t_-(x), t_+(x)) \ .$$

Hierbei ist $t_-(x) \in [-\infty, 0)$, $t_+(x) \in (0, +\infty]$. Wir nennen $t_-(x)$ den *Anfang von* x und $t_+(x)$ das *Ende von* x.

(3) Setzen wir für jedes $t \in \mathbb{R}$

$$G_t = \big\{x \in G \big| (t,x) \in \Omega\big\} \subseteq G \ ,$$

so ist G_t offen und es folgt aus Definition 7.1. (b), dass $\varphi^t(G_t) = \big\{\varphi^t(x) \big| x \in G_t\big\} = G_{-t}$ ist und es ist

$$\varphi^t : G_t \to G_{-t}$$

ein Diffeomorphismus, denn $\varphi^{-t} : G_{-t} \to G_t$ ist sein Inverses:

$$\varphi^{-t} \circ \varphi^t(x) = \varphi^{-t+t}(x) = \varphi^0(x) = x = \operatorname{id}(x) \ ,$$

$$\varphi^t \circ \varphi^{-t}(x) = \varphi^{t-t}(x) = \varphi^0 = x = \operatorname{id}(x) \ .$$

Man nennt die Familie $(\varphi^t)_{t \in \mathbb{R}}$ von diesen Diffeomorphismen den zugehörigen *Fluss* (oder *Strömung*) von φ. Die Verträglichkeitsbedingung liest sich dann gerade so:

$$\varphi^s \circ \varphi^t = \varphi^{s+t}$$

(dort, wo beide Seiten definiert sind).

(4) Besonders schön ist die Situation, wenn $I(x) = \mathbb{R}$ (also $t_-(x) = -\infty$ und $t_+(x) = +\infty$) ist, für alle $x \in G$, d.h. $G_t = G$ für alle $t \in \mathbb{R}$. In diesem Fall ist $\Omega = \mathbb{R} \times G$ und man spricht von einem *globalen dynamischen System* (oder von einem *globalen Fluss*) auf G. In diesem Fall ist also

$$\mathbb{R} \to \operatorname{Diff}(G) \ , \quad t \mapsto \varphi^t$$

ein Gruppenhomomorphismus von $(\mathbb{R}, +)$ nach $(\operatorname{Diff}(G), \circ)$ wobei

$$\operatorname{Diff}(G) = \{f : G \to G \mid f \text{ Diffeomorphismus}\} \ .$$

(Umgekehrt definiert ein solcher (stetig differenzierbarer) Homomorphismus ein globales dynamisches System auf G.)

(5) Dynamische Systeme sind also geeignet, Prozesse zu beschreiben, die durch drei Eigenschaften gekennzeichnet sind:

Endlich-Dimensionalität: Jeder mögliche Zustand wird durch (nur) endlich viele (reelle) Parameter eindeutig bestimmt

Determinismus: Die Zukunft wie auch die Vergangenheit eines Zustandes sind durch den (augenblicklichen) Zustand vollständig festgelegt

Differenzierbarkeit: Die Änderung der Zustände geschieht sowohl in der Zeit, als auch in der Abhängigkeit des Anfangswertes in stetig differenzierbarer Weise

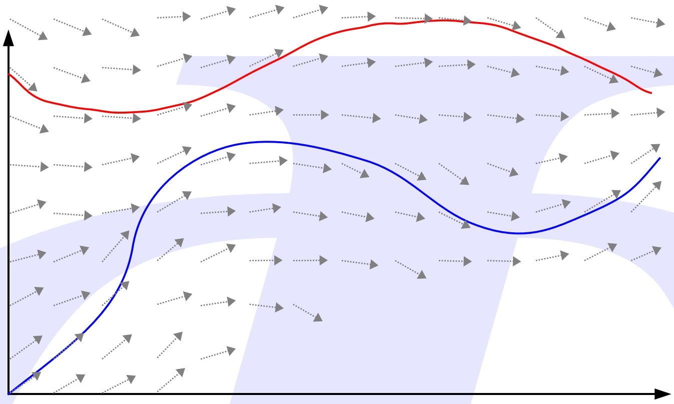
(6) Man schreibt auch – ungeachtet der Zweideutigkeit – statt $\varphi^t(x) = x(t)$. Die Verträglichkeitsbedingung lautet dann

$$x(t)(s) = x(t+s)$$

Definition 7.2.
Sei φ ein dynamisches System auf $G \subseteq \mathbb{R}^n$. Es heißt dann

$$f : G \to \mathbb{R}^n \quad , \; f(x) := \frac{d}{dt}\Big|_{t=0} \varphi^t(x) \; \big(= \dot{x}(0) \; \big)$$

das *zu* φ *gehörende Vektorfeld auf* G.



7.1.2 Bemerkung

Sei φ ein dynamisches System auf $G \subseteq \mathbb{R}^n$ und $f : G \to \mathbb{R}^n$ ihr Vektorfeld. Dann gilt für alle $x \in G$ und alle $t \in I(x)$ (nicht für $t = 0$):

$$\dot{x}(t) = f\big(x(t)\big) \; .$$

¡? Beweis ¿!

Sei $x \in G$ und $t \in I(x)$ beliebig. Mit $y := x(t)$ gilt nun:

$$f(x(t)) = f(y) = \frac{d}{ds}\Big|_{s=0} y(s) = \frac{d}{ds}\Big|_{s=0} x(t)(s) \stackrel{!}{=} \frac{d}{ds}\Big|_{s=0} x(t+s) = \frac{d}{ds}\Big|_{s=t} x(s) = \dot{x}(t) \; .$$

QED. 18.01.2005

7.1.3 Kommentar

(1) Es erfüllt also jede der Kurven $I(x) \to G \,, \quad t \mapsto x(t) \;\; (x \in G)$ des dynamischen Systems das System von n Gleichungen

$$\begin{array}{rcl} \dot{x}_1(t) & = & f_1\big(x_1(t), \ldots, x_n(t)\big) \\ & \vdots & \\ \dot{x}_n(t) & = & f_n\big(x_1(t), \ldots, x_n(t)\big) \end{array}$$

oder kurz

$$\dot{x} = f(x) \; .$$

Man nennt ein solches System aus n Gleichungen ein *System gewöhnlicher Differentialgleichungen*, weil in den Gleichungen sowohl die Funktionen $x_1, \ldots, x_n$ als auch ihre Ableitungen $\dot{x}_1, \ldots, \dot{x}_n$ auftreten. „Gewöhnlich“, weil die Funktionen $x_1, \ldots, x_n$ nur von einer reellen Veränderlichen t abhängen (und nicht von mehreren Veränderlichen $\hookrightarrow$ partielle Differentialgleichungen).

(2) In den meisten Fällen ist es nun so, dass man das dynamische System φ auf G gar nicht genau kennt – es aber sehr gerne kennen würde – wohl aber das zugehörige Vektorfeld $f : G \to \mathbb{R}^n$. Es stellt sich die fundamentale Frage:
In wie weit legt das zugehörige Vektorfeld f auf $G \subseteq \mathbb{R}^n$ das dynamische System φ fest? Der zentrale Satz wird sein: Jedes (stetig differenzierbare) Vektorfeld $f : G \to \mathbb{R}^n$ legt genau ein (maximales) dynamisches System φ auf G fest! (siehe 7.2.4)

In der gewöhnlichen Differenzialgleichung

$$\dot{x} = f(x)$$

ist dann also nach einer Funktion $x : I(x) \to G$ (insbesondere auch $I(x) \subseteq \mathbb{R}$) gesucht, die $\dot{x} = f(x)$ und $x(0) = x$ erfüllt; das Vektorfeld f ist gegeben.

7.1.4 Beispiel

In der klassischen Mechanik ordnet man jedem Materieteil eine positive Zahl zu: seine (träge) Masse $m > 0$. Eines der grundlegenden Newton'schen Gesetze besagt (bekanntlich): Bewegt sich ein Teilchen der Masse $m > 0$ in einem Gebiet $D \subseteq \mathbb{R}^3$ unter dem Einfluss der Kraft $F : D \to \mathbb{R}^3$ (das ist ein Vektorfeld auf D), so befolgt die Bahn des Teilchens $t \mapsto x(t)$ die Gleichung:

$$m\ddot{x} = F(x) \ .$$

Setzen wir $G := D \times \mathbb{R}^3 \subseteq \mathbb{R}^6$ (ein Gebiet im $\mathbb{R}^6$) und $f : G \to \mathbb{R}^6$

$$f(x, y) = \left(y, \frac{1}{m}F(x)\right)$$

so wie $z : I \to G$,

$$z(t) = \big(x(t), \dot{x}(t)\big)$$

so befolgt z die folgende Gleichung

$$\dot{z}(t) = \big(\dot{x}(t), \ddot{x}(t)\big) = \left(\dot{x}(t), \frac{1}{m}F\big(x(t)\big)\right) = f\big(x(t), \dot{x}(t)\big) = f\big(z(t)\big) \ .$$

Bestimmt man zum Vektorfeld $f : G \to \mathbb{R}^6$ das zugehörige dynamische System φ (d.h. alle Lösungskurven $z : I(z) \to G$), so erhält man durch Projektion auf die ersten drei Koordinaten alle Teilchenbahnen $x : I(z) \to D$ (bei Vorgabe von $z \in G$) im Kraftfeld F.

7.2 Anfangswertproblem

Definition 7.3.
Sei $G \subseteq \mathbb{R}^n$ ein Gebiet und $f : G \to \mathbb{R}^n$ ein stetig differenzierbares Vektorfeld auf G.

(a) Sei $I \subseteq \mathbb{R}$ ein offenes Intervall. Man nennt eine stetig differenzierbare Kurve $x : I \to G \subseteq \mathbb{R}^n \quad t \mapsto x(t)$, eine *Lösung der Differenzialgleichung*

$$\dot{x} = f(x) \ , \tag{7.1}$$

wenn für alle $t \in I$ gilt: $\dot{x}(t) = f\big(x(t)\big)$.

(b) Ist $I \subseteq \mathbb{R}$ ein offenes Intervall mit $0 \in I$, so nennt man eine Lösung $x : I \to G$ von Gleichung (7.1) eine Lösung zum Anfangswert $x_0 \in G$, wenn $x(0) = x_0$ ist.

7.2.1 Kommentar

Man verdeutlicht sich das Vektorfeld $f : G \to \mathbb{R}^n$ oft durch das sogenannte *Phasendiagramm* von f. Eine Lösungskurve $x : I \to G$ wird auch als *Integralkurve* bezeichnet, weil man in einfachen Fällen die Lösungskurven durch den Prozess von Stammfunktionbildung erhalten kann (7.2.9). Beim Lösen spricht man auch vom *Integrieren der gewöhnlichen Differenzialgleichung*.

7.2.2 Beispiele

(I) Der freie Fall

In der Nähe der Erdoberfläche zeigt die Beobachtung, dass die Erdanziehungskraft F auf ein Teilchen der (schweren) Masse $m > 0$ nahezu unabhängig vom Ort ist und proportional zu m:

$$F(x) = -mg\hat{e}_3 \qquad g > 0 \ , \ \hat{e}_3 = (0,0,1)$$

bei geeigneter Wahl der Koordinaten $x = (x_1, x_2, x_3)$. 20.01.2005

Die Erde erzeugt gewissermaßen dort nicht ein Kraftfeld, sondern ein Gravitationsfeld G mit $G(x) = -g\hat{e}_3$. Das Kraftfeld, welches auf das Teilchen (der schweren) Masse $m > 0$ wirkt, ist dann $F(x) = mG(x)$, hängt also vom Teilchen selbst ab (Gravitationsgesetz).
Unser „Zustandsraum“ ist damit $D = \mathbb{R}^2 \times \mathbb{R}_+$ und wir müssen die Gleichung

$$m\ddot{x} = -mg\hat{e}_3$$

also

$$\ddot{x} = -g\hat{e}_3 \ ,$$

(wegen träge Masse = schwere Masse) integrieren, d.h. auf $G = D \times \mathbb{R}^3 \subseteq \mathbb{R}^6$ das System

$$\begin{aligned} \dot{x}_1 &= y_1 \\ \dot{x}_2 &= y_2 \\ \dot{x}_3 &= y_3 \\ \dot{y}_1 &= 0 \\ \dot{y}_2 &= 0 \\ \dot{y}_3 &= -g \ . \end{aligned}$$

Die Lösungen sieht man hier unmittelbar durch Integration:

$$\begin{aligned} x_1(t) &= b_1 t + a_1 \\ x_2(t) &= b_2 t + a_2 \\ x_3(t) &= -\frac{1}{2}gt^2 + b_3 t + a_3 \\ y_1(t) &= b_1 \\ y_2(t) &= b_2 \\ y_3(t) &= -gt + b_3 \end{aligned}$$

mit (Integrations-) Konstanten $a_1, \ldots, b_3 \in \mathbb{R}$. Es ist dann $x(0) = a$ und $\dot{x}(0) = b$, also das zu $f : G \to \mathbb{R}^6 \quad f(x,y) = (y, -g\hat{e}_3)$ gehörende dynamisches System φ gegeben durch

$$\varphi^t(x,y) = \left(-\frac{1}{2}gt^2\hat{e}_3 + yt + x, -gt\hat{e}_3 + y \right)$$

(und $\varphi : \Omega \to G$ mit $\Omega = \left\{ (t,x,y) \in \mathbb{R} \times G \middle| -\frac{1}{2}gt^2 + ty_3 + x_3 > 0 \right\}$) und damit die Lösung von $m\ddot{x} = F(x)$ auf $D \subseteq \mathbb{R}^3$ zum Anfangswert $(x,y) \in G$:

$$\psi^t(x,y) = -\frac{1}{2}gt^2\hat{e}_3 + ty + x \quad (\ = x(t) \ )$$

($\hookrightarrow$ Parabelbahn).

(II) Das Hooke'sche Gesetz:

Ein Teilchen der Masse $m > 0$ bewege sich vertikal unter dem Einfluss der Kraft einer Feder. Das Hooke'sche Gesetz besagt: Die Kraft, die auf den Körper einwirkt, ist proportional zur Auslenkung und entgegengesetzt orientiert,

$$F(x) = -kx$$

mit $k > 0$ (Federkonstante).
Die gewöhnliche Differenzialgleichung, die also zu lösen ist, ist

$$m\ddot{x} = -kx \ .$$

Setzt man

$$\omega := \sqrt{\frac{k}{m}} > 0 \quad , G = \mathbb{R} \times \mathbb{R} \quad , \quad f(x,y) = \big(y, -\omega^2 x\big) \quad ,$$

so muss man also das dynamische System von

$$\begin{aligned} \dot{x} &= y \\ \dot{y} &= -\omega^2 x \end{aligned}$$

bestimmen. Es ist gegeben durch (systematische Herleitung folgt später)

$$\begin{aligned} x(t) &= x \cos \omega t + \frac{y}{\omega} \sin \omega t \\ y(t) &= -\omega x \sin \omega t + y \cos \omega t \ . \end{aligned}$$

Definition 7.4.
Sei $m \in \mathbb{N}$. Ein System gewöhnlicher Gifferenzialgleichungen m-ter Ordnung auf einem Gebiet $D \subseteq \mathbb{R}^n$ ist gegeben durch eine stetige differenzierbare Abbildung

$$f : D \times \underbrace{\mathbb{R}^n \times \cdots \times \mathbb{R}^n}_{(m-1)\text{ -mal}} \to \mathbb{R}^n$$

und der Gleichung

$$x^{(m)} = f\left(x, \dot{x}, \ldots, x^{(m-1)}\right) \ . \tag{7.2}$$

7.2.3 Kommentar

(1) Unter einer Lösung von Gleichung (7.2) versteht man eine stetig differenzierbare Kurve, $x : I \to D \subseteq \mathbb{R}$, $I \subseteq \mathbb{R}$ ein offenes Intervall, so dass für alle $t \in I$ gilt:

$$x^{(m)}(t) = f\left(x(t), \dot{x}(t), \ldots, x^{(m-1)}(t)\right) \ .$$

(2) Eine Lösung $x : I \to D$ heißt eine *Lösung zum Anfangswert* $(x_0, y_1, \ldots, y_{m-1}) \in G := D \times \mathbb{R}^{n(m-1)}$, wenn gilt:

$$\begin{aligned} x(0) &= x_0 \\ \dot{x}(0) &= y_1 \\ &\vdots \\ x^{(m-1)}(0) &= y_{m-1} \ . \end{aligned}$$

7.2.4 Bemerkung

Man kann die Lösung eines Systems $m - ter$ Ordnung auf einem Gebiet $D \subseteq \mathbb{R}^n$,

$$x^{(m)} = f\left(x(t), \dot{x}(t), \ldots, x^{(m-1)}(t)\right) ,$$

stets auf die Lösung eines System 1. Ordnung auf $G = D \times \mathbb{R}^{n(m-1)} \subseteq \mathbb{R}^{mn}$ zurückführen. Man definiert dazu das Vektorfeld $g : G \to \mathbb{R}^{mn}$ durch

$$g(x, y_1, \ldots, y_{m-1}) = \big(y_1, \ldots, y_{m-1}, f(x, y_1, \ldots, y_{m-1})\big)$$

und löst

$$\dot{z} = g(z)$$

mit $z = (x, y_1, \ldots, y_{n-1}) \in G$.

¡? Beweis ¿!

Ist $x : I \to D \subseteq \mathbb{R}^n \quad t \mapsto x(t)$ Lösung von $x^{(m)} = f\left(x(t), \dot{x}(t), \ldots, x^{(m-1)}(t)\right)$, so ist $z : I \to G \subseteq \mathbb{R}^{nm}$

$$z(t) = \left(x(t), \dot{x}(t), \ldots, x^{(m)}(t)\right) ,$$

Lösung von $\dot{z} = g(z)$.
Ist $z : I \to G \quad t \mapsto z(t)$ Lösung von $\dot{z} = g(z)$ und $z(t) = (x(t), y(t), \ldots, y_{m-1}(t))$, so ist $x : I \to D \quad t \mapsto x(t)$ Lösung von

$$x^{(m)} = f\big(x, \ldots, x^{(m-1)}\big) .$$

QED.

Definition 7.5.
Sei $I \subseteq \mathbb{R}$ ein offenes Intervall. Ein (stetig differenzierbares) *zeitabhängiges Vektorfeld* f auf einem Gebiet $D \subseteq \mathbb{R}^n$ ist eine (stetig differenzierbare) Abbildung $f : I \times D \to \mathbb{R}^n \quad (t, x) \mapsto f(t, x)$. Man nennt dann das System gewöhnlicher Differenzialgleichungen

$$\dot{x} = f(t, x) \tag{7.3}$$

nicht-autonom, wenn f explizit von $t \in I$ abhängt. Ist f unabhängig von t (wie bisher), so heißt $\dot{x} = f(x)$ ein *autonomes System.*

7.2.5 Kommentar

(1) Sei $J \subseteq I$ ein offenes Teilintervall und $\dot{x} = f(t, x)$ ein (nicht-autonomes) System auf $D \subseteq \mathbb{R}^n$. Eine stetig differenzierbare Kurve $x : J \to D$ heißt eine Lösung von Gleichung (7.3), wenn für alle $t \in J$ gilt:

$$\dot{x}(t) = f\big(t, x(t)\big) .$$

(2) Sei $J \subseteq I$ und $x : J \to D$ eine Lösung von Gleichung (7.3) und $t_0 \in J$ und $x_0 \in D$. Es heißt dann x eine Lösung von Gleichung (7.3) zum *Anfangswert* $x_0 \in D$ *zur Anfangszeit* $t_0 \in J$, wenn gilt

$$x(t_0) = x_0 .$$

7.2.6 Bemerkung

Man kann die Lösung eines nicht-autonomen Systems $\dot{x} = f(t, x)$ auf einem Gebiet $D \subseteq \mathbb{R}^n$ auf die Lösung eines autonomen Systems auf $G := I \times D$ wie folgt zurückführen:
Man setze $g : G \to \mathbb{R}^{n+1} = \mathbb{R} \times \mathbb{R}^n$

$$g(s, x) = \big(1, f(s, x)\big) ,$$

$z = (s, x) \in G$ und löse dann $\dot{z} = g(z)$ auf G.

¡? Beweis ¿!

Sei $x : J \to D$ eine Lösung von Gleichung (7.3) zum Anfang $(t_0, x_0) \in G$, also $x(t_0) = x_0$. Es ist dann $z : J \to G = I \times D$

$$z(t) = \big(t, x(t)\big)$$

Lösung von $\dot{z} = g(z)$ mit Anfangsbedingungen $z(t_0) = (t_0, x_0)$. Ist umgekehrt $z : J \to G$ Lösung von $\dot{z} = g(z)$ mit Anfang $z(t_0) = (t_0, x_0) =: z_0$, so ist $z(t) = \big(t, x(t)\big)$ und $x : J \to D$ dann eine Lösung von Gleichung (7.3) zum Anfang x_0 bei Anfangszeit $t_0 \in I$.

QED.

7.2.7 Kommentar

Ist $f : G \to \mathbb{R}^n$ ein stetig differenzierbares Vektorfeld auf einem Gebiet $G \subseteq \mathbb{R}^n$, so spielt die Zeit $t_0 = 0$ bei der Lösung des Anfangswertproblems

$$\text{(AWP)} \begin{cases} \dot{x} & = f(x) \\ x(0) & = x_0 \end{cases}$$

keine besondere Rolle. In vollkommen analoger Weise spricht man bei einer stetig differenzierbaren Kurve $x : I \to G$ ($I \subseteq \mathbb{R}$ offen) von einer *Lösung des Anfangswertproblems*

$$\text{(AWP)} \begin{cases} \dot{x} & = f(x) \\ x(t_0) & = x_0 \end{cases} ,$$

wenn $x : I \to G$ eine Lösung von $\dot{x} = f(x)$ und $x(t_0) = x_0$ ist. Hierbei ist $t_0 \in \mathbb{R}$ eine beliebig vorgegebene Zeit.

25.01.2005

7.2.8 Beispiele

(I) Keplers Gleichung

Nimmt man die Sonne als (wegen ihrer im Vergleich zu den Planeten riesigen Masse) als unbeweglich im Ursprung eines Koordinatensystems $x = (x_1, x_2, x_3)$ an, so bewegt sich ein Planet (unter Vernachlässigung des Einflusses der anderen Planeten) vermöge folgender Gleichung:

$$\ddot{x} = -1 \cdot \frac{1}{|x|^2} \frac{x}{|x|} = -\frac{x}{|x|^3} , \quad x \in \mathbb{R}^3 \backslash \{0\} .$$

Die Einheiten sind so gewählt, dass für die Konstanten gilt: $M = 1$ (Sonnenmasse) und $\gamma = 1$ (Gravitationskonstante).

Der Phasenraum ist also hier $G = \big(\mathbb{R}^3 \backslash \{0\}\big) \times \mathbb{R}^3 \subseteq \mathbb{R}^6$ und das Vektorfeld $f : G \to \mathbb{R}^6$ ist

$$f(x, y) = \left(y, -\frac{x}{|x|^3} \right) .$$

Das Gravitationsfeld $G(x) = -\frac{x}{|x|^3}$, welches die Sonne erzeugt, ist also zur Sonne gerichtet und umgekehrt proportional zum Quadrat des Abstandes. Man kann zeigen: Ist für die Anfangsbedingung $(x, y) \in G$ die folgende Größe

$$L(x, y) = x \times y \neq 0 ,$$

so existiert die Lösung $t \mapsto x(t)$ für alle Zeiten $I(x, y) = \mathbb{R}$ und beschreibt eine Ellipse, Parabel oder Hyperbel, je nachdem ob die Größe

$$E(x, y) = \frac{1}{2} |y|^2 - \frac{1}{|x|} = \begin{cases} < & 0 \\ = & 0 \\ > & 0 \end{cases}$$

ist. Die Abbildungen $L : G \to \mathbb{R}^3$ und $E : G \to \mathbb{R}$ sind *Bewegungsinvarianten* für $\ddot{x} = -\frac{x}{|x|^3}$, d.h. für alle $(x, y) \in G$ gilt für die zugehörige Bahn $t \mapsto \big(x(t), y(t)\big) \in G$:

$$L\big(x(t), y(t)\big) = L(x, y) \quad , \quad E\big((x(t), y(t)\big) = E(x, y) \ .$$

(II) Das n-Körperproblem:
Bewegen sich n Körper ($n \in \mathbb{N}$) mit den Massen $m_1, \ldots, m_n > 0$ im 3-dim Raum unter dem Einfluss ihrer gegenseitigen Gravitation, so erfüllen die Koordinaten $x_1, \ldots, x_n \in \mathbb{R}^3$ ihres Ortes die Gleichungen (Gravitationsgesetz):

$$m_j \ddot{x}_j \quad = \quad -\sum_{\substack{i=1 \\ i \neq j}}^{n} m_i m_j \frac{x_i - x_j}{|x_i - x_j|^3}, \quad (j = 1, \ldots, n) \ .$$

Der Phasenraum ist also hier

$$G = \big(\mathbb{R}^{3n} \backslash \Delta\big) \times \mathbb{R}^{3n} \subseteq \mathbb{R}^{6n} \ ,$$

wobei Δ die „Diagonale“

$$\Delta = \big\{(x_1, \ldots, x_n) \in \mathbb{R}^{3n} | x_i = x_j \text{ für ein Paar } (i, j), \ i \neq j\big\}$$

bezeichnet. Für das zugehörige Vektorfeld $f : G \to \mathbb{R}^{6n}$ erhält man, dass

$$f(x, y) = \big(y, -\mathrm{grad}\,(U)(x)\big)$$

ist, wobei $U : \mathbb{R}^{3n} \backslash \Delta \to \mathbb{R}$ das *Potential*

$$U(x) = -\sum_{i<j} \frac{m_i m_j}{|x_i - x_j|}$$

bezeichnet (sogenanntes *konservatives Kraftfeld*). Für $n \geq 3$ sind die Lösungskurven (für alle $(x, y) \in G$) nicht bekannt!

7.2.9 Kommentar

(1) Im Allgemeinen hat man keine Chance ein System gewöhnlicher Differenzialgleichungen

$$\dot{x} = f(x)$$

auf einem Gebiet $G \subseteq \mathbb{R}^n$ bei Vorgabe von $f : G \to \mathbb{R}^n$ zu integrieren, d.h. das zugehörge dynamische System φ explizit zu bestimmen. (In Theorem 7.9. wird nur die abstrakte Existenz und Eindeutigkeit gezeigt.) Man ist daher (etwa beim 3-Körperproblem) schon damit zufrieden, bloße qualitative Aussagen über die Bahn $t \mapsto x(t)$ (in Abhänggkeit von der Anfangslage $x(0) = x_0$) zu bekommen, z.B. Antworten auf die Fragen

- Ist $t \mapsto x(t)$ periodisch, d.h.: es gibt ein $T > 0$, so dass $x(t + T) = x(t)$ für alle $t \in \mathbb{R} = I(x)$?
- Konvergiert $t \mapsto x(t)$ für $t \longrightarrow \infty$ gegen eine *Gleichgewichtslage* $y \in G$, d.h. ein Punkt mit $I(y) = \mathbb{R}$ und $y(t) = y$, $\forall t \in \mathbb{R}$?
- Oder auch nur, für welche $x \in G$ ist überhaupt $I(x) = \mathbb{R}$?

Nach Poincaré nennt man dieses Studium die sogenannte „Qualitative Theorie gewöhnlicher Differenzialgleichungen“. Es ist ein eigenständiges Forschungsgebiet der Mathematik, meistens unter dem Titel „Dynamische Systeme“ bekannt.

(2) Hat man spezielle Informationen über f (oder auch nur G), so kann man freilich $\dot{x} = f(x)$ in manchen Fällen doch explizit integrieren, z.B. in folgender Sitution:

7.3 Lösungsansätze

Proposition 7.6.
Sei $G \subseteq \mathbb{R}^1$ ein Intervall und $f : G \to \mathbb{R}$ eine stetig differenzierbare Funktion und $x_0 \in G$ mit $f(x_0) \neq 0$. Seien weiter $\delta_1, \delta_2 > 0$ derart, dass $(x_0 - \delta_1, x_0 + \delta_2) \subseteq G$ und $f(x) \neq 0$ für alle $x \in (x_0 - \delta_1, x_0 + \delta_2)$. Sei dann $\tau : (x_0 - \delta_1, x_0 + \delta_2) \to \mathbb{R}$ gegeben durch

$$\tau(x) = \int_{x_0}^{x} \frac{\mathrm{d}u}{f(u)}$$

Dann gilt: Ist $I = \mathrm{Bild}\,(\tau) \subseteq \mathbb{R}$, so ist $I \subseteq \mathbb{R}$ ein offenes Intervall $\tau : (x_0 - \delta_1, x_0 + \delta_2) \to I$ ein Diffeomorphismus und seine Umkehrung $\tau^{-1} =: \varphi$ ist Lösungskurve von $\dot{x} = f(x)$ auf G zum Anfang x_0.

¡? Beweis ¿!

Es ist $\tau'(x) = \frac{1}{f(x)} \neq 0$ und nach dem Zwischenwertsatz für alle $x \in (x_0 - \delta_1, x_0 + \delta_2)$ von gleichem Vorzeichen, also τ streng monoton. Wiederum nach dem Zwischenwertsatz ist daher τ bijektiv auf sein Bild I und nach dem Umkehrsatz ist $I \subseteq \mathbb{R}$ dann auch offen und für die Umkehrung $\varphi : I \to (x_0 - \delta_1, x_0 + \delta_2)$ gilt

$$\dot{\varphi}(t) = \frac{1}{\tau'(\varphi(t))} = \frac{1}{\frac{1}{f(\varphi(t))}} = f(\varphi(t)) \; ,$$

für alle $t \in I$. Also ist τ Diffeomorphisumus, $\varphi : I \to G$ Lösung von $\dot{x} = f(x)$ und hat Anfang (wegen $\tau(x_0) = 0$):

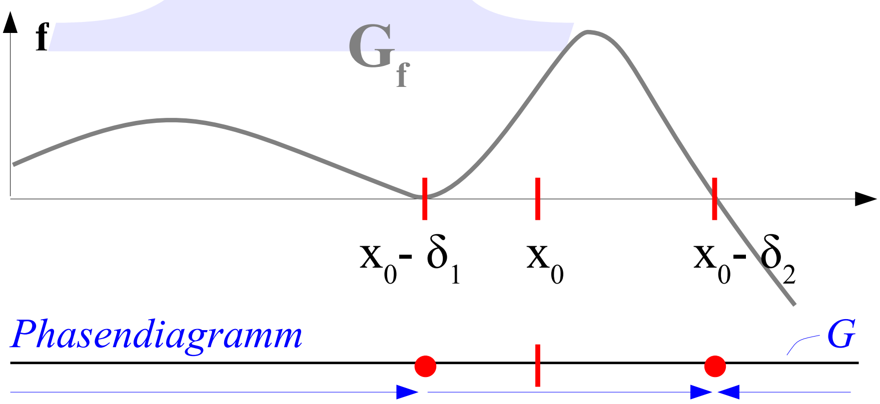
$$\varphi(0) = x_0 \; .$$

QED.

7.3.1 Kommentar

(1) Man kann also eindimensionale Gleichungen durch „Quadratur“ lösen, d.h. durch die Prozesse der „Stammunktionbildung“ und „Invertierung“.

(2) Ganz allgemein sind die Punkte $x_0 \in G$ wo ein gegebenes Vektorfeld $f : G \to \mathbb{R}^n$ ($G \subseteq \mathbb{R}^n$ ein Gebiet) verschwindet, $f(x_0) = 0$, für das dynamische System $\dot{x} = f(x)$ besonders einfach: Es ist nämlich $I(x_0) = \mathbb{R}$ und $x_0(t) = x_0 \;\; \forall\, t \in \mathbb{R}$ (sogenannte *Gleichgewichtslagen*). Damit ist das dynamische System φ von $\dot{x} = f(x)$ auf $G \subseteq \mathbb{R}$ vollständig bestimmt. Sein Phasendiagramm sieht so aus:



27.01.2005

In den Nullstellen von f ruht das System, zwischen ihnen läuft es (monoton) von einer zur nächsten Nullstelle (und braucht dafür unendlich viel Zeit). An den Rändern des Intervalls G können die Flusskurven in endlicher Zeit „entweichen“(vgl. 7.4.6).

(3) Der Beweis des Satzes zeigt die Berechtigung der folgenden „Physikerrechnung“: *Integration durch Trennung der Variablen*

$$\dot{x} = f(x) \quad \Rightarrow \quad \frac{\mathrm{d}x}{\mathrm{d}t} = f(x) \quad \Rightarrow \quad \frac{\mathrm{d}x}{f(x)} = \mathrm{d}t \quad \Rightarrow \quad \int_{x_0}^{x} \frac{\mathrm{d}x}{f(x)} = \int_0^t \mathrm{d}t = t(x)$$

7.3.2 Beispiel

(I) Sei $a \in \mathbb{R}$ und $f : \mathbb{R} \to \mathbb{R}$, $x \mapsto ax$ (linear). Die Lösung von $\dot{x} = ax$ auf $G = \mathbb{R}$ zum Anfangswert $x_0 > 0$ für $a \neq 0$ ist nach (Proposition 7.6.) für $x > 0$ gegeben durch

$$\begin{aligned} t(x) &= \int_{x_0}^{x} \frac{\mathrm{d}u}{au} = \frac{1}{a} \ln u \Big|_{x_0}^{x} = \frac{1}{a} \ln\left(\frac{x}{x_0}\right) \\ \Rightarrow \quad at &= \ln\left(\frac{x}{x_0}\right) \quad \Rightarrow \quad e^{at} = \frac{x}{x_0} \\ \Rightarrow \quad x(t) &= x_0 \cdot e^{at} \quad , \ \forall t \in \mathbb{R} \end{aligned}$$

Diese Formel gilt auch für alle $x_0 \in \mathbb{R}$, für alle $a \in \mathbb{R}$ (insbesondere ist der Fluss global).

(II) Sei $f : \mathbb{R} \to \mathbb{R}$, $f(x) = 1 + x^2$ (quadratisch). Die Lösung zur Anfangslage $x_0 = 0$ ist gegeben durch:

$$\begin{aligned} t(x) &= \int_0^x \frac{\mathrm{d}u}{1+u^2} = \arctan u|_0^x = \arctan x \\ \Rightarrow \quad x(t) &= \tan t \end{aligned}$$

Beachte: $I(0) = \big(t_-(0), t_+(0)\big) = \left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$, d.h. $t \mapsto x(t)$ ist „in endlicher Zeit im Unendlichen“.

7.4 Lipschitz-Stetigkeit

Definition 7.7.
Sei $G \subseteq \mathbb{R}^n$ ein Gebiet und $f : G \to \mathbb{R}^n$ ein Vektorfeld.

(a) Es heißt f *Lipschitz-stetig*, wenn es ein $L > 0$ gibt, so dass für alle $x, y \in G$ gilt:

$$\|f(x) - f(y)\| \leq L \cdot \|x - y\| \ .$$

Es heißt dann L eine *Lipschitz-Konstante* für f.

(b) Es heißt f *lokal Lipschitz-stetig*, wenn jedes $x \in G$ eine offene Umgebung $U \subseteq G$ besitzt, so dass $f|_U : U \to \mathbb{R}^n$ Lipschitz-stetig ist.
(f Lipschitz-stetig $\Leftrightarrow \exists L > 0 \ \ \forall x, y \in U : \|f(x) - f(y)\| \leq L \cdot \|x - y\|$)

7.4.1 Bemerkung

Sei $f : G \to \mathbb{R}^n$ ein stetig differenzierbares Vektorfeld auf einem Gebiet $D \subseteq \mathbb{R}^n$. Dann ist f lokal Lipschitz-stetig.

¡? Beweis ¿!

Sei $x_0 \in G$ beliebig und $r > 0$ so klein, dass $K := \overline{B_r}(x_0) \subseteq G$ ist. Sei $U := B_r(x_0)$ und weiter $\|.\|$ die Operatornorm auf $\mathrm{Hom}(\mathbb{R}^n, \mathbb{R}^n)$ (bzgl. der euklidischen Norm (vgl. 4.1.2)). Weil nun $K \to \mathbb{R}\,,\ x \mapsto \|Df(x)\|$ stetig ist und K kompakt, existiert ein $L > 0$ mit $\|Df(x)\| \leq L$ für alle $x \in K$. Sei nun $x, y \in U$ beliebig. Dann ist die Verbindungsgerade $\{(1-t)x + ty | t \in [0,1]\}$ ganz in U und deshalb nach dem Schrankensatz (Korollar 4.13.):

$$\|f(x) - f(y)\| \leq L \cdot \|x - y\|$$

Also ist $f|_U$ Lipschitz-stetig.
QED.

Proposition 7.8.
Sei $G \subseteq \mathbb{R}^n$ ein Gebiet und $f : G \to \mathbb{R}^n$ lokal Lipschitz-stetig. Ist nun $K \subseteq G$ ein beliebiges Komaktum, so gilt:

$$f\big|_K : K \to \mathbb{R}^n$$

ist Lipschitz-stetig.

¡? Beweis ¿!

Angenommen nicht.
$\Rightarrow \quad \forall n \in \mathbb{N}\ \exists (x_n, y_n) \in K \times K$, so dass

$$\|f(x_n) - f(y_n)\| > n\,\|x_n - y_n\|$$

ist. Nach Übergang zu einer Teilfolge von (x_n, y_n) dürfen wir annehmen, dass $(x_n) \longrightarrow p$ und $(y_n) \longrightarrow q$ mit $p, q \in K$ (Satz 3.3. Bolzano-Weierstraß).
Behauptung: $p = q$. Denn ist $\varepsilon > 0$ beliebig, wähle nun $n_0 \in \mathbb{N}$ so groß, dass $\|x_n - p\| < \frac{\varepsilon}{3}$ und $\|y_n - p\| < \frac{\varepsilon}{3}$ für $n \geq n_0$ ist. Dann ist:

$$\begin{aligned} \|p - q\| &= \|(p - x_n) + (x_n - y_n) + (y_n - q)\| \\ &\leq \|p - x_n\| + \|x_n - y_n\| + \|y_n - q\| \\ &< \frac{\varepsilon}{3} + \frac{1}{n} \underbrace{\|f(x_n) - f(y_n)\|}_{\leq \|f(x_n)\| + \|f(y_n)\|} + \frac{\varepsilon}{3}\ , \end{aligned}$$

für alle $n \geq n_0$. Sei weiter $M > 0$ so groß, dass $\|f(x)\| \leq M$ für alle $x \in K$ (vgl. Satz 3.9.) und $n \geq n_0$ so groß, dass $\frac{2M}{n} \leq \frac{\varepsilon}{3}$ ist. Mit diesem n ist dann:

$$\|p - q\| \leq \frac{\varepsilon}{3} + \frac{2M}{n} + \frac{\varepsilon}{3} \leq \varepsilon\ .$$

Für alle $\varepsilon > 0$! Also muss $\|p - q\| = 0$ sein, also $p = q$.
Aber f ist lokal Lipschitz-stetig. Daher existiert eine offene Umgebung $U \subseteq G$ von p und $L > 0$, so dass für alle $x, y \in U$ gilt:

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

Ist nun $n \in \mathbb{N}$ so groß, dass $n \geq L$ ist und $x_n, y_n \in U$, so ist:

$$\|f(x_n) - f(y_n)\| > n\,\|x_n - y_n\| > L\,\|x_n - y_n\|\ .$$

↯ Widerspruch ↯
QED.

Theorem 7.9. Existenz- und Eindeutigkeits-Satz von Picard-Lindelöf
Sei $G \subseteq \mathbb{R}^n$ ein Gebiet, $f : G \to \mathbb{R}^n$ ein stetig differenzierbares Vektorfeld und $x_0 \in G$. Dann existiert ein $\delta > 0$, so dass es genau eine stetig differenzierbare Abbildung $x : (-\delta, \delta) \to G$ gibt, die Lösung von $\dot{x} = f(x)$ zum Anfangswert x_0 ist,

$$\begin{aligned} \dot{x} &= f\big(x(t)\big) \\ x(0) &= x_0 \ . \end{aligned}$$

7.4.2 Kommentar

Die Grundidee beim Existenzbeweis (und dann auch für die Eindeutigkeit) ist folgende: Ist $x : (-\delta, \delta) \to G$ Lösung, so folgt durch Integration mit dem Hauptsatz der Differenzial- und Integralrechnung

$$x(t) - x(0) = \int_0^t x'(s) \, \mathrm{d}s = \int_0^t f\big(x(s)\big) \, \mathrm{d}s$$

also

$$x(t) = x_0 + \int_0^t f\big(x(s)\big) \, \mathrm{d}s \tag{7.4}$$

und damit ein Fixpunkt der Abbildung Φ, die jeder Funktion φ die Funktion $\Phi[\varphi]$ zuordnet, die gegeben ist durch

$$\Phi[\varphi](t) := x_0 + \int_0^t f\big(\varphi(s)\big) \, \mathrm{d}s \ .$$

Man versucht nun deshalb den Raum, auf dem Φ definiert ist, gerade so zu wählen, dass man den Banach'schen Fixpunktsatz (Satz 2.4.) anwenden kann.

01.02.2005

¡? Beweis ¿! des Satzes von Picard-Lindelöf

- **Schritt 1 (Fixpunkt)**

  Da G offen ist, existiert ein $r > 0$, so dass $K := \overline{B_r}(x_0) \subseteq G$. Da K kompakt und $f|_K$ stetig, existiert ein $M > 0$, so dass $\forall x \in K$ gilt: $\|f(x)\| \leq M$. Nach Bemerkung 7.4 und Proposition 7.8. existiert ein $L > 0$, so dass $f|_K$ Lipschitz-stetig mit Lipschitz-Konstante L ist. Setze nun

  $$\delta = \min\left\{\frac{1}{L}, \frac{r}{M}\right\}$$

  Sei nun $\tilde{\delta} < \delta$ beliebig. Wir betrachten den Vektorraum

  $$X = \Big\{\varphi : [-\tilde{\delta}, \tilde{\delta}] \to \mathbb{R}^n \big| \varphi \text{ stetig } \Big\} = \mathcal{C}([-\tilde{\delta}, \tilde{\delta}], \mathbb{R}^n)$$

  mit der Maximumsnorm

  $$\|\varphi\|_\infty = \max_{t \in [-\tilde{\delta}, \tilde{\delta}]} \|\varphi(t)\| \ .$$

  Es ist dann $(X, \|.\|_\infty)$ ein Banachraum, weil die gleichmäßigen Limiten stetiger Funktionen wieder stetig sind. Wir betrachten die Teilmenge

  $$A = \Big\{\varphi \in X \big| \varphi(0) = x_0, \varphi([-\tilde{\delta}, \tilde{\delta}]) \subseteq K\Big\} \ .$$

  Es ist dann A abgeschlossen. Schließlich betrachten wir für $\varphi \in A$ die neue Funktion

  $$\Phi[\varphi] : [-\tilde{\delta}, \tilde{\delta}] \to \mathbb{R}^n$$

  gegeben durch

  $$\Phi[\varphi](t) = x_0 + \int_0^t f\big(\varphi(s)\big) \, \mathrm{d}s \ .$$

Es ist dann $\Phi[\varphi]$ wieder stetig (sogar stetig differenzierbar) und es ist auch $\Phi[\varphi](0) = x_0$. Es ist auch $\Phi[\varphi](t) \in K \quad \forall t \in [-\tilde{\delta}, \tilde{\delta}]$, denn

$$\|\Phi[\varphi](t) - x_0\| \leq \left\| \int_0^t f\big(\varphi(s)\big)\, \mathrm{d}s \right\| \overset{\text{Lemma 4.14.}}{\leq} \left| \int_0^t \underbrace{\|f\big(\underbrace{\varphi(s)}_{\in K}\big)\|}_{\leq M}\, \mathrm{d}s \right| \leq \tilde{\delta} M \leq \frac{r}{M} \cdot M = r \; .$$

Damit ist Φ eine Abbildung von A nach A, $\Phi : A \to A$. Schließlich behaupten wir, dass Φ eine Kontraktion mit Konstante $\vartheta = \tilde{\delta} L \; , \; 0 < \vartheta < 1$ ist:

$$\begin{aligned}
\|\Phi[\varphi] - \Phi[\psi]\|_\infty &= \max_{|t| \leq \tilde{\delta}} \left\| \int_0^t \big(f(\varphi(s)) - f(\psi(s))\, \mathrm{d}s \big) \right\| \\
&\overset{\text{Lemma 4.14.}}{\leq} \max_{|t| \leq \tilde{\delta}} \int_0^t \big(\|f(\varphi(s)) - f(\psi(s))\|\; \mathrm{d}s \big) \\
&\leq \int_0^{\tilde{\delta}} L\, \|\varphi(s) - \psi(s)\|\; \mathrm{d}s \\
&\leq L \int_0^{\tilde{\delta}} \|\varphi - \psi\|_\infty\; \mathrm{d}s \\
&= L\, \|\varphi - \psi\|_\infty\, \tilde{\delta} \\
&= \vartheta\, \|\varphi - \psi\|_\infty \; ,
\end{aligned}$$

für alle $\varphi, \psi \in A$. Nach dem Banach'schen Fixpunktsatz folgt, dass Φ genau einen Fixpunkt $\varphi \in A$ hat.

- **Schritt 2 (Existenz)**

  Wir hatten für jedes $\tilde{\delta} < \delta$ ein φ mit $\Phi_{\tilde{\delta}}[\varphi] = \varphi$ gefunden. Setze nun $x : (-\delta, \delta) \to G \, , \; x(t) = \varphi(t)$ wo für $t \in (-\delta, \delta)$ erst ein $\tilde{\delta} > 0$ mit $0 \leq |t| < \tilde{\delta} < \delta$ gewählt sei und dann φ der Fixpunkt von $\Phi_{\tilde{\delta}}$ aus Schritt 1 sei.

  Die Eindeutigkeit in Schritt 1 zeigt, dass $x(t)$ nicht von der Wahl von $\tilde{\delta}$ abhängt. Es folgt, dass x stetig ist und $\forall t \in (-\delta, \delta)$:

  $$x(t) = x_0 + \int_0^t f\big(x(s)\big)\, \mathrm{d}s \; .$$

  Daher ist x sogar stetig differenzierbar (weil die rechte Seite es ist) mit

  $$\dot{x} = f(x) \; ,$$

  also ist x Lösung des Anfangswertproblemes.

- **Schritt 3 (Eindeutigkeit)**

  Ist $x : (-\delta, \delta) \to G$ eine Lösung von $\dot{x} = f(x)$ mit $x(0) = x_0$, so ist für alle $t \in (-\delta, \delta)$

  $$x(t) = x_0 + \int_0^t f\big(\varphi(s)\big)\, \mathrm{d}s$$

  also für jedes $0 < \tilde{\delta} < \delta$ ist $x|_{[-\tilde{\delta}, \tilde{\delta}]}$ ein Fixpunkt der Abbildung $\Phi_{\tilde{\delta}}$. Daraus folgt, dass x eindeutig festgelegt ist auf $[-\tilde{\delta}, \tilde{\delta}]$, also auch auf $(-\delta, \delta)$.

QED.

7.4.3 Kommentar

(1) Man beachte, dass der Existenzsatz nur die Existenz der Lösung für „kurze Zeiten“ $(-\delta, \delta)$ liefert (*Kurzzeitexistenz*).

(2) Der Beweis gibt eine untere Schranke für die Lebensdauer der Lösung,

$$\delta(x_0) = \min\left\{\frac{1}{L}, \frac{r}{M}\right\} ,$$

wobei $r(x_0)$ so ist, dass $K := \overline{B_r}(x_0) \subseteq G$ ist, M eine Schranke für $\|f\|$ auf K und L eine Lipschitz-Konstante für f (z.B. eine Schranke auf $\|Df\|$) auf K.

Definition 7.10.
Sei $G \subseteq \mathbb{R}^n$ ein Gebiet und $f : G \to \mathbb{R}^n$ stetig differenzierbar. Eine Lösung $x : I \to G$ ($I \subseteq \mathbb{R}$ ein Intervall) von $\dot{x} = f(x)$ auf G heißt *maximal*, wenn gilt: Ist $\tilde{I} \subseteq \mathbb{R}$ ein Intervall mit $\tilde{I} \supseteq I$ und $\tilde{x} : \tilde{I} \to G$ eine Lösung von $\dot{x} = f(x)$ und $\tilde{x}|_I = x$, so gilt bereits $\tilde{I} = I$ (und $\tilde{x} = x$).

Proposition 7.11.
Sei $G \subseteq \mathbb{R}^n$ ein Gebiet, ein f ein stetig differenzierbares Vektorfeld und seien $x : I \to G$ und $y : J \to G$, $I, J \subseteq \mathbb{R}$ offene Intervalle mit $0 \in I \cap J$, Lösungen von $\dot{x} = f(x)$ mit Anfangswert $x(0) = y(0) = x_0 \in G$. Dann gilt für alle $t \in I \cap J$:

$$x(t) = y(t) \; .$$

¡? Beweis ¿!

Sei $A = \{t \in I \cap J | x(t) = y(t)\}$. Wir behaupten, dass $A = I \cap J$. Sei dazu $I = (t_-, t_+)$, $J = (s_-, s_+)$ und **o.B.d.A.** $t_+ \leq s_+$. Ähnlich wie in (Korollar 4.15.) setzen wir

$$t_0 = \sup\left\{\tau \in (0, t_+) \middle| \; \forall 0 \leq t < \tau \; \middle| \; x(t) = y(t)\right\} \; .$$

Nach dem Eindeutigkeitssatz existiert ein $\delta > 0$, so dass $x(t) = y(t) \;\; \forall \; 0 \leq t < \delta$. Also ist $t_+ \geq t_0 \geq \delta > 0$. Angenommen $t_0 < t_+$. Aus Stetigkeitsgründen folgt dann $x(t_0) = y(t_0) =: \tilde{x_0}$. Mit dem Eindeutigkeitssatz folgt weiter: Es existiert ein $\delta_{\tilde{x_0}} > 0$, so dass $x(t) = y(t) \;\; \forall t \in (t_0 - \delta_{\tilde{x_0}}, t_0 + \delta_{\tilde{x_0}})$
$\lightning$ Widerspruch $\lightning$
Daraus folgt $t_0 = t_+$. Argumentiere ähnlich für t_-, also insgesamt $A = I \cap J$.
QED.

Satz 7.12.
Sei $G \subseteq \mathbb{R}^n$ ein Gebiet, $f : G \to \mathbb{R}^n$ stetig differenzierbar, $x_0 \in G$. Dann existiert genau eine maximale Lösung $x : I \to G$ von $\dot{x} = f(x)$ mit $x(0) = x_0$.

03.02.2005

¡? Beweis ¿!

Ist $y : J \to G$ Lösung von $\dot{x} = f(x)$ mit $y(0) = x_0$ $(0 \in J)$, so notieren wir diese mit (y, J). Sei nun

$$I := \bigcup_{\substack{(y,J) \\ \text{ist Lösung}}} J \subseteq \mathbb{R} \; .$$

Dann ist $I \subseteq \mathbb{R}$ offenes Intervall und ist $t \in I$ beliebig, so wählen wir (y, J) so, dass $t \in J$ ist und setzen $x(t) = y(t)$, $x : I \to G$. Wegen Proposition 7.11. hängt das nicht von der Wahl von (y, J) ab. Es ist dann x offenbar Lösung des Anfangswertproblemes, die maximal ist und nach Konstruktion und Proposition 7.11. die einzige maximale Lösung.
QED.

7.4.4 Kommentar

(1) Wir schreiben für das Definitionsintervall I der maximalen Lösung

$$I(x_0) = \big(t_-(x_0), t_+(x_0)\big) \ ,$$

wobei $t_-(x_0) \in [-\infty, 0)$ und $t_+(x_0) \in (0, \infty]$ ist.

(2) Wir setzen nun weiter

$$\Omega = \big\{(t,x) \in \mathbb{R} \times G \big| t \in I(x)\big\}$$

und

$$\varphi : \Omega \to G \ \ , \ (t,x) \mapsto \varphi^t(x) = x(t) \ ,$$

wobei $x(.)$ die maximale Lösung zum Anfangswertproblem mit Anfangswert x ist.

(3) Man kann nun zeigen (siehe z.B. `Hirsch, Smale: Differential Equations, dynamical systems and linear algebra`), dass Ω offen ist und φ stetig differenzierbar (Nachweis stetig: einfach; stetig differenzierbar: schwer).

(4) φ definiert ein dynamisches System auf G (Definition 7.1. (b)).

Behauptung

$$\begin{array}{rcl} \varphi^0(x) & = & x \\ s \in I\left(\varphi^t(x)\right) & \Leftrightarrow & s+t \in I(x) \\ \varphi^{s+t}(x) & = & \varphi^s\left(\varphi^t(x)\right) \quad \forall\, t \in I(x)\ \forall\, s \in I\big(\varphi^t(x)\big) \end{array}$$

¡? Beweis ¿!

Offenbar ist $\varphi^0(x) = x \ \ \forall x \in G$ und ist $x \in G$ und $t \in I(x)$, so lösen

$$\begin{array}{rlll} & \Phi : & \big(t_-(x) - t, t_+(x) - t\big) \to G, & \Phi(s) = \varphi^{s+t}(x) \ , \\ \text{und} & \Psi : & \Big(t_-\big(\varphi^t(x)\big), t_+\big(\varphi^t(x)\big)\Big) \to G, & \Psi(s) = \varphi^s\Big(\varphi^t(x)\Big) \ , \end{array}$$

die Differenzialgleichung $\dot{x} = f(x)$ zum Anfangswert $\varphi^t(x) : \Phi(0) = \varphi^t(x) = \Psi(0)$,

$$\begin{array}{rcl} \Phi'(s) & = & \dfrac{\mathrm{d}}{\mathrm{d}\sigma}\Big|_{\sigma=s+t} \varphi^\sigma(x) = f\big(\varphi^{s+t}(x)\big) = f\big(\Phi(s)\big) \ , \\ \Psi'(s) & = & \dfrac{\mathrm{d}}{\mathrm{d}\sigma}\Big|_{\sigma=s} \varphi^\sigma\left(\varphi^t(x)\right) = f\Big(\varphi^s\left(\varphi^t(x)\right)\Big) = f\big(\Psi(s)\big) \ . \end{array}$$

Außerdem sind beide Lösungen maximal und daher gilt für alle $t \in I(x)$:

$$s \in I\big(\varphi^t(x)\big) \ \Leftrightarrow \ \ s+t \in I(x)$$

und für diese gilt:

$$\Phi(s) = \Psi(s) \ ,$$

also $\varphi^s\big(\varphi^t(x)\big) = \varphi^{s+t}(x)$ für alle $s \in I\big(\varphi^t(x)\big)$.
QED.

(5) Ein dynamisches System $\varphi : \Omega \to G$ heißt *maximal*, wenn gilt: Ist $\tilde{\varphi} : \tilde{\Omega} \to G$ ein dynamisches System mit $\tilde{\Omega} \supseteq \Omega$ und $\tilde{\varphi}|_\Omega = \varphi$, so muss bereits $\tilde{\Omega} = \Omega$ sein (und damit $\tilde{\varphi} = \varphi$).

(6) Das dynamische System $\varphi : \Omega \to G$, das wir in (1) und (3) zu einem Vektorfeld $f : G \to \mathbb{R}^n$ konstruiert haben, ist offenbar maximal. Sein *assoziiertes Vektorfeld* (nach 7.1.2) ist offenbar gerade f.

(7) Ist umgekehrt $\varphi : \Omega \to G$ ein maximales dynamisches System auf G und $f : G \to \mathbb{R}^n$ das assoziierte Vektorfeld, so ist da zu f gehörende dynamische System nach (2) und (3) gerade wieder φ.

Wir habe damit gezeigt (bis auf die Lücke (3)):

Theorem 7.13.
Sei $G \subseteq \mathbb{R}^n$ ein Gebiet. Es gibt dann eine 1:1-Beziehung zwischen den maximalen dynamischen Systemen φ auf G und den stetig differenzierbaren Vektorfeldern f auf G, die jedem φ sein assoziiertes Vektorfeld bzw. jedem f die Gesamtheit der maximalen Lösungen von $\dot{x} = f(x)$ zuordnet.

$$\begin{array}{rcl} \{\text{maximale dynamische Systeme } \varphi \text{ auf } G\} & \overset{1:1}{\leftrightarrow} & \{\text{stetig differenzierbare Vektorfelder auf } G\} \\ \varphi & \mapsto & f = \frac{\mathrm{d}}{\mathrm{d}t}\Big|_{t=0}(\varphi^t) \\ \text{maximale Lösungen von } \dot{x} = f(x) & \leftarrow & f \end{array}$$

7.4.5 Kommentar

Der Übergang $\varphi \mapsto f$ ist einfach („Differenzieren kann jeder"), der Übergang $f \mapsto \varphi$ schwer („Integrieren ist eine Kunst"). Im allgemeinen weiss man nicht einmal, ob bei gegebenen $f : G \to \mathbb{R}^n$ und $x \in G$ das Ende $t_+(x) \in (0, \infty]$ endlich oder unendlich ist (z.B. beim 3-Körper-Problem). Der folgende Satz besagt aber immerhin, dass die Bahnkurve $t \mapsto x(t)$ eines Punktes $x \in G$ mit $t_+(x) < \infty$ jedes Kompaktum K in G verlassen muss, wenn $t \longrightarrow t_+(x)$ geht, d.h. $t \mapsto x(t)$ strebt zum Rand von G oder (betragsweise) nach Unendlich, für $t \longrightarrow t_+(x)$. Sozusagen: $t \mapsto x(t)$ kann sich in endlicher Zeit „nicht einfach in Luft auflösen".

Satz 7.14.
Sei $G \subseteq \mathbb{R}^n$ ein Gebiet, f ein stetig differenzierbares Vektorfeld auf G und $x \in G$ derart, dass $t_+(x) < \infty$. Ist $K \subseteq G$ ein Kompaktum, so gibt es ein $0 < \tau < t_+(x)$, so dass für alle $\tau < t < t_+(x)$ gilt: $x(t) \notin K$.

¡? Beweis ¿!

Die Abbildung

$$\begin{array}{rcl} \mathrm{dist}(\,.\,, \partial G) & : & G \to (0, \infty)\ , \qquad (7.5) \\ \mathrm{dist}(x, \partial G) & = & \inf\{\|x - y\|\ |y \in \partial G\}\ , \qquad (7.6) \end{array}$$

ist stetig. Da K kompakt, nimmt $h := \mathrm{dist}(., \partial G)|_K$ nach Weierstraß (Satz 3.9.) ihr Minimum an, sagen wir mit $\rho > 0$. Es ist also mit $r := \frac{\rho}{2} > 0$,

$$\overline{B_r}(x) \subseteq G \quad \forall x \in K\ .$$

Seien weiter $M > 0$ und $L > 0$ Schranken für $\|f\|$ bzw. $\|Df\|$ auf $\overline{B_r}(k) = \bigcup_{x \in K} \overline{B_r}(x)$ (erneut Weierstraß). Dann besagt Kommentar (7.4.3): Ist $t \in (0, t_+(x))$ mit $x(t) \in K$, so ist $t_+(x) = t_+\big(x(t)\big) \geq \delta + t$ mit $\delta = \min\left\{\frac{1}{L}, \frac{r}{m}\right\}$. Für alle $t \in \big(\tau, t_+(x)\big)$ mit $\tau := t_+(x) - \delta$ muss also gelten: $x(t) \notin K$.

QED.

7.4.6 Kommentar

(1) Eine entsprechende Aussage gilt, wenn $t_-(x) > -\infty$ ist.

(2) Bleibt eine Lösungskurve $t \mapsto x(t)$ in K, z.B. wenn $\lim\limits_{t \longrightarrow t_+(x)} x(t) = p$ ist, für ein $p \in G$, so muss $t_+(x) = \infty$ sein (vgl. 7.3.1 (3)). In diesem Fall muss p dann eine Gleichgewichtslage sein. ✎ *Übung* ✐

Kapitel 8

Lineare Systeme

03.02.2005

8.1 Lineare Systeme

Definition 8.1.
Sei $I \subseteq \mathbb{R}$ ein offenes Intervall und $A : I \to \mathrm{Mat}_n(\mathbb{R})$ eine stetig differenzierbare Abbildung.

(a) Man nennt die Differenzialgleichung

$$\dot{x} = A(t)x$$

ein (nicht-autonomes) *homogenes lineares System.*

(b) Ist $b : I \to \mathbb{R}^n$ stetig differenzierbar, so heißt

$$\dot{x} = A(t)x + b(t)$$

ein (nicht autonomes) *inhomogenes lineares System.*

8.1.1 Kommentar

(1) Nach 7.2.6 kann man aus dem nicht-autonomen System $\dot{x} = A(t)x + b(t)$ auf $\mathbb{R}^n$ durch

$$\begin{aligned} \dot{s} &= 1 \\ \dot{x} &= A(s)x + b(s) \end{aligned} \tag{8.1}$$

ein autonomes System auf

$$G = I \times \mathbb{R}^n \subseteq \mathbb{R}^{n+1}$$

machen. Zu $(t_0, x_0) \in G$ gibt es dann nach Satz 7.12. eine maximale Lösung $x : J \to \mathbb{R}^n$ von Gleichung (8.1) mit dem Anfangswert $\varphi(t_0) = (t_0, x_0)$. Es muss dann $\varphi(t) = \big(t, x(t)\big)$ und damit $J \subseteq I$ sein. Es gibt also zu $(t_0, x_0) \in I \times \mathbb{R}^n$ genau eine maximale Lösung $x : J \to \mathbb{R}^n$ von $\dot{x} = A(t)x + b(t)$ mit $J \subseteq I$ (und $t_0 \in J$).

(2) Hier ist es unzweckmäßig von $\dot{x} = A(t)x + b(t)$ auf Gleichung (8.1) zu wechseln, weil dadurch die (affin-) lineare Struktur der rechten Seite verloren geht.

(3) Wir sprechen im folgenden von *globaler Existenz*, wenn die maximale Lösung auf $J = I$ gegeben ist.

08.02.2005

Lemma 8.2. Gronwall-Lemma
Sei $0 < b$ und $u : [0,b] \to [0,\infty)$ eine stetige Funktion. Es gebe weiterhin Konstanten $L, C \geq 0$, so dass für alle $t \in [0,b]$ gilt:

$$u(t) \leq C + L \int_0^t u(s)\ \mathrm{d}s\ .$$

Dann gilt für alle $t \in [0,b]$

$$u(t) \leq Ce^{Lt}\ .$$

8.1.2 Kommentar

Man beachte, dass $x : \mathbb{R} \to \mathbb{R}^n,\ t \mapsto Ce^{Lt}$ die Lösung von $\dot{x} = Lx$ auf $\mathbb{R}^n$ zum Anfang $x(0) = C$ ist. Wäre u stetig differenzierbar mit $u(0) \leq C$ und $\dot{u} \leq Lu$, so würde Lemma 8.2. gerade ausdrücken, dass $u(t) \leq x(t)$ bleibt, für alle $t \in [0,b]$, denn aus $u(0) \leq C$ und $\dot{u} \leq Lu$ folgt durch Integration:

$$u(t) = \underbrace{u(0)}_{\leq C} + \int_0^t \underbrace{\dot{u}(s)}_{\leq Lu(s)}\ \mathrm{d}s \quad \leq \quad C + \int_0^t Lu(s)\ \mathrm{d}s\ .$$

¡? Beweis ¿!

- 1.Fall $C > 0$

  Setze $U : [0,b] \to [0,\infty)$,

  $$U(t) := C + L \int_0^t u(s)\ \mathrm{d}s\ .$$

  Dann ist U stetig differenzierbar und $\dot{U}(t) = Lu(t) \geq 0\,,\ \forall t \in [0,b]$, also ist U monoton steigend. Außerdem ist $U(0) = C > 0$. Also ist $U(t) \geq C\ \ ,\ \forall t \in [0,b]$. Es folgt

  $$\frac{\mathrm{d}}{\mathrm{d}t} \ln U(t) = \frac{\dot{U}(t)}{U(t)} = \frac{Lu(t)}{U(t)} \leq L\ ,$$

  weil $u \leq U$ nach Varaussetzung. Also ist

  $$\begin{aligned} \ln U(t) - \ln C \quad &= \quad \ln U(t) - \ln U(0) = \int_0^t \frac{\mathrm{d}}{\mathrm{d}s} (\ln U(s))\ \mathrm{d}s \\ &\leq \quad \int_0^t L\ \mathrm{d}s = Lt \end{aligned}$$

  und damit

  $$\frac{U(t)}{C} = exp\left(\ln \frac{U(t)}{C}\right) \overset{\substack{\text{exp monoton}\\ \text{wachsend}}}{\leq} exp(Lt) \quad \Rightarrow \quad u(t) \leq U(t) \leq Ce^{Lt}$$

- 2.Fall $C = 0$

  Wähle eine Folge (c_n) in $\mathbb{R}$ mit $c_n > 0$ und $(c_n) \searrow 0$. Der 1.Fall liefert dann:

  $$u(t) \leq c_n e^{Lt}\ \ ,\ \forall t \in [0,b]\ \ ,\ \forall n \in \mathbb{N}$$

  und damit auch

  $$u(t) \leq \lim_{n\to\infty} c_n e^{Lt} = 0\ \ ,\ \forall t \in [0,b]\ .$$

  Es ist also $u = 0$.

QED.

Satz 8.3. Langzeitexistenz
Sei $I \subseteq \mathbb{R}$ offen, $A : I \to \mathrm{Mat}_n(\mathbb{R})$ und $b : I \to \mathbb{R}^n$ seien stetig differenzierbar. Ist dann $t_0 \in I$, $x_0 \in \mathbb{R}^n$ und $x : J \to \mathbb{R}^n$ die maximale Lösung von $\dot{x} = A(t)x + b(t)$ mit $x(t_0) = x_0$ (und $J \subseteq I, t_0 \in J$), so gilt bereits:

$$J = I \ .$$

¡? Beweis ¿!

Sei $I = (t_-, t_+)$ und $J = (s_-, s_+)$ mit $-\infty \leq t_- \leq s_- < t_0 < s_+ \leq t_+ \leq +\infty$. Wir zeigen $s_+ = t_+$ (analog sieht man $s_- = t_-$, also $J = I$). Angenommen $s_+ < t_+$.

O.B.d.A. sei $t_0 = 0$ (sonst verschiebe alles um t_0). Ist $J \to \mathbb{R}^n,\ t \mapsto x(t)$ die maximale Lösung von $\dot{x} = A(t)x + b(t)$ mit $x = x_0$, so ist $J \to I \times \mathbb{R}^n, t \mapsto \big(t, x(t)\big)$ maximale Lösung von

$$\begin{aligned} \dot{s} &= 1 \\ \dot{x} &= A(s)x + b(s) \end{aligned}$$

mit Anfang

$$\begin{aligned} s(t_0) &= t_0 \\ x(t_0) &= x_0 \ . \end{aligned}$$

Wegen $s_+ < t_+$ (insbesondere $s_+ < \infty$), muss wegen Satz 7.14. die Kurve $[0, s_+) \to I \times \mathbb{R}^n,\ t \mapsto \big(t, x(t)\big)$ jedes Kompaktum von $I \times \mathbb{R}^n$ verlassen, also $\|x(t)\| \longrightarrow \infty$, für $t \mapsto s_+$.
Andererseits: Ist

$$L := \max\{\|A(t)\| \,|t \in [0, s_+]\} < \infty$$

und

$$M := \max\{\|b(t)\| \,|t \in [0, s_+]\} < \infty \ ,$$

so gilt für die stetige Funktion

$$u : [0, s_+ - \varepsilon] \to [0, \infty) \quad u(t) := \|x(t)\|$$

($\varepsilon > 0$), dass

$$\begin{aligned} u(t) &= \|x(t)\| = \left\| x(0) + \int_0^t x'(s)\ \mathrm{d}s \right\| \\ &\leq \|\underbrace{x(0)}_{=x_0}\| + \int_0^t \|x'(s)\|\ \mathrm{d}s \\ &= u(0) + \int_0^t \|A(s)x(s) + b(s)\|\ \mathrm{d}s \\ &\leq u(0) + \int_0^t \left(\underbrace{\|A(s)\|}_{\leq L} \underbrace{\|x(s)\|}_{=u(s)} + \|b(s)\| \right) \mathrm{d}s \\ &= \underbrace{u(0) + \int_0^{s_+} \underbrace{\|b(s)\|}_{\leq M}\ \mathrm{d}s}_{\leq u(0) + s_+ M =: C} + L \int_0^t u(s)\ \mathrm{d}s \leq C + L \int_0^t u(s)\ \mathrm{d}s \ . \end{aligned}$$

Also ist nach dem Lemma von Gronwall

$$u(t) \leq Ce^{Lt} \leq Ce^{Ls_+}$$

für alle $t \in [0, s_+ - \varepsilon]$ (für alle $0 < \varepsilon < s_+$), also für alle $t \in [0, s_+)$. Es ist also $t \mapsto \|x(t)\|$ beschränkt für $t \to s_+$.
↯ Widerspruch ↯
QED.

8.1.3 Kommentar

Sei $t_0 \in I$ fest. Definiere nun (vgl. 7.1.1) für jedes $t \in I$ die Abbildung

$$\varphi^t : \mathbb{R}^n \to \mathbb{R}^n \quad \varphi^t(x) = x(t) \ ,$$

$I \to \mathbb{R}^n$ wo $t \mapsto x(t)$ die (maximale) Lösungekurve von $\dot{x} = A(t)x + b(t)$ mit $x(t_0) = x$ ist. Man beachte, dass hier wegen Satz 8.3. die Abbildungen φ^t allesamt den gleichen Definitionsbereich haben (i.a. hängt dieser von t ab).

Es ist allerdings

$$\varphi : I \to \mathcal{C}^1(\mathbb{R}^n, \mathbb{R}^n) \ , \ t \mapsto \varphi^t$$

wegen der Nicht-Autonomie von $\dot{x} = A(t)x + b(t)$ i.a. kein Fluss mehr im bisherigen Sinne, weil i.a. nicht $\varphi^t \circ \varphi^s = \varphi^{t+s}$ gilt (für $t, s \in I$, so dass $t + s \in I$ ist).

Satz 8.4.
Sei $I \subseteq \mathbb{R}$ offen und $A : I \to \mathrm{Mat}_n(\mathbb{R})$ stetig differenzierbar. Sei $L_{(h)} \subseteq \mathcal{C}^1(I, \mathbb{R}^n)$ die Menge aller maximalen Lösungen des homogenen linearen Systems

$$\dot{x} = A(t)x \tag{8.2}$$

auf $\mathbb{R}^n$. Dann gilt

(a) $L_{(h)}$ ist n-dimensionaler Unterraum von $\mathcal{C}^1(I, \mathbb{R}^n)$.

(b) Ist $t_0 \in I$, $r \in \mathbb{N}$ und $\xi_1, \ldots, \xi_r \in \mathbb{R}^n$, so seien $x_1, \ldots, x_r \in \mathcal{C}(I, \mathbb{R}^n)$ die Lösungen mit $x_j(t_0) = \xi_j$ $(j = 1, \ldots, r)$. Dann ist äquivalent:

(i) $(x_1, \ldots, x_r)$ sind linear unabhängig in $L_{(h)}$,

(ii) $\big(x_1(t), \ldots, x_r(t)\big)$ sind linear unabhängig in $\mathbb{R}^n$ für alle $t \in I$,

(iii) $(\xi_1, \ldots, \xi_r)$ sind linear unabhängig in $\mathbb{R}^n$.

¡? Beweis ¿!

(a) Seien x, y Lösungen von Gleichung (8.2), $x, y \in L_{(h)}$. Also ist $\dot{x} = A(t)x(t)$ und $y(t) = A(t)y(t)$ für alle $t \in I$. Dann gilt $forallt \in I$:

$$(x+y)\dot{}(t) = \dot{x}(t) + \dot{y}(t) = A(t)x(t) + A(t)y(t) \overset{!}{=} A(t)\left(x(t) + y(t)\right) = A(t)\left(x + y\right)(t) \ ,$$

also ist auch $x + y \in L_{(h)}$. Ähnlich ist für $x \in L_{(h)}$, $\lambda \in \mathbb{R}$

$$(\lambda x)\dot{} = \lambda \dot{x} = \lambda A(t)x = A(t)(\lambda x) \ ,$$

also auch $\lambda x \in L_{(h)}$. Damit ist $L_{(h)} \subseteq \mathcal{C}^1(I, \mathbb{R}^n)$ ein Unterraum.

(b) **(bi)⇒(bii)** Sei $r \in \mathbb{N}$ und $(x_1, \ldots, x_r)$ in $L_{(h)}$ linear unabhängig. Sei $t_1 \in I$ beliebig. und $\lambda_1, \ldots, \lambda_r \in \mathbb{R}$ gegeben mit

$$\lambda_1 x_1(t_1) + \cdots + \lambda_r x_r(t_1) = 0 \text{ (in } \mathbb{R}) \ . \tag{8.3}$$

Wir setzen $x : I \to \mathbb{R}^n$,

$$x(t) = \lambda_1 x_1(t) + \cdots + \lambda_r x_r(t) \ .$$

Nach (a) ist dann x Lösung von Gleichung (8.2), $x \in L_{(h)}$. Es ist $x(t_1) = 0$ nach Gleichung (8.3).

Aber $y : I \to \mathbb{R}^n$, $y(t) = 0$ $\forall t \in I$ ist auch Lösung von Gleichung (8.3) und $y(t_1) = 0$. Dann ist nach dem Eindeutigkeitssatz $x(t) = y(t)$ $\forall t \in I$, also $x = y = 0$. Also ist

$$\lambda_1 x_1 + \cdots + \lambda_r x_r = 0 \ (\text{ in } \mathcal{C}^1(I, \mathbb{R}^n) \ ) \ .$$

10.02.2005

Wegen der linearen Unabhängigkeit von $(x_1, \ldots, x_r)$ folgt daher: $\lambda_1 = \cdots = \lambda_r = 0$. Also ist $\big(x_1(t_1), \ldots, x_r(t_1)\big)$ linear unabhängig in $\mathbb{R}^n$.

(bii)⇒(biii) $t = t_0$.

(biii)⇒(bi) Klar, denn ist $\lambda_1 x_1, \dots, \lambda_r x_r = 0$ in $\mathcal{C}^1(I, \mathbb{R}^n)$, so gilt insbesondere bei $t = t_0$

$$0 = \lambda_1 x_1(t_0) + \cdots + \lambda_r x_r(t_0) = \lambda_1 \xi_1 + \cdots + \lambda_r \xi_r \ .$$

Ist also $(\xi_1, \dots, \xi_r)$ in $\mathbb{R}^n$ linear unabhängig und ist $x_j \in L_{(h)}$ die Lösung von Gleichung (8.2) mit $x_j(t_0) = \xi$ $(j = 1, \dots, r)$, so ist auch $(x_1, \dots, x_r)$ in $L_{(h)}$ linear unabhängig.

Schließlich: Wähle $(\xi_1, \dots, \xi_r)$ linear unabhängig im $\mathbb{R}^n$. Sind $x_1, \dots, x_n \in L_{(h)}$ mit $x_j(t_0) = \xi_j$, so folgt $(x_1, \dots, x_n)$ ist linear unabhängig in $L_{(h)}$, also $\dim L_{(h)} \geq n$. Anderseits: Ist $(x_1, \dots, x_r)$ in $L_{(h)}$ linear unabhängig, so ist $(x_1(t_0), \dots, x_r(t_0))$ linear unabhängig (nach (b)), also $r \leq \dim \mathbb{R}^n = n$. Insgesamt ist damit

$$\dim L_{(h)} = n \ .$$

QED.

Korollar 8.5.
Ist $A : I \to \mathrm{Mat}_n(\mathbb{R})$ stetig differenzierbar und $t_0 \in I$, so sei $\varphi^t : \mathbb{R}^n \to \mathbb{R}^n$ gegeben durch $\varphi^t(x) = x(t)$, wobei $t \mapsto x(t)$ die Lösung von $\dot{x} = A(t)x$ mit Anfang $x(t_0) = x$ sei. Dann ist φ^t für jedes $t \in I$ ein linearer Isomorphismus des $\mathbb{R}^n$, $\varphi^t \in Aut(\mathbb{R}^n)$

¡? Beweis ¿!

Sind $x, y : I \to \mathbb{R}^n$ Lösungen von $\dot{x} = A(t)x$ mit Anfang $x(t_0) = x_0$ und $y(t_0) = y_0$, so ist $z := x + y : I \to \mathbb{R}^n$ ebenfalls Lösung mit $z(t_0) = x_0 + y_0$. Das zeigt:

$$\varphi^t(x_0 + y_0) = z(t) = x(t) + y(t) = \varphi^t(x_0) + \varphi^t(y_0) \ .$$

Ähnlich sieht man

$$\varphi^t(\lambda x_0) = \lambda \varphi^t(x_0) \quad \forall \lambda \in \mathbb{R}\,, \forall x_0, y_0 \in \mathbb{R}^n\ , \ \forall t \in I \ .$$

Also ist φ^t linear. φ^t ist sogar Isomorphismus, denn ist $(\xi_1, \dots, \xi_n)$ Basis von $\mathbb{R}^n$, so ist auch $\big(\varphi^t(\xi_1), \dots, \varphi^t(\xi_n)\big)$ Basis von $\mathbb{R}^n$, weil mit $\varphi^t(\xi_j) = x_j(t)$, x_j Lösung mit $x_j(t_0) = \xi_j$, linear unabhängig bleibt.
QED.

8.2 Fundamental – Lösungen

Definition 8.6.
Ist $A : I \to \mathrm{Mat}_n(\mathbb{R})$, so nennt man ein n-Tupel $(x_1, \dots, x_n)$ in $\mathcal{C}^1(I, \mathbb{R}^n)$ ein *Lösungs-Fundamental-System* von $\dot{x} = A(t)x$, wenn $(x_1, \dots, x_n)$ eine Basis von

$$L_{(h)} = \big\{x \in \mathcal{C}^1(I, \mathbb{R}^n) \big|\ \dot{x} = A(t)x\big\}$$

ist.

8.2.1 Kommentar

(1) Für die Gesamtheit aller Lösungen von $\dot{x} = A(t)x$ braucht man also nur n spezielle Lösungen $x_1, \dots, x_n$ zu kennen, nämlich solche, die zu einem (beliebigen) Zeitpunkt $t_0 \in I$ linear unabhängig sind: $(x_1(t_0), \dots, x_n(t_0))$ ist eine Basis des $\mathbb{R}^n$.

(2) Nimmt man speziell die Lösungen $(x_1, \dots, x_n)$, die zur Zeit $t_0 \in I$ auf der kanonischen Basis $(\hat{e}_1, \dots, \hat{e}_n)$ im $\mathbb{R}^n$ sitzen, $x_j(t_0) = \hat{e}_j$ $(j = 1, \dots, n)$, und schreibt man die Koordinatenfunktion φ_{ij} von x_j, also $x_j = \sum\limits_{i=1}^{n} \varphi_{ij} \hat{e}_i$, in die j-te Spalte einer Matrix $\Phi(t) = \big(\varphi_{ij}(t)\big) \in \mathrm{Mat}_n(\mathbb{R})$, so

beschreibt $\Phi(t)$ gerade den Automorphismus $\varphi^t : \mathbb{R}^n \to \mathbb{R}^n$ aus Korollar 8.5. bzgl. der Basis $(\hat{e}_1, \ldots, \hat{e}_n)$, denn ist $\xi \in \mathbb{R}^n$, $\xi = \sum_{i=1}^{n} \xi_i \hat{e}_i$, so ist:

$$\varphi^t(\xi) = \varphi^t\left(\sum_{i=1}^{n} \xi_i \hat{e}_i\right) = \sum_{i=1}^{n} \xi_i \varphi^t(\hat{e}_i) = \sum_{i=1}^{n} \xi_i \big(x_i(t)\big) = \Phi(t)\xi \ .$$

(3) Die Kurve $I \to Gl_n(\mathbb{R})$, $t \to \Phi(t)$ löst selber Lösung einer Differenzialgleichung auf $\mathrm{Mat}_n(\mathbb{R})$ $(\cong \mathbb{R}^{n^2})$, nämlich

$$\dot{\Phi} = A(t)\Phi \ ,$$

und zwar zum Anfangswert $\Phi(t_0) = \mathbb{1}$, denn:

$$\dot{\Phi}(t)x = \Big(\Phi(t)x\Big)^{\cdot} = \Big(\varphi^t(x)\Big)^{\cdot} = \dot{x}(t) = A(t)x(t) = A(t)\Big(\varphi^t(x)\Big) = A(t)\Big(\Phi(t)\Big)x \ ,$$

für alle $x \in \mathbb{R}^n$, also $\dot{\Phi} = A(t)\Phi$. Kennt man die <u>eine</u> Lösung von $\dot{\Phi} = A(t)\Phi$ zum Anfangswert $\Phi(t_0) = \mathbb{1}$, so kennt man <u>alle</u> Lösungen von $\dot{x} = A(t)x$ auf $\mathbb{R}^n$, denn ist $x(t_0) = x_0 \in \mathbb{R}^n$, so ist offenbar

$$x(t) = \Phi(t)x_0 \ .$$

(Denn: $\dot{x}(t) = \dot{\Phi}(t)x_0 = \big(A(t)\Phi(t)\big)x_0 = A(t)x(t)$ und $x(t_0) = \underbrace{\Phi(t_0)}_{\mathbb{1}} x_0 = x_0$.)

Satz 8.7.
Sei $I \subseteq \mathbb{R}$ offen, $A : I \to \mathrm{Mat}_n(\mathbb{R})$ und $b : I \to \mathbb{R}^n$ seien stetig differenzierbar. Mit $L_{(i)}$ werde die Teilmenge von $\mathcal{C}^1(I, \mathbb{R}^n)$ aller Lösungen der inhomogenen Gleichung

$$\dot{x} = A(t)x + b(t)$$

bezeichnet,

$$L_{(i)} = \left\{x \in \mathcal{C}^1(I, \mathbb{R}^n) \middle| \ x \text{ Lösung von } \dot{x} = A(t)x + b(t)\right\} \ .$$

Ist $L_{(h)} \subseteq \mathcal{C}(I, \mathbb{R}^n)$ der Lösungsraum der Lösungen von $\dot{x} = A(t)x$, so gilt:

(a) Ist $\psi : I \to \mathbb{R}^n$ eine spezielle Lösung der inhomogenen Gleichung, so gilt:

$$L_{(i)} = \psi + L_{(h)} = \left\{\psi + x \middle| x \in L_{(h)}\right\} \subseteq \mathcal{C}^1(I, \mathbb{R}^n) \ .$$

(b) *Variation der Konstanten*
Sei $(\varphi_1, \ldots, \varphi_n)$ in $L_{(h)}$ Fundamentalsystem von $\dot{x} = A(t)x$ und $\Phi : I \to Gl_n(\mathbb{R})$, $\Phi(t) = \big(\varphi_1(t), \ldots, \varphi_n(t)\big)$. Setzt man nun $u : I \to \mathbb{R}^n$ $\ u(t) = \int_{t_0}^{t} \Phi^{-1}(s)b(s)\, \mathrm{d}s$ mit einem $t_0 \in I$, so ist $\psi : I \to \mathbb{R}^n$

$$\psi(t) = \Phi(t)u(t)$$

eine spezielle Lösung von $\dot{x} = A(t)x + b(t)$.

<u>¡? Beweis ¿!</u>

(a) Ist ψ spezielle Lösung von $\dot{x} = A(t)x + b(t)$, so ist $\tilde{\psi}$ genau dann eine weitere Lösung, wenn $\varphi := \tilde{\psi} - \psi$ eine Lösung von $\dot{x} = A(t)x$ ist:

$$\begin{array}{ccc} \dot{\tilde{\psi}} & = & A(t)\tilde{\psi} + b(t) \\ & \Updownarrow & \\ \dot{\varphi} = (\tilde{\psi} - \psi)^{\cdot} = \Big(A(t)\tilde{\psi} + b(t)\Big) & - & \Big(A(t)\psi + b(t)\Big) = A(t)(\tilde{\psi} - \psi) = A(t)\varphi \end{array}$$

(b) Sei also $\Phi = (\varphi_1, \dots, \varphi_n)$ ein Fundamentalsystem von $\dot{x} = A(t)x$ und $\psi(t) = \Phi(t)u(t)$ für eine stetig differenzierbare Abbildung $u : I \to \mathbb{R}^n$ Dann gilt:

$$\dot{\psi} = \dot{\Phi}u + \Phi\dot{u} = \Big(A(t)\Phi\Big)u + \Phi\dot{u} = A(t)\psi + \Phi\dot{u} \ ,$$

denn $\dot{\Phi} = A(t)\Phi$. Also ist ψ genau dann eine Lösung von $\dot{x} = A(t)x + b(t)$, wenn gilt:

$$\begin{aligned} & \Phi\dot{u} = b \\ \Leftrightarrow \quad & \dot{u} = \Phi^{-1}b \\ \Leftrightarrow \quad & u(t) = \int_{t_0}^{t} \Phi^{-1}(s)b(s)\ \mathrm{d}s + \mathrm{const}\,, \quad t_0 \in I \end{aligned}$$

QED.

8.2.2 Kommentar

(1) Die Menge aller Lösungen eines inhomogenen Systems $\dot{x} = A(t)x + b(t)$ auf $\mathbb{R}^n$ trägt also die Struktur eines *n-dimensionalen affinen Unterraums* von $\mathcal{C}^1(I, \mathbb{R}^n)$.

(2) Ist man in der Lage das homogene System $\dot{x} = A(t)x$ zu lösen (z.B. durch Lösen von $\dot{\Phi} = A(t)\Phi$ auf $\mathrm{Mat}_{\,n}(\mathbb{R})$ mit $\Phi(t_0) = \mathbb{1}$), so besagt Satz 8.7., dass man dann auch das inhomogene System $\dot{x} = A(t)x + b(t)$ vollständig lösen kann.

(3) Ist $(\varphi_1, \dots, \varphi_n)$ ein Fundamentalsystem für $\dot{x} = A(t)x$, so gilt mit dem $u = (u_1, \dots, u_n) : I \to \mathbb{R}^n$ aus Satz 8.7. (b), für die spezielle Lösung $\psi = \Phi u : I \to \mathbb{R}^n$, $\psi(t) = \sum\limits_{i=1}^{n} u_j(t)\varphi_j(t)$ oder kurz $\psi = \sum u_j\varphi_j$. Man betrachtet die Funktionen $u_j : I \to \mathbb{R}$ $(j = 1, \dots, n)$ als *„variierende Konstanten“*.

17.02.2005

8.2.3 Beispiel

Betrachte die Differenzialgleichung

$$\dot{x} = 2t \cdot x + t^3$$

auf $\mathbb{R}$. Mit $a(t) = 2t$ und $b(t) = t^3$ wird das zu $\dot{x} = a(t)x + b(t)$. Eine Lösung der homogenen Gleichung $\dot{x} = a(t)x$ ist mit der Methode der Trennung der Variablen (Proposition 7.6.):

$$\varphi(t) = \exp\left(\int_{t_0}^{t} a(s)\ \mathrm{d}s\right) \ ,$$

also hier zum Beispiel mit $t_0 = 0$:

$$\varphi(t) = \exp\left(\int_{0}^{t} 2s\ \mathrm{d}s\right) = \exp(t^2 - 0) = \exp\left(t^2\right) \ ,$$

was eine Basis von $L_{(h)}$ ist, da $\dim L_{(h)} = 1$ ist.

Die Methode der Variation der Konstanten liefert nun die spezielle Lösung: $\psi(t) = u(t) \cdot \varphi(t)$, wobei (mit Substitution $s^2 = \tau \ \Rightarrow \ \ 2s \,\mathrm{d}s = \,\mathrm{d}\tau$)

$$\begin{aligned}
u(t) &= \int_0^t \exp(-s^2)\cdot s^3 \,\mathrm{d}s \\
&= \int_0^{t^2} \exp(-\tau)\cdot \tau \cdot \frac{1}{2}\,\mathrm{d}\tau \\
&= \left[\frac{1}{2}\left[-\exp(-\tau)\tau\right]_0^{t^2} - \int_0^{t^2} \left(-\exp(-\tau)\cdot 1\right)\mathrm{d}\tau\right] \\
&= \frac{1}{2}\left[-\exp(-t^2)t^2 + \left(-\exp(-\tau)\Big|_0^{t^2}\right)\right] \\
&= \frac{1}{2}\left(-t^2\exp(-t^2) - \exp(-t^2) + 1\right) \\
&= \frac{1}{2} - \frac{1}{2}\left(1+t^2\right)\exp(-t^2)
\end{aligned}$$

zunächst nur für $t > 0$, dann aber auch für $t \leq 0$, also

$$\psi(t) \quad = \quad \frac{1}{2}\exp(t^2) - \frac{1}{2}(1+t^2)$$

Die allgemeine Lösung ist damit:

$$x(t) = \left(\frac{1}{2} + c\right)\exp(t^2) - \frac{1}{2}\left(1+t^2\right) \qquad (x(0) = c \in \mathbb{R}) \ .$$

8.2.4 Kommentar

Außer im Fall n=1 kann man das System $\dot{x} = A(t)\cdot x$ im Allgemeinen nicht durch *Quadratur* lösen. Es ergeben sich nämlich durchaus neue Funktionen (das heißt nicht-elementar in dem Sinn, dass sie nicht Verkettung bekannter Funktionen sind und auch nicht Verkettung ihrer Inversen oder Stammfunktionen), selbst wenn $a_{ij}(t)$ für $1 \leq i,j \leq n$ allesamt elementar sind.

8.2.5 Problem

Kann man eine *explizite Integration* von $\dot{x} = A(t)\cdot x$ wenigstens dann angeben, wenn A unabhängig von t ist, also $A \in \mathrm{Mat}_n(\mathbb{R})$ konstant,

$$\dot{x} = A\cdot x \quad ?$$

Beachte: In diesem Fall ist die Fundamentallösung $\Phi : \mathbb{R} \to \mathrm{Gl}_n(\mathbb{R})$ auf ganz $\mathbb{R}$ definiert und ist ein Fluss, also ein Homomorphismus von $(\mathbb{R},+)$ nach $(\mathrm{Gl}_n(\mathbb{R}),\cdot)$:

$$\Phi(t+s) = \Phi(t)\cdot\Phi(s)$$

für alle $t, s \in \mathbb{R}$.

8.2.6 Beispiel

Seien $\lambda_1,\ldots,\lambda_n \in \mathbb{R}$ und $A = \begin{pmatrix} \lambda_1 & \ldots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \ldots & \lambda_n \end{pmatrix} = \mathrm{diag}\Big(\lambda_1,\ldots,\lambda_n\Big)$. Dann wird $\dot{x} = Ax$ auf $\mathbb{R}^n$ zu einem *entkoppeltem System*

$$\begin{aligned}
\dot{x}_1 &= \lambda_1 x_1 \\
&\vdots \\
\dot{x}_n &= \lambda_n x_n
\end{aligned}$$

und hat das Fundamentalsystem

$$\Phi(t) = \begin{pmatrix} \exp(\lambda_1 t) & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \exp(\lambda_n t) \end{pmatrix} = \operatorname{diag}\Big(\exp(\lambda_1 t), \dots, \exp(\lambda_n t)\Big)$$

8.2.7 Bemerkung

Sei $A \in \mathrm{Mat}_n(\mathbb{R})$ und $S \in \mathrm{Gl}_n(\mathbb{R})$. Es ist $\varphi : \mathbb{R} \to \mathbb{R}^n$ genau dann Lösung von $\dot{x} = A \cdot x$, wenn $\psi : \mathbb{R} \to \mathbb{R}^n$, $\psi(t) = S^{-1} \cdot \varphi(t)$ Lösung von $\dot{y} = By$ ist, wo $B = S^{-1} \cdot A \cdot S$ ist.

¡? Beweis ¿!

Es ist mit $\varphi = S \cdot \psi$:

$$\dot{\varphi} = A \cdot \varphi \quad \Leftrightarrow \quad S \cdot \dot{\psi} = (S \cdot \psi)^{\cdot} = \dot{\varphi} = A \cdot \varphi = A \cdot S \cdot \psi \quad \Leftrightarrow \quad \dot{\psi} = S^{-1} \cdot A \cdot S \cdot \psi \quad \Leftrightarrow \quad \dot{\psi} = B \cdot \psi$$

QED.

Proposition 8.8.
Sei $A \in \mathrm{Mat}_n(\mathbb{R})$ eine (über $\mathbb{R}$) diagonalisierbare Matrix mit den Eigenwerten $\lambda_1, \dots, \lambda_n \in \mathbb{R}$ (mit aufgezählten Vielfachheiten). Sei $\Psi : \mathbb{R} \to \mathrm{Gl}_n(\mathbb{R})$, $\Psi(t) = \operatorname{diag}\big(\exp(\lambda_1 t), \dots, \exp(\lambda_n t)\big)$. Dann existiert ein $S \in \mathrm{Gl}_n(\mathbb{R})$, so dass $\Phi : \mathbb{R} \to \mathrm{Gl}_n(\mathbb{R})$, $\Phi(t) = S\Psi(t)$ ein Fundamentalsystem von $\dot{x} = Ax$ ist.

¡? Beweis ¿!

Diagonalisierbarkeit von A bedeutet: Es gibt eine Basis $(s_1, \dots, s_n)$ von $\mathbb{R}^n$ von Eigenvektoren von A, $A \cdot s_j = \lambda_j \cdot s_j$ für $\lambda_j \in \mathbb{R}$. Für $B = S^{-1} \cdot A \cdot S$ gilt dann $B = \operatorname{diag}\big(\lambda_1, \dots, \lambda_n\big)$ und $\dot{y} = B \cdot y$ hat nach 8.2.6 das Fundamentalsystem Ψ. Nach 8.2.7 ist dann $\Phi = S \cdot \psi$ Fundamentalsystem von $\dot{x} = A \cdot x$.
QED.

8.2.8 Kommentar

(1) Um die Normalformentheorie von Matrizen über $\mathbb{C}$ benutzen zu können, betrachtet man die Differenzialgleichung $\dot{x} = Ax$ statt auf $\mathbb{R}^n$ auf $\mathbb{C}^n$ und schreibt (auf $\mathbb{C}^n$)

$$\dot{z} = Az \; .$$

Die Menge $L_{(h)} \subseteq \mathcal{C}^1(I, \mathbb{C}^n)$ der Lösungen von $\dot{z} = Az$ bildet dann einen komplexen Unterraum der (komplexe) Dimension n in $\mathcal{C}(I, \mathbb{C}^n)$. Eine $\mathbb{C}$-Basis $\Phi = \big(\varphi_1, \dots, \varphi_n\big)$ von $L_{(h)}$ heißt dann ein *komplexes Fundamentalsystem* von $\dot{z} = Az$.

(2) Proposition 8.8. hat dann die folgende komplexe Version: Ist $A \in \mathrm{Mat}_n(\mathbb{C})$ über $\mathbb{C}$ diagonalisierbar mit Eigenwerten $\lambda_1, \dots, \lambda_n \in \mathbb{C}$ und $\Psi : \mathbb{R} \to \mathrm{Gl}_n(\mathbb{C})$, $\Psi(t) = \operatorname{diag}\big(\exp(\lambda_1 \cdot t), \dots, \exp(\lambda_n \cdot t)\big)$, so existiert eine Matrix $S \in \mathrm{Gl}_n(\mathbb{C})$, so dass $\Phi := S\psi : \mathbb{R} \to \mathrm{Gl}_n(\mathbb{C})$ ein komplexes Fundamentalsystem von $\dot{z} = Az$ ist.

8.2.9 Bemerkung

Ist $A \in \mathrm{Mat}_n(\mathbb{R})$ und $\big(\varphi_1, \dots, \varphi_n\big)$ ein komplexes Fundamentalsystem für $\dot{z} = Az$ auf $\mathbb{C}^n$, so ist $\big(\Re(\varphi_1), \Im(\varphi_1)\big), \dots, \big(\Re(\varphi_n), \Im(\varphi_n)\big)$ reelles Erzeugendensystem des Lösungsraums $L_{(h)} \subseteq \mathcal{C}^1(I, \mathbb{R}^n)$ von $\dot{x} = Ax$ auf $\mathbb{R}^n$.

¡? Beweis ¿!

Ist $\varphi : \mathbb{R} \to \mathbb{C}^n$ (komplexe) Lösung von $\dot{z} = Az$, so ist auch $\overline{\varphi} : \mathbb{R} \to \mathbb{C}^n$ eine Lösung, denn

$$\dot{\overline{\varphi}} = \overline{\dot{\varphi}} = \overline{A\varphi} = \overline{A} \cdot \overline{\varphi} = A\overline{\varphi} \ ,$$

weil $A = \overline{A} \in \mathrm{Mat}_n(\mathbb{R})$ ist. Deshalb sind auch $\Re(\varphi) = \frac{1}{2}\big(\varphi + \overline{\varphi}\big)$ und $\Im(\varphi) = \frac{1}{2}\big(\varphi - \overline{\varphi}\big)$ Lösungen und zwar nun reelle Lösungen von $\dot{x} = Ax$ auf $\mathbb{R}^n$. Ist andererseits $x : \mathbb{R} \to \mathbb{R}^n$ beliebige Lösung von $\dot{x} = Ax$, $x \in L_{(h)}$, so ist x auch Lösung von $\dot{z} = Az$ auf $\mathbb{C}^n$, also gibt es $\mu_1, \ldots, \mu_n \in \mathbb{C}$, so dass gilt $x = \mu_1\varphi_1 + \cdots + \mu_n\varphi_n$. Setzt man $\mu_j = a_j + ib_j$ mit $j \in \{1, \ldots, n\}$ und $a_j, b_j \in \mathbb{R}$, so erhält man:

$$\begin{aligned} x &= \Re(x) \\ &= \Re(\mu_1\varphi_1 + \cdots + \mu_n\varphi_n) \\ &= \Re(\mu_1\varphi_1) + \cdots + \Re(\mu_n\varphi_n) \\ &= \Big(a_1\Re(\varphi_1) - b_1\Im(\varphi_1)\Big) + \cdots + \Big(a_n\Re(\varphi_n) - b_n\Im(\varphi_n)\Big) \in \mathrm{span}_{\mathbb{R}}\Big(\Re(\varphi_1), \ldots, \Im(\varphi_n)\Big) \ . \end{aligned}$$

Also ist $\big(\Re(\varphi_1), \ldots, \Im(\varphi_n)\big)$ reelles Erzeugendensystem von $L_{(h)}$.
QED.

8.2.10 Beispiel

Betrachte die Schwingungsgleichung

$$\ddot{x} + \omega^2 \cdot x = 0$$

auf $\mathbb{R}$ (mit $\omega > 0$). Setzen wir $x_1 = x$, $x_2 = \dot{x}$, so wird das zu $\dot{x} = Ax$ mit

$$A = \begin{pmatrix} 0 & 1 \\ -\omega^2 & 0 \end{pmatrix} \ .$$

Das charakteristische Polynom von A ist $\chi_A(\lambda) = \det(\lambda\mathbb{1}_2 - A) = \lambda^2 + \omega^2$, also hat A die (verschiedenen) Eigenwerte

$$\lambda_{1/2} = \pm i\omega \in \mathbb{C} \ .$$

Also ist A über $\mathbb{C}$ diagonalisierbar mit $S = \begin{pmatrix} 1 & 1 \\ i\omega & -i\omega \end{pmatrix} \in \mathrm{Gl}_2(\mathbb{C})$. Damit ist

$$B = S^{-1}AS = \begin{pmatrix} i\omega & 0 \\ 0 & -i\omega \end{pmatrix} \ ,$$

denn $s_1 = (1, i\omega)$ ist Eigenvektor für $\lambda_1 = i\omega$ und $s_2 = (1, -i\omega)$ ein solcher für $\lambda_2 = -i\omega$. Es ist dann also

$$\Phi(t) = S\Psi(t) = \begin{pmatrix} 1 & 1 \\ i\omega & -i\omega \end{pmatrix} \begin{pmatrix} \exp(i\omega t) & 0 \\ 0 & \exp(-i\omega t) \end{pmatrix} = \begin{pmatrix} \exp(i\omega t) & \exp(-i\omega t) \\ i\omega\exp(i\omega t) & -i\omega\exp(-i\omega t) \end{pmatrix}$$

komplexes Fundamentalsystem von $\dot{z} = Az$ auf $\mathbb{C}^2$. Da für $\varphi_1 = \begin{pmatrix} \exp(i\omega t) \\ i\omega\exp(i\omega t) \end{pmatrix}$ sicher $\Re(\varphi) = \begin{pmatrix} \cos(\omega t) \\ -\omega\sin(\omega t) \end{pmatrix}$ und $\Im(\varphi) = \begin{pmatrix} \sin(\omega t) \\ \omega\cos(\omega t) \end{pmatrix}$ auf $\mathbb{R}$ linear unabhängig sind (z.B. ist $\det \begin{pmatrix} \cos(\omega t) & \sin(\omega t) \\ -\omega\sin(\omega t) & \omega\cos(\omega t) \end{pmatrix} = \omega \neq 0$), ist $\big(\Re(\varphi_1), \Im(\varphi_1)\big)$ ein Fundamentalsystem für $\dot{x} = Ax$ auf $\mathbb{R}^2$ und damit $\big(\cos(\omega t), \sin(\omega t)\big)$ Fundamentalsystem für $\ddot{x} + \omega^2 x = 0$ auf $\mathbb{R}$.

Index