

# Analysis II - SS 2013

## Prof. Dr. Wilhelm Singhof

Persönliches Skript der Mitschriften mit Beweisen

L<sup>A</sup>T<sub>E</sub>Xed by Thien Phuong Ngo

Stand 18. Juni 2013

### Inhaltsverzeichnis

§1 Normierte und metrische Räume: Definitionen und Beispiele (09.04.2013)	2
§2 Einige grundlegende topologische Begriffe (12.04.2013)	6
§3 Stetige Abbildungen (23.04.2013)	11
§4 Partielle Ableitungen (26.04.2013)	14
§5 Differenzierbare und stetig differenzierbare Funktionen (30.04.2013)	17
§6 Mittelwertsatz und Taylor-Formel (7.05.2013)	22
§7 Extremwerte und kritische Stellen (14.05.2013)	26
Teil II: Gewöhnliche Differenzialgleichungen	30
§8 Beispiele und Problemstellungen (17.05.2013)	30
§9 Lineare Differenzialgleichungen (28.05.2013)	37
§10 Lineare Differenzialgleichungen mit konstanten Koeffizienten (04.06.2013)	42
§11 Der Fixpunktsatz von Banach (14.06.2013)	48
§12 der lokale Eindeutigkeits- und Existenzsatz (14.06.2013)	50

## §1 Normierte und metrische Räume: Definitionen und Beispiele

In Analysis betrachten wir Funktionen, die auf Teilmengen des  $\mathbb{R}^n$  definiert und Werte in  $\mathbb{R}^m$  haben ( $n, m \in \mathbb{N}$ ). Wir müssen zunächst den Absolutbetrag von  $\mathbb{R}$  auf  $\mathbb{R}^n$  verallgemeinern.

$\mathbb{R}^n = \{(x_1, \dots, x_n) \mid x_1, \dots, x_n \in \mathbb{R}\}$ . Später schreiben wir die Elemente von  $\mathbb{R}^n$  in der Form  $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ .

**Definition 1.** Ein (reeller) metrischer Raum besteht aus einem reellen Vektorraum  $V$  und einer Abbildung  $v \mapsto \|v\|$  von  $V$  in  $\mathbb{R}$ , der Norm, sodass gilt:

- (1)  $\|v\| \geq 0 \quad \forall v \in V$
- (2)  $\|v\| = 0 \Leftrightarrow v = 0$
- (3)  $\|\lambda v\| = |\lambda| \cdot \|v\| \quad \forall \lambda \in \mathbb{R}, v \in V$
- (4)  $\|v + w\| \leq \|v\| + \|w\| \quad \forall v, w \in V$  (Dreiecksungleichung)

**Beispiel 1.** Ist  $V = \mathbb{R}$  und  $\|v\| = |v| \quad \forall v \in \mathbb{R}$ , so ist  $V$  ein normierter Raum.

**Beispiel 2.** Sei  $V = \mathbb{R}^n$ , für  $v = (x_1, \dots, x_n) \in \mathbb{R}^n$  sei  $\|v\|_1 := |x_1| + \dots + |x_n|$  („Taxinorm“). Dann ist  $\|\cdot\|_1$  eine Norm:

(4) Sei  $v = (x_1, \dots, x_n), w = (y_1, \dots, y_n)$

$$\begin{aligned} \|v + w\|_1 &= \|(x_1 + y_1, \dots, x_n + y_n)\|_1 = |x_1 + y_1| + \dots + |x_n + y_n| \\ &\leq |x_1| + |y_1| + \dots + |x_n| + |y_n| = \|v\|_1 + \|w\|_1 \end{aligned}$$

**Beispiel 3.** Sei  $V = \mathbb{R}^n$ , für  $v = (x_1, \dots, x_n) \in \mathbb{R}^n$  sei

$$\|v\|_\infty := \max\{|x_1|, \dots, |x_n|\}.$$

Dann ist  $\|\cdot\|_\infty$  eine Norm:

(4) Sei  $v = (x_1, \dots, x_n), w = (y_1, \dots, y_n)$

$$\|v\|_\infty = \max\{|x_1|, \dots, |x_n|\} = |x_i| \quad , \quad \|w\|_\infty = \max\{|y_1|, \dots, |y_n|\} = |y_j|$$

$$\|v + w\|_\infty = \max\{|x_1 + y_1|, \dots, |x_n + y_n|\}$$

Ist  $1 \leq k \leq n$ , so ist  $|x_k + y_k| \leq |x_k| + |y_k| \leq |x_i| + |y_j| = \|v\|_\infty + \|w\|_\infty$

Deswegen ist  $\max\{|x_1 + y_1|, \dots, |x_n + y_n|\} \leq \|v\|_\infty + \|w\|_\infty$

**Beispiel 4.** Sei  $V = \mathbb{R}^n$ , für  $v = (x_1, \dots, x_n) \in \mathbb{R}^n$  sei

$$\|v\|_2 := (x_1^2 + \dots + x_n^2)^{\frac{1}{2}} \quad (\text{Euklidische Norm})$$

Dann ist  $\|\cdot\|_2$  eine Norm. Für (4) brauchen wir

**Satz 1. Ungleichung von Cauchy-Schwarz** Sind  $v = (x_1, \dots, x_n), w = (y_1, \dots, y_n) \in \mathbb{R}^n$ , so ist:

$$|x_1 y_1 + \dots + x_n y_n| \leq \|v\|_2 \cdot \|w\|_2$$

*Beweis.* Wir gehen aus von der Ungleichung:

$$(*) \quad \sqrt{ab} \leq \frac{1}{2}(a + b) \quad \text{für } a, b \in \mathbb{R}_{\geq 0}$$

Wir können annehmen, dass  $v \neq 0 \neq w$ . Für  $1 \leq k \leq n$  ist

$$\frac{|x_k \cdot y_k|}{\|v\|_2 \cdot \|w\|_2} = \sqrt{\frac{x_k^2}{\|v\|_2^2} \cdot \frac{y_k^2}{\|w\|_2^2}} \stackrel{(*)}{\leq} \frac{1}{2} \cdot \left( \frac{x_k^2}{\|v\|_2^2} + \frac{y_k^2}{\|w\|_2^2} \right)$$

Also

$$\begin{aligned} \frac{|x_1 y_1 + \dots + x_n y_n|}{\|v\|_2 \cdot \|w\|_2} &\leq \frac{|x_1 y_1|}{\|v\|_2 \cdot \|w\|_2} + \dots + \frac{|x_n y_n|}{\|v\|_2 \cdot \|w\|_2} \\ &\leq \frac{1}{2} \left( \frac{x_1^2 + \dots + x_n^2}{\|v\|_2^2} + \frac{y_1^2 + \dots + y_n^2}{\|w\|_2^2} \right) \\ &= \frac{1}{2}(1 + 1) = 1 \end{aligned}$$

□

### Nachweis der Dreiecksungleichung für $\|\cdot\|_2$ :

Seien  $v, w \in \mathbb{R}^n$ . Wir müssen zeigen:  $\|v + w\|_2^2 \leq (\|v\|_2 + \|w\|_2)^2$ .

Es ist

$$(\|v\|_2 + \|w\|_2)^2 = \|v\|_2^2 + 2\|v\|_2 \cdot \|w\|_2 + \|w\|_2^2$$

und

$$\begin{aligned} \|v + w\|_2^2 &= (x_1 + y_1)^2 + \dots + (x_n + y_n)^2 \\ &= x_1^2 + \dots + x_n^2 + 2(x_1 y_1 + \dots + x_n y_n) + y_1^2 + \dots + y_n^2 \\ &= \|v\|_2^2 + 2(x_1 y_1 + \dots + x_n y_n) + \|w\|_2^2 \\ &\leq \|v\|_2^2 + 2|x_1 y_1 + \dots + x_n y_n| + \|w\|_2^2 \\ &\stackrel{\text{S. 1}}{\leq} \|v\|_2^2 + 2\|v\|_2 \cdot \|w\|_2 + \|w\|_2^2 \end{aligned}$$

**Bemerkung 1.** Sei  $v = (x_1, \dots, x_n) \in \mathbb{R}^n$ . Dann ist

$$\|v\|_\infty = \max\{|x_1|, \dots, |x_n|\} \leq \|v\|_2 = \sqrt{x_1^2 + \dots + x_n^2} \leq \|v\|_1 = |x_1| + \dots + |x_n| \leq n \cdot \|v\|_\infty$$

Deswegen ist es für viele Zwecke einerlei welche Norm man nimmt.

Es gilt sogar: Ist  $V$  ein endlich dimensionaler Vektorraum und sind  $\|\cdot\|, \|\cdot\|'$  zwei Normen auf  $V$ , so sind sie äquivalent im folgendem Sinn:

Es gibt Zahlen  $a, A \in \mathbb{R}_{>0}$ , sodass  $a \cdot \|\cdot\| \leq \|\cdot\|' \leq A \cdot \|\cdot\| \quad \forall v \in V$

### Zwei wichtige Beispielklassen

1.  $V = \mathbb{R}^n$ , für  $v = (x_1, \dots, x_n) \in \mathbb{R}^n$  sei

$$\begin{aligned} \|v\|_\infty &:= \max\{|x_1|, \dots, |x_n|\} \\ \|v\|_1 &:= |x_1| + \dots + |x_n| \\ \|v\|_2 &:= (x_1^2 + \dots + x_n^2)^{\frac{1}{2}} \\ \|v\|_\infty &\leq \|v\|_2 \leq \|v\|_1 \leq n \cdot \|v\|_\infty \end{aligned}$$

Man kann zeigen: Ist  $p \in \mathbb{R}$ ,  $p \geq 1$ , so erhält man eine Norm  $\|\cdot\|_p$  auf  $\mathbb{R}^n$  durch

$$\|v\|_p := (|x_1|^p + \dots + |x_n|^p)^{\frac{1}{p}}$$

Alle Normen auf einem endlich-dimensionalen Vektorraum  $V$  sind „äquivalent“.

2.  $V$  ist ein Funktionenraum (vgl. Aufg. 3 und 4)

z.B.:  $V = C[0, 1] =$  der Raum aller stetigen Funktionen  $f : [0, 1] \rightarrow \mathbb{R}$

Sei  $f \in C[0, 1]$ .

$$\|f\|_{\infty} := \max\{|f(x)| \mid x \in [0, 1]\}$$

$$\|f\|_1 := \int_0^1 |f(x)| \, dx$$

$$\|f\|_2 := \left( \int_0^1 f(x)^2 \, dx \right)^{\frac{1}{2}}$$

Damit erhält man Normen auf  $C[0, 1]$ , allgemeiner erhält man für  $p \geq 1$  eine Norm  $\|\cdot\|_p$  auf  $C[0, 1]$  durch  $\|f\|_p := \left( \int_0^1 |f(x)|^p \, dx \right)^{\frac{1}{p}}$

Ist  $X$  eine Menge, so ist  $X \times X = \{(x, y) \mid x, y \in X\}$

**Definition 2.** Ist  $X$  eine Menge, so ist eine Metrik auf  $X$  eine Abbildung

$$d : X \times X \longrightarrow \mathbb{R}$$

mit folgenden Eigenschaften:

$$(I) \quad d(x, y) \geq 0 \quad \forall x, y \in X$$

$$(II) \quad d(x, y) = 0 \Leftrightarrow x = y$$

$$(III) \quad d(x, y) = d(y, x) \quad \forall x, y \in X$$

$$(IV) \quad \text{Dreiecksungleichung: } d(x, z) \leq d(x, y) + d(y, z) \quad \forall x, y, z \in X$$

Ist  $d$  eine Metrik auf  $X$ , so nennt man  $(X, d)$  einen metrischen Raum.

Man sagt oft: „Sei  $X$  ein metrischer Raum“, statt „sei  $(X, d)$  ein metrischer Raum“.

**Beispiel 5.** Ist  $\|\cdot\|$  eine Norm auf dem Vektorraum  $V$ , so erhält man eine Metrik  $d$  auf  $V$  durch

$$d(x, y) := \|x - y\|$$

**Beispiel 6.** Ist  $(X, d)$  ein metrischer Raum und  $Y \subseteq X$ , so definiere  $d' := Y \times Y \longrightarrow \mathbb{R}$  durch  $d'(x, y) := d(x, y) \quad \forall x, y \in Y$ . Dann ist  $(Y, d')$  ein metrischer Raum.

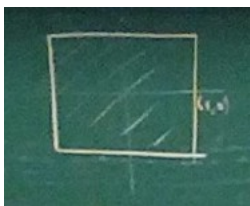
**Definition 3.** Sei  $X$  ein metrischer Raum,  $a \in X$ ,  $r \in \mathbb{R}_{>0}$ .

$$B_r(a) := \{x \in X \mid d(a, x) < r\} \quad \text{„offene Kugel um } a \text{ mit Radius } r$$

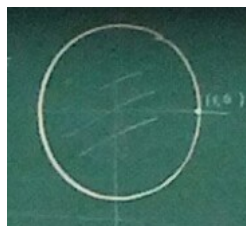
$$\overline{B}_r(a) := \{x \in X \mid d(a, x) \leq r\} \quad \text{„abgeschlossene Kugel um } a \text{ mit Radius } r$$

Manchmal notiert man dabei auch die Metrik  $d$  oder die Norm  $\|\cdot\|$ , wenn  $d$  wie in obigem Beispiel von  $\|\cdot\|$  herkommt.

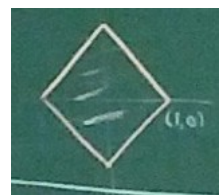
**Beispiel 7.**



1.1  $B_1(0, \|\cdot\|_{\infty})$



1.2  $B_1(0, \|\cdot\|_2)$



1.3  $B_1(0, \|\cdot\|_1)$

**Beispiel 8.** Sei  $X$  eine beliebige Menge.

Definiere  $D : X \times X \longrightarrow \mathbb{R}$  durch  $d(x, y) := \begin{cases} 0 & , \text{falls } x = y \\ 1 & , \text{falls } x \neq y \end{cases}$  Das ist eine Metrik.

$$B_r(a) := \begin{cases} \{a\} & , r = 1 \\ X & , r > 1 \end{cases}$$

$$\overline{B}_r(a) = \begin{cases} \{a\} & , r < 1 \\ X & , r \geq 1 \end{cases}$$

## §2 Einige grundlegende topologische Begriffe

**Motivation:** Ist  $A \subseteq \mathbb{R}^n$  und  $f : A \rightarrow \mathbb{R}^m$  eine Funktion, so kann man nur dann davon reden, dass  $f$  differenzierbar ist, wenn  $A$  eine „offene Teilmenge von  $\mathbb{R}^n$ “ ist.

**Definition 1.** Sei  $X$  ein metrischer Raum und  $A \subseteq X$ . Dann heißt  $A$  offen in  $X$ , falls gilt:  
Ist  $a \in A$ , so gibt es ein  $r \in \mathbb{R}_{>0}$  mit  $B_r(a) \subseteq A$

**Satz 1.** Sei  $X$  ein metrischer Raum, sei  $a \in X$  und  $r \in \mathbb{R}_{>0}$ . Dann ist  $A := B_r(a)$  offen in  $X$ . („Offene Kugeln sind offene Teilmengen“.)

*Beweis.* Sei  $b \in A = B_r(a)$ . Dann ist  $d(a, b) < r$ . Sei  $\rho := r - d(a, b) > 0$ . Dann ist  $B := B_\rho(b) \subseteq A$ . Sei  $x \in B$ . Dann ist  $d(x, b) < \rho$ , also

$$d(x, a) \stackrel{(IV)}{\leq} d(x, b) + d(b, a) < \rho + d(b, a) = (r - d(a, b) + d(b, a)) = r$$

□

**Satz 2.** Sei  $X$  ein metrischer Raum. Dann gilt:

- (1)  $X$  und  $\emptyset$  sind offen in  $X$
- (2) Sei  $\Lambda$  eine beliebige Menge und für jedes  $\lambda \in \Lambda$  sei eine offene Teilmenge  $A_\lambda$  von  $X$  gegeben.  
Dann ist  $A := \bigcup_{\lambda \in \Lambda} A_\lambda$  offen in  $X$
- (3) Ist  $n \in \mathbb{N}$  und sind  $A_1, \dots, A_n$  offen in  $X$ , so ist  $B := A_1 \cap \dots \cap A_n$  offen in  $X$

**Beispiel 1.** Sei  $X = \mathbb{R}$  mit  $d(x, y) = |x - y|$ . Für  $r \in \mathbb{R}_{>0}$  ist  $B_r(0) = ]-r, r[$  offen in  $\mathbb{R}$ . Aber  $\bigcap_{r \in \mathbb{R}_{>0}} B_r(0) = \{0\}$ , und das ist nicht offen in  $\mathbb{R}$

*Beweis.* von Satz 3.

- (1) ✓
- (2) Sei  $a \in A = \bigcup_{\lambda \in \Lambda} A_\lambda$ . Dann gibt es ein  $\lambda \in \Lambda$  mit  $a \in A_\lambda$ .  
Weil  $A_\lambda$  offen in  $X$  ist, gibt es ein  $r > 0$  mit  $B_r(a) \subseteq A_\lambda \subseteq A$
- (3) Sei  $a \in A_1 \cap \dots \cap A_n$ .  
Dann ist  $a \in A_i$  für  $i = 1, \dots, n$  und weil  $A_i$  offen in  $X$  ist, gibt es ein  $r_i > 0$  mit  $B_{r_i}(a) \subseteq A_i$   
Sei  $r := \min\{r_1, \dots, r_n\} > 0$ . Dann ist  $B_r(a) \subseteq B_{r_i}(a) \subseteq A_i \quad \forall i \Rightarrow B_r(a) \subseteq A_1 \cap \dots \cap A_n = B$

□

**Definition 2.** Sei  $X$  ein metrischer Raum,  $x \in X$  und  $U \subseteq X$ . Dann heißt  $U$  eine Umgebung von  $x \in X$ , falls es eine offene Teilmenge  $A$  von  $X$  gibt mit  $x \in A \subseteq U$

**Beispiel 2.** Sei  $X = \mathbb{R}$  mit der üblichen Metrik  $d(x, y) = |x - y|$ . Dann ist  $[-2, -1] \cup \mathbb{R}_{>0}$  eine Umgebung von 1

### Eigenschaften von Umgebungen

- (1) ist  $x \in X$  und  $U \subseteq X$ , so sind äquivalent:

- (a)  $U$  ist Umgebung von  $x$
- (b) Es gibt ein  $r > 0$  mit  $B_r(x) \subset U$

*Beweis.* (a)  $\Rightarrow$  (b) folgt direkt aus den Definitionen

(b)  $\Rightarrow$  (a) folgt aus Satz 1.

□

- (2) Eine Teilmenge  $A$  von  $X$  ist genau dann offen in  $X$ , wenn  $A$  Umgebung von jedem Punkt von  $A$  in  $X$  ist.

*Beweis.* Ist  $A$  offen in  $X$  und  $x \in A$ , so ist  $A$  Umgebung von  $x$  in  $X$  nach Definition einer Umgebung.

Wenn  $A$  Umgebung von jedem Punkt  $x \in A$  in  $X$  ist, so gibt es für jedes  $x \in A$  nach (1) ein  $r > 0$  mit  $B_r(x) \subseteq A$ , und deswegen ist  $A$  offen in  $X$ .  $\square$

- (3) Ist  $U$  eine Umgebung von  $x$  in  $X$  und ist  $U \subseteq V \subseteq X$ , so ist  $V$  eine Umgebung von  $x$  in  $X$

- (4) Sind  $U_1, \dots, U_n$  endlich viele Umgebungen von  $x$  in  $X$ , so ist  $U_1 \cap \dots \cap U_n$  eine Umgebung von  $x$  in  $X$ .

*Beweis.* Es gibt offene Teilmengen  $A_1, \dots, A_n$  von  $X$  mit  $x \in A_i \subseteq U_i$  für  $i = 1, \dots, n$ .

Nach Satz 2.c) ist  $A_1 \cap \dots \cap A_n$  offen in  $X$  und  $x \in A_1 \cap \dots \cap A_n \subseteq U_1 \cap \dots \cap U_n$ .

Deswegen ist  $U_1 \cap \dots \cap U_n$  eine Umgebung von  $x$  in  $X$ .  $\square$

**Beispiel 3.** Wir betrachten den  $\mathbb{R}^n$  mit den Metriken, die zu den Normen  $\|\cdot\|_1, \|\cdot\|_2, \|\cdot\|_\infty$  gehören. Wegen  $\|v\|_\infty \leq \|v\|_2 \leq \|v\|_1 \leq n \cdot \|v\|_\infty$  besitzen diese 3 Metriken dieselben offenen Mengen!

**Definition 3.** Sei  $X$  ein metrischer Raum,  $A \subseteq X$  und  $x \in X$

- a) Dann heißt  $x$  ein Häufungspunkt von  $A$ , wenn gilt:

In jeder Umgebung von  $x$  in  $X$  liegt ein von  $x$  verschiedener Punkt von  $A$ .

- b)  $x$  heißt Berührungspunkt von  $A$ , wenn gilt:

In jeder Umgebung von  $x$  liegt ein Punkt von  $A$ .

**Bemerkung 1.**  $x$  ist Berührungspunkt von  $A \Leftrightarrow x \in A$  oder  $x$  ist Häufungspunkt von  $A$ .

**Beispiel 4.** Sei  $X = \mathbb{R}$ .

$A_1 := ]0, 1[$  Menge der Häufungspunkte von  $A_1$  in  $X$  ist  $[0, 1]$ .

Das ist auch die Menge der Berührungspunkte von  $A_1$  in  $X$ .

$A_2 := \{1\} \cup ]2, 3[$  Menge der Häufungspunkte:  $[2, 3]$ .

Menge der Berührungspunkte:  $\{1\} \cup [2, 3]$

**Satz 3.** und Definition. Sei  $X$  ein metrischer Raum und  $A \subseteq X$ . Dann sind äquivalent:

- (1)  $A$  enthält alle Berührungspunkte von  $A$
- (2)  $A$  enthält alle Häufungspunkte von  $A$
- (3) Die Teilmenge  $X \setminus A = \{x \in X \mid x \notin A\}$  von  $X$  ist offen in  $X$

Wenn  $A$  diese Eigenschaften erfüllt, so heißt  $A$  abgeschlossen in  $X$ .

*Beweis.* Die Äquivalenz von (1) und (2) folgt aus der Bemerkung 1.

(1)  $\Rightarrow$  (3): Wir wollen Eigenschaft (2) der Umgebungen benutzen. Sei also  $x \in X \setminus A$ .

Dann müssen wir zeigen:  $X \setminus A$  ist Umgebung von  $x$  in  $X$ .

Weil  $x \notin A$  ist  $x$  nach (1) kein Berührungspunkt von  $A$ .

Es gibt also eine Umgebung  $U$  von  $x$  in  $X$  mit  $U \cap A = \emptyset$ , also  $x \in U \subseteq X \setminus A$ .

Also ist  $X \setminus A$  Umgebung von  $x$  nach Eigenschaft (3).

(3)  $\Rightarrow$  (1): Wir setzen jetzt voraus, dass  $X \setminus A$  offen in  $X$  ist.

Sei  $x$  Berührungspunkt von  $A$ . Wir müssen zeigen:  $x \in A$ .

Angenommen  $x \in X \setminus A$ . Dann ist  $X \setminus A$  eine Umgebung von  $x$ , weil  $X \setminus A$  offen ist.

Aber  $(X \setminus A) \cap A = \emptyset$  Widerspruch zur Tatsache, dass  $x$  Berührungspunkt von  $A$  ist.  $\square$

**Satz 4.** Sei  $X$  ein metrischer Raum.

- (a)  $X$  und  $\emptyset$  sind abgeschlossen in  $X$
- (b) Der Durchschnitt von beliebig vielen abgeschlossenen Teilmengen ist abgeschlossen in  $X$
- (c) Die Vereinigung von endlich vielen abgeschlossenen Teilmengen ist abgeschlossen in  $X$

*Beweis.* Satz 2 und die Charakterisierung (3) von abgeschlossenen Mengen □

**Beispiel 5.** Für  $n \in \mathbb{N}$ ,  $n \geq 3$ , ist  $[\frac{1}{n}, 1 - \frac{1}{n}]$  abgeschlossen in  $\mathbb{R}$

$\bigcup_{n=3}^{\infty} [\frac{1}{n}, 1 - \frac{1}{n}] = ]0, 1[$  ist nicht abgeschlossen.

**Satz 5.** Sei  $X$  ein metrischer Raum und  $A$  eine endliche Teilmenge von  $X$ . Dann ist  $A$  abgeschlossen in  $X$ .

*Beweis.* Nach Satz 4.c) genügt es zu zeigen: Ist  $a \in X$ , so ist  $\{a\}$  abgeschlossen in  $X$ .

Sei  $b \in X$ ,  $b \neq a$ . Wir wollen zeigen:  $b$  ist kein Berührungspunkt von  $\{a\}$ .

Sei  $r := d(a, b) > 0$ . Dann ist  $B_r(b)$  eine Umgebung von  $b$  mit  $a \notin B$  □

**Definition 4.** Sei  $X$  ein metrischer Raum, sei  $(x_n)_{n \in \mathbb{N}}$  eine Folge in  $X$ . Sei  $a \in X$ .

Dann heißt  $a$  Grenzwert der Folge  $(x_n)$ , in Zeichen:  $a = \lim_{n \rightarrow \infty} x_n$ , wenn eine der folgenden 4 äquivalenten Bedingungen erfüllt ist:

- (1) Für jedes  $\varepsilon > 0$  gibt es ein  $N \in \mathbb{N}$  mit  $d(x_n, a) < \varepsilon$  für  $n \geq N$
- (2)  $\lim_{n \rightarrow \infty} d(x_n, a) = 0$  im Sinne von Analysis I
- (3) Für jedes  $\varepsilon > 0$  gibt es ein  $N \in \mathbb{N}$  mit  $x_n \in B_\varepsilon(a)$  für  $n \geq N$
- (4) Für jede Umgebung  $U$  von  $x$  in  $X$  gibt es ein  $N \in \mathbb{N}$  mit  $x_n \in U$  für  $n \geq N$

Eine Folge besitzt höchstens einen Grenzwert. Wenn  $(x_n)$  den Grenzwert  $x_0$  besitzt, so sagt man, dass  $(x_n)$  gegen  $x_0$  konvergiert. und schreibt  $\lim_{n \rightarrow \infty} x_n = x_0$  oder  $x_n \rightarrow x_0$  (für  $n \rightarrow \infty$ ).

**Beispiel 6.** Sei  $X = \mathbb{R}^n$  mit einer der Normen  $\|\cdot\|_p$ ,  $p = 1, 2, \infty$ .

Sei  $x^k = (\xi_1^k, \dots, \xi_n^k)$  und sei  $x^0 = (\xi_1^0, \dots, \xi_n^0)$ .

Genau dann gilt  $\lim_{k \rightarrow \infty} x^k = x^0 \Leftrightarrow \lim_{k \rightarrow \infty} \xi_\mu^k = \xi_\mu^0$  für alle  $\mu = 1, \dots, n$

*Beweis.* Wegen  $\|v\|_\infty \leq \|v\|_2 \leq \|v\|_1 \leq n \cdot \|v\|_\infty$  genügt es, die Aussage für  $\|\cdot\|_\infty$  zu beweisen.

$$\begin{aligned} \lim_{k \rightarrow \infty} x^k = x^0 &\Leftrightarrow \|x^k - x^0\|_\infty \rightarrow 0 \quad \text{für } k \rightarrow \infty \\ &\Leftrightarrow \max_{\mu} |\xi_\mu^k - \xi_\mu^0| \rightarrow 0 \quad \text{für } k \rightarrow \infty \\ &\Leftrightarrow |\xi_\mu^k - \xi_\mu^0| \rightarrow 0 \quad \text{für } k \rightarrow \infty \text{ und } \mu = 1, \dots, n \\ &\Leftrightarrow \lim_{k \rightarrow \infty} \xi_\mu^k = \xi_\mu^0 \quad \text{für } \mu = 1, \dots, n \end{aligned}$$

□

**Bemerkung 2.** Sei  $X$  ein metrischer Raum,  $A \subseteq X$ ,  $x \in X$ .

- (a)  $x$  ist ein Berührungspunkt von  $A \Leftrightarrow \exists (x_n)$  in  $A$  mit  $\lim_{n \rightarrow \infty} x_n = x$
- (b)  $x$  ist ein Häufungspunkt von  $A \Leftrightarrow \exists (x_n)$  in  $A$  mit  $x_n \neq x$  für alle  $n \in \mathbb{N}$  und  $\lim_{n \rightarrow \infty} x_n = x$
- (c)  $A$  ist eine abgeschlossene Teilmenge in  $X \Leftrightarrow$  Ist  $(x_n)$  eine Folge in  $A$ , sodass  $x_0 = \lim_{n \rightarrow \infty} x_n$  existiert, so ist  $x_0 \in A$



*Beweis.*

- (a) „ $\Leftarrow$ “: zz.:  $x$  ist ein Berührungspunkt. Sei  $U$  eine Umgebung von  $x$ .  
 zz.:  $U \cap A \neq \emptyset$ . Nach Voraussetzung existiert eine Folge  $(x_n)$  mit  $\lim_{n \rightarrow \infty} x_n = x$ . Also existiert ein  $N$  mit  $x_n \in U$  für alle  $n \geq N$ .

„ $\Rightarrow$ “: Da  $x$  ein Berührungspunkt von  $A$  ist existiert für jedes  $n \in \mathbb{N}$  ein Element  $x_n \in B_{\frac{1}{n}}(x) \cap A$ .

Nach Definition konvergiert die Folge  $(x_n)$  gegen  $x$ , denn  $d(x_n, x) \leq \frac{1}{n} \rightarrow 0$  für  $n \rightarrow \infty$

- (b) Analog zu (a)

- (c) folgt aus (a) und der Definition von „ $A$  abgeschlossen in  $X$ “

□

**Definition 5.** Sei  $X$  ein metrischer Raum,  $A \subseteq X$  und  $x \in X$ . Dann heißt  $x$  innerer Punkt von  $A$ , wenn  $A$  eine Umgebung von  $x$  ist (d.h. es existiert ein  $r > 0$  mit  $B_r(x) \subseteq A$ ).

Sei  $\mathring{A} = \{x \in X \mid x \text{ ist innerer Punkt von } A\}$  die Menge der inneren Punkte von  $A$  (man sagt auch:  $\mathring{A}$  ist das innere von  $A$ ).

**Bemerkung 3.**

1.  $\mathring{A} \subseteq A$
2.  $\mathring{A} = A \Leftrightarrow A$  offen in  $X$

**Satz 6.** Sei  $X$  ein metrischer Raum und  $A \subseteq X$ . Dann ist  $\mathring{A}$  die größte offene Teilmenge von  $X$ , die in  $A$  enthalten ist.

*Beweis.*

1. Schritt Zeige  $\mathring{A}$  ist offen. Sei  $x \in \mathring{A}$ . Dann ist  $A$  eine Umgebung von  $x$ . Daher gibt es eine offene Menge  $B$  mit  $x \in B \subseteq A$ .  
 Für  $y \in B$  ist auch  $y \in B \subseteq A$ , also ist  $A$  eine Umgebung von  $y$   
 $\Rightarrow B \subseteq \mathring{A}$ , d.h.  $\mathring{A}$  ist offen
2. Schritt Sei  $B$  eine offene Menge von  $X$  mit  $B \subseteq A$ . Zeige  $B \subseteq \mathring{A}$ .  
 Für jedes  $x \in B$  ist  $x \in B \subseteq A$ , d.h.  $x$  ist innerer Punkt von  $A$ .  
 $\Rightarrow B \subseteq \mathring{A}$

□

**Satz 7.** Sei  $V$  ein normierter Vektorraum. Sei  $a \in V$ ,  $r > 0$ , setze  $A = \overline{B}_r(a) = \{x \in V \mid \|x - a\| \leq r\}$ . Dann ist  $\mathring{A} = B_r(a)$ .

*Beweis.*  $B_r(a)$  ist offen in  $V$  nach Satz 1, also  $B_r(a) \subseteq \mathring{A}$ .

Sei  $x \in \overline{B}_r(a) \setminus B_r(a)$ , d.h.  $\|x - a\| = r$ . Zu zeigen:  $x \notin \mathring{A}$ .

Dafür muss man zeigen: Für jedes  $\varepsilon > 0$  enthält  $B_\varepsilon(x)$  ein Element  $y \in V$ , das nicht in  $A$  liegt.

$$y = x + \frac{\varepsilon}{2r}(x - a)$$

Dann gilt:

- $y \in B_\varepsilon(x)$ , denn  $\|x - y\| = \frac{\varepsilon}{2r} \cdot r = \frac{\varepsilon}{2} < \varepsilon$
- $y \notin \overline{B}_r(a)$ , denn  $\|a - y\| = r + \frac{\varepsilon}{2} > r$

□

**Definition 6.** Sei  $X$  ein metrischer Raum. Sei  $A \subseteq X$ . Sei  $\bar{A}$  die Menge der Berührungspunkte von  $A$  in  $X$ .  
 $\bar{A}$  heißt der Abschluss von  $A$ .

**Bemerkung 4.**

1.  $A \subseteq \bar{A}$
2.  $A = \bar{A} \Leftrightarrow A$  abgeschlossen

**Satz 8.**

- (1)  $X \setminus \bar{A} = (X \setminus A)^\circ$
- (2)  $\bar{A}$  ist die kleinste abgeschlossene Teilmenge von  $X$ , die  $A$  umfasst

*Beweis.*

(1)

$$\begin{aligned}
 x \in X \setminus \bar{A} &\Leftrightarrow x \text{ ist kein Berührungspunkt von } A \\
 &\Leftrightarrow \text{es existiert eine Umgebung } U \text{ von } x \text{ mit } U \cap A = \emptyset \\
 &\Leftrightarrow \text{es existiert eine Umgebung } U \text{ von } x \text{ mit } U \subseteq (X \setminus A) \\
 &\Leftrightarrow X \setminus A \text{ ist eine Umgebung von } x \\
 &\Leftrightarrow x \in (X \setminus A)^\circ
 \end{aligned}$$

(2) folgt aus (1) und Satz 6

□

**Definition 7.** Sei  $X$  ein metrischer Raum,  $A \subseteq X$ ,  $x \in X$ .  $x$  heißt Randpunkt von  $A$ , wenn  $x$  eine Berührungspunkt von  $A$  und von  $X \setminus A$  ist.

Sei  $\delta A$  die Menge der Randpunkte von  $A$ . D.h.  $\delta A = \bar{A} \cap \overline{(X \setminus A)}$

**Bemerkung 5.**  $\delta A$  ist abgeschlossen in  $X$ . Außerdem ist  $X$  die disjunkte Vereinigung von  $\overset{\circ}{A}$ ,  $\delta A$  und  $(X \setminus A)^\circ$

**Beispiel 7.**

1. Sei  $X = \mathbb{R}^2$ .  $A = \{x \in \mathbb{R}^2 \mid \|x\|_2 < 1\} \cup \{(1, 0)\}$   
 $\overset{\circ}{A} = \{x \mid \|x\|_2 < 1\}$ ,  $\bar{A} = \{x \mid \|x\|_2 \leq 1\}$   
 $(X \setminus A)^\circ = \{x \mid \|x\|_2 > 1\}$ ,  $\overline{(X \setminus A)} = \{x \mid \|x\|_2 \geq 1\}$   
 $\delta A = \{x \mid \|x\|_2 = 1\}$
2. Sei  $X = \mathbb{R}$ .  $A = \mathbb{Q}$ ,  $\overset{\circ}{A} = \emptyset$ ,  $\bar{A} = \mathbb{R}$ ,  $(X \setminus A)^\circ = \emptyset$ ,  $\overline{(X \setminus A)} = \mathbb{R}$ ,  $\delta A = \mathbb{R}$

### §3 Stetige Abbildungen

**Definition 1.** Seien  $(X, d)$  und  $(Y, d')$  metrische Räume,  $f : X \rightarrow Y$  eine Abbildung. Sei  $x_0 \in X$ . Dann heißt  $f$  stetig im Punkt  $x_0$ , wenn eine der 3 folgenden Äquivalenten Bedingungen erfüllt ist:

- (1) Zu jedem  $\varepsilon > 0$  gib es ein  $\delta > 0$ , sodass gilt: Ist  $x \in X$  mit  $d(x, x_0) < \delta$ , so ist  $d'(f(x), f(x_0)) < \varepsilon$
- (2) Zu jedem  $\varepsilon > 0$  gibt es ein  $\delta > 0$  mit  $f(B_\delta(x_0)) \subseteq B_\varepsilon(f(x_0))$
- (3) Zu jeder Umgebung  $V$  von  $f(x_0)$  in  $Y$  gibt es eine Umgebung  $U$  von  $x_0$  in  $X$  mit  $f(U) \subseteq V$

$f$  heißt stetig, wenn  $f$  in jedem Punkt  $x_0$  von  $X$  stetig ist.

**Erinnerung:** Eine Teilmenge  $U$  von  $X$  heißt Umgebung von  $x_0$  in  $X$ , wenn es eine offene Teilmenge  $A$  von  $X$  gibt mit  $x_0 \in A \subseteq U$ .

Das ist genau dann der Fall, wenn es ein  $r > 0$  gibt mit  $B_r(x_0) \subseteq U$ .

**Beispiel 1.**

1. Ist  $X$  ein metrischer Raum, so ist  $\text{id}_X : X \rightarrow X$  stetig
2. Eine konstante Abbildung zwischen metrischen Räumen ist stetig
3. Sind  $X, Y, Z$  metrische Räume und sind  $f : X \rightarrow Y$  und  $g : Y \rightarrow Z$  stetig, so ist  $g \circ f : X \rightarrow Z$  stetig.

*Beweis.* Sei  $x_0 \in X$ . Sei  $W$  eine Umgebung von  $g(f(x_0))$  in  $Z$ .

Weil  $g$  stetig ist, gibt es eine Umgebung  $V$  von  $f(x_0)$  mit  $g(V) \subseteq W$ .

Weil  $f$  stetig ist, gibt es eine Umgebung  $U$  von  $x_0$  mit  $f(U) \subseteq V$ .

Dann ist  $g \circ f(U) = g(f(U)) \subseteq g(V) \subseteq W$  □

**Erinnerung:** Sind  $X, Y$  Mengen und ist  $f : X \rightarrow Y$  eine Abbildung und  $A \subseteq Y$ , so ist  $f^{-1}(A) := \{x \in X \mid f(x) \in A\}$  (das Urbild von  $A$  unter  $f$ ).

**Satz 1.** Seien  $X, Y$  metrische Räume,  $f : X \rightarrow Y$  eine Abbildung. Dann sind äquivalent:

- (1)  $f$  ist stetig
- (2) Ist  $A$  offen in  $Y$ , so ist  $f^{-1}(A)$  offen in  $X$  („Urbilder offener Mengen sind offen“)
- (3) Ist  $B$  abgeschlossen in  $Y$ , so ist  $f^{-1}(B)$  abgeschlossen in  $X$
- (4) Ist  $(x_n)$  eine konvergente Folge in  $X$ , so ist  $(f(x_n))$  konvergent und

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

*Beweis.* (1)  $\Rightarrow$  (2): Sei  $A$  offen in  $Y$ . Sei  $x_0 \in f^{-1}(A)$ . Zu zeigen ist:  $f^{-1}(A)$  ist Umgebung von  $x_0$ .

Es ist  $f(x_0) \in A$  und  $A$  ist Umgebung von  $f(x_0)$ . Nach Definition der Stetigkeit (3) gibt es eine Umgebung  $U$  von  $x_0$  in  $X$  mit  $f(U) \subseteq A$ , also  $x_0 \in U \subseteq f^{-1}(A)$ .

Deswegen ist  $f^{-1}(A)$  Umgebung von  $x_0$ .

(2)  $\Rightarrow$  (1): Sei  $x_0 \in X$ . Zeige, dass  $f$  stetig im Punkt  $x_0$  ist.

Sei  $V$  eine Umgebung von  $f(x_0)$  in  $Y$ . Es gibt eine offene Teilmenge  $A$  von  $Y$  mit  $f(x_0) \in A \subseteq V$ .

Nach Voraussetzung ist  $f^{-1}(A)$  offen in  $X$ , und  $x_0 \in f^{-1}(A)$ .

Setzt man  $U := f^{-1}(A)$ , so ist  $U$  eine Umgebung von  $x_0$  in  $X$  mit  $f(U) \subseteq A \subseteq V$

(2)  $\Rightarrow$  (3): Sei  $B$  abgeschlossene Teilmenge von  $Y$ ,  $A := Y \setminus B$ . Dann ist  $A$  offen in  $Y$ .

Nach Voraussetzung (2) ist  $f^{-1}(A)$  offen in  $X$ , und  $f^{-1}(B) = f^{-1}(Y \setminus A) = X \setminus f^{-1}(A)$  ist abgeschlossen in  $X$ .

(3)  $\Rightarrow$  (2): geht genauso

(1)  $\Rightarrow$  (4): Sei  $(x_n)$  konvergente Folge in  $X$ ,  $x_0 := \lim_{n \rightarrow \infty} x_n$ . Sei  $y_n := f(x_n)$  und  $y_0 := f(x_0)$ .

Wir müssen zeigen:  $y_0 = \lim_{n \rightarrow \infty} y_n$ . Sei  $\varepsilon > 0$ . Es gibt ein  $\delta > 0$  mit  $f(B_\delta(x_0)) \subseteq B_\varepsilon(y_0)$ .

Weil  $(x_n)$  gegen  $x_0$  konvergiert, gibt es ein  $N \in \mathbb{N}$  mit  $x_n \in B_\delta(x_0) \quad \forall n \geq N$ .

Ist  $n \geq N$ , so ist  $y_n = f(x_n) \in B_\varepsilon(y_0)$

(4)  $\Rightarrow$  (1): Sei  $x_0 \in X$ . Sei  $\varepsilon > 0$ . Wir müssen zeigen: Es gibt ein  $\delta > 0$  mit  $f(B_\delta(x_0)) \subseteq B_\varepsilon(f(x_0))$ .

Andernfalls wäre  $f(B_{\frac{1}{n}}(x_0)) \not\subseteq B_\varepsilon(f(x_0)) \quad \forall n \in \mathbb{N}$ . Es gibt also  $x_n \in B_{\frac{1}{n}}(x_0)$  mit  $f(x_n) \notin B_\varepsilon(f(x_0))$ .

Deswegen ist  $\lim_{n \rightarrow \infty} x_n = x_0$ , aber  $(f(x_n))$  konvergiert nicht gegen  $f(x_0)$ .  $\nrightarrow$  Widerspruch!  $\square$

**Satz 2.** Seien  $X, Y$  metrische Räume und seien  $f, g : X \rightarrow Y$  stetig. Dann ist  $A := \{x \in X \mid f(x) = g(x)\}$  abgeschlossen in  $X$ .

*Beweis.* Sei  $(x_n)$  eine Folge in  $A$ , die gegen ein Punkt  $x_0 \in X$  konvergiert. Wir müssen zeigen, dass  $x_0 \in A$ . Es ist  $f(x_n) = g(x_n)$ . Nach Satz 1 ist  $f(x_0) = \lim_{n \rightarrow \infty} f(x_n) = \lim_{n \rightarrow \infty} g(x_n) = g(x_0)$   $\square$

**Satz 3.** Sei  $X$  ein metrischer Raum und seien  $f, g : X \rightarrow \mathbb{R}$  stetig. Dann ist  $\{x \in X \mid f(x) \leq g(x)\}$  abgeschlossen in  $X$ .

*Beweis.* Wie Satz 2.  $\square$

**Bemerkung 1.** Sei  $X$  ein metrischer Raum und sei  $f : X \rightarrow \mathbb{R}^m$  eine Abbildung. Dann ist  $f(X) = \{(f_1(x), \dots, f_m(x)) \mid x \in X\}$  mit Abbildungen  $f_1, \dots, f_m : X \rightarrow \mathbb{R}$ .

Wir übersehen  $\mathbb{R}^m$  mit einer der Normen  $\|\cdot\|_2, \|\cdot\|_1, \|\cdot\|_\infty$ .

Genau dann ist  $f$  stetig, wenn  $f_1, \dots, f_m$  stetig sind.

*Beweis.* Sei  $(x_n)$  eine Folge in  $X$  mit  $\lim_{n \rightarrow \infty} x_n = x_0$ .

$$\begin{aligned} f \text{ stetig} &\Leftrightarrow \underbrace{(f(x_n))}_{((f_1(x_n), \dots, f_m(x_n)))} \text{ konvergiert gegen } \underbrace{f(x_0)}_{(f_1(x_0), \dots, f_m(x_0))} \\ &\stackrel{\S 2}{\Leftrightarrow} (f_j(x_n)) \text{ konvergent gegen } f_j(x_0) \text{ für } j = 1, \dots, m \Leftrightarrow f_j \text{ ist stetig für } j = 1, \dots, m \end{aligned}$$

$\square$

**Definition 2.** Eine Teilmenge  $A$  eines normierten Raumes heißt beschränkt, wenn es ein  $M \in \mathbb{R}_{>0}$  gibt mit  $\|a\| \leq M \quad \forall a \in A$ .

**Satz 4.** Sei  $X$  eine beschränkte, abgeschlossene Teilmenge von  $\mathbb{R}^n$  und sei  $f : X \rightarrow \mathbb{R}$  stetig. Dann ist  $f(X) = \{f(x) \mid x \in X\}$  eine beschränkte, abgeschlossene Teilmenge von  $\mathbb{R}$ . Insbesondere nimmt  $f$  auf  $X$  sein Maximum und sein Minimum an.

*Beweis.*

1. Zeige, dass  $f(X)$  beschränkt ist: Andernfalls gibt es für jedes  $m \in \mathbb{N}$  ein  $x_m \in X$  mit  $|f(x_m)| > m$ .

Sei  $x_m = (x_m^1, \dots, x_m^n) \in X \subseteq \mathbb{R}^n$ .

Da  $X$  beschränkt ist, ist die Folge  $(x_m^1)_m$  beschränkt, besitzt also nach Analysis I eine konvergente Teilfolge. Indem wir zu einer Teilfolge von  $(x_m)$  übergehen, können wir annehmen, dass  $(x_m^1)_m$  gegen  $x_0^1 \in \mathbb{R}$  konvergiert. Dann können wir annehmen, dass auch  $(x_m^2)_m$  gegen  $x_0^2 \in \mathbb{R}$  konvergiert, usw.

Schließlich können wir annehmen, dass  $(x_m)$  gegen  $x_0$  konvergiert,  $x_0 \in X$ . Daher konvergiert  $(f(x_m))$  gegen  $f(x_0)$ .  $\nrightarrow$  Widerspruch zur Voraussetzung.

2. Zeige:  $f(A)$  ist eine abgeschlossene Teilmenge von  $\mathbb{R}$ . Sei  $y_0 \in \mathbb{R}$  ein Berührungspunkt von  $f(A)$ .  
Wir müssen zeigen:  $y_0 \in f(A)$ .  
Es gibt eine Folge  $(y^{(m)})$  in  $f(A)$ , die gegen  $y_0$  konvergiert.  
Sei  $x^{(m)} \in A$  mit  $f(x^{(m)}) = y^{(m)}$ ,  $m \in \mathbb{N}$ .  
Indem wir zu einer Teilfolge übergehen, können wir annehmen, dass  $(x^{(m)})$  gegen  $x_0 \in A$  konvergiert. Weil  $f$  stetig ist, ist  $f(x_0) = \lim_{n \rightarrow \infty} f(x^{(m)}) = \lim_{n \rightarrow \infty} y^{(m)} = y_0$ . Also  $y_0 \in f(A)$

□

## §4 Partielle Ableitungen

**Definition 1.** Sei  $U$  offene Teilmenge von  $\mathbb{R}^n$ , sei  $f : U \rightarrow \mathbb{R}$  eine Funktion und sei  $x = (x_1, \dots, x_n) \in U$  ein fester Punkt. Für  $i = 1, \dots, n$  sei

$$U_i := \{\xi \in \mathbb{R} \mid (x_1, \dots, x_{i-1}, \xi, x_{i+1}, \dots, x_n) \in U\}$$

Dann ist  $U_i$  eine offene Teilmenge von  $\mathbb{R}$ .

Definiere  $f_i : U_i \rightarrow \mathbb{R}$  durch  $f_i(\xi) := f(x_1, \dots, x_{i-1}, \xi, x_{i+1}, \dots, x_n)$ .

(Beachte:  $U_i$  und  $f_i$  hängt von  $x$  ab).

Wenn für  $i = 1, \dots, n$  die Funktion  $f_i$  an der Stelle  $x_i$  differenzierbar ist, so heißt  $f$  an der Stelle  $x$  partiell differenzierbar. Man schreibt dann:

$$D_i f(x) := \frac{df}{dx_i}(x) := \frac{d}{dx_i} f(x) := f'_i(x_i)$$

Man nennt dies die  $i$ -te partielle Ableitung von  $f$  an der Stelle  $x$ .

Wenn  $f$  an jeder Stelle von  $U$  partiell differenzierbar ist, so heißt  $f$  partiell differenzierbar.

In diesem Fall hat man Abbildungen  $D_i f = \frac{df}{dx_i} : U \rightarrow \mathbb{R}$ .

Für  $n = 2$  schreibt man meist  $(x, y)$  statt  $(x_1, x_2)$  und  $\frac{df}{dx} := \frac{df}{dx_1}$ ,  $\frac{df}{dy} := \frac{df}{dx_2}$ .

Für  $n = 3$  schreibt man meist  $(x, y, z)$  statt  $(x_1, x_2, x_3)$  und ....

**Beispiel 1.** Sei  $U = \mathbb{R}^2$ , und  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$  sei gegeben durch  $f(x, y) := e^{xy}$ .

$$\frac{df}{dx}(x, y) = ye^{xy} \quad , \quad \frac{df}{dy}(x, y) = xe^{xy}$$

**Beispiel 2.** Definiere  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$  durch  $f(x, y) = \begin{cases} \frac{xy}{x^2 + y^2} & \text{für } (x, y) \neq (0, 0) \\ 0 & \text{für } (x, y) = (0, 0) \end{cases}$ .

An jeder Stelle  $(x, y) \neq (0, 0)$  ist  $f$  partiell differenzierbar.

$f$  ist auch an der Stelle  $(0, 0)$  partiell differenzierbar:

$$\begin{aligned} f_1(\xi) &= f(\xi, 0) = 0 \quad \forall \xi \in \mathbb{R} \quad \text{Also } f'_1(0) = 0 \\ f_2(\xi) &= f(0, \xi) = 0 \quad \forall \xi \in \mathbb{R} \quad \text{Also } f'_2(0) = 0 \\ &\Rightarrow D_1 f(0, 0) = 0, \quad D_2 f(0, 0) = 0 \end{aligned}$$

Für  $\xi \neq 0$  ist  $f(\xi, \xi) = \frac{\xi^2}{\xi^2 + \xi^2} = \frac{1}{2}$ . Deswegen ist  $f$  nicht stetig an der Stelle  $(0, 0)$ .

Ist  $\lambda \in \mathbb{R}$  und  $\xi \neq 0$ , so ist  $f(\xi_1, \lambda\xi) = \frac{\lambda \cdot \xi^2}{\xi^2 + \lambda^2 \xi^2} = \frac{\lambda}{1 + \lambda^2}$

**Definition 2.** Sei  $U$  offen in  $\mathbb{R}^n$  und sei  $f : U \rightarrow \mathbb{R}$  partiell differenzierbar.

Dann haben wir Funktionen  $D_1 f, \dots, D_n f : U \rightarrow \mathbb{R}$ .

Wenn diese Funktionen wieder partiell differenzierbar sind, so sagt man, dass  $f$  zweimal partiell differenzierbar ist und schreibt:

$$D_j(D_i f) =: \frac{d^2 f}{dx_j dx_i} \quad , \quad D_i(D_i f) = D_i^2 f =: \frac{d^2 f}{dx_i^2}$$

usw. Wenn  $f$   $k$ -mal partiell differenzierbar ist und wenn alle partiellen Ableitungen der Ordnung  $\leq k$  stetig sind (dazu gehört, dass  $f$  selbst als partielle Ableitung 0-ter Ordnung von  $f$  stetig ist), so heißt  $f$  von der Klasse  $C^k$ .

Wenn  $f$  für jedes  $k \in \mathbb{N}$  von der Klasse  $C^k$  ist, so heißt  $f$  von der Klasse  $C^\infty$  oder glatt.  
 $f$  heißt von der Klasse  $C^0$ , wenn  $f$  stetig ist.

**Beispiel 3.**

$$f(x, y) = e^{xy}, \quad \frac{df}{dx}(x, y) = ye^{xy}, \quad \frac{df}{dy}(x, y) = xe^{xy}$$

$$\frac{d^2 f}{dx^2}(x, y) = \frac{d}{dx}(ye^{xy}) = y^2 e^{xy}, \quad \frac{d^2 f}{dy^2}(x, y) = \frac{d}{dy}(xe^{xy}) = x^2 e^{xy}$$

$$\frac{d^2 f}{dx dy}(x, y) = \frac{d}{dx}(xe^{xy}) = e^{xy} + xy \cdot e^{xy}$$

$$\frac{d^2 f}{dy dx}(x, y) = \frac{d}{dy}(ye^{xy}) = e^{xy} + xy \cdot e^{xy}$$

**Satz 1. Satz von H.A. Schwarz** Sei  $U$  offen in  $\mathbb{R}^n$  und sei  $f : U \rightarrow \mathbb{R}$  von der Klasse  $C^2$ . Dann ist  $D_i D_j f = D_j D_i f$  für alle  $i, j \in \{1, \dots, n\}$

Für den Beweis nehmen wir  $n = 2$  an. Wir nehmen ferner an, dass  $(0, 0) \in U$  und zeigen, dass  $D_1 D_2 f(0, 0) = D_2 D_1 f(0, 0)$ . Es gibt ein  $r > 0$  mit  $B := \{(x, y) \in \mathbb{R}^2 \mid |x| \leq r \text{ und } |y| \leq r\} \subseteq U$

**Zwischenbehauptung.** Ist  $(x, y) \in B$  mit  $x \neq 0$ ,  $y \neq 0$ , so gibt es  $(\xi, \eta) \in B$  und  $(\xi', \eta') \in B$  mit  $|\xi| = |x|$ ,  $|\xi'| \leq |x|$ ,  $|\eta| \leq |y|$ ,  $|\eta'| \leq |y|$ , sodass  $D_1 D_2 f(\xi', \eta') = D_2 D_1 f(\xi, \eta)$

1. Aus der Zwischenbehauptung folgt der Satz: Wir lassen den Punkt  $(x, y)$  die Folge  $(\frac{1}{n}, \frac{1}{n})$  für großes  $n \in \mathbb{N}$  durchlaufen. Dann bilden die zugehörigen Punkte  $(\xi, \eta)$  und  $(\xi', \eta')$  zwei Nullfolgen sind. Weil  $D_1 D_2 f$  und  $D_2 D_1 f$  stetig sind, folgt:

$$D_1 D_2 f(0, 0) = D_2 D_1 f(0, 0)$$

2. *Beweis.* der Zwischenbehauptung. Definiere  $F : B \rightarrow \mathbb{R}$  durch

$$F(x, y) := f(x, y) - f(x, 0) - f(0, y) + f(0, 0)$$

Für  $|y| < r$  definieren wir  $F_y : ]-r, r[ \rightarrow \mathbb{R}$  durch  $F_y(x) := f(x, y) - f(x, 0)$

$$\Rightarrow F(x, y) = F_y(x) - F_y(0)$$

$$F'_y(\xi) = D_1 f(\xi, y) - D_1 f(\xi, 0)$$

□

Nach dem Mittelwertsatz existiert ein  $\xi$  mit  $|\xi| \leq |x|$ , sodass

$$F_y(x) - F_y(0) = F'_y(\xi) \cdot x = (D_1 f(\xi, y)) = D_1 f(\xi, 0) \cdot x$$

Wieder nach dem Mittelwertsatz existiert ein  $\eta$  mit  $|\eta| \leq |y|$ , sodass

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

Insgesamt:  $F(x, y) = F_y(x) - F_y(0) = D_2 D_1 f(\xi, \eta) \cdot xy$ .

Ebenso findet man  $\xi', \eta'$  mit  $F(x, y) = D_1 D_2 f(\xi', \eta') \cdot xy$

$$\Rightarrow D_2 D_1 f(\xi, \eta) \cdot xy = D_1 D_2 f(\xi', \eta') \cdot xy$$

Weil  $x \neq 0$ ,  $y \neq 0$ , folgt  $D_2 D_1 f(\xi, \eta) = D_1 D_2 f(\xi', \eta')$

Von nun an schreiben wir die Elemente von  $\mathbb{R}^n$  als Spalten.

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = (x_1, \dots, x_n)^T \quad (\text{T: „transponiert“})$$

Ist  $X$  eine Menge und  $f : X \rightarrow \mathbb{R}^m$  eine Abbildung, so ist  $f$  von der Form:

$$f = \begin{pmatrix} f_1 \\ \vdots \\ f_m \end{pmatrix} = (f_1, \dots, f_m)^T \quad \text{mit Abbildungen } f_i : X \rightarrow \mathbb{R}$$

**Definition 3.** Sei  $U$  offen in  $\mathbb{R}^n$  und sei  $f : U \rightarrow \mathbb{R}$  partiell differenzierbar.

$$\text{grad } f(x) := \nabla f(x) := \left( \frac{df}{dx_1}(x), \dots, \frac{df}{dx_n}(x) \right)^T \in \mathbb{R}^n \quad \text{für } x \in U$$

Damit hat man eine Abbildung  $\text{grad } f = \nabla f : U \rightarrow \mathbb{R}^n$ , sie heißt der Gradient von  $f$ .  $\nabla$  wird „Nabla“ gelesen.

**Definition 4.** Sei  $U$  offen in  $\mathbb{R}^n$  und  $f = (f_1, \dots, f_m)^T : U \rightarrow \mathbb{R}^m$  eine Abbildung. Dann heißt  $f$  partiell differenzierbar, wenn alle  $f_i$ , partiell differenzierbar sind. Dann schreibt man für  $x \in U$ :

$$Df(x) := \begin{pmatrix} \frac{df_1}{dx_1}(x) & \dots & \frac{df_1}{dx_n}(x) \\ \vdots & & \vdots \\ \frac{df_m}{dx_1}(x) & \dots & \frac{df_m}{dx_n}(x) \end{pmatrix}$$

$Df(x)$  ist eine reelle  $m \times n$ -Matrix.

Sie heißt die Funktionalmatrix, Jacobimatrix oder Ableitung von  $f$  an der Stelle  $x$ .

Die  $i$ -te Zeile von  $Df(x)$  ist der transponierte Gradient von  $f_i$ .

$f$  heißt von der Klasse  $C^k$ , wenn jedes  $f_i$  von der Klasse  $C^k$  ist.



## §5 Differenzierbare und stetig differenzierbare Funktionen

**Vorbemerkung:** Sei  $U$  offen in  $\mathbb{R}$  und  $f : U \rightarrow \mathbb{R}$  eine Funktion. Sei  $x \in U$ .

Die Ableitung von  $f$  an der Stelle  $x$  ist definiert durch

$$f'(x) = \lim_{\xi \rightarrow 0} \frac{f(x + \xi) - f(x)}{\xi}, \quad \text{falls dieser Grenzwert existiert}$$

Das lässt sich nicht direkt auf Funktionen mehrerer Veränderlicher verallgemeinern, weil man nicht durch Elemente von  $\mathbb{R}^n$  für  $n \geq 2$  teilen kann.

Außweg:

1. Ist  $f$  an der Stelle  $x$  differenzierbar, so setzen wir

$$\varphi(\xi) := \left( f(x + \xi) - f(x) \right) - f'(x) \cdot \xi$$

$$\text{Dann ist } \lim_{\xi \rightarrow 0} \frac{\varphi(\xi)}{|\xi|} = 0$$

2. Umgekehrt: gibt es ein  $a \in \mathbb{R}$ , sodass mit

$$\varphi(\xi) := \left( f(x + \xi) - f(x) \right) - a \cdot \xi$$

$$\text{gilt, dass } \lim_{\xi \rightarrow 0} \frac{\varphi(\xi)}{|\xi|} = 0 \text{ ist, so ist } f \text{ differenzierbar an der Stelle } x \text{ mit } f'(x) = a.$$

Man kann versuchen dies auf Funktionen von mehreren Veränderlichen zu verallgemeinern, indem man dabei  $|\xi|$  durch  $\|\xi\|$  ersetzt.

**Satz 1.** Sei  $U$  offen in  $\mathbb{R}^n$  und  $f : U \rightarrow \mathbb{R}^m$  partiell differenzierbar,  $f = (f_1, \dots, f_m)^T$ .

Wir setzen voraus, dass alle  $D_i f_j : U \rightarrow \mathbb{R}$  stetig sind. Sei  $x \in U$  fest

und definiere  $\varphi(\xi) := \left( f(x + \xi) - f(x) \right) - Df(x) \cdot \xi \in \mathbb{R}^m$  für alle  $\xi \in \mathbb{R}^n$  mit  $x + \xi \in U$ .

$$\text{Dann ist } \lim_{\xi \rightarrow 0} \frac{\varphi(\xi)}{\|\xi\|} = 0.$$

(Das heißt: Für jedes  $\varepsilon > 0$  gibt es ein  $r > 0$ , sodass  $\frac{\|\varphi(\xi)\|}{\|\xi\|} < \varepsilon$  für alle  $\xi \in \mathbb{R}^n$  mit  $\|\xi\| < r$ )

**Bemerkung 1.** Für  $n \geq 2$  kann man die Voraussetzung, dass die  $D_i f_j$  stetig sind, nicht weglassen.

$$\text{Betrachte } f : \mathbb{R}^2 \rightarrow \mathbb{R}, \quad f(\xi_1, \xi_2) = \begin{cases} \frac{\xi_1 \xi_2}{\xi_1^2 + \xi_2^2} & \text{für } (\xi_1, \xi_2) \neq (0, 0) \\ 0 & \text{für } (\xi_1, \xi_2) = (0, 0) \end{cases}$$

$f$  ist partiell differenzierbar mit  $Df(0, 0) = (0, 0)$ . Wir wählen  $x = (0, 0)$

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

$$\varphi(\xi_1, \xi_2) = \frac{1}{2} \quad \text{für alle } \xi_1 \in \mathbb{R} \text{ mit } \xi_1 \neq 0$$

Deswegen existiert  $\lim_{\xi \rightarrow 0} \frac{\varphi(\xi)}{\|\xi\|}$  nicht!

*Beweis.* von Satz 1. Sei o.B.d.A.  $m = 1$ , also  $f : U \rightarrow \mathbb{R}$

$$Df(x) \cdot \xi = \sum_{j=1}^n D_j f(x) \cdot \xi_j$$

Sei  $\varepsilon > 0$ . Es gibt ein  $r > 0$  mit:

- Ist  $\|\xi\|_\infty = \max_i |\xi_i| < r$ , so ist  $x + \xi \in U$
- Ist  $\|\xi\|_\infty < r$ , so ist  $|D_i f(x + \xi) - D_i f(x)| < \varepsilon$  für  $i = 1, \dots, n$

Wir werden zeigen: Ist  $\|\xi\|_\infty < r$ , so ist  $|\varphi(\xi)| < \varepsilon \cdot \|\xi\|_1 = \varepsilon \cdot \sum_j |\xi_j|$

Für  $i = 0, \dots, n$  sei  $z_i := x + (\xi_1, \dots, \xi_i, 0, \dots, 0)$ , insbesondere  $z_0 = x$ ,  $z_n = x + \xi$ .

Nach dem Mittelwertsatz gibt es Punkte  $y_j = x + (\xi_1, \dots, \xi_{j-1}, \theta_j \xi_j, 0, \dots, 0)$

mit  $0 \leq \theta_j \leq 1$ ,  $j = 1, \dots, n$ , sodass  $f(z_j) - f(z_{j-1}) = D_j f(y_j) \cdot \xi_j$  für  $j = 1, \dots, n$

$$\|y_j - x\|_\infty < r$$

$$\begin{aligned} \Rightarrow f(x + \xi) - f(x) &= f(z_n) - f(z_0) = \sum_{j=1}^n \left( f(z_j) - f(z_{j-1}) \right) = \sum_{j=1}^n D_j f(y_j) \cdot \xi_j \\ \Rightarrow \varphi(\xi) &= \sum_{j=1}^n D_j f(y_j) \cdot \xi_j - \sum_{j=1}^n D_j f(x) \cdot \xi_j = \sum_{j=1}^n \left( D_j f(y_j) - D_j f(x) \right) \cdot \xi_j \\ \Rightarrow |\varphi(\xi)| &\leq \sum_{j=1}^n |D_j f(y_j) - D_j f(x)| \cdot |\xi_j| < \varepsilon \cdot \sum_{j=1}^n |\xi_j| = \varepsilon \cdot \|\xi\|_1 \end{aligned}$$

□

**Definition 1.** Sei  $U$  offen in  $\mathbb{R}^n$  und  $f : U \rightarrow \mathbb{R}^m$  eine Abbildung. Sei  $x \in U$ .

Dann heißt  $f$  differenzierbar an der Stelle  $x$ , wenn  $f$  an der Stelle  $x$  partiell differenzierbar ist und wenn gilt: Setzt man

$$\varphi(\xi) := \left( f(x + \xi) - f(x) \right) - Df(x) \cdot \xi$$

für alle  $\xi \in \mathbb{R}^n$  mit  $x + \xi \in U$ , so ist  $\lim_{\xi \rightarrow 0} \frac{\varphi(\xi)}{\|\xi\|} = 0$

**Bemerkung 2.** Für  $n = m = 1$  ist das, das Übliche.

**Bemerkung 3.** Satz 1 besagt: Wenn  $f$  partiell differenzierbar ist und alle partiellen Ableitungen stetig sind, so ist  $f$  differenzierbar.

**Bemerkung 4.** „Differenzierbarkeit“ ist ein ziemlich komplizierter Begriff und nicht so wichtig wie  $C^1$  oder  $C^\infty$ .

$C^1$  heißt:  $f$  ist partiell differenzierbar und alle partiellen Ableitungen sind stetig.

$$C^\infty \Rightarrow \dots \Rightarrow C^2 \Rightarrow C^1 \Rightarrow \text{differenzierbar} \Rightarrow \begin{cases} \Rightarrow \text{partiell differenzierbar} \\ \Rightarrow \text{stetig} \end{cases}$$

**Satz 2.** Sei  $U$  offen in  $\mathbb{R}^n$  und  $f : U \rightarrow \mathbb{R}^m$  sei differenzierbar. Dann ist  $f$  stetig.

**Lemma 1.** Sei  $A$  eine reelle  $m \times n$ -Matrix. Definiere  $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$  durch  $f(x) := A \cdot x$ .

a) Es gibt ein  $\alpha \geq 0$  mit  $\|A \cdot x\| \leq \alpha \cdot \|x\| \quad \forall x \in \mathbb{R}^n$

b)  $f$  ist stetig.

*Beweis.* des Lemmas. Sei  $A = (a_{ij})$  und  $\alpha := \max_{i,j} |a_{ij}|$ .

$$\begin{aligned} \text{a) Ist } x = (x_1, \dots, x_n)^T \in \mathbb{R}^n, \text{ so ist } \left| \sum_{j=1}^n a_{ij} \cdot x_j \right| &\leq \sum_{j=1}^n |a_{ij}| \cdot |x_j| \leq \alpha \cdot \sum_{j=1}^n |x_j| = \alpha \cdot \|x\|_1 \text{ für } i = 1, \dots, m \\ \Rightarrow \|A \cdot x\|_\infty &\leq \alpha \cdot \|x\|_1 \quad \forall x \in \mathbb{R}^n \end{aligned}$$

b) Sind  $x, y \in \mathbb{R}^n$ , so ist  $\|f(x) - f(y)\|_\infty = \|A \cdot x - A \cdot y\|_\infty = \|A \cdot (x - y)\|_\infty \stackrel{a)}{\leq} \alpha \cdot \|x - y\|_1$

□

*Beweis.* von Satz 2. Sei  $x \in U$ . Sei  $(\xi_k)$  eine Nullfolge in  $\mathbb{R}^n$  mit  $x + \xi_k \in U \quad \forall k, \xi_k \neq 0$ .

Zu zeigen:  $\lim_{k \rightarrow \infty} f(x + \xi_k) = f(x)$ .

Es ist  $f(x + \xi_k) - f(x) = Df(x) \cdot \xi_k + \varphi(\xi_k)$ .

Nach dem Lemma ist  $\lim_{k \rightarrow \infty} Df(x) \cdot \xi_k = 0$ . Für großes  $k$  ist  $\|\varphi(\xi_k)\| \leq \|\xi_k\|$ .

Also  $\lim_{k \rightarrow \infty} \left( f(x + \xi_k) - f(x) \right) = 0$ .

Folgerung aus Satz 1 und Satz 2: Ist  $f$  partiell differenzierbar und sind die partiellen Ableitungen von  $f$  stetig, so ist  $f$  stetig, also von der Klasse  $C^1$  (d.h. stetig differenzierbar).

Allgemeiner gilt: Ist  $f$   $k$ -mal partiell differenzierbar und sind die partiellen Ableitungen der Ordnung  $k$  stetig, so ist  $f$  von der Klasse  $C^k$ . □

**Beispiel 1.** Sei  $A$  eine reelle  $m \times n$ -Matrix und definiere  $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$  durch  $f(x) := A \cdot x$ .

$$f = \begin{pmatrix} f_1 \\ \vdots \\ f_m \end{pmatrix} \text{ mit } f_i(x) = \sum_{j=1}^n a_{ij} \cdot x_j.$$

Dann ist  $f$  partiell differenzierbar mit  $Df(x) = A \quad \forall x \in \mathbb{R}^n$ .

Nach Satz 1 folgt:  $f$  ist differenzierbar.

Alle höheren partiellen Ableitungen von  $f$  sind 0

**Satz 3.** Sei  $U$  offen in  $\mathbb{R}^n$ , sei  $f : U \rightarrow \mathbb{R}^m$  eine Funktion und  $x \in U$ .

Es gebe eine  $m \times n$ -Matrix  $A$ , sodass gilt: Definiert man

$$\varphi(\xi) := f(x + \xi) - f(x) - A \cdot \xi$$

für alle  $\xi \in \mathbb{R}^n$  mit  $x + \xi \in U$ , so ist  $\lim_{\xi \rightarrow 0} \frac{\varphi(\xi)}{\|\xi\|} = 0$ .

Dann ist  $f$  an der Stelle  $x$  differenzierbar mit  $Df(x) = A$ .

*Beweis.* Sei  $e_1 := (1, 0, 0, \dots, 0)^T$ ,  $e_2 := (0, 1, 0, \dots, 0)^T, \dots, e_n := (0, \dots, 0, 1)^T$ .

Wir nehmen  $m = 1$  an, also  $A = (a_1, \dots, a_n)$ . Betrachte die Funktion  $F_i(t) := f(x_1, \dots, x_{i-1}, t, x_{i+1}, \dots, x_n)$ .

Um zu zeigen, dass  $f$  an der Stelle  $x$  partiell differenzierbar ist und die richtige Ableitung hat, muss man zeigen:

$$\lim_{h \rightarrow 0} \frac{1}{h} (F_i(x_i + h) - F_i(x_i)) = a_i$$

$$\begin{aligned} \frac{1}{h} (F_1(x_i + h) - F_i(x_i)) &= \frac{1}{h} \left( f(x + h e_i) - f(x) \right) \\ &= \frac{1}{h} \left( f(x + h e_i) - f(x) - A \cdot (h e_i) \right) + A \cdot e_i \\ &= \frac{1}{h} \varphi(h e_i) + a_i & \|\cdot\| = \|\cdot\|_1, \|\cdot\|_2 \text{ oder } \|\cdot\|_\infty \\ &= \pm \frac{\varphi(h e_i)}{\|h e_i\|} + a_i \longrightarrow 0 + a_i = a_i \quad \text{für } h \rightarrow 0 \end{aligned}$$

□

## Anschauliche Bedeutung der Differenzierbarkeit und der Ableitung:

Man möchte beliebige (differenzierbare) Funktionen  $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$  möglichst gut mittels linearer Abbildungen  $\mathbb{R}^n \rightarrow \mathbb{R}^m$  approximieren.

Jede lineare Abbildungen von  $\mathbb{R}^n$  in  $\mathbb{R}^m$  ist von der Form  $\xi \mapsto A \cdot \xi$  mit einer  $m \times n$ -Matrix  $A$ .

Präziser: Man möchte die Funktion  $\xi \mapsto f(x + \xi) - f(x)$  bei festgehaltenem  $x$  für kleine  $\xi$  möglichst gut durch eine lineare Abbildung  $A$  approximieren. Wenn dies in dem Sinn geht, dass

$$\frac{f(x + \xi) - f(x) - A \cdot \xi}{\|\xi\|} \rightarrow 0 \quad \text{für } \xi \rightarrow 0$$

so ist  $f$  an der Stelle  $x$  differenzierbar mit Ableitung  $A =: Df(x)$ .

$A$  ist diejenige lineare Abbildung:  $\mathbb{R}^n \rightarrow \mathbb{R}^m$ , deren Graph die „Tangentialebene“ an den Graph der Funktion  $\xi \mapsto f(x + \xi) - f(x)$  an der Stelle  $(0, 0)$  ist.

**Bemerkung 5.** Sei  $U$  offen in  $\mathbb{R}^n$  und seien  $f, g : U \rightarrow \mathbb{R}^m$  differenzierbar. Sei  $\lambda \in \mathbb{R}$ . Dann sind  $f + g$  und  $\lambda \cdot f$  differenzierbar mit  $(D(f + g))(x) = Df(x) + Dg(x)$  und  $(D(\lambda \cdot f))(x) = \lambda \cdot Df(x)$

**Satz 4. Kettenregel.** Sei  $U$  offen in  $\mathbb{R}^n$  und  $V$  offen in  $\mathbb{R}^m$ .

Seien  $g : U \rightarrow \mathbb{R}^m$  und  $f : V \rightarrow \mathbb{R}^p$  differenzierbar mit  $g(U) \subseteq V$ . Dann ist  $f \circ g : U \rightarrow \mathbb{R}^p$  differenzierbar und

$$(D(f \circ g))(x) = Df(g(x)) \cdot Dg(x) \quad \forall x \in U$$

(Daher bedeutet  $\cdot$  die Matrizenmultiplikation.)

Sind  $f$  und  $g$  von der Klasse  $C^k$  mit  $k \in \mathbb{N} \cup \{\infty\}$ , so ist auch  $f \circ g$  von der Klasse  $C^k$ .

Wenn man diese Formel mit den üblichen partiellen Ableitungen schreiben will, ist es zweckmäßig, die Variablen im  $\mathbb{R}^n$  und  $\mathbb{R}^m$  mit verschiedenen Buchstaben zu bezeichnen. Seien etwa  $x_1, \dots, x_n$  die Variablen im  $\mathbb{R}^n$  und  $y_1, \dots, y_m$  die Variablen im  $\mathbb{R}^m$ . Ist  $p = 1$ , so gilt:

$$\frac{\partial(f \circ g)}{\partial x_i}(x) = \sum_j^m \frac{\partial f}{\partial y_j}(g(x)) \cdot \frac{\partial g_j}{\partial x_i}(x) \quad \forall x = (x_1, \dots, x_n) \in U$$

*Beweis.* Sei  $x \in U$ . Wir wollen Satz 3 benutzen und deswegen zeigen: Schreibt man:

$$f \circ g(x + \xi) - f \circ g(x) - Df(g(x)) \cdot Dg(x) \cdot \xi =: \chi(\xi)$$

so ist  $\lim_{\xi \rightarrow 0} \frac{\chi(\xi)}{\|\xi\|} = 0$ .

Sei  $0 < \varepsilon < 1$ . Es gibt  $r > 0$ , sodass für alle  $\xi \in \mathbb{R}^n$  mit  $\|\xi\| < r$  und alle  $\eta \in \mathbb{R}^m$  mit  $\|\eta\| < r$  gilt: Ist

$$\begin{aligned} \varphi(\xi) &= g(x + \xi) - g(x) - Dg(x) \cdot \xi, \\ \psi(\eta) &= f(g(x) + \eta) - f(g(x)) - Df(g(x)) \cdot \eta \end{aligned}$$

, so ist  $\|\varphi(\xi)\| \leq \varepsilon \cdot \|\xi\|$  und  $\|\psi(\eta)\| \leq \varepsilon \cdot \|\eta\|$

Nach dem Lemma gibt es  $a, b \geq 0$  mit

$$\begin{aligned} \|Dg(x) \cdot \xi\| &\leq a \cdot \|\xi\| \quad \text{und} \quad \|Df(g(x)) \cdot \eta\| \leq b \cdot \|\eta\| \quad \forall \xi, \eta \\ \|Dg(x) \cdot \xi + \varphi(\xi)\| &\leq a \cdot \|\xi\| + \varepsilon \cdot \|\xi\| = (a + \varepsilon) \cdot \|\xi\| \quad \text{für } \|\xi\| \leq r \end{aligned}$$

$$\begin{aligned} f \circ g(x + \xi) &= f(g(x) + Dg(x) \cdot \xi + \varphi(\xi)) \\ &= f(g(x)) + Df(g(x)) \cdot \left( Dg(x) \cdot \xi + \varphi(\xi) \right) + \psi(Dg(x) \cdot \xi + \varphi(\xi)) \\ &= f(g(x)) + Df(g(x)) \cdot Dg(x) \cdot \xi + Df(g(x)) \cdot \varphi(\xi) + \psi\left( Dg(x) \cdot \xi + \varphi(\xi) \right) \end{aligned}$$

$$\Rightarrow \chi(\xi) = Df(g(x)) \cdot \varphi(\xi) + \psi(Dg(x) \cdot \xi + \varphi(\xi)).$$

Ist  $\|\xi\| \leq \frac{r}{a+1}$ , so ist

$$\begin{aligned} \|\xi(Dg(x) \cdot \xi + \varphi(\xi))\| &\leq \varepsilon \cdot \|Dg(x) \cdot \xi + \varphi(\xi)\| \leq \varepsilon \cdot (a+1) \cdot \|\xi\| \\ \|Df(g(x)) \cdot \varphi(\xi)\| &\leq b \cdot \|\varphi(\xi)\| \leq \varepsilon \cdot b \cdot \|\xi\|, \text{ also} \end{aligned}$$

$$\|\chi(\xi)\| \leq \varepsilon \cdot b \cdot \|\xi\| + \varepsilon \cdot (a+1) \cdot \|\xi\| = \varepsilon \cdot (a+b+1) \cdot \|\xi\|.$$

$$\text{Also } \frac{\|\chi(\xi)\|}{\|\xi\|} \leq \varepsilon \cdot (a+b+1) \text{ f\"ur } \|\xi\| \leq \frac{r}{a+1}, \text{ also } \lim_{\xi \rightarrow 0} \frac{\|\chi(\xi)\|}{\|\xi\|} = 0. \quad \square$$

**Definition 2.** Sei  $U$  offen in  $\mathbb{R}^n$  und  $f : U \rightarrow \mathbb{R}$  sei eine Funktion. Sei  $x \in U$  und  $v \in \mathbb{R}^n$ . Dann ist  $U_v := \{t \in \mathbb{R} \mid x + t \cdot v \in U\}$  eine offene Teilmenge von  $\mathbb{R}$  mit  $0 \in U_v$ .

Sei  $F_v : U_v \rightarrow \mathbb{R}$  definiert durch  $F_v(t) := f(x + t \cdot v)$ .

Wenn  $F_v$  an der Stelle 0 differenzierbar ist, so heit

$$D_v f(x) := F'_v(0) = \lim_{t \rightarrow 0} \frac{1}{t} (f(x + t \cdot v) - f(x))$$

die Richtungsableitung von  $f$  im Punkt  $x$  in Richtung  $v$

**Beispiel 2.**  $D_{e_i} f(x) = D_i f(x) = \frac{\partial f}{\partial x_i}(x)$

**Bezeichnung 1.** Seien  $v, w \in \mathbb{R}^n$ ,  $v = (v_1, \dots, v_n)^T$ ,  $w = (w_1, \dots, w_n)^T$ , so sei

$$\begin{aligned} \langle v, w \rangle &= v^T \cdot w = \sum_{j=1}^n v_j \cdot w_j \in \mathbb{R} \\ \|v\|_2 &= \sqrt{\langle v, v \rangle} \end{aligned}$$

Cauchy-Schwarz:  $|\langle v, w \rangle| \leq \|v\|_2 \cdot \|w\|_2$ .

Leicht zu sehen. Dabei gilt das Gleichheitszeichen genau dann, wenn die beiden Vektoren  $v, w$  reelle Vielfache voneinander sind.

**Satz 5.** Sei  $U$  offen in  $\mathbb{R}^n$  und sei  $f : U \rightarrow \mathbb{R}$  differenzierbar. Sei  $x \in U$  und  $v \in \mathbb{R}^n$ . Dann existiert die Richtungsableitung von  $f$  im Punkt  $x$  an der Stelle  $v$ , und es ist

$$D_v f(x) = \langle v, \nabla f(x) \rangle$$

*Beweis.* Definiere  $g : U_v \rightarrow \mathbb{R}^n$  durch  $g(t) := x + t \cdot v$ ,  $g = (g_1, \dots, g_n)^T$ ,  $g_j(t) = x_j + t \cdot v_j$ .

Dann ist  $F_v = f \circ g$ .  $g$  ist differenzierbar mit  $g'_j(0) = v_j$ .

Nach Kettenregel ist  $F_v$  differenzierbar mit

$$F'_v(0) = \sum_{j=1}^n D_j f(x) \cdot g'_j(0) = \sum_{j=1}^n D_j f(x) \cdot v_j = \langle \nabla f(x), v \rangle$$

□

## Anschauliche Interpretation des Gradienten

Sei  $\nabla f(x) \neq 0$  und sei  $v \in \mathbb{R}^n$  mit  $\|v\|_2 = 1$ . Dann ist

$$|D_v f(x)| = |\langle v, \nabla f(x) \rangle| \stackrel{\text{C.S.}}{\leq} \|\nabla f(x)\|_2, \quad \text{dabei gilt}$$

Gleichheit genau dann, wenn  $v$  ein reelles Vielfaches von  $\nabla f(x)$  ist.

Das heit:  $\text{grad } f(x)$  gibt die Richtung des steilsten Anstiegs von  $f$  an.

## §6 Mittelwertsatz und Taylor-Formel

Mittelwertsatz der Analysis I: Sei  $f : [a, b] \rightarrow \mathbb{R}$  stetig; auf  $]a, b[$  sei  $f$  differenzierbar. Dann gibt es ein  $\xi \in ]a, b[$  mit

$$(*) \quad f(b) - f(a) = f'(\xi) \cdot (b - a)$$

In dieser Form lässt sich  $(*)$  nicht auf vektorwertige Funktionen verallgemeinern.

Betrachte  $f : \mathbb{R} \rightarrow \mathbb{R}^2$ ,  $f(x) := \begin{pmatrix} \cos x \\ \sin x \end{pmatrix} \Rightarrow Df(x) = \begin{pmatrix} -\sin x \\ \cos x \end{pmatrix}$

Es ist  $f(0) = \begin{pmatrix} 1 \\ 0 \end{pmatrix} = f(2\pi)$ , aber  $Df(x) \neq 0 \quad \forall x \in \mathbb{R}$ .

Aus  $(*)$  folgt: Wenn es ein  $M \in \mathbb{R}_{\geq 0}$  gibt mit  $|f'(\xi)| \leq M \quad \forall \xi \in ]a, b[$ , so ist

$$(**) \quad |f(b) - f(a)| \leq M \cdot |b - a|$$

Dies lässt sich auf vektorwertige Funktionen verallgemeinern.

**Definition 1.** Sei  $[a, b]$  ein kompaktes Intervall in  $\mathbb{R}$  und sei  $f : [a, b] \rightarrow \mathbb{R}^n$  eine Funktion,  $f = (f_1, \dots, f_n)^T$ . Wenn alle  $f_i$  integrierbar sind, so heißt  $f$  integrierbar. Dann schreibt man:

$$\int_a^b f(t) dt = \left( \int_a^b f_1(t) dt, \dots, \int_a^b f_n(t) dt \right)^T \in \mathbb{R}^n$$

**Bemerkung 1.** Sei  $\|\cdot\|$  eine der Normen  $\|\cdot\|_1, \|\cdot\|_2, \|\cdot\|_\infty$ . Ist  $f$  integrierbar, so ist auch  $\|f\|$  integrierbar und

$$(***) \quad \left\| \int_a^b f(t) dt \right\| \leq \int_a^b \|f(t)\| dt$$

**Satz 1. „Mittelwertsatz“.** Sei  $U$  offen in  $\mathbb{R}^n$  und  $f : U \rightarrow \mathbb{R}^m$  von der Klasse  $C^1$ .

Seien  $x, \xi \in \mathbb{R}^n$ , sodass die Strecke  $\{x + t \cdot \xi \mid 0 \leq t \leq 1\}$  in  $U$  enthalten ist.

Dann gibt es ein  $M_{\geq 0}$  mit

$$\|Df(x + t \cdot \xi) \cdot v\| \leq M \cdot \|v\| \quad \forall t \in [0, 1], \quad \forall v \in \mathbb{R}^n$$

und für jedes derartige  $M$  gilt:

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

*Beweis.*

1. Im Lemma von §5 haben wir gesehen: Ist  $A = (a_{ij})$  eine  $m \times n$ -Matrix und ist  $\alpha := \max_{i,j} |a_{ij}|$ , so ist

$$\|A \cdot v\|_\infty \leq \alpha \cdot \|v\|_1 \quad \forall v \in \mathbb{R}^n$$

Für  $t \in [0, 1]$  schreiben wir  $Df(x + t \cdot \xi) = (a_{ij}(t))$ .

Nach Kettenregel sind die Funktionen  $a_{ij} : [0, 1] \rightarrow \mathbb{R}$  stetig.

Es gibt also ein  $M \geq 0$  mit  $|a_{ij}(t)| \leq M \quad \forall t \in [0, 1], \quad \forall i, j$ .

Dann ist  $\|Df(x + t \cdot \xi) \cdot v\|_\infty \leq M \cdot \|v\|_1 \quad \forall t, \forall v$

2. Nun sei ein solches  $M$  gegeben. Es gibt ein offenes Intervall  $I \supset [0, 1]$ , sodass

$$g(t) := x + t \cdot \xi \in U \quad \forall t \in I$$

Dann ist  $g$  von der Klasse  $C^\infty$  mit  $Dg(t) = \xi \quad \forall t \in I$  (also  $n \times 1$ -Matrix).

Definiere  $\varphi : I \rightarrow \mathbb{R}^m$  durch  $\varphi(t) := f(x + t \cdot \xi) = f(g(t))$ .

Nach Kettenregel ist  $\varphi$  von der Klasse  $C^1$  und

$$\begin{aligned} D\varphi(t) &= Df(x + t \cdot \xi) \cdot Dg(t) = Df(x + t \cdot \xi) \cdot \xi \\ \Rightarrow f(x + \xi) - f(x) &= \varphi(1) - \varphi(0) = \int_0^1 D\varphi(t) dt = \int_0^1 Df(x + t \cdot \xi) \cdot \xi dt \\ \Rightarrow \|f(x + \xi) - f(x)\| &\stackrel{(***)}{\leq} \int_0^1 \|Df(x + t \cdot \xi) \cdot \xi\| dt \leq \int_0^1 M \cdot \|\xi\| dt = M \cdot \|\xi\| \end{aligned}$$

□

**Satz 2.** Sei  $U$  offen in  $\mathbb{R}^n$  mit folgender Eigenschaft: zu je zwei Punkten von  $U$  gebe es einen Streckenzug, der diese beiden Punkte verbindet und ganz in  $U$  liegt.

Sei  $f : \mathbb{R} \rightarrow \mathbb{R}^n$  partiell differenzierbar mit  $Df(x) = 0 \quad \forall x \in U$ . Dann ist  $f$  konstant.

*Beweis.* Seien  $x, x + \xi \in U$ . Zu zeigen:  $f(x) = f(x + \xi)$ . O.B.d.A. gehöre die Strecke zwischen  $x$  und  $x + \xi$  zu  $U$ . Dann können wir Satz 1 mit  $M = 0$  anwenden. □

## Erinnerung an die Taylor-Formel für Funktionen von einer Veränderlichen

Sei  $I$  ein offenes Intervall in  $\mathbb{R}$  und sei  $f : I \rightarrow \mathbb{R}$  von der Klasse  $C^{k+1}$ .

Seien  $a, x \in I$ . Dann ist

$$f(x) = \sum_{m=0}^k \frac{f^{(m)}(a)}{m!} \cdot (x-a)^m + R_{k+1}(x) \quad \text{mit}$$

$$R_{k+1}(x) = \frac{1}{k!} \cdot \int_a^x (x-t)^k \cdot f^{(k+1)}(t) dt$$

**Bezeichnung 2.** Sei  $U$  offen in  $\mathbb{R}^n$  und  $f : U \rightarrow \mathbb{R}$  sei eine Funktion.

a) Die Elemente  $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{N}_0^n$  heißen  $n$ -Multi-Indizes.

Für ein solches  $\alpha$  sei:

$$\begin{aligned} |\alpha| &:= \alpha_1 + \dots + \alpha_n \\ \alpha! &:= \alpha_1! \cdot \dots \cdot \alpha_n! \end{aligned}$$

b) Sei  $f$  von der Klasse  $C^{|\alpha|}$  für ein  $\alpha \in \mathbb{N}_0^n$ . Dann sei

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

Dabei ist  $D_i^0 f := f$  zu setzen.

c) Ist  $x = (x_1, \dots, x_n) \in \mathbb{R}^n$  und  $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{N}_0^n$ , so sei  $x^\alpha := x_1^{\alpha_1} \cdot \dots \cdot x_n^{\alpha_n}$

**Lemma 1.** Sei  $U$  offen in  $\mathbb{R}^n$  und  $f : U \rightarrow \mathbb{R}$  sei von der Klasse  $C^k$ . Seien  $x, \xi \in \mathbb{R}^n$ , sodass  $x$  und  $x + \xi$  in  $U$  liegen.

Sei  $I$  ein offenes Intervall mit  $I \supset [0, 1]$ , sodass  $x + t \cdot \xi \in U \quad \forall t \in I$ . Sei  $g(t) := x + t \cdot \xi$ .

Definiere  $\varphi : I \rightarrow \mathbb{R}$  durch  $\varphi(t) := f(g(t)) = f(x + t \cdot \xi)$ .

a)  $\varphi$  ist von der Klasse  $C^k$  mit

$$\varphi^{(k)}(t) = \sum_{i_1, \dots, i_k=1}^n D_{i_k} \dots D_{i_1} f(x + t \cdot \xi) \cdot \xi_{i_1} \dots \xi_{i_k}$$

b)

$$\varphi^{(k)}(t) = \sum_{\substack{\alpha \in \mathbb{N}_0^n \\ |\alpha|=k}} \frac{k!}{\alpha!} \cdot D^\alpha f(x + t \cdot \xi) \cdot \xi^\alpha$$

*Beweis.*

a) Nach Kettenregel ist  $\varphi'(t) = Df(x + t \cdot \xi) \cdot \xi = \sum_{i=1}^n D_i f(x + t \cdot \xi) \cdot \xi_i$

Dann macht man vollständige Induktion.

- b) Nach dem Satz von H.A. Schwarz gilt: Kommt unter den Zahlen  $i_1, \dots, i_k$  die Zahl genau  $\alpha_1$ -mal, die Zahl 2 genau  $\alpha_2$ -mal, ..., die Zahl  $n$  genau  $\alpha_n$ -mal vor, so ist  $D_{i_k} \dots D_{i_1} f(x + t \cdot \xi) \cdot \xi_{i_1} \dots \xi_{i_k} = D^\alpha f(x + t \cdot \xi) \cdot \xi^\alpha$ .

Die Anzahl der  $k$ -Tupel  $(i_1, \dots, i_k)$ , die zum gleichen Multi-Index  $\alpha$  führen, ist  $\frac{k!}{\alpha!}$ :

Dann ist  $(i_1, \dots, i_k)$  ein solches  $k$ -Tupel, so entstehen alle anderen derartigen  $k$ -Tupel, indem man die Zahlen  $i_1, \dots, i_k$  permutiert.

Die Gruppe der Permutationen von  $k$  Objekten enthält genau  $k!$  Elemente.

Aber nicht alle so erhaltenen  $k$ -Tupel sind verschieden: Wenn man z.B. die  $\alpha_1$  Stellen, an denen eine 1 steht, untereinander permutiert, so ändert sich nichts usw. Insgesamt erhält man

$$\frac{k!}{\alpha_1! \dots \alpha_n!} = \frac{k!}{\alpha!} \quad \text{verschiedene solche } k\text{-Tupel}$$

□

**Satz 3. Taylor-Formel.** Sei  $U$  offen in  $\mathbb{R}^n$  und  $f : U \rightarrow \mathbb{R}$  von der Klasse  $C^{k+1}$ :

- a) Seien  $x, \xi \in \mathbb{R}^n$ , sodass die Strecke zwischen  $x$  und  $x + \xi$  in  $U$  liegt. Dann ist:

$$f(x + \xi) = \sum_{|\alpha| \leq k} \frac{D^\alpha f(x)}{\alpha!} \cdot \xi^\alpha + (k+1) \cdot \sum_{|\alpha|=k+1} \int_0^1 (1-t)^k \cdot \frac{D^\alpha f(x + t\xi)}{\alpha!} \cdot \xi^\alpha dt$$

- b) Ist  $x \in U$  und definiert man

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

für alle  $\xi \in \mathbb{R}^n$  mit  $x + \xi \in U$ , so ist

$$\lim_{\xi \rightarrow 0} \frac{R(\xi)}{\|\xi\|^{k+1}} = 0$$

*Beweis.*

- a) Wir benutzen die Bezeichnungen von Lemma und wenden die Taylor-Formel mit einer Veränderlichen auf die Funktion  $\varphi$  an:

$$\begin{aligned} f(x + \xi) &= \varphi(1) = \sum_{m=0}^k \frac{\varphi^{(m)}(0)}{m!} + \frac{1}{k!} \cdot \int_0^1 (1-t)^k \cdot \varphi^{(k+1)}(t) dt \\ &\stackrel{\text{L}}{=} \sum_{|\alpha| \leq k} \frac{D^\alpha f(\xi)}{\alpha!} \cdot \xi^\alpha + \frac{1}{k!} \cdot \int_0^1 (1-t)^k \cdot \left( \sum_{|\alpha|=k+1} \frac{(k+1)!}{\alpha!} \cdot D^\alpha f(x + t\xi) \cdot \xi^\alpha \right) dt \\ &= \sum_{|\alpha| \leq k} \frac{D^\alpha f(x)}{\alpha!} \cdot \xi^\alpha + (k+1) \cdot \sum_{|\alpha|=k+1} \int_0^1 (1-t)^k \frac{D^\alpha f(x + t\xi)}{\alpha!} \cdot \xi^\alpha dt \end{aligned}$$

- b) Es ist  $\int_0^1 (1-t)^k dt = \frac{1}{k+1}$  (\*).

$$\begin{aligned} R(\xi) &\stackrel{\text{a)}}{=} (k+1) \cdot \sum_{|\alpha|=k+1} \int_0^1 (1-t)^k \cdot \frac{D^\alpha f(x + t\xi)}{\alpha!} \cdot \xi^\alpha dt - \sum_{|\alpha|=k+1} \frac{D^\alpha f(x)}{\alpha!} \cdot \xi^\alpha \\ &\stackrel{(*)}{=} (k+1) \cdot \sum_{|\alpha|=k+1} \int_0^1 \frac{(1-t)^k}{\alpha!} \left( D^\alpha f(x + t\xi) - D^\alpha f(x) \right) \cdot \xi^\alpha dt \\ &\Rightarrow |R(\xi)| \leq (k+1) \cdot \sum_{|\alpha|=k+1} \int_0^1 \frac{(1-t)^k}{\alpha!} \cdot |D^\alpha f(x + t\xi) - D^\alpha f(x)| \cdot |\xi^\alpha| dt \end{aligned}$$



Da  $f$  von der Klasse  $C^{k+1}$  ist, gibt es für jedes  $\varepsilon > 0$  ein  $r > 0$ , sodass

$$|D^\alpha f(x + t\xi) - D^\alpha f(x)| < \varepsilon$$

für  $\|\xi\| < r$ ,  $0 \leq t \leq 1$ ,  $|\alpha| = k + 1$ .

$$|\xi^\alpha| = |\xi_1|^{\alpha_1} \dots |\xi_n|^{\alpha_n} \leq \|\xi\|_\infty^{\alpha_1 + \dots + \alpha_n} = \|\xi\|_\infty^{|\alpha|} = \|\xi\|_\infty^{k+1} \quad \text{für } |\alpha| = k + 1$$

Also gilt für  $\|\xi\| \leq r$ :

$$\begin{aligned} |R(\xi)| &\leq (k+1) \cdot \sum_{|\alpha|=k+1} \int_0^1 \frac{(1-t)^k}{\alpha!} \cdot \varepsilon \cdot \|\xi\|_\infty^{k+1} dt \\ &= (k+1) \cdot \sum_{|\alpha|=k+1} \frac{1}{\alpha!} \cdot \varepsilon \cdot \|\xi\|_\infty^{k+1} \cdot \int_0^1 (1-t)^k dt = \sum_{|\alpha|=k+1} \frac{1}{\alpha!} \cdot \varepsilon \cdot \|\xi\|_\infty^{k+1} \end{aligned}$$

Also ist  $\lim_{\xi \rightarrow 0} \frac{|R(\xi)|}{\|\xi\|_\infty^{k+1}} = 0$ .

□

## §7 Extremwerte und kritische Stellen

**Definition 1.** Sei  $X$  ein metrischer Raum,  $f : X \rightarrow \mathbb{R}$  eine Funktion,  $x_0 \in X$ . Man sagt:  $f$  besitzt in  $x_0$  ein lokales Maximum (bzw. lokales Minimum), wenn es eine Umgebung  $U$  von  $x_0$  gibt, sodass

$$\begin{aligned} f(x_0) &\geq f(x) \quad \forall x \in U \\ (\text{bzw. } f(x_0) &\leq f(x) \quad \forall x \in U) \end{aligned}$$

$f$  besitzt in  $x_0$  ein striktes lokales Maximum (bzw. striktes lokales Minimum), wenn es eine Umgebung  $U$  von  $x_0$  gibt, sodass:

$$\begin{aligned} f(x_0) &> f(x) \quad \forall x \in U \setminus \{x_0\} \\ (\text{bzw. } f(x_0) &< f(x) \quad \forall x \in U \setminus \{x_0\}) \end{aligned}$$

Wenn  $f$  in  $x_0$  ein (striktes) lokales Maximum oder Minimum besitzt, so sagt man:  $f$  besitzt in  $x_0$  ein (striktes) lokales Extremum.

**Definition 2.** Sei  $U$  offen in  $\mathbb{R}^n$  und sei  $f : U \rightarrow \mathbb{R}$  partiell differenzierbar. Sei  $x_0 \in U$ . Wenn  $\text{grad } f(x_0) = 0$ , so heißt  $x_0$  eine kritische Stelle von  $f$ .

**Satz 1.** Sei  $U$  offen in  $\mathbb{R}^n$ , sei  $f : U \rightarrow \mathbb{R}$  partiell differenzierbar.  $f$  besitze in  $x_0$  ein lokales Extremum. Dann ist  $x_0$  eine kritische Stelle von  $f$ .

*Beweis.* Sei  $x_0 = (x_1, \dots, x_n)$ . Für  $i = 1, \dots, n$  betrachte die Funktion  $F_i$ , die definiert ist durch

$$F_i(t) := (x_1, \dots, x_{i-1}, t, x_{i+1}, \dots, x_n)$$

in einer Umgebung von  $x_i$  in  $\mathbb{R}$ . Die Funktion  $F_i$  besitzt an der Stelle  $x_i$  ein lokales Extremum, also ist  $F_i'(x_i) = 0$  nach Analysis I.

Nach Definition der partiellen Ableitung ist  $F_i'(x_i) = D_i f(x_0)$ . □

**Beispiel 1.** Definiere  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$  durch  $f(x, y) = x^2 - y^2$ . („Sattelfläche“)

$$\text{grad } f(x, y) = (2x, -2y)$$

Also  $\text{grad } f(x, y) = (0, 0) \Leftrightarrow (x, y) = (0, 0)$ .  $(0, 0)$  ist die einzige kritische Stelle von  $f$ .

Auch in  $(0, 0)$  besitzt  $f$  kein lokales Extremum, denn in jeder Umgebung von  $(0, 0)$  nimmt  $f$  positive (z.B. in  $(x, 0)$ ,  $x \neq 0$ ) und negative Werte (z.B. in  $(0, y)$ ,  $y \neq 0$ ) an.

**Definition 3.** Sei  $U$  offen in  $\mathbb{R}^n$  und sei  $f : U \rightarrow \mathbb{R}$  von der Klasse  $C^2$ . Für  $x \in U$  sei  $Hf(x)$  die  $n \times n$ -Matrix mit  $Hf(x) = (D_i D_j f(x))_{1 \leq i, j \leq n}$ .  $Hf(x)$  heißt die Hesse-Matrix von  $f$  an der Stelle  $x$ .

**Bemerkung 1.** Nach dem Satz von Schwarz ist  $D_i D_j f(x) = D_j D_i f(x)$ , d.h. die Matrix  $Hf(x)$  ist symmetrisch.

Im obigen Beispiel 1 ist  $Hf(x, y) = \begin{pmatrix} 2 & 0 \\ 0 & -2 \end{pmatrix} \quad \forall (x, y) \in \mathbb{R}^2$

**Lemma 1.** Sei  $U$  offen in  $\mathbb{R}^n$ ,  $f : U \rightarrow \mathbb{R}$  von der Klasse  $C^2$  und  $x \in U$ . Schreibt man

$$f(x + \xi) = f(x) + \langle \text{grad } f(x), \xi \rangle + \frac{1}{2} \langle Hf(x) \cdot \xi, \xi \rangle + R(\xi)$$

für alle  $\xi \in \mathbb{R}^n$  mit  $x + \xi \in U$ , so ist  $\lim_{\xi \rightarrow 0} \frac{R(\xi)}{\|\xi\|^2} = 0$

*Beweis.* Die Taylor-Formel besagt: Schreibt man

$$f(x + \xi) = f(x) + \sum_{|\alpha|=1} \frac{D^\alpha f(x)}{\alpha!} \cdot \xi^\alpha + \sum_{|\alpha|=2} \frac{D^\alpha f(x)}{\alpha!} \cdot \xi^\alpha + R(\xi),$$

so ist  $\lim_{\xi \rightarrow 0} \frac{R(\xi)}{\|\xi\|^2} = 0$ . Wir müssen zeigen:

1.  $\sum_{|\alpha|=1} \frac{D^\alpha f(x)}{\alpha!} \cdot \xi^\alpha = \langle \text{grad } f(x), \xi \rangle$
2.  $\sum_{|\alpha|=2} \frac{D^\alpha f(x)}{\alpha!} \cdot \xi^\alpha = \frac{1}{2} \langle Hf(x) \cdot \xi, \xi \rangle$

Die  $\alpha \in \mathbb{N}_0^n$  mit  $|\alpha| = 2$  sind von der Form (2A)  $\alpha = (0, \dots, 0, 2, 0, \dots, 0) = 2e_i$  oder von der Form (2B)  $\alpha = (0, \dots, 0, 1, 0, \dots, 0, 1, 0, \dots, 0) = e_i + e_j$  mit  $i < j$ .

Im Fall (2A) ist  $\frac{D^\alpha f(x)}{\alpha!} \cdot \xi^\alpha = \frac{D_i^2 f(x)}{2} \cdot \xi_i^2$ ; im Fall (2B) ist  $\frac{D^\alpha f(x)}{\alpha!} \cdot \xi^\alpha = \frac{D_i D_j f(x)}{1} \cdot \xi_i \xi_j$ . Also

$$\begin{aligned}
\sum_{|\alpha|=2} \frac{D^\alpha f(x)}{\alpha!} \cdot \xi^\alpha &= \frac{1}{2} \sum_{i=1}^n D_i^2 f(x) \cdot \xi_i^2 + \sum_{1 \leq i < j \leq n} D_i D_j f(x) \cdot \xi_i \xi_j \\
&= \frac{1}{2} \sum_{i=1}^n D_i^2 f(x) \cdot \xi_i^2 + \frac{1}{2} \sum_{i \neq j} D_i D_j f(x) \xi_i \xi_j \\
&= \frac{1}{2} \sum_{i,j} D_i D_j f(x) \cdot \xi_i \xi_j \\
&= \frac{1}{2} \langle Hf(x) \cdot \xi, \xi \rangle
\end{aligned}$$

□

**Definition 4.** Sei  $A$  eine symmetrische reelle  $n \times n$ -Matrix.

- $A$  heißt positiv definit, falls  $\langle A\xi, \xi \rangle > 0 \quad \forall \xi \in \mathbb{R}^n \setminus \{0\}$
- $A$  heißt negativ definit, falls  $\langle A\xi, \xi \rangle < 0 \quad \forall \xi \in \mathbb{R}^n \setminus \{0\}$
- $A$  heißt indefinit, wenn es ein  $\xi \in \mathbb{R}^n$  und ein  $\eta \in \mathbb{R}^n$  gibt mit:

$$\langle A\xi, \xi \rangle > 0 \quad \text{und} \quad \langle A\eta, \eta \rangle < 0$$

**Beispiel 2.**

- Eine  $1 \times 1$ -Matrix, also eine reelle Zahl  $a$ , ist positiv definit, genau dann, wenn  $a \cdot \xi^2 > 0$  für alle  $\xi \in \mathbb{R} \setminus \{0\}$ , also genau dann, wenn  $a > 0$ .  
(Entsprechend für negativ definit)
- Die Matrix  $A = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$  ist indefinit, denn  $\begin{cases} \langle Ae_1, e_1 \rangle = \langle 2e_1, e_1 \rangle = 2 > 0 \\ \langle Ae_2, e_2 \rangle = \langle -2e_2, e_2 \rangle = -2 < 0 \end{cases}$

**Lemma 2.** Ist  $A$  eine positiv definit symmetrische  $n \times n$ -Matrix, so gibt es ein  $a \in \mathbb{R}_{>0}$ , sodass

$$\langle Ax, x \rangle \geq a \cdot \|x\|^2 \quad \forall x \in \mathbb{R}^n$$

*Beweis.* Sei  $q(x) := \langle Ax, x \rangle$ . Dann ist  $q$  eine Abbildung von  $\mathbb{R}^n \rightarrow \mathbb{R}$ . Sie ist stetig.

Sei  $S := \{x \in \mathbb{R}^n \mid \|x\| = 1\}$ . Dann ist  $S$  eine beschränkte, abgeschlossene Teilmenge von  $\mathbb{R}^n$ ; also nimmt  $q$  auf  $S$  sein Minimum an. Es gibt also ein  $a \in \mathbb{R}_{>0}$ , sodass  $q(x) \geq a \quad \forall x \in S$

Wir zeigen, dass  $q(x) \geq a \cdot \|x\|^2 \quad \forall x \in \mathbb{R}^n$ . Das ist klar für  $x = 0$ . Ist  $x \neq 0$ , so ist  $\frac{x}{\|x\|} \in S$ , also

$$a \leq q\left(\frac{x}{\|x\|}\right) = \left\langle A \frac{x}{\|x\|}, \frac{x}{\|x\|} \right\rangle = \frac{1}{\|x\|} \cdot \langle Ax, x \rangle = \frac{1}{\|x\|^2} \cdot q(x)$$

□

**Satz 2.** Sei  $U$  offen in  $\mathbb{R}^n$  und  $f : U \rightarrow \mathbb{R}$  von der Klasse  $C^2$ .  
Sei  $x_0 \in U$  eine kritische Stelle von  $f$ .

- a) Ist  $Hf(x_0)$  positiv definit, so besitzt  $f$  an der Stelle  $x_0$  ein striktes lokales Minimum.
- b) Ist  $Hf(x_0)$  negativ definit, so besitzt  $f$  an der Stelle  $x_0$  ein striktes lokales Maximum.
- c) Ist  $Hf(x_0)$  indefinit, so besitzt  $f$  an der Stelle  $x_0$  kein lokales Extremum.

*Beweis.* von a). Definiere  $q : \mathbb{R}^n \rightarrow \mathbb{R}$  durch  $q(\xi) = \langle Hf(x)\xi, \xi \rangle$ . Da  $x_0$  eine kritische Stelle von  $f$  ist, folgt aus Lemma 1: Schreibt man  $f(x_0 + \xi) = f(x_0) + \frac{1}{2}q(\xi) + R(\xi)$ , so ist  $\lim_{\xi \rightarrow 0} \frac{R(\xi)}{\|\xi\|^2} = 0$ .

Schreibt man  $\rho(\xi) := \frac{R(\xi)}{\|\xi\|^2}$ , so ist  $\lim_{\xi \rightarrow 0} \rho(\xi) = 0$  und

$$\frac{f(x_0 + \xi) - f(x_0)}{\|\xi\|^2} = \frac{1}{2} \underbrace{q\left(\frac{\xi}{\|\xi\|}\right)}_{> \frac{1}{2}a \in \mathbb{R}_{>0} \text{ nach L. 2} \quad \forall \xi \neq 0} + \underbrace{\rho(\xi)}_{\xrightarrow{0} \text{ für } \xi \rightarrow 0}$$

Deswegen ist  $\frac{f(x_0 + \xi) - f(x_0)}{\|\xi\|^2} > 0$  für hinreichend kleines  $\xi$ , und  $f$  besitzt in  $x_0$  ein striktes lokales Minimum. □

*Beweis.* von b). Wende a) auf  $-f$  an. □

*Beweis.* von c). Seien  $\xi, \eta \in \mathbb{R}^n$  mit  $q(\xi) > 0, q(\eta) < 0$ . Ist  $\lambda \in \mathbb{R}_{>0}$ , so ist

$$q\left(\frac{\lambda\xi}{\|\lambda\xi\|}\right) = q\left(\frac{\xi}{\|\xi\|}\right) = \frac{1}{\|\xi\|^2} \cdot q(\xi) > 0$$

Setzt man in  $\frac{f(x_0 + \xi) - f(x_0)}{\|\xi\|^2} = \frac{1}{2}q\left(\frac{\xi}{\|\xi\|}\right) + \rho(\xi)$  für  $\xi$  nun  $\lambda\xi$  mit kleinem  $\lambda$  ein, so ist die rechte Seite  $> 0$ ;

setzt man für  $\xi$  aber  $\lambda\eta$  mit kleinem  $\lambda$  ein, so ist die rechte Seite  $< 0$ .

Also besitzt  $f$  in  $x_0$  kein lokales Extremum. □

### Exkurs über Lineare Algebra:

Sei  $A$  eine reelle  $n \times n$ -Matrix. Eine komplexe Zahl  $\lambda$  heißt Eigenwert von  $A$ , wenn es ein  $x \in \mathbb{C}^n \setminus \{0\}$ , sodass  $Ax = \lambda x$ .

Das charakteristische Polynom  $\chi_A$  von  $A$  ist das Polynom vom *grad*  $n$ , das definiert ist durch  $\chi_A(t) :=$

$$\det(tI - A), \quad I = \begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{pmatrix}.$$

- Die Eigenwerte von  $A$  sind genau die Nullstellen von  $\chi_A$
- Ist  $A$  symmetrisch, so sind alle Eigenwerte von  $A$  reell und es gilt:

$A$  positiv definit  $\Leftrightarrow$  alle Eigenwerte von  $A$  sind positiv

$A$  negativ definit  $\Leftrightarrow$  alle Eigenwerte von  $A$  sind negativ

$A$  indefinit  $\Leftrightarrow$  es gibt mindestens einen positiven und einen negativen Eigenwert von  $A$

**Kriterium von Horwitz:** Sei  $A = (a_{ij})$  eine symmetrische  $n \times n$ -Matrix. Für  $k = 1, \dots, n$  sei:

$$\Delta_k := \det \begin{pmatrix} a_{11} & \dots & a_{1k} \\ \vdots & & \vdots \\ a_{k1} & \dots & a_{kk} \end{pmatrix}$$

$A$  positiv definit  $\Leftrightarrow \Delta_k > 0$  für  $k = 1, \dots, n$

$A$  negativ definit  $\Leftrightarrow (-1)^k \Delta_k > 0$  für  $k = 1, \dots, n$

**Beispiel 3.** Betrachte  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$  mit  $f(x, y) = x^3 - y^3$ .

$$\text{grad } f(x, y) = (3x^2, -3y^2)$$

$\text{grad } f(x, y) = 0 \Leftrightarrow (x, y) = (0, 0)$ .  $(0, 0)$  ist also die einzige kritische Stelle von  $f$ .

$Hf(x, y) = \begin{pmatrix} 6x & 0 \\ 0 & -6y \end{pmatrix}$ , also  $Hf(0, 0) = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$ , weder positiv, noch negativ definit, noch indefinit.

$(0, 0)$  ist kein lokales Extremum, weil in jeder Umgebung von  $(0, 0)$  die Funktion  $f$  sowohl positive als auch negative Werte annimmt.

**Beispiel 4.**  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$  mit  $f(x, y) := x^3 + y^3 - 3xy$ .

$$\text{grad } f(x, y) = (3x^2 - 3y, 3y^2 - 3x)$$

$(x, y)$  ist kritische Stelle  $\Leftrightarrow y = x^2$  und  $x = y^2 \Rightarrow x = x^4 \Rightarrow x = 0$  oder  $x = 1$ .

Die kritischen Stellen von  $f$  sind  $(0, 0), (1, 1)$ .

$Hf(x, y) = \begin{pmatrix} 6x & -3 \\ -3 & 6y \end{pmatrix}$ , also  $Hf(0, 0) = \begin{pmatrix} 0 & -3 \\ 3 & 0 \end{pmatrix}$  und  $Hf(1, 1) = \begin{pmatrix} 6 & -3 \\ -3 & 6 \end{pmatrix}$

$Hf(1, 1)$  ist positiv definit nach Horwitz  $\Rightarrow f$  besitzt striktes lokales Minimum in  $(1, 1)$ .

$Hf(0, 0)$  ist indefinit, denn  $\chi(t) = \det \begin{pmatrix} t & 3 \\ 3 & t \end{pmatrix} = t^2 - 9 = (t - 3)(t + 3)$ .

## §8 Beispiele und Problemstellungen

Sei  $U$  offen in  $\mathbb{R}^2$  und sei  $f : U \rightarrow \mathbb{R}$  stetig.

Sei  $I$  ein offenes Intervall in  $\mathbb{R}$  und sei  $\varphi : I \rightarrow \mathbb{R}$  differenzierbar mit

- $(x, \varphi(x)) \in U \quad \forall x \in I$
- $\varphi'(x) = f(x, \varphi(x)) \quad \forall x \in I$

Dann heißt  $\varphi$  eine Lösung der Differenzialgleichung

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

Man nennt  $(*)$  eine explizite gewöhnliche Differenzialgleichung 1. Ordnung.

(„1. Ordnung“, weil keine Ableitung höherer Ordnung vorkommen),

„gewöhnlich“, weil keine partiellen Ableitungen vorkommen,

„explizit“, weil die Gleichung nach  $y'$  aufgelöst ist.)

**Beispiel 1.** Sei  $U = \mathbb{R}^2$  und  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$  gegeben durch  $f(x, y) := x$ . Das heißt wir betrachten die Differenzialgleichung:  $y' = x$ .

Ihre Lösungen sind die differenzierbaren Funktionen  $\varphi : I \rightarrow \mathbb{R}$  ( $I$  ein offenes Intervall) mit  $\varphi'(x) = x \quad \forall x \in I$ .

Alle derartigen Funktionen  $\varphi$  sind von der Form  $\varphi(x) = \frac{1}{2}x^2 + c$  mit einer festen Zahl  $c$ .

Allgemeiner: Sei  $J$  ein offenes Intervall,  $g : I \rightarrow \mathbb{R}$  stetig,  $U = J \times \mathbb{R}$ .

Definiere  $f : J \times \mathbb{R} \rightarrow \mathbb{R}$  durch  $f(x, y) := g(x)$ , das heißt wir betrachten die Differenzialgleichung

$$y' = g(x)$$

Eine Lösung dieser Differenzialgleichung ist eine Funktion  $\varphi : I \rightarrow \mathbb{R}$  mit

- $(x, \varphi(x)) \in J \times \mathbb{R} \quad \forall x \in I \quad (\text{d.h. } I \subseteq J)$
- $\varphi'(x) = g(x) \quad \forall x \in I \quad (\text{d.h. } \varphi \text{ ist Stammfunktion von } g \text{ auf } I)$

Es gibt unendlich viele auf ganz  $J$  definierte Lösungen von  $y' = g(x)$ , die sich alle um eine additive Konstante unterscheiden.

Jede weitere Lösung von  $y' = g(x)$  ist die Einschränkung einer auf ganz  $J$  definierten Lösung auf ein kleineres Intervall.

**Beispiel 2.** Sei  $U = \mathbb{R}^2$  und  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$  gegeben durch  $f(x, y) := y$ , d.h. wir haben die Differenzialgleichung

$$y' = y$$

Ist  $c \in \mathbb{R}$  und definiert man  $\varphi_c : \mathbb{R} \rightarrow \mathbb{R}$  durch  $\varphi_c(x) = c \cdot e^x$ , so ist  $\varphi_c$  eine Lösung von  $y' = y$ .

Sei  $\varphi : I \rightarrow \mathbb{R}$  eine Lösung von  $y' = y$ . Definiere  $g : I \rightarrow \mathbb{R}$  durch  $g(x) = \varphi(x) \cdot e^{-x}$ .

$$g'(x) = \varphi'(x) \cdot e^{-x} = \varphi(x)e^{-x} = (\varphi'(x) - \varphi(x)) \cdot e^{-x} = 0, \text{ weil } \varphi' = \varphi$$

Deswegen existiert ein  $c \in \mathbb{R}$  mit  $g(x) = c \quad \forall x \in I$ , also  $\varphi(x) = c \cdot e^x = \varphi_c(x) \quad \forall x \in I$ .

Es gibt also unendlich viele auf ganz  $\mathbb{R}$  definierte Lösungen, die sich alle um eine multiplikative Konstante unterscheiden; jede weitere Lösung ist die Einschränkung einer solchen Lösung auf ein kleineres Intervall.

In der allgemeinen Situation sei  $(x_0, y_0) \in U$  gegeben. Wir suchen oft Lösungen  $\varphi : I \rightarrow \mathbb{R}$  von  $(*)$ , für die gilt:  $x_0 \in I$  und  $\varphi$  „erfüllt die Anfangsbedingung  $\varphi(x_0) = y_0$ “.

Im Beispiel 1 und 2 gilt: Ist  $(x_0, y_0) \in U$  beliebig, so gibt es genau eine Lösung  $\varphi : J \rightarrow \mathbb{R}$  (Beispiel 1) bzw.  $\varphi : \mathbb{R} \rightarrow \mathbb{R}$  (Beispiel 2) mit  $\varphi(x_0) = y_0$ .

**Beispiel 3.** Sei  $U = \mathbb{R}^2$ ,  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$  gegeben durch  $f(x, y) = y^2$ , d.h. wir betrachten die Differentialgleichung  $y' = y^2$ . Definiert man  $\varphi_0 : \mathbb{R} \rightarrow \mathbb{R}$  durch  $\varphi_0(x) := 0 \quad \forall x \in \mathbb{R}$ , so ist  $\varphi_0$  eine Lösung. Sei  $\varphi : I \rightarrow \mathbb{R}$  eine Lösung mit  $\varphi(x) \neq 0 \quad \forall x \in I$ .

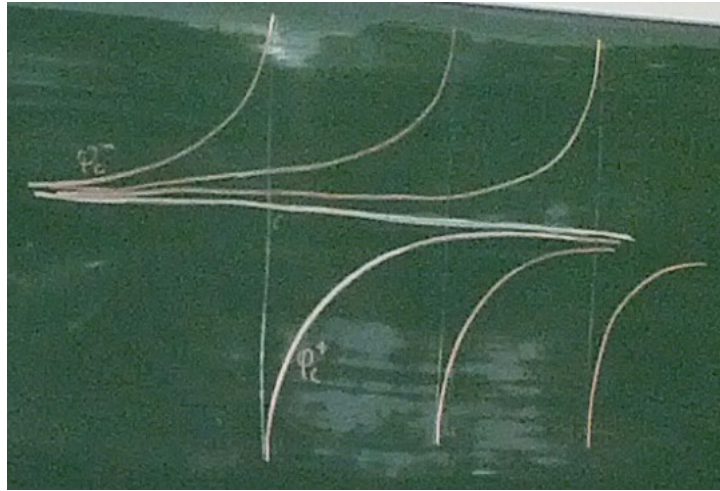
$$g(x) := -\frac{1}{\varphi(x)} \Rightarrow g'(x) = \frac{\varphi'(x)}{\varphi(x)^2} = 1 \quad \forall x \in I$$

$\Rightarrow$  existiert  $c \in \mathbb{R}$  mit  $g(x) = x - c \quad \forall x \in I$ , also  $\varphi(x) = \frac{1}{c-x} \quad \forall x \in I$ .

Deswegen ist  $c \notin I$ . Definiere für  $c \in \mathbb{R}$  die Funktion  $\varphi_c^+ : ]c, \infty[ \rightarrow \mathbb{R}$  durch  $\varphi_c^+ := \frac{1}{c-x}$

und  $\varphi_c^- : ]-\infty, c[ \rightarrow \mathbb{R}$  durch  $\varphi_c^-(x) := \frac{1}{c-x}$ .

Dann sind  $\varphi_c^+, \varphi_c^-$  Lösungen von  $y' = y^2$ , die nicht zu Lösungen auf ganz  $\mathbb{R}$  fortgesetzt werden können.



Wir haben gesehen: Ist  $\varphi : I \rightarrow \mathbb{R}$  eine Lösung mit  $\varphi(x) \neq 0 \quad \forall x \in I$ , so gibt es ein  $c \in \mathbb{R} \setminus I$  mit  $\varphi(x) = \frac{1}{c-x} \quad \forall x \in I$ .

Weitere Lösungen gibt es nicht: Sei  $\varphi : I \rightarrow \mathbb{R}$  eine Lösung, und es gebe  $x_1, x_2 \in I$  mit  $\varphi(x_1) \neq 0$ ,  $\varphi(x_2) = 0$ .

O.B.d.A. sei  $x_2 < x_1$ . Sei  $x_3 := \max\{x \in [x_1, x_2] \mid \varphi(x) = 0\}$ , d.h.  $x_3$  ist die größte Nullstelle von  $\varphi$ , die kleiner als  $x_1$  ist.

Auf  $]x_3, x_1[$  ist  $\varphi$  Lösung ohne Nullstellen, also  $\varphi(x) = \frac{1}{c-x}$  und  $c \neq ]x_3, x_1[$ .

Es ist  $\lim_{x \searrow x_3} \varphi(x) \neq 0$ ,  $\nexists$  Widerspruch.

Insgesamt gilt: Für jedes  $(x_0, y_0) \in U = \mathbb{R}^2$  gibt es genau eine Lösung  $\varphi$  von  $y' = y^2$  mit Anfangsbedingung  $\varphi(x_0) = y_0$  und maximalem Definitionsbereich.

Das ist die „normale Situation“ bei gewissen expliziten Differentialgleichungen 1. Ordnung.

**Beispiel 4.** Sei  $U = \mathbb{R}^2$  und  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$  gegeben durch  $f(x, y) = 3y^{\frac{2}{3}} = 3 \cdot \sqrt[3]{y^2}$ .

Für  $c \in \mathbb{R}$  definiere  $\varphi_c : \mathbb{R} \rightarrow \mathbb{R}$  durch  $\varphi_c(x) := (x-c)^3$ . Dann ist  $\varphi'_c = 3(x-c)^2 = 3\varphi_c(x)^{\frac{2}{3}}$ .

Für  $a, b \in \mathbb{R} \cup \{-\infty, \infty\}$  mit  $a < b$  definiere  $\varphi_{a,b} : \mathbb{R} \rightarrow \mathbb{R}$  durch

$$\varphi_{a,b}(x) := \begin{cases} (x-a)^3 & \text{für } x \leq a \\ 0 & \text{für } a < x < b \\ (x-b)^3 & \text{für } x \geq b \end{cases}$$

Dann ist  $\varphi_{a,b}$  differenzierbar und ist Lösung von  $y' = 3y^{\frac{2}{3}}$ .

Für jede Anfangsbedingung gibt es also unendlich viele verschiedene Lösungen!

**Beispiel 5. DGL mit getrennten Variablen.** Seien  $I, J$  zwei offene Intervalle und seien  $g : I \rightarrow \mathbb{R}$  und  $h : J \rightarrow \mathbb{R}$  stetig mit  $h(y) \neq 0 \quad \forall y \in J$ .

Sei  $U := I \times J$  und definiere  $f : U \rightarrow \mathbb{R}$  durch  $f(x, y) = g(x) \cdot h(y)$  d.h. betrachte die DGL  $y' = g(x) \cdot h(y)$ .

Heuristische Vorgehensweise:

$$\begin{aligned}\frac{dy}{dx} &= g(x) \cdot h(y) \\ \frac{dy}{h(y)} &= g(x) dx \\ \int \frac{dy}{h(y)} &= \int g(x) dx + c\end{aligned}$$

Die linke Seite ist Funktion von  $y$ , die rechte Seite Funktion von  $x$ .

Löse die Gleichung nach  $y$  auf, auf diese Weise erhält man die Lösungen von  $y' = g(x) \cdot h(y)$

Konkretes Beispiel:  $y' = x(y^2 + 1)$ .  $I = \mathbb{R} = J$ ,  $U = \mathbb{R}^2$ ,  $g(x) = x$ ,  $h(y) = y^2 + 1$

$$\begin{aligned}\int \frac{dy}{y^2 + 1} &= \int x dx + c \\ \arctan y &= \frac{x^2}{2} + c \\ y &= \tan\left(\frac{x^2}{2} + c\right) \quad \text{Beachte: tan ist nicht auf ganz } \mathbb{R} \text{ definiert!}\end{aligned}$$

Exakte Vorgehensweise bei der DGL mit getrennten Variablen:

Sei  $(x_0, y_0) \in U = I \times J$ .

Definiere  $G : I \rightarrow \mathbb{R}$  durch  $G(x) := \int_{x_0}^x g(t) dt$  und  $H : J \rightarrow \mathbb{R}$  durch  $H(y) := \int_{y_0}^y \frac{dt}{h(t)}$ .

**Behauptung 1.** Es gibt ein offenes Intervall  $I'$  mit  $x_0 \in I' \subseteq I$  und eine eindeutig bestimmte Lösung  $\varphi : I' \rightarrow \mathbb{R}$  von  $y' = g(x) \cdot h(y)$  mit  $\varphi(x_0) = y_0$ , und es ist  $H(\varphi(x)) = G(x) \quad \forall x \in I'$

*Beweis.*  $G(x_0) = 0$ ,  $H(y_0) = 0$ ,  $G'(x) = g(x) \quad \forall x \in I$ ,  $H'(y) = \frac{1}{h(y)} \neq 0 \quad \forall y \in J$ .

Also ist  $H$  monoton und differenzierbar  $\Rightarrow H$  besitzt eine Umkehrfunktion  $K : J' \rightarrow \mathbb{R}$  mit  $0 \in J'$  und  $K(0) = y_0$ ,  $K$  ist differenzierbar und  $K(H(y)) = y \quad \forall y \in J$ .

$$K'(z) = \frac{1}{H'(K(z))} \quad \forall z \in J'$$

Es gibt ein offenes Intervall  $I'$  mit  $x_0 \in I'$  und  $G(I') \subseteq J'$ .

Setze  $\varphi(x) := K(G(x)) \quad \forall x \in I'$ . Dann ist  $\varphi$  differenzierbar,  $\varphi(x_0) = K(G(x_0)) = K(0) = y_0$ .

$$\varphi'(x) = K'(G(x)) \cdot G'(x) = \frac{1}{H'(K(G(x)))} \cdot g(x) = \frac{1}{H'(\varphi(x))} \cdot g(x) = h(\varphi(x)) \cdot g(x)$$

Deswegen ist  $\varphi$  eine Lösung mit der Anfangsbedingung  $\varphi(x_0) = y_0$ .

Umgekehrt: Ist  $\varphi : I' \rightarrow \mathbb{R}$  Lösung mit  $\varphi(x_0) = y_0$ , so ist

$$\begin{aligned}\varphi'(x) &= g(x) \cdot h(\varphi(x)) \Rightarrow \frac{\varphi'(x)}{h(\varphi(x))} = g(x) \\ \Rightarrow G(x) &= \int_{x_0}^x g(t) dt = \int_{x_0}^x \frac{\varphi'(t)}{h(\varphi(t))} dt \stackrel{\text{Subst. } s = \varphi(t)}{=} \int_{y_0}^{\varphi(x)} \frac{ds}{h(s)} = H(\varphi(x))\end{aligned}$$

Verallgemeinerung der Ausgangssituation:

Sei  $U$  offen in  $\mathbb{R} \times \mathbb{R}^n$  und sei  $f : U \rightarrow \mathbb{R}^n$  stetig,  $f = (f_1, \dots, f_n)$  mit  $f_i : U \rightarrow \mathbb{R}$ .

Wir betrachten das DGL-System

$$(**) \begin{cases} y_1' = f_1(x, y_1, \dots, y_n) \\ \vdots \\ y_n' = f_n(x, y_1, \dots, y_n) \end{cases}$$



Eine Lösung von (\*\*) ist eine Funktion  $\varphi = (\varphi_1, \dots, \varphi_n) : I \rightarrow \mathbb{R}^n$ , wobei  $I$  ein offenes Intervall ist und die  $\varphi_i : I \rightarrow \mathbb{R}$  differenzierbar sind mit:

- $(x, \varphi_1(x), \dots, \varphi_n(x)) = (x, \varphi(x)) \in U$
- $\varphi'_i(x) = f_i(x, \varphi_1(x), \dots, \varphi_n(x))$  für  $i = 1, \dots, n$ ,  
also  $\varphi'(x) = f(x, \varphi(x)) \quad \forall x \in I$

Statt (\*\*) schreibt man wieder einfach  $y' = f(x, \varphi(x))$  und nennt dies eine gewöhnliche explizite DGL 1. Ordnung.  $\square$

Eine gewöhnliche explizite DGL 1. Ordnung ist eine Gleichung der Form

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

wobei  $f : U \rightarrow \mathbb{R}^n$  eine gegebene stetige Funktion ist,  $U$  offen in  $\mathbb{R} \times \mathbb{R}^n$ .

Eine Lösung dieser DGL ist eine differenzierbare Funktion  $\varphi : I \rightarrow \mathbb{R}^n$ , wobei  $I$  ein offenes Intervall in  $\mathbb{R}$  ist und wobei gilt:

- $(x, \varphi(x)) \in U \quad \forall x \in I$
- $\varphi'(x) = f(x, \varphi(x)) \quad \forall x \in I$

Ist  $(x_0, y_0) \in U$  gegeben und ist  $\varphi(x_0) = y_0$ , so sagt man, dass  $\varphi$  der Anfangsbedingung  $y(x_0) = y_0$  genügt.

Wir werden 3 wichtige Sätze zeigen, die wir ab sofort benutzen:

**Extremwertsatz von Peano:** Sei  $U$  offen in  $\mathbb{R} \times \mathbb{R}^n$  und sei  $f : U \rightarrow \mathbb{R}^n$  stetig.

Ist  $(x_0, y_0) \in U$ , so gibt es eine Lösung  $\varphi : I \rightarrow \mathbb{R}^n$  der DGL  $y' = f(x, y)$  mit  $x_0 \in I$  und  $\varphi(x_0) = y_0$ .

**Definition 1.** Seien  $X, Y$  metrische Räume und  $f : X \rightarrow Y$  eine Abbildung.

- $f$  heißt Lipschitz-stetig wenn es ein  $L \in \mathbb{R}_{\geq 0}$  gibt mit  $d(f(x), f(y)) \leq L \cdot d(x, y) \quad \forall x, y \in X$
- $f$  heißt lokal Lipschitz-stetig, wenn es für jedes  $x_0 \in X$  eine Umgebung  $U$  von  $x_0$  gibt, sodass die Einschränkung  $f|_U$  Lipschitz-stetig ist.

**Bemerkung 1.** Lipschitz-stetig  $\Rightarrow$  lokal Lipschitz-stetig  $\Rightarrow$  stetig.

**Beispiel 6.** Definiere  $f : \mathbb{R} \rightarrow \mathbb{R}$  durch  $f(x) = x^2$ . Dann ist  $f$  lokal Lipschitz-stetig, aber nicht Lipschitz-stetig.

**Bemerkung 2.** Ist  $X$  offen in  $\mathbb{R}^n$  und  $f : X \rightarrow \mathbb{R}^n$  von der Klasse  $C^1$ , so ist  $f$  lokal Lipschitz-stetig nach dem Mittelwertsatz.

**Definition 2.** Seien  $X, Y, Z$  metrische Räume, sei  $U \subseteq X \times Y$  und  $f : U \rightarrow Z$  eine Abbildung

- $f$  heißt Lipschitz-stetig im 2. Argument, wenn es ein  $L \in \mathbb{R}_{\geq 0}$  gibt mit

$$d(f(x, y), f(x, \tilde{y})) \leq L \cdot d(y, \tilde{y}) \quad \forall (x, y), (x, \tilde{y}) \in U$$

- $f$  heißt lokal Lipschitz-stetig im 2. Argument, wenn es für jedes  $(x_0, y_0) \in U$  eine Umgebung  $V$  von  $(x_0, y_0)$  in  $U$  gibt, sodass  $f|_V$  Lipschitz-stetig im 2. Argument ist.

**Bemerkung 3.** Sei  $U$  offen in  $\mathbb{R}^n \times \mathbb{R}^m$  und sei  $f : U \rightarrow \mathbb{R}^k$  eine Abbildung. Wir bezeichnen die partiellen Ableitungen von  $f$  (falls sie existieren) mit  $\frac{\partial f}{\partial x_1}, \dots, \frac{\partial f}{\partial x_n}, \frac{\partial f}{\partial y_1}, \dots, \frac{\partial f}{\partial y_m}$ .

Wenn  $\frac{\partial f}{\partial y_1}, \dots, \frac{\partial f}{\partial y_m}$  existieren und wenn sie stetige Abbildungen von  $U$  in  $\mathbb{R}^k$  sind, so ist  $f$  lokal Lipschitz-stetig im 2. Argument.

## Lokale Existenz- und Eindeigkeitssatz von Picard-Lindelöf

Sei  $U$  offen in  $\mathbb{R} \times \mathbb{R}^n$  und  $f : U \rightarrow \mathbb{R}^n$  sei stetig und lokal Lipschitz-stetig im 2. Argument. Sei  $(x_0, y_0) \in U$ .

Dann gibt es eine Lösung  $\varphi : I \rightarrow \mathbb{R}^n$  von  $y' = f(x, y)$  mit  $x_0 \in I$  und:

- $\varphi(x_0) = y_0$
- Ist  $\psi : J \rightarrow \mathbb{R}^n$  eine Lösung von  $y' = f(x, y)$  mit  $x_0 \in J$  und  $\psi(x_0) = y_0$ , so ist  $J \subseteq I$  und  $\psi = \varphi|_J$ .

(Insbesondere ist  $\varphi$  eindeutig bestimmt.) **Globaler Existenz- und Eindeigkeitssatz.** Sei  $I$  ein offenes Intervall in  $\mathbb{R}$  und  $U := I \times \mathbb{R}^n$ . Sei  $f : U \rightarrow \mathbb{R}^n$  stetig.

Für jedes kompakte Teilintervall  $K$  von  $I$  sei  $f|_{(K \times \mathbb{R}^n)}$  Lipschitz-stetig im 2. Argument. Sei  $(x_0, y_0) \in U$ .

Dann gibt es eine eindeutige bestimmte Lösung  $\varphi : I \rightarrow \mathbb{R}^n$  von  $y' = f(x, y)$  mit  $\varphi(x_0) = y_0$ .

**Beispiel 7. Homogene lineare DGL.** Sei  $I$  ein offenes Intervall und  $a : I \rightarrow \mathbb{R}$  stetig. Definiere  $f : I \times \mathbb{R} \rightarrow \mathbb{R}$  durch  $f(x, y) := a(x) \cdot y$ . Die DGL  $y' = f(x, y)$  lautet also:

$$y' = a(x) \cdot y$$

Dann sind die Voraussetzungen des globalen Existenz- und Eindeigkeitssatz erfüllt.

Sei  $K$  ein kompaktes Teilintervall von  $I$ . Weil  $a$  stetig ist, gibt es ein  $M_{\geq 0}$  mit  $|a(x)| \leq M \quad \forall x \in K$ .

Ist  $x \in K$  und sind  $y, \tilde{y} \in \mathbb{R}$ , so ist  $|f(x, y) - f(x, \tilde{y})| = |a(x)y - a(x)\tilde{y}| = M \cdot |y - \tilde{y}|$ .

Also ist  $f$  Lipschitz-stetig im 2. Argument auf  $K \times \mathbb{R}$ .

Für jedes  $(x_0, y_0) \in I \times \mathbb{R}$  gibt es also genau eine Lösung  $\varphi : I \rightarrow \mathbb{R}$  von  $y' = a(x)y$  mit  $\varphi(x_0) = y_0$ .

Diese Lösung ist gegeben durch:

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

**Beispiel 8. Lineare DGL.** Sei  $I$  ein offenes Intervall und seien  $a, b : I \rightarrow \mathbb{R}$  stetig. Definiere  $f : I \times \mathbb{R} \rightarrow \mathbb{R}$  durch  $f(x, y) := a(x)y + b(x)$ , d.h. wir betrachten die DGL

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

Nach dem globalen Existenz und Eindeigkeitssatz gibt es für jedes  $(x_0, y_0) \in I \times \mathbb{R}$  genau eine Lösung  $\psi : I \rightarrow \mathbb{R}$  mit  $\psi(x_0) = y_0$ .

Sei  $\varphi(x) := \exp \cdot \int_{x_0}^x a(t) dt$  für  $x \in I$ . Dann ist  $\varphi(x) \neq 0 \quad \forall x \in I$ , und  $\varphi$  ist Lösung der „zugehörigen homogenen DGL“  $y' = a(x)y$ . Weil  $\varphi(x) \neq 0 \quad \forall x$ , gibt es eine  $C^1$ -Funktion

$$u : I \rightarrow \mathbb{R} \text{ mit } \psi = \varphi \cdot u \Rightarrow \psi' = \varphi' \cdot u + \varphi \cdot u' = a\varphi u + \varphi u' = a\psi + \varphi u'$$

$$\psi \text{ ist Lösung von } y' = ay + b \Leftrightarrow \varphi u' = b \Leftrightarrow u' = \frac{b}{\varphi} \Leftrightarrow u(x) = \int_{x_0}^x \frac{b(t)}{\varphi(t)} dt + \text{const.}$$

$$y_0 = \psi(x_0) = \varphi(x_0) \cdot u(x_0) = u(x_0) = \text{const}$$

$$\Rightarrow \boxed{\psi(x) = \varphi(x) \cdot \left( y_0 + \int_{x_0}^x \frac{b(t)}{\varphi(t)} dt \right)}$$

„Methode der Variation der Konstanten“.

## Differenzialgleichung höherer Ordnung

### Beispiel einer DGL 2. Ordnung

$$y'' = y' + xy + 1 \tag{1}$$

Man setzt  $z := y'$ . Dann ist die Gleichung (1) äquivalent mit dem System (2) erster Ordnung:

$$\begin{cases} y' = z \\ z' = z + xy + 1 \end{cases} \quad (2)$$

Allgemein: Sei  $U$  offen in  $\mathbb{R} \times \mathbb{R}^n$  und sei  $f : U \rightarrow \mathbb{R}$  stetig.

Sei  $I$  ein offenes Intervall und  $\varphi : I \rightarrow \mathbb{R}$  eine  $n$ -mal differenzierbare Funktion mit

- $(x, \varphi(x), \varphi'(x), \dots, \varphi^{(n-1)}(x)) \in U \quad \forall x \in I$
- $\varphi^{(n)} = f(x, \varphi(x), \varphi'(x), \dots, \varphi^{(n-1)}(x)) \quad \forall x \in I$

Dann heißt  $\varphi$  eine Lösung der DGL

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

und man nennt  $(***)$  eine explizite gewöhnliche DGL der Ordnung  $n$ .

### Reduktion auf ein System von DGLn 1. Ordnung

Definiere  $F : U \rightarrow \mathbb{R}^n$  durch  $F(x, y_n, \dots, y_{n-1}) := (y_1, \dots, y_{n-1}, f(x, y_0, \dots, y_{n-1}))$ .

Mit  $Y := (y_0, \dots, y_{n-1})^T$  lautet das System  $Y' = F(x, Y)$  ausgeschrieben:

$$\begin{aligned} y_0' &= y_1 \\ y_1' &= y_2 \\ &\vdots \\ y_{n-2}' &= y_{n-1} \\ y_{n-1}' &= f(x, y_0, \dots, y_{n-1}) \end{aligned}$$

1. Ist  $\varphi$  eine Lösung von  $y^{(n)} = f(x, y, y', \dots, y^{(n-1)})$ , so ist  $(\varphi, \varphi', \dots, \varphi^{(n-1)})$  eine Lösung von  $Y' = F(x, Y)$
2. Ist  $\Phi = (\varphi, \varphi_1, \varphi_2, \dots, \varphi_{n-1})$  eine Lösung von  $Y' = F(x, Y)$ , so ist  $\varphi$  eine Lösung von  $y^{(n)} = f(x, y, y', \dots, y^{(n-1)})$

### Folgerung aus dem Lokalen Existenz- und Eindeutigkeitssatz

Sei  $U$  offen in  $\mathbb{R} \times \mathbb{R}^n$  und  $f : U \rightarrow \mathbb{R}$  sei stetig und lokal Lipschitz-stetig im 2. Argument. Sei  $(x_0, y_0, \dots, y_{n-1}) \in U$ .

Dann existiert eine Lösung  $\varphi : I \rightarrow \mathbb{R}$  der DGL  $y^{(n)} = f(x, y, y', \dots, y^{(n-1)})$  mit  $x_0 \in I$  und

•

$$\begin{aligned} \varphi(x_0) &= y_0, \\ \varphi'(x_0) &= y_1, \\ &\vdots \\ \varphi^{(n-1)}(x_0) &= y_{n-1} \end{aligned}$$

„Anfangsbedingung“

- Ist  $\psi : J \rightarrow \mathbb{R}$  Lösung von  $y^n = f(x, y, y', \dots, y^{(n-1)})$  mit  $x_0 \in J$  und  $\psi(x_0) = y_0, \psi'(x_0) = y_1, \psi^{(n-1)}(x_0) = y_{n-1}$ , so ist  $J \subseteq I$  und  $\psi = \varphi|_J$

**Beispiel 9.**  $y'' = y$ . Für  $a, b \in \mathbb{R}$  definiere  $\varphi_{a,b} : \mathbb{R} \rightarrow \mathbb{R}$  durch

$$\varphi_{a,b}(x) := a \cdot \cos(x) + b \cdot \sin(x)$$

Dann ist  $\varphi_{a,b}$  eine Lösung. Das sind Alle Lösungen:

Sei  $(x_0, y_0, y_1) \in \mathbb{R}^3$ . Wir wollen zeigen, dass es  $a, b \in \mathbb{R}$  gibt mit

$$\varphi_{a,b}(x_0) = y_0, \quad \varphi'_{a,b}(x_0) = y_1$$

Dazu ist das folgende Gleichungssystem nach  $a$  und  $b$  aufzulösen:

$$\begin{aligned} a \cdot \cos x_0 + b \cdot \sin x_0 &= y_0, \\ -a \cdot \sin x_0 + b \cdot \cos x_0 &= y_1 \end{aligned}$$

Dieses System besitzt eine Lösung, weil  $\det \begin{pmatrix} \cos x_0 & \sin x_0 \\ -\sin x_0 & \cos x_0 \end{pmatrix} = \cos^2 x_0 + \sin^2 x_0 = 1 \neq 0$

## §9 Lineare Differenzialgleichungen

Ein lineares DGL-System für 2 gesuchte Funktionen ist von der Form

$$(1) \begin{cases} y_1' = \alpha(x) \cdot y_1 + \beta(x) \cdot y_2 + \gamma(x) \\ y_2' = \delta(x) \cdot y_1 + \varepsilon(x) \cdot y_2 + \eta(x) \end{cases}$$

mit gegebenen stetigen Funktionen  $\alpha, \beta, \dots, \eta$ .

Man fasst  $y_1$  und  $y_2$  zu  $y = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$  zusammen und schreibt

$$A(x) := \begin{pmatrix} \alpha(x) & \beta(x) \\ \delta(x) & \varepsilon(x) \end{pmatrix} \in M(2, 2; \mathbb{R}) \quad \text{für jedes } x$$

$$b(x) := \begin{pmatrix} \gamma(x) \\ \eta(x) \end{pmatrix} \in \mathbb{R}^2 \quad \text{für jedes } x$$

Dann lautet (1):

$$(2) \quad y' = A(x) \cdot y + b(x)$$

Es ist zweckmäßig, lineare DGL-Systeme für komplexwertige Funktionen zu betrachten.

**Bemerkung 1.** Ein komplexer normierter Raum ist ein  $\mathbb{C}$ -Vektorraum  $V$  mit einer Abbildung  $\|\cdot\| : V \rightarrow \mathbb{R}$  mit folgenden Eigenschaften

1.  $\|v\| \geq 0 \quad \forall v \in V$
2.  $\|v\| = 0 \Leftrightarrow v = 0$
3.  $\|\lambda \cdot v\| = |\lambda| \cdot \|v\| \quad \forall \lambda \in \mathbb{C}, v \in V$
4.  $\|v + w\| \leq \|v\| + \|w\| \quad \forall v, w \in V$

**Beispiel 1.**  $\mathbb{C}^n$  wird ein komplexer normierter Raum mit

$$\|(z_1, \dots, z_n)\|_\infty := \max\{|z_1|, \dots, |z_n|\},$$

$$\|(z_1, \dots, z_n)\|_1 := |z_1| + \dots + |z_n|,$$

$$\|(z_1, \dots, z_n)\|_2 := (|z_1|^2 + \dots + |z_n|^2)^{\frac{1}{2}}$$

Ein komplexer normierter Raum wird zu einem metrischen Raum mit

$$d(v, w) := \|v - w\|$$

Zwei Normen  $\|\cdot\|, \|\cdot\|'$  auf einem endlich-dimensionalen  $\mathbb{C}$ -Vektorraum  $V$  sind äquivalent im folgenden Sinn. Es gibt Konstanten  $a, A \in \mathbb{R}_{>0}$  mit

$$a \cdot \|v\| \leq \|v\|' \leq A \cdot \|v\| \quad \forall v \in V$$

**Bemerkung 2.** Sei  $\mathbb{K} = \mathbb{R}$  oder  $\mathbb{K} = \mathbb{C}$ . Sei  $M(m, n; \mathbb{K})$  die Menge der  $m \times n$ -Matrizen und Einträgen in  $\mathbb{K}$ . Dann ist  $M(m, n; \mathbb{K})$  die Menge der  $\mathbb{K}$ -linearen Abbildungen von  $\mathbb{K}^n$  in  $\mathbb{K}^m$ . Wir wählen auf  $\mathbb{K}^n$  und  $\mathbb{K}^m$  Normen, die wir beide mit  $\|\cdot\|$  bezeichnen. Ist  $A \in M(m, n; \mathbb{K})$ , so sei

$$\|A\| := \max\{\|A \cdot x\| \mid x \in \mathbb{K}^n \text{ mit } \|x\| = 1\}$$

(Beachte:

- Ist  $V$  ein normierter Raum, so ist  $\|\cdot\| : V \rightarrow \mathbb{R}$  stetig.
- Die Abbildung  $A : \mathbb{K}^n \rightarrow \mathbb{K}^m$  ist stetig.

- Deswegen ist die Abbildung  $x \mapsto \|A \cdot x\|$  von  $\mathbb{K}^n$  in  $\mathbb{R}$  stetig; sie nimmt also auf der beschränkten und abgeschlossenen Teilmenge  $\{x \mid \|x\| = 1\}$  von  $\mathbb{K}^n$  ihr Maximum an.)

Es gilt:

1. Damit wird  $M(m, n; \mathbb{K})$  zu einem normierten Raum.
2. Für alle  $x \in \mathbb{K}^n$  und  $A \in M(m, n; \mathbb{K})$  ist  $\|A \cdot x\| \leq \|A\| \cdot \|x\|$ .

*Beweis.* Das ist klar für  $x = 0$ . Ist  $x \neq 0$ , so sei  $x_0 := \frac{x}{\|x\|}$ . Dann ist  $\|x_0\| = 1$

$$\Rightarrow \|A \cdot x_0\| \leq \|A\|$$

$$\Rightarrow \|A \cdot x\| = \|A \cdot (\|x_0\| \cdot x_0)\| = \|x\| \cdot \|A \cdot x_0\| \leq \|x\| \cdot \|A\|$$

□

Ist  $n = m$ , so nimmt man auf  $\mathbb{K}^n$  und  $\mathbb{K}^m$  dieselbe Norm. Dann gilt:

3. Sind  $A, B \in M(n, n; \mathbb{K})$ , so ist  $\|A \cdot B\| \leq \|A\| \cdot \|B\|$

*Beweis.* Sei  $x \in \mathbb{K}^n$ . Dann ist  $\|AB \cdot x\| \stackrel{(2)}{\leq} \|A\| \cdot \|Bx\| \stackrel{(2)}{\leq} \|A\| \cdot \|B\| \cdot \|x\|$

$$\Rightarrow \underbrace{\max_{\|x\|=1} \|ABx\|}_{\|AB\|} \leq \|A\| \cdot \|B\|$$

□

4. Ist  $X$  ein metrischer Raum und  $f : X \rightarrow M(m, n; \mathbb{K})$  eine Abbildung,

$$\text{also } f(x) = \begin{pmatrix} f_{11}(x) & \dots & f_{1n}(x) \\ \vdots & & \vdots \\ f_{m1}(x) & \dots & f_{m,n}(x) \end{pmatrix} \text{ mit } f_{ij} : X \rightarrow \mathbb{K}, \text{ so gilt:}$$

$$f \text{ ist stetig} \Leftrightarrow \text{alle } f_{ij} \text{ sind stetig}$$

**Definition 1.** Sei  $I$  ein offenes Intervall und seien  $A : I \rightarrow M(n, n; \mathbb{R})$  und  $b : I \rightarrow \mathbb{R}^n$  zwei stetige Abbildungen. Wir definieren  $f : I \times \mathbb{R}^n \rightarrow \mathbb{R}^n$  durch

$$f(x, y) = A(x)y + b(x)$$

Dann heißt die DGL  $y' = f(x, y)$  ein System von linearen DGLn oder kurz eine lineare DGL. Ist  $b = 0$ , so heißt das DGL-System homogen.

Ist  $A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix}$  mit  $a_{ij} : I \rightarrow \mathbb{R}$ , so bedeutet das System

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

ausgeschrieben:

$$\begin{aligned} y_1' &= a_{11}(x)y_1 + \dots + a_{1n}(x)y_n + b_1(x) \\ &\vdots \\ y_n' &= a_{n1}(x)y_1 + \dots + a_{nn}(x)y_n + b_n(x) \end{aligned}$$

**Satz 1.** Sei  $I$  ein offenes Intervall und seien  $A : I \rightarrow M(n, n; \mathbb{R})$  und  $b : I \rightarrow \mathbb{R}^n$  stetig. Sei  $(x_0, y_0) \in I \times \mathbb{R}^n$ . Dann gibt es genau eine Lösung  $\varphi : I \rightarrow \mathbb{R}^n$  des DGL  $y' = A(x)y + b(x)$  mit  $\varphi(x_0) = y_0$ .

(In Zukunft seien Lösungen einer solchen linearen DGL immer auf ganz  $I$  definiert).

*Beweis.* Wir benutzen den Globalen Existenz- und Eindeigkeitssatz: Betrachte  $f : I \times \mathbb{R}^n \rightarrow \mathbb{R}^n$  mit  $f(x, y) = A(x)y + b(x)$ .

$f$  ist stetig. Wir müssen zeigen: Ist  $K$  ein kompaktes Teilintervall von  $I$ , so ist  $f$  Lipschitz-stetig im 2. Argument auf  $K \times \mathbb{R}^n$ .

Weil  $A$  stetig ist, gibt es ein  $M \geq 0$  mit  $\|A(x)\| \leq M \quad \forall x \in K$ .

Für  $x \in K$ ,  $y \in \mathbb{R}^n$  ist

$$\begin{aligned} \|f(x, y) - f(x, \tilde{y})\| &= \|A(x)y - A(x)\tilde{y}\| = \|A(x) \cdot (y - \tilde{y})\| \\ &\stackrel{(2)}{\leq} \|A(x)\| \cdot \|y - \tilde{y}\| \leq M \cdot \|y - \tilde{y}\| \end{aligned}$$

□

**Bemerkung 3.** Wir identifizieren  $\mathbb{C}^n$  mit  $\mathbb{R}^{2n}$  vermöge

$$(z_1, \dots, z_n) \longleftrightarrow (\operatorname{Re}(z_1), \dots, \operatorname{Re}(z_n), \operatorname{Im}(z_1), \dots, \operatorname{Im}(z_n))$$

Eine  $\mathbb{C}$ -lineare Abbildung  $\mathbb{C}^n \rightarrow \mathbb{C}^n$ , also eine komplexe Matrix  $A \in M(n, n; \mathbb{C})$  wird dabei identifiziert mit einer  $\mathbb{R}$ -linearen Abbildung  $\mathbb{R}^{2n} \rightarrow \mathbb{R}^{2n}$ , also mit einer Matrix  $A_{\mathbb{R}} \in M(2n, 2n; \mathbb{R})$

**Beispiel 2.**  $n = 1$ : Sei  $a = b + i \cdot c \in \mathbb{C}$  mit  $b, c \in \mathbb{R}$ .

Ist  $z = x + i \cdot y \in \mathbb{C}$ , so ist  $a \cdot z = (b + i \cdot c) \cdot (x + i \cdot y) = (bx - cy) + i \cdot (by + cx)$ .

Die komplexe  $1 \times 1$ -Matrix  $A = (a)$  wird identifiziert mit

$$A_{\mathbb{R}} = \begin{pmatrix} b & -c \\ c & b \end{pmatrix}$$

Allgemein: Sei  $A \in M(n, n; \mathbb{C})$ ,  $A = B + i \cdot C$  mit  $B, C \in M(n, n; \mathbb{R})$ . Dann ist  $A_{\mathbb{R}} = \begin{pmatrix} A & -C \\ C & B \end{pmatrix}$

**Definition 2.** Sei  $I$  ein offenes Intervall und seien  $A : I \rightarrow M(n, n; \mathbb{C})$  und  $b : I \rightarrow \mathbb{C}^n$  zwei stetige Abbildungen. Wir definieren  $f : I \times \mathbb{C}^n \rightarrow \mathbb{C}^n$  durch  $f(x, y) = A(x)y + b(x)$ , indem wir  $\mathbb{C}^n$  mit  $\mathbb{R}^{2n}$  identifizieren, wird  $y' = f(x, y)$  zu einem reellen System von  $2n$  linearen DGLn. Wir nennen:

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

eine komplexe lineare DGL oder kurz eine lineare DGL.

**Satz 2.** Sei  $\mathbb{K} = \mathbb{R}$  oder  $\mathbb{K} = \mathbb{C}$ . Sei  $I$  ein offenes Intervall und sei  $A : I \rightarrow M(n, n; \mathbb{K})$  eine stetige Abbildung. Sei  $L$  die Menge aller Lösungen der DGL  $y' = A(x)y$ .

a)  $L$  ist ein  $\mathbb{K}$ -Vektorraum der Dimension  $n$ .

b) Man erhält einen Isomorphismus von  $L$  auf  $\mathbb{K}^n$  durch

$$\varphi \mapsto \varphi(x_0) \quad , \quad \text{wobei } x_0 \in I \text{ fest gewählt ist}$$

*Beweis.* Sind  $\varphi_1, \varphi_2$  Lösungen von  $y' = A(x)y$ , so ist auch  $\varphi_1 + \varphi_2$  Lösung von  $y' = A(x)y$ .

Ist  $\varphi$  Lösung von  $y' = A(x)y$  und  $\lambda \in \mathbb{K}$ , so ist  $\lambda \cdot \varphi$  Lösung.

Deswegen ist  $L$  ein  $\mathbb{K}$ -Vektorraum.

Nach Satz 1 ist die Abbildung  $\varphi \mapsto \varphi(x_0)$  ein Isomorphismus von  $L$  auf  $\mathbb{K}^n$ .

Deswegen hat  $L$  die Dimension  $n$ . □

**Beispiel 3.**  $\begin{matrix} y'_1 = & y_2 \\ y'_2 = & -y_1 \end{matrix}$  In Matrixschreibweise  $y' = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \cdot y$

Zwei Lösungen dieses Systems sind  $\varphi_1(x) = \begin{pmatrix} \sin x \\ \cos x \end{pmatrix}$  und  $\varphi_2(x) = \begin{pmatrix} \cos x \\ -\sin x \end{pmatrix}$ .

Die Beiden Lösungen sind linear unabhängig, denn für jedes  $x_0 \in \mathbb{R}$  ist

$$\det \begin{pmatrix} \sin x_0 & \cos x_0 \\ \cos x_0 & -\sin x_0 \end{pmatrix} = -\sin^2 x_0 - \cos^2 x_0 = -1 \neq 0$$

Daher bilden  $\varphi_1$  und  $\varphi_2$  eine Basis des Lösungsraums unseres DGL-Systems.

**Bemerkung 4.** In der Situation von Satz 2 handelt es sich darum, eine Basis  $\varphi_1, \dots, \varphi_n$  des Lösungsraums  $L$  zu finden. Alle Lösungen sind dann von der Form

$$\alpha \cdot \varphi_1 + \dots + \alpha_n \cdot \varphi_n \quad \text{mit } \alpha_1, \dots, \alpha_n \in \mathbb{K}$$

Hat man  $n$  Lösungen  $\varphi_1, \dots, \varphi_n$  gefunden, so bilden diese genau dann eine Basis von  $L$ , wenn für ein (und damit jedes)  $x_0 \in I$  die Vektoren  $\varphi_1(x_0), \dots, \varphi_n(x_0)$  eine Basis von  $\mathbb{K}^n$  bilden, d.h. wenn  $\det(\varphi_1(x_0), \dots, \varphi_n(x_0)) \neq 0$ .

Es gibt kein allgemeines Verfahren, um eine solche Basis zu finden (für  $n \geq 2$ )!

**Satz 3.** Sei  $\mathbb{K} = \mathbb{R}$  oder  $\mathbb{K} = \mathbb{C}$ , sei  $I$  ein offenes Intervall und seien  $A : I \longleftarrow M(n, n; \mathbb{K})$  und  $b : I \longrightarrow \mathbb{K}^n$  stetig.

Sei  $L$  der Lösungsraum der homogenen DGL  $y' = A(x)y$  und sei  $M$  die Menge der Lösungen von  $y' = A(x)y + b(x)$ .

Ist  $\psi_0 \in M$ , so ist  $M = \{\varphi + \psi_0 \mid \varphi \in L\} =: L + \psi_0$

*Beweis.*

- $L + \psi_0 \subseteq M$ : Sei  $\varphi \in L$ . Es ist  $\varphi' = A \cdot \varphi$  und  $\psi_0' = A\psi_0 + b$   
 $\Rightarrow (\varphi + \psi_0)' = \varphi' + \psi_0' = A\varphi + A\psi_0 + b = A \cdot (\varphi + \psi_0) + b \Rightarrow \varphi + \psi_0 \in M$
- $M \subseteq L + \psi_0$ : Sei  $\psi \in M$ . Sei  $\varphi := \psi - \psi_0$ . Dann ist  $\varphi \in L$ , denn  $\psi' = A\psi + b$  und  $\psi_0' = A\psi_0 + b$ , also  $\varphi' = (A\psi + b) - (A\psi_0 + b) = A(\psi - \psi_0) = A\varphi$

□

**Bemerkung 5.** Hat man eine Basis des Lösungsraums der homogenen Gleichung  $y' = A(x)y$  gefunden, so erhält man eine Lösung von  $y' = A(x)y + b(x)$  durch Variation der Konstanten.

Sei  $\varphi_1, \dots, \varphi_n$  Basis von  $L$ , also  $\varphi_i : I \longrightarrow \mathbb{K}^n$ . Schreibe  $\varphi_i(x)$  als Spalte und setze

$$\Phi(x) := (\varphi_1(x), \dots, \varphi_n(x)) \in M(n, n; \mathbb{K})$$

Für jedes  $x \in I$  ist  $\Phi(x)$  eine invertierbare  $n \times n$ -Matrix.

Jede Lösung  $\psi$  von  $y' = A(x)y + b(x)$  ist also von der Form  $\psi(x) = \Phi(x) \cdot u(x)$  mit einer differenzierbaren Funktion  $u : I \longrightarrow \mathbb{K}^n$ .

$$\psi'(x) = \Phi'(x) \cdot u(x) + \Phi(x) \cdot u'(x) = A(x) \cdot \Phi(x) \cdot u(x) = A\psi + \Phi u'$$

$\psi \in M$  bedeutet  $\psi' = A\psi + b$ .

$$\text{Also:} \quad \psi \in M \Leftrightarrow \Phi \cdot u' = b \Leftrightarrow u' = \Phi^{-1} \cdot b \Leftrightarrow u(x) = \int_{x_0}^x \Phi(t)^{-1} \cdot b(t) \, dt + \text{const.}$$

**Beispiel 4.**

$$\begin{aligned} u(x) &= \int_0^x \Phi(t)^{-1} \cdot b(t) dt = \int_0^x \begin{pmatrix} \sin t & \cos t \\ \cos t & -\sin t \end{pmatrix} \cdot \begin{pmatrix} 0 \\ t \end{pmatrix} dt \\ &= \int_0^x \begin{pmatrix} t \cdot \cos t \\ -t \cdot \sin t \end{pmatrix} dt = \begin{pmatrix} \cos x + \sin x - 1 \\ -\sin x + x \cdot \cos x \end{pmatrix} \end{aligned}$$

$$\psi(t) = \Phi(x) \cdot u(x) = \begin{pmatrix} \sin x & \cos x \\ \cos x & -\sin x \end{pmatrix} \cdot \begin{pmatrix} \cos x + x \cdot \sin x - 1 \\ -\sin x + x \cdot \cos x \end{pmatrix} = \begin{pmatrix} x - \sin x \\ 1 - \cos x \end{pmatrix}$$

Damit haben wir eine Lösung gefunden und damit alle Lösungen.

**Definition 3.** Sei  $I$  ein offenes Intervall und seien  $a_0, a_1, \dots, a_{n-1}, b : I \longrightarrow \mathbb{K}$  stetige Funktionen. Dann heißt

$$y^{(n)} = a_0(x)y + a_1(x)y' + \dots + a_{n-1}(x)y^{(n-1)} + b(x)$$

eine lineare DGL  $n$ -ter Ordnung. Ist  $b = 0$ , so heißt die DGL homogen.



**Satz 4.**

a) Sei  $L$  die Menge der Lösungen der homogenen linearen DGL

$$y^{(n)} = a_0(x)y + a_1(x)y' + \dots + a_{n-1}(x)y^{(n-1)}$$

Dann ist  $L$  ein Vektorraum der Dimension  $n$

b) Sind  $\varphi_1, \dots, \varphi_n \in L$ , so bilden sie genau dann eine Basis von  $L$ , wenn für ein und damit für jedes  $x \in I$  gilt:

$$\det \begin{pmatrix} \varphi_1(x) & \dots & \varphi_n(x) \\ \varphi_1'(x) & \dots & \varphi_n'(x) \\ \vdots & & \vdots \\ \varphi_1^{(n-1)} & \dots & \varphi_n^{(n-1)}(x) \end{pmatrix} \neq 0$$

c) Ist  $M$  die Menge der Lösungen von

$$y^{(n)} = a_0(x)y + a_1(x)y' + \dots + a_{n-1}(x)y^{(n-1)} + b(x)$$

und ist  $\psi_0 \in M$ , so ist  $M = \psi_0 + L$

## §10 Lineare Differenzialgleichungen mit konstanten Koeffizienten

**Definition 1.** Sei  $\mathbb{K} = \mathbb{R}$  oder  $\mathbb{K} = \mathbb{C}$  und  $n \in \mathbb{N}$ . Ist  $a \in M(n, n; \mathbb{K})$ , so heißt

$$y' = A \cdot y$$

eine homogene lineare DGL mit konstanten Koeffizienten.  
Ihre Lösungen sind auf ganz  $\mathbb{R}$  definiert.

**Bemerkung 1.** Ist  $n = 1$ , also  $A = (a)$ , so sind alle Lösungen von  $y' = A \cdot y$  von der Form

$$\varphi(x) = e^{xa} \cdot y_0 \quad \text{mit } y_0 \in \mathbb{R}$$

Auch für  $n > 1$  möchte man die Lösungen von  $y' = A \cdot y$  in der Form

$$\varphi(x) = e^{xA} \cdot y_0 \quad \text{mit } y_0 \in \mathbb{R}^n, e^{xA} \in M(n, n; \mathbb{R}) \text{ schreiben}$$

Dafür müssen wir definieren, was  $e^A$  ist, falls  $A$  eine  $n \times n$ -Matrix ist.

**Definition 2.** Sei  $X$  ein metrischer Raum

- a) Eine Folge  $(x_n)$  in  $X$  heißt eine Cauchy-Folge, wenn es für jedes  $\varepsilon > 0$  ein  $N \in \mathbb{N}$  gibt, sodass  $d(x_n, x_m) < \varepsilon$  für alle  $n, m \geq N$
- b) Der metrische Raum  $X$  heißt vollständig, wenn jede Cauchy-Folge in  $X$  konvergent ist.
- c) Ein normierter Raum  $V$  heißt Banach-Raum, wenn  $V$  (aufgefasst als metrischer Raum) vollständig ist.

**Bemerkung 2.** Jeder endlich-dimensionaler normierter Raum ist ein Banach-Raum.

**Definition 3.** Sei  $V$  ein normierter Raum und  $(a_n)$  eine Folge in  $V$ .

- a) Wenn die Folge  $\left(\sum_{n=1}^k a_n\right)_{k \geq 1}$  konvergiert, so sagt man, dass die Reihe  $\sum_{n=1}^{\infty} a_n$  konvergiert und schreibt  $\sum_{n=1}^{\infty} a_n$  für den Grenzwert.
- b) Wenn die Reihe  $\sum_{n=1}^{\infty} \|a_n\|$  reeller Zahlen konvergiert, so sagt man, dass die Reihe  $\sum_{n=1}^{\infty} a_n$  absolut konvergiert.

**Bemerkung 3.** In einem Banach-Raum ist jede absolut konvergente Reihe konvergent. . . .

**Definition 4.** Sei  $A \in M(n, n; \mathbb{K})$ . Dann ist die Reihe  $\sum_{k=0}^{\infty} \frac{1}{k!} \cdot A^k$  in  $M(n, n; \mathbb{K})$  absolut konvergent, also konvergent. Schreibe  $e^A := \exp(A) := \sum_{k=0}^{\infty} \frac{1}{k!} \cdot A^k$ .

(Begründung dafür, dass  $\sum_{k=0}^{\infty} \frac{1}{k!} \cdot A^k$  absolut konvergent ist:

Wir wissen, dass  $\|A \cdot B\| \leq \|A\| \cdot \|B\|$ . Deswegen ist  $\|A^2\| \leq \|A\|^2$ , also  $\|A^k\| \leq \|A\|^k \quad \forall k \in \mathbb{N}_0$ .  
Die Reihe  $\sum_k \frac{1}{k!} \cdot \|A\|^k$  ist also konvergente Majorante für  $\sum_k \left\| \frac{1}{k!} \cdot A^k \right\|$ .)

$e^A = I_n + A + \frac{1}{2!} \cdot A^2 + \frac{1}{3!} \cdot A^3 + \dots$  Beachte  $A^0 = I \quad \forall A \in M(n, n; \mathbb{K})$ .  
Schreibe  $M_n := M(n, n; \mathbb{K})$

**Beispiel 1.**

a)  $\exp(0_n) = I_n$

b)  $\exp(I) = I + I + \frac{1}{2!} \cdot I + \frac{1}{3!} \cdot I^3 + \dots = e \cdot I$

c) Allgemeiner: Ist  $A$  eine Diagonalmatrix  $A = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} = \text{diag}(\lambda_1, \dots, \lambda_n)$ , so ist  $A^k = \text{diag}(\lambda_1^k, \dots, \lambda_n^k) \Rightarrow \exp(A) = \text{diag}(e^{\lambda_1}, \dots, e^{\lambda_n})$

d) Noch Allgemeiner: Ist  $A = \begin{pmatrix} A_1 & & 0 \\ & \ddots & \\ 0 & & A_m \end{pmatrix}$  mit quadratischen Matrizen  $A_j$ ,  
so ist  $\exp(A) = \begin{pmatrix} \exp(A_1) & & \\ & \ddots & \\ & & \exp(A_m) \end{pmatrix}$

e) Sind  $A, S \in M_n$ , wobei  $S$  invertierbar ist, so gilt:

$$(S \cdot A \cdot S^{-1})^k = (SAS^{-1}) \cdot (SAS^{-1}) \dots (SAS^{-1}) = S \cdot A^k \cdot S^{-1}$$

$$\Rightarrow \exp(SAS^{-1}) = \sum_{k=0}^{\infty} \frac{1}{k!} \cdot SA^k S^{-1} = S \cdot \left( \sum_{k=0}^{\infty} \frac{1}{k!} \cdot A^k \right) \cdot S^{-1} = S \cdot \exp(A) \cdot S^{-1}$$

Eine Matrix  $A$  heißt diagonalisierbar, wenn es eine invertierbare Matrix  $S$  und eine Diagonalmatrix  $\Delta = (\lambda_1, \dots, \lambda_n)$  gibt, mit  $A = S\Delta S^{-1}$ . Dies ist genau der Fall, wenn es eine Basis in  $\mathbb{K}^n$  gibt, die aus Eigenvektoren von  $A$  besteht, genauer: Wenn es eine Basis  $v_1, \dots, v_n$  von  $\mathbb{K}^n$  gibt mit

$$Av_i = \lambda_i \cdot v_i \quad \text{mit } i = 1, \dots, n$$

$$\text{Dann ist } \exp(A) \stackrel{e)}{=} S \cdot e^{\Delta} \cdot S^{-1} \stackrel{c)}{=} S \cdot \text{diag}(e^{\lambda_1}, \dots, e^{\lambda_n}) \cdot S^{-1}$$

f) Sei  $A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$  und  $B = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$

$$e^A \stackrel{c)}{=} \begin{pmatrix} e & 0 \\ 0 & 1 \end{pmatrix} \quad B^2 = 0 = B^3 = B^4 = \dots \Rightarrow e^B = I + B = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

$$e^A \cdot e^B = \begin{pmatrix} e & e \\ 0 & 1 \end{pmatrix}$$

$$\begin{aligned} A + B &= \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} = (A + B)^2 = (A + B)^3 = \dots \Rightarrow e^{A+B} = I + \sum_{k=1}^{\infty} \frac{1}{k!} \cdot (A + B)^k \\ &= I + \sum_{k=1}^{\infty} \frac{1}{k!} \cdot \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} = I + \left( \sum_{k=0}^{\infty} \frac{1}{k!} \right) \cdot \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} - \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} \\ &= I + e \cdot \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} - \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} e & e-1 \\ 0 & 1 \end{pmatrix} \end{aligned}$$

Es ist  $e^{A+B} \neq e^A \cdot e^B$  !

$(A + B)^2 = A^2 + AB + BA + B^2$  Bei Matrizen ist i.A.:  $AB \neq BA$ .

g) Sind  $A, B \in M_n$  mit  $AB = BA$ , so ist  $e^{A+B} = e^A \cdot e^B$

h) Spezialfall von g): Ist  $A \in M_n$  und sind  $s, t \in \mathbb{K}$ , so ist  $(sA)(tA) = stA^2 = (tA)(sA)$

$$\text{also } e^{(s+t) \cdot A} = e^{sA} \cdot e^{tA}$$

i) Insbesondere  $e^A \cdot e^{-A} = e^{(1-1) \cdot A} = e^{0A} = e^0 = I$

Für jedes  $A$  ist  $e^A$  invertierbar und  $(e^A)^{-1} = e^{-A}$

$$j) \text{ Sei } N = \begin{pmatrix} 0 & 1 & 0 \\ \vdots & \ddots & 1 \\ 0 & \dots & 0 \end{pmatrix}$$

$$\Rightarrow N^2 = \begin{pmatrix} 0 & 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 0 & 1 & & 0 \\ \vdots & \vdots & \vdots & \ddots & \ddots & 1 \\ 0 & 0 & 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & 0 & \dots & 0 \end{pmatrix}$$

$$N^3 = \begin{pmatrix} 0 & 0 & 0 & 1 & 0 \\ \vdots & \ddots & \ddots & \ddots & 1 \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \dots & \dots & \dots & 0 \end{pmatrix}, \quad N^{n-1} = \begin{pmatrix} 0 & \dots & 0 & 1 \\ 0 & \dots & 0 & 0 \\ \vdots & & \vdots & \vdots \\ 0 & \dots & 0 & 0 \end{pmatrix}$$

$$N^n = 0, \quad N^{n+1} = 0, \dots$$

$$\Rightarrow e^N = \sum_{k=0}^{n-1} \frac{1}{k!} \cdot N^k, \text{ allgemeiner } e^{xN} = \sum_{k=0}^{n-1} \frac{x^k}{k!} \cdot N^k =: P_n(x).$$

$$P_n(x) := \begin{pmatrix} 1 & x & \frac{x^2}{2!} & \dots & \frac{x^{n-1}}{(n-1)!} \\ 0 & \ddots & \ddots & & \vdots \\ \vdots & & \ddots & \ddots & \frac{x^2}{2!} \\ \vdots & & & \ddots & x \\ 0 & \dots & \dots & \dots & 1 \end{pmatrix}$$

$$k) \quad B := \begin{pmatrix} \lambda & 1 & 0 \\ & \ddots & 1 \\ 0 & & 1 \end{pmatrix} \Rightarrow xB = x\lambda I + xN \xrightarrow{g)} e^{xB} = e^{x\lambda I} \cdot e^{xN} = e^{x\lambda} \cdot e^{xN} \xrightarrow{j)} e^{x\lambda} \cdot P_n(x)$$

**Allgemeine Vorgehensweise zur Berechnung von  $e^A$  für  $A \in M_n(\mathbb{K})$**

- Man kann  $\mathbb{K} = \mathbb{C}$  annehmen.
- Man bestimmt die Eigenwerte und die Eigenvektoren von  $A$ .
- Man transformiert  $A$  auf die Jordansche Normalform, d.h. man sucht eine invertierbare Matrix  $S$  mit  $A = S^{-1}JS$ , wobei

$$(*) \quad J = \begin{pmatrix} J_1 & & 0 \\ & \ddots & \\ 0 & & J_m \end{pmatrix} \text{ mit } J_k \in M_{n_k} \quad J_k = \begin{pmatrix} \lambda_k & 1 & 0 \\ & \ddots & 1 \\ 0 & & \lambda_k \end{pmatrix}$$

Die  $\lambda_k$  sind die Eigenwerte von  $A$ .

$J$  ist dabei bis auf die Reihenfolge der  $J_k$  eindeutig bestimmt.

Man nennt die  $J_k$  die „Jordan-Blöcke“ von  $A$ .

- Mit Beispiel k) berechnet man  $e^{J_k}$
- Mit Beispiel d) erhält man  $e^J$
- Mit Beispiel e) erhält man  $e^A$

**Bemerkung 4.** Aus (\*) sieht man  $\boxed{\det(e^A) = e^{\text{Spur } A}}$ .

Ist  $A = (a_{ij})$ , so ist  $\text{Spur } A = \sum_{i=1}^n a_{ii} = \text{Summe aller Eigenwerte von } A, \text{ gezählt mit Vielfachheit.}$

**Lemma 1.** Ist  $A \in M_n(\mathbb{K})$  und definiert man  $\Phi : \mathbb{R} \rightarrow M_n(\mathbb{K})$  durch  $\Phi(x) := e^{xA}$ , so ist  $\Phi$  differenzierbar mit  $\Phi'(x) = A \cdot e^{xA} = A\Phi(x) \quad \forall x \in \mathbb{R}$ .

*Beweis.*

$$\begin{aligned}\Phi(x+h) - \Phi(x) &= e^{(x+h)A} - e^{xA} \stackrel{h}{=} e^{hA} \cdot e^{xA} - e^{xA} = (e^{hA} - I)e^{xA} \\ &= \left( \sum_{k=1}^{\infty} \frac{1}{k!} h^k A^k \right) \cdot e^{xA} = h \cdot \left( \sum_{k=1}^{\infty} \frac{1}{k!} h^{k-1} A^k \right) \cdot e^{xA}\end{aligned}$$

$$\Rightarrow \text{Für } h \neq 0 \text{ ist } \frac{1}{h}(\Phi(x+h) - \Phi(x)) = \left( \sum_{k=1}^{\infty} \frac{1}{k!} h^{k-1} A^k \right) \cdot e^{xA} \xrightarrow{h \rightarrow 0} A \cdot e^{xA}$$

Aus dem Lemma und §9 folgt: □

**Satz 1.** Sei  $A \in M_n(\mathbb{K})$  und  $y_0 \in \mathbb{K}^n$ . Die einzige Lösung  $\varphi$  der DGL

$$y' = Ay$$

mit  $\varphi(0) = y_0$  ist gegeben durch  $\varphi(x) = e^{xA} \cdot y_0$ .

**Bemerkung 5.** Ist  $v_0 \in \mathbb{K}^n \setminus \{0\}$  ein Eigenvektor von  $A$  zum Eigenwert  $\lambda$ , also  $Av_0 = \lambda v_0$ , so ist  $A^2 v_0 = \lambda^2 v_0$ , allgemein  $A^k v_0 = \lambda^k v_0 \quad \forall k \in \mathbb{N}_0 \Rightarrow \exp(A) \cdot v_0 = e^\lambda \cdot v_0$ .

Allgemeiner:  $\exp(xA) \cdot v_0 = e^{\lambda x} v_0$ . Also ist  $x \mapsto e^{\lambda x} v_0$  eine Lösung von  $y' = Ay$ .

Ist  $A$  diagonalisierbar, so gibt es eine Basis  $v_1, \dots, v_n$  von  $\mathbb{K}^n$  und  $\lambda_1, \dots, \lambda_n \in \mathbb{K}$  mit  $Av_k = \lambda_k v_k$ .

Dann bilden die Funktionen  $x \mapsto e^{\lambda_k x} v_k, k = 1, \dots, n$  eine Basis des Lösungsraumes von  $y' = Ay$ .

**Beispiel 2.**  $\begin{smallmatrix} y'_1 = 5y_1 + 3y_2 \\ y'_2 = -6y_1 - 4y_2 \end{smallmatrix}$ , also  $y' = Ay$  mit  $A = \begin{pmatrix} 5 & 3 \\ -6 & -4 \end{pmatrix}$

$$\begin{pmatrix} 5 & 3 \\ -6 & -4 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ -1 \end{pmatrix} = \begin{pmatrix} 2 \\ -2 \end{pmatrix} = 2 \cdot \begin{pmatrix} 1 \\ -1 \end{pmatrix} \Rightarrow \begin{pmatrix} 1 \\ -1 \end{pmatrix} \text{ ist Eigenvektor zum Eigenwert } 2$$

$$\begin{pmatrix} 5 & 3 \\ -6 & -4 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ -2 \end{pmatrix} = \begin{pmatrix} -1 \\ 2 \end{pmatrix} = -1 \cdot \begin{pmatrix} 1 \\ -2 \end{pmatrix} \Rightarrow \begin{pmatrix} 1 \\ -2 \end{pmatrix} \text{ ist Eigenvektor zum Eigenwert } -1$$

Jede Lösung ist Linearkombination von  $x \mapsto e^{2x} \cdot \begin{pmatrix} 1 \\ -1 \end{pmatrix}$  und  $x \mapsto e^{-x} \cdot \begin{pmatrix} 1 \\ -2 \end{pmatrix}$

Jede Lösung ist also von der Form  $y = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$  mit  $y_2 = \frac{y_1 - \alpha \cdot e^{2x} + \beta \cdot e^{-x}}{-\alpha \cdot e^{2x} - 2\beta \cdot e^{-x}}$  mit  $\alpha, \beta \in \mathbb{K}$

**Beispiel 3.**  $\begin{smallmatrix} y'_1 = y_1 + y_2 \\ y'_2 = y_2 \end{smallmatrix}$ , also  $y' = Ay$  mit  $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$

$$e^{xA} = \exp(xI + \begin{pmatrix} 0 & x \\ 0 & 0 \end{pmatrix}) = \exp(xI) \cdot \exp \begin{pmatrix} 0 & x \\ 0 & 0 \end{pmatrix} = e^x (I + \begin{pmatrix} 0 & x \\ 0 & 0 \end{pmatrix}) = e^x \cdot \begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix}$$

Jede Lösung ist also von der Form  $x \mapsto e^x \begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = e^x \cdot \begin{pmatrix} \alpha + \beta x \\ \beta \end{pmatrix}$ , also  $\begin{smallmatrix} y_1 = \alpha \cdot e^x + \beta x \cdot e^x \\ y_2 = \beta \cdot e^x \end{smallmatrix}$  mit  $\alpha, \beta \in \mathbb{K}$

**Satz 2.** Sei  $A \in M_n(\mathbb{C})$  und  $\varphi(\varphi_1, \dots, \varphi_n) : \mathbb{R} \rightarrow \mathbb{C}^n$  eine Lösung der DGL  $y' = Ay$ . Dann ist jedes  $\varphi_j$  eine komplexe Linearkombination der Funktionen

$$x \mapsto x^k \cdot e^{\lambda x},$$

wobei  $\lambda$  ein Eigenwert von  $A$  und  $k \in \mathbb{N}_0$  kleiner als die algebraische Vielfachheit des Eigenvektors  $\lambda$  von  $A$  (sogar kleiner als die Größe des größten Jordan-Blocks zum Eigenwert  $\lambda$ ) ist.

*Beweis.* Nach Satz 1 gibt es ein  $y_0 \in \mathbb{C}^n$  mit  $\varphi(x) = e^{xA} \cdot y_0$ .

Es gibt eine invertierbare Matrix  $S$  mit  $S^{-1}AS = \begin{pmatrix} J_1 & & 0 \\ & \ddots & \\ 0 & & J_n \end{pmatrix}$ ,  $J_k = \begin{pmatrix} \lambda_1 & 1 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ 0 & & & \lambda_i \end{pmatrix} \in$

$M_{k_i}(\mathbb{C})$

$$\Rightarrow \varphi(x) = e^{xA}y_0 = S(S^{-1}e^{xA}S)S^{-1}y_0 = Se^{xS^{-1}AS}S^{-1}y_0$$

Es ist  $e^{xS^{-1}AS} = \begin{pmatrix} B_1(x) & & 0 \\ & \ddots & \\ 0 & & B_m(x) \end{pmatrix}$  mit  $B_i(x) = e^{xJ_i} \stackrel{j)}{=} e^{\lambda_i x} \cdot P_{k_i}(x)$ , wobei die Einträge in  $P_{k_i}(x)$

Monome in  $x$  vom Grad  $\leq k_i - 1$  sind.  $\square$

**Bemerkung 6.** Ist  $A \in M_n(\mathbb{R})$  und ist  $\varphi$  eine komplexe Lösung von  $y' = Ay$ , so sind  $\operatorname{Re}(\varphi)$  und  $\operatorname{Im}(\varphi)$  reelle Lösungen von  $y' = Ay$ .

Ist  $\lambda = a + bi$  mit  $a, b \in \mathbb{R}$ , so ist  $x^k e^{\lambda x} = x^k e^{ax+bi x} = x^k e^{ax} (\cos bx + i \cdot \sin bx)$ .

Ist  $\lambda = a + bi$  Eigenwert von  $A$ , so auch  $\bar{\lambda} = a - bi$ , und zwar mit derselben Vielfachheit. Damit haben wir gesehen:

**Satz 3.** Sei  $A \in M_n(\mathbb{R})$  und  $\varphi = (\varphi_1, \dots, \varphi_n) : \mathbb{R} \rightarrow \mathbb{R}^n$  eine Lösung der DGL  $y' = Ay$ . Dann ist jedes  $\varphi_j$  eine reelle Linearkombination der Funktionen

$$x \mapsto x^k e^{ax} \cos bx \quad \text{und} \quad x \mapsto x^k e^{ax} \cdot \sin bx,$$

wobei  $a + bi$  die (komplexen) Eigenwerte von  $A$  mit  $b \geq 0$  durchläuft und  $k \in \mathbb{N}_0$  mit  $0 \leq k < \text{Vielfachheit des Eigenwerts } a + bi \text{ von } A$ .

## Lineare Differenzialgleichungen höherer Ordnung mit konstanten Koeffizienten

Sei  $n \in \mathbb{N}$  und seien  $a_0, a_1, \dots, a_{n-1} \in \mathbb{K}$ . Betrachte die DGL

$$(*) \quad y^{(n)} + a_{n-1}y^{(n-1)} + \dots + a_1y' + a_0y = 0$$

Sie ist äquivalent zu dem System von  $n$  DGLn 1. Ordnung

$$(**) \quad Y' = AY$$

mit

$$A = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \\ -a_0 & -a_1 & -a_2 & \dots & -a_{n-1} \end{pmatrix}$$

Man erhält einen Isomorphismus vom Lösungsraum von  $(**)$  auf den Lösungsraum von  $(*)$  durch  $\varphi(\varphi_0, \varphi_1, \dots, \varphi_{n-1}) \mapsto \varphi_0$ .

Das charakteristische Polynom von  $A$  ist:

$$g(x) = \begin{vmatrix} x & -1 & 0 & \dots & 0 & 0 \\ 0 & x & -1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & x & -1 \\ a_0 & a_1 & a_2 & \dots & a_{n-1} & x + a_{n-1} \end{vmatrix} = x^n + a_{n-1}x^{n-1} + \dots + a_1x + a_0,$$

wie man durch Entwickeln nach der letzten sieht.

**Beispiel 4.**  $n=2$ :  $g(x) = \begin{vmatrix} x & -1 \\ a_0 & x + a_1 \end{vmatrix} = x^2 + a_1x + a_0$

Sei  $g(x) = \prod_{j=1}^n (x - \lambda_j)^{k_j}$  mit paarweise verschiedenen  $\lambda_j \in \mathbb{C}$ ,  $k_j \geq 1$ .

Ist  $\varphi = (\varphi_0, \dots, \varphi_{n-1})$  Lösung von  $(**)$ , so ist  $\varphi_0$  nach Satz 2 eine komplexe Linearkombination der Funktionen  $x \mapsto x^k e^{\lambda_j x}$ ,  $j = 1, \dots, m$ ,  $0 \leq k < k_j$ .

Das sind genau  $k_1 + \dots + k_m = n$  Funktionen.

Andererseits hat der Lösungsraum von  $(*)$  die Dimension  $n$ .

Deswegen bilden diese Funktionen  $x \mapsto x^k e^{\lambda_j x}$  eine Basis des Lösungsraumes von  $(*)$

**Satz 4.** Seien  $a_0, a_1, \dots, a_{n-1} \in \mathbb{C}$ . Es sei

$$g(x) := x^n + a_{n-1}x^{n-1} + \dots + a_1x + a_0 = \prod_{j=1}^m (x - \lambda_j)^{k_j}$$

mit paarweisen verschiedenen  $\lambda_j \in \mathbb{C}$ ,  $k_j \geq 1$ . Dann bilden die Funktionen

$$\begin{aligned} & e^{\lambda_1 x}, x e^{\lambda_1 x}, \dots, x^{k_1-1} e^{\lambda_1 x} \\ & \vdots \\ & e^{\lambda_m x}, x e^{\lambda_m x}, \dots, x^{k_m-1} e^{\lambda_m x} \end{aligned}$$

eine Basis des Lösungsraumes der DGL

$$y^{(n)} + a_{n-1}y^{(n-1)} + \dots + a_1y' + a_0y = 0$$

**Bemerkung 7.** Sind  $a_0, \dots, a_{n-1} \in \mathbb{R}$  und ist  $\varphi$  eine komplexe Lösung von  $(*)$ , so sind  $\operatorname{Re}(\varphi)$  und  $\operatorname{Im}(\varphi)$  reelle Lösungen von  $(*)$ . Damit hat man:

**Satz 5.** Seien  $a_0, a_1, \dots, a_{n-1} \in \mathbb{R}$  und  $g(x) := x^n + a_{n-1}x^{n-1} + \dots + a_1x + a_0$ .

Seien  $\lambda_1, \dots, \lambda_r$  die paarweise verschiedenen reellen Nullstellen von  $g$

und seien  $\lambda_{r+1}, \dots, \lambda_s$  die paarweise verschiedenen Nullstellen von  $g$  mit  $\operatorname{Im}(\lambda_j) > 0$ .

Für  $j = 1, \dots, s$  sei  $k_j$  die Vielfachheit der Nullstelle  $\lambda_j$  von  $g$ .

Für  $j = r+1, \dots, s$  sei  $\lambda_j = \mu_j + i \cdot v_j$  mit  $\mu_j, v_j \in \mathbb{R}$ . Dann bilden die Funktionen

$$\begin{aligned} & x^r e^{\lambda_j x} \quad \text{mit } 1 \leq j \leq r, \quad 0 \leq p < k_j, \\ & x^p e^{\mu_0 x} \cos v_j x \\ & x^p e^{\mu_j x} \sin v_j x \quad \text{mit } r+1 \leq j \leq s, \quad 0 \leq p < k_j \end{aligned}$$

eine Basis des reellen Vektorraums der Lösungen von

$$y^{(n)} + a_{n-1}y^{(n-1)} + \dots + a_1y' + a_0y = 0$$

**Beispiel 5.** Die DGL der gedämpften Schwingung  $y'' + 2\mu y' + \omega_0^2 = 0$  mit  $\mu \geq 0$ ,  $\omega_0 > 0$ . ( $2\mu$  heißt der Dämpfungsfaktor,  $\omega_0$  heißt die Frequenz der ungedämpften Schwingung).

$$g(x) = x^2 + 2\mu x + \omega_0^2$$

1. Fall  $\mu < \omega_0$ :  $g$  hat die beiden Nullstellen  $-\mu \pm i\omega$  mit  $\omega := \sqrt{\omega_0^2 - \mu^2} > 0$ .

Eine Basis des Lösungsraums ist gegeben durch  $\varphi_1(x) = e^{-\mu x} \cos \omega x$ ,  $\varphi_2(x) = e^{-\mu x} \sin \omega x$

2. Fall  $\mu = \omega_0$ :  $g$  hat die Nullstelle  $-\mu$  mit Vielfachheit 2.

Eine Basis des Lösungsraums ist gegeben durch  $\varphi_1(x) = e^{-\mu x}$ ,  $x e^{-\mu x}$ .

3. Fall  $\mu > \omega_0$ : Eine Basis des Lösungsraums ist gegeben durch  $\varphi_1(x) = e^{-\mu_1 x}$ ,  $\varphi_2(x) = e^{-\mu_2 x}$ , wobei  $\lambda_{1,2} = -\mu \pm \sqrt{\mu^2 - \omega_0^2}$  die Nullstellen von  $g$  sind und  $\mu_j = -\lambda_j > 0$

## §11 Der Fixpunktsatz von Banach

**Definition 1.** Sei  $X$  eine Menge und  $f : X \rightarrow X$  eine Abbildung. Ein Element  $x \in X$  heißt Fixpunkt von  $f$ , wenn  $f(x) = x$  ist.

**Definition 2.** Sei  $X$  ein metrischer Raum und  $f : X \rightarrow X$  eine Abbildung.  $f$  heißt kontrahierend, wenn es ein  $C \in \mathbb{R}$  gibt mit  $0 \leq C < 1$ , sodass

$$d(f(x), f(y)) \leq C \cdot d(x, y) \quad \forall x, y \in X$$

**Bemerkung 1.** Eine kontrahierende Abbildung ist (Lipschitz-)stetig.

**Erinnerung:** Sei  $X$  ein metrischer Raum:

- Eine Folge  $(x_n)$  in  $X$  heißt Cauchy-Folge, wenn es für jedes  $\varepsilon > 0$  ein  $N \in \mathbb{N}$  gibt, sodass  $d(x_n, x_m) \leq \varepsilon \quad \forall n, m \geq N$ .
- $X$  heißt vollständig, wenn jede Cauchy-Folge in  $X$  konvergent ist.

**Beispiel 1.** Eine abgeschlossene Teilmenge eines vollständigen metrischen Raumes ist vollständig.

**Satz 1. Fixpunktsatz von Banach.** Sei  $X$  ein vollständiger metrischer Raum und  $f : X \rightarrow X$  eine kontrahierende Abbildung. Dann besitzt  $f$  genau einen Fixpunkt.

**Beispiel 2.** Sei  $X := [1, \infty[$  ( $\subset \mathbb{R}$ ). Dann ist  $X$  vollständig. Definiere  $f : X \rightarrow X$  durch  $f(x) := x + \frac{1}{x}$ .

Dann besitzt  $f$  keinen Fixpunkt. Sind  $x, y \in X$ , so gibt es nach dem Mittelwertsatz ein  $\xi$ , das zwischen  $x$  und  $y$  liegt, sodass

$$\begin{aligned} f(y) - f(x) &= f'(\xi) \cdot (y - x), \\ \Rightarrow |f(y) - f(x)| &= f'(\xi) \cdot |y - x| \\ 0 < f'(\xi) &= 1 - \frac{1}{\xi^2} \quad \text{für } \xi \geq 1 \end{aligned}$$

Also  $|f(y) - f(x)| < |y - x|$  für alle  $x, y \in X$  mit  $x \neq y$ .

Das zeigt, dass man bei der Definition von „kontrahierend“ ein  $C < 1$  braucht.

*Beweis.* von Satz 1.

a) Eindeutigkeit: Seien  $x, y$  zwei Fixpunkte von  $f$ .

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

$$d(x, y) = d(f(x), f(y)) \leq C \cdot d(x, y). \text{ Aus } C < 1 \text{ folgt: } d(x, y) = 0, \text{ also } x = y$$

b) Existenz: Wähle ein beliebiges  $x_0 \in X$ . Sei  $f^n := f \circ \dots \circ f$  ( $n$  Faktoren) für  $n \in \mathbb{N}_0$ .

Wir werden zeigen:  $(f(x_0))_{n \in \mathbb{N}_0}$  ist eine Cauchy-Folge.

Wenn das gezeigt ist, so existiert  $\lim_{n \rightarrow \infty} f^n(x_0) =: x$ , weil  $X$  vollständig ist.

Es ist  $f(x) = f(\lim_{n \rightarrow \infty} f^n(x_0)) \stackrel{f \text{ stetig}}{=} f(f^n(x_0)) = \lim_{n \rightarrow \infty} f^{n+1}(x_0) = \lim_{n \rightarrow \infty} f^n(x_0) = x$ , d.h.  $x$  ist ein Fixpunkt von  $f$ .

Nachweis der Behauptung, dass  $(f^n(x_0))_n$  Cauchy-Folge ist:

Sei  $N \in \mathbb{N}$  und  $N \leq n < m$ .

$$\begin{aligned} d(f^m(x_0), f^n(x_0)) &\leq C \cdot d(f^{m-1}(x_0), f^{n-1}(x_0)) \leq C^2 \cdot d(f^{m-2}(x_0), f^{n-2}(x_0)) \\ &\leq C^n \cdot d(f^{m-n}(x_0), x_0) \leq C^N \cdot d(x_0, f^{m-n}(x_0)) \\ &\leq C^N \cdot (d(x_0, f(x_0)) + d(f(x_0), f^2(x_0)) + \dots + d(f^{m-n-1}(x_0), f^{m-n}(x_0))) \\ &\leq C^N \cdot (d(x_0, f(x_0)) + C \cdot d(x_0, f(x_0)) + \dots + C^{m-n-1} \cdot d(x_0, f(x_0))) \\ &= C^N \cdot d(x_0, f(x_0)) \cdot (1 + C + C^2 + \dots + C^{m-n-1}) \\ &\leq C^N \cdot d(x_0, f(x_0)) \cdot (1 + C + C^2 + C^3 + \dots) \\ &= \frac{C^N}{1 - C} \cdot d(x_0, f(x_0)) \end{aligned}$$



Ist  $\varepsilon > 0$  gegeben, so gibt es ein  $N \in \mathbb{N}$  mit  $\frac{C^N}{1-C} \cdot d(x_0, f(x_0)) < \varepsilon$ . Fertig.

□

Manchmal braucht man eine Variante von Satz 1.

**Satz 2.** Sei  $X$  ein vollständiger metrischer Raum,  $x_0 \in X$  und  $R > 0$ .

Sei  $B := B_R(x_0) = \{x \in X \mid d(x, x_0) < R\}$

Sei  $G : B \rightarrow X$  eine Abbildung. Es gebe ein  $C$  mit  $0 \leq C < 1$ , sodass gilt:

- $d(G(x), G(y)) < C \cdot d(x, y) \quad \forall x, y \in B$
- $d(x_0, G(x_0)) < R \cdot (1 - C)$

Dann gibt es genau ein  $x \in B$  mit  $G(x) = x$ .

Der Beweis geht genau wie der von Satz 1, man startet mit dem Mittelpunkt  $x_0$  von  $B$  und überlegt, dass  $G^n(x_0) \in B \quad \forall n \in \mathbb{N}$

## §12 Der lokale Eindeutigkeits- und Existenzsatz

Sei  $U$  offen in  $\mathbb{R} \times \mathbb{R}^n$  und  $f : U \rightarrow \mathbb{R}^n$  stetig. Betrachte die DGL

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

Sei  $(x_0, y_0) \in U$ . Eine Lösung von  $(*)$  mit der Anfangsbedingung  $y(x_0) = y_0$  ist eine differenzierbare Funktion  $\varphi : I \rightarrow \mathbb{R}^n$  mit:

1.  $I$  ist ein offenes Intervall mit  $x_0 \in I$
2. Für alle  $x \in I$  ist  $(x, \varphi(x)) \in U$
3. Für alle  $x \in I$  ist  $\varphi'(x) = f(x, \varphi(x))$
4.  $\varphi(x_0) = y_0$

Formuliere  $(*)$  um in ein Fixpunktproblem.

**Lemma 1.** Sei  $I$  ein offenes Intervall, sei  $H$  offen in  $\mathbb{R}^n$ , sei  $f : I \times H \rightarrow \mathbb{R}^n$  stetig und  $(x_0, y_0) \in I \times H$ .

Sei  $J$  ein offenes Teilintervall von  $I$  mit  $x_0 \in J$ .

Für eine stetige Funktion  $\varphi : J \rightarrow \mathbb{R}^n$  mit  $\varphi(J) \subseteq H$  definieren wir eine stetige Funktion  $G(\varphi) : J \rightarrow \mathbb{R}^n$  durch

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

Für ein solches  $\varphi$  sind äquivalent:

1.  $\varphi$  ist Lösung von  $y' = f(x, y)$  mit  $\varphi(x_0) = y_0$
2.  $G(\varphi) = \varphi$

*Beweis.* 1)  $\Rightarrow$  2): Sei  $\varphi : J \rightarrow H$  differenzierbar mit  $\varphi'(x) = f(x, \varphi(x)) \quad \forall x \in J$  und  $\varphi(x_0) = y_0$ . Dann ist  $(G(\varphi))(x) = y_0 + \int_{x_0}^x \varphi'(t) \, dt = y_0 + (\varphi(x) - \varphi(x_0)) = \varphi(x) \quad \forall x \in J$ , also  $G(\varphi) = \varphi$ .

2)  $\Rightarrow$  1): Sei  $\varphi : J \rightarrow \mathbb{R}^n$  stetig mit  $G(\varphi) = \varphi$ , also

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

Dann ist  $\varphi(x_0) = y_0$ . Weil  $f$  stetig ist, ist  $\varphi$  differenzierbar mit  $\varphi'(x) = f(x, \varphi(x)) \quad \forall x \in J$  □

**Satz 1.** Sei  $I$  ein kompaktes Intervall und  $C(I)$  der Vektorraum aller stetigen Funktionen  $f : I \rightarrow \mathbb{R}$ . Auf  $C(I)$  betrachten wir die Norm

$$\|f\| := \max\{|f(x)| \mid x \in I\}$$

Damit wird  $C(I)$  zu einem Banach-Raum (d.h. aufgefasst als metrischer Raum ist  $C(I)$  vollständig).

**Bemerkung 1.** Konvergenz im normierten Raum  $C(I)$  bedeutet gleichmäßige Konvergenz.

**Erinnerung:** Seien  $f_n (n \in \mathbb{N})$  und  $f$  Funktionen von  $I$  in  $\mathbb{R}$ .

Wir sagen: Die Folge  $(f_n)$  konvergiert gleichmäßig gegen  $f$ , wenn es zu jedem  $\varepsilon > 0$  ein  $N \in \mathbb{N}$  gibt, sodass  $|f_n(x) - f(x)| < \varepsilon \quad \forall n \geq N, \forall x \in I$ .

Wir wissen: Ist  $(f_n)$  eine Folge stetiger Funktionen, die gleichmäßig gegen  $f$  konvergiert, so ist  $f$  stetig.

*Beweis.* von Satz 1. Wir müssen zeigen: Ist  $(f_n)$  eine Cauchy-Folge in  $C(I)$  so konvergiert  $(f_n)$ . Für jedes  $x \in I$  ist  $(f_n(x))_n$  eine Cauchy-Folge in  $\mathbb{R}$ , konvergiert also gegen eine Zahl  $f(x) \in \mathbb{R}$ . Damit haben wir eine Funktion  $f : I \rightarrow \mathbb{R}$  gefunden.

Die Folge  $(f_n)$  konvergiert gleichmäßig gegen  $f$ :

Sei  $\varepsilon > 0$ . Es gibt ein  $N$ , sodass  $|f_n(x) - f_m(x)| < \frac{\varepsilon}{2} \quad \forall n, m \geq N, x \in I$ .

Sei  $x \in I$  und  $n \geq N$ . Es gibt ein  $m \geq N$  mit  $|f_m(x) - f(x)| < \frac{\varepsilon}{2}$ , weil  $(f_m(x))_m$  gegen  $f(x)$  konvergiert.

Dann ist  $|f_n(x) - f(x)| \leq |f_n(x) - f_m(x)| + |f_m(x) - f(x)| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$ .

Deswegen ist  $f$  stetig, d.h.  $f \in C(I)$ . Und  $(f_n)$  konvergiert in  $C(I)$  gegen  $f$ .  $\square$

**Bemerkung 2.** Für  $n \in \mathbb{N}$  sei  $C(I; \mathbb{R}^n)$  der  $\mathbb{R}$ -Vektorraum der stetigen Funktionen  $I \rightarrow \mathbb{R}^n$ .

Man wählt eine Norm  $\|\cdot\|$  auf  $\mathbb{R}^n$  und definiert  $\|f\| := \max_{x \in I} \|f(x)\|$  für  $f \in C(I)$ .

Damit wird  $C(I; \mathbb{R}^n)$  ein Banach-Raum.

**Satz 2. Lokale Existenz- und Eindeutigkeitsatz von Picard-Lindelöf.** Sei  $U$  offen in  $\mathbb{R} \times \mathbb{R}^n$  und  $f : U \rightarrow \mathbb{R}^n$  sei stetig und lokal Lipschitz-stetig im 2. Argument. Sei  $(x_0, y_0) \in U$ . Dann gibt es eine Lösung  $\varphi : I \rightarrow \mathbb{R}^n$  der DGL  $y' = f(x, y)$  mit folgenden Eigenschaften:

1.  $\varphi(x_0) = y_0$
2. Ist  $J$  ein offenes Intervall mit  $x_0 \in J$  und ist  $\psi : J \rightarrow \mathbb{R}^n$  eine Lösung der DGL mit  $\psi(x_0) = y_0$ , so ist  $J \subseteq I$  und  $\psi = \varphi|_J$

*Beweis.* Für den Beweis vom Satz benutzen wir den Banachschen Fixpunktsatz in der Version von §11, Satz 2:

Sei  $X$  ein vollständiger metrischer Raum, sei  $R > 0$ ,  $y_0 \in X$  und  $B := B_R(y_0)$ . Sei  $G : B \rightarrow X$  eine Abbildung mit:

1.  $D(G(x), G(y)) \leq C \cdot d(x, y)$  mit einem  $C$  mit  $0 \leq C < 1$  für alle  $x, y \in B$ .
2.  $d(x_0, G(y_0)) < R \cdot (1 - C)$

Dann gibt es genau ein  $x \in B$  mit  $G(x) = x$ .  $\square$

**Bemerkung 3.**  $\varphi$  ist Lösung von  $y' = f(x, y)$  mit  $\varphi(x_0) = y_0 \Leftrightarrow \varphi(x) = y_0 + \int_{x_0}^x f(t, \varphi(t)) dt$ .

Genauer: Lemma 1: Sei  $I$  offenes Intervall,  $H$  sei offen in  $\mathbb{R}^n$ ,  $f : I \times H \rightarrow \mathbb{R}^n$  stetig.

Sei  $(x_0, y_0) \in I \times H$ . Für eine stetige Funktion  $\varphi : I \rightarrow H$  sei  $G(\varphi) : I \rightarrow \mathbb{R}^n$  definiert durch  $(G(\varphi)) := y_0 + \int_{x_0}^x f(t, \varphi(t)) dt$ .

Dann ist  $\varphi : I \rightarrow \mathbb{R}^n$  genau dann Lösung von  $y' = f(x, y)$  mit  $\varphi(x_0) = y_0$ , wenn  $G(\varphi) = \varphi$

**Lemma 2.** Sei  $(x_0, y_0) \in \mathbb{R} \times \mathbb{R}^n$ , sei  $I$  kompaktes Intervall mit  $x_0 \in I$ , sei  $R > 0$  und  $H := \{y \in \mathbb{R}^n \mid \|y - y_0\| < R\}$ .

Sei  $f : I \times H \rightarrow \mathbb{R}^n$  stetig. Seien  $L, M \in \mathbb{R}_{>0}$  mit

$$\|f(x, y) - f(x, \tilde{y})\| \leq L \cdot \|y - \tilde{y}\| \quad \forall x \in I, y, \tilde{y} \in H$$

$$\|f(x, y)\| \leq M \quad \forall x \in I, y \in H$$

Sei  $0 < r < \frac{R}{M + LR}$  und  $J := [x_0 - r, x_0 + r] \cap I$ . Sei  $X = C(J; \mathbb{R}^n)$ ,  $B := \{\psi \in X \mid \|\psi - y_0\| < R\}$ .

Definiere  $G : B \rightarrow X$  durch  $(G(\varphi))(x) = y_0 + \int_{x_0}^x f(t, \varphi(t)) dt$ . Dann gibt es genau ein  $\varphi \in B$  mit  $G(\varphi) = \varphi$ .

*Beweis.* von Lemma 2: Wir wenden §11, Satz 2 an:

Prüfe die Voraussetzung 1) und 2) dieses Satzes nach:

1. Seien  $\varphi, \psi \in B$ . Dann ist

$$\begin{aligned}
\|G(\varphi) - G(\psi)\| &= \max_{x \in J} \|(G(\varphi))(x) - (G(\psi))(x)\| \\
&= \max_{x \in J} \left\| \int_{x_0}^x (f(t, \varphi(t)) - f(t, \psi(t))) dt \right\| \leq \max_{x \in J} \left| \int_{x_0}^x \|f(t, \varphi(t)) - f(t, \psi(t))\| dt \right| \\
&\leq \max_{x \in J} \left| \int_{x_0}^x L \cdot \|\varphi(t) - \psi(t)\| dt \right| \leq \max_{x \in J} \left| \int_{x_0}^x L \cdot \|\varphi - \psi\| dt \right| = \max_{x \in J} L \cdot \|\varphi - \psi\| \cdot |x - x_0| \\
&\leq L \cdot r \cdot \|\varphi - \psi\| \leq \underbrace{\frac{LR}{M + LR}}_{=: C < 1} \cdot \|\varphi - \psi\| = C \cdot \|\varphi - \psi\|
\end{aligned}$$

2.

$$\begin{aligned}
\|G(y_0) - y_0\| &= \max_{x \in J} \|(G(y_0))(x) - y_0\| = \max_{x \in J} \left\| \int_{x_0}^x f(t, y_0) dt \right\| \\
&\leq r \cdot M < \frac{RM}{M + LR} = R \cdot (1 - C), \\
\text{denn } 1 - C &= 1 - \frac{LR}{M + LR} = \frac{M + LR - LR}{M + LR} = \frac{M}{M + LR}
\end{aligned}$$

□

Aus Lemma 1 und Lemma 2 folgt:

**Lemma 3.** Sei  $U$  offen in  $\mathbb{R} \times \mathbb{R}^n$ , sei  $f : U \rightarrow \mathbb{R}^n$  stetig und Lipschitz-stetig im 2. Argument. Sei  $(x_0, y_0) \in U$ . Dann gibt es ein offenes Intervall  $I$  mit  $x_0 \in I$  und genau eine Lösung  $\varphi : I \rightarrow \mathbb{R}^n$  von  $y' = f(x, y)$  mit  $\varphi(x_0) = y_0$ .

**Lemma 4.** Bezeichnungen und Voraussetzungen seien wie in Lemma 3. Ferner seien  $I_1, I_2$  zwei offene Intervalle mit  $x_0 \in I_1 \cap I_2 =: I_0$  und seien  $\varphi_i : I_i \rightarrow \mathbb{R}^n$  ( $i = 1, 2$ ). Lösungen von  $y' = f(x, y)$  mit  $\varphi_i(x_0) = y_0$ . Dann ist  $\varphi_1|_{I_0} = \varphi_2|_{I_0}$ .

*Beweis.*  $A := \{x \in I_0 \mid \varphi_1(x) = \varphi_2(x)\} \ni x_0$ . Zu zeigen:  $A = I_0$ .

Weil  $\varphi_i$  stetig sind, ist  $A$  abgeschlossen in  $I_0$ .

Ist  $A \neq I_0$  so gibt es ein nicht-leeres offenes Intervall  $]a, b[$  mit  $]a, b[ \subseteq I_0 \setminus A$ .

O.B.d.A. sei  $x_0 < a$ . Sei  $x_1 := \sup\{x \in I_0 \mid x \leq a \text{ und } x \in A\}$ . Weil  $A$  abgeschlossen in  $I_0$  ist, ist  $x_1 \in A$  Mengebild einfügen, also  $\varphi_1(x_1) = \varphi_2(x_1)$ . In jeder Umgebung von  $x_1$  liegen Punkte  $x$  mit  $\varphi_1(x) \neq \varphi_2(x) =: y_1$  im Widerspruch zu Lemma 2, angewandt auf  $(x_1, y_1)$  statt  $(x_0, y_0)$ . □

*Beweis.* von Satz 2. Sei  $\mathcal{M}$  die Menge aller offenen Intervalle  $J$  mit  $x_0 \in J$ , sodass es eine Lösung  $\varphi_J : J \rightarrow \mathbb{R}^n$  mit  $\varphi_J(x_0) = y_0$  gibt. Nach Lemma 3 ist  $\mathcal{M} \neq \emptyset$ , und nach Lemma 4 ist  $\varphi_J$  durch  $J$  eindeutig bestimmt.  $I := \bigcup_{J \in \mathcal{M}} J$  und definiere  $\varphi : I \rightarrow \mathbb{R}^n$  durch  $\varphi(x) := \varphi_J(x)$ , falls  $x \in J \in \mathcal{M}$ .

Dies ist wohldefiniert nach Lemma 4 und hat die gewünschten Eigenschaften. □