

Vorlesungsmitschrift

## **DIFF II**

**Prof. Dr. Dorothea Bahns**

Henry Ruben Fischer

Auf dem Stand vom 7. Juli 2020

---

**Disclaimer**

Nicht von Professor Bahns durchgesehene Mitschrift, keine Garantie auf Richtigkeit ihrerseits.

# Inhaltsverzeichnis

<b>1. Metrische Räume</b>	<b>6</b>
1.I. Charakterisierung topologischer Grundbegriffe in metrischen Räumen . . . . .	17
1.II. Vollständigkeit . . . . .	19
1.III. Betrachtungen in vollständigen metrischen Räumen . . . . .	21
1.IV. Stetige Abbildungen auf metrischen Räumen . . . . .	27
1.V. Kompaktheit . . . . .	30
1.VI. Äquivalenz von Metriken . . . . .	36
<b>2. Normierte Vektorräume</b>	<b>38</b>
2.I. Stetige Abbildungen in normierten Vektorräumen . . . . .	44
2.I.1. Lineare Abbildungen . . . . .	44
2.II. Vektorräume mit Skalarprodukt . . . . .	49
<b>3. Differenzierbarkeit in <math>\mathbb{R}^n</math></b>	<b>55</b>
3.I. Geometrische Anschauung, partielle Ableitung . . . . .	59
3.II. Beispiele und Erläuterungen . . . . .	63
3.III. Implizite Funktionen . . . . .	72
3.IV. Der Satz von der Umkehrabbildung . . . . .	81
3.V. Lokale Extrema unter Nebenbedingungen . . . . .	88
3.VI. Höhere Ableitungen, Taylorformel . . . . .	92
3.VII Der Laplace-Operator . . . . .	95
3.VIII Taylor-Formel, lokale Extrema . . . . .	100
3.IX. Lokale Extrema . . . . .	103
<b>4. Untermannigfaltigkeiten des <math>\mathbb{R}^n</math></b>	<b>109</b>
4.I. Tangential- und Normalraum . . . . .	130
4.II. Flächenbemessung auf Untermannigfaltigkeiten . . . . .	138
<b>5. Differentialgleichungen</b>	<b>143</b>
5.I. Geometrische Interpretation . . . . .	143
5.II. Existenz- und Eindeutigkeitssatz . . . . .	145
5.III. Lineare Differentialgleichungen . . . . .	160
5.IV. Lineare DGL-Systeme mit konstanten Koeffizienten . . . . .	172

*Inhaltsverzeichnis*

---

<b>6. Lebesgue-Integration</b>	<b>182</b>
6.I. Etwas Maßtheorie . . . . .	190
6.II. Weitere Folgerungen . . . . .	192
6.III. Messbare Funktionen . . . . .	196
6.IV. Zum Verhältnis von Lebesgue- / Riemann-Integral . . . . .	197
6.V. Produkt-Maße . . . . .	200
6.VI. Der Transformationssatz . . . . .	207

# **Vorlesungsverzeichnis**

1.	Mo 20.04. 10:15 . . . . .	6
2.	Do 23.04. 10:15 . . . . .	16
3.	Mo 27.04. 10:15 . . . . .	25
4.	Do 30.04. 10:15 . . . . .	36
5.	Mo 04.05. 10:15 . . . . .	44
6.	Do 07.05. 10:15 . . . . .	55
7.	Mo 11.05. 10:15 . . . . .	63
8.	Do 14.05. 10:15 . . . . .	74
9.	Mo 17.05. 10:15 . . . . .	80
10.	Do 21.05. 10:15 . . . . .	88
11.	Mo 25.05. 10:15 . . . . .	100
12.	Do 28.05. 10:15 . . . . .	109
13.	Do 04.06. 10:15 . . . . .	122
14.	Mo 08.06. 10:15 . . . . .	130
15.	Do 11.06. 10:15 . . . . .	143
16.	Mo 15.06. 10:15 . . . . .	153
17.	Do 18.06. 10:15 . . . . .	162
18.	Mo 15.06. 10:15 . . . . .	171
19.	Do 25.10. 10:15 . . . . .	182
20.	Mo 29.06. 10:15 . . . . .	190
21.	Do 02.07. 10:15 . . . . .	199
22.	Mo 06.07. 10:15 . . . . .	207

# Kapitel 1.

## Metrische Räume

### Vorlesung 1

Mo 20.04. 10:15

**Ziel.** Konvergenz, Stetigkeit ... sollten in einem allgemeineren Rahmen konzeptualisiert werden.

**Erinnerung (DIFF I).** Eine Folge  $(a_n)_{n \in \mathbb{N}} \subset \mathbb{R}$  konvergiert gegen den Grenzwert  $a$

$$\iff \forall \varepsilon > 0 \ \exists N \in \mathbb{N} \text{ s. d. } |a_n - a| < \varepsilon \ \forall n \geq N$$

$(a - \varepsilon, a + \varepsilon)$  wird auch  $\varepsilon$ -Umgebung von  $a$  in  $\mathbb{R}$  genannt. Somit lautet die obige Definition in Worten: In jeder noch so kleinen  $\varepsilon$ -Umgebung von  $a$  befinden sich alle bis auf endlich viele Folgenglieder.

Man benötigt für die Formulierung der Definition also lediglich einen Begriff von „(kleine) Umgebung“. Diesen Begriff möchten wir nun verallgemeinern.

**Definition 1.1.** Sei  $X$  eine Menge. Ein System  $\mathcal{T}$  von Teilmengen von  $X$  heißt Topologie auf  $X$  falls gilt:

- $\emptyset, X \in \mathcal{T}$ .
- Sind  $U$  und  $V \in \mathcal{T}$ , so gilt  $U \cap V \in \mathcal{T}$ .
- Ist  $I$  eine Indexmenge und  $U_i \in \mathcal{T}$  für alle  $i \in I$ , so gilt auch  $\bigcup_{i \in I} U_i \in \mathcal{T}$ .

**Notation.** Ein topologischer Raum ist ein Tupel  $(X, \mathcal{T})$ , wobei  $X$  Menge ist und  $\mathcal{T}$  eine Topologie auf  $X$ .

Eine Teilmenge  $U \subset X$  heißt *offen*, falls gilt  $U \in \mathcal{T}$ . Eine Teilmenge  $A \subset X$  heißt *abgeschlossen* falls ihr Komplement  $X \setminus A$  offen ist.

**Beispiele 1.2.** i)  $X = \text{beliebige Menge. } \mathcal{T} = \{ \emptyset, X \}$ .

*Beweis.* 1.1.a) klar

1.1.b)  $\emptyset \cap X = \emptyset \in \mathcal{T}$ ,  $X \cap X = X \in \mathcal{T}$ ,  $\emptyset \cap \emptyset = \emptyset \in \mathcal{T}$

1.1.c)  $\bigcup_{i \in I} U_i = \begin{cases} X & \text{falls eins der } U_i = X \text{ ist} \\ \emptyset & \text{falls nicht} \end{cases}$   $\square$

„Klumpentopologie“

ii)  $X = \mathbb{R}$

$\mathcal{T}$  = alle Teilmengen  $U \subset \mathbb{R}$  mit der Eigenschaft:

$$\forall x \in U \quad \exists \varepsilon > 0 \text{ s. d. } (x - \varepsilon, x + \varepsilon) \subset U$$

Beweis von 1.1.a), 1.1.b) und 1.1.c) als HA (etwas allgemeiner). Hier stellen wir fest, dass insbesondere die offenen Intervalle  $(a, b)$  in diesem Sinne offen (also  $\in \mathcal{T}$ ) sind, halb-abgeschlossene und abgeschlossene dagegen nicht.

*Beweis.* **1. Beh** Zu  $x \in [a, b]$  wähle  $\varepsilon = \min\{\{|x - a|, |x - b|\}\}$

**2. Beh** Zu  $x = a \in [a, b]$  kann man kein  $\varepsilon > 0$  finden s. d.  $(a - \varepsilon, a + \varepsilon) \subset [a, b]$ , denn  $a - \varepsilon/2 \in (a - \varepsilon, a + \varepsilon)$  aber  $a - \varepsilon/2 < a$ , also  $\notin [a, b]$ .  $\square$

Abgeschlossene Intervalle sind in diesem Sinn abgeschlossen, denn  $\mathbb{R} \setminus [a, b]$  ist nach Definition von  $\mathcal{T}$  und Eigenschaft 1.1.c) offen.

Diese Topologie heißt Standard-Topologie auf  $\mathbb{R}$ . Wird nichts anderes gesagt, sehen wir  $\mathbb{R}$  als mit der Standard-Topologie versehen an.

**Definition 1.3.** Sei  $(X, \mathcal{T})$  topologischer Raum. Sei  $x \in X$ . Eine Teilmenge  $V \subset X$  heißt *Umgebung von  $x$* , falls es eine offenen Menge  $U \subset X$  gibt mit  $x \in U$  und  $U \subset V$ .

**Beispiele.** i)  $V = (a, b)$  ist eine Umgebung für jedes  $x \in (a, b)$ , aber *nicht* für  $x = a$ .



ii)  $(a - \varepsilon, a + \varepsilon)$ ,  $\varepsilon > 0$ , ist eine Umgebung von  $a$ .

**Lemma 1.4.** Eine Teilmenge  $V \subset X$  eines topologischen Raumes  $(X, \mathcal{T})$  ist offen gdw für alle  $x \in V$  gilt:  $V$  ist Umgebung von  $x$ .

*Beweis.* „ $\Rightarrow$ “ Ist  $V$  offen, so erfüllt  $U = V$  für jedes  $x$  die Bedingung  $x \in U$  und  $U \subset V \Rightarrow V$  ist Umgebung.

„ $\Leftarrow$ “ Zu  $x \in V$  wähle  $U_x$  s.d.  $x \in U_x$ ,  $U \subset V$ . Dann gilt  $V = \bigcup_{x \in U} U_x$  und das ist offen (nach 1.1.c)).  $\square$

**Definition 1.5 (Konvergenz in topologischen Räumen).** Sei  $(X, \mathcal{T})$  topologischer Raum. Sei  $(x_n)_{n \in \mathbb{N}}$  eine Folge in  $X$ . Dann ist  $(x_n)_n$  konvergent mit Grenzwert  $x$ ,  $x_n \rightarrow x$  in  $(X, \mathcal{T})$ , falls es in jeder Umgebung  $V$  von  $x$  ein  $N \in \mathbb{N}$  gibt, s.d.  $x_n \in V \quad \forall n \geq N$ .

**Beispiele.** i) In der Klumpentopologie konvergieren alle Folgen gegen jedes  $x \in X$ .

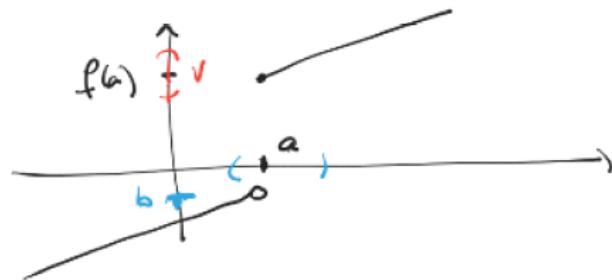
ii) Mit unseren obigen Überlegungen folgern wir, dass Konvergenz in  $\mathbb{R}$  im Sinn von Definition 1.5 mit Konvergenz, wie wir sie in der DIFF I kennengelernt haben.

**Lemma 1.6.** Sei  $(X, \mathcal{T})$  topologischer Raum. Ist  $(X, \mathcal{T})$  ein Hausdorff-Raum, gibt es also zu je zwei Punkten  $x, y \in X$  mit  $x \neq y$  Umgebungen  $U$  von  $x$  und  $V$  von  $y$  mit  $U \cap V = \emptyset$ , so ist der Grenzwert einer konvergenten Folge eindeutig.

*Beweis.* Seien  $x$  und  $y$  Grenzwert einer Folge  $(x_n)_n$ . Angenommen  $x \neq y$ , so wähle  $U$  Umgebung von  $x$ ,  $V$  Umgebung von  $y$  mit  $U \cap V = \emptyset$ . Dann gibt es (wegen der Konvergenz)  $N \in \mathbb{N}$  s.d.  $x_n \in U \quad \forall n \geq N$  und  $M \in \mathbb{N}$  s.d.  $x_n \in V \quad \forall n \geq M$ . Widerspruch zu  $U \cap V = \emptyset$ .  $\square$

**Definition 1.7.** Seien  $(X, \mathcal{T})$  und  $(Y, \tilde{\mathcal{T}})$  topologische Räume. Sei  $f: X \rightarrow Y$  eine Abbildung. Dann heißt  $f$  stetig in  $a \in X$ , falls es zu jeder Umgebung  $V$  von  $f(a) \in Y$  eine Umgebung  $U$  von  $a$  gibt, s.d.  $f(U) \subset V$ .  $f$  heißt stetig (auf  $X$ ), falls  $f$  stetig in allen  $a \in X$  ist.

**Bemerkung.** Wir werden später sehen, dass diese Definition für  $f: \mathbb{R} \rightarrow \mathbb{R}$  mit unserer Definition aus der DIFF I übereinstimmt ( $\varepsilon$ - $\delta$ -Kriterium).



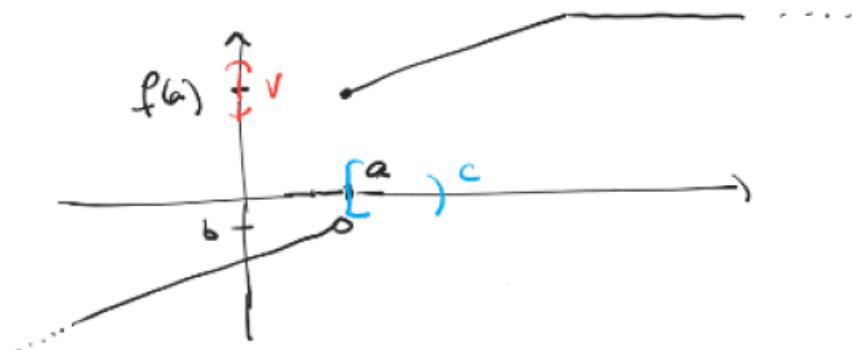
Für jede Umgebung  $U$  von  $a$  gilt:  $f(U)$  enthält auch Punkte  $< b$ , also außerhalb  $V$

**Satz 1.8.** Sei  $f: X \rightarrow Y$  Abbildung zwischen topologischen Räumen. Dann ist  $f$  stetig auf  $X$  gdw für jede offene Teilmenge  $V \subset Y$  das Urbild  $f^{-1}(V)$ , also  $\{x \in X \mid f(x) \in V\}$  offen in  $X$  ist.

*Beweis.* „ $\Rightarrow$ “ Sei  $f$  stetig vorausgesetzt. Sei  $V$  offen in  $Y$ . Ist das Urbild  $f^{-1}(V)$  leer, sind wir fertig.

Sei also  $a \in f^{-1}(V)$ . Dann gibt es nach Voraussetzung eine Umgebung  $U$  von  $a$  s.d.  $f(U) \subset V$ . Also gilt  $U \subset f^{-1}(V)$ . Somit besitzt also jeder Punkt  $a \in f^{-1}(V)$  eine Umgebung  $U$  mit  $U \subset f^{-1}(V)$  und somit ist  $f^{-1}(V)$  selbst Umgebung jedes seiner Elemente  $\stackrel{1.4}{\Rightarrow} f^{-1}(V)$  ist offen.

„ $\Leftarrow$ “ Sei  $a \in X$  beliebig. Sei  $V$  eine Umgebung von  $f(a)$ . Dann gibt es  $\tilde{V}$  offen mit  $f(a) \in \tilde{V}$  und  $\tilde{V} \subset V$ . Nach Voraussetzung ist das Urbild  $U := f^{-1}(\tilde{V})$  offen.  $U$  enthält  $a$ , ist also Umgebung von  $a$  und es gilt  $f(U) = \tilde{V} \subset V \Rightarrow f$  ist stetig in  $a$ .  $\square$



$$f^{-1}(V) = [a, c) \text{ ist nicht offen in } \mathbb{R}$$

**Bemerkung.** Äquivalent:  $f$  ist genau dann stetig, wenn das Bild jeder abgeschlossenen Menge abgeschlossen ist.

### Vorsicht:

Es ist immer Offenheit in  $X$  (bzw.  $Y$ ) gemeint!

**Zur Veranschaulichung:**

Betrachtet man im Beispiel oben als Definitionsbereich  $X = [a, \infty)$ , so ist die Funktion stetig! Dies ist konsistent, da  $[a, c)$  in  $X = [a, \infty)$  versehen mit der Standard-Topologie tatsächlich offen ist:

**Definition / Satz 1.9.** Sei  $(X, \mathcal{T})$  topologischer Raum. Sei  $\tilde{X} \subset X$  eine Teilmenge. Dann induziert  $\mathcal{T}$  auf  $\tilde{X}$  eine Topologie, die sogenannte *Teilraum-Topologie* vermöge

$$\mathcal{T}_{\tilde{X}} := \{ U \cap \tilde{X} \mid U \in \mathcal{T} \}.$$

Den (einfachen) Beweis, dass dies in der Tat eine Topologie definiert, lassen wir weg.

In unserem Beispiel ist  $X = \mathbb{R}$ ,  $\tilde{X} = [a, \infty)$  und da  $(a - \varepsilon, c)$  offen in  $\mathbb{R}$  ist ( $\varepsilon > 0$ ), ist nach Definitionsbereich  $[a, c) = (a - \varepsilon, c) \cap [a, \infty)$  offen in  $[a, \infty)$ .

Dies ist der tiefere Grund, weshalb man bei Funktionen den Raum, in dem sie ihre Werte annehmen (im Beispiel oben  $Y = \mathbb{R}$ ) angeben sollte, nicht ihr Bild.

Denn in  $Y = (-\infty, b) \cup [f(a), \infty)$  wäre das Bild von  $[a - \varepsilon, c] \forall \varepsilon > 0$  in der Tat abgeschlossen, denn sein Komplement

$$Y \setminus ([b - \delta, b) \cup [f(a), f(c)]) = -(-\infty, b - \delta) \cup (f(c), \infty)$$

wäre offen.

Dagegen ist

$$\mathbb{R} \setminus ([b - \delta, b) \cup [f(a), f(c)]) = -(-\infty, b - \delta) \cup [b, f(a)] \cup (f(c), \infty)$$

für kein  $\delta > 0$  offen.

**Definition / Satz 1.10.** Seien  $(X, \mathcal{T}_X)$  und  $(Y, \mathcal{T}_Y)$  topologische Räume. Betrachte das *kartesische Produkt*  $X \times Y = \{ (x, y) \mid x \in X, y \in Y \}$ . Dann nennt man das System

$$\mathcal{T} := \left\{ U \subset X \times Y \mid \begin{array}{l} U = \text{beliebige Vereinigung von Mengen der Form} \\ V \times W, V \in \mathcal{T}_X, W \in \mathcal{T}_Y \end{array} \right\}$$

*Produkttopologie*. Und dies definiert in der Tat eine Topologie auf  $X \times Y$ .

*Beweis.* 1.1.a) klar

1.1.b)

$$\begin{aligned} U &= \bigcup_{\alpha} U_{\alpha} \times W_{\alpha} \\ V &= \bigcup_{\beta} \tilde{V}_{\beta} \times \tilde{W}_{\beta} \\ U \cap V &= \bigcup_{\alpha, \beta} \underbrace{(V_{\alpha} \cap \tilde{V}_{\beta})}_{\text{offen in } X} \times \underbrace{(W_{\alpha} \cap \tilde{W}_{\beta})}_{\text{offen in } Y}. \end{aligned}$$

1.1.c)

$$\bigcup_{\rho} \left( \bigcup_{\alpha} V_{\alpha}^{(\rho)} \times W_{\alpha}^{(\rho)} \right) = \bigcup_{\rho, \alpha} V_{\alpha}^{(\rho)} \times W_{\alpha}^{(\rho)}. \quad \square$$

Wir kommen nun zu einer wichtigen Beispiel-Klasse für Topologien:

**Definition 1.11.** Sei  $X$  eine Menge. Eine *Metrik* auf  $X$  ist eine Abbildung

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

mit den Eigenschaften

- a)  $d(x, y) = 0 \iff x = y$  „ $d$  ist nicht ausgeartet.“
- b)  $d(x, y) = d(y, x) \quad \forall x, y \in X$  „ $d$  ist symmetrisch.“
- c)  $d(x, y) \leq d(x, z) + d(z, y) \quad \forall x, y, z \in X$  „Es gilt die Dreiecksungleichung.“

Ein *metrischer Raum* ist ein Tupel  $(X, d)$ , wobei  $X$  eine Menge ist und  $d$  eine Metrik auf  $X$ . Meist schreibt man nur  $X$ , weil Missverständnisse ausgeschlossen sind.

**Bemerkung.** Aus den Axiomen folgt auch

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

denn

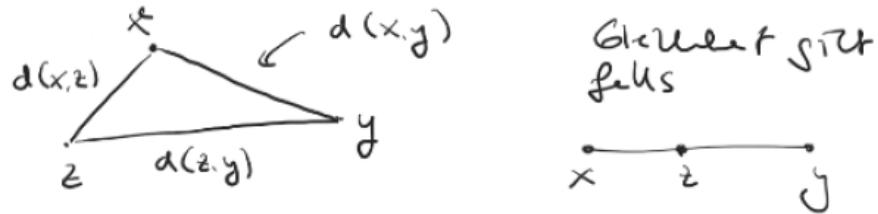
$$0 = d(x, x) \stackrel{\substack{\uparrow \\ 1.11.a)}{\leq} d(x, y) + d(y, x) = \stackrel{\substack{\uparrow \\ \triangle\text{-Ungl.}}}{2d(x, y)} \stackrel{\substack{\uparrow \\ \text{Symm.}}}{= 2d(x, y).$$

**Beispiele.** i)  $\mathbb{R}, d(x, y) = |x - y|$ .

ii)  $X$  Menge,  $d(x, y) = \begin{cases} 1 & x \neq y \\ 0 & x = y \end{cases}$ , „triviale“ oder „diskrete Metrik“.

iii) (aus AGLA I)  $\mathbb{R}^n$ ,  $d(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$ , „Euklidische Metrik“.

Eine Metrik misst den *Abstand* zwischen zwei Punkten. Im zweiten Beispiel sind alle verschiedenen Punkte gleich weit von einander entfernt. Für  $n = 1$  stimmt iii) mit i) überein. Mit iii) wird auch der Name der Dreiecksungleichung klar:



**Definition 1.12.** Sei  $(X, d)$  ein metrischer Raum. Seien  $x \in X$ ,  $\varepsilon > 0$ . Dann nennt man

$$B_\varepsilon(x) := \{ y \in X \mid d(x, y) < \varepsilon \}$$

den (offenen)  $\varepsilon$ -Ball um  $x$ .

**Beispiele.** i)  $B_\varepsilon(x) = (x - \varepsilon, x + \varepsilon)$ .

$$\text{ii)} \quad B_\varepsilon(x) = \begin{cases} x & \varepsilon \leq 1 \\ X & \varepsilon > 1 \end{cases}$$

$$\text{iii)} \quad B_\varepsilon(x) = \text{ (Diagramm eines Kreises mit gestrichelter Linie und Kreuzlinien im Inneren)}$$

**Satz 1.13.** Sei  $(X, d)$  ein metrischer Raum. Dann wird durch

$$\mathcal{T}_d := \{ U \subset X \mid \forall x \in U \ \exists \varepsilon > 0 \text{ s.d. } B_\varepsilon(x) \subset U \}$$

eine Topologie definiert.

*Beweis.* Als Hausaufgabe. □

**Bemerkungen 1.14.** i) 1.2.ii) ist ein Spezialfall dieser Aussage

ii) Die „offenen“  $\varepsilon$ -Bälle sind tatsächlich offen: Zu  $y \in B_\varepsilon(x)$  wähle  $\tilde{\varepsilon} := \varepsilon - d(x, y) > 0$ .



Dann ist  $B_{\tilde{\varepsilon}}(y) = \{ z \mid d(y, z) < \tilde{\varepsilon} \} \subset B_\varepsilon(x)$ . Denn für alle  $z \in B_{\tilde{\varepsilon}}(y)$  ist

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

- iii) Bezuglich der diskreten Metrik ist *jede* Teilmenge offen.
- iv) Die Klumpentopologie wird nicht von einer Metrik erzeugt (wenn  $X$  mehr als 1 Element enthält).

*Beweis.* Seien  $x, y \in X$ ,  $x \neq y$ . Angenommen  $\exists$  Metrik  $d$ .

$$\begin{aligned} \implies d(x, y) &\neq 0 \implies d(x, y) = c > 0 \\ \implies B_c(x) &\text{ ist offen.} \\ \implies B_c(x) &= \emptyset \text{ oder } = X \\ \implies B_c(x) &= X \end{aligned}$$

□

↳ , da  $y \notin B_c(x)$ .

- v) Ein metrischer Raum ist hausdorffsch. → HA.

Wir formulieren nun Konvergenz und Stetigkeit für metrische Räume:

**Bemerkungen 1.15.** Sei  $(X, d)$  metrischer Raum.

- i) [Definition 1.3]  $V \subset X$  heißt Umgebung von  $x \in X$ , falls es  $\varepsilon > 0$  gibt s.d.  $B_\varepsilon(x) \subset U$ .
- ii) [Definition 1.5]  $(x_n)_n$  konvergiert mit Grenzwert  $x$ , falls es zu jedem  $\varepsilon > 0$  ein  $N \in \mathbb{N}$  gibt s.d.  $x_n \in B_\varepsilon(x) \quad \forall n \geq N$ .

- iii) [Definition 1.7] Sei  $(Y, \tilde{d})$  weiterer metrischer Raum,  $f: X \rightarrow Y$  eine Abbildung. Dann ist  $f$  stetig  $a$  gdw :

$$\forall \varepsilon > 0 \quad \exists \delta > 0 \text{ s.d. } f(B_\delta(a)) \subset B_\varepsilon(f(a)).$$

**Bemerkungen.** i) 1.15.iii) ist das  $\varepsilon$ - $\delta$ -Kriterium.

- ii) Die Einschränkung auf  $\varepsilon$ -Bälle in 1.15.ii) und 1.15.iii) (statt allgemeiner Umgebungen) ist keine echte Einschränkung: Gilt etwas für all Umgebungen, so speziell auch für  $\varepsilon$ -Bälle.

Und gilt eine Inklusion für alle  $\varepsilon$ -Bälle (etwa  $x_n \in B_\varepsilon(x) \quad \forall n \geq N(\varepsilon)$ ), so auch für beliebige Umgebungen  $U$  von  $x$ , da es immer einen  $\varepsilon$ -Ball  $B_\varepsilon(x)$  gibt, der ganz in  $U$  enthalten ist.

**Beispiele 1.16.** i)  $\mathbb{R}^m$  mit der Euklidischen Metrik.  $(x_n)_{n \geq 1}$  Folge in  $\mathbb{R}^m$ , also  $n \mapsto x_n = (x_n^{(1)}, \dots, x_n^{(m)}) \in \mathbb{R}^m$ .

$$\text{ii) } x_n = \left( \frac{1}{n} \cos(n), \frac{1}{n} \sin(n), a, \dots, a \right)$$

**Behauptung.**  $x_n \rightarrow (0, 0, a, \dots, a) =: x$ .

*Beweis.* Sei  $\varepsilon > 0$ . Es gilt

$$\begin{aligned} d(x_n, x)^2 &= \sum_{i=1}^m (x_n^{(i)} - x^{(i)})^2 \\ &= \left( \frac{1}{n} \cos(n) - 0 \right)^2 + \left( \frac{1}{n} \sin(n) - 0 \right)^2 + (a - a)^2 + \dots + (a - a)^2 \\ &= \frac{1}{n^2} (\cos(n)^2 + \sin(n)^2) = \frac{1}{n^2} \\ \implies d(x_n, x) &= \frac{1}{n} \\ \implies d(x_n, x) &< \varepsilon \quad \forall n \geq N \text{ mit } N > \frac{1}{\varepsilon} \\ \implies x_n &\in B_\varepsilon(x) \quad \forall n \geq N. \end{aligned}$$

□

- iii)  $X = C([a, b])$ ,  $d(f, g) := \|f - g\|_\infty$  mit  $\|f - g\|_\infty = \sup_{x \in [a, b]} |f(x) - g(x)|$ .

**1. Beh**  $d$  ist eine Metrik auf  $X$ .

*Beweis.* 1.11.a):

$$\begin{aligned} \sup_{x \in [a, b]} |f(x) - g(x)| &= 0 \\ \iff |f(x) - g(x)| &= 0 \quad \forall x \\ \iff f(x) &= g(x) \quad \forall x. \end{aligned}$$

1.11.b):

$$\begin{aligned} |f(x) - g(x)| &= |g(x) - f(x)| \quad \forall x \\ \implies d(f, g) &= d(g, f). \end{aligned}$$

1.11.c):

$$\begin{aligned} |f(x) - g(x)| &= |f(x) - h(x) + h(x) - g(x)| \\ &\leq |f(x) - h(x)| + |h(x) - g(x)| \\ &\quad \uparrow \text{△-Ungl. für } |\cdot| \text{ auf } \mathbb{R} \\ \implies \triangle &\text{-Ungl. für } d. \end{aligned}$$

□

**2. Beh**  $(f_n)_n \subset C([0, 1])$ ,  $f_n(x) = x^n$ , konvergiert nicht (vgl. DIFF I).

*Beweis.* Wir wissen aus der DIFF I, dass wenn Konvergenz vorliegt, der Grenzwert gleich dem punktweisen Grenzwert ist. Dieser ist

$$f(x) = \begin{cases} 1 & x = 1 \\ 0 & \text{sonst} \end{cases}.$$

Aber

$$\sup_{x \in [0, 1]} |f_n(x) - f(x)| = \sup_{x \in [0, 1]} |x^n| = 1.$$

□

iv)  $X = C([0, 1])$ ,  $d(f, g) = \int_0^1 |f(x) - g(x)| dx$ .

**1. Beh**  $d$  ist eine Metrik auf  $C([0, 1])$ .

*Beweis.* HA.

□

**2. Beh**  $(f_n)_n \subset C([0, 1])$ ,  $f_n(x) = x^n$  konvergiert, und zwar gegen  $f(x) = 0 \quad \forall x$ .

*Beweis.*

$$\int_0^1 |f_n(x) - 0| dx \implies d(f_n, f) = \frac{1}{n+1} < \varepsilon \quad \forall n \geq N \text{ mit } N \geq \frac{1}{\varepsilon}. \quad \square$$

sie

**Vorlesung 2**

Do 23.04. 10:15

Bevor wir uns mit offenen und abgeschlossenen Mengen und sogenannten vollständigen metrischen Räumen näher befassen, beweisen wir noch zwei nützliche Lemmata zu Konvergenz und Stetigkeit:

**Lemma 1.17.**  $(X, d)$  sei metrischer Raum.

Eine Folge  $(x_n)_n$  in  $X$  konvergiert in  $X$  gegen  $a \in X$

$$\iff (d(x_n, a))_n \text{ ist Nullfolge (in } \mathbb{R}).$$

*Beweis.*

$$d(x_n, a) = |d(x_n, a) - 0|.$$

Also ist

$$d(x_n, a) < \varepsilon \iff |d(x_n, a) - 0| < \varepsilon. \quad \square$$

**Lemma 1.18.** Seien  $(X, d_x)$  und  $(Y, d_y)$  metrische Räume,  $f: X \rightarrow Y$  eine Abbildung. Dann gilt:

$f$  ist in  $a \in X$  stetig  $\iff$  für jede Folge  $(a_n)_n$  mit  $a_n \rightarrow a$  in  $X$  gilt

$$\lim_{n \rightarrow \infty} f(a_n) = f(\underbrace{\lim_{n \rightarrow \infty} a_n}_{=a}).$$

**Notation.**

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

*Beweis.* „ $\implies$ “ Sei das  $\varepsilon$ - $\delta$ -Kriterium erfüllt (1.15.iii)). Sei  $(x_n)_n$  Folge in  $X$  mit  $x_n \rightarrow a$  in  $X$ . Sei  $\varepsilon > 0$ . Dann  $\exists \delta > 0$  s. d.

$$d_Y(f(x), f(a)) < \varepsilon \quad \forall x \in B_\delta(a) \subset X.$$

Wegen der Konvergenz  $\exists N = N(\delta)$  s. d.

$$\begin{aligned} x_n &\in B_\delta(a) \quad \forall n \geq N \\ \implies f(x_n) &\in B_\varepsilon(f(a)) \subset Y \quad \forall n \geq N. \end{aligned}$$

Also gilt  $f(x_n) \rightarrow f(a)$ .

„ $\Leftarrow$ “ Gelte  $\lim_{x \rightarrow a}(x) = f(a)$ .

Angenommen, das  $\varepsilon$ - $\delta$ -Kriterium wäre verletzt. Dann gäbe es  $\varepsilon > 0$  s. d. für alle  $\delta > 0$  ein  $x \in X$  existierte s. d.

$$\begin{aligned} x \in B_\delta(a) \text{ aber } f(x) \notin B_\varepsilon(f(a)) \\ \text{also } d_y(f(x), f(a)) \geq \varepsilon. \end{aligned}$$

Insbesondere gäbe es zu  $\delta = \frac{1}{n}$  ein solches  $x$ , nennen wir es  $x_n$ . Dann gilt für alle  $n$ :  $d(x_n, a) < \frac{1}{n}$ , aber  $d_y(f(x_n), f(a)) \geq \varepsilon$ , somit  $x_n \rightarrow a$  aber  $f(x_n) \not\rightarrow f(a)$  (wegen 1.17).  $\square$

## 1.I. Charakterisierung topologischer Grundbegriffe in metrischen Räumen

**Lemma 1.19.** Sei  $(X, d)$  metrischer Raum. Dann ist  $A \subset X$  abgeschlossen  $\iff$  für jede Folge  $(a_n)_n$ ,  $a_n \in A$ , die in  $X$  konvergiert, gilt:

$$\lim_{n \rightarrow \infty} a_n \in A.$$

*Beweis.* O.B.d.A.  $\emptyset \neq A \neq X$ .

„ $\Rightarrow$ “ Sei  $(a_n)_n$ ,  $a_n \in A$ , konvergent in  $X$ . Sei  $a = \lim a_n$ . Angenommen  $a \notin A$ . Nach Voraussetzung ist  $X \setminus A$  offen, also ist  $X \setminus A$  Umgebung von  $a \Rightarrow \exists N$  s. d.

$$a_n \in X \setminus A \quad \forall n \geq N \text{ (wegen Konvergenz)}$$

$\not\rightarrow$  zu  $a_n \in A$ .

„ $\Leftarrow$ “ Wir zeigen, dass  $X \setminus A$  offen ist. Sei also  $b \in X \setminus A$ . Es gibt  $\varepsilon > 0$  s. d.  $B_\varepsilon(b) \cap A = \emptyset$ , also  $B_\varepsilon(b) \subset X \setminus A$ .

Denn angenommen es gibt kein solches  $\varepsilon$ . Dann gilt für alle  $\varepsilon > 0$ :  $B_\varepsilon(b) \cap A \neq \emptyset$ , also kann man zu jedem  $k \geq 1$  ein  $x_k \in A$  finden mit  $d(x_k, b) < \frac{1}{k} = \varepsilon$ .

$$\Rightarrow x_k = b \underset{\text{VOR}}{\Rightarrow} b \in A.$$

$\not\rightarrow$  Widerspruch zu  $b \in X \setminus A$ .

Also gibt es ein solches  $\varepsilon > 0$ , also ist  $X \setminus A$  offen.  $\square$

**Definition 1.20.** Sei  $(X, d)$  metrischer Raum,  $M \subset X$ . Ein Punkt  $y \in X$  heißt *Randpunkt* von  $M$ , falls in jeder Umgebung von  $y$  sowohl Punkte von  $M$  als auch  $X \setminus M$  liegen.

**Notation.**  $\partial(M) = \{ \text{Randpunkte von } M \}$

**Beispiel**  $(\mathbb{R}^n, d_{\text{Eucl.}})$ . Kugel im  $\mathbb{R}^m$ :

$$K^n := \{ x \in \mathbb{R}^n \mid \|x - 0\|_E \leq R \} \subset \mathbb{R}^n$$

Sphäre:

$$\partial(K)^n = \{ x \in \mathbb{R}^n \mid \|x\|_E = R \} = S^{n-1}$$

**Beispiel.**  $\mathbb{Q} \subset \mathbb{R}$ .  $\partial(\mathbb{Q}) = \mathbb{R}$ .

**Satz 1.21.** Sei  $(X, d)$  metrischer Raum. Sei  $M \subset X$ . Dann gilt

- i)  $M \setminus \partial(M)$  ist offen.
- ii)  $M \cup \partial(M)$  ist abgeschlossen.
- iii)  $\partial(M)$  ist abgeschlossen.

*Beweis.* 1.21.i):  $a \in M \setminus \partial(M) \implies \exists \varepsilon > 0$  s.d.  $B_\varepsilon(a) \cap X \setminus M = \emptyset$ . Für dieses gilt auch  $B_\varepsilon \cap \partial(M) = \emptyset$  (denn angenommen  $\exists y \in B_\varepsilon(a) \cap \partial(M)$ , dann wäre (da  $y \in \partial(M)$  und  $B_\varepsilon(a)$  eine Umgebung von  $y$ )  $B_\varepsilon(a) \cap (X \setminus M) \neq \emptyset \not\perp \text{ VOR}$ ).

Also gilt  $B_\varepsilon(a) \subset M \setminus \partial(M) \implies$  Beh.

1.21.ii): Es gilt  $\partial(M) = \partial((X \setminus M))$  (nach Definition),  $(X \setminus M) \setminus \partial((X \setminus M))$  ist offen nach 1.21.i)  $\implies$

$$X \setminus ((X \setminus M) \setminus \partial((X \setminus M))) = (X \setminus (X \setminus M)) \cup \partial((X \setminus M)) = M \cup \partial(M)$$

Manipulation mit Mengen

ist offen.



1.21.iii):

$$\begin{aligned}\partial(M) &= (M \cup \partial(M)) \setminus (M \setminus \partial(M)) \\ \implies X \setminus \partial(M) &= X \setminus \left( \underbrace{M \cup \partial(M)}_{\substack{\text{(abgeschl. nach 1.21.ii)} \\ \text{offen}}} \right) \cup \left( \underbrace{M \setminus \partial(M)}_{\substack{\text{offen nach 1.21.i}}} \right).\end{aligned}\quad \square$$

**Notation.** Sei  $M \subset X$ .

$M^\circ := M \setminus \partial(M)$  heißt das *Innere* von  $M$ .

$\overline{M} := M \cup \partial(M)$  heißt der *Abschluss* von  $M$ .

Nach 1.19 können wir  $\overline{M}$  konstruieren, indem wir zu  $M$  noch alle Grenzwerte von Folgen  $(x_n)_n$ ,  $x_n \in M$ , die in  $X$  konvergieren, hinzunehmen.

**Beispiel.**  $M = [a, b)$ ,  $\overline{M} = [a, b]$ .

**Bemerkung (als Hausaufgabe).**

$$M \subset X \text{ offen} \iff M \cap \partial(M) = \emptyset.$$

$$M \subset X \text{ abgeschlossen} \iff \partial(M) \subset M.$$

## 1.II. Vollständigkeit

**Definition 1.22.** Sei  $(X, d)$  ein metrischer Raum. Eine Folge  $(y_n)_n \subset X$  heißt *Cauchy-Folge*, falls gilt

$$\forall \varepsilon > 0 \ \exists N \in \mathbb{N} \text{ s. d. } d(y_n, y_m) < \varepsilon \ \forall n, m \geq N.$$

**Lemma 1.23.** Sei  $(X, d)$  ein metrischer Raum. Eine konvergente Folge in  $X$  ist eine *Cauchy-Folge*.

*Beweis.* Sei  $(y_n)_n$  konvergente Folge mit Grenzwert  $y$  (eindeutig wegen 1.14.v) und 1.6). Sei  $\varepsilon > 0$ .

Dann gibt es  $N \in \mathbb{N}$  s. d.  $d(y_m, y) < \varepsilon \ \forall m \geq N$ .

$$\implies d(y_n, y_m) \underset{\triangle}{\leq} d(y_n, y) + d(y, y_m) < \varepsilon \ \forall n, m \leq N.\quad \square$$

**Bemerkung.** Nicht jede Cauchy-Folge konvergiert:

**Beispiel**  $((\mathbb{Q}, |\cdot|))$ .  $y_{n+1} = \frac{1}{2}y_n + \frac{1}{y_n}$ ,  $y_0 = 1$ .

**Check.** Es gilt für  $n \geq 1$

$$\left[ \frac{1}{y_{n+1}}, y_{n+1} \right] \subset \left[ \frac{1}{y_n}, y_n \right] \quad (*)$$

und für  $l_n := y_n - \frac{1}{y_n}$

$$\begin{aligned} l_{n+1} &\leq \frac{1}{4y_{n+1}}l_n^2 \leq \frac{1}{4}l_n^2 \\ \implies d(y_n, y_m) &= |y_n - y_m| \leq |y_n - \frac{1}{y_n}| = l_n \xrightarrow{\text{O.B.d.A. } m \geq n} 0. \\ &\implies y_m \in \left[ \frac{1}{y_n}, y_n \right] \text{ wg. } (*) \end{aligned} \quad (**)$$

$\mathbb{Q} \subset \mathbb{R}$  und somit  $(y_n)_n \subset \mathbb{R}$ . In  $\mathbb{R}$  konvergiert jede Cauchy-Folge. Nennen wir den Grenzwert  $a \in \mathbb{R}$ . Es gilt dann

$$\underbrace{y_{n+1}}_{\xrightarrow{a}} = \underbrace{\frac{1}{2}y_n}_{\frac{1}{2}a} + \underbrace{\frac{1}{y_n}}_{\frac{1}{a}},$$

also  $a^2 = 2$ . Aber  $\sqrt{2} \notin \mathbb{Q}$ .

**Definition 1.24.** Ein metrischer Raum, in dem jede Cauchy-Folge konvergiert heißt *vollständig*.

**Beispiele 1.25.** i)  $\mathbb{R}, |\cdot|$  ist vollständig (DIFF I).

ii)  $(C([a, b], \mathbb{R}), d_{L^1})$  mit  $d_{L^1}(f, g) = \int_a^b |f(t) - g(t)| dt$  (vgl. HA Blatt 1, A1) ist *nicht* vollständig.

iii)  $(C([a, b], \mathbb{R}), d_{\sup})$ , mit

$$d_{\sup} = \|f - g\|_{\infty} = \sup_{t \in [a, b]} |f(t) - g(t)|,$$

ist vollständig. Den Beweis führen wir später allgemeiner.

Zunächst einige

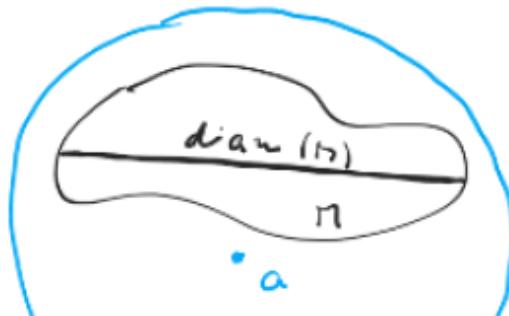
### 1.III. Betrachtungen in vollständigen metrischen Räumen

**Definition 1.26.** Sei  $(X, d)$  metrischer Raum,  $M \subset X$ ,

$$\text{diam}(M) := \sup_{x, y \in M} d(x, y) \text{ „Durchmesser“ (englisch „diameter“).}$$

$M$  heißt *beschränkt*, falls  $\text{diam}(M) < \infty$ .

**Bemerkung.**  $M$  ist beschränkt  $\iff \exists R \geq 0$  und  $a \in X$  s. d.  $M \subset B_R(a)$



**Beispiel.**  $\text{diam}([a, b]) = b - a$

**Satz 1.27 (Schachtelungsprinzip).** Sei  $(X, d)$  ein *vollständiger* metrischer Raum und sei  $A_0 \subset A_1 \subset A_2 \subset \dots$  Eine Familie nicht-leerer abgeschlossener Teilmengen von  $X$  mit

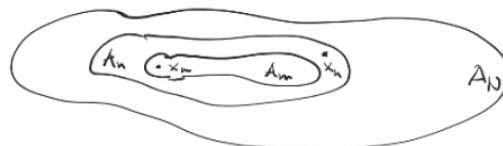
$$\text{diam}(A_k) \rightarrow 0 \text{ (in } \mathbb{R} \text{) für } k \rightarrow \infty.$$

Dann gibt es genau einen Punkt  $a \in X$  der in *allen*  $A_k$  liegt.

*Beweis.* Eindeutigkeit: Angenommen  $\exists x \neq y$  mit  $x \in A_k \quad \forall k$  und  $y \in A_k \quad \forall k$ . Dann kann  $\text{diam}(A_k)$  keine Nullfolge sein, da  $d(x, y) \neq 0$ .

Existenz: Wähle  $x_n \in A_n$ . Dann ist  $(x_n)_n$  eine Cauchy-Folge, denn

$$d(x_n, x_m) \leq \text{diam } A_N \text{ für } n, m \geq N$$



$$d(x_n, x_m) \leq \text{diam } A_N \text{ für}$$

$$\xrightarrow{\substack{\uparrow \\ \text{Vollständigkeit}}} x_n \rightarrow x \text{ in } X,$$

Da  $x_n \in A_k \quad \forall n \geq k$ , folgt mit 1.19:  $x \in A_k \quad \forall k$ .  $\square$

Ein sehr wichtiger Satz, der viele Anwendungen hat ist der folgende:

**Satz 1.28 (Banach'scher Fixpunktsatz).** Sei  $(X, d_X)$  ein *vollständiger* metrischer Raum. Sei  $M \subset X$  eine *abgeschlossene* Teilmenge und  $\Phi: M \rightarrow X$  eine Abbildung mit  $\Phi(M \subset M)$  und es gebe  $0 \leq L < 1$  s. d.

$$d_X(\Phi(X), \Phi(Y)) \leq L d_X(x, y) \quad \forall x, y \in M \quad (\text{„}\Phi\text{ ist Kontraktion“}).$$

Dann gibt es genau ein  $t_*$  s. d.  $\Phi(t_*) = t_*$ . Ein solches  $t_*$  heißt *Fixpunkt* von  $\Phi$ .



$$X = \mathbb{R}, M = [0, 1], \log(2 - x^2), (\text{WolframAlpha})$$

### Beispiel.

*Beweis.* Eindeutigkeit: Seien  $\Phi(t_*) = t_*$ ,  $\Phi(\tilde{t}_*) = \tilde{t}_*$ . Dann gilt

$$\begin{aligned} d(t_*, \tilde{t}_*) &= d(\Phi(t_*), \Phi(\tilde{t}_*)) \\ &\leq d(t_*, \tilde{t}_*) \end{aligned}$$

Da  $L > 1$  ist, folgt  $d(t_*, \tilde{t}_*) = 0$ , also  $t_* = \tilde{t}_*$ .

Existenz: Wir betrachten die Folge  $x_0 \in M$  beliebig,  $x_n := \Phi(x_{n-1})$  für  $n \geq 1$ .

**Behauptung.**  $(x_n)_n$  konvergiert in  $M$  und zwar gegen de Fixpunkt.

*Beweis.* •  $(x_n)_n$  ist Cauchy-Folge:

$$d(x_{n+1}, x_n) \leq L d(x_n, x_{n-1}) \quad \forall n \geq 1.$$

$$\frac{\parallel \Phi(x_n) - \Phi(x_{n-1})}{\parallel \Phi(x_n)}$$

Iteration liefert

$$d(x_{n+1}, x_n) \leq L^2 d(x_{n-1}, x_{n-2}) \leq \dots \leq L^n d(x_1, x_0).$$

Zudem gilt

$$\begin{aligned} d(x_{n+k}, x_n) &\stackrel{\triangle}{\leq} d(x_{n+k}, x_{n+k-1}) \\ &\quad + d(x_{n+k-1}, x_{n+k-2}) \\ &\quad \vdots \\ &\quad + d(x_{n+1}, x_n) \\ &\leq \underbrace{(L^{n+k-1} + L^{n+k-2} + \dots + L^n)}_{= L^n \sum_{r=0}^{k-1} L^r \leq L^n \sum_{r=0}^{\infty} L^r = \frac{L^n}{1-L}} d(x_1, x_0) \\ &\quad \text{geom. Reihe } (L < 1) \end{aligned}$$

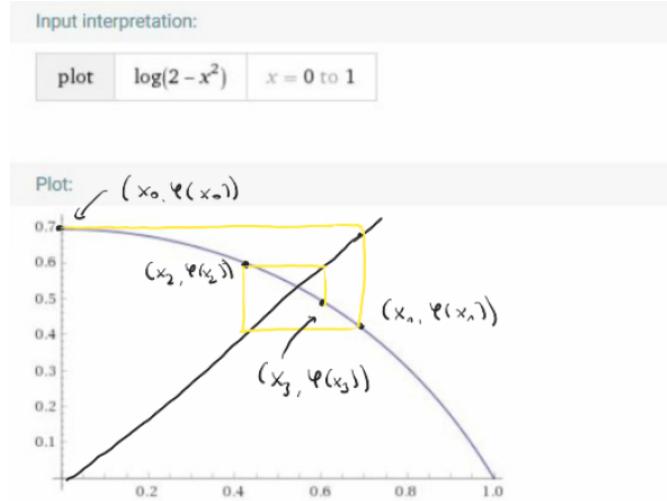
$\implies$  (wegen  $L < 1$ ) Beh.

- Da  $X$  vollständig ist, konvergiert  $(x_n)_n$ . Setze  $t_* = \lim_{n \rightarrow \infty} x_n$ .
- Da  $M$  abgeschlossen ist, ist  $t_* \in M$  nach 1.19.
- $t_*$  ist der gesuchte Fixpunkt:

$$t_* = \lim_{n \rightarrow \infty} x_n = \lim_{n \rightarrow \infty} \Phi(x_{n-1}) = \Phi(t_*)$$

$\square$

Kontraktionen sind stetig und 1.18



$$x_0 = 0, x_1 \ln(2 - (\ln 2)^2) \approx 0.42, x_3 \approx 0.60, x_4 \approx 0.49$$

**Bemerkung.** Kontraktionen sind stetig: Zu  $\varepsilon > 0$  wähle  $\delta = \varepsilon/L$ .

**Bemerkung.** Die Konvergenz ist recht schnell:

$$d(x_n, t_*) \leq \frac{L^n}{1-L} d(x_1, x_0) \quad (L < 1).$$

Alle Voraussetzungen sind notwendig, gilt eine nicht, so gibt es nicht unbedingt einen Fixpunkt (oder keinen eindeutigen).



**Vorlesung 3**

Mo 27.04. 10:15

**Lemma 1.29 (Cauchy-Kriterium für gleichmäßige Konvergenz).**

- a) Sei  $(X, d_x)$  metrischer Raum, sei  $(Y, d_Y)$  ein *vollständiger* metrischer Raum. Sei  $f_n: X \rightarrow Y$  Folge von Funktionen. Dann konvergiert  $f_n$  gegen  $f$  bezüglich

$$d_{\sup}(h, g) := \sup_{t \in X} \underbrace{d_Y(f(t), g(t))}_{\in \mathbb{R}} \quad h, g: X \rightarrow Y$$

$$\iff \forall \varepsilon > 0 \exists N = N(\varepsilon) \in \mathbb{N} \text{ s.d.}$$

$$d_Y(f_n(t), f_m(t)) < \varepsilon \quad \forall t \in X \quad n, m \geq N(\varepsilon). \quad (*)$$

**Notation.** Man spricht von *gleichmäßiger Konvergenz*.**Beachte:**

Der wesentliche Punkt in (\*) ist, dass  $N$  unabhängig von  $t$  gewählt werden kann.

*Beweis.* Wir stellen zunächst fest, dass  $d_{\sup}$  auf

$$\mathcal{F} := \{f: X \rightarrow Y \mid \text{Für je zwei Funktionen gilt: } d_{\sup}(f_1, f_2) < \infty\}$$

eine Metrik definiert (auch wenn  $Y$  nicht vollständig ist).

1.11.a)

$$\begin{aligned} d_{\sup}(f, g) &= 0 \\ \iff d_Y(f(t), g(t)) &= 0 \quad \forall t \in X \\ \iff f(t) &= g(t) \quad \forall t \in X \\ &\quad \uparrow \\ &\quad d_Y \text{ ist Metrik} \end{aligned}$$

1.11.b)

$$d_{\sup}(f, g) = \sup d_Y(f(t), g(t)) = \sup d_Y(g(t), f(t)) = d_{\sup}(g, f)$$

1.11.c)

$$\begin{aligned} d_{\sup}(f, g) &= \sup \underbrace{d_Y(f(t), g(t))}_{\leq d_Y(f(t), h(t)) + d_Y(h(t), g(t))} \\ &\leq \sup d_Y(f(t), h(t)) + \sup d_Y(h(t), g(t)) \\ &= d_{\sup}(f, h) + d_{\sup}(h, g) \end{aligned}$$

**Zum Beweis der Behauptung:**

„ $\implies$ “

$$\sup_t d_Y(f_n(t), f(t)) < \varepsilon \quad \forall n \geq N(\varepsilon)$$

impliziert

$$d_Y(f_n(t), f(t)) < \varepsilon \quad \forall t \in X \quad \forall n \geq N(\varepsilon),$$

somit für alle  $t \in X$

$$\begin{aligned} d_Y(f_n(t), f_m(t)) &\stackrel{\Delta(d_Y)}{\leq} d_Y(f_n(t), f(t)) + d_Y(f(t), f_m(t)) \\ &< 2\varepsilon \quad \forall n, m \geq N(\varepsilon) \end{aligned}$$

„ $\Leftarrow$ “ Gelte (\*). Dann ist für jedes  $t \in X$ , dass  $(f_n(t))_n$  ist Cauchy-Folge in  $Y$ .

Vollständigkeit von  $Y \implies (f_n(t))_n$  konvergiert. Setze  $f(t) := \lim_{n \rightarrow \infty} f_n(t)$ .

Wir zeigen  $f_n$  konvergiert bezüglich  $d_{\sup}$  gegen  $f$ . Sei also  $\varepsilon > 0$ . Wähle in (\*)  $m \geq N(\varepsilon)$  fest. Dann gilt für alle  $t$ :

$$\begin{aligned} \varepsilon &\geq \lim_{n \rightarrow \infty} d_Y(f_n(t), f_m(t)) \\ &= d_Y(f(t), f_m(t)). \end{aligned}$$

$\uparrow$   
 $d_Y$  ist stetig

Das gilt für alle  $m \geq N(\varepsilon)$ ,  $\forall t$ , also auch für das Supremum  $\implies$  Beh.  $\square$

b) Seien  $X, Y$  metrische Räume,  $(f_n)_n$  eine Folge *stetiger* Funktionen  $f_n: X \rightarrow Y$ , die gleichmäßig konvergiere. Dann ist die Grenzfunktion  $f: X \rightarrow Y$ .

*Beweis.* Sei  $a \in X$ . Sei  $\varepsilon > 0$ . Gleichmäßige Konvergenz  $\implies \exists N = N(\varepsilon) \in \mathbb{N}$  s. d.

$$d_Y(f(t), f_n(t)) < \varepsilon \quad \forall t \in X \quad \forall n \geq N$$

$f_n$  stetig in  $a \implies \exists \delta > 0$  s. d.

$$\begin{aligned} d_Y(f_N(t), f_N(a)) &< \varepsilon \quad \forall t \text{ mit } d_X(t, a) < \delta \\ \implies d_Y(f(t), f(a)) &\stackrel{\Delta}{\leq} d_Y(f(t), f_N(t)) + d_Y(f_N(t), f_N(a)) + d_Y(f_N(a), f(a)) \quad \square \\ &< 3\varepsilon \quad \forall t \text{ mit } d_X(t, a) < \delta. \end{aligned}$$

**Folgerung.**

$$(C([a, b], \mathbb{R}), d_{\sup}),$$

Stellt man diese Bedingung, ist automatisch garantiert, dass  $d_{\sup}(f_1, f_2) = \sup_{t \in [a, b]} |f_1(t) - f_2(t)| < \infty$

ist vollständig.

*Beweis.* Sei  $(f_n)_n$  Cauchy-Folge in  $C(D, \mathbb{R})$  bezüglich  $d_{\sup}$ , d. h. zu  $\varepsilon > 0 \exists N = N(\varepsilon)$  s. d.

$$\begin{aligned} d_{\sup}(f_n, f_m) &< \varepsilon \quad \forall n, m \geq N(\varepsilon) \\ \implies d_Y(f_n(t), f_m(t)) &= |f_n(t) - f_m(t)| < \varepsilon \quad \forall n, m \geq N(\varepsilon) \quad \forall t \in D. \end{aligned}$$

$\mathbb{R}$  ist vollständig

$\stackrel{1.29}{\implies} (f_n)_n$  konvergiert bezüglich  $d_{\sup}$  gegen seinen punktweisen Grenzwert

$$f(t) := \lim_{n \rightarrow \infty} f_n(t) \quad (\text{Konvergenz in } \mathbb{R})$$

$\stackrel{1.29.b)}{\implies} t \mapsto f(t)$  ist stetig. □

## 1.IV. Stetige Abbildungen auf metrischen Räumen

**Lemma 1.30.** Seien  $X, Y, Z$  metrische Räume,  $f: X \rightarrow Y$ ,  $g: Y \rightarrow Z$ ,  $f(X) \subset Y$ . Ist  $f$  stetig in  $a \in X$  und  $g$  stetig in  $b = f(a) \in Y$ , so ist  $g \circ f: X \rightarrow Z$  stetig in  $a$ .

*Beweis.* (Über Folgenstetigkeit, Lemma 1.18) Sei  $x_n \rightarrow a \implies \lim f(x_n) = (a) = b$  und  $\lim g(f(x_n)) = g(b) = g(f(a)) \implies \lim g \circ f(x_n) = g \circ f(a)$ .

**Definition 1.31.** Auf dem  $\mathbb{R}^n$  ist durch

$$d_{\max}(x, y) := \max_{i \in \{1, \dots, n\}} |x_i - y_i|.$$

eine Metrik definiert.

**Bemerkungen.** i)  $d_{\max}(x, y) = d_{\sup}(x, y)$ , fasst man  $x$  und  $y$  als Abbildungen

$$x: \{1, \dots, n\} \rightarrow \mathbb{R}$$

auf,  $x(i) = x_i$ .

- ii) Eine Folge  $(x_m)_m \subset \mathbb{R}^n$ ,  $x_m = (x_m^1, \dots, x_m^n)$  konvergiert bezüglich  $d_{\max}$   $\iff$  Alle Komponentenfolgen  $(x_m^i)_m$  ( $1 \leq i \leq n$ ) konvergieren in  $\mathbb{R}$ .

*Beweis.* „ $\implies$ “ Zu  $\varepsilon > 0 \exists N$  s. d.  $\max|x_m^i - a_m^i| < \varepsilon \quad \forall m \geq N$ .

„ $\impliedby$ “ Zu  $\varepsilon > 0 \exists N_i$  s. d.  $|x_m^i - a_m^i| < \varepsilon \quad \forall m \geq N_i$

$$\implies \max_i |x_m^i - a_m^i| < \varepsilon \quad \forall m \geq N = \max\{\{N_1, \dots, N_n\}\}.$$

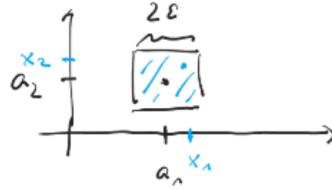
□

- iii) Es folgt:  $(\mathbb{R}^n, d_{\max})$  ist vollständig.

- iv)  $B_\varepsilon(a)$  bezüglich dieser Metrik:

$$\left\{ x \in \mathbb{R}^n \mid \max_{i \in \{1, \dots, n\}} |x_i - a_i| < \varepsilon \right\},$$

Würfel mit Seitenlängen  $2\varepsilon$  um  $a$ .



**Lemma 1.32.** Sei  $(X, d)$  metrischer Raum. Sei  $\mathbb{R}^n$  mit  $d_{\max}$  versehen. Eine Abbildung  $f: X \rightarrow \mathbb{R}^n$ ,  $f = (f_1, \dots, f_n)^T$ ,

$$f(y) = (f_1(y), \dots, f_n(y))^T \in \mathbb{R}^n, y \in X.$$

$f_i: X \rightarrow \mathbb{R}$ ,  $i \in \{1, \dots, n\}$ , „Komponenten-Funktionen“, ist genau dann stetig in  $a \in X$ , falls alle  $f_i$  stetig in  $a$  sind.

*Beweis.* Mit Folgenstetigkeit direkt aus Bemerkung 1.31.ii). Hier nochmals mit  $\varepsilon$ - $\delta$ -Kriterium.

**Notation.**  $\underline{n} = \{1, \dots, n\}$ .

„ $\implies$ “ Sei also  $f: X \rightarrow \mathbb{R}^n$  stetig in  $a$ . Sei  $\varepsilon > 0$ . Dann  $\exists \delta > 0$  s. d.

$$\begin{aligned} & \max_{i \in \underline{n}} |f_i(y) - f_i(a)| < \varepsilon \quad \forall y \in B_\delta(a) \\ & \qquad \qquad \qquad \uparrow \text{bezüglich } d \\ \implies & |f_i(y) - f_i(a)| < \varepsilon \quad \forall y \in B_\delta(a) \quad \forall i \in \underline{n} \\ \implies & f_i \text{ sind stetig in } a. \end{aligned}$$

„ $\Leftarrow$ “ Seien also die  $f_i: X \rightarrow \mathbb{R}$ ,  $i \in \underline{n}$ , stetig in  $a$ . Sei  $\varepsilon > 0$ . Dann  $\exists \delta_i > 0$  s. d.

$$|f_i(y) - f_i(a)| < \varepsilon \quad \forall y \in B_{\delta_i}(a) \subset X.$$

Wähle  $d := \min\{\{\delta_1, \dots, \delta_n\}\}$ . Dann ist

$$\max_{i \in \underline{n}} |f_i(y) - f_i(a)| < \varepsilon \quad \forall y \in B_d(a).$$

□

**Lemma 1.33.** Folgende Abbildungen sind stetig:

$$\begin{aligned} \text{add: } \mathbb{R}^2 &\rightarrow \mathbb{R}, \quad \text{add}(x, y) = x + y \\ \text{mult: } \mathbb{R}^2 &\rightarrow \mathbb{R}, \quad \text{mult}(x, y) = x \cdot y \\ \text{quot: } \mathbb{R} \times \underset{\substack{\parallel \\ \mathbb{R} \setminus \{0\}}}{} \mathbb{R}^* &\rightarrow \mathbb{R}, \quad \text{quot}(x, y) = x/y. \end{aligned}$$

Hierbei sei  $\mathbb{R}^2$ ,  $\mathbb{R} \times \mathbb{R}^*$ , mit  $d_{\max}$  versehen.

*Beweis.* Sei  $((x_m, y_m))_m \subset \mathbb{R}^2$  mit  $(x_m, y_m) \rightarrow (x, y)$  (bezüglich  $d_{\max}$ )

$$\begin{aligned} &\xrightarrow[\text{Bem 1.31.ii)}]{} x_m \rightarrow x \text{ und } x_m \rightarrow y \text{ in } \mathbb{R} \\ &\implies \lim(x_m + y_m) = x + y \\ &\lim(x_m \cdot y_m) = x \cdot y \\ &\lim(x_m / y_m) = x/y \quad \text{falls } y_m \neq 0, y \neq 0. \end{aligned}$$

□

**Folgerung.** Sei  $(X, d)$  metrischer Raum. Seien  $f, g: X \rightarrow \mathbb{R}$  stetig. Dann sind auch

$$\begin{aligned} f + g: X &\rightarrow \mathbb{R}, \quad (f + g)(x) = f(x) + g(x) \text{ und} \\ g \cdot g: X &\rightarrow \mathbb{R}, \quad (g \cdot g)(x) = f(x) \cdot g(x) \end{aligned}$$

stetig. Gilt  $g(x) \neq 0 \quad \forall x \in X$ , so ist auch

$$f/g: X \rightarrow \mathbb{R}, \quad (f/g)(x) = f(x)/g(x)$$

stetig.

*Beweis.*

$$1.32 \implies \begin{pmatrix} f \\ g \end{pmatrix}: X \rightarrow \mathbb{R}^2, \quad \begin{pmatrix} f \\ g \end{pmatrix}(x) = \begin{pmatrix} f(x) \\ g(x) \end{pmatrix}$$

ist stetig.

Es ist

$$\begin{aligned} f + g &= \text{add} \circ \begin{pmatrix} f \\ g \end{pmatrix} \\ f + g &= \text{mult} \circ \begin{pmatrix} f \\ g \end{pmatrix} \\ f/g &= \text{quot} \circ \begin{pmatrix} f \\ g \end{pmatrix} \end{aligned}$$

Mit 1.33 und 1.30 folgt die Behauptung.  $\square$

**Beispiel.** Polynomische Funktionen  $\mathbb{R}^n \rightarrow \mathbb{R}$

$$x \mapsto \sum_{0 \leq k_i \leq r} c_{k_1 \dots k_n} \underbrace{x_1^{k_1} \dots x_n^{k_n}}_{\in \mathbb{R}}$$

sind stetig.

**Bemerkung 1.34.** Wir werden später sehen, dass die Aussage in 1.33 auch gilt, wenn man den  $\mathbb{R}^2$  z. B. mit dem Euklidischen Abstand versieht.

## 1.V. Kompaktheit

**Definition 1.35.** Sei  $(X, d)$  metrischer Raum,  $M \subset X$ . Eine *offene Überdeckung von  $M$*  ist eine Familie  $(U_i)_{i \in I}$  von offenen Teilmengen  $U_i \subset X$  mit  $M \subset \bigcup_{i \in I} U_i$  ( $I$  eine beliebige Indexmenge).

**Definition 1.36.**  $M \subset X$  heißt *kompakt*, falls es zu *jeder* offenen Überdeckung von  $\bigcup_{i \in I} U_i$  von  $M$  *endlich* viele Indizes  $i_1, \dots, i_N$  gibt s. d.

$$M \subset U_{i_1} \cup \dots \cup U_{i_N}.$$

**Achtung.** Ein nicht-kompakter Raum kann eine endliche Überdeckung  $U_1 \cup \dots \cup U_N$  besitzen. Die Aussage der Definition ist, dass man aus *jeder* offenen Überdeckung endlich viele offene Mengen wählen kann, die  $M$  noch ganz überdecken!

**Beispiele 1.37.** i)  $[a, b]$  ist kompakt (Beweis später).

ii)  $(a, b)$  ist nicht kompakt (obwohl etwa  $(a, b)$  eine endliche offene Überdeckung ist!)

*Beweis.*

$$U_j = \left( a + \frac{1}{j}, b \right), \quad j \geq 1$$

$$\bigcup_j U_j = (a, b)$$

aber es gibt *kein*  $N$  s.d.  $\bigcup_{j=1}^N U_j \supset (a, b)$ , denn z.B.  $a + \frac{1}{N+1} \notin \bigcup_{j=1}^N U_j$ .  $\square$

- iii) Sei  $(x_n)_n \subset X$  gegen  $a$  konvergente Folge. Dann ist  $M = \{x_n \mid n \in \mathbb{N}\} \cup \{a\}$  kompakt.

*Beweis.* Sei  $(U_j)_j$  eine offene Überdeckung von  $M$

$$a \in M \implies \exists j_0 \text{ s.d. } a \in U_{j_0}$$

$U_{j_0}$  ist offen, also eine Umgebung von  $a$ .

$$\implies \exists N \text{ s.d. } x_n \in U_{j_0} \quad \forall n \geq N.$$
 $\square$

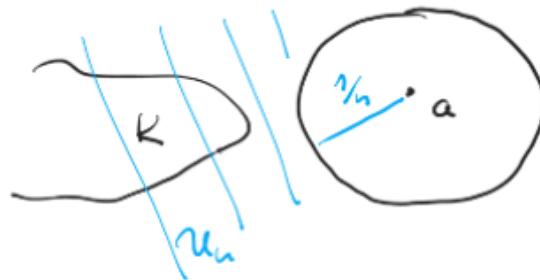
- iv) Sei  $(X_i, d_{\text{discrete}})$ . Dann sind genau die endlichen Mengen kompakt.

*Beweis.* Betrachte  $\bigcup_{x \in M} \{x\}$ .  $\square$

**Satz 1.38.** Sei  $(X, d)$  metrischer Raum,  $K \subset X$  kompakt. Dann ist  $K$  abgeschlossen und beschränkt.

*Beweis.* Abgeschlossen: Sei  $a \in X \setminus K$ . Setze zu  $n \geq 1$

$$U_n := \left\{ y \in X \mid d(y, a) > \frac{1}{n} \right\}$$



$U_n$  ist offen (denn  $X \setminus U_n = \overline{B_{1/n}(a)}$ ) und  $\bigcup_{n=1}^{\infty} U_n = X \setminus \{a\} \supset K$ .  $K$  kompakt  $\implies \exists U_{n_1}, \dots, U_{n_L}$  s.d.  $K \subset U_{n_1} \cup \dots \cup U_{n_L}$ . Setze  $N := \max\{\{n_1, \dots, n_L\}\}$ . Dann ist  $B_{\frac{1}{N}}(a) \subset X \setminus K \implies X \setminus K$  ist offen  $\implies$  Beh.

Beschränktheit: Sei  $a \in X$ . Dann ist  $X = \bigcup_{n=1}^{\infty} B_n(a)$  und somit  $(B_n(a))_n$  eine offene Überdeckung von  $K$ .

$$\begin{aligned} &\implies \exists n_1, \dots, n_k \text{ s. d. } K \subset B_{n_1}(a) \cup \dots \cup B_{n_k}(a) \\ &\implies K \subset B_N(a) \text{ für } N = \max\{\{n_1, \dots, n_k\}\} \\ &\implies \text{diam}(K) \leq 2N. \end{aligned}$$

□

**Folgerung.** Konvergente Folgen sind beschränkt.

**Bemerkung.** Die Umkehrung von 1.38 gilt im Allgemeinen nicht!

$(X, d_{\text{discrete}})$ ,  $X$  habe unendlich viele Elemente. Jede Teilmenge ist abgeschlossen (da jede offen ist) und beschränkt (durch 1), aber nur die *endlichen* sind kompakt.

**Lemma 1.39.** Ist  $K \subset X$  kompakt und  $A \subset K$  ist abgeschlossen, so ist  $A$  kompakt.

*Beweis.* Sei  $(U_j)_j$  offene Überdeckung von  $A$ . Es ist

$$\begin{aligned} &(\underbrace{X \setminus A}_{\text{offen (VOR)}}) \cup \bigcup U_j = X \supset K \\ &\implies \exists j_1, \dots, j_L \text{ s. d. } K \subset (X \setminus A) \cup U_{j_1} \cup \dots \cup U_{j_L} \\ &\implies A \subset U_{j_1} \cup \dots \cup U_{j_L}. \end{aligned}$$

□

**Satz 1.40.** Seien  $X, Y$  metrische Räume und  $f: X \rightarrow Y$  stetig. Ist  $K \subset X$  kompakt, so ist auch  $f(K) \subset Y$  kompakt.

*Beweis.* Sei  $(U_j)_j$  offene Überdeckung von  $f(K)$ .  $f$  stetig  $\xrightarrow{1.8}$  Die Urbilder  $V_j := f^{-1}(U_j)$  sind offen.

Und nach Definition ist  $K \subset \bigcup_j V_j$ .

$$\begin{aligned} &\xrightarrow{\text{VOR}} \exists j_1, \dots, j_N \text{ s. d. } K \subset V_{j_1} \cup \dots \cup V_{j_N} \\ &\implies f(K) \subset U_{j_1} \cup \dots \cup U_{j_N}. \end{aligned}$$

□

**Satz 1.41.** Sei  $\mathfrak{X}$  kompakter metrischer Raum,  $f: \mathfrak{X} \rightarrow \mathbb{R}$  stetig. Dann ist  $f$  beschränkt und nimmt ihr Maximum und Minimum an, d. h.  $\exists a, b \in \mathfrak{X}$

$$f(a) = \sup\{f(x) \mid x \in \mathfrak{X}\}, \quad f(b) = \inf\{f(x) \mid x \in \mathfrak{X}\}.$$

*Beweis.* 1.40  $\implies f(\mathfrak{X})$  ist kompakt. Mit 1.38 folgt:  $f(\mathfrak{X})$  ist beschränkt (somit ist  $f$  beschränkt) und abgeschlossen.

Also sind  $\sup(f(\mathfrak{X}))$  und  $\inf(f(\mathfrak{X}))$  endlich. Zudem gibt es

$$(y_k)_k \subset f(\mathfrak{X}), \quad y_k \rightarrow \sup(f(\mathfrak{X})) \\ (z_k)_k \subset f(\mathfrak{X}), \quad z_k \rightarrow \inf(f(\mathfrak{X})),$$

somit (Abgeschlossenheit!)

$$\sup(f(\mathfrak{X})) \in f(\mathfrak{X}) \\ \inf(f(\mathfrak{X})) \in f(\mathfrak{X})$$

$\implies$  Beh. □

**Beispiel.** Sei  $(\mathfrak{X}, d)$  metrischer Raum.  $M \subset \mathfrak{X}$ . Sei  $x \in \mathfrak{X}$ . Der *Abstand* von  $x$  zu  $M$  ist definiert als

$$d(x, M) := \inf \{ d(x, y) \mid y \in M \}.$$

**Behauptung.**  $x \mapsto d(x, M)$  ist stetig auf  $\mathfrak{X}$ .

*Beweis.* Sei  $\varepsilon > 0$ . Dann ist

$$|d(x, M) - d(\tilde{x}, M)| \leq d(x, \tilde{x}) < \varepsilon \quad \text{falls } d(x, \tilde{x}) < \varepsilon,$$

denn

$$d(x, M) \stackrel{\Delta}{\leq} d(x, \tilde{x}) + d(\tilde{x}, M) \quad \forall x, \tilde{x} \in \mathfrak{X}. \quad \square$$

Definiere zu  $K \subset \mathfrak{X}$

$$d(K, M) := \inf \{ d(x, M) \mid x \in K \}.$$

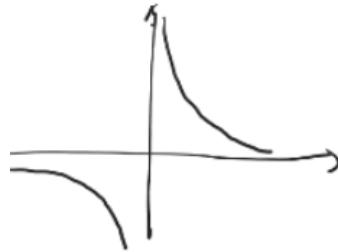
**Behauptung.** Ist  $M$  abgeschlossen,  $K$  kompakt und ist  $M \cap K = \emptyset$ , so gilt  $d(M, K) > 0$ .

*Beweis.*  $x \mapsto d(x, M)$  ist stetig auf  $\mathfrak{X}$ , somit erst recht auf  $K$ .  $K$  ist kompakt  $\stackrel{1.41}{\implies} \exists a \in K$  s. d.  $d(a, M) = d(K, M)$ .  $M$  abgeschlossen  $\implies \exists \varepsilon > 0$  s. d.  $B_\varepsilon(a) \subset X \setminus M \implies d(a, M) \geq \varepsilon$ .

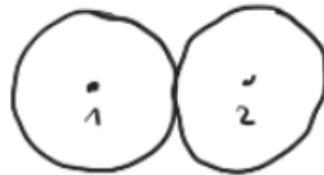
**Achtung.** i) Betrachte

$$M = \{ (x, y) \mid xy = 0 \} \subset N = \{ (x, y) \mid xy = 1 \} \subset \mathbb{R}^2$$

$$d(M, N) = 0.$$



ii) Betrachte  $B_{1/2}(1), B_{1/2}(2) \subset \mathbb{R}^2$ ,  $d_{\text{Euklidisch}}$ . Distanz ist 0.



**Satz / Definition 1.42.** Seien  $\mathfrak{X}, Y$  metrische Räume,  $\mathfrak{X}$  kompakt. Dann ist jede stetig Abbildung  $f: \mathfrak{X} \rightarrow Y$  sogar *gleichmäßig stetig* d. h. im  $\varepsilon$ - $\delta$ -Kriterium kann  $\delta$  unabhängig von  $x$  gewählt werden:

$$\forall \varepsilon > 0 \quad \exists \delta > 0 \text{ s. d. } d_Y(f(x), f(\tilde{x})) < \varepsilon \quad \forall x, \tilde{x}, d_{\mathfrak{X}}(x, \tilde{x}) < \delta.$$

*Beweis.* Sei  $\varepsilon > 0$ . Dann gibt es zu  $a \in \mathfrak{X}$  ein  $\delta(a) > 0$  s. d.

$$d_Y(f(a), f(y)) < \varepsilon \quad \forall y \in B_{\delta(a)}(a).$$

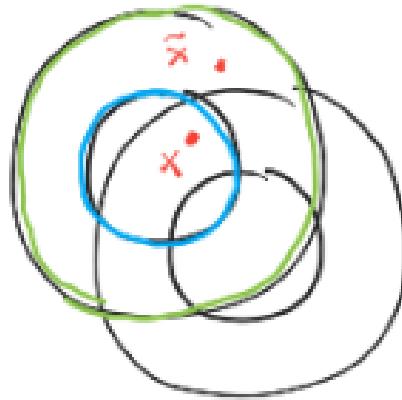
Es gilt  $\bigcup_{a \in X} B_{\frac{\delta(a)}{2}}(a) = \mathfrak{X}$ .

$\mathfrak{X}$  ist kompakt  $\implies \exists a_1, \dots, a_N$  s. d.  $X = \bigcup_{j=1}^N B_{\delta(a_j)/2}(a_j)$ . Setze

$$\delta := \frac{1}{2} \min\{\{\delta(a_1), \dots, \delta(a_N)\}\}.$$

Seien jetzt  $x, \tilde{x}$  beliebig aus  $\mathfrak{X}$  mit  $d_{\mathfrak{X}}(x, \tilde{x}) < \delta$ . Dann gibt es ein  $j \in \{1, \dots, N\}$  s.d.  $x \in B_{\delta(a_j)/2}$  und somit  $\tilde{x} \in B_{\delta(a_j)}(a_j)$

$$\begin{aligned} &\implies d_Y(f(x), f(a_j)) < \varepsilon \quad d_Y(f(\tilde{x}), f(a_j)) < \varepsilon \\ &\implies d_Y(f(x), f(\tilde{x})) < 2\varepsilon \quad \forall x, \tilde{x}, \quad d_{\mathfrak{X}}(x, \tilde{x}) < \delta. \end{aligned} \quad \square$$



**Satz 1.43 (Bolzano-Weierstraß).** Sei  $(X, d)$  metrischer Raum. Sei  $K \subset X$  kompakt. Dann besitzt jede Folge  $(x_n)_n$  in  $K$  eine Teilfolge  $(x_{n_k})_k$ , die gegen einen Punkt  $x \in K$  konvergiert.

*Beweis.* Angenommen,  $\nexists$  Teilfolge, die gegen einen Punkt von  $K$  konvergiert. Dann besitzt jedes  $x \in K$  eine offene Umgebung  $U_x$ , in der nur endlich viele Folgenglieder liegen (sonst könnte man eine gegen  $x$  konvergente Teilfolge konstruieren). Es gilt:  $\bigcup_{x \in K} U_x \supset K$

$$\implies \exists x_1, \dots, x_N \text{ s.d. } \bigcup_{j=1}^N U_{x_j} \subset K \quad \square$$

Aber dann liegen nur endlich viele  $x_k$  in  $K$ ,  $\ntriangleleft$  zur Definition.

**Vorlesung 4**

Do 30.04. 10:15

**1.VI. Äquivalenz von Metriken**

Wir haben gesehen, dass die Eigenschaften derselben Menge sehr verschieden sein können, je nachdem mit welcher Topologie man sie versieht.

**Beispiel.**  $\mathbb{R}$  mit der Standardtopologie  $|x - y|$ :

- $(a, b]$  ist nicht offen,  $[a, b]$  ist kompakt.

$\mathbb{R}$  mit der diskreten Metrik  $d_{\text{disk}}$

- Alle Teilmengen sind offen.
- Nur endliche Teilmengen sind kompakt.
- Konvergiert  $x_n \rightarrow a$  (bezüglich  $d_{\text{disk}}$ ), so muss gelten  $\exists N$  s. d.  $x_n = a \quad \forall n \geq N$  (denn  $\{a\}$  ist Umgebung von  $a$ , oder anders gesagt: damit  $d(x_n, a) < \varepsilon < 1$  wird, muss gelten  $x_n = a$ ).
- Alle Abbildungen  $f: (X, d_{\text{disc}}) \rightarrow (Y, d)$  sind stetig. (Beweis am einfachsten über Folgenstetigkeit).

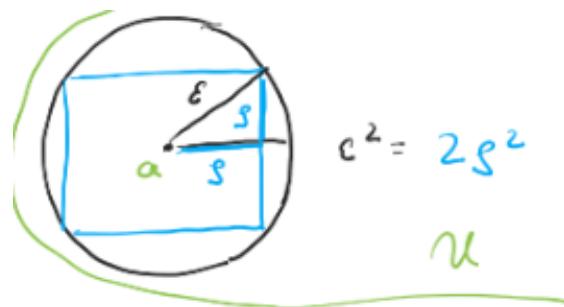
Andererseits gilt:

$U \subset \mathbb{R}^2$  ist offen in  $(\mathbb{R}^2, d_{\text{Eucl}}) \iff U$  ist offen in  $(\mathbb{R}^2, d_{\max})$ .

*Beweis.* „ $\implies$ “ Sei  $a \in U \stackrel{\text{VOR}}{\implies} \exists \varepsilon > 0$  s. d.

$$B_\varepsilon^{d_{\text{E}}} (a) := \left\{ x \in \mathbb{R}^2 \mid d_{\text{Eucl}(x, a)} = \sqrt{(x_1 - a_1)^2 + (x_2 - a_2)^2} < \varepsilon \right\} \subset U$$

Da  $B_\rho^{d_{\max}}(a) \subset B_\varepsilon^{d_{\text{E}}}(a)$  für  $\rho = \frac{\varepsilon}{\sqrt{2}}$ , ist  $U$  auch offen  $(\mathbb{R}^2, d_{\max})$ .



„ $\Leftarrow$ “ Sei  $a \in U \xrightarrow{VOR} \exists \varepsilon > 0$  s. d.

$$B_\varepsilon^{d_{\max}}(a) = \{x \mid d_{\max}(x, a) < \varepsilon\} \subset U.$$

□

Es gilt  $B_\rho^{d_{\text{E}}}(a) \subset B_\varepsilon^{d_{\max}}(a)$  für  $\rho = \varepsilon$ , also ist  $U$  offen in  $(\mathbb{R}^2, d_{\text{Eucl}})$ .



**Definition 1.44.** Sei  $X$  eine Menge, seien  $d$  und  $\tilde{d}$  Metriken auf  $X$ . Dann nennt man  $d$  *stärker* (feiner) als  $\tilde{d}$ , falls jede bezüglich  $\tilde{d}$  offene Menge auch offen bezüglich  $d$  ist, und *schwächer* (größer), falls  $\tilde{d}$  stärker ist als  $d$ . Ist  $d$  sowohl stärker als auch schwächer als  $\tilde{d}$ , so nennt man  $d$  und  $\tilde{d}$  *äquivalent*.

**Beispiel.**  $d_{\max}$  ist äquivalent zu  $d_{\text{Eucl}}$ .  $d_{\text{disk}}$  ist stärker als  $d_{\max}$  und nicht schwächer.

**Bemerkungen 1.45.** Sei  $d$  stärker als  $\tilde{d}$ . Dann gilt:

- i) Konvergiert eine Folge bezüglich der stärkeren Metrik, so auch bezüglich der schwächeren.

(denn: Konvergiere  $x_n \rightarrow a$  (bezüglich  $d$ ). Sei  $\varepsilon > 0$ . Betrachte  $B_\varepsilon^{\tilde{d}}(a) = U$  offen bezüglich  $d \implies U$  Umgebung von  $a$  (bezüglich  $d$ )  $\implies U$  Umgebung von  $a$  (bezüglich  $\tilde{d}$ )  $\implies \exists N$  s. d.  $x_n \in U \quad \forall n \geq N$ .)

- ii) Ist eine Funktion  $f: (X, \tilde{d}) \rightarrow (Y, d_Y)$  stetig, so auch  $f: (X, d) \rightarrow (Y, d_Y)$ .
- iii) Ist eine Funktion  $f: (Y, d_Y) \rightarrow (X, d)$  stetig, so auch  $f: (Y, d_Y) \rightarrow (X, \tilde{d})$ .

*Beweis.*  $f$  stetig  $\iff$  Urbilder offener Mengen sind offen.

1.  $U \subset Y \implies f^{-1}(U)$  offen bezüglich  $\tilde{d} \implies f^{-1}(U)$  offen bezüglich  $d$ .

2. Sei  $U \subset X$  offen bezüglich  $\tilde{d}$ , also auch offen bezüglich  $d \implies f^{-1}(U)$  offen in  $Y$ . □

**Bemerkung.** Sind  $d$  und  $\tilde{d}$  äquivalent, sind die selben Folgen konvergent, die selben Mengen offen, kompakt, die selben Funktionen stetig etc.

## Kapitel 2.

### Normierte Vektorräume

**Definition 2.1.** Sei  $V$  ein reeller Vektorraum. Eine *Norm* auf  $V$  ist eine Abbildung  $\|\cdot\|: V \rightarrow \mathbb{R}$  mit

a)

$$\|x\| = 0 \iff x = 0$$

b)

$$\|\lambda x\| = |\lambda| \|x\| \quad \forall \lambda \in \rho, \quad x \in V$$

c)

$$\|x + y\| \leq \|x\| + \|y\| \quad \forall x, y \in V$$

Dreiecksungleichung.

Ein *normierter VR*  $(V, \|\cdot\|)$  ist ein VR mit einer Norm.

**Beispiele.** •  $\mathbb{R}^n$  mit  $\|\cdot\|_E$  Euklidische Norm auf  $\mathbb{R}^n$

$$\|x\| = \sqrt{\sum_{i=1}^n x_i^2} \quad x = (x_1, \dots, x_n).$$

- $\mathbb{R}^n$  mit  $\|\cdot\|_\infty = \|\cdot\|_{\max}$ ,

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

- $\mathbb{R}^n$  mit  $\|\cdot\|_p$ , „ $p$ -Norm“,  $p \geq 1$ ,  $p \in \mathbb{R}$ .

$$\|x\|_p = \left( \sum_{i=1}^n |x_i|^p \right)^{1/p} \quad \rightarrow \quad \text{Saalübung.}$$

- $C([a, b])$  mit  $\|f\|_{L^1} = \int_a^b |f(t)| dt$ .
- $C([a, b])$  mit  $\|f\|_\infty = \sup_{t \in [a, b]} |f(t)|$ .

**Lemma 2.2.** Sei  $(V, \|\cdot\|)$  normierter VR. Dann wird durch  $d(x, y) := \|x - y\|$  eine Metrik auf  $V$  definiert („induziert“).

*Beweis.* 2.1.a) (Norm)  $\implies$  1.11.a) (Metrik). 1.11.b) (Symmetrie der Metrik): folgt aus  $\|x - y\| = \|y - x\|$ .  $\square$

**Notation.** Wir schreiben  $(V, \|\cdot\|)$  für den *metrischen* Raum, dessen Metrik von  $\|\cdot\|$ , dessen Metrik von  $\|\cdot\|$  induziert wird.

**Bemerkung.** Nicht jede Metrik auf einem Vektorraum wird von einer Norm induziert, denn induzierte Metriken erfüllen  $d(\lambda x, \lambda y) = |\lambda|d(x, y)$ . Die diskrete Metrik erfüllt das nicht.

**Lemma 2.3.** Seien  $d_1$  und  $d_2$  auf  $V$  von Normen  $\|\cdot\|_1$  und  $\|\cdot\|_2$  induziert. Dann ist  $d_2$  stärker als  $d_1$  genau dann, wenn es eine positive Zahl  $C > 0$  gibt s. d.

$$\|x\|_1 \leq C\|x\|_2 \quad \forall x \in V.$$

*Beweis.* Bezeichne  $B_r^j(0)$ ,  $r > 0$ , die offenen Kugeln bezüglich  $d_j$ .

„ $\implies$ “ Nach VOR ist insbesondere  $B_1^1(0)$  offen bezüglich  $d_2 \implies \exists \varepsilon > 0$  s. d.  $B_\varepsilon^2(0) \subset B_1^1(0)$   $\implies$  für  $x \in X$ ,  $x \neq 0$  gilt

$$\begin{aligned} \left\| \frac{\varepsilon x}{2\|x\|_2} \right\|_2 &= \frac{\varepsilon}{2} < \varepsilon \\ \implies \left\| \frac{\varepsilon x}{2\|x\|_2} \right\|_1 &< 1 \\ \implies \|x\|_1 &< \frac{2}{\varepsilon} \|x\|_2. \end{aligned}$$

„ $\Leftarrow$ “ Existiere  $C$  wie oben.  $B_r^2(x) \subset B_{cr}^1(x) \quad \forall x \in X, r > 0$ . Denn  $r > \|x - y\|_2 \geq \frac{1}{c}$ . Sei  $U$  offen bezüglich  $d_1$

$$\begin{aligned} &\implies \forall x \in U \quad \exists \varepsilon > 0 \text{ s. d. } B_\varepsilon^1(x) \subset U \\ &\implies B_{\frac{\varepsilon}{C}}^2(x) \subset B_\varepsilon^1(x) \subset U. \end{aligned} \quad \square$$

**Folgerung.**  $d_2$  ist äquivalent zu  $d_1$

$$\iff \exists C, \tilde{C} \text{ s. d. } \tilde{C}\|x_2\| \leq \|x_1\| \leq C\|x_2\|,$$

(„die Normen sind äquivalent“).

**Bemerkung.** Äquivalenz von Normen ist eine Äquivalenz-Relation (reflexiv, symmetrisch, transitiv).

**Bemerkung.** Für allgemeine Metriken kann man nicht so eine einfache Abschätzung angeben. Etwa ist  $\mathbb{R}^n \times \mathbb{R}^n \ni (x, y) \mapsto \|x - y\|_{\max}$  unbeschränkt, aber  $d_{\text{discrete}}(x, y) = 1 \quad \forall x \neq y$ . Die Beweisrichtung „ $\Leftarrow$ “ gilt noch (also gibt es ein  $C$  wie in 2.3, so ist  $d_2$  stärker als  $d_1$ ). Die Beweisrichtung „ $\Rightarrow$ “ verwendet die Skalar-Multiplikation des zugrunde liegenden Vektorraums und vor allem die „Betrag-Homogenität“ der Norm

$$\left\| \frac{x}{\|x\|_2} \right\| = \frac{1}{\|x\|_2} \|x\|_1.$$

**Satz 2.4.** Auf  $\mathbb{R}^n$  sind alle Normen äquivalent.

*Beweis.* Aufgrund der Transitivität genügt es die Äquivalenz einer beliebigen Norm  $\|\cdot\|$  zu  $\|\cdot\|_\infty$  zu beweisen.

1.  $\|\cdot\|_\infty$  ist stärker als  $\|\cdot\|$ : Denn sei  $x = \sum_{j=1}^n x_j e_j \in \mathbb{R}^n$ ,  $e_j = (0, \dots, \underset{j\text{-te}}{1}, \dots, 0)$ . Dann ist

$$\|x\| = \left\| \sum x_j e_j \right\| \stackrel{\triangle \text{ und Homog.}}{\leq} \left\| \sum |x_j| \cdot e_j \right\| \leq \|x\|_\infty \underbrace{\sum_{j=1}^n \|e_j\|}_{=C}.$$

2.  $\|\cdot\|_\infty$  ist schwächer als  $\|\cdot\|$ :

Betrachte  $M := \{ x \in \mathbb{R}^n \mid \|x\|_\infty = 1 \}$  (Einheits„sphäre“ bezüglich  $\|\cdot\|_\infty$ , also Rand des Einheitswürfels .

**Behauptung.**  $f: M \rightarrow \mathbb{R}$ ,  $x \mapsto \|x\|$  ist stetig bezüglich  $\|\cdot\|_\infty$ .

*Beweis.*

$$|f(x) - f(y)| = \left| \|x\| - \|y\| \right| \stackrel{(*)}{\leq} \|x - y\| \leq C \|x - y\|_\infty$$

(\*) umgekehrte Dreiecksungleichung:

$$\begin{aligned} \|x\| - \|y\| &= \|x + y - y\| - \|y\| \stackrel{\triangle}{\leq} \|x + y\| \\ \|y\| - \|x\| &= \|y + x - x\| - \|x\| \leq \|x + y\|. \end{aligned}$$

□

$M$  ist abgeschlossen bezüglich  $\|\cdot\|_\infty$  (denn  $\mathbb{R}^n \setminus M =$  Urbild der offenen Menge  $\mathbb{R} \setminus \{1\}$  unter der stetigen Abbildung  $x \mapsto \|\cdot\|_\infty$ ).  $M \subset$  abgeschlossenen Quader und dieser ist kompakt in  $(\mathbb{R}^n, \|\cdot\|_\infty)$  (Lemma 2.5)  $\implies M$  ist kompakt (1.39).

Es folgt:  $f$  nimmt sein Minimum  $b$  an und (da  $f > 0$ ) somit ist  $b > 0$ . Nach Definition ist  $\|y\| \geq b \quad \forall y \in M$ . Für alle  $x \in \mathbb{R}^n \setminus \{0\}$  gilt  $\frac{x}{\|x\|_\infty} \in M$ , also ist  $\left\|\frac{x}{\|x\|_\infty}\right\| \geq b$ , also  $\|x\| \geq b\|x\|_\infty$  und für  $x = 0$  gilt dies ohnehin.  $\square$

**Lemma 2.5.** Der Quader  $Q = \{x = (x_1, \dots, x_n) \in \mathbb{R}^n \mid a_j \leq x_j \leq b_j\}$  ist kompakt in  $\mathbb{R}^n, \|\cdot\|_\infty$  ( $a_j \leq b_j$ ).

*Beweis.* Sei  $(U_j)_j$  eine offene Überdeckung von  $Q$ . Angenommen,  $Q$  kann nicht durch endlich viele  $U_j$ 's überdeckt werden.

Wir konstruieren induktiv eine Folge von abgeschlossenen Teilquadern

$$Q_0 \supset Q_1 \supset Q_2 \supset \dots$$

mit

- a)  $Q_n$  kann *nicht* durch endlich viele  $U_j$ 's überdeckt werden
- b)  $\text{diam } Q_m = 2^{-m} \text{ diam } Q$ .

**Beachte:**

$\text{diam } Q = \text{Länge der länsten Seite bezüglich } \|\cdot\|_\infty$ .

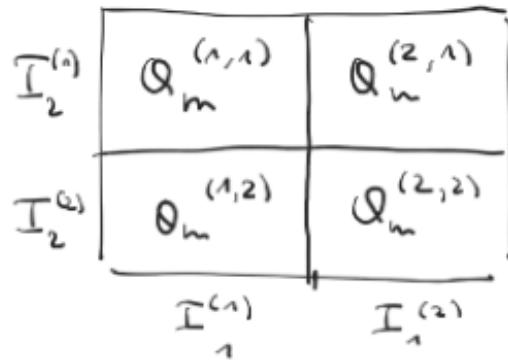


Setze  $Q_0 = Q$ . Sei  $Q_m$  konstruiert. Schreibe  $Q_m = I_1 \times \dots \times I_n$ ,  $I_j$  abgeschlossene Intervalle. Zerlege  $I_j^{(1)} \cup I_j^{(2)}$  in zwei abgeschlossene Intervalle der halben Länge und setze

$$Q^{(s_1, \dots, s_n)} := I_1^{(s_1)} \times \dots \times I_n^{(s_n)}, \quad s_j \in \{1, 2\}.$$

Das ergibt  $2^n$  Quadern mit

$$\bigcup_{s_j \in \{1, 2\}} Q_m^{(s_1, \dots, s_n)} = Q_m$$



Es gibt mindestens einen Quader  $Q_m^{(s_1, \dots, s_n)}$ , der nicht durch endlich viele  $U_j$ 's überdeckt werden kann. Einen solchen wählen wir als  $Q_{m+1}$ . Es gilt per Konstruktion

$$\text{diam}(Q_{m+1}) = \frac{1}{2} \text{diam}(Q_m) = \frac{1}{2^{m+1}} \text{diam}(Q).$$

Nach dem Schachtelungsprinzip  $\exists a \in Q_m \forall m$ . Da  $(U_j)_j Q$  überdeckt  $\exists U_{j_0}$  s. d.  $a \in U_{j_0}$ .  $U_{j_0}$  offen  $\implies \exists \varepsilon > 0$  s. d.  $B_\varepsilon^{\|\cdot\|_\infty}(a) \subset U_{j_0}$ . Wähle  $m$  so groß, dass  $\text{diam}(Q_m) < \varepsilon$ .  $a \in Q_m \implies Q_m \subset B_\varepsilon^{\|\cdot\|_\infty}(a) \subset U_{j_0}$   $\not\subset$  Widerspruch Konstruktion der  $Q_m$ .  $\square$

**Bemerkung 2.6.** Aus 2.4 folgt:  $Q$  ist bezüglich jeder Norm kompakt. Bolzano-Weierstraß (1.43)  $\implies$  In  $(\mathbb{R}^n, \|\cdot\|)$  hat jede beschränkte Folge eine konvergente Teilfolge.

**Bemerkungen 2.7.** Wir haben bereits gesehen:

- i) Auf nicht endlich-dimensionalen Vektor-Räumen sind nicht alle Normen äquivalent:  $(C([a, b]), \|\cdot\|_\infty)$  ist vollständig,  $(C([a, b]), \|\cdot\|_{L^1})$  nicht.
- ii) Auf dem  $\mathbb{R}^n$  sind nicht alle Metriken äquivalent:  $d_{\text{disc}}$  ist stärker als jede Norm (und nicht schwächer).

**Satz 2.8 (Heine-Borel).** Eine Teilmenge  $A \subset \mathbb{R}^n$  ist genau dann kompakt, wenn sie abgeschlossen und beschränkt ist. ( $\mathbb{R}^n$  hier und im Folgenden als normierter VR).

*Beweis.* „ $\implies$ “ Hatten wir letztes Mal (1.38) für Kompakte in metrischen Räumen bewiesen.

„ $\Leftarrow$ “ It  $A$  beschränkt so ist  $A$  in einem Quader enthalten (denn  $\|x - y\|_\infty \leq \|x - y\|$  somit  $\text{diam}_{\|\cdot\|_\infty}(A) \leq C \text{diam}_{\|\cdot\|}(A) < \infty$ ).  $Q$  ist kompakt (bezüglich  $\|\cdot\|_\infty$  somit bezüglich  $\|\cdot\|$ ).  $A$  abgeschlossen  $\implies A$  kompakt (1.39).  $\square$

**Bemerkung.** 2.8 gilt nicht in unendlich-dimensionalen Vektorräumen:

Betrachte in  $\ell_1, \|\cdot\|_{\ell_1} = \sum_{k=0}^{\infty} |x_k|$  die Folge  $(x^n)_n$  wobei  $x^n = (x_k^n)_k$  sei mit  $x_k^n = 0$  für  $n \neq k$  und  $(x^n)_n = 1$ . Dann gilt  $\|x^n\|_{\ell_1} = 1$  und

$$\|x^n - x^m\|_{\ell_1} = 2 \quad \forall m \in \{0, 1, \dots, n-1\}.$$

$\implies$  Die Folge besitzt keine konvergente Teilfolge, kann also (Bolzano-Weierstrass) nicht kompakt sein, obwohl  $\{x^n \mid n \in \mathbb{N}_0\}$  beschränkt und abgeschlossen in  $(\ell_1, \|\cdot\|_{\ell_1})$  ist.

**Vorlesung 5**

Mo 04.05. 10:15

**2.I. Stetige Abbildungen in normierten Vektorräumen****2.I.1. Lineare Abbildungen**

**Satz 2.9.** Seien  $(V, \|\cdot\|_V)$  und  $(W, \|\cdot\|_W)$  normierte Vektorräume. Sei  $A: V \rightarrow W$  linear. Dann sind folgende Aussagen äquivalent:

- a)  $A$  ist stetig
- b)  $A$  ist stetig in 0
- c)  $\|A(x)\|_W \leq C\|x\|_V$ .

*Beweis.* 2.9.a)  $\implies$  2.9.b) ✓

2.9.b)  $\implies$  2.9.c)  $A$  stetig in 0  $\implies$  zu  $\varepsilon = 1 \exists \delta > 0$  s. d.

$$\|A(y) - A(0)\|_W \stackrel{\text{Lin}}{=} \|A(y)\|_W < 1 \quad \forall y \in V \text{ mit } \|y - 0\|_V = \|y\|_V < \delta.$$

Setze  $C := 2/\delta$ . Sei  $x \in V \setminus \{0\}$  beliebig (für  $x = 0$  gilt die Ungleichung 2.9.c) ohnehin). Setze  $\lambda := 1/C\|x\|_V$  und  $y := \lambda x$ .

Dann ist  $\|y\|_V = \frac{1}{C\|x\|_V}\|x\|_V = \delta/2 < \delta$ , also  $\|A(y)\|_W < 1$ .

$$A(y) = A(\lambda x) = \frac{1}{C\|x\|_V}A(x) \implies \text{Beh.}$$

2.9.c)  $\implies$  2.9.a) Es gebe  $C > 0$  s. d.

$$\|A(x)\|_W \leq C\|x\|_V \quad \forall x \in V.$$

Dann gilt insbesondere für  $x = y - a$ .

$$\|A(x)\|_W = \|A(y) - A(a)\| \underset{\substack{\uparrow \\ \text{Linearität}}}{\leq} C\|y - a\|_V.$$

Sei  $\varepsilon > 0$ . Dann ist also

$$\|A(y) - A(a)\|_W < \varepsilon \quad \forall y, a \text{ mit } \|y - a\|_V < \frac{\varepsilon}{C}$$

und somit ist  $A$  sogar gleichmäßig stetig. □

**Beispiele.** i)  $(C([a, b], \mathbb{R}), \|\cdot\|_\infty)$ .

$$I: C([a, b]) \rightarrow \mathbb{R}, \quad I(f) := \int_a^b f(t) dt.$$

$I$  ist linear und es gilt

$$\|I(f)\| \leq (b-a)\|f\|_\infty$$

$\implies I$  ist stetig.

ii)  $D(\cdot): (C^1([a, b]), \|\cdot\|_\infty) \rightarrow (C([a, b]), \|\cdot\|_\infty), D(\cdot): f \mapsto f'$ .

**Behauptung.**  $D(\cdot)$  ist nicht stetig.

*Denn:*  $D$  ist linear ✓, aber die Bedingung aus Satz 2.9 ist verletzt: Betrachte  $f_n \in C^1([0, 2]), f_n = x^n$ . Dann ist  $\|f_n\|_\infty = 1$ , aber  $\|Df_n\|_\infty = n \implies$  es kann kein  $C > 0$  geben s. d.

$$n = \|Df_n\|_\infty \leq C\|f_n\| = C \quad \forall n.$$

□

**Definition.** Seien  $V$  und  $W$  normierte Vektorräume. Sei  $A: V \rightarrow W$  lineare stetige Abbildung. Die *Operatornorm* von  $A$  ist definiert als

$$\|A\|_{\text{op}} := \sup_{\substack{x \in V \\ x \neq 0}} \frac{\|Ax\|_w}{\|x\|_V}.$$

Auf dem VR der stetigen linearen Funktionen  $V \rightarrow W$  ist  $\|\cdot\|_{\text{op}}$  eine Norm.  $\|A\|_{\text{op}}$  ist die kleinste Konstante für die noch die Abschätzungen aus 2.9 gilt und es folgt

**Bemerkung 2.10.** Ein linearer Operator ist genau dann stetig, wenn gilt  $\|A\|_{\text{op}} < \infty$ .

**Beispiel.** Ist  $A: \mathbb{R}^n \rightarrow \mathbb{R}^m$  linear, so gilt

$$A \in \text{Mat}(m \times n, \mathbb{R}) \simeq \mathbb{R}^{m \cdot n}.$$

Daher ist  $\|\cdot\|_{\text{op}}$  in diesem Fall äquivalent zu  $\|\cdot\|_\infty$ ,  $\|A\|_\infty = \max_{i,j} |A_{ij}| < \infty$ , insbesondere also schwächer und somit ist  $A$  stetig.

Konkret gilt: Setze  $V = (\mathbb{R}^n, \|\cdot\|_V)$ ,  $W = (\mathbb{R}^n, \|\cdot\|_W)$ . Sei  $y = Ax \implies y_i = \sum_{j=1}^n A_{ij}x_j$  für  $i = 1, \dots, m$ .

$$\begin{aligned}\|y\|_W &\stackrel{\triangle}{\leq} \sum_{i=1}^m \|y_i e_i\|_W \\ &\stackrel{\triangle}{\leq} \sum_{\text{Hom } i,j} |A_{ij}x_j| \|e_i\|_W \\ &= \sum_{i,j} |A_{ij}| \cdot |x_j| \cdot \|e_i\|_W \\ &\leq \|A\|_\infty \cdot \|x\|_{\ell^1} \cdot \underbrace{\sum_{i=1}^m \|e_i\|_W}_{=c_W} \\ \implies \|A\|_{\text{op}} &= \sup_{x \neq 0} \frac{\|Ax\|_W}{\|x\|_V} \leq \|A\|_\infty \cdot C_W \cdot \underbrace{\sup_{x \neq 0} \frac{\|x\|_{\ell^1}}{\|x\|_V}},\end{aligned}$$

wobei  $C_V$  eine Konstante ist mit

$$\|x\|_{\ell^1} \leq C_V \cdot \|x\|_V \quad \forall x \in V = \mathbb{R}^n.$$

**Bemerkung.** Unsere Beschränkung auf den  $\mathbb{R}^n$  (statt beliebige endlich-dimensionale Vektorräume zuzulassen), bedeutet also keine Einschränkung, da ein Basiswechsel nach der Überlegung oben stetig ist.

**Beispiele 2.11.**  $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$ .

a) Kurven  $\gamma: I \rightarrow \mathbb{R}^n$ ,  $I$  Intervall, stetig.

**Beispiele.**

$\gamma: [0, 2\pi] \rightarrow \mathbb{R}^2$ ,  $t \mapsto (r \cos t, r \sin t)$ ,  $r > 0$ . Stetigkeit: Wir versehen  $\mathbb{R}^2$  mit  $\|\cdot\|_\infty$ . Dann folgt die Stetigkeit von  $\gamma$  aus der Stetigkeit der Komponentenfunktionen  $I \rightarrow \mathbb{R}$ .

$\gamma: \mathbb{R} \rightarrow \mathbb{R}^2$ ,  $t \mapsto (t^2 - 1, t^3 - 1)$  genauso. Spur von  $\gamma = \{ \gamma(t) \mid t \in \mathbb{R} \}$



b) Gebrochen rationale Funktionen:

**Beispiele.**

$$f: \mathbb{R}^2 \rightarrow \mathbb{R}.$$

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

$f$  ist stetig: auf  $\mathbb{R}^2 \setminus \{0\}$  sicherlich als Verknüpfung und Produkt stetiger Funktionen:

$$f = \text{Inv} \circ p_1 \cdot p_2 \quad \text{Inv}(t) = \frac{1}{t}, \quad p_1(x, y) = x^2 + y^2, \quad p_2(x, y) = x^2 y.$$

Stetigkeit in 0: Es gilt  $(x - y)^2 \geq 0$

$$\begin{aligned} &\implies 2|xy| \leq x^2 + y^2 \\ &\implies \left| \frac{x^2 y}{x^2 + y^2} \right| < \frac{|x|}{2} \end{aligned}$$

für  $((x_n, y_n))_n$ ,  $(x_n, y_n) \rightarrow (0, 0)$  (bezüglich irgendeiner Norm) gilt insbesondere  $x_n \rightarrow 0$

$$\implies |f(x_n, y_n) - 0| = \left| \frac{x_n^2 y_n}{x_n^2 + y_n^2} \right| < \frac{|x_n|}{2} \rightarrow 0 \text{ in } \mathbb{R}.$$

$$f: \mathbb{R}^2 \rightarrow \mathbb{R},$$

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

$f$  ist stetig auf  $\mathbb{R}^2 \setminus \{0\}$  (siehe oben).  $f$  ist nicht stetig in 0: Betrachte etwa  $(x_n, y_n) = \left(\frac{1}{n}, \frac{1}{n^2}\right)$ ,  $n \geq 1$ . Dann gilt

$$f(x_n, y_n) = \frac{1}{n^2 n^2} \left( \frac{n^4}{2} \right) = \frac{1}{2} \not\rightarrow 0.$$

**Achtung:**

Es gibt durchaus Folgen  $(x_n, y_n) \rightarrow 0$  s. d.  $f(x_n, y_n) \rightarrow 0$  (für  $n \rightarrow \infty$ ), z. B.  $(x_n, y_n) = \left(0, \frac{1}{n}\right)$ , wo  $f\left(0, \frac{1}{n}\right) = 0 \quad \forall n$  oder  $(x_n, y_n) = (1/n, 1/n)$  wo

$$f(x_n, y_n) = \frac{1}{n^2} \left( \frac{n^2}{1 + 1/n^2} \right) \rightarrow 0.$$

Daher muss man, wenn man Stetigkeit zeigen will, in Argument finden, dass für *alle* Folgen funktioniert.

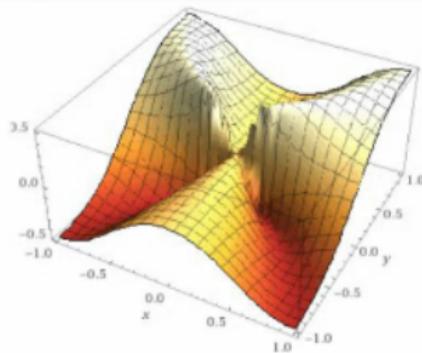
Plot3d[x^2\*y/(x^4+y^2),{x,-1,1},{y,-1,1}]

Extended Keyboard   Upload

Input interpretation:

plot	$x^2 \times \frac{y}{x^4 + y^2}$	$x = -1 \text{ to } 1$
		$y = -1 \text{ to } 1$

3D plot:



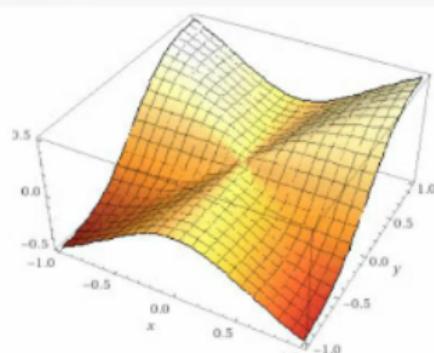
Plot3d[x^2\*y/(x^2+y^2),{x,-1,1},{y,-1,1}]

Extended Keyboard   Upload

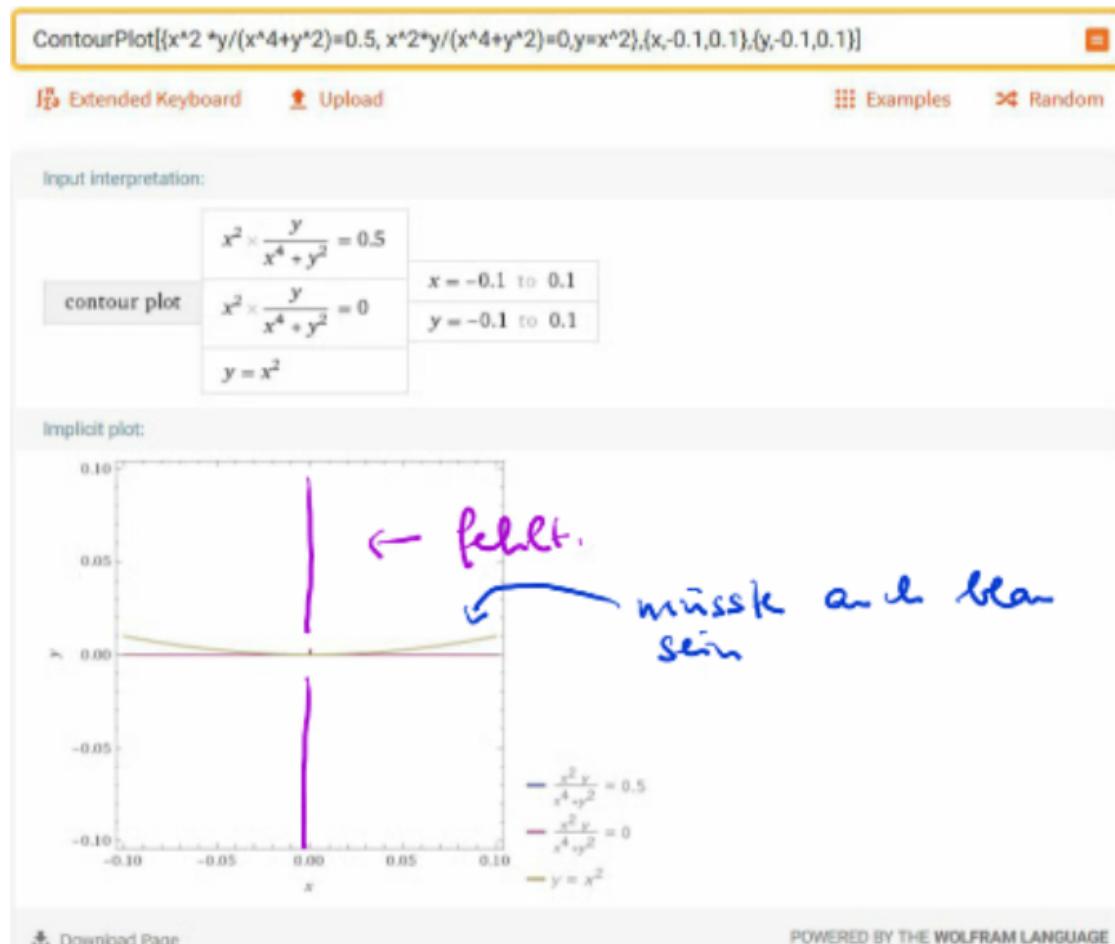
Input interpretation:

plot	$x^2 \times \frac{y}{x^2 + y^2}$	$x = -1 \text{ to } 1$
		$y = -1 \text{ to } 1$

3D plot:



Contour-Plot: Eingezeichnet werden alle  $(x, y)$ , die die gegebene Gleichung erfüllen. Von Wolfram Alpha.



## 2.II. Vektorräume mit Skalarprodukt

Eine spezielle Klasse von Normen sind solche, die von einem sogenannten Skalarprodukt induziert werden.

**Definition 2.12.** Sei  $V$  ein Vektorraum über  $\mathbb{R}$ . Ein *Skalarprodukt* auf  $V$  ist eine Abbildung  $\langle \cdot, \cdot \rangle: V \times V \rightarrow \mathbb{R}$  mit

a) Linear:

$$\langle \lambda x + \mu y, z \rangle = \lambda \langle x, z \rangle + \mu \langle y, z \rangle \quad \forall x, y, z \in V, \lambda, \mu \in \mathbb{R}$$

b) Symmetrisch:

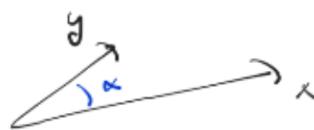
$$\langle x, y \rangle = \langle y, x \rangle \forall x, y \in V$$

c) Positiv definit:

$$\langle x, x \rangle \geq 0 \text{ und } \langle x, x \rangle = 0 \iff x = 0.$$

**Bemerkung.** Mit 2.12.b) folgt auch die Linearität im zweiten Argument.

**Beispiele.** •  $\mathbb{R}^n$ ,  $\langle x, y \rangle = \sum_{i=1}^n x_i y_i$ : Euklidisches Skalarprodukt.



interpretation

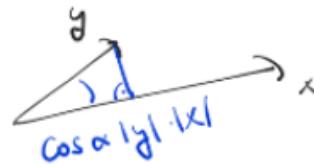
$y$  geht durch Drehstreckung aus  $x \neq 0$  hervor:

$$y = \|y\|_E \begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix} \frac{x}{\|x\|_E}.$$

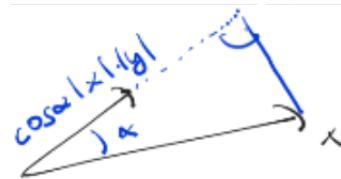
Dann gilt:

$$\begin{aligned} \langle x, y \rangle &= \frac{\|y\|_E}{\|x\|_E} \langle x, (\cos \alpha x_1 - \sin \alpha x_2, \sin \alpha x_1 + \cos \alpha x_2) \rangle \\ &= \frac{\|y\|_E}{\|x\|_E} (x_1^2 \cos \alpha - x_1 x_2 \sin \alpha + x_1 x_2 \sin \alpha + x_2^2 \cos \alpha) \\ &= \|y\|_E \cdot \|x\|_E \cdot \cos \alpha. \end{aligned}$$

Das Skalarprodukt misst die Projektion von  $y$  auf  $x$



und umgekehrt



- $\mathbb{R}^n$  mit  $\langle x, y \rangle_W = \sum_{i=1}^n w_i x_i y_i$ ,  $w = (w_1, \dots, w_n)$  Gewichtsvektor,  $w_i > 0$ .
- $\mathbb{R}^2$  mit  $\langle x, y \rangle := 2x_1 y_1 - x_1 y_2 - x_2 y_1 + 2x_2 y_2$  (zu überprüfen ist die Positive Definitheit).
- Kein Skalarprodukt ist das Minkowski-Produkt:  $\mathbb{R}^{n+1}$  mit  $((x, y)) := x_0 y_0 - \sum_{i=1}^n x_i y_i$ .  
Denn  $((x, x)) = 0 \iff x_0 = \pm \|x\|_E$ ,  $x = (x_1, \dots, x_n)$ .
- $C([a, b])$  mit  $\langle f, g \rangle = \int_a^b f(t)g(t) dt$ .

**Lemma 2.13.** Sei  $V$  VR mit Skalarprodukt  $\langle \cdot, \cdot \rangle$ . Dann ist durch  $\|x\| := \sqrt{\langle x, x \rangle}$  eine Norm auf  $V$  definiert.

*Beweis.* 2.1.a)  $\|x = 0\| \implies \langle x, x \rangle = 0 \stackrel{2.12.c)}{\implies} \|0\| = 0 \checkmark$ .

2.1.b)  $\|\lambda x\| = \sqrt{\lambda^2 \langle x, x \rangle} = |\lambda| \sqrt{\langle x, x \rangle} \forall \lambda \in \mathbb{R}, x \in V$ .

2.1.c)

$$\begin{aligned} \|x + y\|^2 &= \langle x + y, x + y \rangle \\ &= \|x\|^2 + 2\langle x, y \rangle + \|y\|^2 \\ &\leq \|x\|^2 + 2|\langle x, y \rangle| + \|y\|^2 \\ &\leq (\|x\| + \|y\|)^2 \implies \Delta, \\ &\text{siehe (*) unten} \end{aligned}$$

denn die Wurzel ist monoton wachsend.

Es gilt die *Cauchy-Schwarzsche* Ungleichung:

$$|\langle x, y \rangle| \leq \|x\| \cdot \|y\|. \quad (*)$$

*Beweis.*

$$0 \leq \langle x - \lambda y, x - \lambda y \rangle = \|x\|^2 - 2\lambda \langle x, y \rangle + \lambda^2 \|y\|^2 \quad \forall x, y \in V \quad \lambda \in \mathbb{R},$$

also speziell für  $y \neq 0$  (für  $y = 0$  gilt die Ungleichung sowieso) und  $\lambda = \frac{\langle x, y \rangle}{\|y\|^2}$ :

$$0 \leq \|x\|^2 - \frac{\langle x, y \rangle^2}{\|y\|^2}.$$

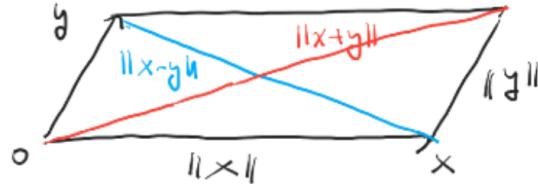
□

Einen Vektorraum mit Skalarprodukt betrachten wir immer als mit der von Skalarprodukt induzierten Norm, also Metrik, also Topologie.

Nicht jede Norm wird von einem Skalarprodukt induziert. Es gilt

**Lemma 2.14.** Sei  $(V, \|\cdot\|)$  normierter VR Dann wird  $\|\cdot\|$  von einem Skalarprodukt induziert genau dann, wenn die Parallelogramm-Gleichung gilt:

$$\|x + y\|^2 - \|x - y\|^2 = 2(\|x\|^2 + \|y\|^2) \quad \forall x, y \in V.$$



Erklärung für den Namen.  $\mathbb{R}^2$ ,  $\|\cdot\| = \|\cdot\|_E$ .

*Beweis.* „ $\implies$ “ Sei  $\|x\| = \sqrt{\langle x, x \rangle}$ . Dann gilt

$$\begin{aligned} \|x + y\|^2 + \|x - y\|^2 &= \langle x + y, x + y \rangle + \langle x - y, x - y \rangle \\ &= 2\|x\|^2 + 0 + 2\|y\|^2. \end{aligned}$$

„ $\impliedby$ “ Erfüllt  $\|\cdot\|$  die Parallelogramm-Gleichung.

**Behauptung.** Durch „Polarisation“, also

$$\langle x, y \rangle := \frac{1}{4}(\|x + y\|^2 - \|x - y\|^2)$$

ist ein Skalarprodukt definiert mit  $\|x\| = \sqrt{\langle x, x \rangle}$ .

*Beweis.* 2. Beh.:  $\langle x, x \rangle = \frac{1}{4} \|2x\|^2 \checkmark$ .

1. Beh.: • Aus der 2. Beh. folgt die positive Definitheit aus der Nichtausgeartetheit und Positivität der Norm.  
• Die Symmetrie folgt sofort aus der Definition.  
• Linearität. Wir zeigen zunächst Additivität:

$$1) \quad \langle x + u, y \rangle + \langle x - u, y \rangle = \langle x, y \rangle$$

denn::

$$\begin{aligned} \text{linke Seite} &= \frac{1}{4} (\|x + u + y\|^2 - \|x + u - y\|^2) \\ &\quad + \|x - u + y\|^2 - \|x - u - y\|^2 \\ &= \frac{1}{2} (\|x + y\|^2 + \|u\|^2 - \|x - y\|^2 - \|u\|^2) \\ &\quad \uparrow \text{Parallelogramm-Gleichung} \\ &= \frac{1}{2} (\langle x + y, x + y \rangle - \langle x - y, x - y \rangle) \\ &= 2\langle x, y \rangle. \end{aligned} \quad \square$$

Damit auch gleich gezeigt:

- 2)  $\langle 2x, y \rangle = 2\langle x, y \rangle$  (setze  $u = x$ ) und mit  $x = u + v, y = u - v$  folgt  
3) Additivität:

$$\begin{aligned} \langle x, y \rangle + \langle y, z \rangle &= \langle u + v, z \rangle + \langle u - v, z \rangle \\ &= 2\langle u, z \rangle \\ &\stackrel{2)}{=} \langle 2u, z \rangle \\ &= \langle x + y, z \rangle \end{aligned}$$

- 4) per Induktion  $\langle nx, y \rangle = n\langle x, y \rangle \forall n \in \mathbb{N}$ , denn

$$\begin{aligned} \langle (n+1)x, y \rangle &= \langle nx + x, y \rangle \\ &\stackrel{3)}{=} \langle nx, y \rangle + \langle x, y \rangle \\ &\stackrel{\text{IV}}{=} n\langle x, y \rangle + \langle x, y \rangle \\ &= (n+1)\langle x, y \rangle. \end{aligned}$$

5) Für  $\lambda \in -\mathbb{N}_0$  gilt

$$\begin{aligned}\lambda \langle x, y \rangle - \langle \lambda x, y \rangle &= \lambda \langle x, y \rangle - \langle |\lambda|(-x), y \rangle \\ &= \lambda \langle x, y \rangle - |\lambda| \langle -x, y \rangle \\ &= \lambda(\langle x, y \rangle + \langle -x, y \rangle) \\ &= 0\end{aligned}$$

6) Für  $\lambda \in \mathbb{Q}$ ,  $\lambda = m/n$ ,  $m, n \in \mathbb{Z}$ :

$$n \left\langle \frac{m}{n} x, y \right\rangle \stackrel{4),5)}{=} \langle mx, y \rangle = m \langle x, y \rangle.$$

7) Für  $\lambda \in \mathbb{R}$  existiert  $(\lambda_n)_n \subset \mathbb{Q}$ ,  $\lambda_n \rightarrow \lambda$ . Da  $\|\cdot\|$  stetig ist, so auch  $\langle \cdot, \cdot \rangle$

$$\begin{aligned}\langle \lambda x, y \rangle &= \langle \lim \lambda_n x, y \rangle \\ &= \lim \langle \lambda_n x, y \rangle \\ &= \lim \lambda_n \langle x, y \rangle \\ &= \lambda \langle x, y \rangle.\end{aligned}$$

Symmetrie  $\implies$  es genügt, das erste Argument zu untersuchen.  $\square$

**Beispiel.**  $\|\cdot\|_{\max}$  wird nicht von einem Skalarprodukt induziert: Sei  $x = e_1$ ,  $y = e_2$ . Dann gilt:

$$\|e_1 + e_2\|_{\max}^2 + \|e_1 - e_2\|_{\max}^2 = 1 + 1 = 2,$$

aber

$$2(\|e_1\|_{\max}^2 + \|e_2\|_{\max}^2) = 4.$$

# Kapitel 3.

## Differenzierbarkeit in $\mathbb{R}^n$

### Vorlesung 6

Do 07.05. 10:15

**Erinnerung.** Approximation einer Funktion  $f: \mathbb{R} \rightarrow \mathbb{R}$ , die in  $a \in \mathbb{R}$  differenzierbar ist, durch eine (affin) lineare Funktion

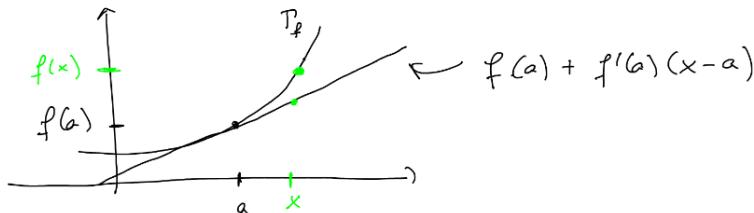
$$f(x) = f(a) + m_a(x - a) + R_a(x)$$

mit  $R_a: \mathbb{R} \rightarrow \mathbb{R}$  und  $\lim_{x \rightarrow a} \frac{R_a(x)}{x - a} = 0$ .

Gibt es ein solches  $R_a$ , so ist  $m_a$  eindeutig festgelegt und es gilt

$$m_a = \lim_{x \rightarrow a} \underbrace{\frac{f(x) - f(a)}{x - a}}_{\text{Differenzenquotient}} .$$

$f'(a) := m_a$  heißt *Ableitung* von  $f$  an der Stelle  $a$ .



Für Abbildungen  $f: \mathbb{R}^n \rightarrow \mathbb{R}^n$  kann man analog definieren:

**Definition 3.1.** Die *Ableitung* einer Funktion  $f: U \rightarrow \mathbb{R}^m$ ,  $U \subset \mathbb{R}^n$  offen, an der Stelle  $a \in U$ , ist, wenn sie existiert, eine Matrix  $Df(a) \in \text{Mat}(m \times n)$ , die eine lineare Approximation von  $f$  ergibt:

$$f(x) = f(a) + Df(a) \cdot \underset{\substack{\uparrow \\ \text{Matrix-Multiplikation}}}{(x - a)} + R_a(x) \quad (*)$$

Sie existiert genau dann, wenn

$$\lim_{\substack{x \rightarrow a \\ \text{in } \mathbb{R}^n}} \frac{R_a(x)}{\|x - a\|} = 0$$

↑  
eine Norm in  $\mathbb{R}^n$

und ist in diesem Fall eindeutig durch (\*) bestimmt und man sagt,  $f$  ist *differenzierbar*.

**Bemerkungen.** i) Für  $n = m = 1$  stimmt die Definition mit der Üblichen überein, da  $\frac{R_a(x)}{x-a} \rightarrow 0 \iff \frac{R_a(x)}{|x-a|} \rightarrow 0$ .

ii) Eindeutigkeit: Sei  $A$  s. d.

$$\begin{aligned} f(x) &= f(a) + A \cdot (x - a) + \tilde{R}_a(x) \\ &= f(a) + Df(a) \cdot (x - a) + R_a(x) \end{aligned}$$

Dann folgt:

$$\lim_{x \rightarrow a} \underbrace{((A - Df(a)) \cdot (x - a) \cdot 1/\|x - a\|)}_{=\frac{1}{\|x-a\|}(R_a(x)-\tilde{R}_a(x))} = 0.$$

$\implies A - Df(a) = 0$  (Nullmatrix), da wegen der Stetigkeit linearer Abbildungen  $\mathbb{R}^n \rightarrow \mathbb{R}^m$  der Grenzwert gleich  $(A - Df(a)) \cdot \lim (x - a)/\|x - a\|$  ist und  $\|\lim (x - a)/\|x - a\|\| = 1 \neq 0$ .

iii) Wie in der DIFF I ist es oft zweckmäßig  $f$  an der Stelle  $a + h$  mit  $f$  an der Stelle  $a$  zu vergleichen ( $x = a + h$ ).

$$f(a + h) = f(a) + Df(a) \cdot h + \underline{R}_h(a)$$

$$\text{mit } \lim_{h \rightarrow 0} (\underline{R}_h(a))/\|h\| = 0.$$

**Beispiele.** i)  $f: \mathbb{R}^n \rightarrow \mathbb{R}^n$ ,  $f(x) = A \cdot x + b$ ,  $A \in \text{Mat}(m \times n, \mathbb{R})$ ,  $b \in \mathbb{R}^m$ .

$$f(a + h) = f(a) + A \cdot h \implies Df(a) = A.$$

Insbesondere verschwindet die Ableitung einer konstanten Funktion.

ii)  $f: \mathbb{R}^n \rightarrow \mathbb{R}$ ,  $f(x) = \langle x, B \cdot x \rangle$ ,  $\langle \cdot, \cdot \rangle$  euklidisches Skalarprodukt,  $B \in \text{Mat}(n \times n, \mathbb{R})$ .

$$\begin{aligned} f(a + h) &= \langle a + h, B \cdot (a + h) \rangle \\ &= \langle a, B \cdot a \rangle + \langle h, B \cdot a \rangle + \langle a, B \cdot h \rangle + \langle h, B \cdot h \rangle \\ &= \langle a, B \cdot a \rangle + \langle (B + B^T) \cdot a, h \rangle + \langle h, B h \rangle. \end{aligned}$$

↑  
CHECK!

Wegen (wähle  $\|\cdot\| = \|\cdot\|_{\text{E}}$ )

$$\frac{\langle a, Bh \rangle}{\|h\|} \underset{\text{C-S}}{\leqslant} \frac{\|h\| \cdot \|Bh\|}{\|h\|} \leqslant \|B\|_{\text{op}} \cdot \|h\| \rightarrow 0 \text{ für } h \rightarrow 0$$

folgt:  $f$  ist in allen  $a \in \mathbb{R}^n$  differenzierbar und

$$\begin{aligned} Df(a) \cdot h &= \langle (B + B^T) \cdot a, h \rangle = (b_1, \dots, b_n) \begin{pmatrix} h_1 \\ \vdots \\ h_n \end{pmatrix} \\ \implies Df(a) &= b = ((B + B^T) \cdot a)^T \in \text{Mat}(1 \times n, \mathbb{R}) \quad \forall h \in \mathbb{R}^n. \end{aligned}$$

Aus der Definition folgt sofort

**Satz 3.2.** Sei  $f: U \rightarrow \mathbb{R}^m$ ,  $U \subset \mathbb{R}^n$  offen, in  $a \in U$  differenzierbar. Dann ist  $f$  in  $a$  stetig.

*Beweis.*

$$\begin{aligned} \lim_{h \rightarrow 0} f(a + h) &= \lim_{h \rightarrow 0} (f(a) + Df(a) + \underline{R}_a(h)) \quad \square \\ &= f(a) + 0 + 0 \\ &\quad \uparrow \quad \uparrow \\ &\quad \left| \begin{array}{l} \text{Es gilt sogar } \underline{R}_a(h)/\|h\| \rightarrow 0 \\ \|Df(a) \cdot h\| \leq \|Df(a)\|_{\text{op}} \cdot \|h\| \end{array} \right. \\ &\quad \uparrow \quad \uparrow \\ &\quad \text{Norm in } \mathbb{R}^m \quad \text{Norm in } \mathbb{R}^n \end{aligned}$$

**Satz 3.3 (Kettenregel).** Seien  $U \subset \mathbb{R}^n$ ,  $V \subset \mathbb{R}^m$  offen,  $g: U \rightarrow \mathbb{R}^m$ ,  $f: V \rightarrow \mathbb{R}^k$ ,  $g(U) \subset V$ . Ist  $g$  in  $a \in U$  differenzierbar und  $f$  in  $b = g(a)$ , so ist die Verkettung  $f \circ g: U \rightarrow \mathbb{R}^k$  in  $a$  differenzierbar und es gilt

$$D(f \circ g)(a) = \underbrace{Df(g(a))}_{\in \text{Mat}(k \times m, \mathbb{R})} \underbrace{Dg(a)}_{\in \text{Mat}(m \times n, \mathbb{R})}.$$

*Beweis.*

$$g(a + u) = g(a) + A \cdot u + \underline{R}_a^g(u) \quad A = Dg(a) \quad (1)$$

$$f(b + v) = f(b) + B \cdot v + \underline{R}_b^f(v) \quad B = Df(b) \quad (2)$$

Setze speziell  $v := g(a+u) - g(a) \stackrel{(1)}{=} A \cdot u + \underline{R}_a^g(u)$ .

$$\begin{aligned} \implies f \circ g(a+u) &= f(g(a+u)) = f(g(a)+v) && \text{Def } v \\ &\stackrel{(2)}{=} f(g(a))+B \cdot v + \underline{R}_b^f(v) && b = g(a) \\ &= f(g(a))+B \cdot A \cdot u + \underbrace{B \cdot \underline{R}_a^g(u) + \underline{R}_b^f(A \cdot u + \underline{R}_a^g(u))}_{\substack{\uparrow \\ \text{Def } v \text{ und (1)}}}. && \text{zu zeigen: } \frac{\cdot}{\|u\|} \xrightarrow{u \rightarrow 0} 0 \text{ für } u \rightarrow 0 \end{aligned}$$

- $\frac{\underline{R}_a^g(u)}{\|u\|} \rightarrow 0 \implies \exists C > 0 \text{ s. d. } \|\underline{R}_a^g(u)\| \leq C \|u\|$ .
- $\frac{\underline{R}_b^f(v)}{\|v\|} \rightarrow 0 \implies \exists r_b^f \text{ s. d. } \underline{R}_b^f(v) = \|v\| r_b^f(v) \text{ mit } r_b^f(v) \rightarrow 0 \text{ (} v \rightarrow 0 \text{)}$ .

$$\begin{aligned} \implies \underline{R}_b^f(A \cdot u + \underline{R}_a^g(u)) &\leq \overbrace{A \cdot u + \underline{R}_a^g(u)}^{\substack{\uparrow \\ \text{Norm in } \mathbb{R}^k}} \cdot \overbrace{\|r_b^f(Au + \underline{R}_a^g(u))\|}^{\substack{\uparrow \\ \text{Norm in } \mathbb{R}^k}} && \xrightarrow{u \rightarrow 0 \text{ für } u \rightarrow 0} \\ &\implies \frac{\underline{R}_b^f(A \cdot u + \underline{R}_a^g(u))}{\|u\|} \rightarrow 0 \quad (u \rightarrow 0). && \square \end{aligned}$$

**Satz 3.4 (Produktregel, Quotientenregel).** Seien  $f, g: U \rightarrow \mathbb{R}$ ,  $U \subset \mathbb{R}^n$  offen, differenzierbar in  $a \in U$ . Dann gilt

i)  $f \cdot g$  ist differenzierbar in  $a$  und es gilt

$$D(f \cdot g)(a) = Df(a) \cdot g(a) + f(a) \cdot Dg(a).$$

ii) Ist  $g(a) \neq 0$  so gilt:  $f/g$  ist auf einer Umgebung von  $a$  definiert und differenzierbar in  $a$  und es gilt

$$D(f/g)(a) = Df(a) \cdot \frac{1}{g(a)} - f(a) \cdot \frac{1}{g(a)^2} \cdot Dg(a).$$

*Beweis.* 3.4.i)

$$\begin{aligned} f(a+h) &= f(a) + Df(a) \cdot h + \underline{R}_a^f(h) \\ g(a+h) &= g(a) + Dg(a) \cdot h + \underline{R}_a^g(h) \\ f \cdot g(a+h) &= (f(a) + Df(a) \cdot h + \underline{R}_a^f(h))(g(a) + Dg(a) \cdot h + \underline{R}_a^g(h)) \\ &= f(a) \cdot g(a) + (\underbrace{Df(a) \cdot g(a)}_{\in \text{Mat}(1 \times n, \mathbb{R})}) \cdot h + (\underbrace{f(a) \cdot Dg(a)}_{\in \text{Mat}(1 \times n, \mathbb{R})}) \cdot h \\ &\quad + \underbrace{(Df(a) \cdot h + \underline{R}_a^f(h))(Dg(a) \cdot h + \underline{R}_a^g(h))}_{\frac{\cdot}{\|h\|} \xrightarrow{h \rightarrow 0} 0 \text{ für } h \rightarrow 0}. \end{aligned}$$

3.4.ii)  $g$  ist in  $a$  stetig  $\implies \exists$  Umgebung von  $a$  s. d.  $g(x) \neq 0 \forall x \in U$  (wie in der DIFF I: Sei o.B.d.A.  $g(a) > 0$ . Sei  $\varepsilon := g(a)/2$ . Sei  $\|\cdot\|$  irgendeine Norm auf  $\mathbb{R}^n$ . Dann gibt es ein  $\delta > 0$  s. d.

$$\begin{aligned} |g(x) - g(a)| &< \varepsilon \quad \forall x \in B_\delta^{\|\cdot\|}(a) \\ \implies -\frac{g(a)}{2} &< g(x) - g(a) < \frac{g(a)}{2} \quad \forall x \in B_\delta^{\|\cdot\|}(a), \end{aligned}$$

also  $0 < \frac{g(a)}{2} < g(x) < \frac{3}{2}\frac{g(a)}{2}$ .)  $\implies$  Auf  $U$  ist  $f/g$  wohldefiniert. Die Berechnung der Ableitung ist analog zu 3.4.i), nachdem man sich überlegt hat, dass

$$\frac{1}{g} = \text{Inv} \circ g, \quad \text{Inv}(t) = \frac{1}{t}$$

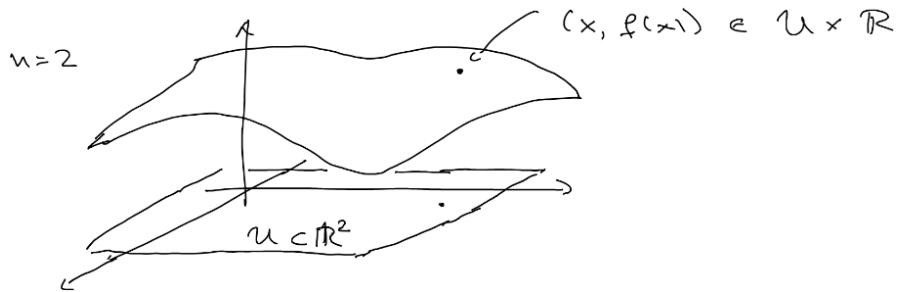
und somit nach der Kettenregel

$$D\left(\frac{1}{g}\right)(a) = D\text{Inv}(g(a)) \cdot Dg(a)$$

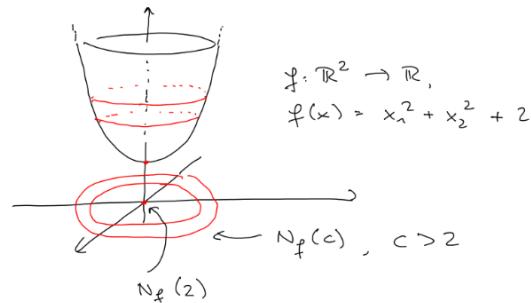
und  $D\text{Inv}(t) = -\frac{1}{t^2}$  (DIFF I).  $\square$

### 3.I. Geometrische Anschauung, partielle Ableitung

DIFF I: Ableitung beschreibt Rate der Veränderung. Höher-dimensional: Sei  $f: U \rightarrow \mathbb{R}$ ,  $U \subset \mathbb{R}^n$ . Betrachte den Graph  $\Gamma_f = \{(x, f(x)) \mid x \in U\}$  (vgl. die Diskussion bei 2.11.b))



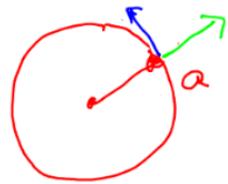
**Definition (Niveau-Mengen).** Zu  $c \in \mathbb{R}$  setze  $N_f(c) := \{x \in U \mid f(x) = c\}$ .



Es gilt  $N_f(c) = \emptyset$  für  $c < 2$ .

**Beispiel.** Entlang der Niveaumengen ist  $f$  konstant. Wie wird das in der Ableitung sichtbar?

Im Beispiel oben ist  $Df(a) = (2a_1, 2a_2)$  (check!). Sei  $a = (r \cos \phi, r \sin \phi)$ ,  $r > 0$ . Betrachte  $f(a + h) = f(a) + Df(a) \cdot h + R_a(h)$ . Ist  $h = (-\varepsilon \sin \phi, \varepsilon \cos \phi)$ ,  $\varepsilon > 0$  (in „Richtung“ der Niveaumenge), so ist  $Df(a) \cdot h = 0$ .

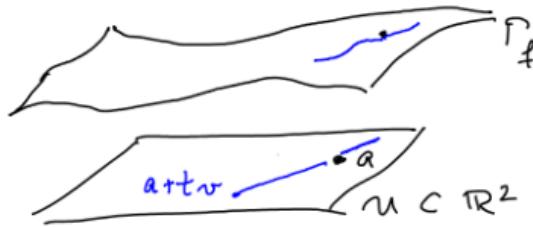


Ist dagegen  $h = (\varepsilon \cos \phi, \varepsilon \sin \phi)$  (von der Niveaumenge weg), so ist  $Df(a) \cdot h = 2r\varepsilon > 0$ . Das wollen wir im Folgenden systematisch studieren.

**Definition.** Sei  $f: U \rightarrow \mathbb{R}^m$ ,  $U \subset \mathbb{R}^n$  offen, gegeben. Für  $a \in U$  und  $v \in \mathbb{R}^n$  heißt der Grenzwert (falls er existiert)

$$\partial_v f(a) := \lim_{\substack{t \rightarrow 0 \\ \text{in } \mathbb{R}^m}} \frac{f(a + tv) - f(a)}{t}$$

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



**Bemerkungen / Beispiele 3.5.** i)  $\partial_v f(a) = \frac{d}{dt} \Big|_{t=0} g_{a,v}(t) = g'_{a,v}(0)$ ,  $g_{a,v}(t) = f(a + tv)$ .

ii)  $f = (f_1, \dots, f_m)^T$ , so gilt

$$\partial_v f(a) = (\partial_v f_1(a), \dots, \partial_v f_m(a))^T$$

iii)  $f: (x_1, x_2) \mapsto x_1^2 + x_2^2 + 2$ . Sei  $a = (r \cos \phi, r \sin \phi)$ . Ist  $v = (\varepsilon \cos \phi, \varepsilon \sin \phi)$ , so ist  
iv)

$$g_{a,v}(t) = ((r + \varepsilon t) \cos \phi)^2 + ((r + \varepsilon t) \sin \phi)^2 + 2$$

und somit  $\partial_v f(a) = g'_{a,v}(0) = 2(r + \varepsilon \cdot 0)\varepsilon = 2r\varepsilon$ .

Ist  $v = (-\varepsilon \sin \phi, \varepsilon \cos \phi)$ , so ist

$$g_{a,v}(t) = (r \cos \phi - \varepsilon t \sin \phi)^2 + (r \sin \phi + \varepsilon t \cos \phi)^2$$

und somit

$$\partial_v f(a) = g'_{a,v}(0) = 2(r \cos \phi - 0)(-\varepsilon \sin \phi) + 2(r \sin \phi + 0)(\varepsilon \cos \phi) = 0.$$

Das ist kein Zufall, denn es gilt

**Satz 3.6.** Sei  $f: U \rightarrow \mathbb{R}^m$ ,  $U \subset \mathbb{R}^n$  offen, in  $a \in U$  differenzierbar. Dann besitzt  $f$  die Richtungsableitungen

$$\partial_v f(a) = Df(a) \cdot v \quad \forall v \in \mathbb{R}^n$$

und  $Df(a)$  hat bezüglich der Standardbasis  $e_1, \dots, e_n$  die Matrix Darstellung

$$\begin{pmatrix} \partial_1 f_1(a) & \dots & \partial_n f_1(a) \\ \partial_1 f_2(a) & \dots & \partial_n f_2(a) \\ \vdots & & \vdots \\ \partial_1 f_m(a) & \dots & \partial_n f_m(a) \end{pmatrix} \in \text{Mat}(m \times n, \mathbb{R}),$$

„Jacobi-Matrix“, wobei  $\partial_j f(a) = \partial_{e_j} f(a) = Df(a) \cdot e_j = j\text{-te Spalte}$ .

*Beweis.* Für  $v = 0$  sind beide Seiten 0 ✓.

Für  $v \neq 0$  betrachte

$$\begin{aligned} & \| (f(a + tv) - f(a))/t - Df(a) \cdot v \|_{\sim} \text{ (Norm auf } \mathbb{R}^m) \\ &= \frac{1}{|t|} \| f(a + tv) - (f(a) + Df(a) \cdot (tv)) \|_{\sim} \\ &= \frac{1}{\|t \cdot v\|} \| f(a + tv) - (f(a) + Df(a) \cdot (tv)) \|_{\sim} \cdot \|v\|. \end{aligned}$$

↑  
Homogenität  
Norm auf  $\mathbb{R}^n$

Differenzierbarkeit  $\implies$  strebt gegen 0 für  $tv \rightarrow 0$ , also für  $t \rightarrow 0 \implies$  Die Richtungsableitungen existieren und

$$\partial_v f(a) = Df(a) \cdot v$$

und bezüglich der kanonischen Basen gilt

$$\begin{aligned} \partial_j f(a) &= Df(a) \cdot e_k = j\text{-te Spalte von } Df(a). & \square \\ & \left( \begin{array}{c} \partial_j f_1(a) \\ \vdots \\ \partial_j f_m(a) \end{array} \right) \end{aligned}$$

**Achtung.** Umgekehrt genügt die Existenz der Richtungsableitungen  $\partial_1 f, \dots, \partial_n f$  nicht, um Differenzierbarkeit zu garantieren!

**Beispiel.**

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

$\partial_1 f(x, y) = \frac{d}{dt} \Big|_{t=0} f(x + t, y), \partial_2 f(x, y) = \frac{d}{dt} \Big|_{t=0} f(x, y + t)$ . Wegen  $f(x, 0) = 0 \quad \forall x, f(0, y) = 0 \quad \forall y$ , ist  $\partial_1 f(0, 0) = 0 = \partial_2 f(0, 0)$ . Aber  $f$  ist in 0 nicht stetig (betrachte etwa  $(1/n, 1/n)$ ), also nicht differenzierbar.

**Vorlesung 7**

Mo 11.05. 10:15

**3.II. Beispiele und Erläuterungen**

Wir hatten letztes Mal gesehen, dass, wenn  $f: U \rightarrow \mathbb{R}^m$ ,  $U \subset \mathbb{R}^n$  offen, in  $a \in U$  differenzierbar ist, dass dann die Ableitung mit Hilfe der *partiellen Ableitungen*, also der Richtungsableitungen in Richtung der kanonischen Basis geschrieben werden kann,

$$Df(a) = \begin{pmatrix} \partial_1 f_1 & \dots & \partial_n f_1 \\ \vdots & & \vdots \\ \partial_1 f_m & \dots & \partial_n f_m \end{pmatrix},$$

dass aber die Existenz der partiellen Ableitungen nicht unbedingt Differenzierbarkeit garantiert.

Konkret kann man also so vorgehen: Man bestimmt die partiellen Ableitungen und überprüft dann, ob Differenzierbarkeit vorliegt.

**Bemerkung 3.7.** Berechnung von partiellen Ableitungen. Es gilt  $\partial_j f(a) = g'_{a_0 e_j}(0)$ , wobei

$$\begin{aligned} g_{a_i e_j}(t) &= f(a + te_j) = f(a_1, \dots, a_j + t, \dots, a_n) \quad t \in (-\nu, \varepsilon) \\ &\text{s. d. } a + te_j \text{ noch im Definitionsbereich von } f \text{ liegt.} \end{aligned}$$

⇒ Um  $\partial_j f(a)$  zu berechnen, kann man die gewöhnlichen Ableitungen bezüglich der  $j$ -ten Koordinate bestimmen (und steht stellt sich die Übrigen als Konstanten vor).

**Beispiele 3.8.** i)  $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ ,  $f(x) = x_1^2 + x_2^2 + 2$ .

$$\partial_1 f(a) = 2a_1 \quad \partial_2 f(a) = 2a_2$$

$f$  ist in der Tat differenzierbar in allen  $a \in \mathbb{R}^2$ , denn

$$\begin{aligned} f(a + h) &= (a_1 + h_1)^2 + (a_2 + h_2)^2 + 2 \\ &= f(a) + A \cdot h + \underbrace{\|h\|_E^2}_{R_a(h)}. \end{aligned}$$

mit  $A = \begin{pmatrix} 2a_1 & 2a_2 \end{pmatrix}$  und

$$\frac{R_a(h)}{\|h\|} \leq C \|h\| \rightarrow 0.$$

ii)  $f: \mathbb{R}^n \rightarrow \mathbb{R}$ ,  $n \geq 2$ ,  $f(x) = \|x\|_{\text{E}} = \sqrt{\sum_{i=1}^n x_i^2}$ .

$a \neq 0$  Mittels Kettenregel aus DIFF I:

$$\partial_j f(a) = \frac{1}{2} \frac{1}{\|x\|_{\text{E}}} \cdot 2x_j \Big|_{x=a} = \frac{a_j}{\|a\|_{\text{E}}}$$

$$a = 0$$

$$\partial_j f(0) = \frac{d}{dt} f(0, \dots, \underset{j-\text{te}}{t}, \dots, 0) = \frac{d}{dt} \Big|_{t=0} |t|.$$

$\implies$  Die partiellen Ableitungen in  $a = 0$  existieren nicht (DIFF I: Für  $h \in \mathbb{R}$  ist  $|h| - 0/h = \pm 1 \not\rightarrow 0$  für  $h \searrow 0$  bzw.  $h \nearrow 0$ ).  $\implies f$  ist *nicht* differenzierbar in  $a = 0$ .

Für  $a \neq 0$  gilt jedoch:  $\sqrt{\cdot}: \mathbb{R}_{>0} \rightarrow \mathbb{R}$  ist differenzierbar (DIFF I) und ebenso die polynomiale Funktion  $x \mapsto \sum_{i=1}^n x_i^2 \implies f$  ist differenzierbar auf  $\mathbb{R}^n \setminus \{0\}$ .

Wir hätten  $Df(a)$ ,  $a \neq 0$  auch mit der höherdimensionalen Kettenregel bestimmen können:

$$Df(a) = Dw(p(a)) \cdot Dp(a) = \frac{1}{2\sqrt{p(a)}} \cdot (2a_1 \dots 2a_n)$$

$$w: \mathbb{R}_{>0} \rightarrow \mathbb{R}, \quad w(t) = \sqrt{t}, \quad p(a) = \sum_{i=1}^n a_i^2.$$

iii) Eine weitere Anwendung der Kettenregel. Betrachte die „Polarkoordinaten“

$$g: \mathbb{R}_{>0} \times \mathbb{R} \rightarrow \mathbb{R}^2$$

$$g(r, \phi) = (r \cos \phi, r \sin \phi)^T$$



und  $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ .

$$D(f \circ g)(r, \phi) = Df(g(r, \phi)) \cdot Dg(r, \phi)$$

$$Dg(r, \phi) = \begin{pmatrix} \partial_1 g(r, \phi) & \partial_2 g(r, \phi) \end{pmatrix} = \begin{pmatrix} \cos \phi & -r \sin \phi \\ \sin \phi & r \cos \phi \end{pmatrix}.$$

Somit

$$\begin{aligned} D(f \circ g)(r, \phi) &= \begin{pmatrix} \partial_1 f(g(r, \phi)) & \partial_2 f(g(r, \phi)) \end{pmatrix} \cdot \begin{pmatrix} \cos \phi & -r \sin \phi \\ \sin \phi & r \cos \phi \end{pmatrix} \\ &= \begin{pmatrix} \partial_1 f(g(r, \phi)) \cdot \cos \phi + \partial_2 f(g(r, \phi)) \cdot \sin \phi \\ -\partial_1 f(g(r, \phi)) \cdot r \sin \phi + \partial_2 f(g(r, \phi)) \cdot r \cos \phi \end{pmatrix}^T. \end{aligned}$$

Man schreibt dafür manchmal

$$\begin{aligned} \partial_r &= \cos \phi \partial_x + \sin \phi \partial_y \\ \partial_\phi &= -r \sin \phi \partial_x + r \cos \phi \partial_y. \end{aligned}$$

Aber oft ist es *viel* übersichtlicher, die partiellen Ableitungen *nicht* nach den Namen der Variablen zu benennen, sondern durchzunummerieren!

**Beispiel.**  $f(x, y) = \frac{xy}{x^2+y^2}$  auf  $U = \mathbb{R}^2 \setminus \{0\}$ .

$$\begin{aligned} Df(x, y) &= \begin{pmatrix} \frac{y(y^2)-x^2}{(x^2+y^2)^2} \\ \frac{x(x^2-y^2)}{(x^2+y^2)^2} \end{pmatrix} \\ D(f \circ g)(r, \phi) &= \frac{1}{r} (\cos^2 \phi - \sin^2 \phi) \begin{pmatrix} \overbrace{-\sin \phi \cos \phi + \cos \phi \sin \phi}^{=0} \\ \underbrace{r \sin^2 \phi + r \cos^2 \phi}_{=r} \end{pmatrix} \\ &= (0, \cos^2 \phi - \sin^2 \phi). \end{aligned}$$

In diesem Fall rechnet man allerdings schneller direkt:

$$D(f \circ g)(r, \phi) = D(\tilde{f})(r, \phi) = \cos \phi \cdot \sin \phi.$$

**Bemerkung / Definition 3.9.** Wir hatten allgemeiner gesehen:

Ist  $f: U \rightarrow \mathbb{R}^m$ ,  $U \subset \mathbb{R}^n$  offen, differenzierbar in  $a$ , so gilt

$$\partial_v f(a) \cdot Df(a) \cdot v \quad \forall v \in \mathbb{R}^n.$$

Ist speziell  $m = 1$ , so definiert man

$$\text{grad } f(a) := Df(a)^T = \begin{pmatrix} \partial_1 f(a) \\ \vdots \\ \partial_n f(a) \end{pmatrix} \in \mathbb{R}^n$$

„Gradient von  $f$  in  $a$ “, und schreibt  $\partial_v f(a) = \langle \text{grad } f(a), v \rangle$ .

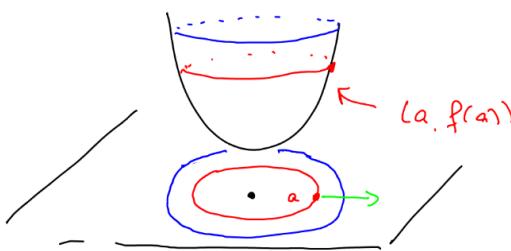
Ist  $\text{grad } f(a) \neq 0$  und  $\|v\|_E = 1$ , so ist

$$\langle \text{grad } f(a), v \rangle = \|\text{grad } f(a)\|_E \cdot \cos \alpha,$$

wobei  $\alpha$  der zwischen  $\text{grad } f(a)$  und  $v$  in  $\mathbb{R}^n$  eingeschlossene Winkel ist (in der durch die beiden Vektoren aufgespannten Ebene).

Es folgt:  $\partial_v f(a)$  ist dann am größten, wenn  $v$  in die selbe Richtung zeigt wie  $\text{grad } f(a)$   
 $\Rightarrow$  der Gradient gibt die Richtung stärksten Anstiegs von  $f$  in  $a$  an.

**Beispiel.**  $f(x) = x_1^2 + x_2^2 + 2$ .



**Beispiel.**  $f(x) = x^2$ .



Ist  $f$  lediglich *partiell differenzierbar*, d. h. die partiellen Ableitungen existieren auf  $U$ , so definiert man dennoch

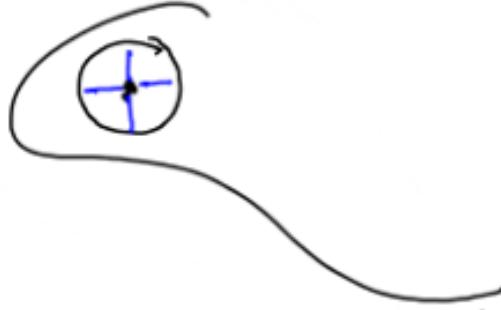
$$\text{grad } f(x) = (\partial_1 f(x), \dots, \partial_n f(x))^T$$

als den Vektor der partiellen Ableitungen bei  $x$ .

**Satz 3.10.** Sei  $f: U \rightarrow \mathbb{R}$ ,  $U \subset \mathbb{R}^n$  offen, partiell differenzierbar auf  $U$ . Sei  $a \in U$  ein lokales Maximum (oder Minimum) von  $f$ , d. h.  $\exists$  Umgebung  $V$  von  $a$  s. d.  $f(x) \leq f(a)$  (oder  $f(x) \geq f(a)$  für alle  $x \in V$ ). Dann gilt

$$\boxed{\text{grad } f(a) = 0}.$$

*Beweis.* Betrachte  $g_i(t) := f(a + te_i)$ ,  $i = 1, \dots, n$ , mit  $t \in (-\varepsilon, \varepsilon)$  s. d.  $B_\varepsilon^{\|\cdot\|}(a) \subset U$ .



Ist  $a$  lokales Extremum (also lokales Maximum oder Minimum) von  $f$ , so ist  $t = 0$  lokales Extremum von  $g_i$ . Die  $g_i$  sind in  $t = 0$  differenzierbar, denn

$$g'_i(0) = \partial_i f(a) \quad (\text{Definition Richtungsableitung})$$

DIFF I  $\implies g'_i(0) = 0 \implies$  Beh. □

**Beispiele.** i)  $f(x) = x_1^2 + x_2^2 + 2$ . Außer in  $a = 0$  kann kein Extremum vorliegen.

ii)  $f(x, y) = \begin{cases} \frac{xy}{x^2+y^2} & (x, y) \neq (0, 0) \\ 0 & (x, y) = (0, 0) \end{cases}$ .  $\partial_1 f(0) = 0 = \partial_2 f(0)$ , also könnte 0 ein Extremum sein. Ist es aber nicht, da für  $x = y = \varepsilon > 0$  gilt  $f(\varepsilon, \varepsilon) = \frac{1}{2} > 0$  und für  $-x = y = \varepsilon > 0$ ,  $f(-\varepsilon, \varepsilon) = -\frac{1}{2} < 0$ .

**Bemerkung.** Hinreichende Kriterien für das Vorliegen lokaler Extremstellen werden wir erst später kennen lernen. Wie in der DIFF I benötigen wir dafür die 2. Ableitung.

**Satz 3.11.** Sei  $f: U \rightarrow \mathbb{R}$ ,  $U \subset \mathbb{R}^n$  offen. Existieren alle partiellen Ableitungen  $\partial_j f(x)$  für alle  $x \in U$ , und sind sie stetig in  $a \in U$ , so ist  $f$  in  $a$  differenzierbar.

*Beweis.* Wir wählen ein Norm  $\|\cdot\|$  auf  $\mathbb{R}^n$ .  $U$  offen  $\implies \exists \delta > 0$  s. d.  $B_\delta^{\|\cdot\|}(a) \subset U$ . Sei  $h \in B_\delta^{\|\cdot\|}(0)$ , also  $a + h \in B_\delta^{\|\cdot\|}(a)$ . Setze

$$x^{(j)} := a + \sum_{i=1}^j h_i e_i \in \mathbb{R}^n \quad j = 0, \dots, n,$$

also  $x^{(0)} = a$ ,  $x^{(1)} = (a_1 + h_1, a_2, \dots, a_n), \dots, x^{(n)} = a + h$ . Es ist  $x^{(j)} - x^{(j-1)} = (0, \dots, 0, h_j, 0, \dots, 0) \implies (\text{MWS, DIFF I}) \exists \eta_j \in [0, 1] \text{ s. d.}$

$$\begin{aligned} f(x) &= \partial_j f(\underbrace{x_1^{(j-1)}, \dots, x_{j-1}^{(j-1)}, x_j^{(j-1)} + \eta_j h_j, a_{j+1}, \dots, a_n}_{x^{j-1} + \eta_j h_j e_j =: y^{(j)}}) \cdot h_j \\ \implies f(a+h) - f(a) &= \sum_{j=1}^n (f(x^{(j)}) - f(x^{(j-1)})) \\ &= \sum_{j=1}^n \partial_j f(y^{(j)}) h_j \\ &\stackrel{!}{=} \left( \partial_1 f(a) \dots \partial_n f(a) \right) \begin{pmatrix} h_1 \\ \vdots \\ h_n \end{pmatrix} + \underline{R}_a(h) \end{aligned}$$

also ist

$$\begin{aligned} \underline{R}_a(h) &= \sum_{j=1}^n (\partial_j f(y^{(j)}) - A_j) h_j \quad A = \partial_j f(a) \\ \stackrel{\text{CS}}{\implies} \frac{|\underline{R}_a(h)|}{\|h\|} &\leq C \|(\partial_1 f(y^{(1)}) - A_1, \dots, \partial_n f(y^{(n)}) - A_n)\|_{\mathbb{E}} \rightarrow 0 \quad h \rightarrow 0, \end{aligned}$$

denn  $\lim_{h \rightarrow 0} y^{(j)} = \lim_{h \rightarrow 0} x^{(j-1)} + \eta_j h_j e_j = a$  (in  $\mathbb{R}^n$ ) und die  $\partial_j f$  sind in  $a$  stetig nach Voraussetzung, s. d.

$$\lim_{h \rightarrow 0} \partial_j f(y^{(j)}) = (\partial_j f)(\lim y^{(j)}) = \partial_j f(a). \quad \square$$

**Bemerkungen 3.12.** i) Man sieht: Sind die  $\partial_j$  auf  $U$  stetig, so ist auch die Ableitung  $x \mapsto Df(x)$  eine stetige Abbildung  $U \rightarrow \text{Mat}(m \times n, \mathbb{R})$ . Man sagt in dem Fall:  $f$  ist *stetig differenzierbar*.

ii) Die Einschränkung auf reellwertige Funktionen ist keine, denn:  $f: U \rightarrow \mathbb{R}^m$  ist genau dann in  $a \in U \subset \mathbb{R}^n$  differenzierbar, wenn alle Komponentenfunktionen  $f_j: U \rightarrow \mathbb{R}$  differenzierbar sind ( $j = 1, \dots, m$ )-

*Beweis.* 3.12.ii)  $f(a+h) = f(a) + A \cdot h + \underline{R}(h)$ ,  $A_{ji} = \partial_j f_j(a)$

$$\iff f_j(a+h) = f_j(a) + (\partial_1 f_j(a) \dots \partial_n f_j(a)) \cdot h + \underline{R}_j(h) \quad \forall j = 1, \dots, m$$

und

$$\frac{\underline{R}(h)}{\|h\|} \rightarrow 0 \text{ in } \mathbb{R}^m \iff \frac{\|\underline{R}(h)\|_{\max}}{\|h\|} \rightarrow 0 \iff \frac{|R|_j(h)}{\|h\|} \rightarrow 0 \quad j = 1, \dots, m.$$

3.12.i) Dito:  $U \rightarrow \text{Mat}(m \times n, \mathbb{R})$  ist stetig  $\iff$  alle Komponentenfunktionen  $A_{ij}: U \rightarrow \mathbb{R}$  sind stetig.  $\square$

**Bemerkung.** Stetig differenzierbar  $\implies$  differenzierbar  $\implies$  partiell differenzierbar.

Die Umkehrungen sind im Allgemeinen falsch.

**Bemerkung (Erinnerung (DIFF I) MWS).**  $f: I \rightarrow \mathbb{R}$  differenzierbar,  $I$  Intervall,  $\exists \epsilon \in [0, 1]$  s. d.

$$f(a + h) - f(a) = f'(a + \eta h) \cdot h.$$

Ist  $f$  stetig differenzierbar, folgt aus dem Hauptsatz der Differenzial- und Integralrechnung eine andere Variante:

$$f(a + h) - f(a) = \int_a^{a+h} f'(u) du = \int_0^1 f'(a + th) dt \cdot h.$$

Eine analoge Aussage wollen wir nun im  $\mathbb{R}^n$  beweisen.

**Definition.** Sei  $A: I \rightarrow \mathbb{R}^k, I \subset \mathbb{R}$  Intervall, stetig. Dann ist das *Integral von A* über  $[a, b] \in I$  definiert als

$$\int_a^b A(t) dt = \begin{pmatrix} \int_a^b A_1(t) dt \\ \vdots \\ \int_a^b A_k(t) dt \end{pmatrix}.$$

Insbesondere ist das Integral einer matrixwertigen, stetigen Funktion  $A: I \rightarrow \text{Mat}(m \times n, \mathbb{R})$  die Matrix, deren Einträge gleich den Integralen der Komponenten von  $A(t)$  ist, also

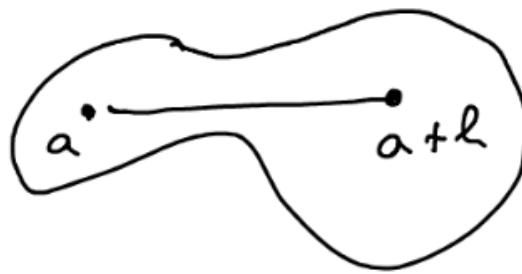
$$\int_a^b A(t) dt = \left( \int_a^b A_{ij}(t) dt \right)_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}}.$$

**Satz 3.13.** Sei  $f: U \rightarrow \mathbb{R}^m, U \subset \mathbb{R}^n$ , stetig differenzierbar. Sei  $a \in U$  und  $h \in \mathbb{R}^n$  s. d. die Verbindungsstrecke

$$\{ a + th \mid t \in [0, 1] \} \subset U.$$

Dann gilt:

$$f(a + h) - f(a) = \left( \int_0^1 Df(a + th) dt \right) \cdot h.$$



*Beweis.* Wir arbeiten zeilenweise, betrachten also die Komponentenfunktionen. Setze  $g_i(t) := f_i(a + th)$ . Dann ist  $g_i$  stetig differenzierbar, denn mit der Kettenregel gilt  $g'_i(t) = Df_i(a + th) \cdot h$ .

$$\begin{aligned}
 f_i(a + h) - f_i(a) &= g_i(1) - g_i(0) \\
 &= \int_0^1 g'_i(t) dt \\
 &= \int_0^1 Df_i(a + th) \cdot h dt \\
 &= \underbrace{\left( \int_0^1 Df_i(a + th) dt \right)}_{\int \text{ linear}} \cdot h \quad 1 \leq i \leq m.
 \end{aligned}$$

Dies sind die Zeilen der Matrix  $\int_0^1 Df(a + th) dt$ . □

**Folgerung 3.14.** Unter den Voraussetzungen von 3.13 gilt  $\|f(x + h) - f(x)\|_E \leq C\|h\|$  mit

$$C = \sup_{t \in [0,1]} \|Df(x + th)\|_{\text{op.}}$$

Der Beweis benötigt noch ein Lemma:

**Lemma 3.15.** Sei  $v: [a, b] \rightarrow \mathbb{R}^m$  stetig. Dann gilt

$$\left\| \int_a^b v(t) dt \right\|_E \leq \int_a^b \|v(t)\|_E dt.$$

*Beweis.* Sei  $\mathbb{R}^m \ni u = \int_a^b v(t) dt$ . Dann gilt

$$\begin{aligned}
\|u\|_E^2 &= \langle u, u \rangle \\
&= \left\langle \int_a^b v(t) dt, u \right\rangle \\
&= \int_a^b \langle v(t), u \rangle dt \\
&\quad \uparrow \text{Linearität} \\
&\leq \int_a^b \|v(t)\|_E \|u\|_E dt \\
&\quad \uparrow \text{Monotonie und C-S} \\
&= \|u\|_E \int_a^b \|v(t)\|_E dt.
\end{aligned}$$
□

**Bemerkungen.** i) Wegen der Äquivalenz aller Normen auf  $\mathbb{R}^n$  gilt diese Abschätzung ebenso wie Folgerung 3.14 auch für beliebige Normen auf  $\mathbb{R}^n$ :

$$\begin{aligned}
\left\| \int v(t) dt \right\| &\leq C_1 \left\| \int v(t) dt \right\|_E \\
&\leq C_1 \int \|v(t)\|_E dt \\
&\leq C_1 C_2 \int \|v(t)\| dt.
\end{aligned}$$

ii) Es folgt, dass für

$$X = (C([a, b], \mathbb{R}^k), \|\cdot\|_\infty),$$

$\sup\|v(t)\| < \infty$ , da  $[a, b]$  kompakt und  $v$  stetig.

$f_a^b: X \rightarrow \mathbb{R}^k$  stetig ist, denn  $f_a^b$  ist linear und

$$\begin{aligned}
\left\| \int_a^b \right\|_{\text{op}} &= \sup_{0 \neq v \in C} \left\| \int_a^b v(t) dt \right\| / \|v\|_\infty \\
&\leq C \sup_{v \neq 0} \left\| \int_a^b v(t) dt \right\|_E / \|v\|_\infty \\
&\leq C \sup_{v \neq 0} \int_a^b \|v(t)\|_E dt / \|v\|_\infty \\
&\leq \tilde{C}(b - a),
\end{aligned}$$

denn  $\|v(t)\|_E \leq \sup_t \|v(t)\|_E \leq C_0 \sup_t \|v(t)\| = C_0 \|v\|_\infty$ .

Beweis von Folgerung 3.14.

$$\begin{aligned}\|f(x+h) - f(x)\|_{\mathbb{E}} &\leq \int_0^1 \|Df(x+th) \cdot h\|_{\mathbb{E}} dt \\ &\leq \int_0^1 \|Df(x+th)\|_{\text{op}} \cdot \|h\|_{\mathbb{E}} dt.\end{aligned}$$

Das Supremum wird angenommen, da  $t \mapsto \|Df(x+th)\|_{\text{op}}$  stetig ist und  $[0, 1]$  kompakt.  $\square$

### 3.III. Implizite Funktionen

Eine Anwendung der Kettenregel

Manchmal ist es einfacher, eine 1-dimensionale Ableitung über eine Ableitung einer höherdimensionalen Funktion zu berechnen. Betrachte etwa zwei differenzierbare Funktionen  $f: U \rightarrow \mathbb{R}$ ,  $U \subset \mathbb{R}^2$  offen, und  $g: I \rightarrow \mathbb{R}$ . Sei  $\Gamma_g \subset U$  und es gelte

$$\underbrace{f(t, g(t))}_{\in \Gamma_g} = 0 \quad \forall t \in I.$$

Dann gilt (Kettenregel!) (mit  $\text{Id}(t) = t$ ):

$$\begin{aligned}0 &= D * f \circ \begin{pmatrix} \text{Id} \\ g \end{pmatrix}(t) \\ &= Df((t, g(t))) \cdot D \begin{pmatrix} \text{Id} \\ g \end{pmatrix}(t) \\ &= (\partial_1 f(t, g(t)) \quad \partial_2 f(t, g(t))) \cdot \begin{pmatrix} 1 \\ g'(t) \end{pmatrix}\end{aligned}$$

$\implies$  Ist  $\partial_2 f(t, g(t)) \neq 0$ , so gilt

$$g'(t) = -\frac{\partial_1 f(t, g(t))}{\partial_2 f(t, g(t))}.$$

**Beispiel.**  $g: (0, 1) \rightarrow \mathbb{R}$ ,  $g(t) = \arcsin \sqrt{1-t^3}$ , also  $\sin g(t) = \sqrt{1-t^3}$ , also  $(\sin g(t))^2 =$

$1 - t^3$ . Also gilt  $f(t, g(t)) = 0$  für

$$\begin{aligned} f(x, y) &= (\sin y)^2 - 1 + x^3 \\ Df(x, y) &= (3x^2, 2 \sin y \cos y) \\ \implies g'(t) &= -\frac{3t^2}{2 \underbrace{\sin g(t) \cos(g(t))}_{>0}} \end{aligned}$$

da  $t \in (0, 1)$  und  $\sqrt{1 - t^3} \in (0, 1)$ , somit  $g(t) \in (0, \pi/2) \implies \cos g(t) = +\sqrt{1 - \underbrace{\sin^2 g(t)}_{1-t^3}} = \sqrt{t^3}$

$$= -\frac{3t^2}{2\sqrt{1-t^3}\sqrt{t^3}}.$$

**Vorlesung 8**

Do 14.05. 10:15

**Satz 3.16.** Sei  $f: B_{r_1}^{\|\cdot\|}(a) \times B_{r_2}^{\|\cdot\|}(b) \rightarrow \mathbb{R}^m$  eine Abbildung mit  $f(a, b) = 0$ , die in  $(a, b)$  differenzierbar sei. Sei zudem  $D_2 f(a, b)$ , definiert über

$$Df(a, b) = (\underbrace{\partial_1 f(a, b) \cdots \partial_k f(a, b)}_{=: D_1 f(a, b) \in \text{Mat}(m \times k)} \underbrace{\partial_{k+1} f(a, b) \cdots \partial_{k+m} f(a, b)}_{=: D_2 f(a, b) \in \text{Mat}(m \times m)}),$$

invertierbar. Sei zudem  $g: B_{r_1}(a) \rightarrow \mathbb{R}^m$  stetig und gelte  $g(B_{r_1}(a)) \subset B_{r_2}(b)$  und  $g(a) = b$  und  $f(x, g(x)) = 0 \forall x \in B_{r_1}(a)$ . Dann ist  $g$  in  $a$  differenzierbar und es gilt

$$Dg(a) = -(D_2 f(a, b))^{-1} D_1 f(a, b) \in \text{Mat}(m \times k, \mathbb{R}).$$

*Beweis.* Es gilt für  $h = (h_1, h_2) \in B_{r_1}(0) \times B_{r_2}(0)$

$$f(\underbrace{(a, b) + h}_{=(a+h_1, b+h_2)}) = 0 + D_1 f(a, b) \cdot h_1 + D_2 f(a, b) \cdot h_2 + \underline{R}_{(a,b)}(h)$$

mit  $\frac{\underline{R}_{(a,b)}(h)}{\|h\|} \rightarrow 0 \forall x \in B_{r_1}(a)$  folgt

$$\begin{aligned} 0 &= f(\underbrace{a + h_1}_{\in B_{r_1}(a)}, g(a + h_1)) = D_1 f(a, b) \cdot h_1 \\ &\quad + D_2 f(a, b) \cdot (g(a + h_1) - b) \\ &\quad + \underline{R}_{(a,b)}(h_1, g(a + h_1) - b) \quad \forall h_1 \in B_{r_1}(0). \end{aligned}$$

Also

$$\begin{aligned} g(a + h_1) &= g(a) - D_2 f(a, b)^{-1} \cdot D_1 f(a, b) \cdot h_1 \\ &\quad - \underbrace{D_2 f(a, b)^{-1} \cdot \underline{R}_{(a,b)}(h_1, g(a + h_1) - b)}_{=: \tilde{R}_{a,b}^g(h_1)} \quad \forall h_1 \in B_{r_1}(0). \end{aligned}$$

Wir sind fertig, wenn wir zeigen können, dass

$$\frac{\tilde{R}_{a,b}^g(h_1)}{\|h_1\|} \rightarrow 0 \text{ für } h_1 \rightarrow 0.$$

Wegen der Differenzierbarkeit von  $f$  in  $(a, b)$   $\exists \tilde{C}$  und  $\delta_1, \delta_2 > 0$ ,  $\delta_i < r_i$ , s. d.

$$\|\underline{R}_{(a,b)}(h_1, h_2)\| \leq \tilde{C}(\|h_1\| + \|h_2\|) \quad \forall h_1 \in B_{\delta_1}(0), h_2 \in B_{\delta_2}(0),$$

also

$$\|\underline{R}_{(a,b)}(h_1, g(a+h_1) - b)\| \leq \tilde{C}(\|h_1\| + \|g(a+h_1) - b\|)$$

für alle  $h_1$  s. d.  $\|h_1\| < \delta_1$  und  $\|g(a+h_1) - b\| < \delta_2$ . Wegen der Stetigkeit von  $g$  in  $a$  gibt es  $\delta > 0, \delta < \delta_1$ ,  $\|\underbrace{g(a+h_1) - g(a)}_{=b}\| < \delta_2$  für alle  $h_1 \in B_\delta(0)$ .  $\square$

- Bemerkungen.**
- i) Ist  $f, g$  wie in Satz 3.16 und  $f$  überall *stetig* differenzierbar. Ist dann  $D_2f(a, b)$  invertierbar. So gibt es eine Umgebung  $V_1 \times V_2$  von  $(a, b)$  s. d.  $D_2f(x, y)$  invertierbar ist für alle  $(x, y) \in V_1 \times V_2$  (denn  $(x, y) \mapsto \det D_2f(x, y)$  ist stetig, da es ein Polynom stetiger Funktionen ist und  $\det D_2f(a, b) \neq 0$ ).
  - ii) Das Beispiel vom letzten Mal ist genau von diesem Typ.

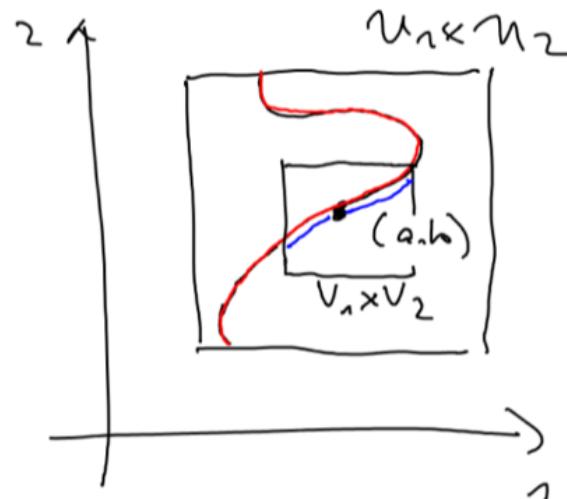
Umgekehrt kann man eine Funktion  $g$  wie oben durch „Auflösen der Gleichung  $f(x, g(x)) = 0$ “ bestimmen (unter gewissen Voraussetzungen):

**Satz 3.17 (Satz von der impliziten Funktion).** Seien  $U_1 \subset \mathbb{R}^k$ ,  $U_2 \subset \mathbb{R}^m$  offen und  $f: U_1 \times U_2 \rightarrow \mathbb{R}^m$  stetig differenzierbar. Sei  $(a, b) \in U_1 \times U_2$  s. d.  $f(a, b) = 0$  und  $D_2f(a, b) \in \text{Mat}(m \times m, \mathbb{R})$  invertierbar. Dann gibt es offene Umgebungen  $V_1 \subset U_1$ ,  $V_2 \subset U_2$  von  $a$  bzw.  $b$  und eine stetige Funktion  $g: V_1 \rightarrow \mathbb{R}^m$ ,  $g(V_1) \subset V_2$ , s. d.  $f(x, g(x)) = 0 \forall x \in V_1$ . Ist  $(x, y) \in V_1 \times V_2$  s. d.  $f(x, y) = 0$ , so ist  $y = g(x)$ .

- Bemerkungen.**
- i) Aus 3.16 und der folgenden Bemerkung folgt, dass  $g$  in einer eventuell verkleinerten Umgebung  $\tilde{V}_1 \subset V_1$  von  $a$  sogar stetig differenzierbar ist und gilt

$$Dg(x) = -D_2f(x, g(x))^{-1} \cdot D_1f(x, g(x)) \quad \forall x \in \tilde{V}_1$$

- ii) Für den Satz ist wichtig, dass  $U_1 \times U_2$  gegebenenfalls verkleinert wird:

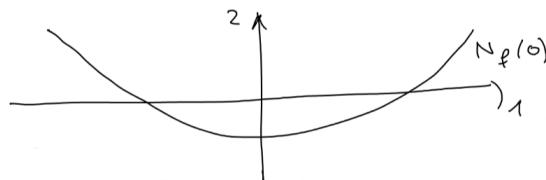


$$\Gamma_g = \{(x, g(x)) \mid x \in V_1\},$$

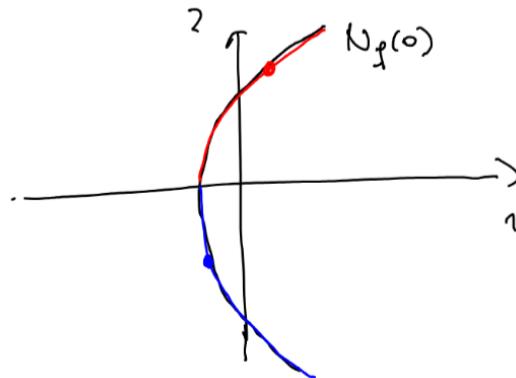
Betrachtete man auch den oberen Teil der Kurve, könnte man  $x$  nicht ein eindeutiges  $y$  zuordnen.

- iii) Die Einschränkung auf Definitionsbeiche der Form  $U_1 \times U_2$  ist keine, sie vereinfacht nur die Notation. Ist  $f$  auf  $U \subset \mathbb{R}^{k+m}$  offen, findet man stets  $U_1 \subset \mathbb{R}^k$ ,  $U_2 \subset \mathbb{R}^m$  s. d.  $U_1 \times U_2 \subset U$ .
- iv) Die Einschränkung auf  $N_f(0)$  ist keine: Will man etwa die Gleichung  $f(x, y) = c$  auflösen, wendet man des Satz auf  $\tilde{f}(x, y) = f(x, y) - c$  an.
- v) Durch Umnummerierung kann man auch andere  $m \times m$ -Untermatrizen von  $Df$  betrachten als die letzten  $m$ .
- vi) Unter den Voraussetzungen von 3.17 sagt man:  $g$  ist durch  $f(x, y) = 0$  implizit gegeben und man löst  $f(x, y) = 0$  nach  $y$  auf.

**Beispiele.** i)  $f(x, y) = 3y - x^2 + 1$  auf  $\mathbb{R}^2$ .  $Df(a, b) = \begin{pmatrix} -2a & 3 \end{pmatrix}$ ,  $3 \neq 0$  ist invertierbar, 3.17  $\implies \exists g: I \rightarrow \mathbb{R}$  s. d.  $f(x, g(x)) = 0 \forall x \in I$ . In diesem Fall sogar  $I = \mathbb{R}$ :  $g(x) = \frac{1}{3}(x^2 - 1)$ .



ii)  $f(x, y) = 3x - y^2 + 1$  auf  $\mathbb{R}^2$ .  $Df(a, b) = \begin{pmatrix} 3 & -2b \end{pmatrix}$ . 3.17  $\Rightarrow$  Zu  $b \neq 0$  gibt es  $g: I \rightarrow \mathbb{R}$ .  $b > 0$ :  $g(x) = +\sqrt{3x+1}$ ,  $x > -\frac{1}{3}$ .  $b < 0$ :  $g(x) = -\sqrt{3x+1}$ .



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

*Beweis von Satz 3.17.* Setze  $B := D_2 f(a, b)$  und definiere eine Abbildung  $h: U_1 \times U_2 \rightarrow \mathbb{R}^m$  vermöge

$$h(x, y) = y - b - B^{-1} f(x, y).$$

Dann gilt

$$D_2 h(x, y) = \mathbb{1} - B^{-1} D_2 f(x, y).$$

$\Rightarrow D_2 h(a, b) = 0 \Rightarrow$  (da alle Ableitungen stetig sind)  $\exists W_1 \subset U_1$ ,  $W_2 \subset U_2$  offene Umgebungen von  $a$  bzw.  $b$  s. d.

$$\|D_2 h(x, y)\| \leq \frac{1}{2} \quad \forall x \in W_1, \quad y \in W_2. \quad (*)$$

Wähle  $r > 0$  s. d.  $V_2 := B_r^{\|\cdot\|}(b) \subset W_2$ . Es ist  $h(a, b) = 0 \Rightarrow$  (da  $h$  differenzierbar ist und somit auch stetig)  $\exists$  offene Umgebung  $V_1 \subset W_1$  von  $a$  s. d.

$$\varepsilon := \sup_{x \in V_1} \|h(x, b)\| < \frac{r}{2} \quad (**)$$

(auf einem Kompaktum  $\subset W_1$  um  $a$  ist  $x \mapsto h(x, b)$  beschränkt und wird auf einem hinreichend kleinen Kompaktum beliebig klein. Um  $V_1$  offen zu erhalten, nehmen wir das Innere eines solchen Kompaktums).

Wir zeigen jetzt: Zu jedem  $x \in V_1$  gibt es höchstens ein  $y \in V_2$  s. d.  $f(x, y) = 0$  also s. d.  $h(x, y) = y - b$ .

Sei also  $x \in V_1$  und seien  $y_1$  und  $y_2$  s. d.  $h(x, y_1) = y_1 - b$  und  $h(x, y_2) = y_2 - b$ .

$$\begin{aligned}
&\implies y_1 - y_2 = h(x, y_1) - h(x, y_2) \\
&\stackrel{\text{MWS und } (*)}{\implies} \|y_1 - y_2\| = \|h(x, y_1) - h(x, y_2)\| \\
&\quad = \|D_2 h(x, \zeta)\| \cdot \|y_1 - y_2\| \\
&\quad \zeta \text{ auf der Verbindungsstrecke zw. } y_1 \text{ und } y_2 \text{ (liegt in } V_2 = B_r(b)) \\
&\quad \leq \frac{1}{2} \|y_1 - y_2\| \\
&\implies \|y_1 - y_2\| = 0 \implies y_1 = y_2.
\end{aligned}$$

1) 2) Wir zeigen nun die Existenz einer Funktion  $g$  wie im Satz behauptet. Setze dazu  $g_0(x) = b$  und definiere rekursiv für  $x \in V_1$ :

$$g_{j+1}(x) := b + h(x, g_j(x)).$$

a) Es gilt

$$\|g_{j+1} - g_j\|_{\infty, V_1} \leq 2^{-j} \varepsilon.$$

↑  
aus (\*\*)

Induktionsanfang:

$$\|g_1 - g_0\|_{\infty, V_1} = \|h(x, b)\|_{\infty, V_1} = \varepsilon.$$

Induktionsschritt: Sei die Behauptung für  $i \leq n$  bewiesen.

$$\begin{aligned}
&g_{n+2}(x) - g_{n+1}(x) = h(x, g_{n+1}(x)) - h(x, g_n(x)) \\
&\stackrel{\text{MWS und } (*)}{\implies} \|g_{n+2} - g_{n+1}\|_{\infty, V_1} \leq \frac{1}{2} \|g_{n+1} - g_n\|_{\infty, V_1}.
\end{aligned}$$

**Bemerkung.** Der MWS darf tatsächlich angewendet werden.  $g_{n+1}(x), g_n(x)$  und somit auch die Verbindungsstrecke zwischen ihnen liegen in  $V_2$ , denn nach Induktionsvoraussetzung gilt für alle  $j \leq n$

$$\|g_{j+1} - b\|_{\infty, V_1} \leq \sum_{i=0}^j \|g_{i+1} - g_i\| \leq 2\varepsilon < r$$

(da  $g_{j+1} - b = \sum_{i=0}^j (g_{i+1} - g_i)$  ist). Somit darf der MWS auf

$$h(x, g_{n+1}(x)) - h(x, g_n(x))$$

angewendet werden.

b) Es folgt  $\|g_n - b\|_{\infty, V_1} < r$  und somit  $g_n(V_1) \subset V_2$ . Denn

$$g_n = \sum_{j=0}^{n-1} (g_{j+1} - g_j) + b \quad (\text{Teleskopsumme}) \implies \|g_n - b\|_{\infty, V_1} \leq \sum_{j=0}^{n-1} 2^{-j} \varepsilon \stackrel{\substack{\uparrow \\ \text{geom. Reihe}}}{\leq} 2\varepsilon < r.$$

c) Zudem gilt:  $\|\sum_{j=0}^{\infty} (g_{j+1} - g_j)\|_{\infty, V_1}$  hat die Majorante  $\sum_{j=0}^{\infty} 2^{-j} \varepsilon$ .  $\implies$  Die Reihe konvergiert gleichmäßig auf  $V_1 \implies$  (DIFF I)

$$g := \lim_{n \rightarrow \infty} g_n = \sum_{j=0}^{\infty} (g_{j+1} - g_j) + b$$

ist stetig auf  $V_1$  und

$$\|g - b\|_{\infty, V_1} \leq 2\varepsilon < r,$$

also  $g(V_1) \subset V_2$ .

Aus der Definition folgt durch Grenzübergang auf beiden Seiten ( $h$  ist stetig)

$$g(x) = b + h(x, g(x)) \quad \forall x \in V_1,$$

also  $g(x) = g(x) - B^{-1}f(x, g(x))$ , also

$$f(x, g(x)) = 0 \quad \forall x \in V_1. \quad \square$$

**Folgerung 3.18.** Sei  $f: U \rightarrow \mathbb{R}$ ,  $U \subset \mathbb{R}^2$  offen, stetig differenzierbar und sei  $(a, b) \in U$ ,  $f(a, b) = c$  und  $\text{grad } f(a, b) \neq 0$ . Dann kann man ein Stück der Höhenlinie  $N_f(c)$  als Graph einer Funktion beschreiben. Denn sei  $\partial_2 f(a, b) \neq 0$ , Satz 3.17, angewandt auf  $f(x, y) = f(x, y) - c$ , impliziert:

$\exists$  Intervalle  $I_1, I_2$ ,  $a \in I_1$ ,  $b \in I_2$ ,  $I_1 \times I_2 \subset U$  und eine stetig differenzierbare Funktion (Bemerkung 3.17.i))  $g: I_1 \rightarrow \mathbb{R}$ , mit  $g(I_1) \subset I_2$  und

$$N_f(c) \cap I_1 \times I_2 = \left\{ (x, y) \in I_1 \times I_2 \mid \tilde{f}(x, y) = 0 \right\} = \{ (x, g(x)) \mid x \in I_1 \} = \Gamma_g.$$

Dito für den Fall, dass  $\partial_1 f(a, b) \neq 0$  ist (mit vertauschten Rollen für  $x, y$ ), also

$$N_f(c) \cap I_1 \times I_2 = \{ (g(x), x) \mid x \in I_2 \}.$$

**Vorlesung 9**

Mo 17.05. 10:15

**Beispiele (Weitere zum Satz über implizite Funktionen).** i)  $f: \mathbb{R}^3 \rightarrow \mathbb{R}^2, f(x, y, z) = \begin{pmatrix} x^2 - y^2 + 1 \\ x^2 - z^2 - 2 \end{pmatrix}$  ist stetig differenzierbar.

$$Df(x, y, z) = \begin{pmatrix} 2x & -2y & 0 \\ 2x & 0 & -2z \end{pmatrix}.$$

Ist  $y_0 \cdot z_0 \neq 0$  (also beide  $\neq 0$ ), so ist  $-2 \begin{pmatrix} y_0 & 0 \\ 0 & z_0 \end{pmatrix}$  invertierbar.

Es gibt also ein offenes Intervall  $I$ ,  $x_0 \in I$  und eine offene Umgebung  $V_2$  von  $\begin{pmatrix} y_0 \\ z_0 \end{pmatrix} \in \mathbb{R}^2 \setminus \text{Achsen}$  und eine stetig differenzierbare Funktion  $g: I \rightarrow \mathbb{R}$  mit  $g(I) \subset V_2$  s. d.

$$f(x, y, z) = C = f(x_0, y_0, z_0) \iff \begin{pmatrix} y \\ z \end{pmatrix} = g(x)$$

und es gilt

$$g'(x) = \frac{1}{2} \begin{pmatrix} \frac{1}{y} & 0 \\ 0 & \frac{1}{z} \end{pmatrix} \begin{pmatrix} 2x \\ 2x \end{pmatrix} = \begin{pmatrix} x/g_1(x) & 0 \\ 0 & x/g_2(x) \end{pmatrix},$$

also  $g'_1(x)g_1(x) = x$  und  $g'_2(x)g_2(x) = x$ . Hauptsatz der Differential- und Integralrechnung

$$\begin{aligned} &\implies \int_{x_0}^x g'_j(t)g_j(t) dt = \frac{x^2}{2} - \frac{x_0^2}{2} \quad j = 1, 2 \\ &\quad \text{(Substitution)} \quad \int_{g_j(x_0)}^{g_j(x)} u du = \frac{1}{2}(g_j^2(x) - g_j^2(x_0)) \\ &\implies x^2 > x_0^2 - y_0^2: g_1(x) = \operatorname{sgn}(y_0) \sqrt{x^2 - x_0^2 + y_0^2} \\ &\quad x^2 > x_0^2 - z_0^2: g_2(x) = \operatorname{sgn}(z_0) \sqrt{x^2 - x_0^2 + z_0^2}. \end{aligned}$$

**Beispiel.** Für  $(x_0, y_0, z_0) = (3, 2, -1)$

$$g: (2\sqrt{2}, \infty) \rightarrow \mathbb{R}^2 \quad g(x) = \begin{pmatrix} \sqrt{x^2 - 5} \\ -\sqrt{x^2 - 8} \end{pmatrix}$$



- ii)  $f: \mathbb{R}^n \times \mathbb{R} \rightarrow \mathbb{R}$ ,  $f(x, t) = t^n + \sum_{j=1}^n x_j t^{n-j}$ ,  $f(x, t) = 0 \implies t$  Nullstelle des Polynoms  $p_x(t) = f(x, t)$ .

Sei  $t_0$  eine einfache Nullstelle. Dann ist  $D_2 f(x, t_0) \neq 0$ , da  $p_x(t) = (t - t_0) q_x(t)$  und  $q_x(t_0) \neq 0$  und somit  $D_2 f(x, t_0) = q_x(t_0) + \cancel{(t_0 - t_0)} q'_x(t_0)$ .

$\implies \exists$  Umgebung  $U$  von  $x_0 \in \mathbb{R}^n$  und ein stetig differenzierbare Funktion  $g: U \rightarrow \mathbb{R}$  mit  $f(x, g(x)) = 0$  und  $g(x_0) = t_0 \implies$  Die einfachen Nullstellen eines Polynoms hängen in stetig differenzierbarer Art von den Koeffizienten ab, d. h. insbesondere hängen sie stetig davon ab, d. h. zu  $\varepsilon > 0$ ,  $\exists \delta > 0$  s. d.  $|t - t_0| = |g(x) - g(x_0)| < \varepsilon$  für  $x \in B_\phi(x_0) \subset U$ .

**Beispiel.**  $3t^2 - t$  Einfache Nullstelle  $t_0 = \frac{1}{3}$ .

$\frac{10}{3}t^2 - t$  Einfache Nullstelle  $t_0 = \frac{3}{10}$ .

- iii) Die Bedingung im Satz über implizite Funktionen ist *nicht* notwendig.  $f(x, y) = y^2$  erfüllt in  $(1, 0)$  die Bedingung  $D_2 f(1, 0)$  invertierbar *nicht*,  $f(x, y) = 0$  besitzt aber die Auflösung  $g(x) = 0 \quad \forall 0$ .

### 3.IV. Der Satz von der Umkehrabbildung

Sei  $f: U \rightarrow \mathbb{R}^n$  stetig differenzierbar,  $U \subset \mathbb{R}^n$ . Gibt es einfach zu prüfende Kriterien, die garantieren, dass  $f$  umkehrbar ist mit stetig differenzierbarer Umkehrfunktion  $g: f(U) \rightarrow \mathbb{R}^n$ ?

Notwendiges Kriterium ist, dass  $Df(x)$  invertierbar ist, da mit der Kettenregel aus  $g \circ f = \text{Id}_u$  folgt  $Dg(f(x)) \cdot Df(x) = \mathbf{1}$ . Tatsächlich ist sogar folgendes hinreichend:

**Satz 3.19.** Sei  $U \subset \mathbb{R}^n$  offen,  $f: U \rightarrow \mathbb{R}^n$ , stetig differenzierbar. Sei  $a \in U$  s. d.  $Df(a)$  invertierbar ist. Dann gibt es eine offene Umgebung  $V \subset U$  von  $a$  und eine offene

Umgebung  $W \in \mathbb{R}^n$  von  $b := f(a)$ , s. d.  $f: V \rightarrow \mathbb{R}^n$ ,  $V$  bijektiv auf  $W$  abgebildet und die Umkehrabbildung  $g: W \rightarrow \mathbb{R}^n$  stetig differenzierbar ist. Es gilt dann

$$Dg(f(x)) = Df(x)^{-1} \quad \forall x \in V.$$

**Bemerkung.** Die letzte Identität, die wir obn schon bewiesen hatten, leuchtet unmittelbar ein, wenn man sich überlegt, dass die Ableitung einer differenzierbare Funktion linear approximiert, und dass die Umkehrabbildung einer invertierbaren linearen Abbildung die inverse Matrix ist.

*Beweis von 3.19.* Betrachte  $F: \mathbb{R}^n \times U \rightarrow \mathbb{R}^n$ ,  $d(z, x) = z - f(x)$ . Dann ist

$$Df(z, x) = \begin{pmatrix} \underbrace{\mathbf{1}_{n \times n}}_{D_1 F(z, x)} & \underbrace{Df(x)}_{D_2 F(z, x)} \end{pmatrix}.$$

Wegen  $F(b, a) = 0$  und  $D_2 F(b, a) = -Df(a)$  invertierbar können wir den Satz von der impliziten Funktion 3.17 auf  $F$  anwenden und  $F$  nach  $x$  auflösen. Es gibt also offene Umgebungen  $V_1 \subset \mathbb{R}^n$  von  $b$  und  $V_2 \subset U$  von  $a$  und eine stetig differenzierbare Funktion (vgl. Bemerkung 3.17.i) nach 3.17)  $g: V_1 \rightarrow \mathbb{R}^n$  mit  $g(V_1) \subset V_2$  und  $F(z, g(z)) = 0 \quad \forall z \in V_1$ , also  $z - f(g(z)) = 0 \quad \forall z \in V_1$  und für all  $(z, x) \in V_1 \times V_2$  mit  $F(z, x) = 0$  gilt  $x = g(z)$ .

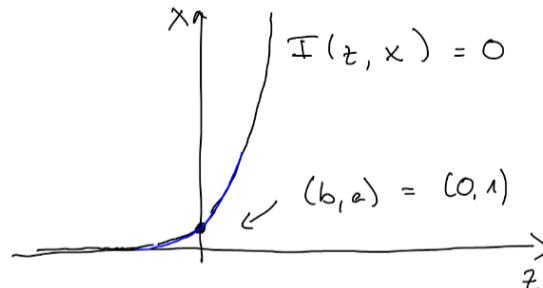
Setze nun

$$\begin{aligned} V &:= V_2 \cap f^{-1}(V_1) \\ &\quad \uparrow \text{Urbild von } V_1 \\ &= \{x \in V_2 \mid f(x) \in V_1\}. \end{aligned}$$

$V$  ist offen (da  $f$  stetig und  $V_1, V_2$  offen) und  $f$  bildet  $V$  bijektiv auf  $W := V_1$  ab und die Umkehrabbildung ist  $g$ .  $\square$

Man sagt, wenn  $f$  wie oben ist: „ $f$  ist bei  $a$  lokal umkehrbar.“

**Beispiele 3.20.** i)  $f(x) = \ln x$ ,  $x > 0$ ,  $a = 1$ ,  $b = 0$ ,  $f'(1) = 1$ .

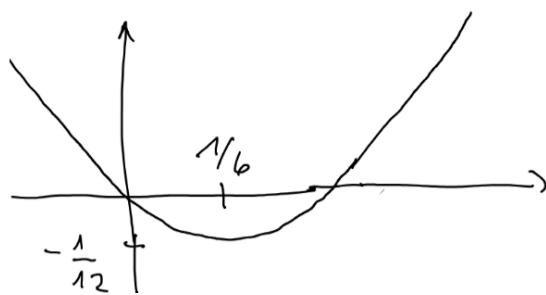


$$\Gamma_g = \{(z, g(z)) \mid z \in V\}$$

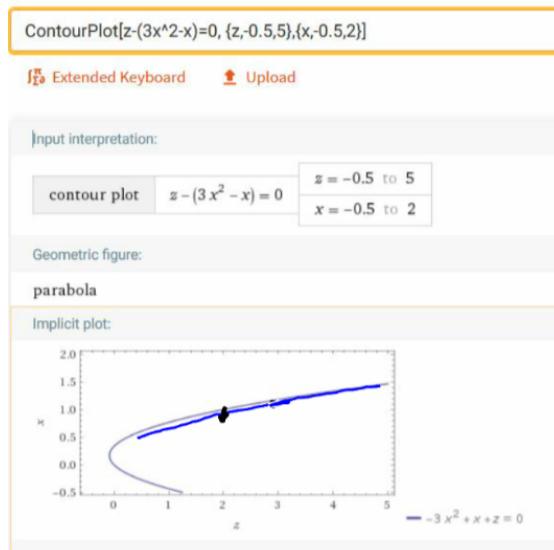
$$\begin{aligned}
 g'(\ln x) &= \left(\frac{1}{x}\right)^{-1} = x \\
 g'(z) &= g(z) \\
 \xrightarrow{\text{(DIFF I)}} g(z) &= C \exp(z),
 \end{aligned}$$

und wegen  $g(0) = 1$  folgt  $C = 1$ .

- ii)  $f(x) = 3x^2 - x$ ,  $f(1) = 2$ ,  $Df(1) = 5 \neq 0$ .



$$F(z, x) = z - f(x) \quad N_f(0) = \{ (f(x), x) \mid x \in U \}$$



$N_F(0)$ ,  $\Gamma_g = \{(z, g(z)) \mid z \in W\}$ , von Wolfram Alpha.  
 $= (f(x), x)$

Die Rekursionsformel  $g_0(z) = a$ ,  $g_{j+1}(z) = a + h(z, g_j(z))$ ,

$$h(z, x) = x - D_2 F(b, a)^{-1} \cdot F(z, x)$$

ist meist nicht sehr nützlich zur Bestimmung von  $g$ . Aber die Ableitung von  $g$  können wir sofort bestimmen:  $Dg(z) = Df(g(z))^{-1}$  oder  $Dg(f(x)) = Df(x)^{-1}$ .

Im Beispiel:  $g'(z) = \frac{1}{6g(z)-1}$  für  $g(z) \neq 1/6$ .  $\Rightarrow$  Ansatz  $g(z) = \alpha + \sqrt{\beta + \gamma z}$

$$\begin{aligned} z &= f(g(z)) = 3(\alpha^2 + \beta + \gamma z + 2a\sqrt{\beta + \gamma z}) - \alpha - \sqrt{\beta + \gamma z} \\ 6\alpha &= 1 \\ \Rightarrow 3\alpha^2 + 3\beta - \alpha &= 0 \Rightarrow g(z) = \frac{1}{6} + \frac{1}{6}\sqrt{+1 + 12z} \\ 3\gamma &= 1 \end{aligned}$$

auf  $(-\frac{1}{12}, \infty)$  definiert  $\Rightarrow$  es muss  $f > -\frac{1}{12}$  sein, also ist das maximale  $V = (\frac{1}{12}, z)$ .

iii)  $f: \mathbb{R}_{>0} \times \mathbb{R} \rightarrow \mathbb{R}^2$ ,  $(r, \phi) \mapsto (r \cos \phi, r \sin \phi)^T$ .

$$\begin{aligned} \det Df(r, \phi) &= \begin{pmatrix} \cos \phi & -r \sin \phi \\ \sin \phi & r \cos \phi \end{pmatrix} \\ \det Df(r, \phi) &= r > 0 \quad \forall (r, \phi). \end{aligned}$$

$\implies$  Bei allen  $r, \phi$  ist  $f$  lokal invertierbar. Es gilt

$$Df(r, \phi)^{-1} = \begin{pmatrix} \cos \phi & \sin \phi \\ -\frac{\sin \phi}{r} & \frac{\cos \phi}{r} \end{pmatrix}.$$

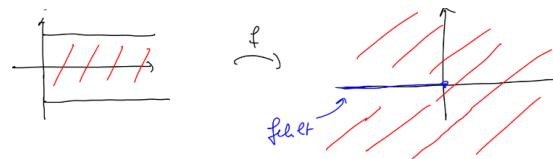
Setze  $f(r, \phi) = \begin{pmatrix} x \\ y \end{pmatrix} \implies r = \sqrt{x^2 + y^2}, \frac{x}{r} = \cos \phi, \frac{y}{r} = \sin \phi$ .

$$\implies Df(r, \phi)^{-1} = Df(f(r, \phi)) = \underbrace{\begin{pmatrix} \frac{x}{\sqrt{x^2+y^2}} & \frac{y}{\sqrt{x^2+y^2}} \\ \frac{-y}{x^2+y^2} & \frac{x}{x^2+y^2} \end{pmatrix}}_{=\begin{pmatrix} x \\ y \end{pmatrix}} = \begin{pmatrix} x \\ y \end{pmatrix}$$

mit  $g$  eine lokale Umkehrung.

**Beachte.**  $f(\mathbb{R}_{>0} \times \mathbb{R}) = \mathbb{R}^2 \setminus \{0\}$ , aber es gibt *keine* globale Umkehrfunktion  $g: \mathbb{R}^2 \setminus \{0\} \rightarrow \mathbb{R}^2$ , denn  $f(r, \phi + k2\pi) = f(r, \phi) \quad \forall k \in \mathbb{Z}$ , also  $f$  nicht injektiv. Man kann maximal Intervalle der Länge  $2\pi$  (in  $\phi$ ) betrachten.

Betrachte etwa  $U = \mathbb{R}_{>0} \times (-\pi, \pi)$ . Das wird unter  $f$  bijektiv auf  $W = \mathbb{R}^2 \setminus \{(x, 0) \mid x \leq 0\}$  abgebildet.



„geschlitzte Ebene“

Wie kommt man darauf, dass  $f$   $V$  auf  $W$  abbildet?  $\phi$  ist entweder  $\neq 0$ , dann ist  $r \sin \phi \in (-r, 0) \cup (0, r)$ , oder  $\phi = 0$ , dann ist  $r \cos \phi = r > 0$ .

$f|_U$  ist injektiv:

Ist  $r \cos \phi = r' \cos \phi'$  und  $r \sin \phi = r' \sin \phi'$ , so ist  $r^2 = r'^2 \implies$  (da,  $r, r' > 0$ )  $r = r'$  und aus  $\cos \phi = \cos \phi'$  folgt zunächst  $\phi = \pm \phi'$ , aber aus  $\sin \phi = \sin \phi' = \pm \sin \phi$  folgt  $\phi = +\phi'$ .

$f|_U$  bildet surjektiv auf  $W$  ab:

Sei  $\begin{pmatrix} x \\ y \end{pmatrix} \in W$ , d. h. es ist  $y \neq 0$  oder  $x > 0 \implies r = \{\sqrt{x^2 + y^2}\} > 0 \implies (\frac{x}{r})^2 + (\frac{y}{r})^2 = 1 \implies |x/r| \leq 1$ . Es ist  $x/r \neq -1$ , denn sonst wäre  $y/r = 0$ , aber  $(-1, 0) \notin W$ . Somit ist  $y = 0 \iff x/r = 1$  und wir legen einen Winkel  $\phi$  fest mit  $\cos \phi = x/r$ .  $\phi \in (-\pi, 0)$ , falls  $y < 0$ , oder  $(0, \pi)$ , falls  $y > 0$ , oder  $\phi = 0$ , falls  $y = 0$ .

**Behauptung.** Für diesen Winkel gilt  $\sin \phi = y/r$ .

*Beweis.*

$$\sin \phi = \operatorname{sgn}(\sin \phi) \sqrt{\sin^2 \phi} = \operatorname{sgn}(y) \sqrt{1 - \cos^2 \phi} = \operatorname{sgn}(y) \sqrt{1 - x^2/r^2} = \operatorname{sgn}(y) |y|/r = y/r. \quad \square$$

Ist man nicht an einer maximalen Umgebung von  $(r_0, \phi_0)$  interessiert, kann man so schneller zum Ziel:

Ist beispielsweise  $-\pi/2 < \phi_0 < \pi/2$  betrachte  $\phi \in (-\pi/2, \pi/2)$ , woraus  $x > 0$  folgt, so dass man recht schnell rät, dass  $f|_V : \mathbb{R}_{>0} \times (-\pi/2, \pi/2) \rightarrow W = \mathbb{R}_{>0} \times \mathbb{R}$  bijektiv auf  $W = \mathbb{R}_{>0} \times \mathbb{R}$  abbildet, indem man sich überlegt, dass die Umkehrfunktion  $g$  durch  $g(x, y) = (\sqrt{x^2 + y^2}, \arctan y/x)^T$  gegeben ist (wohldefiniert für  $x > 0$ ), da

$$g(r \cos \phi, r \sin \phi) = \underbrace{(r, \arctan \frac{\sin \phi}{\cos \phi})}_{\phi = \phi}$$

und es ist klar, dass  $g|_W$  bijektiv auf  $V$  abbildet, da  $\arctan : \mathbb{R} \rightarrow (-\pi/2, \pi/2)$  bijektiv ist, s. d. wenn  $\arctan x/y = \arctan \tilde{x}/\tilde{y}$  gilt  $y = \tilde{y}x$  und dann

$$\sqrt{x^2 + y^2} = \sqrt{x^2 + \frac{\tilde{y}^2 x^2}{\tilde{x}^2}} = \frac{x}{\tilde{x}} \sqrt{\tilde{x}^2 + \tilde{y}^2}$$

nutzt gleich  $\sqrt{\tilde{x}^2 + \tilde{y}^2}$ , wenn  $x = \tilde{x}$  (und somit  $y = \tilde{y}$ ) und  $\mathbb{R}_{>0} \times \mathbb{R} \ni (x, y) \mapsto \sqrt{x^2 + y^2}$  bildet surjektiv auf  $\mathbb{R}_{>0}$  ab (da streng monoton wachsend).

Es folgen einige etwas tiefsinnigere Folgen aus dem Umkehrsatz.

**Satz 3.21.** Sei  $f : U \rightarrow \mathbb{R}^n$  stetig differenzierbar,  $U \subset \mathbb{R}^n$  offen. Sei  $V \subset U$  offen. Ist  $Df(x)$  invertierbar für alle  $x \in V$ , so ist  $f|_V$  eine offene Abbildung, d. h. offene Teilmengen  $\subset V$  auf offene Mengen abgebildet.

**Anti-Beispiel.**  $f: (0, 1) \rightarrow \mathbb{R}$ ,  $f(x) = 1 \quad \forall x$ .  $f'(x) = 0$  nicht invertierbar.  $\{1\}$  nicht offen in  $\mathbb{R}$ .

*Beweis von 3.21.* Sei  $\tilde{V} \subset U$  offen, dann ist  $\tilde{V}$  auch offen  $U$ . Sei  $b \in f(\tilde{V})$ . Wähle ein  $a \in \tilde{V}$  s. d.  $b = f(a)$ . Wende den Satz von der lokalen Umkehrfunktion auf  $f|_{\tilde{V}}$  an  
 $\Rightarrow \exists$  offene Umgebungen  $V_1 \subset \tilde{V}$  und  $W \subset \mathbb{R}^n$  von  $a$  beziehungsweise  $b$  s. d.  $f|_{V_1}$  umkehrbar ist mit  $g: W \rightarrow \mathbb{R}^n$  und insbesondere  $f|_{V_1}$  bijektiv auf  $W$  abbildet. Somit gilt  $W = f(V_1) \subset f(\tilde{V})$  und ist Umgebung von  $b \Rightarrow f(\tilde{V})$  ist offen.  $\square$

**Satz 3.22 (Minimum- / Maximum-Prinzip).** Sei  $f: U \rightarrow \mathbb{R}^n$ ,  $U \subset \mathbb{R}^n$ , stetig differenzierbar. Sei  $V \subset U$  offen und sei  $Df(x)$  invertierbar für alle  $x \in V$ . Dann gilt für die Funktion

$$V \ni x \mapsto \|f(x)\|$$

- i) Sie besitzt kein Maximum.
- ii) Ist  $f(x) \neq 0 \quad \forall x \in V$  besitzt sie kein Minimum.

**Beachte.** Man kann den Beweis dieses Satzes auf 3.10 zurückführen, obwohl  $\|\cdot\|$  nur auf  $\mathbb{R}^n \setminus \{0\}$  partiell differenzierbar ist. Wir wollen hier dennoch anders vorgehen:

*Beweis.* 3.22.i) Angenommen  $a$  ist Maximalstelle von  $\|f\|$ . Dann ist  $f(a) \neq 0$ , da sonst  $\|f(x)\| = 0 \quad \forall x \in V$  und somit  $f|_V = 0$  und  $Df(x) = 0 \quad \forall x \in V$ . 3.21  $\Rightarrow \exists \varepsilon > 0$  s. d.  $B_\varepsilon(f(a)) \subset f(V)$ . Setze  $b := f(a) + \rho f(a)$  mit  $\rho = \varepsilon / 2\|f(a)\|$

$$\begin{aligned} &\Rightarrow \|b - f(a)\| = \rho\|f(a)\| = \varepsilon/2 < \varepsilon \\ &\Rightarrow b \in B_\varepsilon \Rightarrow \exists x_0 \in V \text{ s. d. } b = f(x_0) \\ &\Rightarrow \|f(x_0)\| = \|b\| = (1 + \rho)\|f(a)\| > \|f(a)\|. \end{aligned}$$

3.22.ii) Analog mit  $b = f(a) - \rho f(a)$ .  $\square$

**Vorlesung 10**

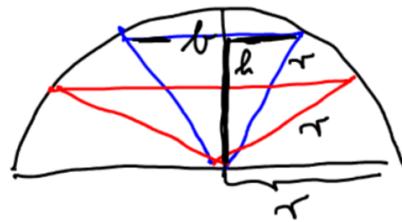
Do 21.05. 10:15

**3.V. Lokale Extrema unter Nebenbedingungen**

Wir betrachten folgende Situation:

Wir suchen Extremstellen einer Funktion  $f$  unter gewissen Nebenbedingungen z. B.

- a) (Mikro-Ökonomie) Es soll das Maximum einer Nutzenfunktion  $f(x, y) = \sqrt{xy}$ ,  $x > 0, y > 0$ , unter der Budgetbedingung  $64 = 2x + 8y$  bestimmt werden.
- b) (Geometrie) Welches Dreieck kopfüber im Halbkreis hat den größten Flächeninhalt?  
Also  $f(b, h) = \frac{1}{2}bh$  unter dr Nebenbedingung  $\frac{1}{4}b^2 + h^2 = r^2$  (Pythagoras).



- c) Berücksichtige, dass eine Bewegung eingeschränkt ist, z. B. auf dem Innenrand einer Kugelschale (Nebenbedingung:  $x^2 + y^2 + z^2 = r^2$ ).

Mathematisch geht es um folgendes Problem:

**Definition 3.23.** Sei  $U \subset \mathbb{R}^n$  offen,  $f: U \rightarrow \mathbb{R}$ , sei  $h: U \rightarrow \mathbb{R}^m$ . Dann sagen wir  $f$  hat in  $a \in U$  ein *lokales Maximum bzw. Minimum* (Extremum) unter der Nebenbedingung  $h(x) = 0$ , wenn:  $a \in N_h(0)$  und es eine Umgebung  $U_0$  von  $a$  gibt s. d.  $f(a) \geq f(x)$  bzw.  $f(a) \leq f(x) \forall x \in U_0 \cap N_h(0)$ .

In einfachen Fällen kann man die Nebenbedingung direkt auflösen und die (3.18) Funktion einsetzen, z. B. in a)

$$y = 8 - \frac{1}{4}x \quad f(x, 8 - \frac{1}{4}x) = \sqrt{x(8 - x/4)}.$$

Allgemeiner: Sei  $f, h$  wie in 3.23 und zudem  $f$  differenzierbar und  $h: U \rightarrow \mathbb{R}^m$ ,  $U \subset \mathbb{R}^{n-m} \times \mathbb{R}^m$ , erfülle in  $a = (a_1, a_2) \in N_h(0)$  die Bedingungen des Satzes über implizite Funktionen. Also gibt es Umgebungen  $V_1$  von  $a_1$  und  $V_2$  von  $a_2$  und eine stetig differenzierbare Funktion  $g: V_1 \rightarrow V_2$  s. d.  $N_h(0) \cap V_1 \times V_2 = \{(u, g(u)) \mid u \in V_1\}$ . Also besitzt

in diesem Fall  $f$  eine lokales Extremum in  $a$  unter der Nebenbedingung  $h = 0$  genau dann, wenn  $u \mapsto f(u, g(u))$  ein lokales Extremum in  $a_1$  besitzt.

Dies führt zur sogenannten

**Satz 3.24 (Lagrange-Multiplikator-Regel).** Sei  $f: U \rightarrow \mathbb{R}$ ,  $U \subset \mathbb{R}^n$  offen, in  $a \in U$  differenzierbar. Sei  $h: U \rightarrow \mathbb{R}^m$  stetig differenzierbar,  $m < n$ . Es sei  $\text{rang } Dh(a) = m$ .

Es besitze  $f$  in  $a$  unter der Nebenbedingung  $h(x) = 0$  ein lokales Extremum. Dann gibt es  $\lambda_1, \dots, \lambda_m \in \mathbb{R}$  (*Lagrange-Multiplikatoren*), s. d.

$$\text{grad } f(a) = \lambda_1 \text{grad } h_1(a) + \dots + \lambda_m \text{grad } h_m(a).$$

*Beweis.* Sei o.B.d.A.  $U = U_1 \times U_2$  mit  $U_1 \subset \mathbb{R}^{n-m}$ ,  $U_2 \subset \mathbb{R}^m$ ,  $a = (a_1, a_2) \in U_1 \times U_2$ , (Notation  $(x, y) \in U_1 \times U_2$ ) und  $Dh(x, y) = (D_1h(x, y) \quad \underbrace{D_2h(x, y)}_{\in \text{Mat}(m \times m, \mathbb{R})})$  mit  $D_2h(a_1, a_2)$  invertierbar.

Satz 3.17  $\implies \exists V_1$  offene Umgebung von  $a_1$ ,  $V_2$  offene Umgebung von  $a_2$ ,  $g: V_1 \rightarrow \mathbb{R}^m$  stetig differenzierbar,  $g(V_1) \subset V_2$ ,  $Dg(x) = -D_2h(x, g(x))^{-1} \cdot D_1h(x, g(x))$  und

$$N_h(0) \cap V_1 \times V_2 = \{(x, g(x)) \mid x \in V_1\}.$$

$a$  ist lokale Extremstelle von  $f$  unter der Nebenbedingung  $h(x) = 0 \implies a_1$  ist lokale Extremstelle von

$$V_1 \ni u \mapsto f(u, g(u)) = f \circ \begin{pmatrix} \text{Id} \\ g \end{pmatrix}(u)$$

$\implies$  (Kettenregel und 3.10)

$$\begin{aligned} 0 &= Df \circ \begin{pmatrix} \text{Id} \\ g \end{pmatrix}(a_1) \\ &= Df(a_1, g(a_1)) \cdot \begin{pmatrix} \mathbb{1}_{n-m} \\ Dg \end{pmatrix}(a_1) \\ &= D_1f(a_1, a_2) - \underbrace{D_2f(a_1, a_2) \cdot D_2h(a_1, a_2)^{-1}}_{(1 \times m) \text{ mal } (m \times m) = (1 \times m) =: (\lambda_1, \dots, \lambda_m)} \cdot D_1h(a_1, a_2) \end{aligned}$$

$$\begin{aligned}
\implies D_1 f(a) &= (\lambda_1, \dots, \lambda_m) \cdot D_1 h(a) \\
&= (\lambda_1, \dots, \lambda_m) \cdot \begin{pmatrix} \partial_1 h_1 & \dots & \partial_{n-m} h_1 \\ \vdots & & \vdots \\ \partial_1 h_m & \dots & \partial_{n-m} h_m \end{pmatrix}(a) \\
&= \left( \sum_{j=1}^m \lambda_j \partial_1 h_j, \sum_{j=1}^m \lambda_j \partial_2 h_j, \dots, \sum_{j=1}^{mm} \lambda_j \partial_{n-m} h_j \right) \\
&= \sum_{j=1}^m \lambda_j D_1 h_j(a).
\end{aligned}$$

Gleichzeitig gilt (Definition der  $\lambda_j$ )

$$D_2 f(a_1, a_2) \cdot D_j h(a_1, a_2)^{-1} =: (\lambda_1, \dots, \lambda_m), \quad (3.1)$$

also (wie oben)

$$\begin{aligned}
D_2 f(a_1, a_2) &= (\lambda_1, \dots, \lambda_m) \cdot D_2 h(a) = \sum_{j=1}^m \lambda_j D_2 h_j(a) \\
\implies (D_1 f(a) &\quad D_2 f(a)) = \sum_{j=1}^m \lambda_j (D_1 h_j(a) \quad D_2 h_j(a)). \quad \square
\end{aligned}$$

Wir nutzt man diesen Satz zur Lösung von Problemen wie oben?

**Bemerkung 3.25.** Unter Voraussetzungen an  $f$  und  $h$  wie oben betrachte das Gleichungssystem

$$\left. \begin{array}{l} \partial_1 f(a) + \lambda_1 \partial_1 h_1(a) + \dots + \lambda_m \partial_1 h_m(a) = 0 \\ \partial_2 f(a) + \lambda_1 \partial_2 h_1(a) + \dots + \lambda_m \partial_2 h_m(a) = 0 \\ \vdots \\ \partial_n f(a) + \lambda_1 \partial_n h_1(a) + \dots + \lambda_m \partial_n h_m(a) = 0 \\ h_1(a) = 0 \\ \vdots \\ h_m(a) = 0 \end{array} \right\} \quad (*)$$

(\*) ist gleichbedeutend mit

$$Df(a) = 0, \quad (**)$$

wenn  $F(a, \lambda) = f(a) + \sum_{j=1}^m \lambda_j h_j(a)$ , denn  $D_1 F = 0$  liefert die ersten  $n$  Zeilen von (\*),  $D_2 F = 0$  liefert die letzten  $m$  Zeilen von (\*).

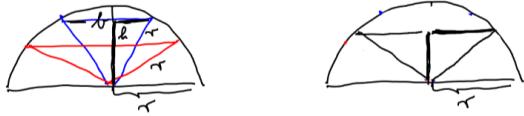
Man versucht dann (\*) oder (\*\*) zu lösen.

**Beispiele.** i) Zunächst b) von oben.  $f: \mathbb{R}_{>0} \times \mathbb{R}_{>0} \rightarrow \mathbb{R}$ ,  $f(b, h) = \frac{1}{2}bh$ ,  $H(b, h) = \frac{1}{4}b^2 + h^2 - r^2$ .  $a = (b, h)$ .  $Df(a) = \left(\frac{1}{2}h, \frac{1}{2}b\right)$ ,  $DH(a) = \left(\frac{1}{2}b, 2h\right)$ .  $h > 0 \implies$  die Voraussetzungen des Satzes sind erfüllt.

$$\begin{aligned} \partial_1 f(a) + \lambda \partial_1 H(a) &= \frac{1}{2}h + \frac{1}{2}b \stackrel{!}{=} 0 \\ \partial_2 f(a) + \lambda \partial_2 H(a) &= \frac{1}{2}b + \lambda 2h \stackrel{!}{=} 0 \\ H(a) &= \frac{1}{4}b^2 + h^2 - r^2 = 0 \end{aligned} \quad (*)$$

$$\implies \lambda = -b/4h \quad h = b^2/4h \implies 4h^2 = b^2$$

$$\stackrel{3. \text{ Zeile}}{\implies} \frac{1}{2}b^2 = r^2 \implies b = \sqrt{2} \quad h = \frac{1}{\sqrt{2}r}.$$



Der Flächeninhalt wird maximal  $1/2r^2$ , denn für  $b \rightarrow 0$  (spitzerwerdendes Dreieck), strebt  $f \rightarrow 0$ , ebenso für  $h \rightarrow 0$  (breiterwerdendes Dreieck).

- ii) Sei  $A \in \text{Mat}(n \times n, \mathbb{R})$  symmetrisch. Sei  $f(x) := \langle x, Ax \rangle$  die zugehörige quadratische Form. Gesucht: Extrema von  $f$  unter der Nebenbedingung  $h(x) = \|x\|_E^2 - 1 = 0$ .  $Dh(x) = 2x \neq 0 \forall x \in N_h(0)$ . Wir können also unseren Satz anwenden

$$\begin{aligned} Df(a) &= 2A \cdot a \\ 2A \cdot a + \lambda 2a &\stackrel{!}{=} 0 \end{aligned} \quad (*)$$

und  $\|a\|_E = 1$ . Das heißt notwendig für das Vorliegen einer Extremstelle ist, dass  $a$  Eigenvektor ist. Da  $A$  symmetrisch, ist  $A$  diagonalisierbar  $\implies \exists$  Eigenvektoren  $v_1, \dots, v_n$ . Wegen  $f(v_j) = \langle v_j, Av_j \rangle = \lambda_j$  Eigenwert folgt:  $f$  wird maximiert von normierten Eigenvektoren zum größten Eigenwert und minimiert von denen zum kleinsten Eigenwert.

- iii) Gesucht ist der Punkt in der Ebene

$$E = \left\{ (x_1, x_2, x_3)^T \in \mathbb{R}^3 \mid x_1 + x_2 - x_3 = 0 \right\},$$

der von Punkt  $(1, 0, 0)^T$  den kleinsten Abstand hat. Da  $\{(1, 0, 0)^T\}$  kompakt ist und  $E$  abgeschlossen, gibt es einen solchen (vgl. unser Beispiel nach Satz 1.41). Wir wollen also  $f(x) = \|x - e_1\|_E^2$  minimieren unter der Neben-Bedingung  $h(x) = x_1 + x_2 - x_3 = 0$ .

$Dh(a) = \begin{pmatrix} 1 & 1 & -1 \end{pmatrix}$  hat maximalen Rang (1) d. h. jeder gesuchte Punkt  $a$  muss erfüllen

$$\begin{aligned} 2(a_1 - 1) + \lambda &= 0 \\ 2a_2 + \lambda &= 0 \\ 2a_3 - \lambda &= 0 \\ a_1 + a_2 - a_3 &= 0, \end{aligned}$$

also ist  $a = \begin{pmatrix} 2/3 & -1/3 & 1/3 \end{pmatrix}$  der einzige Kandidat für ein Extremum. Da es ein Minimum gibt, muss  $a$  dieses sein.

### 3.VI. Höhere Ableitungen, Taylorformel

Ist  $f: U \rightarrow \mathbb{R}^m$ ,  $U \subset \mathbb{R}^n$  offen, stetig differenzierbar, kann man sich fragen, ob auch die Ableitung  $U \ni x \mapsto Df(x)$  als Funktion  $Df: U \rightarrow \text{Mat}(m \times n, \mathbb{R})$  wieder differenzierbar ist (in  $a$ , auf ganz  $U$ , auf  $V \subset U \dots$ ), also ob es eine *lineare* Abbildung  $A: \mathbb{R}^n \rightarrow \text{Mat}(m \times n, \mathbb{R})$  gibt, s. d.

$$Df(a + h) = Df(a) + A(h) * \underline{R}_a^{Df}(h)$$

mit  $\frac{\underline{R}_a^{Df}(h)}{\|h\|} \rightarrow 0$  wenn  $h \rightarrow 0$ . Hier ist  $\underline{R}_a^{Df}: B_\delta(0) \rightarrow \text{Mat}(m \times n, \mathbb{R})$ .

Benennt man  $Df =: g$  sieht man schnell, dass alle Sätze über differenzierbare Abbildungen natürlich auf für  $g$  gelten, insbesondere die Eindeutigkeit der Abbildung, wenn sie existiert, Kettenregel etc. Nur kann man, wenn man nicht den Raum  $\text{Mat}(m \times n, \mathbb{R})$  mit  $\mathbb{R}^{m \times n}$ , die Ausdrücke im Allgemeinen mehr so leicht hinschreiben, etwa  $A(h)$  als Produkt einer Matrix mit dem Vektor  $h$ .

Ist  $m = 1$  oder betrachtet man nur die Komponentenfunktionen einer  $\mathbb{R}^m$ -wertigen Funktion, sieht man jedoch:

$$g := Df: U \rightarrow \text{Mat}(1 \times n, \mathbb{R})$$

nimmt Werte in den  $1 \times n$ -Matrizen an, identifiziert man nun  $\text{Mat}(1 \times n, \mathbb{R}) \simeq \mathbb{R}^n$  (Spaltenvektor), hat man es mit einer Funktion  $g: U \rightarrow \mathbb{R}^n$  zu tun.

Einen *Kandidaten* für  $Dg(a)$  findet man wieder über die partiellen Ableitungen

$$\left( \partial_1 g(a) \dots \partial_n g(a) \right) \in \text{Mat}(n \times n, \mathbb{R}) \quad (3.2)$$

und

$$g(x) = \begin{pmatrix} \partial_1 f(x) \\ \vdots \\ \partial_n f(x) \end{pmatrix}$$

(wenn sie existieren). Insgesamt also

$$\begin{pmatrix} \partial_1^2 f(a) & \partial_2 \partial_1(a) \dots \partial_n \partial_1 f(a) \\ \partial_1 \partial_2 f(a) & \partial_2^2(a) \dots \partial_n \partial_2 f(a) \\ \vdots & \vdots \\ \partial_1 \partial_n f(a) & \partial_2 \partial_n f(a) \dots \partial_n^2 f(a) \end{pmatrix} =: A$$

mit der Notation  $\partial_j \partial_i f(a) = \partial_j g_i(a) = \partial_j(\partial_i f)(a)$  und  $\partial_j^2 = \partial_j \partial_j$ .

Dann berechnet man den Rest wie üblich:

$$g(a + h) = g(a) + A \cdot h + \underline{R}_a^g(h)$$

und überprüft, ob  $\frac{\underline{R}_a^g(h)}{\|h\|} \rightarrow 0$  für  $h \rightarrow 0$ .

Oder man überprüft, ob die partiellen Ableitungen stetig sind. Wenn ja, folgt die Differenzierbarkeit von  $Df$  aus Satz 3.11.

Wenn nein, muss man prüfen, ob  $\underline{R}_a^g$  das richtige Verhalten bei 0 hat. Wir werden allerdings bei den sogenannten höheren Ableitungen (wie der 2. Ableitung oben) meist stetig differenzierbare Funktionen betrachten.

**Definition 3.26.**  $f: U \rightarrow \mathbb{R}$ ,  $U \subset \mathbb{R}^n$  offen, heißt in  $a$   $k$ -fach partiell differenzierbar, wenn alle partiellen Ableitungen

$$\partial_{j_1} f(a), \partial_{j_1} \partial_{j_2} f(a), \dots, \partial_{j_1} \partial_{j_2} \dots \partial_{j_k} f(a)$$

(Notation wie oben)  $j_i \in \{1, \dots, n\}$  existieren.

**Achtung.** Beachte, dass man zur Berechnung einer partiellen Ableitung die Funktion in einer Umgebung des Punktes kennen muss. Um also etwa eine  $k$ -te partielle Ableitung von  $f$  in  $a$  zu berechnen, muss man die  $(k-1)$ -te partielle Ableitung in einer Umgebung von  $a$  kennen.

Ist  $f$  auf  $U$   $k$ -fach partiell differenzierbar und sind alle partiellen Ableitungen der Ordnung  $\leq k$  stetig, so heißt  $f$   $k$ -fach stetig partiell differenzierbar.

**Beispiel.**  $f(x_1, \dots, x_4) = x_1^3 + x_2^2 x_3 - 3x_4 x_2$ .  $\partial_1^2 f(x) = 6x_1$ ,  $\partial_2^2 \partial_4 f(x) = \partial_2^2(-3x_2) = 0$ ,  
 $\partial_2 \partial_4 \partial_3 f(x) = \partial_2 \partial_4(2x_2 x_3 - 3x_4) = \partial_2(-3x_4) = 0$ .

Dass die beiden letzt genannten Ableitungen gleich sind, ist kein Zufall:

**Satz 3.27 (Satz von Schwarz).** Sei  $f: U \rightarrow \mathbb{R}$ ,  $U \subset \mathbb{R}^n$ , zweimal stetig partiell differenzierbar. Dann gilt für alle  $a \in U$  und alle  $i, j \in \{1, \dots, n\}$

$$\partial_j \partial_i f(a) = \partial_i \partial_j f(a).$$

*Beweis.* Sei o.B.d.A.  $n = 2$ ,  $a = (a_1, a_2)$ .  $\exists \delta > 0$  s. d.  $B_\delta(a_1) \times B_\delta(a_2) \subset U$ . Setze zu  
 $y \in B_\delta(a_2)$

$$F_y: B_\delta(a_1) \rightarrow \mathbb{R} \quad F_y(x) = f(x, y) - f(x, a_2).$$

MWS  $\implies \exists \zeta \in B_\delta(a_1) = (a_1 - \delta, a_1 + \delta)$  und zwar zwischen  $a_1$  und  $x$  s. d.

$$F_y(x) - F_y(a_1) = F'_y(\zeta)(x - a_1).$$

Es ist  $F'_y(\zeta) = \partial_1 f(\zeta, y) - \partial_1 f(\zeta, a_2)$ . Betrachte die Funktion  $B_\delta(a_2) \ni y \mapsto \partial_1 f(\zeta, y)$ .  
MWS  $\implies \exists \eta \in B_\delta(a_2) = (a_2 - \delta, a_2 + \delta)$  und zwar zwischen  $a_2$  und  $y$  s. d.

$$\partial_1 f(\zeta, y) - \partial_1 f(\zeta, a_2) = \partial_2(\partial_1 f)(\zeta, \eta)(y - a_2).$$

Insgesamt folgt:

$$f(x, y) - f(x, a_2) - f(a_1, y) + f(a_1, a_2) = \partial_2 \partial_1 f(\zeta, \eta)(y - a_2) \cdot (x - a_1). \quad (*)$$

Dieselbe Konstruktion nehmen wir nun zu festen  $x \in B_\delta(a_1)$  für  $G_x(y) = f(x, y) - f(x, a_1)$  vor.

### MWS:

$\exists \tilde{\eta} \in (a_2 - \delta, a_2 + \delta)$  zwischen  $a_2$  und  $y$  s. d.  $G_x(y) - G_x(a_2) = G'_x(\tilde{\eta})(y - a_2)$ . Es ist  
 $G'_x(\tilde{\eta}) = \partial_2 f(x, \tilde{\eta}) - \partial_2 f(x, a_2)$ .

### MWS:

$\exists \tilde{\zeta} \in (a_1 - \delta, a_1 + \delta)$  zwischen  $a_1$  und  $x$  s. d.  $\partial_2 f(x, \tilde{\eta}) - \partial_2 f(a_1, \tilde{\eta}) = \partial_1(\partial_2 f)(\tilde{\zeta}, \tilde{\eta})(x - a_1)$ .  
 $f(x, y) - f(x, a_2) - f(x, a_1) + f(a_1, a_2) = \partial_1 \partial_2 f(\tilde{\zeta}, \tilde{\eta})(y - a_2) \cdot (x - a_1). \quad (**)$

Aus (\*) und (\*\*) folgt für  $x \neq a_1, y \neq a_2$

$$\partial_2 \partial_1 f(\tilde{\zeta}, \tilde{\eta}) = \partial_1 \partial_2 f(\tilde{\zeta}, \tilde{\eta}),$$

mit  $\zeta, \tilde{\zeta}$  zwischen  $a_1$  und  $x$  und  $\eta, \tilde{\eta}$  zwischen  $a_2$  und  $y$ .

Grenzübergang  $x \rightarrow a_1$  und  $y \rightarrow a_2$  liefert wegen der Stetigkeit von  $\partial_2 \partial_1 f$  und  $\partial_1 \partial_2 f$  die Beh.  $\square$

**Anti-Beispiel.**

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

Es gilt  $\partial_2 \partial_1 f(0, 0) \neq \partial_1 \partial_2 f(0, 0)$ . Denn partielle Ableitung in 0:

$$\partial_1 f(0, 0) = \lim_{h \rightarrow 0} \frac{f(h, 0) - f(0, 0)}{h} = 0 = \partial_2 f(0, 0).$$

Um die zweiten Ableitungen zu bestimmen, müssen wir aber  $\partial_1 f$  auch in der Nähe der 0 kennen. Daher bestimmen wir für  $(x, y) \neq (0, 0)$

$$\begin{aligned} \partial_1 f(x, y) &= \frac{x^4 y + 4x^4 y^3 - y^5}{(x^2 + y^2)^2} \\ \partial_2 f(x, y) &= \frac{x^5 - x y^4 - 4x^3 y^4}{(x^2 + y^2)^2} \\ \implies \partial_2 \partial_1 f(0, 0) &= \lim_{h \rightarrow 0} \frac{1}{h} (\partial_1 f(0, h) - \partial_1 f(0, 0)) \\ &= \lim_{h \rightarrow 0} \frac{1}{h} (-h^5/h^4) = -1 \\ \partial_1 \partial_2 f(0, 0) &= \lim_{h \rightarrow 0} \frac{1}{h} (\partial_2 f(h, 0) - \partial_2 f(0, 0)) \\ &= \lim_{h \rightarrow 0} \frac{1}{h} (h^5/h^4) = +1. \end{aligned}$$

### 3.VII. Der Laplace-Operator

Sei  $f: U \rightarrow \mathbb{R}$ ,  $U \subset \mathbb{R}^n$  offen, zweimal partiell stetig differenzierbar. Man setzt

$$\Delta f(x) := \partial_1^2 f(x) + \partial_2^2 f(x) + \cdots + \partial_n^2 f(x)$$

und nennt  $\Delta = \partial_1^2 + \cdots + \partial_n^2$  den *Laplace-Operator*.

**Beispiele.** i)  $h: \mathbb{R}^n \setminus \{0\} \rightarrow \mathbb{R}$ ,  $h(x) = f(\|x\|_E)$  mit einer auf  $\mathbb{R} \setminus \{0\}$  zweimal stetig partiell differenzierbaren Funktion  $f$ .

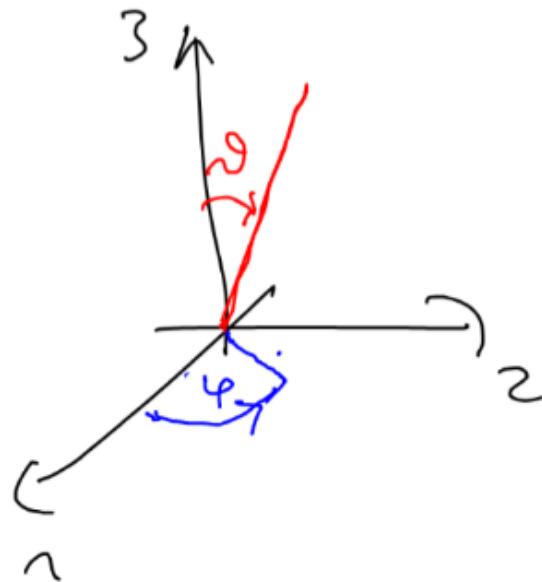
$$\text{grad } f(\|x\|_E) = f'(\|x\|_E) \frac{1}{\|x\|_E} x$$

und

$$\begin{aligned} \partial_i \left( f'(\|x\|_E) \frac{1}{\|x\|_E} x_i \right) &= f''(\|x\|_E) \frac{x_i}{\|x\|_E} \cdot \frac{x_i}{\|x\|_E} + f'(\|x\|_E) \left( -\frac{1}{2} \frac{2x_i}{\sqrt{\sum x_j^2}} \cdot x_i + \frac{1}{x_E} \right) \\ \implies \Delta h(x) &= f''(\|x\|_E) + f'(\|x\|_E) \frac{1}{\|x\|_E} (n - 1) \end{aligned}$$

ii) In Kugelkoordinaten: Sei  $g: \mathbb{R}_{>0} \times \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}^3$

$$g(r, \tau, \phi) = \begin{pmatrix} r \cos \phi \sin \theta \\ r \sin \phi \sin \theta \\ r \cos \theta \end{pmatrix}.$$



$g$  ist zweimal stetig partiell differenzierbar.

Sei  $f$  zweimal stetig differenzierbar auf  $\mathbb{R}^3 \rightarrow \mathbb{R}$ .  $\Delta f$  in Kugelkoordinaten bestimmen, also wenn wir  $\begin{pmatrix} x \\ y \\ z \end{pmatrix} = g(r, \theta, \phi)$  schreiben. Zunächst überlegen wir, was es heißt,  $f$  in Kugelkoordinaten zu schreiben.

**Beispiel.**  $f(x, y, z) = xyz$ .

$$f \circ g(r, \theta, \phi) = f(r \cos \phi \sin \theta, r \sin \phi \sin \theta, r \cos \theta) = r^3 \sin^2 \theta \cos \theta \cos \phi \sin \phi =: F(r, \theta, \phi).$$

Sodann:

$$Dg(r, \theta, \phi) = \begin{pmatrix} \cos \phi \sin \theta & r \cos \phi \cos \theta & -r \sin \phi \sin \theta \\ \sin \phi \sin \theta & r \sin \phi \cos \theta & r \cos \phi \sin \theta \\ \cos \theta & -r \sin \theta & 0 \end{pmatrix}$$

$\det Dg(r, \theta, \phi) = r^2 \sin \theta \implies$  Ist  $v \neq k\pi$ ,  $k \in \mathbb{Z}$ , so ist  $g$  lokal umkehrbar mit stetig differenzierbarer Umkehrfunktion  $h: W \rightarrow V$  und in  $(x, y, z) \in W$  gilt

$$\begin{aligned}
 f(\underbrace{x, y, z}_{=g(r, \theta, \phi)}) &= f \circ g \circ h(x, y, z) =: F \circ h(x, y, z) \\
 \implies Df(x, y, z) &= 1 \times 3\text{-Matrix} \cdot \underbrace{Dh(x, y, z)}_{\substack{\uparrow \\ DF(r, \theta, \phi)}} =: \tilde{F}(r, \theta, \phi) = \tilde{F} \circ h(x, y, z) \\
 \implies (\partial_1^2 + \partial_2^2 + \partial_3^2)f(x, y, z) &= \partial_1(\tilde{F} \circ h)_1(\underbrace{x, y, z}_{=:x}) + \partial_2(\tilde{F} \circ h)_2(x, y, z) + \partial_3(\tilde{F} \circ h)_3(x, y, z) \\
 &= \sum_{j=1}^3 (\partial_j \cdot \tilde{F}_j)(\underbrace{h(x)}_{=(r, \theta, \phi)}) \cdot \underbrace{\partial_j h_j(x)}_{=\frac{1}{\partial_j g_j(r, \theta, \phi)}} \quad (\text{Umkehrsatz}) \\
 &= \sum_{j=1}^3 \partial_j \tilde{F}_j(r, \theta, \phi) \cdot \frac{1}{\partial_j g_j(r, \theta, \phi)}.
 \end{aligned}$$

Eine etwas längere Rechnung, bei der man u. a.  $Dg(r, \theta, \phi)$  invertieren muss, liefert die explizite Formel für die  $\tilde{F}_j$ , etwa

$$\tilde{F}_1 = \begin{pmatrix} \partial_1 F & \partial_2 F & \partial_3 F \end{pmatrix} \cdot \begin{pmatrix} \cos \phi \sin \theta \\ \frac{1}{r} \cos \phi \sin \theta \\ -\frac{1}{r} \sin \phi \end{pmatrix}.$$

↑  
1. Spalte von  $Dg(r, \theta, \phi)^{-1}$

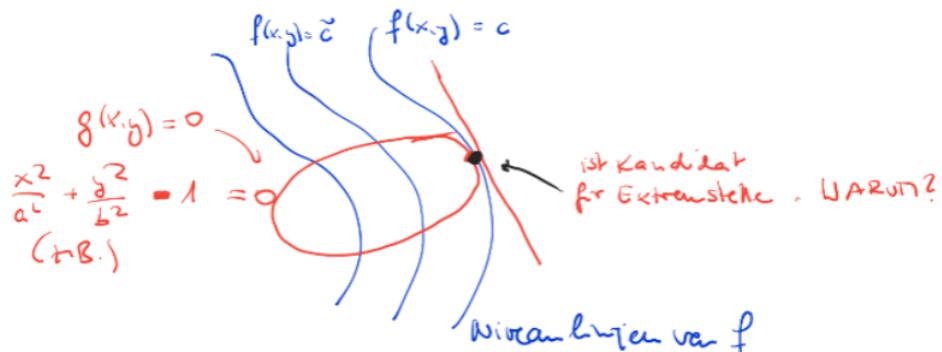
Von diesen berechnet man die jeweilige partielle Ableitung  $\partial_j \tilde{F}_j$  und multipliziert mit  $\frac{1}{\partial_j g_j}$ .

z. B.

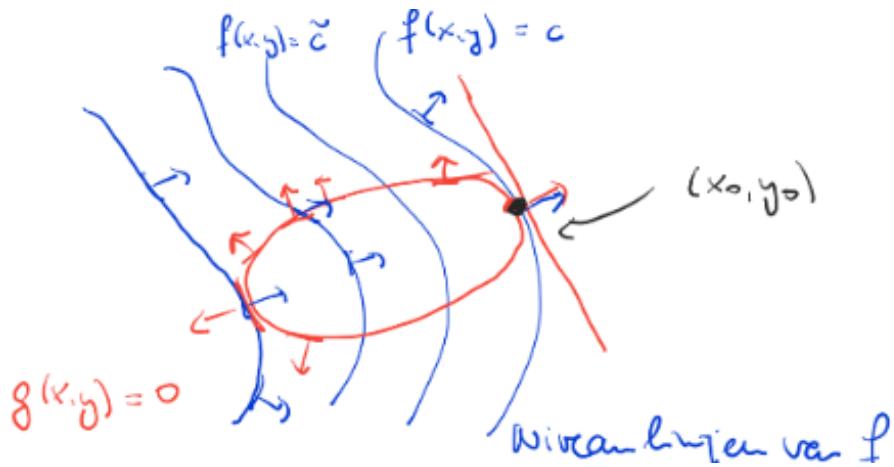
$$\begin{aligned}
 \partial_1 \tilde{F}_1 &= \partial_1^2 F \cdot \cos \phi \in \theta + \partial_1 \partial_2 \cdot \frac{1}{r} \cos \phi \cos \theta - \partial_2 F \cdot \frac{1}{r^2} \cos \phi \cos \theta - \partial_1 \partial_3 \cdot \frac{1}{r \sin \theta} \sin \phi + \partial_3 F \frac{1}{r^2} \frac{\sin \phi}{\sin \theta} \\
 1/\partial_1 g_1 &= \frac{1}{\cos \phi \sin \theta}.
 \end{aligned}$$

Das Ergebnis ist

$$\frac{1}{r^2} \partial_r (r^2 \partial_r F) + \frac{1}{r^2 \sin \theta} \partial_\theta (\sin \theta \partial_\theta F) + \frac{1}{r^2 \sin^2 \theta} \partial_\phi^2 F.$$



**Ergänzung.** Tatsache:  $\text{grad } h(x, y)$  steht senkrecht auf  $h(x, y) = c$ . (Der Gradient zeigt in die Richtung des stärksten Anstiegs der Funktion).



Tatsache 2: Liegt in  $(x_0, y_0) \in \{(x, y) \mid g(x, y) = 0\}$  ein Extremum von  $f$ , so  $\exists \lambda \in \mathbb{R}$  s. d.  $\boxed{\text{grad } f(x_0, y_0) \lambda \text{grad } g(x_0, y_0)}$ . Tatsache 1 + 2 gleichbedeutend  $f(x, y) = c$  schmiegt sich in  $(x_0, y_0)$  tangential an  $\{g(x, y) = 0\}$  an.

Das heißt, wenn wir das Gleichungssystem

$$\left. \begin{array}{l} \text{grad } f(x, y) - \lambda \text{grad } g(x, y) = 0 \\ g(x, y) = 0 \end{array} \right\} \quad (*)$$

lösen, haben wir Kandidaten für Extremstellen gefunden.

$$\begin{aligned} (*) \therefore \partial_x L &= 0 \\ \partial_y L &= 0 \\ \partial_\lambda L &= 0, \end{aligned}$$

$L$  Lagrange-Funktion.

**Vorlesung 11**

Mo 25.05. 10:15

**3.VIII. Taylor-Formel, lokale Extrema**

**Notation 3.28 (Multi-Index-Schreibweise).** Sei  $\alpha := (\alpha_1, \dots, \alpha_n) \in \mathbb{N}_0^n$ . Dann definiert man

$$|\alpha| := \alpha_1 + \alpha_2 + \dots + \alpha_n$$

$$\alpha! := \alpha_1! \alpha_2! \dots \alpha_n!.$$

Ist  $x = (x_1, \dots, x_n)^T \in \mathbb{R}^n$ , so setzt man

$$x^\alpha := x_1^{\alpha_1} x_2^{\alpha_2} \dots x_n^{\alpha_n}.$$

Ist  $f$  eine  $|\alpha|$ -mal stetig differenzierbare Funktion, so ist

$$\partial^\alpha f := \partial_1^{\alpha_1} \partial_2^{\alpha_2} \dots \partial_n^{\alpha_n} f$$

**Lemma 3.29.** Sei  $f: U \rightarrow \mathbb{R}$ ,  $U \subset \mathbb{R}^n$  offen, ine  $k$ -fach stetig differenzierbare Funktion. Sei  $a \in U$  und sei  $h \in \mathbb{R}^n$  so, dass  $\{a + th \mid t \in [0, 1]\}$  ganz in  $U$  liegt.



Dann ist die Funktion  $g: [0, 1] \rightarrow \mathbb{R}$ ,  $g(t) = f(a + th)$   $k$ -mal stetig differenzierbar und es gilt

$$g^{(k)}(t) = \sum_{|\alpha|=k} \frac{k!}{\alpha!} \partial^\alpha f(a + th) h^\alpha.$$

↑  
Summe über alle  $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{N}_0^h$  mit  $\alpha_1 + \alpha_2 + \dots + \alpha_n = k$

*Beweis.* 1. Wir zeigen per Induktion über  $k$ , dass gilt

$$g^{(k)}(t) = \sum_{i_1, \dots, i_k \in \{1, \dots, n\}} \partial_{i_1} \dots \partial_{i_k} f(a + th) h_{i_1} \dots h_{i_k}.$$

Induktionsanfang:  $k = 1$

$$g'(t) = Df(a + th) \cdot h$$

$$\sum_{i=1}^n \partial_i f(a + th) h_i.$$

Kettenregel

$k - 1 \rightarrow k$

$$\frac{d}{dt} \left( \sum_{i_1, \dots, i_{k-1} \in \{1, \dots, n\}} \partial_{i_1} \cdots \partial_{i_{k-1}} f(a + th) h_{i_1} \cdots h_{i_{k-1}} \right) = \sum_{i_1, \dots, i_{k-1} \in \{1, \dots, n\}} \sum_{i=1}^n \partial_i \partial_{i_1} \cdots \partial_{i_{k-1}} f(a + th) h_{i_1} \cdots h_{i_k}$$

2. Kommt unter den Indizes  $(i_1, \dots, i_k)$  der Index 1  $\alpha_1$ -mal vor, der Index 2  $\alpha_2$ -mal, ..., der Index  $n$   $\alpha_n$ -mal, so ist

$$\partial_{i_1} \cdots \partial_{i_k} f(a + th) = \partial_1^{\alpha_1} \cdots \partial_n^{\alpha_n} f(a + th).$$

Es gibt

$$\frac{k!}{\alpha_1! \cdots \alpha_n!}$$

$k$ -Tupel  $(i_1, \dots, i_k)$  von Zahlen  $1 \leq i_j \leq n$ , in denen 1 genau  $\alpha_1$ -mal, 2 genau  $\alpha_2$ -mal, ...,  $n$  genau  $\alpha_n$ -mal vorkommt (ohne Beweis).  $\implies$  Beh.  $\square$

**Satz 3.30 (Taylorsche Formel).** Sei  $f: U \rightarrow \mathbb{R}$ ,  $U \subset \mathbb{R}^n$  offen,  $(k+1)$ -mal stetig differenzierbar. Sei  $a \in U$  und  $h \in \mathbb{R}^n$  s. d.  $\{a + th \mid t \in [0, 1]\} \subset U$ . Dann existiert  $\theta \in [0, 1]$  s. d.

$$f(a + h) = \sum_{\substack{|\alpha| \leq k \\ \alpha \in \mathbb{N}_0^n}} \frac{1}{\alpha!} \partial^\alpha f(a) h^\alpha + \sum_{\substack{|\alpha|=k+1 \\ \alpha \in \mathbb{N}_0^n}} \frac{1}{\alpha!} \partial^\alpha f(a + \theta h) h^\alpha.$$

*Beweis.* Setze  $g: [0, 1] \rightarrow \mathbb{R}$ ,  $g(t) := f(a + th)$ . Taylorsche Formel aus der DIFF I  $\implies \exists \theta \in [0, 1]$  s. d.

$$g(1) = \underbrace{\sum_{j=0}^k \frac{1}{j!} g^{(j)}(0)}_{\text{(3.29)}} + \underbrace{\frac{1}{(k+1)!} g^{(k+1)}(\theta)}_{\sum_{|\alpha|=k+1} \frac{1}{\alpha!} \partial^\alpha f(a + \theta h) h^\alpha} \quad \square$$

**Folgerung 3.31.** Sei  $f: U \rightarrow \mathbb{R}$ ,  $U \subset \mathbb{R}^n$  offen,  $k$ -mal stetig differenzierbar. Sei  $a \in U$  und  $\delta > 0$  s. d.  $B_\delta(a) \subset U$ . Dann gilt für alle  $h \in B_\delta(0)$ :

$$f(a + h) = \sum_{\alpha \leq k} \frac{1}{\alpha!} \partial^\alpha f(a) h^\alpha + R_{k,a}(h)$$

mit  $\frac{R_{k,a}(h)}{\|h\|^k} \rightarrow 0$  für  $h \rightarrow 0$ .

Äquivalent:

$$f(x) = \sum_{\alpha \leq k} \frac{1}{\alpha!} \partial^\alpha f(a) (x - a)^\alpha + R_{k,a}(x - a)$$

für  $x \in B_\delta(a)$ . „Taylor-Entwicklung bis zur Ordnung  $k$  mit Entwicklungspunkt  $a$ “.

*Beweis.*  $\exists \theta = \alpha_{a,h} \in [0, 1]$  s. d.

$$\begin{aligned} f(a+h) &= \sum_{|\alpha| \leq k-1} \frac{1}{\alpha!} f(a) h^\alpha + \sum_{|\alpha|=k} \partial^\alpha f(a+\theta h) h^\alpha \\ &= \sum_{|\alpha| \leq k} \frac{1}{\alpha!} \partial^\alpha f(a) h^\alpha + \sum_{|\alpha|=k} \frac{1}{\alpha!} (\underbrace{\partial^\alpha f(a+\theta h) - \partial^\alpha f(a)}_{:= r_{\alpha,a}(h)}) h^\alpha. \end{aligned}$$

$\partial^\alpha f$  ist stetig  $\implies r_{\alpha,a}(h) \rightarrow r_{\alpha,a}(0) = 0$  für  $h \rightarrow 0$  für  $h \rightarrow 0$ . Setze

$$R_{k,a}(h) := \sum_{|\alpha|=k} \frac{1}{\alpha!} r_{\alpha,a}(h) h^\alpha,$$

dann gilt  $\frac{R_{k,a}(h)}{\|h\|^k} \rightarrow 0$ , denn

$$\frac{|h^\alpha|}{\|h\|_{\max}^k} = \frac{|h_1|^{\alpha_1} \cdots |h_n|^{\alpha_n}}{(\|h\|_{\max})^k} \leq 1$$

□

für  $|\alpha| = \alpha_1 + \cdots + \alpha_n = k$ .

### Notation.

$$P_m(h) := P_{m,a}(h) = \sum_{|\alpha|=m} \frac{1}{\alpha!} \partial^\alpha f(a) h^\alpha$$

ist ein homogenes Polynom von Grad  $m$  (in  $h$ ), das sogenannte *Taylorpolynom* zu  $f$  zum Entwicklungspunkt  $a$ . Es gilt

$$f(a+h) = \sum_{m=0}^k P_{m,a}(h) + R_{k,a}(h).$$

$m=0$   $P_0(h) = f(a)$  konstant.

$m=1$  Die Summe läuft über alle

$$(\alpha_1, \dots, \alpha_n) \in \{(1, 0, \dots, 0), (0, 1, 0, \dots), \dots, (0, \dots, 0, 1)\}.$$

Also gilt

$$P_1(h) = \sum_{j=1}^a \partial_j f(a) h_j = (\partial_1 f(a) \dots \partial_n f(a)) \cdot h.$$

Die Formel aus der Folgerung entspricht also genau der linearen Approximation von  $f$ , die durch Ableitung gegeben ist.

$m = 2$  Die Summe läuft über alle

$$\begin{aligned} (\alpha_1, \alpha_2, \dots, \alpha_n) \in & \{(1, 1, 0, \dots, 0), (1, 0, 1, 0, \dots, 0), \dots \\ & , (0, 1, 1, 0, \dots, 0), (0, 1, 0, 1, 0, \dots, 0), \dots \\ & , \dots \\ & , (2, 0, \dots, 0), (0, 2, 0, \dots, 0), \dots, (0, \dots, 0, 2)\}. \end{aligned}$$

Also ist

$$\begin{aligned} P_2(h) &= \sum_{j=1}^n \underbrace{\sum_{i=j+1}^n \partial_i \partial_j f(a) h_i h_j}_{\substack{= \sum_{j < i} \partial_i \partial_j f(a) h_i h_j \\ \uparrow \text{Schwarz}}} + \sum_{j=1}^n \frac{1}{2} \partial_j^2 f(a) h_j^2 = \frac{1}{2} \sum_{i,j} \partial_j \partial_i f(a) h_i h_j. \end{aligned}$$

Man definiert daher die sogenannte *Hesse-Matrix*

$$H_f(a) := ((\partial_i \partial_j f(a))_{\substack{i \in \{1, \dots, n\} \\ j \in \{1, \dots, n\}}})$$

und schreibt

$$f(a + h) = f(a) + \langle \text{grad } f(a), h \rangle + \frac{1}{2} \langle h, H_f(a) h \rangle + R_{z,a}(h)$$

für  $f: U \rightarrow \mathbb{R}$  zweimal stetig differenzierbar und  $h \in B_\delta(0)$ ,  $\delta > 0$  s. d.  $B_\delta(a) \subset U$ .

### 3.IX. Lokale Extrema

Wir hatten in 3.10 bereits gesehen: Ist  $a \in U$  lokale Extremstell von  $f: U \rightarrow \mathbb{R}$ ,  $U \subset \mathbb{R}^n$  offen, differenzierbar, so ist  $\text{grad } f(a) = 0$ .

**Satz 3.32.** Sei  $f: U \rightarrow \mathbb{R}$  zweimal stetig differenzierbar,  $U \subset \mathbb{R}^n$  offen. Sei  $a \in U$  s. d.  $\text{grad } f(a) = 0$ . Dann gilt:

- a) Ist  $H_f(a)$  positiv definit (d. h.  $\langle h, H_f(a)h \rangle > 0 \forall h \in \mathbb{R}^n, h \neq 0$ ), so hat  $f$  in  $a$  ein lokales (isoliertes) Minimum.
- b) Ist  $H_f(a)$  negativ definit (d. h.  $\langle h, H_f(a)h \rangle < 0 \forall h \neq 0$ ), so hat  $f$  in  $a$  ein lokales (isoliertes) Maximum.
- c) Ist  $H_f(a)$  indefinit (d. h.  $\exists h$  s. d.  $\langle h, H_f(a)h \rangle > 0$  und  $\exists \tilde{h}$  s. d.  $\langle \tilde{h}, H_f(a)\tilde{h} \rangle < 0$ ), so hat  $f$  in  $a$  kein lokales Extremum.

**Bemerkung.** Ist  $H_f(a)$  positiv oder negativ semidefinit s. d.  $\forall h \neq 0$  ist  $\langle h, H_f(a)h \rangle \geq 0$  (bzw.  $\leq 0$ ), so ist keine allgemeine Aussage möglich.

**Beispiel.**  $f_1(x, y) = x^2 + y^4$ ,  $f_2(x, y) = x^2$ ,  $f_3(x, y) = x^2 + y^3$ ,  $H_{f_j}(0) = \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix}$   $\forall j$  ist positiv semidefinit ( $\langle h, H_{f_j}(0)h \rangle = 2h_1^2 \geq 0$ ,  $= 0$  für  $h_1 = 0$ ).

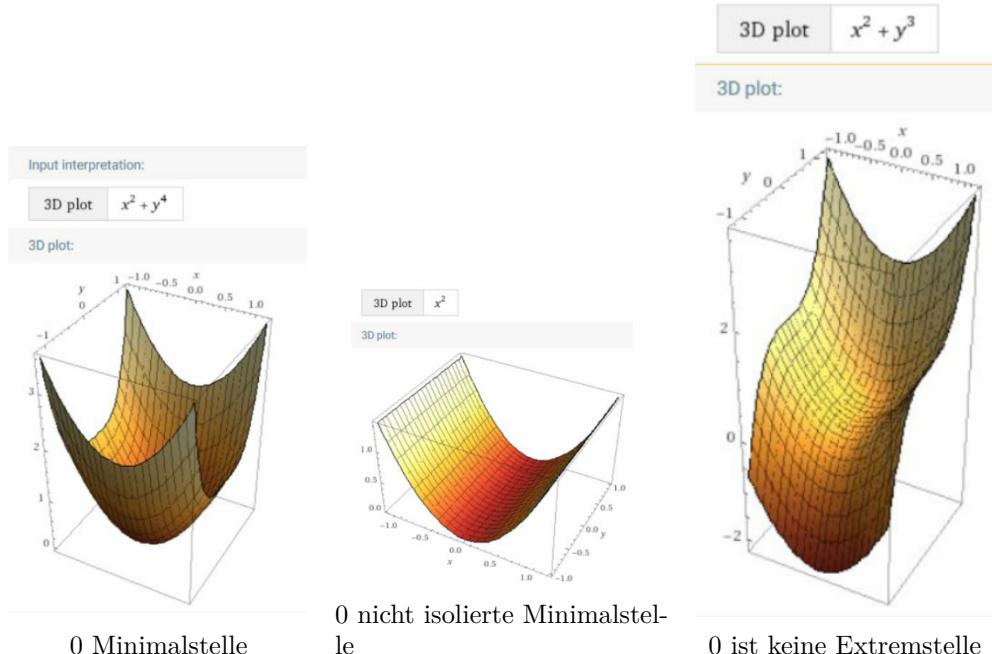


Abbildung 3.1.: Wolfram Alpha

*Beweis von 3.32.* Setze  $A := H_f(a)$ .  $\exists \delta > 0$  s. d.  $B_\delta(a) \subset U$ . Und somit

$$f(a+h) = f(a) + \underbrace{\langle \text{grad } f(a), h \rangle}_{=0} + \frac{1}{2} \langle h, Ah \rangle + R_2(h) \quad \forall h \in B_\delta(0)$$

mit  $\frac{R_2(h)}{\|h\|^2} \rightarrow 0$  ( $h \rightarrow 0$ ).

$$\implies \forall \varepsilon > 0 \quad \exists \tilde{\delta} > 0 \quad (\tilde{\delta} < \delta) \text{ s. d. } |R_2(h)| \leq \varepsilon \|h\|^2 \quad \forall h \in B_{\tilde{\delta}}(0)$$

3.32.a) Sei  $A$  positiv definit. Betrachte  $S = \{ h \in \mathbb{R}^n \mid \|h\|_E = 1 \}$ .  $S$  ist abgeschlossen und beschränkt  $\subset \mathbb{R}^n \implies S$  ist kompakt  $\implies S \ni h \mapsto \langle h, Ah \rangle$  nimmt Maximum und Minimum an (da stetig).

$$\alpha := \inf_{\substack{h \in S \\ > 0 \text{ (VOR)}}} \underbrace{\langle h, Ah \rangle}_{(h \neq 0 \text{ wenn } h \in S)} > 0.$$

**Behauptung.**

$$\langle h, Ah \rangle \geq \alpha \|h\|_E^2 \quad \forall h \in \mathbb{R}^n.$$

*Beweis.*  $h = 0$ : ✓. Sei also  $h \neq 0$ . Setze  $\tilde{h} := \frac{1}{\|h\|_E} h$ .

$$\implies \tilde{h} \in S \implies \langle \tilde{h}, A\tilde{h} \rangle \geq \alpha.$$

Die Behauptung folgt wegen

$$\langle \tilde{h}, A\tilde{h} \rangle = \frac{1}{\|h\|_E^2} \langle h, Ah \rangle. \quad \square$$

Wähle nun  $\tilde{\delta} > 0$  s. d.  $|R_2(h)| \leq \frac{\alpha}{4} \|h\|_E^2$  für  $\|h\| < \tilde{\delta}$ .

$$\implies f(a+h) \geq f(a) + \underbrace{\frac{1}{2} \langle h, Ah \rangle}_{\geq \alpha \|h\|_E^2} - |R_2(h)| \geq f(a) + \underbrace{\frac{1}{4} \alpha \|h\|_E^2}_{\geq 0} > f(a) \quad \forall \|h\| < \tilde{\delta}, \quad h \neq 0.$$

3.32.b) Ist  $A$  negativ definit, betrachte  $-f$  und wende 3.32.a) an.

3.32.c) Ist  $A$  indefinit, gibt es in jeder Umgebung  $V$  von  $a$  Punkte  $x, x' \in V$  s. d.

$$f(x) < f(a) < f'(x).$$

Es gilt  $v \in \mathbb{R}^n$ , s. d.  $\langle v, Av \rangle =: \alpha > 0$ .

$$\implies f(a+tv) = f(a) + \frac{1}{2} t^2 a + R_2(tv)$$

für  $|t|$  klein genug (s. d.  $tv \in B_\delta(0)$ ). Wähle  $|t| > 0$  so klein, dass zudem  $R_2(tv) \leq \frac{\alpha}{2} t^2$

$$\implies f(a+sv) \geq f(a) + \frac{1}{2} s^2 \alpha - |R_2(sv)| > f(a)$$

für  $0 < s < t$  (o.B.d.A.  $t > 0$ ).

Genauso: Es gibt  $v \in \mathbb{R}^n$ , s. d.  $\langle v, Av \rangle =: \alpha < 0$

$$\implies f(a+tv) = f(a) - \frac{1}{2} t^2 |\alpha| + R_2(tv)$$

für  $|t|$  klein genug (s. d.  $tv \in B_\delta(0)$ ). Wähle  $|t| > 0$  so klein, dass zudem  $|R_2(tv)| \leq \frac{\alpha}{4} t^2$

$$\implies f(a+sv) \leq f(a) - \frac{1}{2} s^2 |\alpha| + |R_2(sv)| < f(a) \quad \forall 0 < s < t \quad (\text{o.B.d.A. } t > 0). \quad \square$$

**Beispiele.** i)  $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ ,  $f(x) = x_1^2 + x_2^2 + 2$ ,  $\text{grad } f(a) = (2a_1, 2a_2) = 0$  nur für  $a = 0$ .

$$H_f(0) = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$$

positiv definit  $\implies$  isoliertes Minimum in 0.

ii)  $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ ,  $f(x) = 2 - x_1^2 - x_2^2$ ,  $\text{grad } f(a) = -2a = 0$  nur für  $a = 0$ .

$$H_f(0) = \begin{pmatrix} -2 & 0 \\ 0 & -2 \end{pmatrix}$$

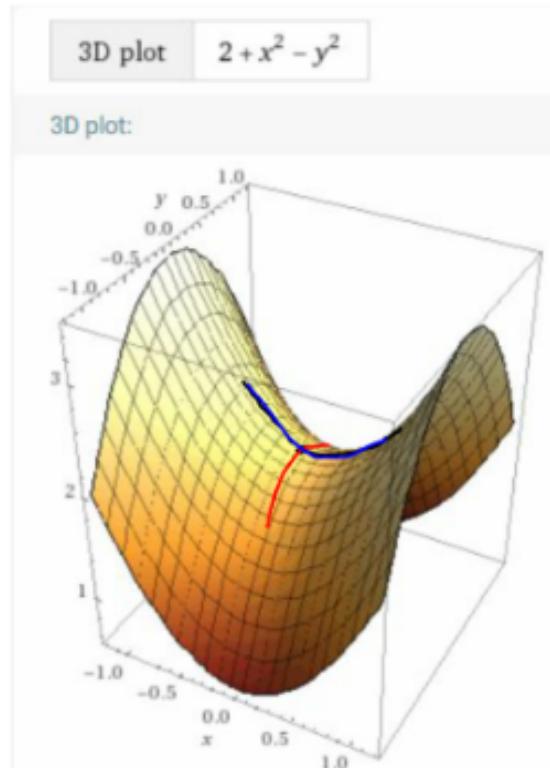
negativ definit  $\implies$  isoliertes Maximum in 0.

iii)  $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ ,  $f(x) = 2 + x_1^2 - x_2^2$ ,  $\text{grad } f(a) = 2a_1 - 2a_2 = 0$  nur für  $a = 0$ .

$$H_f(0) = \begin{pmatrix} 2 & 0 \\ 0 & -2 \end{pmatrix}$$

indefinit:

$$\begin{aligned} \langle e_1, H_f(0)e_1 \rangle &= 2 = -\langle e_2, H_f(0)e_2 \rangle \\ f(0 + te_1) &> f(0) \quad \forall t \neq 0 \\ f(0 + te_2) &< f(0) \quad \forall t \neq 0. \end{aligned}$$



Sattelfläche

**Bemerkung.** Wie untersucht man in komplizierten Fällen  $H_f(a)$  auf Definitheit? Kennt man die Eigenwerte ( $\exists$  wegen Symmetrie!) ist klar: Sind all positiv (negativ), so ist  $H_f(a)$  positiv (negativ) definit, gibt es positive und negative Eigenwerte, so ist sie indefinit. Aber Eigenwerte großer Matrizen sind schwer zu bestimmen. Besser geeignet ist

**Lemma 3.33 (aus der AGLA: Kriterium von Hurwitz).** Sei  $A \in \text{Mat}(n \times n, \mathbb{R})$  symmetrisch. Dann ist  $A$  genau dann positiv definit, wenn

$$\det \begin{pmatrix} A_{11} & \cdots & A_{1k} \\ \vdots & \ddots & \vdots \\ A_{k1} & \cdots & A_{kk} \end{pmatrix} > 0 \quad \forall k \in \{1, \dots, n\}$$

↑  
Haupt-Minoren von  $A$

und genau dann negativ definit, wenn

$$\det(A_{11}) = A_{11} < 0$$

$$\det \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} > 0$$

und dann immer abwechselnd.

# Kapitel 4.

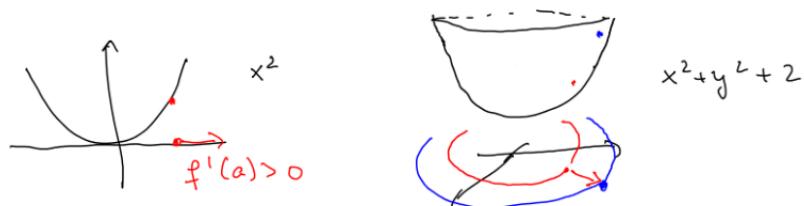
## Untermannigfaltigkeiten des $\mathbb{R}^n$

### Vorlesung 12

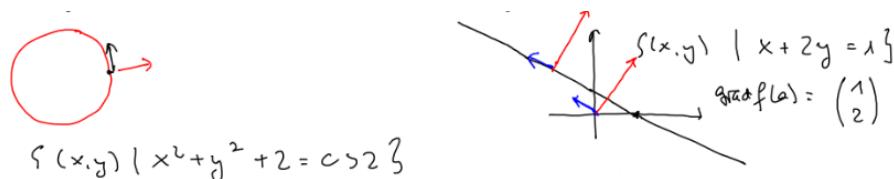
Do 28.05. 10:15

Ziel (u. a.) Geometrische Interpretation der Ableitung.

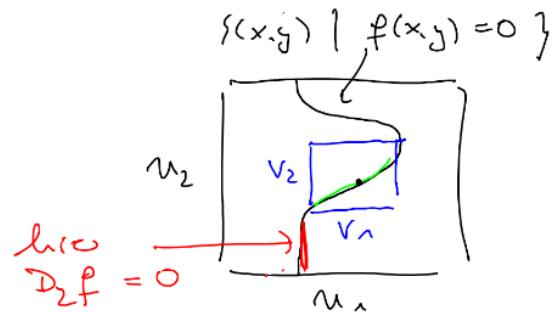
**Erinnerung.** Der Gradient zeigt in die Richtung des stärksten Anstiegs.



In einem Beispiel hatten wir gesehen: Richtungsableitung in Richtung einer Niveaufläche steht  $\perp$  auf  $\text{grad } f$ .

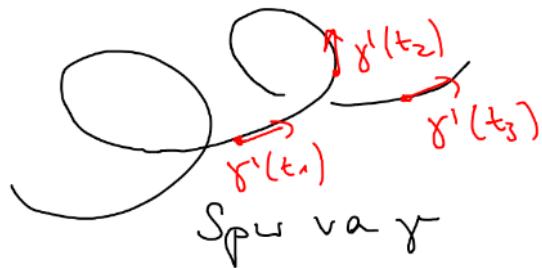


Satz über implizite Funktion:



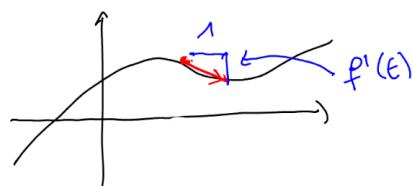
$f: U_1 \times U_2 \rightarrow \mathbb{R}$ . Niveaumenge kann lokal als Graph geschrieben werden, wenn  $D_2 f(a, b)$  invertierbar ist. Weitere Erklärungen → Audio.

**Aufwärm-Übung.** Sei  $\gamma: I \rightarrow \mathbb{R}^3$  differenzierbare Kurve.



Dann ist  $\gamma'(t)$  der *Geschwindigkeitsvektor*, der sich im Punkt  $\gamma(t)$  auf die Kurve schmiegt.

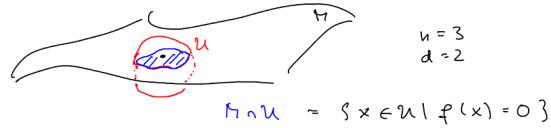
**Beispiel.**  $f: I \rightarrow \mathbb{R}$  differenzierbare Funktion. Betrachte  $\gamma(t) := (t, f(t))$ . Also Spur von  $\gamma = \Gamma_f$ .  $\gamma'(t) = (1, f'(t))$ .



**Definition 4.1.** Eine Teilmenge  $M \subset \mathbb{R}^n$  heißt  $d$ -dimensionale differenzierbare *Untermannigfaltigkeit* (oder reguläre Fläche) des  $\mathbb{R}^n$ , falls es für jedes  $a \in M$  eine Umgebung  $U \subset \mathbb{R}^n$  von  $a$  gibt und eine stetig differenzierbare Abbildung  $f: U \rightarrow \mathbb{R}^{n-d}$  s. d.

- a)  $M \cap U = \{x \in U \mid f(x) = 0\}$  und
- b)  $\text{rang } Df(a) = n - d$  (maximal).

Das heißt  $M$  lässt sich lokal als Nullstellengebilde von  $n - d$   $\mathbb{R}$ -wertigen  $C^1$  Funktionen  $f_j$  schreiben, deren Gradienten linear unabhängig sind.



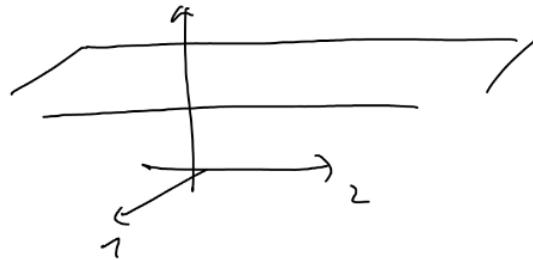
**Beispiele.** i) Die  $n$ -dimensionalen Untermannigfaltigkeiten des  $\mathbb{R}^n$  sind die offenen Teilmengen. Hier ist

$$M \cap U = U = \{x \in U\}.$$

ii) Ebene

$$E = \{p + (s, t, 0, 0) \mid s, t \in \mathbb{R}\} \subset \mathbb{R}^4,$$

$$p \in \mathbb{R}^4.$$

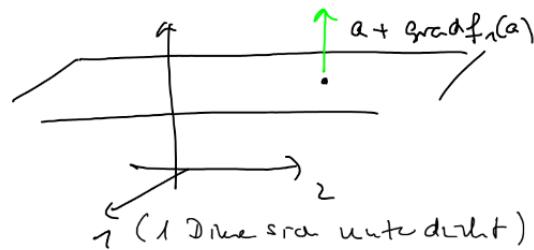


Eine Dimension im Bild unterdrückt.

Zu  $a \in E$  wähle  $U = \mathbb{R}^4$  und setze

$$\begin{aligned} f_1: U &\rightarrow \mathbb{R} & f_1(x) &= \langle x - p, e_3 \rangle = x_3 - p_3 \\ f_2: U &\rightarrow \mathbb{R} & f_2(x) &= \langle x - p, e_4 \rangle = x_4 - p_4. \end{aligned}$$

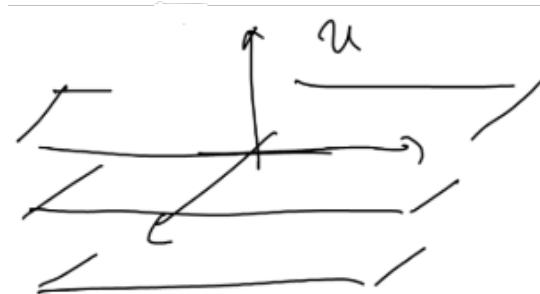
Es gilt  $x \in E \iff f(x) = \begin{pmatrix} f_1(x) \\ f_2(x) \end{pmatrix} = 0$  und  $Df(a) = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$ , rang = 2.



$$\text{grad } f_1(a) = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} \quad \text{grad } f_2(a) = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}.$$

Niveau-Mengen von  $f$ :

$$N_f(c) = \{ x \in \mathbb{R}^n \mid x_3 - p_3 = c, x_4 - p_4 = c \}.$$



$$N_f(0) = E, \text{ Gradient steht senkrecht.}$$

- iii)  $\mathbb{S}^{n-1} := \{ x \in \mathbb{R}^n \mid \|x\|_{\text{E}} = 1 \}$ . Einheits-Sphäre. Zu  $a \in \mathbb{S}^{n-1}$  wähle  $U = \mathbb{R}^n$  und  $f(x) = \|x\|_{\text{E}}^2 - 1$ .

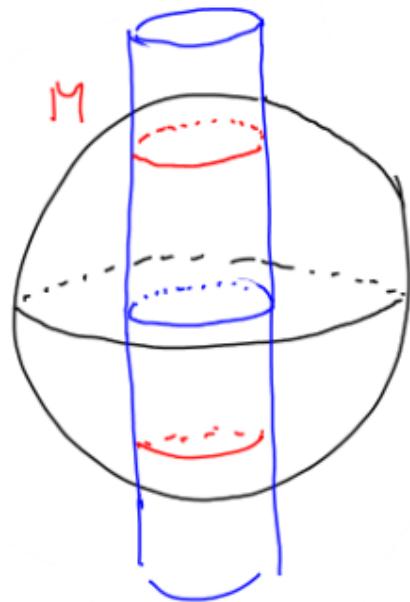
$$Df(a) = (2a_1, \dots, 2a_n) \neq 0.,$$

also  $\text{rang} = 1 \forall a \in \mathbb{S}^{n-1}$ .

- iv) Eine 1-dimensionale differenzierbare Untermannigfaltigkeit wird auch unparametrisierte differenzierbare Kurve genannt.

**Beispiel.**  $f: \mathbb{R}^3 \rightarrow \mathbb{R}^2$ .

$$f(x, y, z) = \begin{pmatrix} x^2 + y^2 + z^2 - R^2 \\ x^2 + y^2 - r^2 \end{pmatrix} \quad R > r > 0.$$



$$M = \{ (x, y, z) \mid f(x, y, z) = 0 \} = N_f(0)$$

$$Df(x, y, z) = \begin{pmatrix} 2x & 2y & 2z \\ 2x & 2y & 0 \end{pmatrix}$$

hat maximalen Rang 2 für  $z \neq 0$  und  $(x, y) \neq 0$ .

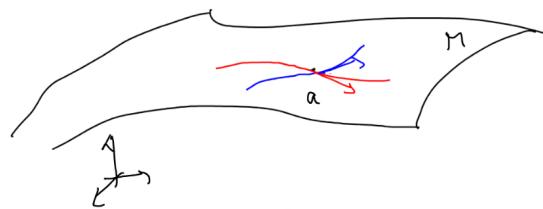
Letzteres ist für alle  $a \in M$  gewährleistet, weil  $(x^2 + y^2 = r^2 > 0)$  gilt und somit ist auch ersteres gewährleistet, da  $z^2 = -r^2 + R^2 > 0$ .

Man sieht auch direkt: Die Funktion, die  $M$  lokal als Nullstellenmenge beschreibt, ist nicht eindeutig: z. B.

$$\tilde{f}(x, y, z) = \begin{pmatrix} z^2 + r^2 - R^2 \\ x^2 + y^2 - r^2 \end{pmatrix}$$

beschreibt im Beispiel oben das selbe  $M$ .

**Definition 4.2.** Ein Vektor  $v \in \mathbb{R}^n$  heißt *Tangentialvektor* der Untermannigfaltigkeit  $M \subset \mathbb{R}^n$  im Punkt  $a$ , falls  $\exists$  differenzierbare Kurve  $\gamma$ ,  $\gamma: (-\varepsilon, \varepsilon) \rightarrow M$ , s. d.  $\gamma(0) = a$  und  $v = \gamma'(0)$ . Die Menge der Tangentialvektoren im Punkt  $a$  bezeichnet man mit  $T_a M$ .



Man zeichnet Tangentialvektoren an den Punkt  $a$  an. Aber eigentlich sind es *keine* affinen Vektoren.

**Beispiele.** i) Ebene

$$E = \{ p + (s, t, 0, 0) \mid s, t \in \mathbb{R} \} \subset \mathbb{R}^4,$$

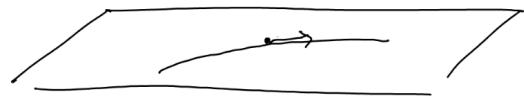
$$p \in \mathbb{R}^4, a \in E.$$

$$\text{Ta } M = \left\{ v \in \mathbb{R}^4 \mid v = v_1 e_2 + v_2 e_2 \right\},$$

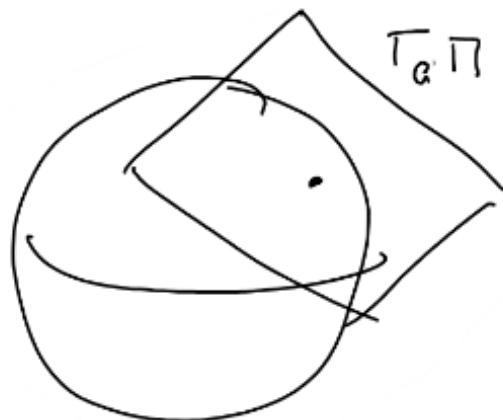
denn jedes  $\gamma: (-\varepsilon, \varepsilon) \rightarrow M$  mit  $\gamma(0)$  ist von der Form

$$\gamma(t) = a + (\tilde{\gamma}_1(t), \tilde{\gamma}_2(t), 0, 0).$$

Man zeichnet den Tangentialraum an der Stelle  $a$  ein, also



ii)  $M = \mathbf{S}^{n-1}$ . Behauptung:



Das erklärt den Namen: Die Tangentialvektoren sind tangential (angeschmiegt) an  $M$  im Punkt  $a$ .

Zum Beweis der Behauptung beweisen wir allgemeiner:

24 min

**Satz 4.3.** Sei  $M$   $d$ -dimensionale differenzierbare Untermannigfaltigkeit des  $\mathbb{R}^n$ . Ist

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

für eine Funktion  $f: U \rightarrow \mathbb{R}^{n-d}$  wie in Definition 4.1, also  $Df(a)$  von maximalem Rang ( $n-d$ ). Dann ist  $T_a M = \ker -Df(a)$ . Insbesondere ist  $T_a M$  Untervektorraum von  $\mathbb{R}^n$ .

*Beweis.* „ $\subset$ “ Sei  $\gamma: (-\varepsilon, \varepsilon) \rightarrow M \cap U$  mit  $\gamma(0) = a$ ,  $\gamma'(0) = v$ . Dann ist

$$0 = \frac{d}{dt} f \circ \gamma(t) \Big|_{t=0} = Df(\gamma(0)) \cdot \gamma'(0) = Df(a) \cdot v \implies v \in \ker -Df(a).$$

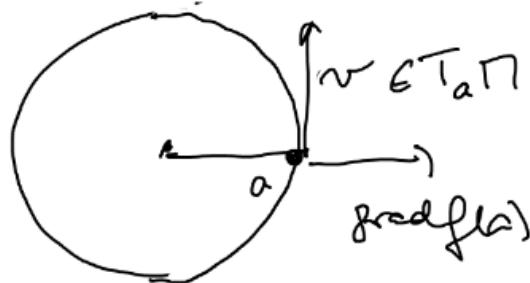
„ $\supset$ “ Beweisen wir später mit Hilfe des folgenden Satzes 4.4.  $\square$

**Interpretation (für  $n-d=1$ ).** Der Gradient zeigt in die Richtung des stärksten Anstiegs von  $f$ . Eine Kurve  $\gamma$  wi in 4.2 (und 4.3) verläuft in einer Niveaumenge von  $f$  (und zwar hier zu  $c=0$ ). In Richtung  $v = \gamma'(0)$  ändert sich  $f$  also nicht und die Richtungsableitung in Richtung  $v$  ist 0:

$$\partial_v f(a) = \langle v, \text{grad } f(a) \rangle = Df(a) \cdot v = 0 \quad (\text{Satz 4.3}),$$

also  $v \perp \text{grad } f(a)$ . „Der Gradient steht auf Niveaumengen senkrecht“.

**Beispiel.**  $f(x, y) = x^2 + y^2 - 2$ .  $M = \text{Niveaumenge } N_f(0) = \{x^2 + y^2 = 2\}$ .  $\text{grad } f(a) = 2(a_1, a_2)$ .



Eine Kurve, die in  $M$  verläuft und für die gilt

$$\gamma(0) = a = \begin{pmatrix} \sqrt{2} \\ 0 \end{pmatrix}$$

ist

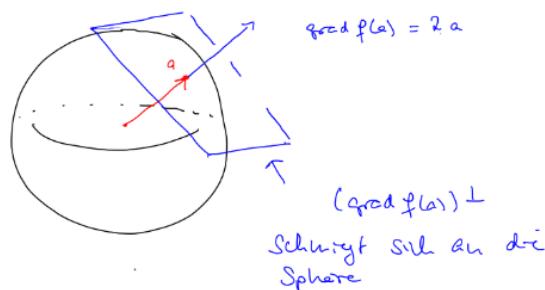
$$\gamma(t) = \begin{pmatrix} \cos t \\ \sin t \end{pmatrix} \cdot \sqrt{2}.$$

Es ist dann

$$\gamma'(0) = \sqrt{2} \begin{pmatrix} 0 \\ 1 \end{pmatrix} \perp \text{grad } f \begin{pmatrix} \sqrt{2} \\ 0 \end{pmatrix} = \begin{pmatrix} 2\sqrt{2} \\ 0 \end{pmatrix}.$$

## Höher-dimensionale Sphären:

$$N_f(0), f(x) = \|x\|_{\mathbb{E}}^2 - 1. Df(a) = 2(a_1, \dots, a_n). \ker -Df(a) = \text{Ebene senkrecht zu } a.$$



Um die Rückrichtung von Satz 4.3 beweisen zu können, benötigen wir eine weitere (äquivalente) Charakterisierung differenzierbarer Untermannigfaltigkeiten.

**Satz 4.4.** Eine Teilmenge  $M \subset \mathbb{R}^n$  ist genau dann  $d$ -dimensionale differenzierbare Untermannigfaltigkeit des  $\mathbb{R}^n$ , wenn es zu jedem Punkt  $a \in M$  (nach eventueller Umnummerierung) Umgebungen  $V_1 \subset \mathbb{R}^d$  von  $a' = (a_1, \dots, a_d)$  und  $V_2 \subset \mathbb{R}^{n-d}$  von  $a'' = (a_{d+1}, \dots, a_n)$  und eine stetig differenzierbare Abbildung  $g: V_1 \rightarrow \mathbb{R}^{n-d}$ ,  $g(V_1) \subset V_2$  gibt, s. d.

$$M \cap (V_1 \times V_2) = \Gamma_q = \{ (x', g(x')) \mid x' \in V_1 \}.$$

*Beweis.* „ $\implies$ “ Sei  $a \in M$ ,  $U$ ,  $f$  wie in Definition 4.1, insbesondere  $\text{rang } Df(a) = n - d$  (maximal)  $\implies$  nach eventueller Umnummerierung ist  $Df(x', x'')$  für  $x = (x', x'') \in U \subset \mathbb{R}^d \times \mathbb{R}^{n-d}$

$$\begin{pmatrix} \partial_1 f_1(x', x'') \cdots \partial_d f_1(x'.x'') & \partial_{d+1} f_1(x', x'') \cdots \partial_n f_1(x'; x'') \cdots \\ \vdots & \vdots \\ \partial_1 f_{n-d}(x', x'') \cdots \partial_d f_{n-d}(x'.x'') & \partial_{d+1} f_{n-d}(x', x'') \cdots \partial_n f_{n-d}(x', x'') \cdots \\ \cdots & D_2 f(x'.x'') \end{pmatrix},$$

s. d.  $D_2 f(a', a'')$  invertierbar ist.

Zudem ist  $f(a) = 0$ . Der Satz über implizite Funktion liefert also eine lokale Auflösung von  $N_f(0) \implies$  Beh.

„ $\Leftarrow$ “ Sei

$$M \cap V_1 \times V_2 = \{ (x', g(x')) \mid x' \in V_1 \},$$

$V_1 \subset \mathbb{R}^d$ ,  $V_2 \subset \mathbb{R}^{n-d}$ . Setze

$$\begin{aligned} f: V_1 \times V_2 &\rightarrow \mathbb{R}^{n-d} \\ f_j(x', x'') &= g_j(x') - x''_j. \end{aligned}$$

Dann ist

$$M \cap V_1 \times V_2 = \{ x \in V_1 \times V_2 \mid f(x) = 0 \}$$

und

$$Df(x) = \left( \begin{array}{ccc|c} \partial_1 g_1(x) & \dots & \partial_d g_1(x) & -1 \\ \vdots & & \vdots & 0 \\ \partial_1 g_{n-d}(x) & \dots & \partial_d g_{n-d}(x) & 0 \end{array} \right)$$

hat maximalen Rang  $(n - d)$ .  $\square$

**Bemerkung.** Man kann also eine  $d$ -dimensionale differenzierbare Untermannigfaltigkeit lokal stets als Graphen einer  $C^1$ -Funktion schreiben.

Für die Beispiele oben gilt

**4.1.i)**  $U \subset \mathbb{R}^n$  offen,  $V_2 = \emptyset$ ,  $V_1 = U$ .

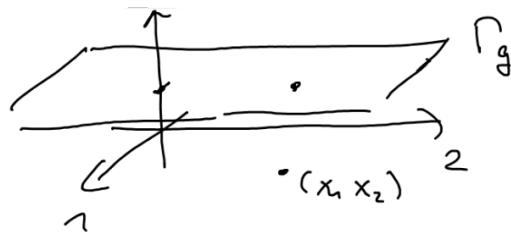
**4.1.ii)** Ebene

$$E = \{ p + se_1 + te_2 \mid s, t \in \mathbb{R} \} \subset \mathbb{R}^4,$$

$$p \in \mathbb{R}^4.$$

$$g: \mathbb{R}^2 \rightarrow \mathbb{R}^2 \quad g(x_1, x_2) = \begin{pmatrix} p_3 \\ p_4 \end{pmatrix}$$

konstant.



**4.1.iii)  $\mathbf{S}^{n-1}$ :** Sei  $a \in \mathbf{S}^{n-1}$  mit  $a_n > 0$ . Sei

$$V_1 = \left\{ x' \in \mathbb{R}^{n-1} \mid \|x'\|_{\mathbb{E}} < 1 \right\}$$

$$g: V_1 \rightarrow \mathbb{R}, \quad g(x_1, \dots, x_{n-1}) = \sqrt{1 - x_1^2 - \dots - x_{n-1}^2}.$$

Dann ist

$$\mathbf{S}^{n-1} \cap V_1 \times (0, \infty) = \Gamma_g.$$



40 min

*Fortsetzung Beweis Satz 4.3.* Wir beweisen nun die zweite Inklusion von Satz 4.3, also  $\text{Ta } M \supset \ker -Df(a)$ . Sei  $g: V_1 \rightarrow \mathbb{R}^{n-d}$  stetig differenzierbar s.d.  $g(V_1) \subset V_2$  und  $a \in V_1 \times V_2 \subset \mathbb{R}^d \times \mathbb{R}^{n-d}$  und

$$M \cap (V_1 \times V_2) = \Gamma_g = \{ (x, g(x)) \mid x \in V_1 \}.$$

Sei  $x_0 \in V_1$  s.d.  $(x_0, g(x_0)) = a$ . Betrachte  $\varphi: V_1 \rightarrow \mathbb{R}^n$ ,  $\varphi(x) = (x, g(x))$ . Dann ist

$$D\varphi(x_0) \cdot v = \frac{d}{dt} \varphi(x_0 + tv) \Big|_{t=0} \quad \forall v \in \mathbb{R}^d.$$

↑  
Kettenregel

$\implies a = \varphi(x - 0)$  und

$$w := \frac{d}{dt} \varphi(x_0 + tv) \Big|_{t=0} \in \text{Ta } M,$$

also  $\text{Im}(D\varphi(x_0)) \subset \text{Ta } M$ . Zudem wissen wir schon:  $\text{Ta } M \subset \ker -Df(a)$ . Es ist  $D\varphi(x_0): \mathbb{R}^d \rightarrow \mathbb{R}^n$ ,  $D\varphi(x_0) = \begin{pmatrix} \mathbb{1}_{d \times d} \\ Dg(x_0) \end{pmatrix}$ , also injektiv  $\implies$  (Dimensionsformel)

$$\dim \text{Im}(D\varphi(x_0)) = d = n - \underbrace{\text{rang } Df(a)}_{=n-d} = \dim \ker -Df(a)$$

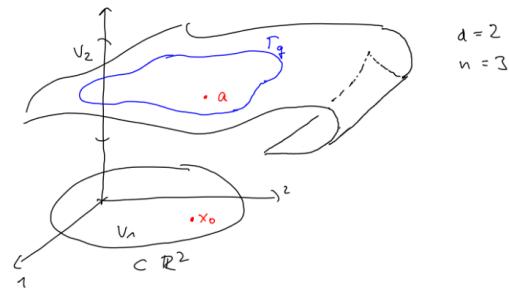
$\implies \text{Ta } M = \ker -Df(a)$ .  $\square$

Wir haben im Beweis auch gesehen:

$$\mathrm{Ta}\, M = \mathrm{Im}\, D\varphi(x_0),$$

wo  $\phi(x) = (x, g(x))^T$  die Graphenabbildung einer lokalen Auflösung  $g$  ist it  $a = (x_0, g(x_0))^T = \phi(x_0)$ .

Die geometrische Interpretation ist:



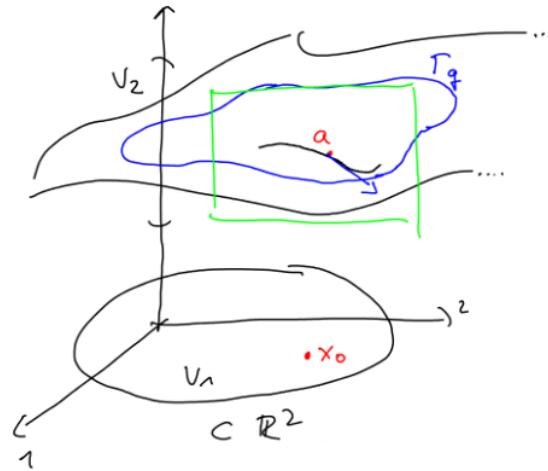
$$D\phi(x_0) = \begin{pmatrix} \mathbb{1}_2 \\ Dg(x_0) \end{pmatrix}.$$

$\uparrow$   
 $\partial_1 g(x_0) \quad \partial_2 g(x_0)$

Die Spalten von  $D\phi(x_0)$

$$\begin{pmatrix} 1 \\ 0 \\ \partial_1 g(x_0) \end{pmatrix} \text{ und } \begin{pmatrix} 0 \\ 1 \\ \partial_2 g(x_0) \end{pmatrix}$$

↑    ↑  
 Veränderung von  $g$  in 1-Richtung                  Veränderung in 2-Richtung



Ebene parallel zur 2-3-Ebene durch  $a = \{x_1 = x_{0,1} \text{ konstant}\}$ .  $\curvearrowright = \{(x_{0,1}, x_{0,2} + t, g(x_{0,1}, x_{0,2} + t)) \mid t \in I \subset \mathbb{R}\}$  ist Kurve in  $M$ .

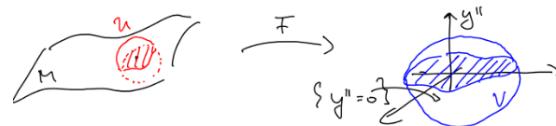
Ableitung

$$\frac{d}{dt} \phi(x_{0,1}, x_{0,2} + t) = (0, 1, \partial_2 g(x_{0,1}, x_{0,2} + t)).$$

Es gilt  $\phi(x_{0,1}, x_{0,2} + t)|_t$  und  $(0, 1, \partial_2 g(x_0))$  ist **Tangentialvektor** bei  $a$ . Genauso für die Ebene  $\{x_2 = x_{0,2}\}$  und die erste Spalte von  $D\phi(x_0)$ . Jede  $d$ -dimensionale Untermannigfaltigkeit lässt sich lokal zu einer Teilmenge einer  $d$ -dimensionalen Ebene im  $\mathbb{R}^n$  „plätzen“,

**Satz 4.5.**  $M \subset \mathbb{R}^n$  ist genau dann  $d$ -dimensionale Untermannigfaltigkeit, wenn es zu jedem Punkt  $a \in M$  offene Umgebungen  $U$  von  $a \subset \mathbb{R}^n$  und  $V$  von  $0 \subset \mathbb{R}^n$  und einen **Diffeomorphismus** von  $U$  nach  $V$ , also eine stetig differenzierbare invertierbare Funktion  $F: U \rightarrow \mathbb{R}^n$  mit stetig differenzierbarer Umkehrfunktion  $G: V \rightarrow \mathbb{R}^n$  s. d.

$$F(M \cap U) = \left\{ (y', y'') \in \mathbb{R}^d \times \mathbb{R}^{n-d} \mid y'' = 0 \right\} \cap V.$$



*Beweis.* „ $\Rightarrow$ “ Sei  $M$  in einer Umgebung von  $a$  als Graph dargestellt ( $\exists$  nach Satz 4.4),

$$M \cap (V_1 \times V_2) = \{(x', g(x')) \mid x' \in V_1\}$$

Setze  $U := V_1 \times V_2$  und  $F(x, x'') = (x', x'' - g(x'))$ .  $F$  ist stetig differenzierbar und die Umkehrfunktion

$$F^{-1}(y', y'') = (y', y'' + g(y'))$$

ist ebenfalls stetig differenzierbar,  $F^{-1}: F(U) \rightarrow \mathbb{R}^n$ .  $F(U)$  ist offen nach Satz 3.21 weil

$$DF(x'.x'') = \begin{pmatrix} \mathbb{1}_{d \times d} & -Dg(x') \\ 0 & \mathbb{1}_{(n-d) \times (n-d)} \end{pmatrix}$$

invertierbar ist für alle  $(x', x'') \in U$ . Es ist  $F(M \cap U) = \{(x', 0) \mid x' \in V_1\}$ .

Es ist

$$F(M \cap U) = \{(x', 0) \mid x' \in V_1\}.$$

„ $\Leftarrow$ “ Sei  $F: U \rightarrow \mathbb{R}^n$  Diffeomorphismus,  $F(U) = V$ , mit

$$F(M \cap U) = \{y \in V \mid y_{d+1} = \dots = y_n = 0\}.$$

Dann ist

$$\begin{aligned} M \cap U &= F^{-1}(\{y \in V \mid y_{d+1} = \dots = y_n = 0\}) \\ &= \{x \in U \mid F_{d+1}(x) = \dots = F_n(x) = 0\}. \end{aligned}$$

Der Ran von

$$\begin{pmatrix} \partial_1 F_{d+1}(a) & \dots & \partial_n F_{d+1}(a) \\ \vdots & & \vdots \\ \partial_1 F_N(a) & \dots & \partial_n F_{d+1}(a) \end{pmatrix}$$

ist maximal (also gleich  $n - d$ ), weil die Ableitungsmatrix  $DF(a) \in \text{Mat}(n \times n, \mathbb{R})$  invertierbar ist  $\implies M$  ist  $d$ -dimensionale Untermannigfaltigkeit.  $\square$

**Beispiel.**  $\mathbb{S}^{n-1}$ ,  $a \in \mathbb{S}^{n-1}$ ,  $a_n > 0$ .  $g: D \rightarrow \mathbb{R}$  mit  $D = \{x \in \mathbb{R}^{n-1} \mid \|x\|_E < 1\}$ ,

$$g(x_1, \dots, x_{n-1}) = \sqrt{1 - \sum_{i=1}^{n-1} x_i^2}.$$

$F: D \times (0, \infty) \rightarrow \mathbb{R}^n$ ,

$$F(x, t) = \left( x, t - \sqrt{1 - \sum_{i=1}^{m_1} x_i^2} \right).$$

$\mathbb{S}^{n-1} \cap D \times (0, \infty) = \text{---} \circlearrowleft$  (ohne den Rand  $\text{---} \circlearrowright$ ) und

$$F(M \cap D \times (0, \infty)) = \text{---} \circlearrowright = D.$$

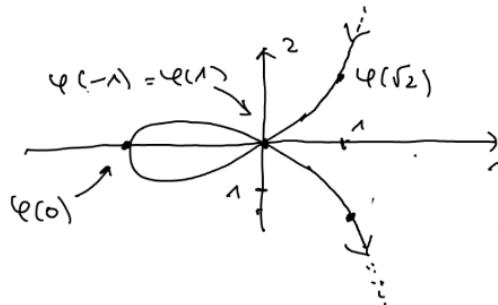
**Vorlesung 13**

Letztes Mal: Untermannigfaltigkeiten *lokal* als Nullstellengebilde (Definition), als Graph und als diffeomorph zu Teilmengen des  $\mathbb{R}^d \times \{0\}$ .

Eine weitere äquivalente lokale Beschreibung mit Hilfe einer Parametrisierung:

**Definition 4.6.** Eine Abbildung  $\varphi: V \rightarrow \mathbb{R}^n$ ,  $V \subset \mathbb{R}^d$  offen,  $d \leq n$ , heißt *Immersion*, falls sie stetig differenzierbar ist und der Rang von  $D\varphi(t)$  für alle  $t \in V$  maximal, also gleich  $d$  ist. Man nennt  $\varphi(V)$  die *Spur* von  $\varphi$ . Ist  $\varphi: V \rightarrow \varphi(V)$  ein Homöomorphismus (also stetig, bijektiv, mit stetigem Inversen), so heißt  $\varphi$  *Einbettung*.

**Beispiele.** i)  $V = \mathbb{R}$ ,  $\phi(t) = (t^2 - 1, t(t^2 - 1))$ .



$$\varphi(1) = (0, 0) = \phi(-1) \quad \text{also keine Einbettung}$$

$$\varphi(0) = (-1, 0)$$

$$\varphi(\sqrt{2}) = (1, \sqrt{2})$$

$$\varphi(-t) = (\phi_1(t), -\phi_2(t))$$

$$\varphi'(t) = \left( \underbrace{3t^2 - 1}_{\neq 0 \text{ für } t \neq \pm \frac{1}{\sqrt{3}}} \right) \neq 0 \implies \text{maximaler Rang.}$$

Eine Immersion  $\varphi: I \subset \mathbb{R} \rightarrow \mathbb{R}^n$  reguläre (parametisierte) Kurve.

**Bemerkung.** 1-dimensional  $\rightarrow I$  kann auch nicht offen gewählt werden, wenn man nur an Differenzierbarkeit interessiert ist. Hier *nicht*.

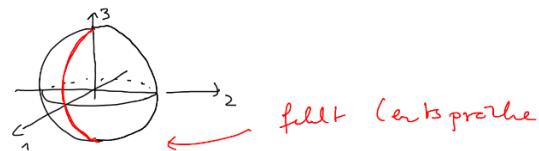
ii)  $V = \{ t \in \mathbb{R}^2 \mid \|t\|_E < 1 \}, \varphi(t_1, t_2) = (t_1, t_2, \sqrt{1 - t_1^2 - t_2^2})^T,$

$$D\varphi(t_1, t_2) = \begin{pmatrix} \mathbb{1}_{2 \times 2} \\ \frac{-t_1}{\sqrt{\dots}} & \frac{-t_2}{\sqrt{\dots}} \end{pmatrix} \quad \text{rang } 2 \quad \forall t \in V.$$

Spur von  $\varphi$ .

iii)  $V = (0, \pi) \times (0, 2\pi),$

$$\varphi(\beta, \alpha) = (\cos \alpha \sin \beta, \sin \alpha \sin \beta, \cos \beta).$$

Spur von  $\varphi$ 

$V = \mathbb{R}^2$ ,  $\varphi$  wie oben:  $\varphi$  schneidet sich  $\infty$  oft, da  $\varphi(\beta + 2\pi\mathbb{Z}, \alpha + 2\pi\mathbb{Z}) = \phi(\beta, \alpha)$  gilt.  $\varphi(\mathbb{R}^2) = \mathbb{S}^2$ .

**Satz 4.7.** Sei  $\varphi: V \rightarrow \mathbb{R}^n$ ,  $V \subset \mathbb{R}^d$ , eine Immersion. Dann gibt es zu jede,  $t_0 \in V$  eine offene Umgebung  $V_0 \subset V$  s.d.  $\varphi|_{V_0}$  eine Einbettung des  $\mathbb{R}^n$  ist und  $\varphi(V_0)$  eine  $d$ -dimensionale Untermannigfaltigkeit des  $\mathbb{R}^n$ .

*Beweis.* Nach eventueller Umnummerierung der Komponenten von  $\varphi$ . Si

$$\begin{pmatrix} \partial_1 \varphi_1 & \dots & \partial_d \varphi_d(t_0) \\ \vdots & & \vdots \\ \partial_1 \varphi_d(t_0) & \dots & \partial_d \varphi_d(t_0) \end{pmatrix}$$

invertierbar.

Dann gibt es nach dem Satz über die Umkehrfunktion Satz 3.19 offene Umgebungen

$$V_0 \subset \mathbb{R}^d \text{ von } t_0 \text{ und } W \subset \mathbb{R}^d \text{ von } \begin{pmatrix} \varphi_1(t_0) \\ \vdots \\ \varphi_d(t_0) \end{pmatrix} \text{ s. d.}$$

$$\tilde{\varphi} = \begin{pmatrix} \varphi_1 \\ \vdots \\ \varphi_d \end{pmatrix}: V_0 \rightarrow W$$

Diffeomorphismus ist (also stetig differenzierbar mit stetig differenzierbarem Inversen). Wir zeigen, dass  $\varphi(V_0)$  diffeomorph ist zu  $V_0 \times \{0\}$ . Dann folgt aus Satz 4.5, dass  $\varphi(V_0)$   $d$ -dimensionale Untermannigfaltigkeit des  $\mathbb{R}^n$  ist.

*Beweis.* Setze dazu  $G: V_0 \times \mathbb{R}^{n-d} \rightarrow W \times \mathbb{R}^{n-d}$ .

$$G(\underbrace{t', t''}_t) := \begin{pmatrix} \varphi_1(t') \\ \vdots \\ \varphi_d(t') \\ \varphi_{d+1}(t') + t_{d+1} \\ \vdots \\ \varphi_{d+1}(t') + t_n \end{pmatrix} = \begin{pmatrix} \varphi_1(t') \\ \vdots \\ \tilde{\varphi}(t') + t'' \end{pmatrix} \quad \tilde{\varphi} = \begin{pmatrix} \varphi_{d+1} \\ \vdots \\ \varphi_n \end{pmatrix}.$$

Dann ist  $G(V_0 \times \mathbb{R}^{n-d}) = W \times \mathbb{R}^{n-d}$ , denn

„ $\subset$ “ ist klar, da  $\tilde{\varphi}: V_0 \rightarrow W$  Diffeomorphismus ist.

„ $\supset$ “ Sei  $(y, y'') \in W \times \mathbb{R}^{n-d}$ . Setze  $t' = \tilde{\varphi}^{-1}(y')$  ( $\tilde{\varphi}: V_0 \rightarrow W$  Diffeomorphismus),  $t'' = y'' - \tilde{\varphi}(t')$ . Dann gilt  $G(t', t'') = \begin{pmatrix} y' \\ y'' \end{pmatrix}$ .

Man sieht:  $G: V_0 \times \mathbb{R}^{n-d} \rightarrow W \times \mathbb{R}^{n-d}$  ist bijektiv. Man sieht auch sofort:  $G$  ist Diffeomorphismus, da  $\tilde{\varphi}$  Diffeomorphismus ist und  $\tilde{\varphi}$  stetig differenzierbar und somit auch die Umkehrfunktion

$$F(y', y'') = \begin{pmatrix} \tilde{\varphi}^{-1}(y') \\ y'' - \tilde{\varphi}(\tilde{\varphi}^{-1}(y)) \end{pmatrix}$$

stetig differenzierbar ist. Wegen  $G(V_0 \times \{0\}) = \varphi(V_0)$   $d$ -dimensionale Untermannigfaltigkeiten.  $\square$

Es folgt auch, dass  $\varphi|_{V_0}$  Einbettung ist:

Denn  $\tilde{\varphi}|_{V_0}$  ist Diffeomorphismus (auf  $W$ ) und somit ist  $\varphi|_{V_0}$  injektiv. Wegen  $\varphi(t) = G(t, 0)$  ist die Umkehrfunktion  $\varphi^{-1}: \varphi(V_0) \rightarrow \mathbb{R}^n$  stetig.  $\square$

**Beispiel (Beispiel i)).** Wegen  $\phi(1) = \varphi(-1)$  ist  $\varphi$  keine Einbettung. Es ist auch kein Untermannigfaltigkeit, was man hier leichter als in Aufgabe 2 auf Blatt 7 sieht: Es gibt 2 linear unabhängige Tangentialvektoren an  $(0, 0)$ , z. B.  $\varphi'(1) = (2, 2)$ ,  $\varphi'(-1) = (-2, 2)$ .

9:15

Wählt man aber etwa  $I_0 = \left(-\frac{1}{2}, \infty\right)$ , so erhält man



ohne Doppelpunkt.

Tatsächlich gilt:

**Satz 4.8.** Eine Teilmenge  $M \subset \mathbb{R}^n$  ist genau dann eine  $d$ -dimensionale Untermannigfaltigkeit des  $\mathbb{R}^n$ , wenn es zu jedem Punkt  $a \in M$  eine offene Umgebung  $U \subset \mathbb{R}^n$  von  $a$  gibt und eine offene Teilmenge  $V \subset \mathbb{R}^d$  und eine Einbettung  $\varphi: V \rightarrow \mathbb{R}^n$  mit  $\varphi(V) = M \cap U$ .

*Beweis.* „ $\Leftarrow$ “ Sei also  $a \in M$  und  $U$ ,  $\varphi: V \rightarrow \mathbb{R}^n$  wie oben,  $a > \varphi(t_0)$ . Dann gibt es (Satz 4.7) eine offene Umgebung  $V_0$  von  $t_0$  s. d.  $\varphi(V_0)$  eine  $d$ -dimensionale Untermannigfaltigkeit ist. Es ist  $\varphi(V_0) \subset M \cap U$  offen, (nach dem Satz von der offenen Abbildung)  
 $\Rightarrow$  Es gibt eine offene Umgebung  $U_0$  von  $\varphi(t_0) = a$  s. d.  $\varphi(V_0) = M \cap U_0 \Rightarrow M \cap U_0$  ist  $d$ -dimensionale Untermannigfaltigkeit. Da man dies für jeden Punkt  $a \in M$  machen kann, folgt:  $M$  ist  $d$ -dimensionale Untermannigfaltigkeit.

„ $\Rightarrow$ “ Wir schreiben  $M$  lokal als Graphen:  $a \in U = U_1 \times U_2 \subset \mathbb{R}^d \times \mathbb{R}^{n-d}$ ,

$$M \cap U = \{ (x', g(x')) \mid x' \in U_1 \}.$$

Setze  $V := U_1$  und  $\varphi: V \rightarrow \mathbb{R}^n$ ,  $\varphi(t) = t, g(t)^T$ . Dann ist  $\varphi$  eine Immersion (denn  $D\varphi(t) = \begin{pmatrix} \mathbf{1}_{d \times d} \\ Dg(t) \end{pmatrix}$  hat maximalen Rang) mit  $\varphi(V) = M \cap U$  (da  $g(V) = U_2$ ).  $\varphi$  ist injektiv und die Umkehrfunktion

$$\varphi^{-1}: \varphi(V) \rightarrow \mathbb{R}^d, \quad \underset{\substack{\parallel \\ U_1 \times U_2}}{(y', y'')} \mapsto y'$$

ist stetig.  $\square$

### Bemerkungen.

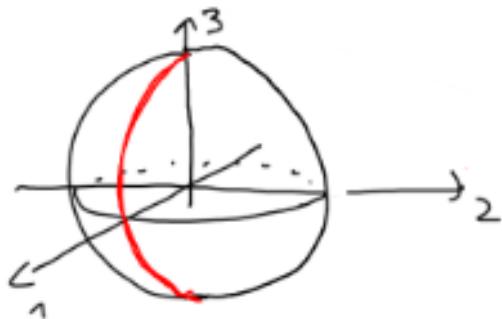
Man nennt  $\varphi$  wie oben eine *lokale Parametrisierung* von  $M$  bei  $a$ .  $V$  heißt *Parameterbereich* und  $M \cap U$  *Koordinatenumgebung* von  $a$ .

Globale Parametrisierungen  $\varphi: V \rightarrow \mathbb{R}^n$  mit  $\varphi: V \rightarrow M$ . Homöomorphismus existieren i. A. nicht.

**Beispiel.**  $M = \mathbb{S}^{n-1} \subset \mathbb{R}^n$ . Es kann keine stetige Abbildung  $\varphi^{-1}: \mathbb{S}^{n-1} \rightarrow V$  gebe, da  $\varphi^{-1}(\mathbb{S}^{n-1})$  kompakt wäre, aber  $V$  ist offen.

Im Beispiel iii) oben, erreicht man daher nicht die ganze  $\mathbb{S}^n$ .  $V = (0, \pi) \times (0, 2\pi)$ ,

$$\varphi(\beta, \alpha) = (\cos \alpha \sin \beta, \sin \alpha \sin \beta, \cos \beta).$$



Spur von  $\varphi$

oder ( $V = \mathbb{R}^2$ ) es liegt kein Homöomorphismus vor.

Für  $a \in \{(\sin \beta, 0, \cos \beta) \mid b \in (0, \pi)\}$  wählt man als lokale Parametrisierung einfach  $V = (0, \pi) \times (-\pi, \pi)$  (dann fehlt ein anderer Halbkreis).

Oder man wählt  $D \ni (y, z) \mapsto (\sqrt{1 - y^2 - z^2}, y, z)$  mit

$$D = \{(y, z) \in \mathbb{R}^2 \mid \|(y, z)\|_E < 1\}.$$

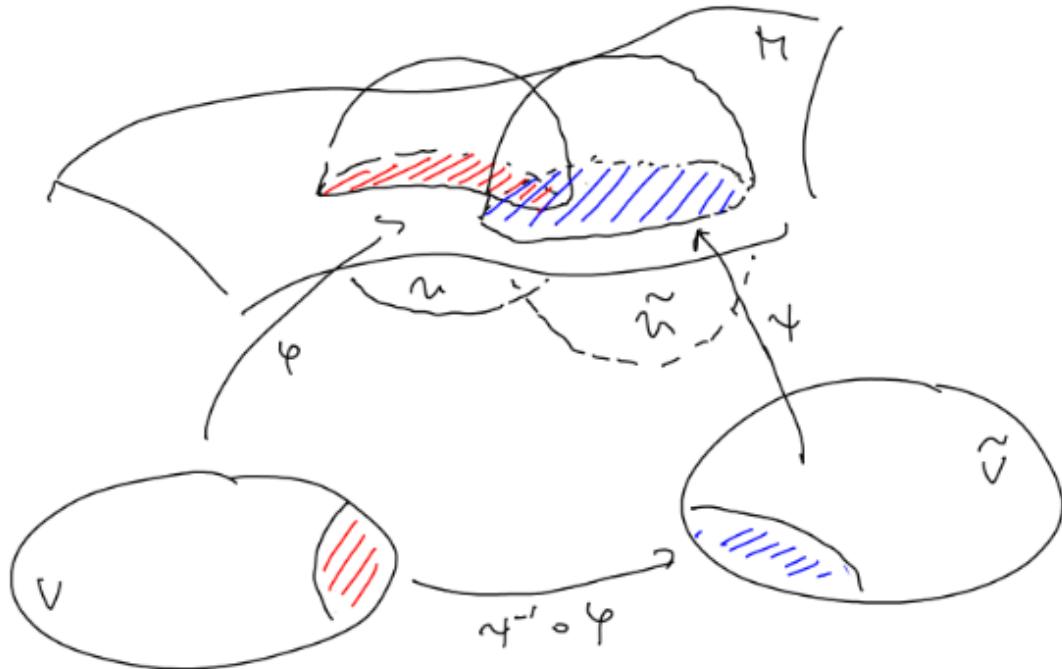
29 min

**Satz 4.9 (Parameterwechsel).** Sei  $M \subset \mathbb{R}^n$  eine  $d$ -dimensionale Untermannigfaltigkeit. Seien  $\varphi: V \rightarrow M \cap U$  und  $\psi: \tilde{V} \rightarrow M \cap \tilde{U}$  zwei lokale Parametrisierungen mit  $M \cap U \cap \tilde{U} \neq \emptyset$ . Dann ist der *Parameterwechsel*

$$\psi^{-1} \circ \varphi: \varphi^{-1}(M \cap U \cap \tilde{U}) \rightarrow \psi^{-1}(M \cap U \cap \tilde{U})$$

ein Diffeomorphismus.

**Bemerkungen.**  $\varphi^{-1}(M \cap U \cap \tilde{U})$  und  $\psi^{-1}(M \cap U \cap \tilde{U})$  sind offen.



*Beweis.* Da  $\varphi$  und  $\psi$  stetig sind, und  $M \cap U$  und  $M \cap \tilde{U}$  offen, sind  $\varphi^{-1}(M \cap U \cap \tilde{U})$  und  $\psi^{-1}(M \cap U \cap \tilde{U})$  offen.

$\psi^{-1} \circ \varphi$  ist als Komposition von Homöomorphismen Homöomorphismus.

Sei  $a \in M \cap U \cap \tilde{U}$ ,  $t_0 := \varphi^{-1}(a)$ ,  $\tilde{t}_0 := \psi^{-1}(a)$ .  $M$  ist lokal diffeomorph zu einem Ebenenstück  $\implies \exists$  offene Umgebung  $U(a)$  von  $a$  und eine offene Teilmenge  $W \subset \mathbb{R}^n$  s. d. und ein Diffeomorphismus  $F: U(a) \rightarrow W$  s. d.

$$F(M \cap U(a)) = \{ y \in W \mid y_{d+1} = \dots = y_n = 0 \}.$$

Ohne Einschränkung nehmen wir an, dass

$$U(a) \subseteq U \cap \tilde{U}$$

gilt. Dann gilt auf  $\varphi^{-1}(M \cap U \cap \tilde{U})$

$$F \circ \varphi = (h_1, \dots, h_d, 0, \dots, 0)$$

und auf  $\psi^{-1}(M \cap U \cap \tilde{U})$

$$F \circ \psi = (f_1, \dots, f_d, 0, \dots, 0)$$

und  $D(F \circ \varphi)$  und  $D(F \circ \psi)$  haben Rang  $d$ , da  $DF$  Rang  $n$  hat und  $D\varphi$  und  $D\psi$  Rang  $d$ .

$$\implies (h, 0): \varphi^{-1}(M \cap U \cap \tilde{U}) \rightarrow \{ y \in W \mid y_j = 0 \quad j \geq d+1 \}$$

und

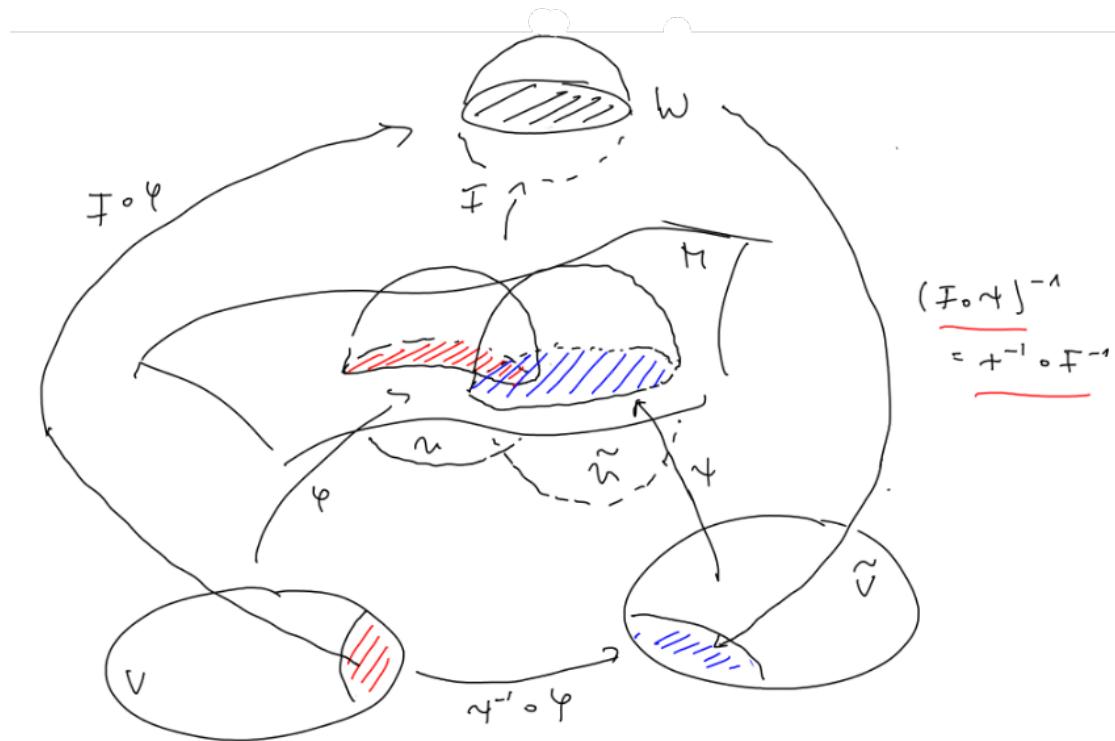
$$(f, 0): \psi^{-1}(M \cap U \cap \tilde{U}) \rightarrow \{ y \in W \mid y_j = 0 \quad j \geq d+1 \}$$

sind Diffeomorphismen. Auf  $\varphi^{-1}(M \cap U \cap \tilde{U})$  gilt

$$\psi^{-1} \circ \varphi = (F \circ \psi)^{-1} \circ (F \circ \varphi) = (f, 0)^{-1} \circ (h, 0),$$

□

also ist  $\psi^{-1} \circ \varphi$  Diffeomorphismus.



**Vorlesung 14**

Mo 08.06. 10:15

**4.I. Tangential- und Normalraum**

**Erinnerung.**  $\text{Ta } M = \text{Kern } Df(a)$ , wo  $f$  eine lokale Beschreibung von  $M$  bei  $a$  als Nullstellengebilde ist.

$$\text{Ta } M = \text{Bild}(D\varphi(x_0)) = \text{Span} \left\{ \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \\ \partial_1 g(x_0) \end{pmatrix}, \dots, \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \\ \partial_d g(x_0) \end{pmatrix} \right\},$$

wo  $\varphi(x) = (x, g(x))$  die zu einer lokalen Beschreibung von  $M$  bei  $a$  als Graph gehörige Graphenabbildung ist und  $a = (x_0, g(x_0))$ .

Es gilt allgemeiner

**Satz 4.10.** Sei  $M$   $d$ -dimensionale Untermannigfaltigkeit des  $\mathbb{R}^n$ . Sei  $a \in M$ ,  $U \subset \mathbb{R}^n$  offene Umgebung von  $a$  und  $\varphi: V \rightarrow M \cap U$  eine lokale Parametrisierung von  $M$  bei  $a$ , also  $\varphi: V \rightarrow \mathbb{R}^n$ ,  $V \subset \mathbb{R}^d$  offen, ist  $C^1$ ,  $D\varphi(t)$  von maximalem Rang  $d$  für alle  $t \in V$ ,  $\varphi: V \rightarrow M \cap U$  Homöomorphismus. Sei  $t_0 \in V$  s. d.  $\varphi(t_0) = a$ . Dann ist  $\text{Ta } M = \text{Bild } D\varphi(t_0)$ , also

$$\text{Ta } M = \text{Span} \underbrace{\{\partial_1 \varphi(t_0), \dots, \partial_d \varphi(t_0)\}}_{\text{Basis von } \text{Ta } M}. \quad (4.1)$$

*Beweis.*

$$\text{rang } D\varphi(t_0) = d \implies \dim \text{Bild } D\varphi(t_0) = d \quad \checkmark.$$

$\text{Ta } M \subset \text{Bild } D\varphi(t_0)$  Sei  $v = v_1 \partial_1 \varphi(t_0) + \dots + v_d \partial_d \varphi(t_0)$ ,  $v_j \in \mathbb{R}$ . Betrachte

$$\begin{aligned} \gamma: (-\varepsilon, \varepsilon) &\rightarrow M \cap U \\ \gamma(s) &:= \varphi(t_0 + s\underline{v}) \end{aligned}$$

mit  $\underline{v} = (v_1, \dots, v_d)^T \in \mathbb{R}^n$  und mit  $\varepsilon$  so klein gewählt, dass  $t_0 + s\underline{v} \in V$  für alle  $s \in (-\varepsilon, \varepsilon)$ . Dann gilt:  $\gamma(0) = \varphi(t_0) = a$  und  $\gamma$  ist stetig differenzierbar und

$$\begin{aligned} \gamma'(0) &= D\varphi(t_0) \cdot \underline{v} \\ &= \partial_1 \varphi(t_0) \cdot v_1 + \dots + \partial_d \varphi(t_0) \cdot v_d \\ &= v. \end{aligned}$$

$\text{Ta } M \supset \text{Bild } D\varphi(t_0)$ . Sei also  $v \in \text{Ta } M$ ,  $\gamma: (-\varepsilon, \varepsilon) \rightarrow M \cap U$  stetig differenzierbar mit  $\gamma(0) = a$ ,  $\gamma'(0) = v$ . Dann ist  $c = \varphi^{-1} \circ \gamma: (-\varepsilon, \varepsilon) \rightarrow V$  stetig differenzierbar (das folgt wie im Beweis von 4.9 azs der Verknüpfung mit einem Diffeomorphismus  $F: \tilde{U} \rightarrow W$  mit

$$F(M \cap \tilde{U}) = \{a \in W \mid y_j = 0 \quad j \geq d+1\}.$$

O.B.d.A.  $U = \tilde{U}$ , somit ist

$$\varphi^{-1} \circ \gamma = (\underbrace{F \circ \varphi}_{=(f,0) \text{ Diffeo}})^{-1} \circ (F \circ \gamma)$$

$C^1$ .) Es gilt (Kettenregel!)

$$\begin{aligned} v &= \gamma'(0) \\ &= (\varphi \circ c)'(0) \\ &= D\varphi(\underbrace{c(0)}_{=\varphi^{-1}(a)=t_0} \cdot c'(0)) \\ &= \partial_1 \varphi(t_0) \cdot \underbrace{c'_1(0)}_{\in \mathbb{R}} + \cdots + \partial_d \varphi(t_0) \cdot \underbrace{c'_d(0)}_{\in \mathbb{R}}, \end{aligned}$$

also  $v \in \text{Span} \{ \partial_1 \varphi(t_0), \dots, \partial_d \varphi(t_0) \}$ .

□



$$T_{\gamma(t)}M = \{ \lambda \gamma'(t) \mid \lambda \in \mathbb{R} \}.$$

**Beispiel.**  $\mathbb{S}^2 \subset \mathbb{R}^3$ ,  $a = (0, 0, 1)^T$ .

a)

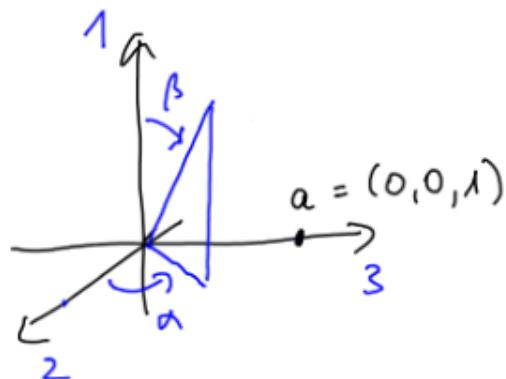
$$\begin{aligned} f(x) &= 1 - x_1^2 - x_2^2 - x_3^2 \\ Df(x) &= -2(x_1, x_2, x_3) \\ \text{Ta } M &= \text{Kern } D(a) = \text{Kern}(-2e_3^T) \\ &= \{ e_3 \}^\perp = \text{Span} \{ e_1, e_2 \}. \end{aligned}$$

b)  $\varphi: D = \{x \in \mathbb{R}^2 \mid \|x\|_E < 1\} \rightarrow \mathbb{R}^3$ ,  $\varphi(x) = (x, \sqrt{1 - x_1^2 - x_2^2})^T$ ,

$$\begin{aligned}\partial_1 \varphi(x) &= \begin{pmatrix} 1 \\ 0 \\ \frac{-x_1}{\sqrt{\dots}} \end{pmatrix} \\ \partial_2 \varphi(x) &= \begin{pmatrix} 0 \\ 1 \\ \frac{-x_2}{\sqrt{\dots}} \end{pmatrix} \\ a &= \varphi(0) \\ \implies \text{Ta } M &= \text{Span} \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \right\}.\end{aligned}$$

c)  $\varphi: (0, \pi) \times (0, 2\pi) \rightarrow \mathbb{R}^3$ ,  $\varphi(\beta, \alpha) = (\cos \beta, \cos \alpha \sin \beta, \sin \alpha \sin \beta)^T$ .

15 min



$$\partial_1 \varphi(\beta, \alpha) = (-\sin \beta, \cos \alpha \cos \beta, \sin \alpha \cos \beta)^T$$

$$\partial_2 \varphi(\beta, \alpha) = (0, -\sin \alpha \sin \beta, \cos \alpha \sin \beta)^T$$

$$\varphi\left(\frac{\pi}{2}, \frac{\pi}{2}\right) = (0, 0, 1)^T = a$$

$$\implies \text{Ta } M = \text{Span} \left\{ \begin{pmatrix} -1 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 \\ -1 \\ 0 \end{pmatrix} \right\} = \text{Span} \{ e_1, e_2 \}.$$

**Definition.** Sei  $M \subset \mathbb{R}^n$   $d$ -dimensionale Untermannigfaltigkeit.  $v \in \mathbb{R}^n$  heißt *Normalenvektor* ab  $M$  im Punkt  $a \in M$ , falls  $\langle v, v \rangle = 0 \forall v \in \text{Ta } M$ , also  $v \in (\text{Ta } M)^\perp =: \text{Na } M$  „*Normalen-Raum*“.

**Bemerkungen.** i) Da  $\text{Ta } M$  ein Vektorraum ist, so ist auch  $\text{Na } M$  ein Vektorraum.

ii)  $\dim \text{Na}(M) = n - d$ , denn  $\text{Ta } M \oplus \text{Ta } M^\perp = \mathbb{R}^n$ .

iii)  $\text{Ta } M = \text{Kern } Df(a)$ , wenn  $f$   $M$  lokal bei  $a$  als Nullstellengebilde beschreibt (also  $\exists U \subset \mathbb{R}^n$  offen,  $a \in U$ , s. d.  $M \cap U = \{x \in U \mid f(x) = 0\}$ ,  $f: U \rightarrow \mathbb{R}^{n-d}$  ist  $C^1$  und  $Df(x)$  hat maximalen Rang  $n - d$  für alle  $x \in U$ .) Anders gesagt:

$$\text{Ta } M = \{v \in \mathbb{R}^n \mid \langle v, \text{grad } f_j(a) \rangle = 0 \quad j = 1, \dots, n - d\}.$$

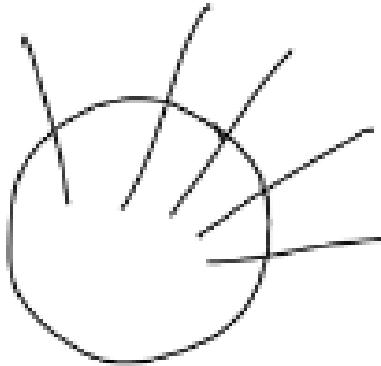
Wegen der linearen Unabhängigkeit der Gradienten folgt hieraus:

$$\text{Na } M = \text{Span}\{\text{grad } f_1(a), \dots, \text{grad } f_{n-d}(a)\}.$$

iv)  $\text{Ta } M = \text{Bild } D\varphi(t_0)$  wenn  $\varphi$  lokal bei  $a$   $M$  parametisiert (vgl. 4.10) und  $\varphi(t_0) = a$ . Die Spalten von  $D\varphi(t_0)$  bilden eine Basis von  $\text{Ta } M$

$$\implies \text{Na } M = \{v \in \mathbb{R}^n \mid \langle v, \partial_j \varphi(t_0) \rangle = 0 \quad j = 1, \dots, d\}.$$

**Beispiel.**  $M = \mathbb{S}^2 \subset \mathbb{R}^3$ .



a)  $\text{Na } M = \{\lambda a \mid \lambda \in \mathbb{R}\}.$

b)  $a \in \mathbb{S}^2$ ,  $a_3 > 0$ .

$$\begin{aligned}\text{Na } M &= \left\{ v \mid v_1 = \frac{a_1}{\sqrt{1 - a_1^2 - a_2^2}} v_3, v_2 = \frac{a_2}{\sqrt{1 - a_1^2 - a_2^2}} v_3 \right\} \\ &= \left\{ \lambda \frac{1}{\sqrt{1 - a_1^2 - a_2^2}} \begin{pmatrix} a_1 \\ a_2 \\ \sqrt{\dots} \end{pmatrix} \mid \lambda \in \mathbb{R} \right\} = \{ \lambda a \mid \lambda \in \mathbb{R} \}.\end{aligned}$$

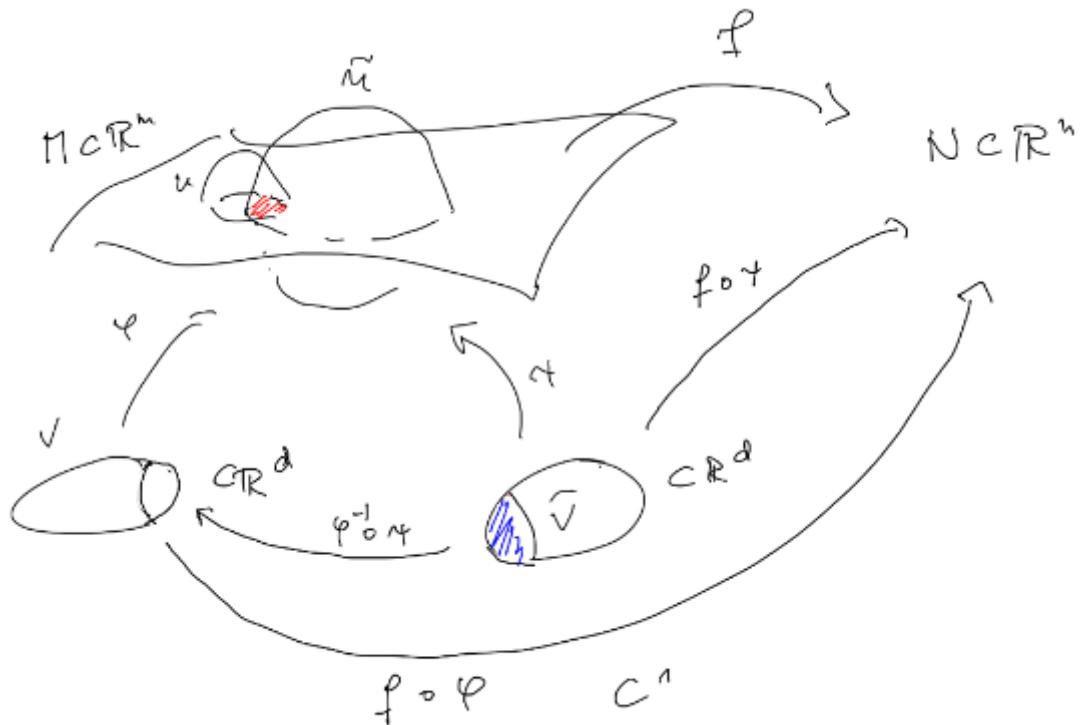
c) Für allgemeinere Punkte als den Nordpol wäre dies mit Rechenaufwand verbunden.

25 min

**Definition 4.11.** Sei  $M$   $d$ -dimensionale Untermannigfaltigkeit des  $\mathbb{R}^m$ , sei  $N$   $d'$ -dimensionale Untermannigfaltigkeit des  $\mathbb{R}^{n'}$ . Eine Abbildung  $f: M \rightarrow N$  heißt *stetig differenzierbar*, falls für alle  $a \in M$  gilt: Wenn  $\varphi: M \cap U$  eine lokale Parametrisierung bei  $a$  ist, so ist  $f \circ \varphi: \mathbb{R}^d \supset V \rightarrow \mathbb{R}^{n'}$  stetig differenzierbar.

**Bemerkungen.** Ist  $X \circ \varphi$  für eine lokale Parametrisierung  $C^1$  und  $\psi: \tilde{V} \rightarrow M \cap U$  weitere Parametrisierung bei  $a$ , so ist auch  $X \circ \psi|_{\psi^{-1}(M \cap U \cap \tilde{U})} C^1$ , denn

$$X \circ \psi|_{\psi^{-1}(M \cap U \cap \tilde{U})} = X \circ \varphi \circ \underbrace{\varphi^{-1} \circ \psi}_{\text{Diffeo (4.10)}} \Big|_{\psi^{-1}(M \cap U \cap \tilde{U})}.$$



$$f \circ \psi|_{\textcolor{blue}{U}} = f \circ \varphi \circ \underbrace{\varphi^{-1} \circ \psi}_{\text{Diffeo}}|_{\textcolor{blue}{U}}, \quad \textcolor{blue}{U} = \psi^{-1}(M \cap U \cap \tilde{U}).$$

**Definition 4.12.** Sei  $M$   $d$ -dimensionale Untermannigfaltigkeit des  $\mathbb{R}^n$ . Ein *Vektorfeld* auf  $M$  ist eine stetig differenzierbare Abbildung

$$X: M \rightarrow \mathbb{R}^n \text{ s.d. } X(a) \in \text{Ta } M \quad \forall a \in M.$$

. Ein *Normalenfeld* ist eine stetig differenzierbare Abbildung

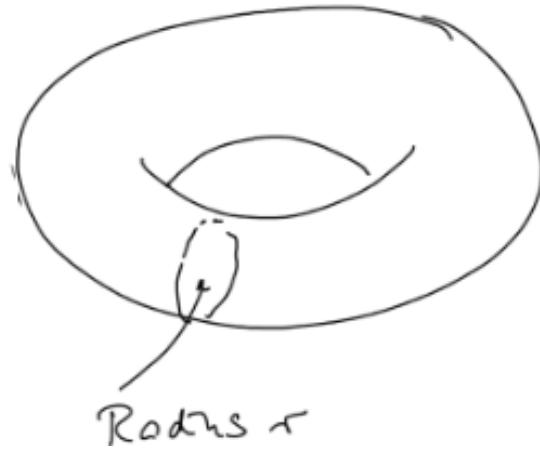
$$\mathcal{N}: M \rightarrow \mathbb{R}^n \text{ s.d. } \mathcal{N}(a) \in \text{Na } M \quad \forall a \in M.$$

**Beispiele.** i)  $M = \mathbb{S}^5$ ,  $\mathcal{N}(a) = a$  offensichtlich stetig differenzierbar.

ii)  $M = \text{Torus } \Pi^2 \subset \mathbb{R}^3$ ,  $r < R$ .

$$f(x, y, z) = (x^2 + y^2 + z^2 + R^2 - r^2)^2 - 4R^2(x^2 + y^2).$$

$f = 0$  definiert eine Untermannigfaltigkeit:



$$\text{grad } f(a) = \begin{pmatrix} 4a_1(a_1^2 + a_2^2 + a_3^2 - R^2 - r^2) \\ 4a_2(a_1^2 + a_2^2 + a_3^2 - R^2 - r^2) \\ 4a_3(a_1^2 + a_2^2 + a_3^2 + R^2 - r^2) \end{pmatrix} \neq 0$$

für  $a \in \Pi^2 \implies Df(a)$  hat maximalen Rang (1).

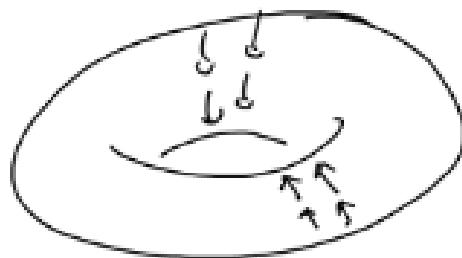
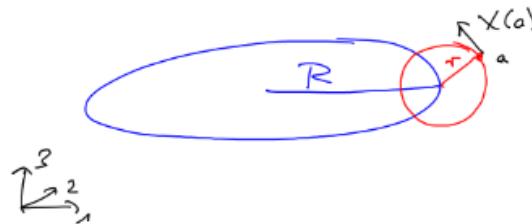
$$\mathcal{N}(a) = \frac{\text{grad } f(a)}{\|\text{grad } f(a)\|_{\mathbb{E}}}$$

ist stetig differenzierbar. *Tangentialraum*  $T_a M$  = Ebene senkrecht  $v_a = \text{grad } f(a)$ . Parametrisierung bei  $a \in \Pi^2$ : Mit geeigneten Intervallen für die Winkel gilt:

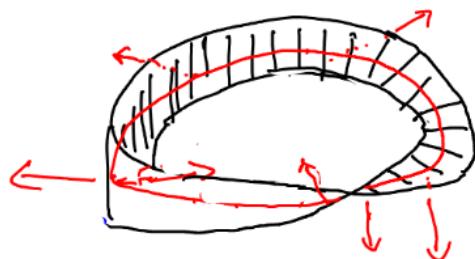
$$a = \varphi(\alpha, \beta) = R \begin{pmatrix} \cos \alpha \\ \sin \alpha \\ 0 \end{pmatrix} + r \begin{pmatrix} \cos \alpha \cos \beta \\ \sin \alpha \cos \beta \\ \sin \beta \end{pmatrix}$$

$$X(\varphi(\alpha, \beta)) = \partial_2 \varphi(\alpha, \beta) = \begin{pmatrix} -r \cos \alpha \sin \beta \\ -r \sin \alpha \sin \beta \\ \cos \beta \end{pmatrix}$$

offensichtlich  $C^1$ .



iii) Möbius-Band.



Langer Streifen Papier, einmal verdrehen, zusammenkleben. Mittellinie einzeichnen. Normalenvektor auf einem Punkt der Mittellinie bestimmen, auf 1 normieren ( $\|v\|_E = 1$ ). Führt man diesen stetig entlang der Mittellinie und verändert dabei seine Länge nicht, sieht man, dass es kein stetiges normiertes Normalenfeld auf dem Möbius-Band geben kann. Eine solche Fläche nennt man nicht orientierbar.

**Definition 4.13.** Sei  $X: M \rightarrow \mathbb{R}^n$  ein Vektorfeld ( $M$   $d$ -dimensionale Untermannigfaltigkeit).

tigkeit). Eine *Integralkurve* von  $X$  ist eine stetig differenzierbare Kurve  $\gamma: I \rightarrow M$  s. d.

$$\gamma'(t) = X(\gamma(t))$$

Im nächsten Kapitel (Differenzialgleichungen) wird es um Existenz und Eindeutigkeit solcher Kurven gehen (aber nur im  $\mathbb{R}^n$ ).

**Definition 4.14.**  $M \subset \mathbb{R}^n$  ist eine  $d$ -dimensionale *Untermannigfaltigkeit mit Rand* falls es zu jedem  $a \in M$  eine lokale Parametrisierung gibt  $\varphi: V \rightarrow \varphi(V) \subset M$  mit  $V \subset \mathbb{R}_-^d$ ,  $\mathbb{R}_-^d = \{x \in \mathbb{R}^d \mid x_1 \leq 0\}$  ( $V$  offen in  $\mathbb{R}_-^d$ , falls  $\exists U$  offen in  $\mathbb{R}^d$  s. d.  $V = \mathbb{R}_-^d \cap U$  (Teilraumtopologie), also z. B. ist  $(\infty, 0) \times (0, 1)^{d-1}$  offen in  $\mathbb{R}_-^d$ ).  $p$  heißt *Randpunkt von*  $M$ , falls  $p = \varphi(t)$  mit  $t \in V \cap \partial\mathbb{R}_-^d$ , wobei

$$\partial\mathbb{R}_-^d = \{ (0, x_2, \dots, x_d) \in \mathbb{R}_-^d \}.$$

Die Menge aller Randpunkte wird mit  $\partial M$  bezeichnet.

**Bemerkung.** Man kann auch  $\mathbb{R}_+^d = \{x \in \mathbb{R}^d \mid x_1 \geq 0\}$  oder auch  $\{x \in \mathbb{R}^d \mid x_d \geq c\}$  bzw.  $\{x \in \mathbb{R}^d \mid x_d \leq c\}$  betrachten.

**Beispiel.**  $\overline{\mathbb{D}} = \{x \in \mathbb{R}^2 \mid \|x\|_{\mathbb{E}} \leq 1\}$ .  $\overline{\partial\mathbb{D}} = \mathbb{S}^1$ . Wähle zu  $a \in \overline{\mathbb{D}}$ ,  $a \neq (1, 0)^T$ . Sei  $V = [0, 1] \times (0, 2\pi)$ . Dann ist  $V$  offen in  $\{x \in \mathbb{R}^2 \mid x_1 \leq 1\}$  und  $V \cap \{x \mid x_1 = 1\} = \mathbb{S}^1 \setminus \{(1, 0)\}$ .

$$\varphi: \underbrace{[0, 1] \times (0, 2\pi)}_{=V} \rightarrow \mathbb{R}^2 \quad \varphi(r, \alpha) = r(\cos \alpha, \sin \alpha)^T.$$

Analog zu  $a = (1, 0)^T$ .

$$\varphi: [0, 1] \times (-\pi, \pi) \rightarrow \mathbb{R}^2 \quad \varphi(r, a) = r(\cos a, \sin a)^T.$$

Ende Audio  
Teil 1

## 4.II. Flächenbemessung auf Untermannigfaltigkeiten

**Definition.**

$$\begin{aligned} I_a &:= \langle \cdot, \cdot \rangle_a: \mathrm{T}_a M \times \mathrm{T}_a M \rightarrow \mathbb{R} \\ I_a(v, w) &= \langle v, w \rangle_a = \langle v, w \rangle_{\mathbb{E}} \quad \forall v, w \in \mathrm{T}_a M \end{aligned}$$

heißt *erste Fundamentalnorm*.

**Bemerkung.** Ta  $M = \text{Span}\{\partial_1\varphi(t_0), \dots, \partial_d\varphi(t_0)\}$ ,  $\varphi: V \rightarrow M \cap U$  Parametrisierung bei  $a$ ,  $\varphi(t_0) = a$ . Es gilt für  $v = \sum v_i \partial_i\varphi(t_0)$ ,  $w = \sum w_i \partial_i\varphi(t_0)$

$$\begin{aligned}\langle v, w \rangle_a &= \left\langle \sum v_i \partial_i\varphi(t_0), w_j \partial_j\varphi(t_0) \right\rangle_a \\ &= \sum_i \sum_j v_i w_j \underbrace{\langle \partial_i\varphi(t_0), \partial_j\varphi(t_0) \rangle_E}_{=: g_{ij}(t_0)}.\end{aligned}$$

$g := (g_{ij})_{\substack{i=1, \dots, d \\ j=1, \dots, d}}: V \rightarrow \mathbb{R}$  heißt *metrischer Tensor* (oder Gramsche Matrix).

**Lemma 4.15.** Unter parameterwechsel (vgl. 4.9) verhält sich der metrische Tensor wie folgt:

$$= D\Phi(s)^T g_\varphi(\Phi(s)) D\Phi(s)$$

↑  
bezgl.  $\psi: \tilde{V} \rightarrow M \cap \tilde{U}$  berechnet  
↑  
bezgl.  $\varphi: V \rightarrow M \cap U$  berechnet

mit

$$\Phi = \varphi^{-1} \circ \psi: \varphi^{-1}(M \cap U \cap \tilde{U}) \rightarrow \varphi^{-1}(M \cap U \cap \tilde{U}).$$

*Beweis.* Betrachte  $\psi = \varphi \circ \Phi$ . Kettenregel liefert

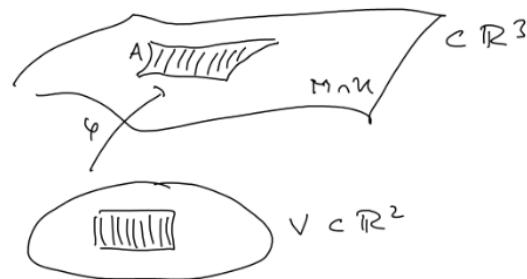
$$\langle \partial_k \psi, \partial_l \psi \rangle = \sum_{i,j} \langle \partial_i \varphi, \partial_j \varphi \rangle \partial_k \Phi_i \partial_l \Phi_j. \quad \square$$

**Frage 4.16.** Was hat  $g$  mit Flächenmessung zu tun?

- a) Wir betrachten eine Hyperfläche ( $M = (n - 1)$ -dimensional, also hier  $d = 2$ )  $M \subset \mathbb{R}^3$ . Sei  $M \cap U \subset \mathbb{R}^3$  parametrisiert von  $\varphi: V \rightarrow M \cap U$ . Dann ist

$$\varphi(t_0 + h) = \varphi(t_0) + D\varphi(t_0) \cdot h + R_{t_0}(h).$$

Wir wollen dies nutzen, um die Fläche von  $A$  abzuschätzen.



- b) Zunächst: Sei  $\mathcal{R} = [a_0, a_0 + l_1] \times [b_0, b_0 + l_2] \subset V$ . Dann ist, wenn man die Restterme vernachlässigt  $\varphi(\mathcal{R})$  ungefähr ( $t_0 = (a_0, b_0)^T$ )

$$\varphi(t_0) + D\varphi(t_0) \cdot \begin{pmatrix} a_0 \\ b_0 \end{pmatrix} + P(l_1 \partial_1 \varphi(t_0), l_2 \partial_2 \varphi(t_0)),$$

wobei  $P(v, w)$  das Parallelotop ist, dass von  $v, w \in \mathbb{R}^3$  aufgespannt wird,

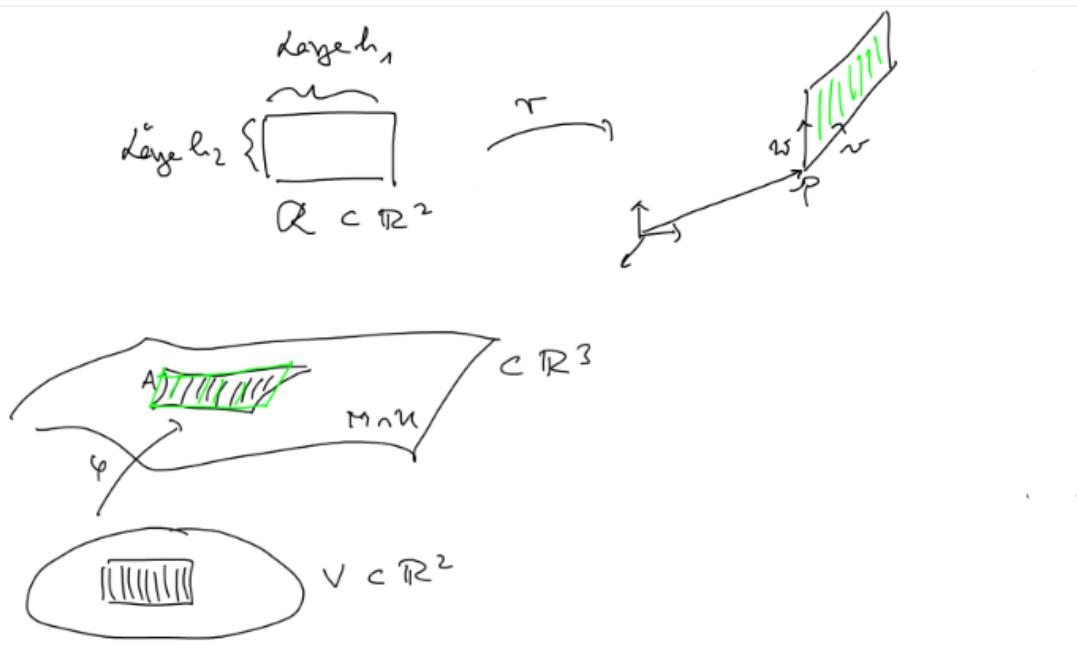
$$P(v, w) = \{ s_1 v + s_2 w \mid 0 \leq s_j \leq 1 \}.$$

Denn für  $h \in \mathbb{R}^2$  gilt

$$D\varphi(t_0) \cdot h = h_1 \partial_1 \varphi(t_0) + h_2 \partial_2 \varphi(t_0) \in \mathbb{R}^3$$

und allgemein gilt für  $v, w \in \mathbb{R}^3$  und

$$\begin{aligned} r_{v,w} = r: \mathcal{R} &\rightarrow \mathbb{R}^3 & r \begin{pmatrix} h_1 \\ h_2 \end{pmatrix} &= h_1 v + h_2 w \\ r(\mathcal{R}) &= \underbrace{r(t_0)}_{=p} + P(l_1 v, l_2 w). \end{aligned}$$



Zur ungefähren Berechnung des Flächeninhaltes von  $\varphi(\mathcal{R})$  berechnen wir also den des Parallelotops  $P(\partial_1 \varphi(t_0) l_1, \partial_2 \varphi(t_0) l_2)$ .

- c) Seien  $v, w \in \mathbb{R}^3$ . Dann ist  $\|v \times w\|$ ,  $v \times w = \begin{pmatrix} v_2w_3 - v_3w_2 \\ v_3w_1 - v_1w_3 \\ v_1w_2 - v_2w_1 \end{pmatrix}$ , „Kreuzprodukt“ der Flächeninhalt des Parallelotops  $P(v, w)$ . Den:

$$\begin{aligned} A &= \|v\| \cdot \|w\| \cdot \sin \beta \\ &= \sqrt{\|v\|^2 \|w\|^2 (1 - \cos(\beta)^2)} \\ &= \sqrt{\|v\|^2 \|w\|^2 - \langle v, w \rangle^2} \\ &= \sqrt{\|v \times w\|^2} = \|v \times w\| \end{aligned}$$

mit  $\times$  wie oben.

- d) Es gilt

$$\|\partial_1 \varphi(t_0) \times \partial_2 \varphi(t_0)\|_{\mathbb{E}} = \sqrt{\det g_\varphi(t_0)} \quad (*)$$

mit  $\det g_\varphi$  = Determinante der Gramschen Matrix. Also ist die GLäche von  $\varphi(R)$  etwa

$$\sqrt{g_\varphi(t_0)} l_1 l_2 = \sqrt{g_\varphi(t_0)} \text{Fläche}(\mathcal{R}).$$

Die Determinante gibt also an, wie sehr die Fläche von  $\mathcal{R}$  gestaucht oder gestreckt wird.

Wir werden uns später überlegen, wie man ausgehend von den hier skizzierten Ideen zu einer Definition des Flächeninhalts kommt, der nicht von der gewählten Parametrisierung abhängt.

Es ist noch (\*) zu zeigen.

**Lemma 4.17.** Sei  $\varphi: V \rightarrow M \cap U$  Parametrisierung bei  $a$ ,  $V \subset \mathbb{R}^d$ ,  $M \subset \mathbb{R}^n$ . Die Determinante von  $g_{ij}$  berechnet sich wie folgt:

$$\det g(t_0) = \sum_{i_1 < \dots < i_d} \left( \det D \begin{pmatrix} \varphi_{i_1} \\ \vdots \\ \varphi_{i_d} \end{pmatrix}(t_0) \right)^2$$

*Beweis.*

$$g(t_0) = (D\varphi)(t_0)^T \cdot D\varphi(t_0)$$

□

und lineare Algebra (Satz von Cauchy-Binet)

**Bemerkung.** Für  $d = 2, n = 3$  gilt nun

$$\begin{aligned}
 \det g_\varphi &= \sum_{i < j} \left( \det D \begin{pmatrix} \phi_i \\ \phi_j \end{pmatrix} \right)^2 \\
 &= \det D \begin{pmatrix} \varphi_1 \\ \varphi_2 \end{pmatrix}^2 + \det D \begin{pmatrix} \varphi_1 \\ \varphi_3 \end{pmatrix}^2 + \det D \begin{pmatrix} \varphi_2 \\ \varphi_3 \end{pmatrix}^2 \\
 &= \underbrace{\langle \partial_1 \varphi \times \partial_2 \varphi, \partial_1 \varphi \times \partial_2 \varphi \rangle}_{\text{Rechnung}} \implies \text{Beh. } (*).
 \end{aligned}$$

# Kapitel 5.

## Differentialgleichungen

### Vorlesung 15

Do 11.06. 10:15

**Definition 5.1.**  $V \subset \mathbb{R}^{n+1}$ ,  $F: V \rightarrow \mathbb{R}^n$  stetig. Dann nennt man

$$x' = F(t, x) \quad (*)$$

eine *Differentialgleichung (DGL)*

1. *Ordnung* ( $n = 1$ ) oder *System von DGL'n 1. Ordnung* ( $n \geq 2$ ). Eine *Lösung* von  $(*)$  ist eine differenzierbare Funktion  $u: I \rightarrow \mathbb{R}^n$ ,  $I \subset \mathbb{R}$  Intervall, mit  $u'(t) = F(t, u(t))$ , also insbesondere  $(t, u(t)) \in V$ , also

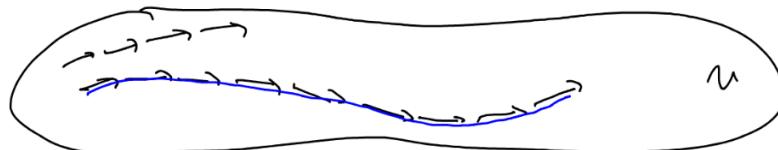
$$\Gamma_u = \{ (t, u(t)) \mid t \in I \} \subset V.$$

### 5.I. Geometrische Interpretation

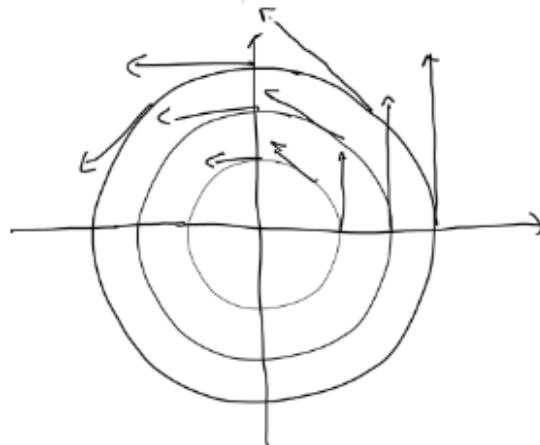
Sei zunächst  $F$  nicht von  $t$  abhängig,  $F: \mathbb{R} * U \rightarrow \mathbb{R}^n$ ,  $U \subset \mathbb{R}^n$  offen,  $T(t, x) = G(x)$ , also

$$x' = G(x). \quad (**)$$

Dann kann man  $G$  als Vektorfeld auf  $U$  ansehen (denn  $\text{Ta } U = \text{Span} \{ e_1, \dots, d_n \}$ ), in diesem Fall ein *stetiges Vektorfeld*, also eine stetige Abbildung  $G: U \rightarrow \text{Ta } U = \mathbb{R}^n$ . Eine Lösung der DGL  $(**)$  ist dann eine differenzierbare Integralkurve (nicht unbedingt stetig differenzierbar).



**Beispiel.**  $G: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ ,  $G(x_1, x_2) = \begin{pmatrix} -x_2 \\ x_1 \end{pmatrix}$ .



Hier raten wir:

$$u(t) = \begin{pmatrix} \cos t \\ \sin t \end{pmatrix} \cdot r \quad r > 0.$$

Tatsächlich

$$\begin{aligned} u'(t) &= \begin{pmatrix} -\sin t \\ \cos t \end{pmatrix} \cdot r \quad r > 0 \\ &= G(u(r)). \end{aligned}$$

Man kann sich vorstellen, während Zeit  $t$  vergeht, wandert ein Punkt entlang von  $u$  und erfüllt zu jedem Zeitpunkt  $t$ , dass seine Geschwindigkeit in Richtung und Größe durch  $G$  vorgegeben ist (und wie wir sehen werden unter gewissen Umständen  $u$  selbst dadurch eindeutig festgelegt wird).

Dieses Bild trägt auch noch, wenn  $F$  explizit von  $t$  abhängt:

In diesem Fall verändert sich das Vektorfeld selbst, während Zeit verstreicht, und der Tangentialvektor, dem  $u'(t)$  zum Zeitpunkt  $t$  gleich sein soll, ist nicht nur von der Position  $\gamma(t)$  abhängig, sondern auch om Zeitpunkt  $t$ .

**Beispiele.** i) Aus dem Satz über implizite Funktionen berechnete Ableitung der Auflösungsfunktion  $g$ , etwa in Vorlesung 9 (1. Beispiel vor Satz 3.19)

$$y' = \begin{pmatrix} \frac{t}{y} & \frac{t}{y} \end{pmatrix}.$$

ii) Wachstumsmodelle:

a) Wachstum proportional zur Populationszahl

$$x' = \alpha x, \quad u(t) = ce^{\alpha t}.$$

b) Logistisches Wachstum (Grenzen des Wachstums wegen Ressourcen-Knappheit).

Wachstum ist proportional zur Populationszahl  $y$  und zum verbleibenden Platz  $\mu(1 - y)$ .

$$x' = \alpha x(1 - x).$$

c) Räuber-Beute-Modelle

iii) Jede Menge Modelle in Physik, Chemie, Biologie, Wirtschaft etc., z. B. freier Fall (2. Ableitung)

$$x'' = -g.$$

Die DGL'n, die wir hier betrachte, werden explizit lösbar sein, d. h. es gibt Lösungen  $u$ , die man explizit angeben kann. Für Räuber-Beute-Modell etwa gibt es solche aber i. A. nicht.

## 5.II. Existenz- und Eindeutigkeitssatz

Sei  $x' = F(t, x)$  eine DGL wie in Definition 5.1,

$$F: V \rightarrow \mathbb{R}^n, V \subset \mathbb{R} \times \mathbb{R}^n.$$

Legt man zusätzlich fest, dass eine Lösung

$$u: I \rightarrow \mathbb{R}^n$$

erfüllen soll, dass  $u(t_0) = u_0$  gelten soll für ein  $(t_0, u_0) \in V$ , so spricht man von einem *Anfangswertproblem* (AWP) und notiert

$$\begin{aligned} x' &= F(t, x) \\ x(t_0) &= u_0 \end{aligned} \tag{*}$$

**Beispiel.**  $x' = \alpha x$ ,  $u(t) = ce^{\alpha t}$  löst die DGL.  $u(t) = u_0 e^{\alpha t}$  löst das AWP  $x' = \alpha x$ ,  $x(0) = u_0$ .

Existenz und Eindeutigkeit der Lösung:

Formuliere das Problem als Fixpunkt-Problem und verwende den Banachschen Fixpunktsatz.

Denn  $u$  ist eine Lösung der DGL  $x' = F(t, x)$

$$\iff u(t) = u(t_0) + \int_{t_0}^t u'(s) ds = u(t_0) + \int_{t_0}^t F(s, u(s)) ds$$

und löst  $u$  das AWP (\*), so ist  $u(t_0) = u_0$ .  $u$  ist also Lösung des AWP  $\iff u$  = Fixpunkt der Abbildung

$$u \mapsto Tu, \quad Tu(t) = u_0 + \int_{t_0}^t F(s, u(s)) ds.$$

**Definition 5.2.** Sei  $V \subset \mathbb{R} \times \mathbb{R}^n$  und  $F: V \rightarrow \mathbb{R}^n$  eine Funktion. Man sagt

- $F$  erfüllt auf  $V$  eine *Lipschitz-Bedingung* mit Lipschitz-Konstante  $L$ , falls es  $L \geq 0$  gibt s. d.
- $$\|F(t, x) - F(t, \tilde{x})\| \leq L \|x - \tilde{x}\| \quad \forall (t, x), (t, \tilde{x}) \in V.$$
- $F$  erfüllt *lokal eine Lipschitz-Bedingung*, falls es zu jedem Punkt  $(t_0, x_0) \in V$  eine offene Umgebung  $\tilde{V}$  und eine Konstante  $L = L_{\tilde{V}} \geq 0$  gibt s. d.  $F$  auf  $V \cap \tilde{V}$  eine Lipschitz-Bedingung mit Konstante  $L$  erfüllt.

**Lemma 5.3.** Ist  $V \subset \mathbb{R} \times \mathbb{R}^n$  offen und ist  $F: V \rightarrow \mathbb{R}^n$  bezüglich der hinteren  $n$  Variablen stetig differenzierbar, so erfüllt  $f$  lokal in  $V$  eine Lipschitz-Bedingung.

*Beweis.* Sei  $(t_0, x_0) \in V$ . Dann  $\exists r > 0$  s. d.

$$K := \{ (t, x) \in V \mid |t - t_0| \leq r \text{ und } \|x - x_0\| \leq r \}.$$

$K$  ist kompakt nach dem Satz von Heine + Borel.

$$\implies L := \sup_{(t,x) \in K} \|D_2 F\| < \infty \quad DF = \begin{pmatrix} \partial_1 f & D_2 F \end{pmatrix}. \quad (5.1) \quad \square$$

MWS  $\implies \|F(t, x) - F(t, \tilde{x})\| \leq L \|x - \tilde{x}\|$ .

**Bemerkung.** Stetige Differenzierbarkeit bezüglich der hinteren  $n$  Variablen bedeutet, dass  $D_2 F$  auch bezüglich der ersten Variablen stetig ist.

**Satz 5.4 (Picard-Lindelöf).** Sei  $V \subset \mathbb{R} \times \mathbb{R}^n$ . Sei  $mF: V \rightarrow \mathbb{R}^n$  stetig und erfülle lokal auf  $V$  eine Lipschitz-Bedingung. Dann hat das Anfangswertproblem

$$\begin{aligned} x' &= F(t, x) \\ x(t_0) &= u_0 \quad (t_0, u_0) \in V. \end{aligned}$$

ine eindeutige Lösung  $u: I \rightarrow \mathbb{R}^n$  mit  $t_0 \in I$  und  $u(t_0) = u_0$ .

*Beweis.* Da  $F$  lokal einer Lipschitz-Bedingung genügt, gibt es  $L \geq 0$  und eine Umgebung  $\tilde{V}$  von  $(t_0, x_0)$  s.d.

$$\|F(t, x) - F(t, \tilde{x})\| < L\|x - \tilde{x}\| \quad \forall (t, x), (t, \tilde{x}) \in \tilde{V} \cap V.$$

Sei  $J \times K \subset \tilde{V} \cap V$  mit  $J$  kompaktes Intervall  $t_0 \in J$ ,  $K \subset \mathbb{R}^n$  kompakt,  $x_0 \in K$ , o.B.d.A.  $K = \overline{B_r(x_0)}$ ,  $r > 0$ .

Wähle  $\delta > 0$  s.d.

$$L\delta < 1 \text{ und } \delta\|F\|_{\infty, J \times K} \leq r.$$

Betrachte  $T: \mathcal{F} \rightarrow C(I, \mathbb{R}^n)$ ,

$$\mathcal{F} = \{u \in C(I, \mathbb{R}^n) \mid u(I) \subset K\}$$

mit  $I = J \cap [t_0 - \delta, t_0 + \delta]$  (kompakt).

$$(Tu)(t) = u_0 + \int_{t_0}^t F(s, u(s)) ds.$$

Es ist tatsächlich  $Tu \in C(I, \mathbb{R}^n)$ . Sei  $u \in I$  und  $t \in I$ . Dann gilt

$$\begin{aligned} \|Tu - u_0\|_\infty &= \left\| \int_{t_0}^t F(s, u(s)) ds \right\|_\infty \\ &\leq \delta\|F\|_\infty \leq r, \end{aligned}$$

somit ist  $Tu \in \mathcal{F} \quad \forall u \in \mathcal{F}$ . Zudem ist

$$\begin{aligned} \|Tu - Tv\|_\infty &= \left\| \int_{t_0}^t (F(s, u(s)) - F(s, v(s))) ds \right\|_\infty \\ &\leq |t - t_0|L\|u - v\|_\infty \\ &\leq \underbrace{L \cdot |I| \cdot \|u - v\|_\infty}_{\leq L \cdot \delta < 1}. \end{aligned} \quad \square$$

$\Rightarrow T: \mathcal{F} \rightarrow \mathcal{F}$  ist Kontraktion. Zudem ist  $\mathcal{F} \subset C(I, \mathbb{R}^n)$  abgeschlossen (bezüglich  $\|\cdot\|_\infty$ )  $\Rightarrow$  Beh. (Banachscher Fixpunktsatz).

**Bemerkung.** Der Beweis ist konstruktiv, d. h. beginnend mit  $u_0(t) = u_0$  (z. B.) erhält man die Lösung als gleichmäßigen Limes der Iteration

$$u_{n+1}(t) = u_0 + \int_{t_0}^t F(s, u_n(s)) ds$$

**Beispiele.** i)  $x' = \begin{pmatrix} -x_2 \\ x_1 \end{pmatrix}$ ,  $x(0) = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ .

$$\begin{aligned} u_0(t) &= \begin{pmatrix} 1 \\ 0 \end{pmatrix} \\ u_1(t) &= \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \begin{pmatrix} -\int_0^t 0 ds \\ \int_0^t 1 ds \end{pmatrix} = \begin{pmatrix} 1 \\ t \end{pmatrix} \\ u_2(t) &= \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \begin{pmatrix} -\int_0^t s ds \\ \int_0^t 1 ds \end{pmatrix} = \begin{pmatrix} 1 - \frac{t^2}{2} \\ t \end{pmatrix}. \end{aligned}$$

**Behauptung.**

$$\begin{aligned} u_{2u}(t) &= \begin{pmatrix} 1 - \frac{t^2}{2} + \cdots + (-1)^n \frac{t^{2n}}{(2n)!} \\ t - \cdots - (-1)^n \frac{t^{2n-1}}{(2n-1)!} \end{pmatrix} \\ u_{2n+1}(t) &= \begin{pmatrix} 1 - \frac{t^2}{2} + \cdots + (-1)^n \frac{t^{2n}}{(2n)!} \\ t - \cdots + (-1)^n \frac{t^{2n+1}}{(2n+1)!} \end{pmatrix}. \end{aligned}$$

*Beweis.* per Induktion □

$$\implies x(t) = \begin{pmatrix} \cos t \\ \sin t \end{pmatrix}$$

ii)  $F(t, x) = 2tx$ ,  $F: \mathbb{R}^2 \rightarrow \mathbb{R}$ ,  $x' = 2tx$ ,  $x(0) = x_0$ .

$$\begin{aligned} u_{k+1}(t) &= x_0 + 2 \int_0^t s u_k(s) ds \\ u_0(t) = x_0 \implies u_1(t) &= x_0 + x_0 t^2 \\ u_2(t) &= x_0 + 2 \int_0^t s(x_0 + x_0 s^2) ds \\ &= x_0 \left( 1 + t^2 + \frac{t^4}{2} \right). \end{aligned}$$

Per Induktion beweist man

$$\begin{aligned} u_k(t) &= x_0 \left( 1 + t^2 + \frac{t^4}{2} + \frac{t^6}{3!} + \cdots + \frac{t^{2k}}{k!} \right) \\ \implies u(t) &= x_0 e^{t^2}. \end{aligned}$$

**Definition.** Sei  $V \subset \mathbb{R} \times \mathbb{R}^{k \cdot n}$ ,  $F: V \rightarrow \mathbb{R}^n$  stetig. Dann nennt man

$$x^{(k)} = F(t, x, x', x'', \dots, x^{(k-1)}) \quad (*)$$

eine *Differentialgleichung (DGL)*  $k$ -ter *Ordnung* ( $n = 1$ ) oder *System von DGL*'n  $k$ -ter *Ordnung* ( $n \geq 2$ ).

**Beispiel.**  $x'' = -(x')^2 + xt$  ( $k = 2$ ) (also  $F(t, x_0, x_1) = -x_1^2 + x_0 t$ ).

Eine Lösung von  $(*)$  ist eine  $n$ -mal differenzierbare Funktion  $\varphi: I \rightarrow \mathbb{R}^n$  mit

$$\varphi^{(k)}(t) = F(t, \varphi(t), \varphi'(t), \varphi''(t), \dots, \varphi^{(k-1)}(t)).$$

Damit dies wohldefiniert ist, muss gelten

$$(t, \varphi(t), \varphi'(t), \varphi''(t), \dots, \varphi^{(k-1)}(t)) \in V \quad \forall t \in I.$$

**Bemerkung.** Wir können in Beweisen o.B.d.A.  $k = 1$  setzen und alle Sätze, die wir für diesen Fall beweisen, auf den Fall  $k \geq 2$  übertragen. Es gilt nämlich:

**Satz 5.5 (Reduktion der Ordnung).** Sei

$$x^{(k)} = F(t, x, x', \dots, x^{(k-1)}) \quad (*)$$

DGL (-System)  $k$ -ter Ordnung. Betrachte

$$F: V \rightarrow \mathbb{R}^{n \cdot k}$$

$$\tilde{F}(t, x_0, x_1, \dots, x_{k-1}) = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_{k-1} \\ F(t, x_0, x_1, \dots, x_{k-1}) \end{pmatrix}$$

und das DGL-System 1. Ordnung

$$\begin{pmatrix} x_0 \\ x_1 \\ \vdots \\ x_{k-1} \end{pmatrix}' = \tilde{F}(t, x_0, \dots, x_{k-1}) = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_{k-1} \\ F(t, x_0, x_1, \dots, x_{k-1}) \end{pmatrix} \quad (**)$$

Sei  $u: I \rightarrow \mathbb{R}^n$  eine Lösung von (\*), gelte also

$$u^{(k)}(t) = F(t, u(t), u'(t), u''(t), \dots, u^{(k-1)}(t)) \quad \forall t \in I,$$

so ist

$$\underline{U} := \begin{pmatrix} u \\ u' \\ \vdots \\ u^{(k-1)} \end{pmatrix}: I \rightarrow \mathbb{R}^{k \cdot n}$$

eine Lösung von (\*\*). Umgekehrt: Sei

$$\underline{U} := \begin{pmatrix} u_0 \\ u_1 \\ \vdots \\ u_{k-1} \end{pmatrix}: I \rightarrow \mathbb{R}^{k \cdot n}$$

Lösung von (\*\*), so ist  $u := u_0: I \rightarrow \mathbb{R}^n$  eine Lösung von (\*).

**Beispiel.**

$$x'' + x = 0 \tag{*}$$

Lösungen: cos, sin.  $F(t, x, x') = -x$ . Reduktion der Ordnung:

$$\begin{aligned} \begin{pmatrix} x_0 \\ x_1 \end{pmatrix}' &= \begin{pmatrix} x_1 \\ F(t, x_0, x_1) \end{pmatrix} = \begin{pmatrix} x_1 \\ -x_0 \end{pmatrix} \\ \begin{pmatrix} \sin \\ \cos \end{pmatrix}'(t) &= \begin{pmatrix} \cos t \\ -\sin t \end{pmatrix}, \quad \begin{pmatrix} \cos \\ -\sin \end{pmatrix}'(t) = \begin{pmatrix} -\sin t \\ -\cos t \end{pmatrix} \end{aligned} \tag{**}$$

$\Rightarrow$  sin und cos sind Lösungen von (\*\*).

**Bemerkung.** Wie man an der Formulierung des Satzes und des Beispiels sieht, ist es zum Lösen einer DGL  $k$ -ter Ordnung in der Regel nicht opportun, die Ordnung zu reduzieren.

*Beweis von Satz 5.5.* Sei  $u: I \rightarrow \mathbb{R}$  Lösung on (\*), d.h.

$$u^{(k)}(t) = F(t, u(t), u'(t), u''(t), \dots, u^{(k-1)}(t)).$$

Dann ist

$$\underline{U}: I \rightarrow \mathbb{R}^{k \cdot n}, \quad \underline{U}(t) := \begin{pmatrix} u(t) \\ u'(t) \\ \vdots \\ u^{(k-1)}(t) \end{pmatrix}$$

eine Lösung von (\*\*), denn

$$\underline{U}'(t) = \begin{pmatrix} u'(t) \\ u''(t) \\ \vdots \\ u^{(k)}(t) \end{pmatrix} = \begin{pmatrix} u'(t) \\ \vdots \\ u^{(k-1)}(t) \\ F(t, u(t), u'(t), u''(t), \dots, u^{(k-1)}(t)) \end{pmatrix}.$$

Sei umgekehrt  $\underline{U}: I \rightarrow \mathbb{R}^k$ ,  $\underline{U}(t) = \begin{pmatrix} u_0(t) \\ \vdots \\ u_{k-1}(t) \end{pmatrix}$  eine Lösung von (\*\*), gelte also

$$\underline{U}'(t) = \begin{pmatrix} u'_0(t) \\ \vdots \\ u'_{k-1}(t) \end{pmatrix} = \begin{pmatrix} u_1(t) \\ \vdots \\ u_{k-1}(t) \\ F(t, u(t), u'(t), u''(t), \dots, u^{(k-1)}(t)) \end{pmatrix},$$

dann ist  $u := u_0$  eine Lösung von (\*), denn

$$\begin{aligned} u'_0(t) &= u_1(t) \\ u'_1(t) &= u_2(t) \text{ und } u'_1(t) = u''_0(t) \\ &\vdots u'_{k-2}(t) && = u_{k-1}(t) \text{ und } u'_{k-2}(t) = \\ u_{k-1}(t) &= F(t, u(t), \dots, u^{(k-1)}(t)) \text{ und } u'_{k-1} = u_0^{(k)}(t). && \square \end{aligned}$$

**Bemerkung (Ergänzend (wie in der „Vorlesung“ vom 11.06. und weitere Bemerkungen (→näheres)**  
Man sieht aus dem Beweis, dass

$$\underline{U}(t) := \begin{pmatrix} u(t) \\ u'(t) \\ \vdots \\ u^{(k-1)}(t) \end{pmatrix}$$

mit  $u$  Lösung von (\*) genau dann Lösung mit Anfangswert  $\underline{U}(t_0) = \begin{pmatrix} v_0 \\ v_1 \\ \vdots \\ v_{k-1} \end{pmatrix}$  ist, wenn

$$\begin{aligned} u(t_0) &= v_0 \\ u'(t_0) &= v_1 \\ &\vdots \\ u^{(k-1)}(t_0) &= v_{k-1}, \end{aligned}$$

also wenn bei der Lösung von (\*) alle Ableitungen bis zur  $(k - 1)$ -sten fest liegen.

**Bemerkung.** Erfüllt  $F$  lokal eine Lipschitz-Bedingung, so auch  $\tilde{F}$ , denn

$$\begin{aligned}\|\tilde{F}(t, x_0, \dots) - F(t, \tilde{x}, \dots)\|_{\mathbb{E}} &:= C \sum_{j=0}^{k-1} \|x_j - \tilde{x}_j\|_{\mathbb{E}} + C \|\tilde{F}(\dots) - F(\dots)\|_{\mathbb{E}} \\ &\leq C \|\tilde{F}(\dots) - F(\dots)\|_{\mathbb{E}}.\end{aligned}$$

Es folgt: Erfüllt  $F$  lokal eine Lipschitz-Bedingung, so hat das AWP

$$\begin{aligned}x^{(k)} &= F(t, x, x', \dots, x^{(k-1)}) \\ x(t_0) &= v_0, x'(t_0) = v_1, \dots, x^{(k-1)} = v_k\end{aligned}$$

mit  $(t_0, v_0, v_1, \dots, v_k) \in V \subset \mathbb{R}^{nk+1}$  eine eindeutige Lösung.

**Vorlesung 16**

Mo 15.06. 10:15

Letztes Mal: Existenz- und Eindeutigkeitssatz.

Die Forderung an  $F$ , lokal einer Lipschitz-Bedingung zu genügen, ist wichtig, wie folgendes Beispiel zeigt:

**Beispiel.**  $x' = x^{\frac{2}{3}}$ ,  $x(t_0) = 0$ . Dieses AWP hat Lösungen

$$\begin{aligned} u(t) &= 0 \quad \forall t \\ u(t) &= \frac{1}{27}(t - t_0)^3 \quad \forall t \\ u(t) &= \begin{cases} \frac{1}{27}(t - a)^3 & t \leq a \\ 0 & a \leq t \leq b \\ \frac{1}{27}(t - b)^3 & t \geq b \end{cases} \end{aligned}$$

für beliebige  $a < t_0 < b$ .

$F(t, x) = x^{2/3}$  erfüllt zwar lokal bei  $(t_0, x_0)$ ,  $x_0 \neq 0$ , eine Lipschitzbedingung ( $\partial_2 f(t, x) = \frac{2}{3}x^{-\frac{1}{3}}$  ist für  $x \neq 0$  stetig) aber nicht bei  $(t_0, 0)$  (für kein  $t_0$ ).

Der Beweis des Satzes von Picard + Lindelof war konstruktiv, d. h. man kann ihn (wenn lokal eine Lipschitz-Bedingung erfüllt ist) verwenden, um Lösungen zu finden (siehe letzte Mal).

In diese Vorlesung lernen wir Verfahren kennen, die auch in anderen Fällen Lösungen liefern.

**Lemma 5.6.** Seien  $G: \tilde{J} \rightarrow \mathbb{R}$ ,  $H: J \rightarrow \mathbb{R}$ , stetig,  $J, \tilde{J}$  Intervalle. Sei  $G(x) \neq 0 \quad \forall x \in \tilde{J}$ . Betrachte

$$\begin{aligned} \tilde{H}(t) &= \int_{t_0}^t H(s) ds \\ \tilde{G}(x) &= \int_{u_0}^x \frac{1}{G(y)} dy. \end{aligned}$$

Ist  $t_0 \in I$  mit  $I \subset J$  und  $\boxed{\tilde{H}(I) \subset \tilde{G}(\tilde{J})}$ . Dann hat die Differentialgleichung

$$x' = F(t, x), \quad F: J \times \tilde{J} \rightarrow \mathbb{R}, \quad F(t, x) = H(t) \cdot G(x)$$

(„Differentialgleichung mit *getrennten Variablen*“) mit der Anfangsbedingung

$$x(t_0) = u_0, \quad (t_0, u_0) \in I \times \tilde{J}$$

eine eindeutige Lösung  $u: I \rightarrow \mathbb{R}$ ,  $I \subset J$ .

*Beweis.* • Ist  $u: I \rightarrow \mathbb{R}$  Lösung des AWP, so gilt

$$\tilde{G}(u(t)) = \tilde{H}(t) \quad \forall t \in I.$$

Denn: Aus  $u'(s) = H(s)G(u(s))$  folgt

$$\int_{t_0}^t \frac{u'(s)}{G(u(s))} ds = \int_{t_0}^t H(s) ds.$$

||

(Substitutionsregel)  $\int_{u_0}^{u(t)} \frac{1}{G(u)} du$

Insbesondere ist also  $I$  derart, dass  $\tilde{H}(I) \subset \tilde{G}(\tilde{J})$ .

- $\tilde{G}'(x) = \frac{1}{G(x)} \neq 0 \implies \tilde{G}$  ist streng monoton  $\implies \tilde{G}$  besitzt  $C^1$  Umkehrfunktion  $\tilde{G}^{-1}: \tilde{G}(I) \rightarrow \mathbb{R} \implies$  mit dem eben bewiesenen

$$u(t) = \tilde{G}^{-1}(\tilde{H}(t)) \tag{*}$$

$\implies$  Wenn sie existiert, ist die Lösung eindeutig.

- Existenz: Wir nehmen (\*) als Definition und zeigen, dass  $u$  dann das AWP löst.  $u$  ist  $C^1$  ✓, und wegen (\*) gilt

$$u(t_0) = \tilde{G}^{-1}(\tilde{H}(t_0)) = \tilde{G}^{-1}(0) = u_0,$$

sowie  $(\tilde{G}(u(t)) = \tilde{H}(t))$

$$\tilde{G}(u(t)) \cdot u'(t) = \frac{1}{G(u(t))} \cdot u'(t) = \tilde{H}'(t) = H(t)$$

$\implies$  Beh.  $\square$

### Merkregel 5.7.

$$\begin{aligned} \frac{dx}{dt} &= H(t)G(x) \\ \frac{dx}{G(x)} &= H(t)dt. \\ \int \frac{1}{G(x)} dx &= \int H(t) dt. \end{aligned}$$

Dem kann man tatsächlich einen Sinn geben — Stichwort Differenzialformen.

**Bemerkung.** Wir haben im Beweis gesehen: Jede Lösung einer DGL mit getrennten Variablen erfüllt

$$\tilde{G}(u(t)) = \tilde{H}(t) \quad \forall t \in I.$$

Das kann man nutzen, um die Lösung zu bestimmen.

**Beispiel.**  $x' = x^2$ ,  $x(0) = c$ ,  $c \in \mathbb{R}$ . Eindeutige Lösung nach Lemma 5.6 mit  $I$  wie dort oder wegen der lokal erfüllten Lipschitz-Bedingung.

$$c = 0 \quad u(t) = 0 \quad \forall t \in \mathbb{R}.$$

$c > 0$  Sei  $u: I \rightarrow \mathbb{R}$  eine Lösung mit  $u(0) = c > 0$ . Dann ist  $u(t) > 0 \quad \forall t \in I$  (denn gäbe es  $t_*$  s.d.  $u(t_*) = 0$ , wäre wegen der Eindeutigkeit  $u(t) = 0 \quad \forall t$  (auch für  $t = t_0$ )). Betrachte also  $F: \mathbb{R} \times \mathbb{R}_{>0} \rightarrow \mathbb{R}$ ,  $F(t, x) = x^2$ . Berechne

$$\begin{aligned}\tilde{H}(t) &= \int_0^t ds = t \\ \tilde{G}(x) &= \int_c^x \frac{1}{y^2} dy = \frac{1}{c} - \frac{1}{x}.\end{aligned}$$

$\tilde{G}(\mathbb{R}_{>0}) = (-\infty, \frac{1}{c})$ . Setze also  $I = (-\infty, \frac{1}{c})$  (s.d.  $\tilde{H}(I) = I \subset \tilde{G}(\mathbb{R}_{>0})$ ). Für eine Lösung der AWP gilt nach Lemma 5.6

$$\frac{1}{c} - \frac{1}{u(t)} = t \quad \forall t \in I.$$

$$\implies u(t) = \frac{c}{1-ct} \text{ und } t < \frac{1}{c}.$$

$c < 0$  Genauso:

$$u(t) = \frac{c}{1-ct}$$

$$\text{und } t > \frac{1}{c}.$$

**Beispiel.** Logistisches Wachstum  $x' = \alpha x(1-x)$ ,  $\alpha > 0$  auf  $\mathbb{R} \times (0, 1)$  ( $G(x) = x(1-x) \neq 0 \quad \forall x \in (0, 1)$ ). Separation der Variablen:

$$\begin{aligned}\int_{x_0}^{u(t)} \frac{1}{y(1-y)} dy &\stackrel{!}{=} \int_{t_0}^t \alpha ds = \alpha(t - t_0). \\ \ln \left| \frac{y}{y-1} \right| \Big|_{x_0}^{u(t)} &\quad (\text{Partialbruchzerlegung: } \frac{1}{y(1-y)} = \frac{1}{y} - \frac{1}{y-1})\end{aligned}$$

Also  $\frac{u(t)}{u(t)-1} = \pm e^{\alpha t + C}$ ,  $C$  Konstante.

$$\implies u(t) = \frac{e^{\alpha t}}{\tilde{C} + e^{\alpha t}} \quad (\tilde{C} = \mp e^C).$$

Anfangsbedingung  $u(0) = x_0$ ,  $x_0 \in (0, 1)$ .  $x_0 \frac{1}{\tilde{C}+1}$ , also  $\tilde{C} = \frac{1-x_0}{x_0} > 0$ . Also ist  $u$  auf ganz  $\mathbb{R}$  definiert.

**Lemma 5.8.** Sei  $f: I \rightarrow \mathbb{R}$ , ( $I \subset \mathbb{R}$  Intervall) und sei

$$V = \left\{ (t, x) \in \mathbb{R}^* \times \mathbb{R} \mid \begin{array}{l} \uparrow \\ t \end{array} \in J \right\}. \quad (5.2)$$

Betrachte das DGL

$$x' = f\left(\frac{x}{t}\right) \quad (*)$$

auf  $V$ . Dann gilt  $u: I \rightarrow \mathbb{R}$ ,  $I \subset \mathbb{R}^*$  Intervall, ist genau dann Lösung von  $(*)$  zur Anfangsbedingung  $x(t_0) = x_0$ ,  $t_0 \in I$ ,  $(t_0, x_0) \in V$ , wenn  $v(t) := \frac{u(t)}{t}$  Lösung des AWP

$$\begin{aligned} y' &= \frac{1}{t}(f(y) - y) \\ y(t_0) &= \frac{x_0}{t_0} \end{aligned}$$

ist.

**Bemerkung.** Man sagt, die zweite DGL gehe durch Substitution aus der ersten hervor. Merkregel:

$$y = \frac{x}{t}, \quad y' = \frac{x'}{t} - \frac{1}{t^2}x = \frac{1}{t}\left(f\left(\frac{x}{t}\right) - \frac{x}{t}\right).$$

Die 2. DGL ist lösbar wie in Lemma 5.6.

*Beweis von 5.8.* „ $\implies$ “ Sei  $u$  Lösung der 1. Gleichung und  $u(t_0) = x_0$ . Dann ist

$$\begin{aligned} v'(t) &= \frac{u'(t)}{t} - \frac{u(t)}{t^2} \\ &= \frac{1}{t}f\left(\frac{u(t)}{t}\right) - \frac{u(t)}{t^2} \end{aligned}$$

und  $v(t_0) = \frac{x_0}{t_0}$ .

„ $\impliedby$ “ Sei  $v$  Lösung der 2. Gleichung und  $v(t_0) = \frac{x_0}{t_0}$ . Dann ist

$$\begin{aligned} u'(t) &= (tv(t))' = v(t) + tv'(t) \\ &= v(t) + (f(v(t)) - v(t)) \\ &= f(u(t)/t) \end{aligned}$$

und  $u(t_0) = x_0$ . □

**Beispiel.**  $x' = 1 + \frac{x}{t} + \left(\frac{x}{t}\right)^2$  auf  $\mathbb{R}_{>0} \times \mathbb{R}$ . Substitution

$$\rightsquigarrow y' = \frac{1}{t}(1 + y^2).$$

Trennung der Variablen:

$$\begin{aligned} \int_{y_0}^{v(t)} \frac{1}{1+z^2} dz &= \int_{t_0}^t \frac{1}{s} ds \\ \implies \arctan v(t) - \arctan y_0 &= \ln \frac{t}{t_0}, \end{aligned}$$

also  $u(t) = t \cdot \tan\left(\ln \frac{t}{t_0} + a\right)$  löst die ursprüngliche Gleichung und ist auf einem Intervall  $I$ , das  $t_0$  enthält, definiert.

**Bemerkung.** Typische DGL aus der klassischen Mechanik (Physik)

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

$f: I \rightarrow \mathbb{R}$ ,  $I \subset \mathbb{R}$  Intervall, stetig („Kraft“). Definiere  $\underline{U}(x) = - \int_a^x f(s) ds$ ,  $\underline{U}: I \rightarrow \mathbb{R}$  („potentielle Energie“). Damit schreibt sich  $(*)$  wie folgt:

$$x'' = -\underline{U}'(x).$$

Ist  $w: J \rightarrow \mathbb{R}$  Lösung, so gilt

$$\begin{aligned} \left(\frac{1}{2}w'(t)\right)^2 &= w'(t) \cdot w''(t) \\ &= w'(t) \cdot (-\underline{U}'(w(t))) \\ &= -\frac{d}{dt}(\underline{U} \circ w)(t), \end{aligned}$$

also

$$\frac{d}{dt}\left(\frac{1}{2}w'(t)^2 + \underline{U}(w(t))\right) = 0$$

(Konstante der Bewegung).  $\implies \exists$  Konstante  $E$  („Gesamtenergie“) s. d.

$$\frac{1}{2}v(t)^2 + \underline{U}(w(t)) = E \quad \forall t \quad (5.3)$$

$(v(t) = w'(t)$  „Geschwindigkeit“)  $\implies$  Lösungen müssen derart sein, dass  $\underline{U}(x(t)) \leq E$  und es gilt dann

$$w'(t) = \pm \sqrt{2(E - \underline{U}(w(t)))}$$

(je nachdem, ob  $w' \geq 0$  oder  $\leq 0$ ).

Wir betrachten also die DGL mit getrennten Variablen

$$x' = \pm \sqrt{2(E - \underline{U}(x))}.$$

Oft sind DGL'n auch *implizit* gegeben, also von der Form

$$f(t, x) + h(t, x)x' = 0.$$

**Satz 5.9 (Exakte Differentialgleichungen).** Sei  $f(t, x) + h(t, x)x' = 0$  gegeben,  $f$  und  $h: V \rightarrow \mathbb{R}$   $C^1$ ,  $V = J_1 \times J_2$  mit  $J_1, J_2$  Intervalle. Es gelte  $\partial_1 h = \partial_2 f$  auf  $V$ . Dann gibt es eine zweimal stetig differenzierbare Funktion  $F: V \rightarrow \mathbb{R}$  mit  $h = \partial_2 F$  und  $f = \partial_1 F$ .

Ist  $u: I \rightarrow \mathbb{R}$  mit  $\Gamma_u \subset V$  eine Lösung der DGL, so gilt  $F(t, u(t)) = C$  ( $C$  durch Anfangsbedingung festgelegt).

Umgekehrt: Ist  $h(t_0, x_0) \neq 0$ , so ist  $F(t, x) = C$  lokal bei  $(t_0, x_0)$  auflösbar nach  $x$  und die Auflösungsfunktion  $u: I \rightarrow \mathbb{R}$  löst die Differentialgleichung mit  $u(t_0) = x_0$ .

*Beweis.* Dass  $\partial_1 h = \partial_2 f$  notwendig ist für die Existenz von  $F$  ist klar. Dass sie auch hinreichend ist, werden wir im letzten Teil der Vorlesung beweisen.

Satz über implizite Funktionen: Betrachte  $F(t, x) - C = 0$ . Ist  $\partial_2 F(t_0, x_0) \neq 0$ , so gibt es lokal bei  $(t_0, x_0)$  eine Auflösungsfunktion  $u: I \rightarrow \mathbb{R}$ ,  $u(t_0) = x_0$  mit

$$u'(t) = -(\partial_2 F(t, u(t)))^{-1} \partial_1 F(t, u(t)) \quad \forall t \in I,$$

also

$$f(t, u(t)) + h(t, u(t)) \cdot u'(t) = 0 \quad \forall t \in I.$$

Ist umgekehrt  $u: I \rightarrow \mathbb{R}$  Lösung der DGL, so gilt (mit  $\varphi(t) := (t, u(t))$ )

$$\begin{aligned} (F \circ \varphi)'(t) &= \begin{pmatrix} \partial_1 F(\varphi(t)) & \partial_2 F(\varphi(t)) \end{pmatrix} \cdot \begin{pmatrix} 1 \\ u'(t) \end{pmatrix} \\ &= f(t, u(t)) + h(t, u(t))u'(t) = 0. \end{aligned} \quad \square$$

### Beispiel.

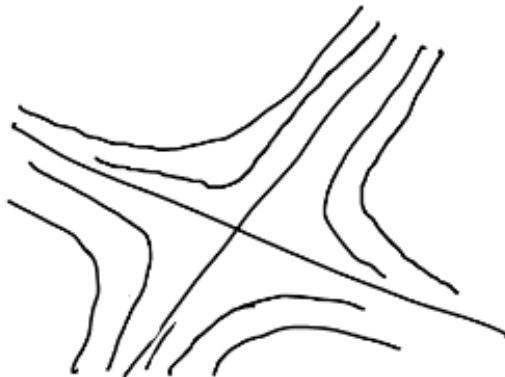
$$\begin{aligned} 7t + 7x + (6t - 2x)x' &= 0 \\ \partial_t(6t - 2x) &= 6 = \partial_x(7t + 6x) \\ \partial_2 F = h &\implies \int h(t, x) dx = F(t, x) + \tilde{C}(t) \\ &\implies F(t, x) = 6tx - x^2 + C(t) \\ \partial_1 F = f &\implies 7t + 6x \stackrel{!}{=} 6x + C'(t) \\ &\implies C(t) = \frac{7}{2}t^2 + C_0. \end{aligned}$$

Allgemeine Lösung gegeben durch

$$F(t, x) = \frac{7}{2}t^2 + 6tx - x^2 = C,$$

also  $u(t) = 3t \pm \sqrt{-C + 25\frac{t^2}{2}}$ , falls  $c > 0$ ,  $t > \frac{\sqrt{2c}}{5}$ .

Hyperbeln mit Hauptachsenrichtungen  $\begin{pmatrix} 2 \\ 1 \end{pmatrix}$  und  $\begin{pmatrix} -1 \\ 2 \end{pmatrix}$ .



**Bemerkung.** Ist eine DGL der Form

$$f(t, x) + h(t, x)x' = 0$$

nicht exakt, kann man nach einer Funktion  $M$  suchen, s. d.

$$M(t, x)f(t, x) + M(t, x)h(t, x)x' = 0$$

exakt ist. Diese DGL hat dieselbe Lösungsmenge wie die ursprüngliche DGL. Eine solche Funktion  $M$  heißt *Eulerscher Multiplikator*.

Er muss erfüllen:

$$\partial_2 M \cdot f + M \partial_2 f = \partial_1 M \cdot h + M \partial_1 h,$$

also

$$\partial_2 M \cdot f - \partial_1 M \cdot h + (\partial_2 f - \partial_1 h)M = 0.$$

**Beispiele.** • Wenn  $\frac{\partial_2 f - \partial_1 h}{h} = \rho(t)$  (unabhängig von  $x$ ), kann man  $M(t) = C e^{\int_{t_0}^t \rho(s) ds}$  wählen.

• Analog wenn  $\frac{\partial_1 h - \partial_2 f}{f} = \rho(x)$ , kann man  $M(x) = C e^{\int_{x_0}^x \rho(y) dy}$  wählen.

### 5.III. Lineare Differentialgleichungen

**Definition 5.10.** Sei  $I \subset \mathbb{R}$  ein Intervall und

$$A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix}: I \rightarrow \text{Mat}(n \times n, \mathbb{R})$$

stetig (wir haben gesehen, ist dies gleichbedeutend mit der Stetigkeit aller  $a_{ij}$ ). Dann heißt

$$x' = A(t)x$$

*homogenes (nicht im Sinne von 5.8) lineares Differentialgleichungssystem* (oder homogene lineare DGL) 1. Ordnung. Sei  $b = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}: I \rightarrow \mathbb{R}^n$  stetig (also alle  $b_i$  stetig). Dann heißt

$$x' = A(t)x + b(t)$$

*inhomogenes lineares Differentialgleichungssystem* (oder inhomogene lineare DGL) 1. Ordnung. Eine Lösung ist eine differenzierbare Funktion

$$u: J \rightarrow \mathbb{R}^n \text{ s. d. } u'(t) = A(t) \cdot \underset{\substack{\uparrow \\ \text{Matrix-Vektor-Multiplikation}}}{u(t)}$$

bzw.  $u'(t) = A(t) \cdot u(t) + b(t)$ .

**Bemerkung.** Es gibt zum AWP

$$x' = A(t)x + b(t), \quad x(t_0) = x_0$$

stets eine eindeutige Lösung

$$u: J \rightarrow \mathbb{R}^n, \quad u(t_0) = x_0.$$

Denn: Für  $F(t, x) = A(t) \cdot x + b(t)$  gilt

$$\|F(t, x) - F(t, \tilde{x})\| = \|A(t) \cdot (x - \tilde{x})\| \leq C \|A(t)\|_{\text{op}} \cdot \|x - \tilde{x}\|$$

$$\sup_{v \neq 0} \frac{\|A(t)v\|}{\|v\|}$$

und  $\sup_{t \in [-\varepsilon + t_0, t_0 + \varepsilon]} \|A(t)\|_{\text{op}} < \infty$ , da  $[-\varepsilon + t_0, t_0 + \varepsilon]$  kompakt ist und  $t \mapsto \|A(t)\|_{\text{op}}$  stetig.

Wir werden sehen, dass eine lineare DGL immer eine Lösung besitzt mit  $J = I$ .

**Bemerkung.** Es ist zweckmäßig, auch komplexe lineare DGL'n zuzulassen, also  $A: I \rightarrow \text{Mat}(n \times n, \mathbb{C})$ ,  $b: I \rightarrow \mathbb{C}^n$ . Um Stetigkeit von  $A$  un  $b$  in diesem Fall zu definieren, identifiziert man  $\mathbb{C} = \mathbb{R}^2$ , betrachtet also  $A$  als Abbildung mit Werten in  $\mathbb{R}^{(2n)^2}$ ,  $b$  als Abbildung mit Werten in  $\mathbb{R}^{2n}$ . Dito für die Differenzierbarkeit von  $u$ .

Bevor wir den allgemeinen Fall studieren, betrachten wir  $n = 1$ .

**Beispiel 5.11.**  $a: I \rightarrow \mathbb{R}$  stetig,  $x' = a(t)x$ ,  $x(t_0) = u_0$ . Das AWP hat die (eindeutige) Lösung

$$u: I \rightarrow \mathbb{R}, \quad u(t) = u_0 \exp\left(\int_{t_0}^t a(s) ds\right).$$

$u(t)$  löst die Gleichung

$$u'(t) = u_0 \cdot a(t) \cdot \exp\left(\int_{t_0}^t a(s) ds\right) \quad \checkmark.$$

**Lemma 5.12 („Variation der Konstanten“).** Betrachte  $x' = a(t)x + b(t)$ ,  $a, b: I \rightarrow \mathbb{R}$  stetig,  $x(t_0) = x_0$ . Sei  $v: I \rightarrow \mathbb{R}$  Lösung der zugehörigen homogenen Gleichung  $x' = a(t)x$ . Sei  $v(t_0) \neq 0$ . Dann ist  $v(t) \neq 0 \quad \forall t$  und wir erhalten die eindeutige Lösung  $u$  in der inhomogenen Gleichung vermöge

$$u(t) := v(t) \cdot w(t)$$

mit

$$w(t) = \int_{t_0}^t v(s)^{-1} b(s) ds + C,$$

wobei  $C$  so angepasst wird, dass  $u(t_0) = x_0$  gilt.

*Beweis.* Wir machen den Ansatz  $u = v \cdot w$  und bestimmen  $w$ . Es ist

$$\begin{aligned} u'(t) &= v'(t)w(t) + v(t)w'(t) \\ &= a(t)v(t)w(t) + v(t)w'(t) \\ &= a(t)v(t)w(t) + b(t) \\ \implies a(t)w'(t) &\stackrel{!}{=} b(t) \implies \text{Beh.} \end{aligned}$$

Denn  $v(t) \neq 0 \quad \forall t$  wegen  $v(t_0) \neq 0$  (denn wäre  $v(t_0) \neq 0$  für ein  $t_*$  so wäre  $r(t) = 0 \quad \forall t$ ).  $\square$

**Vorlesung 17**

Do 18.06. 10:15

**Satz 5.13.** Sei  $I \subset \mathbb{R}$  Intervall und  $\mathbb{K} = \mathbb{R}$  oder  $\mathbb{C}$ . Sei  $A: I \rightarrow \text{Mat}(n \times n, K)$  stetig. Berechne  $\mathcal{L}_H(A)$ , die Menge aller Lösungen von  $x' = A(t) \cdot x$ . Es gilt:  $\mathcal{L}_H$  ist  $n$ -dimensionaler Unterraum über  $\mathbb{K}$ .

Für  $u_1, \dots, u_k \in \mathcal{L}_H(A)$  sind äquivalent

- a)  $u_1, \dots, u_k$  sind linear unabhängig über  $\mathbb{K}$ .
- b)  $\exists t_0 \in I$  s. d.  $u_1(t_0), \dots, u_k(t_0) \in \mathbb{K}^n$  linear unabhängig über  $\mathbb{K}$  sind.
- c)  $u_1(t), \dots, u_k(t) \in \mathbb{K}^n$  sind linear unabhängig über  $\mathbb{K}$  für alle  $t \in I$ .

*Beweis.* •  $\mathcal{L}_H(A)$  ist Vektorraum  $\subset$  VR aller Abbildungen:  $I \rightarrow \mathbb{K}^n$ .

- i)  $0 \in \mathcal{L}_H(A)$  ✓
- ii)  $u, v \in \mathcal{L}_H(A), \lambda \in \mathbb{K}$

$$\begin{aligned} &\implies (u + \lambda v)' = u' + (\lambda v)' = Au + \lambda \cdot Av \\ &= A(u + \lambda v) \\ &\implies u + \lambda v \in \mathcal{L}_H(A). \end{aligned}$$

- 5.13.c)  $\implies$  5.13.b)  $\implies$  5.13.a).

**5.13.a)  $\implies$  5.13.c)**

Seien  $u_1, \dots, u_k$  linear unabhängig. Angenommen,  $\exists t_0$  s. d.  $u_1(t_0), \dots, u_k(t_0) \in \mathbb{K}^n$  linear abhängig sind, d. h.  $\exists \lambda_1, \dots, \lambda_k \in \mathbb{K}$  s. d.

$$\lambda_1 u_1(t_0) + \dots + \lambda_k u_k(t_0) = 0.$$

Aber

$$u := \lambda_1 u_1 + \dots + \lambda_k u_k \in \mathcal{L}_H(A)$$

und

$$u(t_0) = 0 \implies u(t) = 0 \quad \forall t$$

↙ VOR.

- $\dim \mathcal{L}_H(A) = n$ :

Sei  $(e_1, \dots, e_n)$  die kanonische Einheitsbasis von  $\mathbb{K}^n$ .  $\exists u_1, \dots, u_n \in \mathcal{L}_H(A)$  mit  $u_j(t_0) = e_j$ . Wegen 5.13.b)  $\implies$  5.13.a) sind sie linear unabhängig  $\implies \dim \mathcal{L}_H(A) \geq n$ . Es kann aber höchstens  $n$  linear unabhängige Vektoren in  $\mathcal{L}_H(A)$  geben: Denn seien  $v_1, \dots, v_k$  linear unabhängig, dann ist (wegen 5.13.a)  $\implies$  5.13.c))  $v_1(t_0), \dots, v_k(t_0)$  für ein  $t_0 \in I$  linear unabhängig in  $\mathbb{K}^n \implies k \leq n$ .  $\square$

**Definition.** Eine Basis von  $\mathcal{L}_H(A)$  nennt man auch (*Lösungs-*) *Fundamentalsystem* (für  $x' = Ax$ ).

**Bemerkung.** Aus Satz 5.13 folgt: Eine Teilmenge  $\{ u_1, \dots, u_n \} \subset \mathcal{L}_H(A)$  ist genau dann Fundamentalsystem für  $x' = Ax$ , wenn

$$\det -\Phi(t_0) \neq 0 \quad \text{für ein } t_0 \in I,$$

wobei

$$\Phi = \begin{pmatrix} u_{11} & \cdots & u_{n1} \\ \vdots & & \vdots \\ u_{1n} & \cdots & u_{nn} \end{pmatrix}$$

( $u_{ij} = j$ -te Komponentenfunktion der  $i$ -ten Funktion).

**Beispiel.**  $A: \mathbb{R} \rightarrow \text{Mat}(2 \times 2, \mathbb{R})$ .  $A(t) = \begin{pmatrix} 0 & -\alpha \\ \alpha & 0 \end{pmatrix} = \Delta$ ,  $\alpha \in \mathbb{R}$ ,  $x' = A \cdot x$ .

$$u_1(t) = \begin{pmatrix} \cos \alpha t \\ \sin \alpha t \end{pmatrix} \quad u_2(t) = \begin{pmatrix} -\sin \alpha t \\ -\cos \alpha t \end{pmatrix}$$

sind Lösungen. Sie bildet ein Fundamentalsystem, denn

$$\det \begin{pmatrix} u_1(t) & u_2(t) \end{pmatrix} = 1 \neq 0 \quad \forall t.$$

Dies bedeutet: Jede Lösung der DGL ist eine Linearkombination von  $u_1$  und  $u_2$ .

**Beispiel.** Anfangsbedingung  $x(0) = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$ .

$$\begin{aligned} u(t) &= a_1 u_1(t) + a_2 u_2(t) \\ u(0) &= \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \stackrel{!}{=} \begin{pmatrix} 1 \\ 2 \end{pmatrix} \\ \implies u(t) &= \begin{pmatrix} \cos \alpha t - 2 \sin(\alpha t) \\ \sin \alpha t + 2 \cos \alpha t \end{pmatrix}. \end{aligned}$$

**Satz 5.14.** Sei  $I$  Intervall  $\subset \mathbb{R}$ ,  $\mathbb{K} = \mathbb{C}$  oder  $\mathbb{R}$ .  $A: I \rightarrow \text{Mat}(n \times n, \mathbb{K})$  und  $b: I \rightarrow \mathbb{K}^n$  stetig. Bezeichne  $\mathcal{L}_I(A, b)$  die Lösungsmenge des inhomogenen Systems  $x' = A(t) \cdot x + b(t)$ . Ist dann  $v_0: I \rightarrow \mathbb{R}^n$  eine Lösung des inhomogenen Systems, so gilt

$$\mathcal{L}_I(A, b) = v_0 + \mathcal{L}_H(A).$$

**Bemerkung.** Man erhält also *alle* Lösungen als eine Linearkombination von Lösungen der homogenen Gleichung und Addition irgendeiner Lösung der inhomogenen („spezielle Lösung“).

*Beweis.*  $\mathcal{L}_I(A, b) \subset v_0 + \mathcal{L}_H(A)$  Sei also  $v \in \mathcal{L}_I(A, b)$ . Dann gilt

$$(v - v_0)' = v' - v_0' = (Av + b) - (Av_0 + b).$$

Also ist

$$v - v_0 \in \mathcal{L}_H(A) \implies v = v_0 + (v - v_0) \in v_0 + \mathcal{L}_H(A).$$

$\mathcal{L}_I(A, b) \supset v_0 + \mathcal{L}_H(A)$  Sei also  $v = v_0 + u$  mit  $u \in \mathcal{L}_H(A)$

$$\implies v' = v_0' + u' = Av_0 + b + Au = Av + b.$$

□

**Lemma 5.15 (Variation der Konstanten).** Sei eine inhomogene Gleichung wie in Satz 5.14 gegeben. Es sei  $\Phi = (u_1, \dots, u_n)$  eine Fundamentalsystem für  $\mathcal{L}_H(A)$ .

Dann ist  $t \mapsto v(t) = \Phi(t) \cdot w_b(t)$  mit

$$w_b(t) = \int_{t_0}^t \Phi(s)^{-1} \cdot b(s) ds + C$$

eine Lösung der inhomogenen Gleichung. Hier ist . Matrixmultiplikation und  $\Phi^{-1}(s)$  ist die inverse Matrix.

*Beweis.* Ansatz  $v = \Phi \cdot w$ .

$$\begin{aligned} v' &= \Phi' \cdot w + \Phi \cdot w' \\ &= (u'_1, \dots, u'_n) \cdot w + \Phi \cdot w' \\ &= (\underbrace{\Phi}_{Au_1, \dots, Au_n} \cdot A) \cdot w + \Phi \cdot w' \\ &\stackrel{!}{=} A \cdot \Phi \cdot w + b \end{aligned}$$

□

$$\implies \text{Beh. } (w = w_b).$$

**Beispiel.**  $x' = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}x + \begin{pmatrix} 0 \\ t \end{pmatrix}$ .

$$\begin{aligned}\Phi(t) &= \begin{pmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{pmatrix} & \Phi(t)^{-1} &= \begin{pmatrix} \cos t & \sin t \\ -\sin t & \cos t \end{pmatrix} \\ \implies \Phi(s)^{-1} \cdot b(s) &= \begin{pmatrix} s \sin s \\ s \cos s \end{pmatrix} \\ \implies w_b(t) &= \int_0^t \begin{pmatrix} s \sin s \\ s \cos s \end{pmatrix} ds + C \\ &= \begin{pmatrix} \sin t - t \cos t \\ \cos t + t \sin t \end{pmatrix} + C \\ \implies v(t) &= \Phi(t) \cdot w_b(t) = \begin{pmatrix} -t \\ 1 \end{pmatrix}\end{aligned}$$

ist eine Lösung der inhomogenen DGL (wähle  $C = 0$ ). Anfangsbedingung z.B.  $v(0) = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$ . Äquivalente Vorgehensweise:  $v(t)$  wie oben  $+ \lambda_1 u_1 + \lambda_2 u_2$  oder  $w_b(t) = \begin{pmatrix} \sin t & \cdots \\ \dots & \dots \end{pmatrix} + \begin{pmatrix} C_1 \\ C_2 \end{pmatrix}$ . Dann ist  $v(t) = \begin{pmatrix} -t \\ 1 \end{pmatrix} + \underbrace{\Phi \cdot \begin{pmatrix} C_1 \\ C_2 \end{pmatrix}}_{C_1 u_1 + C_2 u_2}$ .

$$v(0) = \begin{pmatrix} 0 \\ 1 \end{pmatrix} + C_1 \begin{pmatrix} 1 \\ 0 \end{pmatrix} + C_2 \begin{pmatrix} 0 \\ 1 \end{pmatrix} \implies C_1 = -1.$$

**Notation.** Man satz  $\begin{pmatrix} -t \\ 1 \end{pmatrix}$  ist eine *spezielle Lösung* des inhomogenen Systems und

$$\begin{pmatrix} -t \\ 1 \end{pmatrix} + C_1 \begin{pmatrix} \cos t \\ \sin t \end{pmatrix} + C_2 \begin{pmatrix} -\sin t \\ \cos t \end{pmatrix}$$

ist die *allgemeine Lösung*.

Mit Hilfe des Satzes von der Reduktion der Ordnung übertragen wir nun das eben bewiesene auf Differentialgleichungen höherer Ordnung.

Sei  $\mathbb{K} = \mathbb{R}$  oder  $\mathbb{C}$ . Eine homogene lineare DGL  $k$ -ter Ordnung ist eine Gleichung der Form

$$x^{(k)} + A_{k-1}(t) \cdot x^{(k-1)} + \cdots + A_1(t) \cdot x' + A_0(t) \cdot x = 0 \quad (*)$$

mit  $A_j: I \rightarrow \text{Mat}(n \times n, \mathbb{K})$  stetig. Eine inhomogene lineare DGL ist von der Form

$$x^{(k)} + A_{k-1}(t) \cdot x^{(k-1)} + \cdots + A_1(t) \cdot x' + A_0(t) \cdot x = b(t) \quad (**)$$

mit  $A_j$  wie oben und  $b: I \rightarrow \mathbb{K}^n$  stetig.

Eine Lösung ist eine  $k$ -fach differenzierbare Funktion  $u: I \rightarrow \mathbb{K}^n$  mit

$$u^{(k)} + A_{k-1}(t) \cdot u^{(k-1)} + \cdots + A_1(t) \cdot u' + A_0(t) \cdot u = 0 \text{ bzw. } b(t) \quad \forall t \in I.$$

**Satz 5.16.** i) Die Menge aller Lösungen eines homogenen linearen Systems wie in (\*),  $\mathcal{L}_{\text{I},k}(A_0, \dots, A_{k-1}, b)$ , ist ein  $n \cdot k$ -dimensionaler Vektorraum.

ii) Für die Menge aller Lösungen eines homogenen linearen Systems wie in (\*),

$$\mathcal{L}_{\text{I},k}(A_0, \dots, A_{k-1}, b),$$

gilt:  $\mathcal{L}_{\text{I},k} = v_0 + \mathcal{L}_{\text{H},k}$  für eine beliebige Lösung  $v_0$  des inhomogenen Systems.

iii) Ein  $k$ -Tupel  $u_1, \dots, u_k \in \mathcal{L}_{\text{H},k}$  ist genau dann linear unabhängig, wenn für ein  $t_0 \in I$  (somit für alle  $t \in I$ ) die *Wronski-Determinante*

$$W(t_0) = \det \begin{pmatrix} u_1(t_0) & u_2(t_0) & \cdots & u_k(t_0) \\ u'_1(t_0) & \vdots & & u'_k(t_0) \\ \vdots & \vdots & & \vdots \\ u_1^{(k-1)}(t_0) & \vdots & & u_k^{(k-1)}(t_0) \end{pmatrix}$$

nicht 0 ist.

*Beweis.* Das System

$$\begin{aligned} x'_0 &= x_1 \\ x'_1 &= x_2 \\ &\vdots \\ x'_{k-2} &= x_{k-1} \\ x'_{k-1} &= -A_0(t)x_0 - \cdots - A_{k-1}(t)x_{n-1} + b(t) \end{aligned}$$

ist nach Satz 5.5 äquivalent zu (\*) bzw. für  $b(t) = 0$  (\*). Jeder Lösung  $u: I \rightarrow \mathbb{K}^n$  von (\*) bzw. (\*\*) entspricht einer Lösung des obigen Systems (mit  $b = \underline{0}$  bzw.  $b \neq \underline{0}$ )

↑  
Nullfunktion

$$\begin{pmatrix} u \\ u' \\ \vdots \\ u^{(k-1)} \end{pmatrix}$$

und umgekehrt ist  $u_0$  Lösung von  $(*)$  bzw.  $(**)$ , wenn

$$\begin{pmatrix} u \\ u' \\ \vdots \\ u^{(k-1)} \end{pmatrix}$$

Lösung des obigen Systems ist (mit  $b = \underline{0}$  bzw.  $b \neq \underline{0}$ ). Somit folgt die Behauptung aus 5.13 und 5.14.  $\square$

**Bemerkung.** Auch 5.15 (Variation der Konstanten) überträgt sich direkt auf  $(**)$ .

**Notation.** Eine Basis von  $\mathcal{L}_{H,k}(A_0, \dots, A_{k-1})$  heißt wieder *Fundamentalsystem*.

**Beispiel.**  $x'' - \frac{1}{2t}x' + \frac{1}{2t^2}x = 0$  auf  $I = \mathbb{R}_{>0}$ . Wir raten 2 Lösungen  $u_1(t) = t$  und  $u_2(t) = \sqrt{t}$ . Probe:

$$\begin{aligned} 0 - \frac{1}{2t}1 + \frac{1}{2t^2}t &= 0 \quad \checkmark \\ \frac{1}{2}\left(-\frac{1}{2t\sqrt{t}}\right) - \frac{1}{2t}\frac{1}{2\sqrt{t}} + \frac{1}{2t^2}\sqrt{t} &= 0 \quad \checkmark. \end{aligned}$$

$$W(t) = \det \begin{pmatrix} u_1(t) & U_2(t) \\ u'_1(t) & u'_2(t) \end{pmatrix} = \det \begin{pmatrix} t & \sqrt{t} \\ 1 & \frac{1}{2\sqrt{t}} \end{pmatrix} = \frac{-\sqrt{t}}{2} \neq 0 \quad \forall t > 0.$$

$\implies$  Die allgemeine Lösung der DGL ist

$$u(t) = c_1 t + c_2 \sqrt{t}$$

mit  $c_j \in \mathbb{R}$  (oder  $\mathbb{C}$ ).

Was tut man, wenn man nur eine Lösung raten kann? Im Fall  $k = 2$  und  $n = 1$  führt folgendes Verfahren auf eine zweite Lösung:

**Lemma 5.17.** Sei  $I$  Intervall,  $a_0, a_1: I \rightarrow \mathbb{K}$  ( $\mathbb{K} = \mathbb{R}$  oder  $\mathbb{C}$ ) stetig. Sei  $u: I \rightarrow \mathbb{K}$  eine Lösung von

$$x'' + a_1(t)x' + a_0(t)x = 0. \tag{*}$$

Auf  $J \subset I$  gelte  $u(t) \neq 0$ . Dann ist  $v: J \rightarrow \mathbb{K}$ ,  $v = u \cdot w$  Lösung von  $(*)$ , wenn für *wgilt*

$$w''(t) + \left(2\frac{u'(t)}{u(t)} + a_1(t)\right)w'(t) = 0. \tag{**}$$

Ist  $w$  auf  $J$  nicht konstant, sind  $v$  und  $u$  linear unabhängig.

**Bemerkung.**  $w'$  löst also die DGL 1. Ordnung

$$y' + \left( 2 \frac{u'(t)}{u(t)} + a_1(t) \right) y = 0,$$

also ist

$$\begin{aligned} w'(t) &= C_0 \exp \left( - \int_{t_0}^t 2 \frac{u'(s)}{u(s)} + a_1(s) ds \right) \\ &= C \frac{1}{u(t)^2} \exp \left( - \int_{t_0}^t a_1(s) ds \right), \end{aligned}$$

woraus man ein  $w$  bestimmen kann, dass  $(**)$  löst.

*Beweis von 5.17.*

$$\begin{aligned} v' &= u'w + uw' \\ v'' &= u''w + 2u'w' + uw'' \\ &\quad = -a_1u'w - a_0uw + 2u'w' + uw'' \\ \implies v'' + a'_1 + a_0v &= 2u'w' + a_1uw' + uw'' = 0, \end{aligned}$$

wenn  $w$   $(**)$  erfüllt.

Die 2. Behauptung ist klar.  $\square$

**Beispiel.** Nochmal

$$x'' - \frac{1}{2t}x' + \frac{1}{2t^2}x = 0$$

auf  $I = \mathbb{R}_{>0}$ .  $u(t) = t$  und bestimmen  $v$  wie oben:

$$\begin{aligned} w'(t) &= \frac{1}{u(t)^2} \exp \left( - \int_{t_0}^t a_1(s) ds \right) \\ &= \tilde{C} \frac{1}{t^2} \sqrt{t} = \tilde{C} \frac{1}{t^{\frac{3}{2}}} \\ w(t) &= \tilde{C} \frac{1}{\sqrt{t}} + C' \implies v = \sqrt{t}. \end{aligned}$$

(Man lässt hier oft die Konstanten einfach weg).

**5.18 (Besondere DGL'n mit  $k = 2, n = 1$ ).** (Achtung: andere Notation als bisher!)

Zu  $n \in \mathbb{N}_0$  betrachten wir die

**Legendresche DGL** für  $-1 < x < 1$

$$(1 - x^2)y'' - 2xy' + n(n + 1)y = 0.$$

**Hermitesche DGL** für  $x \in \mathbb{R}$

$$y'' - 2xy' + 2ny = 0.$$

**Laguerresche DGL** für  $x > 0$

$$xy'' + (1-x)y' + ny = 0.$$

**Lemma 5.19.** Zu  $n \in \mathbb{N}_0$  definiert man

$$\begin{aligned} P_n(x) &= \frac{1}{2^n n!} \frac{d^n}{dx^n} (x^2 - 1)^n && \text{Legendre-Polynom der Ordnung } n \\ H_n(x) &= (-1)^n e^{x^2} \frac{d^n}{dx^n} e^{-x^2} && \text{Hermite-Polynom der Ordnung } n \\ L_n(x) &= e^x \frac{d^n}{dx^n} x^n e^{-x} && \text{Laguerre-Polynom der Ordnung } n \end{aligned}$$

Bei den angegebene Funktionen handelt es sich tatsächlich um Polynome der Ordnung  $n$ . Sie lösen die zugehörigen DGL'n.

*Beweis.* 1. Beh.  $P_n: \frac{d^n}{dx^n} (\text{Polynom der Ordnung } 2n) = \text{Polynom von Ordnung } \leq n$ .

$$\frac{d^n}{dx^n} ((x-1)^n (x+1)^n) = \sum_{l=0}^n \underbrace{\frac{d^l}{dx^l} (x-1)^n}_{\text{Ordnung}=n-l} \underbrace{\frac{d^{n-l}}{dx^{n-l}} (x+1)^n x}_{\text{Ordnung}=l}$$

$H_n$ :

$$\frac{d^n}{dx^n} e^{-x^2} = \frac{d^{n-1}}{dx^{n-1}} (-2xe^{-x^2}) = \frac{d^{n-2}}{dx^{n-2}} \left( (-2x)^2 e^{-x^2} + \frac{d}{dx} (-2x) e^{-x^2} \right)$$

$\implies$  wenn alle  $n$  Ableitungen auf  $e^{-x^2}$  wirken, erhält man  $(-2x)^n e^{-x^2}$ . Die anderen Beiträge sind kleineren Grads multipliziert mit  $e^{-x^2}$ .

$L_n$ : genauso.

2. Beh. Nur für die Hermite-Polynome: Setze  $y(x) = e^{x^2} \frac{d^n}{dx^n} e^{-x^2}$ . Es gilt  $e^{-x^2} y(x) = \frac{d^n}{dx^n} e^{-x^2}$ .

$$\implies \frac{d^2}{dx^2} (e^{-x^2} y(x)) \frac{d^{n+2}}{dx^{n+2}} e^{-x^2} = \frac{d^{n+1}}{dx^{n+1}} (-2xe^{-x^2}).$$

Induktion:

$$(xe^{-x^2})^{(n+1)} = x(e^{-x^2})^{(n+1)} + (n+1)(e^{-x^2})^{(n)}.$$

*Beweis.* IA:  $\frac{d}{dx}(xe^{-x^2}) = e^{-x^2} + x(e^{-x^2})'$ .  $n \rightarrow n+1$ :

$$\begin{aligned} (xe^{-x^2})^{(n+1)} &= \left( (xe^{-x^2})^{(n)} \right)' \\ &\stackrel{\text{IV}}{=} \left( x(e^{-x^2})^{(n)} + n(e^{-x^2})^{(n+1)} \right)' \\ &= \left( e^{-x^2} \right)^{(n)} + x \left( e^{-x^2} \right)^{(n+1)} + n(e^{-x^2})^{(n)}. \end{aligned} \quad \square$$

$$\begin{aligned} \Rightarrow (e^{-x^2}y(x))^{(2)} &= -2x \left( e^{-x^2} \right)^{(n+1)} - 2(n+1) \left( e^{-x^2} \right)^{(n)} \\ &= -2x \frac{d}{dx} \left( e^{-x^2}y(x) \right) - 2(n+1)e^{-x^2}y(x) \\ &= e^{-x^2}(4x^2y(x) - 2y'(x) - 2(n+1)y(x)). \end{aligned}$$

Andererseits:

$$\begin{aligned} (e^{-x^2}y(x))^{(2)} &= e^{-x^2}y''(x) + 2(e^{-x^2})'y'(x) + (e^{-x^2})^{(2)}y(x) \\ &= e^{-x^2}(y''(x) - 4xy'(x) + (4x^2 - 2)y(x)) \\ \Rightarrow y''(x) - 2xy'(x) + 2ny(x) &= 0 - \end{aligned}$$

$L_n, P_n$  ähnlich.  $\square$

Ein Fundamentalsystem für diese Gleichungen (zu gegebenen  $n$ ) findet man z. B. wieder mit Lemma 5.17.

**Beispiel.** Für  $n = 1$  und die Legendre-Gleichung

$$\begin{aligned} w'(x) &= \exp \left( - \int \frac{2}{s} ds + \int \frac{2s}{1-s^2} ds \right) \\ &= \frac{1}{x^2} \frac{1}{(1-x^2)} \\ \Rightarrow w(x) &= -\frac{1}{x} + \frac{1}{2} \ln \frac{1+x}{1-x}. \end{aligned}$$

**Vorlesung 18**

Mo 15.06. 10:15

Letztes Mal: Legendre, Hermite und Laguerre DGL'n. Lösungen waren explizite Funktionen (Polynome,  $\ln$ , ...). Das nicht immer so:

**Besselsche DGL:**

$v \in \mathbb{R}$ .

$$y'' + \frac{1}{x}y' + \left(1 - \frac{v^2}{x^2}\right)y = 0 \quad \text{auf } \mathbb{R}_{>0}.$$

Lösungen heißen *Zylinder-Funktionen* der Ordnung  $v$ . Wir machen folgenden Ansatz (absolute Konvergenz voraussetzend)

$$u(x) = x^\rho \sum_{k=0}^{\infty} a_k x^k.$$

Dann ist

$$\begin{aligned} u'(x) &= \rho x^{\rho-1} \sum_{k=0}^{\infty} a_k x^k + x^{\rho-1} \sum_{k=0}^{\infty} x k a_k x^{k-1} \\ &= x^{\rho-1} \sum_{k=0}^{\infty} (k + \rho) a_k x^k \\ u''(x) &= x^{\rho-2} \sum_{k=0}^{\infty} (k + \rho)(k + \rho - 1) a_k x^k. \end{aligned}$$

Einsetzen (und mal  $x^{-\rho+2}$  nehmen) liefert für  $x^m$  (unter Verwendung von  $1 - \frac{v^2}{x^2} = \frac{x^2 - v^2}{x^2}$ )

$$\begin{aligned} (\rho + m - 1)(\rho + m)a_m + (\rho + m)a_m + a_m(-v^2) + a_{m-2} &= 0 \\ a_m((\rho + m)^2 - v^2) + a_{m-2} &= 0. \end{aligned}$$

Rekursion!

$a_0(\rho^2 - v^2) = 0$ . Wähle  $\rho = v$ , und setze  $a_0 = 1$ .

$a_1 = (2v + 1) = 0$ . Setze  $a_1 = 0$ .

Es muss gelten

$$a_m(2vm + m^2) + a_{m-2} = 0.$$

Sei  $m = 2n + 1$ , dann ist  $a_m = 0$ , denn  $a_1 = 0$  und

$$a_{2n+1} \left( \underbrace{(2n+1)^2}_{\text{ungerade}} \right) \stackrel{!}{=} 0.$$

Ist  $v \in \mathbb{Z}$ : gerade  $2v(2n+1)$

Wähle also  $a_{2n+1} = 0$ .

Ist  $m = 2n$ , so ist (per Induktion beweisen!) ( $\Gamma$  ist die Gammafunktion).

$$a_{2n} = \frac{(-1)^n}{4^n n! \underbrace{(v+n) \cdots (v+1)}_{=\frac{\Gamma(v+n+1)}{\Gamma(v+1)}}}$$

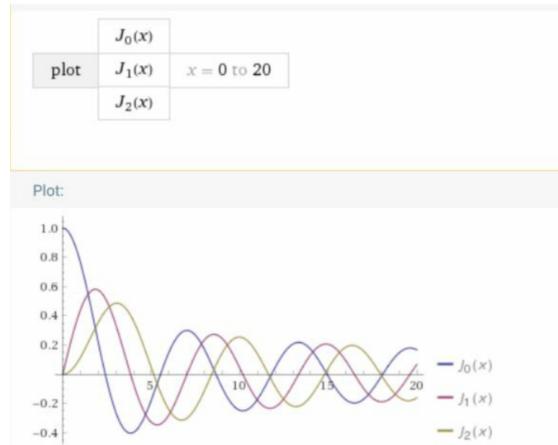
$\Rightarrow$  Lösung:

$$J_v(x) = \left(\frac{x}{2}\right)^v \sum_{n=0}^{\infty} \frac{(-1)^n}{n! \Gamma(v+n+1)} \left(\frac{x}{2}\right)^{2n}$$

„Besselfunktion 1. Art der Ordnung  $v$ “.

Reihe konvergiert absolut, unsere Manipulationen sind also a posteriori gerechtfertigt!

Ist  $v \notin \mathbb{Z}$ , so bilden  $J_v$  und  $J_{\bar{v}}$  ein Fundamentalsystem. Ist  $\gamma \in \mathbb{Z}$ , kann man z. B. die Besselfunktion 2. Art hinzuziehen, die wir hier aber nicht betrachten.



Zum Abschluss:

## 5.IV. Lineare DGL-Systeme mit konstanten Koeffizienten

**Lemma 5.20.** Sei  $A \in \text{Mat}(n \times n, \mathbb{C})$  und sei  $v \in \mathbb{C}^n$  ein Eigenvektor von  $A$  zum Eigenwert  $\lambda \in \mathbb{C}$ . Dann ist

$$u: \mathbb{R} \rightarrow \mathbb{C}^n \quad u(t) = v e^{\lambda t}$$

eine Lösung von  $x' = Ax$ .

*Beweis.*

$$u'(t) = v\lambda e^{\lambda t} = \lambda v e^{\lambda t} = A v e^{\lambda t}.$$

□

**Folgerung.** Ist die Matrix  $A$  diagonalisierbar, besitzt also  $\mathbb{C}^n$  eine Basis aus Eigenvektoren  $v_1, \dots, v_n$  von  $A$ , so bilden die Funktionen  $\{v_1 e^{\lambda_1 t}, \dots, v_n e^{\lambda_n t}\}$  ein Fundamentalsystem (wobei  $Av_j = \lambda_j v_j$   $j = 1, \dots, n$ ). Lineare Unabhängigkeit:  $u_j(0) = v_j$  sind linear unabhängig.

**Beispiel.**  $A = \begin{pmatrix} 0 & -\alpha \\ \alpha & 0 \end{pmatrix} \Rightarrow \lambda = \pm i\alpha$ . Fundamentalsystem  $\left\{ \begin{pmatrix} 1 \\ -i \end{pmatrix} e^{i\alpha t}, \begin{pmatrix} 1 \\ i \end{pmatrix} e^{-i\alpha t} \right\}$ .

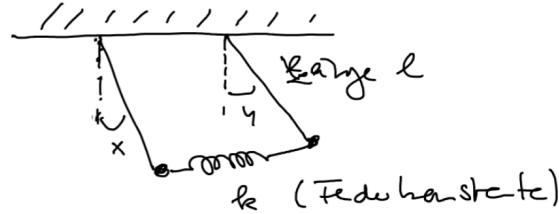
**Bemerkung 5.21.** Ist  $A \in \text{Mat}(n \times n, \mathbb{R})$ , so ist mit  $\lambda \in \mathbb{C}$  Eigenwert auch  $\bar{\lambda}$  Eigenwert und es gilt

$$Av = \lambda v \iff A\bar{v} = \bar{\lambda}\bar{v}.$$

Für  $\lambda \in \mathbb{C} \setminus \mathbb{R}$ . Man kann in diesem Fall also immer ein *reelles* Fundamentalsystem finden  $(v + \bar{v}, v - \bar{v} \in \mathbb{R}^n)$ .

**Beispiel.** Gekoppelte Pendel

$$\begin{aligned} mx'' &= -\frac{mg}{h}x - k(x - y) \\ my'' &= -\frac{mg}{y}y - k(y - x) \quad k, m, g, h > 0 \end{aligned}$$



Reduktion der Ordnung:

$$\begin{pmatrix} x \\ y \\ p \\ q \end{pmatrix}' = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -(w_0^2 + k_0) & k_0 & 0 & 0 \\ k_0 & -(w_0^2 + k_0) & 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ p \\ q \end{pmatrix}.$$

$w_0 = \sqrt{\frac{q}{l}}$ ,  $k_0 = \frac{k}{m}$ . Eigenwerte  $\pm iw_0$ ,  $\pm i\sqrt{w_0^2 + 2k_0} =: \pm iw$ . Eigenvektoren

$$v_1 = \begin{pmatrix} 1 \\ 1 \\ iw_0 \\ iw_0 \end{pmatrix} \quad \bar{v}_1 = \begin{pmatrix} 1 \\ 1 \\ -iw_0 \\ iw_0 \end{pmatrix} \quad v_2 = \begin{pmatrix} 1 \\ -1 \\ iw \\ -iw \end{pmatrix} \quad \bar{v}_2 = \begin{pmatrix} 1 \\ -1 \\ -iw \\ iw \end{pmatrix}.$$

Reelles Fundamentalsystem

$$\begin{aligned} & e^{+iw_0 t} v_1 + e^{-iw_0 t} \bar{v}_1 \\ & -i(e^{+iw_0 t} v_1 - e^{-iw_0 t} \bar{v}_1) \\ & e^{+iwt} v_1 + e^{-iwt} \bar{v}_1 \\ & -i(e^{+iwt} v_1 - e^{-iwt} \bar{v}_1). \end{aligned}$$

**Spezialfall 5.22.** Sei  $A$  diagonalisierbar. Entstehe  $A$  aus DGL  $n$ -ter Ordnung:  $x^{(n)} + a_{n-1}x^{(n-1)} + \dots + a_1x' + a_0x = 0$

$$\begin{pmatrix} x_0 \\ x_1 \\ \vdots \\ x_{n-2} \\ x_{n-1} \end{pmatrix}' = \underbrace{\begin{pmatrix} 0 & 1 & & & & \\ & 0 & \ddots & & & \\ & & \ddots & \ddots & & \\ & & & \ddots & \ddots & \\ & & & & 0 & \ddots \\ -a_0 & -a_1 & -a_2 & \dots & \dots & -a_{n-2} & -a_{n-1} \end{pmatrix}}_{=A \in \text{Mat}(n \times n, \mathbb{K})} \begin{pmatrix} x_0 \\ x_1 \\ \vdots \\ x_{n-2} \\ x_{n-1} \end{pmatrix}$$

Dann ist  $\{x_j e^{\lambda_j t} \mid j = 1, \dots, n\}$  it  $(v_1, \dots, v_n)$  Basis von  $\mathbb{K}^n$  und  $Av_j = \lambda_j v_j$  ein Fundamentalsystem. Die Null-Komponenten  $v_{j,0} e^{\lambda_j t}$  sind dann zwar Lösungen der DGL  $n$ -ter Ordnung, aber nicht unbedingt linear unabhängig: zwar ist  $v_{j,0} \neq 0$  (sonst wäre  $v_j = 0$ ), aber wenn ein Eigenwert mit höherer Multiplizität als 1 vorkommt, wären die Funktionen linear abhängig.

Es gilt jedoch:

**Folgerung 5.23.** Sei eine DGL

$$x^{(n)} + a_{n-1}x^{(n-1)} + \dots + a_1x' + a_0x = 0 \tag{5.4}$$

$a_j \in \mathbb{C}$ , gegeben. Das Polynom

$$T^n + a_{n-1}T^{n-1} + \dots + a_1T + a_0 = 0$$

habe  $n$  paarweise verschiedene Nullstellen  $\lambda_1, \dots, \lambda_n \in \mathbb{C}$ . Dann ist  $\Phi = (e^{\lambda_1 t}, \dots, e^{\lambda_n t})$  ein Fundamentalsystem für die DGL.

*Beweis.* Das charakteristische Polynom der gegebenen Matrix ist

$$\lambda_n + a_{n-1}\lambda^{n-1} + \cdots + a_1\lambda + a_0.$$

□

**Beispiel.**  $x^{(4)} - 6x^{(3)} + 11x^{(2)} - 6x = 0$ .

$$T^4 - 6T^3 + 11T^2 - 6T = T(T-1)(T-2)(T-3)$$

$\Rightarrow (1, e^t, e^{2t}, e^{3t})$  ist ein Fundamentalsystem.

Mehrfache Nullstellen → allgemeiner in Zusammenhang mit triangulierbaren Matrizen.

**Lemma 5.24.** Sei  $A \in \text{Mat}(n \times n, \mathbb{C})$ ,  $S \in \text{GL}_n(\mathbb{C})$ . Dann gilt:  $u: \mathbb{R} \rightarrow \mathbb{C}^n$  ist genau dann Lösung der DGL  $x' = Ax$ , wenn  $\tilde{u} := S^{-1}u: \mathbb{R} \rightarrow \mathbb{C}^n$  Lösung von  $y' := \underbrace{S^{-1}AS}_{B}y$ .

B

*Beweis.*

$$\begin{aligned} u'(t) &= Au(t) \\ \Leftrightarrow S^{-1} \cdot u'(t) &= S^{-1} \cdot Au(t) = (S^{-1}AS)S^{-1}u(t). \quad \square \\ &\parallel \\ (S^{-1}u)'(t) & \end{aligned}$$

**Bemerkung.** Sei  $A \in \text{Mat}(n \times n, \mathbb{C})$ . Dann gibt es  $S \in \text{GL}_n(\mathbb{C})$ , s.d.  $S^{-1}AS$  Jordan-Form besitzt, also von der Form

$$\left( \begin{array}{ccccc} \square & & & & \\ & \square & & & \\ & & \square & & \\ & & & \ddots & \\ & & & & \square \end{array} \right)$$

ist, wobei jedes Kästchen  $\square$  von der Form

$$\boxed{\begin{matrix} \lambda & 1 & 0 & \cdots & 0 \\ 0 & \lambda & 1 & 0 & \cdots & 0 \\ & & \ddots & & & \\ 0 & & & \ddots & \lambda & 1 \\ & & & & \lambda & \end{matrix}}$$

ist.

**Satz 5.25.** Sei  $A \in \text{Mat}(n \times n, \mathbb{C})$ . Sei  $B = S^{-1}AS$  die Jordan-Form von  $A$ . Seien  $J_1, \dots, J_k$  die Jordan-Kästchen von  $B$  zum Eigenwert  $\lambda_1, \dots, \lambda_k$  und seien sie der Größe  $n_1 \times n_1, \dots, n_k \times n_k$ . Dann bildet

$$\begin{aligned}
& e^{\lambda_1 t} \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad e^{\lambda_1 t} \begin{pmatrix} \frac{t}{1!} \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \quad e^{\lambda_1} \begin{pmatrix} \frac{t^2}{2!} \\ \frac{t}{1!} \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \quad \dots, \quad e^{\lambda_1 t} \begin{pmatrix} \frac{t^{n_1-1}}{(n_1-1)!} \\ \vdots \\ t \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \\
& e^{\lambda_2 t} \begin{pmatrix} 0 \\ \vdots \\ u \\ 0 \\ 1 \\ 0 \\ \vdots \\ 1 \end{pmatrix}, \quad e^{\lambda_2 t} \begin{pmatrix} 0 \\ \vdots \\ u \\ 0 \\ t \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \quad \dots \\
& \dots, \quad e^{\lambda_k} \begin{pmatrix} 0 \\ \vdots \\ 0 \\ \frac{t^{n_k-1}}{n_k-1!} \\ \vdots \\ z \\ \vdots \\ t \\ 1 \end{pmatrix}
\end{aligned}$$

ein Fundamentalsystem  $\tilde{\Phi}$  von  $y' = By$  und somit ist  $\Phi = S\tilde{\Phi}$  ein Fundamentalsystem von  $x' = Ax$ .

18:00

*Beweis.* Für ein Jordan-Kästchen  $J$  von  $B$ . Wir betrachten also  $z' = Jz$ ,

$$J = \begin{pmatrix} \lambda & 1 & & & & \\ & \lambda & 1 & & & \\ & & \ddots & \ddots & & \\ & & & \ddots & \ddots & \\ & & & & \ddots & 1 \\ & & & & & \lambda \end{pmatrix} \quad m \times m$$

$$z'_1 = \lambda z_1 + z_2$$

$$\begin{aligned}
z'_2 &= \lambda z_2 + z_3 \\
&\vdots \\
z'_{m-1} &= \lambda z_{m-1} + z_m \\
z'_m &= \lambda z_m \\
\implies v_m(t) &= e^{\lambda t} \\
\implies z'_{m-1} - \lambda z_{m-1} &= e^{\lambda t}
\end{aligned}$$

Mit 5.15 lösen.

$$\begin{aligned}
v_{n-1}(t) &= e^{\lambda t} \int \frac{1}{e^{\lambda t}} e^{\lambda t} dt = e^{\lambda t} \cdot t \\
\implies z'_{m-2} - \lambda z_{m-2} &= te^{\lambda t}.
\end{aligned}$$

Mit 5.15:

$$v_{m-2}(t) = e^{\lambda t} \int t dt = e^{\lambda t} \frac{t^2}{2}$$

$\implies$  Beh. (per Induktion):

$$v_{m-k}(t) = e^{\lambda t} \int \frac{t^{k-1}}{(k-1)!} dt = e^{\lambda t} \frac{t^k}{k!}, \quad \square$$

denn die lineare Unabhängigkeit ist offensichtlich, und  $S\tilde{\Phi}$  ist auch linear unabhängig ( $S \in \text{GL}_n(\mathbb{C})$ ).

**Bemerkung.** Insbesondere folgt Lemma 5.20 und 5.22:

Ist  $B = S^{-1}AS$  diagonal, so ist  $\{e^{\lambda j}e_j \mid j = 1, \dots, n\}$  Fundamentalsystem und  $Se_j = j$ -te Spalte von  $S$  = Eigenvektor zu  $\lambda_j$ .

Die Jordanzerlegung zu berechnen, ist aufwendig (für große  $n$ ). Liegt eine obere Dreiecksmatrix vor, kann man auch vorgehen wie folgt:

**Bemerkung 5.26.** Habe  $B \in \text{Mat}(n \times n, \mathbb{C})$  obere Dreiecksgestalt. Betrachte die DGL  $y' = By$  also

$$\begin{aligned}
y'_1 &= B_{11}y_1 + B_{12}y_2 + \cdots + B_{1,n-1}y_{n-1} + B_{1,n}y_n \\
y'_2 &= B_{22}y_2 + \cdots + B_{2,n-1}y_{n-1} + B_{2,n}y_n \\
&\vdots \\
y'_{n-1} &= B_{n-1,n-1}y_{n-1} + B_{n-1,n}y_n \\
y'_n &= B_{n,n}y_n.
\end{aligned}$$

mit Anfangsbedingung

$$y(t_0) = y_0 = \begin{pmatrix} y_{0,1} \\ \vdots \\ y_{0,n} \end{pmatrix}.$$

Dann erhält man sukzessive eine Lösung  $v: \mathbb{R} \rightarrow \mathbb{C}^n$  von  $y' = By$ :

- 1)  $y'_n = B_{nn}y_n \implies v_n(t) = C_n e^{B_{nn}t}$  ( $C_n e^{B_{nn}t_0} = y_{0,n}$ )
- 2)  $y'_{n-1} - B_{n-1,n-1}v_{n-1} = B_{n-1,n}C_n e^{B_{nn}t}$  inhomogene Gleichung. Mit Lemma 5.15 bestimme ein Fundamentalsystem und eine spezielle Lösung. Passe wieder die Konstanten an...
- k)  $v_n, v_{n-1}, \dots, v_{k+1}$  sind schon bestimmt. Für  $v_k$  ergibt sich

$$y'_k - B_{kk}v_k = \sum_{j=k+1}^n B_{kj}v_j(t)$$

usf.

23:10

Auch Gleichungen wie in 5.22, die mit Reduktion der Ordnung aus einer Gleichung  $n$ -ter Ordnung hervorgingen, lassen sich durch den Umweg über die Jordan-Form (5.25) lösen. Wir gehen direkter vor:

**Lemma 5.27.** Betrachte  $n^{(x)} + a_{n-1}x^{(n-1)} + \dots + a_0x = 0$ . Es gelte

$$p(T) := T^n + a_{n-1}T^{n-1} + \dots + a_1T + a_0 = (T - \lambda_1)^{k_1} \cdots (T - \lambda_r)^{k_r}$$

mit  $\lambda_j$  paarweise verschieden, so bildet

$$u_{j,i}(t) = t^i e^{\lambda_j t} \quad i \in \{0, \dots, k_j - 1\} \quad j \in \{1, \dots, r\}$$

ein Fundamentalsystem.

*Beweis.*  $u_{j,i}(t)$  ist Lösung:

$$(T - \lambda_1)^{k_1} \cdots (T - \lambda_r)^{k_r} = \prod_{s \neq j} (T - \lambda_s)^{k_s} \cdot (T - \lambda_j)^{k_j}.$$

Setze  $\frac{d}{dt}$  für  $T$  ein und wende den Ausdruck auf  $u_{j,i}$  an. D.h., wir prüfen, ob  $u_{j,i}$  die DGL erfüllt. Zunächst eine Hilfs-Aussage: Es gilt

$$\left( \frac{d}{dt} - \lambda \right)^k (f(t) e^{\lambda t}) = f^{(k)}(t) e^{\lambda t} \quad f \in C^k$$

per Induktion.  $k = 0 \checkmark$

$k \rightarrow k + 1$

$$\left( \frac{d}{dt} - \lambda \right) \left( \frac{d}{dt} - \lambda \right)^k (f(t)e^{\lambda t}) = \left( \frac{d}{dt} - \lambda \right) (f^{(k)}(t)e^{\lambda t}) = f^{(k+1)}(t)e^{\lambda t} + f^{(k)}(t)(\lambda e^{\lambda t} - \lambda e^{\lambda t}).$$

□

Also ist

$$\begin{aligned} p\left(\frac{d}{dt}\right) u_{j,i}(t) &= \prod_{j \neq i} \left( \frac{d}{dt} - \lambda_k \right)^{k_l} \cdot \left( \frac{d}{dt} - \lambda_j \right)^{k_j} (t^i e^{\lambda_j t}) \\ &= \prod_{j \neq i} \left( \frac{d}{dt} - \lambda_l \right)^{k_l} \underbrace{\frac{d^{k_j}(t^i)}{dt^{k_j}}} _{=0, \text{ da } i < k_j} \cdot e^{\lambda_j t}. \end{aligned}$$

Lineare Unabhängigkeit

$$W(0) = \det \begin{pmatrix} & & k_1\text{-te} & & (k_1 + k_2)\text{-te} & \\ & 1 & 0 & \dots & 0 & \\ & \lambda_1 & 1 & & \vdots & \\ & \vdots & \lambda_1 & & \vdots & \\ & \vdots & \vdots & 1 & & \\ & \vdots & & \lambda_1 & & \\ k_1\text{-te} & \vdots & & \vdots & & \vdots \\ & \vdots & & \vdots & & \vdots \\ & \vdots & & \vdots & & \vdots \\ & \lambda_1^{n-1} & \lambda_1^{n-2} & \dots & \lambda_1^{n-k_1} & \lambda_2^{n-1} & \dots & \lambda_1^{n-k_2} & \dots \end{pmatrix},$$

denn

$$\underbrace{\frac{d(t^i e^{\lambda t})}{dt^l}}_{\sum_{r=0}^l (t^i)^{(r)} \lambda^r e^{\lambda t}} \Big|_{t=0} = \begin{cases} 0 & l < i \\ 1 & l = i \\ \lambda^{l-i} & l > i. \end{cases}$$

Also

$$W(0) = \det \left( \begin{array}{ccc|c} & & & k_1 \\ & \overbrace{\begin{matrix} 1 & & & \\ \lambda_1 & \ddots & & \\ \vdots & \ddots & \ddots & \\ \lambda_1 & & & 1 \end{matrix}}^{\lambda_1} & & k_1 \\ & & & \vdots \\ & & & k_r \\ & \overbrace{\begin{matrix} 1 & & & \\ \lambda_r & \ddots & & \\ \vdots & \ddots & \ddots & \\ \lambda_r & & & 1 \end{matrix}}^{\lambda_r} & & k_r \\ & & & \vdots \end{array} \right).$$

Da die  $\lambda_j$  paarweise verschieden sind, sind die Spalten linear unabhängig  $\implies W(0) \neq 0$ .  $\square$

**Beispiel.**

$$\begin{aligned} x^{(4)} + 8x^{(2)} + 1 &= 0 \\ T^4 + 8T^2 + 16 &= (T^2 + 4)^2 \\ &= (T - 2i)^2(T + 2i)^2 \\ \lambda_1 &= 2i \quad \lambda_2 = -2i \\ u_{1,0}(t) &= e^{2it}, \quad u_{2,0}(t) = e^{-2it} \\ u_{1,1}(t) &= te^{2it}, \quad u_{2,1}(t) = te^{-2it}. \end{aligned}$$

34:05

**Bemerkung.** Inhomogene Gleichungen löst man über das zugehörige System 1. Ordnung und Variation der Konstanten. In speziellen Fällen kann man auch anders vorgehen:

**Lemma 5.28.** Sei  $b(t) = f(t)e^{\alpha t}$  mit  $\alpha = \text{Nullstelle } k\text{-ter Ordnung } (k \geq 0)$  des Polynoms

$$p(T) = T^n + a_{n-1}T^{n-1} + \cdots + a_1T + a_0$$

und wo  $f$  ein Polynom der Ordnung  $m$  ist. Dann besitzt die DGL

$$p\left(\frac{d}{dt}\right)x = b(t)$$

eine spezielle Lösung  $v: \mathbb{R} \rightarrow \mathbb{C}$  der Gestalt

$$v(t) = h(t)e^{\alpha t},$$

wobei  $h$  ein Polynom von Grad  $m + k$  ist.

*Beweis.*  $p(t) = q(T)(T - \alpha)^k$ ,  $q(\alpha) \neq 0$ . Induktion:

$$m = 0$$

$$p\left(\frac{d}{dt}\right)x = ce^{\alpha t}$$

**Behauptung.**  $v(t) = \frac{c}{k!q(a)}t^ke^{\alpha t}$  ist Lösung.

*Beweis.*

$$\begin{aligned} p\left(\frac{d}{dt}\right)v &= \frac{c}{k!q(a)}q\left(\frac{d}{dt}\right) \cdot \underset{\substack{\uparrow \\ \text{wie im Beweis von 5.27}}}{\left(t^k\right)^{(k)}} \cdot e^{\alpha t} \\ &= \frac{c}{k!q(\alpha)}e^{\alpha t} \\ &= \frac{c}{k!q(\alpha)}k!q(\alpha)e^{\alpha t}. \end{aligned}$$

$\square$

$m - 1 \rightarrow m$  Es gilt

$$\begin{aligned} p\left(\frac{d}{dt}\right)(t^{m+k}e^{\alpha t}) &= q\left(\frac{d}{dt}\right)\left(\frac{d}{dt} - \alpha\right)^k(t^{m+k}e^{\alpha t}) \\ &= q\left(\frac{d}{dt}\right)\left(\frac{(m+k)!}{m!}t^m e^{\alpha t}\right) = \underbrace{g(t)}_{\text{Polynom von Grad } m} e^{\alpha t}. \end{aligned}$$

Für geeignetes  $c \neq 0$  ist daher  $f_1 := f - cg$  ein Polynom von Grad  $m - 1$ . Nach Induktionsvoraussetzung gibt es ein Polynom  $h_1$  von Grad  $m - 1 + k$  s.d.

$$p\left(\frac{dd}{dt}\right)(h_1(t)e^{\alpha t}) = f_1(t)e^{\alpha t}.$$

Dann ist

$$h(t) := h_1(t) + ct^{m+k}$$

Polynom von Grad  $m + k$  und es gilt

$$p\left(\frac{d}{dt}\right)(h(t)e^{\alpha t}) = f_1(t)e^{\alpha t} + \underbrace{p\left(\frac{d}{dt}\right)(ct^{m+k}e^{\alpha t})}_{cg(t)e^{\alpha t}}. \quad \square$$

**Beispiel.**  $x^{(3)} + 2x^{(2)} + x' = t + 2e^{-t}$ . Homogene Gleichung:

$$p(T) = T^3 + 2T^2 + T = T(T+1)^2.$$

$\Rightarrow \lambda_1 = -1$  (Vielfachheit 2),  $\lambda_2 = 0$ .  $\Rightarrow$  Fundamentalsystem  $(e^{-t}, te^{-t}, 1)$ . Inhomogene Gleichung: Wir zerlegen die rechte Seite und lösen

$$p\left(\frac{d}{dt}\right)x = t \tag{1}$$

und

$$p\left(\frac{d}{dt}\right)x = 2e^{-t}. \tag{2}$$

Addition der Lösungen liefert die gesuchte Lösung:

(1) Ansatz:  $v_1(t) = C_1t + C_2t^2$  ( $C_0$  unnötig, da Lösung der homogenen Gleichung) Einsetzen in (1)  $\Rightarrow c_1 = \frac{1}{2}, c_2 = -2$ .

(2)  $v_2(t) = ct^2e^{-t}$  ( $te^{-t}$  und 1 sind Lösung der homogenen). Einsetzen in (2)  $\Rightarrow C = -1$ . Eine spezielle Lösung ist also

$$v(t) = -2t + \frac{1}{2}t^2 - t^2e^{-t}.$$

# Kapitel 6.

## Lebesgue-Integration

### Vorlesung 19

Do 25.10. 10:15

Integralbegriff, s. d.  $\int f_n dx \rightarrow \int f dx$ , wenn  $f_n \rightarrow f$  punktweise  $\rightsquigarrow$  Lebesgue-Integral.  
Unterschiedliche Zugänge. Hier: Riesz (Vilmos Komornik, arXiv: 1805.07289) vgl. auch  
Otto Forster, Analysis 3, Springer Vieweg, vor 2010.

**Definition 6.1.** Eine Teilmenge  $N \subset \mathbb{R}^n$  heißt *Nullmenge*, wenn es zu jedem  $\varepsilon > 0$  eine abzählbare Überdeckung von  $N$  durch offene Quadrate gibt mit Gesamtvolumen  $< \varepsilon$  ( $\text{vol}(I_1 \times \dots \times I_n) = |I_1| \cdots |I_n|$ ). Man sagt eine Eigenschaft träge *fast überall* (f.ü.) zu, wenn sie außerhalb einer Nullmenge zutrifft.

**Beispiele.** •  $N = \bigcup_{i=1}^{\infty} \{x_i\}$  ist Nullmenge  $\subset \mathbb{R}$ .

- $\mathbb{Q}$  ist Nullmenge  $\subset \mathbb{R}$ .
- $f(x) = 1, g(x) = \begin{cases} 1 & c \neq 0 \\ 0 & x = 0 \end{cases}, f = g \text{ f.ü.}$

Im Folgenden wollen wir Funktionen, die f.ü. gleich sind, miteinander identifizieren. Wir bilde also Äquivalenzklassen

$$\bar{f} = \{g \mid f = g \text{ f.ü.}\}.$$

**Notation.** Wir schreiben dennoch  $f$  (nicht  $\bar{f}$ ) und  $f = g, \subseteq g, f_n \rightarrow f, \dots$ , wenn dies nur f.ü. gilt.

**Definition 6.2.** Eine Stufenfunktion ist eine Funktion der Form

$$\varphi = \sum_{j=1}^n c_j \chi_Q,$$

$Q_j = I_{j,1} \times \cdots \times I_{j,n}$ ,  $I_{j,i}$  beschränkte Intervalle.

$$\chi_Q(x) = \begin{cases} 1 & x \in Q \\ 0 & x \notin Q \end{cases}$$

„charakteristische Funktion“. *Integral* von  $\varphi$

$$\int \varphi dx := \sum_{j=1}^m c_j \operatorname{vol}(Q_j)$$

**Lemma 6.3.**  $\mathcal{R}_0$  = Menge aller Stufenfunktionen (bzw. Äquivalenzklassen).

- i)  $\mathcal{R}_0$  ist Vektorraum.
- ii)  $\int \varphi dx$  hängt nicht von der Wahl des Repräsentanten  $\varphi$  ab.
- iii)  $\int : \mathcal{R}_0 \rightarrow \mathbb{R}$ ,  $\varphi \mapsto \int \varphi dx$  ist linear und  $\int \varphi dx \geq 0$ , wenn  $\varphi \geq 0$ .

*Beweis.* vgl. DIFF I (für  $Q = I$ , hier genauso!), aber Achtung: Hier geht es um Äquivalenzklassen.

6.3.ii) folgt aus dem Satz von Borel (siehe unten): nicht-ausgeartete Intervalle ( $\neq \emptyset$ ,  $\neq \{x\}$ ) sind keine Nullmenge in  $\mathbb{R}$ . Also sind nicht-ausgeartete Quader (also solche, die keine ausgearteten Intervalle enthalten) keine Nullmenge in  $\mathbb{R}^n$ . Nur nicht-ausgeartete Quader liefern einen Beitrag zum Integral.

*Beweis.* O.B.d.A.  $I = [0, 1]$  (den  $\{0, 1\}$  ist Nullmenge). Angenommen,  $\exists I_j$  offen s. d.  $I \subset \bigcup I_j$  und  $\sum_{j=1}^{\infty} |I_j| < \varepsilon$  mit  $0 < \varepsilon < 1$ .  $[0, 1]$  ist kompakt  $\implies \exists i_1, \dots, i_k$  s. d.

$$[0, 1] \subset \bigcup_{j=1}^k I_{i_j} \implies 1 = \int_0^1 dx \leq \int \sum_{j=1}^k \chi_{I_{i_j}}(x) dx = \sum_{j=1}^k \int \chi_{I_{i_j}}(x) dx < \varepsilon \not\downarrow. \quad \square$$

$\sum_{j=1}^k \chi_{I_{i_j}}(x) \geq 1$  (DIFF I)      DIFF I

**Notation.**  $f_k(x) \nearrow f(x)$  oder  $f_k \nearrow f$  wenn  $f_k \rightarrow f$  punktweise (f.ü.) und  $f_k \leq f_{k+1}$  (f.ü.).  $f_k(x) \searrow f(x)$  oder  $f_k \searrow f$  wenn  $f_k \rightarrow f$  punktweise (f.ü.) und  $f_k \geq f_{k+1}$  (f.ü.).

**Lemma 6.4.** i)  $(\varphi_k)_k \subset \mathcal{R}_0$ ,  $\varphi_k(x) \searrow 0$ . Dann gilt  $\int \varphi_k dx \searrow 0$ .

ii)  $(\varphi_k)_k \subset \mathcal{R}_0$ ,  $\varphi_k(x) \nearrow f(x)$  und  $\sup \int \varphi_k dx < \infty \implies f$  ist f.ü. endlich.

*Beweis.* 6.4.i)  $\exists [a, b]^n$  und  $\exists M \geq 0$  s. d.  $\varphi_1|_{[a,b]}$  und somit alle  $\varphi_k$  durch  $M$  beschränkt sind und setzen  $\varphi_1$ , somit alle  $\varphi_k$  außerhalb  $[a, b]^n$  gleich 0. Wähle nun Repräsentanten.  $\forall \varepsilon > 0 \exists$  abzählbar viele offene Quadern mit  $\sum_j \text{vol}(Q_j) < \varepsilon$  s. d. alle  $\varphi_k|_{[a,b] \setminus \bigcup_j Q_j}$  stetig sind

$$\implies \|\varphi_k\|_{\infty, [a,b]^n \setminus \bigcup_j Q_j} \rightarrow 0 \implies 0 \leq \int \varphi_k dx \leq M\varepsilon + \varepsilon(b-a)^n$$

für  $k$  groß genug.

6.4.ii) Ersetze  $\varphi_k$  durch  $\tilde{\varphi}_k = \varphi_k - \varphi_1 \implies \tilde{\varphi}_k \geq 0$ . Wähle  $A > 0$  s. d.  $A \geq \sup \int \tilde{\varphi}_k dx$ . Setze zu  $\varepsilon > 0$

$$\mathcal{U}_\varepsilon := \left\{ x \in \mathbb{R}^n \mid f(x) > \frac{A}{\varepsilon} \right\}.$$

Wir sind fertig, wenn wir zeigen, dass  $\mathcal{U}_\varepsilon$  Nullmenge ist. Setze

$$\mathcal{U}_{\varepsilon,k} := \left\{ x \mid \tilde{\varphi}_k(x) > \frac{A}{\varepsilon} \geq \tilde{\varphi}_{k-1} \right\} \quad (\tilde{\varphi}_0 = 0).$$

$\implies \mathcal{U}_{\varepsilon,\parallel} \cap \mathcal{U}_{\varepsilon,j}$  (wegen Monotonie) und  $U_\varepsilon = \bigcup_k \mathcal{U}_{\varepsilon,k}$  (wobei  $\mathcal{U}_{\varepsilon,k}$  eine endliche Vereinigung von Quadern ist). Nur noch zu zeigen:

$$\sum_{j=1}^{\infty} \sum_{j=1}^{l_k} \text{vol}(Q_{k,j}) \leq \varepsilon.$$

Dies gilt dann

$$\frac{A}{\varepsilon} \sum_{k=1}^m \sum_{j=1}^{l_k} \text{vol}(Q_{k,j}) \leq \underbrace{\int_{\mathcal{U}_{\varepsilon,k}} \varphi_k dx}_{\substack{\uparrow \\ \text{Def. } \mathcal{U}_{\varepsilon,k}}} \leq \int \varphi_m dx \leq A. \quad \square$$

### Definition 6.5.

$$\mathcal{R}_1 := \left\{ f: \mathbb{R}^n \rightarrow \overline{\mathbb{R}} = \mathbb{R} \cup \{ \pm\infty \} \mid \exists (\varphi_k)_k \subset \mathcal{R}_0, \quad \varphi_k(x) \nearrow f(x) \right\}$$

zu  $f_1 \in \mathcal{R}_1$  setzte  $\int f dx := \lim_{k \rightarrow \infty} \int \varphi_k dx \in \overline{\mathbb{R}}$ .

- Satz 6.6.**
- i)  $\int f$  wie in 6.5 ist unabhängig von der gewählten Folge  $(\varphi_k)_k \subset \mathcal{R}_0$ .
  - ii) Für  $f \in \mathcal{R}_0$  stimmt der Integralbegriff mit dem aus 6.2 überein.
  - iii) Sind  $f, g \in \mathcal{R}_1$  mit  $f \leq g$  so ist  $\int f \leq \int g$ .
  - iv) Sind  $g, f \in \mathcal{R}_1$ ,  $c \geq 0$  so ist  $cf, f+g \in \mathcal{R}_1$  und  $\int cf dx = c \int f dx$ ,  $\int f+g dx = \int f dx + \int g dx$ .

Für den Beweis benötigen wir ein Lemma.

**Lemma 6.7.**  $(\varphi_k)_k, (\psi_k)_k \subset \mathcal{R}_0$ ,  $\varphi_k \nearrow f$ ,  $\psi_k \nearrow g$ . Ist  $f \leq g$ , so gilt

$$\lim_k \int \varphi_k dx \leq \lim_k \int \psi_k dx.$$

*Beweis.* Es genügt zu zeigen: Für jedes  $m$  gilt

$$\int \varphi_m dx \leq \lim \int \psi_k dx.$$

Das ist äquivalent zu

$$\lim_{k \rightarrow \infty} \int \varphi_m - \psi_k dx \leq 0.$$

Definiere

$$(\varphi_m - \psi_k)^+ := \max\{\varphi_m - \psi_k, 0\} \searrow 0 \quad (k \rightarrow \infty)$$

punktweise. 6.4.i)  $\implies \lim_{k \rightarrow \infty} \int (\varphi_m - \psi_k)^+ dx \searrow 0 \implies$  Beh.

□

33 min

*Beweis von 6.6.* 6.6.i) folgt aus 6.7 mit  $f = g$ .

6.6.iii) folgt aus 6.7.

6.6.ii) Wähle  $\varphi_n = f \in \mathcal{R}_0 \quad \forall n$ .

6.6.iv)  $\varphi_k \nearrow f$ ,  $\psi_k \nearrow g \implies c\varphi_k \nearrow cf$ ,  $\varphi_k + \psi_k \nearrow f + g$ . Der Rest folgt aus der Linearität der Integrals auf  $\mathcal{R}_0$ . □

**Bemerkung.** Achtung! Da wir  $\int f = \infty$  zulassen, kann man nicht einfach Linearität zeigen.

**Beispiel.**  $\int f = \infty$ ,  $\int g = \infty$ ,  $\int(f - g) \neq 0$  i. A.

**Konvention.**  $0 \cdot (\infty) = 0$ ,  $a + \infty = \infty$ ,  $a \cdot (\infty) = \infty$  ( $a > 0$ ),  $a \cdot (\infty) = -\infty$  ( $a < 0$ ). Konsistent mit 6.6.iv).

**Bemerkung.** Sind  $f, g \in \mathcal{R}_1$ , so auch  $\min(f, g)$  und  $\max(f, g)$ .

*Beweis.*

$$\begin{aligned} \min(f, g) &= \frac{1}{2}(f + g - |f - g|) \\ \max(f, g) &= f + g - \frac{1}{2}|f - g|. \end{aligned}$$

□

**Lemma 6.8.**  $(f_k) \subset \mathcal{R}_1$ ,  $f_k \nearrow f$ . Dann ist  $f \in \mathcal{R}_1$  und

$$\int f_k dx \nearrow \int f dx.$$

*Beweis.* Zu  $k$  wähle  $(\varphi_{k,j})_j \subset \mathcal{R}_0$  mit  $\varphi_{k,j} \nearrow_j f_k$ . Dann ist für  $\varphi_i := \sup_{k,j \leq i} \varphi_{k,j}$

$$\varphi_{i+1} \geq \varphi_i$$

und für jedes feste  $k$  gilt

$$\begin{aligned} \varphi_{k,j} &\leq \varphi_j \leq f \quad \forall j \geq k \\ \implies f_k &= \lim_{j \rightarrow \infty} \varphi_{k,j} \leq \lim_{j \rightarrow \infty} \varphi_j \leq f \end{aligned}$$

$\implies$  (Sandwich-Lemma)  $\varphi_j \nearrow f \implies f \in \mathcal{R}_1$  und  $\int \varphi_j dx \rightarrow \int f dx$ . Wegen  $\varphi_k \leq f_k \leq f$  (nach Definition) folgt Beh.  $\square$

42:15

**Bemerkung / Definition 6.9.** Wir erweitern den Integralbegriff noch weiter: Ist  $f \in \mathcal{R}_1$ , definiert man  $\int -f dx := -\int f dx$ .

Allgemein: Sind  $f_1, f_2 \in \mathcal{R}_1$  und  $\int f_1 dx - \int f_2 dx$  ist definiert (d. h. mindestens eines der Integrale ist endlich), dann ist  $f := f_1 - f_2$  wohldefiniert (denn nach 6.4.ii) ist dann eine der Funktionen f.ü. endlich). Die Menge all dieser Funktionen  $f$  nennen wir  $\mathcal{R}_2$  und setzen für  $f = f_1 - f_2 \in \mathcal{R}_2$ ,  $f_i \in \mathcal{R}_1$ :

$$\int f dx := \int f_1 dx - \int f_2 dx.$$

**Bemerkung.** Ohne Einschränkung können wir annehmen, dass  $f_1, f_2 \geq 0$  (sonst wähle  $\varphi, \psi \in \mathcal{R}_0$  mit  $\varphi \leq f_1$  und  $\psi \leq f_2$  und betrachte  $f_1 - \min\{\varphi, \psi\}$  und  $f_2 - \min\{\varphi, \psi\}$ ).



$$f = f_1 - \min\{\varphi, \psi\} - (f_2 - \min\{\varphi, \psi\}) = f_1 - f_2.$$

**Satz 6.10.**  $f \in \mathcal{R}_2$ 

- i)  $\int f dx$  ist unabhängig von der Wahl von  $f_1, f_2$
- ii) Ist  $f \in \mathcal{R}_1 \subset \mathcal{R}_2$ , stimmt  $\int f dx$  mit dem Integral auf  $\mathcal{R}_1$  (6.5) überein.
- iii) Monotonie: Sind  $f, g \in \mathcal{R}_2$ ,  $f \leq g$  so ist  $\int f dx \leq \int g dx$ .
- iv) Sind  $f, g \in \mathcal{R}_2$  und  $\int f dx + \int g dx$  ist definiert, so ist  $f + g \in \mathcal{R}_2$  und  $\int f + g dx = \int f dx + \int g dx$ .
- v) Ist  $f \in \mathcal{R}_2$ ,  $c \in \mathbb{R}$ , so ist  $cf \in \mathcal{R}_2$  und  $\int cf dx = c \int f dx$ .

*Beweis.* 6.10.i) und 6.10.iii) Wir zeigen: Wenn  $f = f_1 - f_2$  und  $g = g_1 - g_2$  und  $f \leq g$ , dann gilt:

$$\int f_1 dx - \int f_2 dx \leq \int g_1 dx - \int g_2 dx \quad (*)$$

(daraus folgt 6.10.i) mit  $f = g$ ). Fallunterscheidungen:

- a) Ist  $\int f_2 dx = \infty$  so gilt (\*).
- b) Ist  $\int f_2 dx < \infty$  und  $\int g_1 dx < \infty$ , so sind  $f_2$  und  $g_2$  endlich f.ü. (6.4.ii))  $\Rightarrow f_1 + g_2 \leq g_1 + f_2$  (wegen  $f_1 - f_2 \leq g_1 - g_2$ )  $\Rightarrow \int f_1 + \int g_2 \leq \int g_1 + \int f_2$  (6.6.iv))  $\Rightarrow (*)$ .
- c) Ist  $\int f_2 dx < \infty$  und  $\int g_2 dx = \infty$ , wähle  $(\varphi_k)_k \subset \mathcal{R}_0$  mit  $\varphi_k \nearrow g_2$ . Dann ist  $f_1 - f_2 \leq g_1 - \varphi_k \quad \forall k$ . Wende b) auf  $f_1, f_2, g_1, \varphi_k$  an  $\Rightarrow \int f_1 - \int f_2 \leq \int g_1 - \int \varphi_k \Rightarrow (*) (k \rightarrow \infty)$ .

6.10.ii) Folgt aus 6.10.i) mit  $f_1 := f$ ,  $f_2 := 0$ .

6.10.v) Folgt aus der Definition in wie in 6.6.iv)

6.10.iv) Schreibe  $f = f_1 - f_2$ ,  $g = g_1 - g_2$  mit  $f_1, f_2, g_1, g_2 \in \mathcal{R}_1$ .

- a) Ist  $\int f_2 dx < \infty$ ,  $\int g_2 dx < \infty \Rightarrow f_1 + g_2, f_2 + g_2 \in \mathcal{R}_1$  (6.6.iv)) und  $\int f_2 + g_2 dx = \int f_2 + g_2 dx < \infty$

$$\Rightarrow f + g = (f_1 + g_2) - (f_2 + g_2) \in \mathcal{R}_2$$

und

$$\begin{aligned} \int f + g dx &= \int f_1 + g_2 dx - \int f_2 + g_2 dx \\ &= \int f_1 + \int g_1 - \int f_2 - \int g_2 \\ &= \int f + \int g. \end{aligned}$$

- b) Ist  $\int f_2 = \infty$  oder  $\int g_2 = \infty$ , so sind *beide Integrale*  $\int f_1, \int g_1$  *endlich*. Denn  $f, g \in \mathcal{R}_2$  und wir hatten vorausgesetzt, dass  $\int f dx + \int g dx$  definiert ist. Wir können also Beweis a) wiederholen. ( $\int f_2 = \infty \implies \int f_1 < \infty$  und  $\int g < \infty$ ).  $\square$

**Satz 6.11 (Verallgemeinerter Satz von Beppo-Levi).** Sei  $(f_k)_k \subset \mathcal{R}_2$ ,  $f_k \nearrow f$ . Ist  $\int f_k dx > -\infty$  für ein  $k$ , so ist  $f \in \mathcal{R}_2$  und

$$\int f_k dx \nearrow \int f dx. \quad (*)$$

**Bemerkung.** Die Voraussetzung ist wichtig:

Betrachte  $f_k = -\chi_{(-\infty, 0)} + \chi_{(0, k)}$ . Dann ist  $f = \text{sgn} \notin \mathcal{R}_2$ ,  $\text{sgn}(x) = \begin{cases} 1 & x > 0 \\ 0 & x = 0 \\ -1 & x < 0. \end{cases}$

Betrachte  $f_k = -\chi_{(k, \infty)}$ . Dann ist  $f = 0 \in \mathcal{R}_2$  aber (\*) trifft nicht zu.

*Beweis von 6.11.* Für  $k \geq k_0$  gilt  $\int f_k > -\infty$  (wegen der Monotonie). Sei  $k \geq k_0$ . Schreibe  $f_k = g_k - h_k$ ,  $g_k, h_k \in \mathcal{R}_1$ . Dann ist  $\int h_k < \infty \implies \exists \varphi_k \in \mathcal{R}_0$  s. d.  $\varphi_k \leq h_k$  und  $\int h_k - \varphi_k dx < 2^{-k}$

$$\implies f_k = (\underbrace{g_k - \varphi_k}_{=\tilde{g}_k}) - (\underbrace{h_k - \varphi_k}_{\tilde{h}_k})$$

und es gilt  $\tilde{h}_k \geq 0$  und  $\int \tilde{h}_k < 2^{-k}$ ,  $\tilde{h}_k, \tilde{g}_k \in \mathcal{R}_1$ . Schreibe nun

$$f_k = (\tilde{h}_1 + \dots + \tilde{h}_{k-1} + \tilde{g}_k) - (\tilde{h}_1 + \dots + \tilde{h}_{k-1} + \tilde{h}_k) = G_k - H_k$$

mit  $H_k, G_k \in \mathcal{R}_1$ ,  $\int H_k \leq 1$ . Lemma 6.8  $\implies G_k \nearrow G$ ,  $H_k \nearrow H$  mit  $G, H \in \mathcal{R}_1$  und  $\int H \leq 1$

$$\implies f = G - H \in \mathcal{R}_2$$

und

$$\int f_k dx = \int G_k dx - \int H_k dx \rightarrow \int G dx - \int H dx = \int f dx. \quad \square$$

**Folgerung 6.12.**  $f, g \in \mathcal{R}_2 \implies \max(f, g), \min(f, g) \in \mathcal{R}_2$ .

*Beweis (nur für max).* Seien  $f = f_1 - f_2$ ,  $g = g_1 - g_2$ .

- a)  $f_1, g_1 \in \mathcal{R}_0 \implies \int f_1 + g_1 dx < \infty$  und

$$\begin{aligned} \max(f_1 - f_2, g_1 - g_2) &= (f_1 + g_1) + \max(-g_1 - f_2, -f_1 - g_2) \\ &= (\underbrace{f_1 + g_1}_{\in \mathcal{R}_1}) - \underbrace{\min g_1 + f_2, f_1 + g_2}_{\in \mathcal{R}_1}. \end{aligned}$$

b)  $\int f_1 + g_1 < \infty \implies \max(f_1 - f_2, g_1 - g_2) \in \mathcal{R}_2$ .  $f_1, g_1 \notin \mathcal{R}_0$ , wähle  $\varphi_k \nearrow f_1$ ,  
 $\leq_k \nearrow g_1$ . 6.12.a)  $\implies \max(\varphi_k - f_2, \psi_k - g_2) \in \mathcal{R}_2$  und

$$\begin{aligned} \max(\varphi_k - f_2, \psi_k - g_2) &\nearrow \max(f_1 - f_2, g_1 - g_2) \\ \int \max(\varphi_k - f_2, \psi_k - g_2) dx &> -\infty \quad \forall k, \end{aligned}$$

weil  $\int f_2 dx < \infty$  oder  $\int g_2 dx < \infty \implies$  Beh. aus Beppo-Levis Satz.

□

**Vorlesung 20**

Mo 29.06. 10:15

**6.I. Etwas Maßtheorie**

4:30

**Definition 6.13.** Sei  $X$  eine Menge,  $\Sigma$  eine Familie von Teilmengen  $E \subset X$  mit  
 $\uparrow$   
 „Sigma“

$$\begin{aligned} X &\in \Sigma \\ E \in \Sigma &\implies X \setminus E \in \Sigma \\ E_k \in \Sigma, \quad k \in I \text{ abzählbar} &\implies \bigcup_{k \in I} E_k \in \Sigma. \end{aligned}$$

$\Sigma$  heißt dann  $\sigma$ -Algebra („Sigma-Algebra“) über  $X$  und  $(X, \Sigma)$  Borel-Raum oder messbarer Raum und  $E \in \Sigma$  nennt man *messbar*.

Gegeben zwei messbare Räume  $(X, \Sigma_X), (Y, \Sigma_Y)$  nennt man  $f: X \rightarrow Y$ , wenn gilt: Ist  $E \subset Y$  messbar (also  $E \in \Sigma_Y$ ), so ist  $f^{-1}(E)$  messbar in  $X$  (also  $f^{-1}(E) \in \Sigma_X$ ). Ein Maß auf einer  $\sigma$ -Algebra über  $X$  ist eine Funktion  $\mu: \Sigma \rightarrow \mathbb{R}_{\geq 0} \cup \{+\infty\}$ , für die gilt  $\mu(\emptyset) = 0$  und die  $\sigma$ -additiv (sigma-additiv) ist, für die also gilt

$$\mu\left(\bigcup_{k=1}^{\infty} E_k\right) = \sum_{k=1}^{\infty} \mu(E_k),$$

wenn  $E_k \cap E_l = \emptyset$  für  $k + l$ .  $(X, \Sigma, \mu)$  heißt *Maßraum*.

**Beispiel.**  $(X, \Sigma, \delta_x)$ ,

$$\delta_x(E) = \begin{cases} 1 & x \in E \\ 0 & x \notin E \end{cases}, E \in \Sigma,$$

$$\text{„Dirac-Maß“. } f: E \rightarrow \mathbb{R}, \text{ „}\int f d\delta_x\text{“} = \begin{cases} f(x) & x \in E \\ 0 & x \notin E. \end{cases}$$

Ein Maß heißt  $\sigma$ -endlich, wenn es zu jedem  $E \in \Sigma$  eine abzählbare Überdeckung  $E \subset \bigcup_{k=1}^{\infty} E_k$ ,  $E_k \in \Sigma$ , gibt s. d. alle  $E_k$  endliches Maß haben, also  $\mu(E_k) < \infty$  gilt für alle  $k$ .

Ein Maß heißt vollständig, wenn gilt: Ist  $N \in \Sigma$  mit  $\mu(N) = 0$  und  $M \subset N$ , so ist  $M \in \Sigma$ .

**Bemerkung.** Aus den Definitionen folgt sofort:

- $\emptyset \in \Sigma$

- Gibt es  $E \in \Sigma$  s. d.  $\mu(E) > 0$ , so ist  $\mu(\emptyset) = 0$  automatisch erfüllt.
- Die Komposition zweier messbarer Funktionen ist messbar.

**Satz 6.14.** Setze  $\mathcal{M} := \{ E \subset \mathbb{R}^n \mid \chi_E \in \mathcal{R}_2 \}$  und  $\mu(E) := \int \chi_E dx$ . Dann ist  $(\mathbb{R}^n, \mathcal{M}, \mu)$  ein  $\sigma$ -endlicher, vollständiger Maßraum.

*Beweis.* Das folgt direkt aus dem verallgemeinerten Satz von Beppo-Levi. Details als Hausaufgabe.  $\square$

**Definition.**

$$L^1 := \left\{ f \in \mathcal{R}_2 \mid \int f dx < \infty \right\}.$$

↑  
eigentlich Äquivalenzklassen  $\{ g \mid g=f \text{ f.ü.} \}$

„Lebesgue-intergrierbare Funktionen“. Wohldefiniertheit 6.17.ii) (Unabhängigkeit von Repräsentanten!).

13 min

**Proposition 6.15.**  $L^1$  ist ein Vektorraum und  $\int: L^1 \rightarrow \mathbb{R}$ ,  $f \mapsto \int f dx$  ist lineare und monotone Abbildung.

*Beweis.* Folgt aus 6.10.iii), 6.10.iv), 6.10.v) (Eigenschaften von  $\mathcal{R}_2$ ). Unabhängigkeit von Repräsentanten folgt aus 6.3.ii).

**Bemerkung.** Es gilt:  $f, g \in L^1 \implies \min(f, g), \max(f, g) \in L^1$ , denn

$$\min(f, g) = \min(f_1 + g_2, g_1 + f_2) - (f_2 + g_1).$$

max analog.

**Satz 6.16 (Beppo-Levi, Satz über monotone Konvergenz).** Sei  $(f_k)_k \subset L^1$ ,  $f_k \nearrow f$ ,  $\sup \int f_k dx < \infty$ . Dann ist  $f \in L^1$  und

$$\int f_k dx \nearrow \int f dx.$$

*Beweis.* Folgt aus dem verallgemeinerten Beppo-Levi 6.11.  $\square$

**Folgerung 6.17.** i)  $\|f\|_{L^1} := \int |f| dx$  definiert eine Norm auf  $L^1$ .  
ii)  $N \subset \mathbb{R}^n$  ist Nullmenge  $\iff N \in \mathcal{M}$  und (s. 6.14)  $\mu(N) = \int \chi_N dx = 0$ .

*Beweis.* 6.17.i)  $\Delta$ -Ungleichung, skalarmultiplikation aus Eigenschaften des Betrags und 6.10.iv) und 6.10.v)  $\checkmark$ .  $f = 0$  (d. h.  $f = 0$  f.ü.), also  $f$  liegt in der Äquivalenzklasse der Nullfunktion, also der Null im Vektorraum  $L^1$ .

$$\implies \|f\|_{L^1} = 0 \checkmark.$$

Sei umgekehrt  $\int |f| dx = 0$ . Betachte  $f_n := n|f|$ . Dann gilt  $\int f_n = 0 \quad \forall n$ .

$$6.16 \text{ (Beppo-Levi)} \implies \begin{matrix} f_n \\ \nearrow \end{matrix} g \in L^1 \implies g = 0 \text{ f.ü.} \implies |f| = 0 \text{ f.ü.} \implies f = 0 \text{ f.ü..}$$

Wichtig:  $f \in L^1 \implies \|f\|_{L^1} < \infty$ , denn  $f \in L^1 \iff |f| \in L^1$ .

„ $\Leftarrow$ “ klar

„ $\implies$ “  $f = f_1 - f_2$  und beide  $\int f_1 < \infty$  und  $\int f_2 < \infty$ .

6.17.ii) Zu zeigen ist:  $\chi_N = 0 \iff \int \chi_N dx = 0$ .

„ $\implies$ “  $\checkmark$

„ $\Leftarrow$ “ wegen 6.17.i)  $\checkmark$ . □

Weitere Erläuterungen  
→22:15

## 6.II. Weitere Folgerungen

**Satz 6.18 (Fatou).** Sei  $(f_k)_k \subset L^1$ ,  $f_k \rightarrow f$ . Ist  $f_k \geq 0 \quad \forall k$  und  $\liminf_{k \rightarrow \infty} \int f_k dx < \infty$ , dann ist  $f \in L^1$  und

$$\int f dx \leq \liminf \int f_n dx.$$

$\liminf$  = limes inferior = der kleinste Häufungspunkt einer Folge.

*Beweis.* Setze  $g_m := \inf \{ f_k \mid k \geq m \}$ . Dann gilt  $g_m \nearrow f$  und (wegen  $\liminf \int f_k < \infty$ )  $\sup \int g_m dx < \infty$ . Also folgt  $f \in L^1$ , wenn gezeigt ist, dass  $g_m \in L^1$  (aus 6.16). Setze zu  $m \in \mathbb{N}$ ,  $l = m, m+1, \dots$

$$g_{m,l} := \min \{ f_k \mid m \leq k \leq l \} \in L^1$$

(wegen Bemerkung nach 6.15).  $g_{m,l} \searrow_l g_m$ ,  $\int g_{m,l} dx < \infty \quad \forall l \implies -g_m \in L^1$ . Es folgt mit 6.16 (monotone Konvergenz) und  $g_m \leq f_m$

$$\begin{aligned}
\int f dx &= \lim_m \int g_m dx \\
&= \lim_m \int \inf_{k \geq m} df_k x \\
&\leq \lim_m \inf_{\substack{k \geq m \\ \uparrow}} \int f_k dx \\
\int \inf_{k \geq m} f_k &\leq \int f_j \quad \forall j \geq m, \text{ also } \int \inf_{k \geq m} f_k \leq \inf_{j \geq m} \int f_j \\
&\leq \liminf_m \int f_m dx. \quad \square
\end{aligned}$$

**Satz 6.19 (Lebesgue, majorisierte Konvergenz).** Sei  $(f_k)_k \subset L^1$ ,  $f_k \rightarrow f$ . Gibt es  $g \in L^1$ , s. d.  $|f_k| \leq g \quad \forall k$ , so ist  $f \in L^1$  und  $\int f_k dx \rightarrow \int f dx$

**Bemerkung.** Voraussetzung  $|f_k| \leq g$  ist wichtig, wie folgendes Anti-Beispiel zeigt:  $f_k = k\chi_{(0, \frac{1}{k})}$ ,  $f_k \rightarrow f = 0$ , aber  $\int f_k dx = 1 \quad \forall k$ .

6.19. Wir wenden Fatou an auf  $(g - f_k)_k$  und  $(g + f_k)_k \implies g - f, g + f \in L^1$  und

$$\begin{aligned}
\int g - f dx &\leq \liminf \int g - f_k dx = \int g dx - \limsup \int f_k dx \quad (*) \\
\int g + f dx &\leq \liminf \int g + f_k dx = \int g dx + \liminf \int f_k dx \quad (**)
\end{aligned}$$

$$f = \frac{1}{2}(g + f) - \frac{1}{2}(g - f) \in L^1$$

$$\implies \limsup \int f_k dx \stackrel{(*)}{\leq} \int f dx \stackrel{(**)}{\leq} \liminf \int f_k dx \quad \square$$

**Beispiel.** Wir berechnen

$$\lim_{n \rightarrow \infty} \int_{[0,1]} \left( n \cdot \sin \left( \frac{x}{n} \right) \right)^n dx,$$

also

$$\lim_{n \rightarrow \infty} \int \underbrace{\chi_{[0,1]}(x) \left( n \sin \frac{x}{n} \right)^n}_{f_n(x)} dx.$$

Für  $x \in [0, 1]$  ist

$$|f_n(x)| = \left| n \sin \frac{x}{n} \right|^n \stackrel{\substack{\uparrow \\ \sin y \leq y, 0 \leq y \leq 1}}{\leq} x^n \leq 1 \quad \forall n$$

und

$$\begin{aligned}\lim f_n(x) &= \lim x^n = \begin{cases} 0 & x \in [0, 1) \\ 1 & x = 1 \end{cases} \\ \implies \lim_{n \rightarrow \infty} \int_{[0,1]} \left( n \cdot \sin\left(\frac{x}{n}\right) \right)^n dx &= 0.\end{aligned}$$

Ein nützliches Kriterium, um zu testen, ob  $f \in L^1$ :

**Lemma 6.20.** Sei  $f: \mathbb{R}^n \rightarrow \mathbb{R} \cup \{\pm\infty\}$  Funktion. Gibt es  $(\varphi_k)_k \subset \mathcal{R}_0$  mit

$$\|f - \varphi_k\|_{L^1} \longrightarrow 0 \quad (k \rightarrow \infty),$$

so gilt

$$\int f dx = \lim_k \int \varphi_k dx < \infty,$$

also  $f \in L^1$ .

Umgekehrt, ist  $f \in L^1$ , so gibt es zu  $\varepsilon > 0$  ein  $\varphi \in \mathcal{R}_0$  s. d.  $\|f - \varphi\|_{L^1} < \varepsilon$ .

*Beweis.* 1. Beh.

$$\begin{aligned}\left| \int \varphi_k dx - \int \varphi_l dx \right| &\leq \int |\varphi_k - \varphi_l| dx = \|\varphi_k - \varphi_l\|_{L^1} \\ &\leq \|\varphi_k - f\|_{L^1} + \|f - \varphi_l\|_{L^1}\end{aligned}$$

$\implies (\int \varphi_k dx)_k$  bilden Cauchy-Folge  $\implies$  Konvergenz in  $\mathbb{R}$ .

$$\left| \int f dx - \int \varphi_k dx \right| \leq \|f - \varphi_k\|_{L^1} \implies \int f dx = \lim \int \varphi_k dx.$$

2. Beh. O.B.d.A.  $f > 0$ .  $\exists (\varphi_k)_k \subset \mathcal{R}_0$ ,  $\varphi x_k > 0$ ,  $\varphi_k \nearrow f$  mit  $\int \varphi_k dx \nearrow \int f dx < \infty$ .

$$\implies \int f dx - \int \varphi_k dx = \int |f - \varphi_k| dx = \|f - \varphi_k\|_{L^1} < \varepsilon$$

für  $k$  groß genug. □

**Satz 6.21.**  $(L^1, \|\cdot\|_{L^1})$  ist vollständig.

*Beweis.* Sei  $(f_k)_k \subset L^1$  Cauchy-Folge bezüglich  $\|\cdot\|_{L^1}$ . Betrachte eine Teilfolge  $(f_{k_j})_j$  mit

$$\implies \|f_{k_{j+1}} - f_{k_j}\|_{L^1} < 2^{-j}$$

und

$$\begin{aligned} f(x) &:= f_{k_1}(x) + \sum_{j=1}^{\infty} (f_{k_{j+1}}(x) - f_{k_j}(x)) \\ g(x) &:= |f_{k_1}|(x) + \sum_{j=1}^{\infty} |f_{k_{j+1}}(x) - f_{k_j}(x)| \end{aligned}$$

sowie die Partialsummen

$$\begin{aligned} S_K^f(x) &:= f_{k_1}(x) + \sum_{j=1}^K (f_{k_{j+1}}(x) - f_{k_j}(x)) \\ S_K^g(x) &:= |f_{k_1}|(x) + \sum_{j=1}^K |f_{k_{j+1}}(x) - f_{k_j}(x)|. \end{aligned}$$

Es gilt

$$\begin{aligned} \|S_K^g(x)\|_{L^1} &\leq \|f_{k_1}\|_{L^1} + \sum_{j=1}^K \|f_{k_{j+1}} - f_{k_j}\|_{L^1} \\ k \rightarrow \infty &\leq \|f_{k_1}\|_{L^1} + \sum_{j=1}^k 2^{-j} \end{aligned}$$

$$\implies g \in L^1 \text{ (monotone Konvergenz, 6.16)}$$

$\implies$  auch die Reihe, die  $f$  definiert, konvergiert punktweise f.ü.

$$\implies f \in L^1.$$

Wir zeigen:  $f_k \rightarrow f$  (bzgl.  $\|\cdot\|_{L^1}$ ). Es gilt (Teleskopsummen-Argument)  $S_{K-1}(x) = f_{k_x}$ .

$$\implies f_{k_K} \rightarrow f \text{ punktweise f.ü. } (K \rightarrow \infty).$$

Zudem gilt

$$\begin{aligned} |f(x) - S_K(f)(x)| &\stackrel{\triangle}{\leq} |f(x)| + |S_K(f)(x)| \\ &\stackrel{\triangle_2}{\leq} |g(x)| \forall k \end{aligned}$$

$$\implies \text{(majorante Konvergenz 6.19)}$$

$$\int |f(x) - S_K(f)(x)| dx \rightarrow \int 0 dx = 0$$

also  $\|f - f_{k_K}\|_{L^1} \rightarrow 0$  ( $k \rightarrow \infty$ ), also:  $\forall \varepsilon > 0 \exists N$  s.d.  $\|f - f_{k_K}\|_{L^1} < \varepsilon \forall k_K \geq N$ . Schließlich:  $(f_k)_k$  ist Cauchy-Folge also  $\forall \varepsilon > 0 \exists M$  s.d.  $\|f_k - f_l\|_{L^1} < \varepsilon \forall k, l > M$   
 $\implies \|f_k - f\|_{L^1} \leq \|f_k - f_{k_K}\|_{L^1} + \|f_{k_K} - f\|_{L^1} < 2\varepsilon$  für  $k, k_K \geq \max\{N, M\}$ .  $\square$

Mann kann  $L^1([a, b], \mathbb{R})$  auch durch Vervollständigung (bzgl.  $\|\cdot\|_{L^1}$ ) von  $C([a, b], \mathbb{R})$  konstruieren.

### 6.III. Messbare Funktionen

**Lemma 6.22.**  $f: \mathbb{R}^n \rightarrow \mathbb{R} \cup \{\pm\infty\}$  ist messbar genau dann wenn  $\exists (\varphi_k)_k \subset \mathcal{R}_0$  mit  $\varphi_k \rightarrow f$  f.ü.

*Beweis.* Hausaufgabe.  $\square$

**Lemma 6.23.** i) Ist  $f \in \mathcal{R}_2$  so ist  $f$  messbar.

ii) Ist  $f$  messbar und  $|f| < g$ ,  $g \in L^1$ , so ist  $f \in L^1$ .

iii) Ist  $f$  messbar und  $f > 0$  so ist  $f \in \mathcal{R}_2$ .

*Beweis.*

i)  $f = f_1 - f_2$ ,  $f_1, f_2 \in \mathcal{R}_1$ . Seien  $(\varphi_k)_k, (\psi_k)_k \subset \mathcal{R}_0$  mit  $\varphi_k \nearrow f_1, \psi_k \nearrow f_2$ . Es gilt  $(\varphi_k - \psi_k)_k \subset \mathcal{R}_0$  und  $\varphi_k - \psi_k \rightarrow f$ .

ii) Sei  $(\varphi_k)_k$  mit  $\varphi_k \rightarrow f$ . Dann sind  $f_k := \text{med}(-g, \varphi_k, g)$  mit

$$\text{med}(x, y, z) = \begin{cases} x & y < x < z \text{ oder } z < x < y \\ y & x < y < z \text{ oder } z < y < x \\ z & x < z < y \text{ oder } y < z < x \end{cases}$$

in  $L^1$  (denn ist etwa  $x < z$ , so ist  $\text{med}(x, y, z) = \max\{x, \min\{y, z\}\}$ ). Wegen  $|f_k| < g \forall k$  folgt die Behauptung aus 6.19.

iii) Wähle  $A_1 \subset A_2 \subset \dots$ ,  $\mu(A_j) < \infty$ ,  $f|_{\mathbb{R}^n \setminus \bigcup_j A_j} = 0$ .  $f_k := \min\{f, k\chi_{A_k}\} \in L^1$  (wegen 6.23.ii)  $f_k \nearrow f$ . Beh. folgt mit 6.16).  $\square$

## 6.IV. Zum Verhältnis von Lebesgue- / Riemann-Integral

Es gibt Funktionen, die sind Lebesgue-integrierbar, aber nicht Riemann-integrierbar.

**Beispiel.**  $\chi_{\mathbb{Q}}$

Riemann-Integration:

Alle Obersummen sind 1, alle Untersummen 0. Nur Funktionen mit endlich vielen Sprungstellen / Unendlichkeiten sind Riemann-integrierbar.

Lebesgue-Integration:

$\mathbb{Q}$  = Nullmenge in  $\mathbb{R} \Rightarrow \int \chi_{\mathbb{Q}} = 0$ . Es sind abzählbar unendlich viele Sprung- / Unendlichkeitsstellen erlaubt.

**Erinnerung.** Bei der Riemann-Integration wurden die Ordinate unterteilt in Partitionen. Lebesgue kümmert sich um das Bild von  $f$  (also die Abszisse).

**Achtung:** Uneigentlich Riemann-integrierbare Funktionen müssen nicht Lebesgue-integrierbar sein.

Denn es gilt: Ist  $f \in L^1$  so ist  $|f| \in L^1$  (6.17.i) Beweis).

Ein Fall wie  $g = \text{Treppenfunktion}$  und  $g|_{[n-1,n]} = (-1)^{n+1} \frac{1}{n}$ ,  $n \in \mathbb{N}$

$$\lim_{m \rightarrow \infty} \int_0^m g(x) dx = \sum (-1)^{n+1} \frac{1}{n} < \infty$$

kann also nicht vorkommen:  $g \notin L^1$  denn  $\int |g| = \sum \frac{1}{n}$ . Allerdings:

**Satz 6.24.** Ist  $g$  Riemann-integrierbar aus  $[a, b]$ , so gibt es eine messbare Funktion  $f$  mit  $g = f$  f.ü. und das Riemann-Integral von  $g$  ist gleich dem Lebesgue-Integral von  $f$ .

(Details: Königsberger Analysis II, Kap. 7, Springer.)

Wir zeigen stattdessen:

**Satz 6.25.** Eine Regelfunktion  $f$  auf  $[a, b]$  ist über  $[a, b]$  Lebesgue-integrierbar und das Regel-Integral ist gleich dem Lebesgue-Integral.

$$\underbrace{\int_{[a,b]} f dx}_{= \int \chi_{[a,b]} f x} = \int_a^b f(x) dx$$

*Beweis.*  $I = [a, b]$ . Für jede Funktion  $h$  auf  $I$  gilt  $|h| < \|h\|_{\infty, I} \cdot \text{Id}_I$

$$\implies \|h_i\|_{L^1} < \|h\|_{\infty, I} \cdot \|\text{Id}_I\|_{L^1} = (b - a)\|h\|_{\infty, I}$$

↑  
als Funktion auf  $\mathbb{R}$ ,  $h_I|_{\mathbb{R} \setminus I} = 0$

Sei  $f$  Regelfunktion auf  $I$  und  $(\varphi_k)_k \subset \mathcal{R}_0$  mit  $\|f - \varphi_k\|_{\infty, I} \rightarrow 0$  ( $k \rightarrow \infty$ ). Dann folgt

$$\begin{aligned} & \|f_I - \varphi_{k,I}\|_{L^1} \rightarrow 0 \\ & \quad \uparrow \\ & \quad \text{als Funktion auf } \mathbb{R}, \quad f_I|_{\mathbb{R} \setminus I} = 0, \quad \varphi_{k,I}|_{\mathbb{R} \setminus I} = 0 \\ \implies & \int_{6.20} [a, b] dfx = \int f_I dx = \lim_k \int \varphi_{k,I} dx = \lim \int_a^b \varphi_k dx = \int_a^b f dx \quad \square \end{aligned}$$

**Vorlesung 21**

Do 02.07. 10:15

**Bemerkung 6.26 (Nullmengen / fast überall).** Letztes Mal hatten wir gesehen

$$\int f dx = 0 \iff f = 0 \text{ f.ü.} \quad (6.17.i))$$

$\implies$  Das Lebesgue-Integral ist auf den Äquivalenzklassen  $\{g \mid g = f \text{ f.ü.}\}$ ,  $\int f < \infty$ , wohldefiniert, denn

$$\int \underbrace{(f - g)(x)}_{=0 \text{ f.ü.}} dx = 0,$$

also  $\int f dx = \int g dx$ .

Es folgte auch:  $N$  Nullmenge  $\iff \int \chi_N = 0$ .

Aus der Definition 6.1 folgt sofort:

Teilmengen von Nullmengen sind Nullmengen und abzählbare Vereinigungen von Nullmengen sind Nullmengen.

Eine direkte Charakterisierung einer wichtigen Klasse von Nullmengen bietet folgender Satz:

**Lemma 6.27.** Sei  $U \subset \mathbb{R}^k$ ,  $f: U \rightarrow \mathbb{R}^{n-k}$  stetig. Dann ist  $\Gamma_f$  Nullmenge  $\subset \mathbb{R}^n$ .

*Beweis (für  $n = 2$ ,  $k = 1$ ).* Sei also  $f: I \rightarrow \mathbb{R}$  stetig,  $I \subset \mathbb{R}$ . Betrachte zunächst  $g = f|_{[a,b]}$ ,  $[a,b] \subset I$ .  $g$  ist gleichmäßig stetig  $\implies \forall \varepsilon > 0 \exists \delta > 0$  s. d.  $|g(x) - g(y)| < \varepsilon \forall x, y \in [a,b]$  mit  $|x - y| < \delta$

Sei als  $\varepsilon > 0$  und sei  $\delta$  wie oben und  $I_1, \dots, I_N$  eine Partition von  $[a, b]$  Intervalle s. d.  $|I_j| < \delta \quad \forall j$  ( $\exists$  da  $[a, b]$  kompakt). Setze  $\alpha_j = \inf_{x \in I_j} g(x)$ ,  $\beta_j = \sup_{x \in I_j} g(x)$

$$\implies \Gamma_{g|_{I_j}} \subset A_j = I_j \times [\alpha_j, \beta_j]$$

und

$$\text{vol}(A_j) < |I_j|(\beta_j - \alpha_j) < |I_j|\varepsilon \quad \forall j.$$

$$\implies \Gamma_g \subset \bigcup_{j=1}^N A_j \text{ und}$$

$$\text{vol}\left(\bigcup A_j\right) = \sum_{j=1}^N |I_j| \cdot \varepsilon = (b-a)\varepsilon.$$

$\implies \Gamma_g$  ist Nullmenge. Da es ein beliebiges Intervall  $I$  als abzählbare Vereinigung von kompakten Intervallen geschrieben werden kann, folgt die Behauptung.  $\square$

## 6.V. Produkt-Maße

Seien  $(X, \Sigma_X, \mu)$ ,  $(Y, \Sigma_Y, \nu)$  Maßräume. Dann ist  $\Sigma_X \otimes \Sigma_Y$  eine  $\sigma$ -Algebra in  $X \times Y$  und  $(\mu \times \nu)(E \times A) := \mu(E) \cdot \nu(A)$  definiert ein Maß auf  $X \times Y$ .

Wir zeigen nun den sehr wichtigen Satz von Fubini + Tonelli in einer sehr allgemeinen Version.

**Satz 6.28 (Fubini).** Sei  $f \in \mathcal{R}_2(X \times Y)$ . Es gilt:

$$\int_{X \times Y} f(x, y) d(x, y) = \int_X \left( \int_Y f(x, y) dy \right) dx.$$

**Bemerkung.** Tatsächlich gilt der Satz noch allgemeiner, wann immer die linke Seite definiert ist.

**Bemerkung.** Da die Rollen von  $X$  und  $Y$  auf der linken Seite vertauscht werden können (Umnummerierung) gilt dann auch

$$\int_{X \times Y} f(x, y) d(x, y) = \int_Y \left( \int_X f(x, y) dx \right) dy.$$

**Bemerkung.** Beachten sie, dass die Integrale den Wert  $\infty$  annehmen dürfen.

Achtung, die Aussage ist nicht trivial!

**Beispiel.**  $\int_R f(x, y) d(x, y)$ ,  $R = [0, 2] \times [0, 1]$ .

$$f(x, y) = \begin{cases} \frac{xy(x^2 - y^2)}{(x^2 + y^2)^3} & (x, y) \neq (0, 0) \\ 0 & \text{sonst.} \end{cases}$$

Sei  $x \neq 0$ .

$$\begin{aligned} A(x) &:= \int_{[0,1]} f(x, y) dy \\ &= \int_0^1 f(x, y) dy \\ &\stackrel{y \mapsto f(x,y) \text{ stetig auf } [0,1]}{=} \\ &= \int_{x^2}^{x^2+1} \frac{x(2x^2 - u)}{2u^3} du \\ &\stackrel{u=x^2+y^2, \frac{du}{dy}=2y}{=} \\ &= \left( -\frac{x^3}{2u^2} + \frac{x}{2u} \right) \Big|_{u=x^2}^{u=x^2+1} = \frac{x}{2(x^2 + 1)^2} \\ x = 0 &\implies f(0, y) = 0 \quad \forall y \end{aligned}$$

$\implies$  Die Formel gilt auch für  $x = 0$ .

$$\int_{[0,2]} A(x) dx = \int_0^2 \frac{x}{2(x^2 + 1^2)} dx \stackrel{u=x^2+1, \frac{du}{dx}=2x}{=} \frac{1}{4} \int_1^5 \frac{1}{u^2} du = \left. \frac{-1}{4u} \right|_1^5 = \frac{1}{5}.$$

Sei nun  $y \neq 0$ .

$$\begin{aligned} B(y) &:= \int_{[0,2]} f(x, y) dx \\ &= \int_0^2 f(x, y) dx \\ &\stackrel{u=x^2+y^2, \frac{du}{dy}=2y}{=} \frac{-2y}{(4+y^2)^2}. \end{aligned}$$

$y = 0 \implies f(x, 0) = 0 \quad \forall y \implies$  Formel ist auch für  $y = 0$  gültig.

$$\int_{[0,1]} B(y) dy = -\frac{1}{20} = \frac{1}{5}.$$

*Komisch!* Die Funktionen, die wir integriert haben, sind sogar stetig in den einzelnen Variablen und  $[0, 1]$  und  $[0, 2]$  sind kompakt. Aber offenbar  $f \notin \mathcal{R}_2$ ! Wäre  $f$  stetig auf  $\mathcal{R}_0$ , wäre  $f \chi_R (\in) \mathcal{R}_2$  ( $f$  ist nicht stetig in 0).

Auflösung:  $f = f_1 - f_2$ ,  $f_1, f_2 \geq 0$ ,  $\int f_1 \rightarrow \infty$  und  $\int f_2 \rightarrow \infty$ .

$$\begin{aligned} \int_{[\varepsilon,1]} \left( \int_{[0,2]} \frac{yx^3}{(x^2+y^2)^2} dx \right) dy &= \int_{[\varepsilon,1]} \frac{4}{y(y^2+4)^2} dy \\ &= \frac{1}{8} \left( \frac{4}{y^2+4} - \ln(y^2+4) + 2\ln y \right) \Big|_{y=\varepsilon}^1 \rightarrow \infty \quad (\varepsilon \rightarrow 0). \end{aligned}$$

Berechnung als uneigentliches Regel-Integral: Das geht, da der Integrand  $\geq 0$  ist.

$$\int_{[0,2]} \left( \int yx^3 d(yx^2 + x^2)^3 \right) dx = \int_{[0,2]} \dots = \frac{1}{8} \left( -\frac{4}{\varepsilon^2+4} - \ln 5 + \ln \left( \frac{4}{\varepsilon^2} + 1 \right) \right) \rightarrow \infty \quad (\varepsilon \rightarrow 0).$$

Für  $\frac{xy^3}{(x^2+y^2)^3}$  genauso.

21:37

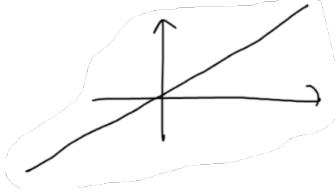
Zum Beweis von 6.28 benötigen wir ein Lemma.

**Lemma 6.29.** Ist  $N \subset X \times Y$  Nullmenge, so sind

$$N_Y(x) := \{ y \in Y \mid (x, y) \in N \}$$

(zu  $x$  fest gewählt). Nullmenge in  $Y$  für fast alle  $x$  (fast alle: gemessen in  $X$ ).

**Beispiel.**  $N = \{ (x, \frac{x}{2}) \in \mathbb{R}^2 \}$  (siehe 6.27) ist Nullmenge  $\subset \mathbb{R}^2$ .



$$N_Y(x) := \{ y \in \mathbb{R} \mid (x, y) \in N \} = \left\{ \frac{x}{2} \right\}$$

ist Nullmenge in  $\mathbb{R}$ .

*Beweis von 6.29.* Zu  $n \in \mathbb{N}$  wählen wir eine Überdeckung von  $N$  der Form  $\bigcup_{i \in I_n} Q_i^{(n)} \times P_i^{(n)}$  durch Quader  $Q_i^{(i)} \subset X$ ,  $P_i^{(i)} \subset Y$ ,

$$\text{vol} \left( \bigcup_{i \in I_n} Q_i^{(n)} \times P_i^{(n)} \right) < 2^{-n} \quad (\text{vol} = \mu \times \nu).$$

Dies liefert eine Überdeckung  $\bigcup_{k \in I} Q_k \times P_k$ ,  $I = \text{disjunkte Vereinigung aller } I_n$  (wieder abzählbar) mit

$$\text{vol} \left( \bigcup_{k \in I} Q_k \times P_k \right) < \sum_{n \in \mathbb{N}} w^{-n} = \frac{1}{1 - \frac{1}{2}} - 1 < \infty.$$

Jeder Punkt in  $N$  wird von dieser abzählbar unendlich oft überdeckt.

Es gilt für  $R_k = Q_k \times P_k$

$$\text{vol}(R_k) = \text{vol}(Q_k) \cdot \text{vol}(P_k).$$

Also

$$\begin{aligned} \mu \times \nu(R_k) &= \int \chi_{R_k}(x, y) d(x, y) \\ &= \int_X \left( \int_Y \chi_{R_k}(x, y) dy \right) dx \\ &\Rightarrow \sum_{k=1}^{\infty} \int_X \left( \int_Y \chi_{R_k}(x, y) dy \right) dx < \infty. \end{aligned}$$

Mit Beppo-Levi (6.16), angewendet auf

$$f_n(x) = \sum_{k=1}^n \int_Y \chi_{R_k}(x, y) dy,$$

folgt:

$$\sum_{k=1}^{\infty} \int_Y \chi_{R_k}(x, y) dy \quad (6.1)$$

konvergent für fast alle  $x$ . Sei  $x_0 \in X$  s. d.  $\lim_{n \rightarrow \infty} f_n(x_0) < \infty$ .

Mit Beppo-Levi, angewendet auf

$$g_n^{x_0} = \sum_{k=1}^n \chi_{R_k}(x_0, y)$$

folgt  $\sum_{k=1}^{\infty} \chi_{R_k}(x_0, y)$  konvergiert für fast alle  $y$ . Sei  $y_0 \in Y$ , s. d.  $\lim_{n \rightarrow \infty} g_n^{x_0}(y_0) < \infty$   
 $\implies (x_0, y_0) \notin N$ , also  $y_0 \notin N_Y(x_0)$ , denn für  $(x_0, y_0) \in N$  ist  $\sum_{k=1}^{\infty} \chi_{R_k}(x_0, y_0) = \infty$ .  $\square$

*Beweis des Satzes von Fubini (6.28).* Zu zeigen ist:

- a)  $f(x, \cdot) \in \mathcal{R}_0(Y)$  für fast alle  $x \in X$ ,
- b)  $\int_Y f(\cdot, y) dy \in \mathcal{R}_2(X)$ ,
- c)  $\int f(x, y) d(x, y) = \int_X (\int_Y f(x, y) dy) dx$ .

Ist  $f = \chi_{P \times Q}$ ,  $P \subset X$ ,  $Q \subset Y$ , so gelten 6.29.a), 6.29.b), 6.29.c) wegen  $\chi_{P \times Q}(x, y) = \chi_P(x)\chi_Q(y) \implies$  6.29.a), 6.29.b), 6.29.c) gelten für Stufen-Funktionen  $\mathcal{R}_0(X \times Y)$  (6.29.a), 6.29.b) mit  $\mathcal{R}_0$  statt  $\mathcal{R}_2$ ) (weil diese Linearkombinationen von Funktionen der Form  $\chi_{P \times Q}$  und das Integral linear).

Sei nun  $f \in \mathcal{R}_1(X \times Y)$ . Wähle  $(\varphi_k)_k \subset \mathcal{R}_0$  und  $N \subset X \times Y$  Nullmenge s. d.

$$\varphi_k(x, y) \nearrow f(x, y) \quad \forall (x, y) \in (X \times Y) \setminus N \implies \int \varphi_k d(x, y) \nearrow \int f d(x, y) \quad (*)$$

Lemma 6.29  $\implies N_y(x_0) = \{y \in Y \mid (x_0, y) \in N\}$  ist Nullmenge für fast alle  $x_0$ . Bezeichne  $\tilde{X} \subset X$  die Menge aller solchen  $x_0$ . Dann gilt für jedes fest gewählte  $x_0 \in \tilde{X}$   $\varphi_k(x_0, y) \nearrow f(x_0, y)$  fast überall in  $Y$  und  $(\varphi_k(x_0, \cdot)) \subset \mathcal{R}_0(Y)$  und

$$\int_Y \varphi_k(x_0, y) dy \nearrow \int_Y f(x_0, y) dy \quad \forall x_0 \in \tilde{X}, \text{ also für fast alle } x \in X.$$

Es ist zudem

$$\phi_k(x) := \int_Y \varphi_k(\cdot, y) dy \in \mathcal{R}_0(X)$$

und somit folgt:

$$g := \int_Y f(\cdot, y) dy \in \mathcal{R}_1(X)$$

und

$$\int_X \varphi_k(x) dx \nearrow \int_X g(x) dx, \quad (6.2)$$

also

$$\int_X \left( \int_Y \varphi_k(x, y) dy \right) dx \nearrow \int_X \left( \int_Y f(x, y) dy \right) dx \quad (**)$$

$\Rightarrow$  6.29.a) und 6.29.b) (mit  $\mathcal{R}_1$  statt  $\mathcal{R}_2$ ) für  $f \in \mathcal{R}_1$ . 6.29.c) folgt ebenso: Es gilt 6.29.c) für  $\varphi_k \in \mathcal{R}_0$ , also folgt die Beh. aus (\*), (\*\*).

Sei nun  $f \in \mathcal{R}_2(X \times Y)$ . Schreibe  $f = f_1 - f_2$ ,  $f_0 \in \mathcal{R}_1(X \times Y)$ ,  $f_j \in \mathcal{R}_1(X \times Y)$ . Dann gilt 6.29.a) und 6.29.b) (mit  $\mathcal{R}_1$  statt  $\mathcal{R}_2$ ) für  $f_1$  und  $f_2$  separat und somit 6.29.a) und 6.29.b) für  $f$ . Auch 6.29.c) gilt für  $f_1$  und  $f_2$  separat und da  $\int_{X \times Y} f_1(x, y) d(x, y)$  oder  $\int_{X \times Y} f_2(x, y) d(x, y)$  endlich ist, erfüllt auch  $f$  6.29.c).  $\square$

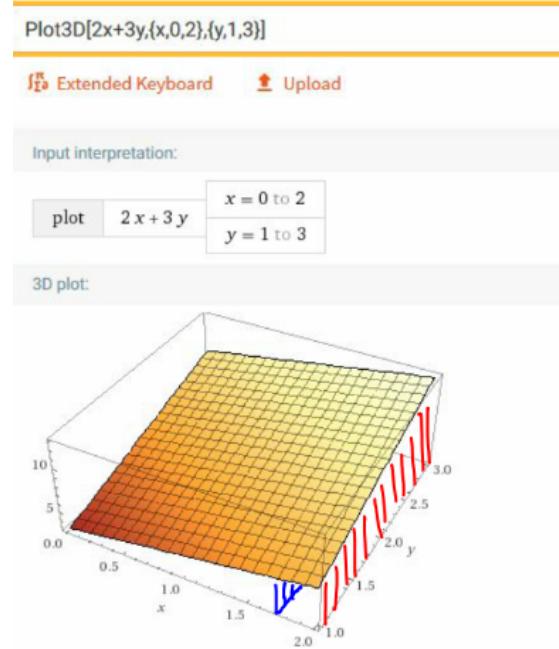
39:15

**Beispiel 6.30 (und geometrische Interpretation des Integrals).**

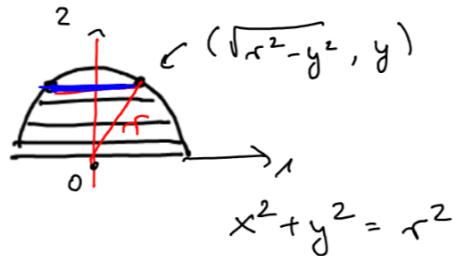
$$\int_{[a,b] \times [c,d]} 2x + 3y d(x, y).$$

$f(x, y) = 2x + 3y$  ist auf  $Q = [a, b] \times [c, d]$  stetig  $\Rightarrow$  beschränkt  $\Rightarrow f \in L^1(Q)$   
 $\Rightarrow$  (mit Fubini)

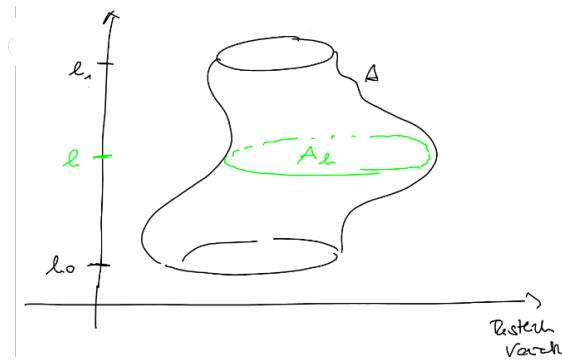
$$\begin{aligned} \int_Q (2x + 3y) d(x, y) &= \int_{[a,b]} \left( \int_{[1,3]} 2x + 3y dy \right) dx \\ &= \int_{[a,b]} \left( -\frac{1}{2} \right) (c-d)(3c+3d+4x) dx \\ &= \frac{1}{2}(b-a)(d-c)(2(a+b)+3(c+d)) \\ &= \text{Volumen unter der Fläche} \end{aligned}$$



Beispiel.  $\text{vol}(\mathcal{D})$



$$\begin{aligned}\Rightarrow \text{vol}(\mathcal{D}) &= \int_{\mathcal{D}} 1 d(x, y) \\ &= \int_{[0,r]} \left( \int_{[-\sqrt{r^2-y^2}, \sqrt{r^2-y^2}]} 1 dx \right) dy \\ &= \int_{[0,r]} 2\sqrt{r^2 - y^2} dy = \frac{1}{2}\pi r^2.\end{aligned}$$



Bemerkung 6.31 (Prinzip von Cavalieri).

$$\text{vol}(A) = \int_{[h_0, h_1]} \text{vol}(A_h) dh$$

Bemerkung 6.32 (Satz von Tonelli).  $f$  messbar,  $f \geq 0 \implies f \in \mathcal{R}_2 \implies$  Fubini gilt.

**Vorlesung 22**

Mo 06.07. 10:15

**6.VI. Der Transformationssatz**

**Ziel.** Eine Verallgemeinerung der Substitutionsregel

$$\int_{g(a)}^{g(b)} f(y) dy = \int_a^b f(g(x))g'(x) dx$$

für  $f: U \rightarrow \mathbb{R}$ ,  $U \subset \mathbb{R}^n$ .

**Satz 6.33 (Transformationssatz).** Seien  $U, V \subset \mathbb{R}^n$  offen und  $\Psi: U \rightarrow \mathbb{R}^n$  eine Diffeomorphismus auf  $\Phi(U) = V$  (d.h.  $\Phi$  ist  $C^1$ , invertierbar mit  $\Phi^{-1} \in C^1$ ). Sei  $f \in L^1(\mathbb{R}^n)$ . Dann gilt

- i)  $U \ni x \mapsto (f \circ \Phi)(x)|\det -D\Phi(x)| \in L^1(\mathbb{R}^n)$  (fortgesetzt mit 0 auf  $\mathbb{R}^n \setminus U$ ).
- ii)
- iii)  $\int_V f(y) dy = \int_U (f \circ \Phi)(v)|\det D\Phi(x)| dx$ .

*Beweis (angelehnt an Diendonné, Grundzüge der modernen Analysis).* Wir gehen vor wie im Satz von Fubini und beweisen die Aussage erst für  $f \in \mathcal{R}_0$ , dann  $f \in \mathcal{R}_1 \cap L^1$ , dann  $f \in \mathcal{R}_2 \cap L^1 = L^1$ .

Zunächst benötigen wir zwei Lemmata:

**Lemma 6.34.** Sei  $N \subset \mathbb{R}^n$  Nullmenge und  $\Phi: N \rightarrow \mathbb{R}^n$  eine lokal Lipschitz-stetige Funktion, also gelte

$$\limsup_{\substack{y \in N \\ y \rightarrow x}} \frac{\|\Phi(x) - \Phi(y)\|}{\|x - y\|} < \infty \quad \forall x \in N.$$

Dann ist  $\Phi(N)$  Nullmenge in  $\mathbb{R}^n$ .

*Beweis.* Wähle  $\|\cdot\| = \|\cdot\|_{\max}$ . Dann ist  $B_r(x)$  der in  $x$  zentrierte Würfel der Kantenlänge  $2r$ . Zu  $k, m \in \mathbb{N}$  setze

$$F_{k,n} := \left\{ x \in N \mid \frac{\|\Phi(x) - \Phi(y)\|}{\|x - y\|} \leq k \quad \forall y \in B_{\frac{1}{m}}(x) \cap N \right\}.$$

Sei  $\varphi > 0$ . Überdecke  $N$  mit offenen Kugeln  $B_j = U_{r_j}(x_j)$  mit  $x_j \in F_{k,m}$  und  $r_j < \frac{1}{m}$  und  $\sum_j \text{vol}(B_j) < \varepsilon$ . Dann ist für  $x \in I_{k,m} \cap B_j$ .

$$\begin{aligned}
& \|x - x_j\| < r_j < \frac{1}{m} \quad \forall x_j \in I_{k,m} \\
& \implies \|\Phi(x) - \Phi(x - j)\| \leq k\|x - x_j\| < kr_j \\
& \implies \Phi(F_{k,m} \cap B_j) \subset U_{k,r_j}(\Phi(x_j)) \\
& \implies \Phi(I_{k,m}) \subset \bigcup U_{k,r_j}(\Phi(x_j)) \\
& \implies \text{vol } \Phi\left(\bigcup_{k,m} F_{k,m}\right) \leq \sum_j \text{vol}(U_{k,r_j}(\Phi(x_j))) = k^n \sum_j \text{vol}(B_j) < k^n \varepsilon \\
& \implies N = \bigcup_{k,m} F_{k,m} \implies \text{Beh.} \quad \square
\end{aligned}$$

**Lemma 6.35.**  $\Phi: \mathbb{R}^n \rightarrow \mathbb{R}^n$ ,  $\Phi(x) = Ax + b$  (affin linear). Dann gilt für jede Borelmenge  $E$

$$\text{vol}(\Phi(E)) = |\det A| \text{vol}(E).$$

*Beweis.* Unabhängigkeit von  $b$ : Folgt aus der Definition, denn das Volumen eines Quaders ändert sich nicht durch Verschiebung. O.B.d.A. nehmen wir an,  $E$  ist ein Quader  $Q$ , dessen eine Ecke im Ursprung liegt. Somit o.B.d.A.  $Q = \{x \mid 0 \leq x_j \leq c_j\}$  mit  $c_j > 0$ .

Jede Seitenfläche von  $Q$  ist Nullmenge (kann als Graph einer Funktion geschrieben werden). Nach Lemma 6.34 ist das Bild einer Seitenfläche unter einer linearen Abbildung wieder Nullmenge.

Sind  $\Phi_1$  und  $\Phi_2$  lineare Abbildungen und gilt die Aussage für  $\Phi_1$  und  $\Phi_2$ , dann auch für die Verknüpfung.

Jede lineare Abbildung ist Komposition von Abbildungen der folgenden Typen:

a) Permutationsmatrizen,

b)  $\text{diag}(\alpha, 1, 1, \dots, 1)$   $\alpha \in \mathbb{R}$ ,

$$\text{c)} \quad \begin{pmatrix} 1 & 0 & & & \\ 1 & 1 & 0 & & \\ & & 1 & \dots & \\ 0 & & \dots & \dots & 1 \end{pmatrix}.$$

**6.35.a)** ändert das Volumen nicht.  $\det = \pm 1$ , also stimmt die Aussage für Permutation.

**6.35.b)** streckt das Intervall in Eins-Richtung um einen Faktor  $\alpha$ . Das neue Volumen unterscheidet sich vom alten durch den Faktor  $|\alpha|$ . Also stimmt auch hier die Aussage des Satzes.

**6.35.c)** wegen 6.35.a) und 6.35.b) o.B.d.A.  $c_1 = c_2 = \dots = c_n = 1$ .  $\Phi$  vom Typ 6.35.c):

$$\Phi(Q) = \{x \mid x_1 \leq x_2 \leq x_1 + 1 \quad 0 \leq x_j \leq 1 \quad j \neq 2\}.$$

Setze  $F_1 := \{x \in \Phi(Q) \mid x_2 \leq 1\}$ ,  $F_2 := \Phi(Q) \setminus F_1$ .

$$\begin{aligned} \Rightarrow \text{vol}(\Phi(Q)) &= \text{vol}(F_1) + \text{vol}(F_2) \\ &= \text{vol}(F_1) + \text{vol}(F_2 - e_2) \\ &\quad \uparrow \text{kanonischer Einheitsektor} \\ &= \text{vol}(Q). \\ &\quad \uparrow \\ Q = F_1 \cup (F_2 - e_2) \end{aligned}$$

Wegen  $\det \begin{pmatrix} 1 & 0 & & 0 \\ 1 & 1 & & \\ & & 1 & \dots \\ 0 & & \dots & \dots & 1 \end{pmatrix} = 1$  folgt auch hier die Beh.  $\square$

Zurück zum Beweis der Transformationsformel 6.33.

Kürze ab  $J(x) = |\det D\Phi(x)|$  ( $J$  für Jacobi). Wie im Beweis des Satzes von Fubini (6.28) zeigen wir die Aussage zunächst für Stufenfunktionen, folgern dann, dass es auch für  $f \in \mathcal{R}_1$  und auch  $\mathcal{R}_2$  gilt, somit erst recht für  $f \in L^1$ .

Wir stellen zunächst fest: Ist  $N$  Nullmenge, so  $\Phi^{-1}(N)$  auch Nullmenge (6.34) und somit ist

$$\int_V \chi_N(y) dy = 0 = \int_Z \underbrace{\chi_{\Phi^{-1}(N)}}_{=\chi_N \circ \Phi} x J(x) dx.$$

Zudem: Die Vorschriften ( $E$  Borelmenge)

$$\begin{aligned} E &\mapsto \int_V \chi_E(y) dy \\ E &\mapsto \int_x (\chi_E \circ \Phi)(x) dI(x) = \int_U \chi_{\Phi^{-1}(E)}(x) I(x) dx \end{aligned}$$

definieren Maße.

Wir betrachten nun eine Treppenfunktion  $f \in \mathcal{R}_0$ . Wegen der Linearität genügt es  $\chi_Q \in \mathcal{R}_0$ ,  $Q$  Quader, zu betrachten.

1.) Die Behauptung stimmt für  $\Phi(x) = A \cdot x + b$  (affin linear):

$$\begin{aligned} \int_V \chi_Q dy &= \mu(Q \cap) \\ &= E := \Phi(Q \cap V)^{-1} = \mu(A(E) + x_0) \\ &\stackrel{6.35}{=} \int \chi_{\Phi^{-1}(Q \cap V)} |\det A| dx. \end{aligned}$$

$$x \in \Phi^\circ(Q \cap V) \iff \Phi(x) \in Q \cap V \iff \Phi(x) \in Q \text{ und } x \in U.$$

2.) Die Behauptung stimmt, wenn  $n = 1$  ist: Dann ist  $U = \bigcup_{k=1}^{\infty} I_k$ ,  $I_k$  Intervall,  $I_k \cap I_j = \emptyset$ ,  $k \neq j$ . Betrachte also o.B.d.A.  $U = [a, b]$ .  $\det D\Phi$  hat auf  $U$  konstantes Vorzeichen. O.B.d.A.  $\det D\Phi > 0$  (also  $\Phi$  monoton wachsend). Dann ist für  $(c, d) \subset \Phi^{-1}((c, \cdot)) = (\Phi^{-1}(c), \Phi^{-1}(d))$ , also

$$\int_V \chi_Q(y) dy = \int_c^d 1 dy = d - c = \int_{\Phi^{-1}(c)}^{\Phi^{-1}(d)} \Phi'(x) dx = \int \underbrace{(\chi_Q \circ \Phi)}_{=\chi_{\Phi^{-1}(Q)}}(x) |\Phi'(x)| dx.$$

3.) Die Behauptung stimmt für  $\Phi$  von der Form

$$\Phi(x) = (x_1, \dots, x_{n-1}, \phi(x_1, \dots, x_n))^T$$

mit  $\varphi \in C^1$  und

$$\det D\Phi(x) = \det \begin{pmatrix} 1 & & & 0 \\ & \ddots & & \vdots \\ & & 1 & \\ & & & 0 \\ \partial_1 \varphi & \dots & \dots & \partial_n \varphi \end{pmatrix} = \partial_n \varphi(x) \neq 0.$$

Ist  $\tilde{x} \in \mathbb{R}^{n-1}$ , s. d.

$$U_{\tilde{x}} := \{x_n \in \mathbb{R} \mid (\tilde{x}, x_n) \in U\} \neq \emptyset,$$

so ist  $\Phi_n: x - n \mapsto \varphi(\tilde{x}, x_n)$  eine Abbildung auf eine offene Teilmenge  $\subset \mathbb{R}$ .

Mit 2.) folgt daher

$$\int_{\Phi_n(U_{\tilde{x}})} \chi_Q((\tilde{x}, x_n)) dx_n = \int_{U_{\tilde{x}}} (\chi_Q \circ \Phi)(\tilde{x}, x_n) \underbrace{\left| \frac{d}{dx_n} \Phi_n(\tilde{x}, x_n) \right|}_{=|\det D\Phi(\tilde{x}, x_n)|} dx_n.$$

Fubini  $\implies$

$$\begin{aligned} \int_V \chi_Q dx &= \int \left( \int_{\Phi_n(U_{\tilde{x}})} \chi_Q(\tilde{x}, x_n) dx_n \right) d\tilde{x} \\ &= \int \int_{U_{\tilde{x}}} (\chi_Q(\tilde{x}, x_n))(\tilde{x}, x_n) J(\tilde{x}, x_n) dx_n d\tilde{x} \\ &\quad \uparrow |\det D\Phi(\tilde{x}, x_n)| \\ &= \int_U (\chi_Q \circ \Phi)(x) J(x) dx. \end{aligned}$$

- 4.) Die Behauptung ist bewiesen, wenn wir zeigen, dass zu jedem  $x_0 \in \exists$  offene Umgebung  $U_0$ , s. d.  $\Phi|_{U_0}$  eine Komposition endlich vieler Abbildungen wie in 1.) und 3.) ist. O.B.d.A.  $x_0 = 0$  (sonst komponiere  $\Phi$  mit Translation). Ersetze  $\Phi$  durch  $D\Phi(0)^{-1} \cdot \Phi$  (Matrixmultiplikation bzw. Komposition mit einer konstanten linearen Abbildung).

Dann gilt  $D\Phi(0) = \mathbb{1}$ . Definiere

$$u_j: U \rightarrow \mathbb{R}^n, \quad u_j(x) = \begin{pmatrix} \Phi_1(x) \\ \vdots \\ \Phi_j(x) \\ x_{j+1} \\ \vdots \\ x_n \end{pmatrix}.$$

Satz von der Umkehrabbildung  $\implies \exists U_0$  Umgebung von 0, s. d.  $u_j|_{U_0}$  Diffeomorphismus auf sein Bild ist.

Es gilt zudem auf  $U_0$

$$\Phi = u_n = (u_n \circ u_{n-1}^{-1}) \circ (u_{n-1} \circ u_{n-2}^{-1}) \circ \cdots \circ (u_2 \circ u_1^{-1}) \circ u_1$$

und

$$u_j \circ u_{j-1}^{-1} = (x_1, \dots, x_{j-1}, \Phi_j(x), x_{j+1}, \dots, x_n). \quad \square$$

Durch Permutation geht diese Abbildung in eine der Form aus 3.) über.

Für  $f \in \mathcal{R}_1$  betrachte  $(\varphi_k)_k \subset \mathcal{R}_0$ ,  $\varphi_k \nearrow f$ . Dann gilt die Identität wegen Beppo-Levi auch für  $f$ . Für  $f = f_1 - f_2 \in \mathcal{R}_2$  gilt sie getrennt für  $f_1$  und  $f_2$  und somit für  $f$ .

Vergleich mit der Substitutionsformel:

Hier tritt der Betrag der Ableitung auf! Warum? Sei z. B.  $\Phi(x) = -x$ ,  $U = (a, b)$ ,

$$\begin{aligned}
 \int_{\Phi(u)} f(y) dy &= \int_{(-b, -a)} f(y) dy \\
 &= \int_{-a}^{-b} f(y) dy \\
 &= \int_a^b f(-x) dx \\
 &\quad \uparrow \text{übliche Substitutionsregel, } \int f(-x)(-1) dx, a, b \\
 &= \int_{(a, b)} f(-x) \cdot 1 dx.
 \end{aligned}$$

**Beispiel (Polarkoordinaten).** Sei  $f \in \mathcal{R}_2(\mathbb{R}^2)$ , sei  $R \in \mathbb{R}_{>0} \cup \{+\infty\}$ ,

$$D_R := \left\{ (x, y) \in \mathbb{R}^2 \mid x^2 + y^2 < R^2 \right\}.$$

Wir wollen  $\int_{D_R} f(x, y) d(x, y)$ . Dieses ist gleich dem über  $V := D_R \setminus \{(x, 0) \mid x \geq 0\}$  (denn  $\{(x, 0) \mid x \geq 0\}$  ist als Graph der stetigen Funktion  $\mathbb{R}_{\geq 0} \ni x \mapsto 0$ ).

$$\Phi: \underbrace{(0, R) \times (0, 2\pi)}_U \rightarrow V, \quad \Phi(r, a) = (r \cos a, r \sin a)$$

ist Diffeomorphismus.

$$\begin{aligned}
 D\Phi(r, a) &= \begin{pmatrix} \cos a & -r \sin a \\ \sin a & r \cos a \end{pmatrix} \\
 \det D\Phi(r, a) &= r > 0 \quad \forall (r, a) \in U \\
 \int_{D_R} f(y) dy &= \int_V f(y) dy = \int_0^R \int_0^{2\pi} f(r \cos a, r \sin a) r da dr.
 \end{aligned}$$

**Beispiel (Kugelkoordinaten).** Sei  $R \in \mathbb{R}_{>0} \cup \{+\infty\}$ ,  $B_r = U_R(0)$ .

$$\begin{aligned}
 \Phi: (0, R) \times (0, 2\pi) \times \left(-\frac{\pi}{2}, \frac{\pi}{2}\right) &\rightarrow \mathbb{R}^3 \setminus \left\{ \begin{pmatrix} x \\ 0 \\ t \end{pmatrix} \mid x \geq 0 \right\} \\
 \Phi(r, \varphi, \theta) &= \begin{pmatrix} r \cos \varphi \cos \theta \\ r \sin \varphi \cos \theta \\ r \sin \theta \end{pmatrix}
 \end{aligned}$$

ist Diffeomorphismus.

$$\det D\Phi(r, \varphi, \theta) = r^2 \cos \theta > 0$$
$$\int_{B_R} f(x_1, x_2, x_3) dx = \int_0^R \int_0^{2\pi} \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} f(r \cos \varphi \cos \theta, \dots, \dots) r^2 d\theta d\varphi dr.$$