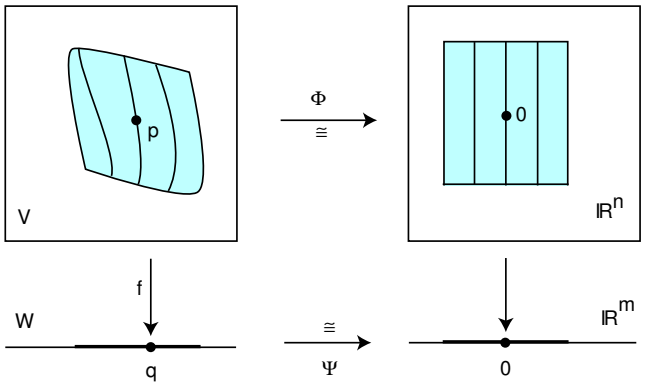
Analysis II

Prof. Dr. Dirk Ferus

Sommersemester 2007



Version vom 10.04.2008

Inhaltsverzeichnis

Literatur

Zur Analysis

Barner/Flohr: Analysis II. Walter de Gruyter, etwa Euro 30.-

Forster: Analysis II. Differential und Integralrechnung in einer Variablen. Vieweg, etwa Euro 16.-

Heuser: Lehrbuch der Analysis, Teil 2. Teubner, etwa Euro 35.-

Hildebrandt: Analysis 2, Springer, etwa Euro 28.-

Königsberger: Analysis 2, Springer, etwa Euro 25.-

Rudin, W.: Principles of Mathematical Analysis, McGraw-Hill 1964/1987

Zur Geschichte der Mathematik (und Analysis)

Moritz Cantor, Vorlesungen über die Geschichte der Mathematik, 4 Bände, um 1900

N. Bourbaki, Elements of the History of Mathematics, Springer

Website: The MacTutor History of Mathematics Archiv, http://turnbull.mcs.st-and.ac.uk/ history/

1 Grundlagen der Topologie

1.1 Topologie in metrischen Räumen

- Bevor wir mit der Analysis von Funktionen mehrerer Variabler beginnen können, müssen wir deren Definitionsbereiche, also höher-dimensionale Räume genauer kennenlernen.
- Wir legen die Grundlagen für die Definition von Konvergenz in solchen Räumen und damit für die Definitionen von Stetigkeit und Differenzierbarkeit von Funktionen auf solchen Räumen.

Metrische Räume

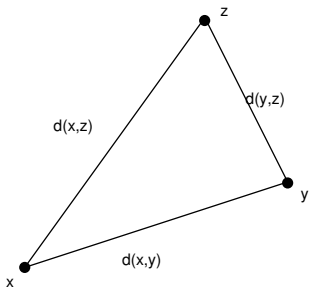
- Wir lernen, was eine Metrik zur Abstandsmessung von Punkten in einem Raum (d.h. in einer Menge) ist
- und betrachten dafür viele sehr verschiedene Beispiele. Das verdeutlicht gleichzeitig die "Universalität"abstrakter mathematischer Begriffsbildungen.

Definition 1. Ein *metrischer Raum* ist ein Paar (X, d) bestehend aus einer Menge X und einer Abbildung (der *Metrik*)

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

mit folgenden Eigenschaften für alle $x, y, z \in X$:

$$d(x, y) \geq 0 \text{ und } d(x, y) = 0 \Leftrightarrow x = y \tag{1}$$
$$d(x, y) = d(y, x) \quad \text{(Symmetrie)} \tag{2}$$
$$d(x, z) \leq d(x, y) + d(y, z) \quad \text{(Dreiecksungleichung)} \tag{3}$$



Beispiel 2. $X = \mathbb{R}$, $d(x, y) = |x - y|$.

□

Beispiel 3 (Standardmetrik auf $\mathbb{R}^n$). Wichtigstes Beispiel für dieses Semester:
Seien $n \in \mathbb{N} \setminus \{0\}$,

$$X = \mathbb{R}^n := \{(x_1, \ldots, x_n) \mid x_i \in \mathbb{R}\}$$

und für $x = (x_1, \ldots, x_n)$, $y = (y_1, \ldots, y_n) \in \mathbb{R}^n$

$$d(x, y) := \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2}.$$

Wir nennen diese Metrik die *Euklidische Metrik* oder die *Standardmetrik* auf $\mathbb{R}^n$. Nur die Dreiecksungleichung $d(x, z) \leq d(x, y) + d(y, z)$ bedarf eines Beweises. Wir verschieben ihn auf das nächste Beispiel. Aber dort ist die Situation etwas allgemeiner und damit komplizierter. Versuchen Sie einen einfacheren Beweis zu finden.

□

Beispiel 4 (l^p-Metrik). Für den $\mathbb{R}^n$ gibt es nicht nur die im letzten Beispiel angegebene Metrik, sondern viele mehr. Zum Beispiel für $p \geq 1$ die sogenannte l^p-Metrik:

$$d^p(x,y) := \sqrt[p]{\sum |x_i - y_i|^p}$$

Für $p = 1$ können Sie die Dreiecksungleichung selbst beweisen, für $p > 1$ ist das etwas komplizierter. An der Stelle $(*)$ benutzen wir die Höldersche Ungleichung aus der Analysis I:

Es gilt mit $q := \frac{p}{p-1}$ für $a_i, b_i \in \mathbb{R}$

$$\begin{aligned}\sum_{i=1}^n |a_i + b_i|^p &= \sum_{i=1}^n |a_i + b_i|\,|a_i + b_i|^{p-1} \leq \sum_{i=1}^n |a_i|\,|a_i + b_i|^{p-1} + \sum_{i=1}^n |b_i|\,|a_i + b_i|^{p-1} \\ &\underset{(*)}{\leq} \left(\sum_{i=1}^n |a_i|^p\right)^{\frac{1}{p}} \left(\sum_{i=1}^n |a_i + b_i|^{q(p-1)}\right)^{\frac{1}{q}} + \ldots \\ &= \left(\left(\sum_{i=1}^n |a_i|^p\right)^{\frac{1}{p}} + \left(\sum_{i=1}^n |b_i|^p\right)^{\frac{1}{p}}\right) \left(\sum_{i=1}^n |a_i + b_i|^p\right)^{\frac{1}{q}}.\end{aligned}$$

Division mit $(\sum_{i=1}^n |a_i + b_i|^p)^{\frac{1}{q}}$ liefert wegen $1 - \frac{1}{q} = \frac{1}{p}$

$$\left(\sum_{i=1}^n |a_i + b_i|^p\right)^{\frac{1}{p}} \leq \left(\sum_{i=1}^n |a_i|^p\right)^{\frac{1}{p}} + \left(\sum_{i=1}^n |b_i|^p\right)^{\frac{1}{p}},$$

falls $\sum_{i=1}^n |a_i + b_i|^p > 0$. Aber die Ungleichung gilt natürlich auch, wenn $\sum_{i=1}^n |a_i + b_i|^p = 0$. Mit

$$a_i := y_i - x_i,\ b_i := z_i - y_i.$$

ergibt sich die Dreiecksungleichung für d^p.

□

Beispiel 5 (l^∞-Metrik). Die sogenannte l^∞-Metrik

$$d^\infty(x,y) := \sup\{|x_i - y_i| \mid 1 \leq i \leq n\}.$$

ist eine weitere Metrik auf dem $\mathbb{R}^n$. Beweisen Sie die Dreiecksungleichung und

$$\lim_{p \to \infty} d^p(x,y) = d^\infty(x,y)$$

zur Rechtfertigung der Bezeichnung d^∞.

□

Ein exotischeres Beispiel:

Beispiel 6 (U-Bahn). Sei X die Menge der Berliner U-Bahnstationen und $d(x,y)$ für $x, y \in X$ die Länge der kürzesten Schienenverbindung zwischen x und y.

□

Beispiel 7 (Spurmetrik). Ist (X,d) ein metrischer Raum, so ist jede Teilmenge $A \subset X$ auf natürliche Weise ein metrischer Raum mit der von d *induzierten Metrik* oder *Spurmetrik*

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

für $x, y \in A$.

□

Beispiel 8 (Diskrete Metrik). Ist X eine Menge, so liefert

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

eine Metrik auf X, die sogenannte *diskrete Metrik*. Beweisen Sie die Dreiecksungleichung.

□

Definition 9 (Beschränktheit). Sei (Y,d) ein metrischer Raum.

(i) $A \subset Y$ heißt *beschränkt*, wenn gilt:

- Zu jedem $y \in Y$ gibt es ein $M \in \mathbb{R}$ mit $d(y,y') \leq M$ für alle $y' \in A$.

Ist $Y \neq \emptyset$, so ist das äquivalent zu folgender Bedingung:

- Es gibt ein $y \in Y$ und ein $M \in \mathbb{R}$ mit $d(y,y') \leq M$ für alle $y' \in A$.

(ii) Für $\emptyset \neq A \subset Y$ heißt

$$\operatorname{diam} A := \sup\left\{ d(y',y'') \;\middle|\; y',y'' \in A \right\}$$

der *Durchmesser* von A. Wir setzen $\operatorname{diam} \emptyset := 0$.

(iii) Eine Abbildung $f : X \to Y$ einer Menge X heißt *beschränkt*, wenn $f(X) \subset Y$ beschränkt ist.

Lemma 10. *Eine Teilmenge A des metrischen Raumes (Y,d) ist genau dann beschränkt, wenn sie endlichen Durchmesser hat.*

Beweis. Sie o.E. $A \neq \emptyset$. Ist A beschränkt, so gibt es $M \in \mathbb{R}$ und $y \in Y$ mit $d(y,y') < M$ für alle $y' \in A$. Also gilt für alle $y', y'' \in A$

$$d(y',y'') \leq d(y',y) + d(y,y'') < 2M.$$

Damit ist $\operatorname{diam} A \leq 2M$.

Ist umgekehrt $M := \operatorname{diam} A < \infty$ und $y \in A$, so gilt für alle $y' \in A$

$$d(y,y') \leq M,$$

also ist A beschränkt. □

Funktionenräume. Wir nehmen nun eine ganz wesentliche Erweiterung unseres Horizontes vor: Neben den Räumen, auf denen unsere Funktionen definiert sind, betrachten wir auch Räume, deren Elemente (Punkte?!) selbst Funktionen sind, sogenannte Funktionenräume. Denn wie zum Beispiel die Theorie der Potenzreihen zeigt, sind wir auch an der Konvergenz von Funktionenfolgen interessiert.

Satz 11 (Supremumsmetrik). *Seien (Y,d) ein metrischer Raum und X eine beliebige Menge, beide $\neq \emptyset$. Sei*

$$\mathcal{B}(X,Y) := \left\{ f : X \to Y \;\middle|\; f \text{ beschränkt} \right\}$$

die Menge der beschränkten Abbildungen von X in Y. Dann definiert

$$d^{sup}(f,g) := \sup\left\{ d(f(x),g(x)) \;\middle|\; x \in X \right\}$$

eine Metrik auf $\mathcal{B}(X,Y)$, die sogenannte Supremumsmetrik.

Beweis. Sei $x_0 \in X$. Dann gibt es zu $f, g \in \mathcal{B}(X, Y)$ ein M mit

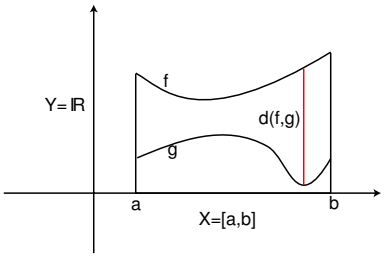
$$d(f(x_0), f(x)) < M \quad \text{und} \quad d(g(x_0), g(x)) < M$$

für alle $x \in X$. Daher ist für alle x

$$d(f(x), g(x)) < 2M + d(f(x_0), g(x_0)),$$

und $\sup_{x \in X} d(f(x), g(x)) \in \mathbb{R}$.

Also ist $d^{sup} : \mathcal{B}(X, Y) \times \mathcal{B}(X, Y) \to \mathbb{R}$ definiert.



(1), (2) sind trivial. Zur Dreiecksungleichung:

$$\begin{aligned} d^{sup}(f, h) &= \sup_x d(f(x), h(x)) \\ &\leq \sup_x \left(d(f(x), g(x)) + d(g(x), h(x))\right) \\ &\leq \sup_x d(f(x), g(x)) + \sup_x d(g(x), h(x)) \\ &= d^{sup}(f, g) + d^{sup}(g, h). \end{aligned}$$

□

Bemerkung: Für $X = \{1, \ldots, n\}$, $Y = \mathbb{R}$ sind $(\mathcal{B}(X, Y), d^{sup})$ und $(\mathbb{R}^n, d^\infty)$ isometrisch isomorph:

$$(\mathcal{B}(X, Y), d^{sup}) \cong (\mathbb{R}^n, d^\infty).$$

Das heißt, es gibt eine Bijektion $\phi : \mathcal{B}(X, Y) \to \mathbb{R}^n$, nämlich

$$\phi : f \mapsto (f(1), \ldots, f(n)),$$

für die

$$d^\infty(\phi(f), \phi(g)) = d^{sup}(f, g).$$

Topologie in metrischen Räumen

- Wir erklären, was *Umgebungen* und was *offene Mengen* in einem metrischen Raum sind und lernen deren wesentliche Eigenschaften kennen.
- Begriffe wie Konvergenz oder Stetigkeit lassen sich allein mit dem Offenheitsbegriff ohne weiteren Rückgriff auf die Metrik definieren. Das ist der Ausgangspunkt der Verallgemeinerung metrischer Räume zu sogenannten *topologischen Räumen*, aber darauf gehen wir in diesem Semester nicht näher ein.

Sei (X, d) ein metrischer Raum.

Definition 12 (Umgebung, offen, abgeschlossen). Sei $a \in X$.

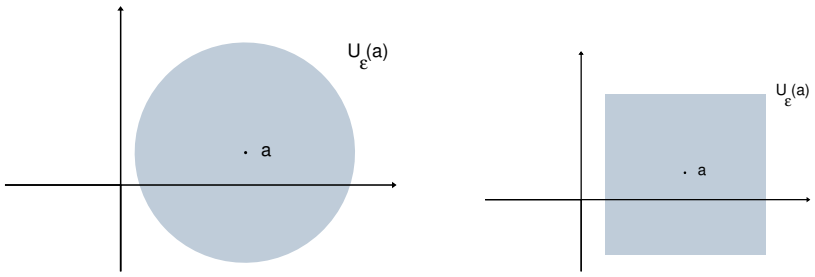
(i) Für $\epsilon > 0$ heißt
$$U_\epsilon(a) := \{x \in X \mid d(x, a) < \epsilon\}$$
die *offene Kugel* vom Radius ϵ um a oder die (offene) ϵ-*Umgebung* von a.

(ii) $U \subset X$ heißt eine *Umgebung* von a, wenn es $\epsilon > 0$ gibt, so dass $U_\epsilon(a) \subset U$.

(iii) $Y \subset X$ heißt *offen*, wenn Y eine Umgebung jedes seiner Punkte $x \in Y$ ist.

(iv) $Y \subset X$ heißt *abgeschlossen*, wenn $X \setminus Y$ offen ist.

Beispiel 13. Für die Standardmetrik bzw. die Metrik d^∞ auf $\mathbb{R}^2$ findet man:



□

Beispiel 14. Für die Supremumsmetrik auf $\mathcal{B}(X, Y)$ besteht $U_\epsilon(f)$ aus allen Funktionen $g : X \to Y$ mit
$$\sup\{d(f(x), g(x)) \mid x \in X\} < \epsilon.$$

□

Beispiel 15. Sei $(X, d) = \mathbb{R}$ mit der Standardmetrik.

$$\begin{array}{llll}]0,1[, &]1,\infty[, &]-\infty,\infty[& \text{sind offen} \\ [0,1], & [1,\infty[, &]-\infty,\infty[& \text{sind abgeschlossen,} \\ & & [0,1[& \text{ist weder offen noch abgeschlossen.} \end{array}$$

□

Beispiel 16. In jedem (X, d) sind X und $\emptyset$ sowohl offen als auch abgeschlossen.

□

Beispiel 17. Bezüglich der diskreten Metrik sind *alle* Teilmengen offen, also *alle* Teilmengen auch abgeschlossen.

□

Beispiel 18. Die offenen Kugeln $U_\epsilon(a)$ sind offen (Dreiecksungleichung).

□

Satz 19. *Die Menge der stetigen Funktionen*

$$C^0([a,b]) := \left\{ f : [a,b] \to \mathbb{R} \;\middle|\; f \text{ stetig} \right\}$$

ist eine abgeschlossene Teilmenge von $(\mathcal{B}([a,b],\mathbb{R}), d^{sup})$.

Beweis. Zunächst sind stetige Funktionen auf einem kompakten Intervall beschränkt. Also ist wirklich $C^0([a,b]) \subset \mathcal{B}([a,b],\mathbb{R})$.

Sei $f \in \mathcal{B}([a,b],\mathbb{R})$ unstetig an der Stelle $x_0 \in [a,b]$. Dann gibt es ein $\epsilon > 0$, so dass für alle $\delta > 0$ gilt:

$$\text{Es gibt ein } x \in [a,b] \text{ mit } |x - x_0| < \delta \text{ und } |f(x) - f(x_0)| \geq \epsilon.$$

Sei nun $g \in U_{\epsilon/3}(f)$, dh. $g \in \mathcal{B}([a,b],\mathbb{R})$ und $\sup |f - g| < \epsilon/3$.

Dann gibt es zu jedem $\delta > 0$ also ein $x \in [a,b]$ mit $|x - x_0| < \delta$ und

$$\epsilon \leq |f(x) - f(x_0)| \leq \underbrace{|f(x) - g(x)|}_{<\epsilon/3} + |g(x) - g(x_0)| + \underbrace{|g(x_0) - f(x_0)|}_{<\epsilon/3}.$$

Also gibt es zu jedem $\delta > 0$ ein $x \in [a,b]$ mit

$$|x - x_0| < \delta \text{ und } |g(x) - g(x_0)| \geq \frac{\epsilon}{3}.$$

Also ist g unstetig und $U_{\epsilon/3}(f)$ besteht nur aus unstetigen Funktionen. Daher ist die Menge der unstetigen Funktionen in $\mathcal{B}([a,b],\mathbb{R})$ offen und das Komplement $C^0([a,b])$ abgeschlossen. □

Satz 20 (Metrische Räume sind Hausdorffsche Räume). *Sind $x, y \in X$ zwei verschiedene Punkte eines metrischen Raumes, so gibt es Umgebungen U von x und V von y, die disjunkt sind.*

Beweis. Sei $\epsilon := d(x,y)$. Wir setzen

$$U := U_{\epsilon/2}(x), \quad V := U_{\epsilon/2}(y).$$

Wäre $z \in U \cap V$, so wäre

$$d(x,y) \leq d(x,z) + d(z,y) < \epsilon/2 + \epsilon/2 = \epsilon.$$

Das ist ein Widerspruch zur Definition von ϵ. Also gibt es kein $z \in U \cap V$: Der Durchschnitt ist leer. □

Satz 21 (Vereinigung und Durchschnitt offener Mengen). *Die Vereinigung von* beliebig *vielen und der Durchschnitt von* endlich *vielen offenen Teilmengen eines metrischen Raumes sind wieder offen.*

Beweis. Sei $(U_i)_{i\in I}$ eine Familie offener Mengen in X. Ist $x \in \bigcup_{i\in I} U_i$, so gibt es ein $i \in I$ mit $x \in U_i$. Weil U_i offen ist, gibt es dazu ein $\epsilon > 0$ mit

$$U_\epsilon(x) \subset U_i \subset \bigcup_{i\in I} U_i.$$

Also ist $\bigcup_{i\in I} U_i$ offen.

Ist andrerseits $x \in \bigcap_{i\in I} U_i$, so gibt es zu jedem $i \in I$ ein $\epsilon_i > 0$ mit

$$U_{\epsilon_i}(x) \subset U_i.$$

Wir wählen zu jedem i ein solches $\epsilon_i > 0$. Ist die Indexmenge I endlich, so ist

$$\epsilon := \min\{\epsilon_i \,|\, i \in I\}$$

positiv, und es gilt für jedes $i \in I$

$$U_\epsilon(x) \subset U_{\epsilon_i}(x) \subset U_i.$$

Daher ist

$$U_\epsilon(x) \subset \bigcap_{i\in I} U_i.$$

□

Für unendliches I klappt das letzte Argument des Beweises nicht und ist die Aussage auch nicht richtig:

Beispiel 22. Der Durchschnitt der unendlich vielen offenen Intervalle $]-\frac{1}{n}, \frac{1}{n}[\subset \mathbb{R}$ ist $\{0\}$, also nicht offen bezüglich der Standardmetrik auf $\mathbb{R}$.

Korollar 23. *Der Durchschnitt von beliebig vielen und die Vereinigung von endlich vielen abgeschlossenen Mengen sind wieder abgeschlossen.*

Beweis. Der Beweis geschieht durch „Dualisieren“. Man benutzt folgende Tatsache: Ist $(A_i)_{i\in I}$ eine *beliebige* Familie von Teilmengen von X, so gilt für $x \in X$:

$$x \in \bigcap_{i\in I} A_i \iff x \notin \bigcup_{i\in I} (X \setminus A_i).$$

Das heißt

$$\boxed{\bigcap_{i\in I} A_i = X \setminus \bigcup_{i\in I} (X \setminus A_i)}$$

und ebenso

$$\boxed{\bigcup_{i\in I} A_i = X \setminus \bigcap_{i\in I} (X \setminus A_i)}.$$

Daraus ergibt sich mit dem vorstehenden Satz die Behauptung. □

Bemerkung. Sei X eine Menge und $\mathcal{T} \subset \mathcal{P}(X)$ eine Teilmenge der Potenzmenge von X. Sind die Vereinigung beliebig vieler und der Durchschnitt endlich vieler Mengen aus $\mathcal{T}$ wieder in $\mathcal{T}$, so nennt man $\mathcal{T}$ eine *Topologie* für X und $(X, \mathcal{T})$ einen *topologischen Raum.*[1] Die Mengen aus $\mathcal{T}$ nennt man dann die *offenen Mengen* von $(X, \mathcal{T})$. Ein beliebiges $U \subset X$ heißt eine *Umgebung* von $x \in X$, wenn es eine offene Menge $V \in \mathcal{T}$ gibt, so dass

$$x \in V \subset U.$$

Topologische Räume sind eine Verallgemeinerung der metrischen Räume. In ihnen kann man Begriffe wie Konvergenz, Stetigkeit usw. einführen.

□

Satz 24 (Spurtopologie). *Seien (X, d) ein metrischer Raum und $A \subset X$ eine beliebige Teilmenge versehen mit der induzierten Metrik d_A. Dann sind die offenen Teilmengen von (A, d_A) genau die Durchschnitte offener Teilmengen von (X, d) mit A:*

$B \subset A$ ist offen in (A, d_A) $\iff$ Es gibt eine offene Teilmenge Y in (X, d) mit $B = A \cap Y$.

Der Satz gilt auch mit "abgeschlossen" statt "offen".

Beweis. Zu ($\Leftarrow$). Selbst.

Zu ($\Rightarrow$). Sei $B \subset A$ offen. Dann gibt es zu jedem $x \in B$ ein $\epsilon(x) > 0$ mit

$$B \supset U^A_{\epsilon(x)}(x) := \left\{ y \in A \;\middle|\; d_A(x, y) = d(x, y) < \epsilon(x) \right\} = A \cap U^X_{\epsilon(x)}(x),$$

wobei

$$U^X_{\epsilon(x)}(x) := \left\{ y \in X \;\middle|\; d_A(x, y) = d(x, y) < \epsilon(x) \right\}.$$

Die Menge

$$Y := \bigcup_{x \in B} U^X_{\epsilon(x)}(x)$$

ist als Vereinigung offener Teilmengen von X offen in X und es gilt $B = A \cap Y$. Damit ist „⇒“ für offene Mengen gezeigt.

Ist $B \subset A$ abgeschlossen in A, so ist $A \setminus B$ offen in A. Also gibt es nach dem eben Bewiesenen eine offene Teilmenge $Y \subset X$ mit $A \setminus B = A \cap X$ und $B = A \cap (X \setminus Y)$ist der Durchschnitt von A mit der abgeschlossenen Teilmenge $X \setminus Y$ von X. □

Beispiel 25. Sei $X =]0, 3]$ mit der Standardmetrik $d(x, y) = |x - y|$. Überlegen Sie, welche Attribute auf welche Teilmengen von X zutreffen, welche nicht:

	$]0, 1]$	$]2, 3]$	$]1, 2]$	$]0, 3]$
offen				
abgeschlossen				

□

[1] Dabei definiert man den "leeren Durchschnitt" als X und die "leere Vereinigung" als $\emptyset$. Will man diese logische Spitzfindigkeit vermeiden, so fordert man noch, dass $X \in \mathcal{T}$ und $\emptyset \in \mathcal{T}$.

Definition 26 (Inneres, abgeschlossene Hülle, Rand). Seien (X, d) ein metrischer Raum und $Y \subset X$. Wir definieren:

(i) Das *Innere von* Y oder der *offene Kern von* Y ist die Menge

$$\overset{\circ}{Y} := Y^0 := \{y \in Y \mid \exists_{\epsilon>0}\, U_\epsilon(y) \subset Y\}.$$

Die Punkte von $\overset{\circ}{Y}$ heißen *innere Punkte* von Y.

(ii) Die *abgeschlossene Hülle von* Y ist die Menge

$$\bar{Y} := X \setminus (X \setminus Y)^0.$$

(iii) Der *Rand von* Y ist die Menge

$$\partial Y := \bar{Y} \setminus \overset{\circ}{Y}.$$

Satz 27. *Sei $Y \subset X$. Dann gilt*

(i) $\overset{\circ}{Y}$ ist die größte offene Menge in Y:

$$\overset{\circ}{Y} = \bigcup_{U \text{ offen}, U \subset Y} U.$$

Insbesondere ist $\overset{\circ}{Y}$ als Vereinigung offener Mengen selber offen.

(ii) $\bar{Y}$ ist die kleinste abgeschlossene Menge, die Y enthält:

$$\bar{Y} = \bigcap_{A \text{ abgeschlossen}, A \supset Y} A.$$

Insbesondere ist $\bar{Y}$ als Durchschnitt abgeschlossener Mengen selber abgeschlossen.

(iii) Die Randpunkte von Y sind charakterisiert dadurch, dass jede ihrer Umgebungen Punkte von Y und Punkte von $X \setminus Y$ enthält:

$$\partial Y = \left\{ x \in X \;\middle|\; \forall_{\epsilon>0}\, U_\epsilon(x) \cap Y \neq \emptyset \text{ und } U_\epsilon(x) \cap (X \setminus Y) \neq \emptyset \right\}.$$

∂Y ist abgeschlossen.

Beweis. Zu (i). Sei

$$V := \bigcup_{U \text{ offen}, U \subset Y} U.$$

Zunächst gilt $\overset{\circ}{Y} \subset V$. Ist nämlich $y \in \overset{\circ}{Y}$, so gibt es $\epsilon > 0$ mit $U := U_\epsilon(y) \subset Y$, und U ist offen. Also ist $y \in V$.

Weiter ist $\overset{\circ}{Y} \supset V$. Ist nämlich $y \in V$, so gibt es ein offenes $U \subset Y$ mit $y \in U$. Da U offen ist, gibt es ein $\epsilon > 0$ mit $U_\epsilon(y) \subset U \subset Y$. Also ist $y \in \overset{\circ}{Y}$.

Damit ist $\overset{\circ}{Y} = V$.

Zu (ii). Das beweisen wir durch „Dualisieren“.

$$\bar{Y} = X \setminus (X \setminus Y)^0 = X \setminus \left(\bigcup_{(X \setminus Y) \supset U \text{ offen}} U \right) = \bigcap_{(X \setminus Y) \supset U \text{ offen}} (X \setminus U)$$
$$= \bigcap_{Y \subset (X \setminus U), U \text{ offen}} (X \setminus U) = \bigcap_{Y \subset A,\ A \text{ abgeschlossen}} A$$

Zu (iii). Nach Definition ist

$$\bar{Y} \setminus Y^0 = (X \setminus (X \setminus Y)^0) \setminus Y^0 = X \setminus \left((X \setminus Y)^0 \cup Y^0 \right).$$

In $(X \setminus Y)^0$ liegen alle Punkte, die eine ganz in $X \setminus Y$ liegende Umgebung besitzen. In Y^0 liegen alle Punkte, die eine ganz in Y liegende Umgebung besitzen. Also besteht $\bar{Y} \setminus Y^0$ genau aus den Punkten, deren sämtliche Umgebungen sowohl Y wie $X \setminus Y$ treffen. Daraus folgt $\partial Y = \bar{Y} \setminus Y^0$ und auch die Abgeschlossenheit von ∂Y.

□

Beispiel 28. Seien

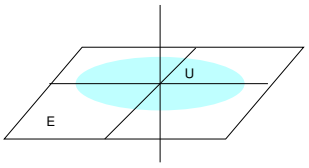
$$U := \left\{ (x, y, 0) \in \mathbb{R}^3 \,\middle|\, x^2 + y^2 < 1 \right\} \subset \mathbb{R}^3$$
$$E := \left\{ (x, y, 0) \in \mathbb{R}^3 \,\middle|\, x, x \in \mathbb{R} \right\} \subset \mathbb{R}^3$$

Wir betrachten den $\mathbb{R}^3$ mit der Standardmetrik d. Dann ist das Innere $\overset{\circ}{U} = \emptyset$ und

$$\partial U = \bar{U} = \left\{ (x, y, 0) \in \mathbb{R}^3 \,\middle|\, x^2 + y^2 \leq 1 \right\}.$$

Betrachtet man hingegen U als Teilmenge von (E, d_E), so ist

$$\overset{\circ}{U} = U,$$
$$\bar{U} = \left\{ (x, y, 0) \in \mathbb{R}^3 \,\middle|\, x^2 + y^2 \leq 1 \right\},$$
$$\partial U = \left\{ (x, y, 0) \in \mathbb{R}^3 \,\middle|\, x^2 + y^2 = 1 \right\}.$$



□

1.2 Konvergenz

- Wir definieren Konvergenz in metrischen Räumen.
- Wir konkretisieren das an sehr verschiedenen Beispielen.
- Wir untersuchen den Zusammenhang mit der Offenheit und Abgeschlossenheit von Mengen.
- Als wichtige Sätze lernen wir das Schachtelungsprinzip und den Banachschen Fixpunktsatz kennen.

Sei (X, d) ein metrischer Raum.

Definition 29 (Konvergente Folgen). Sei $(x_k)_{k\in\mathbb{N}}$ eine Folge in (X, d). Die Folge heißt *konvergent gegen* $a \in X$, und wir schreiben

$$\lim_{k\to\infty} x_k = a \quad \text{oder} \quad x_k \to a,$$

wenn $\lim_{k\to\infty} d(x_k, a) = 0$.

Dann heißt a der *Limes* oder *Grenzwert* der Folge.

Die Folge heißt *konvergent*, wenn es ein $a \in X$ gibt, so dass (x_k) gegen a konvergiert.

Eine nicht konvergente Folge heißt *divergent*.

Sei E eine Eigenschaft, die für alle Glieder einer Folge (x_k) wahr oder falsch ist.
Wie im letzten Semester vereinbaren wir folgende äquivalente Sprechweisen:

- Es gibt ein $k_0 \in \mathbb{N}$, so dass E wahr ist für alle x_k mit $k \geq k_0$.
- E gilt für *fast alle* Folgenglieder (oder für *fast alle* k).
- E gilt für *alle hinreichend großen* k.

Dann ist $\lim_{k\to\infty} x_k = a$, wenn in jeder Umgebung von a fast alle Glieder der Folge (x_k) liegen oder, äquivalent, wenn in jedem $U_\epsilon(a)$ mit $\epsilon > 0$ fast alle Glieder der Folge (x_k) liegen.

Nach der Hausdorff-Eigenschaft ist der Grenzwert einer konvergenten Folge eindeutig bestimmt.

Beispiel 30. Für $\mathbb{R}$ mit der Standardmetrik ist diese Konvergenz die übliche aus Analysis I.

□

Satz 31 (Koordinatenweise Konvergenz). *Wir betrachten eine Folge $(x_k)_{k\in\mathbb{N}}$ im $\mathbb{R}^n$ mit der Standardmetrik und schreiben $x_k = (x_{k1}, \ldots, x_{kn})$. Weiter sei $a = (a_1, \ldots, a_n) \in \mathbb{R}^n$. Dann gilt*

$$\lim_{k\to\infty} x_k = a \iff \lim_{k\to\infty} x_{kj} = a_j \text{ für alle } j.$$

Dasselbe gilt auch für $\mathbb{R}^n$ mit der d^p-Metrik, $1 \leq p \leq \infty$ beliebig.

Beweis. Sei $1 \leq p < \infty$ Für alle $k \in \mathbb{N}$ und $j \in \{1, \ldots, n\}$ gilt

$$|x_{kj} - a_j| \leq \underbrace{\left(\sum_{i=1}^{n} |x_{ki} - a_i|^p\right)^{1/p}}_{d(x_k,a)} \leq n^{1/p} \sup_{1\leq i\leq n} |x_{ki} - a_i|.$$

Daraus folgt die Behauptung. Wie argumentiert man für $p = \infty$? □

Bemerkung. Die Ungleichung im letzten Beispiel hat als einfache Konsequenz die Abschätzung

$$d^\infty(x,y) \le d^p(x,y) \le n^{1/p} d^\infty(x,y). \tag{4}$$

Schließen Sie daraus, dass die offenen Mengen in $(\mathbb{R}^n, d^p)$ für jedes $p \in [1, +\infty[$ dieselben sind wie in $(\mathbb{R}^n, d^\infty)$.

Satz 32 (Folgen-Abgeschlossenheit). *Sei A eine Teilmenge des metrischen Raumes (X,d). Dann sind die beiden folgenden Aussagen äquivalent:*

(i) A ist abgeschlossen.

(ii) Für jede konvergente Folge $(x_k)_{k\in\mathbb{N}}$ mit $\lim x_k = x$ gilt:

$$(\forall_k\, x_k \in A) \implies x \in A.$$

A ist also genau dann abgeschlossen, wenn es bezüglich der Grenzwertbildung abgeschlossen ist.

Beweis. *Zu (i) $\implies$ (ii).* Sei A abgeschlossen. Sei (x_k) eine Folge in A und $\lim x_k = x \in X$. Zu zeigen: $x \in A$. Wäre $x \notin A$, so läge a also in der *offenen* Menge $X \setminus A$, und diese wäre eine Umgebung von x. Dann lägen fast alle x_k in $X \setminus A$. Es liegt aber *kein* x_k in dieser Menge.

Zu (ii) $\impliedby$ (i). Der Grenzwert jeder konvergenten Folge in A liege wieder in A. Wir zeigen $X \setminus A$ ist offen. Andernfalls gibt es nämlich einen Punkt $x \notin A$, so dass kein $U_\epsilon(x)$ ganz in $X \setminus A$ liegt. Dann gibt es zu jedem $k \in \mathbb{N}$ ein $x_k \in A$ mit $x_k \in U_{\frac{1}{k+1}}(x)$. Offenbar konvergieren die x_k gegen $x \in X \setminus A$. Widerspruch! □

Beispiel 33 (Produktmengen). Seien $A_1, \ldots, A_n \subset \mathbb{R}$ abgeschlossen. Dann ist auch $A_1 \times \ldots \times A_n \subset \mathbb{R}^n$ abgeschlossen. Das folgt unmittelbar aus dem vorstehenden Satz in Verbindung mit Satz 31.

□

Definition 34 (Gleichmäßige Konvergenz). Eine Folge $(f_k)_{k\in\mathbb{N}}$ von Abbildungen

$$f_k : X \to Y$$

der Menge $X \neq \emptyset$ in den metrischen Raum (Y,d) heißt *gleichmäßig konvergent* gegen $f : X \to Y$, wenn gilt:

$$\forall_{\epsilon>0}\, \exists_{k_0\in\mathbf{N}}\, \forall_{k\ge k_0}\, \forall_{x\in X}\; d(f_k(x), f(x)) < \epsilon.$$

Das ist äquivalent zu

$$\forall_{\epsilon>0}\, \exists_{k_0\in\mathbf{N}}\, \forall_{k\ge k_0}\, \sup_{x\in X} d(f_k(x), f(x)) < \epsilon.$$

Beispiel 35. Eine Folge (f_k) in $\mathcal{B}(X,Y)$ ist bezüglich der sup-Metrik konvergent gegen $f \in \mathcal{B}(X,Y)$ genau dann, wenn sie *gleichmäßig* gegen f konvergiert.

Der Begriff der gleichmäßigen Konvergenz macht allerdings auch für Folgen unbeschränkter Funktionen einen Sinn.

□

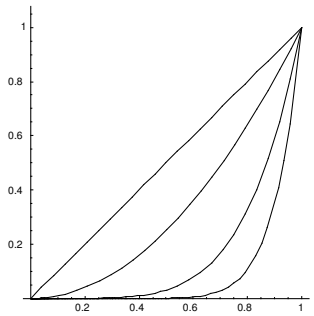
Beispiel 36 (Ungleichmäßige Konvergenz). Sei $f_k : [0,1] \to [0,1], x \mapsto x^k$ und sei

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

Dann konvergiert für jedes $x \in [0,1]$ die Folge $(f_k(x))_{k\in\mathbb{N}}$ gegen $f(x)$.
Man sagt, die Funktionenfolge (f_k) *konvergiert punktweise gegen* f. Aber für jedes k ist

$$\sup_{0\le x\le 1} |f_k(x) - f(x)| = 1,$$

und daher ist die Konvergenz nicht gleichmäßig.



□

Satz 37. *Sei $(f_k : [a,b] \to \mathbb{R})_{k\in\mathbb{N}}$ eine Folge stetiger Funktionen. Ist diese Folge gleichmäßig konvergent gegen $f : [a,b] \to \mathbb{R}$, so ist f stetig.*

Beweis. Als stetige Funktionen auf einem kompakten Intervall sind die f_k beschränkt. Also liegt die Folge in $\mathcal{B}([a,b],\mathbb{R})$. Auch die Grenzfunktion liegt in diesem Raum, weil sie sich von f_k für großes k nur wenig, zum Beispiel weniger als 1 unterscheidet. Also ist sie auch beschränkt. Die Folge (f_k) konvergiert also in $\mathcal{B}([a,b],\mathbb{R})$ gegen f. Nach Satz 19 ist $C^0([a,b])$ abgeschlossen, und nach Satz 32 liegt der Grenzwert f dann auch in $C^0([a,b])$. □

Definition 38 (Cauchyfolge). Eine Folge (x_k) in (X,d) heißt *Cauchyfolge*, wenn gilt:

$$\forall_{\epsilon>0}\ \exists_{k_0\in\mathbf{N}}\ \forall_{k,l>k_0}\ d(x_k,x_l) < \epsilon.$$

Ist jede Cauchyfolge in (X,d) konvergent, so heißt (X,d) *vollständig.*

Beispiel 39. Jede konvergente Folge ist eine Cauchyfolge. Das beweist man wie in der Analysis I.

□

Beispiel 40. Der $\mathbb{R}^n$ mit jeder der Metriken $d^p, 1 \le p$ ist vollständig. Ist nämlich $(x_k)_{k\in\mathbb{N}}$ eine Cauchyfolge, so ist wegen

$$|x_{ki} - x_{li}| \le \left(\sum_j |x_{kj} - x_{lj}|^p\right)^{1/p} = d^p(x_k,x_l)$$

für alle i auch $(x_{ki})_{k\ in\mathbb{N}}$ eine Cauchyfolge, also konvergent. Aber koordinatenweise Konvergenz bedeutet Konvergenz im $(\mathbb{R}^n, d^p)$.

□

Beispiel 41. $(\mathcal{B}(X,Y), d^{sup})$ ist vollständig, falls (Y,d) vollständig ist. (Beweis selbst).

□

Satz 42. *Ist (X, d) vollständig und $A \subset X$ versehen mit der induzierten Metrik d_A, so gilt*

$$(A, d_A) \text{ vollständig } \Leftrightarrow A \text{ abgeschlossen.}$$

Beweis. Zu ($\Longrightarrow$). Sei (A, d_A) vollständig und sei $(a_k)_{k\in\mathbb{N}}$ eine Folge in A, die gegen $x \in X$ konvergiert. Wir müssen zeigen, dass $x \in A$.

Als konvergente Folge ist (a_k) eine d-Cauchyfolge. Damit ist sie aber wegen $d(a_k, a_l) = d_A(a_k, a_l)$ auch eine d_A-Cauchyfolge und nach Voraussetzung konvergent gegen ein $a \in A$. Das bedeutet $\lim d_A(a_k, a) = \lim d(a_k, a) = 0$. Also konvergiert (a_k) auch in X gegen a. Dann ist aber wegen der Eindeutigkeit des Grenzwertes $x = a \in A$.

Zu ($\Longleftarrow$). Seien nun A abgeschlossen und (a_k) eine Cauchyfolge in (A, d_A). Das ist dann auch eine Cauchyfolge in (X, d), also konvergent gegen ein $x \in X$. Weil A abgeschlossen ist, liegt x in A und ist der Grenzwert von (a_k) in (A, d_A). □

Satz 43 (Schachtelungsprinzip). *Sei (X, d) vollständig und*

$$A_0 \supset A_1 \supset \dots \tag{5}$$

eine „absteigende" Folge von abgeschlossenen Mengen $\neq \emptyset$. Es gelte $\lim_{k\to\infty} \operatorname{diam} A_k = 0$. Dann gibt es genau ein $x^* \in X$, das in allen A_k liegt:

$$\{x^*\} = \bigcap_{k=0}^{\infty} A_k.$$

Beweis. Eindeutigkeit. Sind $x_0^*, x_1^* \in \bigcap_{k=0}^{\infty} A_k$, so gilt

$$x_0^*, x_1^* \in A_k \quad \text{für alle } k \in \mathbb{N},$$

und daher

$$d(x_0^*, x_1^*) \leq \operatorname{diam} A_k.$$

Daher ist $d(x_0^*, x_1^*) = 0$, also $x_0^* = x_1^*$.

Existenz. Wähle aus jedem A_k ein a_k. Wir wollen zunächst zeigen, dass die Folge (a_k) eine Cauchyfolge ist. Sei also $\epsilon > 0$. Da $\operatorname{diam} A_k \to 0$ gibt es ein $k_0 \in \mathbb{N}$ mit

$$\operatorname{diam} A_k \leq \epsilon \quad \text{für alle } k \geq k_0.$$

Sind $k, l \geq k_0$, so sind nach (5)

$$a_k, a_l \in A_{k_0}.$$

Damit ist aber

$$d(a_k, a_l) \leq \epsilon$$

und (a_k) ist eine Cauchyfolge. Weil (X, d) vollständig ist, ist sie konvergent gegen ein $x^* \in X$, und es bleibt zu zeigen, dass $x^* \in A_k$ für alle k. Aber die Folge $(a_l)_{l\geq k}$ ist, wiederum wegen (5), eine Folge in der abgeschlossenen Menge A_k, und deshalb liegt ihr Grenzwert in A_k. □

Lemma 44 (Cauchyfolgenkriterium). *Sei $(x_k)_{k\in\mathbb{N}}$ eine Folge im metrischen Raum (X, d). Setze $a_k := d(x_k, x_{k+1})$. Ist $\sum_{k=0}^{\infty} a_k$ konvergent, so ist $(x_k)_{k\in\mathbb{N}}$ eine Cauchyfolge.*

Die Umkehrung gilt aber nicht, finden Sie ein Gegenbeispiel.

Beweis. Nach der Dreiecksungleichung ist

$$d(x_k, x_{k+l}) \leq \sum_{j=k}^{k+l-1} d(x_j, x_{j+1}) = \sum_{j=k}^{k+l-1} a_j = |s_{k+l-1} - s_{k-1}|,$$

wenn s_k die k-te Partialsumme der Reihe $\sum_{k=0}^{\infty} a_k$ bezeichnet. Die bilden aber nach Voraussetzung eine Cauchyfolge. □

Definition 45. Eine Abbildung $f : X \to Y$ zwischen metrischen Räumen $(X, d_X), (Y, d_Y)$ heißt *kontrahierend*, wenn es ein $\lambda \in]0, 1[$ gibt, so dass für alle $x_1, x_2 \in X$

$$d_Y(f(x_1), f(x_2)) \leq \lambda d_X(x_1, x_2).$$

Wichtig ist die echte Ungleichung $\lambda < 1$. Die Zahl λ heißt dann auch ein *Kontraktionsmodul* von f.

Satz 46 (Banachscher Fixpunktsatz). *Seien (X, d) ein nicht-leerer, vollständiger metrischer Raum und $f : X \to X$ eine kontrahierende Abbildung. Dann hat f genau einen Fixpunkt: Es gibt genau ein $x^* \in X$ mit $f(x^*) = x^*$.*
Ist $x_0 \in X$ und definiert man rekursiv $x_{k+1} = f(x_k)$ für alle $k \in \mathbb{N}$, so ist die Iterationsfolge $(x_k)_{k\in\mathbb{N}}$ konvergent gegen x^*.

Beweis. Zur Einzigkeit. Aus $f(x_1^*) = x_1^*$ und $f(x_2^*) = x_2^*$ folgt

$$d(x_1^*, x_2^*) = d(f(x_1^*), f(x_2^*)) \leq \lambda d(x_1^*, x_2^*) < d(x_1^*, x_2^*).$$

Daraus folgt $d(x_1^*, x_2^*) = 0$, also $x_1^* = x_2^*$.

Existenz. Seien $x_0 \in X$ und dazu $(x_k)_{k\in\mathbb{N}}$ definiert wie im Satz. Wir zeigen, dass $(x_k)_{k\in\mathbb{N}}$ eine Cauchyfolge ist. Zunächst ist

$$d(x_k, x_{k+1}) \leq \lambda d(x_{k-1}, x_k) \leq \lambda^2 d(x_{k-2}, x_{k-1}) \leq \ldots \leq \lambda^k d(x_0, x_1)$$

Weil die Reihe $\sum_{k=0}^{\infty} \lambda^k d(x_0, x_1)$ konvergiert, ist nach dem Cauchyfolgenkriterium (x_k) also eine Cauchyfolge und wegen der Vollständigkeit konvergent gegen ein $x^* \in X$.

Es bleibt zu zeigen, dass x^* ein Fixpunkt von f ist. Beachten Sie dazu, dass für alle k

$$\begin{aligned} d(f(x^*), x^*) &\leq d(f(x^*), x_{k+1}) + d(x_{k+1}, x^*) \leq d(f(x^*), f(x_k)) + d(x_{k+1}, x^*) \\ &\leq \lambda \underbrace{d(x^*, x_k)}_{\to 0} + \underbrace{d(x_{k+1}, x^*)}_{\to 0}. \end{aligned}$$

Es folgt $d(f(x^*), x^*) = 0$, also $f(x^*) = x^*$. □

Bemerkung. Die Behauptung des Satzes gilt auch unter der schwächeren(!) Voraussetzung, dass nicht f, aber für ein $m \in \mathbb{N}$ die m-te Iteration $f^m := f \circ \ldots \circ f$ kontrahierend ist. Beweis als Übung.

1.3 Kompaktheit

- Mit der Kompaktheit lernen wir einen zentralen Begriff der Topologie, der Analysis und der Geometrie kennen.
- Die Definition mittels offener Überdeckungen ist logisch nicht ganz einfach, dafür aber an die vielfältigen Verwendungen der Kompaktheit angepasst. Da müssen Sie also durch ....
- Im $\mathbb{R}^n$ gibt es eine einfache Charakterisierung kompakter Mengen (Satz von Heine-Borel), die aber in unendlich-dimensionalen (und vielen anderen) Räumen nicht greift.
- Mit der Hausdorffmetrik sprechen wir kurz den Themenkreis der fraktalen Geometrie an.

Definition 47 (Kompaktheit). Seien (X, d) ein metrischer Raum und $A \subset X$.

(i) Eine *offene Überdeckung* von A ist eine Familie $(U_i)_{i\in I}$ von offenen Teilmengen $U_i \subset X$, so dass $A \subset \bigcup_{i\in I} U_i$.

(ii) A heißt *kompakt*, wenn jede offene Überdeckung von A eine *endliche* Überdeckung von A enthält, d.h. wenn gilt: Ist $(U_i)_{i\in I}$ eine offene Überdeckung von A, so gibt es eine *endliche* Teilmenge $K \subset I$, so dass $A \subset \bigcup_{i\in K} U_i$. Man nennt $(U_i)_{i\in K}$ dann auch eine endliche *Teilüberdeckung*[2] von A.

Beispiel 48. Sei (a_n) eine gegen $a \in X$ konvergente Folge in einem metrischen Raum. Dann ist $A := \{a\} \cup \{a_n \,|\, n \in \mathbb{N}\}$ kompakt. Hat man nämlich eine offenen Überdeckung $(U_i)_{i\in I}$ gegeben, so liegt a in einer der offenen Mengen. In dieser liegen dann aber fast alle a_k, und man braucht nur noch endlich viele weitere U_i, um den Rest „einzufangen“.

□

Beispiel 49. Eine Teilmenge A eines metrischen Raumes (X, d) ist genau dann kompakt, wenn sie als (triviale) Teilmenge von (A, d_A) kompakt ist. Beweisen Sie das!

□

Verallgemeinerung auf topologische Räume. Da die vorstehende Definition nur den Begriff offener Mengen, nicht aber explizit die Metrik benutzt, überträgt sie sich unmittelbar auf topologische Räume. Auch die vorstehenden Beispiele übertragen sich.

Satz 50. *Seien $I_1, \ldots, I_n$ abgeschlossene und beschränkte Intervalle in $\mathbb{R}$, $I_k = [a_k, b_k]$. Dann ist der abgeschlossene Quader*

$$Q := I_1 \times \ldots \times I_n$$

kompakt.

Beweis. Sei $(U_i)_{i\in I}$ eine offene Überdeckung von Q.

Annahme: Keine endliche Teilfamilie von $(U_i)_{i\in I}$ überdeckt ganz Q.

[2] Der Name ist etwas problematisch: „Teil“ heißt nicht, dass nur ein Teil von A überdeckt wird, sondern, dass man nur einen Teil der Familie offener Mengen – genauer: eine Teilmenge der Indizes – benutzt.

Wir zerlegen Q durch Halbieren aller Seiten in 2^n abgeschlossene Teilquader vom halben Durchmesser. Dann gibt es wenigstens eines dieser Teilquader, wir nennen es Q_1, welches nicht durch endlich viele der U_i zu überdecken ist. Wir zerlegen Q_1 durch Halbieren aller Seiten in 2^n abgeschlossene Teilquader vom halben Durchmesser. Dann gibt es wenigstens einen dieser Teilquader, wir nennen es Q_2, welches nicht durch endlich viele der U_i zu überdecken ist. Durch Fortsetzung dieses Verfahrens finden wir eine Folge von abgeschlossenen Quadern

$$Q \supset Q_1 \supset Q_2 \supset \dots$$

mit $\operatorname{diam} Q_k \to 0$, deren keines sich durch endlich viele der U_i überdecken läßt. Nach dem Schachtelungsprinzip gibt es $x \in \bigcap Q_k \subset Q$. Nach Voraussetzung gibt es ein i_0 mit $x \in U_{i_0}$. Weil U_{i_0} offen ist, gibt ein $\epsilon > 0$ mit $U_\epsilon(x) \subset U_{i_0}$. Dann liegt aber jeder Quader Q_k vom Durchmesser $< \epsilon$ ganz in U_{i_0}. Widerspruch! □

Für $n = 1$ liefert der Satz:

Korollar 51. *Intervalle $[a, b]$ mit $-\infty < a \leq b < +\infty$ sind kompakt in $\mathbb{R}$.*

Satz 52. *Eine kompakte Teilmenge eines metrischen Raumes ist abgeschlossen und beschränkt.*

Beweis. Sei $A \subset X$ kompakt. Zum Beweis müssen wir geeignete offene Überdeckungen von A konstruieren und ausnutzen, dass sie endliche Teilüberdeckungen besitzen.

Zur Beschränktheit. Ist $X = \emptyset$ so ist nichts zu zeigen. Andernfalls sei $x \in X$ und für $k \in \mathbb{N}$ sei

$$U_k := U_{k+1}(x).$$

Jedes $a \in A$ liegt dann in U_k, sobald $k + 1 > d(a, x)$. Also bildet $(U_n)_{n\in\mathbb{N}}$ eine offene Überdeckung von A, und wegen der Kompaktheit gibt es ein $n \in \mathbb{N}$ mit

$$A \subset \bigcup_{k=0}^{n} U_k = U_n.$$

Daher ist A beschränkt mit einem Durchmesser $\leq 2(n + 1)$.

Zur Abgeschlossenheit. Sei $x \in X \setminus A$. Zu jedem $a \in A$ sei U_a eine offene Kugel um a mit Radius $\frac{1}{2}\, d(a, x)$. Offenbar bildet $(U_a)_{a\in A}$ eine offene Überdeckung von A, und nach Voraussetzung gibt es also ein $n \in \mathbb{N}$ und $a_0, \dots, a_n$, so dass

$$A \subset \bigcup_{k=0}^{n} U_{a_k}$$

Sei

$$\epsilon := \frac{1}{2} \min\{d(x, a_k) \,|\, 0 \leq k \leq n\}$$

Dann ist $\epsilon > 0$ und für alle $k \in \{0, \dots, n\}$

$$U_\epsilon(x) \cap U_{a_k} = \emptyset.$$

Daher ist $U_\epsilon(x) \subset X \setminus A$, also $X \setminus A$ offen und A abgeschlossen.

□

Satz 53. *Eine abgeschlossene Teilmenge einer kompakten Teilmenge ist kompakt.*

Beweis. Seien $A \subset X$ kompakt und $B \subset A$ eine abgeschlossene Teilmenge. Sei $(U_i)_{i\in I}$ eine offene Überdeckung von B. Wir suchen eine endliche Teilüberdeckung.

Durch Hinzunahme der offenen Menge $U := X \setminus B$ erhält man eine offene Überdeckung von A. Weil A kompakt ist, gibt es eine endliche Teilmenge $K \subset I$ mit

$$A \subset U \cup \bigcup_{k\in K} U_k.$$

Wegen $B \cap U = \emptyset$ ist dann aber

$$B \subset \bigcup_{k\in K} U_k,$$

und wir haben für B eine endliche Teilüberdeckung von $(U_i)_{i\in I}$ gefunden. □

Satz 54 (Heine-Borel). *Eine Teilmenge A des $\mathbb{R}^n$ mit der Standardmetrik ist kompakt genau dann, wenn sie abgeschlossen und beschränkt ist.*

Beweis. Trivial nach den Sätzen 50, 52 und 53. □

Dieser Satz ist falsch in allgemeinen metrischen Räumen.

Beispiel 55. Sei M eine unendliche Menge und $(X, d) = (\mathcal{B}(M, \mathbb{R}), d^{sup})$. Für $m \in M$ sei $f_m \in \mathcal{B}(M, \mathbb{R})$ definiert durch

$$f_m(n) := \begin{cases} 1 & \text{für } m = n, \\ 0 & \text{sonst.} \end{cases}$$

Dann gilt

$$d^{sup}(f_m, f_n) = \begin{cases} 0 & \text{für } m = n, \\ 1 & \text{sonst.} \end{cases} \tag{6}$$

Die Menge

$$A := \left\{ f_m \;\middle|\; m \in M \right\}$$

hat daher den Durchmesser $\operatorname{diam}(M) = 1$ und ist beschränkt.

Betrachtet man A als Teilmenge von A mit der durch d^{sup} induzierten Metrik, so ist A trivialerweise auch abgeschlossen. Nach (6) ist $U_{\frac{1}{2}}(f_m) \cap A = \{f_m\}$, und die offene Überdeckung $\left(U_{\frac{1}{2}}(f_m)\right)_{m\in M}$ von A besitzt deshalb keine endliche Teilübderdeckung. Also ist A als Teilmenge von (A, d_A^{sup}) nicht kompakt, wohl aber abgeschlossen und beschränkt.

Derselbe Beweis klappt für jede unendliche Menge mit der diskreten Metrik.

Nach Satz 49 ist A auch als Teilmenge von $(\mathcal{B}(M, \mathbb{R}), d^{sup})$ nicht kompakt. Es ist aber beschränkt, s. oben, und auch abgeschlossen, wie wir mittels Satz 32 noch zeigen wollen. Eine konvergente Folge in A ist eine Cauchyfolge. Nach (6) und dem Cauchykriterium mit $\epsilon < 1$ sind dann aber fast alle Folgenglieder gleich, also gleich dem Limes, der damit ebenfalls in A liegt.

□

Satz 56 (Bolzano-Weierstraß). *Sei (X, d) ein metrischer Raum, $A \subset X$ kompakt und $(x_n)_{n\in\mathbb{N}}$ eine Folge in A. Dann besitzt (x_n) eine konvergente Teilfolge.*
Bemerkung: Weil A abgeschlossen ist, liegt der Limes dann in A.

Umgekehrt ist eine Teilmenge A eines metrischen Raumes kompakt, wenn in ihr jede Folge eine konvergente Teilfolge besitzt. Der Beweis ist etwas trickreich (Vgl. zum Beispiel *Klaus Jänich, Topologie, Springer Hochschultext, 2. Aufl. p.97*). Wir verzichten darauf.

Beweis. Falls es ein $a \in A$ gibt, so dass $\#\left\{k \;\middle|\; x_k \in U\right\} = \infty$ für jede offene Umgebung U von a, so ist dieses a Grenzwert einer konvergenten Teilfolge: Wähle nämlich $n_0 \in \mathbb{N}$ beliebig und zu jedem $k \in \mathbb{N} \setminus \{0\}$ ein $n_k \in \mathbb{N}$ für das

$$x_{n_k} \in U_{\frac{1}{k}}(a) \text{ und}$$
$$n_k > n_{k-1}.$$

Unter den gemachten Voraussetzungen ist das möglich und liefert die gesuchte Teilfolge.

Gibt es kein solches a, so besitzt andererseits jedes a eine offene Umgebung U_a für die

$$\#\left\{k \;\middle|\; x_k \in U_a\right\} < \infty.$$

Die Familie $(U_a)_{a\in A}$ ist eine offene Überdeckung des kompakten A, also gibt es ein $n \in \mathbb{N}$ und $a_0, \ldots, a_n \in A$, so dass

$$A \subset U_{a_0} \cup \ldots \cup U_{a_n}.$$

Dann ist aber $\#\left\{k \;\middle|\; x_k \in A\right\} < \infty$ im Widerspruch dazu, dass die Folge $(x_n)_{n\in\mathbb{N}}$ eine unendliche Folge ist. Dieser Fall kann also nicht auftreten. □

Korollar 57. *Jeder kompakte metrische Raum ist vollständig.*

Beweis. Ist (X, d) kompakt und $(x_k)_k$ eine Cauchyfolge in X, so hat diese nach Bolzano-Weierstraß eine gegen $x^* \in X$ konvergente Teilfolge $(x_{k_j})_{j\in\mathbb{N}}$. Beweisen Sie mit der Dreiecksungleichung, dass dann die ganze Folge $(x_k)_k$ gegen x^* konvergiert. □

Korollar 58 (Lebesguesche Zahl). *Seien (X, d) ein metrischer Raum, $A \subset X$ kompakt und $(U_i)_{i\in I}$ eine offene Überdeckung von A. Dann gibt es eine positive Zahl δ, so dass gilt:*

$$\textit{Für alle } a \in A \textit{ gibt es ein } i \in I \textit{ mit } U_\delta(a) \subset U_i.$$

Jedes solche δ nennt man eine Lebesguesche Zahl der Überdeckung $(U_i)_{i\in I}$ von A.

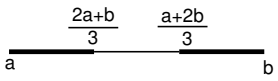
Beweis. Andernfalls gibt es zu jedem $\delta > 0$ ein $a \in A$, so dass $U_\delta(a)$ in keinem einzelnen der U_i enthalten ist. Wir wählen eine Nullfolge $(\delta_k)_{k\in\mathbb{N}}$ und zu jedem δ_k ein solches $a_k \in A$. Weil A kompakt ist, gibt es nach dem Satz von Bolzano-Weierstraß eine konvergente Teilfolge der (a_k) mit Limes $a^* \in A$. Wir können o.E. annehmen, dass $\lim_{k\to\infty} a_k = a^*$. Dann gibt es ein $i_0 \in I$ mit $a^* \in U_{i_0}$.

Weil U_{i_0} offen ist, gibt es ein $\epsilon > 0$ mit $U_\epsilon(a^*) \subset U_{i_0}$. Wähle $k \in \mathbb{N}$ so groß, dass $\delta_k < \frac{\epsilon}{2}$ und $d(a_k, a^*) < \frac{\epsilon}{2}$. Dann ist nach der Dreiecksungleichung $U_{\delta_k}(a_k) \subset U_{i_0}$ im Widerspruch zur Wahl von a_k. □

Beispiel 59 (Cantorsches Diskontinuum). Für ein kompaktes Intervall $[a, b]$ definieren wir

$$c([a, b]) := [a, \frac{2a + b}{3}] \cup [\frac{a + 2b}{3}, b].$$

Die Menge $c([a, b])$ erhält man also aus $[a, b]$, indem man das mittlere Drittel aus dem Intervall herausnimmt.



Sie besteht aus zwei kompakten Intervallen der Länge $\frac{b-a}{3}$. Für die Vereinigung endlich vieler *disjunkter* kompakter Intervalle $I_1 \cup \ldots \cup I_n$ definieren wir

$$c(I_1 \cup \ldots \cup I_n) = c(I_1) \cup \ldots \cup c(I_n).$$

Dann ist also $c(I_1 \cup \ldots \cup I_n)$ wieder die Vereinigung disjunkter kompakter Intervalle. Ist $L := \max_{1 \leq k \leq n}($ Länge von $I_k)$, so ist das längste Teilintervall aus $c(I_1 \cup \ldots \cup I_n)$ von der Länge $L/3$. Wir beginnen nun mit $C_0 := [0, 1]$ und definieren induktiv

$$C_{k+1} := c(C_k).$$

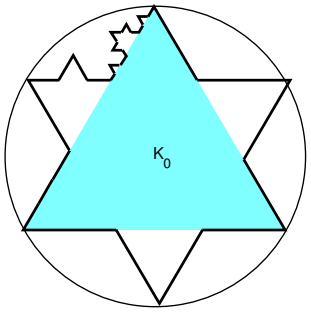
Die Menge $C := \bigcap_k C_k$ heißt das *Cantorsche Diskontinuum*. Als Durchschnitt abgeschlossener Mengen ist C abgeschlossen und offenbar beschränkt, also kompakt und offenbar nicht leer $(0 \in C)$.

Sind x und y zwei Punkte von C mit Abstand $d > 0$, so liegen sie in jedem C_k, insbesondere in einem C_k, dessen Intervall alle kürzer als d sind. Zwischen je zwei verschiedenen Punkten $x, y \in C$ gibt es also einen Punkt aus $\mathbb{R} \setminus C$. Daher rührt der Name *Diskontinuum*.

□

Beispiel 60 (von Kochsche Kurve).

Wir beginnen mit einem gleichseitigen Dreieck K_0, das in der Figur getönt ist. Auf das mittlere Drittel jeder Seite setzen wir ein (gefülltes) gleichseitiges Dreieck und erhalten einen „Stern“ K_1. Offenbar liegt K_1 in der abgeschlossenen Umkreisscheibe U von K_0.



Auf das mittlere Drittel jeder Seitenkante von K_1 setzen wir wieder ein gleichseitiges Dreieck und erhalten K_2. Weil jede Seitenkante auch Seitenkante eines gleichseitigen Dreiecks in U ist, liegen die angesetzten Dreiecke in U. Fortsetzung des Verfahrens liefert eine Folge $(K_j)_{j \in \mathbb{N}}$ von Teilmengen von U. In der obigen Abbildung ist diese Konstruktion nur lokal durchgeführt. Wir setzen

$$K = \bigcup_{j \in \mathbb{N}} K_j.$$

Den Rand ∂K nennt man die *von Kochsche Kurve* oder die *von Kochsche Schneeflocke*. Als Rand einer Menge ist sie abgeschlossen, und weil sie in der kompakten Menge U liegt, ist sie kompakt.

Weil die Länge von ∂K_{k+1} das $\frac{4}{3}$-fache der Länge von ∂K_k ist, ist es plausibel, der von Kochschen Kurve eine unendliche Länge zuzusprechen. Die von ihr eingeschlossene Fläche K ist hingegen offensichtlich von endlichem Flächeninhalt. Allerdings fehlen uns einstweilen exakte Definitionen für *Länge* (von was?) und *Flächeninhalt*.

□

Beispiel 61 (Hausdorffmetrik). Sei $\mathcal{F}X$ die Menge der nicht-leeren kompakten Teilmengen eines vollständigen (kompakten) metrischen Raumes (X, d). Definiere für $A, B \in \mathcal{F}X$

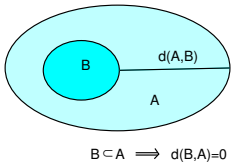
(i) $d(x, B) := \inf\{d(x, y) \mid y \in B\}$ für $x \in A$,

(ii) $d(A, B) := \sup\{d(x, B) \mid x \in A\}$, (wohldefiniert, weil $d(., B)$ stetig auf der kompakten Menge A),

(iii) $h(A, B) := \sup\{d(A, B), d(B, A)\}$.

Dann ist $(\mathcal{F}X, h)$ ein vollständiger (kompakter) metrischer Raum, der *Raum der Fraktale*. h heißt die *Hausdorffmetrik*. Vgl. [M. Barnsley, Fractals everywhere, Academic Press 1988].

Die nebenstehende Abbildung zeigt, dass $d(A, B)$ nicht symmetrisch ist.



Nachweis der Metrik-Eigenschaften.

Die Symmetrie ist klar.

Es gilt

$$h(A, B) = 0 \iff d(A, B) = 0 \text{ und } d(B, A) = 0.$$

Weiter ist

$$d(A, B) = 0 \iff d(x, B) = 0 \text{ für alle } x \in A \iff A \subset B.$$

Also $h(A, B) = 0$ genau dann, wenn $A = B$.

Dreiecksungleichung. Seien $a \in A, b \in B, c \in C$.

$$\begin{aligned} d(a, c) \le d(a, b) + d(b, c) &\implies d(a, C) \le d(a, b) + d(b, C) \\ &\implies d(a, C) \le d(a, b) + d(B, C) \\ &\implies d(a, C) \le d(a, B) + d(B, C) \\ &\implies d(A, C) \le d(A, B) + d(B, C) \\ &\implies d(A, C) \le h(A, B) + h(B, C) \end{aligned}$$

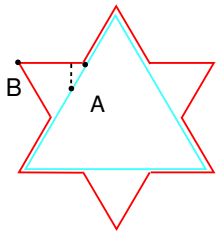
Die rechte Seite ist in A und C symmetrisch, und deshalb folgt auch

$$d(C, A) \le h(A, B) + h(B, C),$$

also $h(A, C) \le h(A, B) + h(B, C)$.

□

Beispiel 62 (Noch einmal die von Kochsche Kurve). Seien $A = \partial K_0$ und $B = \partial K_1$ die beiden ersten Randkurven der von Kochschen Konstruktion:



Sei a die Seitenlänge des Dreiecks A. Dann ist (vgl. Abbildung) bezüglich der Standardmetrik des $\mathbb{R}^2$

$$\begin{aligned} d(A,B) &< \frac{a}{6}, \\ d(B,A) &= \frac{a}{6}\sqrt{3}, \\ h(A,B) &= \frac{a}{6}\sqrt{3}. \end{aligned}$$

Allgemeiner ist für die Randfolge (∂K_j) der von Kochschen Konstruktion $h(\partial K_k, \partial K_{k+1}) = (\sqrt{3}/6)^{k+1}a$. Also ist $(\partial K_k)_{k\in\mathbb{N}}$ eine Cauchyfolge in $\mathcal{F}\mathbb{R}^2$ und man kann zeigen, dass die von Kochsche Kurve ihr Grenzwert ist.

□

Bemerkung. In *Dugundji, Topology* findet man für kompaktes X dazu noch folgende Übungsaufgaben:

(i) Für beliebiges $E \subset X$ ist

$$\{A \in \mathcal{F}X \mid E \subset A\}$$

abgeschlossen.

(ii) Setze für beliebiges $E \subset X$

$$\begin{aligned} I(E) &= \{A \in \mathcal{F}X \mid A \subset E\} \\ J(E) &= \{A \in \mathcal{F}X \mid A \cap E \neq \emptyset\}. \end{aligned}$$

Dann sind $I(E)$ und $J(E)$ mit E offen bzw. abgeschlossen.

(iii) Die Abbildung $\mathcal{F}X \to \mathbb{R}, A \mapsto \operatorname{diam}(A)$ ist stetig.

(iv) Die Abbildung $\mathcal{F}X \times \mathcal{F}X \to \mathbb{R}, (A,B) \mapsto d(A,B)$ ist stetig.

(v) $\mathcal{F}X$ ist kompakt.

1.4 Zusammenhang

- Die Rolle der Intervalle in $\mathbb{R}$ wird in metrischen Räumen übernommen von den sogenannten *zusammenhängenden Mengen*, die wir jetzt kennenlernen.

Jede Menge X mit mindestens zwei Elementen läßt sich trivialerweise schreiben als Vereinigung zweier nicht-leerer disjunkter Teilmengen. Aber nicht jeder metrische Raum läßt sich schreiben als Vereinigung zweier nicht-leerer disjunkter *offener* Mengen.

Definition 63 (Zusammenhang). Sei (X, d) ein metrischer Raum.

(i) X heißt *zusammenhängend*, wenn es nicht die Vereinigung zweier nicht-leerer disjunkter offener Mengen ist:

Für alle offenen $U, V \subset X$ mit

$$U \cap V = \emptyset \quad \text{und} \quad U \cup V = X$$

gilt

$$U = \emptyset \quad \text{oder} \quad V = \emptyset.$$

Das ist äquivalent zur Forderung, dass $\emptyset$ und X die einzigen zugleich offen und abgeschlossen Teilmengen sind.

(ii) Eine Teilmenge $A \subset X$ heißt *zusammenhängend*, wenn (A, d_A) zusammenhängend ist.

(iii) X heißt *wegzusammenhängend*, falls es zu je zwei Punkten $p, q \in X$ einen Weg von p nach q, d.h. eine stetige Abbildung $c : [0, 1] \to X$ mit $c(0) = p$ und $c(1) = q$ gibt.[3]

Satz 64 (Zusammenhängende Teilmengen). *Seien (X, d) ein metrischer Raum und $A \subset X$. Dann ist A genau dann zusammenhängend, wenn gilt:*
Für alle offenen $U, V \subset X$ mit

$$U \cap V = \emptyset \quad \text{und} \quad U \cup V \supset A$$

gilt

$$U \cap A = \emptyset \quad \text{oder} \quad V \cap A = \emptyset.$$

Beweis. *Zu ($\Longrightarrow$).* Sei (A, d_A) zusammenhängend und seien $U, V \subset X$ wie im Satz. Dann sind $U \cap A$ und $V \cap A$ offene Teilmengen von (A, d_A) mit leerem Durchschnitt, deren Vereinigung A ist. Also ist $U \cap A = \emptyset$ oder $V \cap A = \emptyset$.

Zu ($\Longleftarrow$). Wir wollen zeigen, dass (A, d_A) zusammenhängend ist. Seien also $U, V \subset A$ offen in A mit $U \cap V = \emptyset$ und $U \cup V = A$. Dann gibt es offene Mengen $\tilde{U}, \tilde{V}$ von X mit

$$\tilde{U} \cap A = U, \quad \tilde{V} \cap A = V.$$

Aber um die Vorausetzungen anwenden zu können, müssen $\tilde{U}$ und $\tilde{V}$ disjunkt sein. Deshalb müssen wir die Erweiterungen $\tilde{U}$ und $\tilde{V}$ von U und V sorgfältig konstruieren.

Wir wählen zu jedem $x \in U$ ein $\epsilon(x) > 0$ mit $U_{\epsilon(x)}(x) \cap A \subset U$. Das geht, weil U offen ist in A. Wir definieren

$$\tilde{U} := \bigcup_{x \in U} U_{\frac{1}{2}\epsilon(x)}(x).$$

[3] Allerdings haben wir noch gar nicht definiert, was stetige Abbildungen sind. Das ist also eine Definition „auf Vorrat“.

Entsprechend definieren wir $\tilde{V}$. Natürlich sind das offene Teilmengen von (X, d), und sie sind auch disjunkt: Wäre $z \in \tilde{U} \cap \tilde{V}$, so gäbe es $x \in U$ und $y \in V$ mit

$$z \in U_{\frac{1}{2}\epsilon(x)}(x) \cap U_{\frac{1}{2}\epsilon(y)}(y).$$

Sei etwa $\epsilon(y) \leq \epsilon(x)$. Dann ist aber

$$d(x, y) \leq d(x, z) + d(z, y) \leq \frac{1}{2}\epsilon(x) + \frac{1}{2}\epsilon(y) \leq \epsilon(x)$$

Dann wäre aber $y \in U$ im Widerspruch zur Voraussetzung. Also ist $\tilde{U} \cap \tilde{V} = \emptyset$.

Jetzt können wir die Voraussetzungen auf $\tilde{U}$ und $\tilde{V}$ anwenden und erhalten $\tilde{U} \cap A = U = \emptyset$ oder $\tilde{V} \cap A = V = \emptyset$. □

Zeigen Sie (später) entsprechend: Für $A \subset X$ ist (A, d_A) wegzusammenhängend genau dann, wenn es zu allen $p, q \in A$ eine stetige Abbildung $c : [0, 1] \to X$ mit $c(0) = p$ und $c(1) = q$ mit $c([a, b]) \subset A$ gibt.

Bemerkung zur Verallgemeinerung auf topologische Räume. Die obige Definition des Begriffs *zusammenhängend* benutzt nur offene Mengen, nicht explizit die Metrik. Daher läßt sich die Definition ohne Modifikation auf topologische Räume erweitern. Die schwierige Richtung von Satz 64 gilt allerdings nicht für topologische Räume. Ein Gegenbeispiel findet man so: Man nimmt eine Menge X, die wenigstens drei verschiedene Punkte O, P, Q enthält. Man definiert

$$\mathcal{T} = \{\emptyset\} \cup \{Y \subset X \mid O \in Y\}.$$

als System der offenen Menge. $\mathcal{T}$ ist abgeschlossen gegenüber Durchschnitt und Vereinigung, definiert also wirklich eine Topologie auf X. Die Teilmenge $A = \{P, Q\}$ enthält die in A offenen disjunkten Teilmengen $U = \{P\}$ und $V = \{Q\}$. Aber diese lassen sich nicht zu *disjunkten* in X offenen Teilmengen erweitern, weil jede solche den Punkt O enthält.

Satz 65. *Sei $A \subset B \subset \bar{A} \subset X$ und sei A zusammenhängend. Dann ist auch B zusammenhängend. Insbesondere ist auch $\bar{A}$ zusammenhängend.*

Beweis. Seien $U, V \subset X$ offen und disjunkt mit $B \subset U \cup V$. Wir müssen zeigen, dass $U \cap B = \emptyset$ oder $V \cap B = \emptyset$. Weil A zusammenhängend ist, gilt $U \cap A = \emptyset$ oder $V \cap A = \emptyset$. Sei etwa $U \cap A = \emptyset$. Wäre $U \cap B \neq \emptyset$, so gäbe es also ein $b \in U \cap B$ und, weil U offen ist, dazu ein $\epsilon > 0$ mit $U_\epsilon(b) \subset U$. Natürlich ist $b \in \bar{A}$. Also liegen in $U_\epsilon(b)$ auch Punkte von A. Die liegen dann aber in U. Widerspruch zur Annahme $U \cap A = \emptyset$! □

Satz 66. *Die zusammenhängenden Teilmengen von $\mathbb{R}$ sind genau die Intervalle.*

Beweis. Sei J ein Intervall. Seien $U, V \subset \mathbb{R}$ offen mit $U \cap V = \emptyset$ und $J \subset U \cup V$.
Annahme: $p \in U \cap J$, $r \in V \cap J$ und o.E. $p < r$. Wir müssen zeigen, dass dies zum Widerspruch führt. Sei

$$q := \sup\{t \mid [p, t] \subset U\}.$$

Dann gilt nach Voraussetzung $q \leq r < \infty$. Offenbar ist $q \notin U$, denn andernfalls wäre wegen der Offenheit von U auch $[p, q + \epsilon] \subset U$ für kleines $\epsilon > 0$ im Widerspruch zur Wahl von q.

Andrerseits ist $q \notin V$, weil sonst wegen der Offenheit von V auch $q - \epsilon \in V$ für kleines $\epsilon > 0$ im Widerspruch zur Wahl von q.

Damit ist $q \notin U \cup V$. Widerspruch zu $q \in J \subset U \cup V$.

Sei $J \subset \mathbb{R}$ zusammenhängend. Seien $p < q < r$ mit $p, r \in J$. Wäre $q \notin J$, so wäre

$$J \subset]-\infty, q[\cup]q, \infty[,$$

also $J \subset]-\infty, q[$ oder $J \subset]q, \infty[$ im Widerspruch dazu, dass p in der einen, q in der anderen dieser Mengen liegt. $\square$

1.5 Stetige Abbildungen

- Nachdem der Konvergenzbegriff in metrischen Räumen erklärt ist, ist es leicht, auch die Stetigkeit von Abbildungen solcher Räume zu erklären.
- Wir machen uns mit der Bedeutung dieses Begriffes in verschiedenen einfachen Situationen vertraut und formulieren Rechenregeln für stetige Abbildungen.
- Etwas abstrakter ist die Charakterisierung der Stetigkeit mittels offener oder abgeschlossener Mengen.

Definition 67 (Stetigkeit). Seien (X, d_X) und (Y, d_Y) metrische Räume, $G \subset X$ eine Teilmenge und $f : X \supset G \to Y$ eine Abbildung.

(i) f heißt *stetig in* $p \in G$, wenn $\lim_{x\to p} f(x) = f(p)$ ist, d.h. wenn für jede gegen p konvergente Folge $(x_k)_{k\in\mathbb{N}}$ in G auch $\lim_{k\to\infty} f(x_k) = f(p)$ ist.

(ii) f heißt *stetig in* oder *auf* G, wenn es stetig in jedem Punkt $p \in G$ ist.

(iii) Offenbar ist $f : X \supset G \to Y$ im Sinne dieser Definition stetig (in $p \in G$), genau dann, wenn es (in p) stetig ist als Abbildung des metrischen Raumes (G, d_G) nach Y.

Beispiel 68 (Komponentenweise Stetigkeit). Ist (X, d_X) beliebig und $(Y, d_Y) = \mathbb{R}^m$, so ist $f = (f_1, \ldots, f_m)$ genau dann stetig in p, wenn alle Komponentenfunktionen $f_i : X \to \mathbb{R}$ in p stetig sind. Das folgt unmittelbar aus der Definition und Satz 31.

□

Partielle Stetigkeit. Bei einer Abbildung $f : X \to \mathbb{R}^m$ kann man also die Stetigkeit einfach an den (reellwertigen) Komponentenfunktionen untersuchen. $\mathbb{R}^m$ oder $\mathbb{R}$ auf der rechten Seite macht also „keinen großen Unterschied“. Jetzt betrachten wir umgekehrt eine Funktion $f : \mathbb{R}^n \supset G \to Y$. Dann können wir $f(x) = f(x_1, \ldots, x_n)$ als Funktion jeder einzelnen Variablen betrachten, indem wir uns vorstellen, dass die anderen festbleiben. Es stellt sich die naheliegende Frage, ob f in p stetig ist, wenn alle die Funktionen

$$\begin{aligned} x_1 &\mapsto f(x_1, p_2, p_3, \ldots, p_n) \\ x_2 &\mapsto f(p_1, x_2, p_3, \ldots, p_n) \\ x_3 &\mapsto f(p_1, p_2, x_3, \ldots, p_n) \\ &\ldots \\ x_n &\mapsto f(p_1, p_2, p_3, \ldots, x_n) \end{aligned}$$

stetig sind. Man nennt das partielle Stetigkeit, weil man immer nur einen Teil der Variablen - nämlich eine - als variabel betrachtet. Folgt aus partieller Stetigkeit die Stetigkeit? Das ist nicht so:

Partielle Stetigkeit impliziert NICHT Stetigkeit.

Beispiel 69. Sei $f : \mathbb{R}^2 \to \mathbb{R}$ gegeben durch $f(0, 0) := 0$ und

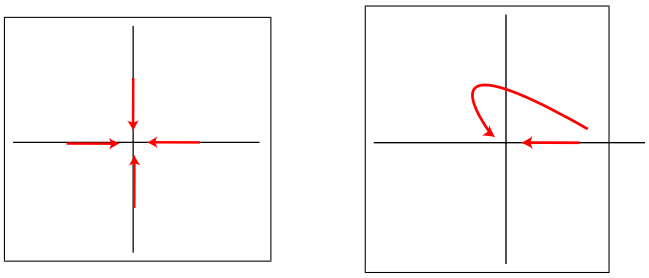
$$f(x, y) := \frac{xy}{x^2 + y^2} \text{ für } (x, y) \neq (0, 0).$$

Für $\lambda \in \mathbb{R}$ geht nämlich die Folge $(\frac{1}{k}, \frac{\lambda}{k})$ gegen $(0, 0)$, aber es ist

$$f(\frac{1}{k}, \frac{\lambda}{k}) = \frac{\lambda}{k^2(\frac{1}{k^2} + \frac{\lambda^2}{k^2})} = \frac{\lambda}{1 + \lambda^2}.$$

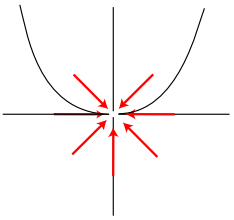
Für $\lambda \neq 0$ und $k \to \infty$ geht das also nicht gegen $0 = f(0,0)$. Andererseits ist f in $(0,0)$ wegen $f(x,0) = 0 = f(0,y)$ aber partiell stetig.

Dieses Beispiel zeigt genauer, warum partielle Stetigkeit viel schwächer ist als „totale“ Stetigkeit: Die Variable x muß sich der Stelle p auf beliebige Weise nähern dürfen. Bei der partiellen Stetigkeit schränkt man sich aber auf achsenparallele Annäherung ein.



In unserem Beispiel ist die Funktion auf allen Geraden durch den Nullpunkt jeweils konstant (Wert $\lambda/(1+\lambda^2)$), nur im Nullpunkt hat sie definitionsgemäß den Wert 0. Der kommt heraus, wenn man auf der x-Achse ($\lambda = 0$) oder auf der y-Achse ($\lambda = \infty$) an den Nullpunkt heranläuft, aber eben nur dann.

Selbst wenn $f(x) \to f(p)$ bei Annäherung auf allen Geraden durch p gilt, folgt daraus nicht die Stetigkeit in p. Ein Gegenbeispiel liefert die Funktion g mit $g(x,y) = 1$, falls $y = x^2 \neq 0$, und $g(x,y) = 0$ sonst. Wie sieht der Graph dieser Funktion aus?



□

Beispiel 70. Die Abbildungen

$$\begin{aligned}
&\alpha : \mathbb{R}^2 \to \mathbb{R}, (x_1, x_2) \mapsto x_1 + x_2 \\
&\mu : \mathbb{R}^2 \to \mathbb{R}, (x_1, x_2) \mapsto x_1 x_2 \\
&\eta : \mathbb{R}^2 \supset \{(x_1, x_2) \,|\, x_2 \neq 0\} \to \mathbb{R}, (x_1, x_2) \mapsto \frac{x_1}{x_2}
\end{aligned}$$

sind stetig.

Wir zeigen das für α. Seien $p = (p_1, p_2) \in \mathbb{R}^2$ und $(x_k = (x_{k1}, x_{k2}))_{k\in\mathbb{N}}$ eine Folge mit $\lim x_k = p$. Dann ist

$$d(\alpha(x_k), \alpha(p)) = |(x_{k1} + x_{k2}) - (p_1 + p_2)| \leq |x_{k1} - p_1| + |x_{k2} - p_2|.$$

Aber nach Satz 31 folgt aus $\lim x_k = p$, dass $\lim x_{ki} = p_i$ für $i = 1, 2$. Daher geht die rechte Seite gegen Null und $\lim \alpha(x_k) = \alpha(p)$.

Beweis für μ und η selbst. Für den letzteren Fall benutzt man die Abschätzung

$$\left|\frac{x_1}{x_2} - \frac{p_1}{p_2}\right| = \left|\frac{x_1 p_2 - x_2 p_1}{x_2 p_2}\right| \leq \frac{|x_1 - p_1||p_2| + |p_1||p_2 - x_2|}{|x_2 p_2|}.$$

□

Beispiel 71. Dieselben Argumente wie im vorstehenden Beispiel zeigen die Stetigkeit der Determinante

$$\det : M(n,\mathbb{R}) \to \mathbb{R}, \quad (x_{ij}) \mapsto \sum_{\sigma} \operatorname{sign}\sigma x_{1\sigma(1)} \dots x_{n\sigma(n)},$$

wenn man den Raum $M(n\mathbb{R})$ der quadratischen n-reihigen Matrizen auf die offensichtliche Weise mit dem $\mathbb{R}^{n^2}$ identifiziert und eine der l^p-Metriken verwendet. Auch die mit der Transponierten $(x_{ij})^T$ gebildete Abbildung

$$M(n,\mathbb{R}) \to M(n,\mathbb{R}), \quad (x_{ij}) \mapsto (x_{ij})(x_{ij})^T$$

gebildete Abbildung ist stetig.

□

Beispiel 72. Seien (X,d) metrischer Raum und $a \in X$. Dann ist

$$d(.,a) : X \to \mathbb{R}, x \mapsto d(x,a)$$

stetig.

□

Beispiel 73. Sei $(X,d) = (C^0([a,b],\mathbb{R}), d^{sup})$. Dann ist die Abbildung

$$\int_a^b : X \to \mathbb{R}, f \mapsto \int_a^b f(x)dx$$

stetig, denn es gilt

$$\begin{aligned}\left|\int_a^b f(x)dx - \int_a^b g(x)dx\right| &= \left|\int_a^b (f(x)-g(x))dx\right| \\ &\le \int_a^b |f(x)-g(x)|dx \\ &\le \int_a^b \sup_{a\le x\le b} |f(x)-g(x)|dx \\ &= |b-a| d^{sup}(f,g).\end{aligned}$$

□

Satz 74. *Seien (X_i,d_i), $i=1,2,3$ metrische Räume und $f_i : X_i \supset G_i \to X_{i+1}$ für $i=1,2$ Abbildungen mit $f_1(G_1) \subset G_2$. Es sei f_1 stetig in $p_1 \in G_1$ und f_2 stetig in $p_2 := f_1(p_1)$. Dann ist $f_2 \circ f_1 : X_1 \supset G_1 \to X_3$ stetig in p_1.*
Kurz: Die Komposition stetiger Abbildungen ist stetig.

Beweis. Leicht. □

Korollar 75. *Ist (X,d) ein metrischer Raum, und sind $f,g : X \supset G \to \mathbb{R}$ stetig in $p \in G$, ist ferner $\lambda \in \mathbb{R}$, so sind auch die Abbildungen*

$$f+g, f-g, fg, \lambda f : X \supset G \to \mathbb{R}$$

in p stetig. Insbesondere ist also der Vektorraum

$$C^0(X,\mathbb{R}) := \{f : X \to \mathbb{R} \,|\, f \text{ stetig}\}$$

ein reeller Vektorraum.

Ist $g(p) \neq 0$, so ist $\frac{f}{g} : X \supset \left\{x \in G \;\middle|\; g(x) \neq 0\right\} \to \mathbb{R}$ in p stetig.

Beweis. Die Abbildungen sind vom Typ

$$x \mapsto (f(x), g(x)) \overset{\alpha}{\mapsto} f(x) + g(x),$$

wobei die erste stetige Komponentenfunktionen besitzt. □

Satz 76 (ϵ-δ-Kriterium für Stetigkeit). *Seien $(X, d_X), (Y, d_Y)$ metrische Räume und $f : X \supset G \to Y$. Sei $p \in G$. Dann ist f in p genau dann stetig, wenn es zu jedem $\epsilon > 0$ ein $\delta > 0$ gibt, so dass*

$$d_Y(f(p), f(x)) < \epsilon \text{ für alle } x \in G \text{ mit } d_X(p, x) < \delta.$$

Die letzte Bedingung ist äquivalent zu

$$f(U_\delta(p) \cap G) \subset U_\epsilon(f(p)).$$

Beweis. Zu ($\Longrightarrow$). Sei $\epsilon > 0$. Gäbe es kein δ wie angegeben, so gäbe es insbesondere zu jedem $k \in \mathbb{N}$ ein $x_k \in G$ mit

$$d_X(p, x_k) < \frac{1}{k+1}, \text{ aber } d_Y(f(p), f(x_k)) \geq \epsilon.$$

Dann wäre aber $\lim x_k = p$ und $\lim f(x_k) \neq f(p)$. Widerspruch!

Zu ($\Longleftarrow$). Sei $(x_k)_{k\in\mathbb{N}}$ eine gegen p konvergente Folge. Wir müssen zeigen, dass $\lim f(x_k) = f(p)$. Sei also $\epsilon > 0$ beliebig. Sei $\delta > 0$ dazu gewählt wie im Satz. Dann gibt es ein $k_0 \in \mathbb{N}$ mit $d_X(x_k, p) < \delta$ für alle $k \geq k_0$. Dann ist aber $d_Y(f(x_k), f(p)) < \epsilon$ für alle $k \geq k_0$. □

Satz 77. *(i) Seien $(X, d_X), (Y, d_Y)$ metrische Räume und $f : X \to Y$. Dann ist f stetig genau dann, wenn für jede offene Teilmenge $V \subset Y$ das Urbild $f^{-1}(V) \subset X$ offen ist. D.h. f ist genau dann stetig, wenn das Urbild aller offenen Mengen offen ist.*

(ii) Die Aussage bleibt richtig, wenn man überall „offen“ durch „abgeschlossen“ ersetzt.

(iii) Ist f nicht auf ganz X definiert, sondern nur

$$f : X \supset G \to Y$$

stetig, so sind die Urbilder offener Mengen offen in (G, d_G), aber nicht unbedingt in X.

Oft wird dieser Satz etwas großzügig zitiert als:

Stetige Urbilder(?) offener Mengen sind offen.

Finden Sie Beispiele, die zeigen, dass die Bilder offener Mengen unter stetigen Abbildungen im allgemeinen nicht offen sind.

Beweis. *Zu (i $\Longrightarrow$).* Seien also f stetig und $V \subset Y$ offen. Wir wollen zeigen, dass $f^{-1}(V)$ offen ist. Sei dazu $x \in f^{-1}(V)$. Zu $f(x) \in V$ gibt es dann ein $\epsilon > 0$ mit

$$U_\epsilon(f(x)) \subset V.$$

Dazu gibt es dann ein $\delta > 0$ mit

$$f(U_\delta(x)) \subset U_\epsilon(f(x)) \subset V.$$

Das bedeutet aber $U_\delta(x) \subset f^{-1}(V)$.

Zu (i $\Longleftarrow$). Sei das Urbild jeder offenen Menge offen und sei $p \in X$.

Sei weiter $\epsilon > 0$. Dann ist $U_\epsilon(f(p))$ offen, also ist $f^{-1}(U_\epsilon(f(p)))$ offen und damit eine Umgebung von p. Daher gibt es ein $\delta > 0$ mit $U_\delta(p) \subset f^{-1}(U_\epsilon(f(p)))$, d.h. mit

$$f(U_\delta(p)) \subset U_\epsilon(f(p)).$$

Das war aber zu zeigen.

Zu (ii). Die Aussage über abgeschlossene Mengen folgt, weil $A \subset Y$ genau dann abgeschlossen ist, wenn $Y \setminus A$ offen ist, und weil andrerseits

$$f^{-1}(Y \setminus A) = X \setminus f^{-1}(A).$$

Zu (iii). Die Offenheit in (G, d_G) ist klar nach (i). □

Beispiel 78. Die Einheitssphäre

$$S^n := \{x \in \mathbb{R}^{n+1} \mid \sum x_i^2 = 1\}$$

ist das Urbild von $\{1\} \in \mathbb{R}$ unter der stetigen Abbildung

$$x \mapsto \sum x_i^2.$$

Also ist S^n abgeschlossen und, weil beschränkt, auch kompakt.

□

Beispiel 79 (Matrizengruppen). Wir erinnern an Beispiel 71. Daraus folgt:

(i) Die *allgemeine linearer Gruppe* $\mathbf{GL}(n, \mathbb{R})$, gebildet aus den invertierbaren Matrizen, ist offen in $M(n, \mathbb{R})$:

$$\mathbf{GL}(n, \mathbb{R}) = \det^{-1}(\mathbb{R} \setminus \{0\}).$$

(ii) Die *spezielle lineare Gruppe* $\mathbf{SL}(n, \mathbb{R})$, gebildet aus allen Matrizen der Determinanten $= 1$, ist abgeschlossen in $M(n\mathbb{R})$:

$$\mathbf{SL}(n, \mathbb{R}) = \det^{-1}(\{1\}).$$

(iii) Die Gruppe der *orthogonalen Matrizen* $\mathbf{O}(n)$, gebildet aus allen Matrizen mit

$$(x_{ij})(x_{ij})^T = E(= \text{Einheitsmatrix})$$

ist abgeschlossen in $M(n, \mathbb{R})$. Weil alle ihre Komponenten x_{ij} maximal vom Betrag 1 sind, ist sie auch beschränkt und damit kompakt in $M(n, \mathbb{R}) = \mathbb{R}^{n^2}$.

Die angegebenen Teilmengen von $M(n, \mathbb{R})$ sind tatsächlich *Gruppen* bezüglich der Matrixmultiplikation. Sie haben außerdem eine von $M(n, \mathbb{R})$ geerbte Metrik, in der die Gruppenoperationen stetig sind. Damit sind sie (die wichtigsten) Beispiele sogenannter *Liegruppen.*

□

1.6 Fünf wichtige Sätze über stetige Abbildungen

- Die Sätze und Definitionen in diesem Abschnitt verallgemeinern Sätze und Definitionen, die wir (mit Ausnahme des letzten Satzes) aus der Analysis I für reelle Funktionen schon kennen.
- Die Beweise sind sehr einfach, weil die Definitionen bereits die wesentlichen Eigenschaften erfassen.

Seien $(X, d_X), (Y, d_Y)$ metrische Räume.

Satz 80 (vom Zusammenhang). *Seien $A \subset X$ zusammenhängend und $f : A \to Y$ stetig. Dann ist auch $f(A) \subset Y$ zusammenhängend.*

Bemerkung. Das verallgemeinert den Zwischenwertsatz!

Beweis. Seien $U, V \subset Y$ offen und disjunkt mit $f(A) \subset U \cup V$. Zu zeigen: $U \cap f(A) = \emptyset$ oder $V \cap f(A) = \emptyset$.

Nach Satz 77 sind $f^{-1}(U)$ und $f^{-1}(V)$ offen in (A, d_A). Sie sind weiter disjunkt mit

$$A = f^{-1}(U) \cup f^{-1}(V).$$

Weil A zusammenhängend ist, folgt

$$f^{-1}(U) = \emptyset \text{ oder } f^{-1}(V) = \emptyset,$$

und entsprechend ist $U \cap f(A) = \emptyset$ oder $V \cap f(A) = \emptyset$.

□

Satz 81 (Kompaktheitssatz). *Seien $A \subset X$ kompakt und $f : A \to Y$ stetig. Dann ist auch $f(A) \subset Y$ kompakt.*

Beweis. Sei $(U_i)_{i \in I}$ eine offene Überdeckung von $f(A)$. Dann ist $(f^{-1}(U_i))_{i \in I}$ eine offene Überdeckung von A. Also gibt es eine endliche Teilmenge $J \subset I$, so dass

$$A = \bigcup_{i \in J} f^{-1}(U_i).$$

Aber dann ist

$$f(A) \subset \bigcup_{i \in J} U_i.$$

□

Beispiel 82. Untersuchen Sie die Funktion

$$\frac{\ln x}{x} :]0, \infty[\to \mathbb{R}$$

um zu zeigen, dass stetige Abbildungen im allgemeinen weder beschränkte Mengen in beschränkte Mengen noch abgeschlossene Mengen in abgeschlossene Mengen abbilden.

□

Satz 83 (vom Maximum). *Seien $A \subset X$ kompakt $\neq \emptyset$ und $f : A \to \mathbb{R}$ stetig. Dann nimmt f sein Maximum und Minimum an.*

Beweis. Nach dem Kompaktheitssatz ist $f(A) \subset \mathbb{R}$ kompakt, also insbesondere beschränkt. Daher ist die Funktion $f : A \to \mathbb{R}$ beschränkt. Dann gibt es eine Folge $(x_k)_{k\in\mathbb{N}}$ in A mit $\lim f(x_k) = \sup f$. Nach dem Satz von Bolzano-Weierstraß hat (x_k) in dem metrischen Raum (A, d_A) eine konvergente Teilfolge $(x_{k_j})_{j\in\mathbb{N}}$. Ist $x^* := \lim x_{k_j} \in A$, so folgt

$$f(x^*) = \lim f(x_{k_j}) = \lim f(x_k) = \sup f.$$

Also wird das Supremum angenommen und ist das Maximum von f.

Analog für das Minimum. □

Definition 84. Sei $f : X \supset G \to Y$. f heißt *gleichmäßig stetig auf G*, wenn gilt

$$\forall_{\epsilon>0}\exists_{\delta>0}\forall_{x\in G} f(U_\delta(x) \cap G) \subset U_\epsilon(f(x)).$$

f ist genau dann gleichmäßig stetig auf der Teilmenge $G \subset X$ in diesem Sinne, wenn es auf (G, d_G) gleichmäßig stetig ist.

Offenbar impliziert gleichmäßige Stetigkeit die gewöhnliche Stetigkeit. Die Umkehrung ist nicht richtig:

Beispiel 85. Die Funktion $f = x^2$ ist auf $[0, 2]$ gleichmäßig stetig. Auf $\mathbb{R}$ ist sie stetig, aber nicht gleichmäßig stetig.

□

Satz 86 (von der gleichmäßigen Stetigkeit). *Seien $f : X \supset A \to Y$ stetig und A kompakt. Dann ist f gleichmäßig stetig.*

Beweis. Sei $\epsilon > 0$. Dann gibt es zu jedem $x \in A$ ein $\delta_x > 0$ mit

$$f(U_{\delta_x}(x) \cap A) \subset U_{\frac{1}{2}\epsilon}(f(x)).$$

Wir setzen

$$U_x := U_{\delta_x}(x).$$

Dann ist $(U_x)_{x\in A}$ eine offene Überdeckung von A. Sei $\delta > 0$ eine Lebesgue-Zahl dazu. Dann gilt für alle $x \in A$: Es gibt ein $y \in A$ mit $U_\delta(x) \subset U_y$. Also ist

$$f(U_\delta(x) \cap A) \subset f(U_y \cap A) \subset U_{\frac{1}{2}\epsilon}(f(y))$$

Insbesondere ist dann $f(x) \in U_{\frac{1}{2}\epsilon}(f(y))$ und deshalb

$$f(U_\delta(x) \cap A) \subset U_{\frac{1}{2}\epsilon}(f(y)) \subset U_\epsilon(f(x)).$$

□

Beispiel 87. Sei $f : [a, b] \to \mathbb{R}$ stetig auf dem kompakten Intervall. Dann ist f gleichmäßig stetig, und es gibt zu $\epsilon > 0$ ein $\delta > 0$ mit $|f(x) - f(y)| < \epsilon$ für alle $x, y \in [a, b]$ mit $|x - y| < \delta$. Wähle eine Zerlegung $a = x_0 < x_1 < \ldots < x_n = b$ mit $|x_i - x_{i-1}| < \delta$ und wähle

$\xi_i \in [x_{i-1}, x_i]$ beliebig. Setze dann $\phi(x) := f(\xi_i)$ für alle $x \in [x_{i-1}, x_i[$ und $\phi(b) = f(b)$. Dann ist ϕ also eine Treppenfunktion, und es gilt für alle $x \in [x_{i-1}, x_i[$

$$|f(x) - \phi(x)| = |f(x) - f(\xi_i)| < \epsilon$$

und

$$|f(b) - \phi(b)| = 0 < \epsilon.$$

Die Treppenfunktion approximiert die stetige Funktion f also besser als ϵ.

Auch andere Approximationsresultate für stetige Abbildungen benutzen die gleichmäßige Stetigkeit, vgl. Übungen und den Weierstraßschen Approximationssatz z.B. in *S. Hildebrandt, Analysis I.*

□

Der in der Definition 34 eingeführte Begriff der gleichmäßigen *Konvergenz* einer Funktionenfolge hat mit der gleichmäßigen *Stetigkeit* nichts zu tun. Im Raum der beschränkten Funktionen bedeutete gleichmäßige Konvergenz einfach die Konvergenz bezüglich der Supremumsmetrik, und in dem Zusammenhang haben wir einen Spezialfall des folgenden Satzes schon kennengelernt,vgl. Satz 37.

Satz 88 (von der gleichmäßigen Konvergenz). *Sei die Folge $(f_i : G \to Y)_{i\in\mathbb{N}}$ auf $G \subset X$ gleichmäßig konvergent gegen $f : G \to Y$. Sind alle f_i stetig, so ist auch f stetig.*

Beweis. Wir zeigen die Stetigkeit in $p \in G$. Sei $(x_j)_{j\in\mathbb{N}}$ eine Folge in G mit Grenzwert p. Zu zeigen:

$$\lim f(x_j) = f(p).$$

Sei dazu $\epsilon > 0$. Dann gibt es ein $k_0 \in \mathbb{N}$, so dass für alle $k \geq k_0$ *und alle* $x \in G$

$$d_Y(f_k(x), f(x)) < \epsilon/3.$$

Weil f_{k_0} in p stetig ist, gibt es ein j_0 mit

$$d_Y(f_{k_0}(x_j), f_{k_0}(p)) < \epsilon/3$$

für alle $j \geq j_0$. Dann ist aber für $j \geq j_0$

$$d_Y(f(x_j), f(p)) \leq d_Y(f(x_j), f_{k_0}(x_j)) + d_Y(f_{k_0}(x_j), f_{k_0}(p)) + d_Y(f_{k_0}(p), f(p)) < \epsilon.$$

□

Beispiel 89 (von der konstanten Majorante, Weierstraß). Sei $(f_k : G \to \mathbb{R})_{k\in\mathbb{N}}$ eine Folge von Funktionen. Sei $(c_k)_{k\in\mathbb{N}}$ eine Folge reeller Zahlen, so dass

$$|f_k(x)| \leq c_k \text{ für alle } x \text{ und } \sum_{i=0}^{\infty} c_i \text{ konvergent.}$$

Dann ist die Partialsummenfolge

$$s_n := \sum_{n=0}^{n} f_n$$

auf G gleichmäßig konvergent. Zum Beweis setzen wir

$$\gamma_n := \sum_{k=0}^{n} c_k, \quad \gamma := \sum_{k=0}^{\infty} c_k.$$

Zu $\epsilon > 0$ gibt es dann ein N mit $|\gamma - \gamma_m| < \epsilon$ für alle $m \geq N$. Dann ist aber für alle x und m, n mit $N < m < n$

$$|s_n(x) - s_m(x)| = \left| \sum_{k=m+1}^{n} f_k(x) \right| \leq \sum_{k=m+1}^{n} |f_k(x)|$$
$$\leq \sum_{k=m+1}^{n} c_k = \gamma_n - \gamma_m \leq \gamma - \gamma_m < \epsilon. \tag{7}$$

Also ist $(s_k(x))_{k\in\mathbb{N}}$ für jedes $x \in G$ eine Cauchyfolge in $\mathbb{R}$ und deshalb konvergent. Der Grenzwert sei $s(x) = \sum_{k=0}^{\infty} f_k(x)$. Bildet man in (7) den Grenzwert für $n \to \infty$, so folgt

$$|s(x) - s_m(x)| < \epsilon \text{ für alle } x \in G \text{ und } m > N,$$

also die gleichmäßige Konvergenz. Nach dem letzten Satz ist $s : G \to \mathbb{R}$ stetig, wenn alle f_k stetig sind.

Insbesondere kann man das anwenden auf Potenzreihen $\sum_{k=0}^{\infty} a_k(x - x_0)^k$. Ist $R > 0$ der Konvergenzradius dieser Reihe, und $0 < r < R$, so ist

$$\sum_{k=0}^{\infty} |a_k| r^k \text{ konvergent}$$

und

$$|a_k(x - x_0)^k| \leq |a_k| r^k \text{ für alle } x \in [x_0 - r, x_0 + r].$$

Also ist die Potenzreihe auf $[x_0 - r, x_0 + r]$ gleichmäßig konvergent und ihre Grenzfunktion auf jedem solchen Intervall stetig. Also ist sie auf $]x_0 - R, x_0 + R[$ stetig, aber das wussten wir schon. Man sagt auch, Potenzreihen seien gleichmäßig konvergent auf jedem Kompaktum im Inneren ihres Konvergenzbereichs.

□

1.7 Normierte Vektorräume

- Wir lernen normierte Vektorräume kennen, die in der mehrdimensionalen Analysis als Definitions- und Zielbereiche der Funktionen dienen.
- Endlich-dimensionale normierte Vektorräume sind insbesondere *vollständige* metrische Räume, und auf ihnen ist jede lineare Abbildung stetig.

Differentialrechnung beschäftigt sich mit der linearen Approximation von Funktionen. In einem allgemeinen metrischen Raum macht das keinen Sinn, weil man keine *lineare Struktur* hat. Zum Beispiel kann man die U-Bahnstationen aus Beispiel 6 nicht addieren. Den richtigen Rahmen für Linearität bieten *Vektorräume*. Und wenn man außerdem über Konvergenz reden will, braucht man in den Vektorräumen eine Metrik, die sich mit der linearen Struktur verträgt. Das führt zur Klasse der *normierten Vektorräume*, mit denen wir uns jetzt befassen wollen.

Ausblick: Man kann nicht nur in normierten Vektorräumen Differentialrechnung betreiben, sondern auch in Räumen, die sich selber durch lineare Räume approximieren lassen: wie Flächen durch ihre Tangentialräume. Das führt zur Analysis auf sogenannten *differenzierbaren Mannigfaltigkeiten*.

„Vektorraum“ heißt hier immer *reeller* Vektorraum.

Definition 90. Sei V ein Vektorraum. Eine *Norm* für V ist eine Abbildung

$$\|..\| : V \to \mathbb{R}, v \mapsto \|v\|,$$

so dass für alle $u, v \in V$ und $\lambda \in \mathbb{R}$ gilt:

(i) $\|v\| \geq 0$ und $(\|v\| = 0 \iff v = 0)$,

(ii) $\|\lambda v\| = |\lambda| \, \|v\|$,

(iii) $\|u + v\| \leq \|u\| + \|v\|$.

Ein *normierter Vektorraum* $(V, \|.\|)$ ist ein Vektorraum V zusammen mit einer Norm $\|.\|$ auf V. Durch die Definition

$$d(u, v) := \|u - v\| \tag{8}$$

wird daraus ein metrischer Raum, und alle Begriffe, die wir für metrische Räume erklärt haben, sind auch für normierte Vektorrräume erklärt. Wenn wir in normierten Vektorräumen von Konvergenz, Stetigkeit, offenen Mengen etc. sprechen, beziehen wir uns immer auf die Metrik (8). Normierte Vektorräume sind also spezielle metrische Räume.

Beispiel 91 (Der $\mathbb{R}^n$ mit der Standardnorm). Der Vektorraum $V = \mathbb{R}^n$ besitzt eine Norm $\|x\| := \sqrt{\sum x_i^2}$, die wir als die *Standardnorm* oder *Euklidische Norm* des $\mathbb{R}^n$ bezeichnen wollen. Die Axiome (i) und (ii) sind klar, die Dreiecksungleichung folgt, wenn wir beachten, dass

$$\|x - y\| = \sqrt{\sum_{k=1}^{n} (x_k - y_k)^2} = d(x, y)$$

die Standardmetrik aus Beispiel 3 ist. Damit folgt

$$\|x + y\| = d(x, -y) \leq d(x, 0) + d(-y, 0) = \|x - 0\| + \| - y - 0\| = \|x\| + \|y\|.$$

Wenn wir vom $\mathbb{R}^n$ als normiertem Vektorraum sprechen, beziehen wir uns immer auf die vorstehende Norm, obwohl es sehr viele andere gibt. Die Metriken d^p aus Beispiel 4 wie die Metrik d^∞ kommen alle von einer Norm

$$\|x\|_p := d^p(x, 0),$$

der sogenannten l^p-Norm.

□

Beispiel 92. Der Vektorraum der beschränkten Funktionen $V = \mathcal{B}(X, \mathbb{R})$ gestattet die Norm $\|f\| = \|f\|_{sup} = \sup\{|f(x)| \,|\, x \in X\}$, die zur Supremumsmetrik führt.

□

Beispiel 93. Nicht jede Metrik kommt von einer Norm, schon weil metrische Räume im allgemeinen eben keine Vektorräume sind: Beliebige Teilmengen von normierten Vektorräumen sind als Teilmengen von metrischen Räumen wieder metrische Räume, im allgemeinen aber keine normierten Vektorräume.

Aber auch auf „kompletten“ Vektorräumen gibt es Metriken, die nicht von einer Norm kommen. Zum Beispiel kommt die diskrete Metrik auf dem $\mathbb{R}^n$ nicht von einer Norm. Warum nicht?

□

Bemerkung. In der linearen Algebra haben Sie die Norm in einem Euklidischen Vektorraum kennengelernt. Jedes positiv definite Skalarprodukt $\langle ., . \rangle$ induziert eine Norm vermöge

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

Aber nicht jede Norm auf einem reellen Vektorraum kommt von einem Skalarprodukt. Notwendig und hinreichend ist die sogenannte *Parallelogrammgleichung*

$$2(\|x\|^2 + \|y\|^2) = \|x + y\|^2 + \|x - y\|^2.$$

Das die Bedingung notwendig ist, rechnen Sie leicht nach. Dass sie auch hinreichend ist, ist schwieriger zu zeigen. Man definiert

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

und muss dann vor allem die Bilinearität zeigen. Einen Beweis finden Sie zum Beispiel in *W. Klingenberg, Lineare Algebra und Analytische Geometrie, Springer 1984, p. 117.*

Lemma 94. *In einem normierten Vektorraum $(V, \|..\|)$ gilt für alle $u, v \in V$*

$$|\, \|u\| - \|v\| \,| \leq \|u - v\|.$$

Beweis. Es gilt nach der Dreiecksungleichung

$$\|u\| = \|(u - v) + v\| \leq \|u - v\| + \|v\|,$$

und daher

$$\|u\| - \|v\| \leq \|u - v\|.$$

Aus Symmetriegründen ist dann aber auch $\|v\| - \|u\| \leq \|u - v\|$, und daraus folgt die Behauptung. □

Als Folgerung ergibt sich, dass die Funktion

$$\|.\| : V \to \mathbb{R}, x \mapsto \|x\|$$

stetig ist.

Definition 95. Mit $L(V,W)$ bezeichnen wir den *Vektorraum der linearen Abbildungen*

$$F : V \to W$$

vom Vektorraum V in den Vektorraum W.

Satz 96. *Seien $(V, \|..\|_V)$ und $(W, \|..\|_W)$ normierte Vektorräume, und sei $F : V \to W$ linear. Dann ist F genau dann stetig, wenn es ein $C \in \mathbb{R}$ gibt, so dass für alle $v \in V$*

$$\|F(v)\|_W \leq C\|v\|_V.$$

Beweis. Zu ($\Longrightarrow$). Wenn F stetig ist, ist es insbesondere in 0 stetig. Also gibt es zu $\epsilon = 1$ ein $\delta > 0$ mit

$$F(U_\delta(0)) \subset U_\epsilon(F(0)) = U_1(0).$$

Mit anderen Worten:

$$\|v\|_V < \delta \implies \|F(v)\|_W < 1.$$

Dann gilt aber für alle $v \neq 0$

$$1 > \|F(\frac{\delta}{2\|v\|_V}\, v)\|_W = \frac{\delta}{2\|v\|_V}\|F(v)\|_W.$$

Das impliziert

$$\|F(v)\|_W \leq \frac{2}{\delta}\, \|v\|_V$$

auch für $v = 0$. Also können wir $C = \frac{2}{\delta}$ wählen.

Zu ($\Longleftarrow$). Gibt es ein C wie im Satz, so ist für alle $v_1, v_2 \in V$

$$\|F(v_1) - F(v_2)\|_W = \|F(v_1 - v_2)\|_W \leq C\|v_1 - v_2\|_V.$$

Daraus folgt die (gleichmäßige) Stetigkeit von F. □

Wir verzichten im weiteren auf den Index am Normsymbol.

Definition 97. Sie $n \in \mathbb{N}$. Ein (reeller) Vektorraum V heißt *n-dimensional*, wenn es einen Isomorphismus

$$\Phi : \mathbb{R}^n \to V$$

gibt. Dabei ist ein *Isomorphismus* eine *bijektive* Abbildung Φ, so dass Φ und Φ^{-1} linear sind. Ein Vektorraum heißt *endlich-dimensional*, wenn er n-dimensional für ein $n \in \mathbb{N}$ ist.

Satz 98. *Seien V, W normierte Vektorräume und V endlich-dimensional. Dann ist jede lineare Abbildung $F : V \to W$ stetig.*

Beweis. 1. Fall: $V = \mathbb{R}^n$. Zunächst gilt für $x \in \mathbb{R}^n$:

$$x = (x_1, \ldots, x_n) = \sum_{i=1}^{n} x_j e_j,$$

wobei $e_j = (0, \ldots, 0, 1, 0, \ldots, 0)$ den j-ten Vektor der sogenannten Standardbasis des $\mathbb{R}^n$ bezeichnet, der in der j-ten Komponente eine 1 und sonst lauter 0 hat. Daher ist

$$\|F(x)\| = \|F(\sum_{j=1}^{n} x_j e_j)\| = \|\sum_{j=1}^{n} x_j F(e_j)\| \leq \sum_{j=1}^{n} |x_j|\, \|F(e_j)\| \leq M \sum_{j=1}^{n} |x_j|$$

mit $M := \max_j \|F(e_j)\|$. Wegen $|x_j| \leq \sqrt{x_1^2 + \ldots + x_n^2}$ folgt $\sum_{j=1}^n |x_j| \leq n\sqrt{x_1^2 + \ldots + x_n^2}$, also

$$\|F(x)\| \leq Mn\|x\|,$$

und F ist stetig.

2. Fall: V beliebig, endlich-dimensional. Sei $\Phi : \mathbb{R}^n \to V$ ein Isomorphismus. Dann ist Φ nach dem 1. Fall stetig, und es gibt ein $C \in \mathbb{R}$ mit $\|\Phi(x)\| \leq C\|x\|$ für alle x. Wir zeigen, dass es auch ein $B > 0$ gibt, so dass

$$B\|x\| \leq \|\Phi(x)\| \leq C\|x\| \text{ für alle } x \in \mathbb{R}^n. \tag{9}$$

Die Funktion $\|\Phi\| : \mathbb{R}^n \to \mathbb{R}$ ist als Komposition stetiger Funktionen stetig und nimmt deshalb auf der kompakten Menge

$$S^{n-1} := \{x \in \mathbb{R}^n \mid \|x\| = 1\}$$

ihr Minimum B in einem Punkt $x^* \in S^{n-1}$ an. Weil $x^* \neq 0$ und Φ ein Isomorphismus ist, ist

$$B := \|\Phi(x^*)\| > 0.$$

Für alle $x \neq 0$ gilt

$$\|\Phi(x)\| = \left\|\Phi(\|x\|\frac{x}{\|x\|})\right\| = \|x\| \left\|\Phi(\frac{x}{\|x\|})\right\| \geq B\|x\|.$$

und die Ungleichung $\|\Phi(x)\| \geq B\|x\|$ gilt offenbar auch für $x = 0$. Damit ist (9) bewiesen. Es folgt

$$\|\Phi^{-1}(v)\| \leq \frac{1}{B}\|v\| \text{ für alle } v \in V.$$

Schließlich ist nach dem 1. Fall die lineare Abbildung $F \circ \Phi : \mathbb{R}^n \to W$ stetig mit

$$\|F \circ \Phi(x)\| \leq A\|x\| \text{ für alle } x \in \mathbb{R}^n.$$

Damit erhalten wir

$$\|F(v)\| = \|F \circ \Phi(\Phi^{-1}(v))\| \leq A\|\Phi^{-1}(v)\| \leq AB\|v\|.$$

□

Korollar 99 (Die Operatornorm auf $L(V,W)$). *Seien V und W normierte Vektorräume und $V \neq \{0\}$ endlich-dimensional. Dann definiert*

$$\|F\| := \sup_{v \neq 0} \frac{\|F(v)\|}{\|v\|} \quad \textit{für } F \in L(V,W)$$

eine Norm auf dem Vektorraum $L(V,W)$.

Beweis. Nach dem Satz gibt es ein $C \in \mathbb{R}$, so dass für alle $v \neq 0$

$$\frac{\|F(v)\|}{\|v\|} \leq C.$$

Daher ist $\|F\| \in \mathbb{R}$. Die Norm-Eigenschaften sind leicht zu verifizieren. □

Korollar 100. *Sei V ein* endlich-dimensionaler *normierter Vektorraum mit zwei Normen $\|..\|_1, \|..\|_2$. Dann gilt*

(i) Es gibt $c, C > 0$, so dass für alle $v \in V$ gilt

$$c \|v\|_1 \leq \|v\|_2 \leq C \|v\|_1.$$

Man sagt: Je zwei Normen auf einem endlich-dimensionalen Vektorraum sind äquivalent.

(ii) Eine Menge $G \subset V$ ist genau dann bezüglich $\|..\|_1$ offen, wenn sie bezüglich $\|..\|_2$ offen ist. Daher sind auch Begriffe wie Konvergenz, Kompaktheit, Stetigkeit usw. unabhängig von der in V verwendeten Norm.

(iii) $\|..\|_1$-Cauchyfolgen sind auch $\|..\|_2$-Cauchyfolgen. Also ist auch der Begriff Cauchyfolge unabhängig von der in V verwendeten Norm.

Beweis. *Zu (i).* Weil die Identität $\mathrm{id} : (V, \|.\|_2) \to (V, \|.\|_1)$ linear, also stetig ist, gibt es ein $M > 0$ mit

$$\|v\|_1 \leq M \|v\|_2,$$

also

$$\frac{1}{M} \|v\|_1 \leq \|v\|_2$$

für alle v. Die Stetigkeit von id in der anderen Richtung liefert die zweite Ungleichung.

Zu (ii). Ist G offen bezüglich $\|.\|_1$ und betrachtet man wieder id als stetige Abbildung von $(V, \|.\|_2)$ nach $(V, \|.\|_1)$, so ist auch

$$G = \mathrm{id}^{-1}(G)$$

offen. Die umgekehrte Richtung folgt aus Symmetriegründen.

Zu (iii). Folgt leicht aus (i). □

Beispiel 101. Für die l^p-Normen auf $\mathbb{R}^n$ aus Beispiel 91 gilt: Ist $1 \leq p \leq q \leq +\infty$ und $x \in \mathbb{R}^n$, so ist

$$\|x\|_q \leq \|x\|_p \leq n^{\frac{1}{p}-\frac{1}{q}} \|x\|_q, \tag{10}$$

wobei $\frac{1}{+\infty} := 0$.

Beweis: Sei zunächst $q < +\infty$. Die linke Abschätzung ist leicht: Man kann o.E. annehmen, dass

$$1 = (\|x\|_q)^q = \sum_i |x_i|^q.$$

Insbesondere sind dann alle $|x_i| \leq 1$ und daher $|x_i|^p \geq |x_i|^q$. Damit ist $\sum_i |x_i|^p \geq 1$ und

$$\|x\|_p = \left(\sum_i |x_i|^p \right)^{1/p} \geq 1 = \|x\|_p.$$

Die rechte Ungleichung beweisen wir später im Beispiel 185.

Für den Fall $q = +\infty$ vergleiche (4).

□

Definition 102. Ein vollständiger normierter Vektorraum heißt ein *Banachraum*.

Satz 103. *Jeder* endlich-dimensionale *normierte Vektorraum* $(V, \|.\|)$ *ist ein Banachraum.*

Beweis. 1. Fall: $V = \mathbb{R}^n$ mit der Standardnorm. Das haben wir bereits im Beispiel 40 gezeigt.

2. Fall: V beliebiger endlich-dimensionaler $\mathbb{R}$-Vektorraum. Sie $(v_k)_{k\in\mathbb{N}}$ eine Cauchyfolge in V. Nach unserer Definition (oder nach Linearer Algebra) gibt es einen Isomorphismus $\Phi : \mathbb{R}^n \to V$. Dann ist auch $\Phi^{-1} : V \to \mathbb{R}^n$ linear, also stetig, und es gibt ein $C \in \mathbb{R}$ mit

$$\|\Phi^{-1}(v_j) - \Phi^{-1}(v_k)\| = \|\Phi^{-1}(v_j - v_k)\| \leq C\|v_j - v_k\|$$

für alle $j, k \in \mathbb{N}$. Also ist auch $(x_k = \Phi^{-1}(v_k))_{k\in\mathbb{N}}$ eine Cauchyfolge in $\mathbb{R}^n$. Sie ist nach dem 1. Fall konvergent gegen ein $x^* \in \mathbb{R}^n$. Wegen der Stetigkeit von Φ ist deshalb

$$\lim_{k\to\infty} v_k = \lim_{k\to\infty} \Phi(x_k) = \Phi(x^*).$$

□

Beispiel 104. Vergleiche Beispiele 71 und 79. Sei $(V, \|.\|)$ ein n-dimensionaler Banachraum. Die Wahl einer Basis von V liefert nach linearer Algebra einen Isomorphismus

$$\Phi : L(V, V) \to M(n, \mathbb{R})$$

zwischen dem Raum der linearen Abbildungen von V in sich und dem Raum der $(n \times n)$-Matrizen. Wenn wir $L(V, V)$ mit der Operatornorm und $M(n, \mathbb{R})$ zum Beispiel mit dem $\mathbb{R}^{(n^2)}$ identifizieren und mit der entsprechenden Norm ausstatten, ist Φ nach Satz 98 ein Homöomorphismus. Die Determinanten ist nach Beispiel 71 stetig auf $M(n, \mathbb{R})$, und weil die Determinante der linearen Abbildung $F \in L(V, V)$ nach linearer Algebra gerade die Determinante der Matrix $\Phi(F)$ ist, ist auch die Determinantenfunktion auf $L(V, V)$ stetig. Damit ist das Urbild von $\mathbb{R} \setminus \{0\}$, also die invertierbaren Endomorphismen von V, eine *offene* Teilmenge $\mathbf{GL}(V)$, die unter Φ der Menge der invertierbaren Matrizen $\mathbf{GL}(n, \mathbb{R})$ entspricht. Für invertierbare Matrizen sind die Komponenten der Inversen durch gebrochenrationale Funktionen der originalen Komponenten gegeben, also insbesondere stetig. Daher ist die Inversenbildung auf $\mathbf{GL}(n, \mathbb{R})$ und wegen der Φ-Invarianz auch auf $\mathbf{GL}(V)$ eine stetige Abbildung.

□

Wir halten noch einmal das Ergebnis aus dem Korollar 100 fest:

Ein (abstrakter) endlich-dimensionaler $\mathbb{R}$-Vektorraum hat unendlich viele Basen, aber keine von diesen ist besonders ausgezeichnet. Ebenso besitzt er unendlich viele Normen, aber keine von diesen ist besonders ausgezeichnet. Allerdings sind sie alle äquivalent: Die durch sie definierten Metriken liefern alle dieselben offenen Mengen, dieselben konvergenten Folgen, dieselben stetigen Abbildungen. Um über Offenheit, Konvergenz oder Stetigkeit in endlich-dimensionalen $\mathbb{R}$-Vektorräumen zu sprechen, kann man eine beliebige Norm wählen. Weil es aber egal ist, welche man wählt, kann man eben unabhängig von einer solchen Wahl über Offenheit, Konvergenz oder Stetigkeit reden.
Der $\mathbb{R}^n$ besitzt eine Standardbasis und eine Standardnorm, die die Standardmetrik d^2 liefert. Natürlich kann man davon Gebrauch machen, oft muss man aber nicht ...
Mehr zu diesem Thema gleich in der Vorbemerkung zum nächsten Abschnitt und im Abschnitt 2.7 über die klassische Vektoranalysis.

2 Grundlagen der mehrdimensionalen Differentiation

Wir werden die Differentialrechnung in endlich-dimensionalen Banachräumen entwickeln. Nach dem vorangehenden Abschnitt sind diese isomorph zu einem $\mathbb{R}^n$, und man könnte sich auch auf die letzteren beschränken.

Der Vorteil wäre, dass man im $\mathbb{R}^n$ eine ausgezeichnete Basis und damit ausgezeichnete Koordinaten hat. Dadurch wird die Theorie konkreter. Man könnte die Differentialrechung auf dem Begriff der partiellen Ableitung, also der Ableitung nach einer einzelnen Variablen, aufbauen.

Der Nachteil wäre, dass man im $\mathbb{R}^n$ eine ausgezeichnete Basis und damit ausgezeichnete Koordinaten hat. Diese verschleiern die Tatsache, dass die Konzepte der Differentialrechung geometrischer Natur sind und mit speziellen Koordinaten nichts zu tun haben, vielleicht aber sehr viel mit anderen Strukturen, die auf dem $\mathbb{R}^n$ auch noch so selbstverständlich vorkommen, dass wir sie gar nicht bemerken.

Zum Beispiel ist $V := \{(x, y, z) \in \mathbb{R}^3 \,|\, x + y + z = 0\}$ ein zweidimensionaler Untervektorraum des $\mathbb{R}^3$. Hat man darauf eine differenzierbare Funktion $f : V \to \mathbb{R}$ gegeben, so ist es zunächst unklar, was ihre partiellen Ableitungen sein sollen. Erst wenn man in V eine Basis gewählt hat und damit eine Isomorphie von V auf $\mathbb{R}^2$, macht der Begriff der partiellen Ableitungen von f einen Sinn. Allerdings für jede Basiswahl einen anderen. Und es gibt keine „kanonische“ Weise, eine Basis zu wählen. Hingegen kann man den viel wichtigeren Begriff des *Gradienten* ohne partielle Ableitungen definieren, braucht dafür aber ein Skalarprodukt. Und das Skalarprodukt des $\mathbb{R}^3$ liefert auf ganz kanonische Weise eines für den Untervektorraum V. (Vgl. Abschnitt 2.7.1.)

2.1 Die Ableitung

- Wir lernen die Ableitung als lineare Approximation einer Abbildung in der Nähe eines Punktes kennen.
- Wir berechne die Ableitung in einfachsten Fällen.

Im folgenden seien V, W endlich-dimensionale Banachräume[4] und G eine offene Teilmenge von V.

Definition 105. Sei $f : V \supset G \to W$ eine Abbildung der offenen Menge G.

(i) f heißt *differenzierbar in* $p \in G$, wenn es eine lineare Abbildung $F : V \to W$ gibt, so dass für die durch

$$f(x) = f(p) + F(x - p) + R(x) \tag{11}$$

definierte „Restfunktion“ $R : G \to W$ gilt

$$\lim_{x \to p} \frac{R(x)}{\|x - p\|} = 0. \tag{12}$$

F ist dann eindeutig bestimmt, vgl. Lemma 106, und wir nennen es *die Ableitung* oder *das Differential von* f *in* p.

Notation:

$$F = D_p f = d_p f.$$

(ii) f heißt *differenzierbar (auf* G*)*, falls f in allen $p \in G$ differenzierbar ist.

[4] Im folgenden genügt es, wenn V endlich-dimensional ist. Aber da wir keine konkreten Anwendungen für unendlich-dimensionales W im Sinn haben, sei der Einfachheit halber auch W endlich-dimensional.

Bemerkungen.

- *Analytisch* gesprochen ist D_pf die lineare Approximation von f in der Nähe von p. Schreibt man x statt p und Δx für $x - p$, so erhält man

$$\Delta f := f(x + \Delta x) - f(x) \approx D_x f(\Delta x).$$

- Die Notation $f'(p)$ für die Ableitung finde ich weniger empfehlenswert, weil die Ableitung eine lineare Abbildung ist, so dass man dann $f'(p)(v)$ schreiben müßte. Wir heben uns diese Schreibweise daher auf für den Fall, wo D_pf auf kanonische Weise durch eine *Matrix* gegeben ist, vgl. Beispiel 114.

Lemma 106. *Ist f in p differenzierbar und F wie in der Definition, so gilt für alle v in V:*

$$F(v) = \lim_{t \to 0} \frac{f(p + tv) - f(p)}{t}.$$

Beachten Sie: Weil der Definitionsbereich G von f offen ist, liegt für jedes $v \in V$ und hinreichend kleines $|t|$ der Punkt $p + tv$ in G. Deshalb ist der Limes sinnvoll. Die Definition der Differenzierbarkeit kann man auch für Abbildungen von nicht-offenen Teilmengen hinschreiben, aber die Ableitung ist dann im allgemeinen nicht mehr eindeutig.

Beweis. Ist F wie in der Definition, so folgt

$$\frac{f(p + tv) - f(p)}{t} = \frac{f(p) + F(tv) + R(p + tv) - f(p)}{t} = F(v) + \frac{R(p + tv)}{t}.$$

Aber

$$\frac{R(p + tv)}{t} = \underbrace{\frac{R(p + tv)}{\|p + tv - p\|}}_{\to 0} \underbrace{\frac{|t| \, \|v\|}{t}}_{= \pm \|v\|}.$$

□

Differenzierbarkeit und das Differential hängen wegen Korollar 100 nicht ab von den gewählten Normen auf V und W. Wir werden deshalb die Norm oft auch nicht spezifizieren. Wenn man eine braucht, nimmt man eine.

Beispiel 107 (Der Fall $\mathbb{R} \to \mathbb{R}$). Wie hängt die neue Ableitungsdefinition mit der aus dem letzten Semester zusammen?

Die einzigen linearen Abbildungen von $\mathbb{R}$ nach $\mathbb{R}$ sind die Abbildungen $x \mapsto ax$ mit einem festen $a \in \mathbb{R}$. Eine Abbildung $f : \mathbb{R} \supset G \to \mathbb{R}$ ist deshalb differenzierbar im Sinne der Analysis I genau dann, wenn sie auch nach der neuen Definition differenzierbar ist. Dann gilt für $p \in G$ und $v \in \mathbb{R}$

$$D_pf(v) = f'(p)v,$$

d.h.

$$\boxed{f'(p) = D_pf(1)} \tag{13}$$

oder verbal:

Neue Ableitung = Multiplikation mit der alten Ableitung.

□

Beispiel 108. Sei $f : \mathbb{R}^2 \to \mathbb{R}, (x, y) \mapsto 1+3x+4y+5xy^2$. Dann ist f in $(0, 0)$ differenzierbar mit

$$D_{(0,0)}f(u, v) = 3u + 4v.$$

Es ist nämlich

$$f(x, y) = 1 + 3(x - 0) + 4(y - 0) + 5xy^2,$$

und weil $5xy^2$ in $(x - 0)$ und $(y - 0)$ „kubisch ist“, geht der Rest für $(x, y) \to (0, 0)$ gegen null:

$$\left|\frac{5xy^2}{\sqrt{x^2+y^2}}\right| = 5\left|\frac{x}{\sqrt{x^2+y^2}}\, y^2\right| \leq 5y^2.$$

Durch Nachrechnen können Sie bestätigen, dass

$$f(x, y) = 32 + 23(x - 1) + 24(y - 2) + \underbrace{20(x - 1)(y - 2) + 5(y - 2)^2 + 5(x - 1)(y - 2)^2}_{=:R(x,y)}.$$

Der Rest dividiert durch $\sqrt{(x-1)^2 + (y-2)^2}$ geht für $(x, y) \to (1, 2)$ wieder gegen null. Damit ist f auch in $(x, y) = (1, 2)$ differenzierbar und

$$D_{(1,2)}f(u, v) = 23u + 24v.$$

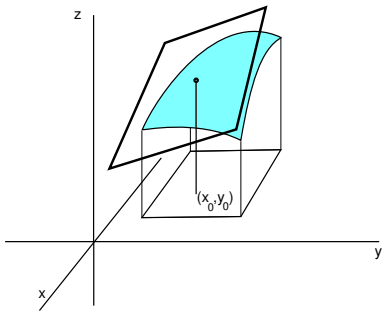
(Die Umrechnung von f auf den Punkt $(x, y) = (1, 2)$ geschieht erst für x und dann für y mit der Taylorformel aus Analysis I. Vgl. auch Satz 148).

□

Geometrische Interpretation.
Die *geometrische Interpretation* ist am einfachsten im Fall $f : \mathbb{R}^2 \supset G \to \mathbb{R}$. Dann ist der Graph von

$$x \mapsto f(p) + D_p f(x - p)$$

eine Ebene im $\mathbb{R}^3$, die *Tangentialebene* an den Graphen von f.



Beispiel 109. Die (offene) obere Einheits-Halbkugel ist gegeben durch den Graphen von

$$f(x, y) = \sqrt{1 - x^2 - y^2}$$

über der offenen Menge $\{(x, y) \,|\, x^2 + y^2 < 1\}$.

Die Tangentialebene im Punkt $(x_0, y_0, \sqrt{1 - x_0^2 - y_0^2})$ ist gegeben durch den Graphen der linearen Abbildung

$$D_{(x_0,y_0)}f : \mathbb{R}^2 \to \mathbb{R},$$

falls f differenzierbar ist. Aber das wissen wir noch nicht, und wir wissen auch noch nicht, wie wir $D_{(x_0,y_0)}f$ berechnen sollen.

□

Berechnung von Ableitungen. Das ist offenbar ein wichtiges Problem, dem wir noch länger nachgehen werden. Wir beginnen mit zwei ganz trivialen Fällen:

Beispiel 110 (Konstante Abbildungen). Sei $f : V \supset G \to W$ konstant. Dann ist $D_p f = 0$ für alle $p \in G$.

□

Beispiel 111 (Lineare Abbildungen). Sei $f : V \to W$ linear. Dann ist

$$f(x) = f(p) + f(x - p) + 0.$$

Also ist f in jedem Punkt $p \in V$ differenzierbar und $D_p f = f$. Zum Beispiel ist die Additionsabbildung

$$\alpha : V \times V \to V, (x, y) \mapsto \alpha(x, y) = x + y$$

vom Vektorraum $V \times V$ der Paare in den Vektorraum V linear:

$$\begin{aligned}\alpha(\lambda_1(x_1, y_1) + \lambda_2(x_2, y_2)) &= \alpha((\lambda_1 x_1 + \lambda_2 x_2, \lambda_1 y_1 + \lambda_2 y_2)) \\ &= \lambda_1 x_1 + \lambda_2 x_2 + \lambda_1 y_1 + \lambda_2 y_2 \\ &= \lambda_1 \alpha(x_1, y_1) + \lambda_2 \alpha(x_2, y_2).\end{aligned}$$

Also ist α differenzierbar, und für alle (x, y) und (u, v) in $V \times V$ ist

$$D_{(x,y)}\alpha(u, v) = u + v.$$

□

Nun ein etwas anspruchsvolleres Beispiel.

Beispiel 112 (Skalarmultiplikation). Die Abbildung der Skalarmultiplikation

$$\mu : \mathbb{R} \times V \to V, (\lambda, x) \mapsto \lambda x$$

ist differenzierbar in jedem $(\lambda_0, x_0) \in \mathbb{R} \times V$. Es ist nämlich

$$\mu(\lambda, x) = \lambda x = \underbrace{\lambda_0 x_0}_{\mu(\lambda_0, x_0)} + \underbrace{(\lambda - \lambda_0)x_0 + \lambda_0(x - x_0)}_{=: F(\lambda - \lambda_0, x - x_0)} + \underbrace{(\lambda - \lambda_0)(x - x_0)}_{=: R(\lambda, x)}.$$

Diese Gleichung rechnet man leicht nach. Es bleibt zu zeigen, dass

$$\lim_{(\lambda,x)\to(\lambda_0,x_0)} \frac{R(\lambda, x)}{\|(\lambda, x) - (\lambda_0, x_0)\|} = 0.$$

Dazu braucht man eine Norm auf $\mathbb{R} \times V$. Wir nehmen an, dass auf V eine Norm $\|.\|$ gegeben ist, und definieren

$$\|(\lambda, x)\| := |\lambda| + \|x\|.$$

Rechnen Sie nach, dass das wirklich eine Norm definiert. Damit gilt dann:

$$\frac{|R(\lambda, x)|}{\|(\lambda, x) - (\lambda_0, x_0)\|} = \frac{|\lambda - \lambda_0|\|x - x_0\|}{|\lambda - \lambda_0| + \|x - x_0\|} \leq \|x - x_0\| \to 0$$

für $(\lambda, x) \to (\lambda_0, x_0)$. Daraus folgt die Behauptung. Wir halten fest:

$$\boxed{D_{(\lambda_0,x_0)}\mu(\lambda, x) = \lambda_0 x + \lambda x_0.}$$

Das ist eine Art Produktregel, auf die wir noch zurückkommen.

□

Die beiden folgenden Beispiele sind überaus wichtig! Sie stellen einen ersten Schritt zur expliziten praktischen Berechnung von Ableitungen dar.

Beispiel 113 (Komponentenweise Differentiation). Sei

$$f = (f_1, \ldots, f_m) : V \supset G \to \mathbb{R}^m$$

mit Komponentenfunktionen $f_i : G \to \mathbb{R}$. Dann ist f genau dann in p differenzierbar, wenn alle f_i in p differenzierbar sind. In diesem Fall gilt

$$D_p f(v) = (D_p f_1(v), \ldots, D_p f_m(v)) \text{ für alle } v \in V. \tag{14}$$

Beweis. In Komponenten sieht die Gleichung (11) so aus:

$$f_i(x) = f_i(p) + F_i(x - p) + R_i(x).$$

Nun ist F linear genau dann, wenn alle Komponenten F_i linear sind. Und weil Konvergenz im $\mathbb{R}^n$ einfach komponentenweise Konvergenz ist, folgt die Behauptung durch Betrachtung der Komponenten $\frac{R_i(x)}{\|x-p\|}$ von $\frac{R(x)}{\|x-p\|}$.

Dieses Beispiel gestattet eine Verallgemeinerung auf folgende Situation:
Seien V und $W_1, \ldots, W_m$ endlich-dimensionale Banachräume, $G \subset V$ offen und seien

$$f_i : V \supset G \to W_i$$

für $i \in \{1, \ldots, m\}$ Abbildungen. Wir definieren

$$\begin{aligned} f : V \supset G \to & W_1 \times \ldots \times W_m \\ & x \mapsto (f_1(x), \ldots, f_m(x)). \end{aligned}$$

Dann ist f genau dann differenzierbar in p, wenn alle f_i in p differenzierbar sind, und es gilt wieder die Gleichung (14).

□

Beispiel 114 (Funktionalmatrix). $f : \mathbb{R}^n \supset G \to \mathbb{R}^m$ sei differenzierbar in $p \in G$. Dann ist $D_p f : \mathbb{R}^n \to \mathbb{R}^m$ eine lineare Abbildung, und eine solche wird nach Linearer Algebra dargestellt durch eine Matrix, die wir mit $f'(p)$ bezeichnen und die *Jacobimatrix* oder *Funktionalmatrix* von f in p nennen. Die Spalten sind gerade die Bilder der Basisvektoren $e_1, \ldots, e_n$:

$$f'(p) = (D_p f(e_1) \ldots D_p f(e_n)) = (D_p f_i(e_j)) = \begin{pmatrix} D_p f_1(e_1) & \ldots & D_p f_1(e_n) \\ \vdots & & \vdots \\ D_p f_m(e_1) & \ldots & D_p f_m(e_n) \end{pmatrix}.$$

Die Formel im Lemma 106 liefert eine Möglichkeit, die $D_p f_m(e_j)$ zu berechnen. Wir kommen im Abschnitt 2.3 darauf zurück.

□

Beispiel 115 (Kurven). Eine Abbildung $f : \mathbb{R} \supset]a, b[\to W$ nennt man eine Kurve in W. Ist f in $t \in]a, b[$ differenzierbar, so ist für alle $\lambda \in \mathbb{R}$

$$D_t f(\lambda) = \lambda D_t f(1),$$

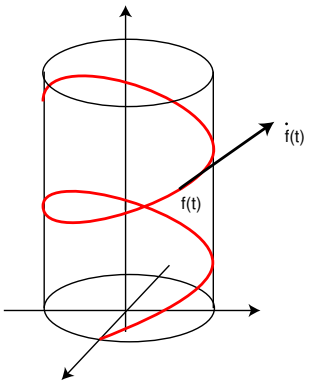
Also ist $D_t f : \mathbb{R} \to W$ durch den *Tangentialvektor* $\dot{f}(t) := D_t f(1)$ eindeutig bestimmt.
Ist $W = \mathbb{R}^m$ und $f = (f_1, \ldots, f_m)$, so ist

$$\dot{f}(t) = \left(\dot{f}_1(t), \ldots, \dot{f}_m(t)\right).$$

Dabei ist nach Lemma 106

$$\dot{f}_i(t) = D_t f_i(1) = \lim_{\tau \to 0} \frac{f_i(\tau + t) - f_i(t)}{\tau}.$$

$\dot{f}_i(t)$ ist also die gewöhnliche Ableitung der Analysis I. Insbesondere können wir die Definition von $\dot{f}$ auch auf den Fall kompakter Intervalle $[a, b]$ ausdehnen.



Konkret: Die Kurve $f : \mathbb{R} \to \mathbb{R}^3, \quad t \mapsto (\cos t, \sin t, t)$ ist eine Spiralkurve. Sie hat den Geschwindigkeitsvektor

$$\dot{f}(t) = (-\sin t, \cos t, 1).$$

Und es gilt zum Beispiel

$$D_2 f(-5) = (5 \sin 2, -5 \cos 2, -5).$$

□

2.2 Rechenregeln für differenzierbare Abbildungen

- Differenzierbare Abbildungen sind stetig.
- Die wichtigsten Hilfsmittel zur Berechnung von Ableitungen sind wie in der Analysis I die Kettenregel und die Produktregel, die wir hier kennenlernen.
- Wir betrachten viele Beispiele für multilineare Abbildungen (Produkte) und sehr wichtige Beispiele von Ableitungen.
- Es lohnt sich, die Formeln (21), (23), (24) und (25) auswendig zu wissen.

Satz 116. *Seien V, W endlich-dimensionale Banachräume. Ist $f : V \supset G \to W$ in $p \in G$ differenzierbar, so ist es dort auch stetig.*

Beweis. Die Behauptung folgt aus

$$f(x) = f(p) + D_p f(x - p) + R(x),$$

weil für $x \to p$ das Restglied gegen 0 geht, und weil die lineare Abbildung $D_p f$ auf einem endlich-dimensionalen Banachraum stetig ist. □

Satz 117 (Kettenregel). *Seien U, V, W endlich-dimensionale Banachräume, $G \subset U$ und $H \subset V$ offen, $g : G \to V$ und $f : H \to W$ Abbildungen mit $g(G) \subset H$. Sei g differenzierbar in $p \in G$ und f differenzierbar in $q = g(p) \in H$. Dann ist die Abbildung $f \circ g : G \to W$ differenzierbar in p, und es gilt:*

$$D_p(f \circ g) = D_{g(p)} f \circ D_p g.$$

Beweis. Die definierenden Gleichungen

$$\begin{aligned} g(x) &= g(p) + D_p g(x - p) + R(x) \\ f(y) &= f(q) + D_q f(y - q) + S(y) \end{aligned}$$

implizieren

$$\begin{aligned} f(g(x)) &= f(g(p)) + D_q f(g(x) - g(p)) + S(g(x)) \\ &= f(g(p)) + D_q f(D_p g(x - p)) + \underbrace{D_q f(R(x)) + S(g(x))}_{=:T(x)}. \end{aligned}$$

Es bleibt zu zeigen, dass

$$\lim_{x \to p} \frac{T(x)}{\|x - p\|} = 0. \tag{15}$$

Im folgenden benutzen wir die im Korollar 99 definierte Operatornorm.

Zunächst gilt

$$\frac{\|D_q f(R(x))\|}{\|x - p\|} \leq \|D_q f\| \frac{\|R(x)\|}{\|x - p\|} \to 0 \tag{16}$$

für $x \to p$.

Schwieriger ist der zweite Summand von $T(x)$. Die Behauptung

$$\lim_{y\to q}\frac{S(y)}{\|y-q\|}=0$$

ist äquivalent zur Behauptung:

$$\forall_{\epsilon>0}\exists_{\delta>0}(\|y-q\|<\delta \implies \|S(y)\|\le\epsilon\|y-q\|).$$

Sei $\epsilon>0$ und sei $\delta>0$ dazu wie vorstehend gewählt. Weil g stetig ist in p, gibt es ein $\eta>0$, so dass

$$\|x-p\|<\eta \implies \|g(x)-g(p)\|<\delta.$$

Für $\|x-p\|<\eta$ ist dann also

$$\|S(g(x))\|\le\epsilon\|g(x)-g(p)\|=\epsilon\|D_pg(x-p)+R(x)\|\le\epsilon(\|D_pg\|\,\|x-p\|+\|R(x)\|).$$

Weil $\lim_{x\to p}\frac{R(x)}{\|x-p\|}=0$, kann man annehmen, dass $\eta>0$ so klein ist, dass

$$\frac{\|R(x)\|}{\|x-p\|}<1 \quad \text{für } 0<\|x-p\|<\eta.$$

Dann folgt

$$\|S(g(x))\|\le\epsilon(\|D_pg\|\,\|x-p\|+\|x-p\|)=\epsilon(\|D_pg\|+1)\|x-p\|.$$

Wir haben also zu jedem $\epsilon>0$ ein $\eta>0$ gefunden, so dass

$$0<\|x-p\|<\eta \implies \frac{\|S(g(x))\|}{\|x-p\|}\le\epsilon(\|D_pg\|+1).$$

Das bedeutet aber

$$\lim_{x\to p}\frac{\|S(g(x))\|}{\|x-p\|}=0. \tag{17}$$

Aus (16) und (17) folgt (15) und damit die Behauptung. □

Die Skalarmultiplikation $\mathbb{R}\times V\to V,(\lambda,v)\mapsto\lambda v$ eines Vektorraums ist in jedem der beiden Argument linear, man nennt das *bilinear*. Eine Verallgemeinerung sind die *multilinearen* oder *k-linearen* Abbildungen, zum Beispiel die Determinante. Der folgende Satz verallgemeinert das Beispiel 112 auf multilineare Abbildungen.

Satz 118 (Produktregel). *Seien $V_1,\dots,V_k,W$ endlich-dimensionale Banachräume und*

$$\mu: V_1\times\ldots\times V_k\to W$$

eine k-lineare *Abbildung, d.h. μ ist in jedem seiner k Argumente linear. Dann ist μ differenzierbar und es gilt*

$$D_{(p_1,\dots,p_k)}\mu(v_1,\dots,v_k)=\sum_{i=1}^{k}\mu(p_1,\dots,p_{i-1},v_i,p_{i+1},\dots,p_k).$$

Bemerkung: Der erste Summand ist zu interpretieren als $\mu(v_1,p_2,\dots,p_k)$, der letzte entsprechend.

Beweis. A. Wir zeigen zunächst die Stetigkeit von μ, genauer: Es gibt C mit

$$\|\mu(x_1,\dots,x_k)\|\le C\|x_1\|\cdot\ldots\cdot\|x_k\| \text{ für alle } x_i\in V_i. \tag{18}$$

Ist $e_1, \ldots, e_n$ eine Basis von V, so sind die Koordinatenabbildungen

$$x = \sum_j x_j e_j \;\mapsto\; x_i$$

linear, also stetig, und es gibt zu jedem i eine Konstante C_i mit $|x_i| \le C_i \|x\|$ für alle x.

Wir wählen nun Basen $e_{i1}, \ldots, e_{in_i}$ für V_i und schreiben $x_i = \sum_j x_{ij} e_{ij} \in V_i$. Aus der Multilinearität folgt dann

$$\begin{aligned}
\|\mu(x_1, \ldots, x_k)\| &= \| \sum_{j_1, \ldots, j_k} x_{1j_1} \cdot \ldots \cdot x_{kj_k} \mu(e_{1j_1}, \ldots, e_{k,j_k}) \| \\
&\le \sum_{j_1, \ldots, j_k} C_{1j_1} \|x_1\| \cdot \ldots \cdot C_{kj_k} \|x_k\| \cdot \|\mu(e_{1j_1}, \ldots, e_{kj_k})\| \\
&= \underbrace{\left(\sum_{j_1, \ldots, j_k} C_{1j_1} \ldots C_{kj_k} \|\mu(e_{1j_1}, \ldots, e_{kj_k})\| \right)}_{=:C} \|x_1\| \cdot \ldots \cdot \|x_k\|.
\end{aligned}$$

B. Nun zum eigentlichen Beweis. Dazu müssen wir den Restterm

$$R(x_1, \ldots, x_k) = \mu(x_1, \ldots, x_k) - \mu(p_1, \ldots, p_k) - \sum_{i=1}^{k} \mu(p_1, \ldots, p_{i-1}, x_i - p_i, p_{i+1}, \ldots, p_k)$$

berechnen. Dann müssen wir eine Norm $\|.\|$ auf $V_1 \times \ldots \times V_k$ wählen und zeigen, dass

$$\lim_{(x_1, \ldots, x_k) \to (p_1, \ldots, p_k)} \frac{R(x_1, \ldots, x_k)}{\|(x_1, \ldots, x_k) - (p_1, \ldots, p_k)\|} = 0.$$

Eigentlich müssen wir den Restterm natürlich gar nicht berechnen, sondern wir müssen ihn in einer Form schreiben, die es ermöglicht, den Grenzwert zu berechnen. Dazu führen wir folgende Schreibweise ein:

$$\sum_{1 \le j_1 < \ldots < j_m \le k} \mu_{j_1 \ldots j_m}(p, x) := \sum_{1 \le j_1 < \ldots < j_m \le k} \mu(p_1, \ldots, x_{j_1} - p_{j_1}, \ldots, x_{j_m} - p_{j_m}, \ldots, p_k), \quad (19)$$

wobei über alle Produkte summiert wird, die aus $\mu(p_1, \ldots, p_k)$ entstehen, indem man die p_{j_i} ersetzt durch $x_{j_i} - p_{j_i}$. Der Restterm ist dann also

$$R(x_1, \ldots, x_k) = \mu(x_1, \ldots, x_k) - \mu(p_1, \ldots, p_k) - \sum_{1 \le j_1 \le k} \mu_{j_1}(p, x).$$

Wir zeigen gleich in einem Lemma, dass dann

$$R(x_1, \ldots, x_k) = \sum_{m=2}^{k} \sum_{1 \le j_1 < \ldots < j_m \le k} \mu_{j_1 \ldots j_m}(p, x). \quad (20)$$

Jeder Summand der rechten Seite enthält also mindestens zwei "Faktoren" der Form $(x_j - p_j)$ und geht deshalb für $(x_1, \ldots, x_k) \to (p_1, \ldots, p_k)$ mindestens quadratisch gegen 0.

Genauer: Ist $\|.\|_i$ eine Norm auf V_i, so definiert

$$\|(x_1, \ldots, x_k)\| := \|x_1\|_1 + \ldots + \|x_k\|_k$$

eine Norm auf $V_1 \times \ldots \times V_k$, und weil nach (18)

$$\begin{aligned}&\frac{\|\mu(p_1, \ldots, x_{j_1} - p_{j_1}, \ldots, x_{j_2} - p_{j_2}, \ldots, p_k)\|}{\|((x_1, \ldots, x_k) - (p_1, \ldots, p_k)\|} \\ &\leq C\|p_1\| \ldots \underbrace{\frac{\|x_{j_1} - p_{j_1}\|}{\|p_1\| + \ldots + \|x_{j_1} - p_{j_1}\| + \ldots + \|p_k\|}}_{\leq 1} \ldots \underbrace{\|x_{j_2} - p_{j_2}\|}_{\to 0} \ldots \|p_k\|,\end{aligned}$$

geht das Restglied gegen 0. □

Die Restgliedformel (20) folgt aus

Lemma 119. *Für jede k-lineare Abbildung $\mu : V_1 \times \ldots \times V_k \to W$ und alle $(x_1, \ldots, x_k)$ und $(p_1, \ldots, p_k) \in V_1 \times \ldots \times V_k$ gilt unter der Verwendung von* (19)

$$\mu(x_1, \ldots, x_k) = \mu(p_1, \ldots, p_k) \quad + \sum_{m=1}^{k} \sum_{1 \leq j_1 < \ldots < j_m \leq k} \mu_{j_1 \ldots j_m}(p, x)$$

Die p-Terme auf der rechten Seite heben sich also weg.

Beweis. Wir zeigen das durch vollständige Induktion über k.

$\underline{k = 1.}$ Die Formel

$$\mu(x_1) = \mu(p_1) + \mu(x_1 - p_1)$$

folgt aus der 1-Linearität.

$\underline{k \to k+1.}$ Sei also $\mu : V_1 \times \ldots \times V_{k+1} \to W$ eine k-lineare Abbildung. Dann ist

$$\begin{aligned}(*) :=&\mu(p_1, \ldots, p_{k+1}) + \sum_{m=1}^{k+1} \sum_{1 \leq j_1 < \ldots < j_m \leq k+1} \mu_{j_1 \ldots j_m}(p, x) \\ &= \mu(p_1, \ldots, p_{k+1}) + \sum_{m=1}^{k+1} \sum_{1 \leq j_1 < \ldots < j_m \leq k} \mu_{j_1 \ldots j_m}(p, x) \\ &\quad + \sum_{m=1}^{k+1} \sum_{1 \leq j_1 < \ldots < j_m = k+1} \mu_{j_1 \ldots j_m}(p, x).\end{aligned}$$

Im mittleren Term kann $m = k + 1$ nicht vorkommen. Deshalb können wir fortfahren

$$\begin{aligned}(*) = &\mu(p_1, \ldots, p_{k+1}) + \sum_{m=1}^{k} \sum_{1 \leq j_1 < \ldots < j_m \leq k} \mu_{j_1 \ldots j_m}(p, x) \\ &+ \mu(p_1, \ldots, p_k, x_{k+1} - p_{k+1}) + \sum_{m=2}^{k+1} \sum_{1 \leq j_1 < \ldots < j_m = k+1} \mu_{j_1 \ldots j_m}(p, x).\end{aligned}$$

Wir definieren nun zwei k-lineare Abbildungen auf $V_1 \times \ldots \times V_k$ durch

$$\begin{aligned}\mu^0(x_1, \ldots, x_k) &:= \mu(x_1, \ldots, x_k, p_{k+1}), \\ \mu^1(x_1, \ldots, x_k) &:= \mu(x_1, \ldots, x_k, x_{k+1}).\end{aligned}$$

Wir erhalten

$$\begin{aligned}
(*) &= \mu^0(p_1,\dots,p_k) + \sum_{m=1}^{k} \sum_{1\le j_1<\dots<j_m\le k} \mu^0_{j_1\dots j_m}(p,x) \\
&\quad + \mu^1(p_1,\dots,p_k) - \mu^0(p_1,\dots,p_k) \\
&\quad + \sum_{m=2}^{k+1} \sum_{1\le j_1<\dots<j_{m-1}\le k} \mu^1_{j_1\dots j_m}(p,x) - \sum_{m=2}^{k+1} \sum_{1\le j_1<\dots<j_{m-1}\le k} \mu^0_{j_1\dots j_m}(p,x) \\
&= \sum_{m=1}^{k} \sum_{1\le j_1<\dots<j_m\le k} \mu^0_{j_1\dots j_m}(x,p) + \mu^1(p_1,\dots,p_k) \\
&\quad + \sum_{m=1}^{k} \sum_{1\le j_1<\dots<j_m\le k} \mu^1_{j_1\dots j_m}(p,x) - \sum_{m=1}^{k} \sum_{1\le j_1<\dots<j_m\le k} \mu^0_{j_1\dots j_m}(p,x) \\
&= \mu^1(p_1,\dots,p_k) + + \sum_{m=1}^{k} \sum_{1\le j_1<\dots<j_m\le k} \mu^1_{j_1\dots j_m}(p,x) \\
&\underset{Ind.Vor.}{=} \mu^1(x_1,\dots,x_k) = \mu(x_1,\dots,x_{k+1}).
\end{aligned}$$

□

Die Produktregel aus der Analysis I ist eine Kombination aus der vorstehenden Produktregel mit der Kettenregel. Das erklären wir genauer:

Beispiel 120 (Alte Produktregel). Seien $J \subset \mathbb{R}$ ein offenes Intervall, $p \in J$ und seien $f, g : J \to \mathbb{R}$ differenzierbare Funktionen. Sei weiter $\mu : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ die Multiplikationsabbildung $(x, y) \mapsto xy$ und sei

$$(f, g) : J \to \mathbb{R} \times \mathbb{R}, \quad t \mapsto (f(t), g(t)).$$

Wir betrachten die Komposition

$$h := \mu \circ (f, g) : t \mapsto f(t)g(t).$$

Dann gilt

$$\begin{aligned}
h'(p) &\underset{(13)}{=} D_p h(1) \\
&\underset{\text{Kettenregel}}{=} D_{(f(p),g(p))}\mu \circ D_p(f,g)(1) \\
&\underset{(14)}{=} D_{(f(p),g(p))}\mu \circ (D_p f(1), D_p g(1)) \\
&\underset{(13)}{=} D_{(f(p),g(p))}\mu \circ (f'(p), g'(p)) \\
&\underset{\text{Produktregel}}{=} \mu(f'(p), g(p)) + \mu(f(p), g'(p)) \\
&= f'(p)g(p) + f(p)g'(p).
\end{aligned}$$

□

Beispiel 121. Hier sind wichtige multilineare Produkte. Überlegen Sie, was in jedem einzelnen Fall die Produktregel besagt.

(i) Das normale Produkt reeller Zahlen hatten wir gerade

$$\mathbb{R} \times \mathbb{R} \to \mathbb{R}, \ (x, y) \mapsto xy.$$

(ii) Matrix mal Vektor: Sei $M(m \times n, \mathbb{R})$ der Vektorraum der reellen $(m \times n)$-Matrizen.

$$M(m \times n, \mathbb{R}) \times \mathbb{R}^n \to \mathbb{R}^m,\ (A, x) \mapsto Ax.$$

(iii) Allgemeiner

$$L(V, W) \times V \to W,\ (f, v) \mapsto f(v)$$

(iv) Das kanonische Skalarprodukt im $\mathbb{R}^n$

$$\mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R},\ (x, y) \mapsto \langle x, y \rangle = \sum_{i=1}^{n} x_i y_i.$$

(v) Allgemeiner jedes Skalarprodukt

$$V \times V \to \mathbb{R},\ (x, y) \mapsto \langle x, y \rangle$$

auf einem Euklidischen Vektorraum V, oder noch allgemeiner jede Bilinearform.

(vi) Das Kreuzprodukt

$$\mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}^3,\ (x, y) \mapsto x \times y$$

auf dem $\mathbb{R}^3$.

(vii) Die Determinante als Funktion der Spalten:

$$\mathbb{R}^n \times \ldots \times \mathbb{R}^n \to \mathbb{R},\ (x_1, \ldots, x_n) \mapsto \det(x_1, \ldots, x_n).$$

Dieses führen wir in einem Beispiel weiter unten aus.

□

Beispiel 122. Wir betrachten einen Euklidischen Vektorraum $(V, \langle ., . \rangle)$ und dazu: die Produktabbildung

$$\mu : V \times V \to \mathbb{R}, \quad (x, y) \mapsto \langle x, y \rangle ,$$

die lineare "Diagonal-Abbildung"

$$\delta : V \to V \times V, \quad x \mapsto (x, x),$$

und die "Norm-Abbildung"

$$r : V \to \mathbb{R}, \quad x \mapsto \|x\| = \sqrt{\langle x, x \rangle}$$

Dann ist

$$r = \sqrt{\ } \circ \mu \circ \delta.$$

Seien $x, p \in V$ und $p \neq 0$. Wir wollen zeigen, dass r in p differenzierbar ist und das Differential ausrechnen.

- δ ist als lineare Abbildung differenzierbar und

$$D_p \delta(x) = \delta(x) = (x, x).$$

- μ ist nach der Produktregel differenzierbar und

$$D_{(p,p)} \mu(x, x) = \mu(x, p) + \mu(p, x) = 2\mu(x, p).$$

- Die Wurzel $\sqrt{\ }$ ist nach Analysis I differenzierbar und

$$(\sqrt{\ })'(\tilde{t}) = \frac{1}{2\sqrt{\tilde{t}}}.$$

Das bedeutet

$$D_{\tilde{t}}\sqrt{\ }(t) = \frac{1}{2\sqrt{\tilde{t}}}\, t.$$

Nimmt man die vorstehenden Ergebnisse zusammen, so sieht man, dass r in p nach der Kettenregel differenzierbar ist und

$$\boxed{D_p r(x) = \frac{1}{2\sqrt{\mu(p,p)}}\, 2\mu(x,p) = \frac{\langle x,p\rangle}{r(p)}.} \tag{21}$$

□

Beispiel 123. Für Vektoren $x_1, \ldots, x_n \in \mathbb{R}^n$ schreiben wir

$$X := (x_1, \ldots, x_n)$$

für die Matrix mit den Spalten x_j. Ist $(e_1, \ldots, e_n)$ die kanonische Basis des $\mathbb{R}^n$, so ist also

$$E := (e_1, \ldots, e_n)$$

die n-reihige Einheitsmatrix. So identifizieren wir den Raum $\mathbb{R}^n \times \ldots \times \mathbb{R}^n$ mit dem Raum $M(n \times n, \mathbb{R})$ der quadratischen n-reihigen Matrizen. Die Determinante – nehmen Sie das als Definition, wenn Sie in der Linearen Algebra noch nicht so weit sind – ist mit dieser Identifikation eine n-lineare Abbildung

$$\det : \mathbb{R}^n \times \ldots \mathbb{R}^n \to \mathbb{R}, \quad (x_1, \ldots, x_n) \mapsto \det(x_1, \ldots, x_n)$$

mit folgenden zusätzlichen Eigenschaften:

$$\begin{gathered} \det(x_1, \ldots, x_n) = 0, \text{falls zwei der } x_j \text{ gleich sind,} \\ \det(e_1, \ldots, e_n) = 1. \end{gathered} \tag{22}$$

Damit gilt für $A = (a_1, \ldots, a_n) \in M(n \times n, \mathbb{R})$ und entsprechendes B nach der Produktregel

$$\begin{aligned} D_A \det(B) &= \sum_{j=1}^{n} \det(a_1, \ldots, a_{j-1}, b_j, a_{j+1}, \ldots, a_n) \\ &= \sum_{j=1}^{n} \det(a_1, \ldots, a_{j-1}, \sum_{k=1}^{n} b_{kj} e_k, a_{j+1}, \ldots, a_n) \\ &= \sum_{k=1}^{n} \sum_{j=1}^{n} b_{kj} \underbrace{\det(a_1, \ldots, a_{j-1}, e_k, a_{j+1}, \ldots, a_n)}_{=:a^v_{jk}} \\ &= \sum_{k=1}^{n} \sum_{j=1}^{n} b_{kj} a^v_{jk}. \end{aligned}$$

Definiert man also die *Adjungte* $\operatorname{adj}(A)$ der Matrix A durch

$$\operatorname{adj}(A) := (a^v_{jk})_{j,k=1,\ldots,n} = (\det(a_1, \ldots, a_{j-1}, e_k, a_{j+1}, \ldots, a_n))_{j,k=1,\ldots,n},$$

so ist $D_A \det(B)$ gerade die Summe der Diagonalelemente der Matrix $B \operatorname{adj}(A)$, die sogenannte *Spur* dieser Matrix:

$$D_A \det(B) = \operatorname{Spur}(B \operatorname{adj}(A)).$$

Wir merken noch an:

1. Ist $A = E$, so ist nach (22)

$$e^v_{jk} = \det(e_1, \ldots, e_{j-1}, e_k, e_{j+1}, \ldots, e_n) = \delta_{jk},$$

also $\operatorname{adj}(E) = E$ und

$$\boxed{D_E \det(B) = \operatorname{Spur}(B).} \tag{23}$$

2. Allgemein gilt nach (22)

$$\begin{aligned}\sum_{k=1}^{n} a^v_{ik} a_{kj} &= \det(a_1, \ldots, a_{i-1}, \sum_{k=1}^{n} a_{kj} e_k, a_{i+1}, \ldots, a_n) \\ &= \det(a_1, \ldots, a_{i-1}, a_j, a_{i+1}, \ldots, a_n) = \delta_{ij} \det(A),\end{aligned}$$

also $\operatorname{adj}(A)A = \det(A)E$. Ist $\det A \neq 0$, so ist A also invertierbar und

$$\operatorname{adj}(A) = \det(A)A^{-1}.$$

In diesem Fall ist

$$\boxed{D_A \det(B) = \det(A) \operatorname{Spur}(BA^{-1}).} \tag{24}$$

□

Beispiel 124. Sei V ein endlich-dimensionaler Banachraum. Wir wollen zeigen

- $\mathbf{GL}(V) := \{A \in L(V,V) \mid A \text{ invertierbar}\}$ ist offen in $L(V,V)$,
- die Abbildung
$$\operatorname{inv} : \mathbf{GL}(V) \to L(V,V), \quad A \mapsto A^{-1}$$
ist differenzierbar und
- ihre Ableitung ist
$$\boxed{D_A \operatorname{inv}(B) = -A^{-1}BA^{-1}.} \tag{25}$$

Beachten Sie: Für $V = \mathbb{R}$ ist

$$\lambda(x^{-1})' = D_x(x^{-1})(\lambda) = -x^{-1}\lambda x^{-1} = \lambda(-x^{-2})$$

genau die aus der Schule bekannte Formel für die Ableitung von $\frac{1}{x}$.

Wir benutzen auf $L(V,V)$ die Operatornorm und die Ungleichung

$$\|AB\| \leq \|A\| \, \|B\| \tag{26}$$

für die Norm der Komposition von A und B (Beweis?).

Sei $A \in \mathbf{GL}(V)$ und $B \in L(V,V)$ mit

$$\|A - B\| < \frac{1}{\|A^{-1}\|}. \tag{27}$$

Aus (26) folgt dann also

$$\|A^{-1}(A - B)\| < 1. \tag{28}$$

Nun benutzen wir

$$B = A - (A - B) = A(E - A^{-1}(A - B)),$$

wobei E die identische Abbildung von V ist, und denken an die geometrische Reihe. Wir definieren

$$S_n := \left(\sum_{k=0}^{n} \left(A^{-1}(A-B) \right)^k \right) A^{-1}.$$

Aus

$$\|S_{n+m} - S_n\| = \left\| \left(\sum_{k=n+1}^{n+m} \left(A^{-1}(A-B) \right)^k \right) A^{-1} \right\| \leq \left(\sum_{k=n+1}^{n+m} \left\| A^{-1}(A-B) \right\|^k \right) \|A^{-1}\|$$

folgt mit (28), dass die S_n eine Cauchyfolge bilden. Weil $L(V,V)$ endlich-dimensional, also ein Banachraum ist, existiert $S := \lim S_n$. Aus

$$\begin{aligned} S_n A &= E + A^{-1}(A-B) + \ldots + (A^{-1}(A-B))^n \\ S_n A(A^{-1}(A-B)) &= \quad A^{-1}(A-B) + \ldots + (A^{-1}(A-B))^n + (A^{-1}(A-B))^{n+1} \end{aligned}$$

folgt durch Subtraktion:

$$S_n B = E - (A^{-1}(A-B))^{n+1}.$$

Der letzte Term geht für $n \to \infty$ gegen Null, also ist $SB = E$, d.h. $S = B^{-1}$. Verschärft man (27) zu

$$\|A-B\| \leq \frac{1}{2\|A^{-1}\|}, \tag{29}$$

so folgt mit der Dreiecksungleichung und (26)

$$\|S_n\| \leq \sum_{k=0}^{n} \|A^{-1}(A-B)\|^k \|A^{-1}\| \leq \frac{1}{1-\frac{1}{2}} \|A^{-1}\| = 2\|A^{-1}\|.$$

Aus (29) folgt also $\|B^{-1}\| \leq 2\|A^{-1}\|$ und damit

$$\|B^{-1} - A^{-1}\| = \|B^{-1}(A-B)A^{-1}\| \leq \|B^{-1}\|\|A-B\|\|A^{-1}\| \leq 2\|A^{-1}\|^2 \|A-B\|.$$

Das impliziert die Stetigkeit von inv. Schließlich untersuchen wir den Restterm

$$\begin{aligned} R(B) &= \mathrm{inv}(B) - \mathrm{inv}(A) + A^{-1}(B-A)A^{-1} = B^{-1} - A^{-1} + A^{-1}(B-A)A^{-1} \\ &= -A^{-1}(B-A)B^{-1} + A^{-1}(B-A)A^{-1} = A^{-1}(B-A)(A^{-1} - B^{-1}). \end{aligned}$$

Dann ist nach (26)

$$\frac{\|R(B)\|}{\|B-A\|} \leq \|A^{-1}\|\|A^{-1} - B^{-1}\|.$$

Wegen der Stetigkeit von inv geht das für $B \to A$ gegen 0 und inv ist differenzierbar mit der angegebenen Ableitung.

□

Im vorstehenden Beispiel haben wir eigentlich nur benutzt, dass $L(V,V)$ ein Banachraum mit einer "Multiplikation" AB ist, für die (26) gilt, eine sogenannte *Banachalgebra*. Dass es sich bei den Elementen um lineare Abbildungen handelt, spielte keine Rolle: Wir haben einen Satz aus der Theorie der Banachalgebren bewiesen.

2.3 Richtungsableitungen, partielle Ableitungen

- Nun endlich die Differentiation für bescheidenere Ansprüche!
- Richtungs- und insbesondere partielle Ableitungen kann man mit Methoden der Analysis I ausrechnen.
- Aber der Zusammenhang mit der (totalen) Differenzierbarkeit ist nicht ganz trivial.

Seien V, W endlich-dimensionale Banachräume, $G \subset V$ offen, $p \in G$ und $f : V \supset G \to W$ eine Abbildung.

Definition 125. (i) Für $v \in V$ definiere die *Richtungsableitung* von f in p in Richtung v durch

$$\partial_v f(p) := \lim_{t \to 0} \frac{f(p+tv) - f(p)}{t},$$

falls dieser Limes existiert.

(ii) Ist $V = \mathbb{R}^n$ mit der kanonischen Basis $e_1, \ldots, e_n$, so nennt man die Richtungsableitungen $\partial_{e_i}(p)$ die *partiellen Ableitungen* von f in p.

Notation:

$$\frac{\partial f}{\partial x_i}(p) = \partial_i f(p) = \partial_{e_i} f(p).$$

Statt x_i auch andere Variablennamen.

Es gilt also

$$\partial_i f(p) = \lim_{t \to 0} \frac{f(p_1, \ldots, p_i + t, \ldots, p_n) - f(p_1, \ldots, p_n)}{t}$$

Das ist (für $W = \mathbb{R}$) die Ableitung von f nach der i-ten Variablen im Sinne der Analysis I.

Beispiel 126. Ist f in p differenzierbar, so existieren dort alle Richtungsableitungen und es gilt

$$\boxed{\partial_v f(p) = D_p f(v).}$$

Speziell gilt also im Fall $V = \mathbb{R}^n$

$$\boxed{\partial_i f(p) = D_p f(e_i).}$$

□

Beispiel 127 (Funktionalmatrix zu Fuß). Ist weiter

$$f : \mathbb{R}^n \supset G \to \mathbb{R}^m,$$

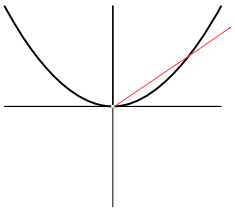
differenzierbar, so ist nach Beispiel 114 die Darstellungsmatrix von $D_p f$, also die Funktionalmatrix, mit Methoden der Analysis I zu berechnen:

$$\boxed{f'(p) = (\partial_j f_i(p))_{\substack{i=1,\ldots,m \\ j=1,\ldots,n}} = \left(\frac{\partial f_i}{\partial x_j}(p)\right)_{\substack{i=1,\ldots,m \\ j=1,\ldots,n}}.}$$

□

Beispiel 128. Vgl. Beispiel 69.

Existieren in p alle Richtungsableitungen, so muß f in p nicht differenzierbar, ja nicht einmal stetig sein, wie man an



$$f(x,y) := \begin{cases} 1 & \text{falls } y = x^2 > 0 \\ 0 & \text{sonst.} \end{cases}$$

sieht.

□

Satz 129 (Differenzierbarkeit und partielle Differenzierbarkeit). *Existieren auf ganz G alle Richtungsableitungen (oder im Fall $V = \mathbb{R}^n$ auch nur alle partiellen Ableitungen) und sind diese stetig, so ist f in G differenzierbar.*

Dieser Satz ist ein überaus nützliches Kriterium, weil oft die Berechnung von partiellen Ableitungen nach *Analysis I* sehr einfach und die Stetigkeit der Ableitungen offensichtlich ist.

Beweis. Wir führen den Beweis nur für $V = \mathbb{R}^n, W = \mathbb{R}$. Mittels komponentenweiser Differentiation bzw. partieller Differentiation folgt daraus der Satz für $V = \mathbb{R}^n, W = \mathbb{R}^m$. Sind schließlich $\Phi : \mathbb{R}^n \to V$ und $\Psi : \mathbb{R}^m \to W$ Isomorphismen und setzt man $\tilde{f} := \Psi \circ f \circ \Phi^{-1}$, so ist f genau dann differenzierbar bzw. partiell differenzierbar, wenn das entsprechende für $\tilde{f}$ gilt. Daraus folgt der Satz dann für beliebige V, W.

Seien also $V = \mathbb{R}^n, W = \mathbb{R}$. Für $p \in G$ definiere

$$F_p(x_1, \ldots, x_n) := \sum_{j=0}^{n} x_j \partial_j f(p).$$

Dann ist $F_p : \mathbb{R}^n \to \mathbb{R}$ linear und der offensichtliche Kandidat für die Ableitung an der Stelle p. Wir betrachten eine offene ϵ-Kugel $U = U_\epsilon(p)$, die ganz in der offenen Menge G liegt, und beschränken uns im folgenden auf $x \in U$. Beachten Sie, dass dann auch die Punkte $(p_1, \ldots, p_j, x_{j+1}, \ldots, x_n)$ in U und damit im Definitionsbereich von f liegen. Für $x \in U$ gilt daher

$$\begin{aligned} f(x) - f(p) &= f(x_1, \ldots, x_n) - f(p_1, \ldots, p_n) \\ &= f(x_1, \ldots, x_n) - f(p_1, x_2, \ldots, x_n) \\ &\quad + f(p_1, x_2, \ldots, x_n) - f(p_1, p_2, x_3, \ldots, x_n) \\ &\quad \vdots \\ &\quad + f(p_1, \ldots, p_{n-1}, x_n) - f(p_1, \ldots, p_n). \end{aligned}$$

Wir wenden auf jede Zeile den Mittelwertsatz an.

$$\begin{aligned} f(x) - f(p) = \partial_1 f(\xi_1, x_2, \ldots, x_n)(x_1 - p_1) \\ + \partial_2 f(p_1, \xi_2, \ldots, x_n)(x_2 - p_2) \\ \vdots \\ + \partial_n f(p_1, \ldots, p_{n-1}, \xi_n)(x_n - p_n) \end{aligned}$$

mit ξ_i zwischen x_i und p_i. Daraus folgt

$$\frac{f(x)-f(p)-F_p(x-p)}{\|x-p\|}=\sum_{i=1}^{n}\underbrace{\frac{x_i-p_i}{\|x-p\|}}_{beschränkt}\underbrace{(\partial_i f(p_1,\ldots,\xi_i,\ldots,x_n)-\partial_i f(p))}_{\to 0}$$

und mit der Stetigkeit der partiellen Ableitungen die Behauptung. □

Beispiel 130. Die Abbildung $f:\mathbb{R}^2\to\mathbb{R}^3$ mit

$$f(x,y)=(\sin x\cos y,\sin x\sin y,\cos y)$$

hat die folgende Matrix partieller Ableitungen:

$$(\partial_j f_i(x,y))=\begin{pmatrix}\cos x\cos y & -\sin x\sin y\\ \cos x\sin y & \sin x\cos y\\ 0 & -\sin y\end{pmatrix}.$$

Die partiellen Ableitungen sind offensichtlich stetig. Daher ist die Funktion differenzierbar und das Differential $D_{(x,y)}f:\mathbb{R}^2\to\mathbb{R}^3$ wird durch die obige Matrix $f'(x,y)$ repräsentiert.

□

Beispiel 131 (Das „totale Differential“). Die Koordinaten-Abbildungen

$$x_i:\mathbb{R}^n\to\mathbb{R},(v_1,\ldots,v_n)\mapsto v_i$$

sind linear. Deshalb ist für alle p in $\mathbb{R}^n$

$$D_p x_i=x_i.$$

Jedes $v\in\mathbb{R}^n$ läßt sich schreiben als

$$v=\sum x_i(v)e_i=\sum D_p x_i(v)e_i.$$

Ist f in $p\in G\subset\mathbb{R}^n$ differenzierbar, so folgt

$$D_p f(v)=D_p f(\sum D_p x_i(v)e_i)=\sum D_p x_i(v)D_p f(e_i)=\sum \partial_i f(p)D_p x_i(v).$$

Das schreibt man auch so:

$$Df=\sum \partial_i f\,Dx_i \tag{30}$$

oder - gebräuchlicher -

$$df=\sum\frac{\partial f}{\partial x_i}dx_i.$$

Man nennt diesen Ausdruck das „totale Differential“ von f im Gegensatz zu den einzelnen partiellen Differentialen $\frac{\partial f}{\partial x_i}$. Bei Lichte besehen ist das totale Differential an der Stelle p aber einfach nur die Ableitung $D_p f$.

□

Beispiel 132 (Kettenregel in partiellen Ableitungen). Für differenzierbare Abbildungen zwischen den Standardräumen sieht die Kettenregel in partiellen Ableitungen folgendermaßen aus:

Aus $D_p(f\circ g)=D_{g(p)}f\circ D_p g$ folgt nach linearer Algebra

$$(f\circ g)'(p)=f'(g(p))g'(p),$$

wobei rechts das Produkt der Matrizen steht. Also, wenn wir die Variablen im Definitionsraum von f mit y_j und im Definitionsraum von g mit x_k bezeichnen,

$$\left(\frac{\partial(f\circ g)_i}{\partial x_k}(p)\right) = \left(\frac{\partial f_i}{\partial y_j}(g(p))\right)\left(\frac{\partial g_j}{\partial x_k}(p)\right)$$

oder

$$\frac{\partial(f\circ g)_i}{\partial x_k} = \sum_j \frac{\partial f_i}{\partial y_j}(g(p))\frac{\partial g_j}{\partial x_k}(p).$$

Abgekürzte Notation:

$$\frac{\partial f_i}{\partial x_k} = \sum_j \frac{\partial f_i}{\partial y_j}\frac{\partial y_j}{\partial x_k}.$$

□

2.4 Höhere Ableitungen

- Die Ableitung einer Funktion von mehreren Variablen ist nicht eine Zahl, sondern eine Lineare Abbildung. Entsprechend werden erst recht die höheren Ableitungen solcher Funktionen kompliziertere Gebilde, nämlich multilineare Abbildungen.
- Wir lernen, wie man sie "trotzdem" effektiv berechnen kann.
- Wir lernen den Satz von Schwarz über die Symmetrie höherer Ableitungen, der manche Rechnung vereinfacht, aber auch wichtige Anwendungen auf Differentialgleichungen hat.

Vorbemerkung. Wir erinnern daran, dass $L(V,W)$ den Vektorraum aller linearen Abbildungen von V nach W bezeichnet. Sind V und W endlich-dimensional, so ist auch $L(V,W)$ endlich-dimensional und es gilt $\dim L(V,W) = (\dim V)(\dim W)$, vgl. Lineare Algebra.

Definition 133 (Zweite Ableitung). Sei $f : V \subset G \to W$ auf G differenzierbar. Dann ist

$$Df : G \to L(V,W),\ p \mapsto D_p f.$$

Ist diese Abbildung differenzierbar in $p \in G$, so heißt f in p zweimal differenzierbar und

$$D_p^2 f := D_p(Df) : V \to L(V,W)$$

die zweite Ableitung von f in p.

Wir haben also für $v_1, v_2 \in V$

$$\begin{aligned} D_p^2 f(v_1) &\in L(V,W), \\ D_p^2 f(v_1)(v_2) &\in W. \end{aligned}$$

Beispiel 134. Sei $(V, \langle .,. \rangle)$ ein Euklidischer Vektorraum und sei $r : V \setminus \{0\} \to \mathbb{R}$ gegeben durch $r(x) = \sqrt{< x, x >}$. Wir haben im Beispiel 122 ausgerechnet, dass

$$D_p r(v) = \frac{1}{r(p)} < p, v >,$$

also

$$Dr : x \mapsto D_x r = \frac{1}{r(x)} < x, . > \in L(V, \mathbb{R}).$$

Das ist das Produkt der Abbildung $\frac{1}{r} : V \setminus \{0\} \to \mathbb{R}$ mit der Abbildung

$$\begin{array}{ccc} g : V & \to & L(V,\mathbb{R}) \\ x & \mapsto & \langle x, . \rangle \end{array}$$

1. Faktor: Weil r differenzierbar ist, ist nach der Kettenregel auch $\frac{1}{r}$ differenzierbar, und es ist

$$D_p \frac{1}{r}(v) = -\frac{1}{r^2(p)} D_p r(v) = -\frac{1}{r^3(p)} < p, v > .$$

2. Faktor: Die Abbildung $g : V \to L(V,\mathbb{R})$ ist linear, also auch differenzierbar und

$$D_p g(v) = g(v) = \langle v, . \rangle .$$

Produktregel: Daher ist Dr nach der Produktregel differenzierbar, und es gilt

$$\begin{aligned} D_p(Dr)(v) &= -\frac{1}{r^3(p)} < p, v >< p, . > + \frac{1}{r(p)} < v, . > \\ &= \frac{1}{r^3(p)} \left(r^2(p) < v, . > - < p, v >< p, . > \right) \\ &= \frac{1}{r^3(p)} \left(< p, p >< v, . > - < p, v >< p, . > \right), \end{aligned}$$

also

$$D_p^2 r(v)(w) = \frac{1}{r^3(p)} \left(< p, p >< v, w > - < p, v >< p, w > \right).$$

□

Es ist klar, wie man **höhere als 2. Ableitungen** definiert. Dabei entsteht allerdings ein kleines Problem: Wir erhalten $D_p^3 f \in L(V, L(V, L(V, W)))$, und den Zielraum der 7. Ableitung mag man nicht mehr hinschreiben. Dieses Problem vermeiden wir folgendermaßen:

Wir definieren

$$D_p^2 f(v, w) := D_p^2(v)(w).$$

Dann ist

$$D_p^2 f : V \times V \to W$$

eine *bilineare Abbildung* von V nach W. Im obigen Beispiel ist also

$$D_p^2 r(v, w) := \frac{1}{r^3(p)} \left(< p, p >< v, w > - < p, v >< p, w > \right). \tag{31}$$

Bezeichnen wir mit $L^k(V, W)$ den Vektorraum der k-linearen Abbildungen von V nach W, so haben wir allgemeiner einen kanonischen Isomorphismus

$$j_k : L(V, L^{k-1}(V, W)) \to L^k(V, W)$$

mit

$$j_k(\Phi)(v_1, \ldots, v_k) := \Phi(v_1)(v_2, \ldots, v_k).$$

(Beweisen Sie das! Es folgt, dass $\dim L^k(V, W) = (\dim V)^k (\dim W) < \infty$, wenn V und W endlich-dimensional sind.)

Damit definieren wir induktiv die k-Ableitung $D_p^k f$ einer Funktion an der Stelle p wir folgt:

Definition 135 (Höhere Ableitungen). Ist $f : V \supset G \to W$ bereits $(k-1)$ mal differenzierbar und ist die $(k-1)$-te Ableitung

$$D^{k-1} f : G \to L^{k-1}(V, W), x \mapsto D_x^{k-1} f$$

in $p \in G$ differenzierbar, so heißt f in p k-mal differenzierbar und die k-te Ableitung in p ist gegeben durch

$$D_p^k f(v_1, \ldots, v_k) := j_k(D_p(D^{k-1} f))(v_1, \ldots, v_k) = D_p(D^{k-1} f)(v_1)(v_2, \ldots, v_k).$$

Die k-te Ableitung einer k-mal differenzierbaren Abbildung $f : V \supset G \to W$ an einer Stelle p ist also eine k-lineare Abbildung

$$\boxed{D_p^k f \in L^k(V, W).}$$

Lemma 136. *Ist f in p k-mal differenzierbar so gilt für $v_1, \dots, v_k \in V$*

$$D_p^k f(v_1, \dots, v_k) = \partial_{v_1} \dots \partial_{v_k} f(p).$$

Insbesondere existiert die rechte Seite.

Beweis. Durch Induktion über k.

$\underline{k = 1.}$ $D_p f(v) = \partial_v f(p)$ wissen wir schon.

$\underline{(k-1) \to k.}$ Die Abbildung

$$\begin{array}{rccc} g: & L^{k-1}(V,W) & \to & W \\ & \mu & \mapsto & \mu(v_2, \dots, v_k). \end{array}$$

ist linear. Nach Kettenregel und Voraussetzung ist daher $g \circ D^{k-1} f$ differenzierbar mit

$$\begin{aligned} D_p(g \circ D^{k-1} f)(v_1) &= g\left(D_p(D^{k-1} f)(v_1)\right) = D_p(D^{k-1} f)(v_1)(v_2, \dots, v_k) \\ &= D_p^k f(v_1, \dots, v_k). \end{aligned}$$

Andrerseits ist nach Induktionsvoraussetzung

$$g \circ D^{k-1} f = D^{k-1} f(v_2, \dots, v_k) = \partial_{v_2} \dots \partial_{v_k} f$$

und deshalb nach dem Fall $k = 1$

$$D_p(g \circ D^{k-1} f)(v_1) = \partial_{v_1} \dots \partial_{v_k} f.$$

□

Dieses Lemma impliziert insbesondere folgende Vereinfachung: Um $D_p^2 f(v_1, v_2)$ zu berechnen, muß ich nicht die schwerer vorzustellende Abbildung $Df : G \to L(V,W)$ differenzieren, sondern ich kann $Df(v_2) : G \to W$ in Richtung v_1 differenzieren: Ich darf *vor* der zweiten Ableitung den Vektor v_2 einsetzen.

$$D_p^2 f(v_1, v_2) = \partial_{v_1} \partial_{v_2} f(p) = D_p(Df(v_2))(v_1).$$

Dabei muß man auf die Reihenfolge der Vektoren achten – bis wir gleich gezeigt haben, dass sie keine Rolle spielt!

Beispiel 137 (Höhere Ableitungen auf dem $\mathbb{R}^n$). Ist $f : \mathbb{R}^n \supset G \to W$ in $p \in G$ k-mal differenzierbar, und hat man k Vektoren

$$v_j = (v_{1j}, \dots, v_{nj}) \in \mathbb{R}^n, \quad j \in \{1, \dots, k\},$$

gegeben, so gilt

$$\boxed{D_p^k f(v_1, \dots, v_k) = \sum_{i_1, \dots, i_k = 1}^{n} \partial_{i_1} \dots \partial_{i_k} f(p) v_{i_1 1} \dots v_{i_k k}.} \tag{32}$$

Also läßt sich die k-te Ableitung von f mittels k-facher partieller Ableitungen ausrechnen.

Konkret betrachen wir die Normfunktion $r(x) = \sqrt{\sum_{i=1}^n x_i^2}$ auf $\mathbb{R}^n \setminus \{0\}$. Wir finden

$$\partial_j r = \frac{x_j}{r}, \quad \partial_i \partial_j r = \frac{r\delta_{ij} - x_j \frac{x_i}{r}}{r^2} = \frac{1}{r^3}\left(r^2 \delta_{ij} - x_i x_j\right)$$

und damit

$$D_p^2 r(v,w) = \frac{1}{r^3(p)} \left(r^2(p) \sum_{i=1}^n v_i w_i - \left(\sum_{i=1}^n p_i v_i \right) \left(\sum_{j=1}^n p_j w_j \right) \right).$$

Vergleichen Sie das mit (31).

□

Für spätere Verwendung zeigen wir hier noch das folgende

Lemma 138. *Für k-mal differenzierbares $f : V \supset G \to W$, $p \in G$ und $v_1, \ldots, v_k \in V$ gilt*

$$D_p^k f(v_1, \ldots, v_k) = D_p^{k-1}(Df)(v_2, \ldots, v_k)(v_1).$$

Beweis. Ich beweise das für $V = \mathbb{R}^n$ mit den Koordinatenprojektionen $x_j : \mathbb{R}^n \to \mathbb{R}$. Der allgemeine Fall geht nach Wahl einer Basis genauso, nur treten an die Stelle der x_j dann die dualen Basisvektoren. Es ist

$$\begin{aligned} D_p^{k-1}(Df)(v_2, \ldots, v_k)(v_1) &= D_p^{k-1}(\sum_{j=1}^n (\partial_j f)\, x_j)(v_2, \ldots, v_k)(v_1) \\ &= \sum_{i_2, \ldots, i_k, j=1}^n \partial_{i_2} \ldots \partial_{i_k} \partial_j f(p) v_{i_2 2} \ldots v_{i_k k} x_j(v_1) \\ &= \sum_{j, i_2, \ldots, i_k=1}^n \partial_{i_2} \ldots \partial_{i_k} \partial_j f(p) v_{j1} v_{i_2 2} \ldots v_{i_k k}. \end{aligned}$$

Vergleich mit (32) liefert die Behauptung. □

Für die Frage, **ob f zweimal differenzierbar** ist, gibt es ebenfalls ein gutes Kriterium mittels partieller Ableitungen: Sei f differenzierbar (z.B. weil es überall stetige partielle Ableitungen besitzt). Nach (30) ist Df wegen der Konstanz der Abbildungen $p \mapsto D_p x_i = x_i$ genau dann differenzierbar, wenn die partiellen Ableitungen $\partial_i f$ alle differenzierbar sind. Das läßt sich wieder mittels partieller Ableitungen testen, und man erhält: Existieren alle zweiten partiellen Ableitungen von f auf G und sind sie dort stetig, so ist f zweimal differenzierbar.

Entsprechendes gilt für höhere Ableitungen.

Definition 139 (C^k-Funktionen). Ist $f : V \supset G \to W$ k-mal differenzierbar und die Abbildung

$$D^k f : G \to L^k(V, W), \quad p \mapsto D_p^k f$$

stetig, so heißt f *k-mal stetig differenzierbar*. Nach der vorstehenden Bemerkung ist das für $V = \mathbb{R}^n$ äquivalent dazu, dass alle partiellen Ableitungen k-ter Ordnung von f existieren und stetig sind. Wir schreiben dafür

$$f \in C^k(G, W) \text{ oder kurz } f \in C^k.$$

Schließlich soll $f \in C^\infty$ bedeuten, dass f beliebig oft differenzierbar ist – die Stetigkeit der Ableitungen folgt dann natürlich von selbst.

Ist f in p zweimal differenzierbar und sind u, v „hinreichend kleine“ Vektoren, so gilt

$$f(p+u+v) - f(p+u) - f(p+v) + f(p) \approx D_{p+u} f(v) - D_p f(v) \approx D_p^2 f(u, v).$$

Die linke Seite ist also eine Approximation für die 2. Ableitung, die zum Beispiel in der diskreten Mathematik wichtig ist. Wir berechnen damit die 2. Ableitung:

Lemma 140. *Sei $f : V \supset G \to W$ differenzierbar und in $p \in G$ zweimal differenzierbar. Dann gilt für alle $u, v \in V$*

$$\boxed{D_p^2 f(u,v) = \lim_{t\to 0} \frac{f(p+tu+tv) - f(p+tu) - f(p+tv) + f(p)}{t^2}.}$$

Weil der Zähler rechts in u und v symmetrisch ist, folgt daraus der wichtige

Satz 141 (H.A. Schwarz). *Sei $f : V \supset G \to W$ in G 2-mal differenzierbar. Dann gilt für alle $u, v \in V$ und $p \in G$*

$$D_p^2 f(u,v) = D_p^2 f(v,u).$$

Beweis des Lemmas 140. 1. Schritt. Es genügt der Beweis für den Fall $W = \mathbb{R}$, weil $f = \sum f_i e_i$, wo die e_i eine Basis von W und die f_i reellwertige Funktionen sind.

2. Schritt. In Anlehnung an die heuristische Betrachtung oben definieren wir eine Funktion

$$F(t) := f(p+u+tv) - f(p+tv).$$

Dabei seien $\|u\|$ und $\|v\|$ hinreichend klein, so dass $p+u+tv$ und $p+tv$ für $0 \le t \le 1$ in G liegen. Dann ist nach dem Mittelwertsatz für ein $\tau \in]0,1[$

$$\begin{aligned} f(p+u+v) - f(p+u) - f(p+v) + f(p) &= F(1) - F(0) \\ &= F'(\tau) \\ &= D_{p+u+\tau v} f(v) - D_{p+\tau v} f(v) \\ &= (D_{p+u+\tau v} f - D_p f)(v) - (D_{p+\tau v} f - D_p f)(v) \\ &=: (*) \end{aligned}$$

Wir wenden jetzt auf die Funktion Df die Definition der Differenzierbarkeit an der Stelle p an, und erhalten für $x = p+u+\tau v$ bzw. $x = p+\tau v$

$$\begin{aligned} (*) = &= D_p^2 f(u+\tau v)(v) + R(p+u+\tau v)(v) \\ &\quad - D_p^2 f(\tau v)(v) - R(p+\tau v)(v) \\ &= D_p^2 f(u,v) + (R(p+u+\tau v) - R(p+\tau v))(v). \end{aligned}$$

Zu gegebenem $\epsilon > 0$ gibt es ein $\delta > 0$, so dass

$$\|R(x)\| \le \epsilon \|x-p\|, \text{ falls } \|x-p\| < \delta. \tag{33}$$

Beachten Sie, dass $R(x) \in L(V,W)$. Für die Norm von $R(x)$ verwenden wir daher wie üblich die Operatornorm auf $L(V,W)$.

Für $\|u\| + \|v\| < \delta$ ist dann

$$\begin{aligned} \|R(p+u+\tau v)\| \le \epsilon \|u+\tau v\| \le \epsilon(\|u\| + \|v\|), \\ \|R(p+\tau v)\| \le \epsilon \|\tau v\| \le \epsilon(\|u\| + \|v\|), \end{aligned}$$

also

$$\|(*) - D_p^2 f(u,v)\| = \|(R(p+u+\tau v) - R(p+\tau v))(v)\| \le 2\epsilon(\|u\| + \|v\|)\|v\|.$$

Nun seien $u, v \in V$ beliebig und $t_0 > 0$ bei vorgegebenem $\epsilon > 0$ so klein gewählt, dass $\|t_0 u\| + \|t_0 v\| < \delta$ ist, vgl. (33). Dann folgt für alle t mit $|t| \leq t_0$

$$\begin{aligned}
&\left\| \frac{f(p+tu+tv) - f(p+tu) - f(p+tv) + f(p)}{t^2} - D_p^2 f(u,v) \right\| \\
&= \left\| \frac{f(p+tu+tv) - f(p+tu) - f(p+tv) + f(p) - D_p^2 f(tu,tv)}{t^2} \right\| \\
&\leq 2\epsilon \frac{(\|tu\| + \|tv\|)\|tv\|}{t^2} = 2\epsilon(\|u\| + \|v\|)\|v\|
\end{aligned}$$

Daraus folgt die Behauptung. □

Korollar 142 (zum Satz von Schwarz). *Ist $f : \mathbb{R}^n \supset G \to W$ in G 2-mal differenzierbar (oder 2-mal partiell differenzierbar mit stetigen zweiten partiellen Ableitungen), so gilt für alle i, j*

$$\partial_i \partial_j f = \partial_j \partial_i f.$$

Beispiel 143 (Wichtige Anwendung: Integrabilitätskriterium). Seien $f_1, \ldots, f_n : \mathbb{R}^n \supset G \to \mathbb{R}$. Die elementarste Frage der Theorie der partiellen Differentialgleichungen ist die, ob es eine Funktion $y : G \to \mathbb{R}$ gibt, so dass für alle i gilt:

$$\partial_i y = f_i. \tag{34}$$

Haben die f_i stetige partielle Ableitungen, so hat ein solches y, falls es existiert, stetige partielle Ableitungen bis zur Ordnung 2. Also ist eine *notwendige* Bedingung für die Lösbarkeit von (34), dass

$$\partial_j f_i = \partial_j \partial_i y = \partial_i \partial_j y = \partial_i f_j,$$

d.h.

$$\partial_j f_i = \partial_i f_j \quad \text{für alle } i, j.$$

□

Aus dem Satz von Schwarz in Verbindung mit dem Lemma 136 ergibt sich unmittelbar die folgende Verallgemeinerung:

Korollar 144 (zum Satz von Schwarz). *Ist $f : V \supset G \to W$ k-mal differenzierbar in G, so gilt für jede Permutation $(i_1, \ldots, i_k)$ von $(1, \ldots, k)$ und für alle $p \in G$ und $v_1, \ldots, v_k \in V$*

$$D_p^k f(v_1, \ldots, v_k) = D_p^k f(v_{i_1}, \ldots, v_{i_k})$$

bzw.

$$\partial_{v_1} \ldots \partial_{v_k} f(p) = \partial_{v_{i_1}} \ldots \partial_{v_{i_k}} f(p).$$

Bemerkungen. In der Literatur (z.B. im Buch von Rudin) findet man den Satz von Schwarz häufig in folgender Form: *Existieren alle partiellen Ableitungen 2. Ordnung von f und sind sie stetig, so gilt*

$$\partial_i \partial_j f = \partial_j \partial_i f.$$

Aus dem Vergleich mit unserer Version ergeben sich zwei Fragen:

1. Gibt es Funktionen, die zweimal differenzierbar, aber nicht zweimal stetig differenzierbar sind? Dann ist die Version des Satzes 141 stärker als die oben zitierte.

   Antwort: Ja. Die Funktion

   $$g : \mathbb{R} \to \mathbb{R}, \quad x \mapsto x^4 \sin \frac{1}{x}$$

   ist auf $\mathbb{R}$ zweimal differenzierbar, aber die 2. Ableitung ist in 0 nicht stetig. Entsprechendes gilt dann auch für die durch $f(x, y) := g(x)$ definierte Funktion auf $\mathbb{R}^2$.

2. Gibt es Funktionen mit (nicht stetigen) partiellen Ableitungen 2. Ordnung, für die der Satz von Schwarz nicht gilt?

   Antwort: Ja. Die Funktion $f : \mathbb{R}^2 \to \mathbb{R}$ mit

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

   besitzt überall stetige 1. partielle Ableitungen und (in 0 unstetige) 2. partielle Ableitungen. Für sie gilt

   $$\partial_1 \partial_2 f(0, 0) = 1 \neq -1 = \partial_2 \partial_1 f(0, 0).$$

Beweise als Übung.

2.5 Schrankensatz, Satz von Taylor

- Schrankensatz und Satz von Taylor kann man auch für Funktionen mehrerer Variablen formulieren.

In höherdimensionalen Räumen wird die naive Verallgemeinerung des Mittelwertsatzes

$$f(b) - f(a) \stackrel{?}{=} D_\xi f(b-a) \text{ für geeignetes } \xi \text{ zwischen(?) } a \text{ und } b$$

falsch. Zum Beispiel gilt für die Funktion $f : [0, 2\pi] \to \mathbb{R}^3$ mit $f(t) = (\cos t, \sin t, t)$, dass

$$f(2\pi) - f(0) = (0, 0, 2\pi).$$

Aber für alle $\xi \in]0, 2\pi[$ ist $D_\xi f(2\pi - 0) = 2\pi(-\sin t, \cos t, 1) \neq (0, 0, 2\pi)$.

Der Schrankensatz, den wir jetzt beweisen, liefert einen Ersatz für den Mittelwertsatz.

Satz 145 (Schrankensatz). *Seien V, W endlich-dimensionale Banachräume und sei*

$$f : V \supset G \to W$$

eine differenzierbare Abbildung. Seien $a, b \in G$, so dass die Strecke

$$\overline{ab} := \{a + t(b-a) \,|\, 0 \leq t \leq 1\}$$

in G enthalten ist. Dann gilt

$$\|f(b) - f(a)\| \leq \sup_{x \in \overline{ab}} \|D_x f\| \, \|b - a\|.$$

Zusatz. Wenn f in den Endpunkten von $\overline{ab}$ nicht differenzierbar, aber stetig ist, gilt dieselbe Behauptung, wobei das Supremum über alle Punkte von $\overline{ab} \setminus \{a, b\}$ zu nehmen ist.

Beweis. Sei $K := \sup_{x \in \overline{ab}} \|D_x f\|$ und sei $\epsilon > 0$. Sei o.E. $K < \infty$, sonst ist nichts zu zeigen. Wir definieren

$$A := \{t \in [0,1] \,|\, \|f(a + t(b-a)) - f(a)\| \leq t(K + \epsilon)\|b - a\|\}.$$

Weil beide Seiten der Ungleichung in der Definition von A in $t \in [0,1]$ stetig sind, ist A abgeschlossen. Wegen $0 \in A$ ist $A \neq \emptyset$. Insbesondere gilt

$$\sup A =: s \in A \subset [0,1].$$

Wir zeigen $s = 1$, d.h.

$$\|f(b) - f(a)\| \leq (K + \epsilon)\|b - a\|.$$

Weil das für alle $\epsilon > 0$ gilt, folgt daraus die Behauptung.

Annahme: $s < 1$. Die Funktion f ist in $p = a + s(b-a)$ differenzierbar, und wir haben

$$f(a + t(b-a)) = f(a + s(b-a)) + D_p f((t-s)(b-a)) + R(a + t(b-a))$$

mit $\lim_{x \to p} \frac{R(x)}{\|x-p\|} = 0$, also $\lim_{t \to s} \frac{R(a+t(b-a))}{|t-s| \cdot \|b-a\|} = 0$.

(Beachten Sie $\|(a + t(b-a)) - (a + s(b-a))\| = \|(t-s)(b-a)\| = |t-s|\|b-a\|$.)

Daher gibt es $\delta > 0$, so dass für $s < t < s + \delta$ gilt

$$\|R(a + t(b - a))\| \leq \epsilon(t - s) \cdot \|b - a\|.$$

Aus der Dreiecksungleichung folgt für $s < t < s + \delta$

$$\begin{aligned}\|f(a + t(b - a)) - f(a + s(b - a))\| &\leq \|D_p f\|\,(t - s) \cdot \|b - a\| + \epsilon(t - s) \cdot \|b - a\| \\ &\leq (t - s)(K + \epsilon)\|b - a\|.\end{aligned}$$

und weiter

$$\begin{aligned}\|f(a + t(b - a)) - f(a)\| &\leq \|f(a + t(b - a)) - f(a + s(b - a))\| + \|f(a + s(b - a)) - f(a)\| \\ &\leq (t - s)(K + \epsilon)\|b - a\| + s(K + \epsilon)\|b - a\| \\ &= t(K + \epsilon)\|b - a\|.\end{aligned}$$

Das ist ein Widerspruch zur Definition von s. Also ist $s = 1$ und *(i)* bewiesen.

Zum Zusatz. Ist f nur im „Inneren“ der Strecke $\overline{ab}$ differenzierbar, so gilt nach dem Beweisenen für alle $0 < t_1 < t_2 < b$

$$\begin{aligned}\|f(a + t_2(b - a)) - f(a + t_1(b - a))\| &\leq \sup_{0<t<1} \|D_{a+t(b-a)}\|\,\|(a + t_2(b - a)) - (a + t_1(b - a))\| \\ &= \sup_{0<t<1} \|D_{a+t(b-a)}\|\,(t_2 - t_1)\|b - a\|.\end{aligned}$$

Durch Grenzübergang $t_1 \searrow 0$ und $t_2 \nearrow 1$ folgt mit der Stetigkeit von f die Behauptung. □

Berechnung der Operatornorm. Bisher hatte die Operatornorm nur eine Hilfsfunktion. Der Schrankensatz macht es wünschenswert, sie explizit zu berechnen. Das ist ein Problem der linearen Algebra. Wir geben die Resultate für zwei einfache Fälle.

1. $F : \mathbb{R}^n \to \mathbb{R}$ sei bezüglich der Standardbasen gegeben durch die $m \times 1$-Matrix $A = (a_1, \ldots, a_n)$. Dann ist die Operatornorm bezüglich der Euklidischen Norm von $\mathbb{R}^n$ und dem Betrag $|.|$ auf $\mathbb{R}$ gegeben durch

   $$\|A\| = \sqrt{\sum_{i=1}^{n} a_i^2}.$$

   Im Falle des Schrankensatzes, ist $F = Df$ gegeben durch die Matrix

   $$f' = (\partial_1 f, \ldots, \partial_n f)$$

   und

   $$\|Df\| = \sqrt{\sum_{i=1}^{n} (\partial_i f)^2}.$$

2. $f : \mathbb{R}^n \to \mathbb{R}^m$ sei bezüglich der Standardbasen gegeben durch die $m \times n$-Matrix A. Die Normen auf $\mathbb{R}^n$ und $\mathbb{R}^m$ seien die üblichen Euklidischen Normen. Mit A^T bezeichnen wir die transponierte Matrix. Dann ist $A^T A$ eine symmetrische (=selbstadjungierte) $n \times n$-Matrix und $\|A\|$ ist die Wurzel aus dem Maximum der Eigenwerte von $A^T A$.

Korollar 146. *Sei $G \subset V$ offen und zusammenhängend und sei $f : G \to \mathbb{R}$ differenzierbar mit $D_p f = 0$ für alle $p \in G$. Dann ist f konstant.*

Beweis. Sei $p \in G$ und

$$U := \{x \in G \,|\, f(x) = f(p)\}.$$

Ist $a \in U$, so gibt es $r > 0$ mit $U_r(a) \subset G$. Aus dem Schrankensatz folgt dann für alle $b \in U_r(a)$

$$\|f(b) - f(p)\| = \|f(b) - f(a)\| \leq 0 \cdot \|b - a\|,$$

also $b \in U$. Deshalb ist U offen. Andrerseits ist

$$G \setminus U = f^{-1}(\mathbb{R} \setminus \{f(p)\})$$

ebenfalls offen. Weil G zusammenhängend ist, folgt dass entweder $U = \emptyset$ oder $G \setminus U = \emptyset$. Wegen $p \in U$ folgt $U = G$. □

Für spätere Verwendung "verfeinern" wir dieses Korollar noch etwas:

Korollar 147. *Seien $G \subset V_1 \times V_2$ offen und $f : G \to W$ differenzierbar. Es gelte für alle $(p_1, p_2) \in G$, dass*

$$0 = D_{(p_1,p_2)} f(0, .) : V_2 \to W.$$

Offenbar gibt es zu jedem $(p_1, p_2) \in G$ ein $\epsilon > 0$ mit

$$U_\epsilon(p_1) \times U_\epsilon(p_2) \subset G.$$

Dafür gilt dann

$$f(q_1, q_2) = f(q_1, r_2) \text{ für alle } q_1 \in U_\epsilon(p_1) \text{ und } q_2, r_2 \in U_\epsilon(p_2).$$

Mit andern Worten: Ist die V_2-Ableitung von f Null, so hängt f lokal nicht von der V_2-Variablen ab.

Beweis. Betrachte die Funktion

$$g : V_2 \supset U_\epsilon(p_2) \to W, t \mapsto f(q_1, t).$$

Für diese gilt nach der Kettenregel angewendet auf $t \mapsto (q_1, t) \mapsto f(q_1, t)$

$$D_t g(v_2) = D_{(q_1,t)} f(0, v_2) = 0,$$

und aus dem Schrankensatz angewendet auf g folgt $g(q_2) = g(r_2)$ und damit die Behauptung. □

Satz 148 (Taylorformel). *Sei $f : V \supset G \to W$ n-mal differenzierbar.*
Dann läßt sich f darstellen als

$$f(x) = \left(\sum_{k=0}^{n} \frac{1}{k!} D_p^k f(\underbrace{x - p, \ldots, x - p}_{k-mal}) \right) + R(x), \tag{35}$$

wobei für die dadurch definierte Restfunktion R gilt

$$\lim_{x \to p} \frac{R(x)}{\|x - p\|^n} = 0.$$

Zusatz: *Ist f sogar $(n+1)$-mal differenzierbar und reellwertig(!) und ist $\overline{px} \subset G$, so gibt es $q \in \overline{px}$, so dass*

$$R(x) = \frac{1}{(n+1)!} D_q^{n+1} f(\underbrace{x - p, \ldots, x - p}_{(n+1)-mal}).$$

Beweis. Der Zusatz folgt direkt aus dem 1-dimensionalen Fall: Wir setzen $v := x - p$ und

$$g : [0,1] \to \mathbb{R}, t \mapsto g(t) := f(p + tv).$$

Dann gibt es $\tau \in [0,1]$, so dass

$$\begin{aligned} f(x) &= g(1) \\ &= \left(\sum_{k=0}^{n} \frac{1}{k!} g^{(k)}(0) \right) + \frac{1}{(n+1)!} g^{(n+1)}(\tau) \\ &= \left(\sum_{k=0}^{n} \frac{1}{k!} \underbrace{\partial_v \dots \partial_v}_{k-mal} f(p) \right) + \frac{1}{(n+1)!} \underbrace{\partial_v \dots \partial_v}_{(n+1)-mal} f(p + \tau v) \\ &= \sum_{k=0}^{n} \frac{1}{k!} D_p^k f(\underbrace{x-p, \dots, x-p}_{k-mal}) + \frac{1}{(n+1)!} D_q^{n+1} f(\underbrace{x-p, \dots, x-p}_{(n+1)-mal}) \end{aligned}$$

mit $q = p + \tau v$.

Nun zum Beweis der ersten Taylorformel. Die kann man nicht einfach auf den eindimensionalen Fall zurückführen. Die Mehrdimensionalität von W ist dabei sekundär. Aber ein Ansatz wie oben führt nur zu Informationen über $\lim_{t \to 0} \frac{R(p+tv)}{\|tv\|^n}$, und das ist eine deutlich eingeschränkte Aussage.

Beweis durch vollständige Induktion über n.

$\underline{n = 1.}$ Das ist einfach die Definition der Differenzierbarkeit.

$\underline{(n-1) \to n.}$ Die Induktionsvoraussetzung angewendet auf die $(n-1)$-mal differenzierbare Funktion Df liefert

$$D_x f = \left(\sum_{k=0}^{n-1} \frac{1}{k!} D_p^k (Df)(x-p, \dots, x-p) \right) + \tilde{R}(x) \tag{36}$$

mit

$$\lim_{x-p \to 0} \frac{\tilde{R}(x)}{\|x-p\|^{n-1}} = 0. \tag{37}$$

Nun berechnen wir die Ableitung von

$$R(x) = f(x) - \sum_{k=0}^{n} \frac{1}{k!} D_p^k (x-p, \dots, x-p).$$

Der erste Term unter dem Summenzeichen ist $f(p)$, fällt bei der Differentiation also weg. Unter Benutzung der Produktregel, des Satzes von Schwarz und des Lemmas 138

$$\begin{aligned} D_x R(v) &= D_x f(v) - \sum_{k=1}^{n} \frac{1}{k!} k D_p^k f(x-p, \dots, x-p, v) \\ &= D_x f(v) - \sum_{k=1}^{n} \frac{1}{(k-1)!} D_p^{k-1} (Df)(x-p, \dots, x-p)(v). \end{aligned}$$

Also ist

$$D_x R = D_x f - \sum_{k=0}^{n-1} \frac{1}{k!} D_p^k (Df)(x-p, \dots, x-p) = \tilde{R}(x).$$

Zu jedem $\epsilon > 0$ gibt es ein $\delta > 0$ mit $U_\delta(p) \subset G$, so dass für alle $x \in U_\delta(p)$

$$\|\tilde{R}(x)\| \leq \epsilon \|x - p\|^{n-1} \tag{38}$$

Wegen $R(p) = 0$ folgt aus dem Schrankensatz dann für alle $x \in U_\delta(p)$

$$\|R(x)\| = \|R(x) - R(p)\| \leq \sup_{y \in U_\delta(p)} \|D_y R\| \, \|x - p\| \leq \epsilon \|x - p\|^n,$$

also

$$\lim_{x \to p} \frac{R(x)}{\|x - p\|^n} = 0.$$

□

2.6 Lokale Extrema

- Wir wenden die Taylorformel auf Extremalprobleme an.

Definition 149. Sei $l \in L^k(V, \mathbb{R})$ eine k-lineare Abbildung. Wenn der Zielraum $\mathbb{R}$ ist, nennt man solche Abbildungen auch *k-Linearformen*, insbesondere für $k = 2$ *Bilinearformen*.

l heißt

(i) *positiv definit*, wenn $l(v, \dots, v) > 0$ für alle $v \neq 0$.

(ii) *positiv-semidefinit*, wenn $l(v, \dots, v) \geq 0$ für alle v.

(iii) *negativ definit*, wenn $l(v, \dots, v) < 0$ für alle $v \neq 0$.

(iv) *negativ-semidefinit*, wenn $l(v, \dots, v) \leq 0$ für alle v.

(v) *indefinit*, wenn $v \mapsto l(v, \dots, v)$ das Vorzeichen wechselt.

Lemma 150. *Sei $l \in L^k(V, \mathbb{R})$ eine* symmetrische *k-Linearform, d.h. es gelte*

$$l(v_1, \dots, v_k) = l(v_{i_1}, \dots, v_{i_k})$$

für jede Permutation $(i_1, \dots, i_k)$ von $(1, \dots, k)$. Dann gilt:

(i) Ist $l(v, \dots, v) = 0$ für alle $v \in V$, so ist $l = 0$.

(ii) Ist k ungerade, so ist $l = 0$ oder l indefinit.

Beweis. Zu (i). Vollständige Induktion über k.

$k = 1$. Trivial.

$(k-1) \to k$. Dann gilt

$$\begin{aligned} 0 &= l(tv + w, \dots, tv + w) \\ &= \sum_{i=0}^{k} \binom{k}{i} l(\underbrace{tv, \dots, tv}_{i-mal}, w, \dots, w) \\ &= \sum_{i=0}^{k} \binom{k}{i} t^i \, l(\underbrace{v, \dots, v}_{i-mal}, w, \dots, w). \end{aligned}$$

Ein Polynom verschwindet aber nur dann identisch, wenn alle Koeffizienten=0 sind. Daher ist insbesondere $l(v, \dots, v, w) = 0$ für alle v, w. Bei festem w ist $l(., \dots, ., w)$ symmetrisch und $(k-1)$-linear. Deshalb ist nach Induktionsvoraussetzung $l(v_1, \dots, v_{k-1}, w) = 0$ für alle v_i, w.

Zu (ii). Gibt es ein v mit $l(v, \dots, v) \neq 0$, etwa > 0, so ist $l(-v, \dots, -v) < 0$ und l indefinit. Andernfalls ist $l = 0$ nach (i). □

Definition 151. Die Funktion $f : V \supset G \to \mathbb{R}$ hat in p ein *strenges lokales Maximum*, wenn es ein $\epsilon > 0$ gibt, so dass

$$\forall_{x \in G} \left(0 < \|x - p\| < \epsilon \implies f(x) < f(p)\right).$$

Analog erklärt man strenge lokale Minima.

Satz 152 (Lokale Extrema). *Sei $f : V \supset G \to \mathbb{R}$ k-mal differenzierbar, $k \geq 1$, $p \in G$ und*

$$D_p f = 0, \ldots, D_p^{k-1} f = 0,$$
$$D_p^k f \neq 0.$$

Dann gilt:

(i) Ist $D_p^k f$ negativ definit, so hat f in p ein strenges lokales Maximum.

(ii) Ist $D_p^k f$ positiv definit, so hat f in p ein strenges lokales Minimum.

(iii) Ist $D_p^k f$ indefinit, insbesondere k ungerade, so hat f in p kein lokales Extremum, sondern einen sogenannten Sattelpunkt.

Im semidefiniten Fall wird keine Aussage gemacht.

Aus *(iii)* folgt insbesondere die wichtige notwendige Bedingung:

Hat f in p ein lokales Extremum, so ist $D_p f = 0$.

Beweis des Satzes. Die Idee des Beweises ergibt sich aus der Taylorformel. Es gilt

$$f(x) - f(p) = \frac{1}{k!} D_p^k f(x-p, \ldots, x-p) + R(x), \tag{39}$$

mit

$$\lim_{x \to p} \frac{R(x)}{\|x-p\|^k} = 0. \tag{40}$$

Also haben $f(x) - f(p)$ und $D_p^k f(x-p, \ldots, x-p)$ dasselbe Vorzeichen, vorausgesetzt, man kann das Restglied vernachlässigen. Letzteres zu zeigen, ist das technisches Problem des Beweises.

Die Einheitssphäre $S := \left\{ v \in V \;\middle|\; \|v\| = 1 \right\}$ ist abgeschlossen und beschränkt, nach dem Satz von Heine-Borel (der in endlich-dimensionalen Banachräumen ebenso gilt wie im Standard-$\mathbb{R}^n$. Warum?) also kompakt. Daher existieren

$$m := \min_{v \in S} D_p^k f(v, \ldots, v) \text{ und } M := \max_{v \in S} D_p^k f(v, \ldots, v).$$

Für beliebiges $v \in V$ folgt daraus

$$m\|v\|^k \leq D_p^k f(v, \ldots, v) \leq M\|v\|^k.$$

Ist $D_p^k f$ positiv definit, negativ definit oder indefinit, so ist $\epsilon := \frac{1}{2k!} \min(|m|, |M|) > 0$. Wegen (40) gibt es ein $\delta > 0$ mit $U_\delta(p) \subset G$ und

$$|R(x)| \leq \epsilon \|x-p\|^k \text{ für alle } x \in U_\delta(p).$$

<u>*Zu (i).*</u> Ist $D_p^k f$ negativ definit, also $M < 0$, so folgt für $x \in U_\delta(p)$

$$\frac{1}{k!} D_p^k f(x-p, \ldots, x-p) + R(x) \leq \frac{M}{k!}\|x-p\|^k - \frac{M}{2k!}\|x-p\|^k = \frac{M}{2k!}\|x-p\|^k \leq 0$$

mit Gleichheit nur für $x = p$. Aus (39) folgt die Behauptung *(i)*.

Zu (ii). Analog.

Zu (iii). Im indefiniten Fall ist $m < 0 < M$, und es gibt $v_1, v_2 \in S$ mit

$$D_p^k f(v_1, \dots, v_1) = m, \quad D_p^k f(v_2, \dots, v_2) = M.$$

Dann ist $x_j := p + \frac{\delta}{2} v_j \in U_\delta(p)$ und es gilt

$$\frac{1}{k!} D_p^k f(x_1 - p, \dots, x_1 - p) + R(x_1) \le \frac{m}{k!} \|x_1 - p\|^k - \frac{m}{2k!} \|x_1 - p\|^k = \frac{m}{2k!} \|x_1 - p\|^k < 0,$$
$$\frac{1}{k!} D_p^k f(x_2 - p, \dots, x_2 - p) + R(x_2) \ge \frac{M}{k!} \|x_2 - p\|^k - \frac{M}{2k!} \|x_2 - p\|^k = \frac{M}{2k!} \|x_2 - p\|^k > 0.$$

Aus (39) folgt die Behauptung *(iii)*. □

Bemerkung. Die Frage, wann eine symmetrische k-Linearform zum Beispiel positiv definit ist, ist eine Frage an die (multi)lineare Algebra. Ein häufiger Spezialfall ist $k = 2$. Wir wollen überdies annehmen, dass $V = \mathbb{R}^n$ ist. Dann ist

$$D_p^2 f(u, v) = \sum_{i,j=1}^{n} \partial_i \partial_j f(p) u_i v_j.$$

Die (symmetrische) Matrix

$$H = \begin{pmatrix} \partial_1 \partial_1 f(p) & \dots & \partial_1 \partial_n f(p) \\ \vdots & & \vdots \\ \partial_n \partial_1 f(p) & \dots & \partial_n \partial_n f(p) \end{pmatrix}$$

der zweiten partiellen Ableitungen heißt auch die *Hessesche Matrix* von f. Für sie gilt also

$$D_p^2 f(u, v) = \langle Hu, v \rangle$$

mit dem kanonischen Skalarprodukt $\langle u, v \rangle = \sum_{i=1}^n u_i v_i$. Es ist also eine interessante Frage, wann die durch eine symmetrische Matrix A gegebene Bilinearform $\langle Au, v \rangle$ positiv definit ist. In der Linearen Algebra lernt man (z.B. im Zusammenhang mit der Hauptachsentransformation), dass dies genau dann gilt, wenn alle Eigenwerte von A positiv sind. Dann nennt man auch A positiv definit. In der linearen Algebra lernt man auch, wie man die Eigenwerte bestimmt, und hat damit eine Methode, um im Fall $k = 2$ positive Definitheit nachzuprüfen.

Ein anderes Kriterium ist das folgende:

Lemma 153 (Hauptminorenkriterium). *Eine symmetrische $(n \times n)$-Matrix*

$$A = (a_{ij})_{i,j=1,\dots,n}$$

ist genau dann positiv definit, wenn alle Hauptminoren *positiv sind. Dabei sind* Hauptminoren *oder* Hauptabschnittsdeterminanten *die Determinanten der Matrizen*

$$A_k := (a_{ij})_{i,j=1,\dots,k}$$

A ist genau dann negativ definit*, wenn die Hauptminoren wechselndes Vorzeichen beginnend mit $a_{11} < 0$ haben.*

Man findet dieses Kriterium oft in der Literatur zitiert (als Kriterium von Sylvester oder Hurwitz), aber selten bewiesen. Wir geben deshalb einen Beweis im Anhang.

Im Falle $n = 2$ ist die Hessematrix gegeben durch

$$\begin{pmatrix} \partial_x^2 f & \partial_y \partial_x f \\ \partial_x \partial_y f & \partial_y^2 f \end{pmatrix}$$

und wir erhalten folgendes Kriterium für lokale Extrema:

Satz 154. *Sei $f : \mathbb{R}^2 \supset G \to \mathbb{R}$ zweimal differenzierbar auf der offenen Menge G und sei $p \in G$. Dann gilt:*

(i) Hat f in p ein lokales Extremum, so ist $D_p f = 0$.

(ii) Ist $D_p f = 0$ und gilt

$$\partial_x^2 f(p) \partial_y^2 f(p) - (\partial_x \partial_y f(p))^2 > 0,$$

so hat f in p ein strenges lokales Extremum, und zwar

- *ein Maximum, falls $\partial_x^2 f(p) < 0$,*
- *ein Minimum, falls $\partial_x^2 f(p) > 0$.*

(iii) Ist

$$\partial_x^2 f(p) \partial_y^2 f(p) - (\partial_x \partial_y f(p))^2 < 0,$$

so hat f in p kein lokales Extremum. (Sattelpunkt)

Wir geben dafür noch einen direkten Beweis ohne weiteren Bezug auf die lineare Algebra:

Beweis. Wir bezeichnen die Funktionalmatrix kurz mit

$$A = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$$

Dann ist

$$D_p^2 f(\begin{pmatrix} x \\ y \end{pmatrix}, \begin{pmatrix} x \\ y \end{pmatrix}) = ax^2 + 2bxy + cy^2 =: \phi(x, y).$$

Wählt man $y = 0$, so sieht man, dass $a > 0$ bzw. $a < 0$ notwendig für die positive bzw. negative Definitheit ist. Die ist in diesem Fall dann aber äquivalent dazu, dass

$$0 < x^2 + 2\frac{b}{a}xy + \frac{c}{a}y^2 = \left(x + \frac{b}{a}y\right)^2 + \frac{c}{a}y^2 - \frac{b^2}{a^2}y^2 = \left(x + \frac{b}{a}y\right)^2 + \frac{ac - b^2}{a^2}\, y^2.$$

Wählt man nun $y \neq 0$ und $x = -\frac{b}{a}y$, so ergibt sich $ac - b^2$ als weitere notwendige Bedingung. Diese ist aber auch hinreichend: Die rechte Seite ist dann ≥ 0, und verschwindet nur für $y = 0$ und $x = 0$. □

2.7 Differentialoperatoren der klassischen Vektoranalysis

- Wir interpretieren Abbildungen - wie in der Physik - als Vektor- oder Skalarenfelder.
- Gradient, Divergenz, Rotation und Laplaceableitung sind Felder, die mit Hilfe von Differentiationsprozessen aus anderen Feldern entstehen. Diese Operationen haben sich in der Physik als wichtig erwiesen.
- Wir lernen elementare Definitionen dieser Operationen im $\mathbb{R}^n$, bemühen uns aber auch um Definitionen in abstrakten Vektorräumen um zu klären, welche zusätzlichen Strukturen ggf. noch erforderlich sind.

Definition 155. Sei $G \subset V$ eine offene Teilmenge des endlich-dimensionalen Banachraums V. Für diese Vorlesung vereinbaren wir folgende Sprechweisen:

- Ein *Vektorfeld* auf G ist eine Abbildung $X : G \to V$.
- Ein *skalares Feld* auf G ist eine reellwertige Funktion $f : G \to \mathbb{R}$.

Wir wollen im folgenden eine kurze Einführung der klassischen Differentialoperatoren geben. Wir geben jeweils zwei Definitionen, eine elementare im $\mathbb{R}^n$ und eine abstrakte, die etwas mehr lineare Algebra voraussetzt und die aufzeigt, welche „Hintergrundstrukturen“ in die Definition einfließen.

2.7.1 Gradient

- **Naiv.** Mit $\langle x, y\rangle = \sum_{i=1}^n x_i y_i$ bezeichnen wir das kanonische Skalarprodukt auf $\mathbb{R}^n$. Der *Gradient* eines differenzierbaren skalaren Feldes $f : \mathbb{R}^n \supset G \to \mathbb{R}$ ist das folgende Vektorfeld:
$$\operatorname{grad} f : G \to \mathbb{R}^n, p \mapsto \operatorname{grad}_p f := (\partial_1 f(p), \dots, \partial_n f(p)).$$
Fundamentale Eigenschaften:
  (i) $\langle \operatorname{grad}_p f, v\rangle = D_p f(v)$ für alle $v \in V$.
  (ii) Der Gradient steht senkrecht auf den Niveaus von f. Genauer gilt für eine differenzierbare Kurve $c :]a, b[\to G$
$$f \circ c \text{ konstant} \iff \langle \operatorname{grad}_{c(t)} f, \dot c(t)\rangle = 0 \text{ für alle } t. \tag{41}$$
  Das folgt aus der Kettenregel, weil $\langle \operatorname{grad}_{c(t)} f, \dot c(t)\rangle = D_{c(t)} f(\dot c(t)) = \frac{d}{dt}(f \circ c)$.
  (iii) Der Gradient ist ein linearer Differentialoperator:
  Für $\alpha, \beta \in \mathbb{R}$ und $f, g : G \to \mathbb{R}$ ist
$$\operatorname{grad}(\alpha f + \beta g) = \alpha \operatorname{grad} f + \beta \operatorname{grad} g.$$
  (iv) Der Gradient gibt die Richtung und Größe des stärksten Wachstums der Funktion f an:
  Ist $\|v\| = 1$ und ϕ der Winkel zwischen dem Gradienten und der Richtung v, so ist
$$\partial_v f(p) = \|\operatorname{grad}_p f\| \cos\phi.$$

- **Für Fortgeschrittene.** Ist $l : V \times V \to \mathbb{R}$ eine (nicht notwendig symmetrische) Bilinearform, so liefert

$$j_l : V \to V^* = L(V, \mathbb{R}), v \mapsto l(v, .)$$

eine lineare Abbildung von V in V^*. Ist diese Abbildung ein Isomorphismus, so heißt l *nicht-degeneriert*. Ist l nicht-degeneriert, so kann man den l-Gradienten eines differenzierbaren skalaren Feldes $f : G \to \mathbb{R}$ definieren durch

$$\operatorname{grad}_p^l f := j_l^{-1}(D_p f),$$

d.h. durch die Gleichung

$$\boxed{l(\operatorname{grad}_p^l f, v) = D_p f(v) \quad \text{für alle } v \in V.}$$

Er ist ebenfalls ein linearer Differentialoperator und die obigen Eigenschaften (i), (ii) gelten mit l statt $\langle ., . \rangle$.

Beispiel 156 (Euklidischer Gradient). Seien $V = \mathbb{R}^n$ und $l(x, y) = \langle x, y \rangle = \sum x_i y_i$ das übliche Skalarprodukt. Das liefert den „naiven“ Gradienten wie oben. Allgemeiner gibt es in jedem *Euklidischen Vektorraum* einen kanonischen Gradienten.

□

Beispiel 157 (Vierergradient). Sei $V = \mathbb{R}^4$ und

$$L(x, y) = x_1 y_1 + x_2 y_2 + x_3 y_3 - x_4 y_4$$

das sogenannte *Lorentzprodukt*. Der zugehörigen Gradient, der sogenannte *Vierergradient* ist gegeben durch

$$\operatorname{grad}^L f = (\partial_1 f, \partial_2 f, \partial_3 f, -\partial_4 f).$$

Er spielt – wie das Lorentzprodukt – eine große Rolle in der Relativitätstheorie.

□

Beispiel 158 (Symplektischer Gradient). Sei $V = \mathbb{R}^{2n}$ und

$$\sigma(x, y) = x_{n+1} y_1 + \ldots + x_{2n} y_n - x_1 y_{n+1} - \ldots - x_n y_{2n}$$

das sogenannte *symplektische Skalarprodukt*. Der entsprechende *symplektische Gradient* ist gegeben durch

$$\operatorname{grad}^\sigma f = (-\partial_{n+1} f, \ldots, -\partial_{2n} f, \partial_1 f, \ldots, \partial_n f).$$

Er spielt eine wichtige Rolle in der Hamilton-Jacobi-Theorie der klassischen Mechanik, vgl. Beispiel 160.

□

2.7.2 Divergenz

- **Naiv.** Sei $V = \mathbb{R}^n$ und $X = (X_1, \ldots, X_n) : G \to \mathbb{R}^n$ ein differenzierbares Vektorfeld. Dann ist die *Divergenz* von X das folgende skalare Feld:

$$\operatorname{div} X : G \to \mathbb{R}, p \mapsto \operatorname{div}_p X := \sum_{i=1}^{n} \partial_i X_i(p).$$

Beachte: $\operatorname{div} X$ ist gerade die Summe der Diagonalelemente der Jacobimatrix $(\partial_j X_i)$ von X.

- **Für Fortgeschrittene.** Ist V ein beliebiger endlich-dimensionaler Vektorraum und $X : G \to V$ ein differenzierbares Vektorfeld, so ist für p in G das Differential D_pX ein Endomorphismus von V. Man definiert

$$\boxed{\operatorname{div} X = \operatorname{Spur}(D_pX).}$$

Der Satz von Gauß (Spezialfall des in der Analysis III zu beweisenden Stokesschen Integralsatzes) gibt eine Interpretation des Divergenz als "Quellstärke"" des Feldes X. Das hat damit zu tun, dass die Spur die Ableitung der Determinante ist und die Determinante Volumina misst.

2.7.3 Rotation

- **Naiv.** Sei $V = \mathbb{R}^3$ und $X : G \to \mathbb{R}^3$ ein differenzierbares Vektorfeld. Die *Rotation* von X ist das folgende Vektorfeld:

$$\operatorname{rot} X : G \to \mathbb{R}^3$$

mit

$$\operatorname{rot}_p X := \begin{pmatrix} \partial_2 X_3(p) - \partial_3 X_2(p) \\ \partial_3 X_1(p) - \partial_1 X_3(p) \\ \partial_1 X_2(p) - \partial_2 X_1(p) \end{pmatrix}.$$

- **Für Fortgeschrittene.** Für zwei Vektoren $a, b \in \mathbb{R}^3$ ist das Vektorprodukt $a \times b$ charakterisiert durch die Bedingungen

(i) $a \times b = 0$, falls a, b linear abhängig,

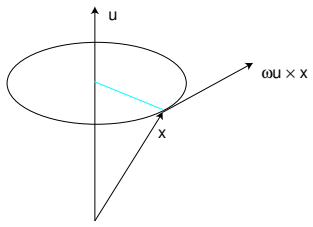
und andernfalls

(ii) $\|a \times b\| = \|a\|\|b\| \sin \angle(a, b)$,

(iii) $\langle a \times b, a\rangle = \langle a \times b, b\rangle = 0$ und $(a, b, a \times b)$ ist eine positiv orientierte Basis des $\mathbb{R}^3$.

Durch diese Bedingungen lässt sich ein Vektorprodukt in jedem orientierten 3-dimensionalen Euklidischen Vektorraum definieren.

Wir erklären nun zwei Methoden, um Achsrotationen in einem orientierten 3-dimensionalen Euklidische Vektorraum V zu beschreiben. Die Achse sei gegeben durch einen Einheitsvektor u. Das Geschwindigkeitsfeld der Drehung in einem Punkt $x \in V$ muss dann senkrecht zu x und u stehen und mit dem Abstand von der Achse linear anwachsen.



Das wird geleistet

1. durch ein Feld

$$x \mapsto \omega u \times x,$$

wo $\omega \in \mathbb{R}$ die sogenannte Winkelgeschwindigkeit bezeichnet, oder

2. durch ein Feld

$$x \mapsto Ax,$$

wobei A ein schiefadjungierter (=schiefsymmetrischer) Endomorphismus von V mit $\operatorname{Kern}(A) = \mathbb{R}u$ ist. (Jeder schiefadjungierte Endomorphismus $\neq 0$ eines dreidimensionalen Raumes hat einen 1-dimensionalen Kern. Warum?)

Der Zusammenhang zwischen diesen beiden Methoden ist einfach: Für $a \in V$ ist $A : x \mapsto a \times x$ schiefadjungiert, weil

$$\langle a \times x, y \rangle = - \langle x, a \times y \rangle ,$$

und die Abbildung

$$a \mapsto a \times \ldots$$

liefert eine Isomorphismus(!) von V auf den Vektorraum der schiefadjungierten Endomorphismen von V.

Das Differential D_pX eines differenzierbaren Vektorfeldes an der Stelle p ist im allgemeinen weder schiefsymmetrisch noch symmetrisch, aber man kann es in einen schiefsymmetrischen Anteil (= Rotationsanteil) und in einen symmetrischen Anteil zerlegen:

$$D_pX = \frac{1}{2}(D_pX - D_pX^*) + \frac{1}{2}(D_pX + D_pX^*).$$

(Der * bezeichnet die Adjungierte oder transponierte Matrix.) Dann gilt (Nachrechnen!)

$$\boxed{\operatorname{rot}_p X \times \ldots = D_pX - D_pX^*.}$$

In diesem Sinne ist die Rotation $\operatorname{rot} X$ der doppelte Rotationsanteil von DX.

Es gibt eine Verallgemeinerung der Rotation auf Vektorräume beliebiger Dimension, aber nicht mehr für Vektorfelder, sondern für kompliziertere Objekte, die sogenannten Differentialformen vgl. (Analysis III).

Satz 159. *Für zweimal differenzierbare Felder gilt*

$$\begin{aligned} \operatorname{rot} \operatorname{grad} f &= 0, \\ \operatorname{div} \operatorname{rot} X &= 0. \end{aligned}$$

Das gibt also notwendige Bedingungen dafür, dass sich ein differenzierbares Vektorfeld als Gradient (eines Potentials) oder Rotation (eines Vektorpotentials) schreiben läßt: die Rotation bzw. Divergenz muß verschwinden. Lokal, nicht aber global, sind diese Bedingungen auch hinreichend, vgl. Analysis III.

Beweis. Stures Nachrechnen unter Benutzung des Satzes von Schwarz über die Vertauschbarkeit der zweiten partiellen Ableitungen. □

2.7.4 Laplaceoperator

Für zweimal differenzierbare skalare Felder auf $G \subset \mathbb{R}^n$ (oder in einem Euklidischen Vektorraum) ist der *Laplaceoperator* definiert durch

$$\boxed{\Delta f = \operatorname{div} \operatorname{grad} f.}$$

In Koordinaten bedeutet das

$$\Delta_p f = \sum_{i=1}^{n} \partial_i^2 f(p).$$

Funktionen mit $\Delta f = 0$ heißen *harmonische Funktionen.*

Der Laplaceoperator spielt eine fundamentale Rolle für die Beschreibung sehr vieler physikalischer Phänomene (Wärmeleitungsgleichung, Wellengleichung, Schrödingergleichung). Zum Beispiel ist die Amplitude f einer Welle in einem homogenen 3-dimensionalen Medium eine Funktion der Raumkoordinaten x_i und der Zeit t und genügt der Gleichung

$$\Delta_p f = \sum_{i=1}^{3} \frac{\partial^2 f}{\partial x_i^2} = \frac{1}{c^2} \frac{\partial^2}{\partial t^2}.$$

Normiert man die Ausbreitungsgeschwindigkeit auf $c = 1$ und verwendet im $\mathbb{R}^4$ den Vierergradienten, so schreibt man den entsprechenden Laplaceoperator auch als $\Box f := \operatorname{div} \operatorname{grad} f$, und die Wellengleichung wird einfach

$$\Box f = 0.$$

2.8 Ein Kapitel Newtonsche Mechanik

Die Differentialrechnung verdankt ihre Entstehung ganz wesentlich den Bemühungen um das Verständnis der Gesetze der Mechanik. Daher ist es (auch für angehende Finanzmathematikerinnen) nicht unangemessen, ein wenig über die mathematischen Modelle der Mechanik zu lernen.

- Wir lernen die Newtonschen Bewegungsgleichungen und zeigen die Erhaltungssätze für Energie und Drehimpluls.
- Wir leiten die Keplerschen Planetengesetze aus Newtons Graviatitonsgesetz und Bewegungsgesetz her.

Die Bewegung eines *Massenpunktes* der Masse m im 3-dimensionalen Euklidischen Raum wird beschrieben durch eine Kurve

$$x : \mathbb{R} \supset J \to \mathbb{R}^3, t \mapsto x(t),$$

wobei wir die in der Physik übliche Bezeichnungsweise verwenden. Die Geschwindigkeit der Punktes ist

$$\dot{x} = \frac{dx}{dt} = Dx(1) : J \to \mathbb{R}^3$$

und seine Beschleunigung gegeben durch

$$\ddot{x} = \frac{d^2x}{dt^2} : J \to \mathbb{R}^3.$$

Das Newtonsche Bewegungsgesetzt besagt nun, dass diese Bewegung bestimmt wird durch die Kraft, die auf den Massenpunkt wirkt, und zwar durch die Formel "Kraft = Masse mal Beschleunigung":

$$m\ddot{x} = F(x).$$

Dabei ist die Kraft gegeben durch ein Vektorfeld $F : \mathbb{R}^3 \to \mathbb{R}^3$, wenn wir uns auf den Fall beschränken, dass die Kraft nur vom Ort und nicht auch von der Zeit abhängt (auf *autonome Systeme* würde der Physiker sagen). In der Physik ist es üblich, den Impuls $p := m\dot{x}$ als "dummy-Variable" einzuführen und die Bewegung des Massenpunktes als eine Kurve im 6-dimensionalen sogenannten *Phasenraum* zu verstehen. Ein Vorteil dieser Beschreibung ist, dass bei bekannter Kraft F die Bewegung des Punktes bekannt ist, wenn man weiß, wo im Phasenraum er sich zu einem Zeitpunkt t_0 befindet. Die Newtonsche Bewegungsgleichung im Phasenraum ist dann das folgende Differentialgleichungssystem:

$$\begin{aligned} \dot{x} &= m^{-1}p, \\ \dot{p} &= F(x). \end{aligned} \tag{42}$$

Lösungen $t \mapsto (x(t), p(t))$ heißen auch *Phasenkurven*. Ihre ersten 3 Komponenten liefern also die Bahn des Massenpunktes im Ortsraum, die zweiten 3 dagegen den Impuls.

Beispiel 160 (Energieerhaltungssatz). Wir nehmen an, dass die Kraft F ein Potential $U : \mathbb{R}^3 \to \mathbb{R}$ besitzt, d.h. dass

$$F(x) = -\operatorname{grad}_x U.$$

Wir definieren dann eine Funktion $H : \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}$ auf dem Phasenraum durch

$$H(x,p) = U(x) + \frac{1}{2m}\langle p, p\rangle .$$

H ist die Summe aus potentieller und kinetischer Energie und heißt auch die *Hamiltonfunktion*. Die Bewegungsgleichungen lauten dann

$$\dot{x}_i = \frac{\partial H}{\partial p_i}, \quad \dot{p}_i = -\frac{\partial H}{\partial x_i},$$

oder, unter Verwendung des symplektischen Gradienten aus Beispiel 158,

$$(\dot{x}, \dot{p}) = -\operatorname{grad}^{\sigma}_{(x,p)} H.$$

Weil für die symplektische Bilinearform aber $\sigma(v, v) = 0$ für alle $v \in \mathbb{R}^{2n}$, folgt daraus

$$\sigma(\operatorname{grad}^{\sigma}_{(x(t),p(t))} H(\dot{x}(t), \dot{p}(t))) = 0.$$

Nach (41) ist also H auf den Phasenkurven $(x(t), y(t))$ konstant. Wir haben den Energieerhaltungssatz bewiesen.

□

Beispiel 161 (Drehimpulserhaltung). Wir nehmen nun an, dass F ein zentrales Feld ist, d.h. dass für alle $x \neq 0$

$$F(x) = f(x)x.$$

Wir definieren eine Funktion

$$J : \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}^3, \quad x(x, p) \mapsto x \times p,$$

die der *Drehimpuls* heißt. Für eine Phasenkurve $t \mapsto (x(t), p(t))$ erhalten wir

$$\frac{d}{dt} J(x, p) = \dot{x} \times p + x \times \dot{p} = -m^{-1} p \times p - f(x) x \times x = 0.$$

Also ist der Drehimpuls J auf jeder Phasenkurve konstant. Wegen $J \perp x$ liegt die zugehörige Ortskurve in einer Ebene senkrecht zum konstanten J.

□

Das letzte Beispiel dieses Abschnitts dokumentiert eine der ganz großen Leistungen in der Geschichte der Naturwissenschaften und einen phantastischen frühen Erfolg der neu entdeckten Differentiallrechnung.

Beispiel 162 (Keplersche Gesetze als Konsequenz der Newtonschen Bewegungsgleichung und des Gravitationsgesetzes). Die Keplerschen Gesetze für die Bewegung der Planeten in einem Zentralfeld besagen:

1. Die Planetenbahnen sind Ellipsen mit der Sonne im Brennpunkt.
2. Der Fahrstrahl überstreicht in gleichen Zeiten gleiche Flächen.
3. Die Quadrate der Umlaufzeiten verhalten sich wie die Kuben der großen Halbachsen.

Kepler (1571-1630) hatte diese Gesetze aus umfassenden astronomischen Beobachtungen (von Tycho Brahe und ihm selbst) errechnet. Das ist eine staunenswerte Leistung experimenteller Naturwissenschaft, vor allem, wenn man bedenkt, dass Kepler eines der ersten Fernrohre konstruierte. Eine Generation später führten sie Newton zur Entdeckung seines Gravitationsgesetzes

$$F(x) = -\gamma m M \frac{x}{r^3}, \quad r = \|x\|, \tag{43}$$

aus dem sich in Verbindung mit den Bewegungsgleichungen die Keplerschen Gesetze herleiten, wie wir nun zeigen wollen.

Die Bewegungsgleichung sieht so aus:

$$\dot{x} = m^{-1}p \tag{44}$$

$$\dot{p} = -\gamma m M \frac{x}{r^3}. \tag{45}$$

Weil das Gravitationsfeld zentralsymmetrisch ist, ist der Drehimpuls $J = x \times p$ längs jeder Lösungskurve (x, p) konstant, und die Bewegung verläuft in einer Ebene senkrecht zu J.

Zwischenrechnung. Wir betrachten eine Lösung $t \mapsto (x(t), p(t))$ der Bewegungsgleichungen und erhalten

$$\frac{d}{dt}(J \times p) = J \times \dot{p} = (x \times p) \times \dot{p} = \langle x, \dot{p}\rangle\, p - \langle p, \dot{p}\rangle\, x = -\gamma m^2 M \left(\frac{\dot{x}}{r} - \frac{\langle \dot{x}, x\rangle}{r^3} x\right).$$

Wenn man Erfahrung im Differenzieren von Vektorfeldern hat, kommt einem der Klammerausdruck bekannt vor, vgl. auch Beispiel 122. Nach (21) ist nämlich

$$\frac{d}{dt}\,\frac{x}{r} = \frac{\dot{x}}{r} - \frac{1}{r^2}\,\frac{dr}{dt} = \frac{\dot{x}}{r} - \frac{\langle x, \dot{x}\rangle}{r^3}\,\frac{dr}{dt}.$$

Wir definieren deshalb

$$A(x, p) := \frac{J \times p}{\gamma m^2 M} + \frac{x}{r}.$$

Dann ist auch der sogenannte *Lenzsche Vektor* A eine Erhaltungsgröße, d.h. $t \mapsto A(x(t), p(t))$ ist längs jeder Phasenkurve konstant.

Wir nehmen jetzt an, dass J in Richtung der z-Achse zeigt. Dann liegen $J \times p$ und x in der xy-Ebene. Also liegt auch A in der xy-Ebene, und wir nehmen an, dass das konstante(!) A in Richtung der positiven x-Achse zeigt. Wir schreiben $x = r(\cos\phi, \sin\phi, 0)$ in Zylinderkoordinaten. Mit $\|A\| =: \epsilon$ ist dann

$$\langle A, x\rangle = \epsilon\, r \cos\phi.$$

Andrerseits ist

$$\langle A, x\rangle = \frac{\langle J \times p, x\rangle}{\gamma m^2 M} + r = -\frac{\langle J, x \times p\rangle}{\gamma m^2 M} + r = -\underbrace{\frac{\langle J, J\rangle}{\gamma m^2 M}}_{=:\eta} + r$$

Aus den beiden Gleichungen folgt

$$r(1 - \epsilon\cos\phi) = \eta. \tag{46}$$

Das ist die Polarkoordinaten-Gleichung eines Kegelschnitts mit Brennpunkt im Ursprung und für $\epsilon < 1$ eine Ellipse mit den Halbachsen

$$a = \frac{\eta}{1 - \epsilon^2}, \quad b = a\sqrt{1 - \epsilon^2}. \tag{47}$$

Das findet man in jeder besseren Formeltafel. Wir geben eine kurze Herleitung:

> Hier sollen $x = r\cos\phi$ und $y = r\sin\phi$ die kartesischen Koordinaten des Punktes $x(t)$ bezeichnen. Aus (46) folgt
>
> $$r = \epsilon x + \eta,$$

und nach Quadrieren

$$x^2 + y^2 = \epsilon^2 x^2 + 2\epsilon\eta x + \eta^2$$
$$x^2(1-\epsilon^2) - 2\epsilon\eta x + y^2 = \eta^2$$
$$x^2 - 2\epsilon \underbrace{\frac{\eta}{1-\epsilon^2}}_{=:a} x + \frac{y^2}{1-\epsilon^2} = \eta\,\frac{\eta}{1-\epsilon^2} = \eta a = (1-\epsilon^2)a^2$$
$$(x-\epsilon a)^2 + \frac{y^2}{1-\epsilon^2} = (1-\epsilon^2)a^2 + \epsilon^2 a^2 = a^2$$

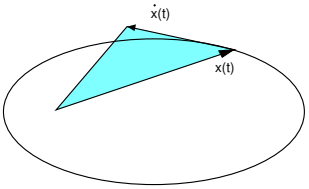
Nach Division mit a^2 folgt schließlich – falls $\epsilon < 1$ – die Gleichung für eine in Richtung der x-Achse verschobene Ellipse mit den Halbachsen a und $b = a\sqrt{1-\epsilon^2}$:

$$\frac{(x-\epsilon a)^2}{a^2} + \frac{y^2}{a^2(1-\epsilon^2)} = 1.$$

Damit ist das 1. Keplersche Gesetz bewiesen.

Der Flächeninhalt des Dreieck zwischen $x(t)$ und $\dot{x}(t)$ ist gegeben durch

$$\frac{1}{2}\|x \times \dot{x}\| = \frac{1}{2m}\|x \times p\| = \frac{1}{2m}\|J\|.$$



In einem kleinen Zeitintervall Δt überstreicht der Fahrstrahl in erster Näherung die Fläche $\frac{1}{2}\|x \times \Delta t\dot{x}\|$, zwischen t_0 und t_1 also die Fläche

$$\int_{t_0}^{t_1} \frac{1}{2}\|x \times \dot{x}\| dt = \frac{t_1 - t_0}{2m}\|J\|.$$

Das ist das 2. Keplersche Gesetz.

Ist T die Umlaufzeit, so ist die Fläche der Ellipse $F = T\,\frac{\|J\|}{2m}$. Andrerseits gilt für Ellipsen, dass $F = \pi ab$. Daher erhalten wir

$$\frac{T^2}{4m^2}\|J\|^2 = \pi^2 a^4(1-\epsilon^2) = \pi^2 a^3 \eta = \pi^2 a^3 \frac{\|J\|^2}{\gamma m^2 M}.$$

Also

$$T^2 = \frac{4\pi^2}{\gamma M}\, a^3.$$

Das ist das 3. Keplersche Gesetz.

□

3 Mehrdimensionale Differentialrechnung: Die großen Sätze

3.1 Der Umkehrsatz

- Weil die Ableitung einer differenzierbaren Abbildung diese lokal sehr gut approximiert, gibt sie zum Beispiel Auskunft auf die Frage nach der lokalen Umkehrbarkeit der Funktion.

Lemma 163. *Für endlich-dimensionale Banachräume V, W gleicher Dimension sei*

$$\mathrm{Iso}(V,W) := \left\{ A \in L(V,W) \;\middle|\; A \text{ invertierbar} \right\}.$$

Dann gilt:

(i) $\mathrm{Iso}(V,W)$ *ist offen in* $L(V,W)$.

(ii) Die Inversenabbildung

$$\mathrm{inv} : \mathrm{Iso}(V,W) \to L(W,V), A \mapsto A^{-1}$$

ist differenzierbar mit

$$D_A \,\mathrm{inv}(B) = -A^{-1}BA^{-1}$$

(iii) Für $A \in Iso(V,W)$ *und* $v \in V$ *gilt*

$$\|A(v)\| \geq \frac{1}{\|A^{-1}\|}\|v\|, \tag{48}$$

wobei $\|A^{-1}\|$ *die Operatornorm bezeichnet.*

Beweis. Die Behauptungen *(i), (ii)* folgen aus dem Beispiel 124 mit Hilfe eines Isomorphimus $\Phi : V \to W$.

Zu (iii). Es ist

$$\|v\| = \|A^{-1}(A(v))\| \leq \|A^{-1}\| \cdot \|A(v)\|.$$

Daraus folgt (48). □

Satz 164 (Umkehrsatz). *Seien $G \subset V$ offen und $f : V \supset G \to W$* stetig *differenzierbar, d.h. Df existiert und ist stetig. Sei $p \in G$ und sei*

$$D_p f : V \to W \quad \text{invertierbar.}$$

Dann ist f bei p lokal invertierbar mit stetig differenzierbarem Inversen.
Genauer: Es gibt eine offene Umgebung U von p in G, so dass gilt

(i) $f|_U$ *ist injektiv,*

(ii) $f(U)$ *ist offen in* W,

(iii) $(f|_U)^{-1} : f(U) \to V$ *ist stetig differenzierbar und für alle* $x \in U$ *gilt*

$$D_{f(x)}(f|_U)^{-1} = (D_x f)^{-1}.$$

Bemerkung. Aus der letzten Formel folgt: Ist f sogar k-mal stetig differenzierbar, so ist auch die lokale Umkehrung k-mal stetig differenzierbar.

Definition 165. Eine k-mal stetig differenzierbare Abbildung mit einem k-mal stetig differenzierbaren Inversen heißt ein *C^k-Diffeomeorphismus*.

Eine stetig differenzierbare Abbildung mit invertierbarem Differential ist also lokal ein C^1-Diffeomorphismus.

Beweis des Umkehrsatzes. *Zu (i).* Lokale Injektivität von f bei p.

Wir setzen $F := D_p f$ und $\beta := \frac{1}{\|F^{-1}\|}$.

Idee: Seien x, y nah bei p. Dann ist

$$\begin{aligned} \|f(y) - f(x)\| &= \|(f(y) - f(p)) - (f(x) - f(p))\| \\ &\approx \|D_p f(y-p) - D_p f(x-p)\| = \|F(y-x)\| \underset{(48)}{\geq} \beta \|y - x\|. \end{aligned} \tag{49}$$

Aus $y \neq x$ „folgt“ dann also $f(y) \neq f(x)$.

Um das zu präzisieren, müssen wir das $\approx$-Zeichen quantitativ kontrollieren. Der Approximationsfehler ist

$$\|f(y) - f(x) - F(y-x)\| = \|(f(y) - F(y)) - (f(x) - F(x))\| = \|\phi(y) - \phi(x)\|$$

mit

$$\phi(x) := f(x) - F(x).$$

Offenbar ist ϕ stetig differenzierbar und $D_p\phi = D_p f - F = 0$. Also gibt es ein $\delta > 0$, so dass

$$U = U_\delta(p) \subset G$$

und

$$\|D_\xi \phi\| \leq \frac{\beta}{3} \text{ für alle } \xi \in U.$$

(Hier genügte im Augenblick auch $\|D_\xi \phi\| < \beta$, aber im Hinblick auf den Beweis von (ii) fordern wir die schärfere Abschätzung.) Dann ist nach dem Schrankensatz

$$\|\phi(y) - \phi(x)\| \leq \sup_{\xi \in U} \|D_\xi f - F\| \, \|y - x\| \leq \frac{\beta}{3} \|y - x\|.$$

Der Approximationsfehler in (49) ist also maximal $\frac{1}{3}$ der rechten Seite. Also ist

$$\|f(y) - f(x)\| \geq \frac{2}{3} \beta \, \|y - x\| \tag{50}$$

und $f|_U$ injektiv. Nach dem Lemma ist weiter

$$D_x f \text{ invertierbar für } x \in U. \tag{51}$$

Zu (ii). Offenheit von $f(U)$.

Seien U wie oben und $x \in U$. Wir müssen zeigen, dass es ein $\epsilon > 0$ gibt, so dass

$$U_\epsilon(f(x)) \subset f(U).$$

Wähle zunächst $r > 0$ mit

$$K := \{y \in V \mid \|y - x\| \leq r\} \subset U.$$

Nach (50) gilt

$$\|y - x\| = r \implies \|f(y) - f(x)\| \geq \frac{2}{3}\beta r, \tag{52}$$

d.h. die Randpunkte von K werden durch f auf Punkte abgebildet, die mindestens den Abstand $\frac{2}{3}\beta r$ von $f(x)$ haben. Wir wollen zeigen, dass

$$U_{\frac{1}{3}\beta r}(f(x)) \subset f(K) \subset f(U). \tag{53}$$

Sei also $z \in U_{\frac{1}{3}\beta r}(f(x))$. Sei $y^* \in K$ ein Punkt, in dem die stetige Funktion $\|f(y) - z\|$ auf dem kompakten K ihr Minimum annimmt. Wir wollen zeigen, dass $f(y^*) = z$; dann ist (53) bewiesen.

Zunächst ist

$$\|y^* - x\| < r. \tag{54}$$

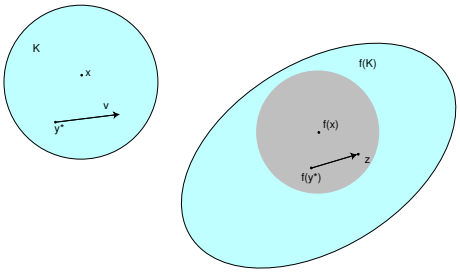
Sonst wäre nach Definition von K nämlich $\|y^* - x\| = r$, und nach (52) folgte mit der Dreiecksungleichung

$$\|f(y^*) - z\| \geq \frac{1}{3}\beta r.$$

Aber das steht wegen $\|z - f(x)\| < \frac{1}{3}\beta r$ im Widerspruch zur Wahl von y^*.

Wir nehmen nun an, dass

$$f(y^*) \neq z. \tag{55}$$



Wegen der Invertierbarkeit von $D_{y^*}f$ gibt es dann ein $v \neq 0$ mit

$$D_{y^*}f(v) = z - f(y^*) \neq 0.$$

Geht man von y^* in Richtung v, so bleibt man für eine Weile in K, und das f-Bild bewegt sich in Richtung $z - f(y^*)$, also in Richtung auf z zu. Daher liegt für kleine positive t der Punkt $f(y^* + tv)$ näher an z als $f(y^*)$, und wir erhalten einen Widerspruch zur Wahl von y^*.

Wir präzisieren das:

Wähle $\delta_1 > 0$ so klein, dass $y^* + tv \in K$ für alle $t \in [0, \delta_1]$. Dann ist

$$\begin{aligned} f(y^* + tv) - z &= f(y^*) - z + tD_{y^*}f(v) + R(y^* + tv) \\ &= (f(y^*) - z)(1 - t) + R(y^* + tv). \end{aligned}$$

Es gibt ein $t \in]0, \delta_1[$, so dass

$$\|R(y^* + tv)\| \leq \frac{\|z - f(y^*)\|}{2\|v\|} \|tv\| = \frac{t}{2} \|z - f(y^*)\|,$$

also

$$\|f(y^* + tv) - z\| \leq (1 - t)\|f(y^*) - z\| + \frac{t}{2} \|z - f(y^*)\| < \|f(y^*) - z\|$$

im Widerspruch zur Wahl von y^*. Damit war die Annahme (55) falsch, und es gilt $f(y^*) = z$, also $U_{\frac{1}{3}\beta r}(f(x)) \subset U$.

Zu (iii). Stetige Differenzierbarkeit der lokalen Umkehrabbildung.

Sei $g := (f|_U)^{-1} : f(U) \to V$. Seien $z, w \in f(U)$ und $x := g(z)$. Dann haben wir

$$\underbrace{f(g(w))}_{=w} = \underbrace{f(g(z))}_{=z} + D_x f(g(w) - g(z)) + R(g(w))$$

oder

$$D_x f(g(w) - g(z)) - (w - z) = -R(g(w)).$$

mit $\lim_{y\to x} \frac{R(y)}{\|y-x\|} = 0$. Wegen (51) ist $D_x f$ invertierbar. Es folgt

$$g(w) = g(z) + (D_x f)^{-1}(w - z) \underbrace{-(D_x f)^{-1}(R(g(w)))}_{=:\tilde{R}(w)}.$$

Wir wollen zeigen, dass

$$\lim_{w\to z} \frac{\tilde{R}(w)}{\|w - z\|} = 0. \tag{56}$$

Wegen der Injektivität von g und nach (50) gilt für $w \neq z$

$$0 < \|g(w) - g(z)\| \leq \frac{3}{2\beta}\|w - z\|.$$

Insbesondere ist g stetig, und aus

$$\frac{\tilde{R}(w)}{\|w - z\|} = -(D_x f)^{-1} \left(\underbrace{\frac{R(g(w))}{\|g(w) - g(z)\|}}_{\to 0 \text{ für } w\to z} \right) \underbrace{\frac{\|g(w) - g(z)\|}{\|w - z\|}}_{\leq \frac{3}{2\beta}}$$

folgt die Behauptung (56). Also ist g differenzierbar und

$$D_{f(x)} g = D_z g = (D_x f)^{-1}.$$

Schließlich ist $z \mapsto g(z) \mapsto D_{g(z)} f \mapsto (D_{g(z)} f)^{-1}$ als Komposition stetiger Abbildungen wieder stetig. Damit haben wir die stetige Differenzierbarkeit der Umkehrabbildung gezeigt. □

Bemerkung. Die Formel für die Ableitung folgt auch aus

$$(f|_U)^{-1} \circ f|_U = \mathrm{id}$$

mit der Kettenregel:

$$D_{f(x)}(f|_U)^{-1} \circ D_x(f|_U) = D_x \,\mathrm{id} = \mathrm{id}\,.$$

Beispiel 166. Die Abbildung $f : \mathbb{R}^2 \setminus \{0\} \to \mathbb{R}^2$ mit $f(x, y) := (x^2 - y^2, 2xy)$ hat die Funktionalmatrix

$$f'(x, y) = \begin{pmatrix} 2x & -2y \\ 2y & 2x \end{pmatrix}.$$

Sie ist deshalb stetig differenzierbar und $D_{(x,y)}f$ ist für alle $(x,y) \in \mathbb{R}^2 \setminus \{0\}$ invertierbar. Also besitzt f um jeden Punkt lokal ein stetig differenzierbares Inverses. Aber f ist nicht global invertierbar, weil z.B. $f(1,1) = f(-1,-1)$. Es ist

$$f(\{(x,y) \,|\, x > 0 \text{ und } y > 0\}) = \{(x,y) \,|\, y > 0\}$$

und

$$g(x,y) = \frac{1}{\sqrt{2}} \left(\sqrt{\sqrt{x^2+y^2}+x}, \sqrt{\sqrt{x^2+y^2}-x} \right), \quad y > 0$$

ist das Inverse von $f|_{\{(x,y) \,|\, x>0 \text{ und } y>0\}}$. Die Formel für die Ableitung der Inversen liefert

$$g'(f(x,y)) = (f'(x,y))^{-1} = \frac{1}{2(x^2+y^2)} \begin{pmatrix} x & y \\ -y & x \end{pmatrix}.$$

Zum Beispiel ergibt sich für $x = y = 1$

$$g'(1,1) = \frac{1}{4} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}.$$

□

Beispiel 167 (Stetige Polarkoordinaten). Die Polarkoordinaten in der Ebene sind nicht eindeutig, die Winkelkoordinate ist nur bis auf ein ganzzahliges Vielfaches von 2π bestimmt. Und wenn man die Eindeutigkeit mit "Gewalt" erzwingt, indem man zum Beispielverlangt, dass $\phi \in [-\pi, \pi[$, so wird die Winkelkoordinate auf der negativen x-Achse unstetig.

Wir wollen aber überlegen: Eine stetige Kurve $c : [a,b] \to \mathbb{R}^2 \setminus \{0\}$ kann man auch in Polarkoordinaten mit stetiger Winkelfunktion beschreiben: Ist $c(a) = \|c(a)\|(\cos\phi_0, \sin\phi_0)$, so gibt es genau eine stetige Funktion $\Phi : [a,b] \to \mathbb{R}$ mit $\Phi(a) = \phi_0$ und

$$\begin{aligned} c(t) &= \|c(t)\|(\cos\Phi(t), \sin\Phi(t)) \\ &= \|c(t)\| e^{i\Phi(t)} \text{ in komplexer Notation.} \end{aligned} \tag{57}$$

Wir betrachten die Abbildung

$$\begin{aligned} f : \mathbb{R}^2 \supset G := \left\{ (r,\phi) \;\middle|\; r > 0 \right\} &\to \mathbb{R}^2 \setminus \{0\} \\ (r,\phi) &\mapsto (r\cos\phi, r\sin\phi) \end{aligned}$$

Dann ist

$$f'(r,\phi) = \begin{pmatrix} \cos\phi & -r\sin\phi \\ \sin\phi & r\cos\phi \end{pmatrix}.$$

Rechnen Sie nach, dass das für alle $(r,\phi) \in G$ invertierbar ist. Also ist f nach dem Umkehrsatz lokal invertierbar. Wir wissen natürlich mehr: Die Abbildung f ist surjektiv auf $\mathbb{R}^2 \setminus \{0\}$, und mittels Arcus-Funktionen lassen sich lokale Umkehrabbildungen explizit hinschreiben.Weil das wegen der erforderlichen Fallunterscheidungen mühsam ist, wählen wir nun zu jedem $p = (r,\phi) \in G$ eine offene Umgebung U_p, die von f diffeomorph auf eine offene Menge $V_p := f(U_p) \subset \mathbb{R}^2 \setminus \{0\}$ abgebildet wird. Dann ist $(V_p)_{p\in G}$ eine offene Übderdeckung von $\mathbb{R}^2 \setminus \{0\}$, und wegen der Stetigkeit von c ist $\left(c^{-1}(V_p)\right)_{p\in G}$ eine offene Überdeckung von $[a,b]$. Nach dem Lebesgue-Lemma gibt es eine Zerlegung

$$a = t_0 < t_1 < \ldots < t_n = b,$$

so dass jedes $[t_{j-1}, t_j]$ in einem der $c^{-1}(V_p)$ enthalten ist. Wir wählen zu jedem j ein solches p, und schreiben $f_j := f|_{U_p}$.

Wir definieren nun rekursiv

$$\begin{aligned}\Phi(a) &:= \phi_0\\ \Phi(t) &:= \left(f_j^{-1}(c(t))\right)_2 - \left(f_j^{-1}(c(t_{j-1}))\right)_2 + \Phi(t_{j-1}) \text{ für } t \in]t_{j-1}, t_j].\end{aligned}$$

Dabei bedeutet der untere Index $(.)_2$ die 2. Komponente (eben die ϕ-Komponente). Offenbar ist dann $\Phi|_{]t_{j-1},t_j]}$ stetig, und weil außerdem

$$\lim_{t \searrow t_{j-1}} \Phi(t) = \Phi(t_{j-1}),$$

ist $\Phi : [a, b] \to \mathbb{R}$ stetig. Wir zeigen, dass (57) gilt. Nehmen wir an, dass das bereits für $t \le t_{j-1}$ erfüllt ist. Dann folgt für $t_{j-1} < t \le t_j$:

$$\begin{aligned}\|c(t)\|e^{i\Phi(t)} &= \|c(t)\| \exp\left(i\left(\left(f_j^{-1}(c(t))\right)_2 - \left(f_j^{-1}(c(t_{j-1}))\right)_2 + \Phi(t_{j-1})\right)\right)\\ &= \|c(t)\| \exp\left(i\left(f_j^{-1}(c(t)\right)_2\right) \frac{\|c(t_{j-1})\| \exp\left(i\Phi(t_{j-1})\right)}{\|c(t_{j-1})\| \exp\left(i\left(f_j^{-1}(c(t_{j-1}))\right)_2\right)}\\ &= c(t)\frac{c(t_{j-1})}{c(t_{j-1})} = c(t).\end{aligned}$$

Zur Eindeutigkeit von Φ. Wir nehmen an, dass

$$\|c(t)\|e^{i\Phi(t)} = c(t) = \|c(t)\|e^{i\tilde{\Phi}(t)} \text{ für alle } t \in [a, b].$$

Dann folgt

$$e^{i(\Phi(t)-\tilde{\Phi}(t))} = 1 \text{ für alle } t \in [a, b],$$

also

$$\Phi(t) - \tilde{\Phi}(t) \in \left\{2k\pi \;\middle|\; k \in \mathbb{Z}\right\} \text{ für alle } t \in [a, b].$$

Wenn Φ und $\tilde{\Phi}$ stetig sind mit $\Phi(a) = \phi_0 = \tilde{\Phi}(a)$, so folgt daraus $\Phi = \tilde{\Phi}$.

□

3.2 Implizite Funktionen

- Ist F linear, so ist $F(x,y) = F((x,0)+(0,y)) = F(x,0)+F(0,y)$ und die Frage, ob sich die Gleichung $F(x,y) = 0$ nach $y = y(x)$ auflösen lässt, ist einfach die Frage nach der Umkehbarkeit von $F(0,.)$. Wir lernen im Satz über implizite Funktionen die Antwort auf die entsprechende Frage für differenzierbares F.

Problem: Seien V_0, V_1, W endlich-dimensionale Banachräume und $f : V_0 \times V_1 \to W$. *Unter welchen Voraussetzungen hat die Gleichung*

$$f(x,y) = 0 \tag{58}$$

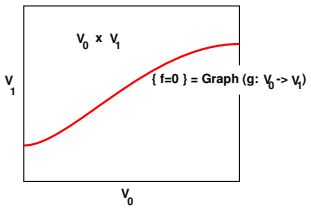
für jedes $x \in V_0$ genau eine Lösung $y \in V_1$?

Unter diesen Umständen gibt es dann eine eindeutig bestimmte Funktion $g : V_0 \to V_1$, für die für alle $x \in V_0$ gilt

$$f(x, g(x)) = 0. \tag{59}$$

Man sagt dann auch, dass (58) sich nach einer Funktion $y = g(x)$ eindeutig auflösen läßt oder dass g durch (59) *implizit definiert* wird.

Geometrisch bedeutet das, dass man das Niveau $f = 0$ als Graphen $\{(x, g(x)) \mid x \in V_0\}$ einer Funktion $g : V_0 \to V_1$ beschreibt, also durch V_0 parametrisiert: Jeder Punkt auf dem 0-Niveau liegt über genau einem Punkt von V_0.



Im Fall $W = \mathbb{R}^m$ hat f die Komponentenfunktionen $f_1, \ldots, f_m$. Man hat also m Gleichungen, die Dimension von W ist die Anzahl der gegebenen Gleichungen. Ebenso kann man die Dimension von V_1 als die Anzahl der gesuchten Unbekannten y_i ansehen. Es ist also wohl vernünftig, $\dim V_1 = \dim W$ zu wählen.

Beispiel 168. Sei $f = F : V_0 \times V_1 \to W$ linear und sei $\dim V_1 = \dim W$. Dann hat man

$$F(x,y) = F((x,0) + (0,y)) = F(x,0) + F(0,y),$$

d.h. F liefert zwei lineare Abbildungen

$$F(.\,,0) : V_0 \to W,$$
$$F(0,.) : V_1 \to W.$$

Dann ist (58) genau dann für jedes $x \in V_0$ eindeutig lösbar, wenn die lineare Abbildung $F(0,.) : V_1 \to W$ invertierbar ist. Die Gleichung

$$0 = F(x,y) = F(x,0) + F(0,y)$$

ist nämlich äquivalent zu

$$F(0,y) = -F(x,0).$$

Das ist höchstens dann eindeutig lösbar, wenn $F(0,.)$ injektiv ist. Nach der Dimensionsvoraussetzung ist in diesem Fall aber $F(0,.)$ bijektiv und die Gleichung tatsächlich für jedes x eindeutig lösbar. Man findet

$$g(x) = -F(0,.)^{-1}(F(x,0)).$$

Im Fall $V_0 = \mathbb{R}^n$ und $V_1 = W = \mathbb{R}^m$ ist F gegeben durch eine $m \times (n+m)-$Matrix der Form

$$(\underbrace{F^{(1)}}_{n} \mid \underbrace{F^{(2)}}_{m}),$$

und $F(0,.)$ wird repräsentiert durch die quadratische $m \times m-$Matrix $F^{(2)}$, die also invertierbar sein muß.

□

Wenn wir dieses Ergebnis von linearen Abbildungen auf differenzierbare Abbildungen verallgemeinern wollen, ist es plausibel, dass wir nur ein *lokales* Ergebnis erhalten. Experimentieren Sie ein bißchen mit dem Fall $V_0 = V_1 = W = \mathbb{R}$ und

$$f(x,y) := x - y^2.$$

Satz 169 (über implizite Funktionen). *Seien V_0, V_1, W endlich-dimensionale Banachräume, $G \subset V_0 \times V_1$ offen und $f : V_0 \times V_1 \supset G \to W$ stetig differenzierbar.*
Sei $(p,q) \in G$ mit

$$f(p,q) = 0, \tag{60}$$

$$D_{(p,q)}f(0,.) : V_1 \to W \text{ invertierbar.} \tag{61}$$

Beachte, dass damit $\dim V_1 = \dim W$.
Dann läßt sich

$$f(x,y) = 0 \tag{62}$$

in einer Umgebung von (p,q) eindeutig nach einer stetig differenzierbaren Abbildung $y = g(x)$ auflösen.

Genauer:
Es gibt offene Umgebungen U_0 von p in V_0 und U_1 von q in V_1 mit folgenden Eigenschaften:

(i) $U_0 \times U_1 \subset G$, und zu jedem $x \in U_0$ gibt es genau ein $y \in U_1$ mit

$$f(x,y) = 0.$$

(ii) Die nach (i) eindeutig bestimmte Funktion $g : U_0 \to U_1$ mit

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

ist stetig differenzierbar.

(iii) Für alle $x \in U_0$ ist $D_{(x,g(x))}f(0,.) : V_1 \to W$ invertierbar und für $v \in V_0$ ist

$$D_x g(v) = -\left(D_{(x,g(x))}f(0,.)\right)^{-1} \circ D_{(x,g(x))}f(v,0). \tag{63}$$

Bemerkung. Im Fall $V_0 = \mathbb{R}^n, V_1 = W = \mathbb{R}^m$ werden die linearen Abbildungen

$$D_{(x,y)}f(0,.) : \mathbb{R}^m \to \mathbb{R}^m \quad \text{bzw} \quad D_{(x,y)}f(.,0) : \mathbb{R}^n \to \mathbb{R}^m$$

repräsentiert durch die Matrizen

$$\left(\frac{\partial f_i}{\partial y_j}(x,y)\right)_{i,j=1,\ldots,m} \quad \text{bzw.} \quad \left(\frac{\partial f_i}{\partial x_j}(x,y)\right)_{i=1,\ldots,m;j=1,\ldots,n}$$

Beweis zum Satz über implizite Funktionen.

Die Idee. Die Gleichung $f(x,y)=0$ ist genau dann eindeutig nach y auflösbar, wenn dasselbe für die Gleichung

$$h(x,y) := (x, f(x,y)) = (x,0)$$

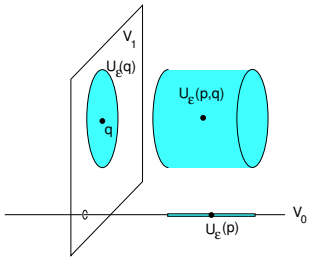
gilt. h erweist sich nach dem Umkehrsatz als lokal invertierbar, und die gesuchte Lösungsfunktion g ist dann gegeben durch

$$(x, g(x)) = h^{-1}(x,0),$$

also durch die zweite Komponente von $h^{-1}(.,0)$.

A. Vorbemerkung. Da $V_0 \times V_1$ endlich-dimensional ist, sind alle Normen äquivalent, und wir verwenden der Einfachheit halber die Norm

$$\|(v,w)\| = \sup(\|v\|, \|w\|).$$



Das hat den Vorteil, dass

$U_\epsilon((p,q)) = U_\epsilon(p) \times U_\epsilon(q)$ für $(p,q) \in V_0 \times V_1$ und $\epsilon > 0$.

Analog verfahren wir gleich mit dem Raum $V_0 \times W$.

B. Reduktion auf den Umkehrsatz. Wir setzen die obige Beweisidee um und definieren die Abbildung

$$h : V_0 \times V_1 \supset G \to V_0 \times W, (x,y) \mapsto (x, f(x,y))$$

zwischen gleich-dimensionalen Vektorräumen. Es gilt

$$D_{(x,y)}h(v,w) = (v, D_{(x,y)}f(v,w)), \tag{64}$$

und deshalb ist mit f auch h stetig differenzierbar. Weiter ist

$$D_{(p,q)}h \text{ invertierbar}, \tag{65}$$

denn

$$\begin{aligned} 0 = D_{(p,q)}h(v,w) &\underset{(64)}{\iff} v = 0 \text{ und } D_{(p,q)}f(v,w) = 0 \\ &\iff v = 0 \text{ und } D_{(p,q)}f(0,w) = 0 \iff v = 0 \text{ und } w = 0 \end{aligned}$$

nach Voraussetzung.

C. Anwendung des Umkehrsatzes. Nach dem Umkehrsatz gibt es $\epsilon > 0$, so dass

$$\begin{gathered} U := U_\epsilon(p) \times U_\epsilon(q) \subset G, \\ h|U \text{ injektiv}, \\ h(U) \text{ offen}, \\ (h|U)^{-1} \text{ stetig differenzierbar}. \end{gathered}$$

Da $h(U)$ offen und

$$(p,0) = (p, f(p,q)) = h(p,q) \in h(U) \subset V_0 \times W,$$

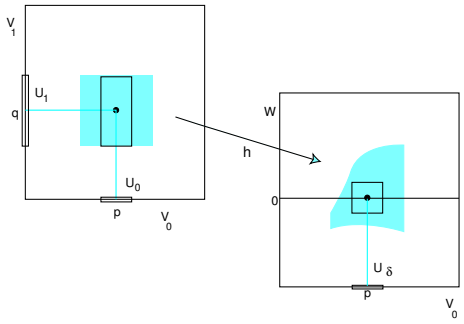
gibt es $\delta > 0$ mit $\delta < \epsilon$ und

$$U_\delta(p) \times U_\delta(0) = U_\delta((p,0)) \subset h(U) \subset V_0 \times W.$$

Wir setzen nun

$$U_0 := U_\delta(p), \quad U_1 := U_\epsilon(q),$$

und behaupten, dass diese das Gewünschte leisten.



Zu (i). Zunächst ist

$$U_0 \times U_1 \subset U_\epsilon(p) \times U_\epsilon(q) = U \subset G. \tag{66}$$

Ist $x \in U_0$, so ist $(x, 0) \in U_\delta((p, 0)) \subset h(U)$. Darum gibt es nach dem Umkehrsatz genau ein $(\tilde{x}, y) \in U$ mit $h(\tilde{x}, y) = (x, 0)$.

Offenbar ist

- $\tilde{x} = x$,
- $y \in U_1$ nach Definition von U und U_1, und
- $f(x, y) = 0$ nach Definition von h.

Also gibt es zu jedem $x \in U_0$ ein $y \in U_1$ mit $f(x, y) = 0$. Wir bezeichnen dieses y mit $g(x)$.

Sind $y_1, y_2 \in U_1$ mit $f(x, y_1) = 0 = f(x, y_2)$, so folgt $h(x, y_1) = (x, 0) = h(x, y_2)$, also $y_1 = y_2$.

Damit ist (i) bewiesen.

Zu (ii). Für $x \in U_0$ haben wir eben gezeigt, dass

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

Bezeichnen wir also mit $\pi : V_0 \times V_1 \to V_1, (v, w) \mapsto w$ die Projektion, so ist

$$g(x) = \pi \circ (h|U)^{-1}(x, 0).$$

Daher ist g stetig differenzierbar.

Zu (iii). Nach (65) ist $D_{(x,y)}h$ für $(x, y) \in U$ invertierbar. Nach (62) haben wir dann

$$D_{(x,y)}h(0, w) = (0, D_{(x,y)}f(0, w)) = 0 \iff w = 0.$$

Daraus folgt, dass für $(x, y) \in U$ auch $D_{(x,y)}f(0, .)$ injektiv und damit invertierbar ist. Insbesondere ist also für alle $x \in U_0$

$$D_{(x,g(x))}f(0, .) \text{ invertierbar.} \tag{67}$$

Nun differenzieren wir

$$\phi : U_0 \ni x \underset{\alpha}{\mapsto} (x, g(x)) \mapsto f(x, g(x))$$

nach der Kettenregel. Wir erhalten

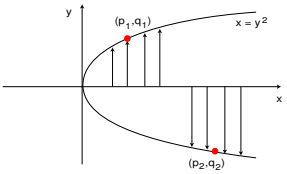
$$\begin{aligned} D_x\phi(v) &= D_{(x,g(x))}f(D_x\alpha(v)) \\ &= D_{(x,g(x))}f(v, D_xg(v)) \\ &= D_{(x,g(x))}f(v, 0) + D_{(x,g(x))}f(0, D_xg(v)). \end{aligned} \tag{68}$$

Andererseit ist $\phi = 0$, also $D_x\phi = 0$. Damit folgt aus (67) und (68) die Formel in (iii). □

Beispiel 170. Seien $V_0 = V_1 = W = \mathbb{R}$ und $f : V_0 \times V_1 = G \to \mathbb{R}, \quad (x, y) \mapsto x - y^2$. Dann ist

$$D_{(p,q)} f(v, w) = v - 2qw.$$

In $(p, q) = (0, 0)$ ist die Voraussetzung über die Invertierbarkeit der Ableitung also nicht erfüllt, wohl aber in allen Punkten (q^2, q) mit $q \neq 0$. In der Nähe dieser Punkte läßt sich $f^{-1}(\{0\}) = \{(x, y) \,|\, x - y^2 = 0\}$ also lokal als Graph schreiben.



□

3.3 Der Rangsatz

- Der Rangsatz beinhaltet in gewisser Weise die Quintessenz der linearen Approximation differenzierbarer Abbildungen.

In der Linearen Algebra betrachtet man folgendes Problem: Eine lineare Abbildung

$$F : V \to W$$

zwischen zwei $\mathbb{R}$-Vektorräumen der Dimensionen n und m kann man durch eine $(m \times n)$-Matrix darstellen, *nachdem man in V und W Basen gewählt hat*. Die Darstellungsmatrix hängt wesentlich von den gewählten Basen ab, und man kann fragen, ob man sie durch geschickte Wahl der Basen besonders einfach gestalten kann. Tatsächlich kann man immer die folgende Form erreichen

$$\begin{pmatrix} 1 & \dots & 0 & 0 & \dots & \dots & 0 \\ \vdots & \ddots & \vdots & \vdots & & & \vdots \\ 0 & \dots & 1 & 0 & \dots & \dots & 0 \\ 0 & \dots & 0 & 0 & \dots & \dots & 0 \\ \vdots & & \vdots & \vdots & & & \vdots \\ 0 & \dots & 0 & 0 & \dots & \dots & 0 \end{pmatrix}.$$

Die Zahl der Einsen ist dabei der Rang r der linearen Abbildung, d.h. die Dimension von $F(V)$. Dieses Resultat kann man auch so formulieren:

Ist $F : V \to W$ wie oben, so gibt es Isomorphismen $\Phi : V \to \mathbb{R}^n$ und $\Psi : W \to \mathbb{R}^m$, so dass

$$\Psi \circ F \circ \Phi^{-1} : \mathbb{R}^n \to \mathbb{R}^m$$

gegeben ist durch

$$\Psi \circ F \circ \Phi^{-1}(x_1, \dots, x_n) = (x_1, \dots, x_r, 0 \dots, 0).$$

Die Isomorphismen Φ und Ψ nennt man auch *Koordinaten*. In geeigneten Koordinaten sieht also jede lineare Abbildung vom Rang r aus wie

$$(x_1, \dots, x_n) \mapsto (x_1, \dots, x_r, 0 \dots, 0).$$

Wir übertragen das nun lokal auf C^k-Abbildungen, vgl. Definition 139. An die Stelle der *linearen* Koordinatenabbildung $\Phi : V \to \mathbb{R}^n$ tritt jetzt ein C^k-Diffeomorphismus

$$\Phi : V \supset U \to \tilde{U} \subset \mathbb{R}^n$$

(also eine bijektive C^k-Abbildung mit C^k-Inversem) zwischen offenen Umgebungen und U von $p \in G$ und $\tilde{U}$ von $\Phi(p) = 0$ in $\mathbb{R}^n$ und analog für Ψ. Diese Diffeomorphismen nennt man ebenfalls *(krummlinige) Koordinaten*.

Satz 171 (Rangsatz). *Seien V, W Banachräume der Dimensionen n und m, $G \subset V$ offen, und*

$$f : V \supset G \to W \quad k\text{-mal stetig differenzierbar, } 1 \leq k \leq +\infty.$$

f sei von konstantem Rang r, d.h. der Rang von $D_x f : V \to W$ sei $= r$ unabhängig von $x \in G$. Dann gilt: Zu jedem $p \in G$ gibt es C^k-Diffeomorphismen

$$\Phi : V \supset U_1 \to \tilde{U}_1 \subset \mathbb{R}^n$$

und

$$\Psi : W \supset U_2 \to \tilde{U}_2 \subset \mathbb{R}^m$$

offener Umgebungen von p bzw. $f(p)$ auf offene Umgebungen von $0 = \Phi(p)$ in $\mathbb{R}^n$ bzw. von $\Psi(f(p)) = 0 \in \mathbb{R}^m$, so dass

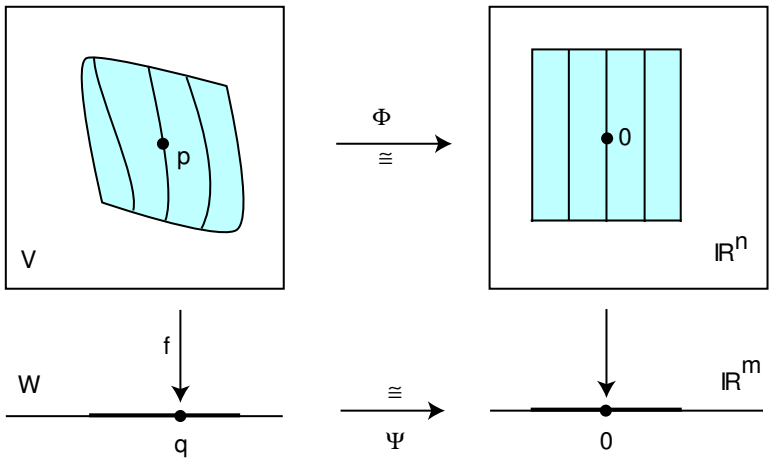
$$f(U_1) \subset U_2$$

und

$$\Psi \circ f \circ \Phi^{-1}(x_1, \ldots, x_n) = (x_1, \ldots, x_r, 0, \ldots, 0) \quad \text{für alle } x \in \tilde{U}_1.$$

C^k-Abbildungen von konstantem Rang r sind also in geeigneten C^k-Koordinaten von der Form

$$(x_1, \ldots, x_n) \mapsto (x_1, \ldots, x_r, 0 \ldots, 0).$$



Konvention. Um die Notation übersichtlich zu halten, schreiben wir zum Beispiel:

$$\textit{Sei } g : V \to W \textit{ ein lokaler Diffeomorphismus bei } p,$$

wenn g auf einer offenen Umgebung von $p \in V$ (nicht notwendig aber auf ganz V) definiert und C^k-differenzierbar ist, und eine (eventuell kleinere) offene Umgebung von p diffeomorph auf eine offene Umgebung von $g(p)$ in W abbildet.

Wegen des zu Beginn dieses Abschnittes angeführten Satzes aus der linearen Algebra genügt es, folgende Version des Rangsatzes zu beweisen:

Satz 172 (Rangsatz, 2. Version). *Seien V, W endlich-dimensionale Banachräume, sei $G \subset V$ offen, $p \in G$ und sei*

$$f : V \supset G \to W \in C^k \quad \text{und von konstantem Rang } r.$$

Dann gibt es lokale C^k-Diffeomorphismen

$$\begin{aligned} \phi &: V \to V \text{ bei } p \text{ mit } \phi(p) = p, \\ \psi &: W \to W \text{ bei } f(p) \text{ mit } \psi(f(p)) = f(p), \end{aligned}$$

für die auf einer offenen Umgebung von $0 \in V$

$$\psi \circ f \circ \phi^{-1} = D_p f$$

gilt. In geeigneten lokalen Koordinaten um p und $f(p)$ sieht f also aus wie seine Ableitung, d.h. wie eine lineare Abbildung vom Rang r.

Beweis. Vorbereitung. Durch Translationen (also C^∞-Diffeomorphismen) in V und W können wir erreichen, dass $p = 0$ und $f(p)$=0. Das setzen wir im folgenden voraus. Um den Beweis übersichtlich zu halten, benutzen wir die obige Konvention und verzichten auf die explizite Kontrolle der Definitionsbereiche.

Wir definieren

$$V_2 := \operatorname{Kern} D_0 f, \quad W_1 := \operatorname{Bild} D_0 f = D_0 f(V),$$

und wählen zu V_2 und W_1 komplementäre Unterräume, so dass also

$$V = V_1 \oplus V_2, \quad W = W_1 \oplus W_2.$$

Das Differential $D_0 f$ bildet dann also den r-dimensionalen Raum V_1 isomorph auf W_1 ab. Entsprechend der Zerlegung bezeichnen wir für $x \in V$ die Komponenten in V_1 bzw. V_2 mit x_1 bzw. x_2 und entsprechend für $y \in W$. Insbesondere ist $f = f_1 + f_2$ mit $f_i : G \to W_i$.

1. Schritt: Konstruktion von ϕ. Die Komponentenabbildungen sind linear, und deshalb ist

$$F := D_0(f_1|_{V_1}) = (D_0 f|_{V_1})_1 : V_1 \to W_1$$

ein Isomorphismus. Daher ist nach dem Umkehrsatz

$$f_1|_{V_1} : V_1 \to W_1$$

ein lokaler Diffeomorphismus[5]. Dann ist auch $F^{-1} \circ f_1 : V_1 \to V_1$ ein lokaler Diffeomorphismus. Wir definieren $\phi : V \to V$ durch

$$\phi(x) := F^{-1}(f_1(x)) + x_2 \text{ für } x = x_1 + x_2.$$

Dann ist

$$D_0\phi(v_1 + v_2) = F^{-1}(D_0 f_1(v_1 + v_2)) + v_2 = v_1 + v_2,$$

also

$$D_0\phi = \mathrm{id}_V, \tag{69}$$

und deshalb ist ϕ ein lokaler Diffeomorphismus.

[5] Genauer: "... ein lokaler C^k-Diffeomorphismus bei 0.", aber das unterdrücken wir in Zukunft: Alle unsere lokalen Diffeomorphismen und Abbildungen sind "bei 0" und k-mal stetig differenzierbar.

Aus der Definition folgt für die V_1-Komponente $\phi_1(x) = F^{-1} \circ f_1(x)$, also $\phi_1(\phi^{-1}(x)) = F^{-1} \circ f_1(\phi^{-1}(x))$ und

$$f_1(\phi^{-1}(x)) = F(x_1). \tag{70}$$

2. Schritt: Konstruktion von ψ. Nun definieren wir $\psi : W \to W$ durch

$$\psi(y_1 + y_2) = y_1 + y_2 - f_2 \circ \phi^{-1} \circ F^{-1}(y_1).$$

Dafür gilt

$$D_0\psi(w_1 + w_2) = w_1 + \underbrace{w_2 - D_0(f_2 \circ \phi^{-1} \circ F^{-1})(w_1)}_{\in W_2} = 0 \quad \iff w_1 = 0 \text{ und } w_2 = 0.$$

Also $D_0\psi$ injektiv und damit bijektiv, und ψ ist ein lokaler Diffeomorphismus. Wir erhalten

$$\begin{aligned}
\psi \circ f \circ \phi^{-1}(x) &= \psi(f \circ \phi^{-1}(x)) \\
&\underset{(70)}{=} \psi(F(x_1) + f_2 \circ \phi^{-1}(x)) \\
&= F(x_1) + f_2 \circ \phi^{-1}(x) - f_2 \circ \phi^{-1} \circ F^{-1}(F(x_1)) \\
&= F(x_1) + f_2 \circ \phi^{-1}(x) - f_2 \circ \phi^{-1}(x_1)
\end{aligned}$$

Wir sind also fertig, wenn wir zeigen können, dass für $x \in V$ nah bei 0

$$f_2 \circ \phi^{-1}(x) = f_2 \circ \phi^{-1}(x_1). \tag{71}$$

3. Schritt: Nachweis von (71). Das ist das eigentliche Herzstück des Beweises. Aus (70) folgt

$$\begin{aligned}
D_x f\left(D_{\phi(x)}\phi^{-1}(v_1 + v_2)\right) &= D_{\phi(x)}(f \circ \phi^{-1})(v_1 + v_2) \\
&= D_{\phi(x)}(f_1 \circ \phi^{-1} + f_2 \circ \phi^{-1})(v_1 + v_2) \\
&= F(v_1) + D_{\phi(x)}f_2(v_1 + v_2).
\end{aligned} \tag{72}$$

Wir betrachten nun die Projektion $\pi_1 : W \to W_1, y \mapsto y_1$. Aus der letzten Gleichung folgt

$$\pi_1\left(D_x f(V)\right) \supset \pi_1\left(D_x f(D_{\phi(x)}\phi^{-1}(V_1)\right) = F(V_1) = W_1.$$

Damit ist $\operatorname{Rang}(D_x f) \geq \dim W_1 = r$ für alle Punkte x nah bei $p = 0$.

Gäbe es $v_2 \in V_2$ mit $D_{\phi(x)}(f_2 \circ \phi^{-1})(v_2) = w_2 \neq 0$, so wäre

$$w_2 \underset{(72)}{=} D_x f(D_{\phi(x)}\phi^{-1}(v_2)) \in D_x f(V) \quad \text{ und } \quad \pi_1(w_2) = D_{\phi(x)}(\pi_1 \circ f_2 \circ \phi^{-1})(v_2) = 0.$$

Also wäre $w_2 \in \operatorname{Kern} \pi_1|_{D_x f(V)}$ und nach Linearer Algebra

$$\dim D_x f(V) = \dim \operatorname{Kern}(\pi_1) + \dim \operatorname{Bild}(\pi_1) \geq r + 1$$

im Widerspruch zur Rangvoraussetzung über f, die wir hier zu ersten Mal benutzen. Es folgt

$$D_x(f_2 \circ \phi^{-1})|_{V_2} = 0,$$

d.h. $f_2 \circ \phi^{-1}$ ist nach Korollar 147 lokal unabhängig von der V_2-Komponenten und

$$f_2 \circ \phi^{-1}(x) = f_2 \circ \phi^{-1}(x_1).$$

□

Wir halten noch ein Ergebnis aus diesem Beweis fest: Im letzten Schritt haben wir – ohne Benutzung der Konstanz des Ranges – gezeigt, dass für alle Punkte x nah bei p

$$D_p f(V) = W_1 \subset \pi_1(D_x f(V)),$$

Also ist der Rang von Df in Nachbarpunkten von p mindestens so groß wie in p. Man sagt, er ist *unterhalb-stetig*. Damit erhalten wir:

Lemma 173. *Für jede stetig differenzierbare Funktion ist der Rang unterhalb-stetig.*

Beispiel 174. Auf der Menge der reellen invertierbaren $n \times n$ Matrizen betrachten wir die Abbildung

$$f : M(n \times n, \mathbb{R}) \supset \mathbf{GL}(n, \mathbb{R}) \to M(n \times n, \mathbb{R})$$

mit $f(A) = AA^T$, wobei A^T die transponierte Matrix bezeichnet. Dafür gilt

$$D_A f(B) = BA^T + AB^T,$$

und diese Matrix ist symmetrisch(=selbstadjungiert)! Ist andererseits $C \in M(n \times n, \mathbb{R})$ symmetrisch, so folgt

$$D_A f\left(\frac{1}{2}C(A^{-1})^T\right) = \frac{1}{2}C(A^{-1})^T A^T + \frac{1}{2}A(A^{-1})C^T = C.$$

Also ist für alle $A \in \mathbf{GL}(n, \mathbb{R})$ das Bild von $D_A f$ der $\frac{n(n+1)}{2}$-dimensionale Raum aller symmetrischen Matrizen und f ist von konstantem Rang $\frac{n(n+1)}{2}$.

□

Korollar 175. *Sei $f : V \subset G \to W$ stetig differenzierbar.*

(i) Ist f eine Immersion, d.h. $D_p f$ für alle p injektiv, so ist f lokal injektiv.

(ii) Ist f eine Submersion, d.h. $D_p f$ für alle p surjektiv, so ist f eine offene Abbildung, d.h. f bildet offene Mengen in offene Mengen ab.

Beweis. Selbst. □

4 Mannigfaltigkeiten

- Wir lernen mit den Mannigfaltigkeiten eine Verallgemeinerung des Flächenbegriffs auf beliebige Dimension (und Kodimension) kennen.
- Beispiele sind vor allem die "Niveaus" von Abbildungen, wie die höherdimensionalen Sphären, aber auch viel abstraktere Räume, wie etwa die orthogonalen Matrizen.
- Der Tangentialraum ist eine lineare Approximation der Mannigfaltigkeit und ermöglicht, auch für Funktionen auf Mannigfaltigkeiten die Ableitung als lineare Abbildung zu definieren.
- Als Anwendung behandeln wir Extrema unter Nebenbedingungen.

Eine m-dimensionale Mannigfaltigkeit im Banachraums V ist eine Teilmenge $M \subset V$, die *in geeigneten krummlinigen Koordinaten (für V!) lokal so aussieht wie ein m-dimensionaler Untervektorraum*:

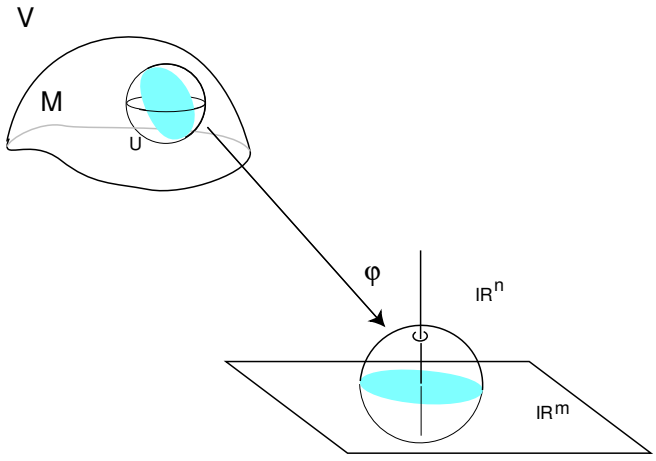
Definition 176. Seien $m, k \in \mathbb{N}$, $k > 0$. Eine Teilmenge $M \subset V$ eines n-dimensionalen Banachraums heißt eine *m-dimensionale C^k-(Unter)mannigfaltigkeit*, wenn es zu jedem Punkt $p \in M$ eine offene Umgebung U von p in V und einen C^k-Diffeomorphismus $\phi : U \to \phi(U)$ auf eine offene Teilmenge $\phi(U) \subset \mathbb{R}^n$ gibt, so dass gilt:

$$M \cap U = \phi^{-1}(\mathbb{R}^m \cap \phi(U)),$$

d.h.

$$M \cap U = \left\{ x \in U \;\middle|\; \phi_{m+1}(x) = \ldots = \phi_n(x) = 0 \right\}. \tag{73}$$

Dabei betrachten wir also $\mathbb{R}^m \subset \mathbb{R}^n$ als den Unterraum aller Punkte, deren letzte $n - m$ Koordinaten verschwinden.



Eine große Klasse von Beispielen liefert der folgende

Satz 177 (Gleichungsdefinierte Untermannigfaltigkeiten). *Seien V und W Banachräume endlicher Dimension. Seien $G \subset V$ offen und $g : G \to W \in C^k$, $k > 0$, vom konstanten Rang r, $0 < r < n := \dim V$ und $q \in g(G)$. Dann ist*

$$M := g^{-1}(\{q\})$$

eine $n - r$-dimensionale C^k-Mannigfaltigkeit.

Im Fall $g : \mathbb{R}^2 \to \mathbb{R}$ bzw. $g : \mathbb{R}^3 \to \mathbb{R}$ ist M also eine Niveaukurve bzw. -fläche. Gleichungsdefinierte Untermannigfaltigkeiten kann man also auch als *Niveaumannigfaltigkeiten* bezeichnen.

Beweis. Sei $p \in M$. Nach dem Rangsatz gibt es C^k-Diffeomorphismen

$$\Phi : V \supset U_1 \to \tilde{U}_1 \subset \mathbb{R}^n$$

und

$$\Psi : W \supset U_2 \to \tilde{U}_2 \subset \mathbb{R}^m \quad (m = \dim W)$$

offener Umgebungen von p bzw. $q = g(p)$ auf offene Umgebungen von $\Phi(p) = 0$ in $\mathbb{R}^n$ bzw. von $\Psi(q) = 0$ in $\mathbb{R}^m$, so dass

$$g(U_1) \subset U_2$$

und

$$\Psi \circ g \circ \Phi^{-1}(x_1, \ldots, x_n) = (x_1, \ldots, x_r, 0, \ldots, 0) \quad \text{für alle } x \in \tilde{U}_1. \tag{74}$$

Dann gilt für $p' \in U := U_1$

$$\begin{aligned} p' \in M &\iff g(p') = q \\ &\iff \Psi(g(p')) = 0 \\ &\iff \Psi \circ g \circ \Phi^{-1} \circ \Phi(p') = 0 \\ &\underset{(74)}{\iff} \Phi_1(p') = \ldots = \Phi_r(p') = 0. \end{aligned}$$

Bis auf die Nummerierung der Koordinatenfunktionen ist das die Definitionsgleichung (73). □

Beispiel 178. Die Abbildung

$$g : \mathbb{R}^{n+1} \to \mathbb{R}, \quad (x_1, \ldots, x_{n+1}) \mapsto (\sum_{i=1}^{n+1} x_i^2) - 1$$

hat die Funktionalmatrix

$$g'(x_1, \ldots, x_{n+1}) = 2(x_1, \ldots, x_{n+1}),$$

und weil $\mathbb{R}$ eindimensional ist, ist $D_x g$ surjektiv für alle $x \neq 0$. Daher ist die Einheitssphäre

$$S^n := \{x \mid g(x) = 0\} = \left\{(x_1, \ldots, x_{n+1}) \;\middle|\; \sum x_i^2 = 1\right\}$$

eine n-dimensionale C^∞-Untermannigfaltigkeit des $\mathbb{R}^{n+1}$.

□

Beispiel 179. Wir betrachten im n^2-dimensionalen Vektorraum der quadratischen n-reihigen Matrizen die Menge

$$\mathbf{O}(n) = \{A \in M(n \times n) \mid AA^t = E\}$$

der *orthogonalen Matrizen*. Nach Beispiel 174 ist das eine C^∞-Untermannigfaltigkeit der Dimension $n^2 - \frac{n(n+1)}{2} = \frac{n(n-1)}{2}$. Die orthogonalen Matrizen bilden außerdem bezüglich der Matrixmultiplikation eine Gruppe. Die Gruppenoperationen sind offenbar differenzierbar und $\mathbf{O}(n)$ ist eine sogenannte *Liegruppe*.

□

Definition 180 (Tangentialraum). Sei M eine m-dimensionale Mannigfaltigkeit im n-dimensionalen Banachraum V, sei $p \in M$ und $\phi : U \to \mathbb{R}^n$ ein Koordinatensystem dazu wie in der Definition 176. Dann ist also

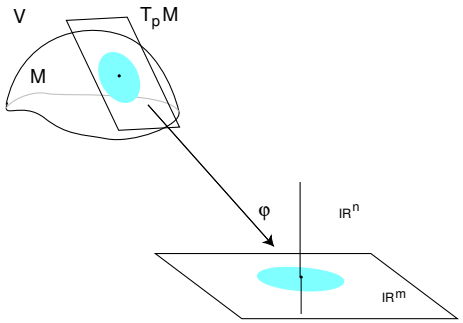
$$M \cap U = \phi^{-1}(\mathbb{R}^m \cap \phi(U)),$$

und wir definieren den Tangentialraum T_pM an M in p durch

$$\boxed{T_pM := D_{\phi(p)}\phi^{-1}(\mathbb{R}^m).}$$

Das ist also ein m-dimensionaler Vektorraum und eine lineare Approximation für M in der Nähe von p.

Auf dem nebenstehenden Bild ist eigentlich nicht T_pM dargestellt, sondern der nach p verschobene Tangentialraum, weil das unserer anschaulichen Vorstellung eher entspricht. Zum Rechnen ist natürlich der Vektorunterraum T_pM angenehmer als der parallele affine Unterrraum.



Damit der Tangentialraum wohldefiniert ist, müssen wir zeigen, dass er nicht vom gewählten Koordinatensystem abhängt. Sei also $\tilde{\phi} : \tilde{U} \to \mathbb{R}^n$ ein weiteres Koordinatensystem um p wie in der Definition 176. Wir können o.E. annehmen, dass $U = \tilde{U}$. Weil ϕ und $\tilde{\phi}$ lokale Diffeomorphismen sind, folgt aus $\tilde{\phi}^{-1}(\mathbb{R}^m \cap \tilde{\phi}(U)) = M \cap U = \phi^{-1}(\mathbb{R}^m \cap \phi(U))$, dass

$$\tilde{\phi} \circ \phi^{-1}(\mathbb{R}^m \cap \phi(U)) \subset \mathbb{R}^m$$

und deshalb

$$D_p\tilde{\phi} \circ D_{\phi(p)}\phi^{-1}(\mathbb{R}^m) \subset \mathbb{R}^m,$$

also

$$D_{\phi(p)}\phi^{-1}(\mathbb{R}^m) \subset \left(D_p\tilde{\phi}\right)^{-1}(\mathbb{R}^m) = D_{\tilde{\phi}(p)}(\tilde{\phi}^{-1})(\mathbb{R}^m).$$

Durch Vertauschen von ϕ und $\tilde{\phi}$ ergibt sich die umgekehrte Inklusion, also Gleichheit der Räume.

Beispiel 181. Ist $M = g^{-1}(\{q\}) \subset V$ eine gleichungsdefinierte Untermannnigfaltigkeit wie im Satz 177, so gilt für $p \in M$ und ein Koordinatensystem $\phi : U \to \mathbb{R}^n$ um p, dass $M \cap U = \phi^{-1}(\mathbb{R}^m \cap \phi(U))$, also $g \circ \phi^{-1}(\mathbb{R}^m \cap \phi(U)) = \{q\}$ und daher

$$D_pg(T_pM) = D_pg(D_{\phi(p)}\phi^{-1}(\mathbb{R}^m)) = 0.$$

Weil der Rang von g aber gerade $\dim V - \dim M$ ist, folgt

$$\boxed{T_pM = \operatorname{Kern} D_pg = (D_pg)^{-1}(\{0\}).}$$

Das ist die linearisierte Version von

$$M = g^{-1}(\{q\}).$$

□

Auf Mannigfaltigkeiten kann man "Analysis treiben", insbesondere die Differenzierbarkeit von Funktionen erklären. Das Differential an einer Stelle $p \in M$ ist dann eine lineare Abbildung *auf dem Tangentialraum* T_pM.

Wir betrachten dazu nur ein

Beispiel 182 (Extrema auf Mannigfaltigkeiten). Seien $G \subset V$ offen und $M \subset G$ ein Mannigfaltigkeit. Sei $f : G \to \mathbb{R}$ eine differenzierbare Funktion. Wir suchen lokale Extrema der Funktion $f|_M : M \to \mathbb{R}$. Sei $p \in M$ und $\phi : U \to \mathbb{R}^n$ ein Koordinatensystem um p wie in der Mannigfaltigkeitsdefinition, $\phi(p) = 0$. Dann ist $M \cap U = \phi^{-1}(\mathbb{R}^m)$. Hat also $f|_M$ in p ein lokales Extremum, so hat $f \circ \phi^{-1}|_{\mathbb{R}^m \cap \phi(U)}$ in 0 ein lokales Extremum. Deshalb ist

$$D_0(f \circ \phi^{-1})(\mathbb{R}^m) = D_p f(T_pM) = 0. \tag{75}$$

Notwendig für lokale Extrema der Einschränkung $f|_M$ von f ist also das Verschwinden der Einschränkung der Ableitung auf den Tangentialraum an M.

Ist $M = g^{-1}(\{q\})$ gleichungsdefiniert, so bedeutet (75), dass

$$\operatorname{Kern} D_p g \subset \operatorname{Kern} D_p f. \tag{76}$$

□

Die im Beispiel zuletzt betrachtete Situation ist unter dem Namen *Extremwerte unter Nebenbedingungen* berühmt. Sei $g : V \supset G \to W$ stetig differenzierbar und sei $q \in g(G)$. Sei weiter $f : G \to \mathbb{R}$ differenzierbar. Wir suchen lokale Extrema der Funktion f unter der Nebenbedingung $g = q$, d.h. lokale Extrema von $f|_{g^{-1}(\{q\})}$. Die Menge

$$\tilde{G} := \left\{ p \in G \;\middle|\; D_p g \text{ ist surjektiv} \right\}$$

ist nach Lemma 173 eine offene Teilmenge und $g^{-1}(\{q\}) \cap \tilde{G}$ eine Mannigfaltigkeit M der Dimension $\dim V - \dim W$. Hat $f|_{g^{-1}(\{q\})}$ ein lokales Extremum in $p \in M$, so gilt dort also die notwendige Bedingung (76). Typischerweise ist in den Anwendungen die Menge $g^{-1}(\{q\}) \setminus M$ der sogenannten singulären Punkte eine endliche Punktmenge, die man dann noch gesondert untersuchen muss.

Wir geben noch eine Variante von (76), die für die explizite Berechnung lokaler Extrema unter Nebenbedingungen hilfreich ist:

Es ist ein Standardproblem der lineare Algebra, den Kern einer linearen Abbildung zu bestimmen, also zu prüfen, ob (76) gilt. Aber meistens kennt man p gar nicht, sondern will die Extremalstellen erst finden. Das führt in der Regel auf nicht-lineare Gleichungssysteme, die schwer zu lösen sind. Bei der Bestimmung der Punkte vom zweiten Typ ist aber das folgende Lemma hilfreich:

Lemma 183 (Lagrange-Multiplikatoren). *Sei G offen in $V = \mathbb{R}^n$ und seien $f : G \to \mathbb{R}$ und $g = (g_1, \ldots, g_m) : G \to \mathbb{R}^m$ differenzierbar bzw. stetig differenzierbar.*

Dann ist (76) *äquivalent dazu, dass es reelle Zahlen $\lambda_1, \ldots, \lambda_m \in \mathbb{R}$ gibt (sog. Lagrange-Multiplikatoren), so dass für alle $j \in \{1, \ldots, n\}$*

$$\partial_j f(p) = \sum_{i=1}^{m} \lambda_i \partial_j g_i(p). \tag{77}$$

Beweis. Bezeichnen wir die Funktionalmatrizen mit $f'(p)$ bzw. $g'(p)$, die Transposition mit $(\ldots)^T$ und setzen wir $\lambda := (\lambda_1, \ldots, \lambda_n)$, so ist (77) äquivalent zu

$$f'(p) = \lambda g'(p) \quad \text{oder} \quad f'(p)^T = g'(p)^T \lambda^T.$$

Dieses lineare Gleichungssystem ist genau dann lösbar, wenn die erweiterte Matrix $(g'(p)^T, f'(p)^T)$ denselben Rang wie $g'(p)^T$ hat, wenn also die Matrix $\begin{pmatrix} g'(p) \\ f'(p) \end{pmatrix}$ denselben Rang wie die Matrix $g'(p)$ hat. Weil beide dieselbe Anzahl von Spalten haben, ist das genau dann der Fall, wenn die Kerne dieser beiden Matrizen gleiche Dimension haben. Weil aber

$$\operatorname{Kern} g'(p) \supset \operatorname{Kern} \begin{pmatrix} g'(p) \\ f'(p) \end{pmatrix} = \operatorname{Kern} g'(p) \cap \operatorname{Kern} f'(p),$$

ist das genau dann der Fall, wenn $\operatorname{Kern} D_p g \subset \operatorname{Kern} D_p f$. □

Rezept. Zur Bestimmung der Kandidaten p für Stellen lokaler Extrema von

$$f : \mathbb{R}^n \supset G \to \mathbb{R}$$

unter der Nebenbedingungen $g = 0$ mit $g : G \to \mathbb{R}^m$ sucht man

1. alle Punkte p mit $g(p) = 0$, in denen $D_p g(\mathbb{R}^n) \neq \mathbb{R}^m$ (singuläre Punkte),
2. alle Lösungen p, λ von

$$g_1(p) = 0, \ldots, g_m(p) = 0,$$
$$\partial_j f(p) = \sum_{i=1}^{m} \lambda_i \partial_j g_i(p), \quad j = 1, \ldots, m.$$

Das sind $m + n$ Gleichungen für die $n + m$ Variablen $p_1, \ldots, p_n, \lambda_1, \ldots, \lambda_m$.
Die λ's kann man wieder vergessen.

In typischen Problemen ist $m < n$, und die so gefundene Kandidatenmenge diskret oder sogar endlich.

Beispiel 184. Wir betrachten das Problem,

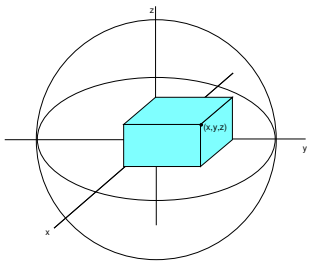
$$f(x, y, z) = xyz$$

unter der Nebenbedingung

$$x^2 + y^2 + z^2 \leq 1$$

zu maximieren, also das größte achsenparallele Quader in der Einheitskugel B zu finden. (Dessen Volumen ist dann $8|xyz|$, vgl. Abbildung.)

Beachten Sie, dass hier die Nebenbedingung durch eine *Ungleichung* gegeben ist. Die Kugel B ist kompakt, und weil f stetig ist, nimmt es auf B sein Maximum an. Das kann nicht in einem inneren Punkt geschehen, weil wir sonst alle Seiten des Quaders ein wenig vergrößern können und immer noch in der Kugel B bleiben:



Wenn $x^2 + y^2 + z^2 < 1$, ist auch $|x|^2 + |y|^2 + |z|^2 < 1$ und es gibt $\epsilon > 0$ mit

$$(|x| + \epsilon)^2 + (|y| + \epsilon)^2 + (|z| + \epsilon)^2 < 1.$$

Dafür ist aber

$$f(|x| + \epsilon, |y| + \epsilon, |z| + \epsilon) = (|x| + \epsilon)(|y| + \epsilon)(|z| + \epsilon) > |xyz| \geq xyz = f(x, y, z).$$

Ein anderes Argument liefert dasselbe Ergebnis: Läge das Maximum in einem inneren Punkt (x, y, z) so wäre

$$f'(x, y, z) = (yz\ xz\ xy) = (0\ 0\ 0).$$

Dann wäre aber $f(x, y, z) = 0$ das Maximum. Jedoch nimmt f offenbar auch positive Werte an.

Also wird das Maximum auf dem Rand angenommen, ist also ein Maximum unter der Nebenbedingung

$$g(x, y, z) := x^2 + y^2 + z^2 - 1 = 0.$$

Die Funktionalmatrix

$$g'(x, y, z) = (2x\ 2y\ 2z)$$

ist $\neq (0\ 0\ 0)$ für alle Punkte, die die Nebenbedingung erfüllen. Daher gibt es keine singulären Punkte.

Wir lösen nach dem Rezept:

$$x^2 + y^2 + z^2 - 1 = 0,$$

und

$$\begin{aligned} yz &= \lambda\, 2x, \\ xz &= \lambda\, 2y, \\ xy &= \lambda\, 2z. \end{aligned}$$

Multipliziert man diese letzteren Gleichungen mit x, y, z, addiert und verwendet die Nebenbedingung, so hat man

$$3xyz = 2\lambda.$$

Einsetzen von λ in die obigen Gleichungen liefert

$$x^2 = y^2 = z^2 = \frac{1}{3}$$

oder zwei Koordinaten sind 0, die dritte dann wegen der Nebenbedingung ± 1. Die letzteren Punkte liefern aber $f = 0$ und scheiden daher für ein Extremum aus. Mögliche Extrema liegen also in den Punkten

$$(\pm\frac{1}{\sqrt{3}}, \pm\frac{1}{\sqrt{3}}, \pm\frac{1}{\sqrt{3}})$$

mit voneinander unabhängigen Vorzeichen. Die entsprechenden Funktionswerte sind

$$f = \pm\frac{1}{3\sqrt{3}}.$$

Die positiven sind die Maxima, die negativen die Minima.

□

Als eine weiteres Anwendung für die Methode der Lagrange-Multiplikatoren beweisen wir im nächsten Beispiel die früher behauptete Abschätzung der l^p-Normen gegeneinander, vgl. Beispiel 101.

Beispiel 185. Wir erinnern an die Definition der l^p-Norm auf $\mathbb{R}^n$:

$$\|x\|_p := \left(\sum_{i=1}^{n} |x_i|^p\right)^{\frac{1}{p}}$$

Wir zeigen: Für $1 \leq p \leq q$ und alle $x \in \mathbb{R}^n$ gilt

$$\boxed{\|x\|_q \leq \|x\|_p \leq n^{\frac{1}{p}-\frac{1}{q}} \|x\|_q.} \tag{78}$$

Wir zeigen das durch vollständige Induktion über n.

$\underline{n = 1.}$ Trivial.

$\underline{n-1 \implies n.}$ Es genügt zu zeigen: Für alle $x = (x_1, \ldots, x_n)$ gilt

$$\sum_{i=1}^{n} |x_i|^q = 1 \quad \implies \quad 1 \leq \|x\|_p \leq n^{\frac{1}{p}-\frac{1}{q}}.$$

Die Voraussetzung impliziert $|x_i| \leq 1$ und deshalb $|x_i|^q \leq |x_i|^p$ für alle i. Also folgt die linke Ungleichung, wir müssen nur noch die rechte beweisen. Offenbar können wir uns dabei auf die kompakte Menge

$$\left\{ x = (x_1, \ldots, x_n) \;\middle|\; \sum_{i=1}^{n} x_i^q = 1 \text{ mit } x_i \geq 0 \text{ für alle } i \right\}$$

beschränken. Ist wenigstens ein $x_i = 0$, so liegt x in einem $\mathbb{R}^{n-1}$. Nach Induktionsvoraussetzung gilt für solche x also

$$\|x\|_p \leq (n-1)^{\frac{1}{p}-\frac{1}{q}} \leq n^{\frac{1}{p}-\frac{1}{q}}.$$

Daher genügt es zu zeigen, dass die differenzierbare Funktion $f(x) := \sum_{i=1}^{n} x_i^p$ auf der Menge $\left\{ x = (x_1, \ldots, x_n) \;\middle|\; x_i > 0 \right\}$ unter der Nebenbedingung

$$g(x) := \sum_{i=1}^{n} x_i^q = 1$$

das Maximum $\left(n^{\frac{1}{p}-\frac{1}{q}}\right)^p$ besitzt. Die notwendige Bedingung für ein Extremum ist die Existenz eines λ mit

$$\frac{\partial f}{\partial x_j} = p x_j^{p-1} = \lambda \frac{\partial g}{\partial x_j} = \lambda q x_j^{q-1}$$

oder

$$x_j^{p-q} = \lambda \frac{q}{p}$$

für alle j. Daraus folgt $x_1 = \ldots = x_n$, nach der Nebenbedingung also $x_1 = \ldots = x_n = \frac{1}{n^{1/q}}$. Der Funktionswert an dieser Stelle ist

$$f(x) = n \frac{1}{n^{p/q}} = n^{1-\frac{p}{q}} = \left(n^{\frac{1}{p}-\frac{1}{q}}\right)^p$$

und damit das (eindeutig bestimmte) Maximum von f.

□

5 Differentialgleichungen

5.1 Existenz- und Eindeutigkeit

- Was ist eine Differentialgleichung? Was ist eine Lösung einer Differentialgleichung?
- Die Differentialgleichung $y' = f$ hat für stetiges f auf einem Intervall viele Lösungen, nämlich die Stammfunktionen von f. Durch Vorgabe des Funktionswertes $y(a)$ an einer Stelle wird daraus *eine* eindeutige Lösung ausgewählt. Der Satz von Picard-Lindelöf verallgemeinert das zu einem Existenz- und Eindeutigkeitssatz für eine große Klasse von Differentialgleichungen.

Definition 186. Seien V ein endlich-dimensionaler Banachraum, $G \subset \mathbb{R} \times V$ offen, und

$$f : G \to V, (t, x) \mapsto f(t, x)$$

eine Abbildung.

(i) Die Gleichung

$$\dot{x} = f(t, x), \tag{79}$$

heißt *eine gewöhnliche Differentialgleichung erster Ordnung in expliziter Form, (kurz eine Differentialgleichung)* oder ein *dynamisches System*. In physikalischen Anwendungen ist t oft eine Zeitvariable, daher die Namenswahl und die Verwendung des Punktes $\dot{}$ anstelle des Strichs $'$.

(ii) Ist f in der ersten Variablen konstant, so kann man f auffassen als eine Abbildung $f : V \supset G \to V$. In diesem Fall nennt man

$$\dot{x} = f(x) \tag{80}$$

eine *autonome Differentialgleichung* oder ein *autonomes System*.

(iii) Eine auf einem Intervall $J \subset \mathbb{R}$ mit nicht-leerem Inneren $\overset{\circ}{J} \neq \emptyset$ definierte differenzierbare Funktion $x : J \to V$ heißt *eine Lösung von* (79), wenn für alle $t \in J$

$$(t, x(t)) \in G \text{ und } \dot{x}(t) = f(t, x(t)).$$

(iv) Sei $(t_0, x_0) \in G$. Das Gleichungssystem

$$\dot{x} = f(t, x), \quad x(t_0) = x_0 \tag{81}$$

heißt ein *Anfangswertproblem*.

(v) Eine Lösung $x : J \to V$ von (79) heißt *eine Lösung des Anfangswertproblems* (81), wenn außerdem $t_0 \in J$ und $x(t_0) = x_0$ ist.

(vi) Das Anfangswertproblem (81) heißt *eindeutig lösbar*, wenn es eine Lösung gibt, und wenn je zwei Lösungen $x_1 : J_1 \to V$ und $x_2 : J_2 \to V$ auf $J_1 \cap J_2$ übereinstimmen.

Beispiel 187. Die Bahn $x(t)$ eines Punktes der Masse m in einem zeit-, raum- und geschwindigkeitsabhängigen Kraftfeld im 3-dimensionalen Raum ist gegeben durch das Newtonsche Bewegungsgesetz

$$m\ddot{x} = F(t, x, \dot{x}).$$

Nach Einführung des Impulses $p = m\dot{x}$ als zusätzlicher Variabler nimmt dieses mit $V = \mathbb{R}^3 \times \mathbb{R}^3 = \mathbb{R}^6$ die Form (79) an:

$$\begin{aligned} \dot{x} &= m^{-1}p \\ \dot{p} &= F(t, x, m^{-1}p). \end{aligned}$$

□

Im einfachsten Fall hängt f nicht von x ab, sondern ist nur eine Funktion von t. Dann ist das Problem, die Differentialgleichung

$$\dot{x} = f(t)$$

zu lösen, einfach(?) das Problem, f zu integrieren. In der Theorie der Differentialgleichungen betrachtet man dieses Problem als „trivial".

Jenseits von diesem einfachsten Fall gibt es aber nur noch in sehr speziellen Fällen Verfahren zur Lösung einer Differentialgleichung im naiven Sinne. Das bedeutet, dass man im Einzelfall allenfalls mit speziellen Tricks Lösungen finden und/oder mit numerischen Verfahren berechnen kann. In anderen Fällen kann man sich eventuell wichtige Informationen über die Lösungen verschaffen, ohne diese explizit zu kennen. Zum Beispiel sind sicher alle Lösungen von $\dot{x} = 1 + x^2 + x^{14}$ streng monoton wachsend.

Gerade in dieser Situation ist es wichtig zu wissen, *ob* eine Differentialgleichung Lösungen hat und *wieviele* sie gegebenenfalls hat: Dann weiß man wenigstens, wonach man sucht. Eine weitere Hilfe können Informationen über die Struktur der Lösungsmenge liefern. Zum Beispiel kann man bei manchen Differentialgleichungen schon gefundene Lösungen benutzen, um weitere zu finden.

Diese Überlegungen unterstreichen die Bedeutung des folgenden Satzes:

Satz 188 (Existenz- und Eindeutigkeitssatz). *Sei $f : \mathbb{R} \times V \supset G \to V$ stetig auf der offenen Menge G, und sei $(t_0, x_0) \in G$. Dann ist das Anfangswertproblem*

$$\dot{x} = f(t, x), \quad x(t_0) = x_0 \tag{82}$$

lösbar. Ist f nach x stetig differenzierbar, so ist die Lösung eindeutig.
Insbesondere gibt es eine Lösung auf einem Intervall der Form $J =]t_0 - \epsilon, t_0 + \epsilon[$.

Bemerkung. Der Existenzsatz bei stetiger rechter Seite stammt von Peano, der Existenz- und Eindeutigkeitssatz bei zusätzlicher lokaler Lipschitz-Stetigkeit der rechten Seite bezüglich x von Picard und Lindelöf. Die hier gemachten Voraussetzungen sind etwas zu scharf, dafür bequem zu formulieren. Die Beweisidee werden wir im nächsten Abschnitt für den Spezialfall linearer Differentialgleichungen kennenlernen, den allgemeinen Fall und andere Details überlassen wir der Vorlesung über Gewöhnliche Differentialgleichungen.

(Gegen)beispiele. Die folgenden Beispiele sollen die Voraussetzungen des Satzes von Picard-Lindelöf illustrieren.

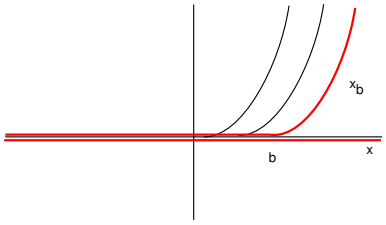
Beispiel 189. Für stetiges f besitzt (82) immer Lösungen (Satz von Peano), aber die sind nicht unbedingt eindeutig.

Das Anfangswertproblem

$$\dot{x} = \sqrt[3]{x^2}, \quad x(0) = 0$$

hat unendlich viele Lösungen, z.B. $x = 0$ oder, für $b > 0$, die Lösungen

$$x_b(t) = \begin{cases} 0 & \text{für } t \leq b \\ \frac{1}{27}(t-b)^3 & \text{für } t \geq b. \end{cases}$$



□

Beispiel 190. Für unstetige rechte Seite muß (82) keine Lösung haben. Ist $G = \mathbb{R} \times \mathbb{R}$ und

$$f(t,x) := \begin{cases} 0 \text{ für } t \leq 0 \\ 1 \text{ für } t > 0 \end{cases},$$

so hat (82) mit der Anfangsbedingung $x(0) = 0$ keine Lösung, weil die Ableitung einer differenzierbaren Funktion keine Sprungstellen hat (Satz von Dini).

□

Der Satz 188 macht keine Ausage über die maximale Größe des Definitonsbereichs einer Lösung.

Beispiel 191. Die Funktion $f(t,x) := 1 + x^2$ ist auf ganz $\mathbb{R} \times \mathbb{R}$ definiert. Die Lösung von

$$\dot{x} = 1 + x^2, \quad x(0) = 0$$

existiert aber auf keinem größeren Intervall als $]-\frac{\pi}{2}, \frac{\pi}{2}[$, wo sie durch $x = \tan$ gegeben ist.

□

5.2 Lineare Differentialgleichungen.

- Lineare Differentialgleichungen sind einfach lineare Gleichungssysteme, wenn auch auf Vektorräumen aus Funktionen. Darum kennen wir die Struktur des Lösungsraumes aus der linearen Algebra: er ist ein affiner Raum, gegeben durch eine konkrete Lösung, plus die Lösung des zugehörigen homogenen Systems, also den Kern der linearen Abbildung, die dem System zugrunde liegt.
- Weil die beteiligten Funktionenräume aber unendliche Dimension haben, sind die Existenz von Lösungen und die Dimension des Kerns nicht so klar. Wir klären das im nächsten Abschnitt.

Definition 192. Eine lineare Differentialgleichung 1. Ordnung auf einem offenen Intervall $J \subset \mathbb{R}$ ist eine Differentialgleichung der Form

$$\dot{x} = F(t)x + g(t), \tag{83}$$

wobei $F : J \ni t \mapsto F(t) \in L(V,V)$ und $g : J \to V$ stetig sind. Jede Lösung x ist dann offenbar stetig differenzierbar. Wir bezeichnen mit $C^k(J,V)$ den Vektorraum der k-mal stetig differenzierbaren Abbildungen von J nach V und definieren

$$\mathrm{L} : C^1(J,V) \to C^0(J,V), x \mapsto \dot{x} - F(t)x.$$

Dann ist L eine lineare Abbildung. Die auf J definierten Lösungen der *zugehörigen homogenen linearen Differentialgleichung*

$$\dot{x} = F(t)x \tag{84}$$

bilden deshalb einen Vektorraum Kern(L), und alle Lösungen von (83) auf J erhält man, indem man zu einer Lösung $\tilde{x}$ von (83) alle Lösungen der homogenen Gleichung addiert.

Beispiel 193. Wir betrachten das Gleichungssystem

$$\begin{aligned} \dot{x}_1 = &\quad x_1 + \quad 3x_2 + \quad 2\cos^2 t \\ \dot{x}_2 = &\quad 3x_1 + \quad x_2 + \quad 2\sin^2 t \end{aligned}$$

oder

$$\frac{d}{dt}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 & 3 \\ 3 & 1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{pmatrix} 2\cos^2 t \\ 2\sin^2 t \end{pmatrix}$$

Das zugehörige homogene System

$$\frac{d}{dt}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 & 3 \\ 3 & 1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

hat Lösungen:

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = a_1\begin{pmatrix} e^{4t} \\ e^{4t} \end{pmatrix} + a_2\begin{pmatrix} e^{-2t} \\ -e^{-2t} \end{pmatrix}, \qquad a_1, a_2 \in \mathbb{R}.$$

(Nachrechnen! Im Abschnitt 5.2.2 wird erklärt, wie man die finden kann.) Wir zeigen gleich, dass das *alle* Lösungen sind. *Eine* Lösung für das inhomogene System werden wir weiter unten konstruieren, nämlich

$$\frac{1}{4}\begin{pmatrix} \sin 2t + \cos 2t - 1 \\ -\sin 2t - \cos 2t - 1 \end{pmatrix}.$$

Die „ allgemeine Lösung“ der inhomogenen Gleichung ist daher

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \frac{1}{4}\begin{pmatrix} \sin 2t + \cos 2t - 1 \\ -\sin 2t - \cos 2t - 1 \end{pmatrix} + a_1\begin{pmatrix} e^{4t} \\ e^{4t} \end{pmatrix} + a_2\begin{pmatrix} e^{-2t} \\ -e^{-2t} \end{pmatrix}, \qquad a_1, a_2 \in \mathbb{R}.$$

□

5.2.1 Der Hauptsatz über lineare Differentialgleichungen

- Wir lernen, wie man eine Differentialgleichung in eine Integralgleichung umschreibt und diese mit Hilfe des Banachschen Fixpunktsatzes löst.

Satz 194 (Hauptsatz über lineare Differentialgleichungen). *Sei $n := \dim V$. Mit den obigen Bezeichnungen gilt für $t_0 \in J, x_0 \in V$:*

(i) Das Anfangswertproblem

$$\begin{aligned} \dot{x} &= F(t)x + g(t) \\ x(t_0) &= \xi_0 \end{aligned} \tag{85}$$

hat genau eine auf ganz J definierte Lösung.

(ii) Der Lösungsraum Kern(L) *der zugehörigen homogenen Gleichung*

$$\dot{x} = F(t)x \tag{86}$$

ist n-dimensional.
Funktionen $x_1, \dots, x_n \in$ Kern(L) sind genau dann linear unabhängig, wenn ihre Werte $x_1(t), \dots, x_n(t) \in V$ an einer (und dann an jeder) Stelle $t \in J$ linear unabhängig sind. In diesem Fall ist jede Lösung der homogenen Gleichung (86) von der Form

$$x(t) = \sum_{i=1}^{n} c_i x_i(t), \quad c_i \in \mathbb{R}, \tag{87}$$

und für jedes n-Tupel $(c_1, \dots, c_n)$ reeller Zahlen ist das eine Lösung.

(iii) Ist $(x_1, \dots, x_n)$ eine Basis von von Kern L*, so ist jede Lösung von (85) von der Form*

$$x(t) = \sum_{i=1}^{n} c_i(t) x_i(t), \tag{88}$$

mit geeigneten Funktionen $c_1, \dots, c_n \in C^1(J, \mathbb{R})$.

Die Funktion (88) ist genau dann eine Lösung von (85), wenn die c_i die folgende Differentialgleichung erfüllen:

$$\sum_{i=1}^{n} \dot{c}_i(t) x_i(t) = g(t). \tag{89}$$

Variation der Konstanten. *Ist $x_1, \dots, x_n$ eine Basis von* Kern(L)*, so erhält man durch Lösen des gewöhnlichen linearen Gleichungssystems (89) die $\dot{c}_i(t)$. Anschließende Integration liefert die $c_i(t)$ und damit eine Lösung von (85) in der Form (88). Diese Methode nennt man* Variation der Konstanten.

Beweis. Die Idee. Ist $x : J \to V$ eine Lösung des Anfangswertproblems, so gilt nach Integration der Differentialgleichung:

$$x(t) = \underbrace{\xi_0 + \int_{t_0}^{t} (F(\tau)x(\tau) + g(\tau))d\tau}_{=:\Phi(x)(t)}. \tag{90}$$

Die Lösung x ist also ein Fixpunkt von Φ. Umgekehrt folgt aus $x = \Phi(x)$ mit stetigem x sofort die (sogar stetige) Differenzierbarkeit von x, die Anfangsbedingung $x(t_0) = \xi_0$ und durch Differenzieren

$$\dot{x}(t) = F(t)x(t) + g(t).$$

Wir wollen deshalb Φ als Abbildung auf einem metrischen Raum stetiger Funktionen x auffassen und mit dem Banachschen Fixpunktsatz zeigen, dass Φ genau einen Fixpunkt hat. Dann ist das Anfangswertproblem eindeutig lösbar. Überdies liefert der Fixpunktsatz ein Iterationsverfahren zur Berechnung der Lösung.

Dabei braucht man allerdings offenbar das Integral von Funktionen mit Werten in einem Banachraum V mit Norm $\|.\|_V$, das ich hier ebenso wenig erklären will, wie die Formel

$$\left\| \int_{t_0}^{t} f(t)dt \right\|_V \leq \left| \int_{t_0}^{t} \|f(t)\|_V dt \right|,$$

die wir im Beweis benötigen. (Die Absolutstriche auf der rechten Seite braucht man, wenn man auch $t < t_0$ zulassen will.) Sie können sich einfach vorstellen, dass $V = \mathbb{R}^n$ und komponentenweise integrieren, oder im Anhang Genaueres darüber finden.

Der Standardbeweis des Existenz- und Eindeutigkeitssatzes von Picard und Lindelöf benutzt dieselbe Idee.

Zu (i). 1. Schritt: Existenz und Eindeutigkeit für kompaktes J. Wir nehmen zunächst an, dass J ein kompaktes Intervall ist. Wir wählen ein $\alpha \geq 0$, über das wir später verfügen wollen, und erklären auf dem Vektorraum $C^0(J, V)$ der auf J stetigen Abbildungen nach $(V, \|.\|)$ eine Norm durch

$$\|x\|_\alpha := \sup_{t \in J} \|x(t)\| e^{-\alpha|t-t_0|}.$$

Offenbar ist $\|x\|_0$ die normale Supremumsnorm. Überlegen Sie, dass man auch für $\alpha > 0$ eine Norm erhält, für die

$$\min_{t \in J} \left(e^{-\alpha|t-t_0|} \right) \|x\|_0 \leq \|x\|_\alpha \leq \max_{t \in J} \left(e^{-\alpha|t-t_0|} \right) \|x\|_0.$$

Die α-Normen sind also zur Supremumsnorm äquivalent, und $(C^0(J, V), \|.\|_\alpha)$ ist mit jeder α-Norm vollständig. Für $x \in C^0(J, V)$ definieren wir nun $\Phi : C^0(J, V) \to C^0(J, V)$ durch

$$\Phi(x)(t) := \xi_0 + \int_{t_0}^{t} (F(\tau)x(\tau) + g(\tau))d\tau.$$

Weil F stetig ist, ist $C := \sup_{t \in J} \|F(t)\| < \infty$. Damit gilt für $x, y \in C^0(J, V)$, $t \in J$ und positives α

$$\begin{aligned}
\|\Phi(x)(t) - \Phi(y)(t)\| = \left\| \int_{t_0}^{t} F(\tau)(x(\tau) - y(\tau))d\tau \right\| &\leq \left| \int_{t_0}^{t} \|F(\tau)(x(\tau) - y(\tau))\| \, d\tau \right| \\
&\leq \left| \int_{t_0}^{t} C \, \|(x(\tau) - y(\tau))\| \, e^{-\alpha|\tau-t_0|} e^{\alpha|\tau-t_0|} d\tau \right| \\
&\leq C\|x - y\|_\alpha \left| \int_{t_0}^{t} e^{\alpha|\tau-t_0|} d\tau \right| \leq C\|x - y\|_\alpha \frac{e^{\alpha|t-t_0|}}{\alpha}.
\end{aligned}$$

Die letzte Ungleichung ergibt sich im Fall $t < t_0$ wie folgt:

$$\left| \int_{t_0}^{t} e^{\alpha|\tau - t_0|} d\tau \right| = \int_{t}^{t_0} e^{\alpha(t_0 - \tau)} d\tau = e^{\alpha t_0} \left. \frac{e^{-\alpha\tau}}{-\alpha} \right|_{t}^{t_0} = e^{\alpha t_0} \frac{e^{-\alpha t} - e^{-\alpha t_0}}{\alpha} \leq \frac{e^{\alpha|t - t_0|}}{\alpha}.$$

Den Fall $t_0 \leq t$ können Sie selbst machen. Wir erhalten

$$\|\Phi(x)(t) - \Phi(y)(t)\| e^{-\alpha|t - t_0|} \leq \frac{C}{\alpha} \|x - y\|_\alpha$$

und damit

$$\|\Phi(x) - \Phi(y)\|_\alpha \leq \frac{C}{\alpha} \|x - y\|_\alpha.$$

Wählen wir also $\alpha > C$. so ist Φ kontrahierend und besitzt nach dem Bachnachschen Fixpunktsatz 46 genau einen Fixpunkt $x \in C^0(J, V)$. Also besitzt das Anfangswertproblem genau eine Lösung auf J.

2. Schritt: Existenz und Eindeutigkeit für nicht-kompaktes J. Ist J nicht kompakt, so gibt es eine Folge kompakter Intervalle $(J_i)_{i \in \mathbb{N}}$ mit

$$t_0 \in J_0 \subset J_1 \subset \ldots$$

und

$$J = \bigcup_{i=0}^{\infty} J_i.$$

Dazu gibt es eine Folge eindeutig bestimmter Lösungen $x_i : J_i \to V$ des Anfangswertproblems, für die also gilt $x_{i+1}|_{J_i} = x_i$. Setzt man deshalb $x(t) := x_i(t)$, falls $t \in J_i$, so definiert das eine Funktion $x : J \to V$, die das Anfangswertproblem löst.

Sind schließlich $x_1, x_2 : J \to V$ zwei Lösungen des Anfangswertproblems und ist $t \in J \setminus \{t_0\}$, so sei I das kompakte Intervall mit Endpunkten t_0 und t. Dann sind $x_1|_I$ und $x_2|_I$ Lösungen des Anfangswertproblems, nach dem 1. Schritt ist also $x_1|_I = x_2|_I$ und insbesondere $x_1(t) = x_2(t)$. Daraus folgt die Eindeutigkeit.

Zu (ii). Sei $t_1 \in J$. Die Abbildung

$$\mathrm{Kern(L)} \to V, x \mapsto x(t_1)$$

ist linear. Weil das Anfangswertproblem mit der Anfangsbedingung $x(t_1) = x_1$ genau eine Lösung hat, ist diese Abbildung also ein Isomorphismus. Daraus folgt die Behauptung.

Zu (iii). Sei $x_1, \ldots, x_n$ eine Basis von Kern(L) und seien $c_1, \ldots, c_n \in C^1(J, V)$. Wir setzen

$$x(t) = \sum c_i(t) x_i(t).$$

Dann gilt

$$\begin{aligned} \dot{x} - F(t)x(t) &= \frac{d}{dt} \sum c_i(t) x_i(t) - F(t) \sum c_i(t) x_i(t) \\ &= \sum \dot{c}_i x_i + \sum c_i(t)(\dot{x}_i - F(t) x_i(t)) \\ &= \sum \dot{c}_i x_i. \end{aligned}$$

Also ist x genau dann eine Lösung von $\dot{x} = F(t)x + g(t)$, wenn für alle $t \in J$

$$\sum \dot{c}_i(t) x_i(t) = g(t).$$

Umgekehrt sind für jedes $t \in J$ die Vektoren $x_1(t), \ldots, x_n(t) \in V$ linear unabhängig, und daher gibt es eindeutig bestimmte $\dot{c}_i$, die dieses inhomogene lineare Gleichungssystem lösen. Schreibt man das Gleichungssystem in Koordinaten aus, so ist die Lösung eine rationale Funktion in den Koeffizienten. Die sind aber stetig, und daher sind auch die $\dot{c}_i$ stetige Funktionen. Durch Integration findet man C^1-Funktionen c_i und damit eine Lösung der inhomogenen Gleichung. Jede andere unterscheidet sich davon nur durch eine Linearkombination der x_i mit konstanten Koeffizienten, ist also auch von der Form $x(t) = \sum c_i(t)x_i(t)$. □

Bemerkung. Nach dem Hauptsatz ist das Problem, eine lineare Differentialgleichung zu lösen, reduziert auf den homogenen Fall. Wenn $F(t) = A \in L(V, V)$ unabhängig von t ist, spricht man von einer *linearen Differentialgleichung mit konstanten Koeffizienten*. In diesem Fall kann man eine Lösungsbasis für die homogene Gleichung mit Methoden der linearen Algebra bestimmen, vgl. den nächsten Abschnitt 5.2.2.

Beispiel 195. Wir kommen zurück auf das Beispiel 193. Die Lösungen

$$\begin{pmatrix} e^{4t} \\ e^{4t} \end{pmatrix} \text{ und } \begin{pmatrix} e^{-2t} \\ -e^{-2t} \end{pmatrix}$$

der homogenen Gleichung sind linear unabhängig, weil sie an der Stelle 0 linear unabhängig sind. Sie bilden also eine Lösungsbasis für die homogene Differentialgleichung.

Variation der Konstanten mit dem Ansatz

$$x_s(t) = c_1(t)x_1(t) + c_2(t)x_2(t)$$

führt auf das Gleichungssystem

$$\begin{pmatrix} e^{4t} & e^{-2t} \\ e^{4t} & -e^{-2t} \end{pmatrix} \begin{pmatrix} \dot{c}_1 \\ \dot{c}_2 \end{pmatrix} = \begin{pmatrix} 2\cos^2 t \\ 2\sin^2 t \end{pmatrix}.$$

Lösen liefert

$$\dot{c}_1(t) = e^{-4t}, \qquad \dot{c}_2(t) = (\cos^2 t - \sin^2 xt)e^{2t} = \cos 2t\, e^{2t},$$

und Integration

$$c_1(t) = -\frac{1}{4}e^{-4t}, \qquad c_2(t) = \frac{1}{4}(\sin 2t + \cos 2t)e^{2t}.$$

Damit erhalten wir die früher schon angegebene Lösung

$$\begin{aligned} x_s(t) &= -\frac{1}{4}e^{-4t}\begin{pmatrix} e^{4t} \\ e^{4t} \end{pmatrix} + \frac{1}{4}(\sin 2t + \cos 2t)e^{2t}\begin{pmatrix} e^{-2t} \\ -e^{-2t} \end{pmatrix} \\ &= \frac{1}{4}\begin{pmatrix} \sin 2t + \cos 2t - 1 \\ -\sin 2t - \cos 2t - 1 \end{pmatrix} \end{aligned}$$

der inhomogenen Gleichung.

□

5.2.2 Lineare Differentialgleichungen mit konstanten Koeffizienten

- Homogene lineare Differentialgleichungen mit von t unabhängiger rechter Seite kann man explizit lösen.
- Wir lernen im Vorübergehen die Matrix-Exponential-Lösung kennen und betrachten dann genauer die Eigenwertmethode zur Lösung.
- Diese ist besonders einfach für diagonalisierbare Endomorphismen, aber wir diskutieren auch, wie man den allgemeinen Fall bewältigt.

Nach dem vorangehenden Abschnitt gibt es mit der Variation der Konstanten eine Methode zur Lösung inhomogener linearer Differentialgleichungen, wenn man die *zugehörige homogene* Differentialgleichung $\dot{x} = F(t)x$ vollständig gelöst hat. Für das letztere Problem aber gibt es keine allgemeines Verfahren. Nur im Fall konstanter Funktion F kann man eine Lösung explizit hinschreiben. Das wollen wir jetzt erläutern.

Im Fall $V = \mathbb{R}$ hat die Differentialgleichung

$$\dot{x} = ax$$

die Lösungen $x(t) = x(0)\exp(ta)$. Ist nun V ein endlich-dimensionaler Banachraum und $F(t) = A$ eine konstante lineare Abbildung von V in sich, so kann man entsprechend den Ansatz $x(t) = \exp(tA)$ machen. Aber was soll $\exp(tA)$ überhaupt bedeuten?

Nun, Endomorphismen (quadratische Matrizen) kann man miteinander multiplizieren und damit sind die Potenzen A^k definiert. Dann ist für jedes $t \in \mathbb{R}$ die Folge $\left(\sum_{j=0}^{n} \frac{t^j}{j!} A^j\right)_{n\in\mathbb{N}}$ wohldefiniert und konvergent im Banachraum $L(V,V)$, eine konvergente Potenzreihe in $L(V,V)$ gewissermaßen[6]. Sie definiert eine differenzierbare Funktion

$$X : \mathbb{R} \to L(V,V), t \mapsto \exp(tA)$$

mit $\dot{X} = AX$. Und für jeden Vektor $v \in V$ ist dann nach der Produktregel

$$x : t \mapsto X(t)v = \exp(tA)v$$

eine Lösung von

$$\dot{x} = Ax. \tag{91}$$

Mit einer Basis $v_1, \ldots, v_n$ von V erhält man eine Basis

$$x_1(t) = \exp(tA)v_1, \ldots, x_n(t) = \exp(tA)v_n$$

für den Lösungsraum von (91), weil die Funktionswerte für $t = 0$ linear unabhängig sind.

Die Berechnung der verallgemeinerten Exponentialfunktion ist natürlich nicht so leicht ist, aber die Lineare Algebra bietet Hilfe. Wir bezeichnen mit $E : V \to V$ im folgenden die Identität bzw. die Einheitsmatrix. Wenn das charakteristische Polynom $\det(A - \lambda E)$ von A in Linearfaktoren zerfällt, besitzt A eine *Jordansche Normalform*, was in anderen Worten bedeutet: Es gibt eine Basis Basis $v_1, \ldots, v_n$ von V aus *Hauptvektoren* von A. Zu jedem $i \in \{1, \ldots, n\}$ gibt es einen Eigenwert λ_i und ein $k_i \in \mathbb{N} \setminus \{0\}$, so dass

$$(A - \lambda_i E)^{k_i} v_i = 0.$$

Im Idealfall ist $k_i = 1$ für alle i, d.h. A ist diagonalisierbar und die v_i bilden eine Basis aus *Eigenvektoren*.

[6] Diese Konstruktion ist nicht ganz ohne: Weil die Matrixmultiplikation nicht kommutativ ist, ist zum Beispiel meistens $\exp(A + B) \neq \exp(A)\exp(B)$. Gleichheit gilt allerdings, wenn $AB = BA$.

Nun kann man zeigen:

$$\exp(tA) = \exp(t\lambda E + t(A - \lambda E)) = \exp(t\lambda E)\exp(t(A - \lambda E)) = e^{\lambda t}\sum_{j=0}^{\infty}\frac{t^j}{j!}(A - \lambda E)^j.$$

Insbesondere ist also

$$x_i(t) = \exp(tA)v_i = e^{\lambda_i t}\sum_{j=0}^{k_i-1}\frac{t^j}{j!}(A - \lambda E)^j v_i,$$

und man bekommt eine Lösungsbasis mittels endlicher Summen. Ist A diagonalisierbar und $v_1, \ldots, v_n$ eine Basis aus Eigenvektoren mit zugehörigen Eigenwerten $\lambda_1, \ldots, \lambda_n$, so ist also

$$x_1(t) = e^{\lambda_1 t}v_1, \ldots, x_n(t) = e^{\lambda_n t}v_n$$

eine Lösungsbasis für $\dot{x} = Ax$.

In den vorstehenden Überlegungen sind wir mit vektorwertigen Potenzreihen relativ großzügig umgegangen. Wir geben nun einen strengen Beweis für den folgenden

Satz 196. *Seien $A \in L(V, V)$ ein Endomorphismus, $k \in \mathbb{N} \setminus \{0\}$ und $v \in V$ ein Hauptvektor der Stufe k von A zum Eigenwert $\lambda \in \mathbb{R}$, d.h. es gelte*

$$(A - \lambda E)^k v = 0.$$

Dann ist

$$x(t) := e^{\lambda t}\sum_{j=0}^{k-1}\frac{t^j}{j!}(A - \lambda E)^j v$$

eine Lösung von $\dot{x} = Ax$. Diese Lösung ist also von der Form $e^{\lambda t}v(t)$, wobei $v(t)$ ein Polynom in t mit vektoriellen Koeffizienten ist.

Beweis. Nach Voraussetzung ist

$$x(t) = e^{\lambda t}\sum_{j=0}^{k}\frac{t^j}{j!}(A - \lambda E)^j v,$$

wobei wir jetzt bis k summieren. Das macht ja keinen Unterschied. Wir finden

$$\begin{aligned}
\dot{x}(t) &= \lambda x(t) + e^{\lambda t}\sum_{j=1}^{k}\frac{t^{j-1}}{(j-1)!}(A - \lambda E)^j v \\
&= \lambda x(t) + e^{\lambda t}(A - \lambda E)\sum_{j=1}^{k}\frac{t^{j-1}}{(j-1)!}(A - \lambda E)^{j-1} v \\
&= \lambda x(t) + (A - \lambda E)e^{\lambda t}\sum_{j=0}^{k-1}\frac{t^j}{j!}(A - \lambda E)^j v \\
&= Ax(t).
\end{aligned}$$

□

Korollar 197. *Besitzt V eine Basis $v_1, \ldots, v_n$ aus Hauptvektoren von $A \in L(V,V)$ der Stufen $k_1, \ldots, k_n$ und zugehörigen Eigenwerten $\lambda_1, \ldots, \lambda_n$, so liefern die Funktionen*

$$x_i(t) := e^{\lambda_i t} \left(\sum_{j=0}^{k_i - 1} \frac{t^j}{j!} (A - \lambda_i E)^j v_i \right), \quad i \in \{1, \ldots, n\}$$

eine Lösungsbasis von $\dot{x} = Ax$.

Nach linearer Algebra ist die Voraussetzung dieses Satzes genau dann erfüllt, wenn A eine Jordansche Normalform besitzt, d.h. wenn das charakteristische Polynom von A in Linearfaktoren zerfällt.

Beweis. Nach dem Satz sind die x_i Lösungen, und wegen $x_i(0) = v_i$ sind sie linear unabhängig. □

Beispiel 198. Vgl. Beispiel 193. Die Matrix der homogenen Differentialgleichung

$$\frac{d}{dt} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 & 3 \\ 3 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

hat die Eigenwerte 4 und -2 mit Eigenvektoren $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ bzw. $\begin{pmatrix} 1 \\ -1 \end{pmatrix}$. Deshalb ist

$$x_1(t) = e^{4t} \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \; x_2(t) = e^{-2t} \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

eine Lösungsbasis.

□

In diesem Beispiel hat man eine Basis aus Eigenvektoren. Braucht man aber Hauptvektoren höherer Stufe, so wird die Sache mühsam, denn man muß die Matrixpotenzen bis $(A - \lambda E)^k$ bilden. Einfacher geht es bei doppelten Nullstellen des charakteristischen Polynoms, d.h. bei Eigenwerten der algebraischen Vielfachheit 2 und der geometrischen Vielfachheit 1:

Beispiel 199. Wir betrachten $\dot{x} = Ax$. Sei λ ein Eigenwert von A mit der algebraischen Vielfachheit 2 und der geometrischen Vielfachheit 1, d.h. λ ist eine doppelte Nullstelle des charakteristischen Polynoms, aber $\dim \mathrm{Kern}(A - \lambda E) = 1$. Sei v_1 einen Eigenvektor zu λ. Dann gibt es zu λ einen von v_1 linear unabhängigen Hauptvektor v_2. Für den gilt

$$0 = (A - \lambda E)^2 v_2 = (A - \lambda E)((A - \lambda E)v_2) = A(A - \lambda E)v_2 - \lambda(A - \lambda E)v_2.$$

Das heißt, $(A - \lambda E)v_2$ ist ein Eigenvektor zum Eigenwert λ. (Beachte $(A - \lambda E)v_2 \neq 0$, sonst wäre v_2 ja ein von v_1 linear unabhängiger Eigenvektor.) Damit ist $(A - \lambda E)v_2 = av_1$ mit $a \neq 0$ und

$$(A - \lambda E)(\frac{1}{a} v_2) = v_1.$$

Weil es auf Vielfache $\neq 0$ bei Eigen- und Hauptvektoren nicht ankommt, können wir den Faktor $1/a$ vergessen.

Fazit: Das Gleichungssystem

$$(A - \lambda E)v_2 = v_1$$

ist lösbar und liefert uns „den“ fehlenden Hauptvektor zum Eigenwert λ. Die zugehörige Lösung ist dann

$$x(t) = e^{\lambda t}(v_2 + tv_1).$$

□

Beispiel 200. Wir betrachten

$$A = \begin{pmatrix} 0 & 1 & -1 \\ -2 & 3 & -1 \\ -1 & 1 & 1 \end{pmatrix}$$

Diese Matrix hat die Eigenwerte $\lambda_1 = \lambda_2 = 1$ und $\lambda_3 = 2$. Die Gleichung

$$(A - 1\,E)v = \begin{pmatrix} -1 & 1 & -1 \\ -2 & 2 & -1 \\ -1 & 1 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

liefert einen linear unabhängigen Eigenvektor $\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$ zum Eigenwert 1. Daher liefert

$$\begin{pmatrix} -1 & 1 & -1 \\ -2 & 2 & -1 \\ -1 & 1 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$$

einen Hauptvektor $\begin{pmatrix} 0 \\ 0 \\ -1 \end{pmatrix}$ zum Eigenwert 1, der offenbar von dem Eigenvektor linear unabhängig ist. Das Differentialgleichungssystem

$$\dot{x}(t) = Ax$$

hat in diesem Fall eine Lösungsbasis aus

$$x_1(t) = e^t \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \qquad x_2(t) = e^t\left(\begin{pmatrix} 0 \\ 0 \\ -1 \end{pmatrix} + t \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}\right)$$

und einer weiteren Lösung $x_3(t) = e^{2t} \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$ zum Eigenwert 2.

□

Beispiel 201. Die Voraussetzungen des Beispiels 199 sind nötig: Wir betrachten die Matrix

$$A = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 2 & 1 \\ 0 & -1 & 0 \end{pmatrix}.$$

Das charakteristische Polynom ist

$$\det(A - \lambda E) = \begin{vmatrix} 1-\lambda & 1 & 1 \\ 0 & 2-\lambda & 1 \\ 0 & -1 & -\lambda \end{vmatrix} = -(\lambda - 1)^3.$$

Also ist $\lambda = 1$ ein Eigenwert der algebraischen Vielfachheit 3. Die zugehörige Eigenvektorgleichung

$$\begin{pmatrix} 0 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & -1 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = 0$$

hat Rang 1. Sie liefert also zwei linear unabhängige Eigenvektoren, zum Beispiel

$$v_1 = \begin{pmatrix}1\\0\\0\end{pmatrix}, \quad v_2 = \begin{pmatrix}0\\1\\-1\end{pmatrix}.$$

Aber keines der beiden Gleichungssysteme

$$\begin{pmatrix}0 & 1 & 1\\0 & 1 & 1\\0 & -1 & -1\end{pmatrix}\begin{pmatrix}x\\y\\z\end{pmatrix} = \begin{pmatrix}1\\0\\0\end{pmatrix}, \quad \begin{pmatrix}0 & 1 & 1\\0 & 1 & 1\\0 & -1 & -1\end{pmatrix}\begin{pmatrix}x\\y\\z\end{pmatrix} = \begin{pmatrix}0\\1\\-1\end{pmatrix}$$

ist lösbar! Um den fehlenden Hauptvektor zu finden muss man also eine von v_1 und v_2 unabhängige Lösung v_3 von $(A-1E)^2 v = 0$ finden. In diesem Fall ist das trivial, weil die algebraische Vielfachheit 3 ist, die Hauptvektoren also den ganzen $\mathbb{R}^3$ aufspannen. Man kann deshalb *irgendeinen* von v_1 und v_2 unabhängigen Vektor v_3 wählen, z.B. $v_3 = \begin{pmatrix}0\\1\\0\end{pmatrix}$.

Eine Lösungsbasis von $\dot{x} = Ax$ ist in diesem Fall also

$$x_1(t) = e^t \begin{pmatrix}1\\0\\0\end{pmatrix}, x_2(t) = e^t \begin{pmatrix}0\\1\\-1\end{pmatrix}, x_3(t) = e^t \left(\begin{pmatrix}0\\1\\0\end{pmatrix} + t \begin{pmatrix}1\\1\\-1\end{pmatrix} \right).$$

□

Komplexe Eigenwerte. Um Probleme mit der abstrakten Komplexifizierung reeller Vektorräume zu vermeiden, beschränken wir uns im folgenden auf $V = \mathbb{R}^n$ und Matrizen A.

Bekanntlich hat nicht jede reelle Matrix eine reelle Jordansche Normalform, d.h. der $\mathbb{R}^n$ besitzt möglicherweise keine Basis aus Hauptvektoren. Wie kommt man dann zu einer Lösungsbasis von $\dot{x} = Ax$?

Zunächst kann man die obigen Überlegungen ohne wesentliche Änderung auf komplexwertige Lösungen $x : \mathbb{R} \to \mathbb{C}^n$ verallgemeinern. Man stellt fest, dass das Anfangswertproblem eindeutig lösbar und der Lösungsraum von $\dot{x} = Ax$ ein n-dimensionaler $\mathbb{C}$-Vektorraum ist. Und weil über $\mathbb{C}$ jede Matrix eine Jordansche Normalform besitzt, liefert die obige Theorie also eine Methode zur Gewinnung einer Lösungsbasis für komplexe homogene lineare Differentialgleichungssysteme mit konstanten Koeffizienten.

Wenn man aber von einem reellen Problem ausgeht, möche man gern eine reelle Lösungsbasis haben. Seien also die Matrix A reell und $x : \mathbb{R} \to \mathbb{C}^n$ eine komplexe Lösung. Wir bezeichnen mit $\bar{}$ die komponentenweise komplexe Konjugation als Abbildung von $\mathbb{C}^n$ nach $\mathbb{C}^n$. Weil diese Abbildung reell linear ist, folgt aus $\dot{x}(t) = Ax(t)$, dass

$$\dot{\bar{x}}(t) = \overline{\dot{x}(t)} = \overline{Ax(t)} = A\bar{x}(t)$$

ist. Mit jeder komplexen Lösung ist also auch die dazu konjugierte Funktion eine Lösung. Weil Linearkombinationen von Lösungen wieder Lösungen sind, erhält man aus jeder komplexen Lösung $x : \mathbb{R} \to \mathbb{C}^n$ zwei reelle

$$x_{re}(t) := \frac{1}{2}(x(t) + \bar{x}(t)), \quad x_{im}(t) := \frac{1}{2i}(x(t) - \bar{x}(t)).$$

Die komplexen Lösungen $x(t)$ und $\bar{x}(t)$ liefern natürlich dieselben reellen Lösungen, deshalb kann man von jedem konjugiert-komplexen Paar eine Lösung ignorieren. Ist schließlich $x(t)$ eine komplexe Lösung mit reellem Anfangswert $x(t_0) = \xi_0$, so ist x_{re} eine reelle Lösung

mit demselben Anfangswert, x_{im} eine mit Anfangswert $x_{im}(t_0) = 0$, also $x_{im} = 0$. Daher ist $x = x_{re}$ überhaupt reell, und man bekommt auf diese Weise Lösungen für alle reellen Anfangswerte, also alle reellen Lösungen.

Wir betrachten das noch genauer. Ist $u_1, \dots, u_n$ eine Basis des $\mathbb{C}^n$ und $u_k = v_k + iw_k$ mit $u_k, w_k \in \mathbb{R}^n$ und ist weiter $\xi \in \mathbb{R}^n$, so gibt es $\alpha_k = \beta_k + i\gamma_k$ mit $\beta_k, \gamma_k \in \mathbb{R}$, so dass

$$\xi = \sum_{k=1}^{n} \alpha_k u_k = \sum_{k=1}^{n} (\beta_k v_k - \gamma_k w_k) + i \sum_{k=1}^{n} (\beta_k w_k + \gamma_k v_k) = \sum_{k=1}^{n} (\beta_k v_k - \gamma_k w_k).$$

Die Real- und Imaginärteile der u_k bilden also ein Erzeugendensystem von $\mathbb{R}^n$. Berechnet man nun die Eigenwerte des reellen A und dazu mittels Hauptvektoren eine Lösungsbasis $x_1, \dots, x_n$ für den Raum der komplexen Lösungen von $\dot{x} = Ax$, so kann man sich bei konjugiert-komplexen Eigenwerten jeweils auf einen beschränken und für den dazu konjugierten die konjugierten Lösungen verwenden. Spaltet man diese in Real- und Imaginärteil und läßt die doppelt auftretenden Lösungen fort, so erhält man n reelle Lösungen $x_{r1}, \dots, x_{rn}$, deren Werte $x_{r1}(0), \dots, x_{rn}(0)$ nach der vorstehenden Überlegung den $\mathbb{R}^n$ erzeugen und die deshalb linear unabhängig sind.

Wir fassen zusammen:

Gesucht eine reelle Lösungsbasis für $\dot{x} = Ax$ mit reeller $n \times n$-Matrix A.

1. Berechne die Eigenwerte von A, also die reellen und komplexen Nullstellen des charakteristischen Polynoms $\det(A - \lambda E)$. Von den Paaren konjugiert-komplexer Eigenwerten lasse jeweils einen weg.

2. Zu jedem der verbleibenden Eigenwerte λ der algebraischen Vielfachheit k berechne k linear unabhängige Hauptvektoren $v_1, \dots, v_k \in \mathbb{C}^n$ als Lösungen von $(A - \lambda E)^k v = 0$. Sie liefern k Lösungen

$$x_{\lambda,\kappa}(t) = e^{\lambda t} \sum_{j=0}^{k} \frac{t^j}{j!} (A - \lambda E)^j v_\kappa, \quad \kappa \in \{1, \dots, k\}.$$

3. Die entstehenden nicht-reellen Lösungen zerlege in Real- und Imaginärteil. Das liefert insgesamt n linear unabhängige reelle Lösungen und damit eine reelle Lösungsbasis.

Wir schließen mit einem einfachen

Beispiel 202. Wir betrachten $\dot{x} = Ax$ mit $A = \begin{pmatrix} 0 & 5 \\ -2 & 2 \end{pmatrix}$.

Das charakteristische Polynom ist

$$\begin{vmatrix} -\lambda & 5 \\ -2 & 2 - \lambda \end{vmatrix} = \lambda^2 - 2\lambda + 10 = (\lambda - (1 + 3i))(\lambda - (1 - 3i)).$$

Berechnung eines Eigenvektors zu $\lambda = 1 + 3i$:

$$\begin{pmatrix} -1 - 3i & 5 \\ -2 & 1 - 3i \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad \Longleftarrow \quad v = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} 5 \\ 1 + 3i \end{pmatrix}$$

Das liefert die komplexe Lösung

$$\begin{aligned} x(t) &= e^{(1+3i)t} \begin{pmatrix} 5 \\ 1 + 3i \end{pmatrix} \\ &= e^t (\cos 3t + i \sin 3t) \begin{pmatrix} 5 \\ 1 + 3i \end{pmatrix} \\ &= e^t \begin{pmatrix} 5 \cos 3t \\ \cos 3t - 3 \sin 3t \end{pmatrix} + i e^t \begin{pmatrix} 5 \sin 3t \\ 3 \cos 3t + \sin 3t \end{pmatrix} \end{aligned}$$

und die reellen Lösungen

$$x_1(t) = e^t \begin{pmatrix} 5\cos 3t \\ \cos 3t - 3\sin 3t \end{pmatrix}, \quad x_2(t) = e^t \begin{pmatrix} 5\sin 3t \\ 3\cos 3t + \sin 3t \end{pmatrix},$$

die offenbar linear unabhängig sind.

□

5.2.3 Skalare lineare Differentialgleichungen höherer Ordnung.

- Skalare lineare Differentialgleichungen höherer Ordnung treten sehr häufig auf, zum Beispiel in vielen grundlegenden Problemen der Mechanik oder Elektrotechnik.
- Wir wissen schon, wie man sie umschreiben kann in ein lineares System erster Ordnung, aber hier lernen wir, wie man diesen Aufwand vermeiden und direkt Lösungen finden kann.

Problem: Sei $J \subset \mathbb{R}$ ein offenes Intervall und seien $f_1, \ldots, f_n, g : J \to \mathbb{R}$ stetige Funktionen. Wir suchen Lösungen der linearen Differentialgleichung

$$x^{(n)} + f_1(t)x^{(n-1)} + \ldots + f_{n-1}(t)\dot{x} + f_n(t)x = g(t), \tag{92}$$

gegebenenfalls mit den Anfangsbedingungen in $t_0 \in J$

$$x(t_0) = \xi_0, \ \ldots, \ x^{(n-1)}(t_0) = \xi_{n-1}. \tag{93}$$

Wir haben in der Einleitung zu den gewöhnlichen Differentialgleichungen schon am Beispiel der Newtonschen Bewegungsgleichung demonstriert, wie man die Differentialgleichung höherer Ordnung auf eine erster Ordnung in einem höher-dimensionalen Raum übersetzt. Wir wenden das auf das vorstehende Problem an.

Ist x eine Lösung von (92), (93), und setzt man $y_1 = x, y_2 = \dot{x}, \ldots, y_n = x^{(n-1)}$, so folgt mit $y := \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$

$$\dot{y} = \begin{pmatrix} 0 & 1 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & 1 \\ -f_n & -f_{n-1} & \ldots & -f_1 \end{pmatrix} y + \begin{pmatrix} 0 \\ \vdots \\ 0 \\ g(t) \end{pmatrix}, \quad y(t_0) = \begin{pmatrix} \xi_0 \\ \vdots \\ \xi_{n-1} \end{pmatrix} \tag{94}$$

Ist umgekehrt $y : J \to \mathbb{R}^n$ eine Lösung von (94), so ist $x := y_1$ eine solche von (92), (93).

Damit folgt aus dem Hauptsatz 194 über lineare Differentialgleichungen:

Satz 203. *Gegeben sei das Anfangswertproblem*

$$x^{(n)} + f_1(t)x^{(n-1)} + \ldots + f_{n-1}(t)\dot{x} + f_n(t)x = g(t), \tag{95}$$

$$x(t_0) = \xi_0, \ \ldots, \ x^{(n-1)}(t_0) = \xi_{n-1}. \tag{96}$$

mit stetigen Funktionen $f_1, \ldots, f_n, g : J \to \mathbb{R}$ auf einem Intervall J um t_0 und $\xi_0, \ldots \xi_{n-1} \in \mathbb{R}$
Dann gilt

(i) Das Anfangswertproblem hat genau eine auf ganz J definierte Lösung.

(ii) Der Lösungsraum der zugehörigen homogenen linearen Differentialgleichung n-ter Ordnung

$$x^{(n)} + f_1(t)x^{(n-1)} + \ldots + f_{n-1}(t)\dot{x} + f_n(t)x = 0 \tag{97}$$

ist ein n-dimensionaler Untervektorraum von $C^n(J, \mathbb{R})$.

Satz 203 (Fortsetzung). *(iii) Lösungen $x_1, \dots, x_n$ von* (97) *sind genau dann linear unabhängig, wenn die Spaltenvektoren der sogenannten* Wronskimatrix

$$W(t) := \begin{pmatrix} x_1(t) & \dots & x_n(t) \\ \dot{x}_1(t) & \dots & \dot{x}_n(t) \\ \vdots & & \vdots \\ x_1^{(n-1)}(t) & \dots & x_n^{(n-1)} \end{pmatrix}$$

an einer Stelle (und dann an allen Stellen) $t \in J$ linear unabhängig sind.

(iv) (Variation der Konstanten) *Hat man eine Lösungsbasis $x_1, \dots, x_n$ für die homogene Gleichung* (97)*, so erhält man alle Lösungen der inhomogenen Gleichung in der Form*

$$x(t) = \sum_{i=1}^{n} c_i(t) x_i(t),$$

wo die Funktionen $c_i \in C^1(J, \mathbb{R})$ bis auf Konstanten bestimmt sind durch ein lineares Gleichungssystem für ihre Ableitungen:

$$W(t) \begin{pmatrix} \dot{c}_1(t) \\ \vdots \\ \dot{c}_n(t) \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ g(t) \end{pmatrix}.$$

Beispiel 204. Löse das Anfangswertproblem

$$\ddot{x} - 6\dot{x} + 8x = 64t^2$$
$$x(0) = 5,\ \dot{x}(0) = 0$$

mit den geschenkten Lösungen $x_1(t) = e^{2t}, x_2(t) = e^{4t}$ für die zugehörige homogene Gleichung.

0. Schritt. Die Wronskimatrix der beiden Lösungen in t=0 ist

$$W(0) = \begin{pmatrix} e^{2t} & e^{4t} \\ 2e^{2t} & 4e^{4t} \end{pmatrix}_{t=0} = \begin{pmatrix} 1 & 1 \\ 2 & 4 \end{pmatrix}.$$

Offenbar sind die Spalten linear unabhängig und x_1, x_2 bilden daher eine Basis für den Lösungsraum der homogenen Gleichung.

1. Schritt. Lösen des linearen Gleichungssystems

$$\begin{pmatrix} e^{2t} & e^{4t} \\ 2e^{2t} & 4e^{4t} \end{pmatrix} \begin{pmatrix} \dot{c}_1(t) \\ \dot{c}_2(t) \end{pmatrix} = \begin{pmatrix} 0 \\ 64t^2 \end{pmatrix}$$

liefert

$$\dot{c}_1(t) = -32t^2 e^{-2t}, \quad \dot{c}_2(t) = 32t^2 e^{-4t}.$$

2. Schritt. Integrieren mit partieller Integration oder *http://integrals.wolfram.com/* liefert

$$c_1(t) = 8e^{-2t}(1 + 2t + 2t^2), \quad c_2(t) = -e^{-4t}(1 + 4t + 8t^2).$$

3. Schritt. Die allgemeine Lösung der gegebenen Differentialgleichung ist also

$$x(t) = 8(1 + 2t + 2t^2) - (1 + 4t + 8t^2) + a_1 e^{2t} + a_2 e^{4t} = (7 + 12t + 8t^2) + a_1 e^{2t} + a_2 e^{4t}$$

mit beliebigen Konstanten $a_1, a_2 \in \mathbb{R}$.

4. Schritt. Um die Anfangsbedingungen zu erfüllen, berechnen wir

$$x(0) = 7 + a_1 + a_2 = 5$$
$$\dot{x}(0) = 12 + 2a_1 + 4a_2 = 0$$

und erhalten aus diesem linearen Gleichungssystem $a_1 = 2, a_2 = -4$. Damit ist die gesuchte Lösung

$$x(t) = 7 + 12t + 8t^2 + 2e^{2t} - 4e^{4t}.$$

Bemerkung. Variation der Konstanten *ein* Algorithmus zur Ermittlung einer Lösung für die inhomogene Gleichung. Er erfordert Lösen eines linearen Gleichungssystems und Integrationen. Nicht selten kann man durch genaues Hinsehen (oder Erfahrung) auch eine Lösung für die inhomogene Gleichung leichter finden, eventuell sogar einfach hinschreiben. Im obigen Fall liefert die linke Seite der Differentialgleichung, weil die Koeffizienten konstant sind, für jedes eingesetzte Polynom $x(t)$ wieder ein Polynom, und zwar vom gleichen Grad. Weil aber auch die rechte Seite ein Polynom von zweiten Grad ist, kann man versuchen, einfach

$$x(t) = A + Bt + Ct^2$$

anzusetzen und die Koeffizienten (durch Koeffizientenvergleich) so zu bestimmen, dass t^2 herauskommt. Das ist im obigen Fall wesentlich einfacher als die Variation der Konstanten. Probieren Sie es!

□

Wir wollen nun zeigen, wie man eine Lösungsbasis für eine lineare homogene Differentialgleichung finden kann, wenn ihre Koeffizienten *konstant* sind. Wie im Fall der Systeme ist es günstig, dabei auch komplexwertige Lösungen $x : \mathbb{R} \to \mathbb{C}$ zuzulassen und sich später zu überlegen, wie man daraus wieder reellwertige gewinnen kann.

Satz 205. *Wir betrachten also auf $J = \mathbb{R}$ die homogene lineare Differentialgleichung*

$$x^{(n)} + a_1 x^{(n-1)} + \ldots + a_{n-1}\dot{x} + a_n x = 0 \tag{98}$$

mit $a_1, \ldots, a_n \in \mathbb{R}$.

(i) Eine Basis für den Unterraum aller komplexwertigen *Lösungen von* (98) *in $C^n(\mathbb{R}, \mathbb{C})$ ist gegeben durch die Funktionen*

$$x_{ij}(t) = t^j\, e^{t\lambda_i}, \qquad 1 \le i \le m, \quad 0 \le j \le k_i - 1.$$

Dabei sind $\lambda_1, \ldots, \lambda_m \in \mathbb{C}$ die verschiedenen Nullstellen des sogenannten charakteristischen Polynoms

$$\chi(\lambda) = \lambda^n + a_1\lambda^{n-1} + \ldots + a_{n-1}\lambda + a_n,$$

der Differentialgleichung und k_i ist die Vielfachheit der Nullstelle λ_i.

(ii) Sind die a_i reell und besitzt $\chi(\lambda)$ nicht-reelle Nullstellen, so treten diese in konjugierten Paaren gleicher Multiplizität auf. Ist $\lambda = \alpha \pm i\omega$ ein solches Paar mit Multiplizität k, so erhält man daraus reelle *linear unabhängige Lösungen*

$$t^j\, e^{\alpha t}\cos(\omega t) \text{ und } t^j\, e^{\alpha t}\sin(\omega t), \quad 0 \le j \le k - 1.$$

Auf diese Weise erhält man eine Basis des reellen *Lösungsraumes von* (98).

Wir brauchen einige Vorbereitungen für den Beweis. Wenn der Satz richtig ist, sind alle Lösungen beliebig oft differenzierbar. Wir suchen unsere Lösungen deshalb gleich im Raum

$$C^\infty := C^\infty(\mathbb{R}, \mathbb{C})$$

und schreiben D^k zur Abkürzung für $\frac{d^k}{dt^k}$. Für ein Polynom

$$\phi(\lambda) = a_0\lambda^n + a_1\lambda^{n-1} + \ldots + a_{n-1}\lambda + a_n$$

mit möglicherweise komplexen Koeffizienten a_k definieren wir

$$\phi(D) \in L(C^\infty, C^\infty)$$

durch

$$\begin{aligned}\phi(D)x &= a_0D^nx + a_1D^{n-1}x + \ldots + a_{n-1}Dx + a_nx \\ &= a_0x^{(n)} + a_1x^{(n-1)} + \ldots + a_1\dot{x} + a_nx.\end{aligned}$$

Lemma 206. *(i) Die "Einsetzung" $\phi \mapsto \phi(D)$ ist ein Algebra-Homomorphismus von der Algebra der Polynome in die Algebra $L(C^\infty, C^\infty)$. Insbesondere gilt für Polynome $\phi(\lambda), \psi(\lambda)$ und $\rho(\lambda) = \phi(\lambda)\psi(\lambda)$, dass*

$$\rho(D) = \phi(D)\psi(D) : C^\infty \to C^\infty.$$

Daher ist $\phi(D)\psi(D) = \psi(D)\phi(D)$.

(ii) Seien $p \neq 0$ ein Polynom, $k \in \mathbb{N}$ und $\mu, \nu \in \mathbb{C}$.
Wir setzen D in das Polynom $\phi(\lambda) := (\lambda - \mu)^k$ ein. Dann ist

$$(D - \mu)^k \left(p(t)e^{\nu t}\right) = q(t)e^{\nu t}$$

mit einem Polynom q, für das gilt:

$$\begin{aligned}\mu \neq \nu &\implies \operatorname{Grad} q = \operatorname{Grad} p, \\ \mu = \nu &\implies \operatorname{Grad} q = (\operatorname{Grad} p) - k,\end{aligned}$$

Dabei soll $\operatorname{Grad} q < 0$ bedeuten, dass $q = 0$.

Beweis des Lemmas. Zu (i). Seien $\phi(\lambda) = \sum a_i\lambda^i$, $\psi(\lambda) = \sum b_j\lambda^j$. Dann ist

$$\phi(D)\psi(D)x = \phi(D)\sum b_jx^{(j)} = \sum b_j \sum a_ix^{(i+j)}.$$

Zu (ii). Vollständige Induktion über k. Für $k = 0$ ist nichts zu zeigen.

$k \to (k+1)$. Nach Induktionsvoraussetzung ist

$$(D - \mu)^k \left(p(t)e^{\nu t}\right) = q(t)e^{\nu t}$$

mit einem Polynom $q(t)$ vom im Satz beschriebenen Grad. Dann gilt aber

$$(D - \mu)^{k+1} \left(p(t)e^{\nu t}\right) = (D - \mu)q(t)e^{\nu t} = (\dot{q}(t) + \nu q(t) - \mu q(t))e^{\nu t}.$$

Ist $\mu = \nu$, so verkleinert sich der Grad des Polynom-Faktors vor $e^{\nu t}$ um 1, andernfalls bleibt er gleich. □

Beweis des Satzes. *Zu (i).* Ist $\chi(\lambda)$ das im Satz definierte charakteristische Polynom, so ist die Differentialgleichung gegeben durch

$$\chi(D)x = 0.$$

Andrerseits ist nach dem Lemma für $j < k_i$

$$\begin{aligned}\chi(D)x_{ij}(t) &= (D-\lambda_1)^{k_1}\dots(D-\lambda_m)^{k_m}\left(t^j e^{\lambda_i t}\right)\\ &= (D-\lambda_1)^{k_1}\dots(D-\lambda_{i-1})^{k_{i-1}}(D-\lambda_{i+1})^{k_{i+1}}\dots(D-\lambda_m)^{k_m}(D-\lambda_i)^{k_i}\left(t^j e^{\lambda_i}\right)\\ &= 0.\end{aligned}$$

Daher sind die angegebenen Funktionen n Lösungen der Differentialgleichung. Wir zeigen ihre lineare Unabhängigkeit. Sei

$$0 = \sum_{i,j}\alpha_{ij}x_{ij}(t) = \sum_i p_i(t)e^{\lambda_i t}$$

für alle t. Dabei sind die $p_i(t)$ Polynome vom Grad $\leq k_i - 1$, und wir müssen zeigen, dass sie alle 0 sind. Aber nach dem Lemma ist

$$0 = (D-\lambda_2)^{k_2}\dots(D-\lambda_m)^{k_k}\left(\sum_i p_i(t)e^{\lambda_i t}\right) = q_1(t)e^{\lambda_1 t}$$

mit einem Polynom q_1 vom selben Grad wie p_1. Also folgt $p_1 = 0$, und entsprechend für die anderen $p_i(t)$.

Zu (ii). Nach der Eulerschen Identität ist

$$\begin{aligned}t^j e^{\alpha t}\cos\omega t &= \frac{1}{2}t^j e^{(\alpha+i\omega)t} + \frac{1}{2}t^j e^{(\alpha-i\omega)t},\\ t^j e^{\alpha t}\sin\omega t &= \frac{1}{2i}t^j e^{(\alpha+i\omega)t} - \frac{1}{2i}t^j e^{(\alpha-i\omega)t}.\end{aligned}$$

Daher sind die $\cos-\sin-$Lösungen als Linearkombination (mit komplexen Koeffizienten) von Lösungen der homogenen linearen Differentialgleichung auch Lösungen. Weil man aus ihnen die komplexe Lösungsbasis linear kombinieren kann, bilden sie ein Erzeugendensystem für den komplexen Lösungsraum mit n Elementen. Daher sind sie linear unabhängig über den komplexen Zahlen, also erst recht über den reellen Zahlen. □

Beispiel 207. Die charakteristische Gleichung von

$$\ddot{x} - 6\dot{x} + ax = 0$$

hat die Lösungen $\lambda_{1,2} = 3 \pm \sqrt{9-a}$. Für $a = 8, 9, 10$ erhält man als Lösungsbasen also

$$x_1(t) = e^{2t},\ x_2 = e^{4t}$$

bzw.

$$x_1(t) = e^{3t},\ x_2 = te^{3t}$$

bzw.

$$x_1(t) = e^{(3+i)t},\ x_2 = e^{(3-i)t}.$$

Im letzteren Fall ist eine reelle Lösungsbasis gegeben durch

$$x_1(t) = e^{3t}\cos t,\ x_2 = e^{3t}\sin t.$$

□

6 Anhang

6.1 Hauptminorenkriterium

Wir geben einen Beweis (von Udo Jeromin) für dieses Kriterium. Vgl. auch *M. Köcher, Lineare Algebra und analytische Geometrie, Springer.*

Hauptminorenkriterium. *Eine symmetrische $(n \times n)$-Matrix $A = (a_{ij})_{i,j=1,\dots,n}$ ist genau dann positiv definit, wenn alle Hauptminoren positiv sind. Dabei sind* Hauptminoren *die Determinanten der Matrizen*

$$A_k := (a_{ij})_{i,j=1,\dots,k}$$

A ist genau dann negativ definit*, wenn die Hauptminoren wechselndes Vorzeichen beginnend mit $a_{11} < 0$ haben.*

Beweis. Die Behauptung über negative Definitheit folgt aus der über positive Definitheit durch Betrachtung von $-A$. Wir beweisen also nur den ersten Teil des Kriteriums.

Mit $\langle .\,, . \rangle$ bezeichnen wir das kanonische Skalarprodukt der Euklidischen Räume. A ist positiv definit, wenn gilt

$$\forall_{x \in \mathbb{R}^n} (x \neq 0 \implies \langle Ax, x \rangle > 0).$$

Wir benutzen das obige Zitat aus der linearen Algebra

$$A \text{ positiv definit} \implies \text{alle Eigenwerte positiv} \implies \det A > 0$$

und kommen zum Beweis des Lemmas:

Beweis von $\Rightarrow$.

$$\text{Für } x = \begin{pmatrix} x_1 \\ \vdots \\ x_k \end{pmatrix} \in \mathbb{R}^k \text{ mit } k \leq n \text{ sei } x' := \begin{pmatrix} x \\ 0 \end{pmatrix} := \begin{pmatrix} x_1 \\ \vdots \\ x_k \\ 0 \\ \vdots \\ 0 \end{pmatrix} \in \mathbb{R}^n.$$

Ist A positiv definit, so gilt für alle $k \in \{1, \dots, n\}$ und $x \in \mathbb{R}^k \setminus \{0\}$:

$$0 < \langle Ax', x' \rangle = \langle A_k x, x \rangle .$$

Also sind alle A_k ebenfalls positiv definit und haben daher positive Determinante.

Beweis von $\Leftarrow$. Seien nun umgekehrt alle $\det A_k$ positiv. Wir zeigen durch vollständige Induktion über n, dass A positiv definit ist.

Der Beweis benutzt folgende Idee, um die Determinante der ganzen Matrix mit der eines Hauptminors in Verbindung zu bringen:

Ist $\begin{pmatrix} A & B \\ C & D \end{pmatrix}$ eine Blockmatrix mit quadratischem A und D, und ist A invertierbar, so gilt

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix} = \begin{pmatrix} A & 0 \\ C & E \end{pmatrix} \begin{pmatrix} E & A^{-1}B \\ 0 & D - CA^{-1}B \end{pmatrix}.$$

Das kann man im Kopf nachrechnen. Insbesondere ist dann

$$\det \begin{pmatrix} A & B \\ C & D \end{pmatrix} = \det A \, \det(D - CA^{-1}B).$$

$\underline{n = 1.}$ Nichts zu zeigen.

$\underline{(n-1) \to n.}$ Wir nehmen also an, der Satz sei für $n-1$ bereits bewiesen. Dann ist A_{n-1} positiv definit und insbesondere invertierbar.

1. Schritt. Für $y \in \mathbb{R}^{n-1} \setminus \{0\}$ ist nach Voraussetzung

$$\left\langle A \begin{pmatrix} y \\ 0 \end{pmatrix}, \begin{pmatrix} y \\ 0 \end{pmatrix} \right\rangle = \langle A_{n-1} y, y \rangle > 0.$$

2. Schritt. Es genügt zu zeigen, dass

$$f(y) := \left\langle A \begin{pmatrix} y \\ 1 \end{pmatrix}, \begin{pmatrix} y \\ 1 \end{pmatrix} \right\rangle > 0 \text{ für alle } y \in \mathbb{R}^{n-1}.$$

Wegen $\langle A(tx), tx \rangle = t^2 \langle Ax, x \rangle$ folgt dann

$$\langle Ax, x \rangle > 0 \text{ für alle } x \text{ mit } x_n \neq 0,$$

und mit dem ersten Schritt ergibt sich daraus die Behauptung. Schreiben wir

$$A = \begin{pmatrix} A_{n-1} & a \\ a^T & \alpha \end{pmatrix}$$

mit $a = \begin{pmatrix} a_{1n} \\ \vdots \\ a_{n-1\,n} \end{pmatrix}$ und $\alpha = a_{nn}$, so wird

$$f(y) = \langle A_{n-1} y, y \rangle + 2 \langle a, y \rangle + \alpha.$$

3. Schritt. Ist $\lambda > 0$ der kleinste Eigenwert von A_{n-1}, so gilt

$$\langle A_{n-1} y, y \rangle \geq \lambda \langle y, y \rangle,$$

also

$$\langle A_{n-1} y, y \rangle + 2 \langle a, y \rangle + \alpha \geq \lambda \|y\|^2 - 2\|a\| \, \|y\| - |\alpha| = \|y\|(\lambda \|y\| - \|a\|) - |\alpha|.$$

Das wird groß, wenn $\|y\|$ groß wird: Außerhalb einer hinreichend großen Kugel im $\mathbb{R}^{n-1}$ ist daher $f(y) \geq f(0)$. Deshalb nimmt die stetige Funktion f auf $\mathbb{R}^{n-1}$ ihr globales Minimum an. An dieser Stelle verschwindet ihre Ableitung, die wir jetzt berechnen:

$$D_y f(v) = \langle A_{n-1} y, v \rangle + \langle A_{n-1} v, y \rangle + 2 \langle a, v \rangle = 2 \langle A_{n-1} y + a, v \rangle.$$

Das Minimum wird also angenommen an der Stelle $y_* = -A_{n-1}^{-1} a$. Sein Wert ist

$$f(y_*) = \left\langle a, A_{n-1}^{-1} a \right\rangle - 2 \left\langle a, A_{n-1}^{-1} a \right\rangle + \alpha = \alpha - \left\langle a, A_{n-1}^{-1} a \right\rangle.$$

4. Schritt. Nach der Vorüberlegung ist

$$\det A_n = \det A_{n-1} (\alpha - \left\langle a, A_{n-1}^{-1} a \right\rangle).$$

Mit $\det A_n$ und $\det A_{n-1}$ ist also auch $f(y*) = \alpha - \left\langle a, A_{n-1}^{-1} a \right\rangle$ positiv. □

6.2 Vektorwertige Integrale

Wir definieren das Integral für stetige Funktionen $g : [a, b] \to V$ mit Werten in einem endlich-dimensionalen Banachraum V so:

Ist $b_1, \ldots, b_n$ eine Basis von V, so schreibt sich g als

$$g = \sum g_i b_i$$

mit stetigen reellwertigen Funktionen g_i. Wir setzen

$$\int_a^b g(t)dt := \sum_i \left(\int_a^b g_i(t)dt \right) b_i.$$

Man zeigt, dass das von der gewählten Basis unabhängig ist. Falls $V = \mathbb{R}^n$, bedeutet das einfach komponentenweise Integration.

Für das so verallgemeinerte Integral gelten die folgenden vom $\mathbb{R}$-wertigen Fall vertrauten Regeln:

$$\int_a^b (g+h)(t)dt = \int_a^b g(t)dt + \int_a^b h(t)dt, \quad \int_a^b \lambda g(t)dt = \lambda \int_a^b g(t)dt, \tag{99}$$

$$\frac{d}{dt} \int_a^t g(\tau)d\tau = g(t), \tag{100}$$

$$\left\| \int_a^b g(t)dt \right\| \leq \int_a^b \|g(t)\|dt. \tag{101}$$

Die beiden ersten Gleichungen folgen trivial aus der Definition. Die dritte beweise ich nur für den Fall $V = \mathbb{R}^n$ und die Norm zum üblichen Skalarprodukt $\langle x, y \rangle = \sum x_i y_i$.

Aus (99) folgt für $v \in \mathbb{R}^n$

$$\int_a^b < v, g(t) > dt = \int_a^b (\sum v_i g_i(t))dt = \sum v_i \int_a^b g_i(t)dt =< v, \int_a^b g(t)dt > .$$

Setzt man

$$v := \frac{\int_a^b g(t)dt}{\| \int_a^b g(t)dt \|} \in \mathbb{R}^n, \tag{102}$$

so wird

$$\| \int g(t)dt \| =< v, \int g(t)dt >= \int < v, g(t) > dt \underset{(Cauchy-Schwarz)}{\leq} \int \|g(t)\|dt.$$

Ebenso beweist man den allgemeinen Fall, nachdem man zuvor gezeigt hat, dass es zu jedem $v \in V$ (bei uns $v = \int g$) ein $\omega \in L(V, \mathbb{R})$ gibt, für das $\|\omega\| \leq 1$ und $\omega(v) = \|v\|$ ist.)