Analysis II im WS 2011/2012 – Kurzskript

Prof. Dr. C. Löh Wintersemester 2011/2012

Inhaltsverzeichnis

Version vom 10. Februar 2012
clara.loeh@mathematik.uni-regensburg.de
Fakultät für Mathematik, Universität Regensburg, 93040 Regensburg

-1 Literaturhinweise

Die folgenden Listen enthalten eine kleine Auswahl an Literatur zur Analysis II.

Topologie

[1] K. Jänich. *Topologie*, achte Auflage, Springer, 2008.

[2] J.L. Kelley. *General Topology*, Springer, 1975.

[3] J.R. Munkres. *Topology*, zweite Auflage, Pearson, 2003.

[4] L.A. Steen. *Counterexamples in Topology*, Dover, 1995.

(Mehrdimensionale) Analysis

[5] R. Denk, R. Racke. *Kompendium der ANALYSIS*, Vieweg+Teubner, 2011.

[6] O. Forster. *Analysis 1/2/3*, neunte überarbeitete Auflage, Vieweg+Teubner, 2009.

[7] S. Lang. *Undergraduate Analysis*, zweite Auflage, Springer, 2010.

[8] S. Lang. *Real and Functional Analysis*, dritte Auflage, Springer, 1993.

[9] T. Tao. *Analysis I/II*, zweite Auflage, Hindustan Book Agency, 2009.

[10] W. Walter. *Analysis 1/2*, siebte Auflage, Springer, 2009.

... und viele weitere Bücher; je nach eigenen Vorlieben werden Ihnen manche Bücher besser gefallen als andere.

Gewhnliche Differentialgleichungen

Einige der oben genannten Bücher über mehrdimensionale Analysis behandeln auch gewöhnliche Differentialgleichungen.

[11] M. Wilke, J.W. Prüss. *Gewöhnliche Differentialgleichungen und dynamische Systeme*, Birkhäuser/Springer, 2010.

[12] J.C. Robinson. *An Introduction to Ordinary Differential Equations*, Cambridge University Press, 2004.

Weiterführende Themen

[13] M. Aigner, G. Ziegler. *Proofs from the BOOK*, vierte Auflage, Springer, 2009.

[14] A. Engel. *Problem Solving Strategies*, zweite Auflage, Springer, 1999.

0 Einführung

Grob gesagt ist Analysis das Studium lokaler und globaler Eigenschaften reell- oder komplexwertiger Funktionen (insbesondere gehören dazu die Begriffe Konvergenz, Stetigkeit, Differenzierbarkeit, Integration und Approximationseigenschaften).

In der Analysis I haben wir uns im wesentlichen mit Funktionen vom Typ $X \longrightarrow \mathbb{R}$ beschäftigt, wobei $X \subset \mathbb{R}$ ist.

Ziel der Analysis II ist nun, diese Konzepte auf den mehrdimensionalen Fall zu verallgemeinern; dabei werden wir insbesondere wichtige Generalisierungs- und Abstraktionstechniken (einem zentralen Bestandteil der modernen Mathematik) kennenlernen.

1 Metrische Räume und normierte Räume

Der zentrale Begriff in der Analysis ist Approximation bzw. Konvergenz. Bisher haben wir uns im wesentlichen mit Konvergenz in $\mathbb{R}$ beschäftigt. Unser erstes Ziel ist es, den Konvergenzbegriff geeignet zu verallgemeinern. Dies geschieht in zwei Schritten:

- Erster Abstraktionsschritt: metrische Räume (insbesondere normierte Räume).
- Zweiter Abstraktionsschritt: topologische Räume (s. Kapitel 2).

Metrische Räume

Um den Konvergenzbegriff zu verallgemeinern, orientieren wir uns an folgender Strategie zur Generalisierung/Abstraktion:

1. Was sind die zentralen Objekte/Eigenschaften für den betrachteten Aspekt?
2. Wir verwenden diese Objekte/Eigenschaften als Definition und abstrahieren somit von der gegebenen konkreten Situation.
3. Wir überprüfen, ob sich die klassische Theorie (d.h. Sätze, Beispiele, ...) hinreichend gut in der allgemeineren Theorie widerspiegelt.

Im Beispiel des Konvergenzbegriffs (für Konvergenz in $\mathbb{R}$) stellen wir fest, dass der „Abstand“ $|x - y|$ zwischen reellen Zahlen $x, y \in \mathbb{R}$ und der Wertebereich $\mathbb{R}_{\geq 0}$ dieser „Abstandsfunktion“ die zentralen Objekte sind, und dass die zentralen Eigenschaften davon sind, dass der „Abstand“ (Un)Gleichheit von Elementen entdecken kann, dass er symmetrisch ist und dass er die Dreiecksungleichung erfüllt.

Dies führt zum Begriff metrischer Räume:

Definition 1.1 (metrischer Raum)**.** Ein *metrischer Raum* ist ein Paar (X, d), wobei X eine Menge und d eine *Metrik auf* X ist, d.h. d ist eine Abbildung $X \times X \longrightarrow \mathbb{R}_{\geq 0}$ mit folgenden Eigenschaften:

- Für alle $x, y \in X$ ist genau dann $d(x, y) = 0$, wenn $x = y$ ist.
- Für alle $x, y \in X$ ist $d(x, y) = d(y, x)$.
- *Dreiecksungleichung.* Für alle $x, y, z \in X$ ist

$$d(x, z) \leq d(x, y) + d(y, z).$$

Notation 1.2.

- Man bezeichnet die durch

$$\begin{aligned} \mathbb{R} \times \mathbb{R} &\longrightarrow \mathbb{R}_{\geq 0} \\ (x, y) &\longmapsto |x - y| \end{aligned}$$

gegebene Metrik als *Standardmetrik auf* $\mathbb{R}$.

- Ist X eine Menge, so bezeichnet man die durch

$$\begin{aligned} X \times X &\longrightarrow \mathbb{R}_{\geq 0} \\ (x, y) &\longmapsto \begin{cases} 0 & \text{falls } x = y \\ 1 & \text{falls } x \neq y \end{cases} \end{aligned}$$

gegebene Metrik als *diskrete Metrik auf* X.

Notation 1.3 (offene/abgeschlossene Bälle)**.** Sei (X,d) ein metrischer Raum, sei $x \in X$ und sei $r \in \mathbb{R}_{\geq 0}$. Dann ist

$$U(x,r) := \bigl\{y \in X \bigm| d(x,y) < r\bigr\} \text{ bzw.}$$
$$B(x,r) := \bigl\{y \in X \bigm| d(x,y) \leq r\bigr\}$$

der *offene* bzw. *abgeschlossene Ball vom Radius r um x*. Ist der umgebende metrische Raum nicht aus dem Kontext ersichtlich, so schreiben wir auch $U^{(X,d)}(x,r)$ bzw. $B^{(X,d)}(x,r)$.

Caveat 1.4. In der Literatur wird die Notation $B(\dots)$ manchmal auch für offene Bälle verwendet!

Wir können nun die Definition der Konvergenzbegriffe in den Kontext metrischer Räume übersetzen:

Definition 1.5 (Cauchyfolge, Grenzwert, konvergente Folge)**.** Sei (X,d) ein metrischer Raum und sei $(a_n)_{n\in\mathbb{N}}$ eine Folge in X.

– Die Folge $(a_n)_{n\in\mathbb{N}}$ ist eine *Cauchyfolge in* (X,d), wenn

$$\forall_{\varepsilon\in\mathbb{R}_{>0}} \quad \exists_{N\in\mathbb{N}} \quad \forall_{n,m\in\mathbb{N}_{\geq N}} \quad d(a_n,a_m) \leq \varepsilon.$$

  Wenn die Metrik auf X aus dem Kontext klar hervorgeht, sagen wir in diesem Fall auch *Cauchyfolge in X*.

– Ein Element $a \in X$ ist ein *Grenzwert von $(a_n)_{n\in\mathbb{N}}$ in* (X,d), wenn

$$\forall_{\varepsilon\in\mathbb{R}_{>0}} \quad \exists_{N\in\mathbb{N}} \quad \forall_{n\in\mathbb{N}_{\geq N}} \quad d(a_n,a) \leq \varepsilon.$$

– Die Folge $(a_n)_{n\in\mathbb{N}}$ ist *konvergent in* (X,d), wenn sie einen Grenzwert in (X,d) besitzt.

Viele grundlegende Eigenschaften von Cauchyfolgen bzw. konvergenten Folgen in $\mathbb{R}$ bzw. $\mathbb{Q}$ übertragen sich nun ohne weiteres auf die Situation metrischer Räume:

Proposition 1.6 (Grundlegende Eigenschaften von konvergenten Folgen)**.** *Sei (X,d) ein metrischer Raum und sei $(a_n)_{n\in\mathbb{N}}$ eine Folge in X.*

1. *Ist $(a_n)_{n\in\mathbb{N}}$ in X konvergent, so ist $(a_n)_{n\in\mathbb{N}}$ eine Cauchyfolge in X.*
2. *Ist $(a_n)_{n\in\mathbb{N}}$ in X konvergent, so besitzt $(a_n)_{n\in\mathbb{N}}$ genau einen Grenzwert in X und wir bezeichnen diesen mit*
$$\lim_{n\to\infty} a_n.$$
3. *Ist $(a_n)_{n\in\mathbb{N}}$ eine Cauchyfolge in X, so ist die Menge $\{a_n \mid n \in \mathbb{N}\} \subset X$ in X beschränkt; dabei heißt eine Teilmenge $Y \subset X$* beschränkt*, wenn die Menge $\{d(y,y') \mid y,y' \in Y\} \subset \mathbb{R}$ beschränkt ist.*

Caveat 1.7. Im allgemeinen ist nicht in jedem metrischen Raum jede Cauchyfolge bereits konvergent!

Definition 1.8 (vollständig)**.** Ein metrischer Raum (X, d) ist *vollständig*, wenn jede Cauchyfolge in X bereits in X konvergiert.

Bemerkung 1.9 (Vervollständigung metrischer Räume)**.** Analog zur Konstruktion von $\mathbb{R}$ aus $\mathbb{Q}$ kann man mit Hilfe von Äquivalenzklassen von Cauchyfolgen zu jedem metrischen Raum einen „geeigneten“ vollständigen metrischen Raum konstruieren.

Abbildungen zwischen metrischen Räumen

Wie in jeder mathematischen Theorie wollen wir nicht nur Objekte, sondern auch „strukturerhaltende“ Abbildungen zwischen diesen Objekten untersuchen. Je nachdem wie stark die Metrik von Abbildungen zwischen metrischen Räumen respektiert werden soll, erhalten wir verschiedene Begriffe. Wir beginnen mit den stärksten Bedingungen und schwächen diese im Verlauf weiter ab.

Definition 1.10 (Isometrie)**.** Seien (X, d_X) und (Y, d_Y) metrische Räume.

- Eine Abbildung $f\colon X \longrightarrow Y$ ist eine *isometrische Einbettung*, wenn f die Metrik erhält, d.h., wenn

$$\forall_{x,x' \in X} \quad d_Y\bigl(f(x), f(x')\bigr) = d_X(x, x').$$

- Eine Abbildung $f\colon X \longrightarrow Y$ ist eine *Isometrie*, wenn sie eine isometrische Einbettung ist und es eine isometrische Einbettung $g\colon Y \longrightarrow X$ mit $f \circ g = \mathrm{id}_Y$ und $g \circ f = \mathrm{id}_X$ gibt.

Bemerkung 1.11. Seien (X, d_X) und (Y, d_Y) metrische Räume.

1. Jede isometrische Einbettung $X \longrightarrow Y$ ist injektiv.
2. Eine isometrische Einbettung $X \longrightarrow Y$ ist genau dann eine Isometrie, wenn sie bijektiv ist.

Bemerkung 1.12 (Isometriegruppe)**.** Ist (X, d) ein metrischer Raum, so bildet die Menge

$$\mathrm{Isom}(X) := \bigl\{f\colon X \longrightarrow X \bigm| f \text{ ist eine Isometrie}\bigr\}$$

bezüglich Komposition von Abbildungen eine Gruppe, die sogenannte *Isometriegruppe von X*.

Geometrisch interpretiert man Isometrien als „Symmetrien“ von geometrischen Objekten und Isometriegruppen sind historisch einer der Ursprünge für den abstrakten Gruppenbegriff.

Eine erste Abschwächung des Isometriebegriffs führt zu Lipschitz-Abbildungen:

Definition 1.13 (Lipschitz, kontrahierend)**.** Seien (X, d_X) und (Y, d_Y) metrische Räume und sei $f\colon X \longrightarrow Y$ eine Abbildung.

- Eine Zahl $L \in \mathbb{R}_{\geq 0}$ ist eine *Lipschitz-Konstante für f*, falls

$$\forall_{x,x' \in X} \quad d_Y\bigl(f(x), f(x')\bigr) \leq L \cdot d_X(x, x').$$

- Besitzt f eine Lipschitz-Konstante, so ist f *Lipschitz*.

– Besitzt f eine Lipschitz-Konstante, die echt kleiner als 1 ist, so ist f *kontrahierend*.

Bemerkung 1.14.
– Alle Isometrien sind Lipschitz, aber im allgemeinen nicht kontrahierend.
– Alle Funktionen in $C^1([0,1],\mathbb{R})$ sind Lipschitz (bezüglich der Standardmetrik auf $[0,1]$ bzw. $\mathbb{R}$).

Lässt man nicht nur einen gleichmäßigen multiplikativen Fehler, sondern auch einen gleichmäßigen additiven Fehler zu, so gelangt man vom Begriff der Lipschitz-Abbildungen zu den sogenannten quasi-isometrischen Einbettungen bzw. Quasi-Isometrien, die in der geometrischen Gruppentheorie eine wichtige Rolle spielen.

Eine weitere Abschwächung ist der Stetigkeitsbegriff – hierbei bleibt nur noch Konvergenz erhalten. Wir definieren Stetigkeit genauso wie im Fall von Funktionen vom Typ $\mathbb{R} \longrightarrow \mathbb{R}$ über die Verträglichkeit mit Approximation:

Definition 1.15 (stetig)**.** Seien (X, d_X) und (Y, d_Y) metrische Räume. Eine Abbildung $f \colon X \longrightarrow Y$ heißt *stetig*, wenn folgendes gilt: Für alle Folgen $(a_n)_{n\in\mathbb{N}}$ in X, die gegen $a \in X$ konvergieren, ist $(f(a_n))_{n\in\mathbb{N}}$ in Y konvergent und es gilt

$$\lim_{n\to\infty} f(a_n) = f(a).$$

Bemerkung 1.16.
– Für Funktionen vom Typ $X \longrightarrow \mathbb{R}$ mit $X \subset \mathbb{R}$ stimmt dieser Stetigkeitsbegriff (bezüglich der Standardmetrik) mit dem klassischen Stetigkeitsbegriff überein.
– Alle Lipschitz-Abbildungen sind stetig.

Caveat 1.17. Im allgemeinen sind *nicht* alle stetigen Abbildungen zwischen metrischen Räumen Lipschitz!

Wie im klassischen Fall können wir stetige Funktionen auch durch das ε-δ-Kriterium charaktierisieren:

Proposition 1.18 (ε-δ-Kriterium)**.** *Seien (X, d_X) und (Y, d_Y) metrische Räume. Dann ist eine Abbildung $f \colon X \longrightarrow Y$ genau dann stetig, wenn*

$$\forall_{x\in X} \quad \forall_{\varepsilon\in\mathbb{R}_{>0}} \quad \exists_{\delta\in\mathbb{R}_{>0}} \quad \forall_{x'\in X} \quad \bigl(d_X(x,x') < \delta \Longrightarrow d_Y(f(x), f(x')) < \varepsilon\bigr).$$

Auch andere grundlegende Eigenschaften stetiger Funktionen übertragen sich auf den Kontext metrischer Räume; wir gehen hier aber nicht genauer darauf ein, da wir dies dann im noch allgemeineren Kontext der topologischen Räume genauer behandeln werden (s. Kapitel 2).

Der Banachsche Fixpunktsatz

Fixpunktsätze sind ein wichtiges Hilfsmittel, da sie erlauben, auf die Lösbarkeit gewisser Gleichungen zu schließen; ein grundlegendes Beispiel ist der Banachsche Fixpunktsatz:

Satz 1.19 (Der Banachsche Fixpunktsatz)**.** *Sei (X, d) ein nicht-leerer vollständiger metrischer Raum und sei $f\colon X \longrightarrow X$ eine kontrahierende Abbildung.*

1. *Dann besitzt f genau einen Fixpunkt $\widehat{x}$ in X.*
2. *Genauer gilt: Sei $c \in [0, 1)$ eine Lipschitz-Konstante für f und sei $x_0 \in X$. Sei $(x_n)_{n\in\mathbb{N}}$ die induktiv durch $x_{n+1} := f(x_n)$ für alle $n \in \mathbb{N}$ definierte Folge in X. Dann gilt für alle $n \in \mathbb{N}$, dass*

$$d(x_n, \widehat{x}) \leq \frac{c^n}{1-c} \cdot d(x_1, x_0).$$

Man beachte dabei, dass der Banachsche Fixpunktsatz (im Gegensatz zum Brouwerschen Fixpunktsatz) konstruktiv ist und sogar eine Abschätzung angibt, wie weit die konstruierten Punkte höchstens vom gesuchten Fixpunkt entfernt sind.

Anwendungen des Banachschen Fixpunktsatzes sind zum Beispiel:

– Der Banachsche Fixpunktsatz erlaubt es, iterativ approximativ gewissen Gleichungen zu lösen.
– Der Satz über lokale Umkehrbarkeit von Funktionen in mehreren Variablen (Satz 3.49).
– Der Satz von Picard-Lindelöf (Satz 5.5) über die Existenz und Eindeutigkeit von Lösungen gewisser gewöhnlicher Differentialgleichungen.

Normierte Räume

Normierte Räume sind reelle Vektorräume mit einem Längenbegriff, der mit der algebraischen Struktur verträglich ist. Aus Normen erhält man Metriken und somit eine große und wichtige Klasse von Beispielen metrischer Räume.

Definition 1.20 (normierter Raum)**.** Ein *normierter Raum* ist ein Paar $(V, \|\cdot\|)$, wobei V ein $\mathbb{R}$-Vektorraum und $\|\cdot\|$ eine *Norm auf* V ist, d.h. $\|\cdot\|$ ist eine Abbildung vom Typ $V \longrightarrow \mathbb{R}_{\geq 0}$ mit folgenden Eigenschaften:

– Für alle $x \in V$ ist genau dann $\|x\| = 0$, wenn $x = 0$ ist.
– Für alle $x \in V$ und alle $c \in \mathbb{R}$ ist $\|c \cdot x\| = |c| \cdot \|x\|$.
– *Dreiecksungleichung.* Für alle $x, y \in V$ ist

$$\|x + y\| \leq \|x\| + \|y\|.$$

Proposition 1.21 (metrische Räume aus normierten Räumen)**.** *Sei $(V, \|\cdot\|)$ ein normierter Raum. Dann ist*

$$\begin{aligned} V \times V &\longrightarrow \mathbb{R}_{\geq 0} \\ (x, y) &\longmapsto \|x - y\| \end{aligned}$$

eine Metrik auf V, die sogenannte von $\|\cdot\|$ auf V induzierte Metrik.

Bemerkung 1.22 (1-Norm/∞-Norm)**.**

1. Ist $n \in \mathbb{N}$, so sind

$$\begin{aligned} \|\cdot\|_1 \colon \mathbb{R}^n &\longrightarrow \mathbb{R}_{\geq 0} \\ x &\longmapsto \sum_{j=1}^{n} |x_j|, \\ \|\cdot\|_\infty \colon \mathbb{R}^n &\longrightarrow \mathbb{R}_{\geq 0} \\ x &\longmapsto \max_{j \in \{1,\dots,n\}} |x_j| \end{aligned}$$

   Normen auf $\mathbb{R}^n$.

2. Analog sind auch

$$\begin{aligned} \|\cdot\|_1 \colon C([0,1],\mathbb{R}) &\longrightarrow \mathbb{R}_{\geq 0} \\ f &\longmapsto \int_0^1 |f|, \\ \|\cdot\|_\infty \colon C([0,1],\mathbb{R}) &\longrightarrow \mathbb{R}_{\geq 0} \\ f &\longmapsto \sup_{x \in [0,1]} |f(x)| \end{aligned}$$

   Normen auf $C([0,1],\mathbb{R})$.
   Konvergenz bezüglich der von der Norm $\|\cdot\|_\infty$ induzierten Metrik auf $C([0,1],\mathbb{R})$ stimmt mit gleichmäßiger Konvergenz von Funktionen in $C([0,1],\mathbb{R})$ überein. Da gleichmäßige Grenzwerte stetiger Funktionen stetig sind, ist $C([0,1],\mathbb{R})$ bezüglich der von $\|\cdot\|_\infty$ induzierten Metrik vollständig, und somit ein sogenannter *Banachraum.*
   Der Raum $(C([0,1],\mathbb{R}), \|\cdot\|_1)$ ist jedoch *nicht* vollständig; diese Tatsache spielt in der Maßtheorie bei der Konstruktion des Lebesgue-Integrals eine entscheidende Rolle.

3. Ähnlich sind die Abbildungen

$$\begin{aligned} \|\cdot\|_1 \colon \ell_1 &\longrightarrow \mathbb{R}_{\geq 0} \\ (a_n)_{n\in\mathbb{N}} &\longmapsto \sum_{n=0}^{\infty} |a_n|, \\ \|\cdot\|_\infty \colon \ell_\infty &\longrightarrow \mathbb{R}_{\geq 0} \\ (a_n)_{n\in\mathbb{N}} &\longmapsto \sup_{n\in\mathbb{N}} |a_n| \end{aligned}$$

   Normen auf den $\mathbb{R}$-Vektorräumen

$$\begin{aligned} \ell_1 &:= \bigl\{(a_n)_{n\in\mathbb{N}} \bigm| (a_n)_{n\in\mathbb{N}} \text{ ist eine Folge in } \mathbb{R}, \text{ für die } \textstyle\sum_{n=0}^\infty |a_n| \text{ konvergiert}\bigr\} \\ \ell_\infty &:= \bigl\{(a_n)_{n\in\mathbb{N}} \bigm| (a_n)_{n\in\mathbb{N}} \text{ ist eine beschränkte Folge in } \mathbb{R}\bigr\}. \end{aligned}$$

In Anlehnung an den Satz von Pythagoras erhält man die euklidische Norm bzw. Metrik, die dem Längenbegriff unserer Anschauung in der Ebene $\mathbb{R}^2$ bzw. im Raum $\mathbb{R}^3$ entspricht:

Proposition 1.23 (euklidische Norm/Metrik)**.** *Sei $n \in \mathbb{N}$. Dann ist*

$$\begin{aligned} \|\cdot\|_2 \colon \mathbb{R}^n &\longrightarrow \mathbb{R}_{\geq 0} \\ x &\longmapsto \sqrt{\sum_{j=1}^{n} |x_j|^2} \end{aligned}$$

eine Norm auf $\mathbb{R}^n$, die sogenannte euklidische Norm*. Die von $\|\cdot\|_2$ auf $\mathbb{R}^n$ induzierte Metrik heißt* euklidische Metrik.

Um zu zeigen, dass diese Abbildung tatsächlich die Dreiecksungleichung erfüllt drückt man $\|\cdot\|_2$ mit Hilfe des Standardskalarproduktes auf $\mathbb{R}^2$ aus und verwendet dann die folgende Abschätzung:

Lemma 1.24 (Cauchy-Schwarzsche Ungleichung)**.** *Sei V ein reeller Vektorraum und sei $\langle \cdot, \cdot \rangle \colon V \times V \longrightarrow \mathbb{R}$ ein Skalarprodukt auf V. Dann gilt für alle $x, y \in V$, dass*

$$\langle x, y \rangle^2 \leq \langle x, x \rangle \cdot \langle y, y \rangle$$

bzw. $|\langle x, y \rangle| \leq \sqrt{\langle x, x \rangle} \cdot \sqrt{\langle y, y \rangle}$. Ist $y \neq 0$, so gilt genau dann Gleichheit in der obigen Abschätzung, wenn $x = \langle x, y \rangle / \langle y, y \rangle \cdot y$ ist.

Bemerkung 1.25 (2-Norm/p-Norm)**.**

1. Analog kann man zeigen, dass

$$\begin{aligned} \|\cdot\|_2 \colon C([0,1], \mathbb{R}) &\longrightarrow \mathbb{R}_{\geq 0} \\ f &\longmapsto \sqrt{\int_0^1 |f|^2}, \\ \|\cdot\|_2 \colon \ell_2 &\longrightarrow \mathbb{R}_{\geq 0} \\ (a_n)_{n \in \mathbb{N}} &\longmapsto \sqrt{\sum_{n=0}^{\infty} |a_n|^2} \end{aligned}$$

   Normen auf $C([0,1], \mathbb{R}^2)$ bzw.

$$\ell_2 := \bigl\{ (a_n)_{n \in \mathbb{N}} \bigm| (a_n)_{n \in \mathbb{N}} \text{ ist eine Folge in } \mathbb{R}, \text{ für die } \textstyle\sum_{n=0}^{\infty} |a_n|^2 \text{ konvergiert} \bigr\}$$

   sind.

2. Außerdem kann man für jedes $p \in \mathbb{R}_{>1}$ in derselben Art und Weise auf geeigneten Räumen sogenannte p-Normen $\|\cdot\|_p :=$ „$\sqrt[p]{\sum |\cdot|^p}$“ definieren; statt der Cauchy-Schwarzschen Ungleichung benötigt man dann die Minkowskische Ungleichung.

Proposition 1.26 (Vollständigkeit euklidischer Räume)**.** *Ist $n \in \mathbb{N}$, so ist $\mathbb{R}^n$ bezüglich der von der euklidischen Norm induzierten Metrik vollständig.*

Korollar 1.27. *Insbesondere ist die Betragsfunktion*

$$\begin{aligned} |\cdot| \colon \mathbb{C} &\longrightarrow \mathbb{R}_{\geq 0} \\ z &\longmapsto \sqrt{(\operatorname{Re} z)^2 + (\operatorname{Im} z)^2} = \sqrt{z \cdot \overline{z}} \end{aligned}$$

eine Norm auf dem $\mathbb{R}$-Vektorraum $\mathbb{C}$ und $\mathbb{C}$ ist bezüglich der induzierten Metrik vollständig.

Bemerkung 1.28. Man kann zeigen, dass alle Normen auf einem endlich-dimensionalen $\mathbb{R}$-Vektorraum in einem gewissen Sinne äquivalent sind und insbesondere denselben Konvergenzbegriff liefern (Satz 2.44).

Definition 1.29 (Reihen in normierten Vektorräumen)**.** Sei $(V, \|\cdot\|)$ ein normierter Vektorraum und sei $(a_n)_{n\in\mathbb{N}}$ eine Folge in V.

- Die Reihe $\sum_{n=0}^{\infty} a_n$ *konvergiert in V bezüglich* $\|\cdot\|$, wenn die Folge $(\sum_{k=0}^{n} a_k)_{n\in\mathbb{N}}$ der Partialsummen in V bezüglich der von $\|\cdot\|$ induzierten Metrik konvergiert; in diesem Fall nennt man den Grenzwert auch *Wert der Reihe* $\sum_{n=0}^{\infty} a_n$.
- Ist V ein Banachraum und konvergiert $\sum_{n=0}^{\infty} \|a_n\|$ in $\mathbb{R}$, so nennt man $\sum_{n\in\mathbb{N}} a_n$ *absolut konvergent.*

Die Eigenschaften (absolut) konvergenter Reihen in $\mathbb{R}$ übertragen sich nun mehr oder weniger wörtlich auf Reihen in (vollständigen) normierten Räumen.

Bemerkung 1.30 (komplexe Exponentialfunktion)**.** Analog zur reellen Exponentialfunktion definiert man die *komplexe Exponentialfunktion*

$$\begin{aligned} \exp \colon \mathbb{C} &\longrightarrow \mathbb{C} \\ z &\longmapsto \sum_{n=0}^{\infty} \frac{z^n}{n!} \end{aligned}$$

(dass die Exponentialreihe auf ganz $\mathbb{C}$ (absolut) konvergiert, folgt mit denselben Argumenten wie im reellen Fall). Außerdem zeigen Argumente analog zum reellen Fall, dass

$$\exp(x + y) = \exp(x) \cdot \exp(y)$$

für alle $x, y \in \mathbb{C}$ gilt. Zusammen mit der Beobachtung, dass für alle $x \in \mathbb{R}$ die Beziehungen

$$\cos x = \operatorname{Re}\bigl(\exp(i \cdot x)\bigr) \qquad \text{und} \qquad \sin x = \operatorname{Im}\bigl(\exp(i \cdot x)\bigr)$$

gelten, erhält man daraus die Additionstheoreme für Sinus bzw. Kosinus und Periodizitätsaussagen.

Zum Abschluss geben wir noch eine Charakterisierung für Stetigkeit von linearen Abbildungen zwischen normierten Räumen:

Satz 1.31 (Stetigkeit und Beschränktheit)**.** *Seien $(V, \|\cdot\|_V)$ und $(W, \|\cdot\|_W)$ normierte Vektorräume und sei $f \colon V \longrightarrow W$ eine* lineare *Abbildung. Dann sind äquivalent:*

1. *Die Abbildung $f\colon V \longrightarrow W$ ist bezüglich den von den Normen induzierten Metriken stetig.*
2. *Die Abbildung $f\colon V \longrightarrow W$ ist* beschränkt *in dem Sinne, dass*

$$\bigl\{\|f(x)\|_W \bigm| x \in V,\ \|x\|_V \leq 1\bigr\}$$

   in $\mathbb{R}$ beschränkt ist; in diesem Fall heißt

$$\|f\| := \sup\bigl\{\|f(x)\|_W \bigm| x \in V,\ \|x\|_V \leq 1\bigr\}$$

   Norm von f.

Die Funktionalanalysis beschäftigt sich systematisch mit (unendlich-dimensionalen) normierten Vektorräumen und stetigen linearen Abbildungen dazwischen.

2 Topologische Grundlagen

Im vorigen Kapitel haben wir Konvergenz und Stetigkeit auf metrische Räume verallgemeinert. Wir werden nun einen weiteren Abstraktionsschritt in Angriff nehmen – den Schritt zu den topologischen Räumen:

Während man in metrischen Räumen einen quantitativen Begriff von Nähe hat (nämlich durch die Metrik), schwächt man diesen für topologische Räume zu einem qualitativen Begriff von Nähe ab. Anschaulich gesprochen erhält man so eine Art Geometrie, die ohne Begriffe wie Länge und Winkel auskommt und die „prinzipielle“ Form von Objekten beschreibt.

Es wird sich herausstellen, dass sich mit dieser Abstraktion der Stetigkeitsbegriff und damit zusammenhängende Konzepte wie zum Beispiel Kompaktheit eleganter und effizienter formalisieren lässt.

Topologische Räume

Die Grundidee topologischer Räume ist, Nähe nicht durch Abstände, sondern durch Systeme von Teilmengen auszudrücken – den sogenannten offenen Mengen:

Definition 2.1 (topologischer Raum, Topologie)**.** Ein *topologischer Raum* ist ein Paar (X, T), wobei X eine Menge und T eine *Topologie auf* X ist, d.h. T ist eine Teilmenge von $P(X)$ mit folgenden Eigenschaften:

- Es ist $\emptyset \in T$ und $X \in T$.
- Ist $U \subset T$, so ist $\bigcup U \in T$.
- Ist $U \subset T$ endlich, so ist $\bigcap U \in T$.

Die Elemente von T heißen *offene Mengen* (bezüglich T); ist $A \subset X$ und $X \setminus A \in T$, so heißt A *abgeschlossen* (bezüglich T).

Warum man gerade diese Axiome für offene Mengen betrachtet, kann man gut anhand des Beispiels metrischer Räume illustrieren:

Proposition 2.2 (die von einer Metrik induzierte Topologie)**.** *Sei (X, d) ein metrischer Raum. Dann ist*

$$T := \bigl\{U \subset X \bigm| \forall_{x \in U} \quad \exists_{\varepsilon \in \mathbb{R}_{>0}} \quad U(x, \varepsilon) \subset U\bigr\}$$

eine Topologie auf X. Man nennt T die von d auf X induzierte Topologie.

Bemerkung 2.3.

- Der Begriff offener Mengen bezüglich der Standardmetrik auf $\mathbb{R}$ stimmt also mit dem Begriff aus der Analysis I überein.
- Die von der euklidischen Metrik auf $\mathbb{R}^n$ induzierte Topologie auf $\mathbb{R}^n$ heißt *Standardtopologie auf* $\mathbb{R}^n$.
- Ist (X, d) ein metrischer Raum, ist $x \in X$ und ist $\varepsilon \in \mathbb{R}_{>0}$, so ist $U(x, \varepsilon)$ in X offen. D.h. offene Bälle sind tatsächlich offene Mengen bezüglich der von der Metrik induzierten Topologie.

– Analog zum Fall von Teilmengen von $\mathbb{R}$ kann Abgeschlossenheit in metrischen Räumen auch durch Konvergenz von Folgen ausgedrückt werden. Daraus folgt, dass abgeschlossene Bälle tatsächlich bezüglich der von der Metrik induzierten Topologie abgeschlossen sind.

Caveat 2.4. Nicht jeder topologische Raum ist metrisierbar! (Korollar 2.28).

Bemerkung 2.5 (Klumpentopologie, diskrete Topologie)**.** Sei X eine Menge.
- Dann ist $\{\emptyset, X\}$ eine Topologie auf X, die sogenannte *Klumpentopologie auf* X.
- Außerdem ist $P(X)$ eine Topologie auf X, die sogenannte *diskrete Topologie auf* X; iese stimmt mit der von der diskreten Metrik induzierten Topologie überein.

Zwei elementare Konstruktionen topologischer Räume sind Teilräume und Produkte:

Bemerkung 2.6 (Teilraumtopologie)**.** Sei (X, T) ein topologischer Raum und $Y \subset X$. Dann ist
$$\{U \cap Y \mid U \in T\}$$
eine Topologie auf Y, die sogenannte *Teilraumtopologie.* Ist T auf X von einer Metrik d auf X induziert, so stimmt die Teilraumtopologie auf Y mit der von d auf Y induzierten Metrik induzierten Topologie überein.

Bemerkung 2.7 (Produkttopologie)**.** Seien (X, T_X) und (Y, T_Y) topologische Räume. Dann ist
$$\bigl\{U \subset X \times Y \bigm| \forall_{x \in U} \quad \exists_{U_X \in T_X} \quad \exists_{U_Y \in T_Y} \quad x \in U_X \times U_Y \subset U\bigr\}$$
eine Topologie auf $X \times Y$, die sogenannte *Produkttopologie.*

Die Standardtopologie auf $\mathbb{R}^2 = \mathbb{R} \times \mathbb{R}$ stimmt dabei mit der Produkttopologie der Standardtopologie auf $\mathbb{R}$ überein.

Außerdem ist es oft nützlich, die folgenden Begriffe zur Verfügung zu haben:

Definition 2.8 ((offene) Umgebung)**.** Sei (X, T) ein topologischer Raum und $x \in X$.
- Eine Teilmenge $U \subset X$ ist eine *offene Umgebung von* x, wenn U offen ist und $x \in U$ ist.
- Eine Teilmenge $U \subset X$ ist eine *Umgebung von* x, wenn es eine offene Umgebung $V \subset X$ von x mit $V \subset U$ gibt.

Definition 2.9 (Abschluss, Inneres, Rand)**.** Sei (X, T) ein topologischer Raum und sei $Y \subset X$.
- Das *Innere von* Y ist
$$Y^\circ := \bigcup \{U \mid U \in T \text{ und } U \subset Y\}.$$
D.h. Y° ist die (bezüglich Inklusion) größte in X offene Menge, die in Y enthalten ist.

- Der *Abschluss von* Y ist

$$\overline{Y} := \bigcap \{A \mid X \setminus A \in T \text{ und } Y \subset A\}.$$

  D.h. $\overline{Y}$ ist die (bezüglich Inklusion) kleinste in X abgeschlossene Menge, die Y enthält.
- Der *Rand von* Y ist

$$\partial Y := \overline{Y} \cap \overline{(X \setminus Y)}.$$

Stetige Abbildungen

Wir haben bereits gesehen, dass für Abbildungen vom Typ $X \longrightarrow \mathbb{R}$ mit $X \subset \mathbb{R}$ (Folgen)Stetigkeit bzw. das ε-δ-Kriterium dazu äquivalent sind, dass Urbilder offener Mengen offen sind. Im Kontext topologischer Räume machen die ersten beiden Formulierungen im allgemeinen keinen Sinn (da die Topologie im allgemeinen nicht von einer Metrik induziert wird), letztere kann aber formuliert werden und wird daher als Definition für Stetigkeit verwendet. Stetige Abbildungen spielen also die Rolle der strukturverträglichen Abbildungen in der Welt der topologischen Räume:

Definition 2.10 (stetig)**.** Seien (X, T_X) und (Y, T_Y) topologische Räume. Eine Abbildung $f\colon X \longrightarrow Y$ ist *stetig* (bezüglich T_X und T_Y), wenn

$$\forall_{U \in T_Y} \quad f^{-1}(U) \in T_X,$$

d.h., wenn Urbilder offener Mengen offen sind.

Bemerkung 2.11.
- Für Abbildungen vom Typ $X \longrightarrow \mathbb{R}$ mit $X \subset \mathbb{R}$ stimmt dieser Begriff von Stetigkeit mit dem aus der Analysis I überein.
- Dasselbe Argument über das ε-δ-Kriterium zeigt: Für Abbildungen zwischen metrischen Räumen stimmt dieser Begriff von Stetigkeit mit dem bereits zuvor definierten überein.
- Sei X eine Menge und seien T bzw. T' Topologien auf X. Dann ist die Identität $\mathrm{id}_X\colon (X, T) \longrightarrow (X, T')$ genau dann stetig, wenn $T' \subset T$ ist, d.h. wenn T' gröber als T ist.
- Die Abbildungen $+, \cdot, -\colon \mathbb{R}^2 \longrightarrow \mathbb{R}$ und $/\colon \mathbb{R} \times (\mathbb{R} \setminus \{0\}) \longrightarrow \mathbb{R}$ sind bezüglich der Standardtopologie stetig.
- Ist (X, T) ein topologischer Raum und $Y \subset X$, so ist die Inklusion $Y \hookrightarrow X$ bezüglich der Teilraumtopologie auf Y stetig.
- Konstante Abbildungen sind stetig.

Proposition 2.12 (Vererbungseigenschaften stetiger Abbildungen)**.** *Seien* (X, T_X)*,* (Y, T_Y) *und* (Z, T_Z) *topologische Räume und seien* $f\colon X \longrightarrow Y$ *und* $g\colon Y \longrightarrow Z$ *Abbildungen.*
1. *Sind* f *und* g *stetig, so ist auch* $g \circ f\colon X \longrightarrow Z$ *stetig.*
2. *Ist* f *stetig und ist* $A \subset X$*, so ist auch die Einschränkung* $f|_A\colon A \longrightarrow Y$ *stetig (bezüglich der Teilraumtopologie auf* A*).*

3. *Die Abbildung $f\colon X \longrightarrow Y$ ist genau dann stetig, wenn $f\colon X \longrightarrow f(X)$ bezüglich der Teilraumtopologie auf $f(X)$ stetig ist.*

Proposition 2.13 (Verkleben stetiger Funktionen)**.** *Seien (X, T_X) und (Y, T_Y) topologische Räume, seien $A, B \subset X$ abgeschlossene Teilmengen mit $A \cup B = X$ und seien $f\colon A \longrightarrow Y$ und $g\colon B \longrightarrow Y$ stetige Abbildungen (bezüglich der Teilraumtopologie auf A bzw. B) mit $f|_{A\cap B} = g|_{A\cap B}$. Dann ist die (wohldefinierte) Abbildung*

$$\begin{aligned} f \cup_{A\cap B} g\colon X &\longrightarrow Y \\ x &\longmapsto \begin{cases} f(x) & \text{falls } x \in A, \\ g(x) & \text{falls } x \in B \end{cases} \end{aligned}$$

stetig.

Der Isomorphiebegriff in der Kategorie der topologischen Räume ist Homöomorphie:

Definition 2.14 (Homöomorphismus)**.** Seien (X, T_X) und (Y, T_Y) topologische Räume. Eine Abbildung $f\colon X \longrightarrow Y$ ist ein *Homöomorphismus*, wenn sie stetig ist und es eine stetige Abbildung $g\colon Y \longrightarrow X$ mit $g \circ f = \mathrm{id}_X$ und $f \circ g = \mathrm{id}_Y$ gibt. Falls es einen Homöomorphismus $X \longrightarrow Y$ gibt, heißen X und Y *homöomorph*.

Caveat 2.15. *Nicht* jede stetige bijektive Abbildung ist ein Homöomorphismus!

Anschaulich gesprochen sind topologische Räume genau dann homöomorph, wenn man sie durch „verbiegen" und „aufblasen/schrumpfen" ineinander überführen kann, ohne zu „schneiden" oder zu „kleben".

Im allgemeinen ist es nicht leicht bzw. sogar nicht möglich zu entscheiden, ob zwei gegebene topologische Räume homöomorph sind oder nicht – da es im allgemeinen sehr viele stetige Abbildungen gibt. In der algebraischen Topologie studiert man algebraische Invarianten, um Probleme dieser Art (partiell) zu lösen.

Zum Abschluss dieses allgemeinen Abschnitts über stetige Abbildungen geben wir noch ein Beispiel für die allgemeine Strategie, Objekte durch ihre strukturerhaltenden Morphismen zu charakterisieren statt durch ihre „Punkte":

Proposition 2.16 (universelle Eigenschaft der Produkttopologie)**.** *Seien (X, T_X) und (Y, T_Y) topologische Räume und seien $\pi_X\colon X \times Y \longrightarrow X$ bzw. $\pi_Y\colon X \times Y \longrightarrow Y$ die Projektionsabbildungen. Ist (Z, T_Z) ein weiterer topologischer Raum, so ist eine Abbildung $f\colon Z \longrightarrow X \times Y$ genau dann stetig, wenn die Abbildungen $\pi_X \circ f\colon Z \longrightarrow X$ und $\pi_Y \circ f\colon Z \longrightarrow Y$ stetig sind.*

Bemerkung 2.17 (universelle Eigenschaften)**.**
- Analog kann man Produkte beliebig vieler topologischer Räume durch die analoge universelle Eigenschaft charakterisieren.
- Die universelle Eigenschaft der Produkttopologie ist das topologische Analogon der universellen Eigenschaft von Produkten (nicht Summen!) von Vektorräumen, Gruppen, …
- Ähnlich kann man auch andere Konstruktionen topologischer Räume durch entsprechende universelle Eigenschaften charakterisieren (z.B. das Verkleben topologischer Räume, was durch Summen bzw. Quotienten modelliert wird).

(Weg)Zusammenhang

Einer der zentralen Sätze über stetige Funktionen vom Typ $[0,1] \longrightarrow \mathbb{R}$ ist der Zwischenwertsatz. Im allgemeineren Kontext der topologischen Räume kann man dieses Phänomen durch die Begriffe Wegzusammenhang und Zusammenhang beschreiben.

Definition 2.18 (Weg, wegzusammenhängend)**.** Sei (X,T) ein topologischer Raum.

- Ein *Weg in* X ist eine stetige Abbildung $\gamma\colon [0,1] \longrightarrow X$ (bezüglich der Standardtopologie auf $[0,1] \subset \mathbb{R}$). Man nennt $\gamma(0)$ den *Startpunkt* und $\gamma(1)$ den *Endpunkt von* γ. Der Weg γ heißt *geschlossen*, wenn $\gamma(0) = \gamma(1)$ ist.
- Der Raum X ist *wegzusammenhängend*, wenn folgendes gilt: Für alle $x, y \in X$ gibt es einen Weg $\gamma\colon [0,1] \longrightarrow X$ mit $\gamma(0) = x$ und $\gamma(1) = y$.

Bemerkung 2.19.

- Das Einheitsintervall ist (bezüglich der Standardtopologie) wegzusammenhängend.
- Ist $n \in \mathbb{N}$, so ist $\mathbb{R}^n$ wegzusammenhängend (bezüglich der Standardtopologie).
- Ist X eine Menge mit $|X| \geq 2$, so ist X bezüglich der diskreten Topologie *nicht* wegzusammenhängend.

Proposition 2.20 (Stetigkeit und Wegzusammenhang)**.** *Seien (X,T_X) und (Y,T_Y) topologische Räume und sei $f\colon X \longrightarrow Y$ stetig. Ist X wegzusammenhängend, so ist auch $f(X)$ bezüglich der Teilraumtopologie wegzusammenhängend.*

Insbesondere gilt: Sind X und Y homöomorph, so ist X genau dann wegzusammenhängend, wenn Y wegzusammenhängend ist. Mit anderen Worten: Wegzusammenhang ist eine topologische Invariante.

Zum Beispiel kann man diese Eigenschaft (und einen kleinen Trick) verwenden, um zu zeigen, dass $\mathbb{R}$ nur dann zu $\mathbb{R}^n$ homöomorph ist (bezüglich der Standardtopologie), wenn $n = 1$ ist.

Eine Abschwächung des Wegzusammenhangsbegriffs ist Zusammenhang:

Definition 2.21 (zusammenhängend)**.** Ein topologischer Raum (X,T_X) ist *zusammenhängend*, wenn folgendes gilt: Für alle $U, V \in T_X$ mit $U \cap V = \emptyset$ und $U \cup V = X$ ist bereits $U = \emptyset$ oder $V = \emptyset$. (D.h. X lässt sich nur trivial in offene Mengen zerlegen).

Bemerkung 2.22. Das Einheitsintervall $[0,1]$ ist bezüglich der Standardtopologie zusammenhängend.

Proposition 2.23 (Wegzusammenhang impliziert Zusammenhang)**.** *Jeder wegzusammenhängende topologische Raum ist zusammenhängend.*

Caveat 2.24. Es gibt topologische Räume, die zusammenhängend, aber *nicht* wegzusammenhängend sind!

In diesem allgemeinen Kontext lautet der verallgemeinerte Zwischenwertsatz nun wie folgt:

Proposition 2.25 (Verallgemeinerter Zwischenwertsatz)**.** *Seien (X,T_X) und (Y,T_Y) topologische Räume und sei $f\colon X \longrightarrow Y$ stetig. Ist X zusammenhängend, so ist auch $f(X)$ zusammenhängend (bezüglich der Teilraumtopologie).*

Insbesondere: Sind X und Y homöomorph, so ist X genau dann zusammenhängend, wenn Y zusammenhängend ist. Mit anderen Worten: Zusammenhang ist eine topologische Invariante.

Hausdorffräume

Die Klumpentopologie zeigt bereits, dass es viele exotische und unintuitive Topologien gibt. Daher gibt es viele Begriffe für topologische Räume, die sicherstellen, dass Räume hinreichend gutartig sind. Ein Beispiel ist der folgende Begriff, der zu den sogenannten Trennungseigenschaften gehört:

Definition 2.26 (hausdorffsch)**.** Ein topologischer Raum (X,T) ist *hausdorffsch*, wenn folgendes gilt: Für alle $x, y \in X$ mit $x \neq y$ existieren offene Mengen $U, V \subset X$ mit $x \in U$, $y \in V$ und $U \cap V = \emptyset$. (D.h. je zwei Punkte können durch offene Mengen getrennt werden.)

Proposition 2.27 (metrische Räume sind hausdorffsch)**.** *Ist (X,d) ein metrischer Raum, so ist X bezüglich der von der Metrik d auf X induzierten Topologie hausdorffsch.*

Korollar 2.28. *Sei X eine Menge mit $|X| \geq 2$. Dann gibt es keine Metrik auf X, die die Klumpentopologie auf X induziert.*

Bemerkung 2.29. Sind (X,T_X) und (Y,T_Y) homöomorphe topologische Räume, so ist X genau dann hausdorffsch, wenn Y hausdorffsch ist.

Es gibt noch weitere Trennungseignschaften topologischer Räume, sowie sogenannte Abzählbarkeitseigenschaften; die Zusammenhänge zwischen diesen Begriffen sind etwas unübersichtlich und können gut in dem Buch *Counterexamples in Topology* von L.A. Steen und J.A. Seebach Jr nachgelesen werden.

Kompaktheit

Einer der wichtigsten und nützlichsten topologischen Begriffe ist Kompaktheit; grob gesprochen ist Kompaktheit eine Art topologische Endlichkeitsbedingung. Wir beginnen mit der abstrakten Definition als Endlichkeitsbedingung von Überdeckungen und zeigen später, dass dieser Begriff den bisherigen Kompaktheitsbegriff in $\mathbb{R}$ verallgemeinert (Satz 2.38).

Definition 2.30 (kompakt)**.** Ein topologischer Raum (X,T) ist *kompakt*, wenn jede offene Überdeckung von X eine endliche Teilüberdeckung enthält, d.h., wenn folgendes gilt: Für alle $U \subset T$ mit $\bigcup U = X$ gibt es eine endliche Teilmenge $V \subset U$ mit $\bigcup V = X$.

Bemerkung 2.31. Sei X eine Menge. Dann ist X bezüglich der Klumpentopologie kompakt. Außerdem ist X genau dann bezüglich der diskreten Topologie kompakt, wenn X endlich ist.

Wir werden später eine Charakterisierung kompakter Mengen in $\mathbb{R}^n$ geben und zunächst allgemeine Eigenschaften kompakter Mengen aus der Definition ableiten:

Proposition 2.32 (Verallgemeinertes Extremalprinzip). *Seien* (X, T_X) *und* (Y, T_Y) *topologische Räume und sei* $f\colon X \longrightarrow Y$ *stetig. Ist* X *kompakt, so ist auch* $f(X)$ *kompakt (bezüglich der Teilraumtopologie).*

Korollar 2.33 (Kompaktheit ist eine topologische Invariante). *Insbesondere gilt: Sind* (X, T_X) *und* (Y, T_Y) *homöomorphe topologische Räume, so ist* X *genau dann kompakt, wenn* Y *kompakt ist.*

Proposition 2.34 (Abgeschlossenheit und Kompaktheit). *Sei* (X, T) *ein topologischer Raum und sei* $Y \subset X$*.*

1. *Ist* X *kompakt und* Y *in* X *abgeschlossen, so ist* Y *bezüglich der Teilraumtopologie kompakt.*
2. *Ist* X *hausdorffsch und* Y *bezüglich der Teilraumtopologie kompakt, so ist* Y *in* X *abgeschlossen.*

Korollar 2.35. *Sei* (X, T_X) *ein kompakter topologischer Raum, sei* (Y, T_Y) *ein Hausdorffraum und sei* $f\colon X \longrightarrow Y$ *stetig und bijektiv. Dann ist* f *bereits ein Homöomorphismus(!).*

Zum Abschluss des Abschnitts über allgemeine Kompaktheitseigenschaften betrachten wir noch die Verträglichkeit von Kompaktheit mit Produkten:

Proposition 2.36 (Produkt zweier kompakter Räume). *Seien* (X, T_X) *und* (Y, T_Y) *kompakte topologische Räume. Dann ist das Produkt* $X \times Y$ *bezüglich der Produkttopologie kompakt.*

Bemerkung 2.37 (Satz von Tychonoff). Der *Satz von Tychnoff*

> Beliebige (auch unendliche!) Produkte kompakter topologischer Räume sind kompakt.

ist äquivalent zum Auswahlaxiom (!), und damit auch äquivalent zum Zornschen Lemma bzw. dem Wohlordnungssatz.

Der Satz von Heine-Borel

Unser nächstes Ziel ist es, eine einfache Charakterisierung kompakter Mengen im euklidischen Raum $\mathbb{R}^n$ zu geben:

Satz 2.38 (Satz von Heine-Borel). *Sei* $n \in \mathbb{N}$ *und sei* $A \subset \mathbb{R}^n$*. Wir betrachten auf* A *die von der Standardtopologie auf* $\mathbb{R}^n$ *induzierte Teilraumtopologie. Dann sind die folgenden Aussagen äquivalent:*

1. *Die Menge* A *ist kompakt.*
2. *Die Menge* A *ist in* $\mathbb{R}^n$ *bezüglich der euklidischen Metrik beschränkt und abgeschlossen.*
3. *Die Menge* A *ist bezüglich der euklidischen Metrik* folgenkompakt*, d.h. jede Folge in* A *besitzt eine bezüglich der euklidischen Metrik konvergente Teilfolge, deren Grenzwert auch in* A *liegt.*

Der Beweis der Implikation „1. $\Longrightarrow$ 2.“ beruht auf der folgenden Beobachtung:

Proposition 2.39. *Kompakte Mengen in metrischen Räumen sind beschränkt und abgeschlossen.*

Caveat 2.40. Die Umkehrung der obigen Proposition gilt im allgemeinen *nicht*! D.h. es gibt beschränkte und abgeschlossene Mengen in gewissen (sogar vollständigen) metrischen Räumen, die *nicht* kompakt sind (zum Beispiel sind unendliche Mengen bezüglich der diskreten Metrik in sich selbst beschränkt und abgeschlossen, aber nicht kompakt).

Der Beweis der Implikation „2. $\Longrightarrow$ 3.“ ist eine einfache Verallgemeinerung des Satzes von Bolzano-Weierstraß.

Der Beweis der Implikation „3. $\Longrightarrow$ 1.“ beruht darauf, dass $\mathbb{R}^n$ das sogenannte *zweite Abzählbarkeitsaxiom* erfüllt und auf der folgenden Variante des Intervallschachtelungsprinzips:

Proposition 2.41 (Intervallschachtelung in folgenkompakten metrischen Räumen)**.** *Sei (X, d) ein folgenkompakter metrischer Raum und sei $(A_n)_{n\in\mathbb{N}}$ eine Folge nichtleerer abgeschlossener Teilmengen von X mit der Eigenschaft, dass $A_{n+1} \subset A_n$ für alle $n \in \mathbb{N}$ erfüllt ist. Dann ist der Durchschnitt $\bigcup_{n\in\mathbb{N}} A_n$ nicht leer.*

Korollar 2.42. *Der allgemeine Kompaktheitsbegriff stimmt für Teilmengen von $\mathbb{R}$ mit dem Folgenkompaktheitsbegriff (bezüglich der Standardmetrik) überein.*

Insbesondere ist nach dem Satz von Heine-Borel jede kompakte Teilmenge von $\mathbb{R}$ beschränkt und abgeschlossen und besitzt somit (falls sie nicht-leer ist) ein Minimum und ein Maximum. Damit folgt aus dem verallgemeinerten Extremalprinzip:

Korollar 2.43 (Verallgemeinertes Extremalprinzip für reellwertige Funktionen)**.** *Sei (X, T) ein kompakter topologischer Raum und sei $f \colon X \longrightarrow \mathbb{R}$ stetig (bezüglich der Standardtopologie auf $\mathbb{R}$).*

1. *Dann ist $f(X) \subset \mathbb{R}$ kompakt, und damit insbesondere beschränkt und abgeschlossen.*
2. *Insbesondere gilt: Ist $X \neq \emptyset$, so nimmt f auf X Minimum und Maximum an, d.h. es existieren $x_-, x_+ \in X$ mit*

$$\forall_{x\in X} \quad f(x_-) \leq f(x) \leq f(x_+).$$

Außerdem liefert der Satz von Heine-Borel einen eleganten Beweis für die Tatsache, dass alle Normen auf einem gegebenen endlich-dimensionalen $\mathbb{R}$-Vektorraum äquivalent sind:

Satz 2.44 (Äquivalenz von Normen auf $\mathbb{R}^n$)**.** *Sei $n \in \mathbb{N}$ und seien $\|\cdot\|$ und $\|\cdot\|'$ Normen auf $\mathbb{R}^n$. Dann sind $\|\cdot\|$ und $\|\cdot\|'$ im folgenden Sinne äquivalent: Es existiert ein $c \in \mathbb{R}_{>0}$ mit*

$$\forall_{x\in\mathbb{R}^n} \quad \frac{1}{c} \cdot \|x\|' \leq \|x\| \leq c \cdot \|x\|'.$$

Insbesondere induzieren $\|\cdot\|$ und $\|\cdot\|'$ dieselbe Topologie auf $\mathbb{R}^n$.

Korollar 2.45. *Seien* $(V, \|\cdot\|_V)$ *und* $(W, \|\cdot\|_W)$ *normierte Vektorräume und sei* $f\colon V \longrightarrow W$ *linear. Ist* V *endlich-dimensional, so ist* f *stetig.*

Caveat 2.46. Auf unendlich-dimensionalen $\mathbb{R}$-Vektorräumen sind nicht alle Normen äquivalent!

3 Mehrdimensionale Analysis – Differenzierbarkeit

Funktionen vom Typ $\mathbb{R}^n \longrightarrow \mathbb{R}^m$ treten in den Anwendungen häufig auf, z.B. bei der Modellierung der Bewegung eines Objekts im euklidischen Raum, bei der Modellierung von Höhenprofilen oder bei der Modellierung von (Magnet)Feldern.

Wir suchen eine lokale Beschreibung von Funktionen vom Typ $\mathbb{R}^n \longrightarrow \mathbb{R}^m$ durch möglichst einfache (d.h. lineare) Abbildungen bzw. wir möchten verstehen, für welche Funktionen eine solche lokale Beschreibung möglich ist.

Dies führt zum Begriff der (totalen) Differenzierbarkeit und der Ableitung.

Allerdings ist die mehrdimensionale Theorie etwas vielfältiger als die der differenzierbaren Funktionen vom Typ $\mathbb{R} \longrightarrow \mathbb{R}$, da es viele verschiedene interessante und auch für die Anwendungen wichtige Differentialoperatoren gibt.

Wir beginnen mit der Entwicklung der Theorie analog zum eindimensionalen Fall, leiten erste Vererbungseigenschaften her, und untersuchen dann, inwieweit sich mehrdimensionale Differenzierbarkeit in viele eindimensionale Probleme übersetzen lässt (was zur partiellen Differenzierbarkeit führt). Außerdem betrachten wir wie im eindimensionalen Fall Optimierungsprobleme als Anwendung der Differentialrechnung. Langfristig erhalten wir dadurch die Werkzeuge, die nötig sind, um Kurven und höherdimensionale geometrische Objekte wie Flächen etc. besser zu verstehen.

Notation 3.1. Seien $n, m \in \mathbb{N}$.
- Im folgenden sind Elemente von $\mathbb{R}^n$ immer als Spaltenvektoren zu verstehen.
- Wir schreiben $L(\mathbb{R}^n, \mathbb{R}^m)$ für die Menge der $\mathbb{R}$-linearen Abbildungen $\mathbb{R}^n \longrightarrow \mathbb{R}^m$.
- Die Standardbasis von $\mathbb{R}^n$ bezeichnen wir mit $(e_1, \dots, e_n)$.

Totale Differenzierbarkeit

Analog zum eindimensionalen Fall nennen wir eine Funktion vom Typ $\mathbb{R}^n \longrightarrow \mathbb{R}^m$ (total) differenzierbar, wenn sie sich lokal gut durch eine lineare Abbildung beschreiben lässt:

Definition 3.2 ((total) differenzierbar, Ableitung)**.** Seien $n, m \in \mathbb{N}$, sei $X \subset \mathbb{R}^n$ offen und sei $f \colon X \longrightarrow \mathbb{R}^m$ eine Abbildung.
- Sei $x \in X$. Dann ist f *in* x *(total) differenzierbar*, wenn es eine lineare Abbildung $A \in L(\mathbb{R}^n, \mathbb{R}^m)$ mit folgender Eigenschaft gibt: es gibt eine offene Umgebung $U \subset \mathbb{R}^n$ um 0 mit $x + U \subset X$ und eine Abbildung $E \colon U \longrightarrow \mathbb{R}^m$ mit

  $$\forall_{h \in U} \quad f(x+h) = f(x) + A(h) + E(h)$$

  und (bezüglich der euklidischen Metrik)

  $$\lim_{U\setminus\{0\} \ni h \to 0} \frac{1}{\|h\|_2} \cdot E(h) = 0.$$

  Man kann zeigen, dass A in diesem Fall eindeutig bestimmt ist, und wir nennen

  $$f'(x) := A \in L(\mathbb{R}^n, \mathbb{R}^m)$$

  die *Ableitung von* f *in* x.

– Die Abbildung $f\colon X \longrightarrow \mathbb{R}^m$ ist *auf X differenzierbar*, wenn f in jedem Punkt von X differenzierbar ist.

Im pathologischen Fall, dass der Startraum $\mathbb{R}^0$ ist, sind alle Funktionen $\mathbb{R}^0 \longrightarrow \mathbb{R}^m$ differenzierbar und die Ableitung im Punkt $0 \in \mathbb{R}^0$ ist $0 \in L(\mathbb{R}^0, \mathbb{R}^m)$.

Bemerkung 3.3 (Ableitungen von Abbildungen vom Typ $\mathbb{R} \longrightarrow \mathbb{R}$)**.** Im Falle von differenzierbaren Abbildungen $f\colon X \longrightarrow \mathbb{R}$, wobei $X \subset \mathbb{R}$ offen ist, verwenden wir die Identifikation

$$\begin{aligned} L(\mathbb{R},\mathbb{R}) &\longleftrightarrow \mathbb{R} \\ g &\longmapsto g(x) \\ (x \mapsto a \cdot x) &\longleftarrow\!\shortmid a \end{aligned}$$

und fassen so zu $x \in X$ die Ableitung $f'(x) \in L(\mathbb{R}, \mathbb{R})$ als reelle Zahl auf. Diese stimmt dann mit der gewöhnlichen Ableitung im eindimensionalen Fall überein.

Für die Ableitung $f'(x)$ gibt es viele weitere gebräuchliche Notationen, z.B. $Df(x)$, $df(x)$, ...

Bemerkung 3.4 (Lokalität von Differenzierbarkeit)**.** Aus der Definition bzw. der Wohldefiniertheit folgt, dass Differenzierbarkeit eine lokale Eigenschaft ist: Sei $X \subset \mathbb{R}^n$ offen und sei $f\colon X \longrightarrow \mathbb{R}^m$ eine Abbildung. Ist $x \in X$ und ist $U \subset X$ eine offene Umgebung von x, so ist f genau dann in x differenzierbar, wenn $f|_U\colon U \longrightarrow \mathbb{R}^m$ in x differenzierbar ist; und es gilt im differenzierbaren Fall, dass

$$f'(x) = (f|_U)'(x) \in L(\mathbb{R}^n, \mathbb{R}^m).$$

Caveat 3.5 (Differentialquotient)**.** Da es in $\mathbb{R}^n$ im allgemeinen keine vernünftige Divisionsoperation gibt, gibt es auch keine offensichtliche Beschreibung der Ableitung als Differentialquotient.

Proposition 3.6 (Differenzierbarkeit impliziert Stetigkeit)**.** *Sei $X \subset \mathbb{R}^n$ offen, sei $x \in X$ und sei $f\colon X \longrightarrow \mathbb{R}^m$ in x differenzierbar. Dann ist f auch in x stetig (bezüglich der euklidischen Metrik).*

Den Beweis dieser Proposition kann man analog zum eindimensionalen Fall führen, indem man das folgenden Übersetzungsprinzip anwendet:

$\mathbb{R} \longrightarrow \mathbb{R}$	$\rightsquigarrow$	allgemeiner Fall $\mathbb{R}^n \longrightarrow \mathbb{R}^m$
$\frac{1}{h}$		$\frac{1}{\|h\|_2}$
$f'(x) \cdot h$		$\bigl(f'(x)\bigr)(h)$
Stetigkeit der Skalarmultiplikation		Stetigkeit von $f'(x)\colon \mathbb{R}^n \longrightarrow \mathbb{R}^m$

Caveat 3.7. Die Umkehrung gilt im allgemeinen *nicht*, d.h. *nicht* jede stetige Funktion $\mathbb{R}^n \longrightarrow \mathbb{R}^m$ ist bereits differenzierbar.

Caveat 3.8 (komplexe Differenzierbarkeit)**.** Komplexe Differenzierbarkeit von Funktionen $\mathbb{C} \longrightarrow \mathbb{C}$ kann über den komplexen Differentialquotient definiert werden und ist *nicht* dasselbe wie (reelle) Differenzierbarkeit von Funktionen $\mathbb{R}^2 \longrightarrow \mathbb{R}^2$. Die Eigenschaften von komplex differenzierbaren Funktionen sind der Gegenstand der sogenannten Funktionentheorie.

Grundlegende Vererbungseigenschaften von Differenzierbarkeit

Analog zum eindimensionalen Fall gelten auch im mehrdimensionalen entsprechende Vererbungseigenschaften:

Proposition 3.9 (Linearität von Differenzierbarkeit)**.** *Sei $X \subset \mathbb{R}^n$ offen, sei $x \in X$, seien $f, g\colon X \longrightarrow \mathbb{R}^m$ in x differenzierbar, und seien $a, b \in \mathbb{R}$. Dann ist auch $a \cdot f + b \cdot g\colon X \longrightarrow \mathbb{R}^m$ in x differenzierbar und es gilt*

$$(a \cdot f + b \cdot g)'(x) = a \cdot f'(x) + b \cdot g'(x) \in L(\mathbb{R}^n, \mathbb{R}^m).$$

Proposition 3.10 (Kettenregel)**.** *Seien $X \subset \mathbb{R}^n$ und $Y \subset \mathbb{R}^m$ offene Mengen und seien $f\colon X \longrightarrow Y \subset \mathbb{R}^m$ und $g\colon Y \longrightarrow \mathbb{R}^k$ Abbildungen. Außerdem sei $x \in X$. Ist f in x differenzierbar und ist g in $f(x)$ differenzierbar, so ist $g \circ f\colon X \longrightarrow \mathbb{R}^k$ in x differenzierbar und es gilt*

$$(g \circ f)'(x) = g'(f(x)) \circ f'(x) \in L(\mathbb{R}^n, \mathbb{R}^k).$$

Mit Hilfe der Kettenregel und der Tatsache, dass lineare Isomorphismen die Dimension erhalten, folgt die Invarianz der Dimension unter Diffeomorphismen:

Definition 3.11 (Diffeomorphismus)**.** Seien $n, m \in \mathbb{N}$ und seien $X \subset \mathbb{R}^n$ und $Y \subset \mathbb{R}^m$ offen. Eine Abbildung $f\colon X \longrightarrow Y \subset \mathbb{R}^m$ ist ein *Diffeomorphismus*, wenn f auf X differenzierbar ist und es eine differenzierbare Abbildung $g\colon Y \longrightarrow X \subset \mathbb{R}^n$ mit

$$g \circ f = \mathrm{id}_X \qquad \text{und} \qquad f \circ g = \mathrm{id}_Y$$

gibt. Falls es einen Diffeomorphismus $X \longrightarrow Y$ gibt, so sind X und Y *diffeomorph*.

Caveat 3.12. *Nicht* jede bijektive differenzierbare Abbildung ist ein Diffeomorphismus!

Je nach Situation/Literatur kann der Begriff des Diffeomorphismus auch stärkere Differenzierbarkeitsbedingungen enthalten als in obiger Definition.

Korollar 3.13 (Invarianz der Dimension)**.** *Seien $m, n \in \mathbb{N}$ und seien $X \subset \mathbb{R}^n$ und $Y \subset \mathbb{R}^m$ nicht-leere offene Mengen.*

1. *Ist $x \in X$ und $f\colon X \longrightarrow Y$ ein Diffeomorphismus, so ist $f'(x) \in L(\mathbb{R}^n, \mathbb{R}^m)$ eine invertierbare lineare Abbildung.*
2. *Insbesondere: Sind X und Y diffeomorph, so ist $n = m$.*

Weitere Vererbungseigenschaften von Differenzierbarkeit und lokale Umkehrbarkeitskriterien werden wir später kennenlernen.

Partielle Differenzierbarkeit

Bisher haben wir außer der Definition und den ersten Vererbungseigenschaften noch keine praktikablen Kriterien um festzustellen, ob eine Funktion $\mathbb{R}^n \longrightarrow \mathbb{R}^m$ differenzierbar ist, oder gar die Ableitung in einem Punkt zu bestimmen. Wir werden im folgenden untersuchen, inwiefern man dieses mehrdimensionale Problem durch (viele) eindimensionale Probleme ersetzen kann. Dazu führen wir Richtungsableitungen und partielle Differenzierbarkeit ein:

Definition 3.14 (Richtungsableitung, partielle Ableitung)**.** Sei $X \subset \mathbb{R}^n$ offen und sei $f\colon X \longrightarrow \mathbb{R}$ eine Abbildung. Sei $x \in X$.

- Sei $v \in \mathbb{R}^n \setminus \{0\}$. Dann ist f *in x in Richtung v (partiell) differenzierbar*, wenn es ein $\varepsilon \in \mathbb{R}_{>0}$ gibt mit $U(x, \varepsilon/\|v\|_2) \subset X$, so dass

$$\begin{aligned} (-\varepsilon, \varepsilon) &\longrightarrow \mathbb{R} \\ t &\longmapsto f(x + t \cdot v) \end{aligned}$$

  in 0 differenzierbar ist. Die Ableitung dieser Funktion in 0 wir dann mit $\partial_v f(x) \in \mathbb{R}$ bezeichnet und heißt *Richtungsableitung von f in x in Richtung v*.
- Sei $j \in \{1, \dots, n\}$. Dann ist f *in x in Richtung des j-ten Einheitsvektors partiell differenzierbar*, wenn die Richtungsableitung $\partial_{e_j} f(x)$ existiert. Wir schreiben dann kurz

$$\partial_j f(x) := \partial_{e_j} f(x) \in \mathbb{R}.$$

- Ist f in x in Richtung aller n Einheitsvektoren partiell differenzierbar, so ist f *in x partiell differenzierbar.*

Eine Abbildung $g\colon X \longrightarrow \mathbb{R}^m$ ist *in x partiell differenzierbar*, wenn alle m Koordinatenfunktionen $g_1, \dots, g_m\colon X \longrightarrow \mathbb{R}$ von g in x partiell differenzierbar sind.

Caveat 3.15 (Notation partieller Ableitungen)**.** Oft wird für die partiellen Ableitungen $\partial_j f(x)$ auch die Notation $\frac{\partial}{\partial x_j} f(x)$ verwendet; letztere Notation beruht jedoch auf einer impliziten Namensgebung der Variablen(!) und kann daher zu Missverständnissen führen; zum Beispiel ist nicht klar, was

$$\frac{\partial}{\partial x_1} f(x_2, x_1^2)$$

bedeutet ...

Caveat 3.16 (partielle vs. totale Differenzierbarkeit)**.** Partiell differenzierbare Funktionen sind im allgemeinen *nicht* total differenzierbar. Im allgemeinen genügt auch die Existenz aller Richtungsableitungen *nicht*!

Umgekehrt ist jedoch jede differenzierbare Funktion partiell differenzierbar und man kann die Ableitung mit Hilfe der partiellen Ableitungen beschreiben:

Definition 3.17 (Jacobi-Matrix)**.** Sei $X \subset \mathbb{R}^n$ offen, sei $x \in X$ und sei $f\colon X \longrightarrow \mathbb{R}^m$ in x partiell differenzierbar. Dann ist

$$Jf(x) := \begin{pmatrix} \partial_1 f_1(x) & \dots & \partial_n f_1(x) \\ \vdots & \ddots & \vdots \\ \partial_1 f_m(x) & \dots & \partial_n f_m(x) \end{pmatrix} \in \mathbb{R}^{m\times n}$$

die *Jacobi-Matrix von f in x*; hierbei bezeichnen $f_1,\dots,f_m\colon X \longrightarrow \mathbb{R}$ die Komponenten von f und $\mathbb{R}^{m\times n}$ ist die Menge der reellen $m \times n$-Matrizen.

Satz 3.18 (Ableitung und Jacobi-Matrix)**.** *Sei $X \subset \mathbb{R}^n$ offen, sei $x \in X$ und sei $f\colon X \longrightarrow \mathbb{R}^m$ in x differenzierbar. Dann ist f in x partiell differenzierbar und $f'(x)$ ist die lineare Abbildung, die bezüglich der Standardbasis von $\mathbb{R}^n$ bzw. $\mathbb{R}^m$ durch die Jacobi-Matrix $Jf(x)$ dargestellt wird, d.h.*

$$f'(x) = \begin{pmatrix} \mathbb{R}^n & \longrightarrow & \mathbb{R}^m \\ h & \longmapsto & Jf(x)\cdot h \end{pmatrix} \in L(\mathbb{R}^n, \mathbb{R}^m).$$

Wir geben nun noch ein praktikables hinreichendes Kriterium für mehrdimensionale Differenzierbarkeit:

Satz 3.19 (Hinreichendes Kriterium für Differenzierbarkeit)**.** *Sei $X \subset \mathbb{R}^n$ offen, sei $f\colon X \longrightarrow \mathbb{R}$ eine auf X partiell differenzierbare Abbildung und die partiellen Ableitungen $\partial_1 f,\dots,\partial_n f\colon X \longrightarrow \mathbb{R}$ seien stetig. Dann ist f auf X differenzierbar (und die Ableitung wird durch die Jacobi-Matrix beschrieben).*

Korollar 3.20. *Sei $X \subset \mathbb{R}^n$ offen, sei $f\colon X \longrightarrow \mathbb{R}^m$ partiell differenzierbar und die partiellen Ableitungen $\partial_1 f_1,\dots,\partial_n f_m\colon X \longrightarrow \mathbb{R}$ von f seien alle stetig. Dann ist f auf X differenzierbar (und die Ableitung wird durch die Jacobi-Matrix beschrieben).*

Insbesondere sind alle „polynomialen“ Abbildungen $\mathbb{R}^n \longrightarrow \mathbb{R}^m$ differenzierbar und die *Polarkoordinatenabbildung* ist differenzierbar.

Der Mittelwertsatz

Wir behandeln nun eine höherdimensionale Variante des Mittelwertsatzes; analog zum eindimensionalen Fall können wir damit den Kern des Differenzierens bestimmen.

Satz 3.21 (Mittelwertsatz)**.** *Sei $X \subset \mathbb{R}^n$ offen und sei $f\colon X \longrightarrow \mathbb{R}$ eine differenzierbare Funktion. Seien $a, b \in X$ und es gebe einen stetigen Weg $\gamma\colon [0,1] \longrightarrow X \subset \mathbb{R}^n$ mit $\gamma(0) = a$ und $\gamma(1) = b$, der auf $(0,1)$ differenzierbar ist. Dann existiert ein $\xi \in (0,1)$ mit*

$$f(b) = f(a) + f'\bigl(\gamma(\xi)\bigr) \circ \bigl(\gamma'(\xi)\bigr)(1).$$

Korollar 3.22. *Sei $X \subset \mathbb{R}^n$ offen und sei $f\colon X \longrightarrow \mathbb{R}$ differenzierbar. Seien $a, b \in X$ und es gelte*

$$\bigl\{t\cdot b + (1-t)\cdot a \bigm| t \in [0,1]\bigr\} \subset X.$$

Dann existiert ein $\xi \in X$ mit

$$f(b) = f(a) + \bigl(f'(\xi)\bigr)(b-a).$$

Caveat 3.23. Es gibt *keinen* analogen Mittelwertsatz für differenzierbare Funktionen vom Typ $\mathbb{R}^n \longrightarrow \mathbb{R}^m$ mit $m > 1$.

Korollar 3.24 (Kern des Differenzierens, I)**.** *Sei $X \subset \mathbb{R}^n$ offen und zusammnenhängend (bezüglich der Teilraumtopologie), sei $f\colon X \longrightarrow \mathbb{R}$ differenzierbar und es gelte $f'(x) = 0$ für alle $x \in X$. Dann ist f konstant.*

Korollar 3.25 (Kern des Differenzierens, II)**.** *Sei $X \subset \mathbb{R}^n$ offen und zusammenhängend, sei $f\colon X \longrightarrow \mathbb{R}^m$ differenzierbar und es gelte $f'(x) = 0$ für alle $x \in X$. Dann ist f konstant.*

Höhere Ableitungen

Bisher haben wir nur Ableitungen in einem Punkt betrachtet. Um höhere Ableitungen definieren zu können, müssen wir nun die Ableitungen an den einzelnen Punkten wieder geeignet zu einer Abbildung zusammenfassen:

Konvention 3.26. Ist $X \subset \mathbb{R}^n$ offen und $f\colon X \longrightarrow \mathbb{R}^m$ differenzierbar, so erhalten wir durch Ableiten die Abbildung

$$\begin{aligned} f'\colon X &\longrightarrow L(\mathbb{R}^n, \mathbb{R}^m) \\ x &\longmapsto f'(x). \end{aligned}$$

Da wir $L(\mathbb{R}^n, \mathbb{R}^m)$ als $\mathbb{R}$-Vektorraum kanonisch mit $\mathbb{R}^{m\times n} \cong \mathbb{R}^{m\cdot n}$ identifizieren können, können wir f' als Abbildung

$$f'\colon X \longrightarrow R^{m\cdot n}$$

auffassen, wofür wir bereits einen Differenzierbarkeitsbegriff haben.

Definition 3.27 (Funktionenräume C^k)**.** Seien $m, n \in \mathbb{N}$ und sei $X \subset \mathbb{R}^n$ offen.

– Dann schreiben wir

$$C(X; \mathbb{R}^m) := \{f\colon X \longrightarrow \mathbb{R}^m \mid f \text{ ist stetig}\}.$$

– Induktiv definieren wir $C^0(X; \mathbb{R}^m) := C(X; \mathbb{R}^m)$ und

$$\begin{aligned} C^{k+1}(X; \mathbb{R}^m) := \bigl\{\, & f\colon X \longrightarrow \mathbb{R}^m \\ & \bigm| f \text{ ist auf } X \text{ differenzierbar und } f' \in C^k(X, L(\mathbb{R}^n, \mathbb{R}^m)) \bigr\} \end{aligned}$$

  für alle $k \in \mathbb{N}$.

– Außerdem setzen wir

$$C^\infty(X; \mathbb{R}^m) := \bigcap_{k\in\mathbb{N}} C^k(X; \mathbb{R}^m).$$

Sei $k \in \mathbb{N} \cup \{\infty\}$; die Elemente von $C^k(X; \mathbb{R}^m)$ heißen dann *k-fach stetig differenzierbare Abbildungen $X \longrightarrow \mathbb{R}^m$.*

Aus dem hinreichenden Kriterium für Differenzierbarkeit (Satz 3.19) folgt induktiv:

Proposition 3.28 (stetige Differenzierbarkeit vs. stetige partielle Differenzierbarkeit)**.** *Sei $X \subset \mathbb{R}^n$ offen und $k \in \mathbb{N}$. Dann ist*

$$C^k(X;\mathbb{R}^m) := \bigl\{f\colon X \longrightarrow \mathbb{R}^m \bigm| f \text{ ist } k\text{-fach stetig partiell differenzierbar auf } X\bigr\}.$$

Mit den Vererbungseigenschaften erhalten wir, dass Ableiten im folgenden linear ist:

Proposition 3.29 (Linearität des Ableitens)**.** *Sei $X \subset \mathbb{R}^n$ offen und sei $k \in \mathbb{N}$. Dann sind $C^{k+1}(X,\mathbb{R}^m)$ und $C^k(X;L(\mathbb{R}^n,\mathbb{R}^m))$ reelle Vektorräume (bezüglich punktweiser Addition und Skalarmultiplikation) und die Abbildung*

$$\begin{aligned} C^{k+1}(X;\mathbb{R}^m) &\longrightarrow C^k\bigl(X;L(\mathbb{R}^n,\mathbb{R}^m)\bigr)\\ f &\longmapsto f' \end{aligned}$$

ist linear.

Wie im eindimensionalen Fall ist der wichtigste Fall höherer Ableitungen die zweite Ableitung (da sie bei Optimierungsproblemen eine wichtige Rolle spielt):

Proposition 3.30 (Zweite Ableitung als bilineare Abbildung)**.** *Sei $X \subset \mathbb{R}^n$ offen, sei $x \in X$ und sei $f\colon X \longrightarrow \mathbb{R}^m$ auf X differenzierbar und in x zweimal differenzierbar (d.h. f' ist in x differenzierbar). Dann ist die Abbildung*

$$\begin{aligned} \mathbb{R}^n \times \mathbb{R}^n &\longrightarrow \mathbb{R}^m\\ (h,k) &\longmapsto \bigl(f''(x)(k)\bigr)(h) \end{aligned}$$

bilinear.

Der Beweis dieser Proposition beruht auf dem sogenannten Exponentialgesetz für (bi)lineare Abbildungen.

Analog kann man natürlich auch die höheren Ableitungen in einem gegebenen Punkt als multilineare Abbildungen auffassen.

Proposition 3.31 (quadratische Abbildungen)**.** *Sei $b\colon \mathbb{R}^n \times \mathbb{R}^n \longrightarrow \mathbb{R}^m$ bilinear und*

$$\begin{aligned} q\colon \mathbb{R}^n &\longrightarrow \mathbb{R}^m\\ x &\longmapsto b(x,x) \end{aligned}$$

sei die zugehörige quadratische Abbildung. Dann ist $q \in C^\infty(\mathbb{R}^n;\mathbb{R}^m)$ und es gilt:

1. *Für alle $x \in \mathbb{R}^n$ ist*

$$q'(x) = \bigl(h \mapsto b(x,h) + b(h,x)\bigr) \in L(\mathbb{R}^n,\mathbb{R}^m).$$

Insbesondere ist $q'\colon X \longrightarrow L(\mathbb{R}^n,\mathbb{R}^m)$ linear.

2. *Also gilt für alle $x \in \mathbb{R}^n$, dass*

$$q''(x) = \begin{pmatrix} \mathbb{R}^n \times \mathbb{R}^n & \longrightarrow & \mathbb{R}^m \\ (h,k) & \longmapsto & b(k,h) + b(h,k) \end{pmatrix}.$$

3. *Für alle $k \in \mathbb{N}_{\geq 3}$ ist $q^{(k)} = 0$.*

Satz 3.32 (Satz von Schwarz)**.** *Sei $X \subset \mathbb{R}^n$ offen, sei $x \in X$, sei $f\colon X \longrightarrow \mathbb{R}^m$ auf X differenzierbar und in x existiere die zweite Ableitung von f.*

1. *Dann gilt für alle $h, k \in \mathbb{R}^n$, dass*

$$f''(x)(h,k) = f''(x)(k,h).$$

D.h. die bilineare Abbildung $f''(x)$ ist symmetrisch.

2. *Insbesondere gilt für alle $j, k \in \{1,\dots,n\}$ und alle $\ell \in \{1,\dots,m\}$, dass*

$$\partial_j \partial_k f_\ell(x) = \partial_k \partial_j f_\ell(x).$$

Korollar 3.33 (Hesse-Matrix)**.** *Sei $X \subset \mathbb{R}^n$ offen, sei $x \in X$ und sei $f\colon X \longrightarrow \mathbb{R}$ auf X differenzierbar und in x zweimal differenzierbar. Dann ist die* Hesse-Matrix von f in x

$$\mathrm{Hess}_f(x) := \bigl(\partial_j \partial_k f(x)\bigr)_{j,k \in \{1,\dots,n\}} \in \mathbb{R}^{n \times n}$$

symmetrisch und es gilt

$$f''(x) = \begin{pmatrix} \mathbb{R}^n \times \mathbb{R}^n & \longrightarrow & \mathbb{R} \\ (h,k) & \longmapsto & h^\top \cdot \mathrm{Hess}_f(x) \cdot k \end{pmatrix}.$$

Taylorentwicklung und lokale Extrema

Analog zum eindimensionalen Fall können wir auch im höherdimensionalen Fall versuchen, hinreichend oft differenzierbare Funktionen durch ihre höheren Ableitungen polynomial zu approximieren; dies führt zur Taylorentwicklung:

Satz 3.34 (Taylorentwicklung)**.** *Sei $X \subset \mathbb{R}^n$ offen, sei $N \in \mathbb{N}$ und sei $f \in C^{N+1}(X;\mathbb{R})$. Sei $x \in X$ und sei $h \in \mathbb{R}^n$ so klein, dass die Verbindung $\{x + t \cdot h \mid t \in [0,1]\}$ von x nach $x + h$ in X liegt. Dann gilt*

$$f(x+h) = \underbrace{\sum_{j=0}^{N} \frac{1}{j!} \cdot f^{(j)}(x)\,\underbrace{(h,\dots,h)}_{j\text{-mal}}}_{N\text{-tes Taylor-Polynom von } f \text{ um } x} + \underbrace{\frac{1}{N!} \cdot \int_0^1 (1-t)^N \cdot f^{(N+1)}(x + t \cdot h)\,\underbrace{(h,\dots,h)}_{N+1\text{-mal}}\,dt}_{\text{Restglied in Integralform}}.$$

Alternativ kann das Restglied auch durch den Wert von $f^{(N+1)}$ an einer geeigneten Stelle, angewendet auf geeignete Vektoren ausgedrückt werden.

Analog zum eindimensionalen Fall kann man natürlich auch für unendlich oft stetig differenzierbare Funktionen untersuchen, unter welchen Bedingungen die entsprechende *Taylorreihe* konvergiert und ob sie die gegebene Funktion darstellt. Darauf gehen wir nun nicht im Detail ein.

Wenn man mit höheren Taylor-Polynomen rechnen möchte, bietet es sich an, die folgende Notation mit Hilfe von Multi-Indizes zu verwenden:

Notation 3.35 (Multiindizes)**.** Sei $n \in \mathbb{N}$ und sei $\alpha \in \mathbb{N}^n$; man bezeichnet α dann auch als *Multiindex*.

– Wir schreiben kurz

$$\alpha! := \alpha_1! \cdot \dots \cdot \alpha_n!,$$
$$|\alpha| := \alpha_1 + \dots + \alpha_n.$$

– Ist $h \in \mathbb{R}^n$, so schreiben wir

$$h^\alpha := h_1^{\alpha_1} \cdot \dots \cdot h_n^{\alpha_n}.$$

– Ist $f \in C^{|\alpha|}(X;\mathbb{R})$ und $x \in X$ (wobei $X \subset \mathbb{R}^n$ offen ist), so schreiben wir

$$\partial_\alpha f(x) := \partial_1^{\alpha_1} \dots \partial_n^{\alpha_n} f(x).$$

Ist $h \in \mathbb{R}^n$, so steht $\partial_\alpha f(x) \cdot h^\alpha$ also für den Ausdruck $\partial_1^{\alpha_1} \dots \partial_n^{\alpha_n} f(x) \cdot h_1^{\alpha_1} \cdot \dots \cdot h_n^{\alpha_n}$.

Indem man verwendet, dass man die Ableitung (und die höheren Ableitungen) an einem Punkt auch durch entsprechende partielle Ableitungen darstellen kann, den Satz von Schwarz über die Vertauschbarkeit partieller Ableitungen anwendet und Summanden geeignet zusammenfasst, kann man obige Taylorentwicklung auch wie folgt darstellen:

Bemerkung 3.36 (Taylorentwicklung via Multiindizes)**.** Sei $X \subset \mathbb{R}^n$ offen, sei $N \in \mathbb{N}$ und sei $f \in C^{N+1}(X;\mathbb{R})$. Sei $x \in X$ und sei $h \in \mathbb{R}^n$ so klein, dass $\{x + t \cdot h \mid t \in [0,1]\}$ in X liegt. Dann gilt

$$f(x+h) = \sum_{\alpha \in \mathbb{N}^n, |\alpha| \leq N} \frac{1}{\alpha!} \cdot \partial_\alpha f(x) \cdot h^\alpha + \frac{1}{N!} \cdot \int_0^1 (1-t)^N \cdot f^{(N+1)}(x+t \cdot h) \underbrace{(h, \dots, h)}_{N+1\text{-mal}} \, dt.$$

Als Anwendung der Taylorentwicklung zeigen wir wie man mit Hilfe der Differentialrechnung eine Strategie zum Lösen von Optimierungsproblemen erhält:

Satz 3.37 (Stationarität lokaler Extrema)**.** *Sei $X \subset \mathbb{R}^n$ offen, sei $x \in X$ und sei $f \colon X \longrightarrow \mathbb{R}$ in x differenzierbar. Sei x eine* lokale Extremalstelle von f*, d.h. es gibt eine offene Umgebung $U \subset X$ von x mit*

$$\forall_{y \in U} \quad f(y) \leq f(x) \qquad \textit{oder} \qquad \forall_{y \in U} \quad f(y) \geq f(x);$$

in ersterem Fall nennt man x eine lokale Maximalstelle*, in letzterem Fall eine* lokale Minimalstelle *von f. Dann ist x eine* kritische Stelle von f*, d.h. $f'(x) = 0 \in L(\mathbb{R}^n, \mathbb{R})$.*

Im allgemeinen ist dieses notwendige Kriterium über die erste Ableitung *nicht* hinreichend (wie bereits schon im eindimensionalen Fall). Analog zum eindimensionalen Fall liefert die „Positivität/Negativität“ der zweiten Ableitung ein hinreichendes Kriterium:

Satz 3.38 (hinreichendes Kriterium für lokale Extrema)**.** *Sei $X \subset \mathbb{R}^n$ offen und sei $f \in C^2(X;\mathbb{R})$. Sei $x \in X$ mit $f'(x) = 0$. Dann gilt:*

1. *Ist $f''(x)$ bzw.* $\mathrm{Hess}_f(x)$ *positiv definit [negativ definit], so ist x eine lokale Minimalstelle [Maximalstelle] von f.*
2. *Ist $f''(x)$ bzw.* $\mathrm{Hess}_f(x)$ *indefinit, so ist x* keine *lokale Extremalstelle von f.*

Bemerkung 3.39 (positiv/negativ definit, indefinit)**.** Sei $b\colon \mathbb{R}^n \times \mathbb{R}^n \longrightarrow \mathbb{R}$ eine symmetrische Bilinearform. Dann ist b

- *positiv definit* [*negativ definit*], falls ein $a \in \mathbb{R}_{>0}$ existiert mit

$$\forall_{h\in\mathbb{R}^n} \quad b(h,h) \geq a \cdot \|h\|_2^2 \qquad \Big[\text{bzw.} \quad \forall_{h\in\mathbb{R}^n} \quad b(h,h) \leq -a \cdot \|h\|_2^2.\Big]$$

- *indefinit*, falls es $h_+, h_- \in \mathbb{R}^n$ gibt mit

$$b(h_+,h_+) > 0 \qquad \text{und} \qquad b(h_-,h_-) < 0.$$

Sei $A \in \mathbb{R}^{n\times n}$ die (symmetrische) Matrix, die b bezüglich der Standardbasis darstellt, d.h. für alle $j,k \in \{1,\dots,n\}$ ist $A_{jk} = b(e_j,e_k)$ bzw. für alle $h,k \in \mathbb{R}^n$ ist

$$b(h,k) = h^\top \cdot A \cdot k.$$

Dann ist A als reelle symmetrische Matrix über $\mathbb{R}$ diagonalisierbar und alle Eigenwerte von A sind reell. Dabei gilt: Die Bilinearform b ist genau dann

- positiv definit [negativ definit], wenn alle Eigenwerte von A positiv [negativ] sind.
- indefinit, wenn A sowohl einen positiven als auch einen negativen Eigenwert besitzt.

Man kann nun versuchen, Optimierungsprobleme zu lösen, indem man zunächst versucht, alle kritischen Stellen der zu optimierenden Funktion zu bestimmen und dann mit Hilfe des obigen hinreichenden Kriteriums festzustellen, bei welchen dieser kritischen Stellen es sich tatsächlich um lokale Extremalstellen handelt.

Man beachte dabei, dass sich auch Ungleichungen als Optimierungsprobleme formulieren lassen und diese Methode somit einen Ansatz liefert, um gewisse Ungleichungen zu beweisen.

Beispiele für Differentialoperatoren

Wir werden im folgenden kurz ein paar wichtige Beispiele für Differentialoperatoren vorstellen und ihre geometrische Bedeutung bzw. ihre Rolle in den Anwendungen skizzieren.

„Definition“ 3.40 (Differentialoperator)**.** Ein *(linearer) Differentialoperator* ist eine (lineare) Abbildung zwischen geeigenten Funktionenräumen, die durch Komposition/Linearkombination von gewissen partiellen Ableitungen gegeben ist.

Gradient

Der Gradient liefert (bis auf ein Vorzeichen) zu einem Potential (z.B. Graviationspotential, elektrostatisches Potential, …) das zugehörige Vektorfeld (Gravitationsfeld, elektrostatisches Feld, …).

Definition 3.41 (Gradient)**.** Sei $X \subset \mathbb{R}^n$ offen und sei $f\colon X \longrightarrow \mathbb{R}$ auf X differenzierbar. Dann heißt

$$\begin{aligned} \nabla f\colon X &\longrightarrow \mathbb{R}^n \\ x &\longmapsto \begin{pmatrix} \partial_1 f(x) \\ \vdots \\ \partial_n f(x) \end{pmatrix} \end{aligned}$$

Gradient von f; das Symbol „∇" heißt *Nabla*.

Man beachte: Ist $x \in X$, so ist $\nabla f(x) = \bigl(Jf(x)\bigr)^\top \in \mathbb{R}^n$ ein *Spalten*vektor.

Mit Hilfe der Cauchy-Schwarzschen Ungleichung (Lemma 1.24) können wir eine erste geometrische Bedeutung des Gradienten herleiten:

Proposition 3.42 (geometrische Bedeutung des Gradienten)**.** *Sei $X \subset \mathbb{R}^n$ offen, sei $f\colon X \longrightarrow \mathbb{R}$ auf X differenzierbar, sei $x \in X$ und es gelte $\nabla f(x) \neq 0$. Dann ist $\nabla f(x)$ die Richtung der größten lokalen Veränderung von f in x, d.h. für alle $v \in \mathbb{R}^n \setminus \{0\}$ gilt*

$$\frac{|\partial_v f(x)|}{\|v\|_2} \leq \frac{|\partial_{\nabla f(x)}(x)|}{\|\nabla f(x)\|_2},$$

und Gleichheit gilt genau dann, wenn v und $\nabla f(x)$ linear abhängig ist.

Auf dieser Proposition beruht ein numerisches Optimierungsverfahren, das dem „Gradientenfluss" folgt, um lokale Minimalstellen zu finden.

Intuitiv sollte die größte Veränderungsrichtung (also der Gradient) senkrecht auf den Höhenlinien (bzw. allgemeiner den Niveauhyperflächen) stehen; dies ist tatsächlich der Fall:

Proposition 3.43 (der Gradient ist orthogonal zu den Niveaulinien)**.** *Sei $X \subset \mathbb{R}^n$ offen, sei $f\colon X \longrightarrow \mathbb{R}$ auf X differenzierbar, sei $c \in \mathbb{R}$ und sei $\gamma\colon (0,1) \longrightarrow f^{-1}(c) \subset X \subset \mathbb{R}^n$ differenzierbar. Dann gilt für alle $t \in (0,1)$, dass*

$$\bigl\langle \nabla f(\gamma(t)), \gamma'(t) \bigr\rangle = 0.$$

Caveat 3.44. *Nicht* jedes Vektorfeld entsteht als Gradient einer geeigneten Potentialfunktion!

Divergenz

Die Divergenz eines Vektorfeldes beschreibt die „Quellenstärke":

Definition 3.45 (Divergenz eines Vektorfeldes)**.** Sei $X \subset \mathbb{R}^n$ offen, sei $f\colon X \longrightarrow \mathbb{R}^n$ auf X differenzierbar. Dann ist

$$\operatorname{div} f := \sum_{j=1}^n \partial_j f_j\colon X \longrightarrow \mathbb{R}$$

die *Divergenz von f*.

Rotation

Die Rotation eines dreidimensionalen Vektorfeldes beschreibt, wie sehr „Teilchen“ allein durch das Feld in Rotation versetzt werden und um welche Rotationsachse:

Definition 3.46 (Rotation)**.** Sei $X \subset \mathbb{R}^3$ offen und sei $f\colon X \longrightarrow \mathbb{R}^3$ auf X differenzierbar. Dann ist

$$\begin{aligned} \operatorname{rot} f\colon X &\longrightarrow \mathbb{R}^3 \\ x &\longmapsto \begin{pmatrix} \partial_2 f_3(x) - \partial_3 f_2(x) \\ \partial_3 f_1(x) - \partial_1 f_3(x) \\ \partial_1 f_2(x) - \partial_2 f_1(x) \end{pmatrix} \end{aligned}$$

die *Rotation von f*.

Mit Hilfe des Satzes von Schwarz (Satz 3.32) folgt:

Proposition 3.47 (Verträglichkeit von Rotation mit anderen Differentialoperatoren)**.** *Sei $X \subset \mathbb{R}^3$ offen.*

1. *Ist $f\colon X \longrightarrow \mathbb{R}^3$ zweimal differenzierbar, so ist*

$$\operatorname{div} \operatorname{rot} f = 0.$$

2. *Ist $f\colon X \longrightarrow \mathbb{R}$ zweimal differenzierbar, so ist*

$$\operatorname{rot} \nabla f = 0.$$

Der Laplace-Operator

Der Laplace-Operator eines Potentials misst die Änderungsrate des Mittelwerts des Potentials auf Sphären um den betrachteten Punkt bei Änderung des Radius. Daher tritt der Laplace-Operator oft bei der Beschreibung physikalischer Systeme auf; z.B. liefert der Laplace-Operator eines Gravitationspotentials (bzw. eines elektrostatischen Potentials) die zugehörige Massenverteilung (bzw. die zugehörige Ladungsverteilung). Außerdem tritt er auch in der Wärmeleitungsgleichung auf, und in der Differentialgeometrie.

Der Laplace-Operator ist ein linearer Differentialoperator zweiter Ordnung:

Definition 3.48 (Laplace-Operator)**.** Sei $X \subset \mathbb{R}^n$ offen und sei $f\colon X \longrightarrow \mathbb{R}$ auf X zweimal differenzierbar. Dann ist

$$\Delta f := \operatorname{div} \nabla f = \sum_{j=1}^{n} \partial_j \partial_j f\colon X \longrightarrow \mathbb{R}$$

der *Laplace-Operator angewendet auf f*.

Ist $\Delta f = 0$, so heißt f *harmonisch.*

Lokale Umkehrbarkeit

Wir wollen uns im folgenden damit beschäftigen, unter welchen Voraussetzungen wir differenzierbare Funktionen lokal invertieren können. Dabei haben wir folgende Ziele:

1. Aus der Analysis I wissen wir: Ist $f \in C^1(\mathbb{R}, \mathbb{R})$ und ist $x \in \mathbb{R}$ mit $f'(x) \neq 0$, so ist f lokal um x invertierbar und das Inverse ist auch stetig differenzierbar. Gibt es eine Verallgemeinerung auf den mehrdimensionalen Fall?
2. Wir möchten die Geometrie von Niveaulinien von Funktionen vom Typ $\mathbb{R}^n \longrightarrow \mathbb{R}$ besser verstehen (und allgemeiner von Untermannigfaltigkeiten) bzw. Gleichungen dieser Art lokal „auflösen“ können oder zumindest die lokale Lösbarkeit untersuchen.

Der zentrale Satz in diesem Kontext ist der Satz über lokale Umkehrbarkeit:

Satz 3.49 (lokale Umkehrbarkeit)**.** *Sei $X \subset \mathbb{R}^n$ offen, sei $f \in C^1(X, \mathbb{R}^n)$, sei $x \in X$ und sei $f'(x) \in L(\mathbb{R}^n, \mathbb{R}^n)$ invertierbar (d.h. $\det Jf(x) \neq 0$). Dann gibt es eine offene Umgebung $U \subset X$ von x mit folgenden Eigenschaften:*

1. *Die Einschränkung $f|_U : U \longrightarrow \mathbb{R}^n$ ist injektiv.*
2. *Die Menge $f(U) \subset \mathbb{R}^n$ ist offen.*
3. *Die Umkehrabbildung $(f|_U)^{-1} : f(U) \longrightarrow U$ ist stetig differenzierbar.*

D.h. $f|_U : U \longrightarrow f(U)$ ist ein C^1-Diffeomorphismus.

Die grundlegende Idee des Beweises ist es, Invertierbarkeit als ein geeignetes Fixpunktproblem aufzufassen und dann den Banachschen Fixpunktsatz (Satz 1.19) anzuwenden.

Zum Beispiel kann man mit Hilfe des Satzes über lokale Umkehrbarkeit leicht nachweisen, dass die Polarkoordinatenabbildung ein lokaler C^1-Diffeomorphismus ist.

Als Anwendung des Satzes über lokale Umkehrbarkeit geben wir einen ersten Schritt zur Lösung des zweiten Problems: Die Frage nach der Existenz lokaler C^1-Parametrisierungen von Höhenlinien (und ihren Verallgemeinerungen) wird durch den Satz über implizite Funktionen beantwortet:

Um den Satz formulieren zu können, benötigen wir noch eine geeignete Notation für gewisse Teile der Jacobi-Matrix:

Notation 3.50. Seien $n, m \in \mathbb{N}$ und sei $X \subset \mathbb{R}^n \times \mathbb{R}^m$ offen. Ist $x \in X$, so schreiben wir kurz

$$x_{①} := (x_1, \dots, x_n)^\top \in \mathbb{R}^n \qquad \text{und} \qquad x_{②} := (x_{n+1}, \dots, x_{n+m})^\top;$$

dies hängt natürlich von der Zerlegung von $n + m$ in n und m ab – um die Notation nicht zu überfrachten, fügen wir dies aber nicht auch noch zur Notation hinzu.

Ist $f : X \longrightarrow \mathbb{R}^n$ differenzierbar, so schreiben wir für alle $x \in X$ außerdem

$$J_{①} f(x) := \begin{pmatrix} \partial_1 f_1(x) & \dots & \partial_n f_1(x) \\ \vdots & \ddots & \vdots \\ \partial_1 f_n(x) & \dots & \partial_n f_n(x) \end{pmatrix} \in \mathbb{R}^{n \times n},$$

sowie

$$J_{②}f(x) := \begin{pmatrix} \partial_{n+1}f_1(x) & \dots & \partial_{n+m}f_1(x) \\ \vdots & \ddots & \vdots \\ \partial_{n+1}f_n(x) & \dots & \partial_{n+m}f_n(x) \end{pmatrix} \in \mathbb{R}^{n\times m}.$$

Satz 3.51 (Satz über implizite Funktionen)**.** *Sei $X \subset \mathbb{R}^n \times \mathbb{R}^m$ offen und sei $f \in C^1(X;\mathbb{R}^n)$. Sei $x \in X$ mit $f(x) = 0$ und* $\det J_{①}f(x) \neq 0$. *Dann existiert eine offene Umgebung $U \subset \mathbb{R}^m$ von $x_{②}$, so dass es eine lokal um $x_{②}$ eindeutig bestimmte Abbildung $\varphi \in C^1(U;\mathbb{R}^n)$ mit $\varphi(x_{②}) = x_{①}$ und $\{(\varphi(t),t) \mid t \in U\} \subset X$ sowie*

$$\forall_{t\in U} \quad f\bigl(\varphi(t),t\bigr) = 0$$

gibt.

Dieser Satz spielt in der Differentialtopologie eine wichtige Rolle.

Extrema unter Nebenbedingungen

Häufig stößt man auf Optimierungsprobleme der folgenden Form: Gesucht sind die Extrema einer Einschränkung $f|_M$, wobei $f\colon \mathbb{R}^n \longrightarrow \mathbb{R}$ und $M \subset \mathbb{R}^n$ gegeben sind. Oft ist dabei M in der Form $M = g^{-1}(\{0\})$ für eine geeignete Abbildung $g\colon \mathbb{R}^n \longrightarrow \mathbb{R}^m$ gegeben.

Caveat 3.52. Im allgemeinen sind in dieser Situation lokale Extrema der Einschränkung $f|_M\colon M \longrightarrow \mathbb{R}$ *keine* lokalen Extrema der Funktion $f\colon \mathbb{R}^n \longrightarrow \mathbb{R}$.

Eine naive Lösungsstrategie ist, die Menge M explizit zu parametrisieren (d.h. die Gleichung „$g(x) = 0$" aufzulösen) und dann in f einzusetzen. Im allgemeinen ist dies aber natürlich nicht explizit möglich. Mithilfe des Satzes über implizite Funktionen (Satz 3.51) kann man aus dieser Strategie aber ein praktikables notwendiges Kriterium ableiten:

Satz 3.53 (Extrema unter Nebenbedingungen und Lagrange-Multiplikatoren)**.** *Sei $X \subset \mathbb{R}^n$ offen, sei $f \in C^1(X;\mathbb{R})$, sei $g \in C^1(X;\mathbb{R}^m)$ und es gelte für alle $x \in X$, dass* $\operatorname{rg} Jg(x) = m$ *ist (Regularitätsbedingung). Außerdem sei*

$$M := g^{-1}(\{0\}) = \bigl\{x \in X \bigm| \forall_{j\in\{1,\dots,m\}} \quad g_j(x) = 0\bigr\}$$

und sei $x \in M$ eine lokale Extremalstelle von $f|_M$ (d.h. es existiert eine offene Umgebung $U \subset M$ von x in M bezüglich der Teilraumtopologie, so dass x eine Extremalstelle von $f|_U\colon U \longrightarrow \mathbb{R}$ ist). Dann gibt es ein $\lambda \in \mathbb{R}^m$ mit

$$\nabla f(x) + \sum_{j=1}^{m} \lambda_j \cdot \nabla g_j(x) = 0.$$

Man bezeichnet λ dann als Lagrange-Multiplikator.

Kurven und Untermannigfaltigkeiten

Im folgenden geben wir einen kurzen Ausblick wie man mithilfe der mehrdimensionalen Differentialrechnung geometrische Objekte wie Kurven, Flächen etc. untersuchen und beschreiben kann.

Wir beginnen mit (parametrisierten) Kurven und werden dann für höherdimensionale Objekte zu einer lokaleren Perspektive wechseln.

Definition 3.54 ((reguläre) parametrisierte Kurve)**.** Sei $n \in \mathbb{N}$.

- Eine *parametrisierte Kurve in* $\mathbb{R}^n$ ist eine C^1-Abbildung $I \longrightarrow \mathbb{R}^n$, wobei $I \subset \mathbb{R}$ ein Intervall ist.
  Als *Intervall* sind hierbei offene, halboffene und abgeschlossene Intervalle in $\mathbb{R}$ zugelassen, wobei wir bei offenen Grenzen auch „$\pm\infty$" erlauben.
  Eine Abbildung $\gamma \colon I \longrightarrow \mathbb{R}^n$ ist eine *C^1-Abbildung*, wenn γ auf I stetig ist, $\gamma|_{I^\circ}$ auf I° differenzierbar ist und die Ableitung $\gamma'|_{I^\circ}$ stetig auf I fortgesetzt werden kann.
- Eine parametrisierte Kurve $\gamma \colon I \longrightarrow \mathbb{R}^n$ heißt *regulär*, wenn für alle $t \in I^\circ$ gilt, dass $\gamma'(t) \neq 0$ ist.

Analog kann man natürlich auch Kurven betrachten, die öfter differenzierbar bzw. sogar glatt sind.

Definition 3.55 (Äquivalenz parametrisierter Kurven)**.** Sei $n \in \mathbb{N}$ und seien $I_1, I_2 \subset \mathbb{R}$ Intervalle. Zwei parametrisierte Kurven $\gamma_1 \colon I_1 \longrightarrow \mathbb{R}^n$ und $\gamma_2 \colon I_2 \longrightarrow \mathbb{R}^n$ in $\mathbb{R}^n$ heißen *äquivalent*, wenn es einen C^1-Diffeomorphismus $\varphi \colon I_2 \longrightarrow I_1$ mit $\gamma_2 = \gamma_1 \circ \varphi$ gibt.

Ein *C^1-Diffeomorphismus* ist dabei eine C^1-Abbildung $\varphi \colon I_2 \longrightarrow I_1$, für die es eine C^1-Abbildung $\psi \colon I_1 \longrightarrow I_2$ mit $\varphi \circ \psi = \mathrm{id}_{I_1}$ und $\psi \circ \varphi = \mathrm{id}_{I_2}$ gibt.

Bemerkung 3.56. Dieser Begriff von „Äquivalenz" für [reguläre] parametrisierte Kurven in $\mathbb{R}^n$ ist tatsächlich eine Äquivalenzrelation auf der Menge der [regulären] parametrisierten Kurven in $\mathbb{R}^n$.

Indem wir nun jeweils die parametrisierten Kurven einer Äquivalenzklasse identifizieren, erhalten wir den Begriff der Kurve:

Definition 3.57 (Kurve)**.** Sei $n \in \mathbb{N}$. Eine *Kurve in* $\mathbb{R}^n$ ist eine Äquivalenzklasse regulärer parametrisierter Kurven in $\mathbb{R}^n$.

Wir möchten nun geeignete Invarianten für Kurven definieren. Wir beginnen mit einer globalen Invariante, der Länge:

Notation 3.58. Ist $\gamma \colon I \longrightarrow \mathbb{R}^n$ eine parametrisierte Kurve und ist $t \in I^\circ$, so schreiben wir kurz (und etwas schlampig)

$$\gamma'(t) := \begin{pmatrix} \gamma_1'(t) \\ \vdots \\ \gamma_n'(t) \end{pmatrix} \in \mathbb{R}^n.$$

Definition 3.59 (Länge parametrisierter Kurven)**.** Seien $a, b \in \mathbb{R}$ mit $a \leq b$ und sei $\gamma\colon [a, b] \longrightarrow \mathbb{R}^n$ eine parametrisierte Kurve. Dann heißt

$$\ell(\gamma) := \int_a^b \left\|\gamma'(t)\right\|_2 \, dt$$

Länge von γ.

Mit Hilfe dieser Länge von Kurven lässt sich insbesondere der Winkelbegriff auf geeignete Bogenlängen zurückführen.

Satz 3.60 (Länge parametrisierter Kurven via Polygonzüge)**.** *Seien $a, b \in \mathbb{R}$ mit $a \leq b$ und sei $\gamma\colon [a, b] \longrightarrow \mathbb{R}^n$ eine parametrisierte Kurve in $\mathbb{R}^n$. Dann ist*

$$\ell(\gamma) = \sup \bigl\{\ell_P(\gamma) \bigm| P \textit{ ist eine Partition von } [a, b]\bigr\},$$

wobei wir die Abkürzung

$$\ell_P(\gamma) := \sum_{j=1}^{k} \left\|\gamma(t_j) - \gamma(t_{j-1})\right\|_2$$

verwenden, wenn $P = (t_0, \dots, t_k)$ eine Partition von $[a, b]$ ist (dies ist die Länge des Polygonzugs auf γ zur Partition P).

Die Länge einer Kurve hängt nicht von der Geschwindigkeit ab, mit der sie durchlaufen wird, d.h. äquivalente parametrisierte Kurven haben dieselbe Länge:

Proposition 3.61 (Länge von Kurven)**.** *Sei $n \in \mathbb{N}$ und seien $\gamma_1\colon [a_1, b_1] \longrightarrow \mathbb{R}^n$ und $\gamma_2\colon [a_2, b_2] \longrightarrow \mathbb{R}^n$ reguläre parametrisierte Kurven in $\mathbb{R}^n$, die dieselbe Kurve c in $\mathbb{R}^n$ repräsentieren. Dann ist*

$$\ell(\gamma_1) = \ell(\gamma_2).$$

Man nennt in diesem Fall $\ell(c) := \ell(\gamma_1) = \ell(\gamma_2)$ die Länge von c.

Ausgezeichnete Parametrisierungen von Kurven sind die, die mit der Länge von Teilkurven kompatibel sind:

Definition 3.62 (nach Bogenlänge parametrisiert)**.** Sei $I \subset \mathbb{R}$ ein Intervall. Eine parametrisierte Kurve $\gamma\colon I \longrightarrow \mathbb{R}^n$ in $\mathbb{R}^n$ ist *nach Bogenlänge parametrisiert*, wenn $\|\gamma'(t)\|_2 = 1$ für alle $t \in I^\circ$ gilt.

Mit Hilfe des Hauptsatzes der Differential- und Integralrechnung kann man zeigen, dass solche Parametrisierungen immer existieren und im wesentlichen eindeutig sind:

Proposition 3.63 (Existenz und Eindeutigkeit von Parametrisierungen nach Bogenlänge)**.** *Sei $I \subset \mathbb{R}$ ein Intervall und sei $\gamma\colon I \longrightarrow \mathbb{R}^n$ eine reguläre parametrisierte Kurve in $\mathbb{R}^n$.*

1. *Dann gibt es eine nach Bogenlänge parametrisierte Kurve in $\mathbb{R}^n$, die zu γ äquivalent ist.*

2. *Sind $\gamma_1 \colon I_1 \longrightarrow \mathbb{R}^n$ und $\gamma_2 \colon I_2 \longrightarrow \mathbb{R}^n$ nach Bogenlänge parametrisierte Kurven, die zu γ äquivalent sind, so existiert ein $c \in \mathbb{R}$ mit*

$$\forall_{t \in I_2} \quad \gamma_2(t) = \gamma_1(t+c) \qquad \text{oder} \qquad \forall_{t \in I_2} \quad \gamma_2(t) = \gamma_2(-t+c).$$

Die Länge einer Kurve ist eine globale Invariante. Geschwindigkeit und Krümmung sind hingegen lokale Invarianten. Die Krümmung einer Kurve in einem Punkt ist die Änderung der Geschwindigkeit(srichtung):

Definition 3.64 (Krümmung von parametrisierten Kurven)**.** Sei $I \subset \mathbb{R}$ ein Intervall und sei $\gamma \colon I \longrightarrow \mathbb{R}^n$ eine zweimal differenzierbare nach Bogenlänge parametrisierte Kurve in $\mathbb{R}^n$ und sei $t \in I^\circ$. Dann ist

$$\kappa_\gamma(t) := \left\| \gamma''(t) \right\|_2 \in \mathbb{R}$$

die *Krümmung von γ in t*. (Dabei fassen wir $\gamma''(t)$ wie $\gamma'(t)$ als Vektor in $\mathbb{R}^n$ auf).

Im Falle ebener Kurven kann man diese Krümmung noch durch ein Vorzeichen verfeinern. Dies beruht auf der Tatsache, dass der der Krümmung zugrundeliegende Vektor senkrecht auf der Geschwindigkeitsrichtung steht:

Proposition 3.65 (Krümmung vs. Geschwindigkeit)**.** *Sei $I \subset \mathbb{R}$ ein Intervall und sei $\gamma \colon I \longrightarrow \mathbb{R}^n$ eine zweimal differenzierbare nach Bogenlänge parametrisierte Kurve. Dann gilt für alle $t \in I^\circ$, dass*

$$\gamma'(t) \perp \gamma''(t).$$

Definition 3.66 (Einheitsnormalenvektor, signierte Krümmung)**.** Sei $I \subset \mathbb{R}$ ein Intervall, sei $\gamma \colon I \longrightarrow \mathbb{R}^2$ eine zweimal differenzierbare nach Bogenlänge parametrisierte Kurve, und sei $t \in I^\circ$.

– Der *Einheitsnormalenvektor von γ in t* ist

$$\nu_\gamma(t) := \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \cdot \gamma'(t) \in \mathbb{R}^2.$$

– Wegen $\nu_\gamma(t) \perp \gamma'(t)$ und $\gamma''(t) \perp \gamma'(t)$ sowie $\|\nu_\gamma(t)\|_2 = \|\gamma'(t)\|_2 = 1 \neq 0$ existiert also (da $\dim_\mathbb{R} \mathbb{R}^2 = 2$) genau ein $\widetilde{\kappa}_\gamma(t) \in \mathbb{R}$ mit

$$\gamma''(t) = \widetilde{\kappa}_\gamma(t) \cdot \nu_\gamma(t).$$

Man nennt $\widetilde{\kappa}_\gamma(t)$ die *signierte Krümmung von γ in t*.

Nach Definition ist in dieser Situation $|\widetilde{\kappa}_\gamma| = \kappa_\gamma$, die signierte Krümmung ist also eine Verfeinerung der Krümmung.

Analog zu (parametrisierten) Kurven kann man auch (parametrisierte) Flächen in $\mathbb{R}^n$ definieren und Krümmungsbegriffe für Flächen einführen. Die Krümmungsbegriffe für Flächen beruhen auf den Krümmungsbegriffen für Kurven.

Wir ändern nun den Blickwinkel und betrachten Teilmengen von $\mathbb{R}^n$ statt Parametrisierungen von Teilmengen von $\mathbb{R}^n$:

Definition 3.67 (Untermannigfaltigkeit). Seien $k, n \in \mathbb{N}$ mit $k \leq n$. Eine Teilmenge $M \subset \mathbb{R}^n$ ist eine *k-dimensionale C^1-Untermannigfaltigkeit von $\mathbb{R}^n$*, wenn es zu jedem $x \in M$ eine *C^1-Karte um x in M* gibt, d.h., wenn es eine offene Umgebung $U \subset \mathbb{R}^n$ von x in $\mathbb{R}^n$ und einen C^1-Diffeomorphismus $\varphi\colon U \longrightarrow \mathbb{R}^n$ mit

$$\varphi(U \cap M) = \mathbb{R}^k \times \{0\} \subset \mathbb{R}^n$$

gibt. (D.h. lokal sieht $M \subset \mathbb{R}^n$ wie $\mathbb{R}^k \times \{0\} \subset \mathbb{R}^n$ aus).

Bemerkung 3.68 (Urbilder regulärer Werte). Seien $n, k \in \mathbb{N}$ mit $k \leq n$, sei $f \in C^1(\mathbb{R}^n, \mathbb{R}^{n-k})$ und sei $y \in \mathbb{R}^{n-k}$ ein *regulärer Wert von f*, d.h. für alle $x \in f^{-1}(\{y\})$ ist $\operatorname{rg} Jf(x) = n - k$. Dann ist $f^{-1}(\{y\})$ eine k-dimensionale C^1-Untermannigfaltigkeit von $\mathbb{R}^n$.

Untermannigfaltigkeiten können lokal linear durch Tangentialräume approximiert werden:

Definition 3.69 (Tangentialraum). Seien $k \in \mathbb{N}$, sei $M \subset \mathbb{R}^n$ eine C^1-Untermannigfaltigkeit und sei $x \in M$. Dann heißt

$$\begin{aligned} T_x^{\mathbb{R}^n} M := \bigl\{ v \in \mathbb{R}^n \bigm| \; & \text{es gibt ein } \varepsilon \in \mathbb{R}_{>0} \text{ und} \\ & \text{eine parametrisierte Kurve } \gamma\colon (-\varepsilon, \varepsilon) \to \mathbb{R}^n \\ & \text{mit } \gamma(0) = x,\ \gamma((-\varepsilon, \varepsilon)) \subset M, \text{ und } \gamma'(0) = v \bigr\} \subset \mathbb{R}^n \end{aligned}$$

Tangentialraum von M in x.

Proposition 3.70 (Tangentialräume sind Untervektorräume). *Seien $k, n \in \mathbb{N}$ mit $k \leq n$, sei $M \subset \mathbb{R}^n$ eine k-dimensionale C^1-Untermannigfaltigkeit und sei $x \in M$. Dann ist $T_x^{\mathbb{R}^n} M \subset \mathbb{R}^n$ ein Untervektorraum von $\mathbb{R}^n$ und es gilt*

$$\dim_{\mathbb{R}} T_x^{\mathbb{R}^n} M = k.$$

Insbesondere ist die Dimension von (nicht-leeren) C^1-Untermannigfaltigkeiten von $\mathbb{R}^n$ eindeutig bestimmt.

Indem man nun das allgemeine Prinzip verwendet, lokale Eigenschaften via Karten auf globale Eigenschaften von Mannigfaltigkeiten hochzuziehen, kann man die Differentialrechnung auf Mannigfaltigkeiten übertragen. Zum Beispiel kann man auf diese Weise definieren, wann eine Abbildung $M_1 \longrightarrow M_2$ zwischen C^1-Untermannigfaltigkeiten $M_1 \subset \mathbb{R}^{n_1}$ bzw. $M_2 \subset \mathbb{R}^{n_2}$ stetig differenzierbar ist; ist $f\colon M_1 \longrightarrow M_2$ stetig differenzierbar und $x \in M_1$, so erhält man eine lineare Abbildung

$$T_x f\colon T_x^{\mathbb{R}^{n_1}} M_1 \longrightarrow T_{f(x)}^{\mathbb{R}^{n_2}} M_2,$$

die man als Ableitung von f in x auffassen kann.

Im nächsten Abstraktionsschritt definiert man einen intrinsichen Mannigfaltigkeitsbegriff, der ohne einen umgebenden Raum $\mathbb{R}^n$ auskommt, sowie Tangentialbündel, Differenzierbarkeit, Krümmung, Volumen, ... Dies ist Gegenstand der Analysis IV, sowie der Differentialtopologie und Differentialgeometrie. Anwendungen findet diese Theorie zum Beispiel in der Kartographie und in der Physik.

4 Mehrdimensionale Analysis – Integration

Wir geben nun einen kurzen Überblick über die wichtigsten Begriffe und Sätze zur Integration von Funktionen vom Typ $\mathbb{R}^n \longrightarrow \mathbb{R}$; Details werden in den Vorlesungen Analysis III (Maßtheorie und Funktionentheorie), sowie Analysis IV (Analysis auf Mannigfaltigkeiten) behandelt.

Konstruktion des Lebesgue-Integrals – Skizze

Wir skizzieren kurz die Konstruktion des Lebesgue-Integrals auf $\mathbb{R}^n$. Wie beim Riemann-Integral auf $\mathbb{R}$ beginnen wir mit Treppenfunktionen – jedoch müssen wir etwas allgemeinere Grundmengen als Intervalle/Quader für die Stufen zulassen und einen Volumenbegriff für solche Mengen erklären. Der erste Schritt ist daher das Lebesgue-Maß auf $\mathbb{R}^n$:

Das Lebesgue-Maß auf $\mathbb{R}^n$

Wir beginnen mit grundlegenden Begriffen aus der Maßtheorie; ähnlich zu topologischen Räumen formalisiert man Maßtheorie über Mengensysteme:

Definition 4.1 (messbarer Raum, σ-Algebra, messbare Funktion, Maß)**.**

- Ein *messbarer Raum* ist ein Paar (X, S), wobei X eine Menge und S eine *σ-Algebra auf X* ist, d.h. $S \subset P(X)$ erfüllt die folgenden Eigenschaften:
  - Es ist $\emptyset \in S$ und $X \in S$.
  - Für alle Folgen $(A_n)_{n\in\mathbb{N}}$ in S ist $\bigcup_{n\in\mathbb{N}} A_n \in S$.
  - Für alle $A \in S$ ist $X \setminus A \in S$.
- Seien (X, S_X) und (Y, S_Y) messbare Räume. Eine Abbildung $f \colon X \longrightarrow Y$ ist *messbar*, wenn: für alle $A \in S_Y$ ist $f^{-1}(A) \in S_X$.
- Sei (X, S) ein messbarer Raum. Ein *Maß auf* (X, S) ist eine Abbildung

  $$\mu \colon S \longrightarrow \mathbb{R}_{\geq 0} \cup \{\infty\}$$

  mit folgenden Eigenschaften:
  - Es ist $\mu(\emptyset) = 0$ und $\mu(X) \neq 0$.
  - Ist $(A_n)_{n\in\mathbb{N}}$ eine Folge paarweise disjunkter Mengen aus S, so ist

  $$\mu\Bigl(\bigcup_{n\in\mathbb{N}} A_n\Bigr) = \sum_{n=0}^{\infty} \mu(A_n).$$

Konvention 4.2 („Rechnen“ mit ∞)**.** Wir definieren

$$\begin{aligned} \infty + \infty &:= \infty, \\ a + \infty &:= \infty, \\ \infty + a &:= \infty, \\ b \cdot \infty &:= \infty, \\ \infty \cdot b &:= \infty \end{aligned}$$

für alle $a \in \mathbb{R}$ und alle $b \in \mathbb{R}_{>0}$. Man beachte jedoch, dass „$\infty - \infty$“ *nicht* definiert ist.

Definition 4.3 (Borel-σ-Algebra)**.** Sei $n \in \mathbb{N}$. Die bezüglich Inklusion kleinste σ-Algebra auf $\mathbb{R}^n$, die alle bezüglich der Standardtopologie offenen Mengen von $\mathbb{R}^n$ enthält, heißt *Borel-σ-Algebra* auf $\mathbb{R}^n$ und wird mit $B(\mathbb{R}^n)$ bezeichnet.

Satz 4.4 (Lebesgue-Maß)**.** *Sei $n \in \mathbb{N}$. Es gibt genau ein Maß $\lambda^n \colon B(\mathbb{R}^n) \longrightarrow \mathbb{R}_{\geq 0} \cup \{\infty\}$ auf $(\mathbb{R}^n, B(\mathbb{R}^n))$ mit folgender Eigenschaft: Für alle $a, b \in \mathbb{R}^n$ mit $a_j \leq b_j$ für alle $j \in \{1, \dots, n\}$ gilt*

$$\lambda^n\big([a_1, b_1] \times \dots \times [a_n, b_n]\big) = \prod_{j=1}^{n} (b_j - a_j).$$

Man nennt λ^n das Lebesgue-Maß *auf $\mathbb{R}^n$.*

Bemerkung 4.5 (Lebesgue-messbare Mengen)**.** Man kann $B(\mathbb{R}^n)$ bezüglich λ^n zu einer etwas größeren σ-Algebra $\overline{B}(\mathbb{R}^n)$ „vervollständigen“, und man kann λ^n zu einem Maß auf $\overline{B}(\mathbb{R}^n)$ fortsetzen. Wir bezeichnen diese Fortsetzung auch wieder als Lebesgue-Maß λ^n und nennen die Mengen in $\overline{B}(\mathbb{R}^n)$ *Lebesgue-messbar*.

Bemerkung 4.6 (stetige Abbildungen sind messbar)**.** Ist $n \in \mathbb{N}$ und $f \colon \mathbb{R}^n \longrightarrow \mathbb{R}$ stetig, so ist f bezüglich den σ-Algebren $B(\mathbb{R}^n)$ und $B(\mathbb{R})$ (bzw. $\overline{B}(\mathbb{R}^n)$ und $\overline{B}(\mathbb{R})$) messbar.

Caveat 4.7 (exotische Mengen)**.** Man kann zeigen, dass aus dem Auswahlaxiom folgt, dass $\overline{B}(\mathbb{R}^n) \neq P(\mathbb{R}^n)$ gilt, dass es also Mengen in $\mathbb{R}^n$ gibt, die *nicht* Lebesgue-messbar sind; dies hängt mit dem sogenannten Banach-Tarski-Paradoxon zusammen.

Das Lebesgue-Integral auf Treppenfunktionen

Wir setzen nun messbare Stufen zu Treppenfunktionen zusammen und definieren für diese ein Integral:

Definition 4.8 (Treppenfunktionen)**.** Sei $n \in \mathbb{N}$.

- Eine Fuktion $f \colon \mathbb{R}^n \longrightarrow \mathbb{R}$ ist eine *Treppenfunktion*, wenn es ein $k \in \mathbb{N}$ und paarweise disjunkte Mengen $A_1, \dots, A_k \in \overline{B}(\mathbb{R}^n)$ sowie $a_1, \dots, a_k \in \mathbb{R}$ gibt mit: Für alle $j \in \{1, \dots, k\}$ ist $\mu(A_j) < \infty$ und es gilt

  $$f = \sum_{j=1}^{k} a_j \cdot \chi_{A_j} \colon \mathbb{R}^n \longrightarrow \mathbb{R}.$$

  Hierbei bezeichnet zu einer Menge $A \subset \mathbb{R}^n$ das Symbol χ_A die zugehörige *charakteristische Funktion*:

  $$\begin{aligned} \chi_A \colon \mathbb{R}^n &\longrightarrow \mathbb{R} \\ x &\longmapsto \begin{cases} 1 & \text{falls } x \in A \\ 0 & \text{falls } x \in \mathbb{R}^n \setminus A. \end{cases} \end{aligned}$$

– Wir bezeichnen die Menge aller solchen Treppenfunktionen mit $T(\mathbb{R}^n, \lambda^n)$.

Man beachte, dass $T(\mathbb{R}^n, \lambda^n)$ bezüglich punktweiser Addition bzw. Skalarmultiplikation ein reeller Vektorraum ist.

Proposition 4.9 (Integral von Treppenfunktionen)**.** *Sei $n \in \mathbb{N}$ und sei*

$$\begin{aligned} I \colon T(\mathbb{R}^n, \lambda^n) &\longrightarrow \mathbb{R} \\ \sum_{j=1}^{k} a_j \cdot \chi_{A_j} &\longmapsto \sum_{j=1}^{k} a_j \cdot \lambda^n(A_j). \end{aligned}$$

Dann gilt:

1. *Die Abbildung I ist wohldefiniert und $\mathbb{R}$-linear.*
2. *Die Abbildung*

$$\begin{aligned} \|\cdot\|_1 \colon T(\mathbb{R}^n, \lambda^n) &\longrightarrow \mathbb{R}_{\geq 0} \\ f &\longmapsto I\bigl(|f|\bigr) \end{aligned}$$

   ist eine (Halb-)Norm auf $T(\mathbb{R}^n, \lambda^n)$.
3. *Die Abbildung I ist bezüglich $\|\cdot\|_1$ stetig.*

Vervollständigung und Fortsetzung des Integrals

Wir vervollständigen nun den (halb-)normierten Raum der Treppenfunktionen und setzen dann das Integral von Treppenfunktionen auf diese Vervollständigung fort:

Definition 4.10. Sei $n \in \mathbb{N}$. Die Vervollständigung von $T(\mathbb{R}^n, \lambda^n)$ bezüglich $\|\cdot\|_1$ bezeichnen wir mit $L^1(\mathbb{R}^n, \lambda^n)$.

Der folgende zentrale Satz zeigt, dass die Elemente des (zunächst abstrakt definierten) Raumes $L^1(\mathbb{R}^n, \lambda^n)$ in einem gewissen Sinne wieder als Funktionen $\mathbb{R}^n \longrightarrow \mathbb{R}$ aufgefasst werden können:

Satz 4.11 (Lebesgue-integrierbare Funktionen und das Lebesgue-Integral)**.** *Sei $n \in \mathbb{N}$.*

1. *Es gibt einen Isomorphismus*

$$L^1(\mathbb{R}^n, \lambda^n) \longleftrightarrow \frac{\text{gewisse messbare Funktionen } \mathbb{R}^n \to \mathbb{R}}{\bigl\{f \colon \mathbb{R}^n \to \mathbb{R} \bigm| f \text{ ist messbar und } \lambda^n(f^{-1}(\mathbb{R} \setminus \{0\})) = 0\bigr\}}$$

   der die kanonische Abbildung $T(\mathbb{R}^n, \lambda^n) \longrightarrow L^1(\mathbb{R}^n, \lambda^n)$ fortsetzt.
   Man nennt daher die Elemente von $L^1(\mathbb{R}^n, \mathbb{R})$ Lebesgue-integrierbare Funktionen auf $\mathbb{R}^n$ *und fasst sie als Funktionen $\mathbb{R}^n \longrightarrow \mathbb{R}$ auf.*
2. *Die eindeutige stetige lineare Fortsetzung*

$$\int_{\mathbb{R}^n} \cdot \, d\lambda^n \colon L^1(\mathbb{R}^n, \lambda^n) \longrightarrow \mathbb{R}$$

   von $I \colon T(\mathbb{R}^n, \lambda^n) \longrightarrow \mathbb{R}$ auf $L^1(\mathbb{R}^n, \lambda^n)$ heißt Lebesgue-Integral auf $\mathbb{R}^n$*.*

3. Die Abbildung

$$\begin{aligned} \|\cdot\|_1 \colon L^1(\mathbb{R}^n, \lambda^n) &\longrightarrow \mathbb{R}_{\geq 0} \\ f &\longrightarrow \int_{\mathbb{R}^n} |f| \, d\lambda^n \end{aligned}$$

ist eine Norm auf $L^1(\mathbb{R}^n, \lambda^n)$.

Man beachte, dass Elemente aus $L^1(\mathbb{R}^n, \lambda^n)$ nicht eindeutig Funktionen $\mathbb{R}^n \longrightarrow \mathbb{R}$ entsprechen, sondern nur bis auf λ^n-Nullmengen (d.h. Mengen, deren λ^n-Maß 0 ist).

Notation 4.12. Ist $f \in L^1(\mathbb{R}^n, \lambda^n)$ und $A \in \overline{B}(\mathbb{R}^n)$, so schreiben wir

$$\int_A f \, d\lambda^n := \int_{\mathbb{R}^n} f \cdot \chi_A \, d\lambda^n.$$

Wichtige Sätze

Wir stellen nun einen Werkzeugkasten der wichtigsten Sätze über das Lebesgue-Integral zusammen, der es bereits ermöglicht, das Lebesgue-Integral in vielen Fällen zu berechnen:

Satz 4.13 (Positivität des Integrals)**.** *Sei* $n \in \mathbb{N}$ *und sei* $f \in L^1(\mathbb{R}^n, \lambda^n)$ *so dass* λ^n*-fast überall* $f \geq 0$ *gilt (d.h. es ist* $\lambda^n(\{x \in \mathbb{R}^n \mid f(x) < 0\} = 0)$*). Dann ist*

$$\int_{\mathbb{R}^n} f \, d\lambda^n \geq 0.$$

Satz 4.14 (Riemann- vs. Lebesgue-Integral)**.** *Seien* $a, b \in \mathbb{R}$ *mit* $a \leq b$ *und sei* $f \colon [a, b] \longrightarrow \mathbb{R}$ *Riemann-integrierbar. Dann gibt es eine Funktion* $\widetilde{f} \in L^1([a, b], \lambda^1|_{[a,b]})$*, die* λ^n*-fast überall mit* f *übereinstimmt und es gilt*

$$\int_a^b f = \int_{[a,b]} \widetilde{f} \, d\lambda^1.$$

Caveat 4.15. Es gibt jedoch viel mehr Lebesgue-integrierbare Funktionen als Riemann-integrierbare Funktionen!

Satz 4.16 (Integrierbarkeit stetiger Funktionen)**.** *Sei* $n \in \mathbb{N}$*, sei* $f \colon \mathbb{R}^n \longrightarrow \mathbb{R}$ *stetig und sei* $A \subset \mathbb{R}^n$ *kompakt. Dann ist* $\chi_A \cdot f \in L^1(\mathbb{R}^n, \lambda^n)$.

Satz 4.17 (monotone Konvergenz)**.** *Sei* $n \in \mathbb{N}$*. Sei* $(f_k)_{k \in \mathbb{N}}$ *eine* λ^n*-fast überall monoton wachsende Folge von Elementen von* $L^1(\mathbb{R}^n, \lambda^n)$ *mit der Eigenschaft, dass die Menge* $\{\int_{\mathbb{R}^n} f_k \, d\lambda^n \mid k \in \mathbb{N}\} \subset \mathbb{R}$ *beschränkt ist. Dann gibt es ein* $f \in L^1(\mathbb{R}^n, \lambda^n)$*, so dass* $(f_k)_{k \in \mathbb{N}}$ *bezüglich* $\|\cdot\|_1$ *und* λ^n*-fast überall punktweise gegen* f *konvergiert. Insbesondere ist*

$$\lim_{k \to \infty} \int_{\mathbb{R}^n} f_k \, d\lambda^n = \int_{\mathbb{R}^n} f \, d\lambda^n.$$

Satz 4.18 (dominierte Konvergenz). *Sei $n \in \mathbb{N}$ und sei $(f_k)_{k\in\mathbb{N}}$ eine Folge in $L^1(\mathbb{R}^n, \lambda^n)$, die λ^n-fast überall punktweise gegen $f\colon \mathbb{R}^n \longrightarrow \mathbb{R}$ konvergiert. Außerdem gebe es ein $g \in L^1(\mathbb{R}^n, \lambda^n)$ mit der Eigenschaft, dass λ^n-fast überall $|f_k| \leq g$ für alle $k \in \mathbb{N}$ gilt. Dann ist $f \in L^1(\mathbb{R}^n, \lambda^n)$ und*

$$\lim_{k\to\infty} \int_{\mathbb{R}^n} f_k \, d\lambda^n = \int_{\mathbb{R}^n} f \, d\lambda^n.$$

Satz 4.19 (Satz von Fubini). *Seien $n, n_1, n_2 \in \mathbb{N}$ mit $n_1+n_2 = n$, und sei $f\colon \mathbb{R}^n \longrightarrow \mathbb{R}$ messbar. Für $x \in \mathbb{R}^{n_1}$ sei*

$$\begin{aligned} f_x\colon \mathbb{R}^{n_2} &\longrightarrow \mathbb{R} \\ y &\longmapsto f(x, y). \end{aligned}$$

1. *Für λ^{n_1}-fast alle $x \in \mathbb{R}^{n_1}$ sei $f_x \in L^1(\mathbb{R}^{n_2}, \lambda^n)$ und die Abbildung*

$$\begin{aligned} \mathbb{R}^{n_1} &\longrightarrow \mathbb{R} \\ x &\longmapsto \int_{\mathbb{R}^{n_2}} |f_x| \, d\lambda^{n_2} \end{aligned}$$

   sei in $L^1(\mathbb{R}^{n_1}, \lambda^{n_1})$. Dann ist $f \in L^1(\mathbb{R}^n, \lambda^n)$.
2. *Sei $f \in L^1(\mathbb{R}^n, \lambda^n)$. Dann gilt: Für λ^{n_1}-fast alle $x \in \mathbb{R}^{n_1}$ ist $f_x \in L^1(\mathbb{R}^{n_2}, \lambda^{n_2})$ und*

$$\begin{aligned} \mathbb{R}^{n_1} &\longrightarrow \mathbb{R} \\ x &\longmapsto \int_{\mathbb{R}^{n_2}} |f_x| \, d\lambda^{n_2} \end{aligned}$$

   ist in $L^1(\mathbb{R}^{n_1}, \lambda^{n_1})$ und es gilt

$$\begin{aligned} \int_{\mathbb{R}^n} f \, d\lambda^n &= \int_{\mathbb{R}^{n_1}} \int_{\mathbb{R}^{n_2}} f_x(y) \, d\lambda^{n_2}(y) \, d\lambda^{n_1}(x) \\ &= \int_{\mathbb{R}^{n_1}} \int_{\mathbb{R}^{n_2}} f(x, y) \, d\lambda^{n_2}(y) \, d\lambda^{n_1}(x). \end{aligned}$$

Satz 4.20 (Transformationssatz). *Sei $n \in \mathbb{N}$, seien $A, B \subset \mathbb{R}^n$ offene Mengen und sei $\varphi\colon A \longrightarrow B$ ein C^1-Diffeomorphismus. Sei außerdem $f\colon \mathbb{R}^n \longrightarrow \mathbb{R}$ messbar. Dann ist $f \cdot \chi_B$ genau dann Lebesgue-integrierbar, wenn $(f \circ \varphi) \cdot \chi_A$ Lebesgue-integrierbar ist und es gilt*

$$\int_{\varphi(A)} f \, d\lambda^n = \int_A (f \circ \varphi) \cdot |\det J\varphi| \, d\lambda^n.$$

Zum Beispiel kann man mehrdimensionale Integrale auch verwenden, um Integrale von Funktionen $\mathbb{R} \longrightarrow \mathbb{R}$ besser zu bestehen; etwa folgt

$$\int_{\mathbb{R}} e^{-x^2} \, d\lambda^1(x) = \sqrt{\pi}$$

aus der Betrachtung des Integrals einer geeigneten Funktion $\mathbb{R}^2 \longrightarrow \mathbb{R}$.

Via Karten und Differentialformen kann man Integration und Volumina etc. auch auf Mannigfaltigkeiten erklären (Analysis IV).

5 Gewöhnliche Differentialgleichungen

Systeme in den Naturwissenschaften oder Wirtschaftswissenschaften werden oft durch sogenannte Differentialgleichungen modelliert. Außerdem treten Differentialgleichungen auch z.B. in der Differentialgeometrie auf. Grob gesagt beschreibt eine Differentialgleichung die gewünschten Beziehungen zwischen einer Funktion und ihren Ableitungen.

Wir wollen uns im folgenden mit solchen Gleichungen beschäftigen. Wichtige Fragestellungen sind dabei:

- Existiert eine Lösung? Zumindest lokal?
- Sind Lösungen eindeutig?
- (Wie) Kann man Lösungen explizit berechnen?
- Was kann man über die „langfristige“ Entwicklung des Systems aussagen?
- Wie stabil sind solche Gleichungen/Lösungen unter „kleinen Veränderungen“?

Wir beginnen mit der Formalisierung des Begriffs „gewöhnliche Differentialgleichung“ bzw. „Lösung einer gewöhnlichen Differentialgleichung.“

Definition 5.1 (gewöhnliche Differentialgleichung, Lösung davon)**.** Seien $n, m, k \in \mathbb{N}$.

- Eine *gewöhnliche Differentialgleichung der Ordnung* $\leq k$ ist gegeben durch eine Abbildung $D \longrightarrow \mathbb{R}^m$, wobei $D \subset \mathbb{R} \times (\mathbb{R}^n)^{k+1}$ ist.
- Sei $D \subset \mathbb{R} \times (\mathbb{R}^n)^{k+1}$ und sei $f \colon D \longrightarrow \mathbb{R}^m$. Ist $I \subset \mathbb{R}$ ein Intervall, so ist $y \in C^k(I, \mathbb{R}^n)$ eine *Lösung der durch f gegebenen Differentialgleichung*, wenn folgende Bedingungen erfüllt sind:
  - Für alle $t \in I^\circ$ ist $\bigl(t, y(t), y'(t), \dots, y^{(k)}(t)\bigr) \in D$, und
  - für alle $t \in I^\circ$ ist
  $$f\bigl(t, y(t), y'(t), \dots, y^{(k)}(t)\bigr) = 0.$$

  Wir schreiben dann
  $$\begin{aligned} \text{Gesucht:} \quad & y \in C^k\bigl(\langle \mathbb{R}^1 \rangle, \langle \mathbb{R}^n \rangle\bigr) \quad \text{mit} \\ & 0 = f(\,\cdot\,, y, y', \dots, y^{(k)}) \end{aligned}$$
  für das entsprechende Problem.

Bemerkung 5.2 (partielle Differentialgleichungen)**.** Gewöhnliche Differentialgleichungen handeln von Funktionen mit eindimensionalem Definitionsbereich; treten Funktionen mit weiteren Parametern und ihre partiellen Ableitungen auf, so gelangt man zu den sogenannten *partiellen Differentialgleichungen*. Dieses Gebiet ist jedoch deutlich schiweriger und unübersichtlicher als der Zoo der gewöhnlichen Differentialgleichungen.

Gewöhnliche Differentialgleichungen, geometrisch

Um ein besseres Gefühl für Differentialgleichungen zu bekommen, geben wir für zwei spezielle Arten von gewöhnlichen Differentialgleichungen geometrische Veranschaulichungen:

Definition 5.3 (Richtungsfeld)**.** Sei $D \subset \mathbb{R}\times\mathbb{R}$ und sei $f\colon D \longrightarrow \mathbb{R}$. Das *Richtungsfeld* der zugehörigen gewöhnlichen Differentialgleichung

$$\begin{array}{ll} \text{Gesucht:} & y \in C^1\bigl(\langle\mathbb{R}^1\rangle, \langle\mathbb{R}^1\rangle\bigr) \quad \text{mit} \\ & y' \;=\; f(\,\cdot\,, y) \end{array}$$

erster Ordnung ist die Menge

$$\left\{\left((t,x), \begin{pmatrix} 1 \\ f(t,x) \end{pmatrix}\right) \,\middle|\, (t,x) \in D\right\} \subset D \times \mathbb{R}^2.$$

(Diese Differentialgleichung heißt „explizit," da sie bereits nach der ersten Ableitung aufgelöst ist).

Eine Funktion löst eine solche Differentialgleichung genau dann, wenn das Richtungsfeld der Differentialgleichung überall tangential zur Funktion verläuft. Mithilfe von Richtungsfeldern kann man also solche gewöhnlichen Differentialgleichungen graphisch „lösen."

Definition 5.4 (Phasenebene, stationärer Punkt, erstes Integral)**.** Sei $D \subset \mathbb{R}^2$ und sei $f\colon D \longrightarrow \mathbb{R}^2$. Wir betrachten das autonome explizite zweidimensionale System

$$\begin{array}{ll} \text{Gesucht:} & y_1, y_2 \in C^1\bigl(\langle\mathbb{R}^1\rangle, \langle\mathbb{R}^1\rangle\bigr) \quad \text{mit} \\ & y_1' \;=\; f_1(y_1, y_2) \\ & y_2' \;=\; f_2(y_1, y_2) \end{array}$$

erster Ordnung. (Eine solche Differentialgleichung heißt „autonom" da die „rechte" Seite nicht vom Parameter „t" abhängt.)

- Die *Phasenebene* (oder das *Phasenportrait*) dieser gewöhnlichen Differentialgleichung ist die Menge

$$\bigl\{\bigl(x, f(x)\bigr) \bigm| x \in D\bigr\} \subset D \times \mathbb{R}^2.$$

- Ist $x \in D$ mit $f(x) = 0$, so heißt x *stationärer Punkt* (oder *Gleichgewichtspunkt*) dieser Gleichung.
- Eine (nicht-konstante) Abbildung $\varphi\colon D \longrightarrow \mathbb{R}$ heißt *erstes Integral* der obigen Differentialgleichung, falls: Ist $y \in C^1(I, \mathbb{R}^2)$ eine Lösung dieser Differentialgleichung, so ist $\varphi \circ y$ konstant (d.h. alle Lösungen dieser Differentialgleichung verlaufen in den Niveaumengen von φ).

Eine parametrisierte Kurve in $\mathbb{R}^2$ löst genau dann ein solches Differentialgleichungssystem, wenn die Vektoren der Phasenebene überall tangential an die Kurve anliegen. Mithilfe der Phasenebene kann man also solche gewöhnlichen Differentialgleichungen graphisch „lösen."

Man beachte, dass die stationären Punkte gerade den konstanten Lösungen entsprechen, und dass erste Integrale helfen können, Eigenschaften von Lösungen zu finden (z.B. Beschränktheit, ohne die Differentialgleichung selbst lösen zu müssen).

Existenz und Eindeutigkeit

Wir werden uns nun etwas systematischer mit den folgenden Fragen beschäftigen: Existieren Lösungen (zumindest lokal)? Sind sie eindeutig?

Der zentrale Satz in der Theorie der gewöhnlichen Differentialgleichungen ist der Satz von Picard-Lindelöf:

Satz 5.5 (Satz von Picard-Lindelöf)**.** *Sei $n \in \mathbb{N}$, sei $D \subset \mathbb{R} \times \mathbb{R}^n$ offen, sei $t_0 \in \mathbb{R}$ und $x_0 \in \mathbb{R}^n$ mit $(t_0, x_0) \in D$ und sei $f \colon D \longrightarrow \mathbb{R}^n$ stetig und lokal Lipschitz in den letzten n Variablen. Wir betrachten die folgende Differentialgleichung (ein sogenanntes* Anfangswertproblem*):*

$$\begin{array}{lll} \textit{Gesucht:} & y \in C^1\bigl(\langle \mathbb{R}^1 \rangle, \langle \mathbb{R}^n \rangle\bigr) & \quad \textit{mit} \\ & y' \quad = \quad f(\,\cdot\,, y) & \\ & y(t_0) \quad = \quad x_0 & \end{array}$$

Dann gilt:

1. Lokale Existenz. *Es existiert ein $\delta_0 \in \mathbb{R}_{>0}$, so dass dieses Anfangswertproblem eine Lösung in $C^1([t_0 - \delta_0, t_0 + \delta_0], \mathbb{R}^n)$ besitzt.*
2. Eindeutigkeit. *Ist $I \subset \mathbb{R}$ ein Intervall mit $t_0 \in I$ und sind $y, \overline{y} \in C^1(I, \mathbb{R}^n)$ Lösungen des obigen Anfangswertproblems, so ist $y = \overline{y}$.*

Definition 5.6 (lokal Lipschitz)**.** Sei $n \in \mathbb{N}$, sei $D \subset \mathbb{R} \times \mathbb{R}^n$ offen und sei $f \colon D \longrightarrow \mathbb{R}^n$ eine Abbildung. Dann ist f *lokal Lipschitz in den letzten n Variablen*, wenn: Für alle $(t_0, x_0) \in D$ (mit $t_0 \in \mathbb{R}$ und $x_0 \in \mathbb{R}^n$) gibt es ein $\delta \in \mathbb{R}_{>0}$, eine offene Umgebung $U \subset \mathbb{R}^n$ von x_0 mit $[t_0 - \delta, t_0 + \delta] \times U \subset D$, und ein $L \in \mathbb{R}_{\geq 0}$ mit

$$\forall_{t \in [t_0 - \delta, t_0 + \delta]} \quad \forall_{x, \overline{x} \in U} \quad \bigl\| f(t, x) - f(t, \overline{x}) \bigr\|_2 \leq L \cdot \|x - \overline{x}\|_2.$$

Der Beweis des Satzes von Picard-Lindelöf beruht auf dem Banachschen Fixpunktsatz (Satz 1.19); insbesondere erhält man so auch ein Verfahren, um Lösungen iterativ anzunähern (da der Banachsche Fixpunktsatz ein solches Verfahren zum Bestimmen von Fixpunkten liefert).

Das folgende Kriterium erlaubt es in vielen Fällen nachzuweisen, dass die Voraussetzungen des Satzes von Picard-Lindelöf erfüllt sind:

Proposition 5.7 (ein hinreichendes Kriterium für lokal Lipschitz)**.** *Sei $I \subset \mathbb{R}$ offen, sei $X \subset \mathbb{R}^n$ offen und sei $f \colon I \times X \longrightarrow \mathbb{R}^n$ eine Abbildung mit den folgenden Eigenschaften:*

- *Für alle $t \in I$ ist die Abbildung*

$$\begin{aligned} f_t \colon X &\longrightarrow \mathbb{R}^n \\ x &\longmapsto f(t, x) \end{aligned}$$

  stetig differenzierbar.

– *Die Abbildung*

$$\begin{aligned} I \times X &\longrightarrow \mathbb{R}^{n\times n} \\ (t,x) &\longmapsto Jf_t(x) \end{aligned}$$

ist stetig.
Dann ist f lokal Lipschitz bezüglich der letzten n Variablen.

Der Satz von Picard-Lindelöf garantiert zunächst nur lokale Existenz von solchen Anfangswertproblemen. Unter welchen Voraussetzungen kann man solche Lösungen fortsetzen bzw. unter welchen Voraussetzungen existieren globale Lösungen?

Es gibt zwei Phänomene, die die Fortsetzbarkeit von Lösungen verhindern:
– Explosion von Lösungen
– Erreichen des Randes des Definitionsbereichs des Anfangswertproblems
Andernfalls können Lösungen fortgesetzt werden:

Satz 5.8 (Fortsetzbarkeit von Lösungen)**.** *Sei $I \subset \mathbb{R}$ ein offenes Intervall, sei $D \subset \mathbb{R}^n$ offen und sei $f\colon I \times \mathbb{R}^n \longrightarrow \mathbb{R}^n$ stetig und lokal Lipschitz in den letzten n Variablen. Wir betrachten das System*

$$\begin{aligned} \textit{Gesucht:} \quad & y \in C^1\bigl(\langle\mathbb{R}^1\rangle, \langle\mathbb{R}^n\rangle\bigr) \quad \textit{mit} \\ & y' \;=\; f(\,\cdot\,, y) \end{aligned}$$

Seien $a, b \in I$ mit $a < b$ und sei $y \in C^1((a,b), \mathbb{R}^n)$ eine Lösung dieser Differentialgleichung mit folgenden Eigenschaften:
– *Es existiert ein $c \in \mathbb{R}_{\geq 0}$ mit*

$$\forall_{t \in (a,b)} \quad \bigl\|y(t)\bigr\|_2 \leq c.$$

– *Es ist $\overline{y((a,b))} \subset D$.*
Dann existiert eine Fortsetzung der Lösung y zu einer Lösung auf einer Umgebung von (a,b), d.h. es existiert ein $\delta \in \mathbb{R}_{>0}$ und eine Lösung $\widetilde{y} \in C^1((a-\delta, b+\delta), \mathbb{R}^n)$ der obigen Differentialgleichung.

Der Beweis beruht darauf, dass man den Satz von Picard-Lindelöf auf geeignete Anfangswertprobleme am Rand des Intervalls (a,b) anwendet und dann diese Lösung mit der ursprünglichen verklebt.

Unter geeigneten Wachstumsbedingungen lassen sich dann Lösungen sogar zu globalen Lösungen zusammensetzen:

Korollar 5.9 (globale Existenz)**.** *Sei $I \subset \mathbb{R}$ ein Intervall, sei $(t_0, x_0) \in I \times \mathbb{R}^n$ und sei $f\colon I \times \mathbb{R}^n \longrightarrow \mathbb{R}^n$ stetig und lokal Lipschitz in den letzten n Variablen. Außerdem gebe es $\alpha, \beta \in \mathbb{R}_{\geq 0}$ mit*

$$\forall_{(t,x) \in I \times \mathbb{R}^n} \quad \bigl\|f(t,x)\bigr\|_2 \leq \alpha + \beta \cdot \|x\|_2.$$

Dann besitzt das Anfangswertproblem

$$
\begin{array}{lrcl}
\textit{Gesucht:} & y \in C^1\big(\langle\mathbb{R}^1\rangle, \langle\mathbb{R}^n\rangle\big) & & \textit{mit} \\
& y' & = & f(\,\cdot\,, y) \\
& y(t_0) & = & x_0
\end{array}
$$

genau eine Lösung in $C^1(I, \mathbb{R}^n)$.

Dass die Wachstumsbedingung an f verhindert, dass Lösungen solcher Anfangswertprobleme explodieren (und der obige Fortsetzungssatz somit anwendbar ist), wird mit Hilfe des Lemmas von Gronwall gezeigt.

Separable Gleichungen

Wir beschäftigen uns nun mit einer speziellen Art von gewöhnlichen Differentialgleichungen, für die es eine Lösungsstrategie gibt. Viele weitere gewöhnliche Differentialgleichungen werden durch geeignete Transformationen auf diese zurückgeführt.

Satz 5.10 (Trennung der Variablen)**.** *Seien* $I, J \subset \mathbb{R}$ *offene Intervalle, seien* $h\colon I \longrightarrow \mathbb{R}$ *und* $g\colon J \longrightarrow \mathbb{R}$ *stetig und sei* $(t_0, x_0) \in I \times J$. *Wir betrachten das Anfangswertproblem*

$$
\begin{array}{lrcl}
\textit{Gesucht:} & y \in C^1\big(\langle\mathbb{R}^1\rangle, \langle\mathbb{R}^1\rangle\big) & & \textit{mit} \\
& y' & = & h \cdot (g \circ y) \\
& y(t_0) & = & x_0
\end{array}
$$

Dann gilt:

1. *Ist* $g(x_0) = 0$, *so ist die konstante Funktion* $x_0\colon I \longrightarrow \mathbb{R}$ *eine Lösung dieses Anfangswertproblems (im allgeminen ist diese Lösung jedoch* nicht *eindeutig).*
2. *Ist* $g(x_0) \neq 0$, *so existieren* $\delta, \varepsilon \in \mathbb{R}_{>0}$, *so dass*

$$
\begin{aligned}
y\colon (t_0 - \delta, t_0 + \delta) &\longrightarrow \mathbb{R} \\
t &\longmapsto G^{-1} \circ H(t),
\end{aligned}
$$

wobei

$$
\begin{aligned}
G\colon (x_0 - \varepsilon, x_0 + \varepsilon) &\longrightarrow \mathbb{R} \\
x &\longmapsto \int_{x_0}^{x} \frac{1}{g(s)}\, ds, \\
H\colon (t_0 - \delta, t_0 + \delta) &\longrightarrow \mathbb{R} \\
t &\longmapsto \int_{t_0}^{t} h(s)\, ds,
\end{aligned}
$$

wohldefiniert und eine Lösung des obigen Anfangswertproblems ist.
Ist $\overline{y} \in C^1((t_0 - \delta, t_0 + \delta), \mathbb{R})$ *eine Lösung des obigen Anfangswertproblems, so ist* $\overline{y} = y$.

Caveat 5.11. Man beachte dabei, dass die im obigen Satz betrachteten Anfangswertprobleme im allgemeinen *nicht* die Voraussetzungen des Satzes von Picard-Lindelöf (Satz 5.5) erfüllen, da nicht alle stetigen Funktionen lokal Lipschitz sind.

Caveat 5.12. Im allgemeinen kann man die Integrale und das Inverse in der obigen Lösung *nicht* explizit berechnen.

Lineare Differentialgleichungssysteme

Im folgenden betrachten wir die Klasse der linearen gewöhnlichen Differentialgleichungssysteme etwas genauer. Da man Systeme höherer Ordnung durch Einführung neuer Variablen auf Systeme erster Ordnung zurückführen kann, betrachten wir nur (explizite) Systeme erster Ordnung:

Definition 5.13 (lineare gewöhnliche Differentialgleichung erster Ordnung)**.** Sei $n \in \mathbb{N}$, sei $I \subset \mathbb{R}$ ein Intervall und seien $A\colon I \longrightarrow \mathbb{R}^{n\times n}$ und $b\colon I \longrightarrow \mathbb{R}^n$ stetig. Ein Differentialgleichungssystem der Form

$$\begin{array}{lrcl} \text{Gesucht:} & y \in C^1\bigl(\langle\mathbb{R}^1\rangle, \langle\mathbb{R}^n\rangle\bigr) & & \text{mit} \\ & y' & = & A\cdot y + b \end{array}$$

heißt *lineares (gewöhnliches) Differentialgleichungssystem erster Ordnung.*

Ist $b = 0\colon I \longrightarrow \mathbb{R}^n$, so nennt man das System *homogen*, andernfalls *inhomogen.*

Mit dem globalen Existenzsatz (Korollar 5.9) erhält man (indem man zunächst kompakte Intervalle betrachtet und dann die Lösungen zusammensetzt):

Satz 5.14 (Existenz und Eindeutigkeit von Lösungen von linearen Anfangswertproblemen)**.** *Sei $n \in \mathbb{N}$, sei $I \subset \mathbb{R}$ ein Intervall, seien $A\colon I \longrightarrow \mathbb{R}^{n\times n}$ und $b\colon I \longrightarrow \mathbb{R}^n$ stetig und seien $t_0 \in I$, $x_0 \in \mathbb{R}^n$. Dann besitzt das Anfangswertproblem*

$$\begin{array}{lrcl} \textit{Gesucht:} & y \in C^1\bigl(\langle\mathbb{R}^1\rangle, \langle\mathbb{R}^n\rangle\bigr) & & \textit{mit} \\ & y' & = & A\cdot y + b \\ & y(t_0) & = & x_0 \end{array}$$

genau eine Lösung in $C^1(I, \mathbb{R}^n)$.

Satz 5.15 (Lösungsräume linearer Differentialgleichungen)**.** *Sei $n \in \mathbb{N}$, sei $I \subset \mathbb{R}$ ein nicht-leeres Intervall und seien $A\colon I \longrightarrow \mathbb{R}^{n\times n}$ und $b\colon I \longrightarrow \mathbb{R}^n$ stetig. Sei $L \subset C^1(I, \mathbb{R}^n)$ die Menge der Lösungen in $C^1(I, \mathbb{R}^n)$ des homogenen Systems*

$$\begin{array}{lrcl} \textit{Gesucht:} & y \in C^1\bigl(\langle\mathbb{R}^1\rangle, \langle\mathbb{R}^n\rangle\bigr) & & \textit{mit} \\ & y' & = & A\cdot y \end{array}$$

Dann gilt:

1. *Es ist L ein reeller Untervektorraum von $C^1(I, \mathbb{R}^n)$.*
2. *Es ist $\dim_{\mathbb{R}} L = n$. Genauer gilt: Sind $y_{(1)}, \dots, y_{(n)} \in L$ und ist $t_0 \in I$, so ist $(y_{(1)}, \dots, y_{(n)})$ genau dann eine Basis von L, wenn $\bigl(y_{(1)}(t_0), \dots, y_{(n)}(t_0)\bigr)$ eine Basis von $\mathbb{R}^n$ ist.*
3. *Ist $\widetilde{y} \in C^1(I, \mathbb{R}^n)$ eine Lösung des Systems*

$$\begin{aligned} \textit{Gesucht:}\quad & y \in C^1\bigl(\langle\mathbb{R}^1\rangle, \langle\mathbb{R}^n\rangle\bigr) \quad \textit{mit} \\ & y' \;=\; A \cdot y + b \end{aligned}$$

(d.h. $\widetilde{y}$ ist eine sogenannte spezielle Lösung*), so ist*

$$\{y + \widetilde{y} \mid y \in L\}$$

die Menge der Lösungen in $C^1(I, \mathbb{R}^n)$ des obigen inhomogenen Systems.

Es bietet sich an, Basen des Lösungsraums zu matrixwertigen Funktionen zusammenzufassen:

Definition 5.16 (Fundamentalmatrix, Wronski-Determinante)**.** Sei $n \in \mathbb{N}$, sei $I \subset \mathbb{R}$ ein Intervall und sei $A \colon I \longrightarrow \mathbb{R}^{n \times n}$ stetig. Sei $L \subset C^1(I, \mathbb{R}^n)$ der Lösungsraum von

$$\begin{aligned} \text{Gesucht:}\quad & y \in C^1\bigl(\langle\mathbb{R}^1\rangle, \langle\mathbb{R}^n\rangle\bigr) \quad \text{mit} \\ & y' \;=\; A \cdot y \end{aligned}$$

in $C^1(I, \mathbb{R}^n)$.
- Basen von L heißen auch *Fundamentalsysteme* der obigen Differentialgleichung.
- Ist $(y_{(1)}, \dots, y_{(n)})$ eine Basis von L, so nennt man

$$\begin{aligned} I &\longrightarrow \mathbb{R}^{n \times n} \\ t &\longmapsto \bigl(y_{(1)}(t), \dots, y_{(n)}(t)\bigr) \end{aligned}$$

  (d.h., die Spalten der Matrix rechts sind durch die Lösungen $y_{(1)}, \dots, y_{(n)}$ gegeben) eine *Fundamentalmatrix* des obigen Systems.
- Ist $Y \colon I \longrightarrow \mathbb{R}^{n \times n}$ eine Lösungsmatrix des obigen Systems (d.h. die Spaltenfunktionen von Y liegen in L), so heißt

$$\begin{aligned} \det Y \colon I &\longrightarrow \mathbb{R} \\ t &\longmapsto \det\bigl(Y(t)\bigr) \end{aligned}$$

  Wronski-Determinante von Y.

Bemerkung 5.17. Nach Satz 5.15 kann man also an der Wronski-Determinante ablesen, ob eine Lösungsmatrix eine Fundamentalmatrix ist oder nicht.

Kennt man die Lösungen eines homogenen linearen Differentialgleichungssystems, so kann man daraus spezielle Lösungen der zugehörigen inhomogenen Systeme „berechnen“:

Satz 5.18 (Variation der Konstanten)**.** *Sei $n \in \mathbb{N}$, sei $I \subset \mathbb{R}$ ein Intervall und seien $A\colon I \longrightarrow \mathbb{R}^{n\times n}$ und $b\colon I \longrightarrow \mathbb{R}^n$ stetig. Sei $t_0 \in I$ und sei $Y\colon I \longrightarrow \mathbb{R}^{n\times n}$ eine Fundamentalmatrix von*

$$\begin{aligned} \textit{Gesucht:}\quad & y \in C^1\bigl(\langle\mathbb{R}^1\rangle, \langle\mathbb{R}^n\rangle\bigr) \quad \textit{mit} \\ & y' \;=\; A \cdot y. \end{aligned}$$

Dann ist

$$\begin{aligned} I &\longrightarrow \mathbb{R}^n \\ t &\longmapsto Y(t) \cdot \int_{t_0}^{t} \bigl(Y(s)\bigr)^{-1} \cdot b(s)\, ds \end{aligned}$$

wohldefiniert und eine Lösung des Systems

$$\begin{aligned} \textit{Gesucht:}\quad & y \in C^1\bigl(\langle\mathbb{R}^1\rangle, \langle\mathbb{R}^n\rangle\bigr) \quad \textit{mit} \\ & y' \;=\; A \cdot y + b. \end{aligned}$$

Im folgenden beschreiben wir, wie man lineare gewöhnliche Differentialgleichungssysteme mit konstanten Koeffizienten lösen kann.

Dazu betrachtet man die matrixwertige Exponentialfunktion:

Proposition 5.19 (matrixwertige Exponentialfunktion)**.** *Sei $n \in \mathbb{N}$, sei $A \in \mathbb{C}^{n\times n}$. Dann ist*

$$\exp(A) := \sum_{k=0}^{\infty} \frac{1}{k!} \cdot A^k$$

in $\mathbb{C}^{n\times n}$ absolut konvergent. Dabei gilt:

1. *Ist $A \in \mathbb{R}^{n\times n}$, so ist auch $\exp(A) \in \mathbb{R}^{n\times n}$.*
2. *Ist $B \in \mathbb{C}^{n\times n}$ mit $A \cdot B = B \cdot A$, so gilt*

$$\exp(A + B) = \exp(A) \cdot \exp(B).$$

Diese Eigenschaften lassen sich analog zu den entsprechenden Eigenschaften der gewöhnlichen reellen/komplexen Exponentialfunktion beweisen. Ausßerdem folgt analog zum eindimensionalen Fall:

Proposition 5.20 (Fundamentalmatrix via Exponentialfunktion)**.** *Sei $n \in \mathbb{N}$ und sei $A \in \mathbb{R}^{n\times n}$. Dann ist*

$$\begin{aligned} \mathbb{R} &\longrightarrow \mathbb{R}^{n\times n} \\ t &\longmapsto \exp(t \cdot A) = \sum_{k=0}^{\infty} \frac{t^k}{k!} \cdot A^k \end{aligned}$$

stetig differenzierbar und eine Fundamentalmatrix des Systems

$$\begin{aligned} \textit{Gesucht:} \quad & y \in C^1\bigl(\langle\mathbb{R}^1\rangle, \langle\mathbb{R}^n\rangle\bigr) \qquad \textit{mit} \\ & y' \;=\; A \cdot y. \end{aligned}$$

Man berechnet die matrixwertige Exponentialfunktion, indem man die Jordan-Normalform über $\mathbb{C}$ der entsprechenden Matrix betrachtet. Wir beginnen mit einem einfachen Spezialfall:

Bemerkung 5.21 (reell diagonalisierbare Systeme)**.** Sei $n \in \mathbb{N}$ und sei $A \in \mathbb{R}^{n\times n}$ über $\mathbb{R}$ diagonalisierbar. Dann existiert eine Matrix $S \in \mathrm{GL}(n, \mathbb{R})$ und eine Diagonalmatrix $D \in \mathbb{R}^{n\times n}$ mit $A = S^{-1} \cdot D \cdot S$ (d.h. die Spalten von S^{-1} bilden eine Basis aus Eigenvektoren von A von $\mathbb{R}^n$). Seien $\lambda_1, \dots, \lambda_n \in \mathbb{R}$ die Diagonaleinträge von D (d.h. die Eigenwerte von A). Dann gilt

$$\exp(t\cdot A) = S^{-1}\cdot \exp(t\cdot D)\cdot S = S^{-1}\cdot \begin{pmatrix} e^{t\cdot\lambda_1} & & 0 \\ & \ddots & \\ 0 & & e^{t\cdot\lambda_n} \end{pmatrix} \cdot S$$

für alle $t \in \mathbb{R}$. Insbesondere gilt: Ist $(u_1, \dots, u_n)$ eine Basis von $\mathbb{R}^n$ aus Eigenvektoren von A zu den Eigenwerten $\lambda_1, \dots, \lambda_n$, so ist

$$\begin{aligned} \mathbb{R} &\longrightarrow \mathbb{R}^{n\times n} \\ t &\longmapsto \bigl(e^{t\cdot\lambda_1}\cdot u_1, \dots, e^{t\cdot\lambda_n}\cdot u_n\bigr) \end{aligned}$$

eine Fundamentalmatrix des homogenen linearen Differentialgleichungssytems, das durch die Matrix A gegeben ist.

Der Vollständigkeit halber erinnern wir nochmal an die Jordan-Normalform:

Bemerkung 5.22 (Jordan-Normalform)**.** Sei $n \in \mathbb{N}$ und sei $A \in \mathbb{R}^{n\times n}$. Dann existiert eine Matrix $S \in \mathrm{GL}(n, \mathbb{C})$ so dass $S^{-1}\cdot A\cdot S$ in *Jordan-Normalform* über $\mathbb{C}$ ist, d.h.

$$S\cdot A\cdot S^{-1} = \begin{pmatrix} \square & & & 0 \\ & \square & & \\ & & \ddots & \\ 0 & & & \square \end{pmatrix},$$

wobei jedes der Kästchen $\square$ ein *Jordan-Kästchen* ist; d.h. jedes dieser Kästchen $\square$ ist von der Gestalt

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

wobei $\lambda \in \mathbb{C}$ ein Eigenwert von A über $\mathbb{C}$ ist. Dabei gilt:

1. Ist $\lambda \in \mathbb{R}$, so gibt es zu jedem Jordan-Kästchen zum Eigenwert λ eine zugehörige *reelle* Kette von Eigen- bzw. Hauptvektoren von A.
2. Ist $\lambda \in \mathbb{C} \setminus \mathbb{R}$, so gibt es zu jedem Jordan-Kästchen zum Eigenwert λ ein entsprechendes Jordan-Kästchen derselben Größe zum Eigenwert $\overline{\lambda}$; die Konjugierten einer Kette von Eigen- bzw. Hauptvektoren von A zum Eigenwert λ bilden eine Kette von Eigen- bzw. Hauptvektoren von A zum Eigenwert $\overline{\lambda}$.

Bemerkung 5.23 (Reduktion auf Jordan-Kästchen)**.** Wegen der folgenden beiden Eigenschaften der Exponentialfunktion genügt es, die Exponentialfunktion für Jordan-Kästchen zu bestimmen:

1. Für alle $B \in \mathbb{C}^{n \times n}$ und alle $S \in \mathrm{GL}(n, \mathbb{C})$ gilt
$$\exp(S^{-1} \cdot B \cdot S) = S^{-1} \cdot \exp(B) \cdot S.$$
2. Für alle $B \in \mathbb{C}^{m \times m}$ und alle $C \in \mathbb{C}^{k \times k}$ gilt
$$\exp \begin{pmatrix} B & 0 \\ 0 & C \end{pmatrix} = \begin{pmatrix} \exp(B) & 0 \\ 0 & \exp(C) \end{pmatrix}.$$

Proposition 5.24 (Exponentialfunktion von Jordankästchen)**.** *Sei $n \in \mathbb{N}$, sei $\lambda \in \mathbb{C}$ und sei*
$$A = \begin{pmatrix} \lambda & 1 & & 0 \\ & \lambda & \ddots & \\ & & \ddots & 1 \\ 0 & & & \lambda \end{pmatrix}.$$
Dann gilt für alle $t \in \mathbb{R}$, dass
$$\exp(t \cdot A) = e^{t \cdot \lambda} \cdot \begin{pmatrix} 1 & t & \frac{t^2}{2} & \dots & \frac{t^{n-1}}{(n-1)!} \\ 0 & 1 & t & \ddots & \vdots \\ 0 & 0 & 1 & \ddots & \frac{t^2}{2} \\ \vdots & & & \ddots & t \\ 0 & \dots & & 0 & 1 \end{pmatrix}.$$

Korollar 5.25 (Fundamentalmatrizen für reelle Jordankästchen-Systeme)**.** *Sei $n \in \mathbb{N}$, sei $\lambda \in \mathbb{R}$ und sei $A \in \mathbb{R}^n$ mit Jordan-Normalform*
$$\begin{pmatrix} \lambda & 1 & & 0 \\ & \lambda & \ddots & \\ & & \ddots & 1 \\ 0 & & & \lambda \end{pmatrix}.$$

Ist $(u_1, \ldots, u_n)$ *eine zugehörige Kette von reellen Eigen- bzw. Hauptvektoren von* A*, so ist*

$$\begin{aligned}\mathbb{R} &\longrightarrow \mathbb{R}^{n\times n}\\ t &\longmapsto \left(e^{t\cdot\lambda}\cdot u_1, e^{t\cdot\lambda}\cdot(u_2 + t\cdot u_1), \ldots, e^{t\cdot\lambda}\cdot\sum_{k=0}^{n-1}\frac{t^k}{k!}\cdot u_{n-k}\right)\end{aligned}$$

eine Fundamentalmatrix des durch A *gegebenen homogenen linearen Differentialgleichungssystems.*

Sind die Eigenwerte nicht reell, so betrachtet man den Real- bzw. Imaginärteil der zugehörigen komplexen Eigen- bzw. Hauptvektoren zum entsprechenden Eigenwert (damit ist dann auch der dazu konjugierte Eigenwert bereits abgedeckt). Wir formulieren dies nur im folgenden Spezialfall explizit:

Bemerkung 5.26 (Systeme mit komplexen Eigenwerten)**.** Sei $A \in \mathbb{R}^{2\times 2}$ eine Matrix mit Eigenwert $\lambda \in \mathbb{C} \setminus \mathbb{R}$; dann ist auch $\overline{\lambda}$ ein Eigenwert von A. Sei $u \in \mathbb{C}^2$ ein Eigenvektor von A zum Eigenwert λ; dann ist $\overline{u}$ ein Eigenvektor von A zum Eigenwert $\overline{\lambda}$. Wir schreiben

$$\lambda = \alpha + i\cdot\beta \quad \text{und} \quad u = v + i\cdot w$$

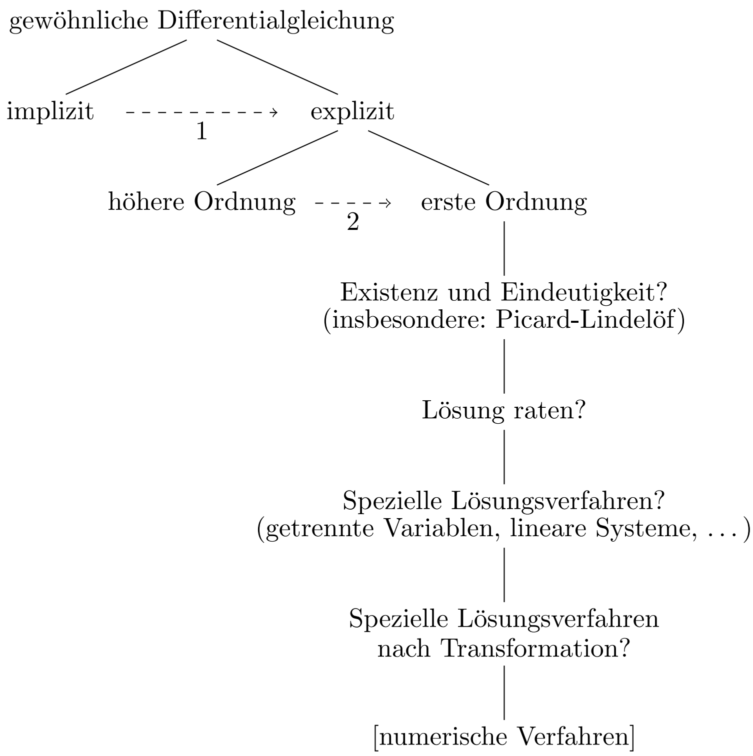
mit $\alpha, \beta \in \mathbb{R}$ und $v, w \in \mathbb{R}^2$. Dann ist

$$\begin{aligned}\mathbb{R} &\longrightarrow \mathbb{R}^{2\times 2}\\ t &\longmapsto \big(\mathrm{Re}(e^{t\cdot\lambda}\cdot u), \mathrm{Im}(e^{t\cdot\lambda}\cdot u)\big)\\ &= \Big(e^{t\cdot\alpha}\cdot\big((\cos(\beta\cdot t))\cdot v - (\sin(\beta\cdot t))\cdot w\big), e^{t\cdot\alpha}\cdot\big((\sin(\beta\cdot t))\cdot v + (\cos(\beta\cdot t))\cdot w\big)\Big)\end{aligned}$$

eine Fundamenatlmatrix für die lineare homogene Differentialgleichung, die durch A gegeben ist.

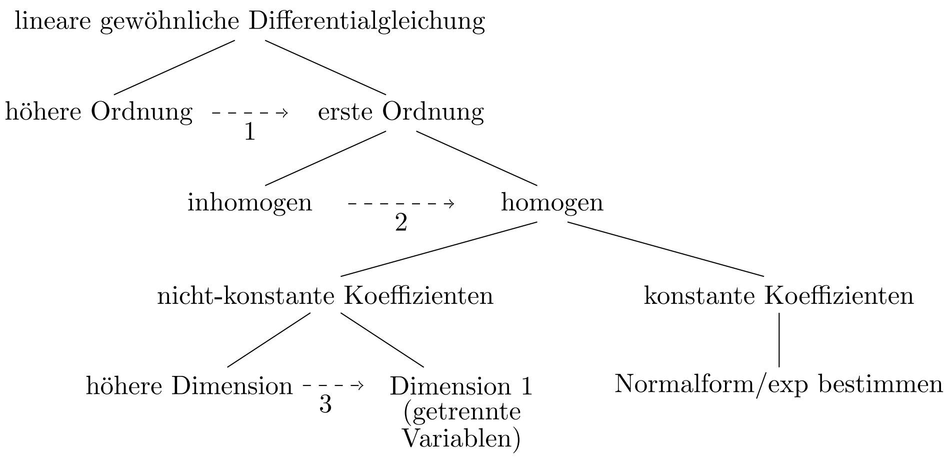
Zusammenfassung der bisherigen Lösungsstrategien

In Abbildung 1 und Abbildung 2 sind die bisher behandelten Lösungsstrategien zusammengestellt.



1 lokales Auflösen (falls hinreichend regulär)
2 Reduktion (Variablen für die höheren Ableitungen einführen)

Abbildung 1: Wie löst man gewöhnliche Differentialgleichungen?



1 Reduktion (Variablen für die höheren Ableitungen einführen)
2 Variation der Konstanten
3 [d'Alembert-Reduktion]

Abbildung 2: Wie löst man lineare gewöhnliche Differentialgleichungen?

Stabilität

Wir beschäftigen uns nun mit der Frage wie Lösungen von Anfangswertproblemen von den Anfangswerten abhängen und insbesondere mit der Stabilität von stationären Punkten. Dadurch werden insbesondere auch die verschiedenen Typen von stabilen Lagen physikalischer/wirtschaftlicher Systeme modelliert.

Definition 5.27 (stabil, instabil, attraktiv, asymptotisch stabil)**.** Sei $n \in \mathbb{N}$, sei $t_0 \in \mathbb{R}$, sei $I \subset \mathbb{R}$ ein Intervall mit $[t_0, \infty) \subset I$ und sei $f\colon I \times \mathbb{R}^n \longrightarrow \mathbb{R}^n$ stetig und lokal Lipschitz in den letzten n Variablen. Wir betrachten die Differentialgleichung

$$\begin{aligned} \text{Gesucht:} \quad & y \in C^1\bigl(\langle\mathbb{R}^1\rangle, \langle\mathbb{R}^n\rangle\bigr) \quad \text{mit} \\ y' \quad &= \quad f(\,\cdot\,, y). \end{aligned}$$

- Eine Lösung $y \in C^1([t_0,\infty),\mathbb{R}^n)$ dieser Differentialgleichung heißt *stabil (bezüglich t_0)*, wenn: Zu jedem $\varepsilon \in \mathbb{R}_{>0}$ gibt es ein $\delta \in \mathbb{R}_{>0}$ mit folgender Eigenschaft: Ist $\widetilde{x}_0 \in B(y(t_0),\delta)$, so existiert die Lösung $\widetilde{y}$ dieser Differentialgleichung mit $\widetilde{y}(t_0) = \widetilde{x}_0$ auf $[t_0,\infty)$ und

$$\forall_{t\in[t_0,\infty)} \quad \bigl\|y(t) - \widetilde{y}(t)\bigr\|_2 < \varepsilon.$$

- Eine Lösung in $C^1([t_0,\infty),\mathbb{R}^n)$ heißt *instabil (bezüglich t_0)*, wenn sie nicht stabil ist.
- Eine Lösung $y \in C^1([t_0,\infty),\mathbb{R}^n)$ heißt *attraktiv (bezüglich t_0)*, wenn: Es gibt ein $\delta \in \mathbb{R}_{>0}$ mit: Ist $\widetilde{x}_0 \in B(y(t_0),\delta)$, so existiert die Lösung $\widetilde{y}$ dieser Differentialgleichung mit $\widetilde{y}(t_0) = \widetilde{x}_0$ und

$$\lim_{t\to\infty} \bigl\|y(t) - \widetilde{y}(t)\bigr\|_2 = 0.$$

- Eine Lösung heißt *asymptotisch stabil (bezüglich t_0)*, wenn sie stabil und asymptotisch stabil bezüglich t_0 ist.

Die Stabilität der trivialen Lösung 0 bezüglich 0 von linearen Differentialgleichungssystemen kann man im wesentlichen anhand der Jordan-Normalform charakterisieren.

Satz 5.28 (Stabilität in linearen Systemen (ohne Beweis))**.** *Sei $n \in \mathbb{N}$ und $A \in \mathbb{R}^{n\times n}$. Wir betrachten das lineare Differentialgleichungssystem*

$$\begin{aligned} \textit{Gesucht:} \quad & y \in C^1\bigl(\langle\mathbb{R}^1\rangle, \langle\mathbb{R}^n\rangle\bigr) \quad \textit{mit} \\ y' \quad &= \quad A \cdot y. \end{aligned}$$

Dann gilt:

1. *Es ist $0 \in C^1([0,\infty),\mathbb{R}^n)$ genau dann stabil (bezüglich 0) für das obige System, wenn gilt: Für alle Eigenwerte $\lambda \in \mathbb{C}$ von A über $\mathbb{C}$ ist $\operatorname{Re}\lambda \le 0$ und für alle Eigenwerte $\lambda \in \mathbb{C}$ von A über $\mathbb{C}$ mit $\operatorname{Re}\lambda = 0$ ist λ für A halbeinfach (d.h. die geometrische und die algebraische Vielfachheit von λ stimmen überein).*

2. *Es ist* $0 \in C^1([0,\infty),\mathbb{R}^n)$ *genau dann asymptotisch stabil (bezüglich* 0*) für das obige System, wenn alle Eigenwerte* $\lambda \in \mathbb{C}$ *von* A *über* $\mathbb{C}$ *erfüllen, dass* $\operatorname{Re}\lambda < 0$ *ist.*

Dies überträgt sich durch lineare Approximation analog auch auf allgemeinere autonome Systeme:

Satz 5.29 (Prinzip der linearisierten Stabilität (ohne Beweis))**.** *Sei* $n \in \mathbb{N}$*, sei* $f \in C^1(\mathbb{R}^n,\mathbb{R}^n)$ *und sei* $x_0 \in \mathbb{R}^n$ *mit* $f(x_0) = 0$ *(d.h.* x_0 *ist ein stationärer Punkt der entsprechenden Differentialgleichung). Sei* $A := Jf(x) \in \mathbb{R}^{n\times n}$*. Wir betrachten die autonome Differentialgleichung*

$$\begin{aligned} \textit{Gesucht:} \quad & y \in C^1\bigl(\langle\mathbb{R}^1\rangle,\langle\mathbb{R}^n\rangle\bigr) \quad \textit{mit} \\ & y' \;=\; f\circ y. \end{aligned}$$

Dann gilt:

1. *Gibt es einen Eigenwert* $\lambda \in \mathbb{C}$ *von* A *über* $\mathbb{C}$ *mit* $\operatorname{Re}\lambda > 0$*, so ist die konstante Lösung* $x_0 \in C^1([0,\infty),\mathbb{R}^n)$ *instabil (bezüglich* 0*) für die obige Differentialgleichung.*
2. *Gilt* $\operatorname{Re}\lambda < 0$ *für alle Eigenwerte* $\lambda \in \mathbb{C}$ *von* A *über* $\mathbb{C}$*, so ist die konstante Lösung* $x_0 \in C^1([0,\infty),\mathbb{R}^n)$ *asymptotisch stabil (bezüglich* 0*) für die obige Differentialgleichung.*

Ein weiteres Hilfsmittel bei Stabilitätsproblemen ist die sogenannte *direkte Methode* mithilfe von *Lyapunov-Funktionen*:

Definition 5.30 ((strikte) Lyapunov-Funktion)**.** Sei $n \in \mathbb{N}$ und sei $f\colon \mathbb{R}^n \longrightarrow \mathbb{R}^n$ lokal Lipschitz. Wir betrachten die autonome Differentialgleichung

$$\begin{aligned} \text{Gesucht:} \quad & y \in C^1\bigl(\langle\mathbb{R}^1\rangle,\langle\mathbb{R}^n\rangle\bigr) \quad \text{mit} \\ & y' \;=\; f\circ y. \end{aligned}$$

- Eine *Lyapunov-Funktion* für diese Differentialgleichung ist eine Funktion $\varphi \in C(\mathbb{R}^n,\mathbb{R})$ mit folgender Eigenschaft: Ist $y \in C^1(I,\mathbb{R}^n)$ eine Lösung dieser Differentialgleichung, so ist $\varphi\circ y\colon I \longrightarrow \mathbb{R}$ monoton fallend.
- Eine *strikte Lyapunov-Funktion* für diese Differentialgleichung ist eine Funktion $\varphi \in C(\mathbb{R}^n,\mathbb{R})$ mit folgender Eigenschaft: Ist $y \in C^1(I,\mathbb{R}^n)$ eine nichtkonstante Lösung dieser Differentialgleichung, so ist $\varphi\circ y\colon I \longrightarrow \mathbb{R}$ streng monoton fallend.

Satz 5.31 (direkte Methode (ohne Beweis))**.** *Sei* $n \in \mathbb{N}$ *und sei* $f\colon \mathbb{R}^n \longrightarrow \mathbb{R}^n$ *lokal Lipschitz. Wir betrachten die autonome Differentialgleichung*

$$\begin{aligned} \textit{Gesucht:} \quad & y \in C^1\bigl(\langle\mathbb{R}^1\rangle,\langle\mathbb{R}^n\rangle\bigr) \quad \textit{mit} \\ & y' \;=\; f\circ y. \end{aligned}$$

Sei $x_0 \in \mathbb{R}^n$ ein stationärer Punkt dieser Differentialgleichung, d.h. $f(x_0) = 0$. Dann gilt:

1. *Ist $\varphi \in C(\mathbb{R}^n, \mathbb{R})$ eine Lyapunov-Funktion dieser Differentialgleichung und x_0 eine strikte lokale Minimalstelle der Funktion φ, so ist die konstante Lösung $x_0 \in C^1([0,\infty), \mathbb{R}^n)$ stabil bezüglich 0 für die obige Differentialgleichung.*
2. *ist $\varphi \in C(\mathbb{R}^n, \mathbb{R})$ eine strikte Lyapunov-Funktion für die obige Differentialgleichung, ist x_0 ein isolierter Punkt von $f^{-1}(\{0\})$ und ist x_0 eine strikte lokale Minimalstelle von φ, so ist die konstante Lösung $x_0 \in C^1([0,\infty), \mathbb{R}^n)$ asymptotisch stabil bezüglich 0 für die obige Differentialgleichung.*