Differential- und Integralrechnung II

Ingo Witt

SS 2008

Inhaltsverzeichnis

Kapitel 7

Topologische Räume

7.1 Grundlegende Begriffe

Bei topologischen Betrachtungen geht es um die Lage von Punkten zueinander im Raum, ohne dass dabei Größenverhältnisse (Längen, Winkel,...) eine Rolle spielen. Das heißt insbesondere, dass topologische Eigenschaften sich bei „stetigen Deformationen“ nicht ändern (Gummibandgeometrie).

Definition 7.1. Ein *topologischer Raum* $(X,\mathcal{U})$ ist eine Menge X und eine Menge $\mathcal{U}$ von Teilmengen von X mit folgenden Eigenschaften:

(i) $\emptyset, X \in \mathcal{U}$,

(ii) $\mathcal{U}$ ist abgeschlossen unter endlichen Durchschnitten,

(iii) $\mathcal{U}$ ist abgeschlossen unter beliebigen Vereinigungen.

Im Allgemeinen schreiben wir X statt $(X,\mathcal{U})$, wenn aus dem Kontext klar ist, was die Menge $\mathcal{U}$ ist. Die Elemente von $\mathcal{U}$ heißen die *offenen Mengen* von X. Die Komplemente in X der Elemente von $\mathcal{U}$ heißen die *abgeschlossenen Mengen* von X. $\mathcal{U}$ heißt auch eine *Topologie* auf X.

Lemma 7.2. *Die Mengen $\emptyset, X$ sind abgeschlossen. Endliche Vereinigungen und beliebige Durchschnitte abgeschlossener Mengen von X sind abgeschlossen.*

Beispiel 7.3.

1. *Diskrete Topologie*: $\mathcal{U}$ besteht aus allen Teilmengen von X. *Triviale Topologie*: $\mathcal{U} = \{\emptyset, X\}$.
2. $X = \{a, b\}$, $\mathcal{U} = \big\{\emptyset, \{a\}, \{a, b\}\big\}$.
3. $X = \mathbb{R}$, $\mathcal{U}$ besteht aus allen offenen Teilmengen von $\mathbb{R}$ (wie im vergangenen Semester eingeführt). (Vergleiche Satz 2.56.)
4. Teilmengen von $\mathbb{R}^n$ bezüglich der induzierten Topologie (siehe unten).
5. Metrische Räume (siehe das nächste Kapitel).

Lemma 7.4. *Sei $(X,\mathcal{U})$ ein topologischer Raum, $A \subseteq X$. Dann gibt es eine größte offene Menge A°, die in A enthalten ist, und eine kleinste abgeschlossene Menge $\bar{A}$, die A enthält.*

Beweis. $A^\circ = \bigcup\limits_{\substack{U \text{ offen}, \\ U \subseteq A}} U, \qquad \bar{A} = \bigcap\limits_{\substack{F \text{ abgeschlossen}, \\ F \supseteq A}} F.$ □

Definition 7.5. A° heißt das *Innere*, $\bar{A}$ der *Abschluss* von A. Darüber hinaus heißt $\partial A = \bar{A} \setminus A^\circ$ der *Rand* von A.

Lemma 7.6. (i) *Eine Menge $A \subseteq X$ ist genau dann abgeschlossen, wenn A alle ihre Randpunkte enthält.*

(ii) *Sei $x \in X$. Dann ist x genau dann ein Randpunkt von A, wenn für alle $U \in \mathcal{U}$ mit $x \in U$*

$$U \cap A \neq \emptyset, \;\; U \cap (X \setminus A) \neq \emptyset$$

gilt.

Beweis. (i) A abgeschlossen heißt gerade $A = \bar{A}$ $(= A \cup \partial A)$.

(ii) Wir haben eine disjunkte Vereinigung

$$X = A^\circ \cup (X \setminus A)^\circ \cup \partial A$$

(weshalb?).

□

7.2 Stetigkeit und Homöomorphie

Seien $(X,\mathcal{U})$, $(Y,\mathcal{V})$ topologische Räume, $f\colon X \to Y$ eine Abbildung.

Definition 7.7. f heißt *stetig* von X nach Y, falls für alle $V \in \mathcal{V}$ die Menge

$$f^{-1}(V) = \left\{x \in X \;\middle|\; f(x) \in V\right\}$$

zu $\mathcal{U}$ gehört (d. h. falls die Urbilder offener Mengen offen sind).

Beispiel 7.8. Sei $I \subseteq \mathbb{R}$ ein Intervall, $f\colon I \to \mathbb{R}$. Dann ist f genau dann im Sinne obiger Definition stetig, wenn f im bisherigen Sinne stetig ist.
Wie in Satz 2.57 beweist man:

Satz 7.9. *Sei $f\colon X \to Y$. Dann sind äquivalent:*

(i) *f ist stetig.*

(ii) *$f^{-1}(G) \subseteq X$ ist abgeschlossen für alle abgeschlossenen Mengen $G \subseteq Y$.*

(iii) *Für alle $A \subseteq X$ gilt: $f(\bar{A}) \subseteq \overline{f(A)}$.*

Satz 7.10. *Die Komposition stetiger Abbildungen ist stetig.*

Beweis. Seien $f\colon X \to Y$ und $g\colon Y \to Z$ stetig. Dann ist $g \circ f\colon X \to Z$ stetig, da für alle offenen Mengen $W \subseteq Z$

$$(g \circ f)^{-1}(W) = f^{-1}\big(g^{-1}(W)\big)$$

ist offen in X gilt. □

Die folgende Definition führt den Begriff der *topologischen Äquivalenz* ein:

Definition 7.11. Zwei topologische Räume X, Y heißen *homöomorph* (topologisch äquivalent, topologisch nicht unterscheidbar), falls es eine stetige Bijektion $f\colon X \to Y$ gibt, so dass $f^{-1}\colon Y \to X$ ebenfalls stetig ist. Ein derartiges f (im Allgemeinen nicht eindeutig bestimmt) heißt eine *Homöomorphie*.

Bemerkung. Homöomorphie ist eine Äquivalenzrelation in dem Sinne, dass

(i) $\mathrm{id}\colon X \to X$ ist eine Homöomorphie für alle X (Reflexivität).

(ii) Ist $f\colon X \to Y$ eine Homöomorphie, so auch $f^{-1}\colon Y \to X$ (Symmetrie).

(iii) Sind $f\colon X \to Y$, $g\colon Y \to Z$ Homöomorphien, so auch $g \circ f\colon X \to Z$ (Transitivität).

Anmerkung: Die hier fehlenden Zeichnungen (die Stellen, an denen ... steht), werden nachgeliefert.

Beispiel 7.12. Sind die folgenden topologischen Räume homöomorph?

(a) ...

(b) ...

(c) ...

(a) Nein, ... ist *kompakt*, ... ist es nicht.

(b) Nein. Wenn wir einen Punkt entfernen, dann zerfällt ... in zwei *zusammenhängende* Mengen (...), ... jedoch nicht (...).

(c) Ja. ... habe die Länge l_1, der verschlungene Knoten die Länge l_2. Wir wählen auf beiden Kurven einen beliebigen Punkt P bzw. Q und einen Umlaufsinn. Der Homöomorphismus f ordnet dann dem Punkt auf ..., der sich in Entfernung θl_1 für $\theta \in [0, 1)$ von P befindet, den Punkt auf ... zu, der sich in Entfernung θl_2 von Q befindet.

7.3 Konstruktion topologischer Räume

7.3.1 Die Unterraumtopologie

Sei $(X,\mathcal{U})$ ein topologischer Raum, $A \subseteq X$. Wir setzen

$$\mathcal{U}_A := \big\{U \cap A \;\big|\; U \in \mathcal{U}\big\}.$$

Lemma 7.13. *$\mathcal{U}_A$ ist eine Topologie auf A.*

Beweis. Selbst. □

Definition 7.14. $\mathcal{U}_A$ heißt die von $\mathcal{U}$ auf A induzierte Unterraumtopologie.

7.3.2 Die koinduzierte Topologie

Sei X eine Menge, $\mathcal{U}$, $\mathcal{V}$ seien zwei Topologien auf X. Dann heißt $\mathcal{U}$ *feiner* oder *stärker* als $\mathcal{V}$ (bzw. $\mathcal{V}$ heißt *gröber* oder *schwächer* als $\mathcal{U}$), falls jede Menge in $\mathcal{V}$ auch zu $\mathcal{U}$ gehört (d. h. wenn $\mathcal{U}$ mehr offene Mengen als $\mathcal{V}$ hat). Die diskrete Topologie ist die feinste Topologie auf X, die triviale Topologie die gröbste.

Bemerkung. Für jede Teilmenge $\mathcal{A}$ der Potenzmenge von X gibt es eine schwächste Topologie auf X, die $\mathcal{A}$ enthält (nämlich der Durchschnitt aller Topologien, die $\mathcal{A}$ enthalten). Diese heißt die *von $\mathcal{A}$ erzeugte Topologie* auf X.

Definition 7.15. Sei $(X,\mathcal{U})$ ein topologischer Raum, Y eine Menge und $f\colon X \to Y$ eine Abbildung. Dann heißt die stärkste Topologie $\mathcal{V}_f$ auf Y, bezüglich der f stetig ist, die von f auf Y *koinduzierte Topologie*.

Lemma 7.16.
$$\mathcal{V}_f = \bigl\{V \subseteq Y \bigm| f^{-1}(V) \text{ ist offen in } X\bigr\}.$$

Beweis. Es genügt zu verifizieren, dass $\mathcal{V}_f$ in der Tat eine Topologie ist. □

Proposition 7.17. *X, Y, Z seien topologische Räume, wobei Y die durch $f\colon X \to Y$ koinduzierte Topologie trage. Dann ist $g\colon Y \to Z$ genau dann stetig, wenn $g \circ f\colon X \to Z$ stetig ist.*

Beweis. $(\Rightarrow)$ Offensichtlich.

$(\Leftarrow)$ Sei $g \circ f\colon X \to Z$ stetig. Für $W \subseteq Z$ offen ist $f^{-1}\bigl(g^{-1}(W)\bigr)$ offen in X, d. h. $g^{-1}(W)$ ist offen in Y. Somit ist $g\colon Y \to Z$ stetig.

□

7.3.3 Produkträume

Sei X_i, $i \in I$, eine Familie topologischer Räume.

Definition 7.18. Die *Produkttopologie* auf $X = \prod_{i\in I} X_i$ ist die von allen Mengen der Form $\prod_{i\in I} U_i$ erzeugte Topologie, wobei $U_i \subseteq X_i$ für alle i offen ist und $U_i = X_i$ *für fast alle* (d. h. mit Ausnahme von höchstens endlich vielen) i gilt.

Die Produkttopologie auf $X = \prod_{i\in I} X_i$ ist offenbar die schwächste Topologie auf X, bezüglich der alle kanonischen Projektionen $\pi_j\colon X \to X_j$ stetig sind.

Satz 7.19. *Sei Z ein topologischer Raum und $f\colon Z \to \prod_{i\in I} X_i$ eine Abbildung. Dann ist f genau dann stetig, wenn alle Abbildungen $\pi_j \circ f\colon Z \to X_j$ stetig sind.*

Beweis. ($\Rightarrow$) Klar.

($\Leftarrow$) Sei $U_i \subseteq X$ für $i \in I$ offen, $U_i = X_i$ für fast alle i. Dann ist

$$f^{-1}\left(\prod_{i\in I} U_i\right) = \bigcap_{i\in I} (\pi_i \circ f)^{-1}(U_i)$$

offen in Z.

□

Beispiel 7.20. (a) $\mathbb{R}^n$ trägt die Topologie des Produkts $\underbrace{\mathbb{R} \times \mathbb{R} \times \cdots \times \mathbb{R}}_{n\text{–mal}}$.

(b) Der Zylinder trägt die Topologie des Produkts $I \times \mathbb{S}^1$, $I = [0,1]$, $\mathbb{S}^1$ ist die Kreislinie.

7.4 Zusammenhängende topologische Räume

7.4.1 Definition und erste Eigenschaften

Definition 7.21. Ein topologischer Raum X heißt *zusammenhängend*, falls er *nicht* in der Form $X = U \cup V$ mit $U, V \subseteq X$ offen, nichtleer und disjunkt geschrieben werden kann.

Beispiel 7.22. (a) Das Intervall $I = \ldots$, die Kreislinie $\mathcal{S}^1 = \ldots$ und der Zylinder ... sind zusammenhängend.

(b) Die disjunkte Vereinigung ... zweier Intervalle ist es nicht.

Allgemeiner gilt, dass die zusammenhängenden Teilmengen von $\mathbb{R}$ gerade die leere Menge sowie die Intervalle sind.

Lemma 7.23. *Die folgenden Bedingungen sind äquivalent:*

(i) *X ist zusammenhängend.*

(ii) *Die einzigen Teilmengen von X, die sowohl offen als auch abgeschlossen sind, sind $\emptyset$ und X.*

(iii) *Jede stetige Abbildung $f\colon X \to Y$, wobei Y die diskrete Topologie trägt, ist konstant.*

Beweis. (i) $\Rightarrow$ (ii). Sei $U \subseteq X$ offen und abgeschlossen. Dann ist $X \setminus U$ ebenfalls offen und abgeschlossen. Wegen $X = U \cup (X \setminus U)$, $U \cap (X \setminus U) = \emptyset$ gilt $U = \emptyset$ oder $X \setminus U = \emptyset$, d. h. $U = X$.
(ii) $\Rightarrow$ (iii). Sei $f\colon X \to (Y, \mathcal{U}_{\text{diskret}})$ stetig. Sei $y^* = f(x^*)$ für ein $x^* \in X$. Dann ist $f^{-1}(y^*)$ offen, abgeschlossen und nichtleer. Also $f^{-1}(y^*) = X$ und $f(x) = y^*$ für alle $x \in X$, d. h. f ist konstant.

(iii) $\Rightarrow$ (i). Sei $X = U \cup V$ mit $U \cap V = \emptyset$. Dann ist die Funktion $f\colon X \to (\{0,1\}, \mathcal{U}_{\text{diskret}})$ mit

$$f(x) = \begin{cases} 0, & x \in U, \\ 1, & x \in V, \end{cases}$$

stetig, also konstant. Damit ist $U = \emptyset$ oder $V = \emptyset$. □

Bemerkung. Der Beweis zeigt, dass wir in (iii) ein allgemeines Y durch die zweielementige Menge $\{0,1\}$ ersetzen können.

Satz 7.24 (Vgl. Satz 2.49)**.** *Stetige Bilder zusammenhängender Mengen sind zusammenhängend.*

Beweis. Sei $f\colon X \to Y$ stetig. Weiterhin seien $V_0, V_1 \in Y$ zwei offene Mengen mit $f(X) \subseteq V_0 \cup V_1$, $f(X) \cap V_0 \cap V_1 = \emptyset$.
Sei $U_0 = f^{-1}(V_0)$, $U_1 = f^{-1}(V_1)$. Dann ist $X = U_0 \cup U_1$, $U_0 \cap U_1 = \emptyset$, U_0, U_1 sind offen. Es folgt $U_0 = \emptyset$ oder $U_1 = \emptyset$, also $f(X) \cap V_0 = \emptyset$ oder $f(X) \cap V_1 = \emptyset$. Damit ist $f(X)$ zusammenhängend. □

Satz 7.25. *Sei X_i zusammenhängend für alle i. Dann ist $\prod\limits_{i \in I} X_i$ zusammenhängend.*

7.4.2 Zusammenhangskomponenten

Satz 7.26. *Sei X ein topologischer Raum. Dann ist $X = \bigsqcup\limits_i X_i$ disjunkte Vereinigung seiner (eindeutig bestimmten) Zusammenhangskomponenten X_i (d. h. jedes X_i ist zusammenhängend und zudem maximal mit dieser Eigenschaft).*

Beweis. Wir führen eine Äquivalenzrelation $\sim$ auf X ein: $x \sim y$, falls es eine zusammenhängende Teilmenge von X gibt, die x und y enthält. Um zu sehen, dass diese Relation transitiv ist, zeigen wir:

Lemma 7.27. *Seien $U, V \subseteq X$ zusammenhängend mit $U \cap V \neq \emptyset$. Dann ist $U \cup V$ zusammenhängend.*

Beweis. Sei $f\colon U \cup V \to \{0,1\}$ stetig, wobei $\{0,1\}$ die diskrete Topologie trägt. Dann ist $f\big|_U\colon U \to \{0,1\}$ stetig, also konstant, und auch $f\big|_V\colon V \to \{0,1\}$ ist stetig, also ebenfalls konstant. Wegen $U \cap V \neq \emptyset$ ist die Konstante dieselbe, also ist $f\colon U \cup V \to \{0,1\}$ konstant. □

Die Äquivalenzklassen von X bezüglich $\sim$ sind gerade die Zusammenhangskomponenten X_i von X. Offenbar gilt $X = \bigsqcup X_i$ und die X_i sind paarweise disjunkt. Weiterhin ist jedes X_i zusammenhängend (Beweis?), während aus Y zusammenhängend und $X_i \subseteq Y$ folgt, dass $X_i = Y$ gilt (Beweis?). □

Bemerkung. Die Anzahl der Zusammenhangskomponenten eines topologischen Raumes X ist offenbar eine *topologische Invariante* (d. h. diese Anzahl ist die gleiche für alle topologischen Räume, die homöomorph zu X sind).
Insbesondere ist ein nichtleerer topologischer Raum genau dann zusammenhängend, wenn diese Anzahl gleich 1 ist. (Aber nicht alle nichtleeren zusammenhängenden topologischen Räume sind homöomorph, z. B. .....!)

Beispiel 7.28. $\mathbb{Q} \subset \mathbb{R}$ mit der induzierten Metrik.

7.4.3 Wegzusammenhängende Räume

Definition 7.29. Eine (stetige) *Kurve* in X ist eine stetige Abbildung $\gamma\colon [0,1] \to X$. $\gamma(0)$ heißt der Anfangs-, $\gamma(1)$ der Endpunkt der Kurve γ.

Definition 7.30. Ein topologischer Raum heißt *wegzusammenhängend*, wenn es für je zwei Punkte x, $y \in X$ eine Kurve γ gibt, die diese beiden Punkte verbindet (d. h. $\gamma(0) = x$, $\gamma(1) = y$).

Beispiel 7.31. Die wegzusammenhängenden Teilmengen von $\mathbb{R}$ sind die leere Menge und die Intervalle.

Satz 7.32. *Das stetige Bild eines wegzusammenhängenden Raumes ist wegzusammenhängend.*

Beweis. Sei $f\colon X \to Y$ stetig, y_0, $y_1 \in f(X)$. Wir wählen x_0, $x_1 \in X$ mit $f(x_0) = y_0$, $f(x_1) = y_1$. Dann gibt es eine Kurve $\gamma\colon [0,1] \to X$ mit $\gamma(0) = x_0$, $\gamma(1) = x_1$. $f \circ \gamma\colon [0,1] \to f(X)$ ist dann eine Kurve in $f(X)$, die y_0 und y_1 verbindet. □

Wie im Fall des Zusammenhangs lassen sich jetzt die *Wegzusammenhangskomponenten* eines topologischen Raumes X definieren. Obiger Satz zeigt uns, dass deren Anzahl wiederum eine topologische Invariante ist.

Satz 7.33. *Jeder wegzusammenhängende topologische Raum ist zusammenhängend.*

Beweis. Sei X wegzusammenhängend und $f\colon X \to \{0,1\}$ stetig, wobei $\{0,1\}$ mit der diskreten Topologie versehen ist. Sei o. B. d. A. $f(x^*) = 0$ für ein $x^* \in X$. Für einen beliebigen Punkt $x \in X$ wählen wir einen Weg $\gamma\colon [0,1] \to X$ von x^* nach x, d. h. $\gamma(0) = x^*$, $\gamma(1) = x$. Dann ist $f \circ \gamma\colon [0,1] \to \{0,1\}$ stetig. Da das Intervall $[0,1]$ zusammenhängend ist, ist $f \circ \gamma$ konstant Null, also $f(x) = f(\gamma(1)) = 0$. Damit ist f konstant. □

Etwas überraschend ist sicherlich, dass die Umkehrung nicht gilt:

Beispiel 7.34. Sei $X \subseteq \mathbb{R}^2$,

$$X = \big\{(0,y) \bigm| |y| \le 1\big\} \cup \left\{ \left(x, \sin\frac{1}{x}\right) \,\middle|\, x > 0 \right\},$$

versehen mit der induzierten Topologie. Dann ist X zusammenhängend, aber nicht wegzusammenhängend.

7.5 Kompakte topologische Räume

Sei X ein topologischer Raum. Eine *offene Überdeckung* von X ist eine Familie $\{U_i\}_{i\in I}$ offener Teilmengen von X, so dass $X = \bigcup_{i\in I} U_i$.

Definition 7.35. Ein topologischer Raum X heißt *kompakt*, falls jede offene Überdeckung von X eine endliche Teilüberdeckung besitzt.

Bemerkung. Insbesondere in der algebraischen Geometrie heißen derartige Räume auch *quasikompakt*. Kompaktheit ist dann Quasikompaktheit plus die Hausdorff-Eigenschaft (siehe den nächsten Abschnitt).

Beispiel 7.36. Die kompakten Teilmengen der reellen Geraden $\mathbb{R}$ sind gerade die abgeschlossenen und beschränkten Teilmengen (Überdeckungssatz von Heine–Borel). Selbiges gilt im $\mathbb{R}^n$, $n \geq 2$, ist jedoch in allgemeinen metrischen Räumen im Allgemeinen falsch.

Satz 7.37. *Eine abgeschlossene Teilmenge eines kompakten Raumes ist kompakt.*

Beweis. Sei X kompakt, $A \subseteq X$ abgeschlossen. Seien U_i, $i \in I$, offene Mengen in X mit $A \subseteq \bigcup_{i\in I} U_i$. Dann ist $X \setminus A$ offen in X und $X = (X \setminus A) \cup \bigcup_{i\in I} U_i$. Folglich existiert eine endliche Teilmegne $J \subseteq I$ mit $X = (X \setminus A) \cup \bigcup_{j\in J} U_j$. Dann $A \subseteq \bigcup_{j\in J} U_j$. □

Satz 7.38. *Das stetige Bild eines kompakten Raumes ist kompakt.*

Beweis. Siehe den Beweis mittels Überdeckungskompaktheit zur Aufgabe 6 der Klausur vom 31.01.2008. □

Dieses Resultat zeigt, dass Kompaktheit eine topologische Eigenschaft ist.

Das folgende Ergebnis ist einer der wichtigsten Sätze der allgemeinen Topologie.

Satz 7.39 (Tychonoff)**.** *Das topologische Produkt einer beliebigen Anzahl kompakter topologischer Räume ist kompakt.*

Beweis (*für den Fall endlich vieler Faktoren*)*.* Es genügt, den Beweis für zwei Faktoren zu führen; die allgemeine Behauptung folgt dann per Induktion.
Seien also X, Y kompakte topologische Räume, sei $X \times Y = \bigcup_i (U_i \times V_i)$ für offene Mengen $U_i \subseteq X$, $V_i \subseteq Y$ (wieso genügt es, sich auf offene Mengen der Form $U_i \times V_i$ zu beschränken?).
Für $y \in Y$ ist die Menge $X \times \{y\} \subset X \times Y$ kompakt als homöomorphes Bild des Raumes X (wieso das?). Wir finden also für jedes $y \in Y$ eine endliche Überdeckung

$$X \times \{y\} \subseteq \bigcup_{r=1}^{m_y} U_r^y \times V_r^y,$$

wobei die $U_r^y \times V_r^y$ aus den $U_i \times V_i$ gewählt sind. Wir betrachten nun die Menge $V^y = \bigcap_{r=1}^{m_y} V_r^y$. Diese Menge ist offen in Y, $y \in V^y$. Es gibt also endlich viele $y_1, \ldots, y_n \in Y$, so dass die $V^{y_1}, \ldots, V^{y_n}$ den Raum Y überdecken, d. h. $Y = \bigcup_{s=1}^{n} V^{y_s}$. Dann gilt

$$X \times Y = \bigcup_{s=1}^{n} \bigcup_{r=1}^{m_{y_s}} U_r^{y_s} \times V_r^{y_s}.$$

Um letztere Beziehung einzusehen, wählen wir einen Punkt $(x^*, y^*) \in X \times Y$. Dann gibt es ein $s \in \{1, \ldots, n\}$ mit $y^* \in V^{y_s}$. Wegen $(x^*, y^s) \in X \times \{y^s\}$ gibt es ein $r \in \{1, \ldots, m_{y^s}\}$ mit $(x^*, y^s) \in U_r^{y_s} \times V_r^{y_s}$. Dann jedoch gilt $x^* \in U_r^{y_s}$ und $y^* \in V^{y_s} \subseteq V_r^{y_s}$, also $(x^*, y^*) \in U_r^{y_s} \times V_r^{y_s}$. □

7.6 Hausdorffräume

Nach Zusammenhangs– und Kompaktheitsbegriffen diskutieren wir jetzt eine sogenannte Trennungseigenschaft, nämlich die Hausdorff-Eigenschaft oder T_2.

Definition 7.40. Ein topologischer Raum X heißt *Hausdorffsch*, falls es für alle $x, y \in X$ mit $x \neq y$ offene Mengen U, V mit $x \in U$, $y \in V$ und $U \cap V = \emptyset$ gibt.

Beispiel 7.41. (i) Metrische Räume sind Hausdorffsch, insbesondere sind $\mathbb{R}$, $\mathbb{R}^n$ für $n \geq 2$ Hausdorffsch.

(ii) Unterräume von Hausdorffräumen sind Hausdorffsch.

(iii) Der Raum $X = \{a, b\}$ mit der Topologie $\{\emptyset, \{a\}, \{a, b\}\}$ aus Beispiel 7.3 ist nicht Hausdorffsch. Die einzige offene Menge, die b enthält, ist $\{a, b\}$, und diese Menge enthält auch a.

Satz 7.42. *Einelementige* (*und damit endliche*) *Mengen in Hausdorff-Räumen sind abgeschlossen.*

Beweis. Gemäß Definition ist $X \setminus \{x\}$ für $x \in X$ offen. □

Satz 7.43. *Sei X ein Hausdorff-Raum, $A \subseteq X$ kompakt. Dann ist A abgeschlossen.*

Beweis. Wir zeigen, dass $X \setminus A$ offen ist. Sei also $x \in X \setminus A$. Für jedes $y \in A$ wählen wir offene Mengen U_y, V_y mit $x \in U_y$, $y \in V_y$ und $U_y \cap V_y = \emptyset$. Dann ist $\{V_y\}_{y \in Y}$ eine offene Überdeckung von A. Da A kompakt ist, existiert eine endliche Teilüberdeckung, d. h. es existieren $y_1, \ldots, y_m \in A$ mit $A \subseteq V_{y_1} \cup \cdots \cup V_{y_m}$. Wir setzen $U = U_{y_1} \cap \cdots \cap U_{y_m}$. Dann ist U offen, $x \in U$ und $U \cap A = \emptyset$, d. h. $U \subseteq X \setminus A$. Da $x \in X \setminus A$ beliebig gewählt war, ist $X \setminus A$ offen. □

Folgerung. Sei $f\colon X \to Y$ bijektiv, stetig, X kompakt, Y Hausdorffsch. Dann ist f eine Homöomorphie.

Beweis. Wir müssen zeigen, dass $f^{-1}\colon Y \to X$ stetig ist. Sei also $A \subseteq X$ abgeschlossen. Dann ist $\left(f^{-1}\right)^{-1}(A) = f(A)$ abgeschlossen in Y, da A als abgeschlossene Teilmenge eines kompakten Raumes nach Satz 7.37 kompakt und somit nach dem vorigen Resultat in Y abgeschlossen ist. □

Satz 7.44. *Sei X Hausdorffsch, A, $B \subset X$ kompakte Teilmengen mit $A \cap B = \emptyset$. Dann gibt es offene Mengen U, V mit $A \subseteq U$, $B \subseteq V$ und $U \cap V = \emptyset$.*

Beweis. (i) Wir behandeln zuerst den Fall, dass A eine einelementige Menge ist, d. h. $A = \{x\}$. Für jedes $y \in B$ wählen wir offene Mengen U_y, V_y mit $x \in U_y$, $y \in V_y$ und $U_y \cap V_y = \emptyset$. Dann ist $\left\{V_y\right\}_{y \in B}$ eine offene Überdeckung von B. Wir wählen eine endliche Teilüberdeckung, also $B \subseteq V_{y_1} \cup \cdots \cup V_{y_m}$. Dann sind $U = U_{y_1} \cap \cdots \cap U_{y_m}$, $V = V_{y_1} \cup \cdots \cup V_{y_m}$ offen, $x \in U$, $B \subseteq V$ und $U \cap V = \emptyset$.

(ii) Sei nun A beliebig. Für jedes $x \in A$ wählen wir gemäß (i) offene Mengen U_x, V_x mit $x \in U_x$, $B \subseteq V_x$ und $U_x \cap V_x = \emptyset$. Dann ist $\left\{U_x\right\}_{x \in A}$ eine offene Überdeckung von A. Wir wählen eine endliche Teilüberdeckung, also $A \subseteq U_{x_1} \cup \cdots \cup U_{x_m}$. Dann sind die Mengen $U = U_{x_1} \cup \cdots \cup U_{x_m}$, $V = V_{x_1} \cap \cdots \cap V_{x_m}$ offen, $A \subseteq U$, $B \subseteq V$ und $U \cap V = \emptyset$.

□

Neben dem Satz von Tychonoff ist der Satz von Urysohn einer der wichtigsten Sätze der allgemeinen Topologie.

Satz 7.45 (Urysohn)**.** *Sei X ein kompakter Hausdorffraum, A, $B \subseteq X$ abgeschlossen, $A \cap B = \emptyset$. Dann existiert eine stetige Funktion $F\colon X \to [0,1]$ mit $F(x) = 0$ für alle $x \in A$ und $F(x) = 1$ für alle $x \in B$.*

Beweis. Sei I die Menge aller rationalen Zahlen der Form $k/2^j$, wobei j, $k \in \mathbb{N}$, $1 \le k < 2^j$. Für $r \in I$ konstruieren wir induktiv offene Mengen $U_r \subseteq X$ mit

- $A \subset U_r$, $\overline{U}_r \subset X \setminus B$,
- $\overline{U}_r \subset U_s$ für $r < s$.

Für $j = 1$ wählen wir unter Benutzung des vorigen Satzes eine offene Menge $U_{1/2}$ mit $A \subset U_{1/2} \subset \overline{U}_{1/2} \subset X \setminus B$.
Für $j = 2$ wählen wir wiederum unter Benutzung des vorigen Satzes offene Mengen $U_{1/4}$, $U_{3/4}$ mit

$$A \subset U_{1/4} \subset \overline{U}_{1/4} \subset U_{1/2} \subset \overline{U}_{1/2} \subset U_{3/4} \subset \overline{U}_{3/4} \subset X \setminus B,$$

usw.
Wir definieren dann

$$F(x) = \begin{cases} \inf\left\{r \in I \;\middle|\; x \in U_r\right\} & \text{für } x \in \bigcup_{r \in I} U_I, \\ 1 & \text{andernfalls.} \end{cases}$$

Dann ist offenbar $F(x) = 0$ für $x \in A$ und $F(x) = 1$ für $x \in B$.
Es bleibt zu zeigen, dass F stetig ist. Dazu bemerken wir, dass für $0 < r < s < 1$

$$\begin{aligned} F(x) > r &\iff \exists r' > r \colon x \notin \overline{U}_{r'} &&\iff x \notin \bigcap_{r'>r} \overline{U}_{r'}, \\ F(x) < s &\iff \exists s' < s \colon x \in U_{s'} &&\iff x \in \bigcup_{s'<s} U_{s'}. \end{aligned}$$

Also sind die Mengen

(a) $F^{-1}\left((r,s)\right) = \left(\bigcup_{s'<s} U_{s'}\right) \setminus \left(\bigcap_{r'>r} \overline{U}_{r'}\right)$,

(b) $F^{-1}\left([0,s)\right) = \bigcup_{s'<s} U_{s'}$,

(c) $F^{-1}\left((r,1]\right) = X \setminus \left(\bigcap_{r'>r} \overline{U}_{r'}\right)$

sämtlich offen in X. $\square$

Kapitel 8

Metrische Räume

8.1 Definition

Definition 8.1. Ein *metrischer Raum* (X, d) ist eine Menge X zusammen mit einer Funktion $d\colon X \times X \to [0, \infty)$ mit folgenden Eigenschaften: Für alle $x, y, z \in X$ gilt

(i) $d(x, y) = 0 \Leftrightarrow x = y$,

(ii) $d(x, y) = d(y, x)$,

(iii) (Dreiecksungleichung) $d(x, y) \le d(x, z) + d(z, y)$.

Die Funktion d heißt auch *Metrik* (oder *Abstandsfunktion*).

Beispiel 8.2. 1. Die durch

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

auf einer beliebigen Menge X gegebene *diskrete Metrik*.

2. $X = \mathbb{R}^n$, $1 \le p \le \infty$,

$$d_p(x, y) = \begin{cases} \left(\sum_i |x_i - y_i|^p \right)^{1/p}, & 1 \le p < \infty, \\ \max\limits_{i=1,\ldots,n} |x_i - y_i|, & p = \infty, \end{cases}$$

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

3. Für einen topologischen Raum Y sei X die Menge aller stetigen Funktionen $f\colon Y \to [0, 1]$ mit der Metrik

$$d(f, g) = \sup_{y \in Y} |f(y) - g(y)|.$$

8.2 Die metrische Topologie

Sei (X, d) ein metrischer Raum.

Definition 8.3. Wir nennen eine Teilmenge $U \subseteq X$ offen, falls es für jedes $x \in U$ ein $r > 0$ gibt, so dass

$$B_r(x) = \{y \in X \mid d(x, y) < r\}$$

in U enthalten ist.

Satz 8.4. *Mit diesem Begriff von Offenheit erhalten wir eine Topologie auf X.*

Beweis. Wir müssen überprüfen, dass der Durchschnitt zweier offener Mengen U, V wieder offen ist. (Allen anderen Axiome einer Topologie sind offensichtlich erfüllt.)
Sei also $x \in U \cap V$ und $B_r(x) \subseteq U$, $B_{r'}(x) \subseteq V$ mit gewissen $r > 0$, $r' > 0$. Dann ist $B_{\min\{r,r'\}}(x) \subseteq U \cap V$. □

Die so definierte Topologie heißt die *durch d auf X induzierte Topologie*.

Lemma 8.5. *Jede offene Kugel $B_r(x)$ ist offen im Sinne obiger Definition.*

Beweis. Sei $y \in B_r(x)$, d. h. $d(x, y) < r$. Dann gilt $B_{r-d(x,y)}(y) \subseteq B_r(x)$, denn $d(y, z) < r - d(x, y)$ impliziert

$$d(x, z) \leq d(x, y) + d(y, z) < d(x, y) + (r - d(x, y)) = r.$$

□

Bemerkung. Die *offenen Kugeln* $B_r(x)$, wobei x ganz X und r alle positiven reellen Zahlen durchläuft, bilden eine *Basis* der durch d auf X induzierten Topologie. Das heißt, dass die offenen Mengen in X gerade die Vereinigungen offener Kugeln sind.

Wir können die metrische Topologie mit Hilfe der Konvergenz von Folgen charakterisieren.

Definition 8.6. (i) Eine Folge $\{x_n\}_{n\in\mathbb{N}}$ ist eine Abbildung $\mathbb{N} \to X$, $n \mapsto x_n$.
(ii) Eine Folge $\{x_n\}_{n\in\mathbb{N}}$ in X heißt *konvergent* gegen eine Stelle $x^* \in X$, falls für jedes $r > 0$ ein $n_0 = n_0(x^*, r) \in \mathbb{N}$ gibt, so dass $x_n \in B_r(x^*)$ für alle $n \geq n_0$ gilt. Wir schreiben dann $x_n \to x^*$ für $n \to \infty$.

Lemma 8.7. *Eine Folge $\{x_n\}_{n\in\mathbb{N}}$ hat höchstens einen Grenzwert.*

Beweis. Angenommen, wir hätten $x_n \to x^*$ für $n \to \infty$ und $x_n \to y^*$ für $n \to \infty$ mit $x^* \neq y^*$. Dann gilt für $0 < r \leq d(x^*, y^*)/2$, dass $B_r(x^*) \cap B_r(y^*) = \emptyset$, was der Definition der Folgenkonvergenz widerspricht. □

Satz 8.8. *Eine Menge A in einem metrischen Raum (X, d) ist dann und nur dann abgeschlossen, wenn die Grenzwerte aller Folgen in A, die konvergieren, in A liegen.*

Beweis. Wir zeigen tatsächlich für eine beliebige Menge $B \subseteq X$:

$$x^* \in \bar{B} \iff \exists \text{ Folge } \{x_n\}_{n\in\mathbb{N}} \text{ in } B \text{ mit } x_n \to x^* \text{ für } n \to \infty.$$

($\Rightarrow$) Für jedes $n \in \mathbb{N}$ wählen wir ein $x_n \in B_{1/n}(x^*) \cap B$. Dann ist offenbar $x_n \to x^*$ für $n \to \infty$.

($\Leftarrow$) Gilt für eine Folge $\{x_n\}$ in B, dass $x_n \to x^*$ für $n \to \infty$, so hat jede offene Kugel $B_r(x^*)$ mit $r > 0$ einen nichtleeren Durchschnitt mit B, d. h. $x^* \in \bar{B}$.

□

Satz 8.9. *Eine Menge $A \subseteq (X, d)$ ist genau dann kompakt, wenn sie abgeschlossen ist und jede Folge in A eine konvergente Teilfolge besitzt. (Äquivalent dazu ist, dass wir aus jeder Folge in A eine in A konvergente Teilfolge auswählen können.)*

Beweis. ($\Rightarrow$) Sei $A \subseteq X$ kompakt. Dann ist A abgeschlossen, da X Hausdorffsch ist.

Sei weiterhin $\{x_n\} \subset A$ eine beliebige Folge. Angenommen, $\{x_n\}$ enthält keine konvergente Teilfolge. Insbesondere kann dann kein Folgenglied unendlich oft vorkommen, so dass wir $x_m \neq x_n$ für $m \neq n$ annehmen dürfen.

Jeder Punkt x_m ist ein isolierter Punkt (d. h. kein Häufungspunkt) der Menge $\{x_n \mid n \in \mathbb{N}\}$. Dann gibt es eine offene Kugel $U_m = B(x_m, r_m)$ mit $x_n \notin U_m$ für alle $n \neq m$. Man sieht leicht, dass die Menge $\{x_n \mid n \in \mathbb{N}\}$ abgeschlossen und somit die Menge $U_0 = X \setminus \{x_n \mid n \in \mathbb{N}\}$ offen in X ist.

Die Mengen U_n für $n \in \mathbb{N}$ zusammen mit U_0 bilden eine offene Überdeckung von A, die keine endliche Teilüberdeckung enthält.

($\Leftarrow$) Nun habe jede Folge in A eine in A konvergente Teilfolge. Wir zeigen zuerst:

Lemma 8.10. *Sei $\{U_i\}_{i\in I}$ eine offene Überdeckung von A. Dann gibt es ein $r > 0$, so dass es für alle $a \in A$ ein $i \in I$ gibt mit*

$$B(a, r) \subseteq U_i.$$

Beweis. Angenommen, dies wäre nicht der Fall. Wir wählen dann eine Folge $\{x_n\} \subset A$ mit $B(x_n, 1/n) \not\subseteq U_i$ für alle i. Sodann wählen wir eine Teilfolge $\{x_{n_k}\}$ von $\{x_n\}$ mit $x_{n_k} \to x^*$ für $k \to \infty$, wobei $x^* \in A$. Es gilt $x^* \in U_{i^*}$ für ein i^* und somit $B(x^*, r^*) \subseteq U_{i^*}$ für ein $r^* > 0$, da U_{i^*} offen ist.
Wir wählen nun ein k so groß, dass $d(x^*, x_{n_k}) \leq r^*/2$ und $1/k \leq r^*/2$ gilt. Dann haben wir für $y \in B(x_{n_k}, 1/n_k)$, dass

$$d(x^*, y) \leq d(x^*, x_{n_k}) + d(x_{n_k}, y) \leq \frac{r^*}{2} + \frac{r^*}{2} = r^*,$$

also $y \in B(x^*, r^*) \subseteq U_{i^*}$. Somit ist $B(x_{n_k}, 1/n_k) \subseteq U_{i^*}$ im Widerspruch zu $B(x_{n_k}, 1/n_k) \not\subseteq U_i$ für alle i. □

Hat nun jede Folge in A eine in A konvergente Teilfolge, so gibt es für jedes $r > 0$ endlich viele $x_1, \ldots, x_m \in A$ mit

$$A \subseteq B(x_1, r) \cup \cdots \cup B(x_m, r).$$

Andernfalls könnten wir ein $x_1 \in A$ beliebig wählen, dann ein $x_2 \in A \setminus B(x_1, r)$, dann ein $x_3 \in A \setminus (B(x_1, r) \cup B(x_2, r))$, usw., und würden auf diese Weise eine Folge $\{x_n\} \subset A$ mit $d(x_m, x_n) \geq r$ für alle $m \neq n$ erhalten, also eine Folge, aus der wir keine konvergente Teilfolge auswählen könnten.
Sei nun also $\{U_i\}_{i \in I}$ eine offene Überdeckung von A. Wir wählen $r > 0$ wie in Lemma 8.10 und wählen $x_1, \ldots, x_m \in A$ mit $A \subseteq B(x_1, r) \cup \cdots \cup B(x_m, r)$. Dann existiert für jedes $k = 1, \ldots, m$ ein $i_k \in I$ mit $B(x_k, r) \subseteq U_{i_k}$. $\{U_{i_k}\}_{k=1}^m$ ist eine endliche Teilüberdeckung von A. □

Satz 8.11. *Seien X, Y metrische Räume, $f\colon X \to Y$ eine Abbildung. Dann ist f genau dann stetig, wenn für alle Folgen $\{x_n\} \subset X$ aus $x_n \to x$ für $n \to \infty$ $f(x_n) \to f(x)$ für $n \to \infty$ folgt.*

Beweis. ($\Rightarrow$) Sei zuerst f stetig und $\{x_n\}$ eine Folge in X mit $x_n \to x$ für $n \to \infty$. Für jedes $\epsilon > 0$ gibt es wegen der Stetigkeit von f ein $\delta > 0$ mit

$$f(B(x, \delta)) \subseteq B(f(x), \epsilon).$$

Wir wählen ein n_0, so dass $x_n \in B(x, \delta)$ für alle $n \geq n_0$. Dann ist $f(x_n) \in B(f(x), \epsilon)$ für alle $n \geq n_0$. Es folgt, dass $f(x_n) \to f(x)$ für $n \to \infty$.

($\Leftarrow$) Nun habe f umgekehrt die in der Formulierung des Satzes genannte Eigenschaft. Sei $B \subseteq Y$ abgeschlossen. Wir wollen zeigen, dass $A = f^{-1}(B)$ ebenfalls abgeschlossen ist. Sei also $\{x_n\} \subset A$ eine konvergente Folge mit $x_n \to x$ für $n \to \infty$. Zu zeigen ist, dass $x \in A$ gilt.

Aus $x_n \to x$ für $n \to \infty$ folgt, dass $f(x_n) \to f(x)$ für $n \to \infty$. Wegen $\{f(x_n)\} \subset B$ und da B abgeschlossen ist, gilt $f(x) \in B$. Damit ist jedoch $x \in f^{-1}(B) = A$.

□

8.3 Vollständige metrische Räume

Sei (X, d) ein metrischer Raum.

Definition 8.12. (i) Eine Folge $\{x_n\}$ in X heißt *Cauchyfolge*, falls

$$\forall \epsilon > 0 \; \exists n_0 = n_0(\epsilon) \; \forall m, n \geq n_0\colon d(x_m, x_n) \leq \epsilon.$$

(ii) X heißt *vollständig*, falls jede Cauchyfolge konvergiert.

Beispiel 8.13. (a) $\mathbb{R}$, $[a, b]$ sind vollständig.

(b) Vollständigkeit ist keine topologische Eigenschaft. So ist beispielsweise das offene Intervall $(0,1)$ homöomorph zu $\mathbb{R}$, aber *nicht* vollständig. ($\{1/n\}_{n\in\mathbb{N}}$ ist eine Cauchyfolge in $(0,1)$, die in $(0,1)$ nicht konvergiert.)

Definition 8.14. Ein metrischer Raum $(\bar{X},\bar{d})$ heißt *Vervollständigung* von (X,d), falls es eine isometrische (d. h. metrikerhaltende) Abbildung

$$i\colon (X,d) \to (\bar{X},\bar{d})$$

gibt, so dass $i(X)$ in $\bar{X}$ dicht ist (d. h. der Abschluss von $i(X)$ ist $\bar{X}$).

Theorem 8.15. *Für jeden metrischen Raum gibt es bis auf Isometrie genau eine Vervollständigung.*

Beispiel 8.16. (i) $\mathbb{Q}$ ist nicht vollständig, seine Vervollständigung ist $\mathbb{R}$.

(ii) Die Vervollständigung von $(0,1)$ ist $[0,1]$.

Beweis von Theorem 8.15. (a) (Eindeutigkeit) Seien $\bar{X}$, $\bar{X}'$ zwei Vervollständigungen. Wir können X sowohl als Unterraum von $\bar{X}$ als auch als Unterraum von $\bar{X}'$ betrachten. Dann setzt sich die identische Abbildung $X \to X$, $x \mapsto x$, in eindeutiger Weise zu einer Isometrie $f\colon \bar{X} \to \bar{X}'$ fort.

(b) (Existenz) Sei $\mathcal{X}$ der Raum aller Cauchyfolgen in X. Auf $\mathcal{X}$ betrachten wir die folgende Relation:

$$\{x_n\} \sim \{y_n\} :\Longleftrightarrow \big\{d(x_n,y_n)\big\} \text{ ist Nullfolge in } \mathbb{R}.$$

$\sim$ ist eine Äquivalenzrelation auf $\mathcal{X}$: $\sim$ ist offenbar reflexiv und symmetrisch. Die Transitivität ergibt sich wie folgt:

Sei $\{x_n\} \sim \{y_n\}$ und $\{y_n\} \sim \{z_n\}$, d. h. $\big\{d(x_n,y_n)\big\}$ und $\big\{d(y_n,z_n)\big\}$ sind Nullfolgen. Dann ist wegen

$$0 \le d(x_n,z_n) \le d(x_n,y_n) + d(y_n,z_n) \to 0 \text{ für } n \to \infty$$

auch $\big\{d(x_n,z_n)\big\}$ eine Nullfolge.

Sei $\bar{X}$ der Raum der Äquivalenzklassen von $\mathcal{X}$ bezüglich $\sim$. Auf $\bar{X}$ definieren wir die Metrik

$$\bar{d}\left([\{x_n\}],[\{y_n\}]\right) := \lim_{n\to\infty} d(x_n,y_n).$$

(i) Zeigen zuerst: $\bar{d}$ ist wohldefiniert.

(a) Seien $\{x_n\}$, $\{y_n\}$ Cauchyfolgen in X. Dann ist $\big\{d(x_n,y_n)\big\}$ wegen

$$\begin{aligned}
&|d(x_m,y_m) - d(x_n,y_n)| \\
&\qquad \le |d(x_m,y_m) - d(x_n,y_n)| + |d(x_m,y_m) - d(x_n,y_n)| \\
&\qquad \le d(x_m,x_n) + d(y_m,y_n) \to 0
\end{aligned}$$

für $m,n \to \infty$ eine Cauchyfolge in $\mathbb{R}$. Also existiert $\lim\limits_{n\to\infty} d(x_n,y_n)$.

(b) (Repräsentantenunabhängigkeit). Sei $\{x'_n\}$ ein weiterer Repräsentant der Klasse $[\{x_n\}]$, d. h. $\{x_n\} \sim \{x'_n\}$. Dann gilt

$$|d(x_n, y_n) - d(x'{}_n, y_n)| \leq d(x_n, x'{}_n) \to 0$$

für $n \to \infty$, d. h. die Folgen $\big\{d(x_n, y_n)\big\}$ und $\big\{d(x'{}_n, y_n)\big\}$ haben denselben Grenzwert.

(ii) Wir zeigen als nächstes, dass der Raum $\bar{X}$ vollständig ist. Dazu wir wählen eine Cauchyfolge $\{\xi^k\}_{k\in\mathbb{N}}$ in $\bar{X}$. Wir repräsentieren jedes Glied ξ^k dieser Folge durch eine Cauchyfolge $\{x_n^k\}_{k\in\mathbb{N}}$ in X, d. h. $\xi^k = \big[\{x_n^k\}\big]$.
Dann wählen wir eine Folge $\{y^k\}$ in X, von der wir zeigen werden:

(a) $\{y^k\}$ ist eine Cauchyfolge,

(b) $\{y^k\}$ repräsentiert den Grenzwert der Folge $\{\xi^k\}$.

Wir setzen $y^k = x_{n_k}^k$, wobei n_k so gewählt ist, dass $d(x_m^k, x_n^k) \leq 1/k$ für $m, n \geq n_k$ gilt.

Zu (a): Sei $\epsilon > 0$. Wir finden dann ein $k_0 \geq 1/\epsilon$, so dass

$$\lim_{n\to\infty} d(x_n^k, x_n^l) \leq \epsilon \text{ für } k, l \geq k_0.$$

Dann gilt für $k, l \geq k_0$

$$\begin{aligned} d(y^k, y^l) &\leq d(x_{n_k}^k, x_n^k) + d(x_n^k, x_n^l) + d(x_n^l, x_{n_l}^l) \\ &\leq \frac{1}{k} + 2\epsilon + \frac{1}{l} \leq 4\epsilon \end{aligned}$$

für ein hinreichend großes n. Also ist $\{y^k\}$ Cauchyfolge in X.

Zu (b): Sei $\epsilon > 0$. Indem wir k_0 wie in (a) wählen, erhalten wir für $k \geq k_0$, dass

$$\begin{aligned} d(y^m, x_m^k) &\leq d(x_{n_k}^m, x_n^m) + d(x_n^m, x_n^k) + d(x_n^k, x_m^k) \\ &\leq \frac{1}{m} + 2\epsilon + \frac{1}{k} \leq 4\epsilon \end{aligned}$$

für $m \geq k_0$ und n hinreichend groß. Insbesondere

$$\lim_{m\to\infty} d(y^m, x_m^k) \leq 4\epsilon \text{ für } k \geq k_0.$$

Folglich gilt $\lim\limits_{k\to\infty} \xi^k = [\{y^m\}]$ in $\bar{X}$.

(iii) Die isometrische Abbildung von X nach $\bar{X}$ ergibt sich, indem jedes $x \in X$ auf die Klasse abgebildet wird, die durch die Folge konstant x repräsentiert ist. Indem wir X mit seinem Bild in $\bar{X}$ unter dieser Einbettung identifizieren, sehen wir, dass X in $\bar{X}$ dicht ist: Es gilt

$$\lim_{n\to\infty} x_n = [\{x_n\}]$$

in $\bar{X}$ für jede Cauchyfolge $\{x_n\}$ in X.

□

Beispiel 8.17 (Beispiel eines vollständigen metrischen Raumes).

$$X = C([0,1],\mathbb{R}) = \big\{f\colon [0,1] \to \mathbb{R} \,\big|\, f \text{ ist stetig}\big\}$$

mit der Metrik

$$d(f,g) = \max_{x\in[0,1]} |f(x) - g(x)|\,.$$

Behauptung. X ist vollständig.

Beweis. Sei dazu $\{f^n\}$ eine Cauchyfolge. Dann ist für jedes $x \in [0,1]$ die Folge $\{f^n(x)\}$ eine Cauchyfolge in $\mathbb{R}$. Somit konvergiert die Folge $\{f^n\}$ punktweise gegen eine Grenzfunktion f, d. h. $\lim\limits_{n\to\infty} f^n(x) = f(x)$ für alle $x \in [0,1]$. Diese Konvergenz ist tatsächlich gleichmäßig, da aus

$$\max_{x\in[0,1]} |f^m(x) - f^n(x)| \le \epsilon$$

für ein gegebenes $\epsilon > 0$ und m, $n \ge n_0$ bei geeeignetem n_0 durch Grenzübergang $m \to \infty$ folgt, dass

$$\sup_{x\in[0,1]} |f(x) - f^m(x)| \le \epsilon$$

für $n \ge n_0$. Damit ist die Funktion f gemäß Satz 6.7 stetig, gehört als zu X.

8.4 Der Banachsche Fixpunktsatz

Sei (X,d) ein vollständiger metrischer Raum (z. B. X kompakt). Sei ferner $T\colon X \to X$ eine Abbildung mit

$$d\,(Tx,Ty) \le \alpha\, d(x,y), \quad \forall x,\, y \in X,$$

für ein $0 \le \alpha < 1$.

Satz 8.18 (Banach)**.** *Die Abbildung $T\colon X \to X$ besitzt genau einen Fixpunkt.*

Beweis. (a) (Eindeutigkeit) Seien x, $y \in X$ Fixpunkte von T. Dann gilt

$$d(x,y) = d\,(Tx,Ty) \le \alpha\, d(x,y),$$

also $(1-\alpha)\, d(x,y) \le 0$, $d(x,y) = 0$ und $x = y$.

(b) (Existenz) Zuerst erhalten wir induktiv, dass

$$d\,(T^n x, T^n y) \le \alpha^n\, d(x,y), \quad \forall x,\, y \in X.$$

Sei nun $x_0 \in X$ beliebig und $x_n = T^n x_0$ für $n \in \mathbb{N}$. Dann gilt

$$d(x_{n+1},x_n) = d\,(T^n x_1, T^n x_0) \le \alpha^n d(x_1,x_0),$$

also für $n,\, k \in \mathbb{N}$

$$\begin{aligned} d(x_{n+k}, x_n) &\le d(x_{n+k}, x_{n+k-1}) + \dots + d(x_{n+2}, x_{n+1}) + d(x_{n+1}, x_n) \\ &\le \left(\alpha^{n+k-1} + \dots + \alpha^{n+1} + \alpha^n\right) d(x_1, x_0) \\ &= \alpha^n \,\frac{1-\alpha^k}{1-\alpha}\, d(x_1, x_0) \le \frac{\alpha^n}{1-\alpha}\, d(x_1, x_0). \end{aligned}$$

Wegen $\frac{\alpha^n}{1-\alpha} \to 0$ für $n \to \infty$ ist $\{x_n\}$ eine Cauchyfolge in X. Wegen der Vollständigkeit von X konvergiert $\{x_n\}$ gegen ein $x^* \in X$. Indem wir in

$$d\left(Tx_{n+1}, x_{n+1}\right) \le \alpha\, d\left(Tx_n, x_n\right)$$

zum Grenzwert für $n \to \infty$ übergehen, erhalten wir

$$d(Tx^*, x^*) \le \alpha\, d(Tx^*, x^*),$$

also $d(Tx^*, x^*) = 0$ und $Tx^* = x^*$. Folglich ist x^* der gesuchte Fixpunkt. □

Bemerkung. In Anwendungen ist es oft wichtig, dass dieser Beweis ein konstruktives Verfahren zur Approximation des gesuchten Fixpunktes x^* liefert: Für ein *beliebiges* x_0 in X gilt $\lim\limits_{n\to\infty} T^n x_0 = x^*$, wobei wir die Fehlerabschätzung

$$d\left(x^*, T^n x_0\right) \le \frac{\alpha^n}{1-\alpha}\, d(Tx_0, x_0)$$

haben.

Beispiel 8.19. Sei $X = C([0,1], \mathbb{R})$ und $k\colon [0,1] \times [0,1] \times \mathbb{R} \to \mathbb{R}$ stetig. Wir betrachten den Integraloperator $K\colon X \to X$,

$$\left(Kf\right)(x) = f(x) + \int_0^1 k\left(x, y, f(y)\right)\, dy, \quad x \in [0,1].$$

Unter der Voraussetzung

$$\left|k\left(x,y,u\right) - k\left(x,y,v\right)\right| \le \alpha\left|u - v\right|$$

für $x, y \in [0,1]$, $u, v \in \mathbb{R}$ und ein $\alpha < 1$ wollen wir zeigen, dass der Operator $K\colon X \to X$ *bijektiv* ist. Dazu wählen wir ein $h \in X$ und formulieren die Gleichung $Kf = h$ für das gesuchte f als Fixpunktgleichung $f = Tf$,

$$Tf(x) = h(x) - \int_0^1 k\left(x, y, f(y)\right)\, dy, \quad x \in [0,1].$$

Wegen

$$\begin{aligned} \left|Tf(x) - Tg(x)\right| &\le \int_0^1 \left|k\left(x,y,f(y)\right) - k\left(x,y,g(y)\right)\right| dy \\ &\le \alpha \int_0^1 \left|f(y) - g(y)\right| dy \le \alpha \max_{x\in[0,1]} \left|f(x) - g(x)\right| \end{aligned}$$

gilt

$$d\,(Tf, Tg) \leq \alpha d(x, y)$$

für alle f, $g \in X$. Damit hat der Operator T genau einen Fixpunkt, d. h. bei gegebenem $h \in X$ hat die Gleichung $Kf = h$ genau eine Lösung $f \in X$.

Kapitel 9

Differentiation von Funktionen mehrerer Veränderlicher

Im folgenden bezeichnet $\mathbb{R}^n$ den *n-dimensionalen Euklidischen Raum* mit seiner Standardtopologie. Stellen im $\mathbb{R}^n$ bezeichnen wir mit $x = (x_1, x_2, \dots, x_n)^T$, d. h. x ist der Spaltenvektor $\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$. Benötigen wir einen Abstandsbegriff im $\mathbb{R}^n$, so arbeiten wir mit der *Euklidischen Metrik*

$$|x - y| = \left(\sum_{k=1}^{n} |x_k - y_k|^2 \right)^{1/2}$$

für $x = (x_1, \dots, x_n)^T$, $y = (y_1, \dots, y_n)^T \in \mathbb{R}^n$. Insbesondere bezeichnet $|x| = \left(\sum_{k=1}^{n} x_k^2 \right)^{1/2}$ den Betrag von x.

Mit den beiden vorangegangenen Kapiteln sind die Begriffe Stetigkeit einer Funktion $f\colon U \to \mathbb{R}^m$ für $U \subseteq \mathbb{R}^n$, Folgenstetigkeit, Kompaktheit, usw. klar. Insbesondere gilt im $\mathbb{R}^n$ der *Satz von Heine-Borel*, wonach eine Teilmenge von $\mathbb{R}^n$ genau dann kompakt ist, wenn sie abgeschlossen und beschränkt ist.

9.1 Partielle Ableitungen

Sei $U \subseteq \mathbb{R}^n$ offen, $f\colon U \to \mathbb{R}$.

Definition 9.1. (i) Die *partielle Ableitung* der Funktion f nach der Veränderlichen x_k an der Stelle $x^0 = (x_1^0, \dots, x_n^0) \in U$ (falls existent) ist die Ableitung der Funktion

$$g(t) = f\left(x_1^0, \dots, x_{k-1}^0, t, x_{k+1}^0, \dots, x_n^0\right)$$

an der Stelle x_k^0. Man schreibt $\frac{\partial f}{\partial x_k}(x^0)$ oder auch $f_{x_k}(x^0)$.

(ii) Die Funktion f heißt in U partiell nach der Veränderlichen x_k differenzierbar, falls $\frac{\partial f}{\partial x_k}$ an jeder Stelle in U existiert.

Die Differentiationsregeln für Summen, Differenzen, Produkte, Quotienten und zusammengesetzte Funktionen übertragen sich direkt vom eindimensionalen auf den mehrdimensionalen Fall.

Beispiel 9.2. Im $\mathbb{R}^2$ schreiben wir auch $(x, y) = (x_1, x_2)$. Die Funktion $f(x, y) = \tan(x^2 + y)$ ist in allen Stellen (x, y) mit $x^2 + y \notin \frac{\pi}{2} + \mathbb{Z}$ sowohl nach x als auch nach y partiell differenzierbar mit

$$\frac{\partial f}{\partial x}(x, y) = \frac{2x}{\cos^2(x^2 + y)}, \qquad \frac{\partial f}{\partial y}(x, y) = \frac{1}{\cos^2(x^2 + y)}.$$

Aus der Tatsache, dass alle partiellen Ableitungen an der Stelle x^0 existieren, folgt nicht, dass f an x^0 stetig ist.

Beispiel 9.3. Die Funktion

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

besitzt partielle Ableitungen auf ganz $\mathbb{R}^2$, ist an $(0, 0)$ aber nicht stetig: Wir haben $f(x, 0) = f(0, y) = 0$, also $\frac{\partial f}{\partial x}(0, 0) = \frac{\partial f}{\partial y}(0, 0) = 0$, jedoch $f(x, x) = \frac{1}{2}$, $f(x, -x) = -\frac{1}{2}$.

9.2 Die totale Ableitung

Sei $U \subseteq \mathbb{R}^n$ offen, $f : U \to \mathbb{R}^m$. In Komponenten schreiben wir dann auch $f = (f_1, \dots, f_m)^T$, wobei $f_j : U \to \mathbb{R}$ für $j = 1, \dots, m$.

Definition 9.4. (i) f heißt an der Stelle x^0 *differenzierbar*, falls es eine Matrix $A \in M(m, n; \mathbb{R})$ gibt, so dass

$$f(x^0 + h) = f(x^0) + Ah + \rho(h)$$

für alle $h \in \mathbb{R}^n$ mit $x^0 + h \in U$, wobei $\lim\limits_{h \to 0} \frac{\rho(h)}{|h|} = 0$.

(ii) f heißt in U differenzierbar, falls f an jeder Stelle von U differenzierbar ist.

Lemma 9.5. *Die Matrix A in* (i) *ist eindeutig bestimmt.*

Beweis. Sei $f(x^0 + h) = f(x^0) + Bh + \rho'(h)$ für alle $h \in \mathbb{R}^n$ mit $x^0 + h \in U$, wobei $\lim\limits_{h \to 0} \frac{\rho'(h)}{|h|} = 0$. Dann gilt $(A - B)h = -\rho(h) + \rho'(h)$, also

$$\lim_{h \to 0} (A - B) \frac{h}{|h|} = 0.$$

Daraus folgt $A - B = 0$ bzw. $A = B$. □

Definition 9.6. Die Matrix A in (i) (falls existent) heißt die (*totale*) *Ableitung* von f an der Stelle x^0. Man schreibt $A = Df(x^0)$.

Satz 9.7. *Ist f an der Stelle x^0 differenzierbar, so ist f an der Stelle x^0 stetig.*

Beweis. Folgt direkt aus der Gleichung

$$f(x^0 + h) = f(x^0) + Df(x^0)h + \rho(h),$$

siehe auch Satz 3.3. □

Satz 9.8. *Ist f an der Stelle x^0 differenzierbar, so existieren die partiellen Ableitungen $\frac{\partial f_j}{\partial x_k}(x^0)$ für $j = 1, \dots, m$, $k = 1, \dots, n$ und*

$$Df = \left(\frac{\partial f}{\partial x_1}, \dots, \frac{\partial f}{\partial x_n} \right) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1} & \cdots & \frac{\partial f_m}{\partial x_n} \end{pmatrix},$$

jeweils an der Stelle x^0, wobei $\frac{\partial f}{\partial x_k}$ der Spaltenvektor $\left(\frac{\partial f_1}{\partial x_k}, \dots, \frac{\partial f_m}{\partial x_k} \right)^T$ ist.

Beweis. Ist $A = Df(x^0)$ und $e_k = (0, \dots, 0, 1, 0, \dots, 0)^T$ mit der 1 an der k-ten Stelle, so ist Ae_k die k-te Spalte von A. Wir erhalten

$$\frac{\partial f}{\partial x_k}(x^0) = \lim_{s \to 0} \frac{f(x^0 + se_k) - f(x^0)}{s} = \lim_{s \to 0} \left[Ae_k + \frac{\rho(se_k)}{s} \right] = Ae_k.$$

□

Definition 9.9. Die Matrix $\begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1} & \cdots & \frac{\partial f_m}{\partial x_n} \end{pmatrix}$ an der Stelle x^0 heißt auch die *Jacobi-Matrix* oder *Jacobische* von f an der Stelle x^0.

Umgekehrt zum vorigen Satz gilt:

Satz 9.10. *Die partiellen Ableitungen $\frac{\partial f_j}{\partial x_k}$ für $j = 1 \dots, m$, $k = 1, \dots, n$ mögen in einer kleinen Umgebung von x^0 existieren und an x^0 stetig sein. Dann ist f an der Stelle x^0 differenzierbar.*

Beweis. O. B. d. A. sei $m = 1$, $f_1 = f$. Wir definieren

$$\rho(h) = f(x^0 + h) - f(x^0) - \sum_{k=1}^{n} \frac{\partial f}{\partial x_k}(x^0) h_k$$

für $h = (h_1, \dots, h_n)^T \in \mathbb{R}^n$ mit $x^0 + h \in U$. Wir setzen weiterhin $x^1 = x^0 + h_1 e_1$, $x^2 = x^1 + h_2 e_2$, …, $x^n = x^{n-1} + h_n e_n = x^0 + h$. Dann gilt

$$\begin{aligned} \rho(h) &= \sum_{k=1}^{n} \left[f(x^k) - f(x^{k-1}) - \frac{\partial f}{\partial x_k}(x^0) h_k \right] \\ &= \sum_{k=1}^{n} \left[\frac{\partial f}{\partial x_k} \left(x^{k-1} + s_k h_k e_k \right) - \frac{\partial f}{\partial x_k}(x^0) \right] h_k \end{aligned}$$

für gewisse $s_k \in (0,1)$ nach dem Mittelwertsatz. Also gilt $\lim\limits_{h\to 0} \frac{\rho(h)}{|h|} = 0$, da $x^{k-1} + s_k h_k e_k \to x^0$ für $h \to 0$. $\square$

Bemerkung. Im Fall $m = 1$ nennt man $\nabla f(x^0) = Df(x^0)$ auch den *Gradienten* von f an der Stelle x^0. Geometrisch zeigt der Gradient in die Richtung des steilsten Anstiegs von f an der Stelle x^0.

Unter der Benutzung der Matrizenschreibweise formuliert sich die *Kettenregel* sehr übersichtlich.

Satz 9.11. *Seien $U \subseteq \mathbb{R}^n$, $V \subseteq \mathbb{R}^m$ offen, $f\colon U \to V$ an $x^0 \in U$ differenzierbar und $g\colon V \to \mathbb{R}^p$ an $y^0 = f(x^0) \in V$ differenzierbar. Dann ist $g \circ f\colon U \to \mathbb{R}^p$ an x^0 differenzierbar und*

$$D\,(g \circ f)\,(x^0) = Dg(y^0)Df(x^0).$$

Beweis. Wir schreiben

$$\begin{aligned} f(x^0 + h) &= f(x^0) + Df(x^0)h + \rho(h), \\ g(y^0 + k) &= g(y^0) + Dg(y^0)k + \sigma(k) \end{aligned}$$

mit $\lim\limits_{h\to 0} \frac{\rho(h)}{|h|} = 0$, $\lim\limits_{k\to 0} \frac{\sigma(k)}{|k|} = 0$. Dann gilt

$$\begin{aligned} &(g \circ f)\,(x^0 + h) = g\left(f(x^0) + Df(x^0)h + \rho(h)\right) \\ &= (g \circ f)\,(x^0) + Dg(y^0)\left(Df(x^0)h + \rho(h)\right) + \rho\left(Df(x^0)h + \rho(h)\right) \\ &= (g \circ f)\,(x^0) + Dg(y^0)Df(x^0)h + \left(Dg(y^0)\rho(h) + \sigma\left(Df(x^0)h + \rho(h)\right)\right) \end{aligned}$$

Berücksichtigen wir, dass es für jedes $A \in M(m,n;\mathbb{R})$ eine Konstante $C \geq 0$ mit $|Ah| \leq C\,|h|$ für alle $h \in \mathbb{R}^n$ gibt (die kleinstmögliche Wahl für C ist $\sup\limits_{|h|=1} |Ah|$), so erhalten wir mit geeigneten Konstanten C, dass

$$\frac{\left|Dg(y^0)\rho(h)\right|}{|h|} \leq C\,\frac{|\rho(h)|}{|h|} \to 0 \text{ für } h \to 0.$$

und

$$\begin{aligned} \frac{\left|\sigma\left(Df(x^0)h + \rho(h)\right)\right|}{|h|} &= \frac{\left|\sigma\left(Df(x^0)h + \rho(h)\right)\right|}{|Df(x^0)h + \rho(h)|} \cdot \frac{\left|Df(x^0)h + \rho(h)\right|}{|h|} \\ &\leq \frac{\left|\sigma\left(Df(x^0)h + \rho(h)\right)\right|}{|Df(x^0)h + \rho(h)|} \cdot \left(C\,\frac{|h|}{|h|} + \frac{|\rho(h)|}{|h|}\right) \to 0 \text{ für } h \to 0. \end{aligned}$$

$\square$

9.3 Höhere Ableitungen

Sei $U \subseteq \mathbb{R}^n$ offen, $f\colon U \to \mathbb{R}$.

Definition 9.12. Sei f in U nach der Veränderlichen x_k partiell differenzierbar, $\frac{\partial f}{\partial x_k}\colon U \to \mathbb{R}$. Ist dann $\frac{\partial f}{\partial x_k}$ an der Stelle $x^0 \in U$ nach x_j differenzierbar, d. h. $\frac{\partial(\partial f/\partial x_k)}{\partial x_j}(x^0)$ existiert, so schreiben wir dafür $\frac{\partial^2 f}{\partial x_j \partial x_k}(x^0)$ oder auch $f_{x_j x_k}(x^0)$. Im Fall $j = k$ schreiben wir $\frac{\partial^2 f}{\partial x_k^2}$.

Induktiv definieren wir höhere Ableitungen $\dfrac{\partial^m f}{\partial x_{k_1} \dots \partial x_{k_m}}(x^0)$, wobei $k_r \in \{1, 2 \dots, n\}$ für $r = 1, \dots, m$.

Satz 9.13 (Schwarz). *Es mögen die Ableitungen $\frac{\partial^2 f}{\partial x_j \partial x_k}$ und $\frac{\partial^2 f}{\partial x_k \partial x_j}$ in einer Umgebung von x^0 existieren und an x^0 stetig sein. Dann gilt*

$$\frac{\partial^2 f}{\partial x_j \partial x_k}(x^0) = \frac{\partial^2 f}{\partial x_k \partial x_j}(x^0).$$

Beweis. Indem man die Veränderlichen x_i mit $i \neq j, k$ als Parameter betrachtet, genügt es den Fall $n = 2$ zu betrachten, $(x, y) = (x_j, x_k)$. Wir behaupten, dass

$$\begin{aligned}
&\frac{\partial^2 f}{\partial x \partial y}(x^0, y^0) \\
&\qquad = \lim_{h \to 0} \frac{f(x^0 + h, y^0 + h) - f(x^0 + h, y^0) - f(x^0, y^0 + h) + f(x^0, y^0)}{h^2} \\
&\qquad = \frac{\partial^2 f}{\partial y \partial x}(x^0, y^0).
\end{aligned}$$

In der Tat ergibt eine Anwendung des Mittelwertsatzes auf die Funktion $y \mapsto f(x^0 + h, y) - f(x^0, y)$ für $h \neq 0$, $|h|$ klein, dass

$$\begin{aligned}
&\frac{f(x^0 + h, y^0 + h) - f(x^0 + h, y^0) - f(x^0, y^0 + h) + f(x^0, y^0)}{h} \\
&\qquad\qquad = f_y(x^0 + h, y^0 + t_h h) - f_y(x^0, y^0 + t_h h)
\end{aligned}$$

für ein $t_h \in (0, 1)$. Nochmalige Anwendung des Mittelwertsatzes, diesmal auf $x \mapsto f_y(x, y^0 + t_h h)$, liefert

$$\begin{aligned}
&\frac{f(x^0 + h, y^0 + h) - f(x^0 + h, y^0) - f(x^0, y^0 + h) + f(x^0, y^0)}{h^2} \\
&\qquad\qquad = f_{xy}(x^0 + s_h h, y^0 + t_h h)
\end{aligned}$$

für ein $s_h \in (0, 1)$. Im Grenzübergang $h \to 0$ folgt, dass der Grenzwert der linken Seite gleich $\frac{\partial^2 f}{\partial x \partial y}(x^0, y^0)$ ist.

Genauso berechnet sich dieser Grenzwert zu $\frac{\partial^2 f}{\partial y \partial x}(x^0, y^0)$, indem man zuerst den Mittelwertsatz auf $x \mapsto f(x, y^0 + h) - f(x, y^0)$ für $h \neq 0$, $|h|$ klein anwendet. $\square$

Bemerkung. Ganz analoge Resultate gelten für höhere partielle Ableitungen, vorausgesetzt, dass diese stetig sind, z. B.

$$\frac{\partial^3 f}{\partial x \partial y \partial x}(x^0) = \frac{\partial^3 f}{\partial x^2 \partial y}(x^0).$$

Analog definiert man höhere Ableitungen im Fall der totalen Ableitung. Sei dazu $U \subseteq \mathbb{R}^n$ offen, $f\colon U \to \mathbb{R}^m$. Ist f in U differenzierbar, so ist $Df\colon U \to \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$, wobei wir mit $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ den *Raum der linearen Abbildungen von* $\mathbb{R}^n$ *nach* $\mathbb{R}^m$ bezeichnen, den wir mit $M(m, n; \mathbb{R})$ identifizieren.

Definition 9.14. Ist Df in $x^0 \in U$ differenzierbar, so bezeichnen wir die Ableitung mit $D^2 f(x^0)$. Analog definieren wir höhere Ableitungen $D^b f(x^0)$ für $b \geq 3$.

Bemerkung. Gemäß Definition ist $D^b f(x^0)$ eine b-multilineare Abbildung[1] $\underbrace{\mathbb{R}^n \times \cdots \times \mathbb{R}^n}_{b-\text{mal}} \to \mathbb{R}^m$, wobei wir beispielsweise im Fall $b = 2$ den Raum $\mathcal{L}(\mathbb{R}^n, \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m))$ mit dem Raum der bilinearen Abbildungen $\mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^m$ identifizieren[2]. Genau gilt

$$\begin{aligned} &D^b f(x^0 + k)(h_1, \ldots, h_b) \\ &\qquad = D^b f(x^0)(h_1, \ldots, h_b) + D^{b+1} f(x^0)(h_1, \ldots, h_b, k) + o(|k|) \end{aligned}$$

für $k \to 0$, wobei $o(|k|)$ eine (nicht weiter spezifizierte) Funktion $\varrho(k)$ mit der Eigenschaft $\lim\limits_{k \to 0} \frac{\varrho(k)}{|k|} = 0$ bezeichnet.

Lemma 9.15. *Die b-multilineare Abbildung*

$$D^b f(x^0)\colon \underbrace{\mathbb{R}^n \times \cdots \times \mathbb{R}^n}_{b-\text{mal}} \to \mathbb{R}^m$$

ist total-symmetrisch, d. h.

$$D^b f(x^0)(h_{\sigma(1)}, \ldots, h_{\sigma(b)}) = D^b f(x^0)(h_1, \ldots, h_b)$$

für alle Permutationen σ von $\{1, \ldots, b\}$.

Beweis. Wie im Beweis des Satzes von Schwarz zeigt man, dass

$$\begin{aligned} D^2 f(x^0)(h, k) &= \lim_{s \to 0} \frac{f(x^0 + sh + sk) - f(x^0 + sh) - f(x^0 + sk) + f(x^0)}{s^2} \\ &= D^2 f(x^0)(k, h). \end{aligned}$$

Analog für höhere Ableitungen. □

[1] Eine Abbildung $B\colon \mathbb{R}^n \times \cdots \times \mathbb{R}^n \to \mathbb{R}^m$ heißt *multilinear*, wenn B linear in jedem Argument ist.

[2] We identifizieren die Abbildung $D \in \mathcal{L}(\mathbb{R}^n, \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m))$ mit der bilinearen Abbildung $B\colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^m$ vermöge

$$B(h, k) = D(h)k, \quad \forall h,\, k \in \mathbb{R}^n,$$

wobei wir den Umstand $D(h) \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ ausnutzen.

Satz 9.16. *Ist $f\colon U \to \mathbb{R}^m$ an der Stelle $x^0 \in U$ b-mal differenzierbar, so ist f an der Stelle x^0 b-mal partiell differenzierbar, und es gilt*

$$\frac{\partial^b f}{\partial x_{k_1} \dots \partial x_{k_b}}(x^0) = \left(D^b f\right)(x^0)(e_{k_1}, \dots, e_{k_b}),$$

wobei e_k den k-ten Einheitsvektor bezeichnet.

Beweis. Der Fall $b = 1$ wurde in Satz 9.8 erledigt. Für $b = 2$ gilt

$$\frac{1}{t}\left[f(x^0 + se_j + te_k) - f(x^0 + se_j) - f(x^0 + te_k) + f(x^0)\right] \to \frac{\partial f}{\partial x_k}(x^0 + se_j) - \frac{\partial f}{\partial x_k}(x^0)$$

für $t \to 0$ und

$$\frac{1}{s}\left[\frac{\partial f}{\partial x_k}(x^0 + se_j) - \frac{\partial f}{\partial x_k}(x^0)\right] \to \frac{\partial^2 f}{\partial x_j \partial x_k}(x^0)$$

für $s \to 0$. Damit folgt das Ergebnis mittels der Formel im Beweis des vorangegangenen Satzes.
Höhere Ableitungen behandelt man analog. □

Bemerkung. Im Fall $m = 1$ heißt die symmetrische Matrix[3]

$$D^2 f(x^0) = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_1 \partial x_n} & \frac{\partial^2 f}{\partial x_2 \partial x_n} & \cdots & \frac{\partial^2 f}{\partial x_n^2} \end{pmatrix}(x^0)$$

die *Hesse-Matrix* von f an der Stelle x^0.

Umgekehrt gilt:

Satz 9.17. *Ist f b-mal partiell differenzierbar in einer Umgebung von $x^0 \in U$ und sind die b-ten partiellen Ableitungen von f an x^0 stetig, so existiert $D^b f(x^0)$.*

Beweis. Mittels Induktion nach b unter Benutzung des Satzes 9.10. □

---

[3]Eine Matrix $A \in M(n, n; \mathbb{R})$ heißt *symmetrisch*, falls $A = A^T$ gilt.
Wir identifizieren den Raum der *symmetrischen Bilinearformen* $a\colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ mit dem Raum symmetrischen Matrizen A vermöge der Abbildung $a \mapsto A = \left(A_{jk}\right)_{j,k=1}^n$, wobei $A_{jk} = a(e_j, e_k)$, und wir die Bilinearform a aus A durch

$$a(h, k) = \sum_{j,k=1}^n A_{jk} h_j h_k$$

für $h = (h_1, \dots, h_n)^T$, $k = (k_1, \dots, k_n)^T \in \mathbb{R}^n$ zurückgewinnen.

9.4 Die Taylorformel

Sei $U \subseteq \mathbb{R}^n$ offen $f\colon U \to \mathbb{R}$.

Satz 9.18. *Sei f $(b+1)$-mal differenzierbar. Liegt dann die Strecke von x^0 nach $x^0 + h$ vollständig in U, so gilt*

$$f(x^0+h) = \sum_{a=0}^{b} \frac{D^a f(x^0)}{a!} \underbrace{(h,\dots,h)}_{a-\text{mal}} + \frac{D^{b+1} f(x^0+\theta h)}{(b+1)!} \underbrace{(h,\dots,h)}_{(b+1)-\text{mal}}$$

für ein $\theta \in (0,1)$.

Beweis. Die Funktion $g\colon [0,1] \to \mathbb{R}$, $t \mapsto f(x^0+th)$ ist $(b+1)$-mal differenzierbar, wobei

$$g^{(a)}(t) = D^a f(x^0+th) \underbrace{(h,\dots,h)}_{a-\text{mal}}$$

für $a \le b+1$. Die Behauptung ergibt sich aus der Taylorformel mit Lagrangeschem Restglied gemäß Satz 6.31:

$$g(1) = \sum_{a=0}^{b} \frac{g^{(a)}(0)}{a!} + \frac{g^{(b+1)}(\theta)}{(b+1)!}$$

für ein $\theta \in (0,1)$. □

Bemerkung. Möchte man vektorwertig arbeiten, also $f\colon U \to \mathbb{R}^m$ mit möglichem $m > 1$, so kann man unter der Voraussetzung, dass f $(b+1)$-mal *stetig* differenzierbar ist, den Satz 6.28 benutzen:

$$\begin{aligned} f(x^0+h) = {} & \sum_{a=0}^{b} \frac{D^a f(x^0)}{a!} (h,\dots,h) \\ & + \frac{1}{b!}\left[\int_0^1 (1-t)^b D^{(b+1)} f(x^0+th)\,dt\right] \underbrace{(h,\dots,h)}_{(b+1)-\text{mal}}. \end{aligned}$$

Dabei ist $\left[\int_0^1 \dots\, dt\right]$ eine $(b+1)$-multilineare Abbildung.
Benötigt man das Restglied nicht explizit, so kann man letzteres auch als

$$f(x^0+h) = \sum_{a=0}^{b} \frac{D^a f(x^0)}{a!} (h,\dots,h) + o\left(|h|^{b+1}\right)$$

für $h \to 0$ schreiben.

Die Formel im vorigen Satz wird typischerweise anders notiert. Dazu einige *Bezeichnungen*: Statt $\frac{\partial f}{\partial x_k}$ bzw. f_{x_k} schreiben wir auch $\partial_{x_k} f$ (bzw. $\partial_k f$). Ist dann $\alpha = (\alpha_1,\dots,\alpha_n) \in \mathbb{N}_0^n$ ein sogenannter *Multiindex*, so bezeichnet $\partial^\alpha f$ (bzw.

$\partial_x^\alpha f$) die partielle Ableitung $\dfrac{\partial^{|\alpha|} f}{\partial x_1^{\alpha_1} \dots \partial x_n^{\alpha_n}}$, wobei $|\alpha| = \alpha_1 + \cdots + \alpha_n$ die Länge des Multiindex α ist. An weiteren Bezeichnungen benötigen wir $h^\alpha = h_1^{\alpha_1} \cdots h_n^{\alpha_n}$ für $h = (h_1, \dots, h_n)^T \in \mathbb{R}^n$ sowie $\alpha! = \alpha_1! \cdots \alpha_n!$.

Satz 9.19. *Sei $f\colon U \to \mathbb{R}$ $(b+1)$-mal stetig (partiell) differenzierbar*[4]. *Die Strecke von x^0 nach $x^0 + h$ möge ganz in U liegen. Dann gilt*

$$f(x^0 + h) = \sum_{|\alpha| \le b} \frac{\partial^\alpha f(x^0)}{\alpha!} h^\alpha + \sum_{|\alpha| = b+1} \frac{\partial^\alpha f(x^0 + \theta h)}{\alpha!} h^\alpha$$

für ein $\theta \in (0, 1)$.

Beweis. Es bleibt lediglich festzustellen, dass für $x \in U$ und $a \le b + 1$

$$D^a f(x) \underbrace{(h, \dots, h)}_{a-\text{mal}} = \sum_{|\alpha| = a} \frac{a!}{\alpha!} \partial^\alpha f(x) h^\alpha$$

gilt. Dies ergibt sich für $h = (h_1, \dots, h_n)^T \in \mathbb{R}^n$ wie folgt:

$$\begin{aligned} D^a f(x)(h, \dots, h) &= D^a f(x) \left(\sum_{k_1=1}^n h_{k_1} e_{k_1}, \dots, \sum_{k_a=1}^n h_{k_a} e_{k_a} \right) \\ &= \sum_{k_1, \dots, k_a = 1}^n h_{k_1} \cdots h_{k_a} D^\alpha f(x)(e_{k_1}, \dots, e_{k_a}) \\ &= \sum_{k_1, \dots, k_a = 1}^n h_{k_1} \cdots h_{k_a} \frac{\partial^a f}{\partial x_{k_1} \dots \partial x_{k_a}}(x) \\ &= \sum_{|\alpha| = a} \frac{a!}{\alpha!} \partial^\alpha f(x) h^\alpha, \end{aligned}$$

wobei die numerische Konstante $\frac{a!}{\alpha!}$ sich kombinatorisch[5] ergibt. □

9.5 Extremwertbestimmung

Im Weiteren sei $U \subseteq \mathbb{R}^n$ offen, $f\colon U \to \mathbb{R}$. Wir möchten lokale Extrema von f bestimmen.

Definition 9.20. f nimmt an der Stelle $x^0 \in U$ ein *lokales Maximum* (*Minimum*) an, falls es ein $\delta > 0$ gibt, so dass $f(x^0) \ge f(x)$ $(f(x^0) \le f(x))$ für alle $x \in U$ mit $|x - x^0| \le \delta$ gilt.

Im Folgenden beweisen wir die Resultate für lokale Maxima, die Resultate für lokale Minima erhält man entsprechend (indem man z. B. f durch $-f$ ersetzt).

[4] Aus den Sätzen 9.16 und 9.17 folgt, dass für f die b-malige stetige Differenzierbarkeit in U und die b-malige stetige partielle Differenzierbarkeit in U äquivalent sind.

[5] $\frac{a!}{\alpha_1! \dots \alpha_n!}$ ist die Anzahl der Permutationen von a Elementen, die in n Gruppen zu jeweils $\alpha_1, \dots, \alpha_n$ gleichen Elementen eingeteilt sind.

9.5.1 Erste Ableitungen

Satz 9.21. *Ist f an der Stelle $x^0 \in U$ differenzierbar und hat f an x^0 ein lokales Maximum (Minimum), so gilt $\nabla f(x^0) = 0$.*

Beweis. Es gilt
$$f(x^0 + h) = f(x^0) + \nabla f(x^0)h + \varrho(h)$$
für $|h|$ klein mit $\lim_{h\to 0} \frac{\varrho(h)}{|h|} = 0$. Ist $\nabla f(x^0) \neq 0$, so gibt es ein $h_0 \in \mathbb{R}^n$, $|h_0| = 1$ mit $\nabla f(x^0)h_0 > 0$. Wir finden dann ein $\lambda_0 > 0$ mit $\left|\frac{\varrho(\lambda h_0)}{\lambda}\right| < \nabla f(x^0)h_0$ für $0 < \lambda \leq \lambda_0$. Es folgt
$$f(x^0 + \lambda h_0) \geq f(x^0) + \lambda \nabla f(x^0)h_0 - |\varrho(\lambda h_0)| > f(x^0)$$
für $0 < \lambda \leq \lambda_0$. Somit hat f an x^0 kein lokales Maximum. □

Beispiel 9.22. Sei $n = 2$ und
$$f(x, y) = x^2 + y^2 + xy^2.$$
Dann ist
$$f_x(x, y) = 2x + y^2, \quad f_y(x, y) = 2(1 + x)y.$$
Folglich liegen lokale Extrema höchstens an den Stellen $x = y = 0$ und $x = -1$, $y = \pm\sqrt{2}$ vor. Ob an einer dieser Stellen tatsächlich ein lokales Extremum vorliegt, werden wir im nächsten Abschnitt entscheiden.

9.5.2 Zweite Ableitungen

Sei jetzt f in U differenzierbar und $\nabla f(x^0) = 0$. Zuerst behandeln wir eine weitere notwendige Bedingung für das Vorliegen einer Extremstelle.

Für reelle symmetrische Matrizen A, $B \in M(n; \mathbb{R})$ schreiben wir $A \geq 0$, falls A nur nichtnegative Eigenwerte hat, und $A \geq B$, falls $A - B \geq 0$.[6]

Definition 9.23. Sei A eine reelle symmetrische Matrix. Dann:

- A heißt *positiv* (*negativ*) *semidefinit*, falls $A \geq 0$ ($A \leq 0$).
- A heißt *positiv* (*negativ*) *definit*, falls $A \geq cI$ ($A \leq -cI$) für ein $c > 0$. Man schreibt $A > 0$ ($A < 0$).
- A heißt *indefinit*, falls A weder positiv noch negativ semidefinit ist.[7]

[6] Eine reelle symmetrische Matrix $A \in M(n; \mathbb{R})$ hat halbeinfache reelle Eigenwerte und ist mittels einer orthogonalen Matrix $U \in O(n; \mathbb{R})$ diagonalisierbar. Das heißt, es gibt eine Matrix U mit $UU^T = I$ und $UAU^T = \begin{pmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{pmatrix}$, wobei die $\lambda_j \in \mathbb{R}$ die Eigenwerte von A sind.

[7] Äquivalente Bedingungen sind:

Satz 9.24. *Ist f an der Stelle x^0 zweimal differenzierbar und hat f an x^0 ein lokales Maximum (Minimum), so ist die Hesse-Matrix $D^2f(x^0)$ negativ (positiv) semidefinit.*

Beweis. Dieser Beweis ist ähnlich dem Beweis des vorigen Satzes.
Es gilt

$$f(x^0+h) = f(x^0) + \frac{1}{2}\,D^2f(x^0)(h,h) + \varrho(h)$$

für $|h|$ klein, wobei $\lim_{h\to 0} \frac{\varrho(h)}{|h|^2} = 0$. Ist $D^2f(x^0) \nleq 0$, so gibt es ein $h_0 \in \mathbb{R}^n$, $|h_0| = 1$ mit $D^2f(x^0)(h_0,h_0) > 0$. Es gibt dann ein $\lambda_0 > 0$ mit $\left|\frac{\varrho(\lambda h_0)}{\lambda^2}\right| < \frac{1}{2}\,D^2f(x^0)(h_0,h_0)$ für $0 < \lambda \leq \lambda_0$. Wir erhalten

$$f(x^0+\lambda h_0) \geq f(x^0) + \frac{\lambda^2}{2}\,D^2f(x^0)(h_0,h_0) - |\varrho(\lambda h_0)| > f(x^0)$$

für $0 < \lambda \leq \lambda_0$. Folglich hat f kein lokales Maximum an x^0. □

Folgerung. Ist $Df(x^0) = 0$, jedoch $D^2f(x^0)$ indefinit, so nimmt f an x^0 keinen lokalen Extremwert an. Man sagt, dass an einer solchen Stelle ein *Sattelpunkt* vorliegt.

Beispiel 9.25. In unserem obigen Beispiel gilt

$$D^2f(x,y) = 2\begin{pmatrix} 1 & y \\ y & 1+x \end{pmatrix}.$$

Die Eigenwerte sind

$$\lambda_\pm = 2 + x \pm \sqrt{x^2+4y^2}.$$

Folglich ist die Hesse-Matrix an $x=-1$, $y=\pm\sqrt{2}$ indefinit (die Eigenwerte sind -2, 4). Somit liegt an diesen Stellen ein Sattelpunkt vor.

Ein hinreichendes Kriterium für das Vorliegen einer Extremstelle ist die Definitheit der Hesse-Matrix.

Satz 9.26. *Ist f an der Stelle x^0 zweimal differenzierbar, ist $\nabla f(x^0) = 0$ und ist $D^2f(x^0)$ negativ definit (positiv definit), so hat f an der Stelle x^0 ein lokales Maximum (Minimum).*

Beweis. Sei $D^2f(x^0)(h,h) \leq -c|h|^2$ für ein $c > 0$. Wir wählen ein $\delta > 0$, so dass $|\varrho(h)| < \frac{c}{2}\,|h|^2$ für $|h| \leq \delta$, wobei $\varrho(h)$ den Rest in der Taylorformel bezeichnet, d. h.

$$f(x^0+h) = f(x^0) + \frac{1}{2}\,D^2f(x^0)(h,h) + \varrho(h).$$

- A ist positiv (negativ) semidefinit, falls $\langle Ah,h\rangle \geq 0$ ($\langle Ah,h\rangle \leq 0$) für alle $h \in \mathbb{R}^n$.
- A ist positiv (negativ) definit, falls $\langle Ah,h\rangle \geq c\,|h|^2$ ($\langle Ah,h\rangle \leq -c\,|h|^2$) für alle $h \in \mathbb{R}^n$ und ein $c > 0$.
- A ist indefinit, falls es h_1, $h_2 \in \mathbb{R}^n$ mit $\langle Ah_1,h_1\rangle > 0$ und $\langle Ah_2,h_2\rangle < 0$ gibt.

Hierbei ist $\langle\ ,\ \rangle\colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ das durch $\langle x,y\rangle = x^Ty$ definierte *Euklidische Skalarprodukt* in $\mathbb{R}^n$. Insbesondere gilt $\langle x,x\rangle = |x|^2$ für alle $x \in \mathbb{R}^n$.

für $|h|$ hinreichend klein. Für $|h| \leq \delta$ erhalten wir dann

$$f(x^0 + h) \leq f(x^0) - \frac{c}{2}|h|^2 + |\varrho(h)| < f(x^0).$$

Folglich ist $f(x^0)$ lokales Maximum an der Stelle x^0. $\square$

Beispiel 9.27. In unserem Beispiel ist die Hesse-Matrix an $x = y = 0$ positiv definit (beide Eigenwerte sind 2), also haben wir an dieser Stelle ein lokales Minimum.

9.5.3 Extremwertbestimmung unter Nebenbedingungen

Sei jetzt zusätzlich eine stetig differenzierbare Abbildung $g\colon U \to \mathbb{R}^r$ für ein $r < n$ gegeben. Wir wollen Extrema der Funktion f unter der *Nebenbedingung* $g(x) = 0$ bestimmen.

Definition 9.28. Sei $x^0 \in U$ mit $g(x^0) = 0$ [8]. Dann hat f an der Stelle x^0 ein *lokales Maximum* (*Minimum*) *unter der Nebenbedingung* $g = 0$, falls es ein $\delta > 0$ gibt, so dass $f(x^0) \geq f(x)$ ($f(x^0) \leq f(x)$) für alle $x \in U$ mit $g(x) = 0$ und $\left|x - x^0\right| \leq \delta$ gilt.

Wir schreiben $g = (g_1, \dots, g_r)^T$.

Satz 9.29. *Sei f an $x^0 \in U$ differenzierbar, $g(x^0) = 0$. Weiter habe f an $x^0 \in U$ ein lokales Maximum (Minimum) unter der Nebenbedingung $g = 0$. Die lineare Abbildung $Dg(x^0) \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^r)$ habe vollen Rang* [9]. *Dann gibt es $\lambda_1, \dots, \lambda_r \in \mathbb{R}$ mit*

$$\nabla f(x^0) = \sum_{k=1}^{r} \lambda_k \nabla g_k(x^0),$$

d. h. der Gradient von f an der Stelle x^0 ist Linearkombination der Gradienten der g_k an der Stelle x^0.

Die $\lambda_1, \dots, \lambda_r$ heißen *Lagrange-Multiplikatoren*.

Beispiel 9.30. Wir betrachten die durch die Gleichung

$$g(x) = |x|^2 - 1 = x_1^2 + \dots + x_n^2 + x_{n+1}^2 - 1 = 0$$

gegebene *n-Sphäre* $\mathbb{S}^n \subset \mathbb{R}^{n+1}$ und die Höhenfunktion

$$f(x) = x_{n+1}.$$

[8] Das heißt, x^0 erfüllt die Nebenbedingung.

[9] Wegen $r < n$ heißt das, dass die lineare Abbildung $Dg(x^0)\colon \mathbb{R}^n \to \mathbb{R}^r$ surjektiv ist bzw. die zugehörige $(r \times n)$–Matrix $\left(\frac{\partial g_k}{\partial x_j}(x^0)\right)_{\substack{k=1,\dots,r,\\ j=1,\dots,n}}$ den maximal möglichen Rang r hat. Das wiederum heißt, dass die r Zeilen $\nabla g_k(x^0)$ für $k = 1, \dots, r$ dieser Matrix linear unabhängig sind.

Es gilt

$$\nabla f(x) = (\underbrace{0, \dots, 0}_{n-\mathrm{mal}}, 1), \quad \nabla g(x) = (2x_1, 2x_2, \dots, 2x_{n+1}) = 2x.$$

Die Bedingung $\nabla f(x) = \lambda \nabla g(x)$ für ein $\lambda \in \mathbb{R}$, d. h.

$$(0, \dots, 0, 1) = 2\lambda x,$$

impliziert $x_1 = \cdots = x_n = 0$ und damit $x_{n+1} = \pm 1$ wegen $|x| = 1$ (dann $\lambda = \pm 1/2$). Somit liegen mögliche Extrema der Höhenfunktion an den Stellen $(0, \dots, 0, 1)$ (Nordpol) und $(0, \dots, 0, -1)$ (Südpol) vor.
Tatsächlich nimmt f auf $\mathbb{S}^n$ an $(0, \dots, 0, 1)$ ein Maximum und an $(0, \dots, 0, -1)$ ein Minimum an.

Wir beginnen den Beweis des Satzes mit folgendem Lemma:

Lemma 9.31. *Es gibt einen Koordinatenwechsel $x \mapsto y = y(x)$ nahe x^0, so dass $y(x^0) = 0$ und $g_k(x) = y_{n-r+k}(x)$ für $k = 1, \dots, r$.*

Wir werden dieses Ergebnis im übernächsten Abschnitt 9.7 beweisen. Dort werden wir auch sehen, dass die im Satz genannten Bedingungen *invariant unter Koordinatenwechseln* sind.

Beweis des Satzes. Wir schreiben $x = (x', x'')^T$ mit $x' = (x_1, \dots, x_{n-r})^T$, $x'' = (x_{n-r+1}, \dots, x_n)^T$. Obiges Lemma bedeutet, dass wir o. B. d. A. $x^0 = 0$ und $g(x) = x''$ annehmen dürfen.
Wir fragen uns, ob die Funktion $h(x') = f(x', 0)$ an $x' = 0$ ein lokales Extremum hat. Notwendig dafür ist $\nabla_{x'} f(0, 0) = \nabla_{x'} h(0) = 0$, also

$$\begin{aligned} D_x f(0,0) = (0, \nabla_{x''} f(0,0)) &= \sum_{k=1}^{r} \frac{\partial f}{\partial x_{n-r+k}}(0,0)\, e_{n-r+k} \\ &= \sum_{k=1}^{r} \lambda_k \nabla_x g_k(0,0), \end{aligned}$$

wobei $\lambda_k = \frac{\partial f}{\partial x_{n-r+k}}(0,0)$.[10] □

Beispiel 9.32. Extremstellen der Funktion $f(x, y) = x^2 + \frac{y^4}{2}$ auf der Ellipse $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$, $a > 0$, $b > 0$. Die Nebenbedingung ist $g = 0$ mit $g(x, y) = \frac{x^2}{a^2} + \frac{y^2}{b^2} - 1$, also wegen

$$\nabla f(x, y) = (2x, 2y^3), \quad \nabla g(x, y) = \left(\frac{2x}{a^2}, \frac{2y}{b^2} \right)$$

ist

$$\left(x, y^3\right) = \lambda \left(\frac{x}{a^2}, \frac{y}{b^2} \right)$$

für ein $\lambda \in \mathbb{R}$ eine notwendige Bedingung für das Vorliegen einer Extremstelle.
Wir unterscheiden die Fälle $x = 0$, dann $y = 0$ und schließlich $x \neq 0$, $y \neq 0$:

[10] Die Formeln zur Berechung der Werte der λ_k sind nicht *invariant* unter Koordinatentransformationen.

Fall 1: $x = 0$. Dann ist $\lambda = b^4$, $y = \pm b$ und $f(0, \pm b) = \frac{b^4}{2}$.

Fall 2: $y = 0$. Dann ist $\lambda = a^2$, $x = \pm a$ und $f(\pm a, 0) = a^2$.

Fall 3: $x \neq 0$, $y \neq 0$. Dann ist $\lambda = a^2$, $x = \pm a\sqrt{1 - \frac{a^2}{b^4}}$, $y = \pm\frac{a}{b}$ (mit voneinander unabhängig gewählten Vorzeichen) und

$$f\left(\pm a\sqrt{1 - \frac{a^2}{b^4}}, \pm\frac{a}{b}\right) = a^2\left(1 - \frac{a^2}{2b^4}\right).$$

Dieser Fall ist nur für $b^4 \geq a^2$ möglich.

Es gilt

$$\begin{aligned} a^2 &> a^2\left(1 - \frac{a^2}{2b^4}\right), \\ \frac{b^4}{2} &\geq a^2\left(1 - \frac{a^2}{2b^4}\right), \end{aligned}$$

wobei im zweiten Fall Gleichheit genau für $a^2 = b^4$ eintritt. Damit ergibt sich folgende Tabelle:

	Minimalstellen	Maximalstellen
$b^4 < a^2$	$(0, \pm b)$	$(\pm a, 0)$
$\frac{b^4}{2} < a^2 \leq b^4$	$\left(\pm a\sqrt{1 - \frac{a^2}{b^4}}, \pm\frac{a}{b}\right)$	$(\pm a, 0)$
$a^2 = \frac{b^4}{2}$	$\left(\pm\frac{a}{\sqrt{2}}, \pm\frac{b}{\sqrt{2}}\right)$	$(\pm a, 0), (0, \pm b)$
$a^2 < \frac{b^4}{2}$	$\left(\pm a\sqrt{1 - \frac{a^2}{b^4}}, \pm\frac{a}{b}\right)$	$(0, \pm b)$

9.6 Der Satz über die implizite Funktion

Wir beginnen mit dem Satz über die lokale Umkehrfunktion.

Satz 9.33 (über die inverse Funktion)**.** *Sei $U \subseteq \mathbb{R}^n$ offen, $f\colon U \to \mathbb{R}^n$ stetig differenzierbar. Sei weiterhin $x^0 \in U$, $Df(x^0) \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^n)$ invertierbar. Dann gibt es offene Umgebungen $U' \subseteq U$ von x^0 und $V' = f(U')$ von $f(x^0)$, so dass f eingeschränkt auf U' invertierbar ist, d. h. $f\big|_{U'}\colon U' \to V' = f(U')$ ist bijektiv. Die Umkehrabbildung $g = \left(f\big|_{U'}\right)^{-1}\colon V' \to U'$ ist stetig differenzierbar und*

$$(Dg)(y) = Df(x)^{-1}, \quad y \in V'.$$

wobei $x = g(y)$.

Beweis. Indem wir f durch $Df(x^0)^{-1}\left(f(x^0 + x) - f(x^0)\right)$ ersetzen, dürfen wir $x^0 = 0$, $f(x^0) = 0$ und $Df(x^0) = I$ annehmen.

1. Wir zeigen zuerst, dass f lokal invertierbar ist. Dazu wählen wir ein $r > 0$ mit $\bar{B}(0,2r) = \{x \in \mathbb{R}^n \mid |x| \leq 2r\} \subseteq U$ und $|(I - Df(x))\, h| \leq |h|/2$ für $|x| \leq 2r$ und $h \in \mathbb{R}^n$.

   Sei nun $y \in \mathbb{R}^n$, $|y| \leq r$. Wir zeigen, dass die Gleichung $f(x) = y$ eine eindeutige Lösung $x \in \mathbb{R}^n$ mit $|x| \leq 2r$ besitzt. Dazu benutzen wir den Banachschen Fixpunktsatz und betrachten

   $$T_y x = x - f(x) + y.$$

   Es gilt $T_y\colon \bar{B}(0,2r) \to \bar{B}(0,2r)$ wegen

   $$|x - f(x)| \leq \int_0^1 |(I - Df(sx))\, x|\, ds \leq \frac{|x|}{2},$$

   also

   $$|T_y x| \leq |x - f(x)| + |y| \leq 2r.$$

   Weiterhin ist T_y eine Kontraktion auf $\bar{B}(0,2r)$ wegen

   $$|T_y x - T_y x'| \leq \int_0^1 |(I - Df(x' + s(x - x')))\,(x - x')|\, ds \leq \frac{|x - x'|}{2}.$$

   Wegen der Vollständigkeit von $\bar{B}(0,2r)$ folgt die Behauptung.

2. Wir setzen nun $V' = B(0,r)$ und $U' = f^{-1}(B(0,r)) \subseteq B(0,2r)$. U' ist eine offene Umgebung der 0 wegen der Stetigkeit von f und $f\big|_{U'}\colon U' \to V'$ ist bijektiv nach Konstruktion.

3. Schließlich zeigen wir, dass $g = \left(f\big|_{U'}\right)^{-1}\colon V' \to U'$ stetig differenzierbar ist und $Dg(y) = Df(x)^{-1}$ für $x = g(y)$ gilt.

   Angenommen, wir wissen bereits, dass g differenzierbar ist. Dann folgt aus $g(f(x)) = x$ für $x \in U'$ und der Kettenregel, dass $Dg(f(x))\, Df(x) = I$ gilt, also $Dg(f(x)) = Df(x)^{-1}$. Da $Df\colon U' \to \mathcal{L}(\mathbb{R}^n, \mathbb{R}^n)$ stetig und, wie wir gerade gezeigt haben, auch überall invertierbar ist, folgt, dass $Dg\colon V' \to \mathcal{L}(\mathbb{R}^n, \mathbb{R}^n)$ ebenfalls stetig ist.

   Somit bleibt zu zeigen, dass g differenzierbar ist. Da die Stelle $x^0 = 0$ in keiner Weise ausgezeichnet ist, genügt es zu zeigen, dass g an 0 differenzierbar ist. Dazu schreiben wir mit $f(0) = 0$ und $Df(0) = I$

   $$f(x) = x + o(|x|)$$

   für $x \to 0$, folglich mit $x = g(y)$

   $$y = g(y) + o(|g(y)|)$$

   beziehungsweise

   $$g(y) = y + o(|y|)$$

   für $y \to 0$. Es folgt, dass g an 0 differenzierbar ist und $Dg(0) = I$ gilt.

□

Wir schreiben jetzt $(x,y)^T$ mit $x \in \mathbb{R}^n$, $y \in \mathbb{R}^k$ für Stellen in $\mathbb{R}^{n+k}$.

Satz 9.34 (über die implizite Funktion)**.** *Sei $U \subseteq \mathbb{R}^{n+k}$ offen, $f\colon U \to \mathbb{R}^k$ stetig differenzierbar, $(x^0, y^0)^T \in U$ mit $f(x^0, y^0) = 0$ und $D_y f(x^0, y^0) \in \mathcal{L}(\mathbb{R}^k, \mathbb{R}^k)$ invertierbar. Dann gibt es offene Umgebungen $V \subseteq \mathbb{R}^n$ von x^0 und $W \subseteq \mathbb{R}^k$ von y^0 mit $V \times W \subseteq U$ und eine stetig differenzierbare Funktion $g\colon V \to W$ mit $g(x^0) = y^0$, so dass*

$$f(x,y) = 0 \text{ für } (x,y)^T \in V \times W \iff y = g(x).$$

Insbesondere gilt $f(x, g(x)) = 0$ für alle $x \in V$.

Beweis. Wir betrachten die Funktion $F\colon U \to \mathbb{R}^{n+k}$, $F(x,y) = (x, f(x,y))^T$. Dann ist

$$DF(x,y) = \begin{pmatrix} I & 0 \\ D_x f(x,y) & D_y f(x,y) \end{pmatrix}$$

an der Stelle $(x^0, y^0)^T$ als lineare Abbildung von $\mathbb{R}^n \times \mathbb{R}^k$ nach $\mathbb{R}^n \times \mathbb{R}^k$ invertierbar. Folglich existieren nach dem Satz über die inverse Funktion offene Umgebungen $\widetilde{U} \subseteq U$ von $(x^0, y^0)^T$ und $\widetilde{V} \subseteq \mathbb{R}^{n+k}$ von $(x^0, 0)^T$ sowie die Umkehrabbildung $G\colon \widetilde{V} \to \widetilde{U}$ zu $F\big|_{\widetilde{U}}\colon \widetilde{U} \to \widetilde{V}$.
G hat notwendigerweise die Gestalt $G(x,y) = (x, h(x,y))^T$ für eine Funktion $h\colon \widetilde{V} \to \mathbb{R}^k$ und es gilt

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

für $(x,y)^T \in \widetilde{V}$.
Nach Konstruktion gilt $f(x,y) = 0$ für $(x,y)^T \in \widetilde{U}$ genau dann, wenn $(x,0)^T \in \widetilde{V}$ und $y = h(x,0)$. Damit können wir jetzt $V \subseteq \{x \in \mathbb{R}^n \mid (x,0)^T \in \widetilde{V}\}$, $W = \{h(x,0) \mid x \in V\}$ und $g(x) = h(x,0)$ für $x \in V$ setzen, wobei wir die Umgebung V von x^0 so klein wählen, dass $V \times W \subseteq U$ gilt. □

Bemerkung. Mit

$$D_x f(x,y) + D_y f(x,y) Dg(x) = 0,$$

wobei $y = g(x)$, berechnet sich die Ableitung von g zu

$$Dg(x) = -D_y f(x,y)^{-1} D_x f(x,y).$$

9.7 Differentiation in krummlinigen Koordinaten

Sei $U \subseteq \mathbb{R}^n$ offen. Eine Funktion $f\colon U \to \mathbb{R}$ heißt von der *Klasse* C^l, wobei $l \in \mathbb{N}_0 \cup \{\infty\}$, falls f in U l-mal stetig differenzierbar ist[11]. Eine Vektorfunktion heißt von der Klasse C^l, falls jede ihrer Komponenten von der Klasse C^l ist.

[11] Für $l = 0$ ist $f\colon U \to \mathbb{R}$ stetig.

Definition 9.35. Ein C^l-*Koordinatensystem in* U, wobei $l \geq 1$, sind n skalare Funktionen $y_k \colon U \to \mathbb{R}$ für $k = 1, \ldots, n$ der Klasse C^l, so dass die Abbildung

$$y = (y_1, \ldots, y_n)^T \colon U \to \mathbb{R}^n$$

injektiv und $Dy(x) \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^n)$ für alle $x \in U$ invertierbar ist.

Dann ist die Abbildung $y \colon U \to \mathbb{R}^n$ bijektiv auf ihr Bild $V = y(U)$, $V \subseteq \mathbb{R}^n$ ist offen und die Umkehrabbildung $V \to U$, $y \mapsto x$, ist ebenfalls von der Klasse C^l. Man nimmt die *Werte* der Funktionen $y_1, \ldots, y_n$ an einer Stelle $x \in U$ als *Koordinaten* von x, d. h. um die Stelle x zu adressieren.

Beispiel 9.36.

1. *Euklidische Koordinaten*: Die Koordinatenfunktionen sind $x_1, \ldots, x_n$.

2. *Polarkoordinaten im* $\mathbb{R}^2$: Euklidische Koordinaten sind x, y, als U nehmen wir den $\mathbb{R}^2$ ohne einen vom Ursprung ausgehenden Strahl, z. B.

$$U = \mathbb{R}^2 \setminus \big\{(x, 0) \in \mathbb{R}^2 \bigm| x \leq 0\big\}.$$

Wir führen neue Koordinaten $(r, \varphi) \in \mathbb{R}_+ \times (-\pi, \pi)$ vermöge

$$x = r\cos\varphi, \quad y = r\sin\varphi$$

ein.

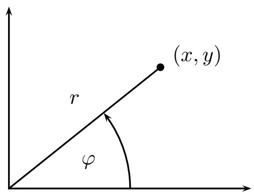

Abbildung 9.1: Polarkoordinaten

Die Abbildung

$$\begin{aligned} \mathbb{R}_+ \times (-\pi, \pi) &\to \mathbb{R}^2 \setminus \big\{(x, 0) \bigm| x \leq 0\big\}, \\ (r, \varphi) &\mapsto (x, y), \end{aligned}$$

ist bijektiv. Zudem gilt

$$\det \begin{pmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \varphi} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \varphi} \end{pmatrix} = \det \begin{pmatrix} \cos\varphi & -r\sin\varphi \\ \sin\varphi & r\cos\varphi \end{pmatrix} = r > 0.$$

Tatsächlich ist $r^2 = x^2 + y^2$ und $\cot\varphi = \frac{\cos\varphi}{\sin\varphi} = \frac{x}{y}$ für $y \neq 0$, also

$$\begin{aligned} r = \sqrt{x^2 + y^2}, \qquad \varphi &= \operatorname{arccot}\frac{x}{y} && \text{für } y > 0 \\ \text{bzw. } \varphi &= -\pi + \operatorname{arccot}\frac{x}{y} && \text{für } y < 0 \end{aligned}$$

sowie $\varphi = 0$ für $x > 0$, $y = 0$.

3. *Zylinderkoordinaten im* $\mathbb{R}^3$: Euklidische Koordinaten sind x, y, z, wir nehmen $U = \mathbb{R}^3 \setminus \{(x, 0, z) \mid x \leq 0\}$. Wir führen neue Koordinaten $(r, \varphi, z) \in \mathbb{R}_+ \times (-\pi, \pi) \times \mathbb{R}$ vermöge
$$x = r\cos\varphi, \quad y = r\sin\varphi, \quad z = z$$
ein.

4. *Kugelkoordinaten im* $\mathbb{R}^3$ Wir nehmen U wie unter 3), die Koordinaten sind $(\varrho, \varphi, \theta) \in \mathbb{R}_+ \times (-\pi, \pi) \times (0, \pi)$ mit
$$x = \varrho\cos\varphi\sin\theta, \quad y = \varrho\sin\varphi\sin\theta, \quad z = \varrho\cos\theta.$$

Den Übergang von einem Koordinatensystem zu einem anderen bezeichnet man als *Koordinatenwechsel* (oder auch als Koordinatentransformation).

Beispiel 9.37. Koordinatenwechsel im $\mathbb{R}^2$, $(x, y) \mapsto (u, v)$

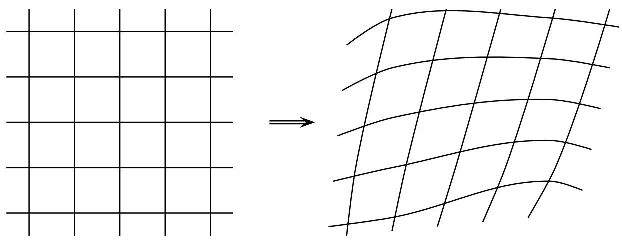

Abbildung 9.2: Übergang zu krummlinige Koordinaten

Lemma 9.38. *Sei* $y\colon U \to \mathbb{R}^n$ *von der Klasse* C^l, $x^0 \in U$ *und* $Dy(x^0) \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^n)$ *invertierbar. Dann kann man die* $y_1, \dots, y_n$ *nahe* x^0 *als* C^l*-Koordinaten nehmen.*

Beweis. Man benutze den Satz über die inverse Funktion. □

Lemma 9.39. *Seien* $y_1, \dots, y_r\colon U \to \mathbb{R}$ *für* $r < n$ C^l*-Funktionen,* $l \geq 1$, *so dass* $Dy(x^0) \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^r)$ *maximalen Rang hat (d. h. die Vektoren* $\nabla y_1(x^0), \dots, \nabla y_r(x^0)$ *sind linear unabhängig). Dann gibt es Funktionen*
$$y_{r+1}, \dots, y_n\colon U \to \mathbb{R},$$
so dass $(y_1, \dots, y_n)$ *ein* C^l*-Koordinatensystem nahe* x^0 *ist.*

Beweis. Es gibt $n - r$ Indizes i_k mit $1 \leq i_1 < i_2 < \cdots < i_{n-r} \leq n$, so dass
$$\det\left(\nabla y(x^0)^T, e_{i_1}, \dots, e_{i_{n-r}}\right) \neq 0.$$
Wir können dann $(y_1, \dots, y_r, x_{i_1}, \dots, x_{i_{n-r}})$ als C^l-Koordinatensystem nahe x^0 wählen. □

Lemma 9.40. *Seien x und y zwei Koordinatensysteme nahe x^0 mit Koordinatenwechseln $x \mapsto y = y(x)$ bzw. $y \mapsto x = x(y)$. Weiter sei f eine Funktion, die in x^0 differenzierbar ist. Dann gilt*

$$\nabla_y f(x^0) = \nabla_x f(x^0) \frac{\partial x}{\partial y} (x^0),$$

wobei $\frac{\partial x}{\partial y}$ die Jacobi-Matrix des Koordinatenwechsels $y \mapsto x$ bezeichnet.

Beweis. Wir schreiben f als Funktion von x, d. h. $f = f(x)$, und x als Funktion von y, d. h. $x = \varphi(y)$. Dann schreibt sich f als Funktion von y als $g(y) = f(\varphi(y))$. Wir erhalten mit $x^0 = \varphi(y^0)$

$$\nabla g(y^0) = \nabla f(x^0) D\varphi(y^0).$$

Das ist die Behauptung. □

Lemma 9.41. *Unter den Voraussetzungen des vorigen Lemmas sei f nahe x^0 stetig differenzierbar, $\nabla f(x^0) = 0$ (diese Bedingung ist unabhängig vom gewählten Koordinatensystem) und f sei an x^0 zweimal differenzierbar. Dann gilt*

$$D_y^2 f(x^0) = \frac{\partial x}{\partial y} (x^0)\, D_x^2 f(x^0) \frac{\partial x}{\partial y} (x^0)^T.$$

Insbesondere sind Definitheitsbedingungen an die Hesse-Matrix koordinateninvariant (unter der Bedingung, dass der Gradient verschwindet).

Beweis. Unter den Bezeichnungen des Beweises des vorigen Lemmas gilt

$$\frac{\partial^2 g}{\partial y_k \partial y_l} (y^0) = \sum_{i,j} \frac{\partial^2 f}{\partial x_i \partial x_j} (x^0) \frac{\partial \varphi_i}{\partial y_k} (y^0) \frac{\partial \varphi_j}{\partial y_l} (y^0),$$

und daraus folgt die Behauptung. □

Wir studieren schließlich das Verhalten von Differentialausdrücken unter Koordinatentransformation.

Definition 9.42. Sei x ein C^l-Koordinatensystem in U. Ein *Differentialausdruck* der Ordnung m in den Koordinaten $x_1, \ldots, x_n$ ist ein Ausdruck der Form

$$\sum_{|\alpha| \le m} a_\alpha(x) \partial_x^\alpha,$$

wobei $\partial_x^\alpha = \dfrac{\partial^{|\alpha|}}{\partial_{x_1}^{\alpha_1} \cdots \partial_{x_n}^{\alpha_n}}$ wie zuvor und die $a_\alpha(x)$ für $|\alpha| \le m$ Funktionen auf U sind.

Es sei jetzt ein weiteres C^l-Koordinatensystem $y_1, \ldots, y_n$ in U gegeben. Unter der Voraussetzung $m \leq l$ können wir dann einen gegebenen Differentialausdruck $\sum\limits_{|\alpha| \leq m} a_\alpha(x) \partial_x^\alpha$ auch als

$$\sum_{|\beta| \leq m} b_\beta(y) \partial_y^\beta$$

mit gewissen resultierenden Funktionen $b_\beta(y)$ auf U schreiben. Die Frage ist, wie sich die b_β aus den a_α berechnen.

Wir betrachten dazu einige Beispiele.

Beispiel 9.43.

1. Wie sich Ableitungen erster Ordnung transformieren, haben wir bereits gesehen:

$$\frac{\partial}{\partial x_j} = \frac{\partial y_1}{\partial x_j} \frac{\partial}{\partial y_1} + \cdots + \frac{\partial y_n}{\partial x_j} \frac{\partial}{\partial y_n}.$$

2. Einer der wichtigsten Differentialoperatoren der mathematischen Physik ist der *Laplace-Operator* $\Delta = \frac{\partial^2}{\partial x_1^2} + \frac{\partial^2}{\partial x_2^2} + \cdots + \frac{\partial^2}{\partial x_n^2}$. Wir wollen im Fall $n = 2$ diesen Operator in Polarkoordinaten schreiben.

   Aus $x = r \cos\varphi$, $y = r \sin\varphi$ folgt

$$\begin{aligned} \frac{\partial}{\partial r} &= \frac{\partial x}{\partial r} \frac{\partial}{\partial x} + \frac{\partial y}{\partial r} \frac{\partial}{\partial y} = \cos\varphi \frac{\partial}{\partial x} + \sin\varphi \frac{\partial}{\partial y} \\ \frac{\partial}{\partial \varphi} &= \frac{\partial x}{\partial \varphi} \frac{\partial}{\partial x} + \frac{\partial y}{\partial \varphi} \frac{\partial}{\partial y} = -r \sin\varphi \frac{\partial}{\partial x} + r \cos\varphi \frac{\partial}{\partial y} \end{aligned}$$

   bzw.

$$\begin{pmatrix} \frac{\partial}{\partial r} & \frac{\partial}{\partial \varphi} \end{pmatrix} = \begin{pmatrix} \frac{\partial}{\partial x} & \frac{\partial}{\partial y} \end{pmatrix} \begin{pmatrix} \cos\varphi & -r \sin\varphi \\ \sin\varphi & r \cos\varphi \end{pmatrix}$$
$$\begin{pmatrix} \frac{\partial}{\partial r} & \frac{\partial}{\partial \varphi} \end{pmatrix} \begin{pmatrix} \cos\varphi & \sin\varphi \\ -\frac{1}{r} \sin\varphi & \frac{1}{r} \cos\varphi \end{pmatrix} = \begin{pmatrix} \frac{\partial}{\partial x} & \frac{\partial}{\partial y} \end{pmatrix}$$

   oder

$$\begin{aligned} \frac{\partial}{\partial x} &= \cos\varphi \frac{\partial}{\partial r} - \frac{1}{r} \sin\varphi \frac{\partial}{\partial \varphi} \\ \partial \partial y &= \sin\varphi \frac{\partial}{\partial r} + \frac{1}{r} \cos\varphi \frac{\partial}{\partial \varphi}. \end{aligned}$$

Es folgt

$$\begin{aligned}\frac{\partial^2}{\partial x^2} &= \left(\cos\varphi\,\frac{\partial}{\partial r} - \frac{1}{r}\,\sin\varphi\,\frac{\partial}{\partial\varphi}\right)\left(\cos\varphi\,\frac{\partial}{\partial r} - \frac{1}{r}\,\sin\varphi\,\frac{\partial}{\partial\varphi}\right)\\
&= \cos^2\!\varphi\,\frac{\partial^2}{\partial r^2} - \frac{1}{r}\,\sin\varphi\cos\varphi\,\frac{\partial^2}{\partial r\partial\varphi} + \frac{1}{r^2}\,\sin\varphi\cos\varphi\,\frac{\partial}{\partial\varphi}\\
&- \frac{1}{r}\,\sin\varphi\cos\varphi\,\frac{\partial^2}{\partial r\partial\varphi} + \frac{1}{r}\sin^2\!\varphi\,\frac{\partial}{\partial r}\\
&+ \frac{1}{r^2}\sin^2\!\varphi\,\frac{\partial^2}{\partial\varphi^2} + \frac{1}{r^2}\,\sin\varphi\cos\varphi\,\frac{\partial}{\partial\varphi}\end{aligned}$$

$$\begin{aligned}\frac{\partial^2}{\partial y^2} &= \left(\sin\varphi\,\frac{\partial}{\partial r} + \frac{1}{r}\,\cos\varphi\,\frac{\partial}{\partial\varphi}\right)\left(\sin\varphi\,\frac{\partial}{\partial r} + \frac{1}{r}\,\cos\varphi\,\frac{\partial}{\partial\varphi}\right)\\
&= \sin^2\!\varphi\,\frac{\partial^2}{\partial r^2} + \frac{1}{r}\,\sin\varphi\cos\varphi\,\frac{\partial^2}{\partial r\partial\varphi} - \frac{1}{r^2}\,\sin\varphi\cos\varphi\,\frac{\partial}{\partial\varphi}\\
&+ \frac{1}{r}\,\sin\varphi\cos\varphi\,\frac{\partial^2}{\partial r\partial\varphi} + \frac{1}{r}\cos^2\!\varphi\,\frac{\partial}{\partial r}\\
&+ \frac{1}{r^2}\cos^2\!\varphi\,\frac{\partial^2}{\partial\varphi^2} - \frac{1}{r^2}\,\sin\varphi\cos\varphi\,\frac{\partial}{\partial\varphi},\end{aligned}$$

also ist

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

der Ausdruck des Laplace-Operators in Polarkoordinaten.

Kapitel 10

Weitere Kapitel

10.1 Komplexe Zahlen

10.1.1 Definition und Eigenschaften

Auf $\mathbb{R}^2$ definieren wir ein Produkt vermöge

$$(a, b) \cdot (c, d) = (ac - bd, ad + bc).$$

Versehen mit diesem Produkt und der üblichen (komponentenweisen) Addition bezeichnen wir $\mathbb{R}^2$ als $\mathbb{C}$.

Satz 10.1. (i) *$\mathbb{C}$ ist ein Körper, wobei Addition und Multiplikation stetig bezüglich der Standardtopologie des $\mathbb{R}^2$ sind.*

(ii) *Der Körper $\mathbb{R}$ bettet in $\mathbb{C}$ vermöge der Abbildung $a \mapsto (a, 0)$ ein, wobei $\mathbb{C}$ auf $\mathbb{R}$ die Standardtopologie induziert.*

Beweis. (i) Nachrechnen. Insbesondere ist $(1, 0)$ die Einheit bezüglich der Multiplikation und $(a, b)^{-1} = \left(\frac{a}{a^2+b^2}, -\frac{b}{a^2+b^2}\right)$ für $(a, b) \neq (0, 0)$.

(ii) Ebenfalls nachrechnen.

$\square$

Mit Hilfe des Ergebnisses in (ii) identifizieren wir Elemente a aus $\mathbb{R}$ mit Elementen $(a, 0)$ aus $\mathbb{C}$. Führen wir für $(0, 1)$ die Bezeichnung i als sogenannte *imaginäre Einheit* ein, so schreibt sich jedes $z \in \mathbb{C}$ in eindeutiger Weise als

$$z = a + b\mathrm{i} \quad \text{mit } a,\, b \in \mathbb{R}$$

(wegen $(a, b) = (a, 0) + (b, 0)(0, 1)$). Diese Darstellung heißt die *arithmetische Form* einer komplexen Zahl. $\mathbb{C}$ heißt der *Körper der komplexen Zahlen*, die Elemente von $\mathbb{C}$ komplexe Zahlen.

Definition 10.2. Sei $z = a + b\mathrm{i}$, $a, b \in \mathbb{R}$. Dann heißen

- $|z| = \sqrt{a^2 + b^2}$ — *Betrag* von z,
- $\operatorname{Re} z = a$ — *Realteil* von z,
- $\operatorname{Im} z = b$ — *Imaginärteil* von z,
- $\bar{z} = a - b\mathrm{i}$ — *konjugiert-komplexe Zahl*.

Lemma 10.3. *Für $z, w \in \mathbb{C}$ gilt*

(i) $|zw| = |z|\,|w|$, $|z + w| \leq |z| + |w|$,

(ii) $|z|^2 = z\bar{z}$, $|z| = |\bar{z}|$,

(iii) $\operatorname{Re} z = \frac{1}{2}(z + \bar{z})$, $\quad \operatorname{Im} z = \frac{1}{2\mathrm{i}}(z - \bar{z})$,

(iv) $\overline{zw} = \bar{z} \cdot \bar{w}$.

Geometrisch stellt man komplexe Zahlen in der *komplexen* oder Gaussschen *Zahlenebene* dar:

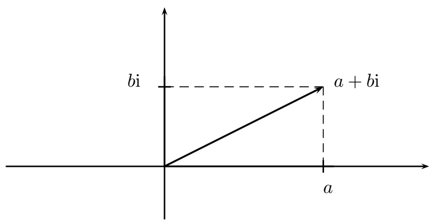

Abbildung 10.1: Die Gausssche Zahlenebene

Addition ist dann einfach Vektoraddition:

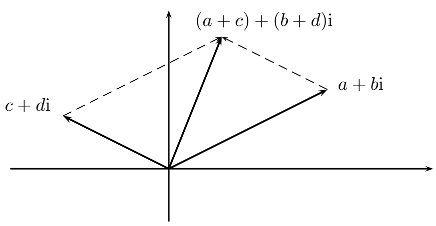

Abbildung 10.2: Addition komplexer Zahlen

Die geometrische Bedeutung der Multiplikation werden wir behandeln, nachdem wir die Exponentialform einer komplexen Zahl eingeführt haben.

Der *Hauptsatz der Algebra* besagt:

Theorem 10.4. *Jedes nicht-konstante komplexe Polynom*

$$p(z) = a_0 z^m + a_1 z^{m-1} + \cdots + a_{m-1} z + a_m \quad (a_0 \neq 0,\, m \geq 1)$$

besitzt wenigstens eine komplexe Nullstelle. Damit zerfällt $p(z)$ über $\mathbb{C}$ vollständig in Linearfaktoren, d. h. es gilt

$$p(z) = a_0 (z - \alpha_1)(z - \alpha_2) \cdots (z - \alpha_m),$$

wobei die $\alpha_j \in \mathbb{C}$ die (nicht notwendig verschiedenen) Nullstellen von $p(z)$ sind. Diese Darstellung ist bis auf die Reihenfolge der Faktoren eindeutig.

Beweis. In der Vorlesung „Funktionentheorie“. □

Beispielsweise gilt $\mathrm{i}^2 = -1$, somit sind $\mathrm{i}, -\mathrm{i}$ die beiden Nullstellen des Polynoms $z^2 + 1 = (z - \mathrm{i})(z + \mathrm{i})$.

10.1.2 Potenzreihen im Komplexen

Die Konvergenz von Folgen und Reihen ist wie im Reellen definiert (bzw. indem man $\mathbb{C}$ mit $\mathbb{R}^2$ identifiziert). Besonders interessant sind Potenzreihen:

Definition 10.5. Eine *Potenzreihe im Komplexen* ist eine Reihe der Form

$$\sum_{k=0}^{\infty} a_k \left(z - z^0\right)^k,$$

wobei $a_k \in \mathbb{C}$ für $k = 0, 1, 2, \ldots,$ $z^0 \in \mathbb{C}$ der Entwicklungspunkt und z die unabhängige Variable ist.

Wie im Reellen zeigt man:

Satz 10.6. (i) *Die Reihe* $\sum\limits_{k=0}^{\infty} a_k \left(z - z^0\right)^k$ *hat den* Konvergenzradius

$$r = \frac{1}{\limsup\limits_{k \to \infty} \sqrt[k]{|a_k|}},$$

d. h. die Reihe konvergiert für $\left|z - z^0\right| < r$ und divergiert für $\left|z - z^0\right| > r$. (Für $\left|z - z^0\right| = r$ ist keine allgemeine Aussage möglich.)

(ii) *Im Inneren $\left|z - z^0\right| < r$ ihres Konvergenzbereiches konvergiert die Reihe $\sum\limits_{k=0}^{\infty} a_k \left(z - z^0\right)^k$ absolut und lokal gleichmäßig gegen eine Grenzfunktion $f(z)$. Die Funktion $f(z)$ ist für $\left|z - z^0\right| < r$ beliebig oft differenzierbar.*[1]

[1] Tatsächlich ist die Funktion $f(z)$ für $\left|z - z^0\right| < r$ beliebig oft *komplex differenzierbar*, was eine wesentlich stärkere Eigenschaft als die bislang behandelte Differenzierbarkeit im Reellen ist. Das wird ebenfalls Thema der Vorlesung „Funktionentheorie“ sein.

10.1.3 Die Exponential- und trigonometrische Funktionen

Definition 10.7. Für $z \in \mathbb{C}$ setzen wir

$$\begin{aligned} \mathrm{e}^z = \exp z &= \sum_{k=0}^{\infty} \frac{z^k}{k!} = 1 + z + \frac{z^2}{2!} + \frac{z^3}{3!} + \dots, \\ \cos z &= \sum_{k=0}^{\infty} (-1)^k \frac{z^{2k}}{(2k)!} = 1 - \frac{z^2}{2!} + \frac{z^4}{4!} \mp \dots, \\ \sin z &= \sum_{k=0}^{\infty} (-1)^k \frac{z^{2k+1}}{(2k+1)!} = z - \frac{z^3}{3!} + \frac{z^5}{5!} \mp \dots. \end{aligned}$$

Die in dieser Definition vorkommenden Reihen konvergieren auf ganz $\mathbb{C}$ (der Konvergenzradius ist jeweils ∞), so gilt z. B.

$$\limsup_{k\to\infty} \sqrt[k]{1/k!} = \lim_{k\to\infty} \frac{1/(k+1)!}{1/k!} = \lim_{k\to\infty} \frac{1}{k+1} = 0.$$

Satz 10.8. *Für $x \in \mathbb{R}$ definieren e^x, $\cos x$ und $\sin x$ die üblichen Exponential-, Kosinus- und Sinusfunktionen.*

Beweis. Wir führen den Beweis für die Exponentialfunktion, die anderen beiden Fälle sind analog.
Für $x \in \mathbb{R}$ und $n \in \mathbb{N}_0$ schreiben wir unter Benutzung der Taylorformel und wegen $(\mathrm{e}^x)' = \mathrm{e}^x$, $\mathrm{e}^0 = 1$

$$\mathrm{e}^x = \sum_{k=0}^{n} \frac{x^k}{k!} + \mathrm{e}^{\vartheta x} \frac{x^{n+1}}{(n+1)!}$$

für ein $\vartheta \in (0,1)$. Somit gilt

$$\left| \mathrm{e}^x - \sum_{k=0}^{n} \frac{x^k}{k!} \right| \le \mathrm{e}^{\max\{x,0\}} \frac{|x|^{n+1}}{(n+1)!} \to 0 \quad \text{für } n \to \infty.$$

Folglich gilt

$$\mathrm{e}^x = \sum_{k=0}^{\infty} \frac{z^k}{k!}$$

für alle $x \in \mathbb{R}$. ☐

Satz 10.9. *Es gelten die üblichen Additionstheoreme, d. h. für z, $w \in \mathbb{C}$ gilt*

$$\begin{aligned} \mathrm{e}^{z+w} &= \mathrm{e}^z \cdot \mathrm{e}^w, \\ \cos(z \pm w) &= \cos z \cos w \mp \sin z \sin w, \\ \sin(z \pm w) &= \sin z \cos w \pm \cos z \sin w. \end{aligned}$$

Zudem ist die Kosinusfunktion gerade und die Sinusfunktion ungerade, d. h. für $z \in \mathbb{C}$ gilt

$$\cos(-z) = \cos z, \quad \sin(-z) = -\sin z.$$

Beweis. Wir zeigen lediglich $e^{z+w} = e^z \cdot e^w$, der Rest des Beweises ist analog. Unter Benutzung des Cauchyproduktes gilt

$$e^{z+w} = \sum_{l=0}^{\infty} \frac{(z+w)^l}{l!} = \sum_{l=0}^{\infty} \sum_{j+k=l} \frac{z^j}{j!} \frac{w^k}{k!} = \left(\sum_{j=0}^{\infty} \frac{z^j}{j!} \right) \left(\sum_{k=0}^{\infty} \frac{z^k}{k!} \right) = e^z \cdot e^w.$$

□

Folgerung. e^z verschwindet nirgends auf $\mathbb{C}$.

Beweis. $e^z \cdot e^{-z} = e^0 = 1$. □

Als Nächstes erhalten wir die wichtige *Eulerformel*:

Satz 10.10. *Für $z \in \mathbb{C}$ gilt*

$$e^{iz} = \cos z + i \sin z.$$

Beweis. Für $z \in \mathbb{C}$ gilt

$$\begin{aligned} e^{iz} &= \sum_{k=0}^{\infty} \frac{(iz)^k}{k!} = \sum_{k=0}^{\infty} \frac{i^{2k} z^{2k}}{(2k)!} + \sum_{k=0}^{\infty} \frac{i^{2k+1} z^{2k+1}}{(2k+1)!} \\ &= \sum_{k=0}^{\infty} (-1)^k \frac{z^{2k}}{(2k)!} + i \sum_{k=0}^{\infty} (-1)^k \frac{z^{2k+1}}{(2k+1)!} \\ &= \cos z + i \sin z. \end{aligned}$$

□

Folgerung. In der Formel

$$\boxed{e^{i\pi} = -1}$$

sind vier der wichtigsten Konstanten der Mathematik, nämlich 1, i, e, π, vereinigt.

Folgerung. Die Exponentialfunktion ist $2\pi i$-periodisch, Kosinus und Sinus sind 2π-periodisch.

Beweis. Es gilt $e^{z+2\pi i} = e^z \cdot e^{2\pi i} = e^z \cdot 1 = e^z$. Weiterhin gilt $e^{iz} = \cos z + i \sin z$, $e^{-iz} = \cos z - i \sin z$, also

$$\cos z = \frac{e^{iz} + e^{-iz}}{2}, \quad \sin z = \frac{e^{iz} - e^{-iz}}{2i}.$$

Es folgt

$$\cos(z + 2\pi) = \frac{e^{iz+2\pi i} + e^{-iz-2\pi i}}{2} = \frac{e^{iz} + e^{-iz}}{2} = \cos z,$$

analog für $\sin(z + 2\pi) = \sin z$. □

Lemma 10.11. *Es gilt* $\cos^2 z + \sin^2 z = 1$.

Beweis.

$$\cos^2 z + \sin^2 z = \left(\frac{\mathrm{e}^{\mathrm{i}z} + \mathrm{e}^{-\mathrm{i}z}}{2}\right)^2 + \left(\frac{\mathrm{e}^{\mathrm{i}z} - \mathrm{e}^{-\mathrm{i}z}}{2\mathrm{i}}\right)^2$$
$$= \frac{\mathrm{e}^{2\mathrm{i}z} + 2 + \mathrm{e}^{-2\mathrm{i}z}}{4} - \frac{\mathrm{e}^{2\mathrm{i}z} - 2 + \mathrm{e}^{-2\mathrm{i}z}}{4} = \frac{4}{4} = 1.$$

□

Bemerkung. Die Eulerformel zusammen mit $\mathrm{e}^{z+w} = \mathrm{e}^z \cdot \mathrm{e}^w$ gibt eine weitere Möglichkeit, die Additionstheoreme für Kosinus und Sinus zu beweisen:

$$\begin{aligned} \cos(z \pm w) + \mathrm{i}\sin(z \pm w) &= \mathrm{e}^{\mathrm{i}(z\pm w)} = \mathrm{e}^{\mathrm{i}z} \cdot \mathrm{e}^{\pm \mathrm{i}w} \\ &= (\cos z \pm \mathrm{i}\sin z)(\cos w \pm \mathrm{i}\sin w) \\ &= (\cos z \cos w \mp \sin z \sin w) + \mathrm{i}(\sin z \cos w \pm \cos z \sin w). \end{aligned}$$

Die Behauptung folgt durch Addition bzw. Subtraktion der so erhaltenen Gleichungen.

Lemma 10.12. *Für* $y \in \mathbb{R}$ *gilt* $\left|\mathrm{e}^{\mathrm{i}y}\right| = 1$.

Beweis. 1. Möglichkeit:

$$\left|\mathrm{e}^{\mathrm{i}y}\right|^2 = \mathrm{e}^{\mathrm{i}y} \cdot \overline{\mathrm{e}^{\mathrm{i}y}} = \mathrm{e}^{\mathrm{i}y} \cdot \mathrm{e}^{-\mathrm{i}y} = 1.$$

2. Möglichkeit: Mit der Eulerformel folgt

$$\operatorname{Re} \mathrm{e}^{\mathrm{i}y} = \cos y, \quad \operatorname{Im} \mathrm{e}^{\mathrm{i}y} = \sin y,$$

also

$$\left|\mathrm{e}^{\mathrm{i}y}\right|^2 = (\operatorname{Re} \mathrm{e}^{\mathrm{i}y})^2 + (\operatorname{Im} \mathrm{e}^{\mathrm{i}y})^2 = \cos^2 y + \sin^2 y = 1.$$

□

Satz 10.13. *Jede komplexe Zahl* $z \neq 0$ *kann in der Form*

$$z = r\mathrm{e}^{\mathrm{i}\theta} \qquad \text{(exponentielle Form)}$$

bzw.

$$z = r(\cos\theta + \mathrm{i}\sin\theta) \qquad \text{(trigonometrische Form)}$$

mit $r > 0$, $\theta \in \mathbb{R}$ *dargestellt werden. Dabei ist* $r = |z|$ *eindeutig bestimmt, während* θ *eindeutig bis auf ganzzahlige Vielfache von* 2π *ist.*

Die Zahl θ heißt das *Argument* $\arg z$ von z.
Schreiben wir zwei komplexe Zahlen z, w als $z = r\mathrm{e}^{\mathrm{i}\theta} = r(\cos\theta + \mathrm{i}\sin\theta)$ und $w = r'\mathrm{e}^{\mathrm{i}\vartheta} = r'(\cos\vartheta + \mathrm{i}\sin\vartheta)$, so

$$zw = rr'\mathrm{e}^{\mathrm{i}(\theta+\vartheta)} = rr'(\cos(\theta+\vartheta) + \mathrm{i}\sin(\theta+\vartheta)).$$

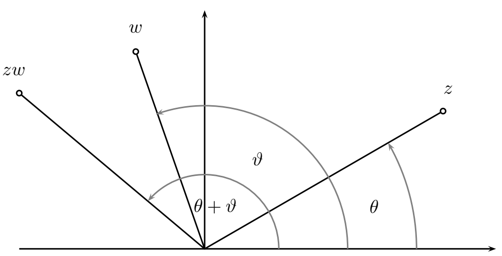

Abbildung 10.3: Geometrische Interpretation der Multiplikation: die Argumente addieren sich

Insbesondere gilt:

Lemma 10.14 (Formel von Moivre)**.** *Ist* $z = re^{i\theta} = r(\cos\theta + i\sin\theta)$ *und* $m \in \mathbb{N}$, *so ist*

$$z^m = r^m e^{im\theta} = r^m\left(\cos(m\theta) + i\sin(m\theta)\right).$$

Folgerung. Sei $w \in \mathbb{C}$, $w \neq 0$, $m \in \mathbb{N}$, wir schreiben $w = \varrho e^{i\vartheta} = \varrho(\cos\vartheta + i\sin\vartheta)$ mit $\varrho > 0$, $\vartheta \in \mathbb{R}$. Dann sind die m Lösungen $z_1, \dots, z_m$ der Gleichung

$$z^m = w$$

gegeben durch

$$z_j = \varrho^{1/m} e^{i\frac{\vartheta + 2\pi j}{m}} = \varrho^{1/m}\left(\cos\left(\frac{\vartheta + 2\pi j}{m}\right) + i\sin\left(\frac{\vartheta + 2\pi j}{m}\right)\right)$$

für $j = 0, 1, \dots, m-1$.

10.2 Vektoralgebra im $\mathbb{R}^3$

Wir diskutieren in diesem Abschnitt verschiedene Vektorprodukte (insbesondere im $\mathbb{R}^3$), wie sie in Anwendungen häufig auftreten. Wir schreiben Vektoren im $\mathbb{R}^n$ jetzt als $\vec{x} = (x_1, \dots, x_n)$, $\vec{y} = (y_1, \dots, y_n)$, usw. Dementsprechend bezeichnen wir die Einheitsvektoren im $\mathbb{R}^3$ mit $\vec{i} = (1,0,0)$, $\vec{j} = (0,1,0)$, $\vec{k} = (0,0,1)$.

10.2.1 Das Skalarprodukt

Wir hatten mit $|\vec{x}| = (x_1^2, \dots, x_n^2)^{1/2}$ den Betrag eines Vektors $\vec{x} = (x_1, \dots, x_n)$ bezeichnet.

Definition 10.15. Das *Skalarprodukt* $\vec{x} \cdot \vec{y}$ zweier Vektoren $\vec{x}, \vec{y}$ ist die reelle Zahl $|\vec{x}|\,|\vec{y}|\cos\varphi$, wobei φ den Winkel zwischen $\vec{x}$ und $\vec{y}$ bezeichnet.

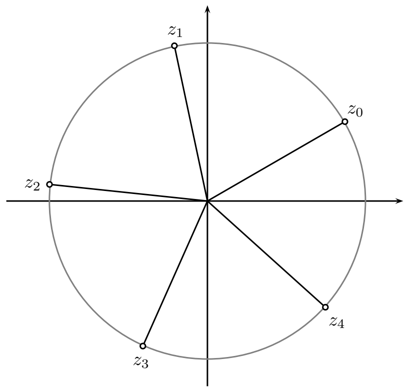

Abbildung 10.4: Die Gleichung $z^5 = w$: die Lösungen $z_0, z_1, \ldots, z_4$ folgen im Abstand von 72° aufeinander

Satz 10.16. (i) *Das Skalarprodukt* $\cdot \colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ *ist bilinear und symmetrisch.*

(ii) $|\vec{x}|^2 = \vec{x} \cdot \vec{x}$ *für alle* $\vec{x} \in \mathbb{R}^n$.

(iii) *Zwei Vektoren* $\vec{x}$, $\vec{y}$ *sind genau dann orthogonal, wenn* $\vec{x} \cdot \vec{y} = 0$.

(iv) *In Koordinaten gilt* $\vec{x} \cdot \vec{y} = x_1 y_1 + x_2 y_2 + \cdots + x_n y_n$.

Beweis. (i) Man benutzt, dass im Fall $\vec{x} \neq 0$ das Skalarprodukt nur von der orthogonalen Projektion von $\vec{y}$ auf die durch $\vec{x}$ gegebene Richtung abhängt: Diese Beobachtung reduziert den Nachweis der Bilinearität des Skalarprodukts auf den Fall, dass $\vec{x}, \vec{y}$ *kollinear* (d. h. parallel) sind (da orthogonale Prokjektionen lineare Abbildungen sind).

(ii), (iii) Sind offensichtlich.

(iv) Das Skalarprodukt ist nach Definition invariant unter Drehungen, d. h. es gilt $A\vec{x} \cdot A\vec{y} = \vec{x} \cdot \vec{y}$ für alle $\vec{x}, \vec{y} \in \mathbb{R}^n$ und alle $A \in \mathrm{O}(n)$.

Da der Ausdruck $x_1 y_1 + x_2 y_2 + \cdots + x_n y_n$ ebenfalls invariant unter Drehungen ist (die Invarianz dieses Ausdrucks ist eine mögliche Definition der orthogonalen Gruppe $\mathrm{O}(n)$), dürfen wir $\vec{x} = (x_1, 0, \ldots, 0)$ und $x_1 > 0$ annehmen. Nach Projektion auf die Richtung $(1, 0, \ldots, 0)$ hat $\vec{y}$ die Form $(y_1, 0, \ldots, 0)$. Doch dann ist die Behauptung offensichtlich,

$$\vec{x} \cdot \vec{y} = |\vec{x}|\,|\vec{y}| \cos\varphi = x_1 y_1,$$

wobei $\varphi = 0$ oder $\varphi = \pi$, je nachdem, ob $\vec{x}, \vec{y}$ in dieselbe oder entgegengesetzte Richtung zeigen.

□

10.2.2 Das Vektor- oder Kreuzprodukt

Definition 10.17. Für $\vec{x}$, $\vec{y} \in \mathbb{R}^3$ definieren wir $\vec{x} \times \vec{y} \in \mathbb{R}^3$ als den Vektor, der senkrecht sowohl auf $\vec{x}$ als auch auf $\vec{y}$ steht (ist einer der Vektoren $\vec{x}$, $\vec{y}$ gleich Null, so auch $\vec{x} \times \vec{y}$), mit diesen beiden Vektoren ein rechtshändiges Koordinatensystem bildet und der die Länge $|\vec{x}|\,|\vec{y}| \sin\varphi$ hat. Hierbei ist φ wiederum der Winkel zwischen $\vec{x}$ und $\vec{y}$.

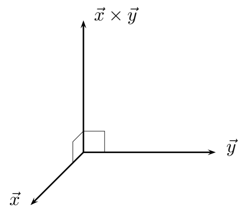

Abbildung 10.5: Zur Definition des Kreuzproduktes

Bemerkung. $|\vec{x} \times \vec{y}|$ ist gleich der Fläche des von $\vec{x}, \vec{y}$ aufgespannten Parallelogramms.

Satz 10.18. (i) *Das Kreuzprodukt* $\times \colon \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}^3$ *ist bilinear und schiefsymmetrisch, d. h.* $\vec{x} \times \vec{y} = -\,\vec{y} \times \vec{x}$.

(ii) *Zwei Vektoren* $\vec{x}$, $\vec{y}$ *sind genau dann kollinear, wenn* $\vec{x} \times \vec{y} = 0$ *gilt.*

(iii) *In Koordinaten gilt*

$$\vec{x} \times \vec{y} = (x_2y_3 - x_3y_2, x_3y_1 - x_1y_3, x_1y_2 - x_2y_1)\,.$$

Beweis. (i) Wir zeigen die Linearität im zweiten Argument. Dazu dürfen wir $\vec{x} \neq 0$ annehmen. Wir beobachten, dass sich das Kreuzprodukt $\vec{x} \times \vec{y}$ nicht ändert, wenn wir $\vec{y}$ orthogonal auf die Gerade senkrecht zu $\vec{x}$ projezieren, die in der von $\vec{x}$, $\vec{y}$ aufgespannten Ebene liegt:

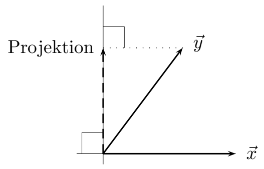

Abbildung 10.6: Zum Beweis von Satz 10.18 (i)

Damit ist der Nachweis der Linearität im zweiten Argument auf einen eindimensionalen Fall zurückgeführt, wo die Linearität jedoch offensichtlich ist.

Die Schiefsymmetrie des Kreuzprodukts ist ebenfalls offensichtlich.

(ii) Folgt direkt aus der Definition.

(iii) Das Kreuzprodukt ist invariant unter orientierungserhaltenden Drehungen, d. h. $A\vec{x} \times A\vec{y} = \vec{x} \times \vec{y}$ für alle $\vec{x}, \vec{y} \in \mathbb{R}^3$ und $A \in \mathrm{SO}(3)$, während orientierungsändernde Drehungen das Vorzeichen ändern, d. h. $B\vec{x} \times B\vec{y} = -\vec{x} \times \vec{y}$ für alle $\vec{x}, \vec{y} \in \mathbb{R}^3$ und $B \in \mathrm{O}(3) \setminus \mathrm{SO}(3)$. Gleiches trifft auf die rechte Seite in (iii) zu (Beweis?). Damit dürfen wir o. B. d. A. $\vec{x} = (x_1, 0, 0)$, $\vec{y} = (y_1, y_2, 0)$ mit $x_1 > 0$, $y_2 > 0$ annehmen. Orthogonale Projektion auf die Richtung $(0, 1, 0)$ vereinfacht schließlich weiter zum Fall $\vec{y} = (0, y_2, 0)$ mit $y_2 > 0$. Dann jedoch ist

$$\vec{x} \times \vec{y} = (0, 0, x_1 y_2),$$

und die Behauptung folgt.

□

Bemerkung. Die Beziehung in (iii) wird häufig als

$$\vec{x} \times \vec{y} = \begin{vmatrix} \vec{i} & \vec{j} & \vec{k} \\ x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \end{vmatrix}$$

geschrieben, wobei $\left| \ldots \right|$ die Determinante einer (in diesem Fall 3×3) Matrix bezeichnet.

10.2.3 Weitere Produkte

Wir betrachten hier mehrfache Produkte, die sich aus Skalar- und Vektorprodukt zusammensetzen:

Definition 10.19. Das *doppelte Vektorprodukt* ist das Produkt $\vec{x} \times (\vec{y} \times \vec{z})$.

Satz 10.20. *$\vec{x} \times (\vec{y} \times \vec{z})$ ist komplanar zu den Vektoren $\vec{y}$, $\vec{z}$. Genau gilt*

$$\vec{x} \times (\vec{y} \times \vec{z}) = (\vec{x} \cdot \vec{z})\, \vec{y} - (\vec{x} \cdot \vec{y})\, \vec{z}.$$

Beweis. Selbst mittels eines geometrischen Arguments oder direktem Nachrechnen unter Benutzung der Formel aus (iii). □

Definition 10.21. Das *gemischte* oder *Spatprodukt* von drei Vektoren $\vec{x}$, $\vec{y}$, $\vec{z}$ ist $(\vec{x} \times \vec{y}) \cdot \vec{z}$. Dafür schreibt man auch $\vec{x}\vec{y}\vec{z}$.

Satz 10.22. (i) *Die Trilinearform $\mathbb{R}^3 \times \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}$, $(\vec{x}, \vec{y}, \vec{z}) \mapsto \vec{x}\vec{y}\vec{z}$ ist total schiefsymmetrisch, d. h. eine Permutation der drei Faktoren ändert das Produkt um das Vorzeichen dieser Permutation.*

(ii) *$\vec{x}\vec{y}\vec{z} = 0$ genau dann, wenn alle drei Vektoren in einer Ebene liegen.*

(iii)
$$\vec{x}\vec{y}\vec{z} = \begin{vmatrix} x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \\ z_1 & z_2 & z_3 \end{vmatrix}.$$

Insbesondere ist $|\vec{x}\vec{y}\vec{z}|$ *gleich dem Inhalt des durch* $\vec{x}, \vec{y}, \vec{z}$ *aufgespannten Parallelepipeds* (*auch Spat genannt*).

Beweis. Wiederum Übungsaufgabe. □

Schließlich möchte ich noch auf folgende Formeln hinweisen:

Satz 10.23. *Für* $\vec{x}$, $\vec{y}$, $\vec{z}$, $\vec{u}$, $\vec{v}$, $\vec{w} \in \mathbb{R}^3$ *gilt*

(i) (*Lagrangesche Identität*)
$$(\vec{x} \times \vec{y})(\vec{z} \times \vec{u}) = (\vec{x}\vec{z})(\vec{y}\vec{u}) - (\vec{y}\vec{z})(\vec{x}\vec{u}),$$

(ii)
$$(\vec{x}\vec{y}\vec{z})(\vec{u}\vec{v}\vec{w}) = \begin{vmatrix} \vec{x}\vec{u} & \vec{x}\vec{v} & \vec{x}\vec{w} \\ \vec{y}\vec{u} & \vec{y}\vec{v} & \vec{y}\vec{w} \\ \vec{z}\vec{u} & \vec{z}\vec{v} & \vec{z}\vec{w} \end{vmatrix}.$$

Beweis. Übungsaufgabe. □

Kapitel 11

Integrale über räumliche Bereiche

11.1 Kurvenintegrale

Wir erinnern daran, dass eine Kurve in einem metrischen Raum X eine stetige Abbildung $\gamma\colon I \to X$ von einem (nichttrivialen) Intervall I nach X ist.

Definition 11.1. Eine Kurve $\gamma\colon I \to \mathbb{R}^n$ heißt *stückweise* C^1, falls $I = \bigcup_{k=1}^{m} I_k$ mit in I abgeschlossenen Intervallen I_k, so dass $I_{k-1} \cap I_k$ für $k = 1, \ldots, m$ aus genau einem Punkt besteht und $\gamma\big|_{I_k}\colon I_k \to \mathbb{R}^n$ für alle $k = 1, \ldots, m$ von der Klasse C^1 ist. Zudem fordern wir $\left(\gamma\big|_{I_k}\right)'(t) \neq 0$ für alle $t \in I_k$.

Ist also I ein Intervall mit Anfangspunkt a_0 und Endpunkt a_m ($a_0 < a_m$, $a_0 = -\infty$ bzw. $a_m = \infty$ ist möglich), so gibt es eine Partition $a_0 < a_1 < \cdots < a_m$, so dass

$$\gamma\big|_{[a_{k-1},a_k]\cap I}\colon [a_{k-1}, a_k] \cap I \to \mathbb{R}^n$$

von der Klasse C^1 ist und dort die Abbildung nirgends verschwindet. Im Allgemeinen werden wir jedoch $\gamma'(a_k - 0) \neq \gamma'(a_k + 0)$ haben.

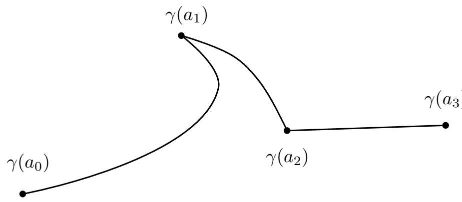

Abbildung 11.1: Stückweise C^1-Kurve

Bemerkung. Für ein (nichttriviales) Intervall $I \subseteq \mathbb{R}$ und eine differenzierbare Abbildung $f\colon I \to \mathbb{R}^n$ können wir $Df(t) \in \mathcal{L}(\mathbb{R}, \mathbb{R}^n)$ mit dem Vektor $f'(t) \in \mathbb{R}^n$ identifizieren, den wir durch komponentenweises Differenzieren erhalten:

$$Df(t)a = af'(t), \quad a \in \mathbb{R}.$$

$\gamma'(t)$ ist der *Tangentialvektor* an die Kurve γ im Punkt $\gamma(t)$. Die Länge $|\gamma'(t)|$ eines Tangentialvektors hängt davon ab, wie „schnell“ die Kurve γ durchlaufen wird.

Definition 11.2. Zwei C^1-Kurven $\gamma\colon [a,b] \to \mathbb{R}^n$ und $\tilde{\gamma}\colon [c,d] \to \mathbb{R}^n$ heißen *äquivalent* (oder durch *Umparametrisierung* auseinander hervorgehend), wenn es eine bijektive C^1-Abbildung

$$\phi\colon [a,b] \to [c,d]$$

gibt mit $\phi'(t) \neq 0$ für alle $t \in [a,b]$, so dass

$$\gamma = \tilde{\gamma} \circ \phi.$$

Gilt dabei $\phi(a) = c$ (äquivalent: $\phi(b) = d$ bzw. $\phi'(t) > 0$ für alle t), so heißt die Umparametrisierung *orientierungserhaltend*, andernfalls *orientierungsändernd*. Diese Begriffe finden auch Anwendung im Fall stückweiser C^1-Kurven $\gamma\colon I \to \mathbb{R}^n$ und $\tilde{\gamma}\colon \tilde{I} \to \mathbb{R}^n$, wenn man Umparametrisierungen $\phi\colon I \to \tilde{I}$ zulässt, die stückweise C^1 sind.

Bemerkung. Offenbar sind Umparametrisierung sowie orientierungserhaltende Umparametrisierung Äquivalenzrelationen auf der Menge der stückweise C^1-Kurven im $\mathbb{R}^n$.

Definition 11.3. Als *Länge* $\ell(\gamma) \in [0,\infty]$ einer Kurve $\gamma\colon [a,b] \to \mathbb{R}^n$ bezeichnet man

$$\ell(\gamma) = \sup\Big\{\sum_{j=1}^{N} |\gamma(t_j) - \gamma(t_{j-1})| \;\Big|\; a = t_0 < t_1 < \cdots < t_{N-1} < t_N = b\Big\}.$$

Ist $\gamma\colon I \to \mathbb{R}^n$ eine Kurve, die auf einem Intervall $I \subseteq \mathbb{R}$ definiert ist, so setzt man

$$\ell(\gamma) = \sup_{[a,b]\subseteq I} \ell\Big(\gamma\big|_{[a,b]}\Big).$$

γ heißt *rektifizierbar*, falls

$$\ell\Big(\gamma\big|_{[a,b]}\Big) < \infty$$

für alle kompakten Intervalle $[a,b] \subseteq I$.

Satz 11.4. *Stückweise C^1-Kurven $\gamma\colon I \to \mathbb{R}^n$ sind rektifizierbar. Zudem gilt*

$$\ell\Big(\gamma\big|_{[a,b]}\Big) = \int_a^b |\gamma'(t)|\, dt$$

für alle $[a,b] \subseteq I$. Die Länge der Kurve $\gamma\big|_{[a,b]}$ ändert sich unter Umparametrisierung nicht.

Beweis. Wir dürfen $\gamma\colon [a,b] \to \mathbb{R}^n$ als C^1-Kurve annehmen.

1. Sei zuerst $a = t_0 < t_1 < \cdots < t_N = b$ eine beliebige Partition des Intervalls $[a,b]$. Dann gilt

$$|\gamma(t_j) - \gamma(t_{j-1})| = \left| \int_{t_{j-1}}^{t_j} \gamma'(s)\, ds \right| \leq \int_{t_{j-1}}^{t_j} |\gamma'(s)|\ ds,$$

also

$$\sum_{j=1}^{N} |\gamma(t_j) - \gamma(t_{j-1})| \leq \sum_{j=1}^{n} \int_{t_{j-1}}^{t_j} |\gamma'(s)|\ ds = \int_a^b |\gamma'(s)|\ ds.$$

Folglich ist γ rektifizierbar und

$$\ell(\gamma) \leq \int_a^b |\gamma'(t)|\ dt.$$

2. Wir wählen $\epsilon > 0$. Dann gibt es ein $\delta > 0$, so dass für jede Partition $a = t_0 < t_1 < \cdots < t_N = b$ mit $t_j - t_{j-1} \leq \delta$ für $j = 1, \ldots, N$

$$\left| \int_a^b |\gamma(t)|\ dt - \sum_{j=1}^{N} |\gamma(s_j)|\,(t_j - t_{j-1}) \right| \leq \epsilon$$

für beliebige Zwischenpunkte $s_j \in [t_{j-1}, t_j]$ gilt. Weiterhin folgt aus dem Mittelwertsatz und der gleichmäßigen Stetigkeit von γ' auf $[a,b]$, dass wir $\delta > 0$ so klein wählen können, dass

$$\left| \frac{|\gamma(t) - \gamma(t')|}{t - t'} - |\gamma'(s)| \right| \leq \frac{\epsilon}{b-a}$$

für $0 < t - t' \leq \delta$ und $s \in [t', t]$ gilt. Dann gilt

$$\begin{aligned} 0 &\leq \int_a^b |\gamma'(t)|\ dt - \sum_{j=1}^{N} |\gamma(t_j) - \gamma(t_{j-1})| \\ &\leq \left| \int_a^b |\gamma'(t)|\ dt - \sum_{j=1}^{N} |\gamma'(s_j)|\,(t_j - t_{j-1}) \right| \\ &\qquad + \sum_{j=1}^{N} \big| |\gamma(t_j) - \gamma(t_{j-1})| - |\gamma'(s_j)|\,(t_j - t_{j-1}) \big| \\ &\leq \epsilon + \frac{\epsilon}{b-a} \sum_{j=1}^{N} (t_j - t_{j-1}) = 2\epsilon \end{aligned}$$

für beliebige markierte Partitionen der Feinheit höchstens δ. Daraus folgt, dass tatsächlich

$$\ell(\gamma) = \int_a^b |\gamma'(t)| \, dt$$

gilt.

3. Da $\ell(\gamma)$ nur vom Bild von γ abhängt, ist die Unabhängigkeit von $\ell(\gamma)$ von der gewählten Parametrisierung offensichtlich.

   Ein davon unabhängiges Argument wäre, dass für eine beliebige C^1-Umparametrisierung $\phi\colon [a,b] \to [c,d]$ mit $\gamma = \tilde{\gamma} \circ \phi$

$$\begin{aligned} \ell(\tilde{\gamma}) &= \int_c^d |\tilde{\gamma}'(s)| \, ds = \int_a^b |\tilde{\gamma}'(\phi(t))| \, |\phi'(t)| \, dt \\ &= \int_a^b |\gamma'(t)| \, dt = \ell(\gamma) \end{aligned}$$

   gilt.

□

Definition 11.5. Die stückweise C^1-Kurve $\gamma\colon I \to \mathbb{R}^n$ heißt bezüglich der *Bogenlänge* parametrisiert, falls $|\gamma'(\sigma)| = 1$ für alle $\sigma \in I$ gilt (mit entsprechender Modifikation an Nicht-C^1-Stellen).

Satz 11.6. *Sei $\gamma\colon I \to \mathbb{R}^n$ eine stückweise C^1-Kurve, $t_0 \in I$. Wir definieren*

$$\sigma(t) = \int_{t_0}^t |\gamma'(s)| \, ds, \quad t \in I.$$

Dann ist σ eine orientierungserhaltende, stückweise C^1-Umparametrisierung von γ, die γ bezüglich der Bogenlänge parametrisiert.

Beweis. O. B. d. A. sei $\gamma\colon [a,b] \to \mathbb{R}^n$ eine C^1-Kurve und $t_0 = a$. Dann ist $\sigma\colon [a,b] \to [0, l(\gamma)]$ wegen $\sigma'(t) = |\gamma'(t)| > 0$ eine orientierungserhaltende Umparametrisierung. Insbesondere existiert $t = \sigma^{-1}\colon [0, l(\gamma)] \to [a,b]$ sowie

$$t'(\sigma) = \frac{1}{\sigma'(t(\sigma))} = \frac{1}{|\gamma'(t(\sigma))|}.$$

Damit gilt mit $\tilde{\gamma} = \gamma \circ t$, dass

$$|\tilde{\gamma}'(\sigma)| = |\gamma'(t(\sigma))| \, |t'(\sigma)| = 1.$$

□

Definition 11.7. Sei $\gamma\colon [a,b] \to \mathbb{R}^n$ stückweise C^1, $f\colon \gamma([a,b]) \to \mathbb{R}$ stetig. Dann definieren wir

$$\int_\gamma f(x) \, |d\sigma| = \int_a^b f(\gamma(t)) \, |\gamma'(t)| \, dt$$

als das *Kurvenintegral erster Art* von f entlang γ.

Satz 11.8. *Dieses Kurvenintegral hat die folgenden Eigenschaften:*

(i) *Es ist linear bezüglich des Integranden f.*

(ii) *Es ist additiv unter der Verkettung von Wegen.*

(iii) *Es ist unabhängig von der Wahl der Parametrisierung.*

(iv) *Es gilt*

$$\left| \int_\gamma f(x)\ |d\sigma| \right| \leq \ell(\gamma) \sup_\gamma |f| .$$

Beweis. (iii) Sei $\phi\colon [a,b] \to [c,d]$ wie oben, $\gamma = \tilde{\gamma} \circ \phi$. Dann gilt

$$\begin{aligned} \int_c^d f\left(\tilde{\gamma}(s)\right) |\tilde{\gamma}'(s)|\ ds &= \int_a^b f\left(\tilde{\gamma}\left(\phi(t)\right)\right) \left|\tilde{\gamma}'\left(\phi(t)\right)\right| |\phi'(t)|\ dt \\ &= \int_a^b f\left(\gamma(t)\right) |\gamma'(t)|\ dt. \end{aligned}$$

(iv) Es gilt

$$\begin{aligned} \left| \int_\gamma f\ |d\sigma| \right| &= \left| \int_a^b f\left(\gamma(t)\right) |\gamma'(t)|\ dt \right| \\ &\leq \sup_\gamma |f| \int_a^b |\gamma'(t)|\ dt = \ell(\gamma) \sup_\gamma |f| . \end{aligned}$$

□

11.2 Mehrdimensionale Integrale

Wir möchten jetzt Integrale der Form $\int_A f(x)\,dx$ für gewisse Mengen $A \subseteq \mathbb{R}^n$ definieren, wobei $f\colon A \to \mathbb{R}$ stetig ist. Dies tun wir, indem wir formal

$$\int_A f(x)\,dx = \int_{\mathbb{R}^n} \chi_A(x) f(x)\,dx$$

setzen, wobei χ_A die durch

$$\chi_A(x) = \begin{cases} 1, & x \in A, \\ 0, & x \notin A, \end{cases}$$

definierte *charakteristische Funktion* von A bezeichnet.

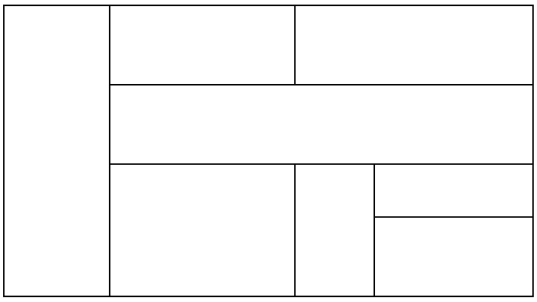

Abbildung 11.2: Eine Partition von B

11.2.1 Riemann-Integrierbarkeit über n-dimensionalen Rechtecken

Definition 11.9. Ein *achsenparalleles n-dimensionales Rechteck* ist eine Menge der Form

$$B = \prod_{j=1}^{n} [a_j, b_j] = \{x \in \mathbb{R}^n \mid a_j \leq x_j \leq b_j \text{ für } j = 1, \ldots, n\},$$

wobei $a_j, b_j \in \mathbb{R}$, $a_j < b_j$. Das *n-dimensionale Volumen* von B ist

$$\mathrm{vol}_n(B) = \prod_{j=1}^{n} (b_j - a_j).$$

Weiterhin setzen wir $\Delta(B) = \max\limits_{j=1,\ldots,n} (b_j - a_j)$.

Das Innere von B ist die Menge

$$B^\circ = \prod_{j=1}^{n} (a_j, b_j) = \{x \in \mathbb{R}^n \mid a_j < x_j < b_j \text{ für } j = 1, \ldots, n\},$$

der Rand von B ist $\partial B = B \setminus B^\circ$.

Definition 11.10. (a) Eine *Partition* P von B ist eine endliche Menge $\{B_k\}_{k=1}^{K}$ von achsenparallelen Rechtecken mit $B = \bigcup\limits_{k=1}^{K} B_k$ und $B_k^\circ \cap B_{k'}^\circ = \emptyset$ für $k \neq k'$. Als *Feinheit* von P definieren wir

$$\Delta(P) = \max_{k=1,\ldots,K} \Delta(B_k).$$

(b) Eine *markierte Partition* von B ist eine Partition P, bei der zusätzlich für jedes $k = 1, \ldots, K$ eine Stelle $c_k \in B_k$ ausgezeichnet wurde.

Lemma 11.11. *Ist $\{B_k\}_{k=1}^K$ eine Partition von B, so gilt*

$$\mathrm{vol}_n(B) = \sum_{k=1}^K \mathrm{vol}_n(B_k).$$

Beweis. (i) Sei $B = \prod_{j=1}^n [a_j, b_j]$ und $d_l \in (a_l, b_l)$ für ein $l \in \{1, \ldots, n\}$. Seien

$$B' = \{x \in B \mid a_l \leq x_l \leq d_l\},$$
$$B'' = \{x \in B \mid d_l \leq x_l \leq b_l\}.$$

Wegen $b_l - a_l = (d_l - a_l) + (b_l - d_l)$ folgt sofort, dass

$$\mathrm{vol}_n(B) = \mathrm{vol}_n(B') + \mathrm{vol}_n(B'')$$

für diese spezielle Partition von B.

(ii) Wir nehmen als nächstes an, dass jedes der Intervalle $[a_j, b_j]$ durch Stellen

$$a_j = d_0^{(j)} < d_1^{(j)} < \cdots < d_{m_j}^{(j)} = b_j$$

unterteilt ist. Dann haben wir für jeden Multiindex $\alpha = (\alpha_1, \ldots, \alpha_n) \in \mathbb{N}^n$ mit $\alpha_j \leq m_j$ das Rechteck

$$B_\alpha = \prod_{j=1}^n \left[d_{\alpha_j - 1}^{(j)}, d_{\alpha_j}^{(j)}\right].$$

Diese B_α bilden eine Partition von B und sukzessive Anwendung von (i) ergibt

$$\mathrm{vol}_n(B) = \sum_\alpha \mathrm{vol}_n(B_\alpha).$$

(iii) Schließlich betrachten wir den allgemeinen Fall. Für jedes j mit $1 \leq j \leq n$ ordnen wir die Endpunkte der Intervalle $[a_j, b_j]$ aller Rechtecke B_k in der Partition $\{B_k\}_{k=1}^K$ von B in aufsteigender Ordnung. Wir erhalten dann eine Partition $\{B_\alpha\}$ von B wie in (ii) (siehe Abbildung 11.3). Außerdem bilden die B_α mit $B_\alpha \subseteq B_k$ eine Partition von B_k. Damit gilt dann

$$\mathrm{vol}_n(B) = \sum_\alpha \mathrm{vol}_n(B_\alpha) = \sum_k \sum_{B_\alpha \subseteq B_k} \mathrm{vol}_n(B_\alpha) = \sum_k \mathrm{vol}_n(B_k).$$

□

Lemma 11.12. *Für je zwei Partitionen P', P'' von B gibt es eine gemeinsame* Verfeinerung *P, d. h. eine Partition P von B, so dass es für alle B_k in dieser Partition ein $B'_{k'}$ in P' mit $B_k \subseteq B'_{k'}$ und ein $B''_{k''}$ in P'' mit $B_k \subseteq B''_{k''}$ gibt.*

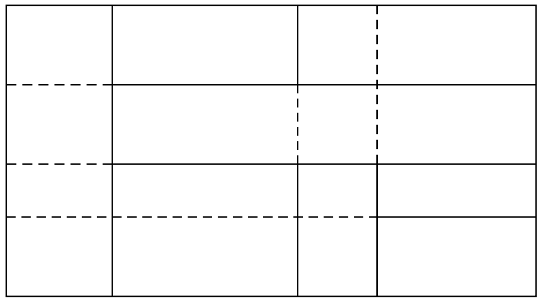

Abbildung 11.3: Verfeinerung der Partition von B aus Abbildung 11.2

Beweis. Wie im vorigen Beweis, indem man auf die Endpunkte der Intervalle $[a_j, b_j]$ für alle Rechtecke $B'_{k'}$ in P' und $B''_{k''}$ in P'' schaut. □

Definition 11.13. Sei $f\colon B \to \mathbb{R}$, P eine markierte Partition. Dann heißt

$$S(f, P) = \sum_{k=1}^{K} f(c_k) \operatorname{vol}_n(B_k)$$

die zu f, P gehörige *Riemann–Summe*.

Definition 11.14. Eine Funktion $f\colon B \to \mathbb{R}$ heißt *Riemann-integrierbar über* B, falls es eine Zahl $L \in \mathbb{R}$ mit folgender Eigenschaft gibt: Für alle $\epsilon > 0$ existiert ein $\delta > 0$, so dass

$$|L - S(f, P)| \leq \epsilon$$

für alle markierten Partitionen P von B mit $\Delta(P) \leq \delta$. Man schreibt

$$L = \int_B f(x)\, dx.$$

Wie in Diff I gilt das Cauchy-Kriterium:

Satz 11.15. *Eine Funktion $f\colon B \to \mathbb{R}$ ist genau dann R-integrierbar, falls es für alle $\epsilon > 0$ ein $\delta > 0$ gibt, so dass für alle markierten Partitionen P_0, P_1 von B mit $\Delta(P_j) \leq \delta$ für $j = 0, 1$ die Ungleichung*

$$|S(f, P_0) - S(f, P_1)| \leq \epsilon$$

gilt.

Wie in Diff I zeigt man:

Satz 11.16. *Das Riemann-Integral hat folgende Eigenschaften:*

(i) *Das Integral ist linear und monoton bezüglich des Integranden.*

(ii) *Ist $f\colon B \to \mathbb{R}$ R-integrierbar, so ist f beschränkt.*

(iii) *Sind $f, g\colon B \to \mathbb{R}$ R-integrierbar und ist $h\colon f(B) \to \mathbb{R}$ stetig, so sind $|f|$, $f \cdot g$, $h \circ f\colon B \to \mathbb{R}$ R-integrierbar. Darüber hinaus gilt*

$$\left| \int_B f(x)\, dx \right| \leq \int_B |f(x)|\, dx \leq \sup_{x \in B} |f(x)| \cdot \mathrm{vol}_n(B).$$

(iv) *Ist $f\colon B \to \mathbb{R}$ R-integrierbar und ist $B' \subseteq B$ ein achsenparalleles Rechteck, so ist $f\big|_{B'}$ R-integrierbar über B'. Umgekehrt gilt: Ist $\{B_k\}_{k=1}^K$ eine Partition von B und ist $f\big|_{B_k}$ R-integrierbar über B_k für alle $k = 1, \ldots, K$, so ist f R-integrierbar und es gilt*

$$\int_B f(x)\, dx = \sum_{k=1}^{K} \int_{B_k} f(x)\, dx.$$

(v) *Stetige Funktionen $f\colon B \to \mathbb{R}$ sind R-integrierbar.*

11.2.2 Jordan-messbare Mengen

Definition 11.17. (i) Eine beschränkte Menge $A \subseteq \mathbb{R}^n$ heißt *Jordan-messbar*, falls χ_A R-integrierbar ist (über einem beliebigen und somit über jedem Rechteck $B \supseteq A$). Wir setzen dann

$$\mathrm{vol}_n(A) = \int_{\mathbb{R}^n} \chi_A(x)\, dx \left(= \int_B \chi_A(x)\, dx \right).$$

(ii) Eine Jordan-messbare Menge A heißt *vernachlässigbar*, falls $\mathrm{vol}_n(A) = 0$ gilt.

Bemerkung. Tatsächlich gilt $\mathrm{vol}_n(B) = \prod_{j=1}^{n} (b_j - a_j)$ für jedes achsenparallele Rechteck $B = \prod_{j=1}^{n} [a_j, b_j]$ (Übungsaufgabe).

Satz 11.18. *Für eine beschränkte Menge $A \subseteq \mathbb{R}^n$ sind folgende Aussagen äquivalent:*

(i) *A ist Jordan-messbar.*

(ii) *∂A ist vernachlässigbar.*

Beweis. ($\Rightarrow$) Sei $\epsilon > 0$. Dann existiert eine Partition $\{B_k\}$ von B mit

$$\sum_{k \in K_a} \mathrm{vol}_n(B_k) - \sum_{k \in K_i} \mathrm{vol}_n(B_k) \leq \epsilon,$$

wobei $K_a = \{k \mid B_k \cap A \neq \emptyset\}$, $K_i = \{k \mid B_k \subseteq A\}$, da

$$\sum_{k \in K_a} \mathrm{vol}_n(B_k) = \sum_{k=1}^{K} \max_{B_k} \chi_A \ \mathrm{vol}_n(B_k),$$

$$\sum_{k \in K_i} \mathrm{vol}_n(B_k) = \sum_{k=1}^{K} \min_{B_k} \chi_A \ \mathrm{vol}_n(B_k)$$

Riemann-Summen für geeignete Markierungen von $\{B_k\}$ sind. Insbesondere gilt

$$\bar{A} \subseteq \bigcup_{k \in K_a} B_k.$$

Wir ersetzen nun die Rechtecke B_k für $k \in K_i$ durch kleinere Rechtecke B'_k, so dass

$$A \supseteq \bigcup_{k \in K_i} B_k \supseteq \bigcup_{k \in K_i} B_k^\circ \supseteq \bigcup_{k \in K_i} B'_k$$

und

$$\sum_{k \in K_a} \mathrm{vol}_n(B_k) - \sum_{k \in K_i} \mathrm{vol}_n(B'_k) \leq 2\epsilon.$$

Es gilt dann

$$A^\circ \supseteq \bigcup_{k \in K_i} B'_k$$

und

$$\partial A = \bar{A} \setminus A^\circ \subseteq \bigcup_{k \in K_a} B_k - \bigcup_{k \in K_i} B'_k.$$

Somit ist $\mathrm{vol}_n(\partial A)$ für jedes $\epsilon > 0$ kleiner als 2ϵ.

($\Leftarrow$) Sei wiederum $\epsilon > 0$. Wir finden dann eine Partition $\{B_k\}$ von B, so dass

$$\partial A \subseteq \bigcup_k B_k^\circ$$

und

$$\sum_k \left(\max_{B_k} \chi_A - \min_{B_k} \chi_A \right) \mathrm{vol}_n(B_k) \leq \epsilon.$$

Für jedes $x \in A \setminus \partial A$ finden wir ein Rechteck $B^{(x)}$ mit $x \in (B^{(x)})^\circ \subseteq B^{(x)} \subseteq A$. Dann gilt insbesondere

$$\max_{B^{(x)}} \chi_A - \min_{B^{(x)}} \chi_A = 0.$$

Da $\bar{A}$ kompakt ist, gibt es eine endliche Teilüberdeckung der Überdeckung

$$\{B_k^\circ\}_{k=1}^{K} \cup \{(B^{(x)})^\circ\}_{x \in A \setminus \partial A}$$

von $\bar{A}$. Wir benutzen nun wiederum die Endpunkte der Koordinatenintervalle der Rechtecke in dieser endlichen Teilüberdeckung, um eine Partition $\{B'_l\}$ von B zu definieren. Für diese Partition gilt nach Konstruktion

$$\sum_{l\in K'_a} \mathrm{vol}_n(B_l) - \sum_{l\in K'_i} \mathrm{vol}_n(B_l) \le \epsilon,$$

wobei K'_a, K'_i dieselbe Bedeutung wie in (i) haben, diesmal mit Bezug auf die Partition $\{B'_l\}$ von B.

□

Folgerung. Sind A, $B \subseteq \mathbb{R}^n$ Jordan-messbar, so auch

$$A\cap B, \quad A\cup B, \quad A\setminus B, \quad A^\circ, \quad \bar{A}.$$

Dabei gilt

$$\mathrm{vol}_n(A\cap B) + \mathrm{vol}_n(A\cup B) = \mathrm{vol}_n(A) + \mathrm{vol}_n(B)$$

und

$$\mathrm{vol}_n(A) = \mathrm{vol}_n(\bar{A}) = \mathrm{vol}_n(A^\circ).$$

Beweis. Es gilt

$$\chi_{A\cap B} = \chi_A\chi_B, \quad \chi_{A\cup B} = \chi_A + \chi_B - \chi_{A\cap B}, \quad \chi_{A\setminus B} = \chi_A - \chi_{A\cap B}$$

sowie

$$A^\circ = A\setminus\partial A, \quad \bar{A} = A\cup\partial A.$$

□

Definition 11.19. Für $A \subseteq \mathbb{R}^n$ bezeichnen wir mit $\mathcal{J}(A)$ den Raum der kompakten, Jordan-messbaren Mengen $K \subseteq A$.

Satz 11.20. *Sei $K \in \mathcal{J}(\mathbb{R}^n)$ und $f\colon K \to \mathbb{R}$ stetig. Dann:*

(i) *f ist R-integrierbar über K.*

(ii) *Ist $g\colon K\to\mathbb{R}$ eine stetige Funktion mit $g(x) \le f(x)$ für alle $x\in K$, so ist $L \in \mathcal{J}(\mathbb{R}^{n+1})$, wobei*

$$L = \big\{(x,y)\in K\times\mathbb{R} \;\big|\; g(x)\le y\le f(x)\big\},$$

mit

$$\mathrm{vol}_{n+1}(L) = \int_K (f(x)-g(x))\,dx.$$

Beweis. Es genügt (i) im Fall $f \ge 0$ zu zeigen. Sei $c = \min\limits_K g$. Dann ist $c \le g(x) \le f(x)$ für alle $x\in K$ und

$$L = \big\{(x,y)\in K\times\mathbb{R} \;\big|\; c\le y\le f(x)\big\} \setminus \big\{(x,y)\in K\times\mathbb{R} \;\big|\; c\le y< g(x)\big\}.$$

Daher genügt es, (ii) im Fall

$$L = \big\{(x,y) \in K \times \mathbb{R} \,\big|\, 0 \leq y \leq f(x)\big\}$$

zu zeigen.
Unter diesen vereinfachenden Annahmen behandeln wir jetzt (i), (ii) simultan. Sei $B \subset \mathbb{R}^n$ ein achsenparalleles Rechteck mit $K \subseteq B$ und sei $\{B_k\}$ eine Partition P von B. Wir setzen $f(x) = 0$ für $x \in B \setminus K$ und betrachten

$$\begin{aligned}
\overline{S}(f,P) &= \sum_k \sup_{B_k} f \ \mathrm{vol}_n(B_k) = \sum_{B_k \cap K \neq \emptyset} \max_{B_k} f \ \mathrm{vol}_n(B_k), \\
\underline{S}(f,P) &= \sum_k \inf_{B_k} f \ \mathrm{vol}_n(B_k) = \sum_{B_k \subseteq K} \min_{B_k} f \ \mathrm{vol}_n(B_k).
\end{aligned}$$

Dann ist $L \subseteq \widetilde{B} = B \times [0, \max_K f]$, wobei $\widetilde{B}$ ein achsenparalleles Rechteck in $\mathbb{R}^{n+1}$ ist. Außerdem gilt

$$L \subseteq \bigcup_{B_k \cap K \neq \emptyset} B_k \times \big[0, \max_{B_k} f\big],$$

und falls $B_k \subseteq K$, so

$$B_k \times \big[0, \min_{B_k} f\big] \subseteq L.$$

Sei $\widetilde{P}$ eine Partition von $\widetilde{B}$, die alle Rechtecke der Form $B_k \times \big[0, \max_{B_k} f\big]$ für $B_k \cap K \neq \emptyset$ und der Form $B_k \times \big[0, \min_{B_k} f\big]$ für $B_k \subseteq K$ verfeinert. Wir erhalten dann

$$\begin{aligned}
\overline{S}\big(\chi_L, \widetilde{P}\big) &\leq \sum_{B_k \cap k \neq \emptyset} \max_{B_k} f \ \mathrm{vol}_n(B_k) = \overline{S}(f,P), \\
\underline{S}\big(\chi_L, \widetilde{P}\big) &\geq \sum_{B_k \subseteq K} \min_{B_k} \ f \, \mathrm{vol}_n(B_k) = \underline{S}(f,B).
\end{aligned}$$

Folglich genügt es zu zeigen, dass die Differenz

$$\begin{aligned}
&\overline{S}(f,P) - \underline{S}(f,P) \\
&\qquad = \sum_{B_k \subseteq K} \big(\max_{B_k} f - \min_{B_k} f\big) \,\mathrm{vol}_n(B_k) + \sum_{\substack{B_k \cap K \neq \emptyset, \\ B_k \setminus K \neq \emptyset}} \max_{B_k} f \ \mathrm{vol}_n(B_k)
\end{aligned}$$

bei geeigneter Wahl von P beliebig klein wird: f ist gleichmäßig stetig auf der kompakten Menge K. Insbesondere existiert für jedes $\epsilon > 0$ ein $\delta > 0$ mit

$$\max_{B_k} f - \min_{B_k} f \leq \epsilon,$$

falls $B_k \subseteq K$ und $\Delta(B_k) \leq \delta$ ist. Dies wird erreicht, indem man die B_k mit $B_k \subseteq K$ gegebenenfalls verfeinert. Weiterhin impliziert die Jordan-Messbarkeit von K, dass ∂K vernachlässigbar ist und folglich

$$\sum_{\substack{B_k \cap K \neq \emptyset, \\ B_k \setminus K \neq \emptyset}} \operatorname{vol}_n(B_k) \leq \epsilon$$

bei geeigneter Wahl von P. Somit gilt für solche P, dass

$$\overline{S}(f,P) - \underline{S}(f,P) \leq \epsilon \operatorname{vol}_n(K) + \epsilon \max_K f,$$

was beliebig klein gewählt werden kann.
Es bleibt zu zeigen, dass die beiden Mengen

$$\begin{aligned} M &= \big\{(x,y) \in K \times \mathbb{R} \bigm| c \leq y \leq g(x)\big\}, \\ M' &= \big\{(x,y) \in K \times \mathbb{R} \bigm| c \leq y < g(x)\big\} \end{aligned}$$

dasselbe $(n+1)$-dimensionale Volumen haben. Dies jedoch folgt aus $\overline{M'} = M$ (bzw. aus $M \setminus M' \subseteq \partial M$ und $\operatorname{vol}_{n+1}(\partial M) = 0$). $\square$

Definition 11.21. Für eine Funktion $f\colon \mathbb{R}^n \to \mathbb{R}$ ist

$$\operatorname{supp} f = \overline{\big\{x \in \mathbb{R}^n \bigm| f(x) \neq 0\big\}}$$

der *Träger von* f. Der Träger von f ist die kleinste abgeschlossene Teilmenge von $\mathbb{R}^n$, außerhalb derer f verschwindet.

Folgerung. Stetige Funktionen $f\colon \mathbb{R}^n \to \mathbb{R}$ mit kompaktem Träger sind R-integrierbar.

Beweis. Es gibt ein achsenparalleles Rechteck B mit $\operatorname{supp} f \subseteq B$. $\square$

Folgerung. Sei $m < n$, $K \in \mathcal{J}(\mathbb{R}^m)$ und $f\colon K \to \mathbb{R}^{n-m}$ stetig. Dann ist der Graph von f,

$$\operatorname{graph} f = \big\{(x, f(x)) \bigm| x \in K\big\},$$

vernachlässigbar im $\mathbb{R}^n$.

Beweis. Wir schreiben $f = (f_1, \dots, f_{n-m-1}, f_{n-m})$. Da die Funktion $x \mapsto (f_1(x), \dots, f_{n-m-1}(x))$ auf K beschränkt ist, gibt es ein achsenparalleles Rechteck $B \subseteq \mathbb{R}^{n-m-1}$ mit $(f_1(x), \dots, f_{n-m-1}(x)) \in B$ für alle $x \in K$. Definiere $\widetilde{f}\colon K \times B \to \mathbb{R}$ durch

$$\widetilde{f}(x,t) = f_{n-m}(x).$$

Dann

$$\begin{aligned} \operatorname{graph} f &= \big\{(x, f(x)) \in \mathbb{R}^n \bigm| x \in K\big\} \\ &\subseteq \big\{(x, t, \widetilde{f}(x,t)) \in \mathbb{R}^n \bigm| (x,t) \in K \times B\big\} \\ &= \operatorname{graph} \widetilde{f}. \end{aligned}$$

Sei $c = \min_{K\times B} \widetilde{f}$. Dann

$$\operatorname{graph} \widetilde{f} \subseteq \partial \left\{(x,t,y) \in \mathbb{R}^n \;\middle|\; (x,t) \in K \times B, c \le y \le \widetilde{f}(x,t)\right\}.$$

Damit folgt die Behauptung aus dem vorigen Satz. □

11.2.3 Iterierte Integrale

Wir möchten jetzt für eine Funktion $f\colon \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}$ das Integral $\int_{\mathbb{R}^{m+n}} f(x,y)\,dxdy$ als iteriertes Integral $\int_{\mathbb{R}^m} \left(\int_{\mathbb{R}^n} f(x,y)\,dy\right) dx$ berechnen. Dies ist nicht immer möglich.

Beispiel 11.22. Sei $f\colon \mathbb{R}^2 \to \mathbb{R}$ definiert durch

$$f(x,y) = \begin{cases} 1/q, & \text{falls } x = p/q,\ p,\ q \in \mathbb{N} \text{ sind relativ prim, } p \le q,\ y \in \mathbb{Q} \cap [0,1], \\ 0, & \text{sonst.} \end{cases}$$

Dann ist f R-integrierbar über $\mathbb{R}^2$ (wobei $\int_{\mathbb{R}^2} f(x,y)\,dxdy = 0$), jedoch für $x = p/q$ wie oben ist die Funktion $y \mapsto f(x,y)$ nicht R-integrierbar über $\mathbb{R}$.

Wir benötigen ein neues Konzept:

Definition 11.23. Sei $B \subseteq \mathbb{R}^n$ ein achsenparalleles Rechteck, $f\colon B \to \mathbb{R}$. Dann definieren wir das *obere* und das *untere Riemann-Integral* von f über B als

$$\overline{\int_B} f(x)\,dx = \inf\left\{\overline{S}(f,P) \;\middle|\; P \text{ ist Partition von } B\right\},$$
$$\underline{\int_B} f(x)\,dx = \sup\left\{\underline{S}(f,P) \;\middle|\; P \text{ ist Partition von } B\right\}.$$

Bemerkung. Ist P' eine Verfeinerung von P, so gilt

$$\underline{S}(f,P) \le \underline{S}(f,P') \le \underline{\int_B} f(x)\,dx \le \overline{\int_B} f(x)\,dx \le \overline{S}(f,P') \le \overline{S}(f,P).$$

Lemma 11.24. *$f\colon B \to \mathbb{R}$ ist genau dann R-integrierbar, falls f beschränkt ist und*

$$\underline{\int_B} f(x)\,dx = \overline{\int_B} f(x)\,dx \left(= \int_B f(x)\,dx\right)$$

gilt.

Beweis. Übungsaufgabe. □

Definition 11.25. Eine beschränkte Funktion $f\colon \mathbb{R}^n \to \mathbb{R}$ heißt *Stufenfunktion*, falls es ein achsenparalleles Rechteck B und eine Partition $\{B_k\}$ von B gibt, so dass $f\big|_{\mathbb{R}^n \setminus B} = 0$ und $f\big|_{B_k^\circ}$ für alle k einen konstanten Wert annimmt. (Die Werte von f auf $\bigcup_k \partial B_k$ spielen oft keine besondere Rolle; typischerweise fordert man, dass f linksseitig stetig ist.)

Satz 11.26. *Die Funktion $f\colon \mathbb{R}^{m+n} \to \mathbb{R}$ habe kompakten Träger und sei R-integrierbar. Dann haben die Funktionen*

$$x \mapsto \overline{\int\limits_{\mathbb{R}^n}} f(x,y)\,dy, \quad x \mapsto \underline{\int\limits_{\mathbb{R}^n}} f(x,y)\,dy$$

kompakten Träger und sind R-integrierbar. Außerdem gilt

$$\begin{aligned}\int_{\mathbb{R}^{m+n}} f(x,y)\,dxdy &= \int_{\mathbb{R}^m} \left(\overline{\int\limits_{\mathbb{R}^n}} f(x,y)\,dy \right) dx \\ &= \int_{\mathbb{R}^m} \left(\underline{\int\limits_{\mathbb{R}^n}} f(x,y)\,dy \right) dx.\end{aligned}$$

Beweis. Sei $B \subset \mathbb{R}^{n+m}$ ein achsenparalleles Rechteck mit $f\big|_{\mathbb{R}^{m+n}\setminus B} = 0$. Seien g, h linksseitig stetige Stufenfunktionen mit Träger in B und $g(x,y) \leq f(x,y) \leq h(x,y)$ für alle $(x,y) \in B$. Dann

$$\int_{\mathbb{R}^n} g(x,y)\,dy \leq \underline{\int\limits_{\mathbb{R}^n}} f(x,y)\,dy \leq \overline{\int\limits_{\mathbb{R}^n}} f(x,y)\,dy \leq \int_{\mathbb{R}^n} h(x,y)\,dy.$$

Weiterhin sind $x \mapsto \int_{\mathbb{R}^n} g(x,y)\,dy$ und $x \mapsto \int_{\mathbb{R}^n} h(x,y)\,dy$ Stufenfunktionen im $\mathbb{R}^m$, und es gilt

$$\begin{aligned}\int_{\mathbb{R}^{m+n}} g(x,y)\,dx\,dy &= \int_{\mathbb{R}^m} \left(\int_{\mathbb{R}^n} g(x,y)\,dy \right) dx \\ &\leq \underline{\int\limits_{\mathbb{R}^m}} \left(\underline{\int\limits_{\mathbb{R}^n}} f(x,y)\,dy \right) dx \leq \overline{\int\limits_{\mathbb{R}^m}} \left(\underline{\int\limits_{\mathbb{R}^n}} f(x,y)\,dy \right) dx \\ &\leq \overline{\int\limits_{\mathbb{R}^m}} \left(\overline{\int\limits_{\mathbb{R}^n}} f(x,y)\,dy \right) dx \leq \int_{\mathbb{R}^m} \left(\int_{\mathbb{R}^n} h(x,y)\,dy \right) dx \\ &= \int_{\mathbb{R}^{m+n}} h(x,y)\,dx\,dy.\end{aligned}$$

Nun gilt

$$\int_{\mathbb{R}^{m+n}} f(x,y)\,dx\,dy = \sup_g \int_{\mathbb{R}^{m+n}} g(x,y)\,dxdy = \inf_h \int_{\mathbb{R}^{m+n}} h(x,y)\,dx\,dy,$$

wobei g, h alle Stufenfunktionen wie oben durchlaufen, also

$$\int_{\mathbb{R}^{m+n}} f(x,y)\,dxdy = \underline{\int\limits_{\mathbb{R}^m}} \left(\underline{\int\limits_{\mathbb{R}^n}} f(x,y)\,dy \right) dx = \overline{\int\limits_{\mathbb{R}^m}} \left(\underline{\int\limits_{\mathbb{R}^n}} f(x,y)\,dy \right) dx,$$

was die Behauptung bezüglich $x \mapsto \underline{\int}_{\mathbb{R}^n} f(x,y)\,dy$ zeigt. Analog argumentiert man bezüglich $x \mapsto \overline{\int}_{\mathbb{R}^n} f(x,y)\,dy$. $\square$

Folgerung. Sei $f\colon \mathbb{R}^{m+n} \to \mathbb{R}$ stetig mit kompaktem Träger. Dann gilt

$$\int_{\mathbb{R}^{m+n}} f(x,y)\,dx\,dy = \int_{\mathbb{R}^m} \left(\int_{\mathbb{R}^n} f(x,y)\,dy \right) dx = \int_{\mathbb{R}^n} \left(\int_{\mathbb{R}^m} f(x,y)\,dx \right) dy.$$

Sei $a_1, b_1 \in \mathbb{R}$, $a_1 \le b_1$, $K_1 = [a_1, b_1]$. Wir nehmen an, dass wir für $j = 2, 3, \dots, n$ induktiv stetige Funktionen

$$a_j, b_j\colon K_{j-1} \to \mathbb{R}, \quad a_j \le b_j,$$

und Mengen $K_j \subset \mathbb{R}^j$ vermöge

$$K_j = \big\{ (x_1, \dots, x_{j-1}, x_j) \in K_{j-1} \times \mathbb{R} \;\big|\; a_j(x_1, \dots, x_{j-1}) \le x_j \le b_j(x_1, \dots, x_{j-1}) \big\}$$

definiert haben.

Satz 11.27. (i) *$K_j \in \mathcal{J}(\mathbb{R}^j)$ für alle $j = 1, \dots, n$.*

(ii) *Für jede stetige Funktion $f\colon K_n \to \mathbb{R}$ gilt*

$$\int_{K_n} f(x)\,dx = \int_{a_1}^{b_1} \int_{a_2(x_1)}^{b_2(x_2)} \cdots \int_{a_n(x_1,\dots,x_{n-1})}^{b_n(x_1,\dots,x_{n-1})} f(x_1, \dots, x_n)\,dx_n \dots dx_2\,dx_1.$$

Beweis. (i) Folgt aus Satz 11.26 mittels Induktion über j.

(ii) Wir schreiben $x = (x', x_n) \in \mathbb{R}^n$ mit $x' = (x_1, \dots, x_{n-1}) \in \mathbb{R}^{n-1}$. Dann

$$\chi_{K_n}(x) = \chi_{K_{n-1}}(x')\ \chi_{[a_n(x'), b_n(x')]}(x_n).$$

Für jedes $x' \in K_{n-1}$ ist die Funktion $x_n \mapsto (\chi_{K_n} f)(x', x_n)$ stetig im Intervall $[a_n(x'), b_n(x')]$, folglich

$$\int_{\mathbb{R}} (\chi_{K_n} f)(x', x_n)\,dx_n = \chi_{K_{n-1}}(x') \int_{a_n(x')}^{b_n(x')} f(x', x_n)\,dx_n.$$

Es ergibt sich

$$\begin{aligned}\int_{K_n} f(x)\,dx &= \int_{\mathbb{R}^n} (\chi_{K_n} f)(x)\,dx \\ &= \int_{\mathbb{R}^{n-1}} \chi_{K_{n-1}}(x') \int_{a_n(x')}^{b_n(x')} f(x', x_n)\,dx_n\,dx' \\ &= \int_{K_{n-1}} \int_{a_n(x')}^{b_n(x')} f(x', x_n)\,dx_n\,dx'.\end{aligned}$$

Der Beweis folgt nun mittels Induktion über n, falls wir zeigen, dass die Funktion

$$x' \mapsto \int_{a_n(x')}^{b_n(x')} f(x', x_n)\, dx_n$$

stetig ist. Dazu definieren wir

$$\begin{aligned}\Delta &= \left\{(a, b, x') \in \mathbb{R}^2 \times K_{n-1} \;\middle|\; (x', a), (x', b) \in K_n,\ a \le b\right\}\\ &= \left\{(a, b, x') \in \mathbb{R}^2 \times K_{n-1} \;\middle|\; a_n(x') \le a \le b \le b_n(x')\right\}\end{aligned}$$

und $F\colon \Delta \to \mathbb{R}$,

$$(a, b, x') \mapsto \int_a^b f(x', x_n)\, dx_n.$$

Dann gilt für (a_0, b_0, x'), $(a_1, b_1, y') \in \Delta$, dass

$$\begin{aligned}&|F(a_0, b_0, x') - F(a_1, b_1, y')|\\ &\quad\le |F(a_0, b_0, x') - F(a_1, b_0, x')| + |F(a_1, b_0, x') - F(a_1, b_1, x')|\\ &\qquad + |F(a_1, b_1, x') - F(a_1, b_1, y')|\\ &\quad\le |a_1 - a_0| \max_{K_n} |f| + |b_1 - b_0| \max_{K_n} |f|\\ &\qquad + \int_{a_1}^{b_1} |f(x', x_n) - f(y', x_n)|\, dx_n.\end{aligned}$$

Die Funktion f ist gleichmäßig stetig auf K, insbesondere existiert für alle $\epsilon > 0$ ein $\delta > 0$, so dass $|f(x', x_n) - f(y', x_n)| \le \epsilon$ für $(x', x_n), (y', x_n) \in K_n$, $|x' - y'| \le \delta$. Daher lässt sich das Integral auf der rechten Seite der obigen Abschätzung für $x', y' \in K_{n-1}$, $|x' - y'| \le \delta$ durch $(b_1 - a_1)\epsilon$ nach oben abschätzen. Es folgt die Stetigkeit der Funktion F auf Δ und damit auch die Stetigkeit der Funktion $x' \mapsto F\left(a_n(x'), b_n(x'), x'\right)$ auf K_{n-1}.

□

Beispiel 11.28. Sei $K = \left\{(x, y) \in \mathbb{R}^2 \;\middle|\; x \ge 0,\, y \ge 0,\, x + y \le 1\right\}$, $f\colon K \to \mathbb{R}$ stetig. Dann

$$\int_K f(x, y)\, dx\, dy = \int_0^1 \left[\int_0^{1-x} f(x, y)\, dy\right] dx.$$

11.2.4 Variablenwechsel unter dem Integral

Formulierung des Satzes

Sei $\varphi\colon [a, b] \to [c, d]$ eine bijektive C^1-Abbildung mit $\varphi'(x) \ne 0$ für alle $x \in [a, b]$. Dann gilt für jede R-integrierbare Funktion $f\colon [c, d] \to \mathbb{R}$, dass die Funktion $x \mapsto f\left(\varphi(x)\right) |\varphi'(x)|$ R-integrierbar auf $[a, b]$ ist und

$$\int_c^d f(y)\, dy = \int_a^b f\left(\varphi(x)\right) |\varphi'(x)|\, dx.$$

Diese Formel verallgemeinert sich auf den Fall höherer Dimensionen.

Satz 11.29. *Seien $U, V \subseteq \mathbb{R}^n$ offene Mengen und $\varphi\colon U \to V$ eine bijektive C^1-Abbildung mit $\det D\varphi(x) \neq 0$ für alle $x \in U$. Ferner sei $f\colon V \to \mathbb{R}$ eine R-integrierbare Funktion mit kompaktem Träger. Dann ist die Funktion $x \mapsto f(\varphi(x))\,|\det D\varphi(x)|$ R-integrierbar auf U, hat kompakten Träger und es gilt*

$$\int_V f(y)\,dy = \int_U f(\varphi(x))\,|\det D\varphi(x)|\,dx.$$

Folgerung. Ist $K \in \mathcal{J}(U)$, so $\varphi(K) \in \mathcal{J}(V)$ und

$$\mathrm{vol}_n(\varphi(K)) = \int_K |\det D\varphi(x)|\,dx.$$

Insbesondere ist φ *volumenerhaltend*, falls $|\det D\varphi(x)| = 1$ für alle $x \in U$ gilt.

Zerlegungen der Eins

Wir bereiten jetzt den Beweis des Satzes vor.

Definition 11.30. Sei $K \subset \mathbb{R}^n$ kompakt und $\{O_j\}_{j=1}^J$ eine offene Überdeckung von K, d. h. $K \subseteq \bigcup_{j=1}^J O_j$. Eine stetige *Zerlegung der Eins* über K, die sich $\mathcal{O}$ unterordnet, sind J stetige Funktionen $\chi_j\colon \mathbb{R}^n \to \mathbb{R}$ mit kompaktem Träger, so dass

(i) $0 \leq \chi_j \leq 1$ für alle j,

(ii) $\operatorname{supp}\chi_j \subset O_j$ für alle j,

(iii) $\sum_{j=1}^J \chi_j(x) = 1$ für $x \in K$.

Satz 11.31. *Für jede kompakte Menge $K \subset \mathbb{R}^n$ und jede endliche offene Überdeckung $\mathcal{O}$ von K existiert eine Zerlegung der Eins über K, die sich $\mathcal{O}$ unterordnet.*

Beweis. Für je zwei achsenparallele Rechtecke B, B' mit $B \subseteq (B')^\circ$ existiert eine stückweise lineare (insbesondere stetige) Funktion $g_{BB'}\colon \mathbb{R}^n \to \mathbb{R}$ mit $\operatorname{supp} g_{BB'} = B'$, $0 \leq g_{BB'} \leq 1$ und $g_{BB'}\big|_B = 1$ (siehe Abbildung 11.4).
Wähle nun für jedes $x \in K$ achsenparallele Rechtecke B_x, B'_x mit $x \in B_x^\circ$ und

$$B_x \subset (B'_x)^\circ \subset B'_x \subset O_j$$

für wenigstens ein $j \in \{1, \ldots, J\}$. Die offenen Mengen B_x° für $x \in K$ überdecken K, folglich gibt es endlich viele $x_1, x_2, \ldots, x_L \in K$ mit $K \subseteq \bigcup_{l=1}^L B_{x_l}^\circ$. Wir setzen dann

$$\psi_l = g_{B_{x_l}B'_{x_l}}$$

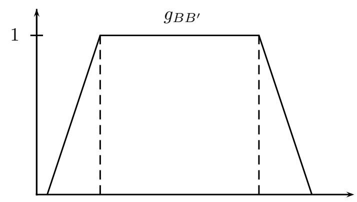

Abbildung 11.4: Eine spezielle Abschneidefunktion

für $l = 1, \ldots, L$ und

$$\widetilde{\chi}_1 = \psi_1, \quad \widetilde{\chi}_{l+1} = (1-\psi_1)(1-\psi_2)\cdots(1-\psi_l)\psi_{l+1}$$

für $l = 1, \ldots, L-1$. Dann ist offenbar

$$\sum_{l=1}^{L'} \widetilde{\chi}_l = 1 - \prod_{l=1}^{L'}(1-\psi_l)$$

für $L' = 1, \ldots, L$. Insbesondere folgt für $L' = L$, dass $\sum\limits_{l=1}^{L} \widetilde{\chi}_l = 1$ auf K (für jedes $x \in K$ gibt es ein l mit $\psi_l(x) = 1$). Wählen wir nun eine Abbildung $\{1, \ldots, L\} \to \{1, \ldots, J\}$, $l \mapsto j(l)$, mit $B'_{x_l} \subset O_{j(l)}$, so bleibt nur

$$\chi_j = \sum_{j(l)=j} \widetilde{\chi}_l$$

zu setzen. □

Approximation R-integrierbarer Funktionen

Sei $K \subset \mathbb{R}^n$ kompakt, $\delta > 0$. Wir setzen

$$K_\delta = \left\{ x \in \mathbb{R}^n \;\middle|\; \exists x' \in K\colon\ |x - x'| \leq \delta \right\}.$$

Lemma 11.32. *K_δ ist kompakt. Ferner existiert für jede offene Menge $U \subseteq \mathbb{R}^n$ mit $K \subset U$ ein $\delta > 0$, so dass $K_\delta \subset U$.*

Beweis. Übungsaufgabe. □

Satz 11.33. *Sei $U \subseteq \mathbb{R}^n$ offen, $f\colon U \to \mathbb{R}^n$ beschränkt mit kompaktem Träger. Dann ist f genau dann R-integrierbar, falls es für jedes $\epsilon > 0$ stetige Funktionen $g, h\colon U \to \mathbb{R}$ mit kompakten Trägern und $g(x) \leq f(x) \leq h(x)$ für alle $x \in U$ gibt, so dass*

$$\int_U (h(x) - g(x))\, dx \leq \epsilon.$$

In diesem Fall gilt

$$\int_U h(x)\, dx - \epsilon \leq \int_U f(x)\, dx \leq \int_U g(x)\, dx + \epsilon.$$

Beweis. Die angegebene Bedingung ist sicherlich hinreichend für die Riemann-Integrierbarkeit von f; wir zeigen, dass sie auch notwendig ist.

(i) Wir behandeln zuerst den Fall $U = \mathbb{R}^n$. Sei also $f\colon \mathbb{R}^n \to \mathbb{R}$ R-integrierbar, $\operatorname{supp} f \subseteq B$ für ein achsenparalleles Rechteck B. Indem wir gegebenenfalls eine Konstante addieren, dürfen wir $f \geq 0$ annehmen.

Dann gibt es zu $\epsilon > 0$ eine Partition $\big\{B_k\big\}_{k=1}^K$ von B mit

$$\sum_k \left(\sup_{B_k} f - \inf_{B_k} f \right) \operatorname{vol}_n(B_k) \leq \epsilon.$$

Wie im vorigen Beweis findet man stückweise lineare (insbesondere stetige) Funktionen $g_k, h_k\colon \mathbb{R}^n \to \mathbb{R}$ mit kompaktem Träger, so dass $g_k(x) \leq \chi_{B_k}(x) \leq h_k(x)$ für alle $x \in \mathbb{R}^n$ und

$$\int_{\mathbb{R}^n} (h_k(x) - g_k(x))\; dx \leq \frac{\epsilon}{(1 + \sup|f|)\, K}.$$

Dann haben die Funktionen

$$\begin{aligned} g(x) &= \sum_k \inf_{B_k} f \cdot g_k(x) \\ h(x) &= \sum_k \sup_{B_k} f \cdot h_k(x) \end{aligned}$$

die gewünschten Eigenschaften (mit ϵ ersetzt durch 3ϵ). In der Tat gilt

$$\begin{aligned} \int (h(x) - g(x))\; dx = &\sum_k \sup_{B_k} f \int (h_k(x) - \chi_{B_k}(x))\; dx \\ &+ \sum_k \inf_{B_k} f \int (\chi_{B_k}(x) - g_k(x))\; dx \\ &+ \sum_k \left(\sup_{B_k} f - \inf_{B_k} f \right) \operatorname{vol}_n(B_k) \\ \leq\; &\epsilon + \epsilon + \epsilon = 3\epsilon. \end{aligned}$$

(ii) Für beliebiges $U \subseteq \mathbb{R}^n$ offen existiert ein $\delta > 0$, so dass $(\operatorname{supp} f)_{2\delta} \subseteq U$. Indem wir in (i) eine Partition $\{B_k\}$ der Feinheit höchstens δ und die h_k mit $\operatorname{supp} h_k \subseteq (B_k)_\delta$ wählen, erhalten wir wie im Beweisschritt (i) eine Funktion h (und g) mit $\operatorname{supp} h \subset U$.

□

Bemerkung. Es lässt sich zeigen, dass eine beschränkte Funktion $f\colon \mathbb{R}^n \to \mathbb{R}$ mit kompaktem Träger genau dann R-integrierbar ist, falls die Menge

$$\big\{ x \in \mathbb{R}^n \;\big|\; f \text{ ist } \textit{nicht stetig} \text{ an } x \big\}$$

n-dimensionales Volumen Null hat.

Beweis des Satzes

Wir beweisen jetzt den Satz über den Variablenwechsel unter dem Integral in fünf Schritten.

Schritt 1: Zurückführen auf stetige Integranden. Wir nehmen an, wir hätten den Satz für stetiges f bereits gezeigt.

Sei dann $f\colon V \to \mathbb{R}$ R-integrierbar. Sei $\epsilon > 0$. Dann existieren stetige Funktionen $g, h\colon V \to \mathbb{R}$ mit kompaktem Träger in V, so dass $g(y) \leq f(y) \leq h(y)$ für alle $y \in V$ und $\int_V (h(y) - g(y))\, dy \leq \epsilon$. Dann sind die Funktionen $x \mapsto g(\phi(x))\, |\det D\phi(x)|$, $x \mapsto h(\phi(x))\, |\det D\phi(x)|$ stetig auf U, haben kompakten Träger und genügen

$$\begin{aligned}\int_U & (h(\phi(x))\, |\det D\phi(x)| - g(\phi(x))\, |\det D\phi(x)|)\, dx \\ &= \int_V (h(y) - g(y))\, dy \leq \epsilon.\end{aligned}$$

Zudem gilt

$$g(\phi(x))\, |\det D\phi(x)| \leq f(\phi(x))\, |\det \phi(x)| \leq h(\phi(x))\, |\det D\phi(x)|$$

für alle $x \in U$. Daraus folgt, dass $x \mapsto f(\phi(x))\, |\det D\phi(x)|$ R-integrierbar ist und

$$\int_V f(y)\, dy = \int_U f(\phi(x))\, |\det D\phi(x)|\, dx.$$

Schritt 2: Reduktion auf den eindimensionalen Fall. Wir benutzen hier den Satz über die implizite Funktion. Sei $x^0 \in U$ fixiert. O. B. d. A. dürfen wir $\frac{\partial \phi_n}{\partial x_n}(x^0) \neq 0$ annehmen. Dann können wir die Gleichung

$$\phi_n(x_1, \ldots, x_{n-1}, x_n) = y_n$$

in einer Umgebung von x^0 nach x_n auflösen,

$$x_n = h_n(x_1, \ldots, x_{n-1}, y_n).$$

Wir betrachten dann σ_n definiert durch

$$\sigma_n(x) = (x_1, \ldots, x_{n-1}, \phi_n(x)).$$

σ_n ist eine bijektive C^1-Abbildung einer hinreichend kleinen Umgebung von x^0 auf eine Umgebung von $\phi(x^0)$. Es gilt

$$\begin{aligned}\left(\phi \circ \sigma_n^{-1}(x_1, \ldots, x_{n-1}, y_n)\right)_n &= \phi_n\left(x_1, \ldots, x_{n-1}, h_n(x_1, \ldots, x_{n-1}, y_n)\right) \\ &= y_n,\end{aligned}$$

wohingegen für $1 \leq j \leq n-1$

$$\left(\phi \circ \sigma_n^{-1}(x_1, \ldots, x_{n-1}, y_n)\right)_j = \phi_j\left(x_1, \ldots, x_{n-1}, h_n(x_1, \ldots, x_{n-1}, y_n)\right).$$

Folglich lässt $\phi \circ \sigma_n^{-1}$ die n-te Koordinate invariant und wirkt als von y_n abhängige bijektive C^1-Abbildung in den restlichen Variablen $y_1, \ldots, y_{n-1}$.

Da wir $\phi = \left(\phi \circ \sigma_n^{-1}\right) \circ \sigma_n$ schreiben können, wird ϕ schließlich in einer kleinen Umgebung von x^0 die Komposition endlich vieler derartiger σ_j für $j = 2, \ldots, n$.

Schritt 3: Lokalisierung. Wir zeigen jetzt, dass das Theorem gilt, falls wir für jedes $x \in \operatorname{supp} f$ eine Umgebung finden, so dass der Satz für stetige Funktionen mit kompaktem Träger in dieser Umgebung gilt.

Sei O_x für $x \in \operatorname{supp} f$ eine derartige Familie offener Mengen mit $x \in O_x$. Endlich viele $O_{x_1}, \ldots, O_{x_J}$ überdecken $\operatorname{supp} f$, zudem finden wir weiterhin eine Zerlegung $\left\{\chi_j\right\}_{j=1}^J$ der Eins über K, die sich dieser offenen Überdeckung unterordnet. Dann gilt nach Voraussetzung

$$\int_V \left(\chi_j f\right)(y)\, dy = \int_U \left(\chi_j f\right)(\phi(x)) \left|\det D\phi(x)\right|\, dx$$

für alle j. Indem wir über die j summieren und $\sum\limits_j \chi_j(x) = 1$ für $x \in \operatorname{supp} f$ ausnutzen, erhalten wir die Behauptung für f.

Schritt 4: Kompositionen Wir nehmen an, wir haben zwei Koordinatenwechsel $\phi\colon U \to V$ und $\psi\colon V \to W$ und das Ergebnis sei für ϕ, ψ richtig. Dann gilt es auch für die Komposition $\psi \circ \phi\colon U \to W$.

Tatsächlich haben wir nach der Kettenregel, dass

$$\begin{aligned} \det D\left(\psi \circ \phi\right)(x) &= \det\left(\left(D\psi\right)(\phi(x))\, D\phi(x)\right) \\ &= \det\left(D\psi\right)(\phi(x)) \cdot \det D\phi(x), \end{aligned}$$

also

$$\begin{aligned} \int_W f(z)\, dz &= \int_V f\left(\psi(y)\right) \left|\det D\psi(y)\right|\, dy \\ &= \int_U f\left(\left(\psi \circ \phi\right)(x)\right) \left|\det D\psi\left(\phi(x)\right)\right| \left|\det D\phi(x)\right|\, dx \\ &= \int_U f\left(\left(\psi \circ \phi\right)(x)\right) \left|\det D\left(\psi \circ \phi\right)(x)\right|\, dx. \end{aligned}$$

Schritt 5: Der eindimensionale Fall Wegen der bisherigen Vorbereitungen dürfen wir jetzt

$$\phi(x) = \left(x_1, \ldots, x_{n-1}, h(x)\right)$$

mit einer C^1-Funktion h und f stetig annehmen. Wegen

$$D\phi(x) = \begin{pmatrix} 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \\ \frac{\partial h}{\partial x_1} & \frac{\partial h}{\partial x_2} & \cdots & \frac{\partial h}{\partial x_{n-1}} & \frac{\partial h}{\partial x_n} \end{pmatrix}$$

gilt $\frac{\partial h}{\partial x_n} \neq 0$. Folglich ist die Abbildung $x_n \mapsto h(x_1, \dots, x_{n-1}, x_n)$ bei festem $(x_1, \dots, x_{n-1})$ strikt monoton (wachsend oder fallend).

In jedem Fall gilt

$$\begin{aligned}
&\int_U f\left(\phi(x)\right) |\det D\phi(x)| \, dx \\
&\quad = \int_U f(x_1, \dots, x_{n-1}, h(x)) \left| \frac{\partial h}{\partial x_n}(x) \right| \, dx \\
&\quad = \int \dots \left(\int f(x_1, \dots, x_{n-1}, h(x)) \left| \frac{\partial h}{\partial x_n}(x) \right| \, dx_n \right) \dots dx_1 \\
&\quad = \int \dots \left(\int f(y_1, \dots, y_n) \, dy_n \right) \dots dy_1 = \int_V f(y) \, dy.
\end{aligned}$$

11.2.5 Absolute Riemann-Integrierbarkeit

Wir wollen jetzt von der Einschränkung wegkommen, dass der Integrand kompakten Träger hat.

Definition 11.34. $f\colon U \to \mathbb{R}$, wobei $U \subseteq \mathbb{R}^n$ offen, heißt *lokal Riemann-integrierbar*, falls für alle $K \in \mathcal{J}(U)$ die Funktion $\chi_K f$ Riemann-integrierbar ist.

Bemerkung. Eine äquivalente Bedingung ist, dass für jede stetige Funktion $h\colon U \to \mathbb{R}$ mit kompaktem Träger die Funktion hf R-integrierbar ist. Insbesondere sind stetige Funktionen $f\colon U \to \mathbb{R}$ lokal R-integrierbar.

Für eine Funktion $f\colon U \to \mathbb{R}$ setzen wir

$$f_+ = \max\{f, 0\}, \quad f_- = \max\{-f, 0\}.$$

Dann gilt

$$f_+ \geq 0, \quad f_- \geq 0, \quad f = f_+ - f_-, \quad |f| = f_+ + f_-.$$

Definition 11.35. Sei $U \subseteq \mathbb{R}^n$ offen, $f\colon U \to \mathbb{R}$. Dann heißt f *absolut R-integrierbar* über U, falls f lokal R-integrierbar und

$$\sup_{K \in \mathcal{J}(U)} \int_K |f(x)| \, dx < \infty$$

ist. In diesem Fall setzen wir

$$\int_U f(x) \, dx = \sup_{K \in \mathcal{J}(U)} \int_K f_+(x) \, dx - \sup_{K \in \mathcal{J}(U)} \int_K f_-(x) \, dx.$$

Bemerkung. Um diese Definition zu rechtfertigen, müssen wir zeigen, dass sich für $f\colon U \to \mathbb{R}$ R-integrierbar mit kompaktem Träger tatsächlich der alte Wert für $\int_U f(x) \, dx$ ergibt.

Beweis. Wir dürfen $f \geq 0$ annehmen. Dann gilt für jedes $K \in \mathcal{J}(U)$, dass

$$\int_K f(x)\,dx = \int_{K \cap \operatorname{supp} f} f(x)\,dx \leq \int_U f(x)\,dx.$$

Weiterhin gilt $\operatorname{supp} f \in \mathcal{J}(U)$ und

$$\int_{\operatorname{supp} f} f(x)\,dx = \int_U f(x)\,dx.$$

□

Bemerkung. Die Klasse der absolut R-integrierbaren Funktionen über U bilden einen $\mathbb{R}$-Vektorraum. Diese Klasse ist jedoch nicht abgeschlossen unter punktweiser Multiplikation. So ist die Funktion $x \mapsto \frac{1}{\sqrt{x}}$ absolut R-integrierbar über $(0,1)$, die Funktion $x \mapsto \left(\frac{1}{\sqrt{x}}\right)^2 = \frac{1}{x}$ jedoch nicht.

Lemma 11.36. *Sei $U \subseteq \mathbb{R}^n$ offen. Dann gibt es eine Folge $\{K_k\}$ in $\mathcal{J}(U)$ mit $\bigcup_{k \in \mathbb{N}} K_k = U$ und $K_k \subset K_{k+1}^\circ$ für alle k.*

Beweis. Für $k \in \mathbb{N}$ setzen wir

$$C_k = \bigl\{x \in U \bigm| |x| \leq k,\ \operatorname{dist}(x, \partial U) \geq 1/k\bigr\},$$

wobei $\operatorname{dist}(x, \partial U) = \inf\bigl\{|x-y| \bigm| y \in \partial U\bigr\}$. Die Mengen C_k sind kompakt (nach dem Satz von Heine-Borel), außerdem überdecken diese Mengen U. Da U offen ist, ist $\operatorname{dist}(x, \partial U) > 0$ für jedes $x \in U$. Weiterhin gilt

$$C_k \subset \bigl\{x \in U \bigm| |x| < k+1,\ \operatorname{dist}(x, \partial U) > 1/(k+1)\bigr\} \subseteq C_{k+1}^\circ.$$

Die Mengen C_k sind möglicherweise nicht Jordan-messbar. Um Mengen in $\mathcal{J}(U)$ zu erhalten, wählen wir für jedes $x \in C_k$ ein achsenparalleles Rechteck in C_{k+1}° mit Mittelpunkt x. Wir wählen dann endlich viele dieser Rechtecke, deren Inneren die Menge C_k überdecken. Sei K_k die Vereinigung dieser endlich vielen Rechtecke. Dann ist $K_k \in \mathcal{J}(U)$ sowie

$$C_k \subset K_k^\circ \subset K_k \subset C_{k+1}^\circ$$

und damit

$$\bigcup_k C_k = \bigcup_k K_k = U.$$

□

Satz 11.37. *Sei $U \subseteq \mathbb{R}^n$ offen, $f\colon U \to \mathbb{R}$ lokal R-integrierbar, $\{K_k\}$ wie im vorigen Lemma. Dann sind äquivalent:*

(i) *f ist absolut Riemann-integrierbar über U.*

(ii) $\left\{\int_{K_k} |f(x)|\, dx\right\}_{k\in\mathbb{N}}$ *ist eine beschränkte Folge.*

Unter diesen Bedingungen ist die Folge $\left\{\int_{K_k} |f(x)|\, dx\right\}_{k\in\mathbb{N}}$ *monoton wachsend und*

$$\lim_{k\to\infty} \int_{K_k} f(x)\, dx = \int_U f(x)\, dx.$$

Beweis. (i)⇒(ii). Da f genau dann absolut R-integrierbar ist, wenn f_+, f_- absolut R-integrierbar sind, dürfen wir $f \geq 0$ annehmen. (ii) folgt dann, da

$$\int_{K_k} f(x)\, dx \leq \sup_{K\in\mathcal{J}(U)} \int_K f(x)\, dx = \int_U f(x)\, dx.$$

(ii)⇒(i). Sei $K \in \mathcal{J}(U)$ beliebig. Die Mengen K_k° überdecken K, folglich überdecken endlich viele dieser K_k° die kompakte Menge K. Wegen $K_k^\circ \subseteq K_{k+1}^\circ$ gibt es somit ein N mit $K \subseteq K_N^\circ$. Es folgt (wiederum unter der Annahme $f \geq 0$)

$$\int_K f(x)\, dx \leq \int_{K_N} f(x)\, dx \leq \lim_{k\to\infty} \int_{K_k} f(x)\, dx.$$

Folglich ist f absolut R-integrierbar über U. Zudem gilt

$$\int_U f(x)\, dx = \sup_{K\in\mathcal{J}(U)} \int_K f(x)\, dx = \lim_{k\to\infty} \int_{K_k} f(x)\, dx.$$

□

Folgerung. Seien $U, V \subseteq \mathbb{R}^n$ offen und sei $\varphi\colon U \to V$ eine bijektive C^1–Abbildung. Dann ist f genau dann absolut R-integrierbar über V, wenn $(f \circ \varphi)\, |\det D\varphi|$ absolut R-integrierbar über U ist. In diesem Fall gilt

$$\int_V f(y)\, dy = \int_U f\left(\varphi(x)\right) |\det D\varphi(x)|\; dx.$$

Beweis. Approximiere f durch R-integrierbare Funktionen mit kompaktem Träger. □

Beispiel 11.38.

$$\int_{-\infty}^{\infty} \mathrm{e}^{-x^2}\, dx = \sqrt{\pi}.$$

Die Funktion $f\colon \mathbb{R}^2 \to \mathbb{R}$, $f(x,y) = \mathrm{e}^{-(x^2+y^2)}$ ist stetig und somit lokal R-integrierbar. Weiterhin gilt in Polarkoordinaten (r,θ) mit $x = r\cos\theta$, $y = r\sin\theta$

$$\begin{aligned}\int_{|x,y|\leq R} f(x,y)\, dx\, dy &= \int_0^{2\pi} \int_0^R r\mathrm{e}^{-r^2}\, dr\, d\theta \\ &= 2\pi \left.\frac{\mathrm{e}^{-r^2}}{2}\right|_0^R = \pi\left(1 - \mathrm{e}^{-R^2}\right) \leq \pi.\end{aligned}$$

Wir schlussfolgern, dass f absolut R-integrierbar über $\mathbb{R}^2$ ist und

$$\int_{\mathbb{R}^2} \mathrm{e}^{-(x^2+y^2)}\, dx\, dy = \lim_{R\to\infty} \int_{|x,y|\le R} f(x,y)\, dx\, dy = \pi.$$

Andererseits gilt

$$\begin{aligned}\int_{\mathbb{R}^2} \mathrm{e}^{-(x^2+y^2)}\, dx\, dy &= \lim_{R\to\infty} \int_{[-R,R]^2} f(x,y)\, dx\, dy \\ &= \lim_{R\to\infty} \left(\int_{-R}^{R} \mathrm{e}^{-x^2}\, dx\right)\left(\int_{-R}^{R} \mathrm{e}^{-y^2}\, dy\right) \\ &= \lim_{R\to\infty} \left(\int_{-R}^{R} \mathrm{e}^{-x^2}\, dx\right)^2 = \left(\int_{-\infty}^{\infty} \mathrm{e}^{-x^2}\, dx\right)^2.\end{aligned}$$

Also gilt (da das Integral positiv ist)

$$\int_{-\infty}^{\infty} \mathrm{e}^{-x^2}\, dx = \sqrt{\pi}.$$

11.2.6 Weitere Eigenschaften des absoluten R-Integrals

Wir beschäftigen uns jetzt u. a. mit der Vertauschbarkeit von Integralen und Limesbildung.

Lemma 11.39. *Sei $K \in \mathcal{J}(\mathbb{R}^n)$ und $\{f_k\}$ eine Folge reeller Funktionen auf K mit folgenden Eigenschaften:*

(i) *$\{f_k(x)\}$ ist monoton fallend für jedes $x \in K$, d. h. $f_{k+1}(x) \le f_k(x)$ für alle $k \in \mathbb{N}$, $x \in K$.*

(ii) *$\lim\limits_{k\to\infty} f_k(x) = 0$ für jedes $x \in K$.*

(iii) *f_1 ist beschränkt.*

Dann gilt

$$\lim_{k\to\infty} \underline{\int_K} f_k(x)\, dx = 0.$$

Beweis. Sei $\epsilon > 0$. Wir finden stetige Funktionen g_k auf K mit $0 \le g_k \le f_k$ und

$$\underline{\int_K} f_k(x)\, dx \le \int_K g_k(x)\, dx + \frac{\epsilon}{2^{k+1}}.$$

Wir definieren dann stetige Funktionen h_k auf K vermöge $h_1 = g_1$ und $h_k = \min\{g_k, h_{k-1}\}$ für $k \ge 2$. Die Folge $\{h_k\}$ ist punktweise monoton fallend und $\lim\limits_{k\to\infty} h_k(x) = 0$ für alle $x \in K$ (da $h_k(x) \le f_k(x)$). Weiterhin

$$\begin{aligned}f_k - h_k &= f_k - \min\{g_k, h_{k-1}\} = \max\{f_k - g_k, f_k - h_{k-1}\} \\ &\le (f_k - g_k) + (f_k - h_{k-1}) \le (f_k - g_k) + (f_{k-1} - h_{k-1}),\end{aligned}$$

also induktiv

$$f_k - h_k \leq \sum_{1 \leq j \leq k} (f_j - g_j)$$

und

$$\begin{aligned}
\underline{\int}_K f_k(x)\,dx &= \int_K h_k(x)\,dx + \underline{\int}_K (f_k - h_k)(x)\,dx \\
&\leq \int_K h_k(x)\,dx + \sum_{j=1}^{k} \underline{\int}_K (f_j - g_j)(x)\,dx \\
&\leq \int_K h_k(x)\,dx + \sum_{j=1}^{k} \frac{\epsilon}{2^{j+1}} \\
&< \int_K h_k(x)\,dx + \frac{\epsilon}{2}.
\end{aligned}$$

Die Folge $\{h_k\}_{k\in\mathbb{N}}$ konvergiert *gleichmäßig* auf K gegen Null (dies war eine Klausuraufgabe!), insbesondere existiert ein k_0 mit

$$\int_K h_k(x)\,dx \leq \frac{\epsilon}{2}$$

für alle $k \geq k_0$. Es folgt

$$0 \leq \underline{\int}_K f_k(x)\,dx < \epsilon$$

für alle $k \geq k_0$. □

Satz 11.40. *Sei $K \in \mathcal{J}(\mathbb{R}^n)$, $\{f_k\}$ eine beschränkte Folge R-integrierbarer Funktionen auf K mit $\lim_{k\to\infty} f_k(x) = f(x)$ für alle $x \in K$. Weiterhin sei f R-integrierbar. Dann gilt*

$$\int_K f(x)\,dx = \lim_{k\to\infty} \int_K f_k(x)\,dx.$$

Beweis. Indem wir f_k durch $f_k - f$ ersetzen, dürfen wir $f = 0$ annehmen. Indem wir $f_k = (f_k)_+ - (f_k)_-$ schreiben, dürfen wir weiterhin $f_k \geq 0$ für alle k annehmen.
Die Funktionen $g_k = \sup\{f_{k+j} \mid j \in \mathbb{N}_0\}$ sind reellwertig (da die Folge $\{f_k\}$ auf K beschränkt ist). Die Folge $\{g_k\}$ ist monoton fallend und konvergiert punktweise gegen Null, außerdem ist g_1 beschränkt. Damit

$$0 \leq \lim_{k\to\infty} \int_K f_k(x)\,dx \leq \lim_{k\to\infty} \underline{\int}_K g_k(x)\,dx = 0.$$

□

Wir erhalten damit folgendes Resultat:

Theorem 11.41 (Arzelà). *Sei $U \subseteq \mathbb{R}^n$ offen, $\{f_k\}$ eine Folge über U absolut R-integrierbarer Funktionen. Es gelte:*

(i) $\lim_{k\to\infty} f_k(x) = f(x)$ *für alle $x \in U$ und f ist absolut R-integrierbar über U.*

(ii) *Es gibt eine absolut R-integrierbare Funktion $g\colon U \to \mathbb{R}$ mit $|f_k(x)| \le g(x)$ für alle $x \in U$ und $k \in \mathbb{N}$.*

Dann gilt

$$\int_U f(x)\,dx = \lim_{k\to\infty} \int_U f_k(x)\,dx.$$

Beweis. Sei $\epsilon > 0$ beliebig. Es gibt dann ein $K \in \mathcal{J}(U)$ mit

$$\int_{U\setminus K} |f(x)|\,dx \le \frac{\epsilon}{3}, \quad \int_{U\setminus K} |f_k(x)|\,dx \le \int_U g(x)\,dx \le \frac{\epsilon}{3}.$$

Gemäß dem vorigen Satz existiert ein k_0, so dass

$$\left|\int_K f(x)\,dx - \int_K f_k(x)\,dx\right| \le \frac{\epsilon}{3}$$

für alle $k \ge k_0$, also

$$\begin{aligned}&\left|\int_U f(x)\,dx - \int_U f_k(x)\,dx\right| \\ &\quad \le \left|\int_K f(x)\,dx - \int_K f_k(x)\,dx\right| + \int_{U\setminus K} |f(x)|\,dx + \int_{U\setminus K} |f_k(x)|\,dx \le \epsilon\end{aligned}$$

für alle $k \ge k_0$. □

Satz 11.42. *Seien $U \subseteq \mathbb{R}^m$, $V \subseteq \mathbb{R}^n$ offen und sei $f\colon U\times V \to \mathbb{R}$ eine Funktion mit folgenden Eigenschaften:*

(i) *Für jedes $x \in U$ ist die Funktion $y \mapsto f(x,y)$ absolut R-integrierbar über V.*

(ii) *Die totale Ableitung $D_x f\colon U \times V \to \mathcal{L}(\mathbb{R}^m,\mathbb{R})$ existiert, wobei für jedes $x \in U$ die Abbildung $y \mapsto (D_x f)(x,y)$ (komponentenweise) absolut R-integrierbar über V und für jedes $y \in V$ die Abbildung $x \mapsto (D_x f)(x,y)$ stetig ist.*

(iii) *Es gibt eine absolut R-integrierbare Funktion $g\colon V \to \mathbb{R}$ mit $|D_x f(x,y)| \le g(y)$ für alle $(x,y) \in U \times V$.*

Dann ist die durch $F(x) = \int_V f(x,y)\,dy$ definierte Funktion $F\colon U \to \mathbb{R}$ differenzierbar und

$$DF(x) = \int_V (D_x f)(x,y)\,dy, \quad x \in U.$$

Beweis. Sei $x^0 \in U$, $h \in \mathbb{R}^m$ mit $x^0 + sh \in U$ für alle $s \in [0,1]$. Nach dem Fundamentalsatz der Differential- und Integralrechung gilt

$$f(x^0 + h, y) - f(x^0, y) = \left(\int_0^1 D_x f(x^0 + sh, y)\, ds \right) h.$$

Daher gilt

$$\begin{aligned} F(x^0 + h) - F(x^0) - &\left(\int_V (D_x f)(x^0, y)\, dy \right) h \\ &= \left(\int_V \big(f(x^0 + h, y) - f(x^0, y) - (D_x f)(x^0, y) \big)\, dy \right) h \\ &= \left(\int_V \int_0^1 \big[(D_x f)(x^0 + sh, y) - (D_x f)(x^0, y) \big]\, dsdy \right) h. \end{aligned}$$

Wir müssen zeigen, dass

$$\lim_{h \to 0} \int_V \int_0^1 \big[(D_x f)(x^0 + sh, y) - (D_x f)(x^0, y) \big]\, dsdy = 0.$$

Dazu wählen wir eine beliebige Folge $\{h_k\} \subset \mathbb{R}^m$ mit $h_k \to 0$ für $k \to \infty$. Dann $(D_x f)(x^0 + sh_k, y) - (D_x f)(x^0, y) \to 0$ für $k \to \infty$ und alle $(s, y) \in [0,1] \times V$. Außerdem gilt

$$\big| (D_x f)(x^0 + sh_k, y) - (D_x f)(x^0, y) \big| \le 2g(y),$$

wobei die Funktion $[0,1] \times V \to \mathbb{R}$, $(s, y) \mapsto 2g(y)$ absolut R-integrierbar ist. Mittels des vorigen Satzes folgt

$$\lim_{k \to \infty} \int_V \int_0^1 \big[(D_x f)(x^0 + sh_k, y) - (D_x f)(x^0, y) \big]\, dsdy = 0,$$

und daraus folgt die Behauptung. □

11.3 Integralsätze

11.3.1 Untermannigfaltigkeiten des $\mathbb{R}^n$

Definition 11.43. Sei $M \subseteq \mathbb{R}^n$, $M \neq \emptyset$, $d \in \mathbb{N}_0$, $d \le n$ und $k \in \mathbb{N} \cup \{\infty\}$. Dann heißt M eine *C^k-Untermannigfaltigkeit* von $\mathbb{R}^n$ der Dimension d, falls für jedes $x^0 \in M$ eine Umgebung $U \subseteq \mathbb{R}^n$ von x^0 und eine C^k-Abbildung f einer offenen Teilmenge $V \subseteq \mathbb{R}^d$ nach $\mathbb{R}^{n-d}$ existieren, so dass

$$M \cap U = \big\{ (y, f(y)) \;\big|\; y \in V \big\}.$$

Hierbei ist möglicherweise die Reihenfolge der Koordinaten $(x_1, x_2, \ldots, x_n)$ im $\mathbb{R}^n$ zu ändern.

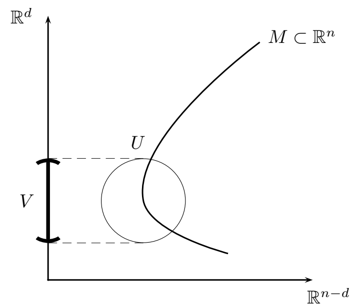

Abbildung 11.5: Zur Definition einer C^k-Untermannigfaltigkeit

Beispiel 11.44. (i) $d = 1$: C^k-Kurven.

(ii) $d = 2$: C^k-Flächen.

(iii) $d = n-1$: C^k-Hyperflächen. So ist $\mathbb{S}^{n-1} = \{x \in \mathbb{R}^n \mid |x| = 1\}$ eine C^∞-Hyperfläche im $\mathbb{R}^n$. Um dies zu sehen, wählen wir einen beliebigen Punkt $x^0 \in S^{n-1}$, beispielsweise $x^0 = (0, \dots, 0, 1)$. Mit $U = \{x \in \mathbb{R}^n \mid x_n > 0\}$ gilt dann

$$\begin{aligned} \mathbb{S}^{n-1} \cap U &= \{x \in \mathbb{R}^n \mid |x| = 1,\, x_n > 0\} \\ &= \{(x', \sqrt{1 - |x'|^2}) \mid x' \in \mathbb{R}^{n-1},\, |x'| < 1\}, \end{aligned}$$

und die Abbildung $f\colon V \to \mathbb{R}$ mit $V = \{x' \in \mathbb{R}^{n-1} \mid |x'| < 1\}$, $f(x') = \sqrt{1 - |x'|^2}$ ist C^∞.

(iv) $d = n$: Offene Teilmengen im $\mathbb{R}^n$ sind C^∞-Untermannigfaltigkeiten des $\mathbb{R}^n$.

Lemma 11.45. *Ist die Abbildung $f\colon V \to \mathbb{R}^{n-d}$ wie im obigen Lemma, so ist die Abbildung $h\colon V \to \mathbb{R}^n$,*

$$h(y) = (y, f(y))$$

injektiv, die Umkehrabbildung $h^{-1}\colon h(V) \to V$ ist stetig und $Dh(y) \in \mathcal{L}(\mathbb{R}^d, \mathbb{R}^n)$ ist injektiv für alle $y \in V$.

Beweis. h ist offenbar injektiv und die Umkehrabbildung $h^{-1}\colon h(V) \to V$, $(y, f(y)) \mapsto y$ ist stetig als Projektion auf den ersten Eintrag. Weiter erhalten wir

$$Dh(y) = \begin{pmatrix} I_d \\ Df(y) \end{pmatrix},$$

wobei I_d die $d \times d$-Einheitsmatrix ist, so dass $Dh(y)$ für $y \in V$ injektiv ist. $\square$

Definition 11.46. (i) Sei $D \subseteq \mathbb{R}^d$ offen, nichtleer, $d < n$. Eine C^k-Abbildung $\varphi\colon D \to \mathbb{R}^n$ mit $k \in \mathbb{N} \cup \{\infty\}$ heißt eine *C^k-Immersion*, falls $D\varphi(y) \in \mathcal{L}(\mathbb{R}^d, \mathbb{R}^n)$ für alle $y \in D$ injektiv ist.

(ii) Eine C^k-Immersion $\varphi\colon D \to \mathbb{R}^n$ heißt *C^k-Einbettung*, falls φ injektiv und die Umkehrabbildung $\varphi^{-1}\colon \varphi(D) \to D$ stetig ist.

Beispiel 11.47. (i) Für jedes $r > 0$ ist die Abbildung

$$\phi\colon \mathbb{R} \to \mathbb{R}^2, \quad \phi(\theta) = (r\cos\theta, r\sin\theta)$$

eine C^∞-Immersion. Das Bild ist die Kreislinie $\left\{x \in \mathbb{R}^2 \;\middle|\; |x| = r\right\}$ mit Mittelpunkt 0 und Radius r. Die Abbildung $\varphi\big|_{(0,2\pi)}\colon (0, 2\pi) \to \mathbb{R}$ ist eine C^∞-Einbettung.

(ii) Die Abbildung $\psi\colon \mathbb{R} \to \mathbb{R}^2$,

$$\psi(t) = (t^2 - 1, t^3 - t)$$

ist eine C^∞-Immersion. Diese Abbildung ist injektiv auf $(-\infty, 1)$ (da $\psi(t) = \psi(s)$ genau dann gilt, wenn $t = s$ oder $t, s \in \{+1, -1\}$), jedoch ist die Umkehrabbildung $\psi^{-1}\colon \psi\left((-\infty, 1)\right) \to (-\infty, 1)$ nicht stetig an $(0, 0)$. Andernfalls gäbe es ein $\delta > 0$ mit

$$|t + 1| = \left|t - \psi^{-1}(0, 0)\right| \leq \frac{1}{2}$$

für $t < 1$, $|\psi(t) - (0, 0)| \leq \delta$. Aus $\lim\limits_{t\to 1-0} \psi(t) = (0, 0)$ erhalten wir einen Widerspruch.

Lokal in der Nähe eines jeden Punktes ist eine Immersion jedoch eine Einbettung:

Satz 11.48. *Sei $V \subseteq \mathbb{R}^d$ offen, $d < n$, $k \in \mathbb{N} \cup \{\infty\}$ und sei $\varphi\colon V \to \mathbb{R}^n$ eine C^k-Abbildung mit $D\varphi(y^0)\colon \mathbb{R}^d \to \mathbb{R}^n$ injektiv für ein $y^0 \in V$. Dann gibt es eine Umgebung $D \subseteq V$ von y^0 mit:*

(i) *$\varphi(D)$ ist eine d-dimensionale Untermannigfaltigkeit von $\mathbb{R}^n$.*

(ii) *Es gibt eine offene Umgebung $U \subseteq \mathbb{R}^n$ von $x^0 = \varphi(y^0)$ und eine bijektive C^k-Abbildung $\Phi\colon U \to \Phi(U) \subseteq \mathbb{R}^n$, so dass für alle $y \in D$*

$$\Phi \circ \varphi(y) = (y, 0) \in \mathbb{R}^d \times \mathbb{R}^{n-d}.$$

Insbesondere ist $\varphi\big|_D\colon D \to \mathbb{R}^n$ injektiv.

Beweis. (i) $D\varphi(y^0)\colon \mathbb{R}^d \to \mathbb{R}^n$ hat Rang d, insbesondere gibt es in der Matrix

$$\begin{pmatrix} \frac{\partial\varphi_1}{\partial y_1} & \cdots & \frac{\partial\varphi_1}{\partial y_d} \\ \vdots & \ddots & \vdots \\ \frac{\partial\varphi_n}{\partial y_1} & \cdots & \frac{\partial\varphi_n}{\partial y_d} \end{pmatrix} (y^0)$$

d Zeilen, die linear unabhängig sind. O. E. d. A. seien dies die ersten d Zeilen. Wir definieren

$$\begin{aligned} g &= (\varphi_1, \dots, \varphi_d)\colon V \to \mathbb{R}^d, \\ h &= (\varphi_{d+1}, \dots, \varphi_n)\colon V \to \mathbb{R}^{n-d}. \end{aligned}$$

Dann ist $Dg(y^0) \in \mathcal{L}(\mathbb{R}^d, \mathbb{R}^d)$ invertierbar. Folglich gibt es eine offene Umgebung $D \subseteq V$ von y^0 und eine offene Umgebung $W \subseteq \mathbb{R}^d$ von $g(y^0)$, so dass $g\big|_D\colon D \to W$ bijektiv ist. Wir nehmen $w = g(y)$ als neue unabhängige Variable. Dann

$$\varphi(D) = \big\{\big(w, \varphi(w)\big) \;\big|\; w \in W\big\}$$

mit $W \subseteq \mathbb{R}^d$ offen und $f = h \circ g^{-1}\colon W \to \mathbb{R}^{n-d}$ eine C^k-Abbildung.

(ii) Wir definieren

$$\begin{aligned} &\psi\colon D \times \mathbb{R}^{n-d} \to W \times \mathbb{R}^{n-d}, \\ &\psi(y, z) = \phi(y) + (0, z) = (g(y), h(y) + z). \end{aligned}$$

Aus (i) folgt, dass ψ eine C^k-Inverse Φ besitzt, nämlich $\Phi(t, w) = (g^{-1}(w), t - f(w))$. Wir setzen dann $U = W \times \mathbb{R}^{n-d}$ und erhalten

$$\begin{aligned} \Phi \circ \varphi(y) &= \Phi(g(y), h(y)) \\ &= \left(g^{-1}\left(g(y)\right), h(y) - f\left(g(y)\right)\right) = (y, 0) \end{aligned}$$

für $y \in D$.

□

Folgerung. Sei $D \subseteq \mathbb{R}^d$ offen, $d < n$, $k \in \mathbb{N} \cup \{\infty\}$. Sei $\varphi\colon D \to \mathbb{R}^n$ eine C^k-Einbettung. Dann ist $\varphi(D)$ eine d-dimensionale C^k-Untermannigfaltigkeit von $\mathbb{R}^n$.

Wir werden folgendes Resultat benötigen:

Lemma 11.49. *Sei M eine d-dimensionale C^k-Untermannigfaltigkeit des $\mathbb{R}^n$. Seien $D, E \subseteq \mathbb{R}^d$ offen und $\varphi\colon D \to \mathbb{R}^n$ und $\psi\colon E \to \mathbb{R}^n$ C^k-Einbettungen mit $\varphi(D) \subseteq M$, $\psi(E) \subseteq M$ und $\varphi(D) \cap \psi(E) \neq \emptyset$. Dann sind die Mengen $\varphi(D)$, $\psi(E)$ offen in M und die Abbildung*

$$\psi^{-1} \circ \varphi\colon \varphi^{-1}\left(\varphi(D) \cap \psi(E)\right) \to \psi^{-1}\left(\varphi(D) \cap \psi(E)\right)$$

ist eine C^k-Bijektion.

Satz 11.50. *Sei $M \subseteq \mathbb{R}^n$, $M \neq \emptyset$. M ist genau dann eine d-dimensionale C^k-Untermannigfaltigkeit von $\mathbb{R}^n$, wenn es C^k-Einbettungen $\varphi_j\colon D_j \to \mathbb{R}^n$ mit $D_j \subseteq \mathbb{R}^d$ offen, $\varphi_j(D_j) \subseteq M$ gibt, so dass die $U_j = \varphi_j(D_j)$ eine offene Überdeckung von M bilden.*

Beweis. ($\Rightarrow$) Dies folgt direkt aus der Definition einer C^k-Untermannigfaltigkeit und Lemma 11.45.

($\Leftarrow$) Jede der Mengen U_j ist eine d-dimensionale C^k-Untermannigfaltigkeit des $\mathbb{R}^n$. Da diese Mengen M überdecken und da der Begriff einer C^k-Untermannigfaltigkeit durch eine lokale Eigenschaft definiert ist, folgt die Behauptung.

$\square$

Definition 11.51. Die $\left(U_j, \varphi_j^{-1}\right)$ wie im vorigen Satz heißen *Koordinatenumgebungen* auf M, die vermöge φ_j^{-1} von $D_j \subseteq \mathbb{R}^d$ nach M transportierten Euklidischen Koordinaten heißen *lokale Koordinaten* auf M (in U_j).
Die Abbildungen

$$\varphi_{j'}^{-1} \circ \varphi_j \colon \varphi_j^{-1}(U_j \cap U_{j'}) \to \varphi_{j'}^{-1}(U_j \cap U_{j'})$$

zwischen offenen Teilmengen des $\mathbb{R}^d$ (für $U_j \cap U_{j'} \neq \emptyset$) heißen *lokale Koordinatenwechsel* auf M.

Um eine Integrationstheorie auf Untermannigfaltigkeiten aufzubauen, benötigen wir auch:

Satz 11.52. *Sei $\{U_j\}$ eine offene Überdeckung von M durch Koordinatenumgebungen. Dann gibt es eine lokal endliche Teilüberdeckung $\{V_l\}$ von $\{U_j\}$ (d. h. für jedes l gibt es ein j mit $V_l = U_j$, zudem ist für jede kompakte Menge $K \subseteq M$ die Menge $\left\{l \;\middle|\; K \cap V_l \neq \emptyset\right\}$ endlich) und eine zugehörige Zerlegung $\{\chi_l\}$ der Eins, d. h. stetige Funktionen $\chi_l \colon M \to \mathbb{R}$ mit $\operatorname{supp} \chi_l \subseteq V_l$ für alle l und*

$$\sum_l \chi_l(x) = 1$$

für alle $x \in M$.[1]

11.3.2 Integration über Untermannigfaltigkeiten

Sei M eine d-dimensionale C^k-Untermannigfaltigkeit des $\mathbb{R}^n$. Wir möchten jetzt über M integrieren. Wir betrachten dazu zuerst das Problem lokal und denken uns M nahe $x^0 \in M$ vermöge einer C^k-Einbettung $\varphi \colon D \to \mathbb{R}^n$ mit $D \subseteq \mathbb{R}^d$ offen, $\varphi(y^0) = x^0$ für ein $y^0 \in D$ und $\varphi(D) \subseteq M$ parametrisiert. Sei weiterhin $f \colon M \to \mathbb{R}$ mit $\operatorname{supp} f \subset \varphi(D)$. Ist dann $f \circ \varphi$ R-integrierbar über D, so können wir das Integral

$$I_\varphi(f) = \int_D (f \circ \varphi)(y)\, dy$$

betrachten. Ist jedoch $\psi \colon E \to \mathbb{R}^n$ eine weitere C^k-Einbettung wie oben mit $\operatorname{supp} f \subset \varphi(D) \cap \psi(E)$, so erhalten wir

$$I_\psi(f) = \int_E (f \circ \psi)(z)\, dz = \int_D (f \circ \varphi)(y) \left|\det D\left(\psi^{-1} \circ \varphi\right)(y)\right| dy,$$

[1] Für jedes $x \in M$ ist die Summe $\sum_l \chi_l(x)$ endlich.

und letzteres ist im Allgemeinen verschieden von $I_\varphi(f)$ (außer im Spezialfall eines volumenerhaltenden Koordinatenwechsels, wenn $\left|\det D\left(\psi^{-1}\circ\varphi\right)(y)\right| = 1$ für alle $y \in D$). Auf diese Weise erhalten wir keinen invarianten Integralbegriff.

Wir erhalten jedoch einen invarianten Integralbegriff, falls wir definieren:

Definition 11.53. (i) Sei $\varphi\colon D \to \mathbb{R}^n$ mit $D \subseteq \mathbb{R}^d$ offen eine C^k-Einbettung, $f\colon \varphi(D) \to \mathbb{R}$ mit kompaktem Träger. Dann heißt f *R-integrierbar über* $M = \varphi(D)$, falls $f \circ \varphi\colon D \to \mathbb{R}$ R-integrierbar über D ist. In diesem Fall setzen wir

$$\int_M f(x)\,dM(x) = \int_D f\left(\varphi(y)\right)\sqrt{\det\bigl(D\varphi(y)^T D\varphi(y)\bigr)}\,dy.$$

(ii) Ist $M \subseteq \mathbb{R}^n$ eine d-dimensionale C^k-Untermannigfaltigkeit des $\mathbb{R}^n$ und hat $f\colon M \to \mathbb{R}$ kompakten Träger, so heißt f *R-integrierbar über* M, falls es eine lokal endliche Überdeckung $\{U_j\}$ von M mit Koordinatenumgebungen und eine zugehörige Partition $\{\chi_j\}$ der Eins gibt, so dass $\chi_j f$ für alle j R-integrierbar über U_j ist. Wir setzen dann

$$\int_M f(x)\,dM(x) = \sum_i \int_{U_j} \left(\chi_j f\right)(x)\,dU_j(x).$$

(iii) *Absolute R-Integrierbarkeit über* M definieren wir jetzt wie im Fall offener Teilmengen $U \subseteq \mathbb{R}^n$.

Für eine lineare Abbildung $A \in \mathcal{L}(\mathbb{R}^d, \mathbb{R}^n)$ ist $A^T A \in \mathcal{L}(\mathbb{R}^d)$ und $\det\left(A^T A\right) \geq 0$.

Definition 11.54. Wir definieren $\omega\colon \mathcal{L}(\mathbb{R}^d, \mathbb{R}^n) \to \mathbb{R}$ durch

$$\omega(A) = \sqrt{\det\left(A^T A\right)}.$$

Die Funktion ω hat folgende Eigenschaften:

(a) $\omega(AB) = |\det B|\,\omega(A)$ für $B \in \mathcal{L}(\mathbb{R}^d)$,

(b) $\omega(CA) = \omega(A)$ für $C \in \mathrm{O}(\mathbb{R}^n)$.

Beweis. (a) Es gilt

$$\begin{aligned}\omega(AB)^2 &= \det\bigl((AB)^T(AB)\bigr) = \det\left(B^T(A^T A)B\right)\\ &= \det B^T \det(A^T A)\det B = (\det B)^2\omega(A)^2.\end{aligned}$$

(b) Es gilt $C^T C = I$ und

$$\begin{aligned}\omega(CA)^2 &= \det\bigl((CA)^T CA\bigr) = \det\left(A^T(C^T C)A\right)\\ &= \det(A^T A) = \omega(A)^2.\end{aligned}$$

□

Für eine C^k-Einbettung $\varphi\colon D \to \mathbb{R}^n$, $D \subseteq \mathbb{R}^d$ offen, $\varphi(D) \subseteq M$ setzen wir jetzt für $y \in D$

$$\omega_\varphi(y) = \omega\left(D\varphi(y)\right) = \sqrt{\det\bigl(D\varphi(y)^T D\varphi(y)\bigr)}.$$

Lemma 11.55. *Seien $\varphi\colon D \to \mathbb{R}^n$, $\psi\colon E \to \mathbb{R}^n$ zwei C^k-Einbettungen wie oben mit $\varphi(D) \cap \psi(E) \neq \emptyset$. Dann gilt für $y \in \varphi^{-1}\left(\varphi(D) \cap \psi(E)\right)$*

$$\omega_\varphi(y) = \left|\det D\left(\psi^{-1} \circ \varphi\right)(y)\right| \omega_\psi(z),$$

wobei $z = \psi^{-1} \circ \varphi(y)$.

Beweis. Mit der Kettenregel gilt

$$D\varphi(y) = D\psi(z) D\left(\psi^{-1} \circ \varphi\right)(y),$$

und es bleibt obiges Resultat mit $A = D\psi(z)$, $B = D\left(\psi^{-1} \circ \varphi\right)(y)$ anzuwenden. □

Lemma 11.56. *Die Integrale in* Definition 11.53 (ii), (iii) *sind wohldefiniert.*

Beweis. Wir behandeln (ii), (iii) folgt dann wie zuvor.
Sei also $\{V_l\}$ eine weitere lokal endliche Überdeckung von M mit Koordinatenumgebungen und $\{\widetilde{\chi}_l\}$ eine zugehörige Zerlegung der Eins. Dann gilt

$$\begin{aligned}
&\sum_j \int_{D_j} (\chi_j f)\left(\varphi_j(y)\right) \omega_{\varphi_j}(y)\, dy \\
&\qquad = \sum_j \int_{D_j} \sum_l \widetilde{\chi}_l\left(\varphi_j(y)\right) \chi_j\left(\varphi_j(y)\right) f\left(\varphi_j(y)\right) \omega_{\varphi_j}(y)\, dy \\
&\qquad = \sum_{j,l} \int_{E_l} \widetilde{\chi}_l\left(\psi_l(z)\right) \chi_j\left(\psi_l(z)\right) f\left(\psi_l(z)\right) \omega_{\varphi_j}\left({\varphi_j}^{-1} \circ \psi_l(z)\right) \\
&\qquad\qquad \times \left|\det D\left({\varphi_j}^{-1} \circ \psi_l(z)\right)\right| dz \\
&\qquad = \sum_l \int_{E_l} \sum_j \chi_j\left(\psi_l(z)\right) \left(\widetilde{\chi}_l f\right)\left(\psi_l(z)\right) \omega_{\psi_l}(z)\, dz \\
&\qquad = \sum_l \int_{E_l} \left(\widetilde{\chi}_l f\right)(z) \omega_{\psi_l}(z)\, dz.
\end{aligned}$$

□

Bemerkung. Für eine 1-dimensionale C^k-Untermannigfaltigkeit parametrisiert durch $\gamma\colon I \to \mathbb{R}^n$ erhalten wir

$$\int_\gamma f(x)\, |d\sigma| = \int_M f(x)\, dM,$$

wobei $M = \gamma(I)$.

Wir haben die üblichen Begriffe, insbesondere heißt eine Menge $A \subseteq M$ *Jordan-messbar in* M, falls die Funktion $\chi_A \colon M \to \mathbb{R}$ R-integrierbar ist. Wir setzen dann

$$\mathrm{vol}_d(A) = \int_M \chi_A(x)\, dM.$$

Eine Teilmenge $A \subseteq M$ heißt *vernachlässigbar* in M, falls A Jordan-messbar in M ist und $\mathrm{vol}_d(A) = 0$.

Satz 11.57. *Seien $\varphi_j \colon D_j \to \mathbb{R}^n$ für $j = 1, \dots, J$ endlich viele C^k-Einbettungen mit $\varphi_j(D_j) \subseteq M$. Weiter sei:*

(i) $\varphi_j(D_j) \cap \varphi_{j'}(D_{j'}) = \emptyset$ *für* $1 \le j < j' \le J$.

(ii) $M \setminus \bigcup_{j=1}^{J} \varphi_j(D_j)$ *ist vernachlässigbar in* M.

Sei $f \colon M \to \mathbb{R}$ eine lokal beschränkte Funktion (d. h. f ist beschränkt auf kompakten Teilmengen von M). Dann ist f absolut R-integrierbar über M genau dann, wenn $(f \circ \varphi)\, \omega_\varphi$ absolut R-integrierbar über D_j für jedes j ist. In diesem Fall gilt

$$\int_M f(x)\, dM = \sum_j \int_{D_j} f\left(\varphi(y)\right) \omega_\varphi(y)\, dy.$$

11.3.3 Integration einer totalen Ableitung

Sei jetzt $\Omega \subseteq \mathbb{R}^n$ eine offene beschränkte Menge. Wir nehmen an, dass $\partial\Omega$ eine C^1-Hyperfläche im $\mathbb{R}^n$ ist und dass Ω einseitig bezüglich seines Randes liegt. In mathematischen Termen bedeutet dies, dass es für jeden Punkt $x^0 \in \partial\Omega$ eine offene Umgebung $U \subseteq R^n$ von x^0 und eine C^1-Funktion $h \colon U \to \mathbb{R}$ mit $\nabla h \neq 0$ auf U gibt, so dass

$$\Omega \cap U = \left\{x \in U \;\middle|\; h(x) < 0\right\}.$$

Bemerkung. Dann gilt $\partial\Omega \cap U = \left\{x \in U \;\middle|\; h(x) = 0\right\}$, $U \setminus \overline{\Omega} = \left\{x \in U \;\middle|\; h(x) > 0\right\}$.

Definition 11.58. Unter diesen Bedingungen definieren wir die *äußere Einheitsnormale* $n = n(x^0)$ an $\partial\Omega$ in $x^0 \in \partial\Omega$ als

$$n = \frac{\nabla h(x^0)^T}{|\nabla h(x^0)|}.$$

Lemma 11.59. *Die äußere Einheitsnormale n ist durch folgende Bedingungen eindeutig festgelegt:*

(i) $|n| = 1$.

(ii) $x^0 + sn \notin \overline{\Omega}$ *für kleine* $s > 0$.

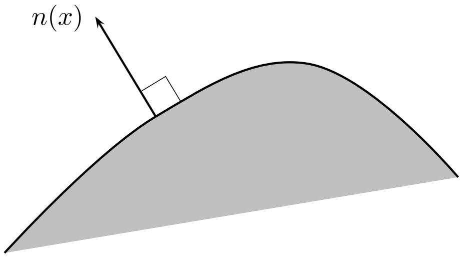

Abbildung 11.6: Die äußere Normale

(iii) *Für jede C^1-Kurve $\gamma\colon (-1,1) \to \mathbb{R}^n$ mit $\gamma(0) = x^0$ und $\gamma((-1,1)) \subseteq \partial\Omega \cap U$ gilt*

$$\langle n, \gamma'(0)\rangle = 0.$$

Beweis. Zu (iii): Es gilt $h(\gamma(t)) = 0$, also $\nabla h(\gamma(t))\,\gamma'(t) = 0$ für alle t. An der Stelle $t = 0$ folgt $\langle n, \gamma'(0)\rangle = 0$. □

Satz 11.60. *Sei $\Omega \subset \mathbb{R}^n$ eine offen, beschränkte Menge mit C^1-Rand $\partial\Omega$, so dass Ω einseitig bezüglich $\partial\Omega$ liegt. Sei $f\colon \Omega \to \mathbb{R}$ eine C^1-Funktion, so dass sich f und $\nabla^T f\colon \Omega \to \mathbb{R}^n$ zu stetigen Funktionen auf $\overline{\Omega}$ fortsetzen lassen (das ist gleichbedeutend damit, dass f und $\nabla^T f$ sind gleichmäßig stetig auf Ω). Dann gilt*

$$\int_\Omega \left(\nabla^T f\right)(x)\,dx = \int_{\partial\Omega} f(x)n(x)\,d(\partial\Omega),$$

wobei $n\colon \partial\Omega \to \mathbb{R}$ die äußere Einheitsnormale bezeichnet. In Komponenten

$$\int_\Omega \frac{\partial f}{\partial x_j}(x)\,dx = \int_{\partial\Omega} f(x)n_j(x)\,d(\partial\Omega)$$

für $j = 1, \dots, n$.

Bemerkung. Der Hauptsatz der Differential– und Integralrechnung, wonach

$$\int_a^b f'(x)\,dx = f(b) - f(a)$$

für eine stetig differenzierbare Funktion $f\colon [a,b] \to \mathbb{R}$ gilt, ist ein Spezialfall.

Beweis. **1. Schritt.** Wir zeigen, dass es für jede Stelle $x \in \overline{\Omega}$ eine Umgebung $U_x \subseteq \mathbb{R}^n$ von x so gibt, dass aus $\operatorname{supp} f \subseteq U_x$ die Gültigkeit der Ausssage des Satzes für f folgt.

Fall 1. $x^0 \in \Omega$. Wir wählen ein achsenparalleles Rechteck $U_{x^0} = \prod_{j=1}^{n}(a_j, b_j)$ in Ω, das x^0 enthält. Dann gilt mit $U' = \prod_{j=1}^{n-1}(a_j, b_j)$ für $\operatorname{supp} f \subset U_{x^0}$

$$\begin{aligned}\int_\Omega \frac{\partial f}{\partial x_n}(x)\,dx &= \int_{U'} \left(\int_{a_n}^{b_n} \frac{\partial f}{\partial x_n}(x', x_n)\,dx_n \right) dx' \\ &= \int_{U'} \left(f(x', b_n) - f(x', a_n) \right) dx' = 0.\end{aligned}$$

Analog erhält man $\int_\Omega \frac{\partial f}{\partial x_j}(x)\,dx = 0$ für $j = 1, \dots, n-1$ und damit die Behauptung.

Fall 2. $x^0 \in \partial\Omega$. Wir zeigen, dass

$$\int_\Omega \nabla f(x) v\,dx = \int_{\partial\Omega} f(x) u(x) \cdot v\,d(\partial\Omega)$$

für jeden Vektor $v \in \mathbb{R}^n$ gilt. Da beide Seiten linear in v sind und sich jeder Vektor im $\mathbb{R}^n$ aus Vektoren mit $u(x^0) \cdot v > 0$ linear kombinieren lässt, genügt es, dies für alle Vektoren v mit $u(x^0) \cdot v > 0$ zu zeigen.
Sei $v \in \mathbb{R}^n$ ein derartiger Vektor. Wir koordinatisieren $\partial\Omega$ in der Nähe von x^0 vermöge einer C^1-Einbettung $\varphi\colon D \to \mathbb{R}^n$ mit $D \subseteq \mathbb{R}^{n-1}$ offen, $0 \in D$, $\varphi(0) = x^0$ und $\varphi(D) \subseteq \partial\Omega$. Wir betrachten dann

$$\Phi\colon D \times (-\epsilon, \epsilon) \to \mathbb{R}^n, \quad (y', y_n) \mapsto \varphi(y') + y_n v.$$

Wegen $\det \Phi(0,0) = \det\left(D\varphi(0), v\right) \neq 0$ ist Φ für D und $\epsilon > 0$ hinreichend klein ein C^1-Koordinatenwechsel. Es gilt

$$\left(\nabla f\right)\left(\Phi(y)\right) v = \left(\nabla f\right)\left(\varphi(y') + y_n v\right) v = \frac{\partial}{\partial y_n}\left(f(\varphi(y') + y_n v)\right).$$

Wir benötigen folgendes Resultat aus der linearen Algebra:

Lemma 11.61. *$A \in \mathcal{L}(\mathbb{R}^{n-1}, \mathbb{R}^n)$ habe Rang $n-1$ und sei $\ker A^T = \left\{\lambda n \mid \lambda \in \mathbb{R}\right\}$ für $n \in \mathbb{R}^n$, $|n| = 1$ (beachte, dass $\dim\ker A^T = 1$). Dann gilt für ein beliebiges $v \in \mathbb{R}^n$*

$$\left|\det\begin{pmatrix} A & v\end{pmatrix}\right| = |n \cdot v|\ \omega(A),$$

wobei $\omega(A)$ in Definition 11.54 *eingeführt wurde.*

Beweis. O. B. d. A. dürfen wir $n = (0, \dots, 0, 1)^T$ annehmen (weshalb?). Dann $A = \binom{B}{0}$, wobei $B \in \mathcal{L}(\mathbb{R}^{n-1})$ invertierbar ist und die 0 in der letzten Zeile für $(0, \dots, 0)$ steht, sowie $\omega(A) = |\det B|$. Wir erhalten mit $v = \binom{v'}{v_n}$, dass

$$\left|\det\begin{pmatrix} A & v\end{pmatrix}\right| = \left|\det\begin{pmatrix} B & v' \\ 0 & v_n\end{pmatrix}\right| = |\det B|\,|v_n| = |n \cdot v|\ \omega(A).$$

□

Indem wir dieses Lemma mit $A = D\varphi(y')$ und $n = n\left(\varphi(y')\right)$ anwenden, erhalten wir

$$|\det \Phi(y',0)| = \left|\det \begin{pmatrix} D\varphi(y') & v \end{pmatrix}\right| = n(\varphi(y')) \cdot v \, \omega_\phi(y').$$

Es ergibt sich für $\operatorname{supp} f \subseteq U_{x^0} = \Phi\left(D \times (-\epsilon,\epsilon)\right)$

$$\begin{aligned}
\int_\Omega \nabla f(x) v \, dx &= \int_D \int_{-\epsilon}^0 \frac{\partial}{\partial y_n} \left(f\left(\varphi(y') + y_n v\right)\right) dy_n \left|\det \begin{pmatrix} D\varphi(y') & v \end{pmatrix}\right| dy' \\
&= \int_D f\left(\varphi(y')\right) n(\varphi(y')) \cdot v \, \omega_\phi(y') \, dy' \\
&= \int_{\partial\Omega} f(x) n(x) \cdot v \, d(\partial\Omega).
\end{aligned}$$

2. Schritt. Sei nun $f\colon \Omega \to \mathbb{R}$ wie in der Formulierung des Satzes, ansonsten beliebig.

$\{U_x\}_{x\in\bar\Omega}$ ist eine offene Überdeckung von $\bar\Omega$. Wir wählen eine endliche Teilüberdeckung aus, also $\bar\Omega \subseteq \bigcup\limits_{k=1}^K U_{x_k}$. Sei $\{\chi_k\}_{k=1}^K$ eine zugehörige C^1-Zerlegung der Eins [2], d. h. die $\chi_k\colon \mathbb{R}^n \to \mathbb{R}$ sind C^1-Funktionen mit $\operatorname{supp} \chi_k \subset U_{x_k}$ und $\sum\limits_{k=1}^K \chi_k(x) = 1$ für alle $x \in \overline{\Omega}$. Dann gilt mit dem 1. Schritt

$$\begin{aligned}
\int_\Omega \nabla^T f(x) \, dx &= \sum_{k=1}^K \int_\Omega \nabla^T \left(\chi_k f\right)(x) \, dx \\
&= \sum_{k=1}^K \int_{\partial\Omega} \left(\chi_k f\right)(x) u(x) \, d(\partial\Omega) \\
&= \int_{\partial\Omega} f(x) u(x) \, d(\partial\Omega).
\end{aligned}$$

□

11.3.4 Eine Verallgemeinerung

In Anwendungen muss man häufig über Gebiete integrieren, deren Ränder nicht C^1 sind, aber „fast so gut“ (z. B. wenn Ω ein polygonales Gebiet ist). Um derartige Situationen behandeln zu können, schwächen wir die Voraussetzungen im vorigen Satz ab.

Definition 11.62. Eine kompakte Menge $S \subset \mathbb{R}^n$ heißt *vernachlässigbar im $(n-1)$-dimensionalen Sinne*, falls

$$\overline{\int_{\mathbb{R}^n}} \chi_{S_\epsilon}(x) \, dx = o(\epsilon) \text{ für } \epsilon \to +0,$$

[2] Der Beweis von Satz 11.31 läßt sich leicht dahingehend abändern, dass er eine C^1-Zerlegung der Eins liefert.

d. h. falls

$$\lim_{\epsilon \to +0} \frac{1}{\epsilon} \overline{\int\limits_{\mathbb{R}^n}} \chi_{S_\epsilon}(x)\, dx = 0$$

gilt. Hierbei ist $S_\epsilon = \left\{x \in \mathbb{R}^n \;\middle|\; |x-y| \le \epsilon \text{ für ein } y \in S\right\}$.

Beispiel 11.63. Ist S endliche Vereinigung von höchstens $(n-2)$-dimensionalen Untermannigfaltigkeiten des $\mathbb{R}^n$, so ist S vernachlässigbar im $(n-1)$-dimensionalen Sinne.

Wir machen jetzt folgende Voraussetzungen an Ω:

(i) $\Omega \subset \mathbb{R}^n$ ist offen, beschränkt.

(ii) Es gibt endlich viele $(n-1)$-dimensionale Untermannigfaltigkeiten $M_1, \dots, M_K$ des $\mathbb{R}^n$ mit $M_k \subseteq \partial\Omega$ für alle k und $M_k \cap M_{k'} = \emptyset$ für $k \neq k'$, so dass $\partial\Omega \setminus \bigcup\limits_{k=1}^{K} M_k$ vernachlässigbar im $(n-1)$-dimensionalen Sinne ist.

(iii) Ω liegt einseitig bezüglich jedes M_k.

Theorem 11.64. *Unter diesen Voraussetzungen sei $f \colon \Omega \to \mathbb{R}$ eine gleichmäßig stetige C^1-Funktion, so dass $\nabla^T f \colon \Omega \to \mathbb{R}^n$ absolut R-integrierbar ist. Dann gilt*

$$\int_\Omega \nabla^T f(x)\, dx = \int_{\partial\Omega} f(x) n(x) d(\partial\Omega),$$

wobei die rechte Seite als $\sum\limits_{k=1}^{K} \int_{M_k} f(x) u(x)\, dM_k$ *interpretiert wird.*

11.3.5 Einige Anwendungen

In den folgenden Beispielen sind Ω und alle auftretenden Funktionen hinreichend gut, damit die vorigen beiden Sätze anwendbar sind.

Beispiel 11.65. 1. (Partielle Integration). Für $f, g \colon \Omega \to \mathbb{R}$ gilt

$$\int_\Omega f(x) \frac{\partial g}{\partial x_j}(x)\, dx = \int_{\partial\Omega} f(x) g(x) n_j(x)\, d(\partial\Omega) - \int_\Omega \frac{\partial f}{\partial x_j}(x) g(x)\, dx$$

für $j = 1, \dots, n$.

Beweis. Es gilt

$$\begin{aligned} \int_\Omega f \frac{\partial g}{\partial x_j}\, dx &= \int_\Omega \frac{\partial}{\partial x_j}\,(f\, g)\; dx - \int_\Omega \frac{\partial f}{\partial x_j} g\, dx \\ &= \int_{\partial\Omega} f\, g\, n_j\, d(\partial\Omega) - \int_\Omega \frac{\partial f}{\partial x_j} g\, dx. \end{aligned}$$

□

2. (Gaußscher Satz). Sei $F\colon \Omega \to \mathbb{R}^n$ ein Vektorfeld. Dann gilt

$$\int_\Omega \operatorname{div} F(x)\, dx = \int_{\partial\Omega} F(x) \cdot n(x)\, d(\partial\Omega).$$

Beweis. Wegen $\operatorname{div} F = \sum\limits_{j=1}^{n} \frac{\partial F_j}{\partial x_j}$ gilt

$$\begin{aligned}\int_\Omega \operatorname{div} F\, dx &= \sum_{j=1}^{n} \int_\Omega \frac{\partial F_j}{\partial x_j}\, dx = \sum_{j=1}^{n} \int_{\partial\Omega} F_j n_j\, d(\partial\Omega)\\ &= \int_{\partial\Omega} F \cdot n\, d(\partial\Omega).\end{aligned}$$

□

3. (1. Greensche Formel). Für $f, g\colon \Omega \to \mathbb{R}$ gilt

$$\begin{aligned}\int_\Omega f(x)\left(\Delta g\right)(x)\, dx = &\int_{\partial\Omega} f(x) \frac{\partial g}{\partial n}(x)\, d(\partial\Omega)\\ &- \int_\Omega \left(\nabla^T f\right)(x) \cdot \left(\nabla^T g\right)(x)\, dx.\end{aligned}$$

Hierbei ist $\frac{\partial g}{\partial n}(x) = \nabla g(x) n$ die sogenannte *Normalenableitung*.

Beweis. Wegen $\Delta = \operatorname{div} \cdot \nabla^T$ gilt

$$\begin{aligned}\int_\Omega f \Delta g\, dx &= \int_\Omega \operatorname{div}\left(f \nabla^T g\right)\, dx - \int_\Omega \left(\nabla^T f\right) \cdot \left(\nabla^T g\right)\, dx\\ &= \int_{\partial\Omega} f \nabla^T g \cdot n\, d(\partial\Omega) - \int_\Omega \left(\nabla^T f\right) \cdot \left(\nabla^T g\right)\, dx.\end{aligned}$$

□

4. (2. Greensche Formel). Für $f, g\colon \Omega \to \mathbb{R}$ gilt

$$\begin{aligned}&\int_\Omega \left(f(x)\left(\Delta g\right)(x) - \left(\Delta f\right)(x) g(x)\right)\, dx\\ &\quad = \int_{\partial\Omega} \left(f(x) \frac{\partial g}{\partial n}(x) - \frac{\partial f}{\partial n}(x) g(x) \right)\, d(\partial\Omega).\end{aligned}$$

Beweis. Man vertausche die Rollen von f, g in 3. und subtrahiere die so erhaltenen Gleichungen. □

Kapitel 12

Gewöhnliche Differenzialgleichungen

12.1 Grundlegende Begriffe

Definition 12.1. (i) Ein *System von M gewöhnlichen Differenzialgleichungen in N unbestimmten Funktionen* ist eine Gleichung der Form

$$F\big(x, u(x), u'(x), u''(x), \ldots, u^{(m)}(x)\big) = f(x), \quad x \in I. \tag{12.2}$$

Hierbei ist $I \subseteq \mathbb{R}$ ein Intervall, $u = \big(u_1, \ldots, u_N\big)^T$ ist die gesuchte Funktion,

$$u' = \frac{du}{dx}, \;\; u'' = \frac{d^2u}{dx^2}, \;\; \ldots, \;\; u^{(m)} = \frac{d^m u}{dx^m}$$

und die stetigen Funktionen

$$F \colon U \subseteq I \times \underbrace{\mathbb{R}^N \times \cdots \times \mathbb{R}^N}_{m+1 \text{ Faktoren}} \to \mathbb{R}^M$$

und $f \colon I \to \mathbb{R}^M$ sind gegeben. m heißt die *Ordnung* des Systems.

(ii) Das System (12.2) heißt *linear*, falls F eine lineare Funktion in $u(x), u'(x), \ldots, u^{(m)}$ ist, also

$$F\big(x, u, u', \ldots, u^{(m)}\big) = a_0(x)u^{(m)} + a_1(x)u^{(m-1)} + \cdots + a_{m-1}(x)u' + a_m(x)u.$$

Hierbei sind die $a_j(x)$ für $j = 0, \ldots, m$ $M \times N$-Matrizen.
Hat F die Form

$$F\big(x, u, u', \ldots, u^{(m)}\big) = a\big(x, u, u', \ldots, u^{(m-1)}\big)\, u^{(m)} + b\big(x, u, u', \ldots, u^{(m-1)}\big),$$

so heißt das System (12.2) *quasilinear* (d. h. das System ist linear in der höchsten Ableitung).
Das allgemeine System (12.2) heißt *voll nichtlinear*.

(ii) Eine *klassische Lösung* des Systems (12.2) ist eine Funktion $u\colon I \to \mathbb{R}^N$ der Klasse C^m mit $\big(x, u(x), u'(x), \ldots, u^{(m)}(x)\big)^T \in U$ für alle $x \in I$ und Gleichung (12.2) ist punktweise erfüllt.

Beispiele für gewöhnliche Differenzialgleichungen

1. *Besselsche Differenzialgleichung*: $u'' + \frac{1}{x}u' + \left(1 - \frac{v^2}{x^2}\right) u = 0$.
2. *Airysche Differenzialgleichung*: $u'' = xu$.
3. *Bernoullische Differenzialgleichung*: $u' = a(x)u + b(x)u^p$, $p \in \mathbb{N}$.
4. *Riccattische Differenzialgleichung*: $u' = a(x)u^2 + b(x)u + c(x)$.

12.2 Ein lokaler Existenzsatz

12.2.1 Funktionalanalytische Vorbereitungen

Definition 12.3. Sei V ein $\mathbb{R}$-(oder $\mathbb{C}$-)Vektorraum. Eine Funktion $\|\cdot\| : V \to [0, \infty)$ heißt eine *Norm auf* V, falls Folgendes für alle u, $v \in V$, $\alpha \in \mathbb{R}$ (oder $\alpha \in \mathbb{C}$) gilt:

(i) $\|u\| = 0$ impliziert $u = 0$,

(ii) $\|\alpha u\| = |\alpha|\, \|u\|$

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

Ein mit einer Norm versehener Vektorraum $(V, \|\cdot\|)$ heißt ein *normierter Raum*.

Beispiel 12.4. $\left(\mathbb{R}^n, |\cdot|_p\right)$ ist ein normierter Raum für $1 \leq p \leq \infty$, wobei

$$|x|_p = \begin{cases} \left(\sum\limits_{j=1}^{n} |x_j|^p\right)^{1/p}, & 1 \leq p < \infty, \\ \max\limits_{j=1,\ldots,n} |x_j|\,, & p = \infty. \end{cases}$$

Lemma 12.5. (i) *Ist* $(V, \|\cdot\|)$ *ein normierter Raum, so ist* $d\colon V \times V \to [0, \infty)$,

$$d(u, v) = \|u - v\|$$

eine Metrik auf V.

(ii) *Die Vervollständigung eines normierten Raumes als metrischer Raum ist in kanonischer Weise ein normierter Raum.*

Beweis. (i) Übungsaufgabe.

(ii) Ist $\{u_m\}$ eine Cauchyfolge in V, so existiert $\lim\limits_{m\to\infty} \|u_m\|$. Ist $\{u'_m\}$ eine weitere Cauchyfolge mit $\lim\limits_{m\to\infty} \|u_m - u'_m\| = 0$, so gilt $\lim\limits_{m\to\infty} \|u_m\| = \lim\limits_{m\to\infty} \|u'_m\|$. Auf diese Weise lässt sich eine Norm auf der Vervollständigung von V als
$$\|[\{u_m\}]\| = \lim_{m\to\infty} \|u_m\|$$
einführen, wobei $[\{u_m\}]$ die Äquivalenzklasse von $\{u_m\}$ bezeichnet.

□

Definition 12.6. Ein normierter Raum $(V, \|\cdot\|)$, der als metrischer Raum vollständig ist, heißt ein *Banachraum.*

Definition 12.7. Sei $I \subset \mathbb{R}$ ein kompaktes Intervall, $m \in \mathbb{N}_0$, $n \in \mathbb{N}$. Wir definieren dann $C^m(I;\mathbb{R}^N)$ als den Raum aller C^m-Funktionen $u\colon I \to \mathbb{R}^N$. Anstelle von $C^0(I;\mathbb{R}^N)$ schreiben wir auch $C(I;\mathbb{R}^N)$.

Offenbar ist $C^m(I;\mathbb{R}^N)$ ein $\mathbb{R}$-Vektorraum und eine $C^m(I;\mathbb{R})$-Modul.

Lemma 12.8. *Der Raum $C^m(I;\mathbb{R}^N)$ versehen mit der Norm*
$$\|u\|_{C^m(I;\mathbb{R}^N)} = \sum_{j=0}^{m} \max_{x\in I} \left|u^{(j)}(x)\right|$$
ist ein Banachraum.

Beweis. $u \mapsto \|u\|_{C^m(I;\mathbb{R}^N)}$ ist offenbar eine Norm auf $C^m(I;\mathbb{R}^N)$.
Wir führen den Vollständigkeitsbeweis zuerst für $m = 0$. Sei also $\{u_k\} \subset C(I;\mathbb{R}^N)$ eine Cauchyfolge. Dann ist $\{u_k(x)\}$ für jedes $x \in I$ eine Cauchyfolge in $\mathbb{R}^N$, also existiert der (punktweise) Grenzwert
$$u(x) = \lim_{k\to\infty} u_k(x), \quad x \in I.$$
Als gleichmäßiger Limes stetiger Funktionen ist die Grenzfunktion u stetig, gehört also zu $C(I;\mathbb{R}^N)$. Damit $u_k \to u$ für $k \to \infty$ in $C(I;\mathbb{R}^N)$.
Für beliebiges m haben wir zuerst $u_k^{(j)} \to v_j$ für $k \to \infty$ in $C(I;\mathbb{R}^N)$ und jedes $0 \le j \le m$. Aus
$$u_k^{(j)}(x) = u_k^{(j)}(x^0) + \int_{x^0}^{x} u_k^{(j+1)}(x')\,dx', \quad x \in I,$$
für ein fixiertes $x^0 \in I$ und $0 \le j < m$ folgt durch Grenzübergang $k \to \infty$
$$v_j(x) = v_j(x^0) + \int_{x^0}^{x} v_{j+1}(x')\,dx', \quad x \in I,$$
also $v_{j+1} = v'_j$. Es ergibt sich $v_j = u^{(j)}$, wobei wir $u = v_0$ gesetzt haben, und $u_k \to u$ für $k \to \infty$ in $C^m(I;\mathbb{R}^N)$. □

12.2.2 Formulierung des Satzes und Beweis

Wir betrachten zuerst ein *Anfangswertproblem* für ein System erster Ordnung der Form

$$\begin{cases} u' = F(x,u), \quad x \in I, \\ u(x^0) = u^0. \end{cases} \tag{12.9}$$

Hierbei ist $u\colon I \to \mathbb{R}^N$ die gesuchte Funktion. Die Funktion F (aus einer Teilmenge) von $I \times \mathbb{R}^N$ nach $\mathbb{R}^N$, $x^0 \in I$ und $u^0 \in \mathbb{R}^N$ sind gegeben.
Annahme 1. Es existieren $a > 0$, $b > 0$, so dass F eine stetige Funktion auf

$$R = \left\{(x,u) \;\middle|\; \left|x - x^0\right| \le a,\ \left|u - u^0\right| \le b\right\}$$

ist. Zudem sei

$$\max_{(x,u)\in R} |F(x,u)| \le M.$$

Lemma 12.10. *Für jede C^1-Lösung $u = u(x)$ obigen Anfangswertproblems auf $\left|x - x^0\right| \le \min\{a, b/M\}$ gilt*

$$\left|u(x) - u^0\right| \le b.$$

Beweis. Sei $\mathcal{F}$ die Menge aller X mit $0 \le X \le \min\{a, b/M\}$, so dass die Abschätzung $\left|u(x) - u^0\right| \le b$ für $\left|x - x^0\right| \le X$ gilt.
Offenbar ist $\mathcal{F}$ nichtleer (da $0 \in \mathcal{F}$) und abgeschlossen. Diese Menge ist aber auch offen in $[0, \min\{a, b/M\}]$, denn ist $X \in \mathcal{F}$ und gilt $0 \le X < \min\{a, b/M\}$, so gilt

$$\left|u(x) - u^0\right| = \left|\int_{x^0}^{x} F\left(x', u(x')\right)\, dx'\right| \le M \left|x - x^0\right| < b$$

für $\left|x - x^0\right| \le X$ und damit ist $X < \sup \mathcal{F}$ (die Funktion $x \mapsto u(x)$ ist stetig). Folglich gilt $\mathcal{F} = [0, \min\{a, b/M\}]$. □

Annahme 2. Für (x,u), $(x,v) \in R$ gilt

$$|F(x,u) - F(x,v)| \le L\, |u - v|$$

für eine Konstante $L > 0$.

Satz 12.11. *Unter diesen Voraussetzungen besitzt das Anfangswertproblem* (12.9) *eine eindeutige C^1-Lösung $u = u(x)$ für $\left|x - x^0\right| \le \min\{a, b/M\}$.*

Beweis. **Existenz.** Wir konstruieren die Lösung u durch sukzessive Approximation. Dazu setzen wir $u_0(x) = u^0$ für $x \in I$, wobei $I = [x^0 - X, x^0 + X]$, $X = \min\{a, b/M\}$, und dann

$$\begin{cases} {u'}_m(x) = F\left(x, u_{m-1}(x)\right), \\ u_m(x^0) = u^0 \end{cases}$$

für $m = 1, 2, 3, \ldots$ Wir erhalten

$$u_m(x) = u^0 + \int_{x^0}^{x} F\left(x', u_{m-1}(x')\right) dx', \quad x \in I.$$

Behauptung. Für alle $m \in \mathbb{N}$ gilt

$$|u_m(x) - u_{m-1}(x)| \leq \frac{ML^{m-1}|x - x^0|^m}{m!} \leq \frac{ML^{-1}}{m!}(LX)^m.$$

Beweis (durch Induktion über m). Für $m = 1$ gilt die Abschätzung wegen

$$|u_1(x) - u_0(x)| \leq \left|u_1(x) - u^0\right| \leq M \left|x - x^0\right|.$$

Der Induktionsschritt $m \to m + 1$ ergibt sich aus

$$\begin{aligned}
|u_{m+1}(x) - u_m(x)| &= \left| \int_{x^0}^{x} \left(F(x', u_m(x')) - F(x', u_{m-1}(x'))\right) dx' \right| \\
&\leq L \int_{x^0}^{x} |u_m(x') - u_{m-1}(x')| \, |dx'| \\
&\leq \frac{ML^m}{m!} \int_{x^0}^{x} \left|x' - x^0\right|^m |dx'| = \frac{ML^m}{(m+1)!} \left|x - x^0\right|^{m+1}.
\end{aligned}$$

□

Aus $\sum_{m=0}^{\infty} \frac{(LX)^m}{m!} < \infty$ folgt jetzt, dass die Folge $\{u_m\}$ gleichmäßig auf I gegen eine Grenzfunktion u konvergiert. Weiterhin ergibt sich $\left|u(x) - u^0\right| \leq b$ für $x \in I$ sowie

$$u(x) = u^0 + \int_{x^0}^{x} F(x', u(x'))\, dx', \quad x \in I.$$

Damit gehört u zu $C^1(I; \mathbb{R}^N)$ und löst das Anfangswertproblem.

Eindeutigkeit. Ist $v = v(x)$ eine weitere Lösung von (12.9), so gilt

$$u(x) - v(x) = \int_{x^0}^{x} \left(F(x', u(x')) - F(x', v(x'))\right) dx',$$

also wegen $|u(x) - v(x)| \leq 2M \left|x - x^0\right|$ wie im Existenzbeweis

$$|u(x) - v(x)| \leq \frac{2ML^{-1}}{m!}(LX)^m$$

für alle m. Für $m \to \infty$ folgt $|u(x) - v(x)| = 0$ für alle $x \in I$, also $u = v$. □

Ist

$$\begin{cases} u^{(m)} = F(x, u, u', \ldots, u^{(m-1)}), \quad x \in I, \\ u^{(j)}(x^0) = u^j, \quad 0 \leq j \leq m-1, \end{cases} \tag{12.12}$$

ein *System höherer Ordnung*, so lässt sich dieses System wie folgt auf ein System erster Ordnung reduzieren:
Wir setzen $U = (u, u', \dots, u^{(m-1)})^T$, $U^0 = (u^0, u^1, \dots, u^{m-1})^T$. Dann

$$\begin{cases} U' = G(x, U), \\ U(x^0) = U^0, \end{cases} \tag{12.13}$$

wobei

$$G(x, U) = \begin{pmatrix} u_1 \\ \vdots \\ u_{m-1} \\ F(x, u_0, \dots, u_{m-1}) \end{pmatrix}$$

mit $U = (u_0, \dots, u_{m-1})^T$.

Lemma 12.14. *Die Systeme* (12.12), (12.13) *sind äquivalent in dem Sinne, dass jede C^1-Lösung $u = u(x)$ von* (12.12) *eine C^1-Lösung $U = U(x)$ von* (12.13) *mittels der angegebenen Transformation ergibt. Ist umgekehrt $U = U(x)$ eine C^1-Lösung von* (12.13)*, so ist die erste Komponente von U eine C^1-Lösung von* (12.12).

Beweis. Selbst. □

12.2.3 Stetige Abhängigkeit vom Anfangswert

Unter den Voraussetzungen wie oben betrachten wir jetzt zusätzlich das Anfangswertproblem

$$\begin{cases} v' = F(x, v), \quad x \in I, \\ v(x^0) = v^0, \end{cases}$$

wobei v^0 hinreichend nahe an u^0 ist. Die (lokal eindeutige) Lösung dieses Problems sei $v = v(x)$.

Satz 12.15. *Für* $\left|x - x^0\right| + \left|u^0 - v^0\right| / M \le \min\{a, b/M\}$ *gilt*

$$|u(x) - v(x)| \le \mathrm{e}^{L|x - x^0|} \left|u^0 - v^0\right|.$$

Beweis. Aus dem Existenz- und Eindeutigkeitssatz folgt, dass $v = v(x)$ auf dem angegebenen Intervall existiert.
Sei $h(x) = |u(x) - v(x)|$. Dann ist h eine C^1-Funktion (es gilt $h(x) > 0$ für alle x, wenn $u^0 \ne v^0$), und wir erhalten

$$\begin{aligned} \frac{1}{2} \frac{d}{dx} \left(h^2(x)\right) &= \frac{1}{2} \frac{d}{dx} \left(|u(x) - v(x)|^2\right) \\ &= \langle u'(x) - v'(x), u(x) - v(x) \rangle \\ &= \langle F(x, u(x)) - F(x, v(x)), u(x) - v(x) \rangle \\ &\le L |u(x) - v(x)|^2 = L\, h^2(x). \end{aligned}$$

Es folgt $h'(x)h(x) \le L\, h^2(x)$, $h'(x) \le L\, h(x)$ und schließlich

$$h(x) \le h(x^0)\, \mathrm{e}^{L\left|x-x^0\right|}.$$

Letzteres ist die Behauptung. □

12.3 Lineare Differenzialgleichungen

Bei der Behandlung linearer Gleichungen ist es bequem, aller auftretenen Funktionen als komplexwertig anzunehmen, was wir ab sofort tun. Die Ergebnisse des vorigen Abschnittes übertragen sich wegen $\mathbb{C} \cong \mathbb{R}^2$ unmittelbar.

12.3.1 Allgemeines

Wir behandeln lineare Differenzialgleichungen der Form

$$\begin{cases} u^{(m)} + a_1(x)u^{(m-1)} + \cdots + a_{m-1}(x)u' + a_m(x)u = f(x), \quad x \in I, \\ u(x^0) = u^0, \quad u'(x^0) = u^1, \quad \ldots, \quad u^{(m-1)}(x^0) = u^{m-1}. \end{cases} \tag{12.16}$$

Hierbei ist $I \subseteq \mathbb{R}$ ein Intervall, $x^0 \in I$ und die rechte Seite $f \in C(I;\mathbb{C})$ bzw. die Anfangsdaten $u^0, u^1, \ldots, u^{m-1} \in \mathbb{C}$ sind gegeben. Die $a_j(x)$ sind stetige komplexwertige Funktionen auf I.

Lemma 12.17. (i) *Unter den genannten Bedingungen existiert eine eindeutige C^m-Lösung u global auf I.*

(ii) *Der Raum der Lösungen u zur homogenen Gleichung* (*d. h. wenn $f \equiv 0$*) *ist ein m-dimensionaler $\mathbb{C}$-Vektorraum. Die Anfangsdaten $u^0, u^1, \ldots, u^{m-1}$ sind lineare Koordinaten in diesen Lösungsraum.*

(iii) *Der Raum der Lösungen u zur inhomogenen Gleichung* (*d. h. mit einer beliebigen rechter Seite f*) *ist ein affiner Raum über $\mathbb{C}$ mit zugehörigem $\mathbb{C}$-Vektorraum wie unter* (ii) *beschrieben.*

Beweis. Dies folgt aus dem analogen Resultat für lineare Systeme erster Ordnung (siehe unten). □

Anstelle (12.16) betrachten wir auch $N \times N$-Systeme erster Ordnung der Form

$$\begin{cases} U' = A(x)U + F(x), \quad x \in I, \\ U(x^0) = U^0, \end{cases} \tag{12.18}$$

wobei $A \in C(I; M(N,\mathbb{C}))$, $F \in C(I;\mathbb{C}^N)$ und $U^0 \in \mathbb{C}^N$. Die Funktion $U \in C^1(I;\mathbb{C}^N)$ ist gesucht.

Die Gleichung (12.16) kann in ein System der Form (12.18) (mit $N = m$) transformiert werden, wobei

$$A = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & 1 \\ -a_m & -a_{m-1} & -a_{m-2} & \cdots & -a_2 & -a_1 \end{pmatrix}, \quad F = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ f \end{pmatrix}$$

und $U^0 = \begin{pmatrix} u^0 \\ u^1 \\ \vdots \\ u^{m-1} \end{pmatrix}$ ist.

Lemma 12.19. (i) *Unter den genannten Bedingungen besitzt das System* (12.18) *eine eindeutige Lösung $U \in C^1(I; \mathbb{C}^N)$.*

(ii) *Der Raum der Lösungen U zur homogenen Gleichung* (*d. h. wenn $F \equiv 0$*) *ist ein N-dimensionaler $\mathbb{C}$-Vektorraum. Das Anfangsdatum $U^0 \in \mathbb{C}^N$ parametrisiert diesen Lösungsraum linear.*

(iii) *Der Raum der Lösungen U zur inhomogenen Gleichung* (*d. h. mit einem beliebigen F*) *ist ein affiner Raum über $\mathbb{C}$ mit zugehörigem $\mathbb{C}$-Vektorraum wie unter* (ii) *beschrieben.*

Beweis. Lediglich (i) bedarf eines Arguments. Wir haben zu zeigen, dass für jedes kompakte Intervall $J \subseteq I$ mit $x^0 \in J$ die Lösung $U = U(x)$ auf J existiert. Dazu genügt es zu zeigen, dass für jedes $x^1 \in J$ die Lösung auf dem Intervall $[x^1 - \frac{1}{2\alpha}, x^1 + \frac{1}{2\alpha}] \cap J$ mit $\alpha = \max_{x \in J} |A(x)|$ [1] existiert.
Der Platz der Funktion $F(x, u)$ aus Abschnitt 12.2. wird von der Funktion $A(x)U + F(x)$ eingenommen. Damit kann $b > 0$ beliebig gewählt werden. Eine mögliche Wahl für M ist $M = \alpha b + \beta$ mit $\beta = \max_{x \in J} \left|A(x)U^0 + F(x)\right|$ (wegen $|A(x)U + F(x)| \leq |A(x)| \left|U - U^0\right| + \left|A(x)U^0 + F(x)\right|$). Die Behauptung folgt dann aus $\frac{b}{M} = \frac{b}{\alpha b + \beta} \to \frac{1}{\alpha}$ für $b \to \infty$. □

Wir betrachten jetzt simultan N Lösungen $U_1, \ldots, U_N$ der homogenen Gleichung, d. h. es gilt

$$U_j' = A(x)U_j, \quad x \in I,$$

für $1 \leq j \leq N$, und fassen diese N Lösungen zu einer $N \times N$-Matrix

$$X(x) = \begin{pmatrix} U_1(x) & \cdots & U_N(x) \end{pmatrix}$$

zusammen. Dann löst $X = X(x)$ die Matrixgleichung

$$X' = A(x)X.$$

[1] Die Matrixnorm ist $|A| = \max_{|u| \leq 1} |Au|$ für $A \in M(N, \mathbb{C})$. Wir bemerken gleich hier für spätere Verwendung, dass $|AB| \leq |A|\,|B|$ für $A, B \in M(N, \mathbb{C})$ gilt.

Die Lösungen $U_1, \ldots, U_N$ sind genau dann linear unabhängig, wenn $\det X(x) \neq 0$ für ein $x \in I$ (und dann für alle $x \in I$) gilt. In diesem Fall heißt $X = X(x)$ eine *Fundamentalmatrix*.

Satz 12.20 (Liouville)**.** *Es gilt*

$$\det X(x) = \det X(x^0)\, \exp\left(\int_{x^0}^{x} \operatorname{tr} A(x')\, dx'\right).$$

Beweis. Es gilt

$$\begin{aligned}\frac{d}{dx}\det X(x) &= \sum_{j=1}^{N} \det\begin{pmatrix} U_1 & \cdots & U_{j-1} & U_j' & U_{j+1} & \cdots & U_N\end{pmatrix}\\ &= \sum_{j=1}^{N} \det\begin{pmatrix} U_1 & \cdots & U_{j-1} & A(x)U_j & U_{j+1} & \cdots & U_N\end{pmatrix}.\end{aligned}$$

Bezeichnen wir für festes $x \in I$

$$F_x\colon \underbrace{\mathbb{C}^N \times \cdots \times \mathbb{C}^N}_{N\text{–mal}} \to \mathbb{C},$$

$$F_x(U_1, \ldots, U_N) = \sum_{j=1}^{N} \det\begin{pmatrix} U_1 & \cdots & U_{j-1} & A(x)U_j & U_{j+1} & \cdots & U_N\end{pmatrix},$$

so ist F_x eine alternierende Multilinearform. Der Raum dieser alternierenden Multilinearformen ist eindimensional, also

$$F_x(U_1, \ldots, U_N) = c_x \det\begin{pmatrix} U_1 & \cdots & U_N\end{pmatrix}$$

mit einer noch zu bestimmenden Konstanten $c_x \in \mathbb{C}$.

Für $U_1 = \begin{pmatrix}1\\0\\\vdots\\0\end{pmatrix}$, $U_2 = \begin{pmatrix}0\\1\\\vdots\\0\end{pmatrix}$, …, $U_N = \begin{pmatrix}0\\0\\\vdots\\1\end{pmatrix}$ ergibt sich jedoch leicht, dass $c_x = \operatorname{tr} A(x)$ gilt. Daher die Behauptung in der äquivalenten Form

$$\frac{d}{dx}\det X(x) = \operatorname{tr} A(x) \cdot \det X(x).$$

□

Für eine skalare Gleichung (12.16) wählt man m Lösungen $u_1 = u_1(x)$, …, $u_m = u_m(x)$. Diese Lösungen sind linear unabhängig, falls die *Wronski–Determinante*

$$W(x) = \det\begin{pmatrix} u_1 & u_2 & \cdots & u_m\\ u_1' & u_2' & \cdots & u_m'\\ \vdots & \vdots & \ddots & \vdots\\ u_1^{(m-1)} & u_2^{(m-1)} & \cdots & u_m^{(m-1)}\end{pmatrix}$$

verschieden von Null ist. In diesem Fall gilt

$$W(x) = W(x^0) \exp\left(-\int_{x^0}^{x} a_1(x')\,dx'\right).$$

Definition 12.21. $X(x, x^0)$ bezeichnet die Fundamentalmatrix mit $X(x^0, x^0) = I_N$, wobei I_N die $N \times N$-Einheitsmatrix ist.

Kennt man die Fundamentalmatrizen $X(x, x^0)$, so kann man das inhomogene Anfangswertproblem lösen.

Satz 12.22. *Die Lösung des AWP* (12.18) *ist durch*

$$U(x) = X(x, x^0)U^0 + \int_{x^0}^{x} X(x, x')F(x')\,dx'$$

gegeben.

Beweis. Durch die angegebene Formel wird offenbar eine C^1-Funktion $U(x)$ definiert. Weiterhin ist $U(x^0) = X(x^0, x^0)U^0 = U^0$ und

$$\begin{aligned} U'(x) &= \frac{\partial X}{\partial x}(x, x^0)U^0 + \int_{x^0}^{x} \frac{\partial X}{\partial x}(x, x')F(x')\,dx' + X(x, x)F(x) \\ &= A(x)\left[X(x, x^0)U^0 + \int_{x^0}^{x} X(x, x')F(x')\,dx'\right] + F(x) \\ &= A(x)U(x) + F(x). \end{aligned}$$

Somit löst $U = U(x)$ das AWP (12.18). Der Eindeutigkeitssatz vervollständigt den Beweis. □

Bemerkung. Für später bemerken wir, daß die Funktion $X(x, x^0)$ auch nach dem zweiten Parameter stetig differenzierbar ist und dass

$$\frac{\partial X}{\partial x^0}(x, x^0) = -X(x, x^0)A(x^0)$$

gilt.

Beweis. Dies folgt durch Differentation der Beziehung $X(x, x^0)X(x^0, x) = I$,[2] also

$$\frac{\partial X}{\partial x^0}(x, x^0)X(x^0, x) + X(x, x^0)A(x^0)X(x^0, x) = 0$$

und damit $\frac{\partial X}{\partial x^0}(x, x^0) + X(x, x^0)A(x^0) = 0$. □

[2] Allgemein gilt $X(x, x')X(x', x^0) = X(x, x^0)$ für alle x, x', $x^0 \in I$.

12.3.2 Lineare Differenzialgleichungen mit konstanten Koeffizienten

Wir betrachten jetzt das AWP (12.18) mit $A(x) = A \in M(N, \mathbb{C})$, also

$$\begin{cases} \quad U' = AU + F(x), \quad x \in I \\ U(x^0) = U^0, \end{cases} \tag{12.23}$$

In diesem Fall lassen sich die Fundamentalmatrizen $X(x, x^0)$ explizit angeben. Dazu führen wir das Exponential e^A einer Matrix $A \in M(N, \mathbb{C})$ ein.

Definition 12.24. Wir definieren

$$\mathrm{e}^A = \sum_{k=0}^{\infty} \frac{A^k}{k!} = I_N + \frac{A}{1!} + \frac{A^2}{2!} + \frac{A^3}{3!} + \dots .$$

Die Reihe $\sum\limits_{k=0}^{\infty} \frac{A^k}{k!}$ konvergiert wegen

$$\sum_{k=0}^{\infty} \left| \frac{A^k}{k!} \right| \leq \sum_{k=0}^{\infty} \frac{|A|^k}{k!} = \mathrm{e}^{|A|} < \infty$$

absolut.

Satz 12.25. *Es gilt*

$$X(x, x^0) = \mathrm{e}^{(x-x^0)A}.$$

Beweis. Wir haben $X(x^0, x^0) = e^{0 I_N} = 1$ und

$$\begin{aligned} \frac{\partial X}{\partial x}(x, x^0) &= \frac{\partial}{\partial x} \left(\sum_{k=0}^{\infty} \frac{(x - x^0)^k}{k!} A^k \right) \\ &= \sum_{k=1}^{\infty} \frac{(x - x^0)^{k-1}}{(k-1)!} A^k = A \sum_{k=0}^{\infty} \frac{(x - x^0)^k}{k!} A^k = AX(x, x^0). \end{aligned}$$

□

Bemerkung. Gilt $\left[A(x), A(x')\right] = 0$ für alle x, $x' \in I$ in (12.18), wobei $[A, B] = AB - BA$ den *Kommutator* zweier Matrizen $A, B \in M(N, \mathbb{C})$ bezeichnet, so ist

$$X(x, x^0) = \mathrm{e}^{\int_{x^0}^{x} A(x')\, dx'}.$$

Es bleibt e^{xA} für $A \in M(N, \mathbb{C})$ zu berechnen. Wir beginnen mit einem Jordanblock der Form

$$\begin{pmatrix} \mu & 1 & & 0 \\ & \mu & \ddots & \\ & & \ddots & 1 \\ 0 & & & \mu \end{pmatrix}, \quad \mu \in \mathbb{C}.$$

Satz 12.26. *Es gilt*

$$e^{xJ} = e^{\mu x} \begin{pmatrix} 1 & \frac{x}{1!} & \frac{x^2}{2!} & \cdots & \frac{x^{N-1}}{(N-1)!} \\ & 1 & \frac{x}{1!} & \cdots & \vdots \\ & & \ddots & \frac{x}{1!} & \frac{x^2}{2!} \\ & 0 & & 1 & \frac{x}{1!} \\ & & & & 1 \end{pmatrix},$$

d. h. auf der j-ten Nebendiagonalen für $j = 0, \ldots, N-1$ (oberhalb der Hauptdiagonalen) haben wir den konstanten Eintrag $\frac{x^j}{j!}$.

Beweis. Bezeichnen wir die rechte Seite mit $Y(x)$, so rechnet man leicht $Y(0) = I$ und $Y'(x) = JY(x)$ nach.
Alternative können wir auch $J^N = 0$ beobachten, dann

$$e^{xJ} = I_N + \frac{x}{1!}J + \frac{x^2}{2!}J^2 + \cdots + \frac{x^{N-1}}{(N-1)!}J^{N-1}.$$

□

Satz 12.27. *Ist $A = P \begin{pmatrix} J_1 & & 0 \\ & \ddots & \\ 0 & & J_b \end{pmatrix} P^{-1}$ für $J_1 \in M(N_1, \mathbb{C}), \ldots, J_b \in M(N_b, \mathbb{C})$ mit $N_1 + \cdots + N_b = N$ und $P \in \mathrm{Gl}(N, \mathbb{C})$ (dies kann beispielsweise die Jordanzerlegung von A sein), so gilt*

$$e^{xA} = P \begin{pmatrix} e^{xJ_1} & & 0 \\ & \ddots & \\ 0 & & e^{xJ_b} \end{pmatrix} P^{-1}.$$

Beweis. Direktes Nachrechnen. □

Wir wollen nun ein entsprechendes Resultat für skalare Differenzialgleichungen m-ter Ordnung mit konstanten Koeffizienten formulieren. Wir betrachten also die Gleichung

$$u^{(m)} + a_1 u^{(m-1)} + \cdots + a_{m-1}u' + a_m u = f(x), \quad x \in I,$$

wobei $a_j \in \mathbb{C}$ für $1 \leq j \leq m$. Transformation auf ein System erster Ordnung führt auf die Matrix

$$A = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & 1 \\ -a_m & -a_{m-1} & -a_{m-2} & \cdots & -a_2 & -a_1 \end{pmatrix}.$$

Satz 12.28. *Es gilt*

$$\det(\mu I_m - A) = \mu^m + a_1\mu^{m-1} + \cdots + a_{m-1}\mu + a_m.$$

Beweis. Durch Induktion nach m.
$\boldsymbol{m = 1}$. Dann ist $\det(\mu + a_1) = \mu + a_1$, d. h. die Behauptung gilt.
$\boldsymbol{m \to m + 1}$. Wir entwickeln die Determinante nach der letzten Spalte und erhalten

$$\begin{aligned}
&\det\begin{pmatrix} \mu & 1 & 0 & \cdots & 0 & 0 \\ 0 & \mu & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \mu & 1 \\ a_{m+1} & a_m & a_{m-1} & \cdots & a_2 & \mu + a_1 \end{pmatrix} \\
&\quad = (\mu + a_1)\det\begin{pmatrix} \mu & 1 & \cdots & 0 & 0 \\ 0 & \mu & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & \mu & 1 \\ 0 & 0 & \cdots & 0 & \mu \end{pmatrix} \\
&\qquad + \det\begin{pmatrix} \mu & 1 & 0 & \cdots & 0 & 0 \\ 0 & \mu & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \mu & 1 \\ a_{m+1} & a_m & a_{m-1} & \cdots & a_3 & \mu + a_2 \end{pmatrix} \\
&\qquad - \det\begin{pmatrix} \mu & 1 & 0 & \cdots & 0 & 0 \\ 0 & \mu & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \mu & 0 \\ a_{m+1} & a_m & a_{m-1} & \cdots & a_3 & \mu \end{pmatrix} \\
&\quad = (\mu + a_1)\mu^m + (\mu^m + a_2\mu^{m-1} + a_3\mu^{m-2} + \cdots + a_m\mu + a_{m+1}) - \mu^m \\
&\quad = \mu^{m+1} + a_1\mu^m + a_2\mu^{m-1} + \cdots + a_m\mu + a_{m+1}.
\end{aligned}$$

□

Folgerung. Es sei

$$\mu^m + a_1\mu^{m-1} + \cdots + a_{m-1}\mu + a_m = \prod_{j=1}^{r} (\mu - \mu_j)^{m_j}$$

mit $\mu_j \neq \mu_{j'}$ für $j \neq j'$, $m_1 + \cdots + m_r = m$. Dann schreibt sich die allgemeine Lösung des homogenen Anfangswertproblems als

$$u(x) = \sum_{j=1}^{r} \sum_{k=0}^{m_{j-1}} c_{jk}\, \mathrm{e}^{\mu_j x}\, \frac{x^k}{k!}$$

mit beliebigen Koeffizienten $c_{jk} \in \mathbb{C}$.

Beweis. Zurückführen auf ein System erster Ordnung und Dimensionsüberlegungen. □

Beispiel 12.29. Die gewöhnliche Differenzialgleichung

$$u^{(3)} - 4u'' + 5u' - 2 = f(x), \quad x \in I,$$

hat das *charakteristische Polynom* $\mu^3 - 4\mu^2 + 5\mu - 2 = (\mu - 1)^2(\mu - 2)$. Ein Fundamentalsystem ist daher durch

$$\mathrm{e}^x, \;\; x\mathrm{e}^x, \;\; \mathrm{e}^{2x}$$

gegeben.

12.3.3 Variation der Konstanten

Wir wollen jetzt eine *spezielle Lösung* des $N \times N$-Systems

$$U' = A(x)\, U + F(x), \quad x \in I, \tag{12.30}$$

finden. (Die allgemeine Lösung dieser Gleichung ist dann diese spezielle Lösung plus die allgemeine Lösung des homogenen Problems.)
Wir beginnen mit folgender Überlegung: Ist $X(x)$ eine Fundamentalmatrix, so ist $X(x)C$ für jedes $C \in \mathrm{Gl}(N, \mathbb{C})$ eine weitere Fundamentalmatrix. Umgekehrt lässt sich jede Fundamentalmatrix in dieser Form schreiben.

Beweis. Um letzteren Punkt einzusehen, nehmen wir eine weitere Fundamentalmatrix $Y(x)$ her und betrachten $X^{-1}(x)Y(x)$. Es gilt dann

$$\begin{aligned} \frac{d}{dx}\left(X^{-1}Y\right) &= -X^{-1}X'X^{-1}Y + X^{-1}Y' \\ &= X^{-1}\left[AXX^{-1}Y - AY\right] = 0. \end{aligned}$$

Folglich ist $X^{-1}(x)Y(x)$ konstant.
Um eine spezielle Lösung zu (12.30) zu finden, machen wir jetzt den Ansatz

$$U(x) = X(x)C(x)$$

mit $C \in C^1(I; \mathbb{C}^N)$. Dann gilt

$$U' = X'C + XC' = AXC + XC'.$$

Letzteres soll gleich $AU + F = AXC + F$ sein. Also benötigen wir $XC' = F$, $C' = X^{-1}F$ bzw.

$$C(x) = \int_{x^0}^{x} X^{-1}(x')F(x')\, dx'$$

für ein $x^0 \in I$. (Eine andere Wahl der Integrationskonstanten führt zu einer anderen speziellen Lösung.) □

Bemerkung. In den Bezeichnungen von Abschnitt 12.3.1. gilt

$$X(x, x') = X(x)X^{-1}(x').$$

Folglich ist die gerade konstruierte Lösung $U = U(x)$ zu (12.30) genau diejenige mit $U(x^0) = 0$.

Wir wollen nun obiges Schema auf *skalare lineare Gleichungen* spezifizieren. Sei also die Gleichung

$$u^{(m)} + a_1(x)u^{(m-1)} + \cdots + a_{m-1}(x)u' + a_m(x) = f(x), \quad x \in I \tag{12.31}$$

vorgelegt. Wir nehmen an, dass wir ein Fundamentalsystem dieser Gleichung, d. h. m linear unabhängige Lösungen $u_1, \dots, u_m$ des homogenen Problems zu (12.31) kennen. Wir suchen dann eine Lösung $u = u(x)$ des inhomogenen Problems in der Form

$$u(x) = c_1(x)u_1(x) + \cdots + c_m(x)u_m(x)$$

mit zu bestimmenden Funktionen $c_1, \dots, c_m$.
Gemäß unseres Schemas haben wir diese Funktionen $c_1, \dots, c_m$ so zu wählen, dass

$$\sum_{k=1}^{m} c_k'(x)u_k^{(j)}(x) = 0, \quad 0 \le j \le m-2,$$
$$\sum_{k=1}^{m} c_k'(x)u_k^{(m-1)}(x) = f(x)$$

gilt. Dies ist ein lineares System in $c_1'(x), \dots, c_m'(x)$ mit Koeffizientendeterminante $W(x) \neq 0$, wobei die Wronski-Determinante $W(x)$ bezüglich $u_1, \dots, u_m$ gebildet wurde.
Bezeichnet $W_k(x)$ die Determinante, die man erhält, indem man in $W(x)$ die k-te Spalte $\begin{pmatrix} u_k \\ u_k' \\ \vdots \\ u_k^{(m-1)} \end{pmatrix}$ durch $\begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ f \end{pmatrix}$ ersetzt, so gilt (siehe die AGLA-Vorlesung)

$$c_k'(x) = \frac{W_k(x)}{W(x)}, \quad k = 1, \dots, m,$$

also beispielsweise

$$c_k(x) = \int_{x^0}^{x} \frac{W_k(x')}{W(x')}\, dx', \quad k = 1, \dots, m,$$

für ein $x^0 \in I$.

Beispiel 12.32. Wir betrachten die Gleichung

$$u'' - 3u' + 2u = \mathrm{e}^{4x}.$$

Wegen $\mu^2 - 3\mu + 2 = (\mu - 1)(\mu - 2)$ ist ein Fundamentalsystem gegeben durch e^x, e^{2x}. Weiterhin gilt

$$\begin{aligned}
W(x) &= \det \begin{pmatrix} \mathrm{e}^x & \mathrm{e}^{2x} \\ \mathrm{e}^x & 2\mathrm{e}^{2x} \end{pmatrix} = \mathrm{e}^{3x}, \\
W_1(x) &= \det \begin{pmatrix} 0 & \mathrm{e}^{2x} \\ \mathrm{e}^{4x} & 2\mathrm{e}^{2x} \end{pmatrix} = -\mathrm{e}^{6x}, \\
W_2(x) &= \det \begin{pmatrix} \mathrm{e}^x & 0 \\ \mathrm{e}^x & \mathrm{e}^{4x} \end{pmatrix} = \mathrm{e}^{5x}.
\end{aligned}$$

Wir erhalten $c_1'(x) = \frac{W_1(x)}{W(x)} = -\mathrm{e}^{3x}$, $c_2'(x) = \frac{W_2(x)}{W(x)} = \mathrm{e}^{2x}$,

$$c_1(x) = -\frac{1}{3}\,\mathrm{e}^{3x}, \quad c_2(x) = \frac{1}{2}\,\mathrm{e}^{2x}.$$

Eine spezielle Lösung ist

$$u(x) = c_1(x)\mathrm{e}^x + c_2(x)\mathrm{e}^{2x} = -\frac{1}{3}\,\mathrm{e}^{4x} + \frac{1}{2}\,\mathrm{e}^{4x} = \frac{1}{6}\,\mathrm{e}^{4x}.$$

Die allgemeine Lösung ergibt sich zu

$$a_1\mathrm{e}^x + a_2\mathrm{e}^{2x} + \frac{1}{6}\,\mathrm{e}^{4x}$$

mit beliebigen Konstanten $a_1, a_2 \in \mathbb{C}$.

12.3.4 Asymptotische Eigenschaften von Fundamentalsystemen: Die Liouville-Greensche Approximation

Heuristik

Wir betrachten eine skalare homogene Gleichung zweiter Ordnung

$$\frac{d^2u}{dx^2} = f(x)u, \quad x \in I, \tag{12.33}$$

wobei f eine zweimal stetig differenzierbare, reellwertige Funktion bezeichnet, *die auf auf I nirgends veschwindet*. Wir möchten diese Gleichung approximativ lösen. In Regionen, in denen f in etwa konstant ist, ist eine erste Näherung

$$u(x) \sim k_1\mathrm{e}^{x\sqrt{f(x)}} + k_2\mathrm{e}^{-x\sqrt{f(x)}}$$

mit beliebigen Konstanten k_1, k_2. Diese Näherung gibt einen Hinweis auf den Charakter der Lösungen (diese sind vom exponentiellen Typ im Fall $f(x) > 0$

und vom trigonometrischen bzw. oszillierenden Typ im Fall $f(x) < 0$), doch oft ist diese Approximation nicht gut geeignet für Betrachtungen global auf I.
Eine bessere Möglichkeit verfährt wie folgt: Wir führen eine neue Koordinate $y = y(x)$ ein und setzen

$$v = \sqrt{\frac{dy}{dx}}\, u.$$

Lemma 12.34. *v genügt der Differenzialgleichung*

$$\frac{d^2v}{dy^2} = \left[\left(\frac{dx}{dy}\right)^2 f(x) + \left(\frac{dx}{dy}\right)^{1/2} \frac{d^2}{dy^2}\left(\left(\frac{dx}{dy}\right)^{-1/2}\right)\right] v. \tag{12.35}$$

Beweis. Nachrechnen. □

Bemerkung. Der Term $(dx/dy)^{1/2} \frac{d^2}{dy^2}\big((dx/dy)^{-1/2}\big)$ kann auch als

$$-\frac{1}{2}\left(\frac{d^3x/dy^3}{dx/dy} - \frac{3}{2}\left(\frac{d^2x/dy^2}{dx/dy}\right)^2\right)$$

geschrieben werden. Letzteres ohne den Faktor -1/2 ist die sogenannte *Schwarzsche Ableitung*.

Wir behandeln zuerst den Fall

$$f(x) > 0, \quad \forall x \in I.$$

Unter der Annahme, dass der zweite Summand in der Klammer $[\ldots]$ in (12.35) klein ist, möchten wir, dass

$$\left(\frac{dx}{dy}\right)^2 f(x) = 1, \quad x \in I,$$

gilt. Eine näherungsweise Lösung der Gleichung (12.35) ist dann

$$v(y) \sim k_1 \mathrm{e}^{y} + k_1 \mathrm{e}^{-y}.$$

Wir setzen also

$$y(x) = \int f^{1/2}(x)\, dx.$$

Dann schreibt sich (12.35) in der Form

$$\frac{d^2v}{dy^2} = (1+\varphi)v \tag{12.36}$$

mit

$$\varphi = \frac{4f(x)f''(x) - 5f'^2(x)}{16 f^{(3)}(x)} = -\frac{1}{f^{3/4}} \frac{d^2}{dx^2}\left(\frac{1}{f^{1/4}}\right).$$

Definition 12.37. Die Näherungslösung

$$u(x) \sim k_1 f^{-1/4} \mathrm{e}^{\int f^{1/2}\,dx} + k_2 f^{-1/4} \mathrm{e}^{-\int f^{1/2}\,dx}$$

zu (12.36) heißt die *Liouville-Greensche Approximation*.

Bemerkung. Betrachtet man stattdessen eine Gleichung der Form

$$\frac{d^2u}{dx^2} = (f(x) + g(x))\,u,$$

so führt obige Transformation auf

$$\frac{d^2v}{dy^2} = \left(1 + \varphi + \frac{g}{f}\right) v$$

mit φ wie oben.

Abschätzung der Fehlerterme

Seien die Annahmen wie oben erfüllt, insbesondere nehme $f(x)$ positive Werte an. Ferner sei $I = (x^0, x^1)$ ein endliches oder unendliches Intervall. Die Funktion g sei stetig auf I (nicht notwendigerweise reellwertig). Wir setzen

$$F(x) = \int \left[\frac{1}{f^{1/4}} \frac{d^2}{dx^2} \left(\frac{1}{f^{1/4}} \right) - \frac{g}{f^{1/2}} \right] dx.$$

Für eine C^1-Funktion h auf I sei die *Variation* von h durch

$$V_{x^0,x}(h) = \int_{x^0}^{x} |h'(x')|\,dx', \quad x^0 \leq x \leq x^1,$$

erklärt, analog für $V_{x^1,x}(h) = V_{x,x^1}(h)$.

Theorem 12.38. *Unter diesen Voraussetzungen besitzt die Gleichung*

$$\frac{d^2u}{dx^2} = (f(x) + g(x))\,u \tag{12.39}$$

zwei C^2-Lösungen u_0, u_1 mit

$$u_0(x) = f^{-1/4}(x) \exp\left(\int f^{1/2}(x)\,dx \right) (1 + \epsilon_0(x)),$$
$$u_1(x) = f^{-1/4}(x) \exp\left(- \int f^{1/2}(x)\,dx \right) (1 + \epsilon_1(x)),$$

wobei

$$|\epsilon_j(x)| \leq \exp\left(1/2 V_{x^j,x}(F)\right) - 1,$$
$$\frac{1}{2} f^{-1/2}(x) \left|\epsilon_j'(x)\right| \leq \exp\left(\frac{1}{2} V_{x^j,x}(F) \right) - 1$$

für $j = 0, 1$ unter der Bedingung, dass $V_{x^j,x}(F) < \infty$ gilt. Ist die Funktion g reellwertig, so sind diese Lösungen reellwertig.

Beweis. Wir schreiben Gleichung (12.39) in der Form

$$\frac{d^2v}{dy^2} = (1+\psi(y))\, v \tag{12.40}$$

mit $u = f^{-1/4}(x)v$ und

$$\psi(y) = \frac{g}{f} - \frac{1}{f^{3/4}} \frac{d^2}{dx^2}\left(\frac{1}{f^{1/4}}\right).$$

Sei $y = y^0$ für $x = x^0$ und $y = y^1$ für $x = x^1$. Dann ist ψ eine stetige Funktion im Intervall (y^0, y^1). Wir setzen $v(y) = \mathrm{e}^y\,(1 + h(y))$ und erhalten

$$h''(y) + 2h'(y) - \psi(y)h(y) = \psi(y).$$

Betrachten wir in dieser Gleichung den Term $-\psi h$ als Korrekturterm, so erhalten wir

$$h(y) = \frac{1}{2}\int_{y^0}^{y} \left(1 - \mathrm{e}^{2(s-y)}\right) \psi(s)\,(1 + h(s))\; ds. \tag{12.41}$$

Wir lösen diese Integralgleichung *iterativ*. Dazu nehmen wir $y^0 > -\infty$ und ψ rechtsseitig stetig an y^0 an. (Der allgemeine Fall ergibt sich analog unter Benutzung von $V_{x^0,x}(F) < \infty$.)
Wir definieren $h_0(y) = 0$ und setzen

$$h_k(y) = \frac{1}{2}\int_{y^0}^{y} \left(1 - \mathrm{e}^{2(s-y)}\right) \psi(s)\,(1 + h_{k-1}(s))\; ds$$

für $k \geq 1$.

Behauptung 12.42. *Für alle $k \geq 1$*

$$|h_k(y) - h_{k-1}(y)| \leq \frac{\Psi^k(y)}{2^k k!}, \tag{12.43}$$

wobei $\Psi(y) = \int_{y^0}^{y} |\psi(s)|\; ds$.

Beweis. Mittels Induktion über k.
$\boldsymbol{k = 1}$. Es gilt

$$h_1(y) = \frac{1}{2}\int_{y^0}^{y} \left(1 - \mathrm{e}^{2(s-y)}\right) \psi(s)\, ds,$$

also $|h_1(y)| \leq \frac{1}{2}\Psi(y)$ wegen $0 \leq 1 - \mathrm{e}^{2(s-y)} < 1$ für $s \in [y^0, y]$.
$\boldsymbol{k \longrightarrow k+1}$. Wir haben

$$h_{k+1}(y) - h_k(y) = \frac{1}{2}\int_{y^0}^{y} \left(1 - \mathrm{e}^{2(s-y)}\right) \psi(s)\,(h_k(s) - h_{k-1}(s))\; ds, \tag{12.44}$$

also mit der Induktionsvoraussetzung

$$|h_{k+1}(y) - h_k(y)| \leq \frac{1}{2^{k+1}k!}\int_{y^0}^{y} |\psi(s)|\,\Psi^k(s)\, ds = \frac{\Psi^{k+1}(y)}{2^{k+1}(k+1)!}.$$

□

Aus (12.43) folgt, dass die Reihe

$$h(y) = \sum_{k=0}^{\infty} (h_{k+1}(y) - h_k(y))$$

gleichmäßig auf kompakten Teilintervallen von $[x^0, x^1)$ konvergiert. Durch Summation von (12.44) finden wir, dass

$$\begin{aligned} h(y) &= h_1(y) + \sum_{k=1}^{\infty} (h_{k+1}(y) - h_k(y)) \\ &= \frac{1}{2} \int_{y^0}^{y} \left(1 - \mathrm{e}^{2(s-y)}\right) \psi(s)\, ds \\ &\qquad + \sum_{k=1}^{\infty} \frac{1}{2} \int_{y^0}^{y} \left(1 - \mathrm{e}^{2(s-y)}\right) \psi(s) \left(h_k(s) - h_{k-1}(s)\right)\, ds \\ &= \frac{1}{2} \int_{y^0}^{y} \left(1 - \mathrm{e}^{2(s-y)}\right) \psi(s) \left(1 + h(s)\right)\, ds, \end{aligned}$$

also genügt die Funktion h der Integralgleichung (12.41) .
Wir zeigen als nächstes, dass h zweimal stetig differenzierbar ist. Dazu müssen wir zeigen, dass die Reihe

$$\sum_{k=0}^{\infty} \left(h''_{k+1}(y) - h''_k(y)\right)$$

lokal gleichmäßig konvergiert.
Wir haben

$$\begin{aligned} h'_1(y) &= \int_{y^0}^{y} \mathrm{e}^{2(s-y)} \psi(s)\, ds \\ h'_{k+1}(y) - h'_k(y) &= \int_{y^0}^{y} \mathrm{e}^{2(s-y)} \psi(s) \left(h_k(s) - h_{k-1}(s)\right)\, ds. \end{aligned}$$

Zusammen mit $\left|\mathrm{e}^{2(s-y)}\right| \leq 1$ und der Abschätzung (12.43) ergibt sich

$$\left|h'_{k+1}(y) - h'_k(y)\right| \leq \frac{\Psi^{k+1}(y)}{2^k (k+1)!}$$

für $k \geq 0$. Folglich ist $\sum\limits_{k=0}^{\infty} \left(h'_{k+1}(y) - h'_k(y)\right)$ lokal gleichmäßig konvergent.
Um die Existenz und Stetigkeit der zweiten Abbildung zu bekommen, benutzen wir die Beziehungen

$$\begin{aligned} h''_1(y) &= -2h'_1(y) + \psi(y), \\ h''_{k+1}(y) - h''_k(y) &= -2\left(h'_{k+1}(y) - h'_k(y)\right) + \psi(y) \left(h_k(y) - h_{k-1}(y)\right). \end{aligned}$$

Summation ergibt, dass h der Differenzialgleichung

$$h''(y) + 2h'(y) = (1 + \psi(y))\, h(y)$$

genügt. Zudem erhalten wir die Abschätzungen

$$|h(y)| \leq \mathrm{e}^{\Psi(y)/2} - 1, \quad 1/2\, |h'(y)| \leq \mathrm{e}^{\Psi(y)/2} - 1.$$

Kehren wir nun zu der ursprünglichen x-Variablen zurück, so ist $F(x) = -\int \psi(y)\, dy$ (wegen $dy = f^{1/2}(x)\, dx$), also $\Psi(y) = V_{x^0,x}(F)$.
Die Behauptungen bezüglich u_0 folgen. Die Herleitungen für u_1 sind analog. □

Bemerkung. Insbesondere erhalten wir

$$\epsilon_0(x) \to 0, \quad f^{-1/2}(x)\epsilon_0'(x) \to 0 \text{ für } x \to x^0 + 0.$$

Analog für ϵ_1 und $x \to x^1 - 0$.
Theorem 12.38 besagt, dass wir eine Lösung u_0 mit

$$u_0(x) \sim f^{-1/4} \exp\left(\int f^{1/2}\, dx\right) \text{ für } x \to x^0 + 0$$

finden. Eine interessante Frage im Fall, dass $\int f^{1/2}\, dx$ für $x \to x^1 - 0$ divergiert, ist, ob es eine Lösung u_2 mit

$$u_2(x) \sim f^{-1/4} \exp\left(\int f^{1/2}\, dx\right) \text{ für } x \to x^1 - 0$$

gibt. Eine erste Antwort darauf gibt der folgende Satz:

Satz 12.45. *Sei zusätzlich $V_{x^0,x^1}(F) < \infty$ und $\int f^{1/2}\, dx \to \infty$ für $x \to x^1 - 0$. Dann*

$$\epsilon_0(x) \to k, \quad f^{-1/2}(x)\epsilon_0'(x) \to 0 \text{ für } x \to x^1 - 0,$$

wobei k eine Konstante ist.

Beweis. Der Beweis benutzt $y^1 = \infty$ und die Konstruktionen des vorigen Beweises. □

Gilt jetzt also

$$\epsilon_0(x^1 - 0) \neq -1,$$

so ist u_0 eine Funktion mit den gewünschten Eigenschaften. Im Allgemeinen findet man eine solche Lösung u_2, indem man die Stelle x^0 hinreichend dicht an x^1 wählt.
Allerdings ist eine Lösung u_2 mit der Eigenschaft

$$u_2(x) \sim f^{-1/4} \exp\left(\int f^{1/2}\, dx\right) \text{ für } x \to x^1 - 0$$

nicht eindeutig bestimmt. Man kann stets ein beliebiges Vielfaches von u_1 addieren und erhält eine weitere Lösung mit derselben Eigenschaft.

Definition 12.46. Nahe x^1 heißt das asymptotische Verhalten von u_2 *dominant*, das von u_1 *subdominant*. Eine analoge Klassifizierung gilt nahe x^0.

Beispiel 12.47. Wir betrachten die Gleichung

$$u'' = (x + \log x)\, u \tag{12.48}$$

für $x \to \infty$. Da das Integral $\int \frac{\log x}{x^{1/2}}\, dx$ für $x \to \infty$ divergiert, können wir *nicht* $f(x) = x$, $g(x) = \log x$ wählen.
Jedoch ist $f(x) = x + \log x$, $g(x) = 0$ eine mögliche Wahl. Man sieht dann, dass

$$f^{-1/4}\big(f^{-1/4}\big)'' = O\big(x^{-5/2}\big) \quad \text{für } x \to \infty,$$

also konvergiert $V(F)$ für $x \to \infty$. (12.48) hat folglich asymptotische Lösungen der Form

$$(x + \log x)^{-1/4} \exp\left(\pm \int (x + \log x)^{1/2}\, dx\right).$$

Dieses Ergebnis lässt sich verbessern. Dazu beobachten wir, dass

$$(x + \log x)^{1/2} = x^{1/2} + 1/2x^{-1/2} \log x + O(x^{-1/2}(\log x)^2),$$

folglich

$$\int (x + \log x)^{1/2}\, dx = \frac{2}{3}x^{3/2} + x^{1/2}\log x - 2x^{1/2} + k + o(1)$$

für $x \to \infty$. Damit besitzt (12.48) eine eindeutige (subdominante) Lösung u_1 mit

$$u_1(x) \sim x^{-1/4-\sqrt{x}} \exp\left(2x^{1/2} - \frac{2}{3}x^{3/2}\right)$$

für $x \to \infty$ und nicht eindeutige (dominante) Lösungen u_2 mit

$$u_2(x) \sim x^{-1/4+\sqrt{x}} \exp\left(2x^{1/2} - \frac{2}{3}x^{3/2}\right)$$

für $x \to \infty$.

Theorem 12.49. *Unter obigen Voraussetzungen*[3] *betrachten wir die Gleichung*

$$\frac{d^2u}{dx^2} = (-f(x) + g(x))\, u. \tag{12.50}$$

Diese besitzt zwei C^2-Lösungen u_0, u_1 mit

$$u_0(x) = f^{-1/4}(x) \exp\left(\mathrm{i} \int f^{1/2}(x)\, dx\right)(1 + \epsilon_0(x))\,,$$
$$u_1(x) = f^{-1/4}(x) \exp\left(-\mathrm{i} \int f^{1/2}(x)\, dx\right)(1 + \epsilon_1(x))$$

[3]Insbesondere nimmt f Werte in den *positiven* reellen Zahlen an.

mit

$$\begin{aligned} |\epsilon_j(x)| &\leq \exp V_{a,x}(F) - 1, \\ f^{-1/2}(x) \left|\epsilon_j'(x)\right| &\leq \exp V_{a,x}(F) - 1 \end{aligned}$$

für $j = 0, 1$ unter der Voraussetzung, dass $V_{x^*,x}(F) < \infty$ für ein fixiertes x^* mit $x^0 \leq x^* \leq x^1$ gilt. Ist g reellwertig, so sind die Lösungen u_0, u_1 komplex-konjugiert.

Beweis. Der Beweis ist analog dem vorigen Beweis. Die Integralgleichung für h lautet diesmal

$$h(y) = \frac{1}{2\mathrm{i}} \int_{y^*}^{y} \left(1 - \mathrm{e}^{2\mathrm{i}(s-y)}\right) \psi(s) \left(1 + h(s)\right) \, ds,$$

wobei $y = y^*$ für $x = x^*$ ist. Der fehlende Faktor von $\frac{1}{2}$ in den Fehlerabschätzungen erklärt sich aus der Abschätzung

$$\left|1 - \mathrm{e}^{2\mathrm{i}(s-y)}\right| \leq 2.$$

□

Die Wahl der Stelle x^* fixiert wegen

$$\epsilon_j(x) \to 0, \quad f^{-1/2}\epsilon_j'(x) \to 0 \text{ für } x \to x^*$$

und $j = 0, 1$ die Wahl der *Anfangsbedingungen.*

Bemerkung. Im Fall, dass g reellwertig ist, lassen sich reellwertige Lösungen u zu (12.50) in der Form

$$u(x) = A f^{-1/4}(x) \left(\sin\left(\int f^{1/2}(x) \, dx + \delta\right) + \epsilon(x)\right)$$

mit gewissen Konstanten A, δ und

$$|\epsilon(x)|, \quad f^{-1/2}(x) \left|\epsilon'(x)\right| \leq \exp V_{x^*,x}(F) - 1$$

schreiben.

Asymptotische Eigenschaften bezüglich eines Parameters

Betrachten wir anstelle von (12.39) die Gleichung

$$\frac{d^2u}{dx^2} = \left(\nu^2 f(x) + g(x)\right) u,$$

wobei $\nu > 0$ ein Parameter ist, so erhalten wir unter den Voraussetzungen von Theorem 12.38 die Existenz zweier C^2-Lösungen $u_0(x,\nu)$, $u_1(x,\nu)$ mit

$$u_j(x,\nu) = f^{-1/4}(x) \exp\left((-1)^j \nu \int f^{1/2}(x) \, dx\right) \left(1 + \epsilon_j(x,\nu)\right),$$

wobei

$$\begin{aligned} |\epsilon_j(x,\nu)| &\leq \exp\left(\frac{V_{x^j,x}(F)}{\nu}\right) - 1, \\ \frac{\left|\epsilon_j'(x,\nu)\right|}{2\nu f^{1/2}(x)} &\leq \exp\left(\frac{V_{x^j,x}(F)}{\nu}\right) - 1 \end{aligned}$$

für $j = 0, 1$. Die rechten Seiten in diesen Fehlerabschätzungen sind $O(\nu^{-1})$ für $\nu \to \infty$ bei fixiertem x.

Theorem 12.51. *Gilt $V_{x^0,x^1}(F) < \infty$, so ist*

$$u_j(x,v) \sim f^{-1/4}(x) \exp\left((-1)^j \nu \int f^{1/2}(x)\,dx\right)$$

für $\nu \to \infty$ gleichmäßig auf (x^0, x^1).

Ein analoges Resultat gilt im Fall der Gleichung

$$\frac{d^2u}{dx^2} = \left(-\nu^2 f(x) + g(x)\right) u,$$

wobei $\nu > 0$ wiederum ein Parameter ist.

12.4 Spezielle Lösungsmethoden

12.4.1 Separation der Variablen

Die Methode der Separation der Variablen arbqeitet für skalare Gleichungen der Form

$$\begin{cases} u' = f(x)g(u), \quad x \in I, \\ u(x^0) = u^0 \end{cases} \tag{12.52}$$

mit f, g stetig.
Unter der Annahme $g(u^0) \neq 0$ ist

$$H(u) = \int_{u^0}^{u} \frac{du'}{g(u')}$$

für u nahe u^0 eine monotone, stetig differenzierbare Funktion. Folglich existiert die Umkehrfunktion H^{-1} nahe 0.

Satz 12.53. *Für $g(u^0) \neq 0$ ist die (lokal eindeutige) Lösung $u = u(x)$ zu* (12.52) *durch*

$$u(x) = H^{-1}\left(\int_{x^0}^{x} f(x')\,dx'\right)$$

gegeben.

Beweis. Sei $u = u(x)$ eine C^1-Lösung zu (12.52). Dann gilt

$$\int_{x^0}^{x} \frac{u'(x')\,dx'}{g\,(u(x'))} = \int_{x^0}^{x} f(x')\,dx',$$

also nach Variablenwechsel $v = u(x)$ für x nahe x^0

$$H\,(u(x)) = \int_{u^0}^{u(x)} \frac{dv}{g(v)} = \int_{x^0}^{x} f(x')\,dx'.$$

Ist umgekehrt $u = u(x)$ durch diese Formel gegeben, so ist u offenbar eine C^1-Funktion, $u(x^0) = u^0$ sowie

$$u'(x) = \frac{f(x)}{H'\left(H^{-1}\left(\int_{x^0}^{x} f(x')\,dx'\right)\right)} = \frac{f(x)}{H'\,(u(x))} = f(x)g(u(x)).$$

Folglich ist u Lösung von (12.52). □

Bemerkung. Formal schreibt man $\frac{du}{g(u)} = f(x)\,dx$ und integriert, also

$$\int \frac{du}{g(u)} = \int f(x)\,dx.$$

Beispiel 12.54. Wir betrachten die Gleichung

$$u' = u(1-u).$$

Für Lösungen $u \not\equiv 0$ und $u \not\equiv 1$ führt Separation der Variablen auf $\int \frac{du}{u(1-u)} = \int dx$, also wegen $\frac{1}{u(1-u)} = \frac{1}{u} + \frac{1}{1-u}$

$$\log|u| + \log|1-u| = x + k$$

mit $k \in \mathbb{R}$. Einfache algebraische Manipulationen führen zur Lösung

$$u(x) = \frac{1}{1 - a\mathrm{e}^{-x}}$$

mit $a \in \mathbb{R}$.

12.4.2 Exakte Differenzialgleichungen und integrierende Faktoren

Definitionen und Beispiele

Definition 12.55. (i) Eine Differenzialgleichung, die in der Form

$$\frac{d}{dx}F(x, u, u', \ldots, u^{(m-1)}) = 0$$

geschrieben werden kann[4], heißt *exakt*.

[4] F ist wenigstens von der Klasse C^1.

(ii) Wird eine Differenzialgleichung durch Multiplikation mit dem Faktor $G(x, u, u', \dots, u^{(m)})$ exakt, so heißt $G(x, u, \dots, u^{(m)})$ ein *integrierender Faktor*.

Die Lösungen einer exakten Differenzialgleichung sind *implizit* durch die Gleichungen

$$F(x, u, u', \dots, u^{(m-1)}) = k$$

gegeben, wobei k eine beliebige Konstante ist.

Beispiel 12.56. 1. Die Gleichung $u'' + xu' + u = 0$ kann als $\frac{d}{dx}(u' + xu) = 0$ geschrieben werden und ist somit exakt. Die Lösungen sind

$$u(x) = \left(k_1 \int_0^x e^{t^2/2}\, dt + k_2\right) e^{-x^2/2}.$$

2. Die Gleichung $u'' + \frac{x+1}{x}\, u' + \frac{x-1}{x^2}\, u = 0$ wird durch Multiplikation mit e^x exakt:

$$\frac{d}{dx}\left(e^x u' + \frac{e^x}{x} u\right) = e^x u'' + e^x\, \frac{x+1}{x}\, u' + e^x\, \frac{x-1}{x^2}\, u = 0.$$

Somit gilt $u' + \frac{u}{x} = k_1\, e^{-x}$ mit der Lösung

$$u(x) = -k_1\, \frac{x+1}{x}\, e^{-x} + \frac{k_2}{x}.$$

Der Fall $m = 1$

Besonders interessant ist der Fall $m = 1$.

Satz 12.57. *Die Differenzialgleichung* $F_0(x, u)u' + F_1(x, u) = 0$ *ist genau dann exakt, wenn*[5]

$$\frac{\partial F_0}{\partial x} = \frac{\partial F_1}{\partial u}.$$

Beweis. ($\Rightarrow$) Ist die Differenzialgleichung exakt, so gibt es ein $F = F(x, u)$ mit $F_0 = \frac{\partial F}{\partial u}$, $F_1 = \frac{\partial F}{\partial x}$. Die Behauptung folgt.

($\Leftarrow$) Gilt umgekehrt $\frac{\partial F_0}{\partial u} = \frac{\partial F_1}{\partial x}$, so finden wir ein derartiges F, indem wir

$$F(x, u) = \int_{(x^0, u^0)}^{(x,u)} F_0(x', u')\, du' + F_1(x', u')\, dx'$$

setzen.[6]

□

[5]Vorausgesetzt, diese partiellen Ableitungen existieren und sind stetig.

[6]Dies ist das Kurvenintegral zweiter Art entlang eines beliebigen stückweise C^1-Weges γ, der im (x', u')-Raum die Stellen (x^0, u^0) und (x, u) verbindet. Man erhält dieses Integral,

Beispiel 12.58.
1. Die Gleichung $(u^2+x)u' + (u-x^2) = 0$ ist exakt. Tatsächlich gilt
$$\frac{\partial}{\partial x}(u^2+x) = \frac{\partial}{\partial u}(u-x^2) = 1.$$
Die Lösungen u sind implizit durch
$$u^3 + 3xu - x^3 = k$$
gegeben.
2. Separable Gleichungen der Form $F_0(u)u' + F_1(x) = 0$ sind wegen $\frac{\partial F_0}{\partial x} = \frac{\partial F_1}{\partial u} = 0$ exakt.
3. Die Gleichung $(1+xu+u^2) + (1+xu+x^2)u' = 0$ ist nicht exakt. Ein integrierender Faktor ist e^{xu}. Man erhält
$$\frac{d}{dx}\bigl(\mathrm{e}^{xu}(x+u)\bigr) = \mathrm{e}^{xu}\left(1+xu+x^2\right)u' + \mathrm{e}^{xu}\left(1+xu+u^2\right) = 0.$$
Folglich gilt
$$(x+u)\,\mathrm{e}^{xu} = k.$$

Ein weiteres wichtiges Beispiel ist:

Lemma 12.59. *Ein integrierender Faktor für die skalare Gleichung*
$$u' + a(x)u = f(x)$$
erster Ordnung ist $\exp\left(\int_{x^0}^x a(x')\,dx'\right)$.

Beweis. Tatsächlich gilt
$$\frac{d}{dx}\left(\exp\left(\int_{x^0}^x a(x')\,dx'\right)u(x)\right) = \exp\left(\int_{x^0}^x a(x')\,dx'\right)\bigl(u'(x) + a(x)u(x)\bigr).$$
□

---

indem man auf der rechten Seite in der Gleichung in Definition 11.7
$$\int_a^b (F_1, F_0)^T\,(\gamma(t))\cdot\gamma'(t)\,dt$$
setzt. Eine mögliche Wahl des Weges γ ist
$$\gamma(t) = \begin{cases} \left(x^0, u^0 + t(u-u^0)\right), & 0 \le t \le 1, \\ \left(x^0 + (t-1)(x-x^0), u\right), & 1 \le t \le 2 \end{cases}$$
(dann $a = 0$, $b = 2$).
Mit orientierter Integration werden wir uns im nächsten Semester in der Vorlesung „Differential- und Integralrechnung III" beschäftigen. Dann werden wir insbesondere sehen, dass obige Definition von F ein Ergebnis liefert, das von der Wahl des verbindenden Weges γ unabhängig ist.

Die Lösung u, die man mit Hilfe dieses integrierenden Faktors unter der Anfangsbedingung $u(x^0) = u^0$ erhält, ist

$$u(x) = \mathrm{e}^{-\int_{x^0}^{x} a(x')\,dx'} u^0 + \int_{x^0}^{x} \mathrm{e}^{-\int_{x'}^{x} a(t)\,dt} f(x')\,dx'.$$

Dies besagt aber nichts anderes, als dass das Fundamentalsystem $v(x, x^0)$ mit $v(x^0, x^0) = 1$ durch

$$v(x, x^0) = \mathrm{e}^{-\int_{x^0}^{x} a(x')\,dx'}$$

gegeben ist.

Bemerkung. Dieses Beispiel zeigt, dass das Finden eines integrierenden Faktors im Fall $m = 1$ zum Lösen der Differenzialgleichung äquivalent und damit genauso schwierig ist.

12.4.3 Autonome Differenzialgleichungen

Definition 12.60. Eine Differenzialgleichung auf $\mathbb{R}$, die invariant unter den Transformationen $x \mapsto x - k$ für $k \in \mathbb{R}$ ist, heißt *autonom*.

Autonome Differenzialgleichungen sind solche, die nicht explizit von der Variablen x abhängen. Sie lassen sich in der Ordnung um Eins reduzieren, indem man $v(x) = u'(x)$ als Funktion von u schreibt.[7] Man erhält:

$$\begin{aligned}
u'(x) &= v(u),\\
u''(x) &= \frac{dv}{dx} = \frac{dv}{du} \cdot \frac{du}{dx} = v'(u)v(u),\\
u^{(3)}(x) &= \frac{d}{dx}\big(v'(u)v(u)\big) = v(u){v'}^2(u) + v^2(u)v''(u),
\end{aligned}$$

usw.

Beispiel 12.61. 1. Die Substitution $v(u) = u'(x)$ vereinfacht die Gleichung $u^{(3)} + uu' = 0$ nach Division durch v zu

$$ {v'}^2 + vv'' + u = 0.$$

2. Die Gleichung

$$uu' = u''u^{(3)}$$

kann durch die Gleichung zweiter Ordnung

$$v{v'}^3 + v^2v'v'' = u$$

ersetzt werden.

[7] Unter der Annahme $u'(x^0) \neq 0$ lässt sich die Abbildung $x \mapsto u(x)$ nahe x^0 lokal umkehren, d. h. u kann als neue unabhängige Variable eingeführt werden.

12.4.4 Eulergleichungen

Definition 12.62. Eine Differenzialgleichung auf $(0, \infty)$, die invariant unter den Transformationen $x \mapsto \frac{x}{k}$ mit $k > 0$ ist, heißt eine *Eulergleichung* (im Englischen: *equidimensional*).

Der Variablenwechsel $\mathbb{R} \to (0, \infty)$, $t \mapsto x = \mathrm{e}^t$ mit

$$x \frac{d}{dx} = \frac{d}{dt},$$
$$x^2 \frac{d}{dx^2} = \frac{d^2}{dt^2} - \frac{d}{dt},$$

usw. überführt eine solche Gleichung in eine autonome Gleichung.

Beispiel 12.63. Die Gleichung $u'' = \frac{uu'}{x}$ ist eine Eulergleichung. Indem wir $x^2 u'' = u(xu')$ schreiben, erhalten wir nach dem Variablenwechsel $x = \mathrm{e}^t$

$$u''(t) - u'(t) = u(t)u'(t).$$

Die Substitution $v(u) = u'(t)$ reduziert diese Gleichung zu

$$vv' - v = uv.$$

Es folgt $v = 0$, also $u(x) = k_0$, oder $v'(u) = u + 1$, also $v(u) = \frac{u^2}{2} + u + k_1$. Aus den Beziehungen $v(u) = u(t)$ und $x = \mathrm{e}^t$ sehen wir schließlich, dass wir die separable Gleichung

$$u'(t) = \frac{u^2(t)}{2} + u(t) + k_1$$

bezüglich $u(x)$ lösen müssen.

Eine *lineare Eulergleichung* ist von der Form

$$a_0 \left(x \frac{d}{dx} \right)^m u + a_1 \left(x \frac{d}{dx} \right)^{m-1} u + \cdots + a_{m-1} \left(x \frac{d}{dx} \right) u + a_m u = 0,$$

wobei $a_0, a_1, \ldots, a_m \in \mathbb{C}$, $a_0 \neq 0$.
Aus der Theorie linearer Gleichungen mit konstanten Koeffizienten erhalten wir:

Satz 12.64. *Es sei*

$$a_0 \mu^m + a_1 \mu^{m-1} + \cdots + a_{m-1} \mu + a_m = \prod_{j=1}^{r} (\mu - \mu_j)^{m_j}$$

mit $\mu_j \neq \mu_{j'}$ für $j \neq j'$. Dann ist

$$x^{\mu_1}, x^{\mu_1} \log x, \ldots, x^{\mu_1} (\log x)^{m_1 - 1}, \ldots, x^{\mu_r}, x^{\mu_r} \log x, \ldots, x^{\mu_r} (\log x)^{m_r - 1}$$

ein Fundamentalsystem für obige lineare Differenzialgleichung.

Beispiel 12.65. Die Gleichung $u'' + \frac{u}{4x^2} = 0$ führt auf die Gleichung $(\mu - \frac{1}{2})^2 = 0$. Die allgemeine Lösung ist deshalb

$$u(x) = k_1 \sqrt{x} + k_2 \sqrt{x} \log x.$$

12.5 Randwert- und Eigenwertprobleme

12.5.1 Grundlegende Eigenschaften

Wir beginnen diesen Abschnitt mit einem Beispiel.
Das Randwertproblem

$$\begin{cases} u'' + \mu u = 0, \\ u(0) = u(\pi) = 0, \end{cases} \tag{12.66}$$

wobei $\mu \in \mathbb{C}$ ein Parameter ist, kann eine von Null verschiedene Lösung besitzen. Tatsächlich ist die allgemeine Lösung der linearen Gleichung $u'' + \mu u = 0$ durch

$$u(x) = k_1 \sin(\sqrt{\mu}x) + k_2 \cos(\sqrt{\mu}x)$$

gegeben.[8] Die Randbedingung $u(0) = 0$ erzwingt $k_2 = 0$, die Randbedingung $u(\pi) = 0$ zusammen mit $k_1 \neq 0$ impliziert $\mu = k^2$ für ein $k \in \mathbb{N}$.[9] Die Werte k^2 für $k \in \mathbb{N}$ heißen die *Eigenwerte* des Randwertproblems (12.66), die Funktionen $u_k(x) = \sin(kx)$ heißen die zugehörigen *Eigenfunktionen.*

Bemerkung. In Kapitel 6 hatten wir 2π-periodische, ungerade Funktionen nach diesen Eigenfunktionen $\sin(kx)$ für $k \in \mathbb{N}$ in *Fourierreihen* entwickelt.

Wir wollen im Weiteren die Frage nach Existenz von *Eigenwerten* allgemein aufgreifen. Dazu betrachten wir ein $N \times N$-System erster Ordnung der Form

$$\begin{cases} U' = A(x)\, U + \mu B(x)\, U, \quad x^0 \leq x \leq x^1, \\ M_0 U(x^0) + M_1 U(x^1) = 0, \end{cases} \tag{12.67}$$

wobei $A, B \in C([x^0, x^1]; M(N; \mathbb{C}))$, $M_0, M_1 \in M(N, \mathbb{C})$ und $x^0 < x^1$. $\mu \in \mathbb{C}$ ein sogenannter *Spektralparameter* ist.
Wir machen weiterhin die Annahme, dass die lineare Abbildung

$$\mathbb{C}^{2N} \to \mathbb{C}^N, \quad \big(U^0, U^1\big) \mapsto M_0 U^0 + M_1 U^1 \tag{12.68}$$

vollen Rang N hat, d. h. surjektiv ist.

Definition 12.69. $\mu^0 \in \mathbb{C}$ heißt ein *Eigenwert* von (12.67), falls es eine Lösung $U = U(x)$ zu (12.67) mit $\mu = \mu^0$ gibt, die nicht identisch Null ist. Ein solches U heißt dann eine *Eigenfunktion* zum Eigenwert μ^0. [10]

Als erstes zeigen wir, dass die Fundamentalmatrix $Y(x, x^0; \mu)$ von (12.67) holomorph vom Parameter μ abhängt.[11]

---

[8] Hier ist $\sqrt{\mu}$ eine der beiden komplexen Zahlen mit $(\sqrt{\mu})^2 = \mu$.

[9] $k = 0$ ergibt die Nulllösung, $-k$ ergibt dasselbe μ wie k.

[10] Wie in der linearen Algebra können Eigenwerte geometrische und algebraische Multiplizitäten größer 1 haben.

[11] Für uns bedeutet *Holomorphie*, dass wir die Fundamentalmatrix in einer konvergenten Potenzreihe bezüglich μ mit Konvergenzradius ∞ entwickeln können. In der Vorlesung „Funktionentheorie“ werden Sie lernen, dass dies gerade die auf *ganz* $\mathbb{C}$ *komplex-differenzierbaren Funktionen*, sogenannte *ganze Funktionen* charakterisiert.

Satz 12.70. *Die Fundamentalmatrix $Y(x,x^0;\mu)$ hat die Form*

$$\sum_{k=0}^{\infty} \mu^k Y_k(x,x^0),$$

wobei $X(x,x^0) = Y_0(x,x^0;0)$ die Fundamentalmatrix des Systems $U' = A(x)U$ ist und die Reihe für alle $\mu \in \mathbb{C}$ gleichmäßig bezüglich x in kompakten Teilintervallen von I absolut konvergiert.

Beweis. Eine formale Manipulation der Gleichungen $Y = \sum\limits_{k=0}^{\infty} \mu^k Y_k$ und $Y' = AY + \mu BY$ ergibt wegen $Y' = \sum\limits_{k=0}^{\infty} \mu^k Y_k'$ die Beziehungen

$$\begin{cases} Y_0' = AY_0, \quad Y_0(x^0) = I, \\ Y_k' = AY_k + BY_{k-1}, \quad Y_k(x^0) = 0 \end{cases} \tag{12.71}$$

für $k \geq 1$.
Das System (12.71) ist eindeutig lösbar, und zwar erhalten wir sukzessiv, dass $Y_0 = X$ und

$$Y_k(x) = \int_{x^0}^{x} X(x,x')B(x')Y_{k-1}(x')\,dx'$$

für $k \geq 1$ gilt.
Sei nun $J \subseteq I$ ein kompaktes Intervall. Sei ferner $a = \max\limits_{x\in J} |A(x)|$ und $b = \max\limits_{x\in J} |B(x)|$. Dann gilt

$$|Y_k(x)| \leq \frac{b^k |x - x_0|^k}{k!}\, \mathrm{e}^{a|x-x^0|}, \quad x \in J, \tag{12.72}$$

für alle $k \in \mathbb{N}$.

Beweis von (12.72)**.** Mittels Induktion nach k.

***k* = 0.** Es gilt $|X(x)| = \max\limits_{|U^0|\leq 1} \left|X(x)U^0\right|$, also müssen wir $|U(x)| \leq \mathrm{e}^{a|x-x^0|} \left|U^0\right|$ für $x \in J$ und jede Lösung $U = U(x)$ zu

$$U' = A(x)U, \quad U(x^0) = U^0$$

zeigen. Für eine solche Lösung gilt

$$\begin{aligned} \frac{d}{dx}\,|U(x)|^2 &= 2\,\mathrm{Re}\left\langle U'(x), \overline{U(x)}\right\rangle \\ &= 2\,\mathrm{Re}\left\langle A(x)U(x), \overline{U(x)}\right\rangle \leq 2a\,|U(x)|^2, \end{aligned}$$

also $|U(x)|^2 \leq \left|U^0\right|^2 \mathrm{e}^{2a|x-x^0|}$ und $|U(x)| \leq \mathrm{e}^{a|x-x^0|} \left|U^0\right|$.

$\boldsymbol{k \longrightarrow k+1}$. Es gilt

$$\begin{aligned}
|Y_{k+1}(x)| &\leq \int_{x^0}^{x} |X(x,x')|\,|B(x')|\,|Y_k(x')|\ |dx'| \\
&\leq \frac{b^{k+1}}{k!} \int_{x^0}^{x} \mathrm{e}^{a|x-x'|}|x'-x^0|^k \mathrm{e}^{a|x'-x^0|}\ |dx'| \\
&= \frac{b^{k+1}}{k!}\,\mathrm{e}^{a|x-x^0|} \int_{x^0}^{x} |x'-x^0|^k\ |dx'| \\
&= \frac{b^{k+1}}{k!}\,\mathrm{e}^{a|x-x^0|}\,\frac{|x-x^0|^{k+1}}{k+1} = \frac{b^{k+1}|x-x^0|^{k+1}}{(k+1)!}\,\mathrm{e}^{a|x-x^0|}.
\end{aligned}$$

Aus (12.72) folgt die absolute Konvergenz der Reihe $\sum\limits_{k=0}^{\infty} \mu^k Y_k(x,x^0)$ gleichmäßig auf J, nämlich

$$\begin{aligned}
\sum_{k=0}^{\infty} \left|\mu^k Y_k(x)\right| &\leq \sum_{k=0}^{\infty} |\mu|^k\,\frac{b^k|x-x^0|^k}{k!}\,\mathrm{e}^{a|x-x^0|} \\
&= \mathrm{e}^{(a+|\mu|b)|x-x^0|} < \infty.
\end{aligned}$$

□

Satz 12.73. $\mu^0 \in \mathbb{C}$ *ist dann und nur dann ein Eigenwert von* (12.67), *wenn*

$$\Delta(\mu) = \det\left(M_0 + M_1 Y(x^1,x^0;\mu)\right)$$

an der Stelle μ^0 *verschwindet.*

Beweis. Sei U eine Lösung zu (12.67) mit $\mu = \mu^0$, die nicht identisch verschwindet. Sei $U^0 = U(x^0)$, $U^1 = U(x^1)$. Dann ist $U^1 = Y(x^1,x^0;\mu^0)U^0$, also

$$\left(M_0 + M_1 Y(x^1,x^0;\mu^0)\right) U^0 = 0. \tag{12.74}$$

Somit ist $U^0 \neq 0$ Eigenvektor zum Eigenwert 0, folglich ist $\Delta(\mu^0) = 0$.

Gilt umgekehrt $\Delta(\mu^0) = 0$, so finden wir ein $U^0 \neq 0$ mit (12.74). Die gesuchte Lösung U von (12.67) ist dann $U(x) = Y(x,x^0;\mu^0)U^0$. □

Wir haben die folgende Folgerung:

Satz 12.75. *Sei* $\Delta(\mu) \not\equiv 0$. *Dann sind die Eigenwerte von* (12.67) *isoliert. Insbesondere können sich die Eigenwerte von* (12.67) *nicht im Endlichen häufen.*

Beweis. $\Delta(\mu)$ ist eine ganze holomorphe Funktion, d. h. $\Delta(\mu)$ ist als eine in ganz $\mathbb{C}$ konvergente Potenzreihe darstellbar. Dann folgt das Ergebnis aus dem Eindeutigkeitssatz 6.2.4 für Potenzreihen, der sinngemäß auch im Komplexen gilt. □

12.5.2 Das adjungierte Randwertproblem

Wir betrachten weiterhin das Problem (12.67). Gleichzeitig werden wir ein sogenanntes adjungiertes Randwertproblem betrachten.
Wir arbeiten mit der Sesquilinearform

$$\big(U, V\big) = \langle U, \overline{V} \rangle, \quad U, V \in \mathbb{C}^N.$$

Die adjungierte Matrix D^* zu einer Matrix $D \in M(N; \mathbb{C})$ ist dann vermöge

$$\big(DU, V\big) = \big(U, D^* V\big), \quad U, V \in \mathbb{C}^N,$$

erklärt.

Definition 12.76. Das zu (12.67) *adjungierte Randwertproblem* ist von der Form

$$\begin{cases} V' = -\left(A^*(x) + \mu' B^*(x)\right) V, & x^0 \le x \le x^1, \\ N_0 V(x^0) + N_1 V(x^1) = 0 \end{cases} \tag{12.77}$$

mit weiter unten zu spezifizierenden Matrizen N_0, $N_1 \in M(N; \mathbb{C})$. Hierbei ist $\mu' \in \mathbb{C}$ der Spektralparameter.

Um die Form der *adjungierten Randbedingungen* zu bestimmen, machen wir eine Reihe von Beoachtungen:

Lemma 12.78. *Die Fundamentalmatrix des adjungierten Problems* (12.77) *ist* $Y^*(x^0, x; \overline{\mu'})$.

Beweis. Dies folgt aus $Y^*(x^0, x^0; \overline{\mu'}) = I^* = I$ und

$$\begin{aligned} \frac{\partial}{\partial x} Y^*(x^0, x; \overline{\mu'}) &= \left(\frac{\partial}{\partial x} Y(x^0, x; \overline{\mu'})\right)^* \\ &= \big(-Y(x^0, x; \overline{\mu'})(A(x) + \overline{\mu'} B(x))\big)^* \\ &= -\left(A^*(x) + \mu' B^*(x)\right) Y^*(x^0, x; \overline{\mu'}, \end{aligned}$$

s. die Bemerkung 12.3.1. □

Lemma 12.79. *Ist* U *eine Lösung des Problems* (12.67) *und* V *eine Lösung des Problems* (12.77) *mit* $\mu' = \overline{\mu}$ *(beides ohne Randbedingungen), so ist die Größe*

$$\big(U(x), V(x)\big), \quad x^0 \le x \le x^1,$$

unabhängig von x.

Beweis. Mit $U^0 = U(x^0)$, $V^0 = V(x^0)$ gilt für $x^0 \le x \le x^1$

$$\begin{aligned} \big(U(x), V(x)\big) &= \big(Y(x, x^0; \mu) U^0, Y^*(x^0, x; \mu) V^0\big) \\ &= \big(Y(x^0, x; \mu) Y(x, x^0; \mu) U^0, V^0\big) = \big(U^0, V^0\big). \end{aligned}$$

□

Die adjungierte Randbedingung ist durch die Bedingung charakterisiert, dass die Größe $\big(U(x), V(x)\big)$ identisch Null ist. Dazu sei

$$\beta\colon \mathbb{C}^{2N} \times \mathbb{C}^{2N} \to \mathbb{C}, \qquad \left(\begin{pmatrix} U^0 \\ U^1 \end{pmatrix}, \begin{pmatrix} V^0 \\ V^1 \end{pmatrix}\right) \mapsto \big(U^1, V^1\big) - \big(U^0, V^0\big)$$

eine *nichtentartete* Sesquilinearform auf $\mathbb{C}^{2N} \cong \mathbb{C}^N \times \mathbb{C}^N$ und

$$\Pi = \big\{(U^0, U^1) \in \mathbb{C}^{2N} \;\big|\; M_0 U^0 + M_1 U^1 = 0\big\}.$$

ein linearer Unterraum von $\mathbb{C}^{2N}$. Laut der Annahme in (12.68) ist

$$\dim \Pi = N.$$

Definition 12.80. Die *adjungierte Randbedingung* in (12.77) ist durch

$$\Pi^\perp = \big\{(V^0, V^1) \in \mathbb{C}^{2N} \;\big|\; N_0 V^0 + N_1 V^1 = 0\big\}$$

gegeben, wobei $\Pi^\perp$ der zu Π orthogonale Unterraum bezüglich β ist.

Wir bemerken, dass $\dim \Pi^\perp = N$.

Lemma 12.81. *Die im vorigen Satz angegebene Bedingung bestimmt die Matrizen N_0, $N_1 \in M(N;\mathbb{C})$ bis auf Linksmultiplikation mit einem Faktor $D \in$ $\mathrm{Gl}(N;\mathbb{C})$.*

Beweis. $\Pi^\perp$ ist der Kern der linearen Abbildung $\mathbb{C}^{2N} \to \mathbb{C}^N$, $(V^0, V^1) \mapsto N_0 V^0 + N_1 V^1$, die wir mit (N_0, N_1) bezeichnen wollen. Diese Abbildung induziert eine lineare Isomorphie $\widetilde{(N_0, N_1)}\colon \mathbb{C}^{2N}/\Pi^\perp \xrightarrow{\cong} \mathbb{C}^N$. Für jedes weitere Paar von Matrizen N_0', $N_1' \in M(N;\mathbb{C})$ mit derselben Eigenschaft erhalten wir analog eine lineare Isomorphie $\widetilde{(N_0', N_1')}\colon \mathbb{C}^{2N}/\Pi^\perp \xrightarrow{\cong} \mathbb{C}^N$. Wir setzen $D = \widetilde{(N_0', N_1')}\widetilde{(N_0, N_1)}^{-1}\colon \mathbb{C}^N \xrightarrow{\cong} \mathbb{C}^N$ und identifizieren D mit einer Matrix in $\mathrm{Gl}(N;\mathbb{C})$. Dann

$$(N_0', N_1') = D \circ (N_0, N_1) = (DN_0, DN_1)$$

als lineare Abbildungen $\mathbb{C}^{2N} \to \mathbb{C}^N$ bzw. $N_j' = DN_j$ für $j = 0, 1$. $\square$

Wir bemerken, dass die Multiplikation von N_1, N_1 mit $D \in \mathrm{Gl}(N;\mathbb{C})$ die Funktion $\Delta^*(\mu')$, nun für das Randwertproblem (12.77) definiert, nur um den konstanten Faktor $\det D \neq 0$ abändert. Insbesondere ändern sich die Nullstellen der Funktion $\Delta^*(\mu')$ dabei nicht.

Beispiel 12.82. Um die adjungierte Randbedingung in einem *Spezialfall* tatsächlich zu bestimmen, nehmen wir an, dass für ein $k \in \{0, 1, \ldots, N\}$ (möglicherweise nach Multiplikation mit einer Matrix $D \in \mathrm{Gl}(N;\mathbb{C})$)

$$M_0 = \begin{pmatrix} M_0^+ & 0 \\ M_0^- & I_{N-k} \end{pmatrix}, \quad M_1 = \begin{pmatrix} I_k & M_1^+ \\ 0 & M_1^- \end{pmatrix}$$

mit $M_0^+ \in M(k;\mathbb{C})$, $M_0^- \in M(N-k,k;\mathbb{C})$, $M_1^+ \in M(k,N_k;\mathbb{C})$ und $M_1^- \in M(N_k;\mathbb{C})$ gilt. (Was bedeutet diese Annahme geometrisch?) Entsprechend zerlegen wir $U^0 = \begin{pmatrix} U_+^0 \\ U_-^0 \end{pmatrix}$ mit $U_+^0 \in \mathbb{C}^k$, $U_-^0 \in \mathbb{C}^{N-k}$, usw. Die Bedingung $M_0U^0 + M_1U^1 = 0$ schreibt sich dann als

$$\begin{aligned} U_-^0 &= -M_0^- U_+^0 - M_1^- U_-^1, \\ U_+^1 &= -M_0^+ U_+^0 - M_1^+ U_-^1. \end{aligned}$$

Lemma 12.83. *Unter diesen Voraussetzungen ist*

$$N_0 = \begin{pmatrix} I_k & -(M_0^-)^* \\ 0 & (M_1^-)^* \end{pmatrix}, \quad N_1 = \begin{pmatrix} (M_0^+)^* & 0 \\ -(M_1^+)^* & I_{N-k} \end{pmatrix}$$

eine mögliche Wahl für N_0, N_1.

Beweis. Die Bedingung, dass $(U^0, V^0) = (U^1, V^1)$ für $(U^0, U^1) \in \Pi$, $(V^0, V^1) \in \Pi^\perp$ ist, schreibt sich als

$$\big(U_+^0, V_+^0\big) - \big(M_0^- U_+^0 + M_1^- U_-^1, V_-^0\big) = -\big(M_0^+ U_+^0 + M_1^+ U_-^1, V_+^1\big) + \big(U_-^1, V_-^1\big).$$

Da die $U_+^0 \in \mathbb{C}^k$, $U_-^1 \in \mathbb{C}^{N-k}$ beliebig sind, folgt daraus, dass

$$\begin{aligned} V_+^0 &= (M_0^-)^* V_-^0 - (M_0^+)^* V_+^1, \\ V_-^1 &= -(M_1^-)^* V_-^0 + (M_1^+)^* V_+^1. \end{aligned}$$

Das ist gerade die Behauptung. □

Die Bedeutung des adjungierten Problems wird mit folgendem Resultat klar:

Satz 12.84. $\mu \in \mathbb{C}$ *ist genau dann Eigenwert von* (12.67)*, wenn* $\bar{\mu}$ *Eigenwert von* (12.77) *ist.*

Beweis. Sei

$$\Sigma = \big\{(U^0, U^1) \in \mathbb{C}^{2N} \;\big|\; U^1 = Y(x^1, x^0; \mu)U^0\big\}$$

der Raum der möglichen Randdaten für das Problem (12.67) und

$$\Sigma^* = \big\{(V^0, V^1) \in \mathbb{C}^{2N} \;\big|\; V^1 = Y^*(x^0, x^1; \mu)V^0\big\}$$

der entsprechende Raum für das Problem (12.77) (mit $\mu' = \overline{\mu}$). Wir bemerken, dass $\dim \Sigma = \dim \Sigma^* = N$. Lemma 12.79 besagt gerade, dass $\Sigma^* = \Sigma^\perp$.
Nun ist μ genau dann ein Eigenwert von (12.67), wenn $\Sigma \cap \Pi \neq \{0\}$, und das gilt genau dann, wenn $\Sigma^* \cap \Pi^\perp \neq \{0\}$, d. h. wenn $\overline{\mu}$ ein Eigenwert von (12.77) ist. □

Gleichzeitig zeigt dieser Beweis, dass die *geometrische Vielfachheit* $\dim(\Sigma \cap \Pi) = \dim\big(\Sigma^* \cap \Pi^\perp\big)$ beider Eigenwerte gleich ist.

12.5.3 Existenz und Eindeutigkeit von Lösungen

Wir betrachten jetzt das *inhomogene Randwertproblem*

$$\begin{cases} U' = (A(x) + \mu B(x))\, U + F(x), \quad x^0 \le x \le x^1, \\ M_0 U(x^0) + M_1 U(x^1) = 0, \end{cases} \tag{12.85}$$

wobei die Annahmen an A, B, M_0, M_1, μ wie im vorigen Abschnitt sind und $F \in C([x^0, x^1]; \mathbb{C}^N)$.

Satz 12.86. (a) *Das inhomogene Randwertproblem* (12.85) *besitzt genau dann eine Lösung, wenn für jede Lösung V des* (*homogenen*) *adjungierten Randwertproblems* (12.77) *mit $\mu' = \overline{\mu}$*

$$\int_{x^0}^{x^1} \big(F(x), V(x)\big)\, dx = 0$$

gilt.
(b) *Insbesondere ist das inhomogene Randwertproblem* (12.85) *genau dann eindeutig lösbar, wenn das adjungierte Randwertproblem* (12.77) *nur die Nulllösung zuläßt.*

Beweis. Sei U eine Lösung zu (12.85), V eine Lösung zu (12.77) mit $\mu' = \overline{\mu}$. Dann gilt

$$\begin{aligned} \int_{x^0}^{x^1} \big(F, V\big)\, dx &= \int_{x^0}^{x^1} \big(U' - \big(A + \mu B\big), V\big)\, dx \\ &= \big(U, V\big)\,\Big|_{x^0}^{x^1} - \int_{x^0}^{x^1} \big(U, V' + \big(A^* + \overline{\mu} B^*\big)V\big)\, dx = 0. \end{aligned}$$

Die umgekehrte Richtung folgt aus Dimensionsgründen. □

Bemerkung. Ein analoges Result gilt im Fall des Problems (12.85) mit einer *inhomogenen Randbedingungen* der Form

$$M_0 U(x^0) + M_1 U(x^1) = U^2,$$

wobei $U^2 \in \mathbb{C}^N$, nur dass sich jetzt die Lösbarkeitsbedingung in (a) etwa im Fall des Beispiels 12.82 als

$$\big(U^2_-, V(x^0)_-\big)_{\mathbb{C}^{N-k}} + \int_{x^0}^{x^1} \big(F(x), V(x)\big)_{\mathbb{C}^N}\, dx = \big(U^2_+, V(x^1)_+\big)_{\mathbb{C}^k}$$

für alle V wie zuvor schreibt. Hierbei bezeichnet V_+ für $V \in \mathbb{C}^N$ den Vektor aus den ersten k Komponenten und V_- den Vektor aus den letzen $N - k$ Komponenten von V.

12.5.4 Eigenwertasymptotik

Schließlich interessieren wir uns für die asymptotische Verteilung der Eigenwerte. Wir betrachten dazu das *Eigenwertproblem*

$$\begin{cases} \dfrac{\partial^2 u}{\partial x^2} = \big(-\nu^2 f(x) + g(x)\big)\, u, \quad x^0 \le x \le x^1, \\ u(x^0) = u(x^1) = 0 \end{cases}$$

in einem endlichen Intervall $[x^0, x^1]$. Die Voraussetzungen an f, g sind wie zuvor, zusätzlich ist g reellwertig.
Wir wissen bereits, dass sich die allgemeine reelle Lösung als

$$u(x,\nu) = A(\nu) f^{-1/4}(x) \left(\sin\left(\nu \int_{x^0}^{x} f^{1/2}(x')\, dx' + \delta(\nu) \right) + \epsilon(x,\nu) \right)$$

mit $A(\nu)$, $\delta(\nu)$ unabhängig von x und

$$|\epsilon(x,\nu)| \le \exp\left(\frac{V_{x^0,x}(F)}{\nu} \right) - 1$$

schreibt.
Sei nun ν ein Eigenwert und $u(x,\nu)$ eine zugehörige Eigenfunktion. Die Randbedingung $u(x^0) = 0$ erzwingt $\sin\delta(\nu) = 0$ wegen $\epsilon(x^0,\nu) = 0$. Daher können wir o. E. d. A. $\delta(\nu) = 0$ annehmen. (Die Wahl von $\delta(\nu) = k\pi$ mit $k \in \mathbb{Z}$ führt zu einer Multiplikation der Lösung mit dem Faktor $(-1)^k$.) Die Randbedingung $u(x^1) = 0$ impliziert dann, dass

$$\sin(D\nu) + \epsilon(x^1,\nu) = 0, \tag{12.87}$$

wobei

$$D = \int_{x^0}^{x^1} f^{1/2}(x)\, dx.$$

Satz 12.88. *Für ν hinreichend groß sind die Eigenwerte durch die asymptotische Beziehung*

$$\nu = \frac{n\pi}{D} + O(n^{-1}), \quad n \in \mathbb{N},$$

für $n \to \infty$ gegeben.

Beweis. Für ν hinreichend groß gilt $\big|\epsilon(x^1,\nu)\big| < 1$. Folglich sind für solche ν die Lösungen zu (12.87) durch

$$D\nu = \pi n + (-1)^{n-1} \arcsin \epsilon(x^1,\nu), \quad n \in \mathbb{N},$$

gegeben. Es bleibt zu bemerken, dass

$$\arcsin \epsilon(x^1,\nu) = O\big(\epsilon(x^1,\nu)\big) = O(\nu^{-1}) = O(n^{-1})$$

für $n \to \infty$, wobei die Beziehung $O(\nu^{-1}) = O(n^{-1})$ aus $\nu \sim \dfrac{\pi n}{D}$ für $n \to \infty$ folgt (was sich wegen $\big|(-1)^{n-1} \arcsin \epsilon(x^1,\nu)\big| < \pi/2$ direkt aus der vorigen Gleichung ergibt). □

Der $O(1/n)$-Term lässt sich noch genauer abschätzen. Wir benötigen dazu das folgende elementare Resultat:

Lemma 12.89 (Jordan-Ungleichung)**.** *Für* $0 \leq \theta \leq \frac{\pi}{2}$ *gilt*

$$\sin\theta \geq \frac{2\theta}{\pi}.$$

Satz 12.90. *Sei* $n \in \mathbb{N}$, $n > 1/2 + Dd/(\pi \log 2)$, *wobei* $d = V_{x^0,x^1}(F)$. *Dann gibt es einen Eigenwert der Form*

$$\nu = \frac{(n+\varrho)\pi}{D},$$

wobei $|\varrho| < 1/2$. *Tatsächlich gilt*

$$|\varrho| \leq \frac{1}{2}\exp\left(\frac{Dd}{(n-1/2)\pi}\right) - \frac{1}{2}.$$

Die zugehörige Eigenfunktion hat die Form

$$f^{-1/4}(x)\left(\sin\left(\frac{(n+\varrho)\pi}{D}\int_{x^0}^{x} f^{1/2}(x')\,dx'\right) + \epsilon_n(x)\right),$$

wobei

$$|\epsilon_n(x)| \leq \exp\left(\frac{D\,V_{x^0,x}(F)}{(n-1/2)\pi}\right) - 1.$$

Beweis. Es gilt

$$\left|\epsilon(x^1,\nu)\right| \leq \mathrm{e}^{d/\nu} - 1.$$

Folglich gilt für $\nu > d/\log 2$ mit der Jordan-Ungleichung

$$|\theta(\nu)| \leq \frac{\pi}{2}\left(\mathrm{e}^{d/\nu} - 1\right), \tag{12.91}$$

wobei $\theta(\nu) = \arcsin\epsilon(x^1,\nu)$. Die Gleichung für die Eigenwerte lautet

$$\xi(\nu) = D\nu - n\pi + (-1)^n\theta(\nu) = 0.$$

Sei nun $n > 1/2 + Dd/(\pi\log 2)$, d. h. $(n-1/2)\pi/D > d/\log 2$. Dann gilt

$$\xi\left(\left(n-\frac{1}{2}\right)\frac{\pi}{D}\right) = -\frac{\pi}{2} + (-1)^n\theta(\nu) < 0,$$
$$\xi\left(\left(n+\frac{1}{2}\right)\frac{\pi}{D}\right) = \frac{\pi}{2} + (-1)^n\theta(\nu) > 0.$$

Da $\theta = \theta(\nu)$ eine stetige Funktion ist [12], gibt es wenigstens einen Eigenwert ν mit

$$\left(n-\frac{1}{2}\right)\frac{\pi}{D} < \nu < \left(n+\frac{1}{2}\right)\frac{\pi}{D}.$$

[12] Strenggenommen müssten wir analog dem Satz 12.15 noch zeigen, dass unter geeigneten Annahmen Lösungen zu gewöhnlichen Differnzialgleichungen stetig von Parametern abhängen.

Wir schreiben dieses ν als

$$\nu = \frac{(n+\varrho)\pi}{D}$$

mit $|\varrho| < 1/2$. Es folgt

$$\varrho\pi = (-1)^{n-1}\theta\left(\frac{(n+\varrho)\pi}{D}\right),$$

also wegen (12.91)

$$|\varrho| \leq \frac{1}{2}\exp\left(\frac{Dd}{(n-1/2)\pi}\right) - \frac{1}{2}.$$

□

Bemerkung. Eine verfeinerte Analysis zeigt, dass die so konstruierten Eigenwerte für hinreichend großes n tatsächlich *einfach* sind.