

Grundkurs Mathematik

LEHRBUCH

Otto Forster

Analysis 2

Differentialrechnung im \mathbb{R}^n ,
gewöhnliche
Differentialgleichungen

10. Auflage



Springer Spektrum

Grundkurs Mathematik

Berater

Prof. Dr. Martin Aigner,
Prof. Dr. Peter Gritzmann,
Prof. Dr. Volker Mehrmann,
Prof. Dr. Gisbert Wüstholtz

Die Reihe „Grundkurs Mathematik“ ist die bekannte Lehrbuchreihe im handlichen kleinen Taschenbuch-Format passend zu den mathematischen Grundvorlesungen, vorwiegend im ersten Studienjahr. Die Bücher sind didaktisch gut aufbereitet, kompakt geschrieben und enthalten viele Beispiele und Übungsaufgaben.

In der Reihe werden Lehr- und Übungsbücher veröffentlicht, die bei der Klausurvorbereitung unterstützen. Zielgruppe sind Studierende der Mathematik aller Studiengänge, Studierende der Informatik, Naturwissenschaften und Technik, sowie interessierte Schülerinnen und Schüler der Sekundarstufe II.

Die Reihe existiert seit 1975 und enthält die klassischen Bestseller von Otto Forster und Gerd Fischer zur Analysis und Linearen Algebra in aktualisierter Neuauflage.

Otto Forster

Analysis 2

Differentialrechnung im \mathbb{R}^n ,
gewöhnliche Differentialgleichungen

10., verbesserte Auflage



Springer Spektrum

Prof. Dr. Otto Forster
Ludwig-Maximilians-Universität
München, Deutschland
forster@mathematik.uni-muenchen.de

ISBN 978-3-658-02356-0

ISBN 978-3-658-02357-7 (eBook)

DOI 10.1007/978-3-658-02357-7

Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über <http://dnb.d-nb.de> abrufbar.

Springer Spektrum

© Springer Fachmedien Wiesbaden 1976 ... 2005, 2006, 2008, 2011, 2013

Das Werk einschließlich aller seiner Teile ist urheberrechtlich geschützt. Jede Verwertung, die nicht ausdrücklich vom Urheberrechtsgesetz zugelassen ist, bedarf der vorherigen Zustimmung des Verlags. Das gilt insbesondere für Vervielfältigungen, Bearbeitungen, Übersetzungen, Mikroverfilmungen und die Einspeicherung und Verarbeitung in elektronischen Systemen.

Die Wiedergabe von Gebrauchsnamen, Handelsnamen, Warenbezeichnungen usw. in diesem Werk berechtigt auch ohne besondere Kennzeichnung nicht zu der Annahme, dass solche Namen im Sinne der Warenzeichen- und Markenschutz-Gesetzgebung als frei zu betrachten wären und daher von jedermann benutzt werden dürften.

Planung und Lektorat: Ulrike Schmickler-Hirzebruch | Barbara Gerlach

Gedruckt auf säurefreiem und chlorfrei gebleichtem Papier

Springer Spektrum ist eine Marke von Springer DE.

Springer DE ist Teil der Fachverlagsgruppe Springer Science+Business Media

www.springer-spektrum.de

Vorwort zur ersten Auflage

Der vorliegende Band stellt den zweiten Teil eines Analysis-Kurses für Studenten der Mathematik und Physik dar.

Das erste Kapitel befaßt sich mit der Differentialrechnung von Funktionen mehrerer reeller Veränderlichen. Nach einer Einführung in die topologischen Grundbegriffe werden Kurven im \mathbb{R}^n , partielle Ableitungen, totale Differenzierbarkeit, Taylorsche Formel, Maxima und Minima, implizite Funktionen und parameterabhängige Integrale behandelt.

Das zweite Kapitel gibt eine kurze Einführung in die Theorie der gewöhnlichen Differentialgleichungen. Nach dem Beweis des Existenz- und Eindeutigkeitssatzes und der Besprechung der Methode der Trennung der Variablen wird besonders auf die Theorie der linearen Differentialgleichungen eingegangen.

Wie im ersten Band wurde versucht, allzu große Abstraktionen zu vermeiden und die allgemeine Theorie durch viele konkrete Beispiele zu erläutern, insbesondere solche, die für die Physik relevant sind.

Bei der Bemessung des Stoffumfangs wurde berücksichtigt, daß die Analysis 2 meist im Sommersemester gelesen wird, in dem weniger Zeit zur Verfügung steht als in einem Wintersemester. Wegen der Kürze des Sommersemesters ist nach meiner Meinung eine befriedigende Behandlung der mehrdimensionalen Integration im 2. Semester nicht möglich, die besser dem 3. Semester vorbehalten bleibt.

Dies Buch ist entstanden aus der Ausarbeitung einer Vorlesung, die ich im Sommersemester 1971 an der Universität Regensburg gehalten habe. Die damalige Vorlesungs-Ausarbeitung wurde von Herrn R. Schimpl angefertigt, dem ich hierfür meinen Dank sage.

Münster, Januar 1977

O. Forster

Vorwort zur 6. Auflage

Nachdem der erste Band der Analysis vor einigen Jahren eine gründliche Überarbeitung erfahren hat, wurde nun auch der zweite Band einer Neubearbeitung unterzogen. Einerseits erhielt der Text durch \TeX -Satz eine schönere äußere Form, was auch künftige Änderungen erleichtert. Zum anderen wurde das Buch auch inhaltlich überarbeitet. Neben kleineren Veränderungen im Text wurde im ersten Teil der Paragraph über implizite Funktionen durch einen Paragraphen über differenzierbare Untermannigfaltigkeiten des \mathbb{R}^n ergänzt. Der zweite Teil über gewöhnliche Differentialgleichungen beginnt nun nicht mehr mit dem allgemeinen Existenz- und Eindeigkeitssatz, sondern es werden zuerst zur Motivation verschiedene elementar lösbare Differentialgleichungen behandelt. Vor die allgemeine Lösungstheorie linearer Differentialgleichungen mit konstanten Koeffizienten wurde ein eigener Paragraph mit einfachen (linearen und nicht-linearen) Differentialgleichungen 2. Ordnung eingefügt, die für die Physik relevant sind.

München, März 2005

Otto Forster

Vorwort zur 10. Auflage

Für die vorliegende 10. Auflage wurden in §16 einige instruktive Bilder über die Lösungskurven von linearen Differentialgleichungen mit konstanten Koeffizienten im \mathbb{R}^2 hinzugefügt. Außerdem wurden Druckfehler korrigiert, die mir dankenswerterweise von vielen aufmerksamen Lesern gemeldet worden sind. Ich bin auch weiterhin allen Leserinnen und Lesern dankbar, die zur Aktualisierung der Errata-Liste beitragen (siehe Seite VIII).

München, März 2013

Otto Forster

Inhaltsverzeichnis

I. Differentialrechnung im \mathbb{R}^n	1
1 Topologie metrischer Räume	1
2 Grenzwerte. Stetigkeit	15
3 Kompaktheit	28
4 Kurven im \mathbb{R}^n	40
5 Partielle Ableitungen	51
6 Totale Differenzierbarkeit	66
7 Taylor-Formel. Lokale Extrema	77
8 Implizite Funktionen	90
9 Untermannigfaltigkeiten	104
10 Integrale, die von einem Parameter abhängen	118
 II. Gewöhnliche Differentialgleichungen	 135
11 Elementare Lösungsmethoden	135
12 Existenz- und Eindeigkeitssatz	149
13 Lineare Differentialgleichungen	165
14 Differentialgleichungen 2. Ordnung	179
15 Lineare Differentialgleichungen mit konstanten Koeffizienten	199
16 Systeme linearer Diff'gleichungen mit konstanten Koeffizienten	213
Literaturhinweise	225
Namens- und Sachverzeichnis	226
Symbolverzeichnis	226

Webseite

Für die Analysis 2 gibt es eine Webseite, die über die Homepage des Verfassers

`http://www.mathematik.uni-muenchen.de/~forster`

erreichbar ist. Dort ist jeweils eine aktuelle Liste der bekannt gewordenen *Errata* abgelegt.

Ich bin allen Leserinnen und Lesern dankbar, die mir per Email an

`forster@mathematik.uni-muenchen.de`

Fehlermeldungen oder sonstige Kommentare zusenden.

Otto Forster

Kapitel I

Differentialrechnung im \mathbb{R}^n

§ 1 Topologie metrischer Räume

Für unsere späteren Untersuchungen von Funktionen mehrerer Veränderlichen brauchen wir u.a. einige topologische Grundbegriffe im \mathbb{R}^n wie Umgebung, offene Menge, abgeschlossene Menge, Rand. Diese Begriffe können alle auf den Begriff des Abstands zurückgeführt werden. Wir betrachten daher gleich allgemeiner metrische Räume, das sind Mengen, auf denen ein gewissen Axiomen genügender Abstandsbe-griff gegeben ist.

Definition. Sei X eine Menge. Unter einer *Metrik* auf X versteht man eine Abbildung

$$d : X \times X \longrightarrow \mathbb{R}, \quad (x, y) \mapsto d(x, y)$$

mit folgenden Eigenschaften:

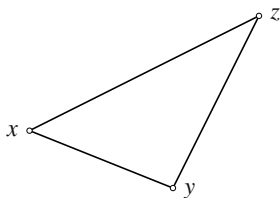
i) $d(x, y) = 0$ genau dann, wenn $x = y$.

ii) *Symmetrie:* Für alle $x, y \in X$ gilt

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

iii) *Dreiecksungleichung* (Bild 1.1): Für alle $x, y, z \in X$ gilt

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

**Bild 1.1**

Ein *metrischer Raum* ist ein Paar (X, d) , bestehend aus einer Menge X und einer Metrik d auf X . Man nennt $d(x, y)$ den *Abstand* oder die *Distanz* der Punkte x und y bzgl. der Metrik d .

Vereinbarung. Sind Missverständnisse ausgeschlossen, schreiben wir kurz X statt (X, d) und $\|x, y\|$ statt $d(x, y)$.

Bemerkung. Aus den Axiomen der Metrik folgt, dass

$$d(x, y) \geq 0 \quad \text{für alle } x, y \in X.$$

Beweis. Wendet man die Dreiecksungleichung auf die Punkte x, y, x an, so ergibt sich unter Verwendung von i) und ii)

$$0 = d(x, x) \leq d(x, y) + d(y, x) = 2d(x, y), \quad \text{q.e.d.}$$

Beispiele

(1.1) Die Menge \mathbb{R} der reellen Zahlen und die Menge \mathbb{C} der komplexen Zahlen werden zu metrischen Räumen, wenn man als Abstand definiert

$$d(x, y) := |x - y| \quad \text{für } x, y \in \mathbb{R} \quad (\text{bzw. } x, y \in \mathbb{C}).$$

(1.2) Sei (X, d) ein metrischer Raum und $A \subset X$ eine Teilmenge von X . Die sog. *induzierte Metrik* auf A ist definiert durch

$$\begin{aligned} d_A : A \times A &\longrightarrow \mathbb{R}, \\ (x, y) &\longmapsto d_A(x, y) := d(x, y). \end{aligned}$$

Damit wird A selbst zu einem metrischen Raum.

(1.3) Ein Beispiel aus der Physik: Sei X ein optisches Medium, also ein lichtdurchlässiger Stoff, der nicht notwendig homogen und isotrop zu sein braucht.

X wird zu einem metrischen Raum, wenn man als Abstand $d(x, y)$ zweier Punkte $x, y \in X$ die Zeit (gemessen in einer vorgegebenen Zeiteinheit) definiert, die ein Lichtstrahl einer gewissen Wellenlänge von x nach y braucht. Die drei Axiome der Metrik folgen aus nichttrivialen physikalischen Aussagen:

Die Eigenschaft i) folgt aus der Endlichkeit der Lichtgeschwindigkeit. Die Symmetrie ist wegen des Satzes von der “Umkehrbarkeit des Lichtweges” erfüllt. Die Dreiecksungleichung folgt aus dem “Fermatschen Prinzip”: Ein Lichtstrahl wählt zwischen zwei Punkten immer den Weg, der die kürzeste Zeit beansprucht.

Seien nämlich x, y, z drei Punkte von X . Sei L_1 der Weg des Lichtstrahls von x nach y und L_2 der Weg des Lichtstrahls von y nach z . Bezeichnet L' den aus L_1 und L_2 zusammengesetzten Weg, so ist der Zeitbedarf des Lichtstrahls für L' gleich $d(x, y) + d(y, z)$. Nach dem Fermatschen Prinzip braucht der Lichtstrahl auf dem tatsächlich gewählten Weg L von x nach z höchstens so lange wie auf dem Weg L' , d.h.

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

Dies Beispiel ist jedoch nicht ganz exakt, u.a. deshalb, weil das Fermatsche Prinzip nur lokal gilt.

(1.4) Auf jeder Menge X kann man eine triviale Metrik einführen durch die Definition

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

Normierte Vektorräume

Die wichtigsten Beispiele metrischer Räume entstehen aus normierten Vektorräumen.

Definition. Sei V ein Vektorraum über dem Körper $\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$. Unter einer *Norm* auf V versteht man eine Abbildung

$$\| \cdot \| : V \longrightarrow \mathbb{R}, \quad x \mapsto \|x\|$$

mit folgenden Eigenschaften:

- i) $\|x\| = 0 \Leftrightarrow x = 0$.
- ii) $\|\lambda x\| = |\lambda| \cdot \|x\|$ für alle $\lambda \in \mathbb{K}$ und $x \in V$.

iii) $\|x+y\| \leq \|x\| + \|y\|$ für alle $x, y \in V$.

Ein *normierter Vektorraum* ist ein Paar $(V, \|\cdot\|)$, bestehend aus einem Vektorraum V und einer Norm $\|\cdot\|$ auf V . Ist klar, um welche Norm es sich handelt, schreibt man meist nur kurz V statt $(V, \|\cdot\|)$.

Satz 1. Sei $(V, \|\cdot\|)$ ein normierter Vektorraum. Dann wird durch

$$d(x, y) := \|x - y\| \quad \text{für } x, y \in V$$

eine Metrik auf V definiert.

Die drei Axiome der Metrik folgen unmittelbar aus den entsprechenden Eigenschaften der Norm.

Beispiele

(1.5) Sei V ein Euklidischer Vektorraum, d.h. ein Vektorraum versehen mit einem symmetrischen, positiv definiten, bilinearen Skalarprodukt $\langle x, y \rangle$. Dann wird durch $\|x\| := \sqrt{\langle x, x \rangle}$ eine Norm auf V definiert.

(1.6) Auf dem \mathbb{R}^n betrachten wir das kanonische Skalarprodukt

$$\langle x, y \rangle := x_1 y_1 + x_2 y_2 + \dots + x_n y_n$$

für Vektoren $(x_1, \dots, x_n) \in \mathbb{R}^n$ und $(y_1, \dots, y_n) \in \mathbb{R}^n$. Man nennt

$$\|x\| := \sqrt{\langle x, x \rangle} = \sqrt{x_1^2 + \dots + x_n^2}$$

die *euklidische Norm* von x . Der daraus abgeleitete Abstand ist

$$d(x, y) = \|x - y\| = \sqrt{(x_1 - y_1)^2 + \dots + (x_n - y_n)^2}.$$

Im Folgenden verwenden wir auf dem \mathbb{R}^n , sofern nicht ausdrücklich etwas anderes vermerkt ist, stets diesen euklidischen Abstand.

(1.7) Eine andere Norm auf dem \mathbb{R}^n ist die *Maximum-Norm*

$$|x| := \max(|x_1|, \dots, |x_n|)$$

für $(x_1, \dots, x_n) \in \mathbb{R}^n$. Zwischen der Maximum-Norm und der euklidischen Norm besteht die Beziehung

$$|x| \leq \|x\| \leq \sqrt{n} |x|.$$

(1.8) Sei X eine beliebige Menge und $\mathcal{B}(X)$ der Vektorraum aller beschränkten reellwertigen Funktionen auf X , d.h. aller Funktionen $f: X \rightarrow \mathbb{R}$, für die

$$\|f\|_X := \sup\{|f(x)| : x \in X\} < \infty.$$

Dann ist $\|\cdot\|_X$ eine Norm auf $\mathcal{B}(X)$. Die Eigenschaften i) und ii) sind klar. Die Dreiecksungleichung sieht man so:

$$\begin{aligned} \|f+g\|_X &= \sup\{|f(x)+g(x)| : x \in X\} \\ &\leq \sup\{|f(x)| + |g(x)| : x \in X\} \\ &\leq \sup\{|f(x)| : x \in X\} + \sup\{|g(x)| : x \in X\} \\ &= \|f\|_X + \|g\|_X. \end{aligned}$$

(1.9) Für ein Intervall $[a, b] \subset \mathbb{R}$ sei $C[a, b]$ der Vektorraum aller stetigen Funktionen $f: [a, b] \rightarrow \mathbb{R}$. Bereits in An. 1, (18.6) haben wir die p -Norm

$$\|f\|_p = \left(\int_a^b |f(x)|^p dx \right)^{1/p}$$

kennengelernt.

Bemerkung. Die Beispiele (1.8) und (1.9) funktionieren natürlich genauso mit komplexwertigen statt reellwertigen Funktionen.

Umgebungen, offene Mengen

Wir führen jetzt einige topologische Grundbegriffe ein, die sich in metrischen Räumen definieren lassen.

Bezeichnung. Sei (X, d) ein metrischer Raum, $a \in X$ ein Punkt und $r > 0$. Dann heißt

$$B_r(a) := \{x \in X : d(a, x) < r\}$$

die *offene Kugel* mit Mittelpunkt a und Radius r bzgl. der Metrik d .

(Der Buchstabe B erinnert an engl. *ball* oder frz. *boule*.)

Definition (Umgebung). Sei X ein metrischer Raum. Eine Teilmenge $U \subset X$ heißt *Umgebung* eines Punktes $x \in X$, falls ein $\varepsilon > 0$ existiert, so dass

$$B_\varepsilon(x) \subset U.$$

Insbesondere ist $B_\varepsilon(x)$ selbst eine Umgebung von x . Man nennt $B_\varepsilon(x)$ die ε -Umgebung von x .

Satz 2 (Hausdorffsches Trennungsaxiom). *Sei X ein metrischer Raum. Dann gibt es zu je zwei Punkten $x, y \in X$ mit $x \neq y$ Umgebungen U von x und V von y , die punktfremd sind, d.h. $U \cap V = \emptyset$.*

Beweis. Sei¹ $\varepsilon := \frac{1}{3}\|x, y\|$. Dann ist $\varepsilon > 0$ und

$$U := B_\varepsilon(x), \quad V := B_\varepsilon(y)$$

sind punktfremde Umgebungen von x bzw. y , vgl. Bild 1.2. Denn gäbe es einen Punkt $z \in U \cap V$, so würde mit der Dreiecksungleichung folgen

$$3\varepsilon = \|x, y\| \leq \|x, z\| + \|z, y\| < \varepsilon + \varepsilon,$$

also $3\varepsilon < 2\varepsilon$, Widerspruch!

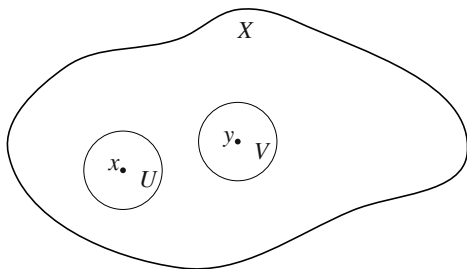


Bild 1.2

Definition (Offene Mengen). Eine Teilmenge U eines metrischen Raumes X heißt *offen*, wenn sie Umgebung jedes ihrer Punkte ist, d.h. wenn zu jedem $x \in U$ ein $\varepsilon > 0$ existiert, so dass

$$B_\varepsilon(x) \subset U.$$

Beispiele

(1.10) Seien $a, b \in \mathbb{R}$, $a < b$. Das Intervall $]a, b[$ ist offen in \mathbb{R} , denn ist $x \in]a, b[$,

¹Es sei daran erinnert, dass wir $\|x, y\|$ als Abkürzung für den Abstand $d(x, y)$ zweier Punkte x, y in einem metrischen Raum (X, d) vereinbart haben. Ist X ein normierter Vektorraum mit Norm $\| \cdot \|$, so ist einfach $\|x, y\| = \|x - y\|$.

so gilt

$$B_\varepsilon(x) \subset]a, b[\quad \text{für} \quad \varepsilon := \min(|a - x|, |b - x|).$$

Ebenso sind die uneigentlichen Intervalle $]a, \infty[$ und $]-\infty, a[$ offen, dagegen sind z.B. die Intervalle $[a, b]$ und $[a, b[$ nicht offen, denn für kein $\varepsilon > 0$ liegt $B_\varepsilon(a)$ ganz in $[a, b]$ oder $[a, b[$.

(1.11) Sei X ein beliebiger metrischer Raum, $a \in X$ und $r > 0$. Dann ist $B_r(a)$ offen im Sinn der obigen Definition. Denn sei $x \in B_r(a)$. Dann ist

$$\varepsilon := r - \|x, a\| > 0$$

und aus der Dreiecksungleichung folgt $B_\varepsilon(x) \subset B_r(a)$, siehe Bild 1.3.

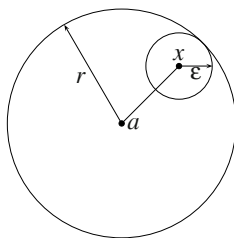


Bild 1.3

(1.12) Bemerkung. Im \mathbb{R}^n erhält man denselben Begriff der offenen Menge, ob man die euklidische Norm oder die Maximum-Norm zugrunde legt. Denn bezeichnet

$$B_\varepsilon(a) := \{x \in \mathbb{R}^n : \|x - a\| < \varepsilon\}$$

die ε -Umgebungen bzgl. der euklidischen Norm und

$$B'_\varepsilon(a) := \{x \in \mathbb{R}^n : |x - a| < \varepsilon\}$$

die ε -Umgebungen bzgl. der Maximum-Norm, so gilt

$$B_{\varepsilon/\sqrt{n}}(a) \subset B_\varepsilon(a) \subset B'_\varepsilon(a).$$

Daraus folgt, dass jede offene Menge bzgl. der euklidischen Norm auch offen bzgl. der Maximum-Norm ist und umgekehrt.

Satz 3. Für die offenen Mengen eines metrischen Raumes X gilt:

a) \emptyset und X sind offen.

- b) Sind U und V offen, so ist auch der Durchschnitt $U \cap V$ offen.
- c) Sei U_i , $i \in I$, eine Familie offener Teilmengen von X . Dann ist auch die Vereinigung $\bigcup_{i \in I} U_i$ offen.

Beweis.

a) Der gesamte Raum X ist offen, da X Umgebung jedes Punktes $x \in X$ ist. Die leere Menge ist offen, da es keinen Punkt $x \in \emptyset$ gibt, zu dem es eine ε -Umgebung $B_\varepsilon(x) \subset \emptyset$ geben müsste.

b) Sei $x \in U \cap V$. Dann gibt es, da U und V offen sind, $\varepsilon_1 > 0$ und $\varepsilon_2 > 0$ mit

$$B_{\varepsilon_1}(x) \subset U \quad \text{und} \quad B_{\varepsilon_2}(x) \subset V.$$

Für $\varepsilon := \min(\varepsilon_1, \varepsilon_2)$ gilt dann $B_\varepsilon(x) \subset U \cap V$, was zeigt, dass $U \cap V$ offen ist.

c) Ist $x \in \bigcup_{i \in I} U_i$, so gibt es einen Index $j \in I$, so dass $x \in U_j$. Da U_j offen ist, existiert ein $\varepsilon > 0$ mit

$$B_\varepsilon(x) \subset U_j \subset \bigcup_{i \in I} U_i, \quad \text{q.e.d.}$$

Bemerkung. Aus b) folgt durch wiederholte Anwendung, dass ein Durchschnitt von endlich vielen offenen Mengen wieder offen ist. Dies gilt nicht mehr für unendliche Durchschnitte. Z.B. sind die Intervalle $]-\frac{1}{n}, 1 + \frac{1}{n}[$, $n \geq 1$, offen in \mathbb{R} , aber ihr Durchschnitt

$$\bigcap_{n=1}^{\infty}]-\frac{1}{n}, 1 + \frac{1}{n}[= [0, 1]$$

ist nicht mehr offen.

Definition (Abgeschlossene Mengen). Eine Teilmenge A eines metrischen Raumes X heißt *abgeschlossen*, wenn ihr Komplement $X \setminus A$ offen ist.

Beispiele

(1.13) Für $a, b \in \mathbb{R}$, $a \leq b$, ist das Intervall $[a, b]$ abgeschlossen, denn sein Komplement

$$\mathbb{R} \setminus [a, b] =]-\infty, a[\cup]b, \infty[$$

ist nach (1.10) und Satz 3c) offen.

(1.14) Sind $A_1 \subset \mathbb{R}^k$ und $A_2 \subset \mathbb{R}^m$ abgeschlossen, so ist auch $A_1 \times A_2 \subset \mathbb{R}^{k+m}$ abgeschlossen. Denn sei $(x, y) \in \mathbb{R}^k \times \mathbb{R}^m$ ein Punkt aus dem Komplement von $A_1 \times A_2$. Dann gilt $x \notin A_1$ oder $y \notin A_2$. Sei etwa $x \notin A_1$. Da A_1 abgeschlossen ist, gibt es ein $\varepsilon > 0$, so dass $B_\varepsilon(x) \subset \mathbb{R}^k \setminus A_1$. Daraus folgt

$$B_\varepsilon((x, y)) \subset \mathbb{R}^{k+m} \setminus A_1 \times A_2,$$

was zeigt, dass das Komplement von $A_1 \times A_2$ offen ist. Also ist $A_1 \times A_2$ abgeschlossen.

Insbesondere folgt daraus, dass jeder Quader

$$Q := \{(x_1, \dots, x_n) \in \mathbb{R}^n : a_i \leq x_i \leq b_i \text{ für } i = 1, \dots, n\},$$

$a_i, b_i \in \mathbb{R}$, $a_i \leq b_i$, abgeschlossen in \mathbb{R}^n ist.

(1.15) In jedem metrischen Raum X sind die Mengen \emptyset und X abgeschlossen, denn ihre Komplemente X und \emptyset sind offen. Es gibt also Teilmengen, die gleichzeitig offen und abgeschlossen sind.

(1.16) Für $a, b \in \mathbb{R}$, $a < b$, ist das Intervall $[a, b[\subset \mathbb{R}$ weder offen noch abgeschlossen.

Topologische Räume

Man kann die Begriffe Umgebung, offene und abgeschlossene Mengen in einen noch abstrakteren Rahmen stellen. Man verzichtet auf eine Metrik und nimmt die offenen Mengen als Grundbegriff.

Definition. Sei X eine Menge. Eine Menge \mathcal{T} von Teilmengen von X heißt *Topologie* auf X , falls gilt:

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

$$\bigcup_{i \in I} U_i \in \mathcal{T}.$$

Ein *topologischer Raum* ist ein Paar (X, \mathcal{T}) , bestehend aus einer Menge X und einer Topologie \mathcal{T} auf X . Eine Teilmenge $U \subset X$ heißt *offen*, wenn sie zu

\mathcal{T} gehört. Eine Teilmenge $A \subset X$ heißt *abgeschlossen*, wenn ihr Komplement $X \setminus A$ offen ist.

Falls klar ist, welche Topologie gemeint ist, schreibt man für einen topologischen Raum nur kurz X statt (X, \mathcal{T}) .

Beispiele

(1.17) Nach Satz 3 bildet das System der offenen Teilmengen eines metrischen Raumes eine Topologie im Sinne der obigen Definition, ein metrischer Raum ist also in natürlicher Weise auch ein topologischer Raum.

(1.18) Der \mathbb{R}^n ist ein metrischer Raum, also auch ein topologischer Raum. Dabei kommt es nach (1.12) nicht darauf an, ob man von der Euklidischen Metrik oder der aus der Maximum-Norm abgeleiteten Metrik ausgeht. Die zugehörige Topologie ist beidesmal dieselbe.

(1.19) Auf jeder Menge X kann man folgende zwei Topologien einführen:

- i) Die grösste Topologie $\mathcal{T}_0 := \{\emptyset, X\}$.
- ii) Die feinste Topologie $\mathcal{T}_1 := \mathfrak{P}(X)$.

Dabei ist $\mathfrak{P}(X)$ die Potenzmenge von X , die aus allen Teilmengen von X besteht. Bzgl. der Topologie \mathcal{T}_1 sind also alle Teilmengen von X offen.

Falls $X \neq \emptyset$, gilt $\mathcal{T}_0 \neq \mathcal{T}_1$, also sind die topologischen Räume (X, \mathcal{T}_0) und (X, \mathcal{T}_1) verschieden, obwohl die unterliegende Menge X gleich ist.

(1.20) Induzierte Topologie. Sei (X, \mathcal{T}) ein topologischer Raum und $Y \subset X$ eine Teilmenge. Dann wird Y auf natürliche Weise wieder ein topologischer Raum mit der sog. induzierten Topologie (oder Relativ-Topologie)

$$\mathcal{T}_1 := \mathcal{T} \cap Y := \{U \cap Y : U \in \mathcal{T}\}.$$

Die Axiome einer Topologie sind für \mathcal{T}_1 leicht nachzuprüfen, also ist (Y, \mathcal{T}_1) wieder ein topologischer Raum. Eine Menge $V \subset Y$ ist genau dann offen bzgl. der induzierten Topologie, wenn es eine offene Menge $U \subset X$ gibt mit $V = U \cap Y$.

Man beachte: Ist Y nicht offen in X , so sind die bzgl. der Relativ-Topologie offenen Teilmengen $V \subset Y$ nicht notwendig offen in X .

Definition. Sei (X, \mathcal{T}) ein topologischer Raum und $x \in X$ ein Punkt. Eine Teilmenge $V \subset X$ heißt *Umgebung* von x , wenn es eine offene Menge $U \in \mathcal{T}$ gibt, so dass

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

Offenbar ist diese Definition im Fall metrischer Räume mit der früher gegebenen äquivalent, da die ε -Umgebungen $B_\varepsilon(x)$ in einem metrischen Raum offen sind.

Was in metrischen Räumen als Definition der offenen Mengen diente (siehe Seite 6), lässt sich in topologischen Räumen nun als Satz beweisen.

Satz 4. Eine Teilmenge V eines topologischen Raumes X ist genau dann offen, wenn sie Umgebung jedes ihrer Punkte ist.

Beweis. a) Ist V offen, so folgt direkt aus der Definition, dass V Umgebung jedes Punktes $x \in V$ ist.

b) Zur Umkehrung. Sei V Umgebung jedes Punktes $x \in V$. Dann gibt es zu jedem $x \in V$ eine offene Menge U_x mit $x \in U_x \subset V$. Daraus folgt

$$\bigcup_{x \in V} U_x = V.$$

Da die Vereinigung beliebig vieler offener Mengen wieder offen ist, ist V offen, q.e.d.

Definition. Ein topologischer Raum (X, \mathcal{T}) heißt *Hausdorff-Raum*, falls in ihm das Hausdorffsche Trennungsaxiom gilt, d.h. zu je zwei Punkten $x, y \in X$, $x \neq y$, existieren Umgebungen U von x und V von y mit $U \cap V = \emptyset$.

Beispiele

(1.21) Nach Satz 2 ist jeder metrische Raum ein Hausdorff-Raum.

(1.22) Sei X die zweipunktige Menge $\{0, 1\}$. Das folgende Mengensystem ist, wie man leicht nachprüft, eine Topologie auf X :

$$\mathcal{T} := \{\emptyset, \{0\}, \{0, 1\}\}.$$

Der topologische Raum (X, \mathcal{T}) ist aber nicht Hausdorffsch, da die Punkte 0 und 1 keine punktfremden Umgebungen besitzen. (Die einzige Umgebung von 1 ist die Menge $\{0, 1\}$.)

Definition (Randpunkt). Sei X ein topologischer Raum und $Y \subset X$ eine Teilmenge und $x \in X$. Der Punkt x heißt *Randpunkt* von Y , wenn in jeder Umgebung von x sowohl ein Punkt von Y als auch ein Punkt von $X \setminus Y$ liegt (Bild 1.4). Die Menge aller Randpunkte von Y heißt der *Rand* von Y und wird mit ∂Y bezeichnet.

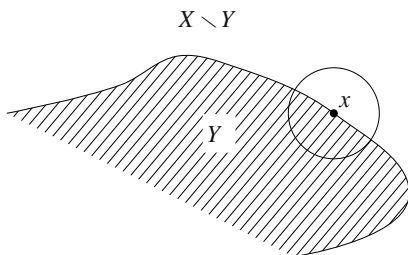


Bild 1.4 Randpunkt

Beispiele

(1.23) Seien $a, b \in \mathbb{R}$, $a < b$, und sei I eines der Intervalle

$$[a, b], [a, b[,]a, b],]a, b[\subset \mathbb{R}.$$

Dann gilt in jedem Fall $\partial I = \{a, b\}$. Dagegen besteht der Rand von $[a, \infty[$ oder $]a, \infty[$ nur aus dem Punkt a .

(1.24) Im \mathbb{R}^n ist der Rand der Einheitskugel

$$K = \{x \in \mathbb{R}^n : \|x\| \leq 1\}$$

die Einheitssphäre

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

(1.25) Der Rand von \mathbb{Q} in \mathbb{R} ist ganz \mathbb{R} , denn in der Umgebung eines jeden Punktes $x \in \mathbb{R}$ liegen sowohl rationale als auch irrationale Zahlen (vgl. An. 1, §9).

Satz 5. Sei X ein topologischer Raum und $Y \subset X$. Dann gilt:

- Die Menge $Y \setminus \partial Y$ ist offen.
- Die Menge $Y \cup \partial Y$ ist abgeschlossen.

c) *Der Rand ∂Y ist abgeschlossen.*

Beweis.

a) Sei $a \in Y \setminus \partial Y$ beliebig. Dann gibt es eine offene Umgebung V von a , so dass

$$V \cap (X \setminus Y) = \emptyset,$$

denn andernfalls wäre a ein Randpunkt von Y . Für dieses V gilt dann auch $V \cap \partial Y = \emptyset$, denn wäre $y \in V \cap \partial Y$, so läge, da y ein Randpunkt von Y und V Umgebung von y ist, in V ein Punkt von $X \setminus Y$, was unserer Annahme über V widerspricht. Insgesamt gilt also

$$V \subset Y \setminus \partial Y.$$

Dies zeigt, dass $Y \setminus \partial Y$ offen ist.

b) Wir setzen $Y' := X \setminus Y$. Aus der Definition des Randes folgt unmittelbar $\partial Y = \partial Y'$. Nach Teil a) ist $Y' \setminus \partial Y'$ offen, also ist

$$X \setminus (Y' \setminus \partial Y') = (X \setminus Y') \cup \partial Y' = Y \cup \partial Y$$

abgeschlossen.

c) Es gilt $\partial Y = (Y \cup \partial Y) \setminus (Y \setminus \partial Y)$, also

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

Nach Teil a) und b) ist dies offen, also ∂Y abgeschlossen.

Definition (Inneres, abgeschlossene Hülle). Ist Y Teilmenge eines topologischen Raumes X , so heißt

$\overset{\circ}{Y} := Y \setminus \partial Y$ das *Innere* oder der *offene Kern* von Y und

$\overline{Y} := Y \cup \partial Y$ die *abgeschlossene Hülle* von Y .

AUFGABEN

1.1. Auf \mathbb{R} werde eine Metrik δ definiert durch

$$\delta(x, y) := \arctan |x - y|.$$

Man zeige, dass δ die Axiome einer Metrik erfüllt und dass die offenen Mengen bzgl. dieser Metrik dieselben sind wie bzgl. der üblichen Metrik $d(x, y) = |x - y|$.

1.2. Sei (X, d) ein metrischer Raum. Man zeige, dass die Abbildung

$$\delta: X \times X \rightarrow \mathbb{R}, \quad \delta(x, y) := \min(d(x, y), 1),$$

eine Metrik auf X ist und dass die Metriken d und δ dieselben offenen Mengen auf X definieren.

1.3. Auf der Menge \mathbb{N} der natürlichen Zahlen werde folgende Topologie eingeführt: Offene Mengen sind außer \emptyset und \mathbb{N} alle Teilmengen $U \subset \mathbb{N}$, so dass $\mathbb{N} \setminus U$ endlich ist. Man zeige, dass die Axiome einer Topologie erfüllt sind, aber das Hausdorffsche Trennungs-Axiom nicht gilt.

1.4. Sei (X, \mathcal{T}) ein topologischer Raum und $Y \subset X$ eine Teilmenge. Man zeige für das Innere und die abgeschlossene Hülle von Y :

- i) $\mathring{Y} = \bigcup \{U : U \subset Y \text{ und } U \text{ offen in } X\},$
- ii) $\bar{Y} = \bigcap \{A : A \supset Y \text{ und } A \text{ abgeschlossen in } X\}.$

1.5. Man beweise: Eine Teilmenge Y eines topologischen Raumes ist genau dann offen, wenn $Y \cap \partial Y = \emptyset$ und genau dann abgeschlossen, wenn $\partial Y \subset Y$.

1.6. Seien $A, B \subset \mathbb{R}$ beliebige Teilmengen. Man zeige, dass für den Rand von $A \times B \subset \mathbb{R}^2$ gilt:

$$\partial(A \times B) = (\partial A \times \bar{B}) \cup (\bar{A} \times \partial B).$$

1.7. Es seien (X_1, d_1) und (X_2, d_2) metrische Räume. Auf dem Produkt $X := X_1 \times X_2$ werde eine Metrik definiert durch

$$d((x_1, x_2), (y_1, y_2)) := \max(d_1(x_1, y_1), d_2(x_2, y_2))$$

für $(x_1, x_2), (y_1, y_2) \in X_1 \times X_2$

- a) Man zeige, dass $d: X \times X \rightarrow \mathbb{R}$ die Axiome einer Metrik erfüllt.
- b) Man beweise: Eine Teilmenge $U_1 \times U_2 \subset X_1 \times X_2$ ist bzgl. dieser Metrik genau dann offen, wenn $U_1 \subset X_1$ und $U_2 \subset X_2$ offen sind.

1.8. Auf der Menge $\overline{\mathbb{R}} := \mathbb{R} \cup \{\pm\infty\}$ werde wie folgt eine Topologie definiert: Eine Teilmenge $U \subset \overline{\mathbb{R}}$ heie offen, wenn folgende Bedingungen erfüllt sind:

- i) $U \cap \mathbb{R}$ ist offen in \mathbb{R} im üblichen Sinn.
- ii) Falls $\infty \in U$, existiert ein $r > 0$ mit $]r, \infty[\subset U$.
- iii) Falls $-\infty \in U$, existiert ein $r > 0$ mit $]-\infty, -r[\subset U$.

Man zeige, dass dadurch eine Topologie auf $\overline{\mathbb{R}}$ definiert wird, mit der $\overline{\mathbb{R}}$ ein Hausdorff-Raum wird.

§ 2 Grenzwerte. Stetigkeit

In diesem Paragraphen wird der Begriff der Konvergenz in metrischen Räumen und die Stetigkeit von Abbildungen zwischen metrischen Räumen eingeführt. Dies verallgemeinert entsprechende Begriffsbildungen für Folgen reeller Zahlen und reelle Funktionen einer Veränderlichen.

Definition (Konvergenz von Folgen). Sei X ein metrischer Raum und $(x_k)_{k \in \mathbb{N}}$ eine Folge von Punkten aus X . Die Folge (x_k) heißt *konvergent* gegen den Punkt $a \in X$, in Zeichen

$$\lim_{k \rightarrow \infty} x_k = a,$$

wenn gilt: Zu jeder Umgebung U von a existiert ein $N \in \mathbb{N}$, so dass

$$x_k \in U \quad \text{für alle } k \geq N.$$

Da in jeder Umgebung eine ε -Umgebung enthalten ist, ist dies gleichbedeutend mit folgender Bedingung: Zu jedem $\varepsilon > 0$ existiert ein $N \in \mathbb{N}$, so dass

$$\|x_k, a\| < \varepsilon \quad \text{für alle } k \geq N.$$

Die Konvergenz von Punktfolgen im \mathbb{R}^n kann einfach auf die Konvergenz von Folgen reeller Zahlen zurückgeführt werden, wie folgender Satz zeigt:

Satz 1. Sei $(x_k)_{k \in \mathbb{N}}$ eine Folge von Punkten im \mathbb{R}^n ,

$$x_k = (x_{k1}, x_{k2}, \dots, x_{kn}), \quad k \in \mathbb{N}.$$

Genau dann konvergiert die Folge (x_k) gegen den Punkt $a = (a_1, a_2, \dots, a_n) \in \mathbb{R}^n$, wenn für $v = 1, 2, \dots, n$ gilt

$$\lim_{k \rightarrow \infty} x_{kv} = a_v.$$

Beweis. a) Es gelte $\lim x_k = a$. Dann gibt es zu vorgegebenem $\varepsilon > 0$ ein $N \in \mathbb{N}$, so dass $\|x_k - a\| < \varepsilon$ für alle $k \geq N$. Daraus folgt für $v = 1, 2, \dots, n$

$$|x_{kv} - a_v| \leq \|x_k - a\| < \varepsilon \quad \text{für } k \geq N.$$

Also ist $\lim_{k \rightarrow \infty} x_{kv} = a_v$.

b) Sei umgekehrt vorausgesetzt, dass $\lim_{k \rightarrow \infty} x_{kv} = a_v$ für $v = 1, \dots, n$. Zu vorgegebenem $\varepsilon > 0$ gibt es dann ein $N_v \in \mathbb{N}$, so dass

$$|x_{kv} - a_v| < \varepsilon' := \frac{\varepsilon}{\sqrt{n}} \quad \text{für alle } k \geq N_v.$$

Für alle $k \geq N := \max(N_1, \dots, N_n)$ gilt dann

$$\|x_k - a\| = \left(\sum_{v=1}^n |x_{kv} - a_v|^2 \right)^{1/2} < \sqrt{n} \varepsilon' = \varepsilon.$$

Also gilt $\lim_{k \rightarrow \infty} x_k = a$.

Mit Hilfe der Konvergenz von Folgen kann man auch die abgeschlossenen Mengen charakterisieren.

Satz 2. Sei X ein metrischer Raum. Eine Teilmenge $A \subset X$ ist genau dann abgeschlossen, wenn gilt: Ist $(x_k)_{k \in \mathbb{N}}$ eine Folge von Punkten $x_k \in A$, die gegen einen Punkt $x \in X$ konvergiert, so liegt x schon in A .

Bemerkung. Dies ist eine Verallgemeinerung von An. 1, §4, Corollar zu Satz 5.

Beweis. a) Sei zunächst A als abgeschlossen vorausgesetzt und $x_k \in A$, $k \in \mathbb{N}$, eine Folge mit $\lim x_k = x$. Angenommen, x läge nicht in A . Da $X \setminus A$ offen ist, ist dann $X \setminus A$ eine Umgebung von x . Nach der Definition der Konvergenz gibt es ein $N \in \mathbb{N}$, so dass $x_k \in X \setminus A$ für alle $k \geq N$. Das ist aber ein Widerspruch.

b) Zur Umkehrung. Das Folgenkriterium sei erfüllt; wir wollen zeigen, dass dann A abgeschlossen, d.h. $X \setminus A$ offen ist. Sei $x \in X \setminus A$ ein beliebiger Punkt. *Behauptung:* Es gibt ein $\varepsilon > 0$ mit $B_\varepsilon(x) \subset X \setminus A$. Wäre dies nicht der Fall, könnten wir zu jedem $k > 0$ ein $x_k \in A$ finden mit $\|x_k, x\| < 1/k$. Dann gilt aber $\lim x_k = x \in A$, was im Widerspruch zu $x \in X \setminus A$ steht. Die Behauptung ist also richtig, was zeigt, dass $X \setminus A$ offen ist.

Cauchyfolgen, Vollständigkeit

Definition. Sei X ein metrischer Raum. Eine Folge $(x_k)_{k \in \mathbb{N}}$ von Punkten aus X heißt *Cauchy-Folge*, wenn gilt: Zu jedem $\varepsilon > 0$ existiert ein $N \in \mathbb{N}$, so dass

$$\|x_k, x_m\| < \varepsilon \quad \text{für alle } k, m \geq N.$$

Bemerkung. Jede konvergente Folge in einem metrischen Raum ist ein Cauchy-Folge.

Beweis. Sei $\lim x_k = x$. Dann gibt es zu vorgegebenem $\varepsilon > 0$ ein $N \in \mathbb{N}$, so dass

$$\|x_k, x\| < \varepsilon/2 \quad \text{für alle } k \geq N.$$

Daraus folgt mit der Dreiecks-Ungleichung für alle $k, m \geq N$

$$\|x_k, x_m\| \leq \|x_k, x\| + \|x, x_m\| < \varepsilon/2 + \varepsilon/2 = \varepsilon, \quad \text{q.e.d.}$$

Definition (Vollständigkeit). Ein metrischer Raum heißt *vollständig*, wenn in ihm jede Cauchy-Folge konvergiert.

Ein vollständiger normierter Vektorraum heißt *Banach-Raum*.

Satz 3. Im \mathbb{R}^n konvergiert jede Cauchy-Folge.

Beweis. Sei $x_k = (x_{k1}, x_{k2}, \dots, x_{kn})$, $k \in \mathbb{N}$, eine Cauchy-Folge in \mathbb{R}^n . Da

$$|x_{kv} - x_{mv}| \leq \|x_k - x_m\|,$$

ist für jedes $v = 1, 2, \dots, n$ die Folge $(x_{kv})_{k \in \mathbb{N}}$ eine Cauchy-Folge in \mathbb{R} , die wegen der Vollständigkeit von \mathbb{R} konvergiert. Nach Satz 1 konvergiert dann die Folge $(x_k)_{k \in \mathbb{N}}$ in \mathbb{R}^n .

Definition (Durchmesser). Für eine Teilmenge A eines metrischen Raumes X wird ihr Durchmesser definiert als

$$\text{diam}(A) := \sup\{\|x, y\| : x, y \in A\}.$$

Die Menge A heißt beschränkt, falls $\text{diam}(A) < \infty$.

Offenbar ist A genau dann beschränkt, wenn A in einer genügend großen Kugel enthalten ist, d.h. wenn ein Punkt $a \in X$ und eine positive reelle Zahl $r > 0$ existiert, so dass $A \subset B_r(a)$. Es gilt

$$\text{diam}(B_r(a)) \leq 2r,$$

wie aus der Dreiecksungleichung folgt.

Der folgende Satz ist eine Verallgemeinerung des Intervallschachtelungs-Prinzips aus An. 1, §5.

Satz 4 (Schachtelungsprinzip). *Sei X ein vollständiger metrischer Raum und*

$$A_0 \supset A_1 \supset A_2 \supset A_3 \supset \dots$$

eine absteigende Folge nichtleerer abgeschlossener Teilmengen mit

$$\lim_{k \rightarrow \infty} \text{diam}(A_k) = 0.$$

Dann gibt es genau einen Punkt $x \in X$, der in allen A_k liegt.

Beweis. Dass es nicht mehr als einen solchen Punkt geben kann, ist klar. Es ist also nur die Existenz zu zeigen. Zu jedem $n \in \mathbb{N}$ wählen wir einen Punkt $x_n \in A_n$. Da

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

ist $(x_n)_{n \in \mathbb{N}}$ eine Cauchy-Folge, konvergiert also gegen ein $x \in X$. Da $x_n \in A_k$ für alle $n \geq k$, folgt aus Satz 2, dass $x \in A_k$, q.e.d.

Stetige Abbildungen

Definition (Stetigkeit). Seien X und Y metrische Räume und $f : X \rightarrow Y$ eine Abbildung. f heißt *stetig* im Punkt $a \in X$, falls

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

d.h. wenn für jede Folge $(x_n)_{n \in \mathbb{N}}$ von Punkten aus X mit $\lim x_n = a$ gilt

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

Die Abbildung f heißt *stetig auf X* , falls f in jedem Punkt $a \in X$ stetig ist.

Satz 5 (Komposition stetiger Abbildungen). *Seien X, Y, Z metrische Räume und*

$$f : X \rightarrow Y, \quad g : Y \rightarrow Z$$

Abbildungen. Ist f stetig im Punkt $a \in X$ und g stetig in $b := f(a) \in Y$, so ist

$$g \circ f : X \rightarrow Z$$

stetig in a .

Beweis. Ist $\lim x_n = a$, so folgt $\lim f(x_n) = f(a) = b$, da f in a stetig ist. Aus

der Stetigkeit von g in b folgt $\lim g(f(x_n)) = g(b) = g(f(a))$, also

$$\lim_{n \rightarrow \infty} (g \circ f)(x_n) = (g \circ f)(a), \quad \text{q.e.d.}$$

Eine Abbildung $f : X \rightarrow \mathbb{R}^n$ mit Werten in \mathbb{R}^n wird durch n Komponenten-Funktionen $f_v : X \rightarrow \mathbb{R}$ gegeben, die durch

$$f(x) = (f_1(x), \dots, f_n(x)) \quad \text{für alle } x \in X$$

definiert sind.

Satz 6. *Sei X ein metrischer Raum. Eine Abbildung*

$$f = (f_1, \dots, f_n) : X \longrightarrow \mathbb{R}^n$$

ist genau dann stetig, wenn alle Komponenten $f_v : X \rightarrow \mathbb{R}$, $v = 1, \dots, n$, stetig sind.

Dies folgt unmittelbar aus Satz 1.

Satz 7. *Folgende Abbildungen sind stetig:*

- a) add: $\mathbb{R} \times \mathbb{R} \longrightarrow \mathbb{R}, \quad (x, y) \mapsto x + y,$
- b) mult: $\mathbb{R} \times \mathbb{R} \longrightarrow \mathbb{R}, \quad (x, y) \mapsto xy,$
- c) quot: $\mathbb{R} \times \mathbb{R}^* \longrightarrow \mathbb{R}, \quad (x, y) \mapsto xy^{-1}.$

Beweis. Sei $((x_k, y_k))_{k \in \mathbb{N}}$ eine Folge von Punkten aus $\mathbb{R} \times \mathbb{R}$ mit

$$\lim_{k \rightarrow \infty} (x_k, y_k) = (x, y).$$

Nach Satz 1 gilt dann $\lim x_k = x$ und $\lim y_k = y$. Daraus folgt

$$\lim_{k \rightarrow \infty} \text{add}(x_k, y_k) = \lim(x_k + y_k) = x + y$$

und

$$\lim_{k \rightarrow \infty} \text{mult}(x_k, y_k) = \lim(x_k y_k) = xy.$$

Gilt zusätzlich $(x_k, y_k) \in \mathbb{R} \times \mathbb{R}^*$ für alle $k \in \mathbb{N}$ und $(x, y) \in \mathbb{R} \times \mathbb{R}^*$, so ist auch

$$\lim_{k \rightarrow \infty} \text{quot}(x_k, y_k) = \lim(x_k y_k^{-1}) = xy^{-1}.$$

Daraus folgt die Stetigkeit von add, mult und quot.

Corollar. Sei X ein metrischer Raum und seien $f, g : X \rightarrow \mathbb{R}$ stetige Funktionen. Dann sind auch die Funktionen

$$f + g : X \rightarrow \mathbb{R} \quad \text{und} \quad fg : X \rightarrow \mathbb{R}$$

stetig. Gilt außerdem $g(x) \neq 0$ für alle $x \in X$, so ist auch

$$\frac{f}{g} : X \rightarrow \mathbb{R}$$

stetig.

Beweis. Nach Satz 6 ist die Abbildung

$$(f, g) : X \longrightarrow \mathbb{R} \times \mathbb{R}$$

stetig. Nun gilt

$$\begin{aligned} f + g &= \text{add} \circ (f, g), \\ fg &= \text{mult} \circ (f, g), \\ f/g &= \text{quot} \circ (f, g). \end{aligned}$$

Aus Satz 7 und Satz 5 folgt nun die Behauptung.

(2.1) Beispiel. Ein Monom vom Grad r auf dem \mathbb{R}^n ist eine Funktion der Gestalt

$$(x_1, \dots, x_n) \mapsto x_1^{k_1} x_2^{k_2} \cdot \dots \cdot x_n^{k_n},$$

wobei k_1, \dots, k_n natürliche Zahlen mit $k_1 + \dots + k_n = r$ sind. Eine Polynomfunktion $F : \mathbb{R}^n \rightarrow \mathbb{R}$ vom Grad $\leq r$ ist eine Linearkombination von Monomen vom Grad $\leq r$,

$$F(x_1, \dots, x_n) = \sum_{k_1 + \dots + k_n \leq r} c_{k_1 \dots k_n} x_1^{k_1} x_2^{k_2} \cdot \dots \cdot x_n^{k_n},$$

$c_{k_1 \dots k_n} \in \mathbb{R}$. Da die Koordinatenfunktionen $(x_1, \dots, x_n) \mapsto x_v$ und die konstanten Funktionen stetig sind, folgt durch wiederholte Anwendung des Corollars, dass alle Polynomfunktionen auf dem \mathbb{R}^n stetig sind.

Satz 8 (ϵ - δ -Kriterium der Stetigkeit). Seien X, Y metrische Räume und $a \in X$ ein Punkt. Eine Abbildung

$$f : X \rightarrow Y$$

ist genau dann in a stetig, wenn zu jedem $\varepsilon > 0$ ein $\delta > 0$ existiert, so dass

$$\|f(x), f(a)\| < \varepsilon \quad \text{für alle } x \in X \text{ mit } \|x, a\| < \delta.$$

(Dieser Satz verallgemeinert den analogen Satz aus An. 1, § 11.)

Beweis.

a) Wir setzen zunächst voraus, dass f in a stetig ist, d.h. $\lim_{x \rightarrow a} f(x) = f(a)$.

Annahme: Das ε - δ -Kriterium ist nicht erfüllt. Dann gibt es ein $\varepsilon > 0$, so dass für jedes $\delta > 0$ ein $x \in X$ existiert mit

$$\|x, a\| < \delta \quad \text{aber} \quad \|f(x), f(a)\| \geq \varepsilon.$$

Insbesondere gibt es zu $\delta = \frac{1}{n}$ ein $x_n \in X$ mit

$$\|x_n, a\| < \frac{1}{n} \quad \text{und} \quad \|f(x_n), f(a)\| \geq \varepsilon. \quad (*)$$

Also ist $\lim x_n = a$, woraus folgt $\lim f(x_n) = f(a)$. Dies steht aber im Widerspruch zu (*). Also ist das ε - δ -Kriterium doch erfüllt.

b) Das ε - δ -Kriterium sei erfüllt.

Sei (x_n) eine Folge in X mit $\lim x_n = a$. Wir müssen zeigen $\lim f(x_n) = f(a)$. Zu vorgegebenem $\varepsilon > 0$ gibt es ein $\delta > 0$, so dass

$$\|x, a\| < \delta \implies \|f(x), f(a)\| < \varepsilon.$$

Es gibt ein $N \in \mathbb{N}$ mit $\|x_n, a\| < \delta$ für alle $n \geq N$. Dann ist $\|f(x_n), f(a)\| < \varepsilon$ für alle $n \geq N$. Also ist $\lim f(x_n) = f(a)$, q.e.d.

(2.2) Beispiel. Sei (X, d) ein metrischer Raum und $x_0 \in X$. Die Funktion $f : X \rightarrow \mathbb{R}$ sei definiert durch

$$f(x) := d(x, x_0), \quad (\text{Abstand vom Punkt } x_0).$$

Diese Funktion ist in jedem Punkt $a \in X$ stetig, denn es folgt aus der Dreiecksungleichung

$$|f(x) - f(a)| = |d(x, x_0) - d(a, x_0)| \leq d(x, a),$$

d.h. $|f(x) - f(a)| < \varepsilon$ für $d(x, a) < \varepsilon$. Beim ε - δ -Kriterium kann man also hier $\delta = \varepsilon$ wählen.

Definition. Seien X, Y topologische Räume und $f : X \rightarrow Y$ eine Abbildung. Die Abbildung f heißt *stetig* im Punkt $a \in X$, wenn zu jeder Umgebung V von $f(a) \in Y$ eine Umgebung U von a existiert mit $f(U) \subset V$. Die Abbildung f heißt *stetig auf X* (oder *stetig schlechthin*), wenn sie in jedem Punkt $x \in X$ stetig ist.

Dies verallgemeinert den Begriff der Stetigkeit von Abbildungen zwischen metrischen Räumen. Denn für metrische Räume X, Y ist die obige Definition nur eine Umformulierung des ε - δ -Kriteriums aus Satz 8.

Satz 9. Seien X, Y zwei topologische Räume. Eine Abbildung $f : X \rightarrow Y$ ist genau dann auf ganz X stetig, wenn das Urbild $f^{-1}(V)$ jeder offenen Menge $V \subset Y$ offen in X ist.

Beweis. Sei zunächst f als stetig vorausgesetzt und sei V offen in Y . Es ist zu zeigen, dass $f^{-1}(V)$ offen in X ist. Sei $a \in f^{-1}(V)$ beliebig. Da V Umgebung von $f(a) \in Y$ ist, gibt es wegen der Stetigkeit von f im Punkt a eine Umgebung U von a mit $f(U) \subset V$. Daraus folgt aber $U \subset f^{-1}(V)$. Deshalb ist $f^{-1}(V)$ Umgebung von a . Da $a \in f^{-1}(V)$ beliebig war, ist $f^{-1}(V)$ offen.

Sei umgekehrt vorausgesetzt, dass das Urbild jeder offenen Menge offen ist und sei $a \in X$ beliebig. Ist V eine Umgebung von $f(a)$, so gibt es eine offene Menge V_1 mit $f(a) \in V_1 \subset V$. Dann ist $U := f^{-1}(V_1)$ offen. U enthält den Punkt a , ist also Umgebung von a , und es gilt $f(U) \subset V$. Also ist f in a stetig, q.e.d.

Bemerkung. Da die abgeschlossenen Mengen gerade die Komplemente der offenen Mengen sind, gilt auch: Eine Abbildung $f : X \rightarrow Y$ ist genau dann stetig, wenn das Urbild jeder abgeschlossenen Menge abgeschlossen ist.

(2.3) Beispiel. Sei X ein topologischer Raum, $f : X \rightarrow \mathbb{R}$ eine stetige Funktion und $c \in \mathbb{R}$. Dann ist die Menge

$$U := \{x \in X : f(x) < c\}$$

offen und die Menge

$$A := \{x \in X : f(x) = c\}$$

abgeschlossen. Denn es gilt $U = f^{-1}(]-\infty, c[)$ und $A = f^{-1}(\{c\})$. Die Menge $]-\infty, c[$ ist offen und die Menge $\{c\}$ abgeschlossen in \mathbb{R} .

Definition (Homöomorphismus). Seien X, Y topologische Räume. Eine bijektive Abbildung $f: X \rightarrow Y$ heißt *Homöomorphismus* (oder *topologische Abbildung*), wenn f stetig ist und die Umkehrabbildung $f^{-1}: Y \rightarrow X$ ebenfalls stetig ist. Zwei topologische Räume heißen *homöomorph*, wenn es einen Homöomorphismus $f: X \rightarrow Y$ gibt.

(2.4) Als *Beispiel* zeigen wir, dass der \mathbb{R}^n zur offenen Einheitskugel

$$B := \{x \in \mathbb{R}^n : \|x\| < 1\}$$

homöomorph ist. Ein Homöomorphismus $f: \mathbb{R}^n \rightarrow B$ wird gegeben durch

$$x \mapsto f(x) := \frac{x}{1 + \|x\|}.$$

Diese Abbildung ist stetig, da $x \mapsto x$ und $x \mapsto 1 + \|x\|$ stetig sind. Es ist leicht nachzuprüfen, dass f bijektiv ist mit der stetigen Umkehrabbildung

$$g := f^{-1}: B \longrightarrow \mathbb{R}^n, \quad x \mapsto \frac{x}{1 - \|x\|}.$$

(2.5) Wir geben noch ein Beispiel einer bijektiven stetigen Abbildung zwischen zwei metrischen Räumen, deren Umkehrung nicht stetig ist.

Sei $X := [0, 2\pi[\subset \mathbb{R}$ und $Y := \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 = 1\} \subset \mathbb{R}^2$, jeweils versehen mit der von \mathbb{R} bzw. \mathbb{R}^2 induzierten Metrik. Die Abbildung

$$f: X \longrightarrow Y, \quad t \mapsto (\cos t, \sin t),$$

ist stetig und bijektiv. Die Umkehrabbildung $f^{-1}: Y \rightarrow X$ ist aber unstetig im Punkt $(1, 0) \in Y$. Die Punktfolge

$$p_k := (\cos(2\pi - 1/k), \sin(2\pi - 1/k)) \in Y, \quad k \geq 1,$$

konvergiert für $k \rightarrow \infty$ gegen den Punkt $(1, 0) = f(0)$, die Folge

$$f^{-1}(p_k) = 2\pi - 1/k, \quad k \geq 1,$$

konvergiert aber nicht gegen 0.

Man kann sogar zeigen (vgl. § 3), dass es überhaupt keinen Homöomorphismus $X \rightarrow Y$ gibt.

Gleichmäßige Konvergenz von Funktionenfolgen

Definition. Seien X eine beliebige Menge, Y ein metrischer Raum, sowie

$$f_n : X \rightarrow Y, \quad n \in \mathbb{N}, \quad \text{und} \quad f : X \rightarrow Y$$

Abbildungen. Man sagt, die Folge $(f_n)_{n \in \mathbb{N}}$ konvergiere *gleichmäßig* gegen f , falls zu jedem $\varepsilon > 0$ ein $N \in \mathbb{N}$ existiert, so dass

$$\|f_n(x), f(x)\| < \varepsilon \quad \text{für alle } x \in X \text{ und alle } n \geq N.$$

Satz 10. Sei X ein topologischer Raum, Y ein metrischer Raum und $f_n : X \rightarrow Y$, $n \in \mathbb{N}$, eine Folge stetiger Funktionen, die gleichmäßig gegen die Funktion $f : X \rightarrow Y$ konvergiere. Dann ist auch f stetig.

Beweis (vgl. An. 1, § 21, Satz 1). Wir zeigen, dass f in einem beliebigen Punkt $a \in X$ stetig ist. Sei $\varepsilon > 0$ vorgegeben. Wegen der gleichmäßigen Konvergenz existiert ein $N \in \mathbb{N}$, so dass

$$\|f_N(x), f(x)\| < \frac{\varepsilon}{3} \quad \text{für alle } x \in X.$$

Da f_N im Punkt a stetig ist, gibt es eine Umgebung U von a , so dass

$$\|f_N(x), f_N(a)\| < \frac{\varepsilon}{3} \quad \text{für alle } x \in U$$

Daher gilt für alle $x \in U$

$$\begin{aligned} \|f(x), f(a)\| &\leq \|f(x), f_N(x)\| + \|f_N(x), f_N(a)\| + \|f_N(a), f(a)\| \\ &< \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon, \quad \text{q.e.d.} \end{aligned}$$

Lineare Abbildungen

Satz 11. Seien V und W normierte Vektorräume (über \mathbb{R} oder \mathbb{C}) und sei

$$A : V \rightarrow W$$

eine lineare Abbildung. A ist genau dann stetig, wenn es eine reelle Konstante $C \geq 0$ gibt, so dass

$$\|A(x)\| \leq C\|x\| \quad \text{für alle } x \in V.$$

Beweis. a) Wir setzen zunächst die Stetigkeit von A im Nullpunkt voraus. Dann gibt es zu $\varepsilon = 1$ ein $\delta > 0$, so dass

$$\|A(z)\| < 1 \quad \text{für alle } z \in V \text{ mit } \|z\| < \delta.$$

Wir setzen $C := 2/\delta$. Sei jetzt $x \in V \setminus 0$ beliebig, $\lambda := (C\|x\|)^{-1}$ und $z := \lambda x$. Dann gilt $\|z\| = |\lambda| \cdot \|x\| = \delta/2 < \delta$, also $\|A(z)\| < 1$. Nun ist

$$A(z) = A(\lambda x) = \lambda A(x) = \frac{1}{C\|x\|} A(x),$$

also folgt $\|A(x)\| \leq C\|x\|$.

b) Es gebe eine Konstante $C \geq 0$ mit $\|A(x)\| \leq C\|x\|$ für alle $x \in V$. Dann gilt

$$\|A(x) - A(x_0)\| = \|A(x - x_0)\| \leq C\|x - x_0\|.$$

Mit Hilfe des ε - δ -Kriteriums folgt daraus die Stetigkeit von A in x_0 .

Bemerkung. Wie aus dem Beweis hervorgeht, ist die lineare Abbildung $A : V \rightarrow W$ genau dann auf ganz V stetig, wenn A im Nullpunkt stetig ist.

Beispiele

(2.6) Sei $C[a, b]$ der Vektorraum aller stetigen Funktionen $f : [a, b] \rightarrow \mathbb{R}$ auf dem Intervall $[a, b] \subset \mathbb{R}$, versehen mit der Supremums-Norm

$$\|f\| := \sup\{|f(x)| : x \in [a, b]\}.$$

Sei $I : C[a, b] \rightarrow \mathbb{R}$ die durch das Integral

$$I(f) := \int_a^b f(x) dx$$

gegebene lineare Abbildung. Dann ist I stetig, denn es gilt die Abschätzung

$$|I(f)| \leq (b - a)\|f\| \quad \text{für alle } f \in C[a, b].$$

(2.7) Sei $C^1[0, 1] \subset C[0, 1]$ der Untervektorraum aller stetig differenzierbaren Funktionen, ebenfalls versehen mit der Supremums-Norm. Sei

$$D : C^1[0, 1] \rightarrow C[0, 1], \quad f \mapsto f',$$

die durch die Differentiation $Df := f'$ gegebene lineare Abbildung.

Behauptung. D ist nicht stetig.

Beweis. Für die Funktionen $f_n \in C^1[0, 1]$, $f_n(x) := x^n$, gilt

$$\|f_n\| = 1 \quad \text{und} \quad \|Df_n\| = n.$$

Daher gibt es keine Konstante $C \geq 0$ mit $\|Df_n\| \leq C\|f_n\|$ für alle n .

Definition (Norm einer linearen Abbildung). Seien V und W normierte Vektorräume und $A : V \rightarrow W$ eine stetige lineare Abbildung. Dann wird ihre *Norm* definiert als

$$\|A\| := \sup\{\|A(x)\| : x \in V \text{ mit } \|x\| \leq 1\}.$$

Bemerkung. Nach Satz 11 ist $\|A\| < \infty$. Es gilt

$$\|A(x)\| \leq \|A\| \cdot \|x\| \quad \text{für alle } x \in V.$$

Dies folgt daraus, dass $\|A(\frac{x}{\|x\|})\| \leq \|A\|$ für alle $x \neq 0$.

Die Menge aller stetigen linearen Abbildungen $A : V \rightarrow W$ bildet in natürlicher Weise einen Vektorraum. Man beweist wie in Beispiel (1.8), dass $A \mapsto \|A\|$ die Axiome einer Norm erfüllt.

(2.8) Beispiel. Sei speziell $V = \mathbb{R}^n$ und $W = \mathbb{R}^m$. Nach (2.1) und Satz 6 ist jede lineare Abbildung $A : \mathbb{R}^n \rightarrow \mathbb{R}^m$ stetig. Bezüglich der kanonischen Basen wird A durch eine $m \times n$ -Matrix

$$(a_{ik})_{1 \leq i \leq m, 1 \leq k \leq n} \in M(m \times n, \mathbb{R})$$

gegeben und man hat folgende Abschätzungen für die Norm von A

$$\max_{i,k} |a_{ik}| \leq \|A\| \leq \sqrt{nm} \max_{i,k} |a_{ik}|.$$

Die erste Abschätzung ist trivial, die zweite sieht man so: Sei $\|x\| \leq 1$ und $y = Ax$. Für die Komponenten von y gilt dann

$$y_i = \sum_{k=1}^n a_{ik}x_k, \quad i = 1, \dots, m.$$

Dies kann man auffassen als das Skalarprodukt des Vektors $(a_{i1}, a_{i2}, \dots, a_{in})$ mit dem Vektor x . Aus der Cauchy-Schwarzschen Ungleichung folgt $|y_i| \leq \sqrt{n} \alpha$, wobei $\alpha = \max_{i,k} |a_{ik}|$. Daraus folgt $\|y\| \leq \sqrt{nm} \alpha$, also $\|A\| \leq \sqrt{nm} \alpha$.

AUFGABEN

2.1. Seien $f, g : X \rightarrow \mathbb{R}$ zwei stetige Funktionen auf dem metrischen Raum X . Für $x \in X$ werde definiert

$$\varphi(x) := \max(f(x), g(x)),$$

$$\psi(x) := \min(f(x), g(x)).$$

Man zeige, dass die Funktionen $\varphi, \psi : X \rightarrow \mathbb{R}$ stetig sind.

2.2. Sei W der offene Würfel im \mathbb{R}^n ,

$$W := \{(x_1, \dots, x_n) \in \mathbb{R}^n : |x_i| < 1 \text{ für } i = 1, \dots, n\}$$

Man konstruiere einen Homöomorphismus von W auf die Einheitskugel

$$B_1(0) = \{x \in \mathbb{R}^n : \|x\| < 1\}.$$

2.3. Man zeige, dass der Vektorraum $C[a, b]$ aller stetigen Funktionen $f : [a, b] \rightarrow \mathbb{R}$ auf dem kompakten Intervall $[a, b] \subset \mathbb{R}$ mit der Supremumsnorm

$$\|f\| := \sup\{|f(x)| : x \in [a, b]\}$$

vollständig ist.

2.4. Auf dem Vektorraum $C^1[a, b]$ aller einmal stetig differenzierbaren Funktionen $f : [a, b] \rightarrow \mathbb{R}$ werde folgende Norm eingeführt:

$$\|f\|_{C^1} := \sup\{|f(x)| + |f'(x)| : x \in [a, b]\}.$$

a) Man zeige, dass $C^1[a, b]$ mit dieser Norm vollständig ist.

b) Man zeige, dass die Abbildung

$$D : C^1[a, b] \longrightarrow C[a, b], \quad f \mapsto f',$$

stetig wird, wenn man $C^1[a, b]$ mit der $\|\cdot\|_{C^1}$ -Norm und $C[a, b]$ mit der Supremums-Norm versieht.

2.5. Sei X ein vollständiger metrischer Raum und $Y \subset X$ eine Teilmenge. Man zeige: Y ist mit der induzierten Metrik genau dann vollständig, wenn Y abgeschlossen in X ist.

§ 3 Kompaktheit

Wir kommen jetzt zu dem sehr wichtigen Begriff der Kompaktheit und studieren das Verhalten stetiger Funktionen auf kompakten Mengen, wie Annahme von Maximum und Minimum und gleichmäßige Stetigkeit. Wir erhalten dabei von neuem von einem abstrakteren Standpunkt aus die schon in Analysis 1 bewiesenen Sätze über stetige Funktionen auf beschränkten abgeschlossenen Intervallen in \mathbb{R} .

Definition. Sei A eine Teilmenge eines topologischen Raumes X . Unter einer *offenen Überdeckung* von A versteht man eine Familie $(U_i)_{i \in I}$ von offenen Teilmengen $U_i \subset X$ mit

$$A \subset \bigcup_{i \in I} U_i.$$

Dabei ist I eine beliebige (endliche oder unendliche) Indexmenge.

Definition. Eine Teilmenge A eines metrischen Raumes (oder allgemeiner eines Hausdorff-Raumes) X heißt *kompakt*, wenn jede offene Überdeckung $(U_i)_{i \in I}$ von A eine endliche Teilüberdeckung besitzt, d.h. endlich viele Indizes $i_1, \dots, i_k \in I$ existieren, so dass

$$A \subset U_{i_1} \cup U_{i_2} \cup \dots \cup U_{i_k}.$$

Ein topologischer Raum X heißt *kompakt*, wenn X Hausdorffsch ist und jede offene Überdeckung von X eine endliche Teilüberdeckung besitzt.

Bemerkung. Dieser Begriff bereitet dem Anfänger erfahrungsgemäß große Schwierigkeiten. Die Definition besagt nicht, dass A kompakt ist, wenn A eine endliche offene Überdeckung besitzt. (Jede Teilmenge von X besitzt eine endliche offene Überdeckung, z.B. die aus der offenen Menge X allein bestehende Überdeckung.) Es wird vielmehr verlangt, dass man aus einer beliebig vorgegebenen offenen Überdeckung eine endliche Teilüberdeckung auswählen kann. Man nennt dies die Heine-Borelsche Überdeckungseigenschaft.

Ein genaues Studium des Beweises des nächsten Satzes und des nachfolgenden Beispiels hilft viel zum Verständnis des Kompaktheitsbegriffs.

Satz 1. Sei X ein metrischer Raum und $(x_n)_{n \in \mathbb{N}}$ eine Punktfolge in X , die gegen den Punkt $a \in X$ konvergiert. Dann ist die Menge

$$A := \{x_n : n \in \mathbb{N}\} \cup \{a\}$$

kompakt.

Beweis. Sei $(U_i)_{i \in I}$ eine offene Überdeckung von A . Da $a \in A$, gibt es einen Index $i^* \in I$, so dass $a \in U_{i^*}$. Weil U_{i^*} offen ist, ist es eine Umgebung von a und wegen $\lim x_n = a$ gibt es ein $N \in \mathbb{N}$, so dass

$$x_n \in U_{i^*} \quad \text{für alle } n > N.$$

Außerdem liegt jedes x_k in einem gewissen U_{i_k} . Es gilt dann

$$A \subset U_{i_0} \cup U_{i_1} \cup \dots \cup U_{i_N} \cup U_{i^*}.$$

Wir haben also eine endliche Teilüberdeckung gefunden.

(3.1) Der Satz gilt i.Allg. nicht mehr, wenn man aus A den Grenzwert der Folge weglässt. Dies zeigt folgendes Beispiel: Sei

$$A := \left\{ \frac{1}{n} : n \in \mathbb{N} \setminus 0 \right\} \subset \mathbb{R}.$$

Behauptung: A ist nicht kompakt.

Beweis. Wir setzen

$$U_1 :=]\frac{1}{2}, 2[\quad \text{und} \quad U_n :=]\frac{1}{n+1}, \frac{1}{n-1}[\quad \text{für } n \geq 2.$$

U_n ist offen, also $(U_n)_{n \geq 1}$ eine offene Überdeckung von A . Jedes U_n enthält genau einen Punkt von A , nämlich $\frac{1}{n}$. Deshalb wird A von keinem endlichen Teilsystem $(U_{n_1}, U_{n_2}, \dots, U_{n_k})$ überdeckt.

Wir wollen die kompakten Teilmengen des \mathbb{R}^n charakterisieren. Dazu beweisen wir zunächst, dass alle abgeschlossenen Quader kompakt sind.

Satz 2 (kompakte Quader). Seien $a_v, b_v \in \mathbb{R}$, $a_v \leq b_v$, $v = 1, 2, \dots, n$. Dann ist der abgeschlossene Quader

$$Q := \{(x_1, \dots, x_n) \in \mathbb{R}^n : a_v \leq x_v \leq b_v\}$$

kompakt in \mathbb{R}^n .

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

Annahme. Q kann nicht durch endlich viele U_{i_k} überdeckt werden.

Um diese Annahme zum Widerspruch zu führen, konstruieren wir durch vollständige Induktion eine Folge von abgeschlossenen Teilquadern

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

mit folgenden Eigenschaften:

- i) Q_m kann nicht durch endlich viele U_{i_k} überdeckt werden.
- ii) $\text{diam}(Q_m) = 2^{-m} \text{diam}(Q)$.

Wir setzen $Q_0 := Q$. Sei Q_m schon konstruiert,

$$Q_m = I_1 \times I_2 \times \dots \times I_n,$$

wobei $I_v \subset \mathbb{R}$ abgeschlossene Intervalle sind. Wir zerlegen I_v in zwei abgeschlossene Intervalle der halben Länge,

$$I_v = I_v^{(1)} \cup I_v^{(2)},$$

und setzen für $s_v \in \{1, 2\}$

$$Q_m^{(s_1, \dots, s_n)} = I_1^{(s_1)} \times I_2^{(s_2)} \times \dots \times I_n^{(s_n)}.$$

Wir erhalten so 2^n Quader mit

$$\bigcup_{s_1, \dots, s_n} Q_m^{(s_1, \dots, s_n)} = Q_m \quad (\text{siehe Bild 3.1}).$$

Q_m	
$Q_m^{(1,2)}$	$Q_m^{(2,2)}$
$Q_m^{(1,1)}$	$Q_m^{(2,1)}$

Bild 3.1

Da Q_m nicht von endlich vielen U_{i_k} überdeckt werden kann, gibt es mindestens einen der Quader $Q_m^{(s_1, \dots, s_n)}$, der nicht von endlich vielen U_{i_k} überdeckt werden kann. Diesen wählen wir als Q_{m+1} . Es gilt

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

Deshalb hat Q_{m+1} wieder die Eigenschaften i) und ii).

Nach dem Schachtelungsprinzip (§2, Satz 4) gibt es einen Punkt a , der in allen Q_m liegt. Da $(U_i)_{i \in I}$ eine Überdeckung von Q ist, gilt $a \in U_{i_0}$ für (mindestens) einen Index $i_0 \in I$. Wegen der Offenheit von U_{i_0} existiert ein $\varepsilon > 0$ mit

$$B_\varepsilon(a) \subset U_{i_0}.$$

Sei nun m so groß, dass $\text{diam}(Q_m) < \varepsilon$. Da $a \in Q_m$, folgt

$$Q_m \subset B_\varepsilon(a) \subset U_{i_0}.$$

Dies ist ein Widerspruch zu i). Deshalb ist die Annahme falsch und der Satz bewiesen.

Satz 3. *Jede kompakte Teilmenge A eines metrischen Raumes X ist beschränkt.*

Beweis. Sei $a \in A$ beliebig. (Falls A leer ist, ist die Behauptung trivial.) Da jeder Punkt aus X einen endlichen Abstand von a hat, gilt

$$\bigcup_{n=1}^{\infty} B_n(a) = X,$$

also ist $(B_n(a))_{n \geq 1}$ eine Überdeckung von A . Weil A kompakt ist, gibt es endlich viele Indizes n_1, \dots, n_k mit

$$A \subset \bigcup_{j=1}^k B_{n_j}(a).$$

Für $n := \max(n_1, \dots, n_k)$ gilt also $A \subset B_n(a)$, d.h. A ist beschränkt.

Corollar. *Jede konvergente Folge in einem metrischen Raum ist beschränkt.*

Dabei heißt eine Folge $(x_n)_{n \in \mathbb{N}}$ beschränkt, wenn die Menge $\{x_n : n \in \mathbb{N}\}$ beschränkt ist.

Das Corollar folgt aus Satz 3 und Satz 1.

Satz 4. *Sei X ein Hausdorff-Raum und $K \subset X$ eine kompakte Teilmenge. Dann ist K abgeschlossen und jede abgeschlossene Teilmenge $A \subset K$ ist kompakt.*

Beweis. a) Wir zeigen zunächst, dass K abgeschlossen ist, d.h. dass $X \setminus K$ offen ist. Sei $b \in X \setminus K$ ein vorgegebener Punkt. Da X Hausdorffsch ist, gibt es zu

jedem Punkt $x \in K$ offene Umgebungen U_x von x und V_x von b mit $U_x \cap V_x = \emptyset$. Da

$$\bigcup_{x \in K} U_x \supset K,$$

und K kompakt ist, gibt es endlich viele Punkte $x_1, \dots, x_s \in K$ mit

$$U_{x_1} \cup U_{x_2} \cup \dots \cup U_{x_s} \supset K.$$

Dann ist

$$V := V_{x_1} \cap V_{x_2} \cap \dots \cap V_{x_s}$$

eine offene Umgebung von b mit $V \cap \bigcup_{i=1}^s U_{x_i} = \emptyset$, also $V \subset X \setminus K$. Da $b \in X \setminus K$ beliebig war, ist $X \setminus K$ Umgebung jedes seiner Punkte, also offen.

b) Sei $(U_i)_{i \in I}$ eine offene Überdeckung von A . Die Menge $U' := X \setminus A$ ist offen und es gilt

$$\bigcup_{i \in I} U_i \cup U' = X \supset K.$$

Da K kompakt ist, gibt es endlich viele Indizes $i_1, \dots, i_k \in I$ mit

$$U_{i_1} \cup \dots \cup U_{i_k} \cup U' \supset K.$$

Daraus folgt

$$U_{i_1} \cup \dots \cup U_{i_k} \supset A, \quad \text{q.e.d.}$$

Satz 5 (Heine-Borel). Eine Teilmenge $A \subset \mathbb{R}^n$ ist genau dann kompakt, wenn sie abgeschlossen und beschränkt ist.

Beweis. Ist A kompakt, so ist es nach den Sätzen 3 und 4 beschränkt und abgeschlossen.

Ist umgekehrt A beschränkt und abgeschlossen, so ist A in einem genügend großen abgeschlossenen Quader Q enthalten, der nach Satz 2 kompakt ist. Nach Satz 4 ist dann A kompakt.

(3.2) Beispiel. Sei $A \subset \mathbb{R}$ kompakt. Da A beschränkt ist, sind $\sup(A)$ und $\inf(A)$ endlich. Es existieren Folgen $x_k \in A$, $y_k \in A$, ($k \in \mathbb{N}$), mit

$$\lim x_k = \sup(A) \quad \text{und} \quad \lim y_k = \inf(A),$$

vgl. An. 1, § 9. Da A abgeschlossen ist, folgt aus §2, Satz 2, dass

$$\sup(A) \in A \quad \text{und} \quad \inf(A) \in A.$$

(3.3) Man beachte: In einem beliebigen metrischen Raum ist eine abgeschlossene und beschränkte Teilmenge nicht notwendig kompakt. Dazu betrachten wir folgendes Beispiel: Sei $X := \mathbb{R}^n$, aber mit einer anderen Metrik als der euklidischen Metrik versehen:

$$\delta : X \times X \rightarrow \mathbb{R}, \quad \delta(x, y) := \min(\|x - y\|, 1).$$

Nach Aufgabe 1.2 sind die offenen Mengen bzgl. dieser Metrik dieselben wie bzgl. der euklidischen Metrik; also sind auch die kompakten Mengen dieselben. Der ganze Raum X ist beschränkt (mit $\text{diam}(X) = 1$) und abgeschlossen, aber nicht kompakt.

Satz 6. Seien X, Y Hausdorff-Räume und $f: X \rightarrow Y$ eine stetige Abbildung. Ist $K \subset X$ kompakt, so ist auch $f(K) \subset Y$ kompakt.

Beweis. Sei $(U_i)_{i \in I}$ eine offene Überdeckung von $f(K)$. Nach § 2, Satz 9, ist $V_i := f^{-1}(U_i)$ offen in X und es gilt $K \subset \bigcup_{i \in I} V_i$. Da K kompakt ist, gibt es endlich viele Indizes $i_1, \dots, i_k \in I$, so dass

$$K \subset V_{i_1} \cup \dots \cup V_{i_k}.$$

Daraus folgt

$$f(K) \subset U_{i_1} \cup \dots \cup U_{i_k} \quad \text{q.e.d.}$$

(3.4) Insbesondere folgt aus Satz 6: Sind zwei Hausdorff-Räume X, Y homöomorph, und ist einer der beiden Räume kompakt, so auch der andere. Zum Beispiel ist die abgeschlossene Einheitskugel

$$K := \{x \in \mathbb{R}^n : \|x\| \leq 1\}$$

weder zur offenen Einheitskugel noch zum ganzen \mathbb{R}^n homöomorph; vgl. dazu Beispiel (2.4). Ebenso folgt (vgl. Beispiel (2.5)), dass das Intervall $[0, 2\pi[\subset \mathbb{R}$ nicht zur 1-Sphäre

$$\mathbb{S}^1 = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 = 1\}$$

homöomorph ist.

Satz 7. Sei X ein kompakter Hausdorff-Raum und $f: X \rightarrow \mathbb{R}$ eine stetige Funktion. Dann ist f beschränkt und nimmt sein Maximum und Minimum an, d.h. es gibt Punkte $p, q \in X$ mit

$$f(p) = \sup\{f(x) : x \in X\}, \quad f(q) = \inf\{f(x) : x \in X\}.$$

Beweis. Nach Satz 6 ist $A := f(X) \subset \mathbb{R}$ kompakt. Die Behauptung folgt deshalb aus der in (3.2) gemachten Bemerkung.

(3.5) Beispiel: Abstandsfunktion. Sei (X, d) ein metrischer Raum, $A \subset X$ eine Teilmenge und $x \in X$. Der Abstand des Punktes x von der Menge A wird definiert als

$$\text{dist}(x, A) := \inf\{d(x, y) : y \in A\}.$$

Die Funktion $x \mapsto \text{dist}(x, A)$ ist stetig auf X , denn wegen

$$\text{dist}(x', A) \leq d(x', x) + \text{dist}(x, A)$$

gilt $|\text{dist}(x, A) - \text{dist}(x', A)| < \varepsilon$ für $d(x, x') < \varepsilon$.

Für eine weitere Teilmenge $K \subset X$ definiert man

$$\text{dist}(K, A) := \inf\{\text{dist}(x, A) : x \in K\} = \inf\{d(x, y) : x \in K, y \in A\}.$$

Behauptung. Ist A abgeschlossen, K kompakt und $A \cap K = \emptyset$, so gilt

$$\text{dist}(K, A) > 0.$$

Beweis. Da K kompakt und die Funktion $x \mapsto \text{dist}(x, A)$ stetig ist, gibt es nach Satz 7 einen Punkt $q \in K$ mit

$$\text{dist}(q, A) = \text{dist}(K, A).$$

Da A abgeschlossen ist, gibt es ein $\varepsilon > 0$ mit $B_\varepsilon(q) \subset X \setminus A$. Daraus folgt $\text{dist}(q, A) \geq \varepsilon$.

Bemerkung. Sind A_1 und A_2 punktfremde abgeschlossene Teilmengen eines metrischen Raumes X , so ist nicht notwendig $\text{dist}(A_1, A_2) > 0$, wie folgendes Beispiel zeigt: In $X = \mathbb{R}^2$ sei

$$A_1 := \{(x, y) \in \mathbb{R}^2 : xy = 0\}$$

das Achsenkreuz und A_2 die Hyperbel

$$A_2 := \{(x, y) \in \mathbb{R}^2 : xy = 1\}$$

Dann gilt $A_1 \cap A_2 = \emptyset$, aber $\text{dist}(A_1, A_2) = 0$.

Als eine weitere Anwendung von Satz 7 beweisen wir nun den sog. *Fundamentalsatz der Algebra*.

Satz 8. *Sei $n \geq 1$ und*

$$P(z) := \sum_{k=0}^n c_k z^k, \quad c_k \in \mathbb{C}, c_n \neq 0,$$

ein Polynom n -ten Grades mit komplexen Koeffizienten. Dann besitzt P mindestens eine komplexe Nullstelle, d.h. es existiert ein $z_0 \in \mathbb{C}$ mit $P(z_0) = 0$.

Beweis. Wir betrachten die stetige Funktion

$$g: \mathbb{C} \cong \mathbb{R}^2 \rightarrow \mathbb{R}, \quad z = x + iy \mapsto g(z) := |P(z)|.$$

Da $\lim_{|z| \rightarrow \infty} |P(z)| = \infty$, existiert ein $R > 0$, so dass

$$g(z) = |P(z)| > |c_0| \quad \text{für alle } |z| > R.$$

Da die Kreisscheibe $K := \{z \in \mathbb{C} : |z| \leq R\}$ kompakt ist, nimmt g sein Minimum auf K in einem gewissen Punkt $z_0 \in K$ an. Da $g(0) = |c_0|$, und g außerhalb von K überall Werte $> |c_0|$ annimmt, wird in z_0 sogar das absolute Minimum der Funktion g auf \mathbb{C} angenommen.

i) Falls $g(z_0) = 0$, gilt $P(z_0) = 0$ und wir sind fertig.

ii) Es bleibt also noch der Fall $g(z_0) = |P(z_0)| > 0$ zu betrachten. Wir werden zeigen, dass dieser Fall nicht auftreten kann. O.B.d.A. können wir annehmen, dass $P(z_0) = 1$ (andernfalls multipliziere man P mit $P(z_0)^{-1}$). Wir führen die Variablen-Substitution $\zeta := z - z_0$ durch. Da

$$z^k = (z_0 + \zeta)^k = \sum_{v=0}^k \binom{k}{v} z_0^{k-v} \zeta^v,$$

ist

$$P(z_0 + \zeta) =: Q(\zeta) = 1 + \sum_{k=1}^n a_k \zeta^k$$

wieder ein Polynom n -ten Grades in ζ , das bei $\zeta = 0$ das absolute Minimum

seines Betrages annimmt. Sei $m \geq 1$ minimal mit $a_m \neq 0$. Damit ist

$$Q(\zeta) = 1 + a_m \zeta^m + R(\zeta) \quad \text{mit} \quad R(\zeta) = \sum_{k=m+1}^n a_k \zeta^k.$$

(Falls $m = n$, ist $R = 0$.) Mit $M := \sum_{k=m+1}^n |a_k|$ gilt

$$|R(\zeta)| \leq M |\zeta|^{m+1} \quad \text{für} \quad |\zeta| \leq 1.$$

Wir können den Koeffizienten a_m schreiben als

$$a_m = A e^{i\alpha} \quad \text{mit} \quad \alpha \in \mathbb{R} \text{ und } A := |a_m| > 0.$$

Sei jetzt speziell

$$\zeta = r e^{i\varphi} \quad \text{mit} \quad r > 0 \text{ und } \varphi := \frac{\pi - \alpha}{m}.$$

Damit ist $a_m \zeta^m = A r^m e^{i\alpha + i m \varphi} = A r^m e^{i\pi} = -A r^m$ und

$$|Q(\zeta)| = |1 - A r^m + R(\zeta)| \leq |1 - A r^m| + M r^{m+1} < 1 = Q(0),$$

falls $r > 0$ genügend klein ist. Dies steht aber im Widerspruch dazu, dass Q an der Stelle 0 sein Minimum annimmt. Der Fall ii) kann also nicht auftreten und Satz 8 ist bewiesen.

Corollar. Jedes Polynom n -ten Grades ($n \geq 1$) mit komplexen Koeffizienten

$$P(z) = z^n + c_{n-1} z^{n-1} + \dots + c_1 z + c_0$$

lässt sich in ein Produkt von Linearfaktoren zerlegen:

$$P(z) = (z - z_1)(z - z_2) \cdot \dots \cdot (z - z_n) \quad \text{mit } z_k \in \mathbb{C}.$$

Beweis durch Induktion nach n . Der Induktions-Anfang $n = 1$ ist trivial.

Induktions-Schritt $(n-1) \rightarrow n$. Nach Satz 8 existiert ein $z_1 \in \mathbb{C}$ mit $P(z_1) = 0$. Mit der Variablen-Substitution $\zeta = z - z_1$ ist $Q(\zeta) := P(z_1 + \zeta)$ ein Polynom in ζ vom Grad n mit $Q(0) = 0$. Daher verschwindet das konstante Glied von Q und es gilt $Q(\zeta) = \zeta \cdot Q_1(\zeta)$ mit einem Polynom Q_1 vom Grad $n-1$. Daraus folgt

$$P(z) = (z - z_1) P_1(z) \quad \text{mit } P_1(z) := Q_1(z - z_1).$$

Auf P_1 lässt sich die Induktions-Voraussetzung anwenden und das Corollar ist bewiesen.

Bemerkung. Da in einem Körper ein Produkt von Elementen genau dann gleich null ist, wenn einer der Faktoren verschwindet, gilt

$$P(z) = (z - z_1)(z - z_2) \cdots (z - z_n) = 0$$

genau dann, wenn $z = z_j$ für ein $j \in \{1, 2, \dots, n\}$. Fasst man gleiche Linearfaktoren zusammen, so erhält man eine Darstellung der Form

$$P(z) = (z - z_1)^{k_1} \cdots (z - z_r)^{k_r}, \quad k_j \geq 1, \quad k_1 + \dots + k_r = n,$$

mit paarweise verschiedenen $z_1, \dots, z_r \in \mathbb{C}$. Man nennt dann z_j eine Nullstelle von P der *Vielfachheit* k_j .

Das Folgende ist eine Verallgemeinerung der entsprechenden Aussage im \mathbb{R}^1 , vgl. An. 1, §5, Satz 6.

Satz 9 (Bolzano-Weierstraß). *Sei A eine kompakte Teilmenge eines metrischen Raumes X und $(x_n)_{n \in \mathbb{N}}$ eine Folge von Punkten $x_n \in A$. Dann gibt es eine Teilfolge $(x_{n_k})_{k \in \mathbb{N}}$, die gegen einen Punkt $a \in A$ konvergiert.*

Beweis. Angenommen, keine Teilfolge von $(x_n)_{n \in \mathbb{N}}$ konvergiert gegen einen Punkt von A . Dann besitzt jeder Punkt $a \in A$ eine offene Umgebung U_a , in der nur endlich viele Folgenglieder liegen (lägen in jeder Umgebung von a unendlich viele Folgenglieder, könnte man eine Teilfolge konstruieren, die gegen a konvergiert). Trivialerweise gilt

$$A \subset \bigcup_{a \in A} U_a.$$

Da A kompakt ist, gibt es endlich viele Punkte $a_1, \dots, a_m \in A$ mit

$$A \subset U_{a_1} \cup \dots \cup U_{a_m}.$$

Dann lägen aber in A nur endlich viele Folgenglieder, Widerspruch!

Corollar. *Jede beschränkte Folge $(x_i)_{i \in \mathbb{N}}$ im \mathbb{R}^n besitzt eine konvergente Teilfolge.*

Beweis. Dies folgt aus Satz 9 und Satz 2, da die beschränkte Folge (x_i) in einem genügend großen abgeschlossenen Quader enthalten ist.

Definition (gleichmäßige Stetigkeit). Seien X, Y metrische Räume. Eine Abbildung $f: X \rightarrow Y$ heißt *gleichmäßig stetig*, wenn zu jedem $\varepsilon > 0$ ein $\delta > 0$

existiert, so dass

$$\|f(x), f(x')\| < \varepsilon \quad \text{für alle } x, x' \in X \text{ mit } \|x, x'\| < \delta.$$

Satz 10. *Seien X, Y metrische Räume und sei X kompakt. Dann ist jede stetige Abbildung $f: X \rightarrow Y$ gleichmäßig stetig.*

Beweis. Sei $\varepsilon > 0$ vorgegeben. Dann gibt es zu jedem $x \in X$ ein $\delta(x) > 0$, so dass

$$\|f(x), f(x')\| < \frac{\varepsilon}{2} \quad \text{für alle } x' \in B_{\delta(x)}(x).$$

Da

$$\bigcup_{x \in X} B_{\delta(x)/2}(x) = X$$

und X kompakt ist, gibt es endlich viele Punkte $x_1, \dots, x_k \in X$ mit

$$B_{\delta(x_1)/2}(x_1) \cup \dots \cup B_{\delta(x_k)/2}(x_k) = X.$$

Wir setzen

$$\delta := \min(\delta(x_1)/2, \dots, \delta(x_k)/2).$$

Sind jetzt $x, x' \in X$ zwei beliebige Punkte mit $\|x, x'\| < \delta$, so gibt es ein $j \in \{1, \dots, k\}$ mit

$$x \in B_{\delta(x_j)/2}(x_j), \quad \text{und deshalb} \quad x' \in B_{\delta(x_j)}(x_j).$$

Es folgt

$$\|f(x_j), f(x)\| < \frac{\varepsilon}{2} \quad \text{und} \quad \|f(x_j), f(x')\| < \frac{\varepsilon}{2},$$

also $\|f(x), f(x')\| < \varepsilon$, q.e.d.

Bemerkung. Der Satz, dass jede stetige Funktion $f: I \rightarrow \mathbb{R}$ auf einem kompakten Intervall $I \subset \mathbb{R}$ gleichmäßig stetig ist (An. 1, § 11, Satz 4), ist ein Spezialfall von Satz 10.

AUFGABEN

3.1. Man zeige, dass die Vereinigung von endlich vielen kompakten Teilmengen eines Hausdorff-Raumes wieder kompakt ist.

3.2. Sei (X, d) ein metrischer Raum mit der trivialen Metrik, vgl. Beispiel (1.4),

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

Man zeige: Eine Teilmenge $A \subset X$ ist genau dann kompakt, wenn sie endlich ist.

3.3. Sei A eine Teilmenge von \mathbb{R}^n . Zu jeder Folge $(x_i)_{i \in \mathbb{N}}$ von Punkten $x_i \in A$ gebe es eine Teilfolge, die gegen einen Punkt $a \in A$ konvergiert. Man zeige, dass A kompakt ist. (Vgl. dazu Satz 9.)

3.4. Sei X ein metrischer Raum, $Y \subset X$ und $x \in X \setminus Y$ ein Punkt mit $\text{dist}(Y, x) = 0$. Man zeige, dass x ein Randpunkt von Y ist.

3.5. Sei K eine kompakte Teilmenge eines metrischen Raumes X und $(U_i)_{i \in I}$ eine offene Überdeckung von K . Man beweise: Es gibt eine Zahl $\lambda > 0$ mit folgender Eigenschaft: Zu jeder Teilmenge $A \subset K$ mit $\text{diam}(A) \leq \lambda$ existiert ein $i \in I$ mit $A \subset U_i$ (*Lebesguesches Lemma*).

3.6. Seien X, Y Hausdorff-Räume, X kompakt und $f: X \rightarrow Y$ eine stetige bijektive Abbildung. Man beweise: Die Umkehrabbildung $f^{-1}: Y \rightarrow X$ ist stetig, d.h. f ist ein Homöomorphismus.

3.7. Man beweise: Jeder kompakte metrische Raum ist vollständig.

3.8. Seien K und L kompakte Teilmengen von \mathbb{R}^n . Man zeige, dass dann auch die Menge $K + L := \{x + y : x \in K, y \in L\}$ kompakt ist.

3.9. Seien $I, J \subset \mathbb{R}$ kompakte Intervalle und $f: I \times J \rightarrow \mathbb{R}$ eine stetige Funktion. Die Funktion $F: I \rightarrow \mathbb{R}$ werde definiert durch

$$F(x) := \sup\{f(x, y) : y \in J\}.$$

Man zeige, dass F stetig ist.

3.10. Eine Funktion $f: X \rightarrow \mathbb{R}$ auf dem topologischen Raum X heißt *halbstetig von unten* (bzw. *von oben*), wenn für jedes $c \in \mathbb{R}$ die Menge $\{x \in X : f(x) > c\}$ (bzw. $\{x \in X : f(x) < c\}$) offen in X ist. Man beweise: Ist X kompakt, so nimmt jede von unten (oben) halbstetige Funktion $f: X \rightarrow \mathbb{R}$ ihr Minimum (Maximum) an.

§ 4 Kurven im \mathbb{R}^n

Nach den bisherigen abstrakten Überlegungen gehen wir jetzt wieder zur Untersuchung konkreter geometrischer Gebilde über, nämlich von Kurven im \mathbb{R}^n . Wir definieren Kurventangenten, Schnittwinkel von Kurven und behandeln den Begriff der Bogenlänge und ihre Berechnung.

In diesem Paragraphen setzen wir immer voraus, dass alle Intervalle aus mehr als einem Punkt bestehen.

Definition. Unter einer *Kurve* im \mathbb{R}^n versteht man eine stetige Abbildung

$$f: I \rightarrow \mathbb{R}^n,$$

wobei $I \subset \mathbb{R}$ ein (eigentliches oder uneigentliches) Intervall ist.

Nach § 2, Satz 6, wird die Kurve f gegeben durch ein n -tupel

$$f = (f_1, f_2, \dots, f_n)$$

stetiger Funktionen $f_k: I \rightarrow \mathbb{R}$, $k = 1, 2, \dots, n$. Die Kurve heißt *differenzierbar* (bzw. *stetig differenzierbar*), wenn alle Funktionen f_k differenzierbar bzw. stetig differenzierbar sind.

Beispiele

(4.1) Sei $r > 0$. Ein *Kreis* vom Radius r in der Ebene wird beschrieben durch die Kurve

$$f: [0, 2\pi] \rightarrow \mathbb{R}^2, \quad t \mapsto (r \cos t, r \sin t).$$

(4.2) Sei $a \in \mathbb{R}^n$ und $v \in \mathbb{R}^n \setminus \{0\}$. Die Abbildung

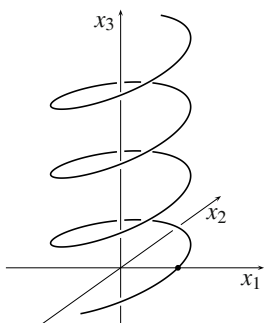
$$f: \mathbb{R} \rightarrow \mathbb{R}^n, \quad t \mapsto a + vt$$

beschreibt eine *Gerade* im \mathbb{R}^n durch den Punkt a mit Richtungsvektor v .

(4.3) Seien $r > 0$ und $c \neq 0$ reelle Zahlen. Die Kurve

$$f: \mathbb{R} \rightarrow \mathbb{R}^3, \quad t \mapsto (r \cos t, r \sin t, ct)$$

ist eine *Schraubenlinie* (Bild 4.1)

**Bild 4.1** Schraubenlinie

(4.4) Sei $\varphi : I \rightarrow \mathbb{R}$ eine stetige reellwertige Funktion auf dem Intervall $I \subset \mathbb{R}$. Der Graph dieser Funktion kann als Kurve im \mathbb{R}^2 aufgefasst werden:

$$f : I \longrightarrow \mathbb{R}^2, \quad t \mapsto (t, \varphi(t)).$$

Bemerkung. Manchmal ist folgende kinematische Interpretation einer Kurve $f : I \rightarrow \mathbb{R}^n$ nützlich: Man fasst die Variable $t \in I$ als Zeit und $f(t) \in \mathbb{R}^n$ als Ort auf. Die Kurve beschreibt dann die zeitliche Bewegung eines Punktes im \mathbb{R}^n .

Definition (Tangentialvektor). Sei $I \subset \mathbb{R}$ ein Intervall und

$$f = (f_1, \dots, f_n) : I \longrightarrow \mathbb{R}^n$$

eine differenzierbare Kurve. Für $t \in I$ heißt

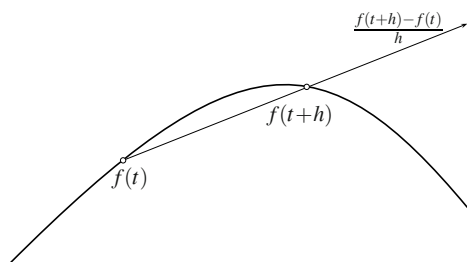
$$f'(t) = (f'_1(t), \dots, f'_n(t)) \in \mathbb{R}^n$$

der *Tangentialvektor* der Kurve f zum Parameterwert t . Falls $f'(t) \neq 0$, heißt der auf den Betrag 1 normierte Vektor $f'(t)/\|f'(t)\|$ *Tangenten-Einheitsvektor*.

Geometrische Interpretation. Der Tangentialvektor $f'(t)$ lässt sich als Limes von Sekanten auffassen, denn

$$f'(t) = \lim_{\substack{h \rightarrow 0 \\ h \neq 0}} \frac{f(t+h) - f(t)}{h},$$

vgl. Bild 4.2.

**Bild 4.2**

Physikalische Interpretation. $f'(t)$ ist der Geschwindigkeitsvektor im Zeitpunkt t der durch $f : I \rightarrow \mathbb{R}^n$ beschriebenen Bewegung (Grenzwert des Quotienten aus Ortsdifferenz und Zeitdifferenz) und

$$\|f'(t)\| = \sqrt{|f'_1(t)|^2 + \dots + |f'_n(t)|^2}$$

ist der Betrag der Geschwindigkeit.

(4.5) Bemerkung. Eine Kurve $f : I \rightarrow \mathbb{R}^n$ braucht nicht notwendig eine injektive Abbildung darzustellen. Gilt $f(t_1) = f(t_2) =: x$ für $t_1 \neq t_2$, so heißt x *Doppelpunkt* der Kurve. Im Punkt x hat dann f i.Allg. zwei verschiedene Tangentialvektoren. Als Beispiel betrachten wir die Kurve $f : \mathbb{R} \rightarrow \mathbb{R}^2$

$$f(t) := (t^2 - 1, t^3 - t).$$

Es gilt, wie man sich leicht überzeugt

$$f(\mathbb{R}) = \{(x, y) \in \mathbb{R}^2 : y^2 = x^3 + x^2\},$$

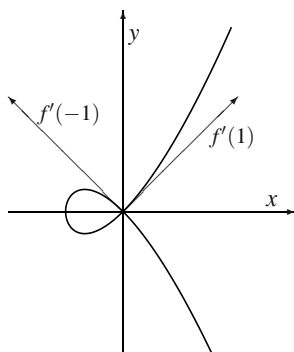
vgl. Bild 4.3. Die Kurve hat einen Doppelpunkt für die Parameterwerte $t = \pm 1$, denn es gilt

$$f(1) = f(-1) = (0, 0).$$

Da $f'(t) = (2t, 3t^2 - 1)$, errechnet man für die Tangenten im Doppelpunkt

$$f'(-1) = (-2, 2), \quad f'(1) = (2, 2).$$

Definition. Sei $f : I \rightarrow \mathbb{R}^n$ eine stetig differenzierbare Kurve. Die Kurve heißt *regulär* oder *nicht-singulär*, falls $f'(t) \neq 0$ für alle $t \in I$. Ein Parameterwert $t \in I$ mit $f'(t) = 0$ heißt *singulär*.

**Bild 4.3**

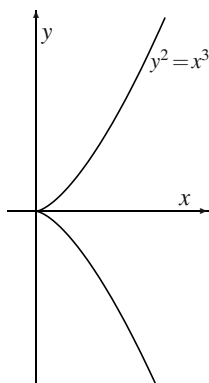
(4.6) Beispiel. Wir betrachten die *Neilsche Parabel*

$$f: \mathbb{R} \longrightarrow \mathbb{R}^2, \quad t \mapsto (t^2, t^3).$$

Für das Bild der Kurve gilt

$$f(\mathbb{R}) = \{(x, y) \in \mathbb{R}^2 : x \geq 0, y = \pm x^{3/2}\},$$

vgl. Bild 4.4. Wegen $f'(t) = (2t, 3t^2)$ liegt für $t = 0$ in der Spitze der einzige singuläre Punkt der Neilschen Parabel vor.

**Bild 4.4**

Schnittwinkel

Seien $f : I_1 \rightarrow \mathbb{R}^n$ und $g : I_2 \rightarrow \mathbb{R}^n$ zwei reguläre Kurven. Für die Parameterwerte $t_1 \in I_1$ und $t_2 \in I_2$ gelte $f(t_1) = g(t_2)$. Dann versteht man unter dem *Schnittwinkel* ϑ der Kurven f und g bei den Parameterwerten t_1, t_2 den Winkel zwischen den Tangentialvektoren $f'(t_1)$ und $g'(t_2)$, siehe Bild 4.5. Der Winkel ϑ ist also bestimmt durch die Gleichung

$$\cos \vartheta = \frac{\langle f'(t_1), g'(t_2) \rangle}{\|f'(t_1)\| \cdot \|g'(t_2)\|} \quad \text{mit } 0 \leq \vartheta \leq \pi.$$

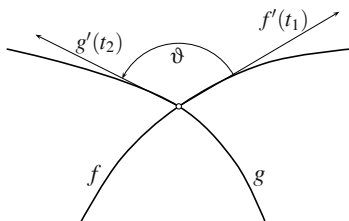


Bild 4.5

Bogenlänge

Sei $[a, b] \subset \mathbb{R}$, $a < b$ ein abgeschlossenes Intervall und $f : [a, b] \rightarrow \mathbb{R}^n$ eine Kurve. Unterteilt man das Intervall

$$a = t_0 < t_1 < \dots < t_k = b$$

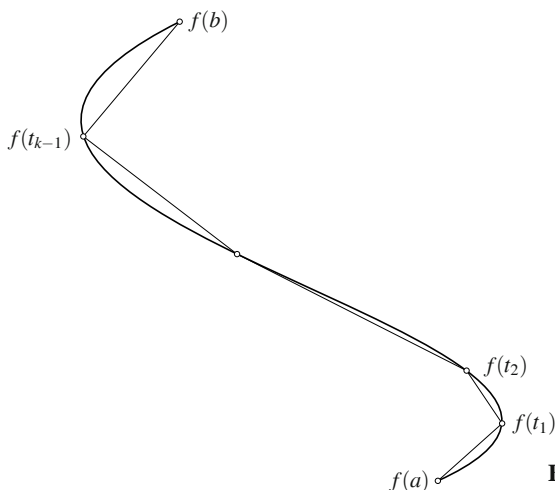
und verbindet die Punkte $f(t_{i-1})$ mit $f(t_i)$, für $i = 1, 2, \dots, k$, geradlinig, so erhält man einen Polygonzug im \mathbb{R}^n , siehe Bild 4.6. Die Länge dieses Polygonzugs ist gleich

$$p_f(t_0, \dots, t_k) := \sum_{i=1}^k \|f(t_i) - f(t_{i-1})\|.$$

Die Länge der Kurve wird nun definiert als Grenzwert der Längen der Polygonzüge bei immer feineren Unterteilungen.

Definition. Eine Kurve $f : [a, b] \rightarrow \mathbb{R}^n$ heißt *rektifizierbar* mit der Länge L , wenn zu jedem $\varepsilon > 0$ ein $\delta > 0$ existiert, so dass für jede Unterteilung

$$a = t_0 < t_1 < \dots < t_k = b$$

**Bild 4.6**

der Feinheit $< \delta$ gilt

$$|p_f(t_0, \dots, t_k) - L| < \varepsilon.$$

Satz 1. Jede stetig differenzierbare Kurve $f : [a, b] \rightarrow \mathbb{R}^n$ ist rektifizierbar, und für ihre Länge L gilt

$$L = \int_a^b \|f'(t)\| dt.$$

Bemerkung. Die stetig Differenzierbarkeit ist nicht notwendig dafür, dass eine Kurve rektifizierbar ist. Es gibt jedoch stetige Kurven, die nicht rektifizierbar sind.

Zum Beweis von Satz 1 benötigen wir einen Hilfssatz.

Hilfssatz. Sei $f : [a, b] \rightarrow \mathbb{R}^n$ stetig differenzierbar. Dann gibt es zu jedem $\varepsilon > 0$ ein $\delta > 0$, so dass

$$\left\| \frac{f(t) - f(\tau)}{t - \tau} - f'(t) \right\| \leq \varepsilon$$

für alle $t, \tau \in [a, b]$ mit $0 < |t - \tau| \leq \delta$.

Beweis.

a) Wir behandeln zunächst den Fall $n = 1$. Die Ableitung $f' : [a, b] \rightarrow \mathbb{R}$ ist nach Voraussetzung stetig, also sogar gleichmäßig stetig. Zu $\varepsilon > 0$ gibt es also ein $\delta > 0$, so dass

$$|f'(t) - f'(s)| \leq \varepsilon \quad \text{für alle } t, s \text{ mit } |t - s| \leq \delta.$$

Sei nun $t, \tau \in [a, b]$ mit $0 < |t - \tau| \leq \delta$. Nach dem Mittelwertsatz gibt es ein s zwischen t und τ , so dass

$$\frac{f(t) - f(\tau)}{t - \tau} = f'(s).$$

Also ist

$$\left| \frac{f(t) - f(\tau)}{t - \tau} - f'(t) \right| = |f'(s) - f'(t)| \leq \varepsilon.$$

b) Sei jetzt n beliebig und $f = (f_1, \dots, f_n)$. Da nach (1.7)

$$\left\| \frac{f(t) - f(\tau)}{t - \tau} - f'(t) \right\| \leq \sqrt{n} \cdot \max_{i=1, \dots, n} \left| \frac{f_i(t) - f_i(\tau)}{t - \tau} - f'_i(t) \right|,$$

folgt die Behauptung aus Teil a).

Beweis von Satz 1. Sei $\varepsilon > 0$ vorgegeben. Nach dem Satz über die Approximation von Integralen durch Riemannsche Summen (An. 1, § 18, Satz 8) existiert ein $\delta_1 > 0$, so dass

$$\left| \int_a^b \|f'(t)\| dt - \sum_{i=1}^k \|f'(t_i)\| (t_i - t_{i-1}) \right| \leq \frac{\varepsilon}{2} \quad (1)$$

für jede Unterteilung

$$a = t_0 < t_1 < \dots < t_k = b \quad (*)$$

des Intervalls $[a, b]$ der Feinheit $\leq \delta_1$. Nach dem Hilfssatz gibt es ein $\delta > 0$ mit $\delta \leq \delta_1$ und folgender Eigenschaft: Hat die Unterteilung $(*)$ eine Feinheit $\leq \delta$, so gilt

$$\left\| \frac{f(t_i) - f(t_{i-1})}{t_i - t_{i-1}} - f'(t_i) \right\| \leq \frac{\varepsilon}{2(b-a)}$$

für $i = 1, \dots, k$. Daraus folgt

$$\left| \|f(t_i) - f(t_{i-1})\| - \|f'(t_i)\| (t_i - t_{i-1}) \right| \leq \frac{t_i - t_{i-1}}{b-a} \cdot \frac{\varepsilon}{2}. \quad (2)$$

Aus (1) und (2) folgt für jede Unterteilung (*) der Feinheit $\leq \delta$

$$\left| \sum_{i=1}^k \|f(t_i) - f(t_{i-1})\| - \int_a^b \|f'(t)\| dt \right| \leq \varepsilon, \quad \text{q.e.d.}$$

Beispiele

(4.7) Sei $\varphi > 0$. Wir betrachten den Kreisbogen

$$f: [0, \varphi] \longrightarrow \mathbb{R}^2, \quad f(t) := (\cos t, \sin t).$$

Es gilt $f'(t) = (-\sin t, \cos t)$, also

$$\|f'(t)\| = \sqrt{\sin^2 t + \cos^2 t} = 1.$$

Deshalb errechnet man für die Bogenlänge

$$L = \int_0^\varphi \|f'(t)\| dt = \int_0^\varphi dt = \varphi.$$

Insbesondere ist der Umfang des Einheitskreises gleich 2π .

(4.8) Die *Zykloide* ist die Kurve

$$f: \mathbb{R} \longrightarrow \mathbb{R}^2, \quad t \mapsto (t - \sin t, 1 - \cos t).$$

Die Zykloide beschreibt die Bahn eines Punktes auf der Peripherie eines Kreises vom Radius 1, der auf der x -Achse der x - y -Ebene abrollt, siehe Bild 4.7. Wir wollen die Länge L des Teils der Zykloide berechnen, der zu den Parameterwerten $0 \leq t \leq 2\pi$ gehört, also den Bogen ABC in Bild 4.7. Es ist

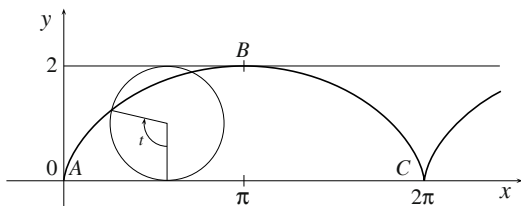


Bild 4.7

$$f'(t) = (1 - \cos t, \sin t),$$

$$\|f'(t)\|^2 = (1 - \cos t)^2 + \sin^2 t$$

$$\begin{aligned}
 &= 1 - 2 \cos t + \cos^2 t + \sin^2 t \\
 &= 2 - 2 \cos t = 4 \sin^2(t/2),
 \end{aligned}$$

also $\|f'(t)\| = 2|\sin(t/2)|$. Damit wird

$$L = \int_0^{2\pi} 2 \left| \sin \frac{t}{2} \right| dt = 4 \int_0^\pi \sin u du = 8.$$

Es ist bemerkenswert, dass sich für die Bogenlänge eine rationale Zahl ergibt.

Parametertransformationen

Sei $f: [a, b] \rightarrow \mathbb{R}^n$ eine Kurve, $[\alpha, \beta] \subset \mathbb{R}$ ein weiteres Intervall und

$$\varphi: [\alpha, \beta] \longrightarrow [a, b]$$

eine bijektive stetige Abbildung. Dann ist die zusammengesetzte Abbildung

$$g := f \circ \varphi: [\alpha, \beta] \longrightarrow \mathbb{R}^n$$

wieder eine Kurve im \mathbb{R}^n . Man sagt, dass die Kurve g aus der Kurve f durch die *Parametertransformation* φ hervorgeht. Sind sowohl φ als auch

$$\varphi^{-1}: [a, b] \rightarrow [\alpha, \beta]$$

stetig differenzierbar, so nennt man φ eine C^1 -Parametertransformation.

Die Kurvenpunkte von f und g sind dieselben, denn es gilt $g(t) = f(\varphi(t))$ für alle $t \in [\alpha, \beta]$, aber sie werden i.Allg. verschieden durchlaufen.

Da $\varphi: [\alpha, \beta] \rightarrow [a, b]$ stetig und bijektiv ist, tritt genau einer der beiden folgenden Fälle auf:

- 1) φ ist auf $[\alpha, \beta]$ streng monoton wachsend. Man nennt die Parametertransformation dann *orientierungstreu*.
- 2) φ ist auf $[\alpha, \beta]$ streng monoton fallend. Dann heißt φ *orientierungsumkehrend*.

Ist φ eine C^1 -Parametertransformation, so ist $\varphi'(t) \neq 0$ für alle $t \in [\alpha, \beta]$, denn aus $\varphi^{-1} \circ \varphi = \text{id}$ folgt mit der Kettenregel

$$(\varphi^{-1})'(\varphi(t)) \cdot \varphi'(t) = 1.$$

φ ist genau dann orientierungstreu, wenn $\varphi'(t) > 0$ für alle t und orientierungsumkehrend, wenn $\varphi'(t) < 0$ für alle t .

Wir untersuchen nun noch das Verhalten von Tangentialvektoren, Schnittwinkel und Bogenlänge bei Parametertransformationen.

a) *Tangentialvektoren.* Sei $f : [a, b] \rightarrow \mathbb{R}^n$ eine differenzierbare Kurve und $\varphi : [\alpha, \beta] \rightarrow [a, b]$ eine C^1 -Parametertransformation. Dann gilt für die Kurve $g = f \circ \varphi : [\alpha, \beta] \rightarrow \mathbb{R}^n$

$$g'(\tau) = f'(\varphi(\tau))\varphi'(\tau).$$

Die Tangentialvektoren an die Kurve g zum Parameterwert $\tau \in [\alpha, \beta]$ und an die Kurve f zum Parameterwert $t := \varphi(\tau) \in [a, b]$ unterscheiden sich also nur um den skalaren Faktor $\varphi'(\tau) \in \mathbb{R}^*$. Die zugehörigen Tangenten-Einheitsvektoren sind also gleich oder entgegengesetzt gleich, je nachdem φ orientierungstreu oder orientierungsumkehrend ist.

b) *Schnittwinkel.* Seien $f_i : [a_i, b_i] \rightarrow \mathbb{R}^n$ zwei reguläre Kurven ($i = 1, 2$). Für $t_i \in [a_i, b_i]$ gelte $f_1(t_1) = f_2(t_2)$ und ϑ sei der Schnittwinkel von f_1 und f_2 bei den Parameterwerten t_1 bzw. t_2 .

Seien nun $\varphi_i : [\alpha_i, \beta_i] \rightarrow [a_i, b_i]$ zwei C^1 -Parametertransformationen und $g_i := f_i \circ \varphi_i$ die transformierten Kurven. Für $\tau_i := \varphi_i^{-1}(t_i)$ gilt dann $g_1(\tau_1) = g_2(\tau_2)$. Wir bezeichnen mit ϑ' den Schnittwinkel von g_1 und g_2 bei den Parameterwerten τ_1 bzw. τ_2 . Aus a) folgt:

- i) $\vartheta' = \vartheta$, falls φ_1 und φ_2 beide orientierungstreu oder beide orientierungsumkehrend sind.
- ii) $\vartheta' = \pi - \vartheta$, falls von den Parametertransformationen φ_1, φ_2 eine orientierungstreu und die andere orientierungsumkehrend ist.

c) *Bogenlänge.* Sei $f : [a, b] \rightarrow \mathbb{R}^n$ eine stetig differenzierbare Kurve, $\varphi : [\alpha, \beta] \rightarrow [a, b]$ eine C^1 -Parametertransformation und

$$g := f \circ \varphi : [\alpha, \beta] \rightarrow \mathbb{R}^n.$$

Dann gilt

$$\int_a^b \|f'(t)\| dt = \int_\alpha^\beta \|g'(\tau)\| d\tau,$$

d.h. f und g haben dieselbe Bogenlänge. Dies folgt aus der Substitutionsregel. Wir führen den Beweis für den (schwierigeren) Fall durch, dass φ orientierungsumkehrend ist, d.h. $\varphi'(\tau) < 0$ für alle $\tau \in [\alpha, \beta]$. Dann ist $\varphi(\alpha) = b$ und $\varphi(\beta) = a$, also

$$\begin{aligned} \int_\alpha^\beta \|g'(\tau)\| d\tau &= \int_\alpha^\beta \|f'(\varphi(\tau))\varphi'(\tau)\| d\tau \\ &= - \int_\alpha^\beta \|f'(\varphi(\tau))\| |\varphi'(\tau)| d\tau \end{aligned}$$

$$= - \int_b^a \|f'(t)\| dt = \int_a^b \|f'(t)\| dt, \quad \text{q.e.d.}$$

Bemerkung. Die Invarianz der Bogenlänge kann man auch direkt aus der Definition als Grenzwert der Längen eingeschriebener Polygonzüge folgern.

AUFGABEN

4.1. Seien $a, b, c, r \in \mathbb{R}$ mit $a < b$, $r > 0$. Man berechne die Bogenlänge der Kurve

$$f: [a, b] \rightarrow \mathbb{R}^3, \quad f(t) := (r \cos t, r \sin t, ct).$$

4.2. Sei $c \in \mathbb{R}^*$ und

$$f: \mathbb{R} \rightarrow \mathbb{R}^2, \quad f(t) := (e^{ct} \cos t, e^{ct} \sin t).$$

Die Kurve f heißt *logarithmische Spirale*.

a) Man skizziere die Kurve für $c = \frac{1}{2\pi}$ im Bereich $-\pi \leq t \leq \pi$.

b) Für $[a, b] \subset \mathbb{R}$ sei $L_{a,b}$ die Bogenlänge der Kurve $f|_{[a,b]}$. Man berechne $L_{a,b}$.

c) Existiert $\lim_{a \rightarrow -\infty} L_{a,0}$?

d) Man zeige, dass die logarithmische Spirale jeden Kreis um den Nullpunkt in genau einem Punkt schneidet und berechne den Schnittwinkel.

4.3. a) Man zeige, dass für jedes $k \in [0, 1]$ das uneigentliche Integral

$$E(k) = \int_0^1 \frac{\sqrt{1 - k^2 t^2}}{\sqrt{1 - t^2}} dt$$

existiert. $E(k)$ heißt vollständiges elliptisches Integral zweiter Gattung.

b) Man drücke die Bogenlänge der Ellipse

$$f: [0, 2\pi] \rightarrow \mathbb{R}^2, \quad t \mapsto (a \cos t, b \sin t),$$

mit den Halbachsen $a, b \in \mathbb{R}_+^*$ mit Hilfe von $E(k)$ aus.

§ 5 Partielle Ableitungen

In diesem Paragraphen definieren wir die partiellen Ableitungen von Funktionen mehrerer Veränderlichen. Die partiellen Ableitungen sind nichts anderes als die gewöhnlichen Ableitungen von Funktionen einer Veränderlichen, die man erhält, wenn man alle Veränderlichen bis auf eine festhält. Mithilfe der partiellen Ableitungen werden wichtige Differential-Operatoren wie Gradient, Divergenz, Rotation und Laplace-Operator definiert.

Die im letzten Paragraphen besprochenen Kurven waren Abbildungen von Teilmengen von \mathbb{R} in den \mathbb{R}^n . Wir betrachten jetzt umgekehrt Abbildungen von Teilmengen $U \subset \mathbb{R}^n$ nach \mathbb{R} ,

$$f : U \longrightarrow \mathbb{R}, \quad (x_1, \dots, x_n) \mapsto f(x_1, \dots, x_n).$$

Der *Graph* von f ist die Menge

$$\Gamma_f := \{(x, y) \in U \times \mathbb{R} : y = f(x)\} \subset \mathbb{R}^{n+1}.$$

Im Fall $n = 2$ kann man sich den Graphen von f als Fläche im dreidimensionalen Raum vorstellen (Bild 5.1).

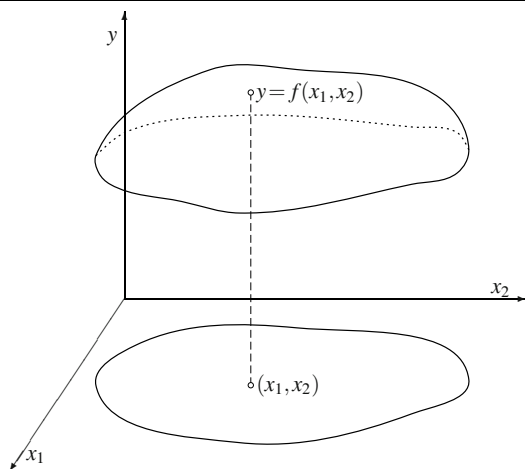
Eine Funktion $f : U \rightarrow \mathbb{R}$ ist auch festgelegt durch die Schar $N_f(c)$, $c \in \mathbb{R}$, ihrer *Niveaumengen*. Dabei ist

$$N_f(c) := \{x \in U : f(x) = c\} \subset \mathbb{R}^n$$

die Menge der Punkte, in denen f den Wert c annimmt. Im Fall $n = 2$ nennt man die Niveaumengen auch *Höhenlinien* (Bild 5.2). Man beachte jedoch, dass für eine beliebige Funktion zweier Veränderlichen die Niveaumengen keine "Linien" im anschaulichen Sinn zu sein brauchen. Wir werden auf diese Frage noch einmal in § 8 zurückkommen.

Definition (Partielle Ableitung). Sei $U \subset \mathbb{R}^n$ eine offene Teilmenge und

$$f : U \longrightarrow \mathbb{R}$$

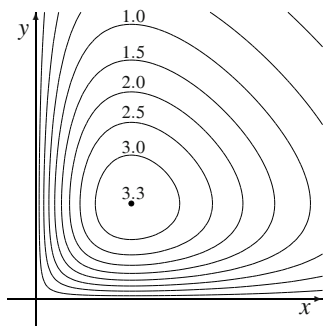
**Bild 5.1**

eine reelle Funktion. f heißt im Punkt $x \in U$ *partiell differenzierbar* in der i -ten Koordinatenrichtung, falls der Limes

$$D_i f(x) := \lim_{h \rightarrow 0} \frac{f(x + h e_i) - f(x)}{h}$$

existiert. Dabei ist $e_i \in \mathbb{R}^n$ der i -te Einheitsvektor,

$$e_i = (0, \dots, \underbrace{0, 1, 0, \dots, 0}_{i\text{-te Stelle}})$$

**Bild 5.2** Höhenlinien

und für den Limes $h \rightarrow 0$ hat man sich auf solche $h \in \mathbb{R}$ zu beschränken, für die $h \neq 0$ und $x + he_i \in U$.

$D_i f(x)$ heißt i -te partielle Ableitung von f in x .

(Bemerkung. Damit man die i -te partielle Ableitung definieren kann, ist nicht unbedingt notwendig, dass U offen ist. Es genügt, dass wenigstens eine Folge $(h_k)_{k \in \mathbb{N}}$ mit $\lim h_k = 0$ existiert, so dass $h_k \neq 0$ und $x + h_k e_i \in U$ für alle k . Dies ist z.B. auch der Fall, wenn $U = I_1 \times \dots \times I_n$ mit abgeschlossenen Intervallen $I_k \subset \mathbb{R}$, die mehr als einen Punkt enthalten.)

Die partiellen Ableitungen einer Funktion $f : U \rightarrow \mathbb{R}$ kann man als gewöhnliche Ableitungen von Funktionen einer Veränderlichen interpretieren:

Sei $x = (x_1, \dots, x_n) \in U$ ein fester Punkt. Für $i = 1, \dots, n$ betrachten wir die Funktionen

$$\xi \mapsto f_i(\xi) := f(x_1, \dots, x_{i-1}, \xi, x_{i+1}, \dots, x_n)$$

Aus der Definition der partiellen Ableitung folgt

$$D_i f(x) = \lim_{h \rightarrow 0} \frac{f_i(x_i + h) - f_i(x_i)}{h} = f'_i(x_i).$$

Die partielle Ableitung in der i -ten Koordinatenrichtung ist also nichts anderes als die gewöhnliche Ableitung nach der i -ten Variablen bei Festhaltung der übrigen $n - 1$ Veränderlichen. Deshalb gelten für die partiellen Ableitungen analoge Rechenregeln wie für die gewöhnlichen Ableitungen.

Definition. Sei $U \subset \mathbb{R}^n$ offen. Eine Funktion $f : U \rightarrow \mathbb{R}$ heißt *partiell differenzierbar*, falls $D_i f(x)$ für alle $x \in U$ und $i = 1, \dots, n$ existiert. f heißt *stetig partiell differenzierbar*, falls zusätzlich alle partiellen Ableitungen $D_i f : U \rightarrow \mathbb{R}$ stetig sind.

Schreibweise. Statt $D_i f$ schreibt man auch $\frac{\partial f}{\partial x_i}$. Entsprechend auch

$$D_i f(x) = \frac{\partial f}{\partial x_i}(x) = \frac{\partial f(x)}{\partial x_i}.$$

Beispiele

(5.1) Als erstes einfaches Beispiel berechnen wir die partiellen Ableitungen der Funktion zweier Veränderlichen

$$F : \mathbb{R}^2 \longrightarrow \mathbb{R}, \quad (x, y) \mapsto F(x, y) := e^{x^2+y^2}$$

Da für die Funktion einer Variablen $f(x) := e^{x^2+c^2}$ (wobei c eine reelle Konstante ist), gilt $f'(x) = 2x e^{x^2+c^2}$, folgt

$$\frac{\partial}{\partial x} e^{x^2+y^2} = 2x e^{x^2+y^2}$$

und ebenso

$$\frac{\partial}{\partial y} e^{x^2+y^2} = 2y e^{x^2+y^2}$$

(5.2) Wir betrachten die Funktion $r : \mathbb{R}^n \rightarrow \mathbb{R}$,

$$r(x) := \|x\| = \sqrt{x_1^2 + \dots + x_n^2} \quad (\text{Abstand vom Nullpunkt})$$

für $x = (x_1, \dots, x_n) \in \mathbb{R}^n$. Die Niveaumengen

$$N_f(c) = \{x \in \mathbb{R}^n : r(x) = c\}$$

sind für $c > 0$ Sphären vom Radius c .

Behauptung. Die Funktion r ist in $\mathbb{R}^n \setminus 0$ partiell differenzierbar und es gilt

$$\frac{\partial r}{\partial x_i}(x) = \frac{x_i}{r(x)} \quad \text{für } x = (x_1, \dots, x_n) \neq 0.$$

Dies folgt daraus, dass die Funktion einer Variablen

$$\xi \mapsto \sqrt{x_1^2 + \dots + \xi^2 + \dots + x_n^2}$$

differenzierbar ist. Mit der Kettenregel für Funktionen einer Veränderlichen erhält man ($x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n$ sind dabei als Konstanten zu betrachten)

$$\begin{aligned} \frac{\partial r}{\partial x_i} &= \frac{\partial}{\partial x_i} (x_1^2 + \dots + x_i^2 + \dots + x_n^2)^{1/2} \\ &= \frac{1}{2} (x_1^2 + \dots + x_i^2 + \dots + x_n^2)^{-1/2} \cdot 2x_i = \frac{x_i}{r}. \end{aligned}$$

(5.3) Sei $f : \mathbb{R}_+^* \rightarrow \mathbb{R}$ eine beliebige differenzierbare Funktion. Dann ist die

zusammengesetzte Funktion $x \mapsto f(r(x))$, die meist kurz mit $f(r)$ bezeichnet wird, auf $\mathbb{R}^n \setminus 0$ definiert und dort partiell differenzierbar. Aus der Kettenregel für Funktionen einer Veränderlichen folgt

$$\frac{\partial}{\partial x_i} f(r) = f'(r) \frac{\partial r}{\partial x_i} = f'(r) \frac{x_i}{r}.$$

(5.4) Sei $n \geq 2$. Wir betrachten die wie folgt definierte Funktion $F : \mathbb{R}^n \rightarrow \mathbb{R}$

$$F(x) = \begin{cases} \frac{x_1 x_2 \cdots x_n}{r(x)^n} & \text{für } x \neq 0, \\ 0 & \text{für } x = 0. \end{cases}$$

In $\mathbb{R}^n \setminus 0$ ist F partiell differenzierbar, wie aus dem vorigen Beispiel und der Produktregel für Funktionen einer Veränderlichen folgt. Für die partielle Ableitung nach x_1 berechnet man

$$\begin{aligned} \frac{\partial F(x)}{\partial x_1} &= \frac{x_2 \cdots x_n}{r^n} + x_1 x_2 \cdots x_n \frac{\partial}{\partial x_1} (r^{-n}) \\ &= \frac{x_2 \cdots x_n}{r^n} - n \frac{x_1^2 x_2 \cdots x_n}{r^{n+2}}. \end{aligned}$$

Die partielle Ableitung in i -ter Koordinatenrichtung ergibt sich daraus durch Vertauschen der Rollen von x_1 und x_i , da F völlig symmetrisch von x_1, \dots, x_n abhängt.

Die Funktion F ist aber auch an der Stelle $x = 0$ partiell differenzierbar mit

$$\frac{\partial F}{\partial x_i}(0) = \lim_{h \rightarrow 0} \frac{F(h e_i) - F(0)}{h} = 0,$$

da $F(h e_i) = 0$ für alle $h \in \mathbb{R}$. Daraus folgt, dass F in ganz \mathbb{R}^n partiell differenzierbar ist. F ist aber im Nullpunkt nicht stetig. Betrachten wir etwa die Werte von F auf der gegen 0 konvergierenden Punktfolge

$$a_k := \left(\frac{1}{k}, \dots, \frac{1}{k} \right), \quad k \geq 1.$$

Es gilt $r(a_k) = \sqrt{n}/k$, also

$$F(a_k) = \frac{(1/k)^n}{(\sqrt{n}/k)^n} = n^{-n/2} \quad (\text{unabhängig von } k).$$

Da $\lim_{k \rightarrow \infty} F(a_k) = n^{-n/2}$ und $F(0) = 0$, ist F im Nullpunkt nicht stetig.

Für Funktionen mehrerer Veränderlichen folgt also, im Gegensatz zum Fall $n = 1$, aus der partiellen Differenzierbarkeit nicht die Stetigkeit.

Bemerkung. Wir werden im nächsten Paragraphen eine Verallgemeinerung des Differenzierbarkeitsbegriffs auf Funktionen mehrerer Veränderlichen kennen lernen, welche die Stetigkeit nach sich zieht. Insbesondere wird sich ergeben (§ 6, Corollar zu Satz 2): Eine stetig partiell differenzierbare Funktion ist auch stetig.

Definition (Gradient). Sei $U \subset \mathbb{R}^n$ offen und $f : U \rightarrow \mathbb{R}$ eine partiell differenzierbare Funktion. Dann heißt der Vektor

$$\operatorname{grad} f(x) = \left(\frac{\partial f}{\partial x_1}(x), \dots, \frac{\partial f}{\partial x_n}(x) \right)$$

der *Gradient* von f im Punkt $x \in U$.

Beispiele

(5.5) Für die in Beispiel (5.2) definierte Funktion r gilt für $x \in \mathbb{R}^n$

$$\operatorname{grad} r = \frac{x}{r}.$$

Der Vektor $\frac{x}{r}$ hat den Betrag 1 und die Richtung x . Mit den Bezeichnungen von (5.3) gilt

$$\operatorname{grad} f(r) = f'(r) \frac{x}{r}, \quad \text{z.B.} \quad \operatorname{grad} \frac{1}{r} = -\frac{x}{r^3}.$$

(5.6) Seien $f, g : U \rightarrow \mathbb{R}$ zwei partiell differenzierbare Funktionen. Dann gilt die Produktregel

$$\operatorname{grad}(fg) = g \cdot \operatorname{grad} f + f \cdot \operatorname{grad} g.$$

Dies folgt aus der Produktregel für Funktionen einer Veränderlichen, denn

$$\frac{\partial}{\partial x_i}(fg) = \frac{\partial f}{\partial x_i} g + f \frac{\partial g}{\partial x_i}.$$

Schreibweise. Statt $\operatorname{grad} f$ schreibt man auch ∇f , gesprochen Nabla f . Man hat ∇ als vektorwertigen Differentialoperator aufzufassen,

$$\nabla = \left(\frac{\partial}{\partial x_1}, \dots, \frac{\partial}{\partial x_n} \right).$$

Definition. Sei $U \subset \mathbb{R}^n$ eine offene Teilmenge von \mathbb{R}^n . Unter einem *Vektorfeld* auf U versteht man eine Abbildung

$$v : U \longrightarrow \mathbb{R}^n.$$

Jedem Punkt $x \in U$ wird also ein Vektor $v(x) \in \mathbb{R}^n$ zugeordnet.

Der Gradient ∇f einer partiell differenzierbaren Funktion $f : U \rightarrow \mathbb{R}$ ist ein spezielles Vektorfeld.

Definition (Divergenz). Sei $U \subset \mathbb{R}^n$ eine offene Menge und

$$v = (v_1, \dots, v_n) : U \longrightarrow \mathbb{R}^n$$

ein partiell differenzierbares Vektorfeld (d.h. alle Komponenten $v_i : U \rightarrow \mathbb{R}$ seien partiell differenzierbar). Dann heißt die Funktion

$$\operatorname{div} v := \sum_{i=1}^n \frac{\partial v_i}{\partial x_i}$$

die *Divergenz*¹ des Vektorfeldes v .

Bemerkung. Formal kann man die Divergenz von v als Skalarprodukt des Differentialoperators ∇ mit dem Vektor v schreiben,

$$\operatorname{div} v = \langle \nabla, v \rangle = \sum_{i=1}^n \frac{\partial}{\partial x_i} v_i.$$

Die Produktregel liefert für die Divergenz die folgende Rechenregel: Auf einer offenen Menge $U \subset \mathbb{R}^n$ sei $f : U \rightarrow \mathbb{R}$ eine partiell differenzierbare Funktion und

$$v : U \longrightarrow \mathbb{R}^n$$

ein partiell differenzierbares Vektorfeld. Dann gilt

$$\frac{\partial}{\partial x_i} (f v_i) = \frac{\partial f}{\partial x_i} \cdot v_i + f \cdot \frac{\partial v_i}{\partial x_i}.$$

Summation über i ergibt

$$\operatorname{div}(fv) = \langle \operatorname{grad} f, v \rangle + f \operatorname{div} v.$$

¹dies ist natürlich etwas ganz anderes als der Begriff Divergenz im Sinne von Nicht-Konvergenz

Mit Hilfe des Nabla-Operators schreibt sich diese Formel folgendermaßen:

$$\langle \nabla, f v \rangle = \langle \nabla f, v \rangle + f \langle \nabla, v \rangle.$$

(5.7) Beispiel. Wir betrachten das Vektorfeld

$$F: \mathbb{R}^n \setminus 0 \longrightarrow \mathbb{R}^n, \quad F(x) := \frac{x}{r}, \quad r = \|x\|.$$

Da $\operatorname{div} x = \sum_{i=1}^n \frac{\partial x_i}{\partial x_i} = n$ und $\langle x, x \rangle = r^2$, folgt mit (5.5)

$$\operatorname{div} \frac{x}{r} = \left\langle \operatorname{grad} \frac{1}{r}, x \right\rangle + \frac{1}{r} \operatorname{div} x = \left\langle -\frac{x}{r^3}, x \right\rangle + \frac{n}{r} = \frac{n-1}{r}.$$

Höhere Ableitungen

Sei $U \subset \mathbb{R}^n$ offen und $f: U \rightarrow \mathbb{R}$ eine partiell differenzierbare Funktion. Sind alle partiellen Ableitungen $D_i f: U \rightarrow \mathbb{R}$ selbst wieder partiell differenzierbar, so heißt f *zweimal partiell differenzierbar*. Man kann dann die partiellen Ableitungen 2. Ordnung $D_j D_i f$ bilden.

Allgemeiner definiert man durch vollständige Induktion: Die Funktion $f: U \rightarrow \mathbb{R}$ heißt $(k+1)$ -mal *partiell differenzierbar*, wenn sie k -mal partiell differenzierbar ist und alle partiellen Ableitungen k -ter Ordnung

$$D_{i_k} \dots D_{i_2} D_{i_1} f: U \longrightarrow \mathbb{R}$$

partiell differenzierbar sind.

Die Funktion $f: U \rightarrow \mathbb{R}$ heißt k -mal *stetig partiell differenzierbar*, wenn sie k -mal partiell differenzierbar ist und alle partiellen Ableitungen der Ordnung $\leq k$ stetig sind.

A priori ist nicht klar, dass die Anwendung der Differentialoperatoren D_i und D_j , ($i \neq j$), auf eine zweimal partiell differenzierbare Funktion $f: U \rightarrow \mathbb{R}$ vertauschbar ist, d.h. $D_j D_i f = D_i D_j f$. Dies ist im Allgemeinen auch falsch, siehe das Gegenbeispiel in Aufgabe 5.2. Der folgende Satz von H.A. Schwarz sagt jedoch, dass es bei stetig partiell differenzierbaren Funktionen auf die Reihenfolge der Differentiation nicht ankommt.

Satz 1 (Schwarz). Sei $U \subset \mathbb{R}^n$ offen und $f: U \rightarrow \mathbb{R}$ eine zweimal stetig partiell differenzierbare Funktion. Dann gilt für alle $a \in U$ und alle $i, j = 1, 2, \dots, n$

$$D_j D_i f(a) = D_i D_j f(a).$$

Beweis. Es bedeutet keine Einschränkung der Allgemeinheit anzunehmen, dass $n = 2$, $i = 1, j = 2$ und $a = 0$. Statt (x_1, x_2) schreiben wir zur Vereinfachung (x, y) . Es gibt ein $\delta > 0$, so dass

$$\{(x, y) \in \mathbb{R}^2 : |x| < \delta, |y| < \delta\} \subset U.$$

Für festes $|y| < \delta$ sei $F_y :]-\delta, \delta[\rightarrow \mathbb{R}$ die Funktion einer Veränderlichen

$$F_y(x) := f(x, y) - f(x, 0).$$

Nach dem Mittelwertsatz (An. 1, § 16, Corollar 1 zu Satz 2) gibt es ein $\xi \in \mathbb{R}$ mit $|\xi| \leq |x|$, so dass

$$F_y(x) - F_y(0) = F'_y(\xi)x.$$

Nun ist $F'_y(\xi) = D_1 f(\xi, y) - D_1 f(\xi, 0)$. Der Mittelwertsatz, angewendet auf die Funktion $y \mapsto D_1 f(\xi, y)$ liefert ein η mit $|\eta| \leq |y|$, so dass

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

Insgesamt ergibt sich also

$$f(x, y) - f(x, 0) - f(0, y) + f(0, 0) = D_2 D_1 f(\xi, \eta)xy. \quad (1)$$

Wir wiederholen jetzt diese Überlegungen unter Vertauschung der Rollen von x und y . Zunächst wenden wir den Mittelwertsatz auf die Funktion

$$G_x(y) := f(x, y) - f(0, y)$$

an und erhalten

$$G_x(y) - G_x(0) = G'_x(\tilde{\eta})y$$

mit $|\tilde{\eta}| \leq |y|$. Weiter folgt

$$G'_x(\tilde{\eta}) = D_2 f(x, \tilde{\eta}) - D_2 f(0, \tilde{\eta}) = D_1 D_2 f(\tilde{\xi}, \tilde{\eta})x$$

mit $|\tilde{\xi}| \leq |x|$, also

$$f(x, y) - f(x, 0) - f(0, y) + f(0, 0) = D_1 D_2 f(\tilde{\xi}, \tilde{\eta})xy. \quad (2)$$

Aus (1) und (2) folgt für $xy \neq 0$

$$D_2 D_1 f(\xi, \eta) = D_1 D_2 f(\tilde{\xi}, \tilde{\eta}),$$

wobei natürlich (ξ, η) und $(\tilde{\xi}, \tilde{\eta})$ von (x, y) abhängen. Lässt man nun (x, y) gegen $(0, 0)$ streben, so gilt auch $(\xi, \eta) \rightarrow 0$ und $(\tilde{\xi}, \tilde{\eta}) \rightarrow 0$. Aus der Stetigkeit der Ableitungen $D_1 D_2 f$ und $D_2 D_1 f$ folgt

$$D_2 D_1 f(0, 0) = D_1 D_2 f(0, 0), \quad \text{q.e.d.}$$

Corollar. Sei $U \subset \mathbb{R}^n$ offen und $f: U \rightarrow \mathbb{R}$ eine k -mal stetig partiell differenzierbare Funktion. Dann gilt

$$D_{i_k} \dots D_{i_2} D_{i_1} f = D_{i_{\pi(k)}} \dots D_{i_{\pi(2)}} D_{i_{\pi(1)}} f$$

für alle $i_1, i_2, \dots, i_k \in \{1, 2, \dots, n\}$ und jede Permutation π der Zahlen $1, \dots, k$.

Der Beweis erfolgt durch vollständige Induktion über k unter Verwendung der Tatsache, dass sich jede Permutation aus Vertauschungen benachbarter Glieder zusammensetzen lässt.

Schreibweise. Man verwendet auch die Schreibweisen

$$D_j D_i f = \frac{\partial^2 f}{\partial x_j \partial x_i}, \quad D_i D_i f = D_i^2 f = \frac{\partial^2 f}{\partial x_i^2} = \left(\frac{\partial}{\partial x_i} \right)^2 f,$$

$$D_{i_k} \dots D_{i_1} f = \frac{\partial^k f}{\partial x_{i_k} \dots \partial x_{i_1}}, \quad \text{usw.}$$

(5.8) Beispiel. Sei U eine offene Menge im \mathbb{R}^3 . Für ein partiell differenzierbares Vektorfeld $v: U \rightarrow \mathbb{R}^3$ definiert man ein neues Vektorfeld $\text{rot } v: U \rightarrow \mathbb{R}^3$, die *Rotation* von v , folgendermaßen:

$$\text{rot } v = \left(\frac{\partial v_3}{\partial x_2} - \frac{\partial v_2}{\partial x_3}, \frac{\partial v_1}{\partial x_3} - \frac{\partial v_3}{\partial x_1}, \frac{\partial v_2}{\partial x_1} - \frac{\partial v_1}{\partial x_2} \right).$$

Aus Satz 1 folgt nun: Ist $f: U \rightarrow \mathbb{R}$ eine zweimal stetig differenzierbare Funktion, so gilt

$$\text{rot grad } f = 0.$$

Für die erste Komponente von $\operatorname{rot} \operatorname{grad} f$ erhält man nach Definition

$$\frac{\partial^2 f}{\partial x_2 \partial x_3} - \frac{\partial^2 f}{\partial x_3 \partial x_2} = 0.$$

Die anderen Komponenten ergeben sich daraus durch zyklische Vertauschung der Indizes $1 \rightarrow 2 \rightarrow 3 \rightarrow 1$.

Damit ein stetig partiell differenzierbares Vektorfeld $v: U \rightarrow \mathbb{R}^3$ sich als Gradient einer Funktion $f: U \rightarrow \mathbb{R}$ darstellen lässt, ist also notwendig $\operatorname{rot} v = 0$.

Bemerkung. Für zwei Vektoren $x = (x_1, x_2, x_3) \in \mathbb{R}^3$ und $y = (y_1, y_2, y_3) \in \mathbb{R}^3$ ist das *Vektorprodukt* definiert als

$$x \times y := (x_2 y_3 - x_3 y_2, x_3 y_1 - x_1 y_3, x_1 y_2 - x_2 y_1) \in \mathbb{R}^3.$$

Deshalb lässt sich die Rotation eines Vektorfelds $v: U \rightarrow \mathbb{R}^3$ formal auffassen als Vektorprodukt des Nabla-Operators ∇ mit v ,

$$\operatorname{rot} v = \nabla \times v.$$

Die Formel $\operatorname{rot} \operatorname{grad} f = 0$ schreibt sich damit $\nabla \times \nabla f = 0$ und ist so einfach zu merken, denn für jeden Vektor $x \in \mathbb{R}^3$ gilt $x \times x = 0$.

Der Laplace-Operator

Sei $U \subset \mathbb{R}^n$ offen und $f: U \rightarrow \mathbb{R}$ eine zweimal stetig partiell differenzierbare Funktion. Man setzt

$$\Delta f := \operatorname{div} \operatorname{grad} f = \frac{\partial^2 f}{\partial x_1^2} + \dots + \frac{\partial^2 f}{\partial x_n^2}.$$

Man nennt

$$\Delta := \frac{\partial^2}{\partial x_1^2} + \dots + \frac{\partial^2}{\partial x_n^2}$$

den *Laplace-Operator*. Er lässt sich formal als Skalarprodukt des Nabla-Operators mit sich selbst auffassen:

$$\langle \nabla, \nabla \rangle = \sum_{i=1}^n \frac{\partial}{\partial x_i} \cdot \frac{\partial}{\partial x_i} = \Delta.$$

(Deshalb findet man manchmal, vor allem in der physikalischen Literatur, auch die Bezeichnung ∇^2 für den Laplace-Operator.) Der Laplace-Operator spielt

eine wichtige Rolle in den Differentialgleichungen der Mathematischen Physik (dort ist meist $n = 3$ oder $n = 2$). Die Gleichung

$$\Delta f = 0$$

heißt *Potentialgleichung*; ihre Lösungen heißen harmonische Funktionen. (Zum Beispiel genügt das elektrostatische Feld im ladungsfreien Raum der Potentialgleichung.)

Außer der Menge $U \subset \mathbb{R}^n$, die als räumlicher Bereich aufgefasst werde, sei noch ein Intervall $I \subset \mathbb{R}$ gegeben, das als Zeitintervall interpretiert werde. Die Koordinaten in $U \times I \subset \mathbb{R}^{n+1}$ seien mit

$$(x, t) = (x_1, \dots, x_n, t)$$

bezeichnet. Für Funktionen

$$f : U \times I \longrightarrow \mathbb{R}, \quad (x, t) \mapsto f(x, t),$$

heißt

$$\Delta f - \frac{1}{c^2} \frac{\partial^2 f}{\partial t^2} = 0$$

die *Wellengleichung* und

$$\Delta f - \frac{1}{k} \frac{\partial f}{\partial t} = 0$$

die *Wärmeleitungsgleichung*. Dabei wirkt der Laplace-Operator auf f als Funktion des Ortes x . (Die Konstanten $c > 0$ und $k > 0$ bedeuten die Wellenausbreitungsgeschwindigkeit bzw. die Temperatur-Leitfähigkeit.)

Beispiele

(5.9) Sei $f: \mathbb{R}_+^* \rightarrow \mathbb{R}$ eine zweimal stetig differenzierbare Funktion. Wir betrachten die Wirkung des Laplace-Operators auf die rotationssymmetrische Funktion

$$\mathbb{R}^n \setminus 0 \longrightarrow \mathbb{R}, \quad x \mapsto f(r), \quad r = \|x\|.$$

Nach (5.5) gilt

$$\operatorname{grad} f(r) = f'(r) \frac{x}{r},$$

also nach der Produktregel und (5.7)

$$\begin{aligned}\Delta f(r) &= \operatorname{div} \operatorname{grad} f(r) = \operatorname{div} \left(f'(r) \frac{x}{r} \right) \\ &= \left\langle \operatorname{grad} f'(r), \frac{x}{r} \right\rangle + f'(r) \operatorname{div} \frac{x}{r} \\ &= \left\langle f''(r) \frac{x}{r}, \frac{x}{r} \right\rangle + f'(r) \frac{n-1}{r},\end{aligned}$$

d.h.

$$\Delta f(r) = f''(r) + \frac{n-1}{r} f'(r).$$

Daraus ergibt sich z.B.

$$\Delta \frac{1}{r^{n-2}} = 0.$$

Insbesondere für $n = 3$ ist also $\frac{1}{r}$ eine Lösung der Potentialgleichung in $\mathbb{R}^3 \setminus 0$. Dies ist bis auf einen konstanten Faktor das sog. *Newton-Potential* mit einer Singularität im Nullpunkt, welches das Gravitationsfeld eines im Nullpunkt befindlichen Massenpunktes darstellt (und ebenso das elektrostatische Feld einer Punktladung im Nullpunkt).

Speziell für $n = 2$ ergibt sich noch

$$\Delta \log r = 0 \quad \text{in } \mathbb{R}^2 \setminus 0.$$

Die Funktion $\log r$ ist das zweidimensionale Analogon des Newton-Potentials.

(5.10) Wir wollen zeigen, dass die Funktion $F: (\mathbb{R}^3 \setminus 0) \times \mathbb{R} \rightarrow \mathbb{R}$,

$$F(x, t) := \frac{\cos(r - ct)}{r}, \quad r = \|x\|,$$

eine Lösung der Wellengleichung

$$\left(\Delta - \frac{1}{c^2} \frac{\partial^2}{\partial t^2} \right) F(x, t) = 0$$

im dreidimensionalen Raum ist. (Natürlich gibt es noch viele andere Lösungen.) Nach der Formel in (5.9) gilt

$$\Delta F = \left(\frac{\partial^2}{\partial r^2} + \frac{2}{r} \frac{\partial}{\partial r} \right) \frac{\cos(r - ct)}{r}.$$

Nun ist

$$\frac{\partial}{\partial r} \frac{\cos(r-ct)}{r} = -\frac{\sin(r-ct)}{r} - \frac{\cos(r-ct)}{r^2},$$

$$\frac{\partial^2}{\partial r^2} \frac{\cos(r-ct)}{r} = -\frac{\cos(r-ct)}{r} + 2\frac{\sin(r-ct)}{r^2} + 2\frac{\cos(r-ct)}{r^3},$$

also

$$\Delta \frac{\cos(r-ct)}{r} = -\frac{\cos(r-ct)}{r}.$$

Andrerseits ist

$$\frac{\partial^2}{\partial t^2} \frac{\cos(r-ct)}{r} = -c^2 \frac{\cos(r-ct)}{r},$$

woraus die Behauptung folgt.

Bemerkung. Genauso zeigt man, dass die Funktion $\frac{\sin(r-ct)}{r}$ der Wellengleichung genügt. Wegen $e^{i\varphi} = \cos \varphi + i \sin \varphi$ kann man beide Funktionen zusammenfassen zu einer komplexwertigen Lösung der Wellengleichung:

$$\left(\Delta - \frac{1}{c^2} \frac{\partial^2}{\partial t^2} \right) \frac{e^{i(r-ct)}}{r} = 0.$$

Die komplexe Form lässt sich auch leicht direkt nachrechnen.

AUFGABEN

5.1. Man untersuche, an welchen Stellen die Funktion

$$f: \mathbb{R}^2 \rightarrow \mathbb{R}^2, \quad (x, y) \mapsto y\sqrt{2x^2 + y^2},$$

(einmal) partiell differenzierbar ist und berechne dort ihre partiellen Ableitungen.

5.2. Die Funktion $F: \mathbb{R}^2 \rightarrow \mathbb{R}$ sei definiert durch

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

$$F(0, 0) := 0.$$

Man zeige, dass F überall zweimal partiell differenzierbar ist, dass aber

$$D_1 D_2 F(0, 0) \neq D_2 D_1 F(0, 0).$$

5.3. Sei $U \subset \mathbb{R}^3$ offen und $v: U \rightarrow \mathbb{R}^3$ ein zweimal stetig partiell differenzierbares Vektorfeld. Man zeige, dass

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

5.4. Sei $U \subset \mathbb{R}^3$ offen und seien $f, g: U \rightarrow \mathbb{R}$ zweimal stetig partiell differenzierbare Funktionen. Man beweise die Formel

$$\Delta(fg) = g\Delta f + 2\langle \nabla f, \nabla g \rangle + f\Delta g.$$

5.5. Man zeige: Die Funktion

$$F: \mathbb{R}^n \times \mathbb{R}_+^* \rightarrow \mathbb{R}, \quad F(x, t) := t^{-n/2} e^{-\|x\|^2/4t},$$

ist eine Lösung der Wärmeleitungsgleichung

$$\Delta F - \frac{\partial F}{\partial t} = 0.$$

5.6. Sei $c > 0$, $k \in \mathbb{R}^n$ und

$$\omega := \|k\|c.$$

Weiter sei $f: \mathbb{R} \rightarrow \mathbb{R}$ eine beliebige, zweimal stetig differenzierbare Funktion. Man zeige: Die Funktion

$$F: \mathbb{R}^n \times \mathbb{R} \rightarrow \mathbb{R}, \quad F(x, t) := f(\langle k, x \rangle - \omega t)$$

ist eine Lösung der Wellengleichung

$$\Delta F - \frac{1}{c^2} \frac{\partial^2 F}{\partial t^2} = 0.$$

§ 6 Totale Differenzierbarkeit

In diesem Paragraphen definieren wir die totale Differenzierbarkeit von Abbildungen einer offenen Teilmenge des \mathbb{R}^n in den \mathbb{R}^m als gewisse Approximierbarkeit durch lineare Abbildungen. Im Gegensatz zur partiellen Differenzierbarkeit braucht man sich dabei nicht auf die einzelnen Koordinaten zu beziehen; auch ist eine total differenzierbare Abbildung von selbst stetig. Ganz einfach aus der Definition lässt sich die Kettenregel für differenzierbare Abbildungen ableiten.

Definition. Sei $U \subset \mathbb{R}^n$ eine offene Menge und $f: U \rightarrow \mathbb{R}^m$ eine Abbildung. f heißt im Punkt $x \in U$ *total differenzierbar* (oder *differenzierbar schlechthin*) falls es eine lineare Abbildung

$$A: \mathbb{R}^n \longrightarrow \mathbb{R}^m$$

gibt, so dass in einer Umgebung von x gilt

$$f(x + \xi) = f(x) + A\xi + \varphi(\xi), \quad (*)$$

wobei φ eine in einer Umgebung von $0 \in \mathbb{R}^n$ definierte Funktion mit Werten in \mathbb{R}^m ist mit

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

Mit dem in An. 1, § 12, eingeführten Landau-Symbol o (das sich in naheliegender Weise auf die hier vorliegende höherdimensionale Situation überträgt), lässt sich die letzte Bedingung auch schreiben als

$$\varphi(\xi) = o(\|\xi\|).$$

Bemerkungen.

a) Für $n = m = 1$ liefert dies die übliche Definition der Differenzierbarkeit von Funktionen einer Veränderlichen (An. 1, § 15, Satz 1).

b) Die lineare Abbildung $A: \mathbb{R}^n \rightarrow \mathbb{R}^m$ kann bzgl. der kanonischen Basen von \mathbb{R}^n bzw. \mathbb{R}^m durch eine $m \times n$ -Matrix

$$(a_{ij}) \in M(m \times n, \mathbb{R}), \quad (1 \leq i \leq m, 1 \leq j \leq n)$$

beschrieben werden. Fassen wir die Elemente von \mathbb{R}^n bzw. \mathbb{R}^m als Spaltenvektoren auf, so wird die Abbildung einfach durch Matrizen-Multiplikation von

links gegeben,

$$A\xi = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix}.$$

Im Folgenden identifizieren wir die lineare Abbildung $A: \mathbb{R}^n \rightarrow \mathbb{R}^m$ mit der sie beschreibenden Matrix.

Seien

$$f = \begin{pmatrix} f_1 \\ \vdots \\ f_m \end{pmatrix} \quad \text{und} \quad \varphi = \begin{pmatrix} \varphi_1 \\ \vdots \\ \varphi_m \end{pmatrix}$$

die Komponenten-Darstellungen von f und φ . Dann schreibt sich die Gleichung (*) ausführlich als

$$f_i(x + \xi) = f_i(x) + \sum_{j=1}^n a_{ij} \xi_j + \varphi_i(\xi), \quad i = 1, \dots, m.$$

Daran sieht man auch, dass die Abbildung $f: U \rightarrow \mathbb{R}^m$ genau dann im Punkt x differenzierbar ist, wenn alle $f_i: U \rightarrow \mathbb{R}$ in x differenzierbar sind.

(6.1) Beispiel. Sei $C = (c_{ij}) \in M(n \times n, \mathbb{R})$ eine symmetrische $n \times n$ -Matrix und

$$f(x) := \langle x, Cx \rangle = \sum_{i,j=1}^n c_{ij} x_i x_j$$

die zugehörige quadratische Form $f: \mathbb{R}^n \rightarrow \mathbb{R}$. Für $x, \xi \in \mathbb{R}^n$ gilt

$$\begin{aligned} f(x + \xi) &= \langle x + \xi, Cx + C\xi \rangle \\ &= \langle x, Cx \rangle + \langle \xi, Cx \rangle + \langle x, C\xi \rangle + \langle \xi, C\xi \rangle \\ &= \langle x, Cx \rangle + 2\langle Cx, \xi \rangle + \langle \xi, C\xi \rangle \\ &= f(x) + \langle a, \xi \rangle + \varphi(\xi). \end{aligned}$$

mit $a := 2Cx$ und $\varphi(\xi) = \langle \xi, C\xi \rangle$. Da

$$\|\varphi(\xi)\| \leq \|\xi\| \cdot \|C\xi\| \leq \|C\| \cdot \|\xi\|^2,$$

gilt $\lim_{\xi \rightarrow 0} \frac{\varphi(\xi)}{\|\xi\|} = 0$. Dies zeigt, dass f in x differenzierbar ist.

Satz 1. Sei $U \subset \mathbb{R}^n$ offen und $f: U \rightarrow \mathbb{R}^m$ eine Abbildung, die im Punkt $x \in U$ differenzierbar ist, und zwar gelte

$$f(x + \xi) = f(x) + A\xi + o(\|\xi\|)$$

mit der Matrix $A = (a_{ij}) \in M(m \times n, \mathbb{R})$. Dann gilt:

- a) f ist im Punkt x stetig.
- b) Alle Komponenten $f_i: U \rightarrow \mathbb{R}$, $1 \leq i \leq m$, von f sind in x partiell differenzierbar mit

$$\frac{\partial f_i}{\partial x_j} = a_{ij}.$$

Bemerkung. Aus b) folgt insbesondere, dass die Matrix A durch die differenzierbare Abbildung f eindeutig bestimmt ist. Man nennt A das *Differential* oder die *Funktional-Matrix* (oder auch *Jacobi-Matrix*) von f im Punkte x , Schreibweisen:

$$(Df)(x) := J_f(x) := \frac{\partial(f_1, \dots, f_m)}{\partial(x_1, \dots, x_n)}(x) := \left(\frac{\partial f_i}{\partial x_j}(x) \right)_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}}.$$

Die i -te Zeile der Funktional-Matrix ist der Gradient der Funktion f_i :

$$\left(\frac{\partial f_i}{\partial x_1}(x), \dots, \frac{\partial f_i}{\partial x_n}(x) \right) = \text{grad } f_i(x)$$

Für eine skalarwertige differenzierbare Funktion $f: U \rightarrow \mathbb{R}$ ist also die Funktional-Matrix einfach der Gradient.

Beweis.

- a) Da $\lim_{\xi \rightarrow 0} A\xi = 0$ und $\lim_{\xi \rightarrow 0} o(\|\xi\|) = 0$, folgt

$$\lim_{\xi \rightarrow 0} f(x + \xi) = f(x),$$

was zeigt, dass f in x stetig ist.

- b) Für $i = 1, \dots, m$ ist

$$f_i(x + \xi) = f_i(x) + \sum_{j=1}^n a_{ij} \xi_j + \varphi_i(\xi)$$

mit $\varphi_i(\xi) = o(\|\xi\|)$, also

$$f_i(x + h e_j) = f_i(x) + h a_{ij} + \varphi_i(h e_j).$$

Daraus folgt

$$\frac{\partial f_i}{\partial x_j}(x) = \lim_{h \rightarrow 0} \frac{f_i(x + h e_j) - f_i(x)}{h} = a_{ij} + \lim_{h \rightarrow 0} \frac{\varphi_i(h e_j)}{h} = a_{ij}.$$

Satz 2. Sei $U \subset \mathbb{R}^n$ offen und $f: U \rightarrow \mathbb{R}$ eine in U partiell differenzierbare Funktion. Alle partiellen Ableitungen $D_k f$ seien im Punkt $x \in U$ stetig. Dann ist f in x total differenzierbar.

Beweis. Da U offen ist, gibt es ein $\delta > 0$, so dass die Kugel um x mit Radius δ ganz in U liegt. Sei nun $\xi = (\xi_1, \dots, \xi_n) \in \mathbb{R}^n$ ein Vektor mit $\|\xi\| < \delta$. Wir definieren Punkte

$$z^{(k)} := x + \sum_{i=1}^k \xi_i e_i, \quad k = 0, 1, \dots, n.$$

Es gilt $z^{(0)} = x$ und $z^{(n)} = x + \xi$. Die Punkte $z^{(k-1)}$ und $z^{(k)}$ unterscheiden sich nur in der k -ten Koordinate. Nach dem Mittelwertsatz für differenzierbare Funktionen einer Veränderlichen gibt es deshalb ein $\theta_k \in [0, 1]$, so dass

$$f(z^{(k)}) - f(z^{(k-1)}) = D_k f(y^{(k)}) \xi_k \quad \text{mit} \quad y^{(k)} := z^{(k-1)} + \theta_k \xi_k e_k.$$

Daraus folgt

$$f(x + \xi) - f(x) = \sum_{k=1}^n D_k f(y^{(k)}) \xi_k.$$

Setzt man $a_k := D_k f(x)$, so gilt

$$f(x + \xi) = f(x) + \sum_{k=1}^n a_k \xi_k + \varphi(\xi)$$

mit

$$\varphi(\xi) = \sum_{k=1}^n (D_k f(y^{(k)}) - a_k) \xi_k.$$

Für $\xi \rightarrow 0$ strebt $y^{(k)}$ gegen x , also folgt aus der Stetigkeit der partiellen Ableitungen $\lim_{\xi \rightarrow 0} D_k f(y^{(k)}) = D_k f(x) = a_k$, woraus folgt

$$\lim_{\xi \rightarrow 0} \frac{\varphi(\xi)}{\|\xi\|} = 0, \quad \text{q.e.d.}$$

Corollar. Sei $U \subset \mathbb{R}^n$ offen und $f: U \rightarrow \mathbb{R}$ eine stetig partiell differenzierbare Funktion. Dann ist f in U stetig.

Denn eine (total) differenzierbare Funktion ist stetig.

Bemerkung. Es gelten also die Implikationen

$$\begin{array}{c} \text{stetig partiell differenzierbar} \\ \Downarrow \\ \text{(total) differenzierbar} \\ \Downarrow \\ \text{partiell differenzierbar} \end{array}$$

Die Umkehrungen gelten i.Allg. nicht.

Wegen dieser Zusammenhänge nennt man eine stetig partiell differenzierbare Funktion kurz *stetig differenzierbar*.

Die Kettenregel

In der Differentialrechnung mehrerer Veränderlichen ist (wie im eindimensionalen Fall) die Kettenregel ein wichtiges Hilfsmittel. Sie macht eine Aussage darüber, wie sich die Funktional-Matrix einer zusammengesetzten Abbildung aus den Funktional-Matrizen der einzelnen Abbildungen berechnen lässt.

Satz 3 (Kettenregel). Seien $U \subset \mathbb{R}^n$ und $V \subset \mathbb{R}^m$ offene Mengen sowie

$$f: U \longrightarrow \mathbb{R}^m \quad \text{und} \quad g: V \longrightarrow \mathbb{R}^k$$

Abbildungen mit $f(U) \subset V$. Die Abbildung f sei im Punkt $x \in U$ differenzierbar und die Abbildung g im Punkt $y := f(x) \in V$ differenzierbar. Dann ist die zusammengesetzte Abbildung

$$g \circ f: U \longrightarrow \mathbb{R}^k$$

im Punkt x differenzierbar und für ihr Differential gilt

$$D(g \circ f)(x) = (Dg)(f(x)) \cdot Df(x).$$

Beweis. Sei $A := Df(x)$ und $B := Dg(y)$. Es ist zu zeigen, dass $D(g \circ f)(x) = BA$. Nach Voraussetzung gilt

$$\begin{aligned} f(x + \xi) &= f(x) + A\xi + \varphi(\xi), \\ g(y + \eta) &= g(y) + B\eta + \psi(\eta) \end{aligned}$$

mit $\varphi(\xi) = o(\|\xi\|)$ und $\psi(\eta) = o(\|\eta\|)$. Setzt man speziell

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

so ergibt sich

$$\begin{aligned} (g \circ f)(x + \xi) &= g(f(x + \xi)) = g(f(x) + \eta) \\ &= g(f(x)) + BA\xi + B\varphi(\xi) + \psi(A\xi + \varphi(\xi)) \\ &= (g \circ f)(x) + BA\xi + \chi(\xi) \end{aligned}$$

mit

$$\chi(\xi) := B\varphi(\xi) + \psi(A\xi + \varphi(\xi)).$$

Der Satz ist bewiesen, wenn wir zeigen können, dass $\chi(\xi) = o(\|\xi\|)$. Dies sieht man so: Da nach Voraussetzung

$$\lim_{\xi \rightarrow 0} \frac{\varphi(\xi)}{\|\xi\|} = 0, \quad \text{gilt auch} \quad \lim_{\xi \rightarrow 0} \frac{B\varphi(\xi)}{\|\xi\|} = 0.$$

Außerdem gibt es ein $\delta > 0$, so dass

$$\|\varphi(\xi)\| \leq \|\xi\| \quad \text{für alle } \xi \text{ mit } \|\xi\| < \delta.$$

Wegen $\psi(\eta) = o(\|\eta\|)$ gilt

$$\psi(\eta) = \|\eta\| \psi_1(\eta) \quad \text{mit} \quad \lim_{\eta \rightarrow 0} \psi_1(\eta) = 0.$$

Daraus folgt

$$\|\psi(A\xi + \varphi(\xi))\| \leq (\|A\| + 1)\|\xi\| \cdot \|\psi_1(A\xi + \varphi(\xi))\|,$$

also

$$\lim_{\xi \rightarrow 0} \frac{\|\psi(A\xi + \varphi(\xi))\|}{\|\xi\|} = 0, \quad \text{d.h.} \quad \chi(\xi) = o(\|\xi\|), \quad \text{q.e.d.}$$

Corollar. Seien $U \subset \mathbb{R}^n$ und $V \subset \mathbb{R}^m$ offene Mengen, $f: V \rightarrow \mathbb{R}$, $x \mapsto f(x)$, eine differenzierbare Funktion sowie

$$\varphi = \begin{pmatrix} \varphi_1 \\ \vdots \\ \varphi_m \end{pmatrix} : U \longrightarrow \mathbb{R}^m, \quad t \mapsto x = \varphi(t),$$

eine differenzierbare Abbildung mit $\varphi(U) \subset V$. Dann ist die Funktion

$$F := f \circ \varphi : U \longrightarrow \mathbb{R}, \quad t \mapsto f(\varphi(t))$$

partiell differenzierbar und es gilt für $i = 1, \dots, n$

$$\frac{\partial F}{\partial t_i}(t_1, \dots, t_n) = \sum_{j=1}^m \frac{\partial f}{\partial x_j}(\varphi_1(t), \dots, \varphi_m(t)) \frac{\partial \varphi_j}{\partial t_i}(t_1, \dots, t_n).$$

Beweis. Die Funktional-Matrizen von F , f und φ sind

$$\begin{aligned} DF(t) &= \left(\frac{\partial F}{\partial t_1}(t), \dots, \frac{\partial F}{\partial t_n}(t) \right), \\ Df(x) &= \left(\frac{\partial f}{\partial x_1}(x), \dots, \frac{\partial f}{\partial x_m}(x) \right), \\ D\varphi(t) &= \begin{pmatrix} \frac{\partial \varphi_1}{\partial t_1}(t) & \dots & \frac{\partial \varphi_1}{\partial t_n}(t) \\ \vdots & & \vdots \\ \frac{\partial \varphi_m}{\partial t_1}(t) & \dots & \frac{\partial \varphi_m}{\partial t_n}(t) \end{pmatrix}. \end{aligned}$$

Die Behauptung ergibt sich deshalb durch Matrix-Multiplikation aus der Gleichung

$$DF(t) = (Df)(\varphi(t)) \cdot D\varphi(t).$$

Definition (Richtungsableitung). Sei $U \subset \mathbb{R}^n$ offen und $f: U \rightarrow \mathbb{R}$ eine Funktion. Weiter sei $x \in U$ ein Punkt und $v \in \mathbb{R}^n$ ein Vektor mit $\|v\| = 1$. Unter der *Richtungsableitung* von f im Punkt x in Richtung v versteht man (im Fall der Existenz) den Differentialquotienten

$$D_v f(x) := \left. \frac{d}{dt} f(x + tv) \right|_{t=0} = \lim_{t \rightarrow 0} \frac{f(x + tv) - f(x)}{t}.$$

Für $v = e_i$ ist also $D_{e_i} f$ gleich der i -ten partiellen Ableitung $D_i f$.

Satz 4. Sei $U \subset \mathbb{R}^n$ offen und $f: U \rightarrow \mathbb{R}$ eine stetig differenzierbare Funktion. Dann gilt für jedes $x \in U$ und jeden Vektor $v \in \mathbb{R}^n$ mit $\|v\| = 1$

$$D_v f(x) = \langle v, \text{grad } f(x) \rangle.$$

Beweis. Sei $\varphi: \mathbb{R} \rightarrow \mathbb{R}^n$ definiert durch

$$\varphi(t) := x + tv = (x_1 + tv_1, \dots, x_n + tv_n).$$

(Dies ist die Parameterdarstellung einer Geraden durch x in Richtung v .) Für genügend kleines $\varepsilon > 0$ gilt $\varphi(]-\varepsilon, \varepsilon[) \subset U$, also ist

$$F := f \circ \varphi:]-\varepsilon, \varepsilon[\longrightarrow \mathbb{R}$$

definiert. Nach Definition der Richtungsableitung ist

$$D_v f(x) = \left. \frac{d}{dt} f(x + tv) \right|_{t=0} = \frac{dF}{dt}(0).$$

Aus der Kettenregel folgt

$$\frac{dF}{dt}(t) = \sum_{i=1}^n \frac{\partial f}{\partial x_i}(\varphi(t)) \frac{d\varphi_i(t)}{dt}.$$

Da

$$\frac{d\varphi_i(t)}{dt} = \frac{d}{dt}(x_i + tv_i) = v_i,$$

folgt weiter

$$\frac{dF}{dt}(0) = \sum_{i=1}^n \frac{\partial f}{\partial x_i}(x) v_i = \langle v, \text{grad } f(x) \rangle \quad \text{q.e.d.}$$

Bemerkung. Ist $\text{grad } f(x) \neq 0$, so ist der Winkel θ zwischen den Vektoren v und $\text{grad } f(x)$ definiert und es gilt

$$D_v f(x) = \langle v, \text{grad } f(x) \rangle = \|\text{grad } f(x)\| \cos \theta.$$

Daraus folgt: Die Richtungsableitung $D_v f(x)$ ist maximal, falls v und $\text{grad } f(x)$ die gleiche Richtung haben. Der Vektor $\text{grad } f(x)$ gibt also die Richtung des stärksten Anstiegs von f an.

Mittelwertsatz

Der Mittelwertsatz für Funktionen einer Veränderlichen lässt sich so formulieren (vgl. An. 1, § 16, Corollar 1 zu Satz 2): Ist $f: I \rightarrow \mathbb{R}$ eine differenzierbare Funktion auf dem Intervall $I \subset \mathbb{R}$ und sind $x, x + \xi \in I$, so gibt es ein $\theta \in]0, 1[$,

so dass

$$f(x + \xi) - f(x) = f'(x + \theta\xi) \cdot \xi.$$

Setzt man zusätzlich voraus, dass f stetig differenzierbar ist, so folgt aus dem Fundamentalsatz der Differential- und Integralrechnung

$$f(x + \xi) - f(x) = \int_x^{x+\xi} f'(u) du = \left(\int_0^1 f'(x + t\xi) dt \right) \cdot \xi.$$

Gegenüber der obigen Form wird also der Wert der Ableitung an einer Zwischenstelle $f'(x + \theta\xi)$ ersetzt durch den Mittelwert $\int_0^1 f'(x + t\xi) dt$ der Ableitung auf der Strecke von x nach $x + \xi$. In dieser integrierten Form lässt sich der Mittelwertsatz auf vektorwertige Funktionen mehrerer Veränderlichen übertragen. Da das Differential einer solchen Funktion eine Matrix ist, benötigen wir den Begriff des Integrals einer matrixwertigen Funktion: Sei $A = (a_{ij})$ eine Matrix, deren Koeffizienten a_{ij} stetige Funktionen auf dem Intervall $[t_0, t_1] \subset \mathbb{R}$ seien. Dann versteht man unter dem Integral

$$\int_{t_0}^{t_1} A(t) dt$$

die Matrix mit den Koeffizienten

$$\int_{t_0}^{t_1} a_{ij}(t) dt.$$

Satz 5 (Mittelwertsatz). Sei $U \subset \mathbb{R}^n$ offen und

$$f : U \longrightarrow \mathbb{R}^m$$

eine stetig differenzierbare Abbildung. Sei $x \in U$ und $\xi \in \mathbb{R}^n$ ein Vektor derart, dass die ganze Strecke $x + t\xi$, $0 \leq t \leq 1$, in U liegt. Dann gilt

$$f(x + \xi) - f(x) = \left(\int_0^1 Df(x + t\xi) dt \right) \cdot \xi.$$

Beweis. Es seien $f_i : U \rightarrow \mathbb{R}$ die Komponenten von f . Wir betrachten die Funktionen

$$g_i : [0, 1] \longrightarrow \mathbb{R}, \quad t \mapsto g_i(t) := f_i(x + t\xi).$$

Es gilt

$$\begin{aligned} f_i(x + \xi) - f_i(x) &= g_i(1) - g_i(0) = \int_0^1 g'_i(t) dt \\ &= \int_0^1 \left(\sum_{j=1}^n \frac{\partial f_i}{\partial x_j}(x + t\xi) \xi_j \right) dt = \sum_{j=1}^n \left(\int_0^1 \frac{\partial f_i}{\partial x_j}(x + t\xi) dt \right) \xi_j. \end{aligned}$$

Da Df die Matrix mit den Koeffizienten $\frac{\partial f_i}{\partial x_j}$ ist, folgt die Behauptung.

Corollar. Die Bezeichnungen von Satz 5 seien beibehalten. Es sei

$$M := \sup_{0 \leq t \leq 1} \|Df(x + t\xi)\|.$$

Damit gilt

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

(Die Norm einer Matrix wurde in § 2 definiert.)

Beweis. Nach Satz 5 ist

$$\begin{aligned} \|f(x + \xi) - f(x)\| &= \left\| \int_0^1 (Df(x + t\xi) \xi) dt \right\| \\ &\leq \int_0^1 \|Df(x + t\xi)\| \cdot \|\xi\| dt \leq M \|\xi\|. \end{aligned}$$

Dabei wurde folgender Hilfssatz benutzt:

Hilfssatz. Sei $v: [a, b] \rightarrow \mathbb{R}^m$ eine stetige vektorwertige Funktion auf dem Intervall $[a, b] \subset \mathbb{R}$. Dann gilt

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

Beweis. Wir setzen $u := \int_a^b v(t) dt \in \mathbb{R}^m$ und $K := \|u\|$. Dann ist

$$\begin{aligned} K^2 = \langle u, u \rangle &= \left\langle \int_a^b v(t) dt, u \right\rangle = \int_a^b \langle v(t), u \rangle dt \\ &\leq \int_a^b \|v(t)\| \cdot \|u\| dt = K \int_a^b \|v(t)\| dt. \end{aligned}$$

Kürzung durch K liefert die Behauptung.

AUFGABEN

6.1. Man berechne die Jacobi-Matrix der Abbildung $F: \mathbb{R}^3 \rightarrow \mathbb{R}^3$,

$$F(r, \vartheta, \varphi) := (r \sin \vartheta \cos \varphi, r \sin \vartheta \sin \varphi, r \cos \vartheta).$$

6.2. (Darstellung des Laplace-Operators bzgl. ebener Polarkoordinaten)

Es sei p die wie folgt definierte Abbildung

$$p: \mathbb{R}_+^* \times \mathbb{R} \rightarrow \mathbb{R}^2, \quad p(r, \varphi) := (r \cos \varphi, r \sin \varphi).$$

Man zeige: Ist $u: G \rightarrow \mathbb{R}$ eine auf der offenen Menge $G \subset \mathbb{R}^2$ zweimal stetig differenzierbare Funktion, so gilt auf der Menge $p^{-1}(G)$ die Gleichung

$$(\Delta u) \circ p = \frac{\partial^2(u \circ p)}{\partial r^2} + \frac{1}{r} \frac{\partial(u \circ p)}{\partial r} + \frac{1}{r^2} \frac{\partial^2(u \circ p)}{\partial \varphi^2}.$$

6.3. Sei $U \subset \mathbb{R}^n$ eine offene Kugel und $f: U \rightarrow \mathbb{R}^m$ eine stetig differenzierbare Abbildung mit beschränktem Differential, d.h. es gebe eine Konstante $K \in \mathbb{R}_+$, so dass

$$\|Df(x)\| \leq K \quad \text{für alle } x \in U.$$

Man zeige, dass f in U gleichmäßig stetig ist.

6.4. Sei $U \subset \mathbb{R}^n$ offen und $f: U \rightarrow \mathbb{R}$ eine stetig differenzierbare Funktion. Sei $x \in U$ und $f(x) =: c$. Man zeige, dass der Gradient $\text{grad } f(x)$ auf der Niveaufäche

$$N_f(c) = \{z \in U : f(z) = c\}$$

senkrecht steht, d.h. folgendes gilt: Ist

$$\varphi:]-\varepsilon, \varepsilon[\longrightarrow \mathbb{R}^n, \quad (\varepsilon > 0),$$

eine beliebige stetig differenzierbare Kurve mit

$$\varphi(0) = x \quad \text{und} \quad \varphi(]-\varepsilon, \varepsilon[) \subset N_f(c),$$

so folgt

$$\langle \varphi'(0), \text{grad } f(x) \rangle = 0.$$

§ 7 Taylor-Formel. Lokale Extrema

Das Differential einer differenzierbaren Funktion f liefert eine Approximation von f durch eine affin-lineare Funktion. Die Taylor-Formel gibt in Verallgemeinerung davon (falls f genügend oft differenzierbar ist) eine Approximation von f bis zu beliebig hoher Ordnung. Mithilfe der Approximation bis zur zweiten Ordnung werden wir in diesem Paragraphen außerdem die lokalen Extrema von Funktionen mehrerer Veränderlichen untersuchen.

Bezeichnungen. Für den Differentialkalkül mehrerer Veränderlichen sind folgende Abkürzungen nützlich:

Für ein n -tupel $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{N}^n$ sei

$$\begin{aligned} |\alpha| &:= \alpha_1 + \alpha_2 + \dots + \alpha_n, \\ \alpha! &:= \alpha_1! \alpha_2! \cdot \dots \cdot \alpha_n! \end{aligned}$$

Ist f eine $|\alpha|$ -mal stetig differenzierbare Funktion, so setzt man

$$D^\alpha f := D_1^{\alpha_1} D_2^{\alpha_2} \dots D_n^{\alpha_n} f = \frac{\partial^{|\alpha|} f}{\partial x_1^{\alpha_1} \dots \partial x_n^{\alpha_n}}$$

wobei

$$D_i^{\alpha_i} = \underbrace{D_i D_i \dots D_i}_{\alpha_i\text{-mal}}$$

Für $x = (x_1, \dots, x_n) \in \mathbb{R}^n$ sei

$$x^\alpha := x_1^{\alpha_1} x_2^{\alpha_2} \cdot \dots \cdot x_n^{\alpha_n}.$$

Satz 1. Sei $U \subset \mathbb{R}^n$ offen und $f : U \rightarrow \mathbb{R}$ eine k -mal stetig differenzierbare Funktion. Sei $x \in U$ und $\xi \in \mathbb{R}^n$ ein Vektor derart, dass die Strecke $x + t\xi$, $0 \leq t \leq 1$, ganz in U liegt. Dann ist die Funktion

$$g : [0, 1] \longrightarrow \mathbb{R}, \quad g(t) := f(x + t\xi),$$

k -mal stetig differenzierbar und es gilt

$$\frac{d^k g}{dt^k}(t) = \sum_{|\alpha|=k} \frac{k!}{\alpha!} (D^\alpha f)(x + t\xi) \xi^\alpha.$$

Beweis.

a) Wir zeigen zunächst durch Induktion über k , dass

$$\frac{d^k g}{dt^k}(t) = \sum_{i_1, \dots, i_k=1}^n D_{i_k} \dots D_{i_1} f(x + t\xi) \xi_{i_1} \dots \xi_{i_k}.$$

Für $k = 1$ ergibt sich aus der Kettenregel

$$\frac{dg}{dt}(t) = \frac{d}{dt} f(x_1 + t\xi_1, \dots, x_n + t\xi_n) = \sum_{i=1}^n D_i f(x + t\xi) \xi_i.$$

Induktionsschritt $k - 1 \rightarrow k$.

$$\begin{aligned} \frac{d^k g}{dt^k}(t) &= \frac{d}{dt} \left(\sum_{i_1, \dots, i_{k-1}=1}^n D_{i_{k-1}} \dots D_{i_1} f(x + t\xi) \xi_{i_1} \dots \xi_{i_{k-1}} \right) \\ &= \sum_{j=1}^n D_j \left(\sum_{i_1, \dots, i_{k-1}=1}^n D_{i_{k-1}} \dots D_{i_1} f(x + t\xi) \xi_{i_1} \dots \xi_{i_{k-1}} \right) \xi_j \\ &= \sum_{i_1, \dots, i_k=1}^n D_{i_k} \dots D_{i_1} f(x + t\xi) \xi_{i_1} \dots \xi_{i_k}. \end{aligned}$$

b) Kommt unter den Indizes (i_1, \dots, i_k) der Index 1 genau α_1 -mal, der Index 2 genau α_2 -mal, \dots , der Index k genau α_k -mal vor, so ist nach § 5, Corollar zu Satz 1

$$D_{i_k} \dots D_{i_1} f(x + t\xi) \xi_{i_1} \dots \xi_{i_k} = D_1^{\alpha_1} \dots D_n^{\alpha_n} f(x + t\xi) \xi_1^{\alpha_1} \dots \xi_n^{\alpha_n}.$$

Da es aber $\frac{k!}{\alpha_1! \dots \alpha_n!}$ k -tupel (i_1, \dots, i_k) von Zahlen $1 \leq i_k \leq n$ gibt, bei denen der Index v genau α_v -mal vorkommt ($v = 1, \dots, n$, $\alpha_1 + \dots + \alpha_n = k$), folgt

$$\begin{aligned} \frac{d^k g}{dt^k}(t) &= \sum_{i_1, \dots, i_k=1}^n D_{i_k} \dots D_{i_1} f(x + t\xi) \xi_{i_1} \dots \xi_{i_k} \\ &= \sum_{|\alpha|=k} \frac{k!}{\alpha_1! \dots \alpha_n!} D_1^{\alpha_1} \dots D_n^{\alpha_n} f(x + t\xi) \xi_1^{\alpha_1} \dots \xi_n^{\alpha_n} \\ &= \sum_{|\alpha|=k} \frac{k!}{\alpha!} (D^\alpha f)(x + t\xi) \xi^\alpha, \quad \text{q.e.d.} \end{aligned}$$

Satz 2 (Taylorsche Formel). Sei $U \subset \mathbb{R}^n$ offen, $x \in U$ ein Punkt und $\xi \in \mathbb{R}^n$ ein Vektor derart, dass die Strecke $x + t\xi$, $0 \leq t \leq 1$, ganz in U liegt. Weiter sei

$f: U \rightarrow \mathbb{R}$ eine $(k+1)$ -mal stetig differenzierbare Funktion. Dann existiert ein $\theta \in [0, 1]$, so dass

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

Beweis. Wir betrachten die Funktion

$$g: [0, 1] \longrightarrow \mathbb{R}, \quad t \mapsto g(t) := f(x+t\xi).$$

Nach Satz 1 ist g eine $(k+1)$ -mal stetig differenzierbare Funktion und nach der Taylor-Formel für Funktionen einer Veränderlichen (An. 1, § 22, Satz 2) existiert ein $\theta \in [0, 1]$, so dass

$$g(1) = \sum_{m=0}^k \frac{g^{(m)}(0)}{m!} + \frac{g^{(k+1)}(\theta)}{(k+1)!}.$$

Nach Satz 1 ist für $m = 1, \dots, k$

$$\frac{g^{(m)}(0)}{m!} = \sum_{|\alpha|=m} \frac{D^\alpha f(x)}{\alpha!} \xi^\alpha$$

und

$$\frac{g^{(k+1)}(\theta)}{(k+1)!} = \sum_{|\alpha|=k+1} \frac{D^\alpha f(x+\theta\xi)}{\alpha!} \xi^\alpha,$$

woraus die Behauptung folgt.

Corollar 1. Sei $U \subset \mathbb{R}^n$ offen und

$$f: U \longrightarrow \mathbb{R}$$

eine k -mal stetig differenzierbare Funktion. Dann gilt für jedes $x \in U$

$$f(x+\xi) = \sum_{|\alpha| \leq k} \frac{D^\alpha f(x)}{\alpha!} \xi^\alpha + o(\|\xi\|^k) \quad \text{für } \xi \rightarrow 0.$$

Beweis. Da U offen ist, gibt es ein $\delta > 0$, so dass $B_\delta(x) \subset U$. Nach Satz 2 gibt es zu jedem $\xi \in \mathbb{R}^n$ mit $\|\xi\| < \delta$ ein $\theta \in [0, 1]$ mit

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

$$= \sum_{|\alpha| \leq k} \frac{D^\alpha f(x)}{\alpha!} \xi^\alpha + \sum_{|\alpha|=k} r_\alpha(\xi) \xi^\alpha,$$

wobei

$$r_\alpha(\xi) = \frac{D^\alpha f(x + \theta \xi) - D^\alpha f(x)}{\alpha!}.$$

Wegen der Stetigkeit von $D^\alpha f$ gilt $\lim_{\xi \rightarrow 0} r_\alpha(\xi) = 0$. Da

$$|\xi^\alpha| = |\xi_1^{\alpha_1} \dots \xi_n^{\alpha_n}| \leq \|\xi\|^{\alpha_1} \dots \|\xi\|^{\alpha_n} = \|\xi\|^k \quad \text{für } |\alpha| = k,$$

folgt

$$\sum_{|\alpha|=k} r_\alpha(\xi) \xi^\alpha = o(\|\xi\|^k), \quad \text{q.e.d.}$$

Bemerkung. Mit den obigen Bezeichnungen sei

$$P_m(\xi) := \sum_{|\alpha|=m} \frac{D^\alpha f(x)}{\alpha!} \xi^\alpha.$$

Dann ist $P_m(\xi)$ ein homogenes Polynom m -ten Grades in $\xi = (\xi_1, \dots, \xi_n)$ und es gilt

$$f(x + \xi) = \sum_{m=0}^k P_m(\xi) + o(\|\xi\|^k).$$

Wir wollen die Polynome P_m für die Fälle $m = 0, 1, 2$ genauer betrachten.

a) $m = 0$. Da es nur ein n -tupel $\alpha \in \mathbb{N}^n$ mit $|\alpha| = 0$ gibt, nämlich $\alpha = (0, \dots, 0)$, gilt

$$P_0(\xi) = \frac{D^0 f(x)}{0!} \xi^0 = f(x).$$

P_0 ist also eine Konstante, nämlich der Funktionswert von f im Entwicklungspunkt x .

b) $m = 1$. Die einzigen n -tupel $\alpha \in \mathbb{N}^n$ mit $|\alpha| = 1$ sind die n Einheitsvektoren

$$e_i = (0, \dots, \underbrace{0, 1, 0, \dots, 0}_{i\text{-te Stelle}})$$

Es gilt $D^{e_i} = D_i$ und $e_i! = 1$ sowie $\xi^{e_i} = \xi_i$, also

$$P_1(\xi) = \sum_{i=1}^n D_i f(x) \xi_i = \langle \text{grad } f(x), \xi \rangle.$$

Für einmal stetig differenzierbare Funktionen lautet also die Formel von Corollar 1

$$f(x + \xi) = f(x) + \langle \text{grad } f(x), \xi \rangle + o(\|\xi\|),$$

was genau die Approximierbarkeit von f durch eine lineare Funktion in der Definition der Differenzierbarkeit darstellt.

c) $m = 2$. Es gibt zwei Arten von n -tupeln $\alpha \in \mathbb{N}^n$ mit $|\alpha| = 2$. Erstens die Vektoren

$$2e_i = (0, \dots, 0, 2, 0, \dots, 0), \quad 1 \leq i \leq n,$$

wobei die 2 an i -ter Stelle steht. Zweitens die Vektoren

$$e_i + e_j = (0, \dots, 1, \dots, 1, \dots, 0) \quad 1 \leq i < j \leq n,$$

mit Einsen an den Stellen i und j . Es gilt

$$\begin{aligned} D^{2e_i} f &= D_i^2 f, & (2e_i)! &= 2, \\ D^{e_i + e_j} f &= D_i D_j f, & (e_i + e_j)! &= 1 \quad \text{für } i \neq j. \end{aligned}$$

Daraus folgt

$$P_2(\xi) = \frac{1}{2} \sum_{i=1}^n D_i^2 f(x) \xi_i^2 + \sum_{i < j} D_i D_j f(x) \xi_i \xi_j.$$

Da $D_i D_j f = D_j D_i f$, gilt für die zweite Summe

$$\sum_{i < j} D_i D_j f(x) \xi_i \xi_j = \frac{1}{2} \sum_{i \neq j} D_i D_j f(x) \xi_i \xi_j$$

und man erhält insgesamt

$$P_2(\xi) = \frac{1}{2} \sum_{i,j=1}^n D_i D_j f(x) \xi_i \xi_j,$$

wobei jetzt i und j unabhängig voneinander von 1 bis n laufen. (Dies kann man auch direkt aus Teil a) des Beweises von Satz 1 sehen.)

P_2 ist also eine quadratische Form mit der Matrix $(\frac{1}{2} D_i D_j f(x))$. Man nimmt dies zum Anlass für folgende Definition.

Definition (Hessesche Matrix). Sei $U \subset \mathbb{R}^n$ offen und $f: U \rightarrow \mathbb{R}$ eine zweimal stetig differenzierbare Funktion. Unter der *Hesseschen Matrix* von f im Punkt $x \in U$ versteht man die $n \times n$ -Matrix

$$(\text{Hess } f)(x) := \left(D_i D_j f(x) \right)_{\substack{1 \leq i \leq n \\ 1 \leq j \leq n}}$$

Die Hessesche Matrix ist symmetrisch, da $D_i D_j f = D_j D_i f$.

Wegen der obigen Bemerkungen ist das Folgende ein Spezialfall von Corollar 1:

Corollar 2 (Approximation zweiter Ordnung). Sei $U \subset \mathbb{R}^n$ offen, $x \in U$ und $f: U \rightarrow \mathbb{R}$ eine zweimal stetig differenzierbare Funktion. Dann gilt

$$f(x + \xi) = c + \langle a, \xi \rangle + \frac{1}{2} \langle \xi, A \xi \rangle + o(\|\xi\|^2),$$

wobei

$$c := f(x), \quad a := \text{grad } f(x), \quad A := (\text{Hess } f)(x).$$

Lokale Extrema

Sei $U \subset \mathbb{R}^n$ eine offene Menge und $f: U \rightarrow \mathbb{R}$ eine Funktion. Ein Punkt $x \in U$ heit *lokales Maximum* bzw. *lokales Minimum* von f , falls eine Umgebung $V \subset U$ von x existiert, so dass

$$f(x) \geq f(y) \quad (\text{bzw. } f(x) \leq f(y)) \quad \text{fr alle } y \in V.$$

Tritt in dieser Definition der Fall $f(x) = f(y)$ nur fr $x = y$ ein, so spricht man von einem *strikten* lokalen Maximum oder Minimum.

Ein *lokales Extremum* ist ein lokales Maximum oder Minimum.

Satz 3 (notwendige Bedingung fr lokales Extremum). Sei $U \subset \mathbb{R}^n$ offen und $f: U \rightarrow \mathbb{R}$ eine partiell differenzierbare Funktion. Besitzt f in $x \in U$ ein lokales Extremum, so gilt

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

Beweis. Fr $i = 1, \dots, n$ betrachten wir die Funktionen

$$g_i(t) := f(x + t e_i).$$

Die Funktion g_i ist auf einem gewissen Intervall $] -\varepsilon, \varepsilon[\subset \mathbb{R}$, $\varepsilon > 0$, definiert und dort differenzierbar. Hat f in x ein lokales Extremum, so hat g_i in 0 ein lokales Extremum, also gilt (An.1, § 16, Satz 1)

$$g'_i(0) = 0.$$

Da $g'_i(0) = D_i f(x)$, folgt

$$\text{grad } f(x) = (D_1 f(x), \dots, D_n f(x)) = 0, \quad \text{q.e.d.}$$

Um neben dieser notwendigen Bedingung auch hinreichende Bedingungen für die Existenz von lokalen Extrema einer Funktion herleiten zu können, müssen wir die Hessesche Matrix von f genauer betrachten. Dazu erinnern wir zunächst an folgende Definition aus der Theorie der quadratischen Formen.

Definition. Sei $A \in M(n \times n, \mathbb{R})$ eine symmetrische $n \times n$ -Matrix.

a) Die Matrix A heißt *positiv definit*, falls

$$\langle \xi, A\xi \rangle > 0 \quad \text{für alle } \xi \in \mathbb{R}^n \setminus 0.$$

b) Die Matrix A heißt *positiv semidefinit*, falls

$$\langle \xi, A\xi \rangle \geq 0 \quad \text{für alle } \xi \in \mathbb{R}^n.$$

c) Die Matrix A heißt *negativ definit* (bzw. *negativ semidefinit*), falls die Matrix $-A$ positiv definit (bzw. positiv semidefinit) ist.

d) Die Matrix A heißt *indefinit*, falls es Vektoren $\xi, \eta \in \mathbb{R}^n$ gibt, so dass

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

Bemerkung. Bekanntlich gibt es zu jeder symmetrischen $n \times n$ -Matrix $A \in M(n \times n, \mathbb{R})$ eine Orthonormalbasis $v_1, \dots, v_n \in \mathbb{R}^n$ von Eigenvektoren:

$$Av_i = \alpha_i v_i, \quad \langle v_i, v_j \rangle = \delta_{ij}$$

Die Eigenwerte α_i sind alle reell. Entwickelt man einen Vektor $\xi \in \mathbb{R}^n$ bezüglich dieser Orthogonalbasis,

$$\xi = \sum_{i=1}^n \lambda_i v_i$$

so wird

$$\langle \xi, A\xi \rangle = \sum_{i=1}^n \alpha_i \lambda_i^2.$$

Damit erkennt man: A ist genau dann positiv (negativ) definit, falls alle Eigenwerte von A positiv (negativ) sind. A ist genau dann positiv (negativ) semidefinit, wenn alle Eigenwerte ≥ 0 (bzw. ≤ 0) sind. A ist genau dann indefinit, wenn A mindestens einen positiven und einen negativen Eigenwert besitzt.

Falls die Eigenwerte bekannt sind, ist es also einfach zu entscheiden, ob die Matrix positiv definit (bzw. negativ definit, usw.) ist. Für $n \geq 3$ ist es aber im Allgemeinen schwierig, die Eigenwerte einer n -reihigen Matrix zu berechnen. Wir geben deshalb ohne Beweis noch ein einfaches Kriterium von Jacobi/Hurwitz für die positive Definitheit einer symmetrischen Matrix an. (Ein Beweis findet sich z.B. in [1], Abschnitt 5.7.7)

Satz (Determinanten-Kriterium für Definitheit). Sei

$$A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix} \in M(n \times n, \mathbb{R})$$

eine reelle symmetrische $n \times n$ -Matrix. A ist genau dann positiv definit, wenn für $k = 1, \dots, n$ gilt

$$\det \begin{pmatrix} a_{11} & \dots & a_{1k} \\ \vdots & & \vdots \\ a_{k1} & \dots & a_{kk} \end{pmatrix} > 0.$$

Beispielsweise ist die 3×3 -Matrix

$$A := \begin{pmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{pmatrix}$$

genau dann positiv definit, wenn folgende drei Bedingungen erfüllt sind:

$$(i) \ a_1 > 0, \quad (ii) \ a_1 b_2 - a_2 b_1 > 0, \quad (iii) \ \det A > 0.$$

Satz 4 (hinreichende Bedingung für lokales Extremum). Sei $U \subset \mathbb{R}^n$ offen, $f: U \rightarrow \mathbb{R}$ eine zweimal stetig differenzierbare Funktion und $x \in U$ ein Punkt mit

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

- a) Ist $(\text{Hess } f)(x)$ positiv definit, so besitzt f in x ein striktes lokales Minimum.
- b) Ist $(\text{Hess } f)(x)$ negativ definit, so besitzt f in x ein striktes lokales Maximum.

c) Ist $(\text{Hess } f)(x)$ indefinit, so besitzt f in x kein lokales Extremum.

Beweis. Sei $A := (\text{Hess } f)(x)$. Nach Corollar 2 zu Satz 2 gilt in einer Umgebung von x

$$f(x + \xi) = f(x) + \frac{1}{2} \langle \xi, A\xi \rangle + \varphi(\xi) \quad (1)$$

mit $\varphi(\xi) = o(\|\xi\|^2)$. Es gibt also zu jedem $\varepsilon > 0$ ein $\delta > 0$, so dass

$$|\varphi(\xi)| \leq \varepsilon \|\xi\|^2 \quad \text{für } \|\xi\| < \delta.$$

a) Sei jetzt vorausgesetzt, dass $A = (\text{Hess } f)(x)$ positiv definit ist. Sei

$$S := \{\xi \in \mathbb{R}^n : \|\xi\| = 1\}$$

die Sphäre vom Radius 1. Da S kompakt ist, nimmt die stetige Funktion

$$\xi \mapsto \langle \xi, A\xi \rangle$$

auf S ihr Minimum an. Da $\langle \xi, A\xi \rangle > 0$ für alle $\xi \in S$, ist

$$\alpha := \min\{\langle \xi, A\xi \rangle : \xi \in S\} > 0.$$

Behauptung.

$$\langle \xi, A\xi \rangle \geq \alpha \|\xi\|^2 \quad \text{für alle } \xi \in \mathbb{R}^n \quad (2)$$

Für $\|\xi\| = 1$ ist das trivial. Der allgemeine Fall wird darauf zurückgeführt, indem man den Vektor ξ schreibt als

$$\xi = \lambda \xi_* \quad \text{mit } \lambda = \|\xi\| \text{ und } \|\xi_*\| = 1.$$

Wir wählen jetzt $\delta > 0$ so klein, dass

$$|\varphi(\xi)| \leq \frac{\alpha}{4} \|\xi\|^2 \quad \text{für } \|\xi\| < \delta.$$

Dann folgt aus (1) und (2)

$$f(x + \xi) \geq f(x) + \frac{\alpha}{4} \|\xi\|^2,$$

also $f(x + \xi) > f(x)$ für $0 < \|\xi\| < \delta$. Daher hat f in x ein striktes lokales Minimum.

b) Ist $A = (\text{Hess } f)(x)$ negativ definit, so betrachte man anstelle von f die Funktion $-f$ und man ist auf den Fall a) zurückgeführt.

c) Sei $A = (\text{Hess } f)(x)$ indefinit. Es ist zu zeigen, dass in jeder Umgebung von

x Punkte y', y'' existieren mit

$$f(y'') < f(x) < f(y').$$

Da A indefinit ist, gibt es mindestens einen Vektor $\xi \in \mathbb{R}^n \setminus 0$ mit

$$\langle \xi, A\xi \rangle =: \alpha > 0.$$

Dann ist nach (1) für kleine reelle Zahlen t

$$f(x + t\xi) = f(x) + \frac{1}{2} \langle t\xi, At\xi \rangle + \varphi(t\xi) = f(x) + \frac{\alpha}{2} t^2 + \varphi(t\xi).$$

Es gibt ein $\delta > 0$, so dass $|\varphi(t\xi)| \leq \frac{\alpha}{4} t^2$ für $|t| < \delta$, also

$$f(x + t\xi) > f(x) \quad \text{für } 0 < |t| < \delta.$$

Ebenso zeigt man: Ist $\eta \in \mathbb{R}^n \setminus 0$ ein Vektor mit $\langle \eta, A\eta \rangle < 0$, so gilt für genügend kleine $t \neq 0$

$$f(x + t\eta) < f(x), \quad \text{q.e.d.}$$

Beispiele

Wir geben einige typische Beispiele für das Auftreten bzw. Nichtauftreten von lokalen Extrema. Dabei betrachten wir Funktionen

$$f: \mathbb{R}^2 \longrightarrow \mathbb{R}, \quad (x, y) \mapsto f(x, y)$$

zweier Veränderlichen, die wir, um Indizes zu sparen, mit (x, y) statt (x_1, x_2) bezeichnen.

(7.1) Sei $f(x, y) := c + x^2 + y^2$.

Die Funktion hat im Nullpunkt ein striktes lokales Minimum, da $\text{grad } f(x) = (0, 0)$ und die Hessesche Matrix

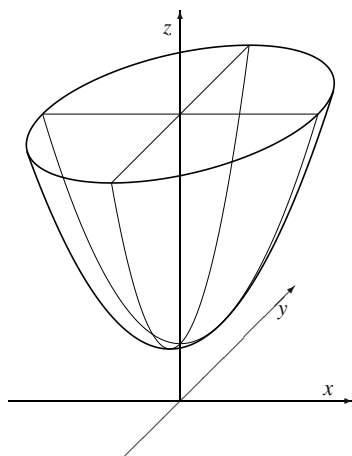
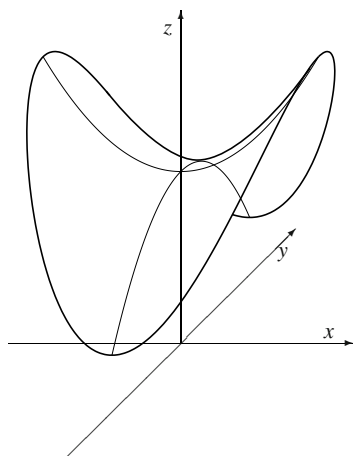
$$\text{Hess } f = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$$

positiv definit ist. (Die Funktion f hat im Nullpunkt sogar ein globales Minimum, wie man direkt sieht.) Der Graph von f ,

$$\Gamma_f = \{(x, y, z) \in \mathbb{R}^3 : z = c + x^2 + y^2\}$$

ist ein nach oben geöffnetes Paraboloid, wenn man sich die z -Achse nach oben gerichtet denkt (Bild 7.1).

(7.2) Sei $g(x, y) := c - x^2 - y^2$.

**Bild 7.1** Paraboloid**Bild 7.2** Sattelfläche

Hier liegt im Nullpunkt ein striktes Maximum vor; die Hessesche Matrix

$$\text{Hess } g = \begin{pmatrix} -2 & 0 \\ 0 & -2 \end{pmatrix}$$

ist negativ definit. Der Graph von g ,

$$\Gamma_g = \{(x, y, z) \in \mathbb{R}^3 : z = c - x^2 - y^2\}$$

ist ein nach unten geöffnetes Paraboloid.

(7.3) Sei $h(x, y) := c + x^2 - y^2$.

Der Gradient von h ,

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

verschwindet im Nullpunkt. Es ist

$$\text{Hess } h = \begin{pmatrix} 2 & 0 \\ 0 & -2 \end{pmatrix}.$$

Die Hessesche Matrix ist also indefinit, es liegt deshalb weder ein lokales Maximum noch Minimum vor. Der Graph von h ,

$$\Gamma_h = \{(x, y, z) \in \mathbb{R}^3 : z = c + x^2 - y^2\}$$

ist eine sog. *Sattelfläche* (Bild 7.2). Längs der x -Achse ($y = 0$) steigen die Werte von h vom Nullpunkt aus an, längs der y -Achse ($x = 0$) fallen sie ab.

(7.4) Ist die Hessesche Matrix in einer Nullstelle des Gradienten semidefinit, so lassen sich keine allgemeinen Aussagen machen, wie folgende Beispiele zeigen:

$$f_1(x, y) := x^2 + y^4,$$

$$f_2(x, y) := x^2,$$

$$f_3(x, y) := x^2 + y^3.$$

Für alle drei Funktionen verschwindet der Gradient im Nullpunkt und es gilt

$$(\text{Hess } f_k)(0) = \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix}, \quad k = 1, 2, 3,$$

die Hessesche Matrix ist also positiv semidefinit. Die drei Funktionen zeigen aber verschiedenes Verhalten:

- a) Die Funktion f_1 hat im Nullpunkt ein striktes lokales Minimum.
- b) Die Funktion f_2 hat im Nullpunkt ein lokales Minimum, das aber nicht strikt ist (denn in allen Punkten der y -Achse hat f_2 denselben Wert wie im Nullpunkt).
- c) Die Funktion f_3 hat im Nullpunkt weder ein lokales Minimum noch ein lokales Maximum.

AUFGABEN

7.1. Man bestimme die Taylor-Entwicklung der Funktion

$$f: \mathbb{R}_+^* \times \mathbb{R}_+^* \rightarrow \mathbb{R}, \quad f(x, y) := \frac{x-y}{x+y},$$

im Punkt $(1, 1)$ bis einschließlich den Gliedern 2. Ordnung

7.2. Man bestimme die lokalen Extrema der Funktion

$$f: \mathbb{R}^2 \rightarrow \mathbb{R}, \quad f(x, y) := (4x^2 + y^2)e^{-x^2 - 4y^2}.$$

7.3. Sei $P: \mathbb{R}^n \rightarrow \mathbb{R}$ das folgende homogene Polynom k -ten Grades:

$$P(x) = \sum_{|\alpha|=k} c_\alpha x^\alpha, \quad \alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{N}^n, \quad x = (x_1, \dots, x_n) \in \mathbb{R}^n.$$

Man beweise:

- a) Ist $\beta \in \mathbb{N}^n$ ein n -tupel mit $|\beta| = k$, so gilt

$$D^\beta P(x) = \beta! c_\beta.$$

- b) Gilt $P(x) = 0$ für alle x aus einer gewissen Umgebung des Nullpunkts, so folgt $c_\alpha = 0$ für alle $\alpha \in \mathbb{N}^n$ mit $|\alpha| = k$.
 c) Es gilt $P(x) = o(\|x\|^m)$ für alle $m < k$.
 d) Gilt $P(x) = o(\|x\|^k)$, so folgt $P(x) = 0$ für alle $x \in \mathbb{R}^n$.

7.4. Sei $U \subset \mathbb{R}^n$ offen, $f: U \rightarrow \mathbb{R}$ eine Funktion und $x \in U$ ein Punkt. In einer Umgebung von x gelte

$$f(x + \xi) = \sum_{|\alpha| \leq k} c_\alpha \xi^\alpha + \varphi(\xi)$$

und

$$f(x + \xi) = \sum_{|\alpha| \leq k} \tilde{c}_\alpha \xi^\alpha + \tilde{\varphi}(\xi)$$

mit $\varphi(\xi) = o(\|\xi\|^k)$ und $\tilde{\varphi}(\xi) = o(\|\xi\|^k)$.

Man zeige, dass $c_\alpha = \tilde{c}_\alpha$ für alle $\alpha \in \mathbb{N}^n$ mit $|\alpha| \leq k$.

7.5. Sei U eine offene Teilmenge des \mathbb{R}^n und $C_b^k(U)$ die Menge aller k -mal stetig differenzierbaren Funktionen $f: U \rightarrow \mathbb{R}$, für die $D^\alpha f$ beschränkt in U ist für alle $\alpha \in \mathbb{N}^n$ mit $|\alpha| \leq k$. Für $f \in C_b^k(U)$ werde definiert

$$\|f\|_k := \sum_{|\alpha| \leq k} \frac{1}{\alpha!} \sup\{|D^\alpha f(x)| : x \in U\}.$$

Man beweise:

- a) Die Abbildung

$$\| \cdot \|_k : C_b^k(U) \rightarrow \mathbb{R}, \quad f \mapsto \|f\|_k,$$

ist eine Norm auf dem Vektorraum $C_b^k(U)$.

- b) Für $f, g \in C_b^k(U)$ gilt

$$\|fg\|_k \leq \|f\|_k \|g\|_k.$$

- c) Der normierte Vektorraum $C_b^k(U)$ ist vollständig.

§ 8 Implizite Funktionen

Auf einer Teilmenge $U \subset \mathbb{R}^2$ sei eine Funktion $F: U \rightarrow \mathbb{R}$, $(x, y) \mapsto F(x, y)$, gegeben. Unter gewissen Voraussetzungen gibt es zu jedem x -Wert aus einem geeigneten Intervall $I \subset \mathbb{R}$ genau ein y , so dass $(x, y) \in U$ und $F(x, y) = 0$. Dadurch wird dann eine Funktion $y = g(x)$ bestimmt, für die $F(x, g(x)) = 0$ für alle $x \in I$. Man sagt in diesem Fall, die Funktion g werde durch die Gleichung $F(x, y) = 0$ implizit definiert. In diesem Paragraphen beschäftigen wir uns genauer mit den Bedingungen für die Existenz und Differenzierbarkeit impliziter Funktionen. Als Anwendung davon untersuchen wir die Umkehrung von differenzierbaren Abbildungen.

Eine Anwendung der Kettenregel

Die Kettenregel für die Differentiation von Funktionen mehrerer Veränderlichen kann dazu dienen, in manchen Fällen die Ableitung einer Funktion einer Veränderlichen einfach auszurechnen. Betrachten wir etwa die folgende Situation: Sei $U \subset \mathbb{R}^2$ offen und

$$F: U \longrightarrow \mathbb{R}, \quad (x, y) \mapsto F(x, y)$$

eine differenzierbare Funktion. Außerdem sei eine differenzierbare Funktion einer Veränderlichen

$$g: I \longrightarrow \mathbb{R}, \quad x \mapsto g(x),$$

auf einem Intervall $I \subset \mathbb{R}$ vorgegeben. Der Graph von g sei in U enthalten und es gelte

$$F(x, g(x)) = 0 \quad \text{für alle } x \in I.$$

Differenzieren wir diese Gleichung nach x mit Hilfe der Kettenregel, so ergibt sich

$$D_1 F(x, g(x)) + D_2 F(x, g(x)) g'(x) = 0.$$

Unter der Voraussetzung, dass $D_2 F(x, g(x)) \neq 0$, gilt also

$$g'(x) = -\frac{D_1 F(x, g(x))}{D_2 F(x, g(x))}.$$

Beispiele

(8.1) Auf dem Intervall $-a < x < a$ betrachten wir die Funktion

$$g(x) := \sqrt{a^2 - x^2}.$$

Es gilt $g(x)^2 = a^2 - x^2$, also

$$x^2 + g(x)^2 - a^2 = 0 \quad (*)$$

Mit den obigen Bezeichnungen ist also hier $F(x, y) = x^2 + y^2 - a^2$. Differentiation von (*) nach x ergibt

$$2x + 2g(x)g'(x) = 0,$$

d.h.

$$g'(x) = -\frac{x}{g(x)} = \frac{-x}{\sqrt{a^2 - x^2}}.$$

(8.2) Es sei $g :]0, 1[\rightarrow \mathbb{R}$ die Funktion

$$g(x) := \arcsin \sqrt{1 - x^3}.$$

Setzen wir zur Abkürzung $y = g(x)$, so ergibt sich

$$\sin y = \sqrt{1 - x^3}$$

und weiter

$$\sin^2 y + x^3 - 1 = 0.$$

Durch Differentiation nach x erhält man

$$2(\sin y)(\cos y)y' + 3x^2 = 0.$$

Da aus $0 < x < 1$ folgt $0 < \sqrt{1 - x^3} < 1$, gilt $0 < y = g(x) < \pi/2$, also $\cos y > 0$. Daraus folgt

$$\cos y = +\sqrt{1 - \sin^2 y} = \sqrt{x^3} = x\sqrt{x},$$

also

$$y' = g'(x) = \frac{-3x^2}{2\sin y \cos y} = \frac{-3x}{2\sqrt{x(1 - x^3)}}.$$

Natürlich kann man dies auch direkt nachrechnen, aber die vorgeführte Rechnung erspart einem zumindest, dass man die Ableitung von \arcsin auswendig kennen muss.

Satz 1 (Banachscher Fixpunktsatz). *Sei A eine abgeschlossene Teilmenge eines Banachraums, d.h. eines vollständigen normierten Vektorraums $(V, \|\cdot\|)$. Die Abbildung $\Phi: A \rightarrow A$ sei eine Kontraktion, d.h. es gebe eine Konstante θ mit $0 < \theta < 1$, so dass*

$$\|\Phi(f) - \Phi(g)\| \leq \theta \|f - g\| \quad \text{für alle } f, g \in A.$$

Dann besitzt Φ genau einen Fixpunkt, d.h. es gibt ein eindeutig bestimmtes Element $f_ \in A$ mit*

$$\Phi(f_*) = f_*.$$

Für einen beliebigen Anfangswert $f_0 \in A$ konvergiert die durch $f_k := \Phi(f_{k-1})$ rekursiv definierte Folge $(f_k)_{k \in \mathbb{N}}$ gegen den Fixpunkt f_ .*

Beweis. a) Zur Eindeutigkeit. Seien $f_*, g_* \in A$ zwei Fixpunkte von Φ . Dann gilt

$$\|f_* - g_*\| = \|\Phi(f_*) - \Phi(g_*)\| \leq \theta \|f_* - g_*\|,$$

woraus folgt $\|f_* - g_*\| = 0$, d.h. $f_* = g_*$.

b) Zur Existenz. Sei $f_0 \in A$ beliebig und $f_k := \Phi(f_{k-1})$ für alle $k \geq 1$. Es gilt

$$\begin{aligned} \|f_{k+1} - f_k\| &= \|\Phi(f_k) - \Phi(f_{k-1})\| \\ &\leq \theta \|f_k - f_{k-1}\| \leq \theta^2 \|f_{k-1} - f_{k-2}\| \leq \dots \\ &\leq \theta^k \|f_1 - f_0\| =: \theta^k c. \end{aligned}$$

Daraus folgt für $m > k$

$$\|f_m - f_k\| = \left\| \sum_{i=k}^{m-1} (f_{i+1} - f_i) \right\| \leq \sum_{i=k}^{m-1} \|f_{i+1} - f_i\| \leq \sum_{i=k}^{m-1} \theta^i c \leq \frac{\theta^k c}{1 - \theta}.$$

Dies zeigt, dass die Folge $(f_k)_{k \in \mathbb{N}}$ eine Cauchy-Folge in V ist, also wegen der Vollständigkeit von V gegen ein Element $f_* \in V$ konvergiert. Weil A abgeschlossen ist, liegt f_* sogar in A . Aus der Gleichung

$$f_{k+1} = \Phi(f_k)$$

folgt durch Grenzübergang $k \rightarrow \infty$

$$f_* = \Phi(f_*), \quad \text{q.e.d.}$$

Zum Beweis des nächsten Satzes werden wir den Banachschen Fixpunktsatz anwenden auf den Vektorraum $C_b(U, \mathbb{R}^m)$ aller beschränkten stetigen Funktio-

nen $f: U \rightarrow \mathbb{R}^m$ auf einer Teilmenge $U \subset \mathbb{R}^k$, versehen mit der Supremums-Norm

$$\|f\| := \sup\{\|f(x)\| : x \in U\}.$$

Eine Cauchyfolge bzgl. dieser Norm konvergiert gleichmäßig auf U . Nach §2, Satz 10, ist die Grenzfunktion wieder stetig, besitzt also einen Limes in $C_b(U, \mathbb{R}^m)$. Daher ist $C_b(U, \mathbb{R}^m)$ ein Banachraum.

Satz 2 (über implizite Funktionen). *Seien $U_1 \subset \mathbb{R}^k$ und $U_2 \subset \mathbb{R}^m$ offene Teilmengen und*

$$F : U_1 \times U_2 \longrightarrow \mathbb{R}^m, \quad (x, y) \mapsto F(x, y),$$

eine stetig differenzierbare Abbildung. Sei $(a, b) \in U_1 \times U_2$ ein Punkt mit

$$F(a, b) = 0.$$

Die $m \times m$ -Matrix

$$\frac{\partial F}{\partial y} := \frac{\partial(F_1, \dots, F_m)}{\partial(y_1, \dots, y_m)} := \begin{pmatrix} \frac{\partial F_1}{\partial y_1} & \cdots & \frac{\partial F_1}{\partial y_m} \\ \vdots & & \vdots \\ \frac{\partial F_m}{\partial y_1} & \cdots & \frac{\partial F_m}{\partial y_m} \end{pmatrix}$$

sei im Punkt (a, b) invertierbar. Dann gibt es eine offene Umgebung $V_1 \subset U_1$ von a , eine Umgebung $V_2 \subset U_2$ von b sowie eine stetig differenzierbare Abbildung $g: V_1 \rightarrow V_2 \subset \mathbb{R}^m$ mit $g(a) = b$, so dass

$$F(x, g(x)) = 0 \quad \text{für alle } x \in V_1.$$

Ist $(x, y) \in V_1 \times V_2$ ein Punkt mit $F(x, y) = 0$, so folgt $y = g(x)$.

Bemerkungen

- 1) Der einfachste (aber schon nicht-triviale) Fall ist $k = m = 1$. Dann ist $\frac{\partial F}{\partial y}$ die gewöhnliche partielle Ableitung. Der Leserin sei empfohlen, beim ersten Studium des Satzes an diesen Fall zu denken.
- 2) Man sagt, die Abbildung g entstehe durch Auflösen der Gleichung

$$F(x, y) = 0$$

nach y . Für die Gültigkeit des Satzes ist wesentlich, dass U_1 und U_2 verkleinert werden; in ganz $U_1 \times U_2$ könnte es zu einem gegebenen x meh-

rere y -Werte (oder auch gar keinen) geben, die der Gleichung $F(x, y) = 0$ genügen, vgl. Bild 8.1

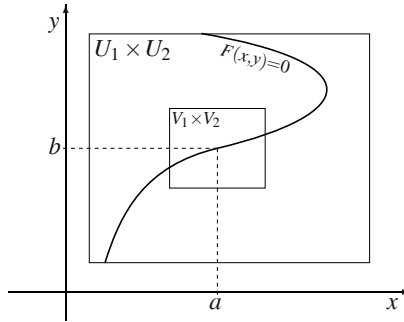


Bild 8.1

- 3) Ist $\frac{\partial F}{\partial y}$ invertierbar im Punkt (a, b) , so ist es auch invertierbar in einer gewissen Umgebung von (a, b) . Dies sieht man so: Die Funktion

$$\delta: U_1 \times U_2 \longrightarrow \mathbb{R}, \quad \delta(x, y) := \det \left(\frac{\partial F}{\partial y}(x, y) \right)$$

ist stetig in $U_1 \times U_2$, da sie ein Polynom in den stetigen Funktionen $\partial F_i / \partial y_j$ ist. Da $\delta(a, b) \neq 0$, gilt auch $\delta(x, y) \neq 0$ für alle (x, y) , die nahe genug bei (a, b) liegen.

- 4) Differenziert man die i -te Komponente der Gleichung $F(x, g(x)) = 0$ partiell nach x_j , so erhält man mit der Kettenregel (§ 6, Corollar zu Satz 3)

$$\frac{\partial F_i}{\partial x_j}(x, g(x)) + \sum_{\mu=1}^m \frac{\partial F_i}{\partial y_\mu}(x, g(x)) \frac{\partial g_\mu}{\partial x_j}(x) = 0, \quad \left(\begin{array}{l} i = 1, \dots, m \\ j = 1, \dots, k \end{array} \right),$$

oder in Matrizen-Schreibweise

$$\frac{\partial F}{\partial x}(x, g(x)) + \frac{\partial F}{\partial y}(x, g(x)) \frac{\partial g}{\partial x}(x) = 0$$

mit den Abkürzungen

$$\begin{aligned} \frac{\partial F}{\partial x} &= \frac{\partial(F_1, \dots, F_m)}{\partial(x_1, \dots, x_k)} = \left(\frac{\partial F_i}{\partial x_j} \right)_{\substack{1 \leq i \leq m \\ 1 \leq j \leq k}}, \\ \frac{\partial F}{\partial y} &= \frac{\partial(F_1, \dots, F_m)}{\partial(y_1, \dots, y_m)} = \left(\frac{\partial F_i}{\partial y_\mu} \right)_{\substack{1 \leq i \leq m \\ 1 \leq \mu \leq m}}, \end{aligned}$$

$$\frac{\partial g}{\partial x} = \frac{\partial(g_1, \dots, g_m)}{\partial(x_1, \dots, x_k)} = \left(\frac{\partial g_\mu}{\partial x_j} \right)_{\substack{1 \leq \mu \leq m \\ 1 \leq j \leq k}}$$

Ist die Matrix $\frac{\partial F}{\partial y}$ im Punkt $(x, g(x))$ invertierbar, erhält man für die Funktional-Matrix der Abbildung g

$$\frac{\partial g}{\partial x}(x) = - \left(\frac{\partial F}{\partial y}(x, g(x)) \right)^{-1} \frac{\partial F}{\partial x}(x, g(x)).$$

Beweis von Satz 2. Wir gehen in mehreren Schritten vor.

a) *Vorbereitungen.* O.B.d.A. sei $(a, b) = (0, 0)$. Wir setzen

$$B := \frac{\partial F}{\partial y}(0, 0) \in GL(m, \mathbb{R})$$

und definieren die Abbildung $G: U_1 \times U_2 \rightarrow \mathbb{R}^m$ durch

$$G(x, y) := y - B^{-1}F(x, y) \quad (1)$$

Da $\frac{\partial G}{\partial y}(x, y) = \mathbb{1} - B^{-1} \frac{\partial F}{\partial y}(x, y)$, wobei $\mathbb{1}$ die m -reihige Einheitsmatrix bezeichnet, folgt

$$\frac{\partial G}{\partial y}(0, 0) = 0.$$

Da alle Komponenten der Matrix $\frac{\partial G}{\partial y}$ stetige Funktionen sind, gibt es Nullumgebungen $W_1 \subset U_1$ und $W_2 \subset U_2$, so dass

$$\left\| \frac{\partial G}{\partial y}(x, y) \right\| \leq \frac{1}{2} \quad \text{für alle } (x, y) \in W_1 \times W_2 \quad (2)$$

Wir wählen ein $r > 0$, so dass

$$V_2 := \{y \in \mathbb{R}^m : \|y\| \leq r\} \subset W_2.$$

Da $G(0, 0) = 0$, gibt es eine offene Nullumgebung $V_1 \subset W_1$, so dass

$$\sup_{x \in V_1} \|G(x, 0)\| =: \varepsilon \leq \frac{r}{2} \quad (3)$$

Aus der Definition (1) folgt

$$F(x, y) = 0 \quad \Longleftrightarrow \quad y = G(x, y),$$

wir haben also die Lösung der Gleichung $F(x, y) = 0$ in ein Fixpunkt-Problem verwandelt. Aus der Abschätzung (2) folgt für alle $x \in V_1$ und $y, \eta \in V_2$

$$\|G(x, y) - G(x, \eta)\| \leq \frac{1}{2} \|y - \eta\| \quad (4)$$

Setzt man $\eta = 0$, so ergibt sich zusammen mit (3) für alle $x \in V_1$

$$\|y\| \leq r \implies \|G(x, y)\| \leq r \quad (5)$$

b) Für jedes feste $x \in V_1$ ist die Abbildung

$$V_2 \ni y \mapsto G(x, y) \in \mathbb{R}^m$$

wegen (5) eine Abbildung der abgeschlossenen Kugel $V_2 \subset \mathbb{R}^m$ in sich, die nach (4) eine Kontraktion ist, also nach dem Banachschen Fixpunktsatz genau einen Fixpunkt hat. Es gibt also zu jedem $x \in V_1$ genau ein $y = g(x) \in V_2$, so dass $G(x, y) = y$, d.h. $F(x, g(x)) = 0$.

c) Wir zeigen jetzt, dass die in b) konstruierte Abbildung $g: V_1 \rightarrow V_2$ sogar stetig ist. Dazu wenden wir den Banachschen Fixpunktsatz auf den Banachraum $C_b(V_1, \mathbb{R}^m)$ aller stetigen und beschränkten Abbildungen $\varphi: V_1 \rightarrow \mathbb{R}^m$ an. Falls

$$\|\varphi\| := \sup\{\|\varphi(x)\| : x \in V_1\} \leq r,$$

so gilt für die durch

$$V_1 \ni x \mapsto \psi(x) := G(x, \varphi(x)) \in \mathbb{R}^m$$

definierte stetige Abbildung $\psi: V_1 \rightarrow \mathbb{R}^m$ nach (5) ebenfalls $\|\psi\| \leq r$, die Zuordnung $\varphi \mapsto \psi$ liefert also eine Abbildung Φ der abgeschlossenen Teilmenge

$$A := \{\varphi \in C_b(V_1, \mathbb{R}^m) : \|\varphi\| \leq r\} = \{\varphi \in C_b(V_1, \mathbb{R}^m) : \varphi(V_1) \subset V_2\}$$

in sich. Aus (4) folgt für $\varphi_1, \varphi_2 \in A$

$$\begin{aligned} \|\Phi(\varphi_1) - \Phi(\varphi_2)\| &= \sup_{x \in V_1} \|G(x, \varphi_1(x)) - G(x, \varphi_2(x))\| \\ &\leq \frac{1}{2} \sup_{x \in V_1} \|\varphi_1(x) - \varphi_2(x)\| = \frac{1}{2} \|\varphi_1 - \varphi_2\|, \end{aligned}$$

die Abbildung $\Phi: A \rightarrow A$ ist also eine Kontraktion und besitzt deshalb genau einen Fixpunkt $g \in A \subset C_b(V_1, \mathbb{R}^m)$. Diese stetige Abbildung $g: V_1 \rightarrow V_2$ erfüllt $G(x, g(x)) = g(x)$, d.h.

$$F(x, g(x)) = 0 \quad \text{für alle } x \in V_1$$

und stimmt natürlich wegen der Eindeutigkeit mit der in b) konstruierten Abbildung überein.

d) Nach evtl. Verkleinerung von V_1 können wir annehmen, dass die Matrix $\frac{\partial F}{\partial y}$ in jedem Punkt $(x, g(x))$, $x \in V_1$, invertierbar ist. Wir zeigen jetzt, dass die Abbildung $g: V_1 \rightarrow \mathbb{R}^m$ differenzierbar ist. An der in Bemerkung 4) angegebenen Formel für die Funktionalmatrix von g sieht man dann, dass g sogar stetig differenzierbar ist.

Es genügt, den Beweis der Differenzierbarkeit von g im Nullpunkt $0 \in V_1 \subset \mathbb{R}^k$ durchzuführen (für die anderen Punkte geht der Beweis analog). Wir setzen

$$A := \frac{\partial F}{\partial x}(0, 0) \in M(m \times k, \mathbb{R}),$$

$$B := \frac{\partial F}{\partial y}(0, 0) \in GL(m, \mathbb{R}).$$

Aus der Definition der Differenzierbarkeit von F im Punkt $(0, 0)$ folgt

$$F(x, y) = Ax + By + \varphi(x, y) \quad \text{mit } \varphi(x, y) = o(\|(x, y)\|).$$

Es gilt $F(x, g(x)) = 0$ für alle $x \in V_1$, d.h.

$$g(x) = -B^{-1}Ax - B^{-1}\varphi(x, g(x)) \quad (6)$$

Für die Differenzierbarkeit von g im Nullpunkt ist also nur zu beweisen, dass

$$\psi(x) := -B^{-1}\varphi(x, g(x)) = o(\|x\|) \quad (7)$$

Dazu zeigen wir zunächst: Es gibt eine Umgebung $V'_1 \subset V_1$ von 0 und eine Konstante $K > 0$, so dass

$$\|g(x)\| \leq K\|x\| \quad \text{für alle } x \in V'_1 \quad (8)$$

Beweis hierfür. Wir setzen

$$c_1 := \|B^{-1}A\|, \quad c_2 := \|B^{-1}\|.$$

Wegen $\varphi(x, y) = o(\|(x, y)\|)$ gibt es zu $\varepsilon := \frac{1}{2c_2}$ eine Umgebung $V' \subset V_1 \times V_2$ von $(0, 0)$, so dass

$$\|\varphi(x, y)\| \leq \varepsilon\|(x, y)\| \leq \frac{1}{2c_2}(\|x\| + \|y\|) \quad \text{für alle } (x, y) \in V'.$$

Wegen der Stetigkeit von g gibt es eine Nullumgebung $V'_1 \subset V_1$, so dass der Graph von $g|_{V'_1}$ ganz in V' enthalten ist. Dann gilt für alle $x \in V'_1$

$$\|\varphi(x, g(x))\| \leq \frac{1}{2c_2}(\|x\| + \|g(x)\|)$$

Die Gleichung (6) liefert nun

$$\|g(x)\| \leq c_1\|x\| + c_2\|\varphi(x, g(x))\| \leq (c_1 + \frac{1}{2})\|x\| + \frac{1}{2}\|g(x)\|,$$

also

$$\|g(x)\| \leq (2c_1 + 1)\|x\| =: K\|x\|.$$

Damit ist (8) bewiesen. Mit (8) ergibt sich für $x \rightarrow 0$

$$\varphi(x, g(x)) = o(\|(x, g(x))\|) = o(\|x\| + \|g(x)\|) = o(\|x\|),$$

woraus die Beziehung (7) folgt. Damit ist Satz 2 vollständig bewiesen.

Höhenlinien

Sei $U \subset \mathbb{R}^2$ eine offene Menge und

$$f : U \longrightarrow \mathbb{R}, \quad (x, y) \mapsto f(x, y),$$

eine stetig differenzierbare Funktion. In § 5 haben wir die “Höhenlinien”

$$N_f(c) := \{(x, y) \in U : f(x, y) = c\}, \quad c \in \mathbb{R},$$

definiert. Wir wollen jetzt zeigen, wie man die Höhenlinien in der Umgebung eines Punktes, in dem $\text{grad } f$ nicht verschwindet, genauer beschreiben kann.

Sei also $(a, b) \in U$, $f(a, b) =: c$ und

$$(\text{grad } f)(a, b) \neq (0, 0),$$

d.h. $\frac{\partial f}{\partial y}(a, b) \neq 0$ oder $\frac{\partial f}{\partial x}(a, b) \neq 0$ (oder beides). Wir wenden auf $F(x, y) := f(x, y) - c$ den Satz 2 an, wobei wir im Fall $\frac{\partial f}{\partial x} \neq 0$ die Rollen von x und y zu vertauschen haben. Es ergibt sich:

a) Falls $\frac{\partial f}{\partial y}(a, b) \neq 0$, existieren offene Intervalle $I_1, I_2 \subset \mathbb{R}$ mit

$$(a, b) \in I_1 \times I_2 \subset U$$

und eine stetig differenzierbare Funktion $\varphi : I_1 \rightarrow I_2$ so dass

$$N_f(c) \cap (I_1 \times I_2) = \{(x, y) \in I_1 \times I_2 : y = \varphi(x)\}.$$

b) Falls $\frac{\partial f}{\partial x}(a, b) \neq 0$, existieren offene Intervalle $J_1, J_2 \subset \mathbb{R}$ mit

$$(a, b) \in J_1 \times J_2 \subset U$$

und eine stetig differenzierbare Funktion $\psi: J_2 \rightarrow J_1$ so dass

$$N_f(c) \cap (J_1 \times J_2) = \{(x, y) \in J_1 \times J_2 : x = \psi(y)\}.$$

Die Höhenlinien lassen sich also in der Umgebung eines Punktes, in dem der Gradient nicht verschwindet, stets als Graph einer Funktion darstellen; entweder y als Funktion von x oder x als Funktion von y , siehe Bild 8.2

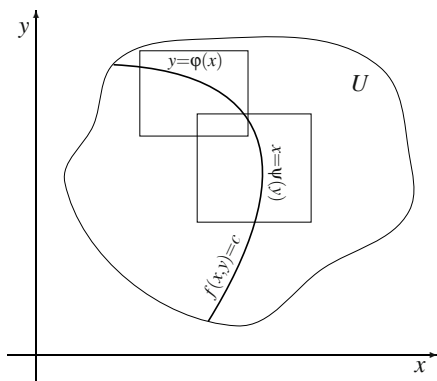


Bild 8.2

(8.3) Beispiel. Betrachten wir die Funktion

$$f: \mathbb{R}^2 \longrightarrow \mathbb{R}, \quad f(x, y) := x^2 + y^2.$$

Es gilt

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

d.h. der Gradient verschwindet nur für $(x, y) = (0, 0)$. Die durch den Nullpunkt gehende “Höhenlinie”

$$N_f(0) = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 = 0\} = \{(0, 0)\}$$

besteht nur aus einem Punkt, ist also keine Linie im eigentlichen Sinn. Für $c < 0$ ist $N_f(c)$ leer, während für $c > 0$ die Menge $N_f(c)$ einen Kreis um den Nullpunkt mit Radius \sqrt{c} darstellt und sich als Vereinigung der folgenden vier Graphen darstellen lässt:

$$\Gamma_1 := \{(x, y) \in \mathbb{R}^2 : -\sqrt{c} < x < \sqrt{c}, y = \sqrt{c - x^2}\},$$

$$\Gamma_2 := \{(x, y) \in \mathbb{R}^2 : -\sqrt{c} < x < \sqrt{c}, y = -\sqrt{c - x^2}\},$$

$$\Gamma_3 := \{(x, y) \in \mathbb{R}^2 : -\sqrt{c} < y < \sqrt{c}, x = \sqrt{c - y^2}\},$$

$$\Gamma_4 := \{(x, y) \in \mathbb{R}^2 : -\sqrt{c} < y < \sqrt{c}, x = -\sqrt{c - y^2}\}.$$

Umkehrabbildungen

Es seien $U_1, U_2 \subset \mathbb{R}^n$ offene Mengen und

$$f : U_1 \longrightarrow U_2$$

eine stetig differenzierbare Abbildung. Wir interessieren uns jetzt für die Frage, wann die Abbildung bijektiv ist und eine stetig differenzierbare Umkehrabbildung $g : U_2 \rightarrow U_1$ besitzt. Eine notwendige Bedingung dafür ist leicht abzuleiten: Sei $a \in U_1$ und $b := f(a) \in U_2$. Aus der Beziehung $g \circ f = \text{id}_{U_1}$ folgt mit Hilfe der Kettenregel

$$Dg(b)Df(a) = \mathbb{1},$$

da die Funktionalmatrix der identischen Abbildung die Einheitsmatrix $\mathbb{1}$ ist. Die Funktionalmatrix $Df(a)$ muss also invertierbar sein. Der nächste Satz zeigt, dass diese Bedingung für die Umkehrbarkeit auch hinreichend ist, wenn man eine Verkleinerung von U_1 und U_2 erlaubt.

Satz 3 (über die Umkehrabbildung). *Sei $U \subset \mathbb{R}^n$ offen und*

$$f : U \longrightarrow \mathbb{R}^n$$

eine stetig differenzierbare Abbildung. Sei $a \in U$ und $b := f(a)$. Die Jacobi-Matrix $Df(a)$ sei invertierbar. Dann gibt es eine offene Umgebung $U_0 \subset U$ von a und eine offene Umgebung V_0 von b , so dass f die Menge U_0 bijektiv auf V_0 abbildet und die Umkehrabbildung

$$g = f^{-1} : V_0 \longrightarrow U_0$$

stetig differenzierbar ist. Es gilt $Dg(b) = (Df(a))^{-1}$.

Beweis. Wir führen diesen Satz auf den Satz über implizite Funktionen zurück und definieren dazu die Funktion

$$F : \mathbb{R}^n \times U \longrightarrow \mathbb{R}^n, \quad (x, y) \mapsto F(x, y) := x - f(y).$$

Es gilt $F(b, a) = 0$. Da $\frac{\partial F}{\partial y}(x, y) = -Df(y)$ und $Df(a)$ invertierbar ist, können wir Satz 2 anwenden. Es gibt also eine offene Umgebung V' von b , eine Umgebung $U' \subset U$ von a und eine stetig differenzierbare Abbildung $g : V' \rightarrow U'$ mit folgenden Eigenschaften:

i) $0 = F(x, g(x)) = x - f(g(x))$, d.h. $f(g(x)) = x$ für alle $x \in V'$.

ii) Ist $(x, y) \in V' \times U'$ mit $F(x, y) = 0$, d.h. $x = f(y)$, so folgt $y = g(x)$.

Aufgrund der Stetigkeit von f gibt es eine offene Umgebung $U_0 \subset U'$ von a mit $f(U_0) \subset V'$. Wegen ii) gilt

$$V_0 := f(U_0) = g^{-1}(U_0).$$

Da g stetig ist, ist V_0 eine offene Umgebung von b . Nach Konstruktion ist $f: U_0 \rightarrow V_0$ bijektiv mit der Umkehrung g , q.e.d.

Bezeichnung. Sind $U, V \subset \mathbb{R}^n$ offene Mengen und

$$f: U \rightarrow V$$

eine bijektive stetig differenzierbare Abbildung, so dass die Umkehrabbildung

$$g := f^{-1}: V \rightarrow U$$

ebenfalls stetig differenzierbar ist, so nennen wir f C^1 -invertierbar.

Ist f sogar 2-mal stetig differenzierbar, so folgt aus der Formel

$$(Dg)(y) = \left((Df)(g(y)) \right)^{-1},$$

dass alle Koeffizienten der Matrix Dg stetig differenzierbar sind, also g ebenfalls 2-mal stetig differenzierbar ist. Durch Induktion beweist man: Ist die C^1 -invertierbare Abbildung s -mal stetig differenzierbar ($s \geq 1$), so auch ihre Umkehrabbildung. Eine solche Abbildung heißt C^s -invertierbar. Dabei kann auch $s = \infty$ sein, falls f und damit f^{-1} beliebig oft stetig differenzierbar sind. Eine C^s -invertierbare Abbildung $f: U \rightarrow V$ nennt man auch *Diffeomorphismus* der Klasse C^s .

(8.4) Beispiel (ebene Polarkoordinaten).

Wir betrachten die Abbildung

$$f: \mathbb{R}_+^* \times \mathbb{R} \rightarrow \mathbb{R}^2, \quad (r, \varphi) \mapsto (r \cos \varphi, r \sin \varphi).$$

Die Funktional-Matrix dieser Abbildung ist

$$Df(r, \varphi) = \frac{\partial(f_1, f_2)}{\partial(r, \varphi)} = \begin{pmatrix} \frac{\partial f_1}{\partial r} & \frac{\partial f_1}{\partial \varphi} \\ \frac{\partial f_2}{\partial r} & \frac{\partial f_2}{\partial \varphi} \end{pmatrix} = \begin{pmatrix} \cos \varphi & -r \sin \varphi \\ \sin \varphi & r \cos \varphi \end{pmatrix}.$$

Da $\det(Df(r, \varphi)) = r > 0$, ist $Df(r, \varphi)$ in allen Punkten $(r, \varphi) \in \mathbb{R}_+^* \times \mathbb{R}$ inver-

tierbar, die Abbildung f also lokal umkehrbar. Es gilt

$$(Df(r, \varphi))^{-1} = \begin{pmatrix} \cos \varphi & -r \sin \varphi \\ \sin \varphi & r \cos \varphi \end{pmatrix}^{-1} = \begin{pmatrix} \cos \varphi & \sin \varphi \\ -\frac{\sin \varphi}{r} & \frac{\cos \varphi}{r} \end{pmatrix}.$$

Setzt man $f(r, \varphi) = (x, y)$, so ist

$$r = \sqrt{x^2 + y^2}, \quad \frac{x}{r} = \cos \varphi, \quad \frac{y}{r} = \sin \varphi.$$

Daher folgt

$$(Df(r, \varphi))^{-1} = \begin{pmatrix} \frac{x}{\sqrt{x^2 + y^2}} & \frac{y}{\sqrt{x^2 + y^2}} \\ \frac{-y}{x^2 + y^2} & \frac{x}{x^2 + y^2} \end{pmatrix} = Dg(x, y),$$

wobei g eine lokale Umkehrung von f ist. In unserem Fall kann man eine solche Umkehrung explizit angeben. Sei etwa $-\pi/2 < \varphi < \pi/2$. Dann folgt $x > 0$. Setzt man

$$V := \mathbb{R}_+^* \times \left] -\frac{\pi}{2}, \frac{\pi}{2} \right[\quad \text{und} \quad V' = \mathbb{R}_+^* \times \mathbb{R},$$

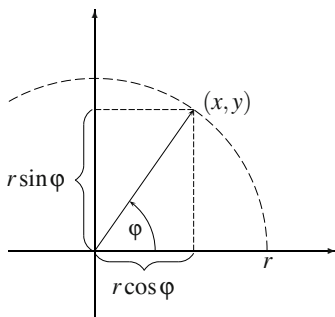
so sind V bzw. V' offene Umgebungen von (r, φ) bzw. (x, y) und die Abbildung $f: V \rightarrow V'$ ist bijektiv mit der Umkehrung $g: V' \rightarrow V$,

$$g(\xi, \eta) = \left(\sqrt{\xi^2 + \eta^2}, \arctan \frac{\eta}{\xi} \right).$$

Durch Berechnung der partiellen Ableitungen der beiden Komponenten von g kann man die oben abgeleitete Formel für die Funktional-Matrix von g direkt bestätigen.

Bild 8.3

Ebene Polarkoordinaten



Die Abbildung f bildet $\mathbb{R}_+^* \times \mathbb{R}$ auf $\mathbb{R}^2 \setminus \{0\}$ ab, sie ist aber nicht global injektiv,

da $f(r, \varphi) = f(r, \varphi + 2k\pi)$ für alle $k \in \mathbb{Z}$. Ist $f(r, \varphi) = (x, y)$, d.h.

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

so heißen (r, φ) die *Polarkoordinaten* des Punktes (x, y) . Dabei ist $r = \sqrt{x^2 + y^2}$ gleich dem Abstand des Punktes (x, y) vom Nullpunkt und φ der (bis auf ganzzahlige Vielfache von 2π eindeutig bestimmte) Winkel zwischen der x -Achse und dem Ortsvektor von (x, y) , siehe Bild 8.3.

AUFGABEN

8.1. Es sei $F: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ die durch

$$F(x, y) := (x^2 - y^2, 2xy)$$

definierte Abbildung. Man berechne die Funktional-Matrix von F und, wo sie existiert, ihre Inverse. Man zeige, dass F surjektiv ist und dass jeder Punkt $(x, y) \in \mathbb{R}^2$, $(x, y) \neq (0, 0)$, genau zwei Urbildpunkte besitzt.

8.2. Man diskutiere die Höhenlinien der Funktion

$$F: \mathbb{R}_+^* \times \mathbb{R}_+^* \rightarrow \mathbb{R}, \quad (x, y) \mapsto xye^{-x-y}$$

und untersuche insbesondere, in welchen Rechtecken $I \times J \subset \mathbb{R}_+^* \rightarrow \mathbb{R}$ die Mengen

$$\{(x, y) \in I \times J : F(x, y) = c\}$$

sich in der Form

$$\{(x, y) \in I \times J : y = \varphi(x)\}$$

bzw.

$$\{(x, y) \in I \times J : x = \psi(y)\}$$

mit differenzierbaren Funktionen $\varphi: I \rightarrow J$ bzw. $\psi: J \rightarrow I$ darstellen lassen.

8.3. Sei $F: \mathbb{R}^3 \rightarrow \mathbb{R}$ die Funktion

$$F(x, y, z) := z^3 + 2xy - 4xz + 2y - 1.$$

Man zeige, dass durch $F(x, y, z) = 0$ in einer Umgebung U von $(x, y) = (1, 1)$ eine differenzierbare Funktion $z = \varphi(x, y)$ mit $\varphi(1, 1) = 1$ implizit definiert ist und berechne die partiellen Ableitungen $\frac{\partial \varphi}{\partial x}$ und $\frac{\partial \varphi}{\partial y}$ im Punkt $(1, 1)$.

§ 9 Untermannigfaltigkeiten

In der Differentialrechnung mehrerer Veränderlichen sind die k -dimensionalen Untermannigfaltigkeiten des \mathbb{R}^n das krummlinige Analogon der k -dimensionalen affinen Unterräume in der Linearen Algebra. Lokal kann eine k -dimensionale Untermannigfaltigkeit im \mathbb{R}^n entweder durch eine Parameterdarstellung mit k reellen Parametern beschrieben werden oder als Nullstellengebilde von $n - k$ unabhängigen differenzierbaren Funktionen. In diesem Paragraphen besprechen wir auch Tangential- und Normalen-Vektoren an Untermannigfaltigkeiten und leiten die Methode der Lagrangeschen Multiplikatoren zur Bestimmung von Extrema unter Nebenbedingungen her.

Definition (Immersion). Sei $T \subset \mathbb{R}^k$ eine offene Teilmenge. Eine stetig differenzierbare Abbildung

$$\varphi: T \longrightarrow \mathbb{R}^n, \quad (t_1, \dots, t_k) \mapsto \varphi(t_1, \dots, t_k),$$

heißt *Immersion*, wenn der Rang der Funktional-Matrix

$$D\varphi = \frac{\partial(\varphi_1, \dots, \varphi_n)}{\partial(t_1, \dots, t_k)} := \left(\frac{\partial \varphi_i}{\partial t_j} \right)_{\substack{1 \leq i \leq n \\ 1 \leq j \leq k}} = \begin{pmatrix} \frac{\partial \varphi_1}{\partial t_1} & \frac{\partial \varphi_1}{\partial t_2} & \dots & \frac{\partial \varphi_1}{\partial t_k} \\ \frac{\partial \varphi_2}{\partial t_1} & \frac{\partial \varphi_2}{\partial t_2} & \dots & \frac{\partial \varphi_2}{\partial t_k} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial \varphi_n}{\partial t_1} & \frac{\partial \varphi_n}{\partial t_2} & \dots & \frac{\partial \varphi_n}{\partial t_k} \end{pmatrix}$$

in jedem Punkt $t \in T$ gleich k ist.

Bemerkungen.

- 1) Die Funktional-Matrix $D\varphi$ ist eine $n \times k$ -Matrix. Wenn der Rang gleich k sein soll, ist notwendig $n \geq k$.
- 2) Die Bedingung $\text{Rang}(D\varphi)(t) = k$ ist gleichbedeutend damit, dass die Vektoren

$$\frac{\partial \varphi}{\partial t_1}(t), \dots, \frac{\partial \varphi}{\partial t_k}(t) \in \mathbb{R}^n$$

linear unabhängig sind.

- 3) Aus der Linearen Algebra ist ferner folgendes Determinanten-Kriterium bekannt: $\text{Rang}(D\varphi)(t) = k$ gilt genau dann, wenn Indizes

$$1 \leq i_1 < i_2 < \dots < i_k \leq n$$

existieren, so dass

$$\det \frac{\partial(\varphi_{i_1}, \dots, \varphi_{i_k})}{\partial(t_1, \dots, t_k)}(t) \neq 0.$$

Da diese Determinante eine stetige Funktion von t ist, folgt daraus: Ist der Rang von $D\varphi$ in einem Punkt $t_* \in T$ gleich k , so auch in allen Punkten t einer kleinen Umgebung von t_* .

- 4) Sei speziell $k = 1$ und $T \subset \mathbb{R}^1$ ein offenes Intervall. In diesem Fall ist eine Immersion $\varphi: T \rightarrow \mathbb{R}^n$ nichts anderes als eine nicht-singuläre Kurve, vgl. die Definition in § 4.
- 5) Ist $k = n$, also T offen in \mathbb{R}^n , so ist eine Immersion $\varphi: T \rightarrow \mathbb{R}^n$ nach dem Satz über die Umkehrabbildung (§ 8, Satz 3) eine lokal umkehrbare Abbildung.

Satz 1. Sei $T \subset \mathbb{R}^k$ offen und $\varphi: T \rightarrow \mathbb{R}^n$ eine Immersion. Dann gibt es zu jedem Punkt $t \in T$ eine offene Umgebung $V \subset T$, so dass die Beschränkung

$$\varphi|_V \rightarrow \varphi(V) \subset \mathbb{R}^n$$

injektiv ist und einen Homöomorphismus von V auf $\varphi(V)$ darstellt.

Beweis. Für $k = n$ ist dies im Satz über die Umkehrabbildung enthalten. Wir können deshalb $k < n$ annehmen. Nach der Vorbemerkung 3) gibt es zu einem vorgegebenen Punkt $t_* \in T$ Indizes $1 \leq i_1 < i_2 < \dots < i_k \leq n$, so dass die $k \times k$ -Matrix

$$\frac{\partial(\varphi_{i_1}, \dots, \varphi_{i_k})}{\partial(t_1, \dots, t_k)}(t_*)$$

invertierbar ist. Um die Schreibweise zu vereinfachen, nehmen wir o.B.d.A. an, dass $(i_1, i_2, \dots, i_k) = (1, 2, \dots, k)$. Auf die Abbildung

$$\tilde{\varphi} := (\varphi_1, \dots, \varphi_k) : T \longrightarrow \mathbb{R}^k$$

können wir jetzt den Satz über die Umkehrabbildung anwenden. Es gibt also eine offene Umgebung $V \subset T$ von t_* und eine offene Umgebung $U \subset \mathbb{R}^k$ von $\tilde{\varphi}(t_*)$, so dass

$$\tilde{\varphi} : V \longrightarrow U$$

bijektiv ist und eine stetig differenzierbare Umkehrabbildung

$$\tilde{\psi}: U \longrightarrow V \subset \mathbb{R}^k$$

besitzt. Die Beschränkung von φ auf V lässt sich schreiben als

$$\varphi = (\tilde{\varphi}, \check{\varphi}): V \rightarrow U \times \mathbb{R}^{n-k} \subset \mathbb{R}^n \quad \text{mit } \check{\varphi} := (\varphi_{k+1}, \dots, \varphi_n).$$

Da $\tilde{\varphi}: V \rightarrow U$ bijektiv ist, ist auch $\varphi: V \rightarrow \varphi(V) \subset U \times \mathbb{R}^{n-k}$ bijektiv und sogar ein Homöomorphismus, denn es besitzt die stetige Umkehrabbildung

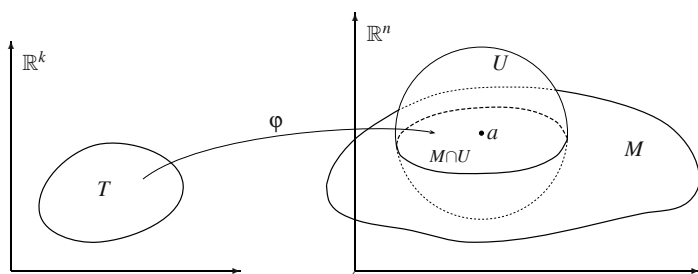
$$\psi: \varphi(V) \longrightarrow V, \quad \psi(x_1, \dots, x_k, \dots, x_n) := \tilde{\psi}(x_1, \dots, x_k).$$

Damit ist Satz 1 bewiesen.

Definition (Untermannigfaltigkeit). Eine Teilmenge $M \subset \mathbb{R}^n$ heißt *k-dimensionale Untermannigfaltigkeit* von \mathbb{R}^n , wenn es zu jedem Punkt $a \in M$ eine offene Umgebung $U \subset \mathbb{R}^n$ gibt, sowie eine offene Teilmenge $T \subset \mathbb{R}^k$ und eine Immersion

$$\varphi: T \longrightarrow \mathbb{R}^n,$$

so dass φ die Menge T homöomorph auf $\varphi(T) = M \cap U$ abbildet, siehe Bild 9.1.



↙ **Bild 9.1**

Man nennt dann

$$\varphi: T \longrightarrow M \cap U$$

eine *Parameterdarstellung* oder *lokales Koordinatensystem* der Untermannigfaltigkeit M in einer Umgebung von a . Ist $(t_1, \dots, t_k) \in T$ und

$$p := \varphi(t_1, \dots, t_k),$$

so heißen t_1, \dots, t_k die lokalen Koordinaten des Punktes $p \in M$ (bzgl. φ).

Die Zahl $n - k$ heißt die *Codimension* der Untermannigfaltigkeit. Untermannigfaltigkeiten der Codimension 1 nennt man auch *Hyperflächen*.

Beispiel

(9.1) Rotationsflächen. Sei $I \subset \mathbb{R}$ ein offenes Intervall und

$$\alpha: I \rightarrow \mathbb{R}^2, \quad t \mapsto \alpha(t) = (\alpha_1(t), \alpha_2(t))$$

eine stetig differenzierbare ebene Kurve, die wir uns in der (x, z) -Ebene des \mathbb{R}^3 mit Koordinaten x, y, z vorstellen, d.h. $x(t) = \alpha_1(t)$, $z(t) = \alpha_2(t)$. Wir setzen außerdem voraus, dass die Kurve regulär ist, d.h. $\alpha'(t) \neq (0, 0)$ für alle $t \in I$. Wird die Kurve um die z -Achse rotiert, so entsteht eine Fläche mit der Parameterdarstellung

$$F: I \times \mathbb{R} \longrightarrow \mathbb{R}^3, \quad (t, \varphi) \mapsto F(t, \varphi) = \begin{pmatrix} \alpha_1(t) \cos \varphi \\ \alpha_1(t) \sin \varphi \\ \alpha_2(t) \end{pmatrix}.$$

Für die Funktionalmatrix von F ergibt sich

$$DF(t, \varphi) = \frac{\partial(F_1, F_2, F_3)}{\partial(t, \varphi)} = \begin{pmatrix} \alpha'_1(t) \cos \varphi & -\alpha_1(t) \sin \varphi \\ \alpha'_1(t) \sin \varphi & \alpha_1(t) \cos \varphi \\ \alpha'_2(t) & 0 \end{pmatrix}.$$

Es ist leicht nachzuprüfen, dass $DF(t, \varphi)$ genau dann den Rang 2 hat, falls $\alpha_1(t) \neq 0$. Falls also die gegebene Kurve α die z -Achse nicht schneidet, ist F eine Immersion und liefert eine Parameterdarstellung der Rotationsfläche.

Wir geben zwei Beispiele.

a) Die *Sphäre* vom Radius $r > 0$ im \mathbb{R}^3 .

Sie entsteht durch Rotation der Kreislinie

$$\mathbb{R} \ni \vartheta \mapsto \begin{pmatrix} x(\vartheta) \\ z(\vartheta) \end{pmatrix} = \begin{pmatrix} r \sin \vartheta \\ r \cos \vartheta \end{pmatrix}$$

um die z -Achse. Die zugehörige Parameterdarstellung der Rotationsfläche ist

$$(\vartheta, \varphi) \mapsto F(\vartheta, \varphi) = \begin{pmatrix} r \sin \vartheta \cos \varphi \\ r \sin \vartheta \sin \varphi \\ r \cos \vartheta \end{pmatrix}.$$

Damit DF den Rang 2 hat, beschränken wir den Parameter ϑ auf den Bereich $0 < \vartheta < \pi$. Dadurch wird nur der ‘Nordpol’ $(0, 0, r)$ und der ‘Südpol’ $(0, 0, -r)$ der Sphäre (die zu $\vartheta = 0$ bzw. $\vartheta = \pi$ gehören würden) ausgeschlossen. (ϑ, φ) sind die Polarkoordinaten auf der Sphäre. Der Parameter ϑ heißt Poldistanz, der Parameter φ ist die ‘geographische Länge’. Die ‘geographische Breite’ β steht zu ϑ in der Beziehung $\beta = \pi/2 - \vartheta$.

b) Der *Torus*.

Sei $R > r > 0$. Der Torus mit Radien r, R entsteht durch Rotation der Kreislinie

$$\mathbb{R} \ni t \mapsto \begin{pmatrix} x(t) \\ z(t) \end{pmatrix} = \begin{pmatrix} R + r \cos t \\ r \sin t \end{pmatrix}$$

um die z -Achse. Da dieser Kreis die z -Achse nicht schneidet, ist

$$F : \mathbb{R} \times \mathbb{R} \longrightarrow \mathbb{R}^3, \quad (t, s) \mapsto F(t, s) = \begin{pmatrix} (R + r \cos t) \cos s \\ (R + r \cos t) \sin s \\ r \sin t \end{pmatrix}$$

eine Immersion; das Bild $F(\mathbb{R} \times \mathbb{R})$ ist der Torus. Natürlich ist F periodisch in beiden Variablen mit der Periode 2π .

Man kann Untermannigfaltigkeiten des \mathbb{R}^n auch auf andere Weise als durch Parameterdarstellungen beschreiben. Der folgende Satz beschreibt einige andere nützliche Möglichkeiten.

Satz 2. *Eine Teilmenge $M \subset \mathbb{R}^n$ ist genau dann eine k -dimensionale Untermannigfaltigkeit, wenn eine der folgenden äquivalenten Bedingungen erfüllt ist:*

- a) (Beschreibung durch Gleichungen) *Zu jedem Punkt $a \in M$ gibt es eine offene Umgebung $U \subset \mathbb{R}^n$ und $n - k$ stetig differenzierbare Funktionen*

$$f_j : U \longrightarrow \mathbb{R}, \quad j = 1, \dots, n - k,$$

so dass

$$M \cap U = \{x \in U : f_1(x) = \dots = f_{n-k}(x) = 0\}$$

und

$$\text{Rang} \frac{\partial(f_1, \dots, f_{n-k})}{\partial(x_1, \dots, x_n)}(x) = n - k \quad \text{für alle } x \in M \cap U.$$

- b) (Darstellung als Graph) Zu jedem Punkt $a \in M$ gibt es nach evtl. Umnummerierung der Koordinaten offene Umgebungen $U' \subset \mathbb{R}^k$ von $a' := (a_1, \dots, a_k)$ und $U'' \subset \mathbb{R}^{n-k}$ von $a'' := (a_{k+1}, \dots, a_n)$ und eine stetig differenzierbare Abbildung

$$g : U' \longrightarrow U'' \subset \mathbb{R}^{n-k}$$

so dass

$$M \cap (U' \times U'') = \{(x', x'') \in U' \times U'' : x'' = g(x')\}.$$

- c) (Transformation in Ebene) Zu jedem Punkt $a \in M$ gibt es eine offene Umgebung $U \subset \mathbb{R}^n$, eine offene Menge $V \subset \mathbb{R}^n$ und eine C^1 -invertierbare Abbildung $\Phi : U \rightarrow V$, so dass

$$\Phi(M \cap U) = E_k \cap V.$$

Dabei bezeichnet $E_k \subset \mathbb{R}^n$ die k -dimensionale Ebene

$$E_k := \{x = (x_1, \dots, x_n) \in \mathbb{R}^n : x_{k+1} = \dots = x_n = 0\}.$$

Beweis. Wir bezeichnen mit (P) die Bedingung aus der Definition einer k -dimensionalen Untermannigfaltigkeit:

Zu jedem Punkt $a \in M$ existiert eine offene Umgebung $U \subset \mathbb{R}^n$, sowie eine offene Teilmenge $T \subset \mathbb{R}^k$ und eine Immersion

$$\varphi : T \longrightarrow \mathbb{R}^n,$$

so dass φ die Menge T homöomorph auf $\varphi(T) = M \cap U$ abbildet.

Sei $t_* \in T$ der Punkt mit $\varphi(t_*) = a$.

Wir beweisen Satz 2 nach folgendem Schema:

$$(P) \implies (b) \implies (a) \implies (c) \implies (P)$$

(P) \Rightarrow (b) Wir können o.B.d.A. annehmen, dass

$$\det \frac{\partial(\varphi_1, \dots, \varphi_k)}{\partial(t_1, \dots, t_k)}(t_*) \neq 0.$$

Nach dem Satz über die Umkehrabbildung bildet dann

$$\tilde{\varphi} := (\varphi_1, \dots, \varphi_k)$$

eine Umgebung $T_1 \subset T$ von t_* bijektiv und C^1 -invertierbar auf eine offene Teilmenge $U_1 \subset \mathbb{R}^k$ ab. Sei $\psi: U_1 \rightarrow T_1$ die Umkehrabbildung von $\tilde{\varphi}$. Dann hat

$$G := \varphi \circ \psi: U_1 \longrightarrow \mathbb{R}^n$$

die Gestalt

$$G(x_1, \dots, x_k) = (x_1, \dots, x_k, g_{k+1}(x'), \dots, g_n(x')), \quad x' = (x_1, \dots, x_k).$$

Die Abbildung

$$g := (g_{k+1}, \dots, g_n): U' \rightarrow \mathbb{R}^{n-k}$$

hat dann für eine geeignete Verkleinerung $U' \subset U_1$ die in (b) geforderten Eigenschaften.

(b) \Rightarrow (a) Wird M in einer Umgebung von a als Graph der Funktion $g = (g_{k+1}, \dots, g_n)$ dargestellt, so ist M dort Lösungsmenge der $n - k$ Gleichungen

$$f_j(x_1, \dots, x_n) := x_{k+j} - g_{k+j}(x_1, \dots, x_k) = 0, \quad j = 1, \dots, n - k.$$

Da

$$\frac{\partial(f_1, \dots, f_{n-k})}{\partial(x_{k+1}, \dots, x_n)}$$

die Einheitsmatrix ist, ist die Rangbedingung von (a) automatisch erfüllt.

(a) \Rightarrow (c) Nach Umnummerierung der Koordinaten können wir annehmen, dass

$$\det \frac{\partial(f_1, \dots, f_{n-k})}{\partial(x_{k+1}, \dots, x_n)}(a) \neq 0.$$

Dann ist auch die Funktionalmatrix

$$\frac{\partial(x_1, \dots, x_k, f_1, \dots, f_{n-k})}{\partial(x_1, \dots, x_k, x_{k+1}, \dots, x_n)}$$

in einer Umgebung von a invertierbar, also bildet

$$\Phi := (x_1, \dots, x_k, f_1, \dots, f_{n-k})$$

eine Umgebung von a C^1 -invertierbar auf eine offene Menge in \mathbb{R}^n ab. Das Nullstellengebilde von f_1, \dots, f_{n-k} wird dabei auf die Ebene

$$\{(x_1, \dots, x_n) : x_{k+1} = \dots = x_n = 0\}$$

abgebildet.

(c) \Rightarrow (P) Sei $\Psi : V \rightarrow U$ die Umkehrabbildung von $\Phi : U \rightarrow V$. Dann liefert

$$(t_1, \dots, t_k) \mapsto \varphi(t_1, \dots, t_k) := \Psi(t_1, \dots, t_k, 0, \dots, 0)$$

eine Parameter-Darstellung von M in einer Umgebung von a .

Tangential- und Normalenvektoren

Unter einem Tangentialvektor an eine Untermannigfaltigkeit versteht man einen Tangentialvektor einer in der Untermannigfaltigkeit verlaufenden Kurve; ein Normalenvektor ist ein Vektor, der auf der Untermannigfaltigkeit senkrecht steht. Dies lässt sich so präzisieren:

Definition. Sei $M \subset \mathbb{R}^n$ eine Untermannigfaltigkeit und $a \in M$. Ein Vektor $v \in \mathbb{R}^n$ heißt *Tangentialvektor* an M im Punkt a , wenn es eine stetig differenzierbare Kurve

$$\alpha :]-\varepsilon, \varepsilon[\longrightarrow M \subset \mathbb{R}^n, \quad (\varepsilon > 0)$$

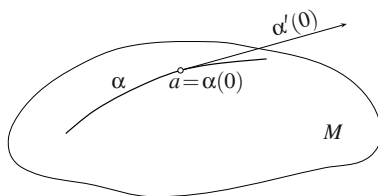
gibt mit

$$\alpha(0) = a \quad \text{und} \quad \alpha'(0) = v,$$

siehe Bild 9.2.

Die Gesamtheit aller Tangentialvektoren an M in a werde mit $T_a(M)$ bezeichnet. Ein *Normalenvektor* von M in a ist ein Vektor $w \in \mathbb{R}^n$, der auf allen Tangentialvektoren $v \in T_a(M)$ senkrecht steht (bzgl. der kanonischen euklidischen Metrik von \mathbb{R}^n). Die Menge aller Normalenvektoren von M in a werde mit $N_a(M)$ bezeichnet.

Satz 3. Sei $M \subset \mathbb{R}^n$ eine k -dimensionale Untermannigfaltigkeit und $a \in M$ ein Punkt. Dann gilt:

**Bild 9.2** Tangentialvektor

a) $T_a(M)$ ist ein k -dimensionaler Vektorraum. Eine Basis von $T_a(M)$ lässt sich so erhalten: Sei

$$\varphi : V \longrightarrow M \subset \mathbb{R}^n$$

ein lokales Koordinatensystem von M in der Umgebung von a , d.h. $V \subset \mathbb{R}^k$ offen und φ eine Immersion, die V homöomorph auf $M \cap U$ abbildet, wobei $U \subset \mathbb{R}^n$ eine offene Umgebung von a ist. Sei $t_* \in V$ der Punkt mit $\varphi(t_*) = a$. Dann bilden die Vektoren

$$\frac{\partial \varphi}{\partial t_1}(t_*), \dots, \frac{\partial \varphi}{\partial t_k}(t_*)$$

eine Basis von $T_a(M)$.

b) $N_a(M)$ ist ein $(n - k)$ -dimensionaler Vektorraum. Eine Basis von $N_a(M)$ lässt sich so erhalten: Seien $f_1, \dots, f_{n-k} : U \longrightarrow \mathbb{R}$ stetig differenzierbare Funktionen in einer offenen Umgebung $U \subset \mathbb{R}^n$ von a mit

$$M \cap U = \{x \in U : f_1(x) = \dots = f_{n-k}(x) = 0\}$$

und

$$\text{Rang} \frac{\partial(f_1, \dots, f_{n-k})}{\partial(x_1, \dots, x_n)}(a) = n - k.$$

Dann bilden die Vektoren

$$\text{grad } f_j(a), \quad j = 1, \dots, n - k,$$

eine Basis von $N_a(M)$.

Beweis. Wir bezeichnen mit V den Vektorraum, der von den Vektoren

$$\frac{\partial \varphi}{\partial t_1}(t_*), \dots, \frac{\partial \varphi}{\partial t_k}(t_*)$$

aufgespannt wird und mit W den Vektorraum, der von den Vektoren

$$\operatorname{grad} f_j(a), \quad j = 1, \dots, n-k,$$

aufgespannt wird und zeigen

$$(a) \quad V \subset T_a(M), \quad (b) \quad W \subset N_a(M).$$

Aus Dimensionsgründen folgt dann in (a) und (b) sogar Gleichheit.

(a) Sei $v \in V$, also

$$v = \sum_{i=1}^k c_i \frac{\partial \varphi}{\partial t_i}(t_*) \quad \text{mit} \quad c_i \in \mathbb{R}.$$

Wir definieren die Kurve

$$\alpha :]-\varepsilon, \varepsilon[\longrightarrow M \subset \mathbb{R}^n$$

durch

$$\alpha(s) := \varphi(t_* + (c_1, \dots, c_k)s).$$

Dies ist definiert für $|s| < \varepsilon$, wenn $\varepsilon > 0$ genügend klein gewählt worden ist. Mit der Kettenregel folgt

$$\frac{d\alpha}{ds}(0) = \sum_{i=1}^k c_i \frac{\partial \varphi}{\partial t_i}(t_*) = v \in T_a(M).$$

(b) Aufgrund seiner Definition ist $N_a(M)$ ein Vektorraum. Es genügt also zu beweisen, dass $\operatorname{grad} f_j(a) \in N_a(M)$ für alle j . Dazu müssen wir zeigen, dass $\operatorname{grad} f_j(a)$ auf jedem Tangentialvektor senkrecht steht. Sei also $v \in T_a(M)$ beliebig. Dann ist $v = \alpha'(0)$ mit einer ganz in M verlaufenden stetig differenzierbaren Kurve

$$\alpha :]-\varepsilon, \varepsilon[\longrightarrow M \quad \text{mit} \quad \alpha(0) = a.$$

Es gilt $f_j(\alpha(t)) = 0$ für alle $t \in]-\varepsilon, \varepsilon[$. Differenziert man diese Beziehung nach

t , erhält man

$$0 = \sum_{i=1}^n \frac{\partial f_j}{\partial x_i}(a) \alpha'_i(0) = \langle \text{grad } f_j(a), \alpha'(0) \rangle = \langle \text{grad } f_j(a), v \rangle$$

also sind $\text{grad } f_j(a)$ und v orthogonal, q.e.d.

Extrema mit Nebenbedingungen

Satz 4. Sei $U \subset \mathbb{R}^n$ offen und $M \subset U$ eine r -codimensionale Untermannigfaltigkeit,

$$M = \{x \in U : g_1(x) = \dots = g_r(x) = 0\},$$

mit stetig differenzierbaren Funktionen $g_j : U \rightarrow \mathbb{R}$ und

$$\text{Rang} \frac{\partial(g_1, \dots, g_r)}{\partial(x_1, \dots, x_n)}(x) = r \quad \text{für alle } x \in M.$$

Weiter sei $F : U \rightarrow \mathbb{R}$ eine stetig differenzierbare Funktion, so dass $F|_M$ in einem Punkt $a \in M$ ein lokales Maximum (Minimum) besitzt, d.h. es gibt eine Umgebung $V \subset U$ von a mit

$$F(x) \leq F(a) \quad (\text{bzw. } F(x) \geq F(a)) \quad \text{für alle } x \in M \cap V.$$

Dann ist $\text{grad } F(a)$ ein Normalenvektor von M in a , d.h. es existieren Konstanten $\lambda_1, \dots, \lambda_r \in \mathbb{R}$, so dass

$$\text{grad } F(a) = \sum_{j=1}^r \lambda_j \text{grad } g_j(a).$$

Man sagt, der Punkt a sei ein lokales Extremum von F unter der Nebenbedingung $\{g_1 = \dots = g_r = 0\}$. Die λ_j werden *Lagrangesche Multiplikatoren* genannt.

Beweis. Ist

$$\alpha :]-\varepsilon, \varepsilon[\longrightarrow M \subset \mathbb{R}^n$$

eine stetig differenzierbare Kurve in M mit $\alpha(0) = a$, so hat die Funktion

$$t \mapsto F(\alpha(t))$$

bei $t = 0$ ein lokales Extremum, also gilt

$$0 = \frac{d}{dt} F(\alpha(t)) \Big|_{t=0} = \sum_{i=1}^n \frac{\partial F}{\partial x_i}(a) \alpha'_i(0) = \langle \text{grad} F(a), \alpha'(0) \rangle,$$

d.h. $\text{grad} F(a)$ steht senkrecht auf allen Tangentialvektoren von M in a , q.e.d.

(9.2) Beispiel. Sei $A = (a_{ij}) \in M(n \times n, \mathbb{R})$ eine symmetrische $n \times n$ -Matrix und $F: \mathbb{R}^n \rightarrow \mathbb{R}$ die zugehörige quadratische Form, d.h.

$$F(x) = \langle x, Ax \rangle = \sum_{i,j} a_{ij} x_i x_j.$$

Wir wollen die Extrema von F unter der Nebenbedingung $\|x\| = 1$ untersuchen. Dazu setzen wir

$$g(x) := \langle x, x \rangle - 1 = \sum_i x_i^2 - 1,$$

$$M := \{x \in \mathbb{R}^n : g(x) = 0\} = \{x \in \mathbb{R}^n : \|x\| = 1\}$$

und wenden Satz 4 an. Da $\partial g / \partial x_k = 2x_k$, gilt

$$\text{grad} g(x) = 2x \neq 0 \quad \text{für alle } x \in M.$$

Weiter berechnet man

$$\begin{aligned} \frac{\partial F}{\partial x_k}(x) &= \frac{\partial}{\partial x_k} \sum_{i,j} a_{ij} x_i x_j \\ &= \sum_{i,j} a_{ij} \frac{\partial x_i}{\partial x_k} x_j + \sum_{i,j} a_{ij} x_i \frac{\partial x_j}{\partial x_k} \\ &= \sum_j a_{kj} x_j + \sum_i a_{ik} x_i = 2 \sum_i a_{ki} x_i, \end{aligned}$$

da $a_{ki} = a_{ik}$. Das bedeutet

$$\text{grad} F(x) = 2Ax.$$

Daher lautet die notwendige Bedingung aus Satz 4 für die Existenz eines lokalen Extremums von F auf M im Punkt $a \in M$

$$Aa = \lambda a \quad \text{für ein geeignetes } \lambda \in \mathbb{R},$$

d.h. a ist ein Eigenvektor von A und der Lagrangesche Multiplikator ist der zugehörige Eigenwert. Andererseits wissen wir, da M kompakt ist, dass die ste-

tige Funktion F auf M ihr Maximum in einem gewissen Punkt $a \in M$ annimmt. Dieses a muss nach dem gerade Gesagten ein Eigenvektor von A sein. Da

$$F(a) = \langle a, Aa \rangle = \langle a, \lambda a \rangle = \lambda,$$

ist der Funktionswert an dieser Stelle gleich dem Eigenwert von a . Damit folgt also, dass die Funktion F auf M ihr absolutes Maximum in einem Eigenvektor zum größten Eigenwert annimmt. (Analoges gilt für das absolute Minimum.) Gleichzeitig ist damit bewiesen, dass jede reelle symmetrische Matrix mindestens einen reellen Eigenvektor mit einem reellen Eigenwert besitzt. (Daraus kann man dann durch Induktion schließen, dass alle Eigenwerte reell sind.)

AUFGABEN

9.1. Die Funktionen $f, g: \mathbb{R}^3 \rightarrow \mathbb{R}$ seien definiert durch

$$f(x, y, z) := x^2 + xy - y - z, \quad g(x, y, z) := 2x^2 + 3xy - 2y - 3z.$$

Man zeige, dass

$$C := \{(x, y, z) \in \mathbb{R}^3 : f(x, y, z) = g(x, y, z) = 0\}$$

eine eindimensionale Untermannigfaltigkeit des \mathbb{R}^3 ist, und dass

$$\varphi: \mathbb{R} \rightarrow \mathbb{R}^3, \quad \varphi(t) = (t, t^2, t^3)$$

eine globale Parameterdarstellung von C ist.

9.2. Die Funktionen $f_i: \mathbb{R}^4 \rightarrow \mathbb{R}$, $i = 1, 2, 3$, seien definiert durch

$$f_1(x_1, \dots, x_4) = x_1x_3 - x_2^2,$$

$$f_2(x_1, \dots, x_4) = x_2x_4 - x_3^2$$

$$f_3(x_1, \dots, x_4) = x_1x_4 - x_2x_3$$

Man zeige, dass $M := \{x \in \mathbb{R}^4 \setminus \{0\} : f_1(x) = f_2(x) = f_3(x) = 0\}$ eine 2-dimensionale Untermannigfaltigkeit des \mathbb{R}^4 ist.

9.3. Die Menge $M(n \times n, \mathbb{R})$ aller reellen $n \times n$ -Matrizen

$$X = \begin{pmatrix} x_{11} & x_{12} & \dots & x_{1n} \\ x_{21} & x_{22} & \dots & x_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \dots & x_{nn} \end{pmatrix}, \quad x_{ij} \in \mathbb{R},$$

werde mit dem \mathbb{R}^{n^2} mit Koordinaten $x_{11}, x_{12}, \dots, x_{nn}$ identifiziert.

Man zeige: Die *spezielle lineare Gruppe*

$$SL(n, \mathbb{R}) := \{A \in M(n \times n, \mathbb{R}) : \det A = 1\}$$

ist eine 1-codimensionale Untermannigfaltigkeit von $M(n \times n, \mathbb{R})$.

9.4. Wie in Aufgabe 9.3 werde $M(n \times n, \mathbb{R})$ mit \mathbb{R}^{n^2} identifiziert. Es sei

$$O(n) = \{A \in M(n \times n, \mathbb{R}) : A^\top A = E\}$$

die Menge aller orthogonalen $n \times n$ -Matrizen. (A^\top bezeichne die zu A transponierte Matrix.)

Man zeige, dass $O(n)$ eine kompakte Untermannigfaltigkeit von $M(n \times n, \mathbb{R})$ der Dimension $\frac{n(n-1)}{2}$ ist.

9.5. Seien $U \subset \mathbb{R}^n$ und $V \subset \mathbb{R}^n$ offene Mengen und $\Phi: U \rightarrow V$ ein Diffeomorphismus. Weiter sei $M \subset U$ eine k -dimensionale Untermannigfaltigkeit. Man beweise:

a) $M' := \Phi(M) \subset V$ ist eine k -dimensionale Untermannigfaltigkeit.

b) Sei $a \in M$ und $b := \Phi(a) \in M'$. Dann gilt

$$D\Phi(a)(T_a(M)) = T_b(M').$$

c) Gilt eine zu b) analoge Aussage auch für den Normalenvektorraum?

9.6. Man bestimme die Maxima und Minima der Funktion

$$f: \mathbb{R}^2 \rightarrow \mathbb{R}, \quad f(x, y) := 4x^2 - 3xy,$$

auf der Kreisscheibe

$$K := \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \leq 1\}.$$

Anleitung. Man berechne zunächst die lokalen Extrema von f im Innern von K und dann auf dem Rand von K , d.h. unter der Nebenbedingung $x^2 + y^2 = 1$.

§ 10 Integrale, die von einem Parameter abhängen

In diesem Paragraphen beschäftigen wir uns mit folgendem Problem: Sei $f(x, y)$ eine Funktion von zwei Variablen x, y . Für einen festen y -Wert werde die Funktion über ein Intervall $a \leq x \leq b$ integriert. Das Integral hängt dann vom gewählten y -Wert ab, es entsteht also eine Funktion φ des “Parameters” y . Es interessiert nun, unter welchen Voraussetzungen an f die Funktion φ stetig bzw. differenzierbar von y abhängt. Die erhaltenen Ergebnisse werden wir benutzen, um die sog. Eulerschen Differentialgleichungen der Variationsrechnung abzuleiten.

Hilfssatz 1. *Sei $[a, b] \subset \mathbb{R}$ ein kompaktes Intervall und $U \subset \mathbb{R}^m$ eine beliebige Teilmenge. Die Funktion*

$$f : [a, b] \times U \longrightarrow \mathbb{R}, \quad (x, y) \mapsto f(x, y),$$

sei stetig. Weiter sei $(y_k)_{k \in \mathbb{N}}$ eine konvergente Punktfolge in U mit

$$c := \lim_{k \rightarrow \infty} y_k \in U.$$

Dann konvergieren die Funktionen

$$F_k : [a, b] \longrightarrow \mathbb{R}, \quad x \mapsto F_k(x) := f(x, y_k),$$

für $k \rightarrow \infty$ gleichmäßig gegen die Funktion

$$F : [a, b] \longrightarrow \mathbb{R}, \quad x \mapsto F(x) := f(x, c).$$

Beweis. Nach § 3, Satz 1, ist die Menge

$$Q := \{y_k : k \in \mathbb{N}\} \cup \{c\}$$

kompakt, also abgeschlossen und beschränkt. Daher ist auch $[a, b] \times Q \subset \mathbb{R}^{1+m}$ abgeschlossen und beschränkt (vgl. (1.14)), also kompakt. Aus § 3, Satz 10, folgt, dass die Beschränkung

$$f \Big| [a, b] \times Q \longrightarrow \mathbb{R}$$

gleichmäßig stetig ist. Sei nun $\varepsilon > 0$ vorgegeben. Dann existiert ein $\delta > 0$, so dass für alle $(x, y), (x', y') \in [a, b] \times Q$ gilt

$$\|(x, y) - (x', y')\| < \delta \implies |f(x, y) - f(x', y')| < \varepsilon.$$

Da $\lim y_k = c$, existiert ein $N \in \mathbb{N}$ mit

$$\|c - y_k\| < \delta \quad \text{für alle } k \geq N.$$

Daher gilt für $k \geq N$

$$|f(x, c) - f(x, y_k)| < \varepsilon \quad \text{für alle } x \in [a, b].$$

Das bedeutet aber

$$|F(x) - F_k(x)| < \varepsilon \quad \text{für alle } x \in [a, b] \text{ und } k \geq N,$$

d.h. die Folge $(F_k)_{k \in \mathbb{N}}$ konvergiert gleichmäßig gegen F , q.e.d.

Satz 1 (stetige Abhängigkeit vom Parameter). *Sei $[a, b] \subset \mathbb{R}$ ein kompaktes Intervall, $U \subset \mathbb{R}^m$ eine beliebige Teilmenge und*

$$f : [a, b] \times U \longrightarrow \mathbb{R}$$

eine stetige Funktion. Für $y \in U$ werde gesetzt

$$\varphi(y) := \int_a^b f(x, y) dx$$

Dann ist die dadurch definierte Funktion $\varphi : U \rightarrow \mathbb{R}$ stetig.

Beweis. Sei $c \in U$ ein beliebiger Punkt und $y_k \in U$, $k \in \mathbb{N}$, eine Punktfolge mit $\lim y_k = c$. Dann gilt mit den Bezeichnungen des obigen Hilfssatzes

$$\varphi(y_k) = \int_a^b F_k(x) dx \quad \text{und} \quad \varphi(c) = \int_a^b F(x) dx.$$

Da die Folge $(F_k)_{k \in \mathbb{N}}$ auf $[a, b]$ gleichmäßig gegen F konvergiert, kann man nach An. 1, § 21, Satz 4, Integration und Limesbildung vertauschen, d.h. es gilt

$$\lim_{k \rightarrow \infty} \int_a^b F_k(x) dx = \int_a^b F(x) dx,$$

also $\lim_{k \rightarrow \infty} \varphi(y_k) = \varphi(c)$, q.e.d.

Hilfssatz 2. Seien $I, J \subset \mathbb{R}$ kompakte Intervalle und

$$f : I \times J \longrightarrow \mathbb{R}, \quad (x, y) \mapsto f(x, y),$$

eine stetige Funktion, die nach der zweiten Variablen stetig partiell differenzierbar sei. Weiter sei $c \in J$ ein Punkt und $y_k \in J$, $k \in \mathbb{N}$, eine Punktfolge mit

$$\lim_{k \rightarrow \infty} y_k = c \quad \text{und} \quad y_k \neq c \quad \text{für alle } k.$$

Wir definieren Funktionen $F_k, F : I \rightarrow \mathbb{R}$ durch

$$F_k(x) := \frac{f(x, y_k) - f(x, c)}{y_k - c}, \quad (k \in \mathbb{N}),$$

$$F(x) := \frac{\partial f}{\partial y}(x, c).$$

Dann konvergiert die Folge $(F_k)_{k \in \mathbb{N}}$ auf I gleichmäßig gegen F .

Beweis. Sei $\varepsilon > 0$ vorgegeben.

Die stetige Funktion $D_2 f : I \times J \rightarrow \mathbb{R}$ ist nach § 3, Satz 10, gleichmäßig stetig, es gibt also ein von $x \in I$ unabhängiges $\delta > 0$, so dass aus $y, y' \in J$, $|y - y'| < \delta$ folgt

$$|D_2 f(x, y) - D_2 f(x, y')| < \varepsilon.$$

Nach dem Mittelwertsatz der Differentialrechnung für Funktionen einer Veränderlichen gibt es zu jedem $x \in I$ und $k \in \mathbb{N}$ ein η_k zwischen c und y_k mit

$$F_k(x) = D_2 f(x, \eta_k).$$

Sei N so groß, dass $|c - y_k| < \delta$ für alle $k \geq N$. Dann gilt auch $|c - \eta_k| < \delta$ für $k \geq N$ und es folgt

$$|F(x) - F_k(x)| = |D_2 f(x, c) - D_2 f(x, \eta_k)| < \varepsilon$$

für alle $x \in I$ und $k \geq N$, q.e.d.

Satz 2 (differenzierbare Abhängigkeit vom Parameter). Seien $I, J \subset \mathbb{R}$ kompakte Intervalle und

$$f : I \times J \longrightarrow \mathbb{R}, \quad (x, y) \mapsto f(x, y),$$

eine stetige Funktion, die nach der Variablen y stetig differenzierbar ist. Für

$y \in J$ werde gesetzt

$$\varphi(y) := \int_I f(x, y) dx.$$

Dann ist die Funktion $\varphi: J \rightarrow \mathbb{R}$ stetig differenzierbar und es gilt

$$\frac{d\varphi(y)}{dy} = \int_I \frac{\partial f(x, y)}{\partial y} dx.$$

Man drückt dies auch so aus: Man darf “unter dem Integral differenzieren”.

Beweis. Sei $c \in J$ und $y_k \in J$, $k \in \mathbb{N}$, eine beliebige Punktfolge mit

$$\lim_{k \rightarrow \infty} y_k = c \quad \text{und} \quad y_k \neq c \quad \text{für alle } k.$$

Die Funktionen F_k und F seien wie in Hilfssatz 2 definiert. Wegen der gleichmäßigen Konvergenz von F_k gegen F gilt nach An. 1, § 21, Satz 4,

$$\lim_{k \rightarrow \infty} \frac{\varphi(y_k) - \varphi(c)}{y_k - c} = \lim_{k \rightarrow \infty} \int_I \frac{f(x, y_k) - f(x, c)}{y_k - c} dx = \int_I \frac{\partial f}{\partial y}(x, c) dx.$$

Daher existiert $\varphi'(c)$ für jedes $c \in J$ und es gilt

$$\varphi'(c) = \int_I \frac{\partial f}{\partial y}(x, c) dx.$$

Die Funktion φ' ist nach Satz 1 stetig, da nach Voraussetzung die Funktion $\partial f / \partial y$ auf $I \times J$ stetig ist.

(10.1) Beispiel. Wir wollen das Integral

$$\int_0^a x^2 \cos x dx$$

berechnen. Man könnte das Integral mittels zweimaliger partieller Integration auswerten (was der Leserin zur Übung empfohlen sei!), es gibt aber auch eine andere Möglichkeit, die auf der Berechnung eines parameterabhängigen Integrals beruht. Dazu betrachten wir die Funktion

$$F(y) := \int_0^a \cos(xy) dx$$

etwa auf dem Intervall $\frac{1}{2} \leq y \leq 2$. Nach Satz 2 gilt

$$F'(y) = \int_0^a \frac{\partial}{\partial y} \cos(xy) dx = - \int_0^a x \sin(xy) dx,$$

$$F''(y) = - \int_0^a x^2 \cos(xy) dx.$$

Das gesuchte Integral ergibt sich also als

$$\int_0^a x^2 \cos x dx = -F''(1).$$

Andererseits kann man die Funktion F direkt ausrechnen:

$$F(y) = \int_0^a \cos(xy) dx = \frac{\sin(xy)}{y} \Big|_{x=0}^{x=a} = \frac{\sin(ay)}{y}.$$

Differenzieren ergibt

$$F'(y) = -\frac{\sin(ay)}{y^2} + \frac{a \cos(ay)}{y},$$

$$F''(y) = \frac{2 \sin(ay)}{y^3} - \frac{2a \cos(ay)}{y^2} - \frac{a^2 \sin(ay)}{y},$$

also

$$\int_0^a x^2 \cos x dx = -F''(1) = (a^2 - 2) \sin a + 2a \cos a.$$

Bemerkung. Aus Satz 2 folgt, da die partiellen Ableitungen nichts anderes sind als gewöhnliche Ableitungen bei festgehaltenen übrigen Variablen:

Sei $[a, b] \subset \mathbb{R}$ ein kompaktes Intervall, $U \subset \mathbb{R}^n$ offen und

$$f: [a, b] \times U \longrightarrow \mathbb{R}, \quad (x, y_1, \dots, y_n) \mapsto f(x, y_1, \dots, y_n),$$

eine stetige Funktion, die bzgl. der Variablen y_1, \dots, y_n stetig partiell differenzierbar ist. Dann gilt für jedes $i \in \{1, \dots, n\}$

$$\frac{\partial}{\partial y_i} \int_a^b f(x, y_1, \dots, y_n) dx = \int_a^b \frac{\partial f}{\partial y_i}(x, y_1, \dots, y_n) dx.$$

(10.2) Beispiel. Sei $U = \{x \in \mathbb{R}^n : \|x\| < r\}$ die offene Kugel vom Radius $r > 0$ um den Nullpunkt des \mathbb{R}^n und

$$v = (v_1, \dots, v_n) : U \rightarrow \mathbb{R}^n$$

ein stetig differenzierbares Vektorfeld in U . Wir suchen eine stetig differenzierbare Funktion

$$f: U \rightarrow \mathbb{R} \quad \text{mit} \quad \text{grad } f = v,$$

d.h.

$$\frac{\partial f}{\partial x_i} = v_i \quad \text{für } i = 1, \dots, n.$$

Falls eine solche Funktion f existiert, ist sie zweimal stetig differenzierbar und nach dem Satz von Schwarz (§5, Satz 1) gilt

$$\frac{\partial^2 f}{\partial x_j \partial x_i} = \frac{\partial^2 f}{\partial x_i \partial x_j} \implies \frac{\partial v_i}{\partial x_j} = \frac{\partial v_j}{\partial x_i} \quad \text{für alle } i \neq j.$$

Eine notwendige Bedingung für die Lösbarkeit der Gleichung $\text{grad } f = v$ ist also $\partial v_i / \partial x_j = \partial v_j / \partial x_i$ für alle $i \neq j$. Wir zeigen nun, dass diese Bedingung auch hinreichend ist. Dazu definieren wir die Funktion $f: U \rightarrow \mathbb{R}$ durch

$$f(x) := \sum_{i=1}^n \left(\int_0^1 v_i(tx) dt \right) x_i.$$

Behauptung. Es gilt $\frac{\partial f}{\partial x_j} = v_j$ für $j = 1, \dots, n$.

Beweis. Nach der Produktregel und Satz 2 ist

$$\begin{aligned} \frac{\partial f(x)}{\partial x_j} &= \sum_i \left(\frac{\partial}{\partial x_j} \int_0^1 v_i(tx) dt \right) x_i + \sum_i \left(\int_0^1 v_i(tx) dt \right) \frac{\partial x_i}{\partial x_j} \\ &= \sum_i \left(\int_0^1 t D_j v_i(tx) dt \right) x_i + \int_0^1 v_j(tx) dt \\ &= \int_0^1 \left(t \sum_i D_j v_i(tx) x_i + v_j(tx) \right) dt. \end{aligned}$$

Nun ist (bei festem $x \in U$):

$$\begin{aligned} \frac{d}{dt}(tv_j(tx)) &= v_j(tx) + t \frac{d}{dt} v_j(tx) \\ &= v_j(tx) + t \sum_i D_i v_j(tx) x_i \\ &= v_j(tx) + t \sum_i D_j v_i(tx) x_i, \end{aligned}$$

also

$$\frac{\partial f(x)}{\partial x_j} = \int_0^1 \frac{d}{dt}(tv_j(tx))dt = tv_j(tx) \Big|_{t=0}^{t=1} = v_j(x), \quad \text{q.e.d.}$$

Damit ist insbesondere gezeigt: Dafür, dass ein auf einer offenen Kugel $U \subset \mathbb{R}^3$ stetig differenzierbares Vektorfeld $v: U \rightarrow \mathbb{R}^3$ sich als Gradient einer stetig differenzierbaren Funktion $f: U \rightarrow \mathbb{R}$ darstellen lässt, ist notwendig und hinreichend, dass $\operatorname{rot} v = 0$.

Doppelintegrale

Seien $[a, b] \subset \mathbb{R}$ und $[c, d] \subset \mathbb{R}$ kompakte Intervalle und

$$f: [a, b] \times [c, d] \longrightarrow \mathbb{R}$$

eine stetige Funktion. Nach Satz 1 ist die Funktion

$$F(y) := \int_a^b f(x, y) dx$$

stetig auf $[a, b]$, kann also wieder integriert werden. Man bezeichnet

$$\int_c^d F(y) dy = \int_c^d \left(\int_a^b f(x, y) dx \right) dy$$

als Doppelintegral.

Man kann die Funktion f jedoch auch erst bzgl. y und dann bzgl. x integrieren. Der folgende Satz sagt, dass sich dabei derselbe Wert ergibt.

Satz 3. *Mit den obigen Bezeichnungen gilt*

$$\int_c^d \left(\int_a^b f(x, y) dx \right) dy = \int_a^b \left(\int_c^d f(x, y) dy \right) dx.$$

Beweis. Wir definieren eine Funktion $\varphi: [c, d] \rightarrow \mathbb{R}$ durch

$$\varphi(y) := \int_a^b \left(\int_c^y f(x, t) dt \right) dx.$$

Es gilt $\varphi(c) = 0$ und nach Satz 2 ist φ differenzierbar mit

$$\varphi'(y) = \int_a^b \frac{\partial}{\partial y} \left(\int_c^y f(x, t) dt \right) dx = \int_a^b f(x, y) dx.$$

Daraus folgt

$$\int_c^d \left(\int_a^b f(x, y) dx \right) dy = \int_c^d \varphi'(y) dy = \varphi(d) = \int_a^b \left(\int_c^d f(x, y) dy \right) dx,$$

was zu beweisen war.

Bemerkung. Ein analoger Satz gilt natürlich auch für n -fache Integrale stetiger Funktionen, die auf einem Quader $I_1 \times \dots \times I_n \subset \mathbb{R}^n$ definiert sind.

Geometrische Interpretation. Anschaulich bedeutet das Doppelintegral

$$V := \int_c^d \int_a^b f(x, y) dx dy$$

über eine nicht-negative Funktion $f: [a, b] \times [c, d] \rightarrow \mathbb{R}$ das Volumen des Körpers unter dem Graphen von f , d.h. der Menge

$$K := \{(x, y, z) \in [a, b] \times [c, d] \times \mathbb{R} : 0 \leq z \leq f(x, y)\}.$$

Die systematische Theorie der Volumenmessung wird erst im 3. Band der Analysis [6] dargestellt. Wir bringen hier nur ein einfaches Beispiel.

(10.3) Wir wollen das Volumen der Kugel vom Radius $r > 0$ im \mathbb{R}^3 berechnen. Dazu betrachten wir die folgende stetige Funktion

$$f: [-r, r] \times [-r, r] \rightarrow \mathbb{R}, \quad f(x, y) := \begin{cases} \sqrt{r^2 - x^2 - y^2} & \text{für } x^2 + y^2 \leq r^2, \\ 0 & \text{sonst.} \end{cases}$$

Der Körper unter dem Graphen von f ist dann (bis auf einen Teil der Höhe 0) die obere Hälfte der Kugel

$$K(r) = \{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 + z^2 \leq r^2\}.$$

Daher ist

$$V := \int_{-r}^r \int_{-r}^r f(x, y) dx dy$$

gleich der Hälfte des Volumens von $K(r)$.

Für festes $y \in [-r, r]$ sei $\rho = \rho(y) := \sqrt{r^2 - y^2}$. Damit ist

$$\begin{aligned} F(y) &= \int_{-r}^r f(x, y) dx = \int_{-\rho}^{\rho} \sqrt{\rho^2 - x^2} dx = \quad [\text{Subst. } x = \rho \sin t] \\ &= \rho^2 \int_{-\pi/2}^{\pi/2} \cos^2 t dt = \rho^2 \frac{\pi}{2} = (r^2 - y^2) \frac{\pi}{2} \end{aligned}$$

und

$$V = \int_{-r}^r F(y) dy = \frac{\pi}{2} \int_{-r}^r (r^2 - y^2) dy = \frac{\pi}{2} (2r^3 - 2r^3/3) = \frac{2}{3} \pi r^3.$$

Es folgt also für das Volumen der dreidimensionalen Kugel vom Radius r die wohlbekannte Formel

$$\text{Vol}(K(r)) = 2V = \frac{4\pi}{3} r^3.$$

Eulersche Differentialgleichungen der Variationsrechnung

Wir betrachten folgende Situation:

Sei $I := [a, b] \subset \mathbb{R}$ ein kompaktes Intervall und

$$L : I \times \mathbb{R} \times \mathbb{R} \longrightarrow \mathbb{R}, \quad (t, y, p) \mapsto L(t, y, p)$$

eine zweimal stetig partiell differenzierbare Funktion. Wir bezeichnen mit $\mathcal{C}^2[a, b]$ den Vektorraum aller zweimal stetig differenzierbaren Funktionen $f: [a, b] \rightarrow \mathbb{R}$ und mit \mathcal{K} die Teilmenge

$$\mathcal{K} := \{\varphi \in \mathcal{C}^2[a, b] : \varphi(a) = c_1, \varphi(b) = c_2\},$$

wobei $c_1, c_2 \in \mathbb{R}$ vorgegebene Konstanten sind. Wir definieren jetzt ein Funktional (d.h. eine Abbildung)

$$S : \mathcal{K} \longrightarrow \mathbb{R}, \quad \varphi \mapsto S(\varphi) := \int_a^b L(t, \varphi(t), \varphi'(t)) dt.$$

Das Problem der Variationsrechnung besteht nun darin, $S(\varphi)$ zu minimieren, d.h. ein $\varphi \in \mathcal{K}$ zu finden, so dass

$$S(\varphi) = \inf\{S(\psi) : \psi \in \mathcal{K}\}.$$

Satz 4. *Mit den obigen Bezeichnungen gilt: Eine notwendige Bedingung dafür, dass $S(\varphi) = \inf_{\psi \in \mathcal{K}} S(\psi)$ ist das Bestehen der Eulerschen Differentialgleichung*

$$\frac{d}{dt} \frac{\partial L}{\partial p}(t, \varphi(t), \varphi'(t)) - \frac{\partial L}{\partial y}(t, \varphi(t), \varphi'(t)) = 0.$$

Beweis. Sei $\varphi \in \mathcal{K}$ eine Funktion mit

$$S(\varphi) \leq S(\psi) \quad \text{für alle } \psi \in \mathcal{K}$$

und $g \in C^2[a, b]$ eine beliebige Funktion mit

$$g(a) = g(b) = 0.$$

Für jedes $\varepsilon \in \mathbb{R}$ ist dann $\varphi + \varepsilon g \in \mathcal{K}$ und damit

$$S(\varphi) \leq S(\varphi + \varepsilon g).$$

Wir definieren jetzt die Funktion $F: \mathbb{R} \rightarrow \mathbb{R}$ durch

$$F(\varepsilon) := S(\varphi + \varepsilon g).$$

Diese Funktion F besitzt für $\varepsilon = 0$ ein Minimum, also gilt

$$\frac{dF}{d\varepsilon}(0) = 0.$$

Mit Satz 2 erhält man

$$\begin{aligned} \frac{dF}{d\varepsilon}(\varepsilon) &= \int_a^b \frac{\partial}{\partial \varepsilon} L(t, \varphi(t) + \varepsilon g(t), \varphi'(t) + \varepsilon g'(t)) dt \\ &= \int_a^b \left(\frac{\partial L}{\partial y}(\dots) g(t) + \frac{\partial L}{\partial p}(\dots) g'(t) \right) dt. \end{aligned}$$

Dabei stehen die Pünktchen (\dots) für $(t, \varphi(t) + \varepsilon g(t), \varphi'(t) + \varepsilon g'(t))$. Partielle Integration liefert

$$\int_a^b \frac{\partial L}{\partial p} g'(t) dt = \frac{\partial L}{\partial p} \cdot g(t) \Big|_a^b - \int_a^b g(t) \frac{d}{dt} \left(\frac{\partial L}{\partial p} \right) dt = - \int_a^b g(t) \frac{d}{dt} \left(\frac{\partial L}{\partial p} \right) dt.$$

Damit erhält man schließlich

$$\begin{aligned} 0 = \frac{dF}{d\varepsilon}(0) &= \int_a^b \left(\frac{\partial L}{\partial y}(t, \varphi, \varphi') g(t) + \frac{\partial L}{\partial p}(t, \varphi, \varphi') g'(t) \right) dt \\ &= \int_a^b \left(\frac{\partial L}{\partial y}(t, \varphi, \varphi') - \frac{d}{dt} \frac{\partial L}{\partial p}(t, \varphi, \varphi') \right) g(t) dt. \end{aligned}$$

(Zur Vereinfachung haben wir nur φ, φ' statt $\varphi(t), \varphi'(t)$ geschrieben.) Mit dem

anschließend bewiesenen Hilfssatz 3 folgt daraus

$$\frac{\partial L}{\partial y}(t, \varphi, \varphi') - \frac{d}{dt} \frac{\partial L}{\partial p}(t, \varphi, \varphi') = 0, \quad \text{q.e.d.}$$

Hilfssatz 3. Sei $a, b \in \mathbb{R}$, $a < b$, und

$$f : [a, b] \longrightarrow \mathbb{R}$$

eine stetige Funktion. Für jede zweimal stetig differenzierbare Funktion

$$g : [a, b] \longrightarrow \mathbb{R} \quad \text{mit } g(a) = g(b) = 0$$

gelte

$$\int_a^b f(t)g(t) dt = 0.$$

Dann gilt $f(t) = 0$ für alle $t \in [a, b]$.

Beweis. Wegen der Stetigkeit von f genügt es zu zeigen, dass f auf dem offenen Intervall $]a, b[$ identisch verschwindet.

Angenommen, es sei $f(t) \neq 0$ für ein $t \in]a, b[$. Ohne Beschränkung der Allgemeinheit ist $f(t) =: \varepsilon > 0$. Dann existiert ein $\delta > 0$, so dass

$$[t - \delta, t + \delta] \subset]a, b[$$

und

$$f(x) \geq \frac{\varepsilon}{2} \quad \text{für alle } x \in [t - \delta, t + \delta].$$

Man kann nun eine zweimal stetig differenzierbare nicht-negative Funktion $g : [a, b] \rightarrow \mathbb{R}$ konstruieren mit

$$g(t) > 0 \quad \text{und} \quad g(x) = 0 \quad \text{für alle } x \notin [t - \delta, t + \delta],$$

vgl. Bild 10.1

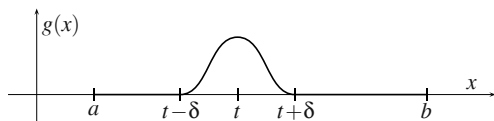


Bild 10.1

Nun folgt

$$0 = \int_a^b f(x)g(x) dx = \int_{t-\delta}^{t+\delta} f(x)g(x) dx \geq \frac{\varepsilon}{2} \int_{t-\delta}^{t+\delta} g(x) dx > 0,$$

Widerspruch! Also ist doch $f(t) = 0$ für alle $t \in]a, b[$.

(10.4) Beispiel. Seien $a, b, c_1, c_2 \in \mathbb{R}$, $a < b$, und

$$\mathcal{K} = \{\varphi \in C^2[a, b] : \varphi(a) = c_1, \varphi(b) = c_2\}.$$

Mit $S(\varphi)$ werde die Bogenlänge der Kurve

$$[a, b] \longrightarrow \mathbb{R}, \quad t \mapsto (t, \varphi(t)),$$

d.h. des Graphen von φ , bezeichnet, siehe Bild 10.2.

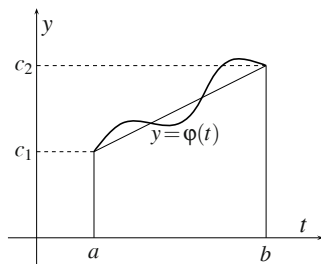


Bild 10.2

Nach § 4, Satz 1, gilt

$$S(\varphi) = \int_a^b \sqrt{1 + \varphi'(t)^2} dt.$$

Wir wollen $S(\varphi)$ minimieren, suchen also die kürzeste Verbindungslinie zwischen den Punkten (a, c_1) und (b, c_2) . Man kann Satz 4 anwenden, wobei hier die Funktion L eine besonders einfache Gestalt hat:

$$L(t, y, p) = \sqrt{1 + p^2}.$$

L hängt also gar nicht explizit von t und y ab, es gilt

$$\frac{\partial L}{\partial y} = 0, \quad \text{und} \quad \frac{\partial L}{\partial p}(t, y, p) = \frac{p}{\sqrt{1 + p^2}}.$$

Die Eulersche Differentialgleichung lautet daher

$$\frac{d}{dt} \frac{\varphi'(t)}{\sqrt{1 + \varphi'(t)^2}} = 0,$$

d.h.

$$\frac{\varphi'(t)}{\sqrt{1 + \varphi'(t)^2}} = \text{const.}$$

Dies ist aber gleichbedeutend mit $\varphi'(t) = \text{const.}$ Die Funktion φ muss also ein Polynom 1. Grades sein:

$$\varphi(t) = \alpha + \beta t,$$

wobei die Konstanten $\alpha, \beta \in \mathbb{R}$ noch so zu bestimmen sind, dass die Randbedingungen $\varphi(a) = c_1$ und $\varphi(b) = c_2$ erfüllt sind. Das bedeutet: Wenn unser Problem überhaupt eine Lösung besitzt, wird es durch die Verbindungsgerade der Punkte (a, c_1) und (b, c_2) gelöst. Dass die geradlinige Verbindung tatsächlich minimale Bogenlänge aufweist, kann hier durch einfache geometrische Betrachtungen gezeigt werden. Im Allgemeinen ist es bei Variationsproblemen aber schwierig zu zeigen, dass das Minimum tatsächlich angenommen wird.

Bemerkung. Satz 4 lässt sich leicht wie folgt auf höhere Dimensionen verallgemeinern: Gegeben sei eine zweimal stetig differenzierbare Funktion

$$L : [a, b] \times \mathbb{R}^n \times \mathbb{R}^n \longrightarrow \mathbb{R},$$

$$(t, y_1, \dots, y_n, p_1, \dots, p_n) \mapsto L(t, y_1, \dots, y_n, p_1, \dots, p_n).$$

Mit \mathcal{K} sei die Menge aller zweimal stetig differenzierbaren vektorwertigen Funktionen

$$\varphi = (\varphi_1, \dots, \varphi_n) : [a, b] \longrightarrow \mathbb{R}^n$$

mit $\varphi(a) = C_1$ und $\varphi(b) = C_2$ bezeichnet, wobei $C_1, C_2 \in \mathbb{R}^n$ vorgegebene Vektoren sind. Ein reelles Funktional $S : \mathcal{K} \rightarrow \mathbb{R}$ werde definiert durch

$$S(\varphi) := \int_a^b L(t, \varphi_1(t), \dots, \varphi_n(t), \varphi'_1(t), \dots, \varphi'_n(t)) dt.$$

Ist dann $\varphi \in \mathcal{K}$ eine Funktion mit

$$S(\varphi) = \inf\{S(\psi) : \psi \in \mathcal{K}\},$$

so gelten die Eulerschen Differentialgleichungen

$$\frac{d}{dt} \frac{\partial L}{\partial p_i}(t, \varphi(t), \varphi'(t)) - \frac{\partial L}{\partial y_i}(t, \varphi(t), \varphi'(t)) = 0, \quad (i = 1, \dots, n).$$

(10.5) Anwendung in der Physik. Die Eulerschen Differentialgleichungen werden angewendet beim sog. *Hamiltonschen Prinzip der kleinsten Wirkung*.

Hier ist t die Zeitkoordinate, die Funktionen $\varphi_1, \dots, \varphi_n$ beschreiben den Zustand eines physikalischen Systems als Funktion der Zeit. $L(t, \varphi, \varphi')$ ist die sog. *Lagrange-Funktion*; bei mechanischen Systemen, in denen die Reibung keine Rolle spielt, ist L gleich der Differenz $T - U$ aus kinetischer Energie T und potentieller Energie U .

$$S(\varphi) = \int_a^b L(t, \varphi(t), \varphi'(t)) dt$$

ist das sog. Wirkungsintegral; das Hamiltonsche Prinzip sagt aus, dass für den tatsächlich ablaufenden Vorgang $S(\varphi)$ minimal wird.

Betrachten wir etwa die Bewegung eines Massenpunktes im \mathbb{R}^3 unter dem Einfluss eines nur vom Ort abhängigen Potentials

$$U : \mathbb{R}^3 \longrightarrow \mathbb{R}, \quad (x_1, x_2, x_3) \mapsto U(x_1, x_2, x_3).$$

Die Bewegung des Massenpunktes wird beschrieben durch eine vektorwertige Funktion

$$x(t) = (x_1(t), x_2(t), x_3(t)),$$

seine Geschwindigkeit zur Zeit t ist

$$v(t) = \frac{dx(t)}{dt} = (\dot{x}_1(t), \dot{x}_2(t), \dot{x}_3(t)).$$

(In der Physik ist es seit Newton üblich, die Ableitung nach der Zeit durch einen Punkt zu bezeichnen.)

Die kinetische Energie des Massenpunktes ist

$$T = \frac{1}{2} m \|v\|^2 = \frac{m}{2} \sum_{i=1}^3 v_i^2 = \frac{m}{2} \sum_{i=1}^3 \dot{x}_i^2,$$

wobei m die Masse bezeichnet. Für die Lagrange-Funktion $L = T - U$ ergibt

sich also

$$L(x, v) = \frac{m}{2}(v_1^2 + v_2^2 + v_3^2) - U(x_1, x_2, x_3),$$

sie hängt nicht explizit von der Zeit ab. (Hier sind x und v die Variablen, die wir früher mit y und p bezeichnet haben.) Man hat

$$\frac{\partial L}{\partial x_i} = -\frac{\partial U}{\partial x_i} \quad \text{und} \quad \frac{\partial L}{\partial v_i} = mv_i.$$

Die Eulerschen Differentialgleichungen lauten daher

$$\frac{d}{dt}(m\dot{x}_i(t)) + \frac{\partial U}{\partial x_i}(x(t)) = 0,$$

oder (falls die Masse m zeitlich konstant ist)

$$m\ddot{x}_i(t) = -\frac{\partial U}{\partial x_i}(x(t)), \quad (i = 1, 2, 3).$$

Man kann dies in vektorieller Form zusammenfassen zu

$$m\ddot{x} = -\text{grad} U(x) \quad (*)$$

Wir werden später (in § 16) sehen, wie man für eine spezielle Potentialfunktion U ein solches Differentialgleichungssystem lösen kann.

Wir wollen hier nur noch zeigen, wie aus der Gleichung $(*)$ die Konstanz der Gesamtenergie $E = T + U$ folgt. Zur Zeit t beträgt die Gesamtenergie des Massenpunkts

$$E(t) = \frac{m}{2} \sum_{i=1}^3 \dot{x}_i(t)^2 + U(x_1(t), x_2(t), x_3(t)).$$

Ableitung nach der Zeit ergibt mit der Kettenregel

$$\begin{aligned} \frac{dE(t)}{dt} &= m \sum_{i=1}^3 \dot{x}_i(t) \ddot{x}_i(t) + \sum_{i=1}^3 \frac{\partial U}{\partial x_i}(x(t)) \dot{x}_i(t) \\ &= m \langle \ddot{x}(t), \dot{x}(t) \rangle + \langle \text{grad} U(x(t)), \dot{x}(t) \rangle \\ &= \langle m\ddot{x}(t) + \text{grad} U(x(t)), \dot{x}(t) \rangle. \end{aligned}$$

Wegen $(*)$ ist dies gleich Null, d.h. E ist konstant.

AUFGABEN

10.1. Man berechne das Integral

$$\int_0^x t^n e^{-t} dt$$

durch Differenzieren des Parameter-abhängigen Integrals

$$F(y) := \int_0^x e^{-ty} dt.$$

10.2. Sei $I \subset \mathbb{R}$ ein offenes Intervall, $a \in I$ und

$$f : I \times I \rightarrow \mathbb{R}, \quad (x, y) \mapsto f(x, y)$$

eine stetige, nach der zweiten Variablen stetig partiell differenzierbare Funktion. Man zeige, dass die durch

$$F(y) := \int_a^y f(x, y) dx$$

definierte Funktion $F : I \rightarrow \mathbb{R}$ differenzierbar ist, und dass für alle $y \in I$ gilt

$$F'(y) = f(y, y) + \int_a^y \frac{\partial f}{\partial y}(x, y) dx.$$

Anleitung: Man beweise, dass die durch

$$G(y, z) := \int_a^z f(x, y) dx$$

definierte Funktion $G : I \times I \rightarrow \mathbb{R}$ stetig partiell differenzierbar ist und wende die Kettenregel an.

10.3. Sei $g : \mathbb{R}^2 \rightarrow \mathbb{R}$ die Funktion

$$g(x, y) = \begin{cases} \frac{xy^3}{(x^2 + y^2)^2} & \text{für } (x, y) \neq (0, 0) \\ 0 & \text{für } (x, y) = (0, 0) \end{cases}$$

Man zeige, dass für jedes $y \in \mathbb{R}$ die Integrale

$$f(y) := \int_0^1 g(x, y) dx \quad \text{und} \quad f^*(y) := \int_0^1 \frac{\partial g}{\partial y}(x, y) dx$$

wohldefiniert sind und dass die Funktion $f : \mathbb{R} \rightarrow \mathbb{R}$ differenzierbar ist, jedoch $f'(0) \neq f^*(0)$ gilt.

10.4. Es sei $f: [0, 1] \times [-1, 1] \rightarrow \mathbb{R}$ die wie folgt definierte Funktion:

$$f(x, y) := \begin{cases} \sqrt{x^2 - y^2} & \text{für } |y| \leq x, \\ 0 & \text{sonst.} \end{cases}$$

Man berechne das Doppelintegral

$$V := \int_{-1}^1 \int_0^1 f(x, y) dx dy.$$

10.5. Sei $r > 0$ und $f: [-r, r]^3 \rightarrow \mathbb{R}$ die wie folgt definierte Funktion dreier Variablen:

$$f(x_1, x_2, x_3) := \begin{cases} \sqrt{r^2 - x_1^2 - x_2^2 - x_3^2} & \text{falls } x_1^2 + x_2^2 + x_3^2 \leq r^2, \\ 0 & \text{sonst.} \end{cases}$$

Man berechne das dreifache Integral

$$V := \int_{-r}^r \int_{-r}^r \int_{-r}^r f(x_1, x_2, x_3) dx_1 dx_2 dx_3.$$

Bemerkung. $2V$ ist das Volumen der 4-dimensionalen Kugel vom Radius r .

Kapitel II

Gewöhnliche Differentialgleichungen

§ 11 Elementare Lösungsmethoden

Eine Differentialgleichung (erster Ordnung) ist eine Bedingungsgleichung für eine zu bestimmende Funktion, in der die Ableitung als Funktion des Arguments und des Wertes der Funktion dargestellt wird. Geometrisch bedeutet das die Vorgabe eines Richtungsfelds; es wird dann eine Funktion gesucht, deren Graph sich an dieses Richtungsfeld anschmiegt. In diesem Paragraphen behandeln wir einige einfache Beispiele, in denen man die Lösungen einer Differentialgleichung explizit bestimmen kann.

Definition. Sei G eine Teilmenge des \mathbb{R}^2 und

$$f : G \longrightarrow \mathbb{R}, \quad (x, y) \mapsto f(x, y)$$

eine stetige Funktion. Dann nennt man

$$y' = f(x, y) \tag{1}$$

eine (*gewöhnliche*) *Differentialgleichung erster Ordnung*. Unter einer *Lösung* von (1) versteht man eine auf einem Intervall $I \subset \mathbb{R}$ definierte differenzierbare Funktion

$$\varphi : I \longrightarrow \mathbb{R}$$

mit folgenden Eigenschaften:

a) Der Graph von φ ist in G enthalten, d.h.

$$\Gamma_\varphi := \{(x, y) \in I \times \mathbb{R} : y = \varphi(x)\} \subset G.$$

b) Es gilt

$$\varphi'(x) = f(x, \varphi(x)) \quad \text{für alle } x \in I.$$

(Die Bedingung a) ist zu stellen, damit man b) überhaupt formulieren kann.)

Bemerkung. Würde die gegebene Funktion f nur von x und nicht von y abhängen, so schriebe sich b) als

$$\varphi'(x) = f(x)$$

und man könnte die Lösungen durch einfache Integration gewinnen:

$$\varphi(x) = c + \int_{x_0}^x f(t) dt.$$

Im allgemeinen Fall folgt aus b) zwar

$$\varphi(x) = c + \int_{x_0}^x f(t, \varphi(t)) dt,$$

wobei $c = \varphi(x_0)$, aber man kann mit dieser Formel nicht unmittelbar die Lösung gewinnen, da unter dem Integral bereits die gesuchte Funktion vorkommt. (Wir werden aber im nächsten Paragraphen sehen, wie man trotzdem mittels dieser Integralgleichung die Differentialgleichung lösen kann.)

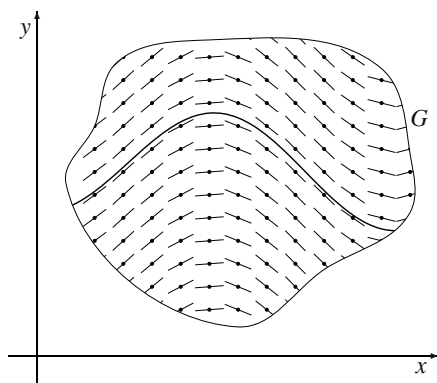


Bild 11.1

Geometrische Interpretation

Eine Differentialgleichung $y' = f(x, y)$ in einem Gebiet $G \subset \mathbb{R}^2$ bestimmt ein *Richtungsfeld*: In jedem Punkt $(x, y) \in G$ wird durch $y' = f(x, y)$ eine Steigung

vorgegeben, vgl. Bild 11.1. Gesucht ist eine differenzierbare Funktion, deren Graph in jedem seiner Punkte die vorgegebene Steigung hat.

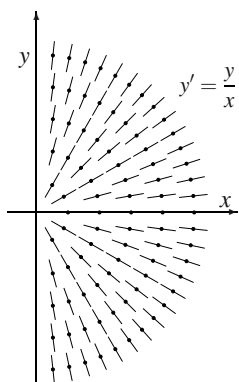


Bild 11.2

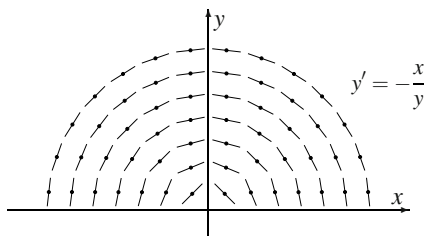


Bild 11.3

In einfachen Fällen kann man aus dem Richtungsfeld bereits die Lösungen der Differentialgleichung ansehen. So besitzt die Differentialgleichung

$$y' = \frac{y}{x} \quad \text{in } \mathbb{R}_+^* \times \mathbb{R}, \quad (\text{Bild 11.2})$$

die Geraden $y = cx$ als Lösungen und die Differentialgleichung

$$y' = -\frac{x}{y} \quad \text{in } \mathbb{R} \times \mathbb{R}_+^*, \quad (\text{Bild 11.3})$$

die Halbkreise $y = \sqrt{c - x^2}$, $|x| < \sqrt{c}$.

Differentialgleichungen mit getrennten Variablen

Seien $I, J \subset \mathbb{R}$ offene Intervalle und

$$f: I \longrightarrow \mathbb{R}, \quad g: J \longrightarrow \mathbb{R}$$

zwei stetige Funktionen. Es werde vorausgesetzt, dass $g(y) \neq 0$ für alle $y \in J$. Die Differentialgleichung

$$y' = f(x)g(y) \quad \text{in } I \times J$$

heißt Differentialgleichung mit getrennten Variablen.

Satz 1. Die obigen Bezeichnungen seien beibehalten. Sei $(x_0, y_0) \in I \times J$ ein Punkt. Wir definieren Funktionen $F: I \rightarrow \mathbb{R}$ und $G: J \rightarrow \mathbb{R}$ durch

$$F(x) := \int_{x_0}^x f(t) dt, \quad G(y) := \int_{y_0}^y \frac{dt}{g(t)}.$$

Es sei $I' \subset I$ ein Intervall mit $x_0 \in I'$ und $F(I') \subset G(J)$. Dann existiert genau eine Lösung $\varphi: I' \rightarrow \mathbb{R}$ der Differentialgleichung

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

mit der "Anfangsbedingung" $\varphi(x_0) = y_0$. Diese Lösung genügt der Beziehung

$$G(\varphi(x)) = F(x) \quad \text{für alle } x \in I'. \quad (*)$$

Beweis. a) Wir zeigen zunächst: Ist $\varphi: I' \rightarrow \mathbb{R}$ irgend eine Lösung der Differentialgleichung $y' = f(x)g(y)$ mit $\varphi(x_0) = y_0$, so gilt (*). Aus

$$\varphi'(x) = f(x)g(\varphi(x))$$

folgt nämlich

$$\int_{x_0}^x \frac{\varphi'(t)}{g(\varphi(t))} dt = \int_{x_0}^x f(t) dt.$$

Führt man im linken Integral die Substitution $u = \varphi(t)$ durch, erhält man

$$\int_{y_0}^{\varphi(x)} \frac{du}{g(u)} = \int_{x_0}^x f(t) dt.$$

Dies bedeutet aber $G(\varphi(x)) = F(x)$.

b) Da $G'(y) = \frac{1}{g(y)} \neq 0$, ist G streng monoton, besitzt also eine stetig differenzierbare Umkehrfunktion

$$H: G(J) \longrightarrow \mathbb{R}.$$

Aus (*) folgt daher

$$\varphi(x) = H(F(x)) \quad \text{für alle } x \in I'. \quad (**)$$

Dies bedeutet, dass die Lösung mit der gegebenen Anfangsbedingung, falls sie überhaupt existiert, eindeutig bestimmt ist.

c) Um die Existenz zu beweisen, nehmen wir die Gleichung (**) als Definition für die Funktion $\varphi: I' \rightarrow \mathbb{R}$. Diese Funktion ist stetig differenzierbar und es gilt

$$\varphi(x_0) = H(F(x_0)) = H(0) = y_0,$$

da $F(x_0) = 0 = G(y_0)$. Weiter folgt aus (**) die Beziehung $G(\varphi(x)) = F(x)$, also

$$G'(\varphi(x))\varphi'(x) = \frac{\varphi'(x)}{g(\varphi(x))} = F'(x) = f(x),$$

d.h.

$$\varphi'(x) = f(x)g(\varphi(x)).$$

Daher erfüllt φ die Differentialgleichung $y' = f(x)g(y)$, q.e.d.

Bemerkung. Man führt die Lösung einer Differentialgleichung mit getrennten Variablen oft in einprägsamer, aber etwas nachlässiger Form wie folgt durch: Man schreibt $y' = f(x)g(y)$ als

$$\frac{dy}{dx} = f(x)g(y)$$

und formt dann um:

$$\begin{aligned} \frac{dy}{g(y)} &= f(x)dx, \\ \int \frac{dy}{g(y)} &= \int f(x)dx + \text{const.} \end{aligned}$$

Die zuletzt erhaltene Gleichung hat man nach y aufzulösen, um zu einer Lösung $y = \varphi(x)$ der Differentialgleichung zu gelangen. Satz 1 ist nichts anderes als eine Präzisierung der obigen formalen Rechnung.

Beispiele

(11.1) Wir wollen die Lösung der Differentialgleichung

$$y' = -\frac{y}{x} \quad \text{in } I \times J := \mathbb{R}_+^* \times \mathbb{R}_+^*$$

mit der Anfangsbedingung $y(1) = c$, ($c > 0$), bestimmen. Dies ist eine Differentialgleichung mit getrennten Variablen,

$$y' = f(x)g(y) \quad \text{mit} \quad f(x) := -\frac{1}{x} \quad \text{und} \quad g(y) := y.$$

Nach Satz 1 müssen wir also folgende Funktionen berechnen:

$$F(x) = \int_1^x f(t) dt = - \int_1^x \frac{dt}{t} = -\log x,$$

$$G(y) = \int_c^y \frac{dt}{g(t)} = \int_c^y \frac{dt}{t} = \log t \Big|_c^y = \log(y/c).$$

Da der Logarithmus \mathbb{R}_+^* auf ganz \mathbb{R} abbildet, gilt hier $G(J) = G(\mathbb{R}_+^*) = \mathbb{R}$, wir können also $I' = I = \mathbb{R}_+^*$ wählen, da $F(\mathbb{R}_+^*) = \mathbb{R} = G(\mathbb{R}_+^*)$. Es gibt also eine auf ganz \mathbb{R}_+^* definierte Lösung $\varphi: \mathbb{R}_+^* \rightarrow \mathbb{R}$ mit

$$G(\varphi(x)) = F(x), \quad \text{d.h.} \quad \log \frac{\varphi(x)}{c} = -\log x \quad \text{für alle } x \in \mathbb{R}_+^*.$$

Nimmt man von beiden Seiten die Exponentialfunktion, erhält man

$$\varphi(x) = \frac{c}{x} \quad \text{für alle } x \in \mathbb{R}_+^*.$$

Die Gesamtheit aller Lösungen der Differentialgleichung $y' = -\frac{y}{x}$ in $\mathbb{R}_+^* \times \mathbb{R}_+^*$ wird also durch die Hyperbelschar $xy = c$, $c > 0$, gegeben, siehe Bild 11.4. Man sieht an dem Bild auch, dass durch jeden Punkt $(x_0, y_0) \in \mathbb{R}_+^* \times \mathbb{R}_+^*$ genau eine, auf ganz \mathbb{R}_+^* definierte, Lösungskurve geht.

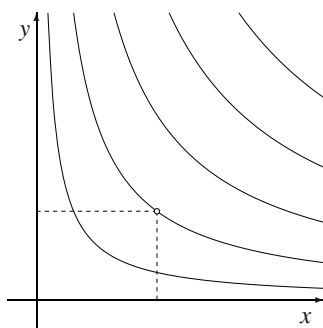


Bild 11.4

(11.2) Wir untersuchen jetzt die Differentialgleichung

$$y' = \frac{x}{y} \quad \text{in } \mathbb{R}_+^* \times \mathbb{R}_+^*.$$

Das Richtungsfeld dieser Differentialgleichung ist zu dem aus dem vorigen Beispiel orthogonal, denn die Richtungsvektoren $(1, -y/x)$ und $(1, x/y)$ stehen für alle $(x, y) \in \mathbb{R}_+^* \times \mathbb{R}_+^*$ aufeinander senkrecht.

Wir bestimmen diesmal die Lösung durch einen Punkt $(x_0, y_0) \in \mathbb{R}_+^* \times \mathbb{R}_+^*$ durch die folgenden informellen Umformungen:

$$\begin{aligned} \frac{dy}{dx} &= \frac{x}{y}, \\ y dy &= x dx, \\ \int_{y_0}^y \eta d\eta &= \int_{x_0}^x \xi d\xi, \\ \frac{1}{2}(y^2 - y_0^2) &= \frac{1}{2}(x^2 - x_0^2), \\ y^2 &= x^2 + (y_0^2 - x_0^2), \\ y = \varphi(x) &= \sqrt{x^2 + c} \quad \text{mit } c := y_0^2 - x_0^2. \end{aligned}$$

Durch Differenzieren kann man direkt nachprüfen, dass die erhaltene Funktion φ die Differentialgleichung erfüllt. Man beachte aber folgende Besonderheit gegenüber dem vorigen Beispiel: Nur für $y_0 \geq x_0$ ist die Lösung φ auf dem ganzen Intervall \mathbb{R}_+^* definiert. Falls $y_0 < x_0$, ist das maximale Definitionsintervall der Lösung gleich $\left] \sqrt{x_0^2 - y_0^2}, +\infty \right[$.

Geometrische Interpretation. Die Lösungen sind die in der positiven Viertel-Ebene gelegenen Teile der Kurvenschar $y^2 - x^2 = c$, $c \in \mathbb{R}$. Diese Kurven sind die “Orthogonal-Trajektorien” zu denen des vorigen Beispiels, siehe Bild 11.5

Lineare Differentialgleichungen

Sei $I \subset \mathbb{R}$ ein Intervall und seien $a, b: I \rightarrow \mathbb{R}$ stetige Funktionen. Dann nennt man

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

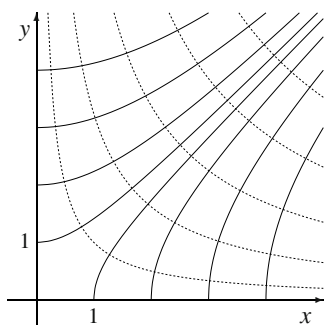


Bild 11.5

eine lineare Differentialgleichung 1. Ordnung, und zwar *homogen*, falls $b = 0$, sonst *inhomogen*. Wir beschäftigen uns zunächst mit der Lösung der homogenen linearen Differentialgleichung

$$y' = a(x)y.$$

Satz 2. *Mit den obigen Bezeichnungen gilt: Sei $x_0 \in I$ und $c \in \mathbb{R}$. Dann gibt es genau eine Lösung $\varphi: I \rightarrow \mathbb{R}$ der Differentialgleichung $y' = a(x)y$, die der Anfangsbedingung $\varphi(x_0) = c$ genügt, nämlich*

$$\varphi(x) := c \exp\left(\int_{x_0}^x a(t) dt\right).$$

Beweis. Die homogene lineare Differentialgleichung ist eine spezielle Differentialgleichung mit getrennten Variablen; man könnte also Satz 1 anwenden (wobei man allerdings eine Fallunterscheidung $c \neq 0, c = 0$ zu machen hätte). Hier sieht man aber direkt aus der Definition von φ , dass $\varphi(x_0) = c$ und

$$\varphi'(x) = a(x)\varphi(x).$$

Die Eindeutigkeit kann man so zeigen: Nach dem gerade Gesagten genügt

$$\varphi_0(x) := \exp\left(-\int_{x_0}^x a(t) dt\right)$$

der Gleichung $\varphi'_0(x) = -a(x)\varphi_0(x)$. Sei jetzt $\psi: I \rightarrow \mathbb{R}$ eine beliebige Lösung der Differentialgleichung $y' = a(x)y$ mit der Anfangsbedingung $\psi(x_0) = c$. Wir

bilden das Produkt $\psi_0(x) := \psi(x)\varphi_0(x)$ und differenzieren mit der Produktregel

$$\begin{aligned}\psi'_0(x) &= \psi'(x)\varphi_0(x) + \psi(x)\varphi'_0(x) \\ &= a(x)\psi(x)\varphi_0(x) - \psi(x)a(x)\varphi_0(x) = 0.\end{aligned}$$

Die Funktion ψ_0 ist also eine Konstante; daher gilt

$$\psi_0(x) = \psi_0(x_0) = \psi(x_0)\varphi_0(x_0) = c \exp(0) = c \quad \text{für alle } x \in I.$$

Es folgt

$$\psi(x) = c\varphi_0(x)^{-1} = c \exp\left(\int_{x_0}^x a(t) dt\right), \quad \text{q.e.d.}$$

Beispiele

(11.3) Die Lösungen $\varphi: \mathbb{R} \rightarrow \mathbb{R}$ der Differentialgleichung

$$y' = ky, \quad (k \in \mathbb{R}),$$

mit der Anfangsbedingung $\varphi(x_0) = c$ haben die Gestalt

$$\varphi(x) = c e^{k(x-x_0)}.$$

(11.4) Die schon in (11.1) behandelte Differentialgleichung

$$y' = -\frac{y}{x}, \quad (x > 0),$$

lässt sich als lineare Differentialgleichung $y' = a(x)y$ mit $a(x) = -1/x$ auffassen. Die Lösung mit der Anfangsbedingung $\varphi(1) = c$ ist also

$$\varphi(x) = c \exp\left(-\int_1^x \frac{dt}{t}\right) = c \exp(-\log x) = \frac{c}{x}.$$

Variation der Konstanten

Wir kommen jetzt zur Behandlung der inhomogenen linearen Differentialgleichung

$$y' = a(x)y + b(x), \tag{I}$$

wo $a, b: I \rightarrow \mathbb{R}$ stetige Funktionen auf dem Intervall $I \subset \mathbb{R}$ sind. Sei $x_0 \in I$ und $c \in \mathbb{R}$. Wir suchen eine Lösung $\varphi: I \rightarrow \mathbb{R}$ der Differentialgleichung mit der Anfangsbedingung $\varphi(x_0) = c$. Dazu gehen wir aus von einer Lösung

$$\varphi_0(x) := \exp\left(\int_{x_0}^x a(t) dt\right)$$

der zugehörigen homogenen linearen Differentialgleichung

$$y' = a(x)y \tag{H}$$

und machen einen als *Variation der Konstanten* bekannten Ansatz

$$\varphi(x) = \varphi_0(x)u(x)$$

für die Lösung φ der inhomogenen Gleichung (I) mit einer noch unbekannten Funktion $u: I \rightarrow \mathbb{R}$. (Da $\varphi_0(x) \neq 0$ für alle $x \in I$, lässt sich jede Funktion φ in dieser Form schreiben.) Wir untersuchen jetzt, welchen Bedingungen u genügen muss, damit φ die Differentialgleichung (I) erfüllt. Es ist

$$\begin{aligned} \varphi' &= \varphi_0' u + \varphi_0 u', \\ a\varphi + b &= a\varphi_0 u + b. \end{aligned}$$

Da $\varphi_0' = a\varphi_0$, gilt $\varphi' = a\varphi + b$ genau dann, wenn

$$\varphi_0 u' = b, \quad \text{d.h.} \quad u(x) = \int_{x_0}^x \varphi_0(t)^{-1} b(t) dt + \text{const.}$$

Damit die Anfangsbedingung $\varphi(x_0) = c$ erfüllt ist, muss $u(x_0) = c$ sein, man hat also die Integrationskonstante $\text{const} = c$ zu setzen. Wir fassen unser Ergebnis zusammen im folgenden Satz.

Satz 3. Sei $I \subset \mathbb{R}$ ein Intervall und seien $a, b: I \rightarrow \mathbb{R}$ stetige Funktionen. Dann gibt es zu vorgegebenen $x_0 \in I$ und $c \in \mathbb{R}$ genau eine Lösung $\varphi: I \rightarrow \mathbb{R}$ der Differentialgleichung

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

mit der Anfangsbedingung $\varphi(x_0) = c$, nämlich

$$\varphi(x) = \varphi_0(x) \left(c + \int_{x_0}^x \varphi_0(t)^{-1} b(t) dt \right),$$

wobei

$$\varphi_0(x) := \exp\left(\int_{x_0}^x a(t) dt\right).$$

Man nennt das hier beschriebene Verfahren zur Bestimmung einer Lösung der inhomogenen Gleichung Variation der Konstanten. Denn die Lösungen der homogenen Gleichung haben die Gestalt $\varphi_0(x)c$ mit einer Konstanten c und man erhält die Lösungen der inhomogenen Gleichung, indem man die Konstante c durch eine “variable” Funktion $u(x)$ ersetzt. Wir werden darauf in allgemeiner Form in § 13 noch einmal zurückkommen.

(11.5) Beispiel. Wir betrachten die Differentialgleichung

$$y' = 2xy + x^3.$$

Die homogene Gleichung $y' = 2xy$ besitzt die Lösung

$$\varphi_0(x) = \exp\left(\int_0^x 2t dt\right) = e^{x^2}.$$

Deshalb bekommt man eine Lösung $\varphi: \mathbb{R} \rightarrow \mathbb{R}$ der inhomogenen Gleichung $y' = 2xy + x^3$ mit der Anfangsbedingung $\varphi(0) = c$ durch

$$\varphi(x) = e^{x^2} \left(c + \int_0^x t^3 e^{-t^2} dt \right).$$

Das Integral kann man mittels der Substitution $s = t^2$ und partieller Integration auswerten, was dem Leser als Übung überlassen sei. Man erhält

$$\int_0^x t^3 e^{-t^2} dt = \frac{1}{2} - \frac{1}{2}(x^2 + 1)e^{-x^2},$$

also

$$\varphi(x) = \left(c + \frac{1}{2}\right)e^{x^2} - \frac{1}{2}(x^2 + 1).$$

Homogene Differentialgleichung

Sei $J \subset \mathbb{R}$ ein Intervall, $f: J \rightarrow \mathbb{R}$ eine stetige Funktion und

$$G := \{(x, y) \in \mathbb{R}^* \times \mathbb{R} : \frac{y}{x} \in J\}.$$

Dann heißt

$$y' = f\left(\frac{y}{x}\right), \quad (x, y) \in G,$$

homogene Differentialgleichung (nicht zu verwechseln mit den homogenen linearen Differentialgleichungen). Der Name kommt daher, dass man die Differentialgleichung auch schreiben kann als

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

mit einer homogenen Funktion h vom Grad 0, d.h. $h(tx, ty) = h(x, y)$ für alle $t \in \mathbb{R}^*$. Setzt man $f(\eta) := h(1, \eta)$, so gilt $h(x, y) = f(y/x)$.

Der nächste Satz zeigt, dass man eine homogene Differentialgleichung auf eine Differentialgleichung mit getrennten Variablen zurückführen kann.

Satz 4. *Die obigen Bezeichnungen seien beibehalten. Weiter sei $I \subset \mathbb{R}^*$ ein Intervall und $(x_0, y_0) \in G$ ein Punkt mit $x_0 \in I$. Dann gilt: Eine Funktion $\varphi: I \rightarrow \mathbb{R}$ ist genau dann Lösung der Differentialgleichung*

$$y' = f\left(\frac{y}{x}\right) \tag{H1}$$

mit der Anfangsbedingung $\varphi(x_0) = y_0$, wenn die Funktion

$$\psi: I \longrightarrow \mathbb{R}, \quad \psi(x) := \frac{\varphi(x)}{x},$$

Lösung der Differentialgleichung

$$z' = \frac{1}{x}(f(z) - z), \quad (x, z) \in \mathbb{R}^* \times J, \tag{H2}$$

mit der Anfangsbedingung $\psi(x_0) = \frac{y_0}{x_0}$ ist.

Beweis. Zunächst verifiziert man leicht, dass der Graph der Funktion $\varphi: I \rightarrow \mathbb{R}$ genau dann in

$$G = \{(x, y) \in \mathbb{R}^* \times \mathbb{R} : y/x \in J\}$$

liegt, wenn der Graph von $\psi: I \rightarrow \mathbb{R}$ Teilmenge von $\mathbb{R}^* \times J$ ist.

a) Sei jetzt vorausgesetzt, dass φ die Differentialgleichung (H1) erfüllt. Dann

gilt

$$\begin{aligned}\psi'(x) &= \frac{d}{dx} \left(\frac{\varphi(x)}{x} \right) = \frac{\varphi'(x)}{x} - \frac{\varphi(x)}{x^2} = \frac{1}{x} f \left(\frac{\varphi(x)}{x} \right) - \frac{\varphi(x)}{x^2} \\ &= \frac{1}{x} (f(\psi(x)) - \psi(x)),\end{aligned}$$

d.h. $z = \psi(x)$ ist Lösung von (H2).

b) Sei umgekehrt vorausgesetzt, dass ψ die Gleichung (H2) erfüllt. Dann gilt

$$\begin{aligned}\varphi'(x) &= \frac{d}{dx} (x\psi(x)) = \psi(x) + x\psi'(x) \\ &= \psi(x) + (f(\psi(x)) - \psi(x)) = f \left(\frac{\varphi(x)}{x} \right),\end{aligned}$$

d.h. $y = \varphi(x)$ ist Lösung von (H1), q.e.d.

Bemerkung. Man drückt dies auch kurz so aus: Die Differentialgleichung (H1) geht durch die Substitution $z = \frac{y}{x}$ in (H2) über. Die Rechnung lautet in etwas salopper Form: Aus $y = xz$ folgt $y' = z + xz'$, also geht die Differentialgleichung $y' = f\left(\frac{y}{x}\right)$ über in

$$z + xz' = f(z), \quad \text{d.h.} \quad z' = \frac{1}{x}(f(z) - z).$$

(11.6) Beispiel. Die Differentialgleichung

$$y' = 1 + \left(\frac{y}{x}\right) + \left(\frac{y}{x}\right)^2 \quad \text{in } \mathbb{R}_+^* \times \mathbb{R}$$

geht durch die Substitution $z = \frac{y}{x}$ über in

$$z' = \frac{1}{x}(1 + z^2),$$

d.h.

$$\frac{dz}{1+z^2} = \frac{dx}{x}.$$

Für die Lösung dieser Differentialgleichung mit der Anfangsbedingung $z(x_0) = z_0 = y_0/x_0$ gilt

$$\arctan z - \arctan z_0 = \log x - \log x_0,$$

d.h.

$$z = \tan\left(\log \frac{x}{x_0} + \alpha\right), \quad \text{wobei } \alpha = \arctan \frac{y_0}{x_0}.$$

Die ursprüngliche Differentialgleichung wird also gelöst durch

$$y = x \tan\left(\log \frac{x}{x_0} + \alpha\right).$$

Diese Lösung ist definiert in einer gewissen Umgebung von x_0 .

AUFGABEN

11.1. Man bestimme die allgemeine Lösung der folgenden Differentialgleichungen, d.h. die Lösung durch einen beliebigen Punkt (x_0, y_0) des Definitionsbereichs.

- a) $y' = e^y \cos x$,
- b) $y' = \sqrt{1-y^2}$, $(|y| < 1)$,
- c) $y' = \frac{1}{y} \sqrt{1-y^2}$, $(0 < y < 1)$,
- d) $y' = (a^2 + x^2)(b^2 + y^2)$, $a, b \in \mathbb{R}$,
- e) $(1-x^2)y' - xy + 1 = 0$, $(|x| < 1)$.

11.2. Man bestimme die allgemeine Lösung der folgenden Differentialgleichungen.

- a) $y' = (x+y)^2$,
- b) $(1+x^2)y' + xy - xy^2 = 0$,
- c) $y' + y + (\sin x + e^x)y^3 = 0$.

Anleitung. Man verwende folgende Substitutionen:

$$\text{a) } z = x + y, \quad \text{b) } z = \frac{1}{y}, \quad \text{c) } z = \frac{1}{y^2}.$$

11.3. Man bestimme eine Lösung der Differentialgleichung

$$y' = \frac{x-y}{x+y}, \quad (x+y > 0),$$

durch den Punkt $(x_0, y_0) = (0, 1)$.

§ 12 Existenz- und Eindeutigkeitssatz

Nachdem wir im vorigen Paragraphen die Lösungen einiger einfacher spezieller Differentialgleichungen studiert haben, beweisen wir jetzt einen allgemeinen Existenz- und Eindeutigkeitssatz. Dabei behandeln wir sogleich Systeme gewöhnlicher Differentialgleichungen 1. Ordnung. Dies liefert gleichzeitig einen Existenz- und Eindeutigkeitssatz für Differentialgleichungen höherer Ordnung, da sich diese auf Systeme von Differentialgleichungen 1. Ordnung zurückführen lassen.

Definition. Sei $G \subset \mathbb{R} \times \mathbb{R}^n$ eine Teilmenge und

$$f : G \longrightarrow \mathbb{R}^n, \quad (x, y) \mapsto f(x, y)$$

eine stetige Funktion. Dann nennt man

$$y' = f(x, y) \tag{1}$$

ein *System* von n Differentialgleichungen erster Ordnung. Unter einer Lösung von (1) versteht man eine auf einem Intervall $I \subset \mathbb{R}$ differenzierbare vektorwertige Funktion

$$\varphi : I \longrightarrow \mathbb{R}^n$$

mit folgenden Eigenschaften:

- a) Der Graph von φ ist in G enthalten.
- b) $\varphi'(x) = f(x, \varphi(x))$ für alle $x \in I$.

Bemerkungen.

1) Rein äußerlich sieht die Gleichung (1) genauso aus wie die im vorigen Paragraphen behandelten Differentialgleichungen für skalarwertige Funktionen; jedoch ist hier y ein Vektor mit n Komponenten. Sind

$$y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \quad \text{und} \quad f = \begin{pmatrix} f_1 \\ \vdots \\ f_n \end{pmatrix}$$

die Komponentendarstellungen von y bzw. f , so lautet (1) ausgeschrieben wie folgt:

$$\begin{cases} y'_1 = f_1(x, y_1, \dots, y_n), \\ y'_2 = f_2(x, y_1, \dots, y_n), \\ \dots \\ y'_n = f_n(x, y_1, \dots, y_n). \end{cases}$$

2) Man beachte, dass wir zwar y von einer auf mehrere Dimensionen verallgemeinert haben, aber x immer noch eindimensional ist, das heißt Funktionen einer Variablen gesucht sind. Man spricht deshalb genauer von *gewöhnlichen* Differentialgleichungen im Gegensatz zu *partiellen* Differentialgleichungen, bei denen Funktionen mehrerer Variablen gesucht werden und die Bedingungengleichungen die partiellen Ableitungen der gesuchten Funktionen enthalten. Die Potentialgleichung, die Wellengleichung und die Wärmeleitungsgleichung (siehe § 5) sind Beispiele partieller Differentialgleichungen.

Satz 1 (Zurückführung auf Integralgleichung). *Sei $G \subset \mathbb{R} \times \mathbb{R}^n$ und $f: G \rightarrow \mathbb{R}^n$ eine stetige Abbildung. Weiter sei ein Punkt $(a, c) \in G$ gegeben. Dann gilt: Eine stetige Funktion*

$$\varphi: I \longrightarrow \mathbb{R}^n,$$

die auf einem Intervall $I \subset \mathbb{R}$ mit $a \in I$ definiert ist, und deren Graph in G enthalten ist, ist genau dann Lösung der Differentialgleichung

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

mit der Anfangsbedingung $\varphi(a) = c$, wenn folgende Integralgleichung gilt:

$$\varphi(x) = c + \int_a^x f(t, \varphi(t)) dt \quad \text{für alle } x \in I.$$

Beweis. a) Sei die Integralgleichung erfüllt. Setzt man darin $x = a$, ergibt sich $\varphi(a) = c$. Da die Funktion $t \mapsto f(t, \varphi(t))$ auf I stetig ist, folgt aus dem Fundamentalsatz der Differential- und Integralrechnung

$$\frac{d}{dx} \int_a^x f(t, \varphi(t)) dt = f(x, \varphi(x)),$$

also ist φ differenzierbar mit $\varphi'(x) = f(x, \varphi(x))$.

b) Sei umgekehrt φ Lösung der Differentialgleichung, also differenzierbar mit $\varphi'(x) = f(x, \varphi(x))$, und es sei die Anfangsbedingung $\varphi(a) = c$ erfüllt. Dann ist

$$\int_a^x f(t, \varphi(t)) dt = \int_a^x \varphi'(t) dt = \varphi(x) - \varphi(a) = \varphi(x) - c,$$

also ist die Integralgleichung erfüllt, q.e.d.

Wir werden jetzt eine Bedingung kennenlernen, mit der man die (lokale) Existenz und Eindeutigkeit der Lösungen von gewöhnlichen Differentialgleichungen beweisen kann.

Definition (Lipschitz-Bedingung). Sei $G \subset \mathbb{R} \times \mathbb{R}^n$ und

$$f : G \longrightarrow \mathbb{R}^n, \quad (x, y) \mapsto f(x, y),$$

eine Funktion. Man sagt, f genüge in G einer *Lipschitz-Bedingung* (bzgl. der Variablen y) mit der Lipschitz-Konstanten $L \geq 0$, wenn

$$\|f(x, y) - f(x, \tilde{y})\| \leq L \|y - \tilde{y}\| \quad \text{für alle } (x, y), (x, \tilde{y}) \in G.$$

Man sagt, f genüge in G *lokal* einer Lipschitz-Bedingung, falls jeder Punkt $(a, b) \in G$ eine Umgebung U besitzt, so dass f in $G \cap U$ einer Lipschitz-Bedingung mit einer gewissen (von U abhängigen) Konstanten $L \in \mathbb{R}_+$ genügt.

Ein einfaches Kriterium für das Erfülltsein einer lokalen Lipschitz-Bedingung liefert der folgende Satz.

Satz 2. Sei $G \subset \mathbb{R} \times \mathbb{R}^n$ *offen* und

$$f : G \longrightarrow \mathbb{R}^n, \quad (x, y) \mapsto f(x, y),$$

eine bzgl. der Variablen $y = (y_1, \dots, y_n)$ stetig partiell differenzierbare Funktion. Dann genügt f in G lokal einer Lipschitz-Bedingung.

Beweis. Sei $(a, b) \in G$. Es gibt ein $r > 0$, so dass

$$V := \{(x, y) \in \mathbb{R} \times \mathbb{R}^n : |x - a| \leq r, \|y - b\| \leq r\}$$

ganz in G liegt. V ist eine kompakte Umgebung von (a, b) . Da nach Voraussetzung alle Komponenten der $(n \times n)$ -Matrix $\frac{\partial f}{\partial y}$ stetige Funktionen sind, gilt

$$L := \sup_{(x, y) \in V} \left\| \frac{\partial f}{\partial y}(x, y) \right\| < \infty.$$

Aus dem Mittelwertsatz (§ 6, Satz 5) folgt nun für alle $(x, y), (x, \tilde{y}) \in V$

$$\|f(x, y) - f(x, \tilde{y})\| \leq L \|y - \tilde{y}\|, \quad \text{q.e.d.}$$

Satz 3 (Eindeutigkeitssatz). *Sei $G \subset \mathbb{R} \times \mathbb{R}^n$ und $f: G \rightarrow \mathbb{R}^n$ eine stetige Funktion, die lokal einer Lipschitz-Bedingung genügt. Seien*

$$\varphi, \psi: I \longrightarrow \mathbb{R}^n$$

zwei Lösungen der Differentialgleichung

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

über einem Intervall $I \subset \mathbb{R}$. Gilt dann

$$\varphi(x_0) = \psi(x_0) \quad \text{für ein } x_0 \in I,$$

so folgt

$$\varphi(x) = \psi(x) \quad \text{für alle } x \in I.$$

Beweis. a) Wir zeigen zunächst Eindeutigkeit im Kleinen: Ist $\varphi(a) = \psi(a)$ für ein $a \in I$, so gibt es ein $\varepsilon > 0$, so dass $\varphi(x) = \psi(x)$ für alle $x \in I$ mit $|x - a| < \varepsilon$.

Aus Satz 1 folgt

$$\varphi(x) - \psi(x) = \int_a^x (f(t, \varphi(t)) - f(t, \psi(t))) dt. \quad (2)$$

Da f lokal einer Lipschitz-Bedingung genügt, gibt es Konstanten $L \geq 0$ und $\delta > 0$, so dass

$$\|f(t, \varphi(t)) - f(t, \psi(t))\| \leq L \|\varphi(t) - \psi(t)\| \quad (3)$$

für alle $t \in I \cap B_\delta(a) = \{t \in I : |t - a| < \delta\}$. Wir können außerdem annehmen, dass φ und ψ in $I \cap B_\delta(a)$ beschränkt sind. Sei

$$\varepsilon := \min(\delta, 1/(2L))$$

und

$$M := \sup\{\|\varphi(t) - \psi(t)\| : t \in I \cap B_\varepsilon(a)\}.$$

Aus (2) und (3) folgt für alle $x \in I \cap B_\varepsilon(a)$

$$\|\varphi(x) - \psi(x)\| \leq L \left| \int_a^x \|\varphi(t) - \psi(t)\| dt \right| \leq L|x-a|M \leq \frac{1}{2}M,$$

also auch $M \leq \frac{1}{2}M$. Dies ist nur möglich, wenn $M = 0$, d.h. wenn φ und ψ in $I \cap B_\varepsilon(a)$ übereinstimmen. Damit ist a) bewiesen.

b) Wir zeigen jetzt

$$\varphi(x) = \psi(x) \quad \text{für alle } x \in I \text{ mit } x \geq x_0.$$

Es sei

$$x_1 := \sup\{\xi \in I : \varphi|_{[x_0, \xi]} = \psi|_{[x_0, \xi]}\}.$$

Falls $x_1 = \infty$ oder x_1 gleich dem rechten Intervallende ist, sind wir fertig. Andernfalls gibt es ein $\delta > 0$, so dass $[x_1, x_1 + \delta] \subset I$. Da φ und ψ stetig sind, gilt $\varphi(x_1) = \psi(x_1)$. Nach a) gibt es ein $\varepsilon > 0$ mit

$$\varphi(x) = \psi(x) \quad \text{für alle } x \in I \cap B_\varepsilon(x_1).$$

Dies steht aber im Widerspruch zur Definition von x_1 . Daher gilt $\varphi(x) = \psi(x)$ für alle $x \in I$ mit $x \geq x_0$.

c) Analog zeigt man $\varphi(x) = \psi(x)$ für alle $x \in I$ mit $x \leq x_0$.

(12.1) Beispiel. Wir geben ein Beispiel einer Differentialgleichung an, für die der Eindeutigkeitssatz nicht gilt:

$$y' = y^{2/3} \quad (\text{definiert in } \mathbb{R} \times \mathbb{R}).$$

Eine spezielle Lösung ist $\varphi_0(x) := 0$ für alle $x \in \mathbb{R}$. Da die Differentialgleichung eine mit getrennten Variablen ist, lassen sich die Lösungen im Bereich $y \neq 0$ mit § 11, Satz 1, bestimmen. Man sieht aber leicht direkt, dass für beliebiges $a \in \mathbb{R}$ die Funktion

$$\psi_a : \mathbb{R} \rightarrow \mathbb{R}, \quad x \mapsto \psi_a(x) := \frac{1}{27}(x-a)^3$$

die Differentialgleichung erfüllt, denn $\psi'_a(x) = \frac{1}{9}(x-a)^2 = \psi_a(x)^{2/3}$. Da

$$\varphi_0(a) = \psi_a(a) = 0,$$

ist der Eindeutigkeitsatz verletzt. Es gibt aber noch unendlich viele andere Lösungen $\psi: \mathbb{R} \rightarrow \mathbb{R}$ mit $\psi(a) = 0$. Seien b, c beliebige reelle Konstanten mit $b \leq a \leq c$ und

$$\psi(x) := \begin{cases} \psi_b(x) & \text{für } x \leq b, \\ 0 & \text{für } b < x < c, \\ \psi_c(x) & \text{für } x \geq c, \end{cases}$$

siehe Bild 12.1. Die Funktion ψ ist auf ganz \mathbb{R} differenzierbar, erfüllt die Differentialgleichung $\psi' = \psi^{2/3}$, und es gilt $\psi(a) = 0$.

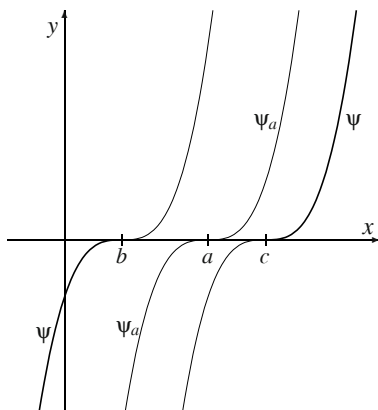


Bild 12.1

Nach Satz 3 kann also die Funktion $f(x, y) = y^{2/3}$ nicht überall einer Lipschitz-Bedingung genügen. Zwar ist f für $y \neq 0$ partiell nach y differenzierbar mit

$$\frac{\partial}{\partial y} f(x, y) = \frac{2}{3} y^{-1/3},$$

und nach Satz 2 genügt deshalb f in $\mathbb{R} \times \mathbb{R}^*$ lokal einer Lipschitz-Bedingung. Jedoch erfüllt f in keiner Umgebung eines Punktes $(a, 0)$ eine Lipschitz-Bedingung.

Satz 4 (Existenzsatz von Picard-Lindelöf). *Sei $G \subset \mathbb{R} \times \mathbb{R}^n$ offen und $f: G \rightarrow \mathbb{R}^n$ eine stetige Funktion, die lokal einer Lipschitz-Bedingung genügt. Dann gibt es zu jedem $(a, c) \in G$ ein $\varepsilon > 0$ und eine Lösung*

$$\varphi: [a - \varepsilon, a + \varepsilon] \longrightarrow \mathbb{R}^n$$

der Differentialgleichung $y' = f(x, y)$ mit der Anfangsbedingung $\varphi(a) = c$.

Bemerkung. Nach Satz 3 ist φ durch die Bedingung $\varphi(a) = c$ eindeutig bestimmt.

Beweis. Da die Differentialgleichung mit der Anfangsbedingung zur Integralgleichung

$$\varphi(x) = c + \int_a^x f(t, \varphi(t)) dt$$

äquivalent ist, beweisen wir den Satz mithilfe des Banachschen Fixpunktsatzes (§ 8, Satz 1). Wir verwenden den Banachraum $C([a - \varepsilon, a + \varepsilon], \mathbb{R}^n)$ aller stetigen Funktionen

$$\Psi : [a - \varepsilon, a + \varepsilon] \longrightarrow \mathbb{R}^n$$

(für ein noch zu bestimmendes $\varepsilon > 0$) mit der Supremumsnorm

$$\|\Psi\| := \sup\{\|\Psi(t)\| : |t - a| \leq \varepsilon\}.$$

Es gibt ein $\delta > 0$ und ein $r > 0$, so dass

$$Q_{\delta, r} := \{(x, y) \in \mathbb{R} \times \mathbb{R}^n : |x - a| \leq \delta, \|y - c\| \leq r\} \subset G$$

und die Funktion f in $Q_{\delta, r}$ einer Lipschitz-Bedingung mit einer gewissen Konstanten $L > 0$ genügt. Da f stetig und $Q_{\delta, r}$ kompakt ist, gibt es eine Konstante $M > 0$ mit

$$\|f(x, y)\| \leq M \quad \text{für alle } (x, y) \in Q_{\delta, r}.$$

Wir setzen jetzt

$$\varepsilon := \min\left(\delta, \frac{r}{M}, \frac{1}{2L}\right)$$

und

$$A := \{\Psi \in C([a - \varepsilon, a + \varepsilon], \mathbb{R}^n) : \|\Psi - c\| \leq r\}.$$

A ist eine abgeschlossene Teilmenge des Banachraums $C([a - \varepsilon, a + \varepsilon], \mathbb{R}^n)$.

Behauptung. Die Abbildung $T: A \rightarrow A$, die einer Funktion $\Psi \in A$ die Funktion $\eta := T(\Psi)$,

$$\eta(x) := c + \int_a^x f(t, \Psi(t)) dt \quad \text{für } x \in [a - \varepsilon, a + \varepsilon],$$

zuordnet, ist (i) wohldefiniert und (ii) eine Kontraktion.

(i) Da $\|\psi - c\| \leq r$, liegt der Graph von ψ in $Q_{\varepsilon, r} \subset G$, also ist $f(t, \psi(t))$ für alle $t \in [a - \varepsilon, a + \varepsilon]$ definiert und als Funktion von t stetig. Für die Funktion $\eta := T(\psi)$ erhält man die Abschätzung

$$\|\eta(x) - c\| = \left\| \int_a^x f(t, \psi(t)) dt \right\| \leq |x - a| \cdot M \leq \varepsilon M \leq r,$$

also liegt η wieder in A .

(ii) Seien $\varphi_1, \varphi_2 \in A$. Dann gilt

$$T(\varphi_1)(x) - T(\varphi_2)(x) = \int_a^x \left(f(t, \varphi_1(t)) - f(t, \varphi_2(t)) \right) dt,$$

also erhält man mit der Lipschitz-Bedingung für alle $x \in [a - \varepsilon, a + \varepsilon]$

$$\begin{aligned} \|T(\varphi_1)(x) - T(\varphi_2)(x)\| &\leq \left\| \int_a^x L \|\varphi_1(t) - \varphi_2(t)\| dt \right\| \\ &\leq L|x - a| \cdot \|\varphi_1 - \varphi_2\| \leq \frac{1}{2} \|\varphi_1 - \varphi_2\|. \end{aligned}$$

Dies zeigt, dass $T: A \rightarrow A$ eine Kontraktion mit der Kontraktionskonstanten $\frac{1}{2}$ ist. Nach dem Banachschen Fixpunktsatz gibt es deshalb ein $\varphi \in A$ mit $T(\varphi) = \varphi$, d.h.

$$\varphi(x) = c + \int_a^x f(t, \varphi(t)) dt \quad \text{für alle } x \in [a - \varepsilon, a + \varepsilon].$$

Dies ist die gesuchte Lösung der Differentialgleichung.

Bemerkung. Selbst wenn die rechte Seite der Differentialgleichung

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

auf ganz $\mathbb{R} \times \mathbb{R}^n$ definiert ist und überall lokal einer Lipschitz-Bedingung genügt, ist die Lösung mit der Anfangsbedingung $\varphi(a) = c$ unter Umständen nur in einer sehr kleinen Umgebung von a definiert. Dazu geben wir folgendes Beispiel:

(12.2) Wir betrachten Differentialgleichung

$$y' = 2xy^2.$$

Die rechte Seite ist auf $\mathbb{R} \times \mathbb{R}$ definiert und stetig partiell nach y differenzierbar, genügt also überall lokal einer Lipschitz-Bedingung. Wir suchen die Lösung

mit der Anfangsbedingung $\varphi(0) = c$. Für $c = 0$ ist das offensichtlich die Funktion $\varphi(x) = 0$ für alle $x \in \mathbb{R}$. Für $c \neq 0$ können wir die Lösung durch Trennung der Variablen berechnen:

$$\begin{aligned}\frac{dy}{y^2} &= 2x dx, \\ \int_c^y \frac{d\eta}{\eta^2} &= \int_0^x 2\xi d\xi, \\ -\frac{1}{y} + \frac{1}{c} &= x^2, \\ y = \varphi(x) &= \frac{1}{\gamma - x^2} \quad \text{mit } \gamma := \frac{1}{c}.\end{aligned}$$

Dies ist die Lösung der Differentialgleichung mit der Anfangsbedingung $\varphi(0) = c$, was man direkt durch Einsetzen verifizieren kann. Falls $c > 0$, ist das maximale Definitions-Intervall dieser Lösung gleich

$$I_c = \{x \in \mathbb{R} : |x| < \sqrt{\gamma} = c^{-1/2}\},$$

vgl. Bild 12.2. Nähert sich x von innen dem Rand des Intervalls, so strebt der Funktionswert der Lösung gegen ∞ . Für $c < 0$ ist die Lösung auf ganz \mathbb{R} definiert.

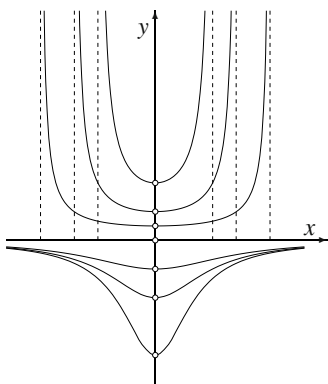


Bild 12.2

Picard-Lindelöfsches Iterationsverfahren

Beim Banachschen Fixpunktsatz für eine Kontraktion $T: A \rightarrow A$ haben wir gesehen, dass für jeden Anfangswert $\varphi_0 \in A$ die rekursiv definierte Folge $\varphi_k := T(\varphi_{k-1})$ gegen den Fixpunkt konvergiert. Angewendet auf das Anfangswertproblem $\varphi(a) = c$ für Lösungen der Differentialgleichung $y' = f(x, y)$ bedeutet dies: Sei φ_0 die konstante Funktion $\varphi_0(x) = c$ und

$$\varphi_k(x) := c + \int_a^x f(t, \varphi_{k-1}(t)) dt.$$

Dann konvergiert die Folge $(\varphi_k)_{k \in \mathbb{N}}$ in einer gewissen Umgebung $[a - \varepsilon, a + \varepsilon]$ von a gleichmäßig gegen eine Lösung der Differentialgleichung mit der Anfangsbedingung $\varphi(a) = c$.

In einfachen Beispielen kann man damit die Lösungen einer Differentialgleichung sogar explizit berechnen.

(12.3) Beispiel. Gegeben sei die Differentialgleichung

$$y' = 2xy \quad \text{in } \mathbb{R} \times \mathbb{R}.$$

Wir suchen eine Lösung φ , die der Anfangsbedingung $\varphi(0) = c$ genügt. Hier lautet die Iterationsvorschrift

$$\varphi_k(x) = c + 2 \int_0^x t \varphi_{k-1}(t) dt.$$

Mit $\varphi_0(x) = c$ ergibt sich

$$\begin{aligned} \varphi_1(x) &= c + 2c \int_0^x t dt = c(1 + x^2), \\ \varphi_2(x) &= c + 2c \int_0^x t(1 + t^2) dt = c \left(1 + x^2 + \frac{x^4}{2} \right). \end{aligned}$$

Allgemein beweist man durch vollständige Induktion

$$\varphi_k(x) = c \left(1 + x^2 + \frac{x^4}{2} + \frac{x^6}{3!} + \dots + \frac{x^{2k}}{k!} \right).$$

Daraus folgt

$$\varphi(x) = \lim_{k \rightarrow \infty} \varphi_k(x) = c \sum_{k=0}^{\infty} \frac{x^{2k}}{k!} = ce^{x^2}.$$

Dass diese Funktion die Differentialgleichung $y' = 2xy$ löst und die Anfangsbedingung $\varphi(0) = c$ erfüllt, sieht man unmittelbar durch Einsetzen.

Bemerkung. Die Differentialgleichung $y' = 2xy$ ist eine homogene lineare Differentialgleichung und kann deshalb auch leicht direkt gelöst werden (§11, Satz 2).

Differentialgleichungen höherer Ordnung

Definition. Sei $G \subset \mathbb{R} \times \mathbb{R}^n$ eine Teilmenge und

$$f : G \longrightarrow \mathbb{R}$$

eine stetige Funktion. Dann heißt

$$y^{(n)} = f(x, y, y', \dots, y^{(n-1)}) \quad (4)$$

eine Differentialgleichung n -ter Ordnung. Unter einer Lösung von (4) versteht man eine auf einem Intervall $I \subset \mathbb{R}$ definierte, n -mal differenzierbare Funktion

$$\varphi : I \longrightarrow \mathbb{R}$$

mit folgenden Eigenschaften:

a) Die Menge

$$\{(x, y_0, y_1, \dots, y_{n-1}) \in I \times \mathbb{R}^n : y_v = \varphi^{(v)}(x) \text{ für } 0 \leq v \leq n-1\}$$

ist in G enthalten.

b) Für alle $x \in I$ gilt

$$\varphi^{(n)}(x) = f(x, \varphi(x), \varphi'(x), \dots, \varphi^{(n-1)}(x)).$$

Eine Differentialgleichung n -ter Ordnung schreibt also eine Bedingung für die n -te Ableitung der gesuchten Funktion φ vor. Diese Ableitung hängt sowohl von x als auch dem Wert von φ und seiner Ableitungen bis zur $(n-1)$ -ten Ordnung im Punkt x ab.

Reduktion auf ein System 1. Ordnung

Man kann eine Differentialgleichung n -ter Ordnung auf ein System von Differentialgleichungen 1. Ordnung zurückführen. Dazu betrachten wir neben der

Differentialgleichung

$$y^{(n)} = f(x, y, y', \dots, y^{(n-1)}), \quad (4)$$

wo $f: G \rightarrow \mathbb{R}$ eine auf einer Teilmenge $G \subset \mathbb{R}^{n+1}$ definierte stetige Funktion ist, das Differentialgleichungs-System

$$\begin{cases} y'_0 = y_1, \\ y'_1 = y_2, \\ \vdots \\ y'_{n-2} = y_{n-1}, \\ y'_{n-1} = f(x, y_0, y_1, \dots, y_{n-1}). \end{cases} \quad (5)$$

Setzt man

$$Y := \begin{pmatrix} y_0 \\ \vdots \\ y_{n-2} \\ y_{n-1} \end{pmatrix} \quad \text{und} \quad F(x, Y) := \begin{pmatrix} y_1 \\ \vdots \\ y_{n-1} \\ f(x, Y) \end{pmatrix} \quad (6)$$

so schreibt sich (5) einfach als

$$Y' = F(x, Y).$$

Sei nun $\varphi: I \rightarrow \mathbb{R}$ eine Lösung von (4), d.h.

$$\varphi^{(n)}(x) = f(x, \varphi(x), \varphi'(x), \dots, \varphi^{(n-1)}(x)).$$

Damit definieren wir nun eine vektorwertige Funktion

$$\Phi = \begin{pmatrix} \varphi_0 \\ \varphi_1 \\ \vdots \\ \varphi_{n-1} \end{pmatrix} : I \longrightarrow \mathbb{R}^n$$

durch

$$\begin{aligned} \varphi_0(x) &:= \varphi(x), \\ \varphi_1(x) &:= \varphi'(x), \\ &\vdots \\ \varphi_{n-1}(x) &:= \varphi^{(n-1)}(x). \end{aligned}$$

Dann ist, wie unmittelbar aus der Definition folgt, Φ eine Lösung des Differentialgleichungs-Systems (5).

Sei nun umgekehrt vorausgesetzt, dass

$$\Phi = \begin{pmatrix} \varphi_0 \\ \varphi_1 \\ \vdots \\ \varphi_{n-1} \end{pmatrix} : I \longrightarrow \mathbb{R}^n$$

eine Lösung von (5) ist.

Behauptung. Dann ist die Funktion

$$\varphi := \varphi_0 : I \longrightarrow \mathbb{R}$$

eine Lösung der Differentialgleichung n -ter Ordnung (4).

Beweis. Aus den $n - 1$ ersten Gleichungen des Systems (5) folgt

$$\begin{aligned} \varphi_1 &= \varphi'_0 = \varphi', \\ \varphi_2 &= \varphi'_1 = \varphi'', \\ &\vdots \\ \varphi_{n-1} &= \varphi'_{n-2} = \varphi^{(n-1)}. \end{aligned}$$

Da φ_{n-1} einmal differenzierbar ist, folgt daraus, dass φ n -mal differenzierbar ist. Die n -te Gleichung von (5) liefert dann

$$\varphi^{(n)} = f(x, \varphi, \varphi', \dots, \varphi^{(n-1)}), \quad \text{q.e.d.}$$

Die Lösungen von (4) und (5) stehen also in ein-eindeutiger Beziehung zueinander. Dies bringt für theoretische Betrachtungen einen großen Vorteil mit sich; man kann sich auf die Betrachtung von Systemen von Differentialgleichungen 1. Ordnung beschränken.

Selbstverständlich können auch Differentialgleichungs-Systeme höherer Ordnung auf solche 1. Ordnung zurück geführt werden. Besonders wichtig sind Differentialgleichungs-Systeme 2. Ordnung, die häufig in der Physik auftreten aufgrund des Newtonschen Gesetzes Kraft = Masse \times Beschleunigung. Beschreibt etwa die vektorwertige Funktion $x = x(t)$ die Bewegung eines Massenpunkts als Funktion der Zeit t , so ist $\dot{x} = dx/dt$ seine Geschwindigkeit und $\ddot{x} = d^2x/dt^2$ seine Beschleunigung. Nimmt man an, dass die auf den Massenpunkt wirkende Kraft nur von der Zeit, dem Ort und der Geschwindigkeit

abhängt, so genügt die Bewegung dem Differentialgleichungs-System

$$m\ddot{x} = F(t, x, \dot{x}), \quad (m = \text{Masse}).$$

Hat x die drei Komponenten x_1, x_2, x_3 , so ist dies ein System von drei Differentialgleichungen 2. Ordnung für die gesuchten Funktionen x_1, x_2, x_3 . Dies System ist äquivalent zu einem System von sechs Differentialgleichungen 1. Ordnung

$$\begin{cases} \dot{x} = v, \\ \dot{v} = \frac{1}{m}F(t, x, v), \end{cases}$$

wobei $v = (v_1, v_2, v_3)$ die Geschwindigkeit bedeutet.

Wir übertragen nun den Eindeutigkeitssatz und den Existenzsatz, den wir für Systeme von Differentialgleichungen 1. Ordnung bewiesen haben, auf eine Differentialgleichung n -ter Ordnung

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

wobei $f: G \rightarrow \mathbb{R}$ eine auf einer Teilmenge $G \subset \mathbb{R} \times \mathbb{R}^n$ definierte stetige Funktion ist. Man sagt, f genüge lokal einer Lipschitz-Bedingung, wenn es zu jedem Punkt $z \in G$ eine Umgebung U und eine Konstante $L \in \mathbb{R}_+$ gibt, so dass

$$|f(x, Y) - f(x, \tilde{Y})| \leq L \|Y - \tilde{Y}\|$$

für alle $(x, Y), (x, \tilde{Y}) \in G$. Dies ist gleichbedeutend damit, dass die f gemäß (6) zugeordnete Funktion $F: G \rightarrow \mathbb{R}^n$ lokal einer Lipschitz-Bedingung genügt. Aus den Sätzen 3 und 4 folgt nun

Satz 5. Sei $G \subset \mathbb{R} \times \mathbb{R}^n$ offen und $f: G \rightarrow \mathbb{R}$ eine stetige Funktion, die lokal einer Lipschitz-Bedingung genügt.

a) (Eindeutigkeit) Seien $\varphi, \psi: I \rightarrow \mathbb{R}$ zwei Lösungen der Differentialgleichung

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

Für einen Punkt $a \in I$ gelte

$$\varphi(a) = \psi(a), \varphi'(a) = \psi'(a), \dots, \varphi^{(n-1)}(a) = \psi^{(n-1)}(a).$$

Dann gilt $\varphi(x) = \psi(x)$ für alle $x \in I$.

b) (Existenz) *Ist umgekehrt ein Punkt $(a, c_0, \dots, c_{n-1}) \in G$ vorgegeben, so gibt es ein $\varepsilon > 0$ und eine Lösung*

$$\varphi: [a - \varepsilon, a + \varepsilon] \longrightarrow \mathbb{R}$$

der Differentialgleichung (), die der Anfangsbedingung*

$$\varphi(a) = c_0, \varphi'(a) = c_1, \dots, \varphi^{(n-1)}(a) = c_{n-1}$$

genügt.

Um die Lösung einer Differentialgleichung n -ter Ordnung eindeutig festzulegen, muss man also nicht nur den Wert der Funktion an einer Stelle a des Definitionsintervalls vorschreiben, sondern auch alle Ableitungen der Ordnung $\leq n - 1$ im Punkt a . Speziell für die Differentialgleichungen zweiter Ordnung, die die Bewegung eines Teilchens als Funktion der Zeit beschreiben, bedeutet das: Die Bewegung ist eindeutig festgelegt, wenn Ort und Geschwindigkeit des Teilchens zu einem Anfangszeitpunkt bekannt sind.

(12.4) Die Differentialgleichung 2. Ordnung

$$y'' + y = 0 \tag{7}$$

besitzt offenbar die auf ganz \mathbb{R} definierten Lösungen

$$y = \cos x \quad \text{und} \quad y = \sin x.$$

Allgemeiner ist mit beliebigen Konstanten $c_0, c_1 \in \mathbb{R}$

$$\varphi_{c_0, c_1}(x) := c_0 \cos x + c_1 \sin x$$

eine Lösung. Da

$$\varphi_{c_0, c_1}(0) = c_0 \quad \text{und} \quad \varphi'_{c_0, c_1}(0) = c_1,$$

ist φ_{c_0, c_1} die eindeutig bestimmte Lösung φ der Differentialgleichung (7) mit den Anfangsbedingungen $\varphi(0) = c_0$ und $\varphi'(0) = c_1$. Andererseits hat jede Lösung $\varphi: \mathbb{R} \rightarrow \mathbb{R}$ der Differentialgleichung an der Stelle $x = 0$ einen gewissen Funktionswert und Wert der Ableitung. Daraus folgt, dass die Gesamtheit aller Lösungen der Differentialgleichung $y'' + y = 0$ durch die Funktionen φ_{c_0, c_1} gegeben wird, wobei c_0 und c_1 ganz \mathbb{R} durchlaufen.

AUFGABEN

12.1. Sei $f: \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$ eine stetige Funktion, die lokal einer Lipschitz-Bedingung genüge. Es gelte

$$f(-x, y) = -f(x, y) \quad \text{für alle } (x, y) \in \mathbb{R}^2.$$

Man beweise: Jede Lösung

$$\varphi:]-r, r[\rightarrow \mathbb{R}, \quad (r > 0),$$

der Differentialgleichung $y' = f(x, y)$ geht bei Spiegelung an der y -Achse in sich über.

12.2. Mit Hilfe des Picard-Lindelöfschen Iterationsverfahrens berechne man die Lösung des Differentialgleichungssystems

$$\begin{cases} y_1' = y_2, \\ y_2' = y_1 \end{cases}$$

mit der Anfangsbedingung $\varphi(0) = \begin{pmatrix} a \\ b \end{pmatrix} \in \mathbb{C}^2$.

12.3. Sei $I \subset \mathbb{R}$ ein Intervall und

$$f: I \times \mathbb{R}^n \longrightarrow \mathbb{R}^n, \quad (x, y) \mapsto f(x, y)$$

eine stetige Funktion, die in $I \times \mathbb{R}^n$ einer globalen Lipschitz-Bedingung mit der Konstanten $L \in \mathbb{R}_+$ genügt. Man beweise:

a) Zu jedem Punkt $(a, c) \in I \times \mathbb{R}^n$ gibt es eine auf dem ganzen Intervall I definierte Lösung $\varphi: I \rightarrow \mathbb{R}^n$ der Differentialgleichung

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

mit der Anfangsbedingung $\varphi(a) = c$.

b) Seien $\varphi, \psi: I \rightarrow \mathbb{R}^n$ zwei Lösungen der Differentialgleichung $y' = f(x, y)$. Für ein $a \in I$ sei $\delta := \|\varphi(a) - \psi(a)\|$. Dann gilt

$$\|\varphi(x) - \psi(x)\| \leq \delta e^{L|x-a|} \quad \text{für alle } x \in I.$$

§ 13 Lineare Differentialgleichungen

Wir behandeln jetzt die allgemeine Theorie der linearen Differentialgleichungen. Die Aussagen gleichen in mancher Hinsicht denen aus der Theorie der linearen Gleichungssysteme in der Linearen Algebra. So bilden die Lösungen einer homogenen linearen Differentialgleichung einen Vektorraum. Man erhält die allgemeine Lösung einer inhomogenen linearen Differentialgleichung, indem man zu einer speziellen Lösung der inhomogenen Differentialgleichung die allgemeine Lösung der zugeordneten homogenen Differentialgleichung addiert.

Definition. Sei $I \subset \mathbb{R}$ ein Intervall und

$$A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix} : I \longrightarrow M(n \times n, \mathbb{R})$$

eine stetige Abbildung (d.h. alle Funktionen $a_{ij}: I \rightarrow \mathbb{R}$ seien stetig). Dann heißt

$$y' = A(x)y \tag{1}$$

ein *homogenes lineares Differentialgleichungssystem* (oder einfach *homogene lineare Differentialgleichung*). Weiter sei

$$b = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} : I \longrightarrow \mathbb{R}^n$$

eine stetige vektorwertige Funktion. Dann heißt

$$y' = A(x)y + b(x) \tag{2}$$

ein *inhomogenes lineares Differentialgleichungssystem*. Das Differentialgleichungssystem (1) heißt das dem System (2) zugeordnete homogene lineare Differentialgleichungssystem.

Es ist im Hinblick auf spätere Betrachtungen zweckmäßig, neben den reellen linearen Differentialgleichungen auch komplexe Differentialgleichungen zu betrachten. Ein komplex-lineares Differentialgleichungssystem hat die Gestalt

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

wobei $A: I \rightarrow M(n \times n, \mathbb{C})$ und $b: I \rightarrow \mathbb{C}^n$ stetige Abbildungen auf dem Intervall $I \subset \mathbb{R}$ sind. Eine Lösung davon ist eine differenzierbare Abbildung $\varphi: I \rightarrow \mathbb{C}^n$, so dass

$$\varphi'(x) = A(x)\varphi(x) + b(x) \quad \text{für alle } x \in I.$$

Da \mathbb{C} mit $\mathbb{R} \times \mathbb{R}$ identifiziert werden kann, kann φ als Abbildung $I \rightarrow \mathbb{R}^{2n}$ aufgefasst werden; die Differenzierbarkeit ist komponentenweise zu verstehen. Das System von n komplexen Differentialgleichungen ist äquivalent zu einem System von $2n$ reellen Differentialgleichungen.

Man beachte, dass wir zwar komplexe Werte zulassen, die unabhängige Variable x aber immer reell ist.

Bezeichnung. Um den reellen und komplexen Fall simultan behandeln zu können, bezeichne im Folgenden \mathbb{K} einen der Körper \mathbb{R} oder \mathbb{C} .

Satz 1 (Existenz und Eindeutigkeit). *Sei $I \subset \mathbb{R}$ ein Intervall und seien*

$$A: I \longrightarrow M(n \times n, \mathbb{K}) \quad \text{und} \quad b: I \longrightarrow \mathbb{K}^n$$

stetige Abbildungen. Dann gibt es zu jedem $x_0 \in I$ und $c \in \mathbb{K}^n$ genau eine Lösung

$$\varphi: I \longrightarrow \mathbb{K}^n$$

der linearen Differentialgleichung

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

mit der Anfangsbedingung $\varphi(x_0) = c$.

Bemerkung. Man beachte, dass die Lösungen einer linearen Differentialgleichung immer über dem ganzen Intervall I existieren, im Gegensatz zum Fall nicht-linearer Differentialgleichungen $y' = f(x, y)$, vgl. Beispiel (12.2).

Beweis von Satz 1. Wir setzen

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

und zeigen zunächst: Ist $J \subset I$ irgend ein kompaktes Teilintervall, so genügt f in $J \times \mathbb{K}^n$ einer globalen Lipschitz-Bedingung. Dann aus der Stetigkeit von A folgt

$$L := \sup \{ \|A(x)\| : x \in J \} < \infty.$$

Damit erhält man für $x \in J$ und $y, \tilde{y} \in \mathbb{K}^n$

$$\|f(x, y) - f(x, \tilde{y})\| = \|A(x)(y - \tilde{y})\| \leq L\|y - \tilde{y}\|.$$

Daraus folgt nach § 12, Satz 3, zunächst die Eindeutigkeit der Lösung. Zum Beweis der Existenz benützen wir das Picard-Lindelöfsche Iterationsverfahren: Wir definieren die Folge von Funktionen $\varphi_k: I \rightarrow \mathbb{K}^n$ durch

$$\begin{aligned}\varphi_0(x) &:= c, \\ \varphi_{k+1}(x) &:= c + \int_{x_0}^x f(t, \varphi_k(t)) dt.\end{aligned}$$

Behauptung. Die Funktionenfolge $(\varphi_k)_{k \in \mathbb{N}}$ konvergiert auf jedem kompakten Teilintervall $J \subset I$ mit $x_0 \in I$ gleichmäßig gegen eine Lösung φ der Differentialgleichung.

Beweis. Es ist klar, dass alle φ_k auf ganz I definiert und dort stetig sind. Da J kompakt ist, ist

$$K := \sup \{ \|\varphi_1(x) - \varphi_0(x)\| : x \in J \} < \infty.$$

Daraus folgt für alle $k \in \mathbb{N}$ und $x \in J$

$$\|\varphi_{k+1}(x) - \varphi_k(x)\| \leq K \frac{L^k |x - x_0|^k}{k!}. \quad (3)$$

Dies zeigt man durch vollständige Induktion nach k . Der Induktionsanfang $k = 0$ gilt nach Definition von K .

Induktionsschritt $k \rightarrow k + 1$. Es gilt

$$\varphi_{k+2}(x) - \varphi_{k+1}(x) = \int_{x_0}^x \left(f(t, \varphi_{k+1}(t)) - f(t, \varphi_k(t)) \right) dt.$$

Dies liefert die Abschätzung

$$\begin{aligned}\|\varphi_{k+2}(x) - \varphi_{k+1}(x)\| &\leq L \left| \int_{x_0}^x \|\varphi_{k+1}(t) - \varphi_k(t)\| dt \right| \\ &\leq KL^{k+1} \left| \int_{x_0}^x \frac{|t - x_0|^k}{k!} dt \right| = K \frac{L^{k+1} |x - x_0|^{k+1}}{(k+1)!}.\end{aligned}$$

Damit ist (3) bewiesen. Wir setzen

$$r := \sup \{ |x - x_0| : x \in J \} < \infty.$$

Es folgt, dass die unendliche Reihe $\sum_{k=0}^{\infty} (\varphi_{k+1} - \varphi_k)$ auf J durch die konvergente Reihe

$$K \sum_{k=0}^{\infty} \frac{L^k J^k}{k!} = K e^{Lr}$$

majorisiert wird, also dort gleichmäßig konvergiert. Daher ist die Funktion

$$\varphi := \lim_{k \rightarrow \infty} \varphi_k = \varphi_0 + \sum_{k=0}^{\infty} (\varphi_{k+1} - \varphi_k)$$

auf ganz I stetig und die Konvergenz ist auf jedem kompakten Teilintervall von I gleichmäßig. Für die Limes-Funktion gilt daher die Integralgleichung

$$\varphi(x) = c + \int_{x_0}^x f(t, \varphi(t)) dt \quad \text{für alle } x \in I,$$

d.h. φ ist eine Lösung der gegebenen Differentialgleichung mit der Anfangsbedingung $\varphi(x_0) = c$, q.e.d.

Homogene Gleichungen

Wie bei linearen Gleichungssystemen ist es auch bei linearen Differentialgleichungen zweckmäßig, sich zunächst mit den Lösungen der homogenen Gleichung zu beschäftigen.

Satz 2. Sei $I \subset \mathbb{R}$ ein nicht-triviales Intervall und

$$A : I \rightarrow M(n \times n, \mathbb{K})$$

eine stetige matrixwertige Funktion. Wir bezeichnen mit L_H die Menge aller Lösungen der homogenen linearen Differentialgleichung

$$y' = A(x)y.$$

Dann ist L_H ein n -dimensionaler Vektorraum über \mathbb{K} .

Für ein k -tupel von Lösungen $\varphi_1, \dots, \varphi_k \in L_H$ sind folgende Aussagen gleichbedeutend:

- i) Die Funktionen $\varphi_1, \dots, \varphi_k$ sind linear unabhängig über \mathbb{K} .
- ii) Es existiert ein $x_0 \in I$, so dass die Vektoren $\varphi_1(x_0), \dots, \varphi_k(x_0) \in \mathbb{K}^n$ linear unabhängig sind.

iii) Für jedes $x_0 \in I$ sind die Vektoren $\varphi_1(x_0), \dots, \varphi_k(x_0) \in \mathbb{K}^n$ linear unabhängig.

Beweis. a) Wir beweisen zunächst, dass L_H ein Untervektorraum des Vektorraums aller Abbildungen $f: I \rightarrow \mathbb{K}^n$ ist. Dazu ist dreierlei zu zeigen:

1. Trivialerweise gilt $0 \in L_H$.
2. Seien $\varphi, \psi \in L_H$, also $\varphi' = A\varphi$ und $\psi' = A\psi$. Dann ist auch $\varphi + \psi$ differenzierbar mit

$$(\varphi + \psi)' = \varphi' + \psi' = A\varphi + A\psi = A(\varphi + \psi),$$

d.h. $\varphi + \psi \in L_H$.

3. Sei $\varphi \in L_H$ und $\lambda \in \mathbb{K}$. Dann ist

$$(\lambda\varphi)' = \lambda\varphi' = \lambda A\varphi = A(\lambda\varphi),$$

d.h. $\lambda\varphi \in L_H$.

b) Wir zeigen jetzt die Äquivalenz der Aussagen i) bis iii). Die Implikationen $\text{iii}) \Rightarrow \text{ii}) \Rightarrow \text{i})$ sind trivial. Es ist daher nur noch die Implikation $\text{i}) \Rightarrow \text{iii})$ zu beweisen.

Seien also $\varphi_1, \dots, \varphi_k \in L_H$ linear unabhängig und $x_0 \in I$. Angenommen, die Vektoren $\varphi_1(x_0), \dots, \varphi_k(x_0)$ wären linear abhängig. Dann gäbe es Konstanten $\lambda_1, \dots, \lambda_k \in \mathbb{K}$, die nicht alle gleich null sind, so dass

$$\lambda_1\varphi_1(x_0) + \dots + \lambda_k\varphi_k(x_0) = 0.$$

Wir betrachten die Lösung

$$\varphi := \lambda_1\varphi_1 + \dots + \lambda_k\varphi_k \in L_H.$$

Es gilt $\varphi(x_0) = 0$. Wegen der Eindeutigkeit der Lösung zu gegebener Anfangsbedingung muss φ identisch null sein. Das heißt aber, dass $\varphi_1, \dots, \varphi_k$ linear abhängig sind, Widerspruch!

c) Wir zeigen jetzt $\dim L_H = n$.

Seien e_1, \dots, e_n die kanonischen Einheitsvektoren des \mathbb{K}^n und $x_0 \in I$. Nach Satz 1 gibt es Lösungen $\varphi_1, \dots, \varphi_n \in L_H$ mit $\varphi_j(x_0) = e_j$. Diese Lösungen sind linear unabhängig, da sie an der Stelle x_0 unabhängig sind. Also gilt $\dim L_H \geq n$.

Andrerseits ist $\dim L_H \leq n$, denn sonst gäbe es $n+1$ linear unabhängige Lösungen $\psi_1, \dots, \psi_{n+1}$ und wegen iii) müssten die Vektoren

$$\psi_1(x_0), \dots, \psi_{n+1}(x_0) \in \mathbb{K}^n$$

linear unabhängig sein, was offensichtlich unmöglich ist.

Damit ist Satz 2 vollständig bewiesen.

Um alle Elemente eines Vektorraums zu kennen, genügt es, eine Basis des Vektorraums zu kennen. Dies gibt Anlass zu folgender Begriffsbildung.

Definition. Unter einem *Lösungs-Fundamentalsystem* der Differentialgleichung $y' = A(x)y$ versteht man eine Basis $(\varphi_1, \dots, \varphi_n)$ des Vektorraums seiner Lösungen.

Schreibt man die Lösungen φ_k als Spaltenvektoren,

$$\varphi_k = \begin{pmatrix} \varphi_{1k} \\ \varphi_{2k} \\ \vdots \\ \varphi_{nk} \end{pmatrix}, \quad k = 1, \dots, n,$$

so ist $\Phi := (\varphi_1, \dots, \varphi_n)$ eine $n \times n$ -Matrix

$$\Phi = \begin{pmatrix} \varphi_{11} & \dots & \varphi_{1n} \\ \vdots & & \vdots \\ \varphi_{n1} & \dots & \varphi_{nn} \end{pmatrix}.$$

Nach Satz 2 sind Lösungen $\varphi_1, \dots, \varphi_n$ genau dann linear unabhängig, wenn für die Matrix $\Phi = (\varphi_1, \dots, \varphi_n)$ gilt

$$\det \Phi(x_0) \neq 0$$

für wenigstens ein $x_0 \in I$. Es gilt dann $\det \Phi(x) \neq 0$ für alle $x \in I$.

Ist $\Phi = (\varphi_1, \dots, \varphi_n)$ ein Lösungs-Fundamentalsystem der Differentialgleichung $y' = A(x)y$, so lässt sich eine beliebige Lösung φ schreiben als

$$\varphi = c_1 \varphi_1 + \dots + c_n \varphi_n$$

mit Konstanten $c_k \in \mathbb{K}$. Man kann dies in Matrizen-Schreibweise zusammenfassen zu

$$\varphi = \Phi c,$$

wobei

$$c = \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix} \in \mathbb{K}^n.$$

Man kann ein Lösungs-Fundamentalsystem $\Phi = (\varphi_1, \dots, \varphi_n)$ selbst als matrixwertige Lösung der Differentialgleichung auffassen, denn

$$\begin{aligned} \Phi' &= (\varphi_1', \dots, \varphi_n'), \\ A\Phi &= (A\varphi_1, \dots, A\varphi_n). \end{aligned}$$

Also gilt $\Phi' = A\Phi$.

(13.1) Beispiel. Wir betrachten das Differentialgleichungssystem

$$\begin{cases} y_1' = -\omega y_2, \\ y_2' = \omega y_1, \end{cases}$$

wobei $\omega \in \mathbb{R}$ eine Konstante ist. Dieses System lautet in Matrizenschreibweise

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix}' = \begin{pmatrix} 0 & -\omega \\ \omega & 0 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}.$$

Man sieht unmittelbar, dass folgende Funktionen $\varphi_k: \mathbb{R} \rightarrow \mathbb{R}^2$ Lösungen sind:

$$\varphi_1(x) := \begin{pmatrix} \cos \omega x \\ \sin \omega x \end{pmatrix}, \quad \varphi_2(x) := \begin{pmatrix} -\sin \omega x \\ \cos \omega x \end{pmatrix}.$$

Die Lösungen sind linear unabhängig, denn für die Matrix

$$\Phi(x) = (\varphi_1(x), \varphi_2(x)) = \begin{pmatrix} \cos \omega x & -\sin \omega x \\ \sin \omega x & \cos \omega x \end{pmatrix}$$

gilt $\det \Phi(x) = 1$ für alle $x \in \mathbb{R}$.

Inhomogene Gleichungen

Wir kommen jetzt zu den inhomogenen Gleichungen. Der folgende Satz zeigt den Zusammenhang zwischen den Lösungen der inhomogenen Gleichung und der zugeordneten homogenen Gleichung.

Satz 3. Sei $I \subset \mathbb{R}$ ein Intervall und seien

$$A: I \longrightarrow M(n \times n, \mathbb{K}) \quad \text{und} \quad b: I \longrightarrow \mathbb{K}^n$$

stetige Abbildungen. Wir bezeichnen mit L_H den Vektorraum aller Lösungen $\varphi: I \rightarrow \mathbb{K}^n$ der homogenen Differentialgleichung

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

und mit L_I die Menge aller Lösungen $\psi: I \rightarrow \mathbb{K}^n$ der inhomogenen Differentialgleichung

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

Dann gilt für beliebiges $\psi_0 \in L_I$

$$L_I = \psi_0 + L_H.$$

Mit anderen Worten: Man erhält die allgemeine Lösung der inhomogenen Gleichung als Summe einer speziellen Lösung der inhomogenen Gleichung und der allgemeinen Lösung der homogenen Gleichung.

Beweis. a) Wir zeigen zunächst $L_I \subset \psi_0 + L_H$.

Sei $\psi \in L_I$ beliebig vorgegeben. Wir setzen $\varphi := \psi - \psi_0$. Dann gilt

$$\varphi' = \psi' - \psi_0' = (A\psi + b) - (A\psi_0 + b) = A(\psi - \psi_0) = A\varphi,$$

d.h. $\varphi \in L_H$. Da $\psi = \psi_0 + \varphi$, folgt $\psi \in \psi_0 + L_H$.

b) Wir zeigen jetzt $\psi_0 + L_H \subset L_I$.

Sei $\psi \in \psi_0 + L_H$, d.h. $\psi = \psi_0 + \varphi$ mit $\varphi \in L_H$. Dann gilt

$$\psi' = \psi_0' + \varphi' = (A\psi_0 + b) + A\varphi = A\psi + b,$$

d.h. $\psi \in L_I$, q.e.d.

Um eine inhomogene lineare Differentialgleichung vollständig zu lösen ist also neben der Lösung der homogenen Gleichung nur die Kenntnis einer einzigen Lösung der inhomogenen Gleichung notwendig. Eine solche kann man sich durch die Methode der Variation der Konstanten beschaffen.

Satz 4 (Variation der Konstanten). *Mit den Bezeichnungen von Satz 3 gilt:*

Sei $\Phi = (\varphi_1, \dots, \varphi_n)$ ein Lösungs-Fundamentalsystem der homogenen Differentialgleichung

$$y' = A(x)y.$$

Dann erhält man eine Lösung $\psi: I \rightarrow \mathbb{K}^n$ der inhomogenen Differentialgleichung

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

durch den Ansatz

$$\psi(x) = \Phi(x)u(x).$$

Dabei ist $u: I \rightarrow \mathbb{K}^n$ eine differenzierbare Funktion mit $\Phi(x)u'(x) = b(x)$, d.h.

$$u(x) = \int_{x_0}^x \Phi(t)^{-1} b(t) dt + \text{const.}$$

Beweis. Aus $\psi = \Phi u$ folgt

$$\psi' = \Phi' u + \Phi u',$$

$$A\psi + b = A\Phi u + b.$$

Da $\Phi' = A\Phi$, gilt $\psi' = A\psi + b$ genau dann, wenn $\Phi u' = b$, q.e.d.

(13.2) Beispiel. Wir behandeln das Differentialgleichungssystem

$$\begin{cases} y_1' = -y_2, \\ y_2' = y_1 + x, \end{cases} \quad (*)$$

oder in Matrizen-Schreibweise

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix}' = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} + \begin{pmatrix} 0 \\ x \end{pmatrix}.$$

Nach Beispiel (13.1) ist

$$\Phi(x) = \begin{pmatrix} \cos x & -\sin x \\ \sin x & \cos x \end{pmatrix}$$

ein Lösungs-Fundamentalsystem des homogenen Systems. Die inverse Matrix ist

$$\Phi(x)^{-1} = \begin{pmatrix} \cos x & \sin x \\ -\sin x & \cos x \end{pmatrix},$$

also

$$\Phi(x)^{-1} b(x) = \begin{pmatrix} \cos x & \sin x \\ -\sin x & \cos x \end{pmatrix} \begin{pmatrix} 0 \\ x \end{pmatrix} = \begin{pmatrix} x \sin x \\ x \cos x \end{pmatrix}.$$

Damit ergibt sich

$$u(x) = \int_0^x \begin{pmatrix} t \sin t \\ t \cos t \end{pmatrix} dt + \text{const.}$$

Nun berechnet man mittels partieller Integration

$$\int x \sin x dx = -x \cos x + \sin x,$$

$$\int x \cos x dx = x \sin x + \cos x,$$

man kann also

$$u(x) = \begin{pmatrix} -x \cos x + \sin x \\ x \sin x + \cos x \end{pmatrix}$$

wählen. Damit ergibt sich eine spezielle Lösung der inhomogenen Gleichung als

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

Dass ψ die Differentialgleichung (*) löst, kann man auch sofort direkt verifizieren. Die allgemeine Lösung von (*) ist daher

$$\phi(x) = \begin{pmatrix} -x \\ 1 \end{pmatrix} + c_1 \begin{pmatrix} \cos x \\ \sin x \end{pmatrix} + c_2 \begin{pmatrix} -\sin x \\ \cos x \end{pmatrix}$$

mit beliebigen Konstanten $c_1, c_2 \in \mathbb{K}$.

Lineare Differentialgleichungen n -ter Ordnung

Wir übertragen jetzt die bewiesenen Resultate über lineare Differentialgleichungssysteme 1. Ordnung auf lineare Differentialgleichungen n -ter Ordnung.

Sei $I \subset \mathbb{R}$ ein Intervall und seien $a_k: I \rightarrow \mathbb{K}$, $0 \leq k \leq n-1$, stetige Funktionen. Dann heißt

$$y^{(n)} + a_{n-1}(x)y^{(n-1)} + \dots + a_1(x)y' + a_0(x)y = 0 \quad (4)$$

homogene lineare Differentialgleichung n -ter Ordnung. Ist $b: I \rightarrow \mathbb{K}$ eine weitere stetige Funktion, so heißt

$$y^{(n)} + a_{n-1}(x)y^{(n-1)} + \dots + a_1(x)y' + a_0(x)y = b(x) \quad (5)$$

inhomogene lineare Differentialgleichung n -ter Ordnung. Die Differentialgleichung (4) heißt die (5) zugeordnete homogene Gleichung.

Satz 5. Die obigen Bezeichnungen seien beibehalten.

- a) Sei L_H die Menge aller Lösungen $\varphi: I \rightarrow \mathbb{K}$ der homogenen linearen Differentialgleichung n -ter Ordnung (4). Dann ist L_H ein n -dimensionaler Vektorraum über \mathbb{K} .
- b) Sei L_I die Menge aller Lösungen $\psi: I \rightarrow \mathbb{K}$ der inhomogenen Differentialgleichung (5). Dann gilt für ein beliebiges $\psi_0 \in L_I$

$$L_I = \psi_0 + L_H.$$

- c) Ein n -tupel $\varphi_1, \dots, \varphi_n \in L_H$ von Lösungen der homogenen Gleichung (4) ist genau dann linear unabhängig, wenn für ein und damit für alle $x \in I$ die "Wronski-Determinante"

$$W(x) := \det \begin{pmatrix} \varphi_1(x) & \varphi_2(x) & \dots & \varphi_n(x) \\ \varphi_1'(x) & \varphi_2'(x) & \dots & \varphi_n'(x) \\ \vdots & \vdots & & \vdots \\ \varphi_1^{(n-1)}(x) & \varphi_2^{(n-1)}(x) & \dots & \varphi_n^{(n-1)}(x) \end{pmatrix}$$

von Null verschieden ist.

Beweis. Die Differentialgleichung (5) ist äquivalent mit dem inhomogenen linearen Differentialgleichungssystem 1. Ordnung (vgl. § 12):

$$\begin{cases} y_0' = y_1, \\ y_1' = y_2, \\ \vdots \\ y_{n-2}' = y_{n-1}, \\ y_{n-1}' = -a_0(x)y_0 - a_1(x)y_1 - \dots - a_{n-1}(x)y_{n-1} + b(x). \end{cases} \quad (6)$$

Jeder Lösung $\varphi: I \rightarrow \mathbb{K}$ von (5) entspricht eine Lösung

$$f = \begin{pmatrix} \varphi \\ \varphi' \\ \vdots \\ \varphi^{(n-1)} \end{pmatrix} : I \longrightarrow \mathbb{K}^n$$

des Systems (6) und umgekehrt. Entsprechendes gilt für die homogene Gleichung ($b = 0$). Damit folgen die Behauptungen aus den Sätzen 2 und 3.

(13.3) Beispiel. Die Differentialgleichung

$$y'' - \frac{1}{2x}y' + \frac{1}{2x^2}y = 0$$

auf dem Intervall $I = \mathbb{R}_+^*$ besitzt die Lösungen

$$\varphi_1(x) := x, \quad \varphi_2(x) := \sqrt{x},$$

wie man sich durch Einsetzen überzeugt. Die Wronski-Determinante von φ_1, φ_2 ist

$$W(x) = \det \begin{pmatrix} \varphi_1(x) & \varphi_2(x) \\ \varphi_1'(x) & \varphi_2'(x) \end{pmatrix} = \det \begin{pmatrix} x & \sqrt{x} \\ 1 & \frac{1}{2\sqrt{x}} \end{pmatrix} = -\frac{\sqrt{x}}{2}.$$

Dies ist eine auf \mathbb{R}_+^* nicht-verschwindende Funktion; φ_1 und φ_2 bilden also ein Lösungs-Fundamentalsystem. Die allgemeine Lösung hat daher die Gestalt

$$\varphi(x) = c_1x + c_2\sqrt{x}$$

mit beliebigen reellen (bzw. komplexen) Konstanten c_1 und c_2 .

Um die inhomogene Differentialgleichung

$$y'' - \frac{1}{2x}y' + \frac{1}{2x^2}y = 1$$

vollständig zu lösen, brauchen wir nur noch eine spezielle Lösung. In diesem einfachen Fall sieht man sofort, dass die linke Seite angewandt auf die Funktion $y = x^2$ eine Konstante ergibt:

$$\left(\frac{d^2}{dx^2} - \frac{1}{2x} \frac{d}{dx} + \frac{1}{2x^2} \right) x^2 = 2 - 1 + \frac{1}{2} = \frac{3}{2}.$$

Daher ist $\psi_0(x) := \frac{2}{3}x^2$ eine spezielle Lösung der inhomogenen Gleichung; die allgemeine Lösung ist

$$\psi(x) = \frac{2}{3}x^2 + c_1x + c_2\sqrt{x}, \quad c_1, c_2 \in \mathbb{K}.$$

AUFGABEN

13.1. Man bestimme alle Lösungen des Differentialgleichungssystems

$$\begin{cases} y_1' = y_2 + 1, \\ y_2' = y_1 + \sin x. \end{cases}$$

13.2. Sei $I \subset \mathbb{R}$ ein Intervall und

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} : I \longrightarrow M(2 \times 2, \mathbb{K})$$

eine stetige Abbildung. Die Differentialgleichung

$$y' = Ay$$

besitze die spezielle Lösung $\varphi = \begin{pmatrix} \varphi_1 \\ \varphi_2 \end{pmatrix} : I \rightarrow \mathbb{K}^2$. Im Teilintervall $J \subset I$ gelte $\varphi_1(x) \neq 0$. Man zeige: Man erhält eine zweite, von φ linear unabhängige Lösung $\psi : J \rightarrow \mathbb{K}^2$ durch den Ansatz

$$\psi(x) = u(x) \begin{pmatrix} \varphi_1(x) \\ \varphi_2(x) \end{pmatrix} + \begin{pmatrix} 0 \\ g(x) \end{pmatrix},$$

wobei $u, g : J \rightarrow \mathbb{K}$ differenzierbare Funktionen sind, die den folgenden Differentialgleichungen genügen:

$$g' = \left(a_{22} - a_{12} \frac{\varphi_2}{\varphi_1} \right) g,$$

$$u' = \frac{a_{12}}{\varphi_1} g.$$

13.3. Man bestimme alle Lösungen des folgenden Differentialgleichungssystems auf \mathbb{R}_+^* :

$$\begin{cases} y_1' = -y_1 + \frac{1}{x}y_2 + \log x + \frac{1}{x}, \\ y_2' = (1-x)y_1 + y_2 + (x-1)\log x. \end{cases}$$

Hinweis. Eine spezielle Lösung des homogenen Systems ist $\varphi(x) = \begin{pmatrix} 1 \\ x \end{pmatrix}$.

13.4. Gegeben sei die Differentialgleichung n -ter Ordnung

$$y^{(n)} + a_{n-1}(x)y^{(n-1)} + \dots + a_0(x)y = 0, \quad (*)$$

deren Koeffizienten stetige Funktionen $a_i: I \rightarrow \mathbb{K}$ auf einem Intervall $I \subset \mathbb{R}$ seien.

Man beweise: Die Wronski-Determinante $W: I \rightarrow \mathbb{K}$ eines Fundamentalsystems von Lösungen von (*) genügt der Differentialgleichung

$$W'(x) + a_{n-1}(x)W(x) = 0.$$

Anleitung. Man beweise dazu folgende Regel für die Differentiation einer Determinante:

Sei $\Phi = (\varphi_{ij})$ eine $n \times n$ -Matrix, deren Koeffizienten differenzierbare Funktionen $\varphi_{ij}: I \rightarrow \mathbb{K}$ sind. Dann gilt

$$\frac{d}{dx} \det \Phi(x) = \sum_{i=1}^n \det \begin{pmatrix} \varphi_{11}(x) & \dots & \varphi_{1n}(x) \\ \vdots & \dots & \vdots \\ \varphi'_{i1}(x) & \dots & \varphi'_{in}(x) \\ \vdots & \dots & \vdots \\ \varphi_{n1}(x) & \dots & \varphi_{nn}(x) \end{pmatrix},$$

wobei im i -ten Summanden nur die i -te Zeile der Matrix differenziert wird.

13.5. Sei $I \subset \mathbb{R}$ ein Intervall und $A: I \rightarrow M(n \times n, \mathbb{K})$ eine matrixwertige Funktion, deren Koeffizienten beliebig oft differenzierbar seien. Man zeige, dass alle Lösungen $\varphi: I \rightarrow \mathbb{R}^n$ der Differentialgleichung

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

beliebig oft differenzierbar sind.

13.6. Sei $r > 0$ und $I :=]-r, r[\subset \mathbb{R}$. Weiter seien $a, b: I \rightarrow \mathbb{R}$ zwei stetige Funktionen. Die Funktion a sei ungerade und b gerade, d.h.

$$a(-x) = -a(x), \quad b(-x) = b(x)$$

für alle $x \in I$. Man zeige: Die Differentialgleichung

$$y'' + a(x)y' + b(x)y = 0$$

besitzt ein Fundamentalsystem von Lösungen, das aus einer geraden und einer ungeraden Funktion besteht.

§ 14 Differentialgleichungen 2. Ordnung

In diesem Paragraphen studieren wir einige spezielle Differentialgleichungen 2. Ordnung, die in der theoretischen Physik eine Rolle spielen. Wir behandeln u.a. die eindimensionale Bewegung in einem Potential, die gedämpfte Schwingung und die Besselsche Differentialgleichung.

Eindimensionale Bewegung

Es sei $J \subset \mathbb{R}$ ein Intervall und

$$f : J \longrightarrow \mathbb{R}, \quad x \mapsto f(x)$$

eine stetige Funktion. Die Differentialgleichung

$$\frac{d^2x}{dt^2} = f(x), \quad (t, x) \in \mathbb{R} \times J,$$

lässt sich physikalisch interpretieren als die Differentialgleichung eines Masseiteilchens (der Masse 1) mit einem Freiheitsgrad, das sich unter dem Einfluss einer nur vom Ort $x \in J$ abhängigen Kraft $f(x)$ bewegt. Dabei bedeutet t die Zeitvariable. Die Ableitung nach der Zeit bezeichnen wir auch durch einen Punkt über der Funktionsvariablen. So ist

$$v(t) := \dot{x}(t) = \frac{dx(t)}{dt}$$

die Geschwindigkeit und $\ddot{x}(t) = d^2x(t)/dt^2$ die Beschleunigung zum Zeitpunkt t . Wir definieren eine Funktion $U : J \rightarrow \mathbb{R}$ durch

$$U(x) := - \int_a^x f(\xi) d\xi,$$

wobei $a \in J$ ein beliebiger Punkt ist. U hat die physikalische Bedeutung einer potentiellen Energie. Die Differentialgleichung geht dann über in

$$\frac{d^2x}{dt^2} = - \frac{dU}{dx}(x). \quad (1)$$

Sei $x = x(t)$ eine Lösung von (1). Wie in § 10, Abschnitt (10.5), erhält man aus (1) durch Multiplikation mit $\dot{x}(t)$

$$\frac{d}{dt} \left(\frac{1}{2} \dot{x}(t)^2 + U(x(t)) \right) = 0.$$

Es gibt also eine Konstante $E \in \mathbb{R}$ (physikalische Bedeutung Gesamt-Energie), so dass

$$\frac{1}{2} v(t)^2 + U(x(t)) = E \quad \text{für alle } t.$$

Da v^2 nicht-negativ ist, sieht man daran schon, dass die Bewegung nur in Bereichen erfolgen kann, in denen $U(x) \leq E$ ist. Die Bewegung genügt dann einer der beiden Differentialgleichungen

$$\dot{x} = \sqrt{2(E - U(x))} \tag{2}$$

oder

$$\dot{x} = -\sqrt{2(E - U(x))}, \tag{2'}$$

je nachdem im betrachteten Zeitintervall $\dot{x}(t) \geq 0$ oder $\dot{x}(t) \leq 0$ ist. Diese Differentialgleichungen gehören zum Typ der getrennten Variablen, wir können also § 11, Satz 1 anwenden. Sei $J_0 \subset J$ ein Intervall mit $U(\xi) < E$ für alle $\xi \in J_0$ und $x_0 \in J_0$ ein fester Punkt. Für $x \in J_0$ definieren wir

$$G(x) := \int_{x_0}^x \frac{d\xi}{\sqrt{2(E - U(\xi))}}.$$

Für die Lösung der Differentialgleichung (2) mit der Anfangsbedingung $x(t_0) = x_0$ gilt dann

$$G(x(t)) = t - t_0, \tag{3}$$

oder

$$x(t) = H(t - t_0),$$

wobei H die Umkehrfunktion von G ist. Die Formel (3) lässt folgende Interpretation zu: Das Integral $G(x)$ gibt die Zeit an, die das Teilchen braucht, um von x_0 nach x zu gelangen.

Betrachten wir genauer folgende spezielle Situation: Es seien $x_A < x_B$ zwei

Punkte mit

$$U(x_A) = U(x_B) = E \quad \text{und} \quad U(\xi) < E \quad \text{für} \quad x_A < \xi < x_B.$$

Ferner sei vorausgesetzt, dass $U'(x_A) \neq 0$ und $U'(x_B) \neq 0$, d.h. $f(x_A) \neq 0$ und $f(x_B) \neq 0$, vgl. Bild 14.1.

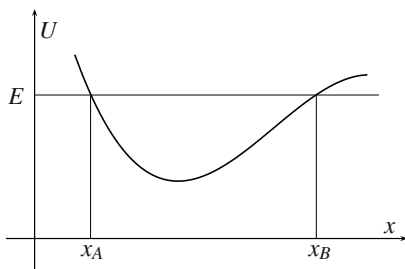


Bild 14.1

Die Bewegung verläuft dann ganz im Intervall $[x_A, x_B]$. Da das uneigentliche Integral

$$\int_0^1 \frac{dx}{\sqrt{x}}$$

existiert (vgl. An. 1, Beispiel 20.2), existiert auch das uneigentliche Integral

$$T := \int_{x_A}^{x_B} \frac{dx}{\sqrt{2(E - U(x))}} \quad (4)$$

und stellt die Zeit dar, die das Teilchen von x_A nach x_B braucht. Im Punkt x_B ist die Geschwindigkeit null und ändert ihr Vorzeichen, denn wir haben vorausgesetzt, dass

$$\frac{dv}{dt}(t_B) = \frac{d^2x}{dt^2}(t_B) = f(x_B) \neq 0 \quad \text{für} \quad x(t_B) = x_B.$$

Das Teilchen läuft anschließend von x_B nach x_A zurück und gehorcht dabei der Differentialgleichung (2'), wobei der Zeitbedarf wieder T ist. Es ergibt sich also eine periodische Bewegung mit der Schwingungsdauer $2T$.

Beispiele

(14.1) Der *harmonische Oszillator* genügt der Differentialgleichung

$$\frac{d^2x}{dt^2} = -kx,$$

wobei k eine positive Konstante ist. Physikalisch beschreibt die Differentialgleichung die eindimensionale Schwingung eines Massenpunktes um den Nullpunkt, wobei die Rückstellkraft proportional zur Auslenkung x ist.

Wir suchen eine Lösung mit der Anfangsbedingung

$$x(t_0) = 0, \quad \dot{x}(t_0) = v_0 > 0.$$

Mit den obigen Bezeichnungen wird

$$U(x) = \int_0^x k\xi d\xi = \frac{k}{2}x^2.$$

Aus der Anfangsbedingung ergibt sich

$$E = \frac{1}{2}\dot{x}(t_0)^2 + U(x(t_0)) = \frac{1}{2}v_0^2.$$

Die Bewegung verläuft daher im Intervall

$$\{x \in \mathbb{R} : U(x) \leq E\} = \{x \in \mathbb{R} : |x| \leq \frac{v_0}{\sqrt{k}}\}.$$

Wir setzen zur Abkürzung

$$A := \frac{v_0}{\sqrt{k}}, \quad \omega := \sqrt{k}.$$

Für $|x| < A$ ist hier

$$G(x) = \int_0^x \frac{d\xi}{\sqrt{2(E - U(\xi))}} = \int_0^x \frac{d\xi}{\sqrt{v_0^2 - k\xi^2}} = \frac{1}{A\omega} \int_0^x \frac{d\xi}{\sqrt{1 - (\xi/A)^2}}.$$

Substituiert man $u = \xi/A$, erhält man weiter

$$G(x) = \frac{1}{\omega} \int_0^{x/A} \frac{du}{\sqrt{1 - u^2}} = \frac{1}{\omega} \arcsin\left(\frac{x}{A}\right),$$

also $\frac{1}{\omega} \arcsin\left(\frac{x(t)}{A}\right) = t - t_0$, d.h.

$$x(t) = A \sin \omega(t - t_0).$$

Diese Beziehung gilt zunächst nur für $|\omega(t - t_0)| < \pi/2$, aber man sieht durch Einsetzen, dass diese Funktion die Differentialgleichung für alle t löst. Die Bewegung ist also eine Schwingung um den Nullpunkt mit Amplitude $A = v_0/\sqrt{k}$ und Schwingungsdauer $2\pi/\omega = 2\pi/\sqrt{k}$.

Bemerkung. Die Differentialgleichung des harmonischen Oszillators lässt sich noch leichter direkt durch einen Ansatz lösen. Siehe den Abschnitt über die gedämpfte Schwingung später in diesem Paragraphen.

(14.2) Das *mathematische Pendel* ist die Idealisierung eines physikalischen Pendels.

Ein Massenpunkt ist an einem (masselosen, inelastischen) Faden aufgehängt und bewegt sich ohne Reibung unter dem Einfluss der Schwerkraft. Die auf den Massenpunkt wirkende Kraft ist

$$K = -mg \sin x,$$

wobei x der Auslenkungswinkel (im Bogenmaß), g die Schwerkraft (Erdbeschleunigung), und m die Masse des Pendels ist, siehe Bild 14.2.

Ist L die Länge des Pendels, so ist die Auslenkung gleich Lx . Die Differentialgleichung lautet also

$$m \frac{d^2}{dt^2}(Lx) = -mg \sin x,$$

d.h.

$$\ddot{x} = -\omega^2 \sin x \quad \text{mit } \omega = \sqrt{g/L}.$$

Für kleine Auslenkungen ist $\sin x \approx x$ und die Differentialgleichung geht in die des harmonischen Oszillators über.

Das Potential für das mathematische Pendel ergibt sich als

$$U(x) = \int_0^x \omega^2 \sin \xi d\xi = \omega^2(1 - \cos x).$$

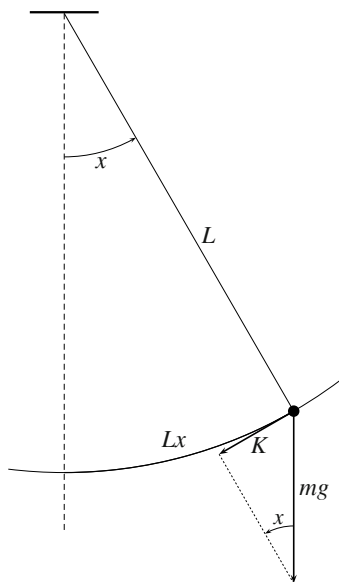


Bild 14.2

Ist a der maximale Ausschlag des Pendels, ($0 < a < \pi/2$), so ist die Gesamtenergie gleich

$$E = U(a) = \omega^2(1 - \cos a),$$

denn im Umkehrpunkt ist die kinetische Energie gleich 0. Nach (4) ist die halbe Periode der Schwingung gleich

$$T = \int_{-a}^a \frac{dx}{\sqrt{2(E - U(x))}} = \frac{2}{\omega} \int_0^a \frac{dx}{\sqrt{2(\cos x - \cos a)}}.$$

Das Integral $\int_0^a \frac{dx}{\sqrt{2(\cos x - \cos a)}}$ ist ein vollständiges elliptisches Integral, es lässt sich wie folgt auf die Normalform

$$E(k) = \int_0^{\pi/2} \frac{dt}{\sqrt{1 - k^2 \sin^2 t}}$$

transformieren (vgl. An.1, Aufgabe 22.5).

Unter Benutzung der Formel $\cos x = 1 - 2 \sin^2(x/2)$ wird

$$\int_0^a \frac{dx}{\sqrt{2(\cos x - \cos a)}} = \frac{1}{2} \int_0^a \frac{dx}{\sqrt{\sin^2(a/2) - \sin^2(x/2)}}.$$

Wir setzen $k := \sin(a/2)$ und führen eine neue Variable t durch

$$\sin(x/2) = k \sin t.$$

ein. Läuft x von 0 bis a , so läuft t von 0 bis $\pi/2$. Durch Differenzieren ergibt sich

$$\begin{aligned} \frac{1}{2} \cos(x/2) dx &= k \cos t dt, \\ \frac{1}{2} \sqrt{1 - \sin^2(x/2)} dx &= k \sqrt{1 - \sin^2 t} dt = \sqrt{k^2 - \sin^2(x/2)} dt, \\ \frac{dx}{\sqrt{k^2 - \sin^2(x/2)}} &= \frac{2 dt}{\sqrt{1 - k^2 \sin^2 t}}, \end{aligned}$$

also

$$\frac{1}{2} \int_0^a \frac{dx}{\sqrt{k^2 - \sin^2(x/2)}} = \int_0^{\pi/2} \frac{dt}{\sqrt{1 - k^2 \sin^2 t}} = E(k).$$

Insgesamt ergibt sich also für die halbe Periode

$$T = \frac{2}{\omega} E(k), \quad k = \sin(a/2), \quad \omega = \sqrt{g/L}.$$

Für $a \rightarrow 0$, d.h. bei kleiner Amplitude, gilt $\lim_{k \rightarrow 0} E(k) = \pi/2$.

Gedämpfte Schwingung

Führt man beim harmonischen Oszillator $\ddot{x} = -\omega_0^2 x$ noch einen Reibungsterm ein, der proportional zur Geschwindigkeit ist, erhält man die Differentialgleichung der gedämpften Schwingung:

$$\ddot{x} + 2\mu\dot{x} + \omega_0^2 x = 0, \quad (\omega_0 \in \mathbb{R}_+^*, \mu \in \mathbb{R}_+).$$

Dabei ist 2μ der Dämpfungsfaktor und ω_0 die Kreisfrequenz der ungedämpften Schwingung.

Die Differentialgleichung ist linear von 2. Ordnung. Um die Gesamtheit aller Lösungen zu kennen, genügt es also, zwei linear unabhängige Lösungen zu kennen. Um eine Lösung φ zu erhalten, machen wir einen “Ansatz”

$$\varphi(t) := e^{\lambda t}$$

mit einer noch zu bestimmenden komplexen Konstanten $\lambda \in \mathbb{C}$. Einsetzen von $x = \varphi(t)$ in die Differentialgleichung liefert

$$\lambda^2 e^{\lambda t} + 2\mu\lambda e^{\lambda t} + \omega_0^2 e^{\lambda t} = (\lambda^2 + 2\mu\lambda + \omega_0^2) e^{\lambda t} = 0,$$

Die Differentialgleichung ist also genau dann erfüllt, wenn λ der “charakteristischen Gleichung”

$$\lambda^2 + 2\mu\lambda + \omega_0^2 = 0$$

genügt. Diese hat die Lösungen

$$\lambda_{1/2} = -\mu \pm \sqrt{\mu^2 - \omega_0^2}.$$

Falls $\mu \neq \omega_0$ (den Fall $\mu = \omega_0$ stellen wir vorläufig zurück), ist $\lambda_1 \neq \lambda_2$ und wir haben die Lösungen

$$\varphi_k(t) := e^{\lambda_k t}, \quad k = 1, 2.$$

Diese Lösungen sind linear unabhängig, bilden also ein Lösungs-Fundamen-

talsystem, denn für die Wronski-Determinante gilt

$$W(0) = \det \begin{pmatrix} \varphi_1(0) & \varphi_2(0) \\ \varphi_1'(0) & \varphi_2'(0) \end{pmatrix} = \det \begin{pmatrix} 1 & 1 \\ \lambda_1 & \lambda_2 \end{pmatrix} = \lambda_2 - \lambda_1 \neq 0.$$

Wir wollen uns nun die Lösungen in Abhängigkeit von μ etwas genauer ansehen.

a) *Ungedämpfte Schwingung*: $\mu = 0$.

Hier ist $\lambda_{1/2} = \pm i\omega_0$, also bilden die beiden Funktionen $\varphi_{\pm}(t) = e^{\pm i\omega_0 t}$ ein Lösungs-Fundamentalsystem. Durch geeignete Linearkombinationen erhält man daraus das reelle Lösungs-Fundamentalsystem

$$\begin{aligned} \varphi_1(t) &= \frac{1}{2}(\varphi_+(t) + \varphi_-(t)) = \cos(\omega_0 t), \\ \varphi_2(t) &= \frac{1}{2i}(\varphi_+(t) - \varphi_-(t)) = \sin(\omega_0 t). \end{aligned}$$

Dass φ_1 und φ_2 linear unabhängig sind, folgt daraus, dass sich umgekehrt φ_+ und φ_- aus φ_1 und φ_2 linear kombinieren lassen.

Für $\mu = 0$ sind also alle Lösungen rein periodisch mit der Periode $2\pi/\omega_0$.

b) *Schwache Dämpfung*: $0 < \mu < \omega_0$.

Wir setzen $\omega := \sqrt{\omega_0^2 - \mu^2}$. Es gilt $0 < \omega < \omega_0$ und

$$\lambda_{1/2} = -\mu \pm i\omega.$$

Ein Lösungs-Fundamentalsystem bilden die Funktionen $\varphi_{\pm}(t) = e^{-\mu t} e^{\pm i\omega t}$. Daraus erhält man das reelle Fundamentalsystem

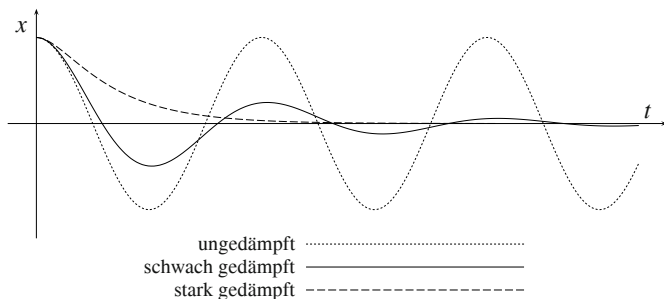
$$\begin{aligned} \varphi_1(t) &= e^{-\mu t} \cos(\omega t), \\ \varphi_2(t) &= e^{-\mu t} \sin(\omega t). \end{aligned}$$

Es findet also immer noch eine Schwingung statt, allerdings mit einer kleineren Frequenz $\omega < \omega_0$ gegenüber dem ungedämpften Fall. Die Amplitude der Schwingung nimmt mit dem Faktor $e^{-\mu t}$ exponentiell ab, vgl. Bild 14.3.

c) *Starke Dämpfung*: $\mu > \omega_0$.

In diesem Fall ist $\kappa := \sqrt{\mu^2 - \omega_0^2} < \mu$ reell, wir erhalten zwei negative Werte

$$\lambda_{1/2} = -\mu \pm \kappa, \quad \lambda_2 < -\mu < \lambda_1 < 0.$$

**Bild 14.3** Gedämpfte Schwingung

Ein Lösungs-Fundamentalsystem wird also gegeben durch die zwei exponentiell abfallenden Funktionen

$$\varphi_1(t) = e^{-(\mu-\kappa)t}, \quad \varphi_2(t) = e^{-(\mu+\kappa)t}.$$

d) *Aperiodischer Grenzfall*. Dies ist der bisher zurückgestellte Fall $\mu = \omega_0$.

Hier fallen die Werte $\lambda_1 = \lambda_2 = -\mu$ zusammen. Unser Ansatz liefert daher nur eine unabhängige Lösung $\varphi_1(t) = e^{-\mu t}$. Um zu einer zweiten, davon unabhängigen Lösung zu gelangen, gehen wir vom Fall starker Dämpfung $\mu > \omega_0$ aus und machen einen Grenzübergang $\mu \rightarrow \omega_0$. Dann strebt $\kappa = \sqrt{\mu^2 - \omega_0^2}$ gegen 0. Falls $\kappa > 0$, ist

$$\psi(\kappa, t) := \frac{1}{2\kappa} (e^{(-\mu+\kappa)t} - e^{(-\mu-\kappa)t}) = \frac{e^{\kappa t} - e^{-\kappa t}}{2\kappa} e^{-\mu t}$$

eine Lösung der Differentialgleichung. Es gilt

$$\lim_{\kappa \rightarrow 0} \frac{e^{\kappa t} - e^{-\kappa t}}{2\kappa} = \frac{d}{d\kappa} \sinh(\kappa t) \Big|_{\kappa=0} = t,$$

also

$$\lim_{\kappa \rightarrow 0} \psi(\kappa, t) = t e^{-\mu t} =: \varphi_2(t).$$

Durch direktes Nachrechnen kann man bestätigen, dass $\varphi_2(t) = t e^{-\mu t}$ tatsächlich eine Lösung der Differentialgleichung

$$\ddot{x} + 2\mu\dot{x} + \mu^2 x = 0$$

ist. Die beiden Lösungen $\varphi_1(t) = e^{-\mu t}$, $\varphi_2(t) = te^{-\mu t}$ sind linear unabhängig, denn ihre Wronski-Determinante im Nullpunkt ist

$$W(0) = \det \begin{pmatrix} \varphi_1(0) & \varphi_2(0) \\ \varphi_1'(0) & \varphi_2'(0) \end{pmatrix} = \det \begin{pmatrix} 1 & 0 \\ -\mu & 1 \end{pmatrix} = 1 \neq 0.$$

Bemerkung. Im nächsten Paragraphen werden wir die hier zur Lösung der Differentialgleichung der gedämpften Schwingung benutzten Methoden auf allgemeine lineare Differentialgleichungen n -ter Ordnung mit konstanten Koeffizienten verallgemeinern.

Legendresche Differentialgleichung

Die *Legendresche Differentialgleichung* ist definiert auf dem offenen Intervall $-1 < x < 1$ durch

$$(1-x^2)y'' - 2xy' + n(n+1)y = 0.$$

Dabei ist $n \geq 0$ ein ganzzahliger Parameter. Man beachte, dass $1-x^2$ auf dem betrachteten Intervall nicht verschwindet, man also mittels Division durch $1-x^2$ den Koeffizienten von y'' zu eins machen kann.

Satz 1. Das Legendre-Polynom n -ter Ordnung

$$P_n(x) := \frac{1}{2^n n!} \left(\frac{d}{dx} \right)^n (x^2 - 1)^n$$

ist eine Lösung der Legendreschen Differentialgleichung mit Parameter n .

Beweis. Wir benutzen folgende Verallgemeinerung der Leibnizschen Regel für die Differentiation eines Produkts zweier n -mal differenzierbaren Funktionen $f, g: I \rightarrow \mathbb{R}$ auf einem Intervall $I \subset \mathbb{R}$:

$$D^n(fg) = \sum_{k=0}^n \binom{n}{k} (D^k f)(D^{n-k} g).$$

Dabei steht D für $\frac{d}{dx}$. Die Regel lässt sich leicht durch vollständige Induktion beweisen, vgl. An. 1, Aufgabe 15.11.

Offensichtlich genügt es zu zeigen, dass die Funktion

$$y := D^n(x^2 - 1)^n$$

der Differentialgleichung genügt. Dazu verwenden wir einen kleinen Trick, indem wir die Funktion

$$\begin{aligned} z &:= (x^2 - 1)D(x^2 - 1)^n = (x^2 - 1)2nx(x^2 - 1)^{n-1} \\ &= 2nx(x^2 - 1)^n \end{aligned}$$

auf zwei Weisen mit der Produktregel differenzieren:

$$\begin{aligned} D^{n+1}z &= D^{n+1}((x^2 - 1)D(x^2 - 1)^n) \\ &= (x^2 - 1)D^{n+2}(x^2 - 1)^n + 2(n+1)x D^{n+1}(x^2 - 1)^n \\ &\quad + 2\binom{n+1}{2}D^n(x^2 - 1)^n \\ &= (x^2 - 1)y'' + 2(n+1)xy' + n(n+1)y. \end{aligned}$$

Andrerseits ist

$$\begin{aligned} D^{n+1}z &= D^{n+1}(2nx(x^2 - 1)^n) \\ &= 2nx D^{n+1}(x^2 - 1)^n + 2n(n+1)D^n(x^2 - 1)^n \\ &= 2nxy' + 2n(n+1)y. \end{aligned}$$

Vergleich ergibt

$$(x^2 - 1)y'' + 2xy' - n(n+1)y = 0, \quad \text{q.e.d.}$$

Reduktion der Ordnung

Mit den Legendre-Polynomen hat man die Legendresche Differentialgleichung aber noch nicht vollständig gelöst. Um die Gesamtheit aller Lösungen zu kennen, benötigt man noch eine zweite linear unabhängige Lösung. Der nächste Satz zeigt, wie man bei Kenntnis einer Lösung einer homogenen linearen Differentialgleichung 2. Ordnung das Problem auf die Lösung einer Differentialgleichung 1. Ordnung reduzieren kann. Es sei wieder \mathbb{K} einer der Körper \mathbb{R} oder \mathbb{C} .

Satz 2. Sei $I \subset \mathbb{R}$ ein Intervall und seien $a, b : I \rightarrow \mathbb{K}$ zwei stetige Funktionen. Weiter sei $\varphi : I \rightarrow \mathbb{K}$ eine Lösung der Differentialgleichung

$$y'' + a(x)y' + b(x)y = 0. \quad (5)$$

Im Intervall $J \subset I$ gelte $\varphi(x) \neq 0$. Dann erhält man über J eine zweite, von φ linear unabhängige Lösung $\psi : J \rightarrow \mathbb{K}$ durch den Ansatz

$$\psi(x) = \varphi(x)u(x),$$

wobei u eine nicht-konstante Lösung der Differentialgleichung

$$u'' + \left(2 \frac{\varphi'(x)}{\varphi(x)} + a(x)\right) u' = 0 \quad (6)$$

ist.

Bemerkung. (6) ist eine Differentialgleichung 1. Ordnung für u' , sie kann nach §11, Satz 2, gelöst werden durch

$$u'(x) = \frac{1}{\varphi(x)^2} \exp\left(-\int_{x_0}^x a(t) dt\right).$$

Man erhält u durch eine weitere Integration.

Beweis. Aus $\psi = \varphi u$ folgt

$$\begin{aligned}\psi' &= \varphi' u + \varphi u', \\ \psi'' &= \varphi'' u + 2\varphi' u' + \varphi u''\end{aligned}$$

Damit wird unter Benutzung von $\varphi'' + a\varphi' + b\varphi = 0$

$$\psi'' + a\psi' + b\psi = \varphi u'' + (2\varphi' + a\varphi)u'.$$

ψ erfüllt also genau dann die Differentialgleichung (4), wenn

$$u'' + \left(2 \frac{\varphi'}{\varphi} + a\right) u' = 0.$$

Ist u nicht-konstant, so ist $\psi = \varphi u$ von φ auf dem Intervall J linear unabhängig.

(14.3) Beispiel. Für $n = 1$ lautet die Legendresche Differentialgleichung

$$y'' - \frac{2x}{1-x^2} y' + \frac{2}{1-x^2} y = 0, \quad (|x| < 1).$$

Sie besitzt die Lösung $\varphi(x) = x$. Infolgedessen erhält man im Intervall $]0, 1[$ eine zweite Lösung durch den Ansatz

$$\psi(x) = xu(x),$$

wobei

$$u'' + \left(\frac{2}{x} - \frac{2x}{1-x^2}\right) u' = 0.$$

Dies kann man lösen durch

$$\begin{aligned} u'(x) &= \exp\left(-\int \frac{2}{x} dx + \int \frac{2x}{1-x^2} dx\right) \\ &= \exp(-2\log x - \log(1-x^2)) \\ &= \frac{1}{x^2(1-x^2)} = \frac{1}{x^2} + \frac{1}{2}\left(\frac{1}{1-x} + \frac{1}{1+x}\right), \end{aligned}$$

also

$$u(x) = \int u'(x) dx = -\frac{1}{x} + \frac{1}{2} \log \frac{1+x}{1-x}.$$

Deshalb ist

$$\psi(x) = xu(x) = \frac{x}{2} \log \frac{1+x}{1-x} - 1$$

eine von ϕ linear unabhängige Lösung der Differentialgleichung. Dies wurde nur im Intervall $]0, 1[$ hergeleitet, aber man sieht direkt, dass ψ die Differentialgleichung im ganzen Intervall $] -1, 1[$ löst.

Hermitesche und Laguerresche Differentialgleichung

Es gibt noch weitere Differentialgleichungen 2. Ordnung, die durch spezielle Polynome gelöst werden.

(14.4) Die *Hermitesche Differentialgleichung* zum Parameter $n \in \mathbb{N}$ ist definiert auf ganz \mathbb{R} durch

$$y'' - 2xy' + 2ny = 0.$$

Das Hermitesche Polynom n -ter Ordnung

$$H_n(x) := (-1)^n e^{x^2} \left(\frac{d}{dx}\right)^n e^{-x^2}$$

ist eine Lösung davon (Aufgabe 14.2).

(14.5) Die *Laguerresche Differentialgleichung* zum Parameter $n \in \mathbb{N}$ ist für $x > 0$ definiert durch

$$xy'' + (1-x)y' + ny = 0.$$

Das Laguerresche Polynom n -ter Ordnung

$$L_n(x) := e^x \left(\frac{d}{dx}\right)^n (x^n e^{-x})$$

ist eine Lösung davon (Aufgabe 14.4).

Besselsche Differentialgleichung

Es ist keinesfalls so, dass sich die Lösungen einer Differentialgleichung immer durch elementare Funktionen ausdrücken lassen. Vielmehr stößt man dabei oft auf neue transzendente Funktionen. Dies ist z.B. der Fall bei der Besselschen Differentialgleichung

$$y'' + \frac{1}{x}y' + \left(1 - \frac{p^2}{x^2}\right)y = 0, \quad (x > 0).$$

Dabei ist p ein reeller Parameter. Ihre Lösungen heißen *Zylinderfunktionen* der Ordnung p . Sie bilden nach der allgemeinen Theorie einen zweidimensionalen Vektorraum. Eine spezielle Basis bildet die sog. *Bessel-Funktion* p -ter Ordnung J_p zusammen mit der *Neumannschen Funktion* p -ter Ordnung N_p . Die Funktionen liegen tabelliert vor, vgl. z.B. Jahnke-Emde-Lösch: Tafeln höherer Funktionen, Teubner, Stuttgart. Die Funktionen sind auch in Computeralgebra-Systemen direkt verfügbar, z.B. in Maple unter den Namen `BesselJ(p,x)` und `BesselY(p,x)`.

Die Bedeutung der Besselschen Differentialgleichung beruht auf ihrem Zusammenhang mit der Wellengleichung der Raumdimension zwei.

Satz 3. Sei $f: \mathbb{R}_+^* \rightarrow \mathbb{C}$ eine zweimal stetig differenzierbare Funktion und p eine ganze Zahl. Mittels ebener Polarkoordinaten (r, φ) ,

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

werde eine Funktion $u: \mathbb{R}^2 \setminus 0 \rightarrow \mathbb{C}$ definiert durch

$$u(x, y) := f(r)e^{ip\varphi}.$$

Dann gilt: Genau dann genügt die Funktion ue^{it} der zweidimensionalen Wellengleichung

$$\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} - \frac{\partial^2}{\partial t^2}\right)u(x, y)e^{it} = 0,$$

wenn f die Besselsche Differentialgleichung zum Parameter p löst, d.h.

$$f''(r) + \frac{1}{r}f'(r) + \left(1 - \frac{p^2}{r^2}\right)f(r) = 0.$$

Beweis. Wir verwenden die Darstellung des Laplace-Operators bzgl. ebener Polarkoordinaten, vgl. Aufgabe 6.2:

$$\Delta = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} = \frac{\partial^2}{\partial r^2} + \frac{1}{r} \frac{\partial}{\partial r} + \frac{1}{r^2} \frac{\partial^2}{\partial \varphi^2}.$$

Damit wird

$$\begin{aligned} \Delta(f(r)e^{ip\varphi}e^{it}) &= \left(\frac{\partial^2}{\partial r^2} + \frac{1}{r} \frac{\partial}{\partial r}\right)f(r)e^{ip\varphi}e^{it} + f(r)\left(\frac{1}{r^2} \frac{\partial^2}{\partial \varphi^2}\right)e^{ip\varphi}e^{it} \\ &= \left(f''(r) + \frac{1}{r}f'(r) - \frac{p^2}{r^2}f(r)\right)e^{ip\varphi}e^{it}. \end{aligned}$$

Andrerseits ist

$$\frac{\partial^2}{\partial t^2}(f(r)e^{ip\varphi}e^{it}) = -f(r)e^{ip\varphi}e^{it}.$$

Da $e^{ip\varphi}e^{it} \neq 0$, ergibt sich daraus die Behauptung.

Wir wollen uns mit dem Fall $p = 0$ etwas genauer beschäftigen (dies entspricht den rotationssymmetrischen Lösungen der Wellengleichung). Um zu einer Lösung der Differentialgleichung

$$y'' + \frac{1}{x}y' + y = 0$$

zu gelangen, machen wir einen Potenzreihen-Ansatz

$$y(x) = \sum_{k=0}^{\infty} a_k x^k. \quad (7)$$

Damit wird

$$\begin{aligned} y'(x) &= \sum_{k=1}^{\infty} k a_k x^{k-1}, \\ y''(x) &= \sum_{k=2}^{\infty} k(k-1) a_k x^{k-2}, \end{aligned}$$

also (mit Index-Verschiebung)

$$\begin{aligned} \frac{1}{x}y'(x) &= \sum_{k=-1}^{\infty} (k+2) a_{k+2} x^k, \\ y''(x) &= \sum_{k=0}^{\infty} (k+2)(k+1) a_{k+2} x^k. \end{aligned}$$

Setzt man dies in die Gleichung $y'' + \frac{1}{x}y' + y = 0$ ein, und vergleicht die Koeffizienten von x^k , so sieht man, dass notwendig $a_1 = 0$ und

$$(k+2)(k+1)a_{k+2} + (k+2)a_{k+2} + a_k = 0 \implies (k+2)^2 a_{k+2} = -a_k$$

für alle $k \geq 0$. Daraus folgt $a_k = 0$ für alle ungeraden k und

$$a_{2k} = (-1)^k \frac{1}{2^{2k}(k!)^2} a_0.$$

Wenn also überhaupt eine Lösung existiert, die sich in eine Potenzreihe (7) entwickeln lässt, hat diese bis auf einen konstanten Faktor die Gestalt

$$y(x) = J_0(x) := \sum_{k=0}^{\infty} (-1)^k \frac{1}{(k!)^2} \left(\frac{x}{2}\right)^{2k}.$$

Tatsächlich stellt man fest (z.B. mit dem Quotientenkriterium), dass die Reihe für alle $x \in \mathbb{R}$ konvergiert (sogar sehr schnell). Man darf deshalb gliedweise differenzieren und sieht, dass die Funktion die Differentialgleichung erfüllt. Nach Definition ist dies die Besselfunktion 0-ter Ordnung. Eine zweite, davon linear unabhängige Lösung ist die Neumannsche Funktion 0-ter Ordnung

$$N_0(x) := \frac{2}{\pi} \left\{ \left(\log \frac{x}{2} + \gamma \right) J_0(x) + \sum_{k=1}^{\infty} (-1)^{k-1} \frac{c_k}{(k!)^2} \left(\frac{x}{2}\right)^{2k} \right\},$$

wobei $\gamma = 0.57721566 \dots$ die Euler-Mascheronische Konstante ist (vgl. An. 1, §20) und

$$c_k = 1 + \frac{1}{2} + \dots + \frac{1}{k}.$$

Wieder kann man sich leicht von der Konvergenz der Reihe und der Erfüllung der Differentialgleichung überzeugen. Dass J_0 und N_0 linear unabhängig sind folgt daraus, dass $\lim_{x \searrow 0} J_0(x) = 1$ und $\lim_{x \searrow 0} N_0(x) = \infty$.

Also ist (J_0, N_0) ein Lösungs-Fundamentalsystem der Besselschen Differentialgleichung der Ordnung $p = 0$, siehe Bild 14.4.

Zylinderfunktionen ganzzahliger Ordnung $p > 0$ kann man auf den Fall $p = 0$ zurückführen, siehe Aufgabe 14.7. Für den Fall halb-ganzer p siehe Aufgabe 14.6.

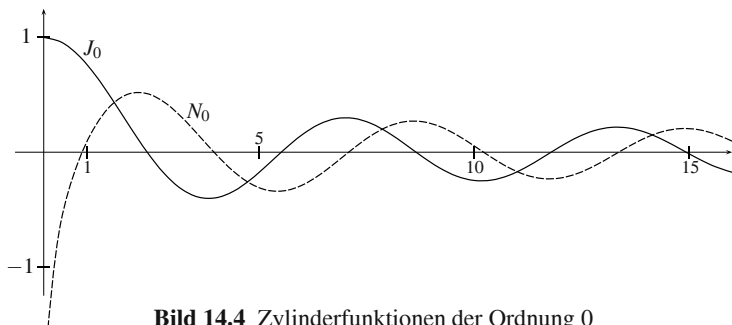


Bild 14.4 Zylinderfunktionen der Ordnung 0

AUFGABEN

14.1. Man löse die Differentialgleichung

$$\frac{d^2 r}{dt^2} = -\frac{c}{r^2}, \quad (r > 0),$$

mit der Anfangsbedingung

$$r(0) = r_0 > 0, \quad \dot{r}(0) = v_0 > 0.$$

Dabei ist c eine positive Konstante.

Man zeige: Es gibt ein $v^* > 0$, so dass für $v_0 \geq v^*$ die Lösung $r(t)$ für $t \rightarrow \infty$ unbegrenzt wächst, während für $v_0 < v^*$ ein $t_1 > 0$ existiert, so dass die Lösung $r(t)$ im Intervall $0 \leq t \leq t_1$ monoton wächst und für $t \geq t_1$ monoton fällt.

Bemerkung. Die Differentialgleichung beschreibt die radiale Bewegung eines Körpers unter dem Einfluss der Schwerkraft eines anderen.

Man berechne die Geschwindigkeit v^* für die Erdanziehung und

$$r_0 = 6370 \text{ km} \quad (\text{Erdradius}).$$

Für die Erde ist $c = gr_0^2$, wobei

$$g = 9.81 \frac{\text{m}}{\text{sec}^2} \quad (\text{Erdbeschleunigung}).$$

14.2. a) Man zeige, dass

$$H_n(x) := (-1)^n e^{x^2} \left(\frac{d}{dx} \right)^n e^{-x^2}$$

ein Polynom n -ten Grades ist und die Hermitesche Differentialgleichung

$$y'' - 2xy' + 2ny = 0$$

löst.

b) Man zeige, dass für $n \neq m$ die Hermiteschen Polynome H_n und H_m orthogonal sind bzgl. des Skalarprodukts

$$\langle f, g \rangle := \int_{-\infty}^{\infty} f(x)g(x)e^{-x^2} dx.$$

14.3. Sei n eine natürliche Zahl. Man zeige, dass für die Lösungen der Differentialgleichung

$$y'' + (2n + 1 - x^2)y = 0$$

gilt $y(x) = e^{-x^2/2}u(x)$, wobei u eine Lösung der Hermiteschen Differentialgleichung

$$u'' - 2xu' + 2nu = 0$$

ist.

14.4. a) Man zeige, dass

$$L_n(x) := e^x \left(\frac{d}{dx} \right)^n (x^n e^{-x})$$

ein Polynom n -ten Grades ist und die Laguerresche Differentialgleichung

$$xy'' + (1-x)y' + ny = 0 \quad (x > 0)$$

löst.

b) Man zeige, dass für $n \neq m$ die Laguerreschen Polynome L_n und L_m orthogonal sind bzgl. des Skalarprodukts

$$\langle f, g \rangle := \int_0^{\infty} f(x)g(x)e^{-x} dx.$$

14.5. Man bestimme alle Lösungen der folgenden Differentialgleichungen.

a) $(2x+1)y'' + (4x-2)y' - 8y = (6x^2+x-3)e^x, \quad (x > -\frac{1}{2}),$

b) $x^2(1-x)y'' + 2x(2-x)y' + 2(1+x)y = x^2 \quad (0 < x < 1).$

Anleitung. Die zugehörige homogene Gleichung besitzt eine spezielle Lösung der Gestalt $y = e^{\alpha x}$ im Fall a) und $y = x^\beta$ im Fall b) mit geeigneten Konstanten $\alpha, \beta \in \mathbb{R}$. Eine weitere Lösung der homogenen Gleichung erhält man mit Satz 2. Eine spezielle Lösung der inhomogenen Gleichung bestimme man durch Zurückführung auf ein System 1. Ordnung und Variation der Konstanten.

14.6. Man bestimme ein Lösungs-Fundamentalsystem der Besselschen Differentialgleichung für $p = \frac{1}{2}$,

$$y'' + \frac{1}{x}y' + \left(1 - \frac{1}{4x^2}\right)y = 0$$

durch den Ansatz $z = \sqrt{x} y$.

14.7. Sei $C^\infty(\mathbb{R}_+^*)$ der Vektorraum aller beliebig oft differenzierbaren Funktionen $f: \mathbb{R}_+^* \rightarrow \mathbb{R}$. Lineare Abbildungen

$$T_p, S_p, B_p: C^\infty(\mathbb{R}_+^*) \rightarrow C^\infty(\mathbb{R}_+^*)$$

seien wie folgt definiert:

$$(T_p f)(x) := f'(x) + \frac{p}{x} f(x),$$

$$(S_p f)(x) := -f'(x) + \frac{p}{x} f(x),$$

$$(B_p f)(x) := f''(x) + \frac{1}{x} f'(x) + \left(1 - \frac{p^2}{x^2}\right) f(x).$$

(Die Besselsche Differentialgleichung lässt sich dann einfach als $B_p y = 0$ schreiben.)

a) Man zeige: Für jedes $f \in C^\infty(\mathbb{R}_+^*)$ gilt

i) $T_{p+1} S_p f = f - B_p f,$

ii) $S_{p-1} T_p f = f - B_p f,$

iii) $T_p B_p f = B_{p-1} T_p f,$

$$\text{iv)} \quad S_p B_p f = B_{p+1} S_p f.$$

b) Sei $V_p := \{f \in C^\infty(\mathbb{R}_+^*) : B_p f = 0\}$ der Vektorraum aller Zylinderfunktionen der Ordnung p . Man zeige:

$$\text{i)} \quad T_p(V_p) \subset V_{p-1}, \quad S_p(V_p) \subset V_{p+1}.$$

ii) Die Abbildungen

$$S_p: V_p \rightarrow V_{p+1} \quad \text{und} \quad T_{p+1}: V_{p+1} \rightarrow V_p$$

sind Isomorphismen und Umkehrungen von einander.

c) Man bestimme mittels b) und Aufgabe 14.6 alle Zylinderfunktionen der Ordnungen $p = 3/2$ und $p = 5/2$.

14.8. a) Seien α, β, γ, p reelle Konstanten, $\beta > 0, \gamma \neq 0$. Man zeige, dass für die Lösungen der Differentialgleichung

$$y'' + \frac{1-2\alpha}{x} y' + \left((\beta \gamma x^{\gamma-1})^2 + \frac{\alpha^2 - p^2 \gamma^2}{x^2} \right) y = 0, \quad (x > 0),$$

gilt $y(x) = x^\alpha u(\beta x^\gamma)$, wobei u eine Lösung der Besselschen Differentialgleichung zum Parameter p ist.

b) Man drücke die Lösungen der folgenden Differentialgleichungen mit Hilfe von Zylinderfunktionen aus ($a, b, m \in \mathbb{R}$):

$$\text{i)} \quad y'' + a^2 x^m y = 0 \quad (a \neq 0, m \neq -2),$$

$$\text{ii)} \quad y'' + \left(1 - \frac{a(a+1)}{x^2} \right) y = 0,$$

$$\text{iii)} \quad y'' + \frac{a}{x} y' + \frac{b^2}{4x} y = 0, \quad (b \neq 0).$$

c) Man löse die Differentialgleichungen i) und iii) in den Ausnahmefällen $m = -2$ und $b = 0$.

§ 15 Lineare Differentialgleichungen mit konstanten Koeffizienten

Für lineare Differentialgleichungen n -ter Ordnung mit konstanten Koeffizienten gibt es eine sehr befriedigende Lösungstheorie. Die Lösung einer solchen Differentialgleichung ist äquivalent mit der Bestimmung der Nullstellen eines Polynoms n -ten Grades.

Polynome von Differentialoperatoren

Wir bezeichnen mit $\mathbb{C}[T]$ die Menge aller Polynome

$$P(T) = a_0 + a_1T + \dots + a_nT^n$$

mit komplexen Koeffizienten a_k in der Unbestimmten T . Ersetzt man hierin die Unbestimmte T durch $D = \frac{d}{dx}$, so erhält man einen "Differentialoperator"

$$P(D) = a_0 + a_1D + \dots + a_nD^n,$$

d.h. eine Abbildung, die einer auf einem Intervall $I \subset \mathbb{R}$ definierten, n -mal differenzierbaren Funktion

$$f : I \longrightarrow \mathbb{C}, \quad x \mapsto f(x),$$

die Funktion

$$P(D)f := a_0f + a_1Df + \dots + a_nD^n f$$

zuordnet. Mit Hilfe dieser Differentialoperatoren schreibt sich eine homogene Differentialgleichung n -ter Ordnung mit konstanten Koeffizienten einfach als

$$P(D)y = 0,$$

wobei $P \in \mathbb{C}[T]$ ein Polynom n -ten Grades mit höchstem Koeffizienten 1 ist.

Wir wollen jetzt zeigen, dass man mit Polynomen von Differentialoperatoren ganz analog rechnen kann wie mit gewöhnlichen Polynomen.

a) *Addition.* Seien $P_1(T), P_2(T) \in \mathbb{C}[T]$ und

$$P(T) := P_1(T) + P_2(T).$$

Dann gilt für jede genügend oft differenzierbare Funktion $f: I \rightarrow \mathbb{C}$

$$P(D)f = P_1(D)f + P_2(D)f.$$

Beweis. Sei

$$P_1(T) = \sum_{k=0}^n a_k T^k \quad \text{und} \quad P_2(T) = \sum_{k=0}^m b_k T^k.$$

Man kann o.B.d.A. annehmen, dass $m = n$. (Falls etwa $m < n$ ergänze man $b_{m+1} = \dots = b_n = 0$.) Dann ist

$$P(T) = \sum_{k=0}^n (a_k + b_k) T^k.$$

Damit ergibt sich

$$\begin{aligned} P(D)f &= \sum_{k=0}^n (a_k + b_k) D^k f = \sum_{k=0}^n a_k D^k f + \sum_{k=0}^n b_k D^k f \\ &= P_1(D)f + P_2(D)f. \end{aligned}$$

b) *Multiplikation.* Seien $P_1(T), P_2(T) \in \mathbb{C}[T]$ und

$$Q(T) := P_1(T)P_2(T).$$

Dann gilt für jede genügend oft differenzierbare Funktion $f: I \rightarrow \mathbb{C}$

$$Q(D)f = P_1(D)(P_2(D)f).$$

Beweis. Ist

$$P_1(T) = \sum_{v=0}^n a_v T^v \quad \text{und} \quad P_2(T) = \sum_{\mu=0}^m b_\mu T^\mu,$$

so folgt

$$Q(T) = \sum_{k=0}^{n+m} c_k T^k \quad \text{mit} \quad c_k = \sum_{v=0}^k a_v b_{k-v}.$$

(Dabei ist $a_v = 0$ für $v > n$ und $b_\mu = 0$ für $\mu > m$ zu setzen.) Damit ergibt sich

$$Q(D)f = \sum_{k=0}^{n+m} c_k D^k f = \sum_{k=0}^{n+m} \left(\sum_{v=0}^k a_v b_{k-v} \right) D^k f$$

$$\begin{aligned}
&= \sum_{k=0}^{n+m} \left(\sum_{\nu+\mu=k} a_\nu b_\mu D^{\nu+\mu} f \right) = \sum_{k=0}^{n+m} \left(\sum_{\nu+\mu=k} a_\nu D^\nu (b_\mu D^\mu f) \right) \\
&= \sum_{\nu=0}^n a_\nu D^\nu \left(\sum_{\mu=0}^m b_\mu D^\mu f \right) \\
&= P_1(D)(P_2(D)f) = Q(D)f, \quad \text{q.e.d.}
\end{aligned}$$

Wir beschäftigen uns jetzt mit der Wirkung von Differentialpolynomen $P(D)$ auf Funktionen der speziellen Gestalt $f(x) = e^{\lambda x}$, wobei λ eine reelle oder komplexe Konstante ist.

Hilfssatz 1. Für jedes Polynom $P(T) \in \mathbb{C}[T]$ und jedes $\lambda \in \mathbb{C}$ gilt

$$P(D)e^{\lambda x} = P(\lambda)e^{\lambda x}.$$

Beweis. Sei $P(T) = \sum_{k=0}^n a_k T^k$.

Da $D e^{\lambda x} = \frac{d}{dx} e^{\lambda x} = \lambda e^{\lambda x}$, folgt $D^k e^{\lambda x} = \lambda^k e^{\lambda x}$ und

$$P(D)e^{\lambda x} = \sum_{k=0}^n a_k D^k e^{\lambda x} = \sum_{k=0}^n a_k \lambda^k e^{\lambda x} = P(\lambda)e^{\lambda x}, \quad \text{q.e.d.}$$

Insbesondere folgt aus Hilfssatz 1: Ist λ eine Nullstelle des Polynoms P , d.h. $P(\lambda) = 0$, so ist die Funktion $\varphi(x) = e^{\lambda x}$ eine Lösung der Differentialgleichung $P(D)y = 0$.

Satz 1. Sei $P(T) = T^n + a_{n-1}T^{n-1} + \dots + a_1T + a_0 \in \mathbb{C}[T]$. Das Polynom P habe n paarweise von einander verschiedene Nullstellen $\lambda_1, \dots, \lambda_n \in \mathbb{C}$. Dann bilden die Funktionen $\varphi_k: \mathbb{R} \rightarrow \mathbb{C}$,

$$\varphi_k(x) := e^{\lambda_k x}, \quad k = 1, \dots, n,$$

ein Fundamentalsystem von Lösungen der Differentialgleichung

$$P(D)y = y^{(n)} + a_{n-1}y^{(n-1)} + \dots + a_1y' + a_0y = 0.$$

Beweis. Dass die Funktionen φ_k Lösungen der Differentialgleichung sind, folgt aus Hilfssatz 1. Es ist nur die lineare Unabhängigkeit zu zeigen. Dazu berechnen wir die Wronski-Determinante W von $\varphi_1, \dots, \varphi_n$. Da

$$\varphi_k^{(v)}(x) = \lambda_k^v e^{\lambda_k x},$$

ergibt sich für den Wert der Wronski-Determinante an der Stelle 0

$$W(0) = \det \begin{pmatrix} 1 & 1 & \dots & 1 \\ \lambda_1 & \lambda_2 & \dots & \lambda_n \\ \vdots & \vdots & & \vdots \\ \lambda_1^{n-1} & \lambda_2^{n-1} & \dots & \lambda_n^{n-1} \end{pmatrix}.$$

Dies ist aber die aus der linearen Algebra bekannte Vandermondesche Determinante, für die gilt

$$W(0) = \prod_{j>k} (\lambda_j - \lambda_k) \neq 0,$$

da die λ_k paarweise von einander verschieden sind. Also sind nach § 13, Satz 5, die Lösungen $\varphi_1, \dots, \varphi_n$ linear unabhängig.

Bemerkung. Wir werden in Satz 2 noch einmal die lineare Unabhängigkeit der Funktionen $e^{\lambda_k x}$ ohne Benutzung der Vandermondeschen Determinante beweisen.

(15.1) Beispiel. Die Differentialgleichung

$$y''' - 2y'' + y' - 2y = 0$$

lässt sich schreiben als $P(D)y = 0$ mit

$$P(T) = T^3 - 2T^2 + T - 2.$$

Das Polynom P zerfällt folgendermaßen in Faktoren

$$P(T) = (T^2 + 1)(T - 2) = (T - i)(T + i)(T - 2),$$

hat also die Nullstellen

$$\lambda_1 = i, \quad \lambda_2 = -i, \quad \lambda_3 = 2.$$

Deshalb bilden die Funktionen $\varphi_k: \mathbb{R} \rightarrow \mathbb{C}$,

$$\varphi_1(x) := e^{ix}, \quad \varphi_2(x) := e^{-ix}, \quad \varphi_3(x) := e^{2x}$$

ein Fundamentalsystem von Lösungen der Differentialgleichung. Durch geeignete Linearkombinationen lässt sich daraus ein reelles Fundamentalsystem gewinnen:

$$\psi_1(x) = \frac{1}{2}(\varphi_1(x) + \varphi_2(x)) = \cos x,$$

$$\psi_2(x) = \frac{1}{2i}(\varphi_1(x) - \varphi_2(x)) = \sin x.$$

Da die Matrix $\frac{1}{2} \begin{pmatrix} 1 & 1 \\ -i & i \end{pmatrix}$ invertierbar ist, lassen sich umgekehrt φ_1 und φ_2 aus ψ_1, ψ_2 linear kombinieren, also bilden die Funktionen

$$\psi_1(x) = \cos x, \quad \psi_2(x) = \sin x, \quad \varphi_3(x) = e^{2x}$$

ein reelles Fundamentalsystem von Lösungen der Differentialgleichung.

Mehrfache Nullstellen

Ein Polynom n -ten Grades

$$P(T) = T^n + a_{n-1}T^{n-1} + \dots + a_1T + a_0 \in \mathbb{C}[T]$$

lässt sich nach dem Fundamentalsatz der Algebra (siehe §3, Corollar zu Satz 8) stets folgendermaßen in Linearfaktoren zerlegen:

$$P(T) = (T - \lambda_1)^{k_1} (T - \lambda_2)^{k_2} \dots (T - \lambda_r)^{k_r}$$

mit paarweise verschiedenen $\lambda_j \in \mathbb{C}$ und natürlichen Zahlen $k_j \geq 1$, $\sum k_j = n$. Dabei ist k_j die Vielfachheit der Nullstelle λ_j . Falls mindestens ein $k_j \geq 2$, erhält man mit den Funktionen $e^{\lambda_j x}$ weniger als n linear unabhängige Lösungen der Differentialgleichung $P(D)y = 0$. Um die noch fehlenden Lösungen zu erhalten, brauchen wir einige Vorbereitungen.

Hilfssatz 2. Sei $\lambda \in \mathbb{C}$ und $k \in \mathbb{N}$. Dann gilt für jede auf einem Intervall $I \subset \mathbb{R}$ k -mal differenzierbare Funktion $f: I \rightarrow \mathbb{C}$

$$(D - \lambda)^k (f(x)e^{\lambda x}) = f^{(k)}(x)e^{\lambda x}.$$

Beweis durch vollständige Induktion nach k .

Für $k = 0$ ist die Aussage trivial. Für $k = 1$ erhalten wir

$$\begin{aligned} (D - \lambda)(f(x)e^{\lambda x}) &= D(f(x)e^{\lambda x}) - \lambda f(x)e^{\lambda x} \\ &= f'(x)e^{\lambda x} + f(x)\lambda e^{\lambda x} - \lambda f(x)e^{\lambda x} = f'(x)e^{\lambda x}. \end{aligned}$$

Induktionsschritt $(k - 1) \rightarrow k$.

$$\begin{aligned} (D - \lambda)^k(f(x)e^{\lambda x}) &= (D - \lambda)(D - \lambda)^{(k-1)}(f(x)e^{\lambda x}) \\ &= (D - \lambda)(f^{(k-1)}(x)e^{\lambda x}) = f^{(k)}(x)e^{\lambda x}, \quad \text{q.e.d.} \end{aligned}$$

Hilfssatz 3. Sei $P(T) \in \mathbb{C}[T]$ ein Polynom und $\lambda \in \mathbb{C}$ mit $P(\lambda) \neq 0$. Ist dann $g: \mathbb{R} \rightarrow \mathbb{C}$ eine Polynomfunktion vom Grad k , so gilt

$$P(D)(g(x)e^{\lambda x}) = h(x)e^{\lambda x},$$

wobei $h: \mathbb{R} \rightarrow \mathbb{C}$ wieder eine Polynomfunktion vom Grad k ist.

Beweis. Man kann das Polynom P nach Potenzen von $T - \lambda$ umordnen:

$$P(T) = \sum_{v=0}^n c_v (T - \lambda)^v, \quad c_v \in \mathbb{C}.$$

Es ist $c_0 = P(\lambda) \neq 0$. Nach Hilfssatz 2 gilt dann

$$\begin{aligned} P(D)(g(x)e^{\lambda x}) &= \sum_{v=0}^n c_v (D - \lambda)^v(g(x)e^{\lambda x}) \\ &= \sum_{v=0}^n c_v g^{(v)}(x)e^{\lambda x} = h(x)e^{\lambda x} \end{aligned}$$

mit

$$h(x) = \sum_{v=0}^n c_v g^{(v)}(x).$$

Wegen $c_0 \neq 0$ hat h denselben Grad wie g , q.e.d.

Wir können jetzt den Hauptsatz über die Lösungen von Differentialgleichungen n -ter Ordnung mit konstanten Koeffizienten beweisen.

Satz 2. Das Polynom

$$P(T) = T^n + a_{n-1}T^{n-1} + \dots + a_1T + a_0 \in \mathbb{C}[T]$$

habe die paarweise von einander verschiedenen Nullstellen $\lambda_j \in \mathbb{C}$ mit den Vielfachheiten k_j , $1 \leq j \leq r$. Dann besitzt die Differentialgleichung

$$P(D)y = 0$$

ein Lösungs-Fundamentalsystem aus folgenden Funktionen:

$$\varphi_{jm}(x) := x^m e^{\lambda_j x}, \quad 1 \leq j \leq r, \quad 0 \leq m \leq k_j - 1.$$

Beweis. a) Alle angegebenen Funktionen lösen die Differentialgleichung. Denn $P(T)$ besitzt den Faktor $(T - \lambda_j)^{k_j}$, d.h.

$$P(T) = Q_j(T)(T - \lambda_j)^{k_j}, \quad Q_j(T) \in \mathbb{C}[T].$$

Also folgt mit Hilfssatz 2

$$\begin{aligned} P(D)\varphi_{jm}(x) &= Q_j(D)(D - \lambda_j)^{k_j}(x^m e^{\lambda_j x}) \\ &= Q_j(D)(D^{k_j} x^m) e^{\lambda_j x} = 0, \quad \text{da } k_j > m. \end{aligned}$$

b) Es ist noch zu zeigen, dass die Funktionen φ_{jm} linear unabhängig sind. Eine Linearkombination der φ_{jm} hat die Gestalt

$$\sum_{j=1}^r g_j(x) e^{\lambda_j x},$$

wobei die g_j Polynome vom Grad $\leq k_j - 1$ sind. Wir müssen beweisen, dass diese Linearkombination nur dann die Nullfunktion darstellt, wenn alle g_j identisch verschwinden. Wir zeigen das durch Induktion nach r .

Induktionsanfang $r = 1$. Falls $g_1(x) e^{\lambda_1 x} = 0$ für alle $x \in \mathbb{R}$, muss g_1 das Nullpolynom sein.

Induktionsschritt $(r - 1) \rightarrow r$. Es gelte

$$\sum_{j=1}^r g_j(x) e^{\lambda_j x} = 0 \quad \text{für alle } x \in \mathbb{R}.$$

Falls eines der Polynome g_j identisch null ist, sind wir nach Induktionsvoraussetzung fertig. Andernfalls wenden wir auf die Gleichung den Differentialoperator $(D - \lambda_r)^{k_r}$ an und erhalten mit den Hilfssätzen 2 und 3

$$\sum_{j=1}^{r-1} h_j(x) e^{\lambda_j x} = 0 \quad \text{für alle } x \in \mathbb{R},$$

wobei die h_j Polynome sind, die ebenfalls nicht identisch verschwinden. Nach Induktionsvoraussetzung ist dies aber unmöglich.

(15.2) Beispiel. Die Differentialgleichung

$$y^{(4)} + 8y'' + 16y = 0$$

gehört zum Differentialoperator $P(D) = D^4 + 8D^2 + 16$. Nun ist

$$T^4 + 8T^2 + 16 = (T^2 + 4)^2 = (T - 2i)^2(T + 2i)^2.$$

Mit den Bezeichnungen von Satz 2 ist also $\lambda_1 = 2i$, $\lambda_2 = -2i$ und $k_1 = k_2 = 2$. Deshalb bilden die folgenden vier Funktionen ein Fundamentalsystem von Lösungen:

$$\begin{aligned}\varphi_{10}(x) &= e^{2ix}, & \varphi_{11}(x) &= xe^{2ix}, \\ \varphi_{20}(x) &= e^{-2ix}, & \varphi_{21}(x) &= xe^{-2ix}.\end{aligned}$$

Daraus lässt sich folgendes reelle Lösungs-Fundamentalsystem erhalten:

$$\begin{aligned}\psi_{10}(x) &= \cos 2x, & \psi_{11}(x) &= x \cos 2x, \\ \psi_{20}(x) &= \sin 2x, & \psi_{21}(x) &= x \sin 2x.\end{aligned}$$

Inhomogene Differentialgleichungen

Es sei

$$P(D) = D^n + a_{n-1}D^{n-1} + \dots + a_1D + a_0$$

ein linearer Differentialoperator mit konstanten Koeffizienten $a_k \in \mathbb{C}$ und $b: I \rightarrow \mathbb{C}$ eine stetige Funktion auf dem Intervall $I \subset \mathbb{R}$. Dann kann man die inhomogene Differentialgleichung

$$P(D)y = b(x) \tag{*}$$

prinzipiell so lösen: Man bestimmt zunächst mittels Satz 2 ein Fundamentalsystem von Lösungen der homogenen Gleichung $P(D)y = 0$, führt dann (*) auf ein System von Differentialgleichungen 1. Ordnung zurück und bestimmt eine spezielle Lösung der inhomogenen Gleichung durch Variation der Konstanten (§ 13, Satz 4). Wir werden jedoch sehen, dass man bei spezieller Form der Funktion b eine Lösung der inhomogenen Gleichung (*) durch einen ein-

fachen Lösungsansatz erhalten kann. Zunächst eine einfache Vorbemerkung:
Ist

$$b(x) = b_1(x) + \dots + b_s(x)$$

eine Summe der Funktionen $b_j: I \rightarrow \mathbb{C}$, und sind die Funktionen $\psi_j: I \rightarrow \mathbb{C}$ Lösungen von $P(D)y = b_j(x)$, so ist die Summe

$$\psi(x) := \psi_1(x) + \dots + \psi_s(x)$$

eine Lösung von $P(D)y = b(x)$.

Wir untersuchen jetzt spezielle rechte Seiten der Gestalt

$$b(x) = f(x)e^{\mu x}, \quad \mu \in \mathbb{C},$$

wobei f ein Polynom in x vom Grad $m \geq 0$ mit komplexen Koeffizienten ist. Dabei zu unterscheiden, ob $P(\mu) = 0$ (sog. Resonanzfall), oder $P(\mu) \neq 0$.

Satz 3 (Inhomogene Gleichung, keine Resonanz). *Sei*

$$P(T) = T^n + a_{n-1}T^{n-1} + \dots + a_1T + a_0 \in \mathbb{C}[T]$$

ein Polynom und $\mu \in \mathbb{C}$ eine Zahl mit $P(\mu) \neq 0$. Dann gilt:

a) Die Differentialgleichung

$$P(D)y = e^{\mu x}$$

besitzt die spezielle Lösung

$$\varphi(x) := \frac{1}{P(\mu)} e^{\mu x}.$$

b) Ist allgemeiner $f: \mathbb{R} \rightarrow \mathbb{C}$ ein Polynom vom Grad m mit komplexen Koeffizienten, so besitzt die Differentialgleichung

$$P(D)y = f(x)e^{\mu x}$$

eine spezielle Lösung der Gestalt

$$\psi(x) = g(x)e^{\mu x},$$

wobei g ein Polynom vom Grad m ist.

Beweis. a) Nach Hilfssatz 1 gilt

$$P(D)e^{\mu x} = P(\mu)e^{\mu x}, \quad \text{also} \quad P(D)\varphi(x) = e^{\mu x}, \quad \text{q.e.d.}$$

b) Wir beweisen die Behauptung durch vollständige Induktion nach m .

Induktionsanfang $m = 0$. Dann ist f eine Konstante, die Behauptung folgt also aus Teil a).

Induktionsschritt $(m-1) \rightarrow m$. Nach Hilfssatz 3 ist

$$P(D)(x^m e^{\mu x}) = f_0(x) e^{\mu x}$$

mit einem Polynom f_0 vom Grad m . Es gibt deshalb eine Konstante $c \in \mathbb{C}$, so dass

$$f_1(x) := f(x) - c f_0(x)$$

ein Polynom vom Grad $\leq m-1$ ist. Nach Induktions-Voraussetzung gibt es deshalb ein Polynom g_1 vom Grad $\leq m-1$ mit

$$P(D)(g_1(x) e^{\mu x}) = f_1(x) e^{\mu x}.$$

Mit $g(x) := c x^m + g_1(x)$ gilt dann

$$P(D)(g(x) e^{\mu x}) = f(x) e^{\mu x}, \quad \text{q.e.d.}$$

Beispiele

(15.3) Wir wollen eine spezielle Lösung der Differentialgleichung

$$(D^3 - 2D^2 - 2D + 2)y = 2 \sin x \tag{1}$$

bestimmen. Da $2 \sin x = \operatorname{Re}(-2ie^{ix})$, betrachten wir zunächst die komplexe Gleichung

$$P(D)y = -2ie^{ix}, \quad P(D) = D^3 - 2D^2 - 2D + 2. \tag{2}$$

Da $P(i) = i^3 - 2i^2 - 2i + 2 = 4 - 3i$, besitzt (2) eine spezielle Lösung

$$\psi(x) = \frac{-2i}{P(i)} e^{ix} = \frac{-2i}{4-3i} e^{ix} = \frac{6-8i}{25} e^{ix}.$$

Da alle Koeffizienten von $P(D)$ reell sind, gilt

$$\operatorname{Re}(P(D)\psi(x)) = P(D)(\operatorname{Re}(\psi(x))),$$

also hat (1) die spezielle Lösung

$$\varphi(x) := \operatorname{Re} \psi(x) = \frac{6}{25} \cos x + \frac{8}{25} \sin x,$$

was man direkt durch Einsetzen verifizieren kann.

(15.4) Wir betrachten die Differentialgleichung

$$y''' - y = x.$$

Hier ist $P(D) = D^3 - 1$. Da

$$P(T) = T^3 - 1 = (T - 1)(T^2 + T + 1) = (T - 1)(T - \rho_1)(T - \rho_2),$$

mit $\rho_{1/2} = -\frac{1}{2} \pm \frac{i}{2}\sqrt{3}$, besitzt die homogene Gleichung ein Fundamentalsystem von Lösungen bestehend aus den drei Funktionen

$$\varphi_{1/2}(x) = e^{-x/2} e^{\pm ix\sqrt{3}/2}, \quad \varphi_3(x) = e^x.$$

Die rechte Seite der Differentialgleichung ist xe^{0x} . Da $P(0) \neq 0$, gibt es also eine spezielle Lösung der inhomogenen Gleichung der Gestalt

$$\psi(x) = a + bx, \quad a, b \in \mathbb{C}.$$

Um die Koeffizienten a, b zu bestimmen, setzen wir ψ in die linke Seite der Differentialgleichung ein:

$$(D^3 - 1)(a + bx) = -a - bx.$$

Damit die Differentialgleichung erfüllt ist, muss $a = 0$ und $b = -1$ sein, d.h.

$$\psi(x) := -x$$

ist eine spezielle Lösung der inhomogenen Gleichung.

Satz 4 (Inhomogene Gleichung, Resonanzfall). Sei

$$P(T) = T^n + a_{n-1}T^{n-1} + \dots + a_1T + a_0 \in \mathbb{C}[T]$$

und $f: \mathbb{R} \rightarrow \mathbb{C}$ ein Polynom vom Grad $m \geq 0$. Die Zahl $\mu \in \mathbb{C}$ sei eine k -fache Nullstelle des Polynoms P . Dann besitzt die inhomogene lineare Differentialgleichung

$$P(D)y = f(x)e^{\mu x}$$

eine spezielle Lösung $\psi: \mathbb{R} \rightarrow \mathbb{C}$ der Gestalt

$$\psi(x) = h(x)e^{\mu x} \quad \text{mit einem Polynom } h(x) = \sum_{j=k}^{m+k} c_j x^j.$$

Beweis. Da μ eine k -fache Nullstelle von P ist, gilt

$$P(T) = Q(T)(T - \mu)^k,$$

wobei Q ein Polynom mit $Q(\mu) \neq 0$ ist. (Falls k gleich dem Grad von P ist, ist Q eine von 0 verschiedene Konstante.) Nach Satz 3 gibt es ein Polynom g vom Grad m , so dass

$$Q(D)(g(x)e^{\mu x}) = f(x)e^{\mu x}.$$

Es gibt ein Polynom $h(x) = \sum_{j=k}^{m+k} c_j x^j$ mit $h^{(k)}(x) = g(x)$, also gilt nach Hilfssatz 2

$$(D - \mu)^k(h(x)e^{\mu x}) = h^{(k)}(x)e^{\mu x} = g(x)e^{\mu x}.$$

Zusammenfassend hat man

$$\begin{aligned} P(D)(h(x)e^{\mu x}) &= Q(D)\left((D - \mu)^k(h(x)e^{\mu x})\right) \\ &= Q(D)(g(x)e^{\mu x}) = f(x)e^{\mu x}, \quad \text{q.e.d.} \end{aligned}$$

(15.5) Beispiel. Die Differentialgleichung

$$\frac{d^2 x}{dt^2} + \omega_0^2 x = a \cos \omega t, \quad \omega_0, \omega > 0, \quad a \in \mathbb{R}^*, \quad (3)$$

beschreibt die Schwingung eines harmonischen Oszillators der Eigenfrequenz ω_0 unter der Wirkung einer periodischen äußeren Kraft $a \cos \omega t$ der Frequenz ω . Zur Vereinfachung betrachten wir wieder die komplexe Differentialgleichung

$$\ddot{x} + \omega_0^2 x = a e^{i\omega t}. \quad (4)$$

In diesem Fall ist

$$P(D) = D^2 + \omega_0^2 = (D - i\omega_0)(D + i\omega_0).$$

Um eine spezielle Lösung der inhomogenen Gleichung zu finden, haben wir zwei Fälle zu unterscheiden.

1. Fall: $\omega \neq \omega_0$.

Man erhält eine Lösung von (4) durch den Ansatz $\psi(t) = ce^{i\omega t}$. Es ist

$$P(D)\psi(t) = c(\omega_0^2 - \omega^2)e^{i\omega t},$$

also ist

$$\psi(t) = \frac{a}{\omega_0^2 - \omega^2} e^{i\omega t}$$

eine Lösung von (4) und

$$\varphi(t) = \operatorname{Re} \psi(t) = \frac{a}{\omega_0^2 - \omega^2} \cos \omega t$$

eine Lösung von (3).

2. Fall: $\omega = \omega_0$.

Man nennt diesen Fall den *Resonanzfall*, da die Frequenz der äußeren Kraft gleich der Eigenfrequenz ist.

Wegen $P(i\omega_0) = 0$ besitzt (4) in diesem Fall eine Lösung der Form

$$\psi(t) = ct e^{i\omega_0 t}.$$

Einsetzen ergibt

$$P(D)(ct e^{i\omega_0 t}) = 2ic\omega_0 e^{i\omega_0 t}.$$

Die Differentialgleichung $P(D)x = ae^{i\omega_0 t}$ ist also erfüllt für die Funktion

$$x = \psi(t) = \frac{a}{2i\omega_0} t e^{i\omega_0 t},$$

also besitzt (3) im Resonanzfall die Lösung

$$\psi(t) = \operatorname{Re} \psi(t) = \frac{a}{2\omega_0} t \sin \omega_0 t.$$

Man sieht, dass die Amplitude der Lösung im Resonanzfall für $t \rightarrow \infty$ unbeschränkt wächst (sog. *Resonanzkatastrophe*). Dies gilt für jede Lösung der inhomogenen Gleichung, da alle Lösungen der homogenen Gleichung beschränkt sind.

AUFGABEN

15.1. Man bestimme ein reelles Fundamentalsystem von Lösungen für die folgenden Differentialgleichungen:

a) $y'' - 4y' + 4y = 0,$

b) $y''' - 2y'' + 2y' - y = 0,$

c) $y''' - y = 0,$

d) $y^{(4)} + y = 0.$

15.2. Man bestimme alle Lösungen der folgenden Differentialgleichungen:

a) $y'' + 3y' + 2y = 2,$

b) $y'' - 5y' + 6y = 4xe^x - \sin x,$

c) $y''' - 2y'' + y' = 1 + e^x \cos 2x,$

d) $y^{(4)} + 2y'' + y = 25e^{2x}.$

15.3. Man bestimme alle reellen Lösungen der Differentialgleichung

$$\ddot{x} + 2\mu\dot{x} + \omega_0^2 x = a \cos \omega t, \quad (\omega_0, \omega, \mu \in \mathbb{R}_+^*, a \in \mathbb{R}^*),$$

und untersuche deren asymptotisches Verhalten für $t \rightarrow \infty$.

15.4. Gegeben sei die Differentialgleichung

$$y'' + \frac{a}{x}y' + \frac{b}{x^2}y = 0, \quad (x > 0), \quad (*)$$

wobei $a, b \in \mathbb{C}$ Konstanten seien.

Man zeige: Eine Funktion $\varphi : \mathbb{R}_+^* \rightarrow \mathbb{C}$ ist genau dann Lösung von (*), wenn die Funktion $\psi : \mathbb{R} \rightarrow \mathbb{C}$, definiert durch

$$\psi(x) := \varphi(e^x)$$

Lösung der Differentialgleichung

$$y'' + (a-1)y' + by = 0$$

ist. Man gebe ein Lösungs-Fundamentalsystem von (*) für alle möglichen Parameterwerte $a, b \in \mathbb{C}$ an.

§ 16 Systeme linearer Differentialgleichungen mit konstanten Koeffizienten

Die Lösungstheorie der Systeme von linearen Differentialgleichungen mit konstanten Koeffizienten beruht auf der Eigenwerttheorie von Matrizen. Die explizite Bestimmung eines Lösungs-Fundamentalsystems läuft auf die Transformation der Matrix des Differentialgleichungssystems auf Normalform hinaus.

Bezeichnungen. Da die Lösungen der im Folgenden behandelten Differentialgleichungssysteme oft dynamisch interpretiert werden als Bewegung eines Punktes im n -dimensionalen Raum, bezeichnen wir meist die unabhängige Variable mit t und die abhängigen Variablen mit x_1, \dots, x_n , die wir zu einem n -dimensionalen Spaltenvektor x zusammenfassen. Die Ableitung nach t wird durch einen Strich oder (in physikalischen Anwendungen) durch einen Punkt bezeichnet: $x' = dx/dt = \dot{x}$.

Der nächste Satz sagt, dass jeder Eigenvektor einer quadratischen Matrix A eine Lösung des Differentialgleichungssystems $x' = Ax$ liefert.

Satz 1. Sei $A \in M(n \times n, \mathbb{C})$ eine $n \times n$ -Matrix mit komplexen Koeffizienten und $a \in \mathbb{C}^n$ ein Eigenvektor von A zum Eigenwert $\lambda \in \mathbb{C}$, d.h. $Aa = \lambda a$. Dann ist die Funktion

$$\varphi: \mathbb{R} \rightarrow \mathbb{C}^n, \quad t \mapsto \varphi(t) := ae^{\lambda t}$$

eine Lösung der Differentialgleichung

$$x' = Ax.$$

Beweis. $\varphi'(t) = \lambda ae^{\lambda t} = Aae^{\lambda t} = A\varphi(t)$.

Corollar. Besitzt die Matrix $A \in M(n \times n, \mathbb{C})$ eine Basis $a_1, \dots, a_n \in \mathbb{C}^n$ von Eigenvektoren zu den Eigenwerten $\lambda_1, \dots, \lambda_n \in \mathbb{C}$, so bilden die Funktionen

$$\varphi_k: \mathbb{R} \rightarrow \mathbb{C}^n, \quad \varphi_k(t) := a_k e^{\lambda_k t}, \quad k = 1, \dots, n,$$

ein Fundamentalsystem von Lösungen der Differentialgleichung

$$x' = Ax.$$

Beweis. Die Lösungen $\varphi_1, \dots, \varphi_n$ sind linear unabhängig, da nach Voraussetzung die Vektoren $\varphi_k(0) = a_k$, $k = 1, \dots, n$, linear unabhängig sind.

Bemerkung. Zu einer Matrix $A \in M(n \times n, \mathbb{C})$ gibt es bekanntlich genau dann eine Basis aus Eigenvektoren, wenn eine Matrix $S \in GL(n, \mathbb{C})$ existiert, so dass die Matrix

$$B = S^{-1}AS$$

Diagonalgestalt hat. Die Spalten von S sind die Eigenvektoren, die Diagonalelemente von B die zugehörigen Eigenwerte. Nicht jede Matrix kann so auf Diagonalgestalt transformiert werden. In jedem Fall kann man aber erreichen, dass B sog. Jordansche Normalform hat, d.h. B setzt sich längs der Diagonalen aus Jordan-Kästchen der Gestalt

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

zusammen.

Satz 2. Sei $A \in M(n \times n, \mathbb{C})$ und $S \in GL(n, \mathbb{C})$. Eine Funktion $\varphi: \mathbb{R} \rightarrow \mathbb{C}^n$ ist genau dann Lösung der Differentialgleichung

$$x' = Ax,$$

wenn die Funktion $\psi := S^{-1}\varphi: \mathbb{R} \rightarrow \mathbb{C}^n$ Lösung der Differentialgleichung

$$y' = (S^{-1}AS)y$$

ist.

Man drückt dies auch so aus: Die Differentialgleichung $x' = Ax$ geht durch die Substitution $y := S^{-1}x$ in $y' = (S^{-1}AS)y$ über.

Beweis. Da S invertierbar ist, ist $\varphi'(t) = A\varphi(t)$ gleichbedeutend mit

$$S^{-1}\varphi'(t) = S^{-1}A\varphi(t) = (S^{-1}AS)S^{-1}\varphi(t),$$

d.h. $\psi'(t) = (S^{-1}AS)\psi(t)$.

Bemerkung. Satz 2 bedeutet, dass man die Lösung einer linearen Differentialgleichung $x' = Ax$ auf den Fall zurückführen kann, wo A Normalform hat.

Selbstverständlich gilt entsprechendes auch für ein System 2. Ordnung $x'' = Ax$. Dazu geben wir ein Beispiel aus der Physik.

(16.1) Sei $U: \mathbb{R}^3 \rightarrow \mathbb{R}$ eine zweimal stetig differenzierbare Funktion. Wir betrachten die Differentialgleichung

$$\frac{d^2x}{dt^2} = -\operatorname{grad} U(x) \quad (1)$$

für die gesuchte Funktion

$$x = x(t) = \begin{pmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \end{pmatrix}.$$

Diese Differentialgleichung beschreibt die Bewegung eines Massenpunkts (der Masse 1) unter dem Einfluss eines Potentials U , vgl. (10.5). Wir wollen die Bewegung in einer kleinen Umgebung eines Minimums des Potentials untersuchen. Ist $a \in \mathbb{R}^3$ ein lokales Minimum von U , so gilt

$$\operatorname{grad} U(a) = 0.$$

Daher ist die konstante Funktion $x(t) := a$ für alle $t \in \mathbb{R}$ eine spezielle Lösung von (1), d.h. ein lokales Minimum des Potentials ist ein Gleichgewichtspunkt.

Wir setzen nun weiter voraus, dass die Hessesche Matrix von U im Punkt a ,

$$A := (\operatorname{Hess} U)(a),$$

positiv definit ist. Nach § 7, Corollar 2 zu Satz 2, gilt

$$U(a + \xi) = U(a) + \frac{1}{2} \langle \xi, A\xi \rangle + o(\|\xi\|^2).$$

Nach einer Translation des Koordinatensystems können wir annehmen, dass $a = 0$. Falls sich die Bewegung in einer hinreichend kleinen Umgebung des Gleichgewichtspunkts abspielt, kann das Restglied $o(\|\xi\|^2)$ vernachlässigt werden. Wir wollen deshalb die Differentialgleichung (1) unter der Voraussetzung lösen, dass

$$U(x) = U(0) + \frac{1}{2} \langle x, Ax \rangle.$$

Daraus folgt $\operatorname{grad} U(x) = Ax$, also lautet die Differentialgleichung (1)

$$\frac{d^2x}{dt^2} = -Ax.$$

Die Matrix A ist symmetrisch, daher gibt es eine orthogonale 3×3 -Matrix S ,

so dass

$$S^{-1}AS = B = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix}.$$

Da A positiv definit war, sind alle Eigenwerte $\lambda_k > 0$. Nach der orthogonalen Koordinaten-Transformation $y = S^{-1}x$ geht die Differentialgleichung über in $\ddot{y} = -By$, d.h.

$$\ddot{y}_k = -\lambda_k y_k \quad \text{für } k = 1, 2, 3.$$

Setzt man $\omega_k := \sqrt{\lambda_k}$, so lautet die allgemeine Lösung davon

$$y_k(t) = \alpha_k \cos \omega_k t + \beta_k \sin \omega_k t$$

mit beliebigen Konstanten $\alpha_k, \beta_k \in \mathbb{R}$.

Differentialgleichungs-System in Jordanform

Nach Satz 2 kann man ein lineares Differentialgleichungssystem durch eine geeignete lineare Transformation auf den Fall zurückführen, wo die Matrix Jordansche Normalform hat. Ein solches System zerfällt wiederum in von einander unabhängige Systeme, deren Matrix ein Jordan-Kästchen

$$J(\lambda) = \lambda E + N$$

einer gewissen Dimension m ist. Dabei ist E die m -reihige Einheitsmatrix, $\lambda \in \mathbb{C}$ und N eine $m \times m$ -Matrix, die auf der Nebendiagonalen die Einträge 1 und sonst lauter Nullen hat,

$$N = \begin{pmatrix} 0 & 1 & & 0 \\ & 0 & 1 & \\ & & \ddots & \ddots \\ & & & 0 & 1 \\ 0 & & & & 0 \end{pmatrix}.$$

Wir untersuchen deshalb jetzt ein Differentialgleichungssystem der Gestalt

$$y' = (\lambda E + N)y, \quad y(t) = \begin{pmatrix} y_1(t) \\ \vdots \\ y_m(t) \end{pmatrix}. \quad (2)$$

Diese Gleichung kann durch den Ansatz

$$y(t) := e^{\mathcal{M}t} z(t)$$

noch vereinfacht werden. Da

$$y'(t) = \lambda e^{\lambda t} z(t) + e^{\lambda t} z'(t) = \lambda y(t) + e^{\lambda t} z'(t),$$

gilt $y'(t) = (\lambda E + N)y(t)$ genau dann, wenn

$$z'(t) = Nz(t). \quad (3)$$

Ausgeschrieben lautet das System (3) wie folgt:

$$\begin{cases} z'_1(t) = z_2(t), \\ z'_2(t) = z_3(t), \\ \vdots \\ z'_{m-1}(t) = z_m(t), \\ z'_m(t) = 0. \end{cases}$$

Beginnend mit der letzten Gleichung lässt sich daraus sofort eine Lösung mit der Anfangsbedingung

$$z(0) = c = \begin{pmatrix} c_1 \\ \vdots \\ c_m \end{pmatrix} \in \mathbb{C}^m$$

bestimmen. Man erhält

$$\begin{aligned} z_m(t) &= c_m, \\ z_{m-1}(t) &= c_{m-1} + c_m t, \\ z_{m-2}(t) &= c_{m-2} + c_{m-1} t + c_m \frac{t^2}{2}, \\ &\vdots \\ z_1(t) &= c_1 + c_2 t + c_3 \frac{t^2}{2} + \dots + c_m \frac{t^{m-1}}{(m-1)!}. \end{aligned}$$

Nun ist $y(t) = e^{\lambda t} z(t)$ eine Lösung der Differentialgleichung (2) mit der Anfangsbedingung $y(0) = c$. Die m Lösungen zu den Anfangsbedingungen $y(0) = e_k$, $k = 1, \dots, m$, ($e_k = k$ -ter Einheitsvektor), bilden dann ein Fundamentalsystem von Lösungen. Zusammenfassend können wir formulieren:

Satz 3. Ein Fundamentalsystem $\Phi: \mathbb{R} \rightarrow M(m \times m, \mathbb{C})$ von Lösungen des Differentialgleichungssystems

$$y' = (\lambda E + N)y$$

mit der Anfangsbedingung $\Phi(0) = E$ wird gegeben durch

$$\Phi(t) = e^{\lambda t} \begin{pmatrix} 1 & t & \frac{t^2}{2} & \cdots & \frac{t^{m-2}}{(m-2)!} & \frac{t^{m-1}}{(m-1)!} \\ 0 & 1 & t & \cdots & \frac{t^{m-3}}{(m-3)!} & \frac{t^{m-2}}{(m-2)!} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & t \\ 0 & 0 & 0 & \cdots & 0 & 1 \end{pmatrix}.$$

Beispiel

(16.2) Gegeben sei das zwei-dimensionale Differentialgleichungs-System

$$\frac{dy}{dt} = Ay$$

mit einer reellen 2×2 -Matrix $A \in M(2 \times 2, \mathbb{R})$.

Es können nun folgende drei Fälle eintreten:

1. Fall. Die Matrix A besitzt zwei linear unabhängige reelle Eigenvektoren $a_1, a_2 \in \mathbb{R}^2$ zu den Eigenwerten $\lambda_1, \lambda_2 \in \mathbb{R}$ (nicht notwendig $\lambda_1 \neq \lambda_2$). Dann bilden die Funktionen

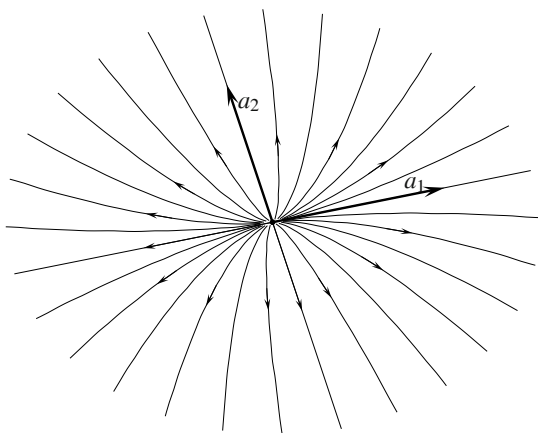
$$\varphi_1(t) = a_1 e^{\lambda_1 t}, \quad \varphi_2(t) = a_2 e^{\lambda_2 t}$$

ein Lösungs-Fundamentalsystem.

Es ist interessant, die Trajektorien der Differentialgleichung im \mathbb{R}^2 zu betrachten, d.h. die Bilder der Lösungskurven

$$\varphi: \mathbb{R} \rightarrow \mathbb{R}^2, \quad t \mapsto \varphi(t) = \alpha_1 \varphi_1(t) + \alpha_2 \varphi_2(t), \quad \alpha_1, \alpha_2 \in \mathbb{R}.$$

Für die Null-Lösung ($\alpha_1 = \alpha_2 = 0$) ergibt sich der Nullpunkt als stationärer Punkt. Falls $\det A \neq 0$, stellen jedoch alle anderen Lösungen injektive Abbildungen $\varphi: \mathbb{R} \rightarrow \mathbb{R}^2$ dar, wie man sich leicht überlegen kann. Aus dem Eindeutigkeitssatz folgt: Haben zwei Trajektorien einen Punkt gemeinsam, sind sie überhaupt identisch. Sind etwa beide Eigenwerte positiv, ergibt sich folgendes Bild 16.1.

**Bild 16.1**

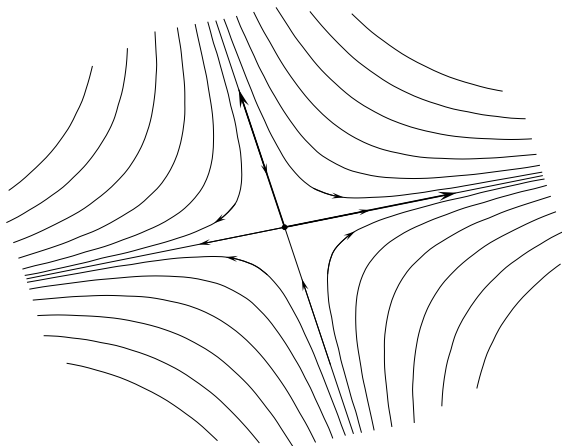
Die positiven Vielfachen der Lösungen φ_1 und φ_2 haben als Trajektorien die Halbstrahlen

$$\mathbb{R}_+^* a_1 = \{\xi a_1 : 0 < \xi < \infty\}$$

bzw. $\mathbb{R}_+^* a_2$. Analog ergeben sich die Halbstrahlen $-\mathbb{R}_+^* a_1$ und $-\mathbb{R}_+^* a_2$ aus negativen Vielfachen von φ_1 bzw. φ_2 . Die anderen Trajektorien sind parabelähnliche Kurven (für $\lambda_1 = \lambda_2$ sind es geradlinige Halbstrahlen).

Sind beide Eigenwerte negativ, ergibt sich das gleiche Bild, nur werden die Trajektorien in entgegengesetzter Richtung durchlaufen.

Ist ein Eigenwert positiv, der andere negativ, sind die Trajektorien hyperbelähnliche Kurven, siehe Bild 16.2.

**Bild 16.2**

2. Fall. Die Matrix A hat zwei konjugiert-komplexe Eigenwerte

$$\lambda_{1,2} = \mu \pm i\omega, \quad \mu \in \mathbb{R}, \quad \omega \in \mathbb{R}^*.$$

Dann sind die zugehörigen Eigenvektoren $c_1, c_2 \in \mathbb{C}^2$ ebenfalls konjugiert-komplex

$$c_{1,2} = a \pm ib, \quad a, b \in \mathbb{R}^2.$$

Da c_1 und c_2 linear unabhängig sind, sind auch a und b linear unabhängig. Aus den komplexen Lösungen

$$\varphi_k(t) = c_k e^{\lambda_k t}, \quad k = 1, 2,$$

lässt sich ein reelles Lösungs-Fundamentalsystem gewinnen:

$$\psi_1(t) := \frac{1}{2}(\varphi_1(t) + \varphi_2(t)) = (a \cos \omega t - b \sin \omega t) e^{\mu t},$$

$$\psi_2(t) := \frac{1}{2i}(\varphi_1(t) - \varphi_2(t)) = (a \sin \omega t + b \cos \omega t) e^{\mu t}.$$

Die Trajektorien der Lösungen haben verschiedene Gestalt, je nachdem der Realteil μ der Eigenwerte verschwindet oder nicht. Im Fall $\mu = 0$ sind die Trajektorien Ellipsen, siehe Bild 16.3, für $\mu \neq 0$ erhält man eine Art logarithmischer Spiralen, siehe Bild 16.4.

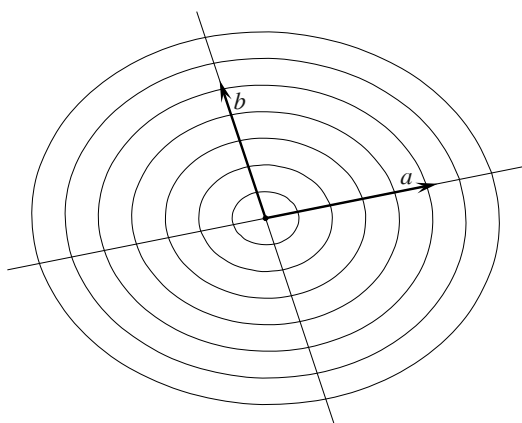


Bild 16.3

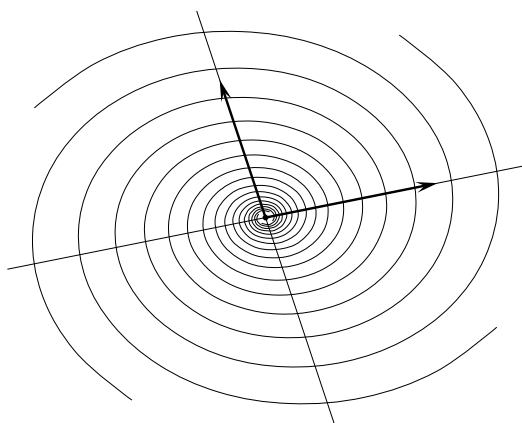


Bild 16.4

3. Fall. Die Matrix A besitzt nur einen Eigenwert $\lambda \in \mathbb{R}$ mit einem eindimensionalen Eigenraum. Dann gibt es eine Matrix $S \in GL(2, \mathbb{R})$ mit

$$B := S^{-1}AS = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}.$$

Durch die Substitution $z = S^{-1}y$ geht die Differentialgleichung über in

$$\frac{dz}{dt} = Bz = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix} z.$$

Nach Satz 3 besitzt dies ein Lösungs-Fundamentalsystem

$$\begin{pmatrix} \psi_1(t) \\ \psi_2(t) \end{pmatrix} = \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix} e^{\lambda t}.$$

Um die ursprüngliche Differentialgleichung $dy/dt = Ay$ zu lösen, hat man die Transformation $z = S^{-1}y$ wieder umzukehren; man erhält das Lösungs-Fundamentalsystem $\varphi_k(t) = S\psi_k(t)$, $k = 1, 2$. Mit

$$S := \begin{pmatrix} v_1 & w_1 \\ v_2 & w_2 \end{pmatrix}, \quad v := \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}, \quad w := \begin{pmatrix} w_1 \\ w_2 \end{pmatrix},$$

wird $\varphi_1(t) = ve^{\lambda t}$, $\varphi_2(t) = (w + tv)e^{\lambda t}$.

Bild 16.5 zeigt die Trajektorien der Lösungen in diesem Fall.

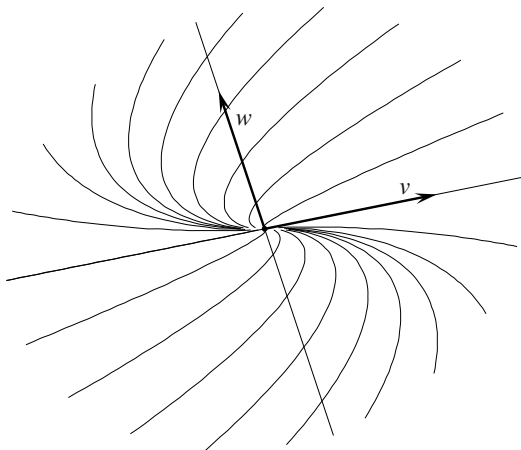


Bild 16.5

AUFGABEN

16.1. Man bestimme ein Fundamentalsystem von Lösungen des Differentialgleichungs-Systems

$$y' = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} y.$$

16.2. Man bestimme ein Fundamentalsystem von Lösungen des Differentialgleichungs-Systems

$$y' = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix} y.$$

16.3. Man bestimme ein Lösungs-Fundamentalsystem für ein System zweier eindimensionaler gekoppelter harmonischer Oszillatoren

$$\begin{cases} \ddot{x} = -\omega^2 x - \gamma(x - y), \\ \ddot{y} = -\omega^2 y + \gamma(x - y). \end{cases}$$

Dabei ist $\omega \in \mathbb{R}_+^*$ die Eigenfrequenz der Oszillatoren und $\gamma \in \mathbb{R}$ die Kopplungskonstante.

Hinweis. Es ist günstig, neue Variable $u := x + y$ und $v := x - y$ einzuführen.

16.4. Sei $U: \mathbb{R}^2 \rightarrow \mathbb{R}$ definiert durch

$$U(x_1, x_2) := \frac{5}{2}x_1^2 + 2x_1x_2 + 4x_2^2.$$

Man bestimme die allgemeine Lösung der Differentialgleichung

$$\frac{d^2x}{dt^2} = -\text{grad } U(x), \quad x = \begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix}.$$

16.5. Man bestimme die Lösung $\varphi: \mathbb{R} \rightarrow \mathbb{R}^2$ der Differentialgleichung

$$y' = \begin{pmatrix} 1 & 2 \\ 3 & 6 \end{pmatrix} y + \begin{pmatrix} x \\ \sin x \end{pmatrix}$$

mit der Anfangsbedingung $\varphi(0) = 0$.

16.6. Sei $A \in M(n \times n, \mathbb{R})$. Man zeige: Die Matrix A ist genau dann schiefsymmetrisch, wenn für jede Lösung $\varphi: \mathbb{R} \rightarrow \mathbb{R}^n$ der Differentialgleichung

$$y' = Ay$$

gilt $\|\varphi(x)\| = \text{const.}$ (d.h. unabhängig von $x \in \mathbb{R}$).

16.7. Man bestimme ein reelles Lösungs-Fundamentalsystem der Differentialgleichung

$$y' = \begin{pmatrix} 0 & -3 & 2 \\ 3 & 0 & -1 \\ -2 & 1 & 0 \end{pmatrix} y.$$

Literaturhinweise

Dieses Buch setzt Vorkenntnisse aus der Differential- und Integralrechnung einer Veränderlichen sowie der Linearen Algebra voraus, wie sie sich z.B. finden in

- [1] *G. Fischer*: Lineare Algebra. Vieweg+Teubner, 17. Aufl. 2009.
- [2] *O. Forster*: Analysis 1. Springer Spektrum, 11. Aufl. 2013
(im Text zitiert als An. 1).

Einige weitere Lehrbücher über Analysis, Topologie und gewöhnliche Differentialgleichungen:

- [3] *B. Aulbach*: Gewöhnliche Differenzialgleichungen. Spektrum Akad. Verlag 2004.
- [4] *M. Barner und F. Flor*: Analysis, Bd. 2. De Gruyter 1999.
- [5] *Th. Bröcker*: Analysis, Bd. 2. Spektrum Akad. Verlag 1995.
- [6] *O. Forster*: Analysis 3. Springer Spektrum, 7. Aufl. 2012.
- [7] *O. Forster und Th. Szymczak*: Übungsbuch zur Analysis 2. Springer Spektrum, 8. Aufl. 2013.
- [8] *H. Heuser*: Lehrbuch der Analysis, Teil 2. Vieweg+Teubner, 14. Aufl. 2008.
- [9] *H. Heuser*: Gewöhnliche Differentialgleichungen. Vieweg+Teubner, 6. Aufl. 2009.
- [10] *K. Jänich*: Topologie. Springer 2005.
- [11] *K. Königsberger*: Analysis 2. Springer 2004.
- [12] *W. Walter*: Gewöhnliche Differentialgleichungen. Springer 2000.

Namens- und Sachverzeichnis

- Abbildung
 stetige, 18
abgeschlossen, 10
abgeschlossene Hülle, 13
abgeschlossene Menge, 8
Ableitung
 partielle, 51
Abstand, 2
Anfangsbedingung, 155
Banach, Stefan (1892 – 1945), 17
Banach-Raum, 17
Banachscher Fixpunktsatz, 92
beschränkt, 17
Bessel, Friedrich Wilhelm (1784–1846), 192
Besselfunktion, 192
Besselsche Differentialgleichung, 192
Bogenlänge, 44
Bolzano, Bernhard (1781–1848), 37
 Satz von B.-Weierstraß, 37
Borel, Emile (1871–1956), 32
Cauchy, Augustin Louis (1789–1857), 16
Cauchy-Folge, 16
definit, positiv (negativ), 83
Diffeomorphismus, 101
Differential, 68
Differentialgleichung
 lineare, 142, 165
 mit getrennten Variablen, 138
differenzierbar, 66
 partiell, 52, 53
 stetig partiell, 69
 total, 66
differenzierbare Kurve, 40
Divergenz, 57
Doppelintegral, 124
Doppelpunkt, 42
Dreiecksungleichung, 1
Durchmesser, 17
ebene Polarkoordinaten, 101
 ε - δ -Kriterium der Stetigkeit, 20
 ε -Umgebung, 5
Euklid (um 300 v.Chr.), 4
euklidische Norm, 4
Euler, Leonhard (1707–1783), 126
Eulersche Differentialgleichung, 126
Fermat, Pierre (1601 – 1655), 3
Fermatsches Prinzip, 3
Fixpunktsatz
 Banachscher, 92
Folge, konvergente, 15
 Cauchy-Folge, 15
Fundamentalsatz der Algebra, 35
Funktional-Matrix, 68
getrennte Variable, 138
gleichmäßig konvergent, 24
gleichmäßig stetig, 37
Gradient, 56
Graph einer Funktion, 51
Hamilton, William Rowan (1805–1860), 131
Hamiltonsches Prinzip, 131
harmonische Funktion, 62
harmonischer Oszillator, 182
Hausdorff, Felix (1868–1942), 6
Hausdorff-Raum, 11
Hausdorffsches Trennungsaxiom, 6
Heine, Eduard (1821–1881), 32

- Heine-Borel, Satz von, 32
Heine-Borelsche Überdeckungseigen-
schaft, 28
Hermite, Charles (1822–1901), 191
Hermiteische Differentialgleichung, 191
Hesse, Otto (1811–1874), 82
Hessesche Matrix, 82
Höhenlinie, 51
Homöomorphismus, 23
homogene Differentialgleichung, 146
homogene lineare Differentialgleichung,
165
Hyperfläche, 107

Immersion, 104
induzierte Metrik, 2
induzierte Topologie, 10
Inneres, 13
 C^1 -invertierbar, 101

Jacobi, Carl Gustav Jakob (1804–
1851), 68
Jacobi-Matrix, 68

Kettenregel, 70
kompakt, 28
konvergent, 15
gleichmäßig, 24
Koordinatensystem
lokales, 107
Kurve, 40
differenzierbare, 40
reguläre, 42
singuläre, 42

Länge einer Kurve, 44
Lagrange, Joseph Louis (1736–1813),
114
Lagrangescher Multiplikator, 114
Laguerre, Edmond (1834–1886), 191

Laguerresche Differentialgleichung,
191
Laplace, Pierre Simon (1736–1813),
61
Laplace-Operator, 61
Lebesgue, Henri (1875–1941), 39
Lebesguesches Lemma, 39
Lindelöf, Ernst (1870–1946), 154
Existenzsatz von Picard-L., 154
Lipschitz, Rudolf (1832–1903), 151
Lipschitz-Bedingung, 151
logarithmische Spirale, 50
lokale Extrema (Maxima, Minima,
82
lokales Koordinatensystem, 107

mathematisches Pendel, 183
Maximum-Norm, 4
Metrik, 1
induzierte, 2
metrischer Raum, 2
Mittelwertsatz, 74
Multiplikator, Lagrangescher, 114

Nabla-Operator, 56
Nebenbedingung
Extremum mit N., 114
Neil, William (1637 – 1670), 43
Neilsche Parabel, 43
Neumann, Carl (1832–1925), 192
Neumannsche Funktion, 192
Newton, Isaac (1643 – 1727), 63
Newton-Potential, 63
Niveaumenge, 51
Norm, 3
einer linearen Abbildung, 26
euklidische, 4
Maximum-Norm, 4
Normalenvektor, 111

- normierter Vektorraum, 3
- offen, 9
- offene Überdeckung, 28
- offener Kern, 13
- orientierungstreu, 48
- Oszillator
 - harmonischer, 182
- Parameterdarstellung, 107
- Parametertransformation, 48
- partiell differenzierbar, 52, 53
- partielle Ableitung, 51
- Pendel
 - mathematisches, 183
- Picard, Emile* (1856–1941), 154
- Picard-Lindelöfsches Iterationsverfahren, 158
- Potentialgleichung, 62
- Prinzip der kleinsten Wirkung, 131
- Produktregel
 - für Divergenz, 57
 - für Gradient, 56
- Rand, 12
- Randpunkt, 12
- reguläre Kurve, 42
- rektifizierbare Kurve, 44
- Relativ-Topologie, 10
- Resonanzfall, 211
- Resonanzkatastrophe, 211
- Richtungsableitung, 72
- Richtungsfeld, 136
- Rotation, 60
- Rotationsfläche, 107
- Schachtelungsprinzip, 18
- Schnittwinkel, 44
- Schraubenlinie, 40
- Schwarz, Hermann Amandus* (1843 – 1921), 59
- Schwarz, Satz von, 59
- singuläre Kurve, 42
- Spirale
 - logarithmische, 50
- stetig, 18, 22
 - gleichmäßig, 37
- Tangentialvektor, 41, 111
- Taylor, Brook* (1685–1731), 78
- Taylorsche Formel, 78
- Topologie, 9
- topologische Abbildung, 23
- topologischer Raum, 9
- Torus, 108
- total differenzierbar, 66
- triviale Metrik, 3
- Überdeckung
 - offene, 28
- Umgebung, 5, 11
- ε -Umgebung, 5
- Untermannigfaltigkeit, 106
- Variation der Konstanten, 145, 172
- Variationsrechnung, 126
- Vektorfeld, 57
- Vektorraum
 - normierter, 3
- vollständig, 17
- Wärmeleitungsgleichung, 62
- Weierstraß, Karl* (1815–1897), 37
- Wellengleichung, 62
- Wirkungsintegral, 131
- Wronski, Josef-Maria* (1778–1853), 175
- Wronski-Determinante, 175
- Zykloide, 47

Symbolverzeichnis

$\mathbb{N} = \{0, 1, 2, 3, \dots\}$	Menge der natürlichen Zahlen
$\mathbb{Z} = \{0, \pm 1, \pm 2, \dots\}$	Menge der ganzen Zahlen
\mathbb{R}	Körper der reellen Zahlen
\mathbb{R}^*	Menge der reellen Zahlen $\neq 0$
\mathbb{R}_+	Menge der reellen Zahlen ≥ 0
\mathbb{R}_+^*	Menge der reellen Zahlen > 0
\mathbb{C}	Körper der komplexen Zahlen
\mathbb{K}	einer der Körper \mathbb{R} oder \mathbb{C}
$M(m \times n, \mathbb{K})$	Vektorraum aller $m \times n$ -Matrizen mit Koeffizienten aus \mathbb{K}
$GL(n, \mathbb{K})$	Gruppe der invertierbaren $n \times n$ -Matrizen mit Koeffizienten aus \mathbb{K}
$\mathbb{C}[T]$	Ring der Polynome in T mit komplexen Koeffizienten
$\operatorname{Re}(z)$	Realteil einer komplexen Zahl z
$\ x\ = \sqrt{x_1^2 + \dots + x_n^2}$	für $x = (x_1, \dots, x_n) \in \mathbb{R}^n$
$\langle x, y \rangle = x_1 y_1 + \dots + x_n y_n$	für $x, y \in \mathbb{R}^n$
$\ \quad\ $	Norm allgemein, 3
$\ x, y\ $	Abstand zweier Punkte in einem metrischen Raum, 2
$\overline{A}, \mathring{A}$	abgeschlossene Hülle, offener Kern von A , 13
∂A	Rand von A , 12
D_i	i -te partielle Ableitung, 52
DF, J_F	Funktionalmatrix, Jacobi-Matrix von F , 68
$\frac{\partial(f_1, \dots, f_m)}{\partial(x_1, \dots, x_m)}$	Funktionalmatrix, 68
$D^\alpha, x^\alpha, \alpha $	Multiindex-Schreibweise, 77
$\operatorname{Hess} f$	Hessesche Matrix, 82
∇	Nabla-Operator, 56
Δ	Laplace-Operator, 61
grad	Gradient, 56
div	Divergenz, 57
rot	Rotation, 60