

Analysis II

—

Vorlesung an der Freien Universität Berlin im
Sommersemester 2010

Alexander Schmitt

Januar 2011



Inhalt

<i>Vorwort</i>	<i>iii</i>
1 <i>Metrische Räume</i>	1
1.1 Metrische Räume	1
1.2 Normierte Vektorräume	2
1.3 Offene Mengen	4
1.4 Topologische Räume	7
1.5 Das Hausdorff-Axiom	8
1.6 Ränder von Mengen	8
2 <i>Konvergenz in metrischen Räumen</i>	11
2.1 Konvergente Folgen und Cauchy-Folgen	11
2.2 Kompakte Mengen	13
2.3 Normen auf \mathbb{R}^n	20
3 <i>Stetige Abbildungen</i>	23
3.1 Definition der Stetigkeit	23
3.2 Neue stetige Abbildungen aus alten	24
3.3 Die ε - δ -Formulierung der Stetigkeit	32
3.4 Stetige Funktionen auf kompakten Mengen	34
3.5 Gleichmäßige Stetigkeit	35
3.6 Folgen stetiger Funktionen	36
3.7 Stetigkeit linearer Abbildungen	38
4 <i>Wege und Kurven in \mathbb{R}^n</i>	43
4.1 Wege	43
4.2 Längenmessung	44
4.3 Differenzierbare Wege	48
4.4 Parametrisierung nach der Bogenlänge	53
5 <i>Partielle Ableitungen</i>	57
5.1 Richtungsableitungen	57
5.2 Der Gradient	61
5.3 Höhere partielle Ableitungen	66
6 <i>Totale Differenzierbarkeit</i>	71
6.1 Differenzierbarkeit	71
6.2 Stetig differenzierbare Funktionen	77
6.3 Die Kettenregel	77

7	<i>Extremwerte differenzierbarer Funktionen</i>	81
7.1	Extrema und Sattelpunkte	81
7.2	Die Taylorformel	84
7.3	Definite Matrizen	87
7.4	Ein hinreichendes Kriterium für isolierte lokale Extrema	89
7.5	Extrema unter Nebenbedingungen	93
7.6	Beispiele	99
8	<i>Der Banachsche Fixpunktsatz</i>	105
8.1	Kontrahierende Abbildungen	105
8.2	Der Banachsche Fixpunktsatz	106
8.3	Nullstellen	107
9	<i>Gewöhnliche Differentialgleichungen</i>	113
9.1	Der Satz von Picard–Lindelöf	113
9.2	Trennung der Variablen	119
9.3	Freier Fall mit Luftwiderstand	122
9.4	Lösung durch Substitution	124
9.5	Lineare Differentialgleichungen	129
9.6	Differentialgleichungen höherer Ordnung	133
10	<i>Der Umkehrsatz</i>	139
10.1	Homöomorphismen, Gebiete und Diffeomorphismen	139
10.2	Umkehrbarkeit differenzierbarer Abbildungen	145
11	<i>Der Satz über implizite Funktionen</i>	153
11.1	Vorbereitende Beispiele	153
11.2	Der Satz über implizite Funktionen	157
11.3	Untermannigfaltigkeiten von \mathbb{R}^n	162
11.4	Tangentialvektoren	167
11.5	Differenzierbare Abbildungen auf Untermannigfaltigkeiten	171
11.6	Extrema unter Nebenbedingungen	173
	Anhang. Abstrakte Mannigfaltigkeiten	176
	<i>Die Aufgabenblätter</i>	181
	<i>Literaturhinweise</i>	203
	<i>Stichwortverzeichnis</i>	205

Vorwort

Das vorliegende Skript enthält den Stoff aus der Vorlesung „Analysis II“ aus dem Sommersemester 2010 und Teilen der Vorlesung „Analysis III“ aus dem Wintersemester 2010-2011. Es basiert im Wesentlichen auf den Lehrbüchern (3), (6) und (7) und soll diese nicht ersetzen.

In der Analysis I haben wir uns vornehmlich mit Funktionen $f: I \rightarrow \mathbb{R}$, $I \subset \mathbb{R}$ ein offenes Intervall, beschäftigt. Zwei Schwerpunkte waren dabei stetige und differenzierbare Funktionen. In vielen Anwendungen ist man gezwungen, Abbildungen $f: \mathbb{R}^m \rightarrow \mathbb{R}^n$ zu betrachten. In der Analysis II wollen wir Stetigkeit und Differenzierbarkeit für solche Abbildungen definieren und untersuchen. Eine Quelle für viele Anwendungen ist die Physik. Wir betrachten zwei Beispiele.

a) Eine Abbildung

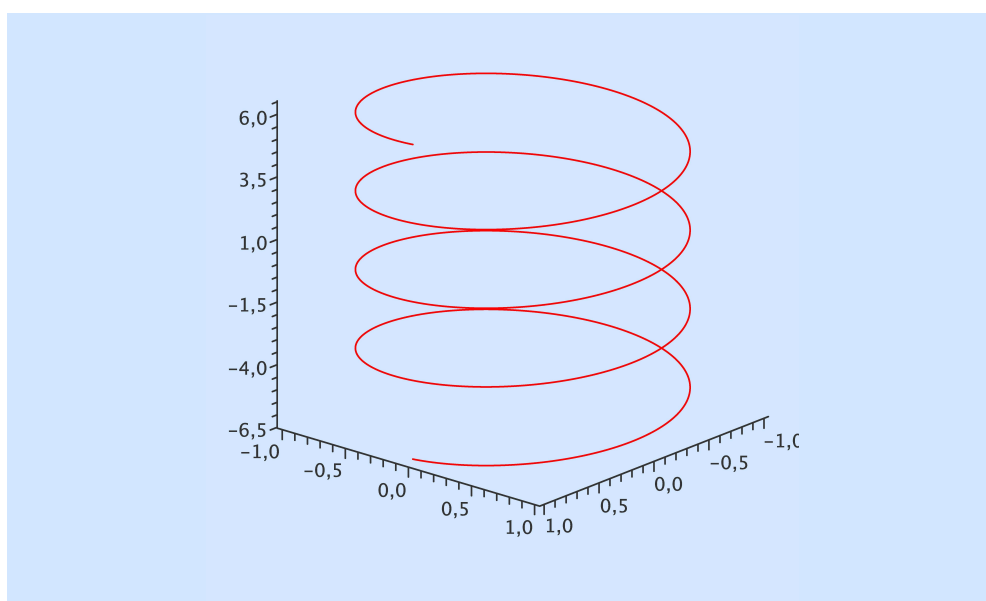
$$\begin{aligned} f: \mathbb{R}^1 &\longrightarrow \mathbb{R}^3 \\ t &\longmapsto (f_1(t), f_2(t), f_3(t)) \end{aligned}$$

kann z.B. die Position eines Teilchens im \mathbb{R}^3 zum Zeitpunkt t beschreiben.

Die Schraubenlinie: Es seien $r > 0$ und $c \neq 0$ fest gewählte reelle Zahlen. Die Abbildung

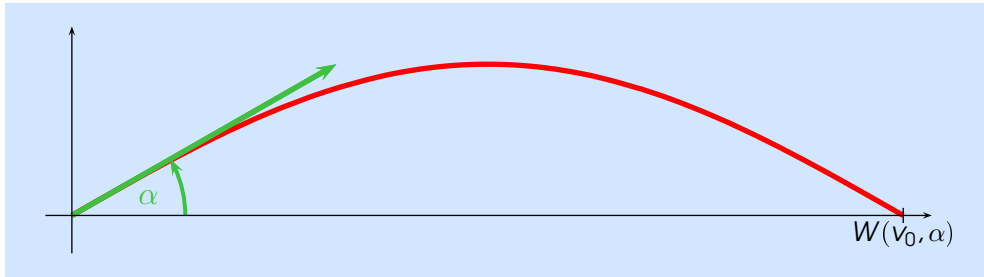
$$\begin{aligned} f: \mathbb{R}^1 &\longrightarrow \mathbb{R}^3 \\ t &\longmapsto (r \cdot \cos(t), r \cdot \sin(t), c \cdot t) \end{aligned}$$

liefert z.B. die folgende Kurve:



Für $c = 0$ erhält man einen Kreis in der (x_1, x_2) -Ebene.

b) **Die Wurfparabel:** Ein Gegenstand wird mit der Geschwindigkeit v_0 unter dem Winkel α gegen die Horizontale abgeworfen. In welchem Abstand trifft der Körper auf dem Boden auf?



Die Wurfweite $W = W(\alpha, v_0)$ hängt von **zwei** Parametern, nämlich α und v_0 , ab. Die genaue Formel (unter idealisierten Bedingungen) lautet:

$$W(\alpha, v_0) = \frac{v_0^2 \cdot \sin(2 \cdot \alpha)}{g}, \quad g \text{ die Erdbeschleunigung.}$$

Die Analysis basiert auf den Begriffen der Konvergenz und des Grenzwerts. In der Analysis I haben wir diese Begriffe mit Hilfe des Absolutbetrags auf den reellen Zahlen erklärt. Allgemeiner sind sie in jedem *metrischen Raum* definiert. Das erste Kapitel führt in die Theorie der metrischen Räume, der normierten Vektorräume und verwandter Konzepte ein.

Das zweite Kapitel bespricht Konvergenz und Grenzwerte in metrischen Räumen. Für die Vorlesung sind die normierten Vektorräume besonders relevant. Dabei ist interessant, dass alle Normen auf \mathbb{R}^n zum selben Konvergenzbegriff führen, während auf unendlichdimensionalen Vektorräumen der Begriff der Konvergenz wirklich von der Norm abhängt.

Sobald man von Grenzwerten sprechen kann, kann man stetige Funktionen wie in der Analysis I definieren. Die grundlegenden Eigenschaften stetiger Funktionen werden im dritten Kapitel entwickelt.

Als erste Verallgemeinerung einer Funktion $f: I \rightarrow \mathbb{R}$ betrachten wir eine Abbildung $F: I \rightarrow \mathbb{R}^n$, *Weg* genannt, $I \subset \mathbb{R}$ ein offenes Intervall. Wir können $F = (F_1, \dots, F_n)$ mit $F_i: I \rightarrow \mathbb{R}$, $i = 1, \dots, n$, schreiben. Wir sagen, dass F *differenzierbar* ist, wenn jede Komponentenfunktion F_i , $i = 1, \dots, n$, differenzierbar ist. Interessant ist hier die Längenmessung solcher Wege.

Im fünften Kapitel beginnen wir mit der Differentialrechnung einer Funktion $f: \mathbb{R}^n \rightarrow \mathbb{R}$. Wir benutzen dafür zunächst Techniken aus der Analysis I, indem wir f auf eine Gerade $G \subset \mathbb{R}^n$ einschränken. Dieser Ansatz führt zum Begriff der *partiellen Differenzierbarkeit*.

Kapitel 6 enthält den Begriff der Differenzierbarkeit, der durch Übertragung der Definition aus der Analysis I in den Kontext von Funktionen mehrerer Veränderlicher unter Verwendung von Normen entsteht. Er wird mit dem Begriff der partiellen Differenzierbarkeit verglichen.

Im siebten Kapitel wird die Differentialrechnung benutzt, um Extremwerte differenzierbarer Funktionen zu untersuchen. Es werden auch Extrema unter einer Nebenbedingung betrachtet.

Das achte Kapitel enthält den Banachschen Fixpunktsatz, eine Aussage über gewisse Selbstabbildungen eines metrischen Raums, in dem jede Cauchy-Folge konvergiert. Eine erste Anwendung ist das Newton-Verfahren für Funktionen mehrerer Variablen.

In der Theorie der Differentialgleichungen versucht man, eine hinreichend oft differenzierbare Funktion einer Veränderlichen aus gewissen Gleichungen,

die die Funktion und ihre Ableitungen enthalten, und Startwerten zu bestimmen. Differentialgleichungen treten in zahlreichen Anwendungen, z.B. in der Physik, auf. Das neunte Kapitel entwickelt die Grundlagen dieser Theorie. Der Banachsche Fixpunktsatz wird dabei in einem unendlichdimensionalen Kontext angewandt, um eine Existenz- und Eindeutigkeitsaussage für die Lösungen gewisser Differentialgleichungen zu erhalten.

In Kapitel 10 wird untersucht, wann sich eine (stetig differenzierbare) Abbildung $f: \mathbb{R}^n \rightarrow \mathbb{R}^n$ zumindest lokal umkehren lässt. Der sogenannte Umkehrsatz ist eine weitere Anwendung des Banachschen Fixpunktsatzes, diesmal in einer endlichdimensionalen Situation.

Im elften Kapitel werden Gleichungssysteme studiert, in denen die Gleichungen durch differenzierbare Funktionen gegeben sind. Als Anwendung des Umkehrsatzes erhält man den Satz über implizite Funktionen, der die lokale Struktur der Lösungsmengen solcher Gleichungssysteme unter geeigneten Voraussetzungen beschreibt. Der Satz über implizite Funktionen führt auf das wichtige Konzept einer *Untermannigfaltigkeit* des \mathbb{R}^n . Mit diesem Konzept kann man die Theorie der Extremwerte differenzierbarer Funktionen unter Nebenbedingungen wesentlich besser verstehen. In einem Anhang wird die Theorie der abstrakten Mannigfaltigkeiten angerissen.

Die Aufgabenblätter zur Vorlesung befinden sich am Ende des Textes.

Ich danke Herrn Nikolai Beck, Herrn Roman Kemmler und Herrn Dr. Norbert Hoffmann für das Korrekturlesen.

Alexander Schmitt
Berlin 2011

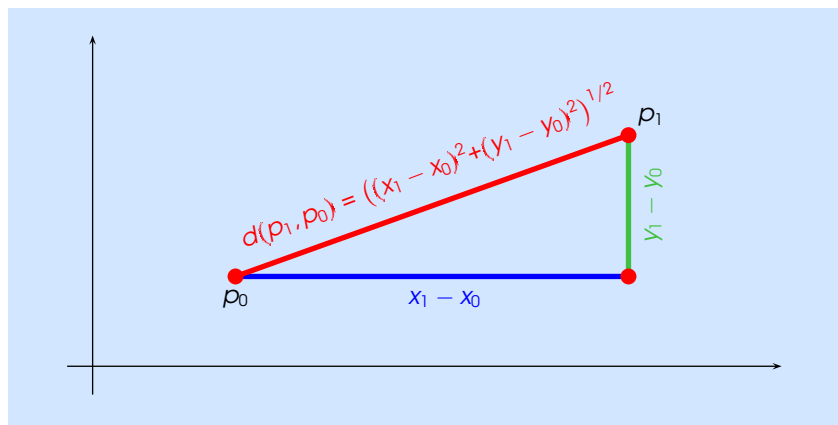
Kapitel 1

Metrische Räume

1.1 Metrische Räume

In der euklidischen¹ Geometrie der Ebene \mathbb{R}^2 verfügen wir über eine Abstandsmessung: Es seien $p_0 = (x_0, y_0)$ und $p_1 = (x_1, y_1)$ Punkte in \mathbb{R}^2 . Der Abstand zwischen p_0 und p_1 ist

$$d(p_1, p_0) := \sqrt{(x_1 - x_0)^2 + (y_1 - y_0)^2}.$$



1.1.1 Eigenschaften. i) $\forall p, q \in \mathbb{R}^2 : d(p, q) = 0 \iff p = q.$

ii) $\forall p, q \in \mathbb{R}^2 : d(p, q) = d(q, p).$ (Symmetrie)

iii) $\forall p, q, r \in \mathbb{R}^2 : d(p, r) \leq d(p, q) + d(q, r).$ (Dreiecksungleichung)

Wie so oft in der Mathematik geschieht nun Folgendes: Nach dem man hinreichend viel Erfahrung, in diesem Fall mit der euklidischen Geometrie der Ebene, gemacht hat, untersucht man kritisch, welche Eigenschaften man wesentlich in den Beweisen benutzt hat. Viele Beweise, vor allem solche, die Konvergenzfragen betreffen, kommen mit den oben aufgelisteten Eigenschaften i) - iii) aus. Aus diesen Eigenschaften macht man dann eine Definition, in unserem Fall die des metrischen Raumes. Man findet viele neue Beispiele von metrischen Räumen. Dabei passieren zwei interessante Dinge: Zum einen gelten alle Gesetzmäßigkeiten aus der euklidischen Geometrie, die sich nur auf die Eigenschaften i) - iii) stützen, in einem beliebigen metrischen Raum, ohne das man sie neu beweisen muss. Zum anderen weisen einige Beispiele metrischer Räume verblüffende Eigenschaften auf, die sich unserer „euklidischen“ Anschauung entziehen.

¹Euklid von Alexandria (ca. 360 v.u.Z. - ca. 280 v.u.Z.), griechischer Mathematiker.

Das Herauskrystallisieren einer guten Definition ist dabei eine der wichtigsten Leistungen in der Mathematik.

1.1.2 Definition. a) Sei X eine Menge. Eine *Metrik (Abstandsfunktion)* auf X ist eine Abbildung

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

mit folgenden Eigenschaften:

i) $\forall x, y \in X: d(x, y) = 0 \iff x = y.$

ii) $\forall x, y \in X: d(x, y) = d(y, x).$

iii) $\forall x, y, z \in X: d(x, z) \leq d(x, y) + d(y, z).$

b) Ein Paar (X, d) , das aus einer Menge X und einer Metrik $d: X \times X \longrightarrow \mathbb{R}$ besteht, heißt *metrischer Raum*.

1.1.3 Beobachtung. Es sei (X, d) ein metrischer Raum. Dann gilt $d(x, y) \geq 0$ für alle $x, y \in X$.

Beweis. Für Punkte $x, y \in X$ haben wir

$$0 = d(x, x) \leq d(x, y) + d(y, x) = 2 \cdot d(x, y).$$

Die erste Gleichung ist in Definition 1.1.2, i), enthalten, die Abschätzung folgt aus iii) und die letzte Gleichung aus ii). \square

1.1.4 Beispiele. i) Die Abbildung

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

ist eine Metrik auf \mathbb{R} .

ii) Es sei X eine Menge. Dann ist

$$\begin{aligned} d_0: X \times X &\longrightarrow \mathbb{R} \\ (x, y) &\longmapsto \begin{cases} 0, & \text{falls } x = y \\ 1, & \text{falls } x \neq y \end{cases} \end{aligned}$$

eine Metrik auf X . Die Metrik d_0 heißt *triviale Metrik*.

iii) Es seien (X, d) ein metrischer Raum und $A \subset X$ eine Teilmenge. Dann ist

$$\begin{aligned} d|_A: A \times A &\longrightarrow \mathbb{R} \\ (a, b) &\longmapsto d(a, b) \end{aligned}$$

eine Metrik auf A , die durch d auf A *induzierte Metrik*.

1.2 Normierte Vektorräume

Ist die zugrundeliegende Menge mit der Struktur eines reellen Vektorraumes ausgestattet, dann betrachtet man meist Längenmessungen, die der Vektorraumstruktur besser angepasst sind, sogenannte *Normen*.

1.2.1 Definition. a) Es sei V (ein möglicherweise unendlichdimensionaler) \mathbb{R} -Vektorraum. Eine *Norm* auf V ist eine Abbildung

$$\begin{aligned} \|\cdot\|: V &\longrightarrow \mathbb{R} \\ v &\longmapsto \|v\|, \end{aligned}$$

so dass gilt:

$$i) \forall v \in V : \|v\| = 0 \iff v = 0.$$

$$ii) \forall \lambda \in \mathbb{R}, v \in V : \|\lambda \cdot v\| = |\lambda| \cdot \|v\|.$$

$$iii) \forall v, w \in V : \|v + w\| \leq \|v\| + \|w\|.$$

b) Ein *normierter Vektorraum* ist ein Paar $(V, \|\cdot\|)$, das aus einem \mathbb{R} -Vektorraum V und einer Norm $\|\cdot\|: V \rightarrow \mathbb{R}$ besteht.

Mit den Definitionen 1.1.2 und 1.2.1 überprüft man sofort:

1.2.2 Lemma. *Es sei $(V, \|\cdot\|)$ ein normierter Vektorraum. Dann ist*

$$\begin{aligned} d: V \times V &\longrightarrow \mathbb{R} \\ (v, w) &\longmapsto \|v - w\| \end{aligned}$$

eine Metrik auf V .

1.2.3 Bemerkung. i) Die Metrik d zu einer Norm $\|\cdot\|$ liefert die Norm über $\|v\| = d(v, 0)$, $v \in V$, zurück.

ii) Es sei V ein \mathbb{R} -Vektorraum. Nicht jede Metrik auf V kommt von einer Norm. Z.B. ist die triviale Metrik d_0 auf einem nichttrivialen Vektorraum nicht durch eine Norm induziert, weil Eigenschaft ii) in Definition 1.2.1 nicht erfüllt sein kann.

1.2.4 Beispiele. i) Die *euklidische Norm* auf \mathbb{R}^n ist wie folgt definiert:

$$\begin{aligned} \|\cdot\|: \mathbb{R}^n &\longrightarrow \mathbb{R} \\ (x_1, \dots, x_n) &\longmapsto \sqrt{x_1^2 + \dots + x_n^2}. \end{aligned}$$

Die euklidische Norm induziert die *euklidische Metrik*

$$\begin{aligned} d: \mathbb{R}^n \times \mathbb{R}^n &\longrightarrow \mathbb{R} \\ ((x_1, \dots, x_n), (y_1, \dots, y_n)) &\longmapsto \sqrt{(x_1 - y_1)^2 + \dots + (x_n - y_n)^2}. \end{aligned}$$

ii) Die *Maximum-Norm* auf \mathbb{R}^n ist durch

$$\begin{aligned} \|\cdot\|_M: \mathbb{R}^n &\longrightarrow \mathbb{R} \\ (x_1, \dots, x_n) &\longmapsto \max\{|x_1|, \dots, |x_n|\}. \end{aligned}$$

festgelegt. Die zugehörige Metrik ist

$$\begin{aligned} d_M: \mathbb{R}^n \times \mathbb{R}^n &\longrightarrow \mathbb{R} \\ ((x_1, \dots, x_n), (y_1, \dots, y_n)) &\longmapsto \max\{|x_1 - y_1|, \dots, |x_n - y_n|\}. \end{aligned}$$

1.2.5 Bemerkung. Man kann die euklidische Norm und die Maximum-Norm miteinander vergleichen: Für alle $p \in \mathbb{R}^n$ gilt

$$\|p\|_M \leq \|p\| \leq \sqrt{n} \cdot \|p\|_M.$$

Wir werden später sehen (Satz 2.3.3), dass man zwei beliebige Normen auf \mathbb{R}^n auf ähnliche Art und Weise miteinander vergleichen kann.

1.2.6 Ein unendlichdimensionales Beispiel. Es sei X eine Menge. Eine Funktion $f: X \rightarrow \mathbb{R}$ heißt *beschränkt*, wenn ein $c \in \mathbb{R}_{\geq 0} := \{r \in \mathbb{R} \mid r \geq 0\}$ existiert, so dass

$$\forall x \in X : |f(x)| \leq c.$$

Dann ist

$$\mathcal{B}(X) = \{f: X \rightarrow \mathbb{R} \mid f \text{ beschränkt}\}$$

ein \mathbb{R} -Vektorraum. Dabei sind die Addition „+“ und die Skalarmultiplikation „ \cdot “ wie folgt erklärt: Für $f, g \in \mathcal{B}(X)$ und $\lambda \in \mathbb{R}$ seien

$$\begin{aligned} f + g: X &\longrightarrow \mathbb{R} \\ x &\longmapsto f(x) + g(x) \\ \lambda \cdot f: X &\longrightarrow \mathbb{R} \\ x &\longmapsto \lambda \cdot f(x). \end{aligned}$$

Bemerkung. Wenn X unendlich viele Elemente enthält, dann ist $\mathcal{B}(X)$ ein unendlichdimensionaler \mathbb{R} -Vektorraum. In der Tat sind die Funktionen

$$\begin{aligned} \delta_y: X &\longrightarrow \mathbb{R} \\ x &\longmapsto \begin{cases} 1, & \text{falls } x = y \\ 0, & \text{sonst} \end{cases}, \quad y \in X, \end{aligned}$$

in $\mathcal{B}(X)$ enthalten und \mathbb{R} -linear unabhängig.

Man setze

$$\begin{aligned} \|\cdot\|: \mathcal{B}(X) &\longrightarrow \mathbb{R} \\ f &\longmapsto \sup\{|f(x)| \mid x \in X\}. \end{aligned}$$

Wir weisen nun nach, dass $\|\cdot\|$ eine Norm auf $\mathcal{B}(X)$ ist. i) Die Implikation $f = 0 \Rightarrow \|f\| = 0$ ist klar. Wenn $\|f\| = 0$ gilt, dann folgt $|f(x)| = 0$, $x \in X$, und f ist die Nullfunktion.

ii) Hier argumentieren wir auf folgende Weise:

$$\begin{aligned} \|\lambda \cdot f\| &= \sup\{|\lambda \cdot f(x)| \mid x \in X\} \\ &= \sup\{|\lambda| \cdot |f(x)| \mid x \in X\} \\ &= |\lambda| \cdot \sup\{|f(x)| \mid x \in X\} \\ &= |\lambda| \cdot \|f\|. \end{aligned}$$

iii) Wir haben

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

1.3 Offene Mengen

Wie wir später sehen werden (Aufgabe A.4.1, b), muss man nur wissen, was die offenen Mengen in einem metrischen Raum sind, um etwa stetige Funktionen zu kennzeichnen. Daher werden wir dieses Konzept jetzt ausführlich erläutern.

1.3.1 Definition. Es sei (X, d) ein metrischer Raum. Für $a \in X$ und $r > 0$ ist

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

der offene Ball vom Radius r um a .

1.3.2 Beispiele. i) Im \mathbb{R}^2 mit der euklidischen Norm $\|\cdot\|$ sind die offenen Bälle die offenen Kreisscheiben (s. Abbildung 1.1).

ii) In \mathbb{R}^2 mit der Maximum-Norm $\|\cdot\|_M$ sind die offenen Bälle (offene) Quadrate (s. Abbildung 1.1).

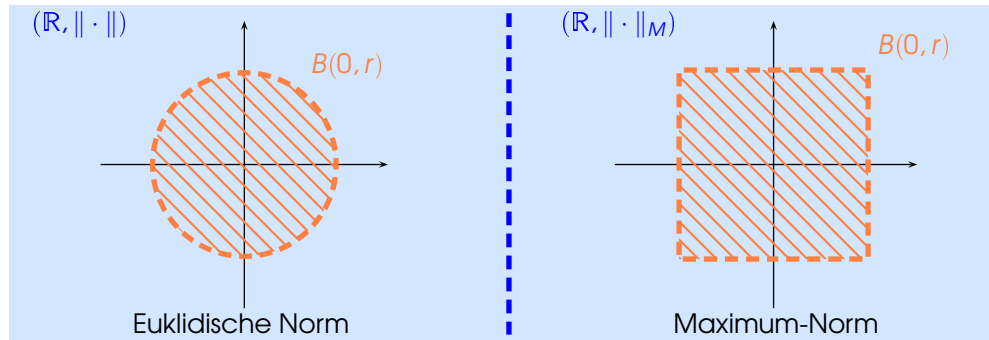


Abbildung 1.1: Bälle in der Ebene bzgl. verschiedener Normen

iii) In einem metrischen Raum (X, d_0) mit trivialer Metrik d_0 sind auch die offenen Bälle nicht besonders interessant: Für $a \in X$ hat man

$$B(a, r) = \begin{cases} \{a\}, & \text{falls } r \leq 1 \\ X, & \text{falls } r > 1 \end{cases}.$$

1.3.3 Definition. Es sei (X, d) ein metrischer Raum.

a) Es seien $U \subset X$ eine Teilmenge und $a \in X$. Die Menge U heißt *Umgebung von a*, wenn eine reelle Zahl $r > 0$ existiert, so dass

$$B(a, r) \subset U.$$

b) Eine Teilmenge $U \subset X$ heißt *offen*, wenn U für jeden Punkt $a \in U$ eine Umgebung von a ist, d.h., wenn zu jedem Punkt $a \in U$ ein $r > 0$ mit $B(a, r) \subset U$ existiert.

c) Eine Teilmenge $Z \subset X$ heißt *abgeschlossen*, wenn ihr Komplement $X \setminus Z$ offen ist.

1.3.4 Beispiele. i) In jedem metrischen Raum (X, d) sind \emptyset und X sowohl offen als auch abgeschlossen.

ii) Es sei (X, d_0) ein metrischer Raum mit trivialer Metrik d_0 . Dann ist jede Menge $U \subset X$ sowohl offen als auch abgeschlossen. (In der Tat gilt für jede Teilmenge $U \subset X$ und jeden Punkt $a \in U$, dass $B(a, 1/2) = \{a\} \subset U$. Nach Definition ist U offen.)

iii) Wir betrachten $(\mathbb{R}, |\cdot|)$, die reellen Zahlen mit dem Absolutbetrag.

- Es seien $a, b \in \mathbb{R}$ mit $a < b$. Dann ist $(a, b) = \{x \in \mathbb{R} \mid a < x < b\}$ offen. (Zu $x \in (a, b)$ wählen wir $0 < \varepsilon < \min\{|a-x|, |b-x|\}$. Aus der Dreiecksungleichung folgt $B(x, \varepsilon) = (x - \varepsilon, x + \varepsilon) \subset (a, b)$.)

- Es sei $a \in \mathbb{R}$. Dann sind die Teilmengen $(a, \infty), (-\infty, a)$ offen.

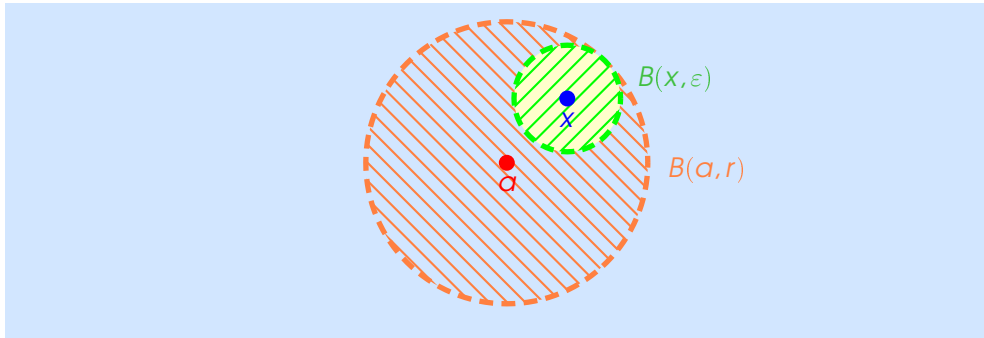
- Es seien $a, b \in \mathbb{R}$ mit $a < b$. Dann ist $[a, b] = \{x \in \mathbb{R} \mid a \leq x \leq b\}$ abgeschlossen. (Das Komplement $X \setminus [a, b] = (-\infty, a) \cup (b, \infty)$ ist offen.)

- Es sei $a \in \mathbb{R}$. Dann sind die Teilmengen $(-\infty, a]$ und $[a, \infty)$ abgeschlossen.

iv) Es seien (X, d) ein metrischer Raum, $a \in X$ und $r > 0$. Wir behaupten, dass $B(a, r)$ eine offene Teilmenge von X ist. Sei $x \in B(a, r)$. Wir setzen $\varepsilon :=$

$r - d(a, x)$. Wir zeigen

$$B(x, \varepsilon) \subset B(a, r).$$



Sei $y \in B(x, \varepsilon)$. Wir finden mit der Dreiecksungleichung

$$d(a, y) \leq d(a, x) + d(x, y) < d(a, x) + \varepsilon = r.$$

Es folgt $y \in B(a, r)$.

1.3.5 Sprechweise. Es seien (X, d) ein metrischer Raum, $x \in X$ und $\varepsilon > 0$. Dann nennt man $B(x, \varepsilon)$ auch ε -Umgebung von x .

Wir fassen einige wichtige Eigenschaften offener Mengen in einem metrischen Raum zusammen:

1.3.6 Beobachtung. Es sei (X, d) ein metrischer Raum.

- i) Die Teilmengen \emptyset und X sind offen.
- ii) Sind $U, V \subset X$ offene Teilmengen, dann ist auch $U \cap V$ eine offene Teilmenge.
- iii) Es seien I eine Indexmenge und $U_i \subset X, i \in I$, offene Teilmengen von X . Dann ist auch die Vereinigungsmenge

$$\bigcup_{i \in I} U_i$$

eine offene Teilmenge von X .

Beweis. Teil i) haben wir schon gesehen.

Teil ii). Für einen Punkt $x \in U \cap V$ existieren positive reelle Zahlen ε_1 und ε_2 , so dass $B(x, \varepsilon_1) \subset U$ und $B(x, \varepsilon_2) \subset V$. Für $\varepsilon := \min\{\varepsilon_1, \varepsilon_2\}$ gilt

$$B(x, \varepsilon) = B(x, \varepsilon_1) \cap B(x, \varepsilon_2) \subset U \cap V.$$

Teil iii). Es sei $x \in \bigcup_{i \in I} U_i$, d.h. es existiert ein $i_0 \in I$ mit $x \in U_{i_0}$. Weil U_{i_0} offen ist, finden wir ein $\varepsilon > 0$, so dass

$$B(x, \varepsilon) \subset U_{i_0} \subset \bigcup_{i \in I} U_i.$$

Damit ist $\bigcup_{i \in I} U_i$ als offene Menge nachgewiesen. □

1.3.7 Bemerkung. Durchschnitte unendlich vieler offener Mengen sind im Allgemeinen **nicht** offen. Man betrachte z.B.

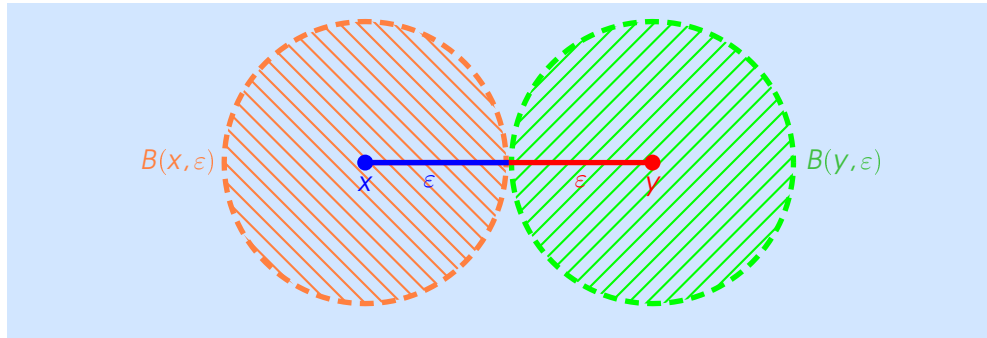
$$\bigcap_{n \geq 1} \left(-\frac{1}{n}, \frac{1}{n}\right) = \{0\}.$$

Die Intervalle $(-1/n, 1/n), n \geq 1$, sind offen, die Menge $\{0\}$ aber nicht.

In metrischen Räumen hat man zusätzlich die folgende „Trennungseigenschaft“:

1.3.8 Satz. Es seien (X, d) ein metrischer Raum und $x \neq y \in X$ zwei verschiedene Punkte. Dann existieren offene Umgebungen U bzw. V von x bzw. y mit

$$U \cap V = \emptyset.$$



Beweis. Es sei $\varepsilon := (1/2) \cdot d(x, y)$. Nach Beispiel 1.3.4 sind $U := B(x, \varepsilon)$ bzw. $V := B(y, \varepsilon)$ offene Umgebungen von x bzw. y . Wir nehmen an, es gebe einen Punkt $z \in X$ mit $z \in U \cap V$.

Dann erhalten wir

$$2 \cdot \varepsilon = d(x, y) \leq d(x, z) + d(y, z) < \varepsilon + \varepsilon = 2 \cdot \varepsilon.$$

Dies führt zu der unmöglichen Ungleichung $\varepsilon < \varepsilon$. Damit muss $U \cap V = \emptyset$ gelten. □

1.4 Topologische Räume

Es stellt sich heraus, dass die in Beobachtung 1.3.6 vorgestellten Eigenschaften offener Mengen in einem metrischen Raum eine gute Definition für eine viel allgemeinere Struktur liefern, die in vielen Bereichen der Mathematik eine große Rolle spielt.

Wir erinnern daran, dass die *Potenzmenge* einer Menge X die Menge

$$\mathcal{P}(X) = \{ W \mid W \subset X \}$$

ist.

1.4.1 Definition. a) Es sei X eine Menge. Eine Teilmenge $\mathcal{T} \subset \mathcal{P}(X)$ der Potenzmenge heißt *Topologie* (= System von offenen Mengen) auf X , wenn gilt:

- i) $\emptyset \in \mathcal{T}, X \in \mathcal{T}$.
- ii) $\forall U \in \mathcal{T}, V \in \mathcal{T}: U \cap V \in \mathcal{T}$.
- iii) Für jede Indexmenge I und jede Familie $U_i \in \mathcal{T}, i \in I$, gilt

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

b) Ein *topologischer Raum* ist ein Paar (X, \mathcal{T}) , das aus einer Menge X und einer Topologie \mathcal{T} auf X besteht.

c) Es sei (X, \mathcal{T}) ein topologischer Raum. Eine Teilmenge $U \subset X$ heißt *offen*, wenn $U \in \mathcal{T}$. Eine Teilmenge $A \subset X$ heißt *abgeschlossen*, wenn $(X \setminus A) \in \mathcal{T}$.

1.4.2 *Beispiele.* i) Für jede Menge X ist $\mathcal{T} = \{\emptyset, X\}$ eine Topologie auf X . Man nennt \mathcal{T} die *Klumpentopologie* auf X .

ii) Für jede Menge X ist $\mathcal{T} = \mathcal{P}(X)$ eine Topologie auf X , die sogenannte *diskrete Topologie*.

iii) In einem metrischen Raum (X, d) ist

$$\mathcal{T} = \{ U \subset X \mid U \text{ ist offen im Sinne von Definition 1.3.3} \}$$

eine Topologie auf X .

iv) Es seien (X, \mathcal{T}) ein topologischer Raum und $A \subset X$ eine Teilmenge. Die Menge

$$\mathcal{T}_A := \{ A \cap U \mid U \in \mathcal{T} \}$$

ist eine Topologie auf A , die *Teilraumtopologie*.

1.4.3 *Bemerkung.* a) Nach Beispiel 1.3.4, ii), ist ii) ein Spezialfall von iii): Die Topologie zur trivialen Metrik auf X ist die diskrete Topologie.

b) Verschiedene Metriken auf einer Menge X können dieselbe Topologie definieren (s. Aufgabe A.1.4).

1.5 Das Hausdorff-Axiom

Allgemeine topologische Räume können sehr merkwürdig aussehen. Sie können Phänomene aufweisen, die sich gänzlich unserer euklidisch geprägten Anschauung entziehen. Daher freut man sich, wenn ein gegebener topologischer Raum zumindest einige vertraute Eigenschaften hat. Z.B. haben wir in Satz 1.3.8 gesehen, dass in einem metrischen Raum eine gewisse Trennungseigenschaft gilt. Wegen ihrer Berühmtheit wollen wir sie in einer Definition festhalten.

1.5.1 Definition. Ein topologischer Raum (X, \mathcal{T}) heißt *hausdorffsch*,² wenn es zu je zwei Punkten $x \neq y \in X$ offene Mengen $U, V \subset X$ mit $x \in U$, $y \in V$ und $U \cap V = \emptyset$ gibt.

1.5.2 *Beispiel.* Es seien (X, d) ein metrischer Raum und \mathcal{T} die zugehörige Topologie. Der topologische Raum (X, \mathcal{T}) ist ein Hausdorff-Raum.

1.6 Ränder von Mengen

In der reellen Analysis einer Veränderlichen haben wir es fast nur mit offenen, abgeschlossenen bzw. halboffenen Intervallen zu tun gehabt. In höherdimensionalen Räumen begegnen wir einem viel größeren Vorrat an wichtigen offenen und abgeschlossenen Mengen und „Dingen dazwischen“. Deshalb beginnen wir hier mit dem Studium der „Feinstruktur“ von Mengen in einem topologischen Raum.

1.6.1 Definition. a) Es seien (X, \mathcal{T}) ein topologischer Raum und $Y \subset X$ eine Teilmenge. Ein Punkt $y \in X$ heißt *Randpunkt* von Y , wenn für jede Umgebung U von y die Bedingung

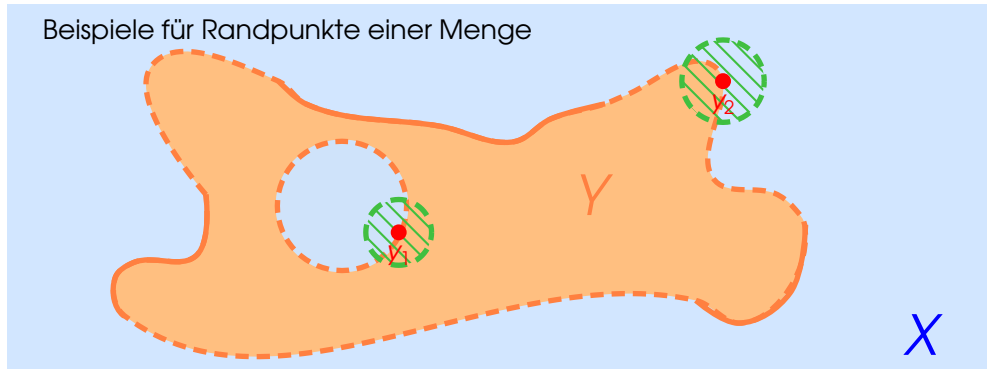
$$U \cap Y \neq \emptyset \quad \text{und} \quad U \cap (X \setminus Y) \neq \emptyset$$

²Felix Hausdorff (1868 - 1942), deutscher Mathematiker. Er wurde von den Nationalsozialisten in den Selbstmord getrieben.

erfüllt ist.

b) Der Rand von Y ist die Menge

$$\partial Y := \{ x \in X \mid x \text{ ist ein Randpunkt von } Y \}.$$



1.6.2 Beispiele. i) In dem normierten Vektorraum $(V, \|\cdot\|)$ sei $Y := B(a, r)$ für ein $a \in V$ und ein $r > 0$. Wir behaupten

$$\partial Y = \{ x \in X \mid d(x, a) = r \}.$$

Die Inklusion „ \subset “ folgt, da ein Punkt $y \in B(a, r)$ eine offene Umgebung U mit $U \subset B(a, r)$ besitzt (Beispiel 1.3.4, iv) und damit kein Randpunkt sein kann. Ein Punkt $y \in V$ mit $\rho := d(a, y) > r$ ist ebenfalls kein Randpunkt. Wir setzen $\varepsilon := \rho - r$. Für jeden Punkt $z \in B(y, \varepsilon)$ gilt nach der Dreiecksungleichung

$$\rho = d(a, y) \leq d(a, z) + d(z, y) < d(a, z) + \varepsilon, \quad \text{so dass } d(a, z) > \rho - \varepsilon = r.$$

Damit gilt $B(y, \varepsilon) \cap B(a, r) = \emptyset$.

Für einen Punkt $y \in V$ mit $d(a, y) = r$ und $\varepsilon > 0$ gilt $y \in (V \setminus B(a, r)) \cap B(y, \varepsilon)$. Es sei z ein Punkt auf der Gerade

$$L := \{ \lambda \cdot a + (1 - \lambda) \cdot y \mid \lambda \in [0, 1] \},$$

die a und y verbindet. Sei $\lambda_0 \in [0, 1]$, so dass $z = \lambda_0 \cdot a + (1 - \lambda_0) \cdot y$. Dann berechnen wir

$$d(a, z) = \| (1 - \lambda_0) \cdot a - (1 - \lambda_0) \cdot y \| = (1 - \lambda_0) \cdot \| a - y \| = (1 - \lambda_0) \cdot d(a, y)$$

und

$$d(z, y) = \| \lambda_0 \cdot a - \lambda_0 \cdot y \| = \lambda_0 \cdot \| a - y \| = \lambda_0 \cdot d(a, y).$$

Wählen wir $\lambda_0 \in (0, 1]$, so dass

$$(0 <) \lambda_0 \cdot d(a, y) < \varepsilon,$$

dann gilt $z \in B(a, r) \cap B(y, \varepsilon)$. Wir haben gezeigt, dass y ein Randpunkt von $B(a, r)$ ist. Man beachte $\partial Y \cap Y = \emptyset$.

ii) Wir betrachten den metrischen Raum $(\mathbb{R}, |\cdot|)$ und die Teilmenge $Y = \mathbb{Q}$. Dann gilt $\partial Y = \mathbb{R}$. Denn für jeden Punkt $a \in \mathbb{R}$ und jedes $\varepsilon > 0$ gilt

$$\begin{aligned} (a - \varepsilon, a + \varepsilon) \cap \mathbb{Q} &\neq \emptyset \\ (a - \varepsilon, a + \varepsilon) \cap (\mathbb{R} \setminus \mathbb{Q}) &\neq \emptyset. \end{aligned}$$

1.6.3 Satz. Es seien (X, \mathcal{T}) ein topologischer Raum und $Y \subset X$ eine Teilmenge.

- i) Die Menge $Y \setminus \partial Y$ ist offen.
- ii) Die Menge $Y \cup \partial Y$ ist abgeschlossen.
- iii) Die Menge ∂Y ist abgeschlossen.

Beweis. i) Sei $y \in Y \setminus \partial Y$. Da $y \notin \partial Y$, existiert eine offene Umgebung U von y , so dass a) $U \cap Y = \emptyset$ oder b) $U \cap (X \setminus Y) = \emptyset$. Wegen $y \in Y$ ist a) unmöglich. Es gilt b), d.h. $U \subset Y$. Es ist noch zu zeigen, dass $U \cap \partial Y = \emptyset$. Für $z \in U \cap \partial Y$ ist U eine offene Umgebung von z . Wegen $z \in \partial Y$ folgt $U \cap (X \setminus Y) \neq \emptyset$. Das hatten wir bereits ausgeschlossen.

ii) Sei $Y' = X \setminus Y$. Es gilt offenbar $\partial Y' = \partial Y$. Nach i) ist $Y' \setminus \partial Y'$ offen. Daher ist

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

abgeschlossen.

iii) In jedem topologischen Raum gilt wegen Definition 1.4.1, a), ii): Sind A und B abgeschlossen, dann ist auch $A \cap B$ abgeschlossen. (Denn man hat $X \setminus (A \cap B) = (X \setminus A) \cup (X \setminus B)$, und diese Menge ist offen.) Wieder sei $Y' := X \setminus Y$. Die Mengen $Y \cup \partial Y$ und $Y' \cup \partial Y' = Y' \cup \partial Y$ sind nach ii) abgeschlossen. Deshalb ist auch

$$(Y \cup \partial Y) \cap (Y' \cup \partial Y) = \partial Y$$

abgeschlossen. □

1.6.4 Definition. Es seien (X, \mathcal{T}) ein topologischer Raum und $Y \subset X$ eine Teilmenge.

a) Die Menge

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

ist der *offene Kern* von Y .

b) Die Menge

$$\bar{Y} := Y \cup \partial Y$$

heißt der *Abschluss* von Y .

Kapitel 2

Konvergenz in metrischen Räumen

2.1 Konvergente Folgen und Cauchy-Folgen

In einem metrischen Raum lassen sich die Begriffe der konvergenten und der Cauchy-Folgen wie in \mathbb{R} formulieren. Die Konvergenz einer Folge in \mathbb{R}^n mit der euklidischen Metrik kann in Termen konvergenter Folgen reeller Zahlen beschrieben werden. In einem allgemeinen metrischen Raum bestehen interessante Verbindungen zwischen den topologischen Eigenschaften von Teilmengen und der Konvergenz gewisser Folgen.

2.1.1 Definition. Es seien (X, d) ein metrischer Raum, $(x_k)_{k \in \mathbb{N}}$ eine Folge von Punkten in X und $a \in X$. Wir sagen, $(x_k)_{k \in \mathbb{N}}$ konvergiert gegen a oder a ist Grenzwert der Folge $(x_k)_{k \in \mathbb{N}}$, wenn gilt:

$$\forall \varepsilon > 0 \exists K \in \mathbb{N} \forall k \geq K : d(a, x_k) < \varepsilon.$$

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

2.1.2 Bemerkung. Die Folge $(x_k)_{k \in \mathbb{N}}$ konvergiert genau dann gegen a , wenn es für jede offene Umgebung U von a ein $K \in \mathbb{N}$ gibt, so dass

$$\forall k \geq K : x_k \in U.$$

2.1.3 Satz. Es seien $X = \mathbb{R}^n$, $\|\cdot\|$ die euklidische Norm und d die zugehörige Metrik. Eine Folge von Punkten $(x_k)_{k \in \mathbb{N}}$ in X mit

$$x_k = (x_{1k}, \dots, x_{nk}), \quad k \in \mathbb{N},$$

konvergiert genau dann gegen $a = (a_1, \dots, a_n) \in \mathbb{R}^n$, wenn

$$\lim_{k \rightarrow \infty} x_{ik} = a_i, \quad i = 1, \dots, n.$$

Damit haben wir Konvergenz von Folgen in \mathbb{R}^n bezüglich der euklidischen Metrik auf die Konvergenz von Folgen in \mathbb{R} zurückgeführt, mit der wir uns im letzten Semester ausführlich beschäftigt haben.

Beweis. a) Wir nehmen an, dass die Folge $(x_k)_{k \in \mathbb{N}}$ gegen a konvergiert. Für $\varepsilon > 0$ existiert ein $K \in \mathbb{N}$, so dass

$$\varepsilon > d(a, x_k) = \sqrt{|a_1 - x_{1k}|^2 + \dots + |a_n - x_{nk}|^2} \geq |a_i - x_{ik}|, \quad i = 1, \dots, n, \quad k \geq K.$$

Es folgt, dass $(x_{ik})_{k \in \mathbb{N}}$ gegen a_i konvergiert, $i = 1, \dots, n$.

b) Es gelte

$$\lim_{k \rightarrow \infty} x_{ik} = a_i, \quad i = 1, \dots, n.$$

Für $\varepsilon > 0$ finden wir ein $K \in \mathbb{N}$ mit

$$|a_i - x_{ik}| < \frac{\varepsilon}{\sqrt{n}}, \quad i = 1, \dots, n, k \geq K.$$

Damit gilt

$$d(a, x_k) = \sqrt{\sum_{i=1}^n |a_i - x_{ik}|^2} \leq \sqrt{n} \cdot \max\{|a_i - x_{ik}| \mid i = 1, \dots, n\} < \varepsilon, \quad k \geq K.$$

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

In metrischen Räumen lässt sich die Abgeschlossenheit einer Menge über Folgen charakterisieren:

2.1.4 Satz. Es seien (X, d) ein metrischer Raum und $A \subset X$ eine Teilmenge. Dann sind folgende Aussagen äquivalent:

i) Die Menge A ist abgeschlossen.

ii) Für jeden Punkt $x \in X$ und jede Folge $(x_k)_{k \in \mathbb{N}}$ in X , so dass $x_k \in A$, $k \in \mathbb{N}$, und $\lim_{k \rightarrow \infty} x_k = x$, gilt $x \in A$.

Beweis. i) \Rightarrow ii): Die Menge $X \setminus A$ ist offen. Seien $x \in X \setminus A$ und $(x_k)_{k \in \mathbb{N}}$ eine Folge mit $x_k \in A$, $k \in \mathbb{N}$, und $\lim_{k \rightarrow \infty} x_k = x$. Es gibt ein $\varepsilon > 0$ mit $B(x, \varepsilon) \subset X \setminus A$, d.h. $B(x, \varepsilon) \cap A = \emptyset$. Auf der anderen Seite gibt es ein K mit $x_k \in B(x, \varepsilon)$ für $k \geq K$, so dass $x_k \notin A$ für $k \geq K$. Das ist aber unmöglich.

ii) \Rightarrow i): Nach Satz 1.6.3, ii), müssen wir $\partial A \subset A$ zeigen. Sei $x \in \partial A$. Wir wählen Punkte

$$x_k \in B\left(x, \frac{1}{k}\right) \cap A, \quad k \in \mathbb{N}.$$

(Die Durchschnitte sind nicht leer, weil x ein Randpunkt von A ist.) Offenbar haben wir $\lim_{k \rightarrow \infty} x_k = x$. Nach Voraussetzung ii) gilt somit $x \in A$. □

2.1.5 Bemerkung. Die Punkte in $\bar{A} = A \cup \partial A$ sind also die Grenzwerte von Folgen in A .

Die in Beispiel 1.6.2, ii), beobachtete Gleichung $\bar{\mathbb{Q}} = \mathbb{R}$ sagt also nur aus, dass man jede reelle Zahl als Grenzwert einer Folge rationaler Zahlen erhalten kann.

2.1.6 Definition. Es sei (X, d) ein metrischer Raum.

a) Eine Folge $(x_k)_{k \in \mathbb{N}}$ in X heißt *Cauchy-Folge*, wenn zu jedem $\varepsilon > 0$ ein $K \in \mathbb{N}$ existiert, so dass

$$\forall k, l \geq K: d(x_k, x_l) < \varepsilon.$$

b) Der metrische Raum (X, d) heißt *vollständig*, wenn jede Cauchy-Folge in X konvergiert.

c) Ein normierter \mathbb{R} -Vektorraum $(V, \|\cdot\|)$ heißt *vollständig*, wenn der zugehörige metrische Raum vollständig ist.

d) Ein vollständiger normierter \mathbb{R} -Vektorraum heißt *Banach-Raum*.

2.1.7 Beispiele. i) Der metrische Raum (\mathbb{R}, d) mit der Metrik $d(x, y) = |x - y|$ ist vollständig. Dies wurde in Analysis I ((12), Satz 2.3.12) nachgewiesen.

ii) Der metrische Raum $(\mathbb{Q}, d_{|\mathbb{Q}})$ ist nicht vollständig.

iii) Jede konvergente Folge in einem metrischen Raum ist eine Cauchy-Folge. (Es sei $x = \lim_{k \rightarrow \infty} x_k$. Für $\varepsilon > 0$ gibt es ein K mit $d(x, x_k) < \varepsilon/2$, $k \geq K$. Für $k, l \geq K$ gilt daher $d(x_k, x_l) \leq d(x_k, x) + d(x, x_l) = d(x, x_k) + d(x, x_l) < \varepsilon/2 + \varepsilon/2 = \varepsilon$.)

2.1.8 Satz. Der Raum $(\mathbb{R}^n, \|\cdot\|)$, $\|\cdot\|$ die euklidische Norm, ist vollständig.

Beweis. Es sei $(x_k)_{k \in \mathbb{N}}$ eine Cauchy-Folge in \mathbb{R}^n . Wir schreiben $x_k = (x_{1k}, \dots, x_{nk})$, $k \in \mathbb{N}$. Sei $\varepsilon > 0$. Dann gibt es ein $K \in \mathbb{N}$, so dass

$$\forall i = 1, \dots, n \forall k, l > K : |x_{ik} - x_{il}| \leq \|x_k - x_l\| < \varepsilon.$$

Folglich ist $(x_{ik})_{k \in \mathbb{N}}$ eine Cauchy-Folge in \mathbb{R} und konvergiert somit gegen eine reelle Zahl x_i , $i = 1, \dots, n$. Aus Satz 2.1.3 folgt

$$\lim_{k \rightarrow \infty} x_k = (x_1, \dots, x_n).$$

Die Folge $(x_k)_{k \in \mathbb{N}}$ ist daher konvergent. □

2.2 Kompakte Mengen

Die kompakten Mengen in einem metrischen Raum spielen vom Standpunkt der Konvergenz von Folgen dieselbe Rolle wie abgeschlossene Intervalle $[a, b] \subset \mathbb{R}$ in den reellen Zahlen.

2.2.1 Definition. Es seien (X, d) ein metrischer Raum und $A \subset X$ eine Teilmenge.

a) Eine Familie $(U_i)_{i \in I}$ von offenen Teilmengen in X heißt (*offene*) *Überdeckung* von A , wenn

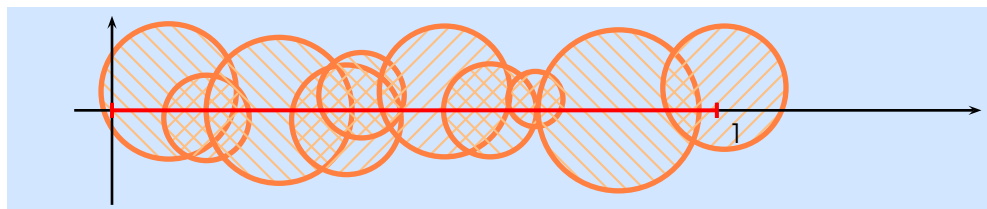
$$A \subset \bigcup_{i \in I} U_i.$$

b) Die Teilmenge A heißt *kompakt*, wenn es zu **jeder** Überdeckung $\bigcup_{i \in I} U_i$ von A eine **endliche Teilmenge** $\{i_1, \dots, i_m\} \subset I$ gibt, so dass

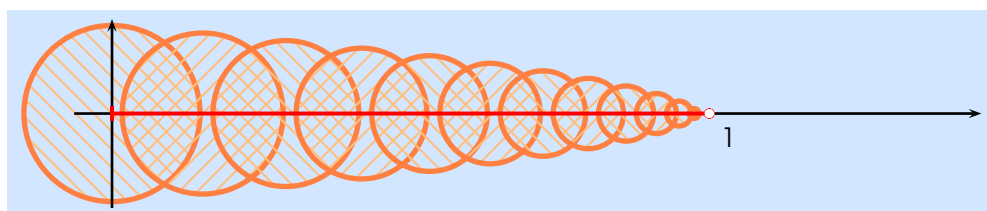
$$A \subset \bigcup_{j=1}^m U_{i_j}.$$

2.2.2 Bemerkung. Es seien (X, τ) ein topologischer Raum und $A \subset X$ eine Teilmenge. Eine (*offene*) *Überdeckung* von A definiert man genau wie in Definition 2.2.1, a). Erfüllt A die Bedingung aus Definition 2.2.1, b), dann nennt man A *quasi-kompakt*. Wenn (X, τ) zusätzlich ein Hausdorff-Raum ist (Definition 1.5.1), dann sagt man, dass A *kompakt* ist.

2.2.3 Beispiele. i) Das abgeschlossene Intervall $I = [0, 1] \subset \mathbb{R}^2$ ist kompakt (s. Satz 2.2.4).



ii) Das halboffene Intervall $I = [0, 1) \subset \mathbb{R}^2$ ist **nicht** kompakt.



iii) Es seien (X, d) ein metrischer Raum und $(x_k)_{k \in \mathbb{N}}$ eine konvergente Folge in X mit Grenzwert a . Dann ist

$$\{x_k \mid k \in \mathbb{N}\} \cup \{a\} =: A$$

kompakt.

Beweis. Es sei $(U_j)_{j \in I}$ eine Überdeckung von A . Dann existiert ein $i_* \in I$ mit $a \in U_{i_*}$. Weiter gibt es ein $K \in \mathbb{N}$, so dass

$$\forall k \geq K + 1: x_k \in U_{i_*}.$$

Außerdem gibt es i_0, \dots, i_K , so dass

$$x_j \in U_{i_j}, \quad j = 0, \dots, K,$$

also

$$A \subset U_{i_*} \cup U_{i_0} \cup U_{i_1} \cup \dots \cup U_{i_K}.$$

iv) Die Teilmenge

$$A := \left\{ \frac{1}{k} \mid k \in \mathbb{N} \setminus \{0\} \right\}$$

ist **nicht** kompakt. (Der Grenzwert 0 fehlt. Man kann also in Beispiel iii) den Punkt a nicht weglassen.)

Beweis. Seien

$$U_1 := \left(\frac{1}{2}, 2 \right),$$

$$U_k := \left(\frac{1}{k+1}, \frac{1}{k-1} \right), \quad k \geq 2.$$

Dann gilt $1/k \in U_k, k \in \mathbb{N} \setminus \{0\}$, so dass

$$A \subset \bigcup_{k \geq 1} U_k.$$

Wegen

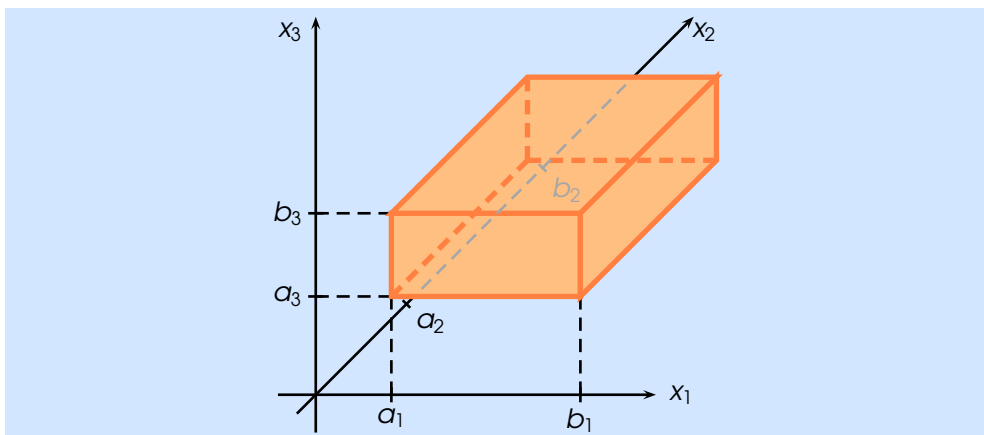
$$U_k \cap A = \left\{ \frac{1}{k} \right\} \quad \text{und} \quad \#A = \infty$$

besitzt die Überdeckung $(U_k)_{k \geq 1}$ keine endliche Teilüberdeckung.

2.2.4 Satz. Gegeben seien $a_i, b_i \in \mathbb{R}$ mit $a_i \leq b_i, i = 1, \dots, n$. Dann ist der abgeschlossene Quader

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

kompakt.



Vorbereitungen

2.2.5 Definition. Es seien (X, d) ein metrischer Raum und $A \subset X$ eine Teilmenge.

a) Die Zahl

$$\text{diam}(A) := \sup\{d(x, y) \mid x, y \in A\} \in \mathbb{R} \cup \{\infty\}$$

heißt der *Durchmesser* von A .

b) Die Menge A heißt *beschränkt*, wenn

$$\text{diam}(A) < \infty$$

gilt.

2.2.6 Beispiel. Aus der Dreiecksungleichung folgt

$$\text{diam}(B(x, r)) \leq 2r, \quad x \in X, r > 0.$$

Für die euklidische Metrik des \mathbb{R}^n gilt „=“.

2.2.7 Bemerkung. Die Menge A ist genau dann beschränkt, wenn $x \in X$ und $r > 0$ mit $A \subset B(x, r)$ existieren (vgl. Beispiel 2.2.6).

2.2.8 Satz (Schachtelungsprinzip). *Es seien (X, d) ein **vollständiger** metrischer Raum und $(A_k)_{k \geq 0}$ eine Familie **beschränkter, nichtleerer abgeschlossener** Teilmengen, so dass für $k \geq 0$*

$$A_{k+1} \subset A_k$$

gilt, d.h. $A_0 \supset A_1 \supset \dots \supset A_k \supset A_{k+1} \supset \dots$, und

$$\lim_{k \rightarrow \infty} \text{diam}(A_k) = 0. \quad (2.1)$$

Dann existiert ein $x \in X$ mit

$$\{x\} = A_\infty := \bigcap_{k \geq 0} A_k.$$

*Beweis. **Schritt 1.** Wir zeigen zunächst, dass A_∞ höchstens ein Element enthält. Seien $x, y \in A_\infty$. Es reicht zu zeigen, dass für alle $\varepsilon > 0$*

$$d(x, y) < \varepsilon$$

gilt. Wegen (2.1) gibt es ein K mit $\text{diam}(A_k) < \varepsilon$, $k \geq K$. Da $x, y \in A_\infty \subset A_k$, $k \geq K$, haben wir

$$d(x, y) \leq \sup\{d(a, b) \mid a, b \in A_k\} = \text{diam}(A_k) < \varepsilon, \quad k \geq K.$$

Schritt 2. Es ist noch zu zeigen, dass A_∞ nicht leer ist. Sei $(x_k)_{k \in \mathbb{N}}$ eine Folge in X mit $x_k \in A_k$, $k \in \mathbb{N}$. Wir beweisen, dass $(x_k)_{k \in \mathbb{N}}$ eine Cauchy-Folge ist. Seien dazu $\varepsilon > 0$ und $K \in \mathbb{N}$, so dass

$$\text{diam}(A_k) < \varepsilon, \quad k \geq K.$$

Für $k, l \geq K$ gilt

$$x_k \in A_k \subset A_K \quad \text{und} \quad x_l \in A_l \subset A_K.$$

Daher findet man

$$d(x_k, x_l) \leq \text{diam}(A_K) < \varepsilon.$$

Da X vollständig ist, konvergiert die Folge $(x_k)_{k \in \mathbb{N}}$. Sei $x := \lim_{k \rightarrow \infty} x_k$. Wir müssen $x \in A_\infty$ zeigen, d.h. $x \in A_\kappa$ für alle $\kappa \in \mathbb{N}$. Für $\kappa \in \mathbb{N}$ konvergiert die Folge $(x_k)_{k \geq \kappa}$ ebenfalls gegen x . Da $x_k \in A_\kappa$ für $k \geq \kappa$ und A_κ abgeschlossen ist, folgt $x \in A_\kappa$ aus Satz 2.1.4. \square

Beweis von Satz 2.2.4. Wir nehmen an, dass der Quader Q nicht kompakt ist. Dann existiert eine Überdeckung $(U_i)_{i \in I}$ von Q , die keine endliche Teilüberdeckung zulässt.

Behauptung. Unter der obigen Annahme existiert eine Folge

$$Q =: Q_0 \supset Q_1 \supset \dots \supset Q_k \supset Q_{k+1} \supset \dots$$

von abgeschlossenen Quadern, so dass für alle $k \in \mathbb{N}$ gilt:

- Die Überdeckung $(U_i)_{i \in I}$ von Q_k besitzt keine endliche Teilüberdeckung.
- $\text{diam}(Q_k) \leq \frac{\text{diam}(Q_0)}{2^k}$.

Diese Behauptung führt sofort zum Widerspruch: Nach dem Schachtelungsprinzip gibt es einen Punkt $x \in Q$ mit

$$\{x\} = \bigcap_{k \geq 0} Q_k.$$

Sei $i_0 \in I$, so dass $x \in U_{i_0}$. Es gibt ein $\varepsilon > 0$, so dass

$$B(x, \varepsilon) \subset U_{i_0}.$$

Sei k , so dass $\text{diam}(Q_k) < \varepsilon$. (Das existiert wegen $\lim_{k \rightarrow \infty} \text{diam}(Q_k) = 0$.) Dann gilt

$$Q_k \subset B(x, \varepsilon) \subset U_{i_0}.$$

Das ist ein Widerspruch zu der ersten Eigenschaft der Folge $(Q_k)_{k \in \mathbb{N}}$.

Beweis der Behauptung. Wir konstruieren die Folge $(Q_k)_{k \in \mathbb{N}}$ durch Rekursion:

- $Q_0 := Q$.
- Q_0, \dots, Q_k seien bereits konstruiert. Wir finden abgeschlossene Intervalle $I_1, \dots, I_n \subset \mathbb{R}$, so dass $Q_k = I_1 \times \dots \times I_n$.

Wir unterteilen I_j in zwei gleichgroße Intervalle:

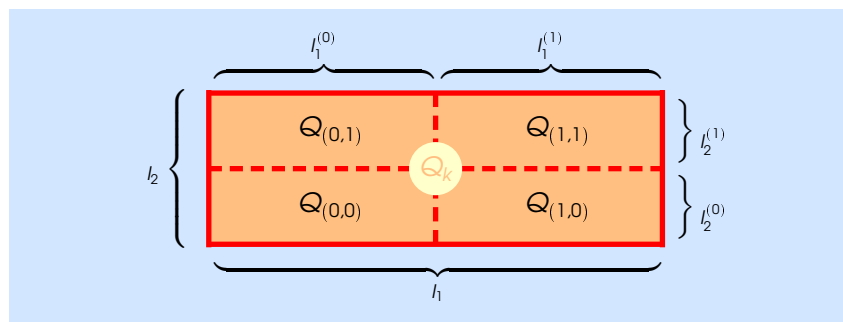
$$I_j = I_j^{(0)} \cup I_j^{(1)}, \quad j = 1, \dots, n.$$

Seien

$$K = \{ \underline{\nu} = (\nu_1, \dots, \nu_n) \mid \nu_j \in \{0, 1\}, j = 1, \dots, n \}$$

und

$$Q_{\underline{\nu}} := I_1^{(\nu_1)} \times \dots \times I_n^{(\nu_n)}, \quad \underline{\nu} = (\nu_1, \dots, \nu_n) \in K.$$



Nach Konstruktion gilt

$$Q_k = \bigcup_{\underline{\nu} \in K} Q_{\underline{\nu}}.$$

Es muss ein Indextupel $\underline{\nu}_0 \in K$ geben, so dass die Überdeckung $(U_i)_{i \in I}$ von $Q_{\underline{\nu}_0}$ keine endliche Teilüberdeckung besitzt.

Wir setzen $Q_{k+1} := Q_{\frac{1}{2^k}}$. Offenbar gilt

$$\text{diam}(Q_{k+1}) \leq \frac{1}{2} \cdot \text{diam}(Q_k) \leq \frac{1}{2} \cdot \frac{\text{diam}(Q_0)}{2^k}.$$

Die so erhaltene Folge $(Q_k)_{k \in \mathbb{N}}$ hat damit die geforderten Eigenschaften. \square

2.2.9 Bemerkung. Einen alternativen Beweis für die Kompaktheit von Quadern, der sich auf die Charakterisierung von Kompaktheit in Satz 2.2.14 stützt, sollen Sie in Aufgabe A.3.3 führen.

Unser nächstes Ziel ist die Charakterisierung kompakter Mengen im \mathbb{R}^n nach Heine¹ und Borel. Diese Charakterisierung beschreibt die schwer zu überprüfende Eigenschaft der Kompaktheit mit zwei einfachen Eigenschaften.

2.2.10 Satz. Sei (X, d) ein metrischer Raum. Eine kompakte Teilmenge $A \subset X$ ist abgeschlossen und beschränkt.

Beweis. **Beschränktheit.** Es sei $a \in A$. Offensichtlich gilt

$$A \subset \bigcup_{k=1}^{\infty} B(a, k).$$

Damit ist $(B(a, k))_{k \in \mathbb{N}}$ eine offene Überdeckung von A . Es gibt natürliche Zahlen $k_1 < \dots < k_m$, so dass

$$A \subset B(a, k_1) \cup \dots \cup B(a, k_m) = B(a, k_m) \quad (\text{s. Bemerkung 2.2.7}).$$

Abgeschlossenheit. Wir zeigen, dass das Komplement $X \setminus A$ offen ist. Sei $x \in X \setminus A$. Wir setzen

$$U_k := \left\{ y \in X \mid d(x, y) > \frac{1}{k} \right\}, \quad k \geq 1.$$

Die Folge $(U_k)_{k \geq 1}$ hat folgende Eigenschaften:

- U_k ist offen, $k \in \mathbb{N}$.
- $\bigcup_{k \geq 1} U_k = X \setminus \{x\} \supset A$.
- $U_k \subset U_l$, $k \leq l$.

Damit ist $(U_k)_{k \in \mathbb{N}}$ eine offene Überdeckung von A . Da A kompakt ist, gibt es natürliche Zahlen $k_1 < \dots < k_m$, so dass

$$A \subset U_{k_1} \cup \dots \cup U_{k_m} = U_{k_m}.$$

Wir sehen, dass $B(x, 1/k_m) \subset X \setminus A$. \square

2.2.11 Satz. Es seien (X, d) ein metrischer Raum, $K \subset X$ eine kompakte und $A \subset X$ eine abgeschlossene Teilmenge. Wenn $A \subset K$ gilt, dann ist A ebenfalls kompakt.

Beweis. Es sei $(U_i)_{i \in I}$ eine offene Überdeckung von A . Die Menge $X \setminus A$ ist offen, und wir haben

$$(X \setminus A) \cup \bigcup_{i \in I} U_i = X \supset K.$$

Weil K kompakt ist, gibt es Indizes $i_1, \dots, i_m \in I$, so dass

$$K \subset X \setminus A \cup U_{i_1} \cup \dots \cup U_{i_m}.$$

¹Heinrich Eduard Heine (1821 - 1881), deutscher Mathematiker.

Ferner haben wir $A \subset K$ und $A \cap (X \setminus A) = \emptyset$. Daher ergibt sich

$$A \subset U_{i_1} \cup \dots \cup U_{i_m},$$

und der Satz ist bewiesen. \square

2.2.12 Satz (Heine–Borel). *In dem normierten Vektorraum $(\mathbb{R}^n, \|\cdot\|)$, $\|\cdot\|$ die euklidische Norm, gilt: Eine Teilmenge $A \subset \mathbb{R}^n$ ist genau dann kompakt, wenn sie abgeschlossen und beschränkt ist.*

Beweis. Auf Grund von Satz 2.2.10 ist lediglich zu zeigen, dass eine abgeschlossene und beschränkte Teilmenge kompakt ist.

Sei

$$d := \text{diam}(A).$$

Da A beschränkt ist, ist d eine nichtnegative reelle Zahl. Wir fixieren $a = (a_1, \dots, a_n) \in A$. Für jeden Punkt $x = (x_1, \dots, x_n) \in A$ folgt

$$|a_i - x_i| \leq \sqrt{|a_1 - x_1|^2 + \dots + |a_n - x_n|^2} = d(a, x) \leq d, \quad i = 1, \dots, n.$$

Damit erkennen wir, dass

$$A \subset Q := [a_1 - d, a_1 + d] \times \dots \times [a_n - d, a_n + d].$$

Gemäß Satz 2.2.4 ist Q kompakt, so dass A nach Satz 2.2.11 kompakt ist. \square

2.2.13 Bemerkung. i) Der Satz von Heine–Borel gilt **nicht** in jedem metrischen Raum. Seien X eine Menge mit unendlich vielen Elementen und d_0 die triviale Metrik auf X . Für jeden Punkt $x \in X$ gilt $X = B(x, 2)$. Die Menge X ist daher abgeschlossen und beschränkt. Wir wissen bereits (Bemerkung 1.4.3), dass jede Teilmenge von X offen ist. Also ist $(U_x)_{x \in X}$ mit $U_x := \{x\}$, $x \in X$, eine offene Überdeckung von X . Sie hat keine endliche Teilüberdeckung, weil wir $\#X = \infty$ vorausgesetzt haben.

ii) In einem normierten Vektorraum $(V, \|\cdot\|)$ gilt der Satz von Heine–Borel genau dann, wenn V endlichdimensional ist (s. Aufgabe A.3.4).

Schließlich charakterisieren wir die Kompaktheit einer Menge über Folgenkonvergenz.

2.2.14 Satz. *Es seien (X, d) ein metrischer Raum und $A \subset X$ eine Teilmenge. Die Menge A ist genau dann kompakt, wenn jede Folge $(x_k)_{k \in \mathbb{N}}$ mit $x_k \in A$, $k \geq 0$, eine konvergente Teilfolge besitzt.*

2.2.15 Bemerkung. Die Implikation „Ist A kompakt, dann besitzt jede Folge in A eine konvergente Teilfolge.“ ist auch als Satz von Bolzano–Weierstraß bekannt. Man vergleiche Satz 2.3.9 in (12).

Beweis. Wir nehmen an, dass $A \subset X$ eine kompakte Teilmenge ist und dass $(x_k)_{k \in \mathbb{N}}$ eine Folge in A ist, die keine konvergente Teilfolge besitzt. (Man beachte, dass $\#\{x_k \mid k \in \mathbb{N}\} = \infty$ gelten muss.) Für jeden Punkt $a \in A$ gibt es dann eine offene Umgebung U_a von a , die nur endlich viele Glieder der Folge $(x_k)_{k \in \mathbb{N}}$ enthält. Wegen $a \in U_a$, $a \in A$, gilt

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

Nun ist A kompakt, so dass es Punkte $a_1, \dots, a_m \in A$ mit

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

gibt. Insbesondere folgt

$$\{x_k \mid k \in \mathbb{N}\} \subset U_{\alpha_1} \cup \dots \cup U_{\alpha_m}.$$

Unsere Wahl der offenen Mengen U_α , $\alpha \in A$, impliziert allerdings

$$\#\left(\{x_k \mid k \in \mathbb{N}\} \cap (U_{\alpha_1} \cup \dots \cup U_{\alpha_m})\right) < \infty.$$

Dieser Widerspruch zeigt, dass unsere Annahme falsch war und jede Folge in A eine konvergente Teilfolge hat.

Jetzt sei $A \subset X$, so dass für jede Folge $(x_k)_{k \in \mathbb{N}}$ mit $x_k \in A$, $k \in \mathbb{N}$, eine konvergente Teilfolge $(x_{k_j})_{j \in \mathbb{N}}$ existiert. Weiter sei $(U_i)_{i \in I}$ eine Überdeckung von A .

Behauptung. *Es gibt ein $\varepsilon > 0$ mit der Eigenschaft:*

$$\forall x \in A \exists i(x) \in I : B(x, \varepsilon) \subset U_{i(x)}.$$

Diese Behauptung begründen wir folgendermaßen: Wenn die Behauptung falsch wäre, dann gäbe es zu $k \geq 1$ einen Punkt $x_k \in A$ mit

$$B\left(x_k, \frac{1}{k}\right) \not\subset U_i, \quad i \in I.$$

Wir bilden die Folge $(x_k)_{k \geq 1}$ in A . Nach Voraussetzung besitzt diese Folge eine konvergente Teilfolge $(x_{k_j})_{j \in \mathbb{N}}$. Sei

$$x := \lim_{l \rightarrow \infty} x_{k_j}.$$

Wir fixieren einen Index $i_0 \in I$ mit $x \in U_{i_0}$ und ein $r > 0$ mit $B(x, r) \subset U_{i_0}$.

Da die Folge $(x_{k_j})_{j \in \mathbb{N}}$ konvergiert, gibt es ein $L \in \mathbb{N}$ mit:

- $x_{k_j} \in B(x, r/2)$, $j \geq L$,
- $1/k_j < r/2$, $j \geq L$.

Für $j \geq L$ und $y \in B(x_{k_j}, 1/k_j)$ schätzen wir ab:

$$d(x, y) \leq d(x, x_{k_j}) + d(x_{k_j}, y) < \frac{r}{2} + \frac{1}{k_j} < r.$$

Wir folgern

$$B\left(x_{k_j}, \frac{1}{k_j}\right) \subset B(x, r) \subset U_{i_0}.$$

Dies widerspricht der Konstruktion der Folge $(x_k)_{k \geq 1}$. Damit ist der Beweis der Behauptung geführt. ✓

Wir argumentieren weiter durch Widerspruch und nehmen an, dass die Überdeckung $(U_i)_{i \in I}$ keine endliche Teilüberdeckung hat.

Sei $\varepsilon > 0$ wie in der Behauptung. Dann ist auch $(B(x, \varepsilon))_{x \in A}$ eine offene Überdeckung von A , die keine endliche Teilüberdeckung zulässt. Ein Element $x_0 \in X$ können wir somit zu einer Folge $(x_k)_{k \in \mathbb{N}}$ ergänzen, so dass

$$x_{k+1} \in X \setminus \bigcup_{i=1}^k B(x_i, \varepsilon), \quad k \geq 1.$$

Es sei $(x_{k_j})_{j \in \mathbb{N}}$ eine konvergente Teilfolge von $(x_k)_{k \in \mathbb{N}}$. Eine konvergente Folge ist eine Cauchy-Folge (Beispiel 2.1.7, iii), so dass ein $L \in \mathbb{N}$ mit

$$d(x_{k_j}, x_{k_m}) < \varepsilon, \quad j, m \geq L,$$

existiert.

Für $l \geq L$ haben wir daher

$$x_{k_{l+1}} \in B(x_{k_l}, \varepsilon).$$

Wegen $k_{l+1} > k_l$ widerspricht dies der Bedingung $x_{k_{l+1}} \notin \bigcup_{i=1}^{k_{l+1}-1} B(x_i, \varepsilon)$. \square

2.3 Normen auf \mathbb{R}^n

Wir wollen Analysis auf dem \mathbb{R}^n betreiben. Für uns wird die Abstandsmessung durch eine Norm gegeben sein. Die Norm führt zu einer Metrik und damit zu einem Konvergenzbegriff. Wir haben bereits Beispiele für verschiedene Normen auf \mathbb{R}^n gesehen. Überraschenderweise führen alle möglichen Normen zu demselben Konvergenzbegriff. Zu diesem Resultat gelangt man über das Konzept der Äquivalenz von Normen.

2.3.1 Definition. Es sei V ein reeller Vektorraum. Zwei Normen $\|\cdot\|_1$ und $\|\cdot\|_2$ auf V sind *äquivalent*, wenn es positive reelle Zahlen c, C gibt, so dass

$$\forall v \in V: \quad c \cdot \|v\|_1 \leq \|v\|_2 \leq C \cdot \|v\|_1.$$

Schreibweise. $\|\cdot\|_1 \sim \|\cdot\|_2$.

2.3.2 Bemerkung. i) Der Sinn und Nutzen dieser Definition ist unmittelbar einleuchtend: Es seien $\|\cdot\|_1$ und $\|\cdot\|_2$ zwei äquivalente Normen auf V . Wenn $(v_k)_{k \in \mathbb{N}}$ eine Folge von Elementen in V ist, dann ist diese Folge genau dann eine Cauchy-Folge bzgl. der Norm $\|\cdot\|_1$, wenn sie eine Cauchy-Folge bzgl. der Norm $\|\cdot\|_2$ ist. Dasselbe gilt für die Konvergenz dieser Folge.

Damit sind $(V, \|\cdot\|_1)$ und $(V, \|\cdot\|_2)$ im Hinblick auf Konvergenzfragen ununterscheidbar. Insbesondere ist $(V, \|\cdot\|_1)$ genau dann vollständig, wenn $(V, \|\cdot\|_2)$ es ist.

Schließlich liefern äquivalente Normen auch dieselbe Topologie (s. Aufgabe A.1.4).

ii) Für „ \sim “ gelten die Axiome einer Äquivalenzrelation: a) Für jede Norm $\|\cdot\|$ gilt $\|\cdot\| \sim \|\cdot\|$. (Man nehme $c = C = 1$.)

b) Für je zwei Normen $\|\cdot\|_1$ und $\|\cdot\|_2$ gilt: $\|\cdot\|_1 \sim \|\cdot\|_2 \Rightarrow \|\cdot\|_2 \sim \|\cdot\|_1$. (Mit $c' := 1/C$ und $C' := 1/c$ haben wir

$$c' \cdot \|v\|_2 \leq c' \cdot C \cdot \|v\|_1 = \|v\|_1 \leq \frac{1}{c'} \cdot \|v\|_2 = C' \cdot \|v\|_2$$

für $v \in V$.)

c) Für drei Normen $\|\cdot\|_1, \|\cdot\|_2$ und $\|\cdot\|_3$ gilt: $\|\cdot\|_1 \sim \|\cdot\|_2$ und $\|\cdot\|_2 \sim \|\cdot\|_3 \Rightarrow \|\cdot\|_1 \sim \|\cdot\|_3$. (Es seien c_1, C_1, c_2, C_2 positive reelle Zahlen, so dass $c_1 \cdot \|v\|_1 \leq \|v\|_2 \leq C_1 \cdot \|v\|_1$ und $c_2 \cdot \|v\|_2 \leq \|v\|_3 \leq C_2 \cdot \|v\|_2$ für jeden Vektor $v \in V$ gilt. Mit $c_3 := c_1 \cdot c_2$ und $C_3 := C_1 \cdot C_2$ erhalten wir

$$c_3 \cdot \|v\|_1 = c_2 \cdot c_1 \cdot \|v\|_1 \leq c_2 \cdot \|v\|_2 \leq \|v\|_3 \leq C_2 \cdot \|v\|_2 \leq C_2 \cdot C_1 \cdot \|v\|_1 = C_3 \cdot \|v\|_1$$

für $v \in V$.)

2.3.3 Satz. Je zwei Normen $\|\cdot\|_1$ und $\|\cdot\|_2$ auf \mathbb{R}^n sind äquivalent.

Beweis. Wir zeigen, dass jede Norm $\|\cdot\|$ auf \mathbb{R}^n zur euklidischen Norm äquivalent ist.

Schritt 1. Es sei $e_i = (0, \dots, 0, 1, 0, \dots, 0) \in \mathbb{R}^n$ der i -te Standardbasisvektor, $i = 1, \dots, n$. Wir setzen

$$C := \|e_1\| + \dots + \|e_n\|.$$

Für $v = v_1 \cdot e_1 + \dots + v_n \cdot e_n \in \mathbb{R}^n$ gilt

$$\begin{aligned} \|v\| &\leq \|v_1 \cdot e_1\| + \dots + \|v_n \cdot e_n\| = |v_1| \cdot \|e_1\| + \dots + |v_n| \cdot \|e_n\| \leq \\ &\leq \max\{|v_i| \mid i = 1, \dots, n\} \cdot C = C \cdot \|v\|_M \leq C \cdot \|v\|. \end{aligned}$$

Schritt 2. Wir nehmen an, dass es keine positive reelle Zahl c gibt, so dass $c \cdot \|v\| \leq \|v\|$ für alle $v \in V$ gilt. In diesem Fall können wir eine Folge $(v_k)_{k \geq 1}$ mit

$$\|v_k\| < \frac{1}{k} \cdot \|v_k\|, \quad k \geq 1,$$

wählen.

Wir setzen

$$w_k := \frac{v_k}{\|v_k\|}, \quad k \geq 1.$$

Für diese Zahlen gilt:

- $\|w_k\| = (1/\|v_k\|) \cdot \|v_k\| = 1, k \geq 1,$
- $\|w_k\| = (1/\|v_k\|) \cdot \|v_k\| < 1/k, k \geq 1.$

Man beachte, dass $(v_k)_{k \in \mathbb{N}}$ eine Nullfolge bezüglich der Norm $\|\cdot\|$ ist.

Auf Grund der ersten Eigenschaft ist die Folge $(w_k)_{k \in \mathbb{N}}$ in der euklidischen Norm beschränkt. Nach dem Satz von Bolzano–Weierstraß 2.2.14 besitzt diese Folge eine bezüglich der euklidischen Norm konvergente Teilfolge. Es gibt also eine in der euklidischen Norm konvergente Folge $(u_l)_{l \in \mathbb{N}}$ mit

- $\|u_l\| = 1, l \in \mathbb{N},$
- $\lim_{l \rightarrow \infty} \|u_l\| = 0.$

Behauptung. Für den Grenzwert $u := \lim_{l \rightarrow \infty} u_l$ bzgl. der **euklidischen** Norm gilt $\|u\| = 0$, d.h. $u = 0$.

Diese Behauptung sehen wir wie folgt ein: Sei $\varepsilon > 0$. Es gibt ein $L \in \mathbb{N}$, so dass

$$\|u_l - u\| < \frac{\varepsilon}{2 \cdot C}, \quad \|u_l\| < \frac{\varepsilon}{2}, \quad l \geq L.$$

Wir fixieren $l_0 \geq L$. Dann sehen wir

$$\|u\| = \|u - u_{l_0} + u_{l_0}\| \leq \|u - u_{l_0}\| + \|u_{l_0}\| < C \cdot \|u - u_{l_0}\| + \frac{\varepsilon}{2} < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Damit gilt $\|u\| < \varepsilon$ für jedes positive ε . Das impliziert $\|u\| = 0$. ✓

Die obige Behauptung ergibt einen Widerspruch: Da $(u_l)_{l \in \mathbb{N}}$ in der euklidischen Norm gegen 0 konvergiert, gibt es z.B. für $\varepsilon = 1/2$ ein L , so dass

$$\frac{1}{2} > \|u_l - u\| = \|u_l - 0\| = \|u_l\| = 1.$$

Das ist offenbar Unsinn. □

2.3.4 Bemerkung. i) Es folgt, dass der Satz von Heine–Borel 2.2.12 ebenso für $(\mathbb{R}^n, \|\cdot\|, \|\cdot\|)$, $\|\cdot\|, \|\cdot\|$ eine beliebige Norm auf \mathbb{R}^n , gilt. In der Tat ist $\|\cdot\|, \|\cdot\|$ nach Satz 2.3.3 zu $\|\cdot\|, \|\cdot\|$ äquivalent. In Aufgabe A.1.4 wird gezeigt, dass beide Normen dieselbe Topologie auf \mathbb{R}^n induzieren. Damit sind die Begriffe „kompakt“ und „abgeschlossen“ für beide Normen gleich. Dass A genau dann bezüglich $\|\cdot\|, \|\cdot\|$ beschränkt ist, wenn es dies bezüglich $\|\cdot\|, \|\cdot\|$ ist, ergibt sich sofort aus der Definition der Äquivalenz von Normen.

ii) In unendlichdimensionalen Vektorräumen gibt es nichtäquivalente Normen (s. Aufgabe A.5.1).

Kapitel 3

Stetige Abbildungen

3.1 Definition der Stetigkeit

Wie in der Analysis einer Veränderlichen ist die Stetigkeit einer Funktion eine notwendige Voraussetzung für die Differenzierbarkeit. Wir erklären jetzt kurz den Begriff einer stetigen Abbildung zwischen metrischen Räumen.

3.1.1 Definition. Es seien (X, d_X) und (Y, d_Y) metrische Räume und $a \in X$ ein Punkt.

a) Eine Abbildung $f: X \rightarrow Y$ heißt *stetig in a* , wenn für jede Folge $(x_k)_{k \in \mathbb{N}}$ in X mit

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

auch die Folge $(f(x_k))_{k \in \mathbb{N}}$ konvergiert und

$$\lim_{k \rightarrow \infty} f(x_k) = f(a)$$

gilt.

b) Die Abbildung $f: X \rightarrow Y$ ist *stetig*, wenn sie in jedem Punkt $a \in X$ stetig ist.

3.1.2 Beispiele. i) Es seien (X, d) ein metrischer Raum und $x_0 \in X$. Dann ist die Abbildung

$$\begin{aligned} f: X &\rightarrow \mathbb{R} \\ x &\mapsto d(x_0, x) \end{aligned}$$

stetig. Dazu seien $a \in X$ und $(y_k)_{k \in \mathbb{N}}$ eine Folge in X , die gegen a konvergiert. Nach Definition heißt das

$$\lim_{k \rightarrow \infty} d(a, y_k) = 0.$$

Wegen der Dreiecksungleichung gilt

$$\forall k \in \mathbb{N} : 0 \leq |d(x_0, a) - d(x_0, y_k)| \leq d(a, y_k).$$

Es folgt

$$\lim_{k \rightarrow \infty} |d(x_0, a) - d(x_0, y_k)| = 0,$$

so dass $\lim_{k \rightarrow \infty} d(x_0, y_k) = d(x_0, a)$.

ii) Für einen normierten Vektorraum $(V, \|\cdot\|)$ folgt, dass die Abbildung

$$\|\cdot\|: V \rightarrow \mathbb{R}.$$

stetig ist.

Auf \mathbb{R} verwenden wir die Metrik $d_{\mathbb{R}}$ mit $d_{\mathbb{R}}(x, y) := |x - y|$, $x, y \in \mathbb{R}$. Für eine Teilmenge $D \subset \mathbb{R}$ erhalten wir die Metrik $d_D := d_{\mathbb{R}|_D}$ durch Einschränkung (s. Beispiel 1.1.4, iii). Wenden wir die obige Definition auf die metrischen Räume (D, d_D) , $(\mathbb{R}, d_{\mathbb{R}})$, eine Funktion $f: D \rightarrow \mathbb{R}$ und einen Punkt $a \in D$ an, dann erhalten wir denselben Begriff, der in (12), Definition 3.4.6, formuliert wurde. Damit liefert die Vorlesung aus dem vorigen Semester eine Fülle von interessanten Beispielen für stetige Abbildungen.

Nun betrachten wir einen metrischen Raum (X, d) (z.B. eine Teilmenge $X \subset \mathbb{R}^m$ mit der durch die euklidische Metrik induzierten Metrik—s. Beispiel 1.1.4, iii) und \mathbb{R}^n mit der euklidischen Metrik.¹ Wir untersuchen eine Abbildung

$$f: X \rightarrow \mathbb{R}^n.$$

Die *Komponentenfunktionen* von f sind wie folgt gegeben:

$$\begin{array}{rcll} f_i: X & \xrightarrow{f} & \mathbb{R}^n & \rightarrow \mathbb{R} \\ x & \mapsto & f(x) & , \quad i = 1, \dots, n. \\ & & (x_1, \dots, x_n) & \mapsto x_i \end{array}$$

Damit können wir

$$\begin{array}{rcl} f: X & \rightarrow & \mathbb{R}^n \\ x & \mapsto & (f_1(x), \dots, f_n(x)). \end{array}$$

schreiben.

3.1.3 Lemma. *In der obigen Situation ist die Funktion f genau dann in a stetig, wenn sämtliche Komponentenfunktionen f_1, \dots, f_n in a stetig sind.*

Beweis. Es sei $(x_k)_{k \in \mathbb{N}}$ eine Folge in X , die gegen a konvergiert. Nach Satz 2.1.3 konvergiert die Folge

$$(f(x_k))_{k \in \mathbb{N}} = (f_1(x_k), \dots, f_n(x_k))_{k \in \mathbb{N}}$$

genau dann gegen $f(a) = (f_1(a), \dots, f_n(a))$, wenn die Folge $(f_i(x_k))_{k \in \mathbb{N}}$ für $i = 1, \dots, n$ gegen $f_i(a)$ konvergiert. Damit ist die Behauptung sofort ersichtlich. \square

3.2 Neue stetige Abbildungen aus alten

Durch Addition, Multiplikation usw. kann man neue stetige Funktionen aus bereits bekannten herstellen. Dies werden wir kurz erläutern und danach einige interessante Abbildungen und Funktionen vorstellen, von denen uns einige mehrmals als nützliche Hilfsmittel wiederbegegnen werden.

3.2.1 Satz. *Es seien (X, d_X) , (Y, d_Y) und (Z, d_Z) metrische Räume, $f: X \rightarrow Y$ und $g: Y \rightarrow Z$ Abbildungen, $a \in X$ ein Punkt und $b := f(a) \in Y$. Wenn die Abbildung f in a stetig ist und die Abbildung g in b , dann ist die Verknüpfung*

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

in a stetig.

Beweis. Der Beweis ist eine leichte Übungsaufgabe. \square

¹Statt mit der euklidischen Metrik können wir auch mit der Metrik zu einer beliebigen Norm auf \mathbb{R}^m bzw. \mathbb{R}^n arbeiten. Dies ergibt sich aus Satz 2.3.3.

3.2.2 Satz. Die folgenden Abbildungen sind stetig:

$$\begin{aligned} \lambda \cdot \text{id}: \mathbb{R} &\longrightarrow \mathbb{R} \\ x &\longmapsto \lambda \cdot x, \quad \lambda \in \mathbb{R}, \\ \text{add}: \mathbb{R}^2 = \mathbb{R} \times \mathbb{R} &\longrightarrow \mathbb{R} \\ (x, y) &\longmapsto x + y, \\ \text{mult}: \mathbb{R}^2 = \mathbb{R} \times \mathbb{R} &\longrightarrow \mathbb{R} \\ (x, y) &\longmapsto x \cdot y, \\ \text{div}: \mathbb{R} \times \mathbb{R} \setminus \{0\} &\longrightarrow \mathbb{R} \\ (x, y) &\longmapsto \frac{x}{y}. \end{aligned}$$

Beweis. Das Resultat folgt aus den Grenzwertsätzen der Analysis (s. (12), Satz 2.2.16). Wir betrachten z.B. die Abbildung div . Es sei $(z_k)_{k \in \mathbb{N}}$ eine Folge in $\mathbb{R} \times \mathbb{R} \setminus \{0\}$ mit $z_k = (x_k, y_k)$, $k \in \mathbb{N}$, die gegen $c = (a, b) \in \mathbb{R} \times \mathbb{R} \setminus \{0\}$ konvergiert. Nach Satz 2.1.3 bedeutet dies, dass $(x_k)_{k \in \mathbb{N}}$ gegen a konvergiert und $(y_k)_{k \in \mathbb{N}}$ gegen b . Wir finden

$$\lim_{k \rightarrow \infty} \text{div}(z_k) = \lim_{k \rightarrow \infty} \frac{x_k}{y_k} = \frac{\lim_{k \rightarrow \infty} x_k}{\lim_{k \rightarrow \infty} y_k} = \frac{a}{b} = \text{div}(c).$$

Daher ist div in c stetig. □

3.2.3 Folgerung. Es seien (X, d) ein metrischer Raum, $f, g: X \rightarrow \mathbb{R}$ zwei Abbildungen, $\lambda \in \mathbb{R}$ und $a \in \mathbb{R}$. Wir nehmen an, dass f und g in a stetig sind. Dann sind auch die folgenden Funktionen stetig in a :

$$\begin{aligned} \lambda \cdot f: X &\longrightarrow \mathbb{R} \\ x &\longmapsto \lambda \cdot f(x), \\ f + g: X &\longrightarrow \mathbb{R} \\ x &\longmapsto f(x) + g(x), \\ f \cdot g: X &\longrightarrow \mathbb{R} \\ x &\longmapsto f(x) \cdot g(x). \end{aligned}$$

Falls zusätzlich $g(x) \neq 0$, $x \in X$, dann ist auch die Funktion

$$\begin{aligned} \frac{f}{g}: X &\longrightarrow \mathbb{R} \\ x &\longmapsto \frac{f(x)}{g(x)} \end{aligned}$$

in a stetig.

Beweis. Wie im Beweis von Lemma 3.1.3 sieht man, dass die Abbildung

$$\begin{aligned} (f, g): X &\longrightarrow \mathbb{R}^2 \\ x &\longmapsto (f(x), g(x)) \end{aligned}$$

in a stetig ist. Wir schreiben

$$f + g: X \xrightarrow{(f, g)} \mathbb{R}^2 \xrightarrow{\text{add}} \mathbb{R}.$$

Die Behauptung folgt jetzt aus Satz 3.2.2 und 3.2.1. □

3.2.4 *Beispiele.* i) Eine Funktion $f: \mathbb{R}^n \rightarrow \mathbb{R}$ heißt *monomiale Funktion vom Grad d* , wenn es natürliche Zahlen k_1, \dots, k_n mit $k_1 + \dots + k_n = d$ gibt, so dass

$$\forall (x_1, \dots, x_n) \in \mathbb{R}^n: f(x_1, \dots, x_n) = x_1^{k_1} \cdot \dots \cdot x_n^{k_n}.$$

Eine monomiale Funktion ist stetig.

ii) Für $n, d \in \mathbb{N}$ sei

$$K(n, d) := \left\{ (k_1, \dots, k_n) \mid k_i \in \mathbb{N}, i = 1, \dots, n, k_1 + \dots + k_n \leq d \right\}.$$

Sei $d \geq 1$. Eine Funktion $f: \mathbb{R}^n \rightarrow \mathbb{R}$ ist eine *polynomiale Funktion vom Grad d* , wenn reelle Zahlen $c_{\underline{k}}$, $\underline{k} \in K(n, d)$, existieren, so dass

$$\forall (x_1, \dots, x_n) \in \mathbb{R}^n: f(x_1, \dots, x_n) = \sum_{\underline{k}=(k_1, \dots, k_n) \in K(n, d)} c_{\underline{k}} \cdot x_1^{k_1} \cdot \dots \cdot x_n^{k_n}$$

und es mindestens einen Index $\underline{k} \in K(n, d) \setminus K(n, d-1)$ mit $c_{\underline{k}} \neq 0$ gibt. Polynomiale Funktionen sind stetig.

iii) Wir betrachten $(\mathbb{R}^n, \|\cdot\|)$, $\|\cdot\|$ die euklidische Norm. Die *Inversion an der Einheitsphäre* ist die Abbildung

$$\begin{aligned} f: \mathbb{R}^n \setminus \{0\} &\longrightarrow \mathbb{R}^n \\ x &\longmapsto \frac{x}{\|x\|^2}. \end{aligned}$$

Eigenschaften.

- Die Funktion f ist stetig (Beispiel 3.1.2, i), und Folgerung 3.2.3).
- Für jedes $x \in \mathbb{R}^n \setminus \{0\}$ liegen x , $f(x)$ und der Ursprung 0 auf einer Gerade.
- Für jeden Punkt $x \in \mathbb{R}^n \setminus \{0\}$ gilt $\|f(x)\| = 1/\|x\|$.
- Für jeden Punkt $x \in \mathbb{R}^n \setminus \{0\}$ gilt $f(x) \in \mathbb{R}^n \setminus \{0\}$, so dass wir $f \circ f$ bilden können. Dabei gilt $f \circ f = \text{id}_{\mathbb{R}^n \setminus \{0\}}$. Man sagt, dass f eine *Involution* ist.

Die zweite und die dritte Eigenschaft bestimmen offenbar die Abbildung f . Die *Einheitssphäre in \mathbb{R}^n* ist

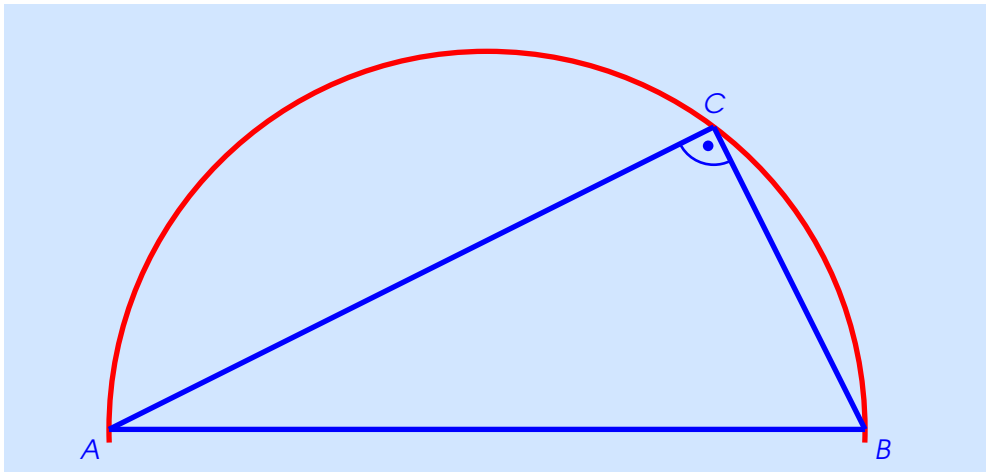
$$S^{n-1} := \{x \in \mathbb{R}^n \mid x_1^2 + \dots + x_n^2 = 1\}.$$

Die Abbildung f hält die Punkte auf S^{n-1} fest, bildet das Innere von S^{n-1} , d.h. $B(0, 1)$, auf das Äußere von S^{n-1} , d.h. $\mathbb{R}^n \setminus \overline{B(0, 1)}$, ab und umgekehrt.

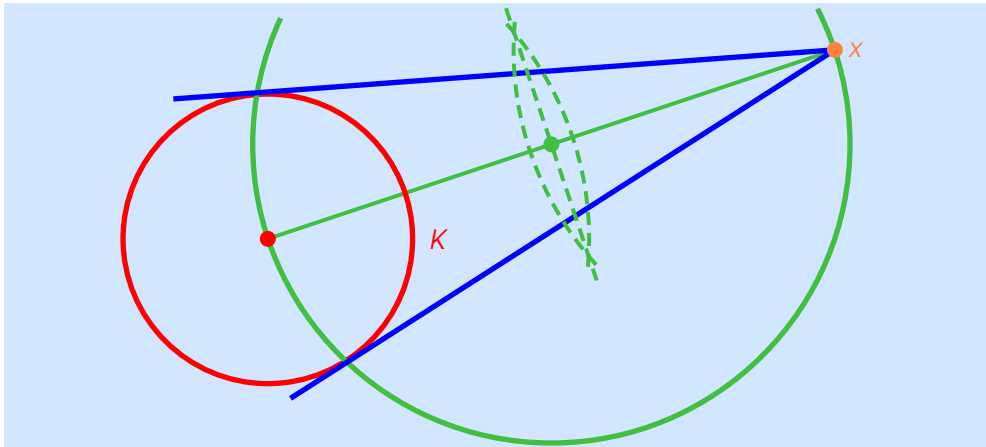
In der Ebene kann der Funktionswert $f(x)$ aus x mit Zirkel und Lineal konstruiert werden. Dazu brauchen wir zwei klassische Sätze aus der Geometrie.

Satz des Thales². *Wenn für ein Dreieck mit den Eckpunkten $A, B, C \in \mathbb{R}^2$ der Punkt C auf dem Kreis mit der Strecke zwischen A und B als Durchmesser liegt, dann hat dieses Dreieck bei C einen rechten Winkel.*

²Thales von Milet (um 624 v.u.Z. - um 546 v.u.Z.), griechischer Naturphilosoph, Staatsmann, Mathematiker, Astronom und Ingenieur.

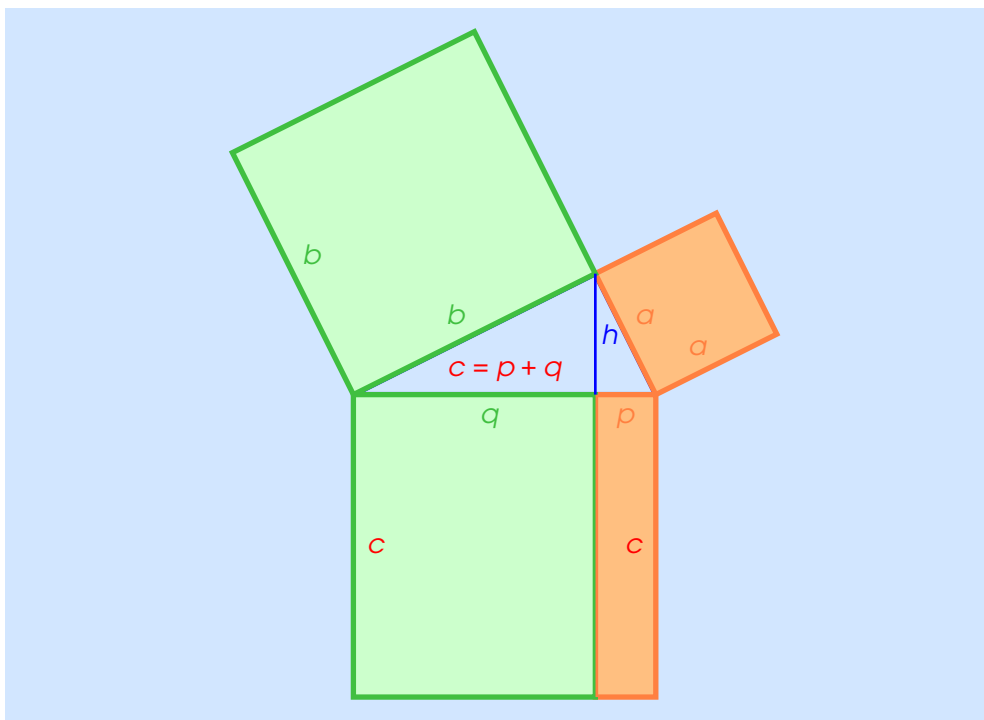


Damit können wir in der Ebene die Tangenten von einem Punkt $x \in \mathbb{R}^2$ an einen Kreis K konstruieren, wenn x nicht auf der durch K begrenzten Kreisscheibe liegt.



Der Kathetensatz des Euklid. Gegeben sei ein rechtwinkliges Dreieck mit der Hypotenuse c und den Katheten a und b . Die Höhe h über der Hypotenuse teilt c in die Abschnitte p bzw. q über a bzw. b . Es gelten dann folgende Identitäten:

- $a^2 = p \cdot c$.
- $b^2 = q \cdot c$.



Beweis. Aus dem Satz des Pythagoras ergeben sich folgende Gleichungen:

$$a^2 + b^2 = c^2 = p^2 + 2pq + q^2 \quad (3.1)$$

$$p^2 + h^2 = a^2 \quad (3.2)$$

$$q^2 + h^2 = b^2. \quad (3.3)$$

Addition von (3.2) und (3.3) ergibt $a^2 + b^2 = p^2 + q^2 + 2h^2$. Durch Vergleich mit (3.1) schließt man

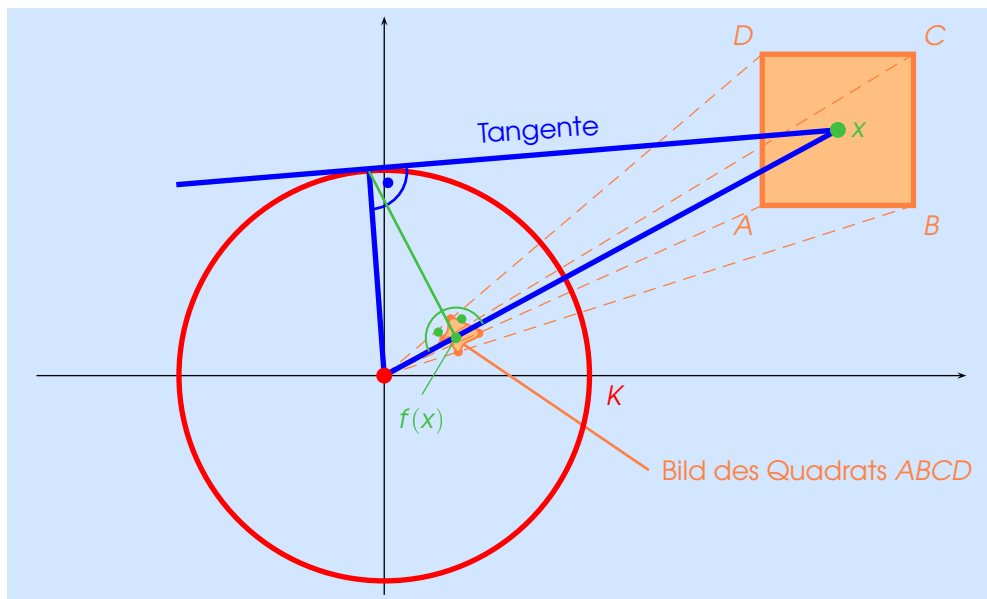
$$h^2 = pq.$$

Setzt man dies wiederum in (3.2) bzw. (3.3) ein, dann berechnet man

$$a^2 = p^2 + pq = p(p + q) = pc \quad \text{bzw.} \quad b^2 = pq + q^2 = q(p + q) = qc$$

und zeigt somit die Behauptung. □

Damit erkennen wir, dass die folgende Zeichnung die Inversion am Einheitskreis darstellt.



iv) Polarkoordinaten. Es sei

$$f: \underbrace{\mathbb{R}_{>0} \times [0, 2\pi)}_{=:D} \longrightarrow \mathbb{R}^2$$

$$(r, \varphi) \longmapsto (r \cdot \cos(\varphi), r \cdot \sin(\varphi)).$$

Die Funktion f ist nach Lemma 3.1.3 und Folgerung 3.2.3 stetig.

Behauptung. Die Funktion f bildet D bijektiv auf $\mathbb{R}^2 \setminus \{0\}$ ab.

Beweis. Wir zeigen zunächst, dass f injektiv ist. Für $x = (r, \varphi)$ berechnen wir

$$\|f(x)\| = \sqrt{r^2 \cdot (\cos^2(\varphi) + \sin^2(\varphi))} \stackrel{\cos^2(\varphi) + \sin^2(\varphi) = 1}{=} r.$$

Damit können wir r aus $f(x)$ zurückgewinnen.

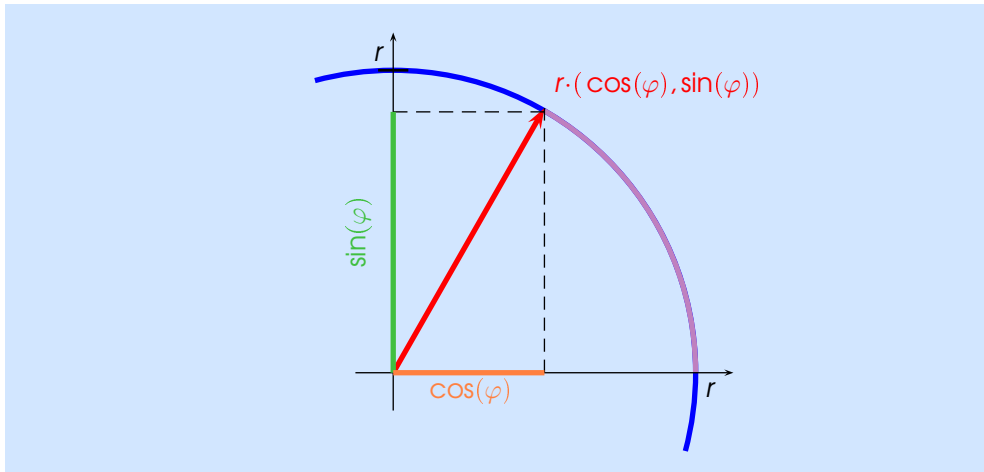
Eigenschaften des Sinus und des Kosinus.

- Der Kosinus ist auf den Intervallen $[0, \pi]$ bzw. $(\pi, 2\pi)$ umkehrbar.
- Für $\varphi \in [0, 2\pi)$ gilt

$$\sin(\varphi) \geq 0 \iff \varphi \in [0, \pi]$$

$$\sin(\varphi) < 0 \iff \varphi \in (\pi, 2\pi).$$

Folglich lässt sich φ aus $\cos(\varphi)$ und dem Vorzeichen von $\sin(\varphi)$ rekonstruieren.



Sei nun $(x, y) \in \mathbb{R}^2 \setminus \{0\}$. Dann gilt $(x, y) = r \cdot (u, v)$ mit $r := \|(x, y)\|$, $u := x/r$ und $v := y/r$. Wir haben weiter $u^2 + v^2 = 1$. Falls $v \geq 0$, dann definieren wir $\varphi \in [0, \pi]$ als die eindeutig bestimmte Zahl mit $\cos(\varphi) = u$. Da $v^2 = 1 - u^2 = 1 - \cos^2(\varphi) = \sin^2(\varphi)$, sehen wir $v = |v| = |\sin(\varphi)|$. Wegen $\varphi \in [0, \pi]$ folgt $\sin(\varphi) \geq 0$, so dass $v = \sin(\varphi)$. Analog legen wir φ im Fall $v < 0$ als die eindeutig bestimmte Zahl im Intervall $(\pi, 2\pi)$ mit $\cos(\varphi) = u$ fest. Wieder überprüfen wir $v = \sin(\varphi)$.

In jedem der betrachteten Fälle haben wir $f(r, \varphi) = (x, y)$. Damit ist die Surjektivität von f nachgewiesen. \square

3.2.5 Definition. Es seien $a \in \mathbb{R}^n$, $U \subset \mathbb{R}^n$ eine offene Umgebung von a und $f: U \setminus \{a\} \rightarrow \mathbb{R}$ eine Funktion. Wir sagen, dass der *Grenzwert*

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

existiert, wenn es eine Zahl $b \in \mathbb{R}$ gibt, so dass für jede Folge $(x_k)_{k \in \mathbb{N}}$ mit $x_k \in U \setminus \{a\}$, $k \in \mathbb{N}$, und $\lim_{k \rightarrow \infty} x_k = a$ die Gleichung

$$\lim_{k \rightarrow \infty} f(x_k) = b$$

erfüllt ist.

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

3.2.6 Bemerkung (Stetige Fortsetzung). Falls f stetig ist und $\lim_{x \rightarrow a} f(x) = b$ gilt, dann ist auch die Funktion

$$\bar{f}: U \rightarrow \mathbb{R} \\ x \mapsto \begin{cases} f(x), & \text{falls } x \neq a \\ b, & \text{falls } x = a \end{cases}$$

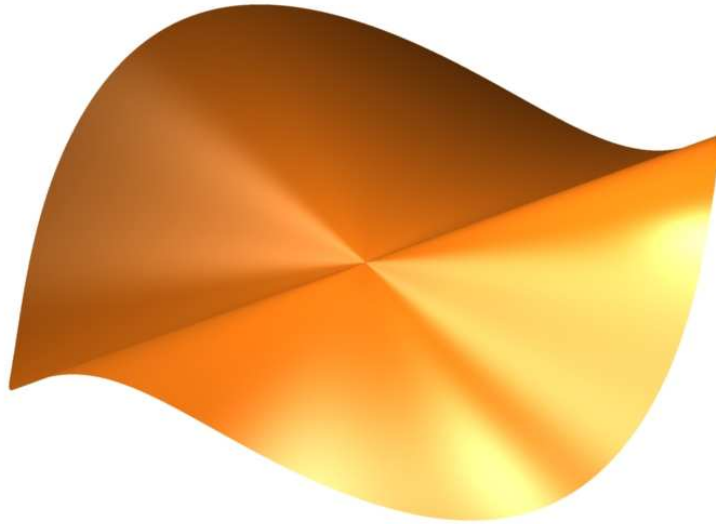
stetig.

3.2.7 Beispiele. i) Es sei

$$f: \mathbb{R}^2 \setminus \{0\} \rightarrow \mathbb{R} \\ (x, y) \mapsto \frac{x^2 \cdot y}{x^2 + y^2}.$$

(Das folgende und andere Bilder wurden erstellt mit „surfer“.³⁾)

³<http://www.imaginary2008.de/surfer.php>



Wir behaupten

$$\lim_{(x,y) \rightarrow 0} f(x,y) = 0.$$

Wir benutzen Polarkoordinaten: Für $(x,y) = (r \cdot \cos(\varphi), r \cdot \sin(\varphi))$ berechnen wir

$$f(x,y) = \frac{r^3 \cdot \cos^2(\varphi) \cdot \sin(\varphi)}{r^2} = r \cdot \cos^2(\varphi) \cdot \sin(\varphi).$$

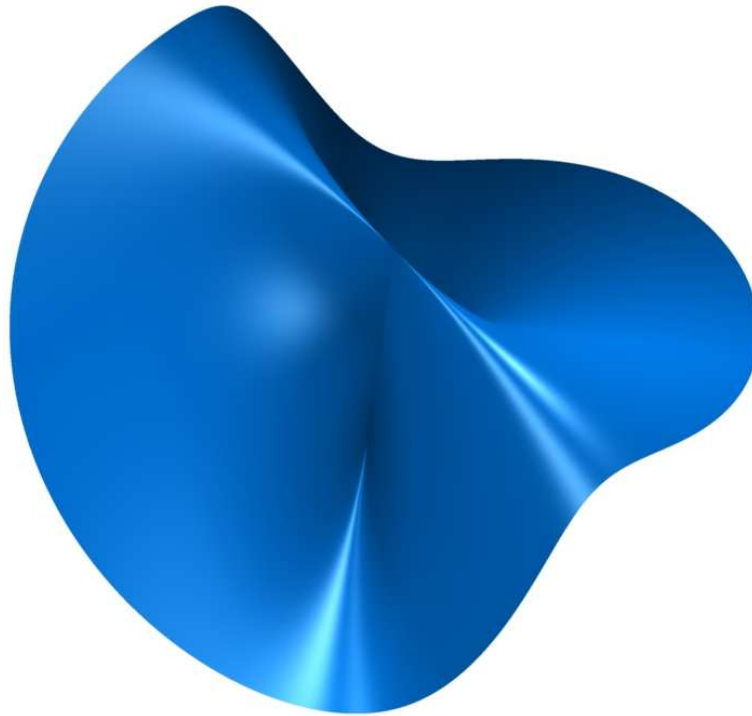
Damit gilt

$$|f(x,y)| = r \cdot |\cos^2(\varphi)| \cdot |\sin(\varphi)| \leq r.$$

Wenn (x,y) gegen 0 strebt, dann strebt auch r gegen 0, und unsere Behauptung folgt.

ii) Diesmal untersuchen wir die Funktion

$$\begin{aligned} f: \mathbb{R}^2 \setminus \{0\} &\longrightarrow \mathbb{R} \\ (x,y) &\longmapsto \frac{x \cdot y}{x^2 + 5 \cdot y^2}. \end{aligned}$$



Für $m \in \mathbb{R}$ haben wir

$$\lim_{x \rightarrow 0} f(x, m \cdot x) = \lim_{x \rightarrow 0} \frac{m \cdot x^2}{x^2 + 5 \cdot m^2 \cdot x^2} = \frac{m}{1 + 5 \cdot m^2}.$$

Für $m = 0$ finden wir 0 als Grenzwert und für $m = 1$ den Wert $1/6$. Damit existiert $\lim_{(x,y) \rightarrow 0} f(x, y)$ **nicht**.⁴

3.3 Die ε - δ -Formulierung der Stetigkeit

Analog zu Ergebnissen in der Analysis I beweisen wir nun die Charakterisierung der Stetigkeit einer Abbildung zwischen metrischen Räumen durch das ε - δ -Kriterium. Es ermöglicht, Stetigkeit ganz in Termen der zugehörigen Topologien zu formulieren. Darauf wird in den Übungsaufgaben eingegangen.

3.3.1 Satz. *Es seien (X, d_X) , (Y, d_Y) metrische Räume, $f: X \rightarrow Y$ eine Abbildung und $a \in X$ ein Punkt. Die Abbildung ist genau dann in a stetig, wenn das folgende Kriterium erfüllt ist:*

$$\forall \varepsilon > 0 \exists \delta > 0 \forall x \in X : d_X(a, x) < \delta \Rightarrow d_Y(f(a), f(x)) < \varepsilon.$$

Beweis. Der Beweis ist eine Übersetzung des Beweises von Satz 3.4.8 in (12) in die Sprache der metrischen Räume.

⁴Im Bild des Funktionsgraphen kann man gut die Gerade erkennen, auf der die Grenzwerte $\lim_{x \rightarrow 0} f(x, m \cdot x)$ für verschiedene Wahlen von $m \in \mathbb{R}$ liegen.

3.3 Die ε - δ -Formulierung der Stetigkeit

Wir nehmen an, dass f in a stetig ist und das ε - δ -Kriterium versagt. Dann gibt es ein $\varepsilon > 0$, so dass für alle $\delta > 0$ ein $x \in X$ mit

$$d_X(a, x) < \delta \wedge d_Y(f(a), f(x)) \geq \varepsilon$$

existiert. Insbesondere gibt es für jedes $k \geq 1$ ein $x_k \in X$ mit

$$d_X(a, x_k) < \frac{1}{k} \wedge d_Y(f(a), f(x_k)) \geq \varepsilon.$$

Die Folge $(x_k)_{k \geq 1}$ konvergiert dann gegen a . Nach Beispiel 3.1.2, i), ist $y \mapsto d_Y(f(a), y)$ eine stetige Funktion, und nach Satz 2.2.18 aus (12) gilt

$$0 = d_Y(f(a), f(a)) = d_Y(f(a), \lim_{k \rightarrow \infty} f(x_k)) = \lim_{k \rightarrow \infty} d_Y(f(a), f(x_k)) \geq \varepsilon.$$

Das ist offenbar unmöglich, so dass die Annahme falsch war und das ε - δ -Kriterium gilt.

Jetzt wird vorausgesetzt, dass das ε - δ -Kriterium erfüllt ist. Es sei $(x_k)_{k \in \mathbb{N}}$ eine Folge in X , die gegen a konvergiert. Um zu zeigen, dass die Folge $(f(x_k))_{k \in \mathbb{N}}$ gegen $f(a)$ konvergiert, müssen wir überprüfen, dass es zu jedem $\varepsilon > 0$ ein $K \in \mathbb{N}$ gibt, so dass

$$\forall k \geq K : d_Y(f(a), f(x_k)) < \varepsilon. \quad (3.4)$$

Seien $\varepsilon > 0$ und $\delta > 0$, so dass

$$\forall x \in X : d_X(a, x) < \delta \Rightarrow d_Y(f(a), f(y)) < \varepsilon. \quad (3.5)$$

Wegen der Konvergenz der Folge $(x_k)_{k \geq 1}$ gegen a existiert ein $K \in \mathbb{N}$ mit der Eigenschaft:

$$\forall k \geq K : d_X(a, x_k) < \delta. \quad (3.6)$$

Die Aussagen (3.5) und (3.6) implizieren zusammen Aussage (3.4). \square

3.3.2 Definition. Es seien (X, d) ein metrischer Raum, $A \subset X$ eine nichtleere Teilmenge und $x \in X$. Die Zahl

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

heißt der *Abstand von x zu A* .

3.3.3 Lemma. Die Funktion

$$\begin{aligned} \text{dist}(\cdot, A) : X &\longrightarrow \mathbb{R} \\ x &\longmapsto \text{dist}(x, A) \end{aligned}$$

ist stetig.

Beweis. Die Behauptung ergibt sich sofort aus:

$$\forall x, x' \in X \forall \varepsilon > 0 : d(x, x') < \varepsilon \Rightarrow |\text{dist}(x, A) - \text{dist}(x', A)| < \varepsilon. \quad (3.7)$$

Für $x, x' \in X$ mit $d(x, x') < \varepsilon$ wählen wir eine Folge $(y_k)_{k \in \mathbb{N}}$ in A mit

$$\lim_{k \rightarrow \infty} \text{dist}(x', y_k) = \text{dist}(x', A).$$

Für jedes $k \in \mathbb{N}$ gilt dann

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

Durch den Grenzübergang $k \rightarrow \infty$ wird daraus

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

Ebenso erhält man

$$\text{dist}(x', A) < \text{dist}(x, A) + \varepsilon.$$

Daraus folgt die behauptete Abschätzung $|\text{dist}(x, A) - \text{dist}(x', A)| < \varepsilon$. \square

3.4 Stetige Funktionen auf kompakten Mengen

In diesem Abschnitt wird gezeigt, dass stetige Funktionen auf kompakten Mengen ein Minimum und ein Maximum annehmen.

3.4.1 Satz. *Es seien (X, d_X) und (Y, d_Y) zwei metrische Räume und $f: X \rightarrow Y$ eine stetige Abbildung. Wenn $K \subset X$ kompakt ist, dann ist auch das Bild $f(K) \subset Y$ kompakt.*

Beweis. Wir verwenden die Definition der Kompaktheit durch die Überdeckungseigenschaft (2.2.1). Es sei $(V_i)_{i \in I}$ eine offene Überdeckung von $f(K)$. Da f stetig ist, ist $U_i := f^{-1}(V_i)$, $i \in I$, eine offene Menge (Aufgabe A.4.1, a). Wir haben

$$K \subset f^{-1}(f(K)) \subset \bigcup_{i \in I} f^{-1}(V_i) = \bigcup_{i \in I} U_i.$$

Da K kompakt ist, gibt es Indizes $i_1, \dots, i_m \in I$, so dass

$$K \subset U_{i_1} \cup \dots \cup U_{i_m}.$$

Dann gilt auch

$$f(K) \subset f(U_{i_1}) \cup \dots \cup f(U_{i_m}) \subset V_{i_1} \cup \dots \cup V_{i_m},$$

und die Behauptung ist gezeigt. □

3.4.2 Folgerung. *Es seien (X, d) ein **kompakter** metrischer Raum und $f: X \rightarrow \mathbb{R}$ eine stetige Abbildung. Dann nimmt f auf X ein Minimum und ein Maximum an, d.h., es gibt Punkte a und b in X , so dass*

$$\forall x \in X: f(a) \leq f(x) \leq f(b).$$

Beweis. Auf Grund des vorherigen Satzes ist die Bildmenge $A := f(X)$ kompakt. Nach Satz 2.2.10 ist A abgeschlossen und beschränkt. Es seien $\inf(A) \in \mathbb{R}$ bzw. $\sup(A) \in \mathbb{R}$ das Infimum bzw. Supremum von A . Wir finden Folgen $(x_k)_{k \in \mathbb{N}}$ und $(y_k)_{k \in \mathbb{N}}$ mit

- $x_k \in A$ und $y_k \in A$, $k \in \mathbb{N}$,
- $\lim_{k \rightarrow \infty} x_k = \inf(A)$ und $\lim_{k \rightarrow \infty} y_k = \sup(A)$,

Da A abgeschlossen ist, folgt aus Satz 2.1.4, dass $\inf(A) \in A$ und $\sup(A) \in A$. Das bedeutet, dass es Punkte $a, b \in X$ mit $f(a) = \inf(A)$ und $f(b) = \sup(A) \in A$ gibt. □

3.4.3 Definition. Es seien (X, d) ein metrischer Raum und $A, B \subset X$ nichtleere Teilmengen. Der *Abstand* von A und B ist

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

3.4.4 Lemma. *Wenn A kompakt und B abgeschlossen ist, dann folgt*

$$\text{dist}(A, B) = 0 \iff A \cap B \neq \emptyset.$$

Beweis. Wenn $A \cap B \neq \emptyset$, dann gilt sicherlich $\text{dist}(A, B) = 0$.

Nun nehmen wir $A \cap B = \emptyset$ an. Die Funktion $x \mapsto \text{dist}(x, B)$ ist nach Lemma 3.3.3 stetig und nimmt wegen Aufgabe A.3.1 nur positive Werte auf A an. Wegen Folgerung 3.4.2 gibt es ein $a_0 \in A$ mit $\text{dist}(A, B) = \text{dist}(a_0, B) > 0$. □

3.4.5 *Beispiel.* Falls A und B nicht kompakt sind, stimmt die Aussage des Lemmas nicht mehr. Wir betrachten z.B. die Funktion

$$\begin{aligned} f: \mathbb{R}^2 &\longrightarrow \mathbb{R} \\ (x, y) &\longmapsto x \cdot y. \end{aligned}$$

Da f stetig ist (Beispiel 3.2.4, i), sind die Teilmengen

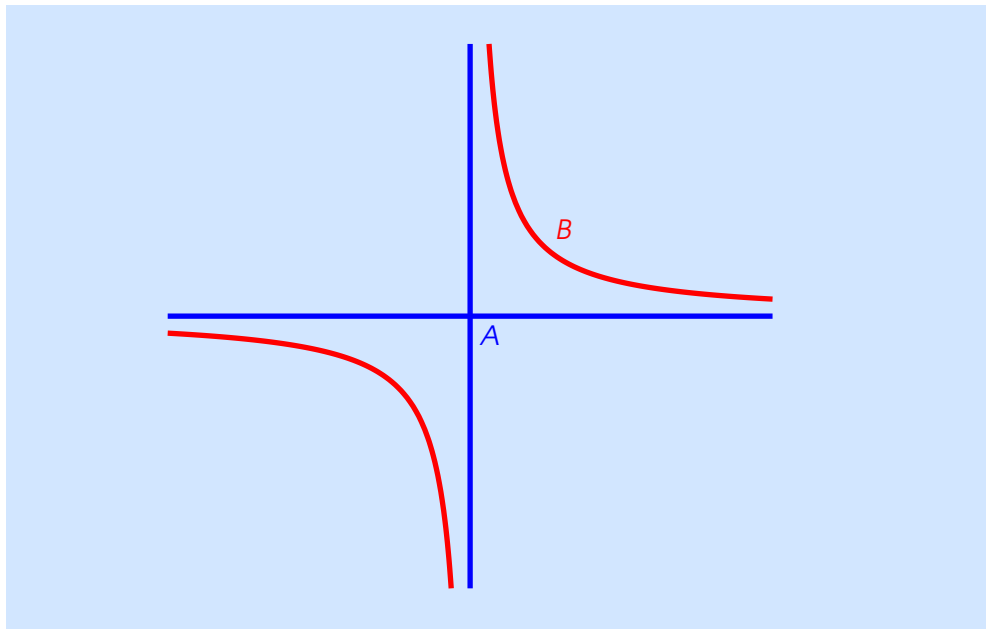
$$A := f^{-1}(0) = \{ (x, y) \in \mathbb{R}^2 \mid x \cdot y = 0 \}$$

und

$$B := f^{-1}(1) = \{ (x, y) \in \mathbb{R}^2 \mid x \cdot y = 1 \}$$

abgeschlossen (Aufgabe A.4.1, a). Man sieht sofort

$$A \cap B = \emptyset \quad \text{und} \quad \text{dist}(A, B) = 0.$$



3.5 Gleichmäßige Stetigkeit

In diesem Abschnitt stellen wir eine Verschärfung des Begriffs der Stetigkeit vor. Stetige Funktionen auf kompakten Mengen erfüllen diese Verschärfung jedoch automatisch.

3.5.1 Definition. Es seien (X, d_X) und (Y, d_Y) metrische Räume. Eine Abbildung $f: X \rightarrow Y$ ist *gleichmäßig stetig*, wenn zu jedem $\varepsilon > 0$ ein $\delta > 0$ existiert, so dass

$$\forall x, x' \in X: d_X(x, x') < \delta \Rightarrow d_Y(f(x), f(x')) < \varepsilon.$$

3.5.2 *Bemerkung.* i) Man vergleiche Behauptung 1 im Beweis von Satz 5.1.9 in (12).

ii) Das ε - δ -Kriterium für Stetigkeit 3.3.1 fordert, dass wir für $\varepsilon > 0$ zu jedem $x \in X$ ein $\delta(x) > 0$ finden können, für das

$$\forall x' \in X: d_X(x, x') < \delta(x) \Rightarrow d_Y(f(x), f(x')) < \varepsilon$$

gilt.

Die gleichmäßige Stetigkeit fordert, dass ein $\delta > 0$ existiert, das für alle $x \in X$ funktioniert.

3.5.3 Beispiel. Es seien (X, d) ein metrischer Raum und $A \subset X$ eine nichtleere Teilmenge. Wie (3.7) zeigt, ist $\text{dist}(\cdot, A): X \rightarrow \mathbb{R}, x \mapsto \text{dist}(x, A)$, gleichmäßig stetig.

3.5.4 Satz. Es seien (X, d_X) und (Y, d_Y) metrische Räume. Wenn X kompakt ist, dann ist jede stetige Funktion $f: X \rightarrow Y$ auch gleichmäßig stetig.

Dieser Satz verallgemeinert Behauptung 1 im Beweis von Satz 5.1.9 in (12).

Beweis. Sei $\varepsilon > 0$. Auf Grund der Stetigkeit der Funktion f existiert für jeden Punkt $x \in X$ ein $\delta(x) > 0$, so dass

$$\forall x' \in X: \quad x' \in B(x, \delta(x)) \quad \Rightarrow \quad d_Y(f(x), f(x')) < \frac{\varepsilon}{2}. \quad (3.8)$$

Wir haben

$$X = \bigcup_{x \in X} B\left(x, \frac{1}{2} \cdot \delta(x)\right).$$

Da X kompakt ist, gibt es $x_1, \dots, x_m \in X$, so dass

$$X = \bigcup_{i=1}^m B\left(x_i, \frac{1}{2} \cdot \delta(x_i)\right).$$

Wir definieren

$$\delta := \frac{1}{2} \cdot \min\{\delta(x_i) \mid i = 1, \dots, m\}.$$

Für Punkte $x, x' \in X$ mit $d_X(x, x') < \delta$, sei $i_0 \in \{1, \dots, m\}$ mit $x \in B(x_{i_0}, (1/2) \cdot \delta(x_{i_0}))$. Damit schätzen wir folgendermaßen ab:

$$d_X(x', x_{i_0}) \leq d_X(x', x) + d_X(x, x_{i_0}) < \delta + \frac{1}{2} \cdot \delta(x_{i_0}) \leq \delta(x_{i_0}),$$

so dass

$$x' \in B(x_{i_0}, \delta(x_{i_0})).$$

Weiter sehen wir mit (3.8)

$$d_Y(f(x), f(x')) \leq d_Y(f(x), f(x_{i_0})) + d_Y(f(x_{i_0}), f(x')) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Damit ist gezeigt, dass f gleichmäßig stetig ist. □

3.6 Folgen stetiger Funktionen

Wir erklären hier, unter welchen Voraussetzungen Folgen stetiger Funktionen gegen eine stetige Funktion konvergieren. Dieses Resultat können wir als Vollständigkeit eines gewissen normierten Raums interpretieren. Damit erhalten wir einen unendlichdimensionalen Banach-Raum.

Es sei (X, d) ein metrischer Raum. Wie in Beispiel 1.2.6 führen wir auf

$$\text{Abb}(X, \mathbb{R}) := \{ f: X \rightarrow \mathbb{R} \mid f \text{ Abbildung} \}$$

die Struktur eines reellen Vektorraums ein. Auf Grund von Folgerung 3.2.3 ist

$$\mathcal{C}^0(X) := \{ f \in \text{Abb}(X, \mathbb{R}) \mid f \text{ ist stetig} \}$$

ein Untervektorraum.

3.6.1 Definition. Die Folge $(f_k)_{k \in \mathbb{N}}$ in $\text{Abb}(X, \mathbb{R})$ konvergiert *gleichmäßig* gegen die Funktion $f \in \text{Abb}(X, \mathbb{R})$, wenn es zu jedem $\varepsilon > 0$ ein $K \in \mathbb{N}$ gibt, für das

$$\sup \{ |f_k(x) - f(x)| \mid x \in X \} < \varepsilon, \quad k \geq K,$$

gilt.

3.6.2 Satz. Es seien (X, d) ein metrischer Raum, $(f_k)_{k \in \mathbb{N}}$ eine Folge von Elementen in $\mathcal{C}^0(X)$ und $f \in \text{Abb}(X, \mathbb{R})$. Wenn $(f_k)_{k \in \mathbb{N}}$ **gleichmäßig** gegen f konvergiert, dann ist $f \in \mathcal{C}^0(X)$, d.h. f ist stetig.

Beweis. Der Beweis von Satz 3.7.9 (12) lässt sich wortwörtlich übertragen. \square

3.6.3 Bemerkung. Wir sagen, dass die Folge $(f_k)_{k \in \mathbb{N}}$ *punktweise* gegen f konvergiert, wenn

$$\forall x \in X: \lim_{k \rightarrow \infty} f_k(x) = f(x).$$

Gleichmäßige Konvergenz impliziert offenbar punktweise Konvergenz.

Wenn $(f_k)_{k \in \mathbb{N}}$ eine Folge stetiger Funktionen auf X ist, die punktweise gegen die Funktion $f: X \rightarrow \mathbb{R}$ konvergiert, dann muss die Grenzfunktion f **nicht** stetig sein. Ein Gegenbeispiel wird in (12), Beispiel 3.7.6, gegeben.

In Beispiel 1.2.6 haben wir den Vektorraum

$$\mathcal{B}(X) := \left\{ f: X \rightarrow \mathbb{R} \mid \sup \{ |f(x)| \mid x \in X \} < \infty \right\}$$

der beschränkten Funktionen auf X eingeführt. Durch

$$\begin{aligned} \|\cdot\|: \mathcal{B}(X) &\rightarrow \mathbb{R} \\ f &\mapsto \sup \{ |f(x)| \mid x \in X \} \end{aligned}$$

ist eine Norm auf $\mathcal{B}(X)$ gegeben.

Sei

$$\mathcal{C}_b^0(X) := \mathcal{C}^0(X) \cap \mathcal{B}(X)$$

der Vektorraum der beschränkten stetigen Funktionen. Mit

$$\|f\|_{\mathcal{C}_b^0(X)} := \|f\|, \quad f \in \mathcal{C}_b^0(X),$$

definieren wir den normierten Vektorraum

$$(\mathcal{C}_b^0(X), \|\cdot\|_{\mathcal{C}_b^0(X)}).$$

3.6.4 Bemerkung. i) Wenn der Raum X kompakt ist, dann impliziert Folgerung 3.4.2, dass

$$\mathcal{C}_b^0(X) = \mathcal{C}^0(X).$$

ii) Nach Satz 2.1.4 ist $\mathcal{C}_b^0(X)$ ein **abgeschlossener** Unterraum von $\mathcal{B}(X)$.⁵

iii) In $\mathcal{B}(X)$ ist die gleichmäßige Konvergenz einer Folge $(f_k)_{k \in \mathbb{N}}$ gegen eine Funktion f gleichbedeutend mit der Konvergenz dieser Folge gegen f in der Supremumsnorm $\|\cdot\|$.

3.6.5 Satz. Der Raum $(\mathcal{C}_b^0(X), \|\cdot\|_{\mathcal{C}_b^0(X)})$ ist ein Banach-Raum.

⁵D.h. $\mathcal{C}_b^0(X)$ ist ein Unterraum und als Teilmenge von $\mathcal{B}(X)$ abgeschlossen.

Beweis. Wir müssen beweisen, dass jede Cauchy-Folge in $\mathcal{C}_b^0(X)$ konvergiert.

Sei also $(f_k)_{k \in \mathbb{N}}$ eine Cauchy-Folge in $\mathcal{C}_b^0(X)$. Zu $\varepsilon > 0$ gibt es daher ein $K \in \mathbb{N}$, so dass

$$\forall k, l \geq K : \sup \{ |f_k(x) - f_l(x)| \mid x \in X \} = \|f_k - f_l\|_{\mathcal{C}_b^0(X)} < \frac{\varepsilon}{2}.$$

Insbesondere hat man

$$\forall x \in X \forall k, l \geq K : |f_k(x) - f_l(x)| < \frac{\varepsilon}{2}. \quad (3.9)$$

Man erkennt, dass die Folge $(f_k(x))_{k \in \mathbb{N}}$ für jeden Punkt $x \in X$ eine Cauchy-Folge von reellen Zahlen ist und damit konvergiert ((12), Satz 2.3.12). Wir definieren daher

$$\begin{aligned} f: X &\longrightarrow \mathbb{R} \\ x &\longmapsto \lim_{k \rightarrow \infty} f_k(x). \end{aligned}$$

Es ist nachzuweisen, dass die Funktion f stetig ist. Wegen Satz 3.6.2 reicht es aus zu zeigen, dass die Folge $(f_k)_{k \in \mathbb{N}}$ gleichmäßig gegen f konvergiert. Mit der Stetigkeit von $|\cdot|$ und (3.9) sehen wir

$$\forall x \in X \forall k \geq K : |f_k(x) - f(x)| = \lim_{l \rightarrow \infty} |f_k(x) - f_l(x)| \leq \frac{\varepsilon}{2} < \varepsilon$$

und damit

$$\forall k \geq K : \sup \{ |f_k(x) - f(x)| \mid x \in X \} < \varepsilon.$$

Dies ist die gleichmäßige Konvergenz. □

3.7 Stetigkeit linearer Abbildungen

In diesem Abschnitt wollen wir die Stetigkeit linearer Abbildungen untersuchen. Es zeigt sich, dass im Fall unendlichdimensionaler Vektorräume lineare Abbildungen nicht automatisch stetig sind. Die in der Analysis I eingeführten Operationen der Ableitung und des Integrals liefern interessante Beispiele für lineare Abbildungen.

3.7.1 Satz. *Es seien $(V, \|\cdot\|_V)$ und $(W, \|\cdot\|_W)$ normierte Vektorräume. Eine lineare Abbildung $A: V \rightarrow W$ ist genau dann stetig, wenn folgende Bedingung erfüllt ist:*

$$\exists C \in \mathbb{R}_{>0} \forall v \in V : \|A(v)\|_W \leq C \cdot \|v\|_V. \quad (3.10)$$

Beweis. Wir setzen zuerst voraus, dass A stetig ist. Weil A dann insbesondere in 0 stetig ist, existiert ein $\delta > 0$ mit der Eigenschaft

$$\forall v \in V : \|v\|_V < \delta \Rightarrow \|A(v)\|_W < 1. \quad (3.11)$$

Sei

$$C := \frac{2}{\delta}.$$

Für $v \in V \setminus \{0\}$ setzen wir weiter

$$\lambda_v := \frac{1}{C \cdot \|v\|_V} \in \mathbb{R}_{>0}.$$

Damit berechnet man

$$\|\lambda_v \cdot v\|_V = \lambda_v \cdot \|v\|_V = \frac{1}{C} = \frac{\delta}{2} < \delta.$$

Mit (3.11) finden wir schließlich

$$\lambda_v \cdot \|A(v)\|_W = \|\lambda_v \cdot A(v)\|_W = \|A(\lambda_v \cdot v)\|_W \leq 1 = C \cdot \|\lambda_v \cdot v\|_V = \lambda_v \cdot C \cdot \|v\|_V.$$

Teilen durch λ_v ergibt die gewünschte Abschätzung.

Jetzt nehmen wir an, dass Bedingung (3.10) gilt, und überprüfen die Stetigkeit von A mit dem ε - δ -Kriterium 3.3.1. Dazu seien $\varepsilon > 0$, $\delta := \varepsilon/C$ und $v_0 \in V$. Wir beobachten:

$$\forall v \in V : \|v - v_0\|_V < \delta \Rightarrow \|A(v) - A(v_0)\|_W = \|A(v - v_0)\|_W \leq C \cdot \|v - v_0\|_V < \varepsilon.$$

Folglich ist A in v_0 stetig. □

3.7.2 Bemerkung. i) Wir können den Satz folgendermaßen umformulieren: Die lineare Abbildung $A: V \rightarrow W$ ist genau dann stetig, wenn

$$\|A\|_{\text{op}} := \sup \left\{ \|A(v)\|_W \mid \|v\|_V = 1 \right\} < \infty.$$

Für eine stetige lineare Abbildung A nennt man $\|A\|_{\text{op}}$ die *Operatornorm* von A .

ii) Aus dem obigen Beweis folgt, dass eine stetige lineare Abbildung gleichmäßig stetig ist.

3.7.3 Beispiel. Jede lineare Abbildung $A: \mathbb{R}^m \rightarrow \mathbb{R}^n$ ist stetig. (Nach Satz 2.3.3 ist die Aussage unabhängig davon, welche Normen wir auf \mathbb{R}^m bzw. \mathbb{R}^n wählen.) Um das mit Satz 3.7.1 einzusehen, fixieren wir die Standardbasen von \mathbb{R}^m und \mathbb{R}^n und definieren

$$M := (m_{ij})_{\substack{i=1,\dots,n \\ j=1,\dots,m}}$$

als die Abbildungsmatrix von A . Für $v = (v_1, \dots, v_m) \in \mathbb{R}^m$ ist $w = (w_1, \dots, w_n) := A(v)$ somit durch

$$w = v \cdot M, \quad \text{d.h.} \quad w_i = \sum_{j=1}^m m_{ji} \cdot v_j, \quad i = 1, \dots, n, \quad (3.12)$$

gegeben. Nun definieren wir

$$C := \sum_{\substack{i=1,\dots,n \\ j=1,\dots,m}} |m_{ij}|.$$

Wie üblich bezeichnen wir mit $\|\cdot\|_M$ die Maximum-Norm auf \mathbb{R}^m und \mathbb{R}^n . Mit (3.12) erkennen wir

$$|w_i| \leq \sum_{j=1}^m |m_{ji}| \cdot |v_j| \leq \left(\sum_{j=1}^m |m_{ji}| \right) \cdot \|v\|_M \leq C \cdot \|v\|_M, \quad i = 1, \dots, n,$$

so dass

$$\|w\|_M \leq C \cdot \|v\|_M$$

für alle $v = (v_1, \dots, v_m) \in \mathbb{R}^m$ und $w = (w_1, \dots, w_n) = A(v) \in \mathbb{R}^n$ folgt.

Für die folgenden Betrachtungen müssen wir unsere Konventionen für differenzierbare Funktionen etwas modifizieren. Für ein Intervall I der Form $[a, b]$ oder $[a, \infty)$ nennen wir a die *linke Intervallgrenze*, und für ein Intervall I der Form $[a, b]$ oder $(-\infty, b]$ ist b die *rechte Intervallgrenze*. Es seien I ein Intervall und - sofern existent - a die linke bzw. b die rechte Intervallgrenze. Wir sagen, dass die Funktion $f: I \rightarrow \mathbb{R}$ in a bzw. b *differenzierbar* ist, wenn sie dort rechtsseitig bzw. linksseitig differenzierbar ist (s. (12), Definition 4.1.1), und setzen kurz

$$f'(a) := f'_+(a) \quad \text{bzw.} \quad f'(b) := f'_-(b).$$

Wir nennen f *differenzierbar*, wenn f in jedem Punkt $x \in I$ differenzierbar ist. Ist zusätzlich die Funktion

$$\begin{aligned} f' : I &\longrightarrow \mathbb{R} \\ t &\longmapsto f'(t) \end{aligned}$$

stetig, so ist f *stetig differenzierbar*.

3.7.4 Bemerkung. Es seien $f: I \rightarrow \mathbb{R}$ eine Funktion und $a \in I$ der linke Randpunkt. Die Funktion f sei im obigen Sinne differenzierbar mit $f'(a) = c$. Wir definieren

$$\begin{aligned} \tilde{f} : (-\infty, a) \cup I &\longrightarrow \mathbb{R} \\ x &\longmapsto \begin{cases} f(x), & \text{falls } x \in I \\ c \cdot (x - a) + f(a), & \text{falls } x \in (-\infty, a) \end{cases} \end{aligned}$$

Die Funktion \tilde{f} ist stetig differenzierbar mit $\tilde{f}|_I = f$. Analog kann man für die rechte Intervallgrenze verfahren. Eine Funktion $f: I \rightarrow \mathbb{R}$ ist mithin genau dann (stetig) differenzierbar, wenn es ein offenes Intervall $J \subset \mathbb{R}$ und eine (stetig) differenzierbare Funktion $\tilde{f}: J \rightarrow \mathbb{R}$ mit $I \subset J$ und $\tilde{f}|_I = f$ gibt.

3.7.5 Beispiele. i) Es seien $I := [0, 1]$ und $(\mathcal{C}^0(I), \|\cdot\|_{\mathcal{C}^0(I)})$ der Vektorraum der stetigen Funktionen auf I mit der Supremumsnorm. Weiter sei

$$\mathcal{C}^1(I) \subset \mathcal{C}^0(I)$$

der Unterraum der stetig differenzierbaren Funktionen. Er erhält durch Einschränkung der Supremumsnorm eine Norm.

Wir behaupten, dass der Differentialoperator

$$\begin{aligned} D : \mathcal{C}^1(I) &\longrightarrow \mathcal{C}^0(I) \\ f &\longmapsto f' \end{aligned}$$

linear aber nicht stetig ist. Die Linearität ist eine Folge der Ableitungsregeln (Satz 4.2.1 in (12)). Wir betrachten die Funktionen $f_k: I \rightarrow \mathbb{R}, x \mapsto x^k, k \in \mathbb{N}$. Für sie gilt:

- $\|f_k\|_{\mathcal{C}^0(I)} = 1, k \in \mathbb{N},$
- $\|f'_k\|_{\mathcal{C}^0(I)} = \|k \cdot f_{k-1}\|_{\mathcal{C}^0(I)} = k, k \geq 1.$

In Aufgabe A.5.1 werden Sie sehen, wie man dieses Problem behebt.

ii) Es seien $I := [0, 1]$ und

$$\begin{aligned} \text{Int} : \mathcal{C}^0(I) &\longrightarrow \mathbb{R} \\ f &\longmapsto \int_0^1 f(x) dx \end{aligned}$$

der Integraloperator. Er ist nach Satz 5.1.13 in (12) linear. Der Leser möge sich von der Abschätzung

$$\left| \int_0^1 f(x) dx \right| \leq \int_0^1 |f(x)| dx \leq \|f\|_{\mathcal{C}^0(I)}$$

überzeugen. Der Integraloperator ist also stetig.

Kapitel 4

Wege und Kurven in \mathbb{R}^n

4.1 Wege

Wir führen den Begriff des Weges ein und erläutern ihn an einigen Beispielen.

In diesem Kapitel steht $\|\cdot\|$ immer für die euklidische Norm.

4.1.1 Definition. a) Ein Weg (in \mathbb{R}^n) ist eine stetige Abbildung $\gamma: I \rightarrow \mathbb{R}^n$, $I \subset \mathbb{R}$ ein Intervall.

b) Für reelle Zahlen $a < b$ und einen Weg $\gamma: [a, b] \rightarrow \mathbb{R}^n$ nennt man $\gamma(a)$ den Anfangspunkt des Weges und $\gamma(b)$ den Endpunkt.

c) Ist $\gamma: I \rightarrow \mathbb{R}^n$ ein Weg, dann bezeichnet man die Menge

$$\text{Spur}(\gamma) := \gamma(I) = \{\gamma(t) \mid t \in I\}$$

als Spur des Weges γ .

4.1.2 Beispiele. i) Die Spur des Weges

$$\begin{aligned} \gamma_1: [0, 2\pi] &\rightarrow \mathbb{R}^2 \\ t &\mapsto (r \cdot \cos(t), r \cdot \sin(t)) \end{aligned}$$

ist der Kreis vom Radius r um den Nullpunkt.

ii) Der Weg

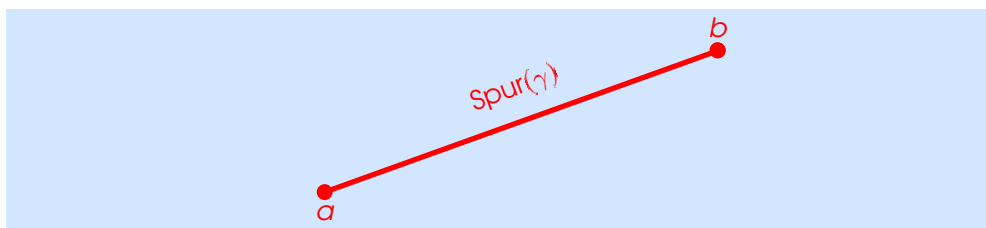
$$\begin{aligned} \gamma_2: [0, 2\pi] &\rightarrow \mathbb{R}^2 \\ t &\mapsto (r \cdot \cos(2t), r \cdot \sin(2t)) \end{aligned}$$

hat ebenfalls den Kreis vom Radius r um den Nullpunkt als Spur. Jetzt wird der Kreis allerdings zweimal durchlaufen.

iii) Für zwei Punkte a, b in \mathbb{R}^n definieren wir den Weg

$$\begin{aligned} \gamma: [0, 1] &\rightarrow \mathbb{R}^n \\ t &\mapsto a + t \cdot (b - a). \end{aligned}$$

Der Anfangspunkt ist $\gamma(0) = a$ und der Endpunkt $\gamma(1) = b$.



Dieser Weg ist sozusagen die direkte Verbindung zwischen a und b .

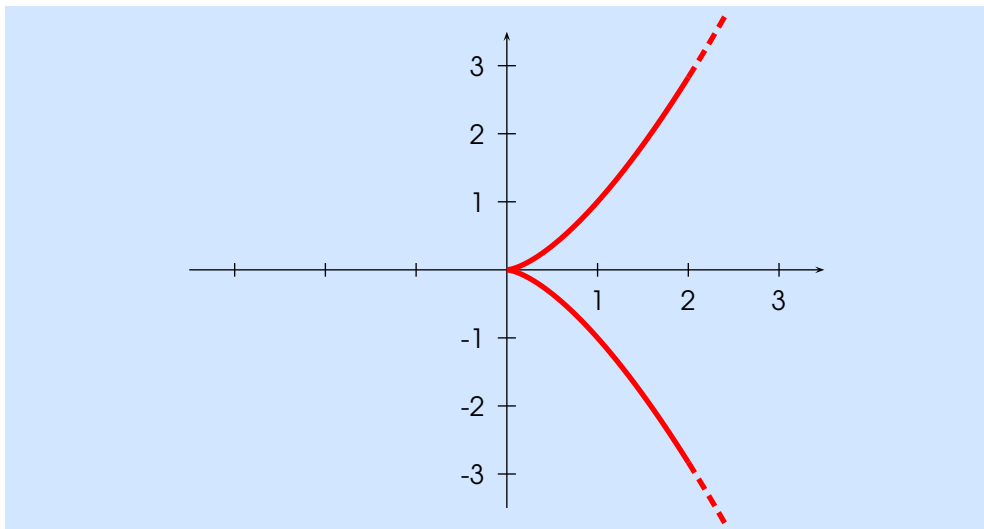
iv) Es sei

$$\begin{aligned}\gamma: \mathbb{R} &\longrightarrow \mathbb{R}^2 \\ t &\longmapsto (t^2, t^3).\end{aligned}$$

Man verifiziert leicht

$$\text{Spur}(\gamma) = \{ (x, y) \in \mathbb{R}^2 \mid x^3 = y^2 \}.$$

Diese Menge nennt man *Neilsche¹ Parabel*.



v) Einer stetigen Funktion

$$\begin{aligned}f: I &\longrightarrow \mathbb{R} \\ t &\longmapsto f(t)\end{aligned}$$

können wir den Weg

$$\begin{aligned}\gamma_f: I &\longrightarrow \mathbb{R}^2 \\ t &\longmapsto (t, f(t))\end{aligned}$$

zuordnen. Die Spur dieses Weges ist der Graph $\Gamma(f)$ der Funktion f (s. (12), Definition 3.1.1).

4.2 Längenmessung

Jetzt werden wir erklären, wie man die Länge eines Weges messen kann, indem man ihn durch stückweise lineare Wege annähert. Es soll also einem Weg $\gamma: I \rightarrow \mathbb{R}^n$ eine nichtnegative reelle Zahl $L(\gamma)$ zugeordnet werden. Für Punkte $a, b \in \mathbb{R}^n$ und den zuvor betrachteten Weg

$$\begin{aligned}\gamma: [0, 1] &\longrightarrow \mathbb{R}^n \\ t &\longmapsto a + t \cdot (b - a).\end{aligned}$$

legt man

$$L(\gamma) = \|b - a\|$$

¹William Neile (1637-1670), englischer Mathematiker

fest. Dabei stehe $\|\cdot\|$ für die euklidische Norm. Wir verfolgen nun den Ansatz, einen beliebigen Weg γ durch solche „Streckenzüge“ zu approximieren.

4.2.1 Definition. Es seien $a < b$ reelle Zahlen.

a) Eine *Teilung* des abgeschlossenen Intervalls $[a, b]$ ist ein Tupel $T = (t_0, \dots, t_N)$ von reellen Zahlen mit

$$a = t_0 < t_1 < \dots < t_{N-1} < t_N = b.$$

b) Es seien $\gamma: [a, b] \rightarrow \mathbb{R}^n$ ein Weg und $T = (t_0, \dots, t_N)$ eine Teilung von $[a, b]$. Man setzt

$$L_T(\gamma) := \sum_{i=1}^N \|\gamma(t_i) - \gamma(t_{i-1})\|.$$

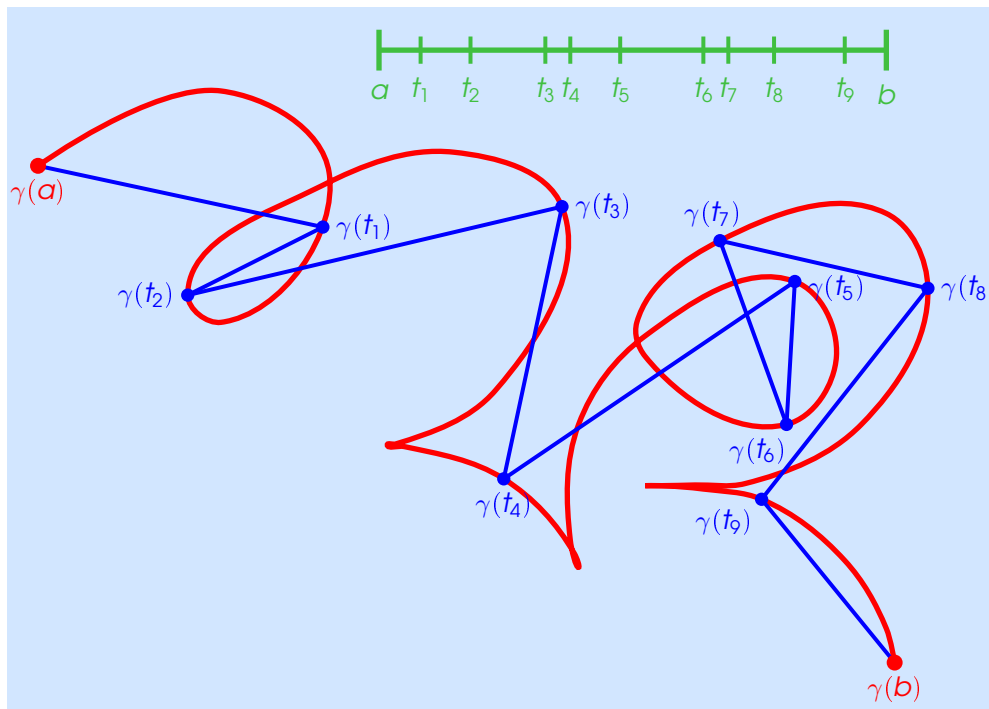
c) Für einen Weg $\gamma: [a, b] \rightarrow \mathbb{R}^n$ ist

$$L(\gamma) := \sup\{L_T(\gamma) \mid T \text{ eine Teilung von } [a, b]\} \in \mathbb{R} \cup \{\infty\}.$$

d) Der Weg $\gamma: [a, b] \rightarrow \mathbb{R}^n$ heißt *rektifizierbar*, wenn

$$L(\gamma) < \infty$$

gilt. In diesem Fall ist $L(\gamma)$ die *Länge* von γ .



4.2.2 Beispiel. Nicht jeder Weg ist rektifizierbar. Wir betrachten z.B. die Funktion

$$f: [0, 1] \rightarrow \mathbb{R} \\ t \mapsto \begin{cases} t \cdot \cos\left(\frac{\pi}{2t}\right), & \text{falls } t > 0 \\ 0, & \text{falls } t = 0 \end{cases}.$$

Diese Funktion ist stetig (s. (12), Beispiel 4.8.16, iii).

Mit dieser Funktion bilden wir den Weg

$$\begin{aligned} \gamma: [0, 1] &\longrightarrow \mathbb{R}^2 \\ t &\longmapsto (t, f(t)). \end{aligned}$$

Weiter benutzen wir die Teilungen $T_N = (t_0^N, \dots, t_{4N}^N)$ mit

$$t_0^N := 0, \quad t_k^N := \frac{1}{4N+1-k}, \quad k = 1, \dots, 4N, \quad N \in \mathbb{N}.$$

Die folgende Wertetabelle gibt die Funktionswerte von f an den Teilungspunkten wieder:

$$\begin{array}{c|cccccccc} t_k^N & 0 & \frac{1}{4N} & \frac{1}{4N-1} & \frac{1}{4N-2} & \cdots & \frac{1}{4} & \frac{1}{3} & \frac{1}{2} & 1 \\ \hline f(t_k^N) & 0 & \frac{1}{4N} & 0 & -\frac{1}{4N-2} & \cdots & \frac{1}{4} & 0 & -\frac{1}{2} & 0 \end{array}$$

Damit berechnen wir

$$L_{T_N}(\gamma) = \sum_{k=1}^{4N} \|(t_k^N, f(t_k^N)) - (t_{k-1}^N, f(t_{k-1}^N))\| \geq \sum_{k=1}^{4N} |f(t_k^N) - f(t_{k-1}^N)| = 2 \cdot \sum_{k=1}^{2N} \frac{1}{2k} = \sum_{k=1}^{2N} \frac{1}{k}.$$

Da die harmonische Reihe bekanntlich divergiert ((12), Beispiel 2.3.13), folgt

$$\lim_{N \rightarrow \infty} L_{T_N}(\gamma) = \infty,$$

so dass

$$L(\gamma) = \infty.$$

In Beispiel 4.1.2 haben wir gesehen, dass es verschiedene Wege mit derselben Spur gibt. Allerdings haben die beiden genannten Wege verschiedene Längen. Wir beschreiben jetzt eine Konstruktion, mit der man durch „Umparametrisieren“ aus einem Weg einen neuen macht, ohne dabei die Spur und die Länge zu ändern. In Satz 4.4.4 werden wir zeigen, dass gewisse Wege besonders schöne Parametrisierungen zulassen.

4.2.3 Definition. a) Es seien $[a, b]$ und $[c, d]$ Intervalle. Eine bijektive stetige Abbildung

$$\varphi: [a, b] \longrightarrow [c, d]$$

heißt *Parametertransformation*.

b) Zwei Wege $\gamma_1: [a, b] \longrightarrow \mathbb{R}^n$ und $\gamma_2: [c, d] \longrightarrow \mathbb{R}^n$ sind *äquivalent*, wenn es eine Parametertransformation $\varphi: [a, b] \longrightarrow [c, d]$ gibt, so dass $\gamma_1 = \gamma_2 \circ \varphi$.

c) Eine *Kurve* (in \mathbb{R}^n) ist eine Äquivalenzklasse² von Wegen³ in \mathbb{R}^n .

4.2.4 Bemerkung. i) Offenbar haben äquivalente Wege dieselbe Spur.

ii) Für eine Parametertransformation $\varphi: [a, b] \longrightarrow [c, d]$ gelten folgende Eigenschaften:

- φ ist streng monoton wachsend und $\varphi(a) = c$ und $\varphi(b) = d$ oder φ ist streng monoton fallend und $\varphi(a) = d$ und $\varphi(b) = c$. (S. (12), Aufgabe 6.10.3)

- Die Umkehrabbildung $\varphi^{-1}: [c, d] \longrightarrow [a, b]$ ist auch eine Parametertransformation. (Dazu muss man zeigen, dass φ^{-1} stetig ist. Übung: Das folgt aus der Kompaktheit abgeschlossener Intervalle.)

²Dabei haben wir vorweggenommen, dass es sich bei „ \sim “ um eine Äquivalenzrelation handelt (s. Bemerkung 4.2.4, iii).

³Hier betrachten wir nur Wege mit einem beschränkten abgeschlossenen Intervall als Definitionsbereich.

iii) Die Relation „ \sim “ ist eine Äquivalenzrelation auf der Menge der Wege in \mathbb{R}^n . Dies überprüft man wie folgt:

- Für $\gamma: [a, b] \rightarrow \mathbb{R}^n$ gilt $\gamma \sim \gamma$. Dafür ziehen wir die Parametertransformation $\varphi = \text{id}: [a, b] \rightarrow [a, b], t \mapsto t$, heran.

- Für Wege $\gamma_1: [a, b] \rightarrow \mathbb{R}^n$ und $\gamma_2: [c, d] \rightarrow \mathbb{R}^n$ mit $\gamma_1 \sim \gamma_2$ findet man $\gamma_2 \sim \gamma_1$. Sei dazu $\varphi: [a, b] \rightarrow [c, d]$ eine Parametertransformation mit $\gamma_1 = \gamma_2 \circ \varphi$. Es folgt $\gamma_2 = \gamma_1 \circ \varphi^{-1}$. Da $\varphi^{-1}: [c, d] \rightarrow [a, b]$ nach Teil i) eine Parametertransformation ist, erhält man $\gamma_2 \sim \gamma_1$.

- Für drei Wege $\gamma_i: [a_i, b_i] \rightarrow \mathbb{R}^n, i = 1, 2, 3$, folgt aus $\gamma_1 \sim \gamma_2$ und $\gamma_2 \sim \gamma_3$, dass $\gamma_1 \sim \gamma_3$. Seien $\varphi_1: [a_1, b_1] \rightarrow [a_2, b_2]$ und $\varphi_2: [a_2, b_2] \rightarrow [a_3, b_3]$ Parametertransformationen mit $\gamma_1 = \gamma_2 \circ \varphi_1$ und $\gamma_2 = \gamma_3 \circ \varphi_2$. Dann ist $\varphi_2 \circ \varphi_1: [a_1, b_1] \rightarrow [a_3, b_3]$ eine Parametertransformation mit

$$\gamma_3 \circ (\varphi_2 \circ \varphi_1) = (\gamma_3 \circ \varphi_2) \circ \varphi_1 = \gamma_2 \circ \varphi_1 = \gamma_1.$$

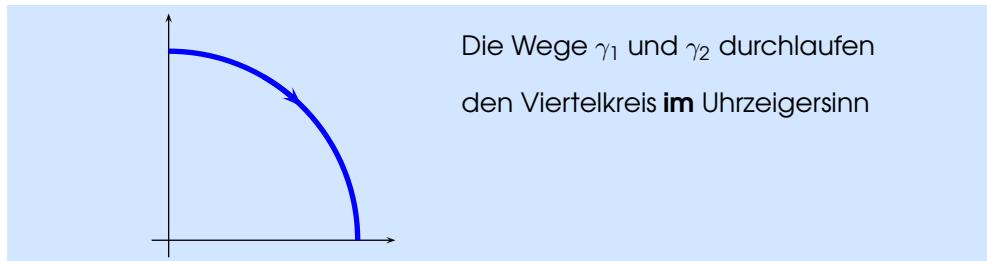
Damit sehen wir $\gamma_1 \sim \gamma_3$.

4.2.5 Beispiel. Wir betrachten den „Viertelkreis“

$$\begin{aligned} \gamma_1: [0, \frac{\pi}{2}] &\rightarrow \mathbb{R}^2 \\ t &\mapsto (\sin(t), \cos(t)) \end{aligned}$$

bzw.

$$\begin{aligned} \gamma_2: [0, 1] &\rightarrow \mathbb{R}^2 \\ t &\mapsto (t, \sqrt{1-t^2}). \end{aligned}$$



Die Abbildung

$$\begin{aligned} \varphi: [0, \frac{\pi}{2}] &\rightarrow [0, 1] \\ t &\mapsto \sin(t) \end{aligned}$$

ist eine Parametertransformation. Es gilt

$$\begin{aligned} \gamma_2 \circ \varphi: [0, \frac{\pi}{2}] &\rightarrow \mathbb{R}^2 \\ t &\mapsto (\sin(t), \sqrt{1-\sin^2(t)}). \end{aligned}$$

Da $1 - \sin^2(t) = \cos^2(t)$ ((12), Folgerung 4.8.10) und $\cos(t) \geq 0$ und damit $\sqrt{\cos^2(t)} = \cos(t)$ für $t \in [0, \pi/2]$ gilt, folgt $\gamma_1 = \gamma_2 \circ \varphi$ und $\gamma_1 \sim \gamma_2$.

Das folgende Lemma zeigt, warum uns die Äquivalenzrelation für die Längenmessung egal sein kann.

4.2.6 Lemma. Gegeben seien zwei Wege $\gamma_i: [a_i, b_i] \rightarrow \mathbb{R}^n, i = 1, 2$. Dann gilt:

$$\gamma_1 \sim \gamma_2 \Rightarrow L(\gamma_1) = L(\gamma_2).$$

Beweis. Sei $\varphi: [a_1, b_1] \rightarrow [a_2, b_2]$ eine Parametertransformation mit $\gamma_1 = \gamma_2 \circ \varphi$.

Wir betrachten zunächst den Fall, dass φ streng monoton wachsend ist (vgl. Bemerkung 4.2.4, ii). Für eine Teilung $T = (t_0, \dots, t_N)$ von $[a_1, b_1]$ ist $T_\varphi = (\varphi(t_0), \dots, \varphi(t_N))$ eine Teilung von $[a_2, b_2]$. Es gilt

$$\begin{aligned} L_T(\gamma_1) &= \sum_{i=1}^N \|\gamma_1(t_i) - \gamma_1(t_{i-1})\| = \sum_{i=1}^N \|(\gamma_2 \circ \varphi)(t_i) - (\gamma_2 \circ \varphi)(t_{i-1})\| \quad (4.1) \\ &= \sum_{i=1}^N \|\gamma_2(\varphi(t_i)) - \gamma_2(\varphi(t_{i-1}))\| = L_{T_\varphi}(\gamma_2) \leq L(\gamma_2). \end{aligned}$$

Somit folgern wir

$$L(\gamma_1) = \sup\{L_T(\gamma_1) \mid T \text{ Teilung von } [a_1, b_1]\} \leq L(\gamma_2).$$

Aus Symmetriegründen folgt $L(\gamma_2) \leq L(\gamma_1)$ und daher $L(\gamma_1) = L(\gamma_2)$.

Falls $\varphi: [a_1, b_1] \rightarrow [a_2, b_2]$ eine streng monoton fallende Parametertransformation ist, dann definiert eine Teilung $T = (t_0, \dots, t_N)$ von $[a_1, b_1]$ die Teilung $T_\varphi = (t'_0, \dots, t'_N)$ von $[a_2, b_2]$ mit $t'_i := \varphi(t_{N-i}), i = 0, \dots, N$. Es gilt weiter

$$\begin{aligned} L_T(\gamma_1) &= \sum_{i=1}^N \|\gamma_1(t_i) - \gamma_1(t_{i-1})\| = \sum_{i=1}^N \|\gamma_1(t_{i-1}) - \gamma_1(t_i)\| \\ &= \sum_{i=1}^N \|\gamma_1(t_{N-i+1}) - \gamma_1(t_{N-i})\| = L_{T_\varphi}(\gamma_2). \end{aligned}$$

Die letzte Gleichung ergibt sich aus einer zu (4.1) analogen Rechnung. Wie zuvor schließen wir auf $L(\gamma_1) = L(\gamma_2)$. \square

4.3 Differenzierbare Wege

Für einen differenzierbaren Weg $\gamma: [a, b] \rightarrow \mathbb{R}^n$ können wir das physikalische Gesetz „Weg = Geschwindigkeit \times Zeit“ als Motivation für die Längenmessung nehmen. Genauer gesagt erhalten wir die zurückgelegte Strecke als Integral des **Betrags** der Geschwindigkeit über das verstrichene Zeitintervall. Schließlich ist die Geschwindigkeit die Ableitung des Weges.

Schreiben wir den Weg $\gamma = (\gamma_1, \dots, \gamma_n)$ in seinen Komponenten, dann ist der *Geschwindigkeitsvektor* zum Zeitpunkt t durch $\gamma'(t) = (\gamma'_1(t), \dots, \gamma'_n(t))$ gegeben. Der Betrag der Geschwindigkeit zum Zeitpunkt t ist folglich $\|\gamma'(t)\|$.

Diese Diskussion motiviert die Definition

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

4.3.1 Definition. Es seien $\gamma = (\gamma_1, \dots, \gamma_n): I \rightarrow \mathbb{R}^n$ ein Weg und $a \in I$. Wir sagen, dass γ *in a differenzierbar* (bzw. *differenzierbar* bzw. *stetig differenzierbar*)⁴ ist, wenn die Funktion γ_i für alle $i \in \{1, \dots, n\}$ in a differenzierbar (bzw. differenzierbar bzw. stetig differenzierbar) ist. Die *Ableitung von γ in a* ist dann der Vektor

$$\gamma'(a) := (\gamma'_1(a), \dots, \gamma'_n(a)).$$

⁴Wir benutzen die Konventionen, die im Anschluss an Beispiel 3.7.3 erklärt wurden.

Für stetig differenzierbare Wege haben wir einen anderen Zugang zur Längenmessung:

4.3.2 Satz. Ist $\gamma: [a, b] \rightarrow \mathbb{R}^n$ ein **stetig differenzierbarer** Weg, dann ist γ rektifizierbar, und es gilt

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

4.3.3 Bemerkung. Da der Weg γ stetig differenzierbar ist, ist die Funktion

$$\begin{array}{ccccc} [a, b] & \longrightarrow & \mathbb{R}^n & \longrightarrow & \mathbb{R} \\ t & \longmapsto & (\gamma'_1(t), \dots, \gamma'_n(t)) & & \\ & & (x_1, \dots, x_n) & \longmapsto & \sqrt{x_1^2 + \dots + x_n^2} \end{array}$$

stetig. Das Integral im Satz existiert nach (12), Satz 5.1.9.

Beweis. Die Funktionen $\gamma'_i: [a, b] \rightarrow \mathbb{R}$, $i = 1, \dots, n$, sind nach Voraussetzung stetig und daher gleichmäßig stetig (Satz 3.5.4 oder (12), Behauptung 1 im Beweis von Satz 5.1.9). Zu einem vorgegebenen $\varepsilon > 0$ existiert ein $\delta > 0$, so dass

$$\forall i = 1, \dots, n \forall t, t' \in [a, b]: |t' - t| < \delta \Rightarrow |\gamma'_i(t') - \gamma'_i(t)| < \frac{\varepsilon}{n \cdot (b - a)}. \quad (4.2)$$

Behauptung. Es sei $T = (t_0, \dots, t_N)$ eine Teilung mit

$$|t_j - t_{j-1}| < \delta, \quad j = 1, \dots, N.$$

Dann gilt

$$\left| L_T(\gamma) - \int_a^b \|\gamma'(t)\| dt \right| \leq \varepsilon.$$

Zum Beweis der Behauptung fixieren wir vorübergehend einen Index $j \in \{1, \dots, N\}$ und setzen

$$\Delta_j := \left| \|\gamma(t_j) - \gamma(t_{j-1})\| - \int_{t_{j-1}}^{t_j} \|\gamma'(t)\| dt \right|.$$

Nach dem Mittelwertsatz der Differentialrechnung ((12), Folgerung 4.3.5) existieren Punkte $\tau_i \in (t_{j-1}, t_j)$, so dass

$$\gamma_i(t_j) - \gamma_i(t_{j-1}) = \gamma'_i(\tau_i) \cdot (t_j - t_{j-1}), \quad i = 1, \dots, n.$$

Der Mittelwertsatz der Integralrechnung ((12), Satz 5.1.17) garantiert die Existenz eines Punkts $\tau \in (t_{j-1}, t_j)$ mit

$$\int_{t_{j-1}}^{t_j} \|\gamma'(t)\| dt = \|\gamma'(\tau)\| \cdot (t_j - t_{j-1}).$$

Damit erarbeiten wir die folgende Abschätzung:

$$\begin{aligned} \Delta_j &= \left| \left\| (\gamma'_1(\tau_1), \dots, \gamma'_n(\tau_n)) \right\| - \left\| (\gamma'_1(\tau), \dots, \gamma'_n(\tau)) \right\| \right| \cdot (t_j - t_{j-1}) \\ &\leq \left\| (\gamma'_1(\tau_1) - \gamma'_1(\tau), \dots, \gamma'_n(\tau_n) - \gamma'_n(\tau)) \right\| \cdot (t_j - t_{j-1}) \\ &\leq \left(\sum_{i=1}^n |\gamma'_i(\tau_i) - \gamma'_i(\tau)| \right) \cdot (t_j - t_{j-1}) \\ &\leq \frac{\varepsilon}{b - a} \cdot (t_j - t_{j-1}). \end{aligned}$$

Für die letzte Ungleichung haben wir $|\tau_i - \tau| \leq |t_j - t_{j-1}| < \delta$ und (4.2) benutzt. Wegen

$$\left| L_T(\gamma) - \int_a^b \|\gamma'(t)\| dt \right| \leq \sum_{j=1}^N \Delta_j \leq \varepsilon$$

folgt die Behauptung.

Nun sei $T = (t_0, \dots, t_N)$ eine beliebige Teilung von $[a, b]$. Durch Einfügen weiterer Teilungspunkte gelangen wir zu einer Teilung $T' = (t'_0, \dots, t'_{N'})$ mit $t'_j - t'_{j-1} < \delta$, $j = 1, \dots, N'$. Die Dreiecksungleichung und die Behauptung führen uns zu dem Ergebnis

$$L_T(\gamma) \leq L_{T'}(\gamma) \leq \int_a^b \|\gamma'(t)\| dt + \varepsilon < \infty.$$

Daran sehen wir bereits, dass γ rektifizierbar ist. Da ε beliebig klein ausfallen kann, folgt auch die Abschätzung

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

Auf der anderen Seite gilt nach der Behauptung

$$L(\gamma) \geq L_T(\gamma) \geq \int_a^b \|\gamma'(t)\| dt - \varepsilon$$

für jede Teilung $T = (t_0, \dots, t_N)$ mit $t_j - t_{j-1} < \delta$, $j = 1, \dots, N$, so dass auch

$$L(\gamma) \geq \int_a^b \|\gamma'(t)\| dt$$

folgt. □

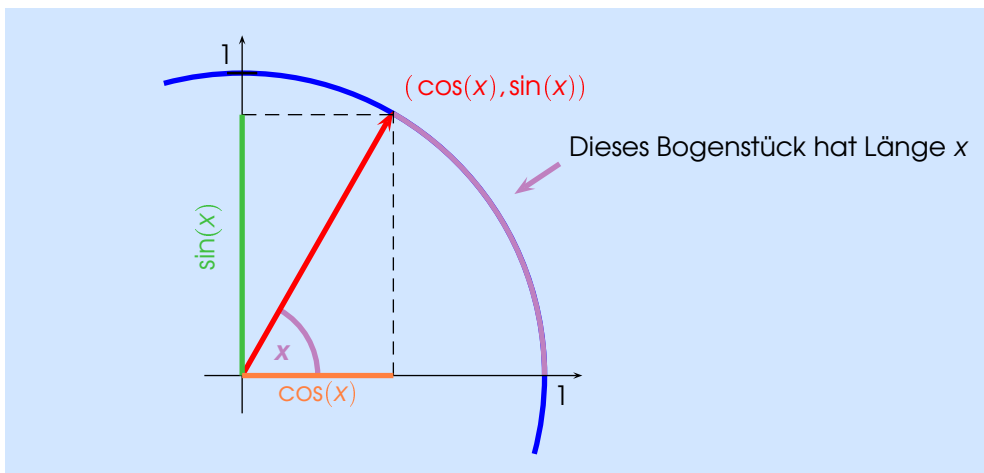
4.3.4 Beispiele. i) Wir beginnen mit dem Kreisweg

$$\begin{aligned} \gamma: [0, x] &\longrightarrow \mathbb{R}^2 \\ t &\longmapsto (r \cdot \cos(t), r \cdot \sin(t)) \end{aligned}$$

vom Radius $r > 0$ für $x \geq 0$. Wir berechnen

$$L(\gamma) = \int_0^x \sqrt{r^2 \cdot \sin^2(t) + r^2 \cdot \cos^2(t)} dt = r \cdot \int_0^x dt = r \cdot x.$$

Damit vervollständigt sich endlich unser Bild von den trigonometrischen Funktionen.

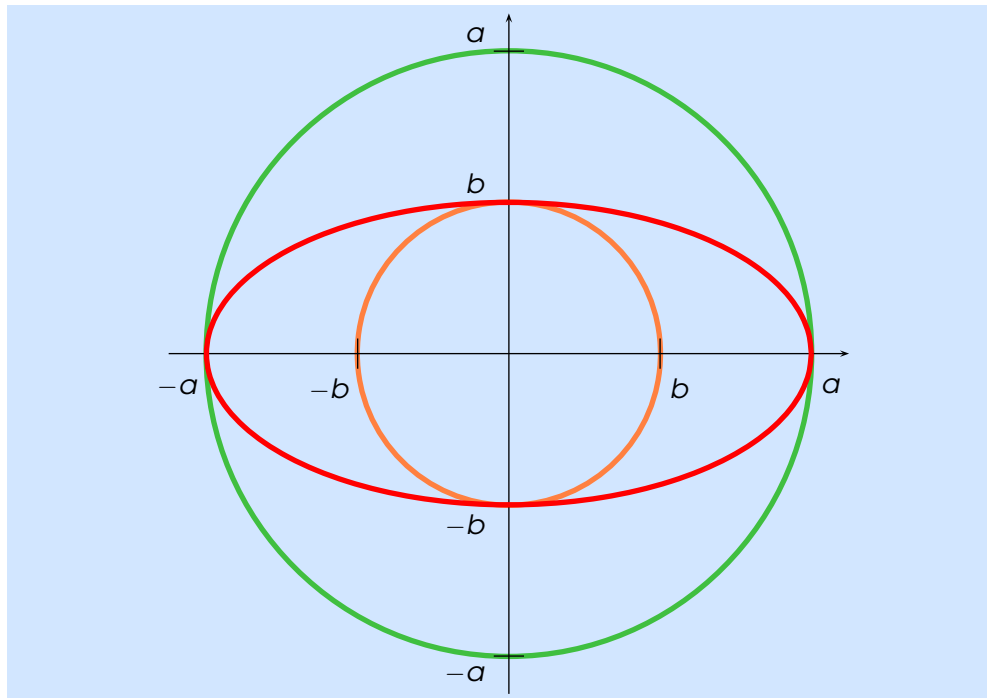


Insbesondere haben wir bewiesen, dass der Kreis vom Radius r den Umfang $2\pi r$ hat.

ii) Es seien $a \geq b > 0$. Die Ellipse mit den Halbachsen a und b ist durch den Weg

$$\begin{aligned} \gamma: [0, 2\pi] &\longrightarrow \mathbb{R}^2 \\ t &\longmapsto (a \cdot \cos(t), b \cdot \sin(t)) \end{aligned}$$

gegeben.



Die Zahl

$$k := \frac{\sqrt{a^2 - b^2}}{a}$$

ist die *numerische Exzentrizität* der Ellipse.

Die Berechnung des Umfangs der Ellipse führt zu dem Integral

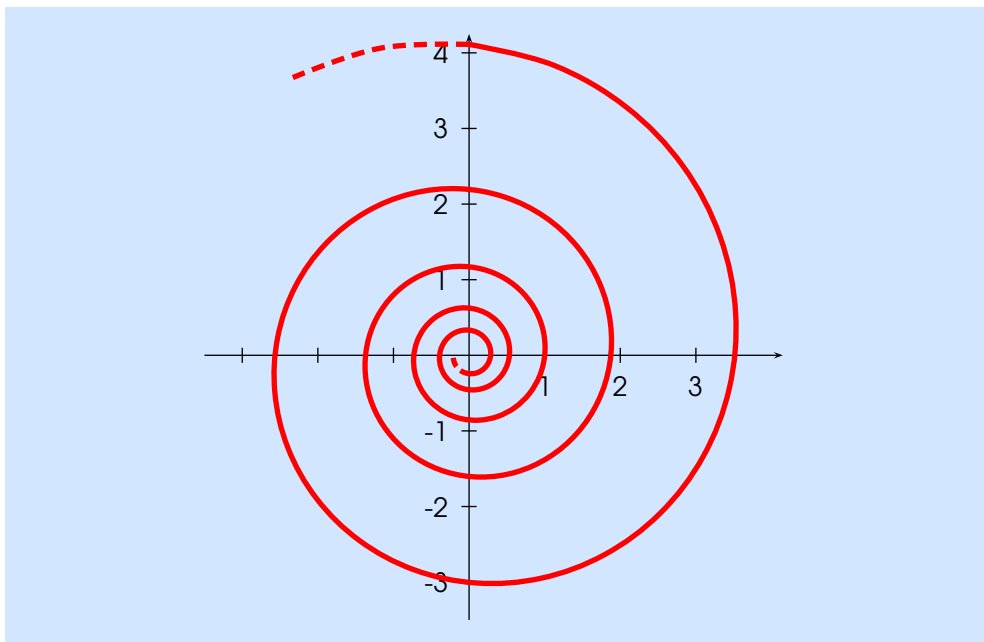
$$\begin{aligned} L(\gamma) &= \int_0^{2\pi} \sqrt{a^2 \cdot \sin^2(t) + b^2 \cdot \cos^2(t)} dt \\ &= \int_0^{2\pi} \sqrt{a^2 - (a^2 - b^2) \cdot \cos^2(t)} dt \\ &= a \cdot \int_0^{2\pi} \sqrt{1 - k^2 \cdot \cos^2(t)} dt \\ &= 4a \cdot \underbrace{\int_0^{\frac{\pi}{2}} \sqrt{1 - k^2 \cdot \cos^2(t)} dt}_{=: E(k)}. \end{aligned}$$

Das Integral $E(k)$ ist für $k \neq 0$ nicht elementar auswertbar.

iii) Die *logarithmische Spirale* zu dem Parameter $c > 0$ wird durch den Weg

$$\begin{aligned} \gamma: \mathbb{R} &\longrightarrow \mathbb{R}^2 \\ t &\longmapsto (\exp(ct) \cdot \cos(t), \exp(ct) \cdot \sin(t)) \end{aligned}$$

definiert.



Für Zahlen $a < b$ erhalten wir

$$\begin{aligned} & \int_a^b \sqrt{\exp(2ct) \cdot (c \cdot \cos(t) - \sin(t))^2 + \exp(2ct) \cdot (c \cdot \sin(t) + \cos(t))^2} dt \\ &= \int_a^b \sqrt{(1+c^2) \cdot \exp(2ct)} dt \\ &= \sqrt{1+c^2} \cdot \int_a^b \exp(ct) dt \\ &= \frac{\sqrt{1+c^2}}{c} \cdot (\exp(cb) - \exp(ca)) =: L_{a,b}(\gamma). \end{aligned}$$

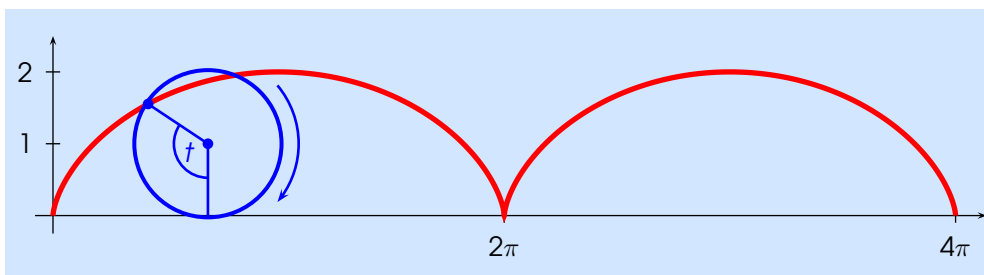
Man beachte⁵

$$\lim_{a \rightarrow \infty} L_{-a,0}(\gamma) = \frac{\sqrt{1+c^2}}{c} < \infty.$$

iv) Die *Zykloide* ist der Weg

$$\begin{aligned} \gamma: \mathbb{R} &\longrightarrow \mathbb{R}^2 \\ t &\longmapsto (t - \sin(t), 1 - \cos(t)). \end{aligned}$$

Dieser Weg beschreibt die Bahn eines Punkts auf dem Einheitskreis beim Abrollen desselben.



⁵Allgemein kann man mit dem Formalismus uneigentlicher Integrale die Länge gewisser Wege mit unbeschränktem Definitionsbereich messen.

Wir berechnen

$$\begin{aligned}
 \int_0^{2\pi} \|\gamma'(t)\| dt &= \int_0^{2\pi} \sqrt{1 - 2 \cdot \cos(t) + \cos^2(t) + \sin^2(t)} dt \\
 &= \int_0^{2\pi} \sqrt{2} \cdot \sqrt{1 - \cos(t)} dt \\
 &= \int_0^{2\pi} \sqrt{2} \cdot \sqrt{2 \cdot \sin^2\left(\frac{t}{2}\right)} dt \\
 &= 2 \cdot \int_0^{2\pi} \left| \sin\left(\frac{t}{2}\right) \right| dt \\
 &\stackrel{t=2u}{=} 4 \cdot \int_0^{\pi} \sin(u) du \\
 &= -4 \cdot \cos(u) \Big|_0^{\pi} = 8.
 \end{aligned}$$

Bei diesen Umformungen haben wir das Additionstheorem ((12), Satz 4.8.8) $\cos(t/2 + t/2) = \cos^2(t/2) - \sin^2(t/2)$ ausgenutzt und die Substitutionsregel ((12), Satz 5.4.1) verwendet.

4.4 Parametrisierung nach der Bogenlänge

Für einen Weg $\gamma: [a, b] \rightarrow \mathbb{R}^n$ haben wir den Zeitparameter $t \in [a, b]$ und die Länge $L(t) := L(\gamma|_{[a,t]})$ des bis zum Zeitpunkt t zurückgelegten Wegs. Falls γ stetig differenzierbar und **regulär** ist, kann man durch geeignete Umparametrisierung $L(t) = t$ erreichen.

4.4.1 Definition. Es sei $\gamma: I \rightarrow \mathbb{R}^n$ ein **stetig differenzierbarer** Weg. Ein Punkt $t \in I$ ist **regulär**, wenn $\gamma'(t) \neq 0$ gilt. Ansonsten heißt t **singulär**.

4.4.2 Beispiele. i) Der Newtonsche Knoten. Wir betrachten den Weg

$$\begin{aligned}
 \gamma: \mathbb{R} &\rightarrow \mathbb{R}^2 \\
 t &\mapsto (t^2 - 1, t^3 - t).
 \end{aligned}$$

Man vergewissert sich, dass

$$\text{Spur}(\gamma) = \{ (x, y) \in \mathbb{R}^2 \mid y^2 = x^3 + x^2 \}$$

gilt. Man beachte

$$\gamma(-1) = (0, 0) = \gamma(1).$$

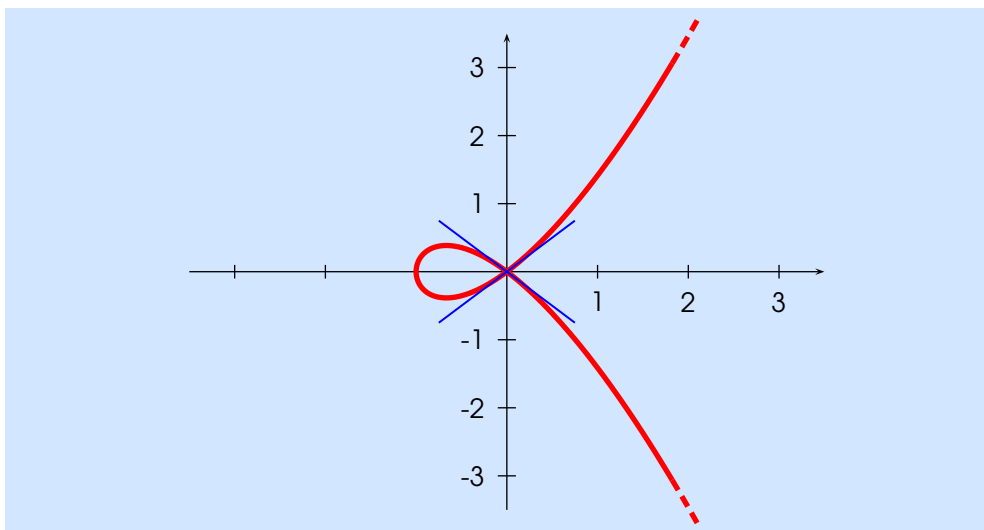
Ansonsten ist γ injektiv, d.h. $\gamma|_{\mathbb{R} \setminus \{\pm 1\}}$ ist injektiv. Es gilt

$$\gamma'(-1) = (-2, 2), \quad \gamma'(1) = (2, 2)$$

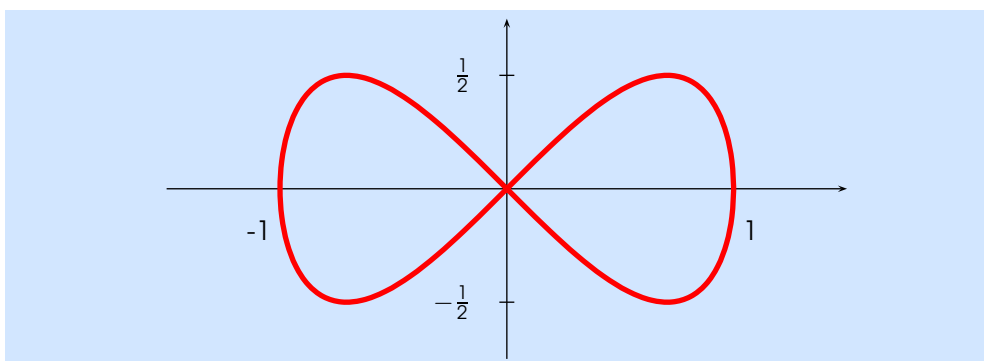
und

$$\gamma'(t) = (2t, 3t^2 - 1) \neq (0, 0), \quad t \in \mathbb{R}.$$

Jeder Punkt in \mathbb{R} ist daher ein regulärer Punkt für γ .



ii) Die Lemniskate von Gerono⁶.



Wir wollen die Menge

$$C := \{ (x, y) \in \mathbb{R}^2 \mid y^2 = x^2(1 - x^2) \}$$

parametrisieren, d.h. wir suchen einen Weg γ mit $\text{Spur}(\gamma) = C$. Für einen Punkt $(x, y) \in C$ mit $x \neq 0$ haben wir

$$\left(\frac{y}{x}\right)^2 = 1 - x^2, \quad \text{d.h.} \quad x^2 + \left(\frac{y}{x}\right)^2 = 1.$$

Deshalb können wir eine Zahl $t \in \mathbb{R}$ mit

$$x = \cos(t) \quad \text{und} \quad \frac{y}{x} = \sin(t)$$

finden. Damit erkennen wir, dass

$$\begin{aligned} \gamma: \mathbb{R} &\longrightarrow \mathbb{R}^2 \\ t &\longmapsto (\cos(t), \sin(t) \cdot \cos(t)) \end{aligned}$$

eine Parametrisierung der Lemniskate ist. Für diese Parametrisierung ist jeder Punkt regulär (Übung).

iii) Die Neilsche Parabel. Wir betrachten nochmals den Weg

$$\begin{aligned} \gamma: \mathbb{R} &\longrightarrow \mathbb{R}^2 \\ t &\longmapsto (t^2, t^3) \end{aligned}$$

⁶Camille-Christophe Gerono (1799 - 1891), französischer Mathematiker

aus Beispiel 4.1.2, iv). Wir finden

$$\gamma'(t) = (2t, 3t^2).$$

Damit ist $t = 0$ der einzige singuläre Punkt für die Neilsche Parabel.

4.4.3 Übung. Untersuchen Sie die Zykloide (Beispiel 4.3.4, iv) auf singuläre Punkte.

4.4.4 Satz (Parametrisierung nach der Bogenlänge). *Es sei $\gamma: I := [a, b] \rightarrow \mathbb{R}^n$ ein **stetig differenzierbarer** Weg, der in jedem Punkt $t \in I$ **regulär** ist. Dann gibt es einen zu γ äquivalenten Weg*

$$\tilde{\gamma}: [0, L(\gamma)] \rightarrow \mathbb{R}^n$$

mit

$$L(\tilde{\gamma}|_{[0,t]}) = t, \quad t \in [0, L(\gamma)].$$

Beweis. Wir betrachten

$$\begin{aligned} \varphi: [a, b] &\longrightarrow [0, L(\gamma)] \\ t &\longmapsto \int_a^t \|\gamma'(u)\| \, du. \end{aligned}$$

Nach (12), Satz 5.3.1, ist φ eine differenzierbare Abbildung mit

$$\varphi'(t) = \|\gamma'(t)\|, \quad t \in I.$$

(Man beachte Bemerkung 3.7.4 für die Randpunkte.) Unsere Voraussetzung impliziert

$$\|\gamma'(t)\| > 0, \quad t \in I.$$

Die Funktion φ ist somit nach (12), Folgerung 4.3.8, iii), streng monoton wachsend und folglich eine Parametertransformation.

Wir setzen nun

$$\tilde{\varphi} := \varphi^{-1}: [0, L(\gamma)] \longrightarrow [a, b]$$

und $\tilde{\gamma} = \gamma \circ \tilde{\varphi}$. Der Weg $\tilde{\gamma}$ ist nach Konstruktion äquivalent zu γ und erfüllt für $t \in [0, L(\gamma)]$ die Gleichung

$$L(\tilde{\gamma}|_{[0,t]}) = L(\gamma|_{[a, \tilde{\varphi}(t)]}) = \int_a^{\tilde{\varphi}(t)} \|\gamma'(u)\| \, du = \varphi(\tilde{\varphi}(t)) = t.$$

Dies ist die gewünschte Eigenschaft. □

Kapitel 5

Partielle Ableitungen

5.1 Richtungsableitungen

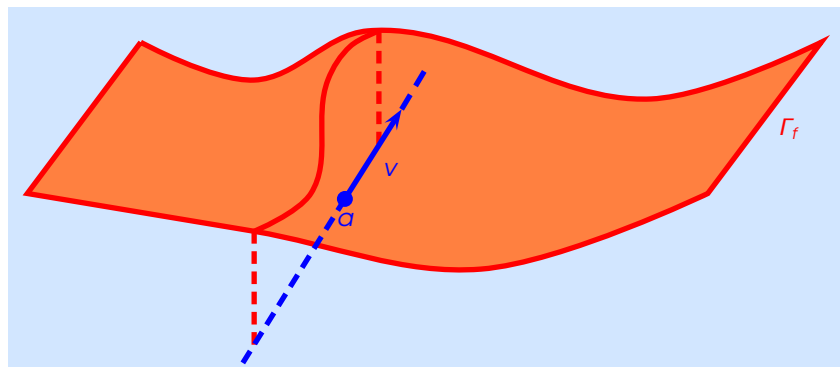
Als nächstes wollen wir für eine Funktion $f: \mathbb{R}^n \rightarrow \mathbb{R}$ den Begriff der Differenzierbarkeit in einem Punkt $a \in \mathbb{R}^n$ entwickeln.

Dafür bieten sich uns die folgenden Möglichkeiten:

i) Wir verwenden den bereits bekannten Begriff der Differenzierbarkeit für reelle Funktionen einer Veränderlichen, indem wir f z.B. auf Geraden in \mathbb{R}^n einschränken. Dieser Ansatz führt zu den Konzepten der Richtungsableitungen und der partiellen Differenzierbarkeit.

ii) Wir betrachten die natürliche Verallgemeinerung der Definition der Differenzierbarkeit. Damit gelangen wir zu der totalen Ableitung und der totalen Differenzierbarkeit der Funktion f .

Wir werden jetzt den ersten Ansatz vorstellen.



5.1.1 Definition. Es seien $B \subset \mathbb{R}^n$ eine offene Teilmenge, $f: B \rightarrow \mathbb{R}$ eine Funktion, $a \in B$ und $v \in \mathbb{R}^n \setminus \{0\}$ ein Vektor. Wir sagen, dass f im Punkt a in die Richtung v differenzierbar ist, wenn der Grenzwert

$$\lim_{t \rightarrow 0} \frac{f(a + t \cdot v) - f(a)}{t}$$

existiert.¹ In diesem Fall heißt der Grenzwert die (Richtungs-)Ableitung von f in Richtung v im Punkt a .

Schreibweise. $D_v f(a) := \lim_{t \rightarrow 0} \frac{f(a + t \cdot v) - f(a)}{t}$.

¹Da B offen ist und $a \in B$, gibt es ein $\varepsilon > 0$, so dass $a + t \cdot v \in B$ für alle $t \in (-\varepsilon, \varepsilon)$.

Genauer gesagt betrachten wir in dieser Definition die Differenzierbarkeit einer Funktion in einer Veränderlichen, die wir auf folgende Weise bilden: Wir haben den Weg

$$\begin{aligned} \gamma: \mathbb{R} &\longrightarrow \mathbb{R}^n \\ t &\longmapsto a + t \cdot v. \end{aligned}$$

Auf der offenen Menge $U := \gamma^{-1}(B) \subset \mathbb{R}$ erhalten wir die Funktion

$$g: U \xrightarrow{\gamma} B \xrightarrow{f} \mathbb{R}.$$

Wir fragen uns, ob die Funktion g in 0 differenzierbar ist.

5.1.2 Definition. Es seien $u = (u_1, \dots, u_n)$ und $v = (v_1, \dots, v_n)$ zwei Vektoren in \mathbb{R}^n . Die Zahl

$$\langle u, v \rangle := \sum_{i=1}^n u_i \cdot v_i$$

ist das *Standardskalarprodukt* von u und v .

5.1.3 Bemerkung. Aus der linearen Algebra ((13), §41) sind folgende Eigenschaften des Standardskalarprodukts bekannt, die sich ohne große Mühe nachprüfen lassen:

i) (Bilinearität) Für eine Zahl $\lambda \in \mathbb{R}$ und Vektoren $u, u', v, v' \in \mathbb{R}^n$ hat man:

- $\langle \lambda \cdot u, v \rangle = \lambda \cdot \langle u, v \rangle = \langle u, \lambda \cdot v \rangle,$
- $\langle u + u', v \rangle = \langle u, v \rangle + \langle u', v \rangle, \quad \langle u, v + v' \rangle = \langle u, v \rangle + \langle u, v' \rangle.$

ii) (Symmetrie) Für Vektoren $u, v \in \mathbb{R}^n$ gilt $\langle u, v \rangle = \langle v, u \rangle$.

iii) (Positive Definitheit) Für $u \in \mathbb{R}^n \setminus \{0\}$ folgt $\langle u, u \rangle > 0$.

5.1.4 Beispiel. Wir untersuchen die Funktion

$$\begin{aligned} f: \mathbb{R}^n &\longrightarrow \mathbb{R} \\ x &\longmapsto 1 - \langle x, x \rangle. \end{aligned}$$

Für $v \in \mathbb{R}^n \setminus \{0\}$ und $a \in \mathbb{R}^n$ berechnet man:

$$\begin{aligned} f(a + t \cdot v) &= 1 - \langle a + t \cdot v, a + t \cdot v \rangle \\ &= 1 - \langle a, a + t \cdot v \rangle - \langle t \cdot v, a + t \cdot v \rangle \\ &= 1 - \langle a, a \rangle - \langle a, t \cdot v \rangle - t \cdot \langle v, a + t \cdot v \rangle \\ &= 1 - \langle a, a \rangle - t \cdot \langle a, v \rangle - t \cdot \langle v, a \rangle - t \cdot \langle v, t \cdot v \rangle \\ &= 1 - \langle a, a \rangle - 2 \cdot t \cdot \langle a, v \rangle - t^2 \cdot \langle v, v \rangle. \end{aligned}$$

Es folgt

$$\lim_{t \rightarrow 0} \frac{f(a + t \cdot v) - f(a)}{t} = \lim_{t \rightarrow 0} (-2 \cdot \langle a, v \rangle - t \cdot \langle v, v \rangle) = -2 \cdot \langle a, v \rangle.$$

5.1.5 Eigenschaften der Richtungsableitung. Es seien $B \subset \mathbb{R}^n$ eine offene Teilmenge, $f, g: B \rightarrow \mathbb{R}$ Funktionen, $\lambda \in \mathbb{R}$, $v \in \mathbb{R}^n \setminus \{0\}$ und $a \in B$.

i) Wenn f in a in Richtung v differenzierbar ist, dann ist dies auch die Funktion $\lambda \cdot f$, und es gilt die Formel

$$D_v(\lambda \cdot f)(a) = \lambda \cdot D_v f(a).$$

ii) Sind f und g in a in Richtung v differenzierbar, dann folgt:

- Die Funktion $f + g$ ist in a in Richtung v differenzierbar mit

$$D_v(f + g)(a) = D_v f(a) + D_v g(a).$$

- Das Produkt $f \cdot g$ ist in a in Richtung v differenzierbar, und man hat

$$D_v(f \cdot g)(a) = (D_v f(a)) \cdot g(a) + f(a) \cdot (D_v g(a)).$$

- Ist weiter $g(x) \neq 0$, $x \in B$, dann ist die Funktion f/g ebenfalls in a in Richtung v differenzierbar, und die Richtungsableitung nimmt den Wert

$$D_v \left(\frac{f}{g} \right) (a) = \frac{(D_v f(a)) \cdot g(a) - f(a) \cdot (D_v g(a))}{g^2(a)}$$

an.

iii) Falls die Funktion f in a in Richtung v differenzierbar ist, dann ist sie auch in Richtung $\lambda \cdot v$ differenzierbar mit Richtungsableitung

$$D_{\lambda \cdot v} f(a) = \lambda \cdot (D_v f(a)).$$

Beweis. Die Eigenschaften i) und ii) folgen sofort aus den Ableitungsregeln (12), Satz 4.2.1. Für Eigenschaft iii) wendet man die Kettenregel auf die Funktionen

$$\begin{aligned} g: \mathbb{R} &\longrightarrow \mathbb{R} \\ t &\longmapsto \lambda \cdot t \end{aligned}$$

und

$$\begin{aligned} h: \mathbb{R} &\longrightarrow \mathbb{R} \\ t &\longmapsto f(a + t \cdot v) \end{aligned}$$

an. □

In der folgenden Definition bezeichne

$$e_i := (0, \dots, 0, \underset{\substack{\uparrow \\ i\text{-te Stelle}}}{1}, 0, \dots, 0)$$

den i -ten Standardbasisvektor von \mathbb{R}^n , $i = 1, \dots, n$.

5.1.6 Definition. Es seien $B \subset \mathbb{R}^n$ eine offene Teilmenge, $f: B \rightarrow \mathbb{R}$ eine Funktion und $a \in B$.

a) Sei $i \in \{1, \dots, n\}$. Wenn f in a in Richtung e_i differenzierbar ist, dann definiert man

$$f_{x_i}(a) := \frac{\partial}{\partial x_i} f(a) := \frac{\partial f}{\partial x_i}(a) := D_{e_i} f(a).$$

b) Die Funktion f ist in a *partiell differenzierbar*, wenn f für $i = 1, \dots, n$ in a in Richtung e_i differenzierbar ist. In diesem Fall ist $(\partial f / \partial x_i)(a)$ die *partielle Ableitung von f nach x_i in a* , $i = 1, \dots, n$.

c) Die Funktion f wird als *partiell differenzierbar* bezeichnet, wenn sie in jedem Punkt $a \in B$ partiell differenzierbar ist.

d) Die Funktion f ist in a *stetig partiell differenzierbar*, wenn sie partiell differenzierbar ist, und die Funktionen

$$\begin{aligned} \frac{\partial f}{\partial x_i}: B &\longrightarrow \mathbb{R} \\ x &\longmapsto \frac{\partial f}{\partial x_i}(x) \end{aligned}$$

für $i = 1, \dots, n$ in a stetig sind.

e) Wir sagen, dass f *stetig partiell differenzierbar* ist, wenn f in jedem Punkt $a \in B$ stetig partiell differenzierbar ist.

5.1.7 *Bemerkung.* Man beachte, dass wir von stetiger partieller Differenzierbarkeit in einem Punkt a nur sprechen können, wenn f zumindest in einer Umgebung von a partiell differenzierbar ist.²

5.1.8 *Beispiel.* i) Die Abstandsfunktion

$$\begin{aligned} r: \mathbb{R}^n &\longrightarrow \mathbb{R} \\ x = (x_1, \dots, x_n) &\longmapsto \|x\| = \sqrt{x_1^2 + \dots + x_n^2} \end{aligned}$$

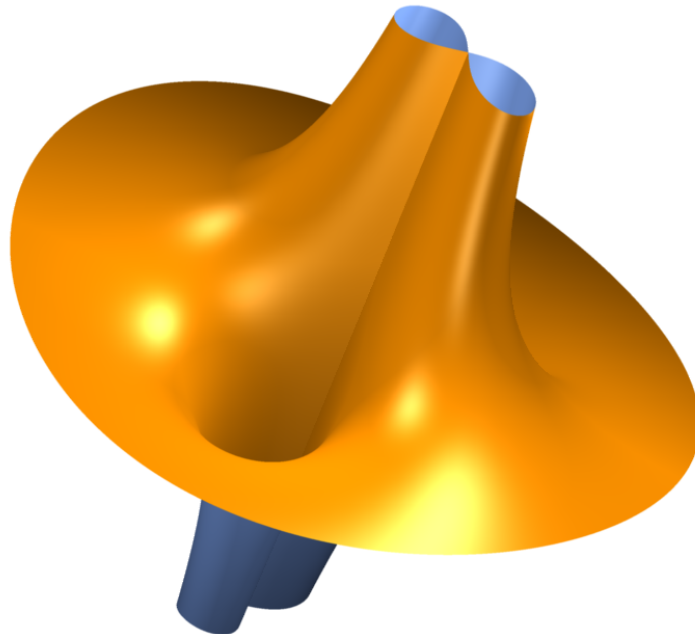
ist auf $\mathbb{R}^n \setminus \{0\}$ partiell differenzierbar. Aus der Kettenregel ((12), Satz 4.2.4) folgt

$$\frac{\partial r}{\partial x_i}(x_1, \dots, x_n) = 2 \cdot x_i \cdot \frac{1}{2 \cdot \sqrt{x_1^2 + \dots + x_n^2}} = \frac{x_i}{r(x)}, \quad i = 1, \dots, n.$$

ii) Wir behaupten, dass die Funktion

$$F: \mathbb{R}^n \longrightarrow \mathbb{R} \\ x \longmapsto \begin{cases} \frac{x_1 \cdot \dots \cdot x_n}{r(x)^{2n}}, & \text{falls } x \neq 0 \\ 0, & \text{falls } x = 0 \end{cases}$$

partiell differenzierbar ist.



²Alternativ könnte man stetige partielle Differenzierbarkeit in a definieren, indem man lediglich fordert, dass f in einer kleinen Umgebung von a partiell differenzierbar ist und die auf dieser kleinen Umgebung definierten partiellen Ableitungen in a stetig sind. Von dieser Möglichkeit wollen wir aber im Folgenden keinen Gebrauch machen.

Für $x = (x_1, \dots, x_n) \in \mathbb{R}^n \setminus \{0\}$ berechnen wir

$$\begin{aligned} \frac{\partial F}{\partial x_i}(x) &= \frac{x_1 \cdots x_{i-1} \cdot x_{i+1} \cdots x_n \cdot r(x)^{2n} - x_1 \cdots x_n \cdot \frac{\partial r}{\partial x_i}(x) \cdot 2n \cdot r(x)^{2n-1}}{r(x)^{4n}} \\ &= \frac{x_1 \cdots x_{i-1} \cdot x_{i+1} \cdots x_n}{r(x)^{2n}} - 2n \cdot \frac{x_1 \cdots x_{i-1} \cdot x_i^2 \cdot x_{i+1} \cdots x_n}{r(x)^{2n+2}}. \end{aligned}$$

Die partielle Differenzierbarkeit in 0 ergibt sich aus der Formel

$$\lim_{t \rightarrow 0} \frac{F(t \cdot e_i) - F(0)}{t} = \lim_{t \rightarrow 0} \frac{0}{t^{2n}} = 0.$$

Überraschenderweise ist die Funktion F in 0 **nicht stetig!** Dazu betrachten wir die Folge $(a_k)_{k \geq 1}$ mit

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

Sie konvergiert gegen 0, und es gilt

$$\begin{aligned} r(a_k)^{2n} &= \left(\frac{\sqrt{n}}{k} \right)^{2n} = \frac{n^n}{k^{2n}}, \\ F(a_k) &= \frac{\left(\frac{1}{k} \right)^n}{\left(\frac{\sqrt{n}}{k} \right)^{2n}} = \left(\frac{k}{n} \right)^n, \quad k \geq 1. \end{aligned}$$

Offenbar folgt

$$\lim_{k \rightarrow \infty} F(a_k) = \infty.$$

Die Vorteile der partiellen Differenzierbarkeit liegen auf der Hand: Wir können alle Techniken und Erkenntnisse aus der Theorie der Funktionen einer Variablen anwenden, um eine Funktion auf partielle Differenzierbarkeit zu untersuchen und die partiellen Ableitungen zu bestimmen. Das obige Beispiel zeigt jedoch, dass der Begriff der partiellen Differenzierbarkeit in gewisser Weise „zu schwach“ ist.³ Einer anderen Merkwürdigkeit werden wir in Beispiel 5.3.3 begegnen. Weiter unten werden wir sehen, dass stetige partielle Differenzierbarkeit sehr gute Eigenschaften hat. Der Begriff der totalen Differenzierbarkeit ist zwischen partieller und stetiger partieller Differenzierbarkeit angesiedelt (s. Bemerkung 6.2.3, ii).

5.2 Der Gradient

Aus den partiellen Ableitungen einer Funktion in einem Punkt bauen wir einen Vektor auf, den sogenannten Gradienten. Anschaulich zeigt der Gradient in Richtung des steilsten Anstiegs bzw. Abfalls der Funktion in dem fraglichen Punkt. Der Gradient stellt auch die Verbindung zwischen partieller und totaler Differenzierbarkeit her.

³Die Tatsache, dass der Begriff der partiellen Differenzierbarkeit von der Auswahl des Koordinatensystems abhängt, sollte einen bereits zweifeln lassen.

5.2.1 Definition. Es seien $B \subset \mathbb{R}^n$ eine offene Teilmenge, $f: B \rightarrow \mathbb{R}$ eine Funktion und $a \in B$. Die Funktion f sei im Punkt a partiell differenzierbar. Der Vektor

$$\text{Grad}f(a) := \left(\frac{\partial f}{\partial x_1}(a), \dots, \frac{\partial f}{\partial x_n}(a) \right)$$

heißt der *Gradient* von f im Punkt a .

5.2.2 Bemerkung (Der Nabla-Operator). Es sei $B \subset \mathbb{R}^n$ eine offene Menge. Wir betrachten die Vektorräume

$$\mathcal{E}^1(B) := \left\{ f: B \rightarrow \mathbb{R} \mid f \text{ stetig partiell differenzierbar} \right\}$$

und

$$\mathcal{E}^0(B, \mathbb{R}^n) = \left\{ h: B \rightarrow \mathbb{R}^n \mid h \text{ stetig} \right\}.$$

Dann ist

$$\begin{aligned} \nabla: \mathcal{E}^1(B) &\longrightarrow \mathcal{E}^0(B, \mathbb{R}^n) \\ f &\longmapsto (\nabla f: x \longmapsto \nabla f(x) := \text{Grad}f(x)) \end{aligned}$$

eine lineare Abbildung, der sogenannte ∇ -Operator.

5.2.3 Satz. Es seien $B \subset \mathbb{R}^n$ eine offene Teilmenge, $f: B \rightarrow \mathbb{R}$ eine Funktion und $a = (a_1, \dots, a_n) \in B$. Die Funktion f sei partiell differenzierbar und in a **stetig** partiell differenzierbar. Dann existiert für jeden Vektor $v \in \mathbb{R}^n \setminus \{0\}$ die Ableitung von f in Richtung v in a , und es gilt die Formel:

$$D_v f(a) = \langle \text{Grad}f(a), v \rangle.$$

5.2.4 Bemerkung. i) Dieser Satz ist ein Spezialfall der Kettenregel, die wir in Satz 6.3.1 in voller Allgemeinheit formulieren werden.

ii) In Beispiel 5.1.8, ii), wurde eine partiell differenzierbare Funktion $F: \mathbb{R}^n \rightarrow \mathbb{R}$ vorgestellt, deren Ableitung in 0 in Richtung $v = (1, \dots, 1)$ nicht existiert. Auf die Stetigkeit der partiellen Ableitung in a kann man in der obigen Formulierung nicht verzichten.

Beweis. Der Beweis des Satzes beruht auf der folgenden Aussage:

Behauptung. Unter den Voraussetzungen des Satzes gilt: Es seien $\gamma = (\gamma_1, \dots, \gamma_n): I \rightarrow B$ ein Weg und $t_0 \in I$ mit $\gamma(t_0) = a$. Wenn γ in t_0 differenzierbar ist, dann ist auch die Verknüpfung $f \circ \gamma: I \rightarrow \mathbb{R}$ in t_0 differenzierbar, und man hat

$$(f \circ \gamma)'(t_0) = \langle \text{Grad}f(a), \gamma'(t_0) \rangle.$$

Unter Annahme der Behauptung verfahren wir wie folgt: Wir betrachten den Weg

$$\begin{aligned} \gamma: \mathbb{R} &\longrightarrow \mathbb{R}^n \\ t &\longmapsto a + t \cdot v. \end{aligned}$$

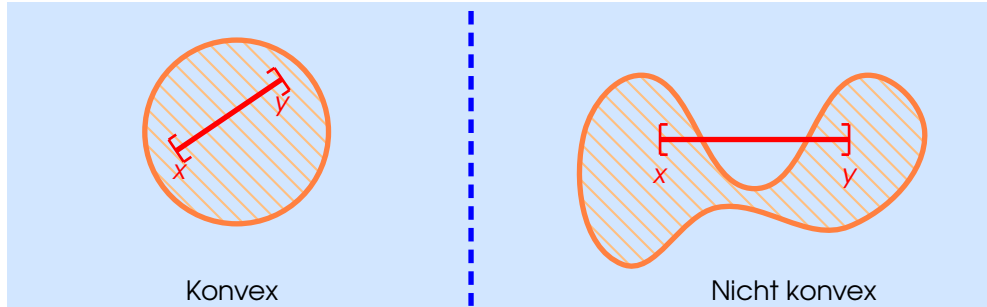
Da γ stetig ist, $\gamma(0) = a \in B$ und B offen ist, gibt es ein $\varepsilon > 0$ mit $\gamma((-\varepsilon, \varepsilon)) \subset B$. Der Weg $\gamma|_{(-\varepsilon, \varepsilon)}: (-\varepsilon, \varepsilon) \rightarrow B, t \mapsto a + t \cdot v$, ist differenzierbar mit $\gamma'|_{(-\varepsilon, \varepsilon)}(0) = v$. Der Satz folgt daher aus der Behauptung.

Jetzt bleibt die Behauptung zu beweisen. Wir können B durch $B(a, \varepsilon)$ für ein geeignetes $\varepsilon > 0$ ersetzen. Der Ball $B(a, \varepsilon)$ ist *konvex*, d.h. für $x, y \in B(a, \varepsilon)$ gilt

$$[x, y] := \left\{ (1-t) \cdot x + t \cdot y \mid t \in [0, 1] \right\} \subset B(a, \varepsilon).$$

Denn für $t \in [0, 1]$ berechnet man

$$\begin{aligned} \|(1-t) \cdot x + t \cdot y - a\| &= \|(1-t) \cdot (x-a) + t \cdot (y-a)\| \\ &\leq (1-t) \cdot \|x-a\| + t \cdot \|y-a\| < \varepsilon. \end{aligned}$$



Sei $x \in B(a, \varepsilon)$. Wir definieren $z_0 := a$, $z_1 := a + (x_1 - a_1) \cdot e_1, \dots, z_n := z_{n-1} + (x_n - a_n) \cdot e_n = x$. Die Vektoren z_0, \dots, z_n sind in $B(a, \varepsilon)$ enthalten, so dass wegen der Konvexität auch

$$[z_{i-1}, z_i] \subset B(a, \varepsilon), \quad i = 1, \dots, n,$$

gilt.

Wir wenden den Mittelwertsatz der Differentialrechnung ((12), Folgerung 4.3.5) auf die Funktion

$$\begin{aligned} h_i: [0, 1] &\longrightarrow \mathbb{R} \\ t &\longmapsto f(z_{i-1} + t \cdot (x_i - a_i) \cdot e_i) \end{aligned}$$

und $i = 1, \dots, n$ an. Wir können Punkte $c_i(x) \in [z_{i-1}, z_i]$ mit

$$f(z_i) - f(z_{i-1}) = \frac{\partial f}{\partial x_i}(c_i(x)) \cdot (x_i - a_i)$$

wählen, $i = 1, \dots, n$. Die Summation all dieser Gleichungen führt zu

$$f(x) - f(a) = \sum_{i=1}^n \frac{\partial f}{\partial x_i}(c_i(x)) \cdot (x_i - a_i).$$

Ferner definieren wir die Funktion

$$\begin{aligned} \delta = (\delta_1, \dots, \delta_n): B(a, \varepsilon) &\longrightarrow \mathbb{R}^n \\ x &\longmapsto \left(\frac{\partial f}{\partial x_1}(c_1(x)) - \frac{\partial f}{\partial x_1}(a), \dots, \frac{\partial f}{\partial x_n}(c_n(x)) - \frac{\partial f}{\partial x_n}(a) \right). \end{aligned}$$

Wir fassen zusammen:

$$\boxed{f(x) = f(a) + \langle \text{Grad}f(a), x - a \rangle + \langle \delta(x), x - a \rangle.} \quad (5.1)$$

Die Stetigkeit der partiellen Ableitungen in a garantiert

$$\lim_{x \rightarrow a} \delta(x) = 0. \quad (5.2)$$

Wegen der Differenzierbarkeit von γ in t_0 gibt es eine Abbildung $\beta: I \longrightarrow \mathbb{R}^n$ mit

$$\gamma(t) - \gamma(t_0) = \gamma'(t_0) \cdot (t - t_0) + \beta(t) \cdot (t - t_0)$$

und

$$\lim_{t \rightarrow t_0} \beta(t) = 0. \quad (5.3)$$

Jetzt berechnen wir

$$\begin{aligned} (f \circ \gamma)(t) - (f \circ \gamma)(t_0) &= f(\gamma(t)) - f(a) \\ &= \langle \text{Grad}f(a), \gamma(t) - \gamma(t_0) \rangle + \langle \delta(\gamma(t)), \gamma(t) - \gamma(t_0) \rangle \\ &= \langle \text{Grad}f(a), \gamma'(t_0) \cdot (t - t_0) + \beta(t) \cdot (t - t_0) \rangle + \\ &\quad + \langle \delta(\gamma(t)), \gamma'(t_0) \cdot (t - t_0) + \beta(t) \cdot (t - t_0) \rangle \\ &= \langle \text{Grad}f(a), \gamma'(t_0) \rangle \cdot (t - t_0) + \\ &\quad + \underbrace{\left(\langle \text{Grad}f(a), \beta(t) \rangle + \langle \delta(\gamma(t)), \gamma'(t_0) + \beta(t) \rangle \right)}_{=: h} \cdot (t - t_0). \end{aligned}$$

Die Grenzwerte (5.2) und (5.3) zeigen, dass

$$\lim_{t \rightarrow t_0} h(t) = 0.$$

Es folgt, dass die Abbildung $f \circ \gamma$ in t_0 differenzierbar ist, und die Ableitung in t_0 durch

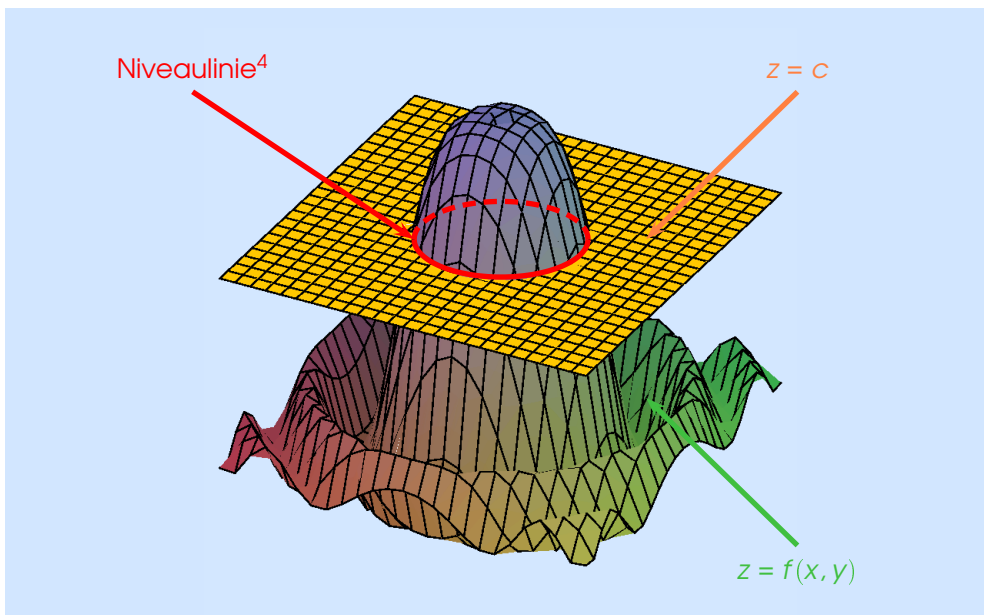
$$(f \circ \gamma)'(t_0) = \langle \text{Grad}f(a), \gamma'(t_0) \rangle$$

gegeben ist. □

Wir betrachten wieder eine offene Menge $B \subset \mathbb{R}^n$ und eine Funktion $f: B \rightarrow \mathbb{R}$. Die Mengen der Form

$$f^{-1}(c) := \{x \in B \mid f(x) = c\}, \quad c \in \mathbb{R},$$

werden die *Niveaumengen* von f genannt.



⁴Die Niveaumenge im Sinne der Definition ist die Projektion dieser Linie auf die (x, y) -Ebene.

5.2.5 Beispiel. Für die Funktion

$$\begin{aligned} f: \mathbb{R}^3 &\longrightarrow \mathbb{R} \\ x &\longmapsto 1 - \langle x, x \rangle \end{aligned}$$

und $c \in \mathbb{R}$ gilt

$$\begin{aligned} f^{-1}(c) &= \left\{ x = (x_1, x_2, x_3) \in \mathbb{R}^3 \mid x_1^2 + x_2^2 + x_3^2 = 1 - c \right\} \\ &= \begin{cases} \emptyset, & \text{falls } c > 1 \\ \{0\}, & \text{falls } c = 1 \\ \text{Sphäre vom Radius } \sqrt{1 - c}, & \text{falls } c < 1 \end{cases}. \end{aligned}$$

5.2.6 Bemerkung. Wenn $f: B \rightarrow \mathbb{R}$ stetig ist, dann sind die Niveaumengen von f abgeschlossene Teilmengen von B (Aufgabe A.4.1).

5.2.7 Satz. Es seien $B \subset \mathbb{R}^n$ eine offene Teilmenge, $f: B \rightarrow \mathbb{R}$ eine **stetig** partiell differenzierbare Funktion und $a \in B$ ein Punkt mit $\text{Grad}f(a) \neq 0$.

i) Der Gradient $\text{Grad}f(a)$ zeigt in die Richtung, in der f am schnellsten wächst bzw. abnimmt⁵. Genauer gesagt wird das Maximum

$$\max \left\{ D_v f(a) \mid \|v\| = 1 \right\}$$

in dem Vektor

$$\frac{\text{Grad}f(a)}{\|\text{Grad}f(a)\|}$$

angenommen.

ii) Der Vektor $\text{Grad}f(a)$ steht auf der Niveaufläche $N := f^{-1}(f(a))$ senkrecht. Das bedeutet, dass für jedes positive ε und jeden differenzierbaren Weg $\gamma: (-\varepsilon, \varepsilon) \rightarrow \mathbb{R}^n$ mit $\gamma(0) = a$ und $\gamma(t) \in N$, $t \in (-\varepsilon, \varepsilon)$,

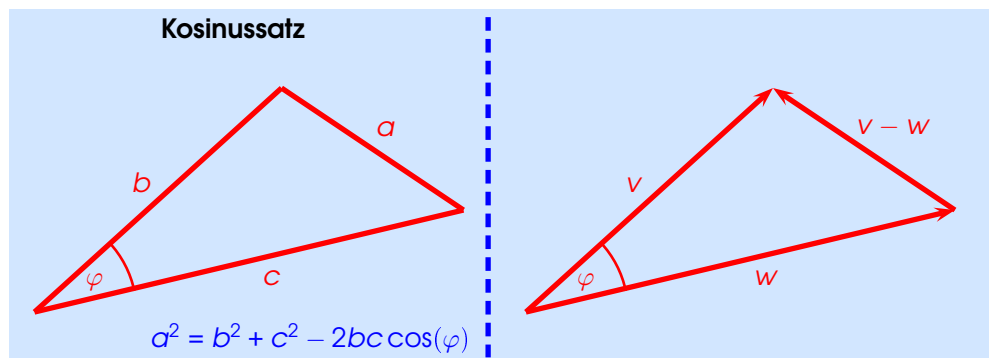
$$\langle \text{Grad}f(a), \gamma'(0) \rangle = 0$$

gilt.

Bevor wir den Beweis des Satzes angeben, erinnern wir an folgende Tatsache: Der Winkel $\varphi \in [0, \pi]$, der von zwei Vektoren $v, w \in \mathbb{R}^n \setminus \{0\}$ eingeschlossen wird, wird mit der Formel

$$\cos(\varphi) = \frac{\langle v, w \rangle}{\|v\| \cdot \|w\|} \quad (5.4)$$

ermittelt. Hintergrund ist der **Kosinussatz**:



⁵Abfall in Richtung v tritt hier als Anstieg in Richtung $-v$ in Erscheinung (s. Eigenschaft 5.1.5, iii).

Er wird auf die Vektoren v , w und $v - w$ angewandt:

$$\|v - w\|^2 = \|v\|^2 + \|w\|^2 - 2 \cdot \|v\| \cdot \|w\| \cdot \cos(\varphi).$$

Da weiter

$$\|v - w\|^2 = \langle v - w, v - w \rangle = \langle v, v \rangle - 2\langle v, w \rangle + \langle w, w \rangle = \|v\|^2 + \|w\|^2 - 2 \cdot \langle v, w \rangle,$$

folgt Gleichung (5.4). Insbesondere stehen zwei nichttriviale Vektoren v und w genau dann aufeinander senkrecht, in Zeichen $v \perp w$, wenn $\varphi = \pi/2$, d.h. $\cos(\varphi) = \langle v, w \rangle = 0$, gilt.

Beweis von Satz 5.2.7. i) Wir wenden Satz 5.2.3 an: Für jeden Vektor $v \in \mathbb{R}^n$ mit $\|v\| = 1$ gilt demnach

$$\begin{aligned} D_v f(a) &= \langle \text{Grad}f(a), v \rangle \\ &\stackrel{(5.4)}{=} \| \text{Grad}f(a) \| \cdot \|v\| \cdot \cos(\varphi) \\ &= \| \text{Grad}f(a) \| \cdot \cos(\varphi). \end{aligned}$$

Dieser Ausdruck wird maximal, wenn $\cos(\varphi) = 1$ gilt. Eingedenk der Tatsache $\varphi \in [0, \pi]$ schließen wir daraus $\varphi = 0$, d.h. $v = \text{Grad}f(a) / \| \text{Grad}f(a) \|$.

ii) Da $(f \circ \gamma)(t) = f(a)$, $t \in (-\varepsilon, \varepsilon)$, $f \circ \gamma$ also konstant ist, folgt

$$0 = (f \circ \gamma)'(0) = \langle \text{Grad}f(a), \gamma'(0) \rangle.$$

Die zweite Gleichung ergibt sich dabei aus der Behauptung im Beweis von Satz 5.2.3 □

5.3 Höhere partielle Ableitungen

In der Analysis einer Veränderlichen spielen Ableitungen höherer Ordnungen eine wichtige Rolle. Sie lassen sich mit partiellen Ableitungen leicht erklären. Ein interessantes neues Phänomen wird in Beispiel 5.3.3 geschildert.

5.3.1 Definition. Es seien $B \subset \mathbb{R}^n$ eine offene Teilmenge, $f: B \rightarrow \mathbb{R}$ eine Abbildung und $a \in B$.

a) Die Funktion f ist *in a zweimal partiell differenzierbar*, wenn f auf B partiell differenzierbar ist und die Funktionen

$$\begin{aligned} \frac{\partial f}{\partial x_i} : B &\rightarrow \mathbb{R} \\ x &\mapsto \frac{\partial f}{\partial x_i}(x) \end{aligned}$$

für $i = 1, \dots, n$ in a partiell differenzierbar sind.

Schreibweise. $\frac{\partial^2}{\partial x_i \partial x_j} f(a) := \frac{\partial}{\partial x_i} \left(\frac{\partial f}{\partial x_j} \right) (a)$, $i, j \in \{1, \dots, n\}$; $\frac{\partial^2}{\partial x_j^2} f(a) := \frac{\partial^2}{\partial x_j \partial x_j} f(a)$, $i = 1, \dots, n$.

b) Die Funktion f ist *zweimal partiell differenzierbar*, wenn f in jedem Punkt $a \in B$ zweimal partiell differenzierbar ist.

c) Wir sagen, dass f in a zweimal stetig partiell differenzierbar ist, wenn f zweimal partiell differenzierbar ist und die Funktionen

$$\begin{aligned} \frac{\partial^2 f}{\partial x_i \partial x_j} : B &\longrightarrow \mathbb{R} \\ x &\longmapsto \frac{\partial^2 f}{\partial x_i \partial x_j}(x) \end{aligned}$$

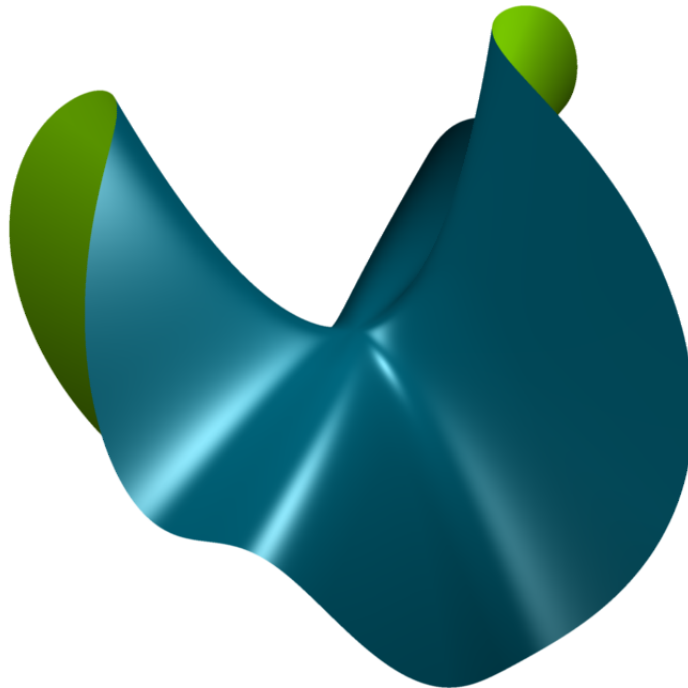
für alle $i, j \in \{1, \dots, n\}$ in a stetig sind.

d) Die Funktion f ist zweimal stetig partiell differenzierbar, wenn sie in jedem Punkt $a \in B$ zweimal stetig partiell differenzierbar ist.

5.3.2 Bemerkung. Induktiv definiert man (stetige) partielle Differenzierbarkeit der Ordnung n (in einem Punkt), $n \geq 1$ (s. Aufgabe A.7.3).

5.3.3 Beispiel. Wir untersuchen die Funktion

$$f: \mathbb{R}^2 \longrightarrow \mathbb{R} \\ (x, y) \longmapsto \begin{cases} x \cdot y \cdot \frac{x^2 - y^2}{x^2 + y^2}, & \text{falls } (x, y) \neq 0 \\ 0, & \text{falls } (x, y) = 0 \end{cases} .$$



Auf $\mathbb{R}^n \setminus \{0\}$ gilt

$$\begin{aligned} \frac{\partial f}{\partial x}(x, y) &= \frac{\partial}{\partial x} \left(\frac{x^3 y - x y^3}{x^2 + y^2} \right) \\ &= \frac{(3x^2 y - y^3) \cdot (x^2 + y^2) - (x^3 y - x y^3) \cdot 2x}{(x^2 + y^2)^2} \\ &= \frac{3x^4 y + 3x^2 y^3 - x^2 y^3 - y^5 - 2x^4 y + 2x^2 y^3}{(x^2 + y^2)^2} \\ &= \frac{x^4 y + 4x^2 y^3 - y^5}{(x^2 + y^2)^2} . \end{aligned}$$

Im Nullpunkt hat man

$$\frac{\partial f}{\partial x}(0,0) = \lim_{t \rightarrow 0} \frac{f(t,0) - f(0,0)}{t} = \lim_{t \rightarrow 0} \frac{0 - 0}{t} = 0.$$

Zur Bestimmung der zweiten partiellen Ableitungen berechnen wir

$$\frac{\partial f}{\partial x}(0,y) = \begin{cases} \frac{-y^5}{y^4}, & \text{falls } y \neq 0 \\ 0, & \text{falls } y = 0 \end{cases} = -y, \quad y \in \mathbb{R}.$$

Damit erkennen wir, dass

$$\frac{\partial^2 f}{\partial y \partial x}(0,0) = -1.$$

Wenn wir die Rollen von x und y vertauschen, so wird alles mit -1 multipliziert. Dies erkennt man an der Definition von f . Daher berechnen wir ohne weitere Mühe

$$\begin{aligned} \frac{\partial f}{\partial y}(x,y) &= \frac{x^5 - 4x^3y^2 - xy^4}{(x^2 + y^2)^2}, \\ \frac{\partial f}{\partial y}(0,0) &= 0, \\ \frac{\partial f}{\partial y}(x,0) &= x, \quad x \in \mathbb{R}, \end{aligned}$$

und

$$\frac{\partial^2 f}{\partial x \partial y}(0,0) = 1.$$

Schließlich sieht man

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

Die Funktion ist f ist zweimal partiell differenzierbar. Bemerkenswert ist hier vor allem die Tatsache

$$\frac{\partial^2 f}{\partial y \partial x}(0,0) = -1 \neq 1 = \frac{\partial^2 f}{\partial x \partial y}(0,0).$$

Das eben beobachtete Phänomen tritt nicht mehr auf, wenn f stärkere Anforderungen an die Differenzierbarkeit erfüllt:

5.3.4 Satz (Schwarz⁶). *Es seien $B \subset \mathbb{R}^n$ eine offene Teilmenge, $f: B \rightarrow \mathbb{R}$ eine zweimal partiell differenzierbare Funktion, $a \in B$ und $i, j \in \{1, \dots, n\}$. Wenn die partiellen Ableitungen $\partial^2 f / \partial x_i \partial x_j$ und $\partial^2 f / \partial x_j \partial x_i$ beide stetig in a sind, dann folgt*

$$\frac{\partial^2 f}{\partial x_i \partial x_j}(a) = \frac{\partial^2 f}{\partial x_j \partial x_i}(a).$$

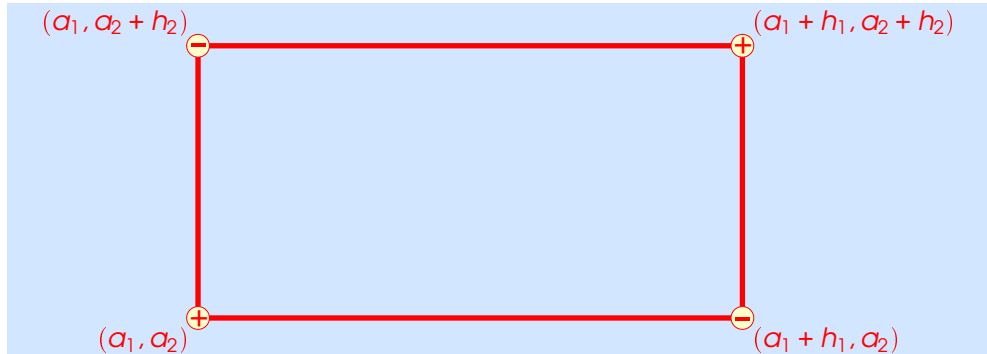
Beweis. Der Beweis ist abermals eine geschickte Anwendung des Mittelwertsatzes der Differentialrechnung ((12), Folgerung 4.3.5). Man sieht leicht ein, dass man ohne Beschränkung der Allgemeinheit $n = 2$, $i = 1$ und $j = 2$ voraussetzen darf. Wir schreiben $a = (a_1, a_2)$. Weiter können wir ein $\varepsilon > 0$ so wählen, dass

$$[a_1 - \varepsilon, a_1 + \varepsilon] \times [a_2 - \varepsilon, a_2 + \varepsilon] \subset B.$$

⁶Hermann Amandus Schwarz (1843 - 1921), deutscher Mathematiker.

Es seien $0 < h_1 < \varepsilon$ und $0 < h_2 < \varepsilon$. Wir setzen

$$F(h_1, h_2) := f(a_1 + h_1, a_2 + h_2) - f(a_1 + h_1, a_2) - f(a_1, a_2 + h_2) + f(a_1, a_2).$$



Wir führen noch die folgenden Funktionen ein:

$$\begin{aligned} \varphi: [a_1, a_1 + h_1] &\longrightarrow \mathbb{R} \\ x_1 &\longmapsto f(x_1, a_2 + h_2) - f(x_1, a_2) \\ \psi: [a_2, a_2 + h_2] &\longrightarrow \mathbb{R} \\ x_2 &\longmapsto f(a_1 + h_1, x_2) - f(a_1, x_2). \end{aligned}$$

(Die Funktionen φ und ψ sind die Zähler der Differenzenquotienten in die beiden Koordinatenrichtungen.) Die Annahme über die partielle Differenzierbarkeit von f impliziert, dass φ und ψ differenzierbare Funktionen sind. Nach dem bereits erwähnten Mittelwertsatz gibt es Zahlen $b \in [a_1, a_1 + h_1]$ und $c \in [a_2, a_2 + h_2]$ mit

$$\varphi'(b) \cdot h_1 = \varphi(a_1 + h_1) - \varphi(a_1) = F(h_1, h_2) \quad (5.5)$$

$$\psi'(c) \cdot h_2 = \psi(a_2 + h_2) - \psi(a_2) = F(h_1, h_2). \quad (5.6)$$

(Diese Ableitungen approximieren die beiden fraglichen partiellen Ableitungen in a . Dies müssen wir jetzt auf präzise Weise zum Ausdruck bringen.) Die Annahme über die zweimalige partielle Differenzierbarkeit von f zeigt, dass auch die Funktionen

$$\begin{aligned} \sigma: [a_1, a_1 + h_1] &\longrightarrow \mathbb{R} \\ x_1 &\longmapsto \frac{\partial f}{\partial x_2}(x_1, c) \\ \tau: [a_2, a_2 + h_2] &\longrightarrow \mathbb{R} \\ x_2 &\longmapsto \frac{\partial f}{\partial x_1}(b, x_2) \end{aligned}$$

differenzierbar sind. Es gibt somit Werte $d \in [a_1, a_1 + h_1]$ und $e \in [a_2, a_2 + h_2]$ mit

$$\begin{aligned} \sigma(a_1 + h_1) - \sigma(a_1) &= \sigma'(d) \cdot h_1 \\ \tau(a_2 + h_2) - \tau(a_2) &= \tau'(e) \cdot h_2. \end{aligned}$$

Jetzt gilt

$$\begin{aligned} \varphi'(b) &= \tau(a_2 + h_2) - \tau(a_2) = \tau'(e) \cdot h_2 = \frac{\partial^2 f}{\partial x_2 \partial x_1}(b, e) \cdot h_2 \\ \psi'(c) &= \sigma(a_1 + h_1) - \sigma(a_1) = \sigma'(d) \cdot h_1 = \frac{\partial^2 f}{\partial x_1 \partial x_2}(d, c) \cdot h_1. \end{aligned}$$

In Verbindung mit (5.5) und (5.6) erhalten wir die Gleichung

$$\frac{\partial^2 f}{\partial x_1 \partial x_2}(d, c) \cdot h_1 \cdot h_2 = F(h_1, h_2) = \frac{\partial^2 f}{\partial x_2 \partial x_1}(b, e) \cdot h_1 \cdot h_2,$$

also

$$\frac{\partial^2 f}{\partial x_1 \partial x_2}(d, c) = \frac{\partial^2 f}{\partial x_2 \partial x_1}(b, e).$$

Wenn wir nun h_1 und h_2 gegen Null streben lassen, dann streben die Zahlen b und d bzw. c und e gegen a_1 bzw. a_2 . Die gewünschte Gleichung

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

folgt somit aus der Stetigkeit der beiden partiellen Ableitungen in a . □

Kapitel 6

Totale Differenzierbarkeit

6.1 Differenzierbarkeit

Jetzt verallgemeinern wir den Begriff der Differenzierbarkeit für Funktionen in einer Variablen auf Abbildungen $f: B \rightarrow \mathbb{R}^n$, $B \subset \mathbb{R}^m$ offen, und diskutieren einige Eigenschaften, darunter den Bezug zur partiellen Differenzierbarkeit, sowie Beispiele.

6.1.1 Definition. Es seien $B \subset \mathbb{R}^m$ eine offene Teilmenge und $f: B \rightarrow \mathbb{R}^n$ eine Abbildung.

a) Sei $a \in B$. Die Abbildung f ist in a (*total*) differenzierbar, falls es eine lineare Abbildung $L: \mathbb{R}^m \rightarrow \mathbb{R}^n$ und eine Abbildung $r: B \rightarrow \mathbb{R}^n$ gibt, so dass gilt:

- $f(x) - f(a) = L(x - a) + r(x)$, $x \in B$,
- $\lim_{x \rightarrow a} \frac{r(x)}{\|x - a\|} = 0$.

Wir nennen L die *Ableitung*¹ von f in a und schreiben

$$Df(a) := L.$$

b) Die Abbildung f heißt (*total*) differenzierbar, wenn sie in jedem Punkt $a \in B$ differenzierbar ist.

6.1.2 Bemerkung. i) Es seien $L_1, L_2: \mathbb{R}^m \rightarrow \mathbb{R}^n$ zwei lineare Abbildungen und $r_1, r_2: B \rightarrow \mathbb{R}^n$ Abbildungen, so dass

$$f(x) = f(a) + L_i(x - a) + r_i(x), \quad x \in B,$$

und

$$\lim_{x \rightarrow a} \frac{r_i(x)}{\|x - a\|} = 0, \quad i = 1, 2.$$

Es folgt offenbar

$$\lim_{x \rightarrow a} \frac{L_1(x - a) - L_2(x - a)}{\|x - a\|} = 0.$$

Zu einem Vektor $v \in \mathbb{R}^m \setminus \{0\}$ bilden wir die Folge $(x_n)_{n \geq 1}$ mit

$$x_n := a + \frac{1}{n} \cdot v, \quad n \geq 1.$$

¹Die Eindeutigkeit der linearen Abbildung L werden wir gleich besprechen.

Sie konvergiert gegen a . Deshalb gilt

$$\begin{aligned} 0 &= \lim_{n \rightarrow \infty} \frac{L_1(x_n - a) - L_2(x_n - a)}{\|x_n - a\|} = \lim_{n \rightarrow \infty} \frac{L_1\left(\frac{1}{n} \cdot v\right) - L_2\left(\frac{1}{n} \cdot v\right)}{\left\|\frac{1}{n} \cdot v\right\|} \\ &= \lim_{n \rightarrow \infty} \frac{\frac{1}{n}}{\frac{1}{n}} \cdot \frac{L_1(v) - L_2(v)}{\|v\|} = \frac{1}{\|v\|} \cdot (L_1(v) - L_2(v)). \end{aligned}$$

Wir schließen $L_1(v) = L_2(v)$. Dies gilt für $v \in \mathbb{R}^m \setminus \{0\}$ und für 0 sowieso. Daher haben wir $L_1 = L_2$. In Bemerkung 6.1.4 wird die Eindeutigkeit noch einmal ganz explizit gezeigt.

ii) Zu der Abbildung r in Definition 6.1.1 existiert eine Abbildung $\rho: B \rightarrow \mathbb{R}^n$ mit

$$r(x) = \rho(x) \cdot \|x - a\|, \quad x \in B,$$

und

$$\lim_{x \rightarrow a} \rho(x) = 0.$$

iii) Die Abbildung $f: B \rightarrow \mathbb{R}^n$ ist genau dann in $a \in B$ differenzierbar, wenn es eine lineare Abbildung $L: \mathbb{R}^m \rightarrow \mathbb{R}^n$ gibt, so dass

$$\lim_{x \rightarrow a} \frac{f(x) - f(a) - L(x - a)}{\|x - a\|} = 0$$

gilt. (Übung.)

iv) Es seien $f = (f_1, \dots, f_n): B \rightarrow \mathbb{R}^n$ eine Abbildung und $a \in B$. Wir nehmen zunächst an, dass f in a differenzierbar ist. Wir schreiben $L = (L_1, \dots, L_n)$ und $r = (r_1, \dots, r_n)$. Dann gilt:

- $L_i: \mathbb{R}^m \rightarrow \mathbb{R}$ ist linear, $i = 1, \dots, n$,
- $f_i(x) - f_i(a) = L_i(x - a) + r_i(x)$, $x \in B$, $i = 1, \dots, n$,
- $\lim_{x \rightarrow a} \frac{r_i(x)}{\|x - a\|} = 0$, $i = 1, \dots, n$. (Dazu benutze man Satz 2.1.3.)

Damit erkennen wir, dass die Funktion f_i in a differenzierbar ist, $i = 1, \dots, n$.

Jetzt setzen wir voraus, dass f_i für $i = 1, \dots, n$ in a differenzierbar ist. Dann gibt es lineare Abbildungen $L_i: \mathbb{R}^m \rightarrow \mathbb{R}$ und Funktionen $r_i: B \rightarrow \mathbb{R}$, $i = 1, \dots, n$, mit

- $f_i(x) - f_i(a) = L_i(x - a) + r_i(x)$, $i = 1, \dots, n$,
- $\lim_{x \rightarrow a} \frac{r_i(x)}{\|x - a\|} = 0$, $i = 1, \dots, n$.

Wir setzen $L := (L_1, \dots, L_n): \mathbb{R}^m \rightarrow \mathbb{R}^n$ und $r = (r_1, \dots, r_n): B \rightarrow \mathbb{R}^n$. Es gilt:

- L ist linear,
- $f(x) - f(a) = L(x - a) + r(x)$, $x \in B$,
- $\lim_{x \rightarrow a} \frac{r(x)}{\|x - a\|} = 0$. (Man benutzt wieder Satz 2.1.3.)

Damit erkennen wir, dass f in a differenzierbar ist.

Eine Abbildung $f = (f_1, \dots, f_n): B \rightarrow \mathbb{R}^n$ ist also genau dann in einem Punkt $a \in B$ differenzierbar, wenn die Komponentenfunktionen f_1, \dots, f_n in a differenzierbar sind.

v) Für $m = 1$ und $n \geq 1$ erhalten wir den Differenzierbarkeitsbegriff aus (12), Definition 4.1.1, und Definition 4.3.1 zurück. (Für $n = 1$ folgt das aus (12), Bemerkung 4.1.2, und für $n > 1$ aus Teil iv) dieser Bemerkung.)

vi) Es sei $L = (L_1, \dots, L_n): \mathbb{R}^m \rightarrow \mathbb{R}^n$ eine lineare Abbildung. Dann existiert eine $(n \times m)$ -Matrix

$$A = (a_{ij})_{\substack{i=1,\dots,n \\ j=1,\dots,m}} = \begin{pmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nm} \end{pmatrix},$$

so dass

$$\forall x = (x_1, \dots, x_m) \in \mathbb{R}^m: \begin{pmatrix} L_1(x) \\ \vdots \\ L_n(x) \end{pmatrix} = A \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix}.$$

Für eine differenzierbare Abbildung $f: B \rightarrow \mathbb{R}^n$, $B \subset \mathbb{R}^m$ offen, können wir die *Ableitung von f als Abbildung*

$$\begin{aligned} Df: \mathbb{R}^m &\rightarrow \mathbb{R}^{n \cdot m} \\ x &\mapsto Df(x) = (a_{ij}(x))_{\substack{i=1,\dots,n \\ j=1,\dots,m}} \end{aligned}$$

erklären. Man beachte, dass im Fall $m > 1$ der Zielraum der Ableitung ein anderer ist als der der ursprünglichen Abbildung, nämlich $\mathbb{R}^{n \cdot m}$ anstatt \mathbb{R}^n .

Das folgende Resultat verallgemeinert (12), Lemma 4.1.3, und stellt den Bezug zum Begriff der partiellen Differenzierbarkeit her.

6.1.3 Satz. *Es seien $B \subset \mathbb{R}^m$ eine offene Teilmenge, $f = (f_1, \dots, f_n): B \rightarrow \mathbb{R}^n$ eine Abbildung und $a \in B$. Die Abbildung f sei in a differenzierbar. Dann gilt:*

- i) *Die Abbildung f ist in a stetig.*
- ii) *Die Funktionen f_1, \dots, f_n sind in a partiell differenzierbar.*

Beweis. i) Wir benutzen die Darstellung

$$f(x) = f(a) + L(x - a) + r(x), \quad x \in B,$$

mit einer linearen Abbildung $L: \mathbb{R}^m \rightarrow \mathbb{R}^n$ und einer Abbildung $r: B \rightarrow \mathbb{R}^n$, so dass $\lim_{x \rightarrow a} r(x)/\|x - a\| = 0$. Da L stetig ist (Bemerkung 3.7.3), gilt

$$\lim_{x \rightarrow a} L(x - a) = L(a - a) = L(0) = 0.$$

Weiter gilt

$$\lim_{x \rightarrow a} r(x) = \lim_{x \rightarrow a} \|x - a\| \cdot \frac{r(x)}{\|x - a\|} = 0 \cdot 0 = 0.$$

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

ii) Nach Bemerkung 6.1.2, iv), ist die Funktion f_i in a differenzierbar, so dass wir eine lineare Abbildung $L_i: \mathbb{R}^m \rightarrow \mathbb{R}$ und eine Funktion $r_i: B \rightarrow \mathbb{R}$ mit $\lim_{x \rightarrow a} r_i(x)/\|x - a\| = 0$ und

$$f_i(x) = f_i(a) + L_i(x - a) + r_i(x), \quad x \in B,$$

finden können, $i = 1, \dots, n$. Für den Standardbasisvektor $e_j \in \mathbb{R}^m$ und $t \in \mathbb{R}$, so dass $a + t \cdot e_j \in B$, berechnen wir somit

$$f_i(a + t \cdot e_j) = f_i(a) + L_i(t \cdot e_j) + r_i(a + t \cdot e_j) = f_i(a) + t \cdot L_i(e_j) + r_i(a + t \cdot e_j).$$

Damit sehen wir sofort ein, dass

$$\frac{\partial f_i}{\partial x_j}(a) = \lim_{t \rightarrow 0} \frac{f_i(a + t \cdot e_j) - f_i(a)}{t} = L_i(e_j) + \lim_{t \rightarrow 0} \frac{r_i(a + t \cdot e_j)}{t} = L_i(e_j)$$

gilt, $j = 1, \dots, m, i = 1, \dots, n$. □

6.1.4 Bemerkung. i) Der Beweis zeigt, dass die Matrix $A = (a_{ij})_{i=1, \dots, n; j=1, \dots, m}$ in Bemerkung 6.1.2, vi), durch

$$a_{ij} = \frac{\partial f_i}{\partial x_j}(a) = L_i(e_j)$$

gegeben ist, $i = 1, \dots, n, j = 1, \dots, m$.

Die Matrix

$$J_f(a) := \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(a) & \cdots & \frac{\partial f_1}{\partial x_m}(a) \\ \vdots & & \vdots \\ \frac{\partial f_n}{\partial x_1}(a) & \cdots & \frac{\partial f_n}{\partial x_m}(a) \end{pmatrix}$$

heißt *Jacobi-Matrix*² von f in a .

ii) Für $n = 1$ gilt

$$J_f(a) = \left(\frac{\partial f}{\partial x_1}(a), \dots, \frac{\partial f}{\partial x_m}(a) \right) = \text{Grad}f(a),$$

und wir können

$$f(x) = f(a) + \langle \text{Grad}f(a), x - a \rangle + r(x), \quad x \in B,$$

mit einer Funktion $r: B \rightarrow \mathbb{R}$, für die

$$\lim_{x \rightarrow a} \frac{r(x)}{\|x - a\|}$$

gilt, schreiben.

iii) Der Beweis von Satz 6.1.3, ii), zeigt allgemeiner, dass f_i an der Stelle a in jede Richtung $v \in \mathbb{R}^m \setminus \{0\}$ differenzierbar ist, und

$$D_v f_i(a) = L_i(v)$$

gilt, $i = 1, \dots, n$. Für $n = 1$ ergibt dies die Formel

$$D_v f(a) = \langle \text{Grad}f(a), v \rangle,$$

deren Gültigkeit wir in Satz 5.2.3 für eine stetig partiell differenzierbare Funktion nachgewiesen hatten.

iv) Für eine offene Menge $B \subset \mathbb{R}^n$ und eine Funktion $f: B \rightarrow \mathbb{R}$ ist der Graph von f die Menge

$$\Gamma_f := \left\{ (x_1, \dots, x_n, y) \in B \times \mathbb{R} \mid y = f(x_1, \dots, x_n) \right\} \subset B \times \mathbb{R}.$$

²Carl Gustav Jacob Jacobi (1804 - 1851), deutscher Mathematiker.

Wenn f im Punkt a differenzierbar ist, dann definieren wir die *Tangentialebene von f in $(a, f(a))$* als

$$T_f(a) := \left\{ (x_1, \dots, x_n, y) \in \mathbb{R}^{n+1} \mid y = f(a) + \langle \text{Grad}f(a), (x_1 - a_1, \dots, x_n - a_n) \rangle \right\}.$$

Mit dem linearen Teilraum

$$E_f(a) := \left\{ (x_1, \dots, x_n, y) \in \mathbb{R}^{n+1} \mid y = \langle \text{Grad}f(a), (x_1, \dots, x_n) \rangle \right\}$$

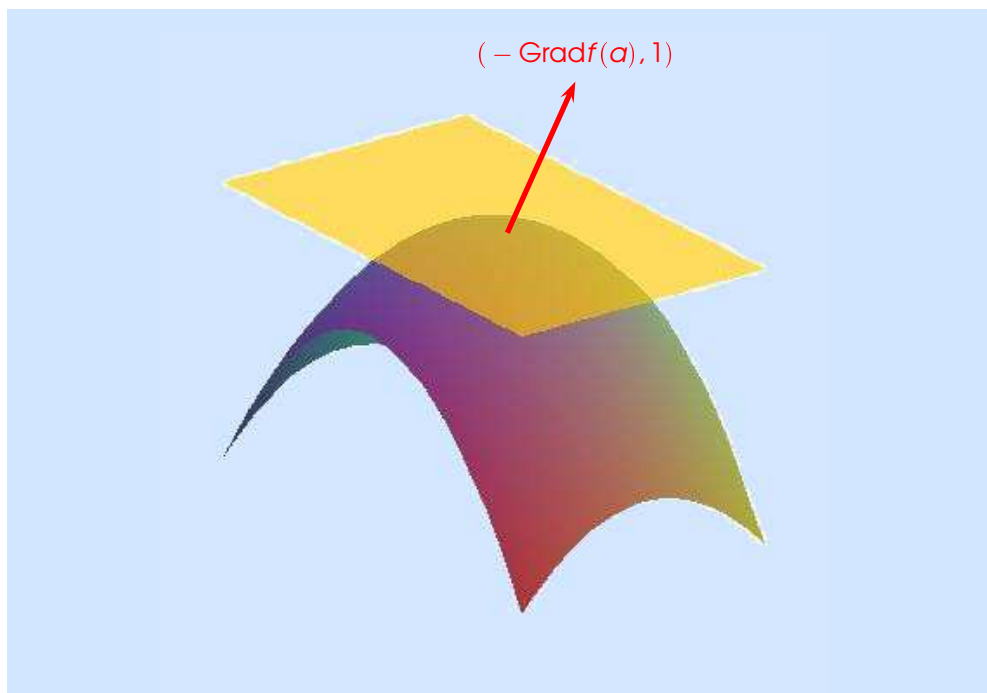
gilt

$$T_f(a) = (a, f(a)) + E_f(a).$$

Man sieht leicht, dass der Vektor

$$(-\text{Grad}f(a), 1)$$

auf der Tangentialebene $T_f(a)$ senkrecht steht (vgl. Seite 66).



6.1.5 Beispiele. i) Zu einem Vektor $c \in \mathbb{R}^n$ definieren wir die *konstante Abbildung*

$$\begin{aligned} f: \mathbb{R}^m &\longrightarrow \mathbb{R}^n \\ x &\longmapsto c. \end{aligned}$$

Für Punkte $a, x \in \mathbb{R}^m$ gilt

$$f(x) = f(a).$$

Die Abbildung f ist in jedem Punkt $a \in \mathbb{R}^m$ differenzierbar, und man hat $D_f(a) = 0$.

ii) Es sei $f: \mathbb{R}^m \longrightarrow \mathbb{R}^n$ eine **lineare** Abbildung. Dann gilt

$$f(x) = f(a) + f(x - a), \quad a, x \in \mathbb{R}^m.$$

Die Abbildung f ist in jedem Punkt $a \in \mathbb{R}^m$ differenzierbar, und die Ableitung in a ist $D_f(a) = f$.

iii) Es sei $(a_{ij})_{i,j=1,\dots,n}$ eine **symmetrische** $(n \times n)$ -Matrix, d.h.

$$A^t = A \quad (\iff a_{ij} = a_{ji}, \quad i, j = 1, \dots, n).$$

Wir definieren die Abbildung

$$f: \mathbb{R}^n \longrightarrow \mathbb{R}$$

$$x = (x_1, \dots, x_n) \longmapsto (x_1, \dots, x_n) \cdot A \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

Für $a, x \in \mathbb{R}^n$ finden wir

$$\begin{aligned} f(x) &= (a + (x - a)) \cdot A \cdot (a + (x - a))^t \\ &= a \cdot A \cdot a^t + a \cdot A \cdot (x - a)^t + (x - a) \cdot A \cdot a^t + (x - a) \cdot A \cdot (x - a)^t \\ &\stackrel{A=A^t}{=} a \cdot A \cdot a^t + 2 \cdot a \cdot A \cdot (x - a)^t + (x - a) \cdot A \cdot (x - a)^t. \end{aligned} \quad (6.1)$$

Cauchy-Schwarz-Ungleichung. Für zwei Vektoren $v, w \in \mathbb{R}^n$ gilt

$$\langle v, w \rangle^2 \leq \langle v, v \rangle \cdot \langle w, w \rangle.$$

Diese Ungleichung folgt wegen $|\cos(\varphi)| \leq 1$ sofort aus Gleichung (5.4). Wir geben auch noch den üblichen Beweis an, der für alle Skalarprodukte funktioniert. Die obige Ungleichung ist nur für den Fall $\langle v, w \rangle \neq 0$ interessant. (Dies impliziert insbesondere $v \neq 0$ und $w \neq 0$.) Für jede Zahl $\lambda \in \mathbb{R}$ gilt

$$0 \leq \langle v + \lambda \cdot w, v + \lambda \cdot w \rangle = \langle v, v \rangle + 2 \cdot \lambda \cdot \langle v, w \rangle + \lambda^2 \cdot \langle w, w \rangle.$$

Für

$$\lambda := -\frac{\langle v, w \rangle}{\langle w, w \rangle}$$

bekommen wir die Ungleichung

$$0 \leq \langle v, v \rangle - 2 \cdot \frac{\langle v, w \rangle^2}{\langle w, w \rangle} + \frac{\langle v, w \rangle^2}{\langle w, w \rangle}.$$

Das ist genau die Cauchy-Schwarz-Ungleichung. ✓

Jetzt gilt

$$\begin{aligned} |(x - a) \cdot A \cdot (x - a)^t| &= \langle x - a, (x - a) \cdot A \rangle \\ &\leq \|x - a\| \cdot \|(x - a) \cdot A\| \\ &\leq \|x - a\|^2 \cdot \|A\|_{\text{op}}. \end{aligned}$$

Offensichtlich haben wir

$$\lim_{x \rightarrow a} \frac{\|A\|_{\text{op}} \cdot \|x - a\|^2}{\|x - a\|} = \lim_{x \rightarrow a} \|A\|_{\text{op}} \cdot \|x - a\| = 0.$$

Aus (6.1) folgern wir nun, dass f differenzierbar ist und für $a \in \mathbb{R}^n$

$$\begin{aligned} Df(a) &= 2 \cdot \langle a \cdot A, \cdot \rangle: \mathbb{R}^n \longrightarrow \mathbb{R} \\ x &\longmapsto 2 \cdot \langle a \cdot A, x \rangle \end{aligned}$$

gilt.

6.2 Stetig differenzierbare Funktionen

Dieser Abschnitt enthält ein wichtiges hinreichendes Kriterium für die Differenzierbarkeit einer Funktion.

6.2.1 Satz. Es seien $B \subset \mathbb{R}^n$ eine offene Menge, $f: B \rightarrow \mathbb{R}$ eine **partiell differenzierbare** Funktion und $a \in B$, so dass die partiellen Ableitungen

$$\frac{\partial f}{\partial x_i}: B \rightarrow \mathbb{R}$$

in a **stetig** sind, $i = 1, \dots, n$. Dann ist f in a differenzierbar.

Beweis. Im Beweis von Satz 5.2.3 haben wir gesehen, dass es eine Funktion $\delta: B \rightarrow \mathbb{R}$ mit $\lim_{x \rightarrow a} \delta(x) = 0$ gibt, so dass

$$f(x) - f(a) = \langle \text{Grad}f(a), x - a \rangle + \langle \delta(x), x - a \rangle, \quad x \in B.$$

Nach der Cauchy–Schwarz-Ungleichung gilt

$$|\langle \delta(x), x - a \rangle| \leq \|\delta(x)\| \cdot \|x - a\|,$$

so dass

$$\lim_{x \rightarrow a} \frac{\langle \delta(x), x - a \rangle}{\|x - a\|} = 0$$

folgt. Wir schließen daraus, dass f in a differenzierbar ist. \square

Wegen dieses Satzes führen wir folgende Terminologie ein:

6.2.2 Definition. Die Funktion f ist in a **stetig differenzierbar**, wenn sie dort stetig partiell differenzierbar ist. Analog bezeichnen wir eine k -mal stetig partiell differenzierbare Funktion als **k -mal stetig differenzierbare Funktion**.

6.2.3 Bemerkung. In Satz 6.2.1 und 6.1.3, ii), haben wir die Implikationen

stetig partiell differenzierbar \Rightarrow differenzierbar \Rightarrow partiell differenzierbar nachgewiesen. Die beiden Umkehrungen sind i.A. falsch: In Beispiel 5.1.8, ii), sind wir einer partiell differenzierbaren Funktion begegnet, die nicht stetig und daher auch nicht differenzierbar ist. In (12), Beispiel 4.9.2, wurde eine differenzierbare Funktion in einer Veränderlichen vorgestellt, deren Ableitung nicht stetig ist.

6.2.4 Folgerung. Unter den Voraussetzungen von Satz 6.2.1 ist f in a stetig.

6.3 Die Kettenregel

Die Verknüpfung differenzierbarer Abbildungen ist wieder differenzierbar, und die Ableitung der Verknüpfung kann aus den Ableitungen der Faktoren berechnet werden.

6.3.1 Satz (Kettenregel). Es seien $B \subset \mathbb{R}^m$ und $D \subset \mathbb{R}^n$ offene Teilmengen und $f: B \rightarrow \mathbb{R}^n$ und $g: D \rightarrow \mathbb{R}^p$ Abbildungen, so dass $f(B) \subset D$. Die Abbildung f sei in $a \in B$ und die Abbildung g in $b := f(a)$ differenzierbar. Dann ist die Verknüpfung $g \circ f: B \rightarrow \mathbb{R}^p$ in a differenzierbar, und die Ableitung wird durch

$$D(g \circ f)(a) = Dg(b) \circ Df(a)$$

berechnet.

Beweis. Die Strategie ist dieselbe wie im Spezialfall von Satz 5.2.3. Wir können

$$\begin{aligned} f(x) &= f(a) + Df(a)(x - a) + \rho(x) \cdot \|x - a\|, \quad x \in B, \\ g(y) &= g(b) + Dg(b)(y - b) + \sigma(y) \cdot \|y - b\|, \quad y \in D, \end{aligned}$$

mit Abbildungen $\rho: B \rightarrow \mathbb{R}^p$ und $\sigma: D \rightarrow \mathbb{R}^p$, die

$$\lim_{x \rightarrow a} \rho(x) = 0 \quad \text{und} \quad \lim_{y \rightarrow b} \sigma(y) = 0$$

erfüllen, schreiben (Bemerkung 6.1.2, ii). Folglich erhalten wir

$$\begin{aligned} g(f(x)) &= g(f(a)) + Dg(f(a))(f(x) - f(a)) + \sigma(f(x)) \cdot \|f(x) - f(a)\| \\ &= g(f(a)) + Dg(f(a))(Df(a)(x - a)) + \underbrace{\|x - a\| \cdot Dg(f(a))(\rho(x)) + \sigma(f(x)) \cdot \|f(x) - f(a)\|}_{=: t(x)}. \end{aligned} \quad (6.2)$$

Wir schätzen

$$\|Dg(f(a))(\rho(x))\| \leq \|Dg(f(a))\|_{\text{op}} \cdot \|\rho(x)\|$$

ab. Damit berechnen wir

$$\lim_{x \rightarrow a} Dg(f(a))(\rho(x)) = 0.$$

Weiter sehen wir

$$\begin{aligned} \|f(x) - f(a)\| &= \|Df(a)(x - a) + \rho(x) \cdot \|x - a\|\| \\ &\leq \|Df(a)(x - a)\| + \|\rho(x)\| \cdot \|x - a\|. \end{aligned}$$

Dabei haben wir wieder

$$\|Df(a)(x - a)\| \leq \|Df(a)\|_{\text{op}} \cdot \|x - a\|.$$

Wir leiten die Abschätzung

$$\frac{\|f(x) - f(a)\|}{\|x - a\|} \leq \|\rho(x)\| + \|Df(a)\|_{\text{op}}$$

ab. Man beachte, dass aus der Stetigkeit von f in a (Satz 6.1.3, ii)

$$\lim_{x \rightarrow a} \sigma(f(x)) = \lim_{y \rightarrow b} \sigma(y) = 0$$

folgt. Zusammengenommen erhalten wir

$$\lim_{x \rightarrow a} \frac{t(x)}{\|x - a\|} = 0.$$

Aus (6.2) folgern wir die Behauptung. □

6.3.2 Bemerkung. Für die Jacobi-Matrizen folgt die Gleichung

$$J_{g \circ f}(a) = J_g(f(a)) \cdot J_f(a) \quad (\text{Matrixprodukt}).$$

6.3.3 Beispiele. i) Es seien Abbildungen $\gamma: \mathbb{R} \rightarrow \mathbb{R}^n$ und $f: \mathbb{R}^n \rightarrow \mathbb{R}$ gegeben. Die Abbildung γ sei in t differenzierbar und die Funktion f in $\gamma(t)$. Dann gilt für die Funktion $f \circ \gamma: \mathbb{R} \rightarrow \mathbb{R}$:

$$(f \circ \gamma)'(t) = \langle \text{Grad} f(\gamma(t)), \gamma'(t) \rangle.$$

Für eine stetig differenzierbare Funktion sind wir dieser Behauptung bereits im Beweis von Satz 5.2.3 begegnet.

ii) Der Gradient der Funktion

$$\begin{aligned} f: \mathbb{R}^2 &\longrightarrow \mathbb{R} \\ (x, y) &\longmapsto \exp(xy - x^2 + y^2) \end{aligned}$$

ist durch

$$\text{Grad}f(x, y) = ((y - 2x) \cdot \exp(xy - x^2 + y^2), (x + 2y) \cdot \exp(xy - x^2 + y^2))$$

gegeben. Auf der anderen Seite können wir diese Funktion als Verknüpfung $f = h \circ g$ mit

$$\begin{aligned} g: \mathbb{R}^2 &\longrightarrow \mathbb{R}^2 \\ (x, y) &\longmapsto (xy, x^2 - y^2) \\ h: \mathbb{R}^2 &\longrightarrow \mathbb{R} \\ (v, w) &\longmapsto \exp(v - w) \end{aligned}$$

darstellen. Als Jacobi-Matrizen erhalten wir

$$\begin{aligned} J_g(x, y) &= \begin{pmatrix} \frac{\partial g_1}{\partial x}(x, y) & \frac{\partial g_1}{\partial y}(x, y) \\ \frac{\partial g_2}{\partial x}(x, y) & \frac{\partial g_2}{\partial y}(x, y) \end{pmatrix} = \begin{pmatrix} y & x \\ 2x & -2y \end{pmatrix} \\ J_h(v, w) &= (\exp(v - w), -\exp(v - w)). \end{aligned}$$

Schließlich überprüft man:

$$\begin{aligned} &J_h(xy, x^2 - y^2) \cdot J_g(x, y) = \\ &= (\exp(xy - x^2 + y^2), -\exp(xy - x^2 + y^2)) \cdot \begin{pmatrix} y & x \\ 2x & -2y \end{pmatrix} = \\ &= ((y - 2x) \cdot \exp(xy - x^2 + y^2), (x + 2y) \cdot \exp(xy - x^2 + y^2)) = \\ &= J_f(x, y). \end{aligned}$$

Kapitel 7

Extremwerte differenzierbarer Funktionen

7.1 Extrema und Sattelpunkte

Wie in der Theorie der reellwertigen Funktionen einer Veränderlichen ((12), Kapitel 4) hilft die Differentialrechnung beim Auffinden von Extremwerten. In diesem Abschnitt führen wir die entsprechenden Begriffe ein und formulieren ein notwendiges Kriterium für lokale Extremstellen. Es verallgemeinert (12), Satz 4.3.2.

7.1.1 Definition. Es seien $M \subset \mathbb{R}^n$ eine Teilmenge, $f: M \rightarrow \mathbb{R}$ eine Funktion und $a \in M$.

a) Wir sagen, dass f in a ein *lokales Minimum* bzw. *Maximum* annimmt, wenn es eine offene Umgebung $U \subset \mathbb{R}^n$ von a gibt, so dass

$$\forall x \in U \cap M: \quad f(a) \leq f(x) \quad \text{bzw.} \quad f(a) \geq f(x).$$

b) Die Funktion f hat in a ein *lokales Extremum*, wenn sie in a ein lokales Minimum oder Maximum hat.

c) Können wir in a) $U = \mathbb{R}^n$ nehmen, dann nimmt f in a ein *globales Minimum* bzw. *Maximum* an.

d) Ein *globales Extremum* von f ist ein globales Minimum oder Maximum.

Es gilt das zu (12), Satz 4.3.2, analoge notwendige Kriterium für das Vorliegen eines lokalen Extremums:

7.1.2 Satz. Es seien $B \subset \mathbb{R}^n$ eine offene Teilmenge, $f: B \rightarrow \mathbb{R}$ eine differenzierbare Funktion und $a \in B$. Wenn f in a ein lokales Extremum annimmt, dann gilt

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

Beweis. Für $n = 1$ besagt die Behauptung $f'(a) = 0$. Wie zuvor bemerkt ist diese Aussage Gegenstand von (12), Satz 4.3.2.

Nun sei $n \geq 1$ beliebig. Es gibt ein $\varepsilon > 0$, so dass $a + t \cdot e_i \in B$, $t \in (-\varepsilon, \varepsilon)$, $i = 1, \dots, n$. Dabei ist e_i wieder der i -te Standardbasisvektor von \mathbb{R}^n , $i = 1, \dots, n$. Die Funktion

$$\begin{aligned} g_i: (-\varepsilon, \varepsilon) &\rightarrow \mathbb{R} \\ t &\mapsto f(a + t \cdot e_i) \end{aligned}$$

hat in 0 ein lokales Extremum. Es folgt

$$\frac{\partial f}{\partial x_i}(a) = g'_i(0) = 0, \quad i = 1, \dots, n.$$

Da

$$\text{Grad}f(a) = \left(\frac{\partial f}{\partial x_1}(a), \dots, \frac{\partial f}{\partial x_n}(a) \right),$$

ist die Aussage hiermit gezeigt. \square

7.1.3 Definition. Es seien $B \subset \mathbb{R}^n$ eine **offene** Teilmenge, $f: B \rightarrow \mathbb{R}$ eine differenzierbare Funktion und $a \in B$.

a) Der Punkt $a \in B$ ist ein *kritischer Punkt* von f , wenn

$$\text{Grad}f(a) = 0$$

gilt.

b) Der Punkt a ist ein *Sattelpunkt* von f , wenn er ein kritischer Punkt von f ist und in **jeder** Umgebung U von a Punkte b, c mit

$$f(b) < f(a) < f(c)$$

zu finden sind.

7.1.4 Bemerkung. Jeder kritische Punkt von f ist entweder ein Sattelpunkt oder ein lokaler Extrempunkt von f .

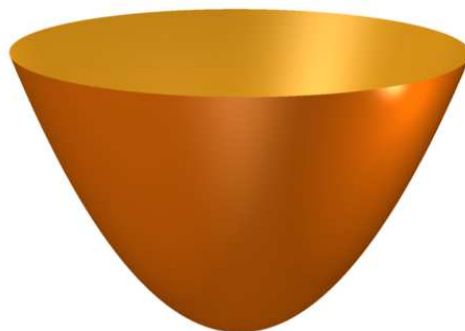
7.1.5 Beispiele. i) Für die Funktion

$$\begin{aligned} f: \mathbb{R}^2 &\longrightarrow \mathbb{R} \\ (x, y) &\longmapsto x^2 + y^2 \end{aligned}$$

gilt

$$\frac{\partial f}{\partial x}(x, y) = 2x, \quad \frac{\partial f}{\partial y}(x, y) = 2y, \quad \text{Grad}f(x, y) = (2x, 2y), \quad (x, y) \in \mathbb{R}^2.$$

Damit ist $a = (0, 0)$ der einzige kritische Punkt von f . In a nimmt f ein **globales Minimum** an.



ii) Es sei

$$\begin{aligned} f: \mathbb{R}^2 &\longrightarrow \mathbb{R} \\ (x, y) &\longmapsto 1 - 3x^2y^4. \end{aligned}$$

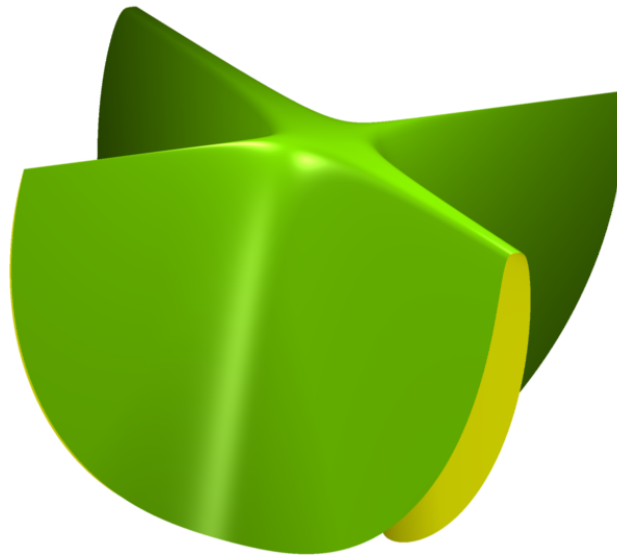
Für $(x, y) \in \mathbb{R}^2$ berechnen wir

$$\frac{\partial f}{\partial x}(x, y) = -6xy^4, \quad \frac{\partial f}{\partial y}(x, y) = -12x^2y^3, \quad \text{Grad}f(x, y) = -6xy \cdot (y^3, 2xy^2).$$

Die Menge der kritischen Punkte ist

$$K := \{ (x, y) \in \mathbb{R}^2 \mid x \cdot y = 0 \}.$$

Kritische Punkte müssen also **nicht isoliert** liegen. In jedem kritischen Punkt $a \in K$ hat f ein globales Maximum.



iii) Für die Funktion

$$\begin{aligned} f: \mathbb{R}^2 &\longrightarrow \mathbb{R} \\ (x, y) &\longmapsto x^2 + 3y^3 \end{aligned}$$

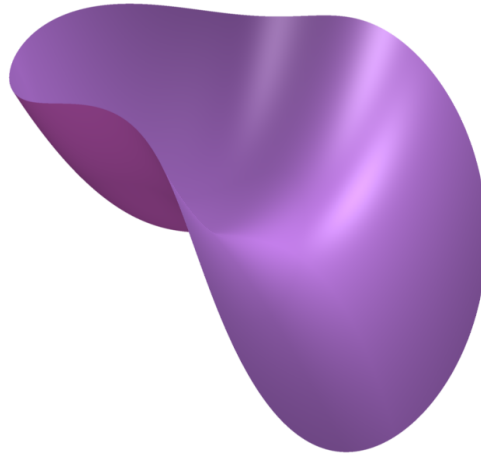
findet man

$$\text{Grad}f(x, y) = (2x, 9y^2).$$

Der Nullpunkt $a = (0, 0)$ ist der einzige kritische Punkt von f . Für jede reelle Zahl $\varepsilon > 0$ gilt

$$\begin{aligned} f(\varepsilon, 0) &= \varepsilon^2 > 0 \\ f(0, -\varepsilon) &= -3\varepsilon^3 < 0. \end{aligned}$$

Wir schließen, dass $a = (0, 0)$ ein Sattelpunkt von f ist.



7.2 Die Taylorformel

Die Taylorformeln für hinreichend oft differenzierbare Funktionen einer Veränderlichen ((12), Abschnitt 5.7) lassen sich auch in den Kontext der Funktionen mehrerer Veränderlicher übertragen. Diese Übertragung wird jetzt ausgeführt. Dabei benötigt man eine effiziente Notation, die einem eine elegante Formulierung ermöglicht.

Für ein Tupel $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{N}^n$ definieren wir

$$\begin{aligned} |\alpha| &:= \alpha_1 + \dots + \alpha_n \\ \alpha! &:= \alpha_1! \cdot \dots \cdot \alpha_n! \end{aligned}$$

Es seien $B \subset \mathbb{R}^n$ eine offene Menge und $k \geq 0$. Wir setzen

$$\mathcal{C}^k(B) := \left\{ f: B \rightarrow \mathbb{R} \mid f \text{ ist } k\text{-mal stetig differenzierbar} \right\}.$$

Wir definieren rekursiv die Differentialoperatoren

$$\begin{aligned} D_j^i: \mathcal{C}^k(B) &\longrightarrow \mathcal{C}^{k-i}(B) \\ f &\longmapsto \begin{cases} \frac{\partial}{\partial x_j} f, & \text{falls } i = 1 \\ \frac{\partial}{\partial x_j} (D_j^{i-1}(f)), & \text{falls } i > 1 \end{cases} \end{aligned}$$

$j \in \{1, \dots, n\}, k \in \mathbb{N}, i \in \{1, \dots, k\}$. Für $\alpha = (\alpha_1, \dots, \alpha_n)$ und $k \geq |\alpha|$ erhalten wir den Differentialoperator

$$D^\alpha := D_1^{\alpha_1} \circ \dots \circ D_n^{\alpha_n}: \mathcal{C}^k(B) \longrightarrow \mathcal{C}^{k-|\alpha|}(B).$$

Schließlich schreiben wir

$$x^\alpha := x_1^{\alpha_1} \cdot \dots \cdot x_n^{\alpha_n}$$

für $x = (x_1, \dots, x_n) \in \mathbb{R}^n$ und $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{N}^n$.

7.2.1 Satz. Es seien $B \subset \mathbb{R}^n$ eine offene Teilmenge, $k \geq 1$ und $f \in \mathcal{C}^k(B)$. Ferner seien $x \in B$ und $\xi \in \mathbb{R}^n$ gegeben, so dass

$$\{x + t \cdot \xi \mid 0 \leq t \leq 1\} \subset B.$$

Dann ist die Funktion

$$\begin{aligned} g: [0, 1] &\longrightarrow \mathbb{R} \\ t &\longmapsto f(x + t \cdot \xi) \end{aligned}$$

k -mal stetig differenzierbar, und es gilt

$$\frac{d^k}{dt^k} g(t) := g^{(k)}(t) = \sum_{\substack{\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{N}^n: \\ |\alpha| = k}} \frac{k!}{\alpha!} \cdot D^\alpha f(x + t \cdot \xi) \cdot \xi^\alpha, \quad t \in [0, 1].$$

Beweis. Wir beweisen die Formel mittels Induktion über k . Für $k = 1$ verwenden wir die Kettenregel (Satz 5.2.3 oder Satz 6.3.1):

$$\frac{d}{dt} g(t) = \langle \text{Grad} f(x + t \cdot \xi), \xi \rangle = \sum_{i=1}^n \frac{\partial f}{\partial x_i}(x + t \cdot \xi) \cdot \xi_i, \quad t \in [0, 1].$$

Für den Schluss $k-1 \rightarrow k$ benötigen wir noch eine Umformulierung der Aussage. Nach dem Satz von Schwarz (Satz 5.3.4 und Aufgabe A.7.3, a) dürfen wir die Reihenfolge der partiellen Ableitungen nach Belieben ändern. Daher gilt

$$\begin{aligned} &\sum_{i_1, \dots, i_k=1}^n \left(\frac{\partial}{\partial x_{i_k}} \circ \dots \circ \frac{\partial}{\partial x_{i_1}} \right) f(x + t \cdot \xi) \cdot \xi_{i_1} \cdot \dots \cdot \xi_{i_{k-1}} \cdot \xi_{i_k} \\ &= \sum_{\substack{\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{N}^n: \\ |\alpha| = k}} c_{\alpha, k} \cdot D^\alpha f(x + t \cdot \xi) \cdot \xi^\alpha \end{aligned}$$

mit¹

$$\begin{aligned} c_{\alpha, k} &= \#\left\{ (i_1, \dots, i_k) \in \{1, \dots, n\}^k \mid \#\{j \mid i_j = l\} = \alpha_l, l = 1, \dots, n \right\} \\ &= \frac{k!}{\alpha_1! \cdot \dots \cdot \alpha_n!} = \frac{k!}{\alpha!}, \quad k \in \mathbb{N}, \alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{N}^n. \end{aligned}$$

Mit dieser Formel berechnen wir:

$$\begin{aligned} g^{(k)}(t) &= \frac{d}{dt} g^{(k-1)}(t) \\ &\stackrel{IV}{=} \frac{d}{dt} \left(\sum_{i_1, \dots, i_{k-1}=1}^n \left(\frac{\partial}{\partial x_{i_{k-1}}} \circ \dots \circ \frac{\partial}{\partial x_{i_1}} \right) f(x + t \cdot \xi) \cdot \xi_{i_1} \cdot \dots \cdot \xi_{i_{k-1}} \right) \\ &\stackrel{V}{=} \sum_{i=1}^n \sum_{i_1, \dots, i_{k-1}=1}^n \left(\frac{\partial}{\partial x_i} \circ \frac{\partial}{\partial x_{i_{k-1}}} \circ \dots \circ \frac{\partial}{\partial x_{i_1}} \right) f(x + t \cdot \xi) \cdot \xi_{i_1} \cdot \dots \cdot \xi_{i_{k-1}} \cdot \xi_i \\ &= \sum_{i_1, \dots, i_k=1}^n \left(\frac{\partial}{\partial x_{i_k}} \circ \dots \circ \frac{\partial}{\partial x_{i_1}} \right) f(x + t \cdot \xi) \cdot \xi_{i_1} \cdot \dots \cdot \xi_{i_{k-1}} \cdot \xi_{i_k} \\ &= \sum_{\substack{\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{N}^n: \\ |\alpha| = k}} \frac{k!}{\alpha!} \cdot D^\alpha f(x + t \cdot \xi) \cdot \xi^\alpha. \end{aligned}$$

¹Um die folgende Formel für $c_{\alpha, k}$ zu verstehen, beachte man, dass ein Tupel in der angegebenen Menge durch eine Permutation aus dem Tupel $i_0 = (i_1^0, \dots, i_k^0)$ erhalten wird, für das $i_1^0 \leq \dots \leq i_k^0$ gilt, bei dem also zunächst α_1 -mal die 1, dann α_2 -mal die 2 usw. steht. Nun gibt es insgesamt $n!$ Permutationen, von denen $\alpha_1! \cdot \dots \cdot \alpha_n!$ das Tupel i_0 nicht ändern. (Welche sind das?)

Damit ist die Behauptung gezeigt. \square

7.2.2 Satz (Taylorformel). *Es seien $B \subset \mathbb{R}^n$ eine offene Teilmenge, $f: B \rightarrow \mathbb{R}$ eine $(k+1)$ -mal stetig differenzierbare Funktion, $x \in B$ und $\xi \in \mathbb{R}^n$, so dass*

$$\{x + t \cdot \xi \mid t \in [0, 1]\} \subset B.$$

Dann existiert eine Zahl $\vartheta \in [0, 1]$ mit

$$f(x + \xi) = \sum_{\substack{\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{N}^n: \\ |\alpha| \leq k}} \frac{D^\alpha f(x)}{\alpha!} \cdot \xi^\alpha + \sum_{\substack{\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{N}^n: \\ |\alpha| = k+1}} \frac{D^\alpha f(x + \vartheta \cdot \xi)}{\alpha!} \cdot \xi^\alpha.$$

Beweis. Der Satz ist uns für $n = 1$ bereits bekannt (s. (12), Satz 5.7.3). Für $n \geq 1$ ergibt sich die Formel durch Anwendung des gerade zitierten Satzes auf die Funktion

$$\begin{aligned} g: [0, 1] &\rightarrow \mathbb{R} \\ t &\mapsto f(x + t \cdot \xi) \end{aligned}$$

und Einsetzen der Ausdrücke aus Satz 7.2.1 für die Ableitungen von g an der Stelle $t = 1$. \square

7.2.3 Folgerung. *Es seien $B \subset \mathbb{R}^n$ eine offene Teilmenge, $f: B \rightarrow \mathbb{R}$ eine k -mal stetig differenzierbare Funktion, $a \in B$ und $\varepsilon > 0$, so dass $B(a, \varepsilon) \subset B$. Dann existiert eine Funktion $r_k: B(a, \varepsilon) \rightarrow \mathbb{R}$ mit*

$$f(x) = \sum_{\substack{\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{N}^n: \\ |\alpha| \leq k}} \frac{D^\alpha f(a)}{\alpha!} \cdot (x - a)^\alpha + r_k(x), \quad x \in B(a, \varepsilon),$$

und

$$\lim_{x \rightarrow a} \frac{r_k(x)}{\|x - a\|^k} = 0.$$

Beweis. Die Taylorformel im Fall $k - 1$ liefert die Darstellung

$$f(x) = \sum_{\substack{\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{N}^n: \\ |\alpha| \leq k}} \frac{D^\alpha f(a)}{\alpha!} \cdot (x - a)^\alpha + r_k(x), \quad x \in B(a, \varepsilon),$$

mit

$$r_k(x) = \sum_{\substack{\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{N}^n: \\ |\alpha| = k}} \frac{1}{\alpha!} \cdot (D^\alpha f(a + \vartheta_x \cdot (x - a)) - D^\alpha f(a)) \cdot (x - a)^\alpha.$$

Aus der vorausgesetzten Stetigkeit der partiellen Ableitungen der Ordnung k folgern wir, dass für $\alpha \in \mathbb{N}^n$ mit $|\alpha| = k$

$$\lim_{x \rightarrow a} D^\alpha f(a + \vartheta_x \cdot (x - a)) - D^\alpha f(a) = 0$$

gilt. Weiter ist für $x \neq a$ jeder Faktor in dem Produkt

$$\frac{|(x - a)^\alpha|}{\|x - a\|^k} = \frac{|x_1 - a_1|^{\alpha_1}}{\|x - a\|^{\alpha_1}} \cdots \frac{|x_n - a_n|^{\alpha_n}}{\|x - a\|^{\alpha_n}}$$

im Intervall $[0, 1]$ enthalten. Fügen wir die beobachteten Tatsachen zusammen, dann erkennen wir, dass $r_k(x)/\|x - a\|^k$ gegen Null strebt, wenn x gegen a läuft. \square

Wir schreiben die Formel für $k = 2$ noch etwas um.

7.2.4 Definition. Es seien $B \subset \mathbb{R}^2$ eine offene Teilmenge und $f: B \rightarrow \mathbb{R}$ eine zweimal partiell differenzierbare Funktion. Für $x \in B$ ist

$$H_f(x) := \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2}(x) & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n}(x) \\ \vdots & & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1}(x) & \cdots & \frac{\partial^2 f}{\partial x_n^2}(x) \end{pmatrix}$$

die *Hesse-Matrix*² von f an der Stelle x .

7.2.5 Bemerkung. Zusätzlich zu den in der Definition genannten Voraussetzungen sei f zweimal **stetig** differenzierbar. Der Satz von Schwarz 5.3.4 impliziert, dass die Hesse-Matrix dann symmetrisch ist:

$$\forall x \in X: H_f(x)^t = H_f(x).$$

7.2.6 Folgerung. Es seien $B \subset \mathbb{R}^n$ eine offene Teilmenge, $f: B \rightarrow \mathbb{R}$ eine **zweimal stetig** differenzierbare Funktion, $a \in B$ und $\varepsilon > 0$, so dass $B(a, \varepsilon) \subset B$. Dann gilt

$$f(x) = f(a) + \langle \text{Grad}f(a), x - a \rangle + \frac{1}{2} \cdot (x - a) \cdot H_f(a) \cdot (x - a)^t + r_2(x), \quad x \in B(a, \varepsilon),$$

und

$$\lim_{x \rightarrow a} \frac{r_2(x)}{\|x - a\|^2} = 0.$$

7.3 Definite Matrizen

Die Formel in Folgerung 7.2.6 legt nahe, dass das Verhalten einer zweimal stetig differenzierbaren Funktion in der Nähe eines kritischen Punkts durch die Hesse-Matrix kontrolliert wird. In diesem Abschnitt wiederholen wir die Konzepte aus der linearen Algebra, die es uns gestatten, diesen Gedanken zu präzisieren.

Es sei $A \in \text{Mat}(n, \mathbb{R})$ eine **symmetrische** ($n \times n$)-Matrix, d.h.

$$A^t = A.$$

Die Abbildung

$$\begin{aligned} \beta_A: \mathbb{R}^n \times \mathbb{R}^n &\longrightarrow \mathbb{R} \\ (x, y) &\longmapsto x \cdot A \cdot y^t \end{aligned}$$

ist dann eine sogenannte *symmetrische Bilinearform* (s. (13), §39), d.h. sie erfüllt folgende Bedingungen:

1. $\forall x, y \in \mathbb{R}^n, \lambda \in \mathbb{R}: \beta_A(\lambda \cdot x, y) = \lambda \cdot \beta_A(x, y) = \beta_A(x, \lambda \cdot y)$.
2. $\forall x, y, z \in \mathbb{R}^n$:

$$\star \beta_A(x + y, z) = \beta_A(x, z) + \beta_A(y, z)$$

²Ludwig Otto Hesse (1811 - 1874), deutscher Mathematiker.

$$\star \beta_A(x, y+z) = \beta_A(x, y) + \beta_A(x, z).$$

$$3. \forall x, y \in \mathbb{R}^n: \beta_A(x, y) = \beta_A(y, x).$$

7.3.1 Definition. a) Wir sagen, dass die Matrix A *positiv* bzw. *negativ definit* ist, wenn

$$\forall x \in \mathbb{R}^n \setminus \{0\}: \beta_A(x, x) > 0 \quad \text{bzw.} \quad \beta_A(x, x) < 0.$$

b) Die Matrix A wird *positiv* bzw. *negativ semidefinit* genannt, wenn

$$\forall x \in \mathbb{R}^n: \beta_A(x, x) \geq 0 \quad \text{bzw.} \quad \beta_A(x, x) \leq 0.$$

c) Die Matrix A ist *indefinit*, wenn es Vektoren $x, y \in \mathbb{R}^n$ mit

$$\beta_A(x, x) < 0 < \beta_A(y, y)$$

gibt.

Die Frage, wann eine gegebene symmetrische Matrix A positiv oder negativ (semi)definit ist, wird in der Linearen Algebra behandelt (s. z.B. (11), Kapitel VII, §5). An dieser Stelle wollen wir den Fall $n = 2$ explizit besprechen. Sei also

$$A := \begin{pmatrix} a & b \\ b & c \end{pmatrix}$$

eine symmetrische (2×2) -Matrix. Ihre Determinante ist

$$\text{Det}(A) = ac - b^2.$$

Wir finden

$$(1, 0) \cdot \begin{pmatrix} a & b \\ b & c \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} = (1, 0) \cdot \begin{pmatrix} a \\ b \end{pmatrix} = a.$$

Wir halten fest:

7.3.2 Lemma. Wenn A *positiv* bzw. *negativ definit* ist, dann gilt $a > 0$ bzw. $a < 0$.

Jetzt betrachten wir einen Vektor $x = (x_1, x_2) \in \mathbb{R}^2$ mit $x_2 \neq 0$. Wegen

$$\beta_A(x, x) = x_2^2 \cdot \beta_A\left(\frac{x}{x_2}, \frac{x}{x_2}\right)$$

dürfen wir $x_2 = 1$ annehmen. Für $t \in \mathbb{R}$ und $x = (t, 1)$ haben wir

$$f(t) := \beta_A(x, x) = a \cdot t^2 + 2b \cdot t + c.$$

Wir studieren jetzt die Funktion $f: \mathbb{R} \rightarrow \mathbb{R}$.

Fall 1: $a > 0$. In diesem Fall gilt genau dann $f(t) > 0$ für alle $t \in \mathbb{R}$, wenn $f(t)$ keine reelle Nullstelle hat. Die (komplexen) Nullstellen von f werden mit der Formel

$$t_{1/2} = \frac{-2b \pm \sqrt{4b^2 - 4ac}}{2a}$$

gefunden. Damit hat f genau dann keine reelle Nullstellen, wenn $b^2 - ac < 0$, d.h. $\text{Det}(A) > 0$.

Fall 2: $a < 0$. Hier überlegt man sich analog, dass genau dann $f(t)$ für jede reelle Zahl t negativ ist, wenn $\text{Det}(A) > 0$.

Fall 3: $a = 0$. In diesem Fall nimmt f genau dann sowohl positive als auch negative Werte an, wenn $b \neq 0$. Da $\text{Det}(A) = ac - b^2 = -b^2$, ist diese Bedingung äquivalent zu $\text{Det}(A) < 0$.

Falls $a \neq 0$ und $\text{Det}(A) < 0$, dann hat f zwei verschiedene reelle Nullstellen, und es ist klar, dass f in diesem Fall sowohl positive als auch negative Werte annimmt. Dies wiederum bedeutet, dass A indefinit ist.

Der folgende Satz fasst unsere Überlegungen zusammen.

7.3.3 Satz. Es sei

$$A := \begin{pmatrix} a & b \\ b & c \end{pmatrix}$$

eine symmetrische (2×2) -Matrix. Dann gilt:

- i) Die Matrix A ist genau dann positiv definit, wenn $a > 0$ und $\text{Det}(A) > 0$.
- ii) Die Matrix A ist genau dann negativ definit, wenn $a < 0$ und $\text{Det}(A) > 0$.
- iii) Die Matrix A ist genau dann indefinit, wenn $\text{Det}(A) < 0$.

7.4 Ein hinreichendes Kriterium für isolierte lokale Extrema

Wir besprechen in diesem Abschnitt die interessante Verallgemeinerung des hinreichenden Kriteriums (12), Folgerung 4.3.9, ii), für die Existenz eines isolierten lokalen Extremums auf reellwertige Funktionen in mehreren Variablen. Zudem finden wir ein hinreichendes Kriterium für die Existenz eines Sattelpunkts, zu dem es kein Analogon in einer Veränderlichen gibt.

7.4.1 Definition. Es seien $B \subset \mathbb{R}^n$ eine offene Teilmenge, $f: B \rightarrow \mathbb{R}$ eine Funktion und $a \in B$. Die Funktion f hat in a ein *isoliertes lokales Extremum* (Minimum/Maximum), wenn sie dort ein lokales Extremum (Minimum/Maximum) besitzt und es ein $\varepsilon > 0$ gibt, so dass

$$\forall x \in B(a, \varepsilon) \setminus \{a\}: f(x) \neq f(a).$$

7.4.2 Satz. Es seien $B \subset \mathbb{R}^n$ eine offene Teilmenge und $f: B \rightarrow \mathbb{R}$ eine **zweimal stetig differenzierbare** Funktion. Für einen **kritischen Punkt** $a \in B$ von f gilt:

- i) Wenn die Hesse-Matrix $H_f(a)$ von f an der Stelle a positiv definit ist, dann liegt in a ein isoliertes lokales Minimum vor.
- ii) Ist die Hesse-Matrix $H_f(a)$ von f an der Stelle a negativ definit, so nimmt f in a ein isoliertes lokales Maximum an.
- iii) Falls die Hesse-Matrix $H_f(a)$ von f an der Stelle a indefinit ist, dann hat f in a einen Sattelpunkt.

7.4.3 Bemerkung. Für das dritte Kriterium gibt es für $n = 1$ keine Entsprechung.

Beweis. i) Nach der Taylorformel in der Fassung von Folgerung 7.2.6 gibt es wegen der Voraussetzung, dass a ein kritischer Punkt ist, also $\text{Grad}f(a) = 0$ gilt, ein $\eta > 0$ und eine Funktion $r_2: B(a, \eta) \rightarrow \mathbb{R}$, so dass

$$f(x) = f(a) + \frac{1}{2} \cdot (x - a) \cdot H_f(a) \cdot (x - a)^t + r_2(x), \quad x \in B(a, \eta),$$

und

$$\lim_{x \rightarrow a} \frac{r_2(x)}{\|x - a\|^2} = 0.$$

Auf Grund der letzten Bedingung gibt es zu einem vorgegebenen $\varepsilon > 0$ ein $\delta > 0$, so dass

$$\forall x \in B(a, \delta) \setminus \{a\}: \frac{|r_2(x)|}{\|x - a\|^2} < \varepsilon,$$

d.h.

$$\forall x \in B(a, \delta): |r_2(x)| \leq \varepsilon \cdot \|x - a\|^2. \quad (7.1)$$

Es sei

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

Zu $A := H_f(a)$ definieren wir die Funktion

$$\begin{aligned} q_A: \mathbb{R}^n &\longrightarrow \mathbb{R} \\ x &\longmapsto \beta_A(x, x) = x \cdot A \cdot x^t. \end{aligned}$$

Die Funktion q_A ist stetig und nimmt nach Voraussetzung auf S^{n-1} nur positive Werte an. Da S^{n-1} abgeschlossen und beschränkt und mithin nach dem Satz von Heine–Borel 2.2.12 kompakt ist, gibt es nach Folgerung 3.4.2 eine reelle Zahl $m > 0$, so dass

$$\forall x \in S^{n-1} : q_A(x) \geq m.$$

Für jeden Vektor $x \in \mathbb{R}^n$ haben wir deshalb

$$q_A(x) \geq m \cdot \|x\|^2.$$

Nun wählen wir $0 < \varepsilon < m/2$ und dazu ein $\delta > 0$, so dass (7.1) gilt. Für $x \in B(a, \delta) \setminus \{a\}$ schätzen wir

$$f(x) = f(a) + \frac{1}{2} \cdot q_A(x - a) + r_2(x) \geq f(a) + \left(\frac{m}{2} - \varepsilon\right) \cdot \|x - a\|^2 > f(a)$$

ab. Dies beweist die behauptete Aussage.

ii) Dieser Teil folgt, indem wir Teil i) auf $-f$ anwenden.

iii) Da die Matrix $A := H_f(a)$ gemäß unserer Voraussetzung indefinit ist, gibt es Vektoren $x, y \in \mathbb{R}^n$ mit

$$q_A(x) < 0 < q_A(y).$$

Wir betrachten eine beliebige offene Umgebung U von a . Es ist zu zeigen, dass es in U Punkte u, v mit $f(u) < f(a) < f(v)$ gibt. Wir wählen $\delta > 0$, so dass $B(a, \delta) \subset U$. Es gibt ein $\varepsilon > 0$, so dass $a + \varepsilon \cdot x \in B(a, \delta)$ und $a + \varepsilon \cdot y \in B(a, \delta)$. Seien $h := \varepsilon \cdot x$ und $k := \varepsilon \cdot y$. Da $B(a, \delta)$ konvex ist, gilt auch

$$a + t \cdot h \in B(a, \delta) \quad \text{und} \quad a + t \cdot k \in B(a, \delta), \quad t \in [0, 1].$$

Die Funktionen

$$\begin{aligned} g_1: [0, 1] &\longrightarrow \mathbb{R} \\ t &\longmapsto f(a + t \cdot h) \end{aligned}$$

und

$$\begin{aligned} g_2: [0, 1] &\longrightarrow \mathbb{R} \\ t &\longmapsto f(a + t \cdot k) \end{aligned}$$

sind zweimal stetig differenzierbar. Dabei gilt wegen Folgerung 7.2.6

$$\begin{aligned} g_1'(0) &= 0 = g_2'(0) \\ g_1''(0) &= q_A(h) = \varepsilon^2 \cdot q_A(x) < 0 \\ g_2''(0) &= q_A(k) = \varepsilon^2 \cdot q_A(y) > 0. \end{aligned}$$

Folglich hat g_1 in 0 ein isoliertes lokales Maximum und g_2 ein isoliertes lokales Minimum ((12), Folgerung 4.3.9). Damit existieren $t_1, t_2 \in [0, 1]$ mit

$$f(a + t_1 \cdot h) = g_1(t_1) < g_1(0) = f(a) = g_2(0) < g_2(t_2) = f(a + t_2 \cdot k),$$

und die Behauptung ist gezeigt. \square

7.4 Ein hinreichendes Kriterium für isolierte lokale Extrema

7.4.4 Beispiele (Fortführung der Beispiele 7.1.5). i) Es sei

$$\begin{aligned} f: \mathbb{R}^2 &\longrightarrow \mathbb{R} \\ (x, y) &\longmapsto x^2 + y^2. \end{aligned}$$

Der Ursprung $(0,0)$ ist der einzige kritische Punkt von f , und die Hesse-Matrix

$$H_f(0,0) = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$$

von f an der Stelle $(0,0)$ ist positiv definit. Nach Teil i) des Satzes liegt in $(0,0)$ ein isoliertes lokales Minimum vor. Das können wir durch Augenschein sofort bestätigen.

ii) Für die Funktion

$$\begin{aligned} f: \mathbb{R}^2 &\longrightarrow \mathbb{R} \\ (x, y) &\longmapsto 1 - 3x^2y^4 \end{aligned}$$

ist

$$H_f(x, y) = \begin{pmatrix} -6y^4 & -24xy^3 \\ -24xy^3 & -36x^2y^2 \end{pmatrix}$$

die Hesse-Matrix an der Stelle (x, y) . Wir haben zuvor bereits beobachtet, dass

$$K = \left\{ (x, y) \in \mathbb{R}^2 \mid x \cdot y = 0 \right\}$$

die Menge der kritischen Punkte ist und dass in jedem kritischen Punkt ein globales Maximum vorliegt. Nach dem vorhergehenden Satz darf die Matrix $H_f(x, y)$ für einen Punkt $(x, y) \in K$ weder definit noch indefinit sein. In der Tat ist

$$H_f(0, y) = \begin{pmatrix} -6y^4 & 0 \\ 0 & 0 \end{pmatrix}, \quad y \in \mathbb{R},$$

negativ semidefinit, und

$$H_f(x, 0) = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \quad x \in \mathbb{R},$$

ist sowohl positiv als auch negativ semidefinit.

iii) Der einzige kritische Punkt der Funktion

$$\begin{aligned} f: \mathbb{R}^2 &\longrightarrow \mathbb{R} \\ (x, y) &\longmapsto x^2 + 3y^3 \end{aligned}$$

ist der Ursprung $(0,0)$. Dort hat f einen Sattelpunkt. Die Hesse-Matrix

$$H_f(0,0) = \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix}$$

von f an der Stelle $(0,0)$ ist positiv semidefinit. Das in Satz 7.4.2, iii), formulierte Kriterium ist somit hinreichend aber nicht notwendig für die Existenz eines Sattelpunkts. Außerdem zeigt sich, dass wir aus der Semidefinitheit der Hesse-Matrix nichts über die Natur des kritischen Punkts ableiten können.

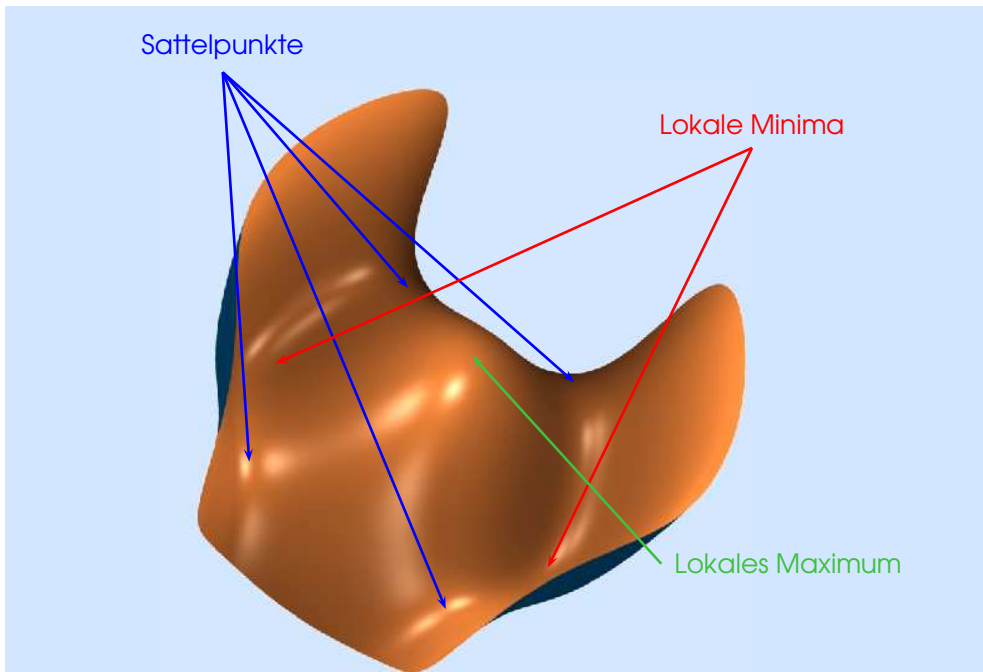
iv) Wir untersuchen die Polynom-Funktion

$$\begin{aligned} f: \mathbb{R}^2 &\longrightarrow \mathbb{R} \\ (x, y) &\longmapsto x^4 - 2x^2 + 2x^2y^2 - y^2. \end{aligned}$$

Die partiellen Ableitungen sind

$$\frac{\partial f}{\partial x}(x, y) = 4x^3 - 4x + 4xy^2 = 4x \cdot (x^2 - 1 + y^2)$$

$$\frac{\partial f}{\partial y}(x, y) = 4x^2y - 2y = 2y \cdot (2x^2 - 1), \quad (x, y) \in \mathbb{R}^2.$$



Wir finden insgesamt sieben kritische Punkte:

$$(0, 0), \quad (-1, 0), \quad (1, 0), \quad \pm \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}} \right), \quad \pm \left(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}} \right).$$

Die Hesse-Matrix von f an der Stelle (x, y) ist

$$H_f(x, y) = \begin{pmatrix} 12x^2 - 4 + 4y^2 & 8xy \\ 8xy & 4x^2 - 2 \end{pmatrix}.$$

Somit gilt

$$H_f(0, 0) = \begin{pmatrix} -4 & 0 \\ 0 & -2 \end{pmatrix}.$$

Dies ist offensichtlich eine negativ definite Matrix. Bei dem kritischen Punkt $(0, 0)$ handelt es sich um ein isoliertes lokales Maximum.

Weiter haben wir

$$H_f(\pm 1, 0) = \begin{pmatrix} 8 & 0 \\ 0 & 2 \end{pmatrix}.$$

Dies ist eine positiv definite Matrix, so dass f an den Stellen $(\pm 1, 0)$ isolierte lokale Minima hat.

Es gilt ferner

$$H_f\left(\pm \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}} \right)\right) = \begin{pmatrix} 4 & 4 \\ 4 & 0 \end{pmatrix}.$$

Wir berechnen

$$\text{Det} \begin{pmatrix} 4 & 4 \\ 4 & 0 \end{pmatrix} = -16 < 0.$$

Nach Satz 7.3.3, iii), ist die Hesse-Matrix indefinit. Ebenso ist die Determinante der Hesse-Matrix

$$H_f \left(\pm \left(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}} \right) \right) = \begin{pmatrix} 4 & -4 \\ -4 & 0 \end{pmatrix}$$

negativ. Die kritischen Punkte

$$\pm \left(\frac{1}{\sqrt{2}}, \pm \frac{1}{\sqrt{2}} \right)$$

sind also Sattelpunkte von f .

7.5 Extrema unter Nebenbedingungen

Man vergewissert sich schnell der Tatsache, dass die Funktion

$$\begin{aligned} f: \mathbb{R}^2 &\longrightarrow \mathbb{R} \\ (x, y) &\longmapsto y \end{aligned}$$

keinerlei kritische Punkte hat, insbesondere auch keine lokalen Extrema. Seien

$$g(x, y) := y - x^2 - 1, \quad (x, y) \in \mathbb{R}^2,$$

und

$$N(g) := \{ (x, y) \in \mathbb{R}^2 \mid g(x, y) = 0 \} = \{ (x, y) \in \mathbb{R}^2 \mid y = x^2 + 1 \}.$$

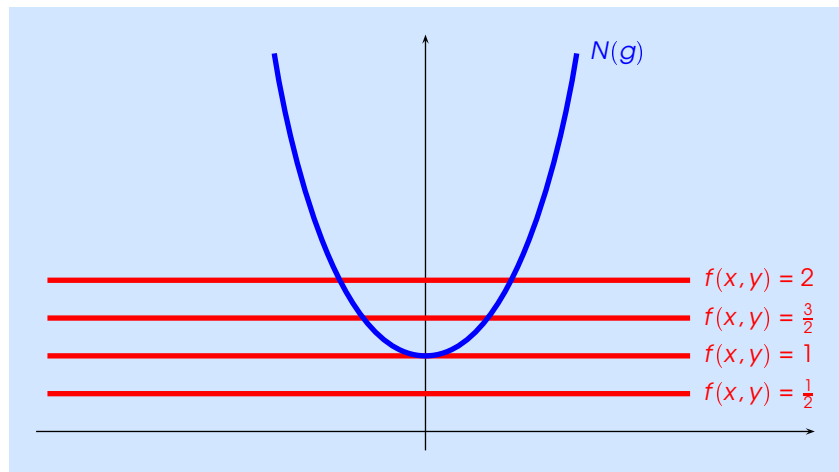
Bei $N(g)$ handelt es sich offenkundig um eine Parabel. Auf der Menge $N(g)$ nimmt f ein Minimum an: Für $(x, y) \in N(g)$ gilt $y = x^2 + 1$, so dass

$$f(x, y) = x^2 + 1 \geq 1,$$

und für $(0, 1) \in N(g)$ hat man

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

Wir sagen, dass f in $(0, 1)$ ein *Minimum unter der Nebenbedingung* $g = 0$ annimmt.



Wir weisen darauf hin, dass im Punkt $(0, 1)$, also in dem Punkt, in dem das Minimum angenommen wird, die Niveaulinie

$$f^{-1}(1) = \{ (x, y) \in \mathbb{R}^2 \mid y = f(x, y) = 1 \}$$

eine **Tangente** an $N(g)$ ist.

In diesem Abschnitt werden wir das obige Beispiel zu einer Problemstellung formalisieren und uns mit der Lösung dieser Problemstellung beschäftigen.

7.5.1 Definition. Es seien $B \subset \mathbb{R}^n$ eine offene Teilmenge, $g: B \rightarrow \mathbb{R}$ eine **stetig differenzierbare** Funktion, so dass

$$\forall x \in B: \text{Grad}g(x) \neq 0 \tag{7.2}$$

und

$$N(g) := \{ x \in B \mid g(x) = 0 \}.$$

Weiter seien $f: B \rightarrow \mathbb{R}$ eine Funktion und $a \in B$ ein Punkt mit $g(a) = 0$.

i) Die Funktion f nimmt in a ein *lokales Minimum (lokales Maximum)* unter der Nebenbedingung $g = 0$ an, wenn es ein $\varepsilon > 0$ gibt, so dass

$$\forall x \in B(a, \varepsilon) \cap N(g): f(x) \geq f(a) \quad (f(x) \leq f(a)).$$

ii) In a liegt ein *globales Minimum (globales Maximum)* der Funktion f unter der Nebenbedingung $g = 0$ vor, wenn

$$\forall x \in N(g): f(x) \geq f(a) \quad (f(x) \leq f(a)).$$

7.5.2 Beispiel. Wir betrachten die Funktion

$$\begin{aligned} f: \mathbb{R}^3 &\longrightarrow \mathbb{R} \\ (x, y, z) &\longmapsto 3x^2 + 3y^2 + z^2. \end{aligned}$$

Die Nebenbedingung wird mit Hilfe der Funktion

$$\begin{aligned} g: \mathbb{R}^3 &\longrightarrow \mathbb{R} \\ (x, y, z) &\longmapsto x + y + z - 1 \end{aligned}$$

erklärt. Wegen

$$\text{Grad}g(x, y, z) = (1, 1, 1), \quad (x, y, z) \in \mathbb{R}^3,$$

ist Bedingung (7.2) erfüllt.

In diesem Beispiel ist es möglich, die Nebenbedingung einzusetzen:

$$\forall (x, y, z) \in \mathbb{R}^3: g(x, y, z) = 0 \iff z = 1 - x - y.$$

Wir müssen also die Funktion

$$\begin{aligned} q(x, y) &= f(x, y, 1 - x - y) \\ &= 3x^2 + 3y^2 + (1 - x - y)^2 \\ &= 3x^2 + 3y^2 + 1 - 2x - 2y + 2xy + x^2 + y^2 \\ &= 4x^2 + 4y^2 + 2xy - 2x - 2y + 1 \end{aligned}$$

auf Extremwerte untersuchen. Der Gradient ist

$$\text{Grad}q(x, y) = (8x + 2y - 2, 8y + 2x - 2).$$

Ein Punkt $(x, y) \in \mathbb{R}^2$ ist genau dann Nullstelle des Gradienten, wenn

$$8x + 2y = 2 = 8y + 2x, \quad \text{i.e.} \quad x = y = \frac{1}{5},$$

gilt. Wir erkennen, dass

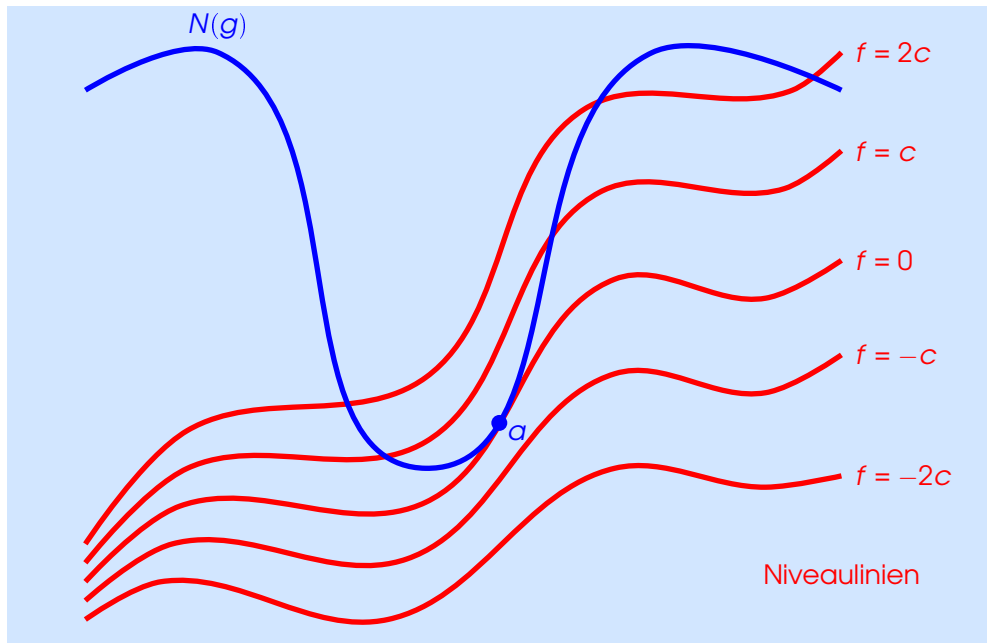
$$\left(\frac{1}{5}, \frac{1}{5}\right)$$

der einzige kritische Punkt von q ist. Die Hesse-Matrix

$$A := H_q\left(\frac{1}{5}, \frac{1}{5}\right) = \begin{pmatrix} 8 & 2 \\ 2 & 8 \end{pmatrix}$$

ist wegen $8 > 0$, $\text{Det}(A) = 60 > 0$ und Satz 7.3.3 positiv definit. Nach Satz 7.4.2, i), hat q in $(1/5, 1/5)$ ein isoliertes lokales Minimum. Somit hat f in $(1/5, 1/5, 3/5)$ ein lokales Minimum unter der Nebenbedingung $g = 0$.

Wir wollen ein notwendiges Kriterium für die Existenz eines lokalen Extremums unter einer Nebenbedingung formulieren. Das folgende Bild unterstützt unsere Anschauung:



In a „berühren“ sich $f^{-1}(f(a))$ und $N(g)$. Nach Satz 5.2.7, ii), stehen der Gradient $\text{Grad}f(a)$ bzw. $\text{Grad}g(a)$ senkrecht auf $f^{-1}(f(a))$ bzw. $N(g)$. Damit müssen $\text{Grad}f(a)$ und $\text{Grad}g(a)$ parallel sein.

7.5.3 Satz (Methode der Lagrangeschen Multiplikatoren). *Es gelten die Voraussetzungen von Definition 7.5.1. Zudem sei die Funktion f in a differenzierbar und besitze dort ein lokales Extremum unter der Nebenbedingung $g = 0$. Dann gibt es eine Zahl $\lambda \in \mathbb{R}$ mit*

$$\text{Grad}f(a) = \lambda \cdot \text{Grad}g(a).$$

7.5.4 Definition. Die Zahl λ heißt *Lagrangescher Multiplikator*.

Beim Beweis des obigen Satzes verwenden wir eine weitere Variante des Mittelwertsatzes:

7.5.5 Satz. Es seien $B \subset \mathbb{R}^n$ eine offene und konvexe Teilmenge und $f: B \rightarrow \mathbb{R}$ eine differenzierbare Funktion. Dann gibt es zu je zwei Punkten $a, b \in B$ ein $t \in [0, 1]$ mit

$$f(b) - f(a) = \langle \text{Grad}f((1-t) \cdot a + t \cdot b), b - a \rangle.$$

Beweis. Wir betrachten den differenzierbaren Weg

$$\begin{aligned} \alpha: [0, 1] &\longrightarrow B \\ t &\longmapsto (1-t) \cdot a + t \cdot b. \end{aligned}$$

Die Verknüpfung $h := f \circ \alpha: [0, 1] \rightarrow \mathbb{R}$ ist eine differenzierbare Funktion. Nach dem Mittelwertsatz in einer Variablen ((12), Folgerung 4.3.5) gibt es ein $t \in (0, 1)$ mit

$$f(b) - f(a) = h(1) - h(0) = h'(t).$$

Die Kettenregel 6.3.1 liefert

$$h'(t) = \langle \text{Grad}f((1-t) \cdot a + t \cdot b), b - a \rangle,$$

so dass sich die gewünschte Aussage ergibt. \square

Beweis von Satz 7.5.3. Wir beginnen mit etwas Linearer Algebra. Es seien $v \in \mathbb{R}^n \setminus \{0\}$ ein nichttrivialer Vektor und $\langle v \rangle \subset \mathbb{R}^n$ der von v aufgespannte eindimensionale Unterraum. Dann ist

$$U := \langle v \rangle^\perp := \{ h \in \mathbb{R}^n \mid \langle v, h \rangle = 0 \}$$

ein $(n-1)$ -dimensionaler Unterraum von \mathbb{R}^n , so dass

$$\mathbb{R}^n = \langle v \rangle \oplus U.$$

Nun sei $w \in \mathbb{R}^n$ ein weiterer Vektor. Wir können $w = \lambda \cdot v + w'$ mit $\lambda \in \mathbb{R}$ und $w' \in U$ schreiben. Es gilt

$$\langle w, w' \rangle = \langle w', w' \rangle.$$

Da $\langle w', w' \rangle > 0$, falls $w' \neq 0$, schließen wir:

$$\exists \lambda \in \mathbb{R}: w = \lambda \cdot v \iff \forall h \in U: \langle w, h \rangle = 0.$$

Auf Grund dieser Vorüberlegungen haben wir zu zeigen:

$$\forall h \in \mathbb{R}^n \setminus \{0\}: \langle \text{Grad}g(a), h \rangle = 0 \implies \langle \text{Grad}f(a), h \rangle = 0.$$

Wir nehmen im Moment an, dass zu jedem Vektor $h \in \mathbb{R}^n \setminus \{0\}$ mit $\langle \text{Grad}g(a), h \rangle = 0$ eine Zahl $\delta > 0$ und eine Abbildung

$$\alpha_h: [-\delta, \delta] \longrightarrow B$$

mit den Eigenschaften

- α_h ist in 0 differenzierbar,
- $\alpha_h(0) = a, \alpha_h'(0) = h,$
- $\forall s \in [-\delta, \delta]: \alpha_h(s) \in N(g),$

existieren.³ Dann ist $f \circ \alpha_h: [-\delta, \delta] \rightarrow \mathbb{R}$ wegen der ersten Eigenschaft in 0 differenzierbar und nimmt wegen der dritten Eigenschaft und der Voraussetzung in 0 ein lokales Extremum an. Nach (12), Satz 4.3.2, und der Kettenregel 6.3.1 haben wir

$$0 = (f \circ \alpha_h)'(0) = \langle \text{Grad}f(a), h \rangle.$$

Das zeigt die Behauptung des Satzes. Es bleibt, die Richtigkeit unserer Annahme zu überprüfen. Dies geschieht in zwei Schritten.

Schritt 1: Reduktion auf eine zweidimensionale Situation. Wegen der Voraussetzung (7.2) ist der Vektor $v := \text{Grad}g(a)$ nicht trivial. Seien weiter $h \in \mathbb{R}^n \setminus \{0\}$ ein Vektor mit $\langle v, h \rangle = 0$ sowie

$$\begin{aligned} w: \mathbb{R}^2 &\longrightarrow \mathbb{R} \\ (s, t) &\longmapsto a + s \cdot h + t \cdot v. \end{aligned}$$

Es gibt ein $\varepsilon > 0$, so dass $w(I \times I) \subset B$, $I := [-\varepsilon, \varepsilon]$. Die Abbildung

$$\begin{aligned} \rho: I \times I &\longrightarrow \mathbb{R} \\ (s, t) &\longmapsto \langle \text{Grad}g(a + s \cdot h + t \cdot v), v \rangle \end{aligned}$$

ist stetig mit

$$\rho(0, 0) = \|\text{Grad}g(a)\|^2 > 0.$$

Wegen der Stetigkeit von ρ können wir nach Verkleinerung von ε annehmen:

$$\forall (s, t) \in I \times I: \rho(s, t) > 0.$$

Für die Abbildung

$$\begin{aligned} \gamma: I \times I &\longrightarrow \mathbb{R} \\ (s, t) &\longmapsto g(a + s \cdot h + t \cdot v) \end{aligned}$$

gilt

- $\{(s, t) \in I \times I \mid \gamma(s, t) = 0\} = \{(s, t) \in I \times I \mid a + s \cdot h + t \cdot v \in N(g)\}$,
- $\gamma(0, 0) = g(a) = 0$,
- $\frac{\partial \gamma}{\partial s}(0, 0) = \langle \text{Grad}g(a), h \rangle = 0$,
- $\frac{\partial \gamma}{\partial t}(s, t) = \langle \text{Grad}g(a + s \cdot h + t \cdot v), v \rangle = \rho(s, t) > 0, (s, t) \in I \times I$.

Schritt 2: Konstruktion des Wegs α_h . Die Funktion

$$\begin{aligned} \Gamma_0: I &\longrightarrow \mathbb{R} \\ t &\longmapsto \gamma(0, t) \end{aligned}$$

erfüllt $\Gamma_0(0) = \gamma(0, 0)$ und ist wegen $\Gamma_0'(t) = (\partial \gamma / \partial t)(0, t) > 0, t \in I$, streng monoton wachsend ((12), Folgerung 4.3.8, iii). Deshalb gilt

$$\gamma(0, -\varepsilon) = \Gamma_0(-\varepsilon) < \Gamma_0(0) = 0 < \Gamma_0(\varepsilon) = \gamma(0, \varepsilon).$$

³Die Tangentialebene an die Menge $N(g)$ im Punkt a besteht gerade aus den Vektoren $a + h$, für die $h \in \mathbb{R}^n$ ein Vektor mit $\langle \text{Grad}g(a), h \rangle = 0$ ist. Unsere Aussage besagt, dass wir zu jedem „Tangentialvektor“ h einen differenzierbaren Weg α_h finden, der ganz in $N(g)$ verläuft und dessen Ableitung („Geschwindigkeit“) in a der gegebene Tangentialvektor h ist. Mit anderen Worten, die Tangentialvektoren sind die Ableitungen differenzierbarer Wege in $N(g)$ durch a an der Stelle 0.

Kapitel 7 Extremwerte differenzierbarer Funktionen

Da γ eine stetige Funktion ist, existiert auch ein $\delta > 0$, so dass

$$\forall s \in [-\delta, \delta]: \quad \gamma(s, -\varepsilon) < 0 < \gamma(s, \varepsilon).$$

Nach dem Zwischenwertsatz ((12), Folgerung 3.5.2) existiert zu einem vorgegebenem Wert $s \in [-\delta, \delta]$ eine Zahl $\alpha(s) \in (-\varepsilon, \varepsilon)$, so dass

$$\gamma(s, \alpha(s)) = 0.$$

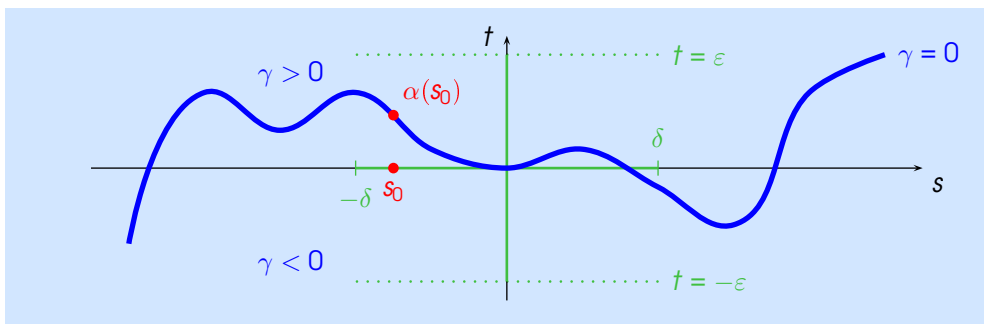
Die Zahl $\alpha(s)$ ist sogar eindeutig bestimmt, weil die Funktion

$$\begin{aligned} \Gamma_s: I &\longrightarrow \mathbb{R} \\ t &\longmapsto \gamma(s, t) \end{aligned}$$

wegen $\Gamma'_s(t) = (\partial\gamma/\partial t)(s, t) > 0$, $s \in [-\delta, \delta]$, $t \in I$, ebenfalls streng monoton wachsend ist. Somit wird durch

$$\begin{aligned} \alpha: [-\delta, \delta] &\longrightarrow I \\ s &\longmapsto \alpha(s) \end{aligned}$$

eine Funktion definiert.



Behauptung. Die Funktion α ist in 0 differenzierbar, und es gilt $\alpha'(0) = 0$.

Wenn die Behauptung stimmt, dann hat die Abbildung

$$\begin{aligned} \alpha_h: [-\delta, \delta] &\longrightarrow \mathbb{R}^n \\ s &\longmapsto a + s \cdot h + \alpha(s) \cdot v \end{aligned}$$

folgende Eigenschaften:

- $\alpha_h(0) = a + \alpha(0) \cdot v = a$,
- $\alpha'_h(0) = h + \alpha'(0) \cdot v = h$,
- $g(\alpha_h(s)) = \gamma(s, \alpha(s)) = 0$, $s \in [-\delta, \delta]$.

Mit anderen Worten: Wir haben die Existenz einer Abbildung mit den für den Beweis des Satzes erforderlichen Eigenschaften nachgewiesen.

Zu guter Letzt zeigen wir die Behauptung. Nach dem Mittelwertsatz 7.5.5 gibt es zu $s \in [-\delta, \delta]$ ein $c \in [0, 1]$ mit

$$0 = \gamma(s, \alpha(s)) - \gamma(0, 0) = \left\langle \text{Grad}\gamma(c \cdot s, c \cdot \alpha(s)), (s, \alpha(s)) \right\rangle,$$

also

$$s \cdot \frac{\partial\gamma}{\partial s}(c \cdot s, c \cdot \alpha(s)) + \alpha(s) \cdot \underbrace{\frac{\partial\gamma}{\partial t}(c \cdot s, c \cdot \alpha(s))}_{>0} = 0.$$

Für $s \neq 0$ folgt

$$\frac{\alpha(s)}{s} = - \frac{\frac{\partial \gamma}{\partial s}(c \cdot s, c \cdot \alpha(s))}{\frac{\partial \gamma}{\partial t}(c \cdot s, c \cdot \alpha(s))}. \quad (7.3)$$

Aus der Konstruktion folgt, dass γ stetig differenzierbar ist. Wegen $(c \cdot s, c \cdot \alpha(t)) \in [-\delta, \delta] \times [-\varepsilon, \varepsilon]$ folgt aus (7.3), dass $\lim_{s \rightarrow 0} \alpha(s) = 0$ ist und zusammen mit der Stetigkeit der partiellen Ableitungen

$$\lim_{s \rightarrow 0} \frac{\alpha(s)}{s} = - \frac{\frac{\partial \gamma}{\partial s}(0,0)}{\frac{\partial \gamma}{\partial t}(0,0)} = 0.$$

Dies ist die Behauptung. □

7.5.6 Bemerkung. i) Es gibt hinreichende Kriterien für die Existenz von lokalen Extrema unter Nebenbedingungen, die Satz 7.4.2 entsprechen. Dazu verweisen wir z.B. auf (9), Satz 24.6, S. 164. In den Beispielen des nächsten Abschnitts werden wir einige Kniffe kennenlernen, mit denen man entscheiden kann, ob an einer Stelle, die man mit dem Verfahren der Lagrangeschen Multiplikatoren gefunden hat, wirklich ein Extremum vorliegt.

ii) In Abschnitt 11.6 werden wir die Problemstellung auf die Situation, in der mehrere Nebenbedingungen gestellt werden, verallgemeinern.

7.6 Beispiele

Dieser Abschnitt enthält verschiedene Beispiele, in denen das Verfahren der Lagrangeschen Multiplikatoren verwendet wird, um Extrema unter Nebenbedingungen zu entdecken. Insbesondere werden Möglichkeiten besprochen, mit denen man feststellen kann, ob an einer gefundenen „kritischen“ Stelle ein Extremum vorliegt.

Beispiel 1 (vgl. Beispiel 7.5.2)

Wir untersuchen die Funktion

$$\begin{aligned} f: \mathbb{R}^3 &\longrightarrow \mathbb{R} \\ (x, y, z) &\longmapsto 3x^2 + 3y^2 + z^2 \end{aligned}$$

unter der Nebenbedingung

$$\begin{aligned} g: \mathbb{R}^3 &\longrightarrow \mathbb{R} \\ (x, y, z) &\longmapsto x + y + z - 1. \end{aligned}$$

Da

$$\forall (x, y, z) \in \mathbb{R}^3: \text{Grad}g(x, y, z) = (1, 1, 1) \neq 0,$$

sind die notwendigen Voraussetzungen für unsere Untersuchungen gegeben. Weiter gilt

$$\text{Grad}f(x, y, z) = (6x, 6y, 2z), \quad (x, y, z) \in \mathbb{R}^3.$$

Wir wollen nun reelle Zahlen a, b, c und λ auffinden, so dass

$$\text{Grad}f(a, b, c) = \lambda \cdot \text{Grad}g(a, b, c) \quad \text{und} \quad g(a, b, c) = 0.$$

Diese Bedingungen entsprechen dem Gleichungssystem

$$\begin{aligned} 6a &= \lambda \\ 6b &= \lambda \\ 2c &= \lambda \\ a + b + c &= 1. \end{aligned}$$

Dies formen wir zu den Gleichungen $a = b = \lambda/6$, $c = \lambda/2$ und

$$\frac{\lambda}{6} + \frac{\lambda}{6} + \frac{\lambda}{2} = 1$$

um. Damit finden wir $\lambda = 6/5$, $a = b = 1/5$ und $c = 3/5$.

Die Probe bestätigt

$$\text{Grad}f\left(\frac{1}{5}, \frac{1}{5}, \frac{3}{5}\right) = \begin{pmatrix} 6 & 6 & 6 \\ 5 & 5 & 5 \end{pmatrix} = \frac{6}{5} \cdot \text{Grad}g\left(\frac{1}{5}, \frac{1}{5}, \frac{3}{5}\right).$$

Es bleibt zu untersuchen, ob an der gefundenen Stelle wirklich ein Extremum vorliegt.

Verfahren 1. Wir kürzen $\alpha := (a, b, c)$ ab und betrachten einen Vektor $h = (h_1, h_2, h_3) \in \mathbb{R}^3 \setminus \{0\}$, so dass

$$g(\alpha + h) = 0.$$

Mit der konkreten Formel für g heißt das

$$h_1 + h_2 + h_3 = 0. \quad (7.4)$$

Nun gilt

$$\begin{aligned} f(\alpha + h) &= 3 \cdot (a + h_1)^2 + 3 \cdot (b + h_2)^2 + (c + h_3)^2 \\ &= 3a^2 + 3b^2 + c^2 + 6ah_1 + 3h_1^2 + 6bh_2 + 3h_2^2 + 2ch_3 + h_3^2 \\ &\stackrel{h \neq 0}{>} f(\alpha) + 6ah_1 + 6bh_2 + 2ch_3 \\ &= f(\alpha) + \frac{6}{5} \cdot (h_1 + h_2 + h_3) = f(\alpha). \end{aligned}$$

Bei der vorletzten Gleichung haben wir $(a, b, c) = (1/5, 1/5, 3/5)$ und bei der letzten (7.4) benutzt. Diese Betrachtungen zeigen, dass in α ein (isoliertes globales) Minimum unter der Nebenbedingung g vorliegt.

Verfahren 2. Wir bemerken, dass

$$f(x, y, z) = 2x^2 + 2y^2 + \|(x, y, z)\|^2 \geq \|(x, y, z)\|^2, \quad (x, y, z) \in \mathbb{R}^3, \quad (7.5)$$

und

$$f\left(\frac{1}{5}, \frac{1}{5}, \frac{3}{5}\right) = \frac{3}{5} < 1 = f(0, 0, 1). \quad (7.6)$$

Seien

$$\begin{aligned} \bar{B} &:= \left\{ (x, y, z) \in \mathbb{R}^3 \mid \|(x, y, z)\| \leq 1 \right\}, \\ N(g) &:= \left\{ (x, y, z) \in \mathbb{R}^3 \mid g(x, y, z) = 0 \right\}. \end{aligned}$$

Da g eine stetige Funktion ist, ist $N(g)$ eine abgeschlossene Teilmenge von \mathbb{R}^3 . Mithin ist $M := \bar{B} \cap N(g)$ eine kompakte Teilmenge von \mathbb{R}^3 (Satz von Heine–Borel 2.2.10). Nach Folgerung 3.4.2 nimmt f auf M ein Minimum an. Aus (7.5) und (7.6) folgt, dass dies ein globales Minimum für f unter der Nebenbedingung g ist. Deshalb **müssen** wir es mit Satz 7.5.3 entdecken. Da dieser Satz uns α als einzigen Kandidaten liefert, folgt, dass in α ein (isoliertes globales) Minimum unter der Nebenbedingung g vorliegt.

7.6.1 Bemerkung. Die als Verfahren 2 vorgestellte Methode, die benutzt, dass eine stetige Funktion auf einer kompakten Menge M ein Minimum bzw. Maximum annimmt (Folgerung 3.4.2), ist oft nützlich. Man muss aber aufpassen, dass das Minimum bzw. Maximum auch in einem Randpunkt $\alpha \in \partial M$ angenommen werden kann. In diesem Fall ist es möglicherweise kein lokales Minimum bzw. Maximum von f unter der gestellten Nebenbedingung. (Daher haben wir bei den obigen Betrachtungen großen Wert darauf gelegt, dass wir wissen, dass das Minimum auf M ein globales Minimum ist.) In der Tat nimmt die Funktion f aus dem Beispiel auf M ja auch ein Maximum an. Mit der Gleichung in (7.5) sieht man leicht, dass dieses Maximum auf dem Rand ∂M angenommen wird und **kein lokales Maximum** von f unter der Nebenbedingung g ist.

Beispiel 2

Es sei $A \in \text{Mat}(n, \mathbb{R})$ eine **symmetrische** ($n \times n$)-Matrix. Wir betrachten die Funktion

$$\begin{aligned} f: \mathbb{R}^n &\longrightarrow \mathbb{R} \\ x &\longmapsto x \cdot A \cdot x^t. \end{aligned}$$

Die Nebenbedingung werde durch

$$\begin{aligned} g: \mathbb{R}^n &\longrightarrow \mathbb{R} \\ x &\longmapsto x \cdot x^t - 1 \end{aligned}$$

definiert. Diesmal ist

$$N(g) = \{ x \in \mathbb{R}^n \mid \|x\| = 1 \} = S^{n-1}$$

kompakt, so dass wir a priori wissen, dass f ein globales Minimum und ein globales Maximum unter der Nebenbedingung g annimmt. Sei weiter

$$U := \mathbb{R}^n \setminus \{0\} \supset S^{n-1}.$$

Für $x \in U$ haben wir

$$\text{Grad}g(x) = 2x \neq 0$$

sowie

$$\text{Grad}f(x) = 2x \cdot A.$$

Gesucht sind nun ein Vektor $x \in \mathbb{R}^n$ und eine reelle Zahl λ mit $\|x\| = 1$ und

$$\text{Grad}f(x) = 2x \cdot A = 2\lambda \cdot x = \lambda \cdot \text{Grad}g(x), \quad \text{i.e.} \quad A \cdot x^t = A^t \cdot x^t = \lambda \cdot x^t. \quad (7.7)$$

Damit wurde gezeigt:

7.6.2 Satz. *Die möglichen Lagrangeschen Multiplikatoren sind die reellen Eigenwerte von A .*

Wie zuvor bemerkt, wissen wir ja, dass es Punkte $m, M \in N(g)$ gibt, so dass f in m ein globales Minimum unter der Nebenbedingung g und in M ein globales Maximum unter der Nebenbedingung g hat. Insbesondere existieren $x \in \mathbb{R}^n$ mit $\|x\| = 1$ und $\lambda \in \mathbb{R}$, für die (7.7) gilt. Damit haben wir den folgenden bekannten Satz aus der linearen Algebra bewiesen (vgl. (13), §46):

7.6.3 Satz. *Die Matrix A besitzt einen reellen Eigenwert.*

7.6.4 Bemerkung. i) Man zeigt leicht induktiv, dass A n reelle Eigenwerte $\lambda_1 \leq \dots \leq \lambda_n$ besitzt (s. (13), §46).

ii) Der Lagrange-Multiplikator zu einem Minimum bzw. Maximum ist der kleinste bzw. größte Eigenwert λ_1 bzw. λ_n der Matrix A .

Beispiel 3

Es sei $U > 0$ eine reelle Zahl. Gesucht sind die Kantenlängen a und b des Rechtecks mit Umfang U , dessen Flächeninhalt maximal unter allen Rechtecken des Umfangs U wird.

Den Flächeninhalt des Rechtecks beschreiben wir mit der Funktion

$$\begin{aligned} f: \mathbb{R}^2 &\longrightarrow \mathbb{R} \\ (x, y) &\longmapsto x \cdot y \end{aligned}$$

und die Nebenbedingung, dass das Rechteck den vorgegebenen Umfang U haben möge, mit der Funktion

$$\begin{aligned} g: \mathbb{R}^2 &\longrightarrow \mathbb{R} \\ (x, y) &\longmapsto 2 \cdot (x + y) - U. \end{aligned}$$

Wegen

$$\forall (x, y) \in \mathbb{R}^2: \text{Grad}g(x, y) = (2, 2) \neq 0,$$

sind die notwendigen Voraussetzungen wieder erfüllt. Weiter hat man

$$\text{Grad}f(x, y) = (y, x), \quad (x, y) \in \mathbb{R}^2.$$

Wir möchten nun reelle Zahlen a , b und λ finden, so dass

$$\text{Grad}f(a, b) = \lambda \cdot \text{Grad}g(a, b) \quad \text{und} \quad g(a, b) = 0,$$

d.h.

$$\begin{aligned} a &= 2\lambda \\ b &= 2\lambda \\ a + b &= \frac{U}{2}. \end{aligned}$$

Man berechnet daraus $a = b = U/4$. (Dies besagt, dass das gesuchte Rechteck das Quadrat ist.)

Warum handelt es sich hierbei um ein Maximum? Dazu greifen wir das zweite Verfahren aus dem ersten Beispiel wieder auf. Die „sinnvollen“ Lösungen liegen in

$$\overset{\circ}{K} := \left(0, \frac{U}{2}\right) \times \left(0, \frac{U}{2}\right) \subset \left[0, \frac{U}{2}\right] \times \left[0, \frac{U}{2}\right] =: K.$$

Für $(x, y) \in N(g) \setminus \overset{\circ}{K}$ findet man $f(x, y) \leq 0$. Auf der anderen Seite gilt

$$f\left(\frac{U}{4}, \frac{U}{4}\right) = \frac{U^2}{16} > 0.$$

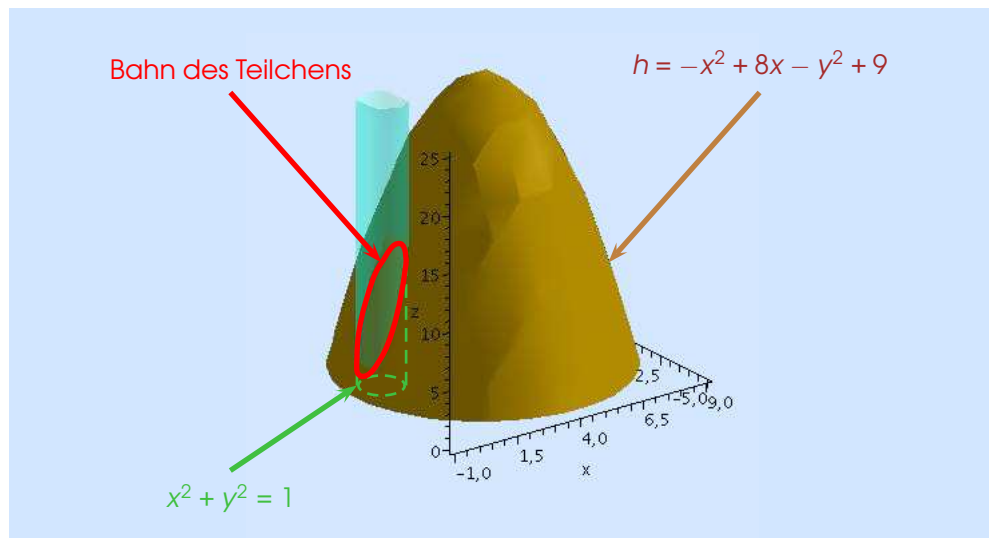
Damit nimmt f ein globales Maximum unter der Nebenbedingung g an, das an einer Stelle in $\overset{\circ}{K}$ auftauchen muss. Diese Stelle findet man mit dem Verfahren der Lagrangeschen Multiplikatoren. Hier haben wir nur einen Kandidaten, nämlich $(U/4, U/4)$.

Beispiel 4

Ein Teilchen der Masse $m > 0$ bewege sich im \mathbb{R}^3 auf der Bahn, die durch die Gleichungen

$$\begin{aligned} h &= -x^2 + 8x - y^2 + 9 \\ x^2 + y^2 &= 1 \end{aligned}$$

beschrieben wird. In welchem Punkt der Bahn ist die potentielle Energie des Teilchens am größten?



Die potentielle Energie wird mit der Formel

$$V = m \cdot g \cdot h, \quad g \text{ die Erdbeschleunigung,}$$

berechnet. Die Aufgabe besteht also darin, das globale Maximum der Funktion

$$\begin{aligned} h: \mathbb{R}^2 &\longrightarrow \mathbb{R} \\ (x, y) &\longmapsto -x^2 + 8x - y^2 + 9 \end{aligned}$$

unter der durch

$$\begin{aligned} g: \mathbb{R}^2 &\longrightarrow \mathbb{R} \\ (x, y) &\longmapsto x^2 + y^2 - 1 \end{aligned}$$

gestellten Nebenbedingung zu bestimmen. Wir setzen $U := \mathbb{R}^2 \setminus \{(0,0)\}$. Es gilt

$$\forall (x, y) \in U: \quad \text{Grad}g(x, y) = (2x, 2y) \neq (0, 0)$$

und

$$N(g) = S^1 \subset U.$$

Weiter haben wir

$$\text{Grad}h(x, y) = (-2x + 8, -2y).$$

Damit müssen wir reelle Zahlen a , b und λ ermitteln, die

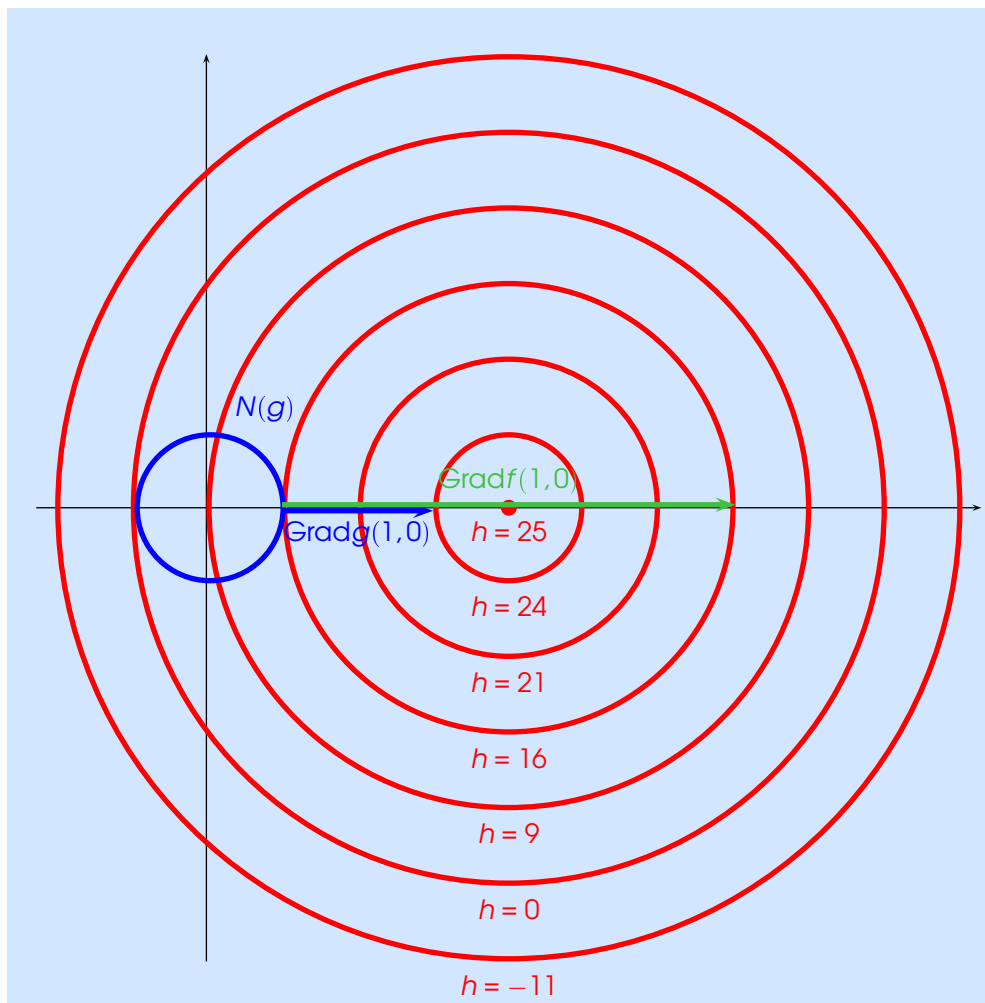
$$\begin{aligned} -2a + 8 &= 2\lambda a \\ -2b &= 2\lambda b \\ a^2 + b^2 &= 1 \end{aligned}$$

Kapitel 7 Extremwerte differenzierbarer Funktionen

erfüllen. Wenn $b \neq 0$ gilt, dann folgt aus der zweiten Gleichung $\lambda = -1$ und aus der ersten $8 = 0$. Deshalb finden wir $b = 0$. Die Gleichungen werden durch $a = 1, b = 0$ und $\lambda = 3$ bzw. $a = -1, b = 0$ und $\lambda = -5$ gelöst. Wir berechnen

$$\begin{aligned}h(1,0) &= 16 \\h(-1,0) &= 0.\end{aligned}$$

Da $N(g)$ kompakt ist, nimmt h auf $N(g)$ ein globales Minimum und ein globales Maximum an. Das globale Minimum liegt an der Stelle $(-1,0)$ und das gesuchte globale Maximum an der Stelle $(1,0)$ vor.



Kapitel 8

Der Banachsche Fixpunktsatz

8.1 Kontrahierende Abbildungen

Kontrahierende Abbildungen von einem metrischen Raum in sich selbst sind solche, die den Abstand zwischen Punkten verringern. Solche Abbildungen sind automatisch stetig. Diesen Begriff stellen wir nun genauer vor.

8.1.1 Definition. Es seien (X, d) ein metrischer Raum und $f: X \rightarrow X$ eine Abbildung. Die Abbildung f heißt *kontrahierend*, wenn es eine reelle Zahl $0 \leq c < 1$ mit der Eigenschaft

$$\forall x, y \in X: d(f(x), f(y)) \leq c \cdot d(x, y)$$

gibt.

8.1.2 Bemerkung. Eine kontrahierende Abbildung ist stetig. (Übung.)

8.1.3 Beispiele. i) Es sei $A: \mathbb{R}^n \rightarrow \mathbb{R}^n$ eine lineare Abbildung. In Bemerkung 3.7.2 haben wir die Operatornorm von A als

$$\|A\|_{\text{op}} := \max \left\{ \|A(x)\| \mid x \in \mathbb{R}^n : \|x\| = 1 \right\}$$

definiert. Die Abbildung A ist genau dann kontrahierend, wenn $\|A\|_{\text{op}} < 1$ gilt.

Für reelle Zahlen λ, μ , die Matrix

$$A := \begin{pmatrix} \lambda & 0 \\ 0 & \mu \end{pmatrix}$$

und $(x, y) \in \mathbb{R}^2$ gilt

$$\left\| A \cdot \begin{pmatrix} x \\ y \end{pmatrix} \right\| = \sqrt{\lambda^2 \cdot x^2 + \mu^2 \cdot y^2} \leq \max\{|\lambda|, |\mu|\} \cdot \|(x, y)\|.$$

Da weiter

$$\left\| A \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right\| = |\lambda| \quad \text{und} \quad \left\| A \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\| = |\mu|$$

ist $\|A\|_{\text{op}} = \max\{|\lambda|, |\mu|\}$, und die Abbildung A ist genau dann kontrahierend, wenn $|\lambda| < 1$ und $|\mu| < 1$.

ii) Es seien $a < b$ reelle Zahlen, $I = [a, b]$ und $f: I \rightarrow I$ eine differenzierbare Funktion. Für Zahlen $a \leq x < y \leq b$ gibt es nach dem Mittelwertsatz (12), Folgerung 4.3.5, eine Stelle $\xi \in (x, y)$, so dass

$$|f(y) - f(x)| = |f'(\xi)| \cdot |y - x|.$$

Wenn also $|f'(\xi)| < 1$ für alle $\xi \in (a, b)$ gilt, dann ist f kontrahierend.

Für

$$\begin{aligned} f: \left[0, \frac{\pi}{3}\right] &\longrightarrow [0, 1] \subset \left[0, \frac{\pi}{3}\right] \\ x &\longmapsto \cos(x) \end{aligned}$$

gilt $f'(x) = -\sin(x)$, $0 \leq x \leq \pi/3$. Für $\xi \in [0, \pi/3]$ hat man $|\sin(\xi)| \leq \sin(\pi/3) < 1$, so dass f eine kontrahierende Abbildung ist.

iii) Die Funktion

$$\begin{aligned} f: [0, 1] &\longrightarrow [0, 1] \\ x &\longmapsto \sqrt{x} \end{aligned}$$

ist **nicht** kontrahierend: Für $0 < \varepsilon < 1$ berechnet man $|f(\varepsilon^2) - f(0)| = \varepsilon > \varepsilon^2 = |\varepsilon^2 - 0|$.

iv) Auf der anderen Seite ist die Funktion

$$\begin{aligned} f: [1, 2] &\longrightarrow [1, 2] \\ x &\longmapsto \sqrt{x} \end{aligned}$$

kontrahierend. Dies folgt aus Teil ii) und

$$f'(x) = \frac{1}{2\sqrt{x}} \leq \frac{1}{2}, \quad x \in [1, 2].$$

8.2 Der Banachsche Fixpunktsatz

Der Banachsche Fixpunktsatz garantiert die Existenz eines Fixpunkts einer kontrahierenden Abbildung auf einem vollständigen metrischen Raum. Wir sind diesem Satz schon in (12), Satz 4.10.1, begegnet. Der Vorteil der allgemeineren Version, die wir im Folgenden erarbeiten werden, liegt darin, dass sie auch in einem unendlichdimensionalen Rahmen funktioniert. Im nächsten Kapitel werden wir den Fixpunktsatz in einem unendlichdimensionalen Banachraum anwenden und so ein hinreichendes Kriterium für die Lösbarkeit gewisser Differentialgleichungen beweisen.

8.2.1 Satz (Banachscher Fixpunktsatz). *Es seien (X, d) ein **vollständiger** metrischer Raum und $f: X \rightarrow X$ eine **kontrahierende** Abbildung. Dann gibt es genau einen Fixpunkt, d.h. einen Punkt $x \in X$ mit*

$$f(x) = x.$$

Angesichts Beispiels 8.1.3, ii), ist dieser Satz eine Verallgemeinerung von (12), Satz 4.10.1. Der Beweis verläuft völlig analog.

Beweis. i) Wir weisen zunächst die Eindeutigkeit nach. Sei $0 \leq c < 1$ eine reelle Zahl, so dass $d(f(y), f(y')) \leq c \cdot d(y, y')$. Für zwei Fixpunkte $x, x' \in X$ ergibt sich

$$d(x, x') = d(f(x), f(x')) \leq c \cdot d(x, x').$$

Die Ungleichung kann nur mit $d(x, x') = 0$ erfüllt werden, d.h. mit $x = x'$.

ii) Um die Existenz zu beweisen, wählen wir einen Punkt $x_0 \in X$ und definieren damit rekursiv die Folge $(x_k)_{k \in \mathbb{N}}$ durch

$$x_{k+1} := f(x_k), \quad k \in \mathbb{N}.$$

Zunächst beobachten wir

$$d(x_k, x_{k+1}) \leq c \cdot d(x_{k-1}, x_k) \leq c^k \cdot d(x_0, x_1)$$

für jede natürliche Zahl k . Seien weiter $l \geq m \geq 1$, i.e. $l = m + p$ mit $p \geq 0$. Wir finden

$$\begin{aligned} d(x_m, x_l) &\leq \sum_{i=0}^{p-1} d(x_{m+i}, x_{m+i+1}) \\ &\leq d(x_0, x_1) \cdot \sum_{i=0}^{p-1} c^{m+i} = d(x_0, x_1) \cdot c^m \cdot \sum_{i=0}^{p-1} c^i \\ &= \frac{1-c^p}{1-c} \cdot c^m \cdot d(x_0, x_1) \leq \frac{c^m}{1-c} \cdot d(x_0, x_1). \end{aligned}$$

Falls $d(x_0, x_1) = 0$ gilt, ist die Folge konstant. Andernfalls können wir wegen $0 \leq c < 1$ zu einem vorgegebenen $\varepsilon > 0$ ein k_0 finden, so dass

$$\forall k \geq k_0 : \frac{c^k}{1-c} < \frac{\varepsilon}{d(x_0, x_1)}.$$

Für $l, m \geq k_0$ ergibt sich aus unseren Betrachtungen, dass

$$d(x_m, x_l) < \varepsilon.$$

Bei der Folge $(x_k)_{k \in \mathbb{N}}$ handelt es sich um eine Cauchy-Folge. Da der metrische Raum (X, d) als vollständig vorausgesetzt wurde, konvergiert sie. Sei x ihr Grenzwert. Nach Bemerkung 8.1.2 ist f stetig. Wir schließen

$$f(x) = \lim_{k \rightarrow \infty} f(x_k) = \lim_{k \rightarrow \infty} x_{k+1} = x.$$

Daher ist x der gesuchte Fixpunkt. \square

8.2.2 Bemerkung. i) Für die Folge $(x_k)_{k \in \mathbb{N}}$ und ihren Grenzwert haben wir die Abschätzung

$$d(x_k, x) = \lim_{l \rightarrow \infty} d(x_k, x_l) \leq \frac{c^k}{1-c} \cdot d(x_0, x_1).$$

Sie gibt uns Auskunft darüber, wie gut das k -te Folgenglied x_k den Fixpunkt x approximiert.

ii) Jede lineare Abbildung hat den Fixpunkt 0. Eine kontrahierende lineare Abbildung auf einem Banachraum hat also den Ursprung als einzigen Fixpunkt.

8.3 Nullstellen

Mit dem Banachschen Fixpunktsatz können wir ein zu (12), Satz 4.10.4, analoges Iterationsverfahren zur Lösung gewisser Gleichungssysteme entwickeln.

Es seien $B \subset \mathbb{R}^n$ eine offene Teilmenge und $f: B \rightarrow \mathbb{R}^n$ eine Abbildung. Wir suchen einen Punkt $x \in B$ mit $f(x) = 0$.

Wir nehmen eine **invertierbare** lineare Abbildung $A: \mathbb{R}^n \rightarrow \mathbb{R}^n$ zu Hilfe und erklären die Abbildung

$$\begin{aligned} g: \mathbb{R}^n &\rightarrow \mathbb{R}^n \\ x &\mapsto x - A(f(x)). \end{aligned}$$

Ein Punkt $x \in B$ erfüllt genau dann die Gleichung $f(x) = 0$, wenn $g(x) = x$ gilt.

8.3.1 Satz. In der obigen Situation seien die folgenden Voraussetzungen erfüllt:

- Die Funktion f ist differenzierbar.
- Es gibt ein $r > 0$, ein $c \in [0, 1)$ und ein $a \in B$, so dass
 - ★ $M := \overline{B(a, r)} \subset B$,
 - ★ $\forall x \in M: \quad \|\text{id}_{\mathbb{R}^n} - A \circ Df(x)\|_{\text{op}} \leq c$,
 - ★ $\|f(a)\| \leq \frac{r \cdot (1 - c)}{\|A\|_{\text{op}}}$.

Dann hat f genau eine Nullstelle x auf der Menge M und für jede Folge $(x_k)_{k \in \mathbb{N}}$ mit

- $x_0 \in M$,
- $x_{k+1} = f(x_k)$, $k \geq 1$,

gilt

$$x = \lim_{k \rightarrow \infty} x_k.$$

Dabei hat man die Fehlerabschätzung

$$\|x - x_k\| \leq \frac{c^k}{1 - c} \cdot \|A\|_{\text{op}} \cdot \|f(x_0)\|, \quad k \in \mathbb{N}.$$

Im Beweis des obigen Satzes benötigen wir eine Verallgemeinerung des Mittelwertsatzes der Differentialrechnung ((12), Folgerung 4.3.5) auf vektorwertige Abbildungen in mehreren Variablen.

8.3.2 Satz (Mittelwertsatz). Es seien $B \subset \mathbb{R}^m$ eine offene Teilmenge, $x, y \in B$ Punkte, so dass die Verbindungsstrecke

$$L := \{x + \lambda \cdot (y - x) \mid \lambda \in [0, 1]\}$$

ganz in B liegt, und $f: B \rightarrow \mathbb{R}^n$ eine differenzierbare Abbildung. Dann existiert eine Zahl $\xi \in (0, 1)$, so dass

$$\|f(y) - f(x)\| \leq \|Df(x + \xi \cdot (y - x))\|_{\text{op}} \cdot \|y - x\|.$$

Beweis. Wir nehmen $f(x) \neq f(y)$ an, weil die Behauptung sonst trivial ist. Es sei $v \in \mathbb{R}^n \setminus \{0\}$ ein Vektor. Wir definieren die Funktion

$$g: [0, 1] \rightarrow \mathbb{R} \\ \lambda \mapsto \langle f(x + \lambda \cdot (y - x)), v \rangle.$$

Aus der Voraussetzung, dass f differenzierbar ist, folgt leicht, dass g differenzierbar ist mit

$$g'(\lambda) = \langle Df(x + \lambda \cdot (y - x))(y - x), v \rangle, \quad \lambda \in [0, 1].$$

Mit dem Mittelwertsatz für differenzierbare Funktionen in einer Veränderlichen ((12), Folgerung 4.3.5) finden wir ein $\xi \in (0, 1)$, so dass

$$\langle f(y) - f(x), v \rangle = g(1) - g(0) = g'(\xi) = \langle Df(x + \xi \cdot (y - x))(y - x), v \rangle. \quad (8.1)$$

Nun wählen wir

$$v := \frac{f(y) - f(x)}{\|f(y) - f(x)\|}.$$

Dieser Vektor hat Norm 1. Einsetzen dieses Vektors in (8.1) führt zu der Gleichung

$$\|f(y) - f(x)\| = \left| \langle Df(x + \xi \cdot (y - x))(y - x), v \rangle \right|.$$

Mit der Cauchy–Schwarz-Ungleichung und $\|v\| = 1$ schließen wir

$$\begin{aligned} \left| \langle Df(x + \xi \cdot (y - x))(y - x), v \rangle \right| &\leq \left\| Df(x + \xi \cdot (y - x))(y - x) \right\| \\ &\leq \left\| Df(x + \xi \cdot (y - x)) \right\|_{\text{op}} \cdot \|y - x\|. \end{aligned}$$

Alles zusammen liefert die behauptete Ungleichung. \square

8.3.3 Bemerkung. Falls $n > 1$ gilt, ist es im Allgemeinen nicht wahr, dass es einen Punkt $\xi \in [0, 1]$ gibt, so dass

$$f(y) - f(x) = Df(x + \xi \cdot (x - y))(y - x).$$

Man betrachte dazu das Beispiel aus Aufgabe A.11.1. Für den Fall $n = 1$ haben wir Satz 7.5.5. Interessant ist auch Satz 12.19 aus (4).

Beweis von Satz 8.3.1. Die Menge M ist konvex. Es seien $y, z \in M$. Nach dem Mittelwertsatz 8.3.2 existiert ein $\xi \in (0, 1)$, so dass

$$\|g(z) - g(y)\| \leq \left\| Dg(y + \xi \cdot (z - y)) \right\|_{\text{op}} \cdot \|z - y\|.$$

Nun gilt nach Definition von g und der Kettenregel 6.3.1, dass

$$Dg(y + \xi \cdot (z - y)) = \text{id}_{\mathbb{R}^n} - A \circ Df(y + \xi \cdot (z - y)).$$

Die Voraussetzung impliziert daher

$$\left\| Dg(y + \xi \cdot (z - y)) \right\|_{\text{op}} \leq c \quad \text{und} \quad \|g(z) - g(y)\| \leq c \cdot \|z - y\|.$$

Sei nun $y \in M$, d.h. $\|y - a\| \leq r$. Mit der eben hergeleiteten Abschätzung folgt

$$\|g(y) - g(a)\| \leq c \cdot \|y - a\| \leq c \cdot r.$$

Nun berechnen wir weiter

$$\begin{aligned} \|g(y) - a\| &\leq \|g(y) - g(a)\| + \|g(a) - a\| \leq c \cdot \|y - a\| + \|(A \circ f)(a)\| \\ &\leq c \cdot \|y - a\| + \|A\|_{\text{op}} \cdot \|f(a)\| \leq c \cdot r + r \cdot (1 - c) = r. \end{aligned}$$

Damit ist $g: M \rightarrow M$ eine kontrahierende Abbildung. Da \mathbb{R}^n vollständig und $M \subset \mathbb{R}^n$ eine kompakte Teilmenge ist, ist M mit der durch die euklidische Metrik induzierten Metrik ein vollständiger metrischer Raum (s. Kapitel 2).

Die Fehlerabschätzung ergibt sich aus der im Banachschen Fixpunktsatz (Bemerkung 8.2.2):

$$\|x - x_k\| \leq \frac{c^k}{1 - c} \cdot \|g(x_0) - x_0\| \leq \frac{c^k}{1 - c} \cdot \|A\|_{\text{op}} \cdot \|f(x_0)\|, \quad k \in \mathbb{N}.$$

Damit ist der Beweis beendet. \square

8.3.4 Bemerkung (Operatornorm). In Beispiel 3.7.3 haben wir eine Abschätzung für die Operatornorm einer Matrix bzgl. der Maximum-Norm ausgerechnet. Jetzt wollen wir sie bzgl. der euklidischen Norm abschätzen. Wir behaupten, dass für eine $(n \times n)$ -Matrix $A = (a_{ij})_{i,j=1,\dots,n}$ und die lineare Abbildung

$$\begin{aligned} L_A: \mathbb{R}^n &\longrightarrow \mathbb{R}^n \\ x &\longmapsto A \cdot x^t \end{aligned}$$

die Ungleichung

$$\|A\|_{\text{op}} \leq \|A\| = \sqrt{\sum_{i,j=1}^n a_{ij}^2}$$

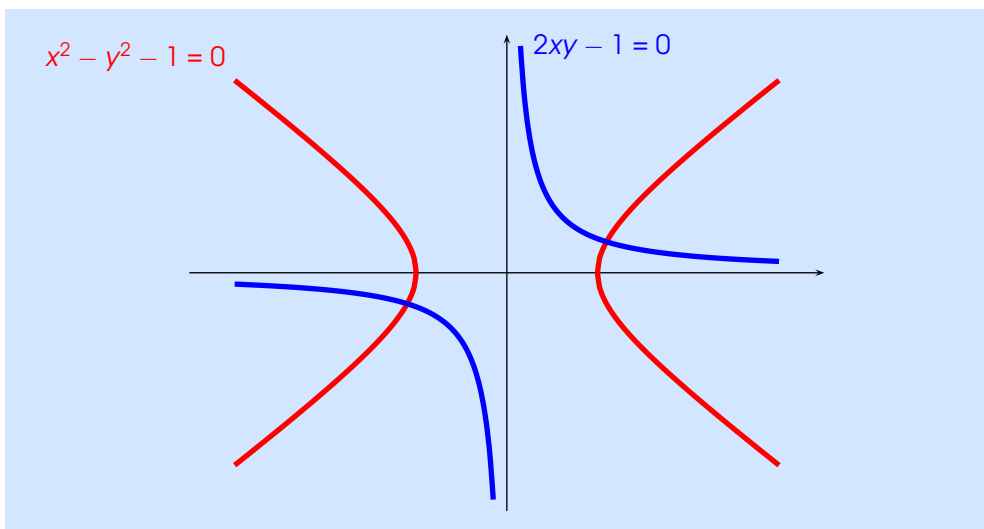
erfüllt ist.

Zum Beweis definieren wir $z_i(A) := (a_{i1}, \dots, a_{in})$ als die i -te Zeile von A , $i = 1, \dots, n$. Jetzt gilt mit der Cauchy-Schwarz-Ungleichung

$$\begin{aligned} \|A\|_{\text{op}}^2 &= \sup \{ \|A \cdot x^t\|^2 \mid \|x\| = 1 \} \\ &= \sup \left\{ \sum_{i=1}^n \langle z_i(A), x \rangle^2 \mid \|x\| = 1 \right\} \\ &\leq \sup \left\{ \sum_{i=1}^n \|z_i(A)\|^2 \cdot \|x\|^2 \mid \|x\| = 1 \right\} \\ &= \sum_{i=1}^n \sum_{j=1}^n a_{ij}^2 = \|A\|^2. \end{aligned}$$

8.3.5 Beispiel. Gesucht ist eine Lösung des Gleichungssystems

$$\begin{aligned} \text{(I)} \quad x^2 - y^2 - 1 &= 0 \\ \text{(II)} \quad 2xy - 1 &= 0. \end{aligned}$$



Wir beschreiten zunächst den elementaren Weg. Die Gleichung (II) impliziert $x \neq 0$ und $y \neq 0$, und wir können sie zu

$$x = \frac{1}{2y}$$

umschreiben. Diesen Ausdruck setzen wir in (I) ein:

$$\frac{1}{4y^2} - y^2 - 1 = 0 \iff 4y^4 + 4y^2 - 1 = 0 \iff y^2 = \frac{-1 \pm \sqrt{2}}{2}.$$

Nur einer der beiden Ausdrücke ist positiv, so dass

$$y^2 = \frac{\sqrt{2} - 1}{2}$$

folgt, d.h.

$$y = \pm \sqrt{\frac{\sqrt{2} - 1}{2}} \approx \pm 0,4550898604.$$

Man erhält

$$x = \pm \sqrt{\frac{\sqrt{2} + 1}{2}} \approx \pm 1,098684114.$$

Alternativ möchten wir das oben beschriebene Iterationsverfahren auf die Abbildung

$$\begin{aligned} f: \mathbb{R}^2 &\longrightarrow \mathbb{R}^2 \\ (x, y) &\longmapsto (x^2 - y^2 - 1, 2xy - 1) \end{aligned}$$

anwenden. Wir arbeiten mit dem Startwert

$$\alpha = \left(1, \frac{1}{2}\right).$$

Es folgt

$$f(\alpha) = \left(-\frac{1}{4}, 0\right), \text{ so dass } \|f(\alpha)\| = \frac{1}{4}.$$

Für die totale Ableitung berechnen wir

$$Df(x, y) = \begin{pmatrix} 2x & -2y \\ 2y & 2x \end{pmatrix} \text{ und } Df\left(1, \frac{1}{2}\right) = \begin{pmatrix} 2 & -1 \\ 1 & 2 \end{pmatrix}.$$

Wir setzen¹

$$A := Df\left(1, \frac{1}{2}\right)^{-1} = \begin{pmatrix} \frac{2}{5} & \frac{1}{5} \\ -\frac{1}{5} & \frac{2}{5} \end{pmatrix}$$

an. Zur Bestimmung der Operatornorm von A berechnen wir für $(x, y) \in \mathbb{R}^2$:

$$\begin{aligned} \left\| A \cdot \begin{pmatrix} x \\ y \end{pmatrix} \right\|^2 &= \frac{1}{25} \cdot \|(2x + y, -x + 2y)\|^2 \\ &= \frac{1}{25} \cdot (4x^2 + 4xy + y^2 + x^2 - 4xy + 4y^2) \\ &= \frac{1}{25} \cdot (5x^2 + 5y^2) = \frac{1}{5} \cdot \|(x, y)\|^2. \end{aligned}$$

Es folgt

$$\|A\|_{\text{op}} = \sqrt{\frac{1}{5}}.$$

¹Diese Wahl entspricht der Wahl im Newtonverfahren in einer Variablen ((12), Satz 4.10.4).

Jetzt wollen wir die Operatornorm von

$$\text{id}_{\mathbb{R}^2} - A \cdot Df(x, y), \quad \text{also von} \quad \begin{pmatrix} 1 - \frac{4}{5} \cdot x - \frac{2}{5} \cdot y & -\frac{2}{5} \cdot x + \frac{4}{5} \cdot y \\ \frac{2}{5} \cdot x - \frac{4}{5} \cdot y & 1 - \frac{4}{5} \cdot x - \frac{2}{5} \cdot y \end{pmatrix} =: B,$$

abschätzen. Nach Bemerkung 8.3.4 genügt es, die euklidische Norm dieser Matrix zu berechnen. Die geforderte Ungleichung

$$\|B\| \leq c$$

wird zu

$$(x-1)^2 + \left(y - \frac{1}{2}\right)^2 \leq \frac{5}{8} \cdot c^2.$$

Die Ungleichung bestimmt eine Scheibe vom Radius $r = \sqrt{5/8} \cdot c$ um $(1, 1/2)$. Ein Wert für c ist geeignet, wenn

$$\|f(a)\| \leq \frac{r \cdot (1-c)}{\|A\|_{\text{op}}}$$

erfüllt ist. Diese Ungleichung lautet

$$\frac{1}{4} \leq \frac{\sqrt{\frac{5}{8}} \cdot c \cdot (1-c)}{\sqrt{\frac{1}{5}}} \iff \frac{\sqrt{2}}{10} \leq c \cdot (1-c).$$

Wir können z.B. $c = 1/5$ wählen und erhalten $r = 1/(2\sqrt{10})$. Unsere Diskussion zeigt, dass die Abbildung

$$g: M \rightarrow M \\ (x, y) \mapsto \left(\frac{3}{5} + x - \frac{2}{5} \cdot x^2 + \frac{2}{5} \cdot y^2 - \frac{2}{5} \cdot xy, \frac{1}{5} + y + \frac{1}{5} \cdot x^2 - \frac{1}{5} \cdot y^2 - \frac{4}{5} \cdot xy \right)$$

mit

$$M := \left\{ (x, y) \in \mathbb{R}^2 \mid \left\| (x, y) - \left(1, \frac{1}{2}\right) \right\| \leq \frac{1}{2\sqrt{10}} \right\}$$

kontrahierend ist. Wie zuvor folgt, dass M mit der durch die euklidische Metrik induzierten Metrik vollständig ist. Schließlich bemerken wir, dass der Term in der Fehlerabschätzung

$$\frac{\sqrt{5}}{16} \cdot \left(\frac{1}{5}\right)^k, \quad k \in \mathbb{N},$$

ist. Die ersten Iterationsschritte lauten:

$$\begin{array}{ll} x_0 = 1 & y_0 = 0,5 \\ x_1 = 1,1 & y_1 = 0,45 \\ x_2 = 1,099 & y_2 = 0,4555 \\ x_3 = 1,0986339 & y_3 = 0,45508855 \\ x_4 \approx 1,098687273 & y_4 \approx 0,4550861548. \end{array}$$

8.3.6 Bemerkung. Natürlich können wir mit dem gegebenen Verfahren zu einem gegebenen Startwert nur eine der beiden Lösungen des Gleichungssystems finden.

Kapitel 9

Gewöhnliche Differentialgleichungen

9.1 Der Satz von Picard–Lindelöf

Wir beginnen nun mit dem Studium von sogenannten gewöhnlichen Differentialgleichungen. Dies sind Gleichungen, in denen eine Funktion und ihre Ableitungen vorkommen. Die Aufgabe besteht darin, die Funktion aus dieser Gleichung und einer sogenannten Anfangsbedingung $f(t_0) = x_0$ zu bestimmen.

Wir sind bereits einfachen Beispielen begegnet: Die Gleichung $f'(t) = 0$, $t \in \mathbb{R}$, hat nur konstante Funktionen als Lösung ((12), Folgerung 4.3.7). Fixieren wir zusätzlich $t_0 \in \mathbb{R}$ und $x_0 \in \mathbb{R}$, dann ist f durch die Bedingung $f(t_0) = x_0$ festgelegt, nämlich $f(t) = x_0$ für alle $t \in \mathbb{R}$. Für $c \in \mathbb{R}^*$ wird die Differentialgleichung $f'(t) = c \cdot f(t)$ durch $t \mapsto k \cdot \exp(c \cdot t)$, $t \in \mathbb{R}$, $k \in \mathbb{R}$, gelöst. Stellen wir z.B. die Anfangsbedingung $f(0) = x_0$ für ein $x_0 \in \mathbb{R}$, dann muss $k = x_0$ gelten ((12), Satz 4.8.5).

Differentialgleichungen beschreiben viele Vorgänge in Natur und Technik. Ihr Studium verbindet daher Theorie und Praxis. Im Folgenden werden wir einige Beispiele dazu kennenlernen.

In diesem Abschnitt beweisen wir ein theoretisches Resultat, das die Existenz und Eindeutigkeit von Lösungen gewisser Differentialgleichungen garantiert. Der Beweis stützt sich auf den Banachschen Fixpunktsatz. Er stellt damit eine interessante Verbindung zwischen der Analysis von reellwertigen Funktionen einer Veränderlichen und der Analysis in unendlichdimensionalen Vektorräumen her.

9.1.1 Definition. a) Ein Ausdruck der Form¹

$$x'(t) = F(t, x(t))$$

heißt *System gewöhnlicher Differentialgleichungen erster Ordnung*. Dabei seien $U \subset \mathbb{R} \times \mathbb{R}^n$ eine offene Teilmenge oder $U = I \times \mathbb{R}^n$, $I = [c, d] \subset \mathbb{R}$ ein kompaktes Intervall, und $F: U \rightarrow \mathbb{R}^n$ eine stetige Abbildung.

b) Seien weiter $(t_0, x_0) \in U$ und $I' \subset \mathbb{R}$ ein Intervall. Eine differenzierbare Abbildung²

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

¹Wie in Kapitel 4 steht $x'(t)$ für den Vektor $(x_1'(t), \dots, x_n'(t))$, $x(t) = (x_1(t), \dots, x_n(t))$.

²Das Intervall I' muss nicht offen sein. Wir erinnern an die Konvention auf Seite 40.

ist eine Lösung des Systems

$$x'(t) = F(t, x(t))$$

gewöhnlicher Differentialgleichungen erster Ordnung zur Anfangsbedingung

$$x(t_0) = x_0,$$

wenn die Bedingungen

$$(t, f(t)) \in U, \quad f'(t) = F(t, f(t)), \quad t \in I,$$

und

$$f(t_0) = x_0$$

erfüllt sind.

c) Die Funktion F erfüllt die *Lipschitz³-Bedingung zur Lipschitz-Konstanten* $L \in \mathbb{R}_{>0}$, wenn

$$\forall t \in I \forall x, y \in \mathbb{R}: \quad \|F(t, y) - F(t, x)\| \leq L \cdot \|y - x\|.$$

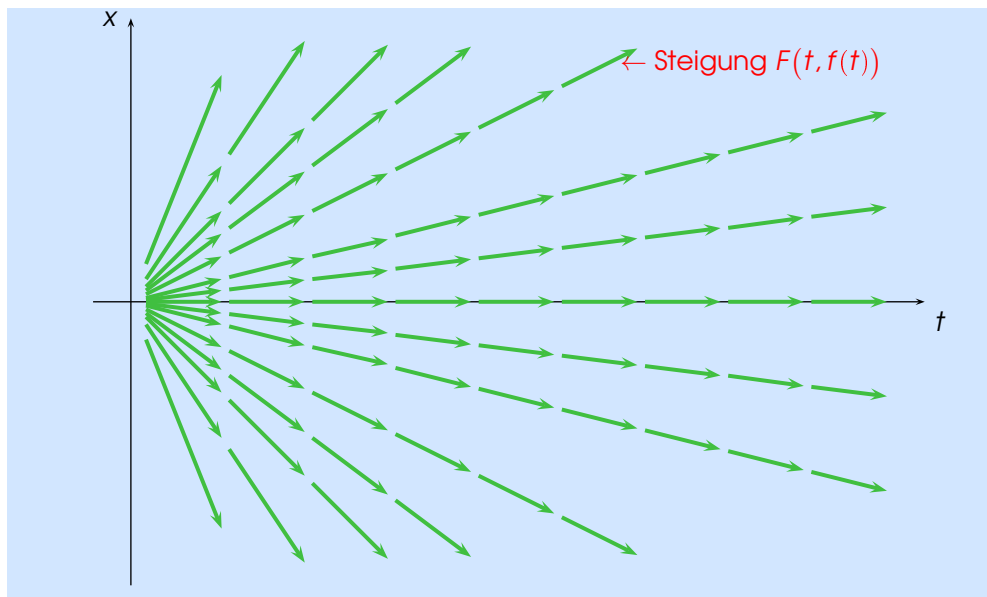
9.1.2 Bemerkung. Wir können uns die Funktion F als **Geschwindigkeitsfeld** vorstellen: Wir interpretieren die erste Variable als Zeit und die restlichen Variablen als Koordinaten des Raums. Die Funktion F ordnet jedem Zeitpunkt t und jedem Raumpunkt x den Geschwindigkeitsvektor $F(t, x)$ zu. Eine Lösung $f: I \rightarrow \mathbb{R}^n$ der Differentialgleichung beschreibt die Bewegung eines Teilchens in \mathbb{R}^n . Die Anfangsbedingung sagt uns, wo sich das Teilchen zum Zeitpunkt t_0 befindet. Die Bewegung des Teilchens soll so verlaufen, dass es zu jedem Zeitpunkt $t \in I$ die Geschwindigkeit $F(t, f(t))$ hat, die durch die Funktion F vorgegeben ist. Dabei ist $f(t)$ die Position des Teilchens zum Zeitpunkt t . Beispiele für diesen Formalismus gibt es viele, z.B. die Feldlinien eines elektrischen Feldes.

9.1.3 Beispiel. Es seien $U = \mathbb{R}_{>0} \times \mathbb{R}$ und

$$\begin{aligned} F: U &\longrightarrow \mathbb{R} \\ (t, x) &\longmapsto \frac{x}{t}. \end{aligned}$$

Das zugehörige Geschwindigkeitsfeld können wir veranschaulichen, indem wir an jedem Punkt $(t, x) \in U$ einen „Geschwindigkeitspfeil“ der Steigung $F(t, x)$ „anbringen“.

³Rudolf Otto Sigismund Lipschitz (1832 - 1903), deutscher Mathematiker.



Für einen Zeitpunkt $t_0 > 0$ stellen wir die Anfangsbedingung $x(t_0) := x_0$ und setzen $c := x_0/t_0$. Die Funktion

$$\begin{aligned} f: I &\longrightarrow \mathbb{R} \\ t &\longmapsto c \cdot t \end{aligned}$$

löst die durch F gegebene Differentialgleichung zu der gestellten Anfangsbedingung.

9.1.4 Satz (Picard⁴–Lindelöf⁵). *Es seien $a < b$ reelle Zahlen, $I := [a, b]$, $F: I \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ eine stetige Funktion und $L \in \mathbb{R}_{>0}$, so dass F die Lipschitz-Bedingung zur Lipschitz-Konstanten L erfüllt. Dann gibt es zu jedem $t_0 \in I$ und jedem $x_0 \in \mathbb{R}^n$ genau eine **stetig differenzierbare** Funktion $f: I \rightarrow \mathbb{R}^n$ mit*

- $f'(t) = F(t, f(t)), t \in I,$
- $f(t_0) = x_0.$

Der Beweis verwendet den Banachschen Fixpunktsatz. Wir müssen zuerst den relevanten vollständigen metrischen Raum vorstellen. Es sei $I \subset \mathbb{R}$ ein abgeschlossenes Intervall wie im Satz von Picard–Lindelöf. Dazu sei

$$\mathcal{C}^0(I, \mathbb{R}^n) := \{ f: I \rightarrow \mathbb{R}^n \mid f \text{ ist stetig} \}.$$

Dies ist ein reeller Vektorraum. Wir definieren

$$\begin{aligned} \|\cdot\|_{\text{sup}}: \mathcal{C}^0(I, \mathbb{R}^n) &\longrightarrow \mathbb{R} \\ f &\longmapsto \|f\|_{\text{sup}} := \sup \{ \|f(t)\| \mid t \in I \}. \end{aligned}$$

Wir behaupten, dass $\|\cdot\|_{\text{sup}}$ eine Norm auf $\mathcal{C}^0(I, \mathbb{R}^n)$ ist und dass der normierte Vektorraum $(\mathcal{C}^0(I, \mathbb{R}^n), \|\cdot\|_{\text{sup}})$ ein Banachraum ist. Für $n = 1$ wurde dies in Satz 3.6.5 nachgewiesen. Für $n > 1$ folgt dies mit den üblichen Argumenten (z.B. Satz 2.1.3). Mit $d: \mathcal{C}^0(I, \mathbb{R}^n) \times \mathcal{C}^0(I, \mathbb{R}^n) \rightarrow \mathbb{R}, (f, g) \mapsto \|f - g\|_{\text{sup}}$ erhalten wir den vollständigen metrischen Raum $(\mathcal{C}^0(I, \mathbb{R}^n), d)$.

⁴Charles Émile Picard (1856 - 1941), französischer Mathematiker.

⁵Ernst Leonard Lindelöf (1870 - 1946), finnischer Mathematiker.

Im Beweis des Satzes von Picard–Lindelöf benötigen wir noch eine Konvention für Integrale. Für eine stetige Funktion $g = (g_1, \dots, g_n) : I \rightarrow \mathbb{R}^n$ definieren wir

$$\int_a^b g(x) dx := \left(\int_a^b g_1(x) dx, \dots, \int_a^b g_n(x) dx \right) \in \mathbb{R}^n.$$

Beweis von Satz 9.1.4. Wir nehmen an, $f : I \rightarrow \mathbb{R}^n$ sei eine stetig differenzierbare Abbildung, für die

$$f'(t) = F(t, f(t)), \quad t \in I, \quad \text{und} \quad f(t_0) = x_0$$

gilt. Da die auftretenden Funktionen stetig sind, können wir integrieren und finden mit dem Hauptsatz der Differential- und Integralrechnung ((12), Satz 5.3.5)

$$f(t) - x_0 = f(t) - f(t_0) = \int_{t_0}^t f'(\tau) d\tau = \int_{t_0}^t F(\tau, f(\tau)) d\tau,$$

also

$$f(t) = x_0 + \int_{t_0}^t F(\tau, f(\tau)) d\tau.$$

Dies motiviert die Definition der Abbildung

$$\phi : \mathcal{C}^0(I, \mathbb{R}^n) \rightarrow \mathcal{C}^0(I, \mathbb{R}^n)$$

über

$$f \in \mathcal{C}^0(I, \mathbb{R}^n) \mapsto \begin{cases} \phi(f) : I \rightarrow \mathbb{R}^n \\ t \mapsto x_0 + \int_{t_0}^t F(\tau, f(\tau)) d\tau \end{cases}.$$

Es gilt offenbar

$$\phi(f)(t_0) = x_0,$$

und nach dem Hauptsatz der Differential- und Integralrechnung

$$\phi(f)'(t) = F(t, f(t)), \quad t \in I.$$

Eine Funktion $f \in \mathcal{C}^0(I, \mathbb{R}^n)$ löst mithin genau dann die Differentialgleichung $x'(t) = F(x, x(t))$ zur Anfangsbedingung $x(t_0) = x_0$, wenn

$$\phi(f) = f$$

gilt. Der Satz ist somit bewiesen, wenn wir zeigen, dass ϕ eine kontrahierende Abbildung ist.

Es seien $f, g \in \mathcal{C}^0(I, \mathbb{R}^n)$. Wir schreiben $F = (F_1, \dots, F_n)$ mit $F_i : I \times \mathbb{R}^n \rightarrow \mathbb{R}$, $i = 1, \dots, n$. Dann gilt für $t \in [a, b]$:

$$\begin{aligned} & \|\phi(f)(t) - \phi(g)(t)\| \\ &= \left\| \int_{t_0}^t (F(\tau, f(\tau)) - F(\tau, g(\tau))) d\tau \right\| \\ &\leq \sqrt{n} \cdot \max \left\{ \left| \int_{t_0}^t (F_i(\tau, f(\tau)) - F_i(\tau, g(\tau))) d\tau \right| \mid i = 1, \dots, n \right\} \\ &\leq \sqrt{n} \cdot \max \left\{ \int_{t_0}^t |F_i(\tau, f(\tau)) - F_i(\tau, g(\tau))| d\tau \mid i = 1, \dots, n \right\} \\ &\leq \sqrt{n} \cdot \int_{t_0}^t \|F(\tau, f(\tau)) - F(\tau, g(\tau))\| d\tau. \end{aligned}$$

Nach Voraussetzung gilt

$$\|F(\tau, f(\tau)) - F(\tau, g(\tau))\| \leq L \cdot \|f(\tau) - g(\tau)\| \leq L \cdot \|f - g\|_{\text{sup}}, \quad \tau \in [a, b].$$

Damit haben wir

$$\begin{aligned} \|\phi(f) - \phi(g)\|_{\text{sup}} &= \sup \left\{ \|\phi(f)(t) - \phi(g)(t)\| \mid t \in I \right\} \\ &\leq \sqrt{n} \cdot L \cdot |t - t_0| \cdot \|f - g\|_{\text{sup}} \leq \sqrt{n} \cdot L \cdot (b - a) \cdot \|f - g\|_{\text{sup}} \end{aligned}$$

gezeigt. Die Abbildung ϕ ist also kontrahierend, wenn

$$b - a < \frac{1}{\sqrt{n} \cdot L}. \tag{9.1}$$

Wir behaupten, dass wir diese Bedingung ohne Beschränkung der Allgemeinheit annehmen können. Dazu überdecken wir I durch endlich viele Teilintervalle J_1, \dots, J_m der Länge $c < 1/(\sqrt{n} \cdot L)$. Ein Intervall J_0 enthält t_0 . Auf J_0 können wir die Differentialgleichung eindeutig lösen, d.h. wir finden eine stetig differenzierbare Funktion $f_0: J_0 \rightarrow \mathbb{R}^n$ mit $f_0'(t) = F(t, f_0(t))$, $t \in J_0$, und $f_0(t_0) = x_0$. Es gibt einen Index $i_1 \in \{1, \dots, m\}$ mit $J_0 \cap J_{i_1} \neq \emptyset$. Sei $t_1 \in J_0 \cap J_{i_1}$. Wir finden eine stetig differenzierbare Funktion $\tilde{f}_1: J_{i_1} \rightarrow \mathbb{R}^n$ mit $\tilde{f}_1'(t) = F(t, \tilde{f}_1(t))$, $t \in J_{i_1}$, und $\tilde{f}_1(t_1) = f_0(t_1)$. Auf Grund der Eindeutigkeitsaussage stimmen \tilde{f}_1 und f_0 auf $J_0 \cap J_{i_1}$ überein und definieren eine stetig differenzierbare Funktion $f_1: J_0 \cup J_{i_1} \rightarrow \mathbb{R}^n$ mit $f_1'(t) = F(t, f_1(t))$, $t \in J_0 \cup J_{i_1}$, und $f_1(t_0) = x_0$. Nach endlich vielen Schritten erhalten wir die stetig differenzierbare Funktion $f: I \rightarrow \mathbb{R}^n$ mit $f(t) = F(t, f(t))$, $t \in I$, und $f(t_0) = x_0$. \square

9.1.5 Beispiel. Es sei

$$\begin{aligned} F: \left[0, \frac{1}{2}\right] \times \mathbb{R} &\longrightarrow \mathbb{R} \\ (t, x) &\longmapsto 2tx. \end{aligned}$$

Diese Funktion erfüllt die Lipschitz-Bedingung zu der Lipschitz-Konstanten $L = 1$:

$$\forall t \in \left[0, \frac{1}{2}\right] \forall x, y \in \mathbb{R}: |F(t, x) - F(t, y)| = 2t \cdot |x - y| \leq |x - y|.$$

Wir fixieren $c \in \mathbb{R}$ und stellen die Anfangsbedingung $f(0) = c$. Abschätzung 9.1 zeigt, dass $\phi: \mathcal{C}^0([0, 1/2], \mathbb{R}) \rightarrow \mathcal{C}^0([0, 1/2], \mathbb{R})$ kontrahierend ist. Die Lösung des Anfangswertproblems ist gemäß dem Beweis des Banachschen Fixpunktsatzes 8.2.1 der Grenzwert f der Funktionenfolge $(f_k)_{k \in \mathbb{N}}$ mit $f_0 = c$ und

$$\begin{aligned} f_{k+1} &:= \phi(f_k): \left[0, \frac{1}{2}\right] \longrightarrow \mathbb{R} \\ t &\longmapsto c + 2 \cdot \int_0^t \tau \cdot f_k(\tau) d\tau, \quad k \in \mathbb{N}. \end{aligned}$$

Diese Iteration liefert

$$\begin{aligned} f_1(t) &= c + 2c \cdot \int_0^t \tau d\tau = c \cdot (1 + t^2), \\ f_2(t) &= c + 2c \cdot \int_0^t (\tau + \tau^3) d\tau = c \cdot \left(1 + t^2 + \frac{t^4}{2}\right) \\ &\vdots \\ f_k(t) &= c \cdot \sum_{i=0}^k \frac{t^{2i}}{i!} \\ &\vdots \end{aligned}$$

Somit gilt

$$f(t) = \lim_{k \rightarrow \infty} f_k(t) = c \cdot \sum_{i=0}^{\infty} \frac{t^{2i}}{i!} = c \cdot \exp(t^2), \quad t \in \left[0, \frac{1}{2}\right].$$

Die Kontrolle zeigt:

$$\begin{aligned} f(0) &= c \cdot \exp(0) = c \\ f'(t) &= c \cdot 2t \cdot \exp(t^2) = 2t \cdot f(t) = F(t, f(t)). \end{aligned}$$

9.1.6 *Bemerkung.* Es seien $a < b$ reelle Zahlen und $I := [a, b]$. Die Funktion

$$\begin{aligned} F: I \times \mathbb{R} &\longrightarrow \mathbb{R} \\ (t, x) &\longmapsto x^2 \end{aligned}$$

erfüllt **nicht** die Lipschitz-Bedingung, so dass der Satz von Picard–Lindelöf in der obigen Version nicht anwendbar ist. Allerdings erfüllt die Funktion $F_{I \times J}: I \times J \rightarrow \mathbb{R}$ für jedes **beschränkte** Intervall $J \subset \mathbb{R}$ die Lipschitz-Bedingung. (Das ist ein Fall, in dem F **lokal einer Lipschitz-Bedingung genügt**.) Man kann das Verfahren von Picard–Lindelöf so modifizieren, dass man zu gegebenem $t_0 \in I$ und $x_0 \in \mathbb{R}$ ein $\varepsilon > 0$ mit $(t_0 - \varepsilon, t_0 + \varepsilon) \subset I$ und eine stetig differenzierbare Funktion $f: (t_0 - \varepsilon, t_0 + \varepsilon) \rightarrow \mathbb{R}$ mit $f(t_0) = x_0$ und $f'(t) = F(t, f(t))$, $t \in (t_0 - \varepsilon, t_0 + \varepsilon)$, finden kann (s. (6), §10).

9.1.7 *Gegenbeispiel.* Wir betrachten

$$\begin{aligned} F: \mathbb{R} \times \mathbb{R} &\longrightarrow \mathbb{R} \\ (t, x) &\longmapsto x^{\frac{2}{3}}. \end{aligned}$$

Die Funktion $f: \mathbb{R} \rightarrow \mathbb{R}$, $t \mapsto 0$, erfüllt offenbar die Differentialgleichung $f'(t) = F(t, f(t))$, $t \in \mathbb{R}$. Es gibt aber noch weitere Lösungen: Zu $a \in \mathbb{R}$ definieren wir

$$\begin{aligned} g_a: \mathbb{R} &\longrightarrow \mathbb{R} \\ t &\longmapsto \frac{1}{27} \cdot (t - a)^3. \end{aligned}$$

Wegen

$$g'_a(t) = \frac{1}{9} \cdot (t - a)^2 = (g_a(t))^{\frac{2}{3}}, \quad t \in \mathbb{R},$$

löst die Funktion g_a ebenfalls die obige Differentialgleichung, $a \in \mathbb{R}$. Es gilt weiter $f(a) = 0 = g_a(a)$, $a \in \mathbb{R}$. Für $c < a < d$ besitzt die durch $F_{[[c,d]] \times \mathbb{R}}: [c, d] \times$

$\mathbb{R} \rightarrow \mathbb{R}$ definierte Differentialgleichung (mindestens) zwei verschiedene Lösungen, die beide der Anfangsbedingung $x(a) = 0$ genügen. Damit kann $F|_{[c,d] \times \mathbb{R}}$ keine Lipschitz-Bedingung erfüllen. In der Tat seien $L > 0$, $0 < \varepsilon < 1/L$ und $\eta := \varepsilon^3$. Es gilt

$$|F(t, \eta) - F(t, 0)| = \varepsilon^2 > L \cdot \varepsilon^3 = L \cdot |\eta - 0|, \quad t \in [c, d].$$

Es gibt einen netten Prozess aus der Alltagswelt, der durch die obige Differentialgleichung modelliert wird und der die Bedeutung des Satzes von Picard–Lindelöf illustriert: das Lutschen eines Bonbons.⁶ Es werde ein kugelrundes Bonbon vom Volumen $V(t)$, $t \geq 0$, gelutscht. Die „Ablutschung“ sei proportional zur Oberfläche $O(t)$ des Bonbons, $t \geq 0$. Es gibt also eine Konstante $k > 0$, so dass

$$\frac{d}{dt}V(t) = -k \cdot O(t), \quad t \geq 0. \tag{9.2}$$

Es sei $r(t)$ der Radius des Bonbons zum Zeitpunkt t , $t \geq 0$. Dann gilt

$$V(t) = \frac{4}{3} \cdot \pi \cdot r(t)^3 \quad \text{und} \quad O(t) = 4 \cdot \pi \cdot r(t)^2 = 4 \cdot \pi \cdot \left(\frac{3}{4 \cdot \pi}\right)^{\frac{2}{3}} \cdot V(t)^{\frac{2}{3}}.$$

Mit $V(0) = V_0$ erhält man eine Anfangsbedingung. Sei t_1 der Zeitpunkt, zu dem das Bonbon „aufgelutscht“ ist. Könnten wir die Differentialgleichung (9.2) zu der Anfangsbedingung $V(t_1) = 0$ eindeutig lösen, dann könnten wir auch die Vergangenheit rekonstruieren, insbesondere das Volumen des ursprünglichen Bonbons. Offenbar können wir aus der Information $V(t_1) = 0$ nicht einmal rekonstruieren, ob ein Bonbon gelutscht wurde oder nicht. Diese einfachen Betrachtungen geben eine wertvolle Verbindung zwischen Alltagserfahrung und Intuition einerseits und mathematischem Formalismus andererseits.

9.2 Trennung der Variablen

In diesem und den folgenden Abschnitten besprechen wir explizite Lösungsmethoden für gewisse einfach strukturierte Differentialgleichungen.

Hier betrachten wir die Situation, in der zwei **offene** Intervalle $I, J \subset \mathbb{R}$ und stetige Funktionen $\varphi: I \rightarrow \mathbb{R}$ und $\psi: J \rightarrow \mathbb{R}$ gegeben sind und die Aufgabe gestellt wird, die Differentialgleichung

$$x'(t) = \varphi(t) \cdot \psi(x(t)) \tag{9.3}$$

zu lösen.⁷

Bei einer Differentialgleichung vom Typ (9.3) sprechen wir von einer *Differentialgleichung mit getrennten Variablen*.

9.2.1 Satz. *In der zuvor beschriebenen Situation sei $\psi(x) \neq 0$ für alle $x \in J$. Zu $t_0 \in I$ und $x_0 \in J$ definieren wir*

$$\begin{aligned} \Phi: I &\longrightarrow \mathbb{R} \\ t &\longmapsto \int_{t_0}^t \varphi(u) du \end{aligned}$$

⁶Aus (8).

⁷Hier sind also $U = I \times J \subset \mathbb{R}^2$ eine **offene** Teilmenge und $F: U \rightarrow \mathbb{R}$, $(t, x) \mapsto \varphi(t) \cdot \psi(x)$.

und

$$\begin{aligned}\psi: J &\longrightarrow \mathbb{R} \\ x &\longmapsto \int_{x_0}^x \frac{1}{\psi(y)} dy.\end{aligned}$$

Weiter sei $I' \subset I$ ein Intervall, so dass $t_0 \in I'$ und $\Phi(I') \subset \psi(J)$. Dann existiert genau eine Lösung

$$f: I' \longrightarrow \mathbb{R}$$

der Differentialgleichung (9.3) zu der Anfangsbedingung $f(t_0) = x_0$. Diese Lösung kann mit Hilfe der Formel

$$\psi(f(t)) = \Phi(t), \quad t \in I', \quad (9.4)$$

bestimmt werden.

Beweis. Wir zeigen zunächst, dass eine Lösung der Differentialgleichung (9.3) der Gleichung (9.4) genügt. Für eine Lösung hat man

$$f'(t) = \varphi(t) \cdot \psi(f(t)), \quad t \in I'.$$

Damit gilt für alle $t \in I'$ auch

$$\int_{t_0}^t \frac{f'(u)}{\psi(f(u))} du = \int_{t_0}^t \varphi(u) du = \Phi(t).$$

Nach der Substitutionsregel ((12), Satz 5.4.1) gilt

$$\Phi(t) = \int_{t_0}^t \frac{f'(u)}{\psi(f(u))} du = \int_{x_0}^{f(t)} \frac{1}{\psi(y)} dy = \psi(f(t)).$$

Jetzt überprüfen wir die Eindeutigkeit der Lösung. Nach dem Hauptsatz der Differential- und Integralrechnung ((12), Satz 5.3.5) ist die Funktion ψ differenzierbar mit Ableitung

$$\psi'(x) = \frac{1}{\psi(x)} \neq 0, \quad x \in J.$$

Damit ist ψ streng monoton wachsend oder fallend. (Genauere Begründung?) Sei

$$\Xi: \psi(J) \longrightarrow \mathbb{R}$$

die Umkehrfunktion. Sie ist stetig differenzierbar. Gleichung (9.4) und die Bedingung $\Phi(I') \subset \psi(J)$ zeigen

$$f(t) = \Xi(\Phi(t)), \quad t \in I'.$$

Mit diesen Erkenntnissen können wir leicht die Existenz der Lösung verifizieren. Auf Grund der Voraussetzung $\Phi(I') \subset \psi(J)$ können wir die Funktion

$$f := \Xi \circ \Phi: I' \longrightarrow \mathbb{R}$$

bilden. Es gilt

$$f(t_0) = \Xi(\Phi(t_0)) = \Xi(0) = x_0.$$

Die Anfangsbedingung ist damit erfüllt. Weiter erfüllt f nach Konstruktion Gleichung (9.4). Differenzieren dieser Gleichung führt zu der Beziehung

$$f'(t) \cdot \frac{1}{\psi(f(t))} = f'(t) \cdot \psi'(f(t)) = \Phi'(t) = \varphi(t).$$

Also löst f die Differentialgleichung (9.3). □

9.2.2 Beispiele. i) Wir untersuchen die Gleichung $x'(t) = x(t)$ zu Startwerten $t_0 \in \mathbb{R}$ und $x_0 \in \mathbb{R}_{>0}$. Wir haben in dieser Situation⁸

$$\begin{aligned} \varphi: \mathbb{R} &\longrightarrow \mathbb{R} \\ t &\longmapsto 1, \\ \psi: \mathbb{R}_{>0} &\longrightarrow \mathbb{R} \\ x &\longmapsto x. \end{aligned}$$

Diese Funktionen ergeben

$$\begin{aligned} \phi(t) &= \int_{t_0}^t \varphi(u) du = t - t_0, \quad t \in \mathbb{R}, \\ \psi(x) &= \int_{x_0}^x \frac{1}{y} dy = \log(x) - \log(x_0), \quad x \in \mathbb{R}_{>0}. \end{aligned}$$

Sowohl ϕ als auch ψ haben den Wertebereich \mathbb{R} , so dass $\phi(\mathbb{R}) \subset \psi(\mathbb{R}_{>0})$. Eine Lösung $f: \mathbb{R} \rightarrow \mathbb{R}$ der Differentialgleichung erfüllt daher

$$\log(f(t)) = t - t_0 + \log(x_0), \quad t \in \mathbb{R},$$

so dass

$$f(t) = x_0 \cdot \exp(t - t_0).$$

ii) Die Differentialgleichung laute $x'(t) = x(t)^2$, und die Anfangswerte seien $t_0 = 0$ und $x_0 \in \mathbb{R} \setminus \{0\}$.⁹

a) Für $x_0 > 0$ benutzen wir

$$\begin{aligned} \varphi: \mathbb{R} &\longrightarrow \mathbb{R} \\ t &\longmapsto 1, \\ \psi: \mathbb{R}_{>0} &\longrightarrow \mathbb{R} \\ x &\longmapsto x^2 \end{aligned}$$

und erhalten

$$\begin{aligned} \phi(t) &= \int_0^t du = t, \quad t \in \mathbb{R}, \\ \psi(x) &= \int_{x_0}^x \frac{1}{y^2} dy = \frac{1}{x_0} - \frac{1}{x}, \quad x \in \mathbb{R}_{>0}. \end{aligned}$$

Der Wertebereich von ψ ist $(-\infty, 1/x_0)$. Daher müssen wir $I' = (-\infty, 1/x_0)$ wählen, um $\phi(I') \subset \psi(\mathbb{R}_{>0})$ sicherzustellen. Die Funktion f wird durch die Gleichung

$$\frac{1}{f(t)} = \frac{1}{x_0} - t = \frac{1 - x_0 \cdot t}{x_0}$$

also durch

$$f(t) = \frac{x_0}{1 - x_0 \cdot t}, \quad t \in \left(-\infty, \frac{1}{x_0}\right),$$

gegeben.

⁸Man beachte, dass wir ein Intervall J mit $\psi(x) \neq 0$ für alle $x \in J$ angeben müssen, so dass $x_0 \in J$.

⁹Der Fall $x_0 = 0$ kann wegen der Forderung $\psi(x) = x^2 \neq 0, x \in J$, nicht mit dem obigen Formalismus behandelt werden.

b) Für $x_0 < 0$ arbeiten wir mit

$$\begin{aligned}\varphi: \mathbb{R} &\longrightarrow \mathbb{R} \\ t &\longmapsto 1, \\ \psi: \mathbb{R}_{<0} &\longrightarrow \mathbb{R} \\ x &\longmapsto x^2,\end{aligned}$$

weil $(t_0, x_0) \in \mathbb{R} \times \mathbb{R}_{<0}$. Die Lösung ist in diesem Fall

$$f(t) = \frac{x_0}{1 - x_0 \cdot t}, \quad t \in \left(\frac{1}{x_0}, \infty\right).$$

9.3 Freier Fall mit Luftwiderstand

Wir stellen eine Differentialgleichung aus der Physik vor, und zwar für die Geschwindigkeit eines Körpers im freien Fall bei Berücksichtigung des Luftwiderstands.

In den folgenden Betrachtungen stehe g für die Erdbeschleunigung, k für den sogenannten Reibungskoeffizienten und m für die Masse des fallenden Körpers. Die Erdbeschleunigung verursacht die Kraft

$$F_1 = m \cdot g$$

und die Reibung

$$F_2 = -k \cdot v^2,$$

v die Geschwindigkeit. Wir wenden das physikalische Gesetz

$\text{Kraft} = \text{Masse} \times \text{Beschleunigung}$

an, d.h.

$$m \cdot a = F_1 + F_2.$$

Damit erhalten wir folgende Differentialgleichung der Geschwindigkeit des fallenden Körpers in Abhängigkeit von der Zeit:

$$m \cdot v'(t) = m \cdot g - k \cdot v(t)^2,$$

also

$$v'(t) = g \cdot \left(1 - \frac{k}{m \cdot g} \cdot v(t)^2\right).$$

Wir setzen

$$f(t) := \sqrt{\frac{k}{m \cdot g}} \cdot v(t).$$

Die Funktion f genügt der Differentialgleichung

$$x'(t) = \sqrt{\frac{g \cdot k}{m}} \cdot (1 - x(t)^2).$$

Die Anfangsbedingung sei $v(0) = 0$,¹⁰ i.e. $f(0) = 0$. Jetzt formalisieren wir folgendermaßen:

$$\begin{aligned} \varphi: \mathbb{R} &\longrightarrow \mathbb{R} \\ t &\longmapsto \sqrt{\frac{g \cdot k}{m}}, \\ \psi: (-1, 1) &\longrightarrow \mathbb{R} \\ x &\longmapsto 1 - x^2. \end{aligned}$$

Diese Funktionen induzieren

$$\begin{aligned} \phi(t) &= \int_0^t \sqrt{\frac{g \cdot k}{m}} dt = \sqrt{\frac{g \cdot k}{m}} \cdot t, \quad t \in \mathbb{R}, \\ \psi(x) &= \int_0^x \frac{1}{1-y^2} dy = \operatorname{Artanh}(x), \quad t \in (-1, 1) \quad (\text{Übung}). \end{aligned}$$

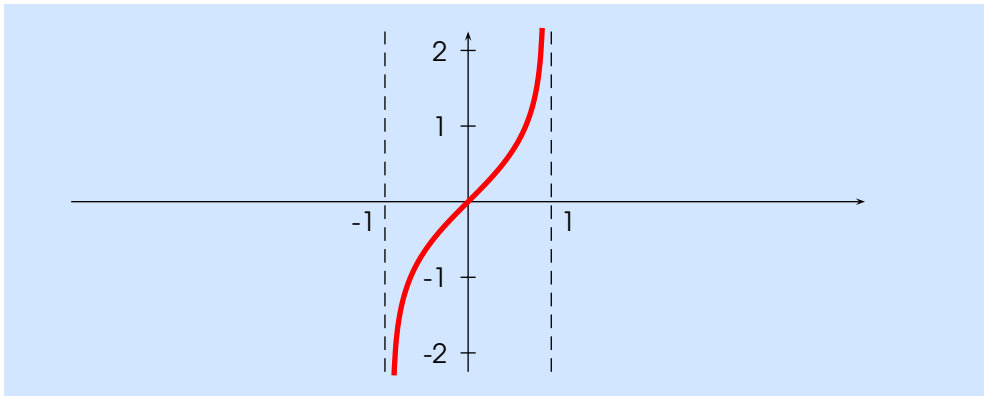


Abbildung 9.1: Areatangens hyperbolicus

Der Wertebereich des Areatangens hyperbolicus ist \mathbb{R} , so dass $\phi(\mathbb{R}) \subset \psi(\mathbb{R})$. Wir erhalten somit die Gleichungen

$$\operatorname{Artanh}(f(t)) = \sqrt{\frac{g \cdot k}{m}} \cdot t \quad \text{und} \quad f(t) = \tanh\left(\sqrt{\frac{g \cdot k}{m}} \cdot t\right), \quad t \in \mathbb{R}.$$

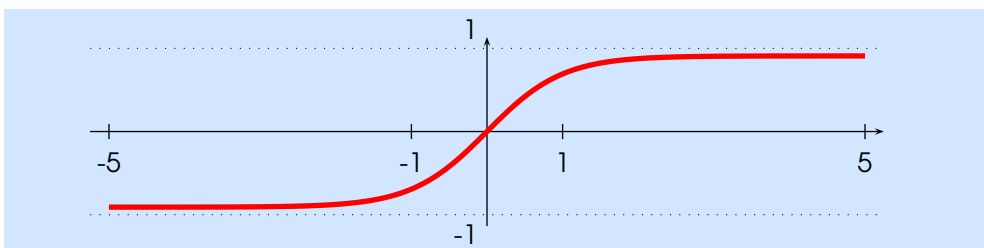


Abbildung 9.2: Tangens hyperbolicus

Schließlich folgern wir für die Geschwindigkeit

$$v(t) = \sqrt{\frac{m \cdot g}{k}} \cdot \tanh\left(\sqrt{\frac{g \cdot k}{m}} \cdot t\right), \quad t \in \mathbb{R}.$$

¹⁰Der Körper wird zum Zeitpunkt $t_0 = 0$ fallen gelassen und hat dabei die Anfangsgeschwindigkeit 0.

9.4 Lösung durch Substitution

Das zuvor besprochene Verfahren der Trennung der Variablen ist eine elegante Methode, geeignete Differentialgleichungen durch Integration zu lösen. In diesem Abschnitt stellen wir zwei weitere Typen von Differentialgleichungen vor, deren Lösung sich mit einem kleinen Kunstgriff auf die Lösung einer Differentialgleichung mit getrennten Variablen zurückführen lässt.

Fall A

Wir nehmen an, dass wir eine offene Teilmenge $U \subset \mathbb{R}^2$, eine stetige Funktion $F: U \rightarrow \mathbb{R}$, $(t, x) \mapsto F(t, x)$, Zahlen $a, b, c \in \mathbb{R}$, ein Intervall $J \subset \mathbb{R}$ und eine stetige Funktion $\alpha: J \rightarrow \mathbb{R}$ vorliegen haben, so dass

- $a \cdot t + b \cdot x + c \in J$, $(t, x) \in U$,
- $F(t, x) = \alpha(a \cdot t + b \cdot x + c)$, $(t, x) \in U$.

Gesucht ist wie üblich eine Lösung der Differentialgleichung

$$x'(t) = F(t, x(t)).$$

Dabei verfolgen wir folgenden Ansatz: Wir führen die Substitution $u(t) = a \cdot t + b \cdot x(t) + c$ durch. Dann gilt

$$u'(t) = a + b \cdot x'(t) = a + b \cdot \alpha(u(t)).$$

Wir müssen folglich die Differentialgleichung

$$x'(t) = a + b \cdot \alpha(x(t))$$

lösen. Dies ist eine Differentialgleichung mit getrennten Variablen. Wir formulieren eine präzise Aussage:

9.4.1 Satz. *In der obigen Situation sei $b \neq 0$. Es seien $(t_0, x_0) \in U$ und $I \subset \mathbb{R}$ ein Intervall, so dass $t_0 \in I$. Wir setzen*

$$\tilde{x}_0 = a \cdot t_0 + b \cdot x_0 + c.$$

Eine Funktion $f: I \rightarrow \mathbb{R}$ genügt genau dann den Bedingungen

- $f(t_0) = x_0$,
- $f'(t) = \alpha(a \cdot t + b \cdot f(t) + c)$, $t \in I$,

wenn die Funktion

$$\begin{aligned} g: I &\rightarrow \mathbb{R} \\ t &\mapsto a \cdot t + b \cdot f(t) + c \end{aligned}$$

die Eigenschaften

- $g(t_0) = \tilde{x}_0$,
- $g'(t) = a + b \cdot \alpha(g(t))$, $t \in I$,

aufweist.

Beweis. Wir nehmen an, f erfülle die genannten Bedingungen. Damit gilt

$$\begin{aligned} g(t_0) &= a \cdot t_0 + b \cdot f(t_0) + c = a \cdot t_0 + b \cdot x_0 + c = \tilde{x}_0, \\ g'(t) &= a + b \cdot f'(t) = a + b \cdot \alpha(a \cdot t + b \cdot f(t) + c) = a + b \cdot \alpha(g(t)), \quad t \in I. \end{aligned}$$

Jetzt habe g die genannten Eigenschaften. Dann gilt

$$a \cdot t_0 + b \cdot f(t_0) + c = g(t_0) = \tilde{x}_0 = a \cdot t_0 + b \cdot x_0 + c.$$

Auf Grund der Voraussetzung $b \neq 0$ folgt $f(t_0) = x_0$. Weiter haben wir

$$a + b \cdot f'(t) = g'(t) = a + b \cdot \alpha(g(t)), \quad t \in I.$$

Mit $b \neq 0$ schließen wir auf

$$f'(t) = \alpha(g(t)) = \alpha(a \cdot t + b \cdot f(t) + c), \quad t \in I.$$

Damit ist der Satz bewiesen. □

9.4.2 Beispiel. Wir betrachten die Differentialgleichung

$$x'(t) = 2t - x(t)$$

für $(t, x) \in \mathbb{R}^2$ mit der Anfangsbedingung $x(t_0) = x_0$. Wir setzen $x_0 > 2(t_0 - 1)$ voraus (s.u.). Hier haben wir

$$\alpha = \text{id}_{\mathbb{R}}: \mathbb{R} \rightarrow \mathbb{R}, \quad a = 2, \quad b = -1, \quad c = 0.$$

Wir lösen zunächst die Differentialgleichung

$$x'(t) = 2 - x(t)$$

zu der Anfangsbedingung $x(t_0) = 2 \cdot t_0 - x_0 =: \tilde{x}_0$. Wegen $x_0 > 2(t_0 - 1)$ gilt $\tilde{x}_0 < 2$. Das Verfahren der Trennung der Variablen führt uns zu den Funktionen

$$\begin{aligned} \phi(t) &= \int_{t_0}^t dx = t - t_0, \quad t \in \mathbb{R}, \\ \psi(x) &= \int_{\tilde{x}_0}^x \frac{1}{2-y} dy = -\log(2-y) \Big|_{\tilde{x}_0}^x = \log(2-\tilde{x}_0) - \log(2-x), \quad x \in (-\infty, 2). \end{aligned}$$

Die Funktionen ϕ und ψ haben beide den Wertebereich \mathbb{R} , so dass $\phi(\mathbb{R}) \subset \psi((-\infty, 2))$. Also gilt

$$t_0 - t = \log(2 - g(t)) - \log(2 - \tilde{x}_0) = \log\left(\frac{g(t) - 2}{\tilde{x}_0 - 2}\right), \quad t \in \mathbb{R},$$

d.h.

$$\begin{aligned} g: \mathbb{R} &\rightarrow \mathbb{R} \\ t &\mapsto (\tilde{x}_0 - 2) \cdot \exp(t_0 - t) + 2. \end{aligned}$$

Jetzt schreiben wir

$$g(t) = 2 \cdot t - f(t),$$

so dass

$$f(t) = 2 \cdot t - g(t) = (2 - \tilde{x}_0) \cdot \exp(t_0 - t) + 2t - 2.$$

Mit $\tilde{x}_0 = 2t_0 - x_0$ ergibt sich letztendlich

$$f(t) = (x_0 - 2t_0 + 2) \cdot \exp(t_0 - t) + 2t - 2, \quad t \in \mathbb{R}.$$

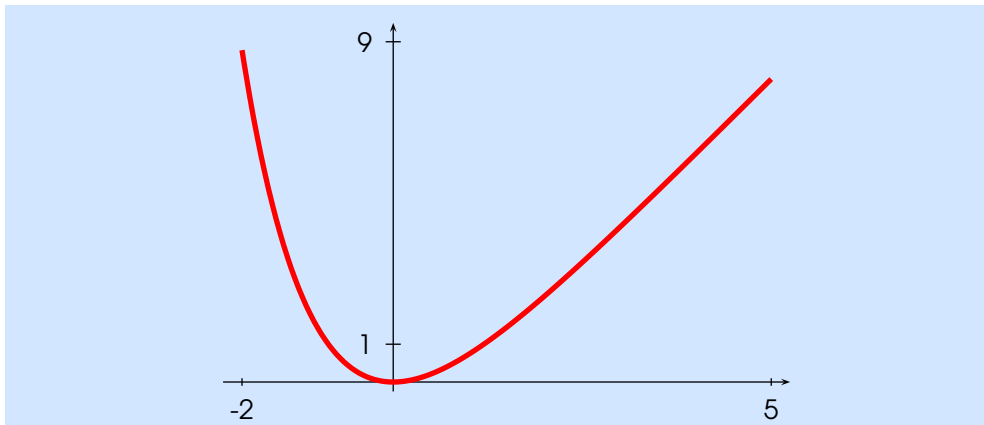


Abbildung 9.3: Die Lösung für $t_0 = 0 = x_0$

Fall B

Diesmal betrachten wir eine offene Menge $U \subset (\mathbb{R} \setminus \{0\}) \times \mathbb{R}$ und eine stetige Funktion $F: U \rightarrow \mathbb{R}$. Dabei nehmen wir an, dass es ein Intervall $J \subset \mathbb{R}$ gibt, so dass

$$\forall (t, y) \in U: \frac{y}{t} \in J$$

sowie eine Funktion

$$\alpha: J \rightarrow \mathbb{R}$$

mit

$$\forall (t, x) \in U: F(t, x) = \alpha\left(\frac{x}{t}\right).$$

Wir substituieren

$$u(t) = \frac{x(t)}{t}.$$

Es folgt

$$u'(t) = \frac{x'(t) \cdot t - x(t)}{t^2} = \frac{\alpha(u(t)) \cdot t - u(t) \cdot t}{t^2} = \frac{\alpha(u(t)) - u(t)}{t}.$$

Daher betrachten wir zunächst die Differentialgleichung

$$x'(t) = \frac{\alpha\left(\frac{x(t)}{t}\right) - x(t)}{t}.$$

Hierbei handelt es sich wiederum um eine Differentialgleichung mit getrennten Variablen. Diese etwas heuristischen Bemerkungen präzisieren wir folgendermaßen:

9.4.3 Satz. *Es seien $(t_0, y_0) \in U$ und $I \subset \mathbb{R} \setminus \{0\}$ ein Intervall mit $t_0 \in I$. Wir definieren $\tilde{x}_0 := x_0/t_0$. Eine Funktion $f: I \rightarrow \mathbb{R}$ hat genau dann die Eigenschaften*

- $f(t_0) = x_0$,
- $f'(t) = \alpha\left(\frac{f(t)}{t}\right), t \in I$,

wenn die Funktion

$$\begin{aligned} g: I &\longrightarrow \mathbb{R} \\ t &\longmapsto \frac{f(t)}{t} \end{aligned}$$

die Gleichungen

- $g(t_0) = \tilde{x}_0$,
- $g'(t) = \frac{\alpha(g(t)) - g(t)}{t}, t \in I$,

löst.

Beweis. Die Funktion f habe obige Eigenschaften. Dann gilt offenbar

$$\begin{aligned} g(t_0) &= \frac{f(t_0)}{t_0} = \frac{x_0}{t_0} = \tilde{x}_0, \\ g'(t) &= \frac{f'(t) \cdot t - f(t)}{t^2} = \frac{\alpha\left(\frac{f(t)}{t}\right) \cdot t - f(t)}{t^2} = \frac{\alpha(g(t)) - g(t)}{t}, \quad t \in I. \end{aligned}$$

Wenn für g die beiden angegebenen Gleichungen gelten, dann folgern wir

$$f(t_0) = t_0 \cdot g(t_0) = t_0 \cdot \tilde{x}_0 = t_0 \cdot \frac{x_0}{t_0} = x_0$$

und

$$f'(t) = g(t) + t \cdot g'(t) = g(t) + \alpha(g(t)) - g(t) = \alpha(g(t)) = \alpha\left(\frac{f(t)}{t}\right), \quad t \in I.$$

Dies beendet den Beweis. □

9.4.4 Beispiele. i) Wir möchten die Differentialgleichung

$$x'(t) = 1 + \left(\frac{x(t)}{t}\right) + \left(\frac{x(t)}{t}\right)^2$$

lösen. Dabei sei $(t, x) \in \mathbb{R}_{>0} \times \mathbb{R}$. Es gilt $\alpha(u) = 1 + u + u^2$ und $u(t) = x/t$, $(t, x) \in \mathbb{R}_{>0} \times \mathbb{R}$. Wir lösen folglich zuerst die Differentialgleichung

$$x'(t) = \frac{x(t)^2 + 1}{t}.$$

Dazu verwenden wir das Verfahren der Trennung der Variablen. Deshalb betrachten wir

$$\begin{aligned} \phi(t) &= \int_{t_0}^t \frac{1}{s} ds = \log(t) - \log(t_0) = \log\left(\frac{t}{t_0}\right), \quad t \in \mathbb{R}_{>0}, \\ \psi(x) &= \int_{\tilde{x}_0}^x \frac{1}{1+y^2} dy = \arctan(x) - \arctan(\tilde{x}_0), \quad x \in \mathbb{R}. \end{aligned}$$

Der Wertebereich von ψ ist

$$\left(-\frac{\pi}{2} - c_0, \frac{\pi}{2} - c_0\right), \quad c_0 := \arctan(\tilde{x}_0).$$

Daher wählen wir

$$I = \left(t_0 \cdot \exp\left(-\frac{\pi}{2} - c_0\right), t_0 \cdot \exp\left(\frac{\pi}{2} - c_0\right) \right).$$

Es gilt nun

$$\arctan(g(t)) = \arctan(\tilde{x}_0) + \log\left(\frac{t}{t_0}\right) = \arctan\left(\frac{x_0}{t_0}\right) + \log\left(\frac{t}{t_0}\right)$$

und damit

$$g(t) = \tan\left(\log\left(\frac{t}{t_0}\right) + \arctan\left(\frac{x_0}{t_0}\right)\right), \quad t \in I.$$

Als Lösung der ursprünglichen Differentialgleichung zur Anfangsbedingung $x(t_0) = x_0$ erhalten wir schließlich

$$f(t) = t \cdot \tan\left(\log\left(\frac{t}{t_0}\right) + \arctan\left(\frac{x_0}{t_0}\right)\right), \quad t \in I.$$

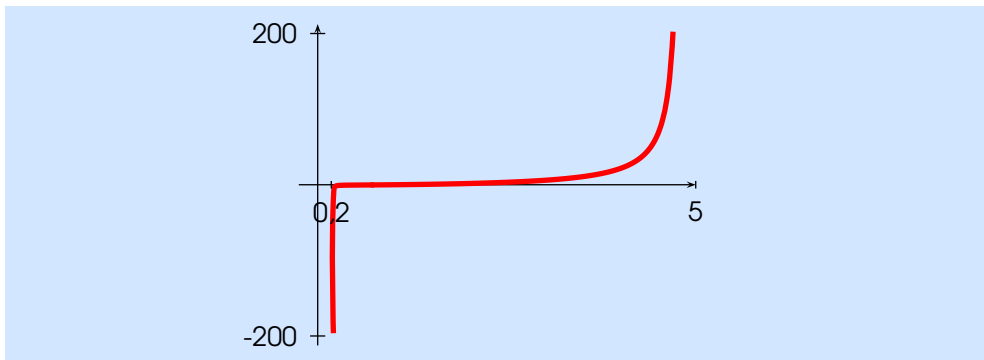


Abbildung 9.4: Die Lösung für $t_0 = 1$ und $x_0 = 0$

ii) Diesmal untersuchen wir

$$x'(t) = \frac{t + 2x(t)}{t} = 1 + 2 \cdot \frac{x(t)}{t}.$$

Dabei sei $(t, x) \in \mathbb{R}_{>0} \times \mathbb{R}$. Wir haben $\alpha(u) = 1 + 2u$, $u = x/t$. Es wird zunächst die Differentialgleichung

$$x'(t) = \frac{1 + x(t)}{t}$$

gelöst. Wir stellen die Anfangsbedingung $f(t_0) = 0$. Diesmal bekommen wir die Funktionen

$$\phi(t) = \log\left(\frac{t}{t_0}\right), \quad t \in \mathbb{R}_{>0},$$

$$\psi(x) = \int_0^x \frac{1}{1+y} dy = \log(1+x), \quad x \in (-1, \infty).$$

Die Funktionen ϕ und ψ haben Wertebereich \mathbb{R} . Wir finden

$$\log(1 + g(t)) = \log\left(\frac{t}{t_0}\right),$$

so dass

$$g(t) = \frac{t}{t_0} - 1, \quad t > 0$$

und

$$f(t) = t \cdot g(t) = \frac{1}{t_0} \cdot t^2 - t, \quad t > 0.$$

9.5 Lineare Differentialgleichungen

Eine weitere Klasse von Differentialgleichungen, die sich direkt durch Integration lösen lassen, bilden die linearen Differentialgleichungen. Wir unterteilen sie in homogene und inhomogene lineare Differentialgleichungen. Die homogenen sind wieder Differentialgleichungen mit getrennten Variablen und lassen sich mit den zuvor entwickelten Methoden behandeln. Eine inhomogene lineare Differentialgleichung kann man lösen, indem man zuerst die zugehörige homogene Differentialgleichung löst und dann eine weitere Integration ausführt.

9.5.1 Definition. Es seien $I \subset \mathbb{R}$ ein offenes Intervall und $F: I \times \mathbb{R} \rightarrow \mathbb{R}$ eine stetige Funktion. Die zugehörige Differentialgleichung ist eine *lineare Differentialgleichung erster Ordnung*, wenn stetige Funktionen $a, b: I \rightarrow \mathbb{R}$ existieren, so dass

$$F(t, x) = a(t) \cdot x + b(t).$$

Wenn $b = 0$ gilt, dann sprechen wir von einer *homogenen* linearen Differentialgleichung, andernfalls von einer *inhomogenen*.

Der homogene Fall

Eine homogene lineare Differentialgleichung erster Ordnung ist ein Spezialfall einer Differentialgleichung mit getrennten Variablen. Wir müssen also nur das Lösungsverfahren aus Abschnitt 9.2 anwenden. Wir stellen eine Anfangsbedingung $x(t_0) = x_0$, $t_0 \in I$, $x_0 \in \mathbb{R}$.

Für $x_0 = 0$ ist $f: \mathbb{R} \rightarrow \mathbb{R}$, $t \mapsto 0$, eine Lösung der Differentialgleichung. Wir behaupten, dass dies die einzige Lösung der Differentialgleichung ist. Für eine beliebige Lösung $g: I \rightarrow \mathbb{R}$ und $t \in I$ sei $I' := [t_0, t]$ bzw. $I' := [t, t_0]$. Es sei $L := \max\{|a(t)| \mid t \in I'\}$. Wir finden

$$|F(t, y) - F(t, x)| = |a(t)| \cdot |y - x| \leq L \cdot |y - x|, \quad x, y \in \mathbb{R}.$$

Daher erfüllt $F|_{I' \times \mathbb{R}}: I' \times \mathbb{R} \rightarrow \mathbb{R}$, $(t, x) \mapsto a(t) \cdot x$, die Lipschitz-Bedingung zur Lipschitz-Konstanten L , und $g|_{I'}: I' \rightarrow \mathbb{R}$ ist die einzige Lösung der Differentialgleichung zur Anfangsbedingung $x(t_0) = 0$. Es folgt $g|_{I'} = 0$, so dass $g(t) = 0$. Damit haben wir $g = 0$ gezeigt.

Für $x_0 \neq 0$ benutzen wir

$$\begin{aligned} \varphi = \alpha: I &\longrightarrow \mathbb{R} \\ t &\longmapsto \alpha(t), \\ \psi: \mathbb{R} &\longrightarrow \mathbb{R} \\ x &\longmapsto x \end{aligned}$$

und

$$\begin{aligned} \phi: I &\longrightarrow \mathbb{R} \\ t &\longmapsto \int_{t_0}^t \alpha(u) du, \\ \psi: J &\longrightarrow \mathbb{R} \\ x &\longmapsto \int_{x_0}^x \frac{1}{y} dy = \log\left(\frac{x}{x_0}\right). \end{aligned}$$

Dabei sei $J := \mathbb{R}_{>0}$, wenn $x_0 > 0$, und $J := \mathbb{R}_{<0}$, falls $x_0 < 0$. Da $\psi(J) = \mathbb{R}$, haben wir immer $\phi(I) \subset \psi(J)$. Wir finden

$$\log\left(\frac{f(t)}{x_0}\right) = \int_{t_0}^t a(u) du$$

und damit

$$f(t) = x_0 \cdot \exp\left(\int_{t_0}^t a(u) du\right), \quad t \in I.$$

Wir haben bewiesen:

9.5.2 Satz. In der obigen Situation gibt es genau eine stetig differenzierbare Funktion $f: I \rightarrow \mathbb{R}$ mit

- $f(t_0) = x_0$,
- $f'(t) = a(t) \cdot f(t)$, $t \in I$.

Sie ist durch

$$f(t) = x_0 \cdot \exp\left(\int_{t_0}^t a(u) du\right), \quad t \in I,$$

gegeben.

9.5.3 Beispiele. i) Die Differentialgleichung $x'(t) = k \cdot x(t)$ mit Anfangsbedingung $x(t_0) = x_0$ hat als einzige Lösung die Funktion¹¹

$$f(t) = x_0 \cdot \exp(k \cdot (t - t_0)), \quad t \in \mathbb{R}.$$

ii) Die eindeutig bestimmte Lösung der Differentialgleichung $x'(t) = 2t \cdot x(t)$ mit der Anfangsbedingung $x(0) = 5$ hat die eindeutig bestimmte Lösung

$$f(t) = 5 \cdot \exp\left(\int_0^t 2u du\right) = 5 \cdot \exp(t^2), \quad t \in \mathbb{R}.$$

iii) Wir betrachten die Differentialgleichung

$$x'(t) = -\frac{1}{t^2} \cdot x(t)$$

mit der Anfangsbedingung

$$x(t_0) = x_0.$$

Sie hat als einzige Lösung

$$f(t) = x_0 \cdot \exp\left(-\int_{t_0}^t \frac{1}{u^2} du\right) = x_0 \cdot \exp\left(\frac{1}{t} - \frac{1}{t_0}\right), \quad t > 0.$$

Der inhomogene Fall

Wir betrachten jetzt eine inhomogene lineare Differentialgleichung nebst Anfangsbedingung:

- $x'(t) = a(t) \cdot x(t) + b(t)$,
- $x(t_0) = x_0$.

¹¹Vgl. (12), Satz 4.8.5.

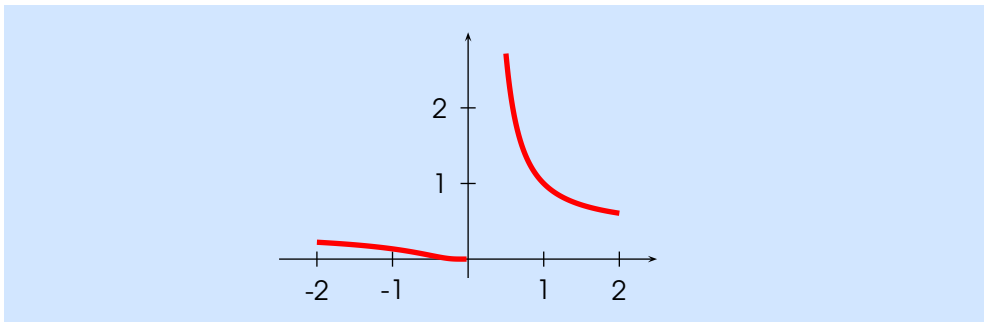


Abbildung 9.5: Die Funktion $t \mapsto \exp((1/t) - 1)$, $t \in \mathbb{R} \setminus \{0\}$

Es gibt höchstens eine Lösung: Für zwei differenzierbare Funktionen $f_1, f_2: I \rightarrow \mathbb{R}$ mit $f_i'(t) = a \cdot f_i(t) + b(t)$, $t \in I$, und $f_i(t_0) = x_0$, $i = 1, 2$, erfüllt die Funktion $g := f_1 - f_2$ die Gleichungen $g'(t) = a(t) \cdot g(t)$, $t \in I$, und $g(t_0) = 0$. Die Nullfunktion $h: I \rightarrow \mathbb{R}$ mit $h(t) = 0$, $t \in I$, erfüllt dieselben Bedingungen. Aus Satz 9.5.2 folgt $g = h = 0$ und damit $f_1 = f_2$.

Wir nehmen an, dass die Funktion $f: I \rightarrow \mathbb{R}$ die homogene Differentialgleichung

$$x'(t) = a(t) \cdot x(t)$$

löst und $f(t_0) = 1$ erfüllt. Diese Funktion wird in Satz 9.5.2 beschrieben. Es folgt:

$$\forall t \in I: f(t) = \exp\left(\int_{t_0}^t a(u) du\right) \neq 0.$$

Eine Lösung $g: I \rightarrow \mathbb{R}$ der inhomogenen Differentialgleichung kann nun in der Form

$$g(t) = f(t) \cdot h(t), \quad t \in I,$$

dargestellt werden. Durch Ableiten finden wir

$$\begin{aligned} g'(t) &= f'(t) \cdot h(t) + f(t) \cdot h'(t) \\ &= a(t) \cdot f(t) \cdot h(t) + f(t) \cdot h'(t) \\ &= a(t) \cdot g(t) + f(t) \cdot h'(t), \quad t \in I. \end{aligned}$$

Die Funktion g löst somit genau dann die Differentialgleichung mit der gestellten Anfangsbedingung, wenn $f(t) \cdot h'(t) = b(t)$, $t \in I$, und $h(t_0) = x_0$. Mit dieser Bedingung können wir h bestimmen:

$$h(t) = \int_{t_0}^t \frac{b(u)}{f(u)} du + x_0, \quad t \in I.$$

Dieses Verfahren nennt man *Variation der Konstanten*. Es ist offenbar durch die Produktregel ((12), Satz 4.2.1, iii) motiviert. Wir fassen unsere Diskussion zusammen:

9.5.4 Satz. *Es gibt genau eine stetig differenzierbare Funktion $g: I \rightarrow \mathbb{R}$ mit*

- $g(t_0) = x_0$,
- $g'(t) = a(t) \cdot g(t) + b(t)$, $t \in I$.

Sie ist durch die Formel

$$g(t) = f(t) \cdot \left(x_0 + \int_{t_0}^t \frac{b(u)}{f(u)} du \right), \quad t \in I,$$

mit

$$f(t) := \exp \left(\int_{t_0}^t a(u) du \right), \quad t \in I,$$

gegeben.

9.5.5 Beispiele. i) Wir studieren die Differentialgleichung $x'(t) = 2tx(t) + t^3$ zu der Anfangsbedingung $x(0) = x_0$. Die homogene lineare Differentialgleichung $x'(t) = 2tx(t)$ mit der Anfangsbedingung $x(0) = 1$ hat die Lösung

$$f(t) = \exp \left(\int_0^t 2u du \right) = \exp(t^2), \quad t \in \mathbb{R}.$$

Es folgt

$$g(t) = \exp(t^2) \cdot \left(x_0 + \int_0^t u^3 \cdot \exp(-u^2) du \right), \quad t \in \mathbb{R}.$$

Jetzt müssen wir das Integral

$$\int_0^t u^3 \cdot \exp(-u^2) du$$

auswerten. Im ersten Schritt substituieren wir $s = u^2$:

$$\int_0^t u^3 \cdot \exp(-u^2) du = \frac{1}{2} \cdot \int_0^{t^2} s \cdot \exp(-s) ds, \quad t \in \mathbb{R}.$$

Jetzt integrieren wir partiell:

$$\begin{aligned} \int_0^{t^2} s \cdot \exp(-s) ds &= -s \cdot \exp(-s) \Big|_0^{t^2} + \int_0^{t^2} \exp(-s) ds \\ &= -t^2 \cdot \exp(-t^2) - \exp(-s) \Big|_0^{t^2} \\ &= 1 - (t^2 + 1) \cdot \exp(-t^2). \end{aligned}$$

Damit folgt

$$g(t) = \left(x_0 + \frac{1}{2} \right) \cdot \exp(t^2) - \frac{1}{2} \cdot (t^2 + 1), \quad t \in \mathbb{R}.$$

ii) Um die Lösung der inhomogenen linearen Differentialgleichung

$$x'(t) = \cos(t) - \frac{x(t)}{t}$$

zu der Anfangsbedingung $x(t_0) = x_0$, $t_0 \in \mathbb{R}_{>0}$, zu bestimmen, ermitteln wir zunächst die Lösung der homogenen Differentialgleichung

$$x'(t) = -\frac{x(t)}{t}$$

zu der Anfangsbedingung $x(t_0) = 1$. Sie ist

$$f(t) = \exp \left(- \int_{t_0}^t \frac{1}{u} du \right) = \exp \left(\log \left(\frac{t_0}{t} \right) \right) = \frac{t_0}{t}, \quad t > 0.$$

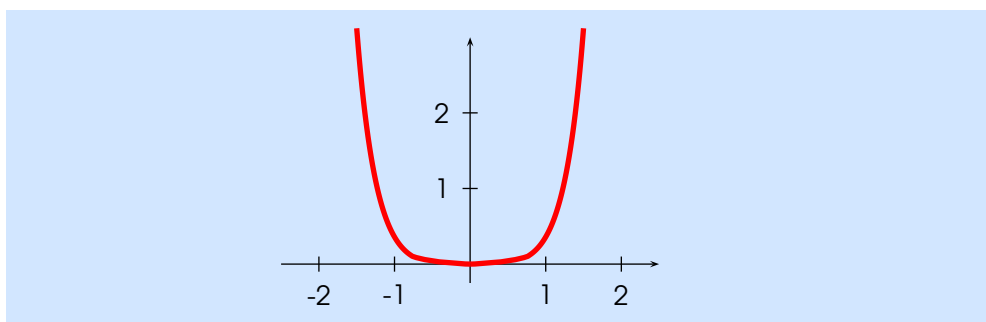


Abbildung 9.6: Die Lösung für $x_0 = 0$

Für die Lösung der inhomogenen Differentialgleichung folgt

$$g(t) = \frac{t_0}{t} \cdot \left(x_0 + \int_{t_0}^t \frac{u \cdot \cos(u)}{t_0} du \right) = \frac{t_0 \cdot x_0}{t} + \frac{1}{t} \cdot \int_{t_0}^t u \cdot \cos(u) du, \quad t > 0.$$

Mittels partieller Integration sehen wir

$$\begin{aligned} \int_{t_0}^t u \cdot \cos(u) du &= u \cdot \sin(u) \Big|_{t_0}^t - \int_{t_0}^t \sin(u) du \\ &= t \cdot \sin(t) - t_0 \cdot \sin(t_0) + \cos(t) - \cos(t_0), \quad t > 0. \end{aligned}$$

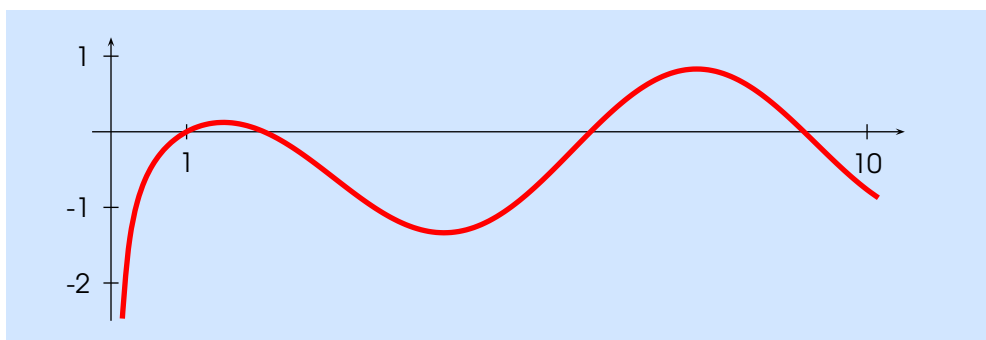


Abbildung 9.7: Die Lösung für $t_0 = 1$ und $x_0 = 0$

Somit haben wir

$$g(t) = \frac{t_0 \cdot x_0 - t_0 \cdot \sin(t_0) - \cos(t_0)}{t} + \sin(t) + \frac{\cos(t)}{t}, \quad t > 0.$$

9.6 Differentialgleichungen höherer Ordnung

Bisher haben wir uns nur mit Systemen von Differentialgleichungen erster Ordnung beschäftigt. In diesem Abschnitt werden wir kurz erklären, wie man eine Differentialgleichung höherer Ordnung in ein System von Differentialgleichungen der Ordnung eins umformt.

Es seien $U \subset \mathbb{R}^{n+1}$ eine offene Teilmenge und $F: U \rightarrow \mathbb{R}$ eine stetige Funktion. Wir möchten die Differentialgleichung

$$x^{(n)}(t) = F(t, x(t), \dots, x^{(n-1)}(t)) \tag{9.5}$$

der Ordnung n lösen, d.h. wir suchen ein Intervall $I \subset \mathbb{R}$ und eine n -mal stetig differenzierbare Funktion $f: I \rightarrow \mathbb{R}$, so dass

- $(t, f(t), \dots, f^{(n-1)}(t)) \in U, t \in I,$
- $f^{(n)}(t) = F(t, f(t), \dots, f^{(n-1)}(t)), t \in I.$

Wir untersuchen folgende Gleichungen

$$\begin{aligned} x_0'(t) &= x_1(t) \\ x_1'(t) &= x_2(t) \\ &\vdots \\ x_{n-2}'(t) &= x_{n-1}(t) \\ x_{n-1}'(t) &= F(t, x_0(t), \dots, x_{n-1}(t)). \end{aligned}$$

Formal definieren wir

$$\begin{aligned} G: U &\rightarrow \mathbb{R}^n \\ (t, x_0, \dots, x_{n-1}) &\mapsto (x_1, \dots, x_{n-1}, F(t, x_0, \dots, x_{n-1})) \end{aligned}$$

und widmen uns dem System

$$x'(t) = G(t, x(t)) \tag{9.6}$$

von Differentialgleichungen erster Ordnung.

Wir nehmen nun an, dass die Funktion $f: I \rightarrow \mathbb{R}$ die Differentialgleichung (9.5) löst, und definieren

$$h(t) := (f(t), f'(t), \dots, f^{(n-1)}(t)), \quad t \in I.$$

Dann haben wir

$$\begin{aligned} h'(t) &= (f'(t), f''(t), \dots, f^{(n-1)}(t), f^{(n)}(t)) \\ &= (f'(t), f''(t), \dots, f^{(n-1)}(t), F(t, f(t), f'(t), \dots, f^{(n-1)}(t))) \\ &= G(t, f(t), \dots, f^{(n-1)}(t)) = G(t, h(t)), \end{aligned}$$

d.h. h ist eine Lösung des Systems (9.6).

Sei umgekehrt

$$h(t) = (h_0(t), \dots, h_{n-1}(t)), \quad t \in I,$$

eine Lösung des Systems (9.6). Wir finden

$$\begin{aligned} h_0'(t) &= h_1(t) \\ &\vdots \\ h_{n-2}'(t) &= h_{n-1}(t) \\ h_{n-1}'(t) &= F(t, h_0(t), \dots, h_{n-1}(t)). \end{aligned}$$

Aus den ersten $n - 1$ Gleichungen leiten wir

$$h_i(t) = h_0^{(i)}(t), \quad i = 1, \dots, n - 1,$$

ab und zusammen mit der letzten Gleichung

$$h_0^{(n)}(t) = h_{n-1}'(t) = F(t, h_0(t), h_1(t), \dots, h_{n-1}(t)) = F(t, h_0(t), h_0'(t), \dots, h_0^{(n-1)}(t)),$$

$t \in I$. Damit löst die Funktion $h_0: I \rightarrow \mathbb{R}$ die Differentialgleichung (9.5).

9.6.1 Fazit. Eine Lösung einer Differentialgleichung der Ordnung n lässt sich aus einer Lösung eines Systems von Differentialgleichungen erster Ordnung gewinnen.

9.6.2 Übung. Überlegen Sie sich, wie die Anfangsbedingung für eine Differentialgleichung höherer Ordnung formuliert werden muss, damit sie unter den obigen Umformungen zu einer Anfangsbedingung für das entsprechende System von Differentialgleichungen erster Ordnung wird.

9.6.3 Beispiel. Das physikalische Gesetz

$$\boxed{\text{Kraft} = \text{Masse} \times \text{Beschleunigung}}$$

führt auf die Differentialgleichung

$$m \cdot x''(t) = F(t, x(t), x'(t))$$

zweiter Ordnung. Diese Differentialgleichung ist äquivalent zu dem System

- $v(t) = x'(t)$
- $v'(t) = (1/m) \cdot F(t, x(t), v(t))$

von Differentialgleichungen erster Ordnung.

Der Satz von Picard–Lindelöf 9.1.4 liefert somit hinreichende Kriterien für die Existenz und Eindeutigkeit von Lösungen von Differentialgleichungen höherer Ordnung.

9.6.4 Beispiel. Wir betrachten die Differentialgleichung

$$x''(t) = -\frac{x'(t)}{t} + \frac{x(t)}{t^2}$$

für $t \neq 0$. Dies entspricht dem System

$$\begin{aligned} x_0'(t) &= x_1(t) \\ x_1'(t) &= -\frac{x_1(t)}{t} + \frac{x_0(t)}{t^2} \end{aligned}$$

von Differentialgleichungen. Es gehört zu der stetigen Funktion

$$\begin{aligned} F: (\mathbb{R} \setminus \{0\}) \times \mathbb{R}^2 &\longrightarrow \mathbb{R}^2 \\ (t, x_0, x_1) &\longmapsto \left(x_1, -\frac{x_1}{t} + \frac{x_0}{t^2} \right). \end{aligned}$$

Wir schreiben die Abbildung F in der Form

$$\begin{pmatrix} F_1(t, x_0, x_1) \\ F_2(t, x_0, x_1) \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ \frac{1}{t^2} & -\frac{1}{t} \end{pmatrix} \cdot \begin{pmatrix} x_0 \\ x_1 \end{pmatrix}, \quad t \in \mathbb{R} \setminus \{0\}, (x_0, x_1) \in \mathbb{R}^2.$$

Nach Bemerkung 8.3.4 haben wir

$$\left\| \begin{pmatrix} 0 & 1 \\ \frac{1}{t^2} & -\frac{1}{t} \end{pmatrix} \right\|_{\text{op}} \leq \sqrt{1 + \frac{1}{t^2} + \frac{1}{t^4}} \leq 1 + \frac{1}{t^2}, \quad t \in \mathbb{R} \setminus \{0\}.$$

Es folgt, dass $F|_{I \times \mathbb{R}^2}$ für jedes kompakte Intervall $I \subset \mathbb{R}_{>0}$ oder $I \subset \mathbb{R}_{<0}$ eine Lipschitz-Bedingung erfüllt. Damit ist der Satz von Picard–Lindelöf anwendbar.

Lineare Differentialgleichungen mit konstanten Koeffizienten

Da wir bisher keine konkreten Lösungsverfahren für **Systeme** von Differentialgleichungen besprochen haben, hilft uns die obige Beobachtung noch nicht bei der expliziten Bestimmung von Lösungen.

9.6.5 Definition. Es seien a_0, \dots, a_n reelle Zahlen. Die zugehörige *homogene lineare Differentialgleichung mit konstanten Koeffizienten* ist

$$a_n \cdot x^{(n)}(t) + \dots + a_1 \cdot x'(t) + a_0 \cdot x(t) = 0.$$

Wir nennen

$$p(\xi) := a_n \cdot \xi^n + \dots + a_1 \cdot \xi + a_0$$

das *charakteristische Polynom* der Differentialgleichung.

Für solch eine Differentialgleichung können wir einen Adhoc-Ansatz verfolgen. Dazu definieren wir zu $\lambda \in \mathbb{R}$

$$\begin{aligned} f_\lambda: \mathbb{R} &\mapsto \mathbb{R} \\ t &\mapsto \exp(\lambda \cdot t). \end{aligned}$$

Aus $f'_\lambda = \lambda \cdot f_\lambda$, $\lambda \in \mathbb{R}$, folgt:

9.6.6 Lemma. Es sei $\lambda \in \mathbb{R}$. Die Funktion f_λ erfüllt genau dann die Differentialgleichung

$$a_n \cdot f_\lambda^{(n)}(t) + \dots + a_1 \cdot f'_\lambda(t) + a_0 \cdot f_\lambda(t) = 0, \quad t \in \mathbb{R},$$

wenn

$$p(\lambda) = 0$$

gilt.

9.6.7 Bemerkung. Da die Differentialgleichung linear und homogen ist, ist zu einer Lösung f_λ , $\lambda \in \mathbb{R}$, auch die Funktion $\alpha \cdot f_\lambda$ für jedes $\alpha \in \mathbb{R}$ eine Lösung.

9.6.8 Beispiel. Wir betrachten die Differentialgleichung

$$x'''(t) - x''(t) + x'(t) - x(t) = 0.$$

Nach dem Lemma suchen wir zunächst die Lösungen der Gleichung

$$\xi^3 - \xi^2 + \xi - 1 = 0.$$

Eine Lösung ist $\lambda = 1$. Da

$$\xi^3 - \xi^2 + \xi - 1 = (\xi - 1) \cdot (\xi^2 + 1),$$

ist $\lambda = 1$ die einzige reelle Lösung der betrachteten Gleichung. Damit liefert unser Verfahren die Lösungen $\alpha \cdot f_1$, $\alpha \in \mathbb{R}$.

Eulersche Differentialgleichungen

9.6.9 Definition. Es seien a_0, \dots, a_n reelle Zahlen. Diese definieren die *Eulersche Differentialgleichung*

$$a_n \cdot t^n \cdot x^{(n)}(t) + \dots + a_1 \cdot t \cdot x'(t) + a_0 \cdot x(t) = 0.$$

Es seien $\lambda \in \mathbb{R}$ und $f: \mathbb{R}_{>0} \rightarrow \mathbb{R}, t \mapsto t^\lambda$. Wegen der Formel $f'(t) = \lambda \cdot t^{\lambda-1}$, $t > 0$, wird die Bedingung

$$a_n \cdot t^n \cdot f^{(n)}(t) + \dots + a_1 \cdot t \cdot f'(t) + a_0 \cdot f(t) = 0, \quad t > 0,$$

zu einer polynomialen Gleichung in λ .

9.6.10 *Beispiel.* Wie wollen Lösungen für die Differentialgleichung

$$2t^2 \cdot x''(t) - t \cdot x'(t) + x(t) = 0$$

finden. Für $f(t) = t^\lambda, t > 0$, finden wir

$$2t^2 \cdot f''(t) - t \cdot f'(t) + f(t) = 2(\lambda \cdot (\lambda - 1)) \cdot t^\lambda - \lambda \cdot t^\lambda + t^\lambda = (2\lambda^2 - 3\lambda + 1) \cdot t^\lambda, \quad t > 0.$$

Wegen

$$2\lambda^2 - 3\lambda + 1 = 2 \cdot (\lambda - 1) \cdot \left(\lambda - \frac{1}{2}\right)$$

sind die Funktionen $f_1: \mathbb{R}_{>0} \rightarrow \mathbb{R}, t \mapsto t$, und $f_2: \mathbb{R}_{>0} \rightarrow \mathbb{R}, t \mapsto \sqrt{t}$, Lösungen der Differentialgleichung. Da es sich um eine homogene lineare Differentialgleichung handelt, sind die Funktionen

$$\alpha_1 \cdot f_1 + \alpha_2 \cdot f_2, \quad \alpha_1, \alpha_2 \in \mathbb{R},$$

ebenfalls Lösungen.

9.6.11 *Bemerkung.* i) Man kann eine Eulersche Differentialgleichung in eine homogene lineare mit konstanten Koeffizienten umformen und den obigen Ansatz aus dem Ansatz für homogene lineare Differentialgleichung mit konstanten Koeffizienten ableiten. Dazu setzen wir die Lösung in der Form

$$f(t) = g(\log(t)), \quad t > 0,$$

an. Dann finden wir mit Produkt- und Kettenregel ((12), Satz 4.2.1 und 4.2.4)

$$\begin{aligned} f'(t) &= \frac{1}{t} \cdot g'(\log(t)) \\ f''(t) &= -\frac{1}{t^2} \cdot g'(\log(t)) + \frac{1}{t^2} \cdot g''(\log(t)), \quad t > 0. \end{aligned}$$

Mittels vollständiger Induktion beweist man die Existenz reeller Zahlen $a_{kl}, l = 0, \dots, k, k \in \mathbb{N}$, mit

$$f^{(k)}(t) = \frac{1}{t^k} \cdot \sum_{l=0}^k a_{kl} \cdot g^{(l)}(\log(t)), \quad k \in \mathbb{N}, t > 0.$$

Damit erhalten wir

$$\sum_{k=0}^n \sum_{l=0}^k a_{kl} \cdot g^{(l)}(\log(t)) = 0, \quad t > 0.$$

Schreiben wir $t = \exp(s), s \in \mathbb{R}$, so finden wir

$$\sum_{k=0}^n \sum_{l=0}^k a_{kl} \cdot g^{(l)}(s) = 0, \quad s \in \mathbb{R}.$$

Damit löst die Funktion g eine homogene lineare Differentialgleichung der Ordnung n mit konstanten Koeffizienten. Mit dem Ansatz $g(s) = \exp(\lambda \cdot s)$, $s \in \mathbb{R}$, für ein $\lambda \in \mathbb{R}$, bekommen wir den Ansatz

$$f(t) = t^\lambda, \quad t > 0.$$

ii) Bei den obigen Betrachtungen haben wir die Frage nach der Vollständigkeit des Systems der gefundenen Lösungen nicht erörtert. Auch sind wir nicht auf Anfangsbedingungen eingegangen. Manchmal waren wir bei den Betrachtungen zum Definitionsbereich etwas sorglos. Die Leserin bzw. der Leser ist herzlich eingeladen, notwendige Präzisierungen und Zusatzüberlegungen selbstständig durchzuführen und die Literatur zu diesem Themenkreis zu studieren, z.B. (2) oder (6).

Kapitel 10

Der Umkehrsatz

10.1 Homöomorphismen, Gebiete und Diffeomorphismen

Es seien I ein offenes Intervall und $f: I \rightarrow \mathbb{R}$ eine Funktion. Die Funktion f sei in $x_0 \in I$ stetig differenzierbar mit $f'(x_0) \neq 0$. Wir betrachten den Fall $f'(x_0) > 0$. Auf Grund der stetigen Differenzierbarkeit von f in x_0 existiert ein $\varepsilon > 0$, so dass f auf $(x_0 - \varepsilon, x_0 + \varepsilon)$ differenzierbar ist. Die Stetigkeit von f' in x_0 und die Bedingung $f'(x_0) > 0$ implizieren, dass wir ein ε' mit $0 < \varepsilon' \leq \varepsilon$ finden können, so dass $f'(x) > 0$ für alle $x \in I' := (x_0 - \varepsilon', x_0 + \varepsilon')$. Die Funktion $f|_{I'}$ ist differenzierbar und streng monoton wachsend. Damit ist sie umkehrbar, und die Umkehrfunktion ist ebenfalls streng monoton wachsend und differenzierbar.

Dieser Satz besitzt eine Verallgemeinerung auf differenzierbare Abbildungen $F: U \rightarrow \mathbb{R}^n$, $U \subset \mathbb{R}^n$ offen. Um diese Verallgemeinerung sauber formulieren zu können, müssen wir unseren „topologischen“ Begriffsapparat ausbauen. Dieser Aufgabe widmen wir uns in diesem Abschnitt.

10.1.1 Definition. Es seien X und Y topologische Räume und $f: X \rightarrow Y$ eine Abbildung. Wir sagen, dass f ein *Homöomorphismus* ist, wenn f stetig ist und eine stetige Abbildung $g: Y \rightarrow X$ mit

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

existiert.

10.1.2 Bemerkung. Eine stetige Abbildung $f: X \rightarrow Y$ ist genau dann ein Homöomorphismus, wenn sie bijektiv ist und die Umkehrabbildung $f^{-1}: Y \rightarrow X$ ebenfalls stetig ist.

10.1.3 Definition. Es seien X, Y topologische Räume und $f: X \rightarrow Y$ eine stetige Abbildung. Die Abbildung f ist eine *offene* Abbildung, wenn gilt:

$$\forall U \subset X: \quad U \text{ offen} \quad \Rightarrow \quad f(U) \text{ offen.}$$

10.1.4 Lemma. Es seien X, Y topologische Räume und $f: X \rightarrow Y$ eine bijektive stetige Abbildung. Dann ist f genau dann ein Homöomorphismus, wenn f bijektiv und offen ist.

Beweis. Wegen Bemerkung 10.1.2 ist zu zeigen, dass f^{-1} genau dann stetig ist, wenn f offen ist. Sei $U \subset X$ eine offene Teilmenge. Wir schreiben¹ $g := f^{-1}$.

¹Die einzige Schwierigkeit im Beweis ist, die beiden Bedeutungen von $(\cdot)^{-1}$ auseinanderzuhalten: Mit $g := f^{-1}$ meinen wir die (mengentheoretische) Umkehrabbildung der bijektiven Abbildung f und mit $g^{-1}(U)$ das Urbild einer Teilmenge $U \subset X$ unter der Abbildung g .

Nach Definition gilt

$$g^{-1}(U) = \{y \in Y \mid g(y) \in U\}.$$

Wenn y in $g^{-1}(U)$ enthalten ist, dann haben wir $y = f(g(y)) \in f(U)$. Damit gilt

$$g^{-1}(U) \subset f(U). \quad (10.1)$$

Wenn umgekehrt y ein Element von $f(U)$ ist, dann gilt $g(y) \in g(f(U)) = U$. Also folgt

$$f(U) \subset g^{-1}(U) \quad \text{und wegen (10.1)} \quad f(U) = g^{-1}(U).$$

Unsere Überlegungen zeigen insbesondere, dass gilt

$$\forall U \subset X \text{ offen: } g^{-1}(U) \text{ ist offen} \iff f(U) \text{ ist offen.}$$

Daraus folgt sofort die Behauptung. □

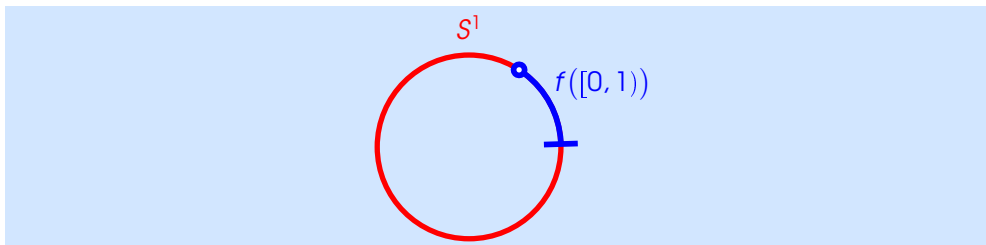
10.1.5 Gegenbeispiel. Die Mengen $[0, 2\pi) \subset \mathbb{R}$ bzw.

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

„erben“ von \mathbb{R} bzw. \mathbb{R}^2 eine Topologie,² und die Abbildung

$$\begin{aligned} f: [0, 2\pi) &\longrightarrow S^1 \\ t &\longmapsto (\cos(t), \sin(t)) \end{aligned}$$

aus Beispiel 3.2.4, iv), ist bijektiv und stetig. Sie ist aber nicht offen: Die Menge $[0, 1) = (-1, 1) \cap [0, 2\pi)$ ist eine offene Teilmenge, ihr Bild in S^1 ist nicht offen, weil jede offene Umgebung von $(1, 0) = f(0)$ sowohl Punkte oberhalb als auch unterhalb der x -Achse enthält. Nach Lemma 10.1.4 ist f kein Homöomorphismus.



Man kann auch folgendermaßen argumentieren: Wir nehmen an, dass $f^{-1}: S^1 \rightarrow [0, 2\pi)$ stetig ist. Nach dem Satz von Heine–Borel 2.2.12 ist S^1 kompakt, $[0, 2\pi)$ aber nicht. Auf der anderen Seite müsste $[0, 2\pi) = f(S^1)$ nach Satz 3.4.1 kompakt sein.

10.1.6 Definition. a) Eine Teilmenge $M \subset \mathbb{R}^n$ heißt *wegzusammenhängend*, wenn es zu je zwei Punkten $x_1, x_2 \in M$ einen Weg, d.h. eine stetige Abbildung,

$$\gamma: [0, 1] \longrightarrow M$$

mit

$$\gamma(0) = x_1 \quad \text{und} \quad \gamma(1) = x_2$$

gibt.

b) Eine offene und wegzusammenhängende Teilmenge $G \subset \mathbb{R}^n$ nennt man ein *Gebiet*.

²S. Beispiel 1.4.2, iv). So ist z.B. eine Teilmenge U von $[0, 2\pi)$ genau dann offen, wenn es eine offene Teilmenge $V \subset \mathbb{R}$ mit $U = V \cap [0, 2\pi)$ gibt.

10.1 Homöomorphismen, Gebiete und Diffeomorphismen

10.1.7 *Beispiele.* i) Jede konvexe Teilmenge $M \subset \mathbb{R}^n$ ist wegzusammenhängend.

ii) Für $a \in \mathbb{R}^n$ und $r > 0$ ist $B(a, r)$ ein Gebiet.

10.1.8 *Bemerkung.* Für jede offene Teilmenge $U \subset \mathbb{R}^n$ gibt es eine abzählbare, i.e. endliche oder abzählbar unendliche, Indexmenge I sowie Gebiete G_i , $i \in I$, so dass

$$U = \bigsqcup_{i \in I} G_i \quad (\text{d.h. } U = \bigcup_{i \in I} G_i \text{ und } G_i \cap G_j = \emptyset, i, j \in I, i \neq j.)$$

Die G_i , $i \in I$, sind die *Zusammenhangskomponenten* von U (vgl. (7), Aufgabe 1.1.27.M).

10.1.9 Definition. Es seien $G_1, G_2 \subset \mathbb{R}^n$ Gebiete und $f: G_1 \rightarrow G_2$ eine Abbildung. Wir nennen f einen *Diffeomorphismus*, wenn f differenzierbar ist und es eine differenzierbare Abbildung $g: G_2 \rightarrow G_1$ mit

$$f \circ g = \text{id}_{G_2} \quad \text{und} \quad g \circ f = \text{id}_{G_1}$$

gibt.

10.1.10 *Bemerkung.* i) Eine differenzierbare Abbildung $f: G_1 \rightarrow G_2$ ist genau dann ein Diffeomorphismus, wenn sie bijektiv ist und die Umkehrabbildung $f^{-1}: G_2 \rightarrow G_1$ ebenfalls differenzierbar ist.

ii) Einen Diffeomorphismus können wir als einen differenzierbaren Koordinatenwechsel auffassen. Diese Sichtweise wird in einigen der folgenden Beispiele im Vordergrund stehen.

10.1.11 *Gegenbeispiel.* Die Abbildung

$$\begin{aligned} f: \mathbb{R} &\longrightarrow \mathbb{R} \\ x &\longmapsto x^3 \end{aligned}$$

ist (beliebig oft) differenzierbar und streng monoton wachsend und somit bijektiv. Die Umkehrfunktion

$$\begin{aligned} f^{-1}: \mathbb{R} &\longrightarrow \mathbb{R} \\ x &\longmapsto \sqrt[3]{x} \end{aligned}$$

ist auf $\mathbb{R} \setminus \{0\}$ differenzierbar, in 0 aber nicht wegen $f'(0) = 0$ (s. (12), Satz 4.2.7). Man beachte, dass f^{-1} im Gegensatz zum Gegenbeispiel 10.1.5 stetig ist. Damit ist f ein **Homöomorphismus aber kein Diffeomorphismus**.

10.1.12 *Beispiele.* i) (Translationen) Für $a \in \mathbb{R}^n$ sei

$$\begin{aligned} T_a: \mathbb{R}^n &\longrightarrow \mathbb{R}^n \\ x &\longmapsto x + a. \end{aligned}$$

Die Abbildung T_a ist offenbar differenzierbar. Die Umkehrabbildung ist T_{-a} und daher differenzierbar. Somit ist T_a ein Diffeomorphismus.

ii) (Lineare Abbildungen) Es sei $A \in \text{Mat}(n, \mathbb{R})$ eine **invertierbare** ($n \times n$)-Matrix. Sie definiert die lineare Abbildung

$$\begin{aligned} f_A: \mathbb{R}^n &\longrightarrow \mathbb{R}^n \\ x &\longmapsto x \cdot A. \end{aligned}$$

Da A invertierbar ist, ist f_A bijektiv. Nach Beispiel 6.1.5, ii), ist f_A differenzierbar. Die Umkehrabbildung zu f_A ist $f_{A^{-1}}$. Sie ist auch differenzierbar. Folglich ist f_A ein Diffeomorphismus.

iii) (Polarkoordinaten) Wir beginnen mit der differenzierbaren Abbildung

$$f: \mathbb{R}_{>0} \times \mathbb{R} \longrightarrow \mathbb{R}^2 \\ (r, \varphi) \longmapsto (r \cdot \cos(\varphi), r \cdot \sin(\varphi)).$$

Das Bild von f ist die offene Menge $\mathbb{R}^2 \setminus \{(0,0)\}$. Allerdings fällt auf, dass f **nicht injektiv** ist:

$$\forall k \in \mathbb{Z} : f(r, \varphi + 2k\pi) = f(r, \varphi).$$

Wie in Beispiel 3.2.4, iv), schränken wir die Abbildung f auf die Menge $\mathbb{R}_{>0} \times [0, 2\pi)$ ein. Wir lassen der Bequemlichkeit halber den Verweis auf die Einschränkung in unserer Bezeichnung weg und erhalten die **bijektive** Abbildung

$$f: \mathbb{R}_{>0} \times [0, 2\pi) \longrightarrow \mathbb{R}^2 \setminus \{(0,0)\}.$$

Nun ist aber $\mathbb{R}_{>0} \times [0, 2\pi)$ **keine offene** Teilmenge von \mathbb{R}^2 und damit auch kein Gebiet, so dass wir die Abbildung f nicht auf „Diffeomorphie“ untersuchen können.

Durch weitere Einschränkung erhalten wir (wieder unter Benutzung einer vereinfachten Bezeichnung) die Abbildung

$$f: \mathbb{R}_{>0} \times (0, 2\pi) \longrightarrow \mathbb{R}^2 \setminus \{(x,0) \mid x \geq 0\}.$$

Dies ist schließlich eine bijektive und differenzierbare Abbildung zwischen Gebieten in \mathbb{R}^2 , und wir können uns endlich der Frage zuwenden, ob die Umkehrabbildung f^{-1} differenzierbar ist. Dazu überdecken wir $\mathbb{R}^2 \setminus \{(x,0) \mid x \geq 0\}$ durch geeignete offene Mengen:

$$\mathbb{R}^2 \setminus \{(x,0) \mid x \geq 0\} = H_+ \cup H_- \cup H_0$$

mit

$$H_+ = \{(x,y) \in \mathbb{R}^2 \mid y > 0\}, \\ H_- = \{(x,y) \in \mathbb{R}^2 \mid y < 0\}, \\ H_0 = \{(x,y) \in \mathbb{R}^2 \mid x < 0\}.$$

Es reicht nachzuweisen, dass die Einschränkung von f^{-1} auf jede dieser offenen Mengen differenzierbar ist. Bevor wir dies tun, erinnern wir an folgende Tatsachen:

1. Das Bild des Arkustangens, der Umkehrfunktion des Tangens, ist das Intervall $(-\pi/2, \pi/2)$ (vgl. (12), Beispiel 3.3.8 und 4.8.16).
2. Für jede reelle Zahl t gilt $\cos(\pi/2 - t) = \sin(t)$ und $\sin(\pi/2 - t) = \cos(t)$ ((12), (4.8)).
3. Für $t \in \mathbb{R}$ hat man³

$$\arctan(t) = \arcsin\left(\frac{t}{\sqrt{1+t^2}}\right) = \arccos\left(\frac{1}{\sqrt{1+t^2}}\right).$$

³Die Gültigkeit dieser Formel prüft man mit den Formeln $\tan = \sin / \cos$ und $\sin^2 + \cos^2 = 1$ nach: So ist z.B. der Wertebereich des Arkussinus das Intervall $[-\pi/2, \pi/2]$ (s. (12), Beispiel 3.3.15). Für $\varphi \in [-\pi/2, \pi/2]$ gilt $\cos(\varphi) \geq 0$ und daher $\cos(\varphi) = \sqrt{1 - \sin^2(\varphi)}$. Für $y \in [-1, 1]$ folgt daher $\tan(\arcsin(y)) = \sin(\arcsin(y)) / \sqrt{1 - \sin^2(\arcsin(y))} = y / \sqrt{1 - y^2}$.

10.1 Homöomorphismen, Gebiete und Diffeomorphismen

Jetzt können wir f^{-1} auf den offenen Mengen H_+ , H_- und H_0 beschreiben. Für einen Vektor $(x, y) \in \mathbb{R}^2$ setzen wir dabei $r(x, y) := \sqrt{x^2 + y^2}$.

a) Beschreibung auf H_+ . Wir bemerken zunächst, dass $(r \cdot \cos(\varphi), r \cdot \sin(\varphi))$, $r \in \mathbb{R}_{>0}$, $\varphi \in (0, 2\pi)$, genau dann in H_+ liegt, wenn $\varphi \in (0, \pi)$. Sei

$$\begin{aligned} \varphi_+ : H_+ &\longrightarrow \mathbb{R} \\ (x, y) &\longmapsto \frac{\pi}{2} - \arctan\left(\frac{x}{y}\right). \end{aligned}$$

Für $(x, y) \in H_+$ gilt $\varphi_+(x, y) \in (0, \pi)$,

$$\begin{aligned} \cos(\varphi_+(x, y)) &= \sin\left(\arctan\left(\frac{x}{y}\right)\right) \\ &= \sin\left(\arcsin\left(\frac{\frac{x}{y}}{\sqrt{1 + \frac{x^2}{y^2}}}\right)\right) \\ &\stackrel{y \geq 0}{=} \sin\left(\arcsin\left(\frac{x}{\sqrt{x^2 + y^2}}\right)\right) \\ &= \frac{x}{\sqrt{x^2 + y^2}} \end{aligned}$$

und

$$\begin{aligned} \sin(\varphi_+(x, y)) &= \cos\left(\arctan\left(\frac{x}{y}\right)\right) \\ &= \cos\left(\arccos\left(\frac{1}{\sqrt{1 + \frac{x^2}{y^2}}}\right)\right) \\ &\stackrel{y \geq 0}{=} \cos\left(\arccos\left(\frac{y}{\sqrt{x^2 + y^2}}\right)\right) \\ &= \frac{y}{\sqrt{x^2 + y^2}}. \end{aligned}$$

Wir sehen:

$$\forall (x, y) \in H_+ : f(r(x, y), \varphi_+(x, y)) = (x, y).$$

Es folgt, dass

$$\begin{aligned} f|_{H_+}^{-1} : H_+ &\longrightarrow \mathbb{R}^2 \\ (x, y) &\longmapsto (r(x, y), \varphi_+(x, y)) \end{aligned}$$

eine differenzierbare Funktion ist.

b) Beschreibung auf H_- . Wir gehen analog zu Fall a) vor. Der Punkt $(r \cdot \cos(\varphi), r \cdot \sin(\varphi))$, $r \in \mathbb{R}_{>0}$, $\varphi \in (0, 2\pi)$, liegt genau dann in H_- , wenn $\varphi \in (\pi, 2\pi)$. Wir definieren

$$\begin{aligned} \varphi_- : H_- &\longrightarrow \mathbb{R} \\ (x, y) &\longmapsto \frac{3}{2} \cdot \pi - \arctan\left(\frac{x}{y}\right). \end{aligned}$$

Für $(x, y) \in H_-$ gilt $\varphi_-(x, y) \in (\pi, 2\pi)$,

$$\begin{aligned} \cos(\varphi_-(x, y)) &= \sin\left(-\pi + \arctan\left(\frac{x}{y}\right)\right) \\ &= -\sin\left(\arcsin\left(\frac{\frac{x}{y}}{\sqrt{1 + \frac{x^2}{y^2}}}\right)\right) \\ &\stackrel{y < 0}{=} -\sin\left(\arcsin\left(\frac{-x}{\sqrt{x^2 + y^2}}\right)\right) \\ &= \frac{x}{\sqrt{x^2 + y^2}} \end{aligned}$$

und entsprechend

$$\sin(\varphi_-(x, y)) = \frac{y}{\sqrt{x^2 + y^2}}.$$

Damit erhalten wir die Beschreibung

$$\begin{aligned} f_{|H_-}^{-1} : H_- &\longrightarrow \mathbb{R}^2 \\ (x, y) &\longmapsto (r(x, y), \varphi_-(x, y)), \end{aligned}$$

die zeigt, dass f^{-1} auf der offenen Menge H_- differenzierbar ist.

c) Beschreibung auf H_0 . Hier liegt $(r \cdot \cos(\varphi), r \cdot \sin(\varphi))$, $r \in \mathbb{R}_{>0}$, $\varphi \in (0, 2\pi)$, genau dann in H_0 , wenn $\varphi \in (\pi/2, (3/2) \cdot \pi)$. Wir definieren

$$\begin{aligned} \varphi_0 : H_0 &\longrightarrow \mathbb{R} \\ (x, y) &\longmapsto \pi - \arctan\left(\frac{y}{x}\right). \end{aligned}$$

Für $(x, y) \in H_0$ finden wir $\varphi_0(x, y) \in (\pi/2, (3/2) \cdot \pi)$,

$$\begin{aligned} \cos(\varphi_0(x, y)) &= -\cos\left(\arctan\left(\frac{y}{x}\right)\right) \\ &= -\cos\left(\arccos\left(\frac{1}{\sqrt{1 + \frac{y^2}{x^2}}}\right)\right) \\ &\stackrel{x < 0}{=} -\cos\left(\arccos\left(\frac{-x}{\sqrt{x^2 + y^2}}\right)\right) \\ &= \frac{x}{\sqrt{x^2 + y^2}} \end{aligned}$$

und analog

$$\sin(\varphi_0(x, y)) = \frac{y}{\sqrt{x^2 + y^2}}.$$

Somit definiert die Vorschrift

$$\begin{aligned} H_0 &\longrightarrow \mathbb{R}^2 \\ (x, y) &\longmapsto (r(x, y), \varphi_0(x, y)), \end{aligned}$$

die Abbildung f^{-1} auf H_0 . Aus dieser Beschreibung folgt, dass f^{-1} auf der offenen Menge H_0 differenzierbar ist.

Insgesamt haben wir bewiesen, dass

$$f: \mathbb{R}_{>0} \times (0, 2\pi) \longrightarrow \mathbb{R}^2 \setminus \{ (x, 0) \mid x \geq 0 \}$$

ein Diffeomorphismus ist.

10.2 Umkehrbarkeit differenzierbarer Abbildungen

Wir wenden uns der bereits angekündigten Verallgemeinerung des Kriteriums für lokale Umkehrbarkeit einer stetig differenzierbaren Funktion einer Veränderlichen zu. Dieses Kriterium erleichtert uns auch den Nachweis, dass gewisse differenzierbare Abbildungen Diffeomorphismen sind (vgl. Beispiel 10.1.12, iii). Im Beweis des Umkehrsatzes spielt der Banachsche Fixpunktsatz wieder eine prominente Rolle.

Wir beginnen mit einer Vorüberlegung: Es seien $G_1, G_2 \subset \mathbb{R}^n$ Gebiete und $f: G_1 \longrightarrow G_2$ ein Diffeomorphismus. Dann ist $f^{-1}: G_2 \longrightarrow G_1$ eine differenzierbare Funktion. Aus $f^{-1} \circ f = \text{id}_{G_1}$, der Kettenregel 6.3.1 und 6.3.2 folgt

$$\forall x \in G_1: \mathbb{E}_n = J_{\text{id}_{G_1}}(x) = J_{f^{-1}}(f(x)) \circ J_f(x). \quad (10.2)$$

Die Matrix $J_f(x)$ ist also für jeden Punkt $x \in G_1$ invertierbar.

10.2.1 Satz (Umkehrsatz, Teil I). *Es seien $G \subset \mathbb{R}^n$ ein Gebiet, $f: G \longrightarrow \mathbb{R}^n$ eine stetig differenzierbare Abbildung, $x_0 \in G$ und $y_0 := f(x_0)$. Wenn*

$$\text{Det}(J_f(x_0)) \neq 0,$$

$J_f(x_0)$ also invertierbar ist, dann existieren offene Mengen $U \subset G$ und $V \subset \mathbb{R}^n$ mit folgenden Eigenschaften:

- $x_0 \in U, y_0 \in V,$
- $\forall x \in U: \text{Det}(J_f(x)) \neq 0,$
- $f: U \longrightarrow V$ ist bijektiv.

Beweis. Wir betrachten die Funktion

$$\mathcal{D}: G \xrightarrow{x \mapsto J_f(x)} \mathbb{R}^{n^2} \cong \text{Mat}(n, \mathbb{R}) \xrightarrow{\text{Det}} \mathbb{R}.$$

Die Zuordnung $x \mapsto J_f(x), x \in G$, ist stetig, weil wir f als stetig differenzierbar voraussetzen. Die Determinantenfunktion ist nach der Leibniz-Regel ((13), §25, Definition 2) eine Polynomfunktion und somit stetig (Beispiel 3.2.4, ii). Also ist \mathcal{D} eine stetige Funktion. Die Menge $U' := \{x \in G \mid \text{Det}(J_f(x)) \neq 0\}$ ist offen und enthält nach Voraussetzung x_0 .

Schritt 1. Wir zeigen, dass es eine offene Menge $U \subset U'$ gibt, so dass $f|_U$ injektiv ist. Wir definieren

$$A := J_f(x_0) \quad \text{und} \quad \lambda := \frac{1}{2 \cdot \|A^{-1}\|_{\text{op}}}. \quad (10.3)$$

Die Funktion

$$\begin{aligned} N: G &\longrightarrow \mathbb{R} \\ x &\longmapsto \|A - J_f(x)\|_{\text{op}} \end{aligned}$$

ist auch stetig. (Warum?) Da $N(x_0) = 0$, können wir eine offene Umgebung $U \subset U'$ von x_0 wählen, so dass

$$\forall x \in U: \|A - J_f(x)\|_{\text{op}} < \lambda.$$

Wir wählen U **konvex**. Sei $y \in \mathbb{R}^n$. Wir setzen

$$\begin{aligned} \varphi_y: \mathbb{R}^n &\longrightarrow \mathbb{R}^n \\ x &\longmapsto x + (y - f(x)) \cdot (A^{-1})^t. \end{aligned}$$

Man beachte:

$$\forall x \in U: y = f(x) \iff \varphi_y(x) = x.$$

Der Ausdruck auf der rechten Seite ist eine **Fixpunktgleichung**. Es genügt zu zeigen, dass φ_y **kontrahierend** ist. Zunächst beobachten wir:

$$\forall x \in U: J_{\varphi_y}(x) = E_n - A^{-1} \cdot J_f(x) = A^{-1} \cdot (A - J_f(x)),$$

so dass

$$\forall x \in U: \|J_{\varphi_y}(x)\|_{\text{op}} = \|A^{-1}\|_{\text{op}} \cdot \|A - J_f(x)\|_{\text{op}} < \frac{1}{2}.$$

Da U konvex ist, können wir den Mittelwertsatz 8.3.2 anwenden. Dabei gilt⁴

$$\|D\varphi_y(x)\|_{\text{op}} = \|(J_{\varphi_y}(x))^t\|_{\text{op}} = \|J_{\varphi_y}(x)\|_{\text{op}}, \quad x \in \mathbb{R}^n.$$

Für Punkte $x_1, x_2 \in U$ ergibt sich

$$\|\varphi_y(x_1) - \varphi_y(x_2)\| \leq \frac{1}{2} \cdot \|x_1 - x_2\|.$$

Daraus folgt leicht,⁵ dass es höchstens einen Punkt $x \in U$ mit $x = \varphi_y(x)$, d.h. $y = f(x)$, gibt. Somit ist $f|_U$ injektiv.

Schritt 2. Wir weisen jetzt nach, dass $V := f(U)$ offen ist. Hierzu werden wir uns des Banachschen Fixpunktsatzes bedienen. Wir betrachten $x_1 \in U$ und $y_1 := f(x_1) \in V$. Da U eine offene Umgebung von x_1 ist, gibt es ein $r > 0$, so dass

$$\bar{B} := \overline{B(x_1, r)} \subset U.$$

Behauptung. Für die Zahl λ aus (10.3) gilt

$$B(y_1, \lambda \cdot r) \subset V.$$

Zum Beweis dieser Behauptung betrachten wir die zu einem Punkt $y \in B(y_1, \lambda \cdot r)$ gehörige Abbildung φ_y . Wir berechnen

$$\begin{aligned} \|\varphi_y(x_1) - x_1\| &= \|(y - y_1) \cdot (A^{-1})^t\| \\ &< \|(A^{-1})^t\|_{\text{op}} \cdot \lambda \cdot r \\ &= \|A^{-1}\|_{\text{op}} \cdot \lambda \cdot r \\ &= \frac{r}{2}. \end{aligned}$$

⁴Für jede Matrix $B \in \text{Mat}(n, \mathbb{R})$ hat man $\|B^t\|_{\text{op}} = \|B\|_{\text{op}}$: Für jeden Vektor $v \in \mathbb{R}^n$ mit $\|v\| = 1$ gilt

$$\|v \cdot B\|^2 = \langle v \cdot B, v \cdot B \rangle = v \cdot B \cdot B^t \cdot v^t = \langle v \cdot B \cdot B^t, v \rangle \leq \|v \cdot B \cdot B^t\| \leq \|v \cdot B\| \cdot \|B^t\|_{\text{op}}.$$

Dabei haben wir die Cauchy-Schwarz-Ungleichung (Seite 76) benutzt. Wir sehen

$$\|B^t\|_{\text{op}} \geq \sup\{\|v \cdot B\| \mid v \in \mathbb{R}^n, \|v\| = 1\} = \|B\|_{\text{op}}.$$

Aus Symmetriegründen folgt $\|B^t\|_{\text{op}} = \|B\|_{\text{op}}$.

⁵Vgl. Beweis von Satz 8.2.1.

10.2 Umkehrbarkeit differenzierbarer Abbildungen

Mit den in Schritt 1 behandelten Abschätzungen finden wir für einen Punkt $x \in \bar{B}$:

$$\begin{aligned} \|\varphi_Y(x) - x_1\| &\leq \|\varphi_Y(x) - \varphi_Y(x_1)\| + \|\varphi_Y(x_1) - x_1\| \\ &\leq \frac{1}{2} \cdot \|x - x_1\| + \frac{r}{2} \\ &= r. \end{aligned}$$

Damit folgern wir $\varphi_Y(x) \in \bar{B}$, und nach Schritt 1 ist $\varphi_{Y|\bar{B}}: \bar{B} \rightarrow \bar{B}$ eine kontrahierende Abbildung. Mit der durch die euklidische Metrik induzierten Metrik wird \bar{B} zu einem vollständigen metrischen Raum. (Warum?) Der Banachsche Fixpunktsatz 8.2.1 zeigt die Existenz eines Punkts $x \in \bar{B}$ mit $\varphi_Y(x) = x$, d.h. $y = f(x)$. Damit ist $y \in f(U) = V$ gezeigt, und wir haben $B(y_1, \lambda \cdot r) \subset V$ bewiesen. \square

10.2.2 Bemerkung. i) Der Beweis liefert ein numerisches Verfahren, um $f^{-1}(y)$, $y \in V$, zu bestimmen.

ii) Der Beweis zeigt, dass eine stetig differenzierbare Abbildung $f: G \rightarrow \mathbb{R}^n$, für die $\text{Det}(J_f(x)) \neq 0$, $x \in G$, gilt, eine offene Abbildung ist.

10.2.3 Satz (Umkehrsatz, Teil II). *In der Situation von Satz 10.2.1 gilt weiter:*

i) Die Umkehrabbildung $f^{-1}: V \rightarrow U$ ist differenzierbar.

ii) Für $y \in V$ und $x = f^{-1}(y)$ gilt

$$J_{f^{-1}}(y) = (J_f(x))^{-1}.$$

Beweis. Teil ii) folgt aus der Kettenregel (vgl. (10.2)), ergibt sich aber auch aus den folgenden Argumenten.

Teil i). Seien $y_1 \in V$ und $x_1 := f^{-1}(y_1) \in U$. Wir möchten zeigen, dass f^{-1} in y_1 differenzierbar ist. Es seien $f_i: U \rightarrow \mathbb{R}$ die Komponentenfunktionen von f , $i = 1, \dots, n$ (s. S. 24). Im Beweis von Satz 5.2.3 wurde die Existenz von Abbildungen $\delta_i: U \rightarrow \mathbb{R}^n$ nachgewiesen, die in x_1 stetig sind und $\delta_i(x_1) = 0$ erfüllen, so dass

$$f_i(x) - f_i(x_1) = (x - x_1) \cdot \text{Grad} f_i(x_1)^t + (x - x_1) \cdot \delta_i(x)^t, \quad x \in U, i = 1, \dots, n.$$

Wir schreiben $\delta_i = (\delta_{i1}, \dots, \delta_{in})$ mit den Komponentenfunktionen δ_{ij} von δ_i und setzen

$$\Delta_{ij}(x) := \frac{\delta f_i}{\delta x_j}(x) + \delta_{ij}(x), \quad x \in U, i, j = 1, \dots, n.$$

Mit diesen Zutaten führen wir die Abbildung

$$\begin{aligned} \Delta: U &\rightarrow \text{Mat}(n, \mathbb{R}) \\ x &\mapsto (\Delta_{ij}(x))_{i,j=1,\dots,n}^t = \begin{pmatrix} \Delta_{11}(x) & \cdots & \Delta_{1n}(x) \\ \vdots & & \vdots \\ \Delta_{n1}(x) & \cdots & \Delta_{nn}(x) \end{pmatrix} \end{aligned}$$

ein. Sie ist in x_1 stetig, und es gilt:

$$\forall x \in U: f(x) - f(x_1) = (x - x_1) \cdot \Delta(x).$$

Damit ist auch die Abbildung

$$d: U \xrightarrow{\Delta} \text{Mat}(n, \mathbb{R}) \xrightarrow{\text{Det}} \mathbb{R}$$

in x_1 stetig. Da

$$d(x_1) = \text{Det}(J_f(x_1)) \neq 0,$$

können wir U verkleinern und annehmen, dass

$$\forall x \in U: \quad d(x) \neq 0.$$

Unter dieser Voraussetzung können wir wiederum die Abbildung

$$\begin{aligned} \Delta^*: U &\longrightarrow \text{Mat}(n, \mathbb{R}) \\ x &\longmapsto (\Delta(x))^{-1} \end{aligned}$$

erklären. Nach der Cramerschen⁶ Regel ((13), §27, Satz 1) sind die Einträge von $(\Delta(x))^{-1}$ Polynome in den Koeffizienten von $\Delta(x)$ dividiert durch $d(x)$. Es folgt, dass Δ^* in x_1 stetig ist. Man überprüft:

$$\forall y \in V: \quad f^{-1}(y) - f^{-1}(y_1) = (y - y_1) \cdot \Delta^*(f^{-1}(y)).$$

Somit können wir endlich

$$\begin{aligned} 0 &\leq \lim_{y \rightarrow y_1} \frac{\|f^{-1}(y) - f^{-1}(y_1) - (y - y_1) \cdot \Delta^*(f^{-1}(y_1))\|}{\|y - y_1\|} \\ &\leq \lim_{y \rightarrow y_1} \|\Delta^*(f^{-1}(y)) - \Delta^*(f^{-1}(y_1))\|_{\text{op}} = 0 \end{aligned} \quad (10.4)$$

folgern. Dabei haben wir benutzt, dass f^{-1} nach Lemma 10.1.4 und Bemerkung 10.2.2, ii), stetig ist, so dass $\lim_{y \rightarrow y_1} f^{-1}(y) = f^{-1}(y_1) = x_1$ folgt, und dass Δ^* in x_1 stetig ist. Gleichung (10.4) zeigt nach Definition 6.1.1, a), dass f^{-1} in y_1 differenzierbar ist. \square

10.2.4 Bemerkung (Höhere Differenzierbarkeit). Es seien $B \subset \mathbb{R}^m$ eine offene Teilmenge, $k \geq 1$ und

$$\begin{aligned} f = (f_1, \dots, f_n): B &\longrightarrow \mathbb{R}^n \\ x &\longmapsto (f_1(x), \dots, f_n(x)) \end{aligned}$$

eine Abbildung. Wir sagen, dass f k -mal stetig differenzierbar ist, wenn die Komponentenfunktionen f_1, \dots, f_n k -mal stetig (partiell) differenzierbar sind (s. Aufgabe A.7.3). Wir vereinbaren auch, dass 0-mal stetig differenzierbar einfach stetig bedeutet. Sei $k \geq 1$. Eine Abbildung f ist genau dann k -mal stetig differenzierbar, wenn sie differenzierbar ist und die matrixwertige Abbildung

$$\begin{aligned} J_f: B &\longrightarrow \text{Mat}(n, m, \mathbb{R}) \cong \mathbb{R}^{n \cdot m} \\ x &\longmapsto J_f(x) \end{aligned}$$

$(k - 1)$ -mal stetig differenzierbar ist.

Seien $B \subset \mathbb{R}^m$ eine offene Teilmenge, $f, g: B \rightarrow \mathbb{R}$ zwei Funktionen und $k \geq 0$. Wenn f und g k -mal stetig differenzierbar sind, dann sind auch die Funktionen $f + g$ und $f \cdot g$ k -mal stetig differenzierbar. Gilt zusätzlich $g(x) \neq 0$ für alle $x \in B$, dann ist auch die Funktion f/g k -mal stetig differenzierbar. (Das beweist man leicht induktiv: Für $k = 0$ ist das Satz 3.2.2. Für $k \rightarrow k + 1$ benutzt man die Summen-, Produkt- bzw. Quotientenregel (Eigenschaften 5.1.5 oder (12), Satz 4.2.1).)

Sei wieder $k \geq 1$. Aus der vorigen Beobachtung, der Definition der Matrixmultiplikation und der Kettenregel folgt, dass die Verknüpfung k -mal stetig differenzierbarer Abbildungen ebenfalls k -mal stetig differenzierbar ist.

Hier interessieren wir uns für folgende Situation: Seien $U, V \subset \mathbb{R}^n$ Gebiete und $f: U \rightarrow V$ eine bijektive differenzierbare Abbildung, so dass

$$\forall x \in B: \quad \text{Det}(J_f(x)) \neq 0.$$

⁶Gabriel Cramer (1704 - 1752), Schweizer Mathematiker.

Behauptung. Sei $k \geq 1$. Wenn f k -mal stetig differenzierbar ist, dann ist auch die Umkehrabbildung f^{-1} k -mal stetig differenzierbar.

Diese Behauptung beweisen wir durch Induktion über k . Der Fall $k = 1$ folgt aus Satz 10.2.1 und 10.2.3.

$k \rightarrow k + 1$. Wir müssen zeigen, dass die Abbildung

$$\begin{aligned} J_{f^{-1}}: V &\longrightarrow \text{Mat}(n, \mathbb{R}) \\ y &\longmapsto J_{f^{-1}}(y) \stackrel{\text{Satz 10.2.3}}{=} \left(J_f(f^{-1}(y)) \right)^{-1} \end{aligned}$$

k -mal stetig differenzierbar ist. Mit

$$\begin{aligned} \tilde{J}_f: U &\longrightarrow \text{Mat}(n, \mathbb{R}) \\ x &\longmapsto (J_f(x))^{-1} \end{aligned}$$

gilt

$$J_{f^{-1}} = \tilde{J}_f \circ f^{-1}. \quad (10.5)$$

Wir wissen, dass J_f k -mal stetig differenzierbar ist. Aus der Leibniz-Formel für Determinanten ((13), §25, Definition 2), der Cramerschen Regel ((13), §27, Satz 1) und unseren Vorbemerkungen ergibt sich, dass \tilde{J}_f auch k -mal stetig differenzierbar ist. Die Induktionsvoraussetzung besagt, dass f^{-1} k -mal stetig differenzierbar ist. Da, wie zuvor dargelegt, die Verknüpfung k -mal stetig differenzierbarer Abbildungen k -mal stetig differenzierbar ist, folgern wir aus (10.5), dass $J_{f^{-1}}$ k -mal stetig differenzierbar ist. \checkmark

10.2.5 Beispiele. i) (Polarkoordinaten) Wir betrachten wieder die Abbildung

$$\begin{aligned} f: \mathbb{R}^2 &\longrightarrow \mathbb{R}^2 \\ (r, \varphi) &\longmapsto (r \cdot \cos(\varphi), r \cdot \sin(\varphi)). \end{aligned}$$

Die Berechnung der Jacobi-Matrix ergibt

$$\forall (r, \varphi) \in \mathbb{R}^2: J_f(r, \varphi) = \begin{pmatrix} \cos(\varphi) & -r \cdot \sin(\varphi) \\ \sin(\varphi) & r \cdot \cos(\varphi) \end{pmatrix}$$

und

$$\text{Det}(J_f(r, \varphi)) = r \cdot (\cos^2(\varphi) + \sin^2(\varphi)) = r.$$

Nach Satz 10.2.1 und 10.2.3 ist die Abbildung

$$\begin{aligned} g: \mathbb{R}_{>0} \times \mathbb{R} &\longrightarrow \mathbb{R}^2 \setminus \{(0, 0)\} \\ (r, \varphi) &\longmapsto (r \cdot \cos(\varphi), r \cdot \sin(\varphi)) \end{aligned}$$

ein *lokaler Diffeomorphismus*, d.h. zu jedem Punkt $(r, \varphi) \in \mathbb{R}_{>0} \times \mathbb{R}$ gibt es eine offene, wegzusammenhängende Umgebung U , so dass $g|_U: U \rightarrow g(U)$ ein Diffeomorphismus zwischen Gebieten in \mathbb{R}^2 ist. Die Abbildung g ist aber **nicht injektiv!**

Da

$$\begin{aligned} h: \mathbb{R}_{>0} \times (0, 2\pi) &\longrightarrow \mathbb{R}^2 \setminus \{(x, 0) \mid x \geq 0\} \\ (r, \varphi) &\longmapsto (r \cdot \cos(\varphi), r \cdot \sin(\varphi)) \end{aligned}$$

zudem bijektiv ist, folgt sofort, dass h^{-1} differenzierbar ist. Wir brauchen also nicht wie in Beispiel 10.1.12, iii), mühsam die Umkehrabbildung zu studieren.

ii) (Zylinderkoordinaten) Auf Grund der Ergebnisse über Polarkoordinaten ist die Abbildung

$$f: \underbrace{\mathbb{R}_{>0} \times (0, 2\pi) \times \mathbb{R}}_{=:G} \longrightarrow \mathbb{R}^3 \setminus \{(x, 0, z) \mid x \geq 0\}$$

$$(r, \varphi, z) \longmapsto (r \cdot \cos(\varphi), r \cdot \sin(\varphi), z)$$

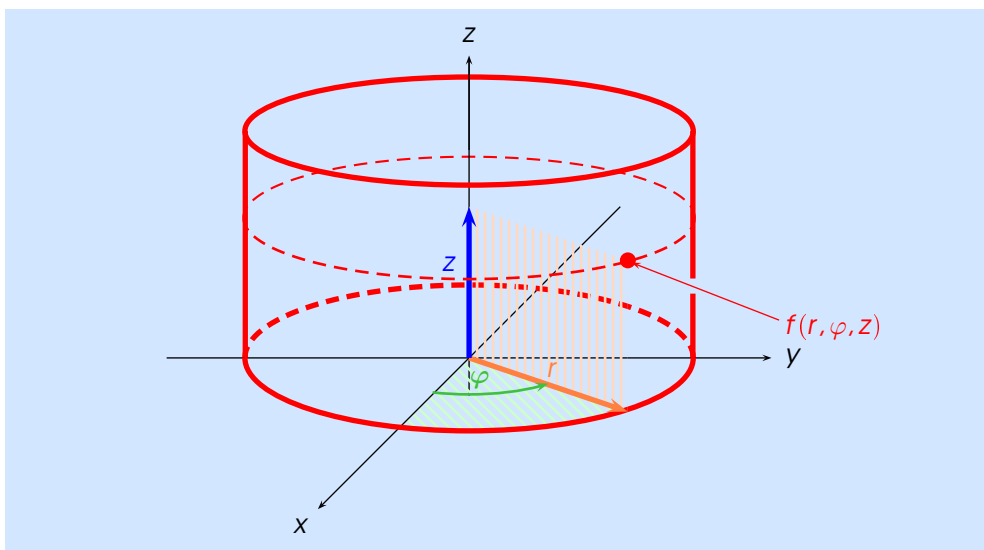
bijektiv. Wir berechnen

$$\forall (r, \varphi, z) \in G: J_f(r, \varphi, z) = \begin{pmatrix} \cos(\varphi) & -r \cdot \sin(\varphi) & 0 \\ \sin(\varphi) & r \cdot \cos(\varphi) & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

und

$$\text{Det}(J_f(r, \varphi, z)) = r \neq 0.$$

Nach Satz 10.2.1 und 10.2.3 ist f^{-1} differenzierbar und somit f ein Diffeomorphismus.

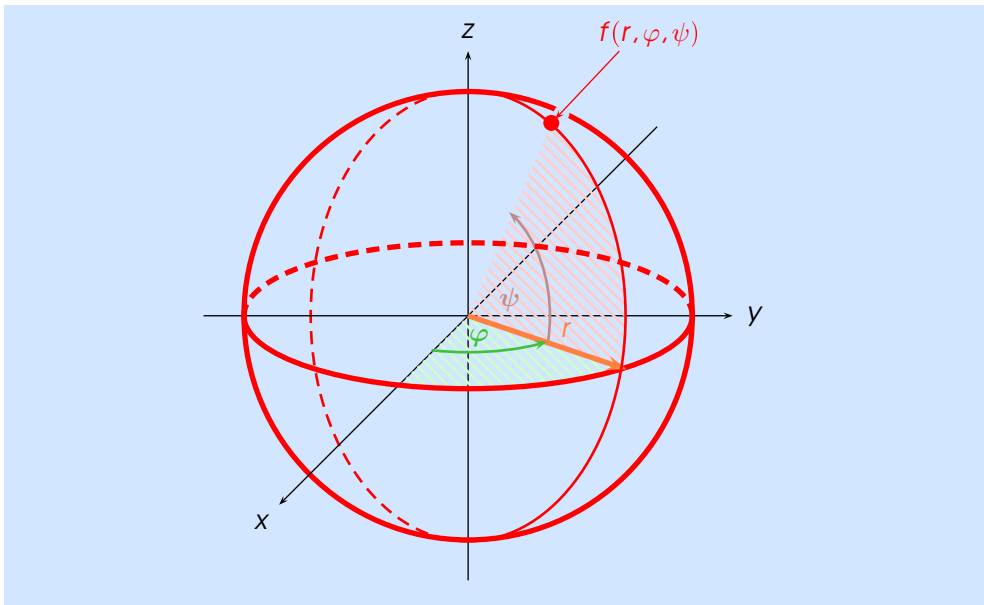


iii) (Kugelkoordinaten) Man vergewissert sich ebenfalls leicht, dass die Abbildung

$$f: \underbrace{\mathbb{R}_{>0} \times (0, 2\pi) \times \left(-\frac{\pi}{2}, \frac{\pi}{2}\right)}_{=:G} \longrightarrow \mathbb{R}^3 \setminus \{(x, 0, z) \mid x \geq 0\}$$

$$(r, \varphi, \psi) \longmapsto (r \cdot \cos(\varphi) \cdot \cos(\psi), r \cdot \sin(\varphi) \cdot \cos(\psi), r \cdot \sin(\psi))$$

bijektiv ist.



Für $(r, \varphi, \psi) \in G$ gilt

$$J_f(r, \varphi, \psi) = \begin{pmatrix} \cos(\varphi) \cdot \cos(\psi) & -r \cdot \sin(\varphi) \cdot \cos(\psi) & -r \cdot \cos(\varphi) \cdot \sin(\psi) \\ \sin(\varphi) \cdot \cos(\psi) & r \cdot \cos(\varphi) \cdot \cos(\psi) & -r \cdot \sin(\varphi) \cdot \sin(\psi) \\ \sin(\psi) & 0 & r \cdot \cos(\psi) \end{pmatrix}.$$

Somit folgt

$$\begin{aligned} \text{Det}(J_f(r, \varphi, \psi)) &= r^2 \cdot \sin(\psi) \cdot \text{Det} \begin{pmatrix} -\sin(\varphi) \cdot \cos(\psi) & -\cos(\varphi) \cdot \sin(\psi) \\ \cos(\varphi) \cdot \cos(\psi) & -\sin(\varphi) \cdot \sin(\psi) \end{pmatrix} + \\ &\quad + r^2 \cdot \cos(\psi) \cdot \text{Det} \begin{pmatrix} \cos(\varphi) \cdot \cos(\psi) & -\sin(\varphi) \cdot \cos(\psi) \\ \sin(\varphi) \cdot \cos(\psi) & \cos(\varphi) \cdot \cos(\psi) \end{pmatrix} \\ &= r^2 \cdot (\sin^2(\varphi) \cdot \sin^2(\psi) \cdot \cos(\psi) + \cos^2(\varphi) \cdot \sin^2(\psi) \cdot \cos(\psi) + \\ &\quad + \cos^2(\varphi) \cdot \cos^3(\psi) + \sin^2(\varphi) \cdot \cos^3(\psi)) \\ &= r^2 \cdot (\sin^2(\psi) \cdot \cos(\psi) + \cos^3(\psi)) \\ &= r^2 \cdot \cos(\psi). \end{aligned}$$

Da $\psi \in (-\pi/2, \pi/2)$, gilt $\cos(\psi) > 0$. Damit folgt

$$\forall (r, \varphi, \psi) \in G: \text{Det}(J_f(r, \varphi, \psi)) > 0,$$

so dass f ein Diffeomorphismus ist.

Kapitel 11

Der Satz über implizite Funktionen

11.1 Vorbereitende Beispiele

In diesem Kapitel werden wir Abbildungen

$$\begin{aligned} f: \mathbb{R}^m &\longrightarrow \mathbb{R}^n \\ x &\longmapsto (f_1(x), \dots, f_n(x)) \end{aligned}$$

mit $m > n$ untersuchen. Sei $a = (a_1, \dots, a_n) \in \mathbb{R}^n$. Wir wollen die Menge

$$f^{-1}(a) := \{x \in \mathbb{R}^m \mid f(x) = a\}$$

studieren.¹

Die Bedingung $x \in f^{-1}(a)$ können wir als Gleichungssystem

$$\begin{aligned} f_1(x_1, \dots, x_m) &= a_1 \\ &\vdots \\ f_n(x_1, \dots, x_m) &= a_n \end{aligned} \tag{11.1}$$

$$\tag{11.2}$$

lesen.

Seien zunächst alle Gleichungen linear. Wenn die Koeffizientenmatrix, die das Gleichungssystem definiert, den maximal möglichen Rang n hat, dann bilden die Lösungen einen affinen Raum, der in der Form $U + p_0$ mit einem $(m - n)$ -dimensionalen Vektorraum $U \subset \mathbb{R}^m$ und einer speziellen Lösung p_0 des Gleichungssystems geschrieben werden kann. Insbesondere können die Lösungen durch \mathbb{R}^{m-n} parametrisiert werden.

Wir wollen im nichtlinearen Fall ein ähnliches Resultat beweisen: Für einen Punkt $p_0 \in f^{-1}(a)$ suchen wir eine offene Umgebung $W \subset \mathbb{R}^m$ von p_0 , Koordinaten $(y_1, \dots, y_k, z_1, \dots, z_n)$, $k := m - n$, eine offene Menge $U \subset \mathbb{R}^k$ und eine Abbildung

$$F: U \subset \mathbb{R}^k \longrightarrow \mathbb{R}^n,$$

so dass

$$f^{-1}(a) \cap W = \left\{ (y_1, \dots, y_k, F(y_1, \dots, y_k)) \mid (y_1, \dots, y_k) \in U \right\}.$$

¹Durch Übergang zu der Abbildung $g: \mathbb{R}^m \longrightarrow \mathbb{R}^n, x \longmapsto f(x) - a$, können wir uns stets auf den Fall $a = 0$ zurückziehen.

Damit haben wir $f^{-1}(a)$ in einer Umgebung von p_0 als Graphen einer Abbildung von einer offenen Teilmenge des \mathbb{R}^k in den \mathbb{R}^n dargestellt. Die Abbildung F ist *implizit durch die Gleichung $f(y, z) = a$ definiert*.

Wie in der linearen Situation werden wir eine gewisse Rangbedingung benötigen, damit die geschilderte Parametrisierung möglich ist.

Beispiel 1

Wir betrachten

$$\begin{aligned} f: \mathbb{R}^2 &\longrightarrow \mathbb{R} \\ (x, y) &\longmapsto x^2 + y^2. \end{aligned}$$

Für $a < 0$ ist $f^{-1}(a)$ leer, es gilt $f^{-1}(0) = (0, 0)$, und für $a > 0$ ist $f^{-1}(a)$ ein Kreis vom Radius \sqrt{a} . Insbesondere ist

$$f^{-1}(1) = \left\{ (x, y) \in \mathbb{R}^2 \mid x^2 + y^2 = 1 \right\} = S^1$$

der Einheitskreis. Es ist sofort klar, dass wir S^1 nicht global als Graphen einer Funktion darstellen können. Lokal ist dies jedoch möglich. Sei dazu $p_0 = (c, d) \in S^1$. Wir nehmen zunächst $c \neq \pm 1$ an. Für $d > 0$ ist

$$W_+ := \left\{ (x, y) \in S^1 \mid y > 0 \right\}$$

eine offene Umgebung von p_0 , und es gilt

$$\forall (x, y) \in \mathbb{R}^2: (x, y) \in W_+ \iff x \in (-1, 1) \wedge y = \sqrt{1 - x^2}.$$

Mit der Funktion

$$\begin{aligned} F_+: (-1, 1) &\longrightarrow \mathbb{R} \\ x &\longmapsto \sqrt{1 - x^2} \end{aligned}$$

erhalten wir

$$W_+ = \left\{ (x, F_+(x)) \mid x \in (-1, 1) \right\}.$$

Wenn $d < 0$ gilt, dann ist (c, d) in der offenen Teilmenge

$$W_- := \left\{ (x, y) \in S^1 \mid y < 0 \right\} \subset S^1$$

enthalten, und mit der Funktion

$$\begin{aligned} F_-: (-1, 1) &\longrightarrow \mathbb{R} \\ x &\longmapsto -\sqrt{1 - x^2} \end{aligned}$$

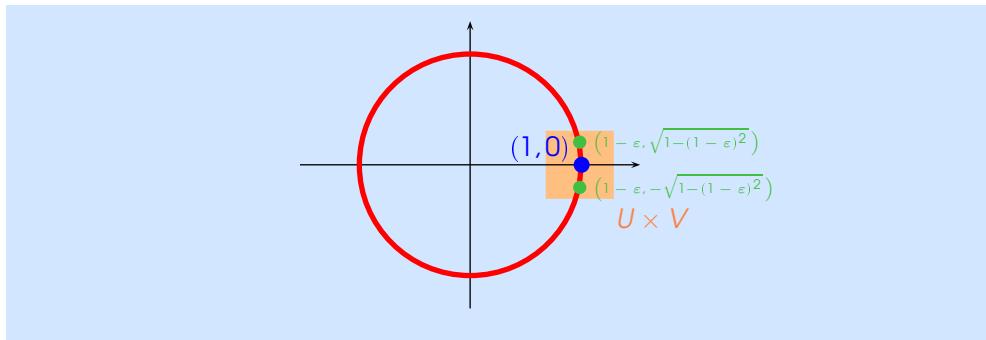
ergibt sich

$$W_- = \left\{ (x, F_-(x)) \mid x \in (-1, 1) \right\}.$$

11.1.1 Bemerkung. In den Punkten $(\pm 1, 0)$ ist die Gleichung $f(x, y) = 0$ nicht nach y auflösbar: Es seien $U, V \subset \mathbb{R}$ offene Intervalle, $W = U \times V$ und $F: U \rightarrow V$ eine Funktion, so dass

$$\left\{ (x, y) \in W \mid f(x, y) = 1 \right\} = \left\{ (x, F(x)) \mid x \in U \right\}. \quad (11.3)$$

Wir behaupten, dass dann $(\pm 1, 0) \notin W$. Wenn z.B. $(1, 0) \in W$, dann ist W eine offene Umgebung von $(1, 0)$. Man sieht leicht, dass es ein $\varepsilon > 0$ gibt, so dass $(1 - \varepsilon, \pm\sqrt{1 - (1 - \varepsilon)^2}) \in W$.



Dann kann aber (11.3) nicht erfüllt sein.

Wenn wir F als differenzierbar voraussetzen, dann können wir auch folgendermaßen argumentieren: Mit (11.3) gilt

$$\forall x \in U: f(x, F(x)) = 1.$$

Mit der Funktion

$$\begin{aligned} g: U &\longrightarrow W \\ x &\longmapsto (x, F(x)) \end{aligned}$$

und der Kettenregel 6.3.1 berechnen wir

$$\begin{aligned} \forall x \in U: 0 &= (f \circ g)'(x) = \langle \text{Grad}f(x, F(x)), g'(x) \rangle && (11.4) \\ &= \frac{\partial f}{\partial x}(x, F(x)) \cdot 1 + \frac{\partial f}{\partial y}(x, F(x)) \cdot F'(x). \end{aligned}$$

Falls $(x, F(x)) = (\pm 1, 0)$, so haben wir

$$\begin{aligned} \frac{\partial f}{\partial x}(\pm 1, 0) &= \pm 2, \\ \frac{\partial f}{\partial y}(\pm 1, 0) &= 0. \end{aligned}$$

Damit erkennt man sofort, dass (11.4) nicht lösbar ist.

Auf der anderen Seite kann man

$$H_{\pm} := \{ (x, y) \in S^1 \mid \pm x > 0 \}$$

und

$$\begin{aligned} G_{\pm}: (-1, 1) &\longrightarrow \mathbb{R} \\ y &\longmapsto \pm \sqrt{1 - y^2} \end{aligned}$$

setzen und erhält

$$H_{\pm} = \{ (G_{\pm}(y), y) \mid y \in (-1, 1) \}.$$

11.1.2 Bemerkung. i) Wir können Bemerkung 11.1.1 auch auf eine beliebige differenzierbare Funktion

$$f: \mathbb{R}^2 \longrightarrow \mathbb{R}$$

ausdehnen. Wir betrachten

$$N(f) := \{ (x, y) \in \mathbb{R}^2 \mid f(x, y) = 0 \}.$$

Das Analogon zu Gleichung (11.4) zeigt folgendes:

- Gelten $(c, d) \in N(f)$, $(\partial f / \partial x)(c, d) \neq 0$ und $(\partial f / \partial y)(c, d) = 0$, dann lässt sich die Gleichung $f(x, y) = 0$ in der Nähe von (c, d) nicht nach y auflösen.
- Gelten $(c, d) \in N(f)$, $(\partial f / \partial y)(c, d) \neq 0$ und $(\partial f / \partial x)(c, d) = 0$, dann lässt sich die Gleichung $f(x, y) = 0$ in der Nähe von (c, d) nicht nach x auflösen.
- Falls in (11.4)

$$\frac{\partial f}{\partial y}(x, F(x)) \neq 0$$

gilt, dann folgt

$$F'(x) = -\frac{\frac{\partial f}{\partial x}(x, F(x))}{\frac{\partial f}{\partial y}(x, F(x))},$$

d.h. die Ableitung der implizit definierten Funktion F kann aus den partiellen Ableitungen der Funktion f berechnet werden. Man spricht in diesem Zusammenhang von *impliziter Differentiation*.

ii) Gelten $(c, d) \in N(f)$ und $\text{Grad}f(c, d) = (0, 0)$, dann lässt sich keine allgemeingültige Aussage treffen: Man betrachte z.B. $f(x, y) = y^2$. Die Menge $N(f)$ ist die x -Achse, also der Graph der konstanten Funktion Null, und $\text{Grad}f(0, 0) = (0, 0)$.

Beispiel 2

Wir wollen die bereits erwähnte implizite Differentiation an einem Beispiel illustrieren. Dazu seien

$$\begin{aligned} f: \mathbb{R}^2 &\longrightarrow \mathbb{R} \\ (x, y) &\longmapsto x^3 + y^3 - 6xy \end{aligned}$$

und

$$N(f) := \left\{ (x, y) \in \mathbb{R}^2 \mid f(x, y) = 0 \right\}$$

das *kartesische Blatt*.

11.1.3 Bemerkung. Wir haben $\text{Grad}f(0, 0) = (0, 0)$. Eine Auflösung der Gleichung um den Ursprung herum ist nicht möglich. Da $N(f)$ invariant unter der Vertauschung von x und y ist,² d.h. $(x, y) \in N(f)$ genau dann gilt, wenn $(y, x) \in N(f)$, reicht es, sich zu überlegen, dass keine Auflösung nach y möglich ist. Zu jedem $\varepsilon > 0$ gibt es drei verschiedene Zahlen, die der Gleichung $y^3 - 6\varepsilon \cdot y + \varepsilon^3 = 0$ genügen.³ Da für $\varepsilon = 0$ die Zahl 0 eine dreifache Nullstelle ist, liegen die Nullstellen für ε nahe bei Null ebenfalls nahe bei Null. Damit ist $N(f) \cap [0, \eta] \times \mathbb{R}$ für kein $\eta > 0$ Graph einer Funktion $F: [0, \eta] \longrightarrow \mathbb{R}$.

Wir nehmen an, dass $U, V \subset \mathbb{R}$ offene Intervalle sind und $F: U \longrightarrow V$ eine differenzierbare Funktion, so dass

$$N(f) \cap (U \times V) = \left\{ (x, F(x)) \mid x \in U \right\}.$$

²Nur sehr wenige Funktionsgraphen haben diese Eigenschaft. Welche sind das?

³Beweis als Übung.

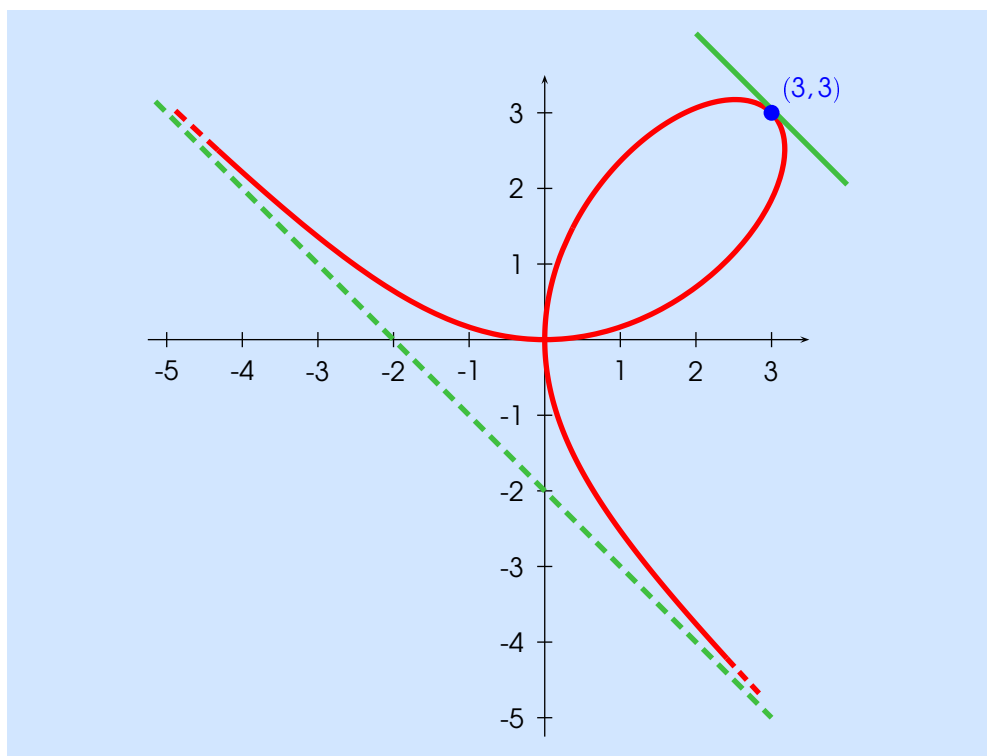


Abbildung 11.1: Das kartesische Blatt und implizite Differentiation

Dann gilt:

$$\forall x \in U: \quad x^3 + F(x)^3 = 6x \cdot F(x).$$

Differentiation nach x liefert

$$3x^2 + 3 \cdot F'(x) \cdot F(x)^2 = 6 \cdot F(x) + 6 \cdot x \cdot F'(x),$$

so dass

$$F'(x) = -\frac{3x^2 - 6 \cdot F(x)}{3 \cdot F(x)^2 - 6x} = -\frac{x^2 - 2 \cdot F(x)}{F(x)^2 - 2x}.$$

Man findet z.B. für den Punkt $(3, 3) \in N(f)$, dass

$$F'(3) = -\frac{9 - 6}{9 - 6} = -1.$$

11.1.4 Bemerkung. Der Vektor $(1, -1)$ ist ein Richtungsvektor der Tangente an die Kurve $N(f)$ im Punkt $(3, 3)$ (s. Abbildung 11.1).

11.2 Der Satz über implizite Funktionen

Wir formulieren nun die Voraussetzungen für den Satz über implizite Funktionen. Es seien $k, m, n \in \mathbb{N}$, so dass $k > 0$ und $m = k + n$. Weiter seien $G \subset \mathbb{R}^m = \mathbb{R}^k \times \mathbb{R}^n$ ein Gebiet und

$$\begin{aligned} f: G &\longrightarrow \mathbb{R}^n \\ p &\longmapsto (f_1(p), \dots, f_n(p)) \end{aligned}$$

eine differenzierbare Abbildung. Wir schreiben

$$p = (x_1, \dots, x_k, y_1, \dots, y_n) =: (x, y), \quad p \in \mathbb{R}^m,$$

und setzen

$$\frac{\partial f}{\partial x}(x, y) := \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(x, y) & \cdots & \frac{\partial f_1}{\partial x_k}(x, y) \\ \vdots & & \vdots \\ \frac{\partial f_n}{\partial x_1}(x, y) & \cdots & \frac{\partial f_n}{\partial x_k}(x, y) \end{pmatrix},$$

$$\frac{\partial f}{\partial y}(x, y) := \begin{pmatrix} \frac{\partial f_1}{\partial y_1}(x, y) & \cdots & \frac{\partial f_1}{\partial y_n}(x, y) \\ \vdots & & \vdots \\ \frac{\partial f_n}{\partial y_1}(x, y) & \cdots & \frac{\partial f_n}{\partial y_n}(x, y) \end{pmatrix},$$

so dass

$$J_f(x, y) = \left(\frac{\partial f}{\partial x}(x, y) \mid \frac{\partial f}{\partial y}(x, y) \right), \quad (x, y) \in G.$$

Schließlich sei

$$N(f) := \left\{ p = (x, y) \in G \mid f(x, y) = 0 \right\}.$$

Die Tangentialvektoren von $N(f)$ an den Punkt $p = (x, y) \in N(f)$ sind die Vektoren $(c, d) \in \mathbb{R}^k \times \mathbb{R}^m$ mit

$$J_f(x, y) \cdot \begin{pmatrix} c^t \\ d^t \end{pmatrix} = 0.$$

Wir nehmen an, der Raum der Tangentialvektoren habe die (minimal mögliche) Dimension k . Wenn es um p herum eine Auflösung der Form

$$y = F(x)$$

gibt, also offene Mengen $U \subset \mathbb{R}^k$, $V \subset \mathbb{R}^m$ und eine differenzierbare Abbildung

$$F: U \rightarrow V,$$

so dass

$$p \in W := U \times V \quad \text{und} \quad N(f) \cap W = \left\{ (x, F(x)) \mid x \in U \right\},$$

dann darf es keine Tangentialvektoren der Form $(0, d)$ mit $d \neq 0$ geben: Man betrachte

$$\begin{aligned} K: U &\rightarrow U \times V \\ x &\mapsto (x, F(x)). \end{aligned}$$

Da $f \circ K$ die Nullabbildung ist, gilt

$$0 = J_{f \circ K}(x_0) = J_f(x_0, F(x_0)) \cdot J_K(x_0) = J_f(x_0, y_0) \cdot \begin{pmatrix} \mathbb{E}_k \\ -J_F(x_0) \end{pmatrix}.$$

Jeder Vektor der Form

$$(c, c \cdot J_F(x_0)^t), \quad c \in \mathbb{R}^k,$$

ist daher ein Tangentialvektor. Diese bilden offenbar einen k -dimensionalen Teilraum von \mathbb{R}^m . Da $J_f(x_0, y_0)$ nach Voraussetzung Rang $m - k$ hat, sind in der Tat alle Tangentialvektoren von der obigen Form.

Wir sehen:

$$\forall d \in \mathbb{R}^n \setminus \{0\}: \quad \frac{\partial f}{\partial y}(p) \cdot d^t \neq 0.$$

Folglich ist $(\partial f / \partial y)(p)$ eine invertierbare $(n \times n)$ -Matrix.

Der Satz über implizite Funktionen kehrt diese Beobachtung um, sofern f **stetig** differenzierbar ist.

11.2.1 Satz (Der Satz über implizite Funktionen). *Es seien $G \subset \mathbb{R}^m = \mathbb{R}^k \times \mathbb{R}^n$ ein Gebiet, $f: G \rightarrow \mathbb{R}^n$ eine **stetig differenzierbare** Abbildung,*

$$N(f) := \left\{ p=(x, y) \in G \mid f(x, y) = 0 \right\}$$

und $p_0=(x_0, y_0) \in N(f)$, so dass die Matrix

$$\frac{\partial f}{\partial y}(x_0, y_0)$$

invertierbar ist. Dann gibt es offene Teilmengen $U \subset \mathbb{R}^k$ und $V \subset \mathbb{R}^n$ mit

- $x_0 \in U$,
- $y_0 \in V$,
- $W := U \times V \subset G$

sowie eine stetig differenzierbare Abbildung

$$F: U \rightarrow V,$$

so dass

- i) $F(x_0) = y_0$,
- ii) $N(f) \cap W = \{ (x, y) \in W \mid y = F(x) \}$,
- iii) für alle $x \in U$ die Matrix $(\partial f / \partial y)(x, y)$ invertierbar ist und

$$J_F(x) = - \left(\frac{\partial f}{\partial y}(x, F(x)) \right)^{-1} \cdot \frac{\partial f}{\partial x}(x, F(x))$$

gilt.

Der Satz über implizite Funktionen wird mit Hilfe des Umkehrsatzes 10.2.1 und 10.2.3 bewiesen. Aus diesem Grund wird die Voraussetzung, dass f stetig differenzierbar ist, benötigt.

Beweis. Wir definieren die Abbildungen

$$\begin{aligned} H: G &\rightarrow \mathbb{R}^m \\ (x, y) &\mapsto (x, f(x, y)), \end{aligned}$$

$$\begin{aligned} \pi_1: \mathbb{R}^m = \mathbb{R}^k \times \mathbb{R}^n &\rightarrow \mathbb{R}^k \\ (x, y) &\mapsto x \end{aligned}$$

und

$$\begin{aligned} \pi_2: \mathbb{R}^m = \mathbb{R}^k \times \mathbb{R}^n &\rightarrow \mathbb{R}^n \\ (x, y) &\mapsto y. \end{aligned}$$

Diese Abbildungen fügen sich in die **kommutativen Diagramme**

$$\begin{array}{ccc} G & \xrightarrow{H} & \mathbb{R}^m, \\ \pi_{1|G} \searrow & & \swarrow \pi_1 \\ & & \mathbb{R}^k \end{array}, \quad \begin{array}{ccc} G & \xrightarrow{H} & \mathbb{R}^m. \\ f \searrow & & \swarrow \pi_2 \\ & & \mathbb{R}^n \end{array} \quad (11.5)$$

Dabei bedeutet „kommutativ“, dass $\pi_{1|G} = \pi_1 \circ H$ bzw. $f = \pi_2 \circ H$ gilt.

Behauptung. Es gibt offene Umgebungen $W \subset G$ von (x_0, y_0) und $Z \subset \mathbb{R}^m$ von $H(x_0, y_0) = (x_0, f(x_0, y_0))$, so dass $H(W) = Z$ und

$$H|_W: W \rightarrow Z$$

ein Diffeomorphismus ist.

Im Beweis dieser Behauptung bringen wir den Umkehrsatz ins Spiel. Nach Definition von H gilt

$$J_H(x_0, y_0) = \left(\begin{array}{c|c} \mathbb{E}_k & 0 \\ \hline \frac{\partial f}{\partial x}(x_0, y_0) & \frac{\partial f}{\partial y}(x_0, y_0) \end{array} \right). \quad (11.6)$$

Es gilt somit

$$\text{Det}(J_H(x_0, y_0)) = \text{Det}(\mathbb{E}_k) \cdot \text{Det}\left(\frac{\partial f}{\partial y}(x_0, y_0)\right) = \text{Det}\left(\frac{\partial f}{\partial y}(x_0, y_0)\right) \neq 0.$$

Deshalb folgt die Behauptung direkt aus Satz 10.2.1 und 10.2.3. √

Wir können nach eventueller Verkleinerung von W annehmen, dass es offene Teilmengen $U' \subset \mathbb{R}^k$ und $V \subset \mathbb{R}^n$ mit $x_0 \in U'$, $y_0 \in V$ und $W = U' \times V$ gibt. Es gilt $(x_0, y_0) \in U' \times V$ und damit $(x_0, 0) = H(x_0, y_0) \in Z$. Da Z offen ist, können wir eine offene Teilmenge $U \subset \mathbb{R}^k$ finden, so dass $x_0 \in U$, $U \subset U'$ und $U \times \{0\} \subset Z$. Sei

$$F: U \xrightarrow{x \mapsto (x,0)} Z \xrightarrow{H^{-1}} W = U' \times V \xrightarrow{\pi_{2|W}} V.$$

Die Abbildung F ist offenbar stetig differenzierbar.

Behauptung.

$$N(f) \cap (U \times V) = \{ (x, F(x)) \mid x \in U \}.$$

Das erste Diagramm in (11.5) und die Definition von F zeigen $(x, F(x)) = H^{-1}(x, 0)$, $x \in U$. Die Inklusion „ \supset “ ergibt sich somit aus dem zweiten Diagramm in (11.5).

Wenn umgekehrt $(x, y) \in N(f) \cap (U \times V)$ gilt, dann folgt $H(x, y) = (x, 0) \in U \times \{0\} \subset Z$. Mit der Definition von F leiten wir

$$(x, y) = H^{-1}(x, 0) = (x, F(x))$$

ab, und die Behauptung ist gezeigt. √

Für $(x, y) \in U \times V$ ist $J_H(x, y)$ invertierbar. Daher ist $(\partial f / \partial y)(x, y)$ für $(x, y) \in U \times V$ invertierbar (vgl. (11.6)). Es bleibt, die Formel für die Jacobi-Matrix nachzuprüfen. Dazu führen wir die Abbildung

$$\begin{aligned} K: U &\rightarrow \mathbb{R}^m \\ x &\mapsto (x, F(x)) \end{aligned}$$

ein. Es gilt:

$$\forall x \in U: (f \circ K)(x) = 0.$$

Damit und mit der Kettenregel 6.3.1 schließen wir

$$\begin{aligned} 0 &= J_{f \circ K}(x) = J_f(x, F(x)) \cdot J_K(x) \\ &= \left(\frac{\partial f}{\partial x}(x, F(x)) \mid \frac{\partial f}{\partial y}(x, F(x)) \right) \cdot \left(\frac{\mathbb{E}_k}{J_F(x)} \right) \\ &= \frac{\partial f}{\partial x}(x, F(x)) + \frac{\partial f}{\partial y}(x, F(x)) \cdot J_F(x). \end{aligned}$$

Diese Gleichung formt man in die behauptete Formel für $J_f(x)$ um. □

11.2.2 *Beispiel.* Wir kommen auf die Lemniskate von Gerono aus Beispiel 4.4.2, ii), zurück. Sie ist als $N(f)$ für die Funktion

$$\begin{aligned} f: \mathbb{R}^2 &\longrightarrow \mathbb{R} \\ (x, y) &\longmapsto x^2 \cdot (1 - x^2) - y^2 = -x^4 + x^2 - y^2 \end{aligned}$$

gegeben. Die Berechnung der partiellen Ableitungen führt zu

$$\begin{aligned} \frac{\partial f}{\partial x}(x, y) &= 2x - 4x^3 = 2x \cdot (1 - 2x^2) \\ \frac{\partial f}{\partial y}(x, y) &= -2y. \end{aligned}$$

Es folgt $\text{Grad}f(0,0) = (0,0)$. Die Menge $N(f)$ lässt sich um den Ursprung herum nicht als Graph einer differenzierbaren Funktion darstellen. Das kann man daran sehen, dass $N(f)$ sowohl zur x - als auch zur y -Achse symmetrisch ist. Für $(x, y) \in N(f)$ mit $(x, y) \neq (0,0)$, $(\pm 1, 0)$ ist die partielle Ableitung nach y nicht null, so dass die Gleichung $f(x, y) = 0$ nach y aufgelöst werden kann. Die Auflösung lautet

$$y = \pm \sqrt{x^2 - x^4} = \pm |x| \cdot \sqrt{1 - x^2}, \quad x \in (-1, 0) \vee x \in (0, 1).$$

In den Punkten $(\pm 1, 0)$ ist die partielle Ableitung nach x ungleich null, und die Gleichung $f(x, y) = 0$ kann nach x aufgelöst werden. Auch hier können wir die Auflösung schnell explizit durchführen. Die Gleichung $f(x, y) = -x^4 + x^2 - y^2 = 0$ können wir als quadratische Gleichung für x^2 auffassen. Mit der Lösungsformel (12), Seite 94, finden wir

$$x^2 = \frac{1 \pm \sqrt{1 - 4y^2}}{2}.$$

Wir schließen

$$x = \pm \frac{1}{\sqrt{2}} \cdot \sqrt{1 \pm \sqrt{1 - 4y^2}}.$$

Dieser Ausdruck ist für $y \in (-1/2, 1/2)$ definiert. Das negative Vorzeichen vor der „großen“ Wurzel ergibt offenbar die Parametrisierung in der Nähe von $(-1, 0)$ und das positive diejenige in der Nähe von $(1, 0)$. Wir müssen allerdings noch das Vorzeichen unter der Wurzel bestimmen. Dazu beachten wir, dass das Einsetzen von $y = 0$ zu $x = -1$ bzw. $x = 1$ führen muss, so dass in beiden Fällen das positive Vorzeichen zu wählen ist.

In einer geeigneten Umgebung von $(1, 0)$ gilt $(x, y) \in N(f)$ genau dann, wenn

$$y \in \left(-\frac{1}{2}, \frac{1}{2}\right) \quad \text{und} \quad x = \frac{1}{\sqrt{2}} \cdot \sqrt{1 + \sqrt{1 - 4y^2}},$$

und in einer geeigneten Umgebung von $(-1, 0)$ gilt $(x, y) \in N(f)$ genau dann, wenn

$$y \in \left(-\frac{1}{2}, \frac{1}{2}\right) \quad \text{und} \quad x = -\frac{1}{\sqrt{2}} \cdot \sqrt{1 + \sqrt{1 - 4y^2}}.$$

Die Leserin bzw. der Leser möge „geeignete“ Umgebungen angeben und sich an der Skizze auf Seite 54 veranschaulichen, welches Stück der Lemniskate jeweils parametrisiert wird.

11.3 Untermannigfaltigkeiten von \mathbb{R}^n

Der Satz über implizite Funktion besagt, dass wir die Lösungsmenge eines Systems von stetig differenzierbaren Gleichungen in \mathbb{R}^m unter gewissen Voraussetzungen lokal mit passenden offenen Mengen in \mathbb{R}^k identifizieren können. Dies legt die Vermutung nahe, dass wir auf diesen Lösungsmengen wieder Analysis betreiben können. Diese Vermutung präzisieren wir über das Konzept der Untermannigfaltigkeit. Dieses Konzept kann ohne große Mühe auf das Konzept der abstrakten Mannigfaltigkeit verallgemeinert werden. Der Begriff der Mannigfaltigkeit geht auf Riemann zurück und ist ein Meilenstein in der Mathematik. So ist z.B. die allgemeine Relativitätstheorie in der Sprache der Mannigfaltigkeiten geschrieben. Es ist daher ratsam, diesen Begriff sorgsam zu begründen.

11.3.1 Definition. Es sei $P \subset \mathbb{R}^k$ ein Gebiet. Ein *glattes parametrisiertes Flächenstück über P* ist eine stetig differenzierbare Abbildung $\varphi: P \rightarrow \mathbb{R}^n$, für die gilt:

1. Die Abbildung φ ist injektiv.
2. Für jeden Punkt $p \in P$ gilt $\text{Rg}(J_\varphi(p)) = k$.
3. Für jeden Punkt $p \in P$ und jede Folge $(p_l)_{l \in \mathbb{N}}$ in P gilt:⁴

$$\lim_{l \rightarrow \infty} \varphi(p_l) = \varphi(p) \quad \Rightarrow \quad \lim_{l \rightarrow \infty} p_l = p.$$

Die Zahl k heißt *Dimension* des glatten parametrisierten Flächenstücks.

11.3.2 Beispiele. i) Es seien $P \subset \mathbb{R}^k$ ein Gebiet und $F: P \rightarrow \mathbb{R}^l$ eine stetig differenzierbare Abbildung. Sei

$$\begin{aligned} \varphi: P &\rightarrow \mathbb{R}^n = \mathbb{R}^k \times \mathbb{R}^l \\ p &\mapsto (p, F(p)). \end{aligned}$$

Wir behaupten, dass φ ein k -dimensionales glattes parametrisiertes Flächenstück ist.

1. Die Abbildung φ ist offensichtlich injektiv.
2. Die Jacobi-Matrix

$$J_\varphi(p) = \begin{pmatrix} \mathbb{E}_k \\ J_F(p) \end{pmatrix}, \quad p \in P,$$

hat Rang k .

3. Für $p \in P$ und eine Folge $(p_s)_{s \in \mathbb{N}}$ mit

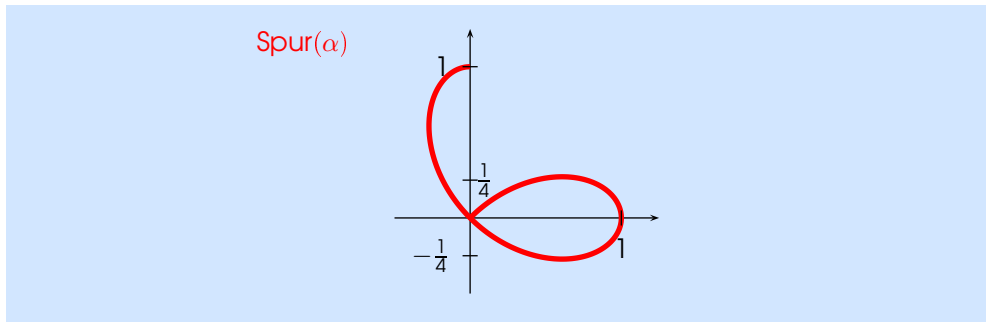
$$\lim_{s \rightarrow \infty} (p_s, F(p_s)) = \lim_{s \rightarrow \infty} \varphi(p_s) = \varphi(p) = (p, F(p))$$

gilt wegen Satz 2.1.3 $\lim_{s \rightarrow \infty} p_s = p$.

- ii) Wir betrachten die stetig differenzierbare Abbildung

$$\begin{aligned} \alpha: I := \left(-\frac{\pi}{2}, \frac{\pi}{4}\right) &\rightarrow \mathbb{R}^2 \\ t &\mapsto (\cos(2t) \cdot \cos(t), \cos(2t) \cdot \sin(t)). \end{aligned}$$

⁴Die angegebene Bedingung impliziert, dass φ injektiv ist. (Warum?) Wir werden aber gleich sehen, dass es Abbildungen $\varphi: P \rightarrow \mathbb{R}^n$ gibt, die 1. und 2. erfüllen, nicht aber 3.



1. Die Abbildung α ist injektiv. (Offenbar kann man $\cos(2t)$ aus $\alpha(t)$ zurückgewinnen. Die Abbildung $t \mapsto \cos(2t)$ ist auf den Intervallen $(-\pi/2, 0]$ und $(0, \pi/4)$ injektiv. Man erkennt $t \in (0, \pi/4)$ daran, dass sowohl die x - als auch die y -Koordinate von $\alpha(t)$ positiv sind.)

2. Man berechnet

$$\alpha'(t) = \begin{pmatrix} -2 \cdot \sin(2t) \cdot \cos(t) - \cos(2t) \cdot \sin(t) \\ -2 \cdot \sin(2t) \cdot \sin(t) + \cos(2t) \cdot \cos(t) \end{pmatrix} \neq \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad t \in I.$$

(Es gilt

$$\begin{aligned} & -\sin(t) \cdot (-2 \cdot \sin(2t) \cdot \cos(t) - \cos(2t) \cdot \sin(t)) + \\ & + \cos(t) \cdot (-2 \cdot \sin(2t) \cdot \sin(t) + \cos(2t) \cdot \cos(t)) = \cos(2t). \end{aligned}$$

Daraus folgt $\alpha'(t) \neq 0$ für $t \neq -\pi/4$, $t \in I$. Zudem findet man $\alpha'(-\pi/4) \neq 0$.)

3. Es sei $(t_k)_{k \in \mathbb{N}}$ eine Folge mit $t_k \in I$, $k \in \mathbb{N}$, und $\lim_{k \rightarrow \infty} t_k = \pi/4$. Dann gilt

$$\lim_{k \rightarrow \infty} \alpha(t_k) = (0, 0) = \alpha\left(-\frac{\pi}{4}\right)$$

aber

$$\lim_{k \rightarrow \infty} t_k \neq -\frac{\pi}{4}.$$

Somit ist α **kein** eindimensionales glattes parametrisiertes Flächenstück.⁵

11.3.3 Definition. Eine Teilmenge $M \subset \mathbb{R}^n$ heißt *k-dimensionale Untermannigfaltigkeit*, wenn es zu jedem Punkt $x \in M$ eine offene Umgebung $U \subset \mathbb{R}^n$ von x , ein Gebiet $P \subset \mathbb{R}^k$, einen Punkt $p \in P$ und ein glattes parametrisiertes Flächenstück $\varphi: P \rightarrow \mathbb{R}^n$ über P mit den Eigenschaften

- $\varphi(p) = x$,
- $\varphi(P) = M \cap U$

gibt.

11.3.4 Satz. Es seien $B \subset \mathbb{R}^n$ eine offene Teilmenge, $f_i: B \rightarrow \mathbb{R}$, $i = 1, \dots, l$, stetig differenzierbare Funktionen und

$$M = \left\{ x \in B \mid f_1(x) = \dots = f_l(x) = 0 \right\}.$$

Für jeden Punkt $x \in M$ seien die Vektoren

$$\text{Grad}f_1(x), \dots, \text{Grad}f_l(x)$$

linear unabhängig.⁶ Dann ist M eine *k-dimensionale Untermannigfaltigkeit* von \mathbb{R}^n , $k := n - l$.

⁵Wir entschuldigen uns für diese seltsame Terminologie.

⁶Daraus folgt $l \leq n$.

Beweis. Wir setzen

$$\begin{aligned} f: B &\longrightarrow \mathbb{R}^n \\ x &\longmapsto (f_1(x), \dots, f_n(x)). \end{aligned}$$

Dann haben wir

$$J_f(x) = \begin{pmatrix} \text{Grad}f_1(x) \\ \vdots \\ \text{Grad}f_l(x) \end{pmatrix}, \quad x \in B.$$

Die Voraussetzung besagt

$$\text{Rg}(J_f(x)) = l, \quad x \in M.$$

Sei $x_0 \in M$. Nach Umsortierung der Koordinaten können wir annehmen, dass die letzten l Spalten von $J_f(x_0)$ linear unabhängig sind.⁷ Wir schreiben $n = k + l$ und $\mathbb{R}^n \ni q = (x, y) = (x_1, \dots, x_k, y_1, \dots, y_l)$. Nach dem Satz über implizite Funktionen 11.2.1 gibt es offene Mengen $U \subset \mathbb{R}^k$ und $V \subset \mathbb{R}^l$ mit $x_0 \in U$, $y_0 \in V$, $q_0 = (x_0, y_0)$, und eine stetig differenzierbare Abbildung $F: U \rightarrow V$, so dass

- $F(x_0) = y_0$,
- $M \cap (U \times V) = \{ (x, F(x)) \mid x \in U \}$.

Wir können ohne Beschränkung der Allgemeinheit annehmen, dass U ein Gebiet ist, z.B. $U = B(x_0, \varepsilon)$ für ein $\varepsilon > 0$. Nach Beispiel 11.3.2, i), ist

$$\begin{aligned} \varphi: U &\longrightarrow \mathbb{R}^n \\ x &\longmapsto (x, F(x)) \end{aligned}$$

ein k -dimensionales glattes parametrisiertes Flächenstück über U . Dann ist $W := U \times V$ eine offene Umgebung von p_0 , und wir haben wie gewünscht $\varphi(x_0) = (x_0, y_0) = q_0$ und $M \cap W = \{ (x, F(x)) = \varphi(x) \mid x \in U \} = \varphi(U)$. \square

11.3.5 Beispiel. Sei

$$\begin{aligned} f: \mathbb{R}^n &\longrightarrow \mathbb{R} \\ (x_1, \dots, x_n) &\longmapsto x_1^2 + \dots + x_n^2 - 1. \end{aligned}$$

Die Menge

$$\begin{aligned} S^{n-1} &:= \{ (x_1, \dots, x_n) \in \mathbb{R}^n \mid f(x_1, \dots, x_n) = 0 \} \\ &= \{ (x_1, \dots, x_n) \in \mathbb{R}^n \mid x_1^2 + \dots + x_n^2 = 1 \} \end{aligned}$$

wird die $(n - 1)$ -dimensionale Einheitskugel genannt. Man beachte:

$$\forall (x_1, \dots, x_n) \in S^{n-1} : \text{Grad}f(x_1, \dots, x_n) = (2x_1, \dots, 2x_n) \neq (0, \dots, 0).$$

Damit ist S^{n-1} eine $(n - 1)$ -dimensionale Untermannigfaltigkeit von \mathbb{R}^n .

11.3.6 Lemma. Es seien $P \subset \mathbb{R}^k$ ein Gebiet, $\varphi: P \rightarrow \mathbb{R}^n$ ein glattes parametrisiertes Flächenstück über P , $S := \varphi(P) \subset \mathbb{R}^n$ und $U \subset P$ eine offene Teilmenge. Dann gibt es eine offene Teilmenge $V \subset \mathbb{R}^n$, so dass

$$\varphi(U) = S \cap V.$$

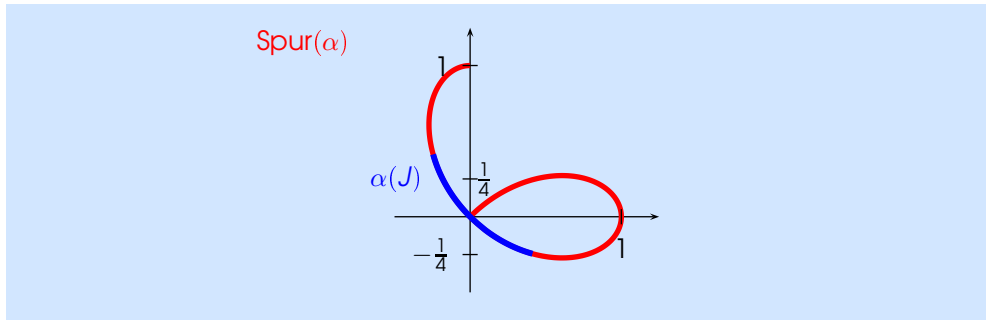
⁷Formal können wir f durch $f \circ A$ ersetzen. Dabei ist $A: B' \rightarrow B$ die Einschränkung einer linearen Abbildung, die die Koordinaten entsprechend vertauscht. Man vergleiche Beispiel 10.1.12, ii).

Die Topologien auf \mathbb{R}^k bzw. \mathbb{R}^n induzieren Topologien auf P bzw. S . Das Lemma besagt, dass $\varphi: P \rightarrow S$ bzgl. dieser Topologien ein Homöomorphismus ist (Lemma 10.1.4).

11.3.7 *Beispiel.* Wir schauen uns nocheinmal Beispiel 11.2.2 an. Sei z.B.

$$J := \left(-\frac{\pi+1}{4}, -\frac{\pi-1}{4} \right) \subset I.$$

Dann ist $\alpha(J)$ nicht offen in $S = \alpha(I)$.



Beweis von Lemma 11.3.6. Wir beginnen mit der folgenden Aussage:

Behauptung. Seien $p \in U$ und $x := \varphi(p)$. Dann existiert ein $\varepsilon > 0$, so dass

$$B(x, \varepsilon) \cap S \subset \varphi(U).$$

Wenn die Behauptung falsch wäre, dann könnte man zu jedem $s \in \mathbb{N}$ einen Punkt

$$x_s \in B\left(x, \frac{1}{s+1}\right) \cap S \text{ finden, so dass } x_s \notin \varphi(U).$$

Die resultierende Folge $(x_s)_{s \in \mathbb{N}}$ konvergierte gegen x . Wir fänden Elemente $p_s \in P$ mit

$$\varphi(p_s) = x_s, \quad s \in \mathbb{N}.$$

Offenbar gälte $p_s \notin U$, $s \in \mathbb{N}$. Definition 11.3.1, 3., beinhaltetete in diesem Fall

$$\lim_{s \rightarrow \infty} p_s = p.$$

Da U eine offene Umgebung von p ist, gäbe es schließlich einen Index s_0 mit $p_s \in U$ für $s \geq s_0$. Dieser Widerspruch zeigt die Gültigkeit der Behauptung. \checkmark

Zu jedem Punkt $p \in U$ wählen wir ein $\varepsilon(p) > 0$, so dass $B(\varphi(p), \varepsilon(p)) \cap S \subset \varphi(U)$. Die Menge

$$V := \bigcup_{p \in U} B(\varphi(p), \varepsilon(p))$$

ist offen. Nach Konstruktion gilt $V \cap S \subset \varphi(U)$, und jeder Punkt $x = \varphi(p)$, $p \in U$, ist in $V \cap S$ enthalten, so dass in der Tat $V \cap S = \varphi(U)$. \square

11.3.8 Satz. Es seien $P \subset \mathbb{R}^k$ ein Gebiet und $\varphi: P \rightarrow \mathbb{R}^n$ ein glattes parametrisiertes Flächenstück über P . Dann gibt es für jeden Punkt $x \in S := \varphi(P)$ eine offene Umgebung $U \subset \mathbb{R}^n$ und eine stetig differenzierbare Abbildung $f: U \rightarrow \mathbb{R}^l$, $l := n - k$, so dass

- $U \cap S = \{x \in U \mid f(x) = 0\}$,

- $\text{Rg}(J_f(x)) = l, x \in U \cap S$.

Beweis. Wir schreiben $\mathbb{R}^n = \mathbb{R}^k \times \mathbb{R}^l$ und definieren

$$\begin{aligned} \pi_1: \mathbb{R}^n &\longrightarrow \mathbb{R}^k \\ (y, z) &\longmapsto y, \\ \pi_2: \mathbb{R}^n &\longrightarrow \mathbb{R}^l \\ (y, z) &\longmapsto z. \end{aligned}$$

Es sei $x = \varphi(p), p \in P$. Wir nehmen ohne Beschränkung der Allgemeinheit an, dass die ersten k Zeilen der Matrix $J_\varphi(p)$ linear unabhängig sind.⁸ Dann ist $J_{\pi_1 \circ \varphi}(p)$ eine invertierbare $(k \times k)$ -Matrix. Nach dem Umkehrsatz 10.2.1 und 10.2.3 gibt es offene Mengen $V, W \subset \mathbb{R}^k$, so dass $p \in V, (\pi_1 \circ \varphi)(p) \in W$ und

$$\pi_1 \circ \varphi: V \longrightarrow W$$

ein Diffeomorphismus ist. Wir setzen

$$\psi := (\pi_1 \circ \varphi)^{-1}: W \longrightarrow V$$

und

$$\begin{aligned} g: W &\longrightarrow \mathbb{R}^l \\ y &\longmapsto (\pi_2 \circ \varphi \circ \psi)(y). \end{aligned}$$

Gemäß Lemma 11.3.6 existiert eine offene Teilmenge $B \subset \mathbb{R}^n$ mit $B \cap S = \varphi(V)$. Für einen Punkt $b = (y, z) \in B$ beobachten wir:

$$\begin{aligned} b \in S &\iff b \in \varphi(V) \\ &\iff \exists q \in V: b = \varphi(q) \\ &\iff \exists q \in V: y = (\pi_1 \circ \varphi)(q), z = (\pi_2 \circ \varphi)(q) && (11.7) \\ &\iff \exists q \in V: y = \psi^{-1}(q), z = g(\psi^{-1}(q)) \\ &\iff y \in W \wedge z = g(y). \end{aligned}$$

Seien

$$U := (W \times \mathbb{R}^l) \cap B$$

und

$$\begin{aligned} f: U &\longrightarrow \mathbb{R}^l \\ (y, z) &\longmapsto z - g(y). \end{aligned}$$

Es gilt

$$J_f(y, z) = (-J_g(y) | E_l), \quad y \in U.$$

Wir müssen also noch

$$U \cap S = \{x \in U \mid f(x) = 0\}$$

nachweisen. Für „ \subset “ argumentieren wir folgendermaßen:

$$\begin{aligned} x = (y, z) \in U \cap S &\Rightarrow x \in B \cap S \\ &\stackrel{(11.7)}{\Rightarrow} y \in W, z = g(y) \\ &\Rightarrow f(x, y) = 0. \end{aligned}$$

Für „ \supset “ gehen wir wie folgt vor: Wegen $x = (y, z) \in U$ gilt $y \in W$ und $(y, z) \in B$. Da $f(y, z) = 0$, haben wir $z = g(y)$ und mit (11.7) $(y, z) \in S$. \square

⁸Formal kann man φ durch $A \circ \varphi$ ersetzen. Dabei ist $A: \mathbb{R}^n \longrightarrow \mathbb{R}^n$ eine lineare Abbildung, die die entsprechenden Koordinaten vertauscht. Man vergleiche Beispiel 10.1.12, ii).

11.3.9 *Bemerkung.* i) Der obige Satz ist eine Umkehrung von Satz 11.3.4: Jede k -dimensionale Untermannigfaltigkeit von \mathbb{R}^n kann lokal als Lösungsmenge eines Systems von l differenzierbaren Gleichungen, $l := n - k$, geschrieben werden, das die „Rangbedingung“ erfüllt.

ii) Aus dem Beweis des Satzes über implizite Funktionen ergibt sich weiter folgende Charakterisierung von Untermannigfaltigkeiten von \mathbb{R}^n :

Satz. *Es sei $M \subset \mathbb{R}^n$ eine Teilmenge von \mathbb{R}^n . Dann ist M genau dann eine Untermannigfaltigkeit der Dimension k , wenn es zu jedem Punkt $x \in M$ offene Teilmengen $U, V \subset \mathbb{R}^n$ mit $x \in U$ und einen Diffeomorphismus*

$$\psi: U \rightarrow V$$

mit

$$\psi(U \cap M) = \{ (x_1, \dots, x_n) \in V \mid x_{k+1} = \dots = x_n = 0 \}$$

gibt.

11.3.10 *Beispiel.* Es seien $S \subset \mathbb{R}^n$ eine k -dimensionale Untermannigfaltigkeit und $I \subset \mathbb{R}$ ein offenes Intervall. Dann ist $S \times I \subset \mathbb{R}^{n+1}$ eine $(k + 1)$ -dimensionale Untermannigfaltigkeit. (Für $(x, t) \in S \times I$ gibt es eine offene Umgebung $U \subset \mathbb{R}^n$ von x und eine stetig differenzierbare Abbildung $f: U \rightarrow \mathbb{R}^l$, $l := n - k$, mit $\text{Rg}(J_f(x)) = l$, $x \in U \cap S$, und $U \cap S = \{ x \in U \mid f(x) = 0 \}$. Wir betrachten weiter

$$\begin{aligned} g: U \times I &\rightarrow \mathbb{R}^l \\ (x, t) &\mapsto f(x). \end{aligned}$$

Dann gilt $\text{Rg}(J_g(x, t)) = \text{Rg}((J_f(x)|0)) = l$, $(x, t) \in (U \times I) \cap (S \times I)$, und $(U \times I) \cap (S \times I) = \{ (x, t) \in U \times I \mid g(x, t) = 0 \}$. Nach Satz 11.3.4 ist $S \times I$ eine $(k + 1)$ -dimensionale Untermannigfaltigkeit von \mathbb{R}^{n+1} .)

11.4 Tangentialvektoren

Wir haben bereits öfters von Tangentialvektoren, Tangentialebenen u.Ä. gesprochen. An dieser Stelle wollen wir diese Begriffe präzisieren und auf verschiedene Weisen charakterisieren.

11.4.1 Definition. Es seien $M \subset \mathbb{R}^n$ eine Untermannigfaltigkeit und $x \in M$ ein Punkt.

a) Ein *Tangentialvektor an M in x* ist ein Vektor $v \in \mathbb{R}^n$, zu dem $\varepsilon > 0$ und ein stetig differenzierbarer Weg

$$\gamma: (-\varepsilon, \varepsilon) \rightarrow \mathbb{R}^n$$

existieren, so dass gilt:

- $\text{Spur}(\gamma) \subset M$,
- $\gamma(0) = x$,
- $\gamma'(0) = v$.

b) Die Menge

$$T_x(M) := \{ v \in \mathbb{R}^n \mid v \text{ ist Tangentialvektor an } M \text{ in } x \}$$

ist der *Tangentialraum an M in x* .

Das obige Konzept von Tangentialvektor ist bereits im Beweis von Satz 7.5.3 aufgetaucht.

11.4.2 *Beispiel.* i) Da wir für beliebiges $\varepsilon > 0$ für γ den konstanten Weg $t \mapsto x$ nehmen können, folgt $0 \in T_x(M)$.

ii) Die Menge $M = \mathbb{R}^n$ ist eine Untermannigfaltigkeit von \mathbb{R}^n . Für jeden Punkt $x \in \mathbb{R}^n$ gilt

$$T_x(\mathbb{R}^n) = \mathbb{R}^n.$$

11.4.3 Satz. *Es seien $M \subset \mathbb{R}^n$ eine k -dimensionale Untermannigfaltigkeit und $x \in M$. Weiter seien $l := n - k$, $U \subset \mathbb{R}^n$ eine offene Umgebung von x , $f: U \rightarrow \mathbb{R}^l$ eine stetig differenzierbare Abbildung, $P \subset \mathbb{R}^k$ ein Gebiet, $p \in P$ und $\varphi: P \rightarrow \mathbb{R}^n$ ein glattes parametrisiertes Flächenstück, so dass gilt:*

- $M \cap U = \{x \in U \mid f(x) = 0\}$,
- $\text{Rg}(J_f(x)) = l$, $x \in U$,
- $\varphi(p) = x$,
- $\varphi(P) = M \cap U$.

Dann sind für jeden Vektor $v \in \mathbb{R}^n$ folgende Aussagen äquivalent:

- i) Der Vektor v ist Tangentialvektor an M in x .
- ii) Es gilt $v \in \text{Ker}(Df(x))$.
- iii) Es gilt $v \in \text{Bild}(D\varphi(x))$.

Die Implikation „ii) \Rightarrow i)“ haben wir in einem Spezialfall im Beweis von Satz 7.5.3 mühsam hergeleitet.

Beweis. „i) \Rightarrow ii)“. Seien $\varepsilon > 0$ und $\gamma: (-\varepsilon, \varepsilon) \rightarrow \mathbb{R}^n$ ein stetig differenzierbarer Weg mit $\gamma(0) = x$, $\gamma'(0) = v$ und $\text{Spur}(\gamma) \subset M$. Die letzte Bedingung impliziert

$$\forall t \in (-\varepsilon, \varepsilon): (f \circ \gamma)(t) = 0.$$

Wir folgern

$$\forall t \in (-\varepsilon, \varepsilon): Df(\gamma(t))(\gamma'(t)) = (f \circ \gamma)'(t) = 0.$$

Für $t = 0$ bedeutet das $Df(x)(v) = 0$, i.e. $v \in \text{Ker}(Df(x))$.

„ii) \Rightarrow iii)“. Hier gilt

$$\forall q \in P: (f \circ \varphi)(q) = 0$$

und somit

$$\forall q \in P: 0 = D(f \circ \varphi)(q) = Df(\varphi(q)) \circ D\varphi(q).$$

Insbesondere gilt

$$Df(x) \circ D\varphi(p) = 0$$

und damit

$$\text{Bild}(D\varphi(p)) \subset \text{Ker}(Df(x)). \tag{11.8}$$

Da φ ein k -dimensionales glattes parametrisiertes Flächenstück über P ist, haben wir

$$\dim(\text{Bild}(D\varphi(p))) = k.$$

Auf der anderen Seite gilt

$$\dim(\text{Ker}(Df(x))) = n - \text{Rg}(J_f(x)) = n - l = k.$$

Deshalb liegt in (11.8) aus Dimensionsgründen Gleichheit vor.
 „iii)⇒i)“. Seien $w \in \mathbb{R}^k$, $v = D\varphi(w)$ und

$$\begin{aligned} \alpha: \mathbb{R} &\longrightarrow \mathbb{R}^k \\ t &\longmapsto p + t \cdot w. \end{aligned}$$

Da α stetig ist, die Menge $P \subset \mathbb{R}^k$ offen ist und $\alpha(0) = p \in P$, existiert ein $\varepsilon > 0$ mit $\alpha((-\varepsilon, \varepsilon)) \subset P$. Damit ist

$$\begin{aligned} \gamma: (-\varepsilon, \varepsilon) &\longrightarrow \mathbb{R}^n \\ t &\longmapsto \varphi(\alpha(t)) \end{aligned}$$

ein stetig differenzierbarer Weg mit $\text{Spur}(\gamma) \subset \varphi(P) \subset M$, $\gamma(0) = \varphi(p) = x$ und

$$\gamma'(0) = D\varphi(\alpha(0))(\alpha'(0)) = D\varphi(p)(w) = v.$$

Diese Gleichung zeigt $v \in T_x(M)$. □

11.4.4 Bemerkung. Die zweite und dritte Charakterisierung zeigen jeweils, dass $T_x(M)$ ein linearer Teilraum von \mathbb{R}^n ist.

11.4.5 Beispiele. i) Wir betrachten die Einheitskugel

$$S^{n-1} = \{x \in \mathbb{R}^n \mid \|x\| = 1\} = \{x \in \mathbb{R}^n \mid \langle x, x \rangle = 1\}.$$

Mit der Abbildung

$$\begin{aligned} f: \mathbb{R}^n &\longrightarrow \mathbb{R} \\ x &\longmapsto \langle x, x \rangle \end{aligned}$$

gilt

$$S^{n-1} = \{x \in \mathbb{R}^n \mid f(x) = 0\}.$$

Für $x \in S^{n-1}$ gilt:

$$\begin{aligned} Df(x): \mathbb{R}^n &\longrightarrow \mathbb{R} \\ v &\longmapsto \langle \text{Grad}f(x), v \rangle. \end{aligned}$$

Wegen $\text{Grad}f(x) = 2x$, $x \in \mathbb{R}^n$, erhalten wir mit Satz 11.4.3

$$\begin{aligned} T_x(S^{n-1}) &= \text{Ker}(Df(x)) \\ &= \{v \in \mathbb{R}^n \mid \langle \text{Grad}f(x), v \rangle = 0\} \\ &= \{v \in \mathbb{R}^n \mid \langle x, v \rangle = 0\}. \end{aligned}$$

ii) Die *orthogonale Gruppe* ist definiert als

$$O_n(\mathbb{R}) := \{A \in \text{Mat}(n, \mathbb{R}) \mid A \cdot A^t = E_n\}.$$

Wir wollen zeigen, dass $O_n(\mathbb{R})$ eine Untermannigfaltigkeit von \mathbb{R}^{n^2} ist und ihre Dimension sowie ihren Tangentialraum in E_n bestimmen. Dazu betrachten wir den reellen Vektorraum

$$\text{Sym}_n(\mathbb{R}) := \{A \in \text{Mat}(n, \mathbb{R}) \mid A = A^t\}$$

der symmetrischen $(n \times n)$ -Matrizen Er hat die Dimension

$$\frac{n \cdot (n+1)}{2}.$$

Ferner führen wir die stetig differenzierbare Abbildung

$$\begin{aligned} f: \text{Mat}(n, \mathbb{R}) &\longrightarrow \text{Sym}_n(\mathbb{R}) \\ A &\longmapsto A \cdot A^t - \mathbb{E}_n \end{aligned}$$

ein. Mit dieser Abbildung erhalten wir die Beschreibung

$$\text{O}_n(\mathbb{R}) = f^{-1}(0)$$

der orthogonalen Gruppe, die wir benutzen können, um die gesteckten Ziele zu erreichen.

Behauptung 1.

$$\forall A, B \in \text{Mat}(n, \mathbb{R}) : Df(A)(B) = B \cdot A^t + A \cdot B^t.$$

Für den Beweis dieser Behauptung beobachten wir zunächst, dass für $A, C \in \text{Mat}(n, \mathbb{R})$

$$f(C) - f(A) = C \cdot C^t - A \cdot A^t = (C - A) \cdot A^t + A \cdot (C - A)^t + (C - A) \cdot (C - A)^t$$

gilt. Weiter gilt

$$0 \leq \lim_{C \rightarrow A} \frac{\|(C - A) \cdot (C - A)^t\|_{\text{op}}}{\|C - A\|_{\text{op}}} \leq \lim_{C \rightarrow A} \frac{\|C - A\|_{\text{op}} \cdot \|(C - A)^t\|_{\text{op}}}{\|C - A\|_{\text{op}}} = 0.$$

Aus den beiden aufgestellten Gleichungen folgt Behauptung 1 (vgl. Definition 6.1.1 und Bemerkung 6.1.2). ✓

Behauptung 2. Für jede Matrix $A \in \text{O}_n(\mathbb{R})$ ist $Df(A)$ surjektiv.

Denn für eine Matrix $S \in \text{Sym}_n(\mathbb{R})$ können wir

$$B := \frac{1}{2} \cdot S \cdot A$$

setzen und bekommen

$$\begin{aligned} Df(A)(B) &= \frac{1}{2} \cdot S \cdot A \cdot A^t + \frac{1}{2} \cdot A \cdot A^t \cdot S^t \\ &\stackrel{A \cdot A^t = \mathbb{E}_n}{=} \frac{1}{2} \cdot S + \frac{1}{2} \cdot S^t \\ &\stackrel{S = S^t}{=} S. \end{aligned}$$

Damit liegt S im Bild von $Df(A)$. ✓

Aus den beiden Behauptungen folgt nach Satz 11.3.8, dass $\text{O}_n(\mathbb{R})$ eine Untermannigfaltigkeit von $\text{Mat}(n, \mathbb{R})$ der Dimension

$$n^2 - \frac{n \cdot (n+1)}{2} = \frac{n \cdot (n-1)}{2}$$

ist. Nach Satz 11.4.3 gilt für den Tangentialraum an $\text{O}_n(\mathbb{R})$ in \mathbb{E}_n :

$$\begin{aligned} T_{\mathbb{E}_n}(\text{O}_n(\mathbb{R})) &= \text{Ker}(Df(\mathbb{E}_n)) \\ &= \left\{ B \in \text{Mat}(n, \mathbb{R}) \mid B + B^t = 0 \right\} \\ &= \left\{ B \in \text{Mat}(n, \mathbb{R}) \mid B = -B^t \right\}. \end{aligned}$$

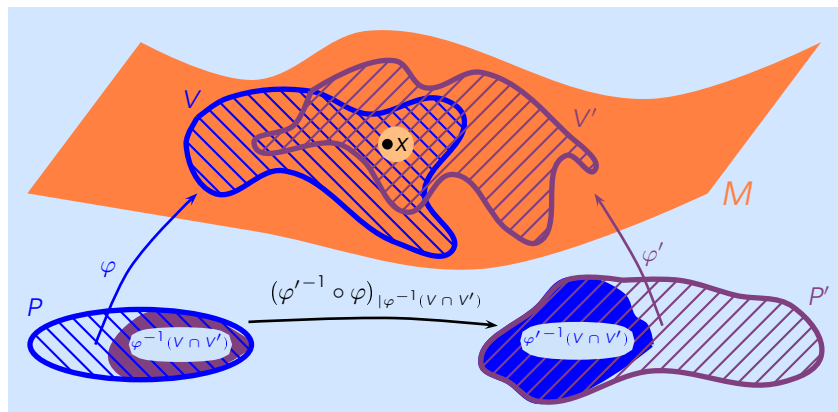
Es handelt sich um den Vektorraum der **schiefsymmetrischen** $(n \times n)$ -Matrizen.

11.5 Differenzierbare Abbildungen auf Untermannigfaltigkeiten

Wir entwickeln nun den ursprünglichen Gedanken, Analysis auf Untermannigfaltigkeiten des \mathbb{R}^n zu betreiben, indem wir erklären, was differenzierbare Abbildungen auf Untermannigfaltigkeiten sind.

Seien also $M \subset \mathbb{R}^n$ eine k -dimensionale Untermannigfaltigkeit und $x \in M$ ein Punkt. Wir betrachten eine Funktion $f: M \rightarrow \mathbb{R}$ und wollen erklären, wann diese Funktion in x differenzierbar ist. Die Idee ist, ein glattes parametrisiertes Flächenstück $\varphi: P \rightarrow \mathbb{R}^n$ zu nehmen, so dass ein Punkt $p \in P$ mit $x = \varphi(p)$ existiert, und f in x differenzierbar zu nennen, wenn $f \circ \varphi$ in p differenzierbar ist. Die Hauptarbeit besteht darin, zu verifizieren, dass dieser Begriff wohldefiniert ist.

Seien also $\varphi: P \rightarrow M$ und $\varphi': P' \rightarrow M$ zwei glatte parametrisierte Flächenstücke, $V := \varphi(P)$, $V' := \varphi'(P')$ und $x \in V \cap V'$.



11.5.1 Satz. In der obigen Situation ist die Abbildung

$$(\varphi'^{-1} \circ \varphi)|_{\varphi^{-1}(V \cap V')} : \varphi^{-1}(V \cap V') \rightarrow \varphi'^{-1}(V \cap V')$$

ein Diffeomorphismus.

Beweis. Es seien $y \in V \cap V'$ und $p \in P$ bzw. $p' \in P'$, so dass $\varphi(p) = y$ bzw. $\varphi'(p') = y$. Wir nehmen ohne Beschränkung der Allgemeinheit an, dass die ersten k -Zeilen von $J_{\varphi'}(p')$ linear unabhängig sind, und definieren

$$F: P' \times \mathbb{R}^{n-k} \rightarrow \mathbb{R}^n \\ (q, t) \mapsto \varphi'(q) + (0, t).$$

Es gilt $F(p', 0) = \varphi'(p') = y$ und

$$J_F(p', 0) = \left(J_{\varphi'}(p') \mid \begin{matrix} 0 \\ \mathbb{E}_{n-k} \end{matrix} \right).$$

Es folgt

$$\text{Det}(J_F(p', 0)) = \text{Det}(\text{erste } k\text{-Zeilen von } J_{\varphi'}(p')) \cdot \text{Det}(\mathbb{E}_{n-k}) \neq 0.$$

Nach dem Umkehrsatz 10.2.1 und 10.2.3 gibt es offene Umgebungen $U' \subset P'$ von p' , $U^* \subset \mathbb{R}^{n-k}$ von 0 und $W \subset \mathbb{R}^n$ von y , so dass

$$F: U' \times U^* \rightarrow W$$

ein Diffeomorphismus ist. Indem wir W falls nötig verkleinern, können wir $M \cap W \subset V \cap V'$ erreichen. Da $\varphi: P \rightarrow V$ nach Lemma 11.3.6 ein Homöomorphismus ist, existiert eine offene Umgebung $U \subset P$ von p mit

$$\varphi(U) = M \cap W.$$

Die Abbildung

$$F^{-1} \circ \varphi: U \rightarrow U' \times U^*$$

ist stetig differenzierbar, und wegen

$$\varphi(U) \subset V \cap V' \subset \varphi'(P') = F(P' \times \{0\})$$

ist das Bild von $F^{-1} \circ \varphi$ in $P' \times \{0\}$ enthalten. Sei $u \in U$. Dann finden wir $u' \in P'$ mit $\varphi'(u') = \varphi(u)$ und berechnen

$$(F^{-1} \circ \varphi)(u) = (F^{-1} \circ \varphi')(u') = (F^{-1} \circ F)(u', 0) = (u', 0) = ((\varphi'^{-1} \circ \varphi)(u), 0).$$

Anhand dieser Formel sieht man ein, dass $\varphi'^{-1} \circ \varphi$ auf U stetig differenzierbar ist.

Der bisherige Beweis zeigt, dass es zu jedem Punkt $p \in \varphi^{-1}(V \cap V')$ eine offene Umgebung U gibt, so dass $(\varphi'^{-1} \circ \varphi)|_U$ stetig differenzierbar ist. Damit ist

$$(\varphi'^{-1} \circ \varphi)|_{\varphi^{-1}(V \cap V')} : \varphi^{-1}(V \cap V') \rightarrow \varphi'^{-1}(V \cap V')$$

stetig differenzierbar. Genauso sieht man ein, dass

$$(\varphi^{-1} \circ \varphi')|_{\varphi'^{-1}(V \cap V')} : \varphi'^{-1}(V \cap V') \rightarrow \varphi^{-1}(V \cap V')$$

stetig differenzierbar ist. □

11.5.2 Bemerkung. Es sei $l \geq 1$ eine natürliche Zahl. Ein glattes parametrisiertes Flächenstück $\varphi: P \rightarrow \mathbb{R}^n$ ist vom Typ \mathcal{C}^l , wenn φ l -mal stetig differenzierbar ist. Wenn $\varphi: P \rightarrow \mathbb{R}^n$ beliebig oft stetig differenzierbar ist, also für alle $l \geq 1$ vom Typ \mathcal{C}^l ist, dann sagen wir, dass φ vom Typ \mathcal{C}^∞ ist.

Der obige Beweis zeigt, dass $(\varphi'^{-1} \circ \varphi)|_{\varphi^{-1}(V \cap V')} : \varphi^{-1}(V \cap V') \rightarrow \varphi'^{-1}(V \cap V')$ l -mal stetig differenzierbar ist, wenn φ und φ' beide vom Typ \mathcal{C}^l sind, $l \geq 1$ (vgl. Bemerkung 10.2.4).

11.5.3 Definition. Es seien $M \subset \mathbb{R}^n$ eine k -dimensionale Untermannigfaltigkeit und $f: M \rightarrow \mathbb{R}^s$ eine Abbildung. Wir sagen, dass f *differenzierbar* bzw. *stetig differenzierbar* ist, wenn es zu jedem Punkt $x \in M$ ein Gebiet P , einen Punkt $p \in P$ und ein glattes parametrisiertes Flächenstück $\varphi: P \rightarrow V \subset M \subset \mathbb{R}^n$ über P gibt, so dass $\varphi(p) = x$ und

$$f \circ \varphi: \mathbb{R}^k \rightarrow \mathbb{R}^s$$

in p differenzierbar bzw. stetig differenzierbar ist.

11.5.4 Bemerkung. Die obige Definition ist sinnvoll, weil sie nicht von der Auswahl des glatten parametrisierten Flächenstücks $\varphi: P \rightarrow \mathbb{R}^n$ abhängt. Denn für ein anderes glattes parametrisiertes Flächenstück $\varphi': P' \rightarrow V' \subset M \subset \mathbb{R}^n$ und einen anderen Punkt $p' \in P'$ mit $\varphi'(p') = x$ gilt

$$(f \circ \varphi')|_{\varphi'^{-1}(V \cap V')} = (f \circ \varphi) \circ (\varphi^{-1} \circ \varphi')|_{\varphi'^{-1}(V \cap V')}.$$

Da $f \circ \varphi$ in $p = (\varphi'^{-1} \circ \varphi)(p')$ (stetig) differenzierbar ist und $(\varphi^{-1} \circ \varphi')|_{\varphi'^{-1}(V \cap V')}$ nach Satz 11.5.1 stetig differenzierbar ist, ist $(f \circ \varphi')$ in p' (stetig) differenzierbar.

11.5.5 *Beispiel.* Sind $M \subset \mathbb{R}^n$ eine Untermannigfaltigkeit und $f: \mathbb{R}^n \rightarrow \mathbb{R}^s$ eine (stetig) differenzierbare Abbildung, dann ist

$$f|_M: M \rightarrow \mathbb{R}^s$$

eine (stetig) differenzierbare Abbildung.

Für die Untermannigfaltigkeit $O_n(\mathbb{R}) \subset \text{Mat}(n, \mathbb{R})$ aus Beispiel 11.4.5, ii), sind die Abbildungen

$$\begin{aligned} m: O_n(\mathbb{R}) \times O_n(\mathbb{R}) &\rightarrow O_n(\mathbb{R}) \subset \text{Mat}(n, \mathbb{R}) \\ (A, B) &\mapsto A \cdot B \end{aligned}$$

und

$$\begin{aligned} \text{inv}: O_n(\mathbb{R}) &\rightarrow O_n(\mathbb{R}) \\ A &\mapsto A^{-1} \end{aligned}$$

stetig differenzierbar. Im ersten Fall ist zu beachten, dass $O_n(\mathbb{R}) \times O_n(\mathbb{R})$ eine Untermannigfaltigkeit von $\text{Mat}(n, \mathbb{R}) \times \text{Mat}(n, \mathbb{R}) \cong \mathbb{R}^{2n^2}$ ist.

11.5.6 *Bemerkung.* Es seien $M \subset \mathbb{R}^n$ eine Untermannigfaltigkeit und $f: M \rightarrow \mathbb{R}^s$ eine Abbildung. Es macht keinen Sinn zu definieren, dass f zweimal stetig differenzierbar ist. Das Argument in Bemerkung 11.5.4 versagt dann, weil wir nicht erwarten können, dass $(\varphi^{-1} \circ \varphi')|_{\varphi^{-1}(V \cap V')}$ zweimal stetig differenzierbar ist.

Wollen wir von Differenzierbarkeit höherer Ordnung sprechen, dann müssen wir den Begriff der Untermannigfaltigkeit entsprechend verschärfen. Sei $l \geq 1$. Eine Teilmenge $M \subset \mathbb{R}^n$ ist eine *k-dimensionale Untermannigfaltigkeit vom Typ \mathcal{C}^l* , wenn es zu jedem Punkt $x \in M$ eine offene Umgebung $U \subset \mathbb{R}^n$ von x , ein Gebiet $P \subset \mathbb{R}^k$, einen Punkt $p \in P$ und ein glattes parametrisiertes Flächenstück $\varphi: P \rightarrow \mathbb{R}^n$ vom Typ \mathcal{C}^l über P mit den Eigenschaften

- $\varphi(p) = x$,
- $\varphi(P) = M \cap U$

gibt. Entsprechend definieren wir eine *Untermannigfaltigkeit vom Typ \mathcal{C}^∞* .

Auf einer Untermannigfaltigkeit vom Typ \mathcal{C}^l können wir von l -mal stetig differenzierbaren Abbildungen sprechen. Dazu verwenden wir Bemerkung 10.2.4.

11.6 Extrema unter Nebenbedingungen

Mit dem bis dato entwickelten Formalismus können wir den Satz über die Lagrange-Multiplikatoren 7.5.3 ohne große Anstrengung auf den Fall mehrerer Nebenbedingungen erweitern. Mit dem Begriff der Untermannigfaltigkeit ergibt sich auch eine andere Betrachtungsmöglichkeit der Problemstellung.

11.6.1 Definition. Es seien $B \subset \mathbb{R}^m$ eine offene Teilmenge, $g = (g_1, \dots, g_n): B \rightarrow \mathbb{R}^n$ eine stetig differenzierbare Abbildung und $f: B \rightarrow \mathbb{R}$ eine stetig differenzierbare Funktion. Es gelte

$$\forall x \in B: \text{Rg}(J_g(x)) = n.$$

Wir setzen

$$M := \left\{ x \in B \mid g_1(x) = \dots = g_n(x) = 0 \right\}.$$

Sei schließlich $a \in M$. Die Funktion f nimmt in a ein *lokales Minimum* bzw. *Maximum unter den Nebenbedingungen* g_1, \dots, g_n an, wenn es eine offene Umgebung $U \subset \mathbb{R}^m$ von a gibt, so dass

$$\forall x \in U \cap M: \quad f(x) \geq f(a) \quad \text{bzw.} \quad f(x) \leq f(a).$$

11.6.2 Bemerkung. Nach Satz 11.3.4 ist M eine $(m - n)$ -dimensionale Untermannigfaltigkeit von \mathbb{R}^m , und $f|_M: M \rightarrow \mathbb{R}$ ist nach Beispiel 11.5.5 eine stetig differenzierbare Funktion auf M . Die Suche nach einem lokalen Minimum bzw. Maximum unter den Nebenbedingungen g_1, \dots, g_n ist damit die Suche nach einem lokalen Minimum bzw. Maximum der Funktion $f|_M$ auf der Untermannigfaltigkeit M . Damit lassen sich die Probleme der Extremwertbestimmung differenzierbarer Funktionen aus Kapitel 7 und der Bestimmung von Extremwerten unter Nebenbedingungen zu dem Problem der Extremwertbestimmung differenzierbarer Funktionen auf Untermannigfaltigkeiten des \mathbb{R}^m zusammenfassen.

11.6.3 Satz (Methode der Lagrangeschen Multiplikatoren). *Die Situation sei dieselbe wie in Definition 11.6.1. Wenn die Funktion f in dem Punkt $a \in M$ ein lokales Minimum oder Maximum unter den Nebenbedingungen g_1, \dots, g_n hat, dann existieren eindeutig bestimmte reelle Zahlen $\lambda_1, \dots, \lambda_n$, so dass*

$$\text{Grad}f(a) = \lambda_1 \cdot \text{Grad}g_1(a) + \dots + \lambda_n \cdot \text{Grad}g_n(a).$$

11.6.4 Definition. Die Zahlen $\lambda_1, \dots, \lambda_n$ heißt *Lagrangesche Multiplikatoren*.

Beweis. Wir wählen ein Gebiet P und ein glattes parametrisiertes Flächenstück $\varphi: P \rightarrow V \subset M \subset \mathbb{R}^m$, so dass ein Punkt $p \in P$ mit $\varphi(p) = a$ existiert. Wenn f ein lokales Minimum bzw. Maximum unter den Nebenbedingungen g_1, \dots, g_n hat, dann nimmt die stetig differenzierbare Funktion $f \circ \varphi: P \rightarrow \mathbb{R}$ an der Stelle p ein lokales Minimum bzw. Maximum an. Nach Satz 7.1.2 gilt

$$0 = \text{Grad}(f \circ \varphi)(p) = \text{Grad}f(\varphi(p)) \cdot J_\varphi(p). \quad (11.9)$$

Auf der anderen Seite hat man

$$\forall q \in P: \quad (g \circ \varphi)(q) = 0,$$

so dass insbesondere

$$0 = J_{g \circ \varphi}(p) = J_g(\varphi(p)) \cdot J_\varphi(p)$$

gilt. Da $\text{Rg}(J_g(\varphi(p))) = n$, sind die Zeilen der Matrix $J_g(\varphi(p))$ linear unabhängig. Diese Zeilen sind

$$\text{Grad}g_1(a), \dots, \text{Grad}g_n(a).$$

Nach Definition 11.6.1 gilt $\text{Rg}(J_\varphi(p)) = m - n$. Deshalb hat der Vektorraum

$$N := \left\{ v \in \mathbb{R}^m \mid v \cdot J_\varphi(p) = 0 \right\}$$

die Dimension n . Wir schließen, dass

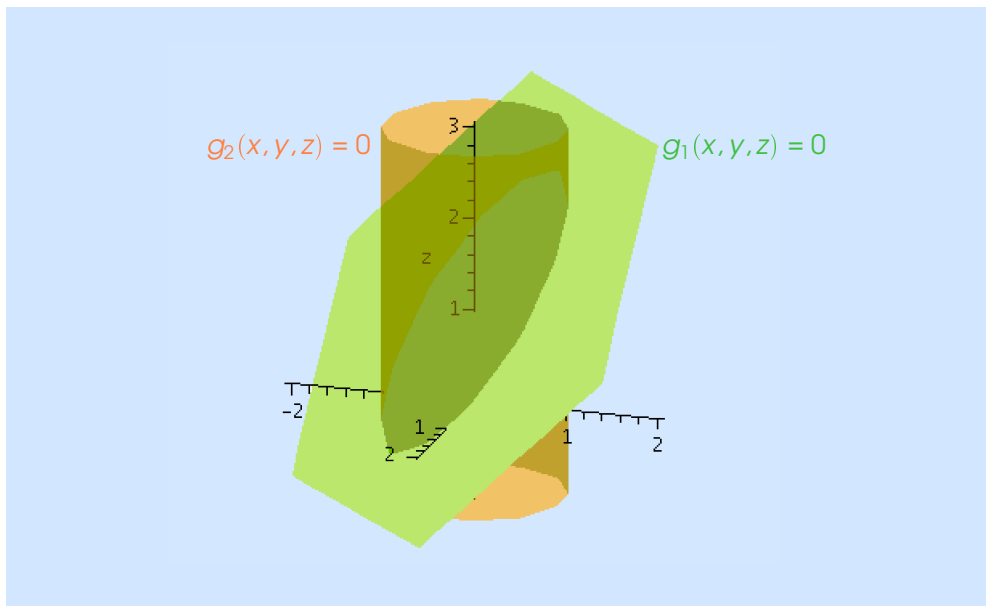
$$\text{Grad}g_1(a), \dots, \text{Grad}g_n(a)$$

eine Basis für N ist. Gemäß (11.9) gilt $\text{Grad}f(p) \in N$. Daraus folgt sofort die Behauptung. \square

Wir sehen, dass der Formalismus der Untermannigfaltigkeiten zu einer starken Vereinfachung des Beweises geführt hat (vgl. Beweis von Satz 7.5.3).

11.6.5 Beispiel. Es seien

$$\begin{aligned} f: \mathbb{R}^3 &\longrightarrow \mathbb{R} \\ (x, y, z) &\longmapsto 3x + 2y + z, \\ g_1: \mathbb{R}^3 &\longrightarrow \mathbb{R} \\ (x, y, z) &\longmapsto x - y + z - 1, \\ g_2: \mathbb{R}^3 &\longrightarrow \mathbb{R} \\ (x, y, z) &\longmapsto x^2 + y^2 - 1. \end{aligned}$$



Für $(x, y, z) \in \mathbb{R}^3$ gilt

$$J_{(g_1, g_2)}(x, y, z) = \begin{pmatrix} 1 & -1 & 1 \\ 2x & 2y & 0 \end{pmatrix}$$

und folglich

$$\forall (x, y, z) \in B := (\mathbb{R}^2 \setminus \{(0, 0)\}) \times \mathbb{R}: \operatorname{Rg}(J_{(g_1, g_2)}(x, y, z)) = 2.$$

Damit ist

$$M := \left\{ (x, y, z) \in \mathbb{R}^3 \mid g_1(x, y, z) = 0 = g_2(x, y, z) \right\} \subset B$$

eine eindimensionale Untermannigfaltigkeit. Da M kompakt ist,⁹ nimmt f nach Folgerung 3.4.2 auf M ein Minimum und ein Maximum an. Wir haben

$$\forall (x, y, z) \in \mathbb{R}^3: \operatorname{Grad} f(x, y, z) = (3, 2, 1).$$

Folglich sind reelle Zahlen x, y, z, λ_1 und λ_2 mit

$$(3, 2, 1) = \lambda_1 \cdot (1, -1, 1) + \lambda_2 \cdot (2x, 2y, 0)$$

⁹ M ist abgeschlossen und beschränkt.

gesucht, d.h.

$$\begin{aligned}\lambda_1 &= 1 \\ 2 &= 2 \cdot x \cdot \lambda_2 \\ 3 &= 2 \cdot y \cdot \lambda_2.\end{aligned}$$

Offenbar muss $\lambda_2 \neq 0$ gelten sowie

$$x = \frac{1}{\lambda_2}, \quad y = \frac{3}{2\lambda_2}, \quad z = -x + y + 1 = 1 + \frac{1}{2\lambda_2}.$$

Bei der letzten Gleichung haben wir die Nebenbedingung $g_1(x, y, z) = 0$ benutzt. Die Nebenbedingung $g_2(x, y, z) = 0$ führt zu der Gleichung

$$1 = \frac{1}{\lambda_2^2} + \frac{9}{4\lambda_2^2} = \frac{13}{4\lambda_2^2}, \quad \text{also} \quad \lambda_2 = \pm \frac{1}{2} \cdot \sqrt{13}.$$

Damit finden wir

$$a_+ = \left(\frac{2}{\sqrt{13}}, \frac{3}{\sqrt{13}}, 1 + \frac{1}{\sqrt{13}} \right)$$

mit

$$f(a_+) = \frac{6}{\sqrt{13}} + \frac{6}{\sqrt{13}} + 1 + \frac{1}{\sqrt{13}} = 1 + \sqrt{13}$$

und

$$a_- = \left(-\frac{2}{\sqrt{13}}, -\frac{3}{\sqrt{13}}, 1 - \frac{1}{\sqrt{13}} \right)$$

mit

$$f(a_-) = -\frac{6}{\sqrt{13}} - \frac{6}{\sqrt{13}} + 1 - \frac{1}{\sqrt{13}} = 1 - \sqrt{13}.$$

Da $f(a_+) > f(a_-)$ und f auf M ein Minimum und ein Maximum annimmt, folgern wir, dass f an der Stelle a_- das globale Minimum unter den Nebenbedingungen g_1, g_2 annimmt und an der Stelle a_+ das globale Maximum.

Anhang. Abstrakte Mannigfaltigkeiten

Es seien $l \geq 1$ und $M \subset \mathbb{R}^n$ eine k -dimensionale Untermannigfaltigkeit vom Typ \mathcal{C}^l bzw. \mathcal{C}^∞ . Nach unseren Betrachtungen aus Abschnitt 11.3 finden wir eine offene Überdeckung $(V_i)_{i \in I}$ von M , offene Teilmengen $U_i \subset \mathbb{R}^k$ und Homöomorphismen $h_i: V_i \rightarrow U_i$,¹⁰ $i \in I$, so dass die Abbildungen

$$h_{ij}: U_i \supset h_i(U_i \cap U_j) \xrightarrow{h_j \circ h_i^{-1}} h_j(U_i \cap U_j) \subset U_j, \quad i, j \in I, \quad (11.10)$$

l -mal bzw. beliebig oft stetig differenzierbar sind.

Damit können wir uns vorstellen, dass wir M erhalten, indem wir die offenen Teilmengen $(U_i)_{i \in I}$ von \mathbb{R}^k vermöge der Abbildungen $h_{ij}, i, j \in I$, miteinander verkleben (vgl. die Illustration auf Seite 171). Bei dieser Vorstellung spielt die Tatsache, dass M eine Teilmenge von \mathbb{R}^n ist, keine Rolle mehr. Wir können somit noch allgemeinere Objekte zulassen, die aus der Verklebung offener Teilmengen von \mathbb{R}^k über differenzierbare Abbildungen entstehen. In diesem Abschnitt wollen wir die heute gebräuchliche Struktur der differenzierbaren Mannigfaltigkeit angeben.

¹⁰Wir drehen hier die Richtung der Pfeile um, um uns auf die übliche Definition einer Mannigfaltigkeit vorzubereiten.

11.A.1 Definition. Es seien X ein topologischer Raum und $(U_i)_{i \in I}$ eine Familie von offenen Teilmengen von X . Die Familie $(U_i)_{i \in I}$ ist eine *Basis der Topologie*, wenn jede offene Teilmenge $V \subset X$ als Vereinigung von Mengen dieser Familie geschrieben werden kann:

$$V = \bigcup_{i \in I(V)} U_i, \quad I(V) := \{i \in I \mid U_i \subset V\}.$$

Ist die Indexmenge I abzählbar, so ist $(U_i)_{i \in I}$ eine *abzählbare Basis der Topologie*.

11.A.2 Beispiel. Es sei $I = \{(a, r) \mid a \in \mathbb{Q}^n, r \in \mathbb{Q}, r > 0\}$. Dies ist eine abzählbare Menge (s. (12), Anhang zu Kapitel I). Die Familie $(B(a, r))_{(a, r) \in I}$ ist die Familie der offenen Bälle in \mathbb{R}^n mit rationalem Zentrum und rationalem Radius. Sie ist eine abzählbare Basis der Topologie von \mathbb{R}^n .

11.A.3 Definition. a) Ein topologischer Raum X ist eine (*topologische*) *Mannigfaltigkeit der Dimension n* , wenn

- X eine abzählbare Basis der Topologie besitzt,
- X hausdorffsch ist und
- zu jedem Punkt $x \in M$ eine offene Umgebung $V \subset M$ von x , eine offene Teilmenge $U \subset \mathbb{R}^n$ und ein Homöomorphismus $h: V \rightarrow U$ existieren.

b) Es sei M eine topologische Mannigfaltigkeit der Dimension n . Eine *Karte von M* ist ein Homöomorphismus $h: V \rightarrow U$ von einer offenen Teilmenge $V \subset M$ auf eine offene Teilmenge $U \subset \mathbb{R}^n$.

c) Es sei M eine topologische Mannigfaltigkeit der Dimension n . Ein *Atlas von M* ist eine Familie $(h_i: V_i \rightarrow U_i)_{i \in I}$ von Karten, so dass $(V_i)_{i \in I}$ eine offene Überdeckung von M ist:

$$M = \bigcup_{i \in I} V_i.$$

11.A.4 Beispiele. i) Natürlich ist \mathbb{R}^n eine topologische Mannigfaltigkeit der Dimension n .

ii) Wir betrachten die Einheitssphäre

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

mit den offenen Teilmengen

$$V_i^\pm := \{x = (x_1, \dots, x_{n+1}) \mid \pm x_i > 0\}, \quad i = 1, \dots, n+1.$$

Es sei

$$D := \{y = (y_1, \dots, y_n) \in \mathbb{R}^n \mid \|y\| < 1\}$$

der offene Ball vom Radius 1. Man überprüft leicht, dass

$$\begin{aligned} h_j^\pm: V_j^\pm &\longrightarrow D \\ (x_1, \dots, x_{n+1}) &\longmapsto (x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n) \end{aligned}$$

ein Homöomorphismus ist, $i = 1, \dots, n+1, j = \pm 1$. Damit ist S^n eine topologische Mannigfaltigkeit, und die Karten $(h_i^\pm: V_i^\pm \rightarrow D)_{i=1, \dots, n+1}$ bilden einen Atlas von S^n .

iii) Allgemeiner ist jede k -dimensionale Untermannigfaltigkeit $M \subset \mathbb{R}^n$ eine topologische Mannigfaltigkeit der Dimension k (Definition 11.3.3 und Lemma 11.A.3).

11.A.5 *Bemerkungen.* i) Wenn M ein topologischer Raum ist, der die zweite und dritte Eigenschaft aus Definition 11.A.3, a), erfüllt, dann können wir ebenso von Karten und Atlanten reden. Die Bedingung, eine abzählbare Topologie zu haben, ist dann äquivalent dazu, einen abzählbaren Atlas zu haben: Wenn $(h_i: V_i \rightarrow U_i)_{i \in I}$ ein abzählbarer Atlas ist, dann bilden die offenen Teilmengen der Form

$$h_i^{-1}(B), \quad i \in I, B \subset U_i \text{ ein offener Ball mit rationalem Zentrum und Radius,}$$

eine abzählbare Basis der Topologie von M .

Wenn umgekehrt $(h_i: V_i \rightarrow U_i)_{i \in I}$ ein beliebiger Atlas ist und $(W_j)_{j \in J}$ eine abzählbare Basis der Topologie, dann ist $(W_j)_{j \in J'}$ mit

$$J' := \{j \in J \mid \exists i \in I : W_j \subset V_i\}$$

immer noch eine abzählbare Basis der Topologie von M (Begründung?). Für jeden Index $j \in J'$ wählen wir einen Index $i(j) \in I$ mit $W_j \subset V_{i(j)}$ und definieren die Karte $k_j := h_{i(j)|W_j}: W_j \rightarrow T_j := h_{i(j)}(W_j)$. Damit ist $(k_j: W_j \rightarrow T_j)_{j \in J'}$ ein abzählbarer Atlas.

ii) Es seien M eine topologische Mannigfaltigkeit und $(h_i: V_i \rightarrow U_i)_{i \in I}$ ein Atlas von M . Damit können wir die „Verklebebeschreibung“ von M explizit machen: Wir erklären die Homöomorphismen $h_{ij}: h_i(U_i \cap U_j) \rightarrow h_j(U_i \cap U_j)$, $i, j \in I$, wie in (11.10), definieren

$$\tilde{M} = \bigsqcup_{i \in I} V_i$$

als diskunkte Vereinigung der Mengen V_i , $i \in I$, und führen folgende Äquivalenzrelation auf \tilde{M} ein:

$$V_i \supset h_i(U_i \cap U_j) \ni x \sim h_{ij}(x) \in h_j(U_i \cap U_j) \subset V_j, \quad i, j \in I.$$

(Es ist eine leichte Übung zu überprüfen, dass „ \sim “ wirklich eine Äquivalenzrelation ist.) Es seien \overline{M} die Menge der Äquivalenzklassen und $p: \tilde{M} \rightarrow \overline{M}$ die Abbildung, die einem Element seine Äquivalenzklasse zuweist. Wir definieren auf \tilde{M} folgendermaßen eine Topologie: Eine Teilmenge

$$W = \bigsqcup_{i \in I} W_i, \quad W_i := W \cap V_i, \quad i \in I,$$

ist *offen*, wenn $W_i \subset V_i$ für alle $i \in I$ offen ist. Die Menge \overline{M} erhält die *Quotiententopologie*: Eine Teilmenge $U \subset \overline{M}$ ist *offen*, wenn $p^{-1}(U) \subset \tilde{M}$ offen ist. Durch die Karten $h_i: V_i \rightarrow U_i$ wird ein Homöomorphismus $\tilde{M} \rightarrow \overline{M}$ induziert.

Damit haben wir gezeigt, dass eine topologische Mannigfaltigkeit der Dimension n durch Verklebung offener Teilmengen von \mathbb{R}^n mittels Homöomorphismen entsteht. Wie Teil i) zeigt, verkleben wir nur abzählbar viele Stücke. Diese technische Bedingung wollen wir nicht weiter erläutern. Die Forderung nach der Hausdorff-Eigenschaft verbietet gewisse ungeschickte Verklebungen (s. (5), Seite 2).

11.A.6 *Beispiel.* Dieses Beispiel macht den Vorzug der Definition einer abstrakten Mannigfaltigkeit gegenüber der einer Untermannigfaltigkeit deutlich. Auf der Einheitssphäre S^n betrachten wir die sogenannte *Antipodenabbildung*

$$\begin{aligned} \alpha: S^n &\longrightarrow S^n \\ x &\longmapsto -x. \end{aligned}$$

Durch

$$x \sim \alpha(x), \quad x \in S^n,$$

ist eine Äquivalenzrelation auf S^n gegeben. Die Menge der Äquivalenzklassen wird mit $\mathbb{R}P^n$ bezeichnet. Wie in der vorigen Bemerkung statfen wir $\mathbb{R}P^n$ mit der Quotiententopologie aus. Man verifiziert leicht, dass $\mathbb{R}P^n$ eine abzählbare Basis der Topologie hat und hausdorffsch ist. Es sei $[x] \in \mathbb{R}P^n$ die Äquivalenzklasse des Punkts $x \in S^n$. Wir definieren

$$V_i := \{ [x] = [x_1, \dots, x_{n+1}] \in \mathbb{R}P^n \mid x_i \neq 0 \}, \quad i = 1, \dots, n+1.$$

Das Urbild von V_i unter der natürlichen Abbildung $S^n \rightarrow \mathbb{R}P^n$ ist $V_i^- \cup V_i^+$ und daher offen, $i = 1, \dots, n+1$. Nach Definition ist V_i offen in $\mathbb{R}P^n$, $i = 1, \dots, n+1$. Die induzierten Abbildungen $V_{i,\pm 1} \rightarrow V_i$ sind Homöomorphismen, so dass wir aus dem in Beispiel 11.A.4, ii), konstruierten Atlas für S^n einen Atlas für $\mathbb{R}P^n$ erhalten. Die induzierten Karten sind

$$\begin{aligned} k_i: V_i &\rightarrow D \\ [x_1, \dots, x_{n+1}] &\rightarrow \text{sign}(x_i) \cdot (x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n), \quad i = 1, \dots, n+1. \end{aligned}$$

Damit ist gezeigt, dass $\mathbb{R}P^n$ eine topologische Mannigfaltigkeit der Dimension n ist. Sie heißt *n-dimensionaler reell projektiver Raum*. Ihre Konstruktion liefert keine Einbettung in einen \mathbb{R}^s . Solche Einbettungen existieren jedoch, z.B. ist

$$\begin{aligned} \mathbb{R}P^2 &\rightarrow \mathbb{R}^4 \\ [x, y, z] &\mapsto \frac{1}{x^2 + y^2 + z^2} \cdot (yz, xz, xy, x^2 + 2y^2 + 3z^2) \end{aligned}$$

eine. Es ist nicht möglich, $\mathbb{R}P^2$ als Untermannigfaltigkeit von \mathbb{R}^3 zu realisieren. Allerdings kann man versuchen, Modelle im \mathbb{R}^4 „möglichst gut“ in den \mathbb{R}^3 zu projizieren, und erhält auf diese Weise schöne Veranschaulichungen der reell projektiven Ebene (1).

Der allgemeine Begriff der Mannigfaltigkeit ermöglicht es z.B., Konstruktionen wie die des projektiven Raums schnell durchzuführen, ohne sich um eine Einbettung in einen \mathbb{R}^s kümmern zu müssen, die oftmals gar nicht interessiert.

11.A.7 Definition. Es seien $k \geq 1$ und M eine topologische Mannigfaltigkeit.

a) Ein Atlas $(h_i: V_i \rightarrow U_i)_{i \in I}$ ist vom Typ \mathcal{E}^k bzw. \mathcal{E}^∞ , wenn die „Kartenwechsel“

$$h_{ij}: h_i(U_i \cap U_j) \rightarrow h_j(U_i \cap U_j)$$

aus (11.10) für alle $i, j \in I$ k -mal bzw. beliebig oft stetig differenzierbar sind.

b) Zwei Atlanten $(h_i: V_i \rightarrow U_i)_{i \in I}$ und $(k_j: W_j \rightarrow T_j)_{j \in J}$ vom Typ \mathcal{E}^k bzw. \mathcal{E}^∞ sind *äquivalent*, wenn der Atlas, der aus allen Karten $(h_i: V_i \rightarrow U_i)_{i \in I}$ und $(k_j: W_j \rightarrow T_j)_{j \in J}$ besteht, ebenfalls vom Typ \mathcal{E}^k bzw. \mathcal{E}^∞ ist.

c) Eine *differenzierbare Struktur vom Typ \mathcal{E}^k bzw. \mathcal{E}^∞ auf M* ist die Äquivalenzklasse eines Atlantes vom Typ \mathcal{E}^k bzw. \mathcal{E}^∞ .

d) Eine *differenzierbare Mannigfaltigkeit vom Typ \mathcal{E}^k bzw. \mathcal{E}^∞* (oder kurz \mathcal{E}^k - bzw. \mathcal{E}^∞ -Mannigfaltigkeit) ist eine topologische Mannigfaltigkeit zusammen mit einer differenzierbaren Struktur vom Typ \mathcal{E}^k bzw. \mathcal{E}^∞ .

11.A.8 Beispiel. Der in Beispiel 11.A.4, ii), bzw. Beispiel 11.A.6 eingeführte Atlas für S^n bzw. $\mathbb{R}P^n$ ist vom Typ \mathcal{E}^∞ .

Definition 11.5.3 können wir sofort verallgemeinern, um differenzierbare Funktionen auf einer \mathcal{E}^1 -Mannigfaltigkeit zu erklären. Damit können wir auf differenzierbaren Mannigfaltigkeiten Analysis betreiben. Dem Studium differenzierbarer Mannigfaltigkeiten widmen sich u.A. die Differentialtopologie (s. z.B. (5)) und -geometrie (s. z.B. (10)).

Die Aufgabenblätter

Blatt 1

Aufgabe A.1.1 (Metrische Räume).

a) Es sei $f: \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}_{\geq 0}$ eine Funktion, die streng monoton wachsend und konkav ist und $f(0) = 0$ erfüllt (Beispiel?). Weiter sei (X, d) ein metrischer Raum. Zeigen Sie, dass

$$\begin{aligned}d_f: X \times X &\longrightarrow \mathbb{R} \\(x, y) &\longmapsto f(d(x, y))\end{aligned}$$

eine Metrik auf X ist.

b) Es seien (X, d_X) und (Y, d_Y) metrische Räume. Wir setzen

$$\begin{aligned}d_1: (X \times Y) \times (X \times Y) &\longrightarrow \mathbb{R} \\((x_1, y_1), (x_2, y_2)) &\longmapsto \max\{d_X(x_1, x_2), d_Y(y_1, y_2)\}, \\d_2: (X \times Y) \times (X \times Y) &\longrightarrow \mathbb{R} \\((x_1, y_1), (x_2, y_2)) &\longmapsto d_X(x_1, x_2) + d_Y(y_1, y_2), \\d_3: (X \times Y) \times (X \times Y) &\longrightarrow \mathbb{R} \\((x_1, y_1), (x_2, y_2)) &\longmapsto \sqrt{d_X(x_1, x_2)^2 + d_Y(y_1, y_2)^2}.\end{aligned}$$

Überprüfen Sie, dass d_1 , d_2 und d_3 Metriken auf $X \times Y$ sind.

c) Es sei \mathbb{R} mit der Metrik $(x, y) \mapsto |x - y|$ versehen. Beschreiben Sie die Metrik d_2 aus Teil b) auf $\mathbb{R}^2 = \mathbb{R} \times \mathbb{R}$ und skizzieren Sie $B(0, 1)$ für diese Metrik.

Aufgabe A.1.2 (Die SNCF-Metrik).

Wir beginnen mit $(\mathbb{R}^2, \|\cdot\|)$, der Ebene \mathbb{R}^2 zusammen mit der euklidischen Norm. Wir definieren

$$\begin{aligned}d: \mathbb{R}^2 \times \mathbb{R}^2 &\longrightarrow \mathbb{R} \\(x, y) &\longmapsto \begin{cases} \|x - y\|, & \text{falls } x, y \text{ und } 0 \text{ auf einer Geraden liegen} \\ \|x\| + \|y\|, & \text{sonst} \end{cases}.\end{aligned}$$

Überprüfen Sie, dass d eine Metrik ist. Skizzieren Sie die offenen Bälle um einen Punkt $x \in \mathbb{R}^2$.

Aufgabe A.1.3 (Über Vierecke).

a) Es sei (X, d) ein metrischer Raum. Zeigen Sie, dass die *Vierecksungleichung*

$$|d(u, v) - d(x, y)| \leq d(u, x) + d(v, y), \quad u, v, x, y \in X,$$

gilt. Was bedeutet diese Ungleichung für ein Viereck in \mathbb{R}^2 ?

b) Es sei $\|\cdot\|$ die euklidische Norm auf \mathbb{R}^n . Verifizieren Sie, dass die *Parallelogrammgleichung*

$$\|x + y\|^2 + \|x - y\|^2 = 2 \cdot (\|x\|^2 + \|y\|^2), \quad x, y \in \mathbb{R}^n,$$

gilt. Erläutern Sie die Bedeutung dieser Gleichung.

c) Zeigen Sie, dass die Parallelogrammgleichung für die Maximum-Norm $\|\cdot\|_M$ auf \mathbb{R}^2 **nicht** gilt.

Aufgabe A.1.4 (Äquivalente Normen und offene Mengen).

Es sei V ein \mathbb{R} -Vektorraum. Zwei Normen $\|\cdot\|_1$ und $\|\cdot\|_2$ auf V heißen *äquivalent*, wenn es reelle Zahlen $r > 0$ und $R > 0$ gibt, so dass

$$\forall v \in V: r \cdot \|v\|_1 \leq \|v\|_2 \leq R \cdot \|v\|_1.$$

Die Norm $\|\cdot\|_i$ definiert eine Metrik d_i und damit eine Topologie τ_i auf V , $i = 1, 2$. Die Normen $\|\cdot\|_1$ und $\|\cdot\|_2$ seien äquivalent. Beweisen Sie $\tau_1 = \tau_2$.

Blatt 2

Aufgabe A.2.1 (Topologische Räume).

a) Zeigen Sie, dass in einem topologischen Raum Vereinigungen zweier abgeschlossener Mengen und beliebige Durchschnitte abgeschlossener Teilmengen ebenfalls abgeschlossen sind.

b) Es sei $X := \{1, 2, 3\}$. Geben Sie die Potenzmenge $\mathcal{P}(X)$ an, und listen Sie alle Topologien auf X auf.

c) Verifizieren Sie, dass

$$\tau := \{U \subset \mathbb{R} \mid U = \emptyset \vee \#(\mathbb{R} \setminus U) < \infty\} \subset \mathcal{P}(\mathbb{R})$$

eine Topologie auf \mathbb{R} ist. Ist \mathbb{R} mit dieser Topologie hausdorffsch?

Aufgabe A.2.2 (Ränder).

a) Bestimmen Sie den Rand ∂A , den Abschluss $\bar{A} = A \cup \partial A$ sowie den offenen Kern $\overset{\circ}{A} = A \setminus \partial A$ für folgende Teilmengen von \mathbb{R} bzw. \mathbb{R}^2 :

1. $A := \mathbb{Q} \cap \{x \in \mathbb{R} \mid 0 < x^2 < 2\} \subset \mathbb{R}$.
2. $A := (2, 5] \times [3, 7) \subset \mathbb{R}^2$.
3. $A := \{x \in \mathbb{R} \mid x \cdot (x + 2) > 0\} \subset \mathbb{R}$.
4. $A := \{(x, y) \in \mathbb{R}^2 \mid \exists n \in \mathbb{N} \setminus \{0\} : y = 1/n \wedge |x| \leq 1/n\} \subset \mathbb{R}^2$.
5. $A := ((0, 1) \times (0, 1)) \setminus \bigcup_{n \geq 2} \{(1/n, y) \mid y < 1/2\} \subset \mathbb{R}^2$.
6. $A := \mathbb{R}^2 \setminus \{(x, y) \in \mathbb{R}^2 \mid x \in \mathbb{Q} \vee y \in \mathbb{Q}\} \subset \mathbb{R}^2$.

(Es wird die Standardtopologie, die durch die euklidische Metrik induziert wird, zugrundegelegt.)

b) Es seien $A \subset \mathbb{R}^m$ und $B \subset \mathbb{R}^n$ Teilmengen. Die Räume \mathbb{R}^m , \mathbb{R}^n und \mathbb{R}^{m+n} seien mit der euklidischen Norm versehen. Beweisen Sie

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

Aufgabe A.2.3 (Abschlüsse und offene Kerne).

Es seien (X, d) ein metrischer Raum und $A \subset X$ eine Teilmenge.

a) Zeigen Sie, dass \bar{A} die kleinste abgeschlossene Teilmenge von X ist, die A enthält, d.h., dass für jede abgeschlossene Teilmenge B von X

$$A \subset B \Rightarrow \bar{A} \subset B$$

gilt.
Beweisen Sie auch

$$\bar{A} = \bigcap_{\substack{B \subset X \\ B \text{ abgeschlossen und } A \subset B}} B.$$

b) Schließen Sie, dass $\overset{\circ}{A}$ die größte offene Teilmenge von X ist, die in A enthalten ist, d.h., dass für jede offene Teilmenge B von X

$$B \subset A \Rightarrow B \subset \overset{\circ}{A}$$

gilt.
Zeigen Sie auch

$$\overset{\circ}{A} = \bigcup_{\substack{B \subset X \\ B \text{ offen und } B \subset A}} B.$$

Aufgabe A.2.4 (Ein unvollständiger metrischer Raum).
Es seien $I := [0, 1] \subset \mathbb{R}$ und $\mathcal{C}(I)$ der Vektorraum aller stetigen Funktionen $f: I \rightarrow \mathbb{R}$ (s. Skript zur Analysis I, Abschnitt 3.7). Man definiere

$$\begin{aligned} \|\cdot\|: \mathcal{C}(I) &\rightarrow \mathbb{R} \\ f &\mapsto \sqrt{\int_0^1 f(x)^2 dx}. \end{aligned}$$

a) Zeigen Sie, dass $\|\cdot\|$ eine Norm auf $\mathcal{C}(I)$ ist.
b) Es sei $(f_n)_{n \geq 2}$ die Folge in $\mathcal{C}(I)$ mit

$$f_n(x) := \begin{cases} 1, & x \in \left[0, \frac{1}{2}\right] \\ 1 - \left(x - \frac{1}{2}\right) \cdot n, & x \in \left(\frac{1}{2}, \frac{1}{2} + \frac{1}{n}\right) \\ 0, & x \in \left[\frac{1}{2} + \frac{1}{n}, 1\right] \end{cases}, \quad n \geq 2.$$

Skizzieren Sie f_2 und f_3 . Weisen Sie nach, dass es sich bei $(f_n)_{n \geq 2}$ um eine Cauchy-Folge handelt, die **nicht** konvergiert.

Blatt 3

Aufgabe A.3.1 (Abstände).
Es seien (X, d) ein metrischer Raum, $x \in X$ und $A \subset X$ eine Teilmenge. Der Abstand von x zu A ist

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

Beweisen Sie folgende Aussage:

Wenn A **abgeschlossen** ist, dann gilt: $\text{dist}(x, A) = 0 \iff x \in A$.

Aufgabe A.3.2 (Das Lemma von Riesz¹¹).
Es seien $(V, \|\cdot\|)$ ein normierter Vektorraum und $W \subset V$ ein **echter, abgeschlossener** Unterraum. Weisen Sie nach, dass es zu jedem $\delta > 0$ einen Vektor $v \in V$ mit

$$\|v\| = 1 \quad \text{und} \quad \text{dist}(v, W) > 1 - \delta$$

gibt.

Hinweis. Es genügt, das Lemma für $0 < \delta < 1$ zu zeigen. Finden Sie zunächst Vektoren $v' \in V \setminus W$ und $w \in W$ mit $\|v' - w\| \leq \text{dist}(v', W)/(1 - \delta)$.

Aufgabe A.3.3 (Quader sind folgenkompakt).

Es seien (X, d) ein metrischer Raum. Eine Teilmenge $A \subset X$ heißt *folgenkompakt*, wenn jede Folge $(x_k)_{k \in \mathbb{N}}$ mit $x_k \in A$, $k \in \mathbb{N}$, eine Teilfolge $(x_{k_i})_{i \in \mathbb{N}}$ besitzt, die gegen ein Element in A konvergiert.

Wir betrachten \mathbb{R}^n mit der euklidischen Metrik. Gegeben seien reelle Zahlen $a_i \leq b_i$, $i = 1, \dots, n$. Zeigen Sie, dass der abgeschlossene Quader

$$Q := [a_1, b_1] \times \cdots \times [a_n, b_n] \subset \mathbb{R}^n$$

folgenkompakt ist.

Bemerkung. In der Vorlesung wird gezeigt, dass die Begriffe „kompakt“ und „folgenkompakt“ für Teilmengen metrischer Räume äquivalent sind. Somit erhalten Sie einen alternativen Beweis dafür, dass Q kompakt ist.

Aufgabe A.3.4 (Kompaktheit der Einheitskugel).

Es seien $(V, \|\cdot\|)$ ein normierter Vektorraum und

$$B := \{v \in V \mid \|v\| \leq 1\}$$

die *Einheitskugel* in V .

a) Zeigen Sie, dass B abgeschlossen ist.

b) Benutzen Sie das Lemma von Riesz, um zu beweisen, dass B genau dann (folgen)kompakt ist, wenn V endlichdimensional ist.

Bemerkung. In jedem normierten **unendlich**dimensionalen Vektorraum gibt es also abgeschlossene und beschränkte Teilmengen, die nicht kompakt sind.

Blatt 4

Aufgabe A.4.1 (Topologische Formulierung der Stetigkeit).

Es seien (X, d_X) und (Y, d_Y) metrische Räume und $f: X \rightarrow Y$ eine Abbildung.

a) Nehmen Sie an, dass f stetig ist. Zeigen Sie, dass dann für jede Teilmenge $V \subset Y$ gilt:

$$\begin{aligned} V \text{ abgeschlossen} &\Rightarrow f^{-1}(V) \text{ abgeschlossen} \\ V \text{ offen} &\Rightarrow f^{-1}(V) \text{ offen} \end{aligned}$$

b) Nun erfülle f folgende Eigenschaft:

$$\forall V \subset Y: V \text{ offen} \Rightarrow f^{-1}(V) \text{ offen.}$$

Beweisen Sie, dass f stetig ist.

Aufgabe A.4.2 (Stetige Abbildungen zwischen topologischen Räumen).

Es seien (X, \mathcal{T}_X) und (Y, \mathcal{T}_Y) topologische Räume. Wir sagen, dass eine Abbildung $f: X \rightarrow Y$ *stetig* ist, wenn gilt:

$$\forall U \in \mathcal{T}_Y: f^{-1}(U) \in \mathcal{T}_X,$$

d.h., wenn Urbilder offener Mengen wieder offen sind (vgl. Aufgabe 1).

a) Es sei X eine Menge. Beweisen Sie, dass es genau eine Topologie \mathcal{T}_X auf X

¹¹Frigyes Riesz (Vorname auch dt. Friedrich oder frz. Frédéric, 1880 - 1956), ungarischer Mathematiker.

gibt, so dass für jeden topologischen Raum (Y, τ_Y) jede Abbildung $f: X \rightarrow Y$ stetig ist.

b) Es sei Y eine Menge. Weisen Sie nach, dass es auf Y genau eine Topologie τ_Y gibt, so dass für jeden topologischen Raum (X, τ_X) jede Abbildung $f: X \rightarrow Y$ stetig ist.

Aufgabe A.4.3 (Additive und lineare Abbildungen).

Es seien $(V, \|\cdot\|_V)$ und $(W, \|\cdot\|_W)$ zwei normierte \mathbb{R} -Vektorräume. Eine Abbildung $f: V \rightarrow W$ heißt *additiv*, wenn

$$\forall v_1, v_2 \in V : f(v_1 + v_2) = f(v_1) + f(v_2)$$

gilt.

a) Zeigen Sie, dass für jede additive Abbildung $f: V \rightarrow W$, jedes $\lambda \in \mathbb{Q}$ und jedes $v \in V$

$$f(\lambda \cdot v) = \lambda \cdot f(v)$$

gilt.

b) Beweisen Sie, dass eine stetige additive Abbildung linear ist.

c) Zeigen Sie, dass eine lineare Abbildung zwischen endlich-dimensionalen \mathbb{R} -Vektorräumen **gleichmäßig** stetig ist.

Aufgabe A.4.4 (Grenzwerte von Funktionen).

a) Zeigen Sie, dass die Funktion

$$f: \mathbb{R}^2 \rightarrow \mathbb{R} \\ (x, y) \mapsto \begin{cases} \frac{\sin(x^2 + y^2)}{x^2 + y^2}, & \text{falls } (x, y) \neq (0, 0) \\ 1, & \text{falls } (x, y) = (0, 0) \end{cases}$$

stetig ist.

b) Weisen Sie nach, dass

$$\lim_{(x,y) \rightarrow (0,0)} \frac{x^2 - y^2}{x^2 + y^2}$$

nicht existiert. (**Hinweis:** Polarkoordinaten sind hier überflüssig.)

c) Auf $\mathbb{R}^2 \setminus \{(0,0)\}$ sei f durch

$$(x, y) \mapsto \frac{xy^3}{x^2 + 4y^6}$$

gegeben. Verifizieren Sie, dass für alle $m \in \mathbb{R}$

$$\lim_{x \rightarrow 0} f(x, mx) = 0$$

gilt, aber

$$\lim_{(x,y) \rightarrow (0,0)} f(x, y)$$

nicht existiert.

Zusatzmaterialien.

Die folgenden Bilder wurden mit dem Program Surfer

<http://www.imaginary2008.de/surfer.php>

bzw. dem Online-Plotter

<http://www.benjoffe.com/code/tools/functions3d/>

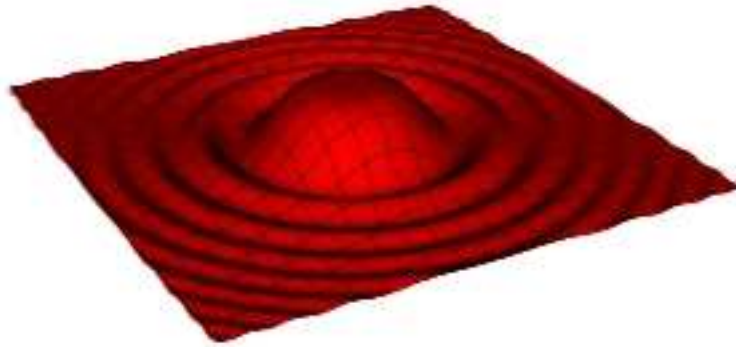
erstellt. Mit diesen Programmen können Sie selbstständig die Graphen solcher

Aufgaben

Funktionen erzeugen und untersuchen.

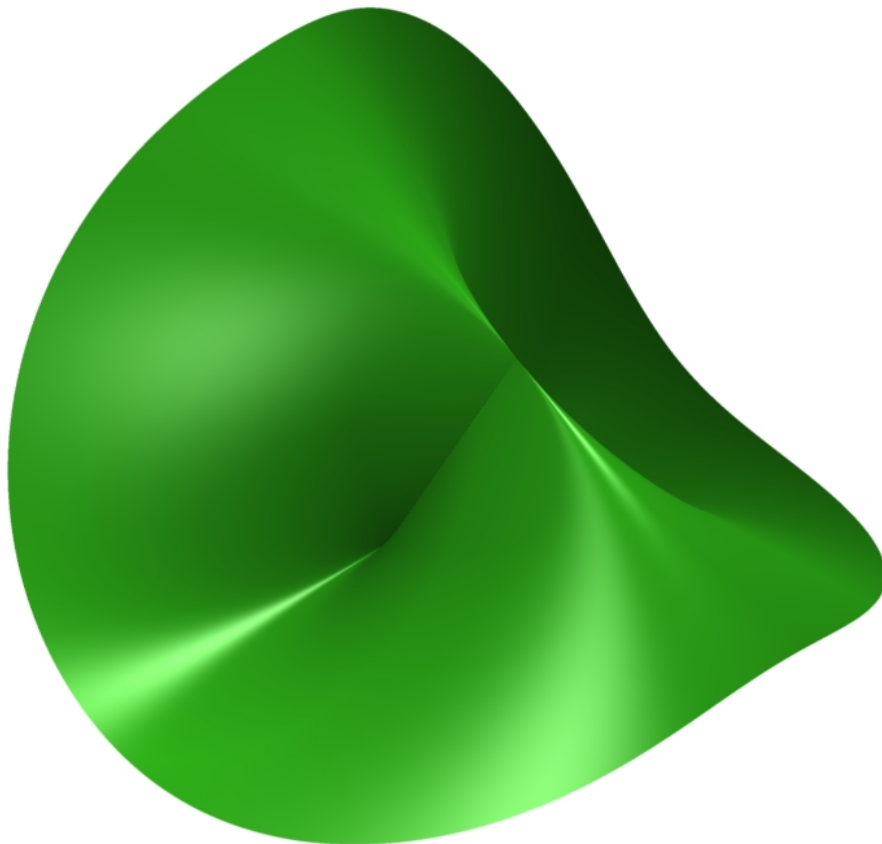
a)

$$\begin{aligned} f: \mathbb{R}^2 \setminus \{(0,0)\} &\mapsto \mathbb{R} \\ (x,y) &\mapsto \frac{\sin(x^2 + y^2)}{x^2 + y^2}. \end{aligned}$$



b)

$$\begin{aligned} f: \mathbb{R}^2 \setminus \{(0,0)\} &\mapsto \mathbb{R} \\ (x,y) &\mapsto \frac{x^2 - y^2}{x^2 + y^2}. \end{aligned}$$



c)

$$f: \mathbb{R}^2 \setminus \{(0,0)\} \mapsto \mathbb{R}$$

$$(x,y) \mapsto \frac{xy^3}{x^2 + 4y^6}.$$



Blatt 5

Aufgabe A.5.1 (Der Differentialoperator).

Es sei $\mathcal{E}^1(I)$ der Vektorraum der stetig differenzierbaren Funktionen auf dem Intervall $I := [a, b]$, $a < b \in \mathbb{R}$. Wir definieren

$$\|\cdot\|_{\mathcal{E}^1}: \mathcal{E}^1(I) \rightarrow \mathbb{R}$$

$$f \mapsto \sup\{|f(x)| + |f'(x)| \mid x \in I\}.$$

a) Beweisen Sie, dass $(\mathcal{E}^1(I), \|\cdot\|_{\mathcal{E}^1(I)})$ ein vollständiger normierter Vektorraum ist.

b) Es sei $(\mathcal{E}^0(I), \|\cdot\|_{\mathcal{E}^0(I)})$ der Raum der stetigen Funktionen auf I mit der Supremumsnorm

$$\|\cdot\|_{\mathcal{E}^0}: \mathcal{E}^0(I) \rightarrow \mathbb{R}$$

$$f \mapsto \sup\{|f(x)| \mid x \in I\}.$$

Zeigen Sie, dass unter diesen Umständen der Differentialoperator

$$\begin{aligned} D: \mathcal{C}^1(I) &\longrightarrow \mathcal{C}^0(I) \\ f &\longmapsto f' \end{aligned}$$

stetig ist.

(Das Beispiel zeigt, dass der Begriff der Stetigkeit bei unendlichdimensionalen Vektorräumen von der Wahl der Norm abhängt. Insbesondere sind die Normen $\|\cdot\|_{\mathcal{C}^0(I)}$ und $\|\cdot\|_{\mathcal{C}^1(I)}$ auf $\mathcal{C}^1(I)$ **nicht äquivalent**.)

Aufgabe A.5.2 (Die Operatornorm).

Es seien $(V, \|\cdot\|_V)$ und $(W, \|\cdot\|_W)$ zwei normierte Vektorräume und $\mathcal{L}_b(V, W)$ der Vektorraum der **stetigen** linearen Abbildungen $A: V \rightarrow W$.

a) Zeigen Sie, dass durch

$$\begin{aligned} \|\cdot\|_{\text{op}}: \mathcal{L}_b(V, W) &\longrightarrow \mathbb{R} \\ A &\longmapsto \sup\{\|A(v)\|_W \mid v \in V: \|v\|_V = 1\} \end{aligned}$$

eine Norm auf $\mathcal{L}_b(V, W)$ gegeben ist.

b) Berechnen Sie die Maximum-Norm, die euklidische Norm und die Operatornorm (bzgl. der euklidischen Norm) der Matrix

$$\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \in \text{Mat}_2(\mathbb{R}) \cong \mathbb{R}^4.$$

c) Weisen Sie nach, dass für $A \in \mathcal{L}_b(V, W)$ die Operatornorm $\|A\|_{\text{op}}$ die kleinste reelle Zahl C ist, so dass

$$\forall v \in V: \|A(v)\|_W \leq C \cdot \|v\|_V.$$

d) Es gelte $V = W$ und $\|\cdot\|_V = \|\cdot\|_W$. Überprüfen Sie, dass für zwei stetige lineare Abbildungen $A, B: V \rightarrow V$ die Abschätzung

$$\|A \circ B\|_{\text{op}} \leq \|A\|_{\text{op}} \cdot \|B\|_{\text{op}}$$

gilt.

e) Wir nehmen jetzt $V = W = \mathbb{R}^n$ an. Unter Verwendung der Standardbasen identifizieren wir $\mathcal{L}_b(V, W)$ mit dem Vektorraum $\text{Mat}_n(\mathbb{R})$ der reellen $(n \times n)$ -Matrizen. Für eine Matrix $A = (a_{ij})_{i,j=1,\dots,n}$ sei

$$\| \|A\| := n \cdot \max\{|a_{ij}| \mid i, j = 1, \dots, n\}.$$

Verifizieren Sie, dass für $\| \cdot \|$ die Eigenschaften aus a) und d) ebenfalls erfüllt sind.

Aufgabe A.5.3 (Das Exponential einer Matrix).

Es sei $V := \text{Mat}_n(\mathbb{R})$ die **Algebra** der $(n \times n)$ -Matrizen, d.h. neben der \mathbb{R} -Vektorraum-Struktur von V betrachten wir auch die multiplikative Struktur, die durch die Matrix-Multiplikation gegeben ist. Für eine Matrix $A \in V$ sei

$$A^i := \underbrace{A \cdots A}_i, \quad i \geq 1, \quad A^0 := \mathbb{E}_n.$$

Wir führen die Folge $(v_k)_{k \in \mathbb{N}}$ über

$$v_k := \sum_{i=0}^k \frac{1}{i!} \cdot A^i, \quad k \in \mathbb{N},$$

ein.

a) Zeigen Sie, dass die Folge $(v_k)_{k \in \mathbb{N}}$ in der Operatornorm konvergiert. Der Grenzwert $\exp(A)$ dieser Folge heißt das *Exponential der Matrix A*.

b) Berechnen Sie die Exponentiale folgender (2×2) -Matrizen:

$$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 1 & a \\ 0 & 1 \end{pmatrix}, a \in \mathbb{R}, \begin{pmatrix} 0 & t \\ -t & 0 \end{pmatrix}, t \in \mathbb{R}.$$

c) Geben Sie (2×2) -Matrizen A und B an, für die

$$\exp(A+B) \neq \exp(A) \cdot \exp(B)$$

gilt.

Blatt 6

Aufgabe A.6.1 (Die Koordinatenachsen).

Geben Sie einen **differenzierbaren** Weg $\gamma: \mathbb{R} \rightarrow \mathbb{R}^2$ an, so dass

$$\text{Spur}(\gamma) = \{ (x, y) \in \mathbb{R}^2 \mid x \cdot y = 0 \}.$$

Aufgabe A.6.2 (Bogenlängen).

a) Es seien

$$\begin{aligned} f: (0, \infty) &\rightarrow \mathbb{R} \\ t &\mapsto \frac{1}{4} \cdot t^2 - \frac{1}{2} \cdot \log(t) \end{aligned}$$

und

$$\begin{aligned} \gamma: [1, e] &\rightarrow \mathbb{R}^2 \\ t &\mapsto (t, f(t)). \end{aligned}$$

Ermitteln Sie die Länge $L(\gamma)$ des Wegs γ .

b) Es seien $a < b \in \mathbb{R}$ und

$$\begin{aligned} \gamma: [a, b] &\rightarrow \mathbb{R}^2 \\ t &\mapsto (t^2, t^3). \end{aligned}$$

Berechnen Sie die Bogenlänge dieses Wegs.

Aufgabe A.6.3 (Eigenschaften der Bogenlänge).

a) Es seien $I := [a, b]$, $a < c < b$, und $\gamma: I \rightarrow \mathbb{R}^n$ ein stetiger, rektifizierbarer Weg. Wir erhalten die Wege $\gamma_1: [a, c] \rightarrow \mathbb{R}^n$, $t \mapsto \gamma(t)$, und $\gamma_2: [c, b] \rightarrow \mathbb{R}^n$, $t \mapsto \gamma(t)$. Beweisen Sie, dass diese beiden Wege ebenfalls rektifizierbar sind und

$$L(\gamma) = L(\gamma_1) + L(\gamma_2)$$

gilt.

b) Zeigen Sie, dass ein stetiger Weg $\gamma: [a, b] \rightarrow \mathbb{R}^n$ genau dann rektifizierbar und von der Länge L ist, wenn es zu jedem vorgegebenen $\varepsilon > 0$ ein $\delta > 0$ gibt, so dass

$$|L_T(\gamma) - L| < \varepsilon$$

für jede Teilung $T = (t_0, \dots, t_N)$ von $[a, b]$ mit

$$|t_j - t_{j-1}| < \delta, \quad j = 1, \dots, N,$$

gilt.

Aufgabe A.6.4 (Die Lemniskate von Bernoulli).

Wir definieren $C \subset \mathbb{R}^2$ als die Menge der Punkte $p = (x, y) \in \mathbb{R}^2$ mit der Eigenschaft, dass das Produkt des Abstands von p zu $(-1, 0)$ mit dem Abstand von p zu $(1, 0)$ gleich eins ist. Geben Sie eine Gleichung für diese Kurve an sowie einen Weg $\gamma: I \rightarrow \mathbb{R}^2$ mit $\text{Spur}(\gamma) = C$. Hat dieser Weg singuläre Punkte?

Aufgabe A.6.5 (Die Kettenlinie).

Wir betrachten den Weg

$$\begin{aligned} \gamma: \mathbb{R} &\longrightarrow \mathbb{R}^2 \\ t &\longmapsto (t, \cosh(t)). \end{aligned}$$

Parametrisieren Sie $\gamma|_{[0, a]}$ für jedes $a > 0$ nach der Bogenlänge.

Blatt 7

Aufgabe A.7.1 (Partielle Ableitungen).

a) Bestimmen Sie den Definitionsbereich der folgenden Funktionen und berechnen Sie dort die ersten partiellen Ableitungen:

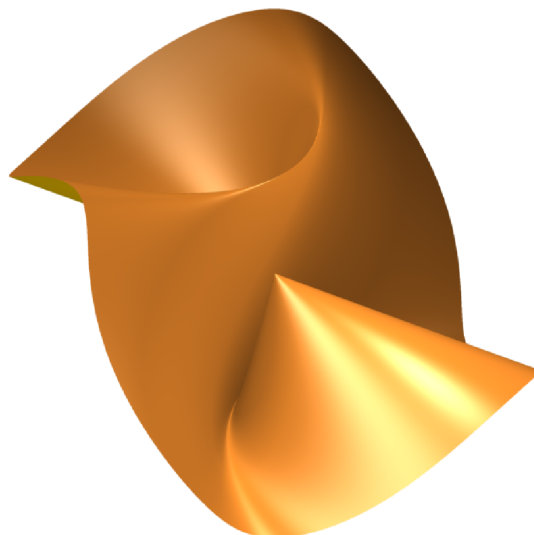
i) $f(x, y) := \log\left(\frac{\sqrt{y}}{x^2}\right)$,

ii) $f(x, y) := \sqrt{2xy - y^2}$,

iii) $f(x, y) := \log(\cos(4x^3 - 2y^2 + 1))$.

b) Es sei

$$\begin{aligned} f: \mathbb{R}^2 &\longrightarrow \mathbb{R} \\ (x, y) &\longmapsto \begin{cases} \frac{xy^2}{x^2 + y^4}, & \text{falls } (x, y) \neq (0, 0) \\ 0, & \text{falls } (x, y) = (0, 0) \end{cases}. \end{aligned}$$



Zeigen Sie, dass für jeden Vektor $v \in \mathbb{R}^2 \setminus \{0\}$ die Richtungsableitung $D_v f(0, 0)$ existiert. Ist f in $(0, 0)$ stetig?

c) Beweisen Sie, dass für die Funktion

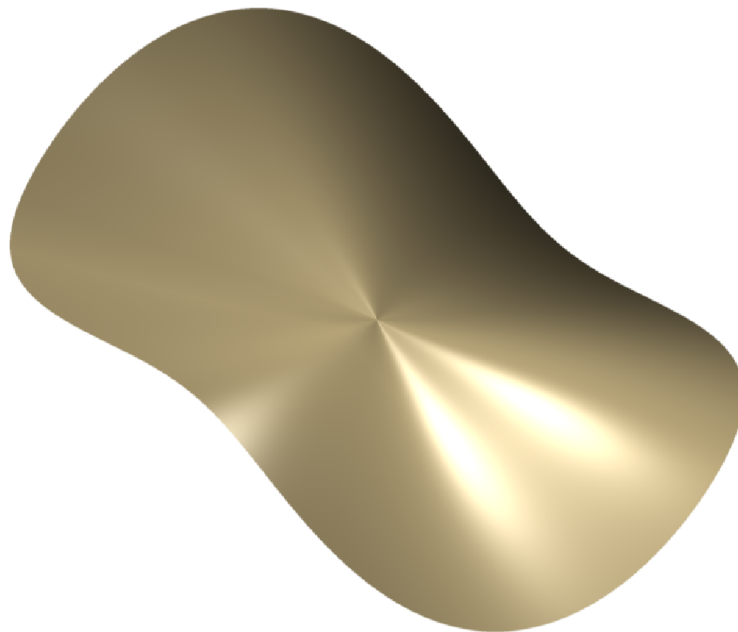
$$f: \mathbb{R}^2 \longrightarrow \mathbb{R}$$

$$(x, y) \longmapsto \begin{cases} \frac{y^3}{x^2 + y^2}, & \text{falls } (x, y) \neq (0, 0) \\ 0, & \text{falls } (x, y) = (0, 0) \end{cases}$$

in $(0, 0)$ alle Richtungsableitungen existieren, im Allgemeinen aber

$$D_v f(0, 0) \neq \langle \text{Grad} f(0, 0), v \rangle$$

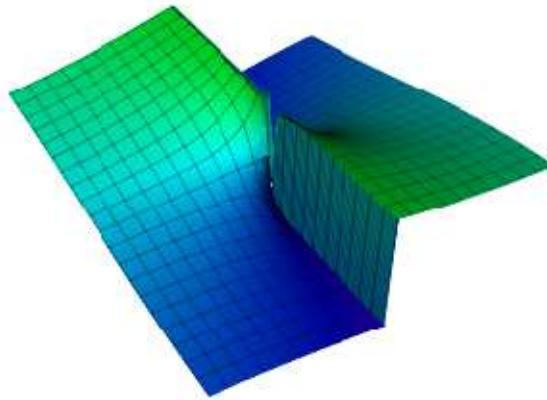
gilt. Was folgt für die Funktion f ? Ist f in $(0, 0)$ stetig?



d) Berechnen Sie auf ganz B die ersten und zweiten partiellen Ableitungen der Funktion

$$f: B := \{ (x, y) \in \mathbb{R}^2 \mid x \neq 0 \} \longrightarrow \mathbb{R}$$

$$(x, y) \longmapsto \arctan\left(\frac{y}{x}\right).$$



Aufgabe A.7.2 (Niveaulinien und Gradienten).
Bestimmen Sie zu $a, b, c \in \mathbb{R}$ die Niveaulinien der Funktion

$$f(x, y) := ax^2 + 2bxy + cy^2$$

auf \mathbb{R}^2 . Berechnen Sie den Gradienten in jedem Punkt. Versuchen Sie, Ihre Ergebnisse graphisch zu veranschaulichen.

Aufgabe A.7.3 (Höhere partielle Ableitungen).

a) **Definition.** Es seien $B \subset \mathbb{R}^n$ eine offene Menge, $f: B \rightarrow \mathbb{R}$ eine Funktion und $k \geq 1$. Man sagt, f ist k -mal stetig (partiell) differenzierbar, wenn gilt:

- f ist $(k - 1)$ -mal stetig (partiell) differenzierbar,
- die $(k - 1)$ -ten partiellen Ableitungen

$$f_{x_{i_1} \dots x_{i_{k-1}}} := \frac{\partial^{k-1}}{\partial x_{i_1} \dots \partial x_{i_{k-1}}} f: B \rightarrow \mathbb{R}$$

$$(x_1, \dots, x_n) \mapsto \frac{\partial^{k-1}}{\partial x_{i_1} \dots \partial x_{i_{k-1}}} f(x_1, \dots, x_n)$$

sind für alle Tupel (i_1, \dots, i_{k-1}) mit $i_j \in \{1, \dots, n\}, j = 1, \dots, k - 1$, stetig partiell differenzierbar. Man setzt

$$f_{x_{i_1} \dots x_{i_k}} := \frac{\partial^k}{\partial x_{i_1} \dots \partial x_{i_k}} f := \frac{\partial}{\partial x_{i_1}} \left(\frac{\partial^{k-1}}{\partial x_{i_2} \dots \partial x_{i_k}} f \right)$$

für ein Tupel (i_1, \dots, i_k) mit $i_j \in \{1, \dots, n\}, j = 1, \dots, k$.

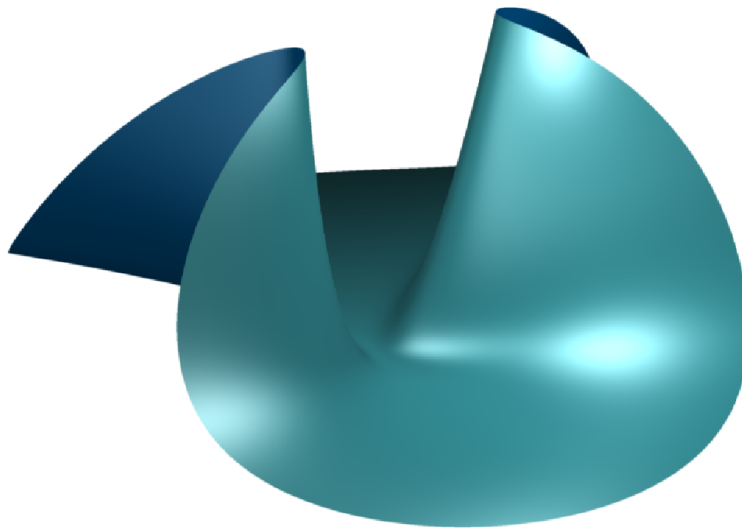
Zeigen Sie, dass für eine k -mal stetig differenzierbare Funktion $f: B \rightarrow \mathbb{R}, B \subset \mathbb{R}^n$ offen, ein Tupel (i_1, \dots, i_k) mit $i_j \in \{1, \dots, n\}, j = 1, \dots, k$, und eine Permutation $\sigma: \{1, \dots, k\} \rightarrow \{1, \dots, k\}$ gilt:

$$\frac{\partial^k}{\partial x_{i_1} \dots \partial x_{i_k}} f = \frac{\partial^k}{\partial x_{i_{\sigma(1)}} \dots \partial x_{i_{\sigma(k)}}} f.$$

b) Bestimmen Sie $f_{xyx} = f_{xyy} = f_{yxx}$ für die Funktion

$$f(x, y) := \frac{x^2}{1 + y^2}$$

auf \mathbb{R}^2 unter Benutzung der verschiedenen Differentiationsreihenfolgen.



Aufgabe A.7.4 (Differentialgleichungen).
 Sei $a \in \mathbb{R}$. Überprüfen Sie, dass für die Funktion

$$\begin{aligned} f: \mathbb{R}^2 &\longrightarrow \mathbb{R} \\ (x, y) &\longmapsto \exp(-\pi^2 \cdot a^2 \cdot y) \cdot \sin(\pi \cdot x) \end{aligned}$$

die *Differentialgleichung*

$$a^2 \cdot f_{xx} - f_y = 0$$

erfüllt ist.

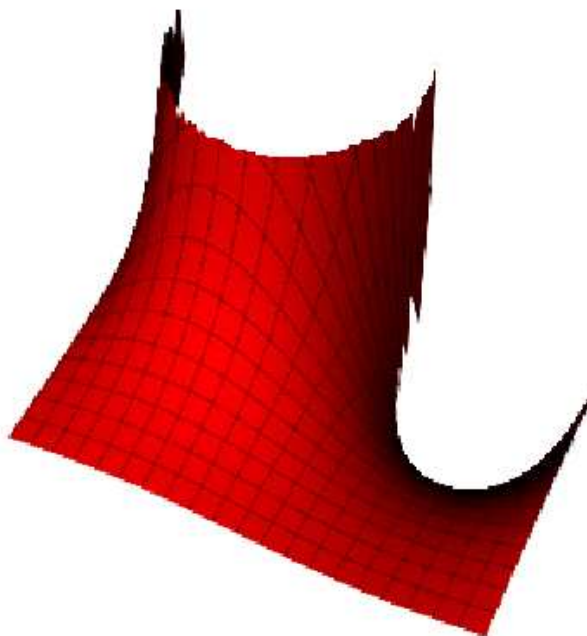


Abbildung A.2: Der Funktionsgraph für $a = 1$

Blatt 8

Aufgabe A.8.1 (Die totale Ableitung).

Es seien $B \subset \mathbb{R}^m$ eine offene Teilmenge, $f, g: \mathbb{R}^m \rightarrow \mathbb{R}^n$ zwei Abbildungen, $\lambda \in \mathbb{R}$ eine reelle Zahl und $a \in B$.

a) Beweisen Sie die folgenden Aussagen:

1. Ist f in a differenzierbar, so ist auch $\lambda \cdot f$ in a differenzierbar mit

$$D(\lambda \cdot f)(a) = \lambda \cdot (Df(a)).$$

2. Wenn f und g in a differenzierbar sind, dann ist auch die Summe $f + g$ in a differenzierbar mit totaler Ableitung

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

b) Es gelte $n = 1$, und f und g seien in a differenzierbar. Formulieren und beweisen Sie die Produktregel für $\text{Grad}(f \cdot g)$.

Aufgabe A.8.2 (Homogene Funktionen).

Es sei $f: \mathbb{R}^n \rightarrow \mathbb{R}$ eine stetig differenzierbare Funktion. Es gebe eine natürliche Zahl p , so dass

$$\forall t \in \mathbb{R}, x \in \mathbb{R}^n: f(t \cdot x) = t^p \cdot f(x).$$

Beweisen Sie:

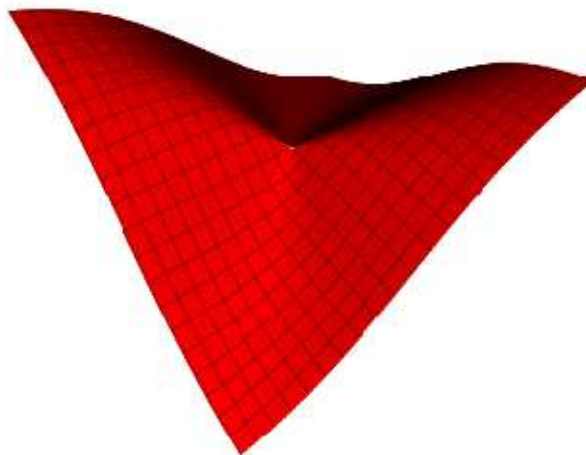
$$\forall x = (x_1, \dots, x_n) \in \mathbb{R}^n: \sum_{i=1}^n x_i \cdot \frac{\partial f}{\partial x_i}(x) = p \cdot f(x).$$

Aufgabe A.8.3 (Stetige, partiell differenzierbare Funktionen).

a) Es seien $B \subset \mathbb{R}^n$ eine offene Menge und $f: B \rightarrow \mathbb{R}$ eine partiell differenzierbare Funktion. Alle partiellen Ableitungen seien auf B beschränkt. Beweisen Sie, dass f stetig ist.

b) Wir definieren

$$f: \mathbb{R}^2 \rightarrow \mathbb{R} \\ (x, y) \mapsto \begin{cases} \frac{xy}{\sqrt{x^2 + y^2}}, & \text{falls } (x, y) \neq (0, 0) \\ 0, & \text{falls } (x, y) = (0, 0) \end{cases}.$$



1. Zeigen Sie, dass f partiell differenzierbar ist und die partiellen Ableitungen auf \mathbb{R}^2 beschränkt sind. (Damit ist f nach Teil a) eine stetige Funktion.)
2. In welchen Punkten ist f differenzierbar?

Aufgabe A.8.4 (Tangentialebenen).

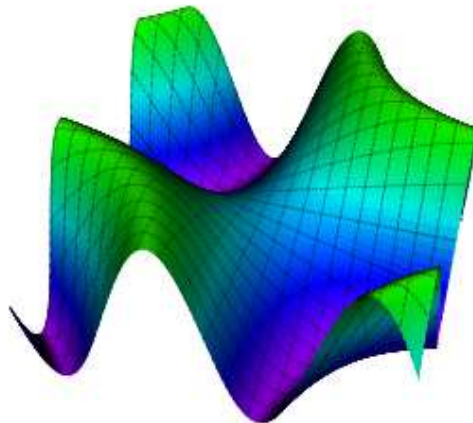
a) Es sei

$$\begin{aligned} f: \mathbb{R}^2 &\longrightarrow \mathbb{R} \\ (x, y) &\longmapsto 5 - 2x^2 - y^2. \end{aligned}$$

Geben Sie die Gleichung für die Tangentialebene von Γ_f im Punkt $(1, 1, 2)$ an.
b) Wir setzen

$$\begin{aligned} f: \mathbb{R}^2 &\longrightarrow \mathbb{R} \\ (x, y) &\longmapsto \sin\left(\frac{\pi}{2} \cdot x \cdot y\right). \end{aligned}$$

Bestimmen Sie die Tangentialebene von Γ_f im Punkt $(3, 5, -1)$.



Blatt 9 (=Analysis III, Blatt 1)

Aufgabe A.9.1 (Zu den Beispielen aus der Vorlesung).

In der Vorlesung (s. Seite 121ff) wurden die folgenden Differentialgleichungen betrachtet:

- $x'(t) = x(t)$,
- $x'(t) = x^2(t)$,

- $x'(t) = \kappa \cdot (1 - x^2(t))$.

Bestimmen Sie jeweils explizit (möglichst große) Intervalle I', J , für die

$$\phi(I') \subset \psi(J)$$

gilt. Führen Sie ggfs. die notwendigen Fallunterscheidungen aus.

Aufgabe A.9.2 (Differentialgleichungen mit getrennten Variablen).

Lösen Sie die folgenden Differentialgleichungen durch Trennung der Variablen:

a) $x'(t) = \exp(t - x(t))$ zu der Anfangsbedingung $x(0) = 1$.

b) $x'(t) = (x(t) + 1) \cdot \sin(t)$ zu der Anfangsbedingung $x(\pi/2) = 4$.

Aufgabe A.9.3 (Lösung durch Substitution).

Geben Sie die allgemeine Lösung der Differentialgleichung

$$x'(t) = 1 + \frac{x^2(t)}{t^2 + t \cdot x(t)}, \quad t > 0,$$

an.

Aufgabe A.9.4 (Lineare Differentialgleichungen).

Bestimmen Sie die Lösungen der folgenden linearen Differentialgleichungen zu den gegebenen Anfangswerten:

a) $x'(t) + 5 \cdot x(t) = \cos(t) \cdot \exp(-5 \cdot t)$, $x(\pi/2) = 0$.

b) $m \cdot v'(t) + k \cdot v(t) = m \cdot g$, $m, k, g \in \mathbb{R}_{>0}$, $v(0) = 0$.¹²

Aufgabe A.9.5 (Eulersche Differentialgleichungen).

Ermitteln Sie die allgemeine Lösung der Eulerschen Differentialgleichung

$$x'''(t) - \frac{2}{t} \cdot x''(t) + \frac{5}{t^2} \cdot x'(t) - \frac{5}{t^3} \cdot x(t) = 0, \quad t > 0.$$

Blatt 10 (=Analysis III, Blatt 2)

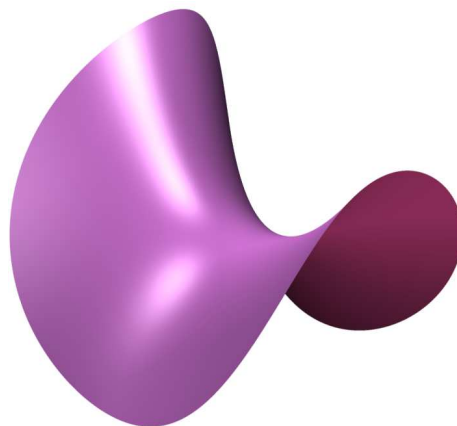
Aufgabe A.10.1 (Semidefinite Matrizen).

Formulieren und beweisen Sie ein notwendiges und hinreichendes Kriterium dafür, dass eine reelle symmetrische (2×2) -Matrix A positiv bzw. negativ **se-midefinit** ist. Was fällt auf?

Aufgabe A.10.2 (Taylorentwicklung).

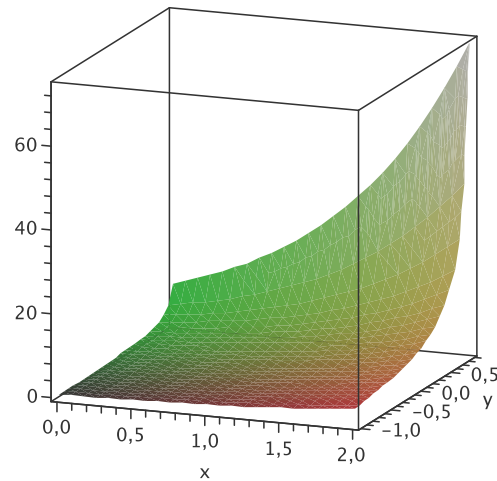
Berechnen Sie die Taylorentwicklung bis einschl. Ordnung 2 der folgenden Funktionen im angegebenen Entwicklungspunkt a :

a) $f: \mathbb{R}_{>0} \times \mathbb{R}_{>0} \rightarrow \mathbb{R}$, $(x, y) \mapsto (x - y)/(x + y)$, $a = (1, 1)$.



¹²Sinkgeschwindigkeit eines Körpers in einer zähen Flüssigkeit

b) $f: \mathbb{R} \times (\mathbb{R} \setminus \{1\}) \rightarrow \mathbb{R}, (x, y) \mapsto \exp(x)/(1 - y), a = (0, 0)$.



Aufgabe A.10.3 (Das Dilemma des Bergsteigers).
Wir betrachten die Funktion

$$\begin{aligned} f: \mathbb{R}^2 &\longrightarrow \mathbb{R} \\ (x, y) &\longmapsto (2x^2 - y) \cdot (y - x^2). \end{aligned}$$

- Zeigen Sie, dass $(0, 0)$ ein kritischer Punkt für f ist, und berechnen Sie $H_f(0, 0)$.
- Überprüfen Sie, dass die Einschränkung von f auf jede Gerade durch $(0, 0)$ ein lokales Maximum in $(0, 0)$ hat.
- Weisen Sie nach, dass f in $(0, 0)$ **kein** lokales Extremum besitzt.

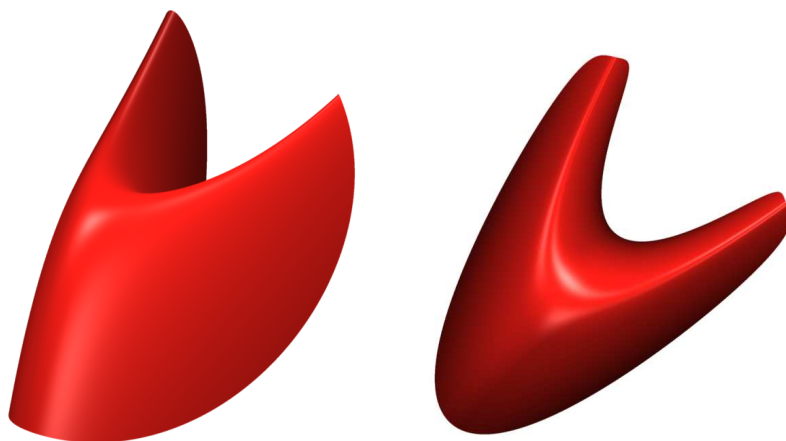
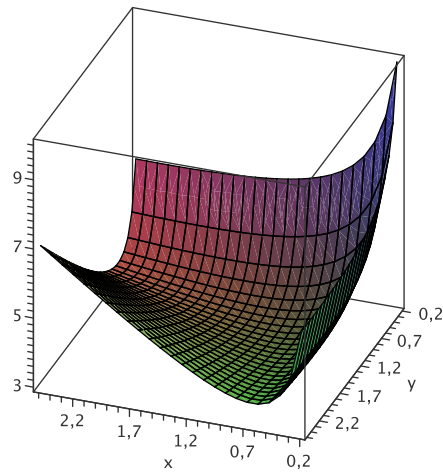


Abbildung A.3: Zwei Ansichten des Funktionsgraphen

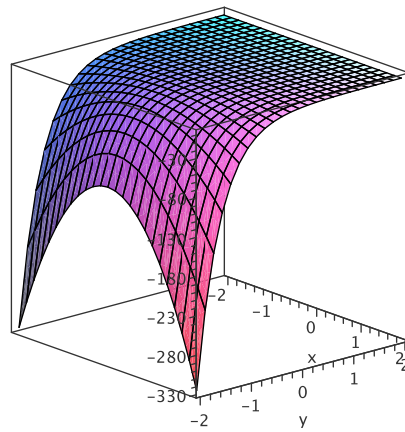
Aufgabe A.10.4 (Lokale Extrema und Sattelpunkte).

a) Untersuchen Sie folgende Funktionen auf lokale Extremwerte:

- $f: \mathbb{R}_{>0} \times \mathbb{R}_{>0} \rightarrow \mathbb{R}, (x, y) \mapsto xy + 1/x + 1/y,$

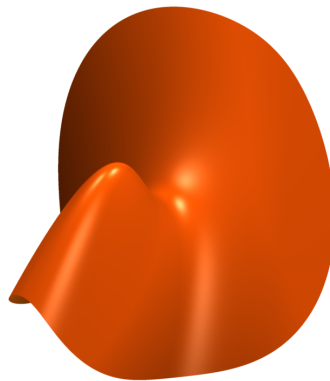


- $f: \mathbb{R}^2 \rightarrow \mathbb{R}, (x, y) \mapsto (y - x^2) \cdot \exp(-2y).$



b) Untersuchen Sie die folgende Funktionen auf lokale Extrema und Sattelpunkte:

- $f: \mathbb{R}^2 \rightarrow \mathbb{R}, (x, y) \mapsto 3xy - x^3 - y^3,$



- $f: \mathbb{R}^3 \rightarrow \mathbb{R}, (x, y, z) \mapsto \cos(x) + \cos(y) + \cos(z).$

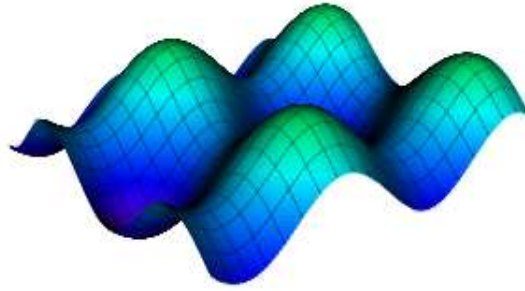


Abbildung A.4: Der Graph der Funktion $f: \mathbb{R}^2 \rightarrow \mathbb{R}^2, (x, y) \mapsto \cos(x) + \cos(y)$

Blatt 11 (=Analysis III, Blatt 3)

Aufgabe A.11.1 (Zum Mittelwertsatz).

Es sei

$$\begin{aligned} \gamma: \mathbb{R} &\rightarrow \mathbb{R}^2 \\ t &\mapsto (\sin(2\pi \cdot t), \cos(2\pi \cdot t)). \end{aligned}$$

Beweisen Sie, dass es kein $\xi \in [0, 1]$ mit

$$\gamma(1) - \gamma(0) = \gamma'(\xi)$$

gibt.

Hintergrund. Man vergleiche die Mittelwertsätze 7.5.5 und 8.3.2: Wir betrachten eine offene Teilmenge $B \subset \mathbb{R}^m$, eine differenzierbare Abbildung $f: B \rightarrow \mathbb{R}^n$ und Punkte $x, y \in B$, so dass

$$L := \{ x + \lambda \cdot (y - x) \mid \lambda \in [0, 1] \} \subset B.$$

Falls $n > 1$, gibt es also im Allgemeinen **keine** Zahl $\xi \in [0, 1]$, so dass

$$f(y) - f(x) = Df(x + \xi \cdot (y - x))(y - x).$$

Aufgabe A.11.2 (Die Jacobi-Matrix).

Es sei

$$\begin{aligned} f: \mathbb{R}^2 &\rightarrow \mathbb{R}^2 \\ (x, y) &\mapsto (x^2 - y^2, 2xy). \end{aligned}$$

a) Bestimmen Sie $f^{-1}(p)$ für jeden Punkt p aus \mathbb{R}^2 . Wieviele Elemente enthält $f^{-1}(p)$?

b) Berechnen Sie die Jacobi-Matrix $J_f(p)$ für alle Punkte $p \in \mathbb{R}^2$. In welchen Punkten ist sie invertierbar?

Aufgabe A.11.3 (Lokale Umkehrbarkeit).

Es sei

$$\begin{aligned} f: \mathbb{R}^2 &\rightarrow \mathbb{R}^2 \\ (x, y) &\mapsto (\exp(x) \cdot \cos(y), \exp(x) \cdot \sin(y)). \end{aligned}$$

a) Zeigen Sie, dass f überall lokal umkehrbar ist. Ist f surjektiv/injektiv?

b) Geben Sie ein Gebiet $G \subset \mathbb{R}^2$ an, so dass $f|_G: G \rightarrow f(G)$ ein Diffeomorphismus ist.

Aufgabe A.11.4 (Die Jacobi-Matrix der Umkehrfunktion).

Es sei

$$\begin{aligned} f: \mathbb{R}_{>0} \times \mathbb{R}_{>0} &\longrightarrow \mathbb{R}^2 \\ (x_1, x_2) &\longmapsto (x_1^2 + x_2^2, x_1^2 - x_2^2). \end{aligned}$$

Zeigen Sie, dass es zu jedem Punkt $(x_1, x_2) \in \mathbb{R}_{>0} \times \mathbb{R}_{>0}$ eine offene Umgebung $U \subset \mathbb{R}_{>0} \times \mathbb{R}_{>0}$ gibt, so dass $f|_U: U \rightarrow f(U)$ ein Diffeomorphismus ist. Geben Sie auch die Jacobi-Matrix

$$J_{(f|_U)^{-1}}(y_1, y_2), \quad (y_1, y_2) := f(x_1, x_2),$$

an.

Aufgabe A.11.5 (Implizite Differentiation).

a) Zeigen Sie, dass man die Gleichung

$$f(x, y) = (x + y^2) \cdot \exp(y - 1) - 1 = 0$$

in einer Umgebung des Punkts $(0, 1)$ nach y auflösen kann, d.h. es offene Intervalle $0 \in I \subset \mathbb{R}$, $1 \in J \subset \mathbb{R}$ und eine Funktion $\varphi: I \rightarrow J$ mit $\varphi(0) = 1$ und

$$\{(x, y) \in \mathbb{R}^2 \mid f(x, y) = 0\} \cap (I \times J) = \{(x, \varphi(x)) \mid x \in I\}$$

gibt.

b) Bestimmen Sie $\varphi'(0)$ und $\varphi''(0)$.

Blatt 12 (=Analysis III, Blatt 4)

Aufgabe A.12.1 (Implizite Funktionen).

Gegeben sei das Gleichungssystem

$$\begin{aligned} 2x_1 + x_2 + x_3 + x_5 - 1 &= 0 \\ x_1x_2^3 + x_1x_3 + x_2^2x_4^2 - x_4x_5 &= 0 \\ x_2x_3x_5 + x_1x_3^2 + x_4x_5^2 &= 0. \end{aligned}$$

a) Es sei $L \subset \mathbb{R}^5$ die Lösungsmenge dieses Gleichungssystems. Überprüfen Sie, dass $a := (0, 1, -1, 1, 1) \in L$ und dass es offene Mengen $(0, 1) \in U \subset \mathbb{R}^2$, $(-1, 1, 1) \in V \subset \mathbb{R}^3$ und eine stetig differenzierbare Abbildung $F: U \rightarrow V$ gibt, so dass

$$L \cap (U \times V) = \{(x_1, x_2, F(x_1, x_2)) \mid (x_1, x_2) \in U\}.$$

b) Bestimmen Sie die Jacobi-Matrix $J_F(0, 1)$.

Aufgabe A.12.2 (Implizite Funktionen und implizite Differentiation).

Man betrachte die Gleichungen

$$\begin{aligned} \exp(y_1) + x_1y_2 - \sin(x_2) + x_3^2 - 2 &= 0 \\ (y_1 - x_1)^2 + (y_2 - x_2)^2 - \exp(x_3) - 1 &= 0 \end{aligned}$$

und den Punkt $(1, 0, 0, 0, 1)$.

a) Weisen Sie die Existenz von offenen Mengen $(1, 0, 0) \in U \subset \mathbb{R}^3$, $(0, 1) \in V \subset \mathbb{R}^2$ und einer stetig differenzierbaren Abbildung $F: U \rightarrow V$ nach, so dass ein Punkt $(x_1, x_2, x_3, y_1, y_2) \in U \times V$ genau dann das obige Gleichungssystem löst,

wenn $(y_1, y_2) = F(x_1, x_2, x_3)$.

b) Es sei $F = (F_1, F_2)$ mit $F_i: U \rightarrow \mathbb{R}, i = 1, 2$. Berechnen Sie die partiellen Ableitungen

$$\frac{\partial F_i}{\partial x_j}(1, 0, 0), \quad i = 1, 2, j = 1, 2, 3.$$

Aufgabe A.12.3 (Ein glattes parametrisiertes Flächenstück).

Wir fixieren reelle Zahlen $0 < a < b$ und setzen

$$\begin{aligned} f: (0, 2\pi) \times (0, 2\pi) &\rightarrow \mathbb{R}^3 \\ (\varphi, \psi) &\mapsto ((b - a \cos(\psi)) \cos(\varphi), (b - a \cos(\psi)) \sin(\varphi), a \sin(\psi)). \end{aligned}$$

a) Zeigen Sie, dass f ein glattes parametrisiertes Flächenstück ist.

b) Skizzieren Sie das Bild von f . (**Hinweis.** Schauen Sie sich die Polar- und Kugelkoordinaten noch einmal an.)

Aufgabe A.12.4 (Tangentialräume).

a) Es sei

$$M := \{ (x, y, z) \in \mathbb{R}^3 \mid x^2 + y^2 - z^2 - 1 = 0 \}.$$

Bestimmen Sie für jeden Punkt $p = (x, y, 0) \in M$ mit $x, y \in \mathbb{R}$ den Tangentialraum von M an p . Überprüfen Sie auch, dass dieser Tangentialraum die z -Achse enthält.

b) Gegeben sei

$$\begin{aligned} f: \mathbb{R}^2 &\rightarrow \mathbb{R} \\ (x, y) &\mapsto 3x^2 - 2y^2 + \cos(\pi \cdot (x - y)). \end{aligned}$$

Bestimmen Sie den Tangentialraum an den Graphen $\Gamma_f \subset \mathbb{R}^3$ von f im Punkt $(1, 2, f(1, 2))$.

c) Es sei

$$M := \{ (x, y, z) \in \mathbb{R}^3 \mid xyz - 1 = 0 \}.$$

Geben Sie den Tangentialraum an M im Punkt $(1, 2, 1/2)$ an.

Blatt 13 (=Analysis III, Blatt 5)

Aufgabe A.13.1 (Untermannigfaltigkeiten).

a) Es seien $M_1 \subset \mathbb{R}^{n_1}$ bzw. $M_2 \subset \mathbb{R}^{n_2}$ Untermannigfaltigkeiten der Dimension k_1 bzw. k_2 . Zeigen Sie, dass $M_1 \times M_2$ eine Untermannigfaltigkeit der Dimension $k_1 + k_2$ von $\mathbb{R}^{n_1+n_2}$ ist.

b) Bei welchen der folgenden Teilmengen handelt es sich um Untermannigfaltigkeiten? (Begründung nicht vergessen!)

- $M := \{ (x, y) \in \mathbb{R}^2 \mid x^2 y + xy^2 = x^4 + y^4 \},$
- $M := \{ (x_1, x_2, x_3) \in \mathbb{R}^3 \mid x_1 - x_2 = 0, 3x_1^3 - 5x_2^3 = 0, x_1^2 + x_2^4 = 0 \},$
- $M := \{ (x_1, x_2, x_3) \in \mathbb{R}^3 \mid x_2^2 - x_1 x_3 + x_1^3 = 0 \},$
- $SL_n(\mathbb{R}) := \{ A \in \text{Mat}(n, \mathbb{R}) \mid \text{Det}(A) = 1 \}.$

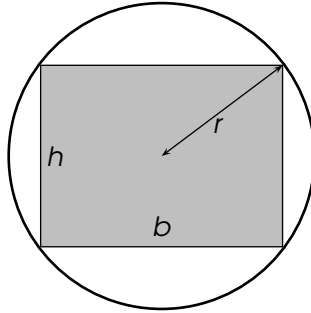
Aufgabe A.13.2 (Extrema unter Nebenbedingungen I).

a) Bestimmen Sie die lokalen Extrema der Funktion $f: \mathbb{R}^2 \rightarrow \mathbb{R}, (x, y) \mapsto -x^2 + 8x - y^2 + 9$ unter der Nebenbedingung $g: \mathbb{R}^2 \rightarrow \mathbb{R}, (x, y) \mapsto x^2 + y^2 - 1$ durch **Einsetzen der Nebenbedingung**.

b) Aus einem langen Baumstamm mit kreisrundem Querschnitt vom Radius r soll ein Balken der Breite b und der Höhe h herausgesägt werden. Dabei sind b und h so zu wählen, dass das **Widerstandsmoment**

$$W := \frac{1}{6} \cdot b \cdot h^2$$

möglichst groß wird. Formulieren Sie die entsprechende Extremwertaufgabe und lösen Sie sie durch **Einsetzen der Nebenbedingung**.



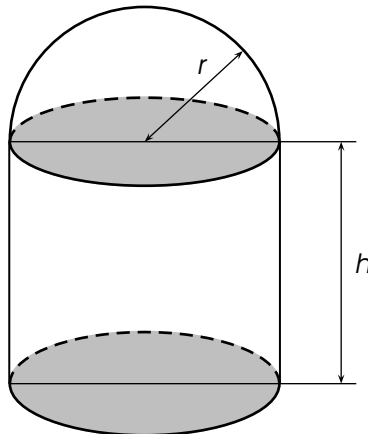
c) Lösen Sie die Extremwertaufgabe aus b) mit Hilfe des **Verfahrens der Lagrangeschen Multiplikatoren**.

Aufgabe A.13.3 (Extrema unter Nebenbedingungen II).

Es seien $f: \mathbb{R}^3 \rightarrow \mathbb{R}, (x, y, z) \mapsto xy + 2xz + 2yz$, und $g: \mathbb{R}^3 \rightarrow \mathbb{R}, (x, y, z) \mapsto xyz - 108$. Bestimmen Sie die lokalen Extrema der Funktion f unter der Nebenbedingung $g(x, y, z) = 0$.

Aufgabe A.13.4 (Extrema unter Nebenbedingungen III).

Auf einen Zylinder mit Radius r und Höhe h werde eine Halbkugel vom Radius r aufgesetzt. Wie sind r und h zu wählen, damit das entstehende Objekt das Volumen $V = 5000\text{cm}^3$ und möglichst kleine Oberfläche hat? Formulieren Sie die entsprechende Extremwertaufgabe und lösen Sie sie mit Hilfe des **Verfahrens der Lagrangeschen Multiplikatoren**.



Literaturhinweise

- (1) F. Apéry, *Models of the real projective plane. Computer graphics of Steiner and Boy surfaces*, Braunschweig/Wiesbaden: Friedr. Vieweg & Sohn, 1987, xi+156 S. ISBN 3-528-08955-5
- (2) T. Arens, F. Hettlich, Ch. Karpfinger, U. Kockelkorn, K. Lichtenegger, H. Stachel, *Mathematik*, Heidelberg: Spektrum Akademischer Verlag, 2008, xiv+1496 S. ISBN 978-3-8274-1758-9/hbk
- (3) M. Barner, F. Flohr, *Analysis II., 2.,* bearb. Aufl., de Gruyter Lehrbuch, Berlin etc.: Walter de Gruyter, 1989, 449 S.
- (4) Ch. Blatter, *Mehrdimensionale Differentialrechnung*, Kapitel 12 auf <http://www.math.ethz.ch/~blatter/>.
- (5) Th. Bröcker, K. Jänich, *Einführung in die Differentialtopologie*, Heidelberger Taschenbücher, Band 143, Berlin–Heidelberg–New York: Springer-Verlag, 1973, vii+168 S.
- (6) O. Forster, *Analysis 2. Differentialrechnung im \mathbb{R}^n , gewöhnliche Differentialgleichungen*, 8. überarb. Aufl., Vieweg Studium: Grundkurs Mathematik, Wiesbaden, Vieweg, 2008, viii+221 S. ISBN 978-3-8348-0575-1/pbk
- (7) K. Fritzsche, *Grundkurs Analysis 2 - Differentiation und Integration in mehreren Veränderlichen. Für Bachelor und Diplom*. Heidelberg: Elsevier/Spektrum Akademischer Verlag, 2006, viii+366 S. ISBN 978-3-8274-1687-2/pbk
- (8) H. Heuser, *Lehrbuch der Analysis. Teil 2., 14.* überarb. Aufl., Studium, Vieweg+Teubner, Wiesbaden, 2008, 737 S. ISBN 978-3-8351-0208-8/pbk
- (9) W. Kabballo, *Einführung in die Analysis. II*, Spektrum Hochschultaschenbuch, Spektrum Akademischer Verlag, Heidelberg, 1997, x+266 S. ISBN 3-8274-0198-4/pbk
- (10) W. Klingenberg, *Eine Vorlesung über Differentialgeometrie*, Heidelberger Taschenbücher, Band 107, Berlin–Heidelberg–New York: Springer-Verlag, 1973, x+135 S.
- (11) F. Lorenz, *Lineare Algebra II., 3.,* überarb. Aufl., B.I. Wissenschaftsverlag, Mannheim, 1992, viii+195 S. ISBN 3-411-15233-8/pbk
- (12) A. Schmitt, *Analysis I*, Vorlesungsskript, <http://userpage.fu-berlin.de/~aschmitt/>
- (13) V. Schulze, *Lineare Algebra*, Vorlesungsskript, <http://page.mi.fu-berlin.de/klarner/index>

Stichwortverzeichnis

Stichwortverzeichnis

- Abbildung
 - konstante —, 75
 - lineare —, 75
 - offene —, 139
 - stetige —, 23
- abgeschlossen, 5, 7
 - er Quader, 14
 - er Unterraum, 37
- Ableitung, 48
 - partielle —, 59, 192
 - Richtungs—, 57
- Abschluss, 10
- Abstand, 1, 34, 183
 - sfunktion, 2
- abzählbare Basis der Topologie, 177
- additiv, 185
- Anfangs
 - bedingung, 114
 - punkt, 43
- Antipodenabbildung, 178
- äquivalent, 20, 46, 179, 182
- Atlas, 177
- Ball, offener —, 4
- Banach
 - Raum, 12
 - scher Fixpunktsatz, 106
- Basis der Topologie, 177
- Bernoulli, 190
 - Lemniskate von —, 190
- beschränkt, 3, 15, 37
- Bilinearform, symmetrische —, 87
- Bonbon, 119
- Borel, 17
- Cauchy
 - Folge, 12
 - Schwarz-Ungleichung, 76
- charakteristisches Polynom, 136
- Cramersche Regel, 148, 149
- definit
 - negativ —, 88
 - positiv —, 88
- Diffeomorphismus, 141
 - lokaler —, 149
- Differential
 - gleichung, 113, 129, 193
 - erster Ordnung, 113
 - mit getrennten Variablen, 119
 - operator, 40, 84, 187
- Differentiation, implizite —, 156
- differenzierbar, 40, 48, 71, 172
 - e Mannigfaltigkeit, 179
 - e Struktur, 179
 - k-mal stetig —, 148
 - partiell —, 59, 192
 - stetig —, 40, 48, 77, 172
- Differenzierbarkeit, höhere —, 148
- Dilemma des Bergsteigers, 197
- Dimension, 162
- diskrete Topologie, 8
- Durchmesser, 15
- Eigenwert, 101
- Einheits
 - kugel, 184
 - sphäre, 164, 177
 - Inversion an der —, 26
- Ellipse, 51
- Endpunkt, 43
- Energie, potentielle —, 103
- Euklid, 1, 27
 - Kathetensatz des —, 27
- euklidische
 - Metrik, 3
 - Norm, 3
- Euler, 136
 - sche Differentialgleichung, 136
- Exponential einer Matrix, 189
- Extremum
 - isoliertes lokales —, 89
 - lokales —, 81
- Exzentrizität, numerische —, 51
- Fixpunktgleichung, 146
- folgenkompakt, 184
- Funktion
 - monomiale —, 26

- polynomiale —, 26
- trigonometrische —, 50
- Gebiet, 140
- Gerono, 54, 161
- Geschwindigkeit
 - sfeld, 114
 - svektor, 48, 114
- glattes parametrisiertes Flächenstück, 162
- gleichmäßig
 - e Konvergenz, 37
 - stetig, 35
- globales
 - Maximum unter einer Nebenbedingung, 94
 - Minimum unter einer Nebenbedingung, 94
- Gradient, 62
- Graph, 44, 74
- Grenzwert, 11, 30
- Gruppe, orthogonale —, 169
- Halbachse, 51
- Hausdorff, 8
- hausdorffsch, 8
- Heine, 17
- Hesse-Matrix, 87
- höhere Differenzierbarkeit, 148
- homogene lineare Differentialgleichung, 129
 - mit konstanten Koeffizienten, 136
- Homöomorphismus, 139
- implizit
 - e Differentiation, 156
 - e Funktion, 159
- indefinit, 88
- induzierte Metrik, 2
- inhomogene lineare Differentialgleichung, 129
- Integraloperator, 41
- Inversion an der Einheitskugel, 26
- isoliertes lokales
 - Extremum, 89
 - Maximum, 89
 - Minimum, 89
- Jacobi, 74
 - Matrix, 74
- Karte, 177
 - nwechsel, 179
- kartesisches Blatt, 156
- Kathetensatz des Euklid, 27
- Ketten
 - linie, 190
 - regel, 77
- Klumpentopologie, 8
- kommutatives Diagramm, 159
- kompakt, 13
- Komponentenfunktion, 24
- konstante Abbildung, 75
- kontrahierend, 105, 146
- Konvergenz, 11
 - gleichmäßige —, 37
 - punktweise —, 37
- Kosinussatz, 65
- Kugelkoordinaten, 150
- Kurve, 46
- Lagrangescher Multiplikator, 95, 174
- Länge, 45
- Leibniz-Formel, 149
- Lemma von Riesz, 183
- Lemniskate
 - von Bernoulli, 190
 - von Gerono, 54, 161
- Lindelöf, 115
 - Satz von Picard–, 115
- lineare
 - Abbildung, 75
 - Differentialgleichung
 - erster Ordnung, 129
 - homogene —, 129
 - inhomogene —, 129
- Lipschitz, 114
 - Bedingung, 114, 118
 - Konstante, 114
- logarithmische Spirale, 51
- lokale
 - Lipschitz-Bedingung, 118
 - r Diffeomorphismus, 149
 - s Extremum, 81
 - s Maximum, 81
 - unter Nebenbedingungen, 94, 174
 - s Minimum, 81
 - unter Nebenbedingungen, 94, 174
- Lösung eines Systems von Differentialgleichungen, 114
- Mannigfaltigkeit, 177
 - differenzierbare —, 179
 - topologische —, 177
- Matrix
 - Hesse- —, 87
 - Jacobi- —, 74

- symmetrische —, 76, 87, 101
 Maximum
 globales — unter einer Nebenbedingung, 94
 isoliertes lokales —, 89
 lokales —, 81
 unter Nebenbedingungen, 94, 174
 Norm, 3
 Metrik, 2
 euklidische —, 3
 induzierte, 2
 triviale —, 2
 metrischer Raum, 2
 Minimum
 globales — unter einer Nebenbedingung, 94
 isoliertes lokales —, 89
 lokales —, 81
 unter Nebenbedingungen, 94, 174
 Mittelwertsatz, 95, 108, 199
 monomiale Funktion, 26

 Nabla-Operator, 62
 Nebenbedingung, 93, 94
 negativ
 definit, 88
 semidefinit, 88, 196
 Neil, 44
 sche Parabel, 44, 54
 Newton, 53
 scher Knoten, 53
 verfahren, 111
 Niveaumenge, 64
 Norm, 2
 euklidische —, 3
 Maximum—, 3
 normierter Vektorraum, 3
 numerische Exzentrizität, 51

 offen, 5, 7, 139, 178
 e Abbildung, 139
 e Menge, 7
 e Überdeckung, 13
 er Ball, 4
 er Kern, 10
 Operatornorm, 39, 110
 orthogonale Gruppe, 169

 Parallelogrammgleichung, 181
 Parametertransformation, 46
 partiell
 differenzierbar, 59, 192
 stetig — —, 59, 192
 zweimal — —, 66
 e Ableitung, 59, 192
 Picard, 115
 Satz von —–Lindelöf, 115
 Polarkoordinaten, 29, 142, 149
 Polynom, charakteristisches —, 136
 polynomiale Funktion, 26
 positiv
 definit, 88
 semidefinit, 88, 196
 potentielle Energie, 103
 Potenzmenge, 7
 projektiver Raum, 179
 punktweise Konvergenz, 37

 Quader, 14
 quasi-kompakt, 13
 Quotiententopologie, 178

 Rand, 9
 punkt, 8
 Raum
 metrischer —, 2
 reell projektiver —, 179
 topologischer —, 7
 reell projektiver Raum, 179
 regulär, 53
 rektifizierbar, 45
 Richtungsableitung, 57
 Riesz, 183
 Lemma von —, 183

 Satz
 des Thales, 26
 über implizite Funktionen, 159
 von Heine–Borel, 18
 von Picard–Lindelöf, 115
 von Schwarz, 68
 Schachtelungsprinzip, 15
 Schwarz, 68
 semidefinit
 negativ —, 88, 196
 positiv —, 88, 196
 singulär, 53
 SNCF-Metrik, 181
 Spur, 43
 Standardskalarprodukt, 58
 stetig, 23, 184
 differenzierbar, 40, 48, 77, 148, 172
 e Fortsetzung, 30
 gleichmäßig —, 35
 partiell differenzierbar, 59, 192
 zweimal — —, 67

- symmetrische
 - Bilinearform, 87
 - Matrix, 76, 87, 101
- System von Differentialgleichungen
 - erster Ordnung, 113

- Tangente, 27
- Tangential
 - ebene, 75
 - raum, 167
 - vektor, 97, 158, 167
- Taylorformel, 86
- Teilung, 45
- Thales, 26
- Topologie, 7
 - abzählbare Basis der —, 177
 - Basis der —, 177
 - diskrete —, 8
 - Klumpen—, 8
- topologische
 - Mannigfaltigkeit, 177
 - r Raum, 7
- trigonometrische Funktion, 50
- triviale Metrik, 2
- Typ, 172, 173, 179

- Überdeckung, offene —, 13
- Umgebung, 5
 - ε —, 6
- Umkehrsatz, 145, 147
- Untermannigfaltigkeit, 163, 173

- Variation der Konstanten, 131
- Vektorraum der beschränkten stetigen Funktionen, 37
- Vierecksungleichung, 181
- Viertelkreis, 47
- vollständig, 12

- Weg, 43
- wegzusammenhängend, 140
- Winkel, 65

- Zusammenhangskomponenten, 141
- zweimal
 - partiell differenzierbar, 66
 - stetig partiell differenzierbar, 67
- Zykloide, 52, 55
- Zylinderkoordinaten, 150