

Analysis 2

Sommersemester 2023

Gandalf Lechner¹

Diese Vorlesung ist die direkte Fortsetzung der Analysis 1 aus dem ersten Semester. Nach einem Nachtrag zum Integrationskapitel der Analysis 1 werden wir uns einerseits vertieft mit metrischen Räumen und andererseits mit der Differentialrechnung in mehreren Variablen beschäftigen.



¹Department Mathematik, FAU Erlangen-Nürnberg, Cauerstr. 11 Erlangen,
gandalf.lechner@fau.de

Inhaltsverzeichnis

7 Uneigentliche Integrale	1
7.1 Definition uneigentlicher Integrale	1
7.2 Anwendungen und absolute Integrierbarkeit	4
8 Normierte Räume	9
8.1 Grundlegende Definitionen	9
8.2 Konvexe Funktionen und fundamentale Ungleichungen	11
8.3 Konvergenz in metrischen Räumen	14
8.4 Die Topologie des \mathbb{K}^d	20
9 Kompaktheit	25
9.1 Motivation und Definition	25
9.2 Eigenschaften kompakter Mengen	27
9.3 Kompakte Teilmengen von \mathbb{K}^d	30
9.4 Stetige Funktionen auf kompakten Mengen	32
10 Stetige Abbildungen auf endlichdimensionalen Vektorräumen	34
10.1 Stetige Funktionen mit mehreren Variablen und mehreren Komponenten . .	34
10.2 Stetigkeit von linearen Abbildungen	37
10.3 Skalarprodukte	42
11 Kurven	45
11.1 Stetige und differenzierbare Kurven	45
11.2 Bogenlänge	47
12 Differentialrechnung in mehreren Variablen	52
12.1 Totale Ableitungen, Richtungsableitungen, partielle Ableitungen	52
12.2 Ableitungsregeln	57
12.3 Vergleich von Differenzierbarkeitsbegriffen und Mittelwertsatz	63
12.4 Höhere Ableitungen und Vektor-/Gradientenfelder	68
12.5 Der Satz von Taylor für mehrere Variablen	75
12.6 Lokale Extrema	81
13 Umkehrfunktionen, implizite Funktionen, Untermannigfaltigkeiten	88
13.1 Umkehrfunktionen linearer Abbildungen	88
13.2 Der Satz über die Umkehrfunktion	91
13.3 Implizite Funktionen	98
13.4 Untermannigfaltigkeiten des \mathbb{R}^d	104
13.5 Extrema unter Nebenbedingungen	111
Literatur	118

7 Uneigentliche Integrale

Dieses Kapitel ist ein Nachtrag zum Kapitel zur Integralrechnung (Kapitel 6) aus dem Analysis-1-Skript [Lec22].

Dort hatten wir das Integral $\int_a^b f$ einer Riemann-integrablen Funktion über ein kompaktes (also beschränktes und abgeschlossenes) Intervall $I = [a, b]$ kennengelernt. Gemäß unserer Definition 6.5 sind Riemann-integrable Funktionen die *beschränkten* Funktionen, deren Ober- und Unterintegral übereinstimmen. In Satz 6.8 hatten wir dann gesehen, dass insbesondere die stetigen Funktionen auf I Riemann integabel sind. Für stetiges $f : I \rightarrow \mathbb{R}$ sagte dann der Hauptsatz der Differential- und Integralrechnung, dass $F : I \rightarrow \mathbb{R}$, $F(x) := \int_a^x f(y)dy$ eine Stammfunktion von f ist.

7.1 Definition uneigentlicher Integrale

In vielen Situationen ist die Einschränkung auf kompakte Intervalle als Integrationsgebiete und beschränkte Funktionen als Integranden zu eng. In diesem Kapitel werden wir diese Einschränkungen mit sogenannten uneigentlichen Integralen aufheben. Zuerst geben wir zwei motivierende Beispiele.

Beispiel 7.1.

a) Sei $p > 1$ und $b > 1$. Dann ist das Integral

$$\int_1^b \frac{1}{x^p} dx = \left[\frac{1}{1-p} x^{1-p} \right]_1^b = \frac{1}{1-p} (b^{1-p} - 1)$$

für alle $b > 1$ wohldefiniert und gibt den Flächeninhalt von $A_b = \{(x, y) : 1 \leq x \leq b, 0 \leq y \leq x^{-p}\}$ an. Da $p > 1$, konvergiert b^{1-p} für $b \rightarrow \infty$ gegen Null. Wir könnten also auf die Idee kommen,

$$\int_1^\infty \frac{1}{x^p} dx := \lim_{b \rightarrow \infty} \int_1^b \frac{1}{x^p} dx = \frac{1}{p-1}$$

zu definieren und als Flächeninhalt der unbeschränkten Fläche $A_\infty = \{(x, y) : 1 \leq x, 0 \leq y \leq x^{-p}\}$ anzusehen.

Dies gibt hier einen endlichen Wert. Für $p = 1$ (oder $p < 1$) wird diese Idee allerdings nicht funktionieren, denn

$$\int_1^b \frac{1}{x} dx = [\log(x)]_1^b = \log(b) \xrightarrow{b \rightarrow \infty} \infty$$

divergiert für $b \rightarrow \infty$.

b) Sei $0 < p < 1$ und $0 < a < 1$. Dann ist das Integral

$$\int_a^1 \frac{1}{x^p} dx = \left[\frac{1}{1-p} x^{1-p} \right]_a^1 = \frac{1}{1-p} (1 - a^{1-p})$$

für alle $0 < a < 1$ wohldefiniert und gibt den Flächeninhalt von $\tilde{A}_a = \{(x, y) : a \leq x \leq 1, 0 \leq y \leq x^{-p}\}$ an. Da $1 - p > 0$, konvergiert a^{1-p} für $a \rightarrow 0$ gegen Null. Wir könnten also auf die Idee kommen,

$$\int_0^1 \frac{1}{x^p} dx := \lim_{a \rightarrow 0} \int_a^1 \frac{1}{x^p} dx = \frac{1}{1-p}$$

zu definieren und als Flächeninhalt der unbeschränkten Fläche $\tilde{A}_0 = \{(x, y) : 0 \leq x \leq 1, 0 \leq y \leq x^{-p}\}$ anzusehen.

Dies gibt hier einen endlichen Wert. Für $p \geq 1$ wird diese Idee allerdings nicht funktionieren, denn z.B. für $p = 1$

$$\int_a^1 \frac{1}{x} dx = [\log(x)]_a^1 = -\log(a) \xrightarrow{a \rightarrow 0} \infty$$

divergiert für $a \rightarrow 0$.

Sei $p \in \mathbb{R}$. Zeigen Sie, dass $\lim_{b \rightarrow \infty} \int_1^b x^{-p} dx$ genau dann existiert, wenn $p > 1$ ist. Zeigen Sie, dass $\lim_{a \rightarrow 0+} \int_a^1 x^{-p} dx$ genau dann existiert, wenn $p < 1$ ist.

Definition 7.2 (Uneigentliche Integrale). Sei $a \in \mathbb{R}$ und $a \leq b \leq \infty$ und $f : [a, b) \rightarrow \mathbb{R}$ eine Funktion, so dass für alle $t \in [a, b)$ die eingeschränkte Funktion $f|_{[a,t]}$ Riemann-integrierbar ist. Eine Funktion mit dieser Eigenschaft nennt man auch auf $[a, b)$ *lokal (Riemann-) integrierbar*. Falls der Grenzwert

$$\lim_{t \rightarrow b-} \int_a^t f(x) dx =: \int_a^b f(x) dx$$

existiert, so nennen wir ihn das *uneigentliche Integral von f über $[a, b)$* . Man sagt dann auch, das uneigentliche Integral $\int_a^b f(x) dx$ konvergiert.

Analog definiert man für $b \in \mathbb{R}$ und $-\infty \leq a \leq b$ das uneigentliche Integral

$$\int_a^b f(x) dx := \lim_{t \rightarrow a+} \int_t^b f(x) dx,$$

falls dieser Grenzwert existiert.

Im allgemeinen Fall sind beide Integrationsgrenzen $a, b \in \mathbb{R} \cup \{\pm\infty\}$ als Grenzwerte zu verstehen: Falls eine Funktion $f : [a, b] \rightarrow \mathbb{R}$ die Eigenschaft hat, dass sie über jedes kompakte Teilintervall $[\alpha, \beta] \subset (a, b)$ Riemann-integrierbar ist, und falls für $c \in (a, b)$ die beiden Grenzwerte

$$\int_a^c f(x) dx := \lim_{\alpha \rightarrow a+} \int_\alpha^c f(x) dx, \quad \int_c^b f(x) dx := \lim_{\beta \rightarrow b-} \int_c^\beta f(x) dx$$

existieren, so heißt das uneigentliche Integral $\int_a^b f(x)dx$ konvergent, und man setzt

$$\int_a^b f(x)dx := \int_a^c f(x)dx + \int_c^b f(x)dx.$$

Diese Definition uneigentlicher Integrale orientiert sich genau an den in dem vorhergehenden Beispiel gemachten Beobachtungen. Beachten Sie, dass für ein uneigentliches Integral $\int_a^b f(x)dx$ der Integrand weder auf (a, b) beschränkt noch am Rand definiert sein muss, d.h. $f(a)$ und/oder $f(b)$ müssen nicht existieren.

Beispiel 7.3.

a) Linke kritische Grenze bei $x = 0$:

$$\int_0^2 \frac{dx}{\sqrt{x}} = \lim_{a \rightarrow 0^+} \int_a^2 \frac{dx}{\sqrt{x}} = \lim_{a \rightarrow 0^+} [2\sqrt{x}]_a^2 = \lim_{a \rightarrow 0^+} 2(\sqrt{2} - \sqrt{a}) = 2\sqrt{2}.$$

b) Rechte kritische Grenze bei $x = 1$:

$$\int_0^1 \frac{dx}{\sqrt{1-x^2}} = \lim_{b \rightarrow 1^-} \int_0^b \frac{dx}{\sqrt{1-x^2}} = \lim_{b \rightarrow 1^-} [\arcsin(x)]_0^b = \lim_{b \rightarrow 1^-} \arcsin(b) = \frac{\pi}{2}.$$

c) Beide Grenzen kritisch:

$$\begin{aligned} \int_{-\infty}^{\infty} \frac{dx}{1+x^2} &= \lim_{a \rightarrow -\infty} \int_a^0 \frac{dx}{1+x^2} + \lim_{b \rightarrow \infty} \int_0^b \frac{dx}{1+x^2} \\ &= \lim_{a \rightarrow -\infty} (\arctan(0) - \arctan(a)) + \lim_{b \rightarrow \infty} (\arctan(b) - \arctan(0)) \\ &= -\left(-\frac{\pi}{2} - 0\right) + \left(\frac{\pi}{2} - 0\right) = \pi. \end{aligned}$$

d) Noch ein Beispiel mit zwei kritischen Grenzen: Das uneigentliche Integral $\int_{-\infty}^{\infty} \frac{x}{1+x^2} dx$ konvergiert *nicht*: Wir haben $\frac{x}{1+x^2} \geq \frac{1}{2x}$ für $x \geq 1$, also (wie in Beispiel 1.1)

$$\int_1^{\infty} \frac{x}{1+x^2} dx \geq \int_1^{\infty} \frac{1}{2x} dx = \infty,$$

wobei wir die Monotonie des Integrals verwendet haben (Satz 6.9). Nichtsdestotrotz gilt

$$\lim_{r \rightarrow \infty} \int_{-r}^r \frac{x}{1+x^2} dx = 0$$

wegen der Antisymmetrie des Integranden. Bei einem uneigentlichen Integral $\int_a^b f(x)dx$ mit zwei kritischen Grenzen ist es also wichtig, dass die beiden Grenzwerte $\lim_{\alpha \rightarrow a^+} \int_{\alpha}^c f(x)dx$ und $\lim_{\beta \rightarrow b^-} \int_c^{\beta} f(x)dx$ beide einzeln (und nicht nur in Summe) existieren.

Konvergieren die uneigentlichen Integrale $\int_0^\infty \frac{\sqrt{x}}{\sqrt{1+x^4}} dx$ bzw. $\int_e^\infty \frac{1}{\log x} dx$?

7.2 Anwendungen und absolute Integrierbarkeit

Eine interessante Anwendung von uneigentlichen Integralen ist das folgende Vergleichskriterium für Reihen.

Satz 7.4 (Integral-Vergleichskriterium für Reihen). Sei $f : [1, \infty) \rightarrow \mathbb{R}_+$ eine monoton fallende Funktion. Dann gilt:

$$\sum_{n=1}^{\infty} f(n) \text{ konvergiert} \Leftrightarrow \int_1^{\infty} f(x) dx \text{ existiert als uneigentliches Riemann-Integral.}$$

Beweis. Wir definieren zwei Treppenfunktionen $\varphi, \psi : [1, \infty) \rightarrow \mathbb{R}$ durch

$$\begin{aligned} \psi(x) &:= f(n) \quad \text{für } n \leq x < n+1, \\ \varphi(x) &:= f(n+1) \quad \text{für } n \leq x < n+1. \end{aligned}$$

Aufgrund der Monotonie von f gilt $\varphi \leq f \leq \psi$. Ebenfalls aufgrund der Monotonie von f ist f lokal Riemann-integrierbar, d.h. die Integrale $\int_1^N f(x) dx$ existieren für alle $N \in \mathbb{N}$. Die Monotonie des Integrals liefert nun

$$\sum_{n=1}^{N-1} f(n+1) = \int_1^N \varphi(x) dx \leq \int_1^N f(x) dx \leq \int_1^N \psi(x) dx = \sum_{n=1}^{N-1} f(n).$$

Falls das uneigentliche Integral $\int_1^\infty f(x) dx = \lim_{b \rightarrow \infty} \int_1^b f(x) dx$ existiert, so erhalten wir die obere Schranke $\sum_{n=2}^N f(n) \leq \int_1^N f(x) dx \leq \int_1^\infty f(x) dx$. In diesem Fall ist die Folge $(\sum_{n=2}^N f(n))_N$ monoton wachsend und nach oben beschränkt, also konvergent, d.h. $\sum_{n=1}^\infty f(n)$ konvergiert.

Ist andererseits $\sum_{n=1}^\infty f(n)$ konvergent, so haben wir die obere Schranke $\int_1^N f(x) dx \leq \sum_{n=1}^N f(n) \leq \sum_{n=1}^\infty f(n)$. Also ist $F : b \mapsto \int_1^b f(x) dx$ monoton wachsend und beschränkt. Dies impliziert, dass der Grenzwert $\lim_{b \rightarrow \infty} F(b)$ existiert: Nach Definition des Supremums $s := \sup_{b \geq 1} F(b)$ existiert zu jedem $\varepsilon > 0$ ein $b \geq 1$ mit $s - \varepsilon < F(b) \leq s$, wegen der Monotonie von F also $s - \varepsilon < F(c) \leq s$ für alle $c \geq b$. Dies impliziert $|F(c) - s| < \varepsilon$ für alle $c \geq b$, also gilt $F(c) \rightarrow s$ für $c \rightarrow \infty$. \square

Beispiel 7.5.

- a) Wir betrachten $p > 0$ und die Reihen $\sum_{n=1}^\infty \frac{1}{n^p}$. Da die Funktion $f : [1, \infty) \rightarrow \mathbb{R}_+$, $f(x) := x^{-p}$ monoton fallend und nicht negativ ist, können wir Satz 7.4 verwenden und sehen, dass diese Reihe genau dann konvergiert, wenn $\int_1^\infty \frac{dx}{x^p}$ existiert. Wir haben bereits in Beispiel 7.1 gesehen, dass dies genau für $p > 1$ der Fall ist.

Das uns bereits aus Analysis 1 bekannte Ergebnis, dass $\sum_{n=1}^\infty n^{-p}$ genau für $p > 1$ konvergiert, erhalten wir so also einfacher per Integral-Vergleichskriterium.^a

b) Für $p = 1$ erhalten wir die divergente harmonische Reihe $\sum_{n=1}^{\infty} \frac{1}{n}$. Mit den Abschätzungen in dem Beweis von Satz 7.4 sehen wir

$$\sum_{n=1}^{N-1} \frac{1}{n+1} \leq \int_1^N \frac{1}{x} dx = \log(N) \leq \sum_{n=1}^{N-1} \frac{1}{n},$$

also $\frac{1}{N} \leq \gamma_N := \sum_{n=1}^N \frac{1}{n} - \log(N) \leq 1$. Die Folge $(\gamma_N)_N$ ist also nach unten (und oben) beschränkt und wegen

$$\gamma_{N-1} - \gamma_N = -\frac{1}{N} + \log(N) - \log(N-1) = \int_{N-1}^N \left(\frac{1}{x} - \frac{1}{N} \right) dx > 0$$

monoton fallend, also konvergent. Dies bedeutet, dass für $N \rightarrow \infty$ die Folge der Partialsummen $\left(\sum_{n=1}^N \frac{1}{n} \right)_{N \in \mathbb{N}}$ genauso schnell (oder eher: genauso langsam) wie $\log(N)$ gegen ∞ geht.

Der Grenzwert

$$\gamma := \lim_{N \rightarrow \infty} \gamma_N$$

wird *Euler-Mascheronische Konstante* genannt und hat einen Näherungswert von $\gamma \cong 0,57721\dots$. Es wird vermutet, dass γ irrational ist, ein Beweis ist aber nicht bekannt.

^aWir können also eine Funktion $\zeta : (1, \infty) \rightarrow \mathbb{R}$, $\zeta(p) := \sum_{n=1}^{\infty} n^{-p}$ definieren, die sogenannte *Riemann'sche Zetafunktion*. Diese Funktion spielt eine große Rolle in der Zahlentheorie und ist Gegenstand der Riemann'schen Vermutung. Für die Formulierung dieser Vermutung ist es allerdings notwendig, ζ auf komplexe Argumente auf geeignete Art und Weise auszudehnen. Solche "analytischen Fortsetzungen" werden in der Funktionentheorie-Vorlesung eingeführt.

Als Korollar zum Beweis von Satz 7.4 notieren wir:

Korollar 7.6. Sei $f : [a, b) \rightarrow \mathbb{R}_+$, $f \geq 0$ eine nicht negative Funktion, so dass $f|_{[a,t]}$ für alle $t \in [a, b)$ Riemann-integrierbar ist. Dann existiert das uneigentliche Integral $\int_a^b f(x) dx$ genau dann, wenn $F : t \mapsto \int_a^t f(x) dx$ auf $[a, b)$ beschränkt ist.

Beweis. Die Funktion F ist monoton wachsend, da $f \geq 0$ (wir haben $F(t_1) = \int_a^{t_1} f \leq \int_a^{t_2} f$ für $t_1 \leq t_2$). Ist F beschränkt, so folgt wie im letzten Teil des vorhergehenden Beweises, dass $F(t)$ für $t \rightarrow b$ konvergiert. Ist F andererseits unbeschränkt, so gilt $F(t) \rightarrow \infty$ für $t \rightarrow b$, in diesem Fall existiert das uneigentliche Integral also nicht. \square

In den meisten Fällen wird die Konvergenz eines uneigentlichen Integrals durch Abschätzung gegen ein bereits bekanntes uneigentliches Integral geprüft. Dazu führen wir kurz den Begriff der absoluten Integrierbarkeit (analog zur absoluten Konvergenz von Reihen, Def. 3.39) ein.

Definition 7.7. Falls $f : [a, b) \rightarrow \mathbb{R}$ lokal integrierbar ist und das uneigentliche Integral $\int_a^b |f(x)| dx$ existiert, so heißt f *absolut integrierbar* (über $[a, b)$).

Erinnern Sie sich, dass für integribles f auch $|f|$ integribel ist (Lemma 6.10). Also ist das Integral $\int_a^b |f(x)| dx$ in dieser Definition wohldefiniert.

Lemma 7.8. *Absolut integribel Funktionen sind integribel.*

Beweis. Wir führen den Beweis für den Fall, dass die Integrale an der oberen Grenze $x = b$ uneigentlich sind. Sei $(b_n)_n$ eine Folge im Definitionsbereich $[a, b)$ der Funktion f , mit $b_n \rightarrow b$ für $n \rightarrow \infty$. Dann haben wir

$$\left| \int_a^{b_n} f - \int_a^{b_m} f \right| = \left| \int_{b_m}^{b_n} f \right| \leq \left| \int_{b_m}^{b_n} |f| \right| = \left| \int_a^{b_m} |f| - \int_a^{b_n} |f| \right|.$$

Da $\int_a^{b_n} |f|$ und $\int_a^{b_m} |f|$ für $n, m \rightarrow \infty$ gegen das uneigentliche Integral $\int_a^b |f|$ konvergieren, geht die rechte Seite der obigen Ungleichung für $n, m \rightarrow \infty$ gegen Null. Demnach ist $\left(\int_a^{b_n} f \right)_{n \in \mathbb{N}}$ eine Cauchyfolge, konvergiert also. Da die Folge (b_n) beliebig war, bedeutet das aber gerade, dass das uneigentliche Integral $\int_a^b f = \lim_{n \rightarrow \infty} \int_a^{b_n} f$ existiert. \square

Korollar 7.9. *Sei $a > 0$ und $f : [a, \infty) \rightarrow \mathbb{R}$ lokal integribel. Dann gilt:*

a) *Existieren $p > 1, c > 0, x_0 \geq a$, so dass*

$$|f(x)| \leq \frac{c}{x^p} \quad \text{für } x \geq x_0,$$

so konvergiert das Integral $\int_a^\infty f(x) dx$ absolut.

b) *Existieren $p \leq 1, c > 0$, und $x_0 \geq a$, so dass*

$$f(x) \geq \frac{c}{x^p} \quad \text{für } x \geq x_0,$$

so divergiert das Integral $\int_a^\infty f(x) dx$.

Beweis. Wir teilen das fragliche Integral gemäß $\int_a^\infty f = \int_a^{x_0} f + \int_{x_0}^\infty f$ auf. Da $\int_a^{x_0} f$ per Annahme (f lokal integribel) existiert, müssen wir uns nur um $\int_{x_0}^\infty f$ kümmern. Wie aus den vorigen Beispielen bekannt ist, existiert $\int_{x_0}^\infty \frac{c}{x^p} dx$ genau dann wenn $p > 1$.

In Teil a) wird $|f|$ also durch eine integribel Funktion dominiert und ist deshalb integribel (Majorantenkriterium). Das heißt, dass f absolut integribel ist, nach Lemma 7.8 ist f dann auch integribel.

In Teil b) ist hingegen f größer als eine nicht integribel Funktion (mit $\int_{x_0}^\infty \frac{c}{x^p} dx = \infty$), also divergiert $\int_a^\infty f(x) dx$. \square

Beispiel 7.10 (Die Gammafunktion). Wir definieren die Gammafunktion

$$\Gamma : (0, \infty) \rightarrow \mathbb{R}, \quad \Gamma(x) := \int_0^\infty t^{x-1} e^{-t} dt.$$

Hier ist das Integral an beiden Grenzen als uneigentlich zu verstehen (Bei $t = \infty$ so-

wieso, und bei $t = 0$, da für $x < 1$ der Integrand $t \mapsto t^{x-1}e^{-t}$ gegen ∞ geht, wenn $t \rightarrow 0+$ geht.). Um die Konvergenz des Integrals zu zeigen, teilen wir es bei $t = 1$ auf und betrachten $\int_0^1 t^{x-1}e^{-t}dt$ und $\int_1^\infty t^{x-1}e^{-t}dt$ separat.

Für das Integral über $(0, 1]$ schätzen wir den Integranden ab gemäß ($t \geq 0$)

$$t^{x-1}e^{-t} \leq \frac{1}{t^{1-x}}.$$

Da $x > 0$ im Definitionsbereich von Γ , sehen wir mit Beispiel 7.1 b) und dem Majorantenkriterium, dass das Integral über $(0, 1]$ existiert.

Für das Integral über $[1, \infty)$ benutzen wir $\lim_{t \rightarrow \infty} t^u e^{-t} = 0$ für alle $u > 0$ (siehe Beispiel 4.23). Somit ist $t^{x+1}e^{-t}$ auf $[1, \infty)$ beschränkt durch eine Konstante c_x , also $t^{x-1}e^{-t} \leq \frac{c_x}{t^2}$. Nun liefert Korollar 7.9 die Existenz von $\int_1^\infty t^{x-1}e^{-t}dt$.

Die Gammafunktion ist eine wichtige nicht elementare Funktion. Ihre zentrale Eigenschaft ist die Funktionalgleichung

$$\Gamma(x+1) = x \cdot \Gamma(x), \quad x > 0.$$

Wir beweisen diese Gleichung durch partielle Integration. Dazu berechnen wir zunächst für $a < b < \infty$:

$$\int_a^b t^x e^{-t} dt = -[t^x e^{-t}]_{t=a}^{t=b} + x \int_a^b t^{x-1} e^{-t} dt.$$

Der Randterm $-e^{x \log b} e^{-b} + e^{x \log a} e^{-a}$ konvergiert für $a \rightarrow 0+$ und $b \rightarrow \infty$ gegen Null. In diesem Limes geht das Integral auf der linken Seite der Gleichung gegen $\Gamma(x+1)$ (vgl. Definition von $\Gamma(x)$) und das auf der rechten Seite gegen $x\Gamma(x)$, was den Beweis der Funktionalgleichung abschließt.

Wir bemerken noch, dass die Gammafunktion die Fakultät interpoliert, d.h.

$$\Gamma(n) = (n-1)!, \quad n \in \mathbb{N}.$$

Um dies zu beweisen, betrachten wir zunächst $n = 1$. Es gilt

$$\Gamma(1) = \int_0^\infty e^{-t} dt = -[e^{-t}]_0^\infty = 1 = (1-1)!.$$

Für $n > 1$ ergibt sich die Behauptung aus der Funktionalgleichung per Induktion, z.B. $\Gamma(4) = 3 \cdot \Gamma(2) = 3 \cdot 2 \cdot \Gamma(1) = (4-1)!$.

Beispiel 7.11 (Fresnel'sches Integral). In diesem Beispiel betrachten wir das uneigentliche *Fresnel'sche Integral*

$$\int_0^\infty \sin(x^2) dx.$$

Zuerst formen wir es mit Hilfe der Substitution $x^2 = t$ um:

$$\begin{aligned}\int_0^\infty \sin(x^2) dx &= \lim_{r \rightarrow \infty} \int_0^r \sin(x^2) dx = \lim_{r \rightarrow \infty} \int_0^r \sin(t) dx \\ &= \lim_{r \rightarrow \infty} \frac{1}{2} \int_0^{r^2} \frac{\sin(t)}{\sqrt{t}} dt = \lim_{r \rightarrow \infty} \frac{1}{2} \int_0^r \frac{\sin(t)}{\sqrt{t}} dt.\end{aligned}$$

Wir wollen nun zeigen, dass dieser Limes existiert. Dazu erinnern wir uns zunächst, dass für jedes $k \in \mathbb{N}$ gilt:

$$\begin{aligned}\sin t &\geq 0 \quad \text{für } 2k\pi \leq t \leq (2k+1)\pi \\ \sin t &\leq 0 \quad \text{für } (2k+1)\pi \leq t \leq (2k+2)\pi.\end{aligned}$$

Außerdem gilt $\sin(t + \pi) = -\sin(t)$ für alle $t \in \mathbb{R}$. Das ergibt

$$-\int_{(2k+1)\pi}^{(2k+2)\pi} \frac{\sin(t)}{\sqrt{t}} dt = -\int_{2k\pi}^{(2k+1)\pi} \frac{\sin(t + \pi)}{\sqrt{t + \pi}} dt = \int_{2k\pi}^{(2k+1)\pi} \frac{\sin(t)}{\sqrt{t + \pi}} dt \leq \int_{2k\pi}^{(2k+1)\pi} \frac{\sin(t)}{\sqrt{t}} dt.$$

Ganz analog zeigt man

$$\int_{2k\pi}^{(2k+1)\pi} \frac{\sin(t)}{\sqrt{t}} dt \leq -\int_{(2k-1)\pi}^{2k\pi} \frac{\sin(t)}{\sqrt{t}} dt.$$

Zusammengefasst zeigen diese Beobachtungen, dass die Folge $a_n := (-1)^n \int_{n\pi}^{(n+1)\pi} \frac{\sin(t)}{\sqrt{t}} dt$ nicht negativ und monoton fallend ist. Weiterhin gilt (Standardabschätzung durch Mittelwertsatz der Integralrechnung)

$$|a_n| \leq \int_{n\pi}^{(n+1)\pi} \frac{1}{\sqrt{t}} dt \leq \frac{\pi}{\sqrt{n\pi}} \xrightarrow{n \rightarrow \infty} 0.$$

Damit sind alle Voraussetzungen für das Leibnizkriterium erfüllt, und wir sehen, dass

$$\sum_{n=0}^{\infty} (-1)^n a_n = \lim_{N \rightarrow \infty} \sum_{n=0}^N (-1)^n a_n = \lim_{N \rightarrow \infty} \int_0^{(N+1)\pi} \frac{\sin(t)}{\sqrt{t}} dt$$

existiert. Dieses Ergebnis ist schon fast die gewünschte Konvergenz von $F(r) := \int_0^r \frac{\sin(t)}{\sqrt{t}} dt$ für $r \rightarrow \infty$: Für $r > 0$ finden wir $n \in \mathbb{N}$, so dass $n\pi \leq r \leq (n+1)\pi$, und erhalten

$$|F(r) - F(n\pi)| = \left| \int_{n\pi}^r \frac{\sin(t)}{\sqrt{t}} dt \right| \leq \int_{n\pi}^{(n+1)\pi} \frac{dt}{\sqrt{t}} \leq \frac{\pi}{\sqrt{n\pi}} \rightarrow 0.$$

Das impliziert, dass $F(r)$ für $r \rightarrow \infty$ konvergiert und beendet den Beweis der Existenz des Fresnel'schen Integrals.

Man kann zeigen, dass der Wert dieses Integrals $\frac{1}{2}\sqrt{\frac{\pi}{2}}$ ist.

Beachten Sie, dass $t \mapsto \frac{\sin(t)}{\sqrt{t}}$ über $(0, \infty)$ *nicht* absolut integrierbar ist. Das liegt im Wesentlichen daran, dass $t^{-1/2}$ nicht schnell genug abfällt (Potenz $p < 1$).

8 Normierte Räume

Ein Hauptthema der Analysis 2 ist das Studium von stetigen und differenzierbaren Funktionen $f : D \subset \mathbb{R}^d \rightarrow \mathbb{R}^k$ zwischen den Vektorräumen \mathbb{R}^d und \mathbb{R}^k beliebiger endlicher Dimensionen $k, d \in \mathbb{N}$. Im Vergleich zur Analysis 1, wo wir uns auf $d = k = 1$, also auf Funktionen $f : D \subset \mathbb{R} \rightarrow \mathbb{R}$, beschränkt haben, werden wir hier viele neue Phänomene antreffen.

Während wir in Analysis 1 zuerst den eindimensionalen Vektorraum (Körper) \mathbb{R} bzw. seine Teilmengen (offene/abgeschlossene Intervalle...) studiert haben, beginnen wir hier unsere Untersuchungen mit einigen Überlegungen zu normierten Vektorräumen, für die \mathbb{R}^d ein typisches Beispiel sein wird. Dies sind insbesondere Vektorräume (hier immer über dem Körper \mathbb{R} oder \mathbb{C}), so dass wir auf Konzepte der Linearen Algebra zurückgreifen werden. Allerdings wird es zusätzlich zu den algebraischen Eigenschaften in der Analysis immer wichtig sein, einen Abstandsbegriff zur Verfügung zu haben, um Konzepte wie Konvergenz zu definieren.

8.1 Grundlegende Definitionen

Erinnern wir dazu zunächst an die Definition eines normierten Vektorraums (vgl. Definition 4.41) und eines metrischen Raums (Definition 4.1):

Definition 8.1. Sei V ein Vektorraum über einem Körper $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$. Eine Abbildung

$$\|\cdot\| : V \rightarrow \mathbb{R}_{\geq 0}$$

heißt *Norm*, falls für beliebige $v, w \in V$, $\lambda \in \mathbb{K}$, die folgenden drei Eigenschaften gelten:

- a) $\|v\| = 0$ ist äquivalent zu $v = 0$.
- b) $\|\lambda v\| = |\lambda| \|v\|$
- c) $\|v + w\| \leq \|v\| + \|w\|$ (Dreiecksungleichung)

Ein Vektorraum mit Norm heißt *normierter Raum* $(V, \|\cdot\|)$.

Eine Norm ist eine Verallgemeinerung der Betragsfunktion $|\cdot|$ auf \mathbb{K} .

Zeigen Sie, dass in jedem normierten Vektorraum $(V, \|\cdot\|)$ die umgekehrte Dreiecksungleichung

$$\|v - w\| \geq \left| \|v\| - \|w\| \right|, \quad v, w \in V,$$

gilt.

Definition 8.2. Ein *metrischer Raum* ist eine Menge X zusammen mit einer *Metrik*, das ist eine Abbildung

$$d : X \times X \rightarrow \mathbb{R}_{\geq 0}$$

mit den Eigenschaften (für beliebige $x, y \in X$)

- a) $d(x, y) = 0$ ist äquivalent zu $x = y$.
- b) $d(x, y) = d(y, x)$ (Symmetrie)
- c) $d(x, z) \leq d(x, y) + d(y, z)$ (Dreiecksungleichung)

Ein normierter Vektorraum ist insbesondere ein metrischer Raum, denn wir können die Metrik

$$d(v, w) := \|v - w\|$$

eingeführen. Metrische Räume sind viel allgemeiner als normierte Vektorräume, sie müssen ja noch nicht einmal Vektorräume sein (vgl. Beispiel 4.2).

Einige Beispiele kennen Sie bereits aus Analysis 1. Wir geben nun einige weitere Beispiele, die wir in der Analysis 2 sehr oft benutzen werden. Im Folgenden steht \mathbb{K} immer für den Körper \mathbb{R} der reellen Zahlen oder den Körper \mathbb{C} der komplexen Zahlen.

Wir betrachten den Vektorraum \mathbb{K}^d mit Dimension $d \in \mathbb{N}$, dessen Elemente wir mit $x = (x_1, \dots, x_d) \in \mathbb{K}$ bezeichnen². Die Zahlen $x_k \in \mathbb{K}$, $k = 1, \dots, d$, bezeichnen wir als die *Komponenten* des Vektors x .

Definition 8.3 (Normen auf \mathbb{K}^d). Auf \mathbb{K}^d definieren wir für beliebiges $p \geq 1$ die sogenannte *p-Norm*

$$\|x\|_p := \left(\sum_{k=1}^d |x_k|^p \right)^{1/p}.$$

Ein spezieller Fall ist die 2-Norm

$$\|x\|_2 := \sqrt{\sum_{k=1}^d |x_k|^2} \quad \text{Euklidische Norm,}$$

und als Grenzfall für $p \rightarrow \infty$ erhalten wir

$$\|x\|_\infty := \max_{k=1, \dots, d} |x_k| \quad \text{Maximumsnorm oder } \infty\text{-Norm.}$$

Für jedes $p \in [1, \infty]$ haben wir also eine Abbildung $\|\cdot\|_p : \mathbb{K}^d \rightarrow \mathbb{R}_{\geq 0}$, die ihrem Namen nach wohl eine Norm sein sollte. Das Überprüfen der Normeigenschaften a) und b) ist ganz einfach – exemplarisch betrachten wir $p = 2$. Dann verschwindet $\|x\|_2 = \sqrt{\sum_{k=1}^d |x_k|^2}$ genau dann, wenn alle $|x_k|$ verschwinden (Summe von nicht-negativen Termen ist nur dann Null, wenn alle Terme Null sind), also genau dann, wenn $x = 0$. Weiterhin ist für $\lambda \in \mathbb{K}$

$$\|\lambda x\|_2 = \sqrt{\sum_{k=1}^d |\lambda x_k|^2} = \sqrt{|\lambda|^2 \sum_{k=1}^d |x_k|^2} = |\lambda| \cdot \|x\|_2.$$

²Es wird für uns nicht nötig sein, zwischen Spalten- und Zeilenvektoren zu unterscheiden.

Schwieriger ist allerdings der Nachweis der Dreiecksungleichung (Eigenschaft c) einer Norm), zumindest für allgemeines p . Wir machen dazu einen kurzen Ausflug zu konvexen Funktionen.

8.2 Konvexe Funktionen und fundamentale Ungleichungen

Gegeben zwei reelle Zahlen x, y , so ist die Menge

$$\{\lambda x + (1 - \lambda)y : 0 \leq \lambda \leq 1\}$$

das abgeschlossene Intervall mit den Endpunkten x und y , also $[x, y]$ falls $x \leq y$ und $[y, x]$ falls $y \leq x$. Mit dem Parameter $\lambda \in [0, 1]$ können Sie den Punkt $\lambda x + (1 - \lambda)y$ von y (für $\lambda = 0$) zu x (für $\lambda = 1$) verschieben. Punkte der Form $\lambda x + (1 - \lambda)y$ werden auch *konvexe Kombination* von x und y genannt.

Definition 8.4. Eine Funktion $f : I \rightarrow \mathbb{R}$ auf einem Intervall $I \subset \mathbb{R}$ heißt *konvex*, wenn für alle $x, y \in I$ und alle $0 \leq \lambda \leq 1$

$$f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y)$$

gilt. Gilt die Ungleichung mit \geq statt \leq , so heißt f *konkav*.

Geometrisch bedeutet Konvexität von f , dass für $x < y$ der Graph von f unterhalb der Sekanten von f durch die Punkte $(x, f(x))$ und $(y, f(y))$ verläuft, d.h. der Graph ist “linksgekrümmt”, wenn Sie ihn von links nach rechts durchlaufen³. Anders formuliert wächst die Steigung des Graphen von links nach rechts. Die genaue Formulierung dieser Tatsache liefert der folgende Satz.

Satz 8.5. Sei $I \subset \mathbb{R}$ ein Intervall und $f : I \rightarrow \mathbb{R}$ differenzierbar. Dann gilt

$$f \text{ konvex} \Leftrightarrow f' \text{ monoton wachsend.}$$

Dies ist insbesondere dann der Fall, wenn f zweimal differenzierbar ist und $f'' \geq 0$ gilt.

Beweis. Angenommen, f' ist monoton wachsend, und $x_1 < x_2$ sind zwei Punkte in I . Für $\lambda \in [0, 1]$ sei $x := \lambda x_1 + (1 - \lambda)x_2 \in [x_1, x_2]$. Für die Konvexität von f müssen wir zeigen, dass $\lambda f(x_1) + (1 - \lambda)f(x_2) = f(x) \leq \lambda f(x_1) + (1 - \lambda)f(x_2)$ gilt, d.h.

$$\lambda(f(x) - f(x_1)) \leq (1 - \lambda)(f(x_2) - f(x)).$$

Nach dem Mittelwertsatz (Satz 5.16 b)) gibt es Punkte $\xi_1 \in (x_1, x)$ und $\xi_2 \in (x, x_2)$ mit

$$f(x) - f(x_1) = f'(\xi_1) \cdot (x - x_1), \quad f(x_2) - f(x) = f'(\xi_2) \cdot (x_2 - x).$$

Wir müssen also nur $\lambda f'(\xi_1)(x - x_1) \leq (1 - \lambda)f'(\xi_2)(x_2 - x)$ zeigen. Da $\xi_1 < x < \xi_2$ haben wir $f'(\xi_1) \leq f'(\xi_2)$ aufgrund des monotonen Wachsens von f' . Weiterhin haben wir

$$\lambda x + (1 - \lambda)x = x = \lambda x_1 + (1 - \lambda)x_2 \Rightarrow \lambda(x - x_1) = (1 - \lambda)(x_2 - x) \geq 0,$$

was die gewünschte Ungleichung impliziert.

Der Beweis der umgekehrten Implikation (f konvex und differenzierbar $\Rightarrow f'$ monoton wachsend) wird als Übungsaufgabe gestellt. \square

³Da f genau dann konvex ist, wenn $-f$ konkav ist, genügt es, konvexe Funktionen zu betrachten.

Beispiel 8.6. Beispiele konvexer Funktionen sind Geraden $x \mapsto ax + b$ (mit 2. Ableitung 0), die Exponentialfunktion \exp (mit 2. Ableitung $\exp > 0$). Die im Zusammenhang mit den p -Normen interessante Potenzfunktion $\mathbb{R}_+ \ni x \mapsto x^p$ hat die 2. Ableitung $x \mapsto p(p-1)x^{p-2}$. Dies ist für $p \geq 1$ und für $p \leq 0$ nicht negativ, d.h. für diese Werte von p ist die Potenzfunktion konvex. Für $0 < p < 1$ ist die Potenzfunktion konkav.

Es empfiehlt sich, die Graphen dieser Funktionen zu skizzieren, um ein Gefühl für Konvexität zu bekommen.

Wir benutzen nun Konvexität, um drei fundamentale Ungleichungen zu beweisen.

Satz 8.7 (Young–Ungleichung). Für $a, b \geq 0$ und $p, q > 1$ mit $\frac{1}{p} + \frac{1}{q} = 1$ gilt

$$ab \leq \frac{a^p}{p} + \frac{b^q}{q}.$$

Beweis. Die Behauptung ist trivial, falls $a = 0$ oder $b = 0$. Seien also $a, b > 0$. Wegen $\frac{1}{p} + \frac{1}{q} = 1$ ist $\frac{1}{p} \log a + \frac{1}{q} \log b$ eine konvexe Kombination von $\log a$ und $\log b$. Mit der Konvexität der Exponentialfunktion erhalten wir so

$$a^{\frac{1}{p}} b^{\frac{1}{q}} = e^{\frac{1}{p} \log a} e^{\frac{1}{q} \log b} = \exp\left(\frac{1}{p} \log a + \frac{1}{q} \log b\right) \leq \frac{1}{p} \exp(\log a) + \frac{1}{q} \exp(\log b) = \frac{a}{p} + \frac{b}{q}.$$

Die behauptete Form der Young–Ungleichung ergibt sich, wenn man a^p statt a und b^q statt b betrachtet. \square

Satz 8.8 (Hölder–Ungleichung). Seien $p, q > 1$ mit $\frac{1}{p} + \frac{1}{q} = 1$. Dann gilt für alle $x = (x_1, \dots, x_d), y = (y_1, \dots, y_d) \in \mathbb{K}^d$ die Ungleichung

$$\sum_{k=1}^d |x_k y_k| \leq \|x\|_p \cdot \|y\|_q.$$

Beweis. Die Behauptung ist trivial für $x = 0$ ($\Leftrightarrow \|x\|_p = 0$) oder $y = 0$ ($\Leftrightarrow \|y\|_q = 0$). Seien also $\|x\|_p, \|y\|_q > 0$. Die nichtnegativen Zahlen

$$a_k := \frac{|x_k|^p}{\|x\|_p^p}, \quad b_k := \frac{|y_k|^q}{\|y\|_q^q}$$

erfüllen $\sum_{k=1}^d a_k = 1 = \sum_{k=1}^d b_k$. Wir verwenden nun die Young–Ungleichung:

$$\frac{|x_k|}{\|x\|_p} \cdot \frac{|y_k|}{\|y\|_q} = a_k^{\frac{1}{p}} \cdot b_k^{\frac{1}{q}} \leq \frac{a_k}{p} + \frac{b_k}{q}.$$

Durch Summation über k ergibt sich

$$\frac{1}{\|x\|_p \|y\|_q} \sum_{k=1}^d |x_k y_k| \leq \frac{1}{p} + \frac{1}{q} = 1,$$

was die Behauptung impliziert. \square

Die Hölder–Ungleichung gilt auch in dem Grenzfall $p = 1$ und $q = \infty$, denn wir haben

$$\sum_{k=1}^d |x_k y_k| \leq \max_{j=1, \dots, d} |y_j| \sum_{k=1}^d |x_k| = \|x\|_1 \cdot \|y\|_\infty.$$

Der wichtigste Fall der Hölder–Ungleichung $p = q = 2$, er impliziert direkt die sogenannte Cauchy-Schwarz–Ungleichung, die Sie aus der Linearen Algebra kennen. (Die linke Seite dieser Ungleichung ist der Betrag des Skalarproduktes $\langle x, y \rangle$.)

Korollar 8.9 (Cauchy-Schwarz–Ungleichung). Für alle $x, y \in \mathbb{K}^d$ gilt

$$\left| \sum_{k=1}^d x_k y_k \right| \leq \sum_{k=1}^d |x_k y_k| \leq \|x\|_2 \|y\|_2.$$

Nach diesen Ungleichungen sind wir soweit, dass wir die Dreiecksungleichung der Normen $\|\cdot\|_p$, $p \geq 1$, nachprüfen können:

Satz 8.10. Für $p \geq 1$ oder $p = \infty$ gilt für alle $x, y \in \mathbb{K}^d$

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

Die Abbildung $\|\cdot\|_p : \mathbb{K}^d \rightarrow \mathbb{R}_{\geq 0}$ ist also eine Norm.

Beweis. Die Ungleichung ist trivial für $x = 0$, $y = 0$, oder $x + y = 0$. Wir dürfen also $\|x\|_p, \|y\|_p, \|x + y\|_p > 0$ annehmen. Sei zunächst $1 < p < \infty$ und definiere q durch $\frac{1}{p} + \frac{1}{q} = 1$. Dann ist $q > 1$, und es gilt $p = pq - q$ und $p - \frac{p}{q} = 1$.

Wir definieren einen Vektor $z \in \mathbb{K}^d$ durch seine Komponenten $z_k := |x_k + y_k|^{p-1}$. Dann gilt

$$z_k^q = |x_k + y_k|^{q(p-1)} = |x_k + y_k|^p \Rightarrow \|z\|_q = \|x + y\|_p^{\frac{p}{q}}.$$

Mit der Hölder–Ungleichung erhalten wir nun

$$\begin{aligned} \|x + y\|_p^p &= \sum_{k=1}^d |x_k + y_k|^{p-1} |x_k + y_k| = \sum_{k=1}^d |z_k| |x_k + y_k| \leq \sum_{k=1}^d |z_k| |x_k| + \sum_{k=1}^d |z_k| |y_k| \\ &\stackrel{\text{Hö.}}{\leq} \|x\|_p \|z\|_q + \|y\|_p \|z\|_q = (\|x\|_p + \|y\|_p) \cdot \|x + y\|_p^{\frac{p}{q}}. \end{aligned}$$

Also haben wir $\|x + y\|_p = \|x + y\|_p^{p - \frac{p}{q}} \leq \|x\|_p + \|y\|_p$, wie behauptet.

Für die beiden Fälle $p = 1$ und $p = \infty$ ergibt sich die Behauptung sofort aus der Dreiecksungleichung $|x_k + y_k| \leq |x_k| + |y_k|$.

Damit ist die Dreiecksungleichung für alle p -Normen als letzte Normeigenschaft bewiesen. \square

Die Dreiecksungleichung der p -Norm heißt auch *Minkowski–Ungleichung*.

8.3 Konvergenz in metrischen Räumen

Die p -Normen geben uns verschiedene Möglichkeiten, Abstände zwischen Punkten in \mathbb{K}^d zu messen. Wie wir gleich sehen werden, kommt es oft nicht darauf an, welche Norm man verwendet, so dass man eine gut zum Problem passende Norm wählen kann.

Zuerst vergleichen wir die p -Normen untereinander.

Lemma 8.11. Für jedes $p, q \in [1, \infty]$ gibt es Konstanten $c, C > 0$, so dass

$$c\|x\|_p \leq \|x\|_q \leq C\|x\|_p$$

für alle $x \in \mathbb{K}^d$ gilt.

Beweis. Für jedes $p \in [1, \infty]$ gilt

$$\|x\|_\infty = \max_k |x_k| = (\max_k |x_k|^p)^{1/p} \leq \left(\sum_{k=1}^d |x_k|^p \right)^{1/p} = \|x\|_p$$

und

$$\|x\|_p^p = \sum_{k=1}^d |x_k|^p \leq d \max_k |x_k|^p = d \|x\|_\infty^p,$$

also $\|x\|_p \leq d^{1/p} \|x\|_\infty$.

Für beliebiges p, q haben wir also

$$d^{-1/p} \|x\|_p \leq \|x\|_\infty \leq \|x\|_q \leq d^{1/q} \|x\|_\infty \leq d^{1/q} \|x\|_p.$$

□

Die p -Normen unterscheiden sich also nicht sehr stark voneinander. In dieser Situation spricht man auch von äquivalenten Normen:

Definition 8.12. Zwei Normen $\|\cdot\|$ und $\|\!\| \cdot \!\|$ auf einem Vektorraum V heißen *äquivalent*, falls es Konstante $c, C > 0$ gibt, so dass $c\|v\| \leq \|\!\|v\!\| \leq C\|v\|$ für alle $v \in V$ gilt.

Alle p -Normen auf \mathbb{K}^d sind also äquivalent.

Zeigen Sie, dass die oben definierte Äquivalenz eine Äquivalenzrelation auf der Menge aller Normen auf einem Vektorraum ist.

Beispiel 8.13. Hier noch zwei Beispiele von Normen auf Vektorraum $C(I)$ der stetigen Funktionen $f : I \rightarrow \mathbb{R}$, wobei $I \subset \mathbb{R}$ ein kompaktes Intervall ist: Zum einen haben wir die aus Analysis 1 bekannte *Supremumsnorm*,

$$\|f\|_I := \sup_{x \in I} |f(x)|.$$

Ein anderes Beispiel einer Norm auf $C(I)$ ist

$$\|f\|_{L^1} := \int_I |f(x)| dx.$$

Alle Normeigenschaften sind leicht nachprüfbar (Übung). Die Normen $\|\cdot\|_I$ und $\|\cdot\|_{L^1}$ sind nicht äquivalent. Wie wir später sehen werden, ist es kein Zufall, dass wir für ein Beispiel zu nicht äquivalenten Normen unendlichdimensionale Vektorräume wie $C(I)$ betrachten (siehe Satz 10.9).

Genau wie die Betragsfunktion $|\cdot|$ in \mathbb{R} (oder \mathbb{C}) zu einem Konvergenzbegriff für Folgen $a : \mathbb{N} \rightarrow \mathbb{K}$ führt, führt eine Metrik d auf einem metrischen Raum X zu einem Konvergenzbegriff für Folgen $a : \mathbb{N} \rightarrow X$. Erinnerung (Def. 4.3): Eine Folge $(a_n)_{n \in \mathbb{N}}$ in einem metrischen Raum (X, d) konvergiert gegen $x \in X$, falls

$$\lim_{n \rightarrow \infty} d(a_n, x) = 0.$$

Ist der vorliegende metrische Raum ein normierter Vektorraum $X = V$ mit der von der Norm stammenden Metrik $d(x, y) = \|x - y\|$, so heißt Konvergenz $v_n \rightarrow v$ in $(V, \|\cdot\|)$ also $\|v_n - v\| \rightarrow 0$ für $n \rightarrow \infty$. In diesem Fall sagt man auch, dass $(v_n)_n$ in der Norm $\|\cdot\|$ konvergiert.

Erinnern Sie auch, dass ein metrischer Raum *vollständig* heißt, falls jede Cauchyfolge in X konvergiert. Ein normierter Vektorraum, der bzgl der durch die Norm induzierten Metrik vollständig ist, heißt *Banachraum*.

Satz 8.14.

- a) Sei V ein Vektorraum und $\|\cdot\|, \|\cdot\|'$ zwei äquivalente Normen auf V . Dann konvergiert eine Folge $(v_n)_{n \in \mathbb{N}} \subset V$ in der Norm $\|\cdot\|$ genau dann, wenn sie in der Norm $\|\cdot\|'$ konvergiert. In diesem Fall stimmen beide Grenzwerte überein.
- b) Sei nun $V = \mathbb{K}^d$ und $p \in [1, \infty]$. Dann konvergiert eine Folge $(x_n) \subset \mathbb{K}^d$ in der Norm $\|\cdot\|_p$ genau dann, wenn alle ihre Komponentenfolgen $x_{1,n}, \dots, x_{d,n}$ konvergieren. Im Konvergenzfall gilt

$$\lim_{n \rightarrow \infty} x_n = \left(\lim_{n \rightarrow \infty} x_{1,n}, \lim_{n \rightarrow \infty} x_{2,n}, \dots, \lim_{n \rightarrow \infty} x_{d,n} \right).$$

- c) \mathbb{K}^d ist ein Banachraum bzgl jeder p -Norm, $p \in [1, \infty]$ (jede Cauchyfolge konvergiert).

Beweis.

- a) Angenommen, (v_n) konvergiert in der Norm $\|\cdot\|$ gegen einen Grenzwert $v \in V$ d.h. $\|v_n - v\| \rightarrow 0$ für $n \rightarrow \infty$. Da $\|\cdot\|'$ äquivalent zu $\|\cdot\|$ ist, gibt es eine Konstante $C > 0$, so dass $\|v_n - v\|' \leq C\|v_n - v\|$ für alle $n \in \mathbb{N}$ gilt. Also konvergiert (v_n) auch in der Norm $\|\cdot\|'$.

- b) Wir betrachten $p = 1$ und nehmen an, dass (x_n) gegen einen Grenzwert $x \in \mathbb{K}^d$ in $\|\cdot\|_1$ konvergiert, d.h.

$$\|x_n - x\|_1 = \sum_{k=1}^d |x_{k,n} - x_k| \rightarrow 0$$

Dann gilt für jedes $j \in \{1, \dots, d\}$

$$|x_{j,n} - x_j| \leq \sum_{k=1}^d |x_{k,n} - x_k|,$$

also konvergiert auch $|x_{j,n} - x_j|$ gegen Null für $n \rightarrow \infty$. Das heißt, dass aus der Konvergenz in $\|\cdot\|_1$ die Konvergenz aller Komponentenfolgen $(x_{j,n})_{n \in \mathbb{N}}$ folgt.

Falls hingegen alle Komponentenfolgen konvergieren, also $x_{j,n} \rightarrow x_j$ für $n \rightarrow \infty$, so betrachten wir den Vektor $x := (x_1, \dots, x_d)$. Nach den Grenzwertsätzen gilt

$$\lim_{n \rightarrow \infty} \sum_{j=1}^d |x_{j,n} - x_j| = 0,$$

also $\|x_n - x\|_1 \rightarrow 0$. Da alle p -Normen äquivalent sind, folgt die Behauptung auch für $p \neq 1$.

- c) Sei (x_n) eine Cauchyfolge in \mathbb{K}^d , und $j \in \{1, \dots, d\}$. Dann ist wegen $|x_{j,n} - x_{j,m}| \leq \|x_n - x_m\|_p$ auch die Komponentenfolge $(x_{j,n})_n$ eine Cauchyfolge. Da \mathbb{K} vollständig ist, gibt es einen Grenzwert $x_j \in \mathbb{K}$, d.h. $x_{j,n} \rightarrow x_j$ für $n \rightarrow \infty$. Mit b) folgt dann $\|x_n - x\|_p \rightarrow 0$, d.h. (x_n) konvergiert in p -Norm. □

Mit diesem Ergebnis können wir Fragen zur Konvergenz von Folgen in \mathbb{K}^d immer auf Fragen zur Konvergenz von Folgen in \mathbb{K} zurückführen. Wir nennen eine Folge in \mathbb{K}^d *konvergent*, wenn alle ihre Komponentenfolgen konvergieren.

Beispiel 8.15. Wir betrachten die Folge

$$s_n := \left(\frac{1}{n} \cos(n), \frac{1}{n} \sin(n)\right), \quad s_{1,n} = \frac{1}{n} \cos(n), \quad s_{2,n} = \frac{1}{n} \sin(n)$$

in \mathbb{R}^2 , die Punkte auf einer Spirallinie beschreibt. Da sowohl $\frac{1}{n} \cos(n)$ als auch $\frac{1}{n} \sin(n)$ konvergieren (beide mit Grenzwert 0), haben wir $s_n \rightarrow (0, 0)$ bzgl. jeder p -Norm auf \mathbb{R}^2 .

Die Folge

$$r_n := \left(\frac{1}{n} \cos(n), \frac{1}{n} \sin(n), n\right)$$

in \mathbb{R}^3 verläuft auf einer sich immer enger zusammenziehenden Schraubenlinie. Sie konvergiert nicht (bzgl. keiner p -Norm), da die dritte Komponentenfolge $r_{3,n} = n$ nicht konvergiert.

Erinnern Sie sich, dass $(C(I), \|\cdot\|_I)$ ein Banachraum ist. Zeigen Sie, dass $(C^1(I), \|\cdot\|_I)$ kein Banachraum ist. Ist $(C^1(I), \|\cdot\|_{C^1})$ mit $\|f\|_{C^1} := \|f\|_I + \|f'\|_I$ ein Banachraum?

Ein wichtiger Satz über Konvergenz in metrischen Räumen ist der folgende Banach'sche Fixpunktsatz.

Satz 8.16 (Banach'scher Fixpunktsatz). Sei (X, d) ein vollständiger metrischer Raum und $T : X \rightarrow X$ eine strikte Kontraktion, d.h. es gebe $0 < q < 1$ so, dass

$$d(T(x), T(y)) \leq q d(x, y), \quad x, y \in X,$$

gilt.

- T hat genau einen Fixpunkt, d.h. es existiert genau ein $x^* \in X$ mit der Eigenschaft $T(x^*) = x^*$.
- Sei $x_0 \in X$ beliebig und definiere eine Folge (x_n) rekursiv durch $x_n := T(x_{n-1})$, $n \in \mathbb{N}$. Dann konvergiert $(x_n)_n$ gegen den Fixpunkt x^* von T .

Beweis. Wir betrachten die in b) rekursiv definierte Folge (x_n) . Nach Voraussetzung gilt

$$d(x_{n+1}, x_n) = d(T(x_n), T(x_{n-1})) \leq q d(x_n, x_{n-1}).$$

Iterieren wir dieses Argument, so ergibt sich

$$d(x_{n+1}, x_n) \leq q^n d(x_1, x_0).$$

Für beliebige $k, n \in \mathbb{N}$ erhalten wir

$$\begin{aligned} d(x_{n+k}, x_n) &\leq d(x_{n+k}, x_{n+k-1}) + d(x_{n+k-1}, x_{n+k-2}) + \dots + d(x_{n+1}, x_n) \\ &= \sum_{j=0}^{k-1} d(x_{n+j+1}, x_{n+j}) \\ &\leq \sum_{j=0}^{k-1} q^{n+j} d(x_1, x_0) \\ &\leq q^n \sum_{j=0}^{\infty} q^j \cdot d(x_1, x_0) \\ &= \frac{q^n}{1-q} \cdot d(x_1, x_0), \end{aligned}$$

wobei wir die geometrische Reihe und $|q| < 1$ verwendet haben. Nun sehen wir $d(x_{n+k}, x_n) \rightarrow 0$ für $n \rightarrow \infty$, für beliebiges $k \in \mathbb{N}$. Also ist $(x_n)_{n \in \mathbb{N}}$ eine Cauchyfolge in X . Da X nach Voraussetzung vollständig ist, konvergiert diese Folge, d.h. es gibt $x^* \in X$ mit $\lim_{n \rightarrow \infty} x_n = x^*$.

Wir zeigen nun, dass x^* ein Fixpunkt von T ist:

$$\begin{aligned} 0 &\leq d(x^*, T(x^*)) \\ &\leq d(x^*, x_{n+1}) + d(x_{n+1}, T(x^*)) \\ &= d(x^*, x_{n+1}) + d(T(x_n), T(x^*)) \\ &\leq d(x^*, x_{n+1}) + q d(x_n, x^*). \end{aligned}$$

Für $n \rightarrow \infty$ geht die rechte Seite gegen Null (wie oben gezeigt). Also gilt $d(x^*, T(x^*)) = 0$, was wegen der Definitheit der Metrik $x^* = T(x^*)$ impliziert. Der Grenzwert x^* ist also tatsächlich ein Fixpunkt von T .

Nun zur Eindeutigkeit: Angenommen, x^* und \tilde{x} sind Fixpunkte von T . Dann gilt

$$d(x^*, \tilde{x}) = d(T(x^*), T(\tilde{x})) \leq q d(x^*, \tilde{x}).$$

Falls $d(x^*, \tilde{x}) > 0$, so können wir diese Ungleichung durch die positive Zahl $d(x^*, \tilde{x}) > 0$ teilen und erhalten $1 \leq q$, ein Widerspruch zu unserer Annahme $q < 1$. Also gilt $d(x^*, \tilde{x}) = 0$, d.h. $x^* = \tilde{x}$, und T hat genau einen Fixpunkt.

Insbesondere konvergiert die rekursiv definierte Folge (x_n) unabhängig von dem beliebig vorgegebenen Startwert x_0 immer gegen den gleichen Grenzwert x^* . \square

Der Banachsche Fixpunktsatz hat eine Vielzahl von wichtigen Anwendungen die Sie im Lauf Ihres Studiums kennenlernen werden. Beachten Sie, dass die Vollständigkeit von X eine wichtige Voraussetzung ist, ohne die der Fixpunktsatz nicht gilt.

Der Banachsche Fixpunktsatz gilt insbesondere in Banachräumen, also vollständigen metrischen Räumen, in denen die Metrik $d(x, y) = \|x - y\|$ von einer Norm stammt; insbesondere also in \mathbb{R}^d und $C(I)$. Wir geben zwei typische Beispiele von Anwendungen dieses Satzes.

Beispiel 8.17 (Das Newton-Verfahren). Das Newton-Verfahren ist eine Methode zur Bestimmung/Approximation von Nullstellen von differenzierbaren Funktionen, die vor allem numerisch von Bedeutung ist.

Die grundlegende Idee ist die folgende: Wir betrachten eine zweimal stetig differenzierbare Funktion $f : \mathbb{R} \rightarrow \mathbb{R}$, die in einer (unbekannten) Stelle x^* eine Nullstelle hat, $f(x^*) = 0$. Um x^* zu finden, "raten" wir einen Startwert x_0 , der idealerweise schon recht nah an x^* liegt. Dann verbessern wir x_0 folgendermaßen: Wir betrachten die Tangente t_{x_0} an den Graphen von f , gegeben durch

$$t_{x_0}(x) = f(x_0) + f'(x_0) \cdot (x - x_0)$$

(Taylorentwicklung bis 1. Ordnung). Die Nullstelle dieser Tangente ist

$$x_1 = x_0 - \frac{f(x_0)}{f'(x_0)},$$

wobei wir $f'(x_0) \neq 0$ voraussetzen. Dieses Verfahren wird dann iteriert, wie im Beweis des Banach'schen Fixpunktsatzes.

Wir betrachten also die Abbildung $T(x) := x - \frac{f(x)}{f'(x)}$ und sehen, dass Fixpunkte von T (also $T(x) = x$) genau Nullstellen von f sind; vorausgesetzt, dass die Ableitung von f nicht verschwindet. Erfüllt T auch die entscheidende Kontraktionseigenschaft? Dazu betrachten wir

$$|T(x) - T(y)| = \left| \int_y^x T'(z) dz \right| \leq q \cdot |x - y|, \quad q := \sup_z |T'(z)|.$$

Die Kontraktionseigenschaft gilt also, falls $\sup_z |T'(z)| < 1$. Eine einfache Rechnung zeigt $T'(z) = \frac{f(z)f''(z)}{f'(z)^2}$. An einer Nullstelle $z = x^*$ von f gilt $T'(z) = 0$, aufgrund der Stetigkeit von T' gilt also in einem hinreichend kleinen Intervall um x^* auch $|T'(z)| \leq q < 1$ für vorgegebenes $0 < q < 1$.

Zusammengefasst: Ist f zweimal stetig differenzierbar mit nichtverschwindender Ableitung mit $q := \|f'\|_{\mathbb{R}} = \sup_z |f'(z)| < 1$ und hat bei x^* eine Nullstelle, so ist $T(x) := x - \frac{f(x)}{f'(x)}$ wohldefiniert und strikt kontrahierend. Da \mathbb{R} bzgl. $|\cdot|$ vollständig ist, können wir den Banachschen Fixpunktsatz verwenden und folgern, dass die iterierte Folge $T^n(x_0)$, für beliebiges $x_0 \in \mathbb{R}$, gegen die eindeutige Nullstelle x^* von f konvergiert.

Schaut man sich den Beweis des Satzes genauer an (\rightarrow Übung), so findet man auch die Fehlerabschätzung

$$|x_n - x^*| \leq \frac{q^n}{1 - q} |x_1 - x_0|$$

für den Abstand der n -ten Iteration des Verfahrens zur Nullstelle x^* .

Häufig braucht man auch eine "lokale Version", in der $f : I \rightarrow \mathbb{R}$ nur auf einem Intervall $I = [a, b]$ definiert ist. Dann kann man zeigen:

- Angenommen, $f : [a, b] \rightarrow \mathbb{R}$ ist zweimal stetig differenzierbar und es existiert $x^* \in (a, b)$ mit $f(x^*) = 0$ und $f'(x^*) \neq 0$.

Dann gibt es ein (hinreichend kleines) abgeschlossenes Intervall J , das x^* enthält und so dass die Abbildung $T : J \rightarrow J$, $T(x) = x - \frac{f(x)}{f'(x)}$, eine strikte Kontraktion ist. Also konvergiert für jedes $x_0 \in J$ die Folge $x_{n+1} := T(x_n)$ gegen x^* .

Hier diskutieren wir eine weitere Anwendung des Banachschen Fixpunktsatzes, die die Nützlichkeit dieses Satzes hervorhebt. Wir betrachten für $f \in C(I)$ (mit $I = [0, 1]$) und $\lambda \in \mathbb{R}$ die Gleichung

$$f(x) = x^2 + \lambda \int_0^1 e^{xy} f(y) dy, \quad x \in I.$$

Gibt es eine Lösung dieser Gleichung, also eine Funktion $f \in C(I)$ mit der gewünschten Eigenschaft? Eine solche Integralgleichung für f ist üblicherweise nicht einfach zu lösen. Um trotzdem etwas über Lösungen sagen zu können, schreiben wir die Frage in eine Fixpunktfrage um. Wir betrachten die Abbildung

$$T : C(I) \rightarrow C(I), \quad T(f)(x) = x^2 + \lambda \int_0^1 e^{xy} f(y) dy$$

(d.h., T bildet Funktionen auf Funktionen ab) und schätzen mit Hilfe des Mittelwertsatzes der Integralrechnung und der Supremumsnorm $\|\cdot\|_I$ ab: ($x_1, x_2 \in I$)

$$|T(f)(x_1) - T(f)(x_2)| \leq |\lambda| \|f\|_I \left| \int_0^1 (e^{x_1 y} - e^{x_2 y}) dy \right| = \|f\|_I |\lambda| \left(\frac{e^{x_1} - 1}{x_1} - \frac{e^{x_2} - 1}{x_2} \right).$$

Da $x \mapsto \frac{e^x - 1}{x}$ stetig ist, sehen wir $T(f)(x_1) - T(f)(x_2) \rightarrow 0$ für $x_1 \rightarrow x_2$. Also ist $T(f)$ stetig, d.h. T ist wirklich eine Abbildung $C(I) \rightarrow C(I)$.

Wir betrachten nun die Kontraktionseigenschaft: Für $f_1, f_2 \in C(I)$ haben wir

$$\begin{aligned} \|T(f_1) - T(f_2)\|_I &= |\lambda| \sup_{x \in I} \left| \int_0^1 (f_1(x) - f_2(x)) e^{xy} dy \right| \\ &\stackrel{\text{MWS-I}}{\leq} |\lambda| \|f_1 - f_2\|_I \cdot \sup_{x \in I} \left| \int_0^1 e^{xy} dy \right| \\ &\leq e|\lambda| \cdot \|f_1 - f_2\|_I. \end{aligned}$$

Falls $q := e|\lambda| < 1$, d.h. falls $|\lambda| < e^{-1}$, ist T also eine Kontraktion. Da $C(I)$ ein vollständiger metrischer Raum bzgl der Supremumsnorm $\|\cdot\|_I$ ist, können wir den Banachschen Fixpunktsatz verwenden und folgern, dass es genau eine stetige Funktion gibt, so dass $f(x) = x^2 + \lambda \int_0^1 e^{xy} f(y) dy$ gilt.

8.4 Die Topologie des \mathbb{K}^d

In diesem Abschnitt wollen wir einige Begriffe aus Analysis 1 (Abschnitt 4.1 und 4.2) wiederholen und erweitern.

Beginnen wir mit Def. 4.4:

- In einem metrischen Raum (X, d) heißt für $x \in X$ und $\varepsilon > 0$

$$U_\varepsilon(x) := \{y \in X : d(x, y) < \varepsilon\}$$

die ε -Umgebung von x .

- Eine Teilmenge $O \subset X$ heißt *offen*, falls es für jedes $x \in O$ ein $\varepsilon > 0$ gibt, so dass $U_\varepsilon(x) \subset O$.
- Eine Teilmenge $A \subset X$ heißt *abgeschlossen*, falls $X \setminus A$ offen ist.
- Eine Teilmenge $U \subset X$ heißt *Umgebung von einem Punkt* $x \in X$, falls $U_\varepsilon(x) \subset U$ für ein $\varepsilon > 0$.

Wir betrachten nun speziell \mathbb{K}^d (also \mathbb{R}^d bzw. \mathbb{C}^d) mit einer p -Norm $\|\cdot\|_p$, $p \in [1, \infty]$. Jede p -Norm definiert ε -Umgebungen $U_\varepsilon^{\|\cdot\|_p}(x)$ von Punkten $x \in \mathbb{K}^d$, die von p abhängen – in Aufgabe P2.2 haben Sie diese Umgebungen für $\varepsilon = 1$ und verschiedene p skizziert.

Nichtdestotrotz haben wir folgenden Satz:

Satz 8.18. Sei V ein Vektorraum und $\|\cdot\|, \|\cdot\|'$ zwei äquivalente Normen auf V (Z.B. $V = \mathbb{R}^d$ mit $\|\cdot\|_{p_1}, \|\cdot\|_{p_2}$, $p_1, p_2 \geq 1$). Eine Teilmenge $A \subset V$ ist *offen* (*abgeschlossen*) bzgl der Metrik $d(v, w) = \|v - w\|$ genau dann wenn sie *offen* (*abgeschlossen*) bzgl der Metrik $d'(v, w) = \|v - w\|'$ ist.

Beweis. Da die beiden Normen äquivalent sind, gibt es Konstanten $c, C > 0$ mit $c\|v\| \leq \|v\|' \leq C\|v\|$ für alle $v \in V$. Das impliziert

$$U_\varepsilon^{\|\cdot\|'}(v) = \{w \in V : \|v - w\|' < \varepsilon\} \supset \{w \in V : \|v - w\| < \varepsilon/C\} = U_{\varepsilon/C}^{\|\cdot\|}(v),$$

und analog $U_\varepsilon^{\|\cdot\|}(v) \supset U_{c\varepsilon}^{\|\cdot\|'}(v)$. Wir sehen also: Jede ε -Umgebung eines Punktes v bzgl der einen Norm enthält eine ε -Umgebung (mit skaliertem ε) von v bzgl der anderen Norm.

Angenommen, $A \subset V$ ist offen bzgl $\|\cdot\|$, d.h. für jedes $a \in A$ gibt es $\varepsilon > 0$, so dass $U_\varepsilon^{\|\cdot\|}(a) \subset A$. Dann enthält A wegen $U_{c\varepsilon}^{\|\cdot\|'}(a) \subset U_\varepsilon^{\|\cdot\|}(a) \subset A$ auch eine ε -Umgebung von a bzgl $\|\cdot\|'$. Also ist A auch offen bzgl $\|\cdot\|'$. Völlig analog sieht man: $(A \text{ offen bzgl } \|\cdot\|') \Rightarrow (A \text{ offen bzgl } \|\cdot\|)$.

Eine Menge $A \subset V$ ist abgeschlossen bzgl $\|\cdot\|$ genau dann wenn $V \setminus A$ bzgl $\|\cdot\|$ offen ist, was äquivalent zu $V \setminus A$ bzgl $\|\cdot\|'$ offen, dh A bzgl $\|\cdot\|'$ abgeschlossen ist. \square

Die Begriffe “offen” und “abgeschlossen” ändern sich also nicht, wenn man von einer Norm zu einer äquivalenten Norm übergeht.

An dieser Stelle sei noch einmal an die grundlegenden Eigenschaften der Familie der offenen Mengen erinnert (Satz 4.6: \emptyset und der ganze Raum sind offen und abgeschlossen, endliche Schnitte von offenen Mengen sind offen, beliebige Vereinigungen von offenen Mengen sind offen).

Anschaulich gesprochen sind die offenen Teilmengen die “ohne Randpunkte”. Um das genauer zu fassen, führen wir einige neue Begriffe ein.

Definition 8.19. Sei (X, d) ein metrischer Raum und $M \subset X$ eine Teilmenge.

- a) Das *Innere von M* ist $M^\circ := \bigcup_{O \subset M, O \text{ offen}} O$. Dies ist die größte offene Menge, die in M enthalten ist.
- b) Der *Abschluss von M* ist $\overline{M} := \bigcap_{A \supset M, A \text{ abgeschlossen}} A$. Dies ist die kleinste abgeschlossene Menge, die M enthält.
- c) Der *Rand von M* ist $\partial M := \overline{M} \setminus M^\circ$.

Beachten Sie, dass M° stets offen ist (als Vereinigung offener Mengen) und $M^\circ \subset M$, und dass \overline{M} stets abgeschlossen ist (als Schnitt abgeschlossener Mengen) und $M \subset \overline{M}$. Aus der Definition ist unmittelbar klar:

- M offen $\Leftrightarrow M = M^\circ$.
- M abgeschlossen $\Leftrightarrow M = \overline{M}$.

Beispiel 8.20. a) Wir betrachten $M = [0, 1) \subset \mathbb{R}$ mit der üblichen Metrik und behaupten $M^\circ = (0, 1)$, $\overline{M} = [0, 1]$, $\partial M = \{0, 1\}$.

Zur ersten Behauptung betrachten wir die offenen Mengen (Intervalle) $(\varepsilon, 1 - \varepsilon)$ für $0 < \varepsilon < 1$, die alle in M enthalten sind. Also

$$M^\circ \supset \bigcup_{0 < \varepsilon < 1} (\varepsilon, 1 - \varepsilon) = (0, 1).$$

Angenommen, $0 \in M^\circ$. Dann $M^\circ = M$, dh dann wäre $M = [0, 1)$ offen. Aber keine ε -Umgebung von 0 ist in M enthalten, also ist dies nicht der Fall. Wir

haben also $M^o = (0, 1)$.

Zur zweiten Behauptung betrachten wir die abgeschlossenen Mengen (Intervalle) $[-\varepsilon, 1 + \varepsilon]$ für $\varepsilon > 0$, die alle M enthalten. Also

$$\overline{M} \subset \bigcap_{\varepsilon > 0} [-\varepsilon, 1 + \varepsilon] = [0, 1].$$

Da $M = [0, 1)$, muss entweder $\overline{M} = [0, 1]$ oder $\overline{M} = [0, 1)$ gelten. Aber $[0, 1)$ ist nicht abgeschlossen (das Komplement $(-\infty, 0) \cup [1, \infty)$ ist nicht offen wegen dem Randpunkt 1). Also $\overline{M} = [0, 1]$.

Nun sehen wir auch $\partial M = [0, 1] \setminus (0, 1) = \{0, 1\}$.

Dieses Beispiel zeigt auch, dass der Rand von M nicht Teilmenge von M sein muss.

- b) Wir betrachten $M = \{4\} \subset \mathbb{R}$ mit der üblichen Metrik. Dann ist $M^o = \emptyset$, da $\{4\}$ keine offenen Mengen außer \emptyset enthält.

Einige Charakterisierungen von Abschluss, Innerem und Rand sind im nächsten Satz zusammengefasst.

Satz 8.21. Sei (X, d) ein metrischer Raum und $M \subset X$ eine Teilmenge.

- $\overline{M} = M \cup \partial M = M^o \cup \partial M$.
- $x \in M^o \Leftrightarrow M$ enthält eine ε -Umgebung von x .
- $x \in \overline{M} \Leftrightarrow$ jede ε -Umgebung von x schneidet M .
- $x \in \partial M \Leftrightarrow$ jede ε -Umgebung von x schneidet M und $X \setminus M$.

Beweis. a) Wegen $M^o \subset M \subset \overline{M}$ gilt per Definition

$$\begin{aligned} \overline{M} &= (\overline{M} \cap M^o) \cup (\overline{M} \setminus M^o) = M^o \cup \partial M \subset M \cup \partial M \\ &= M \cup (\overline{M} \setminus M^o) \subset \overline{M} \cup (\overline{M} \setminus M^o) = \overline{M}, \end{aligned}$$

also sind alle diese Mengen gleich.

b) Falls M eine ε -Umgebung von x enthält, $U_\varepsilon(x) \subset M$, so gilt $x \in U_\varepsilon(x) \subset M^o$ nach Definition von M^o (denn $U_\varepsilon(x)$ ist offen). Falls $x \in M^o$, so liegt x nach Definition von M^o in einer offenen Menge O , die in M enthalten ist. Also gibt es $\varepsilon > 0$ mit $U_\varepsilon(x) \subset O \subset M$, d.h. M enthält eine ε -Umgebung von x .

c) Wir zeigen die äquivalente Aussage: $x \notin \overline{M} \Leftrightarrow$ (es gibt eine ε -Umgebung $U_\varepsilon(x)$ mit $U_\varepsilon(x) \cap M = \emptyset$). Falls $x \notin \overline{M}$, so liegt x in der offenen Menge $X \setminus \overline{M}$. Also gibt es eine Umgebung $U_\varepsilon(x)$ mit $U_\varepsilon(x) \subset X \setminus \overline{M} \subset X \setminus M \Rightarrow U_\varepsilon(x) \cap M = \emptyset$. Existiert andererseits eine ε -Umgebung $U_\varepsilon(x)$ mit $U_\varepsilon(x) \cap M = \emptyset$, so haben wir $M \subset X \setminus U_\varepsilon(x)$. Da $X \setminus U_\varepsilon(x)$ abgeschlossen ist, folgt nach Definition des Abschlusses $\overline{M} \subset X \setminus U_\varepsilon(x)$, insbesondere also $x \notin \overline{M}$.

d) Per Definition gilt $\partial M = \overline{M} \setminus M^\circ$. Also besteht ∂M aus all den Punkten $x \in X$, für die jede ε -Umgebung M schneidet, aber M keine Umgebung von x enthält, d.h. jede ε -Umgebung schneidet auch $X \setminus M$. \square

Beispiel 8.22. Wir betrachten $M = \mathbb{Q} \subset \mathbb{R} = X$ mit der üblichen Metrik. Dann gilt $\mathbb{Q}^\circ = \emptyset$, da jede ε -Umgebung einer rationalen Zahl auch irrationale Zahlen enthält, $\overline{\mathbb{Q}} = \mathbb{R}$, da jede ε -Umgebung einer reellen Zahl \mathbb{Q} schneidet, und $\partial \mathbb{Q} = \mathbb{R}$.

Im Allgemeinen nennt man Mengen M mit $\overline{M} = X$ *dicht*.

 Finden Sie alle Teilmengen $M \subset \mathbb{R}$ mit $\partial M = \emptyset$.

Weitere Beispiele finden sich in den Übungen.

Wir können den Abschluss \overline{M} von M auch mittels Grenzwerten von Folgen charakterisieren:

Satz 8.23. Sei (X, d) ein metrischer Raum und $M \subset X$ eine Teilmenge. Dann gilt

$$\overline{M} = \{x \in X : \text{es gibt eine Folge } (m_n)_n \in M \text{ mit } x = \lim_{n \rightarrow \infty} m_n\}.$$

Beweis. Sei $x \in \overline{M}$. Dann gilt $U_{1/n}(x) \cap M \neq \emptyset$ für jedes $n \in \mathbb{N}$ nach Satz 8.21 c). Wähle zu jedem n ein $x_n \in U_{1/n}(x) \cap M$. Dann liegt die Folge (x_n) in M , und wegen $d(x_n, x) < \frac{1}{n}$ konvergiert sie gegen x . In der behaupteten Gleichheit gilt also “ \subset ”.

Um “ \supset ” zu zeigen, zeigen wir die äquivalente Aussage:

$$x \notin \overline{M} \Rightarrow \exists (x_n)_n \in M : x_n \rightarrow x.$$

Sei also $x \notin \overline{M}$, d.h. x liegt in der offenen Menge $\overline{M}^c = X \setminus \overline{M}$. Nach Satz 8.21 b) gibt es ein $\varepsilon > 0$ mit $U_\varepsilon(x) \subset \overline{M}^c$, also $U_\varepsilon(x) \cap M = \emptyset$. Also gibt es keinen Punkt in M , der Abstand $< \varepsilon$ von x hat. Deshalb gibt es keine in M liegende Folge, die gegen x konvergiert. \square

Dieses nützliche Resultat ergibt sofort einen weiteren Beweis des Satzes 4.17 aus Analysis 1:

- Eine Teilmenge $M \subset X$ eines metrischen Raumes X ist abgeschlossen genau dann, wenn sie die Eigenschaft hat, dass für jede konvergente Folge $(m_n)_n \subset M \subset X$ der Grenzwert $\lim_n m_n \in M$ auch in M liegt.

Eine weitere Konsequenz ist:

- Sei (X, d) ein vollständiger metrischer Raum und $M \subset X$ eine Teilmenge. Dann ist auch $(M, d|_{M \times M})$ ein metrischer Raum. Dieser Raum ist genau dann vollständig, wenn $M \subset X$ abgeschlossen ist.

Der Beweis erfolgt in den Übungen.

Beispiel 8.24. Hier einige weitere Beispiele von Innerem, Abschluss, und Rand. Die Beweise der unten stehenden Behauptungen erfolgen in den Übungen.

a) In $X = \mathbb{R}$ betrachten wir die Teilmenge $M = \{\frac{1}{n} : n \in \mathbb{N}\}$. Dann gilt

$$\overline{M} = M \cup \{0\}, \quad M^\circ = \emptyset, \quad \partial M = \overline{M}.$$

b) Sei $(V, \|\cdot\|)$ ein normierter Vektorraum und $B := \{v \in V : \|v\| < 1\}$ die offene Einheitskugel in V . Dann gilt

$$B^\circ = B, \quad \overline{B} = \{v \in V : \|v\| \leq 1\}, \quad \partial B = \{v \in V : \|v\| = 1\}.$$

c) Im metrischen Raum $(\mathbb{R}^2, \|\cdot\|_2)$ betrachten wir das Quadrat

$$Q := \{(x, y) \in \mathbb{R}^2 : 0 < x \leq 1, 0 \leq y < 1\} = (0, 1] \times [0, 1).$$

Dann gilt

$$\begin{aligned} \overline{Q} &= [0, 1] \times [0, 1], & Q^\circ &= (0, 1) \times (0, 1), \\ \partial Q &= (\{0, 1\} \times [0, 1]) \cup ([0, 1] \times \{0, 1\}). \end{aligned}$$

9 Kompaktheit

Dieses Kapitel widmet sich dem Konzept einer *kompakten* Teilmenge eines metrischen Raums.

9.1 Motivation und Definition

In der Analysis 1 haben wir gesehen, dass Teilmengen $M \subset \mathbb{R}$, die sowohl *abgeschlossen* als auch *beschränkt* sind, gute Eigenschaften haben: Ist $(a_n) \subset A$ eine Folge in A , so hat (a_n) wegen der Beschränktheit von A eine konvergente Teilfolge (Bolzano-Weierstraß), und da A abgeschlossen ist, liegt der Grenzwert dieser Teilfolge auch in A . Wir sagen hierzu kurz: Jede Folge in A hat eine in A konvergente Teilfolge.

Wenn wir diese Begriffe auf einen allgemeinen metrischen Raum (X, d) ausdehnen wollen, liegt es nahe, Beschränktheit folgendermaßen zu definieren:

Definition 9.1. Sei (X, d) ein metrischer Raum und $M \subset X$ eine Teilmenge. Wir definieren den *Durchmesser* von M als

$$\text{diam}M := \sup_{x,y \in M} d(x,y) \in [0, \infty],$$

falls M nicht leer ist, und $\text{diam}\emptyset := 0$.

M heißt *beschränkt*, falls $\text{diam}M < \infty$.

Um den Begriff der Kompaktheit zu motivieren, betrachten wir das folgende Beispiel.

Beispiel 9.2. Wir betrachten den Raum $X := C_b(\mathbb{R})$ der stetigen beschränkten Funktionen $\mathbb{R} \rightarrow \mathbb{R}$ (d.h. $f \in C_b(\mathbb{R})$ genau dann, wenn f stetig ist und es $C > 0$ gibt, so dass $|f(x)| \leq C$ für alle $x \in \mathbb{R}$). Dies ist ein Banachraum bzgl der Supremumsnorm $\|f\|_{\mathbb{R}} := \sup_{x \in \mathbb{R}} |f(x)|$ (siehe Analysis 1).

Die abgeschlossene Einheitskugel

$$B := \{f \in C_b(\mathbb{R}) : \|f\| \leq 1\}$$

ist abgeschlossen (siehe Übung) und beschränkt: Für $f, g \in B$ gilt $\|f - g\|_{\mathbb{R}} \leq \|f\|_{\mathbb{R}} + \|g\|_{\mathbb{R}} \leq 2$ (tatsächlich gilt $\text{diam}B = 2$).

Wir betrachten nun die "Zackenfunktion"

$$f(x) := \begin{cases} 1 - |x| & |x| \leq 1 \\ 0 & \text{sonst} \end{cases}$$

und die Folge $f_n(x) := f(x - 2n)$. Beachten Sie, dass der Graph von f_n aus dem Graphen von f durch Verschiebung um $2n$ nach rechts hervorgeht.

Für beliebiges $n \neq m$ gibt es also keinen Punkt $x \in \mathbb{R}$, an dem beide Funktionen $f_n(x)$ und $f_m(x)$ ungleich Null sind (Skizze!). Das bedeutet

$$\|f_n - f_m\|_{\mathbb{R}} = \sup_{x \in \mathbb{R}} |f_n(x) - f_m(x)| = \sup_{x \in \mathbb{R}} |f_n(x)| = 1.$$

Wir sehen also, dass f_n immer Abstand 1 zu f_m (für $n \neq m$) hat bzgl der Supremumsnorm. Die Folge $(f_n)_n$ konvergiert also nicht (sie ist quasi “maximal nicht-Cauchy”) und hat auch keine konvergente Teilfolge.

Die Lehre, die wir aus diesem Beispiel ziehen können, ist die folgende: In allgemeinen metrischen Räumen haben die Mengen, die abgeschlossen und beschränkt sind, nicht vergleichbar gute Eigenschaften wie die abgeschlossenen und beschränkten Teilmengen von \mathbb{R} .

Wir werden nun einen neuen Begriff einführen, der die Eigenschaft “abgeschlossen und beschränkt” in einer Art und Weise auf allgemeine metrische Räume verallgemeinert, die die guten Eigenschaften von kompakten Intervallen $[a, b] \subset \mathbb{R}$ erhält. Die Definition ist relativ abstrakt.

Definition 9.3. Sei (X, d) ein metrischer Raum.

- a) Eine *offene Überdeckung* einer Teilmenge $M \subset X$ ist eine Familie $\{O_i : i \in I\}$ von offenen Teilmengen $O_i \subset X$, wobei $i \in I$ durch eine beliebige Indexmenge I läuft, so dass

$$M \subset \bigcup_{i \in I} O_i.$$

- b) Eine Teilmenge $M \subset X$ heißt *kompakt*, falls jede offene Überdeckung $\{O_i : i \in I\}$ von M eine endliche Teilüberdeckung besitzt, d.h. es gibt $i_1, \dots, i_n \in I$ (mit $n \in \mathbb{N}$), so dass $M \subset \bigcup_{j=1}^n O_{i_j}$ (die endlich vielen Mengen O_{i_1}, \dots, O_{i_n} überdecken bereits M).

Dies ist eine komplizierte Definition, wahrscheinlich die komplizierteste, die Sie bisher kennengelernt haben. Wir werden Sie uns also Stück für Stück ansehen und beginnen mit einigen Bemerkungen.

Jede Teilmenge $M \subset X$ besitzt eine offene Überdeckung, zB die Überdeckung durch die einzelne offene Menge $O_1 = X$ oder die Überdeckung durch *alle* offenen Mengen $O \subset X$. Eine offene Überdeckung $(O_i)_{i \in I}$ ist oft “sehr redundant”, d.h. Sie können oft viele der Mengen O_i weglassen und die übrigbleibenden Mengen in $(O_i)_{i \in I}$ überdecken M immer noch. Kompaktheit besagt, dass Sie unabhängig davon, mit welcher offenen Überdeckung Sie beginnen, immer alle O_i bis auf endlich viele weglassen können und trotzdem noch eine offene Überdeckung behalten.

Beispiel 9.4.

- Jede endliche Menge $M := \{x_1, \dots, x_n\} \subset X$ ist kompakt: Ist $(O_i)_{i \in I}$ eine offene Überdeckung von M , so gibt es für jedes x_k ein O_{i_k} , so dass $x_k \in O_{i_k}$ (sonst wäre es keine Überdeckung). Die endliche Menge O_{i_1}, \dots, O_{i_n} enthält alle x_k , also ganz M . Somit ist es eine endliche Teilüberdeckung von M .
- Die Menge $M := \{\frac{1}{n} : n \in \mathbb{N}\} \subset \mathbb{R}$ ist nicht kompakt. Zum Beweis betrachten wir die offenen Mengen $O_\varepsilon := (\varepsilon, 2)$, wobei $\varepsilon \in I := \mathbb{R}_+$. Dies ist eine offene Überdeckung von M , denn zu jedem $n \in \mathbb{N}$ existiert ein $\varepsilon > 0$ mit $\frac{1}{n} \in (\varepsilon, 2)$. Aber wenn wir nur endlich viele der Mengen $(\varepsilon, 2)$ auswählen, also

$(\varepsilon_1, 2), \dots, (\varepsilon_r, 2)$ mit $r \in \mathbb{N}$, dann gibt es ein kleinstes $\varepsilon = \min\{\varepsilon_1, \dots, \varepsilon_r\}$. Für $n > \frac{1}{\varepsilon}$ haben wir dann $\frac{1}{n} \notin (\varepsilon_i, 2)$ für alle $i \in \{1, \dots, r\}$, d.h. $(\varepsilon_1, 2), \dots, (\varepsilon_r, 2)$ ist keine Überdeckung von M .

- Sei (x_n) eine konvergente Folge in X , mit Grenzwert $x := \lim_n x_n$. Dann ist $K := \{x_n : n \in \mathbb{N}\} \cup \{x\}$ kompakt. Um dies zu beweisen, wählen wir eine beliebige offene Überdeckung $(O_i)_{i \in I}$ von K . Da dies eine Überdeckung ist, gibt es $i_0 \in I$, so dass der Grenzwert x in O_{i_0} enthalten ist. Nun ist O_{i_0} offen, enthält also eine ε -Umgebung $U_\varepsilon(x)$. Da die Folge (x_n) gegen x konvergiert, gibt es $N \in \mathbb{N}$ so dass $x_n \in U_\varepsilon(x)$ für alle $n \geq N$ (Definition von Konvergenz). Wir wählen also O_{i_0} als Element unserer endlichen Teilüberdeckung, damit haben wir schon “fast alle” (alle bis auf endlich viele) Folgenglieder abgedeckt. Für die verbleibenden endlich vielen Folgenglieder x_1, \dots, x_N wählen wir endlich viele offene Mengen O_{i_k} mit $x_k \in O_{i_k}$ wie im ersten Beispiel aus und haben damit unsere endliche Teilüberdeckung gefunden.

- Zeigen Sie, dass $(0, 1) \subset \mathbb{R}$ nicht kompakt ist. Tipp: Verwenden Sie eine offene Überdeckung, die ähnlich wie in dem zweiten Beispiel in 9.4 definiert ist.
- Zeigen Sie, dass $\mathbb{R} \subset \mathbb{R}$ nicht kompakt ist. Tipp: Verwenden Sie eine offene Überdeckung aus “kleinen” offenen Mengen, so dass endlich viele nicht zur Überdeckung von ganz \mathbb{R} ausreichen. Zum Beispiel $((n, n + 2))_{n \in \mathbb{Z}}$.

Es kann insbesondere passieren, dass in einem metrischen Raum (X, d) die Menge X selbst kompakt ist. In diesem Fall sprechen wir von einem *kompakten metrischen Raum*.

9.2 Eigenschaften kompakter Mengen

Wir machen uns nun daran, den Begriff der Kompaktheit besser zu verstehen, indem wir ihn mit vertrauteren Begriffen vergleichen.

Satz 9.5. *Jede kompakte Teilmenge M eines metrischen Raumes X ist abgeschlossen und beschränkt.*

Beweis. Falls $M = \emptyset$ ist die Aussage richtig. Wir dürfen also $M \neq \emptyset$ annehmen.

Beschränktheit: Sei $M \subset X$ kompakt. Wir beginnen mit der folgenden offenen Überdeckung von M : Wir wählen einen Punkt $m \in M$ und betrachten die offenen Mengen $U_r(m)$ mit beliebigem Radius $r > 0$. Da jeder Punkt $m' \in M$ endlichen Abstand $d(m, m') < \infty$ zu m hat, liegt $m' \in U_r(m)$ für $r > d(m, m')$. Also ist $(U_r(m))_{r>0}$ eine offene Überdeckung von M . Da M kompakt ist, gibt es eine endliche Teilüberdeckung $U_{r_1}(m), \dots, U_{r_n}(m)$, d.h. $M \subset U_{r_1}(m) \cup \dots \cup U_{r_n}(m)$. Sei $r := \max\{r_1, \dots, r_n\}$. Dann gilt $M \subset U_r(m)$ und damit $\text{diam} M \leq 2r < \infty$. Also ist M beschränkt.

Abgeschlossenheit: Wir zeigen, dass M kompakt impliziert, dass $M^c = X \setminus M$ offen ist.

Sei also M kompakt und $z \notin M$ (so ein z gibt es, falls $M \neq X$. Falls $M = X$ ist die Abgeschlossenheit von M klar). Wir betrachten die offenen “Ring”-Mengen $R_r := \{x \in X : d(x, z) > r\}$. Die Menge $(R_r)_{r>0}$ überdeckt $X \setminus \{z\}$, also insbesondere M (da $z \notin M$).

Wegen der Kompaktheit von M gibt es eine endliche Teilüberdeckung $M \subset R_{r_1} \cup \dots \cup R_{r_n}$. Sei $r := \min\{r_1, \dots, r_n\} > 0$. Da $R_a \subset R_b$ für $a \geq b$, gilt $M \subset R_r$. Aber dies bedeutet $U_r(z) \subset M^c$ (Beachten Sie $U_r(z) \cap R_r(z) = \emptyset$). Also gibt es für jeden Punkt $z \in M^c$ eine ganz in M^c liegende offene Umgebung $U_r(z)$, d.h. M^c ist offen. \square

Beispiel 9.6. $\{x \in \mathbb{R}^d : \|x\|_2 < 1\} \subset \mathbb{R}^d$ ist nicht kompakt, da nicht abgeschlossen. $\{x \in \mathbb{R}^d : x_1 \geq 0\} \subset \mathbb{R}^d$ ist nicht kompakt, da nicht beschränkt. $\{x : \|x\|_2 \leq 1\} \subset \mathbb{R}^d$ ist abgeschlossen und beschränkt, könnte also kompakt sein. Später sehen wir, dass diese Menge tatsächlich kompakt ist.

Die Umkehrung des Satzes, also die Aussage “ M abgeschlossen und beschränkt $\Rightarrow M$ kompakt” gilt in allgemeinen metrischen Räumen *nicht*. Wir werden das später explizit sehen. Immerhin haben wir folgende Vererbungseigenschaft von Kompaktheit auf abgeschlossene Teilmengen:

Lemma 9.7. Sei K eine kompakte Teilmenge eines metrischen Raums X und $A \subset X$ eine abgeschlossene Teilmenge mit $A \subset K$. Dann ist auch A kompakt.

Beweis. Sei $\{O_i : i \in I\}$ eine offene Überdeckung von A . Da A abgeschlossen ist, ist das Komplement A^c offen. Wir betrachten nun das vergrößerte System $\{O_i : i \in I\} \cup A^c$ offener Mengen. Wegen

$$K \subset X = A^c \cup A \subset A^c \cup \bigcup_{i \in I} O_i$$

ist $\{O_i : i \in I\} \cup A^c$ eine offene Überdeckung von K . Da K kompakt ist, gibt es eine offene Teilüberdeckung, d.h.

$$A \subset K \subset A^c \cup O_{i_1} \cup \dots \cup O_{i_n}$$

für geeignete O_{i_j} . Also haben wir $A \subset O_{i_1} \cup \dots \cup O_{i_n}$, d.h. wir haben eine endliche Teilüberdeckung von A gefunden. \square

Ein entscheidender Schritt zum Verständnis von Kompaktheit einer Menge K ist der folgende Satz 9.10, der besagt, dass Kompaktheit von K die in der Einleitung zu diesem Kapitel diskutierte Eigenschaft hat, dass jede Folge in K eine in K konvergente Teilfolge hat, wie wir es von den abgeschlossenen und beschränkten Teilmengen von \mathbb{R} kennen.

Als Vorbereitung erinnern wir an die Definition eines Häufungspunktes, die wir aus Analysis 1 bereits für Folgen in \mathbb{R} und \mathbb{C} kennen.

Definition 9.8. Sei (X, d) ein metrischer Raum. Ein Punkt $x \in X$ heißt *Häufungspunkt* einer Folge $(x_n)_{n \in \mathbb{N}} \subset X$, falls es eine Teilfolge $(x_{n_k})_k$ gibt, die gegen x konvergiert.

Diese Definition liest sich identisch zu Def. 3.25, nur der Kontext (allgemeiner metrischer Raum X vs. \mathbb{K}) ist anders. Es verwundert deshalb nicht, dass wir Satz 3.26 a) ganz analog übertragen können. Auch der Beweis ist ganz analog zu dem in Analysis 1.

Lemma 9.9. Sei (X, d) ein metrischer Raum und $(x_n)_n \subset X$ eine Folge. Dann ist ein Punkt $x \in X$ ein Häufungspunkt von $(x_n)_n$ genau dann wenn für jedes $\varepsilon > 0$ die Menge $\{n \in \mathbb{N} : d(x_n, x) < \varepsilon\}$ unendlich ist.

Beweis. Angenommen, x ist Häufungspunkt der Folge. Dann gibt es eine Teilfolge $(x_{n_k})_k$ mit $d(x_{n_k}, x) \rightarrow 0$ für $k \rightarrow \infty$. Für jedes $\varepsilon > 0$ gibt es also $K \in \mathbb{N}$, so dass $d(x_{n_k}, x) < \varepsilon$ für alle $k \geq K$. Also enthält $\{n \in \mathbb{N} : d(x_n, x) < \varepsilon\}$ alle n_k mit $k \geq K$, ist also unendlich.

Angenommen, $I_\varepsilon := \{n \in \mathbb{N} : d(x_n, x) < \varepsilon\}$ ist unendlich für alle $\varepsilon > 0$. Wir betrachten $\varepsilon := \frac{1}{k}$. Wir wählen Indizes $n_1 \in I_1, n_2 \in I_{1/2}, n_k \in I_{1/k}$, so, dass $n_{k+1} > n_k$ für alle k . Dies ist möglich, da die $I_{1/k}$ alle unendlich sind. Die Teilfolge $(x_{n_k})_k$ erfüllt dann $d(x_{n_k}, x) < \frac{1}{k}$, konvergiert also gegen x . Also ist x ein Häufungspunkt. \square

Satz 9.10. Eine Teilmenge M eines metrischen Raumes ist genau dann kompakt, wenn jede Folge in M eine in M konvergente Teilfolge hat.

Beweis. Wir zeigen zuerst: Gibt es in M eine Folge (x_n) ohne in M konvergente Teilfolge, so ist M nicht kompakt. Das heißt also, dass diese Folge keinen Häufungspunkt in M hat. Nach obigem Lemma gibt es also für jedes $x \in M$ ein $\varepsilon(x) > 0$ so, dass $\{n \in \mathbb{N} : d(x_n, x) < \varepsilon(x)\}$ endlich ist. (Dies drückt aus, dass kein x Häufungspunkt der Folge ist.)

Wegen $m \in U_{\varepsilon(m)}(m)$ ist die Familie $\{U_{\varepsilon(m)}(m) : m \in M\}$ eine offene Überdeckung von M . Wäre M kompakt, würde es eine endliche Teilüberdeckung geben, dh

$$M \subset U_{\varepsilon(m_1)}(m_1) \cup \dots \cup U_{\varepsilon(m_r)}(m_r).$$

Aber dies führt auf einen Widerspruch, nämlich

$$\begin{aligned} \mathbb{N} = \{n \in \mathbb{N} : x_n \in M\} &\subset \left\{ n \in \mathbb{N} : x_n \in \bigcup_{k=1}^r U_{\varepsilon(m_k)}(m_k) \right\} \\ &= \bigcup_{k=1}^r \{n \in \mathbb{N} : x_n \in U_{\varepsilon(m_k)}(m_k)\}. \end{aligned}$$

Auf der rechten Seite haben wir eine endliche Vereinigung von endlichen Menge, was endlich ist. Dies kann nicht die unendliche Menge \mathbb{N} enthalten. Also ist M nicht kompakt.

Nun nehmen wir an, dass jede Folge in M eine in M konvergente Teilfolge besitzt und zeigen, dass dann M kompakt ist. Sei also $(O_i)_{i \in I}$ eine beliebige offene Überdeckung von M .

Zuerst zeigen wir folgende Aussage:

$$\exists r > 0 : \forall x \in M : \exists i \in I : U_r(x) \subset O_i.$$

Wir führen einen Widerspruchsbeweis, nehmen also die Negation dieser Aussage an: Für alle $r > 0$ (insbesondere für alle r der Form $r = \frac{1}{n}, n \in \mathbb{N}$) gibt es ein $x = x_n \in M$, so dass $U_{1/n}(x_n) \not\subset O_i$ für alle $i \in I$. Dies gibt uns eine Folge $(x_n)_n \subset M$, die per Annahme eine in M konvergente Teilfolge hat. Sei x der Grenzwert dieser Teilfolge $(x_{n_k})_k$. Da $x \in M$ und $(O_i)_i$ eine offene Überdeckung ist, gibt es $i_0 \in I$ und $\varepsilon > 0$ so, dass $U_\varepsilon(x) \subset O_{i_0}$.

Dies führt auf den gewünschten Widerspruch: Da $x_{n_k} \rightarrow k$, gibt es $K \in \mathbb{N}$ so, dass $d(x, x_{n_k}) < \varepsilon/2$ für $k \geq K$. Wählen wir nun noch k groß genug, so dass $\frac{1}{k} < \varepsilon/2$, so ergibt sich per Dreiecksungleichung $U_{1/k}(x_{n_k}) \subset U_{\varepsilon/2}(x_{n_k}) \subset U_{\varepsilon}(x) \subset O_{i_0}$, ein Widerspruch.

Als nächstes zeigen wir die Aussage: Für jedes $r > 0$ gibt es endlich viele Punkte $x_1, \dots, x_n \in M$ mit $M \subset U_r(x_1) \cup \dots \cup U_r(x_n)$. Wir führen wieder einen Widerspruchsbeweis und nehmen deshalb an, dass es einen Radius $r > 0$ gibt, so dass $M \not\subset U_r(x_1) \cup \dots \cup U_r(x_n)$ für jede endliche Wahl x_1, \dots, x_n von Punkten in M . Dann können wir folgendermaßen eine Folge $(x_m)_m$ definieren: Wir wählen $x_0 \in M$ beliebig, und x_m rekursiv gemäß der Bedingung $x_m \in M \setminus \bigcup_{j=0}^{m-1} U_r(x_j)$. Diese Konstruktion ist möglich, da $M \setminus \bigcup_{j=0}^{m-1} U_r(x_j)$ nach obiger Schlussfolgerung nie leer ist.

Die so konstruierte Folge hat per Konstruktion die Eigenschaft, dass $x_m \notin U_r(x_u)$ für alle $u < m$. Also ist $(x_m)_m$ keine Cauchyfolge, und mehr noch, keine Teilfolge von $(x_m)_m$ ist eine Cauchyfolge. Aber nach Voraussetzung über M hat jede Folge in M eine in M konvergente Folge, die also insbesondere eine Cauchyfolge ist. Das beendet den zweiten Widerspruchsbeweis.

Nun führen wir beide Argumente zusammen und zeigen, dass M kompakt ist. Wir betrachten den in Teil 1 des Beweises gefundenen Radius $r > 0$ und wenden darauf Teil 2 an: Es gibt also endlich viele Punkte x_1, \dots, x_n , so dass $M \subset U_r(x_1) \cup \dots \cup U_r(x_n)$. Nun gibt es nach Teil 1 zu jedem x_j ein O_{i_j} mit $U_r(x_j) \subset O_{i_j}$. Also $M \subset O_{i_1} \cup \dots \cup O_{i_n}$; wir haben eine endliche Teilüberdeckung gefunden und folgern, dass M kompakt ist. \square

9.3 Kompakte Teilmengen von \mathbb{K}^d

Nachdem wir den Begriff der Kompaktheit für allgemeine metrische Räume kennengelernt haben, wenden wir uns nun der konkreteren Frage nach den kompakten Teilmengen von \mathbb{K}^d (wieder mit $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$) zu.

Wir wissen bereits, dass kompakte Mengen stets abgeschlossen und beschränkt sind. Für die folgenden Argumente ist es besonders bequem, sogenannte *Quader* in \mathbb{R}^d zu betrachten. Per Definition soll ein Quader eine Menge der Form

$$Q = [a_1, b_1] \times \dots \times [a_d, b_d]$$

mit $a_l < b_l$, $l = 1, \dots, d$ sein. Für $d = 1$ sind die Quader also genau die abgeschlossenen und beschränkten Intervalle, für $d = 2$ abgeschlossene und beschränkte Rechtecke in der Ebene \mathbb{R}^2 , etc. Offenbar sind Quader in beliebiger Dimension abgeschlossen und beschränkt. Beschränktheit bezieht sich hier auf eine beliebige p -Norm – Beachten Sie, dass es keine Rolle spielt, welche dieser äquivalenten Normen wir betrachten, da sich der Durchmesser von Q nur um einen Faktor ändern würde. Zum Beispiel bezüglich $\|\cdot\|_{\infty}$ haben wir $\text{diam} Q = \max\{|b_l - a_l| : 1 \leq l \leq d\} < \infty$.

Die Quader könnten also kompakt sein. Um zu sehen, dass das tatsächlich der Fall ist, verallgemeinern wir zuerst den Satz von Bolzano-Weierstraß auf höhere Dimensionen. Wir gehen jetzt analog zu Analysis 1 vor, als wir den Satz von Bolzano-Weierstraß von \mathbb{R} nach \mathbb{C} ausgedehnt haben.

Lemma 9.11. Jede beschränkte Folge $(x_n)_n \subset \mathbb{R}^d$ hat eine konvergente Teilfolge.

Beweis. Sei $(x_n)_n \subset \mathbb{R}^d$ eine beschränkte Folge, also $\|x_n\|_\infty \leq C$ für alle $n \in \mathbb{N}$. Dann sind die Komponentenfolgen $x_{n,k}$ ($= k$ -te Komponente von x_n , $k = 1, \dots, d$) ebenfalls beschränkt, denn $|x_{n,k}| \leq \max_k |x_{n,k}| = \|x_n\|_\infty \leq C$. Nach dem Satz von Bolzano-Weierstraß für \mathbb{R}^1 hat die Folge der ersten Komponenten, also $(x_{n,1})_n$, eine konvergente Teilfolge $(x_{n_{k_1},1})_{k_1}$. Die Folge $(x_{n_{k_1}})_k$ hat also die Eigenschaft, dass die Folge ihrer ersten Komponenten konvergiert. Die Folge der zweiten Komponenten $(x_{n_{k_1},2})_{k_1}$ ist beschränkt, hat also eine konvergente Teilfolge $(x_{n_{k_{1l}},2})_{l_1}$. Für die Folge $(x_{n_{k_{1l}}})_{l_1}$ konvergiert also die erste Komponentenfolge (als Teilfolge einer konvergenten Folge) und die zweite Komponentenfolge. Nach d Schritten finden wir so eine Teilfolge von (x_n) , für die alle d Komponentenfolgen konvergieren. Das ist äquivalent zur Konvergenz dieser Folge in \mathbb{R}^d . \square

Korollar 9.12. In \mathbb{R}^d ist jeder Quader Q kompakt.

Beweis. Sei $(x_n)_n \subset Q$ eine Folge in $Q = [a_1, b_1] \times \dots \times [a_d, b_d]$. Da Q beschränkt ist, ist diese Folge beschränkt; hat also eine konvergente Teilfolge. Der Grenzwert g dieser Teilfolge erfüllt $a_j \leq g_j \leq b_j$ für alle $j = 1, \dots, d$, da die Komponentenfolgen $(x_{n_{k_j}})_k \in Q$ diese Eigenschaft haben. Also liegt g in Q . Mit Satz 9.10 folgt nun, dass Q kompakt ist. \square

Quader sind keine besonderen, sondern nur bequeme Beispiele von kompakten Mengen in \mathbb{R}^d . Wir kommen jetzt zum Satz von Heine-Borel, der uns sehr viel mehr kompakte Teilmengen von \mathbb{K}^d an die Hand gibt.

Satz 9.13 (Heine-Borel (Spezialfall)). Sei $d \in \mathbb{N}$ und $p \in [1, \infty]$. Eine Teilmenge K des normierten Raums $(\mathbb{K}^d, \|\cdot\|_p)$ ist genau dann kompakt, wenn sie abgeschlossen und beschränkt ist.

Beweis. In jedem metrischen Raum sind kompakte Mengen abgeschlossen und beschränkt. Wir müssen also nur eine Implikation zeigen.

Ist $K \subset \mathbb{R}^d$ abgeschlossen und beschränkt, so finden wir einen hinreichend großen Quader Q mit $K \subset Q$. Da wir wissen, dass Q kompakt ist, folgt die Kompaktheit von K mit Lemma 9.9.

Im komplexen Fall $K \subset \mathbb{C}^d$ können wir $\mathbb{C}^d \cong \mathbb{R}^{2d}$ identifizieren und erhalten so die gleiche Aussage. \square

Der Satz von Heine-Borel konkretisiert die kompakten Mengen in \mathbb{K}^d , gilt aber nicht in allgemeinen metrischen Räumen oder allgemeinen Banachräumen. Das Beispiel zu Anfang dieses Kapitels demonstriert das: Die Menge $B := \{f \in C_b(\mathbb{R}) : \|f\|_\mathbb{R} \leq 1\}$ ist abgeschlossen und beschränkt in $(C_b(\mathbb{R}), \|\cdot\|_\mathbb{R})$. Sie ist aber nicht kompakt, da sie wie in Beispiel 9.2 gezeigt eine Folge enthält, die keine in B konvergente Teilfolge hat.

Beispiel 9.14. Folgende Teilmengen von \mathbb{K}^d sind kompakt: $\overline{U_\varepsilon^{\|\cdot\|_p}(x)} = \{y \in \mathbb{K}^d : \|x - y\|_p \leq \varepsilon\}$ (da abgeschlossen und beschränkt) und endliche Vereinigungen dieser Mengen.

Sei $K \subset \mathbb{R}$ kompakt und nicht leer. Zeigen Sie: K besitzt ein Minimum und ein Maximum.
Tipp: Wegen Heine-Borel hat K ein Supremum und Infimum. Argumentieren Sie $\sup K \in K$.

Es stellt sich nun die interessante Frage, was über die kompakten Teilmengen von typischen unendlichdimensionalen Banachräumen, insbesondere Funktionenräumen, gesagt werden kann. Was sind zum Beispiel die kompakten Teilmengen von $(C(Q), \|\cdot\|_Q)$, des Raums der stetigen Funktionen $f : Q \rightarrow \mathbb{K}$ auf einem Quader, bzgl. Supremumsnorm $\|f\|_Q = \sup_{x \in Q} |f(x)|$? Der Satz von Arzelà-Ascoli beantwortet diese Frage und ist nützlich, wenn zB gleichmäßig konvergente Teilfolgen einer gegebenen Folge von Funktionen gesucht werden. Sie werden diesen Satz in der Funktionalanalysis kennenlernen. Weitere Untersuchungen von kompakten Mengen (in einem verallgemeinerten Kontext) werden insbesondere in der Topologie-Vorlesung besprochen.

9.4 Stetige Funktionen auf kompakten Mengen

Wir schließen das Kapitel über Kompaktheit mit einigen Betrachtungen zu stetigen Funktionen auf kompakten Mengen ab. Erinnern wir uns zunächst, was eine stetige Funktion zwischen zwei metrischen Räumen ist (Analysis 1).

Dazu betrachten wir zwei metrische Räume (X, d_X) , (Y, d_Y) . Eine Abbildung $f : X \rightarrow Y$ heißt *stetig* in einem Punkt $x_0 \in X$, falls eine der folgenden äquivalenten Bedingungen erfüllt ist:

- $\forall \varepsilon > 0 : \exists \delta > 0 : d_X(x, x_0) < \delta \Rightarrow d_Y(f(x), f(x_0)) < \varepsilon$.
- Für jede Folge (a_n) in X mit $a_n \rightarrow x_0$ gilt $f(a_n) \rightarrow f(x_0)$.
- Für jede offene Menge $O \subset Y$ ist das Urbild $f^{-1}(O) \subset X$ offen.

Der Beweis des folgenden Satz demonstriert, wie hilfreich die abstrakten Konzepte von Kompaktheit und Stetigkeit sind. Kurz formuliert sagt er “stetige Bilder kompakter Mengen sind kompakt”.

Satz 9.15. *Seien X, Y metrische Räume und $f : X \rightarrow Y$ stetig. Dann ist für jedes Kompaktum $K \subset X$ auch das Bild $f(K) \subset Y$ kompakt.*

Beweis. Sei $(O_i)_{i \in I}$ eine offene Überdeckung von $f(K)$. Dann sind die Mengen $f^{-1}(O_i) \subset X$ offen (weil f stetig ist) und bilden eine Überdeckung von K , da jedes $k \in K$ Urbild eines $f(k) \in f(K)$ ist. Da K kompakt ist, gibt es eine endliche Teilüberdeckung $K \subset f^{-1}(O_1) \cup \dots \cup f^{-1}(O_n)$. Also gilt

$$f(K) \subset f(f^{-1}(O_1)) \cup \dots \cup f(f^{-1}(O_n)) \subset O_1 \cup \dots \cup O_n.$$

Damit haben wir eine endliche Teilüberdeckung gefunden, also ist $f(K)$ kompakt. \square

Beispiel 9.16. Kompaktheit überträgt sich durch stetige Funktionen von Mengen auf ihre Bilder, aber nicht anders herum! Nehmen Sie zum Beispiel die Nullfunktion $f : \mathbb{R} \rightarrow \mathbb{R}$, $f(x) = 0$. Dann ist $\{0\} \subset \mathbb{R}$ kompakt, aber das Urbild $f^{-1}(\{0\}) = \mathbb{R}$ nicht.

Eine weitere oft nützliche Eigenschaft von stetigen Funktionen $f : X \rightarrow \mathbb{R}$ ist, dass sie auf Kompakta Maximum und Minimum annehmen. Dies ist eine Verallgemeinerung von Satz 4.27.

Satz 9.17. Sei X ein metrischer Raum, $K \subset X$ eine nicht leere kompakte Menge, und $f : K \rightarrow \mathbb{R}$ stetig. Dann nimmt f Minimum und Maximum an, d.h. es existieren $a, b \in K$ mit $f(a) = \min_{k \in K} f(k)$ und $f(b) = \max_{k \in K} f(k)$.

Beweis. Nach Satz 9.15 ist $f(K) \subset \mathbb{R}$ kompakt, besitzt also ein Maximum und Minimum nach der Übung auf Seite 32. \square

Beispiel 9.18. Sei $f : [0, 1] \times [0, 1] \rightarrow \mathbb{R}$, $f(x, y) = e^{x^2 y} (x^2 - y^3) + 2 \sin(\cos(x^2 y^4))$. Diese Funktion ist stetig (dazu mehr im nächsten Kapitel). Da der Quader $[0, 1] \times [0, 1] \subset \mathbb{R}^2$ kompakt ist, hat f ein Minimum und Maximum.

Noch eine Verallgemeinerung von einem Ergebnis der Analysis 1 (versteckt im Beweis von Lemma 6.7) ist der nächste Satz.

Satz 9.19. Sei K ein kompakter metrischer Raum, Y ein beliebiger metrischer Raum und $f : K \rightarrow Y$ stetig. Dann ist f sogar gleichmäßig stetig^a, d.h. zu jedem $\varepsilon > 0$ gibt es ein $\delta > 0$, so dass für alle $x, y \in K$ mit $d(x, y) < \delta$ stets $d(f(x), f(y)) < \varepsilon$ gilt.

^aDer Unterschied zur bloßen Stetigkeit besteht darin, dass δ unabhängig von x gewählt werden kann.

Beweis. Sei $\varepsilon > 0$ vorgegeben. Aufgrund der Stetigkeit von f existiert dann für jedes $x \in K$ ein $\delta_x > 0$ mit $d(x, y) < 2\delta_x \Rightarrow d(f(x), f(y)) < \frac{\varepsilon}{2}$. Die Mengen $(U_{\delta_x}(x))_{x \in K}$ bilden eine offene Überdeckung von K . Aufgrund der Kompaktheit von K gibt es also eine endliche Teilüberdeckung $U_{\delta_{x_1}}(x_1), \dots, U_{\delta_{x_n}}(x_n)$ von K . Das gesuchte δ aus der Definition der gleichmäßigen Stetigkeit wählen wir nun als $\delta := \min\{\delta_{x_1}, \dots, \delta_{x_n}\} > 0$.

Zum Nachweis der gleichmäßigen Stetigkeit seien nun $x, y \in K$ mit $d(x, y) < \delta$. Da $U_{\delta_{x_1}}(x_1), \dots, U_{\delta_{x_n}}(x_n)$ eine Überdeckung von K ist, liegt y in einer dieser offenen Mengen, d.h. $d(y, x_l) < \delta_{x_l}$ für ein $l \in \{1, \dots, n\}$. Damit erhalten wir

$$d(x, x_l) \leq d(x, y) + d(y, x_l) < \delta + \delta_{x_l} < 2\delta_{x_l},$$

und somit

$$d(f(x), f(y)) \leq d(f(x), f(x_l)) + d(f(x_l), f(y)) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

\square

10 Stetige Abbildungen auf endlichdimensionalen Vektorräumen

Ein Hauptthema der Analysis 2 ist das Studium differenzierbarer Abbildungen $f : \mathbb{R}^d \rightarrow \mathbb{R}^k$. Was “differenzierbar” hier heißen soll, muss noch geklärt werden, wird aber ein Begriff sein, der stärker ist als die uns bereits bekannte Stetigkeit. Wir gehen deshalb in diesem Abschnitt etwas näher auf stetige Abbildungen $\mathbb{R}^d \rightarrow \mathbb{R}^k$ ein.

10.1 Stetige Funktionen mit mehreren Variablen und mehreren Komponenten

Im Folgenden betrachten wir eine Teilmenge $D \subset \mathbb{R}^d$ und Funktionen $f : D \rightarrow \mathbb{R}^k$, wobei die beiden Dimensionen d, k beliebige natürliche Zahlen sind. Eine solche Abbildung hat die Form

$$D \ni x = (x_1, \dots, x_d) \mapsto f(x) = \begin{pmatrix} f_1(x_1, \dots, x_d) \\ f_2(x_1, \dots, x_d) \\ \vdots \\ f_k(x_1, \dots, x_d) \end{pmatrix} \in \mathbb{R}^k.$$

Wir bezeichnen die Komponenten $f(x)_j =: f_j(x)$ als die *Komponentenfunktionen von f* . Jede Komponentenfunktion f_j hängt i.A. von allen Variablen x_1, \dots, x_d ab. Der Definitionsbereich von f_j ist ebenfalls D , d.h. wir haben “skalare” (zahlen- statt vektorwertige) Abbildungen $f_j : D \rightarrow \mathbb{R}$.

Beispiel 10.1. $f : \mathbb{R}^2 \rightarrow \mathbb{R}^3$, $f(x, y) = (\sin(x^2y), x - y + 5, -y)$. Hier ist $D = \mathbb{R}^2$ und $f_1(x, y) = \sin(x^2y)$, $f_2(x, y) = x - y + 5$, $f_3(x, y) = -y$.

Unser allgemeiner Begriff von Stetigkeit von Abbildungen zwischen metrischen Räumen spezialisiert sich auf Abbildungen $f : D \subset \mathbb{R}^d \rightarrow \mathbb{R}^k$. Wir versehen dabei sowohl den Definitionsbereich $D \subset \mathbb{R}^d$ als auch den Bildbereich \mathbb{R}^k mit einer der p -Normen $\|\cdot\|_p$. Folgendes Lemma besagt, dass der sich ergebende Stetigkeitsbegriff nicht von der Wahl dieser Normen abhängt.

Lemma 10.2. *Eine Funktion $f : D \subset \mathbb{R}^d \rightarrow \mathbb{R}^k$ ist stetig bzgl der Normen $\|\cdot\|_p$ auf D und $\|\cdot\|_q$ auf \mathbb{R}^k (mit $p, q \in [1, \infty]$) genau dann wenn alle ihre Komponentenfunktionen $f_j : D \rightarrow \mathbb{R}$ stetig sind (bzgl der üblichen Metrik auf \mathbb{R}).*

Beweis. Die Funktion f ist stetig genau dann, wenn für jedes $x \in D$ und jede Folge $(x_n)_n \subset D$ mit $x_n \rightarrow x$ (bzgl der Norm $\|\cdot\|_p$) auch $f(x_n) \rightarrow f(x)$ (bzgl der Norm $\|\cdot\|_q$) gilt. Nach Satz 8.14 ist das äquivalent zu $f_j(x_n) \rightarrow f_j(x)$ für alle $j \in \{1, \dots, k\}$. Diese Eigenschaft wiederum ist äquivalent zur Stetigkeit aller Komponentenfunktionen f_j . \square

Dieses Lemma besagt also, dass die Stetigkeit einer vektorwertigen Funktion einfach Komponente für Komponente entschieden werden kann. Anders sieht es allerdings mit den mehreren *Variablen* x_1, \dots, x_d aus, wie wir jetzt erklären (für eine skalare Funktion, die wir uns als Komponente einer vektorwertigen Funktion denken können).

Ist $f : D \subset \mathbb{R}^d \rightarrow \mathbb{R}$ stetig, so können wir f in Abhängigkeit von nur einer Variable betrachten. Dazu betrachten wir einen inneren Punkt $x \in D^\circ$ (wobei wir $D^\circ \neq \emptyset$ annehmen) und betrachten die Funktion

$$t \mapsto f(x + te_j) = f(x_1, x_2, \dots, x_j + t, \dots, x_d).$$

Hier und im Folgenden bezeichnet e_j stets den j -ten Basisvektor der Standardbasis von \mathbb{K}^d , d.h. $e_1 = (1, 0, \dots, 0)$, $e_2 = (0, 1, 0, \dots, 0)$, etc.

Da x ein innerer Punkt des Definitionsbereichs ist, liegt eine ε -Umgebung von x in D . Wir wählen diese ε -Umgebung bzgl der Norm $\|\cdot\|_\infty$. Dann sehen wir, dass obige Funktion $t \mapsto f(x + te_j)$ mindestens für $-\varepsilon < t < \varepsilon$ wohldefiniert ist. Weiterhin gilt aufgrund der Stetigkeit von f

$$\lim_{t \rightarrow 0} f(x + te_j) = f(\lim_{t \rightarrow 0} (x + te_j)) = f(x) = f(x + 0e_j),$$

also ist $t \mapsto f(x + te_j)$ auch stetig. Wir nennen den Grenzwert $\lim_{t \rightarrow 0} f(x + te_j)$ den Grenzwert von f in 0 entlang der j -ten Koordinatenachse, bzw den Grenzwert von f in 0 in Richtung von e_j .

Andererseits ist die Stetigkeit der $t \mapsto f(x + te_j)$ nicht ausreichend für die Stetigkeit von f :

Beispiel 10.3. Betrachte die Funktion

$$f : \mathbb{R}^2 \rightarrow \mathbb{R}, \quad f(x, y) := \begin{cases} 1 & x = y, \quad (x, y) \neq (0, 0) \\ 0 & \text{sonst} \end{cases}.$$

Dann ist f nicht stetig in $(0, 0)$ (betrachte die Folge $(1/n, 1/n)$, für die $f(1/n, 1/n) = 1$, aber $f(0, 0) = 0$ gilt). Nichtsdestotrotz existieren die Grenzwerte von f in $(0, 0)$ in Richtung der Koordinatenachsen: Es ist $f(t, 0) = 0$ und $f(0, t) = 0$ für alle $t \in \mathbb{R}$, also insbesondere $f((0, 0) + te_j) \rightarrow 0$ für $t \rightarrow 0$ und $j \in \{1, 2\}$.

Völlig analog können wir statt Grenzwerten von f in einem inneren Punkt $x \in D^\circ$ entlang eines Basisvektors e_j auch den Grenzwert in Richtung eines beliebigen Vektors $v \in \mathbb{R}^d \setminus \{0\}$ betrachten, also den *Richtungsgrenzwert*

$$\lim_{t \rightarrow 0} f(x + t \cdot v).$$

Die Funktion $t \mapsto f(x + t \cdot v)$ beschreibt f eingeschränkt auf die Gerade $\{x + t \cdot v : t \in \mathbb{R}\}$ bzw. auf den Teil dieser Gerade, der im Definitionsbereich von f liegt.

Für stetiges f existiert dieser Grenzwert und stimmt mit $f(x)$ überein. Die Existenz aller Richtungsgrenzwerte $\lim_{t \rightarrow 0} f(x + t \cdot v)$ (für alle $v \neq 0$, nicht nur für die Standardbasisvektoren e_j) ist aber auch nicht genug, um die Stetigkeit von f in x zu folgern, wie das nächste Beispiel zeigt.

Beispiel 10.4. Wir betrachten die drei Funktionen $f, g, h : \mathbb{R}^2 \rightarrow \mathbb{R}$

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

$$h(x, y) := \begin{cases} \frac{x^2y}{x^4+y^2} & (x, y) \neq (0, 0) \\ 0 & (x, y) = (0, 0) \end{cases},$$

Die Frage ist, ob f, g, h bei $(0, 0)$ stetig sind bzw wie ihre Richtungsgrenzwerte aussehen. Wir betrachten zuerst Richtungsgrenzwerte, d.h. einen Vektor $v = (v_1, v_2) \neq (0, 0)$ und (für $t \neq 0$)

$$f(t \cdot v) = f(tv_1, tv_2) = \frac{t^2v_1v_2}{t^2v_1^2 + t^2v_2^2} = \frac{v_1v_2}{v_1^2 + v_2^2}.$$

Dies hängt nicht von t ab, ist allerdings meist $\neq 0$ und hängt von v ab. Für $v_1v_2 \neq 0$ gilt $\lim_{t \rightarrow 0} f(t \cdot v) \neq 0 = f(0, 0)$, also ist f nicht stetig.

Für g und h ergibt die gleiche Betrachtung

$$g(t \cdot v) = g(tv_1, tv_2) = \frac{t^3v_1^2v_2}{t^2v_1^2 + t^2v_2^2} = t \cdot \frac{v_1^2v_2}{v_1^2 + v_2^2} \rightarrow 0 \quad \text{für } t \rightarrow 0$$

$$h(t \cdot v) = h(tv_1, tv_2) = \frac{t^3v_1^2v_2}{t^4v_1^4 + t^2v_2^2} = t \cdot \frac{v_1^2v_2}{t^2v_1^4 + v_2^2} \rightarrow 0 \quad \text{für } t \rightarrow 0.$$

Hier sind alle Richtungsgrenzwerte identisch, und stimmen alle mit $g(0, 0) = h(0, 0) = 0$ überein. Trotzdem ist g stetig in $(0, 0)$ aber h nicht.

Zur Nichtstetigkeit von h betrachten wir die Folge $(\frac{1}{n}, \frac{1}{n^2})$, die für $n \rightarrow \infty$ gegen $(0, 0)$ konvergiert. Aber

$$h(\frac{1}{n}, \frac{1}{n^2}) = \frac{n^{-4}}{n^{-4} + n^{-4}} = \frac{1}{2},$$

was für $n \rightarrow \infty$ nicht gegen $h(0, 0) = 0$ konvergiert. Also ist h in 0 nicht stetig.

Zur Stetigkeit von g in $(0, 0)$ betrachten wir eine *beliebige* gegen den Ursprung konvergente Folge $(x_n, y_n) \rightarrow 0$ mit $(x_n, y_n) \neq (0, 0)$. Dann haben wir

$$|g(x_n, y_n)| = \frac{|y_n| \cdot x_n^2}{x_n^2 + y_n^2} \leq |y_n| \frac{x_n^2 + y_n^2}{x_n^2 + y_n^2} = |y_n| \rightarrow 0,$$

also ist g im Ursprung stetig.

Es empfiehlt sich, einige 3D-plots von diesen und ähnlichen Funktionen zu betrachten, um ein Gefühl für die Thematik zu bekommen. Solche Plots lassen sich zB mit geogebra <https://www.geogebra.org/3d> einfach erstellen.

Wir merken uns also, dass Existenz von Richtungsgrenzwerten an einem Punkt kein hinreichendes Kriterium für Stetigkeit in diesem Punkt ist – auch nicht, wenn die Richtungsgrenzwerte für alle Richtungen übereinstimmen. Dieses Phänomen wird uns später im Zusammen-

hang mit Richtungsableitungen erneut begegnen.

Nun einige Beispiele von stetigen Funktion $\mathbb{R}^d \rightarrow \mathbb{R}$.

Beispiel 10.5. Eine *Polynomfunktion* $f : \mathbb{R}^d \rightarrow \mathbb{R}$ ist eine Funktion der Art

$$f(x_1, \dots, x_d) = \sum_{p_1, \dots, p_d} c_{p_1, \dots, p_d} x_1^{p_1} \cdots x_d^{p_d},$$

wobei alle p_i in \mathbb{N}_0 liegen und die Summe endlich ist.

Polynomfunktionen sind stetig. Um dies zu beweisen, betrachten wir zunächst die Funktion $g_{j,p} : \mathbb{R}^d \ni x \mapsto x_j^p$, die eine Komponente x_j des Arguments auswählt und zur p -ten Potenz erhebt. Mit der Abschätzung $|g_{j,p}(x)| = |x_j^p| \leq \|x\|_p^p$ sehen wir, dass $g_{j,p}$ stetig ist. Da eine allgemeine Polynomfunktion aus den $g_{j,p}$ durch Linearkombinationen und Produktbildungen hervorgeht, und diese Operationen Stetigkeit erhalten (Satz 4.11), folgt, dass Polynomfunktionen stetig sind.

Weitere Beispiele finden sich in den Übungen.

10.2 Stetigkeit von linearen Abbildungen

In Analysis 1 haben wir Differenzierbarkeit von Funktionen $f : \mathbb{R} \rightarrow \mathbb{R}$ unter anderem durch affine Approximierbarkeit charakterisiert, dh wir haben f in der Nähe eines Punktes x_0 durch eine Funktion der Form “linear + Konstante” ($x \mapsto f(x_0) + f'(x_0) \cdot (x - x_0)$) genähert. Dabei war es wichtig, dass die lineare Funktion $x \mapsto f'(x_0) \cdot x$ stetig ist.

Für Funktionen $\mathbb{R}^d \rightarrow \mathbb{R}^k$ werden wir ähnlich vorgehen. Als Vorbereitung gucken wir uns in diesem Abschnitt die Stetigkeitseigenschaften von linearen Abbildungen an. Wir sind hier also in engem Kontakt zur linearen Algebra. Es gibt dann keinen guten Grund, statt \mathbb{R}^d und \mathbb{R}^k nicht gleich zwei beliebige (zumeist endlichdimensionale) normierte Vektorräume V und W zu betrachten. Dies können Vektorräume über \mathbb{R} oder \mathbb{C} sein.

Kurze Erinnerungen an die lineare Algebra:

- Wir schreiben $\text{Hom}(V, W) := \{T : V \rightarrow W : T \text{ linear}\}$ für die Menge aller linearen Abbildungen $V \rightarrow W$ (Vektorraumhomomorphismen). Dies ist ein Vektorraum. Für $V = W$ schreiben wir auch $\text{End}(V) = \text{Hom}(V, V)$ (Vektorraumendomorphismen).
- Wie auch in der linearen Algebra üblich, schreiben wir im Falle einer linearen Abbildung $T : V \rightarrow W$ statt $T(v)$ meistens kürzer Tv .

Wir wollen den Raum aller linearen Abbildungen nun auch zu einem *normierten* Vektorraum machen.

Definition 10.6. Seien $(V, \|\cdot\|_V)$ und $(W, \|\cdot\|_W)$ zwei normierte Vektorräume und $T \in \text{Hom}(V, W)$. Wir definieren die *Operatornorm* von T als

$$\|T\| := \sup_{v \in V \setminus \{0\}} \frac{\|Tv\|_W}{\|v\|_V} = \sup_{v \in V, \|v\|_V=1} \|Tv\|_W \in [0, \infty].$$

Der Raum aller *beschränkten linearen Abbildungen* ist

$$\mathcal{L}(V, W) := \{T \in \text{Hom}(V, W) : \|T\| < \infty\}.$$

Die Operatornorm einer linearen Abbildung gibt ihren “maximalen Streckungsfaktor” an, also um wieviel größer $\|Tv\|_W$ als $\|v\|_V$ sein kann. Die Operatornorm kann endlich (z.B. für $V = W$ und $T = \text{id}_V$ gilt $\|T\| = 1$, viele weitere Beispiele weiter unten) oder unendlich sein (Beispiele später).

Lineare Abbildungen mit endlicher Operatornorm heißen *beschränkt* (s.o.), haben aber unbeschränktes Bild (falls $T \neq 0$)! Denn falls $v \in V$ existiert mit $Tv \neq 0$, also $\|Tv\|_W \neq 0$, so wird $\|T(\lambda v)\|_W = |\lambda| \|Tv\|_W$ beliebig groß, wenn $\lambda \in \mathbb{K}$ nur groß genug gewählt wird. “Beschränkt” für lineare Abbildungen heißt also nur, dass T die Norm nicht um beliebig große Faktoren streckt.

Eine gute Übung ist die folgende:

Zeigen Sie, dass die Operatornorm eine Norm auf $\mathcal{L}(V, W)$ ist.

Der Raum $\text{Hom}(V, W)$ ist im Allgemeinen kein normierter Vektorraum bzgl der Operatornorm, da die Operatornorm unendlich sein kann. Nur auf dem Untervektorraum $\mathcal{L}(V, W)$ ist die Operatornorm eine Norm.

Wie Sie in Funktionalanalysis lernen werden, gibt es auf jedem Vektorraum eine Norm, insbesondere auf $\text{Hom}(V, W)$ für beliebige Vektorräume V, W . Die Idee der Operatornorm ist aber, nicht eine beliebige Norm auf $\text{Hom}(V, W)$ bzw $\mathcal{L}(V, W)$ zu definieren, sondern eine spezielle Norm, die eng mit gegebenen Normen auf V und W verknüpft ist.

Satz 10.7. Seien $(V, \|\cdot\|_V)$ und $(W, \|\cdot\|_W)$ zwei normierte Vektorräume und $T : V \rightarrow W$ eine lineare Abbildung. Dann sind äquivalent:

- T ist stetig (auf ganz V).
- T ist stetig in $v = 0$.
- $\|T\| < \infty$.
- Es existiert eine Konstante $C > 0$ mit $\|Tv\|_W \leq C\|v\|_V$ für alle $v \in V$.

Beweis. c) \Rightarrow d). Sei $v \neq 0$. Dann gilt $\|Tv\|_W = \frac{\|Tv\|_W}{\|v\|_V} \cdot \|v\|_V \leq \|T\| \cdot \|v\|_V$, wir können die Konstante also als $C = \|T\|$ wählen.

d) \Rightarrow a). Für $v, w \in V$ haben wir $\|Tv - Tw\|_W = \|T(v - w)\|_W \leq C\|v - w\|_V$, was sofort die Stetigkeit von T zeigt. a) \Rightarrow b) ist trivial.

b) \Rightarrow c) Sei $\varepsilon > 0$. Aufgrund der Stetigkeit von T bei $v = 0$ gibt es $\delta > 0$ so, dass $\|v\|_V \leq \delta \Rightarrow \|Tv\|_W < \varepsilon$. Für beliebiges $v \in V \setminus \{0\}$ haben wir dann

$$\|Tv\|_W = \frac{1}{\delta} \|v\|_V \|T(\frac{\delta}{\|v\|_V} v)\|_W < \frac{\varepsilon}{\delta} \cdot \|v\|_V,$$

also $\|T\| \leq \frac{\varepsilon}{\delta} < \infty$. □

Beispiel 10.8. Wir betrachten \mathbb{R}^d mit Euklidischer Norm $\|\cdot\|_2$ und eine Diagonalmatrix

$$A = \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_d \end{pmatrix}$$

mit beliebigen Diagonaleinträgen $\lambda_1, \dots, \lambda_d \in \mathbb{R}$. Diese Matrix definiert eine lineare Abbildung $x \mapsto Ax$ (einen Endomorphismus von \mathbb{R}^d), denn wir mit A identifizieren können.

Wir behaupten, dass die Operatornorm (bzgl Euklidischer Norm im Definitions- und Bildbereich)

$$\|A\| = \max\{|\lambda_1|, \dots, |\lambda_d|\}$$

ist. Wegen $\|Ae_l\|_2 = \|\lambda_l \cdot e_l\|_2 = |\lambda_l|$ ist $\|A\| \geq \max\{|\lambda_1|, \dots, |\lambda_d|\}$ klar. Umgekehrt gilt

$$\|Ax\|_2^2 = \lambda_1^2 x_1^2 + \dots + \lambda_d^2 x_d^2 \leq \max\{\lambda_1^2, \dots, \lambda_d^2\} \cdot \|x\|_2^2,$$

was $\|A\| \leq (\max\{\lambda_1^2, \dots, \lambda_d^2\})^{1/2} = \max\{|\lambda_1|, \dots, |\lambda_d|\}$ impliziert.

Was können Sie über die Operatornorm einer orthonormalen Matrix sagen? Was können Sie über die Operatornorm einer symmetrischen Matrix sagen?

Wir möchten diesen Satz nun noch dazu verwenden, zu zeigen, dass auf endlichdimensionalen Vektorräumen alle Normen äquivalent sind. Insbesondere sind auf \mathbb{K}^d alle Normen (nicht nur die p -Normen) äquivalent.

Satz 10.9. Auf einem endlichdimensionalen Vektorraum V sind alle Normen äquivalent.

Beweis. Wir betrachten zunächst den Fall $V = \mathbb{K}^d$.

Sei $\|\cdot\|$ eine beliebige Norm auf \mathbb{K}^d . Wir zeigen, dass $\|\cdot\|$ zu $\|\cdot\|_\infty$ äquivalent ist. Dies impliziert die Behauptung, da Äquivalenz von Normen eine Äquivalenzrelation ist.

Wir müssen also zeigen, dass es Konstante $c, C > 0$ gibt, so dass $c\|x\|_\infty \leq \|x\| \leq C\|x\|_\infty$ für alle $x \in \mathbb{K}^d$ gilt. Die eine Ungleichung ist einfach: Bezeichnet $(e_l)_{l=1\dots d}$ die Standard-Basis von \mathbb{K}^d , so haben wir $x = x_1 \cdot e_1 + \dots + x_d \cdot e_d$ und damit

$$\|x\| = \left\| \sum_{l=1}^d x_l \cdot e_l \right\| \leq \sum_{l=1}^d |x_l| \cdot \|e_l\| \leq \max\{|x_1|, \dots, |x_d|\} \cdot \sum_{l=1}^d \|e_l\| = \|x\|_\infty \cdot \sum_{l=1}^d \|e_l\|.$$

Mit $C := \sum_{l=1}^d \|e_l\| < \infty$ liefert dies $\|x\| \leq C\|x\|_\infty$.

Die andere Ungleichung ($c\|x\|_\infty \leq \|x\|$) ist etwas schwieriger. Zuerst beobachten wir, dass $\|\cdot\| : \mathbb{K}^d \rightarrow \mathbb{R}$ stetig bzgl der Norm $\|\cdot\|_\infty$ ist: Per umgekehrter Dreiecksungleichung gilt ja

$$\left| \|x\| - \|y\| \right| \leq \|x - y\| \leq C\|x - y\|_\infty.$$

Für $\|x - y\|_\infty < \varepsilon/C$ gilt also stets $\| \|x\| - \|y\| \| < \varepsilon$, d.h. $\| \cdot \|$ ist stetig.

Wir betrachten nun die Würfeloberfläche $S := \{x \in \mathbb{K}^d : \|x\|_\infty = 1\}$, die mit dem Rand $\partial U_1^\infty(0)$ der $\varepsilon = 1$ Umgebung von 0 bzgl $\| \cdot \|_\infty$ übereinstimmt. Also ist S abgeschlossen. Da S offensichtlich auch beschränkt ist, ist S kompakt. Die stetige Funktion $\| \cdot \|$ nimmt also ein Minimum c auf S an. Dieses Minimum ist nicht 0, da $\|x\| = 0$ nur für $x = 0$ gilt, und $0 \notin S$.

Nun zu der gesuchten Ungleichung: Für beliebiges $x \neq 0$ gilt $x = \|x\|_\infty \cdot \frac{x}{\|x\|_\infty}$, wobei $\frac{x}{\|x\|_\infty} \in S$. Also erhalten wir

$$\|x\| = \left\| \|x\|_\infty \frac{x}{\|x\|_\infty} \right\| \geq c \|x\|_\infty.$$

Zum Schluss verallgemeinern wir noch von \mathbb{K}^d auf V . Wie aus linearer Algebra bekannt, gibt es zu jedem Vektorraum V über \mathbb{K} mit $d := \dim V < \infty$ eine invertierbare (bijektive) lineare Abbildung $A : \mathbb{K}^d \rightarrow V$. Seien nun $\| \cdot \|$ und $\| \cdot \|'$ zwei Normen auf V . Dann sind $\|x\|_A := \|Ax\|$ und $\|x\|'_A := \|Ax\|'$ zwei Normen auf \mathbb{K}^d (Übung: Prüfen Sie das nach). Also gibt es $c, C > 0$ mit $c\|x\|_A \leq \|x\|'_A \leq C\|x\|_A$ für alle $x \in \mathbb{K}^d$. Aber dann gilt auch

$$c\|v\| = c\|AA^{-1}v\| = c\|A^{-1}v\|_A \leq \|A^{-1}v\|'_A = \|v\|' \leq C\|A^{-1}v\|_A = C\|v\|,$$

was den Beweis abschließt. □

Das folgende Korollar besagt, dass wir uns über die Stetigkeit von linearen Abbildungen *im Endlichdimensionalen* keine Sorgen machen müssen.

Korollar 10.10. *Seien $(V, \| \cdot \|_V)$ und $(W, \| \cdot \|_W)$ zwei normierte (reelle oder komplexe) Vektorräume, V endlichdimensional, und $T : V \rightarrow W$ linear. Dann ist T stetig.*

Beweis. Wir definieren eine neue Norm $\| \cdot \|$ auf V durch $\|v\| := \|v\|_V + \|Tv\|_W$. Dies ist tatsächlich eine Norm (Übung). Also ist sie nach Satz 10.9 zu $\| \cdot \|_V$ äquivalent (hier wird $\dim V < \infty$ benutzt), d.h. es gibt $C > 0$ mit

$$\|v\| = \|v\|_V + \|Tv\|_W \leq C\|v\|_V \quad \forall v \in V.$$

Also haben wir für jeden Vektor $v \in V$

$$\|Tv\|_W \leq \|v\|_V + \|Tv\|_W \leq C\|v\|_V,$$

was nach Satz 10.7 die Stetigkeit von T zeigt. □

Sie werden sich nun vielleicht fragen, ob es überhaupt unstetige lineare Abbildungen gibt. Wir geben ein Beispiel.

Beispiel 10.11. Sei $I = [0, 1]$ das kompakte Einheitsintervall, $V = (C^1(I), \| \cdot \|_I)$ der normierte Vektorraum aller stetig differenzierbaren Funktionen auf I , und $W = (C(I), \| \cdot \|_I)$ der normierte Vektorraum aller stetigen Funktionen auf I . Wir betrachten die Abbildung

$$T : V \rightarrow W, \quad Tf := f',$$

die einer C^1 -Funktion ihre Ableitung $f \in C(I)$ zuordnet. Dies ist eine lineare Abbildung, denn $T(\lambda f + \mu g) = (\lambda f + \mu g)' = \lambda f' + \mu g' = \lambda T f + \mu T g$ gilt nach den bekannten Ableitungsregeln für alle $\lambda, \mu \in \mathbb{R}$ und alle $f, g \in C^1(I)$ (Satz 5.5).

Wir behaupten, dass T unstetig ist. Dazu betrachten wir die Funktionen

$$f_n(x) := \sin(nx), \quad n \in \mathbb{N},$$

die alle in V liegen. Die Supremumsnorm von f_n ist $\|f_n\|_I = \sup_{x \in I} |\sin(nx)| \leq 1$.

Es gilt $(T f_n)(x) = f_n'(x) = n \cos(nx)$, insbesondere also $(T f_n)(0) = n$. Also haben wir $\|T f_n\|_I \geq n$. Wäre T stetig, so würde es nach Satz 10.7 eine Konstante $C > 0$ geben, so dass $\|T f_n\|_I \leq C \|f_n\|_I$ für alle $n \in \mathbb{N}$ gilt. Wegen

$$n \leq \|T f_n\|_I \leq C \|f_n\|_I \leq C$$

ist dies aber unmöglich. Also ist T unstetig.

Noch einmal der Deutlichkeit halber: Lineare Abbildungen zwischen endlichdimensionalen normierten Vektorräumen sind immer stetig (egal, welche Norm Sie betrachten). Sind V, W Vektorräume mit $d := \dim V < \infty$ und $k := \dim W < \infty$, so erhalten wir durch Wahl von Basen – wie aus Linearer Algebra bekannt – Vektorraumisomorphismen (invertierbare lineare Abbildungen) $\varphi : V \rightarrow \mathbb{K}^d$ und $\psi : W \rightarrow \mathbb{K}^k$. Wir können jede lineare Abbildung $T : V \rightarrow W$ dann durch eine $(k \times d)$ -Matrix A darstellen. Genauer gesagt gilt

$$A = \psi \circ T \circ \varphi^{-1} : \mathbb{K}^d \rightarrow \mathbb{K}^k, \quad x \mapsto Ax.$$

Da φ^{-1}, T, ψ alle stetig sind, ist die Matrixmultiplikationsabbildung $x \mapsto Ax$ auch stetig. Zusammenfassend notieren wir:

Korollar 10.12. *Sei V ein endlichdimensionaler Vektorraum.^a*

- a) *Alle Normen auf V sind äquivalent.*
- b) *Alle Normen auf V definieren die gleichen Begriffe von Offenheit, Abgeschlossenheit, Beschränktheit und Kompaktheit.*
- c) *V ist vollständig bzgl jeder Norm auf V .*
- d) *Alle Normen auf V definieren den gleichen Konvergenzbegriff für Folgen in V .*
- e) **Satz von Heine-Borel:** *Eine Teilmenge von V ist kompakt genau dann, wenn sie abgeschlossen und beschränkt bzgl irgendeiner (dann: jeder) Norm auf V ist.*
- f) *Sei (X, d) ein metrischer Raum und $f : V \rightarrow X, g : X \rightarrow V$ Abbildungen. Dann ist f (bzw. g) stetig bzgl einer Norm auf V , wenn f (bzw. g) stetig bzgl jeder Norm auf V ist. Der Begriff der Stetigkeit hängt also nicht von der gewählten Norm auf V ab (wenn die Metrik auf X fixiert ist).*

^aHier, wie überall in diesem Skript, ist V ein Vektorraum über \mathbb{R} oder \mathbb{C} . Die unten stehenden Aussagen stimmen nicht für alle anderen Körper, wir interessieren uns nur für reelle oder komplexe Vektorräume.

Noch einmal zur Betonung: Für all diese Aussagen ist es wesentlich, dass V endlichdimensional ist. Für unendlichdimensionales V sind i.A. alle Aussagen a)–f) falsch.

10.3 Skalarprodukte

In diesem Abschnitt nehmen wir auf ein weiteres Konzept aus der Linearen Algebra Bezug, nämlich auf Skalarprodukte.

Definition 10.13. Sei V ein Vektorraum über $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$. Ein *Skalarprodukt* ist eine Abbildung

$$\langle \cdot, \cdot \rangle : V \times V \rightarrow \mathbb{K},$$

so dass

- a) Für alle $v \in V$ ist $V \rightarrow \mathbb{K}, w \mapsto \langle v, w \rangle$ linear (Linearität im zweiten Argument).
- b) Für alle $v, w \in V$ gilt $\langle w, v \rangle = \overline{\langle v, w \rangle}$. Also ist $\langle \cdot, \cdot \rangle$ im ersten Argument linear (für $\mathbb{K} = \mathbb{R}$, d.h. $\langle \cdot, \cdot \rangle$ ist eine Bilinearform) bzw antilinear (für $\mathbb{K} = \mathbb{C}$, d.h. $\langle \cdot, \cdot \rangle$ ist eine Sesquilinearform).
- c) $\langle v, v \rangle \in [0, \infty)$ für alle $v \in V$, und $\langle v, v \rangle = 0$ genau dann wenn $v = 0$ (positive Definitheit).

Die Konventionen bzgl komplexen Skalarprodukten sind nicht ganz einheitlich; manchmal wird das Skalarprodukt auch im ersten statt im zweiten Argument als linear gewählt. Zur Übersetzung zwischen beiden Konventionen müssen Sie einfach alle Skalarprodukte komplex konjugieren.

Beispiel 10.14.

- a) Das Standardskalarprodukt auf \mathbb{K}^d ist

$$\langle x, y \rangle = \sum_{l=1}^d \overline{x_l} \cdot y_l.$$

Dies ist tatsächlich ein Skalarprodukt (leichte Übung). Es gilt die Abschätzung

$$|\langle x, y \rangle| \leq \|x\|_p \|y\|_q,$$

wobei $p, q \in [1, \infty]$ mit $\frac{1}{p} + \frac{1}{q} = 1$ (Hölder-Ungleichung).

- b) Skalarprodukte gibt es auch auf unendlichdimensionalen Vektorräumen: Auf $V = C(I), I = [0, 1]$ ist durch

$$\langle f, g \rangle := \int_0^1 \overline{f(x)} g(x) dx$$

ein Skalarprodukt gegeben (Übung).

Wie aus Linearer Algebra bekannt, induziert jedes Skalarprodukt eine Norm via

$$\|v\| := \sqrt{\langle v, v \rangle},$$

und es gilt die Cauchy-Schwarz Ungleichung

$$|\langle v, w \rangle| \leq \|v\| \|w\|.$$

Das Standardskalarprodukt auf \mathbb{K}^d induziert die Euklidische Norm $\|x\|_2 = \left(\sum_{l=1}^d |x_l|^2\right)^{1/2}$.

- Ein Vektorraum mit Skalarprodukt, der bzgl der induzierten Norm ein Banachraum (also vollständig) ist, heißt *Hilbertraum*. Hilberträume spielen in der Funktionalanalysis und Quantenphysik eine herausragende Rolle.
- Ein reeller Vektorraum mit Skalarprodukt heißt *Euklidischer Vektorraum*. In einem Euklidischen Vektorraum kann man den *Winkel* zwischen zwei Vektoren $v, w \neq 0$ definieren als

$$\alpha(v, w) := \arccos \frac{\langle v, w \rangle}{\|v\| \|w\|},$$

wobei $\arccos : [-1, 1] \rightarrow [0, \pi]$ die Umkehrfunktion von $\cos : [0, \pi] \rightarrow [-1, 1]$ ist.

Zu jedem Skalarprodukt gibt es eine zugehörige (induzierte) Norm. Gibt es auch zu jeder Norm $\|\cdot\|$ ein Skalarprodukt, so dass $\|v\| = \sqrt{\langle v, v \rangle}$ gilt? Die Antwort auf diese Frage ist Nein. Man überlegt sich das folgendermaßen: Ist $\|\cdot\|$ eine von einem Skalarprodukt $\langle \cdot, \cdot \rangle$ induzierte Norm, so gilt die *Parallelogrammgleichung*

$$\|v + w\|^2 + \|v - w\|^2 = 2(\|v\|^2 + \|w\|^2).$$

Malen Sie ein Parallelogramm in \mathbb{R}^2 mit zwei durch $v, w \in \mathbb{R}^2$ gegebene Seiten, um den Namen dieser Gleichung zu verstehen. Das Überprüfen der Parallelogrammgleichung erfolgt durch Einsetzen der Definition $\|v\|^2 = \langle v, v \rangle$ und Ausmultiplizieren.

Andererseits gilt auch: Ist $\|\cdot\|$ eine Norm, die die Parallelogrammgleichung erfüllt, so ist sie von einem Skalarprodukt induziert, und zwar von

$$\langle v, w \rangle = \begin{cases} \frac{1}{4} (\|v + w\|^2 - \|v - w\|^2) & \mathbb{K} = \mathbb{R} \\ \frac{1}{4} (\|v + w\|^2 - \|v - w\|^2) + \frac{i}{4} (\|v + iw\|^2 - \|v - iw\|^2) & \mathbb{K} = \mathbb{C} \end{cases}$$

Eine schwierigere Übung ist es, zu zeigen, dass dies tatsächlich ein Skalarprodukt ist.

Eine Norm ist also von einem Skalarprodukt induziert genau dann wenn die Parallelogrammgleichung gilt.

Unter den p -Normen auf \mathbb{K}^d ist nur die Euklidische Norm $\|\cdot\|_2$ durch ein Skalarprodukt induziert.

Wir ergänzen unsere Betrachtungen zu Skalarprodukten um die Bemerkung, dass ein Skalarprodukt $\langle \cdot, \cdot \rangle$ stetig in der zugehörigen Norm ist. Dazu müssen wir zuerst $V \times V$ zu einem metrischen Raum machen: Wir definieren für Paare $(v, w) \in V \times V$

$$\|(v, w)\| := \max\{\|v\|, \|w\|\}.$$

Dies ist eine Norm auf $V \times V$ (Übung).

Lemma 10.15. Sei V ein Vektorraum mit Skalarprodukt $\langle \cdot, \cdot \rangle$ und sei $\|\cdot\|$ die von $\langle \cdot, \cdot \rangle$ induzierte Norm auf V . Dann ist

$$\langle \cdot, \cdot \rangle : V \times V \rightarrow \mathbb{K}$$

stetig.

Beweis. Sei $(v_n, w_n) \in V \times V$ eine Folge, die gegen (v, w) konvergiert. Wir schätzen ab per Cauchy-Schwarzscher Ungleichung

$$\begin{aligned} |\langle v_n, w_n \rangle - \langle v, w \rangle| &= |\langle v_n - v, w_n \rangle - \langle v, w_n - w \rangle| \\ &\leq \|v_n - v\| \|w_n\| + \|v\| \|w_n - w\|. \end{aligned}$$

Im Limes $n \rightarrow \infty$ gilt $v_n \rightarrow v$ und $w_n \rightarrow w$ in V . Aufgrund der Stetigkeit von $\|\cdot\|$ gilt weiterhin $\|w_n\| \rightarrow \|w\|$. So sehen wir $\langle v_n, w_n \rangle \rightarrow \langle v, w \rangle$, d.h. das Skalarprodukt ist stetig. \square

11 Kurven

In diesem Kapitel beginnen wir unsere Diskussion von mehrdimensionaler Differentialrechnung. Im Allgemeinen betrachten wir dazu Funktionen $f : G \subset \mathbb{R}^d \rightarrow \mathbb{R}^k$ mit $d, k \in \mathbb{N}$. In diesem einführenden Kapitel beschränken wir uns zunächst auf $d = 1$, also Abbildungen $I \subset \mathbb{R} \rightarrow \mathbb{R}^k$ von Zahlen $t \in I$ zu Vektoren $\gamma(t) \in \mathbb{R}^k$.

Eine nützliche geometrische Anschauung ist, dass $\gamma(t)$ die Position eines Teilchens im k -dimensionalen Raum \mathbb{R}^k zu einem Zeitpunkt t beschreibt. Die Position $\gamma(t)$ kann sich mit der Zeit t ändern, dann haben wir es mit einer Kurve im Raum zu tun.

11.1 Stetige und differenzierbare Kurven

Definition 11.1. Eine *Kurve* ist eine stetige Abbildung $\gamma : [a, b] \rightarrow \mathbb{R}^k$ (mit $a, b \in \mathbb{R}$, $a \leq b$ und $k \in \mathbb{N}$). Der *Anfangspunkt* von γ ist $\gamma(a)$, der *Endpunkt* von γ ist $\gamma(b)$. Falls $\gamma(a) = \gamma(b)$, so heißt γ *geschlossen*. Das Bild $\gamma([a, b]) \subset \mathbb{R}^k$ von γ heißt *Spur* der Kurve.

Wie im vorigen Kapitel besprochen, können wir eine Kurve in ihre Komponentenfunktionen

$$\gamma(t) = (\gamma_1(t), \dots, \gamma_k(t))$$

aufteilen und die Stetigkeit von γ (bzgl irgendeiner Norm auf \mathbb{R}^k) ist gleichbedeutend mit der Stetigkeit aller Komponentenfunktionen γ_j , $j = 1, \dots, k$.

Beispiel 11.2.

- Eine Gerade ist eine Kurve der Form $\gamma : \mathbb{R} \rightarrow \mathbb{R}^k$, $\gamma(t) = x + t \cdot v$. Der Vektor v gibt die Richtung der Geraden an.
- Eine geschlossene Kreiskurve in \mathbb{R}^2 ist durch

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

gegeben. Hierbei bezeichnet $r > 0$ den Radius des Kreises. Beachten Sie, dass die Spur dieser Kurve mit der Spur der anderen Kurve

$$\tilde{\gamma} : [0, 2\pi] \rightarrow \mathbb{R}^2, \quad \tilde{\gamma}(t) = r(\cos(2t), \sin(2t))$$

übereinstimmt. Anschaulich durchläuft $\tilde{\gamma}$ den Kreis mit der doppelten Geschwindigkeit, führt also zwei Umdrehungen aus, während γ nur eine Umdrehung ausführt.

Für differenzierbare Abbildungen $f : [a, b] \rightarrow \mathbb{R}$ gibt die Ableitung $f'(x)$ die Steigung (Richtung) der Tangenten an den Graphen von f im Punkte $(x, f(x))$ an. Wir wollen nun auch für Kurven im mehrdimensionalen Raum Tangenten und ein Konzept von Differenzierbarkeit definieren. Tatsächlich überträgt sich die Definition aus Analysis 1 ohne größere Probleme:

Definition 11.3. Eine Kurve $\gamma : [a, b] \rightarrow \mathbb{R}^k$ heißt in $t \in [a, b]$ *differenzierbar*, falls der Grenzwert

$$\gamma'(t) := \lim_{h \rightarrow 0} \frac{\gamma(t+h) - \gamma(t)}{h}$$

existiert. Der Grenzwert wird dann *Ableitung* von γ in t genannt. Ist γ für alle t differenzierbar, so heißt γ differenzierbar. Ist $\gamma' : [a, b] \rightarrow \mathbb{R}^k$ stetig, so heißt γ *stetig differenzierbar*.

Bemerken Sie, dass der $\gamma'(t)$ definierende Grenzwert ein Grenzwert in \mathbb{R}^k ist. Wegen

$$\frac{\gamma(t+h) - \gamma(t)}{h} = \left(\frac{\gamma_1(t+h) - \gamma_1(t)}{h}, \dots, \frac{\gamma_k(t+h) - \gamma_k(t)}{h} \right)$$

heißt das, dass γ' genau dann differenzierbar ist, wenn alle Komponentenfunktionen γ_j differenzierbar sind. In diesem Falle gilt

$$\gamma'(t) = (\gamma'_1(t), \dots, \gamma'_k(t)).$$

Es ist hier wichtig, dass γ eine *zahlenwertige* (skalare) Variable hat, für einen *Vektor* h wäre das $\frac{1}{h}$ in der Definition der Ableitung bedeutungslos.

Die geometrische Anschauung der Ableitung ist, dass der Vektor $\gamma'(t) \in \mathbb{R}^k$ die *Richtung* der Kurve γ in t angibt (also die Richtung, in der der Punkt $\gamma(t)$ die Spur der Kurve durchläuft, wenn t von a nach b wächst) und die (Euklidische) Norm $\|\gamma'(t)\|_2$ die *Geschwindigkeit* von γ in t angibt.

Beispiel 11.4.

- Die gerade Kurve $\gamma : \mathbb{R} \rightarrow \mathbb{R}^k$, $\gamma(t) = x + t \cdot v$ ist differenzierbar mit Ableitung $\gamma'(t) = v$ für alle $t \in \mathbb{R}$. Hier sind also sowohl die Richtung als auch die Geschwindigkeit der Kurve konstant.
- Die Kreiskurve $\gamma : [0, 2\pi] \rightarrow \mathbb{R}^2$, $\gamma(t) = (r \cos t, r \sin t)$ ist differenzierbar mit Ableitung

$$\gamma'(t) = r \begin{pmatrix} -\sin t \\ \cos t \end{pmatrix}.$$

Also gilt $\|\gamma'(t)\|_2 = r$ für alle t und mit dem Euklidischen Skalarprodukt

$$\langle \gamma(t), \gamma'(t) \rangle = 0.$$

Die Geschwindigkeit ist also konstant, und der Geschwindigkeitsvektor $\gamma'(t)$ steht immer genau senkrecht zur Kurve $\gamma(t)$ (Skalarprodukt = 0).

Der Weg $t \mapsto (r \cos t, -r \sin t)$ beschreibt auch einen Kreis, der aber im Gegensatz zu γ im Uhrzeigersinn durchlaufen wird.

- Welchen Weg beschreibt ein Punkt auf Ihrem Fahrradreifen, wenn Sie fahren?

Eine sogenannte *Zykloide*:

$$\gamma : \mathbb{R} \rightarrow \mathbb{R}^2, \quad \gamma(t) := r \begin{pmatrix} t - \sin t \\ 1 - \cos t \end{pmatrix}.$$

Skizzieren Sie die Spur dieses Weges. Trotz der "Spitzen" im Weg ist γ differenzierbar mit

$$\gamma'(t) = r \begin{pmatrix} 1 - \cos t \\ \sin t \end{pmatrix}.$$

Wie sehen, dass der Punkt genau dann in Ruhe ist (also $\gamma'(t) = 0$), wenn $\cos(t) = 1$ gilt (dann gilt auch automatisch $\sin(t) = 0$). Dies ist genau bei $t = 2\pi n$, $n \in \mathbb{Z}$, der Fall, was wiederum zu $\gamma(t)_2 = 0$ (y -Koordinate gleich Null, der Punkt befindet sich am Boden) äquivalent ist. Die Spur der Kurve hat in diesen Punkten eine Spitze und ist nur deshalb dort differenzierbar, weil sich der Punkt dort in Ruhe (Geschwindigkeitsvektor=0) befindet.

Für eine stetig differenzierbare Kurve γ ist die Ableitung γ' wieder eine Kurve, die wieder differenzierbar sein kann. Die zweite Ableitung γ'' hat die Bedeutung der *Beschleunigung* (Änderung der Geschwindigkeit).

11.2 Bogenlänge

Wir betrachten nun die (Euklidische) Länge einer Kurve.

Definition 11.5. Sei $\gamma : [a, b] \rightarrow \mathbb{R}^k$ eine stetig differenzierbare Kurve. Dann ist die *Länge* von γ

$$L(\gamma) := \int_a^b \|\gamma'(t)\|_2 dt = \int_a^b \sqrt{\gamma_1'(t)^2 + \dots + \gamma_k'(t)^2} dt.$$

- Diese Definition ist mathematisch sinnvoll, da $\gamma' : [a, b] \rightarrow \mathbb{R}^k$ und $\|\cdot\|_2 : \mathbb{R}^k \rightarrow \mathbb{R}$ stetig sind. Also ist der Integrand stetig, das Riemann-Integral existiert also.
- Man kann die Länge auch definieren, wenn γ nur stückweise stetig differenzierbar ist, d.h. wenn es $a = \xi_1 < \dots < \xi_n = b$ gibt, so dass $\gamma|_{[\xi_{j+1}, \xi_j]}$ für alle $j \in \{0, \dots, n-1\}$ stetig differenzierbar ist. In diesem Fall definiert man $L(\gamma)$ als die Summe der Längen dieser Einschränkungen.
- Die Definition von $L(\gamma)$ entspricht unserer geometrischen/physikalischen Anschauung: Die Länge des Weges erhält man, indem man die Geschwindigkeit über den Weg aufsummiert/integriert.
- Statt der Euklidischen Norm könnte man auch eine andere Norm zur Definition der Länge heranziehen, aber $\|\cdot\|_2$ passt zu unserem alltäglichen Längenbegriff.

Wir testen diesen Begriff an einigen Beispielen.

Beispiel 11.6.

a) Länge eines Kreises. Für $\gamma : [0, 2\pi] \rightarrow \mathbb{R}^2$, $\gamma(t) = (r \cos t, r \sin t)$ haben wir

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

b) Länge einer Zykloide über eine Umdrehung: Der schon vorher betrachtete Weg $\gamma : [0, 2\pi] \rightarrow \mathbb{R}^2$, $\gamma(t) := r(t - \sin t, 1 - \cos t)$ hat die Geschwindigkeit $\|\gamma'(t)\|_2 = r\sqrt{(1 - \cos t)^2 + \sin^2 t} = r\sqrt{2 - 2 \cos t}$ Länge

$$L(\gamma) = r \int_0^{2\pi} \sqrt{2 - 2 \cos t} dt = 8r.$$

Die Berechnung des obigen Integrals wird als Übung überlassen.

Unsere geometrische Anschauung ordnet die Länge einer (einfach durchlaufenen) Kurve der *Spur* der Kurve zu, während unsere mathematische Definition die Länge der Kurve selbst, also der Abbildung $\gamma : [a, b] \rightarrow \mathbb{R}^k$ zuordnet. Wir erwarten also, dass Kurven mit derselben (einfach durchlaufenen) Bahn die gleiche Länge haben.

Beispiel 11.7. Wir betrachten die drei Kurven $\gamma, \eta, \xi : [0, 1] \rightarrow \mathbb{R}^2$,

$$\gamma(t) = \begin{pmatrix} \cos(2\pi t) \\ \sin(2\pi t) \end{pmatrix}, \quad \eta(t) = \begin{pmatrix} \cos(2\pi t^2) \\ \sin(2\pi t^2) \end{pmatrix}, \quad \xi(t) = \begin{pmatrix} \cos(6\pi t) \\ \sin(6\pi t) \end{pmatrix}.$$

Die Kurve γ beschreibt einen Kreis mit Radius 1, der einmal im Gegenuhrzeigersinn mit konstanter Geschwindigkeit $\|\gamma'(t)\|_2 = 2\pi$ durchlaufen wird. Also $L(\gamma) = 2\pi$.

Die Kurve η hat die gleiche Spur wie γ , durchläuft den Kreis ebenfalls einmal im Gegenuhrzeigersinn: Wenn t das Intervall $[0, 1]$ durchläuft, durchläuft t^2 auch das Intervall $[0, 1]$, "ohne Umzukehren". Das liegt daran, dass $[0, 1] \ni t \mapsto t^2 \in [0, 1]$ bijektiv und streng monoton wachsend ist. Die Geschwindigkeit von η ist aber nicht konstant, sondern wächst linear mit der Zeit: $\|\eta'(t)\|_2 = 4\pi t$. Die Länge von η ist

$$L(\eta) = \int_0^1 4\pi t dt = 4\pi \left[\frac{t^2}{2} \right]_0^1 = 2\pi,$$

was mit der Länge von γ übereinstimmt.

Die Kurve ξ hingegen durchläuft den Kreis *dreimal* im Gegenuhrzeigersinn, mit konstanter Geschwindigkeit $\|\xi'(t)\|_2 = 6\pi$. Ihre Länge ist also $L(\xi) = 6\pi = 3L(\gamma)$.

Wir formalisieren diese Beobachtungen in der folgenden Definition.

Definition 11.8. Sei $\gamma : [a, b] \rightarrow \mathbb{R}^k$ eine stetig differenzierbare Kurve und $\varphi : [c, d] \rightarrow [a, b]$ bijektiv, streng monoton wachsend, und stetig differenzierbar. Dann ist auch $\gamma \circ \varphi$:

$[c, d] \rightarrow \mathbb{R}^k$ eine stetig differenzierbare Kurve, die wir eine *Umparametrisierung* von γ nennen.

Beachten Sie, dass eine Umparametrisierung insbesondere $\varphi(c) = a$ und $\varphi(d) = b$ erfüllt. Wir zeigen nun, dass sich – wie erwartet – die Länge einer Kurve unter Umparametrisierung nicht ändert.

Satz 11.9 (Parametrisierungsinvarianz der Länge einer Kurve). Sei $\gamma : [a, b] \rightarrow \mathbb{R}^k$ eine stetig differenzierbare Kurve und $\gamma \circ \varphi : [c, d] \rightarrow \mathbb{R}^k$ eine Umparametrisierung. Dann gilt

$$L(\gamma \circ \varphi) = L(\gamma).$$

Beweis. Da die Ableitung einer Kurve komponentenweise gebildet wird, erhalten wir mit der Kettenregel aus Analysis 1 die Beziehung

$$(\gamma \circ \varphi)'(t) = (\varphi'(t)\gamma'_1(\varphi(t)), \dots, \varphi'(t)\gamma'_k(\varphi(t))) = \varphi'(t) \cdot \gamma'(\varphi(t)).$$

Da φ streng monoton wachsend ist, gilt $\varphi'(t) \geq 0$ und damit $\|(\gamma \circ \varphi)'(t)\|_2 = \varphi'(t)\|\gamma'(\varphi(t))\|_2$.

Die Länge der umparametrisierten Kurve ist also

$$L(\gamma \circ \varphi) = \int_c^d \varphi'(t)\|\gamma'(\varphi(t))\|_2 dt.$$

Mit Hilfe der Substitutionsregel (Satz 6.20) erhalten wir nun wegen $\varphi(c) = a$ und $\varphi(d) = b$

$$L(\gamma \circ \varphi) = \int_{\varphi^{-1}(a)}^{\varphi^{-1}(b)} \varphi'(t)\|\gamma'(\varphi(t))\|_2 dt = \int_a^b \|\gamma'(t)\|_2 dt = L(\gamma).$$

□

Wir zeigen noch, dass C^1 -Kurven stets so umparametrisiert werden können, dass die umparametrisierte Kurve $\tilde{\gamma} := \gamma \circ \varphi$

$$\|\tilde{\gamma}'(t)\|_2 = 1$$

erfüllt. Dies nennt man *Parametrisierung durch die Bogenlänge*. Diese Parametrisierung ist geometrisch natürlich und führt mitunter zu einfacheren Formeln.

Lemma 11.10. Sei $\gamma : [a, b] \rightarrow \mathbb{R}^k$ eine stetig differenzierbare Kurve mit $\gamma'(t) \neq 0$ für alle $t \in [a, b]$. Dann kann γ durch die Bogenlänge parametrisiert werden, d.h. es gibt eine Umparametrisierung $\varphi : [c, d] \rightarrow [a, b]$, so dass $\|(\gamma \circ \varphi)'(t)\|_2 = 1$ für alle $t \in [c, d]$ gilt.

Beweis. Wir betrachten die Abbildung (Bogenlänge)

$$\psi(t) := \int_a^t \|\gamma'(t)\|_2 dt.$$

Da der Integrand stetig ist, ist ψ eine stetig differenzierbare streng monoton wachsende Funktion mit $\psi(a) = 0$, $\psi(b) = L(\gamma) > 0$, und $\psi'(t) = \|\gamma'(t)\|_2$. Nach dem Satz über die Ableitung der Umkehrfunktion (Satz 5.9) ist $\psi : [a, b] \rightarrow [0, L(\gamma)]$ bijektiv mit differenzierbarer Umkehrfunktion $\varphi := \psi^{-1}$. Die Ableitung der Umkehrfunktion ist

$$\varphi'(t) = \frac{1}{\psi'(\varphi(t))} = \frac{1}{\|\gamma'(\varphi(t))\|_2}.$$

Also gilt

$$\|(\gamma \circ \varphi)'(t)\|_2 = \frac{1}{\|\gamma'(\varphi(t))\|_2} \cdot \|\gamma'(\varphi(t))\|_2 = 1,$$

d.h. $\gamma \circ \varphi$ ist nach der Bogenlänge parametrisiert. □

Anschaulich gesprochen bewegt sich der Punkt $\gamma(t)$ entlang der Spur von γ mit konstanter Geschwindigkeit $\|\gamma'(t)\|_2 = 1$ fort, wenn γ nach der Bogenlänge parametrisiert ist. Die *Richtung* des Geschwindigkeitsvektors $\gamma'(t)$ kann sich trotzdem mit t ändern. Dies ist ein Ausdruck dafür, wie sehr die Kurve im Punkt $\gamma(t)$ "gekrümmt" ist. Da die Änderung von $\gamma'(t)$ durch die zweite Ableitung $\gamma''(t)$ gegeben ist, wird folgende Definition nahegelegt.

Definition 11.11. Sei $\gamma : I \rightarrow \mathbb{R}^k$ eine nach der Bogenlänge parametrisierte zweimal stetig differenzierbare Kurve. Dann ist die (betragsmäßige) *Krümmung* von γ im Punkt $\gamma(t)$

$$k(t) := \|\gamma''(t)\|_2.$$

Die Parametrisierung auf die Bogenlänge ist nur in wenigen Fällen explizit berechenbar. Hier sind zwei einfache Beispiele, wo es geht.

Beispiel 11.12.

- a) Die Krümmung einer geraden Kurve $\gamma : [a, b] \rightarrow \mathbb{R}^k$, $\gamma(t) = u + t \cdot v$ ist Null: Die Bogenlänge ist $\psi(t) = \int_a^t \|v\| ds = \|v\| \cdot t$, also $\varphi(t) = \frac{t}{\|v\|}$. Nach Bogenlänge parametrisiert haben wir also $(\gamma \circ \varphi)(t) = u + t \cdot \frac{v}{\|v\|}$, was verschwindende zweite Ableitung hat.
- b) Die Krümmung einer Kreiskurve $\gamma(t) = (r \cos \omega t, r \sin \omega t)$ ist konstant: Die Bogenlänge ist $\psi(t) = \int_a^t r \omega ds = r \omega \cdot t$, also $\varphi(t) = \frac{t}{r \omega}$. Nach Bogenlänge parametrisiert haben wir also $(\gamma \circ \varphi)(t) = (r \cos \frac{t}{r}, r \sin \frac{t}{r})$. Die zweite Ableitung des reparametrisierten Weges ist also $(-\frac{1}{r} \cos \frac{t}{r}, -\frac{1}{r} \sin \frac{t}{r})$, mit Länge $\frac{1}{r}$ für alle t .

Gemäß des zweiten Beispiel können Sie sich die Krümmung einer C^2 -Kurve γ in einem Punkt $\gamma(t_0)$ auch als den inversen Radius eines Kreises vorstellen, der die Kurve in $\gamma(t_0)$ tangential berührt (Schmiegekreis).

Weitere Beispiele finden Sie in den Übungen.

Wir werden später etwas mehr zu gekrümmten d -dimensionalen Flächen im \mathbb{R}^k sagen (z.B. eine zweidimensionale Kugeloberfläche im dreidimensionalen Raum), was den Begriff von gekrümmten Kurven ($d = 1$) verallgemeinert.

Gründlich besprochen werden diese Dinge in einer Vorlesung über *Differentialgeometrie*, in der analytische Methoden entwickelt werden, um geometrische Objekte (Kurven, Flächen, allgemeine gekrümmte “Mannigfaltigkeiten”) zu studieren. Dies bildet auch die Grundlage des mathematischen Verständnisses der *Allgemeinen Relativitätstheorie*, die die Raumzeit als eine gekrümmte vierdimensionale Mannigfaltigkeit beschreibt.

12 Differentialrechnung in mehreren Variablen

Wir wenden uns nun einem weiteren Kernthema der Analysis 2 zu, der Differentialrechnung von vektorwertigen Funktionen von mehreren Variablen. Wir werden also Funktionen der Art

$$f : U \subset \mathbb{R}^k \rightarrow \mathbb{R}^d$$

betrachten. Genau wie bei unserer Diskussion von Stetigkeit wird es vor allem darauf ankommen, mit den mehreren *Variablen* (also $k > 1$) umzugehen, während die mehreren *Komponenten* von f (also $d > 1$) keine größeren Schwierigkeiten machen werden.

12.1 Totale Ableitungen, Richtungsableitungen, partielle Ableitungen

Unsere Definition der Differenzierbarkeit von skalaren Funktionen *einer* Variable, also Funktionen $f : I \subset \mathbb{R} \rightarrow \mathbb{R}$, beruhte auf der Forderung nach Existenz des Grenzwertes (für $x \in I$)

$$\lim_{h \rightarrow 0} \frac{1}{h} (f(x+h) - f(x)).$$

Was ändert sich nun, wenn wir es mit $f : U \subset \mathbb{R}^k \rightarrow \mathbb{R}^d$ zu tun haben? Dann sollten x und h Vektoren in \mathbb{R}^k und $f(x+h)$, $f(x)$ Vektoren in \mathbb{R}^d sein. Der Ausdruck $f(x+h) - f(x)$ macht also Sinn als Vektor in \mathbb{R}^d , und wir haben auch einen guten Grenzwertbegriff für Folgen in \mathbb{R}^d .

Schwierig ist allerdings der Vorfaktor $\frac{1}{h}$, der für einen Vektor $h \in \mathbb{R}^k$, $k > 1$, keinen Sinn hat. Wir brauchen deshalb einen anderen Zugang zur Definition der Ableitung. Dazu erinnern wir uns an die affine Approximierbarkeit aus Analysis 1 (siehe S. 115/116 im Analysis 1 Skript): Eine Funktion $f : I \subset \mathbb{R} \rightarrow \mathbb{R}$ ist genau dann in $x \in I$ differenzierbar, wenn es eine Zahl $a_x \in \mathbb{R}$ und eine in einer Umgebung von $h = 0$ definierte Funktion r_x gibt, so dass

$$f(x+h) = f(x) + a_x \cdot h + r_x(h), \quad \lim_{h \rightarrow 0} \frac{r_x(h)}{h} = 0.$$

Die Zahl a_x ist dann eindeutig durch f und x bestimmt und stellt sich als die Ableitung $f'(x)$ an der Stelle x heraus. Der "Restterm" $r_x(h)$ verschwindet zu mindestens zweiter Ordnung für $h \rightarrow 0$, d.h. verschwindet auch noch für $h \rightarrow 0$, wenn wir ihn durch h teilen.

Da $\lim_{h \rightarrow 0} \frac{r_x(h)}{h} = 0$ äquivalent ist zu $\lim_{h \rightarrow 0} \frac{|r_x(h)|}{|h|} = 0$, eröffnet sich hier eine Möglichkeit zur Verallgemeinerung auf Vektoren $h \in \mathbb{R}^k$, wenn wir $|h|$ durch eine Norm $\|h\|$ ersetzen.

Die zweite Verallgemeinerung, die wir durchführen müssen, betrifft den linearen Term $a_x \cdot h$. Für einen Vektor h liegt es nahe, a_x durch eine Matrix, also durch eine lineare Abbildung A_x in $\text{Hom}(\mathbb{R}^k, \mathbb{R}^d)$ zu ersetzen. Das passt auch genau zu unserer geometrischen Vorstellung, dass zB der Graph einer Funktion $\mathbb{R}^2 \rightarrow \mathbb{R}$ nicht durch eine Tangentengerade an einem Punkt $(x, f(x))$ genähert werden kann, wohl aber durch eine Tangentialebene (\rightarrow lineare Abbildung in $\text{Hom}(\mathbb{R}^2, \mathbb{R})$).

Nach dieser Vorrede können wir die Definition von Differenzierbarkeit geben.

Definition 12.1. Sei $U \subset \mathbb{R}^k$ offen, und $x \in U$. Eine Funktion $f : U \rightarrow \mathbb{R}^d$ heißt *in x differenzierbar*, wenn es eine lineare Abbildung $A_x \in \text{Hom}(\mathbb{R}^k, \mathbb{R}^d)$ und eine auf einer Umgebung von 0 definierte Abbildung $r_x : U_\varepsilon(0) \subset \mathbb{R}^k \rightarrow \mathbb{R}^d$ gibt, so dass

$$f(x+h) = f(x) + A_x h + r_x(h), \quad \lim_{h \rightarrow 0} \frac{r_x(h)}{\|h\|} = 0.$$

Die lineare Abbildung A_x heißt die *Ableitung von f (oder: das Differential von f) in x* und wird auch mit $Df(x)$ (oder $df(x)$, oder auch $f'(x)$) bezeichnet^a.

Eine Funktion $f : U \rightarrow \mathbb{R}^d$ heißt *differenzierbar*, wenn sie in jedem $x \in U$ differenzierbar ist.

^aBeim Differenzieren gibt es immer eine Menge unterschiedliche Notationen.

Bemerkungen:

- Die Ableitung $Df(x)$ an einem Punkt x ist also nicht länger eine Zahl, sondern eine lineare Abbildung $Df(x) \in \text{Hom}(\mathbb{R}^k, \mathbb{R}^d)$, die wir nach Wahl von Basen in \mathbb{R}^k und \mathbb{R}^d mit einer $(d \times k)$ -Matrix identifizieren können. Wählen wir jeweils die Standardbasis des \mathbb{R}^k bzw \mathbb{R}^d , wird die darstellende Matrix *Jacobi-Matrix* (von f im Punkt x) genannt und mit $J_x(f)$ bezeichnet.

Für $k = d = 1$ erhalten wir eine (1×1) -Matrix, also eine Zahl - nämlich die bekannte skalare Ableitung $f'(x) \in \mathbb{R}$.

- Im Falle von $k = 1$, also einer Kurve $\gamma : \mathbb{R} \rightarrow \mathbb{R}^d$, ist die Ableitung $\gamma'(x) = D\gamma(x)$ eine $(d \times 1)$ -Matrix, also ein d -dimensionaler (Spalten-)Vektor. Dies passt genau zu unserer Definition der Ableitung einer Kurve.
- Die Ableitung einer differenzierbaren Funktion $f : U \subset \mathbb{R}^k \rightarrow \mathbb{R}^d$ ist also eine Abbildung

$$Df : U \rightarrow \text{Hom}(\mathbb{R}^k, \mathbb{R}^d).$$

Nach Wahl von Basen in \mathbb{R}^k und \mathbb{R}^d können wir Df also als eine matrixwertige Abbildung auffassen.

Da wir $\text{Hom}(\mathbb{R}^k, \mathbb{R}^d) \cong \mathbb{R}^{kd}$ identifizieren können, haben wir auf dem Zielraum $\text{Hom}(\mathbb{R}^k, \mathbb{R}^d)$ Normen zur Verfügung. Es macht also Sinn, über *stetig differenzierbare* Funktionen zu sprechen - das sind die differenzierbaren Funktionen, für die $Df : U \rightarrow \text{Hom}(\mathbb{R}^k, \mathbb{R}^d)$ stetig ist.

Es spielt dabei aufgrund der Äquivalenz aller Normen in endlichdimensionalen Vektorräumen keine Rolle, welche Norm wir zur Definition der Stetigkeit heranziehen. Ebenfalls spielt es keine Rolle, welche Normen auf \mathbb{R}^k und \mathbb{R}^d in Definition 12.1 zugrundegelegt werden.

- Wir wissen aus Lemma 10.2, dass Stetigkeit von Df äquivalent zur komponentenweisen Stetigkeit ist. Also: Eine differenzierbare Funktion $f : U \rightarrow \mathbb{R}^d$ ist stetig differenzierbar genau dann, wenn alle Komponentenfunktionen $J_x(f)_{ij}$, $1 \leq i \leq k$, $1 \leq j \leq d$, der Jacobimatrix stetige Funktionen sind.

- Die Menge aller stetig differenzierbaren Funktion $U \rightarrow \mathbb{R}^d$ bezeichnen wir mit $C^1(U, \mathbb{R}^d)$.
- In Definition 12.1 haben wir uns auf *offene* Definitionsbereiche U beschränkt, im Unterschied zu unserem Vorgehen im eindimensionalen Fall in Analysis 1. Das liegt daran, dass Ränder von Mengen in \mathbb{R}^n viel komplizierter sein können als in \mathbb{R} . Wir definieren hier also nur Differenzierbarkeit in inneren Punkten.

Beispiel 12.2.

- Ohne Schwierigkeiten sehen wir, dass konstante Funktionen $f : \mathbb{R}^k \rightarrow \mathbb{R}^d$ differenzierbar sind mit $Df(x) = 0$.
- Ist $f : \mathbb{R}^k \rightarrow \mathbb{R}^d$ *linear*, so ist f differenzierbar, und es gilt die auf den ersten Blick komisch aussehende Gleichung

$$Df(x) = f, \quad x \in \mathbb{R}^k.$$

Beachten Sie, dass f hier eine lineare Abbildung (Matrix) ist, $f(x)$ ein Vektor, aber $Df(x)$ ebenfalls eine lineare Abbildung (Matrix). Zum Beweis dieser Gleichung müssen wir nur $r_x = 0$ und $A_x = f(x)$ in die definierende Gleichung der Differenzierbarkeit einsetzen; wegen der Linearität von f ergibt sich dann $f(x+h) = f(x) + f(h) = f(x) + A_x h + 0$, genau wie gefordert.

Eine einfache Bemerkung zur Differenzierbarkeit ist, dass diese Eigenschaft komponentenweise geprüft werden kann:

Lemma 12.3. *Sei $U \subset \mathbb{R}^k$ offen, $x \in U$, und $f : U \rightarrow \mathbb{R}^d$ eine Funktion mit Komponentenfunktionen $(f_1, \dots, f_d) = f$. Dann ist f differenzierbar in x genau dann, wenn alle $f_1, \dots, f_d : U \rightarrow \mathbb{R}$ in x differenzierbar sind.*

Beweis. Wir nehmen zunächst an, dass f in x differenzierbar ist, d.h. dass $f(x+h) = f(x) + A_x h + r_x(h)$ wie in Def. 12.1 beschrieben gilt. Durch Skalarproduktbildung mit dem j -ten Standardbasisvektor e_j erhalten wir die j -te Komponente dieser Gleichung (merke $f_j(x) = \langle e_j, f(x) \rangle$), nämlich

$$f_j(x+h) = f_j(x) + \langle e_j, A_x h \rangle + \langle e_j, r_x(h) \rangle.$$

Nun ist $h \mapsto \langle e_j, A_x h \rangle = \langle A_x^T e_j, h \rangle$ eine lineare Abbildung $\mathbb{R}^k \rightarrow \mathbb{R}$, und für den Restterm gilt $\frac{\langle e_j, r_x(h) \rangle}{\|h\|} = \langle e_j, \frac{r_x(h)}{\|h\|} \rangle \rightarrow 0$ für $h \rightarrow 0$ aufgrund der Stetigkeit des Skalarprodukts. Also ist f_j in x differenzierbar.

Nehmen wir nun an, dass f_1, \dots, f_d alle in x differenzierbar sind. Dann gibt es lineare Abbildungen $A_x^{(j)} \in \text{Hom}(\mathbb{R}^k, \mathbb{R})$ und "Restterme" $r_x^{(j)} : U_\varepsilon(0) \rightarrow \mathbb{R}$, so dass

$$f_j(x+h) = f_j(x) + A_x^{(j)} h + r_x^{(j)}(h), \quad \lim_{h \rightarrow 0} \frac{r_x^{(j)}(h)}{\|h\|} = 0.$$

Wir definieren nun $A_x \in \text{Hom}(\mathbb{R}^k, \mathbb{R}^d)$ durch $A_x h := (A_x^{(1)} h, \dots, A_x^{(d)} h)$ und $r_x(h) := (r_x^{(1)} h, \dots, r_x^{(d)} h)$. Dann gilt $f(x+h) = f(x) + A_x h + r_x(h)$, und da Konvergenz einer Vektorfolge äquivalent zur Konvergenz aller ihrer Komponenten ist (Satz 8.14), folgt die Differenzierbarkeit von f in x . \square

An dieser Stelle wäre es noch mühselig, die Differenzierbarkeit von komplizierteren Funktionen wie zB $f(x_1, x_2) = (x_1^2, e^{x_1} + x_2, 1 + \sin(x_1x_2))$ zu prüfen und $Df(x_1, x_2)$ zu bestimmen. Dazu müssen wir erst die Theorie weiter entwickeln. Als ersten Schritt definieren wir nun Richtungsableitungen, ganz analog zu den bei unserer Diskussion von Stetigkeit definierten Richtungsgrenzwerten.

Definition 12.4. Sei $U \subset \mathbb{R}^k$ offen, $x \in U$, $h \in \mathbb{R}^k$ und $f : U \rightarrow \mathbb{R}^d$ eine Funktion. Die *Richtungsableitung von f im Punkt x in Richtung von h* ist der Grenzwert (falls er existiert)

$$(\partial_h f)(x) := \lim_{t \rightarrow 0} \frac{1}{t} (f(x + th) - f(x)).$$

Ist speziell $h = e_j$ ein Standardbasisvektor in \mathbb{R}^k , so heißt die entsprechende Richtungsableitung die *j -te partielle Ableitung von f in x* , und wird mit $(\partial_j f)(x)$ (oder $\frac{\partial f}{\partial x_j}(x)$ oder $(\partial_{x_j} f)(x)$) statt $(\partial_{e_j} f)(x)$ bezeichnet.

Eine Funktion, für die alle partiellen Ableitungen $(\partial_j f)(x)$ an einem Punkt x existieren, heißt *in x partiell differenzierbar*.

Die Limiten in dieser Definition sind genau analog zu unserer Definition der Ableitung einer Kurve. Beachten Sie, dass t hier eine Zahl (kein Vektor) ist.

Beispiel 12.5. Berechnen wir einige partielle Ableitungen. Wir beschränken uns zunächst auf skalare Funktionen $f : U \rightarrow \mathbb{R}$ (also nur eine Komponente, aber mehrere Variable $x = (x_1, \dots, x_k)$). Die Formel für die partielle Ableitung besagt

$$\begin{aligned} (\partial_j f)(x_1, \dots, x_k) &= \lim_{t \rightarrow 0} \frac{1}{t} (f(x + te_j) - f(x)) \\ &= \lim_{t \rightarrow 0} \frac{1}{t} (f(x_1, x_2, \dots, x_j + t, \dots, x_k) - f(x_1, \dots, x_k)). \end{aligned}$$

Dies ist genau die Ableitung von f nach der j -ten Komponente von x , d.h. Sie fixieren alle Variablen x_l , $l \neq j$, betrachten $f(x)$ nur in Abhängigkeit von x_j , und differenzieren nach x_j .

a) $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, $f(x_1, x_2) = x_1^2 e^{-x_2^2}$. Dann gilt

$$(\partial_1 f)(x_1, x_2) = 2x_1 e^{-x_2^2}, \quad (\partial_2 f)(x_1, x_2) = -2x_1^2 x_2 e^{-x_2^2}.$$

b) $g : \mathbb{R}^3 \rightarrow \mathbb{R}$, $g(x_1, x_2, x_3) = x_3 + 1$. Dann gilt

$$(\partial_1 g)(x_1, x_2, x_3) = (\partial_2 g)(x_1, x_2, x_3) = 0, \quad (\partial_3 g)(x_1, x_2, x_3) = 1.$$

Warnung: Partiiell differenzierbare Funktionen sind nicht notwendigerweise differenzierbar! Um diesen Unterschied sprachlich deutlich zu machen, spricht man auch manchmal von *totaler Differenzierbarkeit* (= Differenzierbarkeit) im Gegensatz zu bloßer *partieller Differenzierbarkeit*. Wir werden diesen Unterschied später genau klären.

Satz 12.6. Sei $U \subset \mathbb{R}^k$ offen, $x \in U$, und $f : U \rightarrow \mathbb{R}^d$ eine in x differenzierbare Funktion mit den Komponentenfunktionen $f(x) = (f_1(x), \dots, f_d(x))$, $f_j : U \rightarrow \mathbb{R}$. Dann existieren alle Richtungsableitungen von f in x (insbesondere alle partiellen Ableitungen), und die Jacobimatrix von f ist

$$J_x(f) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(x) & \frac{\partial f_1}{\partial x_2}(x) & \cdots & \frac{\partial f_1}{\partial x_k}(x) \\ \frac{\partial f_2}{\partial x_1}(x) & \frac{\partial f_2}{\partial x_2}(x) & \cdots & \frac{\partial f_2}{\partial x_k}(x) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_d}{\partial x_1}(x) & \frac{\partial f_d}{\partial x_2}(x) & \cdots & \frac{\partial f_d}{\partial x_k}(x) \end{pmatrix}.$$

Die Richtungsableitung von f am Punkt in Richtung $h \in \mathbb{R}^k$ ist

$$(\partial_h f)(x) = Df(x)h.$$

Beweis. Da f in x differenzierbar ist, ist der Grenzwert

$$\begin{aligned} \frac{1}{t}(f(x+th) - f(x)) &= \frac{1}{t}(Df(x)th + r_x(th)) \\ &= Df(x)h + \frac{1}{t}r_x(th) \rightarrow Df(x)h \quad \text{für } t \rightarrow 0. \end{aligned}$$

Dies zeigt die Existenz aller Richtungsableitungen in x als Konsequenz der Differenzierbarkeit in x und, dass die Richtungsableitung in x in Richtung h mit $Df(x)h$ übereinstimmt.

Kommen wir zur Jacobimatrix. Die j -te Spalte der Jacobimatrix erhalten wir durch Anwendung auf den j -ten Basisvektor e_j der Standardbasis, also $J_x(f)e_j = \partial_j f(x)$. Der Eintrag in der i -ten Zeile dieser j -ten Spalte ist $\langle e_i, J_x(f)e_j \rangle = \partial_j f_i(x) = J_x(f)_{ij}$. \square

Bemerkung: Dieser Satz räumt insbesondere eine kleine Schwierigkeit der Definition 12.1 des Differentials aus, die wir bisher übergangen haben: In Def. 12.1 nannten wir die lineare Abbildung A_x *das* Differential von f an der Stelle x und nicht *ein* Differential von f bei x . Das ist nur gerechtfertigt, wenn A_x durch f und x eindeutig festgelegt ist. Satz 12.6 sagt uns, dass für alle $h \in \mathbb{R}^k$ die Beziehung $A_x h = (\partial_h f)(x)$ gilt. Da die Richtungsableitungen $(\partial_h f)(x)$ per Definition eindeutig durch f und x festgelegt sind, gilt dies auch für A_x , und unsere Definition des Differentials hat keine Probleme.

Explizit durch Matrixmultiplikation sehen Jacobimatrizen und Richtungsableitungen so aus: Für $j \in \{1, \dots, d\}$ ist $(Df)(x)e_j$ die j -te Spalte der Jacobimatrix $J_x(f)$. Einen Richtungsvektor $h \in \mathbb{R}^k$ entwickeln wir in die Standardbasis, $h = \sum_{j=1}^k h_j \cdot e_j$, und erhalten so mit dem Standardskalarprodukt auf \mathbb{R}^d

$$\langle e_i, (Df)(x)h \rangle = \sum_{j=1}^d \langle e_i, (Df)(x)e_j \rangle h_j = \sum_{j=1}^d J_x(f)_{ij} h_j = (J_x(f)h)_i.$$

Beispiel 12.7.

a) Die Funktion $f : \mathbb{R}_+^3 \rightarrow \mathbb{R}^3$, $f(x_1, x_2, x_3) = (\sin(x_1 x_2), x_3^{x_2}, x_2)$ hat die Jaco-

bimatrix

$$J_x(f) = \begin{pmatrix} x_2 \cos(x_1 x_2) & x_1 \cos(x_1 x_2) & 0 \\ 0 & \log(x_3) x_3^{x_2} & x_2 x_3^{x_2-1} \\ 0 & 1 & 0 \end{pmatrix}.$$

Beachten Sie, dass die Existenz der Jacobimatrix (partielle Differenzierbarkeit) noch *kein* Beweis der Differenzierbarkeit von f ist. Aktuell können wir nur sagen: *Falls* diese Funktion f differenzierbar ist, so ist ihr Differential in den Standardbasen durch die obenstehende Jacobimatrix gegeben.

- b) Die Funktion $\kappa : \mathbb{R}^2 \rightarrow \mathbb{R}$, $\kappa(x_1, x_2) = (x_1^2 - x_2^2, |x_2| + \sin(x_1))$ ist nicht partiell differenzierbar, da ihre zweite Komponente $\kappa_2(x_1, x_2) = |x_2| + \sin(x_1)$ in $x_2 = 0$ nicht nach x_2 differenzierbar ist. Also ist κ erst recht nicht total differenzierbar.

Zum anschaulichen Verständnis der eingeführten Begriffe betrachten wir eine skalare Funktion $f : U \subset \mathbb{R}^k \rightarrow \mathbb{R}$, deren Graph Sie sich als gekrümmte Fläche ("Gebirge") über U vorstellen können (besonders anschaulich für $k = 2$). Die eindimensionale Ableitung aus Analysis 1 gibt die Steigung eines Funktionsgraphen an. In der hier betrachteten mehrdimensionalen Situation kann man nicht einfach von "der" Steigung sprechen - die Steigung an einem Punkt $x \in U$ wird im Allgemeinen davon abhängen, in welche Richtung h Sie schauen. Die Steigung am Punkt x in Richtung h ist genau $Df(x)h$. Das Differential $Df(x)$ informiert also über die Steigungen in alle Richtungen.

12.2 Ableitungsregeln

Aus Analysis 1 sind uns einige Ableitungsregeln geläufig: Linearität der Ableitung, Produktregel, Kettenregel, Quotientenregel, Ableitung der Umkehrfunktion.

Wir wollen jetzt einige dieser Regeln auf die mehrdimensionale Situation übertragen. Beweistechnisch ist es dabei bequem, eine mehrdimensionale Version unseres Lemmas 5.4 zur affinen Approximierbarkeit zu haben.

Lemma 12.8. Sei $U \subset \mathbb{R}^k$ offen und $x \in U$. Eine Funktion $f : U \rightarrow \mathbb{R}^d$ ist genau dann in x differenzierbar, wenn es eine offene Umgebung $U_\varepsilon(0) \subset \mathbb{R}^k$ von 0 und eine in 0 stetige Abbildung $\Phi : U_\varepsilon(0) \rightarrow \text{Hom}(\mathbb{R}^k, \mathbb{R}^d)$ gibt, so dass

$$f(x+h) = f(x) + \Phi(h)h, \quad h \in U_\varepsilon(0).$$

In diesem Fall gilt $\Phi(0) = (Df)(x)$.

Bemerkung: Genau wie in Lemma 5.4 ist Φ offenbar von x abhängig. Wir unterdrücken diese Abhängigkeit aber in unserer Notation, wenn x fest ist.

Beweis. Angenommen, eine Funktion Φ mit den angegebenen Eigenschaften existiert. Wir definieren $A := \Phi(0)$ und

$$r(h) := (\Phi(h) - \Phi(0))h = \Phi(h)h - Ah.$$

Dann gilt $\Phi(h)h = Ah + r(h)$, also

$$f(x + h) = f(x) + Ah + r(h),$$

und

$$\left\| \frac{r(h)}{\|h\|} \right\| = \frac{\|(\Phi(h) - \Phi(0))h\|}{\|h\|} \leq \|\Phi(h) - \Phi(0)\|,$$

wobei die Norm auf der rechten Seite die Operatornorm auf $\text{Hom}(\mathbb{R}^k, \mathbb{R}^d)$ bezeichnet (siehe Abschnitt 10.2). Da Φ stetig ist (in einer beliebigen Norm auf $\text{Hom}(\mathbb{R}^k, \mathbb{R}^d)$), gilt $\|\Phi(h) - \Phi(0)\| \rightarrow 0$, also $\frac{r(h)}{\|h\|} \rightarrow 0$ für $h \rightarrow 0$. Also ist f in x differenzierbar, und es gilt $(Df)(x) = A = \Phi(0)$.

Angenommen, f ist in x differenzierbar, d.h. $f(x + h) = f(x) + (Df)(x)h + r(h)$ mit $\frac{r(h)}{\|h\|} \rightarrow 0$ für $h \rightarrow 0$. Wir definieren $\Phi(h) \in \text{Hom}(\mathbb{R}^k, \mathbb{R}^d)$ gemäß

$$\Phi(h)v := \begin{cases} (Df)(x)v & h = 0 \\ (Df)(x)v + \langle h, v \rangle \frac{r(h)}{\|h\|^2} & h \neq 0 \end{cases}$$

Diese Definition ist das Analogon unserer im Beweis von Lemma 5.4 gemachten Definition. Wir haben für $h \neq 0$

$$\Phi(h)h = (Df)(x)h + \langle h, h \rangle \frac{r(h)}{\|h\|^2} = (Df)(x)h + r(h) = f(x + h) - f(x).$$

Um zu prüfen, dass Φ in $h = 0$ stetig ist, betrachten wir zuerst den Normabstand von Bildern ($h \neq 0$):

$$\|\Phi(h)v - \Phi(0)v\| = \left\| \langle h, v \rangle \frac{r(h)}{\|h\|^2} \right\| \leq \frac{\|r(h)\|}{\|h\|} \cdot \|v\|,$$

wobei wir die Cauchy-Schwarz Ungleichung verwendet haben. Da diese Ungleichung für alle Vektoren $v \in \mathbb{R}^k$ gilt, erhalten wir für die Operatornorm

$$\|\Phi(h) - \Phi(0)\| \leq \frac{\|r(h)\|}{\|h\|} \rightarrow 0 \text{ für } h \rightarrow 0,$$

was die Stetigkeit von Φ in 0 zeigt. □

Mit dieser Vorbereitung betrachten wir noch einmal die aus Analysis 1 bekannten Ableitungsregeln. Da Funktionen $f : U \subset \mathbb{R}^k \rightarrow \mathbb{R}^d$ linear kombiniert werden können, erwarten wir eine Ableitungsregel über Linearkombinationen. Falls die Definitionsbereiche (insbesondere die Dimensionen) passen, können wir auch Kompositionen $g \circ f$ betrachten; in diesem Fall erwarten wir auch eine Kettenregel.

Was eine Produktregel oder Quotientenregel sein soll, ist im mehrdimensionalen Kontext nicht ohne Weiteres klar, da \mathbb{R}^d keine Produktstruktur hat. Wir beginnen deshalb erstmal mit der Linearität und Kettenregel.

Satz 12.9.

- a) Sei $U \subset \mathbb{R}^k$ offen, $x \in U$, und $f, g : U \rightarrow \mathbb{R}^d$ in x differenzierbar. Dann sind für alle $\lambda, \mu \in \mathbb{R}$ auch die Funktionen $\lambda f + \mu g$ in x differenzierbar, und die Ableitung ist

$$(D(\lambda f + \mu g))(x) = \lambda(Df)(x) + \mu(Dg)(x).$$

- b) Seien $U \subset \mathbb{R}^k$ und $V \subset \mathbb{R}^d$ offen und $f : U \rightarrow \mathbb{R}^d, g : V \rightarrow \mathbb{R}^m$ Funktionen mit $f(U) \subset V$. Ist f in $x \in U$ differenzierbar und g in $f(x) \in V$ differenzierbar, so ist auch $g \circ f : U \rightarrow \mathbb{R}^m$ in x differenzierbar, mit Ableitung

$$(D(g \circ f))(x) = (Dg)(f(x)) \circ (Df)(x).$$

Das Symbol \circ auf der rechten Seite der Kettenregel ist eine Komposition von linearen Abbildungen (= Matrixmultiplikation nach Wahl von Basen), das meistens in der Notation unterdrückt wird, hier aber der Deutlichkeit halber ausgeschrieben wurde. Ausgedrückt in Jacobi-matrizen haben wir (merke: $f : U \subset \mathbb{R}^k \rightarrow \mathbb{R}^d, g : V \subset \mathbb{R}^d \rightarrow \mathbb{R}^m, g \circ f : U \subset \mathbb{R}^k \rightarrow \mathbb{R}^m$)

$$\underbrace{J_x(g \circ f)}_{(m \times k)} = \underbrace{J_{f(x)}(g)}_{(m \times d)} \underbrace{J_x(f)}_{(d \times k)}.$$

In der umgekehrten Reihenfolge ist das Produkt der Jacobimatrizen im Allgemeinen nicht wohldefiniert, weil die Spalten/Zeilenzahl der Matrizen nicht passt.

Beweis. a) Wir verwenden die Charakterisierung von Differenzierbarkeit aus Lemma 12.8. Nach Voraussetzung gibt es also in 0 stetige Funktionen $\Phi, \Psi : U_\varepsilon(0) \rightarrow \mathbb{R}^d$ mit $f(x+h) = f(x) + \Phi(h)h$ und $g(x+h) = g(x) + \Psi(h)h$. Dann ist auch $\Xi(h) := \lambda\Phi(h) + \mu\Psi(h)$ in 0 stetig, und es gilt

$$(\lambda f + \mu g)(x+h) - (\lambda f + \mu g)(x) = \lambda\Phi(h)h + \mu\Psi(h)h = \Xi(h)h.$$

Also ist $\lambda f + \mu g$ in x differenzierbar, mit Ableitung

$$(D(\lambda f + \mu g))(x) = \Xi(0) = \lambda\Phi(0) + \mu\Psi(0) = \lambda(Df)(x) + \mu(Dg)(x).$$

b) Wir verwenden wieder Lemma 12.8. Es gibt also zwei auf Umgebungen von 0 definierte in 0 stetige Abbildungen $\Phi : U_\varepsilon(0) \rightarrow \text{Hom}(\mathbb{R}^k, \mathbb{R}^d)$ und $\Psi : V_\varepsilon(0) \rightarrow \text{Hom}(\mathbb{R}^d, \mathbb{R}^m)$, so dass $f(x+h) = f(x) + \Phi(h)h$ und $g(f(x)+u) = g(f(x)) + \Psi(u)u$.

Für die Verkettung von g und f haben wir also

$$\begin{aligned} (g \circ f)(x+h) &= g(f(x) + (f(x+h) - f(x))) \\ &= g(f(x)) + \Psi(f(x+h) - f(x))(f(x+h) - f(x)) \\ &= g(f(x)) + \Psi(\Phi(h)h)\Phi(h)h. \end{aligned}$$

Hierbei ist $\Psi(\Phi(h)h)\Phi(h)$ ein Element von $\text{Hom}(\mathbb{R}^k, \mathbb{R}^m)$, und wir definieren die Abbildung $\Xi : U_\varepsilon(0) \rightarrow \text{Hom}(\mathbb{R}^k, \mathbb{R}^m), \Xi(h) := \Psi(\Phi(h)h)\Phi(h)$.

Da Φ und Ψ in 0 stetig sind und die Komposition

$$\text{Hom}(\mathbb{R}^k, \mathbb{R}^d) \times \text{Hom}(\mathbb{R}^d, \mathbb{R}^m) \rightarrow \text{Hom}(\mathbb{R}^k, \mathbb{R}^m), \quad (A, B) \mapsto A \circ B$$

stetig ist (Übung, hier braucht man die Operatornormabschätzung $\|A \circ B\| \leq \|A\| \|B\|$), ist auch Ξ in 0 stetig. Die Ableitung von $g \circ f$ in x ist also

$$(D(g \circ f))(x) = \lim_{h \rightarrow 0} \Xi(h) = \lim_{h \rightarrow 0} \Psi(\Phi(h)h)\Phi(h) = \Psi(0)\Phi(0) = (Dg)(f(x))(Df)(x).$$

□

Beispiel 12.10. Wir betrachten die Funktionen $f : \mathbb{R}^2 \rightarrow \mathbb{R}^3$ und $g : \mathbb{R}^3 \rightarrow \mathbb{R}^2$,

$$f(x_1, x_2) = \begin{pmatrix} x_1^2 \\ \cos(x_2 - x_1) \\ x_2 \end{pmatrix}, \quad g(y_1, y_2, y_3) = \begin{pmatrix} y_1 y_2 y_3 \\ y_2 + 1 \end{pmatrix}.$$

Dies sind differenzierbare Funktionen (kein Beweis hier, das sehen wir im nächsten Abschnitt), mit Jacobimatrizen

$$J_x(f) = \begin{pmatrix} 2x_1 & 0 \\ \sin(x_2 - x_1) & -\sin(x_2 - x_1) \\ 0 & 1 \end{pmatrix}, \quad J_y(g) = \begin{pmatrix} y_2 y_3 & y_1 y_3 & y_1 y_2 \\ 0 & 1 & 0 \end{pmatrix}.$$

Die Komposition von g und f ist die Abbildung $g \circ f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$,

$$(g \circ f)(x) = \begin{pmatrix} f_1(x) f_2(x) f_3(x) \\ f_2(x) + 1 \end{pmatrix} = \begin{pmatrix} x_1^2 x_2 \cos(x_2 - x_1) \\ \cos(x_2 - x_1) + 1 \end{pmatrix},$$

mit Jacobimatrix

$$J_x(g \circ f) = \begin{pmatrix} 2x_1 x_2 \cos(x_2 - x_1) + x_1^2 x_2 \sin(x_2 - x_1) & x_1^2 \cos(x_2 - x_1) - x_1^2 x_2 \sin(x_2 - x_1) \\ \sin(x_2 - x_1) & -\sin(x_2 - x_1) \end{pmatrix}.$$

Dies stimmt mit dem Matrixprodukt

$$\begin{aligned} J_{f(x)}(g) J_x(f) &= \begin{pmatrix} f_2(x) f_3(x) & f_1(x) f_3(x) & f_1(x) f_2(x) \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 2x_1 & 0 \\ \sin(x_2 - x_1) & -\sin(x_2 - x_1) \\ 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 2x_1 f_2(x) f_3(x) + \sin(x_2 - x_1) f_1(x) f_3(x) & -\sin(x_2 - x_1) f_1(x) f_3(x) + f_1(x) f_2(x) \\ \sin(x_2 - x_1) & -\sin(x_2 - x_1) \end{pmatrix} \\ &= \begin{pmatrix} 2x_1 x_2 \cos(x_2 - x_1) + x_1^2 x_2 \sin(x_2 - x_1) & x_1^2 \cos(x_2 - x_1) - x_1^2 x_2 \sin(x_2 - x_1) \\ \sin(x_2 - x_1) & -\sin(x_2 - x_1) \end{pmatrix} \end{aligned}$$

überein, genau wie von der Kettenregel behauptet.

Sei $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ differenzierbar. Wie hängt die Ableitung von $\mathbb{R} \ni t \mapsto f(t^2, e^t)$ mit den partiellen Ableitungen von f zusammen?

Beispiel 12.11. Ein wichtiger Spezialfall der Kettenregel ergibt sich für Funktionen einer Variablen. Sind $f : I \subset \mathbb{R} \rightarrow \mathbb{R}^k$ und $g : V \subset \mathbb{R}^k \rightarrow \mathbb{R}^d$, mit $f(I) \subset V$, differenzierbar bei $t \in I$ bzw $f(t) \in V$, so sind f und $g \circ f$ Kurven, und die Ableitung ist

$$(g \circ f)'(t) = J_{f(t)}(g)J_t(f) = \sum_{j=1}^k (\partial_j g)(f(t))f'_j(t).$$

Was Produktregeln angeht, so kennen wir bereits die Produktregel für Ableitungen von differenzierbaren Kurven $\gamma, \eta : I \subset \mathbb{R} \rightarrow \mathbb{R}^d$: Die Funktion $f : I \rightarrow \mathbb{R}$, $f(t) := \langle \gamma(t), \eta(t) \rangle$ ist differenzierbar mit Ableitung

$$f'(t) = \langle \gamma'(t), \eta(t) \rangle + \langle \gamma(t), \eta'(t) \rangle.$$

Allgemeinere Produktregeln erhält man, wenn man eine bilineare Abbildung $b : \mathbb{R}^{d_1} \times \mathbb{R}^{d_2} \rightarrow \mathbb{R}^m$ und differenzierbare Funktionen $f : \mathbb{R}^{k_1} \rightarrow \mathbb{R}^{d_1}$, $g : \mathbb{R}^{k_2} \rightarrow \mathbb{R}^{d_2}$ und das "Produkt" $\mathbb{R}^{k_1} \times \mathbb{R}^{k_2} \ni (x, y) \mapsto b(f(x), g(y))$ betrachtet. Die Details werden als Übung überlassen.

Seien $f : U \subset \mathbb{R}^k \rightarrow \mathbb{R}^d$ und $\varphi : U \rightarrow \mathbb{R}$ differenzierbar. Da φ skalar ist, ist hier das Produkt $\varphi \cdot f$ definiert. Zeigen Sie, dass $\varphi \cdot f$ differenzierbar ist, und beweisen Sie eine Produktregel.

Die Diskussion von Umkehrfunktionen und ihren Differentialen ist ein wichtiger Gegenstand der Vorlesung, auf den wir in einem späteren Kapitel eingehen werden.

Wir schließen diesen Abschnitt mit einigen Bemerkungen zu Gradienten und Höhenlinien.

Definition 12.12 (Gradient). Sei $U \subset \mathbb{R}^k$ offen und $f : U \rightarrow \mathbb{R}$ in $x \in U$ differenzierbar. Der *Gradient von f am Punkt x* ist der Zeilenvektor^a

$$\text{grad}f(x) := \nabla f(x) := J_x(f) = ((\partial_1 f)(x), \dots, (\partial_k f)(x)).$$

^aDas Symbol ∇ wird als "Nabla" ausgesprochen. Die PhysikerInnen unter Ihnen werden diesen "Nablaoperator" sicher schon gut kennen.

Der Gradient ist also nichts anderes als die Jacobimatrix von f im Spezialfall einer skalaren Funktion. Der Gradient hat eine geometrisch anschauliche Bedeutung: Er zeigt an der Stelle x in die Richtung des stärksten Anstiegs des Graphen $\{(x, f(x)) : x \in U\}$.

Lemma 12.13. Sei $f : U \subset \mathbb{R}^k \rightarrow \mathbb{R}$ in $x \in U$ differenzierbar.

a) Die Richtungsableitung von f in Richtung $h \in \mathbb{R}^k$ ist

$$(\partial_h f)(x) = (\nabla f(x))h = \langle (\nabla f)(x)^T, h \rangle.$$

b) Für $(\nabla f)(x) \neq 0$ gilt

$$\max_{h \in \mathbb{R}^k, \|h\|_2=1} (\partial_h f)(x) = (\partial_g f)(x), \quad g := \frac{(\nabla f)(x)^T}{\|(\nabla f)(x)\|_2}.$$

Beweis. a) Das ergibt sich sofort als Spezialfall von Satz 12.6.

b) Nach Teil a) haben wir $(\partial_g f)(x) = \langle \frac{(\nabla f)(x)^T}{\|(\nabla f)(x)\|_2}, (\nabla f)(x)^T \rangle = \|(\nabla f)(x)\|_2$, und für alle Einheitsvektoren $h \in \mathbb{R}^k$ durch Anwendung der Cauchy-Schwarz Ungleichung

$$(\partial_h f)(x) \leq |(\partial_h f)(x)| \leq \|\nabla f(x)\|_2 = (\partial_g f)(x).$$

□

Von Karten bergiger Gebiete kennen Sie *Höhenlinien*, also Linien konstanter Höhe. Dieses Konzept verallgemeinert sich auf skalare Funktionen $f : U \subset \mathbb{R}^k \rightarrow \mathbb{R}$, indem wir die *Höhenlinie der Höhe* $c \in \mathbb{R}$ als

$$H_c(f) := f^{-1}(\{c\}) = \{x \in U : f(x) = c\}$$

definieren. Dies sind nur “Linien” für $k = 2$, und können auch in diesem Fall aus mehreren Komponenten bestehen, leer sein, oder z.B. ein Punkt sein. Im Falle einer echten Höhenlinie steht diese stets senkrecht auf dem Gradienten, wie das folgende Beispiel zeigt.

Beispiel 12.14. Sei $f : U \subset \mathbb{R}^2 \rightarrow \mathbb{R}$ differenzierbar und $\gamma : I \rightarrow U$ ein Weg konstanter Höhe (ein Höhenweg), d.h. es gibt $c \in \mathbb{R}$, so dass $\gamma(t) \in H_c(f)$ für alle $t \in I$. Dann ist die Komposition $f(\gamma(t)) = c$ konstant, hat also verschwindende Ableitung. Per Kettenregel erhalten wir (vgl. Beispiel 12.11)

$$0 = \frac{d}{dt} f(\gamma(t)) = \langle (\nabla f)(\gamma(t))^T, \gamma'(t) \rangle.$$

Das heißt, dass der Gradient (steilster Anstieg) stets senkrecht auf den Höhenlinien (kein Anstieg) steht.

Gradienten spielen in unterschiedlichen Bereichen eine wichtige Rolle. In der Physik modellieren sie konservative Kraftfelder: Gegeben eine skalare differenzierbare Funktion $V : \mathbb{R}^3 \rightarrow \mathbb{R}$ (das Potential), so ist das Kraftfeld $F : \mathbb{R}^3 \rightarrow \mathbb{R}^3$, $F(x) := -\text{grad}V(x)$ ein sogenanntes konservatives Kraftfeld, d.h. es gilt ein Energieerhaltungssatz. Das Vorzeichen in der Definition von F ist hier dadurch motiviert, dass die Kraft in die Richtung des stärksten Abstiegs des Potentials zeigt (ein Ball rollt den Berg unter Einfluss der Gravitationskraft hinab, nicht hinauf). Diese Dinge lernen Sie in Vorlesungen über Differentialgleichungen oder Theoretische Physik.

In der Optimierung ist man daran interessiert, eine Funktion $f : \mathbb{R}^n \rightarrow \mathbb{R}$ zu maximieren/minimieren, zB den Gewinn einer Firma in Abhängigkeit von (vielen) Parametern. Die Grundidee des numerischen Gradientenverfahrens ist es, ausgehend von einem Startpunkt $x \in \mathbb{R}^n$ eine gewisse Schrittweite λ in Richtung des Gradienten von f zu $x + \lambda \text{grad}f(x)$ zu gehen und dieses Verfahren zu iterieren, um ein Maximum von f zu finden. Diese Dinge lernen Sie in Vorlesungen über Optimierung.

12.3 Vergleich von Differenzierbarkeitsbegriffen und Mittelwertsatz

Nachdem wir etwas Erfahrung mit totaler und partieller Differenzierbarkeit gesammelt haben, kommen wir nun zu einem systematischen Vergleich dieser Begriffe. Zuerst zeigen wir, dass totale Differenzierbarkeit Stetigkeit impliziert.

Lemma 12.15. Sei $U \subset \mathbb{R}^k$ offen und $f : U \rightarrow \mathbb{R}^d$ in $x \in U$ differenzierbar. Dann ist f in x stetig.

Beweis. Nach Voraussetzung gibt es eine auf einer Umgebung von $h = 0$ definierte Funktion r_x mit $\frac{r_x(h)}{\|h\|} \rightarrow 0$ für $h \rightarrow 0$ und

$$f(x+h) - f(x) = (Df)(x)h + r_x(h).$$

Die lineare Abbildung $(Df)(x) \in \text{Hom}(\mathbb{R}^k, \mathbb{R}^d)$ ist stetig, da \mathbb{R}^k endlichdimensional ist (Korollar 10.10). Also gilt $(Df)(x)h \rightarrow (Df)(x)0 = 0$ für $h \rightarrow 0$. Weiterhin haben wir $r_x(h) = \|h\| \frac{r_x(h)}{\|h\|} \rightarrow 0$ für $h \rightarrow 0$. Also gilt $f(x+h) - f(x) \rightarrow 0$ für $h \rightarrow 0$, dh f ist stetig in x . \square

Dieses Lemma ist uns als Lemma 5.3 bereits aus der eindimensionalen Situation bekannt. Beachten Sie, dass bloße *partielle* Differenzierbarkeit *nicht* Stetigkeit impliziert! Ein Beispiel ist die Funktion f aus Beispiel 10.4, siehe auch Aufgabe P7.3.

Insbesondere sehen wir an diesem Beispiel, dass partielle Differenzierbarkeit *nicht* totale Differenzierbarkeit impliziert. (Die andere Implikation “total \Rightarrow partiell” gilt nach Satz 12.6).

Verlangt man allerdings, dass alle partiellen Ableitungen existieren *und stetig sind* (eine solche Funktion heißt *stetig partiell differenzierbar*), ergibt sich Differenzierbarkeit als Konsequenz.

Satz 12.16. Sei $U \subset \mathbb{R}^k$ offen, $x \in U$, $f : U \rightarrow \mathbb{R}^d$ partiell differenzierbar, und alle partiellen Ableitungen $\partial_j f_i : U \rightarrow \mathbb{R}$ seien in x stetig. Dann ist f in x (stetig) differenzierbar.

Beweis. Da Differenzierbarkeit komponentenweise geprüft werden kann (Lemma 12.3), dürfen wir ohne Beschränkung der Allgemeinheit $d = 1$ annehmen.

Sei $\varepsilon > 0$ so, dass $U_\varepsilon(x) \subset U$ (mit der ε -Umgebung in 1-Norm), und sei $h \in \mathbb{R}^k$, $\|h\|_1 < \varepsilon$. Mit Hilfe der Entwicklung $h = \sum_{j=1}^k h_j e_j$ definieren wir die Vektoren

$$y_j := x + \sum_{i=1}^j h_i e_i, \quad j = 0, 1, \dots, k.$$

Insbesondere gilt $y_0 = x$ und $y_k = x + h$, und wir haben die Abschätzung in 1-Norm

$$\|x - y_j\|_1 = \left\| \sum_{i=1}^j h_i e_i \right\|_1 \leq \sum_{i=1}^j |h_i| \leq \|h\|_1 < \varepsilon,$$

aus der wir $y_j \in U_\varepsilon(x) \subset U$ schließen. Wir können also die Teleskopsumme

$$f(x+h) - f(x) = f(y_k) - f(y_0) = \sum_{i=1}^k (f(y_{i-1} + h_i e_i) - f(y_{i-1}))$$

betrachten. Die Funktionen $g_i(t) := f(y_{i-1} + te_i)$ sind auf $t \in (-\delta, \delta)$ für hinreichend kleines $\delta > 0$ definiert. Da f partiell differenzierbar ist, sind die g_i differenzierbar, mit Ableitung

$$g_i'(t) = (\partial_i f)(y_{i-1} + te_i).$$

Wir verwenden nun den Mittelwertsatz 5.16, um (von x und h abhängige) Zwischenstellen $\tau_i \in (0, h_i)$ zu finden, so dass

$$f(x+h) - f(x) = \sum_{i=1}^k (g_i(h_i) - g_i(0)) = \sum_{i=1}^k h_i g_i'(\tau_i) = \sum_{i=1}^k h_i (\partial_i f)(y_{i-1} + \tau_i e_i).$$

Jetzt kommen wir zum Beweis der totalen Differenzierbarkeit von f in x . Wir wissen bereits: Falls f in x differenzierbar ist, so muss die Ableitung (Gradient) von f in x durch die partiellen Ableitungen $(\partial_j f)(x)$ gegeben sein. Wir definieren deshalb $\Phi(h) \in \text{Hom}(\mathbb{R}^k, \mathbb{R})$ via

$$\Phi(h)v := \sum_{i=1}^k (\partial_i f)(y_{i-1} + \tau_i e_i) v_i.$$

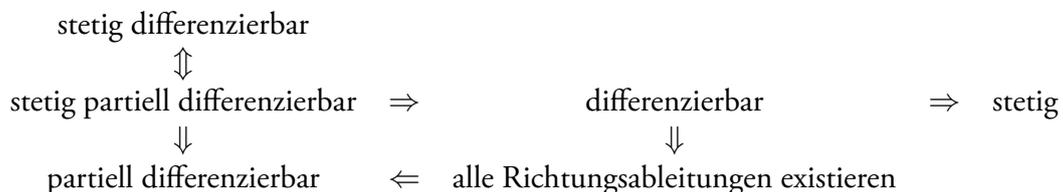
Dieses $\Phi(h)$ ist offenbar eine lineare Abbildung $\mathbb{R}^k \rightarrow \mathbb{R}$ und erfüllt $f(x+h) - f(x) = \Phi(h)h$. Gemäß Lemma 12.8 müssen wir nur zeigen, dass Φ in 0 stetig ist. Hier geht die angenommene Stetigkeit der partiellen Ableitungen ein: Im Limes $h \rightarrow 0$ haben wir $y_i \rightarrow x$ per Definition von y_i , und $\tau_i \rightarrow 0$, da $\tau_i \in (0, h_i)$. Da auch die partiellen Ableitungen stetig sind, erhalten wir so

$$\lim_{h \rightarrow 0} \Phi(h) = \lim_{h \rightarrow 0} \sum_{i=1}^k (\partial_i f)(y_{i-1} + \tau_i e_i) = \sum_{i=1}^k (\partial_i f)(x) = \Phi(0).$$

Also ist Φ in 0 stetig, was den Beweis abschließt. □

Bemerkung: Eine partiell differenzierbare Funktion $f : U \rightarrow \mathbb{R}^d$, für die alle partiellen Ableitungen stetig sind, wird auch stetig partiell differenzierbare Funktion genannt. Wie eben gezeigt, sind solche Funktionen auch total differenzierbar. Da die Ableitung $(Df)(x)$ durch die Jacobimatrix gegeben ist, deren Einträge die partiellen Ableitungen sind, ergibt sich, dass $(Df)(x)$ in diesem Fall stetig von x abhängt. Kurz gesagt: Stetig partiell differenzierbar und stetig total differenzierbar sind ein und derselbe Begriff.

Als Überblick noch einmal die verschiedenen Differenzierbarkeitsbegriffe als Diagramm:



Alle in diesem Diagramm nicht eingetragenen Implikationen sind im Allgemeinen falsch. Es gibt also partiell differenzierbare aber nicht differenzierbare Funktionen, oder nicht differenzierbare Funktionen, für die alle Richtungsableitungen existieren. Beispiele dazu finden sich in den Übungen.

Beispiel 12.17. Die Funktionen in Beispiel 12.7 a) und 12.10 sind stetig differenzierbar, da wir sie als stetig differenzierbar erkennen.

Wir wenden uns nun dem Mittelwertsatz zu. In Analysis 1 hatten wir gelernt, dass für differenzierbares $f : (a, b) \rightarrow \mathbb{R}$ ein $\xi \in (a, b)$ existiert, so dass $f(b) - f(a) = f'(\xi) \cdot (b - a)$ (Satz 5.16 b)).

Wir werden zwei Versionen dieses Satzes im Mehrdimensionalen geben. Die erste Version ist eine direkte Verallgemeinerung, die aber nur für skalare Funktionen gilt.

Satz 12.18 (Mittelwertsatz für skalare Funktionen). Sei $U \subset \mathbb{R}^k$ offen und $f : U \rightarrow \mathbb{R}$ differenzierbar. Sei $x \in U$ und $h \in \mathbb{R}^k$ so, dass die gerade Verbindungsstrecke $x + [0, 1]h = \{x + th : 0 \leq t \leq 1\}$ komplett in U liegt. Dann existiert ein $\tau \in (0, 1)$, so dass

$$f(x + h) - f(x) = (Df)(x + \tau h)h.$$

Beweis. Die Funktion $g : [0, 1] \rightarrow \mathbb{R}$, $g(t) := f(x + th)$ ist differenzierbar (Kettenregel), und hat die Ableitung $g'(t) = (Df)(x + \tau h)h$. Nach dem Mittelwertsatz der Analysis 1 gibt es also ein $\tau \in (0, 1)$ mit

$$(Df)(x + \tau h) = g'(\tau) = g(1) - g(0) = f(x + h) - f(x).$$

□

Dieser Satz ist im Allgemeinen für vektorwertige Funktionen $f : U \rightarrow \mathbb{R}^d$, $d > 1$, falsch! Wir illustrieren dies an einem Beispiel.

Beispiel 12.19. Sei $\gamma : \mathbb{R} \rightarrow \mathbb{R}^2$ der Kreisweg $\gamma(t) = (\cos t, \sin t)$. Dann gilt $\gamma(0) = \gamma(2\pi)$. Die Ableitung ist $\gamma'(t) = (-\sin t, \cos t)$, insbesondere $\|\gamma'(t)\|_2 = 1$, also $\gamma'(t) \neq 0$ für alle t . Also kann $\gamma(0 + 2\pi) - \gamma(0) = 0$ nie mit $\gamma'(\tau 2\pi)2\pi$ übereinstimmen.

Der Mittelwertsatz wird allerdings in den meisten Fällen nur dazu verwendet, um Abschätzungen herzuleiten. Sehr oft geht das in der Form: Schreibe eine Differenz von Funktionswerten $f(x + h) - f(x)$ als Integral über die Ableitung, schätze dann durch Länge des Integrationsintervalls mal Supremum des Betrags der Ableitung ab.

Dieses Vorgehen funktioniert auch in höheren Dimensionen.

Satz 12.20. Sei $U \subset \mathbb{R}^k$ offen und $f : U \rightarrow \mathbb{R}^d$ stetig differenzierbar.

a) Sei $x \in U$ und $h \in \mathbb{R}^k$ so, dass die gerade Verbindungsstrecke $x + [0, 1]h = \{x + th : 0 \leq t \leq 1\}$ komplett in U liegt. Dann gilt

$$f(x + h) - f(x) = \int_0^1 (Df)(x + th)h dt,$$

wobei das Integral über die vektorwertige Funktion komponentenweise zu verstehen ist, d.h. die j -te Komponente von $\int_0^1 (Df)(x + th)h dt$ des \mathbb{R}^d -Vektors ist per Definition gleich $\int_0^1 \langle e_j, (Df)(x + th)h \rangle dt$.

b) (**Schrankensatz**) Gibt es weiterhin eine Konstante $L > 0$, so dass $\|(Df)(x+th)\| \leq L$ (mit irgendeiner Norm auf $\text{Hom}(\mathbb{R}^k, \mathbb{R}^d)$) für alle $0 \leq t \leq 1$, so gibt es $C > 0$, so dass die Abschätzung

$$\|f(x+h) - f(x)\| \leq CL \cdot \|h\|$$

gilt (mit beliebigen Normen auf \mathbb{R}^k und \mathbb{R}^d ; die Konstante C hängt von der Wahl der Normen ab). Sind alle Normen Euklidisch bzw die auf den Euklidischen Normen basierende Operatornorm, so gilt $C = 1$.

Beweis. a) Wir betrachten einen Einheitsvektor $u \in \mathbb{R}^d$, $\|u\|_2 = 1$, und die Hilfsfunktion $g(t) := \langle f(x+th), u \rangle$, die differenzierbar ist mit Ableitung $g'(t) = \langle u, (Df)(x+th)h \rangle$. Es gilt nach dem Hauptsatz der Differential- und Integralrechnung

$$\langle u, f(x+h) - f(x) \rangle = g(1) - g(0) = \int_0^1 g'(t) dt = \int_0^1 \langle u, (Df)(x+th)h \rangle dt.$$

Setzen wir $u = e_j$, so erhalten wir genau die j -te Komponente der behaupteten Gleichung.

b) Da alle Normen äquivalent sind, ist klar, dass es genügt, die Aussage für eine spezielle Wahl von Normen zu zeigen. Wir nehmen nun Beträge und schätzen nacheinander per Dreiecksungleichung für Integrale (Satz 6.11), Cauchy-Schwarz Ungleichung, und der Annahme ab. Dies ergibt

$$\begin{aligned} |\langle u, f(x+h) - f(x) \rangle| &\leq \int_0^1 |\langle u, (Df)(x+th)h \rangle| dt \\ &\leq \int_0^1 \|(Df)(x+th)h\|_2 dt \\ &\leq L \cdot \|h\|_2. \end{aligned}$$

Um den Beweis abzuschließen, bemerken wir, dass für $f(x+h) = f(x)$ nichts zu zeigen ist. Für $f(x+h) \neq f(x)$ betrachten wir den Einheitsvektor $u := \frac{f(x+h) - f(x)}{\|f(x+h) - f(x)\|_2}$. Eingesetzt in die obige Abschätzung ergibt sich

$$\|f(x+h) - f(x)\|_2 = \left\langle \frac{f(x+h) - f(x)}{\|f(x+h) - f(x)\|_2}, f(x+h) - f(x) \right\rangle \leq L \|h\|_2.$$

□

Wenn Sie statt der Euklidischen eine andere Norm erhalten, so bekommen Sie das gleiche Ergebnis, nur mit einer möglicherweise anderen Konstante $C \cdot L$ statt L in der Abschätzung.

Überzeugen Sie sich, dass das Beispiel 12.19 keinen Widerspruch zu diesem Satz bildet.

Der Beweis von Teil b) dieses Satzes enthielt einen guten Abschätzungstrick, den wir hier für spätere Verwendung notieren.

Korollar 12.21. Sei $U \subset \mathbb{R}^k$, $x \in U$, $h \in \mathbb{R}^k$ und $[a, b]$ ein kompaktes Intervall so, dass $\{x + th : t \in [a, b]\} \subset U$. Dann gilt für das (komponentenweise definierte) Integral über eine stetige Funktion $f : U \rightarrow \mathbb{R}^k$

$$\left\| \int_a^b f(x + th) dt \right\|_2 \leq \int_a^b \|f(x + th)\|_2 dt. \quad (12.1)$$

Beweis. Sei $u \in \mathbb{R}^d$, $\|u\|_2 = 1$, ein Einheitsvektor. Dann haben wir per Standard-Integralabschätzung und Cauchy-Schwarz-Ungleichung mit dem Euklidischen Skalarprodukt

$$\begin{aligned} \left| \left\langle u, \int_a^b f(x + th) dt \right\rangle \right| &= \left| \int_a^b \sum_{j=1}^d f_j(x + th) u_j dt \right| \leq \int_a^b |\langle u, f(x + th) \rangle| dt \\ &\leq \int_a^b \|f(x + th)\|_2 dt. \end{aligned}$$

Ist $\int_a^b f(x + th) dt = 0$, so ist nichts zu zeigen. Für $\int_a^b f(x + th) dt \neq 0$ setzen wir den Einheitsvektor $u := \left\| \int_a^b f(x + th) dt \right\|^{-1} \int_a^b f(x + th) dt$ in die obige Abschätzung ein. Da die linke Seite dann gleich $\left| \left\langle u, \int_a^b f(x + th) dt \right\rangle \right| = \left\| \int_a^b f(x + th) dt \right\|_2$ ist, folgt die Behauptung. \square

Beispiel 12.22. Die Euklidischen Normen sind geometrisch am natürlichsten. Für Abschätzungen können auch andere Normen, zB ∞ -Normen, hilfreich sein. Wir illustrieren das hier an einem Beispiel. Dazu verwenden wir die sogenannte Zeilensummennorm auf $\text{Hom}(\mathbb{R}^k, \mathbb{R}^d) \cong \mathbb{R}^{k \times d}$, die durch

$$\|A\|_{\text{zs}} := \max_{i=1..d} \sum_{j=1}^k |A_{ij}|,$$

definiert ist (Übung: Dies ist eine Norm). Dann gilt der vorherige Satz ebenfalls mit $C = 1$ (Übung).

Als konkretes Beispiel betrachten wir nun

$$f : \mathbb{R}^2 \rightarrow \mathbb{R}^2, \quad f(x, y) := \begin{pmatrix} \frac{\arctan(x+1)}{(\sin y+3)^2} \\ \frac{1}{4} e^{\sin(x+y)/3} \end{pmatrix}.$$

Durch Ableiten und Abschätzen können Sie zeigen, dass die Jacobimatrix von f

$$\|J_{(x,y)}(f)\|_{\text{zs}} \leq \frac{3}{4}$$

erfüllt. Damit sagt uns Satz 12.20 $\|f(x, y) - f(x', y')\|_\infty \leq \frac{3}{4} \|(x, y) - (x', y')\|_\infty$ für alle $(x, y), (x', y') \in \mathbb{R}^2$, d.h. f ist eine strikte Kontraktion.

Mit dem Banachschen Fixpunktsatz schließen wir nun, dass es ein eindeutiges $(x, y) \in \mathbb{R}^2$ mit $f(x, y) = (x, y)$ gibt, also eine eindeutige Lösung der Gleichung

$$\arctan(x + 1) = x(\sin y + 3)^2 \quad e^{\sin(x+y)/3} = 4y.$$

Das ist durch direktes Betrachten dieses Gleichungssystems wohl nicht offensichtlich!

12.4 Höhere Ableitungen und Vektor-/Gradientenfelder

Nachdem wir ein gutes Verständnis des Differentials einer differenzierbaren Funktion $f : U \subset \mathbb{R}^k \rightarrow \mathbb{R}^d$ erlangt haben, wenden wir uns nun der Diskussion von höheren Ableitungen zu. Wir beginnen mit einer Diskussion von höheren partiellen Ableitungen.

Definition 12.23. Sei $f : U \subset \mathbb{R}^k \rightarrow \mathbb{R}^d$ eine partiell differenzierbare Funktion.

- a) Wenn die partiellen Ableitungen $\partial_j f : \mathbb{R}^k \rightarrow \mathbb{R}^d$ wieder partiell differenzierbar sind, so nennen wir ihre partiellen Ableitungen

$$(\partial_i \partial_j f)(x) := (\partial_i(\partial_j f))(x) =: \frac{\partial^2 f}{\partial x_i \partial x_j}(x)$$

die *partiellen Ableitungen 2. Ordnung* von f .

- b) Induktiv definieren wir eine n -mal (stetig) partiell differenzierbare Funktion $f : U \subset \mathbb{R}^k \rightarrow \mathbb{R}^d$ als eine Funktion, die $(n - 1)$ -mal partiell differenzierbar ist, so dass alle partiellen Ableitungen $(n - 1)$ ter Ordnung wieder (stetig) partiell differenzierbar sind, dh dass die partiellen Ableitungen n -ter Ordnung

$$(\partial_{i_1} \partial_{i_2} \cdots \partial_{i_n} f)(x)$$

für alle $i_1, \dots, i_n \in \{1, \dots, k\}$ existieren (und stetig sind).

- c) Die Menge aller n -mal stetig partiell differenzierbaren Funktionen $f : U \rightarrow \mathbb{R}^d$ wird mit $C^n(U, \mathbb{R}^d)$ bezeichnet.
d) Die Menge aller *glatten* Funktionen ist

$$C^\infty(U, \mathbb{R}^d) := \bigcap_{n \in \mathbb{N}} C^n(U, \mathbb{R}^d).$$

Es gibt also oft eine große Menge an partiellen Ableitungen.

Beispiel 12.24. Wir betrachten $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, $f(x, y) := xe^y + x^2$. Dann gilt

$$\begin{aligned} (\partial_1 f)(x, y) &= e^y + 2x, & (\partial_2 f)(x, y) &= xe^y, \\ (\partial_1^2 f)(x, y) &= 2, & (\partial_2^2 f)(x, y) &= xe^y, \\ (\partial_1 \partial_2 f)(x, y) &= e^y = (\partial_2 \partial_1 f)(x, y). \end{aligned}$$

In diesem Beispiel existieren sogar alle partiellen Ableitungen *beliebiger* Ordnung (zB

$(\partial_2^7 \partial_1^8 f)(x, y) = 0$ und $(\partial_2^{100} \partial_1 \partial_2 f)(x, y) = e^y$, das heißt $f \in C^\infty(\mathbb{R}^2, \mathbb{R})$.

An dem Beispiel fällt auf, dass die beiden gemischten partiellen Ableitungen 2. Ordnung, also $\partial_1 \partial_2 f$ und $\partial_2 \partial_1 f$, übereinstimmen.

Der Satz von Schwarz besagt, dass dies oft der Fall ist.

Satz 12.25 (Satz von Schwarz). Sei $U \subset \mathbb{R}^k$ offen und $f \in C^2(U, \mathbb{R}^d)$ (zweimal stetig partiell differenzierbar). Dann gilt für alle $i, j \in \{1, \dots, k\}$

$$\partial_i \partial_j f = \partial_j \partial_i f.$$

Beweis. Die behauptete Gleichung kann komponentenweise geprüft werden. Wir dürfen also annehmen, dass $f : U \rightarrow \mathbb{R}$ skalar ist.

Sei $x \in U$. Da U offen ist, gibt es eine δ -Umgebung $U_\delta^\infty(x)$ in der Norm $\|\cdot\|_\infty$, die ganz in U liegt. Das impliziert $x + te_i + se_j \in U$ für alle $s, t \in (-\delta, \delta)$. (Wie immer bezeichnen e_i, e_j die Standardbasisvektoren im \mathbb{R}^k .) Wir betrachten die auf einem Quadrat definierte Funktion

$$\psi : (-\delta, \delta)^2 \rightarrow \mathbb{R}^d, \quad \psi(t, s) := f(x + te_i + se_j).$$

Dann ist ψ zweimal stetig partiell differenzierbar, mit $(\partial_1 \partial_2 \psi)(0, 0) = (\partial_i \partial_j f)(x)$ und $(\partial_2 \partial_1 \psi)(0, 0) = (\partial_j \partial_i f)(x)$. Wir können uns also ganz auf die Funktion ψ konzentrieren und behaupten, dass es einen Punkt $(\sigma, \tau) \in (-\delta, \delta)^2$ gibt, so dass

$$\psi(s, t) - \psi(0, t) - \psi(s, 0) + \psi(0, 0) = ts \cdot (\partial_1 \partial_2 \psi)(\sigma, \tau) \quad (\star)$$

gilt. Diese Behauptung folgt aus dem Mittelwertsatz für Funktionen einer Variablen aus Analysis 1 (Satz 5.16 b): Setzen wir $\zeta_s(t) := \psi(s, t) - \psi(0, t)$, so ist die linke Seite von (\star) gleich $\zeta_s(t) - \zeta_s(0)$. Dies ist eine in t differenzierbare Funktion, wir finden also ein (von s und t abhängiges) τ zwischen 0 und t , so dass $\zeta_s(t) - \zeta_s(0) = t \zeta'_s(\tau) = t(\partial_2 \psi(s, \tau) - \partial_2 \psi(0, \tau))$. Diese Funktion ist in s differenzierbar, dh wir finden durch nochmalige Anwendung des Mittelwertsatzes ein σ zwischen 0 und s , so dass (\star) gleich $ts \cdot (\partial_1 \partial_2 \psi)(\sigma, \tau)$ ist.

Andererseits können wir auch $\xi_t(s) := \psi(s, t) - \psi(s, 0)$ definieren. Dann ist die linke Seite von (\star) gleich $\xi_t(s) - \xi_t(0)$, und wir können wieder zweimal den Mittelwertsatz anwenden: Zuerst finden wir σ' zwischen 0 und s , so dass $\xi_t(s) - \xi_t(0) = s \xi'_t(\sigma') = s(\partial_1 \psi(\sigma', t) - \partial_1 \psi(\sigma', 0))$, und eine zweite Anwendung des MWS ergibt ein τ' zwischen 0 und t , so dass die linke Seite von (\star) gleich $ts \cdot (\partial_2 \partial_1 \psi)(\sigma', \tau')$ ist.

Wir erhalten also zwei Punkte $(\sigma, \tau), (\sigma', \tau') \in (-\delta, \delta)^2$, so dass

$$(\partial_1 \partial_2 \psi)(\sigma, \tau) = (\partial_2 \partial_1 \psi)(\sigma', \tau').$$

Nun benutzen wir die Voraussetzung, dass $\partial_1 \partial_2 \psi$ und $\partial_2 \partial_1 \psi$ stetig sind, da f als zweimal stetig partiell differenzierbar vorausgesetzt wurde. Zu jedem $\varepsilon > 0$ können wir also unser den Definitionsbereich von ψ bestimmendes $\delta > 0$ so klein wählen, dass

$$|\partial_1 \partial_2 \psi(\sigma, \tau) - \partial_1 \partial_2 \psi(0, 0)| < \varepsilon, \quad |\partial_2 \partial_1 \psi(\sigma', \tau') - \partial_2 \partial_1 \psi(0, 0)| < \varepsilon.$$

Das ergibt

$$|\partial_1\partial_2\psi(0,0) - \partial_2\partial_1\psi(0,0)| \leq |\partial_1\partial_2\psi(0,0) - \partial_1\partial_2\psi(\sigma,\tau)| + |\partial_2\partial_1\psi(\sigma',\tau') - \partial_2\partial_1\psi(0,0)| < 2\varepsilon,$$

und damit $\partial_2\partial_1\psi(0,0) = \partial_1\partial_2\psi(0,0)$. \square

Mit einer Beweisvariante lässt sich auch zeigen, dass die Existenz der partiellen Ableitungen $\partial_i f$, $\partial_j f$ und $\partial_i\partial_j f$ sowie die Stetigkeit von $\partial_i\partial_j f$ bereits die Existenz von $\partial_j\partial_i f$ und die Identität $\partial_i\partial_j f = \partial_j\partial_i f$ impliziert.

Es gibt allerdings zweimal partiell differenzierbare (aber nicht zweimal stetig partiell differenzierbare) Funktionen f mit $\partial_i\partial_j f \neq \partial_j\partial_i f$. Die Stetigkeit der zweiten partiellen Ableitungen ist also eine wesentliche Annahme im Satz von Schwarz.

Der Satz von Schwarz findet in der Vektoranalysis vielfältige Anwendung, siehe die Differentialoperatoren div , grad , rot , die in Aufgabe H8.1 definiert sind.

Definition 12.26. Ein *Vektorfeld* ist eine Funktion $f : U \subset \mathbb{R}^k \rightarrow \mathbb{R}^k$. Ein *Gradientenfeld* ist ein Vektorfeld der Form $f = \operatorname{grad}V = (\partial_1 V, \dots, \partial_k V)$ für eine partiell differenzierbare Funktion $V : U \subset \mathbb{R}^k \rightarrow \mathbb{R}$.

Ein Potential eines Gradientenfeldes f spielt in gewissem Sinne die Rolle einer Stammfunktion von f . Die folgenden Beispiele belegen, dass dieser Begriff in mehreren Dimensionen komplizierter ist als in einer Dimension; nicht alle Vektorfelder haben Stammfunktionen (Potentiale).

Beispiel 12.27.

- a) Sei $V : \mathbb{R}^k \setminus \{0\} \rightarrow \mathbb{R}$, $V(x) := \frac{1}{\|x\|}$ (mit der Euklidischen Norm). Dann gilt $\partial_i \frac{1}{\|x\|} = \partial_i (x_1^2 + \dots + x_k^2)^{-1/2} = -x_i (x_1^2 + \dots + x_k^2)^{-3/2}$, also

$$\operatorname{grad}V(x) = -\frac{x}{\|x\|^3}.$$

Potentiale dieser Form spielen in der klassischen Gravitationstheorie und in der Elektrostatik eine große Rolle, hier modelliert $\operatorname{grad}V$ (bis auf einen Vorfaktor) die Gravitationskraft bzw elektrostatische Kraft.

- b) Jedes Gradientenfeld ist ein Vektorfeld, aber für Dimension $k > 1$ ist nicht jedes Vektorfeld ein Gradientenfeld. Angenommen, $f : \mathbb{R}^k \rightarrow \mathbb{R}^k$ ist ein C^1 -Gradientenfeld, d.h. es gibt $V \in C^2(\mathbb{R}^k, \mathbb{R})$, so dass $f_i(x) = \partial_i V(x)$. Dann gilt für alle $i, j \in \{1, \dots, k\}$ nach dem Satz von Schwarz

$$\partial_j f_i(x) = \partial_j \partial_i V(x) = \partial_i \partial_j V(x) = \partial_i f_j(x).$$

Eine notwendige Bedingung für die Existenz eines Potentials ist also

$$\partial_i f_j = \partial_j f_i, \quad i, j \in \{1, \dots, k\}. \quad (**)$$

Ist diese Bedingung auch hinreichend? Im Allgemeinen nicht, wie folgendes Bei-

spiel zeigt. Wir betrachten das auf $\mathbb{R}^2 \setminus \{0\}$ definierte C^1 -Vektorfeld

$$f(x_1, x_2) := \begin{pmatrix} -\frac{x_2}{x_1^2 + x_2^2} \\ \frac{x_1}{x_1^2 + x_2^2} \end{pmatrix}.$$

Eine kleine Rechenübung zeigt $(\partial_1 f_2)(x_1, x_2) = \frac{-x_1^2 + x_2^2}{(x_1^2 + x_2^2)^2} = (\partial_2 f_1)(x_1, x_2)$, die Bedingung $(\star\star)$ ist also erfüllt.

Angenommen, $f = \text{grad}V$ für eine partiell differenzierbare Funktion (Potential) $V : \mathbb{R}^2 \setminus \{0\} \rightarrow \mathbb{R}$, d.h.

$$(\partial_1 V)(x_1, x_2) = -\frac{x_2}{x_1^2 + x_2^2}, \quad (\partial_2 V)(x_1, x_2) = \frac{x_1}{x_1^2 + x_2^2}.$$

Erinnern wir uns daran, dass der Arcustangens $\arctan'(u) = \frac{1}{1+u^2}$ erfüllt. Wegen der für $x_1 \neq 0$ gültigen Gleichung $(\partial_2 V)(x_1, x_2) = \frac{1}{x_1} (1 + (x_2/x_1)^2)^{-1}$ folgt dann $V(x_1, x_2) = \arctan \frac{x_2}{x_1} + c(x_1)$ für $x_1 \neq 0$ mit einer beliebigen differenzierbaren Funktion c , die nur von x_1 abhängt.

Bilden wir nun die partielle Ableitung nach x_1 , so erhalten wir

$$(\partial_1 V)(x_1, x_2) = -\frac{x_2}{x_1^2} \cdot (1 + (x_2/x_1)^2)^{-1} + c'(x_1) = -\frac{x_2}{x_1^2 + x_2^2} + c'(x_1),$$

d.h. wir müssen c konstant wählen (dann $c'(x_1) = 0$). Das sieht erstmal gut aus, aber das Problem ist, dass $V(x_1, x_2)$ für $x_1 = 0$ nicht definiert ist. Das lässt sich auch nicht "reparieren", d.h. V kann nicht stetig auf $\mathbb{R}^2 \setminus \{0\}$ fortgesetzt werden, wie man sich leicht anhand der Grenzwerte von $V(x_1, x_2)$ für $x_1 \rightarrow 0$ überlegt.

Also hat f kein Potential. Allerdings hat f ein Potential, wenn wir f auf den kleineren Definitionsbereich $\{(x_1, x_2) : x_1 \neq 0\}$ einschränken.

Sei $f = \text{grad}V : U \subset \mathbb{R}^k \rightarrow \mathbb{R}^k$ ein Gradientenfeld und $x, y \in U$ so, dass die gerade Verbindungsstrecke von x nach y ganz in U liegt. Dann wissen wir aus Satz 12.20 a)

$$\begin{aligned} V(x) - V(y) &= \int_0^1 \langle \text{grad}V(y + t(x - y)), x - y \rangle dt \\ &= \int_0^1 \langle f(y + t(x - y)), x - y \rangle dt. \end{aligned}$$

Gegeben ein Vektorfeld f ist deshalb $V(x) := \int_0^1 \langle f(y + t(x - y)), x - y \rangle dt$ (für festes $y \in U$) ein Kandidat für ein Potential von f . Um dies genauer zu untersuchen, müssen wir die partiellen Ableitungen dieser Funktion betrachten. Wir stehen also vor der Aufgabe, eine Funktion der Art $x \mapsto \int_a^b g(t, x) dt$ zu differenzieren.

Da dies eine oft auftretende Frage ist, beweisen wir einen Satz zu diesem Sachverhalt.

Satz 12.28 (Differentiation unter dem Integral). Sei $U \subset \mathbb{R}^k$ offen, $[a, b] \subset \mathbb{R}$ ein kompaktes Intervall, und $g : [a, b] \times U \rightarrow \mathbb{R}^d$ stetig.

a) Die Funktion

$$G : U \rightarrow \mathbb{R}^d, \quad G(x) := \int_a^b g(t, x) dt$$

ist stetig. Hierbei ist das Integral wieder komponentenweise zu verstehen, d.h. wir haben $G(x)_j = \int_a^b g_j(t, x) dt$.

b) Wir nehmen nun zusätzlich an, dass die partielle Ableitung $\frac{\partial g}{\partial x_i} : [a, b] \times U \rightarrow \mathbb{R}^d$ für ein $i \in \{1, \dots, k\}$ existiert und stetig ist. Dann ist auch G stetig partiell nach x_i differenzierbar, und die Ableitung kann unter das Integral gezogen werden, d.h.

$$\frac{\partial}{\partial x_i} \int_a^b g(t, x) dt = \int_a^b \frac{\partial g}{\partial x_i}(t, x) dt.$$

Beweis. a) Um zu zeigen, dass G in einem (beliebigen) Punkt $x \in U$ stetig ist, wählen wir zunächst $\delta > 0$ so, dass $\overline{U_\delta(x)} \subset U$. Eingeschränkt auf die abgeschlossene und beschränkte (also kompakte) Menge $[a, b] \times \overline{U_\delta(x)}$ ist g gleichmäßig stetig (Satz 9.19). Gegeben $\varepsilon > 0$ finden wir deshalb $\delta' > 0$ so, dass $\|h\| < \delta' \Rightarrow \|g(t, x+h) - g(t, x)\| < \varepsilon$ für alle $t \in [a, b]$ gilt.

Ist also $\|h\| < \delta'$, so erhalten wir mit der Integralabschätzung aus Korollar 12.21

$$\begin{aligned} \|G(x+h) - G(x)\|_2 &= \left\| \int_a^b (g(t, x+h) - g(t, x)) dt \right\|_2 \\ &\leq \int_a^b \|g(t, x+h) - g(t, x)\|_2 dt < \int_a^b \varepsilon dt = (b-a)\varepsilon. \end{aligned}$$

Da ε beliebig war, folgern wir $G(x+h) \rightarrow G(x)$ für $h \rightarrow 0$, dh G ist stetig in x .

b) Sei $x \in U$. Für die partielle Differenzierbarkeit von G nach x_i und die behauptete Formel für die partielle Ableitung müssen wir zeigen, dass

$$\frac{1}{s} (G(x + se_i) - G(x)) - \int_a^b \frac{\partial g}{\partial x_i}(t, x) dt \quad (\star)$$

für $s \rightarrow 0$ gegen 0 konvergiert. Da (partielle) Differenzierbarkeit komponentenweise geprüft werden kann, dürfen wir annehmen, dass g skalar sind, d.h. nach \mathbb{R} abbildet.

Nach dem Mittelwertsatz der Analysis 1 gibt es eine Zahl σ zwischen 0 und s (abhängig von x, t, i), so dass (\star) identisch zu

$$\int_a^b \left(\frac{1}{s} (g(t, x + se_i) - g(t, x)) - \frac{\partial g}{\partial x_i}(t, x) \right) dt = \int_a^b \left(\frac{\partial g}{\partial x_i}(t, x + \sigma e_i) - \frac{\partial g}{\partial x_i}(t, x) \right) dt$$

ist. Wir können nun wie in Teil a) mit der Integralabschätzung aus Korollar 12.21 abschätzen (hier wird verwendet, dass die partielle Ableitung $\partial_{x_i} g$ stetig, also gleichmäßig stetig auf Kompakta, ist) und folgern, dass (\star) für $s \rightarrow 0$ gegen 0 geht.

Exakt wie in Teil a) folgt dann auch die Stetigkeit der partiellen Ableitung von G . \square

Ausgerüstet mit diesem sehr oft hilfreichen Satz können wir nun die Existenz eines Potentials auf geeigneten Definitionsbereichen beweisen.

Definition 12.29. Eine Menge $U \subset \mathbb{R}^k$ heißt *sternförmig* bzgl eines Punktes $x^* \in U$, falls für jedes $x \in U$ die gerade Verbindungsstrecke von x^* nach x ganz in U liegt.

Beispiel 12.30.

- a) Konvexe Mengen $U \subset \mathbb{R}^k$ sind sternförmig bzgl jedes Punktes $x^* \in U$ nach der Definition von Konvexität.
- b) Sei $p \in [1, \infty]$, $x^* \in \mathbb{R}^k$, und $r > 0$. Dann ist $U_r^{\|\cdot\|_p}(x^*)$ konvex, also sternförmig.
- c) Eine Teilmenge $U \subset \mathbb{R}^2$ der Form \star ist sternförmig bzgl ihres Mittelpunktes, aber nicht konvex.
- d) $\mathbb{R}^k \setminus \{0\}$ ist nicht sternförmig, da für jedes $x^* \in \mathbb{R}^k \setminus \{0\}$ die gerade Verbindungsstrecke zu $-x^*$ den Ursprung 0 enthält, also nicht ganz in $\mathbb{R}^k \setminus \{0\}$ enthalten ist.

Satz 12.31. Sei $U \subset \mathbb{R}^k$ sternförmig bzgl $x^* \in U$ und offen. Ein stetig differenzierbares Vektorfeld $f : U \rightarrow \mathbb{R}^k$ ist ein Gradientenfeld genau dann, wenn $\partial_i f_j = \partial_j f_i$ für alle $i, j \in \{1, \dots, k\}$ gilt.

Beweis. Wir haben bereits gesehen, dass die Symmetriebedingung $\partial_i f_j = \partial_j f_i$ notwendig für die Existenz eines Potentials ist. Um zu zeigen, dass sie in der hier beschriebenen Situation auch hinreichend ist, betrachten wir unseren Kandidaten für ein Potential von f , nämlich

$$V : U \rightarrow \mathbb{R}, \quad V(x) := \int_0^1 \langle f(x^* + t(x - x^*)), x - x^* \rangle dt.$$

Aufgrund der Sternförmigkeit von U bzgl. x^* liegen die Punkte $x^* + t(x - x^*)$ für alle $t \in [0, 1]$ im Definitionsbereich U von f . Aufgrund der Stetigkeit von f und des Skalarproduktes ist der Integrand stetig, das Integral existiert also.

Wir zeigen nun, dass V partiell differenzierbar mit $\text{grad}V = f$ ist. Dazu bemerken wir zunächst, dass

$$g : [0, 1] \times U \rightarrow \mathbb{R}, \quad g(t, x) := \langle f(x^* + t(x - x^*)), x - x^* \rangle$$

nach allen x_i stetig partiell differenzierbar ist, mit partiellen Ableitungen

$$\begin{aligned} \frac{\partial g}{\partial x_i}(t, x) &= \frac{\partial}{\partial x_i} \sum_{j=1}^k f_j(x^* + t(x - x^*))(x_j - x_j^*) \\ &= \sum_{j=1}^k t(\partial_i f_j)(x^* + t(x - x^*))(x_j - x_j^*) + f_i(x^* + t(x - x^*)). \end{aligned}$$

Der wesentliche Schritt des Beweises ist nun, diesen Ausdruck mit Hilfe der Symmetriebe-

dingung $\partial_i f_j = \partial_j f_i$ als eine Ableitung nach t umzuschreiben. Es gilt

$$\begin{aligned} \frac{d}{dt}(t f_i(x^* + t(x - x^*))) &= f_i(x^* + t(x - x^*)) + t(Df_i)(x^* + t(x - x^*))(x - x^*) \\ &= f_i(x^* + t(x - x^*)) + t \sum_{j=1}^k (\partial_j f_i)(x^* + t(x - x^*))(x_j - x_j^*) \\ &= f_i(x^* + t(x - x^*)) + t \sum_{j=1}^k (\partial_i f_j)(x^* + t(x - x^*))(x_j - x_j^*) \\ &= \frac{\partial g}{\partial x_i}(t, x). \end{aligned}$$

Nach dieser Einsicht lässt sich der Beweis ohne größere Schwierigkeiten führen. Nach Satz 12.28 ist V partiell differenzierbar, mit partiellen Ableitungen

$$\begin{aligned} \frac{\partial V}{\partial x_i}(x) &= \int_0^1 \frac{\partial g}{\partial x_i}(t, x) dt = \int_0^1 \frac{d}{dt}(t f_i(x^* + t(x - x^*))) dt \\ &= 1 \cdot f_i(x^* + 1(x - x^*)) + 0 \cdot f_i(x^* + 0(x - x^*)) = f_i(x). \end{aligned}$$

Wir haben also $\text{grad}V = f$, d.h. V ist ein Potential für f und damit ist f ein Gradientenfeld. \square

Beispiel 12.32. Wir betrachten noch einmal Beispiel 12.27 b). Das dort betrachtete Vektorfeld f war auf dem nicht sternförmigen Gebiet $\mathbb{R}^2 \setminus \{0\}$ definiert. Die Voraussetzungen von Satz 12.31 sind also nicht erfüllt und die Tatsache, dass f kein Potential hat, steht deshalb nicht im Widerspruch zu Satz 12.31.

Schränken wir f aber auf ein sternförmiges Gebiet ein, z.B. auf die Halbebene $U = \{(x_1, x_2) : x_1 > 0\}$, so sind die Voraussetzungen von Satz 12.31 erfüllt, und $f|_U$ hat ein Potential – wie in Beispiel 12.27 b) berechnet, ist $V(x_1, x_2) = \arctan \frac{x_2}{x_1}$ ein auf U wohldefiniertes Potential von f .

Ist $f = \text{grad}V : U \rightarrow \mathbb{R}^k$ ein Gradientenvektorfeld mit Potential V , so erinnert die Formel

$$\int_0^1 \langle f(y + t(x - y)), x - y \rangle dt = V(x) - V(y)$$

an den Hauptsatz der Differential- und Integralrechnung. Einerseits können wir das Integral als Auswertung des Potentials am Endpunkt des Weges (x) minus das Potential am Anfangspunkt des Weges (y) schreiben. Andererseits können wir $V(x)$ durch seine “Ableitung” (Gradienten $\text{grad}V = f$) und den Wert $V(y)$ an einem festen Punkt y ausdrücken, also bis auf Konstante aus seiner Änderungsrate rekonstruieren.

Aus diesem Grund werden Potentiale manchmal auch “Stammfunktionen” genannt. Allerdings ist es wichtig, zu betonen, dass der Hauptsatz der Differentialrechnung nicht in äquivalenter Form in mehreren Dimensionen gilt: Wir haben bereits gesehen, dass die Existenz eines Potentials nur unter bestimmten Bedingungen (Symmetriebedingung $\partial_i f_j = \partial_j f_i$ und zB Sternförmigkeit des Definitionsbereiches) gilt.

- a) An dieser Stelle könnte man noch sehr viel mehr über Potentiale und die damit eng verbundenen Wegintegrale sagen, was wir hier aber aus Zeitgründen nicht tun können. Nur soviel: Es gibt einen Begriff eines Wegintegrals $\int_{\gamma} f$, in dem ein Vektorfeld entlang eines Weges integriert wird. Ist f ein Gradientenfeld, so stellt sich heraus, dass dieses Integral nur von dem Anfangs- und Endpunkt des Weges γ , nicht aber von dem genauen Verlauf von γ abhängt. Dieses Thema werden Sie in Vorlesungen über Differentialgleichungen und Funktionentheorie gründlich kennenlernen.
- b) Die Voraussetzung in Satz 12.31, dass der Definitionsbereich sternförmig sein soll, kann abgeschwächt werden. Es gibt den Begriff eines *einfach zusammenhängenden* Gebietes U , der schwächer als sternförmig ist, aber immer noch für den Beweis von Satz 12.31 ausreicht. Anschaulich gesprochen ist ein einfach zusammenhängendes Gebiet ein Gebiet "ohne Löcher". Sie werden diesen Begriff in der Topologie kennenlernen.

12.5 Der Satz von Taylor für mehrere Variablen

Ziel dieses Abschnitts ist es, eine mehrdimensionale Version der Taylorentwicklung anzugeben.

Sie erinnern sich sicher gerne an Taylorentwicklungen in einer Variablen: Ist $f : I \subset \mathbb{R} \rightarrow \mathbb{R}$ eine auf einem Intervall I definierte beliebig oft differenzierbare Funktion, so ist $T_{f,x_0,n}(x) = \sum_{l=0}^n \frac{1}{l!} f^{(l)}(x_0) (x - x_0)^l$ das Taylorpolynom n -ten Grades von f mit Entwicklungspunkt $x_0 \in I$, das unter geeigneten Bedingungen eine gute Näherung von $f(x)$ darstellt. Nach dem Satz 5.24 von Taylor gilt

$$f(x) = T_{f,x_0,n}(x) + \frac{f^{(n+1)}(\xi)}{(n+1)!} (x - x_0)^{n+1}$$

mit einem (von x, x_0 abhängigen) ξ zwischen x und x_0 . Der "Fehlerterm" $\frac{f^{(n+1)}(\xi)}{(n+1)!} (x - x_0)^{n+1}$ der Entwicklung wird das Lagrangesche Restglied genannt.

Wenn wir an eine analoge Taylorentwicklung in mehreren Variablen denken, werden wir es sicher mit höheren partiellen Ableitungen zu tun haben, von denen es jede Menge gibt. Für eine glatte Funktion $f \in C^\infty(\mathbb{R}^k, \mathbb{R})$ zum Beispiel $\partial_1 \partial_3^2 \partial_2^4 \partial_1^2 \partial_3 f$, was, nach dem Satz von Schwarz, mit $\partial_1^3 \partial_2^4 \partial_3^3 f$ übereinstimmt. Es lohnt sich, zuerst eine kompakte Notation einzuführen, die das Zählen der Ordnungen der partiellen Ableitungen erleichtert.

Definition 12.33 (Multiindexnotation). Sei $k \in \mathbb{N}$.

- a) Ein *Multiindex* ist ein k -Tupel $\mu = (\mu_1, \dots, \mu_k) \in \mathbb{N}_0^k$.
- b) $|\mu| := \mu_1 + \dots + \mu_k \in \mathbb{N}_0$ heißt *Ordnung* von μ .
- c) Die Fakultät eines Multiindex ist $\mu! := \mu_1! \cdot \mu_2! \cdot \dots \cdot \mu_k!$.
- d) Für $\mu \in \mathbb{N}_0^k$ ist *das Monom mit Exponent μ* die Funktion

$$\mathbb{R}^k \ni x \mapsto x^\mu := x_1^{\mu_1} \cdot \dots \cdot x_k^{\mu_k} \in \mathbb{R}.$$

e) Sei $U \subset \mathbb{R}^k$ offen und $f \in C^n(U, \mathbb{R}^d)$. Dann setzen wir für $\mu \in \mathbb{N}_0^k$ mit $|\mu| \leq n$

$$\partial^\mu f := \partial_1^{\mu_1} \dots \partial_k^{\mu_k} f, \quad \partial^{(0, \dots, 0)} f := f.$$

Alternative ebenfalls verbreitete Notation für diese höheren partiellen Ableitungen sind $\frac{\partial^{|\mu|} f}{\partial x_1^{\mu_1} \dots \partial x_k^{\mu_k}}$ oder $D^\mu f$.

Beispiel 12.34. Sei $\mu = (0, 3, 1, 2)$. Dann ist $|\mu| = 6$ und $\mu! = 0!3!1!2! = 12$. Für $x \in \mathbb{R}^4$ gilt $x^\mu = x_2^3 x_3 x_4^2$, und für $f : \mathbb{R}^4 \rightarrow \mathbb{R}$ können wir die Ableitung $\partial^\mu f = \partial_2^3 \partial_3 \partial_4^2 f$ betrachten.

Für n -mal stetig partiell differenzierbares $f : U \rightarrow \mathbb{R}^d$ kann jede gemischte partielle Ableitung der Ordnung n in eine partielle Ableitung der Form $\partial^\mu f$ mit $|\mu| = n$ umgeschrieben werden, da die Reihenfolge der partiellen Ableitungen nach dem Satz von Schwarz keine Rolle spielt.

Um uns etwas an den Umgang mit Multiindices zu gewöhnen, besprechen wir eine Multinomialformel, d.h. eine Binomialformel für mehrere Variablen.

Lemma 12.35. Sei $x \in \mathbb{R}^k$ und $n \in \mathbb{N}$. Dann gilt

$$(x_1 + \dots + x_k)^n = n! \sum_{|\mu|=n} \frac{x^\mu}{\mu!}.$$

Die Summe $\sum_{|\mu|=n}$ bedeutet, dass über alle Multiindices $\mu \in \mathbb{N}_0^k$ mit $|\mu| = n$ summiert wird.

Der Beweis dieses Lemmas wird in den Übungen geführt. Machen Sie sich klar, dass die Aussage für $k = 1$ trivial ist, und für $k = 2$ die aus Analysis 1 bekannte binomische Formel (Satz 2.9) ergibt. Für allgemeines k bietet sich ein Beweis per Induktion an.

Beispiel 12.36. Dieses Beispiel sollte Ihnen vor Augen führen, wie übersichtlich und schön Multiindexnotation im Vergleich zu ausgeschriebenen Summen ist.

$$\begin{aligned} (x_1 + x_2 + x_3)^3 &= x_1^3 + 3x_1^2 x_2 + 3x_1^2 x_3 + 3x_1 x_2^2 + 3x_1 x_3^2 + 6x_1 x_2 x_3 \\ &\quad + 3x_2^2 x_3 + 3x_2 x_3^2 + x_2^3 + x_3^3. \end{aligned}$$

Als weitere Übung mit Multiindices betrachten wir Ableitungen von Monomen. Dazu verabreden wir noch zwei weitere Notationen für Multiindices $\mu, \nu \in \mathbb{N}_0^k$.

- Wir schreiben $\mu \leq \nu$ falls $\mu_i \leq \nu_i$ für alle $i \in \{1, \dots, k\}$.
- $\mu - \nu := (\mu_1 - \nu_1, \dots, \mu_k - \nu_k)$. Dies ist ein Multiindex, falls $\mu \geq \nu$.

Lemma 12.37. Sei $\mu \in \mathbb{N}_0^k$ ein Multiindex und $p_\nu : \mathbb{R}^k \rightarrow \mathbb{R}$, $p_\nu(x) := x^\nu$ das Monom

mit Exponent $\nu \in \mathbb{N}_0^k$. Dann gilt

$$\partial^\mu p_\nu(x) = \begin{cases} \frac{\nu!}{(\nu-\mu)!} x^{\nu-\mu} & \mu \leq \nu \\ 0 & \text{sonst} \end{cases}.$$

Beweis. In einer Variablen wissen wir $\frac{d^n}{dx^n} x^m = m(m-1)\cdots(m-(n-1))x^{m-n} = \frac{m!}{(m-n)!} x^{m-n}$ für $m \geq 0$ (und $= 0$ für $m < n$). Die behauptete Version mit mehreren Variablen ergibt sich völlig analog: Für $\mu \leq \nu$ haben wir

$$\begin{aligned} \partial^\mu p_\nu(x) &= \partial_{x_1}^{\mu_1} x_1^{\nu_1} \cdots \partial_{x_k}^{\mu_k} x_k^{\nu_k} = \frac{\nu_1! \cdots \nu_k!}{(\nu_1 - \mu_1)! \cdots (\nu_k - \mu_k)!} x_1^{\nu_1 - \mu_1} \cdots x_k^{\nu_k - \mu_k} \\ &= \frac{\nu!}{(\nu - \mu)!} x^{\nu - \mu}, \end{aligned}$$

und $\partial^\mu p_\nu(x) = 0$, falls $\mu_i > \nu_i$ für ein i , d.h. falls $\mu \not\leq \nu$. □

Ausgerüstet mit Multiindices können wir nun Taylorpolynome in mehreren Variablen definieren.

Definition 12.38. Sei $U \subset \mathbb{R}^k$ offen, $f \in C^k(U, \mathbb{R})$, und $x_0 \in U$. Das *Taylorpolynom der Ordnung n von f mit Entwicklungspunkt x_0* ist das Polynom

$$T_{f,x_0,n}(x) := \sum_{|\mu| \leq n} \frac{1}{\mu!} (\partial^\mu f)(x_0) \cdot (x - x_0)^\mu.$$

Die Summe erstreckt sich hier über alle Multiindices μ der Ordnungen $|\mu| = 0, 1, \dots, n$.

Wir haben uns hier auf skalare Funktionen f beschränkt. Ist man an vektorwertigen Funktionen $f : U \rightarrow \mathbb{R}^d$ interessiert, so kann man Taylorentwicklungen für jede Komponente von f betrachten.

Es lohnt sich, Taylorpolynome niedriger Ordnung genauer anzusehen. Für $n = 0$ haben wir natürlich das konstante Polynom $T_{f,x_0,0}(x) = f(x_0)$, und für $n = 1$ ergibt sich

$$\begin{aligned} T_{f,x_0,1}(x) &= f(x_0) + (\partial_1 f)(x_0) \cdot (x_1 - (x_0)_1) + \cdots + (\partial_k f)(x_0) \cdot (x_k - (x_0)_k) \\ &= f(x_0) + \langle \text{grad} f(x_0), x - x_0 \rangle. \end{aligned}$$

Bemerkung zur Notation: $x_0 \in \mathbb{R}^k$ ist ein Vektor (der Entwicklungspunkt), $(x_0)_j \in \mathbb{R}$ seine j -te Komponente.

Für $n = 2$ haben wir zusätzlich Multiindices zweiter Ordnung, also von der Form $(1, 0, \dots, 1, 0, \dots, 0)$ und $(0, \dots, 2, \dots, 0)$. Das ergibt

$$\begin{aligned} T_{f,x_0,2}(x) &= T_{f,x_0,1}(x) + \frac{1}{2} \sum_{\substack{i,j=1 \\ i \neq j}}^k (\partial_i \partial_j f)(x_0) (x_i - (x_0)_i) (x_j - (x_0)_j) + \frac{1}{2} \sum_{j=1}^k (\partial_j^2 f)(x_0) (x_j - (x_0)_j)^2 \\ &= T_{f,x_0,1}(x) + \frac{1}{2} \sum_{i,j=1}^k (\partial_i \partial_j f)(x_0) (x_i - (x_0)_i) (x_j - (x_0)_j). \end{aligned}$$

Der erste Faktor $\frac{1}{2}$ in der ersten Zeile rührt daher, dass wir zB den Multiindex $(1, 1, 0, \dots, 0)$ nicht doppelt ($i = 1, j = 2$ und $i = 2, j = 1$) zählen dürfen.

Um die Struktur des Taylorpolynoms zweiter Ordnung zu verstehen, machen wir einen Ausflug in die Lineare Algebra und betrachten eine $(k \times k)$ -Matrix M und einen Vektor $v \in \mathbb{R}^k$. Dann gilt mit dem Euklidischen Skalarprodukt $\langle v, Mv \rangle = \sum_{i=1}^k v_i(Mv)_i = \sum_{i,j=1}^k v_i M_{ij} v_j$. Definieren wir also für eine zweimal partiell differenzierbare Funktion $f : U \subset \mathbb{R}^2$ mit Entwicklungspunkt $x_0 \in U$ die *Hesse-Matrix* als

$$H_f(x_0) := \begin{pmatrix} (\partial_1^2 f)(x_0) & (\partial_1 \partial_2 f)(x_0) & \dots & (\partial_1 \partial_k f)(x_0) \\ (\partial_2 \partial_1 f)(x_0) & (\partial_2^2 f)(x_0) & \dots & (\partial_2 \partial_k f)(x_0) \\ \vdots & \vdots & \ddots & \vdots \\ (\partial_k \partial_1 f)(x_0) & (\partial_k \partial_2 f)(x_0) & \dots & (\partial_k^2 f)(x_0) \end{pmatrix},$$

so nimmt das Taylorpolynom zweiter Ordnung die übersichtliche Form

$$T_{f,x_0,2}(x) = f(x_0) + \langle \nabla f(x_0), x - x_0 \rangle + \frac{1}{2} \langle x - x_0, H_f(x_0)(x - x_0) \rangle$$

an. Noch kürzer schreibt es sich, wenn wir den Auswertungspunkt x in der Form $x = x_0 + h$ ausdrücken:

$$T_{f,x_0,2}(x_0 + h) = f(x_0) + \langle \nabla f(x_0), h \rangle + \frac{1}{2} \langle h, H_f(x_0)h \rangle.$$

Dies ist die wichtigste Form der Taylorentwicklung in mehreren Variablen.

Die Hessematrix versammelt alle partiellen Ableitungen zweiter Ordnung. Wenn f sogar C^2 , also zweimal stetig partiell differenzierbar ist, so ist $H_f(x_0)$ nach dem Satz von Schwarz eine symmetrische Matrix,

$$H_f(x_0)_{ij} = (\partial_i \partial_j f)(x_0) = (\partial_j \partial_i f)(x_0) = H_f(x_0)_{ji}.$$

Beispiel 12.39. Sei $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, $f(x) = x_1 \sin(x_2)$. Dies ist eine glatte skalare Funktion mit Gradient und Hessematrix

$$\nabla f(x)^T = \begin{pmatrix} \sin(x_2) \\ x_1 \cos(x_2) \end{pmatrix}, \quad H_f(x) = \begin{pmatrix} 0 & \cos(x_2) \\ \cos(x_2) & -x_1 \sin(x_2) \end{pmatrix}.$$

Das Taylorpolynom zweiter Ordnung von f mit Entwicklungspunkt $x_0 = 0$ ist also (die Terme nullter und erster Ordnung verschwinden hier, $f(0) = 0$, $\nabla f(0) = 0$)

$$T_{f,0,2}(x) = \frac{1}{2} \left\langle x, \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} x \right\rangle = x_1 x_2.$$

Wie im Eindimensionalen ist das Taylorpolynom $T_{f,x_0,n}$ eindeutig dadurch bestimmt, dass es von Ordnung höchstens n ist und alle partiellen Ableitungen bis Ordnung n an der Stelle x_0 mit den partiellen Ableitungen von f übereinstimmen.

Lemma 12.40. Sei $f \in C^n(U, \mathbb{R})$ und $x_0 \in U$. Dann ist $T_{f,x_0,n}$ das eindeutige Polynom $\mathbb{R}^k \rightarrow \mathbb{R}$ der Ordnung höchstens n mit der Eigenschaft

$$(\partial^\mu T_{f,x_0,n})(x_0) = (\partial^\mu f)(x_0), \quad |\mu| \leq n.$$

Beweis. Sei $\mu \in \mathbb{N}_0^k$ ein Multiindex der Ordnung $|\mu| \leq n$. Wir bestimmen die partielle Ableitung ∂^μ des Taylorpolynoms mittels Lemma 12.37. Da $\partial^\mu(x - x_0)^\nu$ ausgewertet bei $x = x_0$ mit $\delta_{\mu,\nu} \nu!$ übereinstimmt, erhalten wir

$$(\partial^\mu T_{f,x_0,n})(x_0) = \sum_{|\nu| \leq n} \frac{1}{\nu!} (\partial^\nu f)(x_0) \cdot \partial^\mu (x - x_0)^\nu \Big|_{x=x_0} = (\partial^\mu f)(x_0).$$

Ist andererseits $p(x) = \sum_{|\nu| \leq n} c_\nu (x - x_0)^\nu$ ein beliebiges Polynom vom Grad höchstens n mit der Eigenschaft $(\partial^\mu p)(x_0) = (\partial^\mu f)(x_0)$ für alle μ mit $|\mu| \leq n$, so erhalten wir durch Ableiten und Auswerten bei $x = x_0$ die Identität $\nu! c_\nu = (\partial^\nu p)(x_0) = (\partial^\nu f)(x_0)$. Also gilt in diesem Fall $p = T_{f,x_0,n}$. \square

Wir kommen nun zum Satz von Taylor. Er sieht hier fast genau wie in einer Dimension aus.

Satz 12.41 (Satz von Taylor in mehreren Variablen). Sei $U \subset \mathbb{R}^k$ offen, $f \in C^{n+1}(U, \mathbb{R})$, und sei $h \in \mathbb{R}^k$ so, dass die gerade Verbindungsstrecke zwischen $x_0 \in U$ und $x = x_0 + h$ ganz in U enthalten ist. Dann existiert $\theta \in (0, 1)$, so dass

$$f(x) = T_{f,x_0,n}(x) + \sum_{|\mu|=n+1} \frac{1}{\mu!} (\partial^\mu f)(x_0 + \theta(x - x_0)) \cdot (x - x_0)^\mu.$$

Beweis. Die Idee des Beweises ist es, den Satz von Taylor für eine Variable zu verwenden (Satz 5.24). Wir betrachten deshalb als Hilfsfunktion $g : [0, 1] \rightarrow \mathbb{R}$, $g(t) := f(x_0 + t \cdot h)$, wobei wir $h := x - x_0$ als Abkürzung schreiben. Diese Funktion liegt in $C^{n+1}([0, 1], \mathbb{R})$. Wir haben $g(0) = f(x_0)$, und die Ableitungen von g sind durch Richtungsableitungen von f in Richtung h gegeben:

$$g^{(l)}(t) = \frac{d^l}{dt^l} f(x_0 + th) = (\partial_h^l f)(x_0 + th), \quad l = 1, \dots, n + 1.$$

Nach dem Satz von Taylor für eine Variable gibt es $\theta \in (0, 1)$, so dass

$$f(x) = g(1) = \sum_{l=0}^n \frac{g^{(l)}(0)}{l!} 1^l + \frac{g^{(n+1)}(\theta)}{(n+1)!}.$$

Wir wollen nun einsehen, wie die mehrfachen Richtungsableitungen durch Taylorpolynome ausgedrückt werden können. Dazu bemerken wir, dass die Richtungsableitung in Richtung h durch $\partial_h f = \sum_{j=1}^k h_j \partial_j f$ gegeben ist. Die zweite Ableitung in Richtung h ist also $\partial_h^2 f = (\sum_{j=1}^k h_j \partial_j)^2 f$, und allgemein $\partial_h^l f = (\sum_j h_j \partial_j)^l f$.

Unter Verwendung von Lemma 12.35 ergibt sich deshalb

$$g^{(l)}(t) = \left(\sum_{j=1}^k h_j \partial_j \right)^l f(x_0 + th) = l! \sum_{|\mu|=l} \frac{1}{\mu!} (\partial^\mu f)(x_0 + th) h^\mu.$$

Eingesetzt in die obige Taylorentwicklung für $g(1)$ haben wir also

$$\begin{aligned} f(x) &= \sum_{l=0}^n \sum_{|\mu|=l} \frac{1}{\mu!} (\partial^\mu f)(x_0) h^\mu + \sum_{|\mu|=n+1} \frac{1}{\mu!} (\partial^\mu f)(x_0 + \theta h) h^\mu \\ &= T_{f,x_0,n}(x) + \sum_{|\mu|=n+1} \frac{1}{\mu!} (\partial^\mu f)(x_0 + \theta(x - x_0)) \cdot (x - x_0)^\mu, \end{aligned}$$

wie behauptet. □

Wie in einer Dimension nennen wir

$$R_{f,x_0,n}(x) := \sum_{|\mu|=n+1} \frac{1}{\mu!} (\partial^\mu f)(x_0 + \theta(x - x_0)) \cdot (x - x_0)^\mu$$

das Lagrangesche Restglied der Taylorentwicklung. Dieses Restglied gibt die Güte der Näherung von $f(x)$ durch $T_{f,x_0,n}(x)$ an. Man ist deshalb interessiert, es abzuschätzen. In dieser Richtung haben wir das folgende Resultat.

Satz 12.42 (Restgliedabschätzung der Taylorentwicklung). Sei $U \subset \mathbb{R}^k$ offen, $f \in C^{n+1}(U, \mathbb{R})$ und $x_0 \in U$. Dann gilt

$$\lim_{x \rightarrow x_0} \frac{|f(x) - T_{f,x_0,n+1}(x)|}{\|x - x_0\|^{n+1}} = 0.$$

Beweis. Nach dem Satz von Taylor gibt es $\theta \in (0, 1)$, so dass

$$\frac{|f(x) - T_{f,x_0,n+1}(x)|}{\|x - x_0\|^{n+1}} = \sum_{|\mu|=n+1} \frac{1}{\mu!} |(\partial^\mu f)(x_0 + \theta(x - x_0)) - (\partial^\mu f)(x_0)| \frac{|(x - x_0)^\mu|}{\|x - x_0\|^{n+1}}.$$

Wir verwenden hier die Euklidische Norm. Dann gilt

$$|(x - x_0)^\mu| = |(x - x_0)_1^{\mu_1}| \cdots |(x - x_0)_k^{\mu_k}| \leq \|x - x_0\|^{\mu_1 + \dots + \mu_k} = \|x - x_0\|^{n+1},$$

d.h. wir können $\frac{|(x-x_0)^\mu|}{\|x-x_0\|^{n+1}} \leq 1$ abschätzen.

Aufgrund der Stetigkeit der partiellen Ableitungen $\partial^\mu f$ geht $|(\partial^\mu f)(x_0 + \theta(x - x_0)) - (\partial^\mu f)(x_0)|$ für $x \rightarrow x_0$ gegen Null. □

Beispiel 12.43. Für die glatte Funktion $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, $f(x_1, x_2) := \sin(x_1 x_2)$, ist das Taylorpolynom dritter Ordnung um $x_0 = 0$ (Übung)

$$T_{f,0,3}(x) = x_1 x_2 - \frac{1}{6} x_1^3 x_2^3.$$

Wir haben also den Fehler

$$\frac{|\sin(x_1 x_2) - x_1 x_2 - \frac{1}{6} x_1^3 x_2^3|}{(x_1^2 + x_2^2)^{3/2}} \rightarrow 0 \text{ für } x_1, x_2 \rightarrow 0.$$



Analog zu unserem Vorgehen in Analysis 1 könnten wir nun hinreichende Bedingungen an f identifizieren, die uns erlauben, die Konvergenz der Taylorreihe $\sum_{\mu} \frac{1}{\mu!} (\partial^{\mu} f)(x_0) (x - x_0)^{\mu}$ (Summation über Multiindices *beliebiger* Ordnung) zu folgern. Dies wird im Folgenden allerdings nicht Schwerpunkt unserer Untersuchungen sein, so dass wir von einer Ausführung absehen.

12.6 Lokale Extrema

In der Analysis 1 haben wir lokale Extrema diskutiert und hinreichende/notwendige Bedingungen für lokale Extrema durch Ableitungen der betrachteten Funktion hergeleitet. Wir verallgemeinern diese Analyse jetzt auf Funktionen mehrerer Variable, wobei sich Taylorentwicklungen als sehr hilfreich herausstellen werden. Die ersten Schritte sind sehr ähnlich zu unserem Vorgehen im letzten Semester:

Definition 12.44. Sei $U \subset \mathbb{R}^k$ und $f : U \rightarrow \mathbb{R}$ eine Funktion.

- Ein Punkt $x \in U$ heißt *lokales Maximum (Minimum)* von f , falls es eine Umgebung $V \subset U$ von x gibt, so dass $f(x) \geq f(y)$ ($f(x) \leq f(y)$) für alle $y \in V$ gilt. Ein *lokales Extremum* ist ein lokales Maximum oder Minimum. Ein lokales Extremum x heißt *isoliert*, wenn sogar $f(x) > f(y)$ bzw. $f(x) < f(y)$ für alle $y \in V$ gilt.
- Ein Punkt $x \in U$ heißt *globales Maximum (Minimum)*, wenn $f(x) \geq f(y)$ ($f(x) \leq f(y)$) für alle $y \in U$ gilt.

Wir haben uns hier auf skalare Funktionen beschränkt, da die Ungleichungen $f(x) \leq f(y)$ für vektorwertige Funktionen nicht definiert sind.

Definition 12.45. Sei $U \subset \mathbb{R}^k$ offen und $f \in C^1(U, \mathbb{R})$. Ein Punkt $x \in U$ heißt *kritischer Punkt* von f , falls $(Df)(x) = 0$.

Ein kritischer Punkt ist also ein Punkt x , an dem der Gradient von f verschwindet, d.h. $(\partial_i f)(x) = 0$ für alle $i \in \{1, \dots, k\}$. Mehr noch, alle Richtungsableitungen $(\partial_v f)(x) = (Df)(x)v = 0$ verschwinden an einem kritischen Punkt. Anschaulich gesprochen sieht der Graph von f bei x also in jede Richtung flach aus (Steigung Null). Das passt gut zu unserer intuitiven Vorstellung von lokalen Extrema (Denken Sie an einen Berggipfel).

Lemma 12.46. *Jedes lokale Extremum einer Funktion $f \in C^1(U, \mathbb{R})$ ist ein kritischer Punkt.*

Beweis. Sei $x \in U$ ein lokales Extremum von f (hier ist U als offen vorausgesetzt). Sei $v \in \mathbb{R}^k$ ein Richtungsvektor, und $g : (-\varepsilon, \varepsilon) \rightarrow \mathbb{R}$, $g : t \mapsto f(x + tv)$ (diese Funktion ist für klein genug gewähltes $\varepsilon > 0$ wohldefiniert, da x ein innerer Punkt ist). Da x ein lokales Extremum von f ist, ist 0 ein lokales Extremum von g . Weiterhin ist g differenzierbar. Aus Analysis 1 wissen wir also $g'(0) = 0$. Aber $0 = g'(0) = (Df)(x)v$; d.h. das Differential $(Df)(x)$ verschwindet auf jedem Vektor $v \in \mathbb{R}^k$ und ist deshalb 0 . Also ist x ein kritischer Punkt von f . \square

Beispiel 12.47.

- a) Sei $f : \mathbb{R}^3 \rightarrow \mathbb{R}$, $f(x_1, x_2, x_3) = x_1^2 + x_2^2 + x_3^2 = \|x\|_2^2$. Dann ist 0 offenbar ein isoliertes lokales (sogar globales) Minimum von f , denn $f(0) = 0$ und $f(x) > 0$ für alle $x \neq 0$. Der Gradient von f ist

$$(\nabla f)(x_1, x_2, x_3) = (2x_1, 2x_2, 2x_3) = 2x,$$

was genau dann verschwindet, wenn $x = 0$ gilt. Also hat f genau einen kritischen Punkt, nämlich 0.

- b) $f : \mathbb{R}^4 \rightarrow \mathbb{R}$, $f(x_1, x_2, x_3, x_4) = x_1^2$ hat an jedem Punkt $(0, x_2, x_3, x_4)$ ein nicht isoliertes lokales Minimum.

- c) Sei $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, $f(x, y) := xy^3 - yx^2$. Dann gilt

$$(\nabla f)(x, y) = (y^3 - 2xy, 3xy^2 - x^2).$$

Was sind die kritischen Punkte von f ? Nullsetzen des Gradienten gibt das Gleichungssystem

$$y(y^2 - 2x) = 0, \quad 3x(y^2 - \frac{1}{3}x) = 0.$$

Die linke Gleichung ist für $y = 0$ erfüllt, und für $y = 0$ ist die rechte Gleichung genau dann erfüllt, wenn $x = 0$. Andererseits ist die rechte Gleichung für $x = 0$ erfüllt, in diesem Fall erzwingt die linke Gleichung auch $y = 0$. Für $x \neq 0, y \neq 0$ erhalten wir $2x = y^2 = \frac{1}{3}x$, ein Widerspruch.

Also hat f genau einen kritischen Punkt, nämlich $(0, 0)$. Der Wert von f an diesem kritischen Punkt ist $f(0, 0) = 0$.

Allerdings ist dieser kritische Punkt kein lokales Extremum: Betrachten wir zB die Einschränkung von f auf die Diagonale $x = y$, so haben wir $f(x, x) = x^3(x-1)$, was das Vorzeichen bei $x = 0$ wechselt. Also nimmt f in jeder ε -Umgebung von $(0, 0)$ sowohl positive als auch negative Funktionswerte an, hat also kein lokales Extremum bei $(0, 0)$.

Das zweite Beispiel führt uns deutlich vor Augen, dass Lemma 12.46 nur eine notwendige, aber keine hinreichende Bedingung für lokale Extrema ist. So etwas kennen wir schon aus Analysis 1, z.B. hat $x \mapsto x^3$ bei $x = 0$ verschwindende Ableitung, aber kein lokales Extremum. Kritische Punkte, die keine lokalen Extrema sind, nennt man auch *Sattelpunkte*.

Wir sind jetzt daran interessiert, auch hinreichende Bedingungen für lokale Extrema zu finden. In der Analysis 1 ist ein solches hinreichendes Kriterium: " x kritischer Punkt und $f''(x) > 0$ ($f''(x) < 0$) $\Rightarrow x$ lokales Minimum (Maximum)."

Auch für mehrere Variablen werden wir ein ähnliches Kriterium haben. Erinnern wir uns an das Taylorpolynom zweiten Grades einer zweimal differenzierbaren Funktion f , nämlich

$$T_{f,x,2}(x+h) = f(x) + \langle \nabla f(x), h \rangle + \frac{1}{2} \langle h, H_f(x)h \rangle,$$

so liegt es nahe, dass dieses Kriterium die Hessematrix von f involvieren wird. An einem kritischen Punkt x gilt $\langle \nabla f(x), h \rangle = 0$, also $T_{f,x,2}(x+h) - f(x) = \frac{1}{2} \langle h, H_f(x)h \rangle$. Da $T_{f,x,2}(x+h)$ im Sinne von Satz 12.41 und Satz 12.42 für kleines $\|h\|$ eine gute Näherung von $f(x+h)$ darstellt, erwarten wir ein lokales Minimum bei x , falls $\langle h, H_f(x)h \rangle > 0$ für alle $h \neq 0$ (denn dann gilt $T_{f,x,2}(x+h) > f(x)$) und ein lokales Maximum, falls $\langle h, H_f(x)h \rangle < 0$ für alle $h \neq 0$.

Dies ist noch kein Beweis, sondern nur eine Motivation der folgenden Untersuchung. Um lokale Extrema genau zu verstehen, erinnern an das Konzept von definiten Matrizen, das Ihnen aus der Linearen Algebra bekannt ist (Kapitel 17 im Skript Lineare Algebra 2):

Definition 12.48. Sei $A = A^T \in M_k(\mathbb{R})$ eine symmetrische reelle $(k \times k)$ -Matrix.

- a) A heißt *positiv definit*, wenn $\langle v, Av \rangle > 0$ für alle $v \in \mathbb{R}^k \setminus \{0\}$ gilt.
- b) A heißt *negativ definit*, wenn $\langle v, Av \rangle < 0$ für alle $v \in \mathbb{R}^k \setminus \{0\}$ gilt.
- c) A heißt *indefinit*, wenn es $v, w \in \mathbb{R}^k$ gibt, so dass $\langle v, Av \rangle > 0$ und $\langle w, Aw \rangle < 0$ gilt.

Die Hessematrix $H_f(x)$ einer C^2 -Funktion f hängt von dem Auswertungspunkt x ab, d.h. wir haben es mit einer matrixwertigen Funktion zu tun. Unsere Vermutung ist, dass positive Definitheit von $H_f(x_0)$ an einem kritischen Punkt x_0 ein lokales Minimum von f bei x_0 impliziert. Wenn wir das per Taylorentwicklung beweisen möchten, werden wir aber mit der Hessematrix an einem nach Satz 12.41 nicht näher bekannten Punkt $x_0 + \theta(x - x_0)$ in der Nähe von x_0 arbeiten müssen und wollen dann gerne sicherstellen, dass die Hessematrix an diesem Punkt ebenfalls positiv definit ist. Dazu betrachten wir zunächst die Menge der positiv definiten Matrizen.

Lemma 12.49. Sei $M_k(\mathbb{R})_{\text{sym}} \subset M_k(\mathbb{R})$ die Menge aller symmetrischen reellen $(k \times k)$ -Matrizen, und sei $M_k(\mathbb{R})_{\pm} \subset M_k(\mathbb{R})_{\text{sym}}$ die Teilmenge aller positiv ($M_k(\mathbb{R})_+$) bzw. negativ ($M_k(\mathbb{R})_-$) definiten Matrizen. Dann ist $M_k(\mathbb{R})_{\pm} \subset M_k(\mathbb{R})_{\text{sym}}$ offen.

Offenheit bezieht sich hier auf eine beliebige Norm auf dem endlichdimensionalen Vektorraum $M_k(\mathbb{R})$.

Beweis. Wir führen den Beweis für positive Definitheit, der Beweis für negative Definitheit ist analog (oder durch die Abbildung $A \mapsto -A$ zu verstehen). Wir verwenden in diesem Beweis die Euklidische Norm auf \mathbb{R}^k und die zugehörige Operatornorm auf $M_k(\mathbb{R})$.

Sei $A \in M_k(\mathbb{R})_+$ positiv definit. Wir müssen $\varepsilon > 0$ finden, so dass die Operatornorm-Umgebung $U_\varepsilon(A) \subset M_k(\mathbb{R})_+$ ganz in der Menge der positiv definiten Matrizen liegt. Dazu betrachten wir zunächst die Einheitssphäre $S := \{v \in \mathbb{R}^k : \|v\| = 1\}$, die abgeschlossen und beschränkt, also kompakt ist (Korollar 10.12). Die Funktion $f : S \rightarrow \mathbb{R}$, $f(v) := \langle v, Av \rangle$ ist stetig, da A stetig ist (lineare Abbildung auf endlichdimensionalem Vektorraum) und das Skalarprodukt stetig ist. Nach Satz 9.17 nimmt f also ein Minimum an. Da A positiv definit ist, gilt $m := \min_{v \in S} \langle v, Av \rangle > 0$.

Wir setzen $\varepsilon := m > 0$ und zeigen nun $U_\varepsilon(A) \subset M_k(\mathbb{R})_+$. Sei also $B \in U_\varepsilon(A)$ beliebig, d.h. $\|A - B\| < \varepsilon$. Durch Abschätzung per Cauchy-Schwarz-Ungleichung und $\|(B - A)v\| \leq$

$\|B - A\| \|v\|$ (siehe Def. der Operatornorm) erhalten wir für alle $v \in S$

$$\begin{aligned} \langle v, Bv \rangle &= \langle v, Av \rangle - \langle v, (A - B)v \rangle \\ &\geq m - \|(A - B)v\| \geq m - \|A - B\| > m - \varepsilon = 0. \end{aligned}$$

Also gilt für alle Vektoren $w \neq 0$ per Skalierung $\langle w, Bw \rangle = \|w\|^2 \langle \frac{w}{\|w\|}, B \frac{w}{\|w\|} \rangle > 0$, d.h. B ist positiv definit. \square

Nun können wir den Beweis unseres vermuteten Kriteriums leicht führen.

Satz 12.50 (Hinreichende Bedingung für lokale Extrema.). Sei $U \subset \mathbb{R}^k$ offen, $f \in C^2(U, \mathbb{R})$ und $x_0 \in U$ ein kritischer Punkt von f . Dann gilt:

- a) Ist $H_f(x_0)$ positiv definit, so hat f in x_0 ein isoliertes lokales Minimum.
- b) Ist $H_f(x_0)$ negativ definit, so hat f in x_0 ein isoliertes lokales Maximum.
- c) Ist $H_f(x_0)$ indefinit, so hat f in x_0 kein lokales Extremum (sondern einen Sattelpunkt).

Beweis. Wir machen eine Taylorentwicklung zweiter Ordnung von f um x_0 . Dazu wählen wir $h \in \mathbb{R}^k$ so klein, dass die gerade Verbindungsstrecke zwischen x_0 und $x_0 + h$ ganz in U enthalten ist. Nach Satz 12.41 existiert $\theta \in (0, 1)$, so dass

$$f(x_0 + h) = f(x_0) + \frac{1}{2} \langle h, H_f(x_0 + \theta h) h \rangle$$

gilt. Hier haben wir bereits $\nabla f(x_0) = 0$ (kritischer Punkt) benutzt.

a) Die Matrix $H_f(x_0)$ ist positiv definit, und $U \ni x \mapsto H_f(x) \in M_k(\mathbb{R})$ ist stetig, da $f \in C^2(U, \mathbb{R})$ (stetige partielle Ableitungen zweiter Ordnung). Da Urbilder offener Mengen unter stetigen Abbildungen offen sind, ist $H_f^{-1}(M_k(\mathbb{R})_+) \subset U$ offen. Es gibt also $\varepsilon > 0$, so dass die Hessematrix $H_f(x)$ für alle x mit $\|x - x_0\| < \varepsilon$ positiv definit ist. Wählen wir h genügend klein, so ist die in dem Restglied der Taylorentwicklung auftretende Hessematrix $H_f(x_0 + \theta h)$ also positiv definit. Damit erhalten wir $f(x_0 + h) - f(x_0) = \frac{1}{2} \langle h, H_f(x_0 + \theta h) h \rangle > 0$, also $f(x_0 + h) > f(x_0)$ für alle h in einer ε -Umgebung von x_0 . Damit ist x_0 als isoliertes lokales Minimum identifiziert.

b) ergibt sich sofort aus a), wenn man $-f$ statt f betrachtet.

c) Sei $H_f(x_0)$ indefinit. Dann gibt es zwei Vektoren $v, w \in \mathbb{R}^k \setminus \{0\}$, so dass $\langle v, H_f(x_0)v \rangle > 0$ und $\langle w, H_f(x_0)w \rangle < 0$. Wir betrachten unsere Funktion f in der Nähe von x_0 in Richtung von v bzw. w . Da U offen ist, finden wir $\varepsilon > 0$ so, dass $x_0 + tv$ und $x_0 + tw$ für alle $t \in (-\varepsilon, \varepsilon)$ stets in U liegen. Die Funktionen $g_v(t) := f(x_0 + tv)$ und $g_w(t) := f(x_0 + tw)$ sind dann auf $(-\varepsilon, \varepsilon)$ wohldefiniert und C^2 . Die ersten Ableitungen sind

$$\begin{aligned} g'_v(t) &= (\partial_v f)(x_0 + tv) = \langle (\nabla f)(x_0 + tv), v \rangle, \\ g'_w(t) &= (\partial_w f)(x_0 + tw) = \langle (\nabla f)(x_0 + tw), w \rangle. \end{aligned}$$

Da x_0 ein kritischer Punkt von f ist, haben wir $(\nabla f)(x_0) = 0$ und deshalb $g'_v(0) = 0$ und $g'_w(0) = 0$. Für die zweiten Ableitungen haben wir

$$g''_v(t) = \frac{d}{dt} \sum_{j=1}^k (\partial_j f)(x_0 + tv) v_j = \sum_{i,j=1}^k (\partial_i \partial_j f)(x_0 + tv) v_j v_i = \langle v, H_f(x_0 + tv)v \rangle$$

und für $t = 0$ insbesondere $g_v''(0) = \langle v, H_f(x_0)v \rangle > 0$. Völlig analog erhalten wir $g_w''(0) = \langle w, H_f(x_0)w \rangle < 0$. Also hat g_v bei $t = 0$ ein isoliertes lokales Minimum und g_w hat bei $t = 0$ ein isoliertes lokales Maximum. Da g_v bzw. g_w die Einschränkung von f auf die Gerade durch x_0 in Richtung v bzw. w ist, kann f in x_0 kein lokales Extremum haben. \square

Wir haben nun ein hinreichendes Kriterium für lokale Extrema, und es wäre wohl Zeit für ein Beispiel. In einem Beispiel werden wir dann vor der Aufgabe stehen, zu entscheiden, ob eine Matrix (die Hessematrix einer Beispielfunktion an einem Punkt) positiv definit, negativ definit, oder indefinit ist. Das prüft man nur selten durch Verifikation der definierenden Ungleichung, so dass wir lieber erst ein Ergebnis aus der linearen Algebra in Erinnerung rufen (siehe Sätze 17.15 und 17.22 aus LA2).

Satz 12.51. Sei A eine symmetrische reelle $(k \times k)$ -Matrix. Dann sind die folgenden Aussagen äquivalent:

- a) A ist positiv definit.
- b) Alle Eigenwerte λ von A erfüllen $\lambda > 0$.
- c) Alle Hauptminoren $\det(A_l) > 0$ sind positiv, $l = 1, \dots, k$ (Hurwitz-Kriterium).

Erinnerung: Der Hauptminor $\det A_l$ ist die Determinante der linken oberen $(l \times l)$ -Untermatrix von A .

Um eine symmetrische Matrix A auf *negative* Definitheit zu prüfen, prüfen Sie einfach $-A$ auf positive Definitheit. Insbesondere ist A negativ definit genau dann, wenn alle Eigenwerte negativ sind. Eine symmetrische Matrix ist *indefinit*, wenn Sie sowohl positive als auch negative Eigenwerte hat (siehe Lineare Algebra).

Sie haben also stets mehrere Möglichkeiten zum Überprüfen einer Matrix auf positive Definitheit. Die Eigenwerte einer $(k \times k)$ -Matrix werden Sie nur selten explizit berechnen können, weil Sie dafür die Nullstellen des charakteristischen Polynoms benötigen. Die anderen beiden Möglichkeiten (Prüfen von $\langle v, Av \rangle > 0$ und Hurwitz-Kriterium) kommen ohne die Kenntnis der Eigenwerte aus.

Beispiel 12.52. Wir möchten die lokalen Extrema der Funktion

$$f : \mathbb{R}^2 \rightarrow \mathbb{R}, \quad f(x, y) := (x^2 + 2y^2)e^{-x^2 - y^2}$$

bestimmen. Dazu bestimmen wir zunächst alle kritischen Punkte, da dies die Kandidaten für lokale Extrema sind. Der Gradient von f ist

$$(\nabla f)(x, y) = \begin{pmatrix} x(1 - x^2 - 2y^2) \\ y(2 - x^2 - 2y^2) \end{pmatrix} \cdot 2e^{-x^2 - y^2}.$$

Nullsetzen des Gradienten liefert die Lösungen $(0, 0)$, $(0, 1)$, $(0, -1)$ für $x = 0$. Für $x \neq 0$ ergibt die erste Gleichung $x^2 + 2y^2 = 1$, was eingesetzt in die zweite Gleichung $y = 0$ impliziert. Das führt dann auf die zwei weiteren Lösungen $(1, 0)$ und $(-1, 0)$. Es gibt also genau fünf kritische Punkte, nämlich $(0, 0)$, $(0, 1)$, $(0, -1)$, $(1, 0)$, $(-1, 0)$.

Nun betrachten wir die Hessematrix von f , die sich als

$$H_f(x, y) = 2e^{-x^2-y^2} \cdot \begin{pmatrix} 1 - 5x^2 - 2y^2 + 2x^4 + 4x^2y^2 & -2xy(3 - x^2 - 2y^2) \\ -2xy(3 - x^2 - 2y^2) & 2 - x^2 - 10y^2 + 2x^2y^2 + 4y^4 \end{pmatrix}$$

berechnet (Prüfen Sie es nach). An den fünf kritischen Punkten haben wir

$$H_f(0, 0) = 2 \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}, \quad H_f(0, 1) = H_f(0, -1) = \frac{2}{e} \begin{pmatrix} -1 & 0 \\ 0 & -4 \end{pmatrix},$$

$$H_f(1, 0) = H_f(-1, 0) = \frac{2}{e} \begin{pmatrix} -2 & 0 \\ 0 & 1 \end{pmatrix}.$$

In diesem Beispiel sind alle Hessematrizen diagonal, so dass wir die Eigenwerte als Diagonaleinträge bequem ablesen können: $H_f(0, 0)$ ist positiv definit, also hat f in $(0, 0)$ ein lokales Minimum. $H_f(0, \pm 1)$ ist negativ definit, in $(0, 1)$ und $(0, -1)$ liegen also lokale Maxima vor. $H_f(\pm 1, 0)$ ist indefinit (je ein positiver und ein negativer Eigenwert), also liegen an diesen Stellen keine lokalen Extrema vor.

Im Allgemeinen sind Hessematrizen aber nicht diagonal, siehe Übungen für weitere Beispiele.

Um unsere Diskussion von lokalen Extrema abzuschließen, betrachten wir noch einmal Satz 12.50. Machen Sie sich klar, dass die dort aufgeführten drei Fälle a)–c) *nicht* alle Möglichkeiten ausschöpfen: Es gibt symmetrische Matrizen, die weder positiv noch negativ definit und auch nicht indefinit sind. Dies sind die semidefiniten aber nicht definiten Matrizen.

Definition 12.53. Eine reelle symmetrische $(k \times k)$ -Matrix A heißt *positiv (negativ) semidefinit*, falls $\langle v, Av \rangle \geq 0$ (bzw. $\langle v, Av \rangle \leq 0$) für alle $v \in \mathbb{R}^k$ gilt.

Aus der Linearen Algebra wissen Sie (Satz 17.15):

Lemma 12.54. Eine symmetrische reelle $(k \times k)$ -Matrix ist genau dann positiv (negativ) semidefinit, wenn ihre Eigenwerte λ alle $\lambda \geq 0$ (bzw. $\lambda \leq 0$) erfüllen.

Das Hurwitzkriterium verallgemeinert sich nicht direkt auf ein Kriterium für positive Semidefinitheit.

In Satz 12.50 brauchen Sie positive/negative Definitheit, nicht nur Semidefinitheit! Das folgende Beispiel macht das klar.

Beispiel 12.55. Die drei Funktionen $f, g, h : \mathbb{R}^2 \rightarrow \mathbb{R}$, $f(x, y) = x^2 + y^4$, $g(x, y) = x^2$, $h(x, y) = x^2 + y^3$ haben alle $(0, 0)$ als kritischen Punkt und haben dort die gleiche Hessematrix

$$H_f(0, 0) = H_g(0, 0) = H_h(0, 0) = \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix},$$

die positiv semidefinit ist (ein Eigenwert $2 > 0$, ein Eigenwert 0). Aber f hat in $(0, 0)$ ein lokales isoliertes Minimum, g hat in $(0, 0)$ ein nicht isoliertes Minimum, und h hat

in $(0, 0)$ einen Sattelpunkt.

Zeigen Sie: Hat $f \in C^2(U, \mathbb{R})$ in x_0 ein lokales Minimum, so ist $H_f(x_0)$ positiv semidefinit.

Beispiel 12.56. Sei $Q : \mathbb{R}^2 \rightarrow \mathbb{R}$, $Q(x, y) = 1 - x^3 - y^2 + x^3y^2$. Dann ist $Q \in C^\infty(\mathbb{R}^2, \mathbb{R})$, mit Gradient

$$(\nabla Q)(x, y) = \begin{pmatrix} 3x^2(-1 + y^2) \\ 2y(-1 + x^3) \end{pmatrix}.$$

Wir bestimmen die kritischen Punkte von Q : Für $x = 0$ verschwindet die erste Komponente des Gradienten, und dann zeigt Betrachtung der zweiten Komponente, dass $y = 0$ gelten muss. Ein kritischer Punkt ist also $(0, 0)$. Ist $y = 0$, so verschwindet die zweite Komponente, und Betrachtung der ersten Komponente führt dann wieder auf $x = 0$. Für $x \neq 0$, $y \neq 0$ sehen wir leicht die Lösungen $y = \pm 1$ und $x = 1$.

Wir haben also die drei kritischen Punkte $(0, 0)$, $(1, 1)$, $(1, -1)$. Die Hessematrix ist

$$H_Q(x, y) = \begin{pmatrix} 6x(-1 + y^2) & 6x^2y \\ 6x^2y & 2(-1 + x^3) \end{pmatrix}.$$

Für $x = y = 0$ haben wir also

$$H_Q(0, 0) = \begin{pmatrix} 0 & 0 \\ 0 & -2 \end{pmatrix},$$

eine negativ semidefinite Hessematrix (ein Eigenwert 0, ein Eigenwert $-2 < 0$). Mit der Übung oben auf dieser Seite folgt, dass $(0, 0)$ kein lokales Minimum ist. Aber wir erfahren aus unseren Kriterien nicht, ob es ein lokales Maximum oder ein Sattelpunkt ist und müssen "per Hand" nachsehen. Dazu gibt es kein allgemeines Rezept. Hier betrachten wir $Q(0, 0) = 1$ und $Q(x, 0) = 1 - x^3$. Wir sehen, dass $(0, 0)$ kein lokales Extremum ist: Für beliebig kleines $x > 0$ gilt $Q(x, 0) < 1 = Q(0, 0)$, und für (betragsmäßig) beliebig kleines $x < 0$ gilt $Q(x, 0) > Q(0, 0)$.

An dem kritischen Punkt $(1, 1)$ haben wir die Hessematrix

$$H_Q(1, 1) = \begin{pmatrix} 0 & 6 \\ 6 & 0 \end{pmatrix}.$$

Der obere linke Eintrag von $H_Q(1, 1)$ (Hauptminor $H_Q(1, 1)_1$) ist 0, also ist $H_Q(1, 1)$ nicht positiv definit. Natürlich ist auch der obere linke Eintrag von $-H_Q(1, 1)$ gleich 0, also ist $H_Q(1, 1)$ auch nicht negativ definit. Das hat also noch nicht viel gebracht. Gucken wir die Eigenwerte an: Das charakteristische Polynom ist $\chi(\lambda) = \lambda^2 - 36$, mit den Nullstellen $\lambda = \pm 6$. Es gibt also einen positiven und einen negativen Eigenwert, d.h. $H_Q(1, 1)$ ist indefinit und damit ist $(1, 1)$ ein Sattelpunkt.

Die Untersuchung von $(1, -1)$ wird als Übung überlassen. Auch dies ist ein Sattelpunkt.

13 Umkehrfunktionen, implizite Funktionen, Untermannigfaltigkeiten

In diesem Kapitel beschäftigen wir uns mit der recht allgemeinen Frage der Lösbarkeit von Gleichungssystemen. Wir werden mehrere Variable $x_1, \dots, x_k \in \mathbb{R}$ (bzw $x = (x_1, \dots, x_k) \in U \subset \mathbb{R}^k$) und mehrere Gleichungen $f_1(x) = y_1, \dots, f_d(x) = y_d$ betrachten, die wir in einer Vektorgleichung $f(x) = y \in \mathbb{R}^d$ zusammenfassen.

Typische Fragen sind: Gibt es Lösungen? Was können wir über die Menge der Lösungen aussagen? Wie hängen Eigenschaften der Lösungsmenge $\{x \in \mathbb{R}^k : f(x) = y\}$ von Eigenschaften von f ab? Können wir eventuell f invertieren und etwas über die Umkehrfunktion $f^{-1}(y) = x$ aussagen?

Zuerst studieren wir Umkehrfunktionen und ihre Differenzierbarkeitseigenschaften. Ausgerüstet mit dem mächtigen Satz über die Umkehrfunktion leiten wir dann den Satz über implizite Funktionen her, der sich mit unterbestimmten Gleichungssystemen (mehr Variable als Gleichungen) befasst und die Lösungsmenge durch eine implizit definiert Funktion beschreibt. Solche Lösungsmengen sind oft sogenannte Untermannigfaltigkeiten des \mathbb{R}^d , die Sie sich als glatte Linien, Flächen und höherdimensionale Verallgemeinerungen davon vorstellen können. Wir werden sehen, wie Untermannigfaltigkeiten durch Funktionen beschrieben werden können. Eine wichtige praktische Anwendung stellt der Abschnitt über Extreme unter Nebenbedingungen dar, in dem Extrema von Einschränkungen von Funktionen $f : U \rightarrow \mathbb{R}$ auf Untermannigfaltigkeiten $M \subset U$ diskutiert werden.

13.1 Umkehrfunktionen linearer Abbildungen

Wir beginnen unsere Betrachtungen zu Lösungen von Gleichungssystemen und Umkehrfunktionen mit dem vergleichsweise einfachen Fall von *linearen* Gleichungssystemen, die Sie gut aus der Linearen Algebra kennen.

Sei also $f : \mathbb{R}^k \rightarrow \mathbb{R}^d$ linear und $y \in \mathbb{R}^d$. Nach Wahl von Basen dürfen wir f mit einer $(d \times k)$ -Matrix A identifizieren, $f(x) = Ax$. Wir betrachten die Gleichung

$$Ax = y.$$

Ist $d > k$, so ist das Gleichungssystem überbestimmt (mehr Gleichungen als Variable). Im Allgemeinen gibt es in diesem Fall keine Lösung, da das Bild von A ein höchstens k -dimensionaler Unterraum von \mathbb{R}^d ist, und y nicht in diesem Unterraum liegen muss.

Im Fall $d = k$ (und nur in diesem Fall) kann die Matrix A invertierbar sein. Dies ist genau dann der Fall, wenn $\det(A) \neq 0$ gilt (siehe Lineare Algebra). Ist A invertierbar, so hat das Gleichungssystem eine eindeutige Lösung, nämlich $x = A^{-1}y$.

Im Fall $d < k$ ist das Gleichungssystem unterbestimmt (weniger Gleichungen als Variable) und deshalb $\dim \ker A > 0$. Wenn es eine Lösung x gibt, ist auch jeder Vektor der Form $x + u$ mit $Au = 0$ eine Lösung.

Wir haben also Grund, uns etwas an invertierbare Matrizen zu erinnern. Wir identifizieren hier und im Folgenden oft lineare Abbildungen mit Matrizen, was durch Wahl von Basen gerechtfertigt ist.

Definition 13.1. Sei $k \in \mathbb{N}$. Die allgemeine lineare Gruppe ist

$$\mathrm{GL}_k(\mathbb{R}) := \{A \in M_k(\mathbb{R}) : A \text{ invertierbar}\} = \{A \in M_k(\mathbb{R}) : \det A \neq 0\}.$$

Die Determinante $\det : M_k(\mathbb{R}) \rightarrow \mathbb{R}$ wird also eine wichtige Rolle spielen. Wir hatten bereits gesehen (Hausaufgabe H5.2):

- \det ist stetig.
- $\mathrm{GL}_k(\mathbb{R}) \subset M_k(\mathbb{R})$ ist offen.

Der Grund für die Stetigkeit von \det ist, dass $\det(A)$ ein Polynom in den Einträgen von A ist. Damit ist \det nicht nur stetig, sondern sogar beliebig oft stetig differenzierbar (glatt).

Wir fügen hier noch zwei weitere Beobachtungen zu invertierbaren Matrizen hinzu. Diese sind zum einen Werkzeuge für den Beweis des Satzes über die Umkehrfunktion, zum anderen für sich genommen interessante und wichtige Resultate.

Satz 13.2. Die Inversion

$$\mathrm{GL}_k(\mathbb{R}) \rightarrow \mathrm{GL}_k(\mathbb{R}), \quad A \mapsto A^{-1}$$

ist eine glatte Funktion.

Beweis. Wir wissen, dass Differenzierbarkeit komponentenweise geprüft werden kann. Nach der Cramer'schen Regel aus der linearen Algebra gilt

$$(A^{-1})_{ij} = \frac{(-1)^{i+j}}{\det A} \det(A_{[ji]}).$$

Hierbei bezeichnet $A_{[ji]}$ die $(k-1) \times (k-1)$ -Matrix, die durch Streichung der j -ten Zeile und i -ten Spalte aus A hervorgeht.

Wir wissen bereits, dass $A \mapsto (-1)^{i+j} \det(A_{[ji]})$ und \det glatte Funktionen (Polynome in den Einträgen von A) sind. Weiterhin hat \det keine Nullstellen auf $\mathrm{GL}_k(\mathbb{R})$. Also ist $A \mapsto \frac{1}{\det A}$ als Verkettung der glatten Funktionen \det und $\mathbb{R} \setminus \{0\} \ni x \mapsto \frac{1}{x}$ glatt. Per Produktregel sehen wir nun, dass $A \mapsto (A^{-1})_{ij}$ glatt ist. \square

Der nächste Satz ist eine Verallgemeinerung der geometrischen Reihe auf Matrizen. Erinnern wir uns an die geometrische Reihe: Für $a \in \mathbb{R}$ mit $|a| < 1$ konvergiert die Reihe $\sum_{n=0}^{\infty} a^n$ absolut, und zwar gegen $\sum_{n=0}^{\infty} a^n = (1-a)^{-1}$.

Für Matrizen (und allgemein lineare Abbildungen) heißt die geometrische Reihe auch *Neumannreihe*. Die in diesem Satz auftretende Norm ist die Operatornorm bzgl einer beliebigen Norm auf \mathbb{R}^k , zB der Euklidischen Norm. Da quadratische Matrizen A potenziert werden können, ist $A^n = A \cdots A$ (n Faktoren) für $n \in \mathbb{N}$ wohldefiniert. Für $n = 0$ setzen wir $A^0 := \mathbb{1}$, die $(n \times n)$ -Einheitsmatrix.

Satz 13.3. Sei $A \in M_k(\mathbb{R})$ mit $\|A\| < 1$. Dann ist $\mathbb{1} - A$ invertierbar und die Neu-

mannreihe

$$\sum_{n=0}^{\infty} A^n = (\mathbb{1} - A)^{-1}$$

konvergiert gegen die zu $(\mathbb{1} - A)$ inverse Matrix.

Beweis. Erinnern Sie sich, dass für zwei $(k \times k)$ -Matrizen A, B

$$\|AB\| \leq \|A\| \|B\|$$

gilt. Beweis: Für beliebiges $v \in \mathbb{R}^k$ gilt

$$\|ABv\| \leq \|A\| \|Bv\| \leq \|A\| \|B\| \|v\|,$$

also $\|AB\| = \sup_{v \neq 0} \frac{\|ABv\|}{\|v\|} \leq \|A\| \|B\|$.

Induktiv erhalten wir so $\|A^n\| \leq \|A\|^n$. Die Folge der Partialsummen $S_N := \sum_{n=0}^N A^n$ erfüllt also (mit $M > N$)

$$\begin{aligned} \|S_M - S_N\| &= \left\| \sum_{n=N+1}^M A^n \right\| \leq \sum_{n=N+1}^M \|A^n\| \leq \sum_{n=N+1}^M \|A\|^n \leq \sum_{n=N+1}^{\infty} \|A\|^n \\ &= \frac{1}{1 - \|A\|} - \sum_{n=0}^N \|A\|^n \xrightarrow{N \rightarrow \infty} 0, \end{aligned}$$

d.h. $(S_N)_N$ ist eine Cauchyfolge in $M_k(\mathbb{R})$ bzgl. Operatornorm. Als endlichdimensionaler normierter Vektorraum ist $M_k(\mathbb{R})$ vollständig (ein Banachraum). Also konvergiert die Cauchyfolge $(S_N)_N$, d.h. ihr Limes $S := \sum_{n=0}^{\infty} A^n$ existiert als wohldefinierte Matrix in $M_k(\mathbb{R})$.

Wir wollen nun $S = (\mathbb{1} - A)^{-1}$ zeigen. Dazu betrachten wir

$$(\mathbb{1} - A)S_N = \sum_{n=0}^N A^n - \sum_{n=1}^{N+1} A^n = \mathbb{1} - A^{N+1},$$

also $\|(\mathbb{1} - A)S_N - \mathbb{1}\| \leq \|A\|^{N+1} \rightarrow 0$. Das zeigt $(\mathbb{1} - A)S = \mathbb{1}$, was $S = (\mathbb{1} - A)^{-1}$ impliziert (Erinnerung an LA: Determinantenbildung ergibt $1 = \det((\mathbb{1} - A)S) = \det(\mathbb{1} - A) \det(S)$, also $\det(\mathbb{1} - A) \neq 0$). Also ist $(\mathbb{1} - A)$ invertierbar, und Multiplikation von links mit $(\mathbb{1} - A)^{-1}$ liefert dann $S = (\mathbb{1} - A)^{-1}$. \square

Beispiel 13.4. Sei

$$A = \frac{1}{2} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Dann gilt $A \begin{pmatrix} x \\ y \end{pmatrix} = \frac{1}{2} \begin{pmatrix} y \\ x \end{pmatrix}$, also in Euklidischer Norm $\left\| A \begin{pmatrix} x \\ y \end{pmatrix} \right\|^2 = \frac{1}{4}(x^2 + y^2)$, d.h. $\|Av\| = \frac{1}{2}\|v\|$ für alle $v \in \mathbb{R}^2$. Insbesondere ist die von der Euklidischen Norm definierte Operatornorm $\|A\| = \frac{1}{2}$.

Um die Neumannreihe zu bestimmen, beachten wir $A^2 = \frac{1}{4} \cdot \mathbb{1}$ (mit der (2×2) -Einheitsmatrix $\mathbb{1}$), und deshalb $A^{2n} = (\frac{1}{4})^n \mathbb{1}$ und $A^{2n+1} = (\frac{1}{4})^n A$. So erhalten wir

$$\begin{aligned} \sum_{n=0}^{\infty} A^n &= \sum_{n=0}^{\infty} A^{2n} + \sum_{n=0}^{\infty} A^{2n+1} = \sum_{n=0}^{\infty} (\frac{1}{4})^n \mathbb{1} + \sum_{n=0}^{\infty} (\frac{1}{4})^n A \\ &= \frac{1}{1 - \frac{1}{4}} (\mathbb{1} + A) = \frac{4}{3} \begin{pmatrix} 1 & \frac{1}{2} \\ \frac{1}{2} & 1 \end{pmatrix}. \end{aligned}$$

Um die Behauptung des Satzes zu prüfen, bemerken wir, dass $\mathbb{1} - A = \begin{pmatrix} 1 & -\frac{1}{2} \\ -\frac{1}{2} & 1 \end{pmatrix}$. Determinante $\det(\mathbb{1} - A) = \frac{3}{4} \neq 0$ hat, also invertierbar ist. Die inverse Matrix ist (siehe LA) $(\mathbb{1} - A)^{-1} = \frac{1}{\det A} \begin{pmatrix} 1 & -\frac{1}{2} \\ -\frac{1}{2} & 1 \end{pmatrix}$, was mit der Berechnung via Neumannreihe übereinstimmt.

13.2 Der Satz über die Umkehrfunktion

Wir wenden uns jetzt allgemeinen (also üblicherweise nicht linearen) Gleichungssystemen zu, d.h. wir betrachten Gleichungen der Form $f(x) = y$, wobei $f : U \subset \mathbb{R}^k \rightarrow \mathbb{R}^d$ stetig differenzierbar sein soll. Beachten Sie, dass selbst wenn $f : U \rightarrow f(U)$ invertierbar (d.h. bijektiv) ist, die Umkehrfunktion $f^{-1} : f(U) \rightarrow U$ nicht differenzierbar sein muss. Ein eindimensionales Beispiel belegt das:

Beispiel 13.5. Die Funktion $f : \mathbb{R} \rightarrow \mathbb{R}$, $f(x) = x^3$ ist glatt (klar, da Polynom) und bijektiv, aber ihre Umkehrfunktion $f^{-1} : \mathbb{R} \rightarrow \mathbb{R}$ ist in $y = 0$ nicht differenzierbar. Beweis: Wäre f^{-1} differenzierbar, so hätten wir wegen $y = f(f^{-1}(y))$ durch Ableiten die Formel $1 = f'(f^{-1}(y)) \cdot (f^{-1})'(y)$, also insbesondere $f'(f^{-1}(y)) \neq 0$. Aber $f'(0) = 3 \cdot 0^2 = 0$.

Wir führen deshalb einen Begriff für invertierbare stetig differenzierbare Funktionen mit stetig differenzierbarer Umkehrfunktion ein.

Definition 13.6. Seien $U \subset \mathbb{R}^k$ und $V \subset \mathbb{R}^d$ offen. Eine Funktion $f : U \rightarrow V$ heißt *Diffeomorphismus*, falls

- f stetig differenzierbar ist,
- f bijektiv ist, und
- $f^{-1} : V \rightarrow U$ ebenfalls stetig differenzierbar ist.

In der Analysis 1 haben wir gelernt (Satz 5.9), dass eine auf einem Intervall I definierte bijektive stetig differenzierbare Funktion f , deren Ableitung keine Nullstellen hat, eine stetig differenzierbare Umkehrfunktion hat, also ein Diffeomorphismus ist. Die Ableitung der Umkehrfunktion berechnet sich dann als $(f^{-1})'(f(x)) = \frac{1}{f'(x)}$. Die Verallgemeinerung dieser Aussage auf mehrere Variablen ist deutlich komplizierter und wird uns im Folgenden beschäftigen.

Zuerst zeigen wir, dass die richtige Verallgemeinerung von "nichtverschwindende Ableitung" "invertierbares Differential" ist.

Lemma 13.7. Seien $U \subset \mathbb{R}^k$ und $V \subset \mathbb{R}^d$ offen und $f : U \rightarrow V$ ein Diffeomorphismus. Dann gilt $k = d$ sowie $\det(Df)(x) \neq 0$ und $(Df)(x)^{-1} = (Df^{-1})(f(x))$ für alle $x \in U$.

Beweis. Da f invertierbar ist, gilt $f^{-1} \circ f = \text{id}_U$, d.h. $f^{-1}(f(x)) = x$ für alle $x \in U$. Sowohl f als auch f^{-1} sind differenzierbar. Also können wir die Kettenregel anwenden (Satz 12.9 b) und erhalten

$$\text{id}_{\mathbb{R}^k} = (Df^{-1})(f(x)) \circ (Df)(x).$$

Dies impliziert, dass $(Df)(x)$ injektiv und $(Df^{-1})(f(x))$ surjektiv ist. Betrachten wir auch $y = f(f^{-1}(y))$, ergibt sich analog $\text{id}_{\mathbb{R}^d} = (Df)(f^{-1}(y)) \circ (Df^{-1})(y)$. Also ist $(Df)(f^{-1}(y))$ auch surjektiv. Für $y = f(x)$ erkennen wir also, dass $(Df)(x)$ bijektiv, d.h. invertierbar ist. Da $(Df)(x)$ eine lineare Abbildung $\mathbb{R}^k \rightarrow \mathbb{R}^d$ ist, folgt $k = d$ (siehe Lineare Algebra) und $(Df)(x)^{-1} = (Df^{-1})(f(x))$. \square

Dieses Lemma zeigt insbesondere, dass Invertierbarkeit des Differentials eine notwendige Eigenschaft eines Diffeomorphismus ist. Das nächste Lemma studiert die Situation, in der die Stetigkeit der Umkehrfunktion bekannt ist und (stetige) Differenzierbarkeit der Umkehrfunktion etabliert werden soll.

Lemma 13.8. Seien $U, V \subset \mathbb{R}^d$ offen, $f : U \rightarrow V$ bijektiv und differenzierbar, und $f^{-1} : V \rightarrow U$ stetig^a. Dann gilt:

- a) Ist $(Df)(x)$ für ein $x \in U$ invertierbar, so ist f^{-1} in $f(x)$ differenzierbar.
- b) Ist $(Df)(x)$ für alle $x \in U$ invertierbar, so ist f^{-1} stetig differenzierbar, d.h. f ist ein Diffeomorphismus.

^aEine stetige bijektive Funktion mit stetiger Umkehrfunktion heißt auch *Homöomorphismus*.

Beweis. a) Wir verwenden die Charakterisierung von Differenzierbarkeit aus Lemma 12.8. Es gibt also eine offene Umgebung $U_\varepsilon(0) \subset \mathbb{R}^d$ von 0 und eine in 0 stetige Funktion $\Phi : U_\varepsilon(0) \rightarrow \text{Hom}(\mathbb{R}^d, \mathbb{R}^d)$, so dass

$$f(x+h) = f(x) + \Phi(h)h$$

und $\Phi(0) = (Df)(x)$. Eine solche Darstellung möchten wir nun auch für f^{-1} bekommen.

Da $\Phi(0)$ invertierbar ist (also $\Phi(0) \in \text{GL}_d(\mathbb{R})$), ist auch $\Phi(h) \in \text{GL}_d(\mathbb{R})$ für hinreichend kleines h (d.h. $\|h\| < \delta$ für geeignetes $\delta > 0$). Hier haben wir die Offenheit von $\text{GL}_d(\mathbb{R})$ und die Stetigkeit von Φ benutzt.

Für $\|h\| < \delta$ können wir dann die Umkehrfunktion f^{-1} benutzen, um

$$f^{-1}(f(x+h)) - f^{-1}(f(x)) = x+h - x = h = \Phi(h)^{-1}(f(x+h) - f(x))$$

zu erhalten. Setzen wir $y := f(x)$ und $y + \tilde{h} := f(x+h)$, so erhalten wir

$$f^{-1}(y + \tilde{h}) - f^{-1}(y) = \Psi(\tilde{h})\tilde{h}, \quad \Psi(\tilde{h}) := \Phi(f^{-1}(y + \tilde{h}) - x)^{-1}.$$

Da f^{-1} als stetig vorausgesetzt wird, ist $\tilde{h} \mapsto f^{-1}(y + \tilde{h}) - x$ stetig. Da die Inversion einer Matrix nach Satz 13.2 ebenfalls stetig ist, erkennen wir Ψ bei $\tilde{h} = 0$ als stetig. Also ist f^{-1} in $y = f(x)$ differenzierbar, mit $(Df^{-1})(f(x)) = \Psi(0) = \Phi(0)^{-1} = (Df)(x)^{-1}$.

b) Wir wissen aus Teil a) bereits, dass f^{-1} an jedem Punkt $y \in V$ differenzierbar ist mit Differential $(Df^{-1})(y) = (Df)(f^{-1}(y))^{-1}$. Da die drei Abbildungen $f^{-1} : V \rightarrow U$, $Df : U \rightarrow \text{GL}_d(\mathbb{R})$ und die Inversion $\text{GL}_d(\mathbb{R}) \rightarrow \text{GL}_d(\mathbb{R})$ alle stetig sind, ist auch Df^{-1} stetig. \square

In einer Variablen gilt, dass eine stetig differenzierbare surjektive Funktion f mit nichtverschwindender Ableitung auch injektiv ist. Denn dann gilt nach dem Zwischenwertsatz entweder $f' > 0$ oder $f' < 0$, und f ist streng monoton, also injektiv. In mehreren Variablen gilt diese Aussage allerdings nicht mehr, wie das folgende Beispiel zeigt.

Beispiel 13.9. Wir betrachten den Koordinatenwechsel auf Polarkoordinaten, d.h. die Funktion

$$f : \mathbb{R}_+ \times \mathbb{R} \rightarrow \mathbb{R}^2 \setminus \{(0, 0)\}, \quad f(r, \varphi) := \begin{pmatrix} r \cos \varphi \\ r \sin \varphi \end{pmatrix}.$$

Diese Funktion ist stetig differenzierbar, mit Jacobimatrix

$$J_{(r, \varphi)}(f) = \begin{pmatrix} \cos \varphi & -r \sin \varphi \\ \sin \varphi & r \cos \varphi \end{pmatrix}.$$

Wegen $\det J_{(r, \varphi)}(f) = r \neq 0$ ist das Differential von f überall invertierbar. f ist zwar surjektiv, aber wegen der 2π -Periodizität der Winkelfunktionen nicht injektiv, es gilt $f(r, 0) = f(r, 2\pi)$. Insbesondere ist f kein Diffeomorphismus.

Allerdings können wir invertieren, wenn wir f nur "lokal betrachten", d.h. den Definitionsbereich auf eine geeignete Umgebung eines Punktes (r_0, φ_0) einschränken. In diesem Beispiel betrachten wir den einen beliebigen Radius $r_0 > 0$ und Winkel $-\frac{\pi}{2} < \varphi_0 < \frac{\pi}{2}$ enthaltenden offenen Definitionsbereich

$$U := \mathbb{R}_+ \times \left(-\frac{\pi}{2}, \frac{\pi}{2}\right).$$

Eingeschränkt auf U ist $f|_U : U \rightarrow f(U)$ natürlich immer noch stetig differenzierbar, per Definition des Bildbereiches surjektiv, und außerdem injektiv (sieht man am Einfachsten durch die Eindeutigkeit der Polardarstellung, siehe Satz 4.48).

Das Bild von $f|_U$ ist die rechte Halbebene $f(U) = \{(x, y) : x > 0\}$, die wegen der echten Ungleichung $x > 0$ offen ist. Die Umkehrfunktion $(f|_U)^{-1} : f(U) \rightarrow U$ der Einschränkung ist also auf einer offenen Menge definiert. Durch Umrechnung der Polarkoordinaten auf kartesische Koordinaten erhalten wir

$$(f|_U)^{-1}(x, y) = \left(\sqrt{x^2 + y^2}, \arctan \frac{y}{x}\right).$$

Beachten Sie, dass wegen der Einschränkung auf $x > 0$ der Arcustangens wohldefiniert ist und die Wurzel nicht an dem problematischen Punkt 0 ausgewertet wird. $(f|_U)^{-1}$ ist also stetig differenzierbar, d.h. $f|_U$ ist ein Diffeomorphismus.

In diesem sehr instruktiven Beispiel haben wir gelernt:

- Für Dimension $k > 1$ garantiert ein invertierbares Differential $(Df)(x)$ nicht, dass eine C^1 -Funktion injektiv ist. Der Grund ist, dass man von einem Punkt $x \in \mathbb{R}^k$ über einen Weg zu x zurückkehren kann, "ohne Umzukehren". In obigem Beispiel können wir einen Kreisweg wählen, der die Periodizität der Winkelfunktionen ausnutzt.
- Immerhin ist wegen Lemma 13.7 die Invertierbarkeit des Differentials eine *notwendige* Eigenschaft eines Diffeomorphismus.
- Im Beispiel haben wir gesehen, dass wir die Nicht-Injektivität von f durch Einschränkung auf einen kleineren Definitionsbereich beseitigen können.

Der Satz über die Umkehrfunktion besagt, dass die im letzten Punkt erwähnte "lokale Invertierbarkeit" für Funktionen mit invertierbarem Differential immer gilt. Er ist ein zentrales Ergebnis der Analysis, das in vielen Gebieten Anwendung findet.

Für eine knappere Sprechweise machen wir zunächst folgende Definition.

Definition 13.10. Eine C^1 -Funktion $f : U \subset \mathbb{R}^d \rightarrow \mathbb{R}^d$ heißt *lokal um* $x_0 \in U$ *invertierbar*, wenn es offene Umgebungen $U_1 \subset U$ und $V_1 \subset \mathbb{R}^d$ gibt, so dass die Einschränkung $f|_{U_1} : U_1 \rightarrow V_1$ ein Diffeomorphismus ist. Die Umkehrfunktion der Einschränkung^a, also

$$(f|_{U_1})^{-1} : V_1 \rightarrow U_1,$$

heißt dann *lokale Umkehrfunktion* von f . Ist f sogar in jedem Punkt $x_0 \in U$ lokal invertierbar, so heißt f *lokaler Diffeomorphismus*.

^aGenaugenommen handelt es sich nicht um die Einschränkung von f , d.h. die Funktion $f|_{U_1} : U_1 \rightarrow \mathbb{R}^d$, sondern um die koeingeschränkte Einschränkung, nämlich $f : U_1 \rightarrow V_1$, die auch im Bildbereich eingeschränkt ist. Im Folgenden werden wir trotzdem einfach von "Einschränkung" sprechen.

Satz 13.11 (Satz von der Umkehrfunktion).

Sei $U \subset \mathbb{R}^d$ offen, $x_0 \in U$, und $f : U \rightarrow \mathbb{R}^d$ eine C^1 -Abbildung. Dann ist f lokal um x_0 invertierbar genau dann wenn $(Df)(x_0)$ invertierbar ist.

Bevor wir uns an den Beweis machen, einige Bemerkungen:

- Punkte x_0 mit invertierbarem $(Df)(x_0)$ heißen auch *reguläre Punkte* von f .
- Sei $f : U \rightarrow \mathbb{R}^d$ stetig differenzierbar, $x_0 \in U$ regulär, und $y_0 = f(x_0)$. Dann besagt der Satz über die Umkehrfunktion, dass das Gleichungssystem $f(x) = y$ lokal eine eindeutige Lösung hat: Zu jedem y nahe y_0 gibt es ein eindeutiges x nahe x_0 , so dass $f(x) = y$.
- Über *globale* Invertierbarkeit, also Invertierbarkeit von f ohne Einschränkung auf ein kleineres Gebiet, wird keine Aussage gemacht. Wir haben bereits gesehen, dass globale Invertierbarkeit im Allgemeinen nicht zu erwarten ist. Es wird auch keine Aussage über *alle* Lösungen der Gleichung $f(x) = y$ gemacht, sondern nur über Lösungen in der Nähe von einem regulären Punkt.

- Es wird keine Methode zur Berechnung der Lösung geliefert. Das ist auch nicht zu erwarten, zB ist es schon für Polynome genügend hohen Grades nicht möglich, die Nullstellen analytisch zu bestimmen.
- Die Voraussetzung (x_0 regulärer Punkt) bringt die lokale Invertierbarkeit von f mit der Invertierbarkeit der linearen Abbildung $(Df)(x)$ in Verbindung. Die Aussage über Existenz einer eindeutigen Lösung erinnert uns weiterhin an den Banach'schen Fixpunktsatz, den wir im Beweis verwenden werden.

Beispiel 13.12. Wir betrachten das Gleichungssystem

$$x_1 e^{x_1 x_2} = y_1, \quad x_2 \cos(x_1^2 + 1) = y_2.$$

Gegeben (y_1, y_2) , gibt es eine (eindeutige?) Lösung (x_1, x_2) ? Für $y_1 = y_2 = 0$ sehen wir ohne große Schwierigkeiten, dass $x_1 = x_2 = 0$ die eindeutige Lösung ist, aber wie sieht es für $(y_1, y_2) \neq (0, 0)$ aus?

Mit der Funktion

$$f : \mathbb{R}^2 \rightarrow \mathbb{R}^2, \quad f(x_1, x_2) = \begin{pmatrix} x_1 e^{x_1 x_2} \\ x_2 \cos(x_1^2 + 1) \end{pmatrix},$$

die wir routinemäßig als C^1 (sogar glatt) erkennen, ist das Gleichungssystem $f(x) = y$. Die Jacobimatrix von f ist

$$J_x(f) = \begin{pmatrix} (1 + x_1 x_2) e^{x_1 x_2} & x_1^2 e^{x_1 x_2} \\ -2x_1 x_2 \sin(x_1^2 + 1) & \cos(x_1^2 + 1) \end{pmatrix},$$

bei $x = 0$ also $J_0(f) = \mathbb{1}$ (Einheitsmatrix). Also ist 0 ein regulärer Punkt von f . Der Satz über die Umkehrfunktion besagt nun, dass es eine offene Umgebung U_1 von $x = 0$ und eine offene Umgebung V_1 von $y = f(0) = 0$ gibt, so dass für $y \in V_1$ eine eindeutige Lösung $x \in U_1$ der Gleichung $f(x) = y$ existiert, und dass die lokale Umkehrfunktion von f stetig differenzierbar ist.

Nach diesen Bemerkungen und Beispielen machen wir uns nun an den Beweis des Satzes über die Umkehrfunktion.

Beweis. Wir haben schon in Lemma 13.7 gesehen, dass Invertierbarkeit von $(Df)(x_0)$ notwendig für die lokale Invertierbarkeit von f ist, müssen aber noch die umgekehrte Implikation zeigen.

(1) Transformation auf vereinfachte Situation. Zuerst wollen wir die Situation etwas vereinfachen, indem wir die auftretenden Daten $x_0, f(x_0), (Df)(x_0)$ auf die spezielle Situation $x_0 = 0, f(x_0) = 0, (Df)(x_0) = \mathbb{1}$ transformieren. Wieso genügt es, diesen speziellen Fall zu betrachten? Angenommen, wir haben den Satz in diesem Spezialfall bewiesen. Für allgemeines $x_0, f(x_0), (Df)(x_0)$ betrachten wir dann

$$g(x) := (Df)(x_0)^{-1}(f(x + x_0) - f(x_0)).$$

Diese Funktion erfüllt $g(0) = 0$, und per Kettenregel erkennen wir g als differenzierbar mit

Differential

$$(Dg)(x) = (Df)(x_0)^{-1}(Df)(x + x_0) \Rightarrow (Dg)(0) = \mathbb{1}.$$

Die Funktion g erfüllt also die Voraussetzungen des Spezialfalls, und per Annahme finden wir offene Umgebungen U_1^g von 0 , V_1^g von 0 , so dass $\hat{g} := g|_{U_1^g} : U_1^g \rightarrow V_1^g$ ein Diffeomorphismus ist. Nun gilt

$$f(x) = (Df)(x_0)g(x - x_0) + f(x_0).$$

Mit den offenen Mengen $U_1 := U_1^g + x_0$ und $V_1 := (Df)(x_0)V_1^g + f(x_0)$ ist dann $f : U_1 \rightarrow V_1$ bijektiv und C^1 . Weiterhin haben wir eine lokale C^1 -Umkehrfunktion von $\hat{g} : U_1^g \rightarrow V_1^g$, d.h. für $y \in V_1^g$ gibt es ein eindeutiges $x = \hat{g}^{-1}(y) \in U_1^g$. Das ergibt die lokale Umkehrfunktion $(f|_{U_1})^{-1}$ von $f : U_1 \rightarrow V_1$,

$$x = (\hat{f}|_{U_1})^{-1}(y) = \hat{g}^{-1}((Df)(x_0)^{-1}(y - f(x_0))) + x_0,$$

an der wir erkennen, dass $(f|_{U_1})^{-1}$ stetig differenzierbar ist.

Es genügt also, den Satz für $x_0 = 0$, $f(x_0) = 0$, $(Df)(x_0) = \mathbb{1}$ zu beweisen.

(2) Verwendung des Banachschen Fixpunktsatzes. Wir nehmen nun $x_0 = 0$, $f(0) = 0$, $(Df)(0) = \mathbb{1}$ an und wollen für "kleine" y (in einer offenen Umgebung von $y = 0$) die Gleichung $f(x) = y$ nach x auflösen. Um den Fixpunktsatz verwenden zu können, schreiben wir dieses Problem etwas um: Mit der Funktion

$$F_y : U \rightarrow \mathbb{R}^d, \quad F_y(x) := y + x - f(x)$$

haben wir $f(x) = y$ genau dann wenn $F_y(x) = x$. Wir müssen also Fixpunkte von F_y untersuchen. Dazu müssen wir insbesondere die im Fixpunktsatz vorausgesetzte strikte Kontraktionseigenschaft auf einem geeigneten Definitionsbereich nachweisen.

Wegen $f(0) = 0$ hat F_0 den Fixpunkt 0 . Das Differential von $F_0(x) = x - f(x)$ ist $(DF_0)(x) = \mathbb{1} - (Df)(x)$, also $(DF_0)(0) = \mathbb{1} - \mathbb{1} = 0$. Da F_0 stetig differenzierbar ist, ist DF_0 stetig und es gibt folglich ein $\varepsilon > 0$, so dass $U_{2\varepsilon}(0) \subset U$ und $\|(DF_0)(x)\| \leq \frac{1}{2}$ für alle x mit $\|x\| \leq \varepsilon$. Wir können also den Schrankensatz (Satz 12.20 b)) anwenden und folgern

$$\|F_0(x) - F_0(x')\| \leq \frac{1}{2}\|x - x'\| \quad \text{für alle } x, x' \text{ mit } \|x\|, \|x'\| \leq \varepsilon. \quad (\star)$$

Hier haben wir uns auf Euklidische Normen spezialisiert, um die Konstante C in Satz 12.20 b) zu vermeiden.

Nun betrachten wir die Menge $X := \{x \in \mathbb{R}^d : \|x\| \leq \varepsilon\}$ und, für $\|y\| \leq \frac{\varepsilon}{2}$, die Abbildung F_y . Wir behaupten, dass wir den Banachschen Fixpunktsatz anwenden können. Zuerst prüfen wir, dass F_y die Menge X in sich selbst abbildet. Dazu bemerken wir $F_y(x) = F_0(x) + y$ und die Abschätzung $\|F_0(x)\| \leq \frac{\varepsilon}{2}$, die wir aus (\star) durch Einsetzen von $x' = 0$ erhalten. Damit folgt für $x \in X$

$$\|F_y(x)\| = \|F_0(x) + y\| \leq \|F_0(x)\| + \|y\| \leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon,$$

also $F_y(x) \in X$. Weiterhin ist X ein vollständiger metrischer Raum mit der Euklidischen Norm, da \mathbb{R}^d vollständig und X abgeschlossen ist. Nun zur Kontraktionseigenschaft: Für $x, x' \in X$ gilt aufgrund von (\star)

$$\|F_y(x) - F_y(x')\| = \|F_0(x) + y - (F_0(x') + y)\| = \|F_0(x) - F_0(x')\| \leq \frac{1}{2}\|x - x'\|.$$

Also ist F_y eine strikte Kontraktion, und wir können den Fixpunktsatz anwenden. Als wesentlichen Schritt in unserem Beweis sehen wir also, dass es genau ein $x \in X$ mit $F_y(x) = x$ (äquivalent: $f(x) = y$) gibt.

(3) Konstruktion der offenen Mengen U_1, V_1 und Diffeomorphismus-Eigenschaft. Wir definieren nun die offenen Umgebungen U_1, V_1 von 0, so dass $f|_{U_1} : U_1 \rightarrow V_1$ ein Diffeomorphismus ist. Wir setzen

$$U_1 := \{x : \|x\| < \varepsilon, \|f(x)\| < \frac{\varepsilon}{2}\} = U_\varepsilon(0) \cap f^{-1}(U_{\frac{\varepsilon}{2}}(0));$$

dies ist eine offene Menge, da endliche Schnitte offener Mengen und Urbilder offener Mengen unter stetigen Funktionen (wie f) offen sind.

Da $U_1 \subset X$, wissen wir bereits, dass $f : U_1 \rightarrow f(U_1)$ bijektiv ist (injektiv nach Teil (2) des Beweises, surjektiv durch Wahl des Bildbereiches). Aber wir müssen noch zeigen, dass $f(U_1)$ offen ist. Dazu betrachten wir $x, x' \in X$ und schätzen mit Hilfe von $x = F_0(x) + f(x)$ ab:

$$\begin{aligned} \|x - x'\| &= \|F_0(x) + f(x) - F_0(x') - f(x')\| \\ &\leq \|F_0(x) - F_0(x')\| + \|f(x) - f(x')\| \\ &\stackrel{\star}{\leq} \frac{1}{2}\|x - x'\| + \|f(x) - f(x')\|. \end{aligned}$$

Durch Umstellen ergibt dies die wichtige Ungleichung

$$\|x - x'\| \leq 2\|f(x) - f(x')\|. \quad (\star\star)$$

Setzen wir $x' = 0$ und beachten $f(0) = 0$, so sehen wir, dass $\|x\| < \varepsilon$ eine Konsequenz von $\|f(x)\| < \frac{\varepsilon}{2}$ ist. Also ist das Bild $f(U_1)$

$$V_1 := f(U_1) = \{f(x) : \|x\| < \varepsilon, \|f(x)\| < \frac{\varepsilon}{2}\} = \{f(x) : \|f(x)\| < \frac{\varepsilon}{2}\} = U_{\varepsilon/2}(0)$$

offen.

Wir haben nun eine Bijektion $f|_{U_1} : U_1 \rightarrow V_1$ zwischen offenen Umgebungen von 0, und betrachten die Umkehrfunktion $(f|_{U_1})^{-1} : V_1 \rightarrow U_1$. Aufgrund von $(\star\star)$ gilt

$$\|(f|_{U_1})^{-1}(x) - (f|_{U_1})^{-1}(x')\| \leq 2\|x - x'\|.$$

Dies impliziert direkt, dass $(f|_{U_1})^{-1}$ stetig ist.

Wie sieht es mit der stetigen Differenzierbarkeit von $(f|_{U_1})^{-1}$ aus? Wir erinnern uns, dass wir gegen Anfang des Beweises unser ε durch die Bedingung $\|(DF_0)(x)\| \leq \frac{1}{2}$ für alle x mit $\|x\| \leq \varepsilon$ gewählt haben. Wegen $(Df)(x) = \mathbb{1} - (DF_0)(x)$ können wir Satz 13.3 verwenden, um die Invertierbarkeit von $(Df_0)(x)$ zu folgern.

Nun sind alle Voraussetzungen von Lemma 13.8 b) gegeben, und wir können schließen, dass $(f|_{U_1})^{-1}$ ein Diffeomorphismus ist. \square

Wir schließen an diesen Beweis zwei unmittelbare Folgerungen an.

Die erste Folgerung betrifft die Offenheit von Bildern. Beachten Sie, dass eine stetige (oder sogar glatte) auf einer offenen Menge $U \subset \mathbb{R}^d$ definierte Funktion im Allgemeinen kein offenes Bild $f(U)$ hat; ein extremes Gegenbeispiel sind konstante Funktionen mit Bild $\{c\}$.

Korollar 13.13. Sei $U \subset \mathbb{R}^d$ offen und $f \in C^1(U, \mathbb{R}^d)$ so, dass $(Df)(x)$ für alle $x \in U$ invertierbar ist. Dann ist $f(U) \subset \mathbb{R}^d$ offen.

Beweis. Sei $y \in f(U)$ und $x \in U$ ein Urbild, d.h. $f(x) = y$. Nach dem Satz über die Umkehrfunktion gibt es offene Umgebungen $U_1 \subset U$ von x und V_1 von y , so dass $f|_{U_1} : U_1 \rightarrow V_1$ ein Diffeomorphismus ist. Insbesondere $V_1 \subset f(U)$. Also ist $f(U)$ offen. \square

Der Satz über die Umkehrfunktion besagt nur die Existenz von *lokalen* Umkehrfunktionen. Ist f allerdings injektiv, so erhalten wir einen Diffeomorphismus auf dem ganzen Definitionsbereich:

Korollar 13.14. Sei $U \subset \mathbb{R}^d$ offen und $f \in C^1(U, \mathbb{R}^d)$. Ist f injektiv und $(Df)(x)$ für alle $x \in U$ invertierbar, so ist $f : U \rightarrow f(U)$ ein Diffeomorphismus.

Beweis. Nach Voraussetzung ist $f : U \rightarrow f(U)$ bijektiv. Die stetige Differenzierbarkeit kann lokal (in der Nähe eines Punktes $x \in U$) geprüft werden und folgt deshalb aus dem Satz über die Umkehrfunktion. \square

Weitere Beispiele zum Satz über die Umkehrfunktion finden Sie in den Übungen.

13.3 Implizite Funktionen

Wir werden uns nun mit impliziten Funktionen befassen. Grob gesprochen ist die Idee, die Nullstellenmenge einer Funktion F von zwei Variablen lokal als Graphen einer Funktion zu schreiben. Die Frage ist: Gibt es eine "implizit definierte" Funktion f , so dass $F(x, y) = 0$ zu $y = f(x)$ äquivalent ist?

Bevor wir das präzisieren, besprechen wir ein einfaches aber instruktives Beispiel.

Beispiel 13.15. Wir betrachten die Funktion

$$F : \mathbb{R}^2 \rightarrow \mathbb{R}, \quad F(x, y) := x^2 + y^2 - 1$$

und ihre Nullstellenmenge

$$N := \{(x, y) : F(x, y) = 0\} = F^{-1}(0).$$

Diese Menge besteht genau aus den Punkten (x, y) mit $x^2 + y^2 = 1$, d.h. N ist der Einheitskreis in der Ebene. Der Einheitskreis ist nicht Graph einer Funktion $f : [-1, 1] \rightarrow \mathbb{R}$, d.h. $N \neq \Gamma(f) = \{(x, f(x)) : x \in [-1, 1]\}$, da z.B. $(0, 1)$ und $(0, -1)$ in N liegen, in einem Funktionsgraphen es aber zu jedem x ein eindeutiges $y = f(x)$ gibt, so dass $(x, y) \in \Gamma(f)$.

Wir können dieses Problem aber beheben, wenn wir N nur lokal, d.h. in der Nähe eines Punktes $(x_0, y_0) \in N$ betrachten: Sei $-1 < x_0 < 1$ und $y_0 = \sqrt{1 - x_0^2}$. Dann haben wir $(x_0, y_0) \in N$, und der Graph von $f : (-1, 1) \rightarrow \mathbb{R}$, $f(x) = \sqrt{1 - x^2}$, ist eine (offene) Teilmenge von N . Eingeschränkt auf Umgebungen von x_0 und y_0 können wir N also durchaus als Graphen einer Funktion schreiben, die (zumindest in diesem Beispiel) eindeutig festgelegt ist. Dieses f nennen wir dann "implizite Funktion".

Probleme gibt es aber an dem Punkt $(1, 0) \in N$. Hier gibt es keine offenen Umgebungen U von $x_0 = 1$ und V von $y_0 = 0$, so dass N lokal Graph einer Funktion ist, denn für alle $\varepsilon > 0$ liegen die *zwei* Punkte $(1 - \varepsilon, \pm\sqrt{1 - (1 - \varepsilon)^2})$ in N , was in einem Funktionsgraphen nicht passieren kann. Das gleiche Problem tritt für $(-1, 0)$ auf.

Was ist an diesen Punkten besonders? Die partielle Ableitung von F nach y ,

$$(\partial_y F)(x, y) = \partial_y(x^2 + y^2 - 1) = 2y$$

verschwindet genau dann, wenn $y = 0$. Anders formuliert: Für alle Punkte $(x_0, y_0) \in N$ mit $(\partial_y F)(x_0, y_0) \neq 0$ gibt es offene Umgebungen U von x_0 und V von y_0 und eine eindeutige Funktion $f : U \rightarrow V$, so dass

$$F(x, y) = 0 \iff f(x) = y.$$

Der Satz über implizite Funktionen verallgemeinert die im Beispiel besprochene Situation. Wir betrachten zwei vektorwertige Variable x, y . Seien dazu $U \subset \mathbb{R}^k$ und $V \subset \mathbb{R}^d$ offen und

$$F : U \times V \rightarrow \mathbb{R}^d, \quad (x, y) \mapsto F(x, y)$$

stetig differenzierbar. Das Gleichungssystem $F(x, y) = 0$ ist also unterbestimmt ($k + d$ Variable $(x, y) = (x_1, \dots, x_k, y_1, \dots, y_d)$, aber nur d Gleichungen). Das Differential von F ist also eine lineare Abbildung $(DF)(x, y) \in \text{Hom}(\mathbb{R}^{k+d}, \mathbb{R}^d)$, d.h. in Basen eine $d \times (k + d)$ Matrix.

Wir teilen das Differential in zwei Teile auf, die zu den Ableitungen nach x bzw y gehören:

$$(DF)(x, y) = ((D_1F)(x, y), (D_2F)(x, y))$$

$$(DF)(x, y) = \begin{pmatrix} (\partial_{x_1} F_1)(x, y) & \dots & (\partial_{x_k} F_1)(x, y) & (\partial_{y_1} F_1)(x, y) & \dots & (\partial_{y_d} F_1)(x, y) \\ (\partial_{x_1} F_2)(x, y) & \dots & (\partial_{x_k} F_2)(x, y) & (\partial_{y_1} F_2)(x, y) & \dots & (\partial_{y_d} F_2)(x, y) \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ (\partial_{x_1} F_d)(x, y) & \dots & (\partial_{x_k} F_d)(x, y) & (\partial_{y_1} F_d)(x, y) & \dots & (\partial_{y_d} F_d)(x, y) \end{pmatrix}.$$

Hier ist $(D_1F)(x, y)$ die $(d \times k)$ -Matrix, die die partiellen Ableitungen nach den x_j enthält und $(D_2F)(x, y)$ die $(d \times d)$ -Matrix, die die partiellen Ableitungen nach den y_l enthält. Beachten Sie, dass die Aufteilung der Variablen so gemacht ist, dass letztere Matrix quadratisch ist.

Bevor wir den Satz über implizite Funktionen formulieren, motivieren wir die dort gemachten Voraussetzungen durch Betrachtung des Spezialfalls, dass F linear ist. Ist die Abbildung $\mathbb{R}^k \times \mathbb{R}^d \ni (x, y) \mapsto F(x, y) \in \mathbb{R}^d$ linear, so muss sie die Form $F(x, y) = F_1(x) + F_2(y)$ mit linearen Abbildungen $F_1 : \mathbb{R}^k \rightarrow \mathbb{R}^d$ und $F_2 : \mathbb{R}^d \rightarrow \mathbb{R}^d$ haben (denn $(x, y) = (x, 0) + (0, y)$).

Wir fragen, wann die Gleichung $F(x, y) = F_1(x) + F_2(y) = 0$ eindeutig nach y aufgelöst werden kann. Dies ist offenbar genau dann der Fall, wenn F_2 invertierbar ist, und in diesem Fall gilt

$$F(x, y) = 0 \iff y = -F_2^{-1}(F_1(x)).$$

Wie passt das mit Beispiel 13.15 zusammen? Da F linear ist, ist das Differential hier $(DF)(x, y) = F = (F_1, F_2)$, also $(D_2F)(x, y) = F_2$. Die Invertierbarkeit von F_2 ist also äquivalent zur Invertierbarkeit des Differentials $(D_2F)(x, y)$.

Satz 13.16 (Satz über implizite Funktionen). Seien $U \subset \mathbb{R}^k$ und $V \subset \mathbb{R}^d$ offen und $F : U \times V \rightarrow \mathbb{R}^d$ stetig differenzierbar, und $(x_0, y_0) \in U \times V$. Falls

$$F(x_0, y_0) = 0 \quad \text{und} \quad (D_2F)(x_0, y_0) \text{ invertierbar ist}$$

(mit der zuvor eingeführten Aufspaltung $(DF)(x, y) = ((D_1F)(x, y), (D_2F)(x, y))$ des Differentials), so gibt es offene Umgebungen $U_1 \subset U$ von x_0 und $V_1 \subset V$ von y_0 und eine eindeutige stetig differenzierbare Funktion $f : U_1 \rightarrow V_1$, für die

$$F(x, y) = 0 \iff y = f(x) \text{ für alle } x \in U_1, y \in V_1$$

gilt^a. Insbesondere gilt $f(x_0) = y_0$.

^aAnders formuliert: Die Nullstellenmenge $N = \{(x, y) : F(x, y) = 0\}$ erfüllt $N \cap (U_1 \times V_1) = \Gamma(f)$.

Die Funktion f wird dann implizite Funktion genannt.

Beweis. Der Beweis beruht ganz wesentlich auf dem Satz über die Umkehrfunktion. Um den benutzen zu können, brauchen wir eine Abbildung zwischen offenen Mengen in Räumen der gleichen Dimension. Um das zu erreichen, erweitern wir F zu der Abbildung

$$G : U \times V \rightarrow \mathbb{R}^k \times \mathbb{R}^d, \quad G(x, y) := (x, F(x, y)).$$

Das Differential von G ist

$$(DG)(x, y) = \begin{pmatrix} \mathbb{1}_k & 0 \\ (D_1F)(x, y) & (D_2F)(x, y) \end{pmatrix},$$

wobei $\mathbb{1}_k$ die $(k \times k)$ -Einheitsmatrix bezeichnet, die von den partiellen Ableitungen von x nach den x_i stammt. Um die Determinante von $(DG)(x, y)$ zu bestimmen, entwickeln wir nach der ersten Zeile $(1, 0, \dots, 0)$, dann nach der zweiten Zeile, etc. bis zur k -ten Zeile. Dies ergibt an dem Punkt (x_0, y_0)

$$\det(DG)(x_0, y_0) = \det(D_2F)(x_0, y_0) \neq 0,$$

da $(D_2F)(x_0, y_0)$ als invertierbar vorausgesetzt wurde. Nach dem Satz über die Umkehrfunktion ist G also lokal um (x_0, y_0) invertierbar: Es gibt eine offene Umgebung $W \subset U \times V$ von (x_0, y_0) , so dass

$$G|_W : W \rightarrow G(W), \quad G(x, y) = (x, F(x, y))$$

ein Diffeomorphismus ist. Für $(x, y) \in W$ und $(a, b) \in G(W)$ gilt dann $(a, b) = G(x, y)$ genau dann, wenn $(x, y) = (G|_W)^{-1}(a, b)$. Wegen $G(x, y) = (x, F(x, y))$ impliziert das $a = x$, d.h. die lokale Umkehrfunktion $(G|_W)^{-1}$ wirkt auf den ersten k Variablen wie die Identität:

$$(G|_W)^{-1}(a, b) = (a, \psi(a, b)) \quad \text{mit einer } C^1\text{-Funktion } \psi : G(W) \rightarrow \mathbb{R}^d.$$

Da wir an $F(x, y) = 0$ interessiert sind, betrachten wir $b = 0$. Wir haben $G_W(x, y) = (a, 0)$ genau dann wenn $(x, y) = (G|_W)^{-1}(a, 0) = (a, \psi(a, 0))$, also genau dann wenn $x = a$ und $y = \psi(x, 0)$. Unsere gesuchte implizite Funktion sollte also durch

$$f : U_0 := \{x \in \mathbb{R}^k : (x, 0) \in G(W)\} \rightarrow \mathbb{R}^d, \quad f(x) := \psi(x, 0)$$

gegeben sein. Der Definitionsbereich U_0 ist offen, da $G(W)$ offen ist, und es gilt wie gewünscht

$$\begin{aligned} (G|_W)^{-1}(x, 0) &= (x, \psi(x, 0)) = (x, f(x)), & x \in U_0 \\ &\Rightarrow (x, 0) = G_W(x, f(x)) = (x, (F(x, f(x)))) \\ &\Rightarrow F(x, f(x)) = 0. \end{aligned}$$

Gilt andererseits $F(x, y) = 0$ für ein $(x, y) \in W$, so ist $G(x, y) = (x, F(x, y)) = (x, 0)$ und deshalb $(x, y) = (G|_W)^{-1}(x, 0) = (x, f(x))$, also $y = f(x)$.

Wir haben also

$$\{(x, y) \in W : F(x, y) = 0\} = \{(x, f(x)) \in W : x \in U_0\}$$

gezeigt. Es verbleibt, die gewünschten offenen Umgebungen $U_1 \subset U$ von x_0 und $V_1 \subset V$ von y_0 zu finden. Zunächst wählen wir klein genügende offene Umgebungen U'_1 von x_0 und V_1 von y_0 , so dass $U'_1 \times V_1 \subset W$ gilt. Wegen der Stetigkeit von f finden wir weiterhin eine offene Umgebung $U_1 \subset U'_1$ von x_0 mit $f(U_1) \subset V_1$. Nach dem Satz über die Umkehrfunktion haben wir $f|_{U_1} \in C^1(U_1, \mathbb{R}^d)$. Wegen

$$\begin{aligned} \{(x, y) \in U_1 \times V_1 : F(x, y) = 0\} &= \{(x, f(x)) \in U_1 \times V_1 : x \in U_0\} \\ &= \{(x, f(x)) : x \in U_1\} \end{aligned}$$

sind nun alle Behauptungen gezeigt, und der Beweis ist abgeschlossen. □

Beispiel 13.17. Wir betrachten das Gleichungssystem

$$\begin{aligned} x^3 + y^3 + z^3 &= 7, \\ xy + yz + xz &= -2, \end{aligned}$$

das wir gerne nach (y, z) auflösen würden. Global oder explizit geht das meistens nicht, aber der Satz über implizite Funktionen liefert trotzdem wichtige Informationen über lokale Lösbarkeit. Wir sehen, dass $(x_0, y_0, z_0) = (2, -1, 0)$ eine Lösung ist. Um die Lösungsmenge in der Nähe dieses Punktes zu parametrisieren, benutzen wir den Satz über implizite Funktionen.

Dazu schreiben wir unser Gleichungssystem zunächst als Nullstellenmenge um, indem

wir die Funktion

$$F : \mathbb{R}^3 \rightarrow \mathbb{R}^2, \quad F(x, y, z) = \begin{pmatrix} x^3 + y^3 + z^3 - 7 \\ xy + yz + xz + 2 \end{pmatrix}$$

betrachten. Wir gruppieren die Variablen als x und (y, z) , denn die zweite Variable sollte aus dem Raum mit der gleichen Dimension wie der Bildbereich sein (hier 2). Lassen Sie sich nicht von den Variablennamen x, y , die wir in der Formulierung des Satzes über implizite Funktionen verwendet haben, verwirren!

Das Differential (Jacobimatrix) von F ist in aufgeteilter Form

$$(DF)(x, y, z) = \begin{pmatrix} 3x^2 & 3y^2 & 3z^2 \\ y + z & x + z & x + y \end{pmatrix},$$
$$(D_1F)(x, y, z) = \begin{pmatrix} 3x^2 \\ y + z \end{pmatrix}, \quad (D_2F)(x, y, z) = \begin{pmatrix} 3y^2 & 3z^2 \\ x + z & x + y \end{pmatrix}.$$

An dem Punkt $(2, -1, 0)$ ist das Differential bzgl. (y, z) also

$$(D_2F)(2, -1, 0) = \begin{pmatrix} 3 & 0 \\ 2 & 1 \end{pmatrix}.$$

Diese Matrix hat Determinante 3 und ist damit invertierbar.

Nach dem Satz über implizite Funktionen gibt es also eine offene Umgebung $U \subset \mathbb{R}$ von $x_0 = 2$ und eine offene Umgebung $V \subset \mathbb{R}^2$ von $(y_0, z_0) = (-1, 0)$ und eine C^1 -Funktion $f : U \rightarrow V$ mit $f(2) = (-1, 0)$, so dass unser Gleichungssystem nahe (x_0, y_0, z_0) genau durch $f(x) = (y, z)$ gelöst wird.

Betrachten Sie das Beispiel 13.17 noch einmal. Können Sie das gegebene Gleichungssystem nach (x, y) statt nach (y, z) auflösen?

Zeigen Sie: Für hinreichend kleine $u, v \in \mathbb{R}$ lässt sich die Gleichung $e^{\sin(uv)} + u^2 = 2v + 1$ nach v auflösen.

Beispiel 13.18. Wir betrachten die Funktion $F : \mathbb{R}^2 \rightarrow \mathbb{R}$, $F(x, y) := (x - y)^2$. Die Nullstellen von F sind offenbar genau die Punkte (x, x) mit $y = x$. Das Differential $(DF)(x, y) = (2(x - y), 2(y - x))$ verschwindet für $x = y$. Insbesondere gilt $(D_2F)(x, x) = (\partial_y F)(x, x) = 0$; die Invertierbarkeitsbedingung aus dem Satz über implizite Funktionen ist also verletzt. Trotzdem lässt sich die Nullstellenmenge $N = \{(x, y) : F(x, y) = 0\} = \{(x, x) : x \in \mathbb{R}\}$ als Graph einer C^1 -Funktion schreiben (sogar ohne Einschränkung auf kleine Umgebungen), nämlich als Graph von $f(x) = x$.

Dieses Beispiel zeigt, dass die Voraussetzungen im Satz über implizite Funktionen hinreichend, aber nicht notwendig sind.

Der Satz über implizite Funktionen informiert uns über die Existenz und stetige Differenzierbarkeit der impliziten Funktion f . Etwas mehr lässt sich noch sagen: Wir können eine Formel für das Differential von f angeben.

Lemma 13.19 (Implizites Differenzieren). *Mit der gleichen Notation und den gleichen Voraussetzungen wie in Satz 13.16 gilt: Das Differential der impliziten Funktion f ist, $x \in U_1$,*

$$(Df)(x) = -(D_2F)(x, f(x))^{-1}(D_1F)(x, f(x)).$$

Beweis. Per Konstruktion erfüllt die implizite Funktion $F(x, f(x)) = 0$, $x \in U_1$. Differentiation per Kettenregel ergibt

$$0 = (DF)(x, f(x)) \circ \begin{pmatrix} \mathbb{1}_k \\ (Df)(x) \end{pmatrix} = (D_1F)(x, f(x)) + (D_2F)(x, f(x))(Df)(x).$$

Da $(D_2F)(x, f(x))$ invertierbar ist (das kommt von unserem lokalen Diffeomorphismus G im Beweis des Satzes über implizite Funktionen), können wir diese Gleichung leicht nach $(Df)(x)$ umstellen, was die behauptete Formel zeigt. \square

Beispiel 13.20.

- a) Wir betrachten noch einmal Beispiel 13.17. Die Ableitung von f bei $x_0 = 2$ ist

$$\begin{aligned} f'(2) &= -(D_2F)(2, -1, 0)^{-1}(D_1F)(2, -1, 0) \\ &= - \begin{pmatrix} 3 & 0 \\ 2 & 1 \end{pmatrix}^{-1} \begin{pmatrix} 12 \\ -1 \end{pmatrix} = -\frac{1}{3} \begin{pmatrix} 1 & 0 \\ -2 & 3 \end{pmatrix} \begin{pmatrix} 12 \\ -1 \end{pmatrix} = \begin{pmatrix} -4 \\ 9 \end{pmatrix}. \end{aligned}$$

- b) Wir betrachten noch einmal unser einführendes Beispiel 13.15 mit dem Kreis, also $F(x, y) = x^2 + y^2 - 1$. Hier können wir die implizite Funktion zB für $x_0 = 0$, $y_0 = 1$ explizit bestimmen, nämlich $f : (-1, 1) \rightarrow \mathbb{R}$, $f(x) = \sqrt{1 - x^2}$. Aber selbst wenn wir sie nicht explizit bestimmt hätten, lernen wir aus dem Lemma über implizite Differentiation wegen $(DF)(x, y) = (2x, 2y)$ doch

$$f'(x) = -\frac{1}{2f(x)} 2x = -\frac{x}{f(x)}.$$

Man sieht leicht, dass $f(x) = \sqrt{1 - x^2}$ tatsächlich diese *Differentialgleichung* löst. Außerdem können Sie f bestimmen (ohne die Lösung vorher zu kennen), indem Sie die Gleichung $f'(x) = -\frac{x}{f(x)}$ lösen. Wie man so etwas systematisch macht, lernen Sie in einer Vorlesung über Differentialgleichungen.

Für die vorliegende Situation können wir so vorgehen: Unsere Gleichung ist

$$-x = f(x)f'(x) = \frac{1}{2} \frac{d}{dx}(f(x)^2).$$

Durch Stammfunktionbildung erhalten wir $f(x)^2 = -x^2 + c$ mit einer Konstanten c . Einsetzen von $x_0 = 0$ soll $f(x_0) = y_0 = 1$ liefern, was die Konstante als

$c = 1$ identifiziert. Nun sehen wir, dass wir x auf das Intervall $(-1, 1)$ beschränken müssen, so dass $f(x)^2 = 1 - x^2$ positiv ist. Wir erhalten dann die implizite Funktion $f(x) = \sqrt{1 - x^2}$.

13.4 Untermannigfaltigkeiten des \mathbb{R}^d

Wir haben uns bereits mit Kurven im \mathbb{R}^d beschäftigt und wollen nun allgemeiner (gekrümmte) Flächen bzw. höherdimensionale Hyperflächen im \mathbb{R}^d betrachten. Diese Objekte werden wir als sogenannte *Untermannigfaltigkeiten* definieren. Die Idee ist, dass wir Teilmengen $M \subset \mathbb{R}^d$ betrachten werden, die “lokal wie ein Stück verbogener \mathbb{R}^k (mit $k \leq d$) aussehen”. In diesem Kapitel werden wir also Methoden der Analysis benutzen, um geometrische Objekte zu studieren. Wir können dabei nur einige Begriffe und als wichtige Anwendung die Diskussion von Extrema unter Nebenbedingungen diskutieren. In einer Vorlesung über *Differentialgeometrie* werden diese Themen viel detaillierter ausgeführt.

Definition 13.21. Es seien $k, d \in \mathbb{N}$ mit $k \leq d$. Eine *k-dimensionale (differenzierbare) Untermannigfaltigkeit des \mathbb{R}^d* ist eine nicht leere Teilmenge $M \subset \mathbb{R}^d$ mit der folgenden Eigenschaft: Für jeden Punkt $x \in M$ gibt es eine offene Umgebung $U \subset \mathbb{R}^d$ von x und eine offene Teilmenge $V \subset \mathbb{R}^d$ und einen Diffeomorphismus $\varphi : U \rightarrow V$ mit

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

Hier ist $\{0\}$ als Teilmenge von \mathbb{R}^{d-k} aufzufassen.

Dies sollte Ihnen als merkwürdige Definition vorkommen. Sie haben schon mehrfach “Unterobjekte” kennengelernt: Zum Beispiel gibt es Vektorräume und Untervektorräume, Gruppen und Untergruppen, Algebren und Unterhalbgebren, etc. Man definiert immer zunächst ein Objekt O mit gewissen Eigenschaften und nennt dann eine Teilmenge von O Unterobjekt, wenn es die gleichen Eigenschaften wie O hat bzw. all diese Eigenschaften von O erbt. Aber hier definieren wir, was eine Untermannigfaltigkeit ist, ohne zu sagen, was eine Mannigfaltigkeit ist. Die Sprechweise legt nahe, dass \mathbb{R}^d eine Mannigfaltigkeit sein sollte, aber wir geben keine Definition einer Mannigfaltigkeit. Das liegt einzig und allein daran, dass uns das hier zu weit führen würde. Aber seien Sie unbesorgt: Es gibt eine gute Definition einer Mannigfaltigkeit, die Sie in der Differentialgeometrie kennenlernen können. Im Sinne sind alle Untermannigfaltigkeiten auch Mannigfaltigkeiten.

Der Diffeomorphismus φ wird *Umgebungskarte* genannt. Wir erläutern den Begriff anhand einiger Beispiele.

Beispiel 13.22.

a) Sei $G \subset \mathbb{R}^k$ offen und $f : G \rightarrow \mathbb{R}^n$ eine C^1 -Funktion. Dann ist der Graph von f

$$M := \Gamma(f) = \{(x, f(x)) : x \in G\}$$

eine k -dimensionale Untermannigfaltigkeit von \mathbb{R}^d , $d = k + n$.

Um das zu beweisen, wählen wir $U := G \times \mathbb{R}^n$ (eine M enthaltene offene Menge). Als Umgebungskarte betrachten wir $\varphi : U \rightarrow \mathbb{R}^d$, $\varphi(x, y) := (x, y - f(x))$.

Diese Funktion ist sicher C^1 . Es gilt $\varphi(U) \subset U$, denn für alle $(x, y) \in U = G \times \mathbb{R}^n$ ist $\varphi(x, y) = (x, y - f(x)) \in G \times \mathbb{R}^n = U$. Wir behaupten, dass φ bijektiv ist mit Umkehrfunktion $\varphi^{-1}(x, y) = (x, y + f(x))$ (Überprüfen als leichte Übung). An der Form von φ^{-1} sehen wir direkt, dass φ^{-1} ebenfalls C^1 ist, d.h. $\varphi : U \rightarrow V := U$ ist ein Diffeomorphismus.

Es gilt

$$\begin{aligned}\varphi(M \cap U) &= \{\varphi(x, f(x)) : x \in G\} = \{(x, 0) : x \in G\} = G \times \{0\} \\ &= (G \times \mathbb{R}^n) \cap (\mathbb{R}^k \times \{0\}) = V \cap (\mathbb{R}^k \times \{0\})\end{aligned}$$

Also ist $\Gamma(f)$ eine k -dimensionale Untermannigfaltigkeit des \mathbb{R}^d .

Das untypische Feature an diesem Beispiel ist, dass wir mit einer einzigen Karte (offene Menge U) die ganze Untermannigfaltigkeit abdecken können.

b) Als konkretes typischeres Beispiel betrachten wir die n -Sphäre

$$S^n = \{x \in \mathbb{R}^{n+1} : \|x\|_2 = 1\}$$

und behaupten, dass S^n eine n -dimensionale Untermannigfaltigkeit des \mathbb{R}^{n+1} ist. Wie schon am Kreisbeispiel (S^1) gesehen, ist S^n nicht Graph einer Funktion.

Wir könnten diese Behauptung hier direkt auf Grundlage von Definition 13.21 beweisen. Stattdessen betrachten wir der Übersichtlichkeit halber nur den Fall $n = 1$, also $S^1 \subset \mathbb{R}^2$. Wir werden später sehen, dass die Behauptung auch für allgemeines n stimmt.

Der obere/untere Halbkreis $S_{\pm}^1 := S^1 \cap \{(x_1, x_2) \in \mathbb{R}^2 : \pm x_2 > 0\}$ ist der Graph der Funktion $f_{\pm} : (-1, 1) \rightarrow \mathbb{R}$, $f_{\pm}(x_1) := \pm\sqrt{1 - x_1^2}$. Nach dem vorigen Beispiel sind S_{\pm}^1 also eindimensionale Untermannigfaltigkeiten von \mathbb{R}^2 , d.h. die Bedingungen aus Def. 13.21 sind für Punkte $x = (x_1, x_2) \in S^1$ mit $x_2 \neq 0$ erfüllt.

Dieses Argument greift aber nicht für die beiden Punkte $(\pm 1, 0)$. Wir betrachten deshalb auch den linken/rechten Halbkreis $\tilde{S}_{\pm}^1 := S^1 \cap \{(x_1, x_2) : \pm x_1 > 0\}$. Betrachten wir die Funktionen $\tilde{f}_{\pm} : (-1, 1) \rightarrow \mathbb{R}$, $\tilde{f}_{\pm}(x_2) := \pm\sqrt{1 - x_2^2}$, so sehen wir

$$\Gamma(\tilde{f}_{\pm}) = \{(x_2, \tilde{f}_{\pm}(x_2)) : -1 < x_2 < 1\} = \{(x_2, x_1) \in \mathbb{R}^2 : (x_1, x_2) \in \tilde{S}_{\pm}^1\},$$

dh $\tilde{S}_{\pm}^1 = \{(\tilde{f}_{\pm}(x_2), x_2) : -1 < x_2 < 1\}$ ist gleich dem Graphen einer C^1 -Abbildung bis auf Vertauschung der Koordinaten. Ganz analog zum ersten Beispiel lässt sich nun zeigen, dass die Bedingungen aus Def. 13.21 erfüllt sind.

Insgesamt haben wir also gezeigt, dass $S^1 \subset \mathbb{R}^2$ eine eindimensionale Untermannigfaltigkeit ist. In diesem Fall haben wir mehrere Kartengebiete benötigt, um die Mannigfaltigkeit abzudecken.

- c) Die Dimension k einer Untermannigfaltigkeit des \mathbb{R}^d ist per Definition eine Zahl in $\{1, \dots, d\}$ (man könnte auch 0-dimensionale Untermannigfaltigkeiten betrachten, aber das ist nicht sehr interessant). Betrachten wir den Fall $k = d$ und behaupten, dass $M \subset \mathbb{R}^d$ genau dann eine d -dimensionale Untermannigfaltigkeit ist, wenn M offen ist.

Beweis: Sei $M \subset \mathbb{R}^d$ offen und $x \in M$. Wir wählen (siehe Def. 13.21) $U := V := M$ und $\varphi := \text{id}_M$. Dies ist offenbar ein Diffeomorphismus, und

$$\varphi(M \cap U) = M = M \cap \mathbb{R}^d$$

(beachten Sie, dass die $\{0\} \subset \mathbb{R}^{d-k}$ in Def. 13.21 hier wegen $k = d$ entfällt). Also ist M eine d -dimensionale Untermannigfaltigkeit.

Sei andersherum $M \subset \mathbb{R}^d$ eine d -dimensionale Untermannigfaltigkeit und $x \in M$. Dann gibt es offene Mengen $x \in U \subset \mathbb{R}^d$ und $V \subset \mathbb{R}^d$ und einen Diffeomorphismus $\varphi : U \rightarrow V$ mit $\varphi(U \cap M) = V \cap \mathbb{R}^d = V = \varphi(U)$. Also $U \cap M = U$, dh wir haben $U \subset M$. Also haben wir eine offene Umgebung U von x gefunden, die ganz in M liegt, dh M ist offen.

Sie können sich den Diffeomorphismus $\varphi : U \rightarrow V$ als Beschreibung “lokaler Koordinaten” für M vorstellen, indem Sie die kartesischen Koordinaten in $V \cap (\mathbb{R}^k \times \{0\})$ (Linien l , auf denen alle x_i bis auf eines konstant sind) zu den Urbildern $\varphi^{-1}(l)$ auf M zurückziehen.

Zeigen Sie, dass jeder affine Unterraum M des \mathbb{R}^d eine Untermannigfaltigkeit des \mathbb{R}^d ist. Die Dimension von M als Untermannigfaltigkeit ist gleich der Dimension von M als affiner Unterraum.

Es ist auch instruktiv, Beispiele von Mengen zu sehen, die keine Untermannigfaltigkeiten sind.

Beispiel 13.23.

- a) Die Gerade $G := \{(x, x - 10, 0) : x \in \mathbb{R}\}$ ist als Graph einer C^1 -Funktion $\mathbb{R} \rightarrow \mathbb{R}^2$ eine eindimensionale Untermannigfaltigkeit von \mathbb{R}^3 , und S^2 ist eine zweidimensionale Untermannigfaltigkeit von \mathbb{R}^3 . Die Vereinigung $M := G \cup S^2$ ist keine Untermannigfaltigkeit von \mathbb{R}^3 . Beachten Sie, dass die Dimension k von einer Untermannigfaltigkeit M für alle Punkte $x \in M$ gleich sein muss.
- b) Die Menge $M := (\{0\} \times \mathbb{R}) \cap (\mathbb{R} \times \{0\})$ (Vereinigung der x -Achse und y -Achse) ist keine Untermannigfaltigkeit des \mathbb{R}^2 . Das Problem ist hier der Schnittpunkt $(0, 0)$.

Durch Betrachtung von Punkten ungleich $(0, 0)$ sehen wir, dass M höchstens eine eindimensionale Untermannigfaltigkeit sein kann. Wäre das so, so gäbe es eine $(0, 0)$ enthaltende offene Menge $U \subset \mathbb{R}^2$, eine weitere offene Menge $V \subset \mathbb{R}^2$ und einen Diffeomorphismus $\varphi : U \rightarrow V$ mit $\varphi(M \cap U) = V \cap (\mathbb{R} \times \{0\})$. Dann enthält U Punkte der Form $(x, 0)$, $(0, y)$ mit $x, y \neq 0$. Wir haben also $\varphi(x, 0) = (a, 0)$, $\varphi(0, y) = (b, 0)$ mit $a, b \neq 0$. Dies ergibt einen Widerspruch zur Bijektivität von φ .

Im Folgenden sind wir vor allem an *Niveaumengen* von C^1 -Funktionen $f : G \subset \mathbb{R}^d \rightarrow \mathbb{R}^n$ interessiert. Damit sind Urbilder der Form $f^{-1}(\{y\})$ gemeint, die die bereits vorher betrachteten Höhenlinien verallgemeinern.

Wir betrachten zunächst kurz den einfachen Fall, dass $f : \mathbb{R}^d \rightarrow \mathbb{R}^n$ linear ist. Sei $y \in \mathbb{R}^n$. Dann ist $f^{-1}(\{y\}) = \{x \in \mathbb{R}^d : f(x) = y\}$ die Lösungsmenge des linearen Gleichungssystems $f(x) = y$. Falls f surjektiv ist (d.h. falls $\text{rang } f = n$), so ist diese Menge sicher nicht leer, und in diesem Fall ist $f^{-1}(\{y\})$ ein affiner Unterraum der Dimension $d - n$, also eine $(d - n)$ -dimensionale Untermannigfaltigkeit des \mathbb{R}^d .

Für nicht lineare C^1 -Abbildungen erwarten wir, dass das Differential $(Df)(x)$ zu einem gewissen Grad die Rolle der linearen Abbildung übernehmen wird. Wir werden deshalb an surjektiven Differentialen (= Differentialen mit vollem Rang) interessiert sein und machen die folgende Definition.

Definition 13.24. Sei $G \subset \mathbb{R}^d$ offen und $f \in C^1(G, \mathbb{R}^n)$.

- a) Ein *regulärer Punkt von f* ist ein Punkt $x \in G$ mit $\text{rang}(Df)(x) = n$. Ein *kritischer Punkt von f* ist hingegen ein Punkt $x \in G$, der nicht regulär ist (also $\text{rang}(Df)(x) < n$ erfüllt).
- b) Ein *regulärer Wert von f* ist ein Punkt $y \in \mathbb{R}^n$ mit $\text{rang}(Df)(x) = n$ für alle $x \in f^{-1}(\{y\})$.

Bemerkungen:

- Beachten Sie, dass alle $y \in \mathbb{R}^n$, die nicht im Bild von f liegen, regulär sind. In diesem Fall ist $f^{-1}(\{y\}) = \emptyset$ leer, enthält also keine kritischen Punkte.
- Da die lineare Abbildung $(Df)(x) \in \text{Hom}(\mathbb{R}^d, \mathbb{R}^n)$ Rang in $\{0, \dots, d\}$ hat, gibt es für $d < n$ keine regulären Punkte.
- Im Zusammenhang mit lokalen Extrema hatten wir in Kapitel 12 differenzierbare Funktionen $f : G \rightarrow \mathbb{R}$ betrachtet und einen Punkt $x \in G$ *kritischen Punkt* genannt, falls $(Df)(x) = 0$ (Def. 12.45). In diesem Fall gilt natürlich $\text{rang}(Df)(x) = 0$, der Punkt ist also auch kritisch im Sinne von Definition 13.24. Da das Differential einer skalaren Funktion $f : G \rightarrow \mathbb{R}$ an einem Punkt x nur Rang 0 oder Rang 1 haben kann, und $\text{rang}(Df)(x) = 0$ genau dann wenn $(Df)(x) = 0$, passen beide Definitionen zusammen. Definition 13.24 ist aber allgemeiner, da sie sich nicht auf skalare Funktionen beschränkt.

Wir zeigen nun, dass die Niveaumengen regulärer Werte von C^1 -Funktionen Untermannigfaltigkeiten sind. Der Beweis ist eine weitere Anwendung des Satzes über die Umkehrfunktion.

Satz 13.25 (Satz vom regulären Wert). Sei $G \subset \mathbb{R}^d$ offen, $f \in C^1(G, \mathbb{R}^n)$, und $y \in f(G)$ ein regulärer Wert. Dann ist $f^{-1}(\{y\})$ eine $(d - n)$ -dimensionale Untermannigfaltigkeit von \mathbb{R}^d .

Beweis. Sei $a \in f^{-1}(\{y\})$. Da y ein regulärer Wert ist, hat $(Df)(a)$ vollen Rang, also $\text{rang}(Df)(a) = n$. Die $(n \times d)$ -Jacobimatrix hat demnach n linear unabhängige Spalten. Da ein Permutieren von Koordinaten die Untermannigfaltigkeits-Eigenschaft nicht ändert,

dürfen wir ohne Beschränkung der Allgemeinheit annehmen, dass die hinteren (rechts stehenden) n Spalten von $J_a(f)$ linear unabhängig sind. Wenn wir $J_a(f) = (A \ B)$ mit einer $(n \times (d - n))$ -Matrix A und einer $(n \times n)$ -Matrix B schreiben, so ist

$$\det B = \det \left(\frac{\partial f_i}{\partial x_j}(a) \right)_{\substack{i=1, \dots, n \\ j=d-n+1, \dots, d}} \neq 0.$$

Wir betrachten die C^1 -Funktion

$$\varphi : G \rightarrow \mathbb{R}^d, \quad \varphi(x) = (x_1, \dots, x_{d-n}, f_1(x) - y_1, \dots, f_n(x) - y_n).$$

An der Stelle $x = a$ ist die $(d \times d)$ -Jacobimatrix von φ

$$J_a(\varphi) = \begin{pmatrix} \mathbb{1}_{d-n} & 0_{n \times n} \\ A & B \end{pmatrix}.$$

Die Determinante der Jacobimatrix ist also $\det(J_a(\varphi)) = \det B \neq 0$. Also dürfen wir den Satz von der Umkehrfunktion (Satz 13.11) verwenden und schließen, dass φ um a lokal invertierbar ist. Es gibt also eine offene Umgebung U von a , so dass $\varphi|_U : U \rightarrow \varphi(U)$ ein Diffeomorphismus ist. Das ist unsere lokale Karte, denn bei $x = a$ verschwinden die letzten n Komponenten von $\varphi(a)$, also

$$\varphi(f^{-1}(\{y\}) \cap U) = \{(x_1, \dots, x_{d-n}, 0, \dots, 0) : x \in U\} = (\mathbb{R}^{d-n} \times \{0\}) \cap \varphi(U).$$

Da $a \in f^{-1}(\{y\})$ beliebig war, folgt, dass $f^{-1}(\{y\})$ eine $(d - n)$ -dimensionale Untermannigfaltigkeit von \mathbb{R}^d ist. \square

Für die Belange von Analysis 2 ist dieser Satz unsere beste Quelle von Untermannigfaltigkeiten des \mathbb{R}^d . Wir verdeutlichen dies an einigen Beispielen (und weiteren Beispielen in den Übungen).

Beispiel 13.26.

- a) Wir hatten schon behauptet, dass die n -Sphäre S^n eine n -dimensionale Untermannigfaltigkeit von \mathbb{R}^{n+1} ist. Für allgemeines n steht ein Beweis noch aus. Aber mit dem Satz vom regulären Wert ist es sehr einfach: Wir betrachten die Funktion^a $\omega : \mathbb{R}^{n+1} \rightarrow \mathbb{R}$, $\omega(x) := \|x\|_2^2 = x_1^2 + \dots + x_n^2$, die offenbar glatt ist (Polynom). Das Differential von ω ist $(D\omega)(x) = 2x$ (als lineare Abbildung heißt das $(D\omega)(x) : v \mapsto 2\langle x, v \rangle$), was nur für $x = 0$ nicht surjektiv ist.

Betrachten wir $r > 0$, so enthält das Urbild $\omega^{-1}(\{r\})$ nicht den Punkt 0. Also ist jedes $r > 0$ ein regulärer Wert von ω . Da $S^n = \omega^{-1}(\{1\})$, folgt mit dem Satz vom regulären Wert, dass S^n eine n -dimensionale Untermannigfaltigkeit von \mathbb{R}^{n+1} ist.

- b) Sei $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, $f(x, y) := x^4 - y^4$. Dann ist das Differential (Gradient) $(Df)(x, y) = (4x^3, -4y^3)$ genau dann 0, wenn $(x, y) = (0, 0)$. Für $c \neq 0$ sehen wir wegen $(0, 0) \notin f^{-1}(\{c\})$, dass c ein regulärer Wert ist. Nach dem Satz vom regulären Wert ist

$$M_c := \{(x, y) : x^4 - y^4 = c\}, \quad c \neq 0,$$

eine eindimensionale Untermannigfaltigkeit von \mathbb{R}^2 .

Wie sieht es für $c = 0$ aus? Wir haben $M_0 = f^{-1}(\{0\}) = \{(x, x) : x \in$

$\mathbb{R}\} \cup \{(x, -x) : x \in \mathbb{R}\}$. Wie in einem vorigen Beispiel besteht diese Menge aus zwei sich schneidenden Geraden und ist deshalb keine Untermannigfaltigkeit von \mathbb{R}^2 .

^aDieser Name sei deshalb gewählt, weil der schöne Buchstabe omega ω (Großbuchstabe: Ω) bisher leider noch nicht in der Analysis-Vorlesung eingesetzt wurde.

Wir haben nun ein grundlegendes Verständnis von Untermannigfaltigkeiten des \mathbb{R}^d . Grob gesprochen sind dies “differenzierbare” Kurven, Flächen, Hyperflächen im \mathbb{R}^d . Differenzierbarkeit hat immer mit affiner Approximierbarkeit zu tun. Insofern wird es nicht verwundern, dass wir Tangentialräume an Untermannigfaltigkeiten legen können.

Wir betrachten hier stets C^1 -Diffeomorphismen und C^1 -(Unter-)mannigfaltigkeiten. Allgemeiner kann man für $n \in \mathbb{N} \cup \{\infty\}$ ganz analog C^n -Diffeomorphismen und C^n -(Unter-)Mannigfaltigkeiten definieren. Je höher n , desto höher die Regularität dieser Objekte. Das natürlichste Setting ist das der glatten Mannigfaltigkeiten (C^∞), auf die wir hier aber nicht weiter eingehen.

Definition 13.27. Sei $M \subset \mathbb{R}^d$ eine Untermannigfaltigkeit von \mathbb{R}^d und $x \in M$. Ein Vektor $v \in \mathbb{R}^d$ heißt *Tangentialvektor von M in x* , falls es ein $\varepsilon > 0$ und eine stetig differenzierbare Kurve $\gamma : (-\varepsilon, \varepsilon) \rightarrow M$ mit $\gamma(0) = x$ und $\gamma'(0) = v$ gibt. Die Menge $T_x(M)$ aller Tangentialvektoren von M in x heißt *Tangentialraum von M in x* .

Die geometrische Anschauung ist, dass der affine Unterraum $x + T_x(M)$ eine Tangente/ Tangentialebene/ höherdimensionale Verallgemeinerung einer Tangente an die Mannigfaltigkeit M im Punkt $x \in M$ ist. Die Situation ist analog zu der des Graphen $M := \Gamma(f)$ einer Funktion $f : \mathbb{R} \rightarrow \mathbb{R}$: Für $x := (x_0, f(x_0)) \in \Gamma(f)$ ist $x + T_x(M) = \{(a, f'(x_0)a) + (x_0, f(x_0)) : a \in \mathbb{R}\}$ die $\Gamma(f)$ in x tangential berührende Gerade, und

$$T_x(M) = \{(a, f'(x_0)a) : a \in \mathbb{R}\}$$

die Gerade, die parallel zur Tangente und durch den Ursprung $(0, 0)$ verläuft.

Zuerst zeigen wir, dass $T_x(M)$ tatsächlich ein Vektorraum ist.

Satz 13.28. Sei M eine k -dimensionale Untermannigfaltigkeit von \mathbb{R}^d , und $x \in M$.

- $T_x(M)$ ist ein k -dimensionaler Untervektorraum von \mathbb{R}^d .
- Ist $\varphi : U \rightarrow V$ eine Umgebungskarte bei $x \in U$, so gilt

$$T_x(M) = (D\varphi)(x)^{-1}(\mathbb{R}^k \times \{0\}).$$

Beweis. Zuerst bemerken wir, dass $(D\varphi)(x)$ invertierbar ist, da φ ein Diffeomorphismus ist (siehe Lemma 13.7). Die in b) behauptete Gleichung $T_x(M) = (D\varphi)(x)^{-1}(\mathbb{R}^k \times \{0\})$ impliziert insbesondere, dass $T_x(M)$ ein k -dimensionaler Untervektorraum von \mathbb{R}^d ist, also Aussage a).

Sei nun $\gamma : (-\varepsilon, \varepsilon) \rightarrow M \cap U$ eine C^1 -Kurve mit $\gamma(0) = x$, und $v := \gamma'(0) \in \mathbb{R}^d$. Nach Definition von $T_x(M)$ bedeutet das $v \in T_x(M)$. Für den transformierten Weg

$$\varphi \circ \gamma : (-\varepsilon, \varepsilon) \rightarrow \varphi(M \cap U) = V \cap (\mathbb{R}^k \times \{0\})$$

haben wir

$$(\varphi \circ \gamma)'(0) = \lim_{t \rightarrow 0} \frac{1}{t} (\varphi(\gamma(t)) - \varphi(\gamma(0))) \in \mathbb{R}^k \times \{0\}.$$

Andererseits gilt nach Kettenregel

$$(\varphi \circ \gamma)'(0) = (D\varphi)(\gamma(0))\gamma'(0) = (D\varphi)(x)v.$$

Da $(D\varphi)(x)$ invertierbar ist, folgt $v \in (D\varphi)(x)^{-1}(\mathbb{R}^k \times \{0\})$. In der Gleichung in b) haben wir also die Inklusion “ \subset ” gezeigt.

Für die umgekehrte Inklusion “ \supset ” betrachten wir $v \in (D\varphi)(x)^{-1}(\mathbb{R}^k \times \{0\})$ und die Kurve $t \mapsto \varphi(x) + t \cdot (D\varphi)(x)v$. Da $\varphi(U)$ offen ist, gibt es $\varepsilon > 0$, so dass $\varphi(x) + t \cdot (D\varphi)(x)v \in \varphi(U)$ für alle $t \in \mathbb{R}$ mit $|t| < \varepsilon$. Wir können also die Kurve

$$\gamma : (-\varepsilon, \varepsilon) \rightarrow M, \quad \gamma(t) := \varphi^{-1}(\varphi(x) + t \cdot (D\varphi)(x)v)$$

betrachten. Da φ ein Diffeomorphismus ist, ist γ eine C^1 -Kurve. Die Ableitung von γ bei $t = 0$ ist (mit Ableitung der Umkehrfunktion, $(D\varphi)(x)^{-1} = (D\varphi^{-1})(\varphi(x))$)

$$\gamma'(0) = (D\varphi^{-1})(\varphi(x) + 0 \cdot (D\varphi)(x)v) \circ (D\varphi)(x)v = (D\varphi)^{-1}(x)(D\varphi)(x)v = v.$$

Das zeigt $v \in T_x(M)$ und schließt den Beweis ab. □

Wir interessieren uns in den nachfolgenden Anwendungen vor allem für Untermannigfaltigkeiten, die Urbilder von regulären Werten von C^1 -Funktionen sind. Deshalb bestimmen wir die Tangentialräume in diesem Fall noch expliziter.

Lemma 13.29. Sei $G \subset \mathbb{R}^d$ offen, $f \in C^1(G, \mathbb{R}^n)$, und $y \in f(G)$ ein regulärer Wert von f . Für $x \in M := f^{-1}(\{y\})$ ist der Tangentialraum

$$T_x(M) = \ker(Df)(x)$$

der Kern des Differentials von f am Punkt x .

Beweis. Sei $v \in T_x(M)$ und $\gamma : (-\varepsilon, \varepsilon) \rightarrow M$ eine Kurve mit $\gamma(0) = x$ und $\gamma'(0) = v$. Da $M = f^{-1}(\{y\})$, haben wir dann $f(\gamma(t)) = y$ für alle t . Differenzieren dieser Gleichung nach t ergibt $0 = (Df)(\gamma(t))\gamma'(t)$, und durch Einsetzen von $t = 0$ erhalten wir $(Df)(x)v = 0$, d.h. $v \in \ker(Df)(x)$. Dies zeigt $T_x(M) \subset \ker(Df)(x)$.

Wir wissen, dass $T_x(M) \subset \ker(Df)(x)$ ein Untervektorraum der Dimension $(d-n)$ ist. Da x ein regulärer Punkt von f ist, ist $(Df)(x) : \mathbb{R}^d \rightarrow \mathbb{R}^n$ surjektiv. Nach der Dimensionsformel gilt also

$$\dim(\ker(Df)(x)) = d - \text{rang}(Df)(x) = d - n = \dim(T_x(M)).$$

Das impliziert $T_x(M) = \ker(Df)(x)$. □

Beispiel 13.30.

- a) Wir bestimmen den Tangentialraum an die n -Sphäre $M = S^n \subset \mathbb{R}^{n+1}$ in einem Punkt $x \in S^n$. Da $S^n = f^{-1}(\{1\})$ mit $f : \mathbb{R}^{n+1} \rightarrow \mathbb{R}$, $f(x) = \|x\|_2^2$ mit $(Df)(x)v = 2\langle x, v \rangle$, erhalten wir

$$T_x(S^n) = \ker(Df)(x) = \{v \in \mathbb{R}^{n+1} : \langle x, v \rangle = 0\} = \{x\}^\perp.$$

Der Tangentialraum ist das orthogonale Komplement des Vektors x , was genau zu unserer geometrischen Anschauung einer Sphäre passt.

- b) Im Beispiel 13.26 b) hatten wir $M = f^{-1}(\{c\})$ für $c \neq 0$ und $f(x, y) = x^4 - y^4$ betrachtet. Das Differential ist $(Df)(x, y) = 4(x^3, -y^3)$, mit Kern

$$T_{(x,y)}(M) = \ker(Df)(x, y) = \{(u, v) \in \mathbb{R}^2 : x^3u - y^3v = 0\}.$$

Dies ist eine Gerade (x, y sind fest, u, v sind die Koordinaten des Punktes auf der Gerade).

13.5 Extrema unter Nebenbedingungen

In diesem letzten Abschnitt wollen wir eine wichtige Anwendung von Untermannigfaltigkeiten auf die Frage von Extrema unter Nebenbedingungen besprechen. In Abschnitt 12.6 hatten wir uns bereits lokale Extrema von differenzierbaren skalaren Funktionen $f : U \rightarrow \mathbb{R}$ angesehen, die auf offenen Teilmengen $U \subset \mathbb{R}^d$ definiert sind. Häufig möchte man aber nicht lokal extreme Funktionswerte $f(x)$ bestimmen, wenn das Argument $x \in U$ beliebig ist, sondern gleichzeitig eine Nebenbedingung an x fordern. Die Form der Nebenbedingung ergibt sich aus der Aufgabenstellung, wird aber jeweils eine Gleichung sein, die x zu erfüllen hat.

Beispiel 13.31. Wir möchten die lokalen Extrema von $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, $f(x_1, x_2) := x_1^3 - x_2^3$ unter der Nebenbedingung $x_1^2 + x_2^2 = 4$ bestimmen. Das heißt, wir schränken f auf die Menge aller (x_1, x_2) ein, die $x_1^2 + x_2^2 = 4$ erfüllen (nach dem Satz vom regulären Wert bilden diese (x_1, x_2) eine eindimensionale Untermannigfaltigkeit M des \mathbb{R}^2) und fragen nach den lokalen Extrema von $f|_M$.

Für besonders einfache Nebenbedingungen, zum Beispiel für $x_1 + x_2^2 = 4$ (statt $x_1^2 + x_2^2 = 4$) lässt sich die Nebenbedingung nach einer der Variablen auflösen, in diesem Fall als $x_1(x_2) = 4 - x_2^2$. Dann müssen wir nur die Nebenbedingung in die Funktion einsetzen und die lokalen Extrema der Funktion

$$g(x_2) := f(x_1(x_2), x_2) = (4 - x_2^2)^3 - x_2^3$$

mit den Kriterien aus Abschnitt 12.6 bestimmen.

In den meisten Fällen ist eine solche Auflösung der Nebenbedingung aber nicht möglich.

Nach diesem motivierenden Beispiel definieren wir genau, was wir unter Extrema unter Nebenbedingungen verstehen.

Definition 13.32. Sei $G \subset \mathbb{R}^d$ offen, $f \in C^1(G, \mathbb{R})$ und M eine G schneidende Untermannigfaltigkeit des \mathbb{R}^d . Ein Punkt $x_0 \in M \cap G$ heißt *lokales Maximum (Minimum) von f unter der Nebenbedingung M* , wenn ein $\delta > 0$ existiert, so dass $f(x) \leq f(x_0)$ ($f(x) \geq f(x_0)$) für alle $x \in M \cap G$ mit $\|x - x_0\| < \delta$ gilt.

Zuerst geben wir ein notwendiges Kriterium für lokale Extrema unter allgemeinen Nebenbedingungen an.

Lemma 13.33. Sei M eine Untermannigfaltigkeit des \mathbb{R}^d , $G \subset \mathbb{R}^d$ offen, $G \cap M \neq \emptyset$, und $f \in C^1(G, \mathbb{R})$. Hat f in $x_0 \in G \cap M$ ein lokales Extremum, so steht der Gradient $(\nabla f)(x_0)$ senkrecht auf dem Tangentialraum $T_{x_0}(M)$.

Beweis. Für beliebiges $v \in T_{x_0}(M)$ müssen wir zeigen, dass $\langle (\nabla f)(x_0), v \rangle = 0$ gilt.

Sei $v \in T_{x_0}(M)$, dh es gibt eine C^1 -Kurve $\gamma : (-\varepsilon, \varepsilon) \rightarrow M \cap G$ mit $\gamma(0) = x_0$ und $\gamma'(0) = v$. Da γ nach $M \cap G$ abbildet und f in x_0 ein lokales Extremum hat, hat auch $f \circ \gamma : (-\varepsilon, \varepsilon) \rightarrow \mathbb{R}$ ein lokales Extremum bei $t = 0$. Also verschwindet dort die Ableitung,

$$0 = (f \circ \gamma)'(0) = \langle (\nabla f)(\gamma(0)), \gamma'(0) \rangle = \langle (\nabla f)(x_0), v \rangle.$$

□

Ein sehr praktisches Resultat erhalten wir, wenn wir annehmen, dass die Nebenbedingung M eine Niveaufäche eines regulären Wertes einer C^1 -Funktion ist.

Satz 13.34 (Methode der Lagrange-Multiplikatoren). Sei $G \subset \mathbb{R}^d$ offen, $f \in C^1(G, \mathbb{R})$ und $g \in C^1(G, \mathbb{R}^n)$ sowie $y \in g(G)$ ein regulärer Wert von g . Falls f in $x_0 \in M := g^{-1}(\{y\})$ ein lokales Extremum unter der Nebenbedingung M hat, so gibt es $\lambda_1, \dots, \lambda_n \in \mathbb{R}$ mit

$$(\nabla f)(x_0) = \sum_{j=1}^n \lambda_j (\nabla g_j)(x_0).$$

Die Zahlen λ_j heißen *Lagrange-Multiplikatoren*.

Beweis. Aus Lemma 13.33 wissen wir, dass der Gradient $(\nabla f)(x_0)$ senkrecht auf $T_{x_0}(M)$ steht. Betrachten wir das orthogonale Komplement $T_{x_0}(M)^\perp = \{v \in \mathbb{R}^d : v \perp T_{x_0}(M)\}$. Dies ist, genau wie $T_{x_0}(M)$, ein Unterraum von \mathbb{R}^d . Die Dimension des orthogonalen Komplements ist (siehe Lineare Algebra)

$$\dim T_{x_0}(M)^\perp = d - \dim T_{x_0}(M) = d - (d - n) = n,$$

wobei wir $\dim T_{x_0}(M) = \dim M = d - n$ verwendet haben (siehe Satz 13.25 und Satz 13.28). Wir zeigen nun, dass die Vektoren $(\nabla g_j)(x_0)$, $j = 1, \dots, n$, eine Basis von $T_{x_0}(M)^\perp$ bilden; dann folgt die Behauptung direkt als Basisentwicklung von $(\nabla f)(x_0) \in T_{x_0}(M)^\perp$.

Aus Lemma 13.29 wissen wir $T_{x_0}(M) = \ker(Dg)(x_0)$. Also liegt ein Vektor $v \in \mathbb{R}^d$ in $T_{x_0}(M)$ genau dann, wenn

$$0 = (Dg)(x_0)v = \begin{pmatrix} (\nabla g_1)(x_0) \\ \vdots \\ (\nabla g_n)(x_0) \end{pmatrix} \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} = \begin{pmatrix} \langle (\nabla g_1)(x_0)^T, v \rangle \\ \vdots \\ \langle (\nabla g_n)(x_0)^T, v \rangle \end{pmatrix};$$

hier haben wir verwendet, dass der Gradient $(\nabla g_j)(x_0)$ der j -ten Komponente von g die j -te Zeile der Jacobimatrix von g ist.

An dieser Rechnung erkennen wir, dass v in $T_{x_0}(M)$ liegt genau dann, wenn $\langle (\nabla g_j)(x_0)^T, v \rangle = 0$ für alle $j \in \{1, \dots, n\}$ gilt. Also haben wir $(\nabla g_j)(x_0) \in T_{x_0}(M)^\perp$ für alle j . Da x_0 ein regulärer Punkt von g ist, hat $(Dg)(x_0)$ vollen Rang n . Das bedeutet (siehe Lineare Algebra), dass die Jacobimatrix $J_{x_0}(g)$ genau n linear unabhängige Zeilen hat. Diese Zeilen sind die $(\nabla g_j)(x_0)$. Da sie alle in $T_{x_0}(M)^\perp$ liegen, bilden sie eine Basis von $T_{x_0}(M)^\perp$ und der Beweis ist abgeschlossen. \square

Wir erläutern die Methode an zwei Beispielen.

Beispiel 13.35. Wir betrachten die Funktion $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, $f(x, y) = x \cdot y$, und fragen nach den Extrema von f auf dem Kreis $S^1 = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 = 1\}$.

Die Mannigfaltigkeit $M := S^1$ ist Urbild des regulären Wertes 1 der Funktion $g(x, y) := x^2 + y^2$. Wir können also Satz 13.34 verwenden und wissen, dass an jedem lokalen Extremum (x_0, y_0) von f unter der Nebenbedingung S^1 ein $\lambda \in \mathbb{R}$ existiert, so dass $(\nabla f)(x_0, y_0) = \lambda(\nabla g)(x_0, y_0)$ (in diesem Fall hat $g = g_1$ nur eine Komponente). Das heißt konkret

$$(\nabla f)(x_0, y_0) = \begin{pmatrix} y_0 \\ x_0 \end{pmatrix} = \lambda(\nabla g)(x_0, y_0) = \begin{pmatrix} 2\lambda x_0 \\ 2\lambda y_0 \end{pmatrix}.$$

Wir erhalten also die drei Gleichungen

$$\begin{aligned} y_0 &= 2\lambda x_0, \\ x_0 &= 2\lambda y_0, \\ x_0^2 + y_0^2 &= 1 \quad (\text{die Nebenbedingung}). \end{aligned}$$

Setzen wir die ersten beiden Gleichungen in die Nebenbedingung ein, so erhalten wir

$$1 = x_0^2 + y_0^2 = (2\lambda y_0)^2 + (2\lambda x_0)^2 = 4\lambda^2(x_0^2 + y_0^2) = 4\lambda^2,$$

was $\lambda = \pm \frac{1}{2}$ impliziert. Nun ergeben die ersten beiden Gleichungen $x_0 = \pm y_0$, eingesetzt in die Nebenbedingung $1 = x_0^2 + y_0^2 = 2x_0^2$, also $|x_0| = |y_0| = \frac{1}{\sqrt{2}}$.

Die Lösungen sind also die vier Punkte

$$\begin{aligned}x_0 &= y_0 = \frac{1}{\sqrt{2}}, \\x_0 &= y_0 = -\frac{1}{\sqrt{2}}, \\x_0 &= -y_0 = \frac{1}{\sqrt{2}}, \\x_0 &= -y_0 = -\frac{1}{\sqrt{2}}.\end{aligned}$$

An diesen vier Punkten *könnten* lokale Extrema von f unter der Nebenbedingung S^1 vorliegen (*müssen* aber im Allgemeinen nicht). Die Funktionswerte an diesen Punkten sind

$$\begin{aligned}f\left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right) &= f\left(-\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right) = \frac{1}{2}, \\f\left(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right) &= f\left(-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right) = -\frac{1}{2}.\end{aligned}$$

Wir wissen, dass $S^1 \subset \mathbb{R}^2$ kompakt ist (also abgeschlossen und beschränkt. Beweis zur Erinnerung: Die Punkte $p = (x, y) \in S^1$ erfüllen $\|p\|_2 = 1$, also ist S^1 offenbar beschränkt. Ist $(p_n) \subset S^1$ eine in \mathbb{R}^2 konvergente Folge, so gilt $\|p_n\|_2 = 1$ für alle n , wegen der Stetigkeit der Norm also auch $\|\lim_n p_n\|_2 = \lim_n \|p_n\|_2 = 1$, d.h. $\lim_n p_n \in S^1$. Also ist S^1 abgeschlossen.)

Da f stetig ist (sogar C^1), nimmt $f|_{S^1}$ ein Minimum und Maximum an (Satz 9.17). Also hat $f|_{S^1}$ mindestens ein lokales (sogar globales) Maximum, und mindestens ein lokales (sogar globales) Minimum. Diese müssen nach der obigen Argumentation an einigen der vier berechneten Punkte angenommen werden. Da die Funktionswerte dort $\pm \frac{1}{2}$ sind, folgt, dass bei $(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}})$ und $(-\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}})$ lokale Maxima und bei $(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}})$ und $(-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}})$ lokale Minima vorliegen.

Praktische Anwendung des Lagrange-Verfahrens: Dieses Verfahren lässt sich am besten so merken: Gegeben eine skalare C^1 -Funktion $f : G \subset \mathbb{R}^d \rightarrow \mathbb{R}$ (mit G offen), so schreiben wir die Nebenbedingung als System von Gleichungen $g_1(x) = 0, \dots, g_n(x) = 0$ (mit C^1 -Funktionen $g_1, \dots, g_n : G \rightarrow \mathbb{R}$, so dass 0 ein regulärer Wert von $g = (g_1, \dots, g_n) : G \rightarrow \mathbb{R}^n$) ist. Dann betrachten wir die von den Variablen und den Lagrange-Multiplikatoren abhängige Hilfsfunktion⁴ $JL : G \times \mathbb{R}^n \rightarrow \mathbb{R}$,

$$JL(x_1, \dots, x_d, \lambda_1, \dots, \lambda_n) := f(x_1, \dots, x_d) + \lambda_1 g_1(x_1, \dots, x_d) + \dots + \lambda_n g_n(x_1, \dots, x_d)$$

und bilden die partiellen Ableitungen $\frac{\partial JL}{\partial x_i}$, $i = 1, \dots, d$ und $\frac{\partial JL}{\partial \lambda_j}$, $j = 1, \dots, n$. Sie setzen dann alle diesen partiellen Ableitungen = 0 (Überzeugen Sie sich, dass dies genau die Bedingung aus Satz 13.34 ist!) und bestimmen die Lösungen dieses Gleichungssystems. Nur an

⁴“JL” zu Ehren von Joseph-Louis Lagrange (1736-1813), dem wir nicht nur den Satz über die Lagrange-Multiplikatoren, sondern noch sehr viel mehr Ergebnisse verdanken, zB das Lagrangesche Restglied der Taylorentwicklung, die Euler-Lagrange-Gleichungen, die Lagrangesche Mechanik, den Satz von Lagrange aus der Gruppentheorie, und die Lagrange-Punkte der Himmelsmechanik.

den sich ergebenden Lösungen (x_1, \dots, x_d) (die Lagrange-Multiplikatoren interessieren dann nicht mehr) können lokale Extrema von f unter der Nebenbedingung g vorliegen. Durch weitere Argumente (zB stetige Funktion auf Kompaktum nimmt Maximum/Minimum an) können Sie dann ggf. prüfen, ob dies tatsächlich lokale Extrema sind.

Beispiel 13.36. Wir fragen uns nach dem Maximum und Minimum (falls es existiert) der Funktion $f : \mathbb{R}^3 \rightarrow \mathbb{R}$, $f(x, y, z) = 5x + y - 3z$ auf dem Schnitt der Ebene $x + y + z = 0$ mit der Kugeloberfläche $x^2 + y^2 + z^2 = 1$.

In diesem Fall haben wir die Nebenbedingungsfunktion

$$g : \mathbb{R}^3 \rightarrow \mathbb{R}^2, \quad g(x, y, z) = \begin{pmatrix} x + y + z \\ x^2 + y^2 + z^2 - 1 \end{pmatrix},$$

die so formuliert ist, dass die Nebenbedingung $g(x, y, z) = 0$ entspricht. Zuerst prüfen wir, ob 0 ein regulärer Wert von g ist: Die Jacobimatrix von g ist

$$(Dg)(x, y, z) = \begin{pmatrix} 1 & 1 & 1 \\ 2x & 2y & 2z \end{pmatrix}.$$

Diese Matrix ist surjektiv, dh hat Rang 2, falls es zwei linear unabhängige Spalten (oder Zeilen) gibt. Wir sehen, dass genau für $x = y = z$ die Zeilen linear abhängig sind, dies sind also die nicht regulären Punkte. Aber im Urbild von 0 gilt $x + y + z = 0$ (erste Nebenbedingung), was für $x = y = z$ einfach $x = y = z = 0$ impliziert. In diesem Fall ist die zweite Nebenbedingung $x^2 + y^2 + z^2 - 1 = 0$ aber verletzt. Deshalb enthält $g^{-1}(\{0\})$ nur reguläre Punkte, d.h. 0 ist ein regulärer Wert.

Damit ist die Methode von Lagrange anwendbar. Wir betrachten also die Hilfsfunktion

$$JL(x, y, z, \lambda_1, \lambda_2) = 5x + y - 3z + \lambda_1(x + y + z) + \lambda_2(x^2 + y^2 + z^2 - 1)$$

und suchen die Lösungen von

$$\begin{aligned} \frac{\partial JL}{\partial x}(x, y, z, \lambda_1, \lambda_2) &= 5 + \lambda_1 + 2\lambda_2 x &= 0, \\ \frac{\partial JL}{\partial y}(x, y, z, \lambda_1, \lambda_2) &= 1 + \lambda_1 + 2\lambda_2 y &= 0, \\ \frac{\partial JL}{\partial z}(x, y, z, \lambda_1, \lambda_2) &= -3 + \lambda_1 + 2\lambda_2 z &= 0, \\ \frac{\partial JL}{\partial \lambda_1}(x, y, z, \lambda_1, \lambda_2) &= x + y + z &= 0, \\ \frac{\partial JL}{\partial \lambda_2}(x, y, z, \lambda_1, \lambda_2) &= x^2 + y^2 + z^2 - 1 &= 0. \end{aligned}$$

Addition der ersten drei Gleichungen gibt in Verbindung mit der vierten Gleichung

$$3 + 3\lambda_1 = 0 \Rightarrow \lambda_1 = -1.$$

Eingesetzt in die ersten beiden Gleichungen gibt das

$$4 + 2\lambda_2 x = 0, \quad 2\lambda_2 y = 0.$$

Also haben wir $\lambda_2 \neq 0$, $x \neq 0$, und $y = 0$. Nun besagen die letzten beiden Gleichungen

$$x = -z, \quad 2x^2 = 1 \quad \Rightarrow \quad x = \pm \frac{1}{\sqrt{2}}.$$

Setzt man diese Werte in die Gleichungen ein, bestätigt sich, dass dies wirklich Lösungen sind. Die Punkte $(\frac{1}{\sqrt{2}}, 0, -\frac{1}{\sqrt{2}})$ und $(-\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}})$ stehen also im Verdacht, lokale Extrema von f unter der Nebenbedingung^a g zu sein.

An diesen Stellen sind die Funktionswerte

$$f\left(\frac{1}{\sqrt{2}}, 0, -\frac{1}{\sqrt{2}}\right) = 4\sqrt{2}, \quad f\left(-\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}\right) = -4\sqrt{2}.$$

Da $g^{-1}(\{0\})$ kompakt ist (diese Menge ist ja der Schnitt der kompakten Sphäre S^2 mit der durch $x+y+z=0$ beschriebenen abgeschlossenen Ebene), nimmt f ein Maximum und Minimum unter der Nebenbedingung g an. Also sind die oben bestimmten Werte die Extrema, d.h.

$$\max_{(x,y,z) \in g^{-1}\{0\}} f(x,y,z) = 4\sqrt{2}, \quad \min_{(x,y,z) \in g^{-1}\{0\}} f(x,y,z) = -4\sqrt{2}.$$

^aDie Sprechweise variiert hier etwas: Man spricht sowohl von einer Nebenbedingung M (mit einer Untermannigfaltigkeit $M \subset \mathbb{R}^d$) als auch von Nebenbedingung g (oder $g(x) = 0$) falls $M = g^{-1}(\{0\})$

Die Methode der Lagrange-Multiplikatoren hat viele wichtige Anwendungen, die Sie zB in Vorlesungen über *Optimierung* erlernen können. Zum Abschluss der Analysis-2-Vorlesung wollen wir aber nicht über Optimierung sprechen, sondern noch einmal die Brücke zur Schwestervorlesung über Lineare Algebra schlagen. Dazu geben wir einen Beweis des Ihnen bekannten Spektralsatzes für symmetrische Matrizen (“symmetrische Matrizen sind diagonalisierbar”) mit Methoden der Analysis.

Satz 13.37. Sei $A \in M_d(\mathbb{R})$ eine reelle symmetrische ($d \times d$)-Matrix.

a) Das Maximum

$$\lambda := \max_{x \in S^{d-1}} \langle x, Ax \rangle$$

existiert und ist ein Eigenwert von A . Jedes $x \in S^{d-1}$ mit $\langle x, Ax \rangle = \lambda$ ist ein zugehöriger Eigenvektor.

b) Es gibt eine aus Eigenvektoren von A bestehende Orthonormalbasis von \mathbb{R}^d .

Bemerkungen:

- Wir könnten das Maximum auch als $\max\{\langle x, Ax \rangle : x \in \mathbb{R}^d, \|x\|_2 = 1\}$ formulieren, heben hier aber den Zusammenhang mit der Untermannigfaltigkeit $S^{d-1} \subset \mathbb{R}^d$ hervor.

- Schon das Ergebnis von Teil a) dieses Lemmas ist bemerkenswert: Die Eigenwerte einer Matrix sind die Nullstellen des charakteristischen Polynoms, und über \mathbb{R} muss ein allgemeines Polynom p keine Nullstellen haben (zB $p(x) = x^2 + 7$).

Beweis. a) Wir betrachten die Funktion $f : \mathbb{R}^d \rightarrow \mathbb{R}$, $f(x) := \langle x, Ax \rangle$. Eingeschränkt auf die kompakte Untermannigfaltigkeit $S^{d-1} \subset \mathbb{R}^d$ hat f ein Maximum, d.h. das im Lemma definierte λ existiert in \mathbb{R} : Es gibt $x_0 \in \mathbb{R}^d$ mit $\lambda = \langle x_0, Ax_0 \rangle$. Nach dem Satz von Lagrange gilt für jede solche Maximalstelle x_0 die Gleichung $(\nabla f)(x_0) = \mu(\nabla g)(x_0)$, wobei $g(x) = \|x\|_2^2 - 1$ die Nebenbedingung beschreibt und der Lagrangemultiplikator hier μ statt λ genannt wird, um Verwechslungen mit dem Eigenwert λ auszuschließen. Wir haben $(\nabla g)(x_0) = 2x_0$ und nach Aufgabe H13.2 a) $(\nabla f)(x_0) = 2Ax_0$. Also existiert $\mu \in \mathbb{R}$ mit

$$\mu x_0 = Ax_0,$$

d.h. x_0 ist ein Eigenvektor von A mit Eigenwert μ (Lagrangemultiplikator). Nun zeigen wir noch $\lambda = \mu$:

$$\lambda = \max_{x \in S^{d-1}} f(x) = f(x_0) = \langle x_0, Ax_0 \rangle = \mu \langle x_0, x_0 \rangle = \mu.$$

b) Diesen Teil beweisen wir einfach wie in Linearer Algebra: Wir schränken die lineare Abbildung A auf das orthogonale Komplement E_λ^\perp des Eigenraums $E_\lambda = \{x \in \mathbb{R}^d : Ax = \lambda x\}$ ein. Das ist möglich, da für $v \in E_\lambda^\perp$ auch $Av \in E_\lambda^\perp$ gilt: Für jedes x mit $Ax = \lambda x$ haben wir ja unter Verwendung der Symmetrie von A

$$\langle Av, x \rangle = \langle v, Ax \rangle = \lambda \langle v, x \rangle = 0.$$

Auf dem orthogonalen Komplement E_λ^\perp können wir nun erneut das Argument aus a) verwenden. Wir finden also einen zu E_λ senkrecht stehenden Eigenraum E_μ . Da $\dim E_\lambda^\perp < \dim \mathbb{R}^d = d$ und d endlich ist, kommen wir nach endlich vielen Schritten bei einem leeren orthogonalen Komplement an. Nun müssen wir nur noch Orthonormalbasen der Eigenräume wählen (beliebig) und sind fertig. \square

Diese Bemerkungen eignen sich für einen kleinen Ausblick auf weitere Vorlesungen: In obigem Argument war es wichtig, dass die Einheitssphäre in \mathbb{R}^d kompakt ist (Teil a)), und dass die Induktion terminiert (Teil b)). Diese beiden Tatsachen liegen in der endlichen Dimension d begründet und stimmen für unendlichdimensionale Vektorräume nicht mehr. Schauen Sie sich Beispiel 9.2 an: Die Einheitssphäre in $C_b(\mathbb{R})$ ist nicht kompakt.

Wenn wir also über Eigenwerte von linearen Abbildungen auf unendlichdimensionalen Vektorräumen sprechen möchten (und es gibt sehr viele gute Gründe, das zu tun (um nur zwei Gründe zu nennen: Lösungen von Differentialgleichungen, Anwendungen in der Quantenphysik), so werden wir uns auf einige Änderungen einstellen müssen. Diese Themen werden in Vorlesungen über *Funktionalanalysis* und *Spektraltheorie* besprochen und stützen sich sehr auf Maßtheorie, was ein wesentlicher Teil von *Analysis 3* ist.

Literatur

- [For16] O. Forster. *Analysis 1*. 12. Aufl. Springer, 2016. URL: <https://link.springer.com/book/10.1007/978-3-658-11545-6>.
- [Heu91] H. Heuser. *Lehrbuch der Analysis II*. Teubner, 1991.
- [Lec22] G. Lechner. *Analysis 1*. 2022.
- [Zor06] V. Zorich. *Analysis 1*. Springer, 2006. URL: <https://link.springer.com/book/10.1007/3-540-33278-2>.

Die Lehrbücher [Zor06; For16; Heu91] sind als Begleitlektüre für die ganze Vorlesung empfehlenswert. Auf die anderen Bücher wird nur an einigen Stellen im Skript verwiesen.