

Bitte beachten (wurde in meinen damaligen VO klar kommuniziert): Der gesamte Analysis-Zyklus ist stark an die (2005-2008 verfügbaren Ausgaben der) Bücher von Forster angelehnt, mit Zusätzen aus Heuser und Rudin im dritten Semester.

Analysis II

Günther Hörmann & David Langer

Fakultät für Mathematik der Universität Wien

Wintersemester 2007

Inhalt

V. FUNKTIONENFOLGEN UND -REIHEN	1
12. Gleichmäßige und punktweise Konvergenz	1
13. Potenzreihen	9
14. Fourier-Reihen	16
VI. TOPOLOGISCHE GRUNDBEGRIFFE	31
15. Metrische und normierte Räume	31
16. Konvergenz und Stetigkeit	43
17. Kompaktheit	55
VII. DIFFERENZIERBARE ABBILDUNGEN $\mathbb{R}^n \rightarrow \mathbb{R}^m$	61
18. Partielle Ableitungen und Differenzierbarkeit	63
19. Taylor-Formel und lokale Extrema	85
20. Implizite Funktionen und Umkehrsatz	99
VIII. KURVEN UND FLÄCHEN — UNTERMANNIGFALTIGKEITEN DES \mathbb{R}^n	113
21. Wege und Kurven	113
22. Vektorfelder, 1-Formen und Kurvenintegrale	120
23. Untermannigfaltigkeiten des \mathbb{R}^n	137

V FUNKTIONENFOLGEN UND -REIHEN

§12. Gleichmäßige und punktweise Konvergenz

12.1. Definition

Es sei K eine Menge und für $n \in \mathbb{N}$ sei eine Funktion $f_n: K \rightarrow \mathbb{C}$ gegeben. Dann erhalten wir für jedes $x \in K$ durch $(f_n(x))_{n \in \mathbb{N}}$ eine Folge komplexer Zahlen.

- 1.) Die Funktionenfolge (f_n) *konvergiert punktweise* gegen die Funktion $f: K \rightarrow \mathbb{C}$, falls für jedes $x \in K$ die Folge $(f_n(x))_{n \in \mathbb{N}}$ gegen $f(x)$ konvergiert:

$$\forall x \in K : f_n(x) \rightarrow f(x) \quad (n \rightarrow \infty),$$

d.h. im Detail

$$(12.1) \quad \forall x \in K \forall \varepsilon > 0 \exists N = N(x, \varepsilon) \in \mathbb{N} \forall n \geq N : |f_n(x) - f(x)| < \varepsilon.$$

(Beachte: N hängt hier von ε und von x ab!)

- 2.) Die Funktionenfolge (f_n) *konvergiert gleichmäßig* gegen die Funktion $f: K \rightarrow \mathbb{C}$, falls gilt:

$$(12.2) \quad \forall \varepsilon > 0 \exists N = N(\varepsilon) \in \mathbb{N} \forall n \geq N : \underbrace{\forall x \in K \quad |f_n(x) - f(x)| < \varepsilon}_{\sup_{x \in K} |f_n(x) - f(x)| \leq \varepsilon}.$$

(Beachte: N hängt hier nicht von x ab!)

- 3.) Es sei $g: K \rightarrow \mathbb{C}$, dann bezeichnen wir

$$(12.3) \quad \|g\|_\infty := \sup_{x \in K} |g(x)|$$

als *Supremumsnorm* oder *Unendlichnorm* von g . Die Funktion g ist beschränkt genau dann, wenn $\|g\|_\infty < \infty$ gilt.

Bemerkungen:

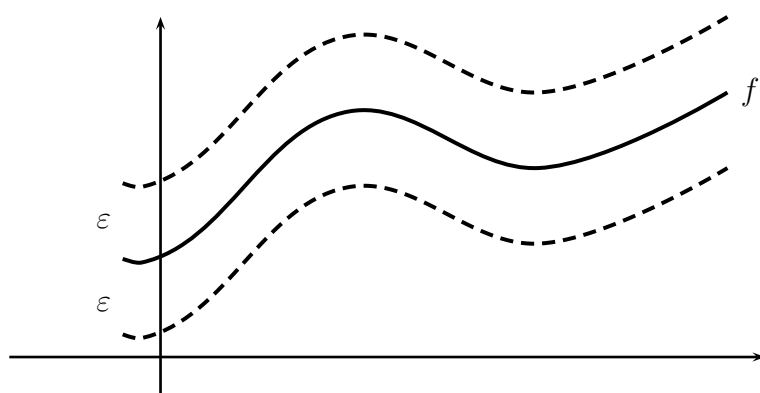
(i) Die Bedingung (12.2) ist stärker als (12.1), daher folgt aus der gleichmäßigen Konvergenz einer Funktionenfolge stets die punktweise Konvergenz.

(ii) Die gleichmäßige Konvergenz $f_n \rightarrow f$ ist gleichbedeutend mit folgender Bedingung

$$(12.2') \quad \forall \varepsilon > 0 \exists N = N(\varepsilon) \in \mathbb{N} \forall n \geq N : \|f_n - f\|_\infty \leq \varepsilon,$$

das heißt $\|f_n - f\|_\infty \rightarrow 0$ für $n \rightarrow \infty$ (insbesondere $\|f_n - f\|_\infty < \infty$ für fast alle n).

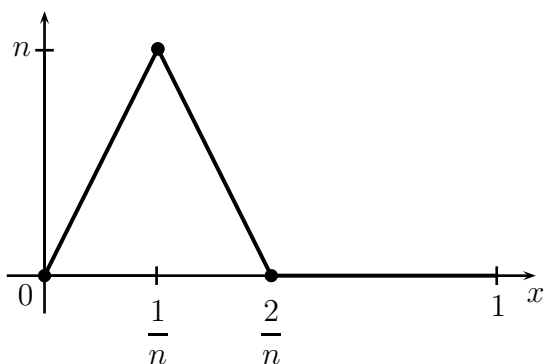
(iii) Die Bedingung (12.2) können wir mit Hilfe von ε -Umgebungen auch so umformulieren: zu jedem $\varepsilon > 0$ existiert ein $N \in \mathbb{N}$, so dass $f_n(x) \in U_\varepsilon(f(x))$ für alle $n \geq N$ und für alle $x \in K$ gilt (äquivalent: $f_n(x) - f(x) \in U_\varepsilon(0)$).



Anschaulich gesprochen heißt letzteres, dass die Graphen der Funktionen f_n schließlich innerhalb eines ε -Schlauches um den Graphen von f bleiben.

12.2. Beispiele

1) Es sei für $n \in \mathbb{N}$ die stückweise (affin) lineare Funktion $f_n: [0, 1] \rightarrow \mathbb{R}$, gegeben durch folgenden Graphen:



[Wir benötigen für die weiteren Überlegungen gar keine expliziten Formeln für $f_n(x)$. Die Skizze macht schneller klar, was passiert, wenn n groß wird: die Dreiecke werden höher und schmaler, wobei die Spitzen und rechten Eckpunkte nach links wandern. Vergleichen Sie mit der Angabe $f_n(x) = n^2x$, wenn $0 \leq x \leq 1/n$, $f_n(x) = -n^2x + 2n$, wenn $1/n < x < 2/n$, und $f_n(x) = 0$, wenn $2/n \leq x \leq 1$.]

Behauptung 1: $f_n \rightarrow 0$ ($n \rightarrow \infty$) punktweise.

Beweis: Zunächst gilt $f_n(0) = 0$ für alle n . Zu $x \neq 0$ gibt es ein $N \geq 2$ mit $2/N \leq x$ (N hängt also von x ab!). Somit folgt für jedes $n \geq N$, dass $2/n \leq x$ und daher $f_n(x) = 0$. Somit gilt für jedes $x \in [0, 1]$, dass $\lim_{n \rightarrow \infty} f_n(x) = 0$ ist.

Behauptung 2: (f_n) ist nicht gleichmäßig konvergent.

Beweis: (Indirekt) Falls (f_n) gleichmäßig konvergiert, so muss der gleichmäßige Limes gleich dem punktweisen Limes sein (klar warum?). Ist f_n gleichmäßig konvergent gegen 0, so folgt aber

$$\|f_n - 0\|_\infty = f_n(1/n) = n \not\rightarrow 0 \quad (n \rightarrow \infty).$$

(Die Folge der Supremumsnormen der f_n ist hier sogar unbeschränkt. Auch mit Spitzen der fixen Höhe 1 ergäbe sich noch keine gleichmäßige Konvergenz.)

2) Sei $m > 0$ beliebig und für $n \in \mathbb{N}$ betrachte $f_n: [-m, m] \rightarrow \mathbb{R}$, gegeben durch

$$f_n(x) := \sum_{k=0}^n \frac{x^k}{k!}.$$

Wir wissen bereits aus Analysis 1, dass f_n punktweise gegen die Exponentialfunktion strebt ($(f_n(x))_{n \in \mathbb{N}}$ ist ja gerade die Partialsummenfolge für $\exp(x)$). Es ist

$$\exp(x) = f_n(x) + R_{n+1}(x),$$

wobei wir für $|x| \leq 1 + n/2$ die Restgliedabschätzung

$$|R_{n+1}(x)| \leq 2 \frac{|x|^{n+1}}{(n+1)!}$$

(in Analysis 1) gezeigt haben. Daher gilt für $n \geq 2(m-1)$

$$\|\exp - f_n\|_\infty = \sup_{|x| \leq m} |\exp(x) - f_n(x)| \leq \frac{2m^{n+1}}{(n+1)!} \rightarrow 0 \quad (n \rightarrow \infty).$$

Mit anderen Worten: (f_n) konvergiert auf dem Intervall $[-m, m]$ gleichmäßig gegen \exp .

12.3. Theorem

Es sei $K \subseteq \mathbb{C}$ und für jedes $n \in \mathbb{N}$ eine stetige Funktion $f_n: K \rightarrow \mathbb{C}$ gegeben. Falls (f_n) gleichmäßig gegen die Funktion $f: K \rightarrow \mathbb{C}$ konvergiert, dann ist auch f stetig.

Beweis:

Es sei $x \in K$ beliebig. Wir müssen zeigen, dass f stetig in x ist.

Sei $\varepsilon > 0$ vorgelegt. Dann gibt es wegen der gleichmäßigen Konvergenz von (f_n) gegen f ein $N \in \mathbb{N}$, so dass gilt

$$\forall \xi \in K : |f_N(\xi) - f(\xi)| < \varepsilon/3.$$

Wegen der Stetigkeit von f_N gibt es ein $\delta > 0$, so dass

$$\forall x' \in K, |x - x'| < \delta : |f_N(x) - f_N(x')| < \varepsilon/3.$$

Somit gilt für alle $x' \in K$ mit $|x - x'| < \delta$ (mittels Einschließen von geeigneten Differenzen)

$$|f(x) - f(x')| \leq |f(x) - f_N(x)| + |f_N(x) - f_N(x')| + |f_N(x') - f(x')| < \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon.$$

□

12.4. Bemerkungen

(i) Sind alle f_n stetig und konvergieren aber nur punktweise gegen die Funktion f , so muss f im Allgemeinen nicht stetig sein. (Ein konkretes Beispiel dafür ist in den Übungsaufgaben versteckt.)

(ii) Das obige Theorem kann natürlich auch für folgende Schlussweise verwendet werden: falls die stetigen Funktionen f_n punktweise gegen eine Funktion f streben und f aber unstetig ist, so kann die Konvergenz nicht gleichmäßig sein.

Als konkrete Anwendung erinnern wir an die (punktweise gültige) Gleichung

$$\sum_{k=1}^{\infty} \frac{\sin(kx)}{k} = \frac{\pi - x}{2} \quad (x \in]0, 2\pi[),$$

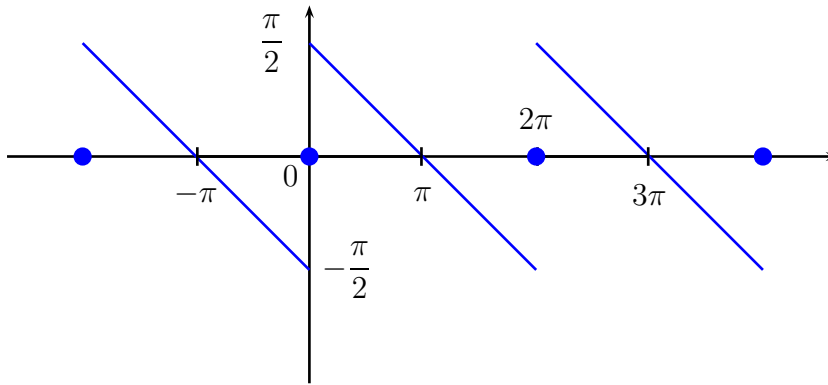
die wir in Analysis 1, Kapitel IV, bewiesen hatten.

Es sei $f: \mathbb{R} \rightarrow \mathbb{R}$ definiert durch $f(0) = 0$, $f(x) = (\pi - x)/2$ für $0 < x < 2\pi$ (die rechte Seite obiger Gleichung) und 2π -periodische Fortsetzung auf ganz \mathbb{R} . Weiters setzen wir $f_n(x) = \sum_{k=1}^n \sin(kx)/k$ ($x \in \mathbb{R}$, $n \in \mathbb{N}$, $n \geq 1$; entspricht also den Partialsummen der linken Seite in obiger Gleichung). Dann bedeutet obige Summenformel gerade, dass

$$\forall x \in \mathbb{R} : f_n(x) \rightarrow f(x) \quad (n \rightarrow \infty),$$

also die punktweise Konvergenz von f_n gegen f .

Die Grenzfunktion f hat aber Sprungstellen, ist also unstetig:



Daher kann f_n (auf \mathbb{R}) nicht gleichmäßig gegen f konvergieren. (Auf keinem Intervall, das einen Punkt $2l\pi$, $l \in \mathbb{Z}$, enthält, kann gleichmäßige Konvergenz stattfinden.)

12.5. Theorem (Satz von Weierstraß)

Es sei K eine Menge und $f_n : K \rightarrow \mathbb{C}$ ($n \in \mathbb{N}$). Falls die Reihe $\sum_{k=0}^{\infty} \|f_k\|_{\infty}$ konvergiert, so ist

die Funktionenreihe $\sum_{k=0}^{\infty} f_k$ absolut konvergent (d.h. $\sum |f_k(x)|$ konvergiert für jedes $x \in K$)

und gleichmäßig konvergent (d.h. $\sum_{k=0}^n f_k$ konvergiert gleichmäßig für $n \rightarrow \infty$).

Beweis: Für $x \in K$ und $k \in \mathbb{N}$ gilt $|f_k(x)| \leq \|f_k\|_{\infty}$, daher ist $\sum \|f_k\|_{\infty}$ eine konvergente Majorante für $\sum |f_k(x)|$. Somit ist für jedes $x \in K$ die Reihe $\sum f_k(x)$ absolut konvergent.

Setze $F(x) := \sum_{k=0}^{\infty} f_k(x)$ für $x \in K$ und $F_n := \sum_{k=0}^n f_k$ für $n \in \mathbb{N}$. Wir zeigen, dass $F_n \rightarrow F$ gleichmäßig (auf K) für $n \rightarrow \infty$.

Sei $\varepsilon > 0$. Wegen der Konvergenz von $\sum \|f_k\|_{\infty}$ gibt es ein $N \in \mathbb{N}$, so dass $\sum_{k=n+1}^{\infty} \|f_k\|_{\infty} < \varepsilon$ für alle $n \geq N$.

Es sei $x \in K$ und $n \geq N$, dann gilt

$$|F(x) - F_n(x)| = \left| \sum_{k=n+1}^{\infty} f_k(x) \right| \leq \sum_{k=n+1}^{\infty} |f_k(x)| \leq \sum_{k=n+1}^{\infty} \|f_k\|_{\infty} < \varepsilon.$$

Wenn wir zum Supremum über $x \in K$ übergehen, erhalten wir daraus

$$\|F - F_n\|_{\infty} \leq \varepsilon \quad \forall n \geq N$$

und somit die Behauptung. □

Eine kleine Warnung: In der Situation von Theorem 12.5 erhalten wir sogar, dass auch die Reihe der Absolutbeträge $\sum |f_k|$ gleichmäßig konvergiert (wende den Satz auf $|f_k|$ statt f_k an). Allerdings folgt im Allgemeinen aus der gleichmäßigen und punktweise absoluten Konvergenz einer Funktionenreihe nicht, dass die Reihe der Absolutbeträge gleichmäßig konvergiert, wie man an Hand des Beispiels (aus [BF00, p.316]) mit $f_k(x) = (-1)^k x^2 / (1+x^2)^k$ ($x \in \mathbb{R}$) studieren kann.

12.6. Beispiele

1) Es sei $m > 0$ und $f_k: [-m, m] \rightarrow \mathbb{R}$ gegeben durch $f_k(x) = \frac{x^k}{k!}$ ($k \in \mathbb{N}$). Dann ist $\|f_k\|_\infty = m^k/k!$ und wegen

$$\sum_{k=0}^{\infty} \|f_k\|_\infty = \sum_{k=0}^{\infty} \frac{m^k}{k!} = e^m < \infty$$

erhalten wir einen weiteren Beweis dafür, dass $\sum \frac{x^k}{k!}$ auf jedem Intervall $[-m, m]$ gleichmäßig konvergent (gegen $\exp(x)$) ist.

2) Es sei $f_k(x) = \frac{\cos(kx)}{k^2}$ ($k \in \mathbb{N}$, $k \geq 1$, $x \in \mathbb{R}$). Wegen $\|f_k\|_\infty = 1/k^2$ ist

$$\sum_{k=1}^{\infty} \|f_k\|_\infty = \sum_{k=1}^{\infty} \frac{1}{k^2} < \infty$$

und daher $\sum_{k=1}^{\infty} \frac{\cos(kx)}{k^2}$ gleichmäßig konvergent (auf \mathbb{R}).

12.7. Proposition (Vertauschung von Limes und Integral)

Es sei $f_n: [a, b] \rightarrow \mathbb{R}$ stetig ($n \in \mathbb{N}$) und (f_n) gleichmäßig konvergent. Dann gilt

$$\int_a^b \left(\lim_{n \rightarrow \infty} f_n(x) \right) dx = \lim_{n \rightarrow \infty} \int_a^b f_n(x) dx.$$

Beweis:

Wir setzen $f(x) := \lim_{n \rightarrow \infty} f_n(x)$ für $x \in [a, b]$. Nach Theorem 12.3 ist $f: [a, b] \rightarrow \mathbb{R}$ stetig, also integrierbar und es gilt

$$\left| \int_a^b f(x) dx - \int_a^b f_n(x) dx \right| \leq \int_a^b |f(x) - f_n(x)| dx \leq (b-a) \|f - f_n\|_\infty \rightarrow 0 \quad (n \rightarrow \infty).$$

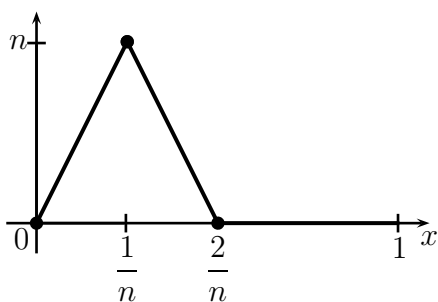
□

12.8. Beispiele

1) Es sei $f_n(x) = \sum_{k=1}^n \frac{\cos(kx)}{k^2}$ ($n \in \mathbb{N}$, $n \geq 1$, $x \in [0, 2\pi]$). Gemäß Beispiel 12.6.2) ist (f_n) gleichmäßig konvergent, daher besagt Proposition 12.7 für jedes $t \in [0, 2\pi]$

$$\int_0^t \sum_{k=1}^{\infty} \frac{\cos(kx)}{k^2} dx = \sum_{k=1}^{\infty} \frac{1}{k^2} \int_0^t \cos(kx) dx = \sum_{k=1}^{\infty} \frac{1}{k^2} \frac{\sin(kx)}{k} \Big|_0^t = \sum_{k=1}^{\infty} \frac{\sin(kt)}{k^3}.$$

2) Punktweise Konvergenz reicht im Allgemeinen nicht aus, um Limes und Integral vertauschen zu dürfen:



für die (Dreiecks-)Funktionen f_n aus Beispiel 12.2.1) ergibt sich $\int_0^1 f_n = \frac{2}{n} \cdot n \cdot \frac{1}{2} = 1$ für alle n ,

aber

$$\int_0^1 \lim f_n(x) dx = \int_0^1 0 dx = 0.$$

12.9. Proposition (Vertauschung von Limes und Differentiation)

Es sei $f_n: [a, b] \rightarrow \mathbb{R}$ stetig differenzierbar ($n \in \mathbb{N}$). Die Folge (f_n) sei punktweise konvergent gegen $f: [a, b] \rightarrow \mathbb{R}$ und die Folge der Ableitungen (f'_n) sei gleichmäßig konvergent. Dann ist f differenzierbar und es gilt

$$\forall x \in [a, b]: f'(x) = \lim_{n \rightarrow \infty} f'_n(x).$$

Insbesondere ist f' stetig.

Beweis:

Wir setzen $g := \lim_{n \rightarrow \infty} f'_n$. Nach Theorem 12.3 ist $g: [a, b] \rightarrow \mathbb{R}$ stetig.

Sei $x \in [a, b]$, dann wird aus der Gleichung $f_n(x) = f_n(a) + \int_a^x f'_n(t) dt$

für $n \rightarrow \infty$ gemäß Proposition 12.7 die Gleichung $f(x) = f(a) + \int_a^x g(t) dt$.

Somit muss $f'(x) = g(x)$ gelten. □

12.10. Bemerkung

Die gleichmäßige Konvergenz der Folge (f_n) allein genügt im Allgemeinen nicht, um Limes und Differentiation vertauschen zu können, d.h. selbst wenn f_n stetig differenzierbar für alle $n \in \mathbb{N}$ ist, $f_n \rightarrow f$ gleichmäßig und f differenzierbar, dann folgt i.a. **nicht** (einmal), dass $f'_n \rightarrow f'$ (punktweise).

Als Beispiel betrachte $f_n(x) = \sin(nx)/n$ ($n \in \mathbb{N}$, $n \geq 1$, $x \in \mathbb{R}$). Wegen $\|f_n\|_\infty = 1/n \rightarrow 0$ gilt $f_n \rightarrow 0$ gleichmäßig ($n \rightarrow \infty$). Aber $f'_n(x) = \cos(nx)$ ist nicht einmal punktweise konvergent (z.B. im Punkt $x = \pi$).

Die Situation wird bedeutend übersichtlicher, wenn wir spezielle Typen von Funktionenfolgen und -reihen betrachten, wie zum Beispiel Potenzreihen im folgenden Abschnitt.

§13. Potenzreihen

Wir hatten in Kapitel IV den Begriff der Taylor-Reihe für eine unendlich oft differenzierbare (reelle) Funktion f auf einem Intervall $I \subseteq \mathbb{R}$ eingeführt und gesehen, dass wir damit in vielen Fällen polynomiale Approximationen an die Funktion gewinnen können. Für einen Entwicklungspunkt $x_0 \in I$ definiert man dazu die Folge

$$c_k = \frac{f^{(k)}(x_0)}{k!} \quad (k \in \mathbb{N})$$

und erhält (für jedes $x \in I$ und $n \in \mathbb{N}$)

$$f(x) = \sum_{k=0}^n c_k (x - x_0)^k + R_{n+1}(x),$$

wobei wir für das Restglied verschiedene konkrete Darstellungen gegeben hatten. Die Taylor-Reihe ist dann gegeben durch

$$\sum_{k=0}^{\infty} c_k (x - x_0)^k$$

und stellt die Funktion f dar, falls das Restglied mit $n \rightarrow \infty$ gegen Null strebt. Wir nehmen nun einen anderen Standpunkt ein, indem wir allgemein Reihen obigen Typs (zunächst) losgelöst von im Vorhinein gegebenen Funktionen betrachten, d.h. für (c_k) irgendeine Zahlenfolge zulassen.

13.1. Definition

Es sei (c_k) eine Folge komplexer Zahlen und $x_0 \in \mathbb{C}$, dann nennen wir $\sum_{k=0}^{\infty} c_k (x - x_0)^k$ ($x \in \mathbb{C}$) eine *Potenzreihe* mit Entwicklungspunkt x_0 .

Wir untersuchen in diesem Abschnitt die Konvergenzeigenschaften von Potenzreihen in Abhängigkeit von x , d.h. wir studieren Folgen von Polynomfunktionen (p_n) , wobei $p_n(x) = \sum_{k=0}^n c_k (x - x_0)^k$.

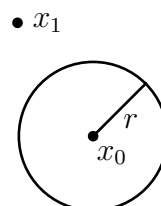
13.2. Proposition

Es sei $\sum c_k (x - x_0)^k$ eine Potenzreihe und $x_1 \in \mathbb{C}$ habe die Eigenschaft, dass $\sum c_k (x_1 - x_0)^k$ konvergent ist.

Für $r \in \mathbb{R}$ mit $0 < r < |x_1 - x_0|$ bezeichne

$$K(x_0, r) := \{x \in \mathbb{C} : |x - x_0| \leq r\}$$

die (abgeschlossene) Kreisscheibe mit Radius r um x_0 .



Dann gilt:

1.) $\sum_{k=0}^{\infty} c_k (x - x_0)^k$ konvergiert absolut und gleichmäßig auf $K(x_0, r)$.

2.) Die (formal¹) gliedweise differenzierte Reihe $\sum_{k=0}^{\infty} c_k k (x - x_0)^{k-1}$ konvergiert ebenfalls absolut und gleichmäßig auf $K(x_0, r)$.

Beweis: 1.) Wir setzen $f_n(x) = c_n (x - x_0)^n$ ($n \in \mathbb{N}$, $x \in K(x_0, r)$). Es ist für jedes $x \in K(x_0, r)$

$$|f_n(x)| = |c_n| |x - x_0|^n = |c_n| |x_1 - x_0|^n \left(\frac{|x - x_0|}{|x_1 - x_0|} \right)^n,$$

wobei $\frac{|x - x_0|}{|x_1 - x_0|} \leq \frac{r}{|x_1 - x_0|} =: \theta < 1$. Weiters müssen wegen der Konvergenz der Reihe $\sum c_n (x_1 - x_0)^n$ die Glieder eine Nullfolge bilden, insbesondere gibt es also ein $M > 0$ mit $|c_n| |x_1 - x_0|^n \leq M$ für alle $n \in \mathbb{N}$. Daher liefert die obige Darstellung für $f_n(x)$ die Abschätzung

$$|f_n(x)| \leq M \cdot \theta^n \quad \forall x \in K(x_0, r) \quad \forall n \in \mathbb{N},$$

und somit

$$\|f_n\|_{\infty} = \sup_{x \in K(x_0, r)} |f_n(x)| \leq M \theta^n.$$

Nach dem Majorantenkriterium ist $\sum \|f_n\|_{\infty}$ konvergent, daher nach dem Satz von Weierstraß 12.5 also die Reihe $\sum f_n$ absolut und gleichmäßig konvergent auf $K(x_0, r)$.

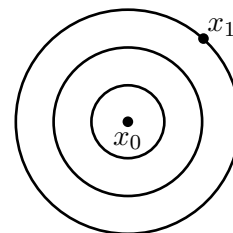
2.) Wir setzen $g_n(x) = c_n n (x - x_0)^{n-1}$ ($= f'_n(x)$) und erhalten wie im Beweis von 1.), dass

$$\|g_n\|_{\infty} \leq n M \theta^{n-1}.$$

Aus dem Quotiententest folgt, dass $\sum n M \theta^{n-1}$ konvergiert. Daher ist wiederum nach dem Satz von Weierstraß die Reihe $\sum g_n$ absolut und gleichmäßig konvergent auf $K(x_0, r)$. \square

13.3. Bemerkungen

(i) Satz 13.2 besagt insbesondere, dass zwischen Punkten $x \in \mathbb{C}$ mit $|x - x_0| = r$ und $x = x_0$ keine Lücken bezüglich Konvergenz auftreten: falls die Potenzreihe für $x = x_1$ konvergent ist, so ist sie in jeder (abgeschlossenen) Kreisscheibe $K(x_0, r)$ mit $r < |x_1 - x_0|$ gleichmäßig konvergent.



In allen Punkten der (offenen) Kreisscheibe $\{x \in \mathbb{C} : |x - x_0| < |x_1 - x_0|\}$ ist die Potenzreihe somit jedenfalls punktweise konvergent. (Für das Konvergenzverhalten in den Randpunkten kann keine allgemeine Aussage gemacht werden.)

¹Für reelle x ist dies natürlich die übliche Ableitung im Sinne der Analysis 1.

(ii) Wir erhalten also durch $f(x) := \sum_{k=0}^{\infty} c_k(x - x_0)^k$ eine Funktion $f: \{x \in \mathbb{C} : |x - x_0| < |x_1 - x_0|\} \rightarrow \mathbb{C}$, die sogar stetig ist: eine genügend kleine Umgebung eines beliebig gegebenen Punktes $x \in \mathbb{C}$ mit $|x - x_0| < |x_1 - x_0|$ ist in einem passenden $K(x_0, r)$ mit $|x - x_0| < r < |x_1 - x_0|$ enthalten, und darauf findet die Konvergenz der stetigen Partialsummen (das sind ja Polynomfunktionen) gleichmäßig statt. Umgekehrt kann man die Frage stellen, welche Funktionen denn durch Potenzreihen gegeben sind. Dies führt auf den Begriff der *analytischen Funktion*, der im Rahmen der Komplexen Analysis (oder ‘Funktionentheorie’) eingehend studiert wird.

Aus dem obigen stellt sich die Frage, ob wir für jede Potenzreihe auf praktikable Weise einen maximalen Konvergenzkreis bestimmen können. Dies führt unmittelbar auf den nächsten Begriff.

13.4. Definition

Es sei $\sum c_n(x - x_0)^n$ eine Potenzreihe, dann heißt

$$(13.1) \quad R := \sup\{r \in [0, \infty[: \sum_{n=0}^{\infty} c_n(x - x_0)^n \text{ konvergiert in } K(x_0, r)\}$$

Konvergenzradius der Potenzreihe. Es ist $0 \leq R \leq \infty$.

13.5. Proposition

Es sei R der Konvergenzradius der Potenzreihe $\sum c_n(x - x_0)^n$. Dann gilt:

- 1.) Im Falle $R = 0$ konvergiert die Reihe nur im Punkt $x = x_0$.
- 2.) Im Falle $R = \infty$ konvergiert die Reihe für alle $x \in \mathbb{C}$ und die Konvergenz ist gleichmäßig auf jeder (abgeschlossenen) Kreisscheibe $K(x_0, r)$ mit $0 \leq r < \infty$.
- 3.) Im Falle $0 < R < \infty$ konvergiert die Reihe für alle $x \in \mathbb{C}$ mit $|x - x_0| < R$ und die Konvergenz ist gleichmäßig auf jeder (abgeschlossenen) Kreisscheibe $K(x_0, r)$ mit $0 \leq r < R$.

Die Reihe ist divergent für alle $x \in \mathbb{C}$ mit $|x - x_0| > R$. (Für die Randpunkte $x \in \mathbb{C}$ mit $|x - x_0| = R$ kann sowohl Divergenz als auch Konvergenz vorliegen [und zwar auch beide Fälle in verschiedenen Randpunkten derselben Reihe].)

- 4.) Es gilt die *Formel von Hadamard*

$$(13.2) \quad \frac{1}{R} = \limsup_{n \rightarrow \infty} |c_n|^{1/n},$$

wobei wir $R = 0$ setzen, falls der Limes superior unendlich ist, und $R = \infty$, falls der Limes superior Null ist.

Beweis:

1.) und 2.) folgen unmittelbar aus der Definition 13.4 und Bemerkung 13.3.

3.) Es sei $x \in \mathbb{C}$ mit $|x - x_0| < R$. Dann gibt es ein $x_1 \in \mathbb{C}$ mit $|x_1 - x_0| > |x - x_0|$, sodass die Reihe $\sum c_n(x_1 - x_0)^n$ konvergiert. Daher konvergiert die Potenzreihe gleichmäßig auf jeder Kreisscheibe $K(x_0, r)$ mit $0 \leq r < |x_1 - x_0|$, insbesondere auch im Punkt x . Weil R als Supremum definiert ist, kann der Punkt x_1 so gewählt werden, dass $|x_1 - x_0|$ beliebig nahe (unterhalb) R ist. Damit folgt der erste Teil der Aussage.

Sei nun $x \in \mathbb{C}$ mit $|x - x_0| > R$. Wäre $\sum c_n(x - x_0)^n$ konvergent, so stünde das im Widerspruch zur Definition von R , weil Proposition 13.2 dann Konvergenz auf einer Kreisscheibe $K(x_0, r)$ mit $R < r < |x - x_0|$ impliziert.

4.) als Übungsaufgabe. (Hinweis: Wurzeltest.) □

13.6. Bemerkung

Die Berechnung des Konvergenzradius ist oft auch mit Hilfe des Quotiententests möglich (bzw. einfacher). Falls $|\frac{c_n}{c_{n+1}}|$ konvergiert, dann gilt

$$(13.3) \quad R = \lim_{n \rightarrow \infty} \left| \frac{c_n}{c_{n+1}} \right|.$$

(Sei nämlich ρ der angegebene Limes, dann haben wir im Quotiententest $|\frac{c_{n+1}(x-x_0)^{n+1}}{c_n(x-x_0)^n}| = |x - x_0| |\frac{c_{n+1}}{c_n}| \rightarrow \frac{|x-x_0|}{\rho}$ für $n \rightarrow \infty$, also Konvergenz für $|x - x_0| < \rho$).

13.7. Beispiele

1) Wir betrachten die Potenzreihe $\sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n} x^n$. [Taylor-Reihe für $\log(1+x)$, Analysis 1]

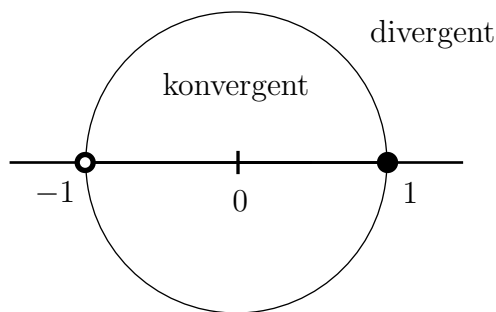
In diesem Beispiel ist also $x_0 = 0$ und $c_0 = 0$, $c_n = (-1)^{n-1}/n$ (für $n \geq 1$).

Für $x = 1$ erhalten wir die Reihe $\sum \frac{(-1)^{n-1}}{n}$, die nach dem Leibniz-Kriterium konvergent ist. Daher konvergiert die Potenzreihe sicher für alle $x \in \mathbb{C}$ mit $|x| < 1$ und gleichmäßig auf jedem Kreis $K(0, r)$ mit $0 \leq r < 1$.

Für $x = -1$ erhalten wir die divergente Reihe $\sum \frac{(-1)^{n-1}}{n} (-1)^n = -\sum \frac{1}{n}$, daher hat der Konvergenzradius den Wert $R = 1$. Insbesondere folgt Divergenz für alle $x \in \mathbb{C}$ mit $|x| > 1$.

Übrigens folgt das Resultat für den Konvergenzradius hier ebenso einfach aus der Formel von Hadamard (weil $(1/n)^{1/n} \rightarrow 1$) oder dem Quotiententest (weil $(n+1)/n \rightarrow 1$).

Zusammenfassend erhalten wir bisher folgende Übersicht: Konvergenz innerhalb des Einheitskreises, Divergenz außerhalb; im Randpunkt $x = -1$ Divergenz und im Randpunkt $x = 1$ Konvergenz. (Man kann zeigen, dass diese Reihe für jeden Randpunkt $x \neq -1$ konvergiert [vgl. etwa [RS02, 4.2, Aufgabe 2] oder [SS03, Chapter 1, Exercises 14 and 19]])



2) $\sum_{n=0}^{\infty} n! \cdot x^n$ ($x_0 = 0$, $c_n = n!$) hat Konvergenzradius

$$R = \lim \left| \frac{c_n}{c_{n+1}} \right| = \lim \frac{1}{n+1} = 0.$$

Daher ist diese Potenzreihe für jedes $x \neq 0$ divergent.

Übrigens folgt hier aus der Tatsache $R = 0$ mit Hilfe der Hadamard-Formel, dass

$$\lim_{n \rightarrow \infty} (n!)^{\frac{1}{n}} = \infty.$$

3) $\sum_{k=0}^{\infty} \frac{(-1)^k}{(2k)!} x^{2k}$ (also die Cosinus-Reihe).

Wir haben $x_0 = 0$ und $c_n = 0$ für ungerades n , $c_n = (-1)^{n/2}/n!$ für gerades n .

Wir wissen bereits aus Analysis 1, dass diese Reihe für jedes $x \in \mathbb{C}$ konvergiert, daher ist $R = \infty$. Ebenso erhalten wir Konvergenzradius $R = \infty$ für die Sinus- und die Exponentialreihe.

Wenn alle Koeffizienten c_n sowie x_0 reell sind, so definiert die Potenzreihe eine reelle Funktion auf dem Durchschnitt ihres Konvergenzkreises mit \mathbb{R} . Es stellt sich heraus, dass wir hierdurch unendlich oft differenzierbare Funktionen erhalten, deren Taylor-Reihen dann gerade durch die Potenzreihen gegeben sind.

13.8. Proposition

Es sei $\sum c_n(x - x_0)^n$ eine Potenzreihe mit $x_0 \in \mathbb{R}$, $c_n \in \mathbb{R}$ für alle $n \in \mathbb{N}$ und Konvergenzradius $R > 0$. Wir setzen $I :=]x_0 - R, x_0 + R[\subseteq \mathbb{R}$ und $f: I \rightarrow \mathbb{R}$,

$$f(x) := \sum_{n=0}^{\infty} c_n(x - x_0)^n \quad (x \in I).$$

Dann gilt:

1.) $f \in \mathcal{C}^\infty(I, \mathbb{R})$, d.h. f ist unendlich oft differenzierbar.

2.) Für alle $x \in I$ gilt: $f'(x) = \sum_{n=0}^{\infty} n c_n (x - x_0)^{n-1}$.

3.) Für alle $a, b \in I$ gilt: $\int_a^b f(x) dx = \sum_{n=0}^{\infty} c_n \frac{(x - x_0)^{n+1}}{n+1} \Big|_a^b$.

Beweis: Nach Proposition 13.2 ist für jedes $0 < r < R$ sowohl die Potenzreihe wie auch die gliedweise differenzierte Reihe gleichmäßig konvergent auf $[x_0 - r, x_0 + r]$. Daher folgt aus Proposition 12.9 die stetige Differenzierbarkeit von f und die Formel für f' in 2.).

Durch sukzessive Anwendung derselben Argumentation auf f', f'', \dots folgt weiters die Behauptung 1.).

3.) ist eine direkte Konsequenz aus den Propositionen 13.2 und 12.7. □

13.9. Korollar

Es sei $f:]x_0 - R, x_0 + R[\rightarrow \mathbb{R}$, $f(x) = \sum c_n (x - x_0)^n$ durch eine Potenzreihe mit reellen Koeffizienten und Entwicklungspunkt sowie Konvergenzradius $R > 0$ gegeben. Dann gilt

$$\forall n \in \mathbb{N}: \quad c_n = \frac{f^{(n)}(x_0)}{n!},$$

d.h. die Taylor-Reihe von f um x_0 ist gleich der Potenzreihe $\sum c_n (x - x_0)^n$.

Insbesondere sind die Koeffizienten einer Potenzreihe also stets eindeutig bestimmt (nämlich durch die Ableitungen von f bei x_0).

Beweis: Nach Proposition 13.8 ist f unendlich oft differenzierbar und sukzessive Anwendung der Ableitungsformel liefert

$$f^{(n)}(x) = \sum_{k=n}^{\infty} c_k \cdot k(k-1) \cdots (k-n+1) (x - x_0)^{k-n}.$$

Für $x = x_0$ folgt daraus $f^{(n)}(x_0) = c_n \cdot n!$. □

13.10. Beispiele

1) In Analysis 1 hatten wir für $0 \leq x \leq 1$ die Gültigkeit der folgenden Taylor-Reihenentwicklung nachgewiesen:

$$\log(1+x) = \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n} x^n.$$

Wir können nun aber alternativ aus Proposition 13.8 herleiten, dass diese Entwicklung im Intervall $] - 1, 1[$ gilt, wodurch dann insgesamt die Gültigkeit in $] - 1, 1[$ bewiesen ist.

In 13.7.1) haben wir schon den Wert $R = 1$ für den Konvergenzradius dieser Potenzreihe ermittelt. Wir setzen $f(x) := \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n} x^n$ für $x \in] - 1, 1[$ und erhalten durch gliedweise Differentiation

$$f'(x) = \sum_{n=1}^{\infty} (-1)^{n-1} x^{n-1} = \sum_{k=0}^{\infty} (-x)^k = \frac{1}{1+x}.$$

Daher folgt auch für alle $t \in] - 1, 1[$

$$f(t) = f(0) + \int_0^t f'(x) dx = 0 + \int_0^t \frac{1}{1+x} dx = \log(1+t).$$

2) Wir bestimmen die Taylor-Reihenentwicklung für die Funktion $f:] - 1, 1[\rightarrow \mathbb{R}$, $f(x) = \arctan(x)$ durch einen ähnlichen Trick wie in 1).

Es gilt

$$f'(x) = \frac{1}{1+x^2} = \frac{1}{1-(-x^2)} = \sum_{k=0}^{\infty} (-1)^k x^{2k},$$

wobei der Konvergenzradius der auftretenden Potenzreihe $R = \lim |(-1)^k|^{1/2k} = 1$ ist (beachte $c_{2k} = (-1)^k$ und $c_{2k+1} = 0$). Daher folgt aus Proposition 13.8 für alle $t \in] - 1, 1[$

$$\arctan(t) = f(t) = f(0) + \int_0^t f'(x) dx = 0 + \sum_{k=0}^{\infty} (-1)^k \int_0^t x^{2k} dx = \sum_{k=0}^{\infty} (-1)^k \frac{t^{2k+1}}{2k+1}.$$

§14. Fourier-Reihen

In der Physik und Elektrotechnik gehört die Zerlegung periodischer Signale oder allgemeiner Schwingungen in Grundschwingungen und Grundfrequenzen zu den Standardmethoden. Mathematisch gesprochen geht es dabei um die Darstellung oder Approximation periodischer Funktionen als Summen bzw. Reihen von Winkelfunktionen bzgl. einer Folge von Frequenzen.

14.1. Periodische Funktionen

Sei $L > 0$; $f: \mathbb{R} \rightarrow \mathbb{C}$ heißt *periodisch* mit Periode L (kurz L -periodisch), wenn

$$(14.1) \quad \forall x \in \mathbb{R}: \quad f(x + L) = f(x).$$

Es folgt dann induktiv $\forall k \in \mathbb{Z}, \forall x \in \mathbb{R}: f(x + kL) = f(x)$. Daher ist eine L -periodische Funktion $f: \mathbb{R} \rightarrow \mathbb{C}$ bereits festgelegt, sobald sie auf einem Intervall der Form $[x_0, x_0 + L[$ bekannt ist für ein beliebiges $x_0 \in \mathbb{R}$.

Es gilt für $f: \mathbb{R} \rightarrow \mathbb{C}$ und $L > 0$

f ist L -periodisch $\Leftrightarrow F(x) := f\left(\frac{L}{2\pi}x\right)$ ist 2π -periodisch.

Daher „genügt“ es also, 2π -periodische Funktionen zu studieren und man kann stets mittels Skalierung rückübersetzen auf L -periodische Funktionen. Wir betrachten deshalb im weiteren oBdA nur noch Periode 2π .

Beispiele: $\sin, \cos: \mathbb{R} \rightarrow \mathbb{R}; x \mapsto \exp(ix), \mathbb{R} \rightarrow \mathbb{C}$ sind 2π -periodisch;

ebenso sind für $k \in \mathbb{Z}$ die Funktionen $e_k: \mathbb{R} \rightarrow \mathbb{C}, e_k(x) := e^{ikx}$ bzw. $x \mapsto \sin(kx), x \mapsto \cos(kx)$ 2π -periodisch;

konstante Funktionen sind periodisch bzgl. jeder Periode $L > 0$.

Definition

Es seien $n \in \mathbb{N}$ und $a_k, b_k \in \mathbb{R}$ ($k = 0, \dots, n$). Dann heißt $p_n: \mathbb{R} \rightarrow \mathbb{C}$ gegeben durch

$$p_n(x) = \frac{a_0}{2} + \sum_{k=1}^n (a_k \cos(kx) + b_k \sin(kx))$$

trigonometrisches Polynom (der Ordnung n).

Die komplexe Version trigonometrischer Polynome der Ordnung $n \in \mathbb{N}$ erhalten wir durch Koeffizienten $c_k \in \mathbb{C}$ ($-n \leq k \leq n$) als Abbildung $q_n: \mathbb{R} \rightarrow \mathbb{C}$ mit der Darstellung

$$q_n(x) = \sum_{k=-n}^n c_k e^{ikx}.$$

Mittels der Euler-Relationen $\cos(kx) = (e^{ikx} + e^{-ikx})/2$ und $\sin(kx) = (e^{ikx} - e^{-ikx})/2i$ erhalten wir nach Koeffizientenvergleich

$$q_n = p_n \iff c_0 = \frac{a_0}{2}, c_k = \frac{1}{2}(a_k - ib_k), c_{-k} = \frac{1}{2}(a_k + ib_k) = \overline{c_k} \quad (k \in \mathbb{N}, k \geq 1).$$

14.2. Fourier-Koeffizienten

Mini-Exkurs über Integration und Differentiation komplexwertiger Funktionen:

Für $f: [a, b] \rightarrow \mathbb{C}$ setze $u(x) := \operatorname{Re}(f(x))$, $v(x) := \operatorname{Im}(f(x))$, somit erhalten wir reelle Funktionen $u, v: [a, b] \rightarrow \mathbb{R}$ mit der Eigenschaft $\forall x \in [a, b]: f(x) = u(x) + iv(x)$.

f heißt Riemann-integrierbar, falls u und v R-integrierbar sind. Wir setzen in diesem Fall

$$\int_a^b f(x) dx := \int_a^b u(x) dx + i \int_a^b v(x) dx.$$

Ebenso verfahren wir beim Begriff der Differenzierbarkeit: f ist differenzierbar, falls u und v es sind und wir setzen $f' := u' + iv'$.

Es sei nun $f: \mathbb{R} \rightarrow \mathbb{C}$ ein trigonometrisches Polynom der Ordnung n , also von der Form $f(x) = \sum_{l=-n}^n c_l e^{ilx}$. Können wir die Koeffizienten c_l aus der Kenntnis von f bestimmen?

Durch Multiplikation mit e^{-ikx} erhalten wir zunächst

$$(\star) \quad e^{-ikx} \cdot f(x) = \sum_{l=-n}^n c_l e^{i(l-k)x}.$$

Bevor wir dies weiter umformen machen wir folgende entscheidende Beobachtung: das Integral $\int_0^{2\pi} e^{imx} dx$ ergibt für $m = 0$ den Wert $\int_0^{2\pi} 1 dx = 2\pi$ und für $m \neq 0$ haben wir eine Stammfunktion e^{imx}/m , somit wegen der 2π -Periodizität also $\int_0^{2\pi} e^{imx} dx = \frac{e^{imx}}{m} \Big|_0^{2\pi} = 0$; zusammenfassend gilt

$$(14.2) \quad \int_0^{2\pi} e^{imx} dx = \begin{cases} 2\pi & m = 0 \\ 0 & m \neq 0. \end{cases}$$

Integrieren wir nun beide Seiten der Gleichung (\star) über das Intervall $[0, 2\pi]$ (und benützen zur vereinfachten Notation das Kronecker-Delta¹: $\delta_{lk} = 0$ für $l \neq k$, $\delta_{kk} = 1$), so erhalten wir

$$\int_0^{2\pi} e^{-ikx} f(x) dx = \sum_{l=-n}^n c_l \int_0^{2\pi} e^{i(l-k)x} dx = \sum_{l=-n}^n c_l \cdot 2\pi \cdot \delta_{lk} = 2\pi \cdot c_k.$$

Folglich ist c_k durch f wie folgt bestimmt:

$$c_k = \frac{1}{2\pi} \int_0^{2\pi} e^{-ikx} f(x) dx.$$

Für allgemeinere 2π -periodische Funktionen erheben wir diese Relation nun zur

Definition: Sei $f: \mathbb{R} \rightarrow \mathbb{C}$ eine 2π -periodische Funktion und \mathbb{R} -integrierbar auf $[0, 2\pi]$. Dann sind

$$(14.3) \quad c_k := \frac{1}{2\pi} \int_0^{2\pi} e^{-ikx} f(x) dx \quad (k \in \mathbb{Z})$$

die (komplexen) *Fourier-Koeffizienten*² von f und

$$(14.4) \quad \sum_{k=-\infty}^{+\infty} c_k e^{ikx}$$

heißt zugeordnete *Fourier-Reihe*.

14.3. Reelle Version der Fourier-Koeffizienten

Ist f reell-wertig, d.h. $\overline{f(x)} = f(x)$ für alle $x \in \mathbb{R}$, dann folgt aus (14.3) auch $\overline{c_k} = c_{-k}$ und somit

$$c_k e^{ikx} + c_{-k} e^{-ikx} = c_k e^{ikx} + \overline{c_k e^{ikx}} = 2 \operatorname{Re}(c_k e^{ikx}) = 2 \operatorname{Re}(c_k) \cos(kx) - 2 \operatorname{Im}(c_k) \sin(kx).$$

¹Leopold Kronecker (*7. 12. 1823 Liegnitz; †29. 12. 1891 Berlin), deutscher Mathematiker

²Jean Baptiste Joseph Fourier (*21. 3. 1768 Auxerre; †16. 5. 1830 Paris) [ʃɔ̃ bap'tist fo'sɛf fu'rʒe], französischer Mathematiker und Physiker

Setzen wir das in die Partialsummen der Fourier-Entwicklung ein, dann ergibt sich (vgl. die reelle Form der trigonometrischen Polynome in 14.1)

$$\begin{aligned} \sum_{k=-n}^n c_k e^{ikx} &= c_0 + \sum_{k=1}^n (c_k e^{ikx} + c_{-k} e^{-ikx}) = \\ &= \underbrace{c_0}_{a_0/2} + \sum_{k=1}^n \left(\underbrace{2 \cdot \operatorname{Re}(c_k)}_{a_k} \cdot \cos(kx) + \underbrace{(-2 \cdot \operatorname{Im}(c_k))}_{b_k} \cdot \sin(kx) \right). \end{aligned}$$

Insgesamt erhalten wir für die reellen Fourier-Koeffizienten a_k, b_k aus (14.3) die Formeln

$$(14.5(a)) \quad a_k = \frac{1}{\pi} \int_0^{2\pi} f(x) \cos(kx) dx \quad (k = 0, 1, 2, \dots)$$

$$(14.5(b)) \quad b_k = \frac{1}{\pi} \int_0^{2\pi} f(x) \sin(kx) dx \quad (k = 1, 2, \dots).$$

Wir lesen aus obigen Integralformeln folgende Eigenschaften ab:

- ist f gerade, d.h. $f(-x) = f(x)$, dann folgt $b_k = 0$ für alle k ,
- ist f ungerade, d.h. $f(-x) = -f(x)$, dann gilt $a_k = 0$ für alle k .

14.4. Beispiel

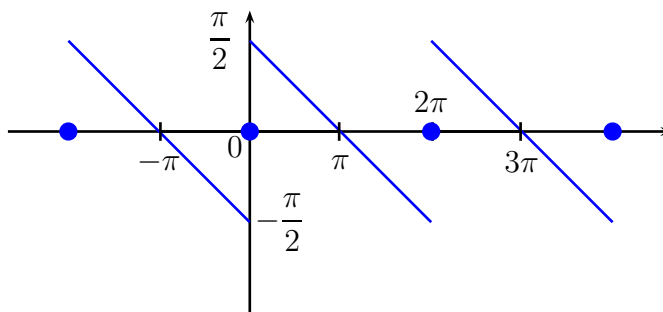
Wir kommen zurück zu der schon in Analysis 1 bewiesenen Relation

$$\forall x \in \mathbb{R}: \quad \sum_{k=1}^{\infty} \frac{\sin(kx)}{k} = f(x),$$

wobei $f: \mathbb{R} \rightarrow \mathbb{R}$ die 2π -periodische Fortsetzung folgender Funktion ist

$$f(x) = \begin{cases} 0 & x = 0 \\ \frac{\pi-x}{2} & x \in]0, 2\pi[. \end{cases}$$

Wie bereits in 12.4(ii) festgestellt, ist die Konvergenz punktweise für alle $x \in \mathbb{R}$, aber nicht gleichmäßig auf \mathbb{R} , weil die Grenzfunktion Unstetigkeitsstellen in den Punkten $2k\pi$ ($k \in \mathbb{Z}$) hat.



Behauptung: Obige Summe ist genau die Fourier-Reihe für f , d.h. in diesem Fall wird die Funktion f (punktweise) durch ihre Fourier-Reihe dargestellt.

Beweis. Wir bestimmen die (komplexen) Fourierkoeffizienten von f : zunächst bemerken wir, dass allgemein der Wert eines R-Integrals $\int_a^b g(x) dx$ gleich bleibt, wenn der Integrand g nur an endlich vielen Stellen abgeändert wird. Daher dürfen wir hier für den Integranden einfach die Funktion $(\pi - x)/2$ verwenden und erhalten für $k = 0$: $2\pi c_0 = \int_0^{2\pi} \frac{\pi-x}{2} dx = 0$; und für $k \neq 0$:

$$2\pi c_k = \int_0^{2\pi} e^{-ikx} \frac{\pi-x}{2} dx = \frac{\pi}{2} \underbrace{\int_0^{2\pi} e^{-ikx} dx}_0 - \frac{1}{2} \int_0^{2\pi} x e^{-ikx} dx =$$

$$[\text{part. Int.}] = \frac{1}{2ik} x e^{-ikx} \Big|_0^{2\pi} - \frac{1}{2ik} \underbrace{\int_0^{2\pi} e^{-ikx} dx}_0 = \frac{\pi}{ik},$$

daher ist $c_k = -i/2k$.

Für die reelle Version der Fourierkoeffizienten erhalten wir daraus $a_k = 2 \operatorname{Re}(c_k) = 0$ und $b_k = -2 \operatorname{Im}(c_k) = 1/k$, d.h. $\sum_{k=1}^{\infty} \frac{\sin(kx)}{k}$ ist tatsächlich genau die Fourier-Reihe von f . \square

Es gilt sogar folgende zusätzliche Aussage: Für jedes $0 < \delta < \pi$ konvergiert die Fourier-Reihe von f auf dem Intervall $[\delta, 2\pi - \delta]$ gleichmäßig gegen f . (Hinweise für einen Beweis finden sich in einem [freiwilligen] Zusatzbeispiel im Proseminar.)

14.5. Bemerkung

Wir haben in der obigen Berechnung bereits verwendet, dass alle Fourierkoeffizienten c_k — und somit auch die Fourier-Reihe — unverändert bleiben, solange f nur an endlich vielen Stellen im Intervall $[0, 2\pi[$ beliebig abgeändert wird. Daher ist klar, dass die Forderung der punktweisen Konvergenz der Fourier-Reihe gegen die Funktionswerte **in allen Punkten**

eines Periodizitätsintervalls im Allgemeinen nicht angemessenen sein kann. Wir werden daher weiter unten eine andere Art von Konvergenz für Fourier-Reihen untersuchen.

Man kann zeigen, dass die Fourier-Reihe einer 2π -periodischen *stückweise stetig differenzierbaren* Funktion f (d.h. es gibt eine Zerlegung $0 = t_0 < t_1 < \dots < t_N = 2\pi$, so dass $f|_{]t_{j-1}, t_j[}$ stetig differenzierbar und stetig fortsetzbar auf $[t_{j-1}, t_j]$ ist) auf jedem endlichen abgeschlossenen Teilintervall innerhalb des Stetigkeitsbereichs gleichmäßig gegen die Funktion konvergiert; an den (endlich vielen) Unstetigkeitsstellen ist der Limes der Fourier-Reihe gleich dem arithmetischen Mittel aus links- und rechtsseitigem Grenzwert der Funktion, d.h. es gilt dann in allen Punkten zumindest

$$\sum_{k=-\infty}^{+\infty} c_k e^{ikx} = \frac{1}{2} \left(\lim_{t \nearrow x} f(t) + \lim_{t \searrow x} f(t) \right)$$

(denn in Stetigkeitspunkten steht rechts einfach $f(t)$). (Vgl. z.B. [Heu04, Sätze 136.3 und 137.2].)

14.6. Beispiel

Wir bestimmen die reelle Version der Fourier-Reihe der Funktion $g : \mathbb{R} \rightarrow \mathbb{R}$, die als 2π -periodische Fortsetzung von

$$g(x) = \frac{(x - \pi)^2}{4} - \frac{\pi^2}{12} \quad (x \in [0, 2\pi[)$$

gegeben ist. Da g gerade ist, sind alle $b_k = 0$. Weiters ist

$$\pi a_0 = \int_0^{2\pi} g(x) dx = \int_0^{2\pi} \frac{(x - \pi)^2}{4} dx - \frac{\pi^2}{12} \cdot 2\pi = \left[\frac{(x - \pi)^3}{12} \right]_0^{2\pi} - \frac{\pi^3}{6} = \frac{\pi^3}{12} + \frac{\pi^3}{12} - \frac{\pi^3}{6} = 0.$$

Für die Berechnung von a_k für $k > 0$ haben wir

$$\begin{aligned} a_k &= \frac{1}{\pi} \int_0^{2\pi} g(x) \cos(kx) dx = \frac{1}{4\pi} \underbrace{\int_0^{2\pi} (x - \pi)^2 \cos(kx) dx}_{\text{[part. Int.]}} - \frac{\pi}{12} \underbrace{\int_0^{2\pi} \cos(kx) dx}_{=0} \\ &= \frac{1}{4\pi} \left(\underbrace{\left. \frac{(x - \pi)^2}{k} \sin(kx) \right|_0^{2\pi}}_{=0} - \frac{2}{k} \underbrace{\int_0^{2\pi} (x - \pi) \sin(kx) dx}_{\text{[part. Int.]}} \right) \\ &= -\frac{1}{2\pi k} \left(\left. (x - \pi) \frac{-\cos(kx)}{k} \right|_0^{2\pi} + \underbrace{\int_0^{2\pi} \frac{\cos(kx)}{k} dx}_{=0} \right) \\ &= \frac{1}{2\pi k^2} (\pi \cos(2k\pi) + \pi \cos(0)) = \frac{1}{2k^2} (1 + 1) = \frac{1}{k^2}. \end{aligned}$$

Daher lautet die Fourier-Reihe von g also

$$\sum_{k=1}^{\infty} \frac{\cos(kx)}{k^2}.$$

Wie wir bereits in 12.6.2) bemerkt haben, ist diese Reihe gleichmäßig konvergent. Es gilt diesem Fall wiederum, dass die Funktion durch ihre Fourier-Reihe (sogar als gleichmäßiger Limes) dargestellt wird, d.h. es gilt im Sinne gleichmäßiger Konvergenz

$$\frac{(x - \pi)^2}{4} - \frac{\pi^2}{12} = \sum_{k=1}^{\infty} \frac{\cos(kx)}{k^2} \quad \forall x \in [0, 2\pi].$$

Dies folgt allgemeiner für stetige stückweise stetig differenzierbare Funktionen (vgl. [For06, §23, Satz 3]), jedoch können wir mit Hilfe der am Ende von Beispiel 14.4 angegebenen Eigenschaft einen direkten Beweis dafür geben: demnach ist für jedes $\delta \in]0, \pi[$ die Reihe

$$\sum_{k=1}^{\infty} \left(\frac{\cos(kx)}{k^2} \right)' = - \sum_{k=1}^{\infty} \frac{\sin(kx)}{k} = \frac{x - \pi}{2}$$

gleichmäßig konvergent im Intervall $[\delta, 2\pi - \delta]$.

Wegen $g'(x) = (x - \pi)/2$ für alle $x \in [0, 2\pi]$ gilt also nach Proposition 12.9, dass g und die Summe der Fourier-Reihe sich nur um eine Konstante c unterscheiden (und zwar durch stetige Ausdehnung auf dem gesamten Bereich). Durch Integration der Differenz von g und der Summenfunktion über das Periodizitätsintervall erhalten wir (dank obiger Berechnung von $a_0 = 0$ und Vertauschung der Summe mit dem Integral [Prop. 12.7]), dass $c = 0$ sein muss. Daher gilt also

$$g(x) = \sum_{k=1}^{\infty} \frac{\cos(kx)}{k^2} \quad \forall x \in [0, 2\pi].$$

- Speziell für $x = 0$ ergibt sich $\sum_{k=1}^{\infty} \frac{1}{k^2} = \frac{\pi^2}{4} - \frac{\pi^2}{12} = \frac{\pi^2}{6}$
- Für $x = \pi$ erhalten wir $\sum_{k=1}^{\infty} \frac{(-1)^k}{k^2} = -\frac{\pi^2}{12}$, d.h. $\sum_{k=1}^{\infty} \frac{(-1)^{k-1}}{k^2} = \frac{\pi^2}{12}$.

14.7. Eine hermitesche Form für R-integrierbare Funktionen

Die auf dem Intervall $[0, 2\pi]$ komplex-wertigen, Riemann-integrierbaren Funktionen $\mathcal{R}[0, 2\pi]$ bilden einen Vektorraum über \mathbb{C} , wenn die Operationen wie üblich punktweise definiert werden, d.h. für $f, g \in \mathcal{R}[0, 2\pi]$ und $\lambda \in \mathbb{C}$ setzen wir

$$(f + g)(x) := f(x) + g(x), \quad (\lambda f)(x) := \lambda \cdot f(x);$$

und sowohl $f + g$ als auch λf sind ebenfalls R-integrierbar.

Aus der Linearen Algebra kennen wir den m -dimensionalen komplexen Vektorraum \mathbb{C}^m als unitären Vektorraum ausgestattet mit dem Skalarprodukt (d.i. eine positiv definite hermitesche Sesquilinearform) definiert für $z = (z_1, \dots, z_m), w = (w_1, \dots, w_m) \in \mathbb{C}^m$ durch

$$\langle w | z \rangle = \sum_{k=1}^m w_k \bar{z}_k$$

und der daraus resultierenden Norm $\|w\| := \sqrt{\langle w | w \rangle}$ (siehe z.B. [Fis03, Kapitel 5]).

In Analogie dazu verwenden wir nun $\int_0^{2\pi} f \bar{g}$, um eine Hermitesche Form auf $\mathcal{R}[0, 2\pi]$ zu erhalten. Übrigens haben wir in Analysis 1 (§9) bewiesen, dass für $f, g \in \mathcal{R}[0, 2\pi]$ auch die R-Integrierbarkeit von $f \bar{g}$ und $|f|^2$ folgt.

Definition: Es sei $\langle \cdot | \cdot \rangle: \mathcal{R}[0, 2\pi] \times \mathcal{R}[0, 2\pi] \rightarrow \mathbb{C}$ gegeben durch

$$(14.6) \quad \langle f | g \rangle := \frac{1}{2\pi} \int_0^{2\pi} f(x) \overline{g(x)} dx \quad \forall f, g \in \mathcal{R}[0, 2\pi]$$

und $\|\cdot\|_2: \mathcal{R}[0, 2\pi] \rightarrow \mathbb{R}$ die (so genannte) 2-Norm

$$(14.7) \quad \|f\|_2 := \sqrt{\langle f | f \rangle} = \left(\frac{1}{2\pi} \int_0^{2\pi} |f(x)|^2 dx \right)^{1/2}$$

(Wegen $|f(x)|^2 \geq 0$ folgt aus der Monotonie des Integrals, dass $\langle f | f \rangle \geq 0$ gilt; daher ist $\|f\|_2$ für jedes $f \in \mathcal{R}[0, 2\pi]$ definiert.)

Proposition: Für $f, g, h \in \mathcal{R}[0, 2\pi], \lambda, \mu \in \mathbb{C}$ gilt:

- 1.) $\langle \lambda f + \mu g | h \rangle = \lambda \langle f | h \rangle + \mu \langle g | h \rangle$ [d.h. $f \mapsto \langle f | h \rangle$ ist linear]
- 2.) $\langle f | \lambda g + \mu h \rangle = \bar{\lambda} \langle f | g \rangle + \bar{\mu} \langle f | h \rangle$ [d.h. $h \mapsto \langle f | h \rangle$ ist konjugiert-linear]

- 3.) $\langle f | g \rangle = \overline{\langle g | f \rangle}$ [d.h. $\langle \cdot | \cdot \rangle$ ist hermitesch]
 4.) $\langle f | f \rangle \geq 0$ (d.h. $\langle \cdot | \cdot \rangle$ ist positiv semidefinit).

Beweis: 1.) und 2.) folgen direkt aus der Linearität des Integrals, 3.) durch Zerlegung in Real- und Imaginärteil. 4.) folgt direkt aus der Definition in (14.7) (bzw. der Klammerbemerkung unmittelbar danach). \square

Bemerkung: $\langle \cdot | \cdot \rangle$ ist kein „richtiges“ Skalarprodukt auf $\mathcal{R}[0, 2\pi]$ (und daher auch $\|\cdot\|_2$ keine „richtige“ Norm), weil $f \neq 0$ (Nullfunktion!) nicht impliziert, dass $\langle f | f \rangle > 0$ gilt, d.h. $\langle \cdot | \cdot \rangle$ ist auf den \mathbb{R} -integrierbaren Funktionen **nicht positiv definit** (es folgt eben i.A. nur $\langle f | f \rangle \geq 0$): als Beispiel betrachten wir die Funktion

$$f(x) = \begin{cases} 0 & x \neq \pi \\ 1 & x = \pi; \end{cases}$$

es gilt $f \in \mathcal{R}[0, 2\pi]$, $f \neq 0$, aber $\int_0^{2\pi} f(x)^2 dx = 0$.

Dieser „Defekt“ stört die Anwendung auf Fourier-Reihen in der Praxis nicht und kann außerdem auch strukturell sozusagen umgangen werden, indem man „Klassen von Funktionen bildet, die sich nur auf solchen Mengen unterscheiden, die bei Integration vernachlässigbar sind“ (\leadsto was sind das für Mengen? [führt auf so genannte Nullmengen, im Jordan- oder Lebesgue-Sinn].)

Auf stetigen Funktionen ergibt $\langle \cdot | \cdot \rangle$ übrigens sehr wohl ein Skalarprodukt, d.h. die Einschränkung auf $\mathcal{C}([0, 2\pi]) \times \mathcal{C}([0, 2\pi])$ ist positiv definit (Übungsaufgabe!).

Für $k \in \mathbb{Z}$ sei $e_k \in \mathcal{R}[0, 2\pi]$ definiert durch

$$(14.8) \quad e_k(x) := e^{ikx} \quad \forall x \in [0, 2\pi].$$

Dann entspricht in dieser Notation der Relation (14.2) die folgende Gleichung

$$(14.9) \quad \langle e_k | e_l \rangle = \delta_{kl} = \begin{cases} 0 & k \neq l \\ 1 & k = l \end{cases} \quad (k, l \in \mathbb{Z}).$$

Die Relation (14.9) besagt, dass $\{e_k : k \in \mathbb{Z}\}$ ein Orthonormalsystem bzgl. $\langle \cdot | \cdot \rangle$ bildet und Formel (14.3) für die Fourierkoeffizienten einer Funktion $f \in \mathcal{R}[0, 2\pi]$ können wir nun einfach so schreiben:

$$c_k = \langle f | e_k \rangle \quad (k \in \mathbb{Z});$$

die Fourier-Reihenentwicklung wird dadurch formal ähnlich zur Darstellung eines Vektors f bzgl. einer Orthonormalbasis

$$\sum_{k=-\infty}^{+\infty} \langle f | e_k \rangle \cdot e_k.$$

(In seiner abstrakten Ausprägung führt dies zur Theorie der Hilberträume und konkret zum Raum der quadratisch Lebesgue-integrierbaren Funktionen $\mathcal{L}^2[0, 2\pi]$; \leadsto Funktionalanalysis bzw. Maß- und Integrationstheorie.)

Lemma: Es seien $f, g \in \mathcal{R}[0, 2\pi]$. Dann gilt:

1.) Die Cauchy-Schwarz-Ungleichung

$$(14.10) \quad |\langle f | g \rangle| \leq \|f\|_2 \cdot \|g\|_2.$$

2.) Die Dreiecksungleichung

$$(14.11) \quad \|f + g\|_2 \leq \|f\|_2 + \|g\|_2$$

(das ist ein Spezialfall der Minkowski-Ungleichung in Abschnitt 15 unten).

Beweis: 1.) Beide Seiten der behaupteten Ungleichung können wir als Limiten von entsprechenden Riemann-Summen schreiben. Daher genügt es, folgende Version der Ungleichung für endliche Summen zu beweisen: Es sei $m \in \mathbb{N}$ und $x_1, \dots, x_m, y_1, \dots, y_m \in \mathbb{C}$, dann gilt

$$(\star) \quad \sum_{k=1}^m |x_k| |y_k| \leq \left(\sum_{k=1}^m |x_k|^2 \right)^{1/2} \cdot \left(\sum_{k=1}^m |y_k|^2 \right)^{1/2}$$

(wegen $|\sum x_k \overline{y_k}| \leq \sum |x_k| |y_k|$ impliziert dies die behauptete Ungleichung.)

Wir setzen $A := \sqrt{\sum |x_k|^2}$ und $B := \sqrt{\sum |y_k|^2}$ und dürfen $A, B > 0$ annehmen, weil andernfalls alle x_k oder alle y_k verschwinden und die Ungleichung dann trivial ist. Mit der Notation $\alpha_k := |x_k|/A$ und $\beta_k := |y_k|/B$ ($k = 1, \dots, m$) ist dann (\star) äquivalent zur Aussage: $\sum \alpha_k \beta_k \leq 1$.

Diese wiederum ergibt sich leicht aus der, in Analysis 1 (§8) bewiesenen, Ungleichung zwischen dem arithmetischen und geometrischen Mittel ($\forall a, b \geq 0: \sqrt{ab} \leq (a + b)/2$) durch folgende Rechnung:

$$\sum \alpha_k \beta_k = \sum \sqrt{\alpha_k^2 \beta_k^2} \leq \sum \frac{\alpha_k^2 + \beta_k^2}{2} = \frac{1}{2} \sum \alpha_k^2 + \frac{1}{2} \sum \beta_k^2 = \frac{1}{2} + \frac{1}{2} = 1,$$

weil ja $\sum \alpha_k^2 = \sum \beta_k^2 = 1$ nach Konstruktion gilt.

2.) Wir verwenden die Cauchy-Schwarz-Ungleichung in der folgenden Rechnung:

$$\begin{aligned} \|f + g\|_2^2 &= \langle f + g | f + g \rangle = \langle f | f \rangle + 2 \operatorname{Re} \langle f | g \rangle + \langle g | g \rangle \\ &\leq \|f\|_2^2 + 2\|f\|_2 \|g\|_2 + \|g\|_2^2 = (\|f\|_2 + \|g\|_2)^2. \end{aligned}$$

□

14.8. Approximation im quadratischen Mittel

Lemma

Sei $f \in \mathcal{R}[0, 2\pi]$ mit den Fourier-Koeffizienten $(c_k)_{k \in \mathbb{Z}}$.

Dann gilt:

$$1.) \quad \forall n \in \mathbb{N}: \quad \left\| f - \sum_{k=-n}^n c_k e_k \right\|_2^2 = \|f\|_2^2 - \sum_{k=-n}^n |c_k|^2$$

$$2.) \quad \sum_{k=-\infty}^{+\infty} |c_k|^2 := \lim_{n \rightarrow \infty} \sum_{k=-n}^n |c_k|^2 \leq \|f\|_2^2 \quad (\text{Bessel-Ungleichung}^3).$$

Beweis: 1.) Setze $f_n := \sum_{k=-n}^n c_k e_k$, dann gilt

$$\langle f | f_n \rangle = \sum_{k=-n}^n \overline{c_k} \underbrace{\langle f | e_k \rangle}_{c_k} = \sum_{k=-n}^n |c_k|^2 = \overline{\langle f | f_n \rangle} = \langle f_n | f \rangle \quad \text{und}$$

$$\langle f_n | f_n \rangle = \sum_{k,l=-n}^n c_k \overline{c_l} \underbrace{\langle e_k | e_l \rangle}_{\delta_{kl}} = \sum_{k=-n}^n |c_k|^2, \text{ somit}$$

$$\begin{aligned} \|f - f_n\|_2^2 &= \langle f - f_n | f - f_n \rangle = \langle f | f \rangle - \overbrace{\langle f_n | f \rangle - \langle f | f_n \rangle}^{-2 \cdot \langle f | f_n \rangle} + \langle f_n | f_n \rangle \\ &= \|f\|_2^2 - 2 \cdot \sum |c_k|^2 + \sum |c_k|^2, \text{ also folgt 1.)} \end{aligned}$$

2.) Aus 1.) folgt $\sum_{k=-n}^n |c_k|^2 \leq \|f\|_2^2$ für alle $n \in \mathbb{N}$, daraus folgt 2.) für $n \rightarrow \infty$. □

Definition: Sei $f \in \mathcal{R}[0, 2\pi]$ und $(f_n)_{n \in \mathbb{N}}$ eine Folge mit $f_n \in \mathcal{R}[0, 2\pi]$; wir sagen, (f_n) konvergiere gegen f im quadratischen Mittel, falls

$$\lim_{n \rightarrow \infty} \|f - f_n\|_2 = 0$$

(man spricht von Konvergenz bzgl. der 2-Norm $\|\cdot\|_2$; äquivalent dazu ist also die Aussage $\int_0^{2\pi} |f(x) - f_n(x)|^2 dx \rightarrow 0$ für $n \rightarrow \infty$).

³Friedrich Wilhelm Bessel (*22. 7. 1784 Minden; †17. 3. 1846 Königsberg) war einer der bekanntesten deutschen Wissenschaftler des 19. Jahrhunderts.

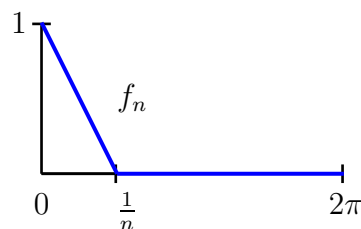
Bemerkung

1.) Falls $f_n \rightarrow f$ gleichmäßig auf $[0, 2\pi]$ und alle f_n stetig sind, dann folgt $\|f_n - f\|_2 \rightarrow 0$ [weil dann auch gilt, dass $|f_n - f|^2 \rightarrow 0$ glm., ist Prop. 12.7 anwendbar].

Die Umkehrung ist jedoch falsch!

Z.B. für f_n , gegeben durch $f_n(x) = 0$ ($\frac{1}{n} \leq x \leq 2\pi$) und $f_n(x) = 1 - nx$ ($0 \leq x \leq \frac{1}{n}$), ergibt sich

$$2\pi \|f_n\|_2^2 = \int_0^{1/n} (1 - nx)^2 dx = \frac{(1 - nx)^3}{-3n} \Big|_0^{1/n} = \frac{1}{3n} \rightarrow 0,$$



während $\|f_n\|_\infty = 1$ für alle n .

2.) Aus Punkt 1.) des obigen Lemmas schließen wir folgende nützliche Äquivalenz:

die Fourier-Reihe konvergiert im quadr. Mittel gegen $f \iff \sum_{k=-\infty}^{\infty} |c_k|^2 = \|f\|_2^2$.

Theorem

Sei $f: \mathbb{R} \rightarrow \mathbb{C}$ 2π -periodisch und Riemann-integrierbar auf $[0, 2\pi]$. Dann konvergiert die Fourier-Reihe von f im quadratischen Mittel gegen f . Sind c_k ($k \in \mathbb{Z}$) die Fourier-Koeffizienten, so gilt

$$(14.12) \quad \sum_{k=-\infty}^{+\infty} |c_k|^2 = \frac{1}{2\pi} \int_0^{2\pi} |f(x)|^2 dx$$

(Parseval-Gleichung⁴ oder Vollständigkeitsrelation)

Beweis: 1. Schritt: Das Theorem gilt für jede charakteristische Funktion der Form

$$f(x) = \mathbf{1}_{[0, a[}(x) = \begin{cases} 1 & 0 \leq x < a \\ 0 & a \leq x < 2\pi \end{cases} \quad (0 \leq a \leq 2\pi)$$

(mit 2π -periodischer Fortsetzung auf \mathbb{R}).

⁴Marc-Antoine Parseval des Chênes (*27. 4. 1755 Rosières-aux-Saline; †16. 8. 1836 Paris) [ˈmark ɑ̃ˈtwan parsəˈval de ʃɛˈne], französischer Mathematiker

Wir berechnen die Fourier-Koeffizienten: $c_0 = \frac{1}{2\pi} \int_0^a dx = \frac{a}{2\pi}$; für $k \neq 0$ erhalten wir $c_k = \frac{1}{2\pi} \int_0^a e^{-ikx} dx = \frac{i}{2\pi k} (e^{-ika} - 1)$, somit

$$|c_k|^2 = \frac{1}{4\pi^2 k^2} \underbrace{(e^{-ika} - 1)(e^{ika} - 1)}_{2 - e^{ika} - e^{-ika} = 2(1 - \cos(ka))} = \frac{1 - \cos(ka)}{2\pi^2 k^2}.$$

Daher gilt zunächst

$$\begin{aligned} \sum_{k=-\infty}^{+\infty} |c_k|^2 &= \frac{a^2}{4\pi^2} + 2 \cdot \sum_{k=1}^{\infty} \frac{1 - \cos(ka)}{2\pi^2 k^2} = \frac{a^2}{4\pi^2} + \frac{1}{\pi^2} \cdot \sum_{k=1}^{\infty} \frac{1}{k^2} - \frac{1}{\pi^2} \cdot \sum_{k=1}^{\infty} \frac{\cos(ka)}{k^2} \stackrel{[14.6]}{=} \\ &= \frac{a^2}{4\pi^2} + \frac{1}{\pi^2} \cdot \frac{\pi^2}{6} - \frac{1}{\pi^2} \left(\frac{(a - \pi)^2}{4} - \frac{\pi^2}{12} \right) = \frac{a^2}{4\pi^2} + \frac{1}{6} - \frac{a^2}{4\pi^2} + \frac{a\pi}{2\pi^2} - \frac{1}{4} + \frac{1}{12} = \frac{a}{2\pi}; \end{aligned}$$

andererseits ist auch

$$\|f\|_2^2 = \frac{1}{2\pi} \int_0^{2\pi} |f(x)|^2 dx = \frac{1}{2\pi} \int_0^a 1 dx = \frac{a}{2\pi},$$

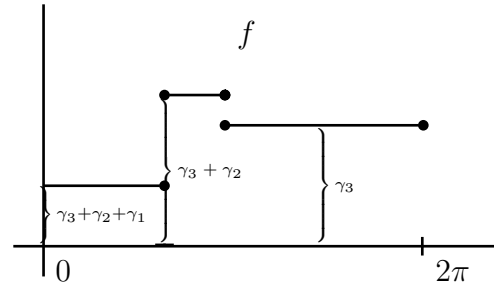
daher konvergiert nach obiger Bemerkung 2.) die Fourier-Reihe im quadratischen Mittel gegen f .

2. Schritt: Das Theorem gilt für Treppenfunktionen (mit 2π -periodischer Fortsetzung).

Es sei $N \in \mathbb{N}$ und $0 \leq a_1 < \dots < a_N \leq 2\pi$, $\gamma_1, \dots, \gamma_N \in \mathbb{R}$ und $f_j = \mathbf{1}_{[0, a_j[}$ ($1 \leq j \leq N$; jedes f_j ist vom Typ wie in Schritt 1).

Für $x \in [0, 2\pi]$ mit $x \neq a_j$ (mögliche Sprungstellen) sei die Treppenfunktion f gegeben durch

$$f(x) = \sum_{j=1}^N \gamma_j f_j(x).$$



Es bezeichne $F[f, n](x) = \sum_{k=-n}^n c_k e^{ikx}$ die n -te Partialsumme der Fourier-Reihe von f und $F[f_j, n]$ ($j = 1, \dots, N$) die entsprechenden Partialsummen für die Fourier-Reihe von f_j .

Es gilt $F[f, n] = \sum_{j=1}^N \gamma_j F[f_j, n]$ (endliche Summe mit fixer Anzahl von Summanden), daher folgt mit Hilfe der Dreiecksungleichung und Schritt 1 für $n \rightarrow \infty$ nun

$$\|f - F[n, f]\|_2 = \left\| \sum_{j=1}^N \gamma_j \cdot (f_j - F[f_j, n]) \right\|_2 \leq \sum_{j=1}^N |\gamma_j| \cdot \|f_j - F[f_j, n]\|_2 \rightarrow 0.$$

3. Schritt: allgemeiner Fall. OBdA ist f reellwertig (betrachte Real- und Imaginärteil getrennt) und $|f(x)| \leq 1$ (andernfalls Division durch $\|f\|_\infty$).

Sei $\varepsilon > 0$ beliebig. Dann gibt es $\varphi, \psi \in \mathcal{T}[0, 2\pi]$ (die wir uns 2π -periodisch fortgesetzt denken) mit folgenden Eigenschaften:

$$(a) \quad -1 \leq \varphi \leq f \leq \psi \leq 1$$

$$(b) \quad \int_0^{2\pi} (\psi(x) - \varphi(x)) \, dx \leq \frac{\pi}{4} \cdot \varepsilon^2.$$

Wir setzen $g := f - \varphi$, dann gilt $F[f, n] = F[g, n] + F[\varphi, n]$.

Gemäß Schritt 2 gibt es ein $n_0 \in \mathbb{N}$: $\forall n \geq n_0$: $\|\varphi - F[\varphi, n]\|_2 \leq \frac{\varepsilon}{2}$.

Aus 14.8, Lemma 1.), folgt zudem $\|g - F[g, n]\|_2^2 \leq \|g\|_2^2$.

Weiters ist $|g|^2 = |f - \varphi|^2 \leq (\psi - \varphi)^2 = \underbrace{(\psi - \varphi)}_{[0 \leq \cdot \leq 2]} \cdot (\psi - \varphi) \leq 2(\psi - \varphi)$, daher

$$\|g - F[g, n]\|_2^2 \leq \|g\|_2^2 \leq \frac{1}{2\pi} \cdot 2 \int_0^{2\pi} (\psi - \varphi)(x) \, dx \stackrel{[(b)]}{\leq} \frac{1}{\pi} \cdot \frac{\pi \varepsilon^2}{4} = \frac{\varepsilon^2}{4}.$$

Durch Kombination erhalten wir (wieder mittels Dreiecksungleichung) für $n \geq n_0$

$$\begin{aligned} \|f - F[f, n]\|_2 &= \|g + \varphi - F[g, n] - F[\varphi, n]\|_2 \\ &\leq \|g - F[g, n]\|_2 + \|\varphi - F[\varphi, n]\|_2 \leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon, \end{aligned}$$

also zusammenfassend $\|f - F[f, n]\|_2 \rightarrow 0$ ($n \rightarrow \infty$), somit wegen Lemma 1.) schließlich die behauptete Parseval-Gleichung

$$\|f\|_2^2 = \sum_{k=-\infty}^{+\infty} |c_k|^2.$$

□

Reelle Version der Parseval-Gleichung: Falls f reellwertig, 2π -periodisch und R-integrierbar auf $[0, 2\pi]$ ist, haben wir für die Koeffizienten der zugeordneten reellen Fourierreihe $c_0 = a_0/2$, und für $k \geq 1$ die Relationen $c_{-k} = \overline{c_k}$ sowie $c_k = (a_k - ib_k)/2$. Somit ist in der Parseval-Gleichung $|c_k|^2 = (a_k^2 + b_k^2)/4$ einzusetzen und wegen $|c_{-k}| = |c_k|$, $|f(x)|^2 = f(x)^2$ also

$$\frac{a_0^2}{4} + 2 \sum_{k=1}^{\infty} \frac{a_k^2 + b_k^2}{4} = \frac{1}{2\pi} \int_0^{2\pi} f(x)^2 \, dx,$$

daher gilt

$$(14.12') \quad \frac{a_0^2}{2} + \sum_{k=1}^{\infty} (a_k^2 + b_k^2) = \frac{1}{\pi} \int_0^{2\pi} f(x)^2 dx.$$

14.9. Beispiel

Wir betrachten nochmals die Fourier-Reihe aus Beispiel 14.6, wobei wir diesmal aber den konstanten Term $\pi^2/12$ der Reihenentwicklung zurechnen, d.h.

$$\underbrace{\frac{\pi^2}{12}}_{a_0/2} + \sum_{k=1}^{\infty} \underbrace{\frac{1}{k^2}}_{a_k} \cdot \cos(kx) = \frac{(x - \pi)^2}{4} \quad (x \in [0, 2\pi]).$$

Die reelle Version der Parseval-Gleichung (mit $a_0 = \pi^2/6$, $a_k = 1/k^2$ für $k \geq 1$, $b_k = 0$) liefert nun

$$\frac{\pi^4}{72} + \sum_{k=1}^{\infty} \frac{1}{k^4} = \frac{1}{\pi} \int_0^{2\pi} \frac{(x - \pi)^4}{16} dx = \frac{1}{16\pi} \cdot \frac{(x - \pi)^5}{5} \Big|_0^{2\pi} = \frac{\pi^5 - (-\pi)^5}{80\pi} = \frac{\pi^4}{40},$$

und daraus erhalten wir

$$\sum_{k=1}^{\infty} \frac{1}{k^4} = \pi^4 \left(\frac{1}{40} - \frac{1}{72} \right) = \frac{\pi^4}{8} \left(\frac{1}{5} - \frac{1}{9} \right) = \frac{\pi^4}{8} \cdot \frac{4}{45} = \frac{\pi^4}{90}.$$

VI TOPOLOGISCHE GRUNDBEGRIFFE

§15. Metrische und normierte Räume

15.1. Definition

Es sei X eine Menge. Eine *Metrik* auf X ist eine Abbildung $d: X \times X \rightarrow \mathbb{R}$, $(x, y) \mapsto d(x, y)$ mit folgenden Eigenschaften:

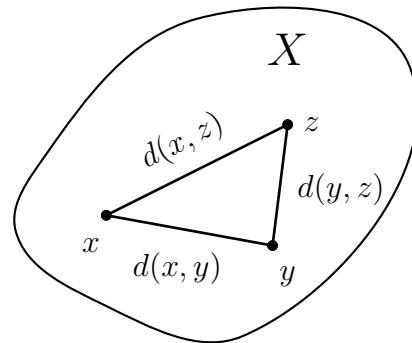
(M1) $d(x, y) \geq 0 \forall x, y \in X$; $d(x, y) = 0 \Leftrightarrow x = y$

(M2) Symmetrie: $\forall x, y: d(x, y) = d(y, x)$

(M3) Dreiecksungleichung: $\forall x, y, z \in X$ gilt

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

Wir nennen (X, d) einen *metrischen Raum* und $d(x, y)$ den *Abstand* oder die *Distanz* der Punkte $x, y \in X$ bzgl. der Metrik d .



15.2. Beispiele

- 1.) \mathbb{R} oder \mathbb{C} mit $d(x, y) := |x - y|$
- 2.) Sei (X, d) ein metrischer Raum und $A \subseteq X$. Dann ist $d_A := d|_{A \times A}$ die so genannte induzierte Metrik auf A und (A, d_A) ist ein metrischer Raum.
- 3.) Sei X eine beliebige Menge. Unter der trivialen oder diskreten Metrik auf X versteht man $d: X \times X \rightarrow \mathbb{R}$, gegeben durch

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

Im Rahmen dieser Vorlesung entstehen die wichtigsten Beispiele metrischer Räume aus Teilmengen von normierten Vektorräumen.

15.3. Definition

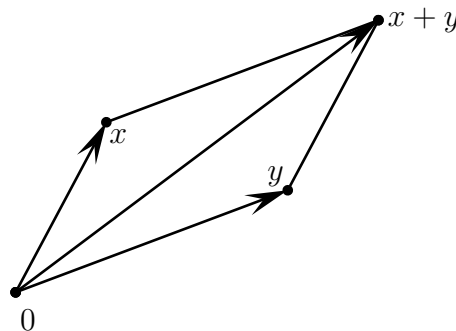
Sei V ein Vektorraum über dem Grundkörper $\mathbb{K} = \mathbb{R}$ oder \mathbb{C} . Eine *Norm* auf V ist eine Abbildung $\|\cdot\|: V \rightarrow \mathbb{R}$, $x \mapsto \|x\|$ mit folgenden Eigenschaften:

(N1) $\forall x \in V: \|x\| \geq 0; \|x\| = 0 \Leftrightarrow x = 0$

(N2) $\forall \lambda \in \mathbb{K}, \forall x \in V: \|\lambda x\| = |\lambda| \cdot \|x\|$

(N3) Dreiecksungleichung:
 $\forall x, y \in V$ gilt: $\|x + y\| \leq \|x\| + \|y\|$

$(V, \|\cdot\|)$ ist ein *normierter (Vektor)-Raum*.



15.4. Proposition

Sei $(V, \|\cdot\|)$ ein normierter Raum. Dann wird durch $d(x, y) := \|x - y\|$ ($x, y \in V$) eine Metrik auf V definiert.

Beweis. Übungsaufgabe. □

15.5. Beispiele: Euklidische Normen

\mathbb{R}^n mit Standardskalarprodukt $\langle x | y \rangle = \sum_{j=1}^n x_j y_j$, wobei $x = (x_j)_{j=1}^n, y = (y_j)_{j=1}^n \in \mathbb{R}^n$. Dann ist

$$(15.1) \quad \|x\| = \|x\|_2 := \sqrt{\langle x | x \rangle} = \left(\sum_{j=1}^n x_j^2 \right)^{1/2}$$

die *euklidische Norm (oder 2-Norm)*. Einen Beweis, dass tatsächlich (N1-3) erfüllt ist, erhält man aus den Überlegungen im Beweis des Lemmas in 14.7 oder als Spezialfall für $p = 2$ in 15.6 unten.

Entsprechend ist

$$d(x, y) := \|x - y\|_2 = \left(\sum_{j=1}^n (x_j - y_j)^2 \right)^{1/2}$$

der *euklidische Abstand* zwischen x und y .

Allgemeiner: Ist V ein euklidischer Vektorraum mit Skalarprodukt $\langle \cdot | \cdot \rangle: V \times V \rightarrow \mathbb{R}$ (symmetrische, positiv definite Bilinearform), dann definiert $\|x\| := \sqrt{\langle x | x \rangle}$ eine Norm auf V (siehe z.B. [Fis03, Kapitel 5]).

15.6. p -Normen

Wir führen hier eine wichtige Klasse von Normen, die sogenannten p -Normen ($1 \leq p \leq \infty$), auf \mathbb{C}^m und für Räume komplexwertiger Funktionen ein.

1.) Maximums- und Supremumsnorm, $p = \infty$:

(a) Sei $x = (x_1, \dots, x_m) \in \mathbb{C}^m$ (oder \mathbb{R}^m), dann setzen wir

$$\|x\|_\infty := \max(|x_1|, \dots, |x_m|) = \max_{1 \leq j \leq m} |x_j|$$

Für $\|\cdot\|_\infty$ auf \mathbb{C}^m sind die Eigenschaften (N1-2) unmittelbar klar, d.h. $\|\lambda x\|_\infty = |\lambda| \cdot \|x\|_\infty$ und $\|x\|_\infty \geq 0$; $\|x\|_\infty = 0 \Leftrightarrow x = 0$. Weiters folgt (N3) direkt aus der Dreiecksungleichung für komplexe Zahlen, indem wir in den Ungleichungen $|x_j + y_j| \leq |x_j| + |y_j|$ ($j = 1, \dots, m$) jeweils zum Maximum übergehen.

Daher gilt: $(\mathbb{R}^m, \|\cdot\|_\infty)$ (bzw. $(\mathbb{C}^m, \|\cdot\|_\infty)$) ist normierter Vektorraum über \mathbb{R} (bzw. \mathbb{C}).

(b) Sei X eine Menge und $f: X \rightarrow \mathbb{C}$ beschränkt, dann setzen wir

$$(15.2) \quad \|f\|_\infty := \sup\{|f(z)| : z \in X\} = \sup_{z \in X} |f(z)| \quad (\text{Supremumsnorm}).$$

Behauptung: Die Supremumsnorm $\|\cdot\|_\infty$ auf den beschränkten Funktionen $\mathcal{B}(X) := \{f: X \rightarrow \mathbb{C} : f \text{ ist beschränkt}\}$ erfüllt die Norm-Axiome (N1-3), mit anderen Worten: $(\mathcal{B}(X), \|\cdot\|_\infty)$ ist ein normierter Vektorraum.

Wir überprüfen die Norm-Axiome:

(N1): $\|f\|_\infty \geq 0$ ist klar; falls $\|f\|_\infty = 0$, dann gilt $|f(x)| = 0$ für alle $x \in X$, also $\forall x \in X: f(x) = 0$ und somit $f = 0$ in $\mathcal{B}(X)$.

(N2): Für $\lambda \in \mathbb{C}, f \in \mathcal{B}(X)$ gilt $\|\lambda f\|_\infty = \sup_{x \in X} |\lambda f(x)| = \sup |\lambda| \cdot |f(x)| = |\lambda| \cdot \|f\|_\infty$.

(N3): $\forall x \in X: \underbrace{|f(x) + g(x)|}_{(f+g)(x)} \leq |f(x)| + |g(x)| \leq \underbrace{\sup_{x \in X} |f(x)| + \sup_{x \in X} |g(x)|}_{\|f\|_\infty + \|g\|_\infty}$, daher folgt

$$\|f + g\|_\infty = \sup_{x \in X} |f(x) + g(x)| \leq \|f\|_\infty + \|g\|_\infty.$$

Somit ist die Behauptung bewiesen.

2.) Die Fälle $1 \leq p < \infty$:

(a) Sei $x = (x_1, \dots, x_m) \in \mathbb{C}^m$ (oder \mathbb{R}^m), dann setzen wir

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

(der Spezialfall $p = 2$ entspricht der euklidischen Norm).

(b) Sei $f \in \mathcal{R}[a, b]$, dann setzen wir

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

(eine Verallgemeinerung von (14.7) — abgesehen vom Normierungsfaktor $1/2\pi$, der dort verwendet worden war).

Für $\|\cdot\|_p$ auf \mathbb{C}^m (bzw. \mathbb{R}^m) sind die Eigenschaften (N1-2) unmittelbar klar, d.h. $\|\lambda x\|_p = |\lambda| \cdot \|x\|_p$ und $\|x\|_p \geq 0$; $\|x\|_p = 0 \Leftrightarrow x = 0$. (N3) folgt aus der Minkowski-Ungleichung (siehe das folgende Lemma).

Für $\|\cdot\|_p$ ($1 \leq p < \infty$) auf \mathbb{R} -integrierbaren Funktionen ist die Eigenschaft (N2) und der erste Teil von (N1) unmittelbar klar, d.h. $\|\lambda f\|_p = |\lambda| \cdot \|f\|_p$, $\|f\|_p \geq 0$; aber $\|f\|_p = 0$ erzwingt nicht $f = 0$.

Allerdings ist auch (N1) gültig, falls wir uns auf die stetigen Funktionen $\mathcal{C}([a, b]) \subseteq \mathcal{R}[a, b]$ beschränken: für f stetig und $f \neq 0$ gibt es x_0 und $\delta > 0$ mit $|f(x)| \geq |f(x_0)|/2 > 0$, wenn $|x - x_0| \leq \delta$; dann folgt $\|f\|_p^p = \int_a^b |f|^p \geq 2\delta |f(x_0)|^p / 2^p > 0$.

(N3) folgt ebenfalls aus dem angekündigten

Lemma

Seien $p, q \in]1, \infty[$ mit $\frac{1}{p} + \frac{1}{q} = 1$ oder $p = 1$ und $q = \infty$.

(i) $\forall x, y \in \mathbb{C}^m$ gilt:

$$\sum_{k=1}^m |x_k y_k| \leq \|x\|_p \cdot \|y\|_q \quad (\text{Hölder-Ungleichung})$$

und

$$\|x + y\|_p \leq \|x\|_p + \|y\|_p \quad (\text{Minkowski-Ungleichung})$$

(ii) $\forall f, g \in \mathcal{R}[a, b]$ gilt:

$$\int_a^b |f(x)g(x)| dx \leq \|f\|_p \cdot \|g\|_q \quad (\text{Hölder-Ungleichung für Integrale})$$

und

$$\|f + g\|_p \leq \|f\|_p + \|g\|_p \quad (\text{Minkowski-Ungleichung für Integrale})$$

Bemerkung: Die Spezialfälle der Hölder-Ungleichungen für $p = q = 2$ ergeben jeweils die entsprechenden Cauchy-Schwarz-Ungleichungen, denn

$$|\langle x | y \rangle| \leq \sum |x_k y_k| \leq \|x\|_2 \cdot \|y\|_2 \quad \text{bzw.} \quad |\langle f | g \rangle| \leq \int |f(x)g(x)| dx \leq \|f\|_2 \cdot \|g\|_2.$$

Beweis

(i) • OBdA ist $x \neq 0$ und $y \neq 0$. Für $p = 1, q = \infty$ ergibt sich die Hölder-Ungleichung direkt wie folgt

$$\sum_{k=1}^m |x_k| |y_k| \leq \sum_{k=1}^m |x_k| \cdot \|y\|_\infty = \|x\|_1 \cdot \|y\|_\infty.$$

• Für $1 < p, q < \infty$ mit $1/p + 1/q = 1$ erinnern wir an die in Analysis 1, §8, bewiesene Ungleichung

$$\forall r, s > 0 : \quad r^{1/p} \cdot s^{1/q} \leq \frac{r}{p} + \frac{s}{q}.$$

Wir setzen $\xi_k := \left(\frac{|x_k|}{\|x\|_p}\right)^p$, $\eta_k := \left(\frac{|y_k|}{\|y\|_q}\right)^q$ ($k = 1, \dots, m$), dann ist $\sum_{k=1}^m \xi_k = \sum_{k=1}^m \eta_k = 1$ und somit

$$\frac{1}{\|x\|_p \|y\|_q} \sum_{k=1}^m |x_k y_k| = \sum_{k=1}^m \xi_k^{1/p} \eta_k^{1/q} \leq \sum_{k=1}^m \left(\frac{\xi_k}{p} + \frac{\eta_k}{q}\right) = \frac{1}{p} \sum_{k=1}^m \xi_k + \frac{1}{q} \sum_{k=1}^m \eta_k = \frac{1}{p} + \frac{1}{q} = 1.$$

• Für $p = 1$ folgt die Minkowski-Ungleichung direkt aus der Dreiecksungleichung für komplexe Zahlen:

$$\sum_{k=1}^m |x_k + y_k| \leq \sum_{k=1}^m (|x_k| + |y_k|) = \|x\|_1 + \|y\|_1.$$

• Sei $p > 1$: Wir setzen $z_k := |x_k + y_k|^{p-1}$ ($k = 1, \dots, m$), dann gilt mit $q \geq 1, 1/p + 1/q = 1$

$$z_k^q = |x_k + y_k|^{q \cdot (p-1)} = |x_k + y_k|^{\frac{p}{p-1} \cdot (p-1)} = |x_k + y_k|^p$$

und weiter $\|z\|_q = \left(\sum_k z_k^q\right)^{1/q} = \|x + y\|_p^{p/q}$. Daraus erhalten wir mit Hilfe der Dreiecksungleichung für komplexe Zahlen und der eben bewiesenen Hölder-Ungleichung

$$\begin{aligned} \|x + y\|_p^p &= \sum_k |x_k + y_k|^p = \sum_k |x_k + y_k| \cdot |z_k| \leq \sum_k |x_k z_k| + \sum_k |y_k z_k| \\ &\leq (\|x\|_p + \|y\|_p) \cdot \|z\|_q = (\|x\|_p + \|y\|_p) \cdot \|x + y\|_p^{p/q}. \end{aligned}$$

Wegen $p - \frac{p}{q} = 1$ folgt also $\|x + y\|_p \leq \|x\|_p + \|y\|_p$.

(ii) In §9 der Analysis 1 haben wir gezeigt, dass mit f und g auch $|f \cdot g|$, $|f|^p$, $|g|^p$ \mathbb{R} -integrierbar (auf $[a, b]$) sind.

Die Hölder- und die Minkowski-Ungleichung gelten für die entsprechenden Riemann-Summen, überträgt sich im Limes daher auch auf die Integrale.

□

Somit erhalten wir zusammenfassend die

Proposition: Es sei $1 \leq p \leq \infty$, dann sind $(\mathbb{R}^m, \|\cdot\|_p)$, $(\mathbb{C}^m, \|\cdot\|_p)$ und $(\mathcal{C}([a, b]), \|\cdot\|_p)$ normierte Vektorräume (über \mathbb{R} bzw. \mathbb{C}).

15.7. Definition

Es sei (X, d) ein metrischer Raum.

1.) Sei $a \in X$, $r > 0$, dann heißt

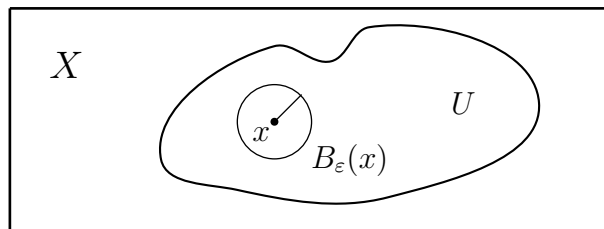
$$(15.3) \quad B_r(x) := \{x \in X : d(a, x) < r\}$$

die (*offene*) Kugel¹ mit Mittelpunkt a und Radius r (bzgl. der Metrik d auf X)

2.) Sei $U \subseteq X$, $x \in U$. Dann ist U eine *Umgebung* von x , falls gilt:

$$\exists \varepsilon > 0: B_\varepsilon(x) \subseteq U$$

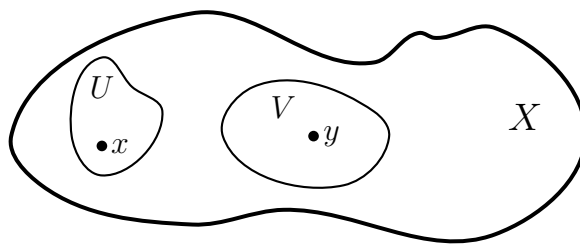
Insbesondere ist $B_\varepsilon(x)$ stets eine Umgebung von x , die so genannte ε -Umgebung von x (bzgl. d).



¹Für Kugeln wird in der Mathematik oftmals der Buchstabe B benutzt, weil in vielen Sprachen das Wort für Kugel mit b beginnt; hier einige Beispiele: englisch ball, spanisch bola, französisch boule.

15.8. Proposition (Hausdorff-Trennungseigenschaft)²

Sei (X, d) ein metrischer Raum, $x, y \in X$ mit $x \neq y$. Dann gibt es Umgebungen U von x und V von y , die disjunkt sind, d. h. $U \cap V = \emptyset$.



Beweis:

Setze $\varepsilon := \frac{d(x,y)}{3} > 0$, $U := B_\varepsilon(x)$, $V := B_\varepsilon(y)$. Dann ist U Umgebung von x und V Umgebung von y .

Weiters ist $U \cap V = \emptyset$, denn

$$\begin{aligned} z \in U \cap V &\Rightarrow d(x, z) < \varepsilon \text{ und } d(z, y) < \varepsilon \\ &\Rightarrow 3\varepsilon = d(x, y) \leq d(x, z) + d(z, y) < 2\varepsilon \quad \text{ein Widerspruch } \zeta. \end{aligned}$$

□

15.9. Definition

Sei (X, d) ein metrischer Raum und $U \subseteq X$. U heißt *offen*, falls $\forall x \in U$ gilt: U ist Umgebung von x . (D.h. U ist Umgebung jedes ihrer Punkte.)

Eine äquivalente Bedingung ist: $\forall x \in U: \exists \varepsilon > 0: B_\varepsilon(x) \subseteq U$.

15.10. Beispiele

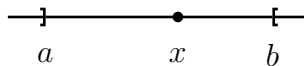
- 1.) Sei (X, d) ein beliebiger metrischer Raum, $a \in X$, $r > 0$. Dann ist $U := B_r(a)$ ist offen:

Sei $x \in B_r(a)$, $\varepsilon := r - d(x, a) > 0$, dann gilt für jedes $y \in B_\varepsilon(x)$:

$$d(y, a) \leq d(y, x) + d(x, a) < \varepsilon + d(x, a) = r, \text{ d. h. } y \in B_r(a); \text{ also } B_\varepsilon(x) \subseteq B_r(a).$$

- 2.) Sei $a < b$, dann ist $]a, b[$ ist offen in \mathbb{R} :

Sei $x \in]a, b[$; setze $\varepsilon := \min(x - a, b - x) > 0$, dann ist $B_\varepsilon(x) =]x - \varepsilon, x + \varepsilon[\subseteq]a, b[$.



Ebenso sieht man, dass $]a, \infty[$ und $] -\infty, a[$ offen sind.

Intervalle der Form $[a, b[$ oder $[a, b]$ sind nicht offen (siehe Übungsaufgaben).

²Felix Hausdorff (*8.11. 1868 Breslau; †26.1. 1942 Bonn), deutscher Mathematiker

15.11. Theorem

Sei (X, d) ein metrischer Raum. Dann gilt:

- 1.) X und \emptyset sind offen
- 2.) $U, V \subseteq X$ offen $\Rightarrow U \cap V$ offen (allgemeiner: endliche Durchschnitte)
- 3.) Sei I eine Menge. Falls $U_i \subseteq X$ offen für alle $i \in I$, dann ist $\bigcup_{i \in I} U_i$ offen.

Beweis:

- 1.) $\forall x \in X, \forall \varepsilon > 0: B_\varepsilon(x) \subseteq X$, also ist X Umgebung jedes Punktes $x \in X$.
Für \emptyset ist die Umgebungsbedingung trivialerweise erfüllt.
- 2.) Sei $x \in U \cap V$.
 U ist offen, daher $\exists \varepsilon_1 > 0: B_{\varepsilon_1}(x) \subseteq U$;
 V ist offen, daher $\exists \varepsilon_2 > 0: B_{\varepsilon_2}(x) \subseteq V$.
Für $\varepsilon := \min(\varepsilon_1, \varepsilon_2)$ ist dann $B_\varepsilon(x) \subseteq U \cap V$.
- 3.) Sei $x \in \bigcup_{i \in I} U_i$. Dann existiert ein $j \in I: x \in U_j$.
 U_j ist offen, daher existiert $\varepsilon > 0: B_\varepsilon(x) \subseteq U_j \Rightarrow B_\varepsilon(x) \subseteq \bigcup_{i \in I} U_i$.

□

15.12. Bemerkung

Eigenschaft 2.) gilt nur für endliche, nicht aber für unendliche Durchschnitte.

Z. B. betrachte in $X = \mathbb{R}$ die Folge offener Intervalle $U_n =] - \frac{1}{n}, 1 + \frac{1}{n} [$ (für $n \in \mathbb{N}, n \geq 1$).
Dann ist $\bigcap_{n \in \mathbb{N}} U_n = [0, 1]$, also nicht offen.

15.13. Definition

Sei (X, d) ein metrischer Raum und $A \subseteq X$. A heißt *abgeschlossen*, wenn $X \setminus A$ offen ist.

Folgende Eigenschaften erhalten wir direkt aus Theorem 15.11:

- 1.) \emptyset und X sind abgeschlossen
(diese Mengen haben beide Eigenschaften: offen und abgeschlossen.)

2.) $A, B \subseteq X$ abgeschlossen $\Rightarrow A \cup B$ abgeschlossen (allg.: endliche Vereinigungen)

3.) Ist $\forall i \in I$ die Menge $A_i \subseteq X$ abgeschlossen, dann ist auch $\bigcap_{i \in I} A_i$ abgeschlossen.

15.14. Beispiele

1.) Sei $a < b$, dann ist $[a, b] \subseteq \mathbb{R}$ abgeschlossen, weil

$$[a, b] = (\mathbb{R} \setminus]-\infty, a[) \cap (\mathbb{R} \setminus]b, \infty[)$$

ein Durchschnitt abgeschlossener Mengen ist.

Hingegen ist $[a, b[$ nicht abgeschlossen (und auch nicht offen).

2.) $A_1 \subseteq \mathbb{R}^k, A_2 \subseteq \mathbb{R}^m$ seien abgeschlossen (jeweils bzgl. der euklidischen Metrik).

Behauptung: $A_1 \times A_2 \subseteq \mathbb{R}^k \times \mathbb{R}^m \cong \mathbb{R}^{k+m}$ ist abgeschlossen.

(Allg.: endliche kartesische Produkte abgeschlossener Mengen sind abgeschlossen.)

Beweis:

Wir zeigen, dass $U := (\mathbb{R}^k \times \mathbb{R}^m) \setminus (A_1 \times A_2)$ offen ist.

Sei $(x, y) \in U$, dann gilt $x \in \mathbb{R}^k \setminus A_1$ oder $y \in \mathbb{R}^m \setminus A_2$.

Angenommen $x \in \mathbb{R}^k \setminus A_1$ (der Fall $y \in \mathbb{R}^m \setminus A_2$ ist dann analog).

$\exists \varepsilon > 0: B_\varepsilon(x) \subseteq \mathbb{R}^k \setminus A_1$ und daher ist auch $B_\varepsilon((x, y)) \subseteq (\mathbb{R}^k \times \mathbb{R}^m) \setminus (A_1 \times A_2)$, denn

$$\begin{aligned} (x', y') \in B_\varepsilon((x, y)) &\Rightarrow \varepsilon^2 > \|(x, y) - (x', y')\|^2 \\ &= \|(x, 0) - (x', 0)\|^2 + \|(0, y) - (0, y')\|^2 \geq \|x - x'\|^2 \Rightarrow x' \notin A_1 \end{aligned}$$

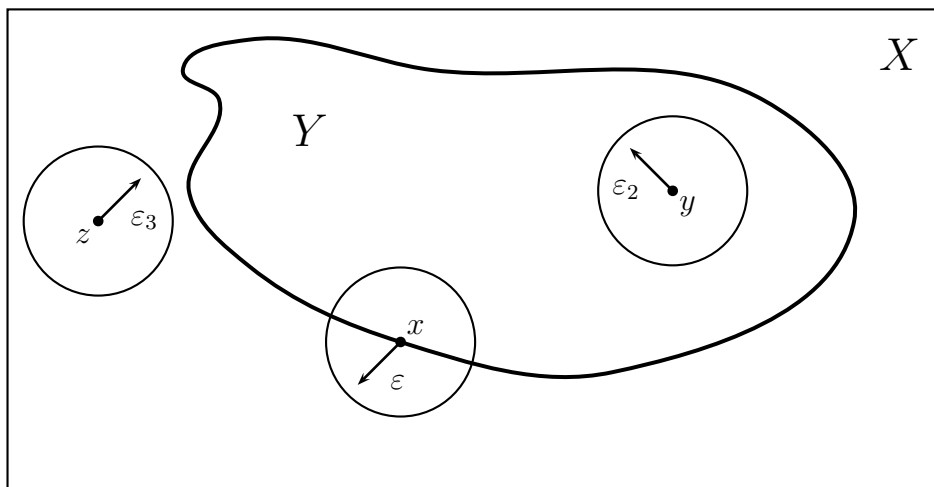
□

Spezialfall: Seien $a_j \leq b_j$ ($j = 1, \dots, n$), dann ist der Quader

$$(15.4) \quad Q := \{(x_1, \dots, x_n) \in \mathbb{R}^n : a_j \leq x_j \leq b_j \ (j = 1, \dots, n)\} = \prod_{j=1}^n [a_j, b_j]$$

abgeschlossen in \mathbb{R}^n .

Lage von Punkten bzgl. einer Teilmenge Y im metrischen Raum X :



y liegt innerhalb von Y : \exists „Schutzkugel“ $B_{\varepsilon_2}(y) \subseteq Y$

z liegt außerhalb von Y : \exists „Schutzkugel“ $B_{\varepsilon_3}(z) \subseteq X \setminus Y$

x liegt weder außen noch innen bzgl. $Y \rightsquigarrow$

15.15. Definition

Sei (X, d) ein metrischer Raum und $Y \subseteq X$. Ein Punkt $x \in X$ heißt *Randpunkt* von Y , wenn gilt:

$$\forall \varepsilon > 0: B_\varepsilon(x) \cap Y \neq \emptyset \text{ und } B_\varepsilon(x) \cap (X \setminus Y) \neq \emptyset.$$

(Äquivalent: Jede Umgebung von x enthält sowohl Punkte aus Y als auch aus $X \setminus Y$.)

Die Menge $\partial Y := \{x \in X: x \text{ ist Randpunkt von } Y\}$ heißt der *Rand* von Y .

15.16. Beispiele

1.) Es sei $a < b$ und $I \subseteq \mathbb{R}$ das Intervall $[a, b]$ oder $[a, b[$ oder $]a, b]$ oder $]a, b[$.

In allen Fällen gilt $\partial I = \{a, b\}$.

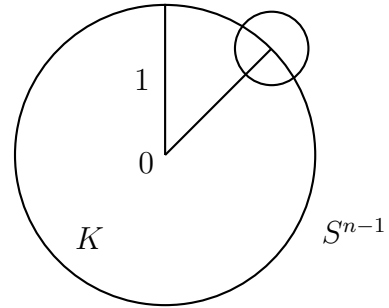
2.) In $X = \mathbb{R}^n$ sei $K := \{x \in \mathbb{R}^n: \|x\| \leq 1\}$. Dann ist $\partial K = \{x \in \mathbb{R}^n: \|x\| = 1\} =: S^{n-1}$, also die (Einheits-)Sphäre.³

Ebenso gilt für die offene Kugel $B_1(0) = \{x \in \mathbb{R}^n: \|x\| < 1\}$, dass $\partial B_1(0) = S^{n-1}$.

³Die $(n-1)$ -Sphäre bildet also die „Oberfläche“ der n -dimensionalen Einheits-Kugel um den Ursprung; der Name kommt von dem griechischen Wort ἡ σφαιρα (Kugel).

• $\varepsilon > 0, y \in S^{n-1} \Rightarrow (1 + \frac{\varepsilon}{2}) \cdot y \in B_\varepsilon(y) \cap (\mathbb{R}^n \setminus K)$
 und der Punkt $(1 - \frac{\varepsilon}{2}) \cdot y$ (falls $\varepsilon < 4$) oder 0 (falls $\varepsilon \geq 4$) gehört jedenfalls zu $K \cap B_\varepsilon(y)$, also gilt $S^{n-1} \subseteq \partial K$.

• $x \notin S^{n-1} \Rightarrow \|x\| > 1$ oder $\|x\| < 1$, also ist $B_\varepsilon(x)$ ganz innerhalb oder außerhalb von K , falls $\varepsilon := |1 - \|x\||/2$ gesetzt wird; daher $\partial K \subseteq S^{n-1}$.



3.) Für $Q \subseteq \mathbb{R}$ ist $\partial Q = \mathbb{R}$ (Q ist also eine echte Teilmenge des eigenen Randes).

[In Analysis 1 hieß es ‘ Q und $\mathbb{R} \setminus Q$ liegen dicht in \mathbb{R} ’ in folgender Bedeutung:
 $\forall x \in \mathbb{R} \forall \varepsilon > 0: B_\varepsilon(x) \cap Q \neq \emptyset$ und $B_\varepsilon(x) \cap (\mathbb{R} \setminus Q) \neq \emptyset$.]

15.17. Proposition

Sei (X, d) ein metrischer Raum, $Y \subseteq X$ Dann gilt:

- 1.) $\overset{\circ}{Y} := Y \setminus \partial Y$ ist offen; $\overset{\circ}{Y}$ heißt das *Innere* von Y (auch *offener Kern* von Y)
- 2.) $\bar{Y} := Y \cup \partial Y$ ist abgeschlossen; \bar{Y} heißt *Abschluss* von Y (auch *abgeschlossene Hülle* von Y)
- 3.) Der Rand ∂Y ist abgeschlossen.
- 4.) $\partial(X \setminus Y) = \partial Y$ (d.h. Y und sein Komplement [in X] haben denselben Rand)

Beweis:

1.) $y \in Y \setminus \partial Y$, d. h. $y \in Y$ und $y \notin \partial Y$

$$\Rightarrow \exists \varepsilon > 0: \underbrace{B_\varepsilon(y) \cap Y = \emptyset}_{\text{unmöglich, weil } y \in Y \cap B_\varepsilon(y)} \quad \text{oder} \quad B_\varepsilon(y) \cap (X \setminus Y) = \emptyset$$

$$\Rightarrow B_\varepsilon(y) \cap (X \setminus Y) = \emptyset, \text{ d. h. } B_\varepsilon(y) \subseteq Y;$$

noch zu zeigen: $B_\varepsilon(y) \cap \partial Y = \emptyset$ [dann fertig, weil $B_\varepsilon(y) \subseteq Y \setminus \partial Y$ folgt]

Indirekt: $z \in B_\varepsilon(y) \cap \partial Y \Rightarrow$ für alle η mit $0 < \eta < \varepsilon - d(y, z)$ gilt (weil z Randpunkt ist und wegen $B_\eta(z) \subseteq B_\varepsilon(y)$)

$$\emptyset \neq (X \setminus Y) \cap B_\eta(z) \subseteq (X \setminus Y) \cap B_\varepsilon(y) = \emptyset, \text{ ein Widerspruch } \zeta.$$

- 4.) $z \in \partial(X \setminus Y) \iff \forall \varepsilon > 0: B_\varepsilon(z) \cap (X \setminus Y) \neq \emptyset$ und $B_\varepsilon(z) \cap Y \neq \emptyset$
 $\iff z \in \partial Y$

$$\begin{aligned}
2.) \quad Y \cup \partial Y &\stackrel{[4.)]}{=} Y \cup \partial(X \setminus Y) = (X \setminus (X \setminus Y)) \cup \partial(X \setminus Y) = \\
&= X \setminus [(X \setminus Y) \setminus \partial(X \setminus Y)] = X \setminus \underbrace{(X \setminus Y)^\circ}_{[1.) \text{ offen}} \text{ ist abgeschlossen}
\end{aligned}$$

$$\begin{aligned}
3.) \quad \partial Y &= (Y \cup \partial Y) \setminus (Y \setminus \partial Y) = \bar{Y} \setminus \overset{\circ}{Y} \\
&\text{und somit ist } X \setminus \partial Y = (X \setminus \bar{Y}) \cup \overset{\circ}{Y} \text{ offen [weil } \bar{Y} \text{ abgeschlossen gem\u00e4\u00df 2.)]} \\
&\Rightarrow \partial Y \text{ abgeschlossen}
\end{aligned}$$

□

§16. Konvergenz und Stetigkeit

16.1. Definition

Sei (X, d) ein metrischer Raum.

- 1.) Eine *Folge* $(x_k)_{k \in \mathbb{N}}$ in X ist eine Abbildung $x: \mathbb{N} \rightarrow X$, wobei $x_k := x(k)$ ($k \in \mathbb{N}$) gesetzt wird.
- 2.) Sei (x_k) eine Folge in X , $a \in X$. Dann heißt (x_k) *konvergent* gegen a , $a = \lim_{k \rightarrow \infty} x_k$, oder auch $x_k \rightarrow a$ ($k \rightarrow \infty$), wenn gilt

$$(16.1) \quad \forall \varepsilon > 0 \exists N \in \mathbb{N} \forall k \geq N : x_k \in B_\varepsilon(a).$$

Äquivalent dazu ist die Bedingung

$$(16.1)' \quad \text{Für jede Umgebung } U \text{ von } a \text{ gilt: } \exists N \in \mathbb{N} \forall k \geq N : x_k \in U$$

oder auch:

$$(16.1)'' \quad \forall \varepsilon > 0 \exists N \in \mathbb{N} \forall k \geq N : d(x_k, a) < \varepsilon.$$

16.2. Definition

Sei (X, d) metrischer Raum, $A \subseteq X$; dann heißt

$$(16.2) \quad \text{diam}(A) := \sup\{d(x, y) : x, y \in A\} \in [0, \infty]$$

der *Durchmesser* von A . (griech.: ῆ διάμετρος).

Falls $\text{diam}(A) < \infty$, so heißt A *beschränkt* (bzgl d). Eine Folge (x_k) in X heißt *beschränkt*, wenn $A := \{x_k : k \in \mathbb{N}\} \subseteq X$ beschränkt ist.

16.3. Proposition

Es sei (X, d) ein metrischer Raum.

- 1.) $r > 0, a \in X \Rightarrow \text{diam } B_r(a) \leq 2r$, insbesondere ist $B_r(a)$ beschränkt
[Bem: für die diskrete Metrik und $r < 1$ ist stets $B_r(x) = \{x\}$ und $\text{diam } B_r(x) = 0$, also echt kleiner als $2r$.]
- 2.) Für $A \subseteq X$ gilt: A beschränkt $\iff \exists r > 0 \exists z \in X : A \subseteq B_r(z)$
- 3.) Jede konvergente Folge in X ist beschränkt.

Beweis:

- 1.) $x, y \in B_r(a) \Rightarrow d(x, y) \leq d(x, a) + d(a, y) \leq 2r$
- 2.) \Leftrightarrow klar, weil $\text{diam } A \leq \text{diam } B_r(z) \leq 2r < \infty$
 \Leftrightarrow • $A = \emptyset \Rightarrow \forall r > 0 \forall z \in X: A \subseteq B_r(z)$
• $A \neq \emptyset$: sei $z \in A$ und $r := 1 + \text{diam}(A)$;
 $x \in A \Rightarrow d(x, z) \leq \text{diam}(A) < r \Rightarrow x \in B_r(z)$
- 3.) Sei $a = \lim(x_k)$; $\varepsilon = 1$ in (16.1) liefert: $\exists N \in \mathbb{N} \forall k \geq N: x_k \in B_1(a)$;
sei $r_0 := \max_{1 \leq j \leq N} d(x_j, a)$, dann gilt $\forall k \in \mathbb{N}: x_k \in B_{\max(r_0, 1)}(a)$.

□

16.4. Theorem

Sei (X, d) ein metrischer Raum und $A \subseteq X$. Dann besteht folgende Äquivalenz:

A ist abgeschlossen \iff Für jede Folge (x_k) in A gilt: falls (x_k) konvergent ist und $x := \lim(x_k)$ (in X), dann folgt $x \in A$.

Beweis:

\Rightarrow Sei (x_k) eine Folge in A , die konvergiert und $x = \lim(x_k) \in X$; z. z.: $x \in A$.

Indirekt: $x \notin A$, d.h. x ist Element der offenen Menge $X \setminus A$

$\Rightarrow \exists \varepsilon > 0: B_\varepsilon(x) \subseteq X \setminus A$, d. h. $B_\varepsilon(x) \cap A = \emptyset$.

Andererseits gilt wegen $x_k \rightarrow x: \exists N: \forall k \geq N: x_k \in B_\varepsilon(x)$, ein Widerspruch \nexists dazu, dass $x_k \in A$.

\Leftarrow z. z.: $X \setminus A$ ist offen

Indirekt: angenommen $\exists x \in X \setminus A$ so, dass $\forall \varepsilon > 0$ gilt $B_\varepsilon(x) \cap A \neq \emptyset$;

zu $\varepsilon = \frac{1}{k}$ ($k \geq 1$) wähle $x_k \in A \cap B_{\frac{1}{k}}(x)$

$\Rightarrow (x_k)$ ist eine Folge in A und $x_k \rightarrow x$, also $x \in A$, ein Widerspruch \nexists

□

16.5. Cauchy-Folgen und Vollständigkeit

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

1.) Eine Folge (x_k) in X heißt *Cauchy-Folge*, falls gilt

$$(16.3) \quad \forall \varepsilon > 0 \exists N \in \mathbb{N} \forall k, m \geq N: \quad d(x_k, x_m) < \varepsilon.$$

2.) (X, d) heißt *vollständig*, wenn jede Cauchy-Folge in X konvergiert.

3.) Ist $(V, \|\cdot\|)$ ein normierter Vektorraum und als metrischer Raum vollständig (bzgl. der Metrik $d(x, y) := \|x - y\|$), so heißt $(V, \|\cdot\|)$ *Banach-Raum*.¹

Beispiel: 1.) aus Analysis 1 wissen wir, dass \mathbb{R} (mit der euklidischen Metrik) vollständig ist.

2.) $X = \mathbb{Q}$ mit $d(x, y) = |x - y|$ für alle $x, y \in \mathbb{Q}$ ist nicht vollständig; sei z. B. $x_k :=$ Dezimalbruchentwicklung von $\sqrt{2}$ bis zur k -ten Stelle

$\Rightarrow (x_k)$ ist Cauchy-Folge in \mathbb{Q} und $x_k \rightarrow \sqrt{2} \in \mathbb{R} \setminus \mathbb{Q}$, d. h. (x_k) ist nicht konvergent in \mathbb{Q} .

Proposition: Wenn (x_k) eine konvergente Folge im metrischen Raum (X, d) ist, dann ist (x_k) eine Cauchy-Folge.

Beweis: Sei $a = \lim(x_k)$ und $\varepsilon > 0$; wähle $N \in \mathbb{N}: \forall k \geq N: d(x_k, a) < \frac{\varepsilon}{2}$,

dann folgt $\forall k, m \geq N: \quad d(x_k, x_m) \leq d(x_k, a) + d(a, x_m) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$ □

16.6. Theorem

Wir betrachten nun \mathbb{R}^n mit der euklidischen Norm $\|x\| = \sqrt{\langle x | x \rangle}$.

1.) Sei $(x_k)_{k \in \mathbb{N}}$ eine Folge in \mathbb{R}^n mit $x_k = (x_{k1}, x_{k2}, \dots, x_{kn})$ für $k \in \mathbb{N}$, weiters sei $a = (a_1, \dots, a_n) \in \mathbb{R}^n$; dann gilt

$$x_k \rightarrow a \quad (k \rightarrow \infty) \text{ in } \mathbb{R}^n \quad \iff \quad j = 1, \dots, n: \quad a_j = \lim_{k \rightarrow \infty} x_{kj} \quad (\text{in } \mathbb{R})$$

2.) $\mathbb{R}^n, \|\cdot\|$ ist ein Banach-Raum.

¹Stefan Banach (*30. 3. 1892 Krakau; †31. 8. 1945 Lemberg) [ˈstefan ˈbanax], polnischer Mathematiker

Beweis:

1.) \Leftrightarrow für alle j gilt: $|x_{kj} - a_j| \leq \|x_k - a\| \rightarrow 0 \quad (k \rightarrow \infty)$

\Leftrightarrow Sei $\varepsilon > 0$; zu jedem $j \in \{1, \dots, n\} \exists N_j \forall k \geq N_j: |a_j - x_{kj}| < \frac{\varepsilon}{\sqrt{n}}$;

sei $N = \max_{1 \leq j \leq n} N_j$ und $k \geq N$, dann ist

$$\|x_k - a\|^2 = \sum_{j=1}^n |x_{kj} - a_j|^2 < n \cdot \frac{\varepsilon^2}{n} = \varepsilon^2.$$

Daher also $\forall k \geq N: d(x_k, a) = \|x_k - a\| < \varepsilon$.

2.) (x_k) Cauchy-Folge in $\mathbb{R}^n \implies \forall j \in \{1, \dots, n\}: |x_{kj} - x_{mj}| \leq \|x_k - x_m\|$

also ist für $j = 1, \dots, n$ jeweils $(x_{kj})_{k \in \mathbb{N}}$ eine Cauchy-Folge in \mathbb{R} .

Daher gilt $\forall j \in \{1, \dots, n\}: \exists a_j := \lim_{k \rightarrow \infty} x_{kj} \in \mathbb{R}$; setze $a = (a_1, \dots, a_n)$, dann folgt nach 1.), dass $x \rightarrow a \quad (k \rightarrow \infty)$.

□

16.7. Stetigkeit:

Definition: Es seien (X, d_1) und (Y, d_2) metrische Räume und $f: X \rightarrow Y$. Die Abbildung f heißt *stetig* im Punkt $a \in X$, falls gilt:

$$\forall \varepsilon > 0 \exists \delta > 0 \forall x \in X: d_1(x, a) < \delta \implies d_2(f(x), f(a)) < \varepsilon.$$

f heißt stetig auf X , wenn f stetig in jedem Punkt $a \in X$ ist.

Wir nennen f einen *Homöomorphismus*, falls f stetig und bijektiv ist und die Umkehrabbildung $f^{-1}: Y \rightarrow X$ ebenfalls stetig ist. X und Y heißen dann *zueinander homöomorph*².

Theorem: Es seien $(X, d_1), (Y, d_2)$ metrische Räume, $f: X \rightarrow Y$ und $a \in X$. Dann sind folgende Aussagen äquivalent:

1.) f ist stetig in a .

2.) Für jede Folge (x_k) in X mit $a = \lim x_k$ (in (X, d_1)) ist $f(a) = \lim f(x_k)$ (in (Y, d_2)).
Man schreibt dann auch $f(a) = \lim_{x \rightarrow a} f(x)$

²Von griechisch $\acute{\omicron}\mu\acute{\omicron}\acute{\omicron}\mu\omicron\rho\phi\omicron\varsigma$, wörtl. „von gleicher Gestalt“.

Beweis:

1.) \Rightarrow 2.): Sei (x_n) eine Folge in X mit $x_n \rightarrow a$.

Zu $\varepsilon > 0$ existiert $\delta > 0$, sodass $\forall x \in X$ mit $d_1(x, a) < \delta$ gilt $d_2(f(x), f(a)) < \varepsilon$.

Wähle $N \in \mathbb{N}$: $d_1(x_n, a) < \delta$ für alle $n \geq N$; daher folgt dann $\forall n \geq N$, dass $d_2(f(x_n), f(a)) < \varepsilon$ gilt; somit $f(x_n) \rightarrow f(a)$.

2.) \Rightarrow 1.): Es ist $\lim_{x \rightarrow a} f(x) = f(a)$;

indirekt: f nicht stetig in $a \Rightarrow$

$$\exists \varepsilon > 0 \forall \delta > 0 \exists x \in X: \quad d_1(x, a) < \delta \quad \text{und} \quad d_2(f(x), f(a)) \geq \varepsilon.$$

Zu $\delta = \frac{1}{n}$ ($n \geq 1$) gibt es $x_n \in X$ mit $d_1(x_n, a) < \frac{1}{n}$ und $d_2(f(x_n), f(a)) \geq \varepsilon$;

nun ist (x_n) eine Folge in X mit $x_n \rightarrow a$ ($n \rightarrow \infty$), aber $f(x_n) \not\rightarrow f(a)$, ein Widerspruch ζ

□

16.8. Beispiele

1.) add: $\mathbb{R}^2 \rightarrow \mathbb{R}, (x, y) \mapsto x + y$ ist stetig.

$$[\lim(x_k, y_k) = (x, y) \Rightarrow \lim(x_k + y_k) = x + y = \text{add}(x, y)]$$

2.) mult: $\mathbb{R}^2 \rightarrow \mathbb{R}, (x, y) \mapsto x \cdot y$ ist stetig.

3.) quot: $\mathbb{R} \times (\mathbb{R} \setminus \{0\}) \rightarrow \mathbb{R}, (x, y) \mapsto x/y$ ist stetig.

16.9. Proposition

Seien X, Y und Z metrische Räume und $f: X \rightarrow Y, g: Y \rightarrow Z$ Abbildungen.

Sei f stetig in $a \in X$ und g stetig in $b := f(a) \in Y$. Dann ist $g \circ f: X \rightarrow Z$ stetig in a .

Beweis: Sei $a = \lim x_k \in X$, dann ist $b = f(a) = \lim f(x_k) \in Y$ und

$$(g \circ f)(a) = g(b) = \lim g(f(x_k)) = \lim (g \circ f)(x_k),$$

also $g \circ f$ ist stetig in a .

□

16.10. Stetige Abbildungen $X \rightarrow \mathbb{R}^n$

Sei $f: X \rightarrow \mathbb{R}^n$, dann ist für jedes $x \in X$ der Funktionswert $f(x) \in \mathbb{R}^n$ gegeben durch ein n -Tupel $f(x) = (f_1(x), \dots, f_n(x))$, wobei $f_1, \dots, f_n: X \rightarrow \mathbb{R}$ die Komponentenfunktionen von f sind.

[Bemerkung: Bezeichne für $1 \leq j \leq n$ jeweils $p_j: \mathbb{R}^n \rightarrow \mathbb{R}$, $(x_1, \dots, x_n) \mapsto x_j$, die j -te (Koordinaten-)Projektion, dann ist $f_j = p_j \circ f$.]

Proposition: Sei X ein metrischer Raum und $f: X \rightarrow \mathbb{R}^n$, dann gilt:

$$f \text{ stetig} \iff \forall j \in \{1, \dots, n\}: f_j: X \rightarrow \mathbb{R} \text{ stetig}$$

Beweis: Sei $X \ni a = \lim x_k$, $b := f(a) \in \mathbb{R}^n$, dann gilt nach Theorem 16.6, 1.):

$$b = f(a) = \lim f(x_k) \iff \forall j: b_j = f_j(a) = \lim_{k \rightarrow \infty} f_j(x_k) \quad \square$$

Korollar: Sei X ein metrischer Raum und $f, g: X \rightarrow \mathbb{R}$ stetig, dann sind auch die Abbildungen $f + g: X \rightarrow \mathbb{R}$, $f \cdot g: X \rightarrow \mathbb{R}$ stetig.

Ist zusätzlich $g(x) \neq 0 \forall x \in X$, dann ist $\frac{f}{g}: X \rightarrow \mathbb{R}$ ebenfalls stetig.

Beweis: $(f, g): X \rightarrow \mathbb{R}^2$, $x \mapsto (f(x), g(x))$ ist stetig nach obiger Proposition; daher sind auch $f + g = \text{add} \circ (f, g)$, $f \cdot g = \text{mult} \circ (f, g)$, $\frac{f}{g} = \text{quot} \circ (f, g)$ also Zusammensetzungen stetiger Abbildungen stetig (siehe 16.8 und 16.9). \square

16.11. Beispiel (Polynomfunktionen auf \mathbb{R}^n)

Es sei $F: \mathbb{R}^n \rightarrow \mathbb{R}$ eine Polynomfunktion in n (reellen) Variablen vom Grad $\leq m$, also gegeben durch

$$F(x_1, \dots, x_n) = \sum_{\substack{k_1, \dots, k_n \in \mathbb{N} \\ k_1 + k_2 + \dots + k_n \leq m}} c_{k_1 \dots k_n} \cdot \underbrace{x_1^{k_1} \cdot x_2^{k_2} \cdot \dots \cdot x_n^{k_n}}_{\text{Monome vom Grad } \leq m},$$

wobei $c_{k_1 \dots k_n} \in \mathbb{R}$.

Zum Beispiel beschreibt der Ausdruck $7x_1^3 + 3x_1x_3^4 - 4x_2^5x_3^2$ eine Polynomfunktion vom Grad 7 auf \mathbb{R}^3 (d.h. in 3 Variablen).

F ist stetig, weil es als Summe von Produkten stetiger Funktionen entsteht ($x \mapsto \text{const}$ und $x \mapsto x_j$ sind stetig).

16.12. Beispiel (Lineare Abbildungen $\mathbb{R}^n \rightarrow \mathbb{R}^m$)

Eine lineare Abbildung $L: \mathbb{R}^n \rightarrow \mathbb{R}^m$ sei gegeben durch die Matrix

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \text{ (bzgl. der Standardbasen), d. h.}$$

$$Lx = \left(\sum_{l=1}^n a_{jl} x_l \right)_{j=1}^m =: \left(L_j(x) \right)_{j=1}^m.$$

Die linearen Abbildungen $L_j: \mathbb{R}^n \rightarrow \mathbb{R}$ sind Spezialfälle von 16.11, nämlich vom Grad ≤ 1 . Also ist jedes L_j stetig ($j = 1, \dots, m$), daher auch L stetig.

16.13. Beispiele

1.) Sei (X, d) ein metrischer Raum, $x_0 \in X$, dann ist $f: X \rightarrow \mathbb{R}$, $f(x) := d(x, x_0)$ stetig.

Sei $a \in X$ und $\varepsilon > 0$. Es ist

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

Mini-Lemma: $|d(y, \xi) - d(z, \xi)| \leq d(y, z)$

„Beweisschen“: kombiniere $d(y, \xi) \leq d(y, z) + d(z, \xi)$ und $d(z, \xi) \leq d(z, y) + d(y, \xi)$ \square

Daher erhalten wir $(\star) \leq d(x, a)$, somit ist mit $\delta := \varepsilon$ die Stetigkeitsbedingung erfüllt.

(Ebenso einfach: $x_n \rightarrow a \Rightarrow d(x_n, a) \rightarrow 0 \Rightarrow |f(x_n) - f(a)| \rightarrow 0$.)

Bemerkung: ähnlich sieht man auch, dass $d: X \times X \rightarrow \mathbb{R}$ stetig ist; denn durch Einschließen geeigneter Terme erhalten wir aus dem Mini-Lemma auch

$$|d(x, y) - d(x_0, y_0)| \leq |d(x, y) - d(x_0, y)| + |d(x_0, y) - d(x_0, y_0)| \leq d(x, x_0) + d(y, y_0) \rightarrow 0,$$

falls $x \rightarrow x_0$ und $y \rightarrow y_0$.

2.) Lineare Abbildung zwischen normierten Räumen:

Seien $(V, \|\cdot\|_1)$, $(W, \|\cdot\|_2)$ normierte Vektorräume und $L: V \rightarrow W$ linear. Wir definieren die *Operatornorm* von L durch

$$(16.4) \quad \|L\|_{\text{op}} := \sup \{ \|Lx\|_2 : x \in V \text{ mit } \|x\|_1 \leq 1 \}.$$

Behauptung: L stetig $\iff \|L\|_{\text{op}} < \infty$

In diesem Fall gilt auch

$$(16.5) \quad \forall x \in V: \quad \|Lx\|_2 \leq \|L\|_{\text{op}} \cdot \|x\|_1$$

Beweis:

⊖ Es ist $\|L0\|_2 = \|0\|_2 = 0$; falls $x \neq 0$, dann gilt

$$\|Lx\|_2 = \|x\|_1 \left\| L \frac{x}{\|x\|_1} \right\|_2 \leq \|L\|_{\text{op}} \|x\|_1,$$

also folgt (16.5).

Somit ist auch $\|Lx - La\|_2 = \|L(x - a)\|_2 \leq \|L\|_{\text{op}} \cdot \|x - a\|_1$,

d. h. zu beliebigem $\varepsilon > 0$ kann $\delta := \frac{\varepsilon}{2 \cdot \max(1, \|L\|_{\text{op}})} > 0$ gewählt werden, sodass schließlich $\|Lx - La\|_2 \leq \frac{\varepsilon}{2} < \varepsilon$ gilt.

⊕ Stetigkeit von L bei 0: zu $\varepsilon = 1 \exists \delta > 0 \forall z \in V: \|z\|_1 < \delta \Rightarrow \|Lz\|_2 < 1$;

Sei $x \in V \setminus \{0\}$, $\bar{x} := \frac{\delta}{2\|x\|_1} \cdot x$, dann ist $\|\bar{x}\|_1 = \frac{\delta}{2} < \delta$ und weiter

$$1 > \|L\bar{x}\|_2 = \frac{\delta}{2\|x\|_1} \cdot \|Lx\|_2 \Rightarrow \frac{2}{\delta} \|x\|_1 \geq \|Lx\|_2$$

d. h. $\forall x \in V$ mit $\|x\|_1 \leq 1$ gilt $\|Lx\|_2 \leq \frac{2}{\delta}$; somit also $\|L\|_{\text{op}} \leq \frac{2}{\delta} < \infty$.

□

Korollar: Sei $L: \mathbb{R}^n \rightarrow \mathbb{R}^m$ linear, dann gilt $\|L\|_{\text{op}} < \infty$ und (jeweils mit euklidischer Norm)

$$\forall x \in \mathbb{R}^n: \quad \|Lx\| \leq \|L\|_{\text{op}} \cdot \|x\|.$$

Beweis: L ist stetig, also folgt dies aus der obigen Behauptung bzw. aus (16.5). □

Bemerkung: Durch direktes Nachrechnen der Normaxiome (N1-3) erhalten wir, dass $\mathcal{L}(V, W) = \{L: V \rightarrow W: L \text{ ist linear und stetig}\}$ ausgestattet mit $\|\cdot\|_{\text{op}}$ zu einem normierten Vektorraum wird.

16.14. Theorem

Es seien X, Y metrische Räume und $f: X \rightarrow Y$. Dann ist äquivalent:

- 1.) Für jeden Punkt $a \in X$ gilt: zu jeder Umgebung V von $f(a)$ (in Y) gibt es eine Umgebung U von a (in X) mit $f(U) \subseteq V$.
- 2.) $\forall W \subseteq Y$ offen: $f^{-1}(W)$ ist offen in X .
- 3.) $\forall B \subseteq Y$ abgeschlossen: $f^{-1}(B)$ ist abgeschlossen in X .
- 4.) f ist stetig auf X .

Beweis:

- 1.) \Leftrightarrow 4.) folgt aus den Definitionen von Umgebung, nämlich mit Hilfe von Kugeln bzgl. der betrachteten Metrik, und Stetigkeit: $\forall \varepsilon > 0: \exists \delta > 0: f(B_\delta(a)) \subseteq B_\varepsilon(f(a))$.
- 2.) \Leftrightarrow 3.) folgt direkt mittels der Äquivalenz ‘ W offen $\Leftrightarrow Y \setminus W$ abgeschlossen’ und allgemeinen Relationen für Abbildungen $f^{-1}(Y \setminus W) = X \setminus f^{-1}(W)$ etc.

Wir zeigen noch 1.) \Leftrightarrow 2.), dann sind wir fertig:

- 1.) \Rightarrow 2.): Sei $W \subseteq Y$ offen; z. z.: $f^{-1}(W)$ offen in X .

Sei $a \in f^{-1}(W)$ beliebig. W ist Umgebung von $f(a)$, daher existiert eine Umgebung U von a (in X): $f(U) \subseteq W$, d. h. $U \subseteq f^{-1}(W)$.

- 2.) \Rightarrow 1.): Sei V Umgebung von $f(a)$, dann $\exists \varepsilon > 0: B_\varepsilon(f(a)) \subseteq V$;

es ist $a \in f^{-1}(B_\varepsilon(f(a))) =: U$ offen, also Umgebung von a , und $f(U) \subseteq V$.

□

16.15. Beispiel

Sei (X, d) ein metrischer Raum, $f: X \rightarrow \mathbb{R}$ eine stetige Funktion, $c \in \mathbb{R}$ beliebig, Dann sind Mengen der Form

$$\{x \in X: f(x) < c\} = f^{-1}(]-\infty, c]) \quad \text{offen}$$

$$\{x \in X: f(x) = c\} = f^{-1}(\{c\}) \quad \text{abgeschlossen}$$

$$\{x \in X: f(x) \leq c\} = f^{-1}(]-\infty, c]) \quad \text{abgeschlossen.}$$

Weiters ist z. B. $K_r(a) := \{x \in X: d(x, a) \leq r\} = f^{-1}([0, r])$ für $f(x) := d(x, a)$, wobei $f: X \rightarrow [0, \infty[$ stetig ist. Somit ist die „abgeschlossene“ Kugel um a mit Radius r auch abgeschlossen im Sinne der Metrik d .

16.16. Der Fixpunktsatz von Banach

Definition:

- 1.) Seien $(X, d_1), (Y, d_2)$ metrische Räume. Eine Abbildung $f: X \rightarrow Y$ heißt *Kontraktion*, falls es eine Konstante $q \in]0, 1[$ gibt, so dass $\forall x, y \in X$ gilt:

$$(16.6) \quad d_2(f(x), f(y)) \leq q \cdot d_1(x, y)$$

(insbesondere folgt, dass f stetig ist).

- 2.) Sei X eine Menge und $f: X \rightarrow X$ eine Abbildung. Ein *Fixpunkt* von f ist ein Punkt $a \in X$ mit der Eigenschaft $f(a) = a$.

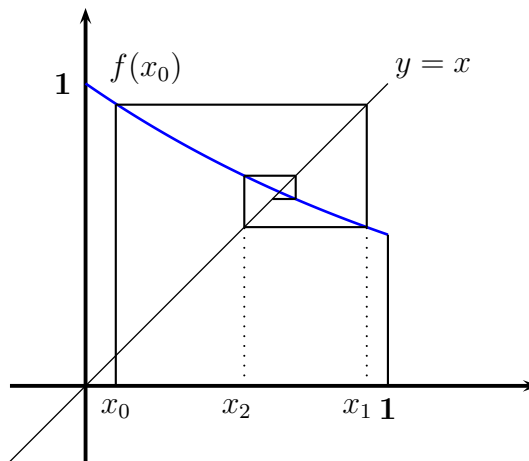
Beispiel: $X = [0, 1] \subseteq \mathbb{R}$. $f: [0, 1] \rightarrow [0, 1]$, $x \mapsto 2^{-x}$, ist eine Kontraktion:

$$\begin{aligned} |f(x) - f(y)| &\leq \sup_{z \in [0, 1]} |f'(z)| \cdot |x - y| \\ &= \sup_{z \in [0, 1]} |\log 2 \cdot 2^{-z}| \cdot |x - y| \leq \log 2 \cdot |x - y|, \end{aligned}$$

wobei $0 < \log 2 < 1$.

? Hat f einen Fixpunkt $a \in [0, 1]$

Idee: wähle $x_0 \in [0, 1]$ beliebig und definiere $x_1 = f(x_0), \dots, x_{n+1} = f(x_n), \dots$ die Skizze gibt Hoffnung auf Konvergenz gegen Fixpunkt



Theorem: Es sei (X, d) ein vollständiger (!) metrischer Raum und $f: X \rightarrow X$ eine Kontraktion. Dann besitzt f genau einen Fixpunkt $a \in X$. Für $x_0 \in X$ (beliebig) konvergiert die Folge (x_n) mit $x_{n+1} := f(x_n)$ ($n \in \mathbb{N}$) gegen a .

Beweis:

Eindeutigkeit: wäre $a \neq b$ und $f(a) = a, f(b) = b$, so folgte

$$0 < d(a, b) = d(f(a), f(b)) \leq q \cdot d(a, b) < d(a, b), \text{ ein Widerspruch } \zeta$$

Existenz: Sei $x_0 \in X$ beliebig und für $n \in \mathbb{N}$: $x_{n+1} := f(x_n)$;

$$\text{dann gilt } \forall n \geq 1: \quad d(x_{n+1}, x_n) = d(f(x_n), f(x_{n-1})) \leq q \cdot d(x_n, x_{n-1});$$

und induktiv: $d(x_{n+1}, x_n) \leq q^{n-k} \cdot d(x_{k+1}, x_k)$ ($0 \leq k < n$); daraus folgt

$$d(x_n, x_k) \leq d(x_n, x_{n-1}) + d(x_{n-1}, x_{n-2}) + \dots + d(x_{k+1}, x_k) \leq$$

$$\leq (q^{n-k-1} + q^{n-k-2} + \dots + 1) \cdot d(x_{k+1}, x_k) = \frac{1-q^{n-k}}{1-q} \cdot d(x_{k+1}, x_k).$$

Außerdem ist $d(x_{k+1}, x_k) \leq q^k \cdot d(x_1, x_0)$ (ersetze oben n durch k und k durch 0), also

$$d(x_n, x_k) \leq \frac{q^k - q^n}{1-q} \cdot d(x_1, x_0) \leq \frac{q^k}{1-q} \cdot d(x_1, x_0).$$

Dies zeigt, dass (x_n) eine Cauchy-Folge ist (denn zu $\varepsilon > 0$ wähle N mit $\frac{q^N}{1-q} \cdot d(x_1, x_0) < \varepsilon$, dann gilt $d(x_n, x_k) < \varepsilon$ für $n, k \geq N$).

(X, d) ist vollständig, daher $\exists a := \lim_{n \rightarrow \infty} x_n$ und es gilt (weil ja f stetig ist)

$$\begin{array}{ccc} x_{n+1} & = & f(x_n) \\ \downarrow & (n \rightarrow \infty) & \downarrow \\ a & & f(a), \end{array}$$

wegen der Eindeutigkeit des Limes also $a = f(a)$. □

Bemerkung: Aus $d(x_n, x_k) \leq \frac{q^k}{1-q} d(x_1, x_0)$ erhalten wir für $n \rightarrow \infty, x_n \rightarrow a$,

die Fehlerabschätzung $d(a, x_k) \leq \frac{q^k}{1-q} \cdot d(x_1, x_0)$ im k -ten Schritt des Verfahrens.

16.17. Gleichmäßige Konvergenz

Definition: Sei X eine Menge, (Y, ρ) ein metrischer Raum, sowie $f: X \rightarrow Y$ und $f_n: X \rightarrow Y$ (für $n \in \mathbb{N}$) eine Folge von Abbildungen.

Die Folge (f_n) konvergiert gleichmäßig gegen f , falls gilt:

$$\forall \varepsilon > 0 \exists N \in \mathbb{N}: \underbrace{\forall x \in X: \rho(f_n(x), f(x)) < \varepsilon}_{\text{[d.h. } \sup_{x \in X} \rho(f_n(x), f(x)) \leq \varepsilon]} \quad \forall n \geq N.$$

Proposition: Seien $(X, d), (Y, \rho)$ metrische Räume, $f_n: X \rightarrow Y$ stetig ($n \in \mathbb{N}$). Wenn (f_n) gleichmäßig gegen $f: X \rightarrow Y$ konvergiert, dann ist f stetig.

Beweis: ① gleichmäßige Konvergenz: $\forall \varepsilon > 0 \exists N \forall n \geq N \forall x \in X: \rho(f_n(x), f(x)) < \frac{\varepsilon}{3}$.

② f_N ist stetig in $a \in X$, daher $\exists \delta > 0: d(x, a) < \delta \Rightarrow \rho(f_N(x), f_N(a)) < \frac{\varepsilon}{3}$.

Somit gilt für alle $x \in X$ mit $d(x, a) < \delta$:

$$\rho(f(x), f(a)) \leq \underbrace{\rho(f(x), f_N(x))}_{\text{①: } < \frac{\varepsilon}{3}} + \underbrace{\rho(f_N(x), f_N(a))}_{\text{②: } < \frac{\varepsilon}{3}} + \underbrace{\rho(f_N(a), f(a))}_{\text{①: } < \frac{\varepsilon}{3}} < \varepsilon.$$

Daher ist f stetig in a ; a war beliebig aus X , also ist f stetig auf X . □

16.18. Korollar

Sei (X, d) ein metrischer Raum und

$$\mathcal{C}_b(X, \mathbb{R}^m) := \{f: X \rightarrow \mathbb{R}^m \text{ stetig} : \|f\|_\infty := \sup_{x \in X} \|f(x)\| < \infty\}.$$

Dann ist $(\mathcal{C}_b(X, \mathbb{R}^m), \|\cdot\|_\infty)$ ein Banach-Raum.

Beweis: Zunächst ist $\|\cdot\|_\infty: \mathcal{C}_b(X, \mathbb{R}^m) \rightarrow \mathbb{R}$ eine Norm [wie in §15].

$$(f_n) \text{ Cauchy-Folge bzgl. } \|\cdot\|_\infty \implies \forall x \in X: (f_n(x)) \text{ ist eine Cauchy-Folge in } \mathbb{R}^m \implies \\ \forall x \in X: \exists f(x) := \lim_{n \rightarrow \infty} f_n(x) \in \mathbb{R}^m$$

Somit ist eine Abbildung $f: X \rightarrow \mathbb{R}^m$ definiert; es bleibt noch zu zeigen, dass $f_n \rightarrow f$ und f stetig und beschränkt ist:

Sei $x \in X, k, n \in \mathbb{N}$

$$\|f_k(x) - f_n(x)\| \leq \|f_k - f_n\|_\infty < \varepsilon \text{ für } k, n \geq N$$

$$\downarrow (n \rightarrow \infty)$$

$$\|f_k(x) - f(x)\|, x \text{ beliebig} \implies \|f_k - f\|_\infty \leq \varepsilon \quad (k \geq N).$$

Daher ist (f_k) gleichmäßig konvergent gegen f und weiters folgt mit 16.17, dass f stetig ist.

$$\text{Schließlich ist } \|f(x)\| \leq \|f(x) - f_N(x)\| + \|f_N(x)\| \leq \underbrace{\|f - f_N\|_\infty}_{\leq \varepsilon} + \|f_N\|_\infty,$$

daher $\|f\|_\infty \leq \varepsilon + \|f_N\|_\infty < \infty$, also ist f beschränkt. □

§17. Kompaktheit

17.1. Definition

Es sei (X, d) ein metrischer Raum und $A \subseteq X$.

1.) Sei I eine Menge und $\forall i \in I : U_i \subseteq X$ offen.

$(U_i)_{i \in I}$ heißt *offene Überdeckung* von A , wenn gilt: $\bigcup_{i \in I} U_i \supseteq A$.

2.) A heißt *kompakt*, wenn jede offene Überdeckung $(U_i)_{i \in I}$ von A eine endliche Teilüberdeckung enthält, d. h. $\exists k \in \mathbb{N} : \exists i_1, \dots, i_k \in I : \bigcup_{l=1}^k U_{i_l} \supseteq A$ [endliche Vereinigung!]

17.2. Proposition

1.) Jede kompakte Teilmenge eines metrischen Raumes ist beschränkt und abgeschlossen.

2.) Jede abgeschlossene Teilmenge einer kompakten Menge ist selbst kompakt.

Beweis:

2.) Sei $K \subseteq X$ kompakt, $A \subseteq K$ abgeschlossen und $(U_i)_{i \in I}$ eine offene Überdeckung von A .

Es ist $X \setminus A$ offen und $(X \setminus A) \cup \bigcup_{i \in I} U_i \supseteq (X \setminus A) \cup A = X \supseteq K$, d. h. durch Hinzufügen von $X \setminus A$ erhalten wir eine offene Überdeckung von K .

Aus der Kompaktheit von K folgt nun: $\exists i_1, \dots, i_k : (X \setminus A) \cup \bigcup_{j=1}^k U_{i_j} \supseteq K \supseteq A$, daher

auch $\bigcup_{j=1}^k U_{i_j} \supseteq A$. Also ist A kompakt.

1.) Sei $A \subseteq X$ kompakt und $A \neq \emptyset$ (andernfalls ist die Aussage trivial).

Sei $a \in A$, dann gilt $\bigcup_{n=1}^{\infty} B_n(a) \supseteq A$, also ist $(B_n(a))_{n \geq 1}$ eine offene Überdeckung von

A . Die Kompaktheit von A garantiert: $\exists n_1, \dots, n_k \in \mathbb{N} \setminus \{0\} : \bigcup_{j=1}^k B_{n_j}(a) \supseteq A$

Wegen $B_n(a) \subseteq B_m(a)$ für $m \geq n$ ist dann $A \subseteq B_N(a)$, wobei $N := \max(n_1, \dots, n_k)$; also ist A beschränkt.

Noch z. z.: A ist abgeschlossen

Wir zeigen: $X \setminus A$ offen

Falls $A = X$, dann $X \setminus A = \emptyset$ offen. Daher nehmen wir im weiteren $A \neq X$ an.

Sei $x \in X \setminus A$; für $n \in \mathbb{N}$, $n \geq 1$ setze

$$U_n := \{y \in X : d(x, y) > \frac{1}{n}\};$$

die Mengen U_n sind offen und $\bigcup_{n=1}^{\infty} U_n = X \setminus \{x\} \supseteq A$, d.h. wir haben damit eine offene Überdeckungen von A konstruiert. Die Kompaktheit von A liefert somit

$$\exists n_1, \dots, n_k: A \subseteq \bigcup_{j=1}^k U_{n_j}, \text{ d.h. } \forall y \in A \exists j: d(x, y) > \frac{1}{n_j}.$$

Setze $N := \max(n_1, \dots, n_k)$, dann ist $\frac{1}{N} \leq \frac{1}{n_j}$ ($j = 1, \dots, k$) und $x \in B_{\frac{1}{N}}(x) \subseteq X \setminus A$. Also ist $X \setminus A$ offen.

□

17.3. Theorem (Satz von Heine-Borel¹)

Für $A \subseteq \mathbb{R}^n$ gilt:

A ist kompakt \iff A ist beschränkt und abgeschlossen

Bemerkung und Warnung: In beliebigen metrischen Räumen müssen beschränkte, abgeschlossene Teilmengen nicht immer kompakt sein (Beispiele in VO zu Topologie und Funktionalanalysis).

Beweis:

⊃ gilt allgemein nach Proposition 17.2.

⊃ Indirekt: Sei $(U_i)_{i \in I}$ eine offene Überdeckung von A , für die keine endliche Teilüberdeckung von A existiert.

A ist beschränkt, daher gibt es einen abgeschlossenen Würfel $W = \prod_{j=1}^n [-\frac{s}{2}, \frac{s}{2}]$ (also mit Kantenlänge $= s > 0$) mit $A \subseteq W$. [Bew: wähle $s > 0$ mit $A \subseteq B_{\frac{s}{2}}(0) \subseteq W$]

Wir zerlegen W in 2^n Teilwürfel mit Kantenlänge $\frac{s}{2}$.

Es muss ein Teilwürfel W_1 darunter sein, so dass $A \cap W_1$ nicht von endlich vielen U_i überdeckt werden kann;

¹Heinrich Eduard Heine (*15. 3. 1821 Berlin; †21. 10. 1881 Halle), deutscher Mathematiker
 Félix Édouard Justin Émile Borel (*7. 1. 1871 Saint-Affrique; †3. 2. 1956 Paris) [fe'liks edu'ar jys'tê e'mil bo'rel], französischer Mathematiker und Politiker

Wiederholung des Verfahrens mit W_1 usw. liefert eine Folge abgeschlossener Würfel $W_1 \supseteq W_2 \supseteq \dots$, wobei W_k Kantenlänge $\frac{s}{2^k}$ hat und

(\star) $\forall k : A \cap W_k$ kann nicht von endlich vielen der U_i ($i \in I$) überdeckt werden.

Für $k \geq 1$ wähle $x_k \in A \cap W_k$ beliebig; für $l \geq k$ gilt dann $x_l, x_k \in A \cap W_k$ und somit $d(x_l, x_k) \leq \frac{s}{2^k}$.

Daher ist (x_k) eine Cauchy-Folge; sei $a := \lim_{k \rightarrow \infty} x_k$. Es ist $a \in A$, weil A abgeschlossen ist. Außerdem $\exists i_0 \in I : a \in U_{i_0}$ (weil $A \subseteq \bigcup_{i \in I} U_i$).

$\exists k_0 \in \mathbb{N} \forall k \geq k_0 : a \in W_k \subseteq U_{i_0}$

(denn $\exists \varepsilon > 0 : B_\varepsilon(a) \subseteq U_{i_0}$ und für $\frac{s}{2^{k_0}} < \varepsilon$ ist $W_k \subseteq W_{k_0} \subseteq B_\varepsilon(a) \forall k \geq k_0$)

Insgesamt folgt nun $\forall k \geq k_0 : A \cap W_k \subseteq U_{i_0}$, ein Widerspruch \nexists zu (\star).

□

Das Konzept der Teilfolgen ist in allgemeinen metrischen Räumen genauso wie in \mathbb{R} :

für eine streng monotone Teilfolge von Indizes $n_0 < n_1 < \dots$ bildet $(x_{n_k})_{k \in \mathbb{N}}$ eine Teilfolge von (x_n) .

17.4. Proposition (Satz von Bolzano-Weierstraß)

Sei A eine kompakte Teilmenge des metrischen Raumes X und (x_n) eine Folge in A (d. h. für jedes n ist $x_n \in A$). Dann gibt es eine Teilfolge $(x_{n_k})_{k \in \mathbb{N}}$, die gegen einen Punkt $a \in A$ konvergiert.

Beweis: Indirekt: angenommen, es gäbe keine konvergente Teilfolge. Dann gibt es zu jedem $a \in A$ eine offene Umgebung U_a von a , in der nur endlich viele Folgenglieder liegen.

Wegen $A \subseteq \bigcup_{a \in A} U_a$ und der Kompaktheit von A gibt es $a_1, \dots, a_N \in A$ so, dass $A \subseteq \bigcup_{j=1}^N U_{a_j}$.

Demnach liegen aber überhaupt nur endlich viele Folgenglieder in A , ein Widerspruch \nexists □

Bemerkung: Es gilt auch die Umkehrung in obigem Satz (vgl. [Heu04, Satz 157.1]).

17.5. Korollar

Jede beschränkte Folge in \mathbb{R}^n besitzt eine konvergente Teilfolge.

Beweis: Sei (x_k) eine beschränkte Folge. Dann ist $\{x_k : k \in \mathbb{N}\}$ enthalten in einer offenen Kugel $B_R(0)$, somit auch in deren Abschluss $\overline{B_R(0)} = \{y \in \mathbb{R}^n : \|y\| \leq R\}$, der nach dem Satz von Heine-Borel kompakt ist. □

17.6. Theorem

Seien X, Y metrische Räume und $f : X \rightarrow Y$ stetig. Ist $K \subseteq X$ kompakt, dann ist auch $f(K) \subseteq Y$ kompakt.

Beweis:

Sei $(V_i)_{i \in I}$ eine offene Überdeckung von $f(K)$. Für jedes i ist $U_i := f^{-1}(V_i)$ offen in X , weil f stetig ist. Außerdem ist laut Konstruktion $K \subseteq \bigcup_{i \in I} U_i$, was wegen der Kompaktheit

von K eine endliche Auswahl von Indizes $\exists i_1, \dots, i_k \in I$ mit $K \subseteq \bigcup_{j=1}^k U_{i_j}$ erlaubt.

Dann ist aber auch $f(K) \subseteq \bigcup_{j=1}^k f(U_{i_j}) \subseteq \bigcup_{j=1}^k V_{i_j}$, somit eine endliche Teilüberdeckung von $(V_i)_{i \in I}$ für $f(K)$ gefunden. \square

17.7. Korollar

Sei X ein metrischer Raum, $K \subseteq X$ eine kompakte Teilmenge und $f : K \rightarrow \mathbb{R}$ stetig. Dann ist die Funktion f beschränkt (d. h. $f(K)$ ist beschränkt) und nimmt ihr Maximum und Minimum an, d. h. $\exists p, q \in K : f(p) = \sup f(K), f(q) = \inf f(K)$.

Beweis:

$A := f(K) \subseteq \mathbb{R}$ ist kompakt, da f stetig ist. Somit ist A beschränkt und abgeschlossen.

Sei $\alpha := \sup A$ und $\beta := \inf A$, dann gibt es Folgen (x_k) und (y_k) in A mit $x_k \rightarrow \alpha$ und $y_k \rightarrow \beta$ ($k \rightarrow \infty$). Da A abgeschlossen ist, gilt $\alpha, \beta \in A$ und $\exists p, q \in K : f(p) = \alpha, f(q) = \beta$. \square

17.8. Proposition

Es seien $(X, d_1), (Y, d_2)$ metrische Räume, $K \subseteq X$ kompakt und $f : K \rightarrow Y$ stetig. Dann ist $f : K \rightarrow Y$ *gleichmäßig stetig*, d. h.

$$\forall \varepsilon > 0 \exists \delta > 0 \forall x, y \in K : d_1(x, y) < \delta \implies d_2(f(x), f(y)) < \varepsilon$$

(δ kann unabhängig von der Stelle x bzw. y gewählt werden).

Beweis: Sei $\varepsilon > 0$; f ist stetig in jedem Punkt von X , daher gilt:

$$\forall x \in K \exists \delta(x) > 0 \forall y \in K : d_1(x, y) < \delta(x) \implies d_2(f(x), f(y)) < \frac{\varepsilon}{2};$$

$\bigcup_{x \in K} B_{\frac{\delta(x)}{2}}(x) \supseteq K$ und K ist kompakt, daher $\exists x_1, \dots, x_k \in K : \bigcup_{j=1}^k B_{\frac{\delta(x_j)}{2}}(x_j) \supseteq K$.

Wir setzen $\delta := \min\left(\frac{\delta(x_1)}{2}, \dots, \frac{\delta(x_k)}{2}\right)$; für $x, y \in K$ beliebig mit $d_1(x, y) < \delta$ gilt:

$\exists j \in \{1, \dots, k\} : x \in B_{\frac{\delta(x_j)}{2}}(x_j)$ und damit auch

$d_1(y, x_j) \leq d_1(y, x) + d_1(x, x_j) < \delta + \frac{\delta(x_j)}{2} \leq \frac{\delta(x_j)}{2} + \frac{\delta(x_j)}{2} = \delta(x_j)$, daher weiter

$d_2(f(x), f(y)) \leq d_2(f(x), f(x_j)) + d_2(f(x_j), f(y)) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$. □

17.9. Anwendung: Äquivalenz von Normen auf \mathbb{R}^n

Definition: Zwei Normen $\|\cdot\|$ und $\|\cdot\|'$ auf dem Vektorraum V (über \mathbb{R} oder \mathbb{C}) heißen *äquivalent*, falls gilt:

$$(17.1) \quad \exists C_1, C_2 > 0 : \forall x \in V : C_1 \|x\| \leq \|x\|' \leq C_2 \|x\|.$$

Bemerkung: 1.) Es ist leicht zu sehen, dass (17.1) eine Äquivalenzrelation (auf der Menge aller Normen auf V) definiert.

2.) Auf unendlichdimensionalen Vektorräumen gibt es inäquivalente Normen (siehe Funktionalanalysis für Beispiele).

Proposition: Seien $\|\cdot\|$ und $\|\cdot\|'$ äquivalente Normen auf V und d bzw. d' die entsprechenden Metriken. Dann definieren d und d' die gleichen offenen Mengen und Umgebungen in V (daher bleiben auch die Begriffe Konvergenz und Stetigkeit dieselben).

(„Die Topologie wird durch Übergang zu einer äquivalenten Norm nicht verändert.“)

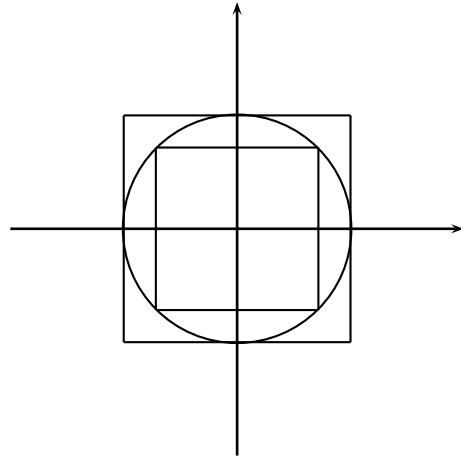
Beweis: Es genügt, die entsprechende Schachtelung von offenen Kugeln zu zeigen; dies zeigt man leicht aus der folgenden Beobachtung:

$$\forall x, y \in V : d(x, y) = \|x - y\| \leq \frac{1}{C_1} \|x - y\|' = \frac{1}{C_1} \cdot d'(x, y) \leq \frac{C_2}{C_1} \|x, y\| = \frac{C_2}{C_1} d(x, y).$$

□

Beispiel: Auf \mathbb{R}^n sind $\|\cdot\|_\infty$ und $\|\cdot\|_2$ äquivalent, denn

$$\begin{aligned} \|x\|_\infty &= \max(|x_1|, \dots, |x_n|) \leq \left(\sum_{j=1}^n |x_j|^2 \right)^{\frac{1}{2}} = \\ &= \|x\|_2 \leq \left(\sum_{j=1}^n \|x\|_\infty^2 \right)^{\frac{1}{2}} = \sqrt{n} \cdot \|x\|_\infty \quad \forall x \in \mathbb{R}^n \end{aligned}$$



Theorem: Je zwei Normen auf \mathbb{R}^n sind äquivalent.

Beweis:

Sei $\|\cdot\|$ eine beliebige Norm auf \mathbb{R}^n und bezeichne $\|\cdot\|_2$ die euklidische Norm.

Sei $e_j = (0, \dots, 0, \underbrace{1}_{j\text{-te Komponente}}, 0, \dots, 0)$ ($j = 1, \dots, n$).

Für jedes $x \in \mathbb{R}^n$ haben wir die Basisdarstellung $x = \sum_{j=1}^n x_j e_j$. Daraus folgt mit der Dreiecksungleichung

$$\|x\| \leq \sum_{j=1}^n |x_j| \cdot \|e_j\| \leq \left(\sum_{j=1}^n |x_j|^2 \right)^{1/2} \cdot \underbrace{\left(\sum_{j=1}^n \|e_j\|^2 \right)^{\frac{1}{2}}}_{=: C_2} = C_2 \cdot \|x\|_2.$$

Somit gilt $|\|x\| - \|y\|| \leq \|x - y\| \leq C_2 \|x - y\|_2$, daher ist die Abbildung $x \mapsto \|x\|$ stetig ($(\mathbb{R}^n, \|\cdot\|_2) \rightarrow (\mathbb{R}, |\cdot|)$).

$S^{n-1} = \{x \in \mathbb{R}^n : \|x\|_2 = 1\}$ ist kompakt (weil beschränkt und abgeschlossen), daher

$$\exists x_0 \in S^{n-1} : \|x_0\| = \min\{\|x\| : x \in S^{n-1}\}.$$

Wir setzen $C_1 := \|x_0\|$; es ist $C_1 > 0$, weil $x_0 \neq 0$.

$$\forall x \neq 0: \frac{x}{\|x\|_2} \in S^{n-1}, \text{ daher } C_1 \leq \left\| \frac{x}{\|x\|_2} \right\| = \frac{\|x\|}{\|x\|_2}, \text{ d.h. } C_1 \cdot \|x\|_2 \leq \|x\|.$$

Insgesamt haben wir gezeigt: $C_1 \cdot \|x\|_2 \leq \|x\| \leq C_2 \cdot \|x\|_2$.

Somit ist jede Norm auf \mathbb{R}^n äquivalent zur 2-Norm, also sind je zwei Normen äquivalent (Transitivität und Symmetrie). \square

VII DIFFERENZIERBARE ABBILDUNGEN $\mathbb{R}^n \rightarrow \mathbb{R}^m$

Grundlegende Notationen und Begriffe

$x \in \mathbb{R}^n$, $x = (x_1, \dots, x_n)$... Zeilenschreibweise (um Platz zu sparen; in Operationen als Spaltenvektoren aufgefasst)

$1 \leq j \leq n$: $e_j = (\delta_{lj})_{l=1}^n = (0, \dots, 0, 1, 0, \dots, 0)$... j -ter Standardbasisvektor (1 in der j -ten Komponente, sonst 0); $\{e_1, \dots, e_n\}$... Standardbasis des \mathbb{R}^n

$U \subseteq \mathbb{R}^n$, $f: U \rightarrow \mathbb{R}^m$ eine Abbildung ... oft in Kurzschreibweise: $f: \mathbb{R}^n \supseteq U \rightarrow \mathbb{R}^m$

Komponentenfunktionen: $\forall x \in U: f(x) = (f_1(x), \dots, f_m(x))$, wobei für $j = 1, \dots, m$ jeweils $f_j: U \rightarrow \mathbb{R}$; es ist $f = (f_1, \dots, f_m)$

Wichtiger Spezialfall: *skalare Funktion* $f: \mathbb{R}^n \supseteq U \rightarrow \mathbb{R}$ *mehrerer Variablen* x_1, \dots, x_n

mit *partiellen Funktionen* bei festem $x \in U$:

$$\tilde{f}_{k,x}: y \mapsto f(x_1, \dots, x_{k-1}, x_k + y, x_{k+1}, \dots, x_n) = f(x + y \cdot e_k) \quad (1 \leq k \leq n)$$

ist definiert für $y \in \mathbb{R}$ mit $(x_1, \dots, x_{k-1}, x_k + y, x_{k+1}, \dots, x_n) \in U$

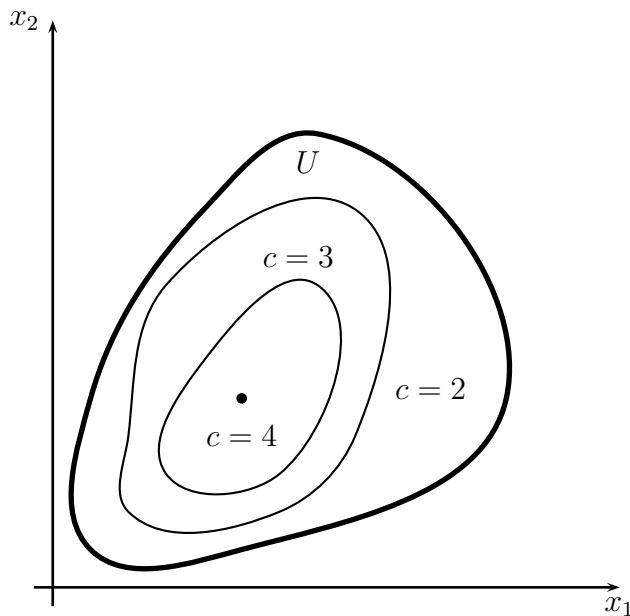
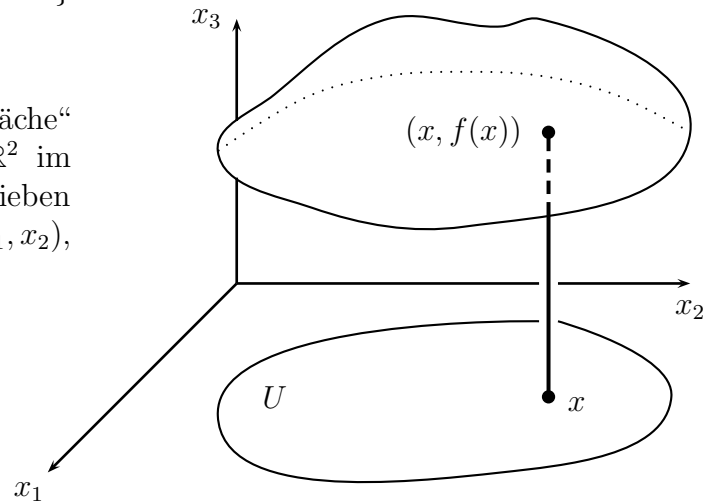
Bemerkung: Wenn U offen, dann existiert ein offenes Intervall $I \subseteq \mathbb{R}$ um 0 derart, dass $\tilde{f}_{k,x}$ auf I definiert ist.

(Beweis: $y \mapsto x + y \cdot e_k$ ist stetig, daher ist das Urbild von U darunter offen in \mathbb{R} .)

Veranschaulichung skalarer Funktionen durch Graphen und Niveaumengen

$G(f) := \{(x, f(x)) \in \mathbb{R}^n \times \mathbb{R} : x \in U\}$
 heißt *Graph von f*

Im Fall $n = 2$ ergibt sich eine „Fläche“
 oder „Landschaft“ über $U \subseteq \mathbb{R}^2$ im
 dreidimensionalen Raum, beschrieben
 in Koordinaten durch $x_3 = f(x_1, x_2)$,
 $x = (x_1, x_2) \in U$.



Niveaumengen (auch „Niveaulinien“ oder
 „Höhenschichtlinien“):

für jedes $c \in \mathbb{R}$ ist

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

§18. Partielle Ableitungen und Differenzierbarkeit

18.1. Definition

Sei $U \subseteq \mathbb{R}^n$ offen, $f: U \rightarrow \mathbb{R}$ und $1 \leq j \leq n$.

Dann heißt f in $x \in U$ *partiell differenzierbar* in der j -ten Koordinatenrichtung, falls

$$(18.1) \quad D_j f(x) := \lim_{h \rightarrow 0} \frac{f(x + h \cdot e_j) - f(x)}{h} \quad [= (\tilde{f}_{j,x})'(0)]$$

existiert (wobei der Limes natürlich für $h \rightarrow 0$ mit $h \neq 0, x + he_j \in U$ zu nehmen ist); in diesem Fall heißt $D_j f(x)$ die *j -te partielle Ableitung* von f bei x .

Wenn $D_j f(x)$ für jedes $x \in U$ existiert, so definiert $x \mapsto D_j f(x)$ wieder eine Funktion: $D_j f: U \rightarrow \mathbb{R}$. Die Funktion f heißt *partiell differenzierbar* (in U), falls $D_j f(x)$ für jedes $x \in U$ und $\forall j \in \{1, \dots, n\}$ existiert. Es sind dann die partiellen Ableitungsfunktionen $D_1 f, \dots, D_n f: U \rightarrow \mathbb{R}$ definiert. f heißt *stetig partiell differenzierbar*, wenn f partiell differenzierbar ist und alle $D_1 f, \dots, D_n f$ stetig sind.

Andere Schreibweisen:

$$D_j f(x) = \partial_j f(x) = \frac{\partial f}{\partial x_j}(x) = \frac{\partial f(x)}{\partial x_j} \quad \text{bzw.} \quad D_j f = \partial_j f = \frac{\partial f}{\partial x_j}$$

- Formal klarer ist die Notation als Operator ohne Variablenbezeichnung
- In Rechnungen mit konkreten Termen sind aber oft Ausdrücke mit Variablenbezeichnungen im Ableitungsoperator selbst vorteilhaft, wie z. B. in Ausdrücken der Form

$$\frac{\partial}{\partial y} \left(e^{-x^2+y^2} \right) = e^{-x^2} \cdot \frac{\partial}{\partial y} \left(e^{y^2} \right) = e^{-x^2} \cdot 2y e^{y^2} = 2y e^{-x^2+y^2}.$$

([nach Michael Grosser:] x wird fürs Ableiten nach y zunächst eingefroren und danach wieder aufgetaut \rightsquigarrow einfache Ableitung nach y mit x als „Konstante“; heraus kommt aber wieder eine Funktion beider Variablen!)

18.2. Beispiele

1.) Sei $r: \mathbb{R}^n \rightarrow \mathbb{R}$, $r(x) := \|x\| = \left(\sum_{j=1}^n x_j^2 \right)^{\frac{1}{2}}$

Für $c > 0$ ist $N_r(c) = \{x \in \mathbb{R}^n : \|x\| = c\} = c \cdot S^{n-1}$, also die Sphäre mit Radius c .

Sei $x \neq 0$ dann ist für $(k = 1, \dots, n)$:

$$D_k r(x) = \frac{1}{2} \left(\sum_{j=1}^n x_j^2 \right)^{-\frac{1}{2}} \cdot 2x_k = \frac{x_k}{r(x)},$$

daher ist r in $\mathbb{R}^n \setminus \{0\}$ stetig partiell differenzierbar.

Etwas allgemeiner: Sei $f:]0, \infty[\rightarrow \mathbb{R}$ stetig differenzierbar, dann ist $g := f \circ r$, d.h. $g(x) = f(r(x)) = f(\|x\|)$, stetig partiell differenzierbar in $\mathbb{R}^n \setminus \{0\}$ und

$$D_k g(x) = f'(\|x\|) \cdot D_k r(x) = f'(\|x\|) \cdot \frac{x_k}{r(x)}.$$

2.) Die Funktion $F: \mathbb{R}^2 \rightarrow \mathbb{R}$, gegeben durch

$$F(x) = \begin{cases} \frac{x_1 \cdot x_2}{\|x\|^2} & \text{für } x \neq 0 \\ 0 & \text{für } x = 0, \end{cases}$$

ist in $\mathbb{R}^2 \setminus \{0\}$ partiell differenzierbar, denn es ist $F(x) = \frac{x_1 x_2}{x_1^2 + x_2^2}$ für $x \neq 0$ und nach der Quotienten-Regel folgt

$$D_1 F(x) = \frac{x_2(x_1^2 + x_2^2) - x_1 x_2 2x_1}{(x_1^2 + x_2^2)^2} = \frac{x_2}{\|x\|^2} - 2 \frac{x_1^2 x_2}{\|x\|^4}.$$

Analog ist $D_2 F(x) = \frac{x_1}{\|x\|^2} - 2 \frac{x_1 x_2^2}{\|x\|^4}$.

? partielle Differenzierbarkeit bei $x = 0$: wegen $\frac{F(0+he_k) - F(0)}{h} = \frac{0-0}{h} = 0$ folgt $D_k F(0) = 0$.

Also ist F partiell differenzierbar auf ganz \mathbb{R}^2 .

ABER F ist nicht stetig in $(0, 0)$! (Übungsaufgabe; betrachte die Folge $(\frac{1}{k}, \frac{1}{k})$)

Somit im Allgemeinen: partiell differenzierbar $\not\Rightarrow$ stetig

3.) Es sei $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ gegeben durch

$$f(x, y) = \begin{cases} \frac{xy^3}{x^2 + y^6} & (x, y) \neq (0, 0) \\ 0 & (x, y) = (0, 0) \end{cases}$$

- $(x, y) \neq (0, 0)$: $t \mapsto \frac{ty^3}{t^2 + y^6}$ ist differenzierbar, die Ableitung bei $t = x$ liefert

$$\begin{aligned} D_1 f(x, y) = D_x f(x, y) &= \frac{y^3}{x^2 + y^6} - xy^3 \cdot \frac{2x}{(x^2 + y^6)^2} \\ &= \frac{y^3}{(x^2 + y^6)^2} \cdot (x^2 + y^6 - 2x^2) = \frac{y^3(y^6 - x^2)}{(x^2 + y^6)^2} \end{aligned}$$

ebenso durch Differentiation von $s \mapsto \frac{xs^3}{x^2 + s^6}$ bei $s = y$

$$\begin{aligned} D_2 f(x, y) = D_y f(x, y) &= \frac{3xy^2}{x^2 + y^6} - xy^3 \cdot \frac{6y^5}{(x^2 + y^6)^2} \\ &= \frac{3xy^2}{(x^2 + y^6)^2} (x^2 + y^6 - 2y^6) = \frac{3xy^2(x^2 - y^6)}{(x^2 + y^6)^2} \end{aligned}$$

- In $(0, 0)$: $f(x, 0) = 0$, daher $D_1 f(0, 0) = 0$;
 $f(0, y) = 0$, daher $D_2 f(0, 0) = 0$;
daher ist f partiell differenzierbar auf \mathbb{R}^2

Es ist aber übrigens $D_1 f$ nicht stetig in $(0, 0)$, denn z.B.

$$D_1 f(2t^3, t) = \frac{t^3(t^6 - 4t^6)}{(4t^6 + t^6)^2} = -\frac{3t^9}{25t^{12}} = -\frac{3}{25t^3}$$

ist nicht konvergent für $t \rightarrow 0$.

- f ist nicht stetig in $(0, 0)$: für $t > 0$ ist

$$f(\pm t^3, t) = \pm \frac{t^3 \cdot t^3}{t^6 + t^6} = \pm \frac{1}{2} \not\rightarrow 0 = f(0, 0) \quad (t \rightarrow 0).$$

18.3. Definition

Sei $U \subseteq \mathbb{R}^n$ offen, $f: U \rightarrow \mathbb{R}$ und $v_1, \dots, v_n: U \rightarrow \mathbb{R}$ seien partiell differenzierbare Funktionen; wir bilden $v := (v_1, \dots, v_n)$, somit ist $v: U \rightarrow \mathbb{R}^n$ ein *Vektorfeld* auf U .

1.) Der *Gradient* von f ist definiert durch $\text{grad } f: U \rightarrow \mathbb{R}^n$ mit $\text{grad } f(x) := \begin{pmatrix} D_1 f(x) \\ \vdots \\ D_n f(x) \end{pmatrix}$

2.) Die *Divergenz* von v ist definiert durch $\text{div}(v): U \rightarrow \mathbb{R}$ mit

$$\text{div}(v)(x) := \sum_{j=1}^n D_j v_j(x)$$

Formal erhalten wir mit dem sogenannten Nabla-Operator¹ $\nabla := \begin{pmatrix} D_1 \\ \vdots \\ D_n \end{pmatrix}$ die Merkgeln

$\text{grad } f = \nabla f$ [wie skalare Multiplikation von ∇ mit f von rechts]

$\text{div}(v) = \langle \nabla | v \rangle = D_1 v_1 + \dots + D_n v_n$ [inneres Produkt von ∇ mit v]

18.4. Beispiele

1.) $r(x) = \|x\|$, $r: \mathbb{R}^n \setminus \{0\} \rightarrow \mathbb{R}$

$$\text{grad } r(x) = (D_1 r(x), \dots, D_n r(x)) = \left(\frac{x_1}{r(x)}, \dots, \frac{x_n}{r(x)} \right) = \frac{x}{r(x)}$$

¹Die Bezeichnung Nabla geht auf William Robertson Smith (1846–1894) zurück und ist vom Namen eines antiken Saiteninstruments abgeleitet, das eine ähnliche Form hatte.

allgemeiner für eine Verknüpfung mit einer differenzierbaren Funktion f , d.h. $g(x) = f(r(x)) = f(\|x\|)$ ist

$$\text{grad } g(x) = \frac{f'(\|x\|)}{\|x\|} \cdot x$$

2.) Für $v(x) := \frac{x}{r(x)}$, $v: \mathbb{R}^n \setminus \{0\} \rightarrow \mathbb{R}^n$ erhalten wir

$$\begin{aligned} \text{div}(v)(x) &= \sum_{j=1}^n D_j \left(\frac{x_j}{\|x\|} \right) = \sum_{j=1}^n \left(\frac{1}{r(x)} - \frac{1}{2} \cdot \frac{2x_j^2}{\|x\|^3} \right) \\ &= \frac{n}{r(x)} - \frac{1}{\|x\|^3} \cdot \sum_{j=1}^n x_j^2 = \frac{n}{\|x\|} - \frac{1}{\|x\|^3} \cdot \|x\|^2 = \frac{n-1}{\|x\|} \end{aligned}$$

18.5. Höhere partielle Ableitungen

Sei $U \subseteq \mathbb{R}^n$ offen, $f: U \rightarrow \mathbb{R}$ partiell differenzierbar und $D_1 f, \dots, D_n f: U \rightarrow \mathbb{R}$ jeweils wieder partiell differenzierbar, d. h.: $\exists D_l(D_j f)$ ($1 \leq l, j \leq n$), dann heißt f *zweimal partiell differenzierbar*.

Induktiv: $f: U \rightarrow \mathbb{R}$ heißt $(k+1)$ -mal *partiell differenzierbar* ($k \in \mathbb{N}$), wenn f k -mal partiell differenzierbar ist und alle partiellen Ableitungen $D_{j_k} \cdots D_{j_2} D_{j_1} f: U \rightarrow \mathbb{R}$ ($1 \leq j_1, \dots, j_k \leq n$) wieder partiell differenzierbar sind.

f heißt k -mal *stetig partiell differenzierbar*, wenn f k -mal partiell differenzierbar ist und alle Ableitungen der Ordnung $\leq k$ stetig sind.

$\boxed{?}$ f zweimal partiell differenzierbar $\stackrel{?}{\implies} D_j D_i f = D_i D_j f$ ($1 \leq i, j \leq n$).

Im Allgemeinen: nein!

(Übungsaufgabe: $f(x, y) = xy^3/(x^2 + y^2)$ für $(x, y) \neq (0, 0)$ und $f(0, 0) = 0$ ergibt $D_x f(0, y) = y$, $D_y f(x, 0) = 0$, daher $D_y D_x f(0, 0) = 1 \neq 0 = D_x D_y f(0, 0)$.)

18.6. Theorem (Satz von Schwarz²)

Es sei $f: \mathbb{R}^n \underset{\text{offen}}{\supseteq} U \rightarrow \mathbb{R}$ zweimal stetig partiell differenzierbar. Dann gilt

$$1 \leq i, j \leq n: \forall a \in U: D_j D_i f(a) = D_i D_j f(a).$$

Beweis:

OBdA: $n = 2$ (nur zwei Koordinatenrichtungen sind betroffen, die restlichen Variablen bleiben konstant) und $a = 0$ (was durch Verschiebung bewerkstelligt werden kann).

²Hermann Amandus Schwarz (*25. 1. 1843 Hermsdorf; †30. 11. 1921 Berlin), deutscher Mathematiker

Wähle $\delta > 0$: $[-\delta, \delta] \times [-\delta, \delta] \subseteq U$.

- Sei $0 \neq x_2 \in [-\delta, \delta]$, definiere $F_{x_2}: [-\delta, \delta] \rightarrow \mathbb{R}$, $F_{x_2}(t) := f(t, x_2) - f(t, 0)$;

Sei $0 \neq x_1 \in [-\delta, \delta]$, dann liefert der Mittelwertsatz der Differentialrechnung [Analysis 1]:
 $\exists \xi_1 \in \mathbb{R}$ mit $|\xi_1| \leq |x_1|$:

$$F_{x_2}(x_1) - F_{x_2}(0) = F'_{x_2}(\xi_1) \cdot x_1 = \underbrace{(D_1 f(\xi_1, x_2) - D_1 f(\xi_1, 0))}_{\text{hier MWS für } s \rightarrow D_1 f(\xi_1, s)} \cdot x_1;$$

weilers $\exists \xi_2$ mit $|\xi_2| \leq |x_2|$: $D_1 f(\xi_1, x_2) - D_1 f(\xi_1, 0) = D_2 D_1 f(\xi_1, \xi_2) \cdot x_2$.

Daher folgt nun mit der Notation $\xi = (\xi_1, \xi_2)$, $x = (x_1, x_2)$

$$(\star) \quad f(x) - f(x_1, 0) - f(0, x_2) + f(0) = F_{x_2}(x_1) - F_{x_2}(0) = D_2 D_1 f(\xi) \cdot x_1 x_2.$$

- Definiere nun $G_{x_1}: [-\delta, \delta] \rightarrow \mathbb{R}$ durch $G_{x_1}(s) := f(x_1, s) - f(0, s)$;

wiederum ergibt zweimalige Anwendung des Mittelwertsatzes der Differentialrechnung,
 dass $\exists \eta_2, \eta_1 \in \mathbb{R}$ mit $|\eta_2| \leq |x_2|$, $|\eta_1| \leq |x_1|$:

$$G_{x_1}(x_2) - G_{x_1}(0) = G'_{x_1}(\eta_2) x_2 = (D_2 f(x_1, \eta_2) - D_2 f(0, \eta_2)) x_2 = D_1 D_2 f(\eta_1, \eta_2) x_1 x_2,$$

d. h. mit $\eta = (\eta_1, \eta_2)$

$$(\star\star) \quad f(x) - f(0, x_2) - f(x_1, 0) + f(0) = D_1 D_2 f(\eta) \cdot x_1 x_2.$$

- Die linken Seiten von (\star) und $(\star\star)$ sind gleich, daher folgt nach Division durch $x_1 x_2$ zunächst

$$D_1 D_2 f(\eta) = D_2 D_1 f(\xi);$$

beachten wir noch, dass $0 \leq \|\xi\|_\infty \leq \|x\|_\infty$ und $0 \leq \|\eta\|_\infty \leq \|x\|_\infty$, dann folgt für $x \rightarrow 0$
 auch $\xi \rightarrow 0, \eta \rightarrow 0$; nach Voraussetzung sind $D_1 D_2 f$ und $D_2 D_1 f$ stetig, daher gilt

$$D_1 D_2 f(0) = D_2 D_1 f(0).$$

□

18.7. Korollar

Sei $f: \mathbb{R}^n \supseteq_{\text{offen}} U \rightarrow \mathbb{R}$ k -mal stetig partiell differenzierbar. Dann gilt für jede Permutation σ der Menge $\{1, \dots, k\}$ und für alle $j_1, \dots, j_k \in \{1, \dots, n\}$, dass

$$D_{j_1} D_{j_2} \cdots D_{j_k} f = D_{j_{\sigma(1)}} \cdots D_{j_{\sigma(k)}} f.$$

Beweis: Jede Permutation ist als Verknüpfung von Transpositionen darstellbar (siehe z.B. [Fis03, Lemma in 3.2.2]) und für jede Transposition ist Theorem 18.6 anwendbar. □

18.8. Bemerkung

Häufig benutzte Schreibweisen für höhere Ableitungen:

$$D_j D_l f = \frac{\partial^2 f}{\partial x_j \partial x_l}, \quad D_j D_j f = D_j^2 f = \left(\frac{\partial}{\partial x_j} \right)^2 f \quad \text{usw.}$$

oder allgemeiner mit Multiindex $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{N}^n$ und $D = (D_1, \dots, D_n)$ in der Form

$$D^\alpha f = D_1^{\alpha_1} D_2^{\alpha_2} \cdots D_n^{\alpha_n} f;$$

hierbei gibt dann $|\alpha| := \alpha_1 + \dots + \alpha_n$ die Ordnung der Ableitung an. (Kommt am Beginn von §19 ausführlicher.)

18.9. Beispiele

- 1.) Auf der Menge der zweimal stetig partiell differenzierbaren Funktionen $f: \mathbb{R}^n \underset{\text{offen}}{\supseteq} U \rightarrow \mathbb{R}$ ist der *Laplace-Operator*³ Δ definiert durch

$$\Delta f := \sum_{i=1}^n D_i^2 f = \operatorname{div}(\operatorname{grad} f) = \langle \nabla \mid \nabla f \rangle.$$

- 2.) Sei $U \subseteq \mathbb{R}^3$ offen, $v = (v_1, v_2, v_3): U \rightarrow \mathbb{R}^3$ ein Vektorfeld und v_j ($j = 1, 2, 3$) partiell differenzierbar. Dann definieren wir die *Rotation* von v als $\operatorname{rot}(v): U \rightarrow \mathbb{R}^3$ mit

$$(18.2) \quad \operatorname{rot}(v) := (D_2 v_3 - D_3 v_2, D_3 v_1 - D_1 v_3, D_1 v_2 - D_2 v_1) = \nabla \times v.$$

Ist speziell $v = \operatorname{grad} f$, wobei $f: U \rightarrow \mathbb{R}$ zweimal **stetig** partiell differenzierbar ist, dann gilt nach dem Satz von Schwarz

$$\operatorname{rot}(\operatorname{grad} f) = (D_2 D_3 f - D_3 D_2 f, \dots) = (0, 0, 0),$$

d.h. in diesem Fall

$$(18.3) \quad v = \operatorname{grad} f \implies \operatorname{rot}(v) = 0.$$

- 3.) Einige partielle Differentialgleichungen aus der Physik:

$$\Delta u = 0 \dots \text{Potentialgleichung für } u: \mathbb{R}^n \supseteq U \rightarrow \mathbb{R}$$

³Pierre-Simon Laplace (* 28. 3. 1749 Beaumont-en-Auge; †5. 3. 1827 Paris) [pjɛrsi'mɔ̃ la'plas], französischer Mathematiker und Astronom

$D_t^2 u - c^2 \cdot \Delta_x u = 0 \dots$ Wellengleichung für $u: \mathbb{R}^n \times \mathbb{R} \rightarrow \mathbb{R}$, $(x, t) \mapsto u(x, t)$

[$c = \text{const}$ hat die Rolle einer Geschwindigkeit]

$D_t u - k \cdot \Delta_x u = 0 \dots$ Wärmeleitungsgleichung [$k = \text{const} \dots$ Leitfähigkeit]

$D_t \psi - i \frac{\hbar}{2} \Delta_x \psi = 0 \dots$ Schrödinger-Gleichung für $\psi: \mathbb{R}^n \times \mathbb{R} \rightarrow \mathbb{C}$, $(x, t) \mapsto \psi(x, t)$

[$\hbar = h/2\pi$, h das Plancksche Wirkungsquantum]

4.) $r(x) = \|x\|$, $r: \mathbb{R}^n \setminus \{0\} \rightarrow \mathbb{R}$; wir haben in 18.4 bereits

$$\text{grad } r(x) = \frac{x}{r(x)} =: v(x)$$

und

$$\text{div}(v)(x) = \frac{n-1}{\|x\|}$$

berechnet. Zusammen ergibt das nun

$$\Delta r(x) = \frac{n-1}{r(x)}.$$

5.) $u(x) := \log \|x\|$, $u: \mathbb{R}^2 \setminus \{0\} \rightarrow \mathbb{R}$; wegen $\log'(t) = 1/t$ erhalten wir aus 18.4

$$\text{grad } u(x) = \frac{1}{\|x\|} \cdot \frac{x}{\|x\|} = \frac{x}{\|x\|^2};$$

nun ist für $j = 1, 2$ zunächst

$$D_j \left(\frac{x_j}{\|x\|^2} \right) = D_j \left(\frac{x_j}{x_1^2 + \dots + x_n^2} \right) = \frac{1}{\|x\|^2} - \frac{x_j \cdot 2x_j}{\|x\|^4},$$

und daraus durch Summenbildung

$$\begin{aligned} \Delta u(x) &= \text{div}(\text{grad } u)(x) = D_1 \left(\frac{x_1}{\|x\|^2} \right) + D_2 \left(\frac{x_2}{\|x\|^2} \right) \\ &= \frac{1}{\|x\|^2} - 2 \frac{x_1^2}{\|x\|^4} + \frac{1}{\|x\|^2} - 2 \frac{x_2^2}{\|x\|^4} = \frac{2}{\|x\|^2} - \frac{2\|x\|^2}{\|x\|^4} = 0. \end{aligned}$$

18.10. Differenzierbarkeit

Wir haben in 18.2 gesehen, dass partielle Differenzierbarkeit (das ist also Differenzierbarkeit entlang jeder Koordinatenrichtung) nicht einmal ausreicht, um die Stetigkeit einer Funktion mehrerer Variablen zu erzwingen. Das legt nahe, dass wir für einen geeigneten Differenzierbarkeitsbegriff im Höherdimensionalen gewissermaßen alle Richtungen „auf einmal“ berücksichtigen müssen.

In Analysis 1 wurde die Ableitung auch (gleichwertig) als lineare Approximation aufgefasst. Nun machen diesen Aspekt zum definierenden Prinzip:

Definition: Es sei $U \subseteq \mathbb{R}^n$ offen. Eine Abbildung $f: U \rightarrow \mathbb{R}^m$ heißt im Punkt $x \in U$ *differenzierbar*, falls es eine lineare Abbildung $A: \mathbb{R}^n \rightarrow \mathbb{R}^m$ gibt mit der Eigenschaft

$$(18.4) \quad \lim_{\substack{0 \neq \|\xi\| \rightarrow 0 \\ x+\xi \in U}} \frac{f(x+\xi) - f(x) - A \cdot \xi}{\|\xi\|} = 0.$$

Anders ausgedrückt: $\exists \delta > 0 \exists r: B_\delta(0) \rightarrow \mathbb{R}^m$ mit $\lim_{\xi \rightarrow 0} r(\xi) = 0$ und

$$(18.5) \quad f(x+\xi) = f(x) + A \cdot \xi + r(\xi) \|\xi\| \quad \forall \xi \in B_\delta(0)$$

[vgl. mit (18.4): $r(\xi) = \frac{f(x+\xi) - f(x) - A\xi}{\|\xi\|}$];

oder wiederum äquivalent dazu:

Es gibt $\delta > 0$ und $\varphi: B_\delta(0) \rightarrow \mathbb{R}^m$ mit $\varphi(\xi) = o(\|\xi\|)$ (für $\xi \rightarrow 0$) und

$$(18.5') \quad f(x+\xi) = f(x) + A \cdot \xi + \varphi(\xi).$$

Bemerkung: 1.) Wir identifizieren die lineare Abbildung A mit ihrer Matrixdarstellung bzgl. der Standardbasen in \mathbb{R}^n bzw. \mathbb{R}^m , d. h.

$$A \cdot \xi = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \cdot \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix},$$

wobei $a_{ij} \in \mathbb{R}$ ($1 \leq i \leq m$, $1 \leq j \leq n$). Dann lautet (18.4) in Komponenten gelesen: für $1 \leq i \leq m$ ist

$$\lim_{0 \neq \xi \rightarrow 0} \frac{f_i(x+\xi) - f_i(x) - \sum_{j=1}^n a_{ij} \xi_j}{\|\xi\|} = 0.$$

Insbesondere folgt für $\xi = h \cdot e_k$ mit $h \in \mathbb{R}$ die Relation

$$0 = \lim_{0 \neq h \rightarrow 0} \frac{f_i(x + he_k) - f_i(x) - h \cdot a_{ik}}{h} = \lim_{0 \neq h \rightarrow 0} \frac{f_i(x + he_k) - f_i(x)}{h} - a_{ik} = D_k f_i(x) - a_{ik};$$

d. h. f ist partiell differenzierbar in x und A ist durch die Bedingung (18.4) eindeutig bestimmt; es ist

$$(18.6) \quad A = \left(D_k f_i(x) \right)_{\substack{1 \leq i \leq m \\ 1 \leq k \leq n}} =: (Df)(x) \quad [(m \times n)\text{-Matrix}].$$

$(Df)(x)$ heißt *Ableitung*, *Differential* oder *Jacobi-Matrix*⁴ von f bei x .

2.) Für den Spezialfall $m = n = 1$ erhalten wir $(Df)(x) = f'(x)$ und somit bedeutet (18.4) genau die Differenzierbarkeit gemäß Analysis 1 (im Falle von Funktionen auf offenen Teilmengen von \mathbb{R}).

⁴Carl Gustav Jacob Jacobi (*10.12. 1804 Potsdam; †18.2. 1851 Berlin), deutscher Mathematiker

Proposition: Es sei $f: \mathbb{R}^n \underset{\text{offen}}{\supseteq} U \rightarrow \mathbb{R}^m$ differenzierbar in $x \in U$. Dann gilt:

- 1.) Für $i = 1, \dots, m$ ist f_i partiell differenzierbar in x .
- 2.) $(Df)(x) = \left(D_k f_i(x) \right)_{\substack{i=1, \dots, m \\ k=1, \dots, n}}$ ist die eindeutige lineare Abbildung mit Eigenschaft (18.4).
- 3.) f ist stetig in x .

Beweis: 1.) und 2.) wurden im Laufe der Herleitung von (18.6) gezeigt.

3.) Nach (18.5') gibt es $\varphi: B_\delta(0) \rightarrow \mathbb{R}^m$ mit $\varphi(\xi) = o(\|\xi\|)$ ($\xi \rightarrow 0$) und

$$f(x + \xi) = f(x) + A \cdot \xi + \varphi(\xi) \rightarrow f(x) + A \cdot 0 + 0 = f(x) \quad (\|\xi\| < \delta, \xi \rightarrow 0).$$

Also gilt

$$\lim_{\eta \rightarrow x} f(\eta) = \lim_{\xi \rightarrow 0} f(x + \xi) = f(x),$$

d. h. f ist stetig in x . □

Beispiele: 1.) Sei $f: \mathbb{R}^n \supseteq U \rightarrow \mathbb{R}$ differenzierbar in $x \in U$; dann ist $Df(x)$ also die $(1 \times n)$ -Matrix

$$(Df)(x) = (D_1 f(x), \dots, D_n f(x)) = {}^t(\text{grad } f(x)).$$

2.) Sei $C = (c_{ij})$ eine symmetrische reelle $(n \times n)$ -Matrix und $f: \mathbb{R}^n \rightarrow \mathbb{R}$ die zugehörige quadratische Form

$$f(x) := \langle x | Cx \rangle = {}^t x \cdot C \cdot x.$$

Es ist

$$\begin{aligned} f(x + \xi) &= \langle x + \xi | C(x + \xi) \rangle = \langle x | Cx \rangle + \langle \xi | Cx \rangle + \langle x | C\xi \rangle + \langle \xi | C\xi \rangle \\ &= f(x) + \langle 2Cx | \xi \rangle + \underbrace{\langle \xi | C\xi \rangle}_{\varphi(\xi)}, \end{aligned}$$

wobei nach der Cauchy-Schwarz-Ungleichung $\left| \frac{\varphi(\xi)}{\|\xi\|} \right| \leq \frac{\|C\xi\| \cdot \|\xi\|}{\|\xi\|} \leq \|C\|_{\text{op}} \cdot \|\xi\| \rightarrow 0$ für $\xi \rightarrow 0$.

Daher ist f differenzierbar in x und $(Df)(x) = {}^t(\text{grad } f(x)) = {}^t(2Cx)$.

(Alternativ könnten wir auch direkt berechnen: $D_k f(x) = D_k \left(\sum_{i=1}^n \sum_{j=1}^n c_{ij} x_i x_j \right) = \dots$)

18.11. Theorem

Es sei $f: \mathbb{R}^n \underset{\text{offen}}{\supseteq} U \rightarrow \mathbb{R}$ partiell differenzierbar und alle partiellen Ableitungen $D_k f$ ($k = 1, \dots, n$) seien stetig in $x \in U$. Dann ist f differenzierbar in x .

Bemerkung: Für $f: U \rightarrow \mathbb{R}^m$, $f = (f_1, \dots, f_m)$ gilt ein analoges Statement. Dabei ist die obige Aussage zunächst auf jede Komponentenfunktion f_j ($j = 1, \dots, m$) anwendbar und die Differenzierbarkeit von f folgt dann leicht mit Hilfe der Tatsache, dass sich die lineare Abbildung $Df(x): \mathbb{R}^n \rightarrow \mathbb{R}^m$ aus den m linearen Funktionalen $Df_1(x), \dots, Df_m(x): \mathbb{R}^n \rightarrow \mathbb{R}$ aufgebaut werden kann, nämlich als Zeilenvektoren in der entsprechenden Matrix (bezüglich der Standardbasen).

Beweis: $\exists \delta > 0: B_\delta(x) \subseteq U$; sei $\xi \in B_\delta(0)$ (somit gilt $x + \xi \in B_\delta(x)$).
Für $k = 1, \dots, n$ setzen wir

$$z^{(k)} := x + \sum_{j=1}^k \xi_j e_j \in B_\delta(x), \quad z^{(0)} := x;$$

dann ist $z^{(n)} = x + \xi$ und für $k \geq 1$ gilt $z^{(k)} - z^{(k-1)} = \xi_k e_k$.

Mit Hilfe der Notation für partielle Funktionen (eingeführt unmittelbar vor §18) können wir schreiben

$$f(z^{(k)}) - f(z^{(k-1)}) = \tilde{f}_{k, z^{(k-1)}}(\xi_k) - \tilde{f}_{k, z^{(k-1)}}(0).$$

Daher gibt es nach dem Mittelwertsatz (bzgl. der k -ten Variable) ein $\theta_k \in [0, 1]$:

$$f(z^{(k)}) - f(z^{(k-1)}) = D_k f(z^{(k-1)} + \theta_k \xi_k e_k) \cdot \xi_k;$$

als Abkürzung setzen wir $y^{(k)} := z^{(k-1)} + \theta_k \xi_k e_k$, dann erhalten wir

$$\begin{aligned} f(x + \xi) - f(x) &= \sum_{k=1}^n (f(z^{(k)}) - f(z^{(k-1)})) \\ &= \sum_{k=1}^n D_k f(y^{(k)}) \cdot \xi_k = \underbrace{\sum_{k=1}^n D_k f(x) \cdot \xi_k}_{\langle \text{grad } f(x) | \xi \rangle} + \underbrace{\sum_{k=1}^n (D_k f(y^{(k)}) - D_k f(x)) \cdot \xi_k}_{=:\varphi(\xi)}. \end{aligned}$$

Für $\xi \rightarrow 0$ folgt $y^{(k)} \rightarrow x$, somit wegen der Stetigkeit von $D_k f$ auch $D_k f(y^{(k)}) - D_k f(x) \rightarrow 0$ ($k = 1, \dots, n$). Daraus folgt

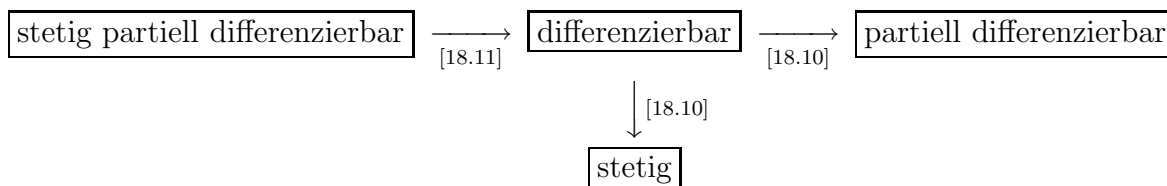
$$\frac{|\varphi(\xi)|}{\|\xi\|} \leq \sum_{k=1}^n \underbrace{|D_k f(y^{(k)}) - D_k f(x)|}_{\rightarrow 0} \underbrace{\frac{|\xi_k|}{\|\xi\|}}_{\text{beschränkt}} \rightarrow 0 \quad (\xi \rightarrow 0),$$

also die Differenzierbarkeit von f bei x . □

Korollar: Ist $f: \mathbb{R}^n \underset{\text{offen}}{\supseteq} U \rightarrow \mathbb{R}$ stetig partiell differenzierbar, dann ist f stetig auf U .

18.12. Bemerkung:

zusammenfassend erhalten wir folgende Übersicht



Vorsicht! Keiner der Implikationspfeile ist (im Allgemeinen) umkehrbar: so kennen wir schon aus Analysis 1 differenzierbare Funktionen, deren Ableitung nicht stetig ist; weiters fanden wir unter den Beispielen 18.2 Funktionen, die partiell differenzierbar, aber nicht einmal stetig sind, daher schon gar nicht differenzierbar; und stetige Funktionen, die nicht differenzierbar sind, kennen wir ebenfalls schon aus Analysis 1.

Wir werden im folgenden meistens die kürzere Sprechweise *stetig differenzierbar* statt *stetig partiell differenzierbar* verwenden. (Eine gewisse Rechtfertigung dafür kann man in obigem Schema sehen.)

18.13. Proposition (Kettenregel)

Seien $U \subseteq \mathbb{R}^n$, $V \subseteq \mathbb{R}^m$ offen und $f: U \rightarrow \mathbb{R}^m$, $g: V \rightarrow \mathbb{R}^p$ mit $f(U) \subseteq V$. Wenn f differenzierbar in $x \in U$ ist und g differenzierbar in $y = f(x) \in V$ ist, dann ist $g \circ f: U \rightarrow \mathbb{R}^p$ differenzierbar in x und es gilt

$$(18.7) \quad D(g \circ f)(x) = Dg(f(x)) \cdot Df(x)$$

(Auf der rechten Seite handelt es sich also um eine Verknüpfung linearer Abbildungen bzw. um eine entsprechende Matrizenmultiplikation.)

Beweis: Wir setzen $A := Df(x)$ und $B := Dg(y)$. Zu zeigen ist: $D(g \circ f)(x) = B \cdot A$.

(18.5') für f, g besagt: $\exists \varphi(\xi) = o(\|\xi\|)$, $\psi(\eta) = o(\|\eta\|)$, sodass (für $\|\xi\|$ und $\|\eta\|$ klein) gilt

$$\begin{aligned} f(x + \xi) &= f(x) + A\xi + \varphi(\xi) \\ g(y + \eta) &= g(y) + B\eta + \psi(\eta). \end{aligned}$$

Für $\eta = f(x + \xi) - f(x) = A\xi + \varphi(\xi)$ gilt wegen der Stetigkeit von f in x :
 $\|\xi\| \rightarrow 0 \implies \|\eta\| \rightarrow 0$.

Es ist

$$\begin{aligned}
 (g \circ f)(x + \xi) &= g(f(x + \xi)) = g(f(x) + \eta) = g(f(x)) + B \cdot \eta + \psi(\eta) \\
 &= g(f(x)) + B \cdot (A\xi + \varphi(\xi)) + \psi(A\xi + \varphi(\xi)) \\
 &= (g \circ f)(x) + (B \cdot A) \xi + \underbrace{B \cdot \varphi(\xi) + \psi(A\xi + \varphi(\xi))}_{=: \chi(\xi)}.
 \end{aligned}$$

Wir müssen zeigen, dass $\chi(\xi) = o(\|\xi\|)$ ($\xi \rightarrow 0$).

Es ist

$$\frac{\chi(\xi)}{\|\xi\|} = B \cdot \frac{\varphi(\xi)}{\|\xi\|} + \underbrace{\frac{\psi(A\xi + \varphi(\xi))}{\|\xi\|}}_{(*)},$$

wobei der erste Term auf der rechten Seite mit $\xi \rightarrow 0$ gegen Null strebt, weil (die lineare Abbildung) B stetig ist und $\varphi(\xi) = o(\|\xi\|)$ ist. Es bleibt also zu zeigen, dass der Ausdruck $(*)$ ebenfalls beliebig klein wird für $\xi \rightarrow 0$.

Zunächst ist für $\|\xi\|$ klein genug (weil $\varphi(\xi) = o(\|\xi\|)$ ist)

$$\|\eta\| = \|A\xi + \varphi(\xi)\| \leq \|A\|_{\text{op}} \|\xi\| + \frac{\|\varphi(\xi)\|}{\|\xi\|} \|\xi\| \leq (\|A\|_{\text{op}} + 1) \|\xi\|.$$

Wegen $\psi(\eta) = o(\|\eta\|)$ gilt auch $\psi(\eta) = \psi_1(\eta) \|\eta\|$ (für $\|\eta\|$ klein),

$$\text{wobei } \lim_{\substack{\eta \rightarrow 0 \\ \eta \neq 0}} \psi_1(\eta) = \lim_{\substack{\eta \rightarrow 0 \\ \eta \neq 0}} \frac{\psi(\eta)}{\|\eta\|} = 0.$$

Wenn $\delta > 0$ klein wird, so wird mit $\|\xi\| < \delta$ auch $\|\eta\| = \|A\xi + \varphi(\xi)\| \leq (\|A\|_{\text{op}} + 1) \cdot \delta$ klein. Daher ist dann

$$\|\underbrace{\psi(A\xi + \varphi(\xi))}_{\eta}\| = \|A\xi + \varphi(\xi)\| \cdot \|\psi_1(A\xi + \varphi(\xi))\| \leq (\|A\|_{\text{op}} + 1) \|\xi\| \|\psi_1(A\xi + \varphi(\xi))\|$$

und somit

$$\|(\ast)\| = \left\| \frac{\psi(A\xi + \varphi(\xi))}{\|\xi\|} \right\| \leq (\|A\|_{\text{op}} + 1) \cdot \|\psi_1(A\xi + \varphi(\xi))\| \rightarrow 0 \quad (\xi \rightarrow 0).$$

Insgesamt haben wir also gezeigt, dass $\lim_{\substack{\xi \rightarrow 0 \\ \xi \neq 0}} \frac{\chi(\xi)}{\|\xi\|} = 0$ gilt. □

18.14. Korollar

Es sei $f: \mathbb{R}^n \underset{\text{offen}}{\supseteq} U \rightarrow \mathbb{R}$ differenzierbar:

- 1.) Koordinatenwechsel: Sei $\varphi: \mathbb{R}^n \underset{\text{offen}}{\supseteq} V \rightarrow \mathbb{R}^n$ differenzierbar mit $\varphi(V) \subseteq U$. Dann gilt für die Ableitung von $g := f \circ \varphi$

$$(18.8) \quad Dg(y) = {}^t \text{grad } f(\varphi(y)) \cdot D\varphi(y) \quad \forall y \in V.$$

- 2.) Ableitung entlang von Kurven: Sei $\gamma: \mathbb{R} \underset{\text{offen}}{\supseteq} I \rightarrow \mathbb{R}^n$ differenzierbar mit $\gamma(I) \subseteq U$. Dann gilt

$$(18.9) \quad D(f \circ \gamma)(t) = {}^t \text{grad } f(\gamma(t)) \cdot \underbrace{\begin{pmatrix} \gamma'_1(t) \\ \vdots \\ \gamma'_n(t) \end{pmatrix}}_{D\gamma(t)=\dot{\gamma}(t)} = \langle \text{grad } f(\gamma(t)) \mid \dot{\gamma}(t) \rangle.$$

18.15. Beispiel

- 1.) Wenn $g, h: \mathbb{R} \rightarrow \mathbb{R}$ und $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ differenzierbar sind, dann gilt

$$\frac{d}{dx}(f(g(x), h(x))) = D_1 f(g(x), h(x)) \cdot g'(x) + D_2 f(g(x), h(x)) \cdot h'(x).$$

- 2.) Für $v \in S^{n-1}$ ($= \{y \in \mathbb{R}^n : \|y\| = 1\}$) und $x_0 \in \mathbb{R}^n$ stellt $\gamma: \mathbb{R} \rightarrow \mathbb{R}^n$, $\gamma(t) = x_0 + t \cdot v$ die Gerade durch x_0 in Richtung v dar.

Für jede differenzierbare Funktion $f: \mathbb{R}^n \rightarrow \mathbb{R}$ gilt dann

$$(f \circ \gamma)'(t) = D(f \circ \gamma)(t) = {}^t \text{grad } f(x_0 + tv) \cdot v = \langle \text{grad } f(x_0 + tv) \mid v \rangle;$$

speziell bei $t = 0$ ergibt sich $(f \circ \gamma)'(0) = \langle \text{grad } f(x_0) \mid v \rangle$.

18.16. Definition

Es sei $f: \mathbb{R}^n \underset{\text{offen}}{\supseteq} U \rightarrow \mathbb{R}$, $x \in U$ und $v \in S^{n-1}$. Die *Richtungsableitung* von f im Punkt x in Richtung v ist

$$(18.10) \quad D_v f(x) := \left. \frac{d}{dt}(f(x + tv)) \right|_{t=0} = \lim_{t \rightarrow 0} \frac{f(x + tv) - f(x)}{t} \quad (\text{falls existent}).$$

Spezialfall: für $v = e_j$ ($j \in \{1, \dots, n\}$) ergibt sich $D_{e_j}f(x) = D_jf(x)$, d.h. die j -te partielle Ableitung ist die Richtungsableitung bzgl. e_j .

Bemerkung: Aus der Existenz aller Richtungsableitungen (d.h. für alle $v \in S^{n-1}$) bei einem Punkt x folgt nicht die Differenzierbarkeit bei x ! (siehe Übungsaufgaben oder auch [BF96, 14.2. Beispiel 5])

18.17. Geometrische Bedeutung des Gradienten

Proposition: Sei $f: \mathbb{R}^n \underset{\text{offen}}{\supseteq} U \rightarrow \mathbb{R}$ differenzierbar; dann gilt:

1.) Für alle $v \in S^{n-1}$ existiert die Richtungsableitung und es gilt

$$D_v f(x) = \langle \text{grad } f(x) \mid v \rangle \quad \forall x \in U.$$

2.) Falls $\text{grad } f(x) \neq 0$, dann ist die Richtungsableitung im Punkt x maximal für die Richtung

$$v_0 := \frac{\text{grad } f(x)}{\|\text{grad } f(x)\|}.$$

d.h. $\text{grad } f(x)$ gibt die Richtung des stärksten Anstiegs von f bei x an.

Beweis: 1.) ist klar nach 18.15.

2.) $\forall v \in S^{n-1}$ gilt nach der Cauchy-Schwarz-Ungleichung

$$\langle \text{grad } f(x) \mid v \rangle \leq \|\text{grad } f(x)\| \cdot \|v\| = \|\text{grad } f(x)\|$$

und für $v = v_0$ erhalten wir

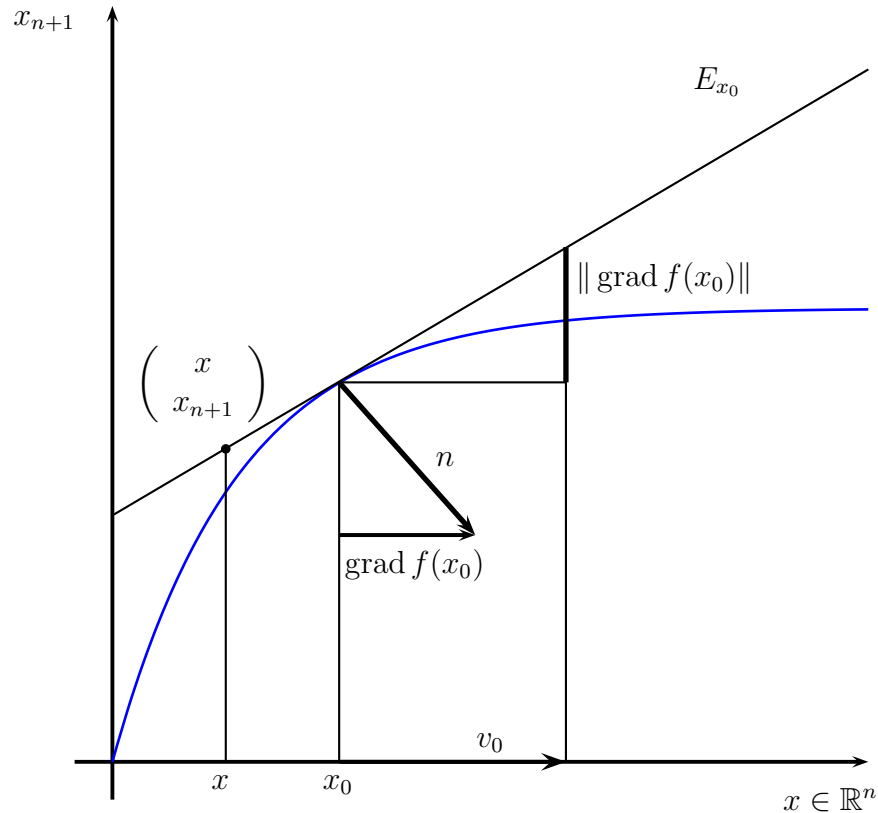
$$\langle \text{grad } f(x) \mid v_0 \rangle = \frac{\langle \text{grad } f(x) \mid \text{grad } f(x) \rangle}{\|\text{grad } f(x)\|} = \frac{\|\text{grad } f(x)\|^2}{\|\text{grad } f(x)\|} = \|\text{grad } f(x)\|.$$

□

Bemerkung: Das Maximum der Richtungsableitungen (in Punkten mit $\text{grad } f(x) \neq 0$) ist sogar eindeutig, weil in der Cauchy-Schwarz-Ungleichung Gleichheit bekanntlich genau dann gilt, wenn die beteiligten Vektoren linear abhängig sind (vgl. [Fis03, Kapitel 5]).

Tangentialhyperebene an den Graphen einer differenzierbaren Funktion:

Es sei $f: \mathbb{R}^n \underset{\text{offen}}{\supseteq} U \rightarrow \mathbb{R}$ differenzierbar und $x_0 \in U$. Wir betrachten den vertikalen Schnitt des Graphen $G(f)$ über der Geraden $t \mapsto x_0 + t \cdot v_0$, wobei $v_0 := \frac{\text{grad } f(x_0)}{\|\text{grad } f(x_0)\|}$ (bzw. $v_0 \in \mathbb{R}^n$ beliebig, falls $\text{grad } f(x_0) = 0$):



Die Funktion hat in dieser Schnittebene den größten Anstieg, nämlich $\|\text{grad } f(x_0)\|$, in (der) Richtung (über) $\text{grad } f(x_0)$. Ein Normalvektor auf die Tangentialhyperebene E_{x_0} über dem Punkt x_0 ist gegeben durch

$$n = \begin{pmatrix} \text{grad } f(x_0) \\ -1 \end{pmatrix}.$$

Daher ist nach der Hesseschen Normalform:

$$\begin{pmatrix} x \\ x_{n+1} \end{pmatrix} \in E_{x_0} \iff \left\langle n \mid \begin{pmatrix} x \\ x_{n+1} \end{pmatrix} - \begin{pmatrix} x_0 \\ f(x_0) \end{pmatrix} \right\rangle = 0,$$

was nach Zerlegung entsprechend (x, x_{n+1}) bedeutet, dass

$$\langle \text{grad } f(x_0) \mid x - x_0 \rangle - (x_{n+1} - f(x_0)) = 0,$$

oder nach kurzer Umformung

$$(*) \quad x_{n+1} = f(x_0) + \langle \text{grad } f(x_0) \mid x - x_0 \rangle.$$

Definition: Die Tangentialhyperebene E_{x_0} an den Graphen von f bei $(x_0, f(x_0))$ ist gegeben als Menge jener Punkte $(x, x_{n+1}) \in \mathbb{R}^{n+1}$, die (*) erfüllen. (Dies entspricht der linearen Approximation von f .)

Beispiel: $f(x, y) = \sqrt{1 - x^2 - y^2} = z$, $U = \{(x, y) : x^2 + y^2 < 1\} \ni (x_0, y_0)$ beliebig und $z_0 = f(x_0, y_0)$.

$\mathbb{R}^3 \supseteq G(f)$ ist die obere Halbsphäre (ohne Äquator) und $\text{grad } f(x_0, y_0) = -\left(\frac{x_0}{z_0}, \frac{y_0}{z_0}\right)$.

Für die Tangentialebene erhalten wir die Gleichung

$$z = z_0 - \frac{1}{z_0} \left\langle \begin{pmatrix} x_0 \\ y_0 \end{pmatrix} \mid \begin{pmatrix} x - x_0 \\ y - y_0 \end{pmatrix} \right\rangle$$

und nach Umformung die sogenannte *Spaltform* für die Sphäre $x^2 + y^2 + z^2 = 1$ als Tangentialebenengleichung

$$x x_0 + y y_0 + z z_0 = x_0^2 + y_0^2 + z_0^2 = 1.$$

18.18. Mittelwertsatz

Erinnere an Analysis 1: sei $I \subseteq \mathbb{R}$ ein Intervall, $f: I \rightarrow \mathbb{R}$ stetig differenzierbar, x und $x + \xi \in I$, dann folgt (durch Substitution $r = x + t\xi$)

$$f(x + \xi) - f(x) = \int_x^{x+\xi} f'(r) dr = \int_0^1 f'(x + t\xi) dt \cdot \xi.$$

Wir geben nun eine entsprechende Verallgemeinerung dieser Relation für $f: \mathbb{R}^n \supseteq U \rightarrow \mathbb{R}^m$ stetig differenzierbar entlang von geradlinigen Strecken innerhalb U . offen

Da nun $(Df)(x+t\xi)$ eine Matrix ist, brauchen wir also eine passende Version von $\int_a^b v(t) dt$, wobei $v: \mathbb{R} \supseteq I \rightarrow \mathbb{R}^p$ vektorwertig und stetig ist (z.B. $p = m \cdot n$ im Falle von $(m \times n)$ -Matrizen, weil der Vektorraum $M(m, n, \mathbb{R})$ all dieser isomorph zu $\mathbb{R}^{m \cdot n}$ ist.)

Ist $v: I \rightarrow \mathbb{R}^p$ stetig mit $v = (v_1, \dots, v_p)$, so definieren wir einfach für $a, b \in I$ das Integral komponentenweise

$$\int_a^b v(t) dt := \left(\int_a^b v_j(t) dt \right)_{j=1}^p$$

Lemma: Sei $v: [a, b] \rightarrow \mathbb{R}^p$ stetig, dann gilt

$$(18.11) \quad \left\| \int_a^b v(t) dt \right\| \leq \int_a^b \|v(t)\| dt.$$

Beweis: Setze $w := \int_a^b v(t) dt \in \mathbb{R}^p$, dann gilt:

$$\begin{aligned} \|w\|^2 = \langle w | w \rangle &= \left\langle \int_a^b v(t) dt \mid w \right\rangle \stackrel{\substack{\text{[Linearität} \\ \text{des Integrals]}}}{=} \int_a^b \langle v(t) \mid w \rangle dt \\ &\leq \int_a^b \|v(t)\| \cdot \|w\| dt = \|w\| \int_a^b \|v(t)\| dt. \end{aligned}$$

Nun ist im Fall $\|w\| = 0$ die Aussage des Lemmas trivial und für $w \neq 0$ folgt sie aus obiger Ungleichung nach Division durch $\|w\|$. \square

Proposition: Es sei $f: \mathbb{R}^n \supseteq U \rightarrow \mathbb{R}^m$ stetig differenzierbar und $x \in U$. Falls für $\xi \in \mathbb{R}^n$ gilt, dass $\{x + t\xi: t \in [0, 1]\} \subseteq U$ (d.h. die ganze Strecke von x nach $x + \xi$ liegt innerhalb U), dann folgt:

$$(18.12) \quad f(x + \xi) - f(x) = \left(\int_0^1 Df(x + t\xi) dt \right) \cdot \xi$$

und mit $M := \max\{\|Df(x + t\xi)\|_{\text{op}}: t \in [0, 1]\}$ weiters

$$(18.13) \quad \|f(x + \xi) - f(x)\| \leq M \cdot \|\xi\|.$$

(Es ist $M < \infty$, weil $t \mapsto \|Df(x + t\xi)\|_{\text{op}}$ stetig ist und $[0, 1]$ kompakt.)

Insbesondere muss f konstant (gleich einem fixen Vektor in \mathbb{R}^m) auf jeder Kugel innerhalb U sein, auf der $Df = 0$ gilt.

Beweis: Seien $f_j: U \rightarrow \mathbb{R}$ die Komponenten von f ($j \in \{1, \dots, m\}$) und setze $g_j: [0, 1] \rightarrow \mathbb{R}$, $g_j(t) := f_j(x + t\xi)$. Dann ist g_j stetig differenzierbar und

$$\begin{aligned} f_j(x + \xi) - f_j(x) = g_j(1) - g_j(0) &= \int_0^1 g_j'(t) dt = \int_0^1 (Df_j(x + t\xi) \cdot \xi) dt \\ &= \left(\int_0^1 Df_j(x + t\xi) dt \right) \cdot \xi. \end{aligned}$$

Es ist $Df = \begin{pmatrix} Df_1 \\ \vdots \\ Df_m \end{pmatrix}$ (mit Df_j als Zeilenvektoren), also gilt (18.12).

Setzen wir $v(t) := Df(x + t\xi) \cdot \xi$ in (18.12), dann folgt aus dem obigen Lemma:

$$\begin{aligned} \|f(x + \xi) - f(x)\| &= \left\| \int_0^1 Df(x + t\xi) \cdot \xi dt \right\| \leq \int_0^1 \|Df(x + t\xi) \cdot \xi\| dt \\ &\leq \int_0^1 \|Df(x + t\xi)\|_{\text{op}} dt \cdot \|\xi\| \leq \int_0^1 M \|\xi\| dt = M \|\xi\|. \end{aligned}$$

Falls $B_\delta(x_0) \subseteq U$ und $Df(y) = 0$ für jedes $y \in B_\delta(x_0)$ gilt, so folgt aus (18.12) (mit $\xi = y - x_0$)

$$f(y) - f(x_0) = 0 \quad \forall B_\delta(x_0).$$

□

18.19. Parameterintegrale

In Anwendungen aus der Physik, aber auch bei innermathematischen Fragen (z.B. bei Kurvenintegralen oder Fourier- und Laplace-Transformation), treffen wir häufig auf Funktionen mehrerer Variabler $(t, x_1, \dots, x_n) \mapsto g(t, x_1, \dots, x_n)$, die bezüglich einer Variable (bei festgehaltenen anderen Variablen) integriert werden: z.B. $(x_1, \dots, x_n) \in \mathbb{R}^n$ fix und $\int_a^b g(t, x_1, \dots, x_n) dt$. Die von der Integration nicht betroffenen Variablen x_1, \dots, x_n werden dann als so genannte („äußere“) Parameter aufgefasst und es stellt sich die Frage, welche Eigenschaften die resultierende Funktion

$$(x_1, \dots, x_n) \mapsto \int_a^b g(t, x_1, \dots, x_n) dt$$

bzgl. Stetigkeit oder Differenzierbarkeit hat.

Proposition: Es sei $[a, b] \subseteq \mathbb{R}$ ein kompaktes Intervall, $U \subseteq \mathbb{R}^n$ und $g: [a, b] \times U \rightarrow \mathbb{R}$ stetig. Wir definieren die Funktion $h: U \rightarrow \mathbb{R}$ durch

$$(18.14) \quad h(x) := \int_a^b g(t, x) dt.$$

dann gilt:

- 1.) h ist stetig
- 2.) Sei U offen. Ist für jedes $t \in [a, b]$ die Funktion $x \mapsto g(t, x)$, $U \rightarrow \mathbb{R}$ partiell differenzierbar und sind die partiellen Ableitungen $D_{x_j} g: [a, b] \times U \rightarrow \mathbb{R}$ stetig ($j = 1, \dots, n$),

dann folgt, dass h stetig differenzierbar auf U ist und für die partiellen Ableitungen von h gelten die Formeln

$$(18.15) \quad D_j h(x) = \int_a^b D_{x_j} g(t, x) dt \quad (j = 1, \dots, n).$$

[mit anderen „Worten“ $\frac{\partial}{\partial x_j} \int_a^b g(t, x_1, \dots, x_n) = \int_a^b \frac{\partial}{\partial x_j} g(t, x_1, \dots, x_n) dt$]

Beweis: Zunächst ist h wohldefiniert, weil für jedes $x \in U$ die Funktion $t \mapsto g(t, x)$ stetig und somit R-integrierbar ist.

ad 1.): Sei $\bar{x} \in U$ und (x_k) eine Folge in U mit $x_k \rightarrow \bar{x}$ ($k \rightarrow \infty$).

$Q := \{x_k \in U : k \in \mathbb{N}\} \cup \{\bar{x}\}$ ist kompakt in \mathbb{R}^n und $[a, b] \times Q$ ist kompakt in \mathbb{R}^{n+1} ; daher ist $g|_{[a, b] \times Q}$ gleichmäßig stetig, d. h.

$$(*) \quad \forall \varepsilon > 0 \exists \delta > 0 \forall (t, x), (t', x') \in [a, b] \times Q: \\ |t - t'| + \|x - x'\| < \delta \Rightarrow |g(t, x) - g(t', x')| < \varepsilon.$$

Wir setzen $g_k(t) := g(t, x_k)$, $g_k: [a, b] \rightarrow \mathbb{R}$ ($k \in \mathbb{N}$).

Behauptung: (g_k) konvergiert gleichmäßig gegen die Funktion \bar{g} , wobei $\bar{g}(t) = g(t, \bar{x})$.

Sei nämlich $\varepsilon > 0$, dann wähle $\delta > 0$ gemäß $(*)$ und $N \in \mathbb{N}$, sodass $\forall k \geq N$ gilt $\|x_k - \bar{x}\| < \delta$; wenden wir $(*)$ auf (t, x_k) und (t, \bar{x}) an, so folgt

$$|g_k(t) - \bar{g}(t)| = |g(t, x_k) - g(t, \bar{x})| < \varepsilon;$$

da diese Abschätzung für beliebiges $t \in [a, b]$ gilt, erhalten wir

$$\sup_{t \in [a, b]} |g_k(t) - \bar{g}(t)| \leq \varepsilon$$

und somit die Behauptung.

Wegen der gleichmäßigen Konvergenz der g_k gilt schließlich

$$\lim_{k \rightarrow \infty} h(x_k) = \lim_{k \rightarrow \infty} \int_a^b g_k(t) dt = \int_a^b \lim_{k \rightarrow \infty} g_k(t) dt = \int_a^b g(t, \bar{x}) dt = h(\bar{x}),$$

daher ist h stetig im Punkt $\bar{x} \in U$.

ad 2.): Sei $x \in U$, $1 \leq j \leq n$ und (h_k) eine Folge in \mathbb{R} mit $h_k \neq 0$, $h_k \rightarrow 0$; wir setzen

$$f_k^{(j)}(t) := \frac{g(t, x + h_k e_j) - g(t, x)}{h_k} \quad (k \in \mathbb{N}).$$

Behauptung: Die Folge $(f_k^{(j)})_{k \in \mathbb{N}}$ konvergiert gleichmäßig gegen \bar{g}_j , wobei $\bar{g}_j(t) = D_{x_j}g(t, x)$.

Sei $r > 0$ so gewählt, dass $\{h_k : k \in \mathbb{N}\} \subseteq [-r, r]$. Die Funktion $D_{x_j}g|_{[a, b] \times \overline{B_r(x)}}$ ist gleichmäßig stetig, d. h.

$$(**) \quad \forall \varepsilon > 0 \exists \delta > 0 \forall (t, y), (t', y') \in [a, b] \times \overline{B_r(x)} : \\ |t - t'| + \|y - y'\| < \delta \Rightarrow |D_{x_j}g(t, y) - D_{x_j}g(t', y')| < \varepsilon.$$

Nach dem Mittelwertsatz der Differentialrechnung (in der j -ten Variable) können wir für jedes k ein $\theta_k \in [0, 1]$ finden, sodass gilt:

$$f_k^{(j)}(t) = D_{x_j}g(t, x + \theta_k h_k e_j).$$

Wir wählen $N \in \mathbb{N}$ so, dass $\|h_k\| < \delta$ für alle $k \geq N$ gilt und wenden $(**)$ auf $(t, x + \theta_k h_k e_j)$ und (t, x) an (es ist ja $\|x + \theta_k h_k e_j - x\| = \theta_k \cdot \|h_k\| \leq \|h_k\| < \delta$). Dies ergibt

$$|f_k^{(j)}(t) - \bar{g}_j(t)| = |D_{x_j}g(t, x + \theta_k h_k e_j) - D_{x_j}g(t, x)| < \varepsilon,$$

und weiter, weil t beliebig war auch

$$\sup_{t \in [a, b]} |f_k^{(j)}(t) - \bar{g}_j(t)| \leq \varepsilon.$$

Also ist die Behauptung bewiesen.

Schließlich erlaubt die gleichmäßige Konvergenz folgende Umformungen

$$\lim_{k \rightarrow \infty} \frac{h(x + h_k e_j) - h(x)}{h_k} = \lim_{k \rightarrow \infty} \int_a^b f_k^{(j)}(t) dt = \int_a^b \bar{g}_j(t) dt = \int_a^b D_{x_j}g(t, x) dt.$$

□

Beispiel einer Anwendung auf Parameterintegrale mit variablen Grenzen: Wir wollen die Differenzierbarkeit eines Integralausdrucks der Form

$$\int_{\varphi(x)}^{\psi(x)} g(t, x) dt$$

bezüglich x untersuchen und die Ableitung bestimmen.

Wir vereinbaren folgende Annahmen: Es seien $\varphi, \psi: [c, d] \rightarrow \mathbb{R}$ differenzierbar. Die Funktion $g: [a, b] \times [c, d] \rightarrow \mathbb{R}$ sei stetig mit der Eigenschaft, dass für jedes $t \in [a, b]$ die partielle Funktion $x \mapsto g(t, x)$ differenzierbar ist und deren Ableitung $D_x g$ stetig auf $[a, b] \times [c, d]$ ist.

Behauptung: Für alle $x \in I :=]c, d[$ gilt

$$(18.16) \quad \frac{d}{dx} \int_{\varphi(x)}^{\psi(x)} g(t, x) dt = g(\psi(x), x) \cdot \psi'(x) - g(\varphi(x), x) \cdot \varphi'(x) + \int_{\varphi(x)}^{\psi(x)} D_x g(t, x) dt.$$

Zum *Beweis* betrachten wir die Hilfsfunktion $f: I^3 \rightarrow \mathbb{R}$, gegeben durch

$$f(x, y, z) := \int_y^z g(t, x) dt.$$

Dann ist also

$$\int_{\varphi(x)}^{\psi(x)} g(t, x) dt = f(x, \varphi(x), \psi(x)).$$

Die Funktion f ist stetig partiell differenzierbar, denn

- bei festem (x, y) ist $z \mapsto f(x, y, z)$ nach dem Hauptsatz der Differential- und Integralrechnung differenzierbar mit Ableitung

$$D_3 f(x, y, z) = g(z, x);$$

die Funktion $(x, y, z) \mapsto D_3 f(x, y, z)$ ist stetig, weil g es ist

- bei festem (x, z) ist $y \mapsto f(x, y, z)$ ebenso nach dem Hauptsatz der Differential- und Integralrechnung differenzierbar mit Ableitung

$$D_2 f(x, y, z) = -g(y, x);$$

die Funktion $(x, y, z) \mapsto D_2 f(x, y, z)$ ist stetig, weil g es ist

- bei festem (y, z) ist $x \mapsto f(x, y, z)$ nach der obigen Proposition differenzierbar mit Ableitung

$$D_1 f(x, y, z) = \int_y^z D_2 g(t, x) dt;$$

wir werden weiter unten den Beweis dafür nachreichen, dass auch die Funktion $(x, y, z) \mapsto D_1 f(x, y, z)$ stetig ist.

Somit ist f also differenzierbar und wir können die Kettenregel anwenden:

$$\begin{aligned} \frac{d}{dx} \left(f(x, \varphi(x), \psi(x)) \right) &= \\ &= D_1 f(x, \varphi(x), \psi(x)) + D_2 f(x, \varphi(x), \psi(x)) \cdot \varphi'(x) + D_3 f(x, \varphi(x), \psi(x)) \cdot \psi'(x) \\ &= \int_{\varphi(x)}^{\psi(x)} D_x g(t, x) dt - g(\varphi(x), x) \cdot \varphi'(x) + g(\psi(x), x) \cdot \psi'(x). \end{aligned}$$

Nun noch der Nachtrag zur Stetigkeit von D_1f : die Ableitung D_2g ist stetig auf der kompakten Menge $[a, b] \times [c, d] \subseteq \mathbb{R}^2$, daher gilt

$$\|D_2g\|_\infty = \sup_{\substack{a \leq t \leq b \\ c \leq x \leq d}} |g(t, x)| < \infty$$

und D_2g ist gleichmäßig stetig, d.h. $\forall \varepsilon > 0 \exists \delta > 0$, sodass

$$\forall (t, x), (t', x') \in [a, b] \times [c, d]: \quad |t - t'| + \|x - x'\| < \delta \quad \Rightarrow \quad |D_2g(t, x) - D_2g(t', x')| < \varepsilon.$$

Nun sei $(x_0, y_0, z_0) \in I^3$ und $\varepsilon > 0$ beliebig. Wir wählen $\delta > 0$ wie oben und zusätzlich kleiner als ε , dann folgt für $(x, y, z) \in I^3$ mit $|x - x_0| + |y - y_0| + |z - z_0| < \delta$ (mittels Einschließen geeigneter Terme)

$$\begin{aligned} |D_1f(x, y, z) - D_1f(x_0, y_0, z_0)| &= \left| \int_y^z D_2g(t, x) dt - \int_{y_0}^{z_0} D_2g(t, x_0) dt \right| \\ &= \left| \int_y^z D_2g(t, x) dt - \int_y^{z_0} D_2g(t, x) dt + \int_y^{z_0} D_2g(t, x) dt - \int_{y_0}^{z_0} D_2g(t, x) dt \right. \\ &\quad \left. + \int_{y_0}^{z_0} D_2g(t, x) dt - \int_{y_0}^{z_0} D_2g(t, x_0) dt \right| \\ &\leq \left| \int_{z_0}^z D_2g(t, x) dt \right| + \left| \int_y^{y_0} D_2g(t, x) dt \right| + \left| \int_{y_0}^{z_0} (D_2g(t, x) - D_2g(t, x_0)) dt \right| \\ &\leq |z - z_0| \cdot \|D_2g\|_\infty + |y_0 - y| \cdot \|D_2g\|_\infty + |z_0 - y_0| \cdot \varepsilon \leq (2\|D_2g\|_\infty + |z_0 - y_0|) \cdot \varepsilon \end{aligned}$$

und somit die Stetigkeit von D_1f bei (x_0, y_0, z_0) .

§19. Taylor-Formel und lokale Extrema

19.1. Höhere Ableitungen

Vorbemerkung: Es seien V_1, V_2, V_3 endlichdimensionale Vektorräume über demselben Grundkörper. Dann gilt

$$L(V_1, L(V_2, V_3)) \cong BL(V_1 \times V_2, V_3),$$

wobei wir allgemein mit $L(V, W)$ den Vektorraum der linearen Abbildungen $V \rightarrow W$ bezeichnen und mit $BL(V_1 \times V_2, V_3)$ den Raum der bilinearen Abbildungen $V_1 \times V_2 \rightarrow V_3$ (vgl. [AE99, Abschnitt VII.4], wo dies sogar etwas allgemeiner für stetige multilineare Abbildungen auf Banachräumen bewiesen wird). [Für einen Beweis betrachte die Abbildung $L(V_1, L(V_2, V_3)) \rightarrow BL(V_1 \times V_2, V_3)$, $A \mapsto \beta_A$ mit $\beta_A(v_1, v_2) := (A \cdot v_1) \cdot v_2$ bzw. als Inverse $\beta \mapsto (v \mapsto \beta(v, \cdot))$.]

Es sei $f : \mathbb{R}^n \underset{\text{offen}}{\supseteq} U \rightarrow \mathbb{R}^m$ eine N -mal stetig (partiell) differenzierbare Abbildung, wobei $N \geq 2$; für $1 \leq k \leq n$, $1 \leq l \leq m$ ist dann $D_k f_l : U \rightarrow \mathbb{R}$ stetig (partiell) differenzierbar.

Daher ist die Abbildung $Df : U \rightarrow L(\mathbb{R}^n, \mathbb{R}^m) \cong M(m, n, \mathbb{R}) \cong \mathbb{R}^{m \cdot n}$ differenzierbar und $\forall x \in U: D(Df)(x) \in L(\mathbb{R}^n, L(\mathbb{R}^n, \mathbb{R}^m))$.

Somit ist $D^2 f(x) := D(Df)(x)$ auffassbar als bilineare Abbildung $\mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}^m$; indem wir so für $p \leq N$ sukzessive fortfahren, erhalten wir das p -fache Differential $D^p f(x) = D(\cdots(Df))(x)$ als p -lineare Abbildung $(\mathbb{R}^n)^p \rightarrow \mathbb{R}^m$ (vgl. [AE99, VII.5]).

Wichtigster Spezialfall $N = 2, m = 1$: Es ist also $f : \mathbb{R}^n \supseteq U \rightarrow \mathbb{R}$ zweimal stetig differenzierbar. Dann ist für $x \in U$ das zweite Differential $D^2 f(x) : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$ eine Bilinearform mit Wirkung $D^2 f(x)(v, w) := (D(Df)(x) \cdot v) \cdot w$ für alle $v, w \in \mathbb{R}^n$.

$$\begin{aligned} (D(Df))(x) \cdot v &\underset{\text{[Lin. der Abl.]}}{=} D(Df \cdot v)(x) = D \left(\underbrace{\sum_{k=1}^n D_k f \cdot v_k}_{\text{als Abb. } F:U \rightarrow \mathbb{R}} \right) (x) \\ &= (D_1 F, \dots, D_n F)(x) = \left(\sum_k D_1 D_k f(x) \cdot v_k, \dots, \sum_k D_n D_k f(x) \cdot v_k \right), \end{aligned}$$

d. h. mit $w = {}^t(w_1, \dots, w_n)$ ist

$$D^2 f(x)(v, w) = \sum_{j=1}^n \sum_{k=1}^n D_j D_k f(x) \cdot v_k w_j = {}^t v \cdot \left(D_j D_k f(x) \right)_{1 \leq j, k \leq n} \cdot w.$$

Nach dem Satz von Schwarz ist $D_j D_k f = D_k D_j f$, daher folgt

$$D^2 f(x)(v, w) = D^2 f(x)(w, v),$$

also ist $D^2 f(x)$ symmetrisch und wird (als quadratische Form auf \mathbb{R}^n) durch die Matrix $(D_j D_k f(x))_{1 \leq j, k \leq n}$ gegeben.

Definition: Es sei $f : \mathbb{R}^n \underset{\text{offen}}{\supseteq} U \rightarrow \mathbb{R}$ zweimal stetig differenzierbar, dann heißt

$$(19.1) \quad H_f(x) := (D_j D_k f(x))_{1 \leq j, k \leq n}$$

*Hesse-Matrix*¹ von f bei x . (Die Hesse-Matrix ist symmetrisch.)

19.2. Multiindex-Notation

Es sei $\alpha \in \mathbb{N}^n$, d.h.. ein n -Tupel $\alpha = (\alpha_1, \dots, \alpha_n)$ mit $\alpha_j \in \mathbb{N}$ ($j = 1, \dots, n$). Im Zusammenhang mit der Differentialrechnung mehrerer Variablen wird so ein n -tupel als sogenannter Multiindex verwendet.

Zunächst heißt

$$|\alpha| := \alpha_1 + \dots + \alpha_n$$

die Ordnung oder Länge von α und zur vereinfachten Notation mit kombinatorischen Faktoren wird

$$\alpha! := \alpha_1! \cdot \alpha_2! \cdot \dots \cdot \alpha_n!$$

gesetzt.

Für $x = (x_1, \dots, x_n) \in \mathbb{R}^n$ führen wir die Notation

$$x^\alpha := x_1^{\alpha_1} \cdot x_2^{\alpha_2} \cdot \dots \cdot x_n^{\alpha_n} \quad (\in \mathbb{R})$$

ein.

Wenn f eine N -mal stetig differenzierbare Funktion in n Variablen ist und $|\alpha| \leq N$, dann ist

$$D^\alpha f := D_1^{\alpha_1} \cdot D_2^{\alpha_2} \cdot \dots \cdot D_n^{\alpha_n} f$$

wobei die Reihenfolge der partiellen Ableitungen gemäß Korollar 18.7 beliebig ist.

¹Ludwig Otto Hesse (*22. 4. 1811 Königsberg; †4. 8. 1874 München), deutscher Mathematiker

19.3. Lemma

Es sei $N \in \mathbb{N}$, $N \geq 1$ und $f : \mathbb{R}^n \supseteq U \rightarrow \mathbb{R}$ N -mal stetig (partiell) differenzierbar. Weiters sei $x \in U$ und für $\xi \in \mathbb{R}^n$ gelte $\{x + t\xi : t \in [0, 1]\} \subseteq U$.

Dann ist die Funktion $g : [0, 1] \rightarrow \mathbb{R}$, $g(t) := f(x + t\xi)$, N -mal stetig differenzierbar und

$$(19.2) \quad g^{(N)}(t) = \sum_{\substack{\alpha \in \mathbb{N}^n \\ |\alpha| = N}} \frac{N!}{\alpha!} (D^\alpha f)(x + t\xi) \cdot \xi^\alpha.$$

Beweis: Wir zeigen zunächst induktiv die Relation

$$(*)_N \quad g^{(N)}(t) = \sum_{j_1, \dots, j_N=1}^n D_{j_N} \cdots D_{j_1} f(x + t\xi) \cdot \xi_{j_1} \cdots \xi_{j_N}.$$

$N = 1$: aus der Kettenregel folgt $g'(t) = Df(x + t\xi) \cdot \xi = \sum_{j=1}^n D_j f(x + t\xi) \cdot \xi_j$, also $(*)_1$;

$$N \mapsto N + 1: g^{(N+1)}(t) = (g^{(N)})'(t) = \sum_{j_1, \dots, j_N=1}^n \underbrace{(D_{j_N} \cdots D_{j_1} f(x + t\xi))'}_{\sum_{j=1}^n D_j D_{j_N} \cdots D_{j_1} f(x + t\xi) \cdot \xi_j} \cdot \xi_{j_1} \cdots \xi_{j_N};$$

wenn wir nun $j = j_{N+1}$ setzen, ergibt das $(*)_{N+1}$.

Wir sortieren nun die Mehrfachsumme in $(*)_N$ um und nützen also aus, dass die Reihenfolge der partiellen Ableitungen $D_{j_N} \cdots D_{j_1}$ beliebig abgeändert werden darf.

Es ist

$$D_{j_N} \cdots D_{j_1} = D_1^{\alpha_1} \cdots D_n^{\alpha_n},$$

wobei für $l = 1, \dots, n$:

$$0 \leq \alpha_l := \underbrace{|\{k \in \{1, \dots, N\} : j_k = l\}|}_{=: A_l(j_1, \dots, j_N)}$$

(α_l gibt an, wie oft l unter j_1, \dots, j_N vorkommt) und mit $\alpha := (\alpha_1, \dots, \alpha_n)$

$$|\{(j_1, \dots, j_N) \in \{1, \dots, n\}^N : |A_l(j_1, \dots, j_N)| = \alpha_l \quad (l = 1, \dots, n)\}| = \frac{N!}{\alpha_1! \cdots \alpha_n!} = \frac{N!}{\alpha!}$$

(Modell: N Zahlen aus $\{1, \dots, n\}$ auf n Kästchen aufteilen, wobei α_1 in das erste Kästchen kommen usw.)

Weiters ist $\xi_{j_1} \cdots \xi_{j_N} = \xi^\alpha$, somit ergibt $(*)_N$ also die Gleichung (19.2). \square

19.4. Theorem (Taylor-Formel)

Die Funktion $f : \mathbb{R}^n \supseteq_{\text{offen}} U \rightarrow \mathbb{R}$ sei $(N+1)$ -mal stetig (partiell) differenzierbar und $x \in U$, $\xi \in \mathbb{R}^n$ so, dass $\{x + t\xi : t \in [0, 1]\} \subseteq U$. Dann gibt es ein $\theta \in [0, 1]$, mit dem gilt

$$(19.3) \quad f(x + \xi) = \sum_{|\alpha| \leq N} \frac{D^\alpha f(x)}{\alpha!} \cdot \xi^\alpha + \sum_{|\alpha|=N+1} \frac{D^\alpha f(x + \theta\xi)}{\alpha!} \cdot \xi^\alpha.$$

Beweis: Wir betrachten die Hilfsfunktion $g : [0, 1] \rightarrow \mathbb{R}$, $g(t) := f(x + t\xi)$. Dann ist g ebenfalls $(N+1)$ -mal stetig differenzierbar und die Taylor-Formel in der Variablen t ergibt: $\exists \theta \in [0, 1]$ so, dass

$$g(1) = \sum_{k=0}^N \frac{g^{(k)}(0)}{k!} 1^k + \frac{g^{(N+1)}(\theta)}{(N+1)!} \stackrel{[19.3]}{=} \underbrace{\sum_{k=0}^N \sum_{|\alpha|=k} \frac{D^\alpha f(x)}{\alpha!}}_{\sum_{|\alpha| \leq N}} \cdot \xi^\alpha + \sum_{|\alpha|=N+1} \frac{D^\alpha f(x + \theta\xi)}{\alpha!} \cdot \xi^\alpha.$$

□

19.5. Korollar

Sei $f : \mathbb{R}^n \supseteq_{\text{offen}} U \rightarrow \mathbb{R}$ eine N -mal stetig differenzierbare Funktion. Dann gilt $\forall x \in U$:

$$(19.4) \quad f(x + \xi) = \sum_{|\alpha| \leq N} \frac{D^\alpha f(x)}{\alpha!} \cdot \xi^\alpha + o(\|\xi\|^N) \quad (\xi \rightarrow 0).$$

Beweis: Zu $x \in U$ existiert $\delta > 0$ mit $B_\delta(x) \subseteq U$.

$\forall \xi \in B_\delta(0) \exists \theta(\xi) \in [0, 1]$:

$$\begin{aligned} f(x + \xi) &= \sum_{|\alpha| \leq N-1} \frac{D^\alpha f(x)}{\alpha!} \xi^\alpha + \sum_{|\alpha|=N} \frac{D^\alpha f(x + \theta\xi)}{\alpha!} \xi^\alpha \\ &= \sum_{|\alpha| \leq N} \frac{D^\alpha f(x)}{\alpha!} \xi^\alpha + \sum_{|\alpha|=N} \underbrace{\frac{D^\alpha f(x + \theta\xi) - D^\alpha f(x)}{\alpha!}}_{=: r_\alpha(\xi)} \cdot \xi^\alpha. \end{aligned}$$

Für $|\alpha| \leq N$ ist $D^\alpha f$ stetig; wegen $|\theta(\xi)| \leq 1$ folgt $x + \theta(\xi)\xi \rightarrow x$ für $\|\xi\| \rightarrow 0$ und weiter $\lim_{\xi \rightarrow 0} r_\alpha(\xi) = 0$.

Außerdem ist $|\xi^\alpha| = |\xi_1|^{\alpha_1} \cdots |\xi_n|^{\alpha_n} \leq \|\xi\|^{\alpha_1} \cdots \|\xi\|^{\alpha_n} = \|\xi\|^{\alpha_1 + \cdots + \alpha_n} = \|\xi\|^{|\alpha|}$,

daher also $\sum_{|\alpha|=N} r_\alpha(\xi) \xi^\alpha = o(\|\xi\|^N)$ ($\xi \rightarrow 0$). □

19.6. Spezialfall: Approximation 2. Ordnung

Sei $f : \mathbb{R}^n \underset{\text{offen}}{\supseteq} U \rightarrow \mathbb{R}$ zweimal stetig differenzierbar, dann gilt $\forall x \in U$:

$$(19.5) \quad f(x + \xi) = f(x) + \langle \text{grad } f(x) \mid \xi \rangle + \frac{1}{2} \langle H_f(x) \xi \mid \xi \rangle + o(\|\xi\|^2) \quad (\xi \rightarrow 0).$$

Beweis: Wir verwenden Korollar 19.5 mit $N = 2$:

$$f(x + \xi) = f(x) + \underbrace{\sum_{|\alpha|=1} \frac{D^\alpha f(x)}{\alpha!} \xi^\alpha}_{\textcircled{1}} + \underbrace{\sum_{|\alpha|=2} \frac{D^\alpha f(x)}{\alpha!} \xi^\alpha}_{\textcircled{2}} + o(\|\xi\|^2).$$

ad $\textcircled{1}$: weil $|\alpha| = 1$ gibt es $l \in \{1, \dots, n\}$: $\alpha_l = 1$ und $\alpha_j = 0$ für $j \neq l$ (d.h. $\alpha = e_l$).

Daher ist $D^\alpha = D_l$, $\alpha! = 1$, $\xi^\alpha = \xi_l$, also einfach

$$\textcircled{1} = \sum_{l=1}^n D_l f(x) \cdot \xi_l = \langle \text{grad } f(x) \mid \xi \rangle.$$

ad $\textcircled{2}$: wir haben $|\alpha| = 2$ und unterscheiden 2 Fälle:

- $\alpha = 2 \cdot e_j = (0, \dots, \underbrace{2}_{j\text{-te Komponente}}, 0, \dots, 0)$, d.h. $D^\alpha = D_j^2$ und $\alpha! = 2! = 2$
- $\alpha = e_l + e_j = (0, \dots, 0, \underbrace{1}_{l\text{-te Stelle}}, 0, \dots, 0, \underbrace{1}_{j\text{-te Stelle}}, 0, \dots, 0)$ mit $l < j$,
d.h. $D^\alpha = D_l D_j = D_j D_l$, $\alpha! = 1$ und $\xi^\alpha = \xi_l \cdot \xi_j$.

Also ist

$$\begin{aligned} \textcircled{2} &= \sum_{j=1}^n \frac{D_j^2 f(x)}{2} \xi_j^2 + \underbrace{\sum_{1 \leq l < j \leq n} \frac{D_j D_l f(x)}{1} \xi_j \xi_l}_{= \frac{1}{2} \sum_{j \neq l}} \\ &= \frac{1}{2} \sum_{l,j} D_j D_l f(x) \xi_l \xi_j = \frac{1}{2} {}^t \xi \cdot \underbrace{(D_j D_l f(x))}_{1 \leq j, l \leq n} \cdot \xi. \end{aligned}$$

□

19.7. Definition

Es sei $f : \mathbb{R}^n \supseteq U \rightarrow \mathbb{R}$. Ein Punkt $x \in U$ heißt *lokales Maximum* (bzw. *Minimum*) von f , falls gilt: Es gibt eine Umgebung V von x , $V \subseteq U$, sodass gilt:

$$\forall y \in V : f(x) \geq f(y) \quad (\text{bzw. } f(x) \leq f(y)).$$

Wir nennen x ein *lokales Extremum*, falls es ein lokales Maximum oder Minimum ist.

Ein Extremum ist *strikt*, falls in obiger Bedingung zusätzlich $\forall y \in V \setminus \{x\}$ gilt $f(x) > f(y)$ (bzw. $f(x) < f(y)$).

19.8. Proposition

Sei $f : \mathbb{R}^n \underset{\text{offen}}{\supseteq} U \rightarrow \mathbb{R}$ partiell differenzierbar und $x \in U$ ein lokales Extremum von f .

Dann gilt $\text{grad } f(x) = 0$

Beweis: Für $j = 1, 2, \dots, n$ betrachten wir die partiellen Funktionen $g_j : t \mapsto f(x + te_j) = \tilde{f}_{j,x}(t)$ (vgl. Einleitung zu Kapitel VII, unmittelbar vor §18); g_j ist definiert auf einem offenen Intervall $0 \in I_j \subseteq \mathbb{R}$ und ist laut Voraussetzung differenzierbar.

Da x ein lokales Extremum für f ist, folgt, dass g_j ein lokales Extremum bei $t = 0$ besitzt. Daher gilt $0 = g'_j(0) = D_j f(x)$. \square

19.9. Bemerkung und Vorbetrachtung

Wie wir schon aus Analysis 1 (d.i. also der Fall $n = 1$) wissen, kann die Bedingung $\text{grad } f(x) = 0$ im Allgemeinen nicht hinreichend sein!

Nun sei f zweimal stetig differenzierbar und besitze ein lokales Extremum bei x . Die quadratische Approximation an f gemäß (19.5) nahe x lautet dann:

$$\begin{aligned} \underbrace{f(x + \xi)}_y &= f(x) + \langle 0 \mid \xi \rangle + \frac{1}{2} \underbrace{\langle H_f(x) \cdot \xi \mid \xi \rangle}_A + o(\|\xi\|^2) \\ &\approx f(x) + \frac{1}{2} \cdot \underbrace{\langle A \cdot (y - x) \mid (y - x) \rangle}_{(y \neq x)} \begin{cases} > 0, & \text{falls } A \text{ positiv definit} \\ < 0, & \text{falls } A \text{ negativ definit} \end{cases} \end{aligned}$$

Aus dieser Heuristik heraus erwarten wir also folgendes:

$H_f(x)$ positiv definit \implies lokales Minimum
 $H_f(x)$ negativ definit \implies lokales Maximum.

Man vergleiche dies wiederum mit dem Fall einer Variable ($n = 1$), in dem z.B. gilt:
 $f''(x) > 0 \implies$ lokales Minimum.

19.10. (Wiederholung) aus der Linearen Algebra

Eine reelle symmetrische ($n \times n$)-Matrix A heißt:

- 1.) *positiv definit*, falls $\forall \xi \in \mathbb{R}^n, \xi \neq 0: \langle A\xi | \xi \rangle > 0$
- 2.) *positiv semidefinit*, falls $\forall \xi \in \mathbb{R}^n: \langle A\xi | \xi \rangle \geq 0$
- 3.) *negativ definit*, falls $-A$ positiv definit ist
- 4.) *negativ semidefinit*, falls $-A$ positiv semidefinit ist
- 5.) *indefinit*, falls gilt $\exists \xi, \eta \in \mathbb{R}^n$ mit $\langle A\xi | \xi \rangle > 0$ und $\langle A\eta | \eta \rangle < 0$.

Da A symmetrisch ist, gibt es eine Orthonormalbasis $\mathcal{B} = \{v_1, \dots, v_n\}$ aus Eigenvektoren;

alle Eigenwerte $\lambda_1 \dots \lambda_n$ sind reell und $\forall \xi \in \mathbb{R}^n$: $A\xi = \sum_{j=1}^n \lambda_j \underbrace{\langle \xi | v_j \rangle}_{j\text{-te Komponente von } \xi \text{ bzgl. } \mathcal{B}} v_j$;

d.h. A ist ähnlich zur Diagonalmatrix $\begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix}$.

Somit ist $\langle A\xi | \xi \rangle = \sum_{j=1}^n \lambda_j \langle \xi | v_j \rangle^2$ und daher gilt:

- (a) A positiv definit $\iff \forall j: \lambda_j > 0$
- (b) A positiv semidefinit $\iff \forall j: \lambda_j \geq 0$
- (c) A negativ definit $\iff \forall j: \lambda_j < 0$
- (d) A negativ semidefinit $\iff \forall j: \lambda_j \leq 0$
- (e) A indefinit $\iff \exists j_1, j_2: \lambda_{j_1} < 0 < \lambda_{j_2}$

Ein Nachteil dieser Kriterien ist allerdings, dass man zu deren Anwendung (außer eventuell im Fall (e)) alle Eigenwerte kennen muss.

Lemma zum Spezialfall $n = 2$: Es sei $A = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$ eine reelle (symmetrische) (2×2)-Matrix, dann gilt:

- 1.) A indefinit $\iff \det A < 0$
- 2.) A positiv definit $\iff \det A > 0$ und $a > 0$
- 3.) A negativ definit $\iff \det A > 0$ und $a < 0$

Bemerkung: Allgemein gilt für symmetrische $(n \times n)$ -Matrizen das folgende Kriterium mittels der sogenannten *Hauptminoren*: sei für $k = 1, \dots, n$ eine $(k \times k)$ -Teilmatrix von A definiert durch $A_k := (a_{ij})_{1 \leq i, j \leq k}$ (die linken oberen k Reihen und Spalten in A), dann gilt (siehe [Fis03, 5.7.7])

$$A \text{ positiv definit} \iff \text{für } k = 1, \dots, n : \det A_k > 0.$$

Beweis: Wir schreiben $z = \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} \in \mathbb{R}^2$

ad 2.) \Leftrightarrow setze $z = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, dann folgt $0 < {}^t z A z = a$;

$$\begin{aligned} \text{Setze } z = \begin{pmatrix} -\frac{b}{a} \\ 1 \end{pmatrix}, \text{ dann folgt } 0 < {}^t z A z &= a \left(-\frac{b}{a}\right)^2 + 2b \left(-\frac{b}{a} \cdot 1\right) + c = \\ &= \frac{b^2}{a} - 2 \cdot \frac{b^2}{a} + c = \frac{ac - b^2}{a} = \frac{\det A}{a} \Rightarrow \det A > 0; \end{aligned}$$

$$\Leftrightarrow {}^t z A z = az_1^2 + 2bz_1z_2 + cz_2^2 = a \left(z_1 + \frac{b}{a}z_2\right)^2 + \underbrace{\left(c - \frac{b^2}{a}\right)}_{\frac{ac - b^2}{a} > 0} z_2^2 \geq 0;$$

$${}^t z A z = 0 \Rightarrow z_2 = 0 \text{ und } 0 = z_1 + \frac{b}{a}z_2 = z_1 \Rightarrow z = 0.$$

ad 3.) Wie in 2.) für $-A$; beachte $\det(-A) = (-1)^2 \cdot \det A = \det A$.

$$\text{ad 1.) } \det \begin{pmatrix} a - \lambda & b \\ b & c - \lambda \end{pmatrix} = (a - \lambda)(c - \lambda) - b^2 = \lambda^2 - \underbrace{(a + c)}_{s := \text{Spur}(A)} \lambda + \underbrace{ac - b^2}_{\det A}$$

hat als Nullstellen $\lambda_{\pm} = \frac{s}{2} \pm \sqrt{\left(\frac{s}{2}\right)^2 - \det A}$; setze $\Delta := \left(\left(\frac{s}{2}\right)^2 - \det A\right)$, dann gilt

$$\det A < 0 \iff \Delta > \left(\frac{s}{2}\right)^2 \iff \lambda_- < 0 < \lambda_+ \iff A \text{ indefinit.}$$

□

19.11. Hinreichende Bedingungen für lokale Extrema

Theorem: Sei $f : \mathbb{R}^n \underset{\text{offen}}{\supseteq} U \rightarrow \mathbb{R}$ zweimal stetig differenzierbar und $x \in U$ mit $\text{grad } f(x) = 0$ (man sagt: x ist eine *kritische Stelle*). Dann gilt:

- 1.) $H_f(x)$ positiv definit $\implies x$ ist striktes lokales Minimum
- 2.) $H_f(x)$ negativ definit $\implies x$ ist striktes lokales Maximum
- 3.) $H_f(x)$ indefinit $\implies x$ ist kein lokales Extremum
(im Fall $n = 2$ spricht man hier von einem Sattelpunkt)

Bemerkung: Die Bedingungen im Theorem sind zwar hinreichend, aber nicht notwendig (vgl. nochmals mit dem Fall $n = 1$). Ist $H_f(x)$ semidefinit, so kann keine allgemeine Aussage getroffen werden: z. B. im \mathbb{R}^2 haben die folgenden Funktionen $\mathbb{R}^2 \rightarrow \mathbb{R}$

$$f_1(x_1, x_2) := x_1^2 + x_2^4, \quad f_2(x_1, x_2) := x_1^2, \quad f_3(x_1, x_2) := x_1^2 + x_2^3$$

jeweils eine kritische Stelle in $(0, 0)$ (mit Funktionswert $f(0, 0) = 0$); die Hesse-Matrix

$H_{f_k}(0, 0) = \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix}$ ist in allen Fällen positiv semidefinit (für $k = 1, 2, 3$), aber

- f_1 hat ein striktes Minimum in $(0, 0)$ $[f(x_1, x_2) > 0 \text{ für } (x_1, x_2) \neq (0, 0)]$
- f_2 hat ein lokales Minimum in $(0, 0)$, das aber nicht strikt ist $[f_2(0, x_2) = 0 \forall x_2]$
- f_3 hat kein Extremum in $(0, 0)$ $[t \mapsto f_3(0, t) = t^3 \text{ nimmt in jeder Umgebung von } (0, 0) \text{ sowohl positive als auch negative Werte an}]$.

Beweis des Theorems: Wir setzen $A := H_f(x)$. Nach (19.5) gilt für $\xi \in \mathbb{R}^n$ mit $\|\xi\| < r$, wenn r klein genug ist ($\Leftrightarrow x + \xi$ nahe x):

$$f(x + \xi) = f(x) + \frac{1}{2} \langle A\xi \mid \xi \rangle + h(\xi),$$

wobei $h(\xi) = o(\|\xi\|^2)$, d.h.

$$(*) \quad \forall \varepsilon > 0 \exists \delta > 0 : \|\xi\| < \delta \implies |h(\xi)| \leq \varepsilon \|\xi\|^2.$$

ad 1.) A ist positiv definit, daher ist $g(\xi) := \langle A\xi \mid \xi \rangle > 0 \quad \forall \xi \neq 0$;

die Abbildung $\xi \mapsto g(\xi)$ ist stetig $S^{n-1} \rightarrow \mathbb{R}$ und S^{n-1} kompakt, somit gilt

$$\alpha := \min\{g(\xi) : \xi \in S^{n-1}\} > 0.$$

Sei $\xi \neq 0$, dann ist

$$\langle A\xi \mid \xi \rangle = \left\langle \frac{A\xi}{\|\xi\|} \mid \frac{\xi}{\|\xi\|} \right\rangle \|\xi\|^2 \geq \alpha \|\xi\|^2$$

und für $\xi = 0$ gilt die resultierende Ungleichung ebenso.

Wähle $\varepsilon := \frac{\alpha}{4}$ in $(*)$ und $\delta > 0$ passend, dann gilt

$$f(x + \xi) \geq f(x) + \frac{\alpha}{2} \|\xi\|^2 - \frac{\alpha}{4} \|\xi\|^2 = f(x) + \frac{\alpha}{4} \|\xi\|^2 > f(x)$$

für $\|\xi\| < \delta$, $\xi \neq 0$; daher hat f ein striktes lokales Minimum in x .

ad 2.) Betrachte $-f$ und wende 1.) an.

ad 3.) A ist indefinit; wähle $\xi_0, \eta_0 \in \mathbb{R}^n \setminus \{0\}$ mit

$$\gamma_1 := \langle A\xi_0 \mid \xi_0 \rangle > 0 \quad \text{und} \quad \gamma_2 := \langle A\eta_0 \mid \eta_0 \rangle < 0.$$

Sei $r > 0$ und so, dass $\forall t \in [-r, r]: t \cdot \xi_0, t \cdot \eta_0 \in B_r(0)$; es ist

$$\begin{aligned} f(x + t\xi_0) &= f(x) + \frac{t^2}{2}\gamma_1 + h(t\xi_0) \\ f(x + t\eta_0) &= f(x) + \frac{t^2}{2}\gamma_2 + h(t\eta_0). \end{aligned}$$

Aus der Bedingung (*) erhalten wir:

$$\exists \delta \in]0, r[: |t| < \delta \Rightarrow |h(t\xi_0)| \leq \gamma_1 t^2/4 \quad \text{und} \quad |h(t\eta_0)| \leq |\gamma_2| t^2/4;$$

somit gilt für $0 < |t| < \delta$:

$$\begin{aligned} f(x + t\xi_0) &\geq f(x) + \frac{t^2\gamma_1}{2} - \frac{t^2\gamma_1}{4} = f(x) + \frac{t^2\gamma_1}{4} > f(x) \\ &> f(x) + \frac{t^2\gamma_2}{4} = f(x) + \frac{t^2\gamma_2}{2} + \frac{t^2|\gamma_2|}{4} \geq f(x + t\eta_0), \end{aligned}$$

d. h. für alle Umgebungen V von x gilt: $\exists \xi, \eta \in V$ mit $f(\xi) > f(x) > f(\eta)$
(z.B. mit t klein genug: $\xi = x + t\xi_0, \eta = x + t\eta_0$).

Also kann in x kein lokales Extremum vorliegen. □

19.12. Beispiel

Wir untersuchen die Funktion $f: \mathbb{R}^2 \rightarrow \mathbb{R}, f(x, y) := (x^2 - y^2) e^{-x^2 - y^2}$ auf lokale Extrema:

$$\text{grad } f(x, y) = e^{-x^2 - y^2} \cdot \begin{pmatrix} 2x(1 - x^2 + y^2) \\ -2y(1 + x^2 - y^2) \end{pmatrix} =: e^{-x^2 - y^2} \begin{pmatrix} a_1 \\ a_2 \end{pmatrix},$$

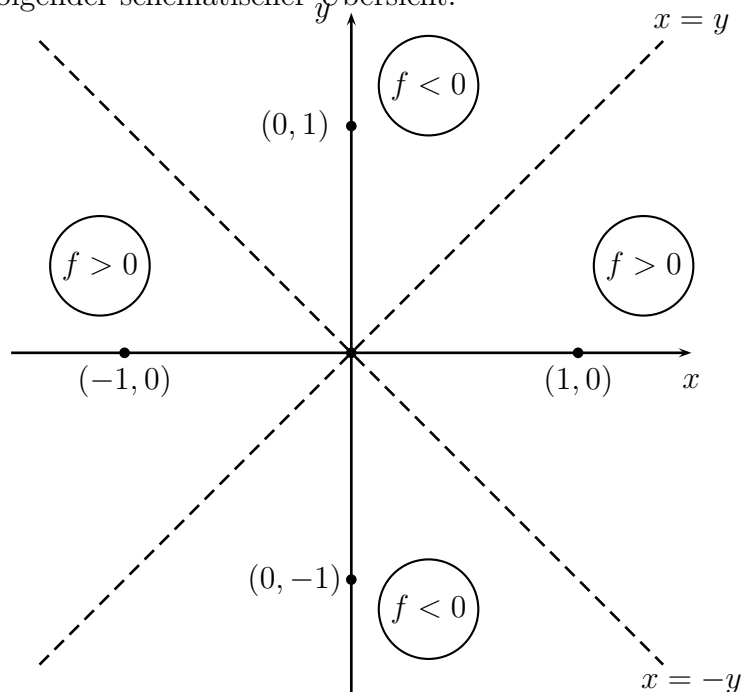
$$H_f(x, y) = e^{-x^2 - y^2} \cdot \begin{pmatrix} 2(1 - x^2 + y^2) - 4x^2 - 2xa_1 & -4xy - 2xa_2 \\ -4xy - 2xa_2 & -2(1 + x^2 - y^2) + 4y^2 - 2ya_2 \end{pmatrix}$$

Um die kritischen Stellen zu finden, lösen wir das Gleichungssystem

$$\begin{aligned} 2x(1 - x^2 + y^2) &= 0 \\ -2y(1 + x^2 - y^2) &= 0. \end{aligned}$$

- $x = 0$: $-2y(1 - y^2) = 0 \Rightarrow y = 0 \vee y = -1 \vee y = 1$
- $x \neq 0$: $1 = x^2 - y^2 \wedge -2y(1 + x^2 - y^2) = 0 \Rightarrow 1 = x^2 - y^2 \wedge -2y(1 + 1) = 0$
 $\Rightarrow y = 0, x = -1 \vee x = 1$

Also haben wir die kritischen Stellen $K_0 = (0, 0)$, $K_1 = (1, 0)$, $K_2 = (0, 1)$, $K_3 = (-1, 0)$ und $K_4 = (0, -1)$. Außerdem lesen wir aus der Definition von f direkt ab, welches Vorzeichen f in verschiedenen Bereichen des \mathbb{R}^2 hat: es ist $f > 0 \Leftrightarrow x^2 > y^2$. Somit kommen wir zunächst zu folgender schematischer Übersicht:



Weiters genügt es wegen der Symmetrien $f(-x, y) = f(x, y) = f(x, -y) = f(-x, -y)$, die Punkte K_0 , K_1 und K_2 zu untersuchen.

$$H_f(0, 0) = \begin{pmatrix} 2 & 0 \\ 0 & -2 \end{pmatrix} \text{ ist indefinit} \Rightarrow \text{Sattelpunkt in } K_0$$

$$H_f(1, 0) = \frac{1}{e} \begin{pmatrix} -4 & 0 \\ 0 & -4 \end{pmatrix} \text{ ist negativ definit} \Rightarrow \text{striktes lokales Maximum in } K_1$$

$$H_f(0, 1) = \frac{1}{e} \begin{pmatrix} 4 & 0 \\ 0 & 4 \end{pmatrix} \text{ positiv definit} \Rightarrow \text{striktes lokales Minimum in } K_2$$

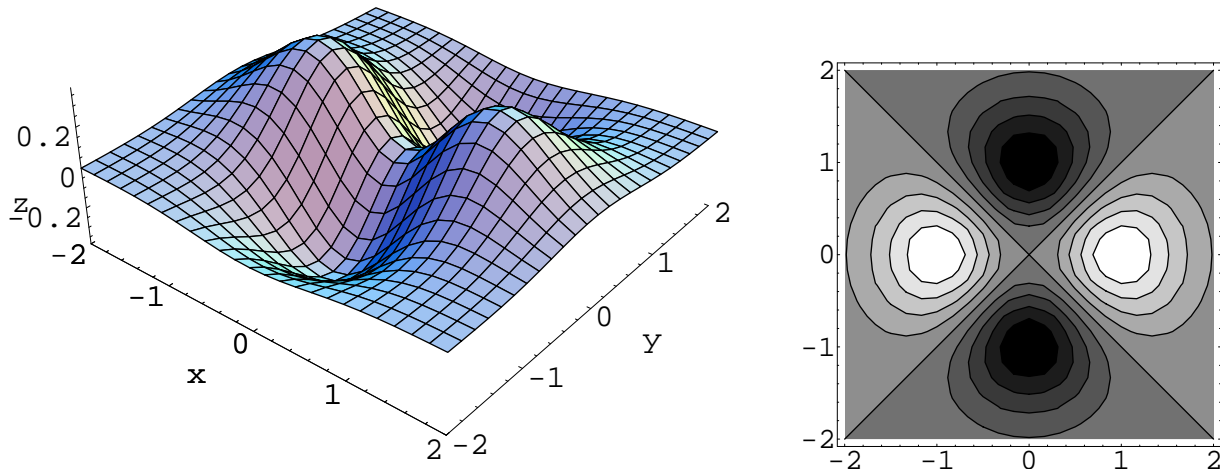
Die entsprechenden Funktionswerte sind $f(0, 0) = 0$, $f(1, 0) = 1/e$ und $f(0, 1) = -1/e$.

Wegen $\lim_{\|z\| \rightarrow \infty} f(z) = 0$ folgt somit, dass

in $(1, 0)$ ein globales Maximum vorliegt, weil $\frac{1}{e} > 0$;

in $(0, 1)$ ein globales Minimum vorliegt, weil $-\frac{1}{e} < 0$.

In den folgenden Darstellungen des Graphen und einiger Niveaulinien der Funktion f mit Hilfe von MATHEMATICA werden die Ergebnisse auch recht gut sichtbar: im rechten Bild bedeuten dunklere Zonen niedrigere Funktionswerte; beachte die Nullniveaulinien, die in diesem Fall genau die Diagonalen $x = y$ und $x = -y$ sind.



Der Funktionsgraph wurde mit der Eingabe

```
Plot3D[(x^2 - y^2) Exp[-x^2 - y^2], {x, -2, 2}, {y, -2, 2},
  AxesLabel -> {x, y, z}, Boxed -> False,
  ViewPoint -> {1.803, -2.246, 1.775}]
```

erzeugt, das Bild der Höhenlinien mit

```
ContourPlot[(x^2 - y^2) Exp[-x^2 - y^2], {x, -2, 2}, {y, -2, 2},
  Contours -> 9, ImageSize -> 180]
```

19.13. Extrema unter Nebenbedingungen

Definition: Es sei $f: \mathbb{R}^n \underset{\text{offen}}{\supseteq} U \rightarrow \mathbb{R}$ stetig differenzierbar und $M \subseteq U$ sowie $a \in M$.

Wir sagen, $f|_M$ besitze ein *lokales Maximum* (bzw. *Minimum*) in a , wenn gilt:

$\exists V \underset{\text{offen}}{\subseteq} U$ mit $a \in V$: $f(x) \leq f(a)$ (bzw. $f(x) \geq f(a)$) $\forall x \in M \cap V$.

Theorem: Es sei $f: \mathbb{R}^n \underset{\text{offen}}{\supseteq} U \rightarrow \mathbb{R}$ stetig differenzierbar. Weiters seien $r \leq n$ und $g_1, \dots, g_r: U \rightarrow \mathbb{R}$ stetig differenzierbare Funktionen mit der Eigenschaft, dass der Rang der Jacobi-Matrix $D(g_1, \dots, g_r)(x)$ gleich r (also maximal) ist für alle $x \in M$, wobei

$$M := \{x \in U: \underbrace{g_1(x) = \dots = g_r(x) = 0}_{r \text{ Nebenbedingungen}}\}.$$

Falls $f|_M$ in $a \in M$ ein lokales Maximum oder Minimum besitzt, dann existieren r reelle Zahlen $\lambda_1, \dots, \lambda_r \in \mathbb{R}$, die so genannten *Lagrange-Multiplikatoren*, sodass die folgende Gleichung gilt

$$(19.6) \quad \text{grad } f(a) = \sum_{j=1}^r \lambda_j \cdot \text{grad } g_j(a).$$

Anstatt eines *Beweises* machen wir hier nur für den Fall einer Nebenbedingungsgleichung von der Form $g(x, y) := y - h(x) = 0$ im \mathbb{R}^2 eine Plausibilitätsüberlegung und geben später in 23.9 einen rein geometrischen Beweis für den allgemeinen Fall. Eine Version des Beweises, die ohne den Begriff der Untermannigfaltigkeit auskommt und direkt auf den Satz über implizite Funktionen (vgl. bei uns 20.2) zugreift findet sich z.B. in [Heu04, Abschnitt 174].

Plausibilitätsüberlegung: Seien also $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ sowie $h: \mathbb{R} \rightarrow \mathbb{R}$ stetig differenzierbar und $g(x, y) := y - h(x)$. Die Lösungsmenge der Nebenbedingung

$$M := \{(x, y) \in \mathbb{R}^2 : g(x, y) = 0\} = \{(x, y) \in \mathbb{R}^2 : y = h(x)\} = \{(x, h(x)) \in \mathbb{R}^2 : x \in \mathbb{R}\}$$

ist dann einfach der Graph von h im \mathbb{R}^2 . Laut Annahme hat f ein lokales Extremum im Punkt $a \in M$. Sei $x_0 \in \mathbb{R}$ mit $(x_0, h(x_0)) = a$. Dann muss die differenzierbare Funktion $x \mapsto f(x, h(x))$ ein lokales Extremum in x_0 besitzen, also verschwindende Ableitung in diesem Punkt haben. Somit folgt nach der Kettenregel

$$0 = \frac{d}{dx} \left(f(x, h(x)) \right) = D_1 f(a) + D_2 f(a) \cdot h'(x_0).$$

Dies lehrt uns aber, dass der Vektor $\text{grad } f(a) \in \mathbb{R}^2$ normal auf $(1, h'(x_0)) \neq 0$ steht, daher also parallel zur Richtung $(-h'(x_0), 1) = \text{grad } g(a)$ sein muss. Mit anderen Worten: es gibt eine reelle Zahl λ so, dass

$$\text{grad } f(a) = \lambda \text{grad } g(a).$$

Methode: Die größte Bedeutung der Lagrange-Multiplikatoren liegt in ihrer Anwendung beim Aufsuchen von Kandidaten für lokale Extrema einer \mathcal{C}^1 -Funktion f unter den Nebenbedingungen $g_1 = \dots = g_r = 0$. Dazu gehen wir wie folgt vor:

wir bilden eine neue Funktion von $n + r$ Variablen durch

$$F(x, \lambda_1, \dots, \lambda_r) := f(x) - \lambda_1 g_1(x) - \dots - \lambda_r g_r(x)$$

und bestimmen Lösungen $(x, \lambda) = (x_1, \dots, x_n, \lambda_1, \dots, \lambda_r) \in U \times \mathbb{R}^r$ des Gleichungssystems $\text{grad}_{(x, \lambda)} F(x, \lambda) = 0$, d.h. in Zeilen gemäß x - und λ -Ableitungen aufgeteilt

$$(19.7) \quad \left\{ \begin{array}{l} \text{grad } f(x) - \lambda_1 \text{grad } g_1(x) - \dots - \lambda_r \text{grad } g_r(x) = 0 \\ g_1(x) = 0 \\ \vdots \\ g_r(x) = 0 \end{array} \right.$$

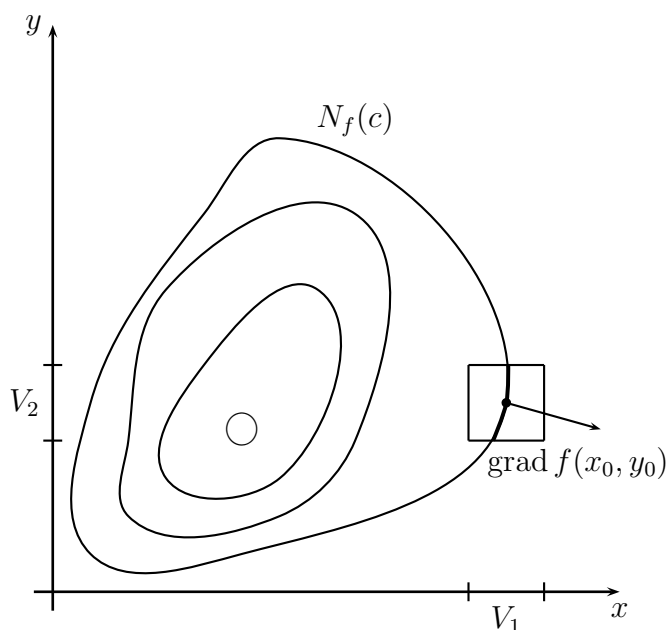
§20. Implizite Funktionen und Umkehrsatz

20.1. Spezialfall: Höhenlinien im \mathbb{R}^2

Es sei $f : \mathbb{R}^2 \underset{\text{offen}}{\supseteq} U \rightarrow \mathbb{R}$ stetig differenzierbar. Für $c \in \mathbb{R}$ ist die entsprechende Niveaumenge im Definitionsbereich gegeben durch

$$N_f(c) = \{(x, y) \in U : f(x, y) = c\},$$

wird also durch „eine Gleichung in zwei Variablen“ beschrieben. Naiv gesehen würden wir erwarten, dass sich aus dieser Gleichung „eine Variable durch die andere ausdrücken lässt“, womit eine explizite Beschreibung einer Höhenlinie erreicht wäre (nämlich beider Koordinaten entlang der Niveaumenge durch einen Parameter).



Um die Fragestellung zu präzisieren, gehen wir von einem fixen Punkt $(x_0, y_0) \in N_f(c)$ aus.

? Gilt nahe (x_0, y_0) :

$$f(x, y) = c \quad [\text{d.h. } (x, y) \in N_f(c)]$$

\Updownarrow

$$x = g(y) \text{ oder } y = h(x).$$

Genauer: es sollen also Umgebungen V_1 von x_0 und V_2 von y_0 existieren, sodass obige Äquivalenz in $V_1 \times V_2$ gilt, wobei $g : V_2 \rightarrow V_1$ bzw. $h : V_1 \rightarrow V_2$ stetig differenzierbar sind.

Um ausgeartete Fälle zu vermeiden nehmen wir zusätzlich an, dass

$$\boxed{\text{grad } f(x_0, y_0) \neq 0}$$

gilt, der Punkt (x_0, y_0) also nicht kritisch ist. Zum Beispiel gehören zu strikten lokalen Extrema stets einpunktige Niveaumengen und in einem Sattelpunkt schneiden einander zwei Niveaulinien (das kann z.B. aus Teil 3.) des Beweises von Theorem 19.11 geschlossen werden.)

Notwendige Bedingungen: Angenommen es gibt eine differenzierbare Funktion $g : V_2 \rightarrow V_1$, sodass $(x_0, y_0) \in V_1 \times V_2 \underset{\text{offen}}{\subseteq} U$ mit $g(y_0) = x_0$ und

$$f(g(y), y) = c \quad \forall y \in V_2.$$

Dann folgt durch Differentiation (nach y) mit der Kettenregel

$$D_1 f(g(y), y) \cdot g'(y) + D_2 f(g(y), y) = 0.$$

$D_1 f(x_0, y_0) = 0$ erzwingt $D_2 f(x_0, y_0) = 0$ und daher auch $\text{grad } f(x_0, y_0) = 0$ — ein Widerspruch ζ

Somit muss gelten:

$$(20.1) \quad D_1 f(x_0, y_0) \neq 0 \quad \text{und} \quad g'(y_0) = -\frac{D_2 f(x_0, y_0)}{D_1 f(x_0, y_0)}.$$

Ähnlich zeigt man: Falls $h : V_1 \rightarrow V_2$ differenzierbar mit $h(x_0) = y_0$, dann folgt aus $f(x, h(x)) = c \forall x \in V_1$, dass

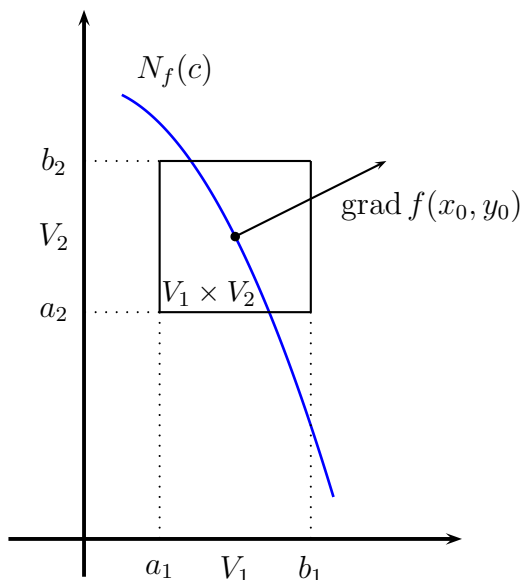
$$(20.2) \quad D_2 f(x_0, y_0) \neq 0 \quad \text{und} \quad h'(x_0) = -\frac{D_1 f(x_0, y_0)}{D_2 f(x_0, y_0)}.$$

Existenz von g oder h : Es sei $\gamma := D_1 f(x_0, y_0) > 0$.

Die Funktion $x \mapsto f(x, y_0)$ ist dann nahe x_0 streng monoton wachsend und wegen der Stetigkeit von $D_1 f$ erhalten wir:

$\exists V_j =]a_j, b_j[$ ($j = 1, 2$) mit $a_1 < x_0 < b_1$ und $a_2 < y_0 < b_2$, sodass gilt

- 1.) $\forall (x, y) \in \bar{V}_1 \times \bar{V}_2: D_1 f(x, y) \geq \frac{\gamma}{2} > 0$
- 2.) $\forall y \in \bar{V}_2 = [a_2, b_2]: f(a_1, y) < c = f(x_0, y_0) < f(b_1, y)$.



⊙ **Konstruktion von $g : V_2 \rightarrow V_1$:**

sei $y \in V_2$: $t \mapsto f(t, y)$ ist streng monoton wachsend $[a_1, b_1] \rightarrow \mathbb{R}$ und es gilt $f(a_1, y) < c < f(b_1, y)$;

daher $\exists! x \in]a_1, b_1[: f(x, y) = c$.

Wir setzen $g(y) := x$.

Die Funktion g ist auf V_2 eindeutig bestimmt und es gilt $f(g(y), y) = c$ nach Konstruktion.

Noch zu zeigen: g ist stetig differenzierbar.

☉ **Stetigkeit von g** : Sei $y \in V_2$ und (y_n) eine Folge in V_2 mit $y_n \rightarrow y$ ($n \rightarrow \infty$).

Setze $x := g(y)$, $x_n := g(y_n)$; wir müssen zeigen, dass $x_n \rightarrow x$ gilt.

Es ist $\underbrace{f(x_n, y_n)}_c - \underbrace{f(x, y_n)}_{\xrightarrow{c}} = \underbrace{D_1 f(\xi_n, y_n)}_{\geq \frac{c}{2} > 0} \cdot (x_n - x)$ für ein ξ_n zwischen x und x_n ,

daher $0 \leq \frac{c}{2} \cdot |x_n - x| \leq |D_1 f(\xi_n, y_n)| |x_n - x| = |c - f(x, y_n)| \rightarrow 0$ ($n \rightarrow \infty$)

und somit $x_n \rightarrow x$ ($n \rightarrow \infty$), d. h. $g(y_n) \rightarrow g(y)$.

☉ **Stetige Differenzierbarkeit von g** : Sei $y \in V_2$ und $0 \neq h$ so klein, dass $y + h \in V_2$ gilt.

- $g(y + h) = g(y) \Rightarrow \frac{g(y+h) - g(y)}{h} = 0$
- $g(y + h) \neq g(y) \Rightarrow f(g(y + h), y) \neq f(g(y), y)$ (wegen strenger Monotonie) und

$$\frac{g(y + h) - g(y)}{h} = \frac{g(y + h) - g(y)}{f(g(y + h), y) - \underbrace{f(g(y), y)}_c} \cdot \frac{\overbrace{f(g(y + h), y) - f(g(y + h), y + h)}^c}{h}$$

[für $h \rightarrow 0$,
 $g(y + h) \rightarrow g(y)$]

$$\downarrow \qquad \qquad \qquad \downarrow$$

$$\frac{1}{D_1 f(g(y), y)} \cdot (-D_2 f(g(y), y))$$

- ist $h_n \neq 0$, $h_n \rightarrow 0$ und $g(y + h_n) = g(y) \forall n$, dann haben wir einfach

$$D_2 f(g(y), y) = \lim_{n \rightarrow \infty} \frac{f(g(y), y + h_n) - f(g(y), y)}{h_n}$$

$$= \lim_{n \rightarrow \infty} \frac{\overbrace{f(g(y + h_n), y + h_n)}^c - \overbrace{f(g(y), y)}^c}{h_n} = \lim_{n \rightarrow \infty} \frac{c - c}{h_n} = 0$$

Insgesamt erhalten wir also:

$$\exists g'(y) = \lim_{h \rightarrow 0} \frac{g(y + h) - g(y)}{h} = -\frac{D_2 f(g(y), y)}{D_1 f(g(y), y)}$$

Diese Formel zeigt auch die Stetigkeit der Ableitung, also ist g stetig differenzierbar.

Analog argumentiert man für die Fälle $D_1 f(x_0, y_0) < 0$ bzw. $D_2 f(x_0, y_0) > 0$ und $D_2 f(x_0, y_0) < 0$, wobei wir bei letzteren zwei Fällen wir lokal $y = h(x)$ erhalten.

Zusammenfassend gilt also folgende

Proposition: Es sei $f : \mathbb{R}^2 \underset{\text{offen}}{\supseteq} U \rightarrow \mathbb{R}$ stetig differenzierbar, $c \in \mathbb{R}$ und $(x_0, y_0) \in N_f(c) = f^{-1}(\{c\})$ mit $\text{grad } f(x_0, y_0) \neq \begin{pmatrix} 0 \\ 0 \end{pmatrix}$.

Falls $\begin{Bmatrix} D_1 f(x_0, y_0) \\ D_2 f(x_0, y_0) \end{Bmatrix} \neq 0$ ist, dann gibt es $V_1, V_2 \subseteq \mathbb{R}$ offen mit $(x_0, y_0) \in V_1 \times V_2 \subseteq U$ und $\begin{Bmatrix} g : V_2 \rightarrow V_1 \\ h : V_1 \rightarrow V_2 \end{Bmatrix}$ stetig differenzierbar und eindeutig mit der Eigenschaft:

$$\forall (x, y) \in V_1 \times V_2 : \quad (x, y) \in N_f(c) \iff \begin{cases} x = g(y) \\ y = h(x) \end{cases}$$

Weiters gilt dann $\begin{cases} g'(y) = -\frac{D_2 f(g(y), y)}{D_1 f(g(y), y)} \\ h'(x) = -\frac{D_1 f(x, h(x))}{D_2 f(x, h(x))} \end{cases}$.

Bemerkung: Bei der Anwendung dieser Proposition wird oft die Schreibweisen $y = y(x)$ statt $y = h(x)$ verwendet und mit so genanntem „impliziten Differenzieren“ gearbeitet:

$$c = f(x, y(x)) \Rightarrow \underbrace{f_x}_{1} \cdot \frac{dx}{dx} + \underbrace{f_y}_{y'} \cdot \frac{dy}{dx} = 0 \Rightarrow y' = -\frac{f_x}{f_y}.$$

Beispiel: $f :]0, \infty[^2 \rightarrow \mathbb{R}$, $f(x, y) := x + y + \log(xy)$, $c = 2$ und $(x_0, y_0) = (1, 1)$;

es ist $f(1, 1) = 1 + 1 + \log(1) = 2$, also $(1, 1)$ Element von $N_f(2)$;

wegen $D_2 f(x, y) = 0 + 1 + \frac{1}{xy} \cdot x = 1 + \frac{1}{y}$, somit $D_2 f(1, 1) = 2 \neq 0$, ist die Gleichung

$$x + y + \log(xy) = 2$$

nahe $x = 1$ nach y auflösbar durch eine stetig differenzierbare Funktion $x \mapsto y(x)$ mit

$$y'(x) = -\frac{D_1 f(x, y)}{D_2 f(x, y)} = -\frac{1 + \frac{1}{x}}{1 + \frac{1}{y}}, \text{ also } y'(1) = -1.$$

Durch weiteres implizites Differenzieren erhalten wir auch

$$y''(x) = -\frac{-\frac{1}{x^2}(1 + \frac{1}{y}) - (1 + \frac{1}{x}) \left(-\frac{y'}{y^2}\right)}{\left(1 + \frac{1}{y}\right)^2}, \text{ also } y''(1) = -\frac{-1 \cdot (2) - 2 \cdot (1)}{2^2} = 1.$$

Z.B. gewinnen wir daraus die Taylorentwicklung für y um $x_0 = 1$:

$$y(x) = y(1) + y'(1) \cdot (x - 1) + \frac{y''(1)}{2} \cdot (x - 1)^2 + O(|x - 1|^3) = 1 - (x - 1) + \frac{(x - 1)^2}{2} + O(|x - 1|^3)$$

20.2. Satz über implizite Funktionen

Es seien $U_1 \subseteq \mathbb{R}^k$, $U_2 \subseteq \mathbb{R}^m$ offen und $F: U_1 \times U_2 \rightarrow \mathbb{R}^m$, eine stetig differenzierbare Abbildung, $F(x, y) = (F_1(x, y), \dots, F_m(x, y))$.

Die Bedingung

$$(*) \quad F(x, y) = 0$$

entspricht einem System von m Gleichungen für x_1, \dots, x_k und y_1, \dots, y_m .

? Können aus den Gleichungen (*) die m Variablen y_1, \dots, y_m durch die k Variablen x_1, \dots, x_k „ausgedrückt“ werden?

Legt (*) durch seine Lösungsmenge $y \in \mathbb{R}^m$ als Funktion von $x \in \mathbb{R}^k$ fest?

Nehmen wir einmal probeweise an, es wäre so; d.h. es sei $y = g(x)$ und g differenzierbar. Für $1 \leq j \leq m$ erhalten wir dann durch Einsetzen in (*) die Gleichung

$$F_j(x, g(x)) = 0$$

und daraus durch partielle Differentiation nach x_l ($1 \leq l \leq k$) gemäß der Kettenregel

$$0 = D_{x_l}(F_j(x, g(x))) = D_{x_l}F_j(x, g(x)) + \sum_{p=1}^m D_{y_p}F_j(x, g(x)) \cdot D_{x_l}g_p(x).$$

Das resultierende Gleichungssystem können wir als linearisierte Form von (*) auffassen. Dazu führen wir noch eine kompaktere Notation für entsprechende Teilmatrizen der Jacobi-Matrix von F ein:

$$(20.3) \quad \begin{cases} \frac{\partial F}{\partial x} = \frac{\partial(F_1, \dots, F_m)}{\partial(x_1, \dots, x_k)} := (D_{x_j}F_i)_{\substack{1 \leq i \leq m \\ 1 \leq j \leq k}} & [(m \times k)\text{-Teilmatrix}] \\ \frac{\partial F}{\partial y} = \frac{\partial(F_1, \dots, F_m)}{\partial(y_1, \dots, y_m)} := (D_{y_j}F_i)_{1 \leq i, j \leq m} & [(m \times m)\text{-Teilmatrix}]. \end{cases}$$

Damit lautet die gewonnene Linearisierung von (*) dann

$$(20.4) \quad \frac{\partial F}{\partial x}(x, g(x)) + \frac{\partial F}{\partial y}(x, g(x)) \cdot Dg(x) = 0.$$

Nun gilt folgende Äquivalenz:

$$(20.4) \text{ ist auflösbar nach } „dy“ = Dg(x) \iff \frac{\partial F}{\partial y}(x, g(x)) \text{ ist invertierbar.}$$

Dass diese Bedingung sogar hinreichend für lokale Auflösbarkeit des ursprünglichen Gleichungssystems (*) nach y ist, beweisen wir im folgenden

Theorem: Es sei $(a, b) \in U_1 \times U_2 \subseteq \mathbb{R}^k \times \mathbb{R}^m$ mit $F(a, b) = 0$. Weiters gelte, dass

$$(20.5) \quad \frac{\partial F}{\partial y}(a, b) \quad \text{invertierbar ist.}$$

Dann gibt es offene Umgebungen V_1 von a mit $V_1 \subseteq U_1$ und V_2 von b mit $V_2 \subseteq U_2$ und eine eindeutige stetig differenzierbare Abbildung $g: V_1 \rightarrow V_2$ mit $g(a) = b$ und der Eigenschaft:

$$\forall (x, y) \in V_1 \times V_2: \quad F(x, y) = 0 \quad \iff \quad y = g(x).$$

Es gilt dann auch

$$(20.6) \quad Dg(x) = - \left(\frac{\partial F}{\partial y}(x, g(x)) \right)^{-1} \cdot \frac{\partial F}{\partial x}(x, g(x)).$$

Bemerkung: Proposition 20.1 ist somit ein Spezialfall dieses Theorems für $k = m = 1$, $F(x, y) = f(x, y) - c$ und $\frac{\partial F}{\partial y} = D_2 f$ (bzw. falls $D_2 f(x_0, y_0) = 0$ ist, vertausche die Rollen von x und y).

Beispiel: $F: \mathbb{R}^1 \times \mathbb{R}^2 \rightarrow \mathbb{R}^2$, $F(x, y, z) = \begin{pmatrix} x^2 + y^2 + z^2 - 6 \\ x^3 + y^3 + z^3 - 10 \end{pmatrix} = \begin{pmatrix} F_1(x, y, z) \\ F_2(x, y, z) \end{pmatrix}$

Für die Ausgangssituation im Theorem haben wir $k = 1$ und $m = 2$.

? Ist innerhalb der Lösungsmenge des Gleichungssystems $F(x, y, z) = 0$ nahe des Lösungspunktes $(1, 2, 1)$ zumindest lokal (y, z) als Funktion von x darstellbar?

Es ist $\frac{\partial F}{\partial (y,z)} = \begin{pmatrix} 2y & 2z \\ 3y^2 & 3z^2 \end{pmatrix}$ und $\frac{\partial F}{\partial x} = \begin{pmatrix} 2x \\ 3x^2 \end{pmatrix}$.

Wir setzen $B := \frac{\partial F}{\partial (y,z)}(1, 2, 1) = \begin{pmatrix} 4 & 2 \\ 12 & 3 \end{pmatrix}$, dann ist $\det B = -12 \neq 0$, also B invertierbar

und $B^{-1} = \frac{-1}{12} \begin{pmatrix} 3 & -2 \\ -12 & 4 \end{pmatrix}$. Weiters setzen wir $A := \frac{\partial F}{\partial x}(1, 2, 1) = \begin{pmatrix} 2 \\ 3 \end{pmatrix}$.

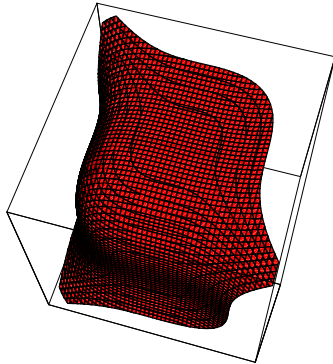
Aus dem Theorem folgt: $\exists V_1 \subseteq \mathbb{R}$ offen, $1 \in V_1$, und $\exists V_2 \subseteq \mathbb{R}^2$ offen, $(2, 1) \in V_2$, sowie $g: V_1 \rightarrow V_2$ stetig differenzierbar mit $g(1) = (2, 1)$ so, dass

$$\forall (x, y, z) \in V_1 \times V_2: \quad F(x, y, z) = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad \iff \quad \begin{pmatrix} y \\ z \end{pmatrix} = g(x)$$

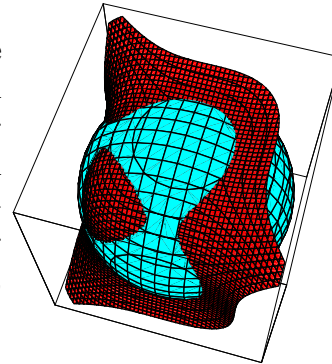
und weiters

$$Dg(1) = -B^{-1} \cdot A = \frac{1}{12} \begin{pmatrix} 3 & -2 \\ -12 & 4 \end{pmatrix} \cdot \begin{pmatrix} 2 \\ 3 \end{pmatrix} = \frac{1}{12} \begin{pmatrix} 0 \\ -12 \end{pmatrix} = \begin{pmatrix} 0 \\ -1 \end{pmatrix}.$$

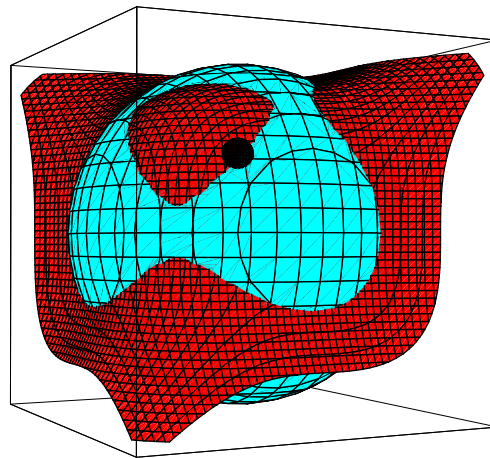
In der geometrischen Interpretation stellt $g: \mathbb{R} \supseteq V_1 \rightarrow \mathbb{R}^2$ lokal die Schnittkurve der beiden Flächen $F_1(x, y, z) = 0$ und $F_2(x, y, z) = 0$ dar, nämlich mittels Parametrisierung in der Form $x \mapsto (x, g(x))$:



$F_1(x, y, z) = 0$ beschreibt eine Sphäre mit Radius $\sqrt{6}$ um den Ursprung. Ein Ausschnitt der durch $F_2(x, y, z) = 0$ gegebenen Teilmenge des \mathbb{R}^3 ist in der Graphik links dargestellt und der Schnitt dieser mit der Sphäre ist im rechten Bild illustriert.



Hier stellen wir die Situation noch einmal in einer anderen Ansicht dar, in der zur Verdeutlichung eine kleine Kugelumgebung um den Punkt $(1, 2, 1)$ eingeschwärzt wurde:



Zur Information: Die obigen Illustrationen wurden in MATHEMATICA mit den folgenden Eingaben erzeugt. Zunächst werden die speziellen Graphik-Pakete eingelesen (z.B. für Niveauflächen von Funktionen dreier Variablen)

```
<< Graphics`ContourPlot3D`
```

Zur Vorbereitung (keine Abbildung oben) erzeugen wir die Sphäre mittels

```
cp1 = ContourPlot3D[x^2 + y^2 + z^2 - 6, {x, -2.5, 2.5}, {y, -2.5, 2.5},
  {z, -2.5, 2.5}, Lighting -> False, ContourStyle -> {RGBColor[0, 1, 1]},
  PlotPoints -> {5, 5}]
```

und dann die durch F_2 gegebene Fläche (Abbildung oben links)

```
cp2 = ContourPlot3D[x^3 + y^3 + z^3 - 10, {x, -2.5, 2.5}, {y, -2.5, 2.5},
  {z, -2.5, 2.5}, Lighting -> False, ContourStyle -> {RGBColor[1, 0, 0]},
  PlotPoints -> {8, 8}, ViewPoint -> {-0.428, 1.498, 3.004}]
```

nun werden die erzeugten Graphiken zusammen dargestellt (Abbildung oben rechts):

```
Show[cp1, cp2, ViewPoint -> {-0.428, 1.498, 3.004}]
```

Die Erzeugung der kleinen schwarzen Kugelumgebung kann z.B. so vorgenommen werden (wir brauchen viele PlotPoints, um bei dem kleinen Radius noch Punkte innerhalb der Sphäre zu erfassen; ohne Abbildung)

```
cp3 = ContourPlot3D[(x - 1)^2 + (y - 2)^2 + (z - 1)^2 - 0.04, {x, -2.5, 2.5},
  {y, -2.5, 2.5}, {z, -2.5, 2.5}, Lighting -> False,
  ContourStyle -> {RGBColor[0, 0, 0]}, PlotPoints -> {12,12},
  ViewPoint -> {-0.428, 1.498, 3.004}]
```

schließlich wurde diese in die kombinierte Ansicht mit angepasster Viewpoint-Option eingefügt

```
Show[cp1, cp2, cp3, ViewPoint -> {1.508, 3.029, -0.026}]
```

Beweis des Theorems:

OBdA ist $(a, b) = (0, 0) \in \mathbb{R}^k \times \mathbb{R}^m$ (andernfalls ist dies stets durch Translation erreichbar, wobei die Ableitung unverändert bleibt).

Wir setzen $A := \frac{\partial F}{\partial x}(0, 0)$ [eine $(m \times k)$ -Matrix] und $B := \frac{\partial F}{\partial y}(0, 0)$ [eine invertierbare $(m \times m)$ -Matrix].

1) Umformulierung in ein Fixpunktproblem mit Parametern:

Wir betrachten die Abbildung $G : U_1 \times U_2 \rightarrow \mathbb{R}^m$, $G(x, y) := y - B^{-1} \cdot F(x, y)$;

dann gilt: $F(x, y) = 0 \iff G(x, y) = y$
(d. h. y ist Fixpunkt von $y \mapsto G(x, y)$ bei gegebenem x)

Es ist

$$\frac{\partial G}{\partial y}(x, y) = \underbrace{I_m - B^{-1} \frac{\partial F}{\partial y}(x, y)}_{\text{Matrix mit stetigen Funktionen als Komponenten}} \quad \left[\text{wobei } I_m = \begin{pmatrix} 1 & & O \\ & \ddots & \\ O & & 1 \end{pmatrix} \right],$$

also ist $(x, y) \mapsto \frac{\partial G}{\partial y}$ stetig und $\frac{\partial G}{\partial y}(0, 0) = 0$. Daher gibt es $W_1 \times W_2 \underset{\text{offen}}{\subseteq} U_1 \times U_2$ mit $(0, 0) \in W_1 \times W_2$ und

$$(*) \quad \left\| \frac{\partial G}{\partial y}(x, y) \right\|_{\text{op}} \leq \frac{1}{2} \quad \forall (x, y) \in W_1 \times W_2;$$

Wähle $r > 0$ so, dass $\overline{B_r(0)} \subseteq W_2$ und setze $V_2 := B_r(0)$.

G ist stetig und $G(0, 0) = 0 - B^{-1} \cdot 0 = 0$, daher gibt es $V_1 \subseteq W_1$ offen mit $0 \in V_1$:

$$\sup_{x \in V_1} \|G(x, 0)\| \leq \frac{r}{2} \quad [\text{weil auch } x \mapsto \|G(x, 0)\| \text{ stetig ist}].$$

Seien $x \in V_1$ und $y, \eta \in \bar{V}_2$, dann gilt:

$$(\Delta) \quad \|G(x, y) - G(x, \eta)\| \leq \sup_{\zeta \in \bar{V}_2} \left\| \frac{\partial G}{\partial y}(x, \zeta) \right\|_{\text{op}} \cdot \|y - \eta\| \stackrel{[(*)]}{\leq} \frac{1}{2} \|y - \eta\|.$$

Mit $\eta = 0$ erhalten wir aus (Δ) zusammen mit obigem weiters

$$\frac{1}{2} \|y\| \geq \|G(x, y) - G(x, 0)\| \geq \|G(x, y)\| - \|G(x, 0)\| \geq \|G(x, y)\| - \frac{r}{2},$$

$$\text{d.h.:} \quad \|G(x, y)\| \leq \frac{1}{2} (\|y\| + r) \quad \forall x \in V_1, \forall y \in V_2 = B_r(0).$$

Insgesamt gilt also

$$(**) \quad \forall x \in V_1: \quad y \in \bar{V}_2 \implies \|G(x, y)\| \leq \frac{1}{2}(r + r) = r,$$

somit bildet $y \mapsto G(x, y)$ die abgeschlossene Kugel $\bar{V}_2 = \overline{B_r(0)}$ in sich ab, falls $x \in V_1$.

2) Entlarvung als Fixpunktproblem im Raum stetiger Funktionen:

$$\mathcal{C}_b(V_1, \mathbb{R}^m) = \{f : V_1 \rightarrow \mathbb{R}^m \text{ stetig} : \|f\|_\infty := \sup_{x \in V_1} \|f(x)\| < \infty\}$$

ist ein Banach-Raum [vgl. Korollar 16.18] und

$$X := \{f \in \mathcal{C}_b(V_1, \mathbb{R}^m) : \|f\|_\infty \leq r\}$$

ist eine abgeschlossene Teilmenge (das ist die abgeschlossene r -Kugel bzgl. der Metrik $d(f_1, f_2) = \|f_1 - f_2\|_\infty$; vgl. Beispiel 16.15).

Daher ist (X, d) ein vollständiger metrischer Raum. [Sei (f_n) Cauchy-Folge in $X \Rightarrow (f_n)$ Cauchy-Folge in $\mathcal{C}_b \Rightarrow (f_n)$ konvergiert in $\mathcal{C}_b \Rightarrow$ Limes von (f_n) in X , weil X abgeschlossen.]

Wir definieren zunächst eine Abbildung $\Phi_0 : \mathcal{C}_b(V_1, \mathbb{R}^m) \rightarrow \mathcal{C}(V_1, \mathbb{R}^m)$ durch $\Phi_0(f)(x) := G(x, f(x))$.

Wegen $(**)$ gilt: $\|f\|_\infty \leq r \Rightarrow \|\Phi_0(f)\|_\infty \leq r$, d.h. $\Phi_0(X) \subseteq X$.

Wir setzen nun $\Phi := \Phi_0|_X : X \rightarrow X$.

Es ist für $f_1, f_2 \in X$:

$$\begin{aligned} \|\Phi(f_1) - \Phi(f_2)\|_\infty &= \sup_{x \in V_1} \|G(x, f_1(x)) - G(x, f_2(x))\| \\ &\stackrel{[(\Delta)]}{\leq} \frac{1}{2} \sup_{x \in V_1} \|f_1(x) - f_2(x)\| = \frac{1}{2} \|f_1 - f_2\|_\infty, \end{aligned}$$

d. h. Φ ist eine Kontraktion auf dem vollständigen metrischen Raum X .

Nach dem Fixpunktsatz von Banach gilt also: $\exists! g \in X: \Phi(g)(x) = g(x) \forall x \in V_1$,
d. h. $g: V_1 \rightarrow B_r(0) \subseteq \overline{V}_2 \subseteq \mathbb{R}^m$ ist stetig mit der Eigenschaft, dass

$$g(x) = G(x, g(x)) = g(x) - B^{-1} \cdot F(x, g(x)),$$

was wiederum äquivalent ist zur Gültigkeit von

$$F(x, g(x)) = 0 \quad \forall x \in V_1.$$

3) Differenzierbarkeit von g :

(*) bedeutet für $(x, y) \in W_1 \times W_2$ und $R := B^{-1} \cdot \frac{\partial F}{\partial y}(x, y)$, dass $\|I_m - R\|_{\text{op}} \leq \frac{1}{2}$ gilt.

Daher ist R injektiv [denn $Rz = 0 \Rightarrow \frac{1}{2}\|z\| \geq \|I_m z - Rz\| = \|z\| \Rightarrow z = 0$], also auch bijektiv [als lineare Abbildung von \mathbb{R}^n in sich].

Somit ist $\frac{\partial F}{\partial y}(x, y)$ invertierbar; insbesondere ist $\frac{\partial F}{\partial y}(x, g(x))$ invertierbar $\forall x \in V_1$.

Sei $x_0 \in V_1$, $y_0 := g(x_0)$ und $A_0 := \frac{\partial F}{\partial x}(x_0, y_0)$, $B_0 := \frac{\partial F}{\partial y}(x_0, y_0)$.

F ist differenzierbar in (x_0, y_0) , daher gibt es nahe (x_0, y_0) eine Abbildung φ mit $\varphi(x, y) = o(\|(x - x_0, y - y_0)\|) = o(\|x - x_0\| + \|y - y_0\|)$ und

$$F(x, y) = \underbrace{F(x_0, y_0)}_0 + \underbrace{DF(x_0, y_0)}_{\substack{(A_0, B_0) \\ [(m \times (k+m))\text{-Matrix}]}} \cdot \begin{pmatrix} x - x_0 \\ y - y_0 \end{pmatrix} + \varphi(x, y),$$

d.h. $F(x, y) = A_0 \cdot (x - x_0) + B_0 \cdot (y - y_0) + \varphi(x, y)$.

Setzen wir $y = g(x)$, so folgt

$$0 = F(x, g(x)) = A_0 \cdot (x - x_0) + B_0 \cdot (g(x) - g(x_0)) + \varphi(x, g(x))$$

und daraus wegen der Invertierbarkeit von B_0 auch

$$g(x) - g(x_0) = -B_0^{-1} \cdot A_0 \cdot (x - x_0) - \underbrace{B_0^{-1} \cdot \varphi(x, g(x))}_{=: \psi(x)}.$$

• Es bleibt zu zeigen: $\psi(x) = o(\|x - x_0\|)$ (Definition der Differenzierbarkeit!)

Behauptung: $\exists V'_1 \subseteq V_1$ offen mit $x_0 \in V'_1$ und $K \geq 0$:

$$\|g(x) - g(x_0)\| \leq K \cdot \|x - x_0\| \quad \forall x \in V'_1.$$

— ist diese Behauptung bewiesen, dann sind wir fertig, denn für $x \rightarrow x_0$ folgt daraus

$$\psi(x) = B_0^{-1} \cdot \varphi(x, g(x)) = o(\|x - x_0\| + K \cdot \|x - x_0\|) = o(\|x - x_0\|).$$

Beweis der Behauptung:

$\forall \varepsilon > 0 \exists V' \subseteq V_1 \times V_2$ offen, $(x_0, y_0) \in V'$, $\forall (x, y) \in V'$: $\|\varphi(x, y)\| \leq \varepsilon \cdot (\|x - x_0\| + \|y - y_0\|)$.

Es ist

$$(***) \quad \|g(x) - g(x_0)\| \leq \underbrace{\|B_0^{-1}A_0\|_{\text{op}}}_{=: c_1} \cdot \|x - x_0\| + \underbrace{\|B_0^{-1}\|_{\text{op}}}_{=: c_2} \cdot \|\varphi(x, g(x))\|.$$

Aus der Stetigkeit von g folgt, dass $x \mapsto (x, g(x))$ stetig ist. Daher $\exists V'_1 \subseteq V_1$ offen mit $x_0 \in V'_1$, sodass der Graph von g , d.h. $\{(x, g(x)) : x \in V'_1\}$, ganz in V' enthalten ist; dann gilt $\forall x \in V'_1$:

$$\|\varphi(x, g(x))\| \leq \varepsilon \cdot (\|x - x_0\| + \|g(x) - g(x_0)\|),$$

somit haben wir vermöge (***) insgesamt

$$\|g(x) - g(x_0)\| \leq (c_1 + \varepsilon \cdot c_2)\|x - x_0\| + \varepsilon \cdot c_2 \cdot \|g(x) - g(x_0)\|.$$

Wählen wir $\varepsilon > 0$ so klein, dass $c_2 \cdot \varepsilon < \frac{1}{2}$, dann folgt somit

$$\frac{1}{2}\|g(x) - g(x_0)\| \leq (c_1 + \frac{1}{2}) \cdot \|x - x_0\|,$$

d.h. die Behauptung gilt mit $K = 2c_1 + 1$.

4) Formel (20.6) und stetige Differenzierbarkeit:

g ist differenzierbar, daher gilt Formel (20.4) und daraus folgt mittels Invertierbarkeit von $\frac{\partial F}{\partial y}(x, g(x))$ ($\forall x \in V_1$) direkt die Gleichung (20.6):

$$Dg(x) = - \left(\frac{\partial F}{\partial y}(x, g(x)) \right)^{-1} \cdot \frac{\partial F}{\partial x}(x, g(x)).$$

Gemäß der Voraussetzungen ist die rechte Seite dieser Gleichung stetig abhängig von x , also ist $x \mapsto Dg(x)$ stetig; somit ist g stetig differenzierbar.

□

20.3. Umkehrabbildungen

Es seien $U_1, U_2 \subseteq \mathbb{R}^n$ offen und $f : U_1 \rightarrow U_2$ stetig differenzierbar und bijektiv.

□ Ist f^{-1} stetig differenzierbar?

Notwendige Bedingung: Falls $f^{-1} : U_2 \rightarrow U_1$ differenzierbar ist, folgt aus $f^{-1} \circ f = \text{id}_{U_1}$ nach der Kettenregel

$$D(f^{-1})(f(x)) \cdot Df(x) = I_n$$

und somit, dass sowohl $(Df)(x)$ als auch $D(f^{-1})(f(x))$ invertierbar sind. Weiters folgt die Gleichung

$$D(f^{-1})(f(x)) = ((Df)(x))^{-1}.$$

Wie wir zeigen werden, ist die Bedingung der Invertierbarkeit von $(Df)(x)$ zumindest lokal auch hinreichend.

Theorem: Es sei $U \subseteq \mathbb{R}^n$ offen, $f: U \rightarrow \mathbb{R}^n$ stetig differenzierbar und $a \in U$ sowie $b := f(a) \in \mathbb{R}^n$. Falls

$$Df(a) \text{ invertierbar ist,}$$

dann gilt: $\exists U_0 \subseteq U$ offen, $a \in U_0$ und $\exists V_0 \subseteq \mathbb{R}^n$ offen, $b \in V_0$ mit der Eigenschaft, dass

$$f|_{U_0}: U_0 \rightarrow V_0 \text{ bijektiv und } g := (f|_{U_0})^{-1}: V_0 \rightarrow U_0 \text{ stetig differenzierbar ist}$$

und es gilt die Gleichung

$$Dg(b) = (Df(a))^{-1}.$$

Beweis: Betrachte die Abbildung $F: \mathbb{R}^n \times U \rightarrow \mathbb{R}^n$, $F(x, y) := x - f(y)$.

Es ist $F(b, a) = 0$ und $\frac{\partial F}{\partial y}(b, a) = -Df(a)$ invertierbar. Nach dem Satz über implizite Funktionen gibt es daher offene Umgebungen $V' \ni b$, $U \supseteq U' \ni a$ und $h: V' \rightarrow U'$ stetig differenzierbar mit der Eigenschaft

$$\forall (x, y) \in V' \times U': \underbrace{F(x, y) = 0}_{\text{d.h. } x=f(y)} \iff y = h(x).$$

Sei $U_0 \subseteq U'$ mit $a \in U_0$, sodass $V_0 := f(U_0) \subseteq V'$. Dann ist $V_0 = h^{-1}(U_0)$, also ist V_0 offen [weil h stetig ist], $b \in V_0$, und $f(U_0) = V_0$, d.h. $f|_{U_0}$ ist surjektiv; $\forall x \in V_0$, $\forall y \in U_0$ gilt:

$$y = h(x) \iff x = f(y)$$

wobei $g := h|_{V_0}$ die Menge V_0 in U_0 abbildet.

Also ist $f|_{U_0}$ bijektiv mit der Umkehrfunktion g , die ebenfalls stetig differenzierbar ist. \square

Definition: Es seien $U, V \subseteq \mathbb{R}^n$ offen und $f: U \rightarrow V$ bijektiv und \mathcal{C}^1 (stetig differenzierbar). Falls $f^{-1}: V \rightarrow U$ ebenfalls stetig differenzierbar ist, so heißt f ein \mathcal{C}^1 -*Diffeomorphismus*.

Allgemein: Sind f und f^{-1} beide k -mal stetig differenzierbar, so heißt die Abbildung \mathcal{C}^k -*Diffeomorphismus*; oft sagt man für $k = \infty$ (oder auch schon für $k = 1$) einfach *Diffeomorphismus*.

Bemerkung: Ist f ein \mathcal{C}^1 -Diffeomorphismus und f k -mal stetig differenzierbar, dann ist f sogar ein \mathcal{C}^k -Diffeomorphismus.

[Beweis induktiv: da in $D(f^{-1})(y) = ((Df)(f^{-1}(y)))^{-1}$ alle Komponentenfunktionen stetig differenzierbar sind, ist $D(f^{-1})$ stetig differenzierbar usw.]

Beispiel: Polarkoordinaten im \mathbb{R}^2

Wir betrachten $f : \mathbb{R}^+ \times \mathbb{R} \rightarrow \mathbb{R}^2, (r, \varphi) \mapsto (r \cdot \cos \varphi, r \cdot \sin \varphi)$.

Es ist $Df(r, \varphi) = \begin{pmatrix} \cos \varphi & -r \cdot \sin \varphi \\ \sin \varphi & r \cdot \cos \varphi \end{pmatrix}$, daher $\det(Df(r, \varphi)) = r > 0$ und somit $Df(r, \varphi)$ für alle $(r, \varphi) \in \mathbb{R}^+ \times \mathbb{R}$ invertierbar.

Also ist f ein lokaler Diffeomorphismus [f ist \mathcal{C}^∞] und

$$D(f^{-1})(x, y) = (Df(\underbrace{f^{-1}(x, y)}_{(r, \varphi)}))^{-1} = \begin{pmatrix} \cos \varphi & -r \cdot \sin \varphi \\ \sin \varphi & r \cdot \cos \varphi \end{pmatrix}^{-1} = \begin{pmatrix} \cos \varphi & \sin \varphi \\ -\frac{\sin \varphi}{r} & \frac{\cos \varphi}{r} \end{pmatrix}.$$

Wegen $x = r \cos \varphi, y = r \sin \varphi$ ist $\sqrt{x^2 + y^2} = r$ und $\frac{x}{r} = \cos \varphi, \frac{y}{r} = \sin \varphi$, somit ist

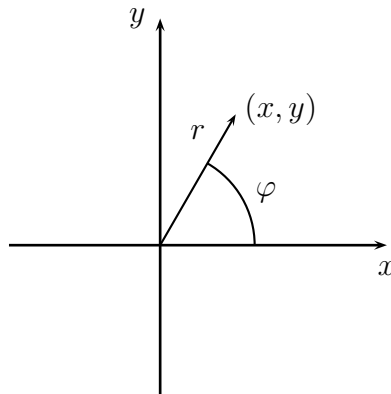
$$D(f^{-1})(x, y) = \begin{pmatrix} \frac{x}{\sqrt{x^2 + y^2}} & \frac{y}{\sqrt{x^2 + y^2}} \\ \frac{-y}{x^2 + y^2} & \frac{x}{x^2 + y^2} \end{pmatrix}.$$

! f ist nicht global bijektiv $\mathbb{R}^+ \times \mathbb{R} \rightarrow \mathbb{R}^2 \setminus \{(0, 0)\}$, denn $f(r, \varphi) = f(r, \varphi + 2k\pi) \quad \forall k \in \mathbb{Z}$.

• Wir geben eine explizite lokale Umkehrung für $-\frac{\pi}{2} < \varphi < \frac{\pi}{2}$ an:

$x = r \cos \varphi > 0$ für $(r, \varphi) \in V := \mathbb{R}^+ \times]-\frac{\pi}{2}, \frac{\pi}{2}[$, daher ist $\frac{y}{x} = \tan \varphi$;

somit ist die Umkehrfunktion $\mathbb{R}^+ \times \mathbb{R} \rightarrow V$ gegeben durch $(x, y) \mapsto (r(x, y), \varphi(x, y)) = (\sqrt{x^2 + y^2}, \arctan \frac{y}{x})$ und die Polarkoordinatenabbildung f ist bijektiv von V nach $\mathbb{R}^+ \times \mathbb{R}$, dort also ein \mathcal{C}^∞ -Diffeomorphismus.



VIII KURVEN UND FLÄCHEN — UNTERMANNIGFALTIGKEITEN DES \mathbb{R}^n

§21. Wege und Kurven

21.1. Definition

Es sei $I \subseteq \mathbb{R}$ ein Intervall.

1) Eine stetige Abbildung $\gamma: I \rightarrow \mathbb{R}^n$ heißt *Weg*. Ist $I = [a, b]$, dann ist γ ein *Weg von* $p := \gamma(a)$ *nach* $q := \gamma(b)$.

2) Es sei $\gamma: I \rightarrow \mathbb{R}^n$ ein differenzierbarer Weg, dann heißt $\dot{\gamma}(t) := D\gamma(t)$ *Tangentenvektor* an γ im Punkt $\gamma(t)$. Falls $\dot{\gamma}(t) \neq 0$ ist, so heißt $\frac{\dot{\gamma}(t)}{\|\dot{\gamma}(t)\|}$ *Tangenteneinheitsvektor*.

3) Ein Weg $\gamma: I \rightarrow \mathbb{R}^n$ heißt *regulär*, wenn γ stetig differenzierbar ist und $\dot{\gamma}(t) \neq 0$ gilt für alle $t \in I$.

4) Ein Weg $\gamma: I \rightarrow \mathbb{R}^n$ heißt *stückweise regulär*, wenn γ stetig ist und es eine Zerlegung $t_0 < t_1 < \dots < t_N$ von I gibt, sodass die Einschränkungen $\gamma|_{[t_j, t_{j+1}]}$ (für $j = 0, \dots, N-1$) jeweils regulär sind.

Kinematische Interpretation: In einer gängigen physikalischen Interpretation von Wegen betrachten wir I als Zeitintervall, $\gamma(t)$ als Ort eines Teilchens (oder eines Körperschwerpunktes) zur Zeit t , $\dot{\gamma}(t)$ als (momentanen) Geschwindigkeitsvektor (engl. *velocity*) und $\|\dot{\gamma}(t)\|$ als (Betrag der momentanen) Geschwindigkeit (engl. *speed*) zur Zeit t .

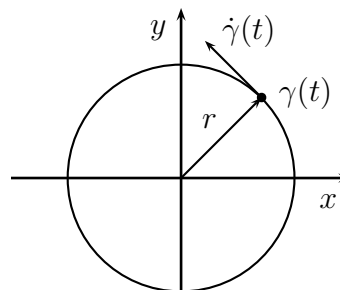
21.2. Beispiele

- 1.) Seien $a, v \in \mathbb{R}^n$, dann beschreibt $\gamma: \mathbb{R} \rightarrow \mathbb{R}^n$ mit $\gamma(t) = a + t \cdot v$ die Gerade durch a in Richtung v . Es ist $\dot{\gamma}(t) = v$ und somit gilt: γ ist regulär $\iff v \neq 0$.

- 2.) Es sei $r > 0$; $\gamma: [0, 2\pi] \rightarrow \mathbb{R}^2$ mit $\gamma(t) = (r \cdot \cos t, r \cdot \sin t)$ beschreibt einen Kreis vom Radius r um den Ursprung. Es gilt

$$\dot{\gamma}(t) = (-r \sin t, r \cos t) \perp \gamma(t).$$

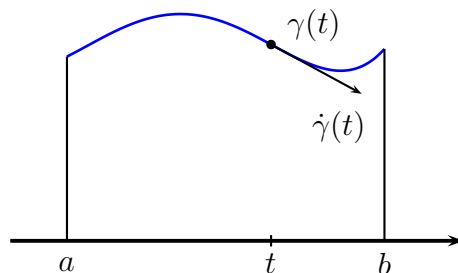
Insbesondere ist γ regulär, weil $\|\dot{\gamma}(t)\| = r > 0$ ist.



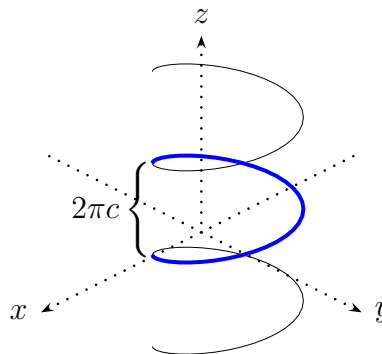
- 3.) Sei $f: I \rightarrow \mathbb{R}$ stetig differenzierbar und $\gamma: I \rightarrow \mathbb{R}^2, t \mapsto (t, f(t))$.

Dann ist $\gamma(I)$ gerade der Graph von f .

Wegen $\dot{\gamma}(t) = (1, f'(t)) \neq (0, 0)$ ist γ stets regulär.



- 4.) Es seien $r > 0$ und $c > 0$; der reguläre Weg $\gamma(t) = (r \cos t, r \sin t, ct)$, $\gamma: [0, 2\pi] \rightarrow \mathbb{R}^3$ beschreibt eine Schraubenlinie mit Ganghöhe $2\pi c$.

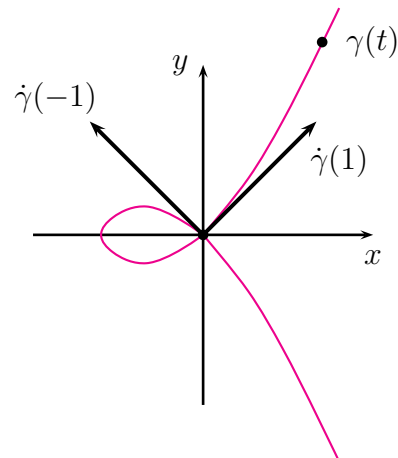


5.) Für $\gamma: \mathbb{R} \rightarrow \mathbb{R}^2$, $\gamma(t) = (t^2 - 1, t^3 - t)$ ist

$$\dot{\gamma}(t) = (2t, 3t^2 - 1)$$

und daher γ regulär.

γ ist nicht injektiv: es gibt einen so genannten *Doppelpunkt* in $(0, 0) = \gamma(-1) = \gamma(1)$, wobei $\dot{\gamma}(-1) = (-2, 2)$ und $\dot{\gamma}(1) = (2, 2)$.



6.) Die Neilsche Parabel $\gamma: \mathbb{R} \rightarrow \mathbb{R}^2$, wobei $\gamma(t) = (t^2, t^3)$; es ist

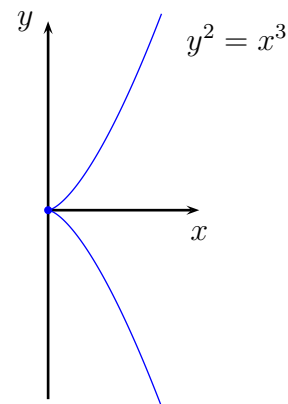
$$\dot{\gamma}(t) = (2t, 3t^2),$$

daher $\dot{\gamma}(0) = (0, 0)$ und γ nicht regulär.

Dieselbe Bildmenge $C := \gamma(\mathbb{R})$ wird durch den stückweise regulären Weg

$$\alpha(t) = \begin{cases} (-t, -|t|^{3/2}) & t < 0, \\ (t, t^{3/2}) & t \geq 0 \end{cases}$$

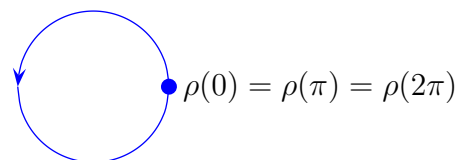
erzeugt.



7.) $\sigma: [0, \pi] \rightarrow \mathbb{R}^2$, $\sigma(\tau) := (r \cos(2\tau), r \sin(2\tau))$ beschreibt wie Beispiel 2.) ebenfalls einen Kreis vom Radius r um den Ursprung, d.h. es gilt $\sigma([0, \pi]) = \gamma([0, 2\pi])$.

Die Durchlaufgeschwindigkeit ist aber wegen $\dot{\sigma}(\tau) = 2 \cdot \dot{\gamma}(2\tau)$ verdoppelt.

8.) Der Weg $\rho: [0, 2\pi] \rightarrow \mathbb{R}^2$, $\rho(\tau) := (r \cos(2\tau), r \sin(2\tau))$ beschreibt denselben Kreis wie in den Beispielen 2.) und 7.). Allerdings wird nun die Bildmenge zweimal durchlaufen.



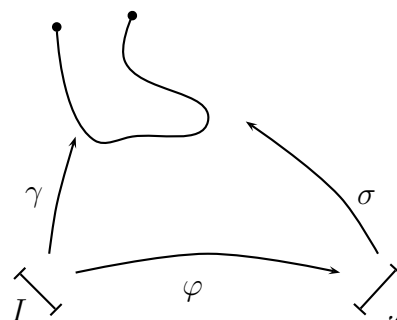
21.3. Definition

Es seien $I, J \subseteq \mathbb{R}$ Intervalle.

1.) Eine *zulässige Parametertransformation* ist eine stetig differenzierbare Abbildung $\varphi: I \rightarrow J$ mit $\varphi'(t) > 0$ für alle $t \in I$.

2.) Zwei Wege $\gamma: I \rightarrow \mathbb{R}^n$ und $\sigma: J \rightarrow \mathbb{R}^n$ heißen *äquivalent*, wenn es eine zulässige Parametertransformation $\varphi: I \rightarrow J$ gibt mit $\sigma \circ \varphi = \gamma$; wir schreiben dafür auch kurz $\gamma \sim \sigma$.

Es ist leicht zu zeigen (nämlich durch entsprechende Verknüpfungen bzw. Inverse von Parametertransformationen), dass dadurch eine Äquivalenzrelation auf der Menge aller Wege im \mathbb{R}^n definiert wird.

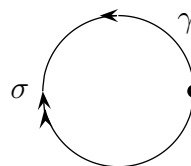
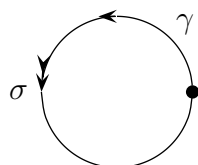


3.) Eine *orientierte stückweise reguläre Kurve* C ist eine Äquivalenzklasse von stückweise regulären Wegen. Jeder Repräsentant γ von C heißt eine *Parametrisierung* von C ; es entspricht dann also C der Klasse aller Wege σ mit $\sigma \sim \gamma$.

21.4. Bemerkung

1.) Die Bedingung $\varphi'(t) > 0$ bedeutet geometrisch, dass γ und σ im selben Sinn durchlaufen werden.

Im Fall $\varphi'(t) < 0$ dreht sich die Orientierung gerade um, wobei die Bildmengen nach wie vor gleich sind.



2.) Wegen $\varphi'(t) > 0$ (für $t \in I$) ist eine zulässige Parametertransformation also streng monoton wachsend und stetig differenzierbar invertierbar (also ein \mathcal{C}^1 -Diffeomorphismus $I \rightarrow \varphi(I)$). Es gilt $(\varphi^{-1})' = \frac{1}{\varphi' \circ \varphi^{-1}} > 0$, daher ist auch φ^{-1} eine zulässige Parametertransformation $\varphi(I) \rightarrow I$.

3.) Für die Änderung der Tangentialvektoren differenzierbarer Wege γ, σ unter einer Parametertransformation gilt

$$\dot{\gamma}(t) = (\sigma \circ \varphi)'(t) = \varphi'(t) \cdot \dot{\sigma}(\varphi(t)).$$

21.5. Definition

Es sei $\gamma: [a, b] \rightarrow \mathbb{R}^n$ ein stückweise regulärer oder stetig differenzierbarer (nicht notwendig regulärer) Weg. Dann heißt

$$(21.1) \quad L(\gamma) := \int_a^b \|\dot{\gamma}(t)\| dt$$

die *Bogenlänge* von γ .

21.6. Bemerkung

- 1.) $t \mapsto \|\dot{\gamma}(t)\|$ ist stetig bis auf höchstens endlich viele Sprungstellen, also Riemann-integrierbar auf $[a, b]$.
- 2.) In der kinematischen Interpretation ist $\|\dot{\gamma}(t)\|$ der Betrag der Momentangeschwindigkeit und somit entspricht $\int_a^b \|\dot{\gamma}(t)\| dt$ dem zurückgelegten Weg.
- 3.) Man kann zeigen: $L(\gamma)$ ist der Limes der Gesamtlängen eingeschriebener Polygonzüge; dies ergibt Riemann-Summen für das Integral in (21.1) und führt auf den Begriff des *rektifizierbaren* Weges (vgl. [For05, §4]).
- 4.) Es sei $\gamma: [0, 1] \rightarrow \mathbb{R}^n$ das Geradenstück $\gamma(t) = a + t \cdot (b - a)$ zwischen a und b im \mathbb{R}^n . Dann gilt $L(\gamma) = \int_0^1 \|b - a\| dt = \|b - a\|$; d.h. für gerade Strecken ist die Bogenlänge genau die euklidische Länge.



21.7. Proposition

- 1.) Sei $I := I_1 \cup I_2$, wobei $I_1 = [a, b]$ und $I_2 = [b, c]$ mit $a < b < c$. Weiters seien $\gamma_1: I_1 \rightarrow \mathbb{R}^n$, $\gamma_2: I_2 \rightarrow \mathbb{R}^n$ stückweise reguläre Wege mit $\gamma_1(b) = \gamma_2(b)$. Wir definieren den stückweise regulären Summenweg $\gamma = \gamma_1 + \gamma_2: I \rightarrow \mathbb{R}^n$ durch

$$\gamma(t) := \begin{cases} \gamma_1(t) & a \leq t \leq b \\ \gamma_2(t) & b \leq t \leq c. \end{cases}$$

Dann gilt $L(\gamma) = L(\gamma_1) + L(\gamma_2)$. (Additivität der Weglänge)

- 2.) Invarianz der Bogenlänge unter zulässiger Parametertransformation: Sind die stückweise regulären Wege $\gamma: [a, b] \rightarrow \mathbb{R}^n$ und $\sigma: [\alpha, \beta] \rightarrow \mathbb{R}^n$ äquivalent ($\gamma \sim \sigma$), dann gilt $L(\gamma) = L(\sigma)$.

Daher ist die Weglänge von (orientierten) stückweise regulären Kurven wohldefiniert.

Beweis: 1.) folgt aus $\int_a^c \|\dot{\gamma}(t)\| dt = \int_a^b \|\dot{\gamma}_1(t)\| dt + \int_b^c \|\dot{\gamma}_2(t)\| dt$.

2.) Sei $\varphi: [a, b] \rightarrow [c, d]$ eine zulässige Parametertransformation und $\sigma \circ \varphi = \gamma$. Wegen $\dot{\gamma} = (\dot{\sigma} \circ \varphi) \cdot \varphi'$ folgt durch Substitution ($\tau = \varphi(t)$)

$$L(\gamma) = \int_a^b \|\dot{\gamma}(t)\| dt = \int_a^b \|\dot{\sigma}(\varphi(t))\| \cdot \underbrace{\varphi'(t)}_{>0} dt = \int_{\varphi(a)}^{\varphi(b)} \|\dot{\sigma}(\tau)\| d\tau = L(\sigma).$$

□

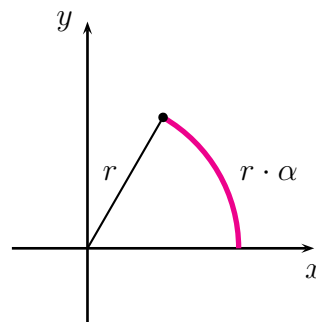
21.8. Beispiele

1.) Für $\alpha > 0$ und $r > 0$ sei $\gamma: [0, \alpha] \rightarrow \mathbb{R}^2$ der Kreisbogen $\gamma(t) = (r \cdot \cos t, r \cdot \sin t)$. Wegen $\dot{\gamma}(t) = (-r \cdot \sin t, r \cdot \cos t)$ ist $\|\dot{\gamma}(t)\| = r$ und daher

$$L(\gamma) = \int_0^\alpha \|\dot{\gamma}(t)\| dt = r \int_0^\alpha dt = r \cdot \alpha.$$

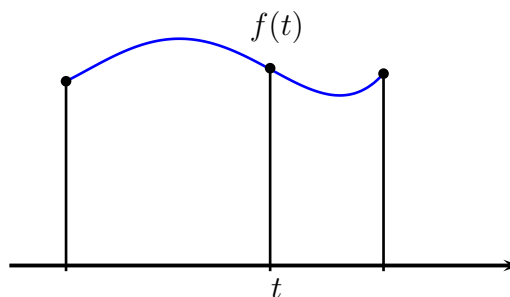
Insbesondere ergibt sich für den Einheitskreisumfang 2π (setze $r = 1$ und $\alpha = 2\pi$).

Der Parameter $t \in [0, \alpha]$ beschreibt hier also den Winkel genau im so genannten Bogenmaß.



2.) Sei $f: [a, b] \rightarrow \mathbb{R}$ stetig differenzierbar, so dass $\gamma(t) := (t, f(t))$ den Graphen von f beschreibt. Dann gilt

$$L(\gamma) = \int_a^b \|(1, f'(t))\| dt = \int_a^b \sqrt{1 + f'(t)^2} dt.$$



21.9. Parametrisierung nach der Bogenlänge

Wir geben nun für orientierte reguläre Kurven eine ausgezeichnete Parametrisierung $\tilde{\gamma}$ an, in der stets $\|\tilde{\gamma}'(s)\| = 1$ gilt, also der Tangentialvektor normiert ist.

Sei $\gamma: [a, b] \rightarrow \mathbb{R}^n$ ein regulärer Weg. Wir suchen ein kompaktes Intervall $J \subseteq \mathbb{R}$ und eine stetig differenzierbare, bijektive Funktion $\varphi: J \rightarrow [a, b]$ mit $\varphi'(s) > 0$ für alle $s \in J$, sodass für $\tilde{\gamma} := \gamma \circ \varphi$ gilt $\|\tilde{\gamma}'(s)\| = 1$ für alle $s \in J$.

Wegen $\dot{\tilde{\gamma}} = (\dot{\gamma} \circ \varphi) \cdot \varphi'$ muss daher $1 = \|\dot{\gamma}(\varphi(s))\| \cdot \varphi'(s)$ gelten. Daraus folgt mit $t := \varphi(s)$ nun

$$(\varphi^{-1})'(t) = \frac{1}{\varphi'(s)} = \|\dot{\gamma}(t)\|$$

und schließlich nach dem Hauptsatz der Differential- und Integralrechnung

$$\varphi^{-1}(t) = \varphi^{-1}(a) + \int_a^t \|\dot{\gamma}(\tau)\| d\tau \quad (a \leq t \leq b).$$

Wir dürfen uns $\varphi^{-1}(a) = 0$ wünschen und erhalten somit $\varphi^{-1}(t) = \int_a^t \|\dot{\gamma}(\tau)\| dt$. Das heißt die gesuchte Parametrisierung $\tilde{\gamma} = \gamma \circ \varphi$ verwendet dann den neuen Parameter

$$s = \varphi^{-1}(t) = L(\gamma|_{[a,t]}) \in [0, L(\gamma)],$$

den so genannten *Bogenlängenparameter*. Der Weg $\tilde{\gamma}: [0, L(\gamma)] \rightarrow \mathbb{R}^n$ heißt *Parametrisierung von γ nach der Bogenlänge*.

Die Berechnung des Bogenlängenparameters gemäß $s = L(\gamma|_{[a,t]})$ ist natürlich auch für stetig differenzierbare (nicht notwendig reguläre) Wege möglich. Es wird i.A. aber dadurch keine zulässige Parametrisierung erzeugt. (Warum?)

Bemerkung: Ist C eine reguläre Kurve und γ ein regulärer Weg, der C repräsentiert, so gilt für seine Parametrisierung $\tilde{\gamma}$ nach der Bogenlänge

$$\|\dot{\tilde{\gamma}}(s)\|^2 = \langle \dot{\tilde{\gamma}}(s) | \dot{\tilde{\gamma}}(s) \rangle = 1 \quad \forall s \in [0, L(\gamma)].$$

Durch Differenzieren nach s erhalten wir daraus $2\langle D\dot{\tilde{\gamma}}(s) | \dot{\tilde{\gamma}}(s) \rangle = 0$, d.h. es gilt mit $\ddot{\tilde{\gamma}}(s) := D\dot{\tilde{\gamma}}(s)$ als *Beschleunigungsvektor* stets

$$\ddot{\tilde{\gamma}}(s) \perp \dot{\tilde{\gamma}}(s).$$

[Warnung: diese Relation gilt i.A. nur für die Parametrisierung nach der Bogenlänge!]

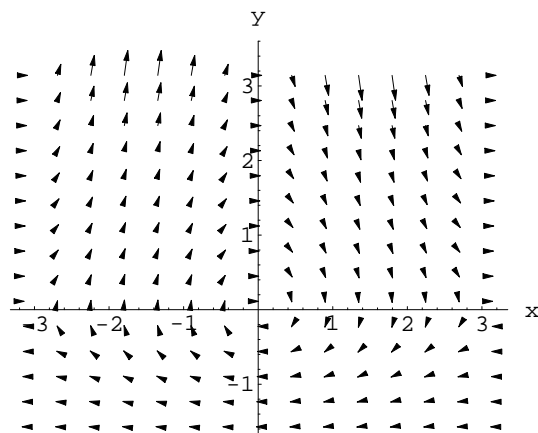
Die Größe $\kappa(s) := \|\ddot{\tilde{\gamma}}(s)\|$ heißt *Krümmung* der Kurve C im Punkt $\tilde{\gamma}(s)$.

§22. Vektorfelder, 1-Formen und Kurvenintegrale

22.1. Definition

Es sei $U \subseteq \mathbb{R}^n$. Eine Abbildung $v: U \rightarrow \mathbb{R}^n$ heißt *Vektorfeld* auf U .

Wir stellen uns dabei die Abbildung v oft so vor, dass an jedem Punkt $p \in U$ der (Tangential-) Vektor $v(p)$ angeheftet wird. In der Physik entspricht das einem (zeitunabhängigen) *Kraftfeld*.



22.2. Beispiel

1.) Obige Graphik zeigt das Vektorfeld $v: \mathbb{R}^2 \rightarrow \mathbb{R}^2$, $v(x, y) = (y, -e^y \sin x)$ auf (einem Ausschnitt von) \mathbb{R}^2 . Der Plot wurde erstellt in MATHEMATICA mit der Eingabe

```
<< Graphics'PlotField'
```

```
PlotVectorField[{y, - Exp[y]Sin[x]}, {x, -Pi, Pi}, {y, -Pi/2, Pi},  
Axes -> True, AxesLabel -> {x, y}]
```

2.) Sei $f: \mathbb{R}^n \supseteq_{\text{offen}} U \rightarrow \mathbb{R}$ eine \mathcal{C}^1 -Funktion, dann definiert $v(p) := \text{grad } f(p)$ (für $p \in U$) ein stetiges Vektorfeld $v: U \rightarrow \mathbb{R}^n$. Wir sagen, v sei ein *Gradientenfeld*. In der Physik nennt man in diesem Fall $-f$ ein *Potenzial* für v .

22.3. Vektorfelder und 1-Formen

Sei v ein Vektorfeld auf der offenen Teilmenge $U \subseteq \mathbb{R}^n$ und bezeichne $\langle \cdot | \cdot \rangle$ das Standard-Skalarprodukt auf \mathbb{R}^n .

Für jedes $p \in U$ definieren wir ein lineares Funktional auf \mathbb{R}^n durch $\omega(p): \mathbb{R}^n \rightarrow \mathbb{R}$, $\omega(p)(h) := \langle v(p) | h \rangle$ für alle $h \in \mathbb{R}^n$; also ist $\omega(p)$ ein Element im Dualraum $(\mathbb{R}^n)^*$.

Vermöge der Zuordnung $p \mapsto \omega(p)$ erhalten wir eine Abbildung $U \rightarrow (\mathbb{R}^n)^*$. In der betrachteten Situation haben wir diese Abbildung aus einem gegebenen Vektorfeld erzeugt. Den allgemeinen Fall fassen wir in die folgende

Definition: Eine 1-Form oder Pfaffsche Form auf U ist eine Abbildung $\omega: U \rightarrow (\mathbb{R}^n)^*$.

Beispiele:

- 1.) Ist $f: U \rightarrow \mathbb{R}$ differenzierbar, so ist $Df: U \rightarrow L(\mathbb{R}^n, \mathbb{R}) = (\mathbb{R}^n)^*$ und definiert also eine 1-Form, die wir mit df bezeichnen und (*äußeres Differential*) von f nennen. [Die Notation $df(p)$ soll den Standpunkt des linearen Funktionals betonen, während wir ja bisher $Df(p)$ zur Vereinfachung meist mit dem Zeilenvektor der partiellen Ableitungen ($(1 \times n)$ -Jacobi-Matrix) identifiziert haben.]

Es gilt für alle $p \in U$ und für alle $h \in \mathbb{R}^n$:

$$df(p)(h) = \langle \text{grad } f(p) \mid h \rangle = Df(p) \cdot h.$$

- 2.) Als Spezialfall von 1.) für $f = pr_j|_U: U \rightarrow \mathbb{R}$, die j -te Koordinatenprojektion mit $pr_j(x_1, \dots, x_n) = x_j$ erhalten wir als Differential dpr_j eine 1-Form, die wir mit dx_j bezeichnen.

Ist $\{e_1, \dots, e_n\}$ die Standardbasis in \mathbb{R}^n , so gilt $\forall p \in U$:

$$dx_j(p)(e_k) = \langle e_j \mid e_k \rangle = \delta_{jk},$$

d.h. $\{dx_1(p), \dots, dx_n(p)\}$ ist die zu $\{e_1, \dots, e_n\}$ duale Basis in $(\mathbb{R}^n)^*$. Insbesondere lautet die Darstellung von $df(p)$ aus 1.) in dieser Basis

$$(22.1) \quad df(p) = \sum_{j=1}^n D_j f(p) \cdot dx_j(p).$$

Wir können ähnlich wie in (22.1) sogar für jede 1-Form ω eine punktweise Basisentwicklung angeben, in der ω durch passende Koeffizientenfunktionen $f_1, \dots, f_n: U \rightarrow \mathbb{R}$ vertreten wird. Zugleich erlaubt dies eine Identifizierung von Vektorfeldern mit 1-Formen (auf der festen offenen Teilmenge $U \subseteq \mathbb{R}^n$).

Proposition

Es sei $U \subseteq \mathbb{R}^n$ offen.

- 1.) Zu jeder 1-Form $\omega: U \rightarrow (\mathbb{R}^n)^*$ gibt es genau eine Abbildung $f = (f_1, \dots, f_n): U \rightarrow \mathbb{R}^n$ mit der Eigenschaft

$$(22.2) \quad \omega(p) = \sum_{j=1}^n f_j(p) \cdot dx_j(p) \quad \forall p \in U.$$

Es gilt $f_j(p) = \omega(p)(e_j)$ für jedes $p \in U$ und $j = 1, \dots, n$.

- 2.) Die Zuordnung $v = (v_1, \dots, v_n) \mapsto \hat{v} = \sum_{j=1}^n v_j \cdot dx_j$ liefert einen Isomorphismus (d. h. eine lineare und bijektive Abbildung) des \mathbb{R} -Vektorraumes von Vektorfeldern mit jenem der 1-Formen auf U .

Beweis: *ad 1.):* Für jedes $p \in U$ ist $\{dx_1(p), \dots, dx_n(p)\}$ eine Basis von $(\mathbb{R}^n)^*$. Daher gilt $\forall p \in U$: es gibt eindeutige reelle Zahlen $\lambda_1, \dots, \lambda_n$ mit

$$\omega(p) = \sum_{j=1}^n \lambda_j dx_j(p).$$

Es ist $\lambda_j = \lambda_j dx_j(p)(e_j) = \omega(p)(e_j)$. Wir setzen $f_j(p) := \lambda_j$.

Somit haben wir Funktionen $f_1, \dots, f_n: U \rightarrow \mathbb{R}$ definiert, für die gilt

$$\omega = \sum_{j=1}^n f_j dx_j.$$

Die Eindeutigkeit von f_1, \dots, f_n ist klar wegen der Eindeutigkeit der skalaren Faktoren in der Basisdarstellung für jedes $p \in U$.

ad 2.): Die Menge aller Vektorfelder auf U sowie die Menge der 1-Formen auf U bilden Vektorräume über \mathbb{R} , wenn die Operation der Vektoraddition und skalaren Multiplikation punktweise ausgeführt werden. [Für $\lambda, \mu \in \mathbb{R}$ und v, w Vektorfelder bzw. 1-Formen ω, ν ist $p \mapsto \lambda \cdot v(p) + \mu \cdot w(p)$ wieder ein Vektorfeld bzw. $p \mapsto \lambda \cdot \omega(p) + \mu \cdot \nu(p)$ eine 1-Form und es gelten alle Vektorraumaxiome.]

Die Linearität der Abbildung $v \mapsto \hat{v}$ ist klar.

Injektivität: falls $\hat{v} = 0$ ist, so folgt daraus $\forall p \in U$, dass $\hat{v}(p) = \sum_{j=1}^n v_j(p) dx_j(p) = 0$; daher muss nach 1.) dann $v_j(p) = 0$ gelten ($j = 1, \dots, n$); da p beliebig war, erzwingt dies $v = 0$.

Surjektivität: sei ω eine 1-Form auf U mit der Darstellung $\omega(p) = \sum_{j=1}^n f_j(p) dx_j(p)$ für alle $p \in U$ gemäß 1.); wir setzen $v := (f_1, \dots, f_n): U \rightarrow \mathbb{R}^n$, dann gilt nach Konstruktion $\hat{v} = \omega$. \square

Definition: Eine 1-Form ω auf der offenen Menge $U \subseteq \mathbb{R}^n$ heißt *stetig* (bzw. *differenzierbar*, *stetig differenzierbar* etc.), wenn alle Komponentenfunktionen $f_1, \dots, f_n: U \rightarrow \mathbb{R}$ in der Darstellung (22.2) es sind.

Bemerkung: Die Identifizierung von Vektorfeldern mit 1-Formen in Punkt 2.) der obigen Proposition baut — bei genauer Betrachtung — auf dem gegebenen Skalarprodukt $\langle \cdot | \cdot \rangle$ im \mathbb{R}^n auf (bzw. hängt von der dadurch vermittelten dualen Basis ab). Ein anderes Skalarprodukt auf \mathbb{R}^n ergäbe in derselben Weise ebenfalls einen Isomorphismus, der aber als lineare Abbildung i.A. verschieden sein wird.

22.4. Wegintegrale

Definition: Sei $U \subseteq \mathbb{R}^n$ offen, ω eine stetige 1-Form auf U und $\gamma: [a, b] \rightarrow U$ ein stetig differenzierbarer Weg. Dann ist

$$(22.3) \quad \int_{\gamma} \omega := \int_a^b \omega(\gamma(t))(\dot{\gamma}(t)) dt$$

das *Wegintegral* von ω über γ .

Wenn ω die Darstellung $\omega = \sum_{j=1}^n f_j \cdot dx_j$ mit stetigen Komponentenfunktionen f_1, \dots, f_n hat, $f = (f_1, \dots, f_n)$ gesetzt wird, und $\gamma = (\gamma_1, \dots, \gamma_n)$ ist, so haben wir für den Integranden in (22.3) den folgenden konkreten Ausdruck

$$\omega(\gamma(t))(\dot{\gamma}(t)) = \sum_{j=1}^n f_j(\gamma(t))\dot{\gamma}_j(t) = \langle f(\gamma(t)) \mid \dot{\gamma}(t) \rangle,$$

woraus auch die Stetigkeit bezüglich t klar ist. Zusammenfassend erhalten wir für die Berechnung des Wegintegrals

$$(22.3') \quad \int_{\gamma} \omega = \sum_{j=1}^n \int_a^b f_j(\gamma(t))\dot{\gamma}_j(t) dt = \int_a^b \langle f(\gamma(t)) \mid \dot{\gamma}(t) \rangle dt.$$

Bemerkung: Falls der Weg γ nur stückweise \mathcal{C}^1 ist, d. h. $\gamma: [a, b] \rightarrow U$ ist stetig und es gibt eine Zerlegung $a = t_0 < t_1 < \dots < t_N = b$, so dass $\sigma_j := \gamma|_{[t_{j-1}, t_j]}$ jeweils \mathcal{C}^1 ist ($j = 1, \dots, N$), dann setzen wir für das Wegintegral

$$\int_{\gamma} \omega := \sum_{j=1}^N \int_{\sigma_j} \omega.$$

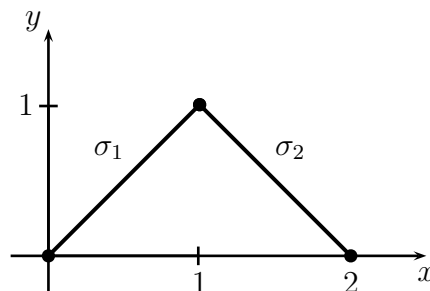
Beispiele: 1.) Sei $\gamma(t) = (\cos t, \sin t)$ für $t \in [0, 2\pi]$ und $\omega(x, y) = -y dx + x dy$ auf \mathbb{R}^2 , dann ist

$$\int_{\gamma} \omega = \int_0^{2\pi} \underbrace{(-\sin t (-\sin t) + \cos t \cos t)}_{=1} dt = 2\pi.$$

2.) Betrachte $\gamma: [0, 2] \rightarrow \mathbb{R}^2$ mit

$$\gamma(t) := \begin{cases} \sigma_1(t) = (t, t) & 0 \leq t \leq 1 \\ \sigma_2(t) = (t, 2-t) & 1 \leq t \leq 2 \end{cases}$$

und $\omega(x, y) = x^2 dx + (x - y) dy$ auf \mathbb{R}^2 .



Wir berechnen mit $\dot{\sigma}_1(t) = (1, 1)$ und $\dot{\sigma}_2(t) = (1, -1)$

$$\begin{aligned} \int_{\gamma} \omega &= \int_{\sigma_1} \omega + \int_{\sigma_2} \omega = \int_0^1 (t^2 \cdot 1 + (t-t) \cdot 1) dt + \int_1^2 (t^2 \cdot 1 + (t-(2-t)) \cdot (-1)) dt \\ &= \int_0^1 t^2 dt + \int_1^2 (t^2 + 2t - 2) dt = \frac{1}{3} + \left(\frac{t^3}{3} + t^2 - 2t \right) \Big|_1^2 \\ &= \frac{1}{3} + \left(\frac{8}{3} + 4 - 4 \right) - \left(\frac{1}{3} + 1 - 2 \right) = \frac{8}{3} + 1 = \frac{11}{3}. \end{aligned}$$

22.5. Lemma (Unabhängigkeit des Wegintegrals von der Parametrisierung)

Sei $U \subseteq \mathbb{R}^n$ offen, ω eine stetige 1-Form, $\gamma: [a, b] \rightarrow U$ ein \mathcal{C}^1 -Weg und $\varphi: [c, d] \rightarrow [a, b]$ eine zulässige Parametertransformation [d.h. $\varphi \in \mathcal{C}^1$ und $\varphi' > 0$]. Dann gilt

$$\int_{\gamma \circ \varphi} \omega = \int_{\gamma} \omega.$$

Bemerkung: Für $\varphi' < 0$, d.h. φ orientierungsumkehrend, so gilt $\int_{\gamma \circ \varphi} \omega = - \int_{\gamma} \omega$.

Beweis: Für den Weg $\sigma := \gamma \circ \varphi$ gilt $\int_{\sigma} \omega = \int_c^d \omega(\sigma(s))(\dot{\sigma}(s)) ds \stackrel{[\dot{\sigma} = \varphi' \cdot \dot{\gamma}]}{=} \int_c^d \omega(\gamma(\varphi(s)))(\dot{\gamma}(\varphi(s)))\varphi'(s) ds$

$$\int_c^d \omega(\gamma(\varphi(s)))(\dot{\gamma}(\varphi(s)))\varphi'(s) ds \stackrel{[\text{Subst. } t=\varphi(s)]}{=} \int_a^b \omega(\gamma(t))(\dot{\gamma}(t)) dt = \int_{\gamma} \omega. \quad \square$$

22.6. Korollar und Definition

Sei $U \subseteq \mathbb{R}^n$ offen, ω eine stetige 1-Form und C eine orientierte reguläre Kurve in U .

Dann ist das *Kurvenintegral* $\int_C \omega$ wohldefiniert wie folgt:

Sei $\gamma: [a, b] \rightarrow U$ ein beliebiger Repräsentant von C , dann ist

$$\int_C \omega := \int_\gamma \omega.$$

(Analog für stückweise reguläre Kurven.)

22.7. Bemerkung (Kurvenintegrale von Vektorfeldern)

Sei $v = (v_1, \dots, v_n): U \rightarrow \mathbb{R}^n$ ein Vektorfeld auf U , dann ist $\hat{v} = \sum_{j=1}^n v_j dx_j$ die entsprechende 1-Form auf U . Symbolisch können wir mit $dx = (dx_1, \dots, dx_n)$ die Identifizierung so schreiben: $\hat{v} = \langle v | dx \rangle$. Somit erhalten wir die Identität

$$\int_C \hat{v} = \int_C \langle v | dx \rangle = \int_a^b \langle v(\gamma(t)) | \dot{\gamma}(t) \rangle dt,$$

deren rechte Seite auch als Definition für das *Kurvenintegral des Vektorfeldes v entlang C* $\int_C \langle v | dx \rangle$ dienen kann.

In der Physik z.B. sind auch die Notationen $d\vec{s} = dx$ und $v \cdot w$ für $\langle v | w \rangle$ üblich. Dann kann der Ausdruck $\int_C v \cdot d\vec{s}$ als Kurvenintegral von v über C im obigen Sinne interpretiert werden.

Eine physikalische Interpretation des Kurvenintegrals fasst dieses als geleistete Arbeit (oder gewonnene Energie — je nach Vorzeichen) im Kraftfeld v bei der Bewegung entlang der Kurve C auf.

22.8. Stammfunktionen

Definition: Sei ω eine stetige 1-Form auf der offenen Teilmenge $U \subseteq \mathbb{R}^n$. Eine *Stammfunktion* von ω ist eine \mathcal{C}^1 -Funktion $F: U \rightarrow \mathbb{R}$ mit $dF = \omega$. Eine 1-Form heißt *exakt*, wenn sie eine Stammfunktion besitzt.

Bemerkung: 1.) Ist F eine Stammfunktion von ω , dann ist für alle $c \in \mathbb{R}$ auch $F + c$ eine Stammfunktion von ω .

2.) Im Fall $n = 1$ können wir stetige 1-Formen stets als $\omega = f dx$ schreiben, wobei f eine stetige Funktion ist. Die Frage nach einer Stammfunktion F für ω führt auf die Gleichung

$$f dx = \omega \stackrel{?}{=} dF = F' dx,$$

daher gilt folgende Äquivalenz:

F Stammfunktion von $\omega \iff F$ ist Stammfunktion von f im Sinne von Analysis 1.

Daraus lesen wir ab: Jede 1-Form auf einem offenem Intervall $U \subseteq \mathbb{R}$ ist exakt.

(Für $x_0 \in U$ fix erhalten wir nämlich stets eine Stammfunktion durch $F(x) := \int_{x_0}^x f(t) dt$.)

[Die Aussage gilt sogar für beliebige offene Teilmengen $U \subseteq \mathbb{R}$; vgl. [AE99, VIII.3.4(c)].]

3.) Sei $v: U \rightarrow \mathbb{R}^n$ ein stetiges Vektorfeld, dann gilt: $\hat{v} = \sum_{j=1}^n v_j \cdot dx_j$ ist exakt

$$\iff \exists F: U \rightarrow \mathbb{R} \mathcal{C}^1 \text{ mit } \sum_{j=1}^n D_j F dx_j = dF = \hat{v} = \sum_{j=1}^n v_j dx_j$$

$$\iff \exists F: U \rightarrow \mathbb{R} \mathcal{C}^1 \text{ mit } v = \text{grad } F.$$

D.h.: v ist ein Gradientenfeld $\iff \hat{v}$ ist exakt

In der Physik sagt man in dem Fall auch, dass v das Potenzial $\Phi = -F$ besitzt.

Folgende Fragen werden wir im Rest dieses Abschnittes behandeln:

- 1) Welchen Nutzen haben Stammfunktionen?
- 2) Welche 1-Formen $\omega: U \rightarrow (\mathbb{R}^n)^*$ besitzen Stammfunktionen?
- 3) Wie findet man Stammfunktionen zu gegebenen 1-Formen?

Proposition: Sei $U \subseteq \mathbb{R}^n$ offen, ω eine exakte stetige 1-Form auf U und F eine Stammfunktion von ω (auf U).

Für jeden stückweisen \mathcal{C}^1 -Weg $\gamma: [a, b] \rightarrow U$ mit Anfangs- und Endpunkten $p := \gamma(a)$, $q := \gamma(b) \in U$ gilt:

$$(22.4) \quad \int_{\gamma} \omega = \int_{\gamma} dF = F(q) - F(p).$$

Insbesondere gilt: wenn γ (zusätzlich) *geschlossen* ist, d.h. $\gamma(a) = \gamma(b)$, dann folgt

$$\int_{\gamma} \omega = 0.$$

Beweis: Zunächst sei γ stetig differenzierbar. Dann ist nach der Kettenregel

$$\frac{d}{dt}(F(\gamma(t))) = \langle \text{grad } F(\gamma(t)) \mid \dot{\gamma}(t) \rangle = dF(\gamma(t))(\dot{\gamma}(t))$$

und daher

$$\int_{\gamma} dF = \int_a^b dF(\gamma(t))(\dot{\gamma}(t)) dt = \int_a^b \frac{d}{dt}(F(\gamma(t))) dt = F(\gamma(t)) \Big|_a^b = F(q) - F(p).$$

Falls γ nur stückweise \mathcal{C}^1 ist, gilt mit einer entsprechenden Zerlegung $a = t_0 < \dots < t_N = b$, $\gamma_k = \gamma|_{[t_{k-1}, t_k]}$:

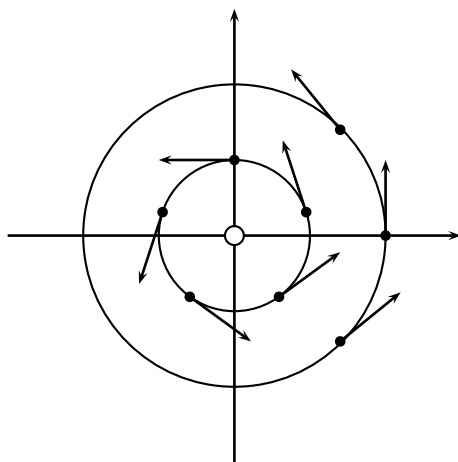
$$\int_{\gamma} dF = \sum_{k=1}^N \int_{\gamma_k} dF = \sum_{k=1}^N (F(\gamma(t_k)) - F(\gamma(t_{k-1}))) = F(\gamma(t_N)) - F(\gamma(t_0)).$$

□

Beispiel: Sei $U := \mathbb{R}^2 \setminus \{0, 0\}$ und

$$\omega(x, y) = \frac{-y}{x^2 + y^2} dx + \frac{x}{x^2 + y^2} dy.$$

Es ist also $\omega = f_1 dx + f_2 dy$ mit $f(x, y) = (f_1(x, y), f_2(x, y)) = \left(\frac{-y}{x^2 + y^2}, \frac{x}{x^2 + y^2} \right)$.



Das Vektorfeld f beschreibt Tangenten an Kreise um den Ursprung, die gegen den Uhrzeigersinn durchlaufen werden.

Sei $\gamma: [0, \alpha] \rightarrow U$ gegeben durch $\gamma(t) = (r \cos t, r \sin t)$ mit $r > 0$ und $\alpha > 0$:

$\dot{\gamma}(t) = (-r \sin t, r \cos t)$ und $\omega(\gamma(t))(\dot{\gamma}(t)) = \left\langle \frac{1}{r^2} \begin{pmatrix} -r \sin t \\ r \cos t \end{pmatrix} \mid \begin{pmatrix} -r \sin t \\ r \cos t \end{pmatrix} \right\rangle = 1$, daher folgt

$$\int_{\gamma} \omega = \int_0^{\alpha} 1 dt = \alpha.$$

Insbesondere ergibt sich für $\alpha = 2\pi$ und $r = 1$ ein geschlossener Kreis und

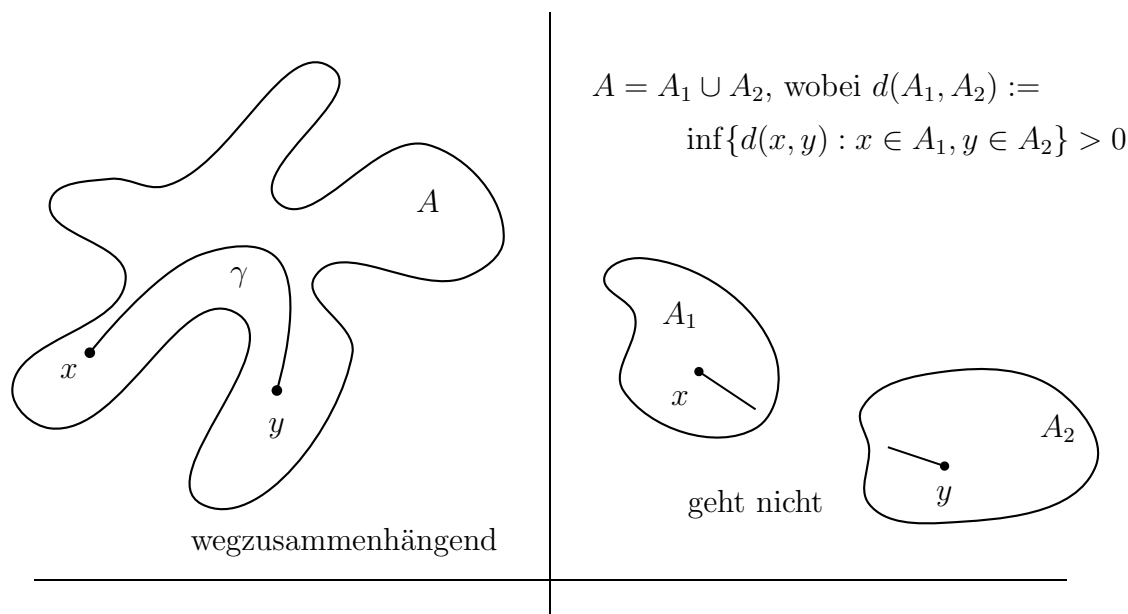
$$\int_{\gamma} \omega = 2\pi \neq 0 :$$

aus obiger Proposition folgt also, dass ω auf U keine Stammfunktion besitzen kann!

Aus den Beobachtungen in 22.8 können wir die Idee gewinnen, die Existenz von Stammfunktionen für eine 1-Form ω auf U durch Integrale über geschlossene Kurven zu testen. Die entscheidende Frage ist dann aber, ob die Bedingung, dass geschlossene Kurvenintegrale allesamt verschwinden müssen auch hinreichend für die Existenz einer Stammfunktion ist?

22.9. Definition

- 1.) Eine Teilmenge $A \subseteq \mathbb{R}^n$ heißt *wegzusammenhängend*, wenn es für alle Punkte $x, y \in A$ einen stetigen Weg von x nach y gibt, d. h. $\gamma: [0, 1] \rightarrow A \subseteq \mathbb{R}^n$ stetig mit $\gamma(0) = x$ und $\gamma(1) = y$.



- 2.) Ein *Gebiet* im \mathbb{R}^n ist eine Teilmenge U , die offen und wegzusammenhängend ist.

[Bemerkung zu einem allgemeineren Begriff: Eine Teilmenge A eines metrischen Raumes heißt *zusammenhängend*, falls es keine zwei offenen, nichtleere und disjunkte Mengen O_1 und O_2 in X geben kann mit $A = (O_1 \cap A) \cup (O_2 \cap A)$.

Im Allgemeinen gilt 'wegzusammenhängend \Rightarrow zusammenhängend', aber nicht umgekehrt (vgl. [vQ01, Kapitel 4]).

Ist nun U eine offene Teilmenge eines normierten Vektorraumes, so gilt allerdings:
 U zusammenhängend $\iff U$ wegzusammenhängend.

Für einen Beweis siehe z.B. [AE02, Kap. III, Korollar 4.11]; d.h. insbesondere hätten wir bei der Definition von Gebiet auch „nur“ zusammenhängend verlangen müssen.]

Die grundlegende Idee der Konstruktion eines Kandidaten für eine Stammfunktion F einer gegebenen stetigen 1-Form auf U ist nun folgende: wir halten einen Punkt $p \in U$ fest und definieren Funktionswerte $F(q)$ für $q \in U$, indem wir jeweils einen Weg γ von p nach q wählen und angeleitet durch Gleichung (22.4) ansetzen: $F(q) := F(p) + \int_{\gamma} \omega$, wobei der fixe Wert $F(p)$ beliebig sein kann.

Ist U ein Gebiet, so wird uns die Existenz eines stetigen Weges γ von p nach q garantiert — für die Kurvenintegrale benötigen wir jedoch stückweise \mathcal{C}^1 -Kurven. Diesen technischen Zusatz verschaffen wir uns im folgenden

22.10. Lemma

- 1.) Sei (X, d) ein metrischer Raum, $K \subseteq X$ kompakt und $A \subseteq X$ abgeschlossen mit $A \cap K = \emptyset$. Dann gilt: $\exists c > 0$ derart, dass

$$\forall x \in K \quad \forall y \in A: \quad d(x, y) \geq c.$$

- 2.) Sei $U \subseteq \mathbb{R}^n$ offen und $\sigma: [0, 1] \rightarrow U$ stetig; wir setzen $p := \sigma(0)$ und $q := \sigma(1)$. Dann gibt es auch einen stückweise stetig differenzierbaren Weg γ in U , der p und q verbindet, d. h. $\exists \gamma: [0, 1] \rightarrow U \subseteq \mathbb{R}^n$ stückweise \mathcal{C}^1 mit $\gamma(0) = p$ und $\gamma(1) = q$.

Beweis: 1.) wurde als Übungsaufgabe zu §17 bewiesen.

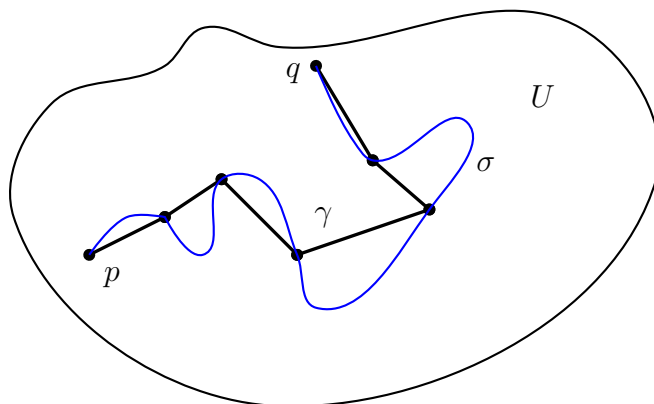
- 2.) Es ist $K := \sigma([0, 1]) \subseteq \mathbb{R}^n$ kompakt und $A := \mathbb{R}^n \setminus U$ abgeschlossen mit $A \cap K = \emptyset$. Nach 1.) gibt es also ein $c > 0$, sodass

$$(*) \quad \forall t \in [0, 1] \quad \forall y \in \mathbb{R}^n \setminus U: \quad \|\sigma(t) - y\| \geq c.$$

Als stetige Abbildung auf einer kompakten Menge ist σ gleichmäßig stetig, daher gibt es eine Zerlegung $0 = t_0 < t_1 < \dots < t_m = 1$ mit

$$\|\sigma(t_j) - \sigma(t_{j-1})\| < c \quad (j = 1, \dots, m).$$

Wir definieren nun $\gamma: [0, 1] \rightarrow U$ als Polygonzug mit den Ecken $\sigma(t_0), \dots, \sigma(t_m)$,



d.h. $\gamma(\lambda t_j + (1 - \lambda)t_{j-1}) := \lambda\sigma(t_j) + (1 - \lambda)\sigma(t_{j-1})$ für $0 \leq \lambda \leq 1, j = 1, \dots, m$.

Dann ist γ stückweise gerade mit stetigen Übergängen an den Zerlegungspunkten, daher also stückweise \mathcal{C}^1 . Weiters gilt $\gamma(0) = \sigma(t_0) = p$ und $\gamma(1) = \sigma(t_m) = q$.

Schließlich gilt auch $\gamma([0, 1]) \subseteq U$, denn zu jedem $t \in [0, 1]$ gibt es $\lambda \in [0, 1]$ und $j \in \{1, \dots, m\}$ mit $t = \lambda t_j + (1 - \lambda)t_{j-1}$ und somit

$$\|\gamma(t) - \sigma(t_{j-1})\| = \lambda \|\sigma(t_j) - \sigma(t_{j-1})\| < 1 \cdot c = c;$$

wegen $\sigma(t_{j-1}) \in \sigma([0, 1]) = K$ und (*) muss daher $\gamma(t) \in U$ gelten. □

22.11. Theorem

Sei $U \subseteq \mathbb{R}^n$ ein Gebiet und ω eine stetige 1-Form auf U . Dann sind folgende Aussagen äquivalent:

- 1.) ω besitzt eine Stammfunktion in U (d.h. ω ist exakt in U).
- 2.) Für jeden geschlossenen stückweise stetig differenzierbaren Weg γ in U gilt

$$(22.5) \quad \int_{\gamma} \omega = 0$$

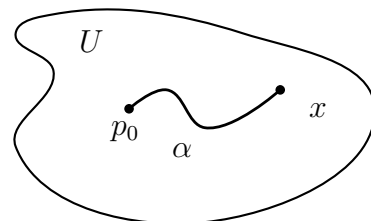
Beweis:

(1.) \Rightarrow (2.): folgt aus Proposition 22.8.

(2.) \Rightarrow (1.): Wir folgen der oben beschriebenen Idee für die Konstruktion einer Stammfunktion F :

sei $p_0 \in U$ beliebig und fix; da U wegzusammenhängend ist, gibt es zu $x \in U$ einen stetigen Weg γ in U von p_0 nach x . Nach Lemma 22.10 gibt es auch einen stückweisen \mathcal{C}^1 -Weg $\alpha: [0, 1] \rightarrow$

U mit $\alpha(0) = p_0$ und $\alpha(1) = x$; wir setzen $F(x) := \int_{\alpha} \omega$.



⊙ F ist wohldefiniert $U \rightarrow \mathbb{R}$:

Sei nämlich $\beta: [0, 1] \rightarrow U$ auch stückweise \mathcal{C}^1 mit $\beta(0) = p_0$ und $\beta(1) = x$, dann bilden wir die Zusammensetzung von α mit β durch den Weg $\gamma: [0, 2] \rightarrow U$ mit

$$\gamma(t) := \begin{cases} \alpha(t) & 0 \leq t \leq 1 \\ \beta(2-t) & 1 \leq t \leq 2 \end{cases}$$

Dann ist γ stückweise \mathcal{C}^1 und geschlossen ($\gamma(0) = p_0 = \gamma(2)$), daher folgt laut Annahme

$$0 = \int_{\gamma} \omega = \int_{\alpha} \omega - \int_{\beta} \omega,$$

wobei wir die Additivität des Kurvenintegrals und 22.5 mit Orientierungsumkehr verwendet haben. Somit gilt $\int_{\alpha} \omega = \int_{\beta} \omega$ und $F(x)$ ist wohldefiniert; daher ist auch die vereinfachte Schreibweise

$$F(x) = \int_{p_0}^x \omega$$

sinnvoll (weil unabhängig vom Weg).

⊙ F ist stetig differenzierbar und es gilt $dF = \omega$:

Es sei $\omega = \sum_{j=1}^n f_j dx_j$. Wir zeigen, dass $f_j = D_j F$ ($1 \leq j \leq n$) gilt, dann sind wir fertig, weil daraus die Stetigkeit aller partiellen Ableitungen von F folgt und weiters auch $dF = \sum D_j dx = \sum f_j dx = \omega$.

Sei wie üblich $\{e_1, \dots, e_n\}$ die Standardbasis des \mathbb{R}^n und $h \in \mathbb{R}$ mit $h \neq 0$ so klein, dass die Strecke $\beta(t) := x + t h e_j$ für $0 \leq t \leq 1$ ganz in U zu liegen kommt. Dann ist

$$F(x + h e_j) - F(x) = \int_{p_0}^{x+h e_j} \omega - \int_{p_0}^x \omega = \int_{\beta} \omega = \int_0^1 \underbrace{\omega(x + t h e_j)(h e_j)}_{f_j(x + t h e_j) \cdot h}$$

und daher

$$D_j F(x) = \lim_{h \rightarrow 0} \frac{F(x + h e_j) - F(x)}{h} = \lim_{h \rightarrow 0} \int_0^1 \underbrace{f_j(x + t h e_j)}_{\rightarrow f_j(x) \text{ [glm.]}} dt = \int_0^1 f_j(x) dt = f_j(x).$$

□

22.12. Integrierbarkeitsbedingungen

Die Bedingung (22.5) ist zwar äquivalent zur Existenz einer Stammfunktion, lässt sich aber keineswegs direkt aus den Koeffizientenfunktionen f_1, \dots, f_n einer stetigen 1-Form ω ablesen. Deshalb untersuchen wir möglichst einfache Folgerungen der Gleichung $\omega = dF$.

Zunächst bedeutet die obige Gleichung wegen $dF = \sum D_j F dx_j$ natürlich, dass $f_j = D_j F$ ($j = 1 \dots, n$) gilt. (F ist aber später eben gesucht — bzw. die Existenz von F in Frage; daher versuchen wir, jegliche Referenz auf Werte von F loszuwerden.)

Angenommen F sei eine \mathcal{C}^2 -Funktion (somit ω also \mathcal{C}^1), dann folgt durch weitere partielle Differentiation mit Hilfe des Satzes von Schwarz 18.6

$$D_i f_j = D_i D_j F = D_j D_i F = D_j f_i \quad (1 \leq i, j \leq n).$$

Durch die Gleichheit der äußersten Terme haben wir also notwendige Bedingung für die Existenz einer zweimal stetig differenzierbaren Stammfunktion vorliegen.

Definition: Sei $U \subseteq \mathbb{R}^n$ offen. Eine stetig differenzierbare 1-Form $\omega = \sum_{j=1}^n f_j dx_j$ auf U heißt *geschlossen*, falls die *Integrabilitätsbedingungen*

$$(22.6) \quad D_j f_i = D_i f_j \quad (1 \leq i, j \leq n)$$

gelten.

Bemerkung:

- 1.) In Dimension $n = 3$ sei ein \mathcal{C}^1 Vektorfeld $v = (f_1, f_2, f_3)$ auf $U \subseteq \mathbb{R}^3$ gegeben: dann bedeutet die Bedingung (22.6) für die entsprechende 1-Form \hat{v} gerade, dass $\text{rot}(v) = 0$ gilt.

Wir hatten bereits in (18.3) gesehen: $v = \text{grad } F \Rightarrow \text{rot } v = 0$; dies übersetzt sich nun mittels $\omega = \hat{v}$ wie folgt: ω exakt $\Rightarrow \omega$ geschlossen.

- 2.) Die Bedingungen (22.6) sind notwendig, aber nicht hinreichend für die Existenz einer Stammfunktion. Um dies zu zeigen, greifen wir zurück auf Beispiel 22.8 mit $U = \mathbb{R}^2 \setminus \{0, 0\}$ und $\omega(x, y) = \frac{-y}{x^2+y^2} dx + \frac{x}{x^2+y^2} dy = f_1(x, y) dx + f_2(x, y) dy$: wir haben dort gezeigt, dass ω keine Stammfunktion in U besitzen kann; dennoch gilt aber

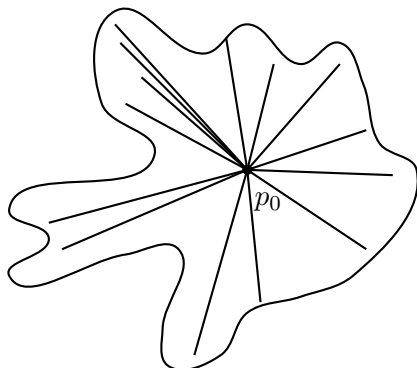
$$\begin{array}{rcl} D_2 f_1(x, y) & = & \frac{-1 \cdot (x^2 + y^2) + y2y}{(x^2 + y^2)^2} = \frac{y^2 - x^2}{(x^2 + y^2)^2} \quad \updownarrow \\ & & \ominus \\ D_1 f_2(x, y) & = & \frac{1 \cdot (x^2 + y^2) - x2x}{(x^2 + y^2)^2} = \frac{y^2 - x^2}{(x^2 + y^2)^2} \quad \updownarrow \end{array}$$

Es zeigt sich, dass der in 22.12, Bemerkung 2.), beobachtete „Defekt“ gewissermaßen in der ungünstigen Gestalt von U liegt: in Gebieten „ohne Löcher“ (so genannte *einfach zusammenhängende* Gebiete) ist für stetig differenzierbare 1-Formen die Bedingung (22.6) gleichwertig mit der Existenz von Stammfunktionen (vgl. [AE99, Kapitel VIII, Theorem 4.8]). Wir werden dies zumindest für eine vereinfachte Variante nachweisen.

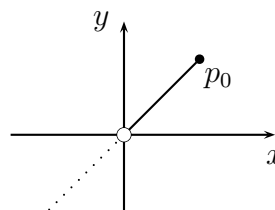
22.13. Sternförmige Teilmengen des \mathbb{R}^n

Definition: Eine Menge $U \subseteq \mathbb{R}^n$ heißt *sternförmig* bezüglich des Punktes $p_0 \in U$, wenn folgendes gilt: $\forall x \in U$ liegt die gesamte Verbindungsstrecke von p_0 nach x in U , d. h. $\{(1-t)p_0 + t \cdot x : 0 \leq t \leq 1\} \subseteq U$.

Mit anderen Worten: „von p_0 aus kann ganz U beleuchtet werden“.



Beispiel: $U = \mathbb{R}^2 \setminus \{0,0\}$ ist nicht sternförmig, denn für jedes $p_0 \in U$ ist stets ein Halbstrahl hinter $(0,0)$ verdeckt.



22.14. Theorem

Sei $U \subseteq \mathbb{R}^n$ ein sternförmiges Gebiet und ω eine stetig differenzierbare 1-Form. Dann sind folgende Aussagen äquivalent:

- 1.) ω ist exakt (d.h. ω besitzt eine Stammfunktion).

- 2.) ω ist geschlossen (d.h. ω erfüllt die Integrationsbedingung (22.6)).

Beweis: 1.) \Rightarrow 2.): folgt aus den Überlegungen am Anfang von 22.12.

2.) \Rightarrow 1.): OBdA ist U sternförmig bzgl. $p_0 = 0$ (andernfalls ist dies durch Translation stetig zu erreichen).

Sei $\omega = \sum_{l=1}^n f_l dx_l$; für $x \in U$ definieren wir den Wert der (prospektiven) Stammfunktion

durch Integration über die Strecke von 0 nach x , d.h. mit $\gamma(t) = tx$ für $0 \leq t \leq 1$ ist

$$F(x) := \int_{\gamma} \omega = \int_0^1 \omega(tx)(x) dt = \int_0^1 \underbrace{\left(\sum_{l=1}^n f_l(tx) x_l \right)}_{g(t,x)} dt.$$

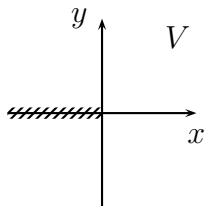
Nach Proposition 18.19 ist F als Parameterintegral stetig differenzierbar und es gilt für $j = 1, \dots, n$

$$\begin{aligned} D_j F(x) &= \int_0^1 D_{x_j} g(t, x) dt = \int_0^1 \left(\left(\sum_{l=1}^n \underbrace{D_j f_l}_{[(22.6)] = D_l f_j} (tx) \cdot tx_l \right) + f_j(tx) \right) dt \\ &= \int_0^1 \underbrace{\left(\sum_{l=1}^n D_l f_j(tx) tx_l \right)}_{\frac{d}{dt}(t f_j(tx)) - f_j(tx)} dt + \int_0^1 f_j(tx) dt = \int_0^1 \frac{d}{dt}(t \cdot f_j(tx)) dt \\ &= 1 \cdot f_j(1 \cdot x) - 0 \cdot f_j(0 \cdot x) = f_j(x), \end{aligned}$$

also ist $dF = \omega$. □

22.15. Beispiel

Wie wir oben gesehen haben, erfüllt die stetig differenzierbare 1-Form $\omega(x, y) = \frac{-y}{x^2+y^2} dx + \frac{x}{x^2+y^2} dy$ auf $U = \mathbb{R}^2 \setminus (0, 0)$ die Integrabilitätsbedingungen und kann auf diesem Gebiet aber keine Stammfunktion haben. Aber auf dem Gebiet $V := \mathbb{R}^2 \setminus \{(x, 0) : x \leq 0\}$ besitzt ω sehr wohl eine Stammfunktion:

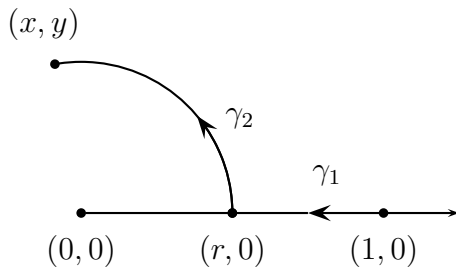


Das Gebiet V entsteht nämlich aus $\mathbb{R}^2 \setminus \{(0, 0)\}$ durch Entfernung der negativen x -Achse und ist somit für jedes $r > 0$ sternförmig bezüglich $(r, 0)$.

Daher gibt es nach dem obigen Theorem eine Stammfunktion $F: V \rightarrow \mathbb{R}$ zu ω .

Wie sieht so eine Stammfunktion konkret aus?

Dazu wählen wir den Punkt $p_0 = (1, 0) \in V$ als Startpunkt für Kurven der Form $\gamma = \gamma_1 + \gamma_2$ nach $(x, y) \in V$ in folgender Art:



mit $r := (x^2 + y^2)^{1/2} > 0$ und $-\pi < \varphi < \pi$ (eindeutig) so, dass $x = r \cos \varphi, y = r \sin \varphi$, setzen wir

- falls $r \leq 1$: $\gamma_1(t) = (r - t, 0)$ ($r - 1 \leq t \leq 0$)
 - falls $r > 1$: $\gamma_1(t) = (r + t, 0)$ ($1 - r \leq t \leq 0$)
- sowie für $(0 \leq t \leq |\varphi|)$
- falls $\varphi \geq 0$: $\gamma_2(t) = (r \cos t, r \sin t)$
 - falls $\varphi < 0$: $\gamma_2(t) = (r \cos t, -r \sin t)$.

Dann ergibt sich

$$\int_{\gamma} \omega = \underbrace{\int_{\gamma_1} \omega}_{[y=0, dy=0]} + \underbrace{\int_{\gamma_2} \omega}_{[\text{Bsp. 22.8}]} = 0 + \varphi = \varphi.$$

Wir betonen nochmals, dass für $(x, y) \in V$ der Polarwinkel φ eindeutig im Bereich $]-\pi, \pi[$ gewählt werden kann und schreiben für die dadurch vermittelte Abbildung $\varphi = \arg(x, y)$, $\arg: V \rightarrow \mathbb{R}$. Somit haben wir also die Stammfunktion

$$F(x, y) = \arg(x, y)$$

zur 1-Form ω auf dem Gebiet V erhalten.

22.16. Zur praktischen Bestimmung einer Stammfunktion

Ein Beispiel im \mathbb{R}^3 : $\omega(x, y, z) = (x + z)dx - (y + z)dy + (x - y)dz = f_1 dx + f_2 dy + f_3 dz$

Wir überprüfen die Integrabilitätsbedingungen:

$$\begin{array}{ccc} \begin{array}{c} x+z \\ \swarrow \quad \searrow \\ \partial y \quad \partial z \\ 0 \qquad \qquad 1 \end{array} & \begin{array}{c} -y-z \\ \swarrow \quad \searrow \\ \partial x \quad \partial z \\ 0 \qquad \qquad -1 \end{array} & \begin{array}{c} x-y \\ \swarrow \quad \searrow \\ \partial x \quad \partial y \\ 1 \qquad \qquad -1 \end{array} \end{array}$$

\leadsto Integrabilitätsbedingung sind erfüllt; da \mathbb{R}^3 sternförmig ist, gibt es eine Stammfunktion, d.h. $F: \mathbb{R}^3 \rightarrow \mathbb{R}$ mit

$$\begin{aligned} \partial_x F &= x + z \\ \partial_y F &= -y - z \\ \partial_z F &= x - y. \end{aligned}$$

(0) Wir versuchen $F^{[1]} := \int f_1 dx = \frac{x^2}{2} + zx$ [$\Rightarrow \partial_x F = f_1$]

und überprüfen: $\partial_y F^{[1]} = 0 \dots$ sollte aber $= f_2 = -y - z$ sein!

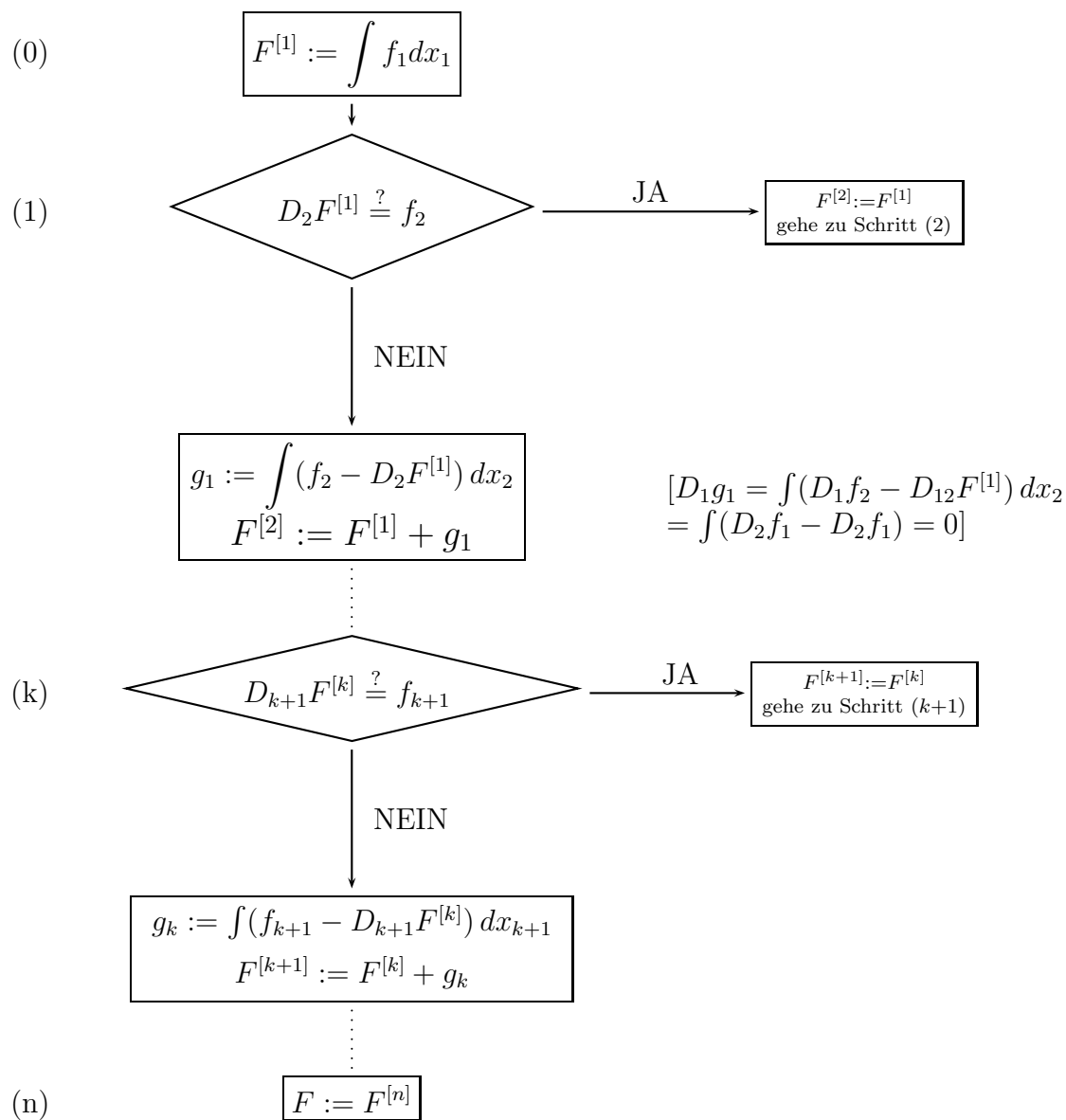
$$(1) \text{ Korrektur } F^{[2]} := F^{[1]} + \int (f_2 - \partial_y F^{[1]}) dy = \frac{x^2}{2} + zx + \int (-y - z) dy \\ = \frac{x^2}{2} + zx - \frac{y^2}{2} - zy$$

$$[\text{Bemerkung: } \partial_x \int (f_2 - \partial_y F^{[1]}) dy = \int (\underbrace{\partial_x f_2}_{\partial_y f_1} - \underbrace{\partial_{xy} F^{[1]}}_{\partial_y f_1}) dy = 0]$$

wir überprüfen: $\partial_z F^{[2]} = x - y = f_3 \dots$ fertig,

d.h. $F(x, y, z) := \frac{x^2 - y^2}{2} + z(x - y)$ ist eine Stammfunktion.

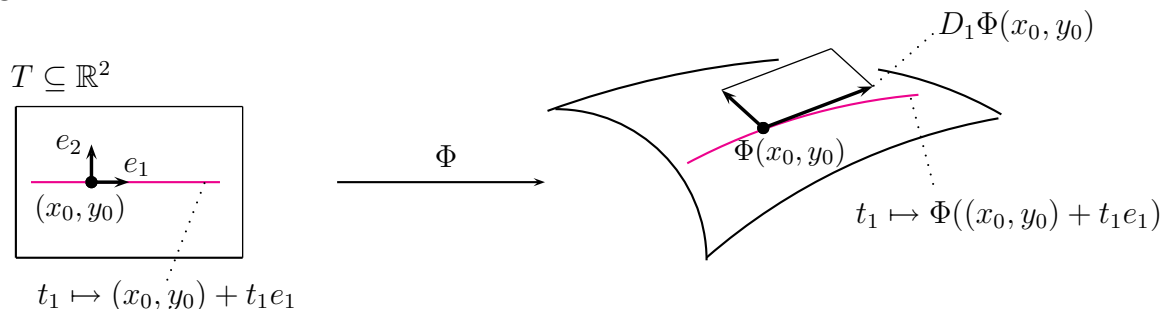
Algorithmus: sei das Gebiet $U \subseteq \mathbb{R}^n$ so beschaffen, dass entlang der Koordinatenrichtungen integriert werden darf und $\omega = f_1 dx_1 + \dots + f_n dx_n$



§23. Untermannigfaltigkeiten des \mathbb{R}^n

23.1. Parametrisierte Flächenstücke und Immersionen

Grundidee: wir stellen uns eine zweidimensionale Fläche im \mathbb{R}^3 als verbogenes Ebenenstück vor, wobei die Verbiegung durch eine stetig differenzierbare Abbildung bewerkstelligt wird:



Mit $\Phi: \mathbb{R}^2 \underset{\text{offen}}{\supseteq} T \rightarrow \mathbb{R}^3$ differenzierbar und der Notation $D_j \Phi = \begin{pmatrix} D_j \Phi_1 \\ D_j \Phi_2 \\ D_j \Phi_3 \end{pmatrix}$ liefert die (komponentenweise gebildete) Taylor-Approximation erster Ordnung eine Linearisierung nahe des Punktes $\Phi(x_0, y_0)$ in der Form

$$\underbrace{\begin{Bmatrix} (x_0, y_0) + t_1 \cdot e_1 \\ (x_0, y_0) + t_2 \cdot e_2 \end{Bmatrix}}_{\text{Basisrichtungen}} \mapsto \underbrace{\Phi(x_0, y_0) + \begin{Bmatrix} t_1 \cdot D_1 \Phi(x_0, y_0) \\ t_2 \cdot D_2 \Phi(x_0, y_0) \end{Bmatrix}}_{\text{Tangentialvektoren an die Fläche } \Phi(T)}$$

Da die Dimension der solcherart infinitesimal linearisierten Fläche gleich 2 sein soll, verlangen wir, dass die Tangentialvektoren linear unabhängig sein sollen, d.h.

$$\dim(\mathbb{R}^2) = 2 = \text{Rang} (D_1 \Phi(x_0, y_0) \ D_2 \Phi(x_0, y_0)) = \text{Rang} D\Phi(x_0, y_0).$$

Definition: Es sei $T \subseteq \mathbb{R}^k$ offen und $\Phi: T \rightarrow \mathbb{R}^n$, $(t_1, \dots, t_k) \mapsto \Phi(t_1, \dots, t_k)$ stetig differenzierbar. Φ heißt *Immersion*, falls gilt:

$$(23.1) \quad \forall t \in T: \quad \text{Rang } D\Phi(t) = k$$

[somit muss übrigens $k \leq n$ gelten]; äquivalent dazu ist

$$(23.1') \quad \forall t \in T: \quad \{D_1 \Phi(t), \dots, D_k \Phi(t)\} \text{ ist linear unabhängig in } \mathbb{R}^n.$$

In diesem Fall wird $F := \Phi(T) \subseteq \mathbb{R}^n$ als *parametrisierte k -Fläche* bezeichnet. Im Spezialfall $k = 1$ erhalten wir reguläre Kurven als 1-Flächen.

Bemerkung: 1.) Falls $k = n$ gilt, ist Φ nach dem Umkehrsatz aus §20 ein lokaler \mathcal{C}^1 -Diffeomorphismus.

2.) Gleichung (23.1) impliziert die folgende wichtige technische Bedingung:
für jedes $t_* \in T$ können wir k von den n Komponentenfunktionen in Φ auswählen, d.h. $1 \leq i_1 < i_2 < \dots < i_k \leq n$, sodass

$$(23.1'') \quad \det \frac{\partial(\Phi_{i_1}, \dots, \Phi_{i_k})}{\partial(t_1, \dots, t_k)}(t) \neq 0$$

in einer offenen Umgebung von t_* gilt [das folgt aus der Rangbedingung zusammen mit der Stetigkeit von $t \mapsto \det \frac{\partial(\Phi_{i_1}, \dots, \Phi_{i_k})}{\partial(t_1, \dots, t_k)}(t)$].

Allgemeiner gesagt gilt also für stetig differenzierbare Abbildungen $f: \mathbb{R}^k \underset{\text{offen}}{\supseteq} W \rightarrow \mathbb{R}^n$ und $t_* \in W$:

$$(23.2) \quad \text{Rang } Df(t_*) = k \implies \exists \text{ Umgebung } U \ni t_*: \text{Rang } Df|_U(t) \geq k \quad \forall t \in U.$$

Proposition: Es sei $\Phi: \mathbb{R}^k \underset{\text{offen}}{\supseteq} T \rightarrow \mathbb{R}^n$ eine Immersion. Dann gilt:

$\forall t \in T \exists V \underset{\text{offen}}{\subseteq} T$ mit $t \in V$ und der Eigenschaft, dass Φ als Abbildung $V \rightarrow \Phi(V)$ aufgefasst ein Homöomorphismus ist (d.h. Φ ist stetig und bijektiv mit stetiger Inverser $\Phi(V) \rightarrow V$ [vgl. Definition 16.7]).

Beweis: Für den Fall $k = n$ folgt die Aussage aus dem Umkehrsatz. Daher nehmen wir also $k < n$ an.

Sei $t_* \in T$, dann gibt es nach (23.1'') Indizes $1 \leq i_1 < \dots < i_k \leq n$, sodass in einer ganzen Umgebung von t_* gilt $\det \frac{\partial(\Phi_{i_1}, \dots, \Phi_{i_k})}{\partial(t_1, \dots, t_k)}(t) \neq 0$. OBdA haben wir $(i_1, \dots, i_k) = (1, \dots, k)$ (sonst nehmen wir einfach eine Umnummerierung der Koordinaten im \mathbb{R}^n vor).

Die Abbildung $G := (\Phi_1, \dots, \Phi_k): T \rightarrow \mathbb{R}^k$ ist dann ein lokaler \mathcal{C}^1 -Diffeomorphismus, daher $\exists V \underset{\text{offen}}{\subseteq} T$ mit $t_* \in V$ und $U \underset{\text{offen}}{\subseteq} \mathbb{R}^k$ mit $G(t_*) \in U$, sodass $\tilde{G} := G|_V: V \rightarrow U$ bijektiv und \mathcal{C}^1 sowie $\tilde{\Psi} := \tilde{G}^{-1}: U \rightarrow V$ ebenfalls \mathcal{C}^1 ist.

Sei weiters $H := (\Phi_{k+1}, \dots, \Phi_n)$, dann ist $\Phi|_V = (G, H): V \rightarrow U \times \mathbb{R}^{n-k}$ injektiv.

Schließlich bezeichne $\bar{\Phi}$ die Abbildung $V \rightarrow \Phi(V) \subseteq U \times \mathbb{R}^{n-k}$ mit $t \mapsto \Phi(t)$. Dann ist $\bar{\Phi}$ bijektiv mit der Umkehrabbildung $\Phi(V) \rightarrow V$ gegeben durch $(x_1, \dots, x_k, x_{k+1}, \dots, x_n) \mapsto \tilde{\Psi}(x_1, \dots, x_k)$. Letztere ist nach Konstruktion auch stetig. \square

Die Bedeutung der Proposition liegt darin, dass wir bei gegebener Immersion $\Phi: T \rightarrow \mathbb{R}^n$ stets durch Verkleinerung von T (also „lokal“) so genannte „Selbstdurchdringungen“ von k -Flächen (das sind Stellen der Nichtinjektivität von Φ , Doppelpunkte usw.) vermeiden können — dies führt auf den Begriff der k -dimensionalen Untermannigfaltigkeiten (oder auch Teilmannigfaltigkeiten) des \mathbb{R}^n : diese sind lokal stets homöomorph zu offenen Teilmengen des \mathbb{R}^k .

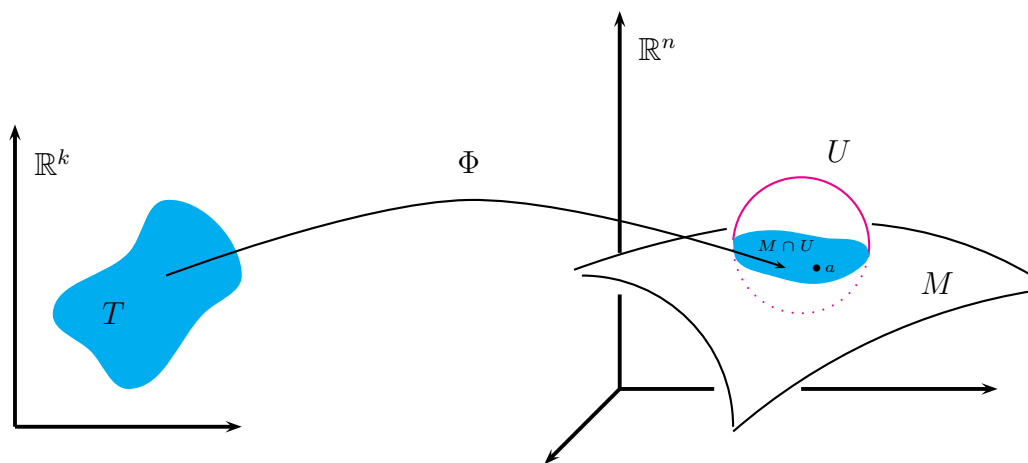
23.2. Definition

1.) $M \subseteq \mathbb{R}^n$ heißt *k-dimensionale Untermannigfaltigkeit*¹ von \mathbb{R}^n , falls gilt:

zu jedem Punkt $a \in M$ gibt es eine offene Umgebung $U \subseteq \mathbb{R}^n$ von a und $T \subseteq \mathbb{R}^k$ offen sowie eine Immersion $\Phi: T \rightarrow \mathbb{R}^n$, so dass gilt:

(a) Φ ist ein Homöomorphismus $T \rightarrow \Phi(T)$

(b) $\Phi(T) = M \cap U$.



Wir nennen $\Phi: T \rightarrow M \cap U$ eine *lokale Parametrisierung* von M nahe a ; die Inverse $\Psi := \Phi^{-1}: M \cap U \rightarrow T \subseteq \mathbb{R}^k$ heißt *Kartenabbildung* (oder nur *Karte*) um a .

Ist $M \ni p = \Phi(t_1, \dots, t_k)$, so sind t_1, \dots, t_k die *lokalen Koordinaten* von p bezüglich Φ .

Die *Codimension* von M ist $n - k$. Untermannigfaltigkeiten der Codimension 1 nennt man auch *Hyperflächen*.

2.) Es seien $M \subseteq \mathbb{R}^m$ und $N \subseteq \mathbb{R}^n$ Untermannigfaltigkeiten. Eine Abbildung $f: M \rightarrow N$ heißt *differenzierbar* (bzw. $\mathcal{C}^k, \mathcal{C}^\infty$), falls gilt:

zu jedem $p \in M$ gibt es $U \subseteq \mathbb{R}^m$ offen mit $p \in U$ und eine differenzierbare (bzw. \mathcal{C}^k -, \mathcal{C}^∞ -) Abbildung $\tilde{f}: U \rightarrow \mathbb{R}^n$ mit $\tilde{f}|_{M \cap U} = f|_{M \cap U}$.

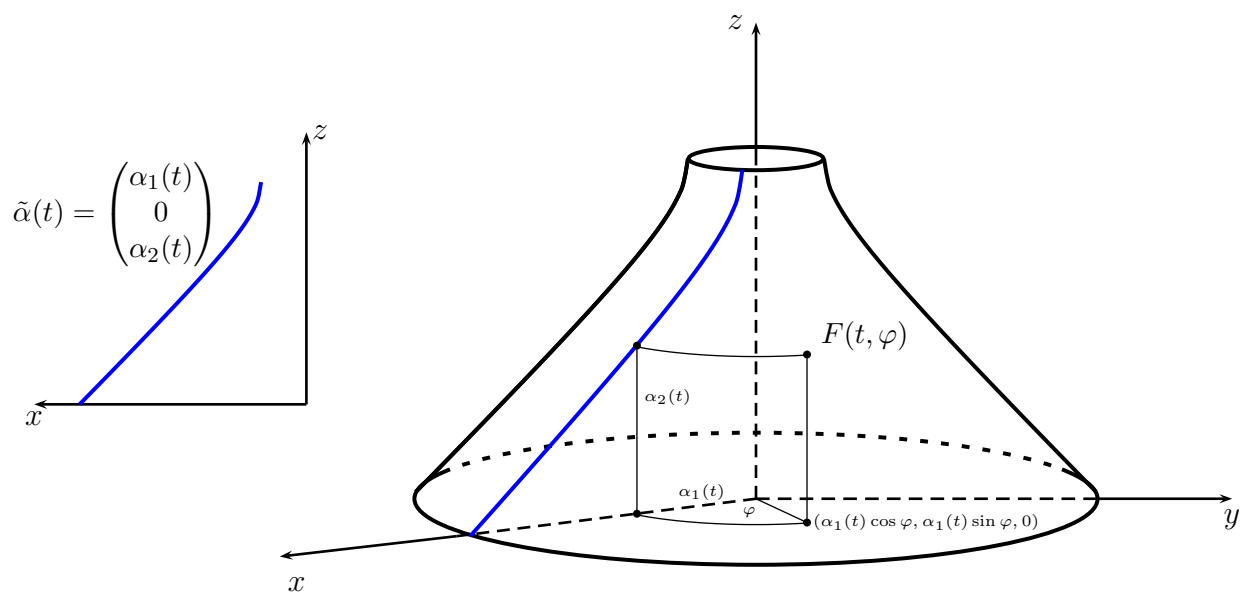
Bemerkung: Die Proposition in 23.1 besagt also folgendes: Ist $\Phi: \mathbb{R}^k \supseteq T \rightarrow \mathbb{R}^n$ eine Immersion und $t_0 \in T$, dann $\exists V \subseteq T$ offen mit $t_0 \in V$, so dass $\Phi(V)$ eine k -dimensionale Untermannigfaltigkeit von \mathbb{R}^n ist. Mit anderen Worten: lokal sind k -Flächen stets k -dimensionale Untermannigfaltigkeiten.

¹Der mathematische Begriff „Mannigfaltigkeit“ wurde 1854 von Bernhard Riemann eingeführt.

23.3. Beispiele

1.) **Rotationsflächen im \mathbb{R}^3 :** Es sei $I \subseteq \mathbb{R}$ ein offenes Intervall und $\alpha: I \rightarrow \mathbb{R}^2$, $\alpha(t) = (\alpha_1(t), \alpha_2(t))$ ein regulärer Weg, dessen Bildmenge wir als Kurve in der xz -Ebene auffassen; dann lassen wir jeden Kurvenpunkt $\tilde{\alpha}(t) := (\alpha_1(t), 0, \alpha_2(t))$ um die z -Achse rotieren und erhalten so eine 2-parametrische Abbildung

$$F: I \times \mathbb{R} \rightarrow \mathbb{R}^3 \quad \text{mit} \quad (t, \varphi) \mapsto \begin{pmatrix} \alpha_1(t) \cos \varphi \\ \alpha_1(t) \sin \varphi \\ \alpha_2(t) \end{pmatrix}.$$



Wir überprüfen zunächst, ob F eine Immersion ist: wegen

$$DF(t, \varphi) = \begin{pmatrix} \alpha_1'(t) \cos \varphi & -\alpha_1(t) \sin \varphi \\ \alpha_1'(t) \sin \varphi & \alpha_1(t) \cos \varphi \\ \alpha_2'(t) & 0 \end{pmatrix}$$

unterscheiden wir zwei Fälle:

$\alpha_1(t) = 0$: es folgt $\text{Rang } DF(t, \varphi) = 1 < 2$

$\alpha_1(t) \neq 0$: • $\alpha_1'(t) \neq 0 \Rightarrow \det \begin{pmatrix} \alpha_1'(t) \cos \varphi & -\alpha_1(t) \sin \varphi \\ \alpha_1'(t) \sin \varphi & \alpha_1(t) \cos \varphi \end{pmatrix} = \alpha_1'(t) \alpha_1(t) \neq 0$, also hat $DF(t, \varphi)$ Rang 2

• $\alpha_1'(t) = 0 \Rightarrow \alpha_2'(t) \neq 0$, weil α regulär ist!

Somit hat $DF(t, \varphi) = \begin{pmatrix} 0 & -\alpha_1(t) \sin \varphi \\ 0 & \alpha_1(t) \cos \varphi \\ \alpha_2'(t) & 0 \end{pmatrix}$ in diesem Fall auch Rang 2.

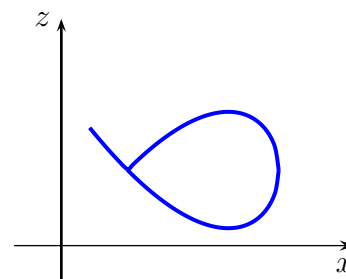
Zusammenfassend haben wir:

F ist Immersion $\iff \forall t \in I: \alpha_1(t) \neq 0$ (kein Schnittpunkt mit der z -Achse)

$\iff M := F(I \times \mathbb{R})$ ist eine 2-Fläche im \mathbb{R}^3

Bemerkungen: (i) Die Abbildung F ist wegen $F(t, \varphi + 2\pi) = F(t, \varphi)$ nicht global injektiv.

(ii) Die Teilmenge $M = F(I \times \mathbb{R}) \subseteq \mathbb{R}^3$ ist im Allgemeinen keine Untermannigfaltigkeit. Wie man sich überlegen kann, tritt dieser Fall z.B. dann ein, wenn die durch α in der xz -Ebene definierte Kurve wie in nebenstehender Skizze aussieht, wobei ein offenes Kurvenende beliebig nahe an einen festen Kurvenpunkt heranführt (kein Doppelpunkt!). Es ist übrigens dann auch das Bild von α keine 1-dimensionale Teilmannigfaltigkeit von \mathbb{R}^2 .

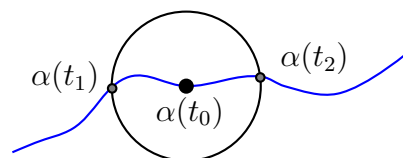


Wir formulieren nun eine zusätzliche Bedingung an α , die garantieren soll, dass die Rotationsfläche M auch eine 2-dimensionale Untermannigfaltigkeit des \mathbb{R}^3 ist. Die topologische Bedeutung ist, dass es um jeden Kurvenpunkt beliebig kleine Umgebungen (im \mathbb{R}^2) geben soll, innerhalb derer die Kurve aus einem zusammenhängenden Stück besteht:

(*) $\forall t_0 \in I \exists \delta_1 > 0 \forall \delta > 0$ mit $0 < \delta < \delta_1$

$\exists t_1, t_2 \in I, t_1 < t_0 < t_2 :$

$$B_\delta(\alpha(t_0)) \cap \alpha(I) = \alpha([t_1, t_2]).$$



Behauptung: Ist zusätzlich zu den obigen Annahmen auch (*) erfüllt, dann ist $M = F(I \times \mathbb{R})$ eine zweidimensionale Untermannigfaltigkeit von \mathbb{R}^3 (also eine Hyperfläche).

Beweis: Sei $p \in M$. Wegen der Rotationsinvarianz dürfen wir OBdA annehmen, dass p in der xz -Ebene liegt; daher gibt es ein $t_0 \in I$, so dass $p = (\alpha_1(t_0), 0, \alpha_2(t_0))$; d.h. $p = F(t_0, 0)$.

Aus Proposition 23.1 folgt, dass F lokal ein Homöomorphismus $T_0 \rightarrow F(T_0)$ ist, wobei T_0 eine geeignete Umgebung von $(t_0, 0)$ ist. Wir wählen $\delta_1 > 0$ gemäß (*). Weiters sei $0 < \delta < \delta_1$ sowie $\varepsilon > 0$ so klein gewählt, dass mit $t_1, t_2 \in I$ wie in (*) auch noch $[t_1, t_2] \times [-\varepsilon, \varepsilon] \subseteq T_0$ gilt.

Wir setzen $T :=]t_1, t_2[\times]-\varepsilon, \varepsilon[$. Dann ist $\Phi: T \rightarrow F(T), (t, \varphi) \mapsto F(t, \varphi)$ ein Homöomorphismus. Außerdem ist

$$U := \{(x_1 \cos \varphi, x_1 \sin \varphi, x_2) : (x_1, x_2) \in B_\delta(\alpha(t_0)), |\varphi| < \varepsilon\}$$

eine offene Umgebung von p (bestehend aus allen Rotationen um Winkel zwischen $-\varepsilon$ und ε der δ -Scheibe um den Punkt p in der xz -Ebene). Nach Konstruktion hat U die Eigenschaft, dass $\Phi(T) = F(T) = M \cap U$ ist. Somit ist Definition 23.2 erfüllt. \square

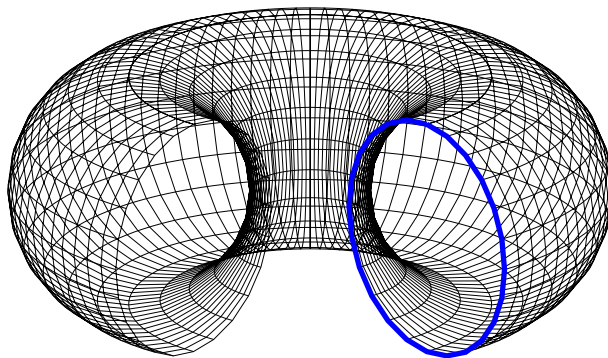
Bemerkung: Die Bedingung (*) bedeutet gerade, dass $\alpha(I)$ eine eindimensionale Untermannigfaltigkeit des \mathbb{R}^2 ist. (Dies folgt aus einem allgemeineren Theorem über 'reguläre Teilmannigfaltigkeiten'; siehe z.B. VO *Differentialgeometrie* von Kollegen Michael Kunzinger, Universität Wien, Sommersemester 2006.)

2.) Der *Torus*² ergibt sich als Rotationsfläche wie in 1.) bezüglich eines Kreises:

mit $0 < r < R$ sei $\alpha: \mathbb{R} \rightarrow \mathbb{R}^2, t \mapsto \begin{pmatrix} x(t) \\ z(t) \end{pmatrix} = \begin{pmatrix} R + r \cos t \\ r \sin t \end{pmatrix}$. Dies erfüllt alle Bedingungen aus 1.) (inklusive (*)) und die Parametrisierung $F: \mathbb{R}^2 \rightarrow \mathbb{R}^3$ lautet explizit

$$F(t, s) = \begin{pmatrix} (R + r \cos t) \cdot \cos(s) \\ (R + r \cos t) \cdot \sin(s) \\ r \sin t \end{pmatrix}.$$

Der Torus $M_{r,R} = F(\mathbb{R}^2)$ ist somit eine zweidimensionale Untermannigfaltigkeit des \mathbb{R}^3 , also eine Hyperfläche.



In vielen konkreten Fragen über Untermannigfaltigkeiten ist es nicht praktikabel, direkt auf die definierenden Eigenschaften in 22.2 zurückzugehen. Deshalb geben wir im folgenden eine Reihe von äquivalenten Bedingungen.

²Das lateinische Wort *torus* bezeichnete ein Kissen in der Form eines Kranzes.

23.4. Theorem

Für eine Teilmenge $M \subseteq \mathbb{R}^n$ sind folgende Aussagen äquivalent:

- 1.) M ist eine k -dimensionale Untermannigfaltigkeit.
- 2.) *Lokale Darstellung als Graph*: $\forall a = (a_1, \dots, a_n) \in M$ gilt: zu $a' := (a_1, \dots, a_k)$ und $a'' := (a_{k+1}, \dots, a_n)$, gibt es (eventuell nach Umnummerierung der Koordinaten) offene Mengen $U' \subseteq \mathbb{R}^k$, $U'' \subseteq \mathbb{R}^{n-k}$ mit $a' \in U'$, $a'' \in U''$ und eine stetig differenzierbare Abbildung $g: U' \rightarrow U''$ mit der Eigenschaft, dass

$$M \cap (U' \times U'') = \{(x', g(x')) \in \mathbb{R}^n : x' \in U'\} = \text{Graph von } g$$

(mit anderen Worten: lokal, d.h. innerhalb der Umgebung $U' \times U''$ von a , ist $x \in M$ gleichwertig mit der Relation $x'' = g(x')$).

- 3.) *Lokale Beschreibung als Nullstellenmenge*: $\forall a \in M$ gibt es eine offene Menge $U \subseteq \mathbb{R}^n$ mit $a \in U$ und stetig differenzierbare Funktionen $f_1, \dots, f_{n-k}: U \rightarrow \mathbb{R}$ mit der Eigenschaft, dass

$$\text{Rang } \frac{\partial(f_1, \dots, f_{n-k})}{\partial(x_1, \dots, x_n)}(x) = n - k \quad \forall x \in M \cap U$$

und

$$M \cap U = \{x \in U : f_1(x) = \dots = f_{n-k}(x) = 0\}.$$

- 4.) *Lokal diffeomorph zu einem k -dimensionalen Untervektorraum des \mathbb{R}^n* : Es bezeichne $E_k := \{(x_1, \dots, x_k, 0, \dots, 0) : x_j \in \mathbb{R} \ (j = 1, \dots, k)\} \subseteq \mathbb{R}^n$; $\forall a \in M$ gibt es eine offene Menge $U \subseteq \mathbb{R}^n$ mit $a \in U$ und $V \subseteq \mathbb{R}^n$ offen sowie einen \mathcal{C}^1 -Diffeomorphismus $F: U \rightarrow V$ mit der Eigenschaft

$$F(M \cap U) = E_k \cap V.$$

Beweis: Wir zeigen 1.) \Rightarrow 2.) \Rightarrow 3.) \Rightarrow 4.) \Rightarrow 1.)

- 1.) \Rightarrow 2.): Sei $a \in M$. Laut Definition 23.2 gibt es eine offene Menge $U \subseteq \mathbb{R}^n$ mit $a \in U$ und $T \subseteq \mathbb{R}^k$ offen sowie eine Immersion $\Phi: T \rightarrow \mathbb{R}^n$, so dass $T \rightarrow \Phi(T) = M \cap U$ ein Homöomorphismus ist.

Sei $t_* \in T$ mit $\Phi(t_*) = a$. Wir können OBdA annehmen, dass $\det \frac{\partial(\Phi_1, \dots, \Phi_k)}{\partial(t_1, \dots, t_k)}(t_*) \neq 0$ gilt. Dann ist $\tilde{\Phi} := (\Phi_1, \dots, \Phi_k): T \rightarrow \mathbb{R}^k$ ein lokaler \mathcal{C}^1 -Diffeomorphismus, d. h. es gibt $T_1 \subseteq T$ mit $t_* \in T_1$ und $U_1 \subseteq \mathbb{R}^k$, so dass $\tilde{\Phi}: T_1 \rightarrow U_1$ ein \mathcal{C}^1 -Diffeomorphismus mit Umkehrabbildung $\Psi: U_1 \rightarrow T_1$ ist.

Wir setzen $G := \Phi \circ \Psi: U_1 \rightarrow \mathbb{R}^n$ und erkennen $G(\underbrace{x_1, \dots, x_k}_{=: x'}) =$

$$= \left(\underbrace{\Phi_1(\Psi(x')), \dots, \Phi_k(\Psi(x'))}_{\tilde{\Phi}(\Psi(x'))=x'}, \underbrace{\Phi_{k+1}(\Psi(x'))}_{=: g_{k+1}(x')}, \dots, \underbrace{\Phi_n(\Psi(x'))}_{=: g_n(x')} \right) = (x', g(x')),$$

wobei $g := (g_{k+1}, \dots, g_n): U_1 \rightarrow \mathbb{R}^{n-k}$ stetig differenzierbar ist. Insbesondere gilt für jede Teilmenge $Y \subseteq U_1$ stets $\text{Graph}(g|_Y) = G(Y) = \Phi(\Psi(Y)) \subseteq M \cap U$.

Sei $U'' \subseteq \mathbb{R}^{n-k}$ offen mit $a'' \in U''$, so dass $U_1 \times U'' \subseteq U$ ist. Wähle weiters $U' \subseteq U_1$ offen mit $a' \in U'$ und $g(U') \subseteq U''$.

Dann gilt $\text{Graph}(g|_{U'}) = G(U') \subseteq U' \times g(U') \subseteq U' \times U''$ sowie $\text{Graph}(g|_{U'}) = G(U') = \Phi(\Psi(U')) \subseteq M \cap U$. Somit ist einerseits

$$\text{Graph}(g|_{U'}) \subseteq M \cap (U' \times U'').$$

Andererseits folgt für $(x', x'') \in M \cap (U' \times U'') \subseteq \Phi(T)$, dass $x' \in U'$ und $(x', x'') = \Phi(\tilde{t})$ mit $\tilde{t} \in T$ gelten muss. Also muss $\tilde{t} \in \Psi(U')$ sein und somit $(x', x'') \in \Phi(\Psi(U')) = G(U') = \text{Graph}(g|_{U'})$; d.h. zusammenfassend

$$\text{Graph}(g|_{U'}) = M \cap (U' \times U'').$$

2.) \Rightarrow 3.): Sei $j = 1, \dots, n-k$; wir setzen

$$f_j(x_1, \dots, x_n) := x_{k+j} - g_{k+j}(x_1, \dots, x_k)$$

für $x \in U' \times U''$ mit $x' = (x_1, \dots, x_k) \in U'$, $x'' = (x_{k+1}, \dots, x_n) \in U''$.

Wegen $\frac{\partial(f_1, \dots, f_{n-k})}{\partial(x_{k+1}, \dots, x_n)} = I_{n-k}$ hat $\frac{\partial(f_1, \dots, f_{n-k})}{\partial(x_1, \dots, x_n)}$ Rang gleich $(n-k)$; weiters folgt direkt

$$M \cap (U' \times U'') = \{(x', g(x')) : x' \in U'\} = \{(x', x'') : f_j(x', x'') = 0 \ (j = 1, \dots, n-k)\}.$$

3.) \Rightarrow 4.): OBdA (d.h. nach eventueller Umnummerierung) können wir annehmen, dass $\det \frac{\partial(f_1, \dots, f_{n-k})}{\partial(x_{k+1}, \dots, x_n)}(a) \neq 0$ gilt.

Wir definieren $F: U \rightarrow \mathbb{R}^n$ durch $F(x_1, \dots, x_n) := (x_1, \dots, x_k, f_1(x), \dots, f_{n-k}(x))$. Dann hat

$$DF = \begin{pmatrix} I_k & 0 \\ * & \frac{\partial(f_1, \dots, f_{n-k})}{\partial(x_{k+1}, \dots, x_n)} \end{pmatrix}$$

nahe a Rang gleich n , ist also invertierbar. Daher ist F ein lokaler \mathcal{C}^1 -Diffeomorphismus, d.h. ein \mathcal{C}^1 -Diffeomorphismus offener Mengen $W \subseteq U$ und $V \subseteq \mathbb{R}^n$ mit $a \in W$.

Für $x \in W$ gilt: $f_1(x) = \dots = f_{n-k}(x) = 0$ ist gleichwertig damit, dass für $y = F(x)$ gilt $y_{k+1} = \dots = y_n = 0$. Daher ist $F(M \cap W) = E_k \cap V$.

4.) \Rightarrow 1.): Laut Voraussetzung ist $F^{-1}: V \rightarrow U$ ein \mathcal{C}^1 -Diffeomorphismus; es bezeichne $p: \mathbb{R}^n \rightarrow \mathbb{R}^k$ die Projektion $(x_1, \dots, x_n) \mapsto (x_1, \dots, x_k)$.

Wir setzen $T := p(V) = p(E_k \cap V) \subseteq \mathbb{R}^k$.

T ist offen: mit der Notation $B_\varepsilon^X(z)$ für die ε -Kugel um z in einem metrischen Raum X können wir die Umgebungseigenschaft von T für jeden Punkt in T auf die Relation $B_\varepsilon^{\mathbb{R}^n}(y) \cap E_k = B_\varepsilon^{\mathbb{R}^k}(p(y)) \times \{0\}$ zurückspielen.

Sei $\Phi: T \rightarrow U$ definiert durch $\Phi(t_1, \dots, t_k) := F^{-1}(t_1, \dots, t_k, 0, \dots, 0)$. Nach Konstruktion ist Φ stetig differenzierbar und eine Immersion. Weiters ist

$$\Phi(T) = F^{-1}(T \times \{0\}) = F^{-1}(E_k \cap V) = M \cap U$$

und es gibt eine stetige Inverse vermöge $M \cap U \xrightarrow{F} V \xrightarrow{p} T$, d.h. Φ ist ein Homöomorphismus auf sein Bild.

□

23.5. Anwendung

Die Funktion $f: \mathbb{R}^n \rightarrow \mathbb{R}$, $f(x) = \|x\|^2 - 1$ ist stetig differenzierbar und

$\text{Rang} \frac{\partial f}{\partial(x_1, \dots, x_n)} = \text{Rang} \text{grad} f(x) = \text{Rang} 2x = 1$ für alle $x \in \mathbb{R}^n$ mit $x \neq 0$.

Somit folgt aus dem obigen Theorem, dass

$$S^{n-1} = \{x \in \mathbb{R}^n: f(x) = 0\}$$

eine Hyperfläche ist, also eine $(n - 1)$ -dimensionale Untermannigfaltigkeit von \mathbb{R}^n .

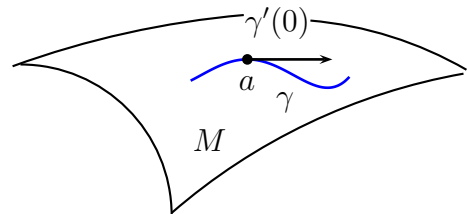
Ein wenig allgemeiner können wir mit dieser Methode eine Fülle von interessanten Untermannigfaltigkeiten als solche erkennen bzw. generieren.

Korollar: Es sei $f: \mathbb{R}^n \supseteq_{\text{offen}} U \rightarrow \mathbb{R}$ stetig differenzierbar und $c \in \mathbb{R}$. Es bezeichne $N_f(c) = f^{-1}(c) \subseteq \mathbb{R}^n$ die Niveaumenge für den Wert c . Falls $\text{grad} f(x) \neq 0$ für alle $x \in N_f(c)$, dann heißt c ein *regulärer Wert* von f und $N_f(c)$ ist eine $(n - 1)$ -dimensionale Untermannigfaltigkeit (also Hyperfläche) im \mathbb{R}^n .

Kurz gesagt: *Niveaumengen zu regulären Werten sind Hyperflächen.*

23.6. Definition

Sei M eine Untermannigfaltigkeit von \mathbb{R}^n und $a \in M$. Ein Vektor $v \in \mathbb{R}^n$ heißt *Tangentialvektor* an M im Punkt a , wenn gilt: es gibt einen stetig differenzierbaren Weg (bzw. Kurve) $\gamma:]-\varepsilon, \varepsilon[\rightarrow M \subseteq \mathbb{R}^n$ mit $\gamma(0) = a$ und $\dot{\gamma}(0) = v$.



$$T_a(M) := \{v \in \mathbb{R}^n : v \text{ ist Tangentialvektor an } M \text{ in } a\} \subseteq \mathbb{R}^n$$

heißt *Tangentialraum* in a an M und

$$N_a(M) := \{w \in \mathbb{R}^n : \forall v \in T_a(M) \text{ ist } \langle w | v \rangle = 0\} \subseteq \mathbb{R}^n$$

heißt *Normalraum* in a an M . [In Kurzschreibweise: $N_a(M) = T_a(M)^\perp$.]

Die Tangential- und Normalräume an alle Punkte einer Untermannigfaltigkeit können formal zusammengefasst werden (man beachte, dass es sich für verschiedene Punkte um verschiedene Mengen handelt): wir nennen

$$T(M) := \bigcup_{a \in M} \{a\} \times T_a(M)$$

das *Tangentialbündel* an M und

$$N(M) := \bigcup_{a \in M} \{a\} \times N_a(M)$$

das *Normalenbündel* an M .

23.7. Proposition

Es sei M eine k -dimensionale Untermannigfaltigkeit von \mathbb{R}^n und $a \in M$.

- 1.) $T_a(M)$ ist ein k -dimensionaler Vektorraum über \mathbb{R} . Für jede lokale Parametrisierung $\Phi: T \rightarrow M$ nahe a mit $\Phi(t_*) = a$ ergibt

$$\{D_1\Phi(t_*), \dots, D_k\Phi(t_*)\}$$

eine Basis von $T_a(M)$.

- 2.) $N_a(M)$ ist ein $(n - k)$ -dimensionaler Vektorraum über \mathbb{R} . Wenn M nahe a durch Gleichungen beschrieben ist wie in 23.4,3.),

$$\text{d. h. } M \cap U = \{x \in U : f_1(x) = \dots = f_{n-k}(x) = 0\},$$

wobei $a \in U \subseteq \mathbb{R}^n$ und $\text{Rang } \frac{\partial(f_1, \dots, f_{n-k})}{\partial(x_1, \dots, x_n)}(a) = n - k$, dann ist

$$\{\text{grad } f_1(a), \dots, \text{grad } f_{n-k}(a)\}$$

eine Basis von $N_a(M)$.

Beweis: Es seien $V := \text{span} \{D_j \Phi(t_*) : j = 1, \dots, k\}$ und $W := \text{span} \{\text{grad } f_j(a) : j = 1, \dots, n - k\}$ (die linearen Erzeugnisse der angegebenen Mengen von Vektoren im \mathbb{R}^n).

Da Φ eine Immersion ist, gilt $\dim V = k$. Die Rangbedingung in 2.) wiederum ergibt $\dim W = n - k$.

Behauptung 1: $V \subseteq T_a(M)$

Sei $v \in V$, d.h. $v = \sum_{j=1}^k c_j D_j \Phi(t_*)$ mit passenden Skalaren $c_j \in \mathbb{R}$. Wir wählen ε klein genug, so dass für $|s| < \varepsilon$ stets $t_* + s \cdot (c_1, \dots, c_k)$ in T liegt. Sodann definieren wir den Weg $\gamma:]-\varepsilon, \varepsilon[\rightarrow M \subseteq \mathbb{R}^n$ durch

$$\gamma(s) = \Phi(t_* + s \cdot (c_1, \dots, c_k)).$$

Dann ist $\gamma(0) = \Phi(t_*) = a$ und nach der Kettenregel

$$\gamma'(0) = \sum_{j=1}^k c_j D_j \Phi(t_* + 0 \cdot (c_1, \dots)) = v,$$

daher $v \in T_a(M)$.

Behauptung 2: $W \subseteq N_a(M)$

$N_a(M)$ ist als orthogonales Komplement einer Teilmenge von \mathbb{R}^n selbst ein Vektorraum. Es genügt also zu zeigen, dass gilt

$$\langle \text{grad } f_j(a) \mid v \rangle = 0 \quad \forall v \in T_a(M).$$

Sei $\gamma:]-\varepsilon, \varepsilon[\rightarrow M$ ein \mathcal{C}^1 -Weg mit $\gamma'(0) = v \in T_a(M)$ und $\gamma(0) = a$.

Nachdem $\gamma(s) \in M$ ist $\forall s \in]-\varepsilon, \varepsilon[$ gilt auch $f_j(\gamma(s)) = 0$. Daher folgt mittels Differentiation nach s aus der Kettenregel

$$0 = \sum_{l=1}^n D_l f_j(\gamma(0)) \cdot \gamma'_l(0) = \langle \text{grad } f_j(a) \mid v \rangle.$$

Aus den Behauptungen 1, 2 folgt nun einerseits

$$n - k \leq \dim N_a(M) \leq \dim V^\perp = n - k \text{ und somit } N_a(M) = W$$

sowie

$$V \subseteq T_a(M) \subseteq N_a(M)^\perp, \text{ wobei } \dim V = k = \dim N_a(M)^\perp \text{ und somit } V = T_a(M) \text{ gilt. } \quad \square$$

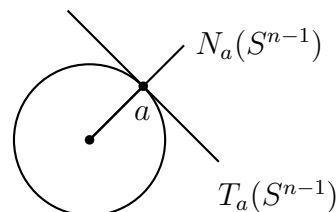
23.8. Beispiel

$S^{n-1} = \{x \in \mathbb{R}^n : f(x) := \|x\|^2 - 1 = 0\}$, daher gilt für jeden Punkt $a \in S^{n-1}$

$$N_a(S^{n-1}) = \text{span}\{\text{grad } f(a)\} = \text{span}\{2 \cdot a\} = \text{span}\{a\}.$$

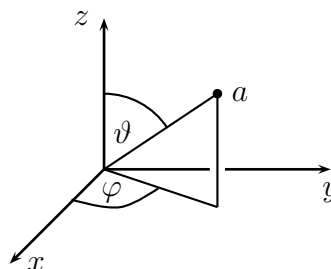
Somit folgt auch

$$T_a(S^{n-1}) = N_a(S^{n-1})^\perp = \{a\}^\perp.$$



Im Spezialfall $n = 3$ vergleichen wir dies mit der Bestimmung des Tangentialraumes $T_a(S^2)$ vermöge einer lokalen Parametrisierung durch Polarkoordinaten

$$a = \Phi(\varphi, \vartheta) = \begin{pmatrix} \cos \varphi \cdot \sin \vartheta \\ \sin \varphi \cdot \sin \vartheta \\ \cos \vartheta \end{pmatrix}.$$



Dann ist

$$D_1\Phi(\varphi, \vartheta) = \begin{pmatrix} -\sin \varphi \sin \vartheta \\ \cos \varphi \sin \vartheta \\ 0 \end{pmatrix} = \sin \vartheta \begin{pmatrix} -\sin \varphi \\ \cos \varphi \\ 0 \end{pmatrix} \perp a$$

und

$$D_2\Phi(\varphi, \vartheta) = \begin{pmatrix} \cos \varphi \cos \vartheta \\ \sin \varphi \cos \vartheta \\ -\sin \vartheta \end{pmatrix} = \cos \vartheta \begin{pmatrix} \cos \varphi \\ \sin \varphi \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ -\sin \vartheta \end{pmatrix} \perp a.$$

23.9. Nachtrag des Beweises von Theorem 19.13

Wir wiederholen hier nochmals die entscheidende Behauptung in 19.13 bezüglich der Methode der Lagrange-Multiplikatoren.

Theorem: Es sei $f: \mathbb{R}^n \supseteq_{\text{offen}} U \rightarrow \mathbb{R}$ stetig differenzierbar. Weiters seien $r \leq n$ und $g_1, \dots, g_r: U \rightarrow \mathbb{R}$ stetig differenzierbare Funktionen mit der Eigenschaft, dass der Rang der Jacobi-Matrix $D(g_1, \dots, g_r)(x)$ gleich r (also maximal) ist für alle $x \in M$, wobei

$$M := \{x \in U : \underbrace{g_1(x) = \dots = g_r(x) = 0}_{r \text{ Nebenbedingungen}}\}.$$

Falls $f|_M$ in $a \in M$ ein lokales Maximum oder Minimum besitzt, dann existieren r reelle Zahlen $\lambda_1, \dots, \lambda_r \in \mathbb{R}$, die so genannten *Lagrange-Multiplikatoren*, sodass die folgende Gleichung gilt

$$\text{grad } f(a) = \sum_{j=1}^r \lambda_j \cdot \text{grad } g_j(a).$$

Beweis: Wir wissen mittlerweile aus 23.4, dass die oben definierte Teilmenge $M \subseteq \mathbb{R}^n$ eine $(n - r)$ -dimensionale Untermannigfaltigkeit ist.

Es sei $\gamma:] - \varepsilon, \varepsilon[\rightarrow M \subseteq \mathbb{R}^n$ ein beliebiger \mathcal{C}^1 -Weg mit $\gamma(0) = a$. Dann hat die Funktion $h:] - \varepsilon, \varepsilon[\rightarrow \mathbb{R}$, definiert durch $h(t) := f(\gamma(t))$, laut Annahme bei $t = 0$ ein lokales Extremum. Daher gilt nach der Kettenregel

$$0 = \frac{d}{dt} (f(\gamma(t))) \Big|_{t=0} = \langle \text{grad } f(\gamma(0)) \mid \dot{\gamma}(0) \rangle.$$

Wenn wir alle möglichen \mathcal{C}^1 -Wege mit $\gamma(0) = a$ einsetzen, so durchläuft $\dot{\gamma}(0)$ den ganzen Tangentialraum $T_a(M)$. Somit gilt $\text{grad } f(a) \in N_a(M)$.

Nach 23.7 ist $N_a(M) = \text{span}\{\text{grad } g_j(a) : j = 1, \dots, r\}$, daher folgt die oben behauptete Darstellung für $\text{grad } f(a)$. \square

Literaturverzeichnis

- [AE99] H. Amann and J. Escher. *Analysis II*. Birkhäuser, Basel, 1999.
- [AE02] H. Amann and J. Escher. *Analysis I*. Birkhäuser, Basel, 2002. 2. Auflage.
- [BF96] M. Barner and F. Flohr. *Analysis II*. Walter de Gruyter, Berlin, 1996. 3. Auflage.
- [BF00] M. Barner and F. Flohr. *Analysis I*. Walter de Gruyter, Berlin, 2000. 5. Auflage.
- [Fis03] G. Fischer. *Lineare Algebra*. Vieweg Verlag, Wiesbaden, 2003. 14. Auflage.
- [For05] O. Forster. *Analysis 2*. Vieweg Verlag, Wiesbaden, 2005. 6. Auflage.
- [For06] O. Forster. *Analysis 1*. Vieweg Verlag, Wiesbaden, 2006. 8. Auflage.
- [Heu04] H. Heuser. *Analysis 2*. BG Teubner, Stuttgart, 2004. 13. Auflage.
- [RS02] R. Remmert and G. Schumacher. *Funktionentheorie 1*. Springer, Berlin, Heidelberg, 2002. 5. Auflage.
- [SS03] E. Stein and R. Shakarchi. *Complex Analysis*. Princeton Lectures in Analysis II. Princeton University Press, Princeton, 2003.
- [vQ01] B. von Querenburg. *Mengentheoretische Topologie*. Springer-Verlag, Berlin, 2001. 4. Auflage.