

Analysis II, WS 2010/11

30. März 2011

Die Nummerierung “Zu ..in Kap. -3 bis 0 bezieht sich auf das AnaSkript von Robert Haller-Dintelmann. Das Skript Mathematik II für ET von Steffen Roch gab die Grundlage für die weiteren Kapitel. Vieles wurde direkt übernommen, insbesondere Beispiele und Figuren. Die meisten Figuren befinden sich im separaten file fig.pdf. Kap. 1,2, 4 sind an das Skript anaII-mehrdim von Prof. Alber angelehnt. Kap. 3, 10, 12-13 wurden aus früheren Skripten zur Linearen Algebra übernommen und auf das für Analysis II Benötigte angepasst. Kap. 15, 16, 17.5 und 20.6 folgen dem Skript AnaII-ss08 von Prof. Hieber.

Die Modulprüfung Analysis II im WS2010/11 bezieht sich auf die im vorliegenden Skript angesprochenen Inhalte. Integration in einer Variablen ist Voraussetzung für die Integration in mehreren Variablen.

-3 Sprungstetige Funktionen und Regelintegrale

Literatur: Königsberger, Analysis I, Kap.11

-3.1 Sprungstetige Funktionen und Regelintegrale

-3.2 Motivation

- Einschließung (Darboux) bzw. Approximation (Riemann) der Fläche F unter dem Graphen von $f : [a, b] \rightarrow \mathbb{R}$, $f \geq 0$ durch Summen von Rechteck-Flächen
- Begriffe in Naturwissenschaft und Technik, die sich aus Abstraktion von Summationen von Produkten wie z.B. Arbeit = Kraft mal Weg ergeben
- Bestimmung der Funktion F (z.B. Weg als Funktion der Zeit t) aus Ableitung F' (Geschwindigkeit) und Anfangswert $F(t_0)$.

-3.3 Zerlegungen und Treppenfunktionen

Zu 30.1-2 Zu einer nichtleeren endlichen Teilmenge Z von \mathbb{R} gibt es genau ein n und eine Liste $x_0 < x_1 < \dots < x_n$ von Elementen von \mathbb{R} so, dass $Z = \{x_0, x_1, \dots, x_n\}$. Z ist *Zerlegung* des Intervalls $I = [a, b]$, falls $a, b \in Z \subseteq I$. Nach der vorangegangenen Bemerkung können wir Z genauso gut als Liste $a = x_0 < x_1 < \dots < x_n = b$ sehen. *Verfeinerung*: $Z' \supseteq Z$. *Treppenfunktion* passend zu Z : $f : I \rightarrow \mathbb{K}$, konstant auf (x_{j-1}, x_j) .

-3.4 Satz von Bolzano Weierstraß

Zu 30.10-13. Wir erinnern an die Begriffe ‘einseitiger Grenzwert’ (Kap.17) und ‘gleichmäßige Konvergenz’ (Kap.21). Wir betrachten Funktionen $f : I = [a, b] \rightarrow \mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$.

- (Cauchy.) f hat an $a < x_0 \leq b$ einen linksseitigen Limes, falls es zu jedem $\varepsilon > 0$ ein $\delta > 0$ gibt so, dass $|f(x) - f(x')| < \varepsilon$ für alle $x, x' \in (x_0 - \delta, x_0) \cap I$

Beweis: Für jede Folge $x_n \in [a, x_0)$ mit $x_n \rightarrow x_0$ ist $f(x_n)$ eine Cauchyfolge, also konvergent. Wie in Satz 17.8 folgt, dass f an x_0 einen linksseitigen Grenzwert besitzt (nämlich den gemeinsamen Limes dieser Cauchyfolgen). \square Dass Cauchyfolgen konvergieren folgt aus

Satz -3.1 *Jede beschränkte reelle Zahlenfolge hat eine konvergente Teilfolge*

Beweis. $x_n \in [a, b]$ für alle n . Setze $a_1 = a$, $b_1 = b$ und $x_{n_1} = x_1$. Sei $c = \frac{1}{2}(a + b)$. Dann $x_n \in [a, c]$ für unendlich viele n oder $x_n \in [c, b]$ für unendlich viele n (oder beides). Tritt das Erste ein, so setze $a_2 = a_1$ und $b_2 = c$ andernfalls $a_2 = c$ und $b_2 = b_1$. Betrachte die Teilfolge der $x_n \in [a_2, b_2]$ und wähle x_{n_2} als das erste Glied dieser Teilfolge. Iteriere diesen Schritt ad infinitum. Das ergibt $x_{n_k} \in [a_k, b_k]$ mit $a_k \leq a_{k+1} \leq b_{k+1} \leq b_k$ und $b_k - a_k \rightarrow 0$, also eine Intervallschachtelung. Sei c die eindeutig bestimmte Zahl mit $c \in [a_k, b_k]$ für alle k , Dann offensichtlich $x_{n_k} \rightarrow c$ für $k \rightarrow \infty$. \square

-3.5 Sprungstetige Funktionen

f ist *sprungstetig* oder *Regelfunktion*, wenn zu jedem $x_0 \in I$ der links- bzw. rechtsseitige Grenzwert existiert sofern nur $x_0 > a$ bzw. $x_0 < b$.

Ist f beschränkt, so ist

$$\|f\| = \sup\{|f(x)| \mid x \in I\} < \infty$$

die *Supremumsnorm* von f - schreibt man $\|f\| < \varepsilon$ so ist die Beschränktheit von f automatisch mit gemeint. Es gilt die Dreiecksungleichung

$$\|f + g\| \leq \|f\| + \|g\|$$

- Die Funktionenfolge f_n konvergiert gleichmäßig gegen f , wenn es zu jedem $\varepsilon > 0$ ein n_0 gibt so, dass $\|f - f_n\| < \varepsilon$ für alle $n \geq n_0$.

Eine *Einschließung* von f wird gegeben durch Treppenfunktionen \underline{f} und \overline{f} so, dass

$$\underline{f} \leq f \leq \overline{f}$$

Es ist eine ε -*Einschließung*, falls $\|\overline{f} - \underline{f}\| < \varepsilon$. Achtung: \overline{f} meint nicht die komplex konjugierte Funktion.

Satz -3.2 vom steten Sprung. *Gegeben sei $f : I \rightarrow \mathbb{K}$. Die folgenden Aussagen sind äquivalent*

- (1) f ist sprungstetig

- (2) Es gibt eine Folge f_n von Treppenfunktionen, die gleichmässig gegen f konvergiert,
- (2') Zu jedem $\varepsilon > 0$ gibt es eine Treppenfunktion ϕ mit $\|f - \phi\| < \varepsilon$
- (3) Es gibt eine Folge $\underline{f}_n, \bar{f}_n$ von Einschließungen von f so, dass $\bar{f}_n - \underline{f}_n$ gleichmässig gegen 0 konvergiert
- (3') Zu jedem $\varepsilon > 0$ gibt es eine Einschließung $\underline{\phi}, \bar{\phi}$ von f mit $\|\underline{\phi} - \bar{\phi}\| < \varepsilon$
- (4) Es gibt eine Folge sprungstetiger Funktionen, die gleichmässig gegen f konvergiert.

In (3) konvergieren \underline{f}_n und \bar{f}_n gleichmässig gegen f . Auch kann man $\underline{f}_n \leq \underline{f}_{n+1}$ und $\bar{f}_{n+1} \leq \bar{f}_n$ fordern. In (2') bzw. (3') genügt es, $\varepsilon = \frac{1}{n}$ zu betrachten.

Beweis. (2) \Leftrightarrow (2'), (3) \Leftrightarrow (3') und (1) \Rightarrow (4) sind offensichtlich. (4) \Rightarrow (1): vgl. Satz 31.12. Gilt (3), so erhält man eine monotone Einschließungsfolge durch

$$\max_{k \leq n} \underline{f}_k, \quad \min_{k \leq n} \bar{f}_k$$

Die gleichmässige Konvergenz folgt z.B. aus $0 \leq \|\underline{f}_n - f\| \leq \|\underline{f}_n - \bar{f}_n\| \rightarrow 0$. Wir bemerken

- Ist der Definitionsbereich von f in zwei Teilintervalle zerlegt und hat man auf jedem eine ε -Einschließung, so kann man diese zu einer ε -Einschließung von f zusammensetzen.

Sei nun angenommen, dass (1) aber nicht (3') gilt. Es gibt also ein $\varepsilon > 0$ so, dass f keine ε -Einschließung besitzt. Durch fortlaufende Halbierung erhalten wir eine Intervallschachtelung $[a_n, b_n]$ so, dass keine Einschränkung $f|_{[a_n, b_n]}$ eine ε -Einschließung besitzt. Sei x_0 der gemeinsame Punkt der Intervallschachtelung, z.B. $a < x_0 < b$. Nach Voraussetzung existieren der links- bzw. rechtsseitige Limes c bzw. d und somit ein $\delta > 0$ so, dass

$$|f(x) - c| < \frac{\varepsilon}{2} \text{ für } x_0 - \delta < x < x_0, \quad |f(x) - d| < \frac{\varepsilon}{2} \text{ für } x_0 < x < x_0 + \delta$$

Nun gibt es aber ein n mit $[a_n, b_n] \subseteq (x_0 - \delta, x_0 + \delta)$. Definiere

$$\underline{\phi} = \begin{cases} c - \frac{\varepsilon}{2} & \text{für } a_n \leq x < x_0 \\ f(x_0) & \text{für } x = x_0 \\ d - \frac{\varepsilon}{2} & \text{für } x_0 < x \leq b_n \end{cases} \quad \bar{\phi} = \begin{cases} c + \frac{\varepsilon}{2} & \text{für } a_n \leq x < x_0 \\ f(x_0) & \text{für } x = x_0 \\ d + \frac{\varepsilon}{2} & \text{für } x_0 < x \leq b_n \end{cases}$$

Dann ist $\underline{\phi}, \bar{\phi}$ eine ε -Einschließung von $f|_{[a_n, b_n]}$. Widerspruch.

Sei (2') angenommen und $a < x_0 \leq b$. Wir zeigen, dass es den linksseitigen Limes gibt. Ist $\varepsilon > 0$ gegeben, so wähle Treppenfunktion ϕ mit $\|\phi - f\| < \frac{\varepsilon}{2}$ und $\delta > 0$ so, dass ϕ auf $(x_0 - \delta, x_0)$ konstant ist. Es folgt

$$|f(x) - f(x')| \leq |f(x) - \phi(x)| + |\phi(x) - f(x')| < \varepsilon \text{ für alle } x, x' \in (x_0 - \delta, x_0)$$

und damit die Existenz des Limes nach Cauchy. \square

30.11-12 Sprungstetiges f ist beschränkt. Beweis im Tutorium. *Montones f ist sprungstetig.* Beweis für $a < x_0$ und monoton wachsendes f : Sei $s = \sup\{f(x) \mid x < x_0\}$. Ist $\varepsilon > 0$

gegeben, so hat man $\delta > 0$ so, dass $s - \varepsilon < f(x) \leq s$ für alle $x \in (x_0 - \delta, x_0)$, also $|f(x) - f(x')| < \varepsilon$ für alle $x, x' \in (x_0 - \delta, x_0)$. Existenz des linken Limes folgt mit Cauchy. \square

Korollar. Sei f stetig und Z_k eine Folge von Zerlegungen $a = x_{k0} < x_{k1} < \dots < x_{kn_k} = b$ mit $w(Z_k) := \max\{x_{kj} - x_{k,j-1} \mid 1 \leq j \leq n_k\} \rightarrow 0$ d.h. mit gegen Null konvergenter Maschenweite. f_k sei zu Z_k passende Treppenfunktion, die auf jedem Segment $(x_{k,j-1}, x_{kj})$ einen Wert $f(\xi)$ mit $\xi \in [x_{k,j-1}, x_{kj}]$ annimmt. Dann konvergiert die Folge f_k gleichmäßig gegen f .

Beweis. Sei $\varepsilon > 0$ gegeben. Da f auf I gleichmäßig stetig ist, gibt es ein $\delta > 0$ so, dass $|f(x) - f(x')| < \varepsilon$ falls $|x - x'| \leq \delta$. Wähle k_0 so, dass $w(Z_k) < \delta$ für alle $k \geq k_0$. Dann folgt mit passendem $\xi \in [x_{k,j-1}, x_{kj}]$

$$|f_k(x) - f(x)| = |f_k(\xi) - f(x)| < \varepsilon \text{ für } x \in [x_{k,j-1}, x_{kj}]$$

also $\|f_k - f\| < \varepsilon$. \square

(3) des Satzes gibt die Rechtfertigung, das Integral als Flächenmaß zu verstehen. Das Korollar rechtfertigt den in Naturwissenschaft und Technik üblichen Übergang von der Betrachtung endlicher Summen $\sum f(x)\Delta x$ zum Integral: auf die Wahl der als konstant angenommenen Werte von $f(x)$ auf Δx kommt es nicht an.

-3.6 Definition und einfache Eigenschaften des Integrals

30.3-6, 30.15-16, 30.7-9, 31.1-3, 31.12

$$\int f := \sum_{j=1}^n c_j(x_j - x_{j-1}), \quad c_j = f(\xi_j), \xi_j \in (x_{j-1}, x_j) \text{ für Treppenfunktion } f$$

$$\int_a^b f := \int f := \lim_{n \rightarrow \infty} \int f_n \text{ mit Treppenfunktionen } f_n \rightarrow f \text{ glm.}$$

Zu zeigen: Wohldefiniert: unabhängig von Zerlegung bzw. Folge f_n . Existenz des Limes.

(*) $\int_a^b f = \int_a^c f + \int_c^b f$ für $c \in [a, b]$ Additivität

(**) $\int_a^b f = \sum_{j=1}^n \int_{x_{j-1}}^{x_j} f$ mit Zerlegung Z

(a) $\int (\lambda f + \mu g) = \lambda \int f + \mu \int g$ Linearität

(b) $|\int f| \leq \int |f|$ Dreiecksungleichung

(c) $|\int f| \leq (b - a) \sup\{|f(x)| \mid x \in I\}$ Standardabschätzung

(d) $f \leq g \Rightarrow \int f \leq \int g$ Monotonie.

(e) $\int f = \lim_{n \rightarrow \infty} \int f_n$ für sprungstetige $f_n \rightarrow f$ glm.

Beweise zunächst für Treppenfunktion f . Wohldefiniertheit mit gemeinsamer Verfeinerung der Zerlegungen. (*) mit Zerlegung $c \in Z$. (**) folgt mit Induktion. Damit kann man alles weitere auf den Fall konstanter Funktionen, also elementare Algebra, reduzieren

Für sprungstetiges f . Existenz des Limes: $\int f_n$ ist Cauchyfolge wegen (a) und glm. Konvergenz. Eindeutigkeit: aus $f_n \rightarrow f$ glm. und $g_n \rightarrow f$ glm. folgt $f_1, g_1, f_2, g_2 \dots \rightarrow f$ glm. (a) folgt leicht - Übung. (b,c) vgl. 31.1.b.

(d) Nun nach (3) des Satzes vom stetigen Sprung $\underline{f}_n \leq f$ und $\bar{g}_n \geq g$ gleichmäßig konvergent gegen f bzw. g . Also $\underline{f}_n \leq f \leq g \leq \bar{g}_m$ für alle n, m also $\int \underline{f}_n \leq \int \bar{g}_m$ und es folgt $\int f = \lim_{n \rightarrow \infty} \int \underline{f}_n \leq \lim_{m \rightarrow \infty} \int \bar{g}_m = \int g$.

(e) $|\int f_n - \int f| = |\int (f_n - f)| \leq (b-a) \|f_n - f\| \rightarrow 0$.

(*) Für $f_n \rightarrow f$ glm. sind auch die Einschränkungen auf $[a, c]$ bzw. $[c, b]$ glm. konvergente Folgen von Treppenfkt. gegen die Einschränkungen von f . \square

-3.7 Riemannsches Integral

Für eine Zerlegung Z von I und Zwischenvektor $\vec{\xi} = (\xi_j)$ mit $\xi_j \in]x_{j-1}, x_j[$ definiert man die *Riemann-Summe*

$$R(Z, \vec{\xi}, f) = \sum_{j=0}^n f(\xi_k)(x_j - x_{j-1})$$

c ist das *Riemann-Integral* von f wenn

$$c = \lim_{k \rightarrow \infty} R(Z, \vec{\xi}_k, f) \text{ für } w(Z_k) \rightarrow 0$$

für jede Folge Z_k von Zerlegungen mit Maschenweite $w(Z_k) \rightarrow 0$ und jede Wahl der Zwischenpunkte $\vec{\xi}_k$. Jede sprungstetige Funktion ist Riemann-integrierbar, aber nicht umgekehrt. Genauer in Kap 23.2.3

-3.8 Integration stetiger Funktionen

Wir fassen diesen wichtigsten Fall nochmal zusammen.

Korollar. (vgl. 31.10) *Sei $f : I \rightarrow \mathbb{K}$ stetig. Dann gibt es ein Funktionen $F : I \rightarrow \mathbb{K}$ mit $F' = f$. Diese heißen Stammfunktion bzw. unbestimmtes Integral von f und sind eindeutig bestimmt bis auf eine Konstante, man schreibt*

$$F(x) = \int f(x) dx + C$$

Für $c, d \in I$ erhält man das bestimmte Integral als

$$\int_c^d f(x) dx = F(d) - F(c) =: F_c^d$$

Beweis. Eindeutigkeit: Aus $F_1' = f = F_2'$ folgt $(F_1 - F_2)' = 0$. Existenz: Setze

$$F(x) = \int_a^x f$$

Wegen der Stetigkeit von f gibt es zu jedem $x_0 \in I$ und $\varepsilon > 0$ ein $\delta > 0$ so, dass für alle $|\Delta x| \leq \delta$ gilt

$$\begin{aligned} f(x_0) - \varepsilon &\leq f(x) \leq f(x_0) + \varepsilon \quad \text{falls } |x - x_0| \leq \delta \\ (f(x_0) - \varepsilon)\Delta x &\leq \Delta F = \int_{x_0}^{x_0+\Delta x} f \leq (f(x_0) + \varepsilon)\Delta x \end{aligned}$$

also

$$\left| \frac{\Delta F}{\Delta x} - f(x_0) \right| < \varepsilon$$

Schließlich

$$\int_c^d f = \int_a^d f - \int_a^c f = F(d) - F(c) \quad \square$$

Satz -3.3 Summation. Sei $f : I \rightarrow \mathbb{K}$ stetig. Sei $W(c, d) \in \mathbb{K}$ für $c \leq d$ in I definiert. Dann sind äquivalent

(1) $W(c, d) = \int_c^d f(x)dx$ für alle $c \leq d$ in I

(2) Es gilt die Additivität

$$W(c, d) = W(c, x) + W(x, d) \quad \text{falls } c \leq x \leq d$$

und für alle $\varepsilon > 0$ gibt es ein $\delta > 0$ so, dass für alle Δx mit $|\Delta x| \leq \delta$ und alle $p \in I$ (2a) bzw (2b) gilt:

(2a) Es gibt $\xi \in [p, p + \Delta x]$ mit $|W(p, p + \Delta x) - f(\xi)\Delta x| \leq \varepsilon|\Delta x|$

(2b) Für alle $\xi \in [p, p + \Delta x]$ gilt $|W(p, p + \Delta x) - f(\xi)\Delta x| \leq \varepsilon|\Delta x|$

Korollar -3.4 Ist f stetig, so f Riemann-integrierbar und das Integral ist das Riemann-Integral.

$$\sum f(\xi_j)\Delta x_j \rightarrow \int_c^d f(x)dx \quad \text{für Zerlegungen } Z \text{ von } [c, d] \text{ mit } \max \Delta x_j \rightarrow 0$$

Beweis. Da f auf I gleichmäßig stetig ist, gibt es zu jedem $\varepsilon > 0$ ein $\delta > 0$ so dass

$$|x - x'| < \delta \Rightarrow |f(x) - f(x')| < \varepsilon$$

In (1) \Rightarrow (2a) haben wir die auf $[x_0, x_0 + \Delta x]$ konstante Funktion k mit Wert $f(\xi)$ und

$$(f(\xi) - \varepsilon)\Delta x = \int (k - \varepsilon) \leq \int_{x_0}^{x_0+\Delta x} f \leq \int (k + \varepsilon) = (f(\xi) + \varepsilon)\Delta x$$

In (2b) \Rightarrow (1) wählen wir Zerlegung Z von $[c, d]$ mit $x_j - x_{j-1} < \delta$ und erhalten

$$|W(x_k, x_{k+1}) - \int_{x_k}^{x_{k+1}} f(x)dx| \leq$$

$$\leq |W(x_k, x_{k+1}) - f(\xi_k)(x_{k+1} - x_k)| + |f(\xi_k)(x_{k+1} - x_k) - \int_{x_k}^{x_{k+1}} f(x)dx| \leq 2\varepsilon(x_{k+1} - x_k)$$

also mit der Additivität und Dreiecksungleichung

$$|W(c, d) - \int_c^d f(x)dx| \leq \sum_{k=0}^{n-1} |W(x_k, x_{k+1}) - \int_{x_k}^{x_{k+1}} f(x)dx| \leq 2\varepsilon(b - a)$$

Dies gilt für alle $\varepsilon > 0$, also $W(c, d) = \int_c^d f(x)dx$. \square

-2 Integrationstechniken

-2.1 Substitutionsregel

Lemma -2.1 Sei $x = x(t)$ auf $[a, b]$ differenzierbar und $f(x)$ auf dem Wertebereich $[c, d]$ von x definiert. Ist $F(x)$ eine Stammfunktion von $f(x)$ so ist $F(x(t))$ eine Stammfunktion von $\phi(t) = f(x(t)) \frac{\partial x}{\partial t} = f(x(t)) \cdot x'(t)$, d.h.

$$\int f(x) dx = \int f(x(t)) \frac{\partial x}{\partial t} dt + C$$

Ist x stetig differenzierbar und $f(x)$ stetig, so gilt

$$\int_{x(a)}^{x(b)} f(x) dx = \int_a^b f(x(t)) \frac{\partial x}{\partial t} dt$$

Beweis. Die erste Aussage folgt aus der Kettenregel. Nun mit dem Korollar des Hauptsatzes

$$\int_{x(a)}^{x(b)} f(x) dx = F(x(b)) - F(x(a)) = \int_a^b f(x(t)) \frac{\partial x}{\partial t} dt \quad \square$$

Beispiel $f(x) = x$

$$\int x(t) \frac{\partial x}{\partial t} dt = \int x dx = \frac{1}{2} (x(t))^2 + C$$

$$\int_0^{\frac{\pi}{2}} \sin t \cos t dt = \frac{1}{2} x^2 \Big|_0^1 = \frac{1}{2} \quad \text{mit } x = \sin t$$

Beispiel Für $f(x) = 1/x$ und $x(u) \neq 0$ erhält man

$$\int \frac{\frac{\partial x}{\partial u}(u)}{x(u)} du = \int \frac{1}{x} dx = \ln |x(u)| + C.$$

$$\int \cot t dt = \int \frac{1}{\sin t} \cos t dt = \ln |\sin t| + C \quad \text{mit } x = \sin t$$

-2.2 Differentiale

Um die Regeln korrekt und in Übereinstimmung mit der Praxis in Natur- und Ingenieurwissenschaften formulieren zu können, bedienen wir uns der Differentiale. Für eine auf $[a, b]$ definierte und differenzierbare Funktion $y = f(t)$ ist das *Differential* an der Stelle $p \in [a, b]$ die homogen lineare Funktion

$$df(p, dt) = \frac{\partial f}{\partial t}(p) \cdot dt = \frac{\partial y}{\partial t} dt \quad dt \in p - a \leq dt \leq b, \quad dt \neq 0$$

oder wenn man die Stelle p nicht explizit erwähnt

$$dy = \frac{\partial y}{\partial t} dt.$$

Das Differential an der Stelle p ist natürlich schon dann bekannt, wenn man es für ein einziges $dt \neq 0$ kennt.

Seien nun $x = x(t)$ und $y = y(t)$ auf $[a, b]$ differenzierbar und sei y eine differenzierbare Funktion von x , also nach der Kettenregel

$$y(t) = y = y(x) = y(x(t)), \quad \frac{\partial y}{\partial t}(p) = \frac{\partial y}{\partial x}(x(p)) \cdot \frac{\partial x}{\partial t}(p)$$

$$dy(p, dt) = \frac{\partial y}{\partial t}(p) \cdot dt = \frac{\partial y}{\partial x}(x(p)) \cdot \frac{\partial x}{\partial t}(p) \cdot dt = \frac{\partial y}{\partial x}(x(p)) \cdot dx(p, dt)$$

d.h. wir können $dy(p)$ auch als Differential bzgl. x verstehen. Dementsprechend haben wir die folgende *Konsistenzvoraussetzung* für den problemlosen Umgang mit Differentialen

- Alle betrachteten Größen sind stetig differenzierbare Funktionen einer vorgegebenen unabhängigen Variablen $t \in [a, b]$.

Dann gilt unzweideutig

$$dy = \frac{\partial y}{\partial x} dx$$

wie auch immer $y = y(x)$ differenzierbare Funktion von x , und

$$\frac{\partial y}{\partial x}(p) = \frac{dy}{dx}(p) \quad \text{falls} \quad \frac{\partial x}{\partial t}(p) \neq 0$$

$$\frac{\partial y}{\partial x}(p) = 0 \quad \text{falls} \quad \frac{\partial x}{\partial t}(p) = 0$$

Natürlich genügt es, wenn man das Differential $dx(p)$ als eine Funktion versteht, die nur für sehr kleine $|dt| \neq 0$ bzw. für infinitesimale dt definiert ist.

-2.3 Integrationsregel

Der folgende Satz fasst die üblichen Integrationsregeln zusammen und zeigt, dass das sogenannte “formale Rechnen” legitim und sinnvoll ist, wenn die Konsistenzbedingung für Differentiale erfüllt ist. Der Vorteil dieser Rechnung ist die intuitive Notation und die Option, die Argumentwerte weitgehend zu unterdrücken (da diese über die Abhängigkeit von t gekoppelt sind).

Satz -2.2 Seien $x = x(t)$, $y = y(t)$ und $z = z(t)$ auf $[a, b]$ differenzierbar und $f(x)$, $g(y)$, $h(z)$ stetige Funktionen auf den jeweiligen Wertebereichen. Für die Differentiale gelte

$$f(x(p))dx(p) = cg(y(p))dy(p) + h(z(p))dz(p) \quad \text{für alle } p \in [a, b]$$

$$\text{kurz} \quad f(x)dx = cg(y)dy + h(z)dz$$

Dann gilt: Für alle Stammfunktionen F , G , H von f , g bzw. h gibt es eine Konstante C mit

$$F(x(t)) = cG(y(t)) + H(z(t)) \quad \text{für alle } t \in [a, b]$$

kurz

$$\int f(x)dx = c \int g(y)dy + \int h(z)dz + C$$

Sind x, y, z stetig differenzierbar, so gilt

$$\int_{x(a)}^{x(b)} f(x)dx = c \int_{y(a)}^{y(b)} g(y)dy + \int_{z(a)}^{z(b)} h(z)dz$$

Linearitäts- und Substitutionsregel sind Spezialfälle. Beweis. Nach Voraussetzung haben wir

$$f(x) \frac{\partial x}{\partial t} dt = c g(y) \frac{\partial y}{\partial t} dt + h(z) \frac{\partial z}{\partial t} dt$$

also

$$\phi(t) := f(x) \frac{\partial x}{\partial t} = c g(y) \frac{\partial y}{\partial t} + h(z) \frac{\partial z}{\partial t}$$

Dann

$$\begin{aligned} \frac{\partial F}{\partial t} &= \frac{\partial F}{\partial x} \frac{\partial x}{\partial t} = f(x) \frac{\partial x}{\partial t} = \phi(t) \\ \frac{\partial(cG + H)}{\partial t} &= c \frac{\partial G}{\partial t} + \frac{\partial H}{\partial t} = c \frac{\partial G}{\partial y} \frac{\partial y}{\partial t} + \frac{\partial H}{\partial z} \frac{\partial z}{\partial t} = c g(y) \frac{\partial y}{\partial t} + h(z) \frac{\partial z}{\partial t} = \phi(t) \end{aligned}$$

Das ist die Aussage für unbestimmte Integrale und die für bestimmte folgt mit der Substitutionsregel.

-2.4 Teilterm als neue Variable

Ist $y = y(x)$ und $f(x)dx = cg(y)dy$ also $f(x) = cg(y) \frac{\partial y}{\partial x}$, so

$$\int f(x)dx = c \int g(y)dy, \quad \int_c^d f(x)dc = c \int_{y(c)}^{y(d)} g(y)dy$$

Beispiel. $\int e^{2 \sin x} \cos x dx$, $y(x) = 2 \sin x$, $g(y) = e^y$

$$e^{2 \sin x} \cos x dx = \frac{1}{2} e^y dy$$

$$\int e^{2 \sin x} \cos x dx = \frac{1}{2} \int e^y dy = \frac{1}{2} e^y + C = \frac{1}{2} e^{\sin x} + C$$

$$\int_0^{\frac{\pi}{2}} e^{2 \sin x} \cos x dx = \frac{1}{2} \int_0^2 e^y dy = \frac{1}{2} (e^2 - e^0)$$

Beispiel. $f(x) = g(\alpha x + \beta)$ mit $\alpha \neq 0$ und $g(y)$ stetig, so ist mit $y = \alpha x + \beta$

$$\frac{\partial y}{\partial x} = \alpha, \quad f(x) = \frac{1}{\alpha} g(y) \frac{\partial y}{\partial x}, \quad f(x)dx = \frac{1}{\alpha} g(y)dy$$

$$\int f(x)dx = \frac{1}{\alpha} \int g(y)dy + C$$

$$\int_c^d f(x) dx = \frac{1}{\alpha} \int_{\alpha c + \beta}^{\alpha d + \beta} g(x) dx + C$$

Beispiel. Auf \mathbb{R} suchen wir $\int \cos x \sin^2 x dx$. Mit $y = \sin x$ und $g(y) = y^2$

$$\sin^2 x \cos x dx = y^2 dy$$

$$\int \sin^2 x \cos x dx = \int y^2 dy = \frac{y^3}{3} + C = \frac{\sin^3 x}{3} + C$$

$$\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \sin^2 x \cos x dx = \int_{-1}^1 y^2 dy = \frac{1^3}{3} - \frac{(-1)^3}{3} = \frac{2}{3}$$

-2.5 Integration durch Umkehrung

Ist $\frac{\partial x}{\partial u}(u) \neq 0$ für alle u und der Wertebereich von $x(u)$ der Definitionsbereich von x , so hat $x = x(u)$ eine Umkehrfunktion $u = u(x)$ und man kann

$$\Phi(u) = \int f(x(u)) \frac{\partial x}{\partial u} du + C$$

nach x auflösen

$$\int f(x) dx = F(x) + C = \Phi(u(x)) + C$$

$$\int_c^d f(x) dx = \Phi(u(d)) - \Phi(u(c))$$

Beweis. Da $\frac{\partial x}{\partial u}$ stetig und nie Null ist, ist $\frac{\partial x}{\partial u}$ nach dem Zwischenwertsatz entweder positiv für alle u oder negativ. Also ist $x(u)$ entweder streng monoton wachsend oder streng monoton fallend. Hieraus folgt die Existenz der Umkehrfunktion $u = u(x)$. Alles weitere folgt aus dem Substitutionslemma. \square

Beispiel. Wir suchen $\int \frac{1}{\sin x} dx$ auf $(0, \pi)$. Die Substitution $x = 2 \arctan u$ führt wegen $\frac{\partial x}{\partial u} = \frac{2}{1+u^2}$ und

$$\begin{aligned} \sin x &= 2 \sin \frac{x}{2} \cos \frac{x}{2} = 2 \tan \frac{x}{2} \cos^2 \frac{x}{2} = 2 \tan \frac{x}{2} \frac{\cos^2 \frac{x}{2}}{\sin^2 \frac{x}{2} + \cos^2 \frac{x}{2}} \\ &= 2 \tan \frac{x}{2} \frac{1}{1 + \tan^2 \frac{x}{2}} \end{aligned}$$

auf das Integral

$$\int \frac{1}{\sin x} dx = \int \frac{1+u^2}{2u} \frac{2}{1+u^2} du = \int \frac{1}{u} du = \ln |u| + C.$$

Rücksubstitution $u = \tan \frac{x}{2}$ liefert

$$\int \frac{1}{\sin x} dx = \ln \left| \tan \frac{x}{2} \right| + C.$$

-2.6 Partielle Integration

Für $w(t) = u(t) \cdot v(t)$ haben wir nach der Produktregel $\frac{\partial w}{\partial t} = v \frac{\partial u}{\partial t} + u \frac{\partial v}{\partial t}$ also das Differential

$$dw = duv = v \frac{\partial u}{\partial t} dt + u \frac{\partial v}{\partial t} dt = v du + u dv$$

(unter Voraussetzung der Konsistenz) und es folgt

$$\begin{aligned}
 uv = w &= \int 1 dw = \int v \frac{\partial u}{\partial t} dt + \int u \frac{\partial v}{\partial t} dt + C = \int v du + \int u dv + C \\
 \int_{w(a)}^{w(b)} 1 dw &= \int_a^b v \frac{\partial u}{\partial t} dt + \int_a^b u \frac{\partial v}{\partial t} dt = \int_{u(a)}^{u(b)} v du + \int_{v(a)}^{v(b)} u dv \\
 u(t)v(t) \Big|_a^b &:= u(b)v(b) - u(a)v(a) = \int_{w(a)}^{w(b)} 1 dw
 \end{aligned}$$

Beispiel. Für $\int x^2 e^{3x} dx$ wenden wir partielle Integration zweimal an:

$$\begin{aligned}
 \int x^2 e^{3x} dx &= x^2 \frac{e^{3x}}{3} - \int \frac{e^{3x}}{3} 2x dx = \frac{1}{3} x^2 e^{3x} - \frac{2}{3} \int x e^{3x} dx \\
 &= \frac{1}{3} x^2 e^{3x} - \frac{2}{3} \left(x \frac{e^{3x}}{3} - \int \frac{e^{3x}}{3} 1 dx \right) \\
 &= \frac{1}{3} x^2 e^{3x} - \frac{2}{9} x e^{3x} + \frac{2}{9} \int e^{3x} dx \\
 &= \left(\frac{1}{3} x^2 - \frac{2}{9} x + \frac{2}{27} \right) e^{3x} + C. \quad \blacksquare
 \end{aligned}$$

Beispiel. Für $\int \cos^2 x dx$ hilft ein einfacher Trick:

$$\begin{aligned}
 \int \cos^2 x dx &= \int \cos x \cdot \cos x dx = \sin x \cos x - \int \sin x (-\sin x) dx \\
 &= \sin x \cos x + \int \sin x \sin x dx \\
 &= \sin x \cos x + \int (1 - \cos^2 x) dx \\
 &= \sin x \cos x + x - \int \cos^2 x dx.
 \end{aligned}$$

Hieraus folgt schließlich

$$2 \int \cos^2 x dx = \sin x \cos x + x$$

bzw.

$$\int \cos^2 x dx = \frac{1}{2} (\sin x \cos x + x) + C.$$

In diesem Beispiel hätte man einfacher benutzen können, dass $\cos^2 x = \frac{1}{2} \cdot (1 + \cos 2x)$ ist. Damit bekommt man sofort

$$\int \cos^2 x \, dx = \frac{1}{2} \int dx + \frac{1}{2} \int \cos 2x \, dx = \frac{1}{2} \left(x + \frac{\sin 2x}{2} \right) + C. \quad \blacksquare$$

Beispiel. Bei $\int \ln t \, dt$ hilft ein Trick: Wir wählen $dt = 1 \, du$

$$\ln t \, dt = \ln t(1 \, dt) = v \, du = duv - u \, dv, \quad u = \int 1 \, dt = t$$

$$\int \ln t \, dt = t \ln t - \int t \frac{1}{t} \, dt = t \ln t - \int 1 \, dt = t \ln t - t + C$$

Gegenbeispiel.

$$\int t \sin t \, dt$$

Mit $\frac{\partial u}{\partial t} = t$ und $v(t) = \sin t$ erhält man

$$\int t \sin t \, dt = \frac{t^2}{2} \sin t - \int \frac{t^2}{2} \cos t \, dt$$

Das Integral $\int t^2 \cos t \, dt$ ist aber komplizierter als das Ausgangsintegral. ■

-1 Integration rationaler Funktionen

-1.1 Polynome und euklidischer Algorithmus

Polynome mit Koeffizienten in \mathbb{K} sind ein Spezialfall der Potenzreihen. Meist braucht man nur einen Körper K und schreibt $K[x]$ für den Ring der Polynome mit Koeffizienten in K . Die Polynomdivision sollten Sie aus der Schule kennen. Genaueres in `supp.pdf`

Korollar -1.1 Zu $p(x) \in K[x]$ und $a \in K$ gibt es eindeutig bestimmte $r \in K$ und $q(x) \in K[x]$ mit

$$p(x) = q(x)(x - a)^k + r(x), \quad \deg r(x) < k$$

Wir definieren nun die Teilbarkeit in $K[x]$ analog zu der in \mathbb{Z}

$$p(x) \text{ teilt } q(x) \Leftrightarrow \exists r(x). \quad q(x) = p(x)r(x)$$

$p(x)$ und $q(x)$ sind *teilerfremd*, wenn ihre einzigen gemeinsamen Teiler Konstanten sind.

Satz -1.2 Bezout. Zu teilerfremden $p(x), q(x) \in K[x]$ gibt es $r(x), s(x) \in K[x]$ mit

$$1 = p(x)r(x) + q(x)s(x) \quad \deg r(x) < \deg q(x), \quad \deg s(x) < \deg p(x)$$

Diese bestimmt man mit dem Euklidischen Algorithmus für Polynome.

Beweis exemplarisch in \mathbb{Z} . Die beiden Startzeilen sind trivial. Dann zieht man immer ein Vielfaches der letzten aktuellen Zeile von der vorletzten ab (Division mit Rest der Einträge in der ersten Spalte). Dabei entsteht wieder eine gültige Relation.

$$\begin{aligned} 98 &= 1 \cdot 98 &+& 0 \cdot 27 \\ 27 &= 0 \cdot 98 &+& 1 \cdot 27 \\ 17 &= 1 \cdot 98 &+& -3 \cdot 27 \\ 10 &= -1 \cdot 98 &+& 4 \cdot 27 \\ 7 &= 2 \cdot 98 &+& -7 \cdot 27 \\ 3 &= -3 \cdot 98 &+& 11 \cdot 27 \\ 1 &= 8 \cdot 98 &+& -29 \cdot 27 \end{aligned}$$

-1.2 Faktorzerlegung

Ein nicht konstantes Polynom $p(x)$ heißt *unzerlegbar* oder *irreduzibel*, wenn in jeder Zerlegung $p(x) = q(x)r(x)$ einer der Faktoren konstant ist. Diese spielen die Rolle der Primzahlen. Man bemerkt, dass *ein irreduzibles $p(x)$ ein Produkt $q(x)r(x)$ nur dann teilt, wenn es mindestens einen der Faktoren teilt*: sonst sind z.B. $p(x)$ und $q(x)$ teilerfremd, also gibt es $a(x)p(x) + b(x)q(x) = 1$ und $p(x)$ teilt $r(x) = a(x)p(x)r(x) + b(x)q(x)r(x)$. Wie für \mathbb{Z} beweist man

Satz -1.3 *Jedes nichtkonstante Polynom $p(x) \in K[x]$ hat eine bis auf die Reihenfolge eindeutige Darstellung mit normierten irreduziblen $p_i(x) \in K[x]$ und $a \in K$*

$$p(x) = ap_1(x) \cdots p_n(x)$$

Beweis: Die Existenz folgt durch Ordnungsinduktion über den Grad: Ist $p(x)$ nicht schon irreduzibel, so $p(x) = q(x)r(x)$ mit Polynomen kleineren Grades, also $q(x) = bq_1(x) \cdots q_k(x)$ und $r(x) = cr_1(x) \cdots r_l(x)$ und somit $p(x) = bcq_1(x) \cdots q_k(x) \cdot r_1(x) \cdots r_l(x)$. Die Eindeutigkeit folgt nun durch Induktion über die Anzahl der Faktoren: Ist

$$p(x) = ap_1(x) \cdots p_n(x) = bq_1(x) \cdots q_m(x)$$

so $a = b$ da die $p_i(x)$ und $q_i(x)$ alle normiert sind. Auch teilt $p_1(x)$ eines der $q_i(x)$, nach Umm Nummerierung also $q_1(x)$ und wegen Normiertheit folgt $p_1(x) = q_1(x)$. Nun kürzt man $p_1(x)$ und beruft sich auf die Induktionsannahme. \square .

-1.3 Körper der rationalen Funktionen

Wie von \mathbb{Z} zu \mathbb{Q} kann man vom Polynomring $K[x]$ zum Körper $K(x)$ der rationalen Funktionen übergehen. Man betrachte Quotienten, d.h. Paare

$$\frac{p(x)}{q(x)} \quad \text{mit } q(x) \neq \mathbf{0}$$

und setzt zwei solche Quotienten gleich

$$\frac{p(x)}{q(x)} = \frac{r(x)}{s(x)} \Leftrightarrow p(x)s(x) = q(x)r(x)$$

Addition und Multiplikation sind wie gewohnt definiert

$$\frac{p(x)}{q(x)} + \frac{r(x)}{s(x)} = \frac{p(x)s(x) + q(x)r(x)}{q(x)s(x)}, \quad \frac{p(x)}{q(x)} \cdot \frac{r(x)}{s(x)} = \frac{p(x)r(x)}{q(x)s(x)}$$

Für $K \subseteq \mathbb{C}$ können wir natürlich auch die Funktion

$$t \mapsto \frac{p(t)}{q(t)} \quad t \in K, \quad q(t) \neq 0$$

betrachten und mit dem Quotienten identifizieren. Diese Funktionen sind auf ihrem Definitionsbereich beliebig oft differenzierbar und integrierbar.

-1.4 Partialbruchzerlegung

Wir wollen rationale Funktionen in einfacherer Weise darstellen. Dazu die folgenden drei Reduktionsschritte

- Ausdividieren $f(x) = q(x)g(x) + r(x)$, $\deg r(x) < \deg g(x)$

$$\frac{f(x)}{g(x)} = q(x) + \frac{r(x)}{g(x)} \quad \deg r(x) < \deg g(x)$$

- Bezout $a(x)h(x) + b(x)g(x) = 1$

$$\frac{f(x)}{g(x)h(x)} = \frac{f(x)a(x)}{g(x)} + \frac{f(x)b(x)}{h(x)} \quad \text{für teilerfremde } g(x), h(x)$$

- Zerlegen $f(x) = q(x)g(x) + r(x)$, $\deg r(x) < \deg g(x)$

$$\frac{f(x)}{g(x)^n} = \frac{r(x)}{g(x)^n} + \frac{q(x)}{g(x)^{n-1}}, \quad \deg r(x) < \deg g(x)$$

Satz -1.4 Zu Polynomen $f(x), g(x) \in K[x]$ mit $g(x) = \prod_{i=1}^n p_i(x)^{k_i}$ mit irreduziblen normierten $p_i(x)$ gibt es eine Darstellung

$$\frac{f(x)}{g(x)} = r(x) + \sum_{i=1}^n \sum_{k=1}^{k_i} \frac{r_{ik}(x)}{p_i(x)^k} \quad \text{mit } \deg r_{ik}(x) < \deg p_i(x)$$

Dabei ist $r(x) = \mathbf{0}$ genau dann, wenn $\deg g(x) < \deg f(x)$. Die Darstellung ist eindeutig.

Beweis. Existenz. Wir führen exemplarisch eine Partialbruchzerlegung von $\frac{71}{60}$ in \mathbb{Q} aus, indem wir folgende Schritte anwenden

$$\frac{a}{b} = q + \frac{r}{b} \quad \text{mit } a = qr + b \quad 0 \leq r < b$$

$$\frac{c}{ab} = \frac{sc}{a} + \frac{rc}{b} \quad \text{mit } ra + sb = 1 \text{ für teilerfremde } a, b$$

$$\frac{a}{p^k} = \frac{q}{p^{k-1}} + \frac{r}{p^k} \quad \text{mit } a = qp + r, \quad 0 \leq r < p$$

Lösung

$$\begin{aligned} \frac{71}{60} &= 1 + \frac{11}{60} \\ 60 &= 5 \cdot 12, \quad 1 = 5 \cdot 5 - 2 \cdot 12, \quad \frac{11}{60} = \frac{55}{12} - \frac{22}{5} = \frac{7}{12} - \frac{2}{5} \\ 12 &= 3 \cdot 4, \quad 1 = 4 - 3, \quad \frac{7}{12} = \frac{28}{12} - \frac{21}{12} = 1 + \frac{1}{3} - \frac{3}{4} \\ 3 &= 1 \cdot 2 + 1, \quad \frac{3}{4} = \frac{2}{4} + \frac{1}{4} = \frac{1}{2} + \frac{1}{4} \\ \frac{71}{60} &= 1 + \frac{11}{60} = 1 + \frac{7}{12} - \frac{2}{5} = 2 - \frac{2}{5} + \frac{1}{3} - \frac{1}{2} - \frac{1}{4} \end{aligned}$$

Hat man eine Partialbruchzerlegung wie im Satz, bringt man die Partialbrüche auf den Hauptnenner $g(x)$ und addiert sie auf, so erhält man im Zähler einen Grad $< \deg g(x)$. Damit ist $r(x)$ das Ergebnis bei der Division mit Rest und eindeutig bestimmt.

Zum weiteren Beweis der Eindeutigkeit dürfen wir also $r(x) = 0$ annehmen. Seien also zwei Zerlegungen gegeben

$$\sum_{i=1}^n \sum_{k=1}^{k_i} \frac{r_{ik}(x)}{p_i(x)^k} = \sum_{i=1}^n \sum_{k=1}^{k_i} \frac{s_{ik}(x)}{p_i(x)^k} \quad \text{mit } \deg r_{ik}(x), \deg s_{ik}(x) < \deg p_i(x)$$

Sei k maximal so, dass für ein i $r_{ik}(x) \neq s_{ik}(x)$. Indem man in der Gleichung alle gleichen Terme streicht und dann mit dem Hauptnenner $N(x)$ multipliziert, erhält man in jedem Summanden einen Faktor $p_i(x)$ im Zähler, ausser in

$$\frac{r_{ik}(x)N(x)}{p_i(x)^k} \quad \text{und} \quad \frac{s_{ik}(x)N(x)}{p_i(x)^k}$$

Also ist $p_i(x)$ Teiler von $h(x)(r_{ik}(x) - s_{ik}(x))$ wobei

$$h(x) = \frac{N(x)}{p_i(x)^k}$$

nach Wahl von k zu $p_i(x)$ teilerfremd ist. Also ist $p_i(x)$ Teiler von $r_{ik}(x) - s_{ik}(x)$, Da $\deg(r_{ik}(x) - s_{ik}(x)) < \deg p_i(x)$ folgt $r_{ik}(x) - s_{ik}(x) = \mathbf{0}$. \square

-1.5 Fundamentalsatz der Algebra

Satz -1.5 Die einzigen normierten irreduziblen Polynome in $\mathbb{C}[x]$ sind die linearen $x - \alpha$.

Korollar -1.6 Ist $p(x) = a_n x^n + \dots + a_0$ ein Polynom mit reellen Koeffizienten a_0, \dots, a_n , so ist eine komplexe Zahl α Nullstelle von $p(x)$ genau dann, wenn auch ihre Konjugierte $\bar{\alpha}$ Nullstelle von $p(x)$ ist; man hat eine Zerlegung

$$p(x) = a_n (x - \alpha_1)^{n_1} \dots (x - \alpha_l)^{n_l} q_{l+1}(x)^{n_{l+1}} \dots q_m(x)^{n_m}$$

in lineare und quadratische reelle Polynome, die den reellen Nullstellen bzw. den Paaren konjugiert komplexer Nullstellen entsprechen. Dabei ist das Polynom zu $\alpha = a + bi$ und $\bar{\alpha} = a - bi$ das folgende

$$q(x) = (x - \alpha)(x - \bar{\alpha}) = (x - (a + bi))(x - (a - bi)) = x^2 - 2ax + a^2 + b^2.$$

Beweis. Die erste Behauptung ergibt sich sofort daraus, dass die Konjugation mit Addition und Multiplikation verträglich ist und reelle Zahlen festlässt. Dann fasst man im Fundamentalsatz die Paare konjugierter zusammen. \square

Korollar -1.7 Über \mathbb{C} gilt bei der Partialbruchzerlegung $\deg p_i(x) = 1$. Über \mathbb{R} hat man $\deg p_i(x) = 1$ bzw. Paare konjugierter komplexer Partialbrüche

$$\frac{A}{(x - \alpha)} + \frac{\bar{A}}{(x - \bar{\alpha})} = \frac{ax + b}{(x^2 + cx + d)}$$

$$c = -(\alpha + \bar{\alpha}), d = \alpha\bar{\alpha}, a = A + \bar{A}, b = -(A\bar{\alpha} + \bar{A}\alpha) \in \mathbb{R}$$

-1.6 Partialbruchzerlegung reeller rationaler Funktionen durch Ansatz

(Weitere Methoden in supp.pdf) Seien Q, R reelle Polynome mit $Q \neq 0$.

1. Schritt Ist der Grad von R größer oder gleich dem von Q , so liefert eine Polynomdivision von R durch Q Polynome P und S mit

$$\frac{R}{Q} = S + \frac{P}{Q},$$

wobei nun der Grad von P kleiner als der von Q ist.

2. Schritt Man zerlegt das Nennerpolynom Q nach dem Fundamentalsatz

$$Q(x) = q \prod_{i=1}^r (x - b_i)^{k_i} \prod_{j=1}^s (x^2 + 2c_j x + d_j)^{m_j} \quad (-1.1)$$

3. Schritt. Ansatz

$$\frac{P(x)}{Q(x)} = \sum_{i=1}^r \sum_{k=1}^{k_i} \frac{A_{ik}}{(x - b_i)^k} + \sum_{j=1}^s \sum_{m=1}^{m_j} \frac{B_{jm}x + C_{jm}}{(x^2 + 2c_j x + d_j)^m} \quad (-1.2)$$

Die Zahlen A_{ik}, B_{jm}, C_{jm} können beispielsweise ermittelt werden, indem man (-1.2) mit Q multipliziert und durch Koeffizientenvergleich ein lineares Gleichungssystem für die gesuchten Größen aufstellt. Auch das Einsetzen der Nullstellen von Q in die entstehenden Polynome kann hilfreich sein.

Achtung: Macht man den Ansatz, ohne darauf zu achten, dass der Grad von P kleiner als der Grad von Q ist, so geben die aus dem Gleichungssystem bestimmten Werte i.A. keine Lösung der Aufgabe, P/Q in Partialbrüche zu zerlegen.

Beispiel Man bestimme Partialbruchzerlegung von

$$\frac{x^4 + 1}{x^4 - x^3 - x + 1}$$

1. Schritt Polynomdivision

$$\frac{x^4 + 1}{x^4 - x^3 - x + 1} = 1 + \frac{x^3 + x}{x^4 - x^3 - x + 1}.$$

2. Schritt Faktorisierung des Nennerpolynoms

$$x^4 - x^3 - x + 1 = (x - 1)(x^3 - 1) = (x - 1)^2(x^2 + x + 1).$$

3. Schritt Partialbruchzerlegung. Der Ansatz

$$\frac{x^3 + x}{x^4 - x^3 - x + 1} = \frac{A_1}{x - 1} + \frac{A_2}{(x - 1)^2} + \frac{Bx + C}{x^2 + x + 1}$$

liefert nach Multiplikation mit $Q(x) = x^4 - x^3 - x + 1 = (x - 1)^2(x^2 + x + 1)$

$$x^3 + x = A_1(x - 1)(x^2 + x + 1) + A_2(x^2 + x + 1) + (Bx + C)(x - 1)^2 \quad (-1.3)$$

bzw. nach Ausmultiplizieren und Zusammenfassen

$$x^3 + x = (A_1 + B)x^3 + (A_2 - 2B + C)x^2 + (A_2 + B - 2C)x + (A_2 - A_1 + C).$$

Ein Vergleich der Koeffizienten auf der linken bzw. rechten Seite ergibt das lineare Gleichungssystem

$$\text{bei } x^3: \quad A_1 + B = 1$$

$$\text{bei } x^2: \quad A_2 - 2B + C = 0$$

$$\text{bei } x^1: \quad A_2 + B - 2C = 1$$

$$\text{bei } x^0: \quad A_2 - A_1 + C = 0.$$

Die Lösung dieses Gleichungssystems ist

$$A_1 = \frac{2}{3}, \quad A_2 = \frac{2}{3}, \quad B = \frac{1}{3}, \quad C = 0.$$

Also

$$\frac{x^4 + 1}{x^4 - x^3 - x + 1} = 1 + \frac{2}{3} \frac{1}{x - 1} + \frac{2}{3} \frac{1}{(x - 1)^2} + \frac{1}{3} \frac{x}{x^2 + x + 1}.$$

Alternativ hätte man z.B. in (-1.3) $x = 1$ einsetzen können und so A_2 sofort gefunden.

Korollar -1.8 *Benutzt man den Ansatz, so bestimmt man den Partialbruch zur höchsten Potenzen eines Linearfaktors durch Einsetzen der Nullstelle im Koeffizientenvergleich*

Weitere Methoden siehe supp.pdf

-1.7 Integrale rationaler Grundfunktionen

Zu allen in -1.2 vorkommenden Brüchen lassen sich durch partielle Integration und Substitution die Stammfunktionen effektiv bestimmen. Einige der folgenden Regeln müssen dazu wiederholt angewandt werden:

$$\int \frac{dx}{(x-b)^k} = \begin{cases} \frac{1}{1-k} (x-b)^{1-k} & \text{falls } k > 1 \\ \ln|x-b| & \text{falls } k = 1, \end{cases}$$

$$\int \frac{dx}{x^2 + 2cx + d} = \frac{1}{\sqrt{d-c^2}} \arctan \frac{x+c}{\sqrt{d-c^2}},$$

$$\int \frac{dx}{(x^2 + 2cx + d)^m} = \frac{x+c}{2(m-1)(d-c^2)(x^2 + 2cx + d)^{m-1}} + \frac{(2m-3)}{2(m-1)(d-c^2)} \int \frac{dx}{(x^2 + 2cx + d)^{m-1}} \quad \text{für } m \geq 2,$$

$$\int \frac{\alpha x + \beta}{x^2 + 2cx + d} dx = \frac{\alpha}{2} \ln(x^2 + 2cx + d) + (\beta - \alpha c) \int \frac{dx}{x^2 + 2cx + d},$$

$$\int \frac{\alpha x + \beta}{(x^2 + 2cx + d)^m} dx = \frac{-\alpha}{2(m-1)(x^2 + 2cx + d)^{m-1}} + (\beta - \alpha c) \int \frac{dx}{(x^2 + 2cx + d)^{m-1}} \quad \text{für } m \geq 2.$$

Beispiel.

$$\begin{aligned} \int \frac{x^4 + 1}{x^4 - x^3 - x + 1} dx &= \int 1 dx + \frac{2}{3} \int \frac{dx}{x-1} + \frac{2}{3} \int \frac{dx}{(x-1)^2} \\ &\quad + \frac{1}{3} \int \frac{x dx}{x^2 + x + 1} \\ &= x + \frac{2}{3} \ln|x-1| - \frac{2}{3} \frac{1}{x-1} + \frac{1}{6} \ln(x^2 + x + 1) \\ &\quad - \frac{1}{3\sqrt{3}} \arctan \frac{2x+1}{\sqrt{3}} + C. \end{aligned}$$

0 Uneigentliche Integrale

Siehe AnaSkript 1 Kap. 32

1 Endlichdimensionale normierte Räume

Die Seitenangaben in Kap 1-4 beziehen sich auf das Skript AnaII-mehrdim von Prof. Alber.

1.1 Affine Räume

Ein *Vektorraum* V über (dem *Skalkörper*) K oder kurz K -*Vektorraum* wird gegeben durch eine Menge V (deren Elemente auch *Vektoren* heissen sollen) mit Operationen

$$(\vec{x}, \vec{y}) \mapsto \vec{x} + \vec{y}, \quad \vec{x} \mapsto -\vec{x},$$

auf V , der Konstante 0 (*Nullektor*) und der *Multiplikation mit Skalaren*

$$(r, \vec{x}) \mapsto \lambda \vec{x} \quad \text{wobei } r \in K, \vec{x} \in V$$

so, dass (V1) - (V8) gelten:

$$(V1) \quad \text{für alle } \vec{u}, \vec{v}, \vec{w} \text{ in } V \text{ gilt } \vec{u} + (\vec{v} + \vec{w}) = (\vec{u} + \vec{v}) + \vec{w}$$

$$(V2) \quad \text{für alle } \vec{u}, \vec{v} \text{ in } V \text{ gilt } \vec{u} + \vec{v} = \vec{v} + \vec{u}$$

$$(V3) \quad \text{für alle } \vec{v} \text{ in } V \text{ gilt } \vec{0} + \vec{v} = \vec{v}$$

$$(V4) \quad \text{für alle } \vec{v} \text{ in } V \text{ gilt } \vec{v} + (-\vec{v}) = \vec{0}$$

$$(V5) \quad \text{für alle } r \text{ in } K \text{ und } \vec{v}, \vec{w} \text{ in } V \text{ gilt } r(\vec{v} + \vec{w}) = r\vec{v} + r\vec{w}$$

$$(V6) \quad \text{für alle } \vec{v} \text{ in } V \text{ gilt } 1\vec{v} = \vec{v}$$

$$(V7) \quad \text{für alle } r, s \text{ in } K \text{ und } \vec{v} \text{ in } V \text{ gilt } (r + s)\vec{v} = r\vec{v} + s\vec{v}$$

$$(V8) \quad \text{für alle } r, s \text{ in } K \text{ und } \vec{v} \text{ in } V \text{ gilt } r(s\vec{v}) = (rs)\vec{v}.$$

Sei X eine Menge und

$$(\vec{v}, P) \mapsto \vec{v} + P, \quad V \times X \rightarrow X$$

eine Abbildung so, dass gilt:

- Zu $P, Q \in X$ gibt es genau ein $\vec{v} \in V$ mit $Q = \vec{v} + P$
- $(\vec{w} + \vec{v}) + P = \vec{w} + (\vec{v} + P), \quad \vec{0} + P = P$

dann ist (X, V) *affiner Raum* mit *Punktmenge* X und Vektorraum V . Es folgt, dass die Abbildung

$$P \mapsto \vec{v} + P, \quad X \rightarrow X$$

bijektiv ist - die Umkehrabbildung ist $P \mapsto \vec{v} + P$.

1.2 Euklidische Vektorräume

Sei V ein \mathbb{R} -Vektorraum. (V, \cdot) ist ein *euklidischer Vektorraum* und \cdot ein *Skalarprodukt* falls für alle $\vec{x}, \vec{y}, \vec{z} \in V, r \in \mathbb{R}$

1. $\vec{x} \cdot \vec{y} = \vec{y} \cdot \vec{x} \in \mathbb{R}$ (*symmetrisch*)
2. $\vec{x} \cdot (\vec{y} + \vec{z}) = \vec{x} \cdot \vec{y} + \vec{x} \cdot \vec{z}, \quad (r\vec{x}) \cdot \vec{y} = r(\vec{x} \cdot \vec{y})$ (*linear*)
3. $\vec{x} \cdot \vec{x} > 0$ für $\vec{x} \neq \vec{0}$ (*positiv definit*)

Man definiert

- $|\vec{x}| = \sqrt{\vec{x} \cdot \vec{x}}$ *euklidische Norm*
- $\vec{x} \perp \vec{y} \Leftrightarrow \vec{x} \cdot \vec{y} = 0$ *orthogonal*

und bemerkt, dass (*): $|r\vec{x}| = |r| \cdot |\vec{x}|$.

Satz 1.1 (Cauchy-Schwarz). *In einem euklidischen Vektorraum gilt*

$$|\vec{x} \cdot \vec{y}| \leq |\vec{x}| \cdot |\vec{y}|$$

Beweis. Wegen (*) o.B.d.A. $|\vec{y}| = 1$. Setze $r = \vec{x} \cdot \vec{y}$. Dann

$$0 \leq |\vec{x} - r\vec{y}|^2 = |\vec{x}|^2 - 2r(\vec{x} \cdot \vec{y}) + r^2|\vec{y}|^2 = |\vec{x}|^2 - r^2$$

also $|r| \leq |\vec{x}|$. \square

1.3 Normierte Vektorräume

Eine *Norm* auf einem \mathbb{R} -Vektorraum V ist eine Abbildung $\vec{x} \mapsto \|\vec{x}\| \in \mathbb{R}_{\geq 0}$ so, dass

1. $\|\vec{x}\| = 0 \Leftrightarrow \vec{x} = \vec{0}$
2. $\|r\vec{x}\| = |r| \cdot \|\vec{x}\|$
3. $\|\vec{x} + \vec{y}\| \leq \|\vec{x}\| + \|\vec{y}\|$

Korollar 1.2 *Die euklidische Norm ist eine Norm.*

Beweis folgt aus Cauchy-Schwarz s.S.4 (AnaII-mehrdim)

1.4 Metrische Räume

Sei X eine Menge und $d : X \times X \rightarrow \mathbb{R}_{\geq 0}$. Dann ist (X, d) ein *metrischer Raum*, falls für alle $P, Q, R \in X$

1. $d(P, Q) = d(Q, P)$
2. $d(P, Q) = 0 \Leftrightarrow P = Q$
3. $d(P, R) \leq d(P, Q) + d(Q, R)$ (*Dreiecksungleichung*)

Satz 1.3 *Sei $(V, \|\cdot\|)$ normierter Vektorraum und (X, V) affiner Raum. Definiere*

$$d(P, Q) = \|\vec{v}\| \quad \text{falls } Q = \vec{v} + P$$

Dann ist (X, d) ein metrischer Raum.

Beweis. Sei $R = \vec{w} + Q$, also $R = (\vec{w} + \vec{v}) + P$ und somit

$$d(P, R) = \|\vec{w} + \vec{v}\| \leq \|\vec{w}\| + \|\vec{v}\| = d(P, Q) + d(Q, R) \quad \square$$

1.5 Konvergenz

Sei (X, d) metrischer Raum. Zu $\varepsilon > 0$ und $Q \in X$ definieren wir die (*offene*) ε -*Umgebung* durch

$$U_\varepsilon(Q) = \{P \in X \mid d(P, Q) < \varepsilon\}$$

$U \subseteq X$ ist *Umgebung* von Q falls es $\varepsilon > 0$ gibt mit $U_\varepsilon(Q) \subseteq U$.

Dann sind für $Q \in X$ und eine Folge $P_k \in X$ offensichtlich die folgenden Aussagen äquivalent

1. $\forall \varepsilon > 0 \exists k_0 \forall k \geq k_0. d(P_k, Q) < \varepsilon$
2. $\forall \varepsilon > 0 \exists k_0 \forall k \geq k_0. P_k \in U_\varepsilon(Q)$
3. Für alle Umgebungen U von Q gilt $\exists k_0 \forall k \geq k_0. P_k \in U$
4. $d(P_k, Q) \rightarrow 0$

Trifft eine/jede der Bedingungen zu, so *konvergiert* die Folge P_k gegen den *Limes* Q (bzgl. d) und wir schreiben

$$Q = \lim_{k \rightarrow \infty} P_k \text{ bzw. } P_k \rightarrow Q$$

und dekorieren das ggF. noch mit der Metrik d oder einer dahinter stehenden Norm $\|\cdot\|$. Es gilt

$$P_k \rightarrow Q \text{ und } P_k \rightarrow Q' \Rightarrow Q = Q'$$

Ist nämlich $Q \neq Q'$, so wähle $\varepsilon = \frac{1}{2}d(Q, Q')$ - dann $U_\varepsilon(Q) \cap U_\varepsilon(Q') = \emptyset$.

1.6 Ortsvektoren

Sei (X, V) affiner Raum. Wählt man einen festen Punkt $O \in X$ (den man dann *Ursprung* nennt), so gibt es eine bijektive Abbildung

$$V \rightarrow X, \vec{p} \mapsto P = \vec{p} + O$$

zwischen *Ortsvektoren* und Punkten. Macht man (V, V) zum affinen Raum durch $(\vec{v}, \vec{p}) \mapsto \vec{v} + \vec{p}$, so wird (X, V) isomorph zu (V, V) . Ist $(V, \|\cdot\|)$ normiert, so

$$d(P, Q) = \|\vec{q} - \vec{p}\| =: d'(\vec{p}, \vec{q}) \text{ mit } \vec{p} + O, Q = \vec{q} + O$$

also ist $\vec{p} \mapsto \vec{p} + O$ eine *Isometrie* von (V, d') auf (X, d) . Insbesondere gilt

$$P_k \rightarrow Q \Leftrightarrow \vec{p}_k \rightarrow \vec{q}$$

1.7 Koordinaten

\mathbb{R}^n ist der \mathbb{R} -Vektorraum aller Spalten

$$\begin{pmatrix} x^{(1)} \\ \vdots \\ x^{(n)} \end{pmatrix}$$

mit

$$\begin{pmatrix} x^{(1)} \\ \vdots \\ x^{(n)} \end{pmatrix} + \begin{pmatrix} y^{(1)} \\ \vdots \\ y^{(n)} \end{pmatrix} = \begin{pmatrix} x^{(1)} + y^{(1)} \\ \vdots \\ x^{(n)} + y^{(n)} \end{pmatrix}, \quad r \begin{pmatrix} x^{(1)} \\ \vdots \\ x^{(n)} \end{pmatrix} = \begin{pmatrix} rx^{(1)} \\ \vdots \\ rx^{(n)} \end{pmatrix}$$

Sei $\alpha : \vec{e}_1, \dots, \vec{e}_n$ Basis des \mathbb{R} -Vektorraums V , d.h. jedes $\vec{x} \in V$ hat eindeutige Darstellung

$$\vec{x} = x^{(1)}\vec{e}_1 + \dots + x^{(n)}\vec{e}_n$$

Definiere die *Koordinaten (-spalte)* von \vec{x}

$$\vec{x}^\alpha = \begin{pmatrix} x^{(1)} \\ \vdots \\ x^{(n)} \end{pmatrix} \in \mathbb{R}^n$$

Dann gilt

$$(\vec{x} + \vec{y})^\alpha = \vec{x}^\alpha + \vec{y}^\alpha, \quad (r\vec{x})^\alpha = r(\vec{x}^\alpha)$$

also ist $\vec{x} \mapsto \vec{x}^\alpha$ ein Isomorphismus von V auf \mathbb{R}^n .

Ist (X, V) ein affiner Raum, so besteht ein *Koordinatensystem* α aus einem Ursprung O_α und einer Basis α . Dann sind $P^\alpha = \vec{p}^\alpha$ die *Koordinaten* von $P = \vec{p} + O_\alpha$.

1.8 Maximumnorm

Sei $\alpha : \vec{e}_1, \dots, \vec{e}_n$ Basis des \mathbb{R} -Vektorraums V . Die zugehörige *Maximumnorm* auf V ist definiert als

$$\|\vec{x}\|_\infty = \|\vec{x}_{\max} = \max\{|x^{(i)}| \mid i = 1, \dots, n\}$$

Satz 1.4 Sei $\|\cdot\|_\infty$ Maximumnorm bzgl. einer Basis α von V . Dann gilt

1. $\|\cdot\|_\infty$ ist eine Norm auf V
2. $\vec{x}_k \rightarrow \vec{p} \Leftrightarrow \forall i. x^{(i)} \rightarrow p^{(i)}$
3. (Bolzano-Weierstraß.) Gilt $\|\vec{x}_k\| \leq c$ für alle k , so hat die Folge \vec{x}_k eine konvergente Teilfolge.

Beweis. Zu 1. Dreiecksungleichung: Zu \vec{x}, \vec{y} gibt es i mit

$$\|\vec{x} + \vec{y}\|_\infty = |x^{(i)} + y^{(i)}| \leq |x^{(i)}| + |y^{(i)}| \leq \|\vec{x}\|_\infty + \|\vec{y}\|_\infty$$

Zu 2.

$$|x_k^{(i)} - p^{(i)}| \leq \|\vec{x}_k - \vec{p}\|_\infty \leq \sum_{j=1}^n |x_k^{(j)} - p^{(j)}|$$

also

$$\forall i. |x_k^{(i)} - p^{(i)}| \rightarrow 0 \Leftrightarrow \|\vec{x}_k - \vec{p}\|_\infty \rightarrow 0$$

Zu 3. Sei z.B. $n = 2$ und die Basis ist orthonormal. Die Folge liegt nach Voraussetzung in einem Quadrat W_1 der Seitenlänge c . Durch fortlaufendes Vierteln erhält man eine Folge von Quadraten $W_m \subseteq W_{m-1}$ der Seitenlänge $\frac{2}{2^m}$, die jeweils unendliche viele Folgenglieder enthalten. Man wähle $k_m > k_{m-1}$ so, dass $\vec{x}_{k_m} \in W_m$. Es folgt

$$|x_{k_l}^{(i)} - x_{k_h}^{(i)}| \leq \frac{2}{2^m} \quad \text{für alle } l, h \geq m$$

also ist $x_{k_m}^{(i)}$ für jedes i eine Cauchyfolge und hat Limes $p^{(i)}$. Nach 2. folgt $\vec{x}_{k_m} \rightarrow \vec{p}$. \square

2 Topologie endlichdimensionaler normierter Räume

Elemente eines Vektorraums schreiben wir nun einfach als $x \in V$ und dürfen die sowohl als Vektoren wie auch als Punkte eines affinen Raums verstehen.

2.1 Äquivalenz von Normen

Normen $\|\cdot\|_1$ und $\|\cdot\|_2$ auf V heißen *äquivalent*, falls es $a, b > 0$ gibt mit

$$\forall x \in V. \|x\|_1 \leq a\|x\|_2, \|x\|_2 \leq b\|x\|_1$$

Gleichwertig ist: es gibt $c, d > 0$ mit

$$c\|x\|_1 \leq \|x\|_2 \leq d\|x\|_1$$

Es handelt sich offenbar um eine Äquivalenzrelation auf der Menge aller Normen auf V .

Lemma 2.1 *Sind die Normen $\|\cdot\|_1$ und $\|\cdot\|_2$ auf V äquivalent, so gilt in $X = V$*

1. $U \subseteq X$ ist Umgebung von p bzgl. $\|\cdot\|_1$ genau dann, wenn U Umgebung von p bzgl. $\|\cdot\|_2$ ist
2. $x_k \rightarrow_1 p \Leftrightarrow x_k \rightarrow_2 p$

Beweis. Es gilt

$$\|p - x\|_1 < \varepsilon \Rightarrow \|p - x\|_2 < a\varepsilon$$

also

$$U_\varepsilon^1(p) \subseteq U_{a\varepsilon}^2(p)$$

Ist nun U eine Umgebung von p bzgl. $\|\cdot\|_2$, so gibt es $\eta > 0$ mit $U_\eta^2(p) \subseteq U$, also

$$U_\varepsilon^1(p) \subseteq U_\eta^2(p) \subseteq U \quad \text{mit } \varepsilon = \frac{\eta}{a}$$

Ebenso mit vertauschten Rollen. Das beweist 1. und 2. folgt sofort. \square

Satz 2.2 Auf einem endlichdimensionalen \mathbb{R} -Vektorraum sind alle Normen äquivalent.

Beweis. Jede Norm auf V ist zu Maximumnorm äquivalent - siehe S.5-6. (AnaII-mehrdim)

Dass etwas ‘‘für \mathbb{R}^n gilt’’, heisst also, dass es für jeden endlichdimensionalen normierten Vektorraum und zugehörigen metrischen affinen Raum gilt.

Korollar 2.3 Sei $(V, \|\cdot\|)$ normiert und α endliche Basis. Dann gilt

1. $x_k \rightarrow p \Leftrightarrow \forall i. x_k^{(i)} \rightarrow p^{(i)}$
2. (Bolzano-Weierstraß.) Jede bzgl. $\|\cdot\|$ beschränkte Folge hat eine konvergente Teilfolge.
3. x_k ist konvergent genau dann, wenn es Cauchyfolge ist, d.h.

$$\forall \varepsilon > 0 \exists k_0 \forall k, l \geq k_0. \|x_k - x_l\| < \varepsilon$$

Beweis. 1. folgt sofort mit dem Lemma. Zu 2. Ist x_k bzgl. $\|\cdot\|$ beschränkt, so auch bzgl. der Maximumnorm, hat also konvergente Teilfolge. Zu 3. Sei x_k Cauchyfolge. Wegen

$$|x_k^{(i)} - x_l^{(i)}| \leq \|x_k - x_l\|_\infty \leq b \|x_k - x_l\|$$

sind die $x_k^{(i)}$ Cauchyfolgen, also konvergent und nach 1. auch x_k konvergent bzgl. $\|\cdot\|$. Die Umkehrung wie in \mathbb{R} . \square

2.2 Topologische Grundbegriffe in metrischen Räumen

Sei (X, d) metrischer Raum und $M \subseteq X$. Definition und Sätzchen

1. x innerer Punkt von $M \Leftrightarrow \exists \varepsilon > 0. U_\varepsilon(x) \subseteq M \Leftrightarrow$ es gibt Umgebung U von x mit $U \subseteq M$
2. $M^\circ := \{x \in M \mid x \text{ innerer Punkt von } M\}$, das Innere oder der offene Kern von M
3. M offen $\Leftrightarrow M = M^\circ$
4. x Häufungspunkt (HP) von $M \Leftrightarrow \forall \varepsilon > 0. U_\varepsilon(x) \cap M \not\subseteq \{x\}$
 \Leftrightarrow für jede Umgebung U von x ist $U \cap M$ unendlich
5. $\overline{M} := M \cup \{x \mid x \text{ HP von } M\}$ abgeschlossene Hülle oder Abschluss
6. M abgeschlossen $\Leftrightarrow M = \overline{M} \Leftrightarrow X \setminus M$ offen
7. $\partial M := \overline{M} \setminus M^\circ$ Rand von M .
 $x \in \partial M \Leftrightarrow U \cap M \neq \emptyset$ und $U \cap X \setminus M \neq \emptyset$ für jede Umgebung U von x
8. Eine Familie U_i ($i \in I$) offener Mengen ist eine offene Überdeckung von M falls $M \subseteq \bigcup_{i \in I} U_i$
9. M ist kompakt, falls es zu jeder offenen Überdeckung U_i ($i \in I$) von M eine endliche Teilüberdeckung gibt, d.h. endliches $J \subseteq I$ mit $M \subseteq \bigcup_{j \in J} U_j$

2.3 Filtrierung in \mathbb{R}^n

Satz 2.4 Seien $\mathbb{R}^n \supseteq A_1 \supseteq A_2 \supseteq A_3 \dots$ abgeschlossen mit $\sup\{\|x - y\| \mid x, y \in A_k\} \rightarrow 0$. Dann gibt es x mit $\{x\} = \bigcap_{k=1}^{\infty} A_k$.

Bew. s.S. 10 (AnaII-mehrdim)

2.4 Kompaktheit in \mathbb{R}^n

Satz 2.5 Sei $M \subseteq \mathbb{R}^n$. Dann sind äquivalent

1. M ist abgeschlossen und beschränkt
2. M ist kompakt
3. Jede unendliche Teilmenge von M hat einen Häufungspunkt in M
4. Jede Folge in M eine Teilfolge mit Limes in M

Außerdem gilt der Zusatz: Jede unendliche beschränkte Menge hat einen Häufungspunkt.

Beweis. 1 \Rightarrow 2 siehe S. 11 (AnaII-mehrdim). Nur hier wird $M \subseteq \mathbb{R}^n$ benutzt. 2 \Rightarrow 3 s.S.12 (AnaII-mehrdim). 3 \Rightarrow 4: O.B.d.A. ist $\{x_k \mid k = 1, 2, \dots\}$ unendlich, hat also HP $x \in M$. Sind die k_l mit $l < m$ schon definiert, so setze $\varepsilon = \frac{1}{m}$ und $k_m > k_{m-1}$ mit $x_{k_m} \in U_\varepsilon(x)$. Das ergibt gegen x konvergente Teilfolge. 4 \Rightarrow 1 wie 3 \Rightarrow 1 auf S. 12 (AnaII-mehrdim). \square

2.5 Berechnung von Skalarprodukten

Sei (V, \cdot) euklidischer Vektorraum. Eine Basis e_1, \dots, e_n von V ist *Orthonormalbasis*, wenn

$$e_i \cdot e_j = 0 \text{ für } i \neq j, \quad e_i \cdot e_i = 1$$

(die Existenz für $\dim V < \infty$ folgt aus dem Verfahren von Gram-Schmidt). Es folgt sofort durch Ausmultiplizieren

$$x \cdot y = \sum_{i=1}^n x^{(i)} \cdot y^{(i)} \quad \text{für } x = \sum_{i=1}^n x^{(i)}, \quad y = \sum_{i=1}^n y^{(i)}$$

Für die zugehörige Norm hat man dann

$$|x| = \sum_{i=1}^n (x^{(i)})^2$$

Definiert man umgekehrt im Vektorraum \mathbb{R}^n der n -Spalten

$$x = \begin{pmatrix} x^{(1)} \\ \vdots \\ x^{(n)} \end{pmatrix}$$

ein Produkt auf diese Weise, so erhält man ein Skalarprodukt und die sog. *Standardbasis* oder *kanonische Basis* mit

$$e_i^{(i)} = 1, \quad e_i^{(j)} = 0 \text{ für } j \neq i$$

ist Orthonormalbasis.

2.6 Kompaktheit in metrischen Räumen

Satz 2.6 *Sei X metrischer Raum und $M \subseteq X$. Dann sind äquivalent*

2. M is kompakt
3. Jede unendliche Teilmenge von M hat einen Häufungspunkt in M
4. Jede Folge in M hat eine Teilfolge mit Limes in M

Jede kompakte Menge M ist beschränkt und abgeschlossen und es gelten

- (a) *Zu jedem $\delta > 0$ gibt es endliches $M_0 \subseteq M$ mit $M \subseteq \bigcup_{x \in M_0} U_\delta(x)$*
- (b) *Zu jeder offenen Überdeckung $U_i (i \in I)$ von M gibt es $\delta > 0$ so, dass es zu jedem $x \in M$ ein $i \in I$ gibt mit $U_\delta(x) \subseteq U_i$.*

Beweis. 2 \Rightarrow 3. Angenommen, $N \subseteq M$ und kein $x \in M$ ist HP von N , d.h. es gibt offene Umgebung U_x von x mit $U_x \cap N \subseteq \{x\}$. Die U_x bilden Überdeckung von M , also gibt es endliches $M_0 \subseteq M$ mit $M \subseteq \bigcup_{x \in M_0} U_x$. Es folgt $N \subseteq M_0$.

3 \Rightarrow 4: O.B.d.A. ist $\{x_k \mid k = 1, 2, \dots\}$ unendlich, hat also HP $x \in M$. Sind die k_l mit $l < m$ schon definiert, so setze $\varepsilon = \frac{1}{m}$ und $k_m > k_{m-1}$ mit $x_{k_m} \in U_\varepsilon(x)$. Das ergibt gegen x konvergente Teilfolge.

Sei 4. angenommen. Ist x HP von M , so gibt es $x_k \in M$ mit $x_k \rightarrow x$. Diese hat nach 4 gegen ein $y \in M$ konvergente Teilfolge. Also $x = y \in M$. Somit $M = \overline{M}$.

4 \Rightarrow (a) - (aus (a) folgt sofort die Beschränktheit). Angenommen, (a) gilt nicht. Also gibt es $\delta > 0$, so dass $M \not\subseteq \bigcup_{x \in N} U_\delta(x)$ für alle endlichen $N \subseteq M$. Mit dem Prinzip der Bedingten Auswahl erhalten wir eine Folge x_k mit

$$x_{k+1} \in M \setminus \bigcup_{i=1}^k U_\delta(x_i)$$

also $d(x_k, x_l) \geq \delta$ für $k \neq l$. Insbesondere ist keine Teilfolge Cauchy, also keine Teilfolge konvergent. Also gilt 4 nicht.

4 \Rightarrow (b): Angenommen 4 gilt, aber (b) gilt nicht. Insbesondere gibt es zu jedem $\delta = \frac{1}{n}$ ein $x_n \in M$ so, dass

$$(*) \quad U_{\frac{1}{n}}(x_n) \not\subseteq U_i \quad \text{für alle } i \in I$$

Wegen $U_{\frac{1}{n}}(p) \subseteq U_{\frac{1}{m}}(p)$ für $m \geq n$ gilt (*) auch für jede Teilfolge und wir haben o.B.d.A. $x_n \rightarrow x \in M$. Da $M \subseteq \bigcup_{i \in I} U_i$ gibt es $i \in I$ mit $x \in U_i$ und $\varepsilon > 0$ mit $U_\varepsilon(x) \subseteq U_i$ da U_i offen ist. Also

$$d(x_k, x) < \frac{\varepsilon}{2}, \quad U_{\frac{1}{k}}(x_k) \subseteq U_\varepsilon(x) \subseteq U_i \quad \text{für } k > \frac{2}{\varepsilon}$$

Widerspruch.

4 \Rightarrow 2. Wähle $\delta > 0$ nach (b) und dann endliches M_0 nach (a). Zu jedem $x \in M_0$ gibt es $i_x \in I$ mit $x \in U_{i_x}$ also

$$M \subseteq \bigcup_{x \in M_0} U_\delta(x) \subseteq \bigcup_{x \in M_0} U_{i_x} \quad \square$$

3 Lineare Abbildungen und Matrizen

3.1 Definition und Beispiele

Sind V, W zwei K -Vektorräume, so ist eine Abbildung $\phi : V \rightarrow W$ *linear* (genauer: K -linear), wenn folgende *Linearitätsbedingungen* gelten

$$\phi(\vec{x} + \vec{y}) = \phi(\vec{x}) + \phi(\vec{y}), \quad \phi(r\vec{x}) = r\phi(\vec{x}) \quad \text{für alle } \vec{x}, \vec{y} \text{ in } V, r \text{ in } K.$$

Ist $V = V'$, so ist ϕ_0 ein *Endomorphismus*.

Wir geben im Folgenden Beispiele linearer Abbildungen $\phi_0 : V \rightarrow V'$ an, wobei V, V' Vektor-Raum bzw. Ebene der anschaulichen Geometrie sind. Sind $\mathcal{P}, \mathcal{P}'$ die zugehörigen Punkträume und wählt man Ursprünge O bzw. O' , so kann man der linearen Abbildung ϕ_0 eine (affine) Abbildung zuordnen

$$\phi : \mathcal{P} \rightarrow \mathcal{P}'. \quad \phi(\vec{x} + O) = \phi_0(\vec{x}) + O'$$

Diese ist der Anschauung meist besser zugänglich und wird dementsprechend illustrativ benannt. In Beispielen und Aufgaben wird die gemeinte lineare Abbildung ϕ_0 häufig auf diese Weise (d.h. durch Hinweise auf das zugehörige ϕ) mitgeteilt und der Leser ist dann aufgefordert, die notwendige Präzisierung selbst zu leisten. In der Regel führt Identifikation von Punkten und Vektoren nach der Zauberformel "Ortsvektor" zu dem vom Aufgabensteller gewünschten Ergebnis. Eine präzise Behandlung der Zusammenhänge folgt hoffentlich in LA. Ebenso die Rechtfertigung von Bezeichnungen wie Drehung, Spiegelung usw. sowie der Nachweis der Linearität der zugehörigen vektoriiellen Abbildungen. Die Begriffe "Eigenraum" E_λ und "Eigenwert" (EV) werden in LAII behandelt.

1. Identische Abbildung $\phi = \text{id}_{\mathcal{P}}$ mit $\phi_0 = \text{id}_V$, Matrix E , $E_1 = V$. bijektiv
3. Punktspiegelung an O mit $\phi_0(\vec{x}) = -\vec{x}$. Matrix $-E$, $E_{-1} = V$, bijektiv
4. Zentrische Streckung an O um r mit $\phi_0(\vec{x}) = r\vec{x}$. Matrix rE , $E_r = V$, bijektiv falls $r \neq 0$, sonst Bild= 0.
5. Parallelprojektion mit Kern $K = \text{Kern}\phi_0$ auf $U + O$ mit $U = \text{Bild}\phi_0$. $E_1 = U$, $E_0 = K$. Bijektiv nur für $U = V$. Matrix bzgl. Basis $\vec{v}_1, \dots, \vec{v}_n$ mit $\vec{v}_1, \dots, \vec{v}_r$ Basis von U und $\vec{v}_{r+1}, \dots, \vec{v}_n$ Basis von K .

$$\begin{pmatrix} E_r & O \\ O & O \end{pmatrix}$$

Bei der Orthogonalprojektion ist $K = \{\vec{x} \mid \vec{x} \perp U\}$ und bzgl. ON-Basis gilt

$$\phi_0(\vec{x}) = \sum_{i=1}^r \langle \vec{v}_i \mid \vec{x} \rangle \vec{v}_i$$

Ist dabei U Hyperebene ($r = n - 1$) und $\vec{n} = \vec{v}_n$ Normalenvektor, so

$$\phi_0(\vec{x}) = \vec{x} - 2\langle \vec{n} \mid \vec{x} \rangle \vec{n}$$

6. Spiegelung an Hyperebene $U + O$ wie in 5. Achse=Normale. Bewegung.

$$\phi_0(\vec{x}) = \vec{x} - 2\langle \vec{n} \mid \vec{x} \rangle \vec{n}, \quad \text{Matrix } \begin{pmatrix} E_{n-1} & O \\ O & -1 \end{pmatrix}, \quad U = E_1, K\vec{n} = E_{-1}$$

7. Drehung in der reellen Ebene mit Zentrum O um Winkel ω - gegen die Uhr. Bewegung. Bzgl. ON-Basis Matrix

$$\begin{pmatrix} \cos \omega & -\sin \omega \\ \sin \omega & \cos \omega \end{pmatrix}$$

Keine reellen EV, komplexe EV von Betrag 1 und Argument $\pm\omega$.

8. Drehung im reellen Raum mit Achse $\mathbb{R}\vec{v}_1 + O$ um Winkel ω - positiv im Sinne der Rechten Hand. Bewegung. Bzgl. ON-Basis $\vec{v}_1, \vec{v}_2, \vec{v}_3$ Matrix

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \omega & -\sin \omega \\ 0 & \sin \omega & \cos \omega \end{pmatrix}, \quad E_1 = \mathbb{R}\vec{v}_1, \quad E_{-1} = E_1^\perp \Leftrightarrow \omega = \pi$$

9. Drehspiegelung im reellen Raum mit Achse $\mathbb{R}\vec{v}_1 + O$ um Winkel $\omega \neq \pi$ - positiv im Sinne der Rechten Hand. Bewegung. Bzgl. ON-Basis $\vec{v}_1, \vec{v}_2, \vec{v}_3$ Matrix

$$\begin{pmatrix} -1 & 0 & 0 \\ 0 & \cos \omega & -\sin \omega \\ 0 & \sin \omega & \cos \omega \end{pmatrix}, \quad E_{-1} = \mathbb{R}\vec{v}_1$$

10. Scherung in der Ebene längs der Achse $K\vec{v}_1 + 0$. Matrix bzgl. Basis \vec{v}_1, \vec{v}_2

$$\begin{pmatrix} 1 & r \\ 0 & 1 \end{pmatrix}, \quad E_1 = K\vec{v}_1$$

11. Zentralprojektion ϕ der Ebene $U + O$ auf die parallele Ebene $U + O'$ des Raumes mit Zentrum $Z \notin U + O \cup U + O'$. Dabei ist $\phi(P)$ der Schnittpunkt der Geraden durch PZ mit der Ebene $U + O'$. Hier $\phi_0 : U \rightarrow U$ und mit Normalenvektor \vec{n} von U

$$\phi_0(\vec{x}) = \frac{s}{r}\vec{x} \quad \text{wobei } r\vec{n} + Z \in U + O, \quad s\vec{n} + Z \in U + O'$$

3.2 Existenz und Eindeutigkeit

Lemma 3.1 *Ist $\phi : V \rightarrow W$ linear, so gilt*

$$\phi(\vec{0}) = \vec{0}, \quad \phi\left(\sum_i r_i \vec{v}_i\right) = \sum_i r_i \phi(\vec{v}_i)$$

Beweis. $\phi(\vec{0}_V) = \phi(0\vec{0}_V) = 0\phi(\vec{0}_V) = \vec{0}_W$ Die zweite Behauptung folgt leicht mit Induktion über n (und der Wohldefiniertheit von ϕ):

$$\phi\left(\sum_{i=1}^n r_i \vec{v}_i\right) = \phi\left(\sum_{i=1}^{n-1} r_i \vec{v}_i\right) + \phi(r_n \vec{v}_n) = \sum_{i=1}^{n-1} r_i \phi(\vec{v}_i) + r_n \phi(\vec{v}_n) = \sum_{i=1}^n r_i \phi(\vec{v}_i) \quad \square$$

Lemma 3.2 *Seien V, W Vektorräume über K , $\vec{v}_1, \dots, \vec{v}_n$ eine Basis von V und $\vec{w}_1, \dots, \vec{w}_n$ Vektoren in W . Dann gibt es genau eine lineare Abbildung $\phi : V \rightarrow W$ mit $\phi(\vec{v}_i) = \vec{w}_i$ für $i = 1, \dots, n$.*

Beweis. Definiere

$$\phi\left(\sum_i r_i \vec{v}_i\right) = \sum_i r_i \vec{w}_i \quad \text{für jede Wahl von } r_i \in K$$

Die Wohldefiniertheit folgt sofort aus der Eindeutigkeit der Darstellung $\sum_i r_i \vec{v}_i$ in V . Zur Linearität:

$$\begin{aligned} \phi\left(\sum_i r_i \vec{v}_i + \sum_i s_i \vec{v}_i\right) &= \phi\left(\sum_i (r_i + s_i) \vec{v}_i\right) = \sum_i (r_i + s_i) \phi(\vec{v}_i) \\ &= \sum_i r_i \phi(\vec{v}_i) + \sum_i s_i \phi(\vec{v}_i) = \phi\left(\sum_i r_i \vec{v}_i\right) + \phi\left(\sum_i s_i \vec{v}_i\right) \\ \phi\left(s \sum_i r_i \vec{v}_i\right) &= \phi\left(\sum_i s r_i \vec{v}_i\right) = \sum_i s r_i \phi(\vec{v}_i) = s \sum_i r_i \phi(\vec{v}_i) = s \phi\left(\sum_i r_i \vec{v}_i\right) \end{aligned}$$

Die Eindeutigkeit von ϕ folgt aus dem vorangehenden Lemma. \square

3.3 Isomorphie

Eine bijektive lineare Abbildung $\phi : V \rightarrow W$ heisst auch ein *Isomorphismus*. Gibt es einen Isomorphismus von V auf W , so heissen V und W isomorph und man schreibt $V \cong W$. Insbesondere ist dann ϕ^{-1} ein Isomorphismus von W auf V .

Bespiele: Der Anschauungsraum ist isomorph zu \mathbb{R}^3 . Der Zeilenraum K^{n*} ist isomorph zum Spaltenraum K^n . Ist σ eine Permutation von $\{1, \dots, m\}$, so erhält man einen Isomorphismus von K^m auf K^m durch

$$\begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix} \mapsto \begin{pmatrix} x_{\sigma 1} \\ \vdots \\ x_{\sigma m} \end{pmatrix}.$$

Bei einem Isomorphismus übertragen sich alle Eigenschaften analog auf die Bilder.

Korollar 3.3 Seien V, W Vektorräume über K , $\vec{v}_1, \dots, \vec{v}_n$ eine Basis von V und $\vec{w}_1, \dots, \vec{w}_n$ eine Basis von W . Dann gibt es genau einen Isomorphismus $\phi : V \rightarrow W$ mit $\phi(\vec{v}_i) = \vec{w}_i$ für $i = 1, \dots, n$.

Beweis. Wir wissen schon, dass es eine lineare Abbildung ϕ gibt. Ist $w \in W$ so hat man eine Darstellung $w = \sum_i r_i \vec{w}_i$ also $w = \phi\left(\sum_i r_i \vec{v}_i\right)$. Also ist $\phi : V \rightarrow W$ surjektiv. Gelte $\phi(\vec{v}) = \phi(\vec{v}')$. Da die $\vec{v}_1, \dots, \vec{v}_n$ Basis sind hat man Linearkombinationen $\vec{v} = \sum_i r^i \vec{v}_i$ und $\vec{v}' = \sum_i r'^i \vec{v}_i$, und es folgt $\sum_i r_i \vec{w}_i = \phi(\vec{v}) = \phi(\vec{v}') = \sum_i r'^i \vec{w}_i$. Dann $r_i = r'^i$, da die $\vec{w}_1, \dots, \vec{w}_n$ Basis sind. Das ergibt $\vec{v} = \vec{v}'$ und die Injektivität von ϕ . \square

Korollar 3.4 Ist $\alpha : \vec{v}_1, \dots, \vec{v}_m$ Basis des K -Vektorraumes V , so hat man den (Koordinaten-) Isomorphismus κ von V auf K^m

$$\vec{x} = \sum x_i \vec{v}_i \mapsto \kappa(\vec{x}) = \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix} = \vec{x}^\alpha$$

3.4 Matrizen

Sei K oder ein anderer Körper. Eine $m \times n$ -Matrix

$$A = (a_{ij}) = (a_{ij})_{m \times n} = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \vdots & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix}$$

über K wird angegeben durch ein (rechteckiges) Schema von Zahlen (*Koeffizienten, Einträgen*) a_{ij} aus K , wobei $i = 1, \dots, m; j = 1, \dots, n$. Oder eine mit den Paaren ij indizierte Liste. Die Einträge a_{i1}, \dots, a_{in} bilden die i -te *Zeile*, die Einträge a_{1j}, \dots, a_{mj} die j -te *Spalte*. Zeilen bzw. Spalten können demnach als Spezialfälle von Matrizen aufgefasst werden. Man lese: n geht nach m -Matrix oder m kreuz n Matrix.

3.5 Summe

Hat man eine weitere $m \times n$ -Matrix $B = (b_{ij})$, so definiert man die Summe

$$A + B = (a_{ij} + b_{ij})_{m \times n}, \quad rA = (ra_{ij})_{m \times n}.$$

Die $m \times n$ -Nullmatrix O hat nur Einträge 0. Es gelten die Regeln (V1-8) entsprechend.

3.6 Matrix mal Spalte

$$\begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix} = \begin{pmatrix} a_{11}x_1 + \dots + a_{1n}x_n \\ \vdots \\ a_{m1}x_1 + \dots + a_{mn}x_n \end{pmatrix} = A\mathbf{x} \text{ in } K^m \text{ für } \mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \text{ in } K^n$$

d.h. die i -te Komponente von Ax erhält man, indem man die i -te Zeile von A komponentenweise mit x multipliziert und dann aufaddiert. Man lese: auf x A angewendet oder A mal x .

	x_1	\dots	x_n
a_{11} \dots a_{1n}	$a_{11}x_1 +$	\dots	$+a_{1n}x_n$
\vdots \vdots \vdots		\vdots	
a_{m1} \dots a_{mn}	$a_{m1}x_1 +$	\dots	$+a_{mn}x_n$

Aus den Regeln (K1-10) für das Rechnen in Körpern erhält man sofort

$$(M1) A(\mathbf{b} + \mathbf{c}) = A\mathbf{b} + A\mathbf{c} \quad (M2) A(r\mathbf{b}) = r(A\mathbf{b}), \quad A\mathbf{0} = \mathbf{0}$$

Bemerkung. Für das zweite Gesetz braucht man das Kommutativgesetz (K10). Das könnte man vermeiden, indem man Matrizen und Skalare auf verschiedene Seiten schreibt, z.B. Abr oder, wenn man von links nach rechts schreibt, $r\mathbf{b}A$ mit \mathbf{b} als Zeile. Dabei erkennt man, dass es nur um das Assoziativgesetz (K5) geht.

3.7 Matrixbeschreibung

Ist $A \in K^{m \times n}$, so ist $x \mapsto Ax$ wegen (M1) und (M2) eine lineare Abbildung von K^n in K^m .

Satz 3.5 Seien V, W K -Vektorräume mit Basen $\alpha : \vec{e}_1, \dots, \vec{e}_n$ bzw. $\beta : \vec{f}_1, \dots, \vec{f}_m$. Dann gibt es eine bijektive Entsprechung zwischen linearen Abbildungen $\phi : V \rightarrow W$ und Matrizen $A \in K^{m \times n}$ vermöge

$$\phi(\vec{x})^\beta = A\vec{x}^\alpha, \quad A = {}_\beta\phi_\alpha \text{ heisst die Matrix von } \phi \text{ bzgl. } \alpha, \beta$$

In den Spalten der Matrix stehen die Koordinaten der Bilder der Basisvektoren

Ist $V = W$ und betrachten wir nur eine Basis so schreiben wir

$$\phi_\alpha = {}_\alpha\phi_\alpha \quad \text{Matrix von } \phi \text{ bzgl. } \alpha$$

Beweis. Ist ϕ gegeben und soll es ein A geben, so muss es das angegebene sein: man setze $\vec{x} = \vec{e}_j$ ein, d.h. $\vec{e}_j^\alpha = \varepsilon_j$ und $A\varepsilon_j$ ist die j -te Spalte von A . Es folgt

$$\phi(\vec{x})^\beta = \sum_j x_j \phi(\vec{e}_j)^\beta = \sum_j x_j A\varepsilon_j = A\left(\sum_j x_j \varepsilon_j\right) = A\vec{x}^\alpha \quad \text{für } \vec{x} = \sum_j x_j \vec{e}_j$$

Ist A gegeben, so erhält man die Linearität von ϕ indem man die Isomorphieeigenschaft der Koordinatenzuordnung benutzt

$$\phi(\vec{x} + \vec{y})^\beta = A(\vec{x} + \vec{y})^\beta = A(\vec{x}^\beta + \vec{y}^\beta) = A\vec{x}^\beta + A\vec{y}^\beta = \phi(\vec{x})^\beta + \phi(\vec{y})^\beta = (\phi(\vec{x}) + \phi(\vec{y}))^\beta$$

$$\text{also } \phi(\vec{x} + \vec{y}) = \phi(\vec{x}) + \phi(\vec{y})$$

$$\phi(r\vec{x})^\beta = A(r\vec{x})^\beta = Ar\vec{x}^\beta = rA\vec{x}^\beta = r\phi(\vec{x})^\beta = (r\phi(\vec{x}))^\beta \quad \text{also } \phi(r\vec{x}) = r\phi(\vec{x})$$

3.8 Matrizenprodukt

Wir definieren das Produkt einer $l \times m$ -Matrix A mit einer $m \times n$ -Matrix B über K als die $l \times n$ -Matrix $D = AB$

$$(a_{ij})_{l \times m} (b_{jk})_{m \times n} = (d_{ik})_{l \times n} = (a_{i1}b_{1k} + \dots + a_{im}b_{mk})_{l \times n}$$

d.h. man rechnet i -te Zeile von A komponentenweise mal k -te Spalte von B und addiert auf, um d_{ik} zu erhalten. Daher müssen die Zeilen von A genauso lang sein wie die Spalten von B . Hat B die Spalten $\mathbf{b}_1, \dots, \mathbf{b}_n$, so hat AB die Spalten $A\mathbf{b}_1, \dots, A\mathbf{b}_n$. Man lese: erst B dann A oder A nach B oder A mal B .

		a_{11}	\dots	a_{1n}	
		\vdots	\vdots	\vdots	
		a_{m1}	\dots	a_{mn}	
a_{11}	\dots	a_{1m}	$a_{11}a_{11} + \dots + a_{1m}a_{m1}$	\dots	$a_{11}a_{1n} + \dots + a_{1m}a_{mn}$
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots
a_{l1}	\dots	a_{lm}	$a_{l1}a_{11} + \dots + a_{lm}a_{m1}$	\dots	$a_{l1}a_{1n} + \dots + a_{lm}a_{mn}$

3.9 Gesetze

Die $n \times n$ -Einheitsmatrix E_n hat Diagonale $e_{ii} = 1$ und sonst überall Null. Es gelten dann die folgenden Regeln - soweit die Operationen für die Matrizen überhaupt ausführbar sind

$$(M3) \quad B(C + C') = BC + BC' \quad (B + B')C = BC + B'C$$

$$(M4) \quad (rB)C = r(BC) = B(rC), \quad (M5) \quad A(BC) = (AB)C$$

$$(M6) \quad E_m A = A = A E_n.$$

Beweis durch Nachrechnen mit Körperaxiomen, wobei man, wenn es geht, nur den Spezialfall betrachtet, dass man Matrix mit Spalte(n) multipliziert. Warnend sei vermerkt, dass für Matrizen im Allgemeinen $AB \neq BA$, und dass man aus $AB = O$ nicht auf $A = O$ oder $B = O$ schliessen kann.

(M3) und (M4) folgen leicht aus (M1) und (M2). Zu (M5) betrachten wir zuerst den Spezialfall, dass $C = \mathbf{c}$ eine Spalte ist. Sei $l \times m$ -Matrix $A = (b_{ij})$ und $m \times n$ -Matrix $B = (b_{jk})$ und $\mathbf{c} \in K^n$. Dann

$$\begin{aligned} A(B\mathbf{c}) &= A \begin{pmatrix} b_{11}c_1 + \dots + b_{1n}c_n \\ \vdots \\ b_{m1}c_1 + \dots + b_{mn}c_n \end{pmatrix} \\ &= \begin{pmatrix} a_{11}(b_{11}c_1 + \dots + b_{1n}c_n) + \dots + a_{1m}(b_{m1}c_1 + \dots + b_{mn}c_n) \\ \vdots \\ a_{l1}(b_{11}c_1 + \dots + b_{1n}c_n) + \dots + a_{lm}(b_{m1}c_1 + \dots + b_{mn}c_n) \end{pmatrix} \\ &= \begin{pmatrix} (a_{11}b_{11} + \dots + a_{1m}b_{m1})c_1 + \dots + (a_{11}b_{1n} + \dots + a_{1m}b_{mn})c_n \\ \vdots \\ (a_{l1}b_{11} + \dots + a_{lm}b_{m1})c_1 + \dots + (a_{l1}b_{1n} + \dots + a_{lm}b_{mn})c_n \end{pmatrix}. \end{aligned}$$

Also

$$A(B\mathbf{c}) = (AB)\mathbf{c}$$

3.10 Komposition

Sind V, W, U K -Vektorräume mit endlichen Basen α, β, γ und ϕ lineare Abbildung von V nach W , ψ von W nach U , so ist $\phi \circ \psi$ lineare Abbildung von V nach U und die Matrix ergibt sich als Produkt der Matrizen

$$\boxed{\gamma[\psi \circ \phi]_{\alpha} = \gamma\psi_{\beta} \cdot \beta\phi_{\alpha}}$$

Beweis. $(\psi(\phi(\vec{x})))^{\gamma} = \gamma\psi_{\beta}\phi(\vec{x})^{\beta} = \gamma\psi_{\beta}\beta\phi_{\alpha}\vec{x}^{\alpha}$.

3.11 Inverse Matrix

Lemma 3.6 Für zwei $n \times n$ -Matrizen A, B über K sind die folgenden Aussagen äquivalent

$$(1) \quad AB = E = BA$$

$$(2) \quad \text{für alle } \mathbf{x}, \mathbf{u} \text{ in } K^n \text{ gilt: } \mathbf{u} = A\mathbf{x} \Leftrightarrow \mathbf{x} = B\mathbf{u}.$$

Wenn zu A ein solches B existiert heisst A invertierbar, regulär oder nichtsingulär. B ist dann durch A eindeutig bestimmt und heisst die Inverse zu A . Man schreibt $B = A^{-1}$.

Beweis. Gelte (1). Wenn $\mathbf{x} = B\mathbf{u}$ dann $A\mathbf{x} = A(B\mathbf{u}) = (AB)\mathbf{u} = E\mathbf{u} = \mathbf{u}$. Ebenso $\mathbf{x} = B\mathbf{u}$ aus $\mathbf{u} = A\mathbf{x}$. Gelte (2). Seien \mathbf{b}_j und \mathbf{e}_j die j -ten Spalten von B bzw. E . Es gilt $\mathbf{b}_j = B\mathbf{e}_j$, also $\mathbf{e}_j = A\mathbf{b}_j$, und folglich $AB = E$. Ebenso $BA = E$. Sei nun auch $AC = E = CA$. Dann $C = CE = C(AB) = (CA)B = EB = B$. \square

Für das Rechnen mit Inversen gelten die folgenden Regeln. Für $n \times n$ -Matrizen gilt:

$$E^{-1} = E; \quad A \text{ invertierbar} \Rightarrow A^{-1} \text{ invertierbar und } (A^{-1})^{-1} = A$$

$$A, B \text{ invertierbar} \Rightarrow AB \text{ invertierbar und } (AB)^{-1} = B^{-1}A^{-1}.$$

Beweis. $B^{-1}A^{-1}AB = B^{-1}B = E$.

3.12 Transponierte Matrix

Zu einer $m \times n$ -Matrix $A = (a_{ij})_{m \times n}$ erhält man die *transponierte* $n \times m$ -Matrix A^t , indem man die Einträge an der Diagonalen "spiegelt":

$$A^t = (a'_{kl})_{n \times m} \quad \text{mit } a'_{kl} = a_{lk}.$$

Es gelten dann die Regeln

$$A^{Tt} = A, \quad (A + B)^t = A^t + B^t, \quad (rA)^t = rA^t, \quad (AB)^t = B^tA^t, \quad E^t = E.$$

$$A \text{ invertierbar} \Rightarrow A^t \text{ invertierbar und } (A^t)^{-1} = (A^{-1})^t.$$

Beweis. $(A^{-1})^tA^t = (AA^{-1})^t = E^t = E$. \square .

3.13 Koordinatentransformation für Vektoren

Seien $\alpha : \vec{v}_1, \dots, \vec{v}_m$ und $\beta : \vec{w}_1, \dots, \vec{w}_m$ zwei Basen des K -Vektorraums V . Wie rechnet man die Koordinaten \vec{x}^α und \vec{x}^β ineinander um? Wir machen den Ansatz

$$\vec{v}^\alpha = {}^\alpha T_\beta \vec{v}^\beta \quad \text{für alle } \vec{v}$$

Die Transformationsmatrix ist dadurch eindeutig bestimmt: wir erhalten ihre Spalten als ${}^\alpha T_\beta \mathbf{e}_j$ indem wir für \vec{v} die Vektoren aus β einsetzen.

Es sei also $S = {}^\alpha T_\beta$ definiert als die *Transformationsmatrix* mit den Koordinatenspalten \vec{w}_j^α der \vec{w}_j bzgl α als Spalten, also

$$S = {}^\alpha T_\beta = (\vec{w}_1^\alpha \dots \vec{w}_m^\alpha), \quad \vec{w}_j = s_{1j}\vec{v}_1 + \dots + s_{mj}\vec{v}_m$$

\vec{a}_2

$$\begin{aligned}\vec{v} &= 2\vec{b}_1 + 1\vec{b}_2 \\ &= 3\vec{a}_1 + 5\vec{a}_2\end{aligned}$$

$$\vec{b}_2 = -1\vec{a}_1 + 3\vec{a}_2$$

$$\vec{b}_1 = 2\vec{a}_1 + 1\vec{a}_2$$

 \vec{a}_1

$$\boxed{\vec{x}^\alpha = {}^\alpha T_\beta \vec{x}^\beta, \quad \vec{x}^\beta = {}^\beta T_\alpha \vec{x}^\alpha \quad \text{mit} \quad {}^\beta T_\alpha = {}^\alpha T_\beta^{-1}}$$

Die Transformationsmatrix $S = {}^\alpha T_\beta$ leistet die Koordinatenumrechnung von der ‘neuen’ Basis β in die ‘alte’ Basis α . In ihren Spalten stehen die Koordinaten der neuen Basisvektoren bzgl. der alten Basis. Für die Umrechnung von ‘alt’ auf ‘neu’ benutzt man die inverse Matrix $S^{-1} = {}^\beta T_\alpha$

Beweis.

$$\vec{x} = \sum_j y_j \vec{w}_j = \sum_j y_j \sum_i s_{ij} \vec{v}_i = \sum_{i,j} s_{ij} y_j \vec{v}_i = \sum_i \left(\sum_j s_{ij} y_j \right) \vec{v}_i$$

Also $\vec{x}^\alpha = {}^\alpha T_\beta \vec{x}^\beta$. Durch Vertauschen der Rollen $\vec{x}^\beta = {}^\beta T_\alpha \vec{x}^\alpha$, also $\vec{x}^\alpha = {}^\alpha T_\beta {}^\beta T_\alpha \vec{x}^\alpha$. Das gilt insbesondere für die $\vec{x}^\alpha = \mathbf{e}_j$, also $E = {}^\alpha T_\beta {}^\beta T_\alpha E$. \square

Ist eine Basis α des m -dimensionalen K -Vektorraumes V gegeben, so kann man jede invertierbare $m \times m$ -Matrix S auf genau eine Weise als Transformationsmatrix $S = {}^\alpha T_\beta$ auffassen, nämlich mit der Basis β deren Koordinaten bzgl. α durch die Spalten von S gegeben sind.

3.14 Orthogonale Transformationsmatrizen

Lemma 3.7 *Ist α Orthonormalbasis eines euklidischen Vektorraumes, so ist die Transformationsmatrix ${}^\alpha T_\beta$ genau dann orthogonal, wenn β ebenfalls Orthonormalbasis ist*

Beweis. Sind die \vec{f}_j die Vektoren von β und berechnet man das Skalarprodukt in den Koordinaten bzgl. α , so

$${}^\alpha T_\beta^t {}^\alpha T_\beta = (\langle \vec{f}_i | \vec{f}_j \rangle)_{1 \leq i, j \leq n}$$

Und β ist Orthonormalbasis genau dann, wenn letztere Matrix E ist. \square

3.15 Transformation von Matrixbeschreibungen

Sind α, β Basen von V und γ, δ Basen von W und $\phi : V \rightarrow W$ linear so gilt

$$\delta\phi_\beta = {}^\delta T_\gamma \gamma\phi_\alpha {}^\alpha T_\beta = {}^\gamma T_\delta^{-1} \gamma\phi_\alpha {}^\alpha T_\beta$$

und, falls $V = W, \alpha = \gamma, \beta = \delta$,

$${}^\beta\phi_\beta = {}^\alpha T_\beta^{-1} \alpha\phi_\alpha {}^\alpha T_\beta$$

Beweis. ${}^\delta\phi_\beta \vec{x}^\beta = (\phi(\vec{x}))^\delta = {}^\delta T_\gamma (\phi(\vec{x}))^\gamma = {}^\delta T_\gamma \gamma\phi_\alpha \vec{x}^\alpha = {}^\delta T_\gamma \gamma\phi_\alpha {}^\alpha T_\beta \vec{x}^\beta$

Für den Spezialfall $W = K$, d.h. für Linearform $\phi : V \rightarrow K$ betrachtet man die Basis $\gamma = \delta$ gegeben durch $1 \in K$. Die Matrix bzgl. Basis α von V ist dann eine Zeile

$$\phi_\alpha$$

Somit haben wir die (*kovariante*) Transformation

$$\phi_\beta = \phi_\alpha {}^\alpha T_\beta$$

Beispiel: Sei V \mathbb{R} -Vektorraum mit Basen

$$\alpha : \vec{e}_1, \vec{e}_2, \quad \beta : \vec{e}_1 + \vec{e}_2, \vec{e}_1 - \vec{e}_2$$

Sei $f : V \rightarrow \mathbb{R}$ und $p \in V$ definiert durch

$$f(x_1\vec{e}_1 + x_2\vec{e}_2) = x_1x_2, \quad p = 2\vec{e}_1 + \vec{e}_2$$

Dann gilt

$$df_\alpha(p) = J_{\alpha,f}(p) = (1 \quad 2)$$

Es folgt

$$df_\beta(p) = J_{\beta,f}(p) = df_\alpha(p) \cdot {}^\alpha T_\beta = (1 \quad 2) \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} = (3 \quad -1)$$

4 Stetige Abbildungen

4.0 Abbildungen

Eine *Abbildung* oder *Funktion* f einer Menge X in eine Menge Y ist eine Teilmenge von $X \times Y$ so, dass gilt

- Sind $(x, y) \in f$ und $(x, y') \in f$ so $y = y'$ (*wohldefiniert*)
- Zu jedem $x \in X$ gibt es y mit $(x, y) \in f$ (*überall definiert*)

Man schreibt: $f : X \rightarrow Y$. Die Wohldefiniertheit erlaubt die Schreibweise

$$f(x) = y \Leftrightarrow (x, y) \in f$$

Dann kann man

$$f = \{(x, y) \in X \times Y \mid f(x) = y\}$$

als den *Graphen* zu der ‘‘Zuordnungsvorschrift’’ f verstehen.

Beispiele: $X = V$ Vektorraum.

1. $f : V \rightarrow V$ mit $f(x) = rx$ wobei $r \in \mathbb{R}$ fest
2. $f : V \times V \rightarrow V$ mit $f((x, y)) = x + y$
3. V normiert, $f : V \rightarrow \mathbb{R}$ mit $f(x) = \|rx\|$ wobei $r \in \mathbb{R}$ fest.
4. (V, \cdot) euklidisch, $f : V \times V \rightarrow \mathbb{R}$ mit $f((x, y)) = x \cdot y$
5. (V, \cdot) euklidisch, $\dim V = 3$ und eine Orientierung von V gegeben. $f : V \times V \rightarrow V$ mit $f((x, y)) = x \times y$ (Vektorprodukt)

Weitere Beispiele s. S.13-15(AnaII-mehrdim) .

Ist f nicht überall definiert, so kann man einen *maximalen Definitionsbereich* $D_{\max} = \{x \in X \mid \exists y (x, y) \in f\}$ und kann jedes $D \subseteq D_{\max}$ als Definitionsbereich der Abbildung $f|_D : D \rightarrow Y$ auswählen. Man schreibt dann einfach $f : D \rightarrow Y$.

Sei $Y = V$ Vektorraum. Dann ist $f : X \rightarrow V$ ein *Vektorfeld*, Ist e_1, \dots, e_m Basis von V , so ist eine Abbildung $f : X \rightarrow Y$ eindeutig bestimmt durch ihre *Komponenten*

$$y = f(x) \Leftrightarrow f_i(x) = y_i (i = 1, \dots, m) \Leftrightarrow f(x) = \sum_i y_i e_i$$

d.h.

$$f(x) \hat{=} \begin{pmatrix} f_1(x) \\ \vdots \\ f_m(x) \end{pmatrix}$$

Ist hier $X = \mathbb{R}$ so kann man f als ‘‘Kurve’’ im affinen Raum zu V' , Vektorfeld längs einer Kurve verstehen oder als vektorwertige Funktion der Zeit verstehen.

Sei V ein \mathbb{R} -Vektorraum und $X = D \subseteq V$. Ist $\alpha : v_1, \dots, v_n$ eine Basis von V . so hat man die *Koordinatenfunktionen*

$$\kappa_i : D_i \rightarrow D \quad \kappa_i(r) = rv_i \text{ falls } rv_i \in D$$

die sind offenbar injektiv. Umgekehrt kann man $x \in D$ mit seiner Koordinatenspalte $x^\alpha \in \mathbb{R}^n$ identifizieren und X als Teilmenge von \mathbb{R}^n auffassen. Ist nun auch $Y = \mathbb{R}$ so hat man eine *reelle Funktion* oder *Skalarenfeld* in n reellen Variablen

$$f : D \rightarrow \mathbb{R}, \quad D \subseteq \mathbb{R}^n, \quad y = f(x_1, \dots, x_n) = f(x), \quad x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in D$$

Den Graphen

$$f = \{(x, f(x)) \mid x \in D\} \hat{=} \{(x_1, \dots, x_n, x_{n+1}) \mid x_{n+1} = f(x_1, \dots, x_n)\}$$

kann man nun als *Hyperfläche* in \mathbb{R}^{n+1} verstehen. Lässt man nur die Variable x_i variieren und setzt $x_j = c_j$, $j \neq i$, konstant, so erhält man eine Funktion

$$\{x_i \in \mathbb{R} \mid (c_1, \dots, x_i, \dots, c_n) \in M\} \rightarrow \mathbb{R}, \quad (c_1, \dots, x_i, \dots, c_n) \mapsto f(c_1, \dots, x_i, \dots, c_n)$$

Für $n = 2$ kann man den Graphen dieser Funktion als Schnitt des Graphen von f mit einer zur $x_i - y$ -Ebene parallelen Ebene verstehen.

Andererseits hat man zu jedem $c \in \mathbb{R}$ die *Niveaumenge* (im Falle $n = 2$ *Höhenlinie*)

$$\{x \in D \mid f(x) = c\}$$

Kennt man zu jedem $c \in \mathbb{R}$ die Niveaumenge von f , so kennt man f .

4.1 Lipschitzstetige und stetige Abbildungen in metrischen Räumen

Seien (X, d) und (X', d') metrische Räume und $f : X \rightarrow X'$ eine Abbildung

- f ist *Lipshitz-stetig* an $p \in X$ wenn es eine Konstante L gibt, so dass gilt

$$d'(f(p), f(x)) \leq L \cdot d(p, x) \quad \text{für alle } x \in X$$

- f ist *stetig* an $p \in X$, falls es zu jeder Umgebung U von $f(p)$ eine Umgebung V von p gibt mit $f(V) \subseteq U$.
- f ist *stetig* bzw. *Lipshitz-stetig*, wenn es das an jedem $p \in X$ ist - mit derselben Konstante L .
- Stammen die metrischen Räume von endlichdimensionalen normierten Räumen ab, so hängt es nicht von den Normen ab, ob f stetig bzw. Lipshitz-stetig ist (welche Lipshitz Konstanten man nehmen kann, hängt dagegen von der Norm ab).

Die Beispiele 1-5 sind alle Lipshitz-stetig, falls V endlichdimensionaler normierter Vektorraum. $V \times V$ wird normierter Vektorraum mit $\|(x, y)\| = \max\{|x|, |y|\}$ - Beweis wie bei Maximumnorm.

Lemma 4.1 *Die folgenden Aussagen sind äquivalent*

1. f ist an p stetig
2. $\forall \varepsilon > 0 \exists \delta > 0. f(U_\delta(p)) \subseteq U_\varepsilon(f(p))$
3. $\forall \varepsilon > 0 \exists \delta > 0 \forall x. d(x, p) < \delta \Rightarrow d'(f(x), f(p)) < \varepsilon$
4. Für alle Folgen $x_k \rightarrow p$ gilt $f(x_k) \rightarrow f(p)$.

Ist f an p Lipshitz-stetig so auch stetig. f ist stetig genau dann, wenn das Urbild $f^{-1}(U)$ jeder offenen (abgeschlossenen) Menge U in (X', d') in (X, d) offen (abgeschlossen) ist.

Beweis. $1 \Rightarrow 2$: Zu $U = U_\varepsilon(f(p))$ gibt es V mit $f(V) \subseteq U$; wähle δ mit $U_\delta(p) \subseteq V$. $2 \Rightarrow 1$: Zu U gibt es $U_\varepsilon(f(p)) \subseteq U$; wähle $V = U_\delta(p)$. $2 \Rightarrow 3$ ist trivial. $3 \Rightarrow 4$: Gelte $x_k \rightarrow p$. Zu $\varepsilon > 0$ bestimme δ nach 3 und dann k_0 so, dass $d(x_k, p) < \delta$ für alle $k \geq k_0$. Dann $d'(f(x_k), f(p)) < \varepsilon$. $4 \Rightarrow 3$: angenommen, 3 gilt nicht. Dann gibt es $\varepsilon > 0$ so, dass es zu jedem $\delta = \frac{1}{n}$ ein x_n gibt mit $d(x_n, p) < \delta$ aber $d'(f(x_n), f(p)) \geq \varepsilon$. Dann $x_n \rightarrow p$ aber $f(x_n) \not\rightarrow f(p)$.

Hat man Lipschitz-Konstante L , so $\delta := \frac{1}{L}\varepsilon$. Ist f stetig und U offen in X' , so gibt es zu $p \in f^{-1}(U)$, d.h. $f(p) \in U$, ein ε mit $U_\varepsilon(f(p)) \subseteq U$, also nach 2 ein δ mit $f(U_\delta(p)) \subseteq U$ also $(U_\delta(p)) \subseteq f^{-1}(U)$. Also ist $f^{-1}(U)$ offen. Umgekehrt folgt Stetigkeit sofort nach Def. Die Aussage über abgeschlossenen Mengen folgt aus $f^{-1}(X' \setminus M) = X \setminus f^{-1}(M)$ und daraus, dass die abgeschlossenen Mengen gerade die Komplemente offener Mengen sind. \square

Korollar 4.2 Seien (X, d) , (X', d') und (X'', d'') metrische Räume und $f : X \rightarrow X'$ und $g : X' \rightarrow X''$ Abbildungen.

1. Ist f an p und g an $f(p)$ stetig, so ist auch die Komposition $g \circ f$ an p stetig.
2. Sind f und g stetig, so auch $g \circ f$.

Beweis. Sei U Umgebung von $g(f(p))$. Dann gibt es Umgebung V von $f(p)$ mit $g(V) \subseteq U$ und Umgebung W von p mit $f(W) \subseteq V$. Also $(g \circ f)(W) \subseteq U$. \square

4.2 Relativtopologie

Ist (X, d) ein metrischer Raum und $D \subseteq X$, so ist auch D ein metrischer Raum mit der auf D eingeschränkten Metrik

$$d|_D(x, y) = d(x, y) \quad x, y, \in D$$

Korollar 4.3 1. U ist offen in $(D, d|_D)$ (relativ offen in D bzgl. (X, d)) genau dann, wenn es offenes V in (X, d) gibt mit $U = D \cap V$.

2. U ist Umgebung von p in $(D, d|_D)$ genau dann, wenn es Umgebung V von p in (X, d) gibt mit $D \cap V \subseteq U$.
3. $M \subseteq D$ ist kompakte Teilmenge von (X, d) genau dann, wenn M in (X, d) kompakte Teilmenge von $(D, d|_D)$ ist - im Falle $M = X$ ist (X, d) kompakter Raum
4. f ist stetige Abbildung von $(D, d|_D)$ in (X', d') genau dann, wenn zu jedem offenen U in (X', d') das Urbild $f^{-1}(U)$ relativ offen in D bzgl. (X, d) ist.

Beweis:

$$U_\varepsilon^{(D, d|_D)}(p) = U_\varepsilon^{(X, d)}(p) \cap D$$

Ist U offen in $(D, d|_D)$, so hat man zu jedem $p \in U$ ein ε und eine ε -Umgebung $U(p)$ in (X, d) mit $U(p) \cap D \subseteq U$. Die Vereinigung V dieser $U(p)$ ist offen in (X, d) und $D \cap V = U$. \square

4.3 Funktionenlimes

Ist p ein Häufungspunkt von D so ist q Limes von $f : D \rightarrow X'$ an der Stelle p

$$q = \lim_{x \rightarrow p} f(x)$$

(und ist dann eindeutig bestimmt), falls eine der äquivalenten Bedingungen gilt

1. $\forall \varepsilon > 0 \exists \delta > 0 \forall x \in D \setminus \{p\}. d(x, p) < \delta \Rightarrow d'(f(x), q) < \varepsilon$
2. Für jede Folge $x_k \in D \setminus \{p\}$ mit $x_k \rightarrow p$ gilt $f(x_k) \rightarrow b$

Korollar 4.4 Sei p Häufungspunkt von D

1. Ist $p \in D$, so ist f an p stetig genau dann, wenn $\lim_{x \rightarrow p} f(x) = f(p)$
2. Ist $p \notin D$ so kann man f genau dann zu einer an p stetigen Funktion erweitern (mit dem Wert q), wenn $q = \lim_{x \rightarrow p} f(x)$ existiert

4.4 Stetige Abbildungen in normierten Räumen

Sei (X, d) metrischer Raum und (X', d') der metrische Raum zu einem normierten Vektorraum V' mit Basis e_1, \dots, e_m . Dann ist eine Abbildung $f : X \rightarrow X'$ eindeutig bestimmt durch ihre *Komponenten*

$$f_i(x) = y^{(i)} (i = 1, \dots, m) \Leftrightarrow y = f(x)$$

d.h.

$$f(x) \hat{=} \begin{pmatrix} f_1(x) \\ \vdots \\ f_m(x) \end{pmatrix}$$

Lemma 4.5 1. f ist stetig an p genau dann, wenn die Komponenten $f_i : X \rightarrow \mathbb{R}$ ($i = 1, \dots, m$) stetige Abbildungen von (X, d) nach \mathbb{R} sind.

2. Sind $f, g : X \rightarrow V'$ stetig (an p), so auch $f + g : X \rightarrow V'$
3. Sind $f : X \rightarrow V'$ und $\phi : X \rightarrow \mathbb{R}$ stetig (an p), so auch $\phi \cdot f : X \rightarrow V'$
4. Sind $f, g : X \rightarrow V'$ stetig (an p), so auch $f \cdot g : X \rightarrow \mathbb{R}$ mit $(f \cdot g)(x) = f(x) \cdot g(x)$ skalares Produkt
5. Sind $f, g : X \rightarrow \mathbb{R}$ stetig (an p) so auch $h : D \rightarrow \mathbb{R}$ mit $D = \{x \in X \mid g(x) \neq 0\}$, $h(x) = \frac{f(x)}{g(x)}$

Beispiele s. S.13-15(AnaII-mehrdim) . Beweis. N ach Kor.2.3

$$f(x_k) \rightarrow f(p) \Leftrightarrow \forall i \ f_i(x_k) \rightarrow f_i(p)$$

also folgt Beh. mit Lemma 4.3 und bekannten Limesregeln. \square

Sei umgekehrt V ein normierter Vektorraum und $D \subseteq V$ mit der von V geerbten Metrik. Ist v_1, \dots, v_n eine Basis von V , so hat man die *Koordinatenfunktionen*

$$\kappa_i : D_i \rightarrow D \quad \kappa_i(r) = rv_i \text{ falls } rv_i \in D$$

die sind offenbar stetig und injektiv. Ist $f : D \rightarrow X'$ stetig, so sind auch die Kompositionen $f \circ \kappa_i$ stetig. Das Umgekehrte zu vermuten, ist abwegig s. Gegenbeispiel auf S.16 (AnaII-mehrdim)

4.5 Stetigkeit und Kompaktheit

Satz 4.6 *Sei f stetige Abbildung zwischen den metrischen Räumen (X, d) und (X', d') . Sei (X, d) kompakt. Dann gilt*

1. *Das Bild $W = f(X)$ ist kompakte Teilmenge von (X', d') .*
2. *Ist $X' = \mathbb{R}$ so gibt es $p, q \in X$ mit $f(p) = \min W$ und $f(q) = \max W$*
3. *f ist gleichmäßig stetig, d.h. $\forall \varepsilon > 0 \exists \delta > 0. \forall x, x'. d(x, x') < \delta \Rightarrow d(f(x), f(x')) < \varepsilon$*
4. *Ist f injektiv, so ist die Umkehrabbildung $f^{-1} : W \rightarrow X$ stetig, d.h. $f : X \rightarrow W$ ein Homöomorphismus*

Beweis. Zu 1. Sei $U_i, (i \in I)$ offene Überdeckung von W . Also gibt es Abbildung $h : X \rightarrow I$ so, dass $f(p)$ in $U_{h(p)}$. Wegen der Stetigkeit gibt es offene Umgebung V_p von p mit $f(V_p) \subseteq U_{h(p)}$. Wegen der Kompaktheit von X gibt es endliches $X_0 \subseteq X$ mit $X = \bigcup_{p \in X_0} V_p$. Es folgt:

$$W = f(X) = \bigcup_{p \in X_0} f(V_p) \subseteq \bigcup_{p \in X_0} U_{h(p)}$$

2. Folgt da hier $W = [a, b]$. Zu 3. Gegeben $\varepsilon > 0$. Zu jedem $p \in X$ gibt es wegen der Stetigkeit ein $\delta_p > 0$ so, dass

$$d'(f(x), f(p)) < \frac{\varepsilon}{42} \text{ für alle } x \in U_{\delta_p}(p)$$

Wegen Kompaktheit gibt es endliches $X_0 \subseteq X$ mit

$$X = \bigcup_{p \in X_0} U_{\frac{\delta_p}{42}}$$

Setze

$$\delta = \frac{1}{42} \min\{\delta_p \mid p \in X_0\}$$

Zu jeden $x \in X$ hat man ein $p \in X_0$ mit $x \in U_{\frac{\delta_p}{42}}(p)$, also $x, x' \in U_{\delta_p}(p)$ falls $d(x, x') < \delta$ und somit

$$d'(f(x), f(x')) \leq d'(f(x), f(p)) + d'(f(p), f(x')) < \varepsilon$$

Zu 4. Das Urbild von $A \subseteq X$ unter f^{-1} ist $f(A)$. Nach Lemma 4.1 genügt es zu zeigen, dass $f(A)$ abgeschlossen ist, falls A abgeschlossen. Sei also A abgeschlossen. Dann ist A auch kompakt: Ist $U_i (i \in I)$ offene Überdeckung von A , so erhält man durch Hinzunahme von $X \setminus A$ eine offene Überdeckung von X , also endliche Teilüberdeckung durch $X \setminus A$ und $U_i (i \in J)$. Dann ist aber $U_i (i \in J)$ Überdeckung von A . Nach 1. folgt nun, dass $f(A)$ kompakt ist, also nach Kor.2.5 abgeschlossen. \square

4.6 Stetigkeit und Zusammenhang

Eine stetige Abbildung γ von $[a, b] \subseteq \mathbb{R}$ in einen metrischen Raum (X, d) ist ein *Weg*. $M \subseteq X$ ist *wegzusammenhängend*, wenn es zu allen $p, q \in M$ einen Weg $\gamma : [a, b] \rightarrow M$ gibt mit $\gamma(a) = p$ und $\gamma(b) = q$. M ist *zusammenhängend*, wenn M nicht disjunkte Vereinigung zweier nichtleerer relativ offener Mengen ist, also wenn $U \cap M = \emptyset$ oder $V \cap M = \emptyset$ für je zwei offene Mengen mit $U \cap V \cap M = \emptyset$.

Satz 4.7 *Jede wegzusammenhängende Menge ist zusammenhängend. Ist M (weg) zusammenhängend und f stetig, so ist auch $f(M)$ (weg)zusammenhängend.*

Bew. s.S. 18-19 (AnaII-mehrdim)

4.7 Gleichmäßige Konvergenz

Sei $(V, \|\cdot\|)$ normierter Vektorraum. $f : D \rightarrow V$ sei beschränkt. Dann ist

$$\|f\|_{\infty} = \sup_{x \in D} \|f(x)\|$$

eine Norm, die *Supremumsnorm* zu $\|\cdot\|$ und somit der Vektorraum aller beschränkten Funktionen $f : D \rightarrow V$ normiert.

Sei nun V endlichdim.. Dann sind alle Supremumsnormen auf V äquivalent (mit den Konstanten, die die Äquivalenz der zugrundeliegenden Normen etablieren) und man darf einfach $B(D, V)$ schreiben. Eine Folge $f_k \in B(D, V)$ heißt *gleichmäßig* konvergent gegen $f \in B(D, V)$ falls

$$\|f_k - f\| \rightarrow 0$$

Satz 4.8 *Es gilt das Cauchy Kriterium. Sind die f_k stetig und $f_k \rightarrow f$ glm, so ist auch f stetig,*

Beweis wie für $V = \mathbb{R}$. Es folgt, dass $B(D, V)$ und der Raum $C(D, V)$ der stetigen beschränkten Abbildungen *vollständig* also *Banachräume* sind

5 Differentiation von Skalarenfeldern

5.1 Lineare Funktionen

Sei V ein K -Vektorraum. Eine skalare Funktion $f : V \rightarrow K$ ist *homogen linear* oder *Linearform*, wenn

$$f(\vec{x} + \vec{y}) = f(\vec{x}) + f(\vec{y}), \quad f(\lambda \vec{x}) = \lambda f(\vec{x}) \quad \text{für alle } \vec{x}, \vec{y} \in V, \lambda \in K$$

Eine skalare Funktion $g : V \rightarrow K$ ist *affin (linear)*, wenn es eine homogen lineare Funktion f und eine Konstante $b \in K$ gibt so, das

$$g(\vec{x}) = f(\vec{x}) + b \quad \text{für alle } \vec{x} \in V$$

Somit

$$b = g(\vec{0}), \quad g(\vec{x} + \vec{p}) - g(\vec{p}) = f(\vec{x})$$

Lemma 5.1 Sei $\dim V < \infty$. Gegeben eine skalare Funktion $g : V \rightarrow K$

1. g ist genau dann affin wenn zu einer/jeder Basis $\vec{e}_1, \dots, \vec{e}_n$ von V (eindeutig bestimmte) $a_1, \dots, a_n, b \in K$ gibt so, dass

$$g(\vec{x}) = a_1x_1 + \dots + a_nx_n + b \quad \text{für alle } \vec{x} = \sum_i x_i \vec{e}_i \in V$$

Identifiziert man V mit K^n , so hat die Funktion g als Graph die affine Hyperebene H (Gerade im Fall $n = 1$, Ebene im Fall $n = 2$) von K^{n+1}

$$H = \{(x_1, \dots, x_n, z) \in K^{n+1} \mid z = a_1x_1 + \dots + a_nx_n + b\} = (p_1, \dots, p_n, q) + H_0$$

mit dem parallelen Untervektorraum (Richtungsraum)

$$H_0 = \{(x_1, \dots, x_n, z) \in K^{n+1} \mid a_1x_1 + \dots + a_nx_n + (-1)z = 0\}$$

beliebigen $p_i \in K$ und $q = f(p_1, \dots, p_n)$.

2. Sei $K = \mathbb{R}$. g ist genau dann affin linear wenn es zu einem/jedem Skalarprodukt $\langle \cdot | \cdot \rangle$ auf V einen (zu diesem eindeutig bestimmten) Vektor \vec{v} , den Gradienten $\mathbf{grad} g$ und $b \in \mathbb{R}$ gibt mit

$$g(\vec{x}) = \langle \vec{v} | \vec{x} \rangle + b \quad \text{für alle } \vec{x} \in V$$

Ist $\vec{e}_1, \dots, \vec{e}_n$ Orthonormalbasis, so sind die a_i aus 1. die Koordinaten von $\mathbf{grad} g$. Identifiziert man V mit \mathbb{R}^n , so steht der Vektor mit Koordinaten $-a_1, \dots, -a_n, 1$ auf H_0 senkrecht, ist also ein Normalenvektor von H_0 und H .

Beweis. Zu 1. g von der angegebenen Form ist offenbar affin. Sei g affin, also $f(\vec{x}) = g(\vec{x}) - b$ Linearform. Setze $a_i = f(\vec{e}_i)$ für die zugrundegelegte Basis $\vec{e}_1, \dots, \vec{e}_n$. Es folgt $f(\vec{x}) = a_1x_1 + \dots + a_nx_n$ aus der Linearität. Fasst man V als K^n auf, so wird H durch eine lineare Gleichung definiert und H_0 durch die zugehörige homogene Gleichung. Also ist H_0 Untervektorraum von Dimension n und $H = \mathbf{h}_s + H_0$ für jede "spezielle Lösung" $\mathbf{h}_s \in H$.

Zum Beweis von 2. wähle $\vec{e}_1, \dots, \vec{e}_n$ als Orthonormalbasis von (V, \cdot) (die dann den kanonischen Einheitsspalten \mathbf{e}_i im \mathbb{R}^n entspricht) und beachte die koordinatenweise Berechnung des Skalarprodukts (Abschn.2.5). \square

Sei V euklidisch, $g : V \rightarrow K$ affin linear und $\vec{v} \in V$ ein Vektor $\neq \vec{0}$. Die Steigung von g an der Stelle p in Richtung \vec{v} ist gegeben durch das Dreieck $(p, g(p)), (\vec{v} + p, g(p)), (\vec{v} + p, g(\vec{v} + p))$

$$d_{\vec{v}}g(p) = \frac{g(\vec{v} + p) - g(p)}{|\vec{v}|}$$

Lemma 5.2 Sei $(V, \langle \cdot | \cdot \rangle)$ endlichdim. euklidisch und g affin linear. $d_{\vec{v}}g$ ist nur von der Richtung von \vec{v} abhängig, insbesondere unabhängig von p . Für $|\vec{v}| = 1$ gilt bzgl. einer Orthonormalbasis \vec{e}_i und $\vec{v} = \sum_i v_i \vec{e}_i$

$$d_{\vec{v}}g = a_1v_1 + \dots + a_nv_n = \langle \mathbf{grad} g | \vec{v} \rangle$$

Die Koordinate a_i von $\mathbf{grad} g$ ist die Steigung in Richtung \vec{e}_i , also der i -ten Koordinatenachse. $d_{\vec{v}}g$ ist maximal genau dann, wenn \vec{v} die Richtung von $\mathbf{grad} g$ hat.

Beweis.

$$d_{\vec{v}}g(p) = \frac{f(\vec{v})}{|\vec{v}|} = f(\vec{v}_0) \text{ mit } \vec{v}_0 = \frac{1}{|\vec{v}|}$$

Für $|\vec{v}| = 1$ folgt $d_{\vec{v}}g(\vec{v}) = f(\vec{v}) = \sum_{i=1}^n a_i v_i$ und damit die weiteren Behauptungen. \square

Geometrische Begründung des Gradienten: Gegeben Ebenen E und F im Raum, die sich in der Geraden g schneiden. Sei $O \in g$, $P \in E \setminus g$ und $Q \in F \setminus g$. Dann sind äquivalent

1. Der Winkel $\angle POQ$ ist maximal
2. Die Ebene H durch POQ steht senkrecht auf E und auf F .

Wir definieren einen der beiden durch E bestimmten Halbräume als den positiven E^+ . Für einen Vektor $\vec{x} \in E$ sei $X = \vec{x} + O$ und $Y \in F$ mit $XY \perp E$. Die zugehörige *Ordinate* ist nun definiert als

$$f(\vec{x}) = \begin{cases} |XY| & \text{falls } Y \in E^+ \\ -|XY| & \text{sonst} \end{cases}$$

Dann gibt es eindeutig bestimmten Vektor $\vec{g}_0 \in E$ mit $|\vec{g}_0| = 1$ und maximalem $f(\vec{g}_0)$, also maximaler Steigung. Es folgt $\vec{g}_0 \in H$. Der *Gradient* ist nun der Vektor

$$\vec{g} = f(\vec{g}_0)\vec{g}_0 \in H \cap E$$

Es folgt mit $\lambda = \langle \vec{g}_0 | \vec{x} \rangle$

$$f(\vec{x}) = f(\lambda\vec{g}_0) = \lambda|\vec{g}| = |g|\langle \vec{g}_0 | \vec{x} \rangle = \langle g | x \rangle$$

da $\lambda\vec{g}_0 + O$ der Fusspunkt des Lotes von $\vec{x} + O$ auf $H \cap E$ ist.

Sei schließlich \vec{n} ein normierter Normalenvektor von E in E^+ . Dann gilt

$$\langle \vec{n} - \vec{g} | \vec{y} \rangle = 0 \quad \text{falls } \vec{y} + O \in F$$

d.h. $\vec{n} - \vec{g}$ ist ein Normalenvektor von F . In der Tat $\langle n | g \rangle = f(\vec{y}) = \langle g | y \rangle$. Schliesslich

- Ist \vec{m} ein Normalenvektor von F mit $\vec{m} = \vec{h} + \vec{n}$ und $\vec{h} \in E$, so $\vec{h} = \vec{g}$

Ist also F der Graph einer linearen Funktion f mit $f(x_1\vec{e}_1 + x_2\vec{e}_2) = a_1x_1 + a_2x_2$ bzgl. einer Orthonormalbasis, so $\vec{g} = a_1\vec{e}_1 + a_2\vec{e}_2$ und $\vec{e}_3 - \vec{g}$ ist Normalenvektor von F .

Lemma 5.3 Sei V endlichdimensional, $D \subseteq V$ und $f : D \rightarrow \mathbb{R}$

1. Gilt $\lim_{n \rightarrow \infty} \frac{f(\vec{x}_n)}{\|\vec{x}_n\|} = 0$ bzgl. einer Norm von V , so auch bzgl. jeder anderen,
2. Ist $f : V \rightarrow \mathbb{R}$ homogen linear und $\frac{f(\vec{x})}{\|\vec{x}\|} \rightarrow 0$ für $\vec{x} \rightarrow \vec{0}$ in einer Umgebung von $\vec{0}$, so ist f konstant 0.

Beweis. 1. Sei die Konvergenz bzgl. einer Norm vorausgesetzt und $\|\cdot\|$ eine weitere Norm. Wegen der Äquivalenz der Normen gibt es $a, b > 0$ mit $a|\vec{x}| \leq \|\vec{x}\| \leq b|\vec{x}|$ und es folgt

$$\frac{1}{b} \frac{|f(\vec{x})|}{|\vec{x}|} \leq \frac{|f(\vec{x})|}{\|\vec{x}\|} \leq \frac{1}{a} \frac{|f(\vec{x})|}{|\vec{x}|}$$

Zu 2. Wähle euklidische Norm und orthonormale Basis. Für $\vec{x} = t\vec{e}_i$ gilt $f(\vec{x})/|\vec{x}| = a_i$, also $a_i = 0$. \square

5.2 Differenzierbarkeit

Wie im Falle der Funktionen einer Variablen, geht es um die Approximation einer Funktion f an einer Stelle p durch eine affin lineare Funktion. Wie vorher, schreiben wir \mathbb{R}^n , meinen aber einen beliebigen endlichdimensionalen normierten Vektorraum mit Norm $\|\cdot\|$, gegebenenfalls mit ausgezeichneteter Basis. In diesem Abschnitt ist $M \subseteq \mathbb{R}^n$ offen und $f : M \rightarrow \mathbb{R}$. Ist $x = (x_1, \dots, x_n) \in M$, so schreiben wir statt $f(x)$ auch $f(x_1, \dots, x_n)$.

Definition 5.4 Sei $M \subseteq \mathbb{R}^n$ offen, $f : M \rightarrow \mathbb{R}$ und $p \in M$ innerer Punkt. Die Funktion f heißt differenzierbar an der Stelle p , wenn es eine homogen lineare Funktion $df(p) = Df(p) = f'(p) : \mathbb{R}^n \rightarrow \mathbb{R}$, die Ableitung von f an der Stelle p , sowie eine in einer Umgebung U von $0 \in \mathbb{R}^n$ (mit $p + U \subseteq M$) definierte Funktion ϕ so gibt, dass

$$f(\Delta\vec{x} + p) = f(p) + df(p)(\Delta\vec{x}) + R(\Delta\vec{x}) \quad (5.1)$$

für alle $\vec{h} = \Delta\vec{x} \in U$ und

$$\lim_{\Delta\vec{x} \rightarrow \vec{0}} \phi(\Delta\vec{x}) = 0 \quad \text{mit} \quad \phi(\Delta\vec{x}) = \frac{1}{\|\Delta\vec{x}\|} R(\Delta\vec{x}) \quad (5.2)$$

Ist f in jedem Punkt $p \in M$ differenzierbar, so heißt f differenzierbar auf M .

Lemma 5.5 Wenn f an der Stelle p differenzierbar ist so ist f an p Lipschitz-stetig und die Ableitung $df(p)$ eindeutig bestimmt und es gibt (zu jedem Skalarprodukt) einen eindeutig bestimmten Vektor $\text{grad} f(p) = \nabla f(p) \in \mathbb{R}^n$, den Gradienten, mit

$$df(p)(\Delta\vec{x}) = \langle \text{grad} f(p) \mid \Delta\vec{x} \rangle$$

Differenzierbarkeit, Ableitung und Gradient hängen nicht von Norm oder Basis ab.

Beweis: Die L -Stetigkeit mit $L = |\text{grad} f(p)| + \varepsilon$ bzgl. einer euklidischen Norm folgt aus: $|f(\Delta\vec{x} + p) - f(p)| \leq |\langle \text{grad} f(p) \mid \Delta\vec{x} \rangle| + |R(\Delta\vec{x})| \rightarrow 0$ für $\Delta\vec{x} \rightarrow \vec{0}$.

Eindeutigkeit: Zu $i = 1, 2$ seien $\alpha_i : \mathbb{R}^n \rightarrow \mathbb{R}$ linear und $R_i : U_i \rightarrow \mathbb{R}$ gegeben mit $f(\Delta\vec{x} + p) = f(p) + \alpha_i(\Delta\vec{x}) + R_i(\Delta\vec{x})$ und $R_i(\Delta\vec{x})/\|\Delta\vec{x}\| \rightarrow 0$. Es folgt $(\alpha_1(\Delta\vec{x}) - \alpha_2(\Delta\vec{x}))/\|\Delta\vec{x}\| \rightarrow 0$ für $\Delta\vec{x} \rightarrow \vec{0}$ in der Umgebung $U_1 \cap U_2$, also $\alpha_1 - \alpha_2 = 0$ nach Lemma 2.

Die Unabhängigkeit von der Norm folgt aus Lemma 2. Die Unabhängigkeit von der Basis ist eh klar. \square

In Folgenden meinen wir mit \mathbb{R}^n einen euklidischen Vektorraum mit dem Skalarprodukt $\langle \vec{x} \mid \vec{y} \rangle = \vec{x} \cdot \vec{y}$ und der euklidischen Norm $|x|$ und identifizieren ihn über eine Orthormalbasis mit \mathbb{R}^n

Also $\vec{x} = x \hat{=} (x_1, \dots, x_n)$ - hier sollten wir eigentlich Spalten schreiben, manchmal deuten wir das durch $x \hat{=} (x_1, \dots, x_n)^T$ an. Wir schreiben \vec{x} statt x , wenn wir darauf hinweisen wollen, dass es eigentlich um einen Vektor geht.

Dann haben wir im Falle der Differenzierbarkeit eindeutig bestimmte a_1, \dots, a_n mit

$$df(p)(\Delta\vec{x}) = a_1 \Delta x_1 + \dots + a_n \Delta x_n = \langle \text{grad} f(p) \mid \Delta\vec{x} \rangle$$

d.h. (a_1, \dots, a_n) ist die Matrix der linearen Abbildung $df(p)$ bzgl. der Basis (also wirklich eine Zeile) und die Koordinaten(spalte) von $\mathbf{grad} f(p)$.

Vernachlässigen wir für kleine $\Delta\vec{x}$ noch den Term $R(\Delta\vec{x})$ in (5.1) (der ja schneller als $\Delta\vec{x}$ gegen $\vec{0}$ strebt), so erhalten wir eine lineare Approximation

$$f(\Delta\vec{x} + p) \approx f(p) + df(p)(\Delta\vec{x}) = f(p) + \langle \mathbf{grad} f(p) | \Delta\vec{x} \rangle \quad (5.3)$$

Ist f an der Stelle p differenzierbar mit Ableitung $df(p)$ bzw. Gradient $\mathbf{grad} f(p)$, so ist die *Tangentialhyperebene* von f an der Stelle p definiert als der Graph der affinen Abbildung $g : \mathbb{R}^n \rightarrow \mathbb{R}$ mit

$$g(\Delta\vec{x}) = f(p) + df(p)(\vec{x}) = f(p) + \langle \mathbf{grad} f(p) | \vec{x} \rangle + f(p)$$

Sind die a_i die Koordinaten von $\mathbf{grad} f(p)$, so haben wir die Gleichung der Tangentialhyperebene

$$z = f(p) + a_1 \Delta x_1 + \dots + a_n \Delta x_n = f(p) + a_1(x_1 - p_1) + \dots + a_n(x_n - p_n)$$

Im Fall $n = 2$ erhalten speziell

$$f(x, y) \approx f(x_0, y_0) + a_1(x - x_0) + a_2(y - y_0) \quad (5.4)$$

für alle (x, y) , die nahe bei (x_0, y_0) liegen. Durch die rechte Seite von (5.4) wird die Ebene

$$z = f(x_0, y_0) + a_1(x - x_0) + a_2(y - y_0)$$

im \mathbb{R}^3 festgelegt. Dies ist die *Tangentialebene* an den Graphen der Funktion $f : M \rightarrow \mathbb{R}$ (den man sich als Fläche über der xy -Ebene vorstellt) im Punkt $(x_0, y_0, f(x_0, y_0))$.

Korollar 5.6 Sei $y = f(\vec{x})$, $f : M \rightarrow \mathbb{R}$ und $p \in M \subseteq \mathbb{R}^n$. Wir schreiben

$$\Delta y = \Delta f = f(p + \Delta x) - f(p) \quad \text{wobei } p + \Delta x \in M$$

Die folgenden Aussagen sind äquivalent - mit $\vec{v} \in \mathbb{R}^n$

1. f ist an p differenzierbar und $\vec{v} = \mathbf{grad} f(p)$, d.h.

$$\Delta y = \langle \vec{v} | \Delta\vec{x} \rangle + R(\Delta\vec{x}) \quad \text{mit } \frac{R(\Delta\vec{x})}{|\Delta\vec{x}|} \rightarrow 0 \text{ für } \Delta\vec{x} \rightarrow \vec{0}$$

2. Es gibt eine auf einer Umgebung U von $\vec{0}$ definierte Funktion $\Phi : U \rightarrow \mathbb{R}^n$

$$\Delta y = \langle \vec{v} + \Phi(\Delta\vec{x}) | \Delta x \rangle, \quad \Phi(\Delta x) \rightarrow \vec{0} \text{ für } \Delta\vec{x} \rightarrow \vec{0}$$

3. Es gibt eine auf einer Umgebung U von $\vec{0}$ definierte Funktion $\vec{a} : U \rightarrow \mathbb{R}^n$ mit

$$\Delta y = \langle \vec{a}(\Delta\vec{x}) | \Delta x \rangle, \quad \vec{a}(\Delta x) \rightarrow \vec{v} \text{ für } \Delta\vec{x} \rightarrow \vec{0}$$

Beweis. 1 \Leftrightarrow 2: mit

$$\Phi(\Delta x) := \frac{R(\Delta \vec{x})}{|\Delta \vec{x}|^2} \Delta \vec{x}$$

gilt

$$\langle \vec{v} | \Delta \vec{x} \rangle + R(\Delta \vec{x}) = \langle \vec{v} | \Delta \vec{x} \rangle + \frac{R(\Delta \vec{x})}{|\Delta \vec{x}|^2} \langle \Delta \vec{x} | \Delta \vec{x} \rangle = \langle \vec{v} + \Phi(\Delta \vec{x}) | \Delta x \rangle$$

$$|\Phi(\Delta x)| = \frac{|R(\Delta \vec{x})|}{|\Delta \vec{x}|^2} |\Delta \vec{x}| = \frac{|R(\Delta \vec{x})|}{|\Delta \vec{x}|}$$

2 \Leftrightarrow 3.

$$\vec{a}(\Delta \vec{x}) = \vec{v} + \Phi(\Delta \vec{x})$$

5.3 Richtungsableitungen

Definition 5.7 Sei $M \subseteq \mathbb{R}^n$ offen, $f : M \rightarrow \mathbb{R}$, $p \in U$ und $\vec{v} \in \mathbb{R}^n$ ein Vektor der Länge $|\vec{v}| = 1$. Man sagt, dass f im Punkt p eine Ableitung in Richtung des Vektors \vec{v} besitzt, wenn der Grenzwert

$$\lim_{t \rightarrow 0} \frac{f(t\vec{v} + p) - f(p)}{t}$$

existiert. Dieser heißt dann die Richtungsableitung von f in Richtung v und wird mit $d_{\vec{v}}f(p)$ oder $f'(p, \vec{v})$ bezeichnet.

Satz 5.8 Seien M, f, p wie in Definition 5.7, und sei f in p differenzierbar. Dann existiert für jeden Einheitsvektor $\vec{v} \hat{=} (v_1, \dots, v_n)^T \in \mathbb{R}^n$ die Ableitung von f in Richtung \vec{v} im Punkt p , und es gilt

$$d_{\vec{v}}f(p) = a_1v_1 + \dots + a_nv_n = \langle \mathbf{grad} f(p) | \vec{v} \rangle \quad (5.5)$$

Die Richtungsableitung ist also die Steigung der Tangentialebene in dieser Richtung.

Beweis Für hinreichend kleines t ist in den Bezeichnungen von Kor.5.6

$$\frac{f(t\vec{v} + p) - f(p)}{t} = \frac{\langle \vec{a}(t\vec{v}) | t\vec{v} \rangle}{t} = \langle \vec{a}(t\vec{v}) | \vec{v} \rangle \rightarrow \langle \mathbf{grad} f(p) | \vec{v} \rangle.$$

■

Führt man wie im \mathbb{R}^2 den Winkel $\varphi \in [0, \pi]$ zwischen zwei Vektoren $x, y \in \mathbb{R}^n \setminus \{0\}$ durch $\langle x | y \rangle = |x||y| \cos \varphi$ ein, so folgt für $(\mathbf{grad} f)(p) \neq 0$ für den Winkel φ zwischen $(\mathbf{grad} f)(p)$ und v

$$d_{\vec{v}}(p) = |(\mathbf{grad} f)(p)| \cdot \cos \varphi.$$

Korollar 5.9 Seien M, f, p wie in Definition 5.7 und sei $(\mathbf{grad} f)(p) \neq 0$. Dann nimmt die Richtungsableitung $d_{\vec{v}}(p)$ ihren größten Wert für $\cos \varphi = 1$ bzw. $\varphi = 0$ an, d.h. wenn $(\mathbf{grad} f)(p)$ und \vec{v} in die gleiche Richtung zeigen. Der Gradient von f zeigt also in Richtung des steilsten Anstiegs von f in p (und steht senkrecht auf der Höhenlinie durch p).

Richtungsableitungen und dann auch partielle Ableitungen wurden mit Bezug auf ein vorgegebenes Skalarprodukt eingeführt. Die Abhängigkeit vom Skalarprodukt gilt sogar für $n = 1$: Sei $f : V = \mathbb{R} \rightarrow W = \mathbb{R}$ die Abbildung $f(x) = x$. Auf $W = \mathbb{R}$ haben wir das

“kanonische Skalarprodukt $(a, b) \mapsto ab$. Auf $V = \mathbb{R}$ sei das Skalarprodukt definiert durch $\langle a | b \rangle = Cab$, $C > 0$ beliebige Konstante. Wir haben die Ableitung

$$df(p)(\Delta x) = \Delta x$$

weil

$$f(p + \Delta x) - f(p) = \Delta x$$

linear ist. Sei $\vec{v} = 1$. Nun hat \vec{v} die Länge C , also ist $\vec{v}_0 = \frac{1}{C}\vec{v} = \frac{1}{C}$ der zugehörige normierte Vektor und wir haben

$$\lim_{t \rightarrow 0} \frac{1}{t} (f(p + t\vec{v}_0) - f(p)) = \lim_{t \rightarrow 0} \frac{1}{C} = \frac{1}{C}$$

Richtungsableitungen lassen sich auch ohne Skalarprodukt, aber mit Bezug auf die zugrundegelegte Norm auf V definieren

$$d_{\vec{v}}f(p) = df(p)(\vec{v}) \quad \text{für } \|\vec{v}\| = 1$$

Sei nun $\vec{e}_1, \dots, \vec{e}_n$ eine Basis von V mit $\|\vec{e}_i\| = 1$ und f an p differenzierbar. Dann

$$d_{\vec{v}}f(p) = \sum_{i=1}^n v_i d_{\vec{e}_i}f(p) \quad \text{für } \vec{v} = \sum_i v_i \vec{e}_i$$

nämlich wegen der Linearität von $df(p)$

$$d_{\vec{v}}f(p) = df(p)(\vec{v}) = \sum_i v_i df(p)(\vec{e}_i) = \sum_i v_i d_{\vec{e}_i}f(p)$$

Will man die Beschränkung auf normierte Vektoren und den Bezug auf die Norm loswerden, so definiert man am besten die *Ableitung von f längs \vec{v}*

$$d_{\vec{v}}f(p) = \lim_{t \rightarrow 0} \frac{f(t\vec{v} + p) - f(p)}{t}$$

und hat im Falle der Differenzierbarkeit

$$d_{\vec{v}}(p) = df(\vec{v})$$

Nämlich

$$\frac{f(t\vec{v} + p) - f(p)}{t} = \frac{1}{t} df(p)(t\vec{v}) + \frac{R(t\vec{v})}{\|t\vec{v}\|} = df(p)(\vec{v}) + \frac{R(t\vec{v})}{\|t\vec{v}\|} \rightarrow df(p)(\vec{v})$$

Das ist dann vom Vektor \vec{v} , nicht nur von der Richtung abhängig. Dafür ist es aber von der Norm unabhängig. Ist $\vec{e}_1, \dots, \vec{e}_n$ eine Basis von V , so folgt sofort aus der Linearität von $df(p)$

$$d_{\vec{v}}f(p) = \sum_{i=1}^n v_i d_{\vec{e}_i}f(p) \quad \text{für } \vec{v} = \sum_i v_i \vec{e}_i$$

5.4 Partielle Ableitungen

Definition 5.10 Die Funktion $f : M \rightarrow \mathbb{R}$ heißt in $p = (p_1, \dots, p_n) \in M$ partiell differenzierbar nach x_k bzgl. einer gegebenen Basis $\vec{e}_1, \dots, \vec{e}_n$, falls die Ableitung längs des Basisvektors \vec{e}_k , d.h. in Richtung der x_k -Koordinatenachse,

$$\begin{aligned} f_{x_k}(p) &= \frac{\partial f}{\partial x_k}(p) = \frac{\partial}{\partial x_k} f(p) = d_k f(p) = d_{\vec{e}_k} f(p) = \lim_{h \rightarrow 0} \frac{f(h\vec{e}_k + p) - f(p)}{h} \\ &= \lim_{h \rightarrow 0} \frac{f(p_1, \dots, p_{k-1}, p_k + h, p_{k+1}, \dots, p_n) - f(p_1, \dots, p_n)}{h} \end{aligned}$$

existiert. Dieser Grenzwert heißt die partielle Ableitung von f nach x_k an der Stelle p . Die Funktion f heißt partiell differenzierbar in p bzgl. der gegebenen Basis, wenn in p alle partiellen Ableitungen existieren.

D.h. man hält alle Variablen bis auf eine konstant und leitet die resultierende Funktion nach eben dieser Variablen ab.

Beispiel 3 Sei $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ gegeben durch $f(x, y) = 2x^2 + y^2$. Beim partiellen Differenzieren nach x betrachten wir y als Konstante. Wir können dann die bekannten Differentiationsregeln für Funktionen einer Veränderlichen anwenden und erhalten

$$f_x(x, y) = 4x \quad \text{sowie} \quad f_y(x, y) = 2y.$$

Analog erhält man für die Funktion $f : \mathbb{R}^n \rightarrow \mathbb{R}$,

$$f(x_1, \dots, x_n) := \sqrt{x_1^2 + \dots + x_n^2} = |x|,$$

dass sie auf $\mathbb{R}^n \setminus \{0\}$ partiell differenzierbar ist und dass dort

$$\frac{\partial f}{\partial x_i}(x) = f_{x_i}(x) = \frac{1}{2}(x_1^2 + \dots + x_n^2)^{-\frac{1}{2}} \cdot 2x_i = \frac{x_i}{|x|}. \quad \blacksquare$$

Beispiel 4 Sei $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ erklärt durch $f(0, 0) = 0$ und

$$f(x, y) = \frac{xy}{(x^2 + y^2)^2} \quad \text{falls } (x, y) \neq (0, 0), \quad f(0, 0) = 0.$$

Für $(x, y) \neq (0, 0)$ finden wir die partiellen Ableitungen sofort:

$$f_x(x, y) = \frac{y}{(x^2 + y^2)^2} - 4 \frac{x^2 y}{(x^2 + y^2)^3}, \quad f_y(x, y) = \frac{x}{(x^2 + y^2)^2} - 4 \frac{xy^2}{(x^2 + y^2)^3}.$$

An der Stelle $(x, y) = (0, 0)$ arbeiten wir mit Definition 5.10:

$$f_x(0, 0) = \lim_{h \rightarrow 0} \frac{f(h, 0) - f(0, 0)}{h} = 0, \quad f_y(0, 0) = 0.$$

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

$$f\left(\frac{1}{n}, \frac{1}{n}\right) = \frac{(1/n^2)}{(2/n^2)^2} = \frac{n^2}{4} \rightarrow \infty \quad \text{für } n \rightarrow \infty$$

ist f aber nicht stetig in $(0, 0)$. Aus der partiellen Differenzierbarkeit folgt also *nicht* die Stetigkeit. \blacksquare

Korollar 5.11 Sei $M \subseteq \mathbb{R}^n$ offen und $f : M \rightarrow \mathbb{R}$ in $p \in M$ differenzierbar. Dann ist f in p partiell differenzierbar, und es gilt

$$\text{grad } f(p) \hat{=} \begin{pmatrix} \frac{\partial f}{\partial x_1}(p) \\ \vdots \\ \frac{\partial f}{\partial x_n}(p) \end{pmatrix}$$

Beweis.

$$d_{\vec{e}_k} f(p) = \langle \text{grad } f(p) \mid \vec{e}_k \rangle \quad \square$$

Ist $f : M \rightarrow \mathbb{R}$ partiell differenzierbar an der Stelle $p \in M$ bzgl. der gegebenen Basis $\alpha : \vec{e}_1, \dots, \vec{e}_n$, so dürfen wir also definieren

$$(\text{grad}^\alpha f)(p) \hat{=} \begin{pmatrix} \frac{\partial f}{\partial x_1}(p) \\ \vdots \\ \frac{\partial f}{\partial x_n}(p) \end{pmatrix}$$

der *Gradient* von f bzgl. α an der Stelle p . Statt $\text{grad } f$ schreibt man auch ∇f (lies: Nabla).
Warnung: nicht nur die Koordinaten, auch der Begriff des Gradienten ist hier basisabhängig.

5.5 Partielle Ableitungen als Funktionen

Existiert die partielle Ableitung $\frac{\partial f}{\partial x_k}(x)$ für alle $x \in M$, so können wir die Funktion

$$\frac{\partial f}{\partial x_k} : M \rightarrow \mathbb{R}$$

betrachten. Ist diese stetig an der Stelle p und M Umgebung von p so ist f an der Stelle p stetig nach x_k differenzierbar. Entsprechend ist f auf M (an p stetig) partiell differenzierbar bzgl. der gegebenen Basis, wenn dies für alle x_k der Fall ist.

Satz 5.12 Ist f in p bzgl. einer Basis α stetig partiell differenzierbar (d.h. die Abbildung $u \mapsto \text{grad}^\alpha f(u)$ ist an der Stelle p stetig), so ist f an der Stelle p differenzierbar

Es genügt, dass die partiellen Ableitungen in einer Umgebung von p beschränkt sind.
Beweis. Wir haben zu zeigen, dass die Differenzierbarkeitsbedingung gilt mit

$$df(p)(\vec{h}) = \langle \text{grad } f(p) \mid \vec{h} \rangle$$

Mit dem Mittelwertsatz der Differentialrechnung gibt es $\eta_k = \eta_k(\vec{h})$ zwischen 0 und h_k so, dass

$$\begin{aligned} f(p + \vec{h}) - f(p) &= \\ &= [f(p + h_1 \vec{e}_1) - f(p)] + [f(p + h_1 \vec{e}_1 + h_2 \vec{e}_2) - f(p + h_1 \vec{e}_1)] + \dots \\ &= \dots + [f(p + h_1 \vec{e}_1 + \dots + h_n \vec{e}_n) - f(p + h_1 \vec{e}_1 + \dots + h_{n-1} \vec{e}_{n-1})] = \\ &= h_1 d_1 f(p + \eta_1 \vec{e}_1) + h_2 d_2 f(p + h_1 \vec{e}_1 + \eta_2 \vec{e}_2) + \dots + h_n d_n f(p + h_1 \vec{e}_1 + \dots + h_{n-1} \vec{e}_{n-1} + \eta_n \vec{e}_n) \\ &= \langle \vec{a}(\vec{h}, \vec{\eta}) \mid \vec{h} \rangle \end{aligned}$$

mit

$$\vec{a}(\vec{h}, \vec{\eta}) = \begin{pmatrix} d_1 f(p + \eta_1 \vec{e}_1) \\ d_2 f(p + h_1 \vec{e}_1 + \eta_2 \vec{e}_2) \\ \vdots \\ d_n f(p + h_1 \vec{e}_1 + \dots + \eta_n \vec{e}_n) \end{pmatrix} \rightarrow \text{grad } f(p) \quad \text{für } \vec{h} \rightarrow \vec{0}$$

wegen der Stetigkeit der partiellen Ableitungen (da auch $\eta_k(\vec{h}) \rightarrow \vec{0}$). \square

Zwischen den eingeführten Differenzierbarkeitsbegriffen bestehen also die folgenden Beziehungen:

$$\begin{array}{ccccc} f \text{ stetig partiell} & \Rightarrow & f \text{ differenzierbar} & \Rightarrow & f \text{ partiell} \\ \text{differenzierbar} & & & & \text{differenzierbar} \\ & & \Downarrow & & \\ & & f \text{ stetig} & & \end{array}$$

5.6 Rechenregeln

Wir kommen nun zu Rechenregeln für Ableitungen.

Satz 5.13 Sei $M \subseteq \mathbb{R}^n$ offen und $f, g : M \rightarrow \mathbb{R}$ in $x \in M$ differenzierbar. Dann ist für $\alpha, \beta \in \mathbb{R}$ auch $\alpha f + \beta g : M \rightarrow \mathbb{R}$ in x differenzierbar, und es gilt

$$(\alpha f + \beta g)'(x) = \alpha f'(x) + \beta g'(x).$$

Beweis. Gemäß Kor.5.6 sei

$$\Delta f = \langle \vec{a}(\Delta \vec{x}) \mid \Delta \vec{x} \rangle, \quad \Delta g = \langle \vec{b}(\Delta \vec{x}) \mid \Delta \vec{x} \rangle$$

Dann

$$\Delta(\alpha f + \beta g) = \alpha \Delta f + \beta \Delta g = \langle \alpha \vec{a}(\Delta \vec{x}) + \beta \vec{b}(\Delta \vec{x}) \mid \Delta \vec{x} \rangle \rightarrow \alpha \text{grad } f(p) + \beta \text{grad } g(p) \square$$

Satz 5.14 (Kettenregel) Seien $U \subseteq \mathbb{R}^n$, $V \subseteq \mathbb{R}$ offen, und $f : U \rightarrow \mathbb{R}$, $g : V \rightarrow \mathbb{R}$ seien Funktionen mit $f(U) \subseteq V$. Ist $y = f(x)$ in $p \in U$ und $z = g(y)$ in $q = g(p)$ differenzierbar, so ist die verkettete Funktion $g \circ f : U \rightarrow \mathbb{R}$ in p differenzierbar, und es gilt

$$d(g \circ f)(p) = g'(q) \cdot df(p) \quad \text{bzw.} \quad \text{grad}(g \circ f) = \frac{\partial g}{\partial y}(q) \text{grad } f(p).$$

Beweis. Wir haben gemäß] Kor.5.6

$$\Delta y = \langle \vec{a}(\Delta \vec{x}) \mid \Delta \vec{x} \rangle, \quad \vec{a}(\Delta \vec{x}) \rightarrow \text{grad } f(p) \quad \text{für } \Delta \vec{x} \rightarrow \vec{0}$$

$$\Delta z = b(\Delta y) \Delta y, \quad b(\Delta y) \rightarrow g'(q) \quad \text{für } \Delta y \rightarrow 0$$

Also

$$\Delta z = b(\Delta y) \langle \vec{a}(\Delta \vec{x}) \mid \Delta \vec{x} \rangle = \langle b(\Delta y) \vec{a}(\Delta \vec{x}) \mid \Delta \vec{x} \rangle \quad \text{mit } b(\Delta y) \vec{a}(\Delta \vec{x}) \rightarrow g'(q) \cdot \text{grad } f(p) \quad \text{für } \Delta \vec{x} \rightarrow \vec{0}$$

da wegen Stetigkeit von f an p auch $\Delta y \rightarrow 0$. \square

Beispiel. Sei $g : \mathbb{R} \rightarrow \mathbb{R}$ gegeben durch $g(z) = e^z$ und $f : (0, \infty) \times (0, 2\pi) \rightarrow \mathbb{R}$ durch $f(x, y) = x \cos y$. Sei $h : (0, \infty) \times (0, 2\pi) \rightarrow \mathbb{R}$ gleich $g \circ f$. Dann

$$\text{grad } f(x, y) = \begin{pmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{pmatrix} = \begin{pmatrix} \cos y \\ -x \sin y \end{pmatrix},$$

Also

$$\text{grad } h(x, y) = e^{x \cos y} \begin{pmatrix} \cos y \\ -x \sin y \end{pmatrix},$$

Satz 5.15 Sei $M \subseteq \mathbb{R}^n$ offen und $f, g : M \rightarrow \mathbb{R}$ in $p \in M$ differenzierbar. Dann sind auch $fg : M \rightarrow \mathbb{R}$ und, falls $f(p) \neq 0$, $1/f$ in p differenzierbar, und es ist

$$d(fg)(p) = g(p)df(p) + f(p)dg(p), \quad d\frac{1}{f}(p) = \frac{-1}{f(p)^2}df(p)$$

Beweis. Die erste Behauptung später. Bei der zweiten haben wir Stetigkeit, also $f(x) \neq 0$ in einer Umgebung von p . Nun Kettenregel mit $g(y) = \frac{1}{y}$. \square

5.7 Differentiale und Fehlerabschätzung

Das *partielle Differential* $d_k f(p)$ von f an der Stelle p in Richtung der x_k -Achse ist die homogene lineare Funktion

$$d_k f(p)(dx_k) = \frac{\partial f}{\partial x_k}(p) \cdot dx_k$$

Das *totale Differential* $df(p)$ an der Stelle p ist dann die homogene lineare Funktion aus (5.3)

$$df(p)(d\vec{x}) = \sum_{k=1}^n d_k f(p)(d\vec{x}) = \frac{\partial f}{\partial x_1}(p) \cdot dx_1 + \dots + \frac{\partial f}{\partial x_n}(p) \cdot dx_n = \langle \text{grad } f(p) \mid d\vec{x} \rangle$$

Das kann wie folgt interpretiert werden: Ändert man die x_k um dx_k , geht also von p zu $d\vec{x} + p$ über, so ändert sich bei linearer Approximation unter Vernachlässigung von Fehlern höherer Ordnung f um df . Das wird oft zur Fehlerabschätzung bei Messprozessen ausgenutzt. Werden statt der wahren Werte x_1, \dots, x_n mit Fehlern behaftete Werte $\hat{x}_k = x_k + \Delta x_k$, $k = 1, \dots, n$, gemessen, so lässt sich der Fehler

$$\Delta f = f(\hat{x}_1, \dots, \hat{x}_n) - f(x_1, \dots, x_n)$$

im Funktionswert abschätzen durch

$$|\Delta f| \approx \sum_{k=1}^n \left| \frac{\partial f}{\partial x_k} \right| |\Delta x_k| = \langle \text{grad } f(x) \mid \Delta \vec{x} \rangle,$$

wobei Fehler höherer Ordnung ignoriert werden. Sind Schranken $|\Delta x_k| \leq s_k$ für die Meßfehler bekannt, erhalten wir

$$|\Delta f| \leq \sum_{k=1}^n \left| \frac{\partial f}{\partial x_k} \right| s_k.$$

Beispiel 6 In einem rechtwinkligen Dreieck mit den Kathedenlängen x und y ist die Länge der Hypotenuse gleich $f(x, y) = \sqrt{x^2 + y^2}$. Wegen

$$f_x(x, y) = \frac{x}{\sqrt{x^2 + y^2}} \quad \text{und} \quad f_y(x, y) = \frac{y}{\sqrt{x^2 + y^2}}$$

wird die Tangentialebene an den Graphen der Funktion $f : (0, \infty) \times (0, \infty) \rightarrow \mathbb{R}$ im Punkt $(x_0, y_0, f(x_0, y_0))$ beschrieben durch die Gleichung

$$z = \sqrt{x_0^2 + y_0^2} + \frac{x_0}{\sqrt{x_0^2 + y_0^2}}(x - x_0) + \frac{y_0}{\sqrt{x_0^2 + y_0^2}}(y - y_0),$$

und das vollständige Differential von f lautet

$$df = \frac{x}{\sqrt{x^2 + y^2}} dx + \frac{y}{\sqrt{x^2 + y^2}} dy.$$

Messen wir statt der wahren Kathedenlängen x_0, y_0 die Werte $x_0 + \Delta x$ und $y_0 + \Delta y$, so ist der resultierende Fehler bei linearer Approximation gleich

$$\Delta f \approx \frac{x_0}{\sqrt{x_0^2 + y_0^2}} \Delta x + \frac{y_0}{\sqrt{x_0^2 + y_0^2}} \Delta y = \frac{x_0 \Delta x + y_0 \Delta y}{\sqrt{x_0^2 + y_0^2}}. \quad \blacksquare$$

5.8 Mittelwertsatz

Der Mittelwertsatz für eine differenzierbare Funktion $f : \mathbb{R} \rightarrow \mathbb{R}$ sagt, dass man für zwei Punkte $a < b$ einen Punkt $\xi \in (a, b)$ findet, so dass

$$f(b) - f(a) = f'(\xi) \cdot (b - a). \quad (5.6)$$

Das folgende Beispiel zeigt, dass für vektorwertige Funktionen der Mittelwertsatz in dieser Form *nicht* gelten kann.

Beispiel 8 Sei $f : \mathbb{R} \rightarrow \mathbb{R}^2$ gegeben durch $f(x) = (\cos x, \sin x)^T$. Dann ist $f'(x) = (-\sin x, \cos x)^T$ und nach Pythagoras

$$\|f'(x)\| = \sqrt{\sin^2 x + \cos^2 x} = 1$$

für alle x . Es gibt deshalb *kein* $\xi \in \mathbb{R}$ mit

$$0 = f(2\pi) - f(0) = f'(\xi) \cdot 2\pi. \quad \blacksquare$$

Für *reellwertige* Funktionen auf \mathbb{R}^n gilt der Mittelwertsatz jedoch in der gewohnten Form.

Satz 5.16 Sei $D \subseteq \mathbb{R}^n$ offen und $f : D \rightarrow \mathbb{R}$ differenzierbar auf D . Weiter sei $p + t\vec{h} \in D$ für alle $t \in [0, 1]$ (d.h. die Verbindungsstrecke von p nach $p + \vec{h}$ liege komplett in D). Dann gibt es ein $\tau \in (0, 1)$ mit

$$f(p + \vec{h}) - f(p) = f'(p + \tau\vec{h})(\vec{h}) = \langle \text{grad} f(p + \tau\vec{h}) | \vec{h} \rangle$$

Zum Beweis schränkt man die Funktion f auf die Strecke $[p, p + \vec{h}] = \{y \in \mathbb{R}^n : y = p + t\vec{h} \text{ mit } t \in [0, 1]\}$ ein, betrachtet die Funktion

$$g : [0, 1] \rightarrow \mathbb{R}, \quad t \mapsto f(p + t\vec{h}) \quad (5.7)$$

und wendet auf diese den Mittelwertsatz (5.6) an. ■

Korollar 5.17 (Schrankensatz.) *Sei D zusätzlich konvex (d.h. D enthält mit je 2 Punkten auch ihre Verbindungsstrecke) und sei $L \geq |\text{grad} f(\vec{x})|$ für alle $\vec{x} \in D$. Dann ist L Lipschitz-Konstante für f .*

Beweis. Mit Cauchy-Schwarz

$$|f(p + \Delta\vec{x}) - f(p)| \leq |(\text{grad} f(p + \tau\Delta\vec{x}) | \Delta\vec{x})| \leq |\text{grad} f(p + \tau\Delta\vec{x})| \cdot |\Delta\vec{x}| \leq L \cdot |\Delta\vec{x}| \quad \square$$

6 Vektorfelder und lineare Approximation

Wir verallgemeinern Kap.5. auf Funktionen von mehreren Veränderlichen, deren Werte Vektoren sind, die also von einer Teilmenge D des \mathbb{R}^n in den \mathbb{R}^m abbilden. Solche Funktionen heißen auch *Vektorfelder*. Als Beispiel stellen wir uns eine durch ein Rohr strömende Flüssigkeit vor. Ordnet man jedem Punkt x im Inneren des Rohres den Geschwindigkeitsvektor des an dieser Stelle befindlichen Flüssigkeitsteilchens zu, so erhält man ein Vektorfeld. Auch das elektrische Feld ist ein Vektorfeld. Ist schließlich $f : \mathbb{R}^n \rightarrow \mathbb{R}$ partiell differenzierbar, so ist die Funktion $F(x) := (\text{grad} f)(x)$ ein Vektorfeld von \mathbb{R}^n nach \mathbb{R}^m . Ein Vektorfeld $F : \mathbb{R}^n \rightarrow \mathbb{R}^m$ schreiben wir oft als

$$F = (F_1, \dots, F_m)^T$$

mit Funktionen $F_k : \mathbb{R}^n \rightarrow \mathbb{R}$.

6.1 Differenzierbarkeit

Definition 6.1 *Sei $D \subseteq \mathbb{R}^n$ offen, $F : D \rightarrow \mathbb{R}^m$ und $p \in D$. Die Funktion F heißt differenzierbar an der Stelle p , wenn es eine lineare Abbildung $A : \mathbb{R}^n \rightarrow \mathbb{R}^m$ sowie eine in einer Umgebung U von $\vec{0} \in \mathbb{R}^n$ definierte Funktion R so gibt, dass*

$$\lim_{\Delta\vec{x} \rightarrow \vec{0}} \frac{R(\Delta\vec{x})}{\|\Delta\vec{x}\|} = \vec{0}. \quad (6.1)$$

wobei

$$R(\Delta\vec{x}) = F(p + \Delta\vec{x}) - F(p) - A\Delta\vec{x} \quad (6.2)$$

für alle $\Delta\vec{x} \in U$ und

Die Ableitung einer Funktion $F : D \rightarrow \mathbb{R}^m$ ist an der Stelle p dann die lineare Abbildung A von \mathbb{R}^n in \mathbb{R}^m . Wir bezeichnen die Ableitung von $F : D \rightarrow \mathbb{R}^m$ in $p \in D$ mit $F'(p)$ oder $(DF)(p)$.

Hat man eine Basis α gewählt, so kann man A einfach als Matrix verstehen, die Jacobimatrix

$$A = J_F(p) = J_{\alpha,p}$$

Ist F in jedem Punkt $p \in D$ differenzierbar, so heißt F differenzierbar auf D . Ist die Abbildung

$$x \rightarrow J_F(x), \quad D \rightarrow \mathbb{R}^{m \times n}$$

stetig, so heißt F auf D stetig differenzierbar.

Analog zu Kor.5.6 gilt

$\vec{y} = F(\vec{x})$ ist an p differenzierbar mit der Ableitung A genau dann, wenn Umgebung U und es zu jedem $\Delta\vec{x}$ mit $p + \Delta\vec{x} \in U$ eine lineare Abbildung $A_{\Delta\vec{x}} : V \rightarrow W$ gibt so, dass

$$\Delta\vec{y} = \Delta F = F(p + \Delta\vec{x}) - F(p) = A_{\Delta\vec{x}}(\Delta\vec{x}), \quad A_{\Delta\vec{x}} \rightarrow A \text{ für } \Delta\vec{x} \rightarrow \vec{0}$$

Beweis. Wir wählen ON-Basis $\vec{e}_1, \dots, \vec{e}_m$ von W und haben die Komponenten

$$F(\vec{x}) = F_1(\vec{x})\vec{e}_1 + \dots + F_m(\vec{x})\vec{e}_m$$

Nach Kor.5.6 gilt

$$\vec{v}_i = \text{grad}F_i(p) \Leftrightarrow \Delta F_i = \langle \vec{a}_{i\Delta\vec{x}} | \Delta\vec{x} \rangle \text{ mit } \vec{a}_{i\Delta\vec{x}} \rightarrow \vec{v}_i \text{ für } \Delta\vec{x} \rightarrow \vec{0}$$

und übersetzt in die zugehörigen Linearformen

$$f_i = dF_i(p) \Leftrightarrow \Delta F_i = f_{i\Delta\vec{x}}(\Delta\vec{x}) \text{ mit } f_{i\Delta\vec{x}} \rightarrow f_i \text{ für } \Delta\vec{x} \rightarrow \vec{0}$$

Also folgt die behauptete Äquivalenz vermöge

$$A(\vec{x}) = \sum_i f_i(\vec{x})\vec{e}_i, \quad A_{\Delta\vec{x}}(\vec{x}) = \sum_i f_{i\Delta\vec{x}}(\vec{x})\vec{e}_i \quad \square$$

6.2 Jacobimatrix und partielle Ableitungen

Satz 6.2 Sei $D \subseteq \mathbb{R}^n$ offen und $F = (F_1, \dots, F_m)^T : D \rightarrow \mathbb{R}^m$. Dann ist F an $p \in D$ differenzierbar genau dann, wenn jede der Funktionen $F_i : D \rightarrow \mathbb{R}$ an der Stelle p differenzierbar ist. In diesem Falle ist die i -te Zeile der Jacobimatrix $J_F(p)$ (bzgl. der gewählten Basis) gerade die Matrix der Ableitung $dF_i(p)$ von F_i also

$$J_F(p) := \begin{pmatrix} \frac{\partial F_1}{\partial x_1}(p) & \dots & \frac{\partial F_1}{\partial x_n}(p) \\ \vdots & & \vdots \\ \frac{\partial F_m}{\partial x_1}(p) & \dots & \frac{\partial F_m}{\partial x_n}(p) \end{pmatrix}. \quad (6.3)$$

Beweis. Seien die R_i die Komponenten von R . Dann gilt

$$|R_i(\Delta\vec{x})| \leq |R(\Delta\vec{x})| \leq \sum_{i=1}^m |R_i(\Delta\vec{x})|$$

und die behauptete Äquivalenz folgt sofort und das Weitere aus Satz 5.8 \square

Insbondre ist $(DF_i)(p) = dF_i(p)$ und $J_{F_i}(p) \hat{=} \text{grad}F_i(p)$. Im Allgemeinen stehen also in der i -ten Zeile von $J_F(p)$ die Koordinaten von $\text{grad}F_i(p)$.

Die Zeilen der Jacobimatrix von F entsprechen den totalen Differentialen der Komponenten F_i

Korollar 6.3 *Differenzierbarkeit und Ableitung hängen nicht von der Norm ab. Ist F an p differenzierbar, so dort auch Lipschitz-stetig. F ist auf D stetig differenzierbar genau dann, wenn alle partiellen Ableitungen der F_i auf D existieren und stetig sind.*

6.3 Differenzierbare Kurven

Eine Abbildung $F(t) = \vec{y}(t)$, $F : U \rightarrow \mathbb{R}^m$, $U \subseteq \mathbb{R}$, ist eine *Kurve* in \mathbb{R}^m und wir schreiben die Ableitung an p (den *Tangentenvektor*) als

$$\frac{\partial \vec{y}}{\partial t}(p) \hat{=} \begin{pmatrix} \frac{\partial y_1}{\partial t}(p) \\ \vdots \\ \frac{\partial y_m}{\partial t}(p) \end{pmatrix}$$

Nach Kor.5.6 gilt, indem man komponentensweise zusammensetzt,

$$\Delta \vec{y} = (\Delta t) \cdot \vec{a}(\Delta t) \quad \text{mit } \vec{a}(\Delta t) \rightarrow \frac{\partial \vec{y}}{\partial t}(p) \text{ für } \Delta t \rightarrow 0$$

Lemma 6.4 (Kettenregel) *Seien $U \subseteq \mathbb{R}$, $V \subseteq \mathbb{R}^m$ offen, und $f : U \rightarrow \mathbb{R}^m$, $g : V \rightarrow \mathbb{R}$ seien Funktionen mit $g(U) \subseteq V$. Ist $\vec{y} = f(t)$ in $p \in U$ und $z = g(\vec{y})$ in $q = f(p)$ differenzierbar, so ist die verkettete Funktion $g \circ f : U \rightarrow \mathbb{R}$ in p differenzierbar, und es gilt*

$$(g \circ f)'(p) = \langle \text{grad}g(q) \mid \frac{\partial \vec{y}}{\partial t}(p) \rangle = \sum_{i=1}^n \frac{\partial g}{\partial y_i}(\vec{y}(p)) \cdot \frac{\partial y_i}{\partial t}(p) = J_g(q) \cdot J_f(p)$$

Beweis: Gemäß Kor.5.6

$$\Delta z = \langle \vec{b}(\Delta \vec{y}) \mid \Delta \vec{y} \rangle \quad \text{mit } \vec{b}(\Delta \vec{y}) \rightarrow \text{grad}g(q) \text{ für } \Delta \vec{0} \rightarrow \vec{0}$$

Also

$$\Delta z = \langle \vec{b}(\Delta \vec{y}) \mid (\Delta t) \cdot \vec{a}(\Delta t) \rangle = \langle \vec{b}(\Delta \vec{y}) \mid \vec{a}(\Delta t)(p) \rangle \Delta t$$

und

$$\langle \vec{b}(\Delta \vec{y}) \mid \vec{a}(\Delta t) \rangle \rightarrow \langle \text{grad}g(q) \mid \frac{\partial \vec{y}}{\partial t} \rangle \quad \text{für } \Delta t \rightarrow 0$$

da wegen der Stetigkeit von f auch $\Delta \vec{y} \rightarrow \vec{0}$. \square

Es folgt, dass der *Gradient senkrecht auf der Höhenlinie steht*: Sei $U \subseteq \mathbb{R}^2$, $f : U \rightarrow \mathbb{R}$ and q differenzierbar und $\vec{x}(t) \in \mathbb{R}^2$ für an $p \in D \subseteq \mathbb{R}$, differenzierbar, $q = \vec{x}(p)$ und $f(\vec{x}(t)) = c$ für $t \in D$. Dann

$$\langle \text{grad}f(q) \mid \frac{\partial \vec{x}}{\partial t}(p) \rangle = 0$$

6.4 Regeln

Satz 6.5 (Kettenregel) Seien $U \subseteq \mathbb{R}^n$, $V \subseteq \mathbb{R}^m$ offen, und $F : U \rightarrow \mathbb{R}^m$, $G : V \rightarrow \mathbb{R}^k$ seien Funktionen mit $F(U) \subseteq V$. Ist F in $p \in U$ und G in $F(p)$ differenzierbar, so ist die verkettete Funktion $G \circ F : U \rightarrow \mathbb{R}^k$ in p differenzierbar, und es gilt

$$J_{G \circ F}(p) = J_G(F(p)) \cdot J_F(p).$$

Die Ableitung einer verketteten Funktion ist also gleich der Verkettung (Hintereinanderausführung) der Ableitungen. Für die Jacobimatrizen ist also das Matrixprodukt zu bilden. **Beweis:** Wir untersuchen, welcher Eintrag in der i -ten Zeile und der l -ten Spalte von $J_{G \circ F}(p)$ steht. Dazu betrachten wir

$$\mathbf{x}(x_l) = \begin{pmatrix} p_1 \\ \vdots \\ x_l \\ \vdots \\ p_n \end{pmatrix} \quad x_l \in \mathbb{R} \text{ in der } l\text{-ten Zeile, die anderen Einträge fest}$$

und die folgende Kurve in \mathbb{R}^m

$$\mathbf{y}(x_l) = \begin{pmatrix} F_1(\mathbf{x}(x_l)) \\ \vdots \\ F_m(\mathbf{x}(x_l)) \end{pmatrix} \quad \text{mit} \quad \frac{\partial \mathbf{y}}{\partial x_l}(p_l) = \begin{pmatrix} \frac{\partial F_1}{\partial x_l}(\mathbf{x}(p_l)) \\ \vdots \\ \frac{\partial F_m}{\partial x_l}(\mathbf{x}(p_l)) \end{pmatrix}$$

Nach dem gerade bewiesenen Fall der Kettenregel haben wir

$$\frac{\partial G_i(\mathbf{y}(p_l))}{\partial x_l} = \langle \mathbf{grad} G_i(\mathbf{y}(p_l)) \mid \frac{\partial \mathbf{y}}{\partial x_l}(p_l) \rangle = J_{G_i}(\mathbf{y}(p_l)) \cdot \frac{\partial \mathbf{y}}{\partial x_l}(p_l)$$

also gerade das Ergebnis der Matrizenmultiplikation.

Es ist aber auch noch die Differenzierbarkeit zu zeigen. Zudem vereinfacht sich auch der Beweis der Aussage über die Jacobimatrix, wenn man die handliche Charakterisierung der Differenzierbarkeit benutzt.

Also sei $\vec{y} = F(\vec{x})$ differenzierbar an p und $\vec{z} = G(\vec{y})$ differenzierbar an $q = F(p)$. Das bedeutet

$$\begin{aligned} \Delta \vec{y} &= A_{\Delta \vec{x}}(\Delta \vec{x}), & A_{\Delta \vec{x}} &\rightarrow A \text{ für } \Delta \vec{x} \rightarrow \vec{0} \\ \Delta \vec{z} &= B_{\Delta \vec{y}}(\Delta \vec{y}), & B_{\Delta \vec{y}} &\rightarrow B \text{ für } \Delta \vec{y} \rightarrow \vec{0} \end{aligned}$$

Es folgt

$$\Delta \vec{z} = B_{\Delta \vec{y}}(A_{\Delta \vec{x}}(\Delta \vec{x})) \text{ mit } B_{\Delta \vec{y}} A_{\Delta \vec{x}} \rightarrow BA \text{ für } \Delta \vec{x} \rightarrow \vec{0}$$

indem man $B_{\Delta \vec{y}}$ und $A_{\Delta \vec{x}}$ als Matrizen auffasst (also auch eine Basis von V wählt) und die Stetigkeit von Addition und Multiplikation reeller Zahlen benutzt. \square

Beispiel. Sei $G : \mathbb{R}^2 \rightarrow \mathbb{R}$ gegeben durch $G(x, y) = e^{xy}$ und $F : (0, \infty) \times (0, 2\pi) \rightarrow \mathbb{R}^2$ durch $F = (F_1, F_2)^T$ mit $F_1(x, y) = x \cos y$ und $F_2(x, y) = x \sin y$. Sei $H : (0, \infty) \times (0, 2\pi) \rightarrow \mathbb{R}$

gleich $G \circ F$ (d.h. $H(x, y) = G(F(x, y))$). Dann sind die Funktionalmatrizen von G und F gleich

$$J_G(x, y) = \left(\frac{\partial G}{\partial x}(x, y), \frac{\partial G}{\partial y}(x, y) \right) = (ye^{xy}, xe^{xy})$$

und

$$J_F(x, y) = \begin{pmatrix} \frac{\partial F_1}{\partial x} & \frac{\partial F_1}{\partial y} \\ \frac{\partial F_2}{\partial x} & \frac{\partial F_2}{\partial y} \end{pmatrix} = \begin{pmatrix} \cos y & -x \sin y \\ \sin y & x \cos y \end{pmatrix},$$

und wegen

$$\begin{aligned} J_G(F(x, y)) &= (x \sin y e^{x^2 \sin y \cos y}, x \cos y e^{x^2 \sin y \cos y}) \\ &= x e^{x^2 \sin y \cos y} (\sin y, \cos y) \end{aligned}$$

ist

$$\begin{aligned} J_H(x, y) &= J_G(F(x, y)) J_F(x, y) \\ &= x e^{x^2 \sin y \cos y} (\sin y, \cos y) \begin{pmatrix} \cos y & -x \sin y \\ \sin y & x \cos y \end{pmatrix} \\ &= x e^{x^2 \sin y \cos y} (2 \sin y \cos y, -x(\sin^2 y - \cos^2 y)). \end{aligned}$$

■

Korollar 6.6 Sei $D \subseteq \mathbb{R}^n$ offen und $F, G : D \rightarrow \mathbb{R}^m$ in $x \in D$ differenzierbar. Dann ist für $\alpha, \beta \in \mathbb{R}$ auch $\alpha F + \beta G : D \rightarrow \mathbb{R}^m$ in x differenzierbar, und es gilt

$$J_{\alpha F + \beta G}(x) = \alpha J_F(x) + \beta J_G(x).$$

Der Beweis von Satz 5.13 folgt nun so: Sei $F_1 = f$ und $F_2 = g$ und $G(x, y) = xy$. Dann ist F in p und G überall differenzierbar und

$$J_F(p) = \begin{pmatrix} \frac{\partial f}{\partial x_1} & \cdots & \frac{\partial f}{\partial x_n} \\ \frac{\partial g}{\partial x_1} & \cdots & \frac{\partial g}{\partial x_n} \end{pmatrix}, \quad J_G\left(\begin{pmatrix} q_1 \\ q_2 \end{pmatrix}\right) = (q_2 \quad q_1)$$

$$H(\vec{x}) = (G \circ F)(\vec{x}) = g(\vec{x}) \cdot f(\vec{x})$$

$$J_H(p) = J_G(f(p), g(p)) \cdot J_F(p) = (g(p) \quad f(p)) \cdot \begin{pmatrix} \frac{\partial f}{\partial x_1} & \cdots & \frac{\partial f}{\partial x_n} \\ \frac{\partial g}{\partial x_1} & \cdots & \frac{\partial g}{\partial x_n} \end{pmatrix} = g(p) J_f(p) + f(p) J_g(p)$$

6.5 Polarkoordinaten und Co

Polarkoordinaten:

$$\sigma(r, \varphi) = (r \cos \varphi, r \sin \varphi)$$

$$J_\sigma(r, \varphi) = \begin{pmatrix} \cos \varphi & -r \sin \varphi \\ \sin \varphi & r \cos \varphi \end{pmatrix}$$

und daher $\det J_\sigma(r, \varphi) = r$.

Häufig benutzt man auch *verallgemeinerte Polarkoordinaten*

$$(x, y) = \sigma(r, \varphi) = (ar \cos \varphi, br \sin \varphi) \quad \text{mit } a, b > 0.$$

Zylinderkoordinaten Für die Transformationsfunktion

$$(x, y, z) = \sigma(r, \varphi, z) = (r \cos \varphi, r \sin \varphi, z)$$

gilt

$$\det \sigma'(r, \varphi, z) = \det \begin{pmatrix} \cos \varphi & -r \sin \varphi & 0 \\ \sin \varphi & r \cos \varphi & 0 \\ 0 & 0 & 1 \end{pmatrix} = r.$$

Kugelkoordinaten Für die Transformationsfunktion

$$(x, y, z) = \sigma(r, \varphi, \vartheta) = (r \cos \varphi \sin \vartheta, r \sin \varphi \sin \vartheta, r \cos \vartheta)$$

gilt

$$\det \sigma'(r, \varphi, \vartheta) = \det \begin{pmatrix} \cos \varphi \sin \vartheta & -r \sin \varphi \sin \vartheta & r \cos \varphi \cos \vartheta \\ \sin \varphi \sin \vartheta & r \cos \varphi \sin \vartheta & r \sin \varphi \cos \vartheta \\ \cos \vartheta & 0 & -r \sin \vartheta \end{pmatrix} = -r^2 \sin \vartheta.$$

In der Tat mit

$$a = \sin \varphi, \quad b = \cos \varphi, \quad c = \sin \vartheta, \quad d = \cos \vartheta$$

gilt

$$a^2 + b^2 = 1, \quad c^2 + d^2 = 1$$

und man erhält die Determinante durch Herausziehen von r und Entwickeln nach der zweiten Spalte als

$$\begin{aligned} \det \begin{pmatrix} bc & -rac & rbd \\ ac & rbc & rad \\ d & 0 & -rc \end{pmatrix} &= r^2 \det \begin{pmatrix} bc & -ac & bd \\ ac & bc & ad \\ d & 0 & -c \end{pmatrix} = r^2 (ac \det \begin{pmatrix} ac & ad \\ d & -c \end{pmatrix} + bc \det \begin{pmatrix} bc & bd \\ d & -c \end{pmatrix}) \\ &= r^2 (ac(-ac^2 - ad^2) + bc(-bc^2 - bd^2)) = r^2 (-a^2c(c^2 + d^2) - b^2c(c^2 + d^2)) \\ &= -r^2(a^2c + b^2c) = -r^2c(a^2 + b^2) = -r^2c \end{aligned}$$

6.6 Mittelwertsatz

Für Funktionen $F : \mathbb{R}^n \supseteq D \rightarrow \mathbb{R}^m$ mit $m \geq 2$ ist die folgende Version des Mittelwertsatzes die nächstbeste.

Satz 6.7 Sei $D \subseteq \mathbb{R}^n$ offen, $F : D \rightarrow \mathbb{R}^m$ differenzierbar auf D und $x + t\vec{h} \in D$ für alle $t \in [0, 1]$. Dann ist

$$F(x + \vec{h}) - F(x) = \int_0^1 DF(x + \tau\vec{h})(\vec{h}) \, d\tau = \left(\int_0^1 DF(x + \tau\vec{h}) \, d\tau \right)(\vec{h}) \quad (6.4)$$

Dabei ist das Integral über die vektorwertige Funktion $\tau \mapsto DF(x + \tau\vec{h})(\vec{h})$ bzw die Matrix $J_F(x + \tau\vec{h})$ komponentenweise zu berechnen. Der Mittelwertsatz in der Form (6.4) ist oft nützlich, wenn man $|F(x + \vec{h}) - f(x)|$ abschätzen möchte. Der Beweis ist wieder ganz einfach: o.B.d.A. ist $m = 1$. Mit der Funktion g aus (5.7) ist nach dem Hauptsatz

$$F(x + \vec{h}) - F(x) = g(1) - g(0) = \int_0^1 g'(\tau) d\tau,$$

und nach der Kettenregel ist $g'(\tau) = DF(x + \tau\vec{h})(\vec{h})$. ■

7 Rechnen mit Taylorreihen in 1 Variablen

Taylorpolynome bzw. -reihen kann man oft aus bekannten durch einfache Rechenoperationen herleiten - der Einfachheit halber diskutieren wir das zuerst für den Fall 1 Variablen. Wir führen dazu folgende zweckdienliche Notation für $f : D \rightarrow \mathbb{R}$ und $p \in D$

$$T_p^n(f) = \sum_{k=0}^n a_k (x-p)^k \quad a_k = \frac{1}{k!} f^{(k)}(p), \quad n \in \mathbb{N}_0 \cup \infty, \quad T_p(f) = T_p^\infty(f)$$

also das Taylorpolynom vom Grad n bzw. die Taylorreihe am Entwicklungspunkt p - vorausgesetzt, dass das existiert. Das ist dann ein Polynom bzw. Potenzreihe in $(x-p)$ und dafür kann man auch Δx oder t schreiben.

7.1 Satz von Taylor-Lagrange

Satz 7.1 (Taylor) Sei $f : [a, b] \rightarrow \mathbb{R}$ n -mal stetig differenzierbar, und auf (a, b) existiere die $n+1$ -te Ableitung. Dann gibt es ein $\xi \in (a, b)$ so, dass

$$f(b) = \underbrace{\sum_{k=0}^n \frac{1}{k!} f^{(k)}(a) (b-a)^k}_{= T_a^n(b)} + \underbrace{\frac{f^{(n+1)}(\xi)}{(n+1)!} (b-a)^{n+1}}_{=: R_n(b, a)}. \quad (7.5)$$

Dabei ist $T_a^n(b)$ das Taylorpolynom vom Grad n von f in a , und $R_n(b, a)$ heißt das Restglied nach Lagrange.

7.2 Näherungsweise Gleichheit

Für zwei Potenzreihen (also auch für Polynome) schreiben wir

$$\Phi = \sum_{k=0}^{\infty} a_k (x-p)^k \approx_p^n \Psi = \sum_{k=0}^{\infty} b_k (x-p)^k \Leftrightarrow a_k = b_k \text{ für } k = 0, \dots, n.$$

Für Potenzreihen verträgt sich das Rechnen mit näherungsweise Gleichheit offenbar mit Addition, Subtraktion, Multiplikation und Substitution d.h.

$$\Phi_1 \approx_p^n \Psi_1 \text{ und } \Phi_2 \approx_p^n \Psi_2 \Rightarrow \lambda \Phi_1 * \Phi_2 \approx_p^n \lambda \Psi_1 * \Psi_2$$

wobei $\lambda \in \mathbb{R}$ und $*$ für jede der Verknüpfungen $+$, $-$, \cdot , \circ steht und \circ die Substitution bedeutet. Z.B. wenn wir $\sin x$ und $\cos x$ missbräuchlich für die zugehörige Reihen schreiben und $\sin x \cos x$ für deren Cauchy-Produkt

$$\sin x \approx_0^3 x - \frac{1}{6}x^3, \quad \cos x \approx_0^3 1 - \frac{1}{2}x^2 \Rightarrow \sin x \cos x \approx_0^3 x - \frac{5}{12}x^3$$

Die Substitution

$$\Psi \circ \Phi \text{ von } \Phi = \sum_{k=0}^{\infty} a_k(x-p)^k \text{ in } \Psi = \sum_{n=0}^{\infty} b_n(x-q)^n$$

ist nur dann definiert, wenn $q = a_0$, und zwar als die Potenzreihe, die sich aus

$$\sum_{n=0}^{\infty} b_n(-q + \sum_{k=0}^{\infty} a_k(x-p)^k)^n = b_0 + \sum_{n=1}^{\infty} b_n \left(\sum_{k=1}^{\infty} a_k(x-p)^k \right)^n$$

durch Ausmultiplizieren und Zusammenfassen nach gleichen Potenzen von x ergibt.

7.3 Allgemeine Ableitungsregeln

Dann haben wir die folgenden *allgemeinen Ableitungsregeln*

$$\begin{aligned} (a) \quad T_p^n(\lambda f) &= \lambda \cdot T_p^n(f) && \lambda \in \mathbb{R} \\ (b) \quad T_p^n(f+g) &= T_p^n(f) + T_p^n(g) \\ (c) \quad T_p^n(f \cdot g) &\approx_p^n T_p^n(f) \cdot T_p^n(g) \\ (d) \quad T_p^n(g \circ f) &\approx_p^n T_q^n(g) \circ (T_p^n(f)) && \text{mit } q = f(p) \\ (e) \quad T_p^n(\int f) &\approx_p^{n+1} \int T_p^n(f) \\ (f) \quad \frac{\partial T_p^n(f)}{\partial x} &\approx_p^{n-1} T_p^n\left(\frac{\partial f}{\partial x}\right) \end{aligned}$$

Wichtig ist: Die Ableitungen müssen existieren, aber Konvergenz der Reihe ist kein Thema.

Achtung: Bei der Kettenregel (d) wird hier $T_p^n(f)$ für x in $T_q^n(g)(x-q) = \sum_{k=0}^n b_k(x-q)^k$ eingesetzt. Schreibt man $T_q^n(g)(t) = \sum_{k=0}^n b_k t^k$ und will in t einsetzen, so muss man $T_p^n(f) - q$ in t einsetzen - das ist genau dasselbe, nur umständlicher formuliert.

Beweis. Indem wir $z := x-p$ und $u := y-q$ als neue Variable nehmen, können wir den Beweis auf den Fall $p = q = 0$ reduzieren. (a) und (b) folgen sofort aus den entsprechenden Regeln für Ableitungen. Wir setzen nun

$$\hat{f} = f - T_0^n(f), \quad \hat{g} = g - T_0^n(g)$$

und erhalten somit

$$\hat{f}^{(k)} = 0 = \hat{g}^{(k)} \quad \text{für } k \leq n$$

Es folgt

$$f \cdot g = T_0^n(f) \cdot T_0^n(g) + h, \quad h = \hat{f}T_0^n(g) + \hat{g}T_0^n(f) + \hat{f}\hat{g}$$

also $h^{(k)} = 0$ für $k \leq n$ und somit (c). Aus (b) und (c) folgt (d) für den Fall, dass g Polynom ist. Im Allgemeinen dann wieder

$$g \circ f = T_0^n(g) \circ f + \hat{g} \circ f$$

$$T_0^n(g \circ f) = T_0^n(T_0^n(g) \circ f) + T_0^n(\hat{g} \circ f) \approx_0^n T_0^n(T_0^n(g) \circ T_0^n(f)) + 0 = T_0^n(g) \circ T_0^n(f)$$

(e) und (f) folgen mit Hauptsatz und formaler gliedweiser Integration bzw. Differentiation.

7.4 Venachlässigung höherer Terme

Wir können jetzt jedoch für beliebige (hinreichend oft differenzierbare) Funktionen definieren

$$f \approx_p^n g \Leftrightarrow T_p^n(f) \approx_p^n T_p^n(g)$$

und haben dann die Verträglichkeit wie bei Potenzreihen

$$f_1 \approx_p^n g_1 \text{ und } f_2 \approx_p^n g_2 \Rightarrow \lambda f_1 * f_2 \approx_p^n \lambda g_1 * g_2$$

wobei im Falle der Substitution $f_2(p) = g_2(p)$ vorausgesetzt werden muss. Wir dürfen dies jedoch nur als Aussagen über die Terme in den Taylorentwicklungen interpretieren und nur dann auf die Funktionen zurückschließen, wenn Konvergenz gesichert ist.

Beispiele.

$$\begin{aligned} x(1+x-\cos x) &\approx_0^2 x(x-\frac{1}{2}x^2) = x^2 - \frac{1}{2}x^2 \approx_0^2 x^2 \\ \cos(x+x^2) &\approx_0^3 1 - \frac{1}{2}(x+x^2)^2 = 1 - \frac{1}{2}x^2 - x^3 - \frac{1}{2}x^4 \approx_0^3 1 - \frac{1}{2}x^2 - x^3 \end{aligned}$$

Auch ohne Kenntnis der Ableitung von $\sqrt{1+x}$ können wir aus

$$\sqrt{1+x^2} = 1+x = 2(1+\frac{1}{2}(x-1)) \approx_1^1 [\sqrt{2}(1+\frac{1}{4}(x-1))]^2$$

schließen, dass

$$\begin{aligned} \sqrt{1+x} &\approx_1^1 \sqrt{2}(1+\frac{1}{4}(x-1)) = \sqrt{2} + \frac{\sqrt{2}}{4}(x-1) \\ \int \sqrt{1+x} &\approx_1^2 \sqrt{2}((x-1) + \frac{1}{8}(x-1)^2) \end{aligned}$$

Aber

$$\frac{\partial \sqrt{1+x}}{\partial x} \approx_1^1 \sqrt{2}(\frac{1}{4} - \frac{1}{32}(x-1)) \not\approx_1^1 \sqrt{2} = \frac{\partial \sqrt{2}(1+\frac{1}{2}(x-1))}{\partial x}$$

8 Taylorentwicklung in n Variablen

8.1 Höhere partielle Ableitungen

Ist $f : M \rightarrow \mathbb{R}$ partiell differenzierbar, so ist $\frac{\partial f}{\partial x_k}$ wieder eine Funktion von M nach \mathbb{R} . Ist diese partiell differenzierbar nach x_ℓ , ergeben sich *partielle Ableitungen zweiter Ordnung*, die wir mit

$$D_l D_k f(x) = \frac{\partial}{\partial x_l} \frac{\partial f}{\partial x_k} = \frac{\partial^2 f}{\partial x_\ell \partial x_k}(x) = f_{x_k x_\ell}(x)$$

bezeichnen (die letzte Notation sollte man vermeiden, weil sie im Widerspruch zu der in der Mathematik üblichen Schreibrichtung steht). Analog werden partielle Ableitungen höherer Ordnung erklärt und bezeichnet.

Beispiel 5 Für die erste Funktion aus Beispiel 3 ist

$$f_{xx}(x, y) = 4, \quad f_{xy}(x, y) = f_{yx}(x, y) = 0, \quad f_{yy}(x, y) = 2.$$

Interessanter ist die durch $f(0,0) = 0$ und

$$f(x, y) := xy \frac{x^2 - y^2}{x^2 + y^2} \quad \text{für } (x, y) \neq (0, 0)$$

auf ganz \mathbb{R}^2 erklärte Funktion. Für $(x, y) \neq (0, 0)$ ist

$$f_x(x, y) = y \frac{x^4 - y^4 + 4x^2y^2}{(x^2 + y^2)^2}, \quad f_y(x, y) = x \frac{x^4 - y^4 - 4x^2y^2}{(x^2 + y^2)^2},$$

und für $(x, y) = (0, 0)$ findet man

$$f_x(0, 0) = \lim_{h \rightarrow 0} \frac{f(h, 0) - f(0, 0)}{h} = 0 \quad \text{und} \quad f_y(0, 0) = 0.$$

Für die gemischten zweiten Ableitungen in $(0, 0)$ erhalten wir schließlich

$$\begin{aligned} f_{xy}(0, 0) &= \lim_{h \rightarrow 0} \frac{f_x(0, h) - f_x(0, 0)}{h} = \lim_{h \rightarrow 0} \frac{-h - 0}{h} = -1, \\ f_{yx}(0, 0) &= \lim_{h \rightarrow 0} \frac{f_y(h, 0) - f_y(0, 0)}{h} = \lim_{h \rightarrow 0} \frac{h - 0}{h} = 1. \end{aligned}$$

■

8.2 Vertauschen der Reihenfolge

Die Reihenfolge der partiellen Ableitungen darf i.Allg. also nicht vertauscht werden. Der folgende Satz gibt Bedingungen an, unter denen dieses Vertauschen erlaubt ist.

Satz 8.1 (H.A. Schwarz) Sei $M \subseteq \mathbb{R}^n$ offen, $f : M \rightarrow \mathbb{R}$ und $p \in M$. Alle partiellen Ableitungen erster Ordnung von f sollen auf M existieren. Außerdem existiere für ein Paar $i \neq j$ die zweite Ableitung $\frac{\partial^2 f}{\partial x_i \partial x_j}$ auf M , und diese sei in p stetig. Dann existiert auch $\frac{\partial^2 f}{\partial x_j \partial x_i}$ in p , und es ist

$$\frac{\partial^2 f}{\partial x_j \partial x_i}(p) = \frac{\partial^2 f}{\partial x_i \partial x_j}(p).$$

Beweis. O.B.d.A. $n = 2$, $i = 2$, $j = 1$. Schreibt man $z = f(x, y)$ so ist $D_1 D_2 f(p) = \frac{\partial^2 z}{\partial x \partial y}(p)$ gegeben und $D_2 D_1 f(p) = \frac{\partial^2 z}{\partial y \partial x}(p)$ gesucht. Indem man $f(x)$ durch $f(x - p)$ ersetzt, hat man o.B.d.A. $p = 0$. Indem man $f(x, y)$ durch $f(x, y) - f(x, 0) - f(0, y) + f(0, 0)$ ersetzt hat man o.B.d.A. $f(x, 0) = f(0, y) = 0$. Indem man $f(x, y)$ durch $f(x, y) - \frac{\partial^2 f}{\partial x \partial y}(\vec{0})xy$ ersetzt hat man o.B.d.A. $\frac{\partial^2 f}{\partial x \partial y}(\vec{0}) = 0$. Wir sollen zeigen

$$0 = \lim_{t \rightarrow 0} \frac{1}{t} \frac{\partial z}{\partial x}(0, t) = \lim_{t \rightarrow 0} \frac{1}{t} \lim_{s \rightarrow 0} \frac{f(s, t)}{s} = \lim_{t \rightarrow 0} \lim_{s \rightarrow 0} \frac{1}{s} \frac{f(s, t)}{t}$$

Nach dem Mittelwertsatz gibt es τ zwischen 0 und t mit

$$\frac{f(s, t)}{t} = \frac{\partial z}{\partial y}(s, \tau)$$

und wegen der Differenzierbarkeit von $\frac{\partial z}{\partial y}(s, \tau)$ nach s gibt es $\sigma = \sigma(\tau)$ zwischen 0 und s mit

$$\frac{1}{s} \frac{\partial z}{\partial y}(s, \tau) = \frac{\partial}{\partial x} \frac{\partial z}{\partial y}(\sigma, \tau)$$

Somit

$$\frac{1}{s} \frac{f(s, t)}{t} = \frac{\partial}{\partial x} \frac{\partial z}{\partial y}(\sigma, \tau) \rightarrow 0 \text{ für } (s, t) \rightarrow (0, 0)$$

da auch $(\sigma, \tau) \rightarrow (0, 0)$ und $\frac{\partial}{\partial x} \frac{\partial z}{\partial y}$ nach Voraussetzung stetig ist. \square

8.3 Multiindices

Um auch mit höheren Ableitungen bequem umgehen zu können, führen wir die folgende Notation ein:

- $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{N}_0^n$ ist ein *Multiindex*
- $|\alpha| := \alpha_1 + \dots + \alpha_n$ ist seine *Ordnung*
- $\alpha! := \alpha_1! \dots \alpha_n!$ ist seine *Fakultät*.
- $\beta \pm \alpha = (\beta_1 \pm \alpha_1, \dots, \beta_n \pm \alpha_n)$.
- $\beta \leq \alpha \Leftrightarrow \beta_i \leq \alpha_i$ für alle i

Zu gegebenen Variablen X_1, \dots, X_n schreiben, definieren und sehen wir

- $X = (X_1, \dots, X_n)$
- $X^\alpha := X_1^{\alpha_1} \dots X_n^{\alpha_n}$ *Monom*
- $X^\alpha \cdot X^\beta = X^{\alpha+\beta}$ falls $X_i X_j = X_j X_i$
- $\sum_{|\alpha| \leq d} a_\alpha X^\alpha \in \mathbb{R}[X_1, \dots, X_n]$ *Polynom* vom Grad $\leq d$ mit *Koeffizienten* $a_\alpha \in \mathbb{R}$

8.4 Ableitungen

Sei $f : U \rightarrow \mathbb{R}$ auf einer offenen Menge $U \subseteq V$ definiert und eine Basis von V gewählt, vermöge derer wir V mit \mathbb{R}^n identifizieren. Unter der Voraussetzung, dass die entsprechenden partiellen Ableitungen existieren, definieren wir induktiv

$$D_i^0 f = f, \quad D_i f = \frac{\partial}{\partial x_i} f = \frac{\partial f}{\partial x_i}, \quad D_i^k f = \frac{\partial^k}{\partial x_i^k} f = \frac{\partial^k f}{\partial x_i^k} = D_i D_i^{k-1} f$$

und mit $D = (D_1, \dots, D_n)$

$$D^\alpha f = D_1^{\alpha_1} \dots D_n^{\alpha_n} f = \frac{\partial^{|\alpha|}}{\partial x_1^{\alpha_1} \dots \partial x_n^{\alpha_n}} f = \frac{\partial^{|\alpha|} f}{\partial x_1^{\alpha_1} \dots \partial x_n^{\alpha_n}}$$

Dabei wird D^α als *Operator* aufgefasst, der der Funktion $f : U \rightarrow \mathbb{R}$ die Funktion $D^\alpha f$ zuordnet

$$f \mapsto D^\alpha f, \quad D^\alpha f : U \rightarrow \mathbb{R} \text{ mit } (D^\alpha f)(p) = D^\alpha f(p) \in \mathbb{R} \text{ (} p \in U \text{)}$$

das Produkt bedeutet also die Hintereinanderausführung.

Wir sagen, dass f k -mal stetig partiell differenzierbar ist und schreiben $f \in C^k(U)$, falls alle $D^\alpha f$ mit $|\alpha| \leq k$ existieren und stetig sind. $C^0(U)$ besteht aus den stetigen $f : U \rightarrow \mathbb{R}$.

Lemma 8.2 $C^k(U)$ hängt weder von der betrachteten Norm noch von der gewählten Basis ab. Die $D^\alpha f$ hängen nicht von der Norm, jedoch von der Basis ab.

Beweis: Normäquivalenz. \vec{v} habe bzgl. Ausgangsbasis die Koordinaten v_1, \dots, v_n und es sei $|\alpha| < k$. Dann ist D^α stetig partiell differenzierbar, also nach Satz 5.12 differenzierbar und nach Satz 5.8 existiert die Ableitung von D^α längs \vec{v} an p

$$d_{\vec{v}} D^\alpha f(p) = \sum_i v_i D_i D^\alpha f(p)$$

und $d_{\vec{v}} D^\alpha f \in C^{k-|\alpha|-1}$, da $D_i D^\alpha f \in C^{k-|\alpha|-1}$. \square

Lemma 8.3 Sei $f \in C^k(U)$ und $|\alpha| + |\beta| \leq k$. Dann gilt

$$D^\alpha D^\beta f = D^\beta D^\alpha f = D^{\alpha+\beta} f$$

Achtung: Die Operatoren D_i kommutieren nicht miteinander. Aber unter den Voraussetzungen des Lemmas darf man mit ihnen so rechnen, als ob sie kommutierten - d.h. sofern keine Terme von Grad $> k$ auftreten. Beweis. Wir betrachten den Fall $\alpha_j = 1$ und $\alpha_i = 0$ für $i \neq j$. Das geht mit Induktion über $|\beta|$. Ist $|\beta| = 0$, so ist trivial. Anderenfalls $k \geq 2$. Sei $\gamma = \beta - (1, 0, \dots, 0)$ und $\delta = \gamma + \alpha$. Dann nach Schwarz und Induktion

$$D_j D^\beta f = D_j D_1 D^\gamma f = D_1 D_j D^\gamma f = D_1 D^\alpha D^\gamma f = D_1 D^\delta f = D^{\alpha+\beta} f$$

Die Behauptung folgt nun mit Induktion über $|\alpha|$: Sei $\gamma = \alpha - (1, 0, \dots, 0)$. Dann

$$D^\alpha D^\beta f = D_1 D^\gamma D^\beta f = D_1 D^{\gamma+\beta} f = D^{\alpha+\beta} f \quad \square$$

Korollar 8.4

$$D^\alpha x^\beta = \frac{\beta!}{(\beta - \alpha)!} x^{\beta - \alpha}$$

$$D^\alpha x^\beta = \begin{cases} \alpha! & \text{falls } \alpha = \beta \\ 0 & \text{falls } \alpha_i > \beta_i \text{ für ein } i \end{cases}$$

8.5 Potenzreihen

Eine Potenzreihe im Entwicklungspunkt p ist gegeben als

$$f(x) = f(p+h) = \sum_{\alpha} a_{\alpha} h^{\alpha}$$

wobei $h = p - x = (h_1, \dots, h_n) = \Delta x = (\Delta x_1, \dots, \Delta x_n)$. Im Falle von $n = 2$ etwa so

$$a_{00} + a_{10}h_1 + a_{01}h_2 + a_{20}h_1^2 + a_{11}h_1h_2 + a_{02}h_2^2 + a_{30}h_1^3 + a_{21}h_1^2h_2 + a_{12}h_1h_2^2 + a_{03}h_2^3 + \dots$$

Ist $a_\alpha = 0$ für alle α mit $|\alpha| > k$ so hat man ein Polynom vom Grad $\leq k$ in der Unbestimmten h . Für Polynome $f(p+h)$ in h hat man

$$D^\alpha f(p) = \alpha! a_\alpha + \sum_{|\beta|>0} a_\beta \frac{\alpha!}{(\beta-\alpha)!} h^{\beta-\alpha}$$

also durch einsetzen von $h = 0$

$$D^\alpha f(p) = \alpha! a_\alpha, \quad a_\alpha = \frac{1}{\alpha!} D^\alpha f(p)$$

8.6 Taylorpolynom

Ist f in einer Umgebung von p k -mal stetig differenzierbar, so definieren wir das k -te Taylorpolynom (in der Unbestimmten h) als

$$T_p^k f(h) = \sum_{|\alpha| \leq k} a_\alpha h^\alpha, \quad \text{mit } a_\alpha = \frac{1}{\alpha!} D^\alpha f(p)$$

Es gelten die Regeln wie bei Taylorpolynome in 1 Variablen (und die Beweise übertragen sich wörtlich). Beispiele

$$f(x, y) = \sin x \cos y \approx_0^3 \left(x - \frac{1}{6}x^3\right) \left(1 - \frac{1}{2}y^2\right) \approx_0^3 x - \frac{1}{6}x^3 - \frac{1}{2}xy^2 \quad \text{Produktregel}$$

$$e^{x^2 + \cos y} = e \cdot e^{x^2 + \cos y - 1} \approx_0^2 e \cdot e^{x^2 - \frac{1}{2}y^2} \approx_0^2 e \left(1 + x^2 - \frac{1}{2}y^2\right) \quad \text{Kettenregel}$$

8.7 Satz von Lagrange-Taylor

Satz 8.5 Sei $f \in C^{k+1}(U)$ und sei $p+th \in U$ für alle $t \in [0, 1]$. Dann gibt es ein $\tau \in (0, 1)$ so, dass

$$f(p+h) = T_p^k(h) + R(h)$$

mit dem Restglied

$$R(h) = \sum_{|\alpha|=k+1} \frac{(D^\alpha f)(x + \tau h)}{\alpha!} h^\alpha$$

Beweis. Wir setzen $g(t) = p + th$, $t \in [0, 1]$ und erhalten nach der Kettenregel für das Taylorpolynom von $(f \circ g)(t)$ an der Stelle 0

$$\begin{aligned} T_\tau^k(f \circ g)(t) &\approx_p^k T_{p+\tau h}^k \circ T_\tau^k(th) = \sum_{|\alpha| \leq k} \frac{1}{\alpha!} (D^\alpha f(p + \tau h))(h^\alpha t^{|\alpha|}) \\ &= \sum_{l \leq k} \left[\sum_{|\alpha|=l} \frac{1}{\alpha!} (D^\alpha f(p + \tau h))(h^\alpha) \right] t^{|\alpha|} \end{aligned}$$

Koeffizientenvergleich ergibt

$$\frac{1}{l!} \left(\frac{\partial^l}{\partial t^l} (f \circ g)(\tau) \right) = \sum_{|\alpha|=l} \frac{1}{\alpha!} (D^\alpha f(p + \tau h))(h^\alpha)$$

Wenden wir den Satz von Taylor in 1 Variablen auf $(f \circ g)(t)$ und k an, so folgt

$$\begin{aligned} f(p+h) &= (f \circ g)(0+1) = T_0^k(f \circ g) + \frac{1}{(k+1)!} (f \circ g)^{(k+1)}(\tau) 1^{k+1} \\ &= T_p^k f(h) + \sum_{|\alpha|=k+1} \frac{1}{\alpha!} (D^\alpha f(p+\tau h))(h^\alpha) \end{aligned}$$

da $T_0^k(th) = h$ und somit $T_p^k f \circ T_0^k(th) = T_p^k f(h)$. \square

Korollar 8.6 Zu $f \in C^k(U)$ gibt es ϕ , in einer Umgebung des Ursprungs definiert, mit

$$f(p+h) = T_p^k f(h) + \phi(h), \quad \lim_{h \rightarrow 0} \frac{\phi(h)}{|h|^k} = 0$$

kurz

$$f(p+h) = T_p^k f(h) + o(|h|^k)$$

Beweis. Setze

$$\begin{aligned} \phi(h) &= f(p+h) - T_p^k f = T_p^k f + R - T_p^k f \\ \frac{|\phi(h)|}{|h|^k} &\leq \sum_{|\alpha|=k+1} \frac{1}{\alpha!} |D^\alpha f(p+\tau h) - D^\alpha f(p)| \frac{|h^\alpha|}{|h|^k} \rightarrow 0 \end{aligned}$$

da aus $|h_i| \leq |h|$ folgt dass $|h^\alpha| = |h_1|^{\alpha_1} \cdots |h_n|^{\alpha_n} \leq |h|^{|\alpha|}$.

8.8 Beispiele

Beispiel 9 Sei $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ gegeben durch $f(x, y) = \sin(x+2y)$. Dann ist

$$f_x(x, y) = \cos(x+2y), \quad f_y(x, y) = 2 \cos(x+2y)$$

$$f_{xx}(x, y) = -\sin(x+2y), \quad f_{xy}(x, y) = -2 \sin(x+2y), \quad f_{yy} = -4 \sin(x+2y),$$

und der Satz von Taylor für die Ordnung 1 liefert

$$f(x+h, y+k) = f(x, y) + f_x(x, y)h + f_y(x, y)k + R$$

bzw.

$$\sin(x+h+2(y+k)) = \sin(x+2y) + \cos(x+2y)h + 2 \cos(x+2y)k + R$$

mit dem Restglied

$$\begin{aligned} R &= \frac{1}{2} (h^2 f_{xx}(\bar{x}, \bar{y}) + 2hk f_{xy}(\bar{x}, \bar{y}) + k^2 f_{yy}(\bar{x}, \bar{y})) \\ &= -\frac{\sin(\bar{x}+2\bar{y})}{2} (h^2 + 4hk + 4k^2) \end{aligned}$$

mit $\bar{x} = x + \tau h$, $\bar{y} = y + \tau k$ mit einem $\tau \in (0, 1)$. Eine grobe Abschätzung mit $|\sin t| \leq 1$ ergibt

$$|R| \leq \frac{1}{2} (h^2 + 4|hk| + 4k^2) \leq 2|h|^2$$

Beispiel 10 Wir bestimmen das Taylorpolynom 2. Ordnung für die Funktion

$$f : \mathbb{R}^2 \rightarrow \mathbb{R}, \quad (x, y) \mapsto e^{x^2 + \cos y}$$

im Punkt $(0, 0)$. Dieses ist gleich (mit $D_1 = \frac{\partial}{\partial x}$ und $D_2 = \frac{\partial}{\partial y}$)

$$\begin{aligned} f(0, 0) + (D_1 f)(0, 0)h + (D_2 f)(0, 0)k + \frac{1}{2}(D_1^2 f)(0, 0) \cdot h^2 \\ + \frac{1}{2}(D_2^2 f)(0, 0)k^2 + (D_1 D_2 f)(0, 0)hk. \end{aligned}$$

Die partiellen Ableitungen von f bis zur 2. Ordnung sind

$$\begin{aligned} (D_1 f)(x, y) &= 2x e^{x^2 + \cos y} && \Rightarrow && (D_1 f)(0, 0) &= 0, \\ (D_2 f)(x, y) &= -\sin y e^{x^2 + \cos y} && \Rightarrow && (D_2 f)(0, 0) &= 0, \\ (D_1^2 f)(x, y) &= (2 + 4x^2)e^{x^2 + \cos y} && \Rightarrow && (D_1^2 f)(0, 0) &= 2e, \\ (D_2^2 f)(x, y) &= (-\cos y + \sin^2 y)e^{x^2 + \cos y} && \Rightarrow && (D_2^2 f)(0, 0) &= -e, \\ (D_1 D_2 f)(x, y) &= -2x \sin y e^{x^2 + \cos y} && \Rightarrow && (D_1 D_2 f)(0, 0) &= 0. \end{aligned}$$

Das gesuchte Taylorpolynom (das wir oben mit der Kettenregel einfacher bestimmt haben) ist also

$$(h, k) \mapsto e + eh^2 - \frac{e}{2}k^2. \quad \blacksquare$$

9 Quadratische Approximation

9.1 Bilinearformen

Im Folgenden sei V ein endlichdimensionaler reeller Vektorraum. Seine Elemente schreiben wir als \vec{v} . Eine *Bilinearform* auf V ist eine Abbildung $\Phi : V \times V \rightarrow \mathbb{R}$ mit

$$\begin{aligned} \Phi(\vec{v} + \vec{w}, \vec{u}) &= \Phi(\vec{v}, \vec{u}) + \Phi(\vec{w}, \vec{u}) & \Phi(\vec{u}, \vec{v} + \vec{w}) &= \Phi(\vec{u}, \vec{v}) + \Phi(\vec{u}, \vec{w}) \\ \Phi(r\vec{v}, \vec{w}) &= r\Phi(\vec{v}, \vec{w}) & \Phi(\vec{v}, r\vec{w}) &= r\Phi(\vec{v}, \vec{w}) \end{aligned}$$

Sei $\alpha : \vec{e}_1, \dots, \vec{e}_n$ eine Basis von V . Dann gibt es eine bijektive Entsprechung zwischen Bilinearformen Φ auf V und Matrizen $A \in \mathbb{R}^{n \times n}$ vermöge

$$\boxed{\Phi(\vec{x}, \vec{y}) = (\vec{x}^\alpha)^t A \vec{y}^\alpha, \quad A = (\Phi(\vec{e}_i, \vec{e}_j))_{n \times n}}$$

(die transponierte Matrix zu X schreiben wir als X^t). $A = \Phi_\alpha$ heisst dann die (*Gram*)-*Matrix* von Φ bzgl. α . Die Form Φ ist *symmetrisch*, d.h.

$$\Phi(x, y) = \Phi(y, x) \text{ für alle } x, y \in V \quad \Leftrightarrow \quad A^t = A$$

Beispiel: Skalarprodukte.

9.2 Quadratische Formen

Die zugehörige *quadratische Form* $Q : V \rightarrow \mathbb{R}$ ist gegeben durch

$$\begin{aligned} Q(\vec{x}) &= \Phi(\vec{x}, \vec{x}) = (\vec{x}^\alpha)^t A \vec{x}^\alpha = \sum_{ij} a_{ij} x_i x_j && \text{mit } \vec{x}^\alpha = (x_1, \dots, x_n)^t \\ &= \sum_i q_{ii} x_i^2 + \sum_{i < j} q_{ij} x_i x_j && \text{mit } q_{ii} = a_{ii} \text{ und } q_{ij} = 2a_{ij} \end{aligned}$$

und muss dann, wenn man die q_{ij} und die Matrix A angeben will beachten, dass

$$a_{ij} = \begin{cases} q_{ii} & \text{falls } i = j \\ \frac{1}{2} q_{ij} & \text{falls } i < j \text{ bzw. } j < i \end{cases}$$

Es gilt

$$Q(\lambda \vec{x}) = \lambda^2 Q(\vec{x})$$

Und kann man Φ aus Q zurückgewinnen:

$$\Phi(\vec{x}, \vec{y}) = \frac{1}{2} (Q(\vec{x} + \vec{y}) - Q(\vec{x}) - Q(\vec{y})).$$

Entsprechend erhält man die Matrix A direkt aus Q

$$a_{ii} = Q(\vec{e}_i), \quad a_{ij} = \frac{1}{2} (Q(\vec{e}_i + \vec{e}_j) - Q(\vec{e}_i) - Q(\vec{e}_j)) \quad i \neq j$$

Eine reelle quadratische Form auf \mathbb{R}^n ist natürlich eine reelle Funktion in n Variablen und man veranschaulicht sich sie durch Niveau-Hyperflächen, für $n = 2$ also durch Höhenlinien.

Lemma 9.1 *Sind Q_1 und Q_2 quadratische Formen, so auch $r_1 Q_1 + r_2 Q_2$ mit*

$$(r_1 Q_1 + r_2 Q_2)(\vec{x}) = r_1 Q_1(\vec{x}) + r_2 Q_2(\vec{x})$$

Beweis. $r_1 \Phi_1 + r_2 \Phi_2$ symmetrisch und bilinear, falls es die Φ_i sind. \square

9.3 Quadratische Abbildungen

Sei V endlichdimensionaler \mathbb{R} -Vektorraum. Eine *quadratische Abbildung* ist von der Form

$$q : V \rightarrow \mathbb{R}, \quad q(\vec{x}) = L(\vec{x}) + Q(\vec{x})$$

mit $L : V \rightarrow \mathbb{R}$ linear und Q quadratische Form.

Lemma 9.2 *Sei V endlichdimensional, $D \subseteq V$ und $f : D \rightarrow \mathbb{R}$*

1. *Gilt $\lim_{n \rightarrow \infty} \frac{f(\vec{x}_n)}{\|\vec{x}_n\|^2} = 0$ bzgl. einer Norm von V , so auch bzgl. jeder anderen,*
2. *Ist $q : V \rightarrow \mathbb{R}$ quadratisch und $\frac{q(\vec{x})}{\|\vec{x}\|^2} \rightarrow 0$ für $\vec{x} \rightarrow \vec{0}$ in einer Umgebung von $\vec{0}$, so ist q konstant 0.*

Beweis. 1. Sei die Konvergenz bzgl. einer Norm $|\cdot|$ vorausgesetzt und $\|\cdot\|$ eine weitere Norm. Wegen der Äquivalenz der Normen gibt es $a, b > 0$ mit $a|\vec{x}| \leq \|\vec{x}\| \leq b|\vec{x}|$ und es folgt

$$\frac{1}{b^2} \frac{|f(\vec{x})|}{|\vec{x}|^2} \leq \frac{|f(\vec{x})|}{\|\vec{x}\|^2} \leq \frac{1}{a^2} \frac{|f(\vec{x})|}{|\vec{x}|^2}$$

Zu 2. Wähle euklidische Norm und orthonormale Basis. Seien (a_1, \dots, a_n) und (a_{ij}) die Matrizen von L und Q . Für $\vec{x} = t\vec{e}_i$ gilt $q(\vec{x})/|\vec{x}| = a_i + ta_{ii} \rightarrow a_i$ für $t \rightarrow 0$, also $a_i = 0$. Nun

$$\frac{Q(\vec{x})}{|\vec{x}|^2} = a_{ii}$$

also $a_{ii} = 0$. Sei schliesslich $\vec{x} = t(\vec{e}_i + \vec{e}_j)$. Dann

$$t^2 a_{ij} = \Phi(t\vec{e}_i, t\vec{e}_j) = \frac{1}{2}(Q(\vec{x}) - Q(t\vec{e}_i) - Q(t\vec{e}_j)) = \frac{1}{2}Q(\vec{x})$$

und somit $a_{ij} = 0$. \square

9.4 Quadratische Approximation von Skalarenfeldern

Sei $p \in D \subseteq \mathbb{R}^n$ offen. $f : D \rightarrow \mathbb{R}$ ist an p *quadratisch approximierbar*, wenn es eine quadratische Abbildung $q : \mathbb{R}^n \rightarrow \mathbb{R}$ so gibt, dass

$$f(p + \Delta\vec{x}) = f(p) + q(\Delta\vec{x}) + R(\Delta\vec{x}) \quad \text{mit} \quad \frac{R(\Delta\vec{x})}{\|\Delta\vec{x}\|^2} \rightarrow 0 \text{ für } \Delta\vec{x} \rightarrow \vec{0}$$

Nach Lemma 2 ist $q = L + Q$ eindeutig bestimmt. Da L eindeutig bestimmt ist, ist es auch Q . Die quadratische Form $2Q$ ist die *Hesseform* $Hf(p) = H_p f$.

Korollar 9.3 *Sei $f \in C^2(U)$ und $p \in U$. Dann ist f an p quadratisch approximierbar und die Hesseform $Hf(p)$ ist bzgl. dieser Basis gegeben durch die Matrix*

$$Hessf(p) = \left(\frac{\partial^2 f}{\partial x_i \partial x_j}(x) \right)_{i,j=1}^n$$

Beweis. Nach dem Korollar zum Satz von Lagrange-Taylor gibt es ϕ mit $\frac{\phi(h)}{\|h\|^2} \rightarrow 0$ für $h \rightarrow 0$ und

$$\begin{aligned} f(p+h) &= f(p) + \sum_i \frac{\partial f}{\partial x_i}(p) h_i + \frac{1}{2} \sum_i \frac{\partial^2 f}{\partial x_i^2}(p) h_i^2 + \sum_{i < j} \frac{\partial^2 f}{\partial x_i \partial x_j}(p) h_i h_j + \phi(h) \\ &= f(p) + \sum_i \frac{\partial f}{\partial x_i}(p) h_i + \frac{1}{2} \sum_{i,j} \frac{\partial^2 f}{\partial x_i \partial x_j}(p) h_i h_j + \phi(h) \\ &= f(p) + \nabla f(p) \cdot h + \frac{1}{2} h^t (Hessf)(p) h + \phi(h) \end{aligned}$$

also

$$f(p+h) = f(p) + df(p)(h) + \frac{1}{2} Hf(p)(h) \quad \square$$

Man kann dies auf $f \in C^k(U)$ verallgemeinern:

$$f(p+h) = f(p) + \sum_{i=1}^k \Phi_k(h, \dots, h) + \phi(h), \quad \frac{\phi(h)}{|h|^k} \rightarrow 0 \text{ für } h \rightarrow 0$$

mit k -linearen und symmetrischen Multilinearformen $\Phi_k : V^k \rightarrow \mathbb{R}$

$$\Phi_k(\vec{v}_1, \dots, \vec{v}_i + r\vec{w}, \dots, \vec{v}_v) = \Phi_k(\vec{v}_1, \dots, \vec{v}_i, \dots, \vec{v}_v) + r\Phi_k(\vec{v}_1, \dots, r\vec{w}, \dots, \vec{v}_v)$$

$$\Phi_k(\vec{v}_{\sigma(1)}, \dots, \vec{v}_{\sigma(n)}) = \Phi_k(\vec{v}_1, \dots, \vec{v}_n) \text{ für jeder Permutation } \sigma$$

9.5 Hesse-Form und zweifache Differenzierbarkeit

Sei Q quadratische Form mit Matrix A . dann ist $2A$ die Hesse-Matrix von Q . Andererseits ist Q beliebig oft stetig partiell differenzierbar mit

$$F(x) := \mathbf{grad}Q(x) \hat{=} 2 \begin{pmatrix} \sum_j a_{1j}x_j \\ \vdots \\ \sum_j a_{nj}x_j \end{pmatrix}$$

und die vektorwertige Abbildung F hat Jacobimatrix

$$J_F(x) = 2A$$

Umgekehrt und allgemeiner sei $f \in C^2(U)$. Dann sind die Abbildungen

$$x \mapsto F(x) = \mathbf{grad}f(x) \mapsto Hf(x)$$

stetig, also $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$ differenzierbar und $DF(x) = Hf(x)$. Insbesondere ist also f zweimal differenzierbar und die Jacobimatrix der zweiten Ableitung ist die Hesse-Matrix - das folgt schon aus der Differenzierbarkeit von F . Diese reicht aber sogar aus, um die Symmetrie zu beweisen. Für die quadratische Approximation von f mussten wir aber $f \in C^2(U)$ voraussetzen und Taylor-Lagrange verwenden.

Korollar 9.4 Sei U offene Teilmenge des normierten Vektorraums V und $p \in U$. Die Hesseform $Hf(p)$ einer Abbildung $f \in C^2(U)$ and hängt nicht von der gewählten Basis von V ab.

Beweis. Sei eine Basis e_1, \dots, e_n gegeben und $v = \sum_j v_j e_j$, $w = \sum_i w_i e_i$. Als Funktion von $x \in U$ haben wir

$$d_v(x) = \sum_j v_j \frac{\partial f}{\partial x_j}(x) \in R$$

die Ableitung von f längs v an x . Die Ableitung von d_v längs w an p ist dann

$$\begin{aligned} d_w(d_v f)(p) &= \sum_i w_i \left(\frac{\partial d}{\partial x_i} d_v f \right)(p) = \sum_i w_i \left(\frac{\partial d}{\partial x_i} \sum_j v_j \frac{\partial f}{\partial x_j} \right)(p) \\ &= \sum_{i,j} w_i v_j \left(\frac{\partial d}{\partial x_i} \frac{\partial f}{\partial x_j} \right)(p) = \sum_{i,j} w_i v_j \left(\frac{\partial^2 f}{\partial x_i \partial x_j} \right)(p) \end{aligned}$$

□.

10 Quadratische Formen

10.1 Transformation

Ist β eine weitere Basis, so werden die Koordinaten von Vektoren transformiert nach dem Ansatz

$$\vec{x}^\alpha = {}^\alpha T_\beta \vec{x}^\beta$$

wobei in den Spalten der Transformationsmatrix ${}^\alpha T_\beta$ die Koordinaten α stehen. Für die Basisbeschreibungen der Bilinearform Φ gilt dann

$$\Phi_\beta = {}^\alpha T_\beta^t \Phi_\alpha {}^\alpha T_\beta$$

Beweis. $\vec{x}^\alpha = {}^\alpha T_\beta \vec{x}^\beta$, $\vec{y}^\alpha = {}^\alpha T_\beta \vec{y}^\beta$, also

$$(\vec{x}^\beta)^t {}^\alpha T_\beta^t \Phi_\alpha {}^\alpha T_\beta \vec{y}^\beta = ({}^\alpha T_\beta \vec{x}^\beta)^t \Phi_\alpha \vec{y}^\alpha = (\vec{x}^\alpha)^t \Phi_\alpha \vec{y}^\alpha = \Phi(\vec{x}, \vec{y})$$

Für die quadratische Form $Q(\vec{x}) = \Phi(\vec{x}, w\vec{x})$ folgt

- Ist A die Matrix von Q bzgl. α , so hat Q bzgl. β die Matrix ${}^\alpha T_\beta^t A {}^\alpha T_\beta$.

10.2 Definitheit

Eine quadratische Form Q heisst

$$\begin{array}{ll} \text{positiv definit} & \Leftrightarrow Q(\vec{x}) > 0 \forall \vec{x} \neq \vec{0} \\ \text{negativ definit} & \Leftrightarrow Q(\vec{x}) < 0 \forall \vec{x} \neq \vec{0}, \\ \text{nichtausgeartet} & \Leftrightarrow Q(\vec{x}) \neq 0 \forall \vec{x} \neq \vec{0} \end{array} \quad \begin{array}{ll} \text{positiv semi-definit} & \Leftrightarrow Q(\vec{x}) \geq 0 \forall \vec{x} \\ \text{negativ semi-definit} & \Leftrightarrow Q(\vec{x}) \leq 0 \forall \vec{x} \\ \text{indefinit} & \Leftrightarrow \exists \vec{x}, \vec{y}. Q(\vec{x}) > 0 > Q(\vec{y}) \end{array}$$

Eine symmetrische Matrix heisst *X-definit*, falls sie bzgl. einer (und dann jeder) Basis eine X-definite Form definiert. Natürlich ist A genau dann negativ (semi) definit, wenn $-A$ positiv (semi) definit ist.

10.3 Symmetrische Diagonalisierung

Satz 10.1 Zu jeder reellen symmetrischen Bilinearform Φ auf einem endlichdimensionalen \mathbb{R} -Vektorraum gibt es eine Basis β so, dass Φ_β Diagonalmatrix ist

$$\Phi(\vec{x}, \vec{y}) = \sum_i d_i x_i y_i$$

Dabei ist die Anzahl p der $d_i > 0$ und die Anzahl q der $d_i < 0$ eindeutig bestimmt (Trägheitssatz von Sylvester).

Zu jedem symmetrischen $A \in \mathbb{R}^{n \times n}$ gibt es invertierbares S so, dass $S^t A S$ diagonal

Warnung: Die d_i in der symmetrischen Diagonalisierung sind in der Regel keine Eigenwerte von Q . Berechnung von Eigenwerten ist eine Aufgabe der Numerik.

Beweis der Existenz durch systematische Transformation: Wir machen jeweils eine Zeilenumformung T gefolgt von der analogen Spaltenumformung T^t bis wir am Ziel sind. Dabei bleibt die Symmetrie zu erhalten. \square

$$\begin{pmatrix} 4 & -2 & 2 \\ -2 & 10 & 2 \\ 2 & 2 & 4 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 4 & 0 & 0 \\ 0 & 9 & 3 \\ 0 & 3 & 3 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 4 & 0 & 0 \\ 0 & 9 & 0 \\ 0 & 0 & 2 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Korollar 10.2 Φ ist

$$\begin{array}{llll} \text{positiv definit} & \Leftrightarrow & p = n & \text{positiv semi-definit} & \Leftrightarrow & q = 0 \\ \text{negativ definit} & \Leftrightarrow & q = n, & \text{negativ semi-definit} & \Leftrightarrow & p = 0 \\ \text{nichtausgeartet} & \Leftrightarrow & p + q = n & \text{indefinit} & \Leftrightarrow & p > 0, q > 0 \end{array}$$

10.4 Symmetrischer Gaußalgorithmus

Man kombiniere das Transformationsschema mit dem Zeilen-Gaußalgorithmus zur Bestimmung einer oberen Stufenform. Man erhält symmetrische A_k , also

$$A_k = \begin{pmatrix} D_k & O \\ O & + \end{pmatrix}, \quad A_{k+1} = T_{k+1} A_k T_{k+1}^t, \quad S_0 = E, \quad S_{k+1} = S_k T_{k+1}^t$$

wobei D_k Diagonalmatrix. Sind alle Hauptminoren $\neq 0$, so gilt stets $a_{k+1,k+1}^{(k)} \neq 0$, d.h. man kann die T_{k+1} als untere und somit alle S_k als obere Dreiecksmatrizen wählen.

► Gegeben sei $A_k = \begin{pmatrix} D_k & O \\ O & \begin{pmatrix} b_{k+1,k+1} & \dots & b_{k+1,n} \\ \vdots & \ddots & \vdots \\ b_{n,k+1} & \dots & b_{nn} \end{pmatrix} \end{pmatrix}$

- Ist $b_{k+1,k+1} \neq 0$, so bewirke durch Zeilenscherungen, dass die neue $k+1$ -te Zeile von 'A' ausserhalb der Diagonalen nur noch Nullen enthält. Ändere die Werte über der Diagonalen von 'A' so, dass wieder eine symmetrische Matrix entsteht
- Ist $b_{k+1,k+1} = 0$ so suche vorher $i > k+1$ mit $b_{i,i} \neq 0$ und vertausche $k+1$ -te und i -te Zeile in 'A', ebenso für Spalten von 'A'
- Ist $b_{ii} = 0$ für alle $i > k$, so suche vorher $j > i > k$ mit $b_{ij} \neq 0$. Addiere in 'A' die j -Zeile zur i -ten und die j -te Spalte zur i -ten. Das ergibt i -ten Diagonaleintrag $b_{ij} + b_{ij} \neq 0$.
- Sind alle $b_{ij} = 0$ für $j > i > k$, so ist es getan.

10.5 Minoren

Der k -te Hauptminor $A_{\leq k}$ einer Matrix A ist die Matrix

$$A_k = \begin{pmatrix} a_{11} & \dots & a_{1k} \\ \vdots & & \vdots \\ a_{k1} & \dots & a_{kk} \end{pmatrix}$$

10.6 Trägheitssatz

Beweis der Eindeutigkeit von p, q . Man kann immer zu einer Basis β übergehen so, dass

$$\Phi_\beta = \begin{pmatrix} E_p & O & O \\ O & -E_q & O \\ O & O & O \end{pmatrix}$$

- man ersetze \vec{v}_i durch $\frac{1}{\sqrt{|d_i|}}$ für $i \leq p+q$. Daher genügt es, folgende Situation zu betrachten.

$$U^t A U = A' = \begin{pmatrix} E_p & O & O \\ O & -E_q & O \\ O & O & O \end{pmatrix}, \quad V^t A V = A'' = \begin{pmatrix} E_k & O & O \\ O & -E_l & O \\ O & O & O \end{pmatrix}$$

mit invertierbaren U, V . Dann $A' = S^t A'' S$ mit $S = V^{-1}U$. Also O.B.d.A. $A = A''$ und $p \leq k$. Multiplikation mit invertierbaren Matrizen ändert den Rang nicht. Somit $r = p + q = k + l$. Seien B, B', T die r -Hauptminoren von A bzw. A' bzw. S . Dann $B' = T^t B T$. Daher o.B.d.A. $r = n$. Sei Q die durch A bzgl. der kanonischen Basis α definierte quadratische Form und $\beta : \mathbf{v}_1, \dots, \mathbf{v}_n$ die Basis mit $S = {}^\alpha T_\beta$. Setze

$$U_1 = \text{Spann}\{\mathbf{e}_1, \dots, \mathbf{e}_k\}, \quad V_2 = \text{Spann}\{\mathbf{v}_{p+1}, \dots, \mathbf{v}_n\}$$

Dann ist Q auf U_1 positiv definit, auf V_2 negativ definit. Es folgt

$$U_1 \cap V_2 = \{0\}$$

und es ist $\mathbf{e}_1, \dots, \mathbf{e}_k, \mathbf{v}_{p+1}, \dots, \mathbf{v}_n$ linear unabhängig. Daher $k \leq p$ und somit $p = k$. \square

10.7 Determinanten

\det ist eine Abbildung, die jeder $n \times n$ -Matrix A den Skalar $\det A$ zuordnet. Es gilt

- \det ist linear als Funktion einer gegebenen Spaltenposition bzw. Zeilenposition.
- $\det B = -\det A$, wenn B aus A hervorgeht, indem man zwei Spalten (bzw. zwei Zeilen) vertauscht.
- $\det A = 0$ falls eine Spalte (Zeile) Linearkombination der anderen ist
- $\det A \neq 0 \Leftrightarrow A$ invertierbar
- $\det(AB) = \det A \cdot \det B$
- $\det A^{-1} = (\det A)^{-1}$ falls A invertierbar
- $\det S^{-1} A S = \det A$ falls S invertierbar
- $\det A^t \det A$
- $\det A = a_{11}a_{22} - a_{21}a_{12}$ falls $n = 2$
- $\det A = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{31}a_{22}a_{13} - a_{32}a_{23}a_{11} - a_{33}a_{21}a_{12}$ falls $n = 3$

10.8 Hauptminorenkriterium

Satz 10.3 Hurwitz. *Für eine symmetrische Matrix A sind gleichwertig:*

- A ist positiv definit.
- Es gibt eine invertierbare Matrix W mit $A = WW^t$.
- Die Hauptminoren $A_{\leq k} = (a_{ij})_{1 \leq i, j \leq k}$ von A haben $\det A_{\leq k} > 0$.

Beweis. Die Äquivalenz von (1) und (2) ist klar nach dem Trägheitssatz. Also ist für positiv definites A die Determinante $\det A = \det(W^tW) = \det W^t \det W = |\det W|^2 > 0$. Andererseits sind für positiv definites A alle Hauptminoren $A_{\leq k}$ positiv definit: für die durch $A_{\leq k}$ definierte quadratische Form $Q_{\leq k}$ gilt: $Q_{\leq k}(x_1, \dots, x_k) = Q(x_1, \dots, x_k, 0, \dots, 0) > 0$ falls ein $x_i \neq 0$. Also $\det A_{\leq k} > 0$ für alle k .

Sei nun (3) vorausgesetzt, insbesondere $a_{11} \neq 0$. Also hat man im symmetrischen Gaußalgorithmus sofort $A \rightsquigarrow B = TAT^t$ und es bleibt nur zu zeigen, dass B wieder die Voraussetzung erfüllt. Man beachte, dass für die Hauptminoren gilt $B_{\leq k} = T_{\leq k}A_{\leq k}T_k^t$ d.h. sie haben Determinante $|\det T_{\leq k}|^2 \det A_{\leq k} > 0$. Ist C der $n-1 \times n-1$ -Minor $C = (b_{ij})_{1 < i, j}$ so ist $C_{\leq k}$ Minor von $B_{\leq k+1}$ und es gilt $\det(B_{\leq k+1}) = a_{11} \det C_{\leq k}$. Also hat C nur positive Hauptminoren und das Verfahren führt zu einer Diagonalmatrix mit nur positiven Einträgen, also schliesslich mit dem Produkt S der Matrizen aus den Einzelschritten zu $SAS^t = E_n$ und $A = WW^t$ mit $W = S^{-1}$.

11 Lokale Extrema

11.1 Definitionen

Sei $p \in U \subseteq V = \mathbb{R}^n$ offen, $f : U \rightarrow \mathbb{R}$ stetig partiell differenzierbar. Sei $\|\cdot\|$ Norm auf V

- f hat an p ein (*striktes*) *lokales Maximum*, falls $f(x) \leq f(p)$ ($f(x) < f(p)$) für alle x in einer Umgebung $U' \subseteq U$ von p
- f hat an p ein (*striktes*) *lokales Minimum*, falls $f(x) \geq f(p)$ ($f(x) > f(p)$) für alle x in einer Umgebung $U' \subseteq U$ von p
- p ist *kritischer* oder *stationärer Punkt* von f , falls $df(p) = 0$
- f ist bei p *lokal positiv definit* falls es $\varepsilon > 0$ und $\lambda > 0$ gibt mit

$$f(p+h) \geq f(p) + \lambda \|h\|^2 \quad \text{für } \|h\| < \varepsilon$$

- f ist bei p *lokal negativ definit* falls es $\varepsilon > 0$ und $\lambda < 0$ gibt mit

$$f(p+h) \leq f(p) + \lambda \|h\|^2 \quad \text{für } \|h\| < \varepsilon$$

- f ist *lokal indefinit* an p falls es $v_+ \neq 0$ und $v_- \neq 0$ gibt mit mit $f|_{\{p + tv_+ \mid t \in \mathbb{R}\}}$ lokal positiv, $f|_{\{p + tv_- \mid t \in \mathbb{R}\}}$ lokal negativ definit

Liegt Maximum oder Minimum vor, so spricht man auch von *Extremum*. Die obigen Aussagen beschreiben das *lokale Extremwertverhalten* von f an p .

Korollar 11.1 1. *Liegt bei p ein lokales Extremum vor, so gilt $df(p) = 0$*

2. *Ist f bei p lokal positiv (negativ) definit, so hat f bei p ein striktes lokales Minimum (Maximum)*

3. *Ist f bei p lokal indefinit, so hat f bei p kein lokales Extremum.*

Beweis. Hat f an p ein lokales Extremum, so hat für jeden Vektor \vec{v} die Funktion

$$t \mapsto f(p + t\vec{v})$$

an 0 ein lokales Extremum, also Ableitung

$$d_{\vec{v}}f(p) = 0$$

Die weiteren Aussagen sind trivial. \square

11.2 Extrema quadratischer Formen

Satz 11.2 *Sei Q quadratische Form.*

1. *Es sind äquivalent*

(i) *Q hat an 0 ein striktes lokales Minimum (Maximum)*

(ii) *Q ist positiv (negativ) definit*

(iii) *Q ist an 0 lokal positiv (negativ) definit*

2. *Q ist indefinit genau dann, wenn Q an 0 lokal indefinit.*

3. *Ist Q nicht ausgeartet, so ist ein Extremum an 0 stets strikt.*

Beweis. Ist Q an 0 lokal X-definit, so offensichtlich auf V X-definit. Zum Beweis von (ii) \Rightarrow (iii) und 3. wähle man Basis nach Satz 10.1 und sortiere so um, dass

$$d_1 \geq \dots d_p > 0 > d_{p+1} \geq \dots \geq d_{p+q}, d_{p+q+1} = \dots = d_n = 0$$

Diese ist orthonormal bzgl. passenden Skalarprodukts und, wegen der Äquivalenz der Normen o.B.d.A. die Norm die euklidische dazu. Setze

$$\lambda_+ = \begin{cases} d_p & \text{falls } p > 0 \\ 1 & \text{sonst} \end{cases} \quad \lambda_- = \begin{cases} d_{p+1} & \text{falls } q > p \\ -1 & \text{sonst} \end{cases}$$

Es folgt

$$Q(\vec{x}) = Q\left(\sum_{i=1}^p x_i \vec{v}_i\right) = \sum_{i=1}^p d_i x_i^2 \geq \sum_{i=1}^p \lambda_+ x_i^2 = \lambda_+ \sum_{i=1}^p x_i^2 = \lambda_+ |\vec{x}|^2 \quad \text{für } \vec{x} = \sum_{i=1}^p x_i \vec{v}_i$$

Entsprechend für λ_- und 0.

Liegt ein striktes lokales Extremum an 0 vor, so muss $p + q = n$ gelten: dieses hat den Wert 0 und man hätte im Falle $p + q < n$ den Wert $0 = Q(t\vec{v}_n)$ für alle t . Umgekehrt gilt im nichtausgearteten Fall $p + q = n$, bei Vorliegen eines Extremums also Definitheit und damit Striktheit. \square

11.3 Hinreichendes Kriterium

Satz 11.3 Sei $f \in C^2(U)$. $p \in U$ ein kritischer Punkt und $Hf(p)$ nicht ausgeartet. Dann hat f an p dasselbe Extremwertverhalten wie die Hesse-Form $Hf(p)$ an 0.

Beweis. Wir haben die quadratische Approximation

$$f(p+h) = f(p) + Hf(p)(h) + \phi(h), \quad \lim_{h \rightarrow 0} \frac{\phi(h)}{\|h\|^2} = 0$$

da $df(p) = 0$. Die Behauptung folgt nun sofort aus dem folgenden. \square

Lemma 11.4 Sei \hat{f} positiv bzw. negativ definit bzw. indefinit an 0. $\hat{f}(0) = 0$ und

$$f(p+h) = f(p) + \hat{f}(h) + \phi(h), \quad \lim_{h \rightarrow 0} \frac{\phi(h)}{\|h\|^2} = 0$$

Dann ist f an p positiv bzw. negativ definit bzw. indefinit.

Beweis für positiv. Es gibt λ_1 und ε_1 mit

$$\frac{1}{\|h\|^2} \hat{f}(h) + \frac{1}{\|h\|^2} \phi(h) \geq \lambda_1 \quad \text{falls } \|h\| < \varepsilon_1$$

für ein $\varepsilon_1 \leq \varepsilon$, da der erste Summand $\geq \lambda$ ist und der zweite gegen 0 geht. Es folgt $f(p+h) \geq f(p) + \lambda_1 \|h\|^2$. \square

11.4 Beispiele

$$f(x, y) = -x^4 - y^4 + 2x^2 + 2y^2$$

$$\text{grad } f = 4 \begin{pmatrix} -x^3 + x \\ -y^3 + y \end{pmatrix}, \quad Hf = 4 \begin{pmatrix} -3x^2 + 1 & 0 \\ 0 & -3y^2 + 1 \end{pmatrix}$$

Kritische Punkte mit $x, y \in \{0, 1, -1\}$.

$$p = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad Hf(p) = 4 \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \text{ positiv definit, Minimum}$$

$$p = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad Hf(p) = 4 \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} \text{ negativ definit, Maximum}$$

$$p = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad Hf(p) = 4 \begin{pmatrix} -2 & 0 \\ 0 & 1 \end{pmatrix} \text{ indefinit, Sattelpunkt}$$

$$f(x, y, z) = \cos x \cos(y+z) \approx_0^2 \left(1 - \frac{1}{2}x^2\right) \left(1 - \frac{1}{2}(y+z)^2\right) \approx_0^2 1 - \frac{1}{2}x^2 - \frac{1}{2}y^2 - yz - \frac{1}{2}z^2$$

$$H = -\frac{1}{2} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \end{pmatrix}$$

ist negativ semidefinit aber nicht definit. Es liegt ein Maximum vor, das kann aber nicht mit dem Kriterium begründet werden.

$$f(x, y, z) = \cos x \cos(y+z) + \cos z \approx_0^2 \left(1 - \frac{1}{2}x^2\right) \left(1 - \frac{1}{2}(y+z)^2\right) + 1 - \frac{1}{2}z^2 \approx_0^2 2 - \frac{1}{2}x^2 - \frac{1}{2}y^2 - yz - z^2$$

$$H = -\frac{1}{2} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 2 \end{pmatrix} \rightsquigarrow -\frac{1}{2} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

negativ definit, also Maximum.

Beispiel 11

- Sei $f(x, y) = \sin x \cdot \sin y$ auf \mathbb{R}^2 . Dann ist

$$f_x\left(\frac{\pi}{2}, \frac{\pi}{2}\right) = \cos \frac{\pi}{2} \sin \frac{\pi}{2} = 0 \quad \text{und} \quad f_y\left(\frac{\pi}{2}, \frac{\pi}{2}\right) = 0.$$

Also ist $p = \left(\frac{\pi}{2}, \frac{\pi}{2}\right)$ ein stationärer Punkt und sogar ein lokales Maximum, da $f_{xx}\left(\frac{\pi}{2}, \frac{\pi}{2}\right) = -\sin \frac{\pi}{2} \cdot \sin \frac{\pi}{2} = -1 < 0$ und

$$\left(f_{xy}\left(\frac{\pi}{2}, \frac{\pi}{2}\right)\right)^2 = \cos^2 \frac{\pi}{2} \cos^2 \frac{\pi}{2} = 0 < 1 = f_{xx}\left(\frac{\pi}{2}, \frac{\pi}{2}\right) f_{yy}\left(\frac{\pi}{2}, \frac{\pi}{2}\right).$$

- Wir suchen alle lokalen Extrema von $f(x, y) = x^3 - 12xy + 8y^3$ auf \mathbb{R}^2 . Wegen $f_x(x, y) = 3x^2 - 12y$ und $f_y(x, y) = 24y^2 - 12x$ ergeben sich alle stationären (also extremwertverdächtigen) Punkte von f aus

$$3x^2 - 12y = 0 \quad \text{und} \quad 24y^2 - 12x = 0.$$

Einsetzen von $y = \frac{1}{4}x^2$ (aus der ersten Gleichung) in die zweite Gleichung ergibt $\frac{1}{8}x^4 = x$. Diese Gleichung hat genau zwei reelle Lösungen, nämlich $x_0 = 0$ und $x_1 = 2$. Hieraus erhält man mit der ersten Gleichung $y_0 = 0$ und $y_1 = 1$. Demnach sind $(0, 0)$ und $(2, 1)$ die einzigen Kandidaten für lokale Extremstellen. Nun ist

$$\left(f_{xy}(0, 0)\right)^2 = (-12)^2 = 144 > 0 = f_{xx}(0, 0) f_{yy}(0, 0)$$

so dass $(0, 0)$ kein lokales Extremum von f ist. Dagegen ist

$$\left(f_{xy}(2, 1)\right)^2 = (-12)^2 = 144 < 576 = 12 \cdot 48 = f_{xx}(2, 1) f_{yy}(2, 1)$$

und $f_{xx}(2, 1) = 12 > 0$, so dass in $(2, 1)$ ein lokales Minimum vorliegt.

- Für die auf \mathbb{R}^3 durch

$$f(x, y, z) = x^2 + y^2 + z^2 + xy - xz + 2$$

definierte Funktion ist an der Stelle $p = (0, 0, 0)$

$$(\text{grad} f)(0, 0, 0) = (2x + y - z, 2y + x, 2z - x) |_{(0,0,0)} = (0, 0, 0).$$

Also ist $(0, 0, 0)$ ein stationärer Punkt. Weiter ist

$$(\text{Hess } f)(x, y, z) = \begin{pmatrix} 2 & 1 & -1 \\ 1 & 2 & 0 \\ -1 & 0 & 2 \end{pmatrix}$$

für alle $(x, y, z) \in \mathbb{R}$, und diese Matrix ist positiv definit, wie man mit dem Hauptminorenkriterium leicht überprüft:

$$\det(2) = 2 > 0, \quad \det \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} = 3 > 0, \quad \det \begin{pmatrix} 2 & 1 & -1 \\ 1 & 2 & 0 \\ -1 & 0 & 2 \end{pmatrix} = 4 > 0.$$

Also liegt in $(0, 0, 0)$ ein lokales Minimum von f vor. ■

Um globale Extrema von $f : D \rightarrow \mathbb{R}$ zu bestimmen, sucht man zunächst die lokalen Extrema im Inneren von D und untersucht dann noch das Verhalten von f in der Nähe des Randes von D .

11.5 Zerlegung

Korollar 11.5 *Zu jeder quadratischen Form Q auf V gibt es eine Basis $\vec{v}_1, \dots, \vec{v}_n$ und zu jeder Norm $\|\cdot\|$ Skalare $\lambda_+ > 0 > \lambda_-$ in \mathbb{R} so, dass*

$$\begin{aligned} Q(\vec{x}) &\geq \lambda_+ \|\vec{x}\|^2 && \text{falls } \vec{x} \in V_+ := \{\vec{x} \mid \sum_{i=1}^p x_i \vec{v}_i\} \\ Q(\vec{x}) &\leq \lambda_- \|\vec{x}\|^2 && \text{falls } \vec{x} \in V_- := \{\vec{x} \mid \sum_{i=p+1}^{p+q} x_i \vec{v}_i\} \\ Q(\vec{x}) &= 0 && \text{falls } \vec{x} \in V_0 := \{\vec{x} \mid \sum_{i=q+1}^n x_i \vec{v}_i\} \end{aligned}$$

Beweis wie Satz 11.2.

$V = V_+ \oplus V_- \oplus V_0$ ist eine direkte Zerlegung von V , eine Zerlegung für Q . Nur das Radikal $V_0 = \{\vec{x} \mid Q(\vec{x}) = 0\}$ und die Dimensionen $\dim V_+$, $\dim V_-$ sind eindeutig bestimmt. Hat man auf V ein Skalarprodukt festgelegt, so gibt es dazu eine eindeutig bestimmte orthogonale Zerlegung für Q , die sich aus der Hauptachsenzerlegung ergibt - V_+ bzw. V_- werden von den Achsen mit positiven Eigenwerten (Momenten) ausgespannt.

Korollar 11.6 *Unter der Voraussetzung von Satz 11.3 gilt für obige Zerlegung*

- $V = V_+ \oplus V_-$
- $f|_{p+V_+}$ ist lokal positiv definit
- $f|_{p+V_-}$ ist lokal negativ definit

Beweis wie Satz 11.3

12 Gaußalgorithmus und inverse Matrix

12.1 Umformung

Sei $K = \mathbb{Q}, \mathbb{R}, \mathbb{C}$ oder ein anderer Körper. Sei ein lineares Gleichungssystem (S) gegeben durch m Gleichungen in n Variablen:

$$\begin{array}{ll} (1) & a_{11}x_1 + a_{12}x_2 + \dots + a_{1j}x_j + \dots + a_{1n}x_n = b_1 \\ (2) & a_{21}x_1 + a_{22}x_2 + \dots + a_{2j}x_j + \dots + a_{2n}x_n = b_2 \\ & \vdots \\ (i) & a_{i1}x_1 + a_{i2}x_2 + \dots + a_{ij}x_j + \dots + a_{in}x_n = b_i \\ & \vdots \\ (m) & a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mj}x_j + \dots + a_{mn}x_n = b_m \end{array}$$

wobei die a_{ij} jeweils feste Zahlen aus K sind. Eine Spalte

$$\mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

aus K^n ist eine *Lösung* von (S) , wenn sie alle Gleichungen (1) – (m) erfüllt. Die Aufgabe, “das Gleichungssystem (S) zu lösen”, bedeutet die Gesamtheit aller Lösungen, den *Lösungsraum* L von (S) , möglichst explizit anzugeben. Im Extremfall besteht diese in der Angabe eines einzigen Lösungsvektors oder in der Mitteilung “unlösbar”

Definition. Eine *elementare Umformung* des Gleichungssystems S in das neue Gleichungssystem (S') : (1') – (m') kann erfolgen durch:

(G1) Subtraktion eines Vielfachen einer Gleichung (k) von einer anderen, (l) :

$$(l') = (l) - r(k) \quad (a_{l1} - ra_{k1})x_1 + \dots + (a_{ln} - ra_{kn})x_n = b_l - rb_k,$$

(G2) Vertauschen der beiden Gleichungen (k) und (l)

(G3) Multiplikation einer Gleichung (k) mit $r \neq 0$ aus K

$$(k') = ra_{k1}x_1 + \dots + ra_{kn}x_n = rb_k, \quad .$$

(G4) Weglassen trivialer Gleichungen $0x_1 + \dots + 0x_n = 0$.

Satz 12.1 *Geht (S') aus (S) durch elementare Umformung hervor, so haben (S) und (S') denselben Lösungsraum.*

Beweis zu (G1). Ist \mathbf{x} Lösung von (k) und (l) , so $(a_{l1} - ra_{k1})x_1 + \dots + (a_{ln} - ra_{kn})x_n = a_{l1}x_1 + \dots + a_{ln}x_n - (ra_{k1}x_1 + \dots + ra_{kn}x_n) = b_l - rb_k$ also auch Lösung von (l') . Ist \mathbf{x} Lösung von (k) und (l') , so $a_{l1}x_1 + \dots + a_{ln}x_n = (a_{l1} - ra_{k1})x_1 + \dots + (a_{ln} - ra_{kn})x_n + r(a_{k1} + \dots + a_{kn}) = b_l - rb_k + rb_k = b_l$.

Die mehr oder weniger systematische und numerisch günstige Verwendung dieser Schritte, mit dem Ziel (S) in ein stufenförmiges System zu überführen, wird *Gaußscher Algorithmus* genannt. Man darf dabei auch Spalten vertauschen, d.h. die Variablen treten dann in anderer Reihenfolge auf. Bei den Lösungsspalten hat man natürlich wieder die ursprüngliche Reihenfolge zu benutzen.

12.2 Stufenform

Für einen Term $0x_i$ im Gleichungssystem dürfen wir auch 0 oder gar nichts schreiben. Ein Gleichungssystem (S) ist *homogen*, falls $b_1 = \dots = b_m = 0$. Das System (S_h) , in dem jedes b_i durch 0 ersetzt wird, heisst das *homogene* System zu (S) . Ein Gleichungssystem (S) in *Stufenform* ist von der Gestalt

$$\begin{array}{cccccccc}
 a_{1j_1}x_{j_1} + & \dots & a_{1j_2}x_{j_2} + & \dots & a_{1j_i}x_{j_i} + & \dots & a_{1j_r}x_{j_r} + & \dots & +a_{1n}x_n = & b_1 \\
 & & a_{2j_2}x_{j_2} + & \dots & a_{2j_i}x_{j_i} + & \dots & a_{2j_r}x_{j_r} + & \dots & +a_{2n}x_n = & b_2 \\
 & & & \ddots & & \vdots & \vdots & \vdots & & \vdots \\
 & & & & a_{ij_i}x_{j_i} + & \dots & a_{ij_r}x_{j_r} + & \dots & +a_{in}x_n = & b_i \\
 & & & & & \ddots & \vdots & \vdots & & \vdots \\
 & & & & & & a_{rj_r}x_{j_r} + & \dots & +a_{rn}x_n = & b_r \\
 & & & & & & & & & 0 = b_{r+1} \\
 & & & & & & & & & \vdots \\
 & & & & & & & & & 0 = b_m
 \end{array}$$

mit Zahlen $a_{ij_i} \neq 0$ rechts neben den Stufenkanten, den *Pivots*. Die entsprechenden Variablen x_{j_i} heissen *Pivotvariable*. r heisst der *Rang* des Systems - die Eindeutigkeit werden wir später beweisen. Für $r = m$ hat man keine Gleichungen $0 = b_l$, für $r = 0$ hat man nur solche.

Scholion. Sei (S) ein stufenförmiges System von m linearen Gleichungen in n Variablen x_1, \dots, x_n mit Koeffizienten in einem Körper K .

a) (S) ist unlösbar genau dann, es mindestens eine Gleichung von der Form $0 = b_l$, $r < l \leq m$, mit einer Zahl $b_l \neq 0$ enthält. Ein homogenes System ist stets lösbar, es hat mindestens die *triviale Lösung* $\mathbf{0}$.

b) Ist (S) lösbar, so hat (S) eine eindeutig bestimmte Lösung \mathbf{x} genau dann, wenn $r = n$, d.h. wenn die Pivots gerade die 'Diagonalkoeffizienten' a_{ii} sind:

$$\begin{array}{cccccccc}
 a_{11}x_1 + & a_{12}x_2 + & \dots & a_{1i}x_i + & \dots & a_{1n}x_n = & b_1 \\
 & a_{22}x_2 + & \dots & a_{2i}x_i + & \dots & a_{2n}x_n = & b_2 \\
 & & \ddots & \vdots & \vdots & & \vdots \\
 & & & a_{ii}x_i + & \dots & a_{in}x_n = & b_i \\
 & & & & \ddots & & \vdots \\
 & & & & & a_{nn}x_n = & b_n \\
 & & & & & & 0 = 0 \\
 & & & & & & \vdots \\
 & & & & & & 0 = 0
 \end{array}$$

Man erhält die Lösung durch Rücksubstitution:

$$x_n = \frac{1}{a_{nn}}b_{nn}, \quad x_{n-1} = \frac{1}{a_{n-1,n-1}}(b_{n-1} - a_{n-1,n}x_n), \dots$$

Bei einem homogenen System ist diese eindeutige Lösung dann $\mathbf{x} = \mathbf{0}$, die triviale Lösung.

12.3 Gauss'scher Algorithmus

Satz 12.2 Zu jedem System von m linearen Gleichungen (in Variablen x_1, \dots, x_n) mit Koeffizienten in einem Körper K gibt es ein dazu gleichwertiges (d.h. mit demselben Lösungsraum) in Stufenform. Man kann ein solches aus dem Ausgangssystem durch wiederholte Anwendung der elementaren Umformungen (G1) und (G2) erhalten, insbesondere auch durch den folgenden Algorithmus. Dabei werden auch die zugehörigen homogenen Systeme ineinander überführt.

Umformungsschritt U_{mn} für ein System (S) von m Gleichungen in n Variablen.

a) Suche das grösste k mit $a_{ij} = 0$ für alle $i = 1, \dots, m$, $j = 1, \dots, k - 1$ (d.h. mit nur Nullen in den den ersten $k - 1$ Spalten und $k = 1$ falls die erste Spalte nicht nur aus Nullen besteht).

b) Falls $a_{1k} = 0$, aber $a_{ik} \neq 0$ für ein i , nimm das erste solche und vertausche die Gleichungen (1) und (i). Nach Umnummerierung fahre mit c) fort.

c) Falls $a_{1k} \neq 0$, lasse die erste Gleichung unverändert und subtrahiere

das $\frac{a_{2k}}{a_{1k}}$ -fache der ersten Gleichung von der zweiten, usw. ...

das $\frac{a_{mk}}{a_{1k}}$ -fache der ersten Gleichung von der m -ten;

gehe also zum folgendem Gleichungssystem über:

$$(1), (2) - \frac{a_{2k}}{a_{1k}}(1), \dots, (m) - \frac{a_{mk}}{a_{1k}}(1).$$

Gauss'scher Algorithmus

1.Schritt. Wende den Umformungsschritt U_{mn} auf das Ausgangssystem an.

2.Schritt. Behalte die erste Gleichung des neuen Systems bei und fasse die restlichen als ein System von $m - 1$ Gleichungen in den $n - k + 1$ Variablen x_k, \dots, x_n auf. Wende auf dieses den Umformungsschritt $U_{m-1, n-k}$ an. Das Ergebnis dieses Schrittes bildet zusammen mit der beibehaltenen ersten Gleichung das neue System nach Schritt 2.

3.Schritt. Behalte aus dem System nach Schritt 2 die ersten beiden Gleichungen bei, fasse die restlichen $m - 2$ als System in $n - k - k_2 + 1$ Variablen auf und wende $U_{m-2, n-k-k_2}$ an.

....und so weiter bis zu höchstens n Schritten.

Bemerkungen. Tritt im Umformungsschritt Fall a) mit $k > 1$ ein, so wird ein waagrechter Teil einer Stufe erzeugt, tritt Fall c) ein, wird eine Stufenkante und ein Pivot erzeugt.

Triviale Gleichungen $0 = 0$ werden ganz nach unten getauscht und zum Schluss weggelassen, man kann sie also auch gleich weglassen.

Tritt eine widersprüchliche Gleichung $0 = b$ mit einer Zahl $b \neq 0$ auf, so kann man abbrechen und feststellen, dass das System unlösbar ist.

Die Beschreibung des Algorithmus ist ein Beweis des Satzes - und von grundsätzlicher Bedeutung in der Linearen Algebra. Offen bleibt vorerst die

Frage: Ist die Anzahl der freien Parameter durch die Koeffizientenmatrix eindeutig bestimmt oder hängt sie davon ab, wie man rechnet?

12.4 Matrixschreibweise für Gleichungssysteme

Matrizen ermöglichen eine komfortable Notation für lineare Gleichungssysteme: Mit $m \times n$ -Matrix A über K , Spalte \mathbf{b} in K^n , und Variablen­spalte \mathbf{x} mit n Komponenten kann man ein Gleichungssystem (S) und das zugehörige homogene System (S_h) so schreiben (A heisst dann die Koeffizientenmatrix des Systems)

$$(S) \mathbf{Ax} = \mathbf{b}, \quad (S_h) \mathbf{Ax} = \mathbf{0}$$

Satz 12.3 *Ist \mathbf{x}_h die allgemeine Lösung bzw. X_h der Lösungsraum des homogenen Systems (S_h) und ist \mathbf{x}_s eine (spezielle) Lösung von (S), so erhält man die allgemeine Lösung bzw. den Lösungsraum von (S) mit*

$$\mathbf{x} = \mathbf{x}_s + \mathbf{x}_h \text{ bzw. } X = \mathbf{x}_s + X_h = \{\mathbf{x}_s + \mathbf{x}_h \mid \mathbf{x}_h \text{ in } X_h\}.$$

Beweis. Sei \mathbf{x}_s eine spezielle Lösung von (S), d.h. $\mathbf{Ax}_s = \mathbf{b}$. Ist nun \mathbf{x}_h eine Lösung von (S_h), d.h. $\mathbf{Ax}_h = \mathbf{0}$, so folgt $A(\mathbf{x}_s + \mathbf{x}_h) = \mathbf{Ax}_s + \mathbf{Ax}_h = \mathbf{b} + \mathbf{0} = \mathbf{b}$, also ist $\mathbf{x}_s + \mathbf{x}_h$ eine Lösung von (S). Hat man umgekehrt eine Lösung \mathbf{x} von (S), so gilt $A(\mathbf{x} - \mathbf{x}_s) = \mathbf{Ax} - \mathbf{Ax}_s = \mathbf{b} - \mathbf{b} = \mathbf{0}$ und somit $\mathbf{x} = \mathbf{x}_s + (\mathbf{x} - \mathbf{x}_s)$ mit Lösung $\mathbf{x} - \mathbf{x}_s$ von (S_h), d.h. \mathbf{x} ist von der behaupteten Gestalt.

12.5 Matrixinversion und Gleichungslösen

Satz 12.4 *Für eine $n \times n$ -Matrix A sind die folgenden Aussagen äquivalent*

- (1) *A is invertierbar*
- (2) *Es gibt eine Matrix X mit $AX = E$*
- (3) *Es gibt eine Matrix Y mit $YA = E$*
- (4) *Jedes Gleichungssystem $\mathbf{Ax} = \mathbf{b}$ ist lösbar*
- (5) *Das Gleichungssystem $\mathbf{Ax} = \mathbf{0}$ hat nur die Lösung $\mathbf{0}$*
- (6) *Bei jeder Umformung von A in eine Stufenform A' ergibt sich eine strikte Dreiecksmatrix.*
- (7) *A lässt sich in eine strikte (obere) Dreiecksmatrix umformen,*

Dabei heisst A' strikte (obere) Dreiecksmatrix, wenn $a'_{ij} = 0$ für $i > j$ und $a'_{jj} \neq 0$.

Korollar 12.5 *Aus $AX = E$ folgt $X = A^{-1}$. Aus $YA = E$ folgt $Y = A^{-1}$. Ist A invertierbar, so hat $\mathbf{Ax} = \mathbf{b}$ die eindeutig bestimmte Lösung $\mathbf{x} = A^{-1}\mathbf{b}$. Eine invertierbare Matrix enthält keine Nullspalte.*

Korollar 12.6 *Gauß-Jordan, Ist A invertierbar, so kann man die Spalten \mathbf{x}_j von $X = A^{-1}$ als Lösungen der Gleichungssysteme $\mathbf{Ax}_j = \mathbf{0}$ bestimmen. Dazu überführt man A durch Zeilenumformungen in die Einheitsmatrix. Führt man beginnend mit der Matrix E statt A dieselben Umformungen aus, so erhält man A^{-1}*

$A = \begin{pmatrix} 1 & 1 & 1 \\ 2 & 2 & 4 \\ 2 & 1 & 1 \end{pmatrix}.$	$\begin{pmatrix} 1x_1 + 1x_2 + 1x_3 = 1u_1 + 0u_2 + 0u_3 \\ 2x_1 + 2x_2 + 4x_3 = 0u_1 + 1u_2 + 0u_3 \\ 2x_1 + 1x_2 + 1x_3 = 0u_1 + 0u_2 + 1u_3 \end{pmatrix}$
	$\Leftrightarrow \begin{pmatrix} 1x_1 + 1x_2 + 1x_3 = 1u_1 + 0u_2 + 0u_3 \\ 0x_1 + 0x_2 + 2x_3 = -2u_1 + 1u_2 + 0u_3 \\ 0x_1 - 1x_2 - 1x_3 = -2u_1 + 0u_2 + 1u_3 \end{pmatrix}$
	$\Leftrightarrow \begin{pmatrix} 1x_1 + 1x_2 + 1x_3 = 1u_1 + 0u_2 + 0u_3 \\ 0x_1 - 1x_2 - 1x_3 = -2u_1 + 0u_2 + 1u_3 \\ 0x_1 + 0x_2 + 2x_3 = -2u_1 + 1u_2 + 0u_3 \end{pmatrix}$
	$\Leftrightarrow \begin{pmatrix} 1x_1 + 1x_2 + 1x_3 = 1u_1 + 0u_2 + 0u_3 \\ 0x_1 - 1x_2 - 1x_3 = -2u_1 + 0u_2 + 1u_3 \\ 0x_1 + 0x_2 + 1x_3 = -1u_1 + \frac{1}{2}u_2 + 0u_3 \end{pmatrix}$
	$\Leftrightarrow \begin{pmatrix} 1x_1 + 1x_2 + 0x_3 = 2u_1 - \frac{1}{2}u_2 + 0u_3 \\ 0x_1 - 1x_2 + 0x_3 = -3u_1 + \frac{1}{2}u_2 + 1u_3 \\ 0x_1 + 0x_2 + 1x_3 = -1u_1 + \frac{1}{2}u_2 + 0u_3 \end{pmatrix}$
	$\Leftrightarrow \begin{pmatrix} 1x_1 + 0x_2 + 0x_3 = -1u_1 + 0u_2 + 1u_3 \\ 0x_1 - 1x_2 + 0x_3 = -3u_1 + \frac{1}{2}u_2 + 1u_3 \\ 0x_1 + 0x_2 + 1x_3 = -1u_1 + \frac{1}{2}u_2 + 0u_3 \end{pmatrix}$
$A^{-1} = \begin{pmatrix} -1 & 0 & 1 \\ 3 & -\frac{1}{2} & -1 \\ -1 & \frac{1}{2} & 0 \end{pmatrix}$	$\Leftrightarrow \begin{pmatrix} 1x_1 + 0x_2 + 0x_3 = -1u_1 + 0u_2 + 1u_3 \\ 0x_1 + 1x_2 + 0x_3 = 3u_1 - \frac{1}{2}u_2 - 1u_3 \\ 0x_1 + 0x_2 + 1x_3 = -1u_1 + \frac{1}{2}u_2 + 0u_3 \end{pmatrix}$

$$\begin{aligned}
 & \left(\begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 \\ 2 & 2 & 4 & 0 & 1 & 0 \\ 2 & 1 & 1 & 0 & 0 & 1 \end{array} \right) \rightsquigarrow \left(\begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 2 & -2 & 1 & 0 \\ 0 & -1 & -1 & -2 & 0 & 1 \end{array} \right) \rightsquigarrow \left(\begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & -1 & -1 & -2 & 0 & 1 \\ 0 & 0 & 2 & -2 & 1 & 0 \end{array} \right) \\
 & \rightsquigarrow \left(\begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & -1 & -1 & -2 & 0 & 1 \\ 0 & 0 & 1 & -1 & \frac{1}{2} & 0 \end{array} \right) \rightsquigarrow \left(\begin{array}{ccc|ccc} 1 & 1 & 0 & 2 & -\frac{1}{2} & 0 \\ 0 & -1 & 0 & -3 & \frac{1}{2} & 1 \\ 0 & 0 & 1 & -1 & \frac{1}{2} & 0 \end{array} \right) \\
 & \rightsquigarrow \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & -1 & 0 & 1 \\ 0 & -1 & 0 & -3 & \frac{1}{2} & 1 \\ 0 & 0 & 1 & -1 & \frac{1}{2} & 0 \end{array} \right) \rightsquigarrow \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & -1 & 0 & 1 \\ 0 & 1 & 0 & 3 & -\frac{1}{2} & -1 \\ 0 & 0 & 1 & -1 & \frac{1}{2} & 0 \end{array} \right)
 \end{aligned}$$

Beweis. Aus (1) folgen (2) und (3) trivialerweise. Aus (2) folgt (4) mit $\mathbf{x} = X\mathbf{b}$. Aus (3) folgt (5): $\mathbf{x} = Y A \mathbf{x} = Y \mathbf{0} = \mathbf{0}$. Aus (4) bzw. (5) folgt (6), wie bei Diskussion der Stufenform festgestellt. Aus (6) folgt (7): Umformung nach Gauß-Algorithmus. Aus (7) folgt (1): Wir bestimmen eine Matrix B mit $A\mathbf{x} = \mathbf{u} \Leftrightarrow B\mathbf{u} = \mathbf{x}$. Dazu betrachten wir das Gleichungssystem $A\mathbf{x} = E\mathbf{u}$, in Matrizenform geschrieben als

$$(A \mid E) \begin{pmatrix} \mathbf{x} \\ -\mathbf{u} \end{pmatrix} = 0$$

Die Umformung von A in die strikte Dreiecksmatrix A' nach (7) ergibt das äquivalente Gleichungssystem

$$(A' \mid C) \begin{pmatrix} \mathbf{x} \\ -\mathbf{u} \end{pmatrix} = 0$$

und indem wir den Gauß-Algorithmus nun von unter nach oben rechnen (wobei wir im rechten Moment die Diagonaleinträge durch Multiplikation in 1 umwandeln), können wir A' in E überführen und das Gleichungssystem in das äquivalente

$$(E \mid B) \begin{pmatrix} \mathbf{x} \\ -\mathbf{u} \end{pmatrix} = 0$$

Somit ist $A\mathbf{x} = \mathbf{u}$ zu $B\mathbf{u} = \mathbf{x}$ äquivalent und B die Inverse von A . Enthält A eine Nullspalte, so können wir eine Stufenform gewinnen, die keine Dreiecksmatrix ist. \square

12.6 Isomorphismen

Korollar 12.7 *Seien V und W K -Vektorräume der endlichen Dimension $n = \dim V = \dim W$ mit Basen α bzw. β und $\phi : V \rightarrow W$ linear mit zugehöriger Matrix A . Dann sind die folgenden Aussagen äquivalent*

- (1) $\phi : V \rightarrow W$ ist surjektiv
- (2) ϕ ist injektiv
- (3) $\phi : V \rightarrow W$ ist ein Isomorphismus
- (4) A ist invertierbar

Die Matrix der Umkehrabbildung ϕ^{-1} bzgl. β und α ist dann A^{-1} .

Beweis. Surjektivität von ϕ ist (4) aus Satz 12.4. (5) ist äquivalent zu $A\mathbf{x} = A\mathbf{u} \Rightarrow \mathbf{x} = \mathbf{u}$ und das ist die Injektivität von ϕ . Also folgt die Äquivalenz von (1)-(4) aus Satz 12.4. Ist B die Matrix von ϕ^{-1} , so $BA = E$ nach Abschn. 3.10, also $B = A^{-1}$. \square

12.7 Spaltenumformungen

Beim Transponieren geht eine Rechnung des Gauß-Algorithmus für Zeilen einer Matrix in eine Rechnung für Spalten über. Die (oberen) Stufen- bzw. Dreiecksmatrizen gehen durch Transponieren in *untere* über. Es folgt:

Korollar. *Jede Matrix A lässt sich durch Spaltenumformungen in eine untere Stufenform überführen. A ist invertierbar genau dann, wenn diese eine (strikte) untere Dreiecksmatrix ist.*

13 Skalar-, Vektor- und Spatprodukt im Raum

13.1 Richtungskomponenten

Wir setzen voraus, dass wir über Längengleichheit $\|PQ\| = \|RS\|$ von Pfeilen reden können und die üblichen Aussagen der Kongruenzgeometrie erfüllt sind. Insbesondere sind also gegenüberliegende Seiten eines Parallelogramms gleich lang und wir können auch darüber reden, dass zwei Vektoren gleich lang sind ($\|\vec{a}\| = \|\vec{b}\|$, und dann auch $\|r\vec{a}\| = \|r\vec{b}\|$). Das passende Gerät um Längengleichheit festzustellen, ist der Zirkel oder ein Band. Das erlaubt uns, rechte Winkel und Orthogonalität von Vektoren einzuführen

- $\angle QPR$ ist *recht* genau dann, wenn $P = Q$ oder wenn es $Q' \neq Q$ auf der Geraden durch PQ gibt mit $\|PQ\| = \|PQ'\|$, $\|RQ\| = \|RQ'\|$
- $\vec{a} \perp \vec{b}$ genau dann, wenn ein/alle $\angle QPR$ mit $\overrightarrow{PQ} = \vec{a}$, $\overrightarrow{PR} = \vec{b}$ recht sind
- $\vec{a} \perp \vec{b}$ genau dann, wenn $\vec{b} \perp \vec{a}$
- Zu jeder Geraden g mit Richtungsvektor \vec{a} und Punkt Q gibt es genau einen Punkt R auf g (den *Fusspunkt des Lotes* von Q auf g) so, dass $\overrightarrow{RQ} \perp \vec{a}$

Damit können wir nun den auch physikalisch grundlegenden Begriff der *Komponente* \vec{c} eines Vektor \vec{b} in der durch einen Vektor \vec{a} gegebenen Richtung definieren durch die Bedingungen

$$\vec{c} = r\vec{a} \text{ für ein } r \in \mathbb{R}, \quad (\vec{b} - \vec{c}) \perp \vec{a}$$

In der Tat, ist $\vec{a} \neq \vec{0}$, so wähle man einen Punkt O und g als die Gerade durch O mit Richtungsvektor \vec{a} und $Q = \vec{b} + O$. Dann ergibt sich \vec{c} eindeutig als $\vec{c} = \overrightarrow{OR}$ mit dem Fusspunkt R des Lotes von Q auf g . Ist $\vec{a} = \vec{0}$, so auch $\vec{c} = \vec{0}$.

13.2 Skalarprodukt

Sei nun zusätzlich eine Längeneinheit festgelegt, z.B. dadurch dass wir einem bestimmten Pfeil OE die Länge 1 zuschreiben. Um die Länge eines beliebigen Vektors $\vec{v} = \overrightarrow{PQ} \neq \vec{0}$ zu bestimmen, konstruieren wir auf der Geraden g durch $O' = P$ und Q einen Punkt E' mit $\|O'E'\| = \|OE\|$ und setzen

$$\|\vec{v}\| = |r| \quad \text{mit } \vec{v} = r\overrightarrow{O'E'} \quad \text{wobei } |r| = \begin{cases} r & \text{falls } r \geq 0 \\ -r & \text{falls } r < 0 \end{cases}$$

Der Nullvektor $\vec{0}$ ist der einzige Vektor der Länge 0. Das *skalare* oder *innere Produkt* $\langle \vec{a} | \vec{b} \rangle$ definieren wir zunächst für einen Spezialfall durch folgende Beziehung

$$\langle \vec{e} | \vec{b} \rangle \vec{e} \text{ ist die Komponente von } \vec{b} \text{ in Richtung } \vec{e} \text{ falls } \|\vec{e}\| = 1$$

$$\vec{f} \qquad \vec{b}$$

$\cos \phi$

$$\vec{c} = \langle \vec{e} | \vec{b} \rangle \vec{e} = (|\vec{b}| \cos \phi) \vec{e}$$

\vec{e}

Schreiben wir

$$\cos \phi := \langle \vec{e} | \vec{f} \rangle \quad \text{falls } \|\vec{f}\| = 1, \quad \vec{b} = \|\vec{b}\| \vec{f}$$

so haben wir

$$\langle \vec{e} | \vec{b} \rangle = \|\vec{b}\| \cos \phi = \|\vec{e}\| \cdot \|\vec{b}\| \cdot \cos \phi$$

Die Bedeutung von ϕ und \cos muss man mit Hilfe der Anschauung aus der Skizze entnehmen. Erwünscht sind für ein allgemein definiertes Skalarprodukt die Regeln

$$\begin{aligned} (E1) \quad & \langle \vec{a} | \vec{b} \rangle = \langle \vec{b} | \vec{a} \rangle \\ (E2) \quad & \langle \vec{a} | s\vec{b} \rangle = s \langle \vec{a} | \vec{b} \rangle \\ (E3) \quad & \langle \vec{a} | \vec{b} + \vec{c} \rangle = \langle \vec{a} | \vec{b} \rangle + \langle \vec{a} | \vec{c} \rangle \\ (E4) \quad & \langle \vec{a} | \vec{a} \rangle > 0 \text{ für } \vec{a} \neq 0 \end{aligned}$$

Diese verifizieren wir leicht falls $\vec{a} = 1$ - und in (E1) auch $\|\vec{b}\| = 1$ - vgl. die Skizzen. Da wir mit $\vec{e} = \frac{1}{\|\vec{a}\|} \vec{a}$ einen Vektor der Länge 1 erhalten, sind wir wegen (E1-2) gezwungen, das Skalarprodukt so zu definieren

$$\langle \vec{a} | \vec{b} \rangle = \|\vec{a}\| \left\langle \frac{1}{\|\vec{a}\|} \vec{a} | \vec{b} \right\rangle$$

und rechnen sofort nach, dass (E1-4) auch dafür gelten. Mit obiger Schreibweise ergibt sich die beliebte Formel

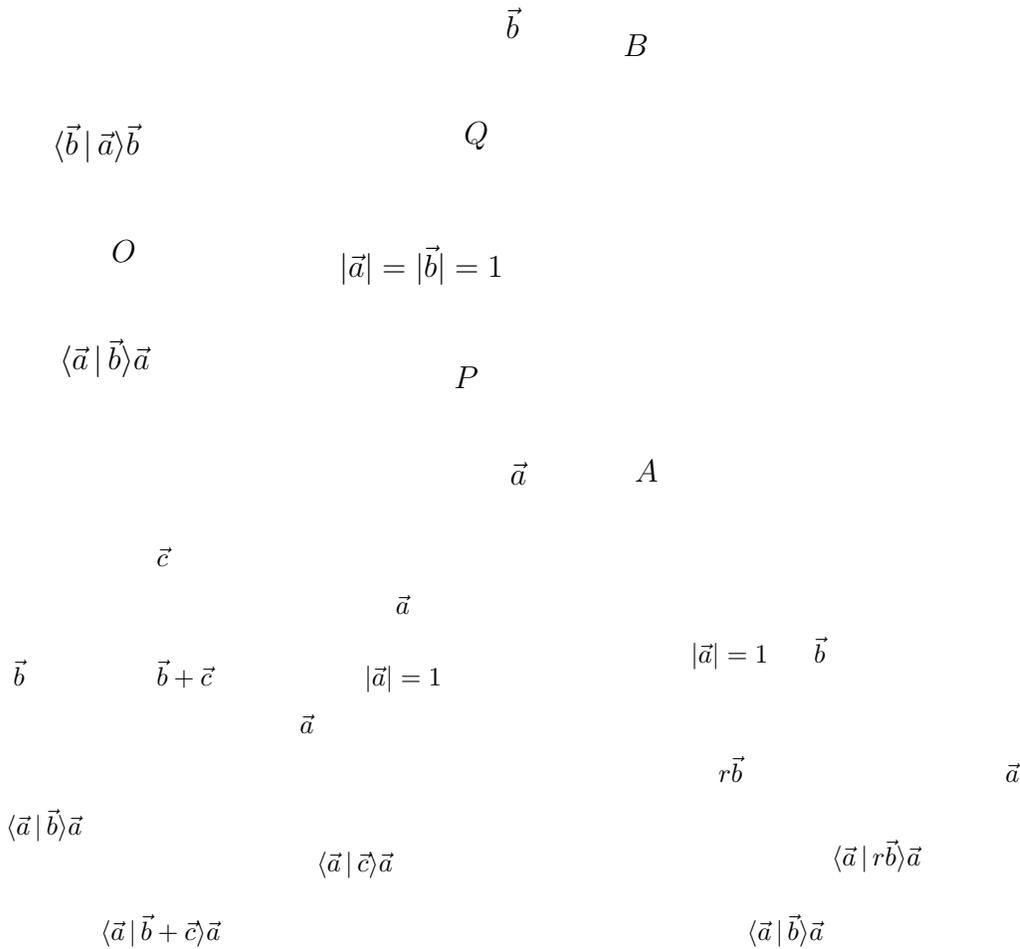
$$\langle \vec{a} | \vec{b} \rangle = \|\vec{a}\| \cdot \|\vec{b}\| \cdot \cos \phi$$

$\vec{a} \cdot \vec{b}$ ist eine andere Schreibweise für das Skalarprodukt. Es folgt

$$\frac{\langle \vec{a} | \vec{b} \rangle}{\|\vec{a}\|^2} \vec{a} \text{ ist die Komponente von } \vec{b} \text{ in Richtung } \vec{a}$$

$$\langle \vec{a} | \vec{a} \rangle = \|\vec{a}\|^2, \quad \|\vec{a}\| = \sqrt{\langle \vec{a} | \vec{a} \rangle}, \quad \langle \vec{a} | \vec{b} \rangle = 0 \text{ genau dann, wenn } \vec{a} \perp \vec{b}$$

Zum Nachweis von (E1), (E3) und (E2) siehe folgende Skizzen



13.3 Ungleichungen

Als Einübung auf späteres rein axiomatisches Vorgehen, wollen wir die folgenden geometrischen Aussagen ausschliesslich mit (V1-8) und (E1-4) herleiten, Insbesondere ist $\|\vec{a}\|$ und \perp wie oben aus dem Skalarprodukt definiert.

Satz 13.1 • $\|\vec{a} - \vec{b}\|^2 = \|\vec{a}\|^2 + \|\vec{b}\|^2 \Leftrightarrow \vec{a} \perp \vec{b}$ Pythagoras

• $(\vec{b} - \vec{c}) \perp \vec{a}$ und $\|\vec{b} - \vec{c}\| = \min\{\|\vec{b} - \lambda\vec{a}\| \mid \lambda \in \mathbb{R}\}$ für $\vec{c} = \frac{\langle \vec{a} | \vec{b} \rangle}{\|\vec{a}\|^2} \vec{a}$ Lot
 $\|\vec{b} - \lambda\vec{a}\| = \|\vec{b} - \vec{c}\|$ genau dann, wenn $\lambda\vec{a} = \vec{c}$

• $|\langle \vec{a} | \vec{b} \rangle| \leq \|\vec{a}\| \cdot \|\vec{b}\|$ Cauchy-Schwarz
 und Gleichheit genau dann, wenn $\vec{a} \parallel \vec{b}$

• $\|\vec{a} + \vec{b}\| \leq \|\vec{a}\| + \|\vec{b}\|$ Dreiecksungleichung
 und Gleichheit genau dann, wenn $\vec{a} = \vec{0}$ oder $\vec{b} = r\vec{a}$ mit $r \geq 0$

$$\begin{array}{ccc}
 & & \vec{a}_2 \\
 & & \vec{a}'_2 \\
 \vec{e}_2 & & \vec{a}_1 \\
 & & \langle \vec{e}_1 | \vec{a}_2 \rangle \vec{e}_1 \\
 \vec{e}_1 & &
 \end{array}$$

Beweis. Pythagoras: $\|\vec{a} - \vec{b}\|^2 = \langle \vec{a} + (-1)\vec{b} | \vec{a} + (-1)\vec{b} \rangle =_{(E3)} \langle \vec{a} | \vec{a} \rangle + \langle \vec{a} | (-1)\vec{b} \rangle + \langle (-1)\vec{b} | \vec{a} \rangle + \langle \vec{b} | \vec{b} \rangle =_{(E1,2,4)} \|\vec{a}\|^2 - 2\langle \vec{a} | \vec{b} \rangle + \|\vec{b}\|^2$. Somit gilt $\|\vec{a} - \vec{b}\|^2 = \|\vec{a}\|^2 + \|\vec{b}\|^2$ genau dann, wenn $2\langle \vec{a} | \vec{b} \rangle = 0$, d.h. wenn $\langle \vec{a} | \vec{b} \rangle = 0$.

Lot: $\langle \vec{a} | \vec{b} - \vec{c} \rangle = \langle \vec{a} | \vec{b} \rangle - \langle \vec{a} | \vec{c} \rangle = 0$ und nun mit Pythagoras $\|\vec{b} - \vec{c}\|^2 \leq \|\vec{b} - \vec{c}\|^2 + \|\vec{c} - \lambda\vec{a}\|^2 = \|\vec{b} - \vec{c} + \vec{c} - \lambda\vec{a}\|^2 = \|\vec{b} - \lambda\vec{a}\|^2$. Und Gleichheit gilt genau dann, wenn $\|\vec{c} - \lambda\vec{a}\|^2 = 0$.

Cauchy-Schwarz: Ist $\vec{a} = \vec{0}$ oder $\vec{b} = \vec{0}$, so ist die Behauptung trivial. Andernfalls können wir nach der Bilinearität $\|\vec{b}\| = 1$ annehmen. Setze $\rho = \langle \vec{a} | \vec{b} \rangle$. Dann $0 \leq \|\vec{a} - \rho\vec{b}\|^2 = \langle \vec{a} - \rho\vec{b} | \vec{a} - \rho\vec{b} \rangle = \|\vec{a}\|^2 - \rho\langle \vec{a} | \vec{b} \rangle - \rho\langle \vec{b} | \vec{a} \rangle + \rho^2\|\vec{b}\|^2 = \|\vec{a}\|^2 - \rho^2$, also $|\rho| \leq \|\vec{a}\|$. Und es gilt $|\rho| = \|\vec{a}\|$ genau dann, wenn $\|\vec{a} - \rho\vec{b}\| = 0$.

Dreiecksungleichung: $\|\vec{a} + \vec{b}\|^2 = \langle \vec{a} | \vec{a} \rangle + \langle \vec{a} | \vec{b} \rangle + \langle \vec{b} | \vec{a} \rangle + \langle \vec{b} | \vec{b} \rangle \leq \|\vec{a}\|^2 + 2\|\vec{a}\|\|\vec{b}\| + \|\vec{b}\|^2 = (\|\vec{a}\| + \|\vec{b}\|)^2$. Da beide Seiten der Dreiecksungleichung ≥ 0 sind, folgt diese durch Wurzelziehen. Gilt die Gleichheit, so folgt $\langle \vec{a} | \vec{b} \rangle = \|\vec{a}\| \cdot \|\vec{b}\|$, also nach Cauchy-Schwarz $\vec{a} = \vec{0}$ oder $\vec{b} = r\vec{a}$. Dann aber $r\|\vec{b}\|^2 = \|\vec{a}\| \cdot \|\vec{b}\| \geq 0$ und somit $r \geq 0$. \square

13.4 Orthonormalbasen

Vektoren $\vec{e}_1, \vec{e}_2, \vec{e}_3$ des Raumes bilden eine *Orthonormalbasis*, wenn sie Länge 1 haben und paarweise aufeinander senkrecht stehen, d.h.

$$\langle \vec{e}_i | \vec{e}_j \rangle = \begin{cases} 1 & \text{falls } i = j \\ 0 & \text{falls } i \neq j \end{cases}$$

Dass sie insbesondere eine Basis α bilden, ist geometrisch klar. Analog für die Ebene. Bilden \vec{a}_1, \vec{a}_2 eine Basis einer Ebene, so erhält man eine ON-Basis mit

$$\vec{e}_1 = \frac{1}{\|\vec{a}_1\|}, \quad \vec{e}_2 = \frac{1}{\|\vec{a}'_2\|} \vec{a}'_2 \quad \text{wobei } \vec{a}'_2 = \vec{a}_2 - \langle \vec{e}_1 | \vec{a}_2 \rangle \vec{e}_1$$

Bilden $\vec{a}_1, \vec{a}_2, \vec{a}_3$ eine Basis des Raumes, so muss man nun \vec{a}_3 orthogonal auf die von \vec{e}_1, \vec{e}_2 aufgepante Ebene projizieren und dann den Lotvektor normieren, um \vec{e}_3 zu erhalten

$$\vec{a}'_3 = \vec{a}_3 - \langle \vec{e}_1 | \vec{a}_3 \rangle \vec{e}_1 - \langle \vec{e}_2 | \vec{a}_3 \rangle \vec{e}_2, \quad \vec{e}_3 = \frac{1}{\|\vec{a}'_3\|} \vec{a}'_3$$

Die Koordinaten von Vektoren, Skalarprodukt und Länge berechnen sich nun für $\vec{x} = x_1\vec{e}_1 + x_2\vec{e}_2 + x_3\vec{e}_3$ und $\vec{y} = y_1\vec{e}_1 + y_2\vec{e}_2 + y_3\vec{e}_3$ einfach so

$$\xi_1 \vec{e}_2 = \langle \vec{e}_2 | \vec{x} \rangle$$

$$\xi_1 \vec{e}_1 = \langle \vec{e}_1 | \vec{x} \rangle$$

$y_i = \langle \vec{e}_i \vec{y} \rangle$ $\langle \vec{x} \vec{y} \rangle = x_1 y_1 + x_2 y_2 + x_3 y_3$ $\ \vec{x}\ = \sqrt{x_1^2 + x_2^2 + x_3^2}$
--

Das zweite rechnet man einfach nach, das dritte folgt dann sofort, ebenso das erste (indem man \vec{e}_i für \vec{x} einsetzt - hier kann man auch geometrisch argumentieren).

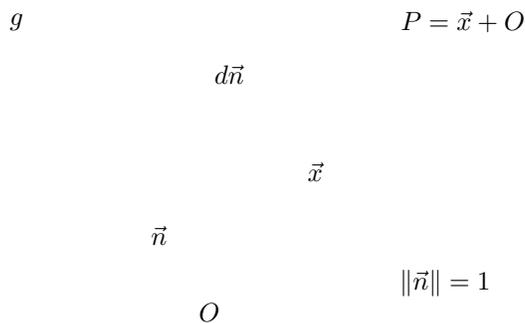
$$\langle x_1 \vec{e}_1 + x_2 \vec{e}_2 + x_3 \vec{e}_3 | y_1 \vec{e}_1 + y_2 \vec{e}_2 + y_3 \vec{e}_3 \rangle = \sum_{i,j} \langle x_i \vec{e}_i | y_j \vec{e}_j \rangle = \sum_{i,j} x_i y_j \langle \vec{e}_i | \vec{e}_j \rangle$$

In der Ebene geht es genauso mit nur 2 Koordinaten.

13.5 Normalenvektoren von Geraden bzw. Ebenen

Sei ein Punkt O der Ebene als Ursprung ausgezeichnet. Zu gegebenem Skalar d und *Normalen-Vektor* $\vec{n} \neq 0$ betrachten wir die Punktmenge

$$g = \{ \vec{x} + O \mid \langle \vec{x} | \vec{n} \rangle = d \}$$



Durch Multiplikation von \vec{n} mit $\pm \frac{1}{\|\vec{n}\|}$ können wir erreichen, dass gilt

$$\|\vec{n}\| = 1, \quad d \geq 0 \quad \text{Hessesche Normal(en)form}$$

Dann ist g die Gerade mit Parameterdarstellung

$$r\vec{v} + A \quad \text{wobei } \vec{0} \neq \vec{v} \perp \vec{n}, \quad \vec{a} = d\vec{n}, \quad A = \vec{a} + O$$

In der Tat, \vec{n}, \vec{v} ist eine Basis, d.h. mit der eindeutigen Darstellung $\vec{x} = x_1\vec{n} + x_2\vec{v}$ gilt $\langle \vec{x} | \vec{n} \rangle = x_1$ und somit $\langle \vec{x} | \vec{n} \rangle = d$ genau dann, wenn $\vec{x} = x_2\vec{v} + \vec{a}$.

Geht man von einer Parameterdarstellung $r\vec{v} + A$ mit $A = \vec{a} + O$ aus, so bestimme man $\vec{0} \neq \vec{v}^\perp \perp \vec{v}$ und dann $\vec{n} = \pm \frac{1}{\|\vec{v}^\perp\|} \vec{v}^\perp$ so, dass $d = \langle \vec{a} | \vec{n} \rangle \geq 0$.

d ist der Abstand der Geraden g von Punkt O . Hat man einen Punkt $P = \vec{p} + O$, so ist $\langle \vec{p} | \vec{n} \rangle \vec{n} + O$ der Fußpunkt des Lotes von P auf die Gerade durch O mit Richtung \vec{n} , also

$$d = \langle \vec{p} | \vec{n} \rangle \quad \text{Abstand des Punktes } P = \vec{p} + O \text{ von der Geraden } g$$

Ist zudem eine Orthonormalbasis \vec{e}_1, \vec{e}_2 der Ebene gegeben, so schreibt sich die Hessesche Normalenform in Koordinaten bzgl. dieser so

$$g = \{x_1\vec{e}_1 + x_2\vec{e}_2 + O \mid x_1n_1 + x_2n_2 = d\} \quad \text{wobei } \vec{n} = n_1\vec{e}_1 + n_2\vec{e}_2$$

Entsprechend geht es mit Ebenen im Raum, zwei zum Normalenvektor \vec{n} senkrechten, nicht parallelen Richtungsvektoren \vec{v}_1, \vec{v}_2 und 3 Koordinaten:

$$\begin{aligned} E &= \{r_1\vec{v}_1 + r_2\vec{v}_2 + A \mid r_1, r_2 \in \mathbb{R}\} = \{\vec{x} + O \mid \langle \vec{x} | \vec{n} \rangle = d\} \\ &= \{x_1\vec{e}_1 + x_2\vec{e}_2 + x_3\vec{e}_3 + O \mid x_1n_1 + x_2n_2 + x_3n_3 = d\} \end{aligned}$$

wobei $\vec{n} = n_1\vec{e}_1 + n_2\vec{e}_2 + n_3\vec{e}_3$ und $d = \langle \vec{a} | \vec{n} \rangle$ mit $A = \vec{a} + O$. Falls $d \geq 0$, so ist $|\langle \vec{p} | \vec{n} \rangle - d|$ der Abstand des Punktes $\vec{p} + O$ von E .

13.6 Orientierung

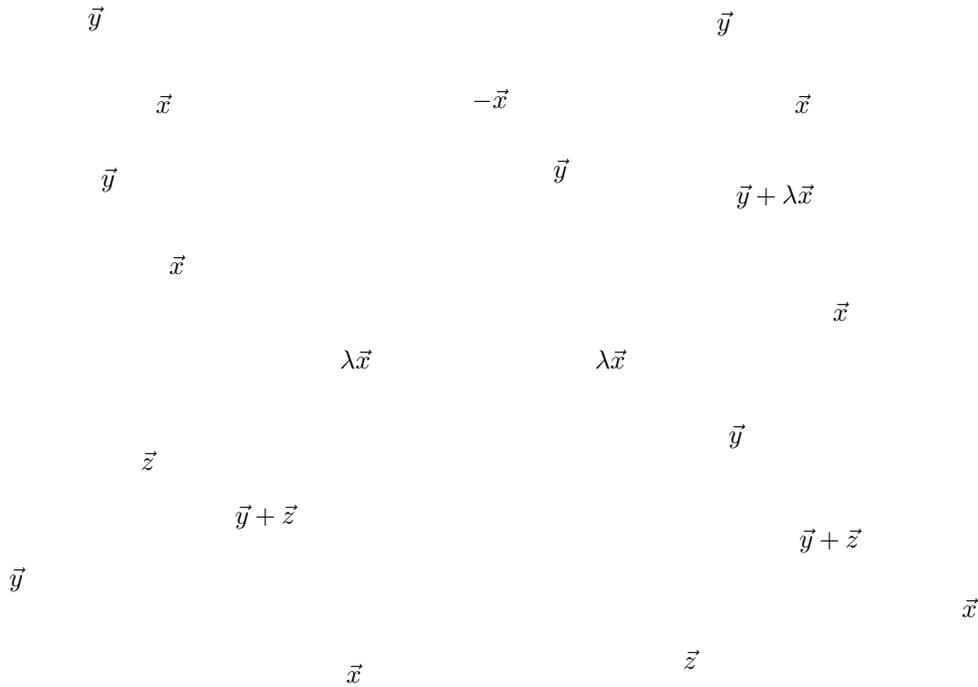
In einer Ebene oder im Raum wird eine Orientierung durch Auszeichnung einer Basis und der zulässigen Deformationen angegeben; dabei kommt beim Anschauungsraum die "Rechte Hand Regel" zur Anwendung. "Orientierte" Flächen bzw. Volumina ergeben sich dann als Determinanten, axiomatisch durch (D1-4) bestimmt. Eine Besonderheit des dreidimensionalen Raumes ist das Vektor- oder äussere Produkt (besonders ist, dass es als Operation im Raum verstanden werden kann). Als Koordinatensysteme benutzen wir hier positiv orientierte Orthonormalbasen. Bezüglich eines solchen lassen sich dann Determinanten und Vektorprodukt koordinatenweise berechnen.

13.7 Flächen

Setzt man für Quadrate mit Seitenlänge 1 den Flächeninhalt 1 fest, so ergibt sich (nach Archimedes) der Flächeninhalt eines Parallelogramms als gh , wobei g die Länge einer (Grund)Seite und h die Länge einer dazu senkrechten Höhe bezeichnet. (Überlegen Sie sich auch mal einen Beweis!) Wir definieren die *Determinante*

$$\det(\vec{a}, \vec{b}) = \epsilon F = \|\vec{a}\| \|\vec{b}\| \sin \phi$$

wobei ϕ der Winkel zwischen (den Ortsvektoren) \vec{a} und \vec{b} ist, F die Fläche des von ihnen aufgespannten Parallelogramms und $\epsilon = 1$ falls $0 < \phi < \pi$ (*Rechtssystem*) bzw. $\epsilon = -1$



falls $\pi < \phi < 2\pi$ (*Linkssystem*). Es gilt für $\vec{a}, \vec{b}, \vec{c}$, in der gegebenen Ebene gilt (vgl Fig.) die Scherungsinvarianz

$$\det(\vec{a}, \vec{b} + r\vec{a}) = \det(\vec{a}, \vec{b}) = \det(\vec{a} + s\vec{b}, \vec{b})$$

d.h. die Grundfläche ändert sich weder nach Betrag noch nach Vorzeichen, wenn man \vec{b} längs einer Parallelen zu \vec{a} schert, Weiterhin

$$(D1) \quad \det(\vec{a}, \vec{b} + \vec{c}) = \det(\vec{a}, \vec{b}) + \det(\vec{a}, \vec{c}), \quad \det(\vec{b} + \vec{c}, \vec{a}) = \det(\vec{b}, \vec{a}) + \det(\vec{c}, \vec{a})$$

$$(D2 - 3) \quad \det(\vec{a}, r\vec{b}) = r \det(\vec{a}, \vec{b}) = \det(r\vec{a}, \vec{b}), \quad \det(\vec{a}, \vec{a}) = 0$$

$$\det(\vec{a}, \vec{b}) = 0 \Leftrightarrow \vec{a}, \vec{b} \text{ parallel}, \quad \det(\vec{b}, \vec{a}) = -\det(\vec{a}, \vec{b})$$

wie man aus (D1-3) leicht beweist. Z.B. $\det(\vec{a}, \vec{b} + r\vec{a}) = \det(\vec{a}, \vec{b}) + \det(\vec{a}, r\vec{a}) = \det(\vec{a}, \vec{b}) + r \det(\vec{a}, \vec{a}) = \det(\vec{a}, \vec{b}) + r \cdot 0$. Bei “ \Rightarrow ” muss man allerdings voraussetzen, dass es überhaupt Vektoren gibt mit $\det(\vec{x}, \vec{y}) \neq 0$. Auch die Scherungsinvarianz lässt sich aus (D1-3) herleiten. Wählt man als Koordinatensystem der Ebene eine Orthonormalbasis \vec{e}_1, \vec{e}_2 , die ein Rechtssystem ist, so gilt

$$(D4) \quad \det(\vec{e}_1, \vec{e}_2) = 1$$

und man kann die Determinante aus den Koordinaten berechnen

$$\det(\vec{a}, \vec{b}) = \begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix} = a_1 b_2 - a_2 b_1$$

nämlich $\det(a_1\vec{e}_1 + a_2\vec{e}_2, b_1\vec{e}_1 + b_2\vec{e}_2) = \det(a_1\vec{e}_1 + a_2\vec{e}_2, b_1\vec{e}_1) + \det(a_1\vec{e}_1 + a_2\vec{e}_2, b_2\vec{e}_2) = \det(a_1\vec{e}_1, b_1\vec{e}_1) + \det(a_2\vec{e}_2, b_1\vec{e}_1) + \det(a_1\vec{e}_1, b_2\vec{e}_2) + \det(a_2\vec{e}_2, b_2\vec{e}_2) = 0 + a_2 b_1 \det(\vec{e}_2, \vec{e}_1) + a_1 b_2 \det(\vec{e}_1, \vec{e}_2) + 0 = -a_2 b_1 + a_1 b_2$.

Umgekehrt wird durch die Auszeichnung einer Orthonormalbasis als Koordinatensystem die Orientierung der Ebene angegeben und ein Flächenmaß eingeführt

$$\vec{a}, \vec{b} \left\{ \begin{array}{l} \text{positiv orientiert} \\ \text{negativ orientiert} \\ \text{linear abhängig} \end{array} \right\} \Leftrightarrow \det(\vec{a}, \vec{b}) = \left\{ \begin{array}{l} > 0 \\ < 0 \\ = 0 \end{array} \right.$$

$$\text{Fläche Parallelogramm}(\vec{a}, \vec{b}) = |\det(\vec{a}, \vec{b})|$$

13.8 Vektorprodukt

Ein Tripel unabhängiger Vektoren bildet ein *Rechtssystem* im Raum, wenn ihre Richtungen (in der gegebenen Reihenfolge) mit den Richtungen von gestrecktem Daumen und Zeigefinger und abgewinkelttem Mittelfinger der rechten Hand identifiziert werden können - Beweglichkeit des Daumens bis zu 180° gegenüber dem Zeigefinger vorausgesetzt. Entsprechend hat man *Linkssysteme* für die linke Hand. Jedes unabhängige Tripel von Vektoren bildet entweder ein Rechts- oder ein Linkssystem. Welche Hand die rechte ist, ist mathematisch gesehen jedoch eine Frage der Definition: es wird irgendein unabhängiges Tripel zum Rechtssystem deklariert und dadurch die *Orientierung* festgelegt. Alle anderen Rechtssysteme ergeben sich hieraus durch Drehung des Tripels insgesamt und stetigen Scherungen eines der Vektoren gegen die beiden anderen, bei denen die drei Vektoren in keinem Stadium in eine gemeinsame Ebene zu liegen kommen. Wir definieren nun das *Vektor-* oder *äussere Produkt* $\vec{c} = \vec{a} \times \vec{b}$ durch die Bedingungen

$$\|\vec{c}\| = |\det(\vec{a}, \vec{b})|, \quad \vec{c} \perp \vec{a}, \vec{b} \quad \text{und} \quad \vec{a}, \vec{b}, \vec{c} \text{ Rechtssystem oder abhängig.}$$

Unmittelbar geometrisch einsichtig sind die Regeln

$$(G1 - 2) \quad \vec{a} \times \vec{b} = \vec{0} \Leftrightarrow \vec{a}, \vec{b} \text{ parallel}, \quad \vec{a} \times \vec{b} = -\vec{b} \times \vec{a}$$

$$G3) \quad r(\vec{a} \times \vec{b}) = (r\vec{a}) \times \vec{b} = \vec{a} \times (r\vec{b})$$

Es folgt wie bei den ebenen Determinanten die Scherungsinvarianz

$$\vec{a} \times (\vec{b} + r\vec{a}) = \vec{a} \times \vec{b} = (\vec{a} + s\vec{b}) \times \vec{b}$$

(d.h. $\det(\vec{a}, \vec{b})$, Normale und Orientierung bleiben unverändert) und daraus dann die Distributivgesetze

$$(G4) \quad \vec{a} \times (\vec{b} + \vec{c}) = \vec{a} \times \vec{b} + \vec{a} \times \vec{c}, \quad (\vec{b} + \vec{c}) \times \vec{a} = \vec{b} \times \vec{a} + \vec{c} \times \vec{a}.$$

Zum Beweis des ersten Distributivgesetzes dürfen wir zunächst annehmen, dass $\|\vec{a}\| = 1$. Nach Scherung darf man $\vec{b} \perp \vec{a}$ annehmen, ebenso $\vec{c} \perp \vec{a}$. Dann liegen \vec{b} und \vec{c} in einer Ebene mit Normale \vec{a} und man erhält $\vec{a} \times \vec{b}$, $\vec{a} \times \vec{c}$, $\vec{a} \times (\vec{b} + \vec{c})$ in dieser Ebene, indem man \vec{b} , \vec{c} und $\vec{b} + \vec{c}$ jeweils um 90° dreht. (vgl. Folie)

Eine typische Anwendung des Vektorprodukts ist es, aus zwei nicht parallelen Richtungsvektoren \vec{b}, \vec{c} einer Ebene einen Normalenvektor $\vec{n} = \vec{b} \times \vec{c}$ zu berechnen. Wählt man als

Koordinatensystem des Raumes ist eine Orthonormalbasis $\alpha : \vec{e}_1, \vec{e}_2, \vec{e}_3$, die Rechtssystem ist, so hat man

$$\vec{e}_1 \times \vec{e}_2 = \vec{e}_3, \quad \vec{e}_2 \times \vec{e}_3 = \vec{e}_1, \quad \vec{e}_3 \times \vec{e}_1 = \vec{e}_2$$

und die Koordinatendarstellung

$$\vec{a} \times \vec{b} = a_1 b_2 \vec{e}_3 + b_1 a_3 \vec{e}_2 + a_2 b_3 \vec{e}_1 - a_3 b_2 \vec{e}_1 - b_3 a_1 \vec{e}_2 - a_2 b_1 \vec{e}_3$$

$$\vec{a} \times \vec{b}^\alpha = \begin{pmatrix} a_2 b_3 - a_3 b_2 \\ a_3 b_1 - a_1 b_3 \\ a_1 b_2 - a_2 b_1 \end{pmatrix} = \begin{pmatrix} \left| \begin{array}{cc} a_2 & b_2 \\ a_3 & b_3 \end{array} \right| \\ - \left| \begin{array}{cc} a_1 & b_1 \\ a_3 & b_3 \end{array} \right| \\ \left| \begin{array}{cc} a_1 & b_1 \\ a_2 & b_2 \end{array} \right| \end{pmatrix} \quad \text{für } \vec{a}^\alpha = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}, \vec{b}^\alpha = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}.$$

Beweis durch Ausmultiplizieren (Übung!)

13.9 Volumen

Nachdem das Volumen des Einheitswürfels als 1 festgelegt ist, ergibt sich das Volumen eines Spats zu “Grundfläche mal Höhe”. Wir definieren nun die *Determinante* oder das *Spatprodukt* als

$$\det(\vec{a}, \vec{b}, \vec{c}) = \langle \vec{a} \times \vec{b} \mid \vec{c} \rangle = \epsilon V,$$

wobei V das Volumen des von (den Ortsvektoren) $\vec{a}, \vec{b}, \vec{c}$ aufgespannten Spats ist und $\epsilon = 1$, falls $\vec{a}, \vec{b}, \vec{c}$ Rechtssystem, $\epsilon = -1$, andernfalls. Wegen der Eigenschaften von Vektor- und Skalarprodukt gelten die Regeln (D1 – 3) und ihre Konsequenzen entsprechend, z.B. $\det(r\vec{a} + s\vec{a}', \vec{b}, \vec{c}) = r \det(\vec{a}, \vec{b}, \vec{c}) + s \det(\vec{a}', \vec{b}, \vec{c})$ d.h. Linearität in der ersten Spalte und entsprechend in den anderen Spalten. Bei zwei gleichen Spalten erhält man in allen 3 möglichen Fällen Determinante 0. Es folgt Vorzeichenwechsel bei Vertauschung zweier Spalten, z.B. $\det(\vec{b}, \vec{a}, \vec{c}) = -\det(\vec{a}, \vec{b}, \vec{c})$. Also bleibt die Determinante bei zyklischen Vertauschungen unverändert.

Eine positiv orientierte ONB hat Determinante 1 und es gilt

$$\det(\vec{a}, \vec{b}, \vec{c}) = \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} = a_1 b_2 c_3 + b_1 c_2 a_3 + c_1 a_2 b_3 - a_3 b_2 c_1 - b_3 c_2 a_1 - c_3 a_2 b_1.$$

Wir schreiben auch

$$\det(\vec{a}, \vec{b}, \vec{c}) = \det A$$

mit der Matrix

$$A = \begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{pmatrix}$$

Ersetzt man hier c_i durch \vec{e}_i , so erhält man das Vektorprodukt $\vec{a} \times \vec{b}$.

13.10 Übersicht

Skalar mal Vektor ergibt Vektor: $r\vec{a}$. Man könnte auch $\vec{a}r = r\vec{a}$ definieren. Dann gelten alle Rechenregeln, es kann aber in einem Produkt nur ein Vektor auftreten und dividieren darf man durch Vektoren auch nicht. Aber man darf kürzen: Aus $r\vec{a} = r\vec{b}$ folgt $\vec{a} = \vec{b}$ falls $r \neq 0$; aus $r\vec{a} = s\vec{a}$ folgt $r = s$ falls $\vec{a} \neq \vec{0}$.

Das skalare Produkt zweier Vektoren ist ein Skalar: $\langle \vec{a} | \vec{b} \rangle$. Das Assoziativgesetz für drei Vektoren: $\vec{a}\langle \vec{b} | \vec{c} \rangle = \langle \vec{a} | \vec{b} \rangle\vec{c}$ gilt nur, wenn \vec{a} und \vec{c} parallel sind. Kürzen darf man Vektoren auch nicht (und erst recht nicht durch Vektoren dividieren): Aus $\langle \vec{a} | \vec{b} \rangle = \langle \vec{a} | \vec{c} \rangle$ folgt nur, dass \vec{a} auf $\vec{b} - \vec{c}$ senkrecht steht. Immerhin hat man noch Kommutativität und Distributivität; und Assoziativität soweit nur zwei Vektoren beteiligt sind.

Das Vektorprodukt zweier Vektoren ist ein Vektor: $\vec{a} \times \vec{b}$ im Raum. Statt Kommutativität haben wir Antikommutativität, statt Assoziativität Grassmann. Immerhin gilt noch das Distributivgesetz, und man kann Skalare herausziehen. Aus $\vec{a} \times \vec{b} = \vec{a} \times \vec{c}$ folgt nur, dass $\vec{a}, \vec{b}, \vec{c}$ linear abhängig sind.

Die Determinante dreier (zweier) Vektoren ist ein Skalar Dabei muss eine Orientierung des Raums (der Ebene) vorgegeben sein. Man hat die Linearitätseigenschaft in jeder Spalte. Aus $\det(\vec{a}, \vec{b}, \vec{c}) = 0$ folgt nur, dass $\vec{a}, \vec{b}, \vec{c}$ linear abhängig sind, d.h. durch Pfeile in einer Ebene repräsentiert werden können,

Die Orientierung der Ebenen bzw. des Raumes wird durch eine ON-Basis festgelegt. Eine weitere Basis \vec{a}_1, \vec{a}_2 bzw. $\vec{a}_1, \vec{a}_2, \vec{a}_3$ ist dann positiv orientiert, wenn $\det(\vec{a}_1, \vec{a}_2) > 0$ bzw. $\det(\vec{a}_1, \vec{a}_2, \vec{a}_3) > 0$, andernfalls ist sie negativ orientiert. Im realen Raum können wir die positive Orientierung durch die "Rechte-Hand-Regel" festlegen, für Ebenen in "Draufsicht" durch die "gegen-die-Uhr-Regel".

14 Determinanten

14.1 Regeln

Seien der Körper K und eine natürliche Zahl n gegeben. Eine (normierte) *Determinantenform* ist eine Abbildung \det , die jeder $n \times n$ -Matrix A

$$A = (\mathbf{a}_1 \dots \mathbf{a}_n)$$

über K einen Skalar

$$|A| = \det A = \det(\mathbf{a}_1, \dots, \mathbf{a}_n)$$

aus K zuordnet so, dass die folgenden Regeln gelten:

$$(D1) \det(\mathbf{a}_1, \dots, \mathbf{a}_j + \mathbf{b}_j, \dots, \mathbf{a}_n) = \det(\mathbf{a}_1, \dots, \mathbf{a}_j, \dots, \mathbf{a}_n) + \det(\mathbf{a}_1, \dots, \mathbf{b}_j, \dots, \mathbf{a}_n)$$

$$(D2) \det(\mathbf{a}_1, \dots, r\mathbf{a}_j, \dots, \mathbf{a}_n) = r \det(\mathbf{a}_1, \dots, \mathbf{a}_j, \dots, \mathbf{a}_n)$$

d.h. \det ist linear in jeder Spalte. Eine Determinante mit zwei benachbarten gleichen Spalten ist Null:

$$(D3) \det(\dots, \mathbf{a}, \mathbf{a}, \dots) = 0.$$

$$(D4) \quad \det E_n = 1.$$

Ob's sowas gibt, wissen wir vorerst nicht, ziehen aber munter unsere Folgerungen über "die Determinante".

Eine Determinante mit zwei gleichen Spalten ist Null:

$$(D3^+) \quad \det(\dots, \mathbf{a}, \dots, \mathbf{a}, \dots) = 0.$$

(D5) *Die Determinante ändert sich nicht, wenn man zu einer Spalte ein Vielfaches einer anderen addiert;*

(D6) *Die Determinante ändert das Vorzeichen, d.h. man muss mit -1 multiplizieren, wenn man zwei Spalten vertauscht.*

Beweis: Wir zeigen zunächst (D5) und (D6) für den Fall, dass die j -te und k -te Spalte benachbart sind.

$$\begin{aligned} \det(\dots, \mathbf{a}_j, \dots, r\mathbf{a}_j + \mathbf{a}_k, \dots) &=_{(D1)} \det(\dots, \mathbf{a}_j, \dots, r\mathbf{a}_j, \dots) + \det(\dots, \mathbf{a}_j, \dots, \mathbf{a}_k, \dots) \\ &=_{(D2)} r \det(\dots, \mathbf{a}_j, \dots, \mathbf{a}_j, \dots) + \det(\dots, \mathbf{a}_j, \dots, \mathbf{a}_k, \dots) =_{(D3)} 0 + \det(\dots, \mathbf{a}_j, \dots, \mathbf{a}_k, \dots). \\ \det(\dots, \mathbf{a}_j, \dots, \mathbf{a}_k, \dots) + \det(\dots, \mathbf{a}_k, \dots, \mathbf{a}_j, \dots) \\ &=_{(D5)} \det(\dots, \mathbf{a}_j, \dots, \mathbf{a}_j + \mathbf{a}_k, \dots) + \det(\dots, \mathbf{a}_k, \dots, \mathbf{a}_k + \mathbf{a}_j, \dots) \\ &=_{(D1)} \det(\dots, \mathbf{a}_j + \mathbf{a}_k, \dots, \mathbf{a}_k + \mathbf{a}_j, \dots) =_{(D3)} 0. \end{aligned}$$

Hat man (D6) für benachbarte Spalten, so folgt (D3⁺) aus (D3) durch Induktion über den Positions-Abstand der zwei gleichen Spalten

$$\det(\dots \mathbf{a} \dots \mathbf{b} \mathbf{a} \dots) = -\det(\dots \mathbf{a} \dots \mathbf{a} \mathbf{b} \dots) = 0$$

Und dann folgen (D5) und (D6) mit obigem Beweis allgemein. \square

14.2 Eindeutigkeit und Berechnung

Satz 14.1 *Zu gegebenem n und K gibt es höchstens eine normierte Determinantenform. Man berechnet $\det A$ durch Umformen nach (D5) und (D6) auf Dreiecksform (mit Berücksichtigung der Vorzeichenwechsel) und dann Produkt über die Diagonale. Es gilt $\det A \neq 0$ genau dann, wenn A invertierbar ist.*

Beweis. Durch (elementare) Spaltenumformungen nach (D5) und (D6) kann man A in eine Matrix A' in unterer Stufenform mit $\det A' = \det A$ oder $\det A' = -\det A$ überführen - und man weiß, welcher Fall vorliegt. Hat A' eine Spalte $\mathbf{0}$, so $\det A' = 0$ nach (D2). Andernfalls handelt es sich wegen der Stufenform um eine untere Dreiecksmatrix mit Diagonaleinträgen $a_{ii} \neq 0$ und man hat nach (D2) $\det A' = a'_{11} \cdot \dots \cdot a'_{nn} \det A''$, wobei A'' untere Dreiecksmatrix mit Diagonaleinträgen 1 ist. Weitere Umformung nach (D5) überführt A'' in E_n , also $\det A'' = \det E_n = 1$.

Nach Gauß-Jordan ist A genau dann invertierbar, wenn A' eine (strikte) untere Dreiecksmatrix (auf der Diagonalen keine Null, darunter nur Nullen) ist, d.h. $\det A = \det A' \neq 0$.

Korollar 14.2 *Eine Abbildung von $K^{n \times n}$ in K , die (D2), (D5) und (D6) erfüllt, ist durch ihren Wert an der Stelle E_n schon eindeutig bestimmt.*

Beispiel 1 Hier mit Zeilen statt Spalten.

$$\det \begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & 2 & 4 & 2 \\ 0 & 3 & 0 & -3 \\ 1 & 1 & 2 & 6 \end{pmatrix} = 2 \cdot 3 \det \begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & 1 & 2 & 1 \\ 0 & 1 & 0 & -1 \\ 1 & 1 & 2 & 6 \end{pmatrix} = 6 \cdot \det \begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & 1 & 2 & 1 \\ 0 & 1 & 0 & -1 \\ 0 & -1 & -1 & 2 \end{pmatrix}$$

$$= 6 \cdot \det \begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & 1 & 2 & 1 \\ 0 & 0 & -2 & -2 \\ 0 & 0 & 1 & 3 \end{pmatrix} = 6 \cdot (-2) \det \begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & 1 & 2 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 3 \end{pmatrix} = -12 \det \begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & 1 & 2 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 2 \end{pmatrix}$$

und nach Satz ist die gesuchte Determinante gleich $-12 \cdot 2 = -24$. ■

14.3 Produktsatz.

$$\det(AB) = \det A \det B$$

Beweis. Sei A fest, B variabel. Die Abbildung $B \mapsto AB$ erfüllt (M1) und (M2), also erfüllen beide Abbildungen

$$B \mapsto \det AB \quad \text{und} \quad B \mapsto \det A \cdot \det B$$

die Bedingungen (D1-3) und $E_n \mapsto \det A$. Daher stimmen nach Kor.14.2 sie überein. \square

Wir vermerken die wichtigen Spezialfälle

Korollar 14.3

$$\det(rA) = r^n \det A, \quad \det A^{-1} = \frac{1}{\det A}, \quad \det(S^{-1}AS) = \det(A).$$

14.4 Transponieren und Zeilenumformungen

Satz 14.4 $\det A = \det A^t$ und die Regeln (D1)- (D6) gelten entsprechend für Zeilenumformungen

Korollar 14.5 $\det(S^tAS) = (\det S)^2 \det A$. $|\det S| = 1$ falls S orthogonal.

Beweis. Wir betrachten jeweils eine (D2), (D5), bzw. (D6) entsprechende Zeilenumformung $A \mapsto z(A)$, d.h. $z(A)$ entsteht aus A indem

- (a) die i -te Zeile mit s multipliziert wird (i und s fest)
- (b) das s -fache der i -ten Zeile zur l -ten addiert wird ($i \neq l$ und s fest)
- (c) die i -te und die l -te Zeile vertauscht werden ($i \neq l$ fest)

Wir behaupten, dass die Abbildung $A \mapsto \det(z(A))$ jeweils (D2), (D5) und (D6) erfüllt. Dazu genügt es zu beobachten, dass in allen drei Fällen

$$\begin{aligned} z(\mathbf{a}_1, \dots, r\mathbf{a}_j, \dots, \mathbf{a}_n) &= (z(\mathbf{a}_1), \dots, rz(\mathbf{a}_j), \dots, z(\mathbf{a}_n)) \\ z(\dots \mathbf{a}_j \dots, r\mathbf{a}_j + \mathbf{a}_k, \dots) &= (\dots z(\mathbf{a}_j) \dots, rz(\mathbf{a}_j) + z(\mathbf{a}_k), \dots) \\ z(\dots, \mathbf{a}_j, \dots, \mathbf{a}_k, \dots) &= (\dots, z(\mathbf{a}_j), \dots, z(\mathbf{a}_k), \dots) \end{aligned}$$

Wir behaupten, dass für jede Umformung z nach (a), (b) bzw. (c) die Abbildung $A \mapsto \det(z(A))$ mit einer (passenden) der Abbildungen

$$A \mapsto s \det(A), \quad A \mapsto \det(A), \quad A \mapsto -\det(A)$$

übereinstimmt. Letztere erfüllen offensichtlich (D2), (D5) und (D6). Wegen Kor.14.2 haben wir daher die Übereinstimmung nur an der Stelle E_n nachprüfen. Das ist aber nicht schwer.

Somit haben wir für $A \mapsto \det(A)$ die (D2),(D5) und (D6) entsprechenden Regeln für Zeilenumformungen bewiesen. Es folgen (D2), (D5) und (D6) für $A \mapsto \det(A^t)$. Für E_n erhalten wir 1 und haben daher $\det(A^t) = \det(A)$ nach Kor.14.2. \square

14.5 Entwicklung

Als zweckmässige Notation führt man ein: Der *Minor* $A^{k\wedge l}$ ist die $n - 1 \times n - 1$ -Matrix, die man aus A durch Weglassen der k -ten Zeile und l -ten Spalte erhält. Dann hat man die Entwicklung nach der j -ten Spalte bzw. der i -ten Zeile

$$\det A = \sum_{j=1}^n (-1)^{i+j} a_{ij} \det A^{i\wedge j}, \quad \det A = \sum_{i=1}^n (-1)^{i+j} a_{ij} \det A^{i\wedge j}$$

Die Faktoren $(-1)^{i+j}$ merkt man sich am besten nach der ‘Schachbretregel’

$$\begin{pmatrix} + & - & + & \dots \\ - & + & - & \dots \\ + & - & + & \dots \\ \vdots & & & \end{pmatrix}$$

Beweis für Spaltenentwicklung - Zeilen durch Transponieren. Seien $\mathbf{a}_1, \dots, \mathbf{a}_n$ die Spalten von A und j fest. Dann $\mathbf{a}_j = a_{1j}\mathbf{e}_1 + \dots + a_{nj}\mathbf{e}_j$, also wegen (D1-2)

$$\det A = a_{1j} \det(\mathbf{a}_1, \dots, \mathbf{e}_1, \dots, \mathbf{a}_n) + \dots + a_{nj} \det(\mathbf{a}_1, \dots, \mathbf{e}_n, \dots, \mathbf{a}_n)$$

wobei jeweils die j -te Spalte ersetzt wurde. Durch Subtraktion des a_{ik} -fachen der j -ten Spalte von jeweils der k -ten werden alle Einträge der i -ten Zeile 0, ausser der 1 in der j -ten Spalte. Durch $i + j$ Zeilen- bzw. Spaltenvertauschungen bringen wir diese 1 in die linke obere Ecke und haben

$$\det(\mathbf{a}_1, \dots, \mathbf{e}_i, \dots, \mathbf{a}_n) = (-1)^{i+j} \det \begin{pmatrix} 1 & O \\ O & A^{i\wedge j} \end{pmatrix} = (-1)^{i+j} \det A^{i\wedge j}.$$

Für den letzten Schluss benutzen wir wieder die Eindeutigkeit, hier für $n - 1$ und die Abbildung

$$B \mapsto \det \begin{pmatrix} 1 & \mathbf{0}^t \\ \mathbf{0} & B \end{pmatrix}, \quad B \in K^{(n-1) \times (n-1)}$$

Beispiel 2 Entwicklung nach der letzten Spalte liefert

$$\begin{aligned} \|\det\|^2 \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} &= \|\det\|^2 \det \begin{pmatrix} 4 & 5 \\ 7 & 8 \end{pmatrix} - 6 \det \begin{pmatrix} 1 & 2 \\ 7 & 8 \end{pmatrix} + 9 \det \begin{pmatrix} 1 & 2 \\ 4 & 5 \end{pmatrix} \\ &= 3 \cdot (-3) - 6(-6) + 9(-3) = 0. \end{aligned}$$

Beispiel 3 Entwicklung jeweils nach der ersten Spalte liefert

$$\begin{aligned} \det \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ 0 & a_{22} & a_{23} & a_{24} \\ 0 & 0 & a_{33} & a_{34} \\ 0 & 0 & 0 & a_{44} \end{pmatrix} &= a_{11} \det \begin{pmatrix} a_{22} & a_{23} & a_{24} \\ 0 & a_{33} & a_{34} \\ 0 & 0 & a_{44} \end{pmatrix} \\ &= a_{11} a_{22} \det \begin{pmatrix} a_{33} & a_{34} \\ 0 & a_{44} \end{pmatrix} = a_{11} a_{22} a_{33} \det(a_{44}) = a_{11} a_{22} a_{33} a_{44}. \end{aligned}$$

14.6 Cramersche Regel

Satz 14.6 Ist $A \in K^{n \times n}$ mit $\det A \neq 0$ so erhält man die eindeutig bestimmte Lösung x von $Ax = b$ durch

$$x_i = \frac{\det(\mathbf{a}_1, \dots, \mathbf{a}_{i-1}, \mathbf{b}, \mathbf{a}_{i+1}, \dots, \mathbf{a}_n)}{\det A}$$

Beweis. $b = \sum_j x_j \mathbf{a}_j$, also

$$\det(\mathbf{a}_1, \dots, \mathbf{a}_{i-1}, \mathbf{b}, \mathbf{a}_{i+1}, \dots, \mathbf{a}_n) = \sum_j x_j \det(\mathbf{a}_1, \dots, \mathbf{a}_{i-1}, \mathbf{a}_j, \mathbf{a}_{i+1}, \dots, \mathbf{a}_n) = \det A$$

Korollar 14.7 Aus $AX = E$ folgt

$$x_{ij} = \frac{\det(\mathbf{a}_1, \dots, \mathbf{a}_{i-1}, \mathbf{e}_j, \mathbf{a}_{i+1}, \dots, \mathbf{a}_n)}{\det A}$$

Beweis

$$A \begin{pmatrix} x_{1j} \\ \vdots \\ x_{nj} \end{pmatrix} = \mathbf{e}_j$$

14.7 Adjugierte Matrix

$$(adA)A = (\det A)E_n \quad \text{mit } adA = ((-1)^{i+j} \det A^{i \wedge j})^t$$

Beweis. Die Berechnung des j -ten Diagonalelements von $(adA)A$ entspricht gerade dem Entwickeln nach der j -ten Spalte. Um zu sehen, dass in Position j, k mit $k \neq j$ eine Null steht, betrachte man die Matrix B , die aus A entsteht, wenn man die k -te Spalte durch die j -te ersetzt - und selbige beibehält. Dann ist $A^{i \wedge k}$ gleich $B^{i \wedge j}$ bis auf eine (für alle i gleiche) Vertauschung und man erhält den Eintrag bis auf das Vorzeichen als die Entwicklung von B nach der j -ten Spalte, was aber wegen zweier gleicher Spalten Null ergibt. \square

Korollar 14.8 *Ist A invertierbar, so gilt*

$$A^{-1} = \frac{1}{\det A} \operatorname{adj} A$$

14.8 Existenz

Im Falle $n = 1$ sei $\det A = a_{11}$. Sei nun \det für $n-1$ schon definiert. Definiere für $n \times n$ Matrizen $\det A$ durch die Entwicklung nach der ersten Zeile. Beim Nachweis von (D1-2) betrachten wir festes j und Matrizen A, B, C mit Spalten $\mathbf{a}_k = \mathbf{b}_k = \mathbf{c}_k$ für alle $k \neq j$, also $A^{1 \wedge j} = B^{1 \wedge j} = C^{1 \wedge j}$. Gilt $\mathbf{c}_j = \mathbf{a}_j + \mathbf{b}_j$, so folgt mit Induktion $\det C^{1 \wedge k} = \det A^{1 \wedge k} + \det B^{1 \wedge k}$ für $k \neq j$ und man berechnet

$$\begin{aligned} \det C &= (-1)^{1+j}(a_{1j} + b_{1j}) \det C^{1 \wedge j} + \sum_{k \neq j} (-1)^{1+k} a_{1k} \det C^{1 \wedge k} \\ &= (-1)^{1+j}(a_{1j} \det A^{1 \wedge j} + b_{1j} \det B^{1 \wedge j}) + \sum_{k \neq j} (-1)^{1+k} a_{1k} \det(A^{1 \wedge k} + B^{1 \wedge k}) \\ &= (-1)^{1+j} a_{1j} \det A^{1 \wedge j} + \sum_{k \neq j} (-1)^{1+k} a_{1k} \det A^{1 \wedge k} \\ &\quad + (-1)^{1+j} b_{1j} \det B^{1 \wedge j} + \sum_{k \neq j} (-1)^{1+k} b_{1k} \det B^{1 \wedge k} \\ &= \det A + \det B \end{aligned}$$

Gilt $\mathbf{c}_j = r \mathbf{a}_j$, so gilt nach Induktion $\det C^{1 \wedge k} = r \det A^{1 \wedge k}$ für $k \neq j$ und somit

$$\begin{aligned} \det C &= (-1)^{1+j} r a_{1j} \det C^{1 \wedge j} + \sum_{k \neq j} (-1)^{1+k} a_{1k} \det C^{1 \wedge k} \\ &= (-1)^{1+j} r a_{1j} \det A^{1 \wedge j} + \sum_{k \neq j} (-1)^{1+k} a_{1k} r \det A^{1 \wedge k} = r \det A \end{aligned}$$

Hat A zwei gleiche Spalten $\mathbf{a}_k = \mathbf{a}_l$, $l = k + 1$, so gilt das auch für alle $A^{1 \wedge j}$ mit $j \neq k, l$, also $\det A^{1 \wedge j} = 0$. Die Minoren $A^{1 \wedge k}$ und $A^{1 \wedge l}$ stimmen überein und treten mit demselben Vorfaktor $a_{1k} = a_{1l}$ aber entgegengesetztem Vorzeichen in der Entwicklung auf. Also $\det A = 0$.

Schliesslich $\det E_n = (-1)^2 \det(E_n)^{1 \wedge i} + \sum_{j \neq i} 0 \det(E_n)^{1 \wedge j} = \det E_{n-1} = 1$.

15 Inversion, Operatornorm und Banachscher Fixpunktsatz

15.1 Stetigkeit der Inversion

Satz 15.1 *Die Abbildung $\det : \mathbb{R}^{n \times n} \rightarrow \mathbb{R}$ ist stetig. Die Menge $\operatorname{GL}_n(\mathbb{R})$ der invertierbaren reellen $n \times n$ -Matrizen ist offen und die Abbildung $A \mapsto A^{-1}$ stetig (bzgl. beliebiger Normen auf $\mathbb{R}^{n \times n}$).*

Beweis. Stetigkeit von \det mit Entwickeln und Induktion. Dann $\operatorname{GL}_n(\mathbb{R})$ offen als Urbild der offenen Menge $\mathbb{R} \setminus \{0\}$. Nun

$$A \mapsto x_{ij} = \frac{\det(\mathbf{a}_1, \dots, \mathbf{a}_{i-1}, \mathbf{e}_j, \mathbf{a}_{i+1}, \dots, \mathbf{a}_n)}{\det A} \in \mathbb{R}$$

stetig, da Zähler und Nenner stetig von A abhängen, also auch $A \mapsto X = A^{-1}$. \square

15.2 Operatornorm

Seien V, W endlichdimensionale euklidische Vektorräume mit euklidischer Norm $|\vec{x}|$ und $A : V \rightarrow W$ eine lineare Abbildung (ein Schelm, wer an eine Matrix denkt). Da $\{\vec{x} \in V \mid |\vec{x}| = 1\}$ kompakt ist, gibt es

$$|A| = \max\{ |A\vec{x}| \mid |\vec{x}| = 1 \}$$

um mit Linearität folgt

$$|A\vec{x}| = \left| |\vec{x}| \cdot A \frac{1}{|\vec{x}|} \vec{x} \right| = |\vec{x}| \cdot \left| A \frac{1}{|\vec{x}|} \vec{x} \right| \leq |A| \cdot |\vec{x}|$$

und dann für die Komposition

$$|BA\vec{x}| \leq |B| \cdot |A\vec{x}| \leq |A| \cdot |B| \cdot |\vec{x}|$$

also

$$|BA| \leq |B| \cdot |A|$$

Offensichtlich gelten

$$|A + B| \leq |A| + |B|, \quad |\lambda A| = |\lambda| \cdot |A|$$

und $|A| = 0$ impliziert $A = 0$, also hat man eine Norm auf

$$L(V, W) \cong R^{m \times n} \cong \mathbb{R}^{mn}$$

Lemma 15.2 *Ist $A : V \rightarrow V$ linear, und $|A| < 1$ so ist $\text{id} - A$ invertierbar.*

Beweis. Ist $\vec{x} - A\vec{x} = \vec{0}$, so o.B.d.A. $|\vec{x}| = 1$ und somit $|A| \geq 1$. \square

Lemma 15.3 *Schranken. Sei $U \subseteq \mathbb{R}^n$ offen, $f : U \rightarrow \mathbb{R}^n$ stetig differenzierbar, $p + t\vec{h} \in U$ und $|DF(p + t\vec{h})| \leq L$ für alle $t \in [0, 1]$. Dann gilt*

$$|f(p + \vec{h}) - f(p)| \leq L \cdot |\vec{h}|$$

Beweis. Nach dem Mittelwertsetz

$$|f(p + \vec{h}) - f(p)| \leq |Ah| \leq L \cdot |\vec{h}| \quad \text{mit } A = \int_0^1 DF(p + t\vec{h}) dt$$

was man jetzt an besten als Matrix versteht. \square

15.3 Vollständige metrische Räume

Definition 15.4 *Ein metrischer Raum (X, d) wird vollständig genannt, wenn jede Cauchy-Folge $(a_n)_{n \in \mathbb{N}}$ in X konvergiert.*

1. \mathbb{Q} mit der üblichen Metrik und $\mathcal{C}([0, 1])$ mit der Metrik d_1 sind nicht vollständig. Auf dem Raum $\mathcal{C}([0, 1])$ aller stetigen Funktionen $f : [0, 1] \rightarrow \mathbb{R}$ betrachten wir die Metrik $d_1(f, g) = \int_0^1 |f(x) - g(x)| dx$. Finden Sie nicht konvergente Cauchy-Folgen (z.B. $f_n(x) = (x + \frac{1}{n})^{-\frac{1}{2}}$).

2. \mathbb{R} mit der üblichen Metrik ist vollständig.
3. Betrachten wir auf \mathbb{R} die Metrik $d(x, y) = |\arctan x - \arctan y|$, dann ist \mathbb{R} nicht vollständig, denn die Folge $a_n = n$ mit $n \in \mathbb{N}$ ist eine Cauchy-Folge für diese Metrik, konvergiert allerdings nicht. Diese Metrik hat aber dieselben offenen Mengen wie die euklidische. Das bedeutet: *Metrische Vollständigkeit ist keine topologische Eigenschaft*. Ebenso ist der Begriff der Cauchy-Folge kein topologischer Begriff. Ob eine Folge eine Cauchy-Folge ist, kann von der speziellen Metrik abhängen.
4. \mathbb{R}^n mit einer euklidischen Metrik ist vollständig.
5. \mathbb{R}^n mit einer durch eine Norm induzierten Metrik ist vollständig.
6. Es sei T eine Menge, X ein vollständiger metrischer Raum. Dann ist der Raum $\mathcal{B}(T, X)$ aller beschränkten Funktionen $f: T \rightarrow X$ mit der Metrik

$$d(f, g) = \sup_{t \in T} d(f(t), g(t)),$$

ebenfalls vollständig.

7. Ein abgeschlossener Teilraum eines vollständigen metrischen Raumes ist ebenfalls vollständig bzgl. der induzierten Metrik.
8. Es sei T ein metrischer Raum und X ein vollständiger metrischer Raum. Die Menge $\mathcal{C}_b(T, X)$ aller stetigen beschränkten Funktionen in $\mathcal{B}(T, X)$ ist abgeschlossen. Es folgt, daß $\mathcal{C}_b(T, X)$ bzgl. der Supremumsmetrik ein vollständiger metrischer Raum ist.
9. $\mathcal{C}([0, 1])$ ist vollständig bzgl. der Metrik d_∞ .

15.4 Banachscher Fixpunktsatz

Es sei X ein *vollständiger* metrischer Raum. Eine Abbildung $f: X \rightarrow X$ wird *kontrahierend* genannt, wenn es eine reelle Zahl $L < 1$ gibt, so daß

$$d(f(x), f(y)) \leq L \cdot d(x, y) \quad \text{für alle } x, y \in X.$$

Die kleinste Zahl L mit dieser Eigenschaft wird auch *Lipschitzkonstante* von f genannt.

Satz 15.5 *Jede kontrahierende Abbildung eines vollständigen metrischen Raumes X in sich hat genau einen Fixpunkt.*

Beweis. Es sei nun $f: X \rightarrow X$ kontrahierend mit Lipschitzkonstante L . Ferner sei $a \in X$ beliebig. Wir betrachten die Folge $(a_n)_{n \in \mathbb{N}}$ mit

$$\begin{aligned} a_0 &= a \\ a_1 &= f(a) \\ a_2 &= f(a_1) = f^2(a) \\ &\vdots \\ a_{k+1} &= f(a_k) = f^{k+1}(a) \\ &\vdots \end{aligned}$$

Es gilt dann

$$\begin{aligned} d(a_n, a_{n+1}) &= d(f(a_{n-1}), f(a_n)) \\ &\leq L \cdot d(a_{n-1}, a_n) \\ &\leq L^2 \cdot d(a_{n-2}, a_{n-1}) \\ &\vdots \\ &\leq L^n \cdot d(a_0, a_1) \end{aligned}$$

also

$$d(a_n, a_{n+1}) \leq L^n \cdot M \quad \text{mit } M = d(a, f(a)) \quad (1)$$

Für $n \leq m$ gilt nun

$$\begin{aligned} d(a_n, a_m) &\leq d(a_n, a_{n+1}) + d(a_{n+1}, a_{n+2}) + \cdots + d(a_{m-1}, a_m) \\ &\leq L^n \cdot M + L^{n+1} \cdot M + \cdots + L^{m-1} \cdot M \\ &= L^n \cdot M \cdot (1 + L + L^2 + \cdots + L^{m-1-n}) \\ &< L^n \cdot M \cdot \left(\sum_{i=1}^{\infty} L^i \right) \end{aligned}$$

und damit

$$d(a_n, a_m) \leq L^n \cdot \frac{M}{1-L} \quad \text{falls } n \leq m \quad (2)$$

Es folgt, daß $(a_n)_{n \in \mathbb{N}}$ eine Cauchy-Folge ist: Zu gegebenem $\varepsilon > 0$ wähle N so, daß $L^N < \varepsilon \cdot \frac{1-L}{M}$. Für alle $n, m \geq N$ folgt dann $d(a_n, a_m) < \varepsilon$.

Also hat die Folge $(a_n)_{n \in \mathbb{N}}$ einen Grenzwert, den wir mit a_∞ bezeichnen wollen. (Hier verwenden wir die Vollständigkeit von X .) Nun gilt

$$d(a_n, a_\infty) = d(a_n, \lim_{m \rightarrow \infty} a_m) = \lim_{m \rightarrow \infty} d(a_n, a_m),$$

also folgt mit (2)

$$d(a_n, a_\infty) \leq L^n \cdot \frac{M}{1-L} \quad (3)$$

Andererseits ist eine kontrahierende Abbildung f auch stetig. Folglich gilt

$$f(a_\infty) = f\left(\lim_{n \rightarrow \infty} a_n\right) = \lim_{n \rightarrow \infty} f(a_n) = \lim_{n \rightarrow \infty} a_{n+1} = a_\infty,$$

d.h. a_∞ ist ein Fixpunkt von f . Dieser Fixpunkt ist eindeutig bestimmt, denn ist b ein weiterer Fixpunkt von f , so gilt

$$d(b, a_\infty) = d(f(b), f(a_\infty)) \leq L \cdot d(b, a_\infty).$$

Da $L < 1$ ist, kann dies nur gelten, wenn $d(b, a_\infty) = 0$, d.h. $a_\infty = b$. \square

Im Gegensatz zum Tarskischen Fixpunktsatz ist dieser Fixpunktsatz konstruktiv: Man startet mit einem beliebigen Anfangspunkt a , wendet f wiederholt an, d.h. setzt $a_0 = a$, $a_1 = f(a)$, \dots , $a_{k+1} = f(a_k)$, \dots , und gewinnt den Fixpunkt als Grenzwert dieser Folge. Die Abschätzung (3) liefert sogar eine Abschätzung der Güte der Approximation des Fixpunktes nach n Schritten.

Bemerkung 15.6 *Geben Sie eine Abbildung f von einem vollständigen metrischen Raum in sich an, welche die Bedingung $d(f(x), f(y)) < d(x, y)$ für alle $x, y \in X$ erfüllt, aber keinen Fixpunkt besitzt.*

Man kann also bei dem Banachschen Fixpunktsatz auf die Voraussetzung $L < 1$ nicht verzichten.

16 Implizite Funktionen und Umkehrabbildungen

Die Umkehrabbildungen, impliziten Funktionen und Extrema unter Nebenbedingungen werden nach dem Skript von Prof. Hieber (Analysis-II-ss08.pdf auf unserer Homepage) behandelt. Die beiden Kapitel 16 und 17 hier geben eine Darstellung, die abstrakte Methoden vermeidet und eher für Nostalgiker geeignet ist.

16.1 Implizite Funktionen

Sei $g : (a, b) \times (c, d) \rightarrow \mathbb{R}$ eine Funktion von zwei reellen Veränderlichen. Wir betrachten das Problem, die Gleichung $g(x, y) = 0$ nach y aufzulösen. Wir suchen also eine Funktion $f : I \rightarrow \mathbb{R}$ auf einem geeigneten Intervall I , so dass $f(I) \subseteq (c, d)$ und

$$g(x, y) = 0 \Leftrightarrow y = f(x) \quad \text{für alle } x \in I, y \in (c, d)$$

Im Falle der Existenz einer solchen Funktion f haben wir durch $y = f(x), x \in I$, eine Auflösung der Gleichung $g(x, y) = 0$ gegeben. Man sagt auch, dass die Funktion f durch die Gleichung $g(x, y)$ *implizit* festgelegt wird.

Satz 16.1 *Sei $g : (a, b) \times (c, d) \rightarrow \mathbb{R}$ eine einmal stetig partiell differenzierbare Funktion. Weiter sei $(x_0, y_0) \in (a, b) \times (c, d)$ mit $g(x_0, y_0) = 0$ und $g_y(x_0, y_0) \neq 0$. Dann gibt es ein Intervall $U \subseteq (a, b)$ um x_0 , ein Intervall $V \subseteq (c, d)$ um y_0 und eine stetig differenzierbare Funktion $f : U \rightarrow V$ mit $f(x_0) = y_0$ und*

$$g(x, y) = 0 \Leftrightarrow y = f(x) \quad \text{für alle } x \in U, y \in V$$

Wählt man U so klein, dass $g_y(x, f(x)) \neq 0$ für alle $x \in U$, so gilt weiter

$$f'(x) = -\frac{g_x(x, f(x))}{g_y(x, f(x))} \quad \text{für } x \in U. \quad (4)$$

Unter den Voraussetzungen des Satzes ist die Gleichung $g(x, y) = 0$ um (x_0, y_0) lokal nach y auflösbar. Es ist dabei bemerkenswert, dass $f'(x_0)$ bestimmt werden kann, ohne f explizit zu kennen:

$$f'(x_0) = -\frac{g_x(x_0, y_0)}{g_y(x_0, y_0)}.$$

Ist Existenz und Differenzierbarkeit von f gegeben, so folgt (4) leicht durch Anwendung der Kettenregel auf die Gleichung $g(x, f(x)) = 0$:

$$g_x(x, f(x)) \cdot 1 + g_y(x, f(x)) \cdot f'(x) = 0.$$

Beispiel 14 Sei $g : \mathbb{R}^2 \rightarrow \mathbb{R}$ gegeben durch $g(x, y) = x^2 + y^2 - 1$. Die Gleichung $g(x, y) = 0$ ist genau dann erfüllt, wenn (x, y) auf der Einheitskreislinie um $(0, 0)$ liegt. Für $x \notin [-1, 1]$ hat $g(x, y)$ daher keine Lösung. Für $x = \pm 1$ gibt es natürlich die eindeutig bestimmte Lösung $y = 0$, man kann aber die Lösungen in einer Umgebung von $x = +1$ bzw. $x = -1$ nicht als eine *Funktion* von x darstellen (Warum?).

Schließlich gibt es für jeden Punkt $x \in (-1, 1)$ zwei Lösungen von $g(x, y) = 0$, nämlich $y_1 = \sqrt{1 - x^2}$ und $y_2 = -\sqrt{1 - x^2}$. Betrachten wir dagegen die Gleichung $g(x, y) = 0$ nur für solche (x, y) , die in einer $\frac{1}{2}$ -Umgebung von $(0, 1)$ liegen so existiert in dieser Umgebung genau eine Lösung, nämlich $f(x) = \sqrt{1 - x^2}$. Man beachte, dass $g_y(0, 1) = 2 \neq 0$ aber $g_y(1, 0) = 0$. Aus (4) erhalten wir

$$f'(x) = -\frac{2x}{2f(x)} = -\frac{x}{\sqrt{1 - x^2}} \quad \text{für } x \in (-1/2, 1/2)$$

was man natürlich auch leicht direkt nachrechnet. Man beachte, dass es für dieses Beispiel einfach war, eine Umgebung von $x_0 = 0$ anzugeben, auf der y durch x eindeutig bestimmt ist (während Satz 16.1 nur die Existenz einer solchen Umgebung behauptet). ■

Beweis. O.B.d.A. $g_y(x_0, y_0) > 0$. Wegen der Stetigkeit von g_y können wir U_0 und V_0 so wählen, dass $g_y(x, y) > 0$ für alle $(x, y) \in U_0 \times V_0$. Dann ist $g(x, y)$ in diesem Bereich streng monoton wachsend und es gibt $y_1 < y_0 < y_2$ mit $g(x_0, y_1) < 0 < g(x_0, y_2)$. Wegen der Stetigkeit gibt es offene Intervalle I_1, I_2 um x_0 mit

$$g(x, y_1) < 0 \text{ für } x \in I_1, \quad g(x, y_2) > 0 \text{ für } x \in I_2$$

Sei $x_0 \in I = [x_1, x_2] \subseteq I_1 \cap I_2$. Dann

$$g(x, y_1) < 0 < g(x, y_2) \quad \text{für } x \in I$$

Also gibt es zu festem $x \in I$ nach dem Zwischenwertsatz und wegen der Monotonie von $g(x, y)$ eindeutig bestimmtes $y \in (y_1, y_2)$ mit $g(x, y) = 0$ und wir setzen $y = f(x)$.

Zur Stetigkeit: Sei \tilde{x} ein innerer Punkt von I , also $\tilde{y} = f(\tilde{x})$ definiert und $y_1 \leq \tilde{y} - \varepsilon < \tilde{y} + \varepsilon \leq y_2$ für alle hinreichend kleinen $\varepsilon > 0$. Wir können die obige Konstruktion eines Rechtecks mit $\tilde{y}_1 = \tilde{y} - \varepsilon$ und $\tilde{y}_2 = \tilde{y} + \varepsilon$ wiederholen und \tilde{x}_1, \tilde{x}_2 so wählen, dass $[\tilde{x}_1, \tilde{x}_2] \times [\tilde{y}_1, \tilde{y}_2]$ ganz in $[x_1, x_2] \times [y_1, y_2]$ enthalten ist. Die dazu gehörige Funktion \tilde{f} stimmt dann auf $[\tilde{x}_1, \tilde{x}_2]$ mit f überein. Andererseits $\tilde{x} \in (\tilde{x}_1, \tilde{x}_2)$ und $|f(\tilde{x}) - f(x)| < \tilde{y}_2 - \tilde{y}_1 \leq 2\varepsilon$ für $x \in [\tilde{x}_1, \tilde{x}_2]$. Also ist f an \tilde{x} stetig.

Zum Beweis der Differenzierbarkeit benutzen wir die Differenzierbarkeit von $g(x, y)$

$$g(x + h, y + k) = g(x, y) + g_x(x, y)h + g_y(x, y)k + \phi(h, k)\sqrt{h^2 + k^2} \quad \phi(h, k) \rightarrow 0$$

Hier sei $y = f(x)$ und $k = f(x + h) - y \rightarrow 0$ wegen der Stetigkeit von f . Es folgt

$$0 = g_x(x, y)h + g_y(x, y)k + \phi(h, k)\sqrt{h^2 + k^2}$$

und mit Division durch $g_y(x, y)h$

$$0 = \frac{g_x(x, y)}{g_y(x, y)} + \frac{k}{h} + \frac{1}{g_y(x, y)}\phi(h, k)\sqrt{1 + \frac{k^2}{h^2}}$$

Also

$$f'(x) = \lim_{h \rightarrow 0} \frac{f(x+h) - f(x)}{h} = \lim_{h \rightarrow 0} \frac{k}{h} = -\frac{g_x(x, y)}{g_y(x, y)} \quad \square$$

16.2 Implizite Vektorfunktionen

Für vektorwertige Funktionen mehrerer Veränderlicher kann man Satz 16.1 wie folgt verallgemeinern.

Satz 16.2 (Satz über implizite Funktionen) *Seien $U \subseteq \mathbb{R}^n$ und $V \subseteq \mathbb{R}^m$ offen, und für $k = 1, \dots, m$ sei $g_k : U \times V \rightarrow \mathbb{R}$ eine einmal stetig partiell differenzierbare Funktion. Sei $(x_0, y_0) \in U \times V$ ein Punkt mit*

$$g_k(x_0, y_0) = 0 \quad \text{für jedes } k$$

und mit

$$\det \frac{\partial(g_1, \dots, g_m)}{\partial(y_1, \dots, y_m)}(x_0, y_0) = \det \begin{pmatrix} \frac{\partial g_1}{\partial y_1}(x_0, y_0) & \dots & \frac{\partial g_1}{\partial y_m}(x_0, y_0) \\ \vdots & & \vdots \\ \frac{\partial g_m}{\partial y_1}(x_0, y_0) & \dots & \frac{\partial g_m}{\partial y_m}(x_0, y_0) \end{pmatrix} \neq 0,$$

so gibt es eine Umgebung $U' \subseteq U$ von x_0 , eine Umgebung $V' \subseteq V$ von y_0 , sowie m stetig differenzierbare Funktionen $f_j : U' \rightarrow \mathbb{R}$ mit $f = (f_1, \dots, f_m)^T : U' \rightarrow V'$ und

$$(f_1(x_0), \dots, f_m(x_0))^T = y_0$$

sowie

$$g_k(x_1, \dots, x_n, f_1(x_1, \dots, x_n), \dots, f_m(x_1, \dots, x_n)) = 0$$

für alle $k = 1, \dots, m$ und alle $(x_1, \dots, x_n) \in U'$.

Beispiel 15 Bei der Transformation kartesischer Koordinaten (x_1, x_2) in Polarkoordinaten (r, φ) ist $x_1 = r \cos \varphi$, $x_2 = r \sin \varphi$. Mit $x = (x_1, x_2)$ und $y = (y_1, y_2) = (r, \varphi)$ haben wir

$$g_1(x, y) = x_1 - r \cos \varphi, \quad g_2(x, y) = x_2 - r \sin \varphi.$$

Für die Funktionaldeterminante gilt

$$\det \frac{\partial(g_1, g_2)}{\partial(r, \varphi)} = \det \begin{pmatrix} -\cos \varphi & r \sin \varphi \\ -\sin \varphi & -r \cos \varphi \end{pmatrix} = r.$$

Für $r \neq 0$ ist demnach (lokal) eine eindeutige Auflösung nach r und φ möglich. ■

16.3 Umkehrfunktionen

Eine wichtige Folgerung aus Satz 16.2 ist der folgende Satz.

Satz 16.3 (Satz über die Umkehrfunktion) Sei $U \subseteq \mathbb{R}^n$ offen und $F : U \rightarrow \mathbb{R}^n$ einmal stetig partiell differenzierbar. Ist $J_F(x_0)$ invertierbar für ein $x_0 \in U$, so ist F um x_0 lokal umkehrbar, d.h. es gibt Umgebungen $U' \subseteq U$ von x_0 und $V' \subseteq \mathbb{R}^n$ von $F(x_0)$, so dass $F|_{U'} : U' \rightarrow V'$ umkehrbar ist, d.h. es gibt $G : V' \rightarrow U'$ mit $F(x) = y \Leftrightarrow G(y) = x$ für alle $x \in U', y \in V'$. Die lokale Umkehrfunktion ist ebenfalls einmal stetig partiell differenzierbar und es gilt $J_G(y) = J_F(x)^{-1}$ für $x \in U'$ und $y = G(x)$.

Für den Beweis wendet man Satz 16.2 auf die Funktion

$$x - F(y)$$

an. ■

16.4 Matrixnorm

Für eine Matrix $A \in \mathbb{R}^{n \times n}$ definieren wir

$$|A| = \sqrt{\sum_{ij} a_{ij}^2}$$

d.h. die Länge von A , wenn wir A als Spalte schreiben. Es gilt

$$|AB| \leq |A| \cdot |B|$$

In der Tat, ist $AB = C = (c_{ij})$ so ist c_{ij} das Skalarprodukt der i -ten Zeile von A mit der j -ten Spalte von B , also nach Cauchy-Schwarz

$$c_{ij}^2 \leq \left(\sum_k a_{ik}^2\right) \left(\sum_k b_{kj}^2\right)$$

und die Behauptung folgt durch Aufsummieren über i, j

16.5 Auflösung

Satz 16.4 Sei $U \subseteq \mathbb{R}^n$ konvex und $F : U \rightarrow \mathbb{R}^n$ stetig differenzierbar. Es gebe eine konstante invertierbare Matrix $A \in \mathbb{R}^{n \times n}$ so, dass

$$|J_F(\vec{x})A^{-1} - E| < 1 \quad \text{für } \vec{x} \in U$$

Dann gilt $F(\vec{p}) \neq F(\vec{x})$ für alle $\vec{p} \neq \vec{x}$ in U

Beweis. Für $\vec{x} \in U$ und beliebige \vec{h}

$$|(J_F(\vec{x}) - A)\vec{h}| = |(J_F(\vec{x})A^{-1}E)(A\vec{h})| \leq |(J_F(\vec{x})A^{-1}E)| \cdot |A\vec{h}| < |A\vec{h}|$$

Angenommen $F(\vec{x}) = F(\vec{p})$ mit $\vec{h} = \vec{x} - \vec{p} \neq \vec{0}$. Nach dem Mittelwertsatz

$$0 = F(\vec{p} + \vec{h}) - F(\vec{p}) = \int_0^1 J_f(\vec{p} + t\vec{h})\vec{h} dt$$

Andererseits $\int_0^1 A\vec{h}dt = A\vec{h}$ und $A\vec{h} \neq \vec{0}$ da A invertierbar. Somit

$$0 \neq |A\vec{h}| = \left| \int_0^1 (A - J_f(\vec{p} + t\vec{h}))\vec{h}dt \right| \leq \int_0^1 |(A - J_f(\vec{p} + t\vec{h}))\vec{h}|dt$$

und das führt zu dem Widerspruch

$$1 \leq \int_0^1 M dt \leq \int_0^1 dt < 1$$

woebi

$$M = \max\left\{ \frac{|(A - J_f(\vec{p} + t\vec{h}))\vec{h}|}{|A\vec{h}|} \mid t \in [0, 1] \right\} < 1.$$

Beweis von Satz 16.2. Der Fall $m = 2$ ist hinreichend allgemein. Wir schreiben

$$G_{\vec{x}}(\vec{y}) = \begin{pmatrix} g_1(\vec{x}, \vec{y}) \\ g_2(\vec{x}, \vec{y}) \end{pmatrix}$$

Nach Voraussetzung $\det J_{G_{\vec{x}_0}}(\vec{y}_0) \neq 0$ und $\det J_{G_{\vec{x}}}(\vec{y})$ stetig als Funktion von (\vec{x}, \vec{y}) . Daher $\det J_{G_{\vec{x}}}(\vec{y}) \neq 0$ auf einer passenden Umgebung, o.B.d.A. auf $U \times V$. Nun

$$J_{G_{\vec{x}}}(y) \rightarrow J_{G_{\vec{x}_0}}(y_0) \text{ für } (\vec{x}, \vec{y}) \rightarrow (\vec{x}_0, \vec{y}_0)$$

also wegen Stetigkeit der Inversion o.B.d.A.

$$|J_{G_{\vec{x}}}(\vec{y}) \cdot (J_{G_{\vec{x}_0}}(\vec{y}_0))^{-1} - E| < 1$$

auf $U_1 = U_\varepsilon(\vec{x}_0)$ und $V' = U_\varepsilon(\vec{y}_0)$ deren Abschluss in U bzw. V enthalten ist. Der vorangehende Satz garantiert, dass es in diesem Bereich zu \vec{x} höchstens ein \vec{y} gibt mit $G_{\vec{x}}(\vec{y}) = \vec{0}$. Es ist also die Existenz zu zeigen. Setze

$$g(\vec{x}, \vec{y}) = g_1(\vec{x}, \vec{y})^2 + g_2(\vec{x}, \vec{y})^2$$

Unter der Bedingung $\|\vec{y} - \vec{y}_0\| = \varepsilon$ nimmt diese Funktion wegen der Kompaktheit ein Minimum an, dieses sei 2μ . Wegen der gleichmässigen Stetigkeit von g kann man $\delta < \varepsilon$ finden mit

$$|g(\vec{x}_0, \vec{y}) - g(\vec{x}, \vec{y})| < \mu \text{ für } \|\vec{x}_0 - \vec{x}\| \leq \delta, \vec{y} \in V'$$

und es folgt

$$|g(\vec{x}, \vec{y}_0)| < \mu \text{ für } \|\vec{x}_0 - \vec{x}\|$$

Also wird

$$\min\{g(\vec{x}, \vec{y}) \mid \|\vec{x} - \vec{x}_0\| \leq \varepsilon\}$$

an (\vec{x}, \vec{y}) mit $\|\vec{x} - \vec{x}_0\| < \varepsilon$ und $\vec{y} \in V'$ angenommen und wir können und auf das notwendige Kriterium für Extrema berufen

$$\text{grad}g(\vec{x}, \vec{0}) = \vec{0}$$

explizit nach der Kettenregel

$$2 \frac{\partial g_1}{\partial y_1}(\vec{x}, \vec{y}) \cdot g_1(\vec{x}, \vec{y}) + 2 \frac{\partial g_2}{\partial y_1}(\vec{x}, \vec{y}) \cdot g_2(\vec{x}, \vec{y}) = 0$$

$$2 \frac{\partial g_1}{\partial y_2}(\vec{x}, \vec{y}) \cdot g_1(\vec{x}, \vec{y}) + 2 \frac{\partial g_2}{\partial y_2}(\vec{x}, \vec{y}) \cdot g_2(\vec{x}, \vec{y}) = 0$$

Fasst man hier $X_1 = g_1(\vec{x}, \vec{y})$ und $X_2 = g_2(\vec{x}, \vec{y})$ als Unbestimmte auf, so hat man ein homogenes lineares Gleichungssystem mit Determinante $\neq 0$ und es folgt $X_1 = X_2 = 0$. Damit wird durch $F(\vec{x}) = \vec{y}$ in der Tat eine Auflösung der Gleichungen angegeben und diese ist, wie oben bemerkt, durch \vec{x} eindeutig bestimmt. Die Stetigkeit von F folgt wie in 30.1.

16.6 Ableitung der Auflösung

Wir zeigen Ableitbarkeit nach x_i und bestimmen die Ableitung, d.h. wir sehen $x_l, l \neq i$ als fest an und brauchen sie nicht zu notieren. Sei als $\vec{y} = F(x_i)$ und $\vec{y} + \vec{k} = F(x_i + h)$. Somit

$$g_k(x_i + h, \vec{y} + \vec{k}) - g_k(x_i, \vec{y}) = 0$$

und nach dem Mittelwertsatz für skalare Funktionen mit passendem $\vec{\xi}$

$$\frac{\partial g_k}{\partial x_i}(\vec{\xi})h + \sum_j \frac{\partial g_k}{\partial y_j}(\vec{\xi})k_j = 0$$

$$-\frac{\partial g_k}{\partial x_i}(\vec{\xi}) = \sum_j \frac{\partial g_k}{\partial y_j}(\vec{\xi}) \frac{k_j}{h}$$

Man hat also m Gleichungen für die m Unbestimmten $\frac{k_j}{h}$ und nach Cramer

$$\frac{k_j}{h} = \frac{\det A_j(\vec{\xi})}{\det A(\vec{\xi})}$$

wobei

$$A(\vec{\xi}) = \left(\frac{\partial g_k}{\partial y_j}(\vec{\xi}) \right)_{k,j=1,\dots,m} = \begin{pmatrix} \frac{\partial g_1}{\partial y_1} & \cdots & \frac{\partial g_1}{\partial y_m} \\ \vdots & & \vdots \\ \frac{\partial g_m}{\partial y_1} & \cdots & \frac{\partial g_m}{\partial y_m} \end{pmatrix}(\vec{\xi})$$

d.h. die "Jacobimatrix" von G bzgl. der y_1, \dots, y_m , und in $A_j(\vec{\xi})$ die j -Spalte ersetzt wird durch

$$\begin{pmatrix} \frac{\partial g_1}{\partial x_i}(\vec{\xi}) \\ \vdots \\ \frac{\partial g_m}{\partial x_i}(\vec{\xi}) \end{pmatrix}$$

Nun folgt

$$\frac{\partial F_j}{\partial x_i} = \lim_{h \rightarrow 0} \frac{k_j}{h} = \lim_{h \rightarrow 0} \frac{\det A_j(\vec{\xi})}{\det A(\vec{\xi})} = \frac{\det A_j(\vec{x}_0, \vec{y}_0)}{\det A(\vec{x}_0, \vec{y}_0)}$$

und damit haben wir stetige Differenzierbarkeit und eine Formel für die Ableitung.

17 Extrema unter Nebenbedingungen

17.1 Vorbemerkungen

Im Abschnitt 11 haben wir lokale Extrema von Funktionen studiert, die auf einer offenen Teilmenge des \mathbb{R}^n definiert sind. Wir betrachten nun Situationen, die in praktischen Problemen viel häufiger auftreten: Wir untersuchen Extrema von Funktionen unter Nebenbedingungen.

Beispiel. Unter allen Rechtecken mit der Fläche 1 suchen wir das mit dem kleinsten Umfang. Bezeichnen wir die Seiten des Rechtecks mit x und y , so wollen wir die Funktion $f(x, y) = 2x + 2y$ unter der *Nebenbedingung* $xy = 1$ bzw. $g(x, y) := xy - 1 = 0$ minimieren. ■

Dabei wollen wir *vermeiden*, die Nebenbedingungen nach einer der Variablen umzuformen und in die zu minimierende Funktion einzusetzen, da dies in der Regel Schwierigkeiten bereitet. Wir lernen nun die *Methode der Lagrange-Multiplikatoren* kennen, mit der die Schwierigkeiten des Umformens nach einer Variablen vermieden werden können.

Zunächst einige Begriffe. Wir suchen Extrema einer Funktion $f : \mathbb{R}^n \rightarrow \mathbb{R}$, wobei die Menge der zulässigen Punkte $x = (x_1, \dots, x_n)$ durch eine Nebenbedingung der Form $g(x_1, \dots, x_n) = 0$ eingeschränkt wird. Man nennt $p = (p_1, \dots, p_n)$ eine *relative Maximalstelle (Minimalstelle) von f unter der Nebenbedingung $g(x) = 0$* , wenn es eine Umgebung U von p gibt, so dass $f(x) \leq f(p)$ (bzw. $f(x) \geq f(p)$) für alle $x \in U$ mit $g(x) = 0$.

17.2 Motivation

Im obigen Beispiel definiert $g(x, y)$ eine Hyperbel. Die Höhenlinien $f(x, y) = c$ sind Geraden von Steigung -1 . Bewegt man sich auf der Hyperbel in Richtung der positiven x -Achse, so fallen die Werte von f , solange $x < 1$ und sie steigen für $x > 1$. Für $x = 1$ ist die Tangente an die Hyperbel gerade eine Höhenlinie (hier $f(x, y) = 4$), allgemeiner ausgedrückt eine Tangente an die Höhenlinie. Dasselbe Phänomen kennen wir von den quadratischen Formen mit der Nebenbedingung $\|x\| = 1$.

Nehmen wir an, dass in der Nähe von p die Punkte (x, y) mit $g(x, y) = 0$ gerade die Punkte auf der Kurve

$$\vec{x}(t) = \begin{pmatrix} x(t) \\ y(t) \end{pmatrix}, \quad \frac{\partial}{\partial t} \vec{x}(t) \neq \vec{0}$$

sind, so folgt aus $g(x, y) = 0$ nach der Kettenregel

$$0 = \frac{\partial g}{\partial t} = \langle \text{grad} g \mid \frac{\partial x}{\partial t} \rangle$$

d.h. der Gradient von g steht auf dem Tangentialvektor senkrecht. Hat man für t_0 ein Extremum von $f(x(t), y(t))$ so folgt

$$0 = \frac{\partial f}{\partial t} = \langle \text{grad} f(p) \mid \frac{\partial x}{\partial t}(t_0) \rangle, \quad p = \begin{pmatrix} x(t_0) \\ y(t_0) \end{pmatrix}$$

Der Tangentialvektor steht also auf beiden Gradienten senkrecht und es folgt
 $\text{grad}f(p)$ ist skalares Vielfaches von $\text{grad}g(p)$

Das ist der Hintergrund des folgenden Satzes. Für das Beispiel haben wir

$$\text{grad}f(x, y) = \begin{pmatrix} 2 \\ 2 \end{pmatrix}, \text{grad}g(x, y) = \begin{pmatrix} y \\ x \end{pmatrix}$$

also $x = y$ und somit $x = y = 1$ aus $xy = 1$.

17.3 Lagrange-Multiplikatoren

Satz 17.1 Sei $U \subseteq \mathbb{R}^n$ offen, und $f : U \rightarrow \mathbb{R}$, $g : U \rightarrow \mathbb{R}$ seien einmal stetig partiell differenzierbar auf U . Hat die Funktion f an der Stelle $p \in U$ ein relatives Extremum unter der Nebenbedingung $g = 0$, und sind nicht alle partiellen Ableitungen von g an der Stelle p gleich 0, so existiert ein $\lambda \in \mathbb{R}$ mit

$$f_{x_i}(p) + \lambda g_{x_i}(p) = 0 \quad \text{für } i = 1, \dots, n. \quad (17.5)$$

Die Gleichungen (17.5) lassen sich zusammenfassen zu

$$(\text{grad}f)(p) + \lambda(\text{grad}g)(p) = 0.$$

Zur praktischen Bestimmung relativer Extrema von f unter der Nebenbedingung $g(x) = 0$ kann man wie folgt vorgehen. Man bildet eine Hilfsfunktion $L : U \times \mathbb{R} \rightarrow \mathbb{R}$, die neben x_1, \dots, x_n von einer reellen Variablen λ abhängt:

$$L(x_1, \dots, x_n, \lambda) := f(x_1, \dots, x_n) + \lambda g(x_1, \dots, x_n).$$

Diese Funktion L heißt auch *Lagrange-Funktion*, und λ heißt *Lagrange-Multiplikator*. Nun betrachtet man das System der $n + 1$ Gleichungen.

$$\begin{aligned} L_{x_1}(x, \lambda) &= f_{x_1}(x) + \lambda g_{x_1}(x) &= 0 \\ L_{x_2}(x, \lambda) &= f_{x_2}(x) + \lambda g_{x_2}(x) &= 0 \\ &\vdots & \\ L_{x_n}(x, \lambda) &= f_{x_n}(x) + \lambda g_{x_n}(x) &= 0 \\ L_\lambda(x, \lambda) &= g(x) &= 0. \end{aligned} \quad (17.6)$$

Auf dieses System wären wir auch gestoßen, wenn wir auf die übliche Weise die lokale Extrema der Lagrange-Funktion L bestimmen wollten. Jede Lösung $(x_1^{(0)}, \dots, x_n^{(0)}, \lambda^{(0)})$ von (17.6) liefert nun einen *Kandidaten* $p = (x_1^{(0)}, \dots, x_n^{(0)})$ für eine relative Extremstelle von f unter der Nebenbedingung $g(x) = 0$. Ob ein solcher Kandidat tatsächlich eine relative Extremstelle unter der Nebenbedingung $g(x) = 0$ ist, muss jeweils gesondert untersucht werden. Der Wert $\lambda^{(0)}$ ist dabei irrelevant.

17.4 Beispiele

Beispiel 16 In unserem Beispiel von oben ist $U = (0, \infty) \times (0, \infty)$, $f(x, y) = 2x + 2y$ und $g(x, y) = xy - 1$. Die Lagrange-Funktion lautet

$$L(x, y, \lambda) = 2x + 2y + \lambda(xy - 1),$$

und partielles Ableiten nach x, y und λ sowie Nullsetzen liefert

$$\begin{aligned} 2 + \lambda x &= 0 \\ 2 + \lambda y &= 0 \\ xy &= 1. \end{aligned}$$

Aus den ersten beiden Gleichungen folgt $x = -\frac{2}{\lambda} = y$, und aus der dritten Gleichung erhalten wir $x = y = 1$. Die einzige extremwertverdächtige Stelle ist also $(x, y) = (1, 1)$. Nun muß man sich noch klarmachen, dass an dieser Stelle tatsächlich ein (sogar globales) Minimum unter der Nebenbedingung $g(x, y) = 0$ vorliegt (z.B. durch die Überlegung, dass die Funktion f auf der Hyperbel $\{(x, y) \in (0, \infty) \times (0, \infty) : xy = 1\}$ für $x \rightarrow \infty$ oder $y \rightarrow \infty$ beliebig große Werte annimmt. ■

Beispiel 17 Gesucht sind Maximum und Minimum der Funktion $f(x, y) = xy$ unter der Nebenbedingung $g(x, y) = x^2 + y^2 - 1 = 0$, d.h. auf der Einheitskreislinie mit dem Mittelpunkt $(0, 0)$. Der Gradient

$$(\text{grad}g)(x, y) = (2x, 2y)$$

ist offenbar in jedem Punkt der Einheitskreislinie ungleich Null, so dass Satz 17.1 anwendbar ist. Wir bilden die Lagrange-Funktion

$$L(x, y, \lambda) = f(x, y) + \lambda g(x, y) = xy + \lambda(x^2 + y^2 - 1)$$

und setzen ihre partiellen Ableitungen Null:

$$\begin{aligned} \frac{\partial L}{\partial x} = 0 &\Rightarrow y + 2\lambda x = 0 \\ \frac{\partial L}{\partial y} = 0 &\Rightarrow x + 2\lambda y = 0 \\ \frac{\partial L}{\partial \lambda} = 0 &\Rightarrow x^2 + y^2 = 1. \end{aligned}$$

Wir multiplizieren die erste erhaltene Gleichung mit x , die zweite mit y und finden, dass $2\lambda x^2 = 2\lambda y^2$ sein muss. Da $\lambda \neq 0$ sein muss, ist $x^2 = y^2$, weshalb zusammen mit der dritten Gleichung für (x, y) die Paare

$$\left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right), \left(-\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right), \left(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right) \quad \text{und} \quad \left(-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)$$

in Frage kommen. Diese lösen tatsächlich das Gleichungssystem (die beiden ersten Paare für $\lambda = -1/2$, die beiden anderen mit $\lambda = 1/2$). Als Funktionswert ergibt sich für die ersten beiden Paare $+1/2$, für die anderen jeweils $-1/2$. Da f als stetige Funktion auf einer kompakten Menge Maximum und Minimum annehmen muss, ist $1/2$ das Maximum

und $-1/2$ das Minimum von f . ■ Das war eh klar: die Extremwerte werden bei den Hauptachsen angenommen und sind maximaler bzw. minimaler Eigenwert.

Beispiel

$$f(x, y, z) = \frac{1}{2}(x^4 + y^4 + z^4), \quad g(x, y, z) = x^2 + y^2 + z^2 - 1$$

$$\text{grad}f(x, y, z) = 2 \begin{pmatrix} x^3 \\ y^3 \\ z^3 \end{pmatrix}, \quad \text{grad}g(x, y, z) = 2 \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

Das Gleichungssystem

$$\begin{pmatrix} x^3 \\ y^3 \\ z^3 \end{pmatrix} + \lambda \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

führt für $x \neq 0$ zu $\lambda = -x^2$, also $y = 0$ oder $y^2 = x^2$, d.h. $y = \pm x$ und ebenso $z = 0$ oder $z = \pm x$. Die weiteren kritischen Punkte ergeben sich analog.

1. $y = z = 0$, Dann $x = \pm 1$ und $f(x, 0, 0) = \frac{1}{2}$.
2. $y = \pm x, z = 0$. Dann $|x| = |y| = \frac{1}{\sqrt{2}}$, $f(x, y, 0) = \frac{1}{4}$
3. $y^2 = z^2 = x^2$. Dann $|x| = |y| = |z| = \frac{1}{\sqrt{3}}$, $f(x, y, z) = \frac{1}{6}$

Da die Einheitskugel kompakt ist, werden Maximum und Minimum angenommen, und wir können schliessen, dass wir in (1) und den analogen Fällen $x = y = 0$ bzw. $x = z = 0$ alle globalen Maxima bestimmt haben. Ebenso mit (3) für Minima. Dass in (2) Sattelpunkte vorliegen, ist mühsamer zu zeigen.

Beispiel 18 Gesucht wird ein Dreieck, das bei gegebenem Umfang u maximalen Flächeninhalt besitzt. Sind $x, y, z > 0$ die Seitenlängen des Dreiecks, so ist $u = x + y + z$, und der Flächeninhalt ist

$$F = \sqrt{\frac{u}{2} \left(\frac{u}{2} - x\right) \left(\frac{u}{2} - y\right) \left(\frac{u}{2} - z\right)}$$

(Heronsche Formel). Da F genau dann maximal wird, wenn F^2 maximal wird, haben wir die Funktion $f: \mathbb{R}^3 \rightarrow \mathbb{R}$,

$$f(x, y, z) = \frac{u}{2} \left(\frac{u}{2} - x\right) \left(\frac{u}{2} - y\right) \left(\frac{u}{2} - z\right),$$

unter der Nebenbedingung

$$g(x, y, z) = x + y + z - u = 0$$

zu maximieren. Die Lagrange-Funktion L lautet

$$L(x, y, z, \lambda) = \frac{u}{2} \left(\frac{u}{2} - x\right) \left(\frac{u}{2} - y\right) \left(\frac{u}{2} - z\right) + \lambda(x + y + z - u) = 0,$$

und Nullsetzen der partiellen Ableitungen ergibt

$$\begin{aligned} L_x(x, y, z, \lambda) &= -\frac{u}{2}\left(\frac{u}{2} - y\right)\left(\frac{u}{2} - z\right) + \lambda = 0 \\ L_y(x, y, z, \lambda) &= -\frac{u}{2}\left(\frac{u}{2} - x\right)\left(\frac{u}{2} - z\right) + \lambda = 0 \\ L_z(x, y, z, \lambda) &= -\frac{u}{2}\left(\frac{u}{2} - x\right)\left(\frac{u}{2} - y\right) + \lambda = 0 \\ L_\lambda(x, y, z, \lambda) &= x + y + z - u = 0. \end{aligned}$$

Subtraktion der 2. von der ersten Gleichung liefert

$$-\frac{u}{2}(x - y)\left(\frac{u}{2} - z\right) = 0.$$

Also ist $\frac{u}{2} - z = 0$ oder $x = y$. Für $\frac{u}{2} - z = 0$ ist $f(x, y, z) = 0$, also mit Sicherheit kein Maximum. Also bleibt nur $x = y$. Aus der zweiten und dritten Gleichung findet man analog $y = z$, und zusammen mit der vierten Gleichung ergibt sich $x_0 = y_0 = z_0 = u/3$ als Kandidat für das Maximum. Diese Lösung entspricht einem gleichseitigen Dreieck, und $f(x_0, y_0, z_0) = f(u/3, u/3, u/3) = u^4/432$. ■

Beispiel 19 Wir suchen die extremwertverdächtigen Punkte der Funktion

$$f : \mathbb{R}^n \rightarrow \mathbb{R}, \quad x \mapsto \langle Ax, x \rangle,$$

wobei A eine symmetrische $n \times n$ -Matrix ist, unter der Nebenbedingung $\|x\| = 1$, die wir als $g(x) = \|x\|^2 - 1 = 0$ schreiben.

Der Gradient von f ist

$$(\text{grad}f)(x) = 2x^T A,$$

und der von g ist

$$(\text{grad}g)(x) = 2x^T$$

(mit der Definition der Ableitung nachrechnen!). Das zu lösende Gleichungssystem ist also

$$\begin{aligned} (\text{grad}f)(x) - \mu(\text{grad}g)(x) &= 2x^T A - \mu 2x^T \stackrel{!}{=} 0, \\ g(x) &= \|x\|^2 - 1 \stackrel{!}{=} 0 \end{aligned}$$

mit dem Lagrange-Parameter $\mu = -\lambda$. Transponieren der ersten Gleichung liefert das System

$$Ax = \mu x \quad \text{mit} \quad \|x\|^2 = 1.$$

Ein Vektor der Länge 1 ist also genau dann extremwertverdächtig, wenn er ein Eigenvektor zum Eigenwert μ ist. Der zugehörige Funktionswert ist

$$f(x) = \langle Ax, x \rangle = \langle \mu x, x \rangle = \mu \langle x, x \rangle = \mu.$$

Folglich wird f in x unter der Nebenbedingung $\|x\| = 1$ genau dann maximal (minimal), wenn x ein Eigenvektor von A zum größten (kleinsten) Eigenwert von A ist. Man beachte, dass die stetige Funktion f auf der kompakten Menge $\{x \in \mathbb{R}^n : \|x\| = 1\}$ ihr Maximum und Minimum annehmen muß. Es ist also

$$\begin{aligned} \max_{x \neq 0} \frac{\langle Ax, x \rangle}{\langle x, x \rangle} &= \text{größter Eigenwert von } A, \\ \min_{x \neq 0} \frac{\langle Ax, x \rangle}{\langle x, x \rangle} &= \text{kleinster Eigenwert von } A. \end{aligned}$$

Das war eh klar: die Extremwerte werden bei den Hauptachsen angenommen und sind maximaler bzw. maximaler Eigenwert

Man kann die Methode der Lagrange-Multiplikatoren auch benutzen, wenn mehrere Nebenbedingungen zu erfüllen sind. Die Lagrange-Funktion lautet dann

$$L(x, \lambda_1, \dots, \lambda_r) = f(x) + \lambda_1 g_1(x) + \dots + \lambda_r g_r(x)$$

für die Nebenbedingungen $g_1(x) = 0, \dots, g_r(x) = 0$. ■

17.5 Untermannigfaltigkeiten und Extrema unter Nebenbedingungen

Siehe Skript Prof. Hieber (Analysis-II-ss08), VIII.3.

18 Wege

18.1 Wege im euklidischen Raum

Wir haben schon vektorwertige reelle Funktionen

$$\vec{x}(t) \in V, \quad (t \in D \subseteq \mathbb{R})$$

behandelt, wobei V ein endlichdimensionaler euklidischer Vektorraum ist, also bzgl. einer Orthomormalbasis mit \mathbb{R}^n identifiziert werden kann

$$\vec{x}(t) = \begin{pmatrix} x_1(t) \\ \vdots \\ x_n(t) \end{pmatrix} \quad (t \in D).$$

Solche Funktionen nennen wir auch *Kurven*. Insbesondere wurden Stetigkeit und Differenzierbarkeit behandelt. Als Ableitung haben wir hier den *Tangentialvektor* zur Zeit t

$$\frac{\partial \vec{x}}{\partial t}(t) = \dot{\vec{x}}(t) = \begin{pmatrix} \frac{\partial x_1}{\partial t}(t) \\ \vdots \\ \frac{\partial x_n}{\partial t}(t) \end{pmatrix}$$

Im Folgenden wird meist $X(t) = \vec{x}(t)$ geschrieben und Zeilen statt Spalten, das Skalarprodukt dann als $X \cdot Y$.

Schon im \mathbb{R}^2 lassen sich Kurven nicht immer als Graph einer *reellen Funktion* $f : [a, b] \rightarrow \mathbb{R}$ darstellen. So müssen beispielsweise für die Einheitskreislinie $\{(x, y) \in \mathbb{R}^2 : x^2 + y^2 = 1\}$ der obere bzw. untere Kreisbogen jeweils durch $y = f_1(x) = +\sqrt{1 - x^2}$ bzw. $y = f_2(x) = -\sqrt{1 - x^2}$ gesondert beschrieben werden. Eine ganz andere Möglichkeit der Beschreibung von Kurven bieten Parameterdarstellungen wie $\{(x, y) \in \mathbb{R}^2 : x = \cos t, y = \sin t \text{ mit } t \in [0, 2\pi]\}$ für die Einheitskreislinie. Hier beschreiben wir die Kreislinie als Bild des Intervalls $[0, 2\pi]$ unter der Abbildung

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

Definition 18.1 Jede stetige Abbildung $X : [a, b] \rightarrow \mathbb{R}^n$ heißt ein Weg in \mathbb{R}^n . Dabei heißen $X(a)$ Anfangs- und $X(b)$ Endpunkt des Weges. Das Bild von $[a, b]$ unter X , also die Menge $\{X(t) : t \in [a, b]\}$, heißt die zu X gehörende Kurve in \mathbb{R}^n .

Man beachte: ein Weg ist eine *Abbildung*, die zugehörige Kurve eine *Punktmenge*. Man kann sich vorstellen, dass durch einen Weg eine *Parametrisierung* der zugehörigen Kurve gegeben ist. Mit den Begriffen „Anfangs- und Endpunkt“ verbindet man die Vorstellung, dass die Kurve von $X(a)$ nach $X(b)$ im Sinne wachsender Parameter $t \in [a, b]$ durchlaufen wird. Man beachte auch, dass zu verschiedenen Wegen gleiche Kurven gehören können. So liefern die Wege

$$\begin{aligned} X_1 & : [0, 2\pi] \rightarrow \mathbb{R}^2, & t & \mapsto (\cos t, \sin t), \\ X_2 & : [0, 4\pi] \rightarrow \mathbb{R}^2, & t & \mapsto (\cos t, \sin t), \\ X_3 & : [0, 2\pi] \rightarrow \mathbb{R}^2, & t & \mapsto (\cos 3t, \sin 3t), \end{aligned}$$

jeweils die gleiche Kurve, nämlich die Einheitskreislinie im \mathbb{R}^2 .

Beispiel 1 Sei $f : [a, b] \rightarrow \mathbb{R}$ eine stetige Funktion. Dann ist

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

ein Weg, und die zugehörige Kurve ist der Graph von f . ■

Beispiel 2 Seien $A, B \in \mathbb{R}^3$. Dann ist

$$X : [0, 1] \rightarrow \mathbb{R}^3, \quad t \mapsto tA + (1 - t)B$$

ein Weg, und die zugehörige Kurve ist die Verbindungsstrecke von A nach B . ■

Beispiel 3 Der für $r > 0$, $h > 0$ und $k \in \mathbb{N}$ durch

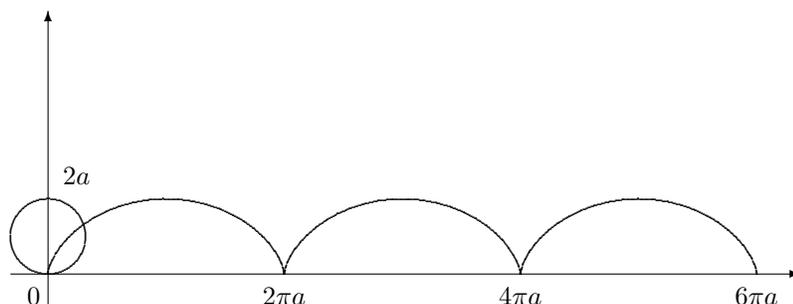
$$X : [0, 2k\pi] \rightarrow \mathbb{R}^3, \quad t \mapsto (r \cos t, r \sin t, \frac{h}{2\pi}t)$$

definierte Weg liefert als Kurve eine Schraubenlinie vom Radius r mit der Ganghöhe h und mit k Windungen. ■

Beispiel 4 Auf einer Kreisscheibe vom Radius a ist auf dem Rand ein Punkt P markiert. Die Kreisscheibe wird nun auf der Ebene abgerollt. Dabei durchläuft P den Weg

$$X : \mathbb{R} \rightarrow \mathbb{R}^2, \quad t \mapsto a(t - \sin t, 1 - \cos t),$$

und die zugehörige Kurve heißt *Rollkurve* oder *Zykloide*.





Der Begriff des Weges ist aber weitaus komplexer, als diese einfachen Beispiele vermuten lassen, und er läßt Beispiele zu, die unserer Anschauung völlig widersprechen. So gibt es einen Weg $X : [0, 1] \rightarrow \mathbb{R}^2$, dessen zugehörige Kurve das gesamte Einheitsquadrat $[0, 1] \times [0, 1]$ ausfüllt (eine sogenannte Peano-Kurve)!

18.2 Stückweise stetig differenzierbare Wege

Unserer Anschauung am nächsten kommen die stetig differenzierbaren und die stückweise stetig differenzierbaren Wege. Ein Weg $X = (X_1, \dots, X_n) : [a, b] \rightarrow \mathbb{R}^n$ ist *stetig differenzierbar*, wenn jede seiner Komponenten X_j auf $[a, b]$ stetig differenzierbar ist. Der Weg $X = (X_1, \dots, X_n) : [a, b] \rightarrow \mathbb{R}^n$ heißt *stückweise stetig differenzierbar*, wenn es eine Zerlegung

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

des Intervalles $[a, b]$ gibt, so dass jede Komponente X_j auf jedem Teilintervall $[t_{i-1}, t_i]$ für sich stetig differenzierbar ist.

Ist $X = (X_1, \dots, X_n) : [a, b] \rightarrow \mathbb{R}^n$ stetig differenzierbar und ist $\dot{X}(t_0) = (\dot{X}_1(t_0), \dots, \dot{X}_n(t_0)) \neq 0$ für ein $t_0 \in (a, b)$, so heißt $\dot{X}(t_0)$ ein *Tangentialvektor* von X im Punkt $X(t_0)$, und die Tangente an die zugehörige Kurve im Punkt $X(t_0)$ wird beschrieben durch (Fig.1.)

$$T(\lambda) = X(t_0) + \lambda \dot{X}(t_0).$$

Mitunter ist es nützlich, dass man Wege addieren kann. Sind $X : [a, b] \rightarrow \mathbb{R}^n$ und $Y : [b, c] \rightarrow \mathbb{R}^n$ zwei Wege mit $X(b) = Y(b)$, so definiert man ihre *Summe* $Z = X \oplus Y$ durch

$$Z : [a, c] \rightarrow \mathbb{R}^n, \quad Z(t) = \begin{cases} X(t) & \text{falls } t \in [a, b] \\ Y(t) & \text{falls } t \in [b, c]. \end{cases}$$

Entsprechend definiert man die Summe endlich vieler Wege. Der zu einem Weg $X : [a, b] \rightarrow \mathbb{R}^n$ *entgegengesetzte Weg* ist durch

$$X^- : [a, b] \rightarrow \mathbb{R}^n, \quad X^-(t) = X(a + b - t)$$

gegeben. Er liefert die gleiche Kurve wie der Ausgangsweg, die jedoch in entgegengesetzter Richtung durchlaufen wird. Jeder stückweise stetig differenzierbare Weg ist eine Summe endlich vieler stetig differenzierbarer Wege. Eine Summe von Wegen wie in Beispiel 2 („Strecken“) heißt auch ein **Polygonzug**.

18.3 Weglänge

Zur Definition der *Länge* eines Weges $X : [a, b] \rightarrow \mathbb{R}^n$ wählen wir eine Zerlegung $Z = \{t_0, t_1, \dots, t_m\}$ mit $a = t_0 < t_1 < \dots < t_m = b$ des Intervalles $[a, b]$ und betrachten den Polygonzug durch die Punkte $X(t_0), X(t_1), \dots, X(t_m)$. Er hat die Länge

$$L(X, Z) := \sum_{i=1}^m \|X(t_i) - X(t_{i-1})\|,$$

und für jede Verfeinerung Z' von Z gilt wegen der Dreiecksungleichung

$$L(X, Z) \leq L(X, Z').$$

(Fig.2)

Definition 18.2 Der Weg $X : [a, b] \rightarrow \mathbb{R}^n$ heißt rektifizierbar, wenn es eine Zahl M gibt mit der Eigenschaft, dass $L(X, Z) \leq M$ für alle Zerlegungen Z von $[a, b]$. Die kleinste Zahl M mit dieser Eigenschaft heißt die Länge $L(X)$ des Weges X .

Es ist also

$$L(X) = \sup_Z L(X, Z),$$

wobei das Supremum über alle Zerlegungen Z von $[a, b]$ zu nehmen ist.

Satz 18.3 Jeder stetig differenzierbare Weg $X : [a, b] \rightarrow \mathbb{R}^n$ ist rektifizierbar, und es gilt

$$L(X) = \int_a^b \|\dot{X}(t)\| dt = \int_a^b \sqrt{\dot{X}_1(t)^2 + \dots + \dot{X}_n(t)^2} dt.$$

Der folgende Satz führt die Weglängenberechnung für stückweise stetig differenzierbare Wege auf die für stetig differenzierbare Wege zurück.

Satz 18.4 Ist der Weg Z die Summe $X^{(1)} \oplus \dots \oplus X^{(m)}$ der stetig differenzierbaren Wege $X^{(j)}$, so ist Z rektifizierbar, und

$$L(Z) = L(X^{(1)}) + \dots + L(X^{(m)}).$$

Es genügt, dass die $X^{(i)}$ rektifizierbar sind. Der Beweis folgt sofort mit der Dreiecksungleichung.

18.4 Parametrisierung durch Weglänge

Ist $\vec{x}(t)$ ($t \in [a, b]$) stetig differenzierbar und gilt $\|\frac{\partial \vec{x}}{\partial t}(t)\| = 1$ für alle $t \in [a, b]$ so haben wir konstante Bahngeschwindigkeit vom Betrag 1 und können die Zeit t als Maß der Weglänge s des Weges zwischen $\vec{x}(a)$ und $\vec{x}(t)$ nehmen. Man sagt auch, dass die Kurve *durch die Weglänge parametrisiert* wird. Wir beweisen obigen Satz für diesen Fall

Lemma 18.5 Ist $\vec{x} : [a, b] \rightarrow \mathbb{R}^n$ stetig differenzierbar mit $\|\frac{\partial \vec{x}}{\partial t}\| = 1$ für alle $t \in [a, b]$, so ist $b - a$ das Supremum der zur Kurve gehörigen Polygonzüge.

Beweis. Nach der Dreiecksungleichung nimmt die Länge der Polygonzüge bei Verfeinerung der Zerlegungen von $[a, b]$ zu. Nach dem Mittelwertsatz der Differentialrechnung, angewendet auf die Koordinatenfunktionen $x_i(t)$ ergibt gilt

$$\sum_{k=1}^m \|\vec{x}(t_k) - \vec{x}(t_{k-1})\| = \sum_{k=1}^m \sqrt{\sum_{i=1}^n [x_i(t_k) - x_i(t_{k-1})]^2} = \sum_{k=1}^m \sqrt{\sum_{i=1}^n \left[\frac{\partial x_i}{\partial t}(\tau_{ik})(t_k - t_{k-1}) \right]^2}$$

mit passenden $\tau_{ik} \in [t_{k-1}, t_k]$. Also genügt es zu zeigen, dass es zu jedem $\varepsilon > 0$ ein $\delta > 0$ gibt so, dass

$$(*) \quad \left| 1 - \sum_{i=1}^n \left[\frac{\partial x_i}{\partial t}(\tau_{ik}) \right]^2 \right| \leq \varepsilon^2 \quad \text{falls } t_k - t_{k-1} \leq \delta.$$

Dann folgt nämlich

$$|b - a - L| = \left| \sum_{k=1}^m (t_k - t_{k-1}) - L \right| \leq \varepsilon \quad \text{mit } L = \sum_{k=1}^m \|\vec{x}(t_k) - \vec{x}(t_{k-1})\|$$

Die Behauptung (*) ergibt sich aber sofort aus dem folgenden Hilfssatz. \square

Lemma 18.6 *Ist $\vec{x}(s)$ auf $[a, b]$ stetig, so gibt es zu jedem $\varepsilon > 0$ ein $\delta > 0$ so, dass für alle $s_0, \dots, s_n \in [a, b]$*

$$\left\| \begin{pmatrix} x_1(s_1) \\ \vdots \\ x_n(s_n) \end{pmatrix} - \vec{x}(s_0) \right\| \leq \varepsilon \quad \text{falls } |s_i - s_j| \leq \delta \text{ für alle } i, j$$

Beweis durch Induktion über n : Da $x_1(s)$ auf $[a, b]$ stetig und somit gleichmäßig stetig ist, gibt es δ_1 so, dass $|x_1(s_1) - x_1(s_0)| \leq \frac{\varepsilon}{2}$ falls $|s_1 - s_0| \leq \delta_1$. Nach Induktion gibt es δ_2 so, dass

$$\left\| \begin{pmatrix} x_2(s_1) \\ \vdots \\ x_n(s_n) \end{pmatrix} - \vec{x}(s_0) \right\| \leq \frac{\varepsilon}{2} \quad \text{falls } |s_i - s_j| \leq \delta_2 \text{ für alle } i, j \neq 1.$$

Wähle $\delta = \min\{\delta_1, \delta_2\}$. Die Behauptung folgt mit der Dreiecksungleichung \square

Korollar 18.7 *Ist $\vec{x}(t)$ ($t \in [a, b]$) stetig differenzierbar ist und $\|\frac{\partial \vec{x}}{\partial t}(t)\| > 0$ für alle $t \in [a, b]$, so ist*

$$\int_a^b \left\| \frac{\partial \vec{x}}{\partial t} \right\| dt$$

die Weglänge, d.h. das Supremum der Längen der Polygonzüge.

Beweis. Sei

$$s = s(t) = \int_a^t \left\| \frac{\partial \vec{x}}{\partial \tau}(\tau) \right\| d\tau$$

Nach dem Hauptsatz ist s differenzierbar mit

$$\frac{\partial s}{\partial t} = \left\| \frac{\partial \vec{x}}{\partial t}(t) \right\| > 0$$

also umkehrbar, d.h. $t = t(s)$. Mit der Kettenregel und der Umkehrregel folgt

$$\frac{\partial \vec{x}}{\partial s} = \frac{\partial \vec{x}}{\partial t} \frac{\partial t}{\partial s} = \frac{\partial \vec{x}}{\partial t} \frac{1}{\frac{\partial s}{\partial t}} = \frac{\partial \vec{x}}{\partial t} \frac{1}{\left\| \frac{\partial \vec{x}}{\partial t} \right\|}$$

$$\left\| \frac{\partial \vec{x}}{\partial s} \right\| = 1$$

Daher haben wir Weglänge $s(b) - s(a) = s(b)$. \square

18.5 Beispiele von Weglängen

Beispiel 5 Der durch die stetige Funktion

$$f : [0, 1] \rightarrow \mathbb{R}, \quad t \mapsto \begin{cases} 0 & \text{für } t = 0 \\ t \cos \frac{\pi}{t} & \text{für } t \in (0, 1] \end{cases}$$

festgelegte Weg $X : [0, 1] \rightarrow \mathbb{R}^2, t \mapsto (t, f(t))$ ist *nicht* rektifizierbar. Für die Punkte $t_n = 1/n$ mit $n \geq 1$ gilt nämlich

$$\begin{aligned} \|X(t_{n+1}) - X(t_n)\| &\geq |f(t_{n+1}) - f(t_n)| = \left| \frac{\cos(n+1)\pi}{n+1} - \frac{\cos n\pi}{n} \right| \\ &= \left| \frac{(-1)^{n+1}}{n+1} - \frac{(-1)^n}{n} \right| = \frac{1}{n+1} + \frac{1}{n} > \frac{1}{n}, \end{aligned}$$

und die Reihe $\sum_{n=1}^{\infty} \frac{1}{n}$ divergiert bekanntlich. Man kann also nicht jedem Weg eine Länge zuschreiben. ■

Beispiel 6 Die Länge des Weges

$$X : [0, 2\pi] \rightarrow \mathbb{R}^2, \quad t \mapsto (r \cos t, r \sin t) \quad \text{mit } r > 0$$

ergibt sich aus

$$L(X) = \int_0^{2\pi} \sqrt{\dot{X}_1(t)^2 + \dot{X}_2(t)^2} dt = \int_0^{2\pi} \sqrt{r^2 \sin^2 t + r^2 \cos^2 t} dt = 2\pi r. \quad (18.7)$$

Das Ergebnis ist natürlich (?) gleich dem Kreisumfang eines Kreises vom Radius r . Doch haben wir in (18.7) tatsächlich die Länge des Kreisumfangs, d.h. einer Kurve, berechnet? Der Weg

$$Y : [0, 4\pi] \rightarrow \mathbb{R}^2, \quad t \mapsto (r \cos t, r \sin t)$$

führt auf die gleiche Kurve wie der Weg X , hat jedoch die doppelte Länge wie X , was natürlich daran liegt, dass Y die Kreislinie zweimal durchläuft. Will man also *Kurvenlängen* berechnen, muß man einen rektifizierbaren Weg wählen, der jeden Punkt der Kurve nur einmal durchläuft. Nur Anfangs- und Endpunkt dürfen gegebenenfalls (wie bei der Kreislinie und anderen geschlossenen Kurven) zusammenfallen. ■

Beispiel 7 Ist $f : [a, b] \rightarrow \mathbb{R}$ stetig differenzierbar und $X : [a, b] \rightarrow \mathbb{R}^2, t \mapsto (t, f(t))$, so ist

$$L(X) = \int_a^b \sqrt{1 + f'(t)^2} dt. \quad (18.8)$$

18.6 Linienelemente

Für stetige differenzierbare Wege mit $\|\frac{\partial \vec{x}}{\partial s}\| = 1$ definieren wir das *Linienelement* an der Stelle p als die lineare Abbildung $d\vec{x}(p) : \mathbb{R}^n \rightarrow \mathbb{R}$ mit

$$d\vec{x}(p)(ds) = ds \frac{\partial \vec{x}}{\partial s}(p)$$

Das Linienelement approximiert die Kurve in der Nähe von p durch eine Strecke

$$\vec{x}(p + \Delta s) \approx \vec{x}(p) + d\vec{x}(p)(\Delta s) = \Delta s \frac{\partial \vec{x}}{\partial s}(p)$$

Ist t differenzierbare Funktion von s so gilt

$$d\vec{x}(p)(dt) = ds \frac{\partial \vec{x}}{\partial t} \frac{\partial t}{\partial s} = dt \frac{\partial \vec{x}}{\partial t}$$

19 Wegintegrale

Bisher haben wir ausschließlich Integrale über Intervallen betrachtet. Ziel dieses Abschnittes ist es, Integrale über Kurven zu erklären und zu berechnen. Besonders interessiert uns dabei die Frage, ob ein solches Integral nur von Anfangs- und Endpunkt der Kurve abhängt.

19.1 Definition

Wenn es existiert, so nennen wir

$$\int_a^b f(\vec{x}(t)) dt$$

das *Integral der skalaren Funktion f längs des Weges $\vec{x}(t), t \in [a, b]$* . Ist $f(\vec{x}(t)) \geq 0$ für alle t , können wir es als Maß für Fläche $\{(\vec{x}(t), y) \mid 0 \leq y \leq f(\vec{x}(t)), t \in [a, b]\}$ verstehen.

19.2 Summationstheorem für Wegintegrale

Satz 19.1 Sei $\vec{x} : [a, b] \rightarrow \mathbb{R}^n$ stetig differenzierbar mit $\|\frac{\partial \vec{x}}{\partial t}(t)\| = 1$ für alle $t \in [a, b]$. Sei $D \subseteq \mathbb{R}^n$ und $f : D \rightarrow \mathbb{R}$ so, dass $f(\vec{x}(t))$ für alle $t \in [a, b]$ definiert und stetig ist. Dann gibt es zu jedem $\varepsilon > 0$ ein $\delta > 0$ so, dass für alle $a \leq c < d \leq b$ mit $d - c < \delta$ und alle $\tau \in [c, d]$ gilt

$$\left| \|\vec{x}(d) - \vec{x}(c)\| f(\vec{x}(\tau)) - \int_c^d f(\vec{x}(t)) dt \right| \leq \varepsilon(d - c)$$

Anschaulich für $f(\vec{x}(t)) \geq 0$: Sind c und d genügend nahe beieinander, so ist die Fläche $\{(\vec{x}(t), y) \mid 0 \leq y \leq f(\vec{x}(t)), t \in [c, d]\}$ näherungsweise gleich der Fläche des Rechtecks $\{(\lambda \vec{x}(c) + (1 - \lambda)\vec{x}(d), h) \mid \lambda \in [0, 1]\}$, dessen Grundseite die Strecke von $\vec{x}(c)$ nach $\vec{x}(d)$ ist und dessen Höhe einer der Werte $f(\vec{x}(\tau))$ mit $\tau \in [c, d]$ ist.

Beweis. Nach dem Satz über die Integrierbarkeit stetiger Funktionen gibt es $\delta_1 > 0$ mit

$$\left| \int_c^d f(\vec{x}(t)) dt - f(\vec{x}(\tau))(d - c) \right| \leq \frac{\varepsilon}{2}(d - c) \text{ für } d - c \leq \delta_1.$$

Wegen der Stetigkeit auf $[a, b]$ existiert

$$M = \max\{|f(\vec{x}(t))| \mid t \in [a, b]\}$$

Wegen der stetigen Differenzierbarkeit mit Betrag 1 gibt es δ_2 mit

$$\left| \frac{\|\vec{x}(d) - \vec{x}(c)\|}{d - c} - 1 \right| \leq \frac{\varepsilon}{2M} \quad \text{für } d - c \leq \delta_2$$

Also

$$\|\vec{x}(d) - \vec{x}(c)\| f(\vec{x}(\tau) - (d - c)f(\vec{x}(\tau))) \leq \frac{\varepsilon}{2}(d - c)$$

und die Behauptung folgt mit $\delta = \min\{\delta_1, \delta_2\}$ und der Dreiecksungleichung. \square

Korollar 19.2 *Unter den Voraussetzungen des Satzes sei $W(c, d) \in \mathbb{R}$ für alle $a \leq c < d \leq b$ definiert und additiv. Weiterhin gelte: Für alle $\varepsilon > 0$ gibt es $\delta > 0$ so, dass für alle Δt mit $|\Delta t| \leq \delta$ und alle $p \in [a, b]$ gilt: Es gibt $\tau \in [p, p + \Delta t]$ mit $W(p, p + \Delta t) - f(\vec{x}(\tau))\Delta t \leq \varepsilon|\Delta t|$. Dann gilt für alle $c \leq d$ in $[a, b]$: $W(c, d) = \int_c^d f(\vec{x}(t)) dt$*

Beweis. Summationstheorem und obiger Satz. \square

Den Fall, dass $\vec{x}(t)$ stetig differenzierbar ist mit $\|\frac{\partial \vec{x}}{\partial t}\| > 0$ können wir wie in 18.1.7 durch Substitution auf diesen Fall zurückführen. Wir haben dann das Wegintegral

$$\int_a^b f(\vec{x}(t)) \left\| \frac{\partial \vec{x}}{\partial t} \right\| dt = \int_{s(a)}^{s(b)} f(\vec{x}(t(s))) ds$$

mit der Weglänge $s = s(t)$ von a nach t und $ds = \frac{\partial s}{\partial t} dt$.

19.3 Wegintegral im Vektorfeld

Der wichtigste Fall: \vec{F} ist ein Vektorfeld, so, dass $\vec{F}(\vec{x}(t))$ für alle $t \in [a, b]$ definiert und stetig ist. Wir setzen

$$f(\vec{x}(t)) = \langle \vec{F}(\vec{x}(t)) | \frac{\partial \vec{x}}{\partial t}(t) \rangle \frac{1}{\|\frac{\partial \vec{x}}{\partial t}\|}$$

Definition 19.3 *Sei $D \subseteq \mathbb{R}^n$, $F : D \rightarrow \mathbb{R}^n$, $\vec{x} : [a, b] \rightarrow D \subseteq \mathbb{R}^n$ ein stetig differenzierbarer Weg und $F(\vec{x}(t))$ auf $[a, b]$ stetig. Dann heißt*

$$\int_{\Gamma} \vec{F} \cdot d\vec{x} = \int_a^b \langle \vec{F}(\vec{x}(t)) | \frac{\partial \vec{x}}{\partial t}(f) \rangle dt = \int_a^b \vec{F}(X(t)) \cdot \dot{X}(t) dt \quad (19.9)$$

das Wegintegral von F entlang des Weges $\vec{x}(t) = X(t)$ ($t \in [a, b]$).

Statt (19.9) findet man oft auch die Schreibweise

$$\int_X F \cdot dX \quad \text{oder} \quad \int_X F_1 dX_1 + \dots + F_n dX_n,$$

wobei F_i bzw. X_i die Komponenten von F bzw. X sind.

Unter der Voraussetzung $\|\frac{\partial \vec{x}}{\partial t}\| > 0$ ist das Wegintegral (auch *Linienintegral*) von der Kurve aber nicht von der Parametrisierung abhängig: Sind $s = s(t)$ und $t = t(s)$ stetig differenzierbar mit nicht verschwindenden Ableitungen so gilt

$$\int_{s(a)}^{s(b)} \langle \vec{F}(\vec{x}(t(s))) | \frac{\partial \vec{x}}{\partial s}(s) \rangle ds = \int_a^b \langle \vec{F}(\vec{x}(t)) | \frac{\partial \vec{x}}{\partial t}(t) \rangle \frac{\partial t}{\partial s} ds = \int_a^b \langle \vec{F}(\vec{x}(t)) | \frac{\partial \vec{x}}{\partial t}(t) \rangle dt$$

Insbesondere können wir stets zum Fall konstanter Bahngeschwindigkeit 1 übergehen. Ist nun \vec{F} ein Kraftfeld und beschreibt $\vec{x}(t)$ die Bewegung einer Masse längs eines Weges, so dürfen wir nach dem Korollar $W(c, d)$ als die auf dem Weg von $\vec{x}(c)$ nach $\vec{x}(d)$ geleistete Arbeit ansehen, weil dies nach allgemeiner Überzeugung für hinreichend kleine Intervalle $[c, d]$ durch die Formel “gerade Wegstrecke mal Komponente in dieser Richtung des als konstant angenommenen Kraftvektors” hinreichend genau angenähert wird.

19.4 Beispiele von Wegintegralen

Beispiel 8 Der Weg X sei gegeben durch

$$X : [0, 2\pi] \rightarrow \mathbb{R}^3, \quad t \mapsto (a \cos t, a \sin t, \frac{h}{2\pi}t),$$

und P bewege sich entlang X von $(a, 0, 0)$ nach $(a, 0, h)$. Dabei wirke eine Kraft $F(\vec{x}) = -\alpha\vec{x}$ mit einer Konstanten $\alpha > 0$, d.h. P wird mit der Kraft $\|F(\vec{x})\| = \alpha\sqrt{x_1^2 + x_2^2 + x_3^2}$ in Richtung des Koordinatenursprungs gezogen. Für die zu leistende Arbeit W finden wir

$$\begin{aligned} W &= \int_0^{2\pi} F(X(t)) \cdot \dot{X}(t) dt \\ &= \int_0^{2\pi} -\alpha(a \cos t, a \sin t, \frac{h}{2\pi}t) \cdot (-a \sin t, a \cos t, \frac{h}{2\pi}) dt \\ &= \alpha \int_0^{2\pi} (a^2 \sin t \cos t - a^2 \sin t \cos t - (\frac{h}{2\pi})^2 t) dt \\ &= -\alpha(\frac{h}{2\pi})^2 \int_0^{2\pi} t dt = -\alpha(\frac{h}{2\pi})^2 \frac{(2\pi)^2}{2} = -\frac{\alpha h^2}{2}. \end{aligned}$$

Wir betrachten bei gleicher Kraft noch einmal die Bewegung von P von $(a, 0, 0)$ nach $(a, 0, h)$, nun aber entlang des Weges

$$X : [0, h] \rightarrow \mathbb{R}^3, \quad t \mapsto (a, 0, t).$$

Dann ist

$$W = \int_0^h -\alpha(a, 0, t) \cdot (0, 0, 1) dt = -\alpha \int_0^h t dt = -\frac{\alpha h^2}{2},$$

d.h. wir gelangen zum gleichen Resultat. ■

Beispiel 9 Sei $F : \mathbb{R}^2 \rightarrow \mathbb{R}^2, (x, y) \mapsto (x^2 + y^2, xy)$ und $X : [0, 1] \rightarrow \mathbb{R}^2, t \mapsto (t, t)$ (die zugehörige Kurve ist die Strecke von $(0, 0)$ nach $(1, 1)$). Dann ist

$$\int_0^1 F(X(t)) \cdot \dot{X}(t) dt = \int_0^1 (2t^2, t^2) \cdot (1, 1) dt = \int_0^1 3t^2 dt = 1.$$

Bei gleichem F finden wir dagegen für das Wegintegral entlang des Weges

$$X : [0, 1] \rightarrow \mathbb{R}^2, \quad t \mapsto (t, t^2)$$

(der entlang einer Parabel von $(0, 0)$ nach $(1, 1)$ führt)

$$\int_X F \cdot dX = \int_0^1 (t^2 + t^4, t^3) \cdot (1, 2t) dt = \int_0^1 (t^2 + 3t^4) dt = \frac{14}{15},$$

also ein vom Weg abhängiges Ergebnis. ■

19.5 Regeln für Wegintegrale

Hier sind einige Regeln für den Umgang mit Wegintegralen.

1. Ist X^- der zu X entgegengesetzte Weg, so gilt

$$\int_{X^-} F \cdot dX = - \int_X F \cdot dX. \quad (19.10)$$

2. Es gilt die Abschätzung

$$\left| \int_X F \cdot dX \right| \leq \max_{t \in [a,b]} \|F(X(t))\| L(X). \quad (19.11)$$

3. Ist $Z = X \oplus Y$ die Summe der Wege X und Y , so ist

$$\int_{X \oplus Y} F \cdot dX = \int_X F \cdot dX + \int_Y F \cdot dY. \quad (19.12)$$

Man benutzt (19.12) auch, um das Wegintegral entlang eines nur *stückweise* stetig differenzierbaren Weges zu definieren.

Wir greifen nun die Beobachtung aus Beispiel 8 und 9 auf und fragen, wann ein Wegintegral *wegunabhängig* ist, d.h. nur vom Anfangs- und Endpunkt des Weges abhängt (vgl. Beispiel 8).

20 Potentiale

20.1 Definition und Anwendung

Definition 20.1 Sei $D \subseteq \mathbb{R}^n$ offen und $F : D \rightarrow \mathbb{R}^n$ ein Vektorfeld. Existiert eine stetig partiell differenzierbare Funktion $\varphi : D \rightarrow \mathbb{R}$ mit

$$F(x) = \text{grad}\varphi(x) \quad \text{für } x \in D,$$

so heißt φ ein Potential (oder eine Stammfunktion) von F .

Korollar 20.2 Besitzt F auf D ein Potential, so ist $F(x(t))$ stetig differenzierbar für jeden Weg $x(t)$ in D .

Beweis. Nach Voraussetzung sind die $F_i(x) = \frac{\partial \varphi}{\partial x_i}(x)$ stetig differenzierbar, also nach der Kettenregel auch $F_i(x(t))$. \square

In Beispiel 8 (Kap.18) ist $\varphi(x_1, x_2, x_3) := -\frac{\alpha}{2}(x_1^2 + x_2^2 + x_3^2)$ ein Potential von $F(x_1, x_2, x_3) = -\alpha(x_1, x_2, x_3)$. Physikalisch interessanter ist folgendes Beispiel.

Beispiel 10 Eine Masse m im Koordinatenursprung übt nach dem Newtonschen Gravitationsgesetz auf einen Massepunkt der Masse 1 in $(x, y, z) \in \mathbb{R}^3$ eine Kraft F der Stärke

$$\|F(x, y, z)\| = \frac{Gm}{\|(x, y, z)\|^2} = \frac{Gm}{x^2 + y^2 + z^2}$$

aus. Diese Kraft weist zum Nullpunkt, hat also die Richtung $-\frac{(x,y,z)}{\|(x,y,z)\|}$. Demnach ist

$$F(x, y, z) = -\frac{Gm}{\|(x, y, z)\|^3}(x, y, z) = -\frac{Gm}{(x^2 + y^2 + z^2)^{3/2}}(x, y, z).$$

Man rechnet leicht nach, dass die Funktion $\varphi : \mathbb{R}^3 \setminus \{0\} \rightarrow \mathbb{R}$,

$$\varphi(x, y, z) = \frac{Gm}{\sqrt{x^2 + y^2 + z^2}}$$

ein Potential für F auf $\mathbb{R}^3 \setminus \{0\}$ ist (Newtonsches Gravitationspotential). ■

Satz 20.3 Sei $D \subseteq \mathbb{R}^n$ ein Gebiet und $F : D \rightarrow \mathbb{R}^n$ ein Vektorfeld mit einem Potential $\varphi : D \rightarrow \mathbb{R}$. Sind A und E zwei beliebige Punkte in D und ist $X : [a, b] \rightarrow D$ ein beliebiger stückweise stetig differenzierbarer Weg in D mit Anfangspunkt A und Endpunkt E , so gilt für das Wegintegral

$$\int_X F \cdot dX = \varphi(E) - \varphi(A). \quad (20.13)$$

Unter den Voraussetzungen des Satzes ist das Wegintegral also wegunabhängig. Man vergleiche (20.13) mit dem Hauptsatz der Differential- und Integralrechnung. Insbesondere ist (wieder unter den Voraussetzungen von Satz 20.3) jedes Wegintegral entlang eines geschlossenen Weges (d.h. eines Weges mit $A = E$) gleich Null.

Beweis. Wir schreiben $X(t) = \vec{x}(t)$ und $z = f(\vec{x})$. Dann mit der Kettenregel und dem Hauptsatz

$$\int_a^b \langle F(\vec{x}(t)) | \frac{\partial \vec{x}}{\partial t} \rangle dt = \int_a^b \sum_{i=1}^n \frac{\partial z}{\partial x_i}(\vec{x}(t)) \frac{\partial x_i}{\partial t} dt = \int_a^b \frac{\partial z(\vec{x}(t))}{\partial t}(t) dt = z(\vec{x}(b)) - z(\vec{x}(a)) \quad \square$$

20.2 Sternförmige Mengen

Wir greifen nun die Beobachtung aus Beispiel 8 und 9 auf und fragen, wann ein Wegintegral *wegunabhängig* ist, d.h. nur vom Anfangs- und Endpunkt des Weges abhängt (vgl. Beispiel 8). Hierfür benötigen wir, dass die Wege in Mengen mit speziellen Eigenschaften verlaufen, die wir nun definieren.

Definition 20.4 1. Eine Menge $M \subseteq \mathbb{R}^n$ heißt wegzusammenhängend, wenn sich je zwei Punkte aus M durch einen Weg verbinden lassen, der komplett durch M verläuft.

2. M heißt konvex, wenn sich je zwei Punkte aus M durch eine Strecke verbinden lassen, die ganz in M liegt.

3. M heißt sternförmig, wenn es einen Punkt $z \in M$ gibt, so dass für jeden Punkt $x \in M$ die Verbindungsstrecke von z nach x ganz in M liegt.

4. M heißt ein Gebiet, wenn M offen und wegzusammenhängend ist.

Konvexe Mengen sind sternförmig, und sternförmige Mengen sind wegzusammenhängend. (Fig.3)

20.3 Integrierbare Vektorfelder

Mer folgende Satz hilft bei der Entscheidung, ob ein gegebenes Vektorfeld ein Potential besitzt.

Satz 20.5 Sei $M \subseteq \mathbb{R}^n$ offen und sternförmig, und $F : M \rightarrow \mathbb{R}^n$ ein Vektorfeld. Sind die Komponenten des Vektorfeldes $F = (F_1, \dots, F_n)$ stetig partiell differenzierbar und gilt

$$\frac{\partial F_i}{\partial x_j} = \frac{\partial F_j}{\partial x_i} \quad \text{für } 1 \leq i < j \leq n \quad \text{Exaktheit} \quad (20.14)$$

auf M , so besitzt F ein Potential auf M .

Die Bedingungen des Satzes 20.5 garantieren also die Existenz eines Potentials und damit die Wegunabhängigkeit des Wegintegrals. Es gilt auch die Umkehrung: besitzt das auf der offenen Menge M stetig partiell differenzierbare Vektorfeld ein Potential, so gilt (20.14) (Satz von Schwarz). Man beachte, dass wir in Satz 20.5 die Sternförmigkeit von M fordern müssen. Ein Potential kann aber auch existieren, wenn diese Eigenschaft nicht vorliegt (Beispiel 10: $\mathbb{R}^n \setminus \{0\}$ ist nicht sternförmig!). Für den Beweis benötigt man die ‘Parameterintegrale’, die im nächsten Kapitel behandelt werden.

Wir sehen uns die Bedingungen (20.14) für $n = 2$ und $n = 3$ genauer an und zeigen, wie man bei Erfülltsein dieser Bedingung ein Potential finden kann.

20.4 Vektorfelder in der Ebene

Fall $n = 2$ Sei $F = (F_1, F_2) : M \rightarrow \mathbb{R}^2$ auf $M =]a, b[\times]c, d[$. Wähle $(x_0, y_0) \in M$. Dann verlaufen zu jedem $(x, y) \in M$ die Wege

$$\Gamma_1 : \vec{x}(t) = (x_0, t) \quad (t \in [y_0, y]), \quad \Gamma_2 : \vec{x}(t) = (t, y) \quad (t \in [x_0, x])$$

in M und wir haben Tangentialvektor $(0, 1)$ bzw. $(1, 0)$. Gibt es ein Potential ϕ , so folgt

$$\phi(x, y) - \phi(x_0, y_0) = \int_{\Gamma_1} F \cdot d\vec{x} + \int_{\Gamma_2} F \cdot d\vec{x} = \int_{y_0}^y F_2(x_0, t) dt + \int_{x_0}^x F_1(t, y) dt$$

und man kann es so bestimmen. Beachte, dass das Potential nur bis auf eine Konstante bestimmt ist. Beispiel

$$F_1(x, y) = x + y, \quad F_2(x, y) = x - 3y$$

Mit $x_0 = y_0 = 0$

$$\phi(x, y) = \int_0^y -3t dt + \int_0^x t + y dt = -\frac{3}{2}y^2 + \frac{1}{2}x^2 + xy$$

Hier hängt der erste Summand nur von y ab. Das gibt den Anlass zu dem Ansatz

$$\phi(x, y) = h(x, y) + \psi(y), \quad \frac{\partial h}{\partial x} = F_1, \quad \frac{\partial \psi}{\partial y} = F_2 - \frac{\partial h}{\partial y}$$

weil $F_1 = \frac{\partial\phi}{\partial x} = \frac{\partial h}{\partial x}$ und $F_2 = \frac{\partial\phi}{\partial y} = \frac{\partial h}{\partial y} + \frac{\partial\psi}{\partial y}$. Somit

$$h(x, y) = \int F_1(x, y)dx, \quad \psi(x, y) = \int \left(\frac{\partial h}{\partial x}(x, y) - F_2(x, y) \right) dy$$

im Beispiel

$$h(x, y) = \int (x + y)dx = \frac{1}{2}x^2 + xy, \quad \frac{\partial h}{\partial y} = x, \quad \psi(y) = \int (x - 3y - x)dy = -\frac{3}{2}y^2$$

Sei nun M eine sternförmige Menge. Hier reduziert sich (20.14) auf die eine Bedingung

$$\frac{\partial F_1}{\partial x_2} = \frac{\partial F_2}{\partial x_1} \quad \text{auf } M. \quad (20.15)$$

Ist dies erfüllt, so versuchen wir, $\varphi : M \rightarrow \mathbb{R}$ aus dem Ansatz

$$\frac{\partial\varphi}{\partial x_1} = F_1, \quad \frac{\partial\varphi}{\partial x_2} = F_2 \quad \text{auf } M$$

zu bestimmen. Integration der ersten Gleichung bzgl. x_1 liefert

$$\varphi(x_1, x_2) = \int F_1(x_1, x_2)dx_1 + g(x_2) \quad (20.16)$$

mit einer nicht von x_1 , aber von x_2 abhängigen „Integrationskonstanten“ g . Wir leiten (20.16) formal nach x_2 ab und erhalten mit (20.15)

$$\frac{\partial\varphi}{\partial x_2} = \frac{\partial}{\partial x_2} \int F_1(x_1, x_2)dx_1 + g'(x_2) \stackrel{!}{=} F_2(x_1, x_2).$$

Hieraus folgt

$$g'(x_2) = F_2(x_1, x_2) - \frac{\partial}{\partial x_2} \int F_1(x_1, x_2)dx_1,$$

und durch unbestimmte Integration bzgl. x_2 gewinnt man g und damit φ . Eine Probe zeigt, ob tatsächlich eine Stammfunktion gefunden wurde.

Beispiel 11 Für $x \neq 0$ und $y \neq \frac{\pi}{2} + k\pi$, $k \in \mathbb{Z}$ sei

$$F_1(x, y) = -\frac{\tan y}{x^2} + 2xy + x^2, \quad F_2(x, y) = \frac{1}{x \cos^2 y} + x^2 + y^2.$$

Dann ist (20.15) erfüllt, und wir wählen den Ansatz

$$\frac{\partial\varphi}{\partial x} = -\frac{\tan y}{x^2} + 2xy + x^2, \quad \frac{\partial\varphi}{\partial y} = \frac{1}{x \cos^2 y} + x^2 + y^2.$$

Integration der ersten Beziehung nach x liefert

$$\varphi(x, y) = \int \left(-\frac{\tan y}{x^2} + 2xy + x^2 \right) dx = \frac{\tan y}{x} + x^2 y + \frac{x^3}{3} + g(y).$$

Wir leiten dies nach y ab und setzen das Ergebnis gleich F_2 :

$$\frac{\partial \varphi}{\partial y} = \frac{1}{x \cos^2 y} + x^2 + g'(y) \stackrel{!}{=} \frac{1}{x \cos^2 y} + x^2 + y^2.$$

Also ist $g'(y) = y^2$ und damit $g(y) = \frac{y^3}{3} + C$ sowie

$$\varphi(x, y) = \frac{\tan y}{x} + x^2 y + \frac{x^3}{3} + \frac{y^3}{3} + C.$$

Sucht man nur *ein* Potential, kann man z.B. $C = 0$ wählen. Die Probe durch Ableiten zeigt, dass tatsächlich ein Potential von F gefunden wurde. ■

20.5 Vektorfelder im Raum

Fall $n = 3$ Dann ergibt (20.14) die drei Bedingungen

$$\frac{\partial F_1}{\partial x_2} = \frac{\partial F_2}{\partial x_1}, \quad \frac{\partial F_1}{\partial x_3} = \frac{\partial F_3}{\partial x_1}, \quad \frac{\partial F_2}{\partial x_3} = \frac{\partial F_3}{\partial x_2}. \quad (20.17)$$

Sind diese erfüllt, machen wir den Ansatz

$$\frac{\partial \varphi}{\partial x_1} = F_1, \quad \frac{\partial \varphi}{\partial x_2} = F_2, \quad \frac{\partial \varphi}{\partial x_3} = F_3$$

für das gesuchte Potential φ . Wir integrieren die erste Gleichung nach x_1

$$\varphi(x_1, x_2, x_3) = \int F_1(x_1, x_2, x_3) dx_1 + g(x_2, x_3) \quad (20.18)$$

und leiten nach x_2 ab:

$$\frac{\partial \varphi}{\partial x_2} = \frac{\partial}{\partial x_2} \int F_1(x_1, x_2, x_3) dx_1 + \frac{\partial g}{\partial x_2} \stackrel{!}{=} F_2(x_1, x_2, x_3).$$

Es ist also

$$\frac{\partial g}{\partial x_2} = F_2(x_1, x_2, x_3) - \frac{\partial}{\partial x_2} \int F_1(x_1, x_2, x_3) dx_1.$$

Die rechte Seite hängt wegen (20.17) nicht von x_1 ab; wir bezeichnen sie mit $h(x_2, x_3)$ und erhalten durch Integration nach x_2

$$g(x_2, x_3) = \int h(x_2, x_3) dx_2 + \ell(x_3).$$

Einsetzen in (20.18) liefert

$$\varphi(x_1, x_2, x_3) = \int F_1(x_1, x_2, x_3) dx_1 + \int h(x_2, x_3) dx_2 + \ell(x_3).$$

Nun leiten wir noch nach x_3 ab und setzen das Ergebnis gleich F_3 :

$$\frac{\partial \varphi}{\partial x_3} = \frac{\partial}{\partial x_3} \int F_1(x_1, x_2, x_3) dx_1 + \frac{\partial}{\partial x_3} \int h(x_2, x_3) dx_2 + \ell'(x_3) \stackrel{!}{=} F_3(x_1, x_2, x_3),$$

also

$$\ell'(x_3) = F_3(x_1, x_2, x_3) - \frac{\partial}{\partial x_3} \int F_1(x_1, x_2, x_3) dx_1 - \frac{\partial}{\partial x_3} \int h(x_2, x_3) dx_2.$$

Die rechte Seite dieser Gleichung hängt nur von x_3 ab, und Integration dieser Gleichung bzgl. x_3 liefert ℓ und damit φ .

Beispiel 12 Sei $F(x, y, z) = (x + z, -y - z, x - y)$. Das Vektorfeld $F : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ erfüllt die Bedingungen (20.14) auf der sternförmigen Menge \mathbb{R}^3 , so dass mit Sicherheit ein Potential existiert. Sei $\varphi : \mathbb{R}^3 \rightarrow \mathbb{R}$ dieses Potential. Aus $\varphi_x = F_1$ folgt durch Integration bzgl. x

$$\varphi(x, y, z) = \int (x + z) dx + g(y, z) = \frac{x^2}{2} + xz + g(y, z).$$

Ableiten nach y ergibt

$$\varphi_y = \frac{\partial g}{\partial y}(y, z) \stackrel{!}{=} -y - z,$$

woraus folgt

$$g(y, z) = -\frac{y^2}{2} - yz + \ell(z).$$

Also ist

$$\varphi(x, y, z) = \frac{x^2}{2} + xz - \frac{y^2}{2} - yz + \ell(z),$$

und nach Ableiten nach z wird

$$\varphi_z = x - y + \ell'(z) \stackrel{!}{=} x - y,$$

d.h. $\ell'(z) = 0$. Es ist also ℓ eine konstante Funktion, d.h.

$$\varphi(x, y, z) = \frac{x^2}{2} + xz - \frac{y^2}{2} - yz + c \quad \text{mit } c \in \mathbb{R}.$$

■

20.6 Rotation und Divergenz

Siehe Skript Prof. Hieber (Analysis-II-ss08), IX.2

21 Parameterabhängige Integrale

21.1 Vertauschung von Integral und partieller Ableitung

Satz 21.1 Sei $M \subseteq \mathbb{R}^n$ offen und $f : M \times [a, b]$ stetig und stetig nach x_i differenzierbar. Dann gilt für jedes $p \in M$

$$\left(\frac{\partial}{\partial x_i} \int_a^b f(x, t) dt \right)(p) = \int_a^b \frac{\partial f}{\partial x_i}(p, t) dt$$

Das Integral $\int_a^b f(x, t) dt$ hängt hier von *Parameter* x ab. Beweis. Wir benutzen die Schreibweise

$$x = y \pm \varepsilon \Leftrightarrow y - \varepsilon \leq x \leq y + \varepsilon$$

O.B.d.A. $n = 1$. Sei $p \in M$ gegeben. Es gibt eine kompakte konvexe Umgebung $M' \subseteq M$ von p (z.B. $U_\varepsilon(p) \subseteq M$ für ein $\varepsilon > 0$, da M offen, und $M' = \{x \mid |x - p| \leq \varepsilon/2\}$). Sei nun $\varepsilon > 0$ und $c \in [a, b]$ gegeben. Wegen der gleichmäßigen Stetigkeit von $\frac{\partial f}{\partial x}(x, t)$ auf $M' \times [a, b]$ gibt es ein $\delta > 0$ so, dass

$$\frac{\partial f}{\partial x}(x, t) = \frac{\partial f}{\partial x}(p, t) \pm \varepsilon \quad \text{für alle } x \in M' \text{ mit } x = p \pm \delta \text{ und } t \in [a, b]$$

Für $|\Delta x| < \delta$ folgt nun mit dem Mittelwertsatz angewandt auf $\frac{\partial f}{\partial x}$ und mit $\xi_{\Delta x} \in M'$

$$\begin{aligned} \frac{1}{\Delta x} \left(\int_a^b f(p + \Delta x, t) dt - \int_a^b f(p, t) dt \right) &= \frac{1}{\Delta x} \int_a^b (f(p + \Delta x, t) - f(p, t)) dt \\ &= \frac{1}{\Delta x} \int_a^b \Delta x \left(\frac{\partial f}{\partial x}(\xi_{\Delta x}, t) \right) dt = \frac{1}{\Delta x} \int_a^b \Delta x \left(\frac{\partial f}{\partial x}(p, t) \pm \varepsilon \right) dt \\ &= \int_a^b \left(\frac{\partial f}{\partial x}(p, t) \pm \varepsilon \right) dt = \int_a^b \frac{\partial f}{\partial x}(p, t) dt \pm \varepsilon(b - a) \end{aligned}$$

21.2 Beispiele

Beispiel 12 Sei $f(x, y) = e^{xy}$ für $(x, y) \in [0, 1] \times [-1, 1]$. Da f stetig ist, ist auch die Funktion

$$g(y) = \int_0^1 e^{xy} dx = \begin{cases} 1 & \text{für } y = 0 \\ \frac{e^y - 1}{y} & \text{für } y \in [-1, 1] \setminus \{0\} \end{cases}$$

stetig. Da außerdem $\frac{\partial f}{\partial y} = x e^{xy}$ auf \mathbb{R}^2 stetig ist, ist g differenzierbar auf $[-1, 1]$ mit

$$\frac{dg}{dy} = \int_0^1 x e^{xy} dx = \begin{cases} 1/2 & \text{für } y = 0 \\ \frac{(y-1)e^y + 1}{y^2} & \text{für } y \in [-1, 1] \setminus \{0\}. \end{cases}$$

■

Beispiel 13 Für $|t| < 1$ berechnen wir das Integral

$$F(t) := \int_0^\pi \ln(1 - 2t \cos x + t^2) dx. \quad (21.19)$$

Um Satz 21.1 benutzen zu können, wählen wir ein $a \in (0, 1)$ mit $t \in [-a, a]$. Die Funktion

$$f(x, t) = \ln(1 - 2t \cos x + t^2)$$

ist nach t stetig partiell differenzierbar mit

$$\frac{\partial f}{\partial t}(x, t) = \frac{2t - 2 \cos x}{1 - 2t \cos x + t^2} \quad \text{auf } [0, \pi] \times [-a, a].$$

Aus Satz 21.1 (c) folgt

$$F'(t) = \int_0^\pi \frac{2t - 2 \cos x}{1 - 2t \cos x + t^2} dx.$$

Zur Berechnung dieses Integrals substituieren wir

$$s := \tan \frac{x}{2} \quad \text{bzw.} \quad x = 2 \arctan s.$$

Dann ist $\frac{dx}{ds} = \frac{2}{1+s^2}$ bzw. (formal) $dx = \frac{2ds}{1+s^2}$ sowie

$$\cos x = \frac{\cos^2 \frac{x}{2} - \sin^2 \frac{x}{2}}{\cos^2 \frac{x}{2} + \sin^2 \frac{x}{2}} = \frac{1 - \tan^2 \frac{x}{2}}{1 + \tan^2 \frac{x}{2}} = \frac{1 - s^2}{1 + s^2}.$$

Das gesuchte Integral geht über in

$$\int_0^\infty \frac{2t - 2\frac{1-s^2}{1+s^2}}{1 - 2t\frac{1-s^2}{1+s^2} + t^2} \cdot \frac{2}{1+s^2} ds = 4 \int_0^\infty \frac{s^2(1+t) - (1-t)}{(1+t)^2 s^4 + 2(1+t^2)s^2 + (1-t)^2} ds.$$

Partialbruchzerlegung liefert

$$4 \frac{s^2(1+t) - (1-t)}{(1+t)^2 s^4 + 2(1+t^2)s^2 + (1-t)^2} = \frac{2}{t} \left(\frac{1}{s^2 + 1} + \frac{t^2 - 1}{(1+t)^2 s^2 + (1-t)^2} \right),$$

und mit dem Grundintegral

$$\int \frac{1}{ax^2 + c} dx = \frac{1}{\sqrt{ac}} \arctan \sqrt{\frac{a}{c}} x \quad \text{für} \quad ac > 0$$

erhalten wir

$$F'(t) = \frac{2}{t} \left(\arctan s - \arctan \left(\frac{1+t}{1-t} s \right) \right) \Big|_0^\infty = 0$$

(man beachte, dass $\arctan 0 = 0$ und $\lim_{s \rightarrow \infty} \arctan s = \frac{\pi}{2}$). Also ist F eine *konstante Funktion*, und wir bestimmen ihren (einzigen) Wert, indem wir in (21.19) $t = 0$ setzen:

$$F(t) = F(0) = \int_0^\pi \ln 1 dx = 0. \quad \blacksquare$$

21.3 Integrierbare Vektorfelder

Zum Beweis von Satz 20.5. O.B.d.A. sieht das sternförmige Gebiet M so aus, dass für jedes p in M die Verbindungsstrecke $\Gamma_p = \{tp \mid t \in [0, 1]\}$ zu M gehört. Das gesuchte Potential definieren als Wegintegral längs dieser Wege

$$f(p) = \int_{\Gamma_p} F \cdot d\vec{x} = \int_0^1 \langle F(pt) \mid \frac{\partial tp}{\partial t} \rangle dt = \int_0^1 \left(\sum_{i=1}^n F_i(tp) p_i \right) dt$$

Zu zeigen ist, dass f stetig partiell differenzierbar ist mit $\frac{\partial f}{\partial x_k} = F_k$. Sei $p \in M$. Für jedes k betrachten wir die k -partielle Ableitung an der Stelle x . Wegen der stetigen partiellen Differenzierbarkeit der $F_i(tx)$ nach x_k gilt für das Parameterintegral nach Satz 21.1

$$\frac{\partial f}{\partial x_k}(x) = \int_0^1 \frac{\partial}{\partial x_k} \left(\sum_{i=1}^n F_i(tx) \cdot x_i \right) dt = \int_0^1 \sum_{i=1}^n \frac{\partial}{\partial x_k} (F_i(tx) \cdot x_i) dt$$

Produkt- und Kettenregel ergeben

$$\frac{\partial}{\partial x_k} F_i(tx) \cdot x_i = \frac{\partial}{\partial x_k} (F_i(tx) \cdot x_i + F_i(tx) \frac{\partial x_i}{\partial x_k}) = \frac{\partial}{\partial x_k} F_i(tx) \cdot tx_i + F_i(tx) \frac{\partial x_i}{\partial x_k}$$

und hier ist der zweite Summand $F_k(tx_k)$ falls $i = k$, andernfalls $= 0$. Also

$$\frac{\partial f}{\partial x_k}(x) = \int_0^1 t \left(\sum_{i=1}^n \frac{\partial}{\partial x_k} F_i(tx) \cdot x_i \right) + F_k(tx) dt$$

Nun gilt nach der Intergrabilitätsbedingung (19.7) und nach der Kettenregel

$$\sum_{i=1}^n \frac{\partial}{\partial x_k} F_i(tx) \cdot x_i = \sum_{i=1}^n \frac{\partial}{\partial x_i} F_k(tx) \frac{\partial}{\partial t} tx_i = \frac{\partial}{\partial t} F_k(tx)$$

also mit der Produktregel und dem Hauptsatz

$$\frac{\partial f}{\partial x_k}(x) = \int_0^1 t \frac{\partial}{\partial t} F_k(tx) + F_k(tx_k) dt = \int_0^1 \frac{\partial}{\partial t} t F_k(tx) dt = [t F_k(tx)]_0^1 = F_k(x) \quad \square$$

Ein durchsichtigerer Beweis ergibt sich wie folgt. Sei M zunächst ein Intervall, d.h. ein achsenparalleles Rechteck. Wähle $(a, b_0) \in M$ und definiere für $(a, b) \in M$

$$\phi(a, b) = \int_{a_0}^a \langle F(t, b_0) | \vec{e}_1 \rangle dt + \int_{b_0}^b \langle F(a, t) | \vec{e}_2 \rangle dt = \int_{a_0}^a F_1(t, b_0) dt + \int_{b_0}^b F_2(a, t) dt$$

Aus dem Hauptsatz folgt sofort

$$\frac{\partial}{\partial x} \int_{a_0}^a F_1(t, b_0) dt = F_1(a, b_0)$$

und mit Vertauschung und Exaktheit

$$\frac{\partial}{\partial x} \int_{b_0}^b F_2(a, t) dt = \int_{b_0}^b \frac{\partial F_2}{\partial x}(a, t) dt = \int_{b_0}^b \frac{\partial F_1}{\partial y}(a, t) dt = F_1(a, b) - F_1(a, b_0)$$

und somit

$$\frac{\partial \phi}{\partial x}(a, b) = F_1(a, b)$$

Der Hauptsatz ergibt sofort

$$\frac{\partial \phi}{\partial y}(a, b) = \frac{\partial}{\partial y} \int_{b_0}^b F_2(a, t) dt = F_2(a, b)$$

Sei nun M sternförmig mit Zentralpunkt (a_0, b_0) . Ist R ein Intervall mit Rand in M , so gehört ganz R zu M (jeder Punkt liegt auf der Verbindungsstecke von (a_0, b_0) mit einem Randpunkt von R), und da man auf R ein Potential hat, sind Intergrale längs Wegen in R wegunabhängig. Dasselbe gilt für jede Vereinigung endlich vieler Rechtecke.

Ein *Haken* sei ein Weg, der aus einer horizontalen und einer vertikalen Strecke besteht (die dürfen auch Punkte sein). Ein *Hakenweg* ist eine Summe von Haken. Zu je zwei Hakenwegen Γ_1, Γ von (a_0, b_0) nach (a, b) in M findet man eine endliche Folge $\Gamma_1, \Gamma_2, \dots, \Gamma_k = \Gamma$

von Hakenwegen so, dass sich die Kurve Γ_{i+1} von der Kurve Γ_i nur dadurch unterscheidet, dass die eine ein bestimmtes Rechteck $[(x, y), (u, v)]$ über (u, y) , die andere über (x, v) durchläuft. Es folgt

$$\int_{\Gamma_1} F \cdot d\vec{x} = \int_{\Gamma_2} F \cdot d\vec{x} = \dots = \int_{\Gamma} F \cdot d\vec{x}$$

Die Existenz eines Hakenwegs folgt aus der Kompaktheit der Strecke von (a_0, b_0) nach (a, b) : Diese ist enthalten in einer Vereinigung endlich vieler offener, in M enthaltener, Kreise mit Mittelpunkt auf der Strecke. Diese Kreise kann man so nummerieren, dass aufeinanderfolgende nichtleeren Schnitt, insbesondere eine gemeinsame Sehne senkrecht zur Strecke haben. Nimmt man als ε das Minimum dieser Sehnenlängen, so erhält man einen abgeschlossenen Streifen von Breite $\frac{2}{3}\varepsilon$ um die Strecke, der ganz in M liegt. Innerhalb dieses Streifens findet man den Hakenweg Γ . Man kann also unabhängig vom Hakenweg Γ definieren

$$\phi(a, b) = \int_{\Gamma} F \cdot d\vec{x}$$

Betrachten wir nun $\frac{\partial\phi}{\partial x}(a, b)$, so geht es um eine Verlängerung der Hakenwegs in horizontaler Richtung, also nach Hauptsatz

$$\begin{aligned} \frac{\partial\phi}{\partial x}(a, b) &= \lim_{\Delta x \rightarrow 0} \frac{1}{\Delta x} \int_a^{a+\Delta x} \langle F(a+t, b) | \vec{e}_1 \rangle dt \\ &= \lim_{\Delta x \rightarrow 0} \frac{1}{\Delta x} \int_a^{a+\Delta x} F_1(t, b) dt = \left(\frac{\partial}{\partial x} \int_a^x F_1(t, b) dt \right)(a) = F_1(a, b) \end{aligned}$$

Analog für $\frac{\partial\phi}{\partial y}$. Dasselbe Vorgehen führt zum Erfolg, wenn jedes von einem geschlossenen Hakenweg in M begrenztes Gebiet ganz zu M gehört, d.h. wenn M *einfach zusammenhängend* ist. \square

23 Integration auf Intervallen im \mathbb{R}^n

In diesem und dem folgenden Kapitel kommen wir zur Definition und den wesentlichen Eigenschaften des Riemann-Integrals einer Funktion von mehreren Veränderlichen. Als Motivation wird uns das Problem der Volumendefinition und -berechnung dienen (genau wie das Problem der Flächenberechnung bei Funktionen einer Veränderlichen).

23.1 Das Riemann-Integral über Intervallen im \mathbb{R}^n

23.1.1 Intervalle

Wir beginnen mit der Integration über den „einfachsten“ Teilmengen des \mathbb{R}^n , nämlich über Intervallen, Rechtecken, Quadern usw., die wir unter dem Namen „Intervall im \mathbb{R}^n “ zusammenfassen. Während wir im \mathbb{R}^1 ausschließlich über Intervalle integriert haben, gibt es beispielsweise im \mathbb{R}^3 wesentlich mehr Mengen, über die man integrieren möchte (Kugeln, Pyramiden, ...). Wir werden daher die in diesem Abschnitt angestellten Überlegungen später auf allgemeinere Mengen übertragen.

Seien $a = (a_1, \dots, a_n)$, $b = (b_1, \dots, b_n)$ zwei Punkte des \mathbb{R}^n mit $a_i \leq b_i$ für $i = 1, \dots, n$. Das *abgeschlossene Intervall* $[a, b]$ im \mathbb{R}^n ist die Menge

$$[a_1, b_1] \times \dots \times [a_n, b_n] = \{(x_1, \dots, x_n) \in \mathbb{R}^n : a_i \leq x_i \leq b_i \text{ für alle } i\},$$

und die Menge

$$]a, b[= (a, b) := (a_1, b_1) \times \dots \times (a_n, b_n) = \{(x_1, \dots, x_n) \in \mathbb{R}^n : a_i < x_i < b_i \text{ für alle } i\}$$

heißt ein *offenes Intervall* im \mathbb{R}^n und ist nicht leer genau dann, wenn $a_i < b_i$ für alle $i = 1, \dots, n$. Ist $I = [a, b]$ oder $I = (a, b)$ ein abgeschlossenes oder offenes Intervall, so erklären wir sein *Maß* $\mu_0(I)$ durch

$$\mu_0(I) := (b_1 - a_1) \cdot \dots \cdot (b_n - a_n).$$

und seine *Weite* durch

$$\max\{(b_1 - a_1), \dots, (b_n - a_n)\}$$

Gilt $b_i - a_i = c$ für $i = 1, \dots, n$ so ist das Intervall *regelmäßig* mit Kantenlänge c . Wir erhalten also für

- $n = 1$ gewöhnliche Intervalle, und $\mu_0(I)$ ist die Länge des Intervalls,
- $n = 2$ Rechtecke, und $\mu_0(I)$ ist der Flächeninhalt,
- $n = 3$ Quader, und $\mu_0(I)$ ist das Volumen.

Intervalle sind offensichtlich beschränkt und wegzusammenhängend.

23.1.2 Rechteckzerlegungen

Eine *Gitter-Zerlegung* Z eines abgeschlossenen Intervalls $I = [a, b] \subseteq \mathbb{R}^n$ ist ein Produkt $Z_1 \times \dots \times Z_n$ von Zerlegungen Z_i der Intervalle $[a_i, b_i]$. Die *Teilintervalle* der Zerlegung Z erhält man, indem man im Produkt $T_1 \times \dots \times T_n$ die T_i alle Teilintervalle der Zerlegung Z_i von $[a_i, b_i]$ durchlaufen läßt. Also: Ist $a_i = x_{i0} < z_{i1} < \dots < z_{im_i} = b_i$ die Zerlegung Z_i von $[a_i, b_i] \subseteq \mathbb{R}$, so hat man die Teilintervalle (Fig.4)

$$\{x \in \mathbb{R}^n \mid z_{ik_i} \leq x_i \leq z_{ik_i+1} \text{ für } i = 1, \dots, n\} \quad (0 \leq k_i < m_i)$$

Allgemeiner kann man eine Rechteckzerlegung Z eines Intervalls I definieren als eine endliche Menge abgeschlossener Intervalle, der *Teilintervalle der Zerlegung*, so, dass

- die zugehörigen offenen Teilintervalle nicht leer sind
- I die Vereinigung der Teilintervalle ist
- keine zwei verschiedenen Teilintervalle einen inneren Punkt gemeinsam haben

Die *Weite* von Z ist das Maximum der Weiten der Teilintervalle. Eine Zerlegung Z' von I ist *feiner* als Z , wenn jedes Teilintervall von Z' ganz in einem Teilintervall von Z enthalten ist.

Lemma 23.1 *Jede Zerlegung Z von I kann zu einer Gitterzerlegung Z' verfeinert werden. Je zwei Zerlegungen von I haben eine gemeinsame Verfeinerung.*

Beweis. Wähle die Rechteckzerlegung Z'_i von $[a_i, b_i]$ so, dass sie zu jedem $[u, v] \in Z$ die Punkte u_i, v_i enthält. \square

23.1.3 Treppenfunktionen

Eine *Treppenfunktion* auf einem Intervall I mit Rechteckzerlegung Z ist eine Funktion $f : I \rightarrow \mathbb{R}$, die auf jedem offenen Teilintervall von Z konstant ist. Das zugehörige *Integral* ist dann definiert als

$$\int_I f dx = \sum_{[u,v] \in Z} f(\xi_{uv}) \mu_0([u,v]) \in \mathbb{R} \quad \text{wobei } \xi_{uv} \in]u, v[$$

f ist auch bzgl. jeder Verfeinerung Z' von Z eine Treppenfunktion und das Integral bzgl. Z' ist dasselbe.

23.1.4 Unter- und Obersummen

Jede beschränkte Teilmenge X von \mathbb{R} hat 19.1.5 ein Supremum. Eine obere Schranke s von X ist genau dann Supremum, wenn es eine Folge $x_n \in X$ gibt mit $x_n \rightarrow s$. Ersetzt man \leq durch \geq , so erhält man Definition, Charakterisierung und Existenz des *Infimums* $\inf X$, der größten unteren Schranke.

Sei $I \subseteq \mathbb{R}^n$ ein abgeschlossenes Intervall, und Z sei eine Rechteckzerlegung von I in Teilintervalle I_1, \dots, I_m . Für jede beschränkte Funktion $f : I \rightarrow \mathbb{R}$ erklären wir

$$m_k := \inf\{f(x) : x \in I_k\} \quad \text{sowie} \quad M_k := \sup\{f(x) : x \in I_k\},$$

und wir nennen

$$U(Z, f) := \sum_{k=1}^m m_k \mu_0(I_k)$$

bzw.

$$O(Z, f) := \sum_{k=1}^m M_k \mu_0(I_k)$$

die *Unter-* bzw. *Obersumme* von f bzgl. der Rechteckzerlegung Z . Ist Z' feiner als Z so gilt offensichtlich

$$U(Z, f) \leq U(Z', f) \leq O(Z', f) \leq O(Z, f)$$

Schließlich sei noch

$$U(f, I) := \sup_Z U(Z, f)$$

und

$$O(f, I) := \inf_Z O(Z, f),$$

wobei das Supremum bzw. das Infimum über alle Rechteckzerlegungen Z von I zu bilden ist.

23.1.5 Riemann-Summe und Integral

Zu einer Rechteckzerlegung Z von I und betrachtet man die *Zwischenvektoren* $\xi = (\xi_{uv} \mid [u, v] \in Z)$ mit $\xi_{uv} \in]u, v[$ und die *Riemannsumme*

$$R(Z, \xi, f) = \sum_{[u,v] \in Z} f(\xi_{uv}) \mu_0([u, v]).$$

Satz 23.2 Sei $I = [a, b]$ ein Intervall in \mathbb{R}^n und $f : I \rightarrow \mathbb{R}$ eine beschränkte Funktion. Für $c \in \mathbb{R}$ sind äquivalent

- (i) Es gibt Treppenfunktionen $\underline{f}_k, \bar{f}_k : I \rightarrow \mathbb{R}$ mit $\underline{f}_k \leq f \leq \bar{f}_k$ für alle k und $\lim_{k \rightarrow \infty} \int_I \underline{f}_k(x) dx = c = \lim_{k \rightarrow \infty} \int_I \bar{f}_k(x) dx$
- (ii) Für alle $\varepsilon > 0$ gibt es $\delta > 0$ so, dass für alle Rechteckzerlegungen Z von I mit Weite $\leq \delta$ und alle Zwischenvektoren ξ gilt: $|c - R(Z, \xi, f)| \leq \varepsilon$
- (iii) Für alle $\varepsilon > 0$ gibt es eine Gitter-Zerlegung Z von I so, dass für alle Zwischenvektoren ξ gilt: $|c - R(Z, \xi, f)| \leq \varepsilon$
- (iv) $U(f, I) = O(f, I) = c$

Es gibt höchstens ein c , das (iii) erfüllt und es gilt $m(b-a) \leq c \leq M(b-a)$ falls $m \leq f(x) \leq M$ für alle $x \in I$.

Definition 23.3 Sei $I \subseteq \mathbb{R}^n$ ein abgeschlossenes Intervall und $f : I \rightarrow \mathbb{R}$ eine beschränkte Funktion. Gilt eine (also alle) der Aussagen (i)-(iv) für ein $c \in \mathbb{R}$, so heißt f Riemann-integrierbar auf I , und der eindeutig bestimmte Wert c heißt Riemann-Integral von f auf I . Als Bezeichnung wählen wir

$$c = \int_I f(x_1, \dots, x_n) d(x_1, \dots, x_n) \quad \text{oder kurz} \quad \int_I f(x) dx.$$

Besser wäre z.B. $\int_I f \prod_i dx_i$ oder einfach $\int_I f(\vec{x})$.

Korollar 23.4 Beschränktes $f : I \rightarrow \mathbb{R}$ ist Riemann-integrierbar genau dann, wenn es Treppenfunktionen $\underline{f}_k \leq f \leq \bar{f}_k$ gibt mit $\int_I \bar{f}_k(x) dx - \int_I \underline{f}_k(x) dx \rightarrow 0$ für $k \rightarrow \infty$.

Beweis. Die Äquivalenz von (i),(ii) und (iii) beweist man wie im Fall $n = 1$. Gilt (iv), so wähle zu $\varepsilon = \frac{1}{k}$ eine Rechteckzerlegung Z mit Untersumme $c - \varepsilon \leq U(Z, f) \leq c$ und \underline{f}_k als die zugehörige Treppenfunktion. Ebenso für Obersummen. Dann folgt (i). Gelte umgekehrt (i) und sei \underline{Z}^k Rechteckzerlegung zu \underline{f}_k . Dann gibt es zu jeder Rechteckzerlegung Z gemeinsame Verfeinerung Z' und es folgt $\int_I \underline{f}_k dx \leq U(Z', f) \leq U(f, I)$. Also $c = U(f, I)$. Entsprechend für Obersummen. Beim Beweis des Korollars kann man $\underline{f}_{k-1} \leq \underline{f}_k \leq \bar{f}_k \leq \bar{f}_{k-1}$ annehmen, indem man \underline{f}_k durch $\max\{\underline{f}_l \mid l \leq k\}$ ersetzt, analog für \bar{f}_k . Damit erhält man eine Intervallschachtelung und das gesuchte c . \square

23.1.6 Beispiel und Eigenschaften des Riemann-Integrals

In \mathbb{R}^2 bzw. \mathbb{R}^3 schreibt man auch

$$\int_I f(x_1, x_2) d(x_1, x_2) = \int_I f(x, y) d(x, y)$$

bzw.

$$\int_I f(x_1, x_2, x_3) d(x_1, x_2, x_3) = \int_I f(x, y, z) d(x, y, z).$$

Als einfaches Beispiel betrachten wir die konstante Funktion $f(x) = c$ auf $I \subseteq \mathbb{R}^n$. Dann ist

$$U(Z, f) = \sum_{k=1}^m c\mu_0(I_k) = c \sum_{k=1}^m \mu_0(I_k) = c\mu_0(I)$$

und analog $O(Z, f) = c\mu_0(I)$, und wir erhalten

$$O(f, I) = U(f, I) = \int_I cd(x_1, \dots, x_n) = c\mu_0(I).$$

Mit der Wahl $c = 1$ wird insbesondere

$$\mu_0(I) = \int_I d(x_1, \dots, x_n),$$

d.h. man kann das Maß von I als Integral über die Funktion $f(x) = 1$ darstellen. Wir kommen hierauf noch zurück.

Die folgenden einfachen Eigenschaften des Riemann-Integrals im \mathbb{R}^n beweist man wie im Fall $n = 1$.

Satz 23.5 *Sei $I \subseteq \mathbb{R}^n$ ein abgeschlossenes Intervall, und $f, g : I \rightarrow \mathbb{R}$ seien Riemann-integrierbar.*

- Für beliebige $\alpha, \beta \in \mathbb{R}$ ist $\alpha f + \beta g$ Riemann-integrierbar, und

$$\int_I (\alpha f(x) + \beta g(x)) dx = \alpha \int_I f(x) dx + \beta \int_I g(x) dx.$$

- Ist $f(x) \leq g(x)$ für alle $x \in I$, so ist

$$\int_I f(x) dx \leq \int_I g(x) dx.$$

- Es ist

$$\left| \int_I f(x) dx \right| \leq \sup_{x \in I} |f(x)| \mu_0(I).$$

23.1.7 Additivität

Lemma 23.6 *Ist $f : I \rightarrow \mathbb{R}$ beschränkt, $J \subseteq I$ und existiert $\int_I f(x) dx$ so existiert auch $\int_J f(x) dx$*

Beweis. Beweis. O.B.d.A. $f(x) \geq 0$ für alle $x \in I$. Nun benutze, dass jede Rechteckzerlegung von I eine Verfeinerung besitzt, die eine passende Verfeinerung von J umfasst. \square

Für Intervalle $[u, v]$, $[w, y]$ und $[c, d]$ in \mathbb{R}^n schreiben wir

$$[u, v] \oplus [w, y] = [c, d]$$

und sprechen von der *Summe*, falls $[u, v] \cup [w, y] = [c, d]$ und falls $[u, v]$ und $[w, y]$ keinen inneren Punkt gemeinsam haben. Eine Abbildung W , die jedem $[c, d] \subseteq I \subseteq \mathbb{R}^n$ ein $W([c, d]) \in \mathbb{R}$ zuordnet ist *additiv* auf I , falls

$$[u, v] \oplus [w, y] = [c, d] \Rightarrow W([u, v]) + W([w, y]) = W([c, d])$$

Lemma 23.7 Jede Riemann-integrierbare beschränkte Funktion $f : I \rightarrow \mathbb{R}$ ist auf I additiv.

Beweis: Ist $[u, v] \oplus [w, y] = [c, d]$, Z Rechteckzerlegung von $[u, v]$ und Z' Rechteckzerlegung von $[w, y]$, so ist $Z \cup Z'$ Rechteckzerlegung von $[c, d]$. \square

23.2 Integrale stetiger Funktionen

23.2.1 Zwischenwertsatz

Lemma 23.8 Sei $M \subseteq \mathbb{R}^n$ wegzusammenhängend und $f : M \rightarrow \mathbb{R}$ stetig. Dann ist $f(M) = \{f(x) \mid x \in M\}$ ein Intervall. Ist dabei M kompakt, so ist $f(M)$ ein kompaktes Intervall.

Beweis. Seien $p, q \in M$, o.B.d.A. $f(p) \leq f(q)$ und $z \in [f(p), f(q)]$. Nach Voraussetzung gibt es einen Weg $x(t)$ ($t \in [a, b]$) im M mit $x(a) = p$ und $x(b) = q$. Nun wende den Zwischenwertsatz an auf die stetige Funktion $g : [a, b] \rightarrow \mathbb{R}$ mit $g(t) = f(x(t))$. Ist M kompakt, so ist $f(M)$ nach Satz 4.6 kompakt. \square

23.2.2 Summationstheorem

Theorem 23.9 Sei $I = [a, b] \subseteq \mathbb{R}^n$ und $f : I \rightarrow \mathbb{R}$ stetig. Dann gilt

- (i) $f : I \rightarrow \mathbb{R}$ ist Riemann-integrierbar und zu jedem $\varepsilon > 0$ gibt es $\delta > 0$ so, dass für alle Intervalle $[c, d] \subseteq I$ von Weite $\leq \delta$ und alle $\xi \in [c, d]$ gilt

$$\left| \int_{[c,d]} f(x) dx - f(\xi) \mu_0([c, d]) \right| \leq \varepsilon \mu_0([c, d])$$

- (ii) Mittelwertsatz: Es gibt $\xi \in I$ mit $\int_I f(x) dx = f(\xi) \mu_0(I)$

- (iii) Sei $W[[c, d]]$ für alle $[c, d] \subseteq I$ definiert und additiv. Gelte weiterhin

- (*) Für alle $\varepsilon > 0$ gibt es ein $\delta > 0$ so, dass für alle (regelmäßigen) $[c, d] \subseteq I$ von Weite $\leq \delta$ gilt

$$\text{Es gibt } \xi \in [c, d] \text{ mit } |W[[c, d]] - f(\xi) \mu_0([c, d])| \leq \varepsilon \mu_0([c, d])$$

Dann gilt für alle $[c, d] \subseteq I$: $W[[c, d]] = \int_{[c,d]} f(x) dx$.

Der Beweis folgt wie für $n = 1$. In (i) benutzt man wieder die gleichmäßige Stetigkeit von f auf dem Kompaktum I und in (ii) den Zwischenwertsatz. In (iii) schreibt man $[c, d]$ als Summe von Teilintervallen einer Rechteckzerlegung Z und lässt die Weite von Z gegen 0 gehen. \square

Bei der ursprünglichen Auffassung von Integralen ist $\int_I f(x) dx$ die reelle Zahl, die von der Summe $\sum_I f(x) dx$ mit infinitesimalem Flächen- oder Volumenelement $dx = (dx)^n$ nur infinitesimalen Abstand hat. Die Voraussetzung (*) des Theorems besagt dann, dass $\alpha = W(p, p + d\vec{x}) - f(\xi) dx$ infinitesimal von zweiter Ordnung ist, d.h. dass $\frac{\alpha}{dx}$ infinitesimal ist. In dieser Form war das Theorem die Grundlage für die Einführung physikalischer Größen als Integrale. In den Natur- und Technikwissenschaften wird das heute so formuliert:

Von der Summe $W = \sum_k f(\xi_k)(\Delta x)_k^n$ mit den Flächen- bzw. Volumenelemente $(\Delta x)_k^n$ geht man mit $(\Delta x)_k \rightarrow 0$ zum Integral $W = \int f(x)dx$ über (unter Bezugnahme auf den Fall stückweise konstanter Funktionen).

Ist die Stetigkeit von $f(x)$ gegeben, so handelt es sich demnach um eine Anwendung des Theorems. In der gängigen mathematischen Literatur, wird bestenfalls bei der Definition einer Größe durch ein Integral eine Motivation im Sinne des Theorems gegeben. Die infinitesimale Form findet man in Lehrbüchern zur Nonstandard-Analysis.

23.2.3 Nachtrag: Riemannintegral in 1 Variablen

Riemannsches Integral

Für eine Zerlegung von I Z und Zwischenvektor $\vec{\xi} = (\xi_k)$ mit $\xi_k \in]z_k, z_{k+1}[$ definiert man die *Riemann-Summe*

$$R(Z, \vec{\xi}, f) = \sum_{k=0}^n f(\xi_k)(\Delta x)_k$$

c ist das *Riemann-Integral* von f wenn

$$c = \lim_{k \rightarrow \text{infy}} R(Z, \vec{\xi}_k, f) \text{ für } w(Z_k) \rightarrow 0$$

für jede Folge Z_k von Zerlegungen mit Maschenweite $w(Z_k) \rightarrow 0$ und jede Wahl der Zwischenpunkte \vec{z}_k . Jede sprungetige Funktion ist Riemann-integrierbar, aber nicht umgekehrt. Genaueres in supp.pdf

Satz. Sei $f : [a, b] \rightarrow \mathbb{R}$ eine beschränkte Funktion. Für $c \in \mathbb{R}$ sind äquivalent

- (i) Es gibt Treppenfunktionen $\underline{f}_n, \bar{f}_n : [a, b] \rightarrow \mathbb{R}$ mit $\underline{f}_n \leq f \leq \bar{f}_n$ für alle n (o.B.d.A. $\underline{f}_{n-1} \leq \underline{f}_n \leq \bar{f}_n \leq \bar{f}_{n-1}$) und $\lim_{n \rightarrow \infty} \int_a^b \underline{f}_n(x)dx = c = \lim_{n \rightarrow \infty} \int_a^b \bar{f}_n(x)dx$
- (ii) Für alle $\varepsilon > 0$ gibt es $\delta > 0$ so, dass für alle Zerlegungen Z von $[a, b]$ mit Maschenweite $\leq \delta$ und alle $\xi_k \in [z_k, z_{k+1}[$ gilt: $|c - \sum_{k=0}^b f(\xi_k)(\Delta x)_k| \leq \varepsilon$
- (iii) Für alle $\varepsilon > 0$ gibt es eine Zerlegung Z von $[a, b]$ so, dass für alle $\xi_k \in [z_k, z_{k+1}[$ gilt: $|c - \sum_{k=0}^n f(\xi_k)(\Delta x)_k| \leq \varepsilon$

Es gibt höchstens ein c , das (iii) erfüllt und es gilt $m(b-a) \leq c \leq M(b-a)$ falls $m \leq f(x) \leq M$ für alle $x \in [a, b]$.

Gilt (i), so heißt $f : [a, b] \rightarrow \mathbb{R}$ (Riemann-integrierbar und $c = \int_a^b f(x)dx$ das *Integral* von f auf dem Intervall $[a, b]$. (ii) und (iii) geben Anlass,

Beschränktes $f : [a, b] \rightarrow \mathbb{R}$ ist Riemann-integrierbar genau dann, wenn es Treppenfunktionen $\underline{f}_n \leq f \leq \bar{f}_n$ gibt mit $\int_a^b \bar{f}_n(x)dx - \int_a^b \underline{f}_n(x)dx \rightarrow 0$ für $n \rightarrow \infty$.

Beweis. (iii) folgt trivial aus (ii). Um (iii) \Rightarrow (i) zu zeigen, sei $\varepsilon = \frac{1}{n}$ und Z die zugehörige Zerlegung. sei $\eta = \varepsilon/(b-a)$. Da f auf $]z_k, z_{k+1}[$ nach oben und unten beschränkt ist, gibt es (dank Archimedes) $\bar{\xi}_k, \underline{\xi}_k \in]z_k, z_{k+1}[$ so, dass

$$f(\underline{\xi}_k) - \eta \leq f(x) \leq f(\bar{\xi}_k) + \eta \text{ für alle } x \in]z_k, z_{k+1}[.$$

Definiere Treppenfunktionen

$$\begin{aligned}\bar{f}_n(x) &= f(\bar{\xi}_k) + \eta & \text{für } z_k \leq x < z_{k+1} \\ \underline{f}_n(x) &= f(\underline{\xi}_k) - \eta & \text{für } z_k \leq x < z_{k+1}\end{aligned}$$

Dann folgt $\underline{f}_n \leq f \leq \bar{f}_n$ und

$$\sum_{k=0}^{n-1} f(\underline{\xi}_k)(\Delta x)_k - \int_a^b \underline{f}_n(x) dx \leq \varepsilon$$

also mit Voraussetzung (iii)

$$\left| c - \int_a^b \underline{f}_n(x) dx \right| \leq 2\varepsilon.$$

Es folgt

$$c = \lim_{n \rightarrow \infty} \int_a^b \underline{f}_n(x) dx$$

und entsprechend $c = \lim_{n \rightarrow \infty} \int_a^b \bar{f}_n(x) dx$. Insbesondere ist c durch f eindeutig bestimmt. Indem man \underline{f}_n durch $\max\{\underline{f}_n, \underline{f}_{n-1}\}$ und \bar{f}_n durch $\min\{\bar{f}_n, \bar{f}_{n-1}\}$ ersetzt, kann man $\underline{f}_{n-1} \leq \underline{f}_n \leq \bar{f}_n \leq \bar{f}_{n-1}$ erreichen.

Sei nun (i) angenommen. Wegen der Beschränktheit von f gibt es m, M mit $m \leq f(x) \leq M$ für alle $x \in [a, b]$. Wir dürfen also auch $m \leq \underline{f}_n(x) \leq \bar{f}_n(x) \leq M$ annehmen (indem wir zu $\max\{m, \underline{f}_n(x)\}$ und $\min\{M, \bar{f}_n(x)\}$ übergehen). Sei $\varepsilon > 0$ gegeben. Nach Voraussetzung gibt es ein n_0 so, dass

$$\int_a^b \bar{f}_{n_0}(x) dx - \int_a^b \underline{f}_{n_0}(x) dx \leq \frac{\varepsilon}{3}.$$

Sei N größer als die Anzahl der Teilpunkte in den gewählten Zerlegungen zu \underline{f}_{n_0} bzw. \bar{f}_{n_0} . Wähle $\delta > 0$ mit

$$\delta(M - m)N \leq \frac{\varepsilon}{3}$$

Seien nun eine Zerlegung Z von Maschenweite $\leq \delta$ und $\xi_k \in [z_k, z_{k+1}[$ gegeben. Definiere die Treppenfunktion g durch

$$g(x) = f(\xi_k) \text{ für } z_k \leq x < z_{k+1}$$

Dann gilt

$$\int_a^b \underline{f}_{n_0}(x) dx - \frac{\varepsilon}{3} \leq \int_a^b g(x) dx \leq \int_a^b \bar{f}_{n_0}(x) dx + \frac{\varepsilon}{3}.$$

Ist nämlich ein Linienelement $[z_k, z_{k+1}[$ von Z ganz in einem von \underline{f}_{n_0} enthalten, so gilt auf diesem $\underline{f}_{n_0}(x) \leq f(x)$. Andererseits gibt es höchstens N Linienelemente von Z , die nicht ganz in einem Linienelement von \underline{f}_{n_0} enthalten sind und für jedes solche gilt

$$|g(\xi_k)(\Delta x)_k - \underline{f}_{n_0}(\xi_k)(\Delta x)_k| \leq \delta(M - m).$$

Entsprechendes gilt für \bar{f}_{n_0} . Es folgt, wie in (ii) behauptet,

$$|c - \int_a^b g(x)dx| \leq \varepsilon.$$

Hat man Treppenfunktionen wie im Kasten, so hat man Treppenfunktionen

$$\underline{g}_n(x) = \max\{\underline{f}_k(x) \mid k \leq n\} \leq f(x) \leq \bar{g}_n(x) = \min\{\bar{f}_k(x) \mid k \leq n\}$$

und somit Intervallschachtelung

$$\left[\int_a^b \underline{g}_n(x)dx, \int_a^b \bar{g}_n(x)dx \right]$$

und diese bestimmt $c = \int_a^b f(x)dx$. \square

Korollar. *Sprungstetige Funktionen sind Riemann-integrierbar und das Riemann-Integral stimmt für diese mit dem oben definierten Integral überein.*

Beweis. (i) und (e). \square

Warnung. Es gibt Riemann-integrierbare Funktionen, die nicht sprungstetig sind.

Summationstheorem

Satz. *Jede stetige Funktion $f : [a, b] \rightarrow \mathbb{R}$ ist Riemann-integrierbar und es gilt*

(i) *Zu jedem $\varepsilon > 0$ gibt es $\delta > 0$ so, dass für alle $a \leq c \leq d \leq b$ mit $d - c \leq \delta$ und alle $\xi \in [c, d]$ gilt*

$$\left| \int_c^d f(x)dx - f(\xi)(d - c) \right| \leq \varepsilon(d - c)$$

Beweis. Wegen der gleichmäßigen Stetigkeit von f kann man zu jedem $\varepsilon > 0$ ein $\delta > 0$ wählen so, dass $|f(x) - f(p)| \leq \frac{\varepsilon}{2}$ falls $|x - p| \leq \delta$. Sei $\varepsilon = \frac{1}{n}$. Wähle eine Zerlegung $a = z_0 < z_1 < \dots < z_m = b$ so, dass $z_{k+1} - z_k \leq \delta$ für alle k und beliebige $\xi_k \in [z_k, z_{k+1}]$. Definiere Treppenfunktionen durch

$$\begin{aligned} \bar{f}_n(x) &= f(\xi_k) + \frac{\varepsilon}{2} && \text{für } z_k < x < z_{k+1} \\ \underline{f}_n(x) &= f(\xi_k) - \frac{\varepsilon}{2} && \text{für } z_k < x < z_{k+1} \end{aligned}$$

und $\bar{f}_n(z_k) = \underline{f}_n(z_k) = f(z_k)$. Dann folgt $\underline{f}_n \leq f \leq \bar{f}_n$ und

$$\int_a^b \bar{f}_n dx - \int_a^b \underline{f}_n dx \leq \varepsilon(b - a) = \frac{b - a}{n}.$$

Also $\lim_{n \rightarrow \infty} \int_a^b \bar{f}_n dx = \lim_{n \rightarrow \infty} \int_a^b \underline{f}_n dx$, d.h. f ist integrierbar. Sei schließlich $0 < d - c \leq \delta$ und $\xi \in [c, d]$. Die Einschränkung g von f auf $[c, d]$ ist integrierbar. Definiere auf $[c, d]$ die Treppenfunktionen

$$\begin{aligned} \bar{g}(x) &= f(\xi) + \frac{\varepsilon}{2} && \text{für } c < x < d \\ \underline{g}(x) &= f(\xi) - \frac{\varepsilon}{2} && \text{für } c < x < d \end{aligned}$$

und $\bar{g}(x) = \underline{g}(x) = g(x)$ für $x = c, d$. Dann $\underline{g} \leq g \leq \bar{g}$ und

$$f(\xi)(d-c) - \frac{\varepsilon(d-c)}{2} = \int_c^d \bar{g} dx \leq \int_c^d g(x) dx \leq \int_c^d \underline{g} dx = f(\xi)(d-c) + \frac{\varepsilon(d-c)}{2}$$

und es folgt die Behauptung (i). \square .

Summationstheorem. Sei $W(a, b) \in \mathbb{R}$ für $a \leq b$ im offenen Intervall $I \subseteq \mathbb{R}$ definiert und gelte die Additivität

$$W(a, b) = W(a, c) + W(c, b) \quad \text{falls } c \text{ zwischen } a \text{ und } b$$

Sei $f : I \rightarrow \mathbb{R}$ auf I stetig und gelte

(*) Für alle $\varepsilon > 0$ gibt es ein $\delta > 0$ so, dass für alle Δx mit $|\Delta x| \leq \delta$ und alle $p \in I$ gilt:

$$- \text{Es gibt } \xi \in [p, p + \Delta x] \text{ mit } |W(p, p + \Delta x) - f(\xi)\Delta x| \leq \varepsilon|\Delta x|$$

Dann gilt für alle $a \leq b$ in I : $W(a, b) = \int_a^b f(x) dx$.

Bei der ursprünglichen Auffassung von Integralen ist $\int_a^b f(x) dx$ die reelle Zahl, die von der Summe $\sum_a^b f(x) dx$ mit infinitesimalen dx nur infinitesimalen Abstand hat. Die Voraussetzung (*) des Theorems besagt dann, dass $\alpha = W(p, p + dx) - f(\xi) dx$ infinitesimal von zweiter Ordnung ist, d.h. dass $\frac{\alpha}{dx}$ infinitesimal ist. In dieser Form war das Theorem die Grundlage für die Einführung physikalischer Größen als Integrale. In den Natur- und Technikwissenschaften wird das heute so formuliert:

Von der Summe $W = \sum_k f(\xi_k)(\Delta x)_k$ mit den Linienelementen $(\Delta x)_k$ geht man mit $(\Delta x)_k \rightarrow 0$ zum Integral $W = \int f(x) dx$ über (unter Bezugnahme auf den Fall stückweise konstanter Funktionen).

Ist die Stetigkeit von $f(x)$ gegeben, so handelt es sich demnach um eine Anwendung des Theorems. In der gängigen mathematischen Literatur, wird bestenfalls bei der Definition einer Größe durch ein Integral eine Motivation im Sinne des Theorems gegeben. Die infinitesimale Form findet man in Lehrbüchern zur Nonstandard-Analysis. Das Theorem verallgemeinert sich in offensichtlicher Weise auf Bereichsintegrale und mit mehr Anstrengung auf Kurven- und Flächenintegrale - das Problem ist die Zuordnung von geradlinigen Strecken- bzw. Flächenelementen zu gekrümmten Kurven- bzw. Flächenstücken.

Beweis. Sei $a \geq b$ und $\varepsilon > 0$ gegeben. Wähle δ gemäß (*) und nach dem vorangehenden Satz so, dass

$$\left| \int_p^{p+\Delta x} f(x) dx - f(\xi)\Delta x \right| \leq \varepsilon|\Delta x| \quad \text{für alle } p \in [a, b], |\Delta x| \leq \delta \text{ und passendes } \xi \in [p, p + \Delta x].$$

Wähle ein Zerlegung $a = z_0 < z_1 < \dots < z_n = b$ von Maschenweite $\leq \delta$ und $\xi_k \in [z_k, z_{k+1}]$ so, dass $|W(z_k, z_{k+1}) - f(\xi_k)(z_{k+1} - z_k)| \leq \varepsilon(z_{k+1} - z_k)$. Dann folgt

$$|W(z_k, z_{k+1}) - \int_{z_k}^{z_{k+1}} f(x) dx| \leq$$

$$\leq |W(z_k, z_{k+1}) - f(\xi_k)(z_{k+1} - z_k)| + |f(\xi_k)(z_{k+1} - z_k) - \int_{z_k}^{z_{k+1}} f(x)dx| \leq 2\varepsilon(z_{k+1} - z_k)$$

also mit der Additivität und Dreiecksungleichung

$$|W(a, b) - \int_a^b f(x)dx| \leq \sum_{k=0}^{n-1} |W(z_k, z_{k+1}) - \int_{z_k}^{z_{k+1}} f(x)dx| \leq 2\varepsilon(b - a)$$

Dies gilt für alle $\varepsilon > 0$, also nach Archimedes $W(a, b) = \int_a^b f(x)dx$. \square

23.3 Mehrfachintegrale

Der folgende Satz ist der Schlüssel zur Berechnung von Riemann-Integralen im \mathbb{R}^n . Durch ihn wird die Berechnung eines Integrales auf die Berechnung von zwei Integralen in Räumen von niedrigerer Dimension als n zurückgeführt. Durch mehrfache Anwendung dieses Satzes führt man schließlich die Berechnung eines Integrals über einem Intervall im \mathbb{R}^n auf die Berechnung von „gewöhnlichen“ Riemann-Integralen über „gewöhnlichen“ Intervallen in \mathbb{R}^1 zurück.

23.3.1 Satz von Fubini

Satz 23.10 (Satz von Fubini) *Seien $I_x \subseteq \mathbb{R}^k$ und $I_y \subseteq \mathbb{R}^\ell$ abgeschlossene Intervalle und sei $I = I_x \times I_y \subseteq \mathbb{R}^k \times \mathbb{R}^\ell = \mathbb{R}^{k+\ell}$. Weiter sei die Funktion $f : I_x \times I_y \rightarrow \mathbb{R}$ Riemann-integrierbar auf I .*

- Existiert für jedes feste $y \in I_y$ das Riemann-Integral

$$g(y) := \int_{I_x} f(x, y)dx,$$

so ist die Funktion g Riemann-integrierbar, d.h. es existiert das iterierte Integral

$$\int_{I_y} \left(\int_{I_x} f(x, y)dx \right) dy,$$

und es ist

$$\int_I f(x, y) d(x, y) = \int_{I_y} \left(\int_{I_x} f(x, y)dx \right) dy.$$

- Existiert für jedes feste $x \in I_x$ das Riemann-Integral

$$h(x) := \int_{I_y} f(x, y)dy,$$

so ist h Riemann-integrierbar, das iterierte Integral

$$\int_{I_x} \left(\int_{I_y} f(x, y)dy \right) dx$$

existiert, und

$$\int_I f(x, y) d(x, y) = \int_{I_x} \left(\int_{I_y} f(x, y)dy \right) dx.$$

Ist f stetig, so existieren g bzw. h und sind stetig.

Man beachte, dass aus der Existenz der beiden iterierten Integrale

$$\int_{I_y} \left(\int_{I_x} f(x, y) dx \right) dy \quad \text{und} \quad \int_{I_x} \left(\int_{I_y} f(x, y) dy \right) dx$$

noch nicht die Existenz des Integrals

$$\int_I f(x, y) d(x, y) \tag{23.20}$$

folgt.

23.3.2 Beweis.

Schreibweise

$$f(C) = f(x), \quad x \in C^o \quad \text{falls } f|_{C^o} \text{ konstant}$$

Wir benutzen Zerlegungen

$$Z_n = \{C \times D \mid C \in Z_{x_n}, D \in Z_{y_n}\}$$

wobei Z_{x_n} Zerlegung von I_x und Z_{y_n} Zerlegung von I_y mit Weite $\rightarrow 0$. Seien

$$\underline{f}_n \leq f \leq \bar{f}_n, \quad \int \bar{f}_n - \int \underline{f}_n \rightarrow 0$$

dazu passende Treppenfunktionen. Dann ist

$$\underline{g}_n(y) := \sum_{C \in Z_{x_n}} \underline{f}_n(C \times D) \mu_0(C) \leq g(y), \quad y \in D \in Z_{y_n}$$

Treppenfunktion auf I_y passend zu Z_{y_n} und

$$\int \underline{f}_n \leq \int \underline{g}_n$$

Ebenso erhält man

$$g \leq \bar{g}_n, \quad \int \bar{g}_n \leq \int \bar{f}_n$$

Also existiert $\int g$ und $\int g = \int f$.

Ist f stetig, so auch g . Dazu sei $n = k + l$. Gegeben $\varepsilon > 0$. Da f auf I gleichmäßig stetig ist, gibt es $\delta > 0$ so, dass für alle $((x, y), (x', y') \in I$

$$\|(x, y) - (x', y')\| \leq \delta \Rightarrow |f(x', y') - f(x, y)| \leq \varepsilon$$

Seien nun $y, y' \in I_y$ mit $\|y - y'\| \leq \frac{\delta}{\sqrt{2}}$. Zerlege $I_x = I_{x_1} \oplus \dots \oplus I_{x_m}$ mit I_{x_j} von Weite $\leq \frac{\delta}{\sqrt{k}}$. Nach dem Mittelwertsatz gibt es $\xi_j, \xi'_j \in I_{x_j}$ mit

$$\int_{I_{x_j}} f(x, y) dx = f(\xi_j, y) \mu_0(I_{x_j}) \quad \text{und} \quad \int_{I_{x_j}} f(x, y') dx = f(\xi'_j, y') \mu_0(I_{x_j})$$

also $\|(\xi_j, y) - (\xi'_j, y')\| \leq \delta$ und somit

$$\left| \int_{I_{x_j}} f(x, y) dx - \int_{I_{x_j}} f(x, y') dx \right| = |f(\xi_j, y) - f(\xi'_j, y')| \mu_0(I_{x_j}) \leq \varepsilon \mu_0(I_{x_j})$$

Da $\mu_0(I_x) = \sum_{j=1}^n \mu_0(I_{x_j})$, folgt

$$(**) \quad |g(y) - g(y')| = \left| \int_{I_x} f(x, y) dx - \int_{I_x} f(x, y') dx \right| \leq \varepsilon \mu_0(I_x)$$

Somit ist g stetig und integrierbar. \square

23.3.3 Beispiele

Beispiel 1 Sei $f(x, y) = \cos(2x + 3y)$ auf $I = [0, \frac{\pi}{4}] \times [0, \frac{\pi}{12}]$. Die Funktion f ist auf I stetig, so dass wir $\int_I f(x, y) d(x, y)$ mit Hilfe iterierter Integrale bestimmen können. Es ist

$$\begin{aligned} \int_I f(x, y) d(x, y) &= \int_0^{\pi/4} \left(\int_0^{\pi/12} \cos(2x + 3y) dy \right) dx \\ &= \int_0^{\pi/4} \frac{1}{3} \sin(2x + 3y) \Big|_0^{\pi/12} dx \\ &= \int_0^{\pi/4} \frac{1}{3} \left(\sin\left(2x + \frac{\pi}{4}\right) - \sin 2x \right) dx \\ &= -\frac{1}{6} \cos\left(2x + \frac{\pi}{4}\right) + \frac{1}{6} \cos 2x \Big|_0^{\pi/4} = \frac{1}{6}(\sqrt{2} - 1). \end{aligned}$$

Das gleiche Resultat hätte natürlich auch die Berechnung von

$$\int_0^{\pi/12} \left(\int_0^{\pi/4} \cos(2x + 3y) dx \right) dy$$

gebracht. ■

Beispiel 2 Sei $f(x, y, z) = x + y + z + xz$ auf $I = [0, 1] \times [1, 2] \times [2, 3]$. Dann ist f stetig auf I , und die zweifache Anwendung des Satzes von Fubini ergibt

$$\begin{aligned}
\int_I f(x, y, z) d(x, y, z) &= \int_0^1 \left(\int_{[1,2] \times [2,3]} f(x, y, z) d(y, z) \right) dx \\
&= \int_0^1 \int_1^2 \int_2^3 (x + y + z + xz) dz dy dx \\
&= \int_0^1 \int_1^2 \left(xz + yz + \frac{z^2}{2} + x \frac{z^2}{2} \right) \Big|_2^3 dy dx \\
&= \int_0^1 \int_1^2 \left(x + y + \frac{5}{2} + \frac{5}{2}x \right) dy dx \\
&= \int_0^1 \left(\frac{7}{2}xy + \frac{y^2}{2} + \frac{5}{2}y \right) \Big|_1^2 dx \\
&= \int_0^1 \left(\frac{7}{2}x + \frac{3}{2} + \frac{5}{2} \right) dx \\
&= \frac{7}{4}x^2 + 4x \Big|_0^1 = \frac{7}{4} + 4 = \frac{23}{4}.
\end{aligned}$$

Das gleiche Resultat hätten wir beispielsweise auch aus dem iterierten Integral

$$\int_2^3 \int_0^1 \int_1^2 (x + y + z + xz) dy dx dz$$

gewonnen. ■

24 Messbarkeit von Mengen und Integrale

24.1 Jordan messbare Mengen

24.1.1 Definition

Sei $B \subseteq \mathbb{R}^n$ nichtleer und beschränkt. Die Funktion

$$\chi_B : \mathbb{R}^n \rightarrow \mathbb{R}, \quad \chi_B(x) = \begin{cases} 1 & \text{wenn } x \in B \\ 0 & \text{wenn } x \notin B \end{cases}$$

heißt die *charakteristische Funktion von B*.

Sei $I \supseteq B$ ein abgeschlossenes Intervall und Z eine Rechteckzerlegung von I . Ist $T \subseteq Z$ so sei

$$\bigcup T = \bigcup_{C \in T} C, \quad \mu_0(T) = \sum_{C \in T} \mu_0(C)$$

Die *innere* bzw. *äußere Approximation* von B in Z ist definiert als

$$U(Z, B) := \sum_{J \in Z, J \subseteq B} \mu_0(J) \leq O(Z, B) := \sum_{J \in Z, J \cap B \neq \emptyset} \mu_0(J)$$

Ist Z' feiner als Z so gilt

$$U(Z, B) \leq U(Z', B) \leq O(Z', B) \leq O(Z, B)$$

Lemma 24.1 Für eine nichtleere beschränkte Menge $B \subseteq \mathbb{R}^n$, abgeschlossenes Intervall $I \subseteq B$ und $c \in \mathbb{R}_{\geq 0}$ sind die folgenden Aussagen äquivalent

$$(i) \quad c = \int_I \chi_B$$

(ii) Es gibt Zerlegungen Z_n von I und $\underline{T}_n \subseteq \{J \in Z \mid J \subseteq B\}$ und $\overline{T}_n \supseteq \{J \in Z \mid J \cap B \neq \emptyset\}$ mit $\mu_0(\underline{T}_n) \leq c \leq \mu_0(\overline{T}_n)$ und $\mu_0(\overline{T}_n) - \mu_0(\underline{T}_n) \rightarrow 0$

(iii) Für alle Zerlegungen Z_n von I mit Weite $\rightarrow 0$ gilt $U(Z_n, B) \leq c \leq O(Z_n, B)$ und $O(Z_n, B) - U(Z_n, B) \rightarrow 0$.

Das ist unabhängig von I .

Insbesondere ist c dann eindeutig bestimmt und wir sagen, dass B (Jordan) *messbar* ist mit (Jordan) *Maß* bzw. *Inhalt* $c = \mu(B)$. Fig.5.

Beweis. (i) \Leftrightarrow (iii). Für

$$\inf \chi_B|_J = 0 \text{ für } J \not\subseteq B, \quad \sup \chi_B|_J = 1 \text{ für } J \cap B \neq \emptyset$$

also

$$U(Z, B) = U(Z, \chi_B), \quad O(Z, B) = O(Z, \chi_B)$$

(iii) \Rightarrow (ii) trivial. (ii) \Rightarrow (i): setze

$$\underline{f}_n|_J = \begin{cases} 1 & \text{falls } J \in \underline{T}_n \\ 0 & \text{falls } J \notin \underline{T}_n \end{cases} \quad \overline{f}_n|_J = \begin{cases} 1 & \text{falls } J \in \overline{T}_n \\ 0 & \text{falls } J \notin \overline{T}_n \end{cases}$$

Dann

$$\underline{f}_n \leq \chi_B \leq \overline{f}_n, \quad \mu_0(\underline{T}_n) = \int \underline{f}_n, \quad \mu_0(\overline{T}_n) = \int \overline{f}_n$$

Zur Unabhängigkeit von I : hat man ein zweites I' , so O.B.d.A. $I' \subseteq I$ und man betrachte Z'_n mit den nichtleeren $J \cap I'$, $J \in Z$. \square

Für Intervalle I folgt sofort $\mu(I) = \mu_0(I)$. Man beachte aber die Abhängigkeit vom Koordinatensystem.

Beispiel 3 Sei $B = \{(x, y) \in \mathbb{R} \times \mathbb{R} : 0 \leq x \leq 1, 0 \leq y \leq 1, x \text{ rational}\}$, und für I wählen wir das Quadrat $[0, 1] \times [0, 1]$. Da \mathbb{Q} in \mathbb{R} dicht liegt, ist für jede Rechteckzerlegung Z von I

$$U(Z, B) = 0, \quad O(Z, B) = 1.$$

Also ist B *nicht* Jordan-messbar, und wir können B keinen Flächeninhalt zuordnen. \blacksquare

Bemerkung 24.2 Eine nichtleere Menge $B \subseteq \mathbb{R}^n$ ist genau dann Jordan-messbar, wenn ihr Rand ∂B Jordan-messbar ist und den Jordan-Inhalt $\mu(\partial B) = 0$ hat.

Eine messbare Menge B ist eine *Nullmenge*, wenn $\mu(B) = 0$. Zum Beweis der Bemerkung benötigt man man den Begriff der Lebesgue-Nullmenge und die Charakterisierung der Riemann-integrierbaren Funktionen.

24.1.2 Vereinigungs- und Schnittmengen

Satz 24.3 Sind $A, B \subseteq \mathbb{R}^n$ Jordan-messbar, so sind auch $A \cup B$, $A \cap B$ und $B \setminus A$ Jordan-messbar und es gilt

$$\mu(A \cup B) + \mu(A \cap B) = \mu(A) + \mu(B). \quad (24.21)$$

$$A \subseteq B \Rightarrow \mu(A) \leq \mu(B)$$

Beweis. Betrachte $I \supseteq A \cup B$ und $\text{Weite}(Z_n) \rightarrow 0$ und $J \in Z_n$. J trifft die Menge $A \cup B$ genau dann, wenn es entweder $A \cap B$ oder nur A oder nur B oder sowohl A wie B aber nicht $A \cap B$ trifft.. Es folgt

$$O(Z_n, A \cup B) + O(Z_n, A \cap B) \leq O(Z_n, A) + O(Z_n, B)$$

Weiterhin

$$U(Z_n, A \cup B) + U(Z_n, A \cap B) \geq U(Z_n, A) + U(Z_n, B)$$

da aus $J \subseteq A$ oder $J \subseteq B$ folgt, dass $J \subseteq A \cup B$, und da $\mu(J)$ sowohl bei A wie bei B genau dann eingeht, wenn $J \subseteq A \cap B$. Es folgt

$$\begin{aligned} 0 &\leq [O(Z_n, A \cup B) - U(Z_n, A \cup B)] + [O(Z_n, A \cap B) - U(Z_n, A \cap B)] \\ &\leq [O(Z_n, A) - U(Z_n, A)] + [O(Z_n, B) - U(Z_n, B)] \rightarrow 0 \end{aligned}$$

Wegen der Nichtnegativität der Summanden folgt

$$[O(Z_n, A \cup B) - U(Z_n, A \cup B)] \rightarrow 0, \quad [O(Z_n, A \cap B) - U(Z_n, A \cap B)] \rightarrow 0$$

und aus (*) die Aussage über die Maße.

Alternativ: $\chi_A \cdot \chi_B = \chi_{A \cap B}$ ist auf I integrierbar (vgl. 24.3.3), also ist $A \cap B$ messbar. $\chi_{I \setminus A} = 1 - \chi_A$, also $I \setminus A$ messbar und somit auch $A \cup B = I \setminus [(I \setminus A) \cap (I \setminus B)]$ sowie $A \setminus B = A \cap (I \setminus B)$. Ist $A \subseteq B$, so $A = B \cup (A \setminus B)$ und $B \cap (A \setminus B) = \emptyset$, also $\mu(B) + \mu(A \setminus B) = \mu(A)$ und daher $\mu(B) \leq \mu(A)$. \square

Korollar 24.4 Seien A, B wie im vorangehenden Satz und $\mu(A \cap B) = 0$. Dann gilt

$$\mu(A \cup B) = \mu(A) + \mu(B). \quad (24.22)$$

24.2 Krumme Zerlegungen

24.2.1 Definition

Sei $B \subseteq \mathbb{R}^n$ beschränkt und im abgeschlossenen Intervall I enthalten. Eine Zerlegung Z von B ist dann eine endliche Menge kompakter Jordan-messbarer Mengen C , den Zellen der Zerlegung, so dass

- $B \subseteq \bigcup_{C \in Z} C \subseteq I$
- $\mu(C \cap C') = 0$ für alle $C \neq C'$ in Z

d.h. die Zellen von Z sind in I enthalten, jeder Punkt von B ist in mindestens einer Zelle enthalten und zwei verschiedene Zellen haben nur eine Nullmenge gemeinsam.

Beispiele. 1. B Kreissektor mit Radius r und Winkel ϕ (im Bogenmaß): Zerlegung Z_n hat n kongruente Dreiecke und n kongruente Trapeze. Es folgt mit Satz 24.6: $\mu(B) = \frac{1}{2}r^2\phi$.

2. B rechtwinkliges Dreieck mit Katheten der Längen g und h . Zerlegung Z_n besteht aus $2n$ rechtwinkligen Dreiecken, deren Hypotenuse zur x -Achse parallel ist und die sich paarweise zur Rechtecken R_n ergänzen - diese haben als Maß dann 'Grundseite mal Höhe. Für die Summen g_n bzw. h_n gilt $g_n \rightarrow g$ und $h_n \rightarrow h$. Die restlichen $n + 1$ -Zellen sind rechtwinklige Dreiecke, deren Hypotenuse Teil der von B ist - messbar via Zerlegung durch Parallele zur x -Achse. Die Summe ihrer Maße geht gegen Null - wie für entsprechende einhüllende Achsenparallele Rechtecke, Es folgt mit Satz 24.6: $\mu(B) = gh$.

24.2.2 Verfeinerung, Weite und Einschränkung

Z' ist *Verfeinerung* von Z , wenn jede Zelle von Z' in einer Zelle von Z enthalten ist. Eine Zerlegung Z hat *Weite* oder *Feinheit* $\leq \delta$ falls es zu jedem $C \in Z$ ein Intervall J der Weite δ gibt mit $C \subseteq J$.

Nach Satz 24.3 erhält man zu je zwei Zerlegungen Z_1, Z_2 eine gemeinsame Verfeinerung, deren Zellen die nichtleeren $C_1 \cap C_2$ mit $C_i \in Z_i$ sind.

Ist B messbar, so ist zu jeder Zerlegung Z_B von B die *Einschränkung*

$$Z_B = \{B \cap C \mid C \in Z\}$$

wieder eine Zerlegung von B .

24.2.3 Messbarkeit

Ist eine Zerlegung Z von B gegeben, so definieren wir

$$U(Z, B) = \sum_{J \in Z, J \subseteq B} \mu J, \quad O(Z, B) = \sum_{J \in Z, J \cap B \neq \emptyset} \mu J$$

Offensichtlich $O(Z, B) \geq U(Z, B)$.

Für Zerlegungen Z, Z' definieren wir die *Differenz*

$$Z' - Z = \{J \in Z' \mid J \not\subseteq C \text{ für alle } C \in Z\}$$

Lemma 24.5 Sei Z Zerlegung für B auf I . Dann gibt es zu jedem $\varepsilon > 0$ eine Gitterzerlegung Z^ε auf I so, dass

$$\sum_{J \in Z^\varepsilon - Z, J \cap B \neq \emptyset} \mu(J) \leq \varepsilon$$

und

$$O(Z^\varepsilon, B) - U(Z^\varepsilon, B) \leq O(Z, B) - U(Z, B) + \varepsilon$$

Beweis. Wähle für jedes $C \in Z$ eine Gitterzerlegung Z_C so, dass

$$O(Z_C, C) - U(Z_C, C) \leq \frac{1}{c} \varepsilon$$

c die Anzahl der Zellen von Z . Wähle Z^ε als gemeinsame Verfeinerung aller (endlich vielen) Z_C . Dann

$$\begin{aligned} \sum_{J \in Z^\varepsilon - Z, J \cap B \neq \emptyset} \mu(J) &\leq \sum_{C \in Z} \sum_{J \in Z^\varepsilon - Z, J \cap C \neq \emptyset} \mu(J) \\ &\leq \sum_{C \in Z} O(Z^\varepsilon, C) - U(Z^\varepsilon, C) \leq \sum_{C \in Z} O(Z_C, C) - U(Z_C, C) \leq c \frac{\varepsilon}{c} = \varepsilon \end{aligned}$$

und es folgt mit $T = \{J \in Z^\varepsilon \mid J \cap B \neq \emptyset, J \not\subseteq B\}$

$$\begin{aligned} O(Z^\varepsilon, B) - U(Z^\varepsilon, B) &= \sum_{J \in T} \mu(J) = \sum_{C \in Z} \sum_{J \in T, J \subseteq C} \mu(J) + \sum_{J \in T, \forall C \in Z: J \not\subseteq C} \mu(J) \\ &\leq \sum_{C \in Z, C \not\subseteq B, C \cap B \neq \emptyset} \mu(C) + \varepsilon = O(Z, B) - U(Z, B) + \varepsilon \end{aligned}$$

Satz 24.6 Sei $B \subseteq \mathbb{R}^n$ eine nichtleere beschränkte Menge. Dann ist B genau dann Jordan-messbar, wenn es ein abgeschlossenes Intervall $I \subseteq B$ und eine Folge Z_n von Zerlegungen gibt so, dass

$$\lim_{n \rightarrow \infty} [O(Z_n, B) - U(Z_n, B)] = 0$$

In diesem Fall gilt

$$\mu(B) = \lim_{n \rightarrow \infty} U(Z_n, B)$$

Beweis. Ist die Folge Z_n gegeben, so wähle man Z_n^ε nach dem vorangehenden Lemma mit $\varepsilon = \frac{1}{n}$. Dann

$$O(Z_n^\varepsilon, B) - U(Z_n^\varepsilon, B) \leq O(Z_n, B) - U(Z_n, B) + \frac{1}{n} \rightarrow 0$$

und damit die Behauptung. \square

24.2.4 Bewegungsinvarianz

Eine *Bewegung* ist eine Abbildung $\phi : \mathbb{R}^n \rightarrow \mathbb{R}^n$ mit $|\phi(x)| = |x|$.

Ist ϕ Bewegung in \mathbb{R}^2 und B messbar, so ist auch $\phi(B)$ messbar und $\mu(\phi(B)) = \mu(B)$

Insbesondere ist die Rechteckfläche = Grundseite mal Höhe. Der Beweis folgt mit Satz 24.6 und 24.2.1 2. Beispiel. \square

24.3 Riemann-Integrale

24.3.1 Treppenfunktionen

Sei $B \subseteq \mathbb{R}^n$ eine messbare Menge $f : B \rightarrow \mathbb{R}$ eine beschränkte Funktion. Zu einer Zerlegung Z von B und betrachtet man die *Zwischenvektoren* $\xi = (\xi_C \mid C \in Z_B)$ mit $\xi_C \in \text{int}C$ im Innern von C und die *Riemannsumme*

$$R(Z, \xi, f) = \sum_{C \in Z_B} f(\xi_C) \mu(C).$$

f ist eine *Treppenfunktion* zu $X = Z$, wenn f für jedes $C \in Z_B$ auf C° konstant ist - den Wert bezeichnen wir mit $f(C)$. In diesem Fall setzt man, unabhängig von der Wahl der Zwischenvektoren

$$\int_B f = R(Z, \xi, f)$$

Für beliebiges beschränktes f definiert man *Unter-* und *Obersummen*

$$U(Z, f) = R(Z, \underline{\xi}, f), \quad O(Z, f) = R(Z, \bar{\xi}, f)$$

mit Zwischenvektoren so, dass

$$\underline{\xi}_C = \inf f(C), \quad \bar{\xi}_C = \sup f(C)$$

Für jede nichtleere Menge $B \subseteq \mathbb{R}^n$ und jede Funktion $f : B \rightarrow \mathbb{R}$ definieren wir die *Fortsetzung* (oder *Erweiterung*) \hat{f} von f durch

$$\hat{f} : \mathbb{R}^n \rightarrow \mathbb{R}, \quad \hat{f}(x) = \begin{cases} f(x) & \text{falls } x \in B \\ 0 & \text{sonst,} \end{cases}$$

d.h. \hat{f} setzt f durch 0 auf ganz \mathbb{R}^n fort.

24.3.2 Riemann-Intergrale auf messbaren Mengen

Satz 24.7 *Sei B eine messbare Menge in \mathbb{R}^n und $f : B \rightarrow \mathbb{R}$ eine beschränkte Funktion. Für $c \in \mathbb{R}$ sind äquivalent*

- (i) *Es gibt Treppenfunktionen $\underline{f}_k, \bar{f}_k : B \rightarrow \mathbb{R}$ mit $\underline{f}_k \leq f \leq \bar{f}_k$ für alle k und $\lim_{k \rightarrow \infty} \int_B \underline{f}_k(x) dx = c = \lim_{k \rightarrow \infty} \int_B \bar{f}_k(x) dx$*
- (ii) *Für alle $\varepsilon > 0$ gibt es $\delta > 0$ so, dass für alle Zerlegungen Z von B mit Weite $\leq \delta$ und alle Zwischenvektoren ξ gilt: $|c - R(Z, \xi, f)| \leq \varepsilon$*
- (iii) *Für alle $\varepsilon > 0$ gibt es eine Gitter-Zerlegung Z (alternativ: Zerlegung mit $Z = Z_B$) von B so, dass für alle Zwischenvektoren ξ gilt: $|c - R(Z, \xi, f)| \leq \varepsilon$*
- (iv) *$\sup_Z U(Z, f) = c = \inf_Z O(Z, f)$, wobei Z über alle Zerlegungen (mit $Z = Z_B$), alternativ alle Gitterzerlegungen läuft.*
- (v) *Es gibt ein abgeschlossenes Intervall $I \supseteq B$, so dass die Fortsetzung \hat{f} auf I Riemann integrierbar ist und $c = \int_I \hat{f} dx$.*

Es gibt höchstens ein c , das (iii) erfüllt und es gilt $m(b-a) \leq c \leq M(b-a)$ falls $m \leq f(x) \leq M$ für alle $x \in B$.

Beweis. "Äquivalenz von (i) bis (iv) nach bekanntem Muster. (v) \Rightarrow (iv) trivial. Gelte (v), sei o.B.d.A. $0 \leq f \leq M$. und sei Z_n Zerlegung. Wähle Z_n^ε nach Lemma 24.5. Dann

$$O(Z_n^\varepsilon, \hat{f}) = O(Z_n, f), \quad U(Z_n^\varepsilon, \hat{f}) = U(Z_n, f) \pm \varepsilon M \quad \square$$

Definition 24.8 Gilt eine (also alle) der Aussagen (i)-(iv) für ein $c \in \mathbb{R}$, so heißt f Riemann-integrierbar auf B , und der eindeutig bestimmte Wert c heißt Riemann-Integral von f auf B . Als Bezeichnung wählen wir

$$c = \int_B f(x_1, \dots, x_n) d(x_1, \dots, x_n) \quad \text{oder kurz} \quad \int_B f(x) dx.$$

Besser wäre z.B. $\int_B f \prod_i dx_i$ oder einfach $\int_B f(\vec{x})$.

Korollar 24.9 Beschränktes $f : B \rightarrow \mathbb{R}$ ist Riemann-integrierbar genau dann, wenn es Treppenfunktionen $\underline{f}_k \leq f \leq \bar{f}_k$ gibt mit $\int_I \bar{f}_k(x) dx - \int_I \underline{f}_k(x) dx \rightarrow 0$ für $k \rightarrow \infty$.

Beweis wie für Intervalle. Beachte $R(Z, \xi, f) = R(Z_B, \xi, f)$. \square

Beispiel: B Einheitskreis, $f(x) = 1 - |x|$. $\int_B f = \frac{\pi}{3}$.

24.3.3 Rechenregeln

Als nächstes geben wir einige Resultate über das Rechnen mit Riemann-Integralen auf Jordan-messbaren Mengen an. Das Analogon zu Satz 23.5 lautet wie folgt.

Satz 24.10 Sei $B \subseteq \mathbb{R}^n$ Jordan-messbar, und $f, g : B \rightarrow \mathbb{R}$ seien Riemann-integrierbar.

- Für beliebige $\alpha, \beta \in \mathbb{R}$ ist $\alpha f + \beta g$ Riemann-integrierbar auf B , und

$$\int_B (\alpha f(x) + \beta g(x)) dx = \alpha \int_B f(x) dx + \beta \int_B g(x) dx.$$

- Ist $f(x) \leq g(x)$ für alle $x \in B$, so ist auch

$$\int_B f(x) dx \leq \int_B g(x) dx.$$

- Es ist

$$\left| \int_B f(x) dx \right| \leq \sup_{x \in B} |f(x)| \cdot \mu(B).$$

- fg ist auf B integrierbar.

Beweis von (d). Wegen (a) o.B.d.A. $0 \leq f$ und $0 \leq g$. Die Integrierbarkeit von f bzw. g sei bezeugt durch Treppenfunktionen mit $\underline{f}_n \leq f \leq \bar{f}_n$ und $\underline{g}_n \leq g \leq \bar{g}_n$ zur Zerlegungsfolge Z_n . Sei $h = fg$. Wir setzen

$$\underline{h}_n(C) = \underline{f}_n(C)\underline{g}_n(C), \quad \bar{h}_n(C) = \bar{f}_n(C)\bar{g}_n(C) \quad C \in Z_n$$

un erhalten

$$\begin{aligned} \int_B \bar{h}_n - \int_B \underline{h}_n &= \sum_{C \in Z_n} (\bar{f}_n(C)\bar{g}_n(C) - \underline{f}_n(C)\bar{g}_n(C) + \bar{g}_n(C)\underline{f}_n(C) - \underline{g}_n(C)\underline{f}_n(C))\mu(C) \\ &\leq \|g\|_\infty \sum_{C \in Z_n} (\bar{f}_n - \underline{f}_n)\mu(C) + \|f\|_\infty \sum_{C \in Z_n} (\bar{g}_n - \underline{g}_n)\mu(C) \rightarrow 0 \end{aligned}$$

Nun sehen wir uns noch an, wie das Integral bei fester Funktion f vom Integrationsbereich abhängt. Dazu vereinbaren wir, dass $\int_\emptyset f dx = 0$.

Satz 24.11 *Ist $A \subseteq B$ messbar und f auf B integrierbar, so ist f auch auf D . Sind $A, B \subseteq \mathbb{R}^n$ Jordan-messbar und ist f auf A und auf B Riemann-integrierbar, so ist f auch auf $A \cup B$, $A \cap B$, $A \setminus B$ Riemann-integrierbar, und es gilt*

$$\int_{A \cup B} f(x) dx + \int_{A \cap B} f(x) dx = \int_A f(x) dx + \int_B f(x) dx.$$

Beweis. $B \setminus A$ ist messbar, also erhalten wir zu jeder Zerlegung Z_n von B Zerlegungen Z'_n von A und Z''_n von $B \setminus A$ mit Zellen

$$C \cap A, \quad \text{bzw } C \cap (B \setminus A), \quad C \in Z_n$$

Für Treppenfunktionen \underline{f}_n und \bar{f}_n , die Integrierbarkeit von f auf B belegen, erhalten wir

$$\sum_{C \in Z_n} (\bar{f}_n(C) - \underline{f}_n(C))\mu(C \cap A) + \sum_{C \in Z_n} (\bar{f}_n(C) - \underline{f}_n(C))\mu(C \cap (B \setminus A)) = \int_B \bar{f}_n - \int_B \underline{f}_n \rightarrow 0$$

also existiert $\int_A f$.

Sei angenommen, dass $\int_B f$ und $\int_A f$ existieren mit zugehörigen Treppenfunktionen $\underline{f}_n^A, \bar{f}_n^A$ zu Z_n^A bzw. $\underline{f}_n^B, \bar{f}_n^B$ zu Z_n^B . Sei zunächst $A \cap B = \emptyset$. Dann erhält man passende Treppenfunktionen auf $A \cup B$ als gemeinsame Fortsetzungen und damit die Behauptung. Indem wir diesen Spezialfall auf den allgemeinen anwenden

$$\begin{aligned} \int_A f + \int_B f &= \int_{A \cap B} f + \int_{A \setminus B} f + \int_{A \cap B} f + \int_{B \setminus A} f \\ &= \int_{A \cap B} f + \int_A f + \int_{B \setminus A} f = \int_{A \cap B} f + \int_{A \cup B} f \quad \square \end{aligned}$$

Korollar 24.12 *Seien A, B, f wie eben und $\mu(A \cap B) = 0$. Dann gilt*

$$\int_{A \cup B} f(x) dx = \int_A f(x) dx + \int_B f(x) dx.$$

24.3.4 Summationstheorem

Ist $W(C)$ für alle kompakten messbaren Teilmengen von B definiert, so nennen wir W *additiv*, falls gilt

$$W(C_1 \cup C_2) = W(C_1) + W(C_2) \quad \text{falls } C_1 \cap C_2 \subseteq \partial C_1 \cap \partial C_2$$

Theorem 24.13 Sei $B \subseteq \mathbb{R}^n$ kompakt und Jordan-messbar und $f : I \rightarrow \mathbb{R}$ stetig. Dann gilt

- (i) $f : B \rightarrow \mathbb{R}$ ist Riemann-integrierbar und zu jedem $\varepsilon > 0$ gibt es $\delta > 0$ so, dass für alle Intervalle $[c, d] \subseteq B$ von Weite $\leq \delta$ und alle messbaren $C \subseteq [c, d]$ und $\xi \in C$ gilt

$$\left| \int_C f(x) dx - f(\xi) \mu(C) \right| \leq \varepsilon \mu(C)$$

- (ii) Mittelwertsatz: Ist B wegzusammenhängend, so gibt $\xi \in B$ mit $\int_B f(x) dx = f(\xi) \mu(B)$

- (iii) Sei $W(C)$ für alle kompakten messbaren $C \subseteq B$ definiert und additiv. Sei Z_n eine Folge von Zerlegungen von B mit $\text{Weite}(Z_n) \rightarrow 0$. Gelte weiterhin

- (*) Für alle $\varepsilon > 0$ gibt es ein n_0 so, dass für alle $n \geq n_0$ und $C \in Z_n$ gilt

$$\text{Es gibt } \xi \in C \text{ mit } |W(C) - f(\xi) \mu(C)| \leq \varepsilon \mu(C)$$

Dann gilt für alle kompakten messbaren $D \subseteq B$: $W(D) = \int_D f(x) dx$.

Der Beweis folgt wie für Intervalle. In (iii) wendet man Lemma 24.5 an auf die Einschränkung Z_{nD} und zerlegt D entsprechend Z_{nD}^ε in Zellen. Dann ist der Beitrag der Zellen $\not\subseteq D$ kleiner als $\varepsilon \max_{x \in B} |f(x)| \rightarrow 0$. Etwas genauer: Wir setzen $W(\emptyset) = 0$. Gegeben $\eta > 0$ wähle $\varepsilon < \frac{1}{27} \eta$ so, dass $\varepsilon M < \frac{\eta}{27} \mu(D)$ für $M = \max\{|f(x)| \mid x \in B\}$. Wähle δ gemäss (i). Wähle n_0 gemäß (*) zu ε und so, dass $\text{Weite}(Z_n) \leq \delta$ für alle $n \geq n_0$. Wir betrachten $n \geq n_0$ schreiben $Z' = Z_{nD}^\varepsilon$. Dann mit der Additivität und mit $\xi_C \in C$

$$\begin{aligned} W(D) &= \sum_{C \in Z'} W(C \cap D) = \sum_{C \in Z', C \subseteq D} f(\xi_C) \mu(C) \pm \varepsilon \mu(C) \pm \sum_{C \in Z', C \not\subseteq D} |W(C)| \\ &= \sum_{C \in Z', C \subseteq D} \left[\int_C f dx \pm 2\varepsilon \mu(C) \right] \pm \sum_{C \in Z', C \cap D \neq \emptyset, C \not\subseteq D} M \cdot \mu(C) \\ &= \sum_{C \in Z', C \subseteq D} \int_C f dx \pm 2\varepsilon \mu(D) \pm \varepsilon M \\ &= \int_D f(x) dx \pm \sum_{C \in Z', C \cap D \neq \emptyset, C \not\subseteq D} \int_C f(x) dx \pm 2\varepsilon \mu(D) \pm \varepsilon M \\ &= \int_D f(x) dx \pm \sum_{C \in Z', C \cap D \neq \emptyset, C \not\subseteq D} M \cdot \mu(C) \pm 2\varepsilon \mu(D) \pm \varepsilon M \\ &= \int_D f(x) dx \pm \varepsilon M \pm 2\varepsilon \mu(D) \pm \varepsilon M = \int_D f(x) dx \pm \eta \cdot \mu(D) \end{aligned}$$

Das gilt für alle $\eta > 0$. Also $W(D) = \int_D f(x) dx$. \square

24.4 Ordinatenmengen und Normalbereiche

24.4.1 Ordinatenmengen

Die *Ordinatenmenge* $M(f)$ einer Funktion $f : \mathbb{R}^n \supseteq B \rightarrow [0, \infty)$ ist die Menge

$$M(f) = \{(x, y) \in \mathbb{R}^n \times \mathbb{R} : x \in B, 0 \leq y \leq f(x)\},$$

d.h. die Menge aller Punkte "zwischen der x -Achse und dem Graphen der Funktion".

Satz 24.14 *Sei $B \subseteq \mathbb{R}^n$ ein messbar, $f : B \rightarrow \mathbb{R}$ beschränkt und $f \geq 0$. Dann ist $M(f) \subseteq \mathbb{R}^{n+1}$ Jordan-messbar genau dann, wenn f Riemann-integrierbar ist und es gilt*

$$\mu(M(f)) = \int_B f(x) dx.$$

In diesem Fall ist der Graph $\{(x, f(x)) \mid x \in B\}$ von f eine Nullmenge.

Für eine Treppenfunktion f schreiben wir wieder $f(C)$, falls das der konstante Wert auf C ist.

Beweis. Sei $f \leq b$. Sei $M(f)$ messbar. Dann gibt es passende Gitterzerlegungen von der Form $Z_n \times Z'_n$, wobei Z_n durch Einschränkung auf B eine Zerlegung Z_{nB} ergibt und $Z'_n : 0 = z_0 < z_1 < \dots < z_{m_n} = b$ Zerlegung von $[0, b]$. Wir definieren Treppenfunktionen $\underline{f}_n \leq f \leq \bar{f}_n$ bzgl. Z_{nB} mit

$$\underline{f}_n(C) = \max\{z_k \mid C \times [0, z_k] \subseteq M(f)\}, \quad \bar{f}_n(C) = \max\{z_k \mid M(f) \cap (C \times \mathbb{R}) \subseteq C \times [0, z_k]\}$$

Für die Zerlegung $\tilde{Z}_n = Z_{nB} \times Z'_n$ von $M(f)$ gilt dann

$$U(\tilde{Z}_n, M(f)) \leq \int \underline{f}_n \leq \int \bar{f}_n \leq O(\tilde{Z}_n, M(f))$$

also f integrierbar und $\int_b f = \mu(M(f))$.

Sei umgekehrt f integrierbar, belegt durch Treppenfunktionen \underline{f}_n und \bar{f}_n bzgl. Zerlegung $Z_n = Z_{nB}$. Dann gibt es Zerlegungen \tilde{Z}_n von $M(f)$ mit Zellen

$$C \times [0, \underline{f}_n(C)] \quad \text{und} \quad C \times [\underline{f}_n(C), \bar{f}_n(C)] \quad (C \in Z_n), \quad C \times [\bar{f}_n(C), b]$$

und es gilt

$$\int \underline{f}_n \leq U(\tilde{Z}_n, M(f)) \leq O(\tilde{Z}_n, M(f)) \leq \int \bar{f}_n$$

Es folgt die Messbarkeit von $M(f)$. Die Zellen zweiter Art machen den Unterschied zwischen O und U aus und ihre Vereinigung umfasst den Graphen - also ist der eine Nullmenge. \square

Sind allgemeiner $f_1, f_2 : B \rightarrow \mathbb{R}$ Funktionen mit $f_1(x) \leq f_2(x)$ für alle $x \in B$, so definieren wir

$$M(f_1, f_2) := \{(x, y) \in \mathbb{R}^n \times \mathbb{R} : x \in B, f_1(x) \leq y \leq f_2(x)\}$$

(24.23)

Korollar 24.15 *Ist B messbar und sind f_1, f_2 Riemann integrierbar, so ist $M(f_1, f_2)$ Jordan-messbar und es gilt*

$$\mu(M(f_1, f_2)) = \int_B (f_2(x) - f_1(x)) dx.$$

Beweis. O.B.d.A. $0 \leq f_1$. Dann $M(f_1, f_2) = M(f_2) \setminus M(f_1)$. \square

Beispiel 4 Sei $(x_0, y_0) \in \mathbb{R}^2$, $r > 0$ und $B = [x_0 - r, x_0 + r]$. Für

$$f_1(x) = y_0 - \sqrt{r^2 - (x - x_0)^2}, \quad f_2(x) = y_0 + \sqrt{r^2 - (x - x_0)^2}$$

ist $M(f_1, f_2)$ gerade die Kreisscheibe mit dem Mittelpunkt (x_0, y_0) und dem Radius r , und für ihren Jordan-Inhalt finden wir

$$\begin{aligned} \mu(M(f_1, f_2)) &= 2 \int_{x_0-r}^{x_0+r} \sqrt{r^2 - (x - x_0)^2} dx = 2 \int_{-r}^r \sqrt{r^2 - t^2} dt \\ &= 2 \left[\frac{1}{2} t \sqrt{r^2 - t^2} + \frac{1}{2} r^2 \arcsin \frac{t}{r} \right]_{-r}^r = \pi r^2. \quad \blacksquare \end{aligned}$$

24.4.2 Relative Normalbereiche

Sei $B \subseteq \mathbb{R}^n$ und seien $h_1, h_2 : B \rightarrow \mathbb{R}$ stetig mit $h_1(x) \leq h_2(x)$ für alle $x \in B$. Dann ist

$$D = \{(x, y) \mid x \in B, y \in \mathbb{R}, h_1(x) \leq y \leq h_2(x)\}$$

ein *relativer Normalbereich* in \mathbb{R}^{n+1} über B . Auf die Anordnung der Variablen kommt es dabei nicht an. Aus Kor.29.3 folgt

Korollar 24.16 *Ist B messbar, so auch D und*

$$\mu(D) = \int_B (h_2(x) - h_1(x)) dx$$

Satz 24.17 *Sei D relativer Normalbereich über messbarem B und und $f : D \rightarrow \mathbb{R}$ integrierbar. Für alle $x \in B$ existiere*

$$g : B \rightarrow \mathbb{R}, \quad g(x) = \int_{h_1(x)}^{h_2(x)} f(x, y) dy$$

$$\int_D f(x, y) d(x, y) = \int_B g(x) dx$$

Beweis. Sei B im Intervall I enthalten und $c \leq h_1 \leq h_2 \leq d$. Sei $R = I \times [c, d]$. Bezeichne \hat{f} die Fortsetzung von f auf R und \hat{g} die Fortsetzung von g auf I , also

$$\int_c^d \hat{f}(x, y) dy = \hat{g}(x)$$

Nach Fubini folgt

$$\int_D f = \int_R \hat{f} = \int_I \left(\int_c^d \hat{f}(x, y) dy \right) dx = \int_I \hat{g}(x) dx = \int_B g(x) dx \quad \square$$

Satz 24.18 *Ist D relativer Normalbereich über B und B wegzusammenhängend bzw. kompakt, so ist es auch D .*

Sei B wegzusammenhängend. Zwischen $(x, y), (x', y') \in D$ hat man den Weg bestehend aus der Strecke von (x, y) nach $(x, h_1(x))$ dem Weg $(t, h_1(x(t))) t \in [u, v]$ von $(x, h_1(x))$ nach $(x', h_1(x'))$ und der Strecke von $(x', h_1(x'))$ nach (x', y') wobei $x(t) t \in [u, v]$ ein Weg in B von x nach x' ist.

Sei B kompakt. Mit

$$a = \min\{h_1(x) \mid x \in B\}, \quad b = \max\{h_2(x) \mid x \in B\}$$

gilt $D \subseteq [a, b] \times B$, also D beschränkt. Sei $(x_n, y_n) \in D$ und $(x, y) = \lim_{n \rightarrow \infty} (x_n, y_n)$. Dann $x \in B$ und wegen Stetigkeit

$$h_1(x) = \lim_{n \rightarrow \infty} h_1(x_n) \leq \lim_{n \rightarrow \infty} y_n \leq \lim_{n \rightarrow \infty} h_2(x_n) = h_2(x)$$

also D abgeschlossen. \square

24.4.3 Normalbereiche

Wir definieren nun induktiv: Die *Normalbereiche* in \mathbb{R} sind die Intervalle $[a, b]$. Ist $B \subseteq \mathbb{R}^n$ ein Normalbereich und $D \subseteq \mathbb{R}^n$ ein relativer Normalbereich über B , so ist D ein *Normalbereich* in \mathbb{R}^{n+1} .

Für \mathbb{R}^2 haben wir demnach 2 Typen von Normalbereichen. Unter einem *Normalbereich bzgl. der x -Achse* versteht man eine Menge B der Gestalt

$$B = \{(x, y) \in \mathbb{R}^2 : a \leq x \leq b, g_1(x) \leq y \leq g_2(x)\}$$

mit stetigen Funktionen $g_1, g_2 : [a, b] \rightarrow \mathbb{R}$, wobei $g_1(x) \leq g_2(x)$ für alle $x \in [a, b]$. Eine solche Menge heißt auch *y -projizierbar*. Ein *Normalbereich bzgl. der y -Achse* (oder eine *x -projizierbare Menge*) ist eine Menge der Gestalt

$$B = \{(x, y) \in \mathbb{R}^2 : c \leq y \leq d, h_1(y) \leq x \leq h_2(y)\}$$

mit stetigen Funktionen $h_1, h_2 : [c, d] \rightarrow \mathbb{R}$, wobei $h_1(y) \leq h_2(y)$ für alle $y \in [c, d]$. Fig.10

Korollar 24.19 *Sei B ein Normalbereich und $f : B \rightarrow \mathbb{R}$ stetig. In den obigen Bezeichnungen gilt*

$$\int_B f(x, y) d(x, y) = \int_a^b \left(\int_{g_1(x)}^{g_2(x)} f(x, y) dy \right) dx,$$

falls B ein Normalbereich bzgl. der x -Achse ist und

$$\int_B f(x, y) d(x, y) = \int_c^d \left(\int_{h_1(y)}^{h_2(y)} f(x, y) dx \right) dy,$$

falls B ein Normalbereich bzgl. der y -Achse ist.

Ist B kein Normalbereich, so versucht man, ihn in endlich viele nicht überlappende Normalbereiche zu zerlegen.

In \mathbb{R}^3 erhalten wir Normalbereiche z.B. so: Sei $B_1 := [a, b]$. Auf B_1 seien g_1, g_2 stetige Funktionen mit $g_1 \leq g_2$, und sei

$$B_2 := \{(x, y) \in \mathbb{R}^2 : x \in B_1, g_1(x) \leq y \leq g_2(x)\}.$$

Auf B_2 seien h_1, h_2 stetige Funktionen mit $h_1 \leq h_2$, und sei

$$B_3 := \{(x, y, z) \in \mathbb{R}^3 : (x, y) \in B_2, h_1(x, y) \leq z \leq h_2(x, y)\}.$$

Dann ist B_3 ein Normalbereich und für jede auf B_3 stetige Funktion f ist

$$\int_{B_3} f(x, y, z) d(x, y, z) = \int_a^b \int_{g_1(x)}^{g_2(x)} \int_{h_1(x, y)}^{h_2(x, y)} f(x, y, z) dz dy dx.$$

24.4.4 Beispiele - nur zur Illustration der Methode

Sinnvoller arbeitet man mit zur Geometrie passenden Zerlegungen.

Beispiel 5 Das Volumen V des Tetraeders mit den Ecken $(0, 0, 0)$, $(a, 0, 0)$, $(0, b, 0)$ und $(0, 0, c)$ läßt sich bestimmen als Volumen der Punktmenge zwischen dem Dreieck in der (x, y) -Ebene mit den Ecken $(0, 0, 0)$, $(a, 0, 0)$ und $(0, b, 0)$ und der durch $\frac{x}{a} + \frac{y}{b} + \frac{z}{c} = 1$, also $z = f(x, y) = c(1 - \frac{x}{a} - \frac{y}{b})$, beschriebenen Ebene. Daher ist

$$\begin{aligned} V &= \int_B f(x, y) d(x, y) = \int_0^a \int_0^{b - \frac{b}{a}x} c(1 - \frac{x}{a} - \frac{y}{b}) dy dx \\ &= \int_0^a c(y - \frac{xy}{a} - \frac{y^2}{2b}) \Big|_0^{b - \frac{b}{a}x} dx = \int_0^a \frac{bc}{2} (1 - \frac{x}{a})^2 dx = \frac{abc}{6}. \end{aligned}$$

Beispiel 6 Sei $B = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \leq r^2, y \geq 0\}$ mit fest gewähltem $r > 0$, und sei $f(x, y) = x^2 y$ auf B . Der Halbkreis B ist ein Normalbereich sowohl bzgl. der x -Achse als auch bzgl. der y -Achse. Wir beschreiben ihn z.B. als

$$B = \{(x, y) \in \mathbb{R}^2 : x \in [-r, r], 0 \leq y \leq \sqrt{r^2 - x^2}\}$$

und erhalten

$$\begin{aligned} \int_B x^2 y d(x, y) &= \int_{-r}^r \int_0^{\sqrt{r^2 - x^2}} x^2 y dy dx \\ &= \int_{-r}^r \frac{x^2 y^2}{2} \Big|_0^{\sqrt{r^2 - x^2}} dx = \frac{1}{2} \int_{-r}^r x^2 (r^2 - x^2) dx \\ &= \frac{1}{2} \left(r^2 \frac{x^3}{3} - \frac{x^5}{5} \right) \Big|_{-r}^r = \frac{2}{15} r^5. \end{aligned}$$

Beispiel 7 Ein gerader Kreiskegel mit Radius R und Höhe h ist ein Normalbereich in \mathbb{R}^3 . Wir haben nämlich $B_1 = [-R, R]$,

$$B_2 = \{(x, y) \in \mathbb{R}^2 : x \in B_1, -\sqrt{R^2 - x^2} \leq y \leq \sqrt{R^2 - x^2}\} \\ (= \text{Grundfläche des Kegels}),$$

$$B_3 = \{(x, y, z) \in \mathbb{R}^3 : (x, y) \in B_2, 0 \leq z \leq h - \frac{h}{R} \sqrt{x^2 + y^2}\}.$$

Für das Volumen V des Kegels finden wir daher

$$\begin{aligned} V &= \int_{B_3} d(x, y, z) = \int_{-R}^R \int_{-\sqrt{R^2-x^2}}^{\sqrt{R^2-x^2}} \int_0^{h-\frac{h}{R}\sqrt{x^2+y^2}} dz dy dx \\ &= \int_{-R}^R \int_{-\sqrt{R^2-x^2}}^{\sqrt{R^2-x^2}} \left(h - \frac{h}{R} \sqrt{x^2 + y^2}\right) dy dx \\ &= h \int_{-R}^R \left(y - \frac{1}{R} \left(\frac{y}{2} \sqrt{x^2 + y^2} + \frac{x^2}{2} \ln(y + \sqrt{x^2 + y^2})\right)\right) \Big|_{-\sqrt{R^2-x^2}}^{\sqrt{R^2-x^2}} dx \\ &= 2h \int_0^R \left(\sqrt{R^2 - x^2} - \frac{x^2}{2R} \ln \frac{R + \sqrt{R^2 - x^2}}{R - \sqrt{R^2 - x^2}}\right) dx. \end{aligned}$$

Wir substituieren $x = R \sin t$ und erhalten

$$\begin{aligned} V &= 2h \int_0^{\pi/2} \left(R \cos t - \frac{R}{2} \sin^2 t \ln \frac{1 + \cos t}{1 - \cos t}\right) R \cos t dt \\ &= 2hR^2 \int_0^{\pi/2} \left(\cos^2 t - \frac{1}{2} \sin^2 t \cos t \ln \frac{1 + \cos t}{1 - \cos t}\right) dt. \end{aligned}$$

Das Integral $\int \sin^2 t \cos t \ln \frac{1+\cos t}{1-\cos t} dt$ kann durch partielle Integration bestimmt werden: der Faktor $\sin^2 t \cos t$ wird integriert und hat $\frac{1}{3} \sin^3 t$ als eine Stammfunktion, und $\ln \frac{1+\cos t}{1-\cos t}$ wird differenziert und ergibt $\frac{-2}{\sin t}$. Eingesetzt erhält man schließlich

$$V = 2hR^2 \left(\frac{\pi}{4} - \frac{1}{6} \sin^3 t \ln \frac{1 + \cos t}{1 - \cos t} \Big|_0^{\pi/2} - \frac{2}{6} \int_0^{\pi/2} \sin^2 t dt\right),$$

also

$$V = \frac{1}{3} \pi h R^2.$$

25 Transformation von Maßen und Integralen

25.1 Unabhängigkeit des Jordan-Maßes

25.1.1 Vorbemerkung

Die Grundlage zur Behandlung des Raumes und physikalischer Größen ist der Begriff des euklidischen Raumes, eines affinen Raums mit zugehörigem Vektorraum, ausgestattet mit einer Längenmessung bzw. Skalarprodukt. Für diese kann man das Konzept einer Orthonormalbasis bzw. kartesischen Koordinatensystems einführen. Ist ein solches gegeben,

kann man Vektoren bzw. Punkte durch Koordinatenspalten in \mathbb{R}^n bezeichnen - also den Raum auf diese Weise mit \mathbb{R}^n jeweils identifizieren. Dabei sollte man aber nicht vergessen, dass diese Koordinatisierung sekundär ist, dass primär die von Koordinaten unabhängige vektorielle und metrische Struktur gegeben ist.

\mathbb{R}^3 kommt heraus, wenn man den Raum mithilfe der Koordinaten bzgl. eines kartesischen Koordinatensystems beschreibt. Es gibt aber unendlich viele solche Koordinatensysteme, also unendlich viele gleichberechtigte Möglichkeiten, den Raum als \mathbb{R}^3 aufzufassen. Berechnet man Längen bzw. Skalarprodukte über solche Koordinaten, so hängt das Ergebnis nicht vom Koordinatensystem ab, weil der Längenbegriff den Koordinaten vorausgeht.

Um Flächen bzw. Volumen zu messen, haben wir Zerlegungen benutzt, die zwei wesentlichen Anforderungen genügen mussten: Das Maß einer Zelle ergibt sich unmittelbar aus zugehörigen Längenmaßen; je zwei Zerlegungen haben eine gemeinsame Verfeinerung. Die zweite Bedingung war der Grund dafür, ein kartesisches Koordinatensystem α auszuwählen und das Jordan-Maß μ_α nur über achsenparallele Rechteck-Zerlegungen bzgl. dieses einen Koordinatensystems zu definieren. Allerdings hätten wir für Intervalle $I = [a, b]$ die Vorgabe $\prod_i (b_i - a_i)$ für das Volumen als $\mu_0(I)$ notieren sollen und uns dann überlegen, dass dann I im Sinne unserer Definition messbar ist und $\mu_\alpha(I) = \mu_0(I)$. Das ist aber klar, weil für jede Gitter-Zerlegung Z von $B = I \subseteq I$ gilt: $\sum_{J \in Z} \mu_0(J) = \mu_0(I)$.

Darauf aufbauend, haben wir allgemeinere Zerlegungen definiert, deren Zellen bzgl. des eingeführten Maßes messbar sind, und gezeigt, dass diese dann zum gleichen Maß führen. Um das anwenden zu können, müssen wir jedoch die Messbarkeit und das Maß geeigneter Zellen herleiten, insbesondere für Rechtecke bzw. Quader in beliebiger Lage. Für diese sollte sich jedoch das Maß als Produkt der Kantenlängen ergeben - d.h. das Maß unabhängig vom gewählten kartesischen Koordinatensystem sein. Wir haben also Folgendes zu zeigen.

25.1.2 Unabhängigkeit vom Koordinatensystem

Satz 25.1 *Sind α und β kartesische Koordinatensysteme eines endlichdimensionalen euklidischen affinen Raumes R mit zugehörigen Jordan-Maßen μ_α bzw. μ_β , so ist eine Punktmenge B genau dann bzgl. μ_α messbar, wenn sie bzgl. μ_β messbar ist und zwar $\mu_\alpha(B) = \mu_\beta(B)$.*

25.1.3 Eulers Dreh

Lemma 25.2 *Ist $\dim R = n \geq 3$ so gibt es zu je zwei ON-Basen α, β von R gibt es ON-Basen β_0, \dots, β_k mit $\alpha = \beta_0$, $\beta = \beta_k$ und so, dass β_{i+1} und β_i eine Achse gemeinsam haben.*

Für $n = 3$ sieht man das so: Nach Spiegelung an einer Koordinatenebene kann man annehmen, dass α und β gleichorientiert sind. Haben α und β einen Vektor gemeinsam, so drehe man das 'Dreibein' α um diesen, bis die beiden anderen Beine auf die von β zu liegen kommen. Um auf ein gemeinsames Bein zu kommen, wähle man ein Bein von α und drehe α um eine seiner anderen Achsen, bis das Bein in eine der Koordinatenebenen

von β zu liegen kommt. Danach drehe man das Bein in dieser Ebene, bis es auf einem Bein von β liegt.

25.1.4 Beweis des Satzes

Beweis. Es ist zu zeigen: Ist I Intervall bzgl. α , so ist I auch bzgl. β messbar und $\mu_\beta(I) = \mu_\alpha(I)$. Das ist offensichtlich, wenn sich α und β nur bzgl. des Ursprungs unterscheiden. Für $n = 2$ haben wir das schon bewiesen (vgl. Folien). Nun geht es mit Induktion weiter.

Wir zeigen, dass sich für jedes kartesische Koordinatensystem α ein n -dimensionalen Quader Q messbar ist mit Maß $\mu_\alpha(Q)$ das sich aus den Seitenlängen a_1, \dots, a_n unabhängig von der Lage ergibt als

$$(+)\quad \mu_\alpha(Q) = a_1 \cdot \dots \cdot a_n$$

Dabei spielt die Wahl des Ursprungs offenbar keine Rolle, es geht also nur um die Orthonormalbasen. Der Fall $n = 1$ ist trivial und $n = 2$ haben wir schon behandelt (siehe Folien): Für ein Dreieck D mit zwei zueinander senkrechten Seiten der Längen a und b gilt $\mu_\alpha(D) = \frac{1}{2}ab$ unabhängig vom Koordinatensystem. Man kann D auch in dreieckige Normalbereiche zerlegen und so $\mu_\alpha(D) = \frac{1}{2}ab$ herleiten. Um $\mu_\alpha(Q) = ab$ für ein Rechteck der Seitenlängen a, b zu zeigen, zerlege man dieses durch die Diagonale in zwei rechtwinklige Dreiecke.

Sei nun $n = 3$. Ist eine Kante des Quaders Q achsenparallel, o.B.d.A. a_3 in z -Richtung, so kann Q als Ordinatenmenge

$$\{(x, y, z) \mid (x, y) \in B, 0 \leq z \leq a_3\}$$

aufgefasst werden und nach dem Fall $n = 2$ hat man $\mu(B) = a_1 a_2$ also

$$\mu(Q) = \int_B a_3 d(x, y) = a_3 \mu(B) = a_1 a_2 a_3$$

(+) folgt nun zunächst für den Fall, dass β aus α durch Drehung um eine Koordinatenachse oder Spiegelung an einer Koordinatenebene von α hervorgeht, nach Euler dann auch allgemein. Für $n > 3$ geht's ähnlich weiter.

Damit ist (+) gezeigt. Seien nun α und β kartesische Koordinatensysteme und B bzgl. α messbar. Um B bzgl. β zu messen, betrachten wir Gitter-Zerlegungen von B bzgl. β , d.h. mit Zellen Q , die zu β achsenparallele Quader sind. Nach (+) sind die Q bzgl. α messbar und $\mu_\alpha(Q) = \mu_\beta(Q)$. Also bilden die Q eine (krumme) Zerlegung bzgl. α und nach Satz 24.6 ist B bzgl. α messbar und $\mu_\alpha(B) = \mu_\beta(B)$. \square

25.1.5 Polyeder

Lemma 25.3 *Konvexe Polyeder sind messbar.*

Beweis. Jeder konvexe Polyeder lässt sich in Simplices zerlegen. Also genügt es, diese zu betrachten. Im Falle $n = 2$ haben wir ein Dreieck und nichts zu tun. Für $n = 3$ haben wir einen Tetraeder T mit Ecken p_1, p_2, p_3, p_4 . Wähle das Koordinatensystem so, dass p_1, p_2, p_3 in einer Koordinatenebene E liegen. Sei q der Fußpunkt des Lotes auf E . Liegt q in Dreieck D mit Ecken p_1, p_2, p_3 , so ist T ein Normalbereich, also messbar. Andernfalls

ist wird das von p_1, p_2, p_3, q aufgespannte konvexe Polygon z.B. durch die Strecke p_2, p_3 in die Dreiecke D und p_2, p_3, q zerlegt und durch T und den Tetraeder T_1 mit Ecken p_2, p_3, p_4, q erhalten wir eine Zerlegung der Pyramide T_2 mit Ecken p_1, p_2, p_3, p_4, q . Diese wird aber andererseits durch die Tetraeder T_3 mit Ecken p_1, p_2, p_4, q und T_4 mit Ecken p_1, p_3, p_4, q zerlegt. T_1, T_3 und T_4 sind nach dem schon behandelten Fall messbar, also auch T_2 und $T_2 \setminus T_1$. Weiter so für $n > 3$. \square

25.2 Affine Abbildungen

Eine äquivalente Sichtweise ist, dass Messbarkeit und Maß einer Menge unter längenerhaltender Abbildung erhalten bleiben. Allgemeiner studieren wir Abbildungen σ , für die $\mu(\sigma(B))$ aus $\mu(B)$ bestimmt werden kann und wenden das auf die Berechnung von Integralen an.

25.2.1 Affine Abbildungen und ihre Determinanten

Eine *affine Abbildung* $\phi : V \rightarrow V$ ist von der Form $\phi(\vec{x}) = \phi_0(\vec{x}) + \vec{c}$ mit einer linearen Abbildung ϕ_0 und konstanten $\vec{c} \in V$, also bezgl. einer Basis α

$$\phi(\vec{x})^\alpha = A\vec{x}^\alpha + \vec{c}^\alpha$$

(wir fassen hier V als affinen Raum mit Ursprung $\vec{0}$ auf). Wir definieren

$$\det(\phi) = \det A$$

was nach dem Produktsatz von der Basis α unabhängig ist. In der Tat $\det({}^\beta T_\alpha A {}^\alpha T_\beta) = \det A$ da ${}^\beta T_\alpha = {}^\alpha T_\beta^{-1}$.

Lemma 25.4 *Affine Abbildungen sind stetig. Eine affine Abbildung ϕ ist umkehrbar genau dann, wenn $\det \phi \neq 0$. Die Umkehrung ϕ^{-1} ist dann auch affin und es gilt $\det(\phi^{-1}) = (\det \phi)^{-1}$.*

Beweis. Die Stetigkeit folgt aus der Matrixbeschreibung. Die Umkehrbarkeit von ϕ ist gleichbedeutend zur Invertierbarkeit von A - und die Inverse hat Matrix A^{-1} . \square

25.2.2 Bewegungen

Eine affine Abbildung mit ϕ_0 d.h. A orthogonal, so handelt es sich um eine *Bewegung*, d.h. es gilt

$$|\phi(x) - \phi(y)| = |x - y| \quad \text{für alle } x, y \in \mathbb{R}^n$$

Nämlich $|Ax + c - (Ay + c)| = |A(x - y)| = |x - y| = |x + c - (y + c)|$. Umgekehrt folgt aus dieser Bedingung, dass ϕ in dieser Form beschrieben werden kann. ϕ hat eine Umkehrabbildung ϕ^{-1} und diese ist ebenfalls eine Bewegung. Hier $|\det \phi| = \pm 1$.

Ist $\phi_0 = \text{id}$ die identische Abbildung, also $\phi(\vec{x}) = \vec{x} + \vec{c}$ so ist ϕ eine *Verschiebung* oder *Translation*.

25.3 Maßtransformation bei affinen Abbildungen

25.3.1 Bewegungsinvarianz des Jordan-Maßes

Ist Z Zerlegung einer Menge $B \subseteq \mathbb{R}^n$ und $\sigma : \mathbb{R}^n \rightarrow \mathbb{R}^n$ eine Abbildung, so sei

$$\sigma(Z) = \{\sigma(C) \mid C \in Z\}$$

Korollar 25.5 *Ist σ eine Bewegung von \mathbb{R}^n und $B \subseteq \mathbb{R}^n$ messbar, so ist auch $\sigma(B)$ messbar und $\mu(\sigma(B)) = \mu(B)$.*

Beweis. Ist μ bzgl. der ON-Basis α definiert, so sei β das Bild von α unter σ . Ist nun Z_n eine Folge von Gitterzerlegungen B bzgl. α mit Weite $\rightarrow 0$, so ist $\sigma(Z_n)$ eine Folge von Gitterzerlegungen von $\sigma(B)$ mit Weite $\rightarrow 0$ und $\mu(\sigma(C)) = \mu(C)$ für $C \in Z_n$, da C und $\sigma(C)$ Quader mit gleichen Kantenlängen sind. Also $\mu_\beta(\sigma(B)) = \mu_\alpha(B)$ und die Behauptung folgt aus der Basisunabhängigkeit des Jordan-Maßes. \square

25.3.2 Volumen eines Spats

Lemma 25.6 *Ist S das von den unabhängigen Vektoren $\vec{a}_1, \dots, \vec{a}_n$ aufgespannte Parallelepiped (Parallelogramm bzw. Spat) und A die Matrix, deren Spalten die Koordinaten der \vec{a}_j sind, so gilt*

$$\mu(S) = \det A$$

Beweis. Wir zeigen

- Entsteht S' aus S durch Vertauschen zweier Vektoren, so gilt $S' = S$.
- Geht S' aus S durch Scherung hervor und ist S' messbar, so ist auch S messbar und es gilt $\mu(S) = \mu(S')$.

Das Erste ist Definition. Beim Zweitem wird in S' ein \vec{a}_i ersetzt durch $\vec{a}'_i = \vec{a}_i + r\vec{a}_j$ mit einem $j \neq i$. O.B.d.A $\vec{a}'_2 = \vec{a}_2 + r\vec{a}_1$, $0 \leq r \leq |\vec{a}_1|$ und Winkel zwischen \vec{a}_1 und \vec{a}'_2 höchstens 90° . Dann wird das Dreieck $D O, r\vec{a}_1, \vec{a}'_2$ bzw. (schiefe) Prisma mit 3 parallelen Kanten in Richtung \vec{a}_3 von S abgeschnitten und nach der Verschiebung τ um \vec{a}_1 auf der anderen Seite wieder angesetzt und S' zu erhalten. Also

$$S = (S' \setminus \tau(D)) \cup D$$

D ist nach Lemma 25.3 messbar, $\mu(D) = \mu(\tau(D))$ und daher

$$\mu(S) = (\mu(S') - \mu(\tau(D))) + \mu(D) = \mu(S')$$

Ist nun A die Matrix, deren Koordinaten die Spalten der \vec{a}_i sind, so ist $\det A \neq 0$ und A kann durch Spalten-Scherungen und Vertauschungen in Diagonalgestalt A' überführt werden mit $|\det A| = |\det A'|$. A' beschreibt dann einen achsenparallelen Quader Q mit Volumen

$$\mu(Q) = a'_{11} \cdot \dots \cdot a'_{nn} = \det A' = \det A$$

Führt man rückwärt diese Scherungen und Vertauschungen von Q ausgehend aus, so erhält man S , die Messbarkeit von S und $\mu(S) = \mu(Q)$. \square

25.3.3 Affine Maßtransformation

Satz 25.7 Ist $\sigma : \mathbb{R}^n \rightarrow \mathbb{R}^n$ eine affine Abbildung mit $\det \sigma \neq 0$, so ist B genau dann messbar, wenn $\sigma(B)$ messbar ist und es gilt

$$\mu(\sigma(B)) = \mu(B) |\det \sigma|$$

Dies kann man auch so interpretieren: misst man bzgl. des schiefwinkligen, durch die Vektoren $\vec{v}_1, \dots, \vec{v}_n$ gegebenen Koordinatensystems β , so gilt

$$\mu_\beta(B) = \mu(B) \det(\vec{v}_1, \dots, \vec{v}_n)$$

Beweis. Wegen $\det \sigma \neq 0$ ist Bild $S = \sigma(C)$ eines Rechtecks bzw. Quaders ein Parallelogramm bzw. Spat der durch unabhängige Vektoren gegeben ist. Nach obigem Lemma ist $\sigma(C)$ messbar mit $\mu(\sigma(C)) = \mu(C) |\det \sigma|$. Sind nun die Z_n Gitter-Zerlegungen von B mit Weite $\rightarrow 0$, so sind die $\sigma(Z_n)$ Zerlegungen von $\sigma(B)$ mit Weite $\rightarrow 0$ und bei der Bestimmung von $\mu(\sigma(B))$ bzgl. dieser Zerlegungen haben wir jeweils den Faktor $|\det \sigma|$. \square

25.4 Integralsubstitution

Sei $B \subseteq \mathbb{R}^d$ messbar und kompakt und Z_n eine Folge von Zerlegungen von B mit Weite $\rightarrow 0$. Seien $\sigma : B \rightarrow \mathbb{R}^d$ Abbildung, $\sigma(u) = x$, und $\tau : B \rightarrow \mathbb{R}$ Abbildung, $\tau(u) \in \mathbb{R}$, so, dass gilt

- (i) $\sigma(B)$ ist messbar und kompakt
- (ii) $\sigma : B \rightarrow \mathbb{R}^d$ und $\tau : B \rightarrow \mathbb{R}$ sind stetig
- (iii) $\sigma(C) \cap \sigma(C')$ ist Nullmenge für alle $C \neq C'$ in Z_n , alle n - das ist in (iv) enthalten
- (iv) Die $\sigma(Z_n)$ bilden Zerlegungen von $\sigma(B)$ mit Weite $\rightarrow 0$ und $O(\sigma(Z_n), \sigma(B)) - U(\sigma(Z_n), \sigma(B)) \rightarrow 0$
- (v) Zu jedem n und $C \in Z_n$ gibt es ξ_{nC} mit $\mu(\sigma(C)) = \tau(\xi_{nC}) \mu(C)$

Dann bestimmen σ, τ eine *Substitution* für B .

Satz 25.8 Ist B messbar und kompakt, bestimmen σ, τ eine Substitution bzgl. passender Z_n und ist $f : \sigma(B) \rightarrow \mathbb{R}$ stetig, so gilt

$$\int_B f(\sigma(u)) \cdot \tau(u) \, du = \int_{\sigma(B)} f(x) \, dx$$

Beweis. Mit $\xi'_{nD} = \sigma(\xi_{nC})$

$$\begin{aligned} \int_B f(\sigma(u)) \cdot \tau(u) \, du &= \lim_{n \rightarrow \infty} \sum_{C \in Z_n, C \subseteq B} f(\sigma(\xi_{nC})) \tau(\xi_{nC}) \mu(C) \\ &= \lim_{n \rightarrow \infty} \sum_{D \in \sigma(Z_n), D \subseteq \sigma(B)} f(\xi'_{nD}) \mu(D) = \int_{\sigma(B)} f(x) \, dx \quad \square \end{aligned}$$

Korollar 25.9 Ist σ affin mit $\det \sigma \neq 0$ so erhält man eine Substitution bzgl. beliebiger Gitterzerlegungen Z durch

$$\tau(\xi_C) = |\det \sigma| \quad \text{und es gilt} \quad \int_{\sigma(B)} f(x) dx = |\det \sigma| \int_B f(\sigma(u)) du$$

Bei allgemeineren Substitutionen σ spricht man auch von *krummlinigen Koordinaten*, wobei in der Ebene ein *Koordinatengitter* gegeben ist durch die "Koordinatenlinien"

$$\{\sigma(u, b) \mid u \in \mathbb{R}, (u, b) \in B\}, \quad \{\sigma(a, v) \mid v \in \mathbb{R}, (a, v) \in B\}$$

mit geeigneten a, b und die "Flächenelemente"

$$\{\sigma(u, v) \mid a_1 \leq u \leq a_2, b_1 \leq v \leq b_2\}$$

welche man sich gern "infinitesimal" denkt. Analog im Raum mit "Volumenelementen". Durch die Abbildung $\sigma(u, v) = (x, y)$ wird ein Rechteck-Netz über B in ein „krummliniges“ Netz über $\sigma(B)$ übersetzt:

Fig.11

25.5 Polarkoordinaten

25.5.1 Polarkoordinaten

Der Zusammenhang zwischen den kartesischen Koordinaten (x, y) und den *Polarkoordinaten* (r, φ) eines Punktes im \mathbb{R}^2 ist gegeben durch (Fig.6)

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

25.5.2 Kreisringsegment

Um die Flächenelemente für Polarkoordinaten zu bestimmen betrachten wir zunächst Kreis-sektoren B mit Radius R , Winkel Φ und Zentrum O . Um die Messbarkeit zu zeigen und die Fläche zu bestimmen, teilen wir zu jeder Zweierpotenz n den Sektor in n kongruente Sektoren B_n mit der zugehörigen Sehne der Länge s_n . Sei h_n die Höhe durch O im Sehnendreieck. Wähle eine Zerlegung von B , in der jedes Sehnendreieck dur äquidistante Parallelen zur der Sehne in n Teile zerlegt wird und zusätzlich den gleichschenkligen Trapezen mit der Sehne s_n als einer Seite, der anderen tangential zum Kreis und senkrecht zu h_n , und mit radialen Schenkeln. Also mit Fläche

$$s_n(R - h_n) + \frac{s_n}{2h_n}(R - h_n)^2$$

Wegen $h_n \rightarrow R$ folgt und $ns_n \rightarrow r\Phi$ (Bogenlänge) folgt

$$O(Z_n, B) - U(Z_n, B) = ns_n(r - h_n) + \frac{ns_n}{2h_n}(R - h_n)^2 \rightarrow 0$$

und für die Summe der Sehnendreiecke

$$U(Z_n, B) = nh_n \frac{s_n}{2} = \frac{1}{2} h_n ns_n \rightarrow \frac{1}{2} R^2 \Phi = \mu(B)$$

Für das Kreisringsektor

$$B = \{(r \cos \phi, r \sin \phi) \mid r_1 \leq r \leq r_2, \phi_1 \leq \phi \leq \phi_2\}$$

folgt

$$\mu(B) = \frac{r_1 + r_2}{2} \Delta r \Delta \phi \quad \text{mit } \Delta r = r_2 - r_1, \Delta \phi = \phi_2 - \phi_1$$

In der Tat, B ist die Differenz zweier Kreissektoren, also $\mu(B) = \frac{1}{2}r_2^2\Delta\pi\phi - \frac{1}{2}r_1^2\Delta\pi\phi = \frac{1}{2}(r_1 + r_2)(r_2 - r_1)\Delta\phi$.

25.5.3 Polarsubstitution

Auf Polarkoordinaten zu transformieren, heißt also, die Substitution

$$(x, y) = \sigma(r, \varphi) = (r \cos \varphi, r \sin \varphi) \quad r \geq 0, \quad \varphi_1 \leq \varphi \leq \varphi_2$$

mit $\varphi_2 - \varphi_1 \leq 2\pi$ vorzunehmen. Ist beispielsweise

$$B = \{(r, \varphi) \in \mathbb{R}^2 : r_1 < r < r_2, \varphi_1 < \varphi < \varphi_2\}$$

ein Rechteck, so ist $\sigma(B)$ ein Teil eines Kreisringes. Fig. 7.

$\sigma(B)$ ist Jordan-messbar und es gilt mit $\Delta r = r_2 - r_1, \Delta \varphi = \varphi_2 - \varphi_1, r = \frac{1}{2}(r_1 + r_2)$

$$\mu(\sigma(B)) = \frac{1}{2}(r_2^2 - r_1^2)(\varphi_2 - \varphi_1) = r \cdot \Delta r \cdot \Delta \phi$$

Korollar 25.10 *Ist σ durch die Polarkoordinaten gegeben und $\tau(r, \phi) = r$, so erhält man eine Substitution. Insbesondere gilt für messbare kompakte B im Definitionsbereich von σ und stetige $f : \sigma(B) \rightarrow \mathbb{R}$*

$$\int_{\sigma(B)} f(u) \, du = \int_B f(r \cos \phi, r \sin \phi) \cdot r \, d(r, \phi)$$

Ist $B = \{(r, \phi) \mid r_1(\phi) \leq r \leq r_2(\phi), \phi_1 \leq \phi \leq \phi_2\}$ ein Normalbereich bzgl. der ϕ -Achse so gilt

$$\int_{\sigma(B)} f(x) \, dx = \int_{\phi_1}^{\phi_2} \left(\int_{r_1(\phi)}^{r_2(\phi)} f(r \cos \phi, r \sin \phi) \cdot r \, dr \right) d\phi$$

Ist $B = \{(r, \phi) \mid r_1 \leq r \leq r_2, \phi_1(r) \leq \phi \leq \phi_2(r)\}$ ein Normalbereich bzgl. der r -Achse so gilt

$$\int_{\sigma(B)} f(x) \, dx = \int_{r_1}^{r_2} \left(\int_{\phi_1(r)}^{\phi_2(r)} f(r \cos \phi, r \sin \phi) \cdot r \, d\phi \right) dr$$

Beispiel 8 Sei $0 \leq \varphi_1 < \varphi_2 \leq 2\pi$, und die Funktion $\rho : [\varphi_1, \varphi_2] \rightarrow (0, \infty)$ sei stetig. Dann ist der Flächeninhalt von $B = \{(r, \varphi) \in \mathbb{R}^2 : \varphi_1 \leq \varphi \leq \varphi_2, 0 \leq r \leq \rho(\varphi)\}$ gleich (Fig.12)

$$\mu(B) = \int_B d(x, y) = \int_{\varphi_1}^{\varphi_2} \int_0^{\rho(\varphi)} r \, dr \, d\varphi = \int_{\varphi_1}^{\varphi_2} \frac{\rho^2(\varphi)}{2} \, d\varphi.$$

25.6 Zylinderkoordinaten

25.6.1 Zylinderkoordinaten

Die Zylinderkoordinaten (r, φ, z) eines Punktes $P \in \mathbb{R}^3$ mit kartesischen Koordinaten (x, y, z) sind gegeben durch

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

(Fig.8) d.h. die Transformationsfunktion ist

$$(x, y, z) = \sigma(r, \varphi, z) = (r \cos \varphi, r \sin \varphi, z)$$

Ist B ein Quader

$$B = [r_1, r_2] \times [\varphi_1, \varphi_2] \times [z_1, z_2] \quad \text{mit } 0 \leq r_1 < r_2, \varphi_1 < \varphi_2, \varphi_2 - \varphi_1 \leq 2\pi$$

so gilt nach Archimedes mit $\Delta z = z_2 - z_1$ und wie oben $r = \frac{1}{2}(r_2 - r_1)$

$$\mu(\sigma(B)) = \frac{1}{2}(r_2^2 - r_1^2)(\varphi_2 - \varphi_1)(z_2 - z_1) = r \cdot \Delta r \cdot \Delta \varphi \cdot \Delta z$$

25.6.2 Zylindersubstitution

Ist B ein Kreisringsektor so ist

$$D = B \times [z_1, z_2], \quad \Delta z = z_2 - z_1 \geq 0$$

ein Hohlzylindersektor und man erhält

$$\mu(D) = \frac{r_1 + r_2}{2} \Delta r \Delta \phi \Delta z \quad \text{mit } \Delta r = r_2 - r_1, \Delta \phi = \phi_2 - \phi_1$$

indem man von den ebenen Zerlegungen mit polygonalen Zellen C zu räumlichen (Prisma) Zellen $C \times [z_1, z_2]$ übergeht.

Korollar 25.11 *Ist σ durch die Zylinderkoordinaten gegeben und $\tau(r, \phi, z) = r$, so erhält man eine Substitution. Insbesondere gilt für messbare kompakte B im Definitionsbereich von σ und stetige $f : \sigma(B) \rightarrow \mathbb{R}$*

$$\int_{\sigma(B)} f(x) dx = \int_B f(r \cos \phi, r \sin \phi, z) \cdot r d(r, \phi, z)$$

Ist B ein Normalbereich, so kann man die rechte Seite entsprechend durch Dreifachintegrale ausdrücken.

25.6.3 Schwerpunkt

Sei $B \subseteq \mathbb{R}^n$ messbar und kompakt. Die Koordinaten des Schwerpunktes von B sind gegeben durch

$$x_{iB} = \frac{1}{\mu(B)} \int_B x_i dx$$

25.6.4 Rotationskörper

Sei $B \subseteq \{(r, z) \mid r \geq 0, z \in \mathbb{R}\}$ kompakt und messbar mit Schwerpunkt r_B bzgl. der r -Achse. Sei $0 \leq \phi_2 - \phi_1 \leq 2\pi$. Durch Rotation (im Sektor von ϕ_1 bis ϕ_2) der Fläche um die z -Achse entsteht der *Rotationskörper*

$$D = \{(r \cos \phi, r \sin \phi, z) \mid (r, z) \in B\}$$

Dieser ist kompakt und messbar und es gilt

$$\mu(D) = r_B \mu(B) (\phi_2 - \phi_1)$$

also Fläche mal Länge des vom Schwerpunkt der Fläche durchlaufenen Kreisbogens.

Zum Beweis betrachten wir Gitter-Zerlegungen Z_n von B mit Weite $\rightarrow 0$. Zu $C \in Z_n$ bezeichne r_C die (nicht polare!) r -Koordinate des Schwerpunkts, also hat der Schwerpunkt von B die r -Koordinate

$$r_B = \frac{1}{\mu(B)} \int_B r d(r, z) = \frac{1}{\mu(B)} \lim_{n \rightarrow \infty} \sum_{C \in Z_n, C \subseteq B} r_C \mu(C)$$

Für den Hohlzylindersektor

$$C' = \{(r \cos \phi, r \sin \phi, z) \mid (r, z) \in C, \phi_1 \leq \phi \leq \phi_2\}$$

haben wir

$$\mu(C') = r_C \mu(C) (\phi_2 - \phi_1)$$

Dann ist $Z'_n = \{C' \mid C \in Z_n\}$ eine Zerlegung von D und es gilt

$$\begin{aligned} O(Z'_n, D) - U(Z'_n, D) &= \sum_{C \in Z_n, C \not\subseteq B, C \cap B \neq \emptyset} r_C \mu(C) (\phi_2 - \phi_1) \\ &= (\phi_2 - \phi_1) \sum_{C \in Z_n, C \not\subseteq B, C \cap B \neq \emptyset} r_C \mu(C) \rightarrow 0 \end{aligned}$$

Also ist D messbar und

$$\mu(D) = \lim_{n \rightarrow \infty} U(Z'_n, D) = (\phi_2 - \phi_1) \sum_{C \in Z_n, C \subseteq B} r_C \mu(C) = (\phi_2 - \phi_1) r_B \mu(B) \quad \square$$

25.7 Kugelkoordinaten

25.7.1 Kugelkoordinaten

Wir kombinieren hier die Polarkoordinaten der x - y -Ebene mit der Polarkoordinaten der Ebene durch die z -Achse (als erster Achse) und den Punkt P

Die *Kugelkoordinaten* (r, ϑ, φ) sind mit den kartesischen Koordinaten (x, y, z) verknüpft durch

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

wobei $r \geq 0$, $0 \leq \vartheta \leq \pi$ und $0 \leq \varphi \leq 2\pi$. Die Substitutionsfunktion ist (Fig.13)

$$\begin{aligned} (x, y, z) &= \sigma(r, \varphi, \vartheta) \\ &= (r \cos \varphi \sin \vartheta, r \sin \varphi \sin \vartheta, r \cos \vartheta) \end{aligned}$$

25.7.2 Kugelschalensektor

$$D = \{(r \cos \varphi \sin \vartheta, r \sin \varphi \sin \vartheta, r \cos \vartheta) \mid r_1 \leq r \leq r_2, \varphi_1 \leq \varphi \leq \varphi_2, \vartheta_1 \leq \vartheta \leq \vartheta_2\}$$

ist Rotationskörper des Kreisringsektors

$$B = \{(r \sin \vartheta, r \cos \vartheta) \mid r_1 \leq r \leq r_2, \vartheta_1 \leq \vartheta \leq \vartheta_2\}$$

Hat der Schwerpunkt von B die Polarkoordinaten (r_B, ϑ_B) , so ist der Abstand von der z -Achse

$$r_B \sin \vartheta_B$$

und es folgt mit $\Delta r = r_2 - r_1$, $\Delta \varphi = \varphi_2 - \varphi_1$ und $\Delta \vartheta = \vartheta_2 - \vartheta_1$

$$\mu(D) = \Delta \varphi \cdot r_B \cdot \sin \vartheta_B \cdot \mu(B) = r_B^2 \sin \vartheta_B \cdot \Delta r \Delta \varphi \Delta \vartheta$$

Ist $\Delta \vartheta$ im Verhältnis zu Δr hinreichend klein, so liegt der Schwerpunkt im Inneren von B . Damit folgt

25.7.3 Substitution

Korollar 25.12 *Ist σ durch Kugelkoordinaten gegeben, so erhält man eine Substitution mit*

$$\tau(r, \varphi, \vartheta) = r^2 \sin \vartheta$$

Insbesondere gilt für messbare kompakte B im Definitionsbereich von σ und stetige $f : \sigma(B) \rightarrow \mathbb{R}$

$$\int_{\sigma(B)} f(x) dx = \int_B f((r \cos \varphi \sin \vartheta, r \sin \varphi \sin \vartheta, r \cos \vartheta)) \cdot r^2 \sin \vartheta d(r, \varphi, \vartheta)$$

Ist B ein Normalbereich, so kann man die rechte Seite entsprechend durch Dreifachintegrale ausdrücken.

Beispiel 10 Die Kugel B mit Mittelpunkt $(0, 0, 0)$ und Radius $R > 0$ ist das Bild des Quaders $T = [0, R] \times [0, 2\pi] \times [0, \pi]$ unter der oben beschriebenen Transformation. Für das Kugelvolumen finden wir daher

$$\begin{aligned} \mu(B) &= \int_B d(x, y, z) = \int_0^R \int_0^{2\pi} \int_0^\pi r^2 \sin \vartheta \, d\vartheta d\varphi dr \\ &= \int_0^R \int_0^{2\pi} -r^2 \cos \vartheta \Big|_0^\pi d\varphi dr \\ &= \int_0^R \int_0^{2\pi} 2r^2 d\varphi dr \\ &= \int_0^R 4\pi r^2 dr = \frac{4}{3} \pi R^3. \end{aligned}$$

25.7.4 Kugelkoordinaten, alternativ

Die Kugelkoordinaten (r, ϑ, φ) sind mit den kartesischen Koordinaten (x, y, z) verknüpft durch

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

wobei $r \geq 0$, $-\frac{\pi}{2} \leq \vartheta \leq \frac{\pi}{2}$ und $0 \leq \varphi < 2\pi$. Die Transformationsfunktion ist (Fig.9)

$$\begin{aligned} (x, y, z) &= h(r, \vartheta, \varphi) \\ &= (r \cos \vartheta \cos \varphi, r \cos \vartheta \sin \varphi, r \sin \vartheta) \end{aligned}$$

Das Kugelvolumen mit Radius R ist $\frac{2}{3} R^3 \int_0^{2\pi} d\varphi = \frac{4}{3} \pi R^3$.

25.8 ε -Substitution

Es genügt, wenn die Bedingungen für eine Substitution nur asymptotisch erfüllt sind (und das kann man dann für eine bekannte hinreichende Bedingung ausnutzen).

Sei $U \subseteq \mathbb{R}^d$ offen, $B \subseteq U$ messbar und kompakt. $\sigma, \tau : U \rightarrow \mathbb{R}^d$ bilden eine ε -Substitution für B falls gilt

1. σ und τ sind stetig
2. $\sigma(B)$ ist messbar
3. Es gibt eine Folge Z_n von Zerlegungen von B mit Weite $\rightarrow 0$ und Zellen $C \subseteq U$ so, dass gilt
 - (a) die $\sigma(Z_n)$ sind Zerlegungen von $\sigma(B)$
 - (b) Weite $\rightarrow 0$ für die $\sigma(Z_n)$
 - (c) zu jedem ε_0 gibt es ein k_0 so, dass für alle $k \geq k_0$ und $C \in Z_n$ ein $\xi_{nC} \in C$ existiert mit

$$|\mu(\sigma(C)) - \tau(\xi_{nC})\mu(C)| \leq \varepsilon \cdot \mu(C)$$

Es folgt mit Satz 4.6, dass $\sigma(B)$ kompakt ist.

Satz 25.13 *Ist B messbar und kompakt, bestimmen $\sigma, \tau : U \rightarrow \mathbb{R}$ eine ε -Substitution für B und ist $f : \sigma(B) \rightarrow \mathbb{R}$ stetig, so gilt*

$$\int_B f(\sigma(u)) \cdot \tau(u) \, du = \int_{\sigma(B)} f(x) \, dx$$

Beweis. Sei $M = \max\{|f(x)| \mid x \in B\}$. Ist $\eta > 0$ gegeben, so setze

$$\varepsilon = \frac{\eta}{M \cdot \mu(B)}$$

und wähle k_0 und ξ_{nC} gemäß (c). Dann folgt

$$\begin{aligned} & \left| \sum_{C \in Z_n, C \subseteq B} f(\sigma(\xi_{nC})) \mu(\sigma(C)) - \sum_{C \in Z_n, C \subseteq B} f(\sigma(\xi_{nC})) \cdot \tau(\xi_{nC}) \cdot \mu(C) \right| \\ & \leq \sum_{C \in Z_n, C \subseteq B} M \cdot \varepsilon \cdot \mu(C) = M \cdot \varepsilon \mu(B) = \eta \end{aligned}$$

Mit Satz 24.7 folgt die Behauptung. \square

25.9 Funktionaldeterminante

25.9.1 Substitutionsregel

Satz 25.14 *Sei $U \subseteq \mathbb{R}^d$ offen, und $\sigma : U \rightarrow \mathbb{R}^d$ sei stetig differenzierbar und injektiv. Sei $\det D\sigma(u) \neq 0$ auf U entweder überall positiv oder überall negativ. Weiter sei B eine kompakte und Jordan-messbare Teilmenge von B , und f sei eine auf $\sigma(B)$ stetige reellwertige Funktion. Dann erhält man eine ε -Substitution für B mit*

$$\sigma \text{ und } \tau(u) = |\det D\sigma(u)|$$

Inbesondere ist dann $\sigma(B)$ Jordan-messbar, f auf $\sigma(B)$ Riemann-integrierbar, und es ist

$$\int_{\sigma(B)} f = \int_B f \circ \sigma \cdot |\det D\sigma| \quad (25.24)$$

$\det D\sigma(u)$ ist die *Funktionaldeterminante* von σ , d.h. = $\det J_\sigma(u)$ unabhängig von der Basis.

Man beachte bei der Anwendung des Satzes, dass man diese Regeln wegen der geforderten Umkehrbarkeit oft nicht direkt anwenden kann, z.B. wenn man die Substitution für Polar- oder Kugelkoordinaten herleiten wollte.

Die Beweisidee ist, dass sich σ in der Nähe jedes Punktes von B durch jeweils eine affine Abbildung approximieren lässt und daher

$$\mu(\sigma(C)) \approx |\det J_\sigma(u)| \mu(C) \text{ für ein } u \in C \in Z_n$$

Der Beweis erfordert eine eingehendere Diskussion des Jordan-Maßes und wird auf Kap. 28 verschoben.

25.9.2 Beispiele

verallgemeinerte Polarkoordinaten

$$(x, y) = \sigma(r, \varphi) = (ar \cos \varphi, br \sin \varphi) \quad \text{mit } a, b > 0.$$

Dann ist $\det J_\sigma(r, \varphi) = abr$

Zylinderkoordinaten

$$(x, y, z) = \sigma(r, \varphi, z) = (r \cos \varphi, r \sin \varphi, z)$$

$$\det \sigma'(r, \varphi, z) = \det \begin{pmatrix} \cos \varphi & -r \sin \varphi & 0 \\ \sin \varphi & r \cos \varphi & 0 \\ 0 & 0 & 1 \end{pmatrix} = r.$$

Kugelkoordinaten

$$(x, y, z) = \sigma(r, \varphi, \vartheta) = (r \cos \varphi \sin \vartheta, r \sin \varphi \sin \vartheta, r \cos \vartheta)$$

$$\det \sigma'(r, \varphi, \vartheta) = \det \begin{pmatrix} \cos \varphi \sin \vartheta & -r \sin \varphi \sin \vartheta & r \cos \varphi \cos \vartheta \\ \sin \varphi \sin \vartheta & r \cos \varphi \sin \vartheta & r \sin \varphi \cos \vartheta \\ \cos \vartheta & 0 & -r \sin \vartheta \end{pmatrix} = -r^2 \sin \vartheta.$$

In der Tat mit

$$a = \sin \varphi, \quad b = \cos \varphi, \quad c = \sin \vartheta, \quad d = \cos \vartheta$$

gilt

$$a^2 + b^2 = 1, \quad c^2 + d^2 = 1$$

und man erhält die Determinante durch Herausziehen von r und Entwickeln nach der zweiten Spalte als

$$\begin{aligned} \det \begin{pmatrix} bc & -rac & rbd \\ ac & rbc & rad \\ d & 0 & -rc \end{pmatrix} &= r^2 \det \begin{pmatrix} bc & -ac & bd \\ ac & bc & ad \\ d & 0 & -c \end{pmatrix} = r^2 (ac \det \begin{pmatrix} ac & ad \\ d & -c \end{pmatrix} + bc \det \begin{pmatrix} bc & bd \\ d & -c \end{pmatrix}) \\ &= r^2 (ac(-ac^2 - ad^2) + bc(-bc^2 - bd^2)) = r^2 (-a^2c(c^2 + d^2) - b^2c(c^2 + d^2)) \\ &= -r^2(a^2c + b^2c) = -r^2c(a^2 + b^2) = -r^2c \end{aligned}$$

26 Integralsätze in der Ebene

In diesem Kapitel ist \mathbb{R}^2 eine Abkürzung für: 2-dimensionaler orientierter euklidischer Vektorraum. Wenn von einem $x_1 - x_2$ -Koordinatensystem gesprochen wird, ist die Basis \vec{e}_1, \vec{e}_2 positiv orientiert.

26.1 Wege

Wir betrachten zunächst *einfache Wege*

$$\Gamma : t \mapsto \vec{x}(t) \in \mathbb{R}^2, \quad t \in [a, b] \subseteq \mathbb{R}$$

die stetig differenzierbar sind. Eine *Umparametrisierung* wird gegeben durch eine bijektive stetige Abbildung

$$u \mapsto t(u), \quad [c, d] \rightarrow [a, b]$$

die im Inneren von $[c, d]$ stetig differenzierbar ist mit $\frac{\partial t}{\partial u} > 0$. Der so entstandene Weg ist

$$\Gamma_u : u \mapsto \vec{x}(t(u)) \quad u \in [c, d]$$

In der Literatur spricht man von regulären C^1 -Wegen und orientierungserhaltender C^1 -Umparametrisierung.

Zwei einfache Wege der Form Γ, Γ_u sind *äquivalent*, geschrieben $\Gamma \sim \Gamma_u$. Das ist eine Äquivalenzrelation auf der Menge der einfachen Wege.

Die *Weglänge* von Γ ist gegeben durch

$$s(t) = \int_a^t \left| \frac{\partial \vec{x}}{\partial t}(\tau) \right| d\tau$$

Ist $\left| \frac{\partial \vec{x}}{\partial t} \right| \neq 0$ für $t \neq a, b$ so erhält man durch die Umkehrfunktion eine Umparametrisierung von Γ

$$t = t(s), \quad s \in [0 = s(a), s(b)]$$

nämlich

$$\frac{\partial s}{\partial t} = \left| \frac{\partial \vec{x}}{\partial t} \right|, \quad \frac{\partial \vec{x}}{\partial s} = \frac{\partial \vec{x}}{\partial t} \cdot \frac{\partial t}{\partial s}$$

Für diese gilt dann

$$\left| \frac{\partial \vec{x}}{\partial s} \right| = 1$$

In diesem Falle sind einfache Wegeäquivalent genau dann, wenn sie dieselbe Umparametrisierung auf Weglänge haben.

Die *Umkehrung* eines einfachen Weges Γ ist

$$\ominus \Gamma : \ominus \vec{x}(t) = \vec{x}(a - t + b) \quad t \in [a, b]$$

Wir definiert nun induktiv: Einfache Wege sind *Wege*. Sind $\Gamma_i : \vec{x}_i(t), t \in [a_i, b_i]$ Wege und gilt

$$\vec{x}_1(b_1) = \vec{x}_2(a_2)$$

so ist auch die *Summe* ein Weg

$$\Gamma_1 \oplus \Gamma_2 : \vec{x}(t) = \begin{cases} \vec{x}_1(t) & t \in [a_1, b_1] \\ \vec{x}_2(t + (a_2 - b_1)) & t \in [b_1, b_1 + b_2 - a_2] \end{cases}$$

In der Literatur: stückweise C^1 -Wege. Wege sind *äquivalent*, $\Gamma \sim \tilde{\Gamma}$, wenn es $\Gamma_i \sim \tilde{\Gamma}_i$ gibt mit

$$\Gamma \sim \bigoplus_i \Gamma_i, \quad \tilde{\Gamma} \sim \bigoplus_i \tilde{\Gamma}_i$$

Eine Äquivalenzklasse L von Wegen ist eine *Plinie* ('P' steht für 'Parametrisierung').

26.2 Wegintegrale

Sei ein Weg Γ mit *Bahn* oder *Spur*

$$\text{Spur}(\Gamma) = \{\vec{x}(t) \mid t \in [a, b]\}$$

und eine stetige reellwertige auf der Bahn definierte Funktion f gegeben. Dann existiert

$$\int_{\Gamma} f(\vec{x}) dx_k := \int_a^b f(\vec{x}(t)) \frac{\partial x_k}{\partial t} dt$$

wobei $\frac{\partial x_k}{\partial t}$ die k -te Komponente von $\frac{\partial \vec{x}}{\partial t}$ ist. Statt dx_k sollte man besser $d\vec{x}^{(k)}$ schreiben, um Konfusion mit dem Differential der unabhängigen Variablen x_k zu vermeiden.

Dieses Integral ist invariant unter Umparametrisierung von Γ , da

$$\frac{\partial x_k}{\partial t} dt = \frac{\partial x_k}{\partial u} \cdot \frac{\partial u}{\partial t} dt = \frac{\partial x_k}{\partial u} du$$

Sei nun

$$\vec{F} = \begin{pmatrix} F_1 \\ F_2 \end{pmatrix} : \{\vec{x}(t) \mid t \in [a, b]\} \rightarrow \mathbb{R}^2$$

stetig. Wir definieren

$$\int_{\Gamma} \vec{F} = \int_{\Gamma} \vec{F}(\vec{x}) = \sum_{k=1}^2 \int_{\Gamma} F_k(\vec{x}) dx_k = \int_a^b F(\vec{x}(t)) \cdot \frac{\partial \vec{x}}{\partial t} dt$$

(wir schreiben hier das Skalarprodukt als $\vec{x} \cdot wy$. Wer sich noch mit infinitesimalen Größen wie $d\vec{x}$ auskennt, darf auch so schreiben

$$\int_{\Gamma} \vec{F}(\vec{x}) \cdot d\vec{x}$$

Also

$$\int_{\Gamma} f(\vec{x}) dx_1 = \int_{\Gamma} \begin{pmatrix} f(\vec{x}) \\ 0 \end{pmatrix}, \quad \int_{\Gamma} f(\vec{x}) dx_2 = \int_{\Gamma} \begin{pmatrix} 0 \\ f(\vec{x}) \end{pmatrix}$$

Wegen der Invarianz unter Umparametrisierung können wir bei einem Wegintegral \int_{Γ} das Γ als Repräsentanten seiner Äquivalenzklasse L auffassen, d.h. wir haben das *Pflichtenintegral*

$$\int_L \vec{F} := \int_{\Gamma} \vec{F} \quad \text{für } \Gamma \in L$$

Insbesondere gilt

- Aus $\Gamma \sim \ominus \tilde{\Gamma}$ folgt

$$\int_{\Gamma} f(\vec{x}) dx_k = - \int_{\tilde{\Gamma}} f(\vec{x}) dx_k$$

Da die Definitionen von Weg, Tangentialvektor und Wegintegral sich nur auf den euklidischen Vektorraum beziehen gilt:

- Wegintegrale sind vom Koordinatensystem abhängig.

Die "halben Wegintegrale" $\int_{\Gamma} f dx_k$ bzgl. x_k -Achse sind dagegen essentiell vom Koordinatensystem abhängig.

26.3 Normalbereiche bzgl. x_1 -Achse

Sei Γ ein Weg so, dass es stückweis stetig differenzierbare

$$\alpha, \beta : [a, b] \rightarrow \mathbb{R}, \quad \alpha \leq \beta, \quad \alpha(t) < \beta(t) \text{ für } a < t < b$$

gibt mit

$$\Gamma \sim \Gamma_1 \oplus \Gamma_2 \oplus \Gamma_3 \oplus \Gamma_4$$

wobei

$$\Gamma_1 : \begin{pmatrix} t \\ \alpha(t) \end{pmatrix}, \quad t \in [a, b]$$

$$\Gamma_2 : \begin{pmatrix} b \\ \alpha(b) + t \end{pmatrix}, \quad t \in [0, \beta(b) - \alpha(b)]$$

$$\ominus \Gamma_3 : \begin{pmatrix} t \\ \beta(t) \end{pmatrix}, \quad t \in [a, b]$$

$$\ominus \Gamma_4 : \begin{pmatrix} a \\ \alpha(a) + t \end{pmatrix}, \quad t \in [0, \beta(a) - \alpha(a)]$$

Dann ist

$$B = B_\Gamma = \left\{ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mid x_1 \in [a, b], \alpha(x_1) \leq x_2 \leq \beta(x_1) \right\}$$

der durch Γ definierte *Normalbereich* bzgl. der x_1 -Achse und Γ bzw. seine Bahn wird etwas missverständlich auch als Rand ∂B notiert. Beachte, dass ∂B durch Γ positiv orientiert durchlaufen wird.

Zu gegebenen α und β sind die Γ_i offensichtlich Wege, also auch Γ . Auch aus Umparametrisierungen von Γ_1 und Γ_3 kann man die ursprünglichen α und β bis auf eine gemeinsame Substitution eindeutig zurückgewinnen und damit auch B . Schliesslich kann man auch aus einer Umparametrisierung von Γ die Summanden Γ_i bis auf Umparametrisierung zurückgewinnen - da Γ_2 und Γ_4 vertikal verlaufen. Es folgt

$$B_\Gamma \text{ hängt nur von der Äquivalenzklasse von } \Gamma \text{ ab}$$

26.4 Halbgrün

Lemma 26.1 Sei (Γ, B_Γ) Normalbereich bzgl. der x_1 -Achse und $f : U \rightarrow \mathbb{R}^2$ stetig differenzierbar, $U \subseteq \mathbb{R}^2$ offen, und $B_\Gamma \subseteq U$. Dann gilt

$$\int_{B_\Gamma} \frac{\partial f}{\partial x_2} = - \int_\Gamma f(\vec{x}) dx_1$$

Beweis. Mit Fubini und Hauptsatz

$$\int_{B_\Gamma} \frac{\partial f}{\partial x_2} = \int_a^b \left(\int_{\alpha(x_1)}^{\beta(x_1)} \frac{\partial f}{\partial x_2} dx_2 \right) dx_1 = \int_a^b (f(x_1, \beta(x_1)) - f(x_1, \alpha(x_1))) dx_1$$

Nun gilt für Γ_1 in der ersten Komponente $\frac{\partial x_1}{\partial t} = 1$, also $dx_1 = dt$ und somit

$$\int_a^b f(x_1, \alpha(x_1)) dx_1 = \int_a^b f(\vec{x}(t)) dt = \int_{\Gamma_1} f(\vec{x}) dx_1$$

Andererseits gilt für Γ_3 in der ersten Komponente $\frac{\partial x_1}{\partial t} = -1$, also $dx_1 = -dt$ und somit

$$\int_a^b f(x_1, \beta(x_1)) dx_1 = \int_a^b f(\vec{x}(t)) dt = - \int_{\Gamma_3} f(\vec{x}) dx_1$$

Für Γ_2 und Γ_4 gilt $\frac{\partial x_1}{\partial t} = 0$, also

$$\int_{\Gamma_2} f(\vec{x}) dx_1 = 0 = \int_{\Gamma_4} f(\vec{x}) dx_1$$

Es folgt

$$\begin{aligned} & \int_a^b (f(x_1, \beta(x_1)) - f(x_1, \alpha(x_1))) dx_1 \\ &= - \int_{\Gamma_1} f(\vec{x}) dx_1 - \int_{\Gamma_2} f(\vec{x}) dx_1 - \int_{\Gamma_3} f(\vec{x}) dx_1 - \int_{\Gamma_4} f(\vec{x}) dx_1 = - \int_{\Gamma} f(\vec{x}) dx_1 \quad \square \end{aligned}$$

26.5 Normalbereich bzgl. der x_2 -Achse

Sei

$$\sigma : \mathbb{R}^2 \rightarrow \mathbb{R}^2, \quad \sigma \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_2 \\ x_1 \end{pmatrix}$$

Der Weg Γ definiert den *Normalbereich* (Γ, B) bzgl. der x_2 -Achse genau dann, wenn $\sigma \circ \Gamma$ den Normalbereich $(\sigma \circ \Gamma, \sigma(B))$ bzgl. der x_1 -Achse definiert. Beachte, dass ∂B von Γ negativ orientiert durchlaufen wird. Es folgt mit f wie oben

$$\int_B \frac{\partial f}{\partial x_1} d\vec{x} = - \int_{\Gamma} f(\vec{x}) dx_2 = \int_{\ominus \Gamma} f(\vec{x}) dx_2$$

26.6 Grüner Bereich

Ist (Γ, B) s Normalbereich bzgl. x_1 -Achse und $(\tilde{\Gamma}, B)$ Normalbereich bzgl. x_2 -Achse und gilt $\tilde{\Gamma} \sim \ominus \Gamma$, so ist (Γ, B) ein *grüner Bereich*.

Lemma 26.2 Sei $B = B_{\Gamma}$ durch α, β als Normalbereich bzgl. x_1 -Achse gegeben. Seien $a \leq a_i \leq b_i \leq b$ so, das

$$\frac{\partial \alpha}{\partial x_1}(t) = \begin{cases} < 0 & a < t < a_1 \\ = 0 & a_1 < t < b_1 \\ > 0 & b_1 < t < b \end{cases} \quad \frac{\partial \beta}{\partial x_1}(t) = \begin{cases} > 0 & a < t < a_2 \\ = 0 & a_2 < t < b_2 \\ < 0 & b_2 < t < b \end{cases}$$

Dann ist (Γ, B) ein grüner Bereich.

Beweis. Setze

$$c = \alpha(a_1), \quad d = \beta(a_2), \quad c_1 = \alpha(a), \quad d_1 = \beta(a), \quad c_2 = \alpha(b), \quad d_2 = \beta(d)$$

Sei ϕ Umkehrung von α auf $[c, c_1]$, von β auf $[d_1, d]$ und konstant $= a$ auf $[c_1, d_1]$. Sei ψ Umkehrung von α auf $[c, c_2]$, von β auf $[d_2, d]$ und konstant $= b$ auf $[c_2, d_2]$. Wenn man in der Definition von Normalbereich bzgl. x_1 -Achse x_1 und x_2 vertauscht, ergibt sich als definierter Bereich wieder B . Und für den zugehörigen Weg $\tilde{\Gamma}$ gilt $\tilde{\Gamma} \sim \ominus\Gamma$. \square

Satz 26.3 (Green) Sei $U \subseteq \mathbb{R}^2$ offen, $\vec{f} = \begin{pmatrix} F_1 \\ F_2 \end{pmatrix} : U \rightarrow \mathbb{R}^2$ stetig differenzierbar, $B \subseteq U$ und (Γ, B) grüner Bereich, Dann gilt

$$\int_{\Gamma} \vec{f} = \int_{\Gamma} \vec{F}(\vec{x}) = \int_B \left(\frac{\partial F_2}{\partial x_1} - \frac{\partial F_1}{\partial x_2} \right)$$

Der Ausdruck

$$\int_{\partial B} \vec{F} \cdot d\vec{x}$$

ist nicht definiert und kann auch nicht ohne Weiteres sinnvoll definiert werden. Daher sollte er nicht benutzt werden. Manche Autoren meinen damit wohl irgendwie ein Wegintegral wie im Satz. Der Ausdruck stammt aus der Infinitesimalen Analysis und hat dort seine exakte Bedeutung: Summation, über die infinitesimalen Linienelemente $d\vec{s}$ der 'Linie' ∂B , der infinitesimalen Skalare $\vec{F} \cdot d\vec{s}$, \vec{F} der Feldvektor am Ort von $d\vec{s}$.

Beweis. Wie oben gezeigt

$$\begin{aligned} \int_{\Gamma} \vec{F}(\vec{x}) d\vec{x} &= \int_{\Gamma} F_1 dx_1 + \int_{\Gamma} F_2 dx_2 = \int_{\Gamma} F_1 dx_1 - \int_{\ominus\Gamma} F_2 dx_2 = - \int_{B_{\Gamma}} \frac{\partial F_1}{\partial x_2} - \left(- \int_{B_{\ominus\Gamma}} \frac{\partial F_2}{\partial x_1} \right) \\ &= - \int_{B_{\Gamma}} \frac{\partial F_1}{\partial x_2} + \int_{B_{\Gamma}} \frac{\partial F_2}{\partial x_1} = \int_B \left(\frac{\partial F_2}{\partial x_1} - \frac{\partial F_1}{\partial x_2} \right) \quad \square \end{aligned}$$

26.7 Orthogonal

In einen 2-dimensionalen euklidischen Vektorraum gibt es eine eindeutig bestimmte Abbildung $\vec{x} \mapsto \vec{x}^{\perp}$ mit

$$\vec{x}^{\perp} \perp \vec{x}, \quad |\vec{x}^{\perp}| = |\vec{x}|, \quad \det(\vec{x}, \vec{x}^{\perp}) > 0 \text{ für } \vec{x} \neq \vec{0}$$

(die 90°-Drehung). Diese Abbildung ist linear und es gilt

$$\vec{x}^{\perp\perp} = -\vec{x}, \quad \vec{x}^{\perp} \cdot \vec{y} = -\vec{x} \cdot \vec{y}^{\perp}$$

26.8 Divergenz

Sei $U \subseteq \mathbb{R}^2$ offen und $\vec{F} = \begin{pmatrix} F_1 \\ F_2 \end{pmatrix} : U \rightarrow \mathbb{R}^2$ stetig differenzierbar. Dann

$$\operatorname{div} \vec{F}(\vec{x}) := \frac{\partial F_1}{\partial x_1}(\vec{x}) + \frac{\partial F_2}{\partial x_2}(\vec{x}) = \operatorname{Spur} D\vec{F}(\vec{x})$$

also unabhängig vom Koordinatensystem. Die physikalische Interpretation ist unmittelbar klar, wenn man zu einem Feldvektor das Koordinatensystem so wählt, das eine Achse zum Vektor parallel ist.

Die Spur einer Matrix $A \in \mathbb{R}^{n \times n}$ ist die Summe der Diagonaleinträge und man rechnet nach, dass $\operatorname{Spur}(AB) = \operatorname{Spur}(BA)$. Es folgt $\operatorname{Spur}(S^{-1}AS) = \operatorname{Spur}(A)$.

26.9 Rotation

$$\operatorname{rot} \vec{F}(\vec{x}) := \frac{\partial F_2}{\partial x_1}(\vec{x}) - \frac{\partial F_1}{\partial x_2}(\vec{x}) = \operatorname{div}(-\vec{F}^\perp)(\vec{x})$$

also ebenfalls unabhängig vom Koordinatensystem und auch die physikalische Bedeutung ist sofort klar.

26.10 Grüner Gauss von Scharfenburg

Satz 26.4 Für grünen Bereich (Γ, B) und $\vec{F} \in C^1(U)$, $U \supseteq B$ offen gilt nach Green bzw. Ostrogradski-Gauß

$$\begin{aligned} \int_B \operatorname{rot} \vec{F} &= \int_\Gamma \vec{F} = \int_a^b \vec{F}(\vec{x}(t)) \cdot \frac{\partial \vec{x}}{\partial t} dt \\ \int_B \operatorname{div} \vec{F} &= \int_\Gamma \vec{F}^\perp dx = \int_a^b \vec{F}^\perp(\vec{x}(t)) \cdot \frac{\partial \vec{x}}{\partial t} dt = \int_a^b \vec{F}(\vec{x}(t)) \cdot -\left(\frac{\partial \vec{x}}{\partial t}\right)^\perp dt \\ &= \int_a^b \vec{F}(\vec{x}(t)) \cdot \vec{n}(t) \left| \frac{\partial \vec{x}}{\partial t} \right| dt \end{aligned}$$

Hier ist

$$\operatorname{div} \vec{F} = \operatorname{rot} \vec{F}^\perp, \quad \vec{N}(t) := -\left(\frac{\partial \vec{x}}{\partial t}\right)^\perp$$

äußer Normalenvektor an Γ und der normierte dazu ist

$$\vec{n}(t) = \begin{cases} \vec{0} & \text{falls } \vec{N}(t) = \vec{0} \\ \frac{1}{|\vec{N}(t)|} \vec{N}(t) & \text{sonst} \end{cases}$$

Für das letzte Integral findet man auch die folgende höchst problematische Schreibweise:

$$\int_{\partial B} \vec{F} \cdot \vec{n} ds$$

In der Infinitesimalen Analysis hat man zu jedem infinitesimalen Linienelement $d\vec{s}$ die normierte Normale \vec{n} , kann das Skalarprodukt bilden und dann über die Linienelemente

aufsummieren und so den Fluss durch den Rand ∂B erhalten. Demgegenüber beschreibt $\operatorname{div} \vec{F}(\vec{x})$ den infinitesimalen Fluss an der Quelle \vec{x} und $\int_B \operatorname{div} F$ die Summe dieser Flüsse an Quellen in B . Das *Divergenztheorem* von Ostrogradski-Gauß besagt die Gleichheit beider Summen.

Korollar 26.5 *Ist (Γ, B) ein grüner Bereich, so gilt*

$$\mu(B) = \int_{\Gamma} \frac{1}{2} \vec{x}^{\perp} d\vec{x}$$

Beweis.

$$1 = \frac{1}{2} \operatorname{div} \vec{x} = \frac{1}{2} \operatorname{rot} \vec{x}^{\perp}$$

26.11 Allgemeinerer Bereiche

Den Begriff *grüner Bereich B mit Rand Γ* können wir so erweitern;

- (B, Γ) ist *grün*, wenn es das bzgl. eines passenden ON-Koordinatensystems ist.
- Ist (B, Γ) grün und $\tilde{\Gamma} \sim \Gamma$ so ist auch $(B, \tilde{\Gamma})$ grün
- Sind (B_i, Γ_i) grün und gilt
 - $\Gamma_i = \Gamma_{i0} \oplus \Gamma_{i2}$
 - $\Gamma_{11} = \ominus \Gamma_{21}$
 - $B_1 \cap B_2$ die Bahn von Γ_{11}
 - $B = B_1 \cup B_2$ und $\Gamma = \Gamma_{12} \oplus \Gamma_{22}$

so ist auch (B, Γ) *grün*

Korollar 26.6 *Der Satz gilt für grüne Bereiche in diesem allgemeineren Sinne.*

Auch eine Erweiterung mit nicht zusammenhängendem Rand ist möglich z.B. so: Sind (B_i, Γ_i) grün und gilt

- $\Gamma_i = \Gamma_{i0} \oplus \Gamma_{i2} \oplus \Gamma_{i3} \oplus \Gamma_{i4}$
- $\Gamma_{12} = \ominus \Gamma_{22}$ und $\Gamma_{14} = \ominus \Gamma_{24}$
- $B_1 \cap B_2$ die Vereinigung der Bahnen von Γ_{12} und Γ_{14}
- $B = B_1 \cup B_2$, $\Gamma^1 = \Gamma_{11} \oplus \Gamma_{23}$, $\Gamma^2 = \Gamma_{13} \oplus \Gamma_{21}$

so ist auch $(B, \Gamma^1 \cup \Gamma^2)$ *grün*.

Hier ist dann im Satz zu lesen

$$\int_{\Gamma^1 \cup \Gamma^2} = \int_{\Gamma^1} + \int_{\Gamma^2}$$

27 Integralsätze im Raum

In diesem Kapitel steht ‘ \mathbb{R}^3 ’ für: 3-dimensionaler orientierter euklidischer Vektorraum. Bei einem Koordinatensystem x_1, x_2, x_3 ist die Basis $\vec{e}_1, \vec{e}_2, \vec{e}_3$ stets als positiv orientiert vorausgesetzt.

Bei Parameterdarstellungen ϕ werden die Komponenten als $\phi^{(i)}$ notiert. Bei Vektorfeldern \vec{F} jedoch als F_i .

27.1 Vektorprodukt

vgl. 13.8. Bzgl. positiv orientierter ON-Basis \vec{e}_i kann man es so ausrechnen: Stehen die Koordinaten von \vec{a}_1 in der ersten, die von \vec{a}_2 in der zweiten Spalte von

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{pmatrix}$$

so bilde man die Matrix A_k durch Streichen der k-ten Zeile und hat dann für $\vec{a}_1 \times \vec{a}_2$ die Koordinaten

$$\begin{pmatrix} \det A_1 \\ -\det A_2 \\ \det A_3 \end{pmatrix}$$

Lemma 27.1

$$(c_2\vec{e}_2 + c_3\vec{e}_3) \cdot (a_{12}\vec{a}_1 - a_{11}\vec{a}_2) = (c_3\vec{e}_2 - c_2\vec{e}_3)(\vec{a}_1 \times \vec{a}_2)$$

Beweis. Ist $c_2 = 1, c_3 = 0$ so ergibt sich links $\det A_3$. Ist $c_2 = 0, c_3 = 1$ so ergibt sich $-\det A_2$. \square .

Lemma 27.2 Ist $\vec{b}_1 = c_{11}\vec{a}_1 + c_{21}\vec{a}_2$ und $\vec{b}_2 = c_{12}\vec{a}_1 + c_{22}\vec{a}_2$ so folgt

$$\vec{b}_1 \times \vec{b}_2 = (\det C)(\vec{a}_1 \times \vec{a}_2)$$

Beweis durch Ausmultiplizieren

$$\vec{b}_1 \times \vec{b}_2 = c_{11}c_{22}(\vec{a}_1 \times \vec{a}_2) + c_{21}c_{12}(\vec{a}_2 \times \vec{a}_1) \quad \square$$

27.2 Einfache Parameterdarstellungen

Sei $G \subseteq \mathbb{R}^2$ offen und $\phi : G \rightarrow \mathbb{R}^3$, stetig differenzierbar. Ist $K \subseteq G$ kompakt und messbar, so ist (ϕ, K) bzw. (ϕ, K) eine *einfache Parameterdarstellung* und $G = \text{dom}\phi$. Die dargestellte “Fläche” ist

$$\text{Spur}(\phi, K) = \{\phi(u) \mid u \in K\}$$

Die einfache Parameterdarstellung (ψ, H) entsteht aus (ϕ, K) durch die *Umparametrisierung* $\sigma : \text{dom}\psi \rightarrow \text{dom}\phi$ falls gilt

1. σ ist Differomorphismus

2. $\sigma(H) = K$
3. $\psi = \phi \circ \sigma$
4. $\det D\sigma(v) > 0$ für alle $v \in H$

Die einfachen Parameterdarstellungen (ϕ, K) und (ψ, H) sind *äquivalent*, wenn (ψ, H) aus (ϕ, K) durch Umparametrisierung entsteht. Das bestimmt offenbar eine Äquivalenzrelation auf der Menge der einfachen Parameterdarstellungen.

27.3 Tangenten und Normalen

Sei eine einfache Parameterdarstellung (ϕ, K) gegeben. Man kann die Spalten der Jacobimatrix $J_\phi(u) \hat{=} D\phi(u)$ als Tangentialvektoren verstehen:

$$\vec{\partial}_k \phi(u) = \begin{pmatrix} \partial_k \phi^{(1)} \\ \partial_k \phi^{(2)} \\ \partial_k \phi^{(3)} \end{pmatrix}, \quad k = 1, 2$$

Der *Normalenvektor* ist dann

$$\vec{N} = \vec{N}_\phi = \vec{\partial}_1 \phi \times \vec{\partial}_2 \phi$$

Er ist $\neq \vec{0}$ genau dann, wenn $D\phi(u)$ Rang 2 hat, d.h. die Tangentialvektoren sind unabhängig. Dann wird er normiert zum *Normaleneinheitsvektor* $\vec{n}_\phi(u)$ und ϕ ist an u *regulär*. Die so bestimmten Vektoren des Raumes hängen nicht vom Koordinatensystem ab.

Korollar 27.3 *Entsteht (ψ, L) aus (ϕ, K) durch die Umparametrisierung σ , so gilt*

$$\vec{N}_\psi(s) = \det D\sigma(s) \vec{N}_\phi(\sigma(s)), \quad \vec{n}_\psi = \vec{n}_\phi \circ \sigma, \quad \text{Spur}(\psi, L) = \text{Spur}(\phi, K)$$

Beweis. Nach der Kettenregel

$$D\psi(s) = D\phi(\sigma(s)) \circ D\sigma(s)$$

also

$$\vec{\partial}_k \psi(s) = c_{1k} \vec{\partial}_1 \phi(s) + c_{2k} \vec{\partial}_2 \phi(s), \quad C = J_\sigma(s)$$

und die Behauptung folgt mit mit Lemma 27.2. \square Im Fall $\det D\sigma < 0$ gilt $\vec{n}_\psi = -\vec{n}_\phi \circ \sigma$.

27.4 Zusammengesetzte Parameterdarstellungen

Induktive Definition: Eine *Parameterdarstellung* ist eine einfache Parameterdarstellung oder eine *Zusammensetzung*, d.h. eine Liste

$$(\phi, K) = (\phi_1, K_1), \dots, (\phi_n, K_n)$$

von Parameterdarstellungen so, dass

$$\mu\{x \in K_i \mid \exists j \neq i \exists y \in K_j. \phi_i(x) = \phi_j(y)\} = 0 \quad \text{für alle } i$$

Wir definieren

$$\text{Spur}(\phi, K) = \text{Spur}(\phi_1, K_1) \cup \dots \cup \text{Spur}(\phi_n, K_n)$$

Die *Äquivalenz* ist definiert als die feinste Äquivalenzrelation, die die Äquivalenz einfacher Parametrisierungen einschließt und mit Zusammensetzung verträglich ist: zugehörige *Umparametrisierung* σ ist dann eine Liste $\sigma_1, \dots, \sigma_n$ von Umparametrisierungen (ϕ_i, K_i) zu (ψ_i, H_i) . Wir schreiben $\psi = \phi \circ \sigma$.

Nach passender Umparametrisierung durch Verschiebung, kann man die K_i oder auch die $\text{dom}\phi_i$ als disjunkt ansehen und K als die Vereinigung der K_i auffassen. Wir schreiben dann

$$(\phi, K) = (\phi_1, K_1) \oplus \dots \oplus (\phi_n, K_n), \quad \text{dom}\phi = \text{dom}\phi_1 \cup \dots \cup \text{dom}\phi_n$$

Die (ϕ_i, K_i) können wir auch als einfach voraussetzen. Dann gibt es zu jedem $u \in K$ einen wohldefinierten Punkt $\phi(u) \in \text{Spur}(\phi, K)$. Ist i mit $\phi(u) \in \text{Spur}(\phi_i, K_i)$ eindeutig bestimmt, so sind auch die Vektoren

$$\vec{\partial}_k \phi(u) = \vec{\partial}_k \phi_i(u), \quad \vec{N}_\phi(u) = \vec{N}_{\phi_i}(u)$$

eindeutig bestimmt. Andernfalls setzen wir sie $= \vec{0}$.

Auf gleiche Weise können wir eine Umparametrisierung auf

$$(\psi, H) = (\psi_1, H_1) \oplus \dots \oplus (\psi_n, H_n)$$

als $\sigma : \text{dom}\psi \rightarrow \text{dom}\phi$ verstehen.

27.5 Pflähen

Die Klassen S äquivalenter Parametrisierungen nennen wir *Pflähen* - insbesondere haben je zwei Parametrisierungen in S dieselbe Spur und wir setzen

$$\text{Spur}S = \text{Spur}(\phi, K) \quad (\phi, K) \in S$$

d.h. anschaulich: wir erhalten dieselbe Fläche im Raum. Tritt, eventuell nach passender Umparametrisierung, der Fall ein, dass ϕ injektiv ist, so können wir jedem $\vec{x} \in \text{Spur}S$ eindeutig den Normaleinheitsvektor zuordnen

$$\vec{n}_\phi(u), \quad \vec{x} = \phi(u)$$

27.6 Pflähenintegrale

Sei nun (ϕ, K) eine Parameterdarstellung $f : \text{Spur}(\phi, K) \rightarrow \mathbb{R}$ stetig. Dann existiert das Integral

$$\int_K f(\phi(u)) |\vec{N}_\phi(u)| = \sum_{i=1}^n \int_{K_i} f(\phi_i(u)) |\vec{N}_{\phi_i}(u)|$$

wobei

$$(\phi, K) = (\phi_1, K_1) \oplus \dots \oplus (\phi_n, K_n)$$

Korollar 27.4 Sind (ϕ, K) und (ψ, H) äquivalent so gilt

$$\int_K f(\phi(u)) |\vec{N}_\phi(u)| = \int_H f(\psi(v)) |\vec{N}_\psi(v)|$$

Beweis. Kor. 27.3 und Substitutionsregel für Bereichsintegrale. \square Wir dürfen also für eine Pflähe S das *Pflächenintegral* der Funktion f definieren als

$$\int_S f = \int_K f(\phi(u)) |\vec{N}_\phi(u)| \quad \text{mit } (\phi, K) \in S$$

und *Flächeninhalt* von S als

$$\int_S 1$$

Zur Motivation: Das von den Vektoren in Tangentialrichtung

$$\Delta u_1 \vec{\partial}_1 \phi, \quad \Delta u_2 \vec{\partial}_2 \phi$$

im Punkt $\phi(u)$ aufgespannte Parallelogramm ($\Delta u_k \geq 0$) hat die Fläche

$$|\Delta u_1 \vec{\partial}_1 \phi \times \Delta u_2 \vec{\partial}_2 \phi| = \Delta u_1 \cdot \Delta u_2 \cdot |N_\phi(u)|$$

die näherungsweise als Flächeninhalt von $\phi([u_1, u_1 + \Delta u_1] \times [u_2, u_2 + \Delta u_2])$ angesehen werden kann.

In alten Zeiten als konnte man die Summation der Skalare $f d\sigma$ mit infinitesimalen Flächenelementen $d\sigma$ verstehen (die man sich als infinitesimale Parallelogramme denken durfte). Von daher datiert die Notation $\int_S f d\sigma$. Die Dekoration mit $d\sigma$ ist unschädlich, solange man sich klar macht, dass das Integral nicht aus der Spur von S definiert ist, sondern dass die Parametrisierung bzw. ihre Äquivalenzklasse notwendig in die Notation des Integrals eingehen muss. Also: $\int_M f d\sigma$ ist für Mengen $M \subseteq \mathbb{R}^3$ nicht definiert, auch dann, wenn M die Spur einer Pflähe ist. Wenn man statt M den Buchstaben S benutzt, wird die Sache nicht besser.

27.7 Rotationsflächen

Satz 27.5 Sei $\Gamma : t \mapsto \vec{x}(t) \in \mathbb{R}^2$ auf offener Umgebung U von $[a, b]$ stetig differenzierbar mit $\frac{\partial \vec{x}}{\partial t} \neq \vec{0}$ und $x_2(t) \geq 0$ auf $[a, b]$. Dann ist (ϕ, K) mit

$$\phi \begin{pmatrix} t \\ u \end{pmatrix} = \begin{pmatrix} x_1(t) \\ \cos(u)x_2(t) \\ \sin(u)x_2(t) \end{pmatrix}, \quad t \in U, \quad u \in [-\varepsilon, 2\pi + \varepsilon], \quad K = [a, b] \times [0, 2\pi]$$

eine Parameterdarstellung mit Flächeninhalt

$$\int_a^b x_2(t) \left| \frac{\partial \vec{x}}{\partial t}(t) \right| dt = L \cdot 2\pi S_2$$

wobei L die Länge und S der Schwerpunkt des Weges Γ ist, d.h. nach Umparametrisierung auf Weglänge

$$S_2 = \frac{1}{L} \int_0^L x_2(t(s)) ds$$

Lemma 27.6 $|\vec{a} \times \vec{b}|^2 = |\vec{a}|^2 \cdot |\vec{b}|^2 - (\vec{a} \cdot \vec{b})^2$

Beweis. O.B.d.A. $|\vec{a}| = 1$. Dann hat man

$$\vec{h} = \vec{b} - (\vec{a} \cdot \vec{b})\vec{a}$$

als Höhe des Parallelogramms mit Grundseite \vec{a} und Seite \vec{b} , nach Pythagoras $|\vec{h}|^2 = |\vec{b}|^2 - (\vec{a} \cdot \vec{b})^2$ und die Behauptung nach Definition des Vektorprodukts. \square

Beweis des Satzes

$$\vec{\partial}_1 \phi(t, u) = \begin{pmatrix} \frac{\partial x_1}{\partial t}(t) \\ \cos(u) \frac{\partial x_2}{\partial t}(t) \\ \sin(u) \frac{\partial x_3}{\partial t}(t) \end{pmatrix}, \quad \vec{\partial}_2 \phi(t, u) = \begin{pmatrix} 0 \\ -x_2(t) \sin u \\ x_2(t) \cos u \end{pmatrix}$$

also

$$|\vec{\partial}_1 \phi(t, u)|^2 = \left| \frac{\partial \vec{x}}{\partial t}(t) \right|^2, \quad |\vec{\partial}_2 \phi(t, u)|^2 = x_2(t)^2, \quad \vec{\partial}_1 \phi \cdot \vec{\partial}_2 \phi = 0$$

also mit dem Lemma

$$|\vec{N}_\phi(t, u)| = |\vec{\partial}_1 \phi(t, u) \times \vec{\partial}_2 \phi(t, u)| = x_2(t) \left| \frac{\partial \vec{x}}{\partial t}(t) \right|$$

Der Flächeninhalt ist nun

$$\begin{aligned} \int_{(\phi, K)} 1 &= \int_{(t, u) \in K} |\vec{N}_\phi(t, u)| = \int_0^{2\pi} \left(\int_a^b x_2(t) \left| \frac{\partial \vec{x}}{\partial t}(t) \right| dt \right) du \\ &= 2\pi \int_a^b x_2(t) \left| \frac{\partial \vec{x}}{\partial t}(t) \right| dt = 2\pi \int_0^L x_2(t(s)) ds = 2\pi L S_2 \end{aligned}$$

nach Umparametrisierung auf Weglänge $s(t)$, d.h. $\left| \frac{\partial \vec{x}(s)}{\partial s} \right| = 1$. \square

27.8 Normalenintegral

Sei (ϕ, K) eine Parametrisierung und \vec{F} ein auf $\text{Spur}(\phi, K)$ definiertes stetiges Vektorfeld. Dann es existiert

$$\int_K \vec{F}(\phi(u)) \cdot \vec{N}_\phi(u)$$

Korollar 27.7 Sind (ϕ, K) und (ψ, H) äquivalente Parametrisierungen

$$\int_K \vec{F}(\phi(u)) \cdot \vec{N}_\phi(u) = \int_H \vec{F}(\psi(v)) \cdot \vec{N}_\psi(v)$$

Beweis. Kor. 27.3 und Substitutionsregel für Bereichsintegrale. \square Wir dürfen also für eine Pfläche S das *Normalenintegral* des Feldes \vec{F} definieren als

$$\int_{S^\perp} \vec{F} = \int_K \vec{F}(\phi(u)) \cdot w N_\phi(u) \quad \text{mit } (\phi, K) \in S$$

Man findet oft folgende Schreibweise

$$\int_S \vec{F} \cdot \vec{n} d\sigma$$

wobei $d\sigma$ auf ein infinitesimales Flächenelement, \vec{n} auf den (ihm zugeordneten??) Normaleinheitsvektor verweist. Nun gilt ja $|\vec{n}d\sigma| = |\vec{N}|$, sinnvoller wäre also allemale

$$\int_S \vec{F} \cdot \vec{N}$$

27.9 Rotation

Zu einem normierten Vektor \vec{a} des Raumes sei

$$U_{\vec{a}} = \{\vec{x} \mid \vec{x} \perp \vec{a}\}$$

und $\pi_{\vec{a}}$ die Orthogonalprojektion des Raumes auf $U_{\vec{a}}$. Die positive Orientierung von $U_{\vec{a}}$ wird durch eine Basis \vec{b}, \vec{c} festgelegt, für die $\vec{b}, \vec{c}, \vec{a}$ im Raum positiv orientiert ist.

Sei nun \vec{F} ein differenzierbares Vektorfeld im Raum. Wir definieren die *Rotation bzgl. der Achse \vec{a}*

$$\text{rot}_{\vec{a}} \vec{F} = (\text{rot}_{U_{\vec{a}}}(\pi_{\vec{a}} \vec{F}|U)) \vec{a}$$

mittels der skalaren Rotation des Feldes $\pi_{\vec{a}} \vec{F}|U$ in $U_{\vec{a}}$. Damit ist klar, dass diese Definition basisunabhängig ist.

Sei ein $\alpha : \vec{e}_1, \vec{e}_2, \vec{e}_3$ positiv orientierte ON-Basis des Raumes. Wir definieren

$$\text{rot}_{\alpha} \vec{F} = \sum_i \text{rot}_{\vec{e}_i} \vec{F} \hat{=} \text{rot}_{J_{\vec{F}}^{1 \wedge 1}} \vec{e}_1 - \text{rot}_{J_{\vec{F}}^{2 \wedge 2}} \vec{e}_2 + \text{rot}_{J_{\vec{F}}^{3 \wedge 3}} \vec{e}_3$$

(beachte, dass $\vec{e}_1, \vec{e}_3, \vec{e}_2$ negativ orientiert ist) vgl. Kap. 14.5. Es folgt

$$(\text{rot}_{\alpha} \vec{F})^{\alpha} = \begin{pmatrix} 0 \\ \frac{\partial F_1}{\partial x_3} \\ -\frac{\partial F_1}{\partial x_2} \end{pmatrix} + \begin{pmatrix} -\frac{\partial F_2}{\partial x_3} \\ 0 \\ \frac{\partial F_2}{\partial x_1} \end{pmatrix} + \begin{pmatrix} \frac{\partial F_3}{\partial x_2} \\ -\frac{\partial F_3}{\partial x_1} \\ 0 \end{pmatrix} \quad \text{für } \vec{F}^{\alpha} = \begin{pmatrix} F_1 \\ F_2 \\ F_2 \end{pmatrix}$$

vgl. Merkregel $\text{rot} \vec{F} = \nabla \times \vec{F}$. Entsteht β aus α durch zyklische Vertauschung der Basisvektoren, so ergeben sich die β -Koordinaten von $\text{rot}_{\beta} \vec{F}$ durch entsprechende Vertauschung aus den α -Koordinaten und es gilt $\text{rot}_{\alpha} \vec{F} = \text{rot}_{\beta} \vec{F}$.

Achtung: Die Definition von $\text{rot}_{\alpha} \vec{F}$ durch Koordinaten (wie hier angegeben) ist basisabhängig. Der Bezug zur basisunabhängigen Rotation um eine Achse ist nur ein Hinweis auf eine mögliche Unabhängigkeit. In der gängigen Literatur zur Analysis wird die Frage nicht diskutiert. Das soll im Folgenden geschehen (vgl. Skript Mathe II für MB von U. Reif).

Im Raum, d.h. einem orientierten 3-dimensionalen euklidischen Vektorraum sei eine lineare Abbildung ϕ gegeben. Ist κ eine positiv orientierter ON-Basis $\vec{e}_1, \vec{e}_2, \vec{e}_3$. und A die zugehörige Matrix von ϕ so definieren wir die lineare Abbildung ρ_A durch

$$\rho_A(\vec{x})^{\kappa} = (A - A^t) \vec{x}^{\kappa}$$

ρ_A hängt aber nicht von der Basis ab: ist λ eine zweite positiv orientierte ON-Basis und $\vec{x}^\kappa = S\vec{x}^\lambda$, so gilt

$$\rho_A(\vec{x})^\lambda = S^t \rho_A(\vec{x})^\kappa = S^t(A - A^t)S\vec{x}^\lambda = (S^tAS - S^tA^tS)\vec{x}^\lambda = (S^tAS - (S^tAS)^t)\vec{x}^\lambda$$

und S^tAS ist die Matrix von ϕ bzgl. λ . Also hat man wohldefiniertes $\rho_\phi = \rho_A$.

Andererseits hat man zu jeden Vektor \vec{r} die lineare Abbildung $\rho_{\vec{r}}$ definiert durch

$$\rho_{\vec{r}}(\vec{x}) = \vec{r} \times \vec{x}$$

und \vec{r} ist durch $\rho_{\vec{r}}$ eindeutig bestimmt: ist $\vec{r} \times \vec{x} = \vec{s} \times \vec{x}$ für alle \vec{x} , so $(\vec{r} - \vec{s}) \times \vec{x} = \vec{0}$ für alle \vec{x} und somit $\vec{r} - \vec{s} = \vec{0}$.

Definieren wir nun bzgl. κ den Vektor $\text{rot}_\kappa A$ durch

$$(\text{rot}_\kappa A)^\kappa = \begin{pmatrix} a_{32} - a_{23} \\ a_{13} - a_{31} \\ a_{21} - a_{12} \end{pmatrix}$$

so gilt, wie man leicht durch Einsetzen der \vec{e}_i nachprüft,

$$\rho_\phi = \rho_A = \rho_{\text{rot}_\kappa A}$$

Also ist $\text{rot}_\kappa A$ nur von ϕ , nicht jedoch von κ abhängig und wir haben

- zu jeder linearen Abbildung ϕ des Raumes einen wohldefinierten Vektor $\text{rot}\phi$ so, dass bzgl. jeder positiv orientierten ON-Basis κ gilt: ist A die Matrix von ϕ bzgl. κ , so gilt

$$(\text{rot}\phi)^\kappa = \begin{pmatrix} a_{32} - a_{23} \\ a_{13} - a_{31} \\ a_{21} - a_{12} \end{pmatrix}$$

Korollar 27.8 Ist \vec{F} ein C^{k+1} -Vektorfeld \vec{F} auf der offenen Menge U des Raumes, so haben wir auf U das wohldefinierte C^k -Vektorfeld

$$(\text{rot}\vec{F})(\vec{x}) = \text{rot}((D\vec{F})(\vec{x}))$$

Ohne Matrizen geht es so: Zu ϕ gibt es die eindeutig bestimmte *adjungierte* Abbildung ϕ^* mit $\phi(\vec{x}) \cdot \phi^*(\vec{y}) = \phi^*(\vec{x}) \cdot \phi(\vec{y})$ für alle \vec{x}, \vec{y} (mit Matrix A^t) und man hat $\rho_\phi = \phi - \phi^*$.

Zur Motivation:

$$\phi = \frac{1}{2}(\phi + \phi^*) + \frac{1}{2}(\phi - \phi^*)$$

ist eine Zerlegung $\phi = \psi + \chi$ mit selbstadjungierten (symmetrischen) ψ und antisymmetrischem χ , d.h. $\chi^* = -\chi$. Für ein stetig differenzierbares Vektorfeld \vec{F} hat man die entsprechende Zerlegung

$$D\vec{F} = D\vec{G} + D\vec{R}$$

und erhält $D\vec{G}$ wegen der Symmetrie als Gradientenfeld. D.h. alle Rotation steckt in $D\vec{R}$.

27.10 Schwarze Produktregel

In der Ebene gilt für ein C^1 -Vektorfeld \vec{F} und C^1 -skalare Funktion h nach der Produktregel

$$\operatorname{div}(h\vec{F}) = \sum_i \frac{\partial h F_i}{\partial x_i} = \sum_i \frac{\partial h}{\partial x_i} F_i + \sum_i h \frac{\partial F_i}{\partial x_i} = (\operatorname{grad} h) \cdot \vec{F} = h \operatorname{div} \vec{F}$$

$$\operatorname{rot}(h\vec{F}) = \operatorname{div}(-h\vec{F}^\perp) = (\operatorname{grad} h) \cdot (-\vec{F}^\perp) + h \operatorname{rot} \vec{F}$$

Mit Schwarz folgt

- Ist \vec{F} ein Gradientenfeld, so gilt $\operatorname{rot} \vec{F} = 0$ und $\operatorname{rot}(h\vec{F}) = (\operatorname{grad} h) \cdot (-\vec{F}^\perp)$

27.11 Stokes

Satz 27.9 Sei $\alpha : \vec{e}_1, \vec{e}_2, \vec{e}_3$ ein positiv orientiertes ON-Koordinatensystem des Raumes und (Γ, K) ein grüner Bereich bzgl. $\{\vec{e}_1, \vec{e}_2\}$ in der von diesen aufgespannten Ebene, (ϕ, K) eine Parametrisierung mit ϕ zweimal stetig partiell differenzierbar. Sei $U \subseteq \mathbb{R}^3$ offen, $\operatorname{Spur}(\phi, K) \subseteq U$ und $\vec{F} : U \rightarrow \mathbb{R}^3$ stetig differenzierbar. Dann gilt

$$\int_{u \in K} \operatorname{rot}_\alpha \vec{F}(\phi(u)) \cdot \vec{N}_\phi(u) = \int_{\phi \circ \Gamma} \vec{F} \cdot d\vec{x}$$

kurz

$$\int_{(\phi, K)^\perp} \operatorname{rot} \vec{F} = \int_{\phi \circ \Gamma} \vec{F}$$

mystifiziert zu

$$\int_S \operatorname{rot} \vec{F} \cdot \vec{n} \, d\sigma = \int_{\partial S} \vec{F} \cdot d\vec{x}$$

Beweis. Sei

$$\vec{F} \hat{=} \begin{pmatrix} f \\ 0 \\ 0 \end{pmatrix}$$

Dann

$$\begin{aligned} & \int_{\phi \circ \Gamma} \vec{F} \cdot d\vec{x} \\ &= \int_a^b \begin{pmatrix} f(\phi(\vec{x}(t))) \\ 0 \\ 0 \end{pmatrix} \cdot \frac{\partial(\phi \circ \vec{x})}{\partial t} dt \quad \text{Definition} \\ &= \int_a^b f(\phi(\vec{x}(t))) \left(\frac{\partial \phi \circ \vec{x}}{\partial t} \right)^{(1)} dt \quad \text{Skalarprodukt} \\ &= \int_a^b f(\phi(\vec{x}(t))) \operatorname{grad} \phi^{(1)}(\vec{x}) \cdot \frac{\partial \vec{x}}{\partial t} dt \quad \text{Kettenregel} \\ &= \int_\Gamma (f \circ \phi) \operatorname{grad} \phi^{(1)}(\vec{x}) d\vec{x} \quad \text{Definition} \end{aligned}$$

$$= \int_{u \in K} \operatorname{rot}((f \circ \phi) \operatorname{grad} \phi^{(1)})(u) \quad \text{Green}$$

Nun

$$\begin{aligned} & \operatorname{rot}((f \circ \phi) \operatorname{grad} \phi^{(1)}) \\ &= \frac{\partial(f \circ \phi)}{\partial x_1} \cdot \frac{\partial \phi^{(1)}}{\partial x_2} - \frac{\partial(f \circ \phi)}{\partial x_2} \cdot \frac{\partial \phi^{(1)}}{\partial x_1} \quad \text{schwarze Produktregel} \\ &= (\operatorname{grad} f \cdot \vec{\partial}_1 \phi) \frac{\partial \phi^{(1)}}{\partial x_2} - (\operatorname{grad} f \cdot \vec{\partial}_2 \phi) \frac{\partial \phi^{(1)}}{\partial x_1} \quad \text{Kettenregel} \\ &= \operatorname{grad} f \cdot \left(\vec{\partial}_1 \phi \frac{\partial \phi^{(1)}}{\partial x_2} - \vec{\partial}_2 \phi \frac{\partial \phi^{(1)}}{\partial x_1} \right) \quad \text{Linearität des Skalarprodukts} \\ &= \begin{pmatrix} 0 \\ \frac{\partial f}{\partial x_3} \\ -\frac{\partial f}{\partial x_2} \end{pmatrix} \cdot (\vec{\partial}_1 \phi \times \vec{\partial}_2 \phi) \quad \text{Lemma 27.1} \\ &= (\operatorname{rot} \vec{F}) \cdot \vec{N}_\phi \quad \text{Definition} \end{aligned}$$

Da Skalarprodukte, Normalenvektor und Integrale vom Koordinatensystem unabhängig sind und rot von positiver Permutation der Koordinatenachsen, folgt die Behauptung ebenso für \vec{F} der Form

$$\vec{F} \hat{=} \begin{pmatrix} 0 \\ f \\ 0 \end{pmatrix}, \quad \text{wie} \quad \vec{F} \hat{=} \begin{pmatrix} 0 \\ 0 \\ f \end{pmatrix}$$

und damit wegen der Linearität für alle \vec{F} . \square

28 Divergenz und Normalenintegrale im Raum

Der Satz von Ostrogradski-Gauß in Raum besagt, dass (unter geeigneten Voraussetzungen) das Normalenintegral (Oberflächenintegral zweiter Art) eines Vektorfeldes über eine Oberfläche gleich dem Integral der Divergenz dieses Feldes über das durch die Fläche begrenzte Volumen ist.

Eine adäquate Behandlung ist auf Untermannigfaltigkeiten und Differentialformen zu gründen. Eine Annäherung dazu findet man in Hildebrandt, Analysis II. Alternativ kann man versuchen, nur solche Volumina und Oberflächen zu betrachten, die in bestimmter Form beschrieben sind. In der Lehrbuchliteratur der letzten 40 Jahre werden solche Versuche unter dem Stichwort “ C^1 -Normalbereiche” gemacht. Die beiden Hauptprobleme, die dabei auftreten sind

- (i) Offensichtlich nicht schlüssige “Beweise” bzw. falsche “Sätze”
- (ii) Mangelnde Anwendbarkeit auf naheliegende Beispiele, auch wenn zusammengesetzte Bereiche benutzt.

Die “Beweise” behandeln korrekt den Fall, dass das Feld nur in einer Komponente von 0 verschieden ist, und benutzen dazu die durch den Normalbereich in dieser Richtung gegebene Parametrisierung der “Oberfläche”. Und stellen dann lakonisch fest, dass alles bewiesen sei. Das Defizit lässt sich leicht durch eine Zusatzforderung beheben: Die 3 sich kanonisch ergebenden Parametrisierungen der Oberfläche müssen als äquivalent vorausgesetzt werden.

Die Voraussetzung, dass die “Deckel” durch C^1 -Funktionen gegeben werden sollen ist ebenso hinderlich wie überflüssig: Knicke in Deckeln sind bei der Diskussion von Beispielen kaum zu vermeiden und sind für den Beweis unschädlich.

Im Folgenden wird versucht, ein Konzept von “Normalbereichen”, genannt “grüne Bereiche” zu entwickeln, das erlauben soll, Standardbeispiele korrekt und vollständig zu diskutieren.

Um eine einfacher zu beweisende Version der Integralsubstitutionsregel anwenden zu können, müssen wir mit der oben eingeführten unhandlichen Begriff von Parametrisierung arbeiten. Einen handlicheren Begriff, der aber mehr Aufwand bei der Substitutionsregel erfordert, findet man in Walter, Analysis 2.

28.1 Diffeomorphismus

Wir erinnern an das folgende Ergebnis aus Kap.16 - vgl. Skript von Prof. Hieber, Analysis-II-ss08.

Korollar 28.1 *Sei $U \subseteq \mathbb{R}^n$ offen und $\sigma : U \rightarrow \mathbb{R}^n$ injektiv und stetig differenzierbar mit $\det D\sigma \neq 0$. Dann ist $\sigma(U)$ offen und $\sigma : U \rightarrow \sigma(U)$ ein Diffeomorphismus und für die Umkehrabbildung gilt $D(\sigma^{-1}) = (D\sigma)^{-1}$. Diese ist insbesondere ebenfalls stetig.*

Beweis. Nach dem Satz über die Umkehrabbildung gibt es zu jedem $x \in U$ offenes $x \in U_x \subseteq U$ so, dass $\sigma(U_x)$ offen ist und $\sigma|_{U_x} : U_x \rightarrow \sigma(U_x)$ Diffeomorphismus. Also $\sigma(U) = \bigcup_{x \in U} \sigma(U_x)$ offen und $\sigma^{-1} = \bigcup_{x \in U} (\sigma|_{U_x})^{-1}$ Diffeomorphismus und $D(\sigma^{-1}) = (D\sigma)^{-1}$ nach der Kettenregel. \square

28.2 Treue Parametrisierungen

Eine einfache Parametrisierung (ϕ, K) ist *treu*, wenn

- ϕ injektiv ist
- $\text{Rang } D\phi(u) = 2$ für $u \in \text{dom } \phi$
- es gibt wegzusammenhängende offene Menge \hat{K} mit $K \subseteq \hat{K} \subseteq \text{dom } \phi$

Beachte, dass $W_1 \cap W_2 \supseteq K$ wegzusammenhängend und offen ist, wenn das für die $W_i \supseteq K$ gilt. Die Definition von *treu* überträgt sich entsprechend auf Zusammensetzungen.

28.3 Umorientierte Pflache

Zu einer Parameterdarstellung (ϕ, K) definieren wir die *umorientierte*

$$\ominus(\phi, K) = (\phi \circ \tau, \tau(K)) \quad \text{mit } \tau \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = \begin{pmatrix} u_1 \\ -u_2 \end{pmatrix}$$

Dann gilt

- $\text{Spur } \ominus(\phi, K) = \text{Spur}(\phi, K)$

- $\vec{N}_{\phi \circ \tau}(\phi \circ \tau(\tau(u))) = -\vec{N}_\phi(\phi(u)) \quad u \in K$

-

$$\int_K \vec{F} \cdot \vec{N}_\phi(\phi(u)) = - \int_{\tau(K)} \vec{F} \cdot \vec{N}_{\phi \circ \tau}(\phi \circ \tau(v))$$

- Sind (ϕ_1, K_1) und (ϕ_2, K_2) aquivalent, so sind auch $\ominus(\phi_1, K_1)$ und $\ominus(\phi_2, K_2)$ aquivalent.

Wir konnen demnach auch fur eine Pflache S die *umorientierte* definieren

$$(\phi, K) \in S \Leftrightarrow \ominus(\phi, K) \in \ominus S$$

und haben

$$\int_{S^\perp} \vec{F} = - \int_{(\ominus S)^\perp} \vec{F}$$

28.4 Graphen und Zylinder

Sei κ ein Koordinatensystem des Raumes mit positiv orientierter ON-Basis $\vec{e}_1, \vec{e}_2, \vec{e}_3$.

Sei U eine wegzusammenhangende offene Teilmenge in der x_1, x_2 -Ebene bzgl. κ und $\beta : U \rightarrow \mathbb{R}$ eine stetig differenzierbare Abbildung

$$\phi \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \\ \beta \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \end{pmatrix}$$

Dann ist $\phi(U)$ ein offener κ -1-2-Graph, und, fur kompaktes $K \subseteq U$, (ϕ, K) eine einfache treue Parametrisierung, die zugehorige *kanonische Parametrisierung* von $\phi(K)$ als κ -1-2-Graph. Mit $U' = \left\{ \begin{pmatrix} x_1 \\ -x_2 \end{pmatrix} \mid \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in U \right\}$ und kompaktem $K' \subseteq U'$ und

$$\psi \begin{pmatrix} x_1 \\ -x_2 \end{pmatrix} = \begin{pmatrix} x_1 \\ -x_2 \\ \beta \begin{pmatrix} x_1 \\ -x_2 \end{pmatrix} \end{pmatrix}$$

erhalt man die *kanonische Parametrisierung* von $\psi(K')$ als κ -1-2-Anti-Graph.

Sei U eine offene Teilmenge in der x_1, x_2 -Ebene bzgl. κ und $\alpha, \beta : U \rightarrow \mathbb{R}$ stetig differenzierbare Abbildungen, $\alpha \leq \beta$. Sei $V \subseteq \mathbb{R}$, $\gamma : V \rightarrow U$ stetig differenzierbar und $\frac{\partial \gamma}{\partial t} \neq 0$ für alle $t \in V$. Sei $\varepsilon > 0$ und

$$W = \left\{ \begin{pmatrix} t \\ z \end{pmatrix} \mid t \in V, \alpha\gamma(t) - \varepsilon < z < \beta\gamma t + \varepsilon \right\}, \quad \zeta \begin{pmatrix} t \\ z \end{pmatrix} = \begin{pmatrix} \gamma t \\ z \end{pmatrix} \quad \left(\begin{pmatrix} t \\ z \end{pmatrix} \in W \right)$$

Dann ist $\zeta(W)$ ein offener κ -1-2-Zylinder. Sei $[a, b] \subseteq V$, $\Gamma = \gamma|[a, b]$. Dann ist

$$(\zeta, K), \quad K = \left\{ \begin{pmatrix} t \\ z \end{pmatrix} \in W \mid t \in [a, b], \alpha\gamma(t) \leq z < \beta\gamma t \right\}$$

eine treue Parametrisierung, die *kanonische Parametrisierung* von $\zeta(K)$ als κ -1-2-Zylinder bzgl. Γ, α, β . Entsprechend werden offene κ -2-3 und κ -3-1 Graphen, Anti-Graphen und Zylinder und ihre kanonischen Parametrisierungen definiert.

Zum Beweis der Treue: Injektivität und Rangbedingung sind offensichtlich erfüllt. Zu \hat{K} : Zylinder kann man problemlos noch oben und unten verlängern. Bei den Graphen wähle man eine Darstellung von U als Vereinigung offener Kreisscheiben. Wegen der Kompaktheit gibt es dazu eine endliche Teilüberdeckung von K , deren Vereinigung ist das gewünschte \hat{K} . \square

28.5 Drittel-grün

Für die Zusammensetzung grüner Bereiche (Γ_i, B_i) der Ebene nach Kap.26.11 benutzen wir die Notation

$$(\Gamma, B) = \bigoplus_i (\Gamma_i, B_i)$$

Gegeben sei nun ein positiv orientiertes ON-Koordinatensystem $\kappa : O, \vec{e}_1, \vec{e}_2, \vec{e}_3$ des Raumes. Sei (Γ, B) eine grüne Teilmenge der κ -1-2-Ebene, $\alpha, \beta : B \rightarrow \mathbb{R}$ stetig, $\alpha \leq \beta$. Das durch (Γ, α, β) definierte *Volumen* ist

$$V(\Gamma, \alpha, \beta) = \left\{ \begin{pmatrix} u_1 \\ u_2 \\ z \end{pmatrix} \mid \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} \in B \right\}$$

Sei (ϕ, K) eine aus einfachen (ϕ_i^j, K_i^j) , ($i \in I_j, j = 1, 2, 3$) zusammengesetzte treue Parametrisierung (ϕ, K) . Dann ist das Quadrupel $((\phi, K), \Gamma, \alpha, \beta)$ κ -1-2-grün, wenn gilt

- $(\Gamma, B) = \bigoplus_{i \in I_2} (\Gamma_i^2, K_i^2)$ mit grünen (Γ_i^2, K_i^2)
- Die (ϕ_i^2, K_i^2) sind kanonische Parametrisierungen von κ -1-2-Graphen β_i mit $\beta|K_i^2 = \beta_i|K_i^2$
- $(\Gamma, B) = \bigoplus_{i \in I_1} (\Gamma_i^1, \tau K_i^1)$ mit grünen $(\Gamma_i^1, \tau K_i^1)$
- Die (ϕ_i^1, K_i^1) sind kanonische Parametrisierungen von κ -1-2-Anti-Graphen α_i mit $\alpha\tau|K_i^1 = \alpha_i\tau|K_i^1$

- $\Gamma = \bigoplus_{i \in I_3} \Gamma_i^3 \oplus \bigoplus \Delta_k$ wobei $\alpha \Delta_k = \beta \Delta_k$ und es zu jedem $i \in I_3$ $i_j \in I_j$ gibt mit Γ_i^3 Teilstück von $\Gamma_{i_j}^j$ für $j = 1, 2$
- Die (ϕ_i^3, K_i^3) sind kanonische Parametrisierungen von κ -1-2-Zylindern bzgl. $\Gamma_i^3, \alpha_i, \beta_i$ mit $\alpha_i \Gamma_i^3 = \alpha \Gamma_i^3$ und $\beta_i \Gamma_i^3 = \beta \Gamma_i^3$

2 – 3 und 3 – 1 grüne Quadrupel und zugehörige Volumina werden analog definiert.

28.6 Mini-Gauß

Lemma 28.2 Sei $((\phi, K), \Gamma, \alpha, \beta)$ ein κ -1-2-grünes Quadrupel mit Volumen V und $\vec{F}(\vec{x}) = f(\vec{x})\vec{e}_3$ stetig differenzierbar auf offener Umgebung von V . Dann gilt

$$\int_V \operatorname{div} F = \int_{(\phi, K)^\perp} \vec{F}$$

Beweis. Wir nehmen Koordinaten bzgl. κ .

$$\vec{F}(\vec{x}) = \begin{pmatrix} 0 \\ 0 \\ f(\vec{x}) \end{pmatrix}, \quad \text{also } \operatorname{div} \vec{F} = \frac{\partial f}{\partial x_3}$$

Mit Fubini und Hauptsatz haben wir

$$\int_V \operatorname{div} \vec{F} = \int_K \int_{\alpha(u)}^{\beta(u)} \frac{\partial f}{\partial x_3} = \int_K (f(u, \beta(u)) - f(u, \alpha(u)))$$

Wir setzen

$$(\phi^j, K^j) = \bigoplus_{i \in I_j} (\phi_i^j, K_i^j)$$

Nun

$$\int_K f(u, \beta(u)) = \int_K \vec{F}(u) \cdot \vec{N}_{\phi_2}(u) = \int_{(\phi_2, K_2)^\perp} \vec{F}$$

da der Normalenvektor 3.Komponente 1 hat (abgesehen von Nullmenge, wo er Null ist). Andererseits

$$- \int_K f(u, \alpha(u)) = \int_{K_1} \vec{F}(v) \cdot \vec{N}_{\phi_1}(v) = \int_{(\phi_1, K_1)^\perp} \vec{F}$$

da hier der Normalenvektor 3.Komponente -1 hat. Schliesslich

$$\int_{(\phi_3, K_3)^\perp} \vec{F} = 0$$

da hier der Normalenvektor 3.Komponente 0 hat. Also

$$\int_V \operatorname{div} \vec{F} = \sum_{i=1}^3 \int_{(\phi_i, K_i)^\perp} \vec{F} = \int_{S^\perp} \vec{F} \quad \square$$

Korollar 28.3 Gegeben sei ein Koordinatensystem des Raumes mit positiv orientierter ON-Basis $\vec{e}_i, \vec{e}_j, \vec{e}_k$ und bzgl. dieser ein i - j -grünes Quadrupel $((\phi, K), \Gamma, \alpha, \beta)$ mit zugehörigem Volumen V . Seien f, g auf offener Umgebung von V stetig differenzierbare skalarwertige Funktionen. Dann gilt

$$\int_V f \frac{\partial g}{\partial x_k} = - \int_V \frac{\partial f}{\partial x_k} g + \int_{(\phi, K)} fg(\vec{N}_\phi)^{(k)}$$

Beweis. Betrachte $\vec{F}(\vec{x}) = f(\vec{x}) \cdot g(\vec{x})\vec{e}_k$. Dann

$$\operatorname{div} \vec{F} = f \frac{\partial g}{\partial x_k} + \frac{\partial f}{\partial x_k} g, \quad \vec{F} \cdot \vec{N}_\phi = fg(\vec{N}_\phi)^{(k)}$$

und der Beweis ergibt sich nun aus dem von Satz 28.5. \square

28.7 Grüne Bereiche

Sei S eine Pfläche und $V \subseteq \mathbb{R}^3$. Dann ist (S, V) bzgl. $\kappa : O, \vec{e}_1, \vec{e}_2, \vec{e}_3$ ein *grüner Elementarbereich* des Raumes mit der *Oberpfläche* S und dem *Volumen* V , wenn es κ - i - j -grüne Qadrupel $(\phi_{ij}, K_{ij}), \Gamma_{ij}, \alpha_{ij}, b_{ij}$ gibt mit Volumen V und $(\phi_{ij}, K_{ij}) \in S$ für alle $(i, j) \in \{(1, 2), (2, 3), (3, 1)\}$

Induktive Definition: (S, V) ist *grüner Bereich* des Raumes

- falls (S, V) grüner Elementarbereich bzgl. eines (passenden) Koordinatensystems des Raumes
- oder falls es grüne Bereiche (S_i, V_i) , $(i \in I)$ gibt mit
 - $S_i = \bigoplus_{j \in I} S_{ij}$, $\phi_{ij}, K_{ij} \in S_{ij}$
 - S_{ij} äquivalent $\ominus S_{ji}$ oder $\mu(K_{ij}) = 0 = \mu(K_{ji})$ für $i \neq j$
 - $\mu(K_{ij} \cap K_{kl}) = 0$ für $(i, j) \neq (k, l)$.
 - $V_i \cap V_j \subseteq \operatorname{Spur} S_{ij}$ und $\mu(V_i \cap V_j) = 0$ für $i \neq j$
 - $(S, V) = (\bigoplus_{i \in I} S_{ii}, \bigcup_{i \in I} V_i)$

Im zweiten und dritten Fall ist (S, V) durch *Verklebung* entstandener grüner Bereich und wir notieren $(S, V) = \bigoplus_{i \in I} (S_i, V_i)$ was aber auch die Erinnerung an die Bereiche (S_i, V_i) und die Wahl der S_{ij} einschließt.

Korollar 28.4 Sei f bzw. \vec{F} auf einer offenen Umgebung von V stetig differenzierbar. Ist (S, V) Verklebung der grünen Bereiche (S_i, V_i) so gilt

$$\int_V f = \sum_{i \in I} \int_{V_i} f, \quad \int_{S^\perp} \vec{F} = \sum_{i \in I} \int_{S_i^\perp} \vec{F}$$

28.8 Gauß

Satz 28.5 *Ist (S, V) ein grüner Bereich des Raumes, $U \subseteq \mathbb{R}^3$ offen, $V \subseteq U$ und $\vec{F} : U \rightarrow \mathbb{R}^3$ stetig differenzierbar, so gilt*

$$\int_V \operatorname{div} \vec{F} = \int_{S^\perp} \vec{F}$$

Beweis. Es genügt grüne Elementarbereiche zu betrachten. Sei also (S, V) grüner Elementarbereich bzgl. $\vec{e}_1, \vec{e}_2, \vec{e}_3$. Wegen der Linearität von div , Skalarprodukt und Integralen genügt es $\vec{F} = f\vec{e}_i$ zu betrachten. Sei z.B. $i = 3$. Da wir zu jeder Koordinatenebene eine Parametrisierung in S haben, können wir uns die passende aussuchen: die durch den grünen Bereich (Γ, K) der Ebene und das grüne 1-2-Quadrupel $(\phi, K), \Gamma, \alpha, \beta$ gegebene mit $(\phi, K) \in S$. Dafür folgt die Behauptung aus Lemma 28.2. \square

28.9 Äquivalenz bei Graphen und Zylindern

Es bleibt die Aufgabe, die Grünheit jeweils nachzuweisen. Dabei hilft das Folgende.

Einfache treue Parametrisierungen. (ϕ, K) und (ψ, H) sind *stark äquivalent*, wenn

- $\operatorname{Spur}(\phi, K) = \operatorname{Spur}(\psi, H)$
- es gibt offene wegzusammenhängende \hat{K} und \hat{H} mit
 - $K \subseteq \hat{K} \subseteq \operatorname{dom}\phi$ und $H \subseteq \hat{H} \subseteq \operatorname{dom}\psi$
 - $\phi(\hat{K}) = \psi(\hat{H})$ κ -i-j-Graph, -Anti-Graph oder -Zylinder bzgl. eines geeigneten Koordinatensystems ist
- $\vec{N}_\phi(u) = \lambda \vec{N}_\psi(v)$ mit $\lambda > 0$ falls für mindestens ein Paar u, v mit $\phi(u) = \psi(v)$.

Entsprechend für Zusammensetzungen.

Lemma 28.6 *Stark äquivalente Parametrisierungen sind äquivalent*

Beweis. Es genügt, dem einfachen Fall zu betrachten. Wir betrachten den Fall eines offenen 1-2-Graphen G von β . Nach Voraussetzung haben wir G als bijektives Bild von \hat{K} unter ϕ . Setze

$$\sigma(u) = \begin{pmatrix} \phi^{(1)}(u) \\ \phi^{(2)}(u) \end{pmatrix} \quad u \in \hat{K}, \quad \text{also } \phi(u) = \begin{pmatrix} \sigma(u) \\ \beta(\sigma(u)) \end{pmatrix}$$

Dann ist σ injektiv und stetig differenzierbar und nach der Kettenregel gilt

$$J_\phi = \begin{pmatrix} J_\sigma \\ J_\beta J_\sigma \end{pmatrix}$$

Wegen $\operatorname{Rang} J_\phi = 2$ folgt $\det J_\sigma \neq 0$ und es gilt

$$(\vec{N}_\phi)^{(3)} = (\vec{\partial}_1 \phi \times \vec{\partial}_2 \phi)^{(3)} = \det J_\sigma = \det D\sigma$$

Da \hat{K} wegzusammenhängend ist, gilt auf ganz \hat{K} entweder $\det D\sigma > 0$ oder $\det D\sigma < 0$.

Entsprechend gibt es τ definiert auf \hat{H} . Dann ist $\rho = \sigma^{-1} \circ \tau$ Diffeomorphismus von \hat{H} auf \hat{K} und für das obige Paar u, v gilt

$$\vec{N}_\psi(v) \det \rho = \vec{N}_\phi(u) = \lambda \vec{N}_\psi(v)$$

also $\det \rho > 0$ und es liegt Äquivalenz vor.

Im Falle der Anti-Graphen benutzt man, dass der 1-2-Antigraph im entgegengesetzt orientierten Koordinatensystem $\vec{e}_2, \vec{e}_1, \vec{e}_3$ zum 2-1-Graphen wird und dass bei diesem Übergang der Normalenvektor \vec{N}_ϕ umgekehrt wird zu $-\vec{N}_\phi$. Dann folgt der Beweis nach dem Graphen-Fall. Im Falle eines 1-2-Zylinders haben wir den Diffeomorphismus σ mit

$$\sigma(\vec{u}) = \begin{pmatrix} t \\ x_3 \end{pmatrix} \Leftrightarrow \phi(\vec{u}) = \zeta \begin{pmatrix} t \\ x_3 \end{pmatrix}$$

Die Äquivalenz folgt nun wie eben. \square

Korollar 28.7 Sei bzgl. eines Koordinatensystems ein für eine Wahl i_0 - j_0 -grünes Quadrupel $((\phi, K), \Gamma, \alpha, \beta)$ mit Volumen V gegeben, und finde man ein neues Koordinatensystem κ so, dass die (ϕ_j^j, K_i^j) bzgl. jeder der 3 Koordinatenebenen in 1-1-Entsprechung zu den Bausteinen eines i - j -grünem Quadrupels stehen so, dass entsprechende Parametrisierungen stark äquivalent sind. Dann ist $((\phi, K), V)$ grüner Elementarbereich.

Korollar 28.8 Seien (ϕ, K) wie im Lemma 28.6 und $U \supseteq \text{Spur}(\phi, K)$ offene Teilmenge des Raumes und f auf U stetig. Seien die Basisvektoren $\vec{e}_i, \vec{e}_j, \vec{e}_k$ aus κ positiv orientiert und $\text{Spur}(\phi, k)$ κ - i - j Graph bzgl. β , Anti-Graph bzgl. α bzw. Zylinder. Dann gilt jeweils

$$\begin{aligned} \int_{u \in K} f(\phi(u)) \vec{e}_3 \cdot \vec{N}_\phi(u) &= \int_{x \in L} f(x, \beta(x)) \\ \int_{u \in K} f(\phi(u)) \vec{e}_3 \cdot \vec{N}_\phi(u) &= - \int_{x \in L} f(x, \alpha(x)) \\ \int_{u \in K} f(\phi(u)) \vec{e}_3 \cdot \vec{N}_\phi(u) &= 0 \end{aligned}$$

Hier wird nur die Stetigkeit von α bzw. β benötigt. Beweis. Für β liefert die Anwendung der Substitutionsregel auf $g(x) = f(x, \beta(x))$ die Behauptung

$$\int_{u \in K} f(\phi(u)) \cdot (\vec{N}_\phi)^{(3)}(u) = \int_{u \in K} f(\sigma(u), \beta(\sigma(u))) \cdot \det D\sigma(u) = \int_{x \in L} f(x, \beta(x))$$

und der Fall des Anti-graphen folgt wegen der Umkehrung der Normalenvektoren. Im Falle eines 1-2-Zylinders betrachten wir für festes t_0 den Weg

$$\vec{y}(x_3) = \phi \sigma^{-1} \begin{pmatrix} t_0 \\ x_3 \end{pmatrix} = \begin{pmatrix} \Gamma(t_0)^{(1)} \\ \Gamma(t_0)^{(2)} \\ x_3 \end{pmatrix}, \quad \alpha(\Gamma(t_0)) < x_3 < \beta(\Gamma(t_0))$$

Hier gilt

$$\frac{\partial \vec{y}}{\partial x_3} = \vec{e}_3 \quad \text{und somit} \quad \vec{e}_3 \cdot \vec{N}_\phi = 0 \quad \square$$

28.10 Anwendbarer Gauß

Will man z.B. die Sphäre durch Kugelkoordinaten parametrisieren und diese Parametrisierung bei der Bestimmung der Oberflächenintegrals 2. Art benutzen, so hat man erstens zu zeigen, dass die Sphäre die genannten Bedingungen des Satzes von Gauss erfüllt, und zweitens, dass das die dazu geeigneten Parametrisierungen zu demselben Wert des Integrals führen.

Sei (S, V) ein grüner Bereich des Raumes und $(\phi, K) = \bigoplus_i (\phi_i, K_i) \in S$ eine treue Parametrisierung so, dass $\phi_i(\hat{K}_i)$ Graph, Anti-Graph oder Zylinder ist (also z.B. zu einem bei der Konstruktion von (S, V) benutzten Elementarbereich gehört). Ist (ψ, H) zu (ϕ, K) stark äquivalent, so ist (ψ, H) eine *grüne Parametrisierung* von (S, V) .

Satz 28.9 Sei (S, V) grüner Bereich des Raumes und \vec{F} auf offener Umgebung von V stetig differenzierbar. Dann gilt für jede grüne Parametrisierung (ψ, H) von (S, V)

$$\int_{(\psi, H)^\perp} \vec{F} = \int_{S^\perp} \vec{F} = \int_V \operatorname{div} \vec{F}$$

Beweis. Es genügt, Elementarbereiche zu betrachten. Nach Lemma 28.6 ist (ψ, H) zur kanonischen Parametrisierung äquivalent. Das ergibt die erste Gleichheit. Die zweite ist Satz 28.5. \square

Korollar 28.10 Sei (S, V) ein grüner Bereich des Raumes und $(\phi, K) \in S$ eine treue Parametrisierung. Sei (ψ, H) eine Parametrisierung mit $\operatorname{Spur}((\psi, H) = V$. Zu jedem $n \in \mathbb{N}$ gebe es Zerlegungen

$$(\phi, K) = (\phi_{n1}, K_{n1}) \oplus (\phi_{n2}, K_{n1}), \quad (\psi, H) = (\psi_{n1}, H_{n1}) \oplus (\psi_{n2}, H_{n1})$$

so, dass gilt

- die (ϕ_{ni}, K_{ni}) sind grün für grüne Bereiche (S_{ni}, V_{ni}) mit

$$(S, V) = (S_{n1}, V_{n1}) \oplus (S_{n2}, V_{n2})$$

- (ψ_{n1}, H_{n1}) ist grüne Parametrisierung von (S_{n1}, V_{n1})

-

$$\lim_{n \rightarrow \infty} \mu(K_{n2}) = \lim_{n \rightarrow \infty} \mu(V_{n2}) = \lim_{n \rightarrow \infty} \mu(H_{n2}) = 0$$

Dann gilt für jedes auf offener Umgebung von V stetig differenzierbare Vektorfeld \vec{F}

$$\int_{(\psi, H)^\perp} \vec{F} = \int_V \operatorname{div} \vec{F}$$

Beweis. Nach grünem Gauß

$$\int_{(\psi_{n1}, H_{n1})^\perp} \vec{F} = \int_{(\phi_{n1}, K_{n1})^\perp} \vec{F} = \int_{V_n} \operatorname{div} \vec{F}$$

Wegen Linearität des Integrals

$$\int_{(\psi, H)^\perp} \vec{F} = \int_{(\psi_{n1}, H_{n1})^\perp} \vec{F} + \int_{(\psi_{n2}, H_{n2})^\perp} \vec{F}$$

$$\int_V \operatorname{div} \vec{F} = \int_{V_{n1}} \operatorname{div} \vec{F} + \int_{V_{n2}} \operatorname{div} \vec{F}$$

Nach Voraussetzung, gehen die Werte der zweiten Summanden jeweils gegen 0 und die Bauhauptung folgt. \square

28.11 Alternativen zur Flächendefinition

Statt bei der Definition einer Parameterdarstellung (ϕ, K) zu verlangen, dass ϕ auf einer offenen Umgebung von K definiert und stetig differenzierbar ist, kann man ϕ nur auf K definiert annehmen und verlangen (vgl. Walter, Analysis II, Kap.8)

- K messbar und beschränkt
- ϕ auf dem Inneren K° von K injektiv und stetig differenzierbar mit $\operatorname{Rang} J_\phi = 2$
- $\phi(K^\circ) \cap \phi(\partial K) = \emptyset$
- ϕ auf K Lipschitz stetig

Äquivalenz wird dann gegeben durch

- $\sigma : H^\circ \rightarrow K^\circ$ Diffeomorphismus
- $\det D\sigma > 0$ oder $\det D\sigma < 0$
- $\psi = \phi \circ \sigma$

und kann für Parametrisierungen mit derselben Spur bewiesen werden. Schwieriger wird dann der Nachweis, dass Oberflächenintegrale erster (bzw. zweiter Art) bei äquivalenten Parametrisierungen übereinstimmen. Zusammensetzungen kann man wie gehabt einführen. Räumliche Normalbereiche bzgl. einer Koordinatenebene kann man auch mittels stetiger Abbildungen α, β einführen, deren Graphen Parametrisierungen erlauben - das einen Graphen entsprechende Oberflächenintegral zweiter Art ist dann (bei gegebener Orientierung) für ein Feld \vec{F} parallel zur Koordinatenebene durch α bzw. β eindeutig bestimmt.

Für beliebige stetig differenzierbare Felder bleibt aber das Problem, dass die bei der Definition des Volumens V als Normalbereich bzgl. jeder der 3 Koordinatenebenen betrachteten 3 Parametrisierungen der Oberfläche in der Tat äquivalent sein müssen - das muss man im Satz von Gauß als Voraussetzung fordern und im Einzelfall nachprüfen.

29 Beweis der Substitutionsregel

Wir haben zu zeigen, dass in der Tat eine ε -Substitution vorliegt. Dazu benötigen wir zunächst eine Charakterisierung der Messbarkeit. (Beachten Sie die Korrekturen in Kap.25 - die Raumdimension sollte stets d sein).

29.1 Charakterisierung messbarer Mengen

vgl. Walter, Analysis 2, Kap.7.5. Wir beziehen uns auf ein beliebiges ON-Koordinatensystem von \mathbb{R}^d .

Lemma 29.1 *Ist $C = [a, b]$ ein Intervall in \mathbb{R}^d und $C \cap B \neq \emptyset$ so gilt entweder $C \subseteq C^\circ$ oder $C \cap \partial B \neq \emptyset$. Gilt $C \subseteq B$, so gilt für jedes $\varepsilon > 0$*

$$C_\varepsilon = [a + \varepsilon e, b - \varepsilon e] \subseteq B^\circ, \quad e + \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$$

Beweis. Sei $B \cap C \neq \emptyset$, $C \not\subseteq B^\circ$ und $C \cap \partial B = \emptyset$. Da \overline{B} disjunkte Vereinigung von B° und ∂B ist, $C \cap B^\circ \neq \emptyset$ und $C \cap \mathbb{R}^d \setminus \overline{B} \neq \emptyset$. Wähle

$$p \in C \cap B^\circ, \quad q \in C \cap \mathbb{R}^d \setminus \overline{B}$$

$$L = \{x(t) \mid t \in [0, 1]\} \text{ mit } x(t) = (1-t)p + tq$$

Da C konvex ist, gilt $L \subseteq C$. Setze

$$s_0 = \sup\{s \in [0, 1] \mid \forall t \leq s. x(t) \in B^\circ\}$$

Dann $x(s_0) \in \partial B$, also $C \cap \partial B \neq \emptyset$, ein Widerspruch.

Sei nun $C \subseteq B$. Ist $p \in C_\varepsilon$, so $U_{\frac{\varepsilon}{42d}}(p) \subseteq C^\circ \subseteq B^\circ$. Also $C_\varepsilon \subseteq B^\circ$. \square

Eine *Würfelzerlegung* der Weite δ ist eine Gitterzerlegung mit Zellen

$$C = p + [0, \delta]^d$$

Ist Z eine Folge von Zerlegungen Z_n feiner Z_{n-1} von B mit Weite $\rightarrow 0$, so existieren nach dem Monotoniekriterium

$$U(B, Z) = \lim_{n \rightarrow \infty} U(B, Z_n) \leq O(B, Z) = \lim_{n \rightarrow \infty} O(B, Z_n)$$

B ist messbar genau dann, wenn $U(B, Z) = O(B, Z)$.

Lemma 29.2 *Ist B messbar, so ist auch jedes A mit $B^\circ \subseteq A \subseteq \overline{B}$ messbar.*

Beweis. Sei Z eine Folge von Gitterzerlegungen Z_n feiner Z_{n-1} mit Weite $\rightarrow 0$.

$$\bigcup_{C \in Z_n, C \cap B \neq \emptyset}$$

enthält B und ist abgeschlossen, enthält also auch \overline{B} . Es folgt

$$O(B, Z_n) = O(\overline{B}, Z_n), \quad O(B, Z) = O(\overline{B}, Z)$$

Ist B Nullmenge, so ist die Behauptung trivial. Andernfalls gibt es zu jedem $\varepsilon > 0$ ein n_0 so, dass

$$0 < U(B, Z_n), \quad O(B, Z_n) - U(B, Z_n) \leq \varepsilon \text{ für alle } n \geq n_0$$

Sei \tilde{Z}_n die gemeinsame Verfeinerung von Z_n und einem Würfelgitter der Weite $\frac{\varepsilon}{42^d}$. Nach Lemma 29.1 liegt eine Zelle $C \subseteq B$ von Z_n bis auch einen beliebig schmalen Randstreifen in B° , also

$$U(B, \tilde{Z}_n) - U(B, Z_n) \leq \varepsilon$$

und es folgt

$$O(B, \tilde{Z}_n) - U(B^\circ, \tilde{Z}_n) \leq 2\varepsilon$$

Somit folgt die Behauptung aus

$$U(B^\circ, \tilde{Z}) = O(\bar{B}, \tilde{Z}) \quad \square$$

Satz 29.3 *Eine beschränkte Menge B ist genau dann messbar, wenn ihr Rand ∂B eine Nullmenge ist.*

Beweis. Sei Z eine Folge von Gitterzerlegungen Z_n feiner Z_{n-1} mit Weite $\rightarrow 0$. Nach Lemma 29.2 gilt

$$U(B^\circ, Z_n) + O(\partial B, Z_n) = O(B, Z_n)$$

und somit

$$U(B^\circ, Z) + O(\partial B, Z) = O(B, Z)$$

Ist B messbar, so nach Lemma 29.2 auch B° und $U(B^\circ, Z) = O(\bar{B}, Z)$ also $O(\partial B, Z) = 0$ und damit ∂B Nullmenge. Ist umgekehrt ∂B Nullmenge, so $O(\partial B, Z) = 0$ also $U(B^\circ, Z) = U(B, Z) = O(B, Z)$ und somit B messbar. \square

29.2 Uniformer Rest

Sei $U \subseteq \mathbb{R}^d$ offen, $\sigma : U \rightarrow \mathbb{R}^d$ stetig differenzierbar und $K \subseteq U$ kompakt. Setze

$$R_\sigma(x, h) = \sigma(x + h) - \sigma(x) - D\sigma(x)(h) \quad x, x + h \in U$$

(1) Zu jeden $\varepsilon > 0$ gibt es ein $\delta > 0$ so, dass

$$\frac{|R_\sigma(x, h)|}{|h|} \leq \varepsilon \quad \text{für alle } |h| \leq \delta \text{ mit } [x, x + h] \subseteq K$$

Beweis. Nach Voraussetzung ist $x \mapsto J_\sigma(x)$ auf B stetig, also wegen Kompaktheit gleichmäßig stetig, d.h. es gibt zu $\varepsilon > 0$ ein $\delta > 0$ so, dass

$$\|J_\sigma(x) - J_\sigma(x + h)\|_\infty \leq \varepsilon \quad \text{falls } |h| \leq \delta$$

bzgl. der Matrixnorm $\|A\|_\infty = \max_{ij} |a_{ij}|$. Es folgt

$$\left\| \int_0^1 J_\sigma(x + \tau h) d\tau - J_\sigma(x) \right\|_\infty \leq \int_0^1 \|J_\sigma(x + \tau h) - J_\sigma(x)\|_\infty d\tau \leq \varepsilon$$

Nun nach dem Mittelwertsatz

$$\sigma(x + h) - \sigma(x) = \left(\int_0^1 J_\sigma(x + \tau h) d\tau \right)(h)$$

Ist nun C eine Konstante so, dass für die Operatornorm $|A|$ nach dem Satz über die Normäquivalenz gilt

$$|A| \leq C|A|_\infty$$

so folgt

$$\frac{|R_\sigma(x, h)|}{|h|} = \left| \left(\int_0^1 J_\sigma(x + \tau h) d\tau \right) - J_\sigma(x) \right| \left(\frac{h}{|h|} \right) \leq C\varepsilon \|A\|_\infty \leq C\varepsilon \quad \square$$

(2) Ist α linear, so $R_{\alpha \circ \sigma}(x, h) = \alpha(R_\sigma(x, h))$

Beweis.

$$\alpha(R_\sigma(x, h)) = \alpha(\sigma(x+h) - \sigma(x) - D\sigma(x)(h)) = (\alpha \circ \sigma)(x+h) - (\alpha \circ \sigma)(x) = D(\alpha \circ \sigma)(h) \quad \square$$

(3) Sei nun $D\sigma(x)$ für alle $x \in U$ invertierbar. Dann gibt es jeden $\varepsilon > 0$ ein $\delta > 0$ so, dass für alle $\alpha = (D\sigma(x))^{-1}$, $x \in B$ gilt

$$\frac{|R_{\alpha \circ \sigma}(x, h)|}{|h|} \leq \varepsilon \text{ für alle } |h| \leq \delta \text{ mit } [x, x+h] \subseteq K$$

Beweis. $x \mapsto D\sigma(x)^{-1}$ ist stetig (Satz 15.1), also existiert M mit $|D\sigma(x)^{-1}| \leq M$ für alle $x \in B$. Es folgt für $\alpha = D\sigma(x)^{-1}$ mit Matrix A sowie ε und $|h| \leq \delta$ aus (1)

$$\frac{|R_{\alpha \circ \sigma}(x, h)|}{|h|} = \frac{|A \cdot R_\sigma(x, h)|}{|h|} \leq |A| \cdot \frac{|R_\sigma(x, h)|}{|h|} \leq M \cdot \varepsilon \quad \square$$

29.3 Bilder von Würfeln

Es gelten die Voraussetzungen von Satz 25.14.

(4) Für messbares $C \subseteq B$ mit messbarem $\sigma(C)$ und für $\alpha = (D\sigma(u))^{-1}$ mit $u \in C$ gelte

$$|\mu(\alpha \circ \sigma(C)) - \mu(C)| \leq \varepsilon \mu(C)$$

Dann folgt

$$|\mu(\sigma(C)) - |\det D\sigma(u)| \mu(C)| \leq \varepsilon \mu(C) \max_{x \in B} |\det D\sigma(x)|$$

Beweis.

$$\begin{aligned} \varepsilon \mu(C) &\geq |\mu(\alpha \circ \sigma(C)) - \mu(C)| = \left| |\det \alpha| \mu(\sigma(C)) - |\det \alpha| \cdot |\det D\sigma(u)| \mu(C) \right| \\ &= |\det \alpha| \cdot |\mu(\sigma(C)) - |\det D\sigma(u)| \mu(C)| \end{aligned}$$

Die Behauptung folgt mit

$$\frac{1}{\det \alpha} = \det D\sigma(u) \quad \square$$

- (5) Für alle $\varepsilon > 0$ gibt es $\delta > 0$ so, dass für alle Würfel $C = p + [0, \delta]^d \subseteq B$ und $u \in C$ gilt: ist $\sigma(C)$ messbar, so

$$|\mu(\sigma(C)) - |\det D\sigma(u)|\mu(C)| \leq \varepsilon\mu(C)$$

Beweis. Sei zunächst $D\sigma(u) = \text{id}$. Nach Verschiebung O.B.d.A. $p = 0$ und $\sigma(0) = 0$. Sei $1 \geq \varepsilon > 0$ gegeben und δ nach (3) bestimmt. Es folgt

$$|\sigma(h) - h| \leq |h|\varepsilon \text{ falls } |h| \leq \delta$$

Also ist $\sigma(C)$ in einem Würfel W der Kantenlänge

$$\delta + 2\sqrt{d}\delta\varepsilon$$

enthalten. und somit

$$\begin{aligned} |\mu(\sigma(C)) - \mu(C)| &\leq (\delta + 2\sqrt{d}\delta\varepsilon)^d - \delta^d = \sum_{k=1}^d \binom{d}{k} \delta^k (2\sqrt{d}\delta\varepsilon)^{d-k} - \delta^d \\ &\leq \varepsilon\delta^d \sum_{k=1}^d \binom{d}{k} (2\sqrt{d})^{d-k} = \varepsilon\mu(C) \cdot \text{Konst.} \end{aligned}$$

Mit (3) und (4) folgt der allgemeine Fall. \square

29.4 Bilder von Nullmengen

- (6) In (5) gilt ohne die Voraussetzung der Messbarkeit:

$$\text{Weite}(\sigma(C)) \leq dM\delta(1 + 2\sqrt{d}\varepsilon)$$

wobei $M \geq |J_\sigma(x)|$ (euklidische Matrix Norm) für alle $x \in B$.

Beweis. Für affines α ist $\alpha(W)$ ein Spat aufgespannt von den $\delta\vec{a}_i$, wobei die \vec{a}_i die Spalten der Matrix A von α sind. Die Raum-Diagonale hat Länge $\leq \sum_i \delta\vec{a}_i \leq d\delta|A|$ und ist eine obere Schranke für die Weite des Spats. Die Behauptung folgt wie in (5). \square

Korollar 29.4 Sei $\sigma : U \rightarrow \mathbb{R}^d$ wie in Kor.28.1. und $B \subseteq U$ kompakt und messbar. Dann gilt: Ist $D \subseteq B$ Nullmenge, so ist $\sigma(D)$ eine Nullmenge.

Beweis. Wir wählen Z_n als Gitterzerlegung für B mit Zellen

$$C = p + [0, \delta_n]^d \subseteq U, \quad \delta_n \rightarrow 0$$

Nach Voraussetzung gilt

$$\sum_{C \in Z_n, C \cap D \neq \emptyset} \mu(C) \rightarrow 0$$

Nach (6) haben wir für $\sigma(D)$ eine Überdeckung mit affinen Bildern der Würfel W_C , $C \in Z_n, C \cap D \neq \emptyset$

$$\mu(W_C) \leq \text{Konst.} \cdot \mu(C)$$

also

$$\sum_{C \in Z_n, C \cap D \neq \emptyset} \mu(W_C) \rightarrow 0$$

Die zu den affinen Abbildungen gehörenden Determinanten sind beschränkt, also $\mu(\sigma(D)) = 0$. \square

Korollar 29.5 Sei $\sigma : U \rightarrow \mathbb{R}^d$ wie in Kor.28.1. Dann gilt: Ist $B \subseteq U$ kompakt und messbar, so ist $\sigma(B)$ kompakt und messbar

Beweis. $\sigma(B)$ ist wegen Stetigkeit auch kompakt. Nach Satz 29.3 ist $\partial B = B \setminus B^\circ$ Nullmenge - $B = \overline{B}$ wegen Kompaktheit. Nach Kor.28.1 ist $\sigma : U \rightarrow \sigma(U)$ Homöomorphismus der topologischen Räume U und $\sigma(U)$ (wegen der Offenheit von U bzw. $\sigma(U)$ in \mathbb{R}^d ist das Innere einer Teilmenge dasselbe, ob mans im Teilraum oder in \mathbb{R}^d nimmt) also gilt

$$\partial\sigma(B) = \sigma(B) \setminus \sigma(B)^\circ = \sigma(B \setminus B^\circ) = \sigma(\partial B)$$

Nach Kor.29.4 ist also auch $\partial\sigma(B)$ Nullmenge und daher nach Satz 29.3 $\sigma(B)$ messbar. \square

29.5 Beweis der Substitutionsregel

Wir haben zu zeigen, dass σ und $\tau = |\det D\sigma|$ aus der Regel für passende Z_n die Bedingungen einer ε -Substitution erfüllen. Zunächst bemerken wir, dass die Voraussetzungen von Kor.28.1 gegeben sind. Nach Kor.29.5 ist $\sigma(B)$ messbar und kompakt und somit Bedingung 2 erfüllt. Bedingung 1: Stetigkeit von σ folgt aus Differenzierbarkeit, die von τ folgt aus der vorausgesetzten Stetigkeit von $u \mapsto D\sigma$, \det und Betrag.

Wir wählen nun Z_n als Gitterzerlegung für B mit Zellen

$$C = p + [0, \delta_n]^d \subseteq U, \quad \delta_n \rightarrow 0$$

Dann sind die C messbar und kompakt, also nach Kor.29.5 auch die $\sigma(C)$. Ist $D \neq C$ in Z_n , so $C \cap D$ beschränkte Nullmenge, also nach Satz 29.4 auch $\sigma(C \cap D)$ Nullmenge und wegen der Injektivität gilt $\sigma(C \cap D) = \sigma(C) \cap \sigma(D)$. Also ist $\sigma(Z_n)$ eine Zerlegung für $\sigma(B)$ und Bedingung 3(a) erfüllt. Die Bedingungen 3 (b) und (c) sind nach (6) und (5) in 28.2 erfüllt. \square

Inhaltsverzeichnis

-3 Sprungstetige Funktionen und Regelintegrale	1
-3.1 Sprungstetige Funktionen und Regelintegrale	1
-3.2 Motivation	1
-3.3 Zerlegungen und Treppenfunktionen	1
-3.4 Satz von Bolzano Weierstraß	2
-3.5 Sprungstetige Funktionen	2
-3.6 Definition und einfache Eigenschaften des Integrals	4
-3.7 Riemannsches Integral	5
-3.8 Integration stetiger Funktionen	5
-2 Integrationstechniken	7
-2.1 Substitutionsregel	7
-2.2 Differentiale	7
-2.3 Integrationsregel	8
-2.4 Teilterm als neue Variable	9
-2.5 Integration durch Umkehrung	10
-2.6 Partielle Integration	11
-1 Integration rationaler Funktionen	12
-1.1 Polynome und euklidischer Algorithmus	12
-1.2 Faktorzerlegung	13
-1.3 Körper der rationalen Funktionen	13
-1.4 Partialbruchzerlegung	14
-1.5 Fundamentalsatz der Algebra	15
-1.6 Partialbruchzerlegung reeller rationaler Funktionen durch Ansatz	16
-1.7 Integrale rationaler Grundfunktionen	18
0 Uneigentliche Integrale	18
1 Endlichdimensionale normierte Räume	18
1.1 Affine Räume	19
1.2 Euklidische Vektorräume	19
1.3 Normierte Vektorräume	20
1.4 Metrische Räume	20
1.5 Konvergenz	21
1.6 Ortsvektoren	21
1.7 Koordinaten	22
1.8 Maximumnorm	22
2 Topologie endlichdimensionaler normierter Räume	23
2.1 Äquivalenz von Normen	23
2.2 Topologische Grundbegriffe in metrischen Räumen	24
2.3 Filtrierung in \mathbb{R}^n	25
2.4 Kompaktheit in \mathbb{R}^n	25
2.5 Berechnung von Skalarprodukten	25

2.6	Kompaktheit in metrischen Räumen	26
3	Lineare Abbildungen und Matrizen	27
3.1	Definition und Beispiele	27
3.2	Existenz und Eindeutigkeit	28
3.3	Isomorphie	29
3.4	Matrizen	30
3.5	Summe	30
3.6	Matrix mal Spalte	30
3.7	Matrixbeschreibung	31
3.8	Matrizenprodukt	31
3.9	Gesetze	32
3.10	Komposition	32
3.11	Inverse Matrix	33
3.12	Transponierte Matrix	33
3.13	Koordinatentransformation für Vektoren	33
3.14	Orthogonale Transformationsmatrizen	34
3.15	Transformation von Matrixbeschreibungen	35
4	Stetige Abbildungen	35
4.0	Abbildungen	35
4.1	Lipshitzstetige und stetige Abbildungen in metrischen Räumen	37
4.2	Relativtopologie	38
4.3	Funktionenlimes	39
4.4	Stetige Abbildungen in normierten Räumen	39
4.5	Stetigkeit und Kompaktheit	40
4.6	Stetigkeit und Zusammenhang	41
4.7	Gleichmäßige Konvergenz	41
5	Differentiation von Skalarenfeldern	41
5.1	Lineare Funktionen	41
5.2	Differenzierbarkeit	44
5.3	Richtungsableitungen	46
5.4	Partielle Ableitungen	48
5.5	Partielle Ableitungen als Funktionen	49
5.6	Rechenregeln	50
5.7	Differentiale und Fehlerabschätzung	51
5.8	Mittelwertsatz	52
6	Vektorfelder und lineare Approximation	53
6.1	Differenzierbarkeit	53
6.2	Jacobimatrix und partielle Ableitungen	54
6.3	Differenzierbare Kurven	55
6.4	Regeln	56
6.5	Polarkoordinaten und Co	57
6.6	Mittelwertsatz	58

7	Rechnen mit Taylorreihen in 1 Variablen	59
7.1	Satz von Taylor-Lagrange	59
7.2	Naherungsweise Gleichheit	59
7.3	Allgemeine Ableitungsregeln	60
7.4	Venachlassigung hoherer Terme	61
8	Taylorentwicklung in n Variablen	61
8.1	Hohere partielle Ableitungen	61
8.2	Vertauschen der Reihenfolge	62
8.3	Multiindices	63
8.4	Ableitungen	63
8.5	Potenzreihen	64
8.6	Taylorpolynom	65
8.7	Satz von Lagrange-Taylor	65
8.8	Beispiele	66
9	Quadratische Approximation	67
9.1	Bilinearformen	67
9.2	Quadratische Formen	68
9.3	Quadratische Abbildungen	68
9.4	Quadratische Approximation von Skalarenfeldern	69
9.5	Hesse-Form und zweifache Differenzierbarkeit	70
10	Quadratische Formen	71
10.1	Transformation	71
10.2	Definitheit	71
10.3	Symmetrische Diagonalisierung	71
10.4	Symmetrischer Gaußalgorithmus	72
10.5	Minoren	72
10.6	Tragheitssatz	73
10.7	Determinanten	73
10.8	Hauptminorenkriterium	74
11	Lokale Extrema	74
11.1	Definitionen	74
11.2	Extrema quadratischer Formen	75
11.3	Hinreichendes Kriterium	76
11.4	Beispiele	76
11.5	Zerlegung	78
12	Gaußalgorithmus und inverse Matrix	79
12.1	Umformung	79
12.2	Stufenform	80
12.3	Gauß'scher Algorithmus	81
12.4	Matrixschreibweise fur Gleichungssysteme	82
12.5	Matrixinversion und Gleichungslosen	82
12.6	Isomorphismen	84

12.7 Spaltenumformungen	84
13 Skalar-, Vektor- und Spatprodukt im Raum	85
13.1 Richtungskomponenten	85
13.2 Skalarprodukt	85
13.3 Ungleichungen	87
13.4 Orthonormalbasen	88
13.5 Normalenvektoren von Geraden bzw. Ebenen	89
13.6 Orientierung	90
13.7 Flächen	90
13.8 Vektorprodukt	92
13.9 Volumen	93
13.10 Übersicht	94
14 Determinanten	94
14.1 Regeln	94
14.2 Eindeutigkeit und Berechnung	95
14.3 Produktsatz.	96
14.4 Transponieren und Zeilenumformungen	96
14.5 Entwicklung	97
14.6 Cramersche Regel	98
14.7 Adjungierte Matrix	98
14.8 Existenz	99
15 Inversion, Operatornorm und Banachscher Fixpunktsatz	99
15.1 Stetigkeit der Inversion	99
15.2 Operatornorm	100
15.3 Vollständige metrische Räume	100
15.4 Banachscher Fixpunktsatz	101
16 Implizite Funktionen und Umkehrabbildungen	103
16.1 Implizite Funktionen	103
16.2 Implizite Vektorfunktionen	105
16.3 Umkehrfunktionen	105
16.4 Matrixnorm	106
16.5 Auflösung	106
16.6 Ableitung der Auflösung	108
17 Extrema unter Nebenbedingungen	109
17.1 Vorbemerkungen	109
17.2 Motivation	109
17.3 Lagrange-Multiplikatoren	110
17.4 Beispiele	111
17.5 Untermannigfaltigkeiten und Extrema unter Nebenbedingungen	114

18 Wege	114
18.1 Wege im euklidischen Raum	114
18.2 Stückweise stetig differenzierbare Wege	116
18.3 Weglänge	116
18.4 Parametrisierung durch Weglänge	117
18.5 Beispiele von Weglängen	119
18.6 Linienelemente	119
19 Wegintegrale	120
19.1 Definition	120
19.2 Summationstheorem für Wegintegrale	120
19.3 Wegintegral im Vektorfeld	121
19.4 Beispiele von Wegintegralen	122
19.5 Regeln für Wegintegrale	123
20 Potentiale	123
20.1 Definition und Anwendung	123
20.2 Sternförmige Mengen	124
20.3 Integrierbare Vektorfelder	125
20.4 Vektorfelder in der Ebene	125
20.5 Vektorfelder im Raum	127
20.6 Rotation und Divergenz	128
21 Parameterabhängige Integrale	128
21.1 Vertauschung von Integral und partieller Ableitung	128
21.2 Beispiele	129
21.3 Integrierbare Vektorfelder	130
23 Integration auf Intervallen im \mathbb{R}^n	132
23.1 Das Riemann-Integral über Intervallen im \mathbb{R}^n	132
23.1.1 Intervalle	132
23.1.2 Rechteckzerlegungen	133
23.1.3 Treppenfunktionen	134
23.1.4 Unter- und Obersummen	134
23.1.5 Riemann-Summe und Integral	134
23.1.6 Beispiel und Eigenschaften des Riemann-Integrals	135
23.1.7 Additivität	136
23.2 Integrale stetiger Funktionen	137
23.2.1 Zwischenwertsatz	137
23.2.2 Summationstheorem	137
23.2.3 Nachtrag: Riemannintegral in 1 Variablen	138
23.3 Mehrfachintegrale	142
23.3.1 Satz von Fubini	142
23.3.2 Beweis.	143
23.3.3 Beispiele	144

24 Messbarkeit von Mengen und Integrale	145
24.1 Jordan messbare Mengen	145
24.1.1 Definition	145
24.1.2 Vereinigungs- und Schnittmengen	147
24.2 Krumme Zerlegungen	147
24.2.1 Definition	147
24.2.2 Verfeinerung, Weite und Einschränkung	148
24.2.3 Messbarkeit	148
24.2.4 Bewegungsinvarianz	149
24.3 Riemann-Integrale	150
24.3.1 Treppenfunktionen	150
24.3.2 Riemann-Intergrale auf messbaren Mengen	150
24.3.3 Rechenregeln	151
24.3.4 Summationstheorem	153
24.4 Ordinatenmengen und Normalbereiche	154
24.4.1 Ordinatenmengen	154
24.4.2 Relative Normalbereiche	155
24.4.3 Normalbereiche	156
24.4.4 Beispiele - nur zur Illustration der Methode	157
25 Transformation von Maßen und Integralen	158
25.1 Unabhängigkeit des Jordan-Maßes	158
25.1.1 Vorbemerkung	158
25.1.2 Unabhängigkeit vom Koordinatensystem	159
25.1.3 Eulers Dreh	159
25.1.4 Beweis des Satzes	160
25.1.5 Polyeder	160
25.2 Affine Abbildungen	161
25.2.1 Affine Abbildungen und ihre Determinanten	161
25.2.2 Bewegungen	161
25.3 Maßtransformation bei affinen Abbildungen	162
25.3.1 Bewegungsinvarianz des Jordan-Maßes	162
25.3.2 Volumen eines Spats	162
25.3.3 Affine Maßtransformation	163
25.4 Integralsubstitution	163
25.5 Polarkoordinaten	164
25.5.1 Polarkoordinaten	164
25.5.2 Kreisringsegment	164
25.5.3 Polarsubstitution	165
25.6 Zylinderkoordinaten	166
25.6.1 Zylinderkoordinaten	166
25.6.2 Zylindersubstitution	166
25.6.3 Schwerpunkt	167
25.6.4 Rotationskörper	167
25.7 Kugelkoordinaten	168
25.7.1 Kugelkoordinaten	168

25.7.2	Kugelschalensektor	168
25.7.3	Substitution	168
25.7.4	Kugelkoordinaten, alternativ	169
25.8	ε -Substitution	169
25.9	Funktionaldeterminante	170
25.9.1	Substitutionsregel	170
25.9.2	Beispiele	171
26	Integralsätze in der Ebene	171
26.1	Wege	172
26.2	Wegintegrale	173
26.3	Normalbereiche bzgl. x_1 -Achse	174
26.4	Halbgrün	174
26.5	Normalbereich bzgl. der x_2 -Achse	175
26.6	Grüner Bereich	175
26.7	Orthogonal	176
26.8	Divergenz	177
26.9	Rotation	177
26.10	Grüner Gauss von Scharfenburg	177
26.11	Allgemeinere Bereiche	178
27	Integralsätze im Raum	179
27.1	Vektorprodukt	179
27.2	Einfache Parameterdarstellungen	179
27.3	Tangenten und Normalen	180
27.4	Zusammengesetzte Parameterdarstellungen	180
27.5	Pflähen	181
27.6	Pflähenintegrale	181
27.7	Rotationsflächen	182
27.8	Normalenintegral	183
27.9	Rotation	184
27.10	Schwarze Produktregel	186
27.11	Stokes	186
28	Divergenz und Normalenintegrale im Raum	187
28.1	Diffeomeorphismus	188
28.2	Treue Parametrisierungen	188
28.3	Umorientierte Pflähe	189
28.4	Graphen und Zylinder	189
28.5	Drittel-grün	190
28.6	Mini-Gauß	191
28.7	Grüne Bereiche	192
28.8	Gauß	193
28.9	Äquivalenz bei Graphen und Zylindern	193
28.10	Anwendbarer Gauß	195
28.11	Alternativen zur Flächendefinition	196

29 Beweis der Substitutionsregel	196
29.1 Charakterisierung messbarer Mengen	197
29.2 Uniformer Rest	198
29.3 Bilder von Würfeln	199
29.4 Bilder von Nullmengen	200
29.5 Beweis der Substitutionsregel	201