Analysis 2

Prof. Dr. Kohnen
Dr. O. Delzeith

WS 1996/97

Inhaltsverzeichnis

Vorwort zur 1. Auflage

Dies ist eine **nichtautorisierte Mitschrift** der Vorlesung „Analysis 2“, die Prof. Dr. Kohnen im WS 1996/97 gehalten hat, für den Inhalt dieser Mitschrift kann nicht gehaftet werden.

Prof. Dr. Kohnen hielt vom SS 96 bis zum SS 97 in Heidelberg die Vorlesungen „Analysis 1-3“. Der Aufbau dieser Vorlesungen war so gut, daß wir uns entschlossen haben, unsere Mitschrift der Vorlesung in geTEXte Form zu bringen. Natürlich kann und soll ein Skript nie den Besuch und die eigenhändige Mitschrift einer Vorlesung ersetzen, es kann höchstens als Ergänzung beim Lernen dienen.

In dieser Mitschrift fehlt das Kapitel 6: „Gewöhnliche Differentialgleichungen – eine Einführung“. Ich habe dies bewußt ausgespart und möchte den Leser hier auf weiterführende Literatur, bzw. das von Herrn Kohnen gehaltene Seminar im SS 97 verweisen.

Zum Schriftsatz möchte ich anmerken, daß ich versucht habe, mich so gut wie möglich an die Schreibweise der Vorlesung zu halten. Komponenten von Funktionen und Vektoren werden durch hochgesetzte Ziffern gekennzeichnet, z.B. $x = \left(x^{(1)}, x^{(2)}, \ldots, x^{(n)}\right) \in \mathbb{R}^n$, was unweigerlich zu Komplikationen z.B. mit der Schreibweise der n-ten Ableitung für Funktionen führt, ich vertraue hier auf den Scharfsinn des Lesers.

Diese Mitschrift wird mit Sicherheit noch Tipfehler aber u.U. auch inhaltliche Fehler enthalten. Ich möchte mich dafür entschuldigen und kann nur jedem raten, alles anhand der eigenen Mitschrift oder anderer Literatur genau zu prüfen. Für Hinweise zur Verbesserung bzw. Berichtigung bin ich immer dankbar.

Dieses File ist in aktuellster Form immer über die Skriptensammlung der Fachschaft MathPhys zu beziehen: http://MathPhys.fsk.uni-heidelberg.de/skripte

Um ein Durcheinander zu verhindern, ist auf dem Internet nur die Postscript-Version erhältlich, ich werde mich bemühen, Änderungen schnellstmöglich vorzunehmen.

Es bleibt mir noch „Danke“ zu sagen:

Danke, Herr Kohnen, für die Mühe, die sie in Ihre Vorlesungen gesteckt haben!
Danke, Oli, für gute Übungsblätter, faire Klausuren und eine gute Betreuung!
Danke, Andrea und Max, für gute Übungsgruppen!
Danke, Petra, für das TEXen von Ana 1 und Ana 3!
Danke, Gerhard, für das TEXen der Sätze aus Ana 2!
Danke, Dir, für viele schöne Stunden, die wir in diesem Sommer zusammen verbracht haben!

Heidelberg, 10.11.97 Hans Boie

Vorwort zur Version 1.1

In der Zwischenzeit wurden noch ein paar Schreibfehler verbessert. Ansonsten freue ich mich weiterhin über Anmerkungen und Verbesserungsvorschläge.

Heidelberg, 20.10.2000 Hans Boie

eMail: hboie@ix.urz.uni-heidelberg.de

Erinnerung

Wichtige Sätze aus der Differentialrechnung, Zusammenhang zwischen Diffenrential- und Integralrechnung:

Definition 1:
Sei $f : M \to \mathbb{R}, M \subset \mathbb{R}$ *mit*

$$U_r(x_0) = \big\{x \in \mathbb{R} \big| |x - x_0| < r\big\} \subset M \qquad \textit{für } r > 0$$

Dann heißt f *differenzierbar in* x_0*, falls*

$$\lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0} =: f'(x_0)$$

existiert

Satz 1:

$$f \textit{ differenzierbar in } x_0 \Longrightarrow f \textit{ stetig in } x_0$$

Satz 2 („Rechenregeln"):
Seien f *und* g *in* x_0 *differenzierbar, dann gilt:*

(i) $f + g$, $cf(c \in \mathbb{R})$ *und* fg *sind in* x_0 *differenzierbar und es gilt*

$$\begin{aligned}(f+g)'(x_0) &= f'(x_0) + g'(x_0)\\ (cf)'(x_0) &= cf'(x_0)\\ (fg)'(x_0) &= f'(x_0)g(x_0) + f(x_0)g'(x_0)\end{aligned}$$

(ii) Ist zusätzlich $g(x_0) \neq 0$, *so ist auch* $\frac{f}{g}$ *in* x_0 *differenzierbar und es gilt:*

$$\left(\frac{f}{g}\right)'(x_0) = \frac{f'(x_0)g(x_0) - f(x_0)g'(x_0)}{g^2(x_0)}$$

Satz 3 („Kettenregel"):
Seien I, I^* *zwei Intervalle, die nicht nur aus einem Punkt bestehen. Seien* $f : I \to \mathbb{R}$, $g : I^* \to \mathbb{R}$ *mit* $f(I) \subset I^*$*. Seien* f *in* $x_0 \in I$ *und* g *in* $y_0 = f(x_0)$ *differenzierbar. Dann ist auch* $g \circ f$ *in* x_0 *differenzierbar, und es gilt:*

$$(g \circ f)'(x_0) = g'\big(f(x_0)\big)f'(x_0)$$

Satz 4 („Mittelwertsatz“):
Sei $I = [a, b]$ mit $a < b$. Sei $f : I \to \mathbb{R}$ auf I stetig und auf (a, b) differenzierbar, dann existiert $\xi \in (a, b)$, so daß

$$f(b) - f(a) = f'(\xi)(b - a)$$

Korollar 1:
Gilt zusätzlich noch $f'(x) = 0$ $\forall x \in (a, b)$, so ist f konstant.

Satz 5:
Sei $I = [a, b]$ mit $a < b$. Sei $f : I \to \mathbb{R}$ stetig und injektiv, es sei f auf (a, b) differenzierbar mit $f'(x) > 0$ $\forall x \in (a, b)$. Sei $f([a, b]) = [\alpha, \beta]$ und sei $g : [\alpha, \beta] \to \mathbb{R}$ eine zu f inverse Funktion. Dann ist g auf (α, β) differenzierbar und es gilt:

$$g'(x) = \frac{1}{f'\big(g(y)\big)} \qquad \forall y \in (\alpha, \beta)$$

Satz 6 („Hauptsatz der Differential- und Integralrechnung“):
Sei $I = [a, b]$ mit $a < b$. Sei $f \in C^0(I)$. Sei $c \in I$. Dann ist die durch

$$F(x) = \int_c^x f(n)dn \qquad (x \in I)$$

definierte Funktion $F : I \to \mathbb{R}$ stetig differenzierbar auf I, und es gilt:

$$F'(x) = f(x) \qquad \forall x \in I$$

(d.h. F ist eine Stammfunktion von f)

Korollar 2:
Sei $I = [a, b]$ mit $a < b$, $f \in C^0(I)$. Dann gilt:

$$\int_a^b f(x)dx = \Phi(b) - \Phi(a) =: \Phi(x)\Big|_a^b$$

wobei Φ irgendeine Stammfunktion von f auf I ist.

Satz 7 („Partielle Integration“):
Seien $f, g \in C^1(I)$. Dann gilt:

$$\int_a^b f'(x)g(x)dx = f(x)g(x)\Big|_a^b - \int_a^b f(x)g'(x)dx$$

Satz 8 („Variablentransformation“):
Seien I bzw. I^* Intervalle mit den Endpunkten α, β bzw. a, b und $\alpha \neq \beta$, $a \neq b$. Sei $f \in C^0(I^*)$ und $\varphi \in C^1(I)$ mit $\varphi(I) \subset I^*$. Dann gilt:

$$\int_{\varphi(\alpha)}^{\varphi(\beta)} f(x)dx = \int_\alpha^\beta f\big(\varphi(n)\big)\varphi'(n)dn$$

(etwas unpräzieser: $x = \varphi(n)$, $dx = d\varphi(n) = \varphi'(n)dn$)

Kapitel 1

Spezielle Funktionen

Ziel: Einführung der sogenannten elementaren Funktionen, Untersuchung ihrer wichtigen Eigenschaften.

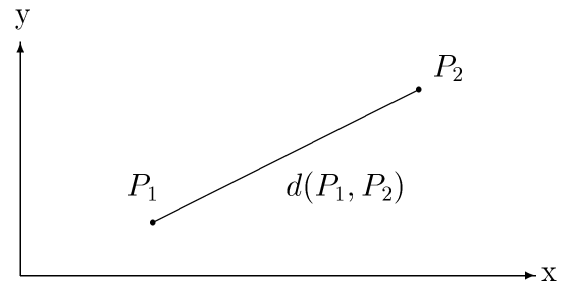
§1 Bogenlänge

Definition 1:
Sei $I = [a, b]$ mit $a < b$ $a, b \in \mathbb{R}$ ein kompaktes Intervall. Sei $f \in C^1(I)$. Dann setzt man

$$L(f) := \int_a^b \sqrt{1 + |f'(x)|^2} dx$$

Die Zahl $L(f)$ heißt **Bogenlänge** *des Graphen $G(f)$ von f.*

Motivation:



Ist $P_1 = (x_1, y_1)$, $P_2 = (x_2, y_2)$, so ist der **euklidische Abstand** von P_1 zu P_2 definiert als:

$$d(P_1, P_2) := \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}$$

(Satz des PYTHAGORAS)

allgemeiner: Seien $P_0, P_1, \ldots, P_n$ $n + 1$ Punkte in $\mathbb{R}^2$, $P_\nu = (x_\nu, y_\nu)$ $(\nu = 0, 1, \ldots, n)$; man verbinde $P_{\nu-1}$ mit P_ν durch ein Geradenstück $(\nu = 1, 2, \ldots, n)$, so erhält man einen sogenannten

„Polygonzug“ mit Eckpunkten $P_0, P_1, \ldots, P_n$.
Man definiert dessen Länge als:

$$\begin{aligned} &\sum_{\nu=1}^{n} d(P_{\nu-1}, P_\nu) \\ =&\sum_{\nu=1}^{n} \sqrt{(x_\nu - x_{\nu-1})^2 + (y_\nu - y_{\nu-1})^2} \end{aligned}$$

Sei $f \in C^1(I)$. Sei $\left(Z_n\right)_{n\in\mathbb{N}}$ eine Folge von Zerlegungen

$$Z_n = \left\{a = x_0^{(n)} < x_1^{(n)} < \cdots < x_{m_n}^{(n)} = b\right\}$$

von $[a, b]$ mit Feinheit $\delta(Z_n) = \max\limits_{1\le\nu\le m_n} \Delta x_\nu^{(n)} \xrightarrow{n\to\infty} 0$ mit $\Delta x_\nu^{(n)} = x_\nu^{(n)} - x_{\nu-1}^{(n)}$.

Man betrachte den „in $G(f)$ einbeschriebenen“ Polygonzug $\mathcal{P}_n$ mit Eckpunkten $P_\nu^{(n)} = (x_\nu^{(n)}, y_\nu^{(n)})$, wobei $y_\nu^{(n)} = f(x_\nu^{(n)})$ $(\nu = 0, 1, \ldots, m_n)$.

Die Länge von $\mathcal{P}_n$ ist dann gleich:

$$L(\mathcal{P}_n) = \sum_{\nu=1}^{m_n} \Delta S_\nu^{(n)}$$

wobei $\Delta S_\nu^{(n)} := \sqrt{\left(\Delta x_\nu^{(n)}\right)^2 + \left(\Delta y_\nu^{(n)}\right)^2}$

Nach dem Mittelwertsatz gibt es $\xi_\nu^{(n)} \in (x_{\nu-1}^{(n)}, x_\nu^{(n)})$ $\forall \nu$ mit:

$$\begin{aligned} f\left(x_\nu^{(n)}\right) - f\left(x_{\nu-1}^{(n)}\right) &= f'\left(\xi_\nu^{(n)}\right)\left(x_\nu^{(n)} - x_{\nu-1}^{(n)}\right) \\ L(\mathcal{P}_n) &= \sum_{\nu=1}^{m_n} \sqrt{1 + f'\left(\xi_\nu^{(n)}\right)^2} \cdot \Delta x_\nu^{(n)} \end{aligned}$$

In anderen Worten: $L(\mathcal{P})$ ist eine RIEMANN'sche Summe zum Integral $\sqrt{1 + f'(x)^2}$.
(Analysis 1, Kap. 6, §4)

Da $\delta(Z_n) \xrightarrow{n\to\infty} 0$ gilt, existiert es nach Satz

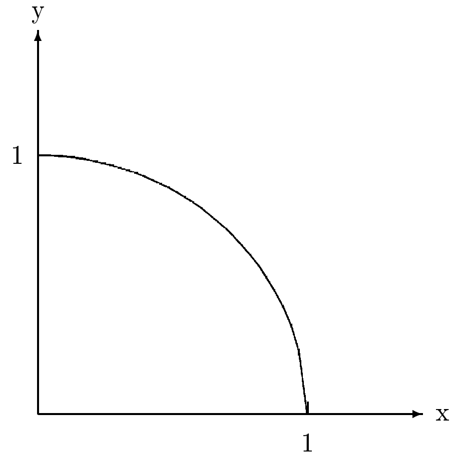
$$\lim_{n\to\infty} L(\mathcal{P}_n) = \int_a^b \sqrt{1 + f'(x)^2} dx$$

Dies rechtfertigt den Namen „Bogenlänge“ für

$$L(f) = \int_a^b \sqrt{1 + \left|f'(x)\right|^2} dx$$

Beispiel 1 (Viertelkreis):

$$V = \left\{(x, y) \in \mathbb{R}^2 \middle| x^2 + y^2 = 1, 0 < x \leq 1, 0 \leq y < 1\right\}$$



Der Viertelkreis kann als Graph der Funktion

$$y \mapsto g(y) = \sqrt{1 - y^2} \qquad (0 \leq y < 1)$$

aufgefaßt werden.
Für $P = (x, y) \in V$ setze man

$$\omega(P) = \int_b^a \sqrt{1 + |g'(u)|^2} du$$

d.h. $\omega(P)$ ist die Länge des Bogens zwischen $P_0(1, 0)$ und P.
Es gilt:

$$g'(y) = \frac{-2y}{2\sqrt{1 - y^2}} = \frac{-y}{\sqrt{1 - y^2}}$$
$$a + |g'(y)|^2 = 1 + \frac{y^2}{1 - y^2} = \frac{1}{1 - y^2}$$

Daher:

$$\omega(P) = \int_0^y \frac{du}{\sqrt{1 - u^2}}$$

Ziel: Definition der trigonometrischen Funktionen Sinus und Cosinus derart, daß für $P = (x, y) \in V$ und $\omega = \omega(P)$ gilt:

$$x = \cos \omega \qquad\qquad y = \sin \omega$$

§2 Die trigonometrischen Funktionen und ihre Umkehrfunktionen

Für $|y| < 1$ setze man

$$A(y) = \int_0^y \frac{du}{\sqrt{1-u^2}}$$

Nach dem Hauptsatz der Differtial- und Integralrechnung ist $A(y)$ eine stetig differenzierbar Funktion und es gilt:

$$A'(y) = \frac{1}{\sqrt{1-y^2}} > 0 \qquad (|y| < 1)$$

Daher ist $A(y)$ streng monoton wachsend.

Es gilt:

$$A(-y) = -A(y)$$

d.h. A ist „ungerade“, denn:

$$\begin{aligned} A(-y) &= \int_0^{-y} \frac{du}{\sqrt{1-u^2}} \underset{(u=-v)}{=} \int_0^{y} \frac{d(-v)}{\sqrt{1-(-v)^2}} \\ &= -\int_0^y \frac{dv}{\sqrt{1-v^2}} = -A(y) \end{aligned}$$

Wir setzen $A(y)$ zu einer stetigen Funktion auf $[-1,1]$ fort durch:

$$\begin{aligned} A(-1) &:= \lim_{y \to -1+} A(y) \\ A(1) &:= \lim_{y \to 1-} A(y) \end{aligned}$$

Behauptung: Die Limites existieren

Beweis:

Wegen $A(-y) = -A(y)$ genügt es zu zeigen, daß $\lim\limits_{y \to 1-} A(y)$ existiert.

Man schreibe:

$$A(y) = \int_0^{\frac{1}{2}} \frac{du}{\sqrt{1-u^2}} + \int_{\frac{1}{2}}^{y} \frac{du}{\sqrt{1-u^2}}$$

daher reicht es, das zweite Integral zu betrachten, es gilt:

$$\int\limits_{\frac{1}{2}}^{y} \frac{du}{\sqrt{1-u^2}} \underset{(v=\sqrt{1-u^2})}{=} \int\limits_{\frac{1}{2}\sqrt{3}}^{\sqrt{1-y^2}} \frac{1}{v} \cdot \frac{-v}{\sqrt{1-v^2}} dv$$
$$= - \int\limits_{\frac{1}{2}\sqrt{3}}^{\sqrt{1-y^2}} \frac{dv}{\sqrt{1-v^2}}$$

Daher gilt:

$$\lim_{y \to 1-} A(y) = \int\limits_0^{\frac{1}{2}} \frac{du}{\sqrt{1-u^2}} + \int\limits_0^{\frac{1}{2}\sqrt{3}} \frac{du}{\sqrt{1-u^2}}$$

$\Longrightarrow$ Limes existiert! □

Definition 1:
Man setzte:

$$\pi := 2A(1)$$

(d.h. 2π ist der Umfang des Einheitskreises.)

Unter Anwendung des Zwischenwertsatzes haben wir also gezeigt: Die Funktion A ist eine streng monoton wachsende, bijektive Abbildung von $[-1,1]$ auf $[-\frac{\pi}{2}, \frac{\pi}{2}]$, welche von der Klasse $C^0\big([-1,1]\big) \cap C^1\big((-1,1)\big)$ ist; sie ist ungerade und erfüllt:

$$A(-1) = -A(1) = -\frac{\pi}{2} \qquad A(0) = 0 \qquad A(1) = \frac{\pi}{2}$$

Man setze:

$$B(x) := \begin{cases} A(\sqrt{1-x^2}) & \text{für } 0 \le x \le 1 \\ \pi - A(\sqrt{1-x^2}) & \text{für } -1 \le x < 0 \end{cases}$$

Dann ist B auf $[-1,1]$ stetig (!) und für $0 < |x| < 1$ stetig differenzierbar mit der Ableitung:

$$B'(x) = \operatorname{sign}(x) A'(\sqrt{1-x^2}) \cdot \frac{-x}{\sqrt{1-x^2}} = \operatorname{sign}(x) \frac{1}{\sqrt{1-(\sqrt{1-x^2})^2}} \cdot \frac{-x}{\sqrt{1-x^2}}$$
$$= \operatorname{sign}(x) \frac{1}{\sqrt{|x|^2}} \cdot \frac{-x}{\sqrt{1-x^2}} = \frac{-1}{\sqrt{1-x^2}} < 0 \qquad \left(\text{denn } \frac{x}{|x|} = \operatorname{sign}(x)\right)$$

Daher ist B streng monoton fallend.
Es gilt:

$$B(-1) = \pi \qquad B(1) = 0$$

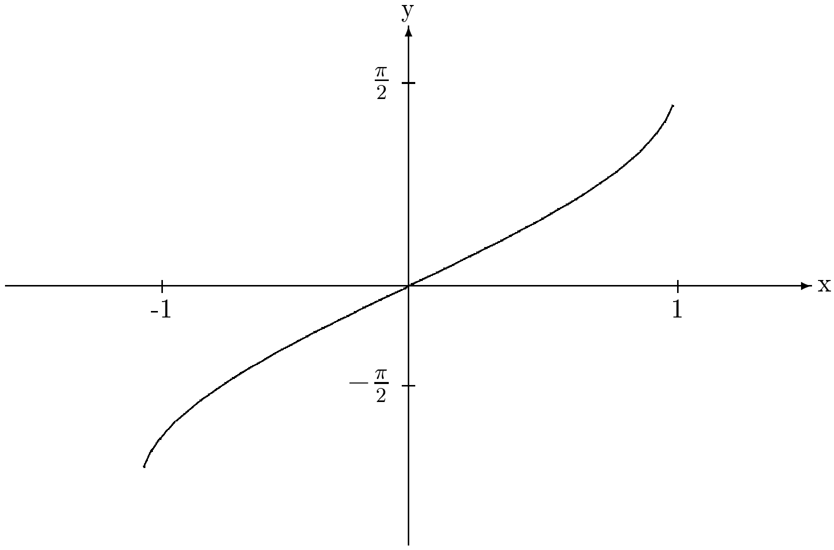


Abbildung 1.1: Die Funktion A (arcsin(x))

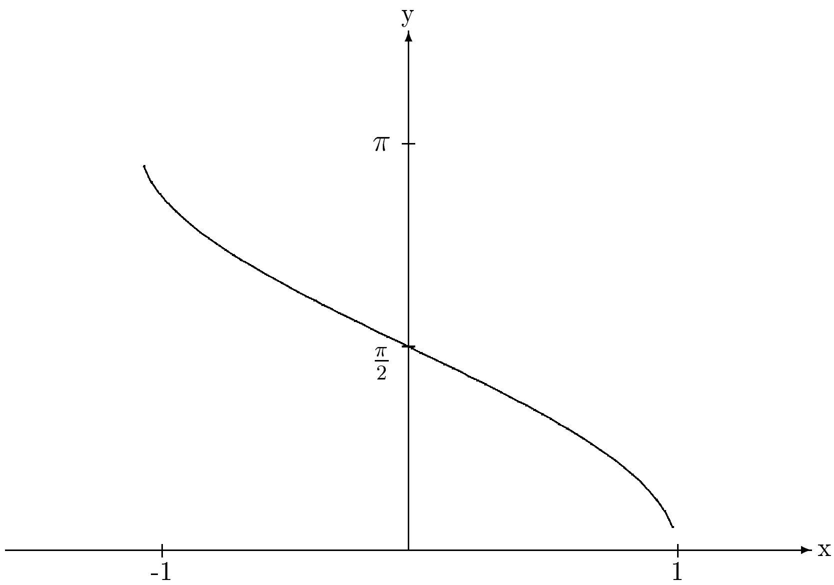


Abbildung 1.2: Die Funktion B (arccos(x))

Lemma 1:
Sei f stetig auf $U_r(x_0)$ und differenzierbar auf $U_r(x_0) \setminus \{x_0\}$. Es gelte $\lim\limits_{x \to x_0} f'(x) = c$ existiert, dann ist f auch in x_0 differenzierbar und $f'(x_0) = c$.

Beweis:
Übungsaufgabe!

(Entsprechendes gilt auch, wenn f nur in einer einseitigen Umgebung von x_0 definiert ist.)
Man wende das Lemma auf B und $x_0 = 0$ an. Es gilt:

$$\lim_{x \to 0} B'(x) = -1$$

also ist B auch in $x_0 = 0$ differenzierbar und $B'(0) = 1$.
Also: B ist eine streng monoton fallende, bijekitve Abbildung von $[-1, 1]$ auf $[0, \pi]$ der Klasse $C^0\big([-1, 1]\big) \cap C^1\big((-1, 1)\big)$. Es gilt:

$$B'(x) = -\frac{1}{\sqrt{1 - x^2}}$$

und:

$$B(-1) = \pi \qquad B(0) = \frac{\pi}{2} \qquad B(1) = 0$$

Seien:

$$S := A^{-1} \qquad C := B^{-1}$$

die Umkehrabbildungen von A bzw. B.
Nach dem Satz über inverse Funktionen gilt:

$$S \in C^0\big([-\frac{\pi}{2}, \frac{\pi}{2}]\big) \cap C^1\big((-\frac{\pi}{2}, \frac{\pi}{2})\big)$$
$$C \in C^0\big([0, \pi]\big) \cap C^1\big((0, \pi)\big)$$

Ferner ist S streng monoton wachsend mit

$$S(-\frac{\pi}{2}) = -1 \qquad S(0) = 0 \qquad S(\frac{\pi}{2}) = 1$$

C ist streng monoton fallend mit

$$C(0) = 1 \qquad C(\frac{\pi}{2}) = 0 \qquad C(\pi) = -1$$

Ferne ist S ungerade.
Beweis:
Es gilt:

$$\begin{aligned} S \circ A(y) = y \quad &\Longrightarrow \quad S(A(-y)) = -y \quad \Longrightarrow \quad -y = S(-A(y)) \qquad \text{(denn } A \text{ ist ungerade)} \\ &\Longrightarrow \quad -S(A(y)) = S(-A(y)) \qquad \text{d.h.} \quad -S(\omega) = S(-\omega) \quad \forall \omega \end{aligned}$$

□

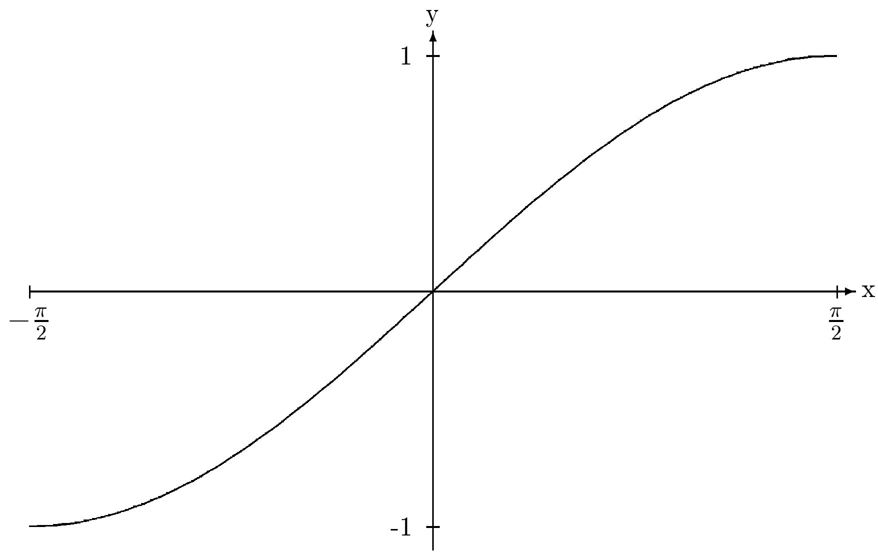


Abbildung 1.3: Die Funktion S ($\sin(x)$)

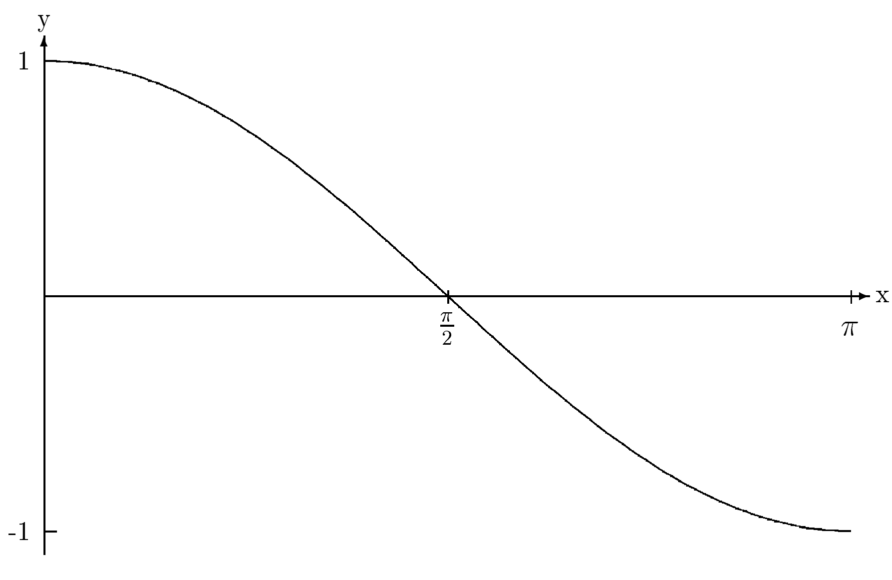


Abbildung 1.4: Die Funktion C ($\cos(x)$)

Ferner gilt:

$$S^2(\omega) + C^2(\omega) = 1 \qquad \forall \omega \in [0, \frac{\pi}{2}]$$

Beweis:
Für $0 \leq x \leq 1$ ist $B(x) = A(\sqrt{1-x^2})$. Daher gilt:

$$\begin{aligned} & S(B(x)) = S(A(\sqrt{1-x^2}) = \sqrt{1-x^2} \\ \Longrightarrow \qquad & 1 = x^2 + S(B(x))^2 \end{aligned}$$

Setze $\omega = B(x)$, also $x = C(\omega)$, dann folgt:

$$1 = C^2(\omega) + S^2(\omega)$$

Sei $0 < \omega < \frac{\pi}{2}$. Mit $x = C(\omega)$, $y = S(\omega)$ gilt nach dem Satz über die inverse Funktion:

$$\begin{aligned} S'(\omega) &= \frac{1}{A'(y)} = \frac{1}{\frac{1}{\sqrt{1-y^2}}} = \sqrt{1 - S^2(\omega)} \\ &= C(\omega) \qquad \text{denn: } S^2(\omega) + C^2(\omega) = 1 \text{ und } C(\omega) > 0 \\ C'(\omega) &= \frac{1}{B'(x)} = \frac{1}{-\frac{1}{\sqrt{1-x^2}}} = \sqrt{1 - C^2(\omega)} = -S(\omega) \end{aligned}$$

$\square$

Bemerkung:
Im Satz über die inverse Funktion kann man natürlich die Voraussetzung $f'(x) > 0\ \forall x$ durch $f'(x) < 0\ \forall x$ ersetzen. Zum Beweis ersetze man f durch $-f$ und beachte:

$$g\Big(-\big(-f(x)\big)\Big) = x$$

Die Formel

$$g'(y) = \frac{1}{f'(g(y))}$$

bleibt also korrekt.

Eine entsprechende Grenzwertbetrachtung (s. Lemma 1, S. 15) zeigt, daß S und C auch in 0 und $\frac{\pi}{2}$ differenzierbar sind und obige Formel auch dort korrekt bleibt.

Wie werden jetzt in 4 Schritten Funktionen $s, c \in C^1(\mathbb{R})$ konstruieren (ausgehend von S, C) derart, daß $s' = c$, $c' = s$ ist, s ungerade, c gerade gilt und s,c die Periode 2π haben:

$$s(x + 2\pi) = s(x) \qquad\qquad c(x + 2\pi) = c(x) \qquad\qquad \forall x \in \mathbb{R}$$

(i) Sei $0 \leq \omega \leq \frac{\pi}{2}$. Man setze:

$$c(\omega) := C(\omega) \qquad s(\omega) := S(\omega)$$

Nach dem oben Bewiesenen gilt dann:

$$s'(\omega) = c(\omega) \qquad c'(\omega) = -s(\omega) \qquad \forall \omega \in [0, \frac{\pi}{2}]$$

(ii) Sei $\frac{\pi}{2} \leq \omega \leq \pi$. Man setze dann:

$$c(\omega) := C(\omega) \qquad s(\omega) := \sqrt{1 - C^2(\omega)}$$

Sei $\frac{\pi}{2} \leq \omega < \pi$, dann gilt:

$$\begin{aligned} c'(\omega) = C'(\omega) &= -\sqrt{1 - C^2(\omega)} \qquad \text{(wie oben berechnet)} \\ &= -s(\omega) \end{aligned}$$

Ferner ist:

$$s'(\omega) = \frac{1}{2\sqrt{1 - C^2(\omega)}} \cdot (-2C(\omega)C'(\omega)) = \frac{1}{s(\omega)} \cdot (-c(\omega)(-s(\omega))) = c(\omega)$$

Durch eine Grenzwertbetrachtung (s. Lemma 1, S. 15) folgt, daß s und c auch in π differenzierbar sind und die obigen Formeln auch dort gelten.

(iii) Sei $-\pi \leq \omega \leq 0$. Dann setze man:

$$s(\omega) := -s(-\omega) \qquad c(\omega) := c(-\omega)$$

Es gilt:

$$\begin{aligned} s'(\omega) &= -(-1)s'(-\omega) = s'(-\omega) = c(-\omega) = c(\omega) \\ c'(\omega) &= -c'(-\omega) = s(-\omega) = -s(\omega) \end{aligned}$$

Bisher wurde gezeigt: Existenz von s, c auf $[-\pi, \pi)$ mit $s' = c$, $c' = -s$, s ungerade, c gerade und

$$s(-\pi) = s(\pi) = 0 \quad s'(-\pi) = s'(\pi) = -1 \quad c(-\pi) = c(\pi) = -1 \quad c'(-\pi) = c'(\pi) = 0$$

Ferner: $s(0) = 0$ und $c(0) = 1$

(iv) Für $\omega = \theta + 2\pi k$ mit $\theta \in [-\pi, \pi]$ und $k \in \mathbb{Z}$ setzt man:

$$s(\omega) := s(\theta) \qquad c(\omega) := c(\theta)$$

Man verifiziert dann sofort, daß alle gewünschten Eigenschaften für s und c gelten.

Man beachte:

$$\begin{aligned} s(x) = 0 \quad &\Longleftrightarrow \quad x = \pi k \qquad \text{mit } k \in \mathbb{Z} \\ c(x) = 0 \quad &\Longleftrightarrow \quad x = \frac{\pi}{2} + \pi k \qquad \text{mit } k \in \mathbb{Z} \end{aligned}$$

Satz 1:
Seien $s, c \in C^1(\mathbb{R})$ mit $s' = c$, $c' = -s$ und $s(0) = 0$, $c(0) = 1$. Dann sind s und c eindeutig bestimmt, und es gilt:

$$s^2(x) + c^2(x) = 1 \qquad \forall x \in \mathbb{R}$$

Ferner ist s ungerade, c gerade, und es gelten die **Additionstheoreme***:*

$$\begin{aligned} s(x+y) &= s(x)c(y) + s(y)c(x) \\ c(x+y) &= c(x)c(y) - s(x)s(y) \end{aligned}$$

Beweis:
Übungsaufgabe!

Man setze für $x \in \mathbb{R}$:

$$\sin x := s(x) \qquad\qquad \cos x := c(x)$$

Dann gilt:

$$\frac{d}{dx}\sin x = \cos x \qquad\qquad \frac{d}{dx}\cos x = -\sin x$$
$$\sin^2 x + \cos^2 x = 1 \quad \ldots \text{ etc.}$$

Man setze:

Tangens: $$\tan x := \frac{\sin x}{\cos x} \qquad \left(x \in \mathbb{R}, x \neq \frac{\pi}{2} + \pi k, k \in \mathbb{Z}\right)$$

Kotangens: $$\cot x := \frac{\cos x}{\sin x} \qquad (x \in \mathbb{R}, x \neq \pi k, k \in \mathbb{Z})$$

Für die Ableitungen gilt:

$$\tan' x = \frac{\sin' x \cos x - \cos' x \sin x}{\cos^2 x} = \frac{\cos^2 x + \sin^2 x}{\cos^2 x} = \frac{1}{\cos^2 x} = 1 + \tan^2 x$$
$$\cot' x = -\frac{1}{\sin^2 x}$$

Man nennt die Funktionen:

$$\arcsin x := A(x) \qquad \arccos x := B(x) \qquad x \in [-1, 1]$$

den **„Hauptzweig“** der „Funktionen“ **Arcus Sinus** bzw. **Arcus Cosinus**.
Dann gilt also:

$$\arcsin = \left(\sin\big|_{[-\frac{\pi}{2}, \frac{\pi}{2}]}\right)^{-1}$$
$$\arccos = \left(\cos\big|_{[0,\pi]}\right)^{-1}$$

Genauer gesagt, besitzt die Einschränkung von sin auf jedes Intervall $[-\frac{\pi}{2}+\pi k, \frac{\pi}{2}+\pi k]$ eine Umkehrfunktion, welche man einen **„Zweig“ von Arcus Sinus** nennt. Ähnliches gilt für den Cosinus. Für die Hauptzweige gilt ($x \in (-1,1)$)

$$\frac{d}{dx}\arcsin x = \frac{1}{\sqrt{1-x^2}}$$
$$\frac{d}{dx}\arccos x = -\frac{1}{\sqrt{1-x^2}}$$

Genauso kann man „Zweige“ und Umkehrfunktionen für Tangens und Kotangens definieren. Man definiert die Hauptzweige durch:

$$\operatorname{arctan} := \left(\tan\big|_{[-\frac{\pi}{2},\frac{\pi}{2}]}\right)^{-1}$$
$$\operatorname{arccot} := \left(\cot\big|_{[0,\pi]}\right)^{-1}$$

Setzt man $y = \arctan x$, $x = \tan x$, so gilt:

$$\tan' y = \frac{1}{\cos^2 y} = \frac{\sin^2 y \cos^2 y}{\cos^2 y} = \frac{\sin^2 y}{\cos^2 y} + 1 = \tan^2 y + 1 = 1 + x^2$$

Für die Ableitungen folgt dann also:

$$\frac{d}{dx}\arctan x = \frac{1}{\tan' y} = \frac{1}{1+x^2}$$
$$\frac{d}{dx}\operatorname{arccot} x = -\frac{1}{1+x^2}$$

§3 Expontential- und Logarithmusfunktion

Definition 1 (Logarithmusfunktion):
Für $x > 0$ setze man:

$$\log x := \int_1^x \frac{du}{u}$$

Die Funktion $\log : \mathbb{R}_+ \to \mathbb{R}$ *heißt* **Logarithmusfunktion** *(zuweilen auch „Logarithmus naturalis“, in Zeichen:* $\ln x$*). Offenbar gilt:*

$$\log 1 = 0$$

Satz 1:
Die Funktion $\log x$ *(*$x > 0$*) ist unendlich oft differenzierbar, und es gilt:*

$$\frac{d}{dx}\log x = \frac{1}{x}$$

Ferner ist das unbestimmte Integral von $\frac{1}{x}$ *(*$x \neq 0$*), d.h. die Menge aller Stammfunktionen gleich:*

$$\int \frac{dx}{x} = \log|x| + const.$$

Beweis:
Nach dem Hauptsatz der Differtial- und Integralrechnung ist:

$$\log' x = \frac{1}{x}$$

Die zweite Aussage ist für $x > 0$ klar, denn dann ist $|x| = x$. Sei $x < 0$, dann ist $|x| = -x$, also:

$$\log|x| = \log(-x)$$

daher:

$$\frac{d}{dx}\log|x| = \frac{1}{-x}\cdot(-1) = \frac{1}{x}$$

□

Satz 2 (Additionstheorem):
Für $a, b \in \mathbb{R}_+$ *gilt:*

$$\log(ab) = \log a + \log b$$

Beweis:
Es gilt:

$$\log(ab) = \int\limits_1^{ab} \frac{du}{u} = \int\limits_1^{a} \frac{du}{u} + \int\limits_a^{ab} \frac{du}{u}$$

Im zweiten Integral substituiert man:

$$u = va \qquad\qquad 1 \leq v \leq b$$

also: $v = \frac{u}{a}$, $\frac{du}{dv} = a$, daher:

$$\int\limits_a^{ab} \frac{du}{u} = \int\limits_1^{b} \frac{a\ dv}{va} = \int\limits_1^{b} \frac{dv}{v} = \log b$$

$$\Longrightarrow \qquad \log(ab) = \int\limits_1^{a} \frac{du}{u} + \int\limits_a^{ab} \frac{du}{u} = \int\limits_1^{a} \frac{du}{u} + \int\limits_1^{b} \frac{du}{u} = \log a + \log b$$

□

Korollar 1:

(i) $\log(a_1 \cdot a_2 \cdot \ldots \cdot a_n) = \log a_1 + \log a_2 + \ldots \log a_n \quad (a_i > 0 \ \forall i = 1, 2, \ldots, n)$

(ii) $\log x^\rho = \rho \log x \quad (x > 0, \rho \in \mathbb{Q})$

Beweis:

(i) klar durch Induktion!

(ii) nach (i) gilt $\log x^n = n \log x \quad (n \in \mathbb{N})$.
Ferner gilt $\log x + \log \frac{1}{x} = \log(x\frac{1}{x}) = \log 1 = 0$, also: $\log \frac{1}{x} = -\log x$.
Daher gilt für $n \in \mathbb{N}$:

$$\log x^{-n} = \log \underbrace{(\frac{1}{x} \cdot \ldots \cdot \frac{1}{x})}_{n\text{-mal}} = n \log \frac{1}{x} = -n \log x$$

Für $n \in \mathbb{Z}$, $n \neq 0$ gilt:

$$\log x = \log(x^{\frac{1}{n}})^n = n \log x^{\frac{1}{n}} \qquad \Longrightarrow \qquad \log x^{\frac{1}{n}} = \frac{1}{n} \log x$$

Hieraus folgt die Behauptung!

$\square$

Satz 3:
Die Logarithmusfunktion ist streng monoton wachsend (insbesondere also injektiv)
Es gilt: $\log \mathbb{R}_+ = \mathbb{R}$ *und:*

$$\lim_{x \to \infty} \log x = \infty \qquad \lim_{x \to 0} \log x = -\infty$$

Beweis:
Wegen $\log' x = \frac{1}{x} > 0$ gilt: Die Funktion log ist streng monoton wachsend. Daher ist

$$\begin{aligned} & \log 2 > \log 1 = 0 \\ \Longrightarrow \quad & \log 2^n = n \log 2 \xrightarrow{n \to \infty} \infty \end{aligned}$$

Wegen der Monotonie folgt also:

$$\lim_{x \to \infty} \log x = \infty$$

Für $x \to \infty$ gilt $\frac{1}{x} \to 0$ und umgekehrt, ferner ist $\log \frac{1}{x} = -\log x \Longrightarrow \lim_{x \to 0} \log x = -\infty$.
Wende den Zwischenwertsatz an, so folgt schließlich:

$$\log \mathbb{R}_+ = \mathbb{R}$$

$\square$

Definition 2 (Exponentialfunktion):
Die Umkehrfunktion

$$\log^{-1} : \mathbb{R} \to \mathbb{R}_+ \subset \mathbb{R}$$

zur Logarithmusfunktion heißt **Exponentialfunktion** *und wird mit* exp *bezeichnet.*

Satz 4:

(i) exp *ist unendlich oft differenzierbar streng monoton wachsend, und es gilt:*

$$\frac{d}{dx}\exp = \exp \qquad\qquad \exp x \begin{cases} > 1 & \textit{für } x > 0 \\ = 1 & \textit{für } x = 0 \\ < 1 & \textit{für } x < 0 \end{cases}$$

(ii) $\exp(x+y) = \exp x \cdot \exp y$

(iii) $\exp(\rho x) = (\exp x)^\rho \qquad$ *für* $\rho \in \mathbb{Q}$

(iv) $\exp 1 = \mathrm{e} \qquad$ *(*„Euler**'sche Zahl**“*)*

Beweis:

(i)

$$\frac{d}{dx}\exp x = \frac{1}{\log'(\exp x)} = \frac{1}{\frac{1}{\exp x}} = \exp x$$

Die übrigen Aussagen unter (i) folgen aus den entsprechenden Aussagen über log. (z.B. $\exp x > 1 \iff \log(\exp x) > 0 \iff x > 0 \ldots$ etc.)

(ii) Sei $a = \exp x$, $b = \exp y$, dann gilt:

$$\begin{aligned} & x + y = \log a + \log b = \log(ab) \\ \Longrightarrow \quad & \exp(x+y) = \exp\big(\log(ab)\big) = ab = \exp x \exp y \end{aligned}$$

(iii) Sei $a = \exp x$. Es gilt

$$\begin{aligned} & \log a^\rho = \rho \log a = \rho x \qquad \text{(Kor. 1 zu Satz 2, S. 22)} \\ \Longrightarrow \quad & \exp(\rho x) = \exp\big(\log a^\rho\big) = a^\rho = (\exp x)^\rho \end{aligned}$$

(iv) Für $x > 0$ gilt:

$$\frac{1}{x} = \log' x = \lim_{h \to 0} \frac{\log(x+h) - \log x}{h}$$

Speziell mit $x = 1$ und $h_n = \frac{1}{n}$ $\quad (n \in \mathbb{N})$ gilt also:

$$\begin{aligned}
1 &= \lim_{n\to\infty} \frac{\log(1+\frac{1}{n}) - \log 1}{\frac{1}{n}} \\
&= \lim_{n\to\infty} n \log(1+\frac{1}{n}) = \lim_{n\to\infty} \log(1+\frac{1}{n})^n
\end{aligned}$$

Nach Definition ist:

$$\mathrm{e} = \lim_{n\to\infty} (1+\frac{1}{n})^n \qquad \text{(s. Analysis 1)}$$

Da log stetig, folgt also:

$$1 = \log \mathrm{e} \qquad \Longrightarrow \qquad \exp 1 = \mathrm{e}$$

$\square$

Bemerkung:
Nach (iii) und (iv) gilt für $\rho \in \mathbb{Q}$:

$$\exp \rho = \exp(\rho \cdot 1) = (\exp 1)^\rho = \mathrm{e}^\rho$$

Daher ist folgende Definition sinnvoll:

Definition 3:

$$\mathrm{e}^x := \exp x \quad (x \in \mathbb{R})$$

Insbesondere gilt also:

$$\mathrm{e}^{x+y} = \mathrm{e}^x\, \mathrm{e}^y \qquad\qquad (\mathrm{e}^x)^\rho = \mathrm{e}^{\rho x} \qquad \forall \rho \in \mathbb{Q}$$

Definition 4:
Für $c > 0$, $x \in \mathbb{R}$ setzt man:

$$c^x := \mathrm{e}^{x \log c}$$

Bemerkung:

(i) Diese Definition von c^x stimmt für $x \in \mathbb{Q}$ mit der alten Definition c^x überein:

$$c^\rho = \exp(\log c^\rho) = \exp(\rho \log c) = \mathrm{e}^{\rho \log c}$$

(ii) Sei $x \in \mathbb{R}$ und sei $\left(\rho_n\right)_{n\in\mathbb{N}}$ eine Folge rationaler Zahlen mit

$$\lim_{n\to\infty} \rho_n = x$$

(s. Analysis 1, Übungsblatt 9, Aufg. 4) Dann gilt:

$$\mathrm{e}^{x \log c} = \mathrm{e}^{\lim\limits_{n\to\infty} \rho_n \log c} \underset{\text{(Stetigkeit)}}{=} \lim_{n\to\infty} \mathrm{e}^{\rho_n \log c}$$

nach $\log x^\rho = \rho \log x$ für $\rho \in \mathbb{Q}$ gilt also:

$$\begin{aligned} \mathrm{e}^{x \log c} &= \lim_{n\to\infty} \mathrm{e}^{\log c^{\rho_n}} \\ \Longrightarrow \quad c^x &= \mathrm{e}^{x \log c} = \lim_{n\to\infty} c^{\rho_n} \end{aligned}$$

Man kann daher die „allgemeine Potenz“ als Grenzwert einer Folge rationaler Potenzen deuten.

Satz 5:

(i) $c^{x+y} = c^x c^y \qquad \forall x, y \in \mathbb{R}$

(ii) $(c^x)^y = c^{xy} \qquad \forall x, y \in \mathbb{R}$

(iii) $\log c^x = x \log c \qquad \forall x \in \mathbb{R}$

(iv) Die Funktion $x \mapsto c^x$ ist unendlich oft differenzierbar und es gilt:

$$\frac{d}{dx} c^x = c^x \log c$$

Beweis:

(i)

$$\mathrm{e}^{x+y} = \mathrm{e}^{(x+y)\log c} = \mathrm{e}^{x \log c + y \log c} = \mathrm{e}^{x \log c} + \mathrm{e}^{y \log c} = c^x c^y$$

(ii) Sei $\left(\rho_n\right)_{n\in\mathbb{N}}$ eine Folge rationaler Zahlen mit

$$\lim_{n\to\infty} \rho_n = y$$

Dann gilt:

$$(c^x)^y = \lim_{n\to\infty} (c^x)^{\rho_n} = \lim_{n\to\infty} (\mathrm{e}^{x \log c})^{\rho_n} = \lim_{n\to\infty} \mathrm{e}^{\rho_n x \log c} \underset{\text{(Stetigkeit)}}{=} \mathrm{e}^{xy \log c} = c^{xy}$$

(iii) $\log c^x = \log \mathrm{e}^{x \log c} = x \log c$, denn exp und log sind zueinander inverse Funktionen.

(iv) $\mathrm{e}^x = \mathrm{e}^{x \log c} \Longrightarrow \frac{d}{dx} c^x = \mathrm{e}^{x \log c} \log c = c^x \log c$

□

Bemerkung:
Sei $c > 0$. Für $c \neq 1$ hat dann $x \mapsto c^x$ eine Umkehrfunktion (denn $\frac{d}{dx} c^x = c^x \log c$ ist positiv für $c > 1$ und negativ für $c < 1$). Diese nennt man den **„Logarithmus zur Basis c“**, in Zeichen $\log_c x$.

Satz 6:
Sei $a \in \mathbb{R}$, dann ist die Abbildung $x \mapsto x^a$ $(x > 0)$ differenzierbar, und es gilt:

$$\frac{d}{dx} x^a = a x^{a-1}$$

Beweis:
Nach Definition ist: $x^a = \mathrm{e}^{a \log x}$, daher gilt:

$$\frac{d}{dx} x^a \underset{\text{(Kettenregel)}}{=} \mathrm{e}^{a \log x}\, a \frac{1}{x} = a x^a \frac{1}{x} = a x^{a-1}$$

□

Satz 7:
Sei $f \in C^1(I)$ mit $I = [a,b]$ mit $a < b$. Sei $f(x) > 0\ \forall x \in I$, dann ist auch $x \mapsto \log f(x)$ von der Klasse $C^1(I)$, und

$$\frac{d}{dx} \log f(x) = \frac{f'(x)}{f(x)}$$

Beweis:
Übungsaufgabe

Korollar 2:
Sei $f \in C^1(I)$ mit $I = [a,b]$ mit $a < b$. Sei $f(x) \neq 0\ \forall x \in I$, dann gilt:

$$\int \frac{f'(x)}{f(x)} dx = \log |f(x)| + const.$$

Beweis:
Übungsaufgabe

Satz 8:

(i) $\mathrm{e}^x = \lim\limits_{h \to 0} (1 + hx)^{\frac{1}{h}} \qquad (x \in \mathbb{R})$

(ii) $\log a = \lim\limits_{h \to 0} \frac{a^h - 1}{h} \qquad (a \in \mathbb{R}_+)$

Beweis:

(i) Es gilt (ohne Einschränkung $x \neq 0$):

$$\frac{1}{u} \log(1+u) = \frac{\log(1+u) - \log 1}{u} \xrightarrow{u \to 0} \log' 1 = 1$$

Setze: $u = hx$, dann folgt:

$$(1+hx)^{\frac{1}{h}} = \mathrm{e}^{\log(1+hx)^{\frac{1}{h}}} = \mathrm{e}^{\frac{1}{h} \log(1+hx)} = \mathrm{e}^{x \frac{1}{u} \log(1+u)} \xrightarrow{h \to 0} \mathrm{e}^{x \cdot 1} = \mathrm{e}^x$$

(ii)

$$\frac{a^h - 1}{h} = \frac{a^h - a^0}{h} \xrightarrow{h \to 0} \frac{d}{dx} a^x \bigg|_{x=0}$$
$$\underset{\text{(Satz 5, (iv))}}{=} \left(a^x \log a\right) \bigg|_{x=0} = a^0 \log a = \log a$$

□

§4 Die Hyperbel-Funktionen

Definition 1:
Für $x \in \mathbb{R}$ setze man:

Sinus hyperbolicus: $$\sinh x := \frac{e^x - e^{-x}}{2}$$

Cosinus hyperbolicus: $$\cosh x := \frac{e^x + e^{-x}}{2}$$

Eine einfache Rechnung zeigt, daß gilt:

$$\cosh^2 x - \sinh^2 x = 1$$

d.h. diese Funktionen parametrisieren die Hyperbel.

Weiter gilt:

$$\sinh' x = \cosh x \qquad \cosh' x = \sinh x$$

Beweis:
Übungsaufgabe!

Kapitel 2

Weiterer Ausbau der Infinitesimalrechnung einer reellen Funktion einer reellen Variablen

§1 Die Regel von de l' Hospital

Ziel: Untersuchung von „unbestimmten Ausdrücken der Form $\frac{\infty}{\infty}$ oder $\frac{0}{0}$“ (z.B. Grenzwertverhalten von $\frac{x}{\log x}$ für $x \to \infty$) mit Hilfe der Ableitung.

Satz 1 (Verallgemeinerung des Mittelwertsatzes):
Seien $a, b \in \mathbb{R}$ mit $a < b$. Seien f, g zwei Funktionen, die auf $[a, b]$ stetig und auf (a, b) differenzierbar sind. Es gelte $g'(x) \neq 0 \; \forall x \in (a, b)$. Dann gilt $g(b) - g(a) \neq 0$ und es existiert $\xi \in (a, b)$, so daß:

$$\frac{f(b) - f(a)}{g(b) - g(a)} = \frac{f'(\xi)}{g'(\xi)}$$

Beweis:
Nach dem Mittelwertsatz angewandt auf g gilt:

$$\exists\ \xi_1 \in (a, b) \text{ mit } g(b) - g(a) = g'(\xi_1)(b - a)$$

Wegen $g'(x) \neq 0 \; \forall x \in (a, b)$ gilt daher: $g(b) - g(a) \neq 0$

Man betrachte die Hilfsfunktion:

$$h(x) := f(x) - f(a) - \frac{f(b) - f(a)}{g(b) - g(a)}\big(g(x) - g(a)\big) \qquad (x \in [a, b])$$

Dann ist h auf $[a, b]$ stetig und auf (a, b) differenzierbar. Es gilt ferner $h(a) = 0$, $h(b) = 0$

Nach dem Satz von Rolle (Analysis 1) gibt es $\xi \in (a, b)$ mit $h'(\xi) = 0$. Das heißt:

$$0 = f'(\xi) - \frac{f(b) - f(a)}{g(b) - g(a)} g'(\xi)$$

Daher folgt die Behauptung. □

Satz 2 (de l' Hospital):
Seien $a, b \in \overline{\mathbb{R}} (= \mathbb{R} \cup \{\infty, -\infty\})$ *mit* $a < b$. *Seien* f, g *auf* (a, b) *differenzierbare Funktionen mit* $g'(x) \neq 0 \ \forall x \in (a, b)$.

Es gelte entweder:

$$f(x) \xrightarrow{x \to a} 0 \quad \wedge \quad g(x) \xrightarrow{x \to a} 0 \qquad (*)$$

oder:

$$f(x) \xrightarrow{x \to a} \infty \quad \wedge \quad g(x) \xrightarrow{x \to a} \infty \qquad (**)$$

Dann gilt folgende Aussage:

Wenn der Grenzwert

$$\lim_{x \to a+} \frac{f'(x)}{g'(x)}$$

existiert, so existiert auch der Grenzwert

$$\lim_{x \to a+} \frac{f(x)}{g(x)}$$

und es gilt:

$$\lim_{x \to a+} \frac{f(x)}{g(x)} = \lim_{x \to a+} \frac{f'(x)}{g'(x)}$$

Bemerkung:
Entsprechende Aussagen gelten auch für $x \to b$ und obige Limites dürfen auch als verallgemeinerte Grenzwerte interpretiert werden.

Beweis:

(i) Es gelte $(*)$. Sei $a > -\infty$. Wegen $f(x) \xrightarrow{x \to a} 0$ und $g(x) \xrightarrow{x \to a} 0$ kann man durch die Forderung $f(a) := 0$, $g(a) := 0$ die Funktionen f und g zu stetigen Funktionen auf $[a, b)$ fortsetzen. Sei $x \in (a, b)$. Wendet man Satz 1 (S. 29) auf f bzw. g eingeschränkt auf $[a, x]$ an, so folgt die Existenz eines $\xi = \xi(x) \in (a, x)$ mit:

$$\frac{f'(\xi)}{g'(\xi)} = \frac{f(x) - f(a)}{g(x) - g(a)} = \frac{f(x)}{g(x)}$$

Es gilt $\xi = \xi(x) \xrightarrow{x \to a} a$, also folgt die Behauptung.

Sei $a = -\infty$. Man substituiere $t = -\frac{1}{x}$. Unter dieser Substitution wird $(-\infty, b)$ abgebildet auf $(0, -\frac{1}{b})$ (ohne Einschränkung kann man voraussetzen, daß $b < 0$). Für $0 < t < -\frac{1}{b}$ setze man:

$$\varphi(t) := f(-\frac{1}{t}) \qquad \psi(t) := g(-\frac{1}{t})$$

Dann sind φ und ψ differenzierbar mit:

$$\varphi'(t) = f'(-\frac{1}{t})t^{-2} \qquad \psi'(t) = g'(-\frac{1}{t})t^{-2}$$

Insbesondere ist $\psi'(t) \neq 0\ \forall t$ (denn $g'(x) \neq 0\ \forall x$ nach Voraussetzung). Es gilt ferner:

$$\varphi(t) \xrightarrow{t\to 0+} 0 \qquad\qquad \psi(t) \xrightarrow{t\to 0+} 0$$

Ferner existiert nach Voraussetzung:

$$\lim_{t\to 0+} \frac{\varphi'(t)}{\psi'(t)} = \lim_{t\to 0+} \frac{f'(-\frac{1}{t})t^{-2}}{g'(-\frac{1}{t})t^{-2}} = \lim_{x\to -\infty} \frac{f'(x)}{g'(x)}$$

Nach dem schon Bewiesenen (angewandt auf $a = 0$, φ und ψ) folgt daher die Existenz von

$$\lim_{t\to 0+} \frac{\varphi(t)}{\psi(t)}$$

und es gilt:

$$\lim_{t\to 0+} \frac{\varphi(t)}{\psi(t)} = \lim_{t\to 0+} \frac{\varphi'(t)}{\psi'(t)}$$

Wegen

$$\lim_{t\to 0+} \frac{\varphi(t)}{\psi(t)} = \lim_{x\to -\infty} \frac{f(x)}{g(x)}$$

gilt daher:

$$\lim_{x\to -\infty} \frac{f(x)}{g(x)} = \lim_{x\to -\infty} \frac{f'(x)}{g'(x)}$$

(ii) Es gelte $(**)$. Man setze:

$$A := \lim_{x\to a+} \frac{f'(x)}{g'(x)}$$

Wir setzen (der Einfachheit halber) voraus, daß $A \in \mathbb{R}$.
Wegen $f(x) \xrightarrow{x\to a} \infty$, $g(x) \xrightarrow{x\to a} \infty$ und der Existenz von A gibt es c_0 mit $c_0 \in (a, b)$, so daß $f(x) > 0,\ g(x) > 0\ \forall x$ mit $a < x < c_0$ ist und

$$\left|\frac{f'(x)}{g'(x)}\right| = \left|\frac{f'(x)}{g'(x)} - A + A\right| \leq \left|\frac{f'(x)}{g'(x)} - A\right| + |A| \leq 1 + |A| \qquad \forall x \text{ mit } a < x < c_0$$

Sei $\varepsilon > 0$. Wegen der Existenz von A gibt es dann $c = c(\varepsilon)$ mit $a < c < c_0$, so daß:

$$\left|A - \frac{f'(x)}{g'(x)}\right| < \frac{\varepsilon}{2} \qquad \forall x \text{ mit } a < x < c$$

Es gilt ferner:

$$\frac{1 - \frac{g(c)}{g(x)}}{1 - \frac{f(c)}{f(x)}} \xrightarrow{x\to a} 1$$

da $g(c) = const.$ und $g(x) \xrightarrow{x \to a} \infty$, also existiert $x_0 = x_0(\varepsilon)$ mit $a < x_0 < c$, so daß

$$\left| \frac{1 - \frac{g(c)}{g(x)}}{1 - \frac{f(c)}{f(x)}} - 1 \right| < \frac{\varepsilon}{2(1 + |A|)} \qquad \forall x \text{ mit } a < x < x_0$$

nach dem verallgemeinerten Mittelwertsatz (Satz 1, S. 29) gibt es für $x \in (a, c)$ ein $\xi \in (x, c)$ mit

$$\frac{f'(\xi)}{g'(\xi)} = \frac{f(x) - f(c)}{g(x) - g(c)} = \frac{f(x)}{g(x)} \cdot \frac{1 - \frac{f(c)}{f(x)}}{1 - \frac{g(c)}{g(x)}}$$

Für $a < x < x_0$ $(< c < c_0)$ gilt daher:

$$\begin{aligned} \left| A - \frac{f(x)}{g(x)} \right| &= \left| A - \frac{f'(\xi)}{g'(\xi)} \cdot \frac{1 - \frac{g(c)}{g(x)}}{1 - \frac{f(c)}{f(x)}} \right| = \left| A - \frac{f'(\xi)}{g'(\xi)} \left(\frac{1 - \frac{g(c)}{g(x)}}{1 - \frac{f(c)}{f(x)}} - 1 \right) - \frac{f'(\xi)}{g'(\xi)} \right| \\ &\le \left| -\frac{f'(\xi)}{g'(\xi)} \right| \cdot \left| \frac{1 - \frac{g(c)}{g(x)}}{1 - \frac{f(c)}{f(x)}} - 1 \right| + \left| A - \frac{f'(\xi)}{g'(\xi)} \right| \le (1 + |A|) \frac{\varepsilon}{2(1 + |A|)} + \frac{\varepsilon}{2} = \varepsilon \end{aligned}$$

□

Beispiel 2:

$$\lim_{x \to \infty} \frac{x}{\log x} = \lim_{x \to \infty} \frac{1}{\frac{1}{x}} = \lim_{x \to \infty} x = \infty$$

Beispiel 3:
Ausdrücke der Form „$0 \cdot \infty$“ schreibt man um in „$\frac{\infty}{\frac{1}{0}} = \frac{\infty}{\infty}$“, z.B.:

$$\lim_{x \to 0+} x \log x = - \lim_{x \to 0+} \frac{-\log x}{\frac{1}{a}} = - \lim_{x \to 0+} \frac{-\frac{1}{x}}{-\frac{1}{x^2}} = - \lim_{x \to 0+} x = 0$$

Beispiel 4:
Ausdrücke der Form „$\infty - \infty$“ schreibt man um in

$$\frac{1}{0} - \frac{1}{0} = \frac{1 \cdot 0 - 1 \cdot 0}{0 \cdot 0} = \frac{0}{0}$$

z.B.:

$$\begin{aligned} \lim_{x \to 0+} \left(\frac{1}{\sin x} - \frac{1}{x} \right) &= \lim_{x \to 0+} \frac{x - \sin x}{x \sin x} = \lim_{x \to 0+} \frac{1 - \cos x}{\sin x + x \cos x} \\ &= \lim_{x \to 0+} \frac{\sin x}{\cos x + \cos x - x \sin x} = \frac{0}{2} = 0 \end{aligned}$$

§2 Der Satz von TAYLOR

Sei $x_0 \in \mathbb{R}$, $\delta > 0$ und sei $f : U_\delta(x_0) \to \mathbb{R}$ eine auf $U_\delta(x_0)$ n-mal stetig differenzierbare Funktion, wobei $n \in \mathbb{N}_0$

Approximationsproblem: Gibt es ein Polynom $p_n(x)$ vom Grad $\leq n$ derart, daß

$$\frac{f(x) - p_n(x)}{(x - x_0)^n} \xrightarrow{x \to x_0} 0 \qquad (*)$$

Wenn ja, ist p_n eindeutig bestimmt?

Wir machen für p_n probeweise den Ansatz

$$p_n(x) = a_0 + a_1(x - x_0) + a_2(x - x_0)^2 + \cdots + a_n(x - x_0)^n$$

mit noch zu bestimmenden Koeffizienten $a_0, a_1, \ldots, a_n$.

Man definiere:

$$F : \begin{cases} U_\delta(0) \to \mathbb{R} \\ F(h) = f(x_0 + h) - p_n(x_0 + h) \end{cases}$$

Dann ist F auf $U_\delta(0)$ n-mal stetig differenzierbar. Es gilt:

$$F(h) = \frac{f(x_0 + h) - p_n(x_0 + h)}{h^n} h^n \xrightarrow{h \to 0} 0$$

wegen $(*)$ und $n \in \mathbb{N}_0$. Da F stetig, folgt $F(0) = 0$. Sei $n \geq 1$. Dann gilt:

$$\frac{F(h)}{h} = \frac{f(x_0 + h) - p_n(x_0 + h)}{h^n} h^{n-1} \xrightarrow{h \to 0} 0$$

wegen $(*)$ und $n - 1 \geq 0$. Andererseits gilt nach dem Mittelwertsatz:

$$\begin{aligned} \frac{F(h)}{h} &= \frac{F(h) - F(0)}{h} = \begin{cases} \dfrac{F(h) - F(0)}{h} & (h > 0) \\ \dfrac{F(0) - F(h)}{-h} & (h < 0) \end{cases} \\ &= F'(h_1) \end{aligned}$$

mit $h_1 \in (0, h)$ für $h > 0$ und $h_1 \in (h, 0)$ für $h < 0$.

Mit $h \to 0$ (also auch $h_1 \to 0$) folgt also $F'(0) = 0$, denn F' ist stetig.

Sei jetzt $n \geq 2$. Wie oben zeigt man dann $\frac{F(h)}{h^2} \xrightarrow{h \to 0} 0$. Andererseits gilt nach dem verallgemeinerten Mittelwertsatz mit $0 < |h_2| < |h_1| < |h|$:

$$\frac{F(h)}{h^2} = \frac{F(h) - F(0)}{h^2 - 0} = \frac{F'(h_1)}{2h_1} = \frac{1}{2} \frac{F'(h_1) - F'(0)}{h_1} = \frac{1}{2} F''(h_2)$$

Daher folgt mit $h \to 0$, daß $F''(0) = 0$. Man fährt so sukzessive fort und erhält:

$$F^{(\nu)}(0) = 0 \qquad \forall \nu,\ 0 \leq \nu \leq n$$

d.h.:

$$f^{(\nu)}(x_0) = p_n^{(\nu)}(x_0) = \nu!\, a_\nu \qquad \Longrightarrow \qquad a_\nu = \frac{f^{(\nu)}(x_0)}{\nu!}$$

Definition 1 (TAYLORpolynom):
Das Polynom

$$p_n(x) := \sum_{\nu=0}^{n} \frac{f^{(\nu)}(x_0)}{\nu!}(x - x_0)^\nu \tag{+}$$

*heißt das n-*TAYLOR**polynom zu f gebildet an der Stelle** x_0.

Satz 1:
Wenn ein Polynom mit (∗) existiert, dann ist es eindeutig bestimmt und gleich (+)

Bemerkung:
Ist f selbst ein Polynom vom Grad $\leq n$, so gilt $f = p_n$; insbesondere läßt sich also jedes Polynom vom Grad $\leq n$ in der Form

$$\sum_{\nu=0}^{n} b_\nu (x - x_0)^n$$

schreiben, für jedes $x_0 \in \mathbb{R}$, wobei die b_ν's von x_0 abhängen.

Beweis:
Setze $q := f - p_n$. Dann ist q ein Polynom vom Grad $\leq n$, und es gilt:

$$q^{(\nu)}(x_0) = 0 \qquad \text{für } \nu = 0, 1, \ldots, n \tag{++}$$

Sei

$$q(x) = \sum_{\nu=0}^{n} c_\nu x^\nu$$

Man fasse (++) als lineares Gleichungssystem für die Koeffizienten $c_0, c_1, \ldots, c_n$ auf.

Genauer ist (++) äquivalent zu:

$$\begin{pmatrix} 0! & x_0 & x_0^2 & \cdots & x_0^n \\ 0 & 1! & 2x_0 & \cdots & n x_0^{n-1} \\ 0 & 0 & 2! & \cdots & n(n-1)x_0^{n-2} \\ \vdots & \vdots & \cdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & n! \end{pmatrix} \begin{pmatrix} c_0 \\ c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

Die Determinante der Matrix links (obere Dreiecksmatrix) ist gleich:

$$\prod_{\nu=0}^{n} \nu! \neq 0$$

also ist die Matrix invertierbar, also folgt:

$$c_0 = c_1 = \cdots = c_n = 0$$

direkt: aus der letzten Zeile folt: $n!\, c_n = 0 \implies c_n = 0$, daraus folgt sukzessive:

$$c_n = c_{n-1} = \cdots = c_0 = 0$$

□

Satz 2 (TAYLOR)**:**
Sei $x_0 \in \mathbb{R}$, $\delta > 0$ und $f : U_\delta(x_0) \to \mathbb{R}$ $(n+1)$-mal differenzierbar auf $U_\delta(x_0)$, wobei $n \in \mathbb{N}_0$ fest vorgegeben. Dann gilt:

$$f(x) = p_n(x) + R_n(x - x_0) \qquad (|x - x_0| < \delta)$$

wobei

$$p_n(x) := \sum_{\nu=0}^{n} \frac{f^{(\nu)}(x_0)}{\nu!}(x - x_0)^\nu$$

das n-**te** TAYLOR**polynom von** f **gebildet an der Stelle** x_0 *ist und das* **Restglied** $R_n(x-x_0)$ *durch*

LAGRANGE*:*
$$R_n(x - x_0) = \frac{f^{(n+1)}(\xi)}{(n+1)!}(x - x_0)^{n+1}$$

*mit $\xi = x_0 + \vartheta h$, $h = x - x_0$ und $\vartheta = \vartheta_x \in (0, 1)$, gegeben wird. (***Form von** LAGRANGE*)*
Gilt also z. B.:

$$\left|\frac{f^{(n+1)}(y)}{(n+1)!}\right| \leq M \qquad \text{für ein } M > 0,\ \forall y \in U_\delta(x_0)$$

so folgt insbesondere:

$$\frac{f(x) - p_n(x)}{(x - x_0)^n} \xrightarrow{x \to x_0} 0$$

(denn $R_n(x - x_0) = f(x) - p_n(x)$), also löst p_n tatsächlich $(*)$
Für das Restglied hat man ferner die Darstellung:

CAUCHY*:*
$$R_n(x - x_0) = R_n(h) = \frac{h^{n+1}}{n!}(1 - \theta)^n f^{(n+1)}(x_0 + \theta h)$$

*mit $h = x - x_0$, $\theta = \theta_x \in (0, 1)$. (*CAUCHY-**Form***)*
Gilt zudem $f \in C^{n+1}(U_\delta(x_0))$, so kann man das Restglied auch in folgender Form darstellen:

$$R_n(x - x_0) = R_n(h) = \int_0^1 \frac{(1 - t)^n}{n!} f^{(n+1)}(x_0 + th)h^{n+1}dt$$

Beweis:

(i) Man setze:

$$R_n(h) := f(x_0+h) - p_n(x_0+h) \qquad |h| < \delta$$

Dann gilt:

$$R_n^{(\nu)}(0) = f^{(\nu)}(x_0) - p_n^{(\nu)}(x_0) = f^{(\nu)}(x_0) - \nu!\,\frac{f^{(\nu)}(x_0)}{\nu!} = 0 \qquad \nu = 0,1,\ldots,n$$

Es gilt:

$$R_n^{(n+1)}(h) = f^{(n+1)}(x_0+h)$$

denn p_n hat Grad $\leq n$, also ist $p_n^{(n+1)}(h) = 0$. Unter Benutzung des verallgemeinerten Mittelwertsatzes findet man nacheinander $h_1, h_2, \ldots, h_{n+1}$ mit

$$0 < h_{n+1} < h_n < \cdots < h_1 < h \qquad \text{für } h > 0$$

bzw.

$$h < h_a < h_2 < \cdots < h_{n+1} < 0 \qquad \text{für } h > 0$$

so daß:

$$\begin{aligned}
\frac{R_n(h)}{h^{n+1}} &= \frac{R_n(h) - R_n(0)}{h^{n+1} - 0} = \frac{1}{n+1}\cdot\frac{R_n'(h_1)}{h_1^n} = \frac{1}{n+1}\cdot\frac{R_n'(h_1) - R_n'(0)}{h_1^n - 0} \\
&= \frac{1}{n(n+1)}\cdot\frac{R_n''(h_2)}{h_2^{n-1}} = \cdots = \frac{1}{n+1!}\cdot\frac{R_n^{(n+1)}(h_{n+1})}{h_{n+1}^0} \\
&= \frac{1}{n+1!}R_n^{(n+1)}(h_{n+1}) = \frac{f^{(n+1)}(x_0+h_{n+1})}{(n+1)!}
\end{aligned}$$

Man schreibe $h_{n+1} = \vartheta h$ mit $\vartheta \in (0,1)$.

Dann folgt also:

$$R_n(h) = \frac{f^{(n+1)}(x_0+\vartheta h)}{(n+1)!}h^{n+1}$$

Dies ist die LAGRANGE-Form des Restgliedes.

(ii) Sei jetzt $f \in C^{n+1}(U_\delta(x_0))$. Mit $h = x - x_0$ fest setze man:

$$F(t) := f(x_0 + th) \qquad t \in [0,1]$$

Man schreibe:

$$F(1) - F(0) = \int_0^1 F'(t)dt \qquad \text{(Hauptsatz)}$$

nach partieller Integration ($u = (t-1) \Longrightarrow u' = 1$ und $v = F'(t) \Longrightarrow v' = F''(t)$):

$$= (t-1)F'(t)\Big|_0^1 - \int_0^1 (t-1)F''(t)dt = F'(0) + \int_0^1 (1-t)F''(t)dt$$

nach paritieller Integration ($u = \dfrac{(1-t)^2}{2!} \Longrightarrow u' = (1-t)$ und $v = F''(t) \Longrightarrow v' = F'''(t)$):

$$= F'(0) + \frac{(1-t)^2}{2!}F''(t)\Big|_0^1 - \int_0^1 \frac{(1-t)^2}{2}F'''(t)dt$$

Durch vollständige Induktion ergibt sich:

$$F(1) - F(0) = F'(0) + \frac{F''(0)}{2!} + \cdots + \frac{F^{(n)}(0)}{n!} + \int_0^1 \frac{(1-t)^n}{n!}F^{(n+1)}(t)dt$$

Es gilt:

$$F(1) = f(x_0 + h) = f(x) \qquad F(0) = f(x_0)$$
$$F^{(\nu)}(0) = f^{(\nu)}(x_0)h^\nu \qquad \text{da } F^{(\nu)}(h) = f^{(\nu)}(x_0 + th)h^\nu \qquad (0 \le \nu \le n+1)$$

Daher ergibt sich die Integralform des Restgliedes mit:

$$f(x) = f(x_0) + \frac{f'(x_0)}{1!}(x-x_0) + \frac{f''(x_0)}{2!}(x-x_0)^2 + \ldots$$
$$\cdots + \frac{f^{(n)}(x_0)}{n!}(x-x_0)^n + \int_0^1 \frac{(1-t)^n}{n!}f^{(n+1)}(x_0+th)h^{n+1}dt$$

(iii) Wendet man auf das Integral den Mittelwertsatz der Integralrechnung (Ana. 1, Kap. 3, §8) an, so ergibt sich:

$$\int_0^1 \frac{(1-t)^n}{n!}f^{(n+1)}(x_0+th)h^{n+1}dt = (1-0)\frac{(1-\theta)^n}{n!}f^{(n+1)}(x_0+\theta h)h^{n+1}$$
$$= \frac{(1-\theta)^n}{n!}f^{(n+1)}(x_0+\theta h)h^{n+1}$$

wobei $\theta = \theta_x \in (0,1)$. Dies ist die CAUCHY-Form des Restgliedes.

(*Anmerkung:* Die CAUCHY-Form des Restgliedes wird hier nicht ohne die Voraussetzung $f \in C^{n+1}(U_\delta(x_0))$ bewiesen, sie ist jedoch auch ohne diese gültig.)

□

Definition 2:
Sei I ein Intervall. Dann sei

$$C^\infty(I) = \bigcap_{n\in\mathbb{N}} C^n(I)$$

der Raum der beliebig oft stetig differenzierbar Funktionen $f : I \to \mathbb{R}$

Sei $f \in C^\infty(U_\delta(x_0))$. Dann kann

$$p_n(x) = \sum_{\nu=0}^{n} \frac{f^{(\nu)}(x_0)}{\nu!}(x - x_0)^\nu$$

für jedes $n \in \mathbb{N}$ gebildet werden. Man kann offenbar $p_n(x)$ als n-te Partialsumme der unendlichen Reihe

$$\sum_{\nu=0}^{\infty} \frac{f^{(\nu)}(x_0)}{\nu!}(x - x_0)^\nu$$

auffassen.

Definition 3 (Taylorreihe):
Die Reihe

$$\sum_{\nu=0}^{\infty} \frac{f^{(\nu)}(x_0)}{\nu!}(x - x_0)^\nu$$

heißt Taylor**reihe** *von f gebildet an der Stelle x_0.*

Probleme:

(i) Für welche x ist die Taylorreihe konvergent?

(ii) Konvergiert sie gegen den Funktionswert $f(x)$?

Hierzu muß man

$$f(x) - p_n(x) = R_n(x - x_0) \qquad \text{für } n \to \infty$$

betrachten.

Beispiel 5:
Sei $f(x) := \mathrm{e}^x$ $(x \in \mathbb{R})$ und $x_0 = 0$. Dann ist $f^{(\nu)}(x) = \mathrm{e}^x$, also gilt nach Taylor:

$$\mathrm{e}^x = 1 + \frac{x}{1!} + \frac{x^2}{2!} + \cdots + \frac{x^n}{n!} + \frac{\mathrm{e}^{\vartheta x}}{(n+1)!}x^{n+1}$$

mit $\vartheta = \vartheta_{x,n} \in (0,1)$. Es gilt:

$$|R_n(x)| = \frac{|x|^{n+1}}{(n+1)!}\,\mathrm{e}^{\vartheta x} \leq \frac{|x|^{n+1}}{n+1!}\,\mathrm{e}^{|x|} \xrightarrow{n\to\infty} 0$$

denn $\sum\limits_{\nu=0}^{\infty} \frac{x^\nu}{\nu!}$ konvergiert $\forall x \in \mathbb{R}$ (Quotientenkriterium, Analysis 1).

Daraus folgt, daß gilt: $\frac{x^\nu}{\nu!} \xrightarrow{\nu\to\infty} 0$.

Die TAYLORreihe von e^x gebildet an der Stelle $x_0 = 0$ konvergiert also $\forall x \in \mathbb{R}$ und ist gleich dem Funktionswert e^x.

Insbesondere gilt mit $x = 1$:

$$\mathrm{e} = \sum_{\nu=0}^{\infty} \frac{1}{\nu!}$$

Bemerkung:
Dies zeigt, daß die Definition von e als $\mathrm{e} = \lim\limits_{n\to\infty}(1+\frac{1}{n})^n$ (Analysis 1, Kap. 2, §6, Bem. zu Satz 3) („offizielle Definition“) und von e als $\mathrm{e} = \sum\limits_{\nu=0}^{\infty} \frac{1}{\nu!}$ (Analysis 1, Kap. 3, §4, Bem. zu Bsp. iii) übereinstimmen.

Beispiel 6:
Sei $f(x) = \sin x$ $(x \in \mathbb{R})$ und $x_0 = 0$. Wegen $\sin' x = \cos x$, $\cos' x = -\sin x$ und $\sin 0 = 0$, $\cos 0 = 1$ findet man für die TAYLORreihe von $\sin x$ um $x_0 = 0$ den Ausdruck:

$$x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} \pm \ldots$$

Wegen $|\sin x| \leq 1$, $|\cos x| \leq 1$ $\forall x \in \mathbb{R}$ findet man ähnlich wie in Beispiel 5, daß

$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} \pm \ldots \qquad (x \in \mathbb{R})$$

Genauso erhält man

$$\cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} \pm \ldots \qquad (x \in \mathbb{R})$$

Beispiel 7:
Sei $f : \mathbb{R} \to \mathbb{R}$ gegeben durch:

$$f(x) := \begin{cases} \mathrm{e}^{-\frac{1}{x^2}} & x \neq 0 \\ 0 & x = 0 \end{cases}$$

Man kann zeigen (Übungsaufgabe), daß f von der Klasse $C^\infty(\mathbb{R})$ ist, und daß

$$f^{(\nu)}(0) = 0 \qquad \forall \nu \in \mathbb{N}_0$$

Daher ist die TAYLORreihe von f in $x_0 = 0$ identisch Null, konvergiert also insbesondere $\forall x \in \mathbb{R}$. Aber $f(x) \neq 0$ $\forall x \neq 0$, also konvergiert die TAYLORreihe nur im Entwicklungspunkt $x_0 = 0$ gegen den Funktionswert $f(x)$.

Satz 3:
Sei $f \in C^\infty(U_\delta(x_0))$. Es gebe ein $c > 0$, so daß

$$\frac{1}{\nu!}|f^{(\nu)}(x)| \leq \frac{c}{\delta^\nu} \qquad \forall \nu \in \mathbb{N}_0 \text{ und } \forall x \in U_\delta(x_0)$$

dann ist die TAYLOR*reihe von f gebildet in x_0 konvergent gegen $f(x)$ $\forall x \in U_\delta(x_0)$.*

Beweis:
Nach dem Satz von TAYLOR gilt für das Restglied:

$$|R_n(x - x_0)| = \frac{f^{(n+1)}\big(x_0 + \vartheta(x - x_0)\big)}{(n+1)!}|x - x_0|^{n+1} \qquad \forall n \in \mathbb{N}_0$$

Nach Voraussetzung gilt:

$$|R_n(x - x_0)| \leq \frac{c}{\delta^{n+1}}|x - x_0|^{n+1} = c\left(\frac{|x - x_0|}{\delta}\right)^{n+1} \xrightarrow{n\to\infty} 0$$

denn für $x \in U_\delta(x_0)$ gilt $\frac{|x-x_0|}{\delta} < 1$. □

§3 Potenzreihen

Definition 1 (Potenzreihe):
Eine unendliche Reihe der Form

$$\sum_{\nu=0}^{\infty} a_\nu (x - x_0)^\nu$$

heißt **Potenzreihe** *mit Entwicklungspunkt x_0. Die Zahlen a_ν heißen* **Koeffizienten** *der Reihe. Die Menge der $x \in \mathbb{R}$, für die die Reihe konvergiert heißt* **Konvergenzbereich**.

Beispiele:

(i) geometrische Reihe $\sum\limits_{n=0}^{\infty} x^n$

(ii) TAYLORreihe

Lemma 1:
Konvergiert die Reihe

$$\sum_{\nu=0}^{\infty} a_\nu (x - x_0)^\nu$$

für $x = x_1$ und setzt man $r_1 := |x_1 - x_0|$, dann konvergiert die Reihe für alle $x \in U_{r_1}(x_0)$. Ferner konvergiert die Reihe sogar gleichmäßig absolut für $|x - x_0| \leq \rho < r_1$, für jedes $\rho < r_1$.

Beweis:
Da $\sum_{\nu=0}^{\infty} a_\nu(x_1-x_0)^\nu$ nach Voraussetzung konvergiert, bilden die Glieder der Reihe notwendigerweise eine Nullfolge, also insbesondere eine beschränkte Folge.
Es gibt also $c > 0$, so daß

$$|a_\nu| \cdot r_1^\nu = |a_\nu(x_1-x_0)^\nu| \le c \qquad \forall \nu \in \mathbb{N}_0$$

d.h. also

$$|a_\nu| \le c \cdot r_1^{-\nu} \qquad \forall \nu \in \mathbb{N}_0$$

Sei $x \in \mathbb{R}$ mit $|x-x_0| \le \rho < r_1$. Dann gilt:

$$|a_\nu(x-x_0)^\nu| = |a_\nu| \cdot |x-x_0|^\nu \le c \cdot r_1^{-\nu} \cdot \rho^\nu = c\left(\frac{\rho}{r_1}\right)^\nu = c \cdot q^\nu \qquad \forall \nu \in \mathbb{N}_0$$

wobei $q := \frac{\rho}{r_1} \in [0,1)$.

Daher ist die geometrische Reihe $c\sum_{\nu=0}^{\infty} q^\nu$ eine Majorante für $\sum_{\nu=0}^{\infty} a_\nu(x-x_0)^\nu$, daher folgt die Behauptung (Analysis 1, Kap. 4, §4, Satz 3). □

Korollar 1:
Wenn $\sum_{\nu=0}^{\infty} a_\nu(x-x_0)^\nu$ *in* $x = x_1$ *nicht absolut konvergiert, so divergiert die Reihe für alle* x *mit* $|x-x_0| > |x_0-x_1|$

Beweis:
klar!

Satz 1:
Zu jeder vorgegebenen Potenzreihe

$$\sum_{n=0}^{\infty} a_n(x-x_0)^n$$

existiert genau ein $r \in \overline{\mathbb{R}}$ *mit* $0 \le r \le \infty$, **Konvergenzradius** *genannt, so daß gilt:*

(i) *Die Reihe ist gleichmäßig absolut konvergent für* $|x-x_0| \le \rho < r$ *für jedes* $\rho < r$.

(ii) *Die Reihe divergiert für* $|x-x_0| > r$.

(iii) *Der Konvergenzradius* r *berechnet sich aus den Koeffizienten der Reihe als:*

Formel von CAUCHY-HADAMAR*:*
$$r = \frac{1}{\limsup\limits_{n\to\infty} \sqrt[n]{|a_n|}}$$

Erinnerung: Ist $\left(b_n\right)_{n\in\mathbb{N}}$ eine folge reeller Zahlen mit $b_n \geq 0\ \forall n \in \mathbb{N}$, und ist $\left(b_n\right)_{n\in\mathbb{N}}$ nach oben beschränkt, so setzt man:

$$\limsup_{n\to\infty} b_n = \sup M$$

wobei M die Menge der Häufungspunkte von $\left(b_n\right)_{n\in\mathbb{N}}$ bezeichnet.

Ist $\left(b_n\right)_{n\in\mathbb{N}}$ nicht nach oben beschränkt, so setzt man:

$$\limsup_{n\to\infty} b_n = \infty$$

Ist $\limsup\limits_{n\to\infty} \sqrt[n]{|a_n|}$ gleich 0 oder ∞, so hat man die Formel als $r = \infty$ bzw. $r = 0$ zu interpretieren. Im ersten Fall konvergiert die Reihe $\forall x \in \mathbb{R}$, im zweiten Fall nur für $x = x_0$. Für $0 < r < \infty$ ist der Konvergenzbereich ein Intervall mit Mittelpunkt $x_0$und der Länge $2r$. Ob die Endpunkte $x_0 + r$ oder $x_0 - r$ zum Konvergenzbereich gehören, muß in jedem Fall gesondert untersucht werden.

Beweis:

(i) folgt aus dem Lemma

(ii) folgt aus dem Lemma

(iii) Sei r durch

$$r := \frac{1}{\limsup\limits_{n\to\infty} \sqrt[n]{|a_n|}}$$

definiert.

1. Fall: $0 < r < \infty$, $r_0 < r$

Sei r_0 beliebig mit $0 < r_0 < r$, dann gilt:

$$\limsup_{n\to\infty} \sqrt[n]{|a_n|} = \frac{1}{r} < \frac{1}{r_0}$$

Nach Definition des Limes Superior bzw. des Häufungspunktes gibt es dann $N \in \mathbb{N}$, so daß:

$$|a_n|^{\frac{1}{n}} < \frac{1}{r_0} \qquad \forall n > N$$

Also gilt:

$$|a_n| < \frac{1}{r_0^n} = r_0^{-n} \qquad \forall n > N$$

Daher gilt:

$$|a_n(x - x_0)^n| < \left(\frac{|x - x_0|}{r_0}\right)^n < 1 \qquad \forall n > 0, \text{ für } |x - x_0| < r_0$$

Also konvergiert die Reihe

$$\sum_{n=0}^{\infty} a_n (x - x_0)^n$$

für $|x - x_0| < r_0$. (Majorantenkriterium und Vergleich mit der geometrischen Reihe), da $0 < r_0 < r$, folgt die Konvergenz für $|x - x_0| < r$.

2. Fall $0 \leq r < \infty$, $r < r_0$

Sei $r < r_0 < \infty$, also

$$\frac{1}{r_0} < \frac{1}{r} = \limsup_{n\to\infty} \sqrt[n]{|a_n|}$$

Nach Definition des Limes superior und des Häufungspunktes existiert daher eine Folge $n_1 < n_2 < n_3 < \ldots$, so daß

$$\frac{1}{r_0} < |a_{n_p}|^{\frac{1}{n_p}} \qquad \forall p \in \mathbb{N}$$

d.h.:

$$|a_{n_p}| > r_0^{-n_p}$$

Sei jetzt $|x - x_0| > r_0$. Dann folgt

$$|a_{n_p}(x - x_0)^{n_p}| = |a_{n_p}| \cdot |x - x_0|^{n_p} \geq r_0^{-n_p} r_0^{n_p} = 1 \qquad \forall p \in \mathbb{N}$$

Daher können die Glieder der Potenzreihe $\sum\limits_{n=0}^{\infty} a_n (x - x_0)^n$ keine Nullfolge bilden, also ist die Reihe divergent.

Da r_0 beliebig mit $r < r_0 < \infty$, divergiert die Reihe also für $|x - x_0| > r$.

$\square$

Bemerkung:

Wenn $\left(b_n\right)_{n\in\mathbb{N}}$ konvergiert, so gilt:

$$\limsup_{n\to\infty} b_n = \lim_{n\to\infty} b_n$$

Beispiel 8:

Die Reihe $\sum\limits_{n=0}^{\infty} n^n x^n$ hat den Konvergenzradius $r = 0$, denn

$$\frac{1}{\limsup\limits_{n\to\infty} \sqrt[n]{|a_n|}} = \frac{1}{\limsup\limits_{n\to\infty} n} = \frac{1}{\infty} = 0$$

Beispiel 9:
Jede der Reihen

$$\sum_{n=0}^{\infty} x^n \qquad \sum_{n=0}^{\infty} \frac{x^n}{n} \qquad \sum_{n=0}^{\infty} \frac{x^n}{n^2}$$

hat $r = 1$, denn $\sqrt[n]{n} \xrightarrow{n\to\infty} 1$. Die erste Reihe ist in keinem der Randpunkte $+1, -1$ konvergent (allgemeines Glied keine Nullfolge), die zweite Reihe ist in $+1$ divergent („harmonische Reihe“), in -1 konvergent („alternierende Reihe“), die dritte konvergiert sowohl in $+1$, als auch in -1.

Beispiel 10:
Wir haben schon gezeigt, daß

$$\sum_{n=0}^{\infty} \frac{x^n}{n!}$$

$\forall x \in \mathbb{R}$ konvergiert (nach dem Quotientenkriterium), also hat diese Reihe einen Konvergenzradius $r = \infty$.

Also gilt:

$$\infty = \frac{1}{\limsup\limits_{n\to\infty} \sqrt[n]{\frac{1}{n!}}} = \limsup_{n\to\infty} \sqrt[n]{n!}$$

Es gilt also:

$$\frac{1}{\limsup\limits_{n\to\infty} \sqrt[n]{n!}} = 0$$

Daher hat $\sum\limits_{n=0}^{\infty} n!\, x^n$ den Konvergenzradius $r = 0$.

Beispiel 11:
Hat

$$\sum_{n=0}^{\infty} a_n (x - x_0)^n$$

den Konvergenzradius r, so hat auch die gliedweise differenzierte Potenzreihe

$$\sum_{n=0}^{\infty} n \cdot a_n (x - x_0)^{n-1}$$

den Konvergenzradius r.

Denn $\sum\limits_{n=0}^{\infty} n \cdot a_n (x - x_0)^n$ hat den Konvergenzradius r, denn $\sqrt[n]{n} \xrightarrow{n\to\infty} 1$, und

$$\left(\sum_{n=0}^{\infty} n \cdot a_n (x - x_0)^{n-1} \right) (x - x_0) = \sum_{n=0}^{\infty} n \cdot a_n (x - x_0)^n$$

Daher konvergiert die linke Reihe genau dann, wenn die rechte Reihe konvergiert.

Satz 2 (Vertauschung von Differentiation und Grenzwert):
Sei $I \subset \mathbb{R}$ ein Intervall, dann gilt:

(i) Sei $\left(f_n\right)_{n\in\mathbb{N}}$ eine Folge von Funktionen, $f_n \in C^1(I)$ mit $f_n \to f$ $(n \to \infty)$ auf I und $f_n' \rightrightarrows g$ $(n \to \infty)$ auf I (gleichmäßige Konvergenz). Dann ist $f \in C^1(I)$ und $f' = g$.

(ii) Sei $\left(h_n\right)_{n\in\mathbb{N}}$ eine Funktionenfolge mit $h_n \in C^1(I)$ derart, daß die unendliche Reihe

$$\sum_{n=1}^{\infty} h_n(x)$$

auf I konvergiert, während die Reihe der Ableitungen

$$\sum_{n=1}^{\infty} h_n'(x)$$

auf I sogar gleichmäßig konvergiert, so ist die Funktion

$$x \mapsto \sum_{n=1}^{\infty} h_n(x)$$

von der Klasse $C^1(I)$ und es gilt:

$$\frac{d}{dx}\left(\sum_{n=1}^{\infty} h_n(x)\right) = \sum_{n=1}^{\infty} h_n'(x) \qquad (x \in I)$$

Beweis:

(i) Da $f_n' \in C^0(I)$, und wegen der gleichmäßigen Konvergenz $f_n' \rightrightarrows g$ $(n \to \infty)$ auf I folgt $g \in C^0(I)$ (Analysis 1, Kap. 5, §4, Satz 2). Seien $c, x \in I$ mit c fest. Dann gilt ferner wegen der gleichmäßigen Konvergenz: (Analysis 1, Kap. 5, §6)

$$\lim_{n\to\infty} \int_c^x f_n'(t)dt = \int_c^x \lim_{n\to\infty} f_n'(t)dt = \int_c^x g(t)dt$$

Nach dem Hauptsatz der Differtial- und Integralrechnung gilt:

$$\int_c^x f_n'(t)dt = f_n(x) - f_n(c)$$

Wegen $f_n \to f$ $(n \to \infty)$ folgt daher:

$$f(x) = f(c) + \int_c^x g(t)dt$$

Nach dem Hauptsatz der Differtial- und Integralrechnung folgt also:

$$f \in C^1(I) \qquad \text{und} \qquad f'(x) = g(x)$$

(ii) folgt aus (i). □

Satz 3:
Sei $\sum\limits_{n=0}^{\infty} a_n(x-x_0)^n$ *eine Potenzreihe mit Konvergenzradius* $r > 0$ *und sei*

$$f(x) := \sum_{n=0}^{\infty} a_n(x-x_0)^n$$

die durch die Potenzreihe dargestellte Funktion $(x \in U_r(x_0))$. *Dann ist* $f \in C^\infty\big(U_r(x_0)\big)$ *und die Ableitungen* $f', f'', \ldots$ *werden durch Potenzreihen mit Konvergenzradius* r *dargestellt, nämlich durch die durch gliedweise Differentiation gewonnene Potenzreihe*

$$\begin{aligned} f'(x) &= \sum_{n=1}^{\infty} n \cdot a_n(x-x_0)^{n-1} \\ f''(x) &= \sum_{n=2}^{\infty} n(n-1) \cdot a_n(x-x_0)^{n-2} \\ &\vdots \\ f^{(\nu)}(x) &= \sum_{n=0}^{\infty} n(n-1) \cdot \ldots \cdot (n-\nu) \cdot a_n(x-x_0)^{n-\nu} \end{aligned}$$

Insbesondere ist:

$$a_n = \frac{f^{(n)}(x_0)}{n!}$$

Beweis:
Nach Beispiel 8 (S. 43) hat

$$\sum_{n=1}^{\infty} n \cdot a_n(x-x_0)^{n-1}$$

wieder den Konvergenzradius r, und nach dem Lemma 1 (S. 40) ist sie für $|x-x_0| \leq \rho < r$ sogar gleichmäßig konvergent.

Nach Satz 2 (ii) (S. 45) mit $h_n = a_n(x-x_0)^n$ folgt daher:

$$f \in C^1(U_r(x_0)) \qquad \text{und} \qquad f'(x) = \sum_{n=1}^{\infty} n \cdot a_n(x-x_0)^{n-1}$$

Man iteriere obigen Schluß bliebig oft, so folgt die Behauptung. Die Formel für a_n folgt, indem man in dem Ausdruck für $f^{(\nu)}$ $x = x_0$ setzt:

$$f^{(\nu)}(x_0) = \sum_{n=0}^{\infty} n(n-1) \cdot \ldots \cdot (n-\nu) \cdot a_n \cdot 0^{n-\nu}$$

$$\Longrightarrow \qquad f^{(\nu)}(x_0) = \nu! \cdot a_n \qquad \Longrightarrow \qquad a_n = \frac{f^{(\nu)}(x_0)}{\nu!}$$

$\square$

Satz 4 (Identitätssatz für Potenzreihen):
Seien $\sum\limits_{\nu=0}^{\infty} a_\nu (x-x_0)^\nu$ *und* $\sum\limits_{\nu=0}^{\infty} b_\nu (x-x_0)^\nu$ *zwei Potenzreihen mit Konvergenzradius* $r > 0$.
Sei $\big(x_n\big)_{n\in\mathbb{N}}$ *eine Folge von Zahlen mit* $x_n \neq x_0$ $\forall n \in \mathbb{N}$, $x_n \xrightarrow{n\to\infty} x_0$. *Es gelte:*

$$\sum_{\nu=0}^{\infty} a_\nu (x_n - x_0)^\nu = \sum_{\nu=0}^{\infty} b_\nu (x_n - x_0)^\nu \qquad \forall n \in \mathbb{N}$$

Dann gilt:

$$a_\nu = b_\nu \quad \forall \nu \in \mathbb{N}_0$$

In anderen Worten: Stimmen zwei Potenzreihen auf einer Punktmenge M *überein, die* x_0 *als Häufungspunkt hat, so sind sie identisch.*

Beweis:
Angenommen es gibt $\nu \in \mathbb{N}_0$ mit $a_\nu \neq b_\nu$. Sei m das kleinste ν mit dieser Eigenschaft. Dann gilt also $a_\nu = b_\nu$ für $0 \leq \nu \leq m-1$, $a_m \neq b_m$. Aus der Voraussetzung folgt dann:

$$\sum_{\nu=m}^{\infty} a_\nu (x_n - x_0)^\nu = \sum_{\nu=m}^{\infty} b_\nu (x_n - x_0)^\nu \qquad \forall n \in \mathbb{N}$$

wegen $x_n \neq x_0$ $\forall n$ ergibt sich:

$$\sum_{\nu=m}^{\infty} a_\nu (x_n - x_0)^{\nu-m} = \sum_{\nu=m}^{\infty} b_\nu (x_n - x_0)^{\nu-m} \tag{$*$}$$

Nach Lemma 1 (S. 40) bzw. Satz 1 (S. 41) konvergieren $\sum\limits_{\nu=0}^{\infty} a_\nu (x-x_0)^\nu$ und $\sum\limits_{\nu=0}^{\infty} b_\nu (x-x_0)^\nu$ für x nahe bei x_0 gleichmäßig und bei gleichmäßig konvergenten Reihen dürfen Limes und Summation vertauscht werden. Daher folgt aus $(*)$ mit $n \to \infty$, wegen $\lim\limits_{n\to\infty} x_n = x_0$, daß

$$a_m = b_m \qquad \lightning \text{Widerspruch!}$$

$\square$

Satz 5 (Umordnungssatz):
Sei $\left(a_{ik}\right)_{i,k\in\mathbb{N}}$ *eine Doppelfolge reeller Zahlen (d. h.:* $a : \mathbb{N}\times\mathbb{N}\to\mathbb{R}$*). Es gebe* $K>0$*, so daß:*

$$\sum_{i,k=1}^{n} |a_{ik}| \le K \quad \forall n\in\mathbb{N} \tag{$*$}$$

Dann gilt:

(i) Es gibt $s\in\mathbb{R}$*, so daß für jede bijektive Abbildung* $t:\mathbb{N}\to\mathbb{N}\times\mathbb{N}$ *die Reihe*

$$\sum_{n=1}^{\infty} a_{t(n)}$$

(absolut) konvergiert und die Summe s *hat. Hierfür schreibt man einfach:*

$$s=\sum_{i,k=1}^{\infty} a_{ik}$$

wobei die Terme der Doppelreihe in irgendeiner Weise anzuordnen sind.

(ii) Die **„Zeilensummen“** *und die* **„Spaltensummen“**

Zeilensummen: $$\sum_{k=1}^{\infty} a_{ik} \qquad (i\in\mathbb{N})$$

Spaltensummen: $$\sum_{i=1}^{\infty} a_{ik} \qquad (k\in\mathbb{N})$$

sind (absolut) konvergente Reihen. Für die Summen:

$$z_i := \sum_{k=1}^{\infty} a_{ik} \quad (i\in\mathbb{N}) \qquad\qquad s_k := \sum_{i=1}^{\infty} a_{ik} \quad (k\in\mathbb{N})$$

gilt:

$$s=\sum_{i=1}^{\infty} z_i = \sum_{k=1}^{\infty} s_k$$

Beweis:

(i) Sei $t_0:\mathbb{N}\to\mathbb{N}\times\mathbb{N}$ eine Bijektion. Dann ist mit $(*)$ die Reihe $\sum\limits_{n=1}^{\infty} a_{t_0(n)}$ absolut konvergent, ihr Grenzwert sei $s\in\mathbb{R}$.
Sei $t:\mathbb{N}\to\mathbb{N}\times\mathbb{N}$ eine weitere Bijektion. Dann ist auch $\sum\limits_{n=1}^{\infty} a_{t(n)}$ eine absolut konvergente Reihe und eine Umordnung von $\sum\limits_{n=1}^{\infty} a_{t_0(n)}$ bzgl. der Bijektion

$$\tau : \begin{cases} \mathbb{N}\to\mathbb{N} \\ \tau := t_0^{-1}\circ t \end{cases}$$

Nach Satz aus der Analysis 1 gilt dann:

$$s = \sum_{n=1}^{\infty} a_{t(n)}$$

(ii) Analog wie in (i) zeigt man, daß die Zeilen- und Spaltensummen absolut konvergieren:

$$z_i := \sum_{k=1}^{\infty} a_{ik} \quad (i \in \mathbb{N}) \qquad\qquad s_k := \sum_{i=1}^{\infty} a_{ik} \quad (k \in \mathbb{N})$$

Man setze:

$$\sigma_n := \sum_{i,k=1}^{n} a_{ik} \qquad\qquad \alpha_n := \sum_{i,k=1}^{n} |a_{ik}| \quad (n \in \mathbb{N})$$

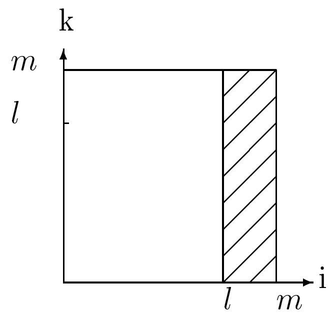
Aus (i) folgt, daß $\big(\sigma_n\big)_{n\in\mathbb{N}}$ konvergent mit dem Grenzwert s ist, aus $(*)$ folgt, daß $\big(\alpha_n\big)_{n\in\mathbb{N}}$ eine monoton wachsende beschränkte Folge ist, also konvergent.

Behauptung:

$$s = \sum_{n=1}^{\infty} z_n = \lim_{N\to\infty} \sum_{n=1}^{N} z_n$$

Sei $\varepsilon > 0$ beliebig vorgegeben. Dann existiert ein $N \in \mathbb{N}$, so daß für $N < l < m$ $(m, l \in \mathbb{N})$ gilt:

$$\alpha_l - \alpha_m = \sum_{\substack{i,k\leq m \\ i>l \ \wedge \ k>l}} |a_{ik}| < \varepsilon$$



Für solche m, l gilt nun

$$\left| \sum_{i=n}^{l} \sum_{k=1}^{m} a_{ik} - \sigma_m \right| \leq \sum_{\substack{l<i\leq m \\ 1\leq k\leq m}} |a_{ik}| \leq \alpha_m - \alpha_l < \varepsilon$$

Für $m \to \infty$ erhält man:

$$\left|\sum i = 1^l \left(\sum_{k=1}^{\infty} a_{ik}\right) - s\right| \leq \varepsilon \quad \Longrightarrow \quad \left|\sum i = 1^l z_i - s\right| \leq \varepsilon \qquad \forall l \in \mathbb{N} \text{ mit } l > N$$

Also gilt die Behauptung.

Für $l < m$ gilt:

$$\sum_{i=1}^{l}\left|\sum_{k=1}^{m} a_{ik}\right| \leq \sum_{i,k \leq m} |a_{ik}| \leq K \qquad \Longrightarrow \qquad \sum_{i=1}^{l} |z_i| \leq K$$

Hieraus folgt die absolute Konvergenz von $\sum\limits_{i=1}^{\infty} z_i$.

Analog für die Spaltensummen. □

Satz 6 (Multiplikation unendlicher Reihen):
Sind $\sum\limits_{i=1}^{\infty} a_i$, $\sum\limits_{i=1}^{\infty} b_i$ absolut konvergente Reihen, so erfüllt die Doppelfolge $\left(a_i b_k\right)_{(i,k)\in\mathbb{N}\times\mathbb{N}}$ die Voraussetzungen von Satz 5 (S. 48). Folglich konvergiert die Doppelreihe

$$\sum_{i,k=1}^{\infty} a_i b_k$$

absolut und es gilt:

$$\sum_{i,k=1}^{\infty} a_i b_k = \left(\sum_{i=1}^{\infty} a_i\right)\left(\sum_{k=1}^{\infty} b_k\right) = \sum_{n=1}^{\infty}\left(\sum_{m=1}^{n} a_m b_{n-m}\right)$$

CAUCHY-**Produkt**

Beweis:
Es gilt für $n \in \mathbb{N}$:

$$\sum_{i,k=1}^{n} |a_i b_k| = \left(\sum_{i=1}^{n} |a_i|\right)\left(\sum_{k=1}^{n} |b_k|\right) \leq \underbrace{\left(\sum_{i=1}^{\infty} |a_i|\right)\left(\sum_{k=1}^{\infty} |b_k|\right)}_{K}$$

Der Rest folgt aus Satz 5. □

Satz 7:
Sind $\sum\limits_{n=0}^{\infty} a_n (x-x_0)^n$, $\sum\limits_{n=0}^{\infty} b_n (x-x_0)^n$ zwei für $0 \leq |x - x_0| < r$ konvergente Potenzreihen $(r > 0)$ mit Summen $f(x)$ und $g(x)$, so gilt:

$$\sum_{n=0}^{\infty}\left(\sum_{n=0}^{m} a_m b_{n-m}\right)(x - x_0)^n$$

ist eine in $U_r(x_0)$ konvergente Potenzreihe mit Summe $f(x)g(x)$.

Beweis:
Folgt aus Satz 1 (S. 41) und Satz 6 (S. 50).

Satz 8 (ABEL**'scher Grenzwertsatz):**
Die Potenzreihe $\sum\limits_{n=0}^{\infty} a_n(x-x_0)^n$ *habe den Konvergenzradius* $r > 0$ *und sei im Punkt* $x = x_0 + r$ *noch konvergent, dann gilt: die Funktion*

$$f(x) = \sum_{n=0}^{\infty} a_n(x-x_0)^n \qquad (x_0 - r < x \leq x_0 + r)$$

ist stetig und es gilt:

$$\lim_{x \to x_0 + r-} f(x) = f(x_0 + r)$$

Beispiel 12:
Es gilt:

$$\arctan x = x - \frac{x^3}{3} + \frac{x^5}{5} \pm \dots \qquad (|x| < 1)$$

Hier konvergiert die rechte Seite für $x = 1$ und nach Satz 8 folgt:

$$\arctan 1 = 1 - \frac{1}{3} + \frac{1}{5} \pm \dots = \frac{\pi}{4}$$

Beweis (von Satz 8):
Man kann ohne Einschränkung $x_0 = 0$ und $r = 1$ annehmen, ansonsten betrachte man:

$$g(x) := f(\frac{x}{r} + x_0) = \sum_{n=0}^{\infty} \frac{a_n}{r^n} x^n \qquad (-1 < x \leq 1)$$

Sei also $x_0 = 0$, $r = 1$ und $\sum\limits_{n=0}^{\infty} a_n$ konvergent.
Man setze

$$s_n := \sum_{k=n}^{\infty} a_k \qquad (n \in \mathbb{N}_0)$$

Es gilt:

(i) $\lim\limits_{n\to\infty} s_n = 0 \implies \exists\ c > 0\ \forall n \in \mathbb{N}_0$ mit $|s_n| \leq c$

(ii) $s_0 = \sum\limits_{n=0}^{\infty} a_n$

(iii) $s_{n+1} - s_n = -a_n \quad (n \in \mathbb{N}_0)$

Behauptung:

$$\lim_{x \to 1-} f(x) = s_0 = f(1)$$

Die Potenzreihe $\sum\limits_{n=0}^{\infty} s_{n+1} x^n$ konvergiert für $|x| < 1$, da $|s_n|$ beschränkt ist.
Für $|x| < 1$ gilt:

$$\begin{aligned}(1-x)\sum_{n=0}^{\infty} s_{n+1}x^n &= \sum_{n=0}^{\infty} s_{n+1}x^n - \sum_{n=0}^{\infty} s_{n+1}x^{n+1} = \sum_{n=0}^{\infty} s_{n+1}x^n - \sum_{n=0}^{\infty} s_n x^n + s_0 \\ &= s_0 + \sum_{n=0}^{\infty}(s_{n+1} - s_n)x^n = f(1) - \sum_{n=0}^{\infty} a_n x^n = f(1) - f(x)\end{aligned}$$

Sei $\varepsilon > 0$ beliebig vorgegeben. Dann existiert ein $N \in \mathbb{N}$, so daß $|s_n| \leq \frac{\varepsilon}{2}$ $\forall n > N$
Man setze:

$$\delta := \min\left(1, \frac{\varepsilon}{2cN}\right)$$

dann gilt für $1 - \delta < x < 1$:

$$\begin{aligned}|f(1) - f(x)| &\leq (1-x)\sum_{n=0}^{N-1} |s_{n+1}| \cdot x^n + (1-x)\sum_{n=N}^{\infty} |s_{n+1}| \cdot x^n \\ &\leq \underbrace{(1-x)}_{<\delta} cN + (1-x)\frac{\varepsilon}{2}\sum_{n=N}^{\infty} x^n \\ &\leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2}(1-x) \underbrace{\sum_{n=0}^{\infty} x^n}_{\substack{\text{geometrische Reihe} \\ \sum_{n=0}^{\infty} x^n = \frac{1}{1-x}}} = \varepsilon\end{aligned}$$

Daraus folgt die Behauptung. $\square$

Kapitel 3

$\mathbb{R}^n$ als topologischer Raum

§1 Grundlegende topologische Begriffe im $\mathbb{R}^n$

Die Menge $\mathbb{R}^n$ der n-Tupel reeller Zahlen ist ein Vektorraum über $\mathbb{R}$ unter der **Addition**:

$$(x_1, x_2, \ldots, x_n) + (y_1, y_2, \ldots, y_n) = (x_1 + y_1, x_2 + y_2, \ldots, x_n + y_n)$$

und der **Skalarmultiplikation**:

$$\lambda(x_1, x_2, \ldots, x_n) = (\lambda x_1, \lambda x_2, \ldots, \lambda x_n) \qquad \lambda \in \mathbb{R}$$

(siehe Lineare Algebra 1)

Definition 1:

(i) Seien $x = (x_1, x_2, \ldots, x_n), y = (y_1, y_2, \ldots, y_n) \in \mathbb{R}^n$, dann setzt man:

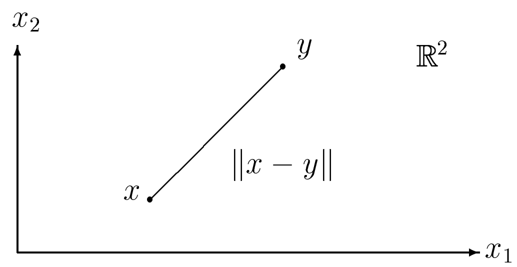
$$xy := \sum_{\nu=1}^{n} x_\nu y_\nu$$

Man nennt xy das **„Skalarprodukt“** *von x und y.*

(ii) Sei $x = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$, dann setzt man

$$\|x\| := \sqrt{xx} = \sqrt{\sum_{\nu=1}^{n} x_\nu^2}$$

Man nennt $\|x\|$ die **„EUKLIDische Norm“** *von x*



Satz 1:

(i) Die Abbildung $\mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$, $(x,y) \mapsto xy$ *ist ein* **Skalarprodukt** *auf* $\mathbb{R}^n$*, d.h.:*

a) $xy = yx \ \forall x, y \in \mathbb{R}^n$ **(Symmetrie)**

b) $x(y+z) = xy + xz \qquad \forall x, y, z \in \mathbb{R}^n$
$x(\lambda y) = \lambda(xy) \qquad \forall x, y \in \mathbb{R}^n \ \forall \lambda \in \mathbb{R}$ **(Linearität)**

c) $xx \geq 0 \qquad \forall x \in \mathbb{R}^n$
$xx = 0 \iff x = 0 \qquad \forall x \in \mathbb{R}^n$ **(positiv Definitheit)**

Beachte: *Aus a) und b) folgt auch, daß das Skalarprodukt linksseitig linear ist.*

(ii) Die Abbildung $\mathbb{R}^n \to \mathbb{R}$, $x \mapsto \|x\|$ *ist eine* **Norm** *auf* $\mathbb{R}^n$*, d.h.:*

a) $\|x\| \geq 0 \qquad \forall x \in \mathbb{R}^n$
$\|x\| = 0 \iff x = 0 \qquad \forall x \in \mathbb{R}^n$

b) $\|\lambda x\| = |\lambda| \cdot \|x\| \qquad \forall x \in \mathbb{R}^n,\ \forall \lambda \in \mathbb{R}$

c) $\|x+y\| \leq \|x\| + \|y\| \qquad \forall x, y \in \mathbb{R}^n$ **(Dreiecksungleichung)**

(iii) Es gilt:

$$|xy| \leq \|x\| \cdot \|y\| \qquad \forall x, y \in \mathbb{R}^n$$

(CAUCHY**'sche Ungleichung)**

Beweis:

(i) Klar, nach elementaren Rechenregeln für reelle Zahlen.

(ii) a) klar!

b) klar!

c) Es gilt:

$$\|x+y\|^2 = (x+y)(x+y) \underset{\text{nach (i)}}{=} xx + xy + yx + yy = \|x\|^2 + 2(xy) + \|y\|^2$$

Daher:

$$\|x+y\|^2 \leq \|x\|^2 + 2|xy| + \|y\|^2 \underset{\text{nach (iii)}}{\leq} \|x\|^2 + 2\|x\| \cdot \|y\| + \|y\|^2 = (\|x\| + \|y\|)^2$$

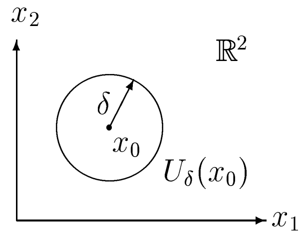
(iii) Wurde schon gezeigt. (Analysis 1, Kap. 2, §6, Satz 4)

□

Definition 2:
Sei $x_0 \in \mathbb{R}^n$*, sei* $\delta > 0$*. Dann verstehen wir unter der* δ*-Umgebung von* x_0 *die Menge:*

$$U_\delta(x_0) := \left\{ x \in \mathbb{R}^n \,\middle|\, \|x - x_0\| < \delta \right\}$$

Beispiel 13:
Im $\mathbb{R}^2$ ist $U_\delta(x_0)$ offenbar das „Innere“ einer Kreisscheibe mit dem Mittelpunkt x_0 und dem Radius δ



Definition 3:
Sei $A \subset \mathbb{R}^n$

(i) $x_0 \in \mathbb{R}^n$ *heißt* **innerer Punkt von** A, *falls es eine δ-Umgebung $U_\delta(x_0)$ von x_0 gibt, so daß*

$$U_\delta(x_0) \subset A$$

(ii) $x_0 \in \mathbb{R}^n$ *heißt* **äußerer Punkt von** A, *falls es eine δ-Umgebung $U_\delta(x_0)$ von x_0 gibt, so daß*

$$U_\delta(x_0) \subset A^c$$

($A^c := \mathbb{R}^n \setminus A$ heißt „das Komplement von A“)

(iii) $x_0 \in \mathbb{R}^n$ *heißt* **Randpunkt von** A, *falls in jeder δ-Umgebung $U_\delta(x_0)$ von x_0 sowohl ein Punkt von A als auch ein Punkt von A^c liegt.*

Definition 4:
Sei $A \subset \mathbb{R}^n$. Die Menge aller inneren Punkte von A heißt **das Innere von** A *und wird mit* $\operatorname{int} A$ *bezeichnet.*

Die Menge aller Randpunkte von A heißt **der Rand von** A *und wird oft mit ∂A bezeichnet.*

Die Menge $A \cup \partial A$ heißt **der Abschluß von** A *und wird mit $\overline{A}$ oder dA bezeichnet.*

Beispiel 14:
Betrachte $A = \mathbb{R}^n$, dann ist

$$\operatorname{int} \mathbb{R}^n = \mathbb{R}^n \qquad \partial \mathbb{R}^n = \emptyset \qquad \overline{\mathbb{R}^n} = \mathbb{R}^n$$

Beispiel 15:
Betrachte $A = U_\delta(x_0)$, es ist $\operatorname{int} U_\delta(x_0) = U_\delta(x_0)$, denn: Sei $x \in U_\delta(x_0)$. Dann ist $\|x - x_0\| < \delta$, also ist $r := \delta - \|x - x_0\| > 0$. Sei $y \in U_r(x)$. Dann gilt:

$$\|y - x_0\| = \|(y - x) + (x - x_0)\| \leq \|y - x\| + \|x - x_0\| \leq r + \|x - x_0\| = \delta$$

Also ist $y \in U_\delta(x_0) \implies U_r(x) \subset U_\delta(x_0)$

Beispiel 16:
Betrachte $A = \{x \in \mathbb{R}^n \,|\, \|x - x_0\| \leq \delta\}$, dann ist

$$\operatorname{int} A = U_\delta(x_0) \qquad\qquad \partial A = \{x \in \mathbb{R}^n \,|\, \|x - x_0\| = \delta\} \quad \text{(Sphäre)}$$

Definition 5:
Eine Teilmenge $A \subset \mathbb{R}^n$ heißt **offen**, *falls $A = \operatorname{int} A$.*
(d.h. jeder Punkt von A ist innerer Punkt von A)

Beispiele: $\emptyset$, $\mathbb{R}^n$, $U_\delta(x_0)$ sind offene Mengen.

Satz 2:

(i) Die Vereinigung **beliebig vieler** *offener Mengen ist wieder offen.*

(ii) Der Durchschnitt **endlich vieler** *offener Mengen ist wieder offen.*

Beweis:

(i) Sei $A_\mu \subset \mathbb{R}^n$ offen ($\mu \in I$, Indexmenge). Sei

$$A := \bigcup_{\mu \in I} A_\mu$$

Sei $x \in A$. Dann gibt es $\nu \in I$ mit $x \in A_\nu$. Da A_ν offen ist, existiert eine δ-Umgebung $U_\delta(x)$ mit $U_\delta(x) \subset A_\nu$. Daher folgt $U_\delta(x) \subset A$.

(ii) Seien $A_1, A_2, \ldots, A_m \subset \mathbb{R}^n$ offen. Sei

$$A := \bigcap_{\mu=1}^{m} A_\mu$$

Sei $x \in A$. Dann ist $x \in A_\mu \ \forall \mu = 1, 2, \ldots, m$. Da A_μ offen, gibt es eine δ_μ-Umgebung $U_{\delta_\mu}(x)$ mit $U_{\delta_\mu}(x) \subset A_\mu \ \forall \mu = 1, 2, \ldots, m$.

Sei $\delta := \min\{\delta_1, \delta_2, \ldots, \delta_m\}$. Dann ist $\delta > 0$ und $U_\delta(x) \subset A_\mu \ \forall \mu = 1, 2, \ldots, m$, denn $U_\delta(x) \subset U_{\delta_\mu}(x) \ \forall \mu$. Daher gilt $U_\delta(x) \subset A$, also ist A offen.

$\square$

Warnung: Der Durchschnitt beliebig vieler offener Mengen braucht im allgemeinen nicht mehr offen zu sein.

z.B. ist $\bigcup_{m \in \mathbb{N}} U_{\frac{1}{m}}(x_0) = \{x_0\}$ nicht offen!

Definition 6:
Eine Teilmenge $A \subset \mathbb{R}^n$ heißt **abgeschlossen**, *falls $A^c = \mathbb{R}^n \setminus A$ offen ist.*

Satz 3:

(i) Der Durchschnitt **beliebig vieler** *abgeschlossener Mengen ist wieder abgeschlossen.*

(ii) Die Vereinigung **endlich vieler** *abgeschlossener Mengen ist wieder abgeschlossen.*

Beweis:
Folgt direkt aus Satz 2 und den Formeln:

$$\left(\bigcap_{\mu\in I} A_\mu\right)^c = \bigcup_{\mu\in I} A_\mu^c \qquad\qquad \left(\bigcup_{\mu\in I} A_\mu\right)^c = \bigcap_{\mu\in I} A_\mu^c$$

□

§2 Folgen im $\mathbb{R}^n$

Definition 1:
Eine Folge $\left(x_m\right)_{m\in\mathbb{N}}$ mit $x_m \in \mathbb{R}^n$ heißt konvergent gegen $x_0 \in \mathbb{R}^n$, wenn zu jedem $\varepsilon > 0$ ein $N = N(\varepsilon) \in \mathbb{N}$ existiert, so daß:

$$\|x_m - x_0\| < \varepsilon \qquad \forall m > N$$

Satz 1:

(i) Der Limes einer konvergenten Folge ist eindeutig bestimmt.

(ii) Gilt

$$\lim_{m\to\infty} x_m = x_0 \qquad \text{und} \qquad \lim_{m\to\infty} y_m = y_0$$

so sind auch die Folgen $\left(x_m + y_m\right)_{m\in\mathbb{N}}$, $\left(\lambda x_m\right)_{m\in\mathbb{N}}$ ($\lambda \in \mathbb{R}$ fest), $\left(x_m y_m\right)_{m\in\mathbb{N}}$ konvergent und es gilt:

$$\lim_{m\to\infty} (x_m + y_m) = x_0 + y_0 \qquad \lim_{m\to\infty} \lambda x_m = \lambda x_0 \qquad \lim_{m\to\infty} x_m y_m = x_0 y_0$$

Beweis:
Ähnlich wie für $n = 1$ in Analysis 1 unter Benutzung der Norm-Eigenschaften von $\|x\|$.

(Anmerkung: schreibe $x_m y_m - x_0 y_0 = (x_m - x_0)y_0 + (y_m - y_0)x_m$)

Satz 1 kann auch mit dem folgenden Satz bewiesen werden.

Satz 2:
Sei $x_m = \left(x_m^{(1)}, x_m^{(2)}, \ldots, x_m^{(n)}\right)$, $x_0 = \left(x_0^{(1)}, x_0^{(2)}, \ldots, x_0^{(n)}\right)$. Dann gilt:

$$\lim_{m\to\infty} x_m = x_0 \iff \lim_{m\to\infty} x_m^{(\nu)} = x_0^{(\nu)} \qquad \forall \nu = 1, 2, \ldots, n$$

Beweis:
Für jeden Vektor $h = \left(h^{(1)}, h^{(2)}, \ldots, h^{(n)}\right) \in \mathbb{R}^n$ gilt:

$$\|h\| \geq \left|h^{(\nu)}\right| \qquad \forall \nu = 1, 2, \ldots, n \tag{$*$}$$

$$\|h\| \leq \sum_{\mu=1}^{n} \left|h^{(\mu)}\right| \tag{$**$}$$

(Beweis durch Quadrieren, denn $\|h\|^2 = \sum\limits_{\mu=1}^{n}(h^{(\mu)})^2 \leq \big(\sum\limits_{\mu=1}^{n}|h^{(\mu)}|\big)^2)$

Insbesondere gilt dies für $f = x_m - x_0$.

Sei $\lim\limits_{m\to\infty} x_m = x_0$. Sei $\varepsilon > 0$, dann existiert also ein $N \in \mathbb{N}$ mit

$$\|x_m - x_0\| < \varepsilon \qquad \forall m > N$$

Nach $(*)$ folgt daher:

$$\left|x_m^{(\nu)} - x_0^{(\nu)}\right| < \varepsilon \qquad \forall m > N$$

für jede $\nu = 1, 2, \ldots, n$. Daher gilt: $\lim\limits_{m\to\infty} x_m^{(\nu)} = x_0^{(\nu)} \qquad \forall \nu$

Umgekehrt: Sei $\lim\limits_{m\to\infty} x_m^{(\nu)} = x_0^{(\nu)}\ \forall \nu$. Sei $\varepsilon > 0$. Dann existiert für jedes ν ein $N_\nu \in \mathbb{N}$, so daß

$$\left|x_m^{(\nu)} - x_0^{(\nu)}\right| < \frac{\varepsilon}{n} \qquad \forall m > N_\nu$$

Sei $N := \max(N_1, N_2, \ldots, N_n)$. Wegen $(**)$ gilt dann:

$$\|x_m - x_0\| < \underbrace{\frac{\varepsilon}{n} + \frac{\varepsilon}{n} + \cdots + \frac{\varepsilon}{n}}_{n\text{-mal}} = \varepsilon \qquad \forall m > N$$

Daher gilt nach Definition: $\lim\limits_{m\to\infty} x_m = x_0$ $\square$

Definition 2:
Eine Folge $(x_m)_{m\in\mathbb{N}}$ mit $x_m \in \mathbb{R}^n$ heißt CAUCHY*-Folge falls zu jedem $\varepsilon > 0$ ein $N = N(\varepsilon) \in \mathbb{N}$ existiert so daß:*

$$\|x_m - x_l\| < \varepsilon \qquad \forall m, l > N$$

Satz 3:
$\mathbb{R}^n$ ist vollständig, d.h. eine Folge ist genau dann CAUCHY*-Folge, wenn sie konvergiert.*

Beweis:
Sei $\lim\limits_{m\to\infty} x_m = x_0$. Sei $\varepsilon > 0$. Dann existiert also ein $N \in \mathbb{N}$, so daß

$$\|x_m - x_0\| < \frac{\varepsilon}{2} \qquad \forall m > N$$

Daher gilt für $m, l > N$:

$$\|x_m - x_l\| \leq \|x_m - x_0\| + \|x_l - x_0\| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$

Daher ist $(x_m)_{m\in\mathbb{N}}$ eine CAUCHY-Folge.

Umgekehrt: Sei $(x_m)_{m\in\mathbb{N}}$ ein CAUCHY-Folge. Nach $(*)$ im Beweis von Satz 2 folgt, daß $(x_m^{(\nu)})_{m\in\mathbb{N}}$ für jedes $\nu = 1, 2, \ldots, n$ eine CAUCHY-Folge ist. Da $\mathbb{R}$ vollständig ist (s. Analysis 1), existiert also $\lim\limits_{m\to\infty} x_m^{(\nu)}$ für jedes ν und es gilt $\lim\limits_{m\to\infty} x_m^{(\nu)} = x_0^{(\nu)}$. Nach Satz 2 gilt daher $\lim\limits_{m\to\infty} x_m = x_0$, also ist $(x_m)_{m\in\mathbb{N}}$ konvergent. $\square$

§3 Limites und Stetigkeit

Definition 1:
Sei $x_0 \in \mathbb{R}^n$. Sei $D \subset \mathbb{R}^n$. Es gelte

$$\left\{x \in \mathbb{R}^n \,\middle|\, 0 < \|x - x_0\| < \delta_1\right\} \subset D$$

für ein $\delta_1 > 0$, d.h. D enthalte die in x_0 gelochte δ_1-Umgebung von x_0.

Sei $f : D \to \mathbb{R}^m$. Man sagt, daß der Limes von $f(x)$ für x gegen x_0 existiert wenn es $A \in \mathbb{R}^m$ gibt, so daß für jede Folge $\big(x_l\big)_{l\in\mathbb{N}}$ mit $x_l \in D$ und $\lim\limits_{l\to\infty} x_l = x_0$ gilt:

$$\lim_{l\to\infty} f(x_l) = A$$

In Zeichen:

$$\lim_{x\to x_0} f(x) = A$$

Satz 1:
Es gilt:

$$\lim_{x\to x_0} f(x) = A$$

genau dann, wenn zu jedem $\varepsilon > 0$ ein $\delta > 0$ existiert, so daß

$$\|f(x) - A\| < \varepsilon \qquad \text{für } 0 < \|x - x_0\| < \delta$$

(Beachte: links steht die Norm im $\mathbb{R}^m$, rechts die Norm im $\mathbb{R}^n$)

Beweis:
Genauso wie im Fall $n = m = 1$

Satz 2:
Es gelte $\lim\limits_{x\to x_0} f(x) = A$, $\lim\limits_{x\to x_0} g(x) = B$. Dann gilt auch:

$$\lim_{x\to x_0} \big(f(x) + g(x)\big) = A + B$$
$$\lim_{x\to x_0} \lambda f(x) = \lambda A \qquad (\lambda \in \mathbb{R})$$
$$\lim_{x\to x_0} f(x)g(x) = AB$$

Beweis:
Folgt aus §2 Satz 1 (S. 57).

Satz 3:
Es gilt $\lim\limits_{x\to x_0} f(x) = A$ genau dann, wenn

$$\lim_{x\to x_0} f^{(\nu)}(x) = A^{(\nu)} \qquad \forall \nu = 1, 2, \ldots, m$$

wobei $f = (f^{(1)}, f^{(2)}, \ldots, f^{(m)})$, $A = (A^{(1)}, A^{(2)}, \ldots, A^{(m)})$.

Beweis:
Folgt aus §2 Satz 2 (S. 57).

Definition 2:
Sei $x_0 \in \mathbb{R}^n$. *Sei* $D \subset \mathbb{R}^n$ *mit* $U_{\delta_1}(x_0) \subset D$ *für ein* $\delta_1 > 0$. *Sei* $f : D \to \mathbb{R}^m$. *Dann heißt* f **stetig in** x_0, *falls* $\lim\limits_{x \to x_0} f(x)$ *existiert und gleich* $f(x_0)$ *ist, also:*

$$\lim_{x \to x_0} f(x) = f(x_0)$$

Nach Definition und dem Bewiesenen gilt:

$$\begin{aligned}
& f(x) \text{ ist stetig in } x_0 \\
\iff\ & \text{für jede Folge } \left(x_m\right)_{m \in \mathbb{N}} \text{ mit } x_m \in D,\ \lim_{m \to \infty} x_m = x_0 \text{ gilt: } \quad \lim_{m \to \infty} f(x_m) = f(x_0) \\
\iff\ & \forall \varepsilon > 0\ \exists\ \delta > 0, \text{ so daß: } \quad \|f(x) - f(x_0)\| < \varepsilon \quad \text{für } \|x - x_0\| < \delta \\
\iff\ & \text{Jede Komponentenfunktion } f^{(\nu)} \text{ von } f,\ (\nu = 1, 2, \ldots, m) \text{ ist in } x_0^{(\nu)} \text{ stetig.}
\end{aligned}$$

Beispiel 17:
Jedes Polynom $p : \mathbb{R}^n \to \mathbb{R}$ in n Variablen

$$p(x) = \sum_{\substack{0 \le \nu_1 \le r_1 \\ \vdots \quad \vdots \quad \vdots \\ 0 \le \nu_n \le r_n}} a_{\nu_1, \nu_2, \ldots, \nu_n} \cdot x_1^{(\nu_1)} \cdot x_2^{(\nu_2)} \cdot \ldots \cdot x_n^{(\nu_n)} \qquad a_{\nu_1, \nu_2, \ldots, \nu_n} \in \mathbb{R}$$

ist stetig in jedem Punkt $x_0 \in \mathbb{R}^n$. (Grenzwertrechenregeln)

Beispiel 18:
Sei $f : \mathbb{R}^2 \to \mathbb{R}$ gegeben mit

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

Sei $(x_m, y_m) = (\frac{1}{m}, 0)$. Dann gilt

$$\lim_{m \to \infty} (x_m, y_m) = (0, 0) \qquad \text{und} \qquad f(x_m, y_m) = 1 \implies \lim_{m \to \infty} f(x_m, y_m) = 1$$

Sei $(x_m, y_m) = (0, \frac{1}{m})$. Dann gilt

$$\lim_{m \to \infty} (x_m, y_m) = (0, 0) \qquad \text{aber} \qquad f(x_m, y_m) = 0 \implies \lim_{m \to \infty} f(x_m, y_m) = 0$$

Daher ist f in $x_0 = (0, 0)$ nicht stetig.

Beispiel 19:
Wenn $\lim\limits_{x\to x_0} f(x) = A$, so ist auch

$$\lim_{\substack{t\to 0\\ t\in\mathbb{R}}} f(x_0+tv) = A \qquad \text{für jeden Vektor } v\in\mathbb{R}^n,\ v\neq 0$$

(da $t_m \xrightarrow{m\to\infty} 0$ gilt, gilt $x_0 + t_m v \xrightarrow{m\to\infty} x_0$)
Man beachte, daß die Punkte $x_0 + tv$ ($t\in\mathbb{R}$) auf der Geraden durch x_0 und x_0+v liegen.

Bemerkung:
Die Umkehrung des obigen Sachverhaltes ist im allgemeinen **nicht wahr**!

Beispiel 20:
Sei $f : \mathbb{R}^2 \setminus \{(0,0)\} \to \mathbb{R}$ gegeben durch

$$f(x,y) = \frac{(y^2-x)^2}{y^4+x^2}$$

Sei $x_0 = (0,0)$. Sei $v = t(h,k)$, $(h,k)\neq(0,0)$. Dann ist

$$\begin{aligned} f(th,tk) = \frac{(t^2k^2-th)^2}{t^4k^4+t^2h^2} &\underset{h\neq 0}{=} \frac{(tk^2-h)^2}{t^2k^4+h^2} \xrightarrow{t\to 0} \frac{h^2}{h^2} = 1\\ &\underset{h=0}{=} \frac{t^2k^4}{t^2k^4} = 1 \xrightarrow{t\to 0} 1 \end{aligned}$$

Andererseits gilt $f(t^2,t) = 0$ also $\lim\limits_{t\to 0} f(t^2,t) = 0$.

§4 Topologische Räume

Definition 1:
Sei X eine Menge. Für jedes $x\in X$ sei eine Familie $\mathcal{F}$ von Teilmengen von X, **Umgebungen** *von x genannt, gegeben mit den folgenden Eigenschaften:*

(i) Jedes $x\in X$ hat mindestens eine Umgebung.

(ii) Jede Umgebung von x enthält x.

(iii) Sind U_1, U_2 Umgebungen von x, so gibt es eine Umgebung U_3 von x mit $U_3\subset U_1\cap U_2$.

(iv) Ist U Umgebung von x und ist $y\in U$, so gibt es eine Umgebung V von y mit $V\subset U$.

Man nennt dann X (zusammen mit dem System von Umgebungen $\{\mathcal{F}_x\}_{x\in X}$) einen **topologischen Raum**.

Topologie (=wörtlich „Ortskunde") Studium topologischer Räum und stetiger Abbildungen.

Bemerkung (zur Notation):
Was hier „Umgebung" heißt, wird oft auch als „Fundamentalumgebung" bezeichnet. Dann heißt eine Teilmenge von X, die eine Fundamentalumgebung enthält, „Umgebung".

Beispiel 21:
Sei $X = \mathbb{R}^n$, die Umgebungen von $x \in X$ seien die δ-Umgebungen

$$U_\delta(x) = \{y \in \mathbb{R}^n \,\big|\, \|y - x\| < \delta\}$$

von x $(\delta > 0)$. Nachprüfen von (i) bis (iv):

(i) + (ii) klar!

(iii) $U_{\delta_1}(x) \cap U_{\delta_2}(x) = U_{\delta_3}(x) \quad$ mit $\delta_3 = \min(\delta_1, \delta_2)$

(iv) wurde in §1, Beispiel 14 gezeigt.

Beispiel 22:
Sei X eine Menge. Jedes x besitze genau eine Umgebung, nämlich X. (sogenannte „Klumpentopologie“, uninteressant)

Beispiel 23:
Sei X eine Menge. Jedes x besitze genau eine Umgebung, nämlich $\{x\}$. (sogenannte „diskrete Topologie“)

Beispiel 24:
Sei X ein normierter Vektorraum über $\mathbb{R}$ mit der Norm $\|.\|$. Für jedes $x \in X$ besitze als Umgebungen die Mengen

$$U_\delta(x) := \{x \in X \,\big|\, \|y - x\| < \delta\} \qquad (\delta > 0)$$

Beispiel 25:
Jeder **metrische Raum** (s. Übungsaufgaben) ist in natürlicher Weise ein topologischer Raum.

Beispiel 26:
Sei X ein topologischer Raum und $A \subset X$. Für jedes $x \in A$ definiere man als Umgebung von x genau die Menge $A \cap U$, wobei U alle Umgebungen von x bzgl. der Topologie von X durchläuft.

Dann ist A ein topologischer Raum mit der sogenannten **„induzierten Topologie“**.

(klar! z.B. gilt (iii), denn $(A \cap U_1) \cap (A \cap U_2) \supset A \cap U_3$ wenn $U_1 \cap U_2 \supset U_3$.)

In jedem topologischen Raum kann man **innere Punkte**, **äußere Punkte** und **Randpunkte** (bzgl. einer vorgegebener Teilmengen $A \subset X$) definieren, genauso wie im Falle $X = \mathbb{R}^n$. Genauso kann man von **offenen** bzw. **abgeschlossenen** Mengen sprechen:

z.B. heißt ein Punkt $x \in X$ **innerer Punkt von** A , wenn es eine Umgebung U von x gibt mit $U \subset A$.

Eine Teilmenge $A \subset X$ heißt **offen**, wenn jeder Punkt $x \in A$ innerer Punkt von A ist.

A heißt **abgeschlossen**, wenn $A^c = X \setminus A$ offen ist.

Genauso wie in §1 (für Fall $X = \mathbb{R}^n$) zeigt man, daß die Vereinigung **beliebig vieler** offener Mengen und der Durchschnitt **endlich vieler** offener Mengen wieder offen ist.

Beispiel 27:
X habe die diskrete Topologie (s. Bsp. 23). Dann ist jede Teilmenge $A \subset X$ offen.

Beispiel 28:
X habe die Klumpentopologie (s. Bsp. 22). Dann sind die offenen Mengen genau $\emptyset$ und X.

Beispiel 29:
Sei X ein topologischer Raum, $x_0 \in X$, dann folgt aus Axiom (iv) der Topologie, daß jede Umgebung U von x_0 offen ist.

Beispiel 30:
Sei X ein topologischer Raum, $A \subset X$ mit der induzierten Topologie (s. Bsp. 26). Dann sind die (bzgl. der induzierten Topologie von A) offenen Teilmengen von A genau die Mengen $A \cap B$, wobei $B \subset X$ offen (bzgl. der Topologie von X).

Definition 2:
Seien X und Y topologische Räume. Sei $f : X \to Y$ eine Abbildung und $x_0 \in X$, dann heißt f **stetig** *in x_0, wenn zu jeder Umgebung V von $f(x_0)$ eine Umgebung U von x_0 mit $f(U) \subset V$ existiert.*

Die Abbildung f heißt **stetig**, *wenn sie in jedem Punkt $x_0 \in X$ stetig ist.*

Beispiel 31:
Sei $X \subset \mathbb{R}^n$ offen. Dann stimmt obige Definition der Stetigkeit für $Y = \mathbb{R}^m$ mit der in §3 gegebenen ε-δ-Definition überein, denn:

$$V = \left\{ y \in \mathbb{R}^m \middle| \|y - f(x_0)\| < \varepsilon \right\} \qquad (\varepsilon > 0)$$
$$U = \left\{ x \in X \middle| \|x - x_0\| < \delta \right\} \qquad (\delta > 0)$$

$f(U) \subset V$ bedeutet dann

$$\|f(x) - f(x_0)\| < \varepsilon \qquad \text{für } \|x - x_0\| < \delta$$

Beispiel 32:
Sei X topologischer Raum mit der diskreten Topologie, Y topologischer Raum. Dann ist jede Abbildung $f : X \to Y$ stetig.

Satz 1:
Seien X und Y topologische Räume und $f : X \to Y$ eine Abbildung. Dann ist f stetig genau dann, wenn das Urbild $f^{-1}(B)$ für jede offene Menge $B \subset Y$ offen ist.

Beweis:

"$\Rightarrow$": Sei f stetig. Sei $B \subset Y$ offen. Z.z: dann ist $f^{-1}(B)$ offen. Sei $x \in f^{-1}(B)$ also ist $f(x) \in B$. Die Abbildung f ist nach Voraussetzung in x stetig, also existiert zu jeder vorgegebenen Umgebung V von $f(x)$ eine Umgebung U von x mit $f(U) \subset V$. Da B nach Voraussetzung offen und $f(x) \in B$, existiert ein solches $V \subset B$, also gilt $f(U) \subset V$, also $U \subset f^{-1}(V) \subset f^{-1}(B)$. Also ist $f^{-1}(B)$ offen.

"$\Leftarrow$": Es gelte: das Urbild jeder offenen Menge ist offen. Z.z: f ist stetig in jedem Punkt $x \in X$. Sei V Umgebung von $f(x)$. Dann ist V offen nach Definition und Axiom (iv) der Topologie. Daher ist nach Voraussetzung auch $f^{-1}(V)$ offen. Es gilt $x \in f^{-1}(V)$, denn $f(x) \in V$. Also ist x innerer Punkt von $f^{-1}(V)$ also existiert eine Umgebung U von x mit $U \subset f^{-1}(V)$, also $f(U) \subset V$. Also ist f in x stetig. Da x beliebig war, ist also f stetig.

$\square$

Korollar 1:
Sei $f : X \to \mathbb{R}$, X topologischer Raum, f stetig. Sei $c \in \mathbb{R}$. Dann sind die Mengen

$$\{x \in X \big| f(x) < c\} \qquad \{x \in X \big| f(x) > c\}$$

offen und die Mengen

$$\{x \in X \big| f(x) \leq c\} \qquad \{x \in X \big| f(x) \geq c\} \tag{$*$}$$

abgeschlossen.

Beweis:
Es gilt

$$\{x \in X \big| f(x) < c\} = f^{-1}\big((-\infty, c)\big)$$

Das Intervall $(-\infty, c)$ ist offen in $\mathbb{R}$, da f stetig ist, ist also $f^{-1}\big((-\infty, c)\big)$ offen.
(ähnlich für $\{x \in X \big| f(x) > c\}$)

Die Aussagen $(*)$ folgen durch Komplementbildung, denn eine Teilmenge $A \subset X$ ist genau dann abgeschlossen, wenn A^c offen ist.

$$\{x \in X \big| f(x) \leq c\}^c = \{x \in X \big| f(x) > c\}$$
$$\{x \in X \big| f(x) \geq c\}^c = \{x \in X \big| f(x) < c\}$$

$\square$

Satz 2:
Seien X, Y, Z topologische Räume und $f : X \to Y$, $g : Y \to Z$ Abbildungen. Es sei f in $x_0 \in X$ stetig und g in $y_0 = f(x_0)$ stetig. Dann ist auch die Komposition $g \circ f : X \to Z$ in x_0 stetig.

Beweis:
Sei V eine Umgebung von $g \circ f(x_0) = g\big(f(x_0)\big) = g(y_0)$. Da g in y_0 stetig ist, gibt es eine Umgebung U von y_0 mit $g(U) \subset V$. Da f in x_0 stetig ist, existiert eine Umgebung W von x_0 mit $f(W) \subset U$. Daher folgt $g \circ f(W) = g\big(f(W)\big) \subset g(U) \subset V$. Daher ist $g \circ f$ stetig in x_0. $\square$

Definition 3:
Seien X, Y topologische Räume und $f : X \to Y$. Dann heißt f **Homöomorphismus**, *wenn f bijektiv und sowohl f als auch f^{-1} stetig sind.*

Bemerkung:
Ist f bijektiv und f stetig, so braucht f kein Homöomorphismus zu sein (s. Übungsaufgabe).

§5 Zusammenhängende topologische Räume

Intuitiv sollte ein „zusammenhängender“ topologischer Raum „aus einem Stück“ bestehen.

Definition 1:
Sei X topologischer Raum, dann heißt X **zusammenhängend**, *falls es nicht möglich ist zu schreiben:*

$$X = A \cup B$$

wobei A und B offene Teilmengen von X sind mit

$$A \neq \emptyset \qquad B \neq \emptyset \qquad A \cap B = \emptyset$$

Eine Teilmenge $S \subset X$ heißt zusammenhängend, falls S mit der induzierten Topologie zusammenhängend ist.

Beispiel 33:
Man kann zeigen, daß die zusammenhängenden Teilmengen $S \subset R$ genau die Intervalle sind.

Satz 1:
Sei $f : X \to Y$ stetig. Ist dann X zusammenhängend, so ist auch $f(X)$ zusammenhängend.

Beweis:
Annahme: $f(X)$ ist nicht zusammenhängend.

Dann kann man nach Definition schreiben: $f(X) = P \cup Q$ mit $P, Q \subset f(X)$ offen (bzgl. der induzierten Topologie), $P \neq \emptyset$, $Q \neq \emptyset$, $P \cap Q = \emptyset$. Da f stetig, ist auch die Abbildung $g : X \to f(X)$, $g(x) := f(x)$ stetig (klar nach Definition!). Nach Satz 1, §4 (S. 63) sind daher:

$$A = g^{-1}(P) = f^{-1}(P) \qquad B = g^{-1}(Q) = f^{-1}(Q)$$

offen. Ferner gilt $A \neq \emptyset$, $B \neq \emptyset$, denn $f(X) = P \cup Q$ und $P \neq \emptyset$, $Q \neq \emptyset$. Wegen $P \cap Q = \emptyset$ gilt $A \cap B = \emptyset$. Ferner folgt aus $f(X) = P \cup Q$, daß $X = f^{-1}(P) \cup f^{-1}(Q) = A \cup B$. Daher folgt, daß X nicht zusammenhängend ist. $\lightning$Widerspruch! zur Voraussetzung. Daher ist $f(X)$ zusammenhängend. $\square$

Definition 2:
Sei X topologischer Raum, $x, y \in X$. Dann versteht man unter einem **Weg** *von x nach y (in X) eine stetige Abbildung*

$$a : [0,1] \to X \qquad \text{mit } a(0) = x,\ a(1) = y$$

Definition 3:
Ein topologischer Raum X heißt **wegzusammenhängend**, *falls je zwei Punkte $x, y \in X$ durch einen Weg verbunden werden können.*

Satz 2:
Es gilt

$$X \text{ ist wegzusammenhängend} \quad \Longrightarrow \quad X \text{ ist zusammenhängend}$$

Beweis:
Übungsaufgabe

Bemerkung:
Umkehrung gilt im allgemeinen nicht!

§6 Kompakte topologische Räume

Wir betrachten zunächst Teilmengen $A \in \mathbb{R}^n$.

Definition 1:

(i) Eine Teilmenge $A \subset \mathbb{R}^n$ heißt **beschränkt**, *falls es $C > 0$ gibt mit $\|x\| \leq C \; \forall x \in A$.*

(ii) Sei $A \subset \mathbb{R}^n$, $x_0 \in \mathbb{R}^n$. Dann heißt x_0 **Häufungspunkt von** *A, falls jede Umgebung U von x_0 unendlich viele Punkte aus A enthält.*

Lemma 1 (Cantor)**:**
Sei $A_1 \supset A_2 \supset A_3 \supset \ldots$ eine Folge von abgeschlossenen nicht-leeren Teilmengen in $\mathbb{R}^n$ mit

$$\operatorname{diam} A_m := \sup\big\{\|x-y\| \big| x, y \in A_m\big\} \qquad \operatorname{diam} A_m \xrightarrow{m\to\infty} 0$$

Dann besteht $\bigcap\limits_{m\in\mathbb{N}} A_m$ aus genau einem Punkt x_0

Beweis:
Sei $x_m \in A_m$. Sei $\varepsilon > 0$. Wegen $\operatorname{diam} A_m \xrightarrow{m\to\infty} 0$ existiert $N \in \mathbb{N}$, so daß $\operatorname{diam} A_N < \varepsilon$.

Seien $l, m \in \mathbb{N}$ mit $l, m > N$, dann ist $x_m, x_l \in A_N$, denn $A_m, A_l \subset A_N$.

Daher folgt $\|x_m - x_l\| < \varepsilon$.

Nach Definition ist daher $\big(x_m\big)_{m\in\mathbb{N}}$ eine Cauchy-Folge. Da $\mathbb{R}^n$ vollständig ist (s. §1), existiert also $x_0 = \lim\limits_{m\to\infty} x_m$.

Sei l fest. Für $m \geq l$ ist dann $x_m \in A_l$, also folgt

$$x_0 = \lim_{m\to\infty} x_m \in \overline{A_l} = A_l$$

da A_l abgeschlossen (s. Übungsaufgabe).

Daher gilt also: $x_0 \in \bigcap\limits_{l\in\mathbb{N}} A_l$.

Sei $x \in \bigcap\limits_{l\in\mathbb{N}} A_l$. Dann gilt

$$\|x - x_0\| \leq \operatorname{diam} A_l \underset{\text{(nach Voraus.)}}{\xrightarrow{l\to\infty}} 0 \qquad \Longrightarrow \qquad x = x_0$$

Dies beweist das Lemma. □

Satz 1 (Bolzano-Weierstrass)**:**
Jede unendliche, beschränkte Teilmenge $A \subset \mathbb{R}^n$ hat mindestens einen Häufungspunkt.

Beweis:
Ähnlich wie im Fall $n = 1$ durch „Intervallschachtelung“:

Da A beschränkt ist, ist A enthalten in einem n-dimensionalen, abgeschlossenen Würfel A_1. Man unterteile diesen Würfel in 2^n n-dimensionale, kongruente abgeschlossene Unterwürfel. Da A unendlich, enthält mindestens einer dieser Unterwürfel unendlich viele Elemente aus A. Dieser sei A_2. Man verfahre induktiv fort und erhält auf diese Weise eine Folge $A_1 \supset A_2 \supset A_3 \supset \ldots$ von abgeschlossenen, nicht-leeren Teilmengen von $\mathbb{R}^n$ mit:

$$\operatorname{diam} A_m \xrightarrow{m\to\infty} 0$$

Die Behauptung des Satzes folgt daher aus Lemma 1. □

Definition 2:
Eine Teilmenge A von $\mathbb{R}^n$ heißt **kompakt**, *falls jede unendliche Teilmenge $S \subset A$ mindestens einen Häufungspunkt x_0 in A besitzt.*

Satz 2:
Eine Teilmenge $A \subset \mathbb{R}^n$ ist kompakt genau dann, wenn sie abgeschlossen und beschränkt ist.

Beweis:

“$\Rightarrow$“: Sei A abgeschlossen und beschränkt. Z.z:A ist kompakt.

Sei $S \subset A$ unendliche Teilmenge von A. Da A beschränkt ist, ist auch S beschränkt. Nach Satz 1 hat also S einen Häufungspunkt x_0.

Annahme: x_0 ist nicht Element von A ($x_0 \notin A$). Dann gilt $x_0 \in A^c$. Da A abgeschlossen, ist A^c offen, also ist x_0 innerer Punkt von A^c, d.h. es existiert eine Umgebung U von x_0 mit $U \subset A^c$. Also ist $A \cap U = \emptyset$, insbesondere $S \cap U = \emptyset$ $\lightning$ Widerspruch! da x_0 ja Häufungspunkt von S ist.

Also gilt $x_0 \in A$. Demnach ist A kompakt.

“$\Leftarrow$“: Sei A kompakt.

Annahme: A ist nicht beschränkt. Dann existiert für jedes $m \in \mathbb{N}$ ein $x_m \in A$ mit $\|x_m\| \geq m$. Sei $S = \{x_1, x_2, \ldots\}$. Dann ist S unendlich. Aber S hat keinen Häufungspunkt, insbesondere keinen solchen in A $\lightning$ Widerspruch! zur Kompaktheit von A. Daraus folgt: A ist beschränkt.

Annahme: A ist nicht abgeschlossen. Dann ist also A^c nicht offen, also nicht jeder Punkt aus A^c ist innerer Punkt von A^c. Also gibt es $x_0 \in A^c$ und zu jedem $m \in \mathbb{N}$ ein $x_m \in A$ mit $\|x_m - x_0\| < \frac{1}{m}$. Sei $S = \{x_1, x_2, \ldots\}$. Dann ist S unendlich, denn $x_0 \notin A$. Andererseits hat S genau einen Häufungspunkt, nämlich x_0. Aber $x_0 \notin A$ $\lightning$ Widerspruch! Daher ist A abgeschlossen.

□

Wir geben im Folgenden eine Charakterisierung von Kompaktheit in Termen von „offenen Überdeckungen“.

Definition 3:
Sei X topologischer Raum, $A \subset X$.

(i) Eine Familie $\mathcal{F} = \{U_i\}_{i\in I}$ (I Indexmenge) von Teilmengen $U_i \subset X$ heißt **Überdeckung** *von A, falls $A \subset \bigcup_{i\in I} U_i$.*

(ii) Ist $\mathcal{F}$ Überdeckung von A und $\mathcal{F}'$ Überdeckung von A mit $\mathcal{F}' \subset \mathcal{F}$, so nennt man $\mathcal{F}'$ **Teilüberdeckung** *von $\mathcal{F}$.*

(iii) Ist $\mathcal{F} = \{U_i\}_{i\in I}$ Überdeckung von A und sind alle U_i ($i \in I$) offen, so nennt man $\mathcal{F}$ eine **offene Überdeckung** *von A.*

Satz 3 (Heine-Borel):
$A \subset \mathbb{R}^n$ ist kompakt genau dann, wenn jede offene Überdeckung von A eine endliche Teilüberdeckung (von A) enthält.

Beweis:

"$\Rightarrow$": Sei A kompakt. Sei $\mathcal{F}$ eine offene Überdeckung von A.

Annahme: Es gibt keine Teilfamilie $\mathcal{F}' \subset \mathcal{F}$, die ganz A überdeckt. Wir konstruieren induktiv eine Folge $A_1 \supset A_2 \supset \dots$ von kompakten Teilmengen des $\mathbb{R}^n$ mit

$$\operatorname{diam} A_l \xrightarrow{l\to\infty} 0$$

derart daß keine endliche Teilfamilie $\mathcal{F}' \subset \mathcal{F}$ die Menge A_l überdeckt, für jedes $l \in \mathbb{N}$. Man setze $A_1 := A$. Da A kompakt ist, ist A beschränkt nach Satz 2 (S. 67). Also ist A enthalten in einem abgeschlossenen n-dimensionalen Würfel I_1. Man unterteile I_1 in 2^n n-dimensionale, abeschlossene, kongruente Unterwürfel $I_{11}, I_{12}, \dots, I_{1m}$ mit $m = 2^n$. Man setze $A_{1\nu} := A_1 \cap I_{1\nu}$ ($\nu = 1, 2, \dots, m$). Dann ist $A_{1\nu}$ kompakt nach Satz 2 (, denn A_1 und $I_{1\nu}$ sind beschränkt und abgeschlossen, also gilt dies auch für den Durchschnitt.)

Wenn für jedes ν, $1 \le \nu \le m$ eine endliche Teilfamilie $\mathcal{F}'_\nu \subset \mathcal{F}$ existierte, die ganz $A_{1\nu}$ überdeckt, so wäre

$$\bigcup_{\nu=1}^{n} \mathcal{F}'_\nu \subset \mathcal{F}$$

eine endliche Teilfamilie von $\mathcal{F}$, die nach Konstruktion ganz A überdecken würde. Dies kann nach Annahme nicht sein. $\lightning$ Widerspruch!

Also existiert mindestens ein μ, $1 \le \mu \le m$ mit der Eigenschaft, daß keine endliche Teilfamilie $\mathcal{F}' \subset \mathcal{F}$ ganz $A_{1\nu}$ überdeckt. Man setze $A_2 := A_\mu$. Dann ist A_2 kompakt, ...

Induktiv erhält man dann die gemischte Folge $A_1 \supset A_2 \supset \dots$ Nach Konstruktion sind alle A_l nichtleer. (Die leere Menge wird von jedem $U \in \mathcal{F}$ überdeckt.) Nach dem Lemma 1 (S. 66) enthält $A_1 \cap A_2 \cap \dots$ genau einen Punkt, nämlich x_0.

Insbesondere ist $x_0 \in A_1 = A$. Da $\mathcal{F}$ eine Überdeckung von A ist, gilt $x_0 \in U$ für $U \in \mathcal{F}$. Da $\mathcal{F}$ eine offene Überdeckung ist, ist U offen. Also ist x_0 innerer Punkt von U, also existiert eine Umgebung V von x_0 mit $V \subset U$. Wegen $\operatorname{diam} A_l \xrightarrow{l\to\infty} 0$ sind daher alle A_l mit l genügend groß in V enthalten, also auch in U enthalten. $\lightning$ Widerspruch! da U offen und $A_l \subset U$ (l groß) ist also $\{U\} \subset \mathcal{F}$ eine endliche Teilüberdeckung von A_l, $\lightning$ Widerspruch! zur Konstruktion der A_l's.

Daher ist die ursprüngliche Annahme falsch, d.h. es existiert eine endliche Teilfamilie $\mathcal{F}' \subset \mathcal{F}$ die ganz A überdeckt.

"$\Leftarrow$": Es gelte: Jede offene Überdeckung von A enthält eine endliche Teilüberdeckung. Z.z: A ist kompakt.

Nach Satz 2 ist also zu zeigen, daß A abgeschlossen und beschränkt ist.

Annahme: A ist nicht beschränkt. Dann ist

$$\mathcal{F} = \{U_m(0)\}_{m\in\mathbb{N}} \qquad \big(U_m(0) := \{x \in \mathbb{R}^n \,|\, \|x\| < m\}\big)$$

eine offene Überdeckung von A, die keine endliche Teilüberdeckung von A enthält. $\lightning$ Widerspruch!

Annahme: A ist nicht abgeschlossen. Dann wäre also A_c nicht offen, also gäbe es $x_0 \in A_c$ und für jedes $m \in \mathbb{N}$ einen Punkt $x_m \in A$ mit $\|x_m - x_0\| < \frac{1}{m}$.
Sei $U_m := \{x \in \mathbb{R}^n \,|\, \|x - x_m\| > \frac{1}{m}\}$ für jedes $m \in \mathbb{N}$. Dann ist U_m offen.
Sei $\mathcal{F} = \{U_m\}_{m\in\mathbb{N}}$. Dann ist $\mathcal{F}$ eine offene Überdeckung von A (denn $\mathcal{F}$ überdeckt ja sogar $\{x \in \mathbb{R}^n \,|\, x \neq x_0\}$ und $x_0 \notin A$), die keine endliche Teilüberdeckung von A enthält (,denn nach Konstruktion existieren ja Punkte aus A beliebig nahe an x_0). $\lightning$ Widerspruch!

$\square$

Definition 4:
Sei X topologischer Raum, sei $A \subset X$. Dann heißt A **kompakt**, *wenn jede offene Überdeckung von A eine endliche Teilüberdeckung enthält.*

Definition 5:
Man nennt $A \subset X$ (X topologischer Raum) **folgenkompakt**, *wenn jede unendliche Teilmenge $S \subset A$ mindestens einen Häufungspunkt in A hat.*

Satz 4:
Es gilt:

$$\textit{Kompaktheit} \quad \Longrightarrow \quad \textit{Folgenkompaktheit}$$

Beweis:
Wird hier nicht ausgeführt.

Bemerkung:
Die Umkehrung gilt im allgemeinen nicht!

Bemerkung:
Was hier also „kompakt" bezeichnet wird, wird in der Literatur auch „quasi-kompakt" genannt.

Satz 5:
Seien X, Y topologische Räume und $f : X \to Y$ eine stetig Abbildung. Sei X kompakt. Dann ist auch $f(X)$ kompakt.

Beweis:
Sei $\mathcal{F} = \{U_i\}_{i\in I}$ (I Indexmenge) eine offene Überdeckung von $F(X)$, d.h. $f(X) \subset \bigcup\limits_{i\in I} U_i$ und jedes U_i ($i \in I$) ist offen.
Es gilt daher

$$X \subset f^{-1}\big(\bigcup_{i\in I} U_i\big) = \bigcup_{i\in I} f^{-1}U_i$$

Da f stetig und U_i offen, ist auch $f^{-1}(U_i)$ offen. Daher ist also $\{f^{-1}(U_i)\}_{i\in I}$ eine offene Überdeckung von X. Da X kompakt, $\{f^{-1}(U_i)\}_{i\in I}$ eine endliche Teilüberdeckung, d.h.

$$X \subset f^{-1}(U_{i_1}) \cup f^{-1}(U_{i_2}) \cup \cdots \cup f^{-1}(U_{i_m})$$

für gewisse, endlich viele Indizes $i_1, i_2, \ldots, i_m$. Also gilt

$$f(X) \subset U_{i_1} \cup U_{i_2} \cup \cdots \cup U_{i_m}$$

Also ist $\{U_{i_1}, U_{i_2}, \ldots, U_{i_m}\}$ eine endliche Teilüberdeckung. Daher ist $f(X)$ kompakt. $\square$

Korollar 1:
Sei X topologischer Raum, X kompakt und $f : X \to \mathbb{R}$ stetig. Dann nimmt f auf X ihr Maximum und ihr Minimum an.

Beweis:
Nach Satz 5 (S. 70) ist $f(X)$ kompakt, also abgeschlossen und beschränkt (Satz 2). Daher hat $f(X)$ ein Infimum a und ein Supremum b (Beschränktheit). Nach Konstruktion (s. Analysis 1) gibt es Folgen $(a_n)_{n\in\mathbb{N}}$ und $(b_n)_{n\in\mathbb{N}}$ mit $a_n, b_n \in X$ $\forall n \in \mathbb{N}$ und $\lim\limits_{n\to\infty} a_n = a$ und $\lim\limits_{n\to\infty} b_n = b$.

Da $f(X)$ abgeschlossen ist, gilt $a, b \in f(X)$ (wäre z.B. $a \notin f(X)$, so wäre $a \in \big(f(X)\big)^c$, aber $\big(f(X)\big)^c$ ist offen, wegen $f(X)$ abgeschlossen, also existiert eine δ-Umgebung von a, die ganz in $\big(f(X)\big)^c$ enthalten ist, $\lightning$ Widerspruch! zu $\lim\limits_{n\to\infty} a_n = a$ und $a_n \in f(X)$ $\forall n$) Dies beweißt die Behauptung. $\square$

Kapitel 4

Differenzierbarkeit reeller Funktionen mehrerer Variabler

§1 Richtungsableitungen, parielle Ableitungen

Ist $I \in \mathbb{R}$ ein offenes Intervall, $x_0 \in I$ und $f : I \to \mathbb{R}$, so definiert man die Ableitung von f in x_0 als

$$f'(x) = \lim_{h \to 0} \frac{f(x_0 + h) - f(x_0)}{h}$$

vorausgesetzt dieser Limes existiert. Ist f eine Funktion mehrerer Variabler, so ist $h \in \mathbb{R}^n$ ein Vektor und obiger Ausdruck macht keinen Sinn. (Division durch Vektoren macht keinen Sinn.)

Ziel: akzeptabler Ersatz

Definition 1:
Ein Einheitsvektor $v \in \mathbb{R}^n$ (d.h. $\|v\| = 1$) heißt **Richtung** *im $\mathbb{R}^n$.*

Ist $x_0 \in \mathbb{R}^n$ und v eine Richtung im $\mathbb{R}^n$, so heißt die Gerade durch x_0 und $x_0 + v$, d.h. die Menge

$$\{x_0 + tv | t \in \mathbb{R}\}$$

die **Gerade durch x_0 in Richtung** *v.*

Bemerkung:
Die Richtungen im $\mathbb{R}^n$ sind also die Punkte der „$(n-1)$-dimensionalen Sphäre“, welche die „n-dimensionale abgeschlossene Einheitskugel“

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

berandet. (Bsp.: Einheitskreis im $\mathbb{R}^2$)

Definition 2:
Sei $D \subset \mathbb{R}^n$ und $x_0 \in D$ ein innerer Punkt von D. Sei $f : D \to \mathbb{R}$. Sei v eine Richtung im $\mathbb{R}^n$. Dann heißt

$$\lim_{\substack{t \to 0 \\ t \in \mathbb{R}}} \frac{f(x_0 + tv) - f(x_0)}{t}$$

die **Richtungsableitung von f in x_0 in Richtung** v.

Da x_0 innerer Punkt von D ist, gibt es eine δ-Umgebung $U_\delta(x_0)$ von x_0 die ganz in D enthalten ist.

Für $|t| < \delta$ ist $x_0 + tv \in D$, denn

$$\|(x_0 + tv) - x_0\| = \|tv\| = |t| \cdot \|v\| = |t| < \delta$$

Für $|t| < \delta$ setze man

$$\varphi(t) := f(x_0 + tv)$$

Dann ist

$$\frac{f(x_0 + tv) - f(x_0)}{t} = \frac{\varphi(t) - \varphi(0)}{t}$$

damit ist die Existenz der Richtungsableitung von f in x_0 in Richtung v äquivalent zur Differentiation von $\varphi(t)$ in $t = 0$

Beispiel 34:
Sei $D = \mathbb{R}^2$ und $f : \mathbb{R}^2 \to \mathbb{R}$ gegeben durch

$$f(x, y) := \begin{cases} \frac{2xy}{x^2+y^2} & \text{für } (x, y) \neq (0, 0) \\ 1 & \text{für } (x, y) = (0, 0) \end{cases}$$

Man untersuche f auf Existenz der Richtungsableitungen in $(0, 0)$.

Sei $v = (\cos\vartheta, \sin\vartheta)$ mit $0 \leq \vartheta < 2\pi$ eine Richtung im $\mathbb{R}^2$. Dann gilt für $t \neq 0$:

$$\begin{aligned} \varphi(t) = f(tv) = f\big(t(\cos\vartheta, \sin\vartheta)\big) &= \frac{2t^2 \cos\vartheta \sin\vartheta}{t^2(\cos^2\vartheta + \sin^2\vartheta)} \\ = 2\cos\vartheta \sin\vartheta \underset{\text{(Additionstheoreme)}}{=} & \sin 2\vartheta \end{aligned}$$

Ferner ist: $\varphi(0) = f\big((0, 0)\big) = 1$

Gilt also $\sin 2\vartheta = 1$ (d.h. $\vartheta = \frac{\pi}{4} \wedge \vartheta = \frac{5}{4}\pi$), so ist $\varphi(t)$ die konstante Funktion mit Wert 1. In diesem Fall ist also φ in $t = 0$ differenzierbar und $\varphi'(0) = 0$

Ist $\sin 2\vartheta \neq 1$, so ist φ wegen $\varphi(0) = 1$ in $t = 0$ nicht stetig, also erst recht nicht differenzierbar.

Daher existiert die Richtungsableitung von f in $(0, 0)$ in Richtung $v = (\cos\vartheta, \sin\vartheta)$ genau dann, wenn $v = (\frac{1}{2}\sqrt{2}, \frac{1}{2}\sqrt{2})$ oder $v = (-\frac{1}{2}\sqrt{2}, -\frac{1}{2}\sqrt{2})$ ist.

Wir bezeichnet mit

$$e_i = (0, 0, \ldots, 0, \underset{\substack{\uparrow \\ i\text{-te Stelle}}}{1}, 0, \ldots, 0) \in \mathbb{R}^n$$

den i**-ten Einheitsvektor** im $\mathbb{R}^n$. Man beachte, daß $\|e_i\| = 1$ und daß $\{e_1, e_2, \ldots, e_n\}$ eine Basis des $\mathbb{R}^n$ ist.

Definition 3:
Sei $D \subset \mathbb{R}^n$, $x_0 \in D$ ein innerer Punkt von D und $f : D \to \mathbb{R}$. Dann nennt man die Richtungsableitung von f in x_0 in Richtung e_i $(i = 1, 2, \ldots, n)$ die i**-te partielle Ableitung** *von f in x_0, vorausgesetzt diese Richtungsableitung existiert. Man schreibt hierfür:*

$$\frac{\partial f}{\partial x_i}(x_0) \qquad \textit{oder} \qquad f_i(x_0)$$

Nach Definition gilt also

$$f_i(x_0) = \lim_{t \to 0} \frac{f\big((x_0^{(1)}, x_0^{(2)}, \ldots, x_0^{(i-1)}, x_0^{(i)} + t, x_0^{(i+1)}, \ldots, x_0^{(n)})\big) - f(x_0)}{t}$$

sofern dieser Limes existiert.

Bemerkung:
Grob gesprochen erhält man die i-te partielle Ableitung von f in x_0, indem man alle Variablen von f außer der i-ten als konstant ansieht und nach der i-ten Variablen differenziert.

Beispiel 35:
Sei $f : \mathbb{R}^3 \to \mathbb{R}$ gegeben durch

$$f(x, y, z) = x^2 + y + \cos(yz)$$

Dann ist

$$\begin{aligned}
\frac{\partial f}{\partial x}(x, y, z) &= f_1(x, y, z) = 2x \\
\frac{\partial f}{\partial y}(x, y, z) &= f_2(x, y, z) = 1 - z\sin(yz) \\
\frac{\partial f}{\partial z}(x, y, z) &= f_3(x, y, z) = -y\sin(yz)
\end{aligned}$$

Definition 4:
f heißt **partiell differenzierbar** *in x_0, wenn alle partiellen Ableitungen f_i $(i = 1, 2, \ldots, n)$ von f in x_0 existieren.*

§2 Differenzierbarkeit

Bei Funktionen einer Variablen bedeutet Differenzierbarkeit geometrisch die Existenz der Tangente. Bei reellen Funktionen mehrerer Variablen wollen wir Differenzierbarkeit so definieren, daß es geometrisch die Existenz der **Tangentialhyperebene** bedeutet.

Dies ist im allgemeinen eine viel schärfere Aussage als nur die Existenz der partiellen Ableitung.

Definition 1:
Sei $D \subset \mathbb{R}^n$, $x_0 \in D$ innerer Punkt von D und $f : D \to \mathbb{R}$. Dann heißt f **differenzierbar** *in x_0 (oder auch total differenzierbar), falls es eine lineare Abbildung $L : \mathbb{R}^n \to \mathbb{R}$ (abhängig von x_0) gibt, so daß*

$$\lim_{h\to 0} \frac{f(x_0+h) - f(x_0) - L(h)}{\|h\|} = 0$$

Erinnerung: Eine Abbildung $L : \mathbb{R}^n \to \mathbb{R}$ heißt **linear**, falls gilt

$$\begin{aligned} L(x+y) &= L(x) + L(y) \qquad && \forall x, y \in \mathbb{R}^n \\ L(\lambda x) &= \lambda L(x) && \forall x \in \mathbb{R}^n \ \forall \lambda \in \mathbb{R} \end{aligned}$$

Man beachte: Eine Abbildung $L : \mathbb{R}^n \to \mathbb{R}$ ist genau dann linear, falls es $a_1, a_2, \ldots, a_n \in \mathbb{R}$ gibt, so daß

$$L(x) = a_1 x^{(1)} + a_2 x^{(2)} + \cdots + a_n x^{(n)} \qquad \forall x = \left(x^{(1)}, x^{(2)}, \ldots, x^{(n)}\right) \in \mathbb{R}^n$$

In der Tat: Jede solche Abbildung ist linear. Umgekehrt: Sei $L : \mathbb{R}^n \to \mathbb{R}$ linear. Man schreibe

$$x = \left(x^{(1)}, x^{(2)}, \ldots, x^{(n)}\right) = e_1 x^{(1)} + e_2 x^{(2)} + \cdots + e_n x^{(n)}$$

Dann folgt wegen der Linearität von L:

$$\begin{aligned} L(x) = L\left(e_1 x^{(1)} + e_2 x^{(2)} + \cdots + e_n x^{(n)}\right) &= L\left(e_1 x^{(1)}\right) + L\left(e_2 x^{(2)}\right) + \cdots + L\left(e_n x^{(n)}\right) \\ &= \underbrace{L(e_1)}_{a_1} x^{(1)} + \underbrace{L(e_2)}_{a_2} x^{(2)} + \cdots + \underbrace{L(e_n)}_{a_n} x^{(n)} \end{aligned}$$

Satz 1:
Sei f differenzierbar in x_0. Dann gilt:

(i) Die Richtungsableitungen von f in x_0 in jede Richtung v existieren. Insbesondere ist f in x_0 partiell differenzierbar.

(ii) Es gilt

$$L(h) = \sum_{i=1}^{n} f_i(x_0) h^{(i)} \qquad \forall h = \left(h^{(1)}, h^{(2)}, \ldots, h^{(n)}\right)$$

Insbesonder ist L eindeutig (durch f und x_0) bestimmt.

Beweis:

(i) Nach Voraussetzung gilt

$$0 = \lim_{h\to 0} \frac{f(x_0+h) - f(x_0) - L(h)}{\|h\|} \tag{$*$}$$

Man wähle für $h = tv$ mit $t \in \mathbb{R}$, $t \to 0$, $t \neq 0$. Es folgt dann:

$$\begin{aligned} 0 &= \lim_{t\to 0}\left|\frac{f(x_0+tv)-f(x_0)-L(tv)}{\|tv\|}\right| \underset{(\text{wegen } \|v\|=1)}{=} \lim_{t\to 0}\left|\frac{f(x_0+tv)-f(x_0)-L(tv)}{|t|}\right| \\ &= \lim_{t\to 0}\left|\frac{f(x_0+tv)-f(x_0)-tL(v)}{t}\right| = \lim_{t\to 0}\left|\frac{f(x_0+tv)-f(x_0)}{t} - L(v)\right| \end{aligned}$$

also gilt:

$$L(v) = \lim_{t\to 0}\frac{f(x_0+tv)-f(x_0)}{t}$$

d.h. die Richtungsableitung von f in x_0 in Richtung v existiert (und ist gleich $L(v)$).

Insbesondere gilt: $L(e_i) = f_i(x_0)$

(ii) Es gilt nach (i):

$$\begin{aligned} L(h) = L\big(h^{(1)}e_1 + h^{(2)}e_2 + \dots h^{(n)}e_n\big) &= h^{(1)}L(e_1) + h^{(2)}L(e_2) + \cdots + h^{(n)}L(e_n) \\ &= h^{(1)}f_1(x_0) + h^{(2)}f_2(x_0) + \cdots + h^{(n)}f_n(x_0) \end{aligned}$$

□

Definition 2:
Sei f differenzierbar in x_0.

(i) Die durch f und x_0 nach Satz 1 (i) eindeutig bestimmte lineare Abbildung L heißt das **Differential** *von f in x_0 und wird mit $df(x_0)$ oder df_{x_0} bezeichnet.*

(ii) Der Vektor $\big(f_1(x_0), f_2(x_0), \dots, f_n(x_0)\big) \in \mathbb{R}^n$ heißt **Gradient** *(„Steigungsvektor“) von f in x_0 und wird mit $\operatorname{grad} f(x_0)$ abgekürzt.*

Man beachte:

$$df(x_0)(h) = \operatorname{grad} f(x_0) \cdot h \qquad \textit{(Skalarprodukt)}$$

Bemerkung:

(i) Ist f differenzierbar in x_0, so liefert das Differential $df(x_0)$ von f in x_0 eine lineare Approximation $df(x_0)(x - x_0)$ an $f(x) - f(x_0)$ für x nahe bei x_0.

Der Fehler dieser Approximation ist gleich

$$f(x) - f(x_0) - df(x_0)(x - x_0)$$

d.h. gleich dem Zähler von $(*)$. Dieser Fehler ist klein im Verhältnis zu $\|x - x_0\|$, wenn x nahe bei x_0 ist.

Geometrisch bedeutet Differenzierbarkeit von f in x_0, daß die Hyperebene im $\mathbb{R}^{n+1}$ gegeben durch die Gleichung

$$z = f(x_0) + \operatorname{grad} f(x_0)(x - x_0)$$

im Punkt $\big(x_0, f(x_0)\big)$ tangential an den Graphen von f ist.

„Hyperebenen“ werden hier als „affin“ verstanden, d.h. aus dem Nullpunkt translatiert. Verschiebt man zurück in den Ursprung, so erhält man also einen linearen Unterraum von $\mathbb{R}^{n+1}$ der Kodimension $(\dim V - \dim U)$.

(*Definition*: Sei V ein endlich dimensionaler Vektorraum, $U \subset V$ ein Unterraum, dann nennt man $\dim V - \dim U$ die Kodimension von U (in V).)

(ii) Formal kann man die Hyperebenengleichung

$$z = f(x_0) + \big(f_1(x_0), f_2(x_0), \ldots, f_n(x_0)\big)(x - x_0)$$

aufschreiben, schon dann, wenn f in x_0 nur partiell differenzierbar ist. Diese „berührt“ jedoch im allgemeinen dann nicht in $\big(x_0, f(x_0)\big)$, egal wie sich x an x_0 annähert.

Inbesondere impliziert im allgemeinen die Existenz der Richtungsableitungen von f in x_0 für alle Richtungen nicht die Differenzierbarkeit von f in x_0, wie Beispiel 38 zeigt.

Beispiel 36:
Ist $n = 1$, so stimmt die Definition der Differenzierbarkeit mit der aus Analysis 1 überein.

Beispiel 37:
Ist $D \subset \mathbb{R}^n$ offen und $f : D \to \mathbb{R}$ auf ganz D partiell differenzierbar und sind zusätzlich alle partiellen Ableitungen $f_1(x), f_2(x), \ldots, f_n(x)$ auf D stetig, so kann man zeigen (s. später), daß f auf ganz D differenzierbar ist.

Beispiel 38:
Sei $f : \mathbb{R}^2 \to \mathbb{R}$ gegeben durch

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

Sei $v = (\cos\vartheta, \sin\vartheta)$ $(0 \leq \vartheta < 2\pi)$ eine Richtung im $\mathbb{R}^2$. Dann gilt (mit $x_0 = (0,0)$):

$$\begin{aligned} \frac{f(x_0+tv) - f(x_0)}{t} &= \frac{f(tv)}{t} = \frac{2t^3\cos\vartheta\sin^2\vartheta}{t(t^2\cos^2\vartheta + t^4\sin^4\vartheta)} \\ &= \frac{2\cos\vartheta\sin^2\vartheta}{\cos^2\vartheta + t^2\sin^4\vartheta} \xrightarrow{t\to 0} \begin{cases} \frac{2\sin^2\vartheta}{\cos\vartheta} & \text{für } \cos\vartheta \neq 0 \\ 0 & \text{für } \cos\vartheta = 0 \end{cases} \end{aligned}$$

Also existiert die Richtungsableitung von f in $x_0 = (0,0)$ in jede Richtung v.

Aber f ist nicht stetig in $x_0 = (0,0)$, denn $f(y^2, y) = 1$, also auch $\lim\limits_{y\to 0} f(y^2,y) = 1 \neq 0 = f(x_0)$, also auch in x_0 nicht differenzierbar, wie Satz 2 zeigt.

Satz 2:
Ist f differenzierbar in x_0, so ist f in x_0 stetig.

Beweis:
Nach Voraussetzung gilt:

$$0 = \lim_{h\to 0} \frac{f(x_0+h) - f(x_0) - df(x_0)(h)}{\|h\|}$$

Nach der Definition des Limes (mit $\varepsilon = 1$) gibt es also ein $\delta_0 > 0$, so daß

$$\left| \frac{f(x_0 + h) - f(x_0) - df(x_0)(h)}{\|h\|} \right| < 1$$

für $0 < \|h\| < \delta_0$. Daher gilt

$$|f(x_0 + h) - f(x_0) - df(x_0)(h)| < \|h\| \qquad \text{für } \|h\| < \delta_0$$

Es gilt (nach der CAUCHY-SCHWARZ'schen Ungleichung):

$$df(x_0)(h) = \operatorname{grad} f(x_0) \cdot h \qquad |df(x_0)(h)| = |\operatorname{grad} f(x_0) \cdot h| \leq \|\operatorname{grad} f(x_0)\| \cdot \|h\|$$

Daher gilt für $\|h\| < \delta_0$:

$$\begin{aligned} |f(x_0 + h) - f(x_0)| &= |f(x_0 + h) - f(x_0) - df(x_0)(h) + df(x_0)(h)| \\ &\leq |f(x_0 + h) - f(x_0) - df(x_0)(h)| + |df(x_0)(h)| \\ &\leq \|h\| + \|\operatorname{grad} f(x_0)\| \cdot \|h\| = C\|h\| \qquad \text{mit } C := 1 + \|\operatorname{grad} f(x_0)\| \end{aligned}$$

Sei $\varepsilon > 0$ frei gewählt. Sei $\delta = \min(\delta_0, \frac{\varepsilon}{C})$. Für $\|h\| < \delta$ gilt dann

$$\|f(x_0 + h) - f(x_0)\| \leq \varepsilon$$

Also ist f in x_0 stetig. □

Lemma 1:
Sei $D \subset \mathbb{R}^n$, $x_0 \in D$, $f : D \to \mathbb{R}$. Seien $h \in \mathbb{R}^n$, $t_0 \in \mathbb{R}$ fest, derart daß $x_0 + t_0 h$ innerer Punkt von D ist (man beachte, daß $x_0 + t_0 h$ auf der Geraden durch x_0 und $x_0 + h$ liegt). Ist dann f in $x_0 + t_0 h$ differenzierbar, so ist auch

$$\varphi(t) := f(x_0 + th) \qquad \textit{(t nahe bei } t_0\textit{)}$$

in $t = t_0$ differenzierbar und es gilt:

$$\varphi'(t_0) = df(x_0 + t_0 h)(h) = \operatorname{grad} f(x_0 + t_0 h)h$$

Beweis:
Ist $h = 0$, so ist $\varphi(t) = f(x_0)$ konstant, also gilt $\varphi'(t_0) = 0$, also ist die Behauptung trivialerweise richtig.

Sei $h \neq 0$. Da f in $x_0 + t_0 h$ differenzierbar ist, gilt

$$0 = \lim_{\eta \to 0} \frac{f(x_0 + t_0 h + \eta) - f(x_0 + t_0 h) - df(x_0 + t_0 h)(\eta)}{\|\eta\|}$$

Man wähle speziell $\eta = th$ mit $t \in \mathbb{R}$, $t \to 0$. Dann folgt

$$\begin{aligned} 0 &= \lim_{t \to 0} \left| \frac{f(x_0 + t_0 h + th) - f(x_0 + t_0 h) - df(x_0 + t_0 h)(th)}{\|th\|} \right| \\ &= \frac{1}{\|h\|} \lim_{t \to 0} \left| \frac{f(x_0 + (t_0 + t)h) - f(x_0 + t_0 h) - t \cdot df(x_0 + t_0 h)(h)}{t} \right| \\ \Longrightarrow \qquad 0 &= \lim_{t \to 0} \left(\frac{f(x_0 + (t_0 + t)h) - f(x_0 + t_0 h)}{t} - df(x_0 + t_0 h)(h) \right) \\ &= \lim_{t \to 0} \left(\frac{\varphi(t + t_0) - \varphi(t_0)}{t} - df(x_0 + t_0 h)(h) \right) \end{aligned}$$

Dies ist gerade die Behauptung. □

Satz 3 (Mittelwertsatz in mehreren Variablen):
Sei $D \subset \mathbb{R}^n$ offen. Sei $x_0 \in D$. Sei $h \in \mathbb{R}^n$ fest, derart daß das Geradenstück zwischen x_0 und $x_0 + h$ ganz in D enthalten ist:

$$\left\{x_0 + th \middle| 0 \le t \le 1\right\} \subset D$$

Sei $f : D \to \mathbb{R}$ in jedem Punkt des Geradenstücks differenzierbar. Dann gibt es $s \in (0,1)$, so daß:

$$f(x_0 + h) - f(x_0) = df(x_0 + sh)(h)$$

Beweis:
Sei $\varphi(t) := f(x_0 + th)$ $(0 \le t \le 1)$. Da f in jedem Punkt des Geradenstücks differenzierbar ist, ist f dort auch stetig (Satz 2), also ist auch $\varphi(t)$ $(0 \le t \le 1)$ stetig, denn die Komposition stetiger Abbildungen ist wiederum stetig.

Nach dem Lemma 1 ist φ auf $(0,1)$ differenzierbar, und

$$\varphi'(t) = df(x_0 + th)(h)$$

Nach dem Mittelwertsatz für Funktionen einer Variablen folgt also die Existenz eines $s \in (0,1)$ mit

$$\varphi(1) - \varphi(0) = \varphi'(s) \cdot (1 - 0) = \varphi'(s)$$

Also gilt:

$$f(x_0 + h) - f(x_0) = df(x_0 + sh)(h)$$

□

Definition 3:
Sei $f : D \to \mathbb{R}$ und $A \subset \operatorname{int}(D)$. Ist dann f in jedem Punkt von A differenzierbar, so nennt man f auf A **differenzierbar**. *Ist D offen und $f : D \to \mathbb{R}$ auf D differenzierbar, so nennt man f eine differenzierbare Funktion (auf D).*

Definition 4:
Eine Teilmenge $K \subset \mathbb{R}^n$ heißt **konvex**, *falls die Verbindungsgerade je zweier Punkte aus K ganz in K enthalten ist, d.h.:*

$$x, y \in K \qquad \Longrightarrow \qquad tx + (1-t)y \in K \qquad \forall t \in [0,1]$$

Satz 4:
Sei $D \subset \mathbb{R}^n$ und $f : D \to \mathbb{R}$. Sei $K \subset D$ konvex derart, daß f auf K differenzierbar ist. Es gebe $C \ge 0$, so daß:

$$\|\operatorname{grad} f(x)\| \le C \qquad \forall x \in K$$

Dann gilt:

$$|f(x) - f(y)| \le C\|x - y\| \qquad \forall x, y \in K$$

Beweis:
Nach dem Mittelwertsatz (Satz 3) gibt es (zu vorgegebenen $x, y \in K$) ein $s \in (0,1)$ mit

$$f(x) - f(y) = df\big(y + s(x-y)\big)(x-y)$$

Daher folgt nach der CAUCHY-SCHWARZ'schen Ungleichung:

$$|f(x) - f(y)| = |\operatorname{grad} f\big(y + s(x-y)\big)(x-y)| \leq \|\operatorname{grad} f(\big(y + s(x-y)\big)\| \cdot \|x-y\|$$

da $y + s(x-y) \in K$ und $\|\operatorname{grad} f(x)\| \leq C \; \forall x \in K$ gilt:

$$|f(x) - f(y)| \leq C\|x-y\|$$

□

Korollar 1:
Sei $D \subset \mathbb{R}^n$ offen und zusammenhängend. Sei $f : D \to \mathbb{R}$ differenzierbar und es gelte:

$$df(x) = 0 \quad \forall x \in D \qquad \textit{bzw.} \qquad \operatorname{grad} f(x) = 0 \quad \forall x \in D$$

Dann ist f konstant.

Beweis:
Ohne Einschränkung sei $D \neq \emptyset$. Sei $x_0 \in D$. Sei

$$D_1 := \big\{x \in D \big| f(x) = f(x_0)\big\}$$

Dann ist $D_1 \neq \emptyset$. Sei $x \in D_1$. Da D offen, ist x innerer Punkt von D, also existiert eine Umgebung U von x mit $U \subset D$. Da aber U konvex ist (δ-Umgebungen sind konvexe Mengen), folgt aus $df(x) = 0 \; \forall x \in D$ und aus Satz 4 (mit C=0), daß

$$f(y) = f(x) = f(x_0) \qquad \forall y \in U$$

Daher gilt $U \in D_1$, also ist D_1 offen.

Es gilt

$$D \setminus D_1 = \big\{x \in D \big| f(x) \neq f(x_0)\big\} = \big\{x \in D \big| f(x) > f(x_0)\big\} \cup \big\{x \in D \big| f(x) < f(x_0)\big\}$$

also ist $D \setminus D_1$ als Vereinigung zweier offener Mengen wieder offen (man beachte, daß f stetig ist) nach Kap. 3, §4, Kor. 1 (S. 64).

Wäre also $D \setminus D_1$ nicht-leer, so wäre D die Vereinigung zweier nicht-leerer, offener, disjunkter Teilmengen, also wäre D nicht zusammenhängend. (beachte: da D offen ist, ist eine Teilmenge von D in der induzierten Topologie genau dann offen, wenn sie als Teilmenge von $\mathbb{R}^n$ offen ist.)
↯ Widerspruch! zur Voraussetzung D zusammenhängend.

Daraus folgt das Korollar. □

§3 Funktionen der Klasse C^q

Definition 1:
Sei $D \subset \mathbb{R}^n$ offen. Sei $f : D \to \mathbb{R}$. Ist f auf D stetig, so nennt man f von der Klasse C^0 (auf D).

Ist f auf D partiell differenzierbar und sind alle partiellen Ableitungen $f_1(x), f_2(x), \ldots, f_n(x)$ auf D stetig, so nennt man f von der Klasse C^1 (auf D).

Satz 1:
Sei $D \subset \mathbb{R}^n$ offen, $f : D \to \mathbb{R}$ von der Klasse C^1 auf D. Dann ist f auf D differenzierbar.

Beweis:
durch vollständige Induktion über n.

$n = 1$: klar!

Sei $n \geq 2$ und sei die Aussage schon für $n-1$ gezeigt. Sei $x_0 \in D$. Da D offen, gibt es $\delta_0 > 0$ mit $U_{\delta_0}(x_0) \in D$.

Wir bezeichnen mit $\hat{x} = \left(x^{(1)}, x^{(2)}, \ldots x^{(n-1)}\right) \in \mathbb{R}^{n-1}$ Koordinaten im $\mathbb{R}^{n-1}$, also gilt $x = \left(\hat{x}, x^{(n)}\right) \in \mathbb{R}^n$. Speziell setzen wir

$$\hat{x_0} = \left(x_0^{(1)}, x_0^{(2)}, \ldots, x_0^{(n-1)}\right) \qquad \text{und} \qquad x_0 = \left(\hat{x_0}, x_0^{(n)}\right)$$

Für $\hat{x} \in U_{\delta_0}(\hat{x_0})$ ist $\left(\hat{x}, x_0^{(n)}\right) \in U_{\delta_0}(x_0)$, da $\|\left(\hat{x}, x_0^{(n)}\right)\| = \|\hat{x} - \hat{x_0}\|$.
Man definiere

$$\varphi : \begin{cases} U_{\delta_0}(\hat{x_0}) \to \mathbb{R} \\ \varphi(\hat{x}) := f\left(\hat{x}, x_0^{(n)}\right) \end{cases}$$

Nach Voraussetzung existieren $f_1, f_2, \ldots, f_n$, also ist auch φ paritell differenzierbar und es gilt

$$\varphi_i(\hat{x}) = f_i\left(\hat{x}, x_0^{(n)}\right) \qquad i = 1, 2, \ldots, n-1$$

Nach Vorraussetzung sind $f_1, f_2, \ldots, f_n$ stetig, also ist auch $\varphi_1, \varphi_2, \ldots, \varphi_{n-1}$ stetig. Also ist φ von der Klasse C^1 auf $U_{\delta_0}(\hat{x_0})$. Nach Induktionsvoraussetzung ist daher φ in $\hat{x_0}$ differenzierbar.

Sei $\varepsilon > 0$. Dann existiert also ein δ_1, $0 < \delta_1 < \delta_0$ mit

$$\left| \varphi(\hat{x_0} + \hat{h}) - \varphi(\hat{x_0}) - \sum_{i=1}^{n-1} \varphi_i(\hat{x_0}) h_i \right| \leq \frac{\varepsilon}{2} \|\hat{h}\| \qquad \text{für } \|\hat{h}\| < \delta_1 \qquad (*)$$

Da f_n in x_0 stetig ist, gibt es δ_2, $0 < \delta_2 < \delta_0$, so daß

$$|f_n(y) - f_n(x_0)| \leq \frac{\varepsilon}{2} \qquad \text{für } \|y - x_0\| \leq \delta_2 \qquad (**)$$

Sei $\delta := \min(\delta_1, \delta_2)$ und im folgenden $h \in \mathbb{R}^n$ mit $\|h\| \leq \delta$

Die Funktion

$$t \mapsto f(\hat{x_0} + \hat{h}, t)$$

ist für $t \in \left[x_0^{(n)}, x_0^{(n)} + h^{(n)}\right]$ definiert (denn für solche t gilt $\left(x_0^{(n)}, x_0^{(n)} + h^{(n)}\right) \in U_\delta(x_0)$), auf $\left[x_0^{(n)}, x_0^{(n)} + h^{(n)}\right]$ stetig und auf $\left(x_0^{(n)}, x_0^{(n)} + h^{(n)}\right)$ differenzierbar nach Voraussetzung über f. Nach dem Mittelwertsatz für eine Variable existiert also $s \in (0,1)$, mit

$$\begin{aligned} f(x_0 + h) - \varphi\big(\hat{x_0} + \hat{h}\big) &= f\big(\hat{x_0} + \hat{h}, x_0^{(n)} + h^{(n)}\big) - f\big(\hat{x_0} + \hat{h}, x_0^{(n)}\big) \\ &= f_n\big(\hat{x_0} + \hat{h}, x_0^{(n)} + sh^{(n)}\big) h^{(n)} \end{aligned}$$

Sei $y = \big(\hat{x_0} + \hat{h}, x_0^{(n)} + sh^{(n)}\big)$, dann gilt:

$$\begin{aligned} \|y - x_0\| &= \big\|\big(\hat{x_0} + \hat{h}, x_0^{(n)} + sh^{(n)}\big) - x_0\big\| \\ &= \big\|\big(\hat{h}, sh^{(n)}\big)\big\| \underset{\text{(wegen } s \in (0,1))}{\leq} \|h\| \leq \delta \leq \delta_2 \end{aligned}$$

Also gilt nach $(**)$:

$$|f_n(y) - f_n(x_0)| \leq \frac{\varepsilon}{2}$$

Es gilt

$$\begin{aligned} f(x_0 + h) - f(x_0) &= \big(f(x_0 + h) - \varphi(\hat{x_0} + \hat{h})\big) + \big(\varphi(\hat{x_0} + \hat{h}) - f(x_0)\big) \\ &= \big(f(x_0 + h) - \varphi(\hat{x_0} + \hat{h})\big) + \big(\varphi(\hat{x_0} + \hat{h}) - \varphi(x_0)\big) \end{aligned}$$

Also ist

$$\begin{aligned} &\left| f(x_0 + h) - f(x_0) - \sum_{i=1}^{n} f_i(x_0) h_i \right| \\ =&\left| \Big(f(x_0 + h) - \varphi(\hat{x_0} + \hat{h}) - f_n(x_0) h_n \Big) + \Big(\varphi(\hat{x_0} + \hat{h}) - \varphi(\hat{x_0}) - \sum_{i=1}^{n-1} \varphi_i(\hat{x_0}) h_i \Big) \right| \\ \leq& |f_n(y) h_n - f_n(x_0) h_n| + \left| \varphi(\hat{x_0} + \hat{h}) - \varphi(\hat{x_0}) - \sum_{i=1}^{n-1} \varphi_i(\hat{x_0}) h_i \right| \\ \leq& \underbrace{\frac{\varepsilon}{2}|h_n|}_{\text{nach } (**)} + \underbrace{\frac{\varepsilon}{2}\|\hat{h}\|}_{\text{nach } (*)} \underset{\text{wegen } h = (\hat{h}, h^{(n)})}{\leq} \frac{\varepsilon}{2}\|h\| + \frac{\varepsilon}{2}\|h\| = \varepsilon\|h\| \end{aligned}$$

für $h \leq \delta$.
Daraus folgt:

$$\lim_{h \to 0} \frac{f(x_0 + h) - f(x_0) - \sum\limits_{i=1}^{n} f_i(x_0) h_i}{\|h\|} = 0$$

Also ist f auf D differenzierbar. $\square$

Die partiellen Ableitungen von f nennt man **partielle Ableitungen erster Ordnung** von f. Hat f_i eine j-te partielle Ableitung, so nennt man diese eine partielle Ableitung zweiter Ordnung von f, man schreibt:

$$f_{ij}(x) \qquad \text{oder} \qquad \frac{\partial^2 f}{\partial x_j \partial x_i}$$

Beispiel 39:
Sei f gegeben durch: $f(x,y) = x^2y^3$ so gilt für die partiellen Ableitungen erster und zweiter Ordnung:

$$\begin{array}{llll} f_1(x,y) = 2xy^3 & f_{11}(x,y) = 2y^3 & f_2(x,y) = 3x^2y^2 & f_{21}(x,y) = 6xy^2 \\ & f_{12}(x,y) = 6xy^2 & & f_{22}(x,y) = 6x^2y \end{array}$$

Entsprechend werden die partiellen Ableitungen von f höherer Ordnung $q = 3, 4, \ldots$ definiert, sofern sie existieren.

Man schreibt:

$$f_{i_1 i_2 \ldots i_q}(x) \qquad (1 \leq i_l \leq n,\ 1 \leq l \leq q)$$

für partielle Ableitungen von f in x zunächst in Richtung e_{i_1}, dann in Richtung $e_{i_2}, \ldots$

Definition 2:
Sei $D \subset \mathbb{R}^n$, $f : D \to \mathbb{R}$. Dann heißt f von der **Klasse** C^q **auf** *D, falls alle partiellen Ableitungen q-ter Ordnung $f_{i_1 i_2 \ldots i_q}(x)$ $(1 \leq i_l \leq n, 1 \leq l \leq q)$ $\forall x \in D$ existieren und auf D stetig sind.*

Satz 2:
Sei $D \subset \mathbb{R}^n$ offen, $f : D \to \mathbb{R}$. Ist dann f von der Klasse C^q $(q \geq 1)$, so ist f auch von der Klasse C^{q-1}.

Beweis:
Nach Voraussetzung existieren $f_{i_1 i_2 \ldots i_{q-1}\nu}(x)$ $(\nu = 1, 2, \ldots, n)$ $\forall x \in D$ und sind auf D stetig. Nach Satz 1 (S. 80) (mit f ersetzt durch $f_{i_1 i_2 \ldots i_{q-1}}$) ist daher $f_{i_1 i_2 \ldots i_{q-1}}$ auf D differenzierbar, also auch auf D stetig (§2, Satz 2 S. 76).

Daher ist f von der Klasse C^{q-1}. □

Warnung: Es kann passieren, daß f_{ij} und f_{ji} $(i \neq j)$ beide existieren, aber voneinander verschieden sind.

Beispiel 40:
Sei $f : \mathbb{R}^2 \to \mathbb{R}$ gegeben durch

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

Mann kann zeigen, daß $f_{12}(0,0)$ und $f_{21}(0,0)$ beide existieren, aber daß gilt: $f_{12}(0,0) \neq f_{21}(0,0)$

In „guten“ Situationen (z.B. f von der Klasse C^2 auf einer offenen Menge D) gilt jedoch in der Regel $f_{ij} = f_{ji}$.

Satz 3:
Sei $D \subset \mathbb{R}^n$ offen, $f : D \to \mathbb{R}$ von der Klasse C^1. Existieren dann f_{ij} und f_{ji} auf D und sind in $x_0 \in D$ stetig, so gilt

$$f_{ij}(x_0) = f_{ji}(x_0)$$

Beweis:
Wir werden dies zunächst für $n = 2$ beweisen. Wir müssen also zeigen, daß unter der Voraussetzung, daß f_{12} und f_{21} in $(x_0, y_0) \in D$ stetig sind, gilt

$$f_{12}(x_0, y_0) = f_{21}(x_0, y_0)$$

Da D offen, gibt es $\delta_0 > 0$ mit $U_{\delta_0}\big((x_0, y_0)\big) \subset D$.
Für $0 < u < \frac{1}{\sqrt{2}}\delta_0$ setze man:

$$A(u) := \frac{1}{u^2}\Big(f(x_0 + u, y_0 + u) - f(x_0, y_0 + u) - f(x_0 + u, y_0) + f(x_0, y_0)\Big)$$

(„Differenzenquotient zweiter Ordnung“)
(Man beachte, daß die Punkte $(x_0 + u, y_0 + u)$, $(x_0, y_0 + u)$, ... in $U_{\delta_0}\big((x_0, y_0)\big)$ liegen, z.B. gilt $\|(x_0 + u, y_0 + u) - (x_0, y_0)\| = \|(u, u)\| = \sqrt{u^2 + u^2} = u\sqrt{2} < \sqrt{2}\frac{\delta_0}{\sqrt{2}} = \delta_0$)
Für festes u und $x \in [x_0, x_0 + u]$ setze man

$$g(x) := f(x, x_0 + u) - f(x, x_0)$$

Dann ist g auf $[x_0, x_0 + u]$ stetig und auf $(x_0, x_0 + u)$ differenzierbar.
Dann gilt:

$$g'(x) = f_1(x, y_0 + u) - f_1(x, y_0)$$

Es gilt dann ferner

$$A(u) = \frac{1}{u^2}\big(g(x_0 + u) - g(x_0)\big)$$

Nach dem Mittelwertsatz existiert ein ξ mit $\xi = \xi(u) \in (x_0, x_0 + u)$, so daß gilt

$$A(u) := \frac{1}{u^2}g'(\xi)(u) = \frac{1}{u}\big(f_1(\xi, y_0 + u) - f_1(\xi, y_0)\big)$$

Für $y \in [y_0, y_0 + u]$ setze man:

$$h(y) := f_1(\xi, y)$$

Dann ist h auf $[y_0, y_0+u]$ stetig und auf (y_0, y_0+u) differenzierbar und es gilt nach Voraussetzung über f:

$$h'(y) = f_{12}(\xi, y)$$

Es gilt dann weiter für $A(u)$:

$$A(u) = \frac{1}{u}\big(f_1(\xi, y_0 + u) - f_1(\xi, y_0)\big) = \frac{1}{u}\big(h(y_0 + u) - h(y_0)\big)$$

Nach dem Mittelwertsatz existiert dann ein η mit $\eta = \eta(u) \in (y_0, y_0 + u)$, so daß gilt:

$$A(u) = \frac{1}{u}h'(\eta)u = h'(\eta) = f_{12}(\xi, \eta)$$

Vertauscht man die Rollen von x und y und argumentiert genauso, so findet man in der selben Weise ein $\xi^* = \xi^*(u) \in (x_0, x_0 + u)$ und $\eta^* = \eta^*(u) \in (y_0, y_0 + u)$, so daß gilt:

$$A(u) = f_{21}(\xi^*, \eta^*)$$

Sei $\varepsilon > 0$, da f_{12} und f_{21} in (x_0, y_0) stetig sind, gibt es ein δ, $0 < \delta < \delta_0$, so daß

$$\begin{array}{lll} & |f_{12}(x, y) - f_{12}(x_0, y_0)| < \varepsilon & \\ \text{und} & |f_{21}(x, y) - f_{21}(x_0, y_0)| < \varepsilon & \text{für } \|(x, y) - (x_0, y_0)\| < \delta \end{array}$$

Es gilt, falls $0 < u < \frac{\delta}{\sqrt{2}}$

$$\|(\xi, \eta) - (x_0, y_0)\| \le \sqrt{u^2 + u^2} = u\sqrt{2} < \sqrt{2}\frac{\delta}{\sqrt{2}} = \delta$$

Genauso gilt für $0 < u < \frac{\delta}{\sqrt{2}}$

$$\|(\xi^*, \eta^*) - (x_0, y_0)\| < \delta$$

Daher gilt:

$$|f_{12}(\xi, \eta) - f_{12}(x_0, y_0)| < \varepsilon \qquad \Longrightarrow \qquad |A(u) - f_{12}(x_0, y_0)| < \varepsilon$$

Genauso:

$$|f_{21}(\xi^*, \eta^*) - f_{21}(x_0, y_0)| < \varepsilon \qquad \Longrightarrow \qquad |A(u) - f_{21}(x_0, y_0)| < \varepsilon$$

Dies bedeutet nach Definition, daß

$$\lim_{u \to 0+} A(u) = f_{12}(x_0, y_0) \qquad \text{und} \qquad \lim_{u \to 0+} A(u) = f_{21}(x_0, y_0)$$

Daher gilt:

$$f_{12}(x_0, y_0) = f_{21}(x_0, y_0)$$

Sei $n \geq 2$ und $i < j$.
Sei $x_0 = \big(x_0^{(1)}, x_0^{(2)}, \ldots, x_0^{(n)}\big)$. Man setze

$$\varphi(x,y) = f\big((x_0^{(1)}, x_0^{(2)}, \ldots, x_0^{(i-1)}, \underset{\substack{\uparrow \\ i\text{-te Stelle}}}{x}, x_0^{(i+1)}, \ldots, x_0^{(j-1)}, \underset{\substack{\uparrow \\ j\text{-te Stelle}}}{y}, x_0^{(j+1)}, \ldots, x_0^{(n)})\big)$$

(für (x,y) nahe bei $(x_0^{(i)}, x_0^{(j)})$).
Dann gilt:

$$f_{ij}(x_0) = \varphi_{12}\big(x_0^{(i)}, x_0^{(j)}\big) \underset{(\text{Fall } n=2)}{=} \varphi_{21}\big(x_0^{(i)}, x_0^{(j)}\big) = f_{ji}(x_0)$$

$\square$

Satz 4 (Satz von TAYLOR):
Sei $x_0 \in \mathbb{R}^n$, sei $\delta > 0$ und $f : U_\delta(x_0) \to \mathbb{R}$ von der Klasse C^{q+1} auf $U_\delta(x_0)$, wobei $q \in \mathbb{N}_0$. Dann gilt für alle $x \in U_\delta(x_0)$ die Darstellung (mit $h = x - x_0$):

$$f(x) = f(x_0) + \sum_{i=1}^{n} f_i(x_0)h^{(i)} + \frac{1}{2!}\sum_{i,j=1}^{n} f_{ij}(x_0)h^{(i)}h^{(j)} + \ldots$$
$$\cdots + \frac{1}{q!}\sum_{i_1,i_2,\ldots,i_q=1}^{n} f_{i_1 i_2 \ldots i_q}(x_0)h^{(i_1)}h^{(i_2)}\cdots h^{(i_q)} + R_{q+1}(x)$$

wobei der Rest $R_{q+1}(x)$ durch

$$R_{q+1}(x) = \frac{1}{(q+1)!}\sum_{i_1,i_2,\ldots,i_{q+1}=1}^{n} f_{i_1 i_2 \ldots i_{q+1}}(x_0 + sh)h^{(i_1)}h^{(i_2)}\cdots h^{(i_{q+1})}$$

mit $s = s(x) \in (0,1)$ gegeben ist.

Beweis:
Sei $x \in U_\delta(x_0)$ fest gewählt und $h = x - x_0$. Ist $h = 0$, so ist nichts zu zeigen. Sei daher $h \neq 0$ und $\delta_0 := \frac{\delta}{\|h\|}$. Dann ist $\delta_0 > 1$, denn $x \in U_\delta(x_0)$.
Sei $t \in (-\delta_0, \delta_0)$. Dann gilt

$$\|(x_0 + th) - x_0\| = \|th\| = |t| \cdot \|h\| < \delta_0\|h\| = \delta$$

Also gilt $x_0 + th \in U_\delta(x_0)$.
Für $t \in (-\delta_0, \delta_0)$ setze man

$$\varphi(t) := f(x_0 + th)$$

Nach Voraussetzung ist f von der Klasse C^1 auf $U_\delta(x_0)$, also auf $U_\delta(x_0)$ differenzierbar. Nach §2 Lemma 1 (S. 77) ist daher auch φ differenzierbar und es gilt:

$$\varphi'(t) = df(x_0 + th)(h) = \sum_{i=1}^{n} f_i(x_0 + th)h^{(i)}$$

Da f von der Klasse C^{q+1} ist, kann man das Lemma 1 sukzessive auf f_i $(i = 1, 2, \ldots, n)$, dann auf f_{ij} $(i, j = 1, 2, \ldots, n)$ anwenden und erhält:

$$\begin{aligned}
\varphi''(t) &= \sum_{i=1}^{n} \left(\sum_{j=1}^{n} f_{ij}(x_0 + th) h^{(j)} \right) h^{(i)} \\
&= \sum_{i,j=1}^{n} f_{ij}(x_0 + th) h^{(i)} h^{(j)} \\
&\vdots \\
\varphi^{(q+1)}(t) &= \sum_{i_1, i_2, \ldots, i_{q+1}=1}^{n} f_{i_1 i_2 \ldots i_{q+1}}(x_0 + th) h^{(i_1)} h^{(i_2)} \cdots h^{(i_{q+1})}
\end{aligned}$$

Wendet man den Satz von TAYLOR für Funktionen einer Variablen auf $\varphi(t)$ an, indem man um $t_0 = 0$ entwickelt und in $t = 1$ auswertet (in alter Notation ist also $x_0 = 0$ und $x = 1$), so erhält man (man beachte daß $\delta_0 > 1$ also $1 \in (-\delta_0, \delta_0)$)

$$\varphi(1) = \varphi(0) + \varphi'(0) + \frac{1}{2!}\varphi''(0) + \cdots + \frac{1}{q!}\varphi^{(q)}(0) + \frac{1}{(q+1)!}\varphi^{(q+1)}(0 + s(1-0))$$

mit $s \in (0, 1)$.

Es gilt aber:

$$\varphi(1) = f(x_0 + h) = f(x) \qquad \begin{aligned} \varphi(0) &= f(x_0) \\ \varphi'(0) &= \sum_{i=1}^{n} f_i(x_0) h^{(i)} \\ &\vdots \end{aligned}$$

□

Beispiel 41:
Sei $f : \mathbb{R}^2 \to \mathbb{R}$ gegeben durch $f(x, y) = x^2 y$. Man entwickle um $(x_0, y_0) = (1, -1)$:

$$\begin{aligned}
&f_1(x,y) = 2xy && f_2(x,y) = x^2 & \\
&f_{11}(x,y) = 2y && f_{12}(x,y) = f_{21}(x,y) = 2x && f_{22}(x,y) = 0 \\
&f_{111}(x,y) = 0 && f_{112}(x,y) = f_{121}(x,y) = f_{211}(x,y) = 2 &
\end{aligned}$$

Alle übrigen partiellen Ableitungen sind Null. Insbesondere ist $R_4(x, y) = 0$ und man erhält:

$$\begin{aligned}
f(x,y) =& f(1,-1) + f_1(1,-1)(x-1) + f_2(1,-1)(y+1) \\
&+ \frac{1}{2}\big(f_{11}(1,-1)(x-1)^2 + 2f_{12}(1,-1)(x-1)(y+1)\big) \\
&+ \frac{1}{6}\big(3f_{112}(1,-1)(x-1)^2(y+1)\big) \\
=& -1 - 2(x-1) + (y+1) - (x-1)^2 + 2(x-1)(y+1) + (x-1)^2(y+1) = x^2 y
\end{aligned}$$

§4 Relative Extrema

Definition 1:
Sei $D \subset \mathbb{R}^n$ offen, $x_0 \in D$, $f : D \to \mathbb{R}$. Gibt es eine Umgebung U von x_0 mit $U \subset D$ derart, daß $f(x) \geq f(x_0)$ $\forall x \in U$ (bzw. $f(x) \leq f(x_0)$ $\forall x \in U$), so nennt man x_0 ein **relatives Minimum** *(bzw.* **relatives Maximum***) von f.*

Gilt $f(x) > f(x_0)$ $\forall x \in U$, $x \neq x_0$ (bzw. $f(x) < f(x_0)$ $\forall x \in U$, $x \neq x_0$), so heißt x_0 ein **striktes relatives Minimum** *(bzw.* **striktes relatives Maximum***) von f.*

Das Wort **„Extremum“** *steht synonym für Minimum oder Maximum.*

Definition 2:
Sei $D \subset \mathbb{R}^n$ offen, $x_0 \in D$, $f : D \to \mathbb{R}$ in x_0 differenzierbar. Gilt dann $df(x_0) = 0$ (d.h. $f_1(x_0) = 0, f_2(x_0) = 0, \ldots, f_n(x_0) = 0$), so heißt x_0 **kritischer Punkt** *von f.*

Satz 1:
Sei $D \subset \mathbb{R}^n$ offen, $x_0 \in D$, $f : D \to \mathbb{R}$ in x_0 differenzierbar. Ist dann x_0 ein relatives Extremum von f, so ist notwendigerweise x_0 ein kritischer Punkt.

Beweis:
Es habe f z.B. in x_0 ein relatives Minimum. Sei $i \in \{1, 2, \ldots, n\}$. Sei

$$\varphi(t) := f(x_0 + te_i) \qquad (t \in \mathbb{R}, \text{ nahe bei } 0)$$

Da f in x_0 ein relatives Minimum hat, hat $\varphi(t)$ in $t = 0$ ebenfalls ein relatives Minimum, d.h.:

$$\varphi(t) - \varphi(0) = f(x_0 + te_i) - f(x_0) \geq 0 \qquad (\text{für } t \text{ nahe bei } 0)$$

Es gilt also:

$$\frac{\varphi(t) - \varphi(0)}{t} \geq 0 \qquad \text{für } t \text{ nahe bei } 0,\ t > 0 \tag{$*$}$$

$$\frac{\varphi(t) - \varphi(0)}{t} \leq 0 \qquad \text{für } t \text{ nahe bei } 0,\ t < 0 \tag{$**$}$$

Es gilt aber:

$$\varphi'(0) = \lim_{t \to 0} \frac{\varphi(t) - \varphi(0)}{t} = f_i(x_0)$$

Wegen $(*)$ gilt also $\varphi'(0) \geq 0$, wegen $(**)$ gilt $\varphi'(0) \leq 0$, also folgt:

$$f_i(x_0) = \varphi'(0) = 0$$

□

Definition 3:
Sei $a_{ij} \in \mathbb{R}$ $(i, j = 1, 2, \ldots, n)$ mit $a_{ij} = a_{ji}$, d.h. die Matrix $A = (a_{ij})_{1 \le i,j \le n}$ ist symmetrisch. Dann heißt die Abbildung

$$\mathcal{Q} : \begin{cases} \mathbb{R}^n \to \mathbb{R} \\ \mathcal{Q}(h) := \sum\limits_{i,j=1}^{n} a_{ij} h^{(i)} h^{(j)} \end{cases}$$

eine **quadratische Form** *(in den Koeffizienten a_{ij})*

Bemerkung:

(i) Es gilt

$$\mathcal{Q}(h) = \sum_{i=1}^{n} a_{ii} {h^{(i)}}^2 + 2 \sum_{1 \le i < j \le n} a_{ij} h^{(i)} h^{(j)}$$

Man kann also $\mathcal{Q}$ als quadratisches Polynom auffassen. Dieses ist homogen vom Grad 2, d.h. es gilt:

$$\mathcal{Q}(\lambda h) = \lambda^2 \mathcal{Q}(h) \qquad \forall h \in \mathbb{R}^n, \forall \lambda \in \mathbb{R}$$

(ii) Die Matrix A heißt Matrix von $\mathcal{Q}$. Es gilt

$$\mathcal{Q}(h) = \left(h^{(1)}, h^{(2)}, \ldots, h^{(n)}\right) A \begin{pmatrix} h^{(1)} \\ h^{(2)} \\ \vdots \\ h^{(n)} \end{pmatrix} \qquad \forall h \in \mathbb{R}^n$$

denn: $h = h^{(1)} e_1 + h^{(2)} e_2 + \cdots + h^{(n)} e_n$, also gilt:

$$\begin{aligned} &\left(h^{(1)}, h^{(2)}, \ldots, h^{(n)}\right) A \begin{pmatrix} h^{(1)} \\ h^{(2)} \\ \vdots \\ h^{(n)} \end{pmatrix} \\ =&(h^{(1)} e_1 + h^{(2)} e_2 + \cdots + h^{(n)} e_n) \cdot A \cdot (h^{(1)} e_1^t + h^{(2)} e_2^t + \cdots + h^{(n)} e_n^t) \\ =&\sum_{i,j=1}^{n} h^{(i)} h^{(j)} e_i A e_j^t \end{aligned}$$

und man hat:

$$e_i A e_j^t = e_i A \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} \leftarrow i\text{-te Stelle} = e_i \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{nj} \end{pmatrix} \leftarrow j\text{-te Spalte von } A = a_{ij}$$

Daraus folgt:

$$\left(h^{(1)}, h^{(2)}, \ldots, h^{(n)}\right) A \begin{pmatrix} h^{(1)} \\ h^{(2)} \\ \vdots \\ h^{(n)} \end{pmatrix} = \sum_{i,j=1}^{n} a_{ij} h^{(i)} h^{(j)}$$

Definition 4:
Eine quadratische Form $\mathcal{Q}$ heißt **positiv semi-definit** *(bzw.* **negativ semi-definit***), falls $\mathcal{Q}(h) \geq 0 \; \forall h \in \mathbb{R}^n$ (bzw. $\mathcal{Q}(h) \leq 0 \; \forall h \in \mathbb{R}^n$). Sie heißt* **positiv definit** *(bzw.* **negativ definit***), falls $\mathcal{Q}(h) > 0 \; \forall h \in \mathbb{R}^n$, $h \neq 0$ (bzw. $\mathcal{Q}(h) < 0 \; \forall h \in \mathbb{R}^n$, $h \neq 0$). Ist $\mathcal{Q}$ weder positiv definit noch negativ definit, so nennt man $\mathcal{Q}$* **indefinit***. Ist $\mathcal{Q}$ positiv definit (bzw. negativ definit), so schreibt man $\mathcal{Q} > 0$ oder $A > 0$ (bzw. $\mathcal{Q} < 0$ oder $A < 0$).*

Beispiel 42:
Sei $n = 2$. Dann sieht $\mathcal{Q}$ folgendermaßen aus:

$$\begin{aligned} \mathcal{Q}(h,k) &= ah^2 + bhk + bkh + ck^2 \\ &= ah^2 + 2bhk + ck^2 \end{aligned}$$

mit $a, b, c \in \mathbb{R}$ fest.

Mit Hilfe quadratischer Ergänzung zeigt man leicht (Übungsaufgabe):

$$\begin{aligned} \mathcal{Q} > 0 \quad &\Longleftrightarrow \quad a > 0 \text{ und } \det A > 0 \\ \mathcal{Q} < 0 \quad &\Longleftrightarrow \quad a < 0 \text{ und } \det A > 0 \end{aligned}$$

mit $A = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$

Definition 5:
Sei $D \subset \mathbb{R}^n$ offen und $f : D \to \mathbb{R}$ von der Klasse C^2 auf D. Sei $x_0 \in D$. Dann heißt die quadratische Form

$$\mathcal{Q}(x_0, .) \begin{cases} \mathbb{R}^n \to \mathbb{R} \\ \mathcal{Q}(x_0, h) := \sum\limits_{i,j=1}^{n} f_{ij}(x_0) h^{(i)} h^{(j)} \end{cases}$$

HESSE-**Form** *von f in x_0.*
Die zugehörige Matrix

$$\left(f_{ij}(x_0)\right)_{1 \leq i,j \leq n}$$

heißt HESSE-**Matrix** *von f in x_0.*
Man beachte, daß $f_{ij}(x_0) = f_{ji}(x_0)$ nach §3 Satz 3 (S. 83).

Satz 2:
Sei $f : D \to \mathbb{R}$ von der Klasse C^2 auf der offenen Menge $D \subset \mathbb{R}^n$. Sei $x_0 \in D$ ein kritischer Punkt von f. Dann gilt:

(i) f hat in x_0 ein relatives Minimum

$$\Longrightarrow \qquad \mathcal{Q}(x_0,.) \geq 0 \qquad \textit{d.h. die } \text{HESSE}\textit{-Form von } f \textit{ in } x_0 \textit{ ist positiv semi-definit}$$

(ii) f hat in x_0 ein relatives Maximum

$$\Longrightarrow \qquad \mathcal{Q}(x_0,.) \leq 0 \qquad \textit{d.h. die } \text{HESSE}\textit{-Form von } f \textit{ in } x_0 \textit{ ist negativ semi-definit}$$

(iii) $\mathcal{Q}(x_0,.) > 0$, d.h. die HESSE*-Form von f in x_0 ist positiv definit.*

$$\Longrightarrow \qquad f \textit{ hat in } x_0 \textit{ ein striktes relatives Minimum}$$

(iv) $\mathcal{Q}(x_0,.) < 0$, d.h. die HESSE*-Form von f in x_0 ist negativ definit.*

$$\Longrightarrow \qquad f \textit{ hat in } x_0 \textit{ ein striktes relatives Maximum}$$

Beweis:

(i) Sei x_0 ein relatives Minimum von f, d.h. es gibt eine Umgebung U von x_0, $U \subset D$ mit $f(x) \geq f(x_0)$ $\forall x \in U$. Man wende den Satz von TAYLOR (§3 Satz 4 S. 85) mit $q = 1$ an. Es gilt also $\forall x \in U$ die Darstellung

$$\begin{aligned} f(x) &= f(x_0) + \sum_{i=1}^{n} f_i(x_0)h^{(i)} + \frac{1}{2}\sum_{i,j=1}^{n} f_{ij}(x_0 + sh)h^{(i)}h^{(j)} \quad \text{mit } h = x - x_0,\ s = s(x) \in (0,1) \\ &= f(x_0) + \frac{1}{2}\mathcal{Q}(x_0 + sh, h) \end{aligned}$$

denn x_0 ist kritischer Punkt.

Wegen $f(x) \geq f(x_0)$ folgt also

$$\mathcal{Q}(x_0 + sh, h) \geq 0 \qquad \forall x \in U \tag{$*$}$$

Annahme: Es gibt $h_0 \in \mathbb{R}^n$ mit

$$\mathcal{Q}(x_0, h_0) < 0$$

Da f von der Klasse C^2 auf D ist, ist $f_{ij}(x)$ auf D stetig, also existiert eine Umgebung U_1 von x_0, $U_1 \subset U$ mit der Eigenschaft:

$$\mathcal{Q}(y, h_0) < 0 \qquad \forall y \in U_1$$

(Anm.: Da $f_{ij}(x)$ auf D stetig ist, gilt auch $x \mapsto \mathcal{Q}(x, h_0)$ ist stetig auf D.)

In der Tat, die Menge $\{x \in D \big| \mathcal{Q}(x, h_0) < 0\}$ ist offen (Kap. 3 §4 Kor. 1 S. 64) und x_0 liegt darin.

Wähle $x = x_0 + ch_0$ mit $c \in \mathbb{R}$, $c \neq 0$, c so klein, daß $x \in U_1$. Setze $y = x_0 + sh$. Da y auf der Verbindungsgeraden zwischen x_0 und x liegt (Merke: $h = x - x_0$) und U_1 konvex ist, gilt auch $y \in U_1$, also hat man

$$\mathcal{Q}(y, h_0) < 0$$

Es folgt:

$$\mathcal{Q}(x_0 + sh, h) = \mathcal{Q}(y, ch_0) = c^2 \mathcal{Q}(y, h_0) < 0 \quad \lightning \text{Widerspruch! zu } (*)$$

Also gilt:

$$\mathcal{Q}(x_0, h_0) \geq 0 \qquad \forall h_0 \in \mathbb{R}^n$$

(ii) folgt aus (i), wenn man f durch $-f$ ersetzt.

(iii) Sei also $\mathcal{Q}(x_0, .) > 0$. Da f_{ij} nach Voraussetzung in x_0 stetig ist, kann man unter Benutzung von Übungsblatt 11, Aufg. 4, Teil ii eine Umgebung U von x_0, $U \subset D$ finden, so daß

$$\mathcal{Q}(y, .) > 0 \qquad \forall y \in U$$

Nach Taylor (wie in (i)) und da x_0 kritischer Punkt ist, folgt $\forall x \in U$

$$f(x) = f(x_0) + \frac{1}{2}\mathcal{Q}(x_0 + sh, h) \qquad \text{mit } h = x - x_0,\ s = s(x) \in (0, 1)$$

Da $x_0, x \in U$, gilt auch $y = x_0 + sh \in U$, also $\mathcal{Q}(x_0 + sh, h) > 0$, falls $h \neq 0$ (d.h. $x \neq x_0$), also gilt

$$f(x) > f(x_0) \qquad \forall x \in U, x \neq x_0$$

d.h. f hat in x_0 ein striktes relatives Minimum.

(iv) folt aus (iii), ersetze f durch $-f$.

□

Beispiel 43:
$n = 2$. Es gilt

$$\mathcal{Q}(x_0, y_0, h, k) = f_{11}(x_0, y_0)h^2 + 2f_{12}(x_0, y_0)hk + f_{22}(x_0, y_0)k^2$$

Unter den Voraussetzungen von Satz 2 ($n = 2$) erhält man das folgende hinreichende Kriterium (s. Übungsblatt 11, Aufg. 1):

Sei (x_0, y_0) ein kritischer Punkt von f, d.h. $f_1(x_0, y_0) = 0$, $f_2(x_0, y_0) = 0$. Dann gilt:

(i) $f_{11}(x_0,y_0) > 0$ und $f_{11}(x_0,y_0)f_{22}(x_0,y_0) - f_{12}^2(x_0,y_0) > 0$

$$\Longrightarrow \qquad f \text{ hat ein striktes relatives Minimum in } (x_0,y_0)$$

(ii) $f_{11}(x_0,y_0) < 0$ und $f_{11}(x_0,y_0)f_{22}(x_0,y_0) - f_{12}^2(x_0,y_0) > 0$

$$\Longrightarrow \qquad f \text{ hat ein striktes relatives Maximum in } (x_0,y_0)$$

Ist (x_0,y_0) ein kritischer Punkt von f und gilt

$$f_{11}(x_0,y_0)f_{22}(x_0,y_0) - f_{12}^2(x_0,y_0) < 0$$

so nennt man (x_0,y_0) einen **Sattelpunkt** von f.

Beispiel 44:
$n=2$. Sei $f:\mathbb{R}^2\to\mathbb{R}$ gegeben durch:

$$f(x,y) = 2y^2 - x(x-1)^2$$

So gilt:

$$f_1(x,y) = -(x-1)(3x-1) \qquad\qquad f_2(x,y) = 4y$$

Also hat f genau zwei kritische Punkte nämlich

$$(x_0,y_0) = (1,0) \qquad\qquad (x_0,y_0) = (\frac{1}{3},0)$$

Es gilt:

$$f_{11}(x,y) = -6x+4 \quad f_{22}(x,y) = 4 \quad f_{12}(x,y) = 0 \quad f_{11}(x,y)f_{22}(x,y) - f_{12}^2(x,y) = 16-24x$$

Man findet:

$(x_0,y_0) = (1,0)$ ist ein relatives Minimum $\qquad$ $(x_0,y_0) = (\frac{1}{3},0)$ ist ein Sattelpunkt

Satz 3 (Hurwitz):
Sei $A = (a_{ij})_{1\le i,j\le n}$ eine symmetrische reelle $n\times n$-Matrix. Dann sind folgende Aussagen äquivalent

(i) $A > 0$*, d.h. A ist positiv definit*

(ii) $\det(A_\nu) > 0\ \forall \nu = 1,2,\ldots,n$*, wobei* $A_\nu = (a_{ij})_{1\le i,j\le\nu}$.
(Man nennt A_ν den ν-ten **Hauptminor** *von A bzw. die ν-te* **Hauptabschnittsmatrix** *von A.)*

(iii) $A = U^tU$*, wobei U eine invertierbare obere Dreiecksmatrix der Größe n ist.*

Beweis:

(i) $\Longrightarrow$ (ii): Sei $A > 0$. Sei

$$x = \begin{pmatrix} x^{(1)} \\ x^{(2)} \\ \vdots \\ x^{(\nu)} \\ 0 \\ \vdots \\ 0 \end{pmatrix} \qquad\qquad x \neq 0$$

Dann gilt:

$$x^t A x > 0 \qquad\qquad \text{da } A > 0$$

$$x^t A x = (x^{(1)}, x^{(2)}, \ldots, x^{(\nu)}, 0, \ldots, 0) A \begin{pmatrix} x^{(1)} \\ x^{(2)} \\ \vdots \\ x^{(\nu)} \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

also gilt: $A > 0$

Es genügt also zu zeigen, daß aus $A > 0$ folgt $\det A > 0$. Nach der Linearen Algebra gilt:

$$\det A = \prod_{\mu=1}^{n} \lambda\mu$$

wobei $\lambda_1, \lambda_2, \ldots, \lambda_n$ die Eigenwerte von A sind.
Sei $\lambda \in \mathbb{C}$ ein Eigenwert von A und $x \in \mathbb{C}^{(n,1)}$ ein zugehöriger Eigenvektor, d.h. x ist eine komplexe n-Spalte, $x \neq 0$ und $Ax = \lambda x$. Es folgt:

$$\overline{x}^t A x = \lambda \overline{x}^t x$$

Man hat

$$\overline{\overline{x}^t A x} = x^t \overline{A} \overline{x} = x^t A \overline{x} \underset{(\text{da } x^t A\overline{x} \in \mathbb{C})}{=} (x^t A \overline{x})^t = \overline{x}^t A^t x \underset{(\text{A symmetrisch})}{=} \overline{x}^t A x$$

Also ist $\overline{x}^t A x$ invariant unter komplexer Konjugation, also reell. Wegen

$$\lambda = \frac{\overline{x}^t A x}{\overline{x}^t x}$$

folgt: $\lambda \in \mathbb{R}$. Daher kann man x reell wählen: $\lambda = \frac{x^t A x}{x^t x}$. Da $A > 0$ nach Voraussetzung gilt: $x^t A x > 0$ wegen $x \neq 0$. Daher folgt $\det A > 0$.

(ii) $\Longrightarrow$ (iii) Induktion über n:

$\underline{n=1:}$ Dann ist $A = a \in \mathbb{R}$, $\det A = a > 0$, setze $U = u = \sqrt{a}$.

$\underline{n \to n+1:}$ Es gelte also: $\det A_1 > 0, \det A_2 > 0, \ldots, \det A_n > 0, \det A_{n+1} > 0$

Man wende die Induktionsannahme für A_n an, also existiert eine Obere Dreiecksmatrix U_n, invertierbar, der Größe n mit $A_n = U_n^t U_n$. Man schreibe:

$$A = \begin{pmatrix} A_n & b \\ b^t & c \end{pmatrix} \qquad \text{wobei } b \text{ } n\text{-Spalte, } c \in \mathbb{R}$$

$$A = C^t \begin{pmatrix} A_n & 0 \\ 0 & c - b^t A_n^{-1} b \end{pmatrix} C \quad \text{wobei } C = \begin{pmatrix} E_n & A_n^{-1} b \\ 0 & 1 \end{pmatrix}, \; E_n \text{ die } n \times n\text{-Einheitsmatrix}$$

Wie man durch einfaches Rechnen sieht:

$$C^t \begin{pmatrix} A_n & 0 \\ 0 & c - b^t A_n^{-1} b \end{pmatrix} C = \begin{pmatrix} E_n & 0 \\ (A_n^{-1} b)^t & 1 \end{pmatrix} \begin{pmatrix} A_n & 0 \\ 0 & c - b^t A_n^{-1} b \end{pmatrix} \begin{pmatrix} E_n & A_n^{-1} b \\ 0 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} E_n & 0 \\ b^t (A_n^{-1})^t & 1 \end{pmatrix} \begin{pmatrix} A_n & b \\ 0 & c - b^t A_n^{-1} b \end{pmatrix} \underset{(A_n^{-1} \text{ symmetrisch})}{=} \begin{pmatrix} E_n & 0 \\ b^t A_n^{-1} & 1 \end{pmatrix} \begin{pmatrix} A_n & b \\ 0 & c - b^t A_n^{-1} b \end{pmatrix}$$

$$= \begin{pmatrix} A_n & b \\ b^t & b^t A_n^{-1} b + c - b^t A_n^{-1} b \end{pmatrix} = \begin{pmatrix} A_n & b \\ b^t & c \end{pmatrix} = A$$

Es folgt also insbesondere

$$\det A = \det C^t \det \begin{pmatrix} A_n & 0 \\ 0 & d \end{pmatrix} \det C \qquad \text{mit } d := c - b^t A_n^{-1} b$$

Da C eine obere Dreiecksmatrix mit 1 auf der Hauptdiagonalen ist, gilt $\det C = 1$ und somit:

$$\det A = d \det A_n \quad \Longrightarrow \quad d > 0 \quad \text{da } \det A > 0 \text{ und } \det A_n > 0$$

Es gilt also:

$$A = C^t \begin{pmatrix} U_n^t U_n & 0 \\ 0 & \sqrt{d}\sqrt{d} \end{pmatrix} C = C^t \begin{pmatrix} U_n^t & 0 \\ 0 & \sqrt{d} \end{pmatrix} \begin{pmatrix} U_n & 0 \\ 0 & \sqrt{d} \end{pmatrix} C$$

Daraus folgt:

$$A = U_{n+1}^t U_{n+1} \qquad \text{mit } U_{n+1} := \begin{pmatrix} U_n & 0 \\ 0 & \sqrt{d} \end{pmatrix} C$$

U_{n+1} ist obere invertierbare Dreiecksmatrix, denn das gleiche gilt für U_n und C.

(iii) $\Longrightarrow$ (i) Es gelte also

$$A = U^t U$$

Sei $x \in \mathbb{R}^{(n,1)}$, $x \neq 0$. Dann gilt:

$$x^t A x = x^t U^t U x = (Ux)^t (Ux) = y^t y = \sum_{i=1}^n y_i^2 > 0 \qquad \text{mit } y := Ux$$

denn $y \neq 0$, wegen $x \neq 0$ und U invertierbar. Also gilt $A > 0$. $\square$

Anmerkung: Mitbewiesen beim Beweis von Satz 3, (i) $\Longrightarrow$ (ii):

Eine reelle symmetrische Matrix hat nur reelle Eigenwerte.

Satz 4:
Sei A eine reelle symmetrische Matrix. Dann gilt:

$$A > 0 \qquad \Longrightarrow \qquad \textit{alle Eigenwerte von } A \textit{ sind positiv}$$

Beweis:
Siehe Lineare Algebra.

(Man zeigt, daß zu A eine orthogonale Matrix U existiert (d.h. $U^tU = E_n$), so daß $U^{-1}AU$ eine Diagonalmatrix ist.)

Kapitel 5

Differenzierbarkeit vektorwertiger Funktionen mehrerer Variabler

§1 Lineare Abbildungen

Erinnerung: Seien $r, n \in \mathbb{N}$. Dann sei $\mathcal{L}(\mathbb{R}^r, \mathbb{R}^n)$ die Menge der Linearen Abbildungen $L : \mathbb{R}^r \to \mathbb{R}^n$. Diese wird zu einem reellen Vektorraum unter der Addition:

$$(L_1 + L_2)(t) := L_1(t) + L_2(t) \qquad \forall t \in \mathbb{R}^r$$

und dem Skalarprodukt:

$$(\lambda L)(t) := \lambda L(t) \qquad \forall \lambda \in \mathbb{R},\ \forall t \in \mathbb{R}^r$$

Sei $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_r$ die Standartbasis des $\mathbb{R}^r$ und $e_1, e_2, \ldots, e_n$ die Standartbasis des $\mathbb{R}^n$.
Schreibt man für $L \in \mathcal{L}(\mathbb{R}^r, \mathbb{R}^n)$:

$$L(\varepsilon_j) = \sum_{i=1}^{n} a_{ij} e_i \qquad \text{mit } a_{ij} \in \mathbb{R}$$

so heißt

$$A = \big(a_{ij}\big)_{\substack{1 \le i \le n \\ 1 \le j \le r}}$$

die **Matrix von** L (bzgl. der genannten Basen).

Die Spalten von A sind gerade die Komponenten von $L(\varepsilon_j)$ $(j = 1, 2, \ldots, r)$, die Zeilen von A sind „die Komponenten“ von $L^{(i)}$ $(i = 1, 2, \ldots, n)$, wobei $L = (L^{(1)}, L^{(2)}, \ldots, L^{(n)})$.

Unter „Komponenten“ einer linearen Abbildung $\Lambda : \mathbb{R}^r \to \mathbb{R}$ verstehen wir dabei das r-Tupel $\big(b^{(1)}, b^{(2)}, \ldots, b^{(r)}\big)$, wenn $\Lambda(x) = b^{(1)} x^{(1)} + b^{(2)} x^{(2)} + \cdots + b^{(r)} x^{(r)}$.

Bemerkung:
$\mathcal{L}(\mathbb{R}^r, \mathbb{R}^n) \cong \mathbb{R}^{(n,r)}$, $L \mapsto A$, also gilt: $\dim \mathcal{L}(\mathbb{R}^r, \mathbb{R}^n) = n \cdot r$

Jede lineare Abbildung $L : \mathbb{R}^r \to \mathbb{R}^n$ ist stetig, denn schreibt man

$$L = \left(L^{(1)}, L^{(2)}, \ldots, L^{(n)}\right)$$

so ist $L^{(i)} : \mathbb{R}^r \to \mathbb{R}$ als lineare Abbildung stetig, also auch L. (s. Kap. 3 §3).
Die Abbildung

$$\mathbb{R}^n \to \mathbb{R} \qquad\qquad x \mapsto \|x\| = \sqrt{\sum_{i=1}^{n} x_i^2}$$

ist stetig, also auch die Abbildung

$$\mathbb{R}^r \to \mathbb{R} \qquad\qquad t \mapsto \|L(t)\|$$

(Komposition stetiger Abbildungen)
Da reellwertige stetige Abbildungen auf kompakten Mengen ihr Maximum annehmen (s. Kap. 3 §6 Kor. 1 S. 70) und $\{t \in \mathbb{R}^r \,|\, \|t\| \leq 1\}$ kompakt ist (abgeschlossen und beschränkt), existiert also:

$$\|L\| := \max_{\|t\| \leq 1} \|L(t)\|$$

Operatornorm genannt.

Satz 1:

(i) Die Abbildung

$$\mathcal{L}(\mathbb{R}^r, \mathbb{R}^n) \to \mathbb{R} \qquad\qquad L \mapsto \|L\| = \max_{\|t\| \leq 1} \|L(t)\|$$

ist eine Norm auf $\mathcal{L}(\mathbb{R}^r, \mathbb{R}^n)$.

(ii) Es gilt:

$$\|L(t)\| \leq \|L\| \cdot \|t\| \qquad \forall L \in \mathcal{L}(\mathbb{R}^r, \mathbb{R}^n),\ \forall t \in \mathbb{R}^r$$

(iii) Es gilt:

$$\|M \circ L\| \leq \|M\| \cdot \|L\| \qquad \forall L, M \in \mathcal{L}(\mathbb{R}^r, \mathbb{R}^n)$$

Beweis:

(i) Man muß die Eigenschaften einer Norm verifizieren. Daß $\|L\| > 0$ und $\|L\| = 0$ für $L = 0$, ist klar. Sei umgekehrt $\|L\| = 0$. Dann gilt also $\|L(t)\| = 0\ \forall t \in \mathbb{R}^r$, $\|t\| \leq 1$. Sei $t \in \mathbb{R}^r$, $t \neq 0$, dann ist

$$L(t) = L\big(\|t\| \frac{t}{\|t\|}\big) = \|t\| \cdot L\big(\frac{t}{\|t\|}\big) = 0$$

denn $\|\frac{t}{\|t\|}\| = 1$. Also ist $L = 0$.

Ferner gilt:

$$\|\lambda L\| = \max_{\|t\|\leq 1}\|(\lambda L)(t)\| = \max_{\|t\|\leq 1}|\lambda| \cdot \|L(t)\| = |\lambda| \cdot \max_{\|t\|\leq 1}\|L(t)\| = |\lambda| \cdot \|L\|$$

Dreiecksungleichung: Sei $\|t\| \leq 1$, dann gilt

$$\|(L_1 + L_2)(t)\| = \|L_1(t) + L_2(t)\| \leq \|L_1(t)\| + \|L_2(t)\| \leq \|L_1\| + \|L_2\|$$

Daraus folgt dann:

$$\|L_1 + L_2\| = \max_{\|t\|\leq 1}\|(L_1 + L_2)(t)\| \leq \|L_1\| + \|L_2\|$$

(ii) Sei ohne Einschränkung $t \neq 0$. Dann ist $\|\frac{t}{\|t\|}\| = 1$, also gilt

$$\|L(\frac{t}{\|t\|})\| \underset{\text{(nach Def.)}}{\leq} \|L\| \underset{\text{(da } L \text{ linear)}}{\Longrightarrow} \|L(t)\| \leq \|L\| \cdot \|t\|$$

(iii) Übungsaufgabe

□

§2 Differenzierbarkeit

Definition 1:
Sei $\Delta \subset \mathbb{R}^r$ und $t_0 \in \Delta$ ein innerer Punkt von Δ. Sei $g : \Delta \to \mathbb{R}^n$. Dann heißt g **differenzierbar** *in t_0, wenn es eine lineare Abbildung $L : \mathbb{R}^r \to \mathbb{R}^n$ gibt, so daß*

$$\lim_{k\to 0}\frac{g(t_0 + k) - g(t_0) - L(k)}{\|k\|} = 0$$

Bemerkung:
Ist $n = 1$, also g reellwertig, so stimmt die obige Definition mit der in Kap. 4 gegebenen überein.

Satz 1:
Sei $g = (g^{(1)}, g^{(2)}, \ldots, g^{(n)})$.

(i) Es gilt

$$g \text{ ist differenzierbar in } t_0 \quad \Longleftrightarrow \quad g^{(i)} \text{ ist differenzierbar in } t_0 \ \forall i = 1, 2, \ldots, n$$

(ii) Ist g differenzierbar in t_0, so ist die Matrix A von L gleich

$$A = \left(g_j^{(i)}(t_0)\right)_{\substack{1\leq i\leq n\\ 1\leq j\leq r}}$$

wobei $g_j^{(i)}(t_0)$ die j-te partielle Ableitung der i-ten Komponentenfunktion $g^{(i)}$ ausgewertet in t_0 ist.

Insbesondere ist L (durch t_0 und g) eindeutig bestimmt.

Beweis:

(i) Sei $L = \left(L^{(1)}, L^{(2)}, \ldots, L^{(n)}\right)$. Sei

$$\varphi(k) := \frac{g(t_0 + k) - g(t_0) - L(k)}{\|k\|} \qquad k \text{ klein}$$

Dann sind die Komponentenfunktionen von φ gleich

$$\varphi^{(i)}(k) = \frac{g^{(i)}(t_0 + k) - g^{(i)}(t_0) - L^{(i)}(k)}{\|k\|}$$

und es gilt nach Kap. 3 §3

$$\varphi(k) \xrightarrow{k \to 0} 0 \qquad \Longleftrightarrow \qquad \varphi^{(i)} \xrightarrow{k \to 0} 0 \quad \forall i = 1, 2, \ldots, n$$

Dies beweist (i).

(ii) Die Zeilen der Matrix A von L sind gerade die Komponenten von $L^{(i)}$ $(i = 1, 2, \ldots, n)$, nach dem Beweis von (i), also gerade die Vektoren $\operatorname{grad} g^{(i)}(t_0)$ $(i = 1, 2, \ldots, n)$. Dies beweist (ii). (s. Kap. 4 §2 Satz 2 S. 76)

□

Definition 2:
Sei $\Delta \subset \mathbb{R}^r$, $t_0 \in \Delta$ ein innerer Punkt von Δ, $g : \Delta \to \mathbb{R}^n$. Dann heißt

$$g_j(t_0) := \lim_{\substack{s \to 0 \\ s \in \mathbb{R}}} \frac{g(t_0 + s\varepsilon_j) - g(t_0)}{s} \qquad (j = 1, 2, \ldots, r)$$

die j-te partielle Ableitung von g in t_0, *sofern diese existiert.*

Bemerkung:
Man zeigt leicht:

$$g \text{ ist differenzierbar in } t_0 \qquad \Longrightarrow \qquad g \text{ ist partiell differenzierbar in } t_0$$

Dann sind die Spalten der Matrix A von L gerade die Vektoren $g_j(t_0)$ $(j = 1, 2, \ldots, r)$.

Definition 3:
Sei $g : \Delta \to \mathbb{R}^n$ in $t_0 \in \Delta$ differenzierbar. Dann heißt die Matrix

$$\left(g_j^{(i)}(t_0)\right)_{\substack{1 \leq i \leq n \\ 1 \leq j \leq r}}$$

von L die **Jacobi-Matrix von g in** t_0.

Die lineare-Abbildung L selbst heißt das **Differential von g in** t_0 *und wird mit $Dg(t_0)$ bezeichnet.*

Definition 4:
Im Falle $r = n$ heißt die Determinante der linearen Abbildung $L = Dg(t_0)$ die JACOBI-**Determinante von g in t_0** *und wird mit*

$$Jg(t_0)$$

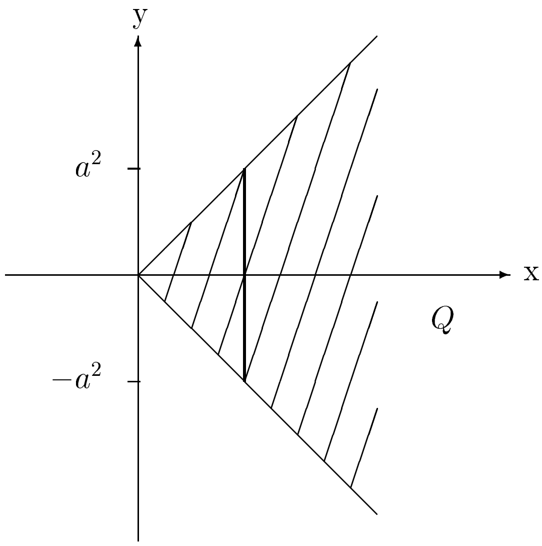
notiert.
(Manchmal auch mit: $\frac{\partial(g^{(1)},g^{(2)},\dots,g^{(n)})}{\partial(t_0^{(1)},t_0^{(2)},\dots,t_0^{(n)})}$)

Beispiel 45:
$n = r = 2$. Sei $g : \mathbb{R}^2 \to \mathbb{R}^2$ gegeben mit

$$g(s,t) = (s^2 + t^2, 2st)$$

Behauptung: Es gilt:

$$g(\mathbb{R}^2) = Q := \left\{(x,y) \in \mathbb{R}^2 \,\middle|\, -x \le y \le x\right\}$$



Schreibt man

$$x = s^2 + t^2 \qquad\qquad y = 2st$$

so gilt:

$$x + y = (s+t)^2 > 0 \qquad\qquad x - y = (s-t)^2 > 0$$

Also gilt: $g(\mathbb{R}^2) \subset Q$

Umgekehrt: Sei C der Kreis in der (s,t)-Ebene vom Radius $q > 0$ mit Mittelpunkt $(0,0)$. Dieser wird durch die Punkte

$$(a\cos\varphi, a\sin\varphi) \qquad 0 \le \varphi \le 2\pi$$

gegeben. Es gilt:

$$g(a\cos\varphi, a\sin\varphi) = \big(a^2(\cos^2\varphi + \sin^2\varphi), 2a^2\cos\varphi\sin\varphi\big) = \big(a^2, a^2\sin(2\varphi)\big)$$

Das Bild $g(C)$ ist also gleich der Verbindungsgeraden zwischen $(a^2, -a^2)$ und (a^2, a^2).

Läßt man daher a alle positiven reellen Zahlen durchlaufen und beachtet man $g(0,0) = (0,0)$, so sieht man $g(\mathbb{R}^2) = Q$.

Offenbar ist g in jedem Punkt (s,t) differenzierbar, denn dies ist wahr für die Komponentenfunktionen:

$$g^{(1)}(s,t) = s^2 + t^2 \qquad\qquad g^{(2)}(s,t) = 2st$$

Die Jacobi-Matrix von g in (s,t) ist gleich

$$\begin{pmatrix} 2s & 2t \\ 2t & 2s \end{pmatrix}$$

die Jacobi-Determinante ist gleich: $4s^2 - 4t^2$

Definition 5:
Sei $\Delta \subset \mathbb{R}^r$ offen und $g : \Delta \to \mathbb{R}^n$.

(i) Sind alle Komponentenfunktionen $g^{(i)}$ $(i = 1, 2, \ldots, n)$ von der Klasse C^q $(q \in \mathbb{N}_0)$ auf Δ, so nennt man g von der **Klasse** C^q *auf Δ.*

(ii) Ist g in jedem Punkt von Δ differenzierbar, so heißt g auf Δ **differenzierbar**.

Satz 2:

(i) Ist g von der Klasse C^1 auf Δ, so ist g auch von der Klasse C^0 auf Δ.

(ii) Ist g von der Klasse C^q auf Δ $(q \in \mathbb{N})$, so auch von der Klasse C^{q-1} auf Δ.

Beweis:
Folgt aus Satz 1 (i) (S. 99) und den entsprechenden Aussagen über reellwertige Funktionen (s. Kap. 4 §2 Satz 2 S. 76 und §3 S. 80).

Lemma 1:
Sei $\Delta \subset \mathbb{R}^r$ und $t_0 \in \Delta$ innerer Punkt von Δ. Sei $g : \Delta \to \mathbb{R}^n$ in t_0 differenzierbar. Sei $\varepsilon > 0$. Dann gibt es eine Umgebung Δ_0 von t_0, $\Delta_0 \subset \Delta$, so daß

$$\|g(t) - g(t_0)\| \le \big(\|Dg(t_0)\| + \varepsilon\big)\|t - t_0\| \qquad \forall t \in \Delta_0$$

Beweis:
Sei $L := Dg(t_0)$ und $\tilde{g}(t) := g(t) - L(t)$ $(t \in \Delta)$. Dann ist $\tilde{g}$ in t_0 differenzierbar und es gilt

$$D\tilde{g}(t_0) = Dg(t_0) - DL(t_0) = L - L = 0 \quad \text{(Nullmatrix)}$$

Nach Definition der Differzierbarkeit (angewandt auf $\tilde{g}$) existiert daher zu dem vorgegebenen $\varepsilon > 0$ eine Umgebung Δ_0 von t_0 mit

$$\|\tilde{g}(t) - \tilde{g}(t_0)\| \leq \varepsilon \|t - t_0\| \qquad \forall t \in \Delta_0$$

Es folgt:

$$\begin{aligned}\|g(t) - g(t_0)\| &= \|\big(L(t) - L(t_0)\big) + \big(\tilde{g}(t) - \tilde{g}(t_0)\big)\| \\ &\leq \|L(t) - L(t_0)\| + \|\tilde{g}(t) - \tilde{g}(t_0)\| = \|L(t - t_0)\| + \|\tilde{g}(t) - \tilde{g}(t_0)\| \\ &\leq \|L\| \cdot \|t - t_0\| + \varepsilon\|t - t_0\| = \big(\|Dg(t_0)\| + \varepsilon\big)\|t - t_0\|\end{aligned}$$

□

Lemma 2:
$\Delta \subset \mathbb{R}^r$ offen, $t_0 \in \Delta$. Sei $g : \Delta \to \mathbb{R}^n$ von der Klasse C^1 auf Δ. Sei $\varepsilon > 0$. Dann gibt es eine Umgebung Δ_0 von t_0, $\Delta_0 \subset \Delta$ mit

$$\|g(s) - g(t)\| \leq \big(\|Dg(t_0)\| + \varepsilon\big)\|s - t\| \qquad \forall s, t \in \Delta_0$$

Beweis:
Seien die Notationen wie im Beweis von Lemma 1, also $\tilde{g} = g - L$, $L = Dg(t_0)$.
Wegen $D\tilde{g}(t_0) = 0$ (Nullmatrix), gilt dann

$$\operatorname{grad} \tilde{g}^{(i)}(t_0) = 0 \qquad \forall i = 1, 2, \ldots, n$$

Nach Voraussetzung sind alle partiellen Ableitungen von g stetig, also sind auch die Funktionen

$$t \mapsto \operatorname{grad} \tilde{g}^{(i)}(t) \qquad i = 1, 2, \ldots, n$$

in t_0 stetig. Daher existiert zu dem vorgegebenen $\varepsilon > 0$ eine Umgebung Δ_0 von t_0, $\Delta_0 \subset \Delta$ mit:

$$\|\operatorname{grad} \tilde{g}^{(i)}(t)\| < \frac{\varepsilon}{n} \qquad \forall i = 1, 2, \ldots, n \ \forall t \in \Delta_0$$

Die Umgebung Δ_0 von t_0 ist konvex. Nach Kap. 4 §2 Satz 4 (S. 78) gilt daher:

$$|\tilde{g}^{(i)}(s) - \tilde{g}^{(i)}(t)| \leq \frac{\varepsilon}{n}\|s - t\| \tag{$*$}$$

Es gilt:

$$\begin{aligned}\|\tilde{g}(s) - \tilde{g}(t)\| &= \|\big(\tilde{g}^{(1)}(s) - \tilde{g}^{(1)}(t), \tilde{g}^{(2)}(s) - \tilde{g}^{(2)}(t), \ldots \tilde{g}^{(n)}(s) - \tilde{g}^{(n)}(t)\big)\| \\ &\leq \sum_{i=1}^{n} |\tilde{g}^{(i)}(s) - \tilde{g}^{(i)}(t)|\end{aligned}$$

Nach $(*)$ gilt also:

$$\|\tilde{g}(s) - \tilde{g}(t)\| \leq \varepsilon\|s - t\|$$

Man fährt dann fort genauso, wie im Beweis von Lemma 1 und erhält die Behauptung. □

Lemma 3:
Sei $r = n$. Sei $\Delta \subset \mathbb{R}^n$ offen und $g : \Delta \to \mathbb{R}^n$ von der Klasse C^1 auf Δ. Sei $t_0 \in \Delta$. Sei das Differential $Dg(t_0)$ invertierbar. Sei $\varepsilon > 0$. Dann gibt es eine Umgebung Δ_0 von t_0, $\Delta_0 \subset \Delta$, so daß:

$$\|g(s) - g(t)\| \geq (c - \varepsilon)\|s - t\| \qquad \forall s, t \in \Delta_0$$

Wobei:

$$c := \frac{1}{\|Dg^{-1}(t_0)\|} > 0$$

Bemerkung:
Es gilt:

$$Dg(t_0) \text{ invertierbar} \qquad \Longleftrightarrow \qquad Jg(t_0) = \det Dg(t_0) \neq 0$$

Beweis:
Sei $L := Dg(t_0) \in \mathrm{GL}(\mathbb{R}^n)$ bzw. $\in \mathrm{End}(\mathbb{R}^n)$. Dann gilt für $s, t \in \mathbb{R}^n$:

$$\|s - t\| = \|(L^{-1} \circ L)(s - t)\| = \|L^{-1}\big(L(s) - L(t)\big)\| \leq \|L^{-1}\| \cdot \|L(s) - L(t)\|$$

Daraus folgt:

$$\|L(s) - L(t)\| \geq \frac{1}{\|L^{-1}\|}\|s - t\| = c\|s - t\| \tag{$*$}$$

Setze $\tilde{g}(t) := g(t) - L(t)$ $(t \in \Delta)$, dann gilt (s. Beweis Lemma 1) wieder $D\tilde{g}(t_0) = 0$.

Nach Lemma 2 angewandt auf $\tilde{g}$ gilt: zu jedem $\varepsilon > 0$ existiert eine Umgebung Δ_0 von t_0, $\Delta_0 \subset \Delta$, so daß gilt:

$$\|\tilde{g}(s) - \tilde{g}(t)\| \leq \big(\|D\tilde{g}(t_0)\| + \varepsilon\big)\|s - t\| = \varepsilon\|s - t\| \qquad \forall s, t \in \Delta_0$$

Dann gilt für $s, t \in \Delta_0$:

$$\begin{aligned}\|L(s) - L(t)\| &= \|\big(g(s) - g(t)\big) - \big(\tilde{g}(s) - \tilde{g}(t)\big)\| \\ &\leq \|g(s) - g(t)\| + \|\tilde{g}(s) - \tilde{g}(t)\| \\ &\leq \|g(s) - g(t)\| + \varepsilon\|s - t\|\end{aligned}$$

Aus $(*)$ folgt dann:

$$c\|s - t\| \leq \|L(s) - L(t)\| \leq \|g(s) - g(t)\| + \varepsilon\|s - t\|$$

Daraus folgt dann:

$$\|g(s) - g(t)\| \geq (c - \varepsilon)\|s - t\|$$

□

Bemerkung:
In dieser Situation ist $g\big|_{\Delta_0}$ injektiv.

Beweis:
Seien $s, t \in \Delta_0$ mit $g(s) = g(t)$, dann gilt:

$$0 = \|0\| = \|g(s) - g(t)\| \geq (c - \varepsilon)\|s - t\| \quad \Longrightarrow \quad 0 \leq \|s - t\| \leq 0 \quad \Longrightarrow \quad s = t$$

□

§3 Komposition von Funktionen und Kettenregel

Satz 1 (Kettenregel):
Sei $\Delta \subset \mathbb{R}^r$ offen, $g : \Delta \to \mathbb{R}^n$ eine Abbildung. $D \subset \mathbb{R}^n$ offen, $f : D \to \mathbb{R}^p$ eine Abbildung mit $g(\Delta) \subset D$. Sind g in $t_0 \in \Delta$ differenzierbar und f in $x_0 = g(t_0) \in D$ differenzierbar, dann gilt:

(i) Die Komposition $F := f \circ g$ ist in t_0 differenzierbar.

(ii) $DF(t_0) = Df\big(g(t_0)\big) \circ Dg(t_0)$

$$\begin{tikzcd}
\Delta \arrow[r, "g"] \arrow[dr, "F = f \circ g"'] & D \subset \mathbb{R}^n \arrow[d, "f"] \\
 & \mathbb{R}^p
\end{tikzcd}
\qquad
\begin{tikzcd}
\mathbb{R}^r \arrow[r, "Dg(t_0)"] \arrow[dr, "{DF(t_0) = Df(x_0) \circ Dg(t_0)}"'] & \mathbb{R}^n \arrow[d, "Df(x_0)"] \\
 & \mathbb{R}^p
\end{tikzcd}$$

Beweis:
Nach der Definition der Differenzierbarkeit (S. 99) gilt:

$$g(t_0 + h) = g(t_0) + Dg(t_0)(h) + \|h\|\varphi(h) \qquad \forall h \in U_\varepsilon(0) \in \mathbb{R}^r, \text{ mit } \lim_{h \to 0} \varphi(h) = 0$$

$$f(x_0 + k) = f(x_0) + Df(x_0)(k) + \|k\|\psi(h) \qquad \forall k \in U_\varepsilon(0) \in \mathbb{R}^n, \text{ mit } \lim_{k \to 0} \psi(k) = 0$$

Zu zeigen:

$$\lim_{h \to 0} \frac{F(t_0 + h) - F(t_0) - DF(t_0)(h)}{\|h\|} = 0$$

Es gilt:

$$F(t_0 + h) = f\big(g(t_0 + h)\big) = f\big(\underbrace{g(t_0)}_{x_0} + \underbrace{Dg(t_0)(h) + \|h\|\varphi(h)}_{k}\big)$$

$$\begin{aligned}
&= f(x_0) + Df(x_0)(k) + \|k\|\psi(k) \\
&= f(x_0) + Df(x_0)\big(Dg(t_0)(h) + \|h\|\varphi(h)\big) + \|Dg(t_0)(h) + \|h\|\varphi(h)\| \cdot \psi\big(Dg(t_0)(h) + \|h\|\varphi(h)\big) \\
&= f(x_0) + Df(x_0)\big(Dg(t_0)(h)\big) + \Phi(h)
\end{aligned}$$

mit: $\Phi(h) = \|h\| Df(x_0)\big(\varphi(h)\big) + \|Dg(t_0)(h) + \|h\|\varphi(h)\| \cdot \psi\big(Dg(t_0)(h) + \|h\|\varphi(h)\big)$

Es gilt:

$$\frac{F(t_0 + h) - F(t_0) - \big(Df(x_0) \circ Dg(t_0)\big)(h)}{\|h\|} = \frac{\Phi(h)}{\|h\|}$$

Es gilt aber:

$$\begin{aligned}\|\Phi(h)\| &\le \|h\|\cdot\|Df(x_0)\big(\varphi(h)\big)\| + \|Dg(t_0)(h)+\|h\|\varphi(h)\|\cdot\big\|\psi\big(Dg(t_0)(h)+\|h\|\varphi(h)\big)\big\| \\ &\le \|h\|\cdot\|Df(x_0)\big(\varphi(h)\big)\| + \big(\|h\|\cdot\|Dg(t_0)\|+\|h\|\cdot\|\varphi(h)\|\big)\cdot\big\|\psi\big(Dg(t_0)(h)+\|h\|\varphi(h)\big)\big\| \\ \Longrightarrow \left\|\frac{\Phi(h)}{\|h\|}\right\| &\le \|Df(x_0)\big(\varphi(h)\big)\| + \big(\|Dg(t_0)\|+\|\varphi(h)\|\big)\cdot\big\|\psi\big(Dg(t_0)(h)+\|h\|\varphi(h)\big)\big\|\end{aligned}$$

Es gilt aber:

(i) $\|Df(x_0)\big(\varphi(h)\big)\| \xrightarrow{h\to 0} 0$ da $\varphi(h) \xrightarrow{h\to 0} 0$

(ii) $\big(\|Dg(t_0)\|+\|\varphi(h)\|\big)$ ist beschränkt

(iii) $\big\|\psi\big(Dg(t_0)(h)+\|h\|\varphi(h)\big)\big\| \xrightarrow{h\to 0} 0$, da ψ stetig an der Stelle 0 und

$$Dg(t_0)(h)+\|h\|\varphi(h) \xrightarrow{h\to 0} 0$$

Daraus folgt:

$$\lim_{h\to 0}\frac{\Phi(h)}{\|h\|} = 0$$

□

Bemerkung:
Die Jacobi-Matrix von $DF(t_0)$ berechnet sich durch entsprechende Multiplikation der Jacobi-Matrizen von $Df(x_0)$ und $Dg(t_0)$, d.h.

$$\begin{pmatrix}\operatorname{grad} F^{(1)}(t_0)\\ \operatorname{grad} F^{(2)}(t_0)\\ \vdots \\ \operatorname{grad} F^{(p)}(t_0)\end{pmatrix} = \begin{pmatrix}\operatorname{grad} f^{(1)}(x_0)\\ \operatorname{grad} f^{(2)}(x_0)\\ \vdots \\ \operatorname{grad} f^{(p)}(x_0)\end{pmatrix}\cdot\begin{pmatrix}\operatorname{grad} g^{(1)}(t_0)\\ \operatorname{grad} g^{(2)}(t_0)\\ \vdots \\ \operatorname{grad} g^{(n)}(t_0)\end{pmatrix}$$

Bzw. nach der Formel:

$$F_j^{(i)}(t_0) = \sum_{k=1}^{n} f_k^{(i)}\big(g(t_0)\big)g_j^{(k)}(t_0) \qquad 1\le i\le q,\ 1\le j\le r$$

Korollar 1:
Sei $p=1$ (d.h. f bzw. F reellwertig) und $F=f\circ g$:

$$F_j = \sum_{i=1}^{n} f_i\big(g(t_0)\big)g_j^{(i)}(t_0)$$

Korollar 2:
Sei $r=1$ und $F=f\circ g$:

$$F'(t_0) = Df\big(g(t_0)\big)g'(t_0)$$

Korollar 3:
Sind f und g von der Klasse C^q ($q \geq 1$), so auch $F = f \circ g$.

Beweis:
Dies liest man ab von der Gleichung:

$$F_j^{(i)}(t_0) = \sum_{k=1}^{n} f_k^{(i)}\big(g(t_0)\big) g_j^{(k)}(t_0)$$

(formal: Induktion)

Beispiel 46:
Sei $f \in C^2(\mathbb{R}^2)$, $h \in C^2(\mathbb{R})$ und g gegeben mit

$$g : \mathbb{R} \to \mathbb{R}^2 \qquad g(t) = \big(t, h(t)\big) \qquad \Longrightarrow \qquad g^{(1)}(t) = t \qquad g^{(2)}(t) = h(t)$$

Betrachtet man $F(t) = (f \circ g)(t)$, so erhält man:

$$\begin{aligned} F'(t) &= f_1\big(t, h(t)\big) + f_2\big(t, h(t)\big)h'(t) \\ F''(t) &= f_{11}\big(t, h(t)\big) + 2f_{12}\big(t, h(t)\big)h'(t) + f_2\big(t, h(t)\big)h''(t) + f_{22}\big(t, h(t)\big)\big(h'(t)\big)^2 \end{aligned}$$

Korollar 4:
Sei $n = r = p$, dann gilt:

$$Jf \circ g(t_0) = Jf\big(g(t_0)\big) \cdot Jg(t_0)$$

Beweis:
Wende auf die Beziehung

$$DF(t_0) = Df\big(g(t_0)\big) \circ Dg(t_0)$$

den Determinantenmultiplikationssatz an (Lineare Algebra). □

§4 Der Satz über die inverse Funktion

Satz 1:
Sei $\Delta \subset \mathbb{R}^n$ offen und $g : \Delta \to \mathbb{R}^n$ von der Klasse C^q auf Δ, wobei $q \geq 1$. Es gelte:

$$Jg(t) = \det\left(g_j^{(i)}(t)\right)_{1 \leq i,j \leq n} \neq 0 \qquad \forall t \in \Delta$$

Sei $t_0 \in \Delta$. Dann gibt es eine Umgebung U von t_0, $U \subset \Delta$, mit folgenden Eigenschaften:

(i) Die Einschränkung $g\big|_U$ von g auf U ist injektiv

(ii) Das Bild $V := g(U)$ von U unter g ist offen

(iii) Die inverse Abbildung f von $g\big|_U$ ist von der Klasse C^q auf V

Beweis:

(i) Nach §2 Lemma 3 gibt es zu vorgegebenem $\varepsilon > 0$ eine Umgebung Δ_0 von t_0, $\Delta_0 \subset \Delta$ derart, daß gilt

$$\|g(s) - g(t)\| \geq (c - \varepsilon)\|s - t\| \qquad \forall s, t \in \Delta_0$$

wobei: $c = \frac{1}{\|Dg^{-1}(t_0)\|}$
Wähle $\varepsilon := \frac{c}{2}$ und setze $U := \Delta_0$, dann gilt also:

$$\|g(s) - g(t)\| \geq (\frac{c}{2})\|s - t\| \qquad \forall s, t \in \Delta_0$$

Ist $g(s) = g(t)$ für $s, t \in U$, so folgt also $s = t$, also ist $g\big|_U$ injektiv.

(ii) *z.z.:* $V := g(U)$ ist offen, d.h. zu jedem $x_1 \in V$ gibt es eine Umgebung V_1 von x_1 mit $V_1 \subset U$.

Nach (i) gilt $x_1 = g(t_1)$ für genau ein $t_1 \in U$. Man wähle eine Umgebung U_1 von t_1, deren Abschluß $\overline{U_1}$ ganz in U enthalten ist. Sei $\Gamma_1 = \partial\overline{U_1}$, der Rand von $\overline{U_1}$.

Es gilt $x_1 \notin g(\Gamma_1)$, wegen der Injektivität von $g\big|_U$. (Wäre $x_1 = g(s)$ mit $s \in \Gamma_1$, also $g(t_1) = g(s)$, so folgt $t_1 = s$ ↯ Widerspruch! da $t_1 \notin \Gamma_1$.)

Die stetige Funktion $x \mapsto \|x - x_1\|$ nimmt auf der kompakten Menge $g(\Gamma_1)$ ihr Minimum m an. (Man beachte: Γ_1 ist kompakt, da g nach Voraussetzung von der Klasse C^1 ist, ist g stetig, also ist auch $g(\Gamma)$ kompakt, Kap. 3 §6 Satz 5 S. 70.)

Es gilt $m > 0$, denn sonst wäre $x_1 \in g(\Gamma_1)$. Sei $\sigma_1 := \frac{m}{2} > 0$ und sei V_1 die Umgebung von x_1 mit Radius σ_1. Wir wollen zeigen, daß $V_1 \subset V$.

Sei $x \in V_1$ fest. Für $t \in \Gamma_1$ gilt nach Definition von m:

$$\begin{aligned} 2\sigma_1 = m &\leq \|x_1 - g(t)\| = \|(x_1 - x) + (x - g(t))\| \\ &\leq \|x_1 - x\| + \|x - g(t)\| \underset{x \in V_1}{<} \sigma_1 + \|x - g(t)\| \\ \Longrightarrow \qquad \sigma_1 &< \|x - g(t)\| \qquad\qquad \forall t \in \Gamma_1 \end{aligned}$$

Die stetige Funktion $\psi : \overline{U_1} \to \mathbb{R}$ gegeben durch:

$$\psi(t) := \|x - g(t)\|^2 = \sum_{i=1}^{n} \left(x^{(i)} - g^{(i)}(t)\right)^2$$

nimmt auf der kompakten Menge $\overline{U_1}$ ihr Minimum an. Es gilt aber:

$$\psi(t_1) = \|x - g(t_1)\|^2 = \|x - x_1\|^2 < \sigma_1^2$$

und

$$\psi(t) = \|x - g(t)\|^2 > \sigma_1^2 \qquad \forall t \in \Gamma_1$$

Daher ist dieses Minimum echt kleiner als σ_1^2 und kann nicht in einem Punkt aus Γ_1 angenommen werden, in anderen Worten das Minimum wird in einem Punkt $t_2 \in U_1$ angenommen.

Die Funktion $\psi\big|_U$ ist nach Voraussetzung differenzierbar, denn g ist nach Voraussetzung mindestens von der Klasse C^1, also ist auch $\psi\big|_U$ von der Klasse C^1 (§3 Kor. 3 S. 107), also auch differenzierbar. Daher ist t_2 ein kritischer Punkt von ψ, d.h. $d\psi(t_2) = 0$ (Kap. 4 §4 Satz 1 S. 87).

Nach der Kettenregel gilt:

$$d\psi(t) = -2\sum_{i=1}^{n}\big(x^{(i)} - g^{(i)}(t)\big)dg^{(i)}(t)$$

Also gilt:

$$0 = \sum_{i=1}^{n} c_i dg^{(i)}(t_2) \qquad\qquad \text{wobei: } c_i = x^{(i)} - g^{(i)}(t_2)$$

Da nach Voraussetzung $Jg(t_2) \neq 0$ gilt, sind die Zeilenvektoren

$$dg^{(1)}(t_2), dg^{(2)}(t_2), \ldots, dg^{(n)}(t_2)$$

linear unabhängig, also folgt $c_i = 0\ \forall i = 1, 2, \ldots, n$.

Daher gilt:

$$x = g(t_2)$$

Also gilt

$$x \in g(U_1) \subset g(U) = V$$

Daher ist V offen.

(iii) Wir müssen zeigen, daß f auf V von der Klasse C^q ist. Sei $x_1 \in V$ und $x_1 = g(t_1)$, $t_1 \in U$, wie unter (ii), also $t_1 = f(x_1)$. Sei $L_1 := Dg(t_1)$.

Wir wollen zunächst zeigen, daß f in x_1 differenzierbar ist und $Df(x_1) = L_1^{-1}$ ist.

Sei $c_1 := \frac{1}{\|L_1^{-1}\|}$, sei $\varepsilon > 0$. Da g in t_1 differenzierbar ist, gibt es eine Umgebung U_2 von t_1, $U_2 \subset U$ derart, daß gilt

$$\|g(t) - g(t_1) - L_1(t - t_1)\| \leq \frac{\varepsilon c_1 c}{2}\|t - t_1\| \qquad \forall t \in U_2 \tag{1}$$

wobei c wie unter (i) gewählt ist, d.h. also

$$\|g(s) - g(t)\| \geq \frac{c}{2}\|s - t\| \qquad \forall s, t \in U \tag{$*$}$$

Wendet man die schon bewiesenen Aussagen (i), (ii) an mit Δ ersetzt durch U_2, g ersetzt durch $g\big|_{U_2}$ und t_0 ersetzt durch t_1, so folgt also die Existenz einer Umgebung W von t_1,

$W \subset U_2$, so daß $g(W)$ offen ist. Wegen $x_1 \in g(W)$ gibt es also eine Umgebung V_2 von x_1 mit $V_2 \subset g(W) \subset g(U_2)$.

Sei $x \in V_2$. Dann gilt also:

$$x = g(t) \qquad \text{mit } t \in U_2$$

Nach $(*)$ gilt:

$$\frac{c}{2}\|t - t_1\| \leq \|g(t) - g(t_1)\| = \|x - x_1\| \tag{2}$$

Ferner gilt:

$$\begin{aligned} L_1\big(f(x) - f(x_1) - L_1^{-1}(x - x_1)\big) &= L_1\big(f(x) - f(x_1)\big) - (x - x_1) \\ =L_1(t - t_1) - \big(g(t) - g(t_1)\big) &= -\Big(\big(g(t) - g(t_1)\big) - L_1(t - t_1)\Big) \end{aligned} \tag{3}$$

Ferner gilt $\forall \tau \in \mathbb{R}^n$:

$$\|\tau\| = \|L_1^{-1}\big(L_1(\tau)\big)\| \leq \|L_1^{-1}\| \cdot \|L_1(\tau)\|$$

also gilt (mit $c_1 = \frac{1}{\|L_1^{-1}\|}$):

$$c_1\|\tau\| \leq \|L_1(\tau)\| \tag{4}$$

Setzt man $\tau = f(x) - f(x_1) - L_1^{-1}(x - x_1)$, dann folgt:

$$\begin{aligned} c_1\|\tau\| &= c_1\|f(x) - f(x_1) - L_1^{-1}(x - x_1)\| \\ \text{(wegen (4))} \quad &\leq \|L_1\big(f(x) - f(x_1) - L_1^{-1}(x - x_1)\big)\| \\ \text{(wegen (3))} \quad &= \|g(t) - g(t_1) - L_1(t - t_1)\| \\ \text{(wegen (1))} \quad &\leq \frac{\varepsilon c_1 c}{2}\|t - t_1\| \\ \text{(wegen (2))} \quad &\leq \varepsilon c_1\|x - x_1\| \end{aligned}$$

Nach Division mit c_1 erhält man:

$$\|f(x) - f(x_1) - L_1^{-1}(x - x_1)\| \leq \varepsilon\|x - x_1\| \qquad \forall x \in V_2$$

Daher ist f in x_1 differenzierbar und $Df(x_1) = L_1^{-1}$.

Wir haben also gezeigt, daß f auf V differenzierbar ist und daß gilt

$$Df(x) = Dg^{-1}\big(f(x)\big)$$

Da f differenzierbar ist, ist f stetig. Da $g_j^{(i)}$ nach Voraussetzung stetig ist, ist auch $g_j^{(i)}\big(f(x)\big)$ stetig. Die Einträge der zu $Dg\big(f(x)\big)$ gehörigen Matrix haben also stetige Komponenten. Ist A eine invertierbare reelle Matrix, so gilt:

$$A^{-1} = \frac{1}{\det A}A^{ad}$$

wobei A^{ad} die „adjungierte Matrix" zu A ist, die Komponenten von A^{ad} sind dabei Polynome in den Komponenten von A.

Wenden wir obige Situation mit $A = Dg\big(f(x)\big)$ an, so folgt also die Stetigkeit von $f_j^{(i)}(x)$. Daher ist f von der Klasse C^1. Induktiv folgt dann Schritt für Schritt, daß f von der Klasse C^q ist, falls g dies ist.

□

Bemerkung:
Wir haben mitbewiesen:

$$\partial f(x) = \frac{1}{\partial g\big(f(x)\big)}$$

(folgt aus $Df(x) = Dg^{-1}\big(f(x)\big)$)

Korollar 1:
Es mögen die Voraussetzungen des Satzes gelten. Dann ist g offen, d.h. das Bild $g(B)$ jeder offenen Menge $B \subset \Delta$ ist wieder offen.

Beweis:
Sei $B \subset \Delta$ offen. Sei $x_0 \in g(B)$. Man schreibe $x_0 = g(t_0)$ mit $t_0 \in B$. Man wende den Satz auf $g\big|_B$ an. Dann existiert also eine Umgebung U von t_0, so daß $U \subset B$ und insbesondere $g(U)$ offen ist. Wegen $x_0 \in g(U)$ gibt es eine Umgebung V von x_0 mit $V \subset g(U) \subset g(B)$. □

Beispiel 47:
$n = 2$. Sei $\Delta = \big\{(s,t) \in \mathbb{R}^2 \big| s > 0\big\}$ und g gegeben durch

$$g : \Delta \to \mathbb{R}^2 \qquad\qquad g(s,t) = (\cosh s \cos t, \sinh s \sin t)$$

(Erinnerung: $\cosh s = \frac{e^s + e^{-s}}{2}$, $\sinh s = \frac{e^s - e^{-s}}{2}$)
Offenbar ist g von der Klasse C^q, $\forall q > 1$. Es gilt

$$\begin{pmatrix} g_1^{(1)} & g_2^{(1)} \\ g_1^{(2)} & g_2^{(2)} \end{pmatrix} = \begin{pmatrix} \sinh s \cos t & -\cosh s \sin t \\ \cosh s \sin t & \sinh s \cos t \end{pmatrix}$$

$$\begin{aligned} Jg(s,t) &= \sinh^2 s \cos^2 t + \cosh^2 s \sin^2 t = \sinh^2 s \cos^2 t + (1 + \sinh^2 s) \sin^2 t \\ &= \sinh^2 s + \sin^2 t \neq 0 \end{aligned}$$

denn für $s > 0$ ist $\sinh s > 0$.

Also sind die Voraussetzungen des Satzes erfüllt, die Funktion ist also „lokal invertierbar", d.h. zu jedem $(s_0, t_0) \in \Delta$ existiert eine Umgebung U von (s_0, t_0), so daß (i), (ii), (iii) gelten.

*Achtung!:*Die Abbildung $g : \Delta \to \mathbb{R}^2$ besitzt *kein* Inverses, denn es gilt $g(s, t + 2\pi) = g(s,t)$, wegen der Periodizität von cos und sin, also ist g (auf Δ) nicht injektiv!

§5 Der Satz über die implizite Funktion

Ziel: Gegeben sei eine Gleichung:

$$\Phi\big(x_1, x_2, \ldots, x_n\big) = 0$$

unter welchen Voraussetzungen kann dann eine der Variablen $x_1, x_2, \ldots, x_n$ als Funktion der übrigen $n-1$ Variablen ausgedrückt werden?

Allgemeiner: Gegeben sei ein Gleichungssystem:

$$\Phi^{(1)}(x) = 0, \Phi^{(2)}(x) = 0, \ldots, \Phi^{(m)}(x) = 0 \qquad x = (x_1, x_2, \ldots, x_n), 1 \leq m < n$$

unter welchen Voraussetzungen können dann m der Variablen $x_1, x_2, \ldots, x_n$ als Funktion der übrigen Variablen ausgedrückt werden?

Man setze $r = n - m$ und schreibe:

$$\hat{x} = (x_1, x_2, \ldots, x_r) \qquad\qquad x = (x_1, x_2, \ldots, x_n)$$

Ist $\Phi = (\Phi^{(1)}, \Phi^{(2)}, \ldots, \Phi^{(m)}) : \Delta \to \mathbb{R}^m$, wobei $\Delta \subset \mathbb{R}^n$ offen, mindestens von der Klasse C^1 auf Δ, so schreiben wir

$$\frac{\partial(\Phi^{(1)}, \Phi^{(2)}, \ldots, \Phi^{(m)})}{\partial(x_{r+1}, \ldots, x_n)}(x)$$

für die Determinante derjenigen Matrix, die gerade aus den letzten m Spalten der Jacobi-Matrix von Φ in x besteht.

Satz 1:
Sei $\Delta \subset \mathbb{R}^n$ offen und $x_0 \in \Delta$. Sei $1 \leq m < n$ und $r = n - m$. Sei

$$\Phi = (\Phi^{(1)}, \Phi^{(2)}, \ldots, \Phi^{(m)}) : \Delta \to \mathbb{R}^m$$

von der Klasse C^q auf Δ, wobei $q \geq 1$. Sei $\Phi(x_0) = 0$ und

$$\frac{\partial(\Phi^{(1)}, \Phi^{(2)}, \ldots, \Phi^{(m)})}{\partial(x_{r+1}, \ldots, x_n)}(x_0) \neq 0$$

Dann gibt es eine Umgebung U von x_0, eine offene Menge $V \subset \mathbb{R}^r$ und eine Funktion $\varphi = (\varphi^{(1)}, \varphi^{(2)}, \ldots, \varphi^{(m)}) : V \to \mathbb{R}^m$ derart, daß gilt:

(i) $\frac{\partial(\Phi^{(1)}, \Phi^{(2)}, \ldots, \Phi^{(m)})}{\partial(x_{r+1}, \ldots, x_n)}(x) \neq 0 \qquad \forall x \in U$

(ii) $\big\{x \in U \big| \Phi(x) = 0\big\} = \big\{\big(\hat{x}, \varphi(\hat{x})\big) \big| \hat{x} \in V\big\}$

Beweis:
Da Φ mindestens von der Klasse C^1 auf Δ ist, ist die Determinante

$$\frac{\partial(\Phi^{(1)}, \Phi^{(2)}, \ldots, \Phi^{(m)})}{\partial(x_{r+1}, \ldots, x_n)}(x)$$

stetig auf Δ. Wegen

$$\frac{\partial(\Phi^{(1)}, \Phi^{(2)}, \ldots, \Phi^{(m)})}{\partial(x_{r+1}, \ldots, x_n)}(x_0) \neq 0$$

gibt es eine Umgebung U_0 von x_0, so daß diese Determinate $\neq 0$ ist, für alle $x \in U_0$.
Sei $f : U_0 \to \mathbb{R}^n$

$$f(x_1, x_2, \ldots, x_n) = \big(x_1, x_2, \ldots, x_r, \Phi^{(1)}(x), \Phi^{(2)}(x), \ldots, \Phi^{(m)}(x)\big) = \big(\hat{x}, \Phi(x)\big)$$

Dann ist f von der Klasse C^q auf U_0.
Es gilt:

$$\left(f_j^{(i)}\right)_{1\leq i,j\leq n} = \begin{pmatrix} 1 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 1 & 0 & \cdots & 0 \\ \Phi_1^{(1)} & \Phi_2^{(1)} & \cdots & \Phi_r^{(1)} & \Phi_{r+1}^{(1)} & \cdots & \Phi_n^{(1)} \\ \Phi_1^{(2)} & \Phi_2^{(2)} & \cdots & \Phi_r^{(2)} & \Phi_{r+1}^{(2)} & \cdots & \Phi_n^{(2)} \\ \vdots & \vdots & \ddots & \vdots & \vdots & & \vdots \\ \Phi_1^{(m)} & \Phi_2^{(m)} & \cdots & \Phi_r^{(m)} & \Phi_{r+1}^{(m)} & \cdots & \Phi_n^{(m)} \end{pmatrix}$$

Nach dem Laplace'schen Entwicklungsatz folgt:

$$Jf(x) = \det \begin{pmatrix} \Phi_{r+1}^{(1)} & \Phi_{r+2}^{(1)} & \cdots & \Phi_n^{(1)} \\ \Phi_{r+1}^{(2)} & \Phi_{r+2}^{(2)} & \cdots & \Phi_n^{(2)} \\ \vdots & \vdots & \ddots & \vdots \\ \Phi_{r+1}^{(m)} & \Phi_{r+2}^{(m)} & \cdots & \Phi_n^{(m)} \end{pmatrix} = \frac{\partial(\Phi^{(1)}, \Phi^{(2)}, \ldots, \Phi^{(m)})}{\partial(x_{r+1}, \ldots, x_n)}(x) \neq 0 \qquad \forall x \in U_0$$

wie schon gezeigt. Nach dem Satz über inverse Funktionen (§4 Satz 1 S. 107) angewandt auf f gibt es also eine Umgebung U von x_0, so daß $f\big|_U$ injektiv, $f(U)$ offen ist und die inverse Abbildung g von $f\big|_U$ auf $f(U)$ von der Klasse C^q. (Man beachte, daß in den Notationen im Vergleich zu §4 die Rollen von f und g vertauscht sind.)

Wegen $f \circ g(y) = y$ und $f^{(1)}(x) = x_1, f^{(2)}(x) = x_2, \ldots, f^{(r)}(x) = x_r$ gilt auch $g^{(1)}(y) = y_1, g^{(2)}(y) = y_2, \ldots, g^{(r)}(y) = y_r$. Sei

$$V := \left\{\hat{x} \in \mathbb{R}^r \,\middle|\, (\hat{x}, 0) \in f(U)\right\}$$

(Anm: $(\hat{x}, 0) = (x_1, x_2, \ldots, x_r, 0, \ldots, 0)$) Es gilt $\hat{x}_0 \in V$, denn $(\hat{x}_0, 0) = \big(\hat{x}_0, \Phi(x_0)\big) = f(x_0)$ und $x_0 \in U$. Ferner ist V offen, denn V ist das Urbild der offenen Menge $f(U)$ unter der stetigen Abbildung $\hat{x} \mapsto (\hat{x}, 0)$.

Für $\hat{x} \in V$ setze man

$$\varphi^{(l)}(\hat{x}) = g^{(r+l)}\big((\hat{x}, 0)\big) \qquad l = 1, 2, \ldots, m$$

Es gilt dann $\varphi = \big(\varphi^{(1)}, \varphi^{(2)}, \dots, \varphi^{(m)}\big) : V \to \mathbb{R}^m$ ist von der Klasse C^q auf V. Ferner gilt

$$x \in U,\ \Phi(x) = 0 \qquad \Longleftrightarrow \qquad \hat{x} \in V,\ f(x) = (\hat{x}, 0)$$

"$\Leftarrow$": $\Phi(x) = 0$ klar, wegen der Definition von f.

$$f(x) = (\hat{x}, 0) \quad \Longrightarrow \quad g \circ f(x) = g\big((\hat{x}, 0)\big) \quad \Longrightarrow \quad x = g\big((\hat{x}, 0)\big) \in U$$

wegen der Definition von g.

Daraus ergibt sich:

$$\begin{array}{lcl} x \in U,\ \Phi(x) = 0 & \Longleftrightarrow & \hat{x} \in V,\ f(x) = (\hat{x}, 0) \\ & \Longleftrightarrow & \hat{x} \in V,\ x = g \circ f(x) = g\big((\hat{x}, 0)\big) \\ & & x = \big(\hat{x}, \varphi(\hat{x})\big) \end{array}$$

$\square$

Bemerkung (Nachtrag zum Satz über die implizite Funktion):
Die Voraussetzung

$$\frac{\partial(\Phi^{(1)}, \Phi^{(2)}, \dots, \Phi^{(m)})}{\partial(x_{r+1}, \dots, x_n)}(x_0) \neq 0$$

$(r = n - m)$, d.h. die letzten m Spalten der Matrix $D\Phi(x_0)$ sind linear unabhängig, kann man durch die Voraussetzung, daß $D\Phi(x_0)$ maximalen Rang (also Rang m) hat, ersetzen.

Haben dann die m linear unabhängigen Spalten die Indizes $j_1, j_2, \dots, j_m$ mit $j_1 < j_2 < \dots < j_m$ und ist

$$\big\{1, 2, \dots, n\big\} \setminus \big\{j_1, j_2, \dots, j_m\big\} = \big\{i_1, i_2, \dots, i_r\big\}$$

mit $i_1 < i_2 < \dots < i_r$, so zeigt man wie vorher die Existenz von U, V und Φ mit

(i) $\frac{\partial(\Phi^{(1)}, \Phi^{(2)}, \dots, \Phi^{(m)})}{\partial(x_{j_1}, x_{j_2}, \dots, x_{j_m})}(x) \neq 0 \qquad \forall x \in U$

(ii) $\big\{x \in U \big| \Phi(x) = 0\big\}$
$= \big\{x \big| \big(x_{i_1}, x_{i_2}, \dots, x_{i_r}\big) \in V, x_{j_l} = \varphi^{(l)}\big(x_{i_1}, x_{i_2}, \dots, x_{i_r}\big)\ \forall l = 1, 2, \dots, m\big\}$

Index