2.2 Totale D'barkeit

Die aus der Analysis I bekannte Schlußweise

$$f \text{ diff'bar} \implies f \text{ stetig}$$

ist nicht verallgemeinerbar auf Funktionen mehrerer Veränderlicher, wenn wir von f nur die partielle Differenzierbarkeit verlangen. Wir müssen einen etwas schärferen Differenzierbarkeitsbegriff für solche Funktionen einführen, damit die obige Schlußweise wieder korrekt ist. Dies führt auf den Begriff der totalen D'barkeit. Um diesen zu erläutern, gehen wir noch einmal zurück zu Funktionen einer reellen Veränderlichen:

Lemma 1: Es sei $I \subset \mathbb{R}$ ein Intervall, $x_0 \in I$ und $f: I \to \mathbb{R}$ eine Funktion. Dann gilt: f ist d'bar in x_0 genau dann, wenn $c \in \mathbb{R}$, $\varepsilon > 0$ und eine Funktion $\varphi: (x_0-\varepsilon, x_0+\varepsilon)$ mit $\lim\limits_{h \to 0} \frac{\varphi(h)}{h} = 0$ existieren, so daß für alle $|h| < \varepsilon$

$$f(x_0+h) = f(x_0) + c \cdot h + \varphi(h). \qquad (*)$$

Bew.: Aus $(*)$ folgt

$$\frac{1}{h}\big(f(x_0+h) - f(x_0)\big) = c + \frac{\varphi(h)}{h} \longrightarrow c \quad (h \to 0),$$

und das ist die Diff'barkeit von f in x_0.

Umgekehrt sei die Diff'barkeit von f in x_0 vorausgesetzt. Dann setzen wir

$$\varphi(h) = f(x_0+h) - f(x_0) - f'(x_0)\cdot h\,,$$

so daß (*) mit $c = f'(x_0)$ gilt. Ferner ist

$$\frac{\varphi(h)}{h} = \frac{f(x_0+h)-f(x_0)}{h} - f'(x_0) \longrightarrow 0 \quad (h \to 0),$$

also ist auch die zusätzliche Eigenschaft von φ erfüllt. $\square$

Die in Lemma 1 formulierte äquivalente Eigenschaft bezeichnet man als "Approximierbarkeit durch eine lineare Abbildung" oder kurz "lineare Approximierbarkeit". Im Ggs. zur traditionellen Definition

$$\exists \lim_{h\to 0} \frac{1}{h}\left(f(x_0+h)-f(x_0)\right),$$

bei der man für h keinen Vektor einsetzen kann, ist sie problemlos verallgemeinerbar für Funktionen mehrerer Veränderlicher:

Def.: Es sei $\Omega \subset \mathbb{R}^n$ offen und $f: \Omega \to \mathbb{R}^m$ eine Funktion. f heißt (total) d'bar in $x_0 \in \Omega$, falls eine lineare Abbildung $A: \mathbb{R}^n \to \mathbb{R}^m$ existiert, so daß

$$f(x_0+h) = f(x_0) + Ah + \varphi(h)$$

gilt. Hierbei ist $\varphi: \mathbb{R}^n \supset B_\varepsilon(0) \to \mathbb{R}^m$ mit der Eigenschaft $\displaystyle\lim_{h\to 0} \frac{\varphi(h)}{|h|} = 0$.

Bem.: (1) Ausgeschrieben als Vektoren

$f = (f_1, \dots, f_m)^T$, $\varphi = (\varphi_1, \dots, \varphi_m)^T$ bzw. als Matrix

$A = (a_{ij})_{\substack{1 \le i \le m \\ 1 \le j \le n}}$ lautet die charakterisierende

Gleichung

$$f_i(x_0 + h) = f_i(x_0) + \sum_{j=1}^{n} a_{ij} h_j + \varphi_i(h), \quad 1 \le i \le m.$$

Also ist f in x_0 (total) diff.'bar genau dann, wenn dies für alle Komponenten von f gilt.

(2) f heißt (total) diff.bar in Ω, wenn f in jedem $x_0 \in \Omega$ (total) diffbar ist.

Bsp.: $f: \mathbb{R}^n \to \mathbb{R}$, $x \mapsto f(x) := \langle x, Mx \rangle = \sum_{i,j=1}^{n} m_{ij} x_i x_j$

mit einer Matrix $M \in M_n(\mathbb{R})$. Dann ist

$$f(x+h) = \langle x+h, M(x+h) \rangle = \langle x, Mx \rangle + \langle h, Mx \rangle + \dots$$

$$\dots + \langle x, Mh \rangle + \langle h, Mh \rangle = f(x) + A \cdot h + \varphi(h),$$

wobei $A = A(x) = Mx + M^T x = (M + M^T)\, x$

und $\varphi(h) = \langle h, Mh \rangle$ mit

$$\lim_{h \to 0} \frac{|\varphi(h)|}{|h|} = \lim_{h \to 0} \frac{1}{|h|} \underbrace{|\langle h, Mh \rangle|}_{\le \|M\|\,|h|} = 0$$

Die Approximierende lineare Abbildung

$A: h \to Ah = \langle (M + M^T)x, h \rangle$ hängt also von x ab!

Lemma 2: Ist $f: \mathbb{R}^n \supset \Omega \to \mathbb{R}^m$ (total) diff.'bar in $x_0 \in \Omega$, so ist f in x_0 stetig.

Bew.: $|f(x_0+h) - f(x_0)| = |Ah + \varphi(h)|$

$$\leq \|A\| \cdot |h| + \frac{|\varphi(h)|}{|h|} \cdot |h| \longrightarrow 0 \quad (h \to 0). \quad \square$$

Welchen Zusammenhang besteht nun zwischen totaler und partieller Differenzierbarkeit?

Satz 1: Es sei $\Omega \subset \mathbb{R}^n$ offen und $f: \Omega \to \mathbb{R}^m$, total d'bar in $x_0 \in \Omega$, es gelte also

$$f(x_0+h) = f(x_0) + Ah + \varphi(h)$$

mit einer Matrix $A = (a_{ij})_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}}$ und einer Funktion $\varphi: \mathbb{R}^n \supset U_\varepsilon(0) \to \mathbb{R}^m$, für die $\lim\limits_{h \to 0} \frac{\varphi(h)}{|h|} = 0$ ist.

Dann ist f in x_0 partiell diff.'bar und es gilt

$$a_{ij} = \frac{\partial f_i}{\partial x_j}(x_0) \qquad (1 \leq i \leq m,\ 1 \leq j \leq n).$$

Bew.: Mit $h = t \cdot e_j$ haben wir für die Komponenten

$$f_i(x_0 + te_j) = f_i(x_0) + t(Ae_j)_i + \varphi_i(te_j) \quad \text{bzw.}$$

$$\frac{1}{t}(f_i(x_0+te_j) - f_i(x_0)) = \underbrace{\sum_{k=1}^{n} a_{ik}\delta_{kj}}_{= a_{ij}} + \frac{\varphi_i(te_j)}{t}$$

Für $t \to 0$ verschwindet der letzte Beitrag, so daß

$$\exists \lim_{t \to 0} \frac{1}{t}(f_i(x_0+te_j) - f(x_0)) = \frac{\partial f_i}{\partial x_j}(x_0) = a_{ij} \quad \square$$

Bem. / Bez.: Die Matrix A ist also durch die Funktion f eindeutig festgelegt. Man nennt sie die "Jacobimatrix" oder auch "Funktionalmatrix" von f. Schreibweisen: $A = Df(x_0) = J_f(x_0)$

bzw. $$A = \left(\frac{\partial f_i}{\partial x_j}(x_0) \right)_{\substack{1 \le i \le m \\ 1 \le j \le n}}$$

Ist f reellwertig, besteht Df nur aus einer einzigen Zeile: $Df(x_0) = \nabla f(x_0)$, falls $f: \Omega \to \mathbb{R}$.

Für vektorwertige Funktionen gilt: Die Zeilen der Jacobi-Matrix sind die Gradienten der Komponentenfunktionen: Ist $f: \Omega \to \mathbb{R}^m$, $m \ge 2$:

$$Df(x_0) = \begin{pmatrix} - \nabla f_1(x_0) - \\ \vdots \\ - \nabla f_m(x_0) - \end{pmatrix}.$$

Wir können die Jacobi-Matrix auch spaltenweise auffassen: Die Spalten der Matrix Df sind gerade die partiellen Ableitungen der vektorwertigen Funktion f:

$$Df(x_0) = \begin{pmatrix} | & & | \\ \frac{\partial f}{\partial x_1}(x_0) & \cdots & \frac{\partial f}{\partial x_n}(x_0) \\ | & & | \end{pmatrix}.$$

Bsp. zur Berechnung einer Jacobi-Matrix:

$$f: \mathbb{R}^3 \longrightarrow \mathbb{R}^2, \quad x = (x_1, x_2, x_3)^T \longmapsto f(x) = \begin{pmatrix} f_1(x) \\ f_2(x) \end{pmatrix}$$

mit $f_1(x) = x_1^2 + x_2^2 + x_3^2$ und $f_2(x) = x_1 x_2 x_3$.

Hierfür ist

$$Df(x) = \left(\frac{\partial f_i}{\partial x_j}(x) \right)_{\substack{1 \le i \le 2 \\ 1 \le j \le 3}} = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(x) & \dots & \frac{\partial f_1}{\partial x_3}(x) \\ \frac{\partial f_2}{\partial x_1}(x) & \dots & \frac{\partial f_2}{\partial x_3}(x) \end{pmatrix}$$

$$= \begin{pmatrix} 2x_1 & 2x_2 & 2x_3 \\ x_2 x_3 & x_1 x_3 & x_1 x_2 \end{pmatrix}$$

Bem.: Beispiel zeigt: Die Abhängigkeit der Matrix Df von x ist in der Regel nichtlinear. Wenn in der Definition der totalen D'barkeit von "linearer Abbildung" die Rede ist, so ist

$$h \longmapsto Df(x) \cdot h \qquad \text{(bei festem } x\text{) gemeint.}$$

(To be continued.)

Kann man anhand der partiellen Ableitungen bereits feststellen, ob eine Funktion total differenzierbar ist? Hierzu gibt ein zwar nicht scharfes, aber zumindest hinreichendes Kriterium:

Satz 2: Sei $\Omega \subset \mathbb{R}^n$ offen und $f: \Omega \to \mathbb{R}^m$ partiell d'bar. Alle partiellen Ableitungen $\frac{\partial f}{\partial x_j}$ seien in $x_0 \in \Omega$ stetig. Dann ist f in x_0 total diff.'bar.

Bew.: O.E. $m=1$. Wir setzen

$$\varphi(h) = f(x_0+h) - f(x_0) - \nabla f(x_0)\cdot h$$

und zeigen, dass $\lim\limits_{h\to 0} \frac{\varphi(h)}{|h|} = 0$. Dazu führen wir

$$z_k = x_0 + \sum_{j=1}^{k} h_j e_j$$

ein, so daß

$z_0 = x_0$, $z_n = x_0+h$ und $z_k - z_{k-1} = h_k e_k$. Dann gilt

$$f(x_0+h) - f(x_0) = \sum_{k=1}^{n} f(z_k) - f(z_{k-1})$$

$$= \sum_{k=1}^{n} f(x_0^{(1)}+h_1, \dots, x_0^{(k-1)}+h_{k-1}, x_0^{(k)}+h_k, x_0^{(k+1)}, \dots, x_0^{(n)})$$

$$- f(x_0^{(1)}+h_1, \dots, x_0^{(k-1)}+h_{k-1}, x_0^{(k)}, \dots, x_0^{(n)})$$

$$= \sum_{k=1}^{n} \frac{\partial f}{\partial x_k}(\xi_k)\cdot h_k \quad , \quad \xi_k \in \{z_{k-1} + \lambda h_k e_k : 0 \le \lambda \le 1\}$$

Hierbei haben wir im letzten Schritt den MWS für Funktionen einer Veränderlicher benutzt.

weil $\nabla f(x_0)\cdot h = \sum_{k=1}^{n} \frac{\partial f}{\partial x_k}(x_0)\cdot h_k$ folgt

$$\frac{\varphi(h)}{|h|} = \sum_{k=1}^{n} \left(\frac{\partial f}{\partial x_k}(\xi_k) - \frac{\partial f}{\partial x_k}(x_0) \right) \cdot \frac{h_k}{|h|} \longrightarrow 0 \quad (|h|\to 0),$$

letzteres wegen $\left|\frac{h_k}{|h|}\right| \leq 1$ und aufgrund der Stetigkeit aller partiellen Ableitungen in x_0. $\square$

Zsf.: Zwischen den Regularitätseigenschaften einer Funktion mehrerer Veränderlicher bestehen also die Implikationen

$$\begin{tikzcd}
 & & \text{stetig} \\
\text{stetig partiell diff'bar} \arrow[r, Rightarrow, "\text{Satz 2}"'] & \text{total diff'bar} \arrow[ur, Rightarrow, "\text{Lemma 2}"'] \arrow[dr, Rightarrow, "\text{Satz 1}"'] & \\
 & & \text{partiell diff'bar}
\end{tikzcd}$$

Insbesondere ist jede stetig partiell diff'bare Funktion stetig. Alle Umkehrungen im obigen Diagramm gelten i. allg. nicht!

Bsp. zu Satz 1: $f(x,y) = \begin{cases} \frac{2xy}{x^2+y^2} & (x,y) \neq (0,0) \\ 0 & (x,y) = (0,0) \end{cases}$ ist im

Nullpunkt zwar partiell d'bar, aber unstetig und somit nicht total diffbar.

Bsp. zu Satz 2: Dieses Phänomen tritt bereits im eindimensionalen Fall auf für $g(x) = x^2\cos(\frac{1}{x})$ $(x \neq 0)$, $g(0) = 0$. Hierfür ist

$$g'(x) = \begin{cases} 2x\cos(\frac{1}{x}) + \sin(\frac{1}{x}) & (x \neq 0) \\ 0 & (x = 0) \end{cases} \quad \text{unstetig in } x = 0.$$

Satz 3 (Kettenregel): Es seien $\Omega \subset \mathbb{R}^n$ und $\Omega' \in \mathbb{R}^k$ offen und

$$f: \Omega \to \mathbb{R}^m \qquad \text{sowie} \qquad g: \Omega' \to \Omega$$

Funktionen, so daß g in $x_0 \in \Omega'$ und f in $g(x_0) \in \Omega$ total diff'bar sind mit Funktionalmatrizen

$$Df(g(x_0)) \qquad \text{bzw.} \qquad Dg(x_0).$$

Dann ist auch $f \circ g: \Omega' \to \mathbb{R}^m$ in $x_0 \in \Omega'$ total d'bar, und es gilt

$$D f\circ g(x_0) = Df(g(x_0)) \cdot Dg(x_0).$$

Bew.: Wir haben

$$f(y+h) = f(y) + Df(y)\cdot h + \varphi(h),$$

wobei $\lim\limits_{h\to 0} \frac{\varphi(h)}{|h|} = 0$, und

$$g(x_0+k) = g(x_0) + Dg(x_0)\cdot k + \psi(k),$$

wobei $\lim\limits_{k\to 0} \frac{\psi(k)}{|k|} = 0$. Hieraus folgt mit $y = g(x_0)$

und $h = Dg(x_0)\cdot k + \psi(k)$

$$f(g(x_0+k)) - f(g(x_0))$$

$$= f(g(x_0) + \underbrace{Dg(x_0)k + \psi(k)}_{h}) - f(g(x_0)) = \ldots$$

$$= Df(g(x_0)) \cdot Dg(x_0) \cdot k + r(k)$$

wobei

$$r(k) = Df(g(x_0)) \cdot \psi(k) + \varphi(Dg(x_0) \cdot k + \psi(k)) =: r_1(k) + r_2(k).$$

Nun ist

$$\frac{|r_1(k)|}{|k|} = \frac{|Df(g(x_0))\,\psi(k)|}{|k|} \le \|Df(g(x_0))\| \, \frac{|\psi(k)|}{|k|} \longrightarrow 0 \quad (|k| \to 0)$$

und

$$\frac{|r_2(k)|}{|k|} = \frac{1}{|k|} \, |\varphi(Dg(x_0) \cdot k + \psi(k))|$$

$$= \frac{|\varphi(Dg(x_0)k + \psi(k))|}{|Dg(x_0)k + \psi(k)|} \cdot \frac{|Dg(x_0)k + \psi(k)|}{|k|}$$

$$\le (\|Dg(x_0)\| + 1) \cdot \frac{|\varphi(Dg(x_0)k + \psi(k))|}{|Dg(x_0) \cdot k + \psi(k)|} \longrightarrow 0 \quad (|k| \to 0)$$

k hinreichend klein

Insgesamt also $\lim_{k \to 0} \frac{r(k)}{|k|} = 0$ und damit die Behauptung. $\square$

Bsp. zur Anwendung der Kettenregel: Wir betrachten die Funktionen

$$f : \mathbb{R}^3 \to \mathbb{R}^2, \quad y = (y_1, y_2, y_3)^T \longmapsto f(y) = \begin{pmatrix} y_1^2 + y_2^2 + y_3^2 \\ y_1 y_2 y_3 \end{pmatrix}$$

aus dem vorigen Bsp. und

$$g : \mathbb{R}^2 \to \mathbb{R}^3, \quad x = (x_1, x_2)^T \longmapsto g(x) := \begin{pmatrix} g_1(x) \\ g_2(x) \\ g_3(x) \end{pmatrix} := \begin{pmatrix} x_1^2 \\ x_2^2 \\ x_1 x_2 \end{pmatrix}.$$

Bereits berechnet haben wir

$$Df(y) = \begin{pmatrix} 2y_1 & 2y_2 & 2y_3 \\ y_2 y_3 & y_1 y_3 & y_1 y_2 \end{pmatrix}$$

und für g erhalten wir

$$Dg(x) = \begin{pmatrix} 2x_1 & 0 \\ 0 & 2x_2 \\ x_2 & x_1 \end{pmatrix}$$

und damit $\quad D f \circ g(x) = Df(g(x)) \cdot Dg(x)$

$$= \begin{pmatrix} 2g_1(x) & 2g_2(x) & 2g_3(x) \\ g_2(x)g_3(x) & g_1(x)g_3(x) & g_1(x)g_2(x) \end{pmatrix} \begin{pmatrix} 2x_1 & 0 \\ 0 & 2x_2 \\ x_2 & x_1 \end{pmatrix}$$

$$= \begin{pmatrix} 2x_1^2 & 2x_2^2 & 2x_1x_2 \\ x_1x_2^3 & x_1^3x_2 & x_1^2x_2^2 \end{pmatrix} \begin{pmatrix} 2x_1 & 0 \\ 0 & 2x_2 \\ x_2 & x_1 \end{pmatrix}$$

$$= \begin{pmatrix} 4x_1^3 + 2x_1x_2^2 & 4x_2^3 + 2x_1^2x_2 \\ 3x_1^2x_2^3 & 3x_1^3x_2^2 \end{pmatrix}$$

Überprüfung: Man berechnet $f \circ g(x)$

$$= f(g(x)) = \begin{pmatrix} g_1^2(x) + g_2^2(x) + g_3^2(x) \\ g_1(x)\,g_2(x)\,g_3(x) \end{pmatrix} = \begin{pmatrix} x_1^4 + x_2^4 + x_1^2x_2^2 \\ x_1^3\,x_2^3 \end{pmatrix}$$

und bestimmt hiervon die Jacobi-Matrix. In der Tat führt dies zum selben Ergebnis.

Bsp.: Spezialfall der Kettenregel. Wir betrachten die folgende Situation:

$$\mathbb{R}^k \supset \Omega' \xrightarrow{g} \underset{\mathbb{R}^n}{\overset{\cap}{\Omega}} \xrightarrow{f} \mathbb{R} \quad (m=1!),$$

dabei sei die totale D'barkeit von f und g vorausgesetzt.

Dann ist $f \circ g$ eine reellwertige Funktion und also

$$D f\circ g(x) = \nabla f\circ g(x) = \left(\frac{\partial f\circ g}{\partial x_1}(x), \dots, \frac{\partial f\circ g}{\partial x_k}(x)\right).$$

Andererseits ergibt die Kettenregel (für $f : \Omega \to \mathbb{R}$):

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

$$= \left(\frac{\partial f}{\partial y_i}(g(x))\right)_{1\le i\le n} \cdot \left(\frac{\partial g_i}{\partial x_j}(x)\right)_{\substack{1\le i\le n\\ 1\le j\le k}}$$

In Komponenten: Für $1\le j\le k$ ist dann

$$\frac{\partial f\circ g}{\partial x_j}(x) = \sum_{i=1}^{n} \frac{\partial f}{\partial y_i}(g(x))\cdot\frac{\partial g_i}{\partial x_j}(x)$$

Weitere Spezialisierung: Es gelte auch $k=1$. Dann ist $\frac{\partial}{\partial x} = \frac{d}{dx}$ die (totale) eindim. Ableitung und

$$\frac{d f\circ g}{dx}(x) = \sum_{i=1}^{n} \frac{\partial f}{\partial y_i}(g(x))\cdot\frac{d g_i}{dx}(x) = \underbrace{\nabla f(g(x))}_{\text{Zeile}}\cdot\underbrace{g'(x)}_{\text{Spalte}}$$

Anwendung (Physik):

f = Komponente eines Feldes, das auf einen Massepunkt am Ort $(x,y,z) \in \mathbb{R}^3$ wirkt,

f sei zeit- und ortsabhängig, also

$$f: (t,x,y,z) \longmapsto f(t,x,y,z)$$

Soweit für einen festen Massepunkt am Ort (x,y,z).

Weitergehende Annahme: Der Massepunkt (Elementarteilchen o.ä.) bewegt sich auf einer Bahn

$$\gamma: \mathbb{R} \to \mathbb{R}^3, \ t \longmapsto \gamma(t) = (x(t), y(t), z(t))$$

Dann wird die Zeitabhängigkeit der Komponente des Feldes, die auf den Massepunkt wirkt beschrieben durch die Funktion

$$\varphi: \mathbb{R} \to \mathbb{R}, \ t \longmapsto \varphi(t) = f(t, x(t), y(t), z(t)).$$

Die Kettenregel in der zuletzt genannten Form ergibt hier

$$\frac{d\varphi}{dt}(t) = \frac{\partial f}{\partial t}(t, x(t), y(t), z(t)) \cdot 1$$

$$+ \frac{\partial f}{\partial x}(t, x(t), .., z(t)) \cdot x'(t) + \frac{\partial f}{\partial y}(t, x(t), .., z(t))\, y'(t) + \frac{\partial f}{\partial z}(t, x(t), ..., z(t))$$

Eine weitere Anwendung der Kettenregel liefert der Begriff der Richtungsableitung:

<u>Def.</u> (Richtungsableitung): Es sei $\Omega \subset \mathbb{R}^n$ offen, $x_0 \in \Omega$ und $f : \Omega \to \mathbb{R}^m$ eine Funktion. $\xi \in \mathbb{R}^n$ sei ein fester Vektor mit $|\xi| = 1$. Falls existent, nennen wir

$$\frac{\partial f}{\partial \xi}(x_0) := \lim_{t \to 0} \frac{1}{t}\left(f(x_0 + t\xi) - f(x_0)\right)$$

die <u>Richtungsableitung</u> von f nach ξ in x_0.

<u>Bem.</u>: (1) Wir nennen f in diesem Fall auch in x_0 nach ξ oder in Richtung von ξ differenzierbar.

(2) Für $\xi = e_\nu$ ist p.d. $\frac{\partial f}{\partial \xi}(x_0) = \frac{\partial f}{\partial x_\nu}(x_0)$. Die partiellen Ableitungen sind also die Richtungsableitungen nach den kanonischen Basisvektoren.

(3) Im Fall total diff'barer, reellwertiger Funktionen kann man Richtungsableitungen mit Hilfe des Gradienten recht einfach berechnen; s.u..

(4) Wird nur definiert für Vektoren ξ mit $|\xi| = 1$. Wir wollen eine Information über die Änderung von f in Richtung von ξ haben, die unabhängig ist von $|\xi|$.

<u>Satz</u> 4: Sei $\Omega \subset \mathbb{R}^n$ offen, $f: \Omega \to \mathbb{R}$ total diff'bar in $x_0 \in \Omega$ und $\xi \in \mathbb{R}^n$ mit $|\xi| = 1$. Dann existiert die Richtungsableitung von f nach ξ in x_0 und es gilt:

$$\frac{\partial f}{\partial \xi}(x_0) = \nabla f(x_0) \cdot \xi = \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}(x_0) \cdot \xi_j .$$

<u>Bew.</u>: Wir setzen (für ein geeignetes $\varepsilon > 0$)

$$\gamma : (-\varepsilon, \varepsilon) \longrightarrow \Omega, \quad t \mapsto \gamma(t) = x_0 + t\xi$$

und $\quad \varphi(t) = f(x_0 + t\xi) = f \circ \gamma(t).$

Nun sind γ d'bar in $t_0 = 0$ und f d'bar in $x_0 = \gamma(t_0)$, also nach der Kettenregel φ in $t_0 = 0$ d'bar und es gilt: Es existiert der Grenzwert

$$\frac{\partial f}{\partial \xi}(x_0) = \lim_{t \to 0} \frac{f(x_0 + t\xi) - f(x_0)}{t} = \lim_{t \to 0} \frac{\varphi(t) - \varphi(0)}{t} = \varphi'(t)\Big|_{t=0}$$

$$= \nabla f(\gamma(0)) \cdot \gamma'(0) = \nabla f(x_0) \cdot \xi \qquad \Box.$$

(vergleiche die obige Anwendung der Kettenregel!)

<u>Bem.</u>: (1) Ohne die Voraussetzung der totalen D'barkeit wird die Aussage des Satzes falsch.

(2) Umgekehrt ist die Existenz aller Richtungsableitungen nicht hinreichend für die totale D'barkeit, nicht einmal für die Stetigkeit.

(Beispiele zu (1) und (2): s. Übung)

(3) Wenn die Formel aus Satz 4 gilt, haben wir

$$\frac{\partial f}{\partial \xi}(x_0) = \nabla f(x_0)\cdot \xi = \langle \nabla f(x_0)^T, \xi \rangle = |\nabla f(x_0)|\cdot \cos(\alpha),$$

wobei α der Winkel zwischen ξ und $\nabla f(x_0)$ ist. Dies wird maximal für $\xi = \frac{\nabla f(x_0)}{|\nabla f(x_0)|}$, was bedeutet: <u>Der Gradient</u> $\nabla f(x_0)$ <u>weist in die Richtung des stärksten Anstiegs der Funktion f.</u>

(4) Diese geometrische Interpretation des Gradienten wird ergänzt durch eine weitere Anwendung der Kettenregel. Gegeben sei eine d'bare Kurve

$$\gamma : \mathbb{R} \supset I \longrightarrow N_f(c) \subset \mathbb{R}^n,$$

wobei $N_f(c) = \{x \in \mathbb{R}^n : f(x) = c\}$ die Niveaufläche von f zum Wert $c \in \mathbb{R}$ ist. Dann ist

$$c = f\circ\gamma(t) \Rightarrow 0 = \frac{d}{dt} f\circ\gamma(t) = \nabla f(\gamma(t))\cdot \gamma'(t),$$

dabei ist $\gamma'(t)$ eine Tangente an $N_f(c)$. Also: <u>Der Gradient steht senkrecht auf den Niveauflächen von f.</u>

Zum Mittelwertsatz: Für eine differenzierbare Fkt.

$$f : [a,b] \to \mathbb{R}, \quad t \mapsto f(t)$$

existiert eine Zwischenstelle $\xi \in (a,b)$, so daß

$$f(b) - f(a) = f'(\xi)(b-a) \qquad \text{(MWS)}.$$

Wir hatten bereits anhand des Beispiels

$$\gamma(t) = \begin{pmatrix} \cos(t) \\ \sin(t) \end{pmatrix} \quad (\Rightarrow |\gamma'(t)| = 1, \ b = 2\pi, \ a = 0 \ !)$$

eingesehen, dass dieser Satz für vektorwertige Funktionen nicht gelten kann. Was aber immerhin gelingt, ist der Beweis einer Ungleichung, die für viele Zwecke ausreicht.

Satz 5: Es seien $\Omega \subset \mathbb{R}^n$ offen und $x, x+h \in \Omega$, so dass die gesamte Strecke $[x, x+h] := \{x + th : 0 \le t \le 1\}$ in Ω liegt. $f : \Omega \to \mathbb{R}^m$ sei total diff'bar. Dann gilt

$$|f(x+h) - f(x)| \le \sup\{\|Df(\xi)\| : \xi \in [x, x+h]\}\,|h|.$$

Bew. Wir setzen $\varphi(t) = f(x+th)$. Dann ist nach dem Hauptsatz der Differenzial und Integralrechnung

$$f(x+h) - f(x) = \varphi(1) - \varphi(0) = \int_0^1 \varphi'(t)\,dt$$

$$\underset{\text{Kettenregel}}{\leadsto} \int_0^1 Df(x+th) \cdot h \, dt$$

und daher

$$|f(x+h)-f(x)| = \left|\int_0^1 Df(x+th)\cdot h\,dt\right|$$

$$\leq \int_0^1 |Df(x+th)\cdot h|\,dt$$ Dreiecksungleichung, Integral ist der Grenzwert von Riemann-Summen.

$$\leq \int_0^1 \|Df(x+th)\|\,|h|\,dt$$

$$\leq \sup\{\|Df(\xi)\| : \xi \in [x,x+h]\}\cdot|h| .$$ $\square$

Mit Hilfe von Satz 5 können wir einige der Folgerungen, die wir in Analysis I aus dem MWS gezogen haben, auf vektorwertige Funktionen verallgemeinern. Dazu benötigen wir einige neue Begriffe:

Def. Eine Teilmenge $M \subset \mathbb{R}^n$ heißt

(a) wegzusammenhängend, wenn zu $x, y \in M$ eine stetige Abbildung $\gamma : [0,1] \to M$ (ein "Weg") existiert mit $\gamma(0) = x$ und $\gamma(1) = y$;

(b) ein Gebiet, wenn sie wegzusammenhängend und offen ist;

(c) sternförmig, wenn $x_0 \in M$ existiert, so daß zu jedem $y \in M$ die Verbindungsstrecke $[x_0, y]$ ganz in M enthalten ist;

(d) konvex, wenn für alle $x, y \in M$ gilt, dass $[x,y] \subset M$. D26

Bem. + Bsp.: (1) Es gelten die Implikationen

konvex $\Rightarrow$ sternförmig $\Rightarrow$ wegzusammenhängend

(2) $\mathbb{R}^2 \setminus \{(x,y) \in \mathbb{R}^2 : x \geq 0 \text{ und } y = 0\}$ ist sternförmig, aber nicht konvex.

(3) $\mathbb{R}^2 \setminus \{(0,0)\}$ ist wegzusammenhängend, aber nicht sternförmig.

Unsere erste Folgerung aus Satz 5 ist der

<u>Schrankensatz</u>: Ist $\Omega \subset \mathbb{R}^n$ ein konvexes Gebiet und $f : \Omega \to \mathbb{R}^m$ total d'bar mit $\sup\{\|Df(\xi)\| : \xi \in \Omega\} < \infty$. Dann ist f Lipschitz-, also insbesondere glm. stetig.

Bew.: $|f(x) - f(y)| \underset{\substack{\text{Satz 5,}\\ [x,y] \subset \Omega,\\ \text{da } \Omega \text{ konvex}}}{\leq} \sup\{\|Df(\xi)\| : \xi \in [x,y]\}\, |x-y|$

$\leq \sup\{\|Df(\xi)\| : \xi \in \Omega\}\, |x-y|.$

Bem.: Der Schrankensatz gilt nicht mehr, wenn wir lediglich die Sternförmigkeit von Ω voraussetzen. Bsp.: $\Omega = \mathbb{R}^2 \setminus \{(x,0) : x \geq 0\}$

$$f(x,y) = \begin{cases} 0 & \text{für } x < 0 \\ \frac{x^2}{1+x^2} & \text{für } x \geq 0 \text{ und } y > 0 \\ -\frac{x^2}{1+x^2} & \text{für } x \geq 0 \text{ und } y < 0 \end{cases}$$

Diese Funktion ist total d'bar in Ω' mit beschränkter Ableitung, aber nicht glm. stetig, denn: Für die Folgen

$u_n = (1, \frac{1}{n})$ und $v_n = (1, -\frac{1}{n})$ gilt

$\lim_{n\to\infty} |u_n - v_n| = 0$, aber $f(u_n) = \frac{1}{2} \neq -\frac{1}{2} = f(v_n)$.

Als zweite Folgerung aus Satz 5 notieren wir:

Ist $\Omega \subset \mathbb{R}^n$ ein sternförmiges Gebiet und $f: \Omega \to \mathbb{R}^m$ total d'bar mit $Df(\xi) = 0 \;\forall \xi \in \Omega$, so ist f konstant.

Bew.: $|f(x) - f(x_0)| \leq \sup\{\|Df(\xi)\| : \xi \in \Omega\}\, |x - x_0| = 0$.

Also $f(x) = f(x_0)$ für alle $x \in \Omega$.

Bem.: Diese Folgerung gilt sogar für alle Gebiete Ω, auch wenn diese nicht sternförmig sind. Ohne die Zusammenhangsvoraussetzung wird die Aussage allerdings falsch. Bsp.

$\Omega = \Omega_1 \cup \Omega_2$, $\Omega_{1,2}$ offen mit $\operatorname{dist}(\Omega_1, \Omega_2) > 0$.

$$f(x) = \begin{cases} 1 & \text{für } x \in \Omega_1 \\ -1 & \text{" } x \in \Omega_2 \end{cases}$$

Dann ist $Df \equiv 0$, aber f nicht konstant.

Für reellwertige Funktionen mehrerer Veränderlicher können (D 21)
wir den folgenden MWS zeigen, mit dem wir diesen Abschnitt abschließen und zugleich Anlauf nehmen auf die Taylorreihe für solche Funktionen.

Satz 6: Es seien $\Omega \subset \mathbb{R}^n$ offen, $x, x+h \in \Omega$ mit $[x, x+h] \subset \Omega$. Die Funktion $f : \Omega \to \mathbb{R}$ sei total diff.'bar. Dann existiert ein $\xi \in [x, x+h]$, so daß

$$f(x+h) - f(x) = \nabla f(\xi) \cdot h.$$

Beweis: Wir setzen $\varphi(t) = f(x + t \cdot h)$, so dass

$$f(x+h) - f(x) = \varphi(1) - \varphi(0) = \varphi'(\vartheta) \quad (\vartheta \in (0,1), \text{ MWS aus Ana I})$$

$$= \nabla f(x + \vartheta h) \cdot h.$$

Mit $\xi = x + \vartheta h$ folgt die Beh. $\square$