Satz (7.11)

Sei $X = \mathbb{R}^m$, $Y = \mathbb{R}^n$ und $U \subseteq X \times Y$ offen. Sei außerdem $f : U \to Y$ eine stetig differenzierbare Abbildung und $(a, b) \in U$ eine Nullstelle von f mit der Eigenschaft, dass $\partial_Y f(a, b)$ invertierbar ist. Dann gibt es Umgebungen $U' \subseteq X$ von a und $U'' \subseteq Y$ von b mit $U' \times U'' \subseteq U$ und eine stetig differenzierbare Abbildung $g : U' \to U''$, so dass die Äquivalenz

$$f(x, y) = 0_Y \;\Leftrightarrow\; y = g(x) \qquad \text{für alle } (x, y) \in U' \times U'' \quad \text{erfüllt ist.}$$

Die Ableitung der impliziten Funktion

Folgerung (7.12)

Die Funktion $g : U' \to U''$ aus dem Satz hat an der Stelle a die Ableitung

$$dg(a) \quad = \quad -\partial_Y f(a,b)^{-1} \circ \partial_X f(a,b).$$

Beweis von Folgerung 7.12

Beh.: $dg(a) = -\partial_Y f(a,b)^{-1} \circ \partial_X f(a,b)$

Betrachte die Funktion $\Phi : U' \to Z$, $x \mapsto f(x, g(x))$.

Es gilt $\Phi = f \circ g_1$ mit $g_1 : U' \to X \times Y$ geg. durch $g_1(x) = (x, g(x))$. Berechne die Ableitung von g_1 komponentenweise, erhalte $dg_1(x) = (\mathrm{id}_X, dg(x))$.

Die mehrdimensionale Kettenregel liefert

$$d\Phi(x) = d(f \circ g_1)(x) = df(g_1(x)) \circ dg_1(x) = df(x, g(x)) \circ (\mathrm{id}_X, dg(x)).$$

weise, erhalte $dg_1(x) = (id_X, dg(x))$

Die mehrdimensionale Kettenregel liefert

Die Definition von $\partial_X f$, $\partial_Y f$ ergibt für $v \in X$

$$d\Phi(x)(v) = df(x, g(x))\big((id_X, dg(x))(v)\big)$$

$$= df(x, g(x))\big(v, dg(x)(v)\big)$$

$$= \partial_X f(x, g(x))(v) + \partial_Y f(x, g(x))\big(dg(x)(v)\big)$$

$$\Rightarrow d\Phi(x) = \partial_X f(x, g(x)) + \partial_Y f(x, g(x)) \circ dg(x)$$

Auf Grund des Zusammenhangs zwischen f und g gilt

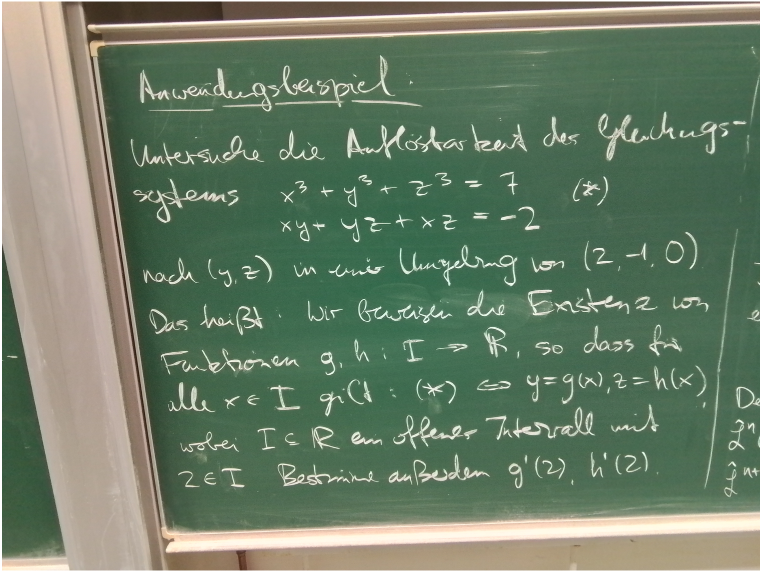
$\Phi(x) = f(x, g(x)) = 0$ für alle $x \in U_1$ $\Rightarrow$

$d\Phi(x) = 0 \ \forall x \in U_1$ Wende dies auf $(a, b) = (a, g(a))$ an

$$\Rightarrow \partial_X f(a, b) + \partial_Y f(a, b) \circ dg(a) = 0$$

$$\Rightarrow \partial_Y f(a, b) \circ dg(a) = -\partial_X f(a, b) \Rightarrow dg(a) = -\partial_Y f(a, b)^{-1} \circ \partial_X f(a, b)$$

$\square$


Anwendungsbeispiel:
Untersuche die Auflösbarkeit des Gleichungs-
systems $x^3 + y^3 + z^3 = 7$ (*)
$xy + yz + xz = -2$
nach (y, z) in einer Umgebung von $(2, -1, 0)$
Das heißt: Wir beweisen die Existenz von
Funktionen $g, h : I \to \mathbb{R}$, so dass für
alle $x \in I$ gilt: (*) $\Leftrightarrow y = g(x), z = h(x)$,
wobei $I \subseteq \mathbb{R}$ ein offenes Intervall mit
$2 \in I$. Bestimme außerdem $g'(2)$, $h'(2)$.

Setze $X = \mathbb{R}$, $Y = Z = \mathbb{R}^2$ und

$f : X \times Y \to Z$, $(x, y_1, y_2) \mapsto (x^3 + y_1^3 + y_2^3 - 7,\ xy_1 + y_1 y_2 + x y_2 + 2)$

totale Ableitung von f:

$$df(x, y_1, y_2) = \left(\underbrace{\begin{matrix} 3x^2 \\ y_1 + y_2 \end{matrix}}_{\partial_x f(x, y_1, y_2)} \quad \underbrace{\begin{matrix} 3y_1^2 & 3y_2^2 \\ x + y_2 & x + y_1 \end{matrix}}_{\partial_y f(x, y_1, y_2)}\right)$$

$$\Rightarrow df(2, -1, 0) = \begin{pmatrix} 12 & 3 & 0 \\ -1 & 2 & 1 \end{pmatrix}, \quad \partial_x f(2, -1, 0) = \begin{pmatrix} 12 \\ -1 \end{pmatrix}$$

und $\partial_y f(2, -1, 0) = \begin{pmatrix} 3 & 0 \\ 2 & 1 \end{pmatrix}$

$\det \partial_y f(2, -1, 0) = 3 \neq 0$, $\partial_y f(2, -1, 0)^{-1} = \frac{1}{3} \begin{pmatrix} 1 & 0 \\ -2 & 3 \end{pmatrix}$

Weil $\partial_y f(2,-1,0)$ invertierbar ist, existieren ein Intervall I und g, h wie angeg.

Folgerung 7.12 liefert

$$\begin{pmatrix} g'(2) \\ h'(2) \end{pmatrix} = -\frac{1}{3}\begin{pmatrix} 1 & 0 \\ -2 & 3 \end{pmatrix}\begin{pmatrix} 12 \\ -1 \end{pmatrix} = \begin{pmatrix} -4 \\ 9 \end{pmatrix}$$

$\Rightarrow g'(2) = -4, \; h'(2) = 9$

Notation:

- V, W endlich-dimensionale, normierte $\mathbb{R}$-Vektorräume
- $U \subseteq V$ offen, $f : U \to W$ diff'bare Funktion

Dann ist auch $\mathscr{L}(V, W)$ ein endlich-dimensionaler $\mathbb{R}$-Vektorraum und die Ableitungsfunktion von f eine Abbildung

$$df : U \longrightarrow \mathscr{L}(V, W) \quad , \quad x \mapsto df(x).$$

Definition der zweimal differenzierbaren Funktionen

Definition (8.1)

Ist die Abbildung df auf U stetig, dann bezeichnen wir f als stetig differenzierbare Funktion. Ist sie darüber hinaus in jedem Punkt $x \in U$ differenzierbar, dann sprechen wir von einer zweimal differenzierbaren Funktion.

geg.: zweimal diff'bare Funktion $f : U \to W$

totale Ableitung: $df : U \to \mathcal{L}(V, W)$

nochmalige Ableitung ergibt eine Funktion

$$d^2 f = d(df) : U \to \mathcal{L}(V, \mathcal{L}(V, W))$$

Ist auch diese Funktion differenzierbar, dann erhalten wir

$$d^3 f = d(d^2 f) : U \to \mathcal{L}(V, \mathcal{L}(V, \mathcal{L}(V, W)))$$

Definiert man rekursiv $\mathbb{R}$-Vektorräume $\hat{\mathcal{L}}^n(V, W)$ durch $\hat{\mathcal{L}}^1(V, W) = \mathcal{L}(V, W)$ und $\hat{\mathcal{L}}^{n+1}(V, W) = \mathcal{L}(V, \hat{\mathcal{L}}^n(V, W))$, dann ist die n-te totale Ableitung von f also eine Funktion

$$df : U \to \hat{\mathcal{L}}^n(V, W).$$

An Stelle von $\hat{\mathcal{L}}^n(V, W)$ arbeitet man aus praktischen Gründen mit $\mathcal{L}^n(V, W)$, dem Raum der multilinearen Abb.

$$\phi : \underbrace{V \times \dots \times V}_{n\text{-mal}} \to W.$$

Eine Hilfsaussage aus der Linearen Algebra

Satz (8.2)

Jedem Element $\hat{\phi} \in \hat{\mathscr{L}}^n(V, W)$ kann durch die Definition

$$\phi(v_1, ..., v_n) \quad = \quad \hat{\phi}(v_1)(v_2) \dots (v_n)$$

ein Element in $\mathscr{L}^n(V, W)$ zugeordnet werden, und die Abbildung

$$\Phi : \hat{\mathscr{L}}^n(V, W) \to \mathscr{L}^n(V, W) \quad , \quad \hat{\phi} \mapsto \phi$$

ist ein Isomorphismus von $\mathbb{R}$-Vektorräumen.

Proposition (8.3)

Sei $U \subseteq V$ offen, $p \in U$ und $f : U \to \mathbb{R}$ eine n-mal in p differenzierbare Funktion. Dann gilt für alle $v_1, ..., v_n \in V$ jeweils

$$d^n f(p)(v_1, ..., v_n) \quad = \quad \partial_{v_1} \cdots \partial_{v_n} f(p).$$

Beispiel für eine zweite totale Ableitung:

$f: \mathbb{R}^2 \to \mathbb{R}^2, \ (x,y) \mapsto (x^2+y^2, 2xy)$

Fixiere einen Punkt $p = (x_0, y_0)$.

erste totale Ableitung: $df(p): \mathbb{R}^2 \to \mathbb{R}^2$

zweite totale Ableitung: $d^2 f(p): \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}^2$

bereits bekannt: Die Ableitung df ist in jedem Punkt geg. durch die Jacobimatrix, d.h.

$$df(x,y) = \begin{pmatrix} \partial_1 f_1(x,y) & \partial_2 f_1(x,y) \\ \partial_1 f_2(x,y) & \partial_2 f_2(x,y) \end{pmatrix} = \begin{pmatrix} 2x & 2y \\ 2y & 2x \end{pmatrix}$$

Betrachte die Basis $B = (B_{11}, B_{12}, B_{21}, B_{22})$ von $M_{2,\mathbb{R}}$

gegeben durch Basismatrizen. Dann ist $\Phi_{\mathcal{B}} \circ df : \mathbb{R}^2 \to \mathbb{R}^4$

geg. durch $(\Phi_{\mathcal{B}} \circ df)(x,y) = \begin{pmatrix} 2x \\ 2y \\ 2y \\ 2x \end{pmatrix}$

$$\Rightarrow d(\Phi_{\mathcal{B}} \circ df)(p) = \begin{pmatrix} 2 & 0 \\ 0 & 2 \\ 0 & 2 \\ 2 & 0 \end{pmatrix}$$

$$\Rightarrow d(\Phi_{\mathcal{B}} \circ df)(p)(e_1) = \begin{pmatrix} 2 \\ 0 \\ 0 \\ 2 \end{pmatrix}, \quad d(\Phi_{\mathcal{B}} \circ df)(p)(e_2) = \begin{pmatrix} 0 \\ 2 \\ 2 \\ 0 \end{pmatrix}$$

Weil $\Phi_{\mathcal{B}}$ eine lineare Abbildung ist, liefert die Kettenregel

$$d(\Phi_{\mathcal{B}} \circ df)(p) = d\Phi_{\mathcal{B}}(df(p)) \circ d^2 f(p) = \Phi_{\mathcal{B}} \circ d^2 f(p)$$

$$\Rightarrow (\Phi_{\mathcal{B}} \circ d^2 f)(p)(e_1) = \begin{pmatrix} 2 \\ 0 \\ 0 \\ 2 \end{pmatrix}, \quad (\Phi_{\mathcal{B}} \circ d^2 f)(p)(e_2) = \begin{pmatrix} 0 \\ 2 \\ 2 \\ 0 \end{pmatrix}$$

$\Rightarrow d^2 f(p)(e_1) = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$, $d^2 f(p)(e_2) = \begin{pmatrix} 0 & 2 \\ 2 & 0 \end{pmatrix}$

Für jedes $v \in \mathbb{R}^2$ kann $A_v = d^2 f(p)(v)$ als lineare Abbildung betrachtet werden, und zwar durch $w \mapsto A_v w$.

$\Rightarrow$ Als Elemente von $\mathbb{R}^2 \to \mathbb{R}^2$ sind $d^2 f(p)(e_1)$, $d^2 f(p)(e_2)$ geg. durch

$$d^2 f(p)(e_1)(w) = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} w$$

$$d^2 f(p)(e_2)(w) = \begin{pmatrix} 0 & 2 \\ 0 & 2 \end{pmatrix} w.$$

Betrachten wir $d^2f(p)$ nun als bilineare Abbildung $\mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}^2$, dann erhalten wir

$$d^2f(p)(e_1, w) = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} w$$

$$d^2f(p)(e_2, w) = \begin{pmatrix} 0 & 2 \\ 2 & 0 \end{pmatrix} w.$$

insbes.:

$$d^2f(p)(e_1, e_1) = \begin{pmatrix} 2 \\ 0 \end{pmatrix}, \quad d^2f(p)(e_1, e_2) = \begin{pmatrix} 0 \\ 2 \end{pmatrix}.$$

$$d^2f(p)(e_2, e_1) = \begin{pmatrix} 0 \\ 2 \end{pmatrix}, \quad d^2f(p)(e_2, e_2) = \begin{pmatrix} 2 \\ 0 \end{pmatrix}$$

Die beiden Komponenten dieser bilinearen Abbildung sind Bilinearformen, mit den Darstellungsmatrizen $\begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$ bzw. $\begin{pmatrix} 0 & 2 \\ 2 & 0 \end{pmatrix}$.

Definition der Hessematrix einer Funktion

Definition (8.4)

Sei $U \subseteq \mathbb{R}^n$ offen, $p \in \mathbb{R}^n$ und $f : \mathbb{R}^n \to \mathbb{R}$ eine in p mindestens zweimal differenzierbare Funktion. Sei $\mathscr{E} = (e_1, ..., e_n)$ die Einheitsbasis des $\mathbb{R}^n$. Dann wird die Darstellungsmatrix $\mathscr{H}(f)(p) = \mathscr{M}_{\mathscr{E}}(d^2 f(p))$ mit den Einträgen

$$a_{ij} \quad = \quad d^2 f(p)(e_i, e_j) \quad = \quad \partial_{ij} f(p)$$

die Hesse-Matrix von f an der Stelle p genannt.

Eindimensionale Taylorpolynome

Erinnerung:
Sei $p \in \mathbb{N}_0$, sei $I \subseteq \mathbb{R}$ ein offenes Intervall, $a \in I$ und $f : I \to \mathbb{R}$ eine in a mindestens m-mal differenzierbare Funktion. Dann ist das m-te Taylorpolynom

$$\tau_m(f,a)(x) \quad = \quad \sum_{k=0}^{m} \frac{1}{k!} f^{(k)}(a)(x-a)^k.$$

von f an der Stelle a definiert.

Satz (8.5)

Sei $p \in \mathbb{N}$, $I \subseteq \mathbb{R}$ ein offenes Intervall und $f : I \to \mathbb{R}$ eine mindestens m-mal differenzierbare Funktion. Seien $a, x \in I$ mit $x > a$. Dann gibt es einen Punkt $\xi \in]a, x[$ mit

$$f(x) \quad = \quad \tau_{m-1}(f,a)(x) + \frac{1}{p!} f^{(m)}(\xi)(x-a)^m.$$

Definition der höherdimensionalen Taylorpolynome

Definition (8.6)

Sei $n \in \mathbb{N}$, $U \subseteq V$ offen, $p \in U$ und $f : U \to \mathbb{R}$ eine im Punkt p mindestens m-mal differenzierbare Funktion. Dann bezeichnet man die Funktion für $x \in V$ gegeben

$$\tau_m(f,p)(x) \quad = \quad f(p) + \sum_{k=1}^{m} \frac{1}{k!} d^k f(p)(\underbrace{x-p, \ldots, x-p}_{k\text{-mal}})$$

als Taylorpolynom m-ten Grades von f an der Stelle p.

wichtige Spezialfälle:

Taylorpolynome ersten und zweiten Grades

$$\tau_1(f,p)(x) = f(p) + df(p)(x-p)$$

$$\tau_2(f,p)(x) = f(p) + df(p)(x-p) + \frac{1}{2} d^2 f(p)(x-p, x-p)$$

$$= f(p) + df(p)(x-p) + \frac{1}{2}\,{}^t(x-p)\,\mathcal{H}(f,p)\,(x-p)$$

<u>Beispiel</u>: zweite Taylorpolynom der Funktion

$$f(x,y) = x^2 + 3y^2 - 7xy + 5x - 3y + 8$$

an der Stelle $(3,5)$

$$\partial_1 f(x,y) = 2x - 7y + 5$$

$$\partial_2 f(x,y) = -7x + 6y - 3$$

$$\partial_{11} f(x,y) = 2, \quad \partial_{22} f(x,y) = 6$$

$$\partial_{12} f(x,y) = \partial_{21} f(x,y) = -7$$

$$f(3,5) = -13, \quad df(3,5) = \begin{pmatrix} -24 & 6 \end{pmatrix}$$

$$\mathcal{H}(f,(3,5)) = \begin{pmatrix} 2 & -7 \\ -7 & 6 \end{pmatrix} \qquad (x,y) - p = (x-3, y-5)$$

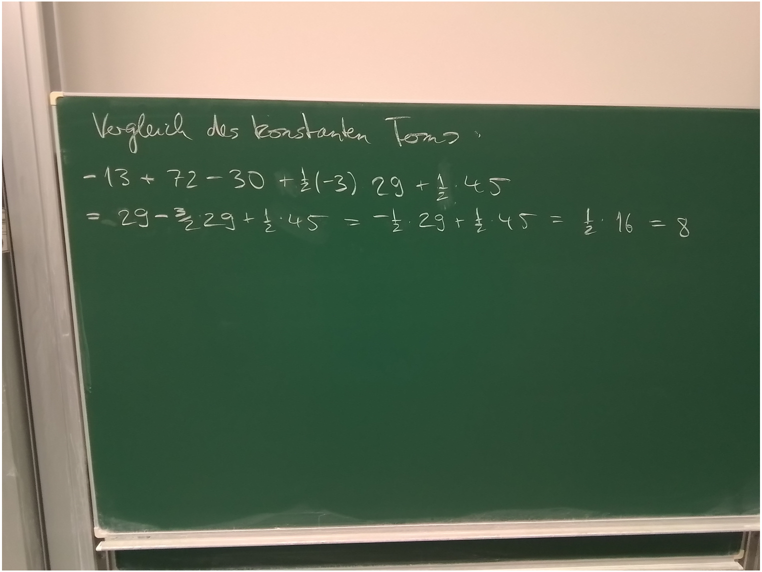
$$\Rightarrow T_2(f,(3,5))(x,y) = -13 + \begin{pmatrix} -24 & 6 \end{pmatrix} \begin{pmatrix} x-3 \\ y-5 \end{pmatrix}$$

$$+ \frac{1}{2} \cdot \begin{pmatrix} x-3 & y-5 \end{pmatrix} \begin{pmatrix} 2 & -7 \\ -7 & 6 \end{pmatrix} \begin{pmatrix} x-3 \\ y-5 \end{pmatrix} =$$

$$-13 + (-24)(x-3) + 6 \cdot (y-5) + \frac{1}{2} \begin{pmatrix} x-3 & y-5 \end{pmatrix} \begin{pmatrix} 2x-7y+29 \\ -7x+6y-9 \end{pmatrix}$$

$$= -13 - 24x + 72 + 6y - 30 + \frac{1}{2}(x-3)(2x-7y+29)$$

$$+ \frac{1}{2}(y-5)(-7x+6y-9) = x^2 + 3y^2 - 7xy + 5x - 3y + 8$$


Vergleich des konstanten Terms:
$-13 + 72 - 30 + \frac{1}{2}(-3)\, 29 + \frac{1}{2} \cdot 45$
$= 29 - \frac{3}{2}\, 29 + \frac{1}{2} \cdot 45 = -\frac{1}{2} \cdot 29 + \frac{1}{2} \cdot 45 = \frac{1}{2} \cdot 16 = 8$

Die Approximationseigenschaft der Taylorpolynome

Satz (8.7)

Sei $U \subseteq \mathbb{R}^n$ offen und $f : U \to \mathbb{R}$ eine m-mal stetig differenzierbare Funktion. Dann gibt es zu jedem Punkt $p \in U$ eine Umgebung U_0 des Nullpunkts und eine Funktion $\psi : U_0 \to \mathbb{R}$ mit

$$f(p+h) \quad = \quad \tau_m(f,p)(p+h) + \psi(h)$$

für alle $h \in U_0$ und

$$\lim_{h \to 0} \frac{\psi(h)}{\|h\|^m} = 0.$$