E. ABLEITUNGEN IN MEHREREN DIMENSIONEN

In diesem Abschnitt, seien V und W Banachräume mit $\dim V = m < \infty$, sei $G \subset V$ offen und sei $f\colon G \to W$ eine Abbildung.

Die Ableitung von f in einem Punkt $p \in G$ ist jetzt kein Vektor in W sondern eine lineare Abblidung $L\colon V \to W$. Weil dies ein bisschen kompliziert wird, fangen wir mit etwas einfacherem an.

E1. Richtungsableitungen

Sei $I \subset \mathbb{R}$ ein Intervall und sei $\gamma\colon I \to G \subset V$ eine Kurve in G. Dann ist $f \circ \gamma$ eine Abbildung $I \to W$ und wir können ihre Ableitungen (im Sinne von §D1) untersuchen.

Definition E1.1. Sei $s \in I$ und sei $p := \gamma(s) \in G$. Die *Ableitung von f in p entlang der Kurve γ* ist

$$(f \circ \gamma)'(s) = \frac{d}{dt}\Big|_{t=s} f(\gamma(t)) = \lim_{h\to 0} \frac{f(\gamma(s+h)) - f(\gamma(s))}{h} \in W.$$

Der wichtigste Fall ist der, wo γ eine Gerade durch p in Richtung $v \in V$ ist, d.h., $\gamma(t) := p + tv$.

Definition E1.2. Seien $p \in G$ und $v \in V$. Die Ableitung entlang der Geraden $\gamma(t) := p+tv$ heißt *Richtungsableitung* von f im Punkt p in Richtung v:

$$\partial_v f(p) := (f \circ \gamma)'(0) = \frac{d}{dt}\Big|_{t=0} f(p+tv) = \lim_{t\to 0} \frac{f(p+tv) - f(p)}{t}.$$

Beispiel E1.3. Sei $f\colon V \to W$ linear. Für alle $p, v \in V$ gilt dann $\partial_v f(p) = f(v)$.

Weiter sei $\gamma\colon I \to G$ eine differenzierbare Kurve mit $\gamma(s) = p$. Die Ableitung von f in p entlang γ können wir wie folgt berechnen, indem wir zunächst die Linearität und danach die Stetigkeit von f benutzen:

$$\begin{aligned}(f \circ \gamma)'(s) &= \lim_{h\to 0} \frac{f(\gamma(s+h)) - f(\gamma(s))}{h} \\ &= \lim_{h\to 0} f\Big(\tfrac{1}{h}(\gamma(s+h) - \gamma(s))\Big) \\ &= f\Big(\lim_{h\to 0} \tfrac{1}{h}(\gamma(s+h) - \gamma(s))\Big) \\ &= f(\gamma'(s)).\end{aligned}$$

Definition E1.4. Eine Abbildung $f\colon V \to W$ heißt *homogen*, falls $f(\lambda v) = \lambda f(v)$ für alle $\lambda \in \mathbb{R}$ und $v \in V$.

Beispiel E1.5. Sei $S := \{v : \|v\| = 1\}$ die Einheitssphäre bezüglich der Norm auf V. Jede Funktion $f\colon S \to W$ mit $f(-x) = -f(x)$ (auch z.B. wenn nirgendwo stetig) hat eine eindeutige Erweiterung zu einer homogenen Abbildung $V \to W$. Diese ist in 0 stetig (aber möglicherweise in keinem anderen Punkt).

Beispiel E1.6. Auf $\mathbb{R}^2$ ist die Funktion

$$f = r\cos 3\theta = \frac{x^3 - 3xy^2}{x^2 + y^2}$$

homogen. Hier existiert $\partial_v f(p)$ für alle p und v.

Lemma E1.7. *Sei $f\colon V \to W$ homogen. Für jedes $v \in V$ gilt dann $\partial_v f(0) = f(v)$.*

Beweis. Es gilt $f(tv) = tf(v)$, insbes. $f(0) = 0$. Damit gilt

$$\partial_v f(0) = \lim_{t\to 0} \tfrac{1}{t}(f(tv) - f(0)) = \lim_{t\to 0} \tfrac{1}{t} tf(v) = f(v). \qquad \square$$

Bemerkung E1.8. Entlang anderen Kurven γ durch 0 muss die Ableitung von einem homogenen f nicht existieren.

Beispiel E1.9. Sei $P := \{(x,y) \in \mathbb{R}^2 : y = x^2 \neq 0\}$ und sei $f = \chi_P : \mathbb{R}^2 \to \mathbb{R}$ die charakteristische Funktion von P, d.h., $f(p) = 1$ für $p \in P$ und $f(p) = 0$ für $p \notin P$. Nun sei $p := (0,0)$. Auf jeder Gerade durch p ist f lokal konstant um p. Deswegen gilt $\partial_v f(p) = 0$ für alle $v \in \mathbb{R}^2$. Hingegen ist f in p unstetig. Entlang der Kurve $\gamma(t) := (t, t^2)$ hat f in p keine Ableitung.

Diese Beispiele dienen um zu zeigen, dass auch die Existenz aller Richtungsableitungen in einem Punkt nicht besonders aussagekräftig ist. Für eine differenzierbare Funktion (definiert unten) sind aber die Richtungsableitungen nützlich, um andere Ableitungen zu berechnen. Im Falle $V = \mathbb{R}^m$ sind die Richtungsableitungen in den Richtungen e_i aus der Standardbasis von besonderem Interesse.

Definition E1.10. Sei $G \subset \mathbb{R}^n$ offen, sei $f\colon G \to W$ eine Abbildung und sei $p \in G$. Die *partielle Ableitungen* von f in p sind

$$\frac{\partial}{\partial x_i} f(p) = \partial_i f(p) := \partial_{e_i} f(p), \quad i = 1, \dots, n.$$

Es gilt also

$$\partial_i f(p) = \frac{d}{dt}\Big|_{t=0} f(p_1, \dots, p_i + t, \dots, p_n).$$

Bemerkung E1.11. Um die partielle Ableitung $\partial_i f(p)$ zu berechnen, nimmt man die Ableitung von f bezüglich der i-ten Variablen x_i (im Sinne von §I.E1 bzw. von §D1); dabei betrachtet man alle anderen Variablen x_j ($j \neq i$) als Konstanten $x_j = p_j$.

Ende der Vorlesung 2009 Januar 12

E2. Ableitungen als lineare Abbildungen

Sei $G \subset V$ offen und sei $f\colon G \to W$ eine Funktion mehrerer Variabler. Wie können wir sinnvoll eine Ableitung von f definieren? Ein Differenzenquotient macht keinen Sinn, weil $h = x - p$ ein Vektor wäre: wir können nicht durch einen Vektor teilen.

Die Ableitung einer Funktion $g\colon I \to W$ versteht man normalerweise als die (Änderungs-)Geschwindigkeit der Werte von g. Dies macht auch wenig Sinn für $f\colon G \to W$. Wir sollten uns eher die Ableitung als „beste lineare Approximation“ vorstellen.

Man kann Lemma D1.6 leicht wie folgt umschreiben:

Lemma E2.1. *Eine Funktion $g\colon I \to W$ hat genau dann eine Ableitung $g'(p)$ in $p \in I$, wenn*

$$g(x) - g(p) = g'(p)(x-p) + R(x), \quad 0 = \lim_{x\to p} \frac{R(x)}{|x-p|}.$$

Hier ist R natürlich das Restglied des ersten Taylorpolynoms von g und dieses Kriterium ist der Fall $n = 1$ der Taylorapproximation (Satz I.F6.1).

Jede lineare Abbildung $L\colon \mathbb{R} \to W$ hat die Form $L(t) = wt$ für einen konstanten Vektor $w \in W$. Das letzte Lemma können wir wie folgt verstehen: $g'(p)(x-p)$ ist eine lineare Funktion von $x-p$, die beste lineare Approximation zur Differenz $g(x)-g(p)$.

In dieser Form können wir die Ableitung auch für Funktionen meherer Variabler definieren:

Definition E2.2. Sei $f\colon G \to W$ eine Abbildung und sei $p \in G$. Falls es eine lineare Abbildung $L\colon V \to W$ so gibt, dass

$$0 = \lim_{x\to p} \frac{R(x)}{\|x-p\|} \quad \text{für das durch} \quad f(x)-f(p)-L(x-p) =: R(x)$$

definierte Restglied $R\colon G \to W$, dann sagen wir, f ist in p *differenzierbar* und $D_p f := L$ ist die *Ableitung* von f in p.

Bemerkung E2.3. Äquivalent ist folgende Umschreibung: Differenzierbarkeit in x heißt,

$$f(x+h) = f(x) + D_x f(h) + \|h\|\, r(h),$$

wobei $r(h) \to 0$ für $h \to 0$. (Hier ist r zunächst auf einer punktierten Umgebung von $h = 0$ definiert.)

Bemerkung E2.4. Dass die Ableitung $D_p f$ eindeutig bestimmt ist (wenn sie existiert), folgt aus dem nächsten Lemma.

Lemma E2.5. *Sei f in p differenzierbar und sei $v \in V$. Dann existiert für jedes $v \in V$ die Richtungsableitung $\partial_v f(p)$ und es gilt*

$$D_p f(v) = \partial_v f(p).$$

Beweis. Es gilt

$$\begin{aligned}\frac{f(p+tv)-f(p)}{t} &= \frac{f(p) + D_p f(tv) + R(p+tv) - f(p)}{t} \\ &= D_p f(v) + \frac{R(p+tv)}{t} \\ &= D_p f(v) \pm \frac{R(p+tv)}{\|tv\|}\|v\|\end{aligned}$$

und damit gilt im Limes $t \to 0$ $\partial_v f(p) = D_p f(v) + 0$. □

Beispiel E2.6. Die Funktion f aus Beispiel E1.6 ist im Ursprung nicht differenzierbar, weil die Richtungsableitungen $\partial_v f(0) = f(v)$ nicht linear in v sind.

Korollar E2.7. *Im Falle $V = \mathbb{R}^n$ ist die Ableitung $D_p f$ (wenn sie existiert) durch die partiellen Ableitungen $\partial_i f(p)$ bestimmt: für $v = (v_1, \dots, v_n)$ gilt*

$$D_p f(v) = \sum_{i=1}^{n} \partial_i f(p)\, v_i.$$

Beweis. Es gilt $v = \sum v_i e_i$ und damit aus der Linearität von $D_p f$ gilt

$$D_p f(v) = \sum v_i D_p f(e_i) = \sum v_i \partial_{e_i} f(p).$$ □

Bemerkung E2.8. Weitere Namen für die partiellen Ablietungen sind z.B.

$$f_{x_i} := \frac{\partial}{\partial x_i} f := \partial_{e_i} f.$$

Nennt man die Koordinaten auf $\mathbb{R}^3$ z.B. x, y, z, dann schreibt man auch etwa

$$f_y := \frac{\partial}{\partial y} f := \partial_y f := \partial_{e_2} f.$$

Lemma E2.9. *Ist $f\colon G \to W$ in $p \in G$ differenzierbar, dann ist f in p stetig.*

Beweis. Wir haben

$$f(x) = f(p) + D_p f(x-p) + R(x).$$

Hier ist R sicherlich in p stetig. ($R(x) \to 0$ für $x \to p$, sogar schneller als $\|x-p\|$.) Weil V endlich-dimensional ist, ist auch die lineare Abbildung $D_p f$ stetig. □

Lemma E2.10. *Sei $W = \mathbb{R}^n$ und sei*

$$f = (f_1, \dots, f_n)\colon G \to \mathbb{R}^n.$$

Diese Abbildung ist genau dann in $p \in G$ differenzierbar, wenn jedes f_i in p differenzierbar ist. In diesem Fall gilt

$$D_p f = (D_p f_1, \dots, D_p f_n) : V \to \mathbb{R}^n.$$

Beweis. (Aufgabe.) □

Bemerkung E2.11. Ähnliches gilt für $W = W_1 \times W_2$ (z.B. mit der Norm $\|(w_1, w_2)\| = \|w_1\| + \|w_2\|$) und $f = (f_1, f_2)$. Auch dann haben wir $D_p f = (D_p f_1, D_p f_2)$.

E3. Rechenregeln

Satz E3.1. *Seien $f, g\colon G \to W$ differenzierbar in p und sei $\lambda \in \mathbb{R}$. Dann ist auch $f + \lambda g$ differenzierbar in p und es gilt $D_p(f + \lambda g) = D_p f + \lambda D_p g$.*

Beweis. (Aufgabe.) □

Satz E3.2 (Kettenregel). *Seien U, V, W Banachräume (U, V endlich-dimensional). Seien $G \subset U$ und $H \subset V$ offen. Seien $f : G \to V$ und $g : H \to W$ Abbildungen mit $f(G) \subset H$, damit die Komposition $g \circ f : G \to W$ existiert. Seien $x \in G$ und $y := f(x) \in H$. Ist f bzw. g differenzierbar in x bzw. y, dann ist $g \circ f$ in x differenzierbar, mit*

$$D_x(g \circ f) = D_y g \circ D_x f.$$

Bemerkung E3.3. Das heißt, die Ableitung der Komposition ist die Komposition der Ableitungen. Dies sollte nicht überraschend sein, wenn man die Ableitungen als die besten linearen Approximationen betrachtet.

Beweis. Mit Bemerkung E2.3 haben wir

$$f(x+h) = f(x) + D_x f(h) + \|h\| r(h),$$
$$g(y+k) = g(y) + D_y g(k) + \|k\| s(k).$$

Wir definieren

$$k(h) := f(x+h) - y = D_x f(h) + \|h\| r(h).$$

Damit gilt

$$\begin{aligned} g(f(x+h)) &= g(y) + D_y g(k(h)) + \|k(h)\|\, s(k(h)) \\ &= g(f(x)) + D_y g(D_x f(h)) + \\ &\qquad \|h\|\, D_y g(r(h)) + \|k(h)\|\, s(k(h)) \\ &= (g \circ f)(x) + (D_y g \circ D_x f)(h) + \|h\|\, u(h), \end{aligned}$$

wobei $u(h) := D_y g(r(h)) + \|k(h)\|\, s(k(h))/\|h\|$. Wir müssen zeigen, $u(h) \to 0$ für $h \to 0$.

Der erste Summand ist einfach: weil $D_y g$ stetig und linear ist, gilt

$$\lim_{h\to 0} D_y g(r(h)) = D_y g\Big(\lim_{h\to 0} r(h)\Big) = D_y g(0) = 0.$$

Für den zweiten, sei $M := \|D_x f\|$ die Operatornorm von $D_x f$. Dann gilt

$$\|k(h)\| \le \|D_x f(h)\| + \|h\|\, \|r(h)\| \le \|h\|\, (M + \|r(h)\|).$$

Insbesondere gilt $k(h) \to 0$ für $h \to 0$ und damit $s(k(h)) \to 0$. Weil $r(h) \to 0$ gilt $r(h) \le 1$ für hinreichend kleines h und damit $\|k(h)\| \le (M+1)\|h\|$. Deshalb haben wir

$$\lim_{h\to 0} \frac{\|k(h)\|}{\|h\|} \|s(k(h))\| \le \lim_{h\to 0} (M+1)\|s(k(h))\| = 0. \qquad \square$$

Definition E3.4. Seien V, V' und W Vektorräume. Eine Abbildung $\mu\colon V \times V' \to W$ heißt dann *bilinear*, falls sie in jedem Argument linear ist, d.h., falls

$$\mu(\lambda v + w, v') = \lambda\mu(v, v') + \mu(w, v')$$

und

$$\mu(v, \lambda v' + w') = \lambda\mu(v, v') + \mu(v, w')$$

für alle $v, w \in V$, $v', w' \in V'$ und $\lambda \in \mathbb{R}$.

Beispiele E3.5. Folgende Abbildungen sind bilinear:

1. das Produkt $\mathbb{R} \times \mathbb{R} \to \mathbb{R}$, $(x, y) \mapsto xy$,
2. jedes Skalarprodukt $V \times V \to \mathbb{R}$, $(v, w) \mapsto \langle v, w\rangle$,
3. die Skalarmultiplikation $\mathbb{R} \times V \to V$, $(\lambda, v) \mapsto \lambda v$,
4. $L(V, W) \times V \to W$, $(f, v) \mapsto f(v)$.

Lemma E3.6. [Vgl. D2.6.] *Seien V und V' endlich-dimensionale Banachräume und sei W ein Banachraum. Zu jedem bilinearen $\mu\colon V \times V' \to W$ gibt es $C > 0$ so, dass*

$$\big\|\mu(v, v')\big\| \le C\|v\|\, \|v'\|$$

für alle $v \in V$, $v' \in V'$.

Beweis. Wir dürfen annehmen, $V = \mathbb{R}^m$, $V' = \mathbb{R}^n$, jeweils mit der ℓ^1-Norm. Sei $w_{ij} := \mu(e_i, e_j)$ für $i = 1, \dots, m$ und $j = 1, \dots, n$ und sei $C := \max_{i,j} \|w_{ij}\|$. Für $v = (v_1, \dots v_m)$ und $v' = (v'_1, \dots, v'_n)$ gilt dann

$$\begin{aligned} \big\|\mu(v, v')\big\| &= \Big\|\sum_{i,j} v_i v'_j w_{ij}\Big\| \le \sum_{i,j} |v_i|\, |v'_j|\, \|w_{ij}\| \\ &\le C \sum_i |v_i| \sum_j |v'_j| = C\|v\|_1\, \|v'\|_1. \end{aligned} \qquad \square$$

Satz E3.7 (Allgemeine Produktregel). *Seien V und V' endlich-dimensionale Banachräume, sei W ein Banachraum und sei $\mu\colon V \times V' \to W$ bilinear. Dann ist μ differenzierbar und es gilt*

$$D_{(p,p')}\mu(v, v') = \mu(p, v') + \mu(v, p').$$

Bemerkung E3.8. Allgemeiner: eine *multilineare* Abbildung $\mu\colon V_1 \times \cdots \times V_k \to W$ ist differenzierbar mit

$$D_p\mu(v) = \sum_{i=1}^{k} \mu(p_1, \dots, p_{i-1}, v_i, p_{i+1}, \dots, p_k).$$

Beweis. (Aufgabe.) $\square$

Ende der Vorlesung 2009 Januar 15

In konkreten Fällen benutzt man die allgemeine Produktregel (z.B. für eine der bilinearen Funktionen aus E3.5 zusammen mit der Kettenregel. Zu E3.5.3 haben wir:

Korollar E3.9. *Seien V und W endlich-dimensionale Banachräume, sei $G \subset V$ offen und seien $f\colon G \to \mathbb{R}$ und $g\colon G \to W$ differenzierbare Abbildungen. Dann ist $fg\colon G \to W$ auch differenzierbar; in $p \in G$ gilt*

$$D_p(fg)(v) = f(p)\, D_p g(v) + D_p f(v)\, g(p).$$

Beweis. Wir schreiben $fg = \mu \circ h$, wobei $\mu\colon \mathbb{R} \times W \to W$ die Skalarmultiplikation ist und $h := (f, g)\colon G \to \mathbb{R} \times W$. Für $p \in G$ seien $t := f(p)$ und $q := g(p)$, sodass $h(p) = (t, q)$. Nach Bemerkung E2.11 gilt $D_P h = (D_p f, D_p g)$; nach der allgemeinen Produktregel gilt

$$D_{(t,q)}\mu(a, w) = \mu(t, w) + \mu(a, q) = tw + aq.$$

Mit der Kettenregel gilt

$$D_p(fg) = D_p(\mu \circ h) = D_{(t,q)}\mu \circ D_p h.$$

Das heißt dann,

$$\begin{aligned} D_p(fg)(v) &= D_{(t,q)}\mu\big(D_p f(v), D_p g(v)\big) \\ &= tD_p g(v) + D_p f(v) q \\ &= f(p)\, D_p g(v) + D_p f(v)\, g(p). \end{aligned} \qquad \square$$

Beispiel E3.10. Sei V ein endlich-dimensionaler Vektorraum mit Skalarprodukt $\langle\cdot,\cdot\rangle$. (Ein Skalarprodukt heißt auch inneres Produkt.) Die Norm $\|v\| := \sqrt{\langle v, v\rangle}$ ist eine Abbildung $r : V \to \mathbb{R}$, $r(v) := \|v\|$. In 0 ist sie nicht differenzierbar. (Warum?) Wir behaupten, sie ist in jedem Punkt $p \neq 0$ differenzierbar mit $D_p r(v) = {}^{\langle v, p\rangle}/_{\|p\|}$. Definieren wir

$$\delta\colon v \mapsto (v, v), \quad \mu\colon (v, w) \mapsto \langle v, w\rangle, \quad \sigma\colon x \mapsto \sqrt{x},$$

gilt dann

$$r = \sigma \circ \mu \circ \delta\colon V \to V \times V \to I \to \mathbb{R},$$

wobei $I := [0, \infty) \subset \mathbb{R}$. Weil δ linear ist und μ bilinear, gelten

$$\begin{aligned} D_p\delta(v) &= \delta(v) = (v, v), \\ D_{(p,p)}\mu(v, v) &= \mu(p, v) + \mu(v, p) = 2\langle v, p\rangle, \\ D_x\sigma(h) &= \frac{1}{2\sqrt{x}}h. \end{aligned}$$

Deshalb gilt nach der Kettenregel

$$\begin{aligned} D_p r(v) &= D_{\langle p,p\rangle}\sigma\Big(D_{(p,p)}\mu(D_p\delta(v))\Big) \\ &= \frac{1}{2\sqrt{\langle p, p\rangle}} 2\langle v, p\rangle = \frac{\langle v, p\rangle}{\|p\|}. \end{aligned}$$

Bemerkung E3.11. Eine Norm $\|\cdot\|$, die nicht aus einem Skalarprodukt kommt, ist natürlich stetig, muss aber nicht differenzierbar sein. Die ℓ^1-Norm ist z.B. in e_i nicht differenzierbar.

Ein wichtiger Fall der Kettenregel sagt uns, dass Ableitungen entlang einer differenzierbaren Kurve einfach Richtungsableitungen sind.

Lemma E3.12. *Seien $\gamma\colon I \to G$ und $f\colon G \to W$ differenzierbar und sei $t \in I$. Die Ableitung von f in $p := \gamma(t)$ entlang der Kurve γ ist*

$$(f \circ \gamma)'(t) = \partial_{\gamma'(t)} f(p).$$

Beweis. Es gilt

$$\begin{aligned} (f \circ \gamma)'(t) &= D_t(f \circ \gamma)(1) = D_p f \circ D_t\gamma(1) \\ &= D_p f(\gamma'(t)) = \partial_{\gamma'(t)} f(p). \end{aligned}$$ □

Bemerkung E3.13. Ist f in p differenzierbar, hängt nach diesem Lemma die Ableitung entlang einer Kurve γ nur von dessen Tangentenvektor ab. Ist f hingegen in p nicht differenzierbar, kann es sein, dass die Ableitungen entlang verschiedenen Kurven mit demselben Tangentenvektor unterschiedlich sind bzw. teilweise nicht existieren. [Vgl. E1.9.]

E4. Existenz der Ableitung

Wollen wir zeigen, eine Abbildung $f\colon G \to W$ ist in $p \in G$ differenzierbar, können wir bisher nur die Definition „per Hand“ anwenden.

Wir kennen Beispiele, wo alle Richtungsableitungen existieren, $D_p f$ aber nicht (Beispiel E2.6). Auch wenn die verschiedenen Richtungsableitungen $\partial_v f(p)$ in p linear in v sind (z.B. wenn alle Null sind wie im Beispiel E1.9) muss $D_p f$ nicht existieren, muss f nicht mal in p stetig sein.

Es ist deshalb schön, ein Kriterium für Differenzierbarkeit zu finden, das nur von den Richtungs- bzw. partiellen Ableitungen abhängt, deren Stetigkeit aber verlangt. Dabei nehmen wir an, dass nicht nur V sondern auch W endlich-dimensional ist.

Satz E4.1. *Sei $U \subset \mathbb{R}^m$ offen und sei $f\colon U \to \mathbb{R}$ eine Funktion. Existieren in jedem Punkt $x \in U$ die m partiellen Ableitungen $\partial_i f(x)$, betrachten wir diese als Funktionen $\partial_i f\colon U \to \mathbb{R}$. Sind diese auf U stetig, dann ist f auf U stetig differenzierbar.*

Beweis. Sei $x = (x_1, \dots, x_m) \in U$. Aus Lemma E2.5 wissen wir, welche lineare Abbildung die Ableitung $D_x f$ sein muss, falls sie existiert. Für $h = \sum h_i e_i \in \mathbb{R}^m$ gälte nämlich

$$D_x f(h) = \sum h_i D_x f(e_i) = \sum h_i \partial_i f(x).$$

Diese lineare Abbildung $h \mapsto \sum h_i \partial_i f(x)$ nennen wir L. Wir müssen nur zeigen, das Restglied $f(x+h) - f(x) - Lh$ wächst (wie in der Definition von Differenzierbarkeit) sublinear.

Weil U offen ist, gibt es $r > 0$ so, dass $B_r(x) \subset U$. Für $\|h\| < r$ definieren wir wie folgt $m + 1$ Punkte $y_0, \dots, y_m$:

$$y_0 := x,\ y_1 := y_0 + h_1 e_1,\ \dots, y_k := y_{k-1} + h_k e_k,\ \dots,\ y_m = x + h.$$

Wir haben

$$f(x + h) - f(x) = \sum_{k=1}^{m} f(y_k) - f(y_{k-1}).$$

Die Punkte y_k (und die geraden Strecken dazwischen) liegen in $B_r(x)$ und deshalb in U.

Auf jede Strecke wenden wir den Mittelwertsatz auf $g(t) := f(y_k + te_k)$ an. Hier liegt t zwischen 0 und h_k; um die Notation zu vereinfachen, nehmen wir an, $h_k > 0$. (Für $h_k = 0$ gilt $y_k = y_{k-1}$ und wir setzen $\xi_k = 0$.) Aus dem Mittelwertsatz gibt es $\xi_k \in [0, h_k]$ so, dass $g(h_k) - g(0) = h_k g'(\xi_k)$. Das heißt,

$$f(y_k) - f(y_{k-1}) = h_k \partial_k f(z_k),$$

wobei wir $z_k := y_{k-1} + \xi_k e_k$ setzen. Dies bedeutet,

$$f(x + h) - f(x) = \sum_{k=1}^{m} h_k \partial_k f(z_k)$$

und damit

$$\begin{aligned} \Big|f(x + h) - f(x) - Lh\Big| &= \Big|\sum_{k=1}^{m} h_k(\partial_k f(z_k) - \partial_k f(x_k))\Big| \\ &\leq \|h\| \sum_{k=1}^{m} \big|\partial_k f(z_k) - \partial_k f(x_k)\big|. \end{aligned}$$

Für $h \to 0$ konvergiert $z_k \to x$ (weil $\|z_k - x\| \leq \|h\|$). Weil jedes $\partial_k f$ stetig ist, konvergiert jeder Summand – und deshalb auch die Summe – gegen 0. □

Ende der Vorlesung 2009 Januar 19

Korollar E4.2. *Seien V und W endlich-dimensionale Banachräume und sei $G \subset V$ offen. Ferner sei $f\colon G \to W$ eine Abbildung und sei $\{v_1, \dots, v_m\}$ eine Basis für V. Existieren in jedem Punkt $p \in G$ die m Richtungsableitungen $\partial_{v_i} f(p)$ und sind diese stetige Funktionen von p, dann ist f auf G stetig differenzierbar.*

Beweis. Die Basis $\{v_i\}$ gibt uns einen Isomorphismus $\mathbb{R}^m \to V$, wodurch die Richtungsableitungen ∂_{v_i} den partiellen Ableitungen ∂_i entsprechen. Um jeden Punkt $p \in G$ gibt es einen offenen Ball $U \subset G$.

Ist $n := \dim W$, dann sind auch W und $\mathbb{R}^n$ isomorph und wir können $f = (f_1, \dots, f_n)$ komponentenweise betrachten. Wir wenden den Satz auf jede Komponente an. □

E5. Ableitungen in Koordinaten

Bemerkung E5.1. Aus der linearen Algebra wissen wir, eine lineare Abbildung $L\colon \mathbb{R}^m \to \mathbb{R}^n$ wird (bezüglich der Standardbasen $\{e_1, \dots, e_m\}$ für $\mathbb{R}^m$ und $\{e'_1, \dots, e'_n\}$ für $\mathbb{R}^n$) durch eine $(m\times n)$-Matrix (L_{ij}) dargestellt. Die Einträge L_{ij} werden durch $L(e_i) = \sum_j L_{ij} e'_j$ definiert. Für jeden Vektor $v = \sum_i v_i e_i \in V$ gilt dann

$$L\Big(\sum_{i=1}^m v_i e_i\Big) = \sum_{i=1}^m v_i L(e_i) = \sum_{i=1}^m v_i \sum_{j=1}^n L_{ij} e'_j.$$

Definition E5.2. Sei $G \subset \mathbb{R}^m$ offen und $f\colon G \to \mathbb{R}^n$ differenzierbar. Für $p \in G$ ist die Ableitung $D_p f$ eine lineare Abbildung $\mathbb{R}^m \to \mathbb{R}^n$. Deren Darstellungsmatrix $J_p f$ heißt die *Jacobimatrix* von f in p.

Lemma E5.3. *Die Einträge der Jacobimatrix $J_p f$ sind die partiellen Ableitungen $\partial_i f_j(p)$ der Komponenten von f. D.h., es gilt*

$$Jf = \big(\partial_i f_j\big) = \begin{pmatrix} \partial_1 f_1 & \cdots & \partial_m f_1 \\ \vdots & & \vdots \\ \partial_1 f_n & \cdots & \partial_m f_n \end{pmatrix}.$$

Beweis. Es gilt $D_p f(e_i) = \sum_j D_p f_j(e_i) e'_j = \sum_j (\partial_i f_j(p)) e'_j$. □

Bemerkung E5.4. Die i-te Spalte von $J_p f$ ist die partielle Ableitung $\partial_i f(p) = D_p f(e_i)$. Die j-te Reihe ist $J_p f_j$, die Darstellungsmatrix von $D_p f_j$. (Diese nennen wir später auch Gradient von f_j.)

Folgendes Lemma kennen wir aus der linearen Algebra.

Lemma E5.5. *Die Darstellungsmatrix der Komposition zweier linearer Abbildungen ist das Matrixprodukt der jeweiligen Darstellungsmatrizen. Das heißt, für $N = M \circ L$ gilt $N_{ik} = \sum_j L_{ij} M_{jk}$.* □

Bemerkung E5.6. Die Kettenregel sagt, die Ableitung einer Komposition ist die Komposition der Ableitungen. Das heißt, die Jacobimatrix der Komposition ist das Matrixprodukt der Jacobimatrizen. Seien $f\colon \mathbb{R}^m \to \mathbb{R}^n$ und $g\colon \mathbb{R}^l \to \mathbb{R}^m$ Abbildungen und sei $h = f \circ g\colon \mathbb{R}^l \to \mathbb{R}^m$ deren Komposition. (Natürlich können wir auch Funktionen mit offenen Teilmengen als Definitionsbereichen betrachten.) Sei $x \in \mathbb{R}^l$. Wir setzen $y := g(x) \in \mathbb{R}^m$ und $z := f(y) \in \mathbb{R}^n$. Dann sagt die Kettenregel, $D_x h = D_y f \circ D_x g$. Für die Jacobimatrizen haben wir deshalb $J_x h = J_y f \cdot J_x g$. Das heißt, $\partial_i h_k = \sum_j \partial_j f_k \, \partial_i g_j$ oder

$$\frac{\partial h_k}{\partial x_i} = \sum_{j=1}^m \frac{\partial f_k}{\partial y_j} \frac{\partial g_j}{\partial x_i}.$$

Wegen $y_j = g_j(x)$ und $z_k = f_k(y) = h_k(x)$ schreibt man auch gern

$$\frac{\partial z_k}{\partial x_i} = \sum_{j=1}^m \frac{\partial z_k}{\partial y_j} \frac{\partial y_j}{\partial x_i}.$$

E6. Höhere Ableitungen

Seien weiter V und W Banachräume mit $\dim V = m < \infty$ und sei $G \subset V$ offen. Bezeichne $L(V, W)$ den Vektorraum aller linearen Abbildungen $V \to W$. Mit der Operatornorm ist $L(V, W)$ auch ein Banachraum.

Ist eine Abbildung $f\colon G \to W$ auf G differenzierbar, dann ist $Df\colon p \mapsto D_p f$ eine Abbildung $G \to L(V, W)$. Falls Df stetig ist, heißt f *stetig differenzierbar*. Dann kann man fragen, ob Df sogar in $p \in G$ differenzierbar ist. Wenn ja, dann ist $D_p(Df)$ eine lineare Abbildung $V \to L(V, W)$. (Wir könnten schreiben $D_p(Df) \in L(V, L(V, W))$.) Das heißt, für $v \in V$ ist $D_p(Df)(v) = \partial_v(Df)(p) \in L(V, W)$ eine lineare Abbildung $V \to W$. Ist weiter $v' \in V$, dann haben wir

$$(D_p(Df)(v))(v') = (\partial_v(Df)(p))(v') \in W.$$

Diese Notation und die Rollen der beiden Vektoren v und v' bleiben dabei ein bisschen unklar.

Es hilft, ein Lemma aus der linearen Algebra zu benutzen.

Definition E6.1. Für $k \in \mathbb{N}$ sei $L^k(V, W)$ der Vektorraum aller k-linearen Abbildungen

$$f\colon V^k = V \times \cdots \times V \to W,$$

d.h. Abbildungen, die in jedem der k Argumente linear sind. Es gelten $L^0(V, W) \cong W$ und $L^1(V, W) = L(V, W)$.

Lemma E6.2. *Für jedes $k \in \mathbb{N}$ gibt es einen natürlichen Isomorphismus*

$$\varphi_k\colon L(V, L^k(V, W)) \to L^{k+1}(V, W)$$

definiert durch

$$(\varphi_k(\mu))(v, v_1, v_2, \dots, v_k) := (\mu(v))(v_1, \dots, v_k) \in W.$$

Beweis. Dass $\varphi_k(\mu)$ eine Abbildung $V^{k+1} \to W$ ist, ist offensichtlich. Dass diese Abbildung multilinear ist, ist eine unkomplizierte Aufgabe. □

Jetzt können wir die höheren Ableitungen $D^k f$ rekursiv so definieren, dass $D^k_p f \in L^k(V,W)$.

Definition E6.3. Sei $f\colon G \to W$ eine Abbildung. Für $k = 0$ definieren wir $D^0_p f = f(p) \in L^0(V,W) \cong W$, d.h., $D^0 f = f$, und wir sagen, jede Abbildung f ist 0-*mal differenzierbar*.

Für $k \in \mathbb{N}^+$ nehmen wir an, f ist $(k-1)$-mal differenzierbar. Falls $D^{k-1} f\colon G \to L^{k-1}(V,W)$ in $p \in G$ differenzierbar ist, sagen wir, f ist *k-mal in p differenzierbar*. Wir definieren die k-te Ableitung von f in p als

$$D^k_p f = \varphi_{k-1}(D_p(D^{k-1} f)) \in L^k(V,W).$$

Falls $D^k_p f$ in jedem $p \in G$ existiert, sagen wir, f ist *k-mal differenzierbar* (auf G). Wir definieren

$$D^k f\colon G \to L^k(V,W), \quad p \mapsto D^k_p f.$$

Bemerkung E6.4. Damit ist die erste Ableitung die Ableitung: $D^1_p f = D_p f$. Die zweite Ableitung in p ist eine bilineare Abbildung $D^2_p f\colon V \times V \to W$.

Wir werden gleich sehen, dass auch höhere Ableitungen sich mittels Richtungsableitungen ausdrücken lassen. Zunächst berechnen wir aber einmal eine zweite Ableitung direkt aus der Definition.

Beispiel E6.5. Sei V ein endlich-dimensionaler Vektorraum mit Skalarprodukt $\langle\cdot,\cdot\rangle$. Es gibt einen natürlichen Isomorphismus

$$\varphi\colon V \to L(V,\mathbb{R}), \quad p \mapsto \langle p, \cdot\rangle .$$

Das heißt, $\varphi(p)(v) := \langle p, v\rangle$.

Sei $r(v) := \|v\| = \sqrt{\langle v, v\rangle}$ die Norm auf V. In Beispiel E3.10 haben wir die Ableitung $D_p r(v) = {}^{\langle v, p\rangle}/_{\|p\|}$ berechnet. Das heißt,

$$D_p r = \frac{1}{r(p)}\varphi(p) \in L(V,\mathbb{R}).$$

Damit ist Dr das Produkt der beiden Funktionen ${}^1/_r$ und φ. Nach der Kettenregel gilt

$$D_p \frac{1}{r} = -\frac{1}{r(p)^2} D_p r = -\frac{1}{r(p)^3}\varphi(p).$$

Weil φ linear ist, gilt natürlich $D_p\varphi = \varphi$. Korollar E3.9 erklärt die Produktregel für Skalarmultiplikation; danach ist

$$\begin{aligned} D_p(Dr)(v) = D_p\Big(\frac{1}{r}\varphi\Big)(v) &= \frac{1}{r(p)}\varphi(v) - \frac{1}{r(p)^3}(\varphi(p)(v))\,\varphi(p) \\ &= \frac{\varphi(v)}{\|p\|} - \frac{\langle p, v\rangle\ \varphi(p)}{\|p\|^2}. \end{aligned}$$

Nach der Definition der zweiten Ableitung gilt dann

$$D^2_p r(v,w) = (D_p(Dr)(v))(w) = \frac{\langle v, w\rangle\,\langle p, p\rangle - \langle p, v\rangle\,\langle p, w\rangle}{\|p\|^3}.$$

Ende der Vorlesung 2009 Januar 22

Mithilfe des folgenden Lemmas können wir höhere Ableitungen leichter berechnen.

Lemma E6.6. *Die Funktion $f\colon G \to W$ sei k-mal differenzierbar in $p \in G$. (Dies setzt voraus, dass f auf G $(k-1)$-mal differenzierbar ist.) Für alle $v_1, \dots, v_k \in V$ gilt dann*

$$D^k_p f(v_1, \dots, v_k) = \partial_{v_1} \cdots \partial_{v_k} f(p).$$

Insbesondere existiert die k-fache Richtungsableitung auf der rechten Seite.

Beweis. Wir benutzen Induktion über k. Der Fall $k = 0$ ist die Definition $D^0_p f = f(p)$. (Und der Fall $k = 1$ ist Lemma E2.5.)

Für den Induktionsschritt definieren wir die lineare Auswertungsabbildung

$$g\colon L^{k-1}(V,W) \to W, \quad \mu \mapsto \mu(v_2, \dots, v_k).$$

Wir betrachten die Komposition

$$g \circ D^{k-1} f\colon G \to W, \quad p \mapsto D^{k-1}_p f(v_2, \dots, v_k).$$

Nach Lemma E2.5 und der Induktionsvoraussetzung gilt dann

$$D_p(g \circ D^{k-1} f)(v_1) = \partial_{v_1}(D^{k-1}_p f(v_2, \dots, v_k)) = \partial_{v_1}\partial_{v_2} \cdots \partial_{v_k} f.$$

Andererseits gilt aus der Kettenregel

$$\begin{aligned} D_p(g \circ D^{k-1} f)(v_1) &= g(D_p(D^{k-1} f)(v_1)) \\ &= (D^{k-1}_p f(v_1))(v_2, \dots, v_k) \\ &= D^k_p f(v_1, \dots, v_k). \end{aligned}$$

Ein Vergleich dieser beiden Gleichungen schließt den Beweis ab. □

Bemerkung E6.7. Sei G offen in $V = \mathbb{R}^m$. Die k-te Ableitung einer Funktion $f\colon G \to W$ lässt sich mittels k-facher patieller Ableitungen berechnen. Seien k Vektoren $v_1, \dots, v_k \in \mathbb{R}^m$ gegeben, wobei

$$v_i = (v_{i,1}, \dots, v_{i,m}) = \sum_{j=1}^{m} v_{i,j} e_j \quad \text{für } i = 1, \dots, k.$$

Ist f in p k-mal differenzierbar, dann gilt

$$\begin{aligned} D^k_p f(v_1, \dots, v_k) &= \sum_{j_1, \dots, j_k = 1}^{m} v_{1,j_1} \cdots v_{k,j_k} D^k_p f(e_{j_1}, \dots e_{j_k}) \\ &= \sum_{j_1, \dots, j_k = 1}^{m} v_{1,j_1} \cdots v_{k,j_k} \partial_{j_1} \cdots \partial_{j_k} f(p). \end{aligned}$$

Definition E6.8. Ist $f\colon G \to W$ k-mal differenzierbar und ist $D^k f\colon G \to L^k(V,W)$ stetig, dann sagen wir f ist (auf G) *k-mal stetig differenzierbar*. Wir schreiben $f \in C^k(G,W)$. Falls $f \in C^k$ für alle $k \in \mathbb{N}$, sagen wir, f ist *beliebig oft differenzierbar*, $f \in C^\infty(G,W)$.

Satz E6.9. *Seien V und W endlich-dimensionalen Banachräume und sei $G \subset V$ offen. Ferner sei $f\colon G \to W$ eine Abbildung und sei $\{v_1, \dots, v_m\}$ eine Basis für V. Existieren in jedem Punkt $p \in G$ alle k-fachen Richtungsableitungen*

$$\partial_{v_{i_1}} \cdots \partial_{v_{i_k}} f(p)$$

und sind diese stetige Funktionen von p, dann ist f auf G k-mal stetig differenzierbar.

Beweisskizze. Wir dürfen annehmen, $V = \mathbb{R}^m$ und $W = \mathbb{R}^n$; wir benutzen die Standardbasen $\{e_i\}$ und $\{e'_j\}$. Seien f_j die Komponenten von $f = \sum_j f_j e'_j$. Dann sind die $(k-1)$-fachen partiellen Ableitungen

$$\partial_{i_1} \cdots \partial_{i_{k-1}} f_j$$

die Komponenten von $D^{k-1} f$ bzgl. einer natürlichen Basis für $L^{k-1}(\mathbb{R}^m, \mathbb{R}^n)$. Existieren deren (einfache) partielle Ableitungen, ist $D^{k-1} f$ nach Korollar E4.2 stetig differenzierbar. □

Bemerkung E6.10. Es gilt

$$\begin{aligned} D_p^2 f(u,v) = \partial_u \partial_v f(p) &= \lim_{s\to 0} \frac{\partial_v f(p+su) - \partial_v f(p)}{s} \\ &= \lim_{s\to 0} \frac{1}{s} \lim_{t\to 0} \Big(\frac{f(p+su+tv) - f(p+su)}{t} - \frac{f(p+tv) - f(p)}{t} \Big) \\ &= \lim_{s\to 0} \lim_{t\to 0} \frac{1}{st} \big(f(p+su+tv) - f(p+su) - f(p+tv) + f(p) \big). \end{aligned}$$

Definition E6.11. Wir nennen

$$\Delta_p^2 f(h,k) := f(p+h+k) - f(p+h) - f(p+k) + f(p)$$

die *Differenz zweiter Ordnung* von f. Es gilt $\Delta_p^2 f(h,k) = \Delta_p^2 f(k,h)$.

Bemerkung E6.12. Für „kleine" Vektoren h, k gilt $\Delta_p^2 f(h,k) \approx D_p^2 f(h,k)$. Genauer gesagt können wir die letzte Bemerkung wie folgt schreiben:

$$D_p^2 f(u,v) = \partial_u \partial_v f(p) = \lim_{s\to 0} \lim_{t\to 0} \frac{\Delta_p^2 f(su,tv)}{st}.$$

Entsprechend gilt

$$D_p^2 f(v,u) = \partial_v \partial_u f(p) = \lim_{t\to 0} \lim_{s\to 0} \frac{\Delta_p^2 f(su,tv)}{st}.$$

Natürlich darf man nicht immer zwei Limites tauschen. Viele wichtige Sätze der Analysis geben Voraussetzungen, unter denen man bestimmte Limites tauschen darf. Hier reicht es (nach einem H.A. Schwarz bzw. A. Clairaut zugeordneten Satz) aus, dass f in p zweimal differenzierbar ist.

Beispiel E6.13. Die existenz der zweifachen Richtungsableitungen ist nicht ausreichend. Sei $p := (0,0) \in \mathbb{R}^2$ und sei $f\colon \mathbb{R}^2 \to \mathbb{R}$ durch $f(p) := 0$ und

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

für $(x,y) \neq p$ definiert. Dann gilt $\partial_x \partial_y f(p) \neq \partial_y \partial_x f(p)$. (Beweis als Aufgabe. Wegen $f(x,y) = -f(y,x)$ ist es klar, dass $\partial_x \partial_y f(x,y) = -\partial_y \partial_x f(y,x)$.)

Satz E6.14. *Sei $f\colon G \to W$ differenzierbar und in $p \in G$ zweimal differenzierbar. Dann gilt*

$$\partial_u \partial_v f(p) = \lim_{t\to 0} \frac{\Delta_p^2 f(tu,tv)}{t^2} = \partial_v \partial_u f(p).$$

für alle $u, v \in V$.

Beweis. Weil der Ausdruck in der Mitte symmetrisch in u, v ist, reicht es aus, die erste Gleichung zu beweisen. Um der Einfachheit willen, nehmen wir an, $\dim W < \infty$. (Für den allgemeinen Fall, siehe z.B. Amman–Escher VII.5.2.) Weil wir dann alles komponentenweise in W betrachten können, dürfen wir $W = \mathbb{R}$ annehmen.

Wir wissen, f ist in p zweimal differenzierbar. Das heißt per Definition, dass $Df\colon G \to W$ existiert und in p differenzierbar ist, d.h., dass

$$R(h) := D_{p+h} f - D_p f - D_p(Df)(h) \in L(V, \mathbb{R})$$

sublinear wächst ($\lim {}^{R(h)}/_{\|h\|} = 0$). Sei $\varepsilon > 0$ gegeben. Dass der Limes 0 ist heißt, es gibt $\delta > 0$ so, dass $\|R(h)\| < \varepsilon \|h\|$ für $\|h\| < \delta$. (Hier ist $\|R\|$ die Operatornorm.) Das heißt, für $\|h\| < \delta$ und alle $v \in V$ gilt

$$\begin{aligned} \big| D_{p+h} f(v) - D_p f(v) - D_p^2 f(h,v) \big| = |R(h)(v)| &\le \|R(h)\| \, \|v\| \\ &\le \varepsilon \|h\| \, \|v\|. \end{aligned}$$

Diese Ungleichung können wir auch wie folgt schreiben:

$$\big| \partial_v f(p+h) - \partial_v f(p) - \partial_h \partial_v f(p) \big| \le \varepsilon \|h\| \, \|v\|, \qquad (*)$$

(Ist nun $h = su$ mit $|s| < {}^{\delta}/_{\|u\|}$, dann folgt

$$\Big| \lim_{t\to 0} \frac{\Delta_p^2 f(su,tv)}{st} - \partial_u \partial_v f(p) \Big| \le \varepsilon \|u\| \, \|v\|,$$

was eine genauere Version der ersten Gleichung in Bemerkung E6.12 ist.)

Nun sei $g(t) := f(p+su+tv) - f(p+tv)$. (Natürlich macht diese Definition nur dann Sinn, wenn s und t hinreichend klein sind, damit die Punkte im Definitionsbereich G von f liegen. Weil wir sowieso am Ende den Limes $s = t \to 0$ betrachten, schreiben wir diese Schranken nicht explizit.) Es gelten $\Delta_p^2 f(su,tv) = g(t) - g(0)$ und $g'(t) = \partial_v f(p+su+tv) - \partial_v f(p_t v)$. Nach dem Mittelwertsatz gibt es zu jedem t ein τ (mit $0 < {}^{\tau}/_{t} < 1$) so, dass

$$\begin{aligned} \frac{1}{t} \Delta_p^2 f(su,tv) = \frac{1}{t}(g(t) - g(0)) &= g'(\tau) \\ &= \partial_v f(p+su+\tau v) - \partial_v f(p+\tau v) \\ &= (\partial_v f(p+su+\tau v) - \partial_v f(p)) \\ &\quad - (\partial_v f(p+\tau v) - \partial_v f(p)). \end{aligned}$$

Jetzt benutzen wir $(*)$ zweimal mit $h = su + \tau v$ bzw. $h = \tau v$. Es gilt

$$\begin{aligned} &\big| g'(\tau) - s \partial_u \partial_v f(p) \big| \\ &\quad \le \big| \partial_v f(p+su+\tau v) - \partial_v f(p) - \partial_{su+\tau v} \partial_v f(p) \big| \\ &\qquad + \big| \partial_v f(p+\tau v) - \partial_v f(p) - \partial_{\tau v} \partial_v f(p) \big| \\ &\quad \le \varepsilon \|v\| \big(\|su + \tau v\| + \|\tau v\| \big). \end{aligned}$$

Damit gilt (natürlich unter der Voraussetzung, dass $\|\tau v\| < \delta$ und $\|su + \tau v\| < \delta$)

$$\Big| \frac{\Delta_p^2 f(su,tv)}{st} - \partial_u \partial_v f(p) \Big| \le \frac{\varepsilon \|v\|}{|s|} \big(\|su + \tau v\| + \|\tau v\| \big).$$

Für $s = t$ (mit t hinreichend klein) gilt dann

$$\left|\frac{\Delta_p^2 f(tu,tv)}{t^2} - \partial_u\partial_v f(p)\right| \le \varepsilon\,\|v\|\left(\left\|u + \tfrac{\tau}{t}v\right\| + \left\|\tfrac{\tau}{t}v\right\|\right) \le \varepsilon\,\|v\|\Big(\|u\| + 2\|v\|\Big).$$

Weil ε beliebig war, konvergiert die linke Seite wie gewünscht gegen 0. □

Ende der Vorlesung 2008 Januar 26

Bemerkung E6.15. Häufig sieht man eine schwächere Version des Satzes, die die Stetigkeit von $D^2 f$ voraussetzt.

Korollar E6.16. *Die Abbildung $f\colon G \to W$ sei k-mal in $p \in G$ differenzierbar. Für jede Permutation*

$$\sigma\colon (1,\dots,k) \mapsto (\sigma 1,\dots,\sigma k)$$

und für alle $v_1,\dots,v_k \in V$ gilt

$$\partial_{v_1}\cdots\partial_{v_k} f(p) = D_p^k f(v_1,\dots,v_k) = D_p^k f(v_{\sigma 1},\dots,v_{\sigma k}) = \partial_{v_{\sigma 1}}\cdots\partial_{v_{\sigma k}} f(p).$$

Das heißt, die k-lineare Abbildung $D_p^k f$ ist symmetrisch.

Beweisidee. Jede Permutation von $(1,\dots,k)$ ist eine Komposition von (höchstens $\binom{k}{2}$) Transpositionen. □

Beispiel E6.17. Sei $U \subset \mathbb{R}^m$ offen und seien $f_1,\dots,f_m\colon U \to \mathbb{R}$ differenzierbare Funktionen. Man kann sich fragen, ob es eine Funktion $g\colon U \to \mathbb{R}$ so gibt, dass $\partial_i g = f_i$ für alle $i = 1,\dots,m$. Der letzte Satz gibt uns eine notwendige (aber nicht hinreichende) Bedingung dafür, dass dieses System partieller Differentialgleichungen eine Lösung g hat: falls g existiert, dann gilt

$$\partial_i f_j = \partial_i\partial_j g = \partial_j\partial_i g = \partial_j f_i$$

für alle i, j.

Definition E6.18. Sei V ein Vektorraum und seien $p, q \in V$. Die *Strecke* $\overline{pq}$ zwischen p und q ist

$$\overline{pq} := \{(1-t)p + tq : t \in [0,1]\} \subset V.$$

Das heißt, $\overline{pq}$ ist die Spur des geraden Weges $\gamma\colon [0,1] \to V$ mit $\gamma(t) := (1-t)p + tq$. Eine Teilmenge $T \subset V$ heißt *konvex*, falls

$$p, q \in T \implies \overline{pq} \subset T.$$

Bemerkung E6.19. Sei $f\colon G \to W$ und sei $\overline{pq} \subset G$. Die Werte von f auf dieser Strecke $\overline{pq}$ können wir als Funktion von t betrachten:

$$g(t) := f(\gamma(t)) = f((1-t)p + tq).$$

Es gilt $\gamma'(t) \equiv q - p$ und damit $g'(t) = \partial_{q-p} f(\gamma(t))$ und per Induktion weiter

$$g^{(k)}(t) = \partial_{q-p}\cdots\partial_{q-p} f(\gamma(t)) = \partial_{q-p}^k f(\gamma(t)) = D_{\gamma(t)}^k f(q-p,\dots,q-p).$$

Auf dieser Weise können wir den Schrankensatz D3.4 auf f anwenden.

Korollar E6.20. *Sei $f\colon G \to W$ differenzierbar und sei $\overline{pq} \subset G$. Setze $M := \sup\{\|D_x f\| : x \in \overline{pq}\}$. Dann gilt*

$$\big\|f(q) - f(p)\big\| \le M\,\|q - p\|.$$

Beweis. (Folgt direkt aus Satz D3.4; Details als Aufgabe.) □

Definition E6.21. Sei $f\colon G \to W$ eine n-mal differenzierbare Abbildung. Das n-te Taylorpolynom von f an der Entwicklungsstelle $p \in G$ ist

$$T_n(x) := T_{n,p}^f(x) := \sum_{k=0}^{n} \frac{\partial_{x-p}^k f(p)}{k!}.$$

Satz E6.22. *Sei $f\colon G \to W$ n-mal differenzierbar. Das n-te Taylorpolynom ist eine gute Approximation für f, im Sinne, dass für das Restglied $R(x) := f(x) - T_n(x)$ gilt*

$$\lim_{x\to p} \frac{\|R(x)\|}{\|x-p\|^n} = 0.$$

Bemerkung E6.23. Der Fall $n = 0$ bzw. $n = 1$ ist einfach die Definition von Stetigkeit bzw. Differenzierbarkeit. Würden wir wie oben alles nur eindimensional entlang der Strecke $\overline{xp}$ betrachten, könnten wir daraus schließen, dass entlang jeder Geraden $R(x)/\|x-p\|^n \to 0$. Dies ist natürlich eine deutlich schwächere Aussage. (Für $n = 1$ vgl. Differenzierbarkeit mit der Existenz aller Richtungsableitungen.)

Beweis. Wir benutzen Induktion über n und haben schon den Induktionsbeginn erledigt. Für den Induktionsschritt, sei $R(x) = f(x) - T_{n,p}^f(x)$ das Restglied zu einem n-mal differenzierbaren f. Betrachten wir (für $k > 0$) nun die Abbildung

$$g\colon x \mapsto \partial_{x-p}^k f(p) = D_p^k f(x-p,\dots,x-p)$$

als Komposition von $x \mapsto (x-p,\dots,x-p)$ mit dem multilinearen $D_p^k f$, ergibt sich nach der allgemeinen Produktregel E3.8 und dem Korollar E6.16:

$$\begin{aligned}\partial_v g(x) = D_x g(v) &= kD_p^k f(x-p,\dots,x-p,v)\\ &= k\Big(D_p^{k-1}(Df)(x-p,\dots,x-p)\Big)(v)\\ &= k\Big(\partial_{x-p}^{k-1}(Df)(p)\Big)(v).\end{aligned}$$

Damit gilt

$$\partial_v R(x) = \partial_v f(x) - \sum_{k=1}^{n} \frac{1}{(k-1)!}\Big(\partial_{x-p}^{k-1}(Df)(p)\Big)(v)$$

und deshalb

$$\begin{aligned}D_x R &= D_x f - \sum_{k=1}^{n} \frac{1}{(k-1)!}\partial_{x-p}^{k-1}(Df)(p)\\ &= D_x f - \sum_{j=0}^{n-1} \frac{1}{j!}\partial_{x-p}^{j}(Df)(p) = D_x f - T_{n-1,p}^{Df}(x),\end{aligned}$$

d.h., $D_x R$ ist das Restglied für das Taylorpolynom von Df. Weil Df eine $(n-1)$-mal differenzierbare Abbildung ist, folgt

aus der Induktionsvoraussetzung, dass $D_x R/\|x-p\|^{n-1} \to 0$. Das heißt, zu jedem $\varepsilon > 0$ gibt es $\delta > 0$ so, dass

$$0 < \|x-p\| < \delta \implies \|D_x R\| < \varepsilon \|x-p\|^{n-1}.$$

Damit gilt für $0 < \|x-p\| < \delta$, dass

$$M_x := \sup_{y \in \overline{xp}} \|D_y R\| < \varepsilon \|x-p\|^{n-1}.$$

Nach Korollar E6.20 und wegen $R(p) = 0$ gilt dann

$$\|R(x)\| = \|R(x) - R(p)\| \le M_x \|x-p\| < \varepsilon \|x-p\|^n.$$

Weil $\varepsilon > 0$ beliebig war, heißt das wie gewünscht, dass $\|R(x)\|/\|x-p\|^n \to 0$ für $x \to p$. □

E7. Klassische Vektoranalysis

In diesem Abschnitt betrachten wir den Raum $V = \mathbb{R}^n$ mit dem euklidischen Skalarprodukt $\langle x, y\rangle = x \cdot y = \sum_i x_i y_i$. Das Skalarprodukt erlaubt uns, die klassischen Differentialoperatoren div, grad und (für $n = 3$) rot zu definieren. Die bekannten eigenschaften dieser Operatoren erwähnen wir hier, ohne die (meistens sehr einfache) Beweise aufzuschreiben.

Definition E7.1. Sei $G \in \mathbb{R}^n$ offen. Aus der Physik kennt man die folgenden Namen: eine reelle Funktion $f\colon G \to \mathbb{R}$ heißt *skalares Feld* auf G; eine Abbildung $X\colon G \to \mathbb{R}^n$ heißt *Vektorfeld* auf G.

a. Gradient

Definition E7.2. Sei f ein differenzierbares skalares Feld auf $G \subset \mathbb{R}^n$. Der *Gradient* von f ist das folgende Vektorfeld $\nabla f = \operatorname{grad} f$ auf G:

$$\operatorname{grad}_p f := \nabla f(p) := (\partial_1 f(p), \dots, \partial_n f(p)).$$

Bemerkung E7.3. Das Symbol ∇ heißt „Nabla“. Als Gedächtnisstütze kann man $\nabla = (\partial_1, \dots, \partial_n)$ schreiben. (Dieser Gedächtnisstütze eine genaue Bedeutung zu geben – in einem Vektorraum von Differentialoperatoren – würde den Rahmen dieser Vorlesung sprengen.)

Bemerkung E7.4. Der Gradient $\nabla f = \operatorname{grad} f$ wird durch

$$\langle v, \operatorname{grad}_p f\rangle = \langle v, \nabla f(p)\rangle = \partial_v f(p) = D_p f(v) \text{ für alle } v \in V$$

charaktisiert.

Bemerkung E7.5. Der Richtungsableitungoperator ∂_v wird oft (besonders in der Physik) als $\langle v, \nabla\rangle$ oder $v \cdot \nabla$ geschrieben:

$$\langle v, \nabla\rangle f := \langle v, \nabla f\rangle = \partial_v f = v \cdot (\nabla f) = (v \cdot \nabla) f.$$

Ende der Vorlesung 2009 Januar 29

Bemerkung E7.6. Der Gradient ∇f gibt die Richtung und die Größe des stärksten Wachstums von f an. Für $\|v\| = 1$ ist

$$\partial_v f = \langle v, \nabla f\rangle = \|\nabla f\| \cos\varphi,$$

wobei φ der Winkel zwischen v und ∇f ist. Insbesondere ist ∇f senkrecht zu den Niveaumengen von f: f ist genau dann konstant entlang einer Kurve $\gamma(t)$ in G, wenn $\nabla f \perp \dot\gamma(t)$ für alle t.

Bemerkung E7.7. In der Physik ist z.B. das (statische) elektrische Feld der Gradient des elektrischen Potentials. (Allgemeiner ist der Gradient einer jenen potentiellen Energie ein Kraftfeld.)

Bemerkung E7.8. Ein skalares Feld f mit $\operatorname{grad} f = 0$ ist lokal konstant.

Beispiel E7.9. Sei $v \in \mathbb{R}^n$. Das skalare Feld $f(x) := \langle v, x\rangle$ hat konstanten Gradienten $\operatorname{grad} f \equiv v$.

b. Divergenz

Definition E7.10. Sei $X = (X_1, \dots, X_n)$ ein differenzierbares Vektorfeld auf $G \subset \mathbb{R}^n$. Die *Divergenz* von X ist das folgende skalare Feld $\langle \nabla, X\rangle = \nabla \cdot X = \operatorname{div} X$ auf G:

$$\operatorname{div}_p X := \langle \nabla, X\rangle(p) := (\nabla \cdot X)(p) := \sum_{i=1}^n \partial_i X_i(p).$$

Bemerkung E7.11. Die *Spur* einer $(n \times n)$-Matrix $A = (A_{ij})$ ist die Summe der Diagonalelemente: $\operatorname{tr} A := \sum_i A_{ii}$. Damit gilt $\operatorname{div}_p X = \operatorname{tr} J_p X$.

Bemerkung E7.12. Die Divergenz eines Vektorfeldes X kann physikalisch als „Quellendichte“ von X interpretiert werden. Die Divergenz des Schwerefeldes z.B. ist die (Massen-) Dichte; die Divergenz des elektrischen Feldes ist die Ladungsdichte. Ein Vektorfeld X mit $\operatorname{div} X = 0$ heißt *divergenzfrei* oder *quellenfrei*; das Magnetfeld ist quellenfrei. Weitere Beispiele sind Strömungsfelder von inkompressiblen Flüssigkeiten.

Bemerkung E7.13. Das Skalarprodukt $\langle v, w\rangle = \langle w, v\rangle$ ist symmetrisch. Die Schreibweisen $\langle v, \nabla\rangle = v \cdot \nabla$ bzw. $\langle \nabla, X\rangle = \nabla \cdot X$ für Richtungsableitungen bzw. Divergenz darf man hingegen nicht verwechseln.

Beispiel E7.14. Das Vektorfeld $X(x) = x$ hat konstante Divergenz $\operatorname{div} X \equiv n$.

c. Rotation

Auf $\mathbb{R}^3$ wird das Vektorprodukt $v \times w \in \mathbb{R}^3$ durch

$$v \times w = (v_2 w_3 - v_3 w_2, v_3 w_1 - v_1 w_3, v_1 w_2 - v_2 w_1)$$

definiert. Dazu verwandt ist die Rotation eines Vektorfeldes.

Definition E7.15. Sei $X = (X_1, X_2, X_3)$ ein differenzierbares Vektorfeld auf $G \subset \mathbb{R}^3$. Die *Rotation* von X ist das folgende Vektorfeld $\nabla \times X = \operatorname{rot} X = \operatorname{curl} X$ auf G:

$$\operatorname{rot}_p X := \operatorname{curl}_p X := (\nabla \times X)(p)$$
$$:= \big(\partial_2 X_3 - \partial_3 X_2, \partial_3 X_1 - \partial_1 X_3, \partial_1 X_2 - \partial_2 X_1\big)(p).$$

Bemerkung E7.16. Sei $\{e_1, e_2, e_3\}$ die Standardbasis für $\mathbb{R}^3$. Als Gedächtnisstütze kann man schreiben:

$$v \times w = \det\begin{pmatrix} e_1 & v_1 & w_1 \\ e_2 & v_2 & w_2 \\ e_3 & v_3 & w_3 \end{pmatrix}, \quad \nabla \times X = \det\begin{pmatrix} e_1 & \partial_1 & X_1 \\ e_2 & \partial_2 & X_2 \\ e_3 & \partial_3 & X_3 \end{pmatrix}.$$

Bemerkung E7.17. Es gibt vier sinnvolle Möglichkeiten, Operatoren aus {div, grad, rot} zu kombinieren: div(rot X) und div(grad f) sind skalare Felder; rot(rot X) und rot(grad f) sind Vektorfelder. Den Fall rot(rot X) betrachten wir nicht weiter.

Satz E7.18. *Seien f und X zweimal differenzierbare Felder auf $G \subset \mathbb{R}^3$. Es gelten*

$$0 = \operatorname{rot}(\operatorname{grad} f) = \nabla \times (\nabla f),$$
$$0 = \operatorname{div}(\operatorname{rot} X) = \langle \nabla, \nabla \times X \rangle .$$

Beweis. (Aufgabe nach dem Satz von Schwarz.) □

Bemerkung E7.19. Um diesen Satz sinnvoll auf höhere Dimensionen zu erweitern, muss man die sogenannte äußere Ableitung von Differentialformen benutzen.

Bemerkung E7.20. Beispiel E6.17 können wir jetzt wie folgt verstehen: Sei X ein Vektorfeld auf $G \subset \mathbb{R}^3$. Wann gibt es ein skalares Feld g mit $\nabla g = X$? Eine notwendige Bedingung dafür ist, dass $\nabla \times X = 0$.

Bemerkung E7.21. Die Rotation $\nabla \times X$ heißt auch *Wirbeldichte*. Ein Vektorfeld X mit $\nabla \times X = 0$ heißt *wirbelfrei*. Für jedes f ist der Gradient ∇f wirbelfrei. Das Gesetz der elektromagnetischen Induktion sagt, die Wirbeldichte des elektrischen Feldes ist die zeitliche Ableitung des Magnetfeldes.

Beispiel E7.22. Sei $v \in \mathbb{R}^3$. Das Vektorfeld $X(x) = v \times x$ beschreibt eine Rotation um die Achse $\overline{0v}$. Es gilt rot $X \equiv 2v$.

d. Laplaceoperator

Definition E7.23. Sei f ein zweimal differenzierbares skalares Feld auf $G \subset \mathbb{R}^n$. Der *Laplaceoperator* $\Delta = \nabla \cdot \nabla$ liefert ein skalares Feld Δf auf G:

$$\Delta f := \operatorname{div}(\operatorname{grad} f) = \langle \nabla, \nabla f \rangle = \nabla \cdot (\nabla f) = \sum_{i=1}^{n} \partial_i^2 f.$$

Bemerkung E7.24. Die symmetrische Matrix $Hf = (\partial_i \partial_j f)$ heißt *Hessematrix* von f. Es gilt $\Delta f = \operatorname{tr}(Hf)$.

Bemerkung E7.25. Der Laplaceoperator ist sehr wichtig in der Physik. Ist Φ ein Potential, dann ist $\nabla\Phi$ ein konservatives (d.h. wirbelfreies) Kraftfeld und $\Delta\Phi$ eine Dichte. Der Laplaceoperator taucht auch z.B. in den partiellen Differentialgleichungen für Wärmeleitung und Wellenausbreitung auf.

Definition E7.26. Eine *harmonische Funktion* auf $G \subset \mathbb{R}^n$ ist ein skalares Feld f mit $\Delta f = 0$.

Beispiel E7.27. Jede affin lineare Funktion $f(x) = \langle x, v \rangle + b$ ist harmonisch. Eine quadratische Funktion f hat $\Delta f \equiv c$, z.B. für $f(x) = \|x\|^2 = \sum x_i^2$ gilt $\Delta f \equiv 2n$.

Beispiel E7.28. In der Elektrostatik ist das elektrische Potential harmonisch in jedem Gebiet, wo die Ladungsdichte verschwindet.

Bemerkung E7.29. Die Operatoren Gradient, Divergenz, Rotation und Laplaceoperator sind linear:

$$\nabla(f + \lambda g) = \nabla f + \lambda \nabla g,$$
$$\langle \nabla, X + \lambda Y \rangle = \langle \nabla, X \rangle + \lambda \langle \nabla, Y \rangle$$
$$\nabla \times (X + \lambda Y) = \nabla \times X + \lambda \nabla \times Y,$$
$$\Delta(f + \lambda g) = \Delta f + \lambda \Delta g,$$

E8. Umkehrsatz

In der Mathematik geht es oft darum, nach den Lösungen einer Gleichung (bzw. eines Gleichungssystems) zu suchen. Schreiben wir die Gleichungen als $f(x) = y$, dann sind die Folgenden vielleicht die ersten Fragen: ob es zu jedem y ein x gibt mit $f(x) = y$, ob dieses x eindeutig ist und ob kleine Änderungen in den Parametern y zu kleinen Änderungen in der Lösung x führen. Das heißt, wir möchten wissen, ob f eine Umkehrabbildung f^{-1} hat und ob diese stetig (und vielleicht auch differenzierbar) ist.

Es geht hier nur um die (theoretische) Existenzfrage und nicht z.B. darum, eine schöne Formel für f^{-1} zu finden. Schon bei den meisten elementaren Funktionen musste man neue Notationen für die Umkehrfunktionen erfinden ($\sqrt[3]{\cdot}$, log, arcsin usw.).

Beispiel E8.1. Der Fall einer linearen Gleichung $y = Ax$ mit $A \in L(\mathbb{R}^m, \mathbb{R}^n)$ kann man mit der linearen Algebra betrachten. Es gibt zu jedem y eine eindeutige Lösung x (d.h., A ist invertierbar) genau dann, wenn $n = m$ und $\det A \neq 0$. In diesem Fall ist A^{-1} auch linear, insbesondere stetig und sogar differenzierbar.

Bemerkung E8.2. Jede differenzierbare Abbildung f kann man lokal gut durch ihre Ableitung $D_p f$ approximieren. Der Umkehrsatz sagt, falls $D_p f$ invertierbar ist, dann ist auch f bei p lokal invertierbar. Wir betrachten die Gleichung $y = f(x)$ etwa als Störung der affin linearen Gleichung $y - f(p) = D_p f(x - p)$. Falls Letzteres invertierbar ist, bleibt diese Eigenschaft unter der Störung erhalten.

Bemerkung E8.3. Die Invertierbarkeit von $D_p f$ ist natürlich keine notwendige Bedingung. Auch wenn z.B. die Ableitung von $x \mapsto x^3$ in 0 verschwindet, ist $x \mapsto x^3$ sogar global invertierbar.

Definition E8.4. Seien $U \subset \mathbb{R}^m$ und $V \subset \mathbb{R}^n$ offen und sei $f\colon U \to V$ eine Bijektion. Mengentheoretisch existiert die Umkehrabbildung $f^{-1}\colon V \to U$. Falls f und f^{-1} stetig sind, heißt f *Homöomorphismus* zwischen U und V. Falls f und f^{-1} stetig differenzierbar sind, heißt f *Diffeomorphismus*.

Beispiele E8.5.

- $x \mapsto x^3$ ist ein Homöomorphismus $\mathbb{R} \to \mathbb{R}$, ist aber kein Diffeomorphismus, weil $\sqrt[3]{x}$ in 0 nicht differenzierbar ist.

- ist f streng monoton auf einem Intervall $I \subset \mathbb{R}$, dann ist f ein Homöomorphismus $I \to f(I)$.
- $x \mapsto \tan x$ ist ein Diffeomorphismus $(-\pi/2, \pi/2) \to \mathbb{R}$.
- ist $f' \neq 0$ auf einem Intervall $I \subset \mathbb{R}$, dann ist f ein Diffeomorphismus $I \to f(I)$.

Ende der Vorlesung 2009 Februar 2

Lemma E8.6. *Seien $U \subset \mathbb{R}^m$ und $V \subset \mathbb{R}^n$ offen und sei $f\colon U \to V$ eine Bijektion mit Umkehrabbildung $g\colon V \to U$. Sei $p \in U$ so, dass f in p differenzierbar ist und g in $q := f(p) \in V$ differenzierbar ist. Dann gelten $m = n$ und $D_q g = (D_p f)^{-1}$, insbesondere ist $D_p f$ invertierbar.*

Beweis. (Aufgabe nach der Kettenregel.) □

Bemerkung E8.7. Dieses Lemma sagt insbesondere, Dimension wird unter Diffeomorphismen erhalten. Dasselbe gilt unter Homöomorphismen, der Beweis dazu ist aber nicht einfach und braucht Techniken aus der algebraischen Topologie. Für je zwei offene Mengen $U \subset \mathbb{R}^m$ und $V \subset \mathbb{R}^n$ gibt es aber eine Bijektion $U \to V$: die Mengen sind gleich groß (gleichmächtig). Für $m \neq n$ sind aber solche Bijektionen nie stetig; sie sind Stoffe eher für die Mengentheorie als für die Analysis.

Definition E8.8. Ein *Automorphismus* des Vektorraumes $\mathbb{R}^n$ ist eine invertierbare lineare Abbildung $A\colon \mathbb{R}^n \to \mathbb{R}^n$. Die Menge aller Automorphismen heißt $\mathsf{GL}(\mathbb{R}^n) \subset L(\mathbb{R}^n, \mathbb{R}^n)$. Sei $\varphi\colon \mathsf{GL}(\mathbb{R}^n) \to \mathsf{GL}(\mathbb{R}^n)$ die Inversenabbildung $A \mapsto A^{-1}$.

Lemma E8.9. *Für $A \in \mathsf{GL}(\mathbb{R}^n)$ und $v \in \mathbb{R}^n$ gilt*

$$\frac{\|v\|}{\|A^{-1}\|} \leq \|A(v)\| \leq \|A\|\,\|v\|.$$

Beweis. Es gilt $\|v\| = \left\|A^{-1}(Av)\right\| \leq \left\|A^{-1}\right\| \|Av\|$. □

Aus der linearen Algebra wissen wir:

Lemma E8.10. *In $L(\mathbb{R}^n, \mathbb{R}^n)$ ist $\mathsf{GL}(\mathbb{R}^n)$ offen. Die Abbildung φ ist stetig.*

Bemerkung E8.11. Bezüglich der Standardbasis wird $A \in L(\mathbb{R}^n, \mathbb{R}^n)$ durch eine $(n \times n)$-Matrix (A_{ij}) dargestellt. Es gilt genau dann $A \in \mathsf{GL}(\mathbb{R}^n)$, wenn $\det(A_{ij}) \neq 0$.

Bemerkung E8.12. Tatsächlich ist φ sogar differenzierbar mit $D_A\varphi(B) = -A^{-1}BA^{-1}$. Nehmen wir die Differenzierbarkeit an, können wir leicht diese Formel bestätigen: Wir betrachten

$$\alpha\colon A \mapsto A\varphi(A) = AA^{-1} = I.$$

Dann gilt

$$0 = D_A\alpha(B) = B\varphi(A) + AD_A\varphi(B) = BA^{-1} + AD_A\varphi(B).$$

Lemma E8.13. *Seien $U, V \subset \mathbb{R}^n$ offen und sei $f\colon U \to V$ ein stetig differenzierbarer Homöomorphismus. Falls $D_p f$ invertierbar ist für jedes $p \in U$, dann ist f ein Diffeomorphismus.*

Beweis. Weil f ein Homöomorphismus ist, ist die Umkehrabbildung $g\colon V \to U$ stetig. Wir behaupten, g ist differenzierbar, d.h., für jedes $p \in U$ existiert $D_{f(p)}g$. Nach Lemma E8.6 können wir dann die Ableitung als $Dg = \varphi \circ Df \circ g$ schreiben. Die Stetigkeit von Dg folgt aus der Stetigkeit von g, von Df und von φ (Lemma E8.10).

Um die Behauptung zu beweisen, sei $p \in U$. Wir wissen, $D_p f$ ist invertierbar und wollen zeigen, g ist in $f(p)$ differenzierbar. Ohne Einschränkung dürfen wir $p = 0$ annehmen, indem wir f mit $x \mapsto f(p + x)$ ersetzen (Translation im Definitionsbereich). Weiter dürfen wir $f(0) = 0$ annehmen, indem wir f mit $x \mapsto f(x) - f(0)$ ersetzen (Translation in der Zielmenge). Schließlich dürfen wir $D_0 f = \mathrm{id}$ annehmen, indem wir f mit $(D_0 f)^{-1} \circ f$ ersetzen (lineare Abbildung in der Zielmenge). Unter diesen Annahmen gilt dann auch $g(0) = 0$. Zudem gilt $D_0 g = \mathrm{id}$, falls $D_0 g$ existiert, was wir jetzt zeigen wollen.

Mit der Definition von Differenzierbarkeit im Auge setzen wir

$$f(h) - h =: r(h), \qquad g(k) - k =: s(k)$$

für hinreichend kleine $h, k \in \mathbb{R}^n$. Für $k = f(h)$ gilt dann

$$s(k) = g(k) - k = h - f(h) = -r(h).$$

Wir wissen, dass $r(h)/\|h\| \to 0$ für $h \to 0$ und müssen zeigen, dass $s(k)/\|k\| \to 0$ für $k \to 0$.

Es gibt $\varepsilon > 0$ so, dass für $\|h\| < \varepsilon$ gilt $\|r(h)\| \leq \|h\|/2$. Weil g stetig ist, gibt es zu diesem ε ein $\delta > 0$ so, dass $\|g(k)\| < \varepsilon$ für $\|k\| < \delta$. Damit gilt $\|s(k)\| \leq \frac{1}{2}\|g(k)\|$. Wegen $g(k) = k + s(k)$ gilt dann nach der Dreiecksungleichung $\|g(k)\| \leq 2\|k\|$. Sei nun $h := g(k)$. Es gelten $\|s(k)\| = \|r(h)\|$ und $\|h\| \leq 2\|k\|$. Für $k \to 0$ konvergiert auch $h \to 0$. In diesem Limes gilt dann

$$\frac{s(k)}{2\|k\|} \leq \frac{r(h)}{\|h\|} \to 0,$$

wie gewünscht. □

Beispiel E8.14. Wir werden den Fixpunktsatz B8.13 von Banach benutzen, um den Umkehrsatz zu beweisen. Wir fangen mit einem einfacheren Beispiel an, wie man eine Gleichung mit dem Fixpunktsatz lösen kann. Sei A eine lineare Abbildung $A \in L(\mathbb{R}^n, \mathbb{R}^n)$ mit Operatornorm $\|A\| < 1$. Wir behaupten, $\mathrm{id} - A$ ist invertierbar, d.h., zu jedem $y \in \mathbb{R}^n$ gibt es ein eindeutiges $x \in \mathbb{R}^n$ mit

$$(\mathrm{id} - A)(x) = x - Ax = y.$$

Sei $F_y\colon \mathbb{R}^n \to \mathbb{R}^n$ die Abbildung $x \mapsto Ax + y$. Sie ist so gewählt, dass ein Fixpunkt von F_y eine Lösung zur Gleichung ist:

$$x = F_y(x) = Ax + y \iff (\mathrm{id} - A)x = x - Ax = y.$$

Wir sehen, F_y ist eine Kontraktion, weil

$$\left\|F_y(x) - F_y(x')\right\| = \left\|Ax - Ax'\right\| \leq \|A\|\,\|x - x'\|$$

mit $\|A\| < 1$. Nach dem Fixpunktsatz hat dann F_y einen eindeutigen Fixpunkt $x = Ax + y$. Das heißt, $\mathrm{id} - A$ ist invertierbar.

Der Fixpunktsatz gibt uns darüber hinaus eine rekursiv definierte Folge (x_n), die gegen den Fixpunkt konvergiert. Mit $x_0 = 0$ und $x_{n+1} = F_y(x_n) = Ax_n + y$ gilt

$$x_n = \sum_{i=0}^{n-1} A^i y \quad \to \quad x = \sum_{i=0}^{\infty} A^i y.$$

Das heißt, wir haben eine Formel für die Umkehrabbildung: $(\mathrm{id} - A)^{-1} = \sum_{i=0}^{\infty} A^i$ als Reihensumme im Banachraum $L(V,V)$.

Satz E8.15 (Umkehrsatz). *Sei $U \subset \mathbb{R}^n$ offen und sei $f\colon U \to \mathbb{R}^n$ stetig differenzierbar. Sei $p \in U$ so, dass $D_p f \in \mathsf{GL}(\mathbb{R}^n)$ invertierbar ist. Dann ist f „lokal bei p ein Diffeomorphismus", das heißt, es gibt eine offene Menge $V \subset U$ mit $V \ni p$ so, dass $f|_V\colon V \to f(V)$ ein Diffeomorphismus ist.*

Beweis. Wie im Lemma dürfen wir ohne Einschränkung annehmen, dass

$$p = 0, \quad f(0) = 0, \quad D_0 f = \mathrm{id},$$

indem wir f mit

$$x \mapsto (D_p f)^{-1}(f(p + x) - f(p))$$

ersetzen.

Weil Df stetig ist und $\mathsf{GL}(\mathbb{R}^n)$ offen ist, dürfen wir annehmen, dass $D_q f$ invertierbar ist für alle $q \in U$. (Wenn nicht, ersetzen wir U mit einer kleineren offenen Umgebung von p.)

Für gegebenes $y \in \mathbb{R}^n$ suchen wir nach einer Lösung der Gleichung $f(x) = y$. Um die Lösung x als Fixpunkt zu sehen, betrachten wir die Abbildung

$$F_y\colon U \to \mathbb{R}^n, \quad F_y(x) := x - f(x) + y.$$

Wir wollen zeigen – für jedes hinreichend kleine y –, dass F_y eine Kontraktion $K \to K$ ist, wobei $K \subset U$ kompakt ist.

Wir bemerken, $D_x F_y = \mathrm{id} - D_x f$. Weil Df stetig ist, ist

$$x \mapsto \|D_x F_y\| = \|\mathrm{id} - D_x f\|$$

eine stetige Abbildung $U \to \mathbb{R}$ mit $0 \mapsto 0$. Das heißt, es gibt $r > 0$ so, dass $\|D_x F_y\| < 1/2$ für $x \in K$, wobei

$$K := \{x \in \mathbb{R}^n : \|x\| \le 2r\}$$

eine (konvexe und kompakte) Teilmenge von U ist. Aus dem Schrankensatz E6.20 gilt dann für $x, x' \in K$

$$\big\|F_y(x) - F_y(x')\big\| \le \sup_{p \in K} \big\|D_p F_y\big\| \, \big\|x - x'\big\| \le \tfrac{1}{2}\big\|x - x'\big\|. \qquad (*)$$

(Insbesondere mit $x' = 0$ und $y = 0$ gilt $\|F_0(x)\| \le 1/2\|x\|$.) Für großes y ist natürlich F_y keine Abbildung $K \to K$. Für $\|y\| < r$ (und $x \in K$) gilt allerdings

$$\|F_y(x)\| \le \|F_0(x)\| + \|y\| \le \frac{1}{2}\|x\| + \|y\| < 2r,$$

d.h., $F_y\colon K \to K$. Diese Abbildung ist dann eine Kontraktion (mit Lipschitzkonstante $1/2$) und hat deshalb einen eindeutigen Fixpunkt $x \in K$.

Das heißt, zu jedem $y \in B_r(0)$ gibt es genau ein $x \in K$ mit $f(x) = y$. Wir setzen $V := \{x \in K : f(x) \in B_r(0)\}$ und sehen, dass $f|_V\colon V \to B_r(0)$ eine Bijektion ist.

Sei $g := (f|_V)^{-1}$. Nach $(*)$ und der Dreiecksungleichung gilt

$$\begin{aligned}\|x - x'\| - \|f(x) - f(x')\| &\le \|x - x' - f(x) + f(x')\| \\ &= \|F_0(x) - F_0(x')\| \le \tfrac{1}{2}\|x - x'\|.\end{aligned}$$

Damit folgt, dass g eine Lipschitzabbildung ist mit Lipschitzkonstante 2:

$$\|g(y) - g(y')\| \le 2\big\|f(g(y)) - f(g(y'))\big\| = 2\|y - y'\|.$$

Insbesondere ist g stetig. Damit ist $f|_V$ ein Homöomorphismus. Weil $D_x f$ invertierbar ist für jedes $x \in V \subset U$, ist nach dem Lemma dieser Homöomorphismus sogar ein Diffeomorphismus. □

Ende der Vorlesung 2009 Februar 5

E9. Implizite Funktionen und Mannigfaltigkeiten

Der Umkehrsatz untersucht die Lösbarkeit eines nichtlinearen Gleichungssystems in dem Fall, wo die Anzahl der Gleichungen mit der Anzahl der Variablen übereinstimmt und damit eine eindeutige Lösung zu erwarten ist. Wenn wir mehr Variablen als Gleichungen haben, erwarten wir mehrere Lösungen. Ist F eine Abbildung $\mathbb{R}^{n+k} \to \mathbb{R}^k$, dann erwarten wir – für ein festes $c \in \mathbb{R}^n$ – eine n-dimensionale Familie von Lösungen zu $F(z) = c$. Wenn wir $z \in \mathbb{R}^{n+k} = \mathbb{R}^n \times \mathbb{R}^k$ als $z = (x, y)$ schreiben (mit $x \in \mathbb{R}^n$ und $y \in \mathbb{R}^k$), dann können wir hoffen, diese Familie wird (mindestens lokal) durch x parametrisiert mit $y = g(x)$. Das heißt, wir suchen nach einer Auflösung der Gleichung $F(x, y) = c$ nach y als Funktion von x. Wir sagen auch, die Gleichung $F(x, y) = c$ *definiert implizit* die Funktion $y = g(x)$.

Hier betrachten wir dan festen Wert $c \in \mathbb{R}^k$ als Konstante. Indem wir F mit $F - c$ ersetzen, können und werden wir durchgehend $c = 0$ annehmen.

Beispiel E9.1. Eine lineare Abbildung $F\colon \mathbb{R}^n \times \mathbb{R}^k \to \mathbb{R}^k$ hat immer die Form

$$F(x, y) = F(x, 0) + F(0, y) =: F'(x) + F''(y),$$

wobei $F'\colon \mathbb{R}^n \to \mathbb{R}^k$ und $F''\colon \mathbb{R}^k \to \mathbb{R}^k$ linear sind. Falls F'' invertierbar ist, hat die Gleichung $F(x, y) = 0$ eine eindeutige Auflösung $y = g(x)$: es gilt nämlich

$$y = g(x) := -(F'')^{-1}(F'(x)).$$

Bemerkung E9.2. Sei $F\colon \mathbb{R}^n \times \mathbb{R}^k \to \mathbb{R}^k$ eine differenzierbare Abbildung. Wir schreiben

$$D'_{(p,q)} F(v) := D_{(p,q)} F(v, 0), \quad D''_{(p,q)} F(w) := D_{(p,q)} F(0, w).$$

Damit ist $D''_{p,q} F\colon \mathbb{R}^k \to \mathbb{R}^k$ linear.

Beispiel E9.3. Sei $n = k$ und sei $F(x,y) := -x + f(y)$ für eine stetig differenzierbare Abbildung $f\colon \mathbb{R}^n \to \mathbb{R}^n$. Die Gleichung $F(x,y) = 0$ heißt dann $x = f(y)$. Der Umkehrsatz sagt, es gibt eine lokale Auflösung der Form $y = g(x)$ (wobei natürlich $g = f^{-1}$), falls $D_q f$ invertierbar ist. Wie sieht $D_{p,q}F$ aus? Wir haben $D_{(p,q)}F(v,w) = -v + D_q f(w)$. Das heißt, die Ableitung von F bezüglich der y-Variablen ist $D''_{(p,q)}F = D_q f$; diese muss invertierbar sein, um den Umkehrsatz anzuwenden.

Beispiel E9.4. Sei $F\colon \mathbb{R}\times\mathbb{R} \to \mathbb{R}$ durch $F(x,y) = x^2+y^2-1$ definiert. Die Lösungen der Gleichung $F(x,y) = 0$ sind natürlich genau die Punkte des Einheitskreises K. Lokal können wir $y = \pm\sqrt{1-x^2} =: g_\pm(x)$ schreiben: Bei einem Punkt $(p,q) \in K$ mit $q > 0$ (bzw. $q < 0$) gilt lokal $y = g_+(x)$ (bzw. $y = g_-(x)$). Bei den beiden Punkten $(\pm 1, 0)$ lässt sich $x^2 + y^2 - 1 = 0$ nicht in der Form $y = g(x)$ auflösen. (Da gilt $\partial_y F = 0$, d.h. $D'f = 0$. Im Allgemeinen gilt $D_{(p,q)}f(v,w) = 2pv + 2qw$.)

Bemerkung E9.5. Sei $y = g(x)$ eine lokale Auflösung der Gleichung $F(x,y) = 0$. Geometrisch können wir sagen, die Niveaumenge $\{(x,y) : F = 0\}$ stimmt lokal mit dem Graphen $\{(x, g(x))\}$ von g überein. Dies ist nicht der Fall bei einem Punkt wo die Niveaumenge „vertikal" ist.

Beispiel E9.6. Sei $f(x,y) = x^2 - y^2$. Die Niveaumenge $\{f = 0\}$ besteht aus zwei Geraden $y = \pm x$. Bei $(0,0)$ lässt sie sich nicht als *einen* Graphen darstellen. Die Ableitung von f verschwindet in $(0,0)$.

Satz E9.7 (Satz über implizite Funktionen). *Sei W eine offene Teilmenge in $\mathbb{R}^n \times \mathbb{R}^k$ und sei $F\colon W \to \mathbb{R}^k$ stetig differenzierbar. Sei $(p,q) \in W$ mit $F(p,q) = 0$. Falls die lineare Abbildung*

$$D''_{(p,q)}F\colon \mathbb{R}^k \to \mathbb{R}^k$$

invertierbar ist, dann lässt sich $F(x,y) = 0$ lokal bei (p,q) nach einer stetig differenzierbaren Abbildung $y = g(x)$ auflösen. Das heißt, es gibt Umgebungen $U \subset \mathbb{R}^n$ von p und $V \subset \mathbb{R}^k$ von q mit $U \times V \subset W$ und eine stetig differenzierbare Abbildung $g\colon U \to V$ so, dass

$$F(x,y) = 0 \text{ mit } x \in U,\ y \in V \iff y = g(x) \text{ mit } x \in U.$$

Ferner ist $D''_{(x,g(x))}F$ invertierbar für jedes $x \in U$ und es gilt

$$D_x g = -(D''_{(x,g(x))}F)^{-1} \circ D'_{(x,g(x))}F.$$

Beweis. Wir wollen den Umkehrsatz auf die stetig differenzierbare Abbildung

$$f\colon W \to \mathbb{R}^n \times \mathbb{R}^k, \quad f(x,y) = (x, F(x,y))$$

anwenden. Es gilt $D_{(x,y)}f(v,w) = (v, D_{(x,y)}F(v,w))$ und damit

$$\begin{aligned} D_{(x,y)}f(v,w) = 0 &\iff v = 0 \text{ und } D_{(x,y)}F(0,w) = 0 \\ &\iff v = 0 \text{ und } D''_{(x,y)}F(w) = 0. \end{aligned}$$

Daraus folgt, dass $D_{(x,y)}f$ genau dann invertierbar ist, wenn $D''_{(x,y)}F$ invertierbar ist. (Hier gilt die obige Gleichung nur für $w = 0$.)

Insbesondere ist $D_{(p,q)}f$ invertierbar. Nach dem Umkehrsatz gibt es dann eine Umgebung $G \subset W$ von (p,q) so, dass f eine stetig differenzierbare Umkehrabbildung

$$(x,y) \mapsto (\varphi(x,y), \psi(x,y))$$

auf $f(G)$ hat. Wegen

$$(x,y) = f(\varphi(x,y), \psi(x,y)) = \big(\varphi(x,y), F(\varphi(x,y), \psi(x,y))\big)$$

gilt $\varphi(x,y) = x$. Für $(x,y) \in G$ gilt deshalb

$$\begin{aligned} F(x,y) = 0 &\iff f(x,y) = (x,0) \\ &\iff (x,y) = (x, \psi(x,0)) \iff y = \psi(x,0). \end{aligned}$$

Wir definieren $\tilde{U} := \{x : (x,0) \in G\} \subset \mathbb{R}^n$ als Schnitt von G und $g\colon \tilde{U} \to \mathbb{R}^k$ durch $g(x) := \varphi(x,0)$. Dann ist g eine Auflösung der Gleichung nach y.

Ende der Vorlesung 2009 Februar 9

Wir müssen nur ggf. eine kleinere Umgebung $U \subset \tilde{U}$ wählen, falls $\tilde{U} \times g(\tilde{U}) \not\subset G$. Dazu wählen wir zunächst $\varepsilon > 0$ hinreichend klein, dass $B_\varepsilon(p) \times B_\varepsilon(q) \subset G$; wir definieren $V := B_\varepsilon(q)$. Weil g stetig ist, gibt es $\delta \in (0, \varepsilon]$ so, dass mit $U := B_\delta(p)$ gilt $g(U) \subset V$. Es gilt auch $U \times V \subset G \subset W$ und wir haben die gewünschte Auflösung.

Um die letzte Aussage zu beweisen, sei $h : U \to G$ die Abbildung $x \mapsto (x, g(x))$. Nach dem Umkehrsatz ist f ein Diffeomorphismus auf G, insbesondere ist $D_{h(x)}f$ invertierbar für jedes $x \in U$. Wie wir oben zeigten, heißt das, $D''_{h(x)}F$ ist invertierbar. Es gilt $0 \equiv F \circ h$ und damit

$$\begin{aligned} 0 &= D_x(F \circ h)(v) = D_{h(x)}F(D_x h)(v) = D_{h(x)}F(v, D_x g(v)) \\ &= D'_{h(x)}F(v) + D''_{h(x)}F(D_x g(v)). \end{aligned}$$

Weil $D''_{h(x)}F$ invertierbar ist, folgt die gewünschte Formel für $D_x g$. □

Bemerkung E9.8. Für $k = 1$ haben wir $F\colon \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}$. Die zu untersuchende Ableitung $D''_{(p,q)}F$ ist einfach die letzte partielle Ableitung von F, die man lieber $\partial_{n+1}F(p,q) = \partial_y F(p,q)$ nennt. Ist diese ungleich Null, können wir lokal $y = g(x)$ schreiben.

Beispiel E9.9. Sei $P = t^2 - 2at + b$ ein quadratisches Polynom. Sind die Nullstellen differenzierbare Funktionen von (a,b)? Wir betrachten P als Abbildung

$$P\colon \mathbb{R}^2 \times \mathbb{R} \to \mathbb{R}, \quad P(a,b,t) = t^2 - 2at + b.$$

Der Satz über implizite Funktionen sagt, wir können $P = 0$ lokal in der Form $t = t(a,b)$ auflösen, falls

$$0 \neq \partial_3 P(a,b,t) = \frac{\partial P}{\partial t}(a,b,t) = 2t - 2a,$$

d.h., falls $t \neq a$. Aus der Schule wissen wir,

$$P(a,b,t) = t^2 - 2at + b = (t - t_+)(t - t_-),$$

wobei

$$t_\pm := t_\pm(a,b) := a \pm \sqrt{a^2 - b}.$$

Wir bemerken,

$$t = a \iff a^2 = b \iff t_+ = t_- \iff P = (t-a)^2.$$

Das heißt, der Satz über implizite Funktionen scheitert genau dann, wenn P eine doppelte Nullstelle hat.

Die Menge $\{(a,b,t) : t^2 - 2at + b = 0\}$ ist eine Fläche in $\mathbb{R}^2 \times \mathbb{R}$. Über dem Gebiet $\{(a,b) : a^2 > b\}$ in der Ebene $\mathbb{R}^2$ hat die Fläche zwei Schichten. Jede Schicht ist der Graph einer differenzierbaren Funktion $t_\pm$. Über der Parabel $\{a^2 = b\}$ wird die Fläche gefaltet; da ist die Tangentenebene vertikal.

Beispiel E9.10. Allgemeiner betrachten wir ein Polynom beliebigen Grades n als Abbildung

$$P\colon \mathbb{R}^{n+1} \times \mathbb{R} \to \mathbb{R}, \quad P(a,t) = P_a(t) = \sum_{i=0}^{n} a_i t^i.$$

Sei (a,c) eine Nullstelle von P, d.h., das Polynom P_a hat c als Nullstelle: $P_a(c) = 0$. Wir können den Satz über implizite Funktionen genau dann anwenden, wenn $\partial_t P_a(c) \neq 0$. Was ist aber diese partielle Ableitung? Nach Satz C4.9 können wir $P_a(t) = (t-c)Q(t)$ schreiben, wobei Q ein Polynom ist. Dann gilt $\partial_t P_a(c) = 0Q'(c) + 1Q(c) = Q(c)$. Das heißt, der Satz über implizite Funktionen scheitert genau dann, wenn c eine Nullstelle von Q ist, d.h., wenn c eine doppelte Nullstelle von P_a ist. Daraus schließen wir, die Nullstellen eines Polynoms sind differenzierbare Funktionen der Koeffizienten, solange jede Nullstelle eine einfache Nullstelle ist.

Eine Mannigfaltigkeit ist ein metrischer Raum, der lokal dem euklidischen Raum $\mathbb{R}^n$ gleicht. Man kann Mannigfaltigkeiten abstrakt definieren (vgl. Vorlesungen zu Differentialtopologie bzw. -geometrie). Man zeigt, jede Mannigfaltigkeit kann in einen höher-dimensionalen euklidischen Raum $\mathbb{R}^{n+k}$ eingebettet werden. Hier betrachten wir nur solche Untermannigfaltigkeiten; lokal gleichen diese dem Untervektorraum $\mathbb{R}^n \subset \mathbb{R}^{n+k}$.

Definition E9.11. Eine Teilmenge $M \subset \mathbb{R}^{n+k}$ heißt *n-dimensionale Untermannigfaltigkeit*, falls es zu jedem Punkt $p \in M$ folgendes gibt:

- eine offene Umgebung $U \subset \mathbb{R}^{n+k}$ von p,
- eine offene Menge $V \subset \mathbb{R}^{n+k}$ und
- einen Diffeomorphismus $F\colon U \to V$ so, dass

$$F(U \cap M) = V \cap (\mathbb{R}^n \times \{0\}).$$

Wir sagen auch, M hat *Kodimension* k in $\mathbb{R}^{n+k}$. Der Diffeomorphismus F heißt *Karte* für M bei p.

Beispiel E9.12. Sei $G \subset \mathbb{R}^n$ offen und sei $g\colon G \to \mathbb{R}^k$ stetig differenzierbar. Der Graph

$$\Gamma := \{(x, g(x)) : x \in G\} \subset G \times \mathbb{R}^k$$

ist dann eine n-dimensionale Untermannigfalitigkeit: unabhängig von $p \in \Gamma$ können wir $U := V := G \times \mathbb{R}^k$ und

$$F(x,y) := (x, y - g(x))$$

setzen. Auch die Umkehrabbildung $F^{-1}(x,y) = (x, y + g(x))$ ist offensichtlich stetig differenzierbar.

Beispiel E9.13. Jede offene Teilmenge $U \subset \mathbb{R}^n$ ist eine n-dimensionale Untermannigfaltigkeit von $\mathbb{R}^n$.

Definition E9.14. Sei $U \subset \mathbb{R}^{n+k}$ offen und sei $f\colon U \to \mathbb{R}^k$ stetig differenzierbar. Ein Punkt $x \in U$ heißt *kritischer Punkt*, falls $D_x f\colon \mathbb{R}^{n+k} \to \mathbb{R}^k$ nicht surjektiv ist. Ein Punkt $y \in \mathbb{R}^k$ heißt *kritischer Wert*, falls es einen kritischen Punkt $x \in U$ gibt mit $y = f(x)$; sonst heißt $y \in \mathbb{R}^k$ *regulärer Wert*.

Beispiele E9.15.

- Für eine konstante Abbildung $f(x) \equiv c \in \mathbb{R}^k$ $(k > 0)$ ist jeder Punkt $x \in U$ kritisch. Jeder Wert $y \neq c$ ist regulär, während c ein kritischer Wert ist.
- Für $f\colon \mathbb{R}^{n+1} \to \mathbb{R}$, $f(x) := \|x\|^2 = \langle x, x\rangle$ ist $x = 0$ der einzige kritische Punkt. Damit ist $y = 0$ der einzige kritische Wert.
- Ein Wert, der nicht angemonnen wird ($y \notin f(U)$), ist regulär.
- Für $k = 1$ ist x genau dann ein reglärer Punkt, wenn $\operatorname{grad}_x f \neq 0$.

Der Satz über implizite Funktionen sagt, die Niveaumenge in einem regulären Wert ist eine Untermannigfaltigkeit. Dazu brauchen wir ein Lemma aus der linearen Algebra.

Lemma E9.16. *Sei $F\colon \mathbb{R}^{n+k} \to \mathbb{R}^k$ eine lineare und surjektive Abbildung. Nach einer Permutation (Umordnung) der Koordinaten in $\mathbb{R}^{n+k}$ können wir $F(x,y) = F'(x) + F''(y)$ schreiben, wobei $F''\colon \mathbb{R}^k \to \mathbb{R}^k$ bijektiv ist.*

Satz E9.17. *Eine Teilmenge $M \subset \mathbb{R}^{n+k}$ ist genau dann eine n-dimensionale Untermannigfaltigkeit, wenn es zu jedem Punkt $p \in M$ eine offene Umgebung $U \subset \mathbb{R}^{n+k}$ und eine stetig differenzierbare Abbildung $h\colon U \to \mathbb{R}^k$ so gibt, dass $M \cap U = h^{-1}(c)$, wobei $c \in \mathbb{R}^k$ ein regulärer Wert von h ist.*

Beweis. „$\Longrightarrow$“: Sei $p \in M$ und sei $F\colon U \to \mathbb{R}^{n+k}$ eine Karte für M bei p. Wir schreiben $F =: (f,h)$, wobei $h\colon U \to \mathbb{R}^k$. Dann gilt $M \cup U = h^{-1}(0)$. Weil F ein Diffeomorphismus ist, ist $D_x F$ in jedem Punkt x surjektiv, daher auch $D_x h$.

„$\Longleftarrow$“: Sei $p \in M$. Weil p ein regulärer Punkt von h ist, können wir die Koordinaten nach dem Lemma so umordnen, dass $D''_p h$ surjektiv ist. Damit können wir den Satz über implizite Funktionen anwenden: lokal ist die Niveaumenge $h^{-1}(c)$ der Graph einer Funktion $g\colon U' \to \mathbb{R}^k$. Nach Beispiel E9.12 ist dieser Graph eine Untermannigfaltigkeit. □

Korollar E9.18. *Sei c ein regulärer Wert von $h\colon U \to \mathbb{R}^k$. Dann ist die Niveaumenge $h^{-1}(c) \subset U \subset \mathbb{R}^{n+k}$ eine Untermannigfaltigkeit.*

Beweis. Zu jedem $p \in M$ benutzen wir im Satz die gegebenen U und h. □

Beispiel E9.19. Die Einheitssphäre $\mathbb{S}^n \subset \mathbb{R}^{n+1}$ ist eine Untermannigfaltigkeit.