Skriptum zur Vorlesung

Analysis II

Hans-Dieter Alber

zuletzt geändert am 05.09.2014

Inhaltsverzeichnis

1 Stetige Abbildungen auf dem $\mathbb{R}^n$

1 a.) Normen auf dem $\mathbb{R}^n$

Sei $n \in \mathbb{N}$. Auf der Menge aller n–Tupel reeller Zahlen

$$\left\{ x := (x_1, x_2, \ldots, x_n) \,\middle|\, x_i \in \mathbb{R},\ \ i = 1, \ldots, n \right\}$$

kann man eine Addition und eine Multiplikation mit reellen Zahlen, (Skalaren), einführen, so daß diese Menge ein reeller Vektorraum wird:

$$\begin{aligned} x + y &:= (x_1 + y_1, \ldots, x_n + y_n) \\ cx &:= (cx_1, \ldots, cx_n)\,. \end{aligned}$$

Man bezeichnet diesen Vektorraum mit $\mathbb{R}^n$. Er hat die Dimension n. Eine Basis bilden zum Beispiel die n Vektoren

$$e_1 = (1, 0, \ldots, 0),\ \ e_2 = (0, 1, 0, \ldots, 0),\ \ldots,\ \ e_n = (0, \ldots, 0, 1)\,.$$

Ist V ein beliebiger Vektorraum, dann heißt eine Abbildung $\|\cdot\| : V \to \mathbb{R}_0^+$ Norm auf V, wenn für alle $x, y \in V$ und $c \in \mathbb{R}$

(i) $\|x\| = 0 \iff x = 0$

(ii) $\|cx\| = |c|\,\|x\|$

(iii) $\|x + y\| \leq \|x\| + \|y\|$ (Dreiecksungleichung)

gilt. Ein Vektorraum, auf dem eine Norm erklärt ist, heißt normierter Raum. Auf dem Vektorraum $\mathbb{R}^n$ kann man Normen auf verschiedene Weise erklären. Zwei wichtige Beispiele will ich betrachten:

1.) Die Maximumsnorm:

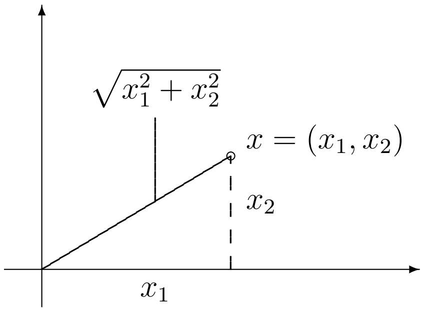
$$\|x\|_\infty := \max\{|x_1|, \ldots, |x_n|\}\,.$$

Um zu zeigen, daß dies eine Norm ist, müssen die Eigenschaften (i) – (iii) nachgewiesen werden. Die Eigenschaften (i) und (ii) sind erfüllt. Es bleibt (iii) zu zeigen.
Es gibt ein i mit $\|x + y\|_\infty = |x_i + y_i|$. Damit folgt

$$\|x + y\|_\infty = |x_i + y_i| \leq |x_i| + |y_i| \leq \|x\|_\infty + \|y\|_\infty\,.$$

2.) Die Euklidsche Norm:

$$|x| := \sqrt{x_1^2 + \ldots + x_n^2}$$



Mit dem **Skalarprodukt**

$$x \cdot y := x_1 \cdot y_1 + \ldots + x_n \cdot y_n$$

gilt hierfür auch

$$|x| = \sqrt{x \cdot x}\,.$$

(i) und (ii) sind klar. Es gleibt (iii) zu zeigen. Hierzu beweist man zuerst die **Cauchy–Schwarzsche Ungleichung**

$$|x \cdot y| \leq |x|\,|y|\,.$$

Beweis der Cauchy–Schwarzschen Ungleichung: Das quadratische Polynom in t

$$|x|^2 t^2 + 2x \cdot yt + |y|^2 = |tx + y|^2 \geq 0$$

kann keine zwei verschiedenen reellen Nullstellen haben, also muß für die Diskriminante

$$(x \cdot y)^2 - |x|^2\,|y|^2 \leq 0$$

gelten. ■

Beweis der Eigenschaft (iii):

$$\begin{aligned} |x+y|^2 &= (x+y) \cdot (x+y) = |x|^2 + 2x \cdot y + |y|^2 \\ &\leq |x|^2 + 2|x \cdot y| + |y|^2 \\ &\leq |x|^2 + 2|x|\,|y| + |y|^2 = (|x| + |y|)^2\,, \end{aligned}$$

also

$$|x + y| \leq |x| + |y|\,.$$

■

Definition: Sei $\|\cdot\|$ eine Norm auf $\mathbb{R}^n$. Eine Folge $\{x_k\}_{k=1}^{\infty}$, $x_k \in \mathbb{R}^n$, heißt konvergent, wenn ein $a \in \mathbb{R}^n$ existiert mit

$$\lim_{k \to \infty} \|x_k - a\| = 0\,.$$

Man schreibt dann $\lim_{k\to\infty} x_k = a$ und bezeichnet a als Grenzwert oder Grenzelement der Folge $\{x_k\}_{k=1}^{\infty}$.

Wie im Fall des $\mathbb{R}^1$ beweist man, daß eine Folge nicht gegen zwei verschiedene Grenzwerte konvergieren kann.

Zur Definition der Konvergenz wird eine Norm benötigt. Trotzdem hängt der Konvergenzbegriff auf dem $\mathbb{R}^n$ nicht von der verwendeten Norm ab. Dies ergibt sich aus den folgenden Resultaten.

Lemma: Eine Folge $\{x_k\}_{k=1}^{\infty}$, $x_k \in \mathbb{R}^n$, konvergiert genau dann bezüglich der Maximumsnorm, wenn jede der n Komponentenfolgen $\{x_k^{(i)}\}_{k=1}^{\infty}$, $i = 1, \ldots, n$, konvergiert.

Beweis: Es gilt

$$|x_k^{(i)} - a^{(i)}| \leq \|x_k - a\|_\infty \leq |x_k^{(1)} - a^{(1)}| + \ldots + |x^{(n)} - a^{(n)}| \,.$$

■

Satz: Sei $\{x_k\}_{k=1}^{\infty}$, $x_k \in \mathbb{R}^n$, eine bezüglich der Maximumsnorm beschränkte Folge. (D. h. es existiert $c > 0$ mit $\|x_k\|_\infty \leq c$ für alle $k \in \mathbb{N}$.) Dann besitzt $\{x_k\}_{k=1}^{\infty}$ eine (bezüglich der Maximumsnorm) konvergente Teilfolge.

Beweis: Da jede der Komponentenfolgen $\{x_k^{(i)}\}_{k=1}^{\infty}$, $i = 1, \ldots, n$, beschränkt ist, besitzt jede dieser Folgen eine konvergente Teilfolge. Sei $\{x_{k(j)}^{(1)}\}_{j=1}^{\infty}$ eine konvergente Teilfolge von $\{x_k^{(1)}\}_{k=1}^{\infty}$. Dann ist $\{x_{k(j)}^{(2)}\}_{j=1}^{\infty}$ eine Teilfolge von $\{x_k^{(2)}\}_{k=1}^{\infty}$, und besitzt eine konvergente Teilfolge $\{x_{k(j(\ell))}^{(2)}\}_{\ell=1}^{\infty}$. Auch $\{x_{k(j(\ell))}^{(1)}\}_{\ell=1}^{\infty}$ ist konvergent als Teilfolge von $\{x_{k(j)}^{(1)}\}_{j=1}^{\infty}$. Setzt man das Verfahren fort so erhält man nach n–Schritten eine Teilfolge $\{x_{k_s}\}_{s=1}^{\infty}$ von $\{x_k\}_{k=1}^{\infty}$, für die alle Komponentenfolgen konvergieren, und die also konvergent ist bezüglich der Maximumsnorm im $\mathbb{R}^n$. ■

Satz: Seien $\|\cdot\|_1$ und $\|\cdot\|_2$ Normen auf $\mathbb{R}^n$. Dann existieren Konstanten $a, b > 0$ so daß für alle $x \in \mathbb{R}^n$ gilt

$$a\|x\|_1 \leq \|x\|_2 \leq b\|x\|_1 \,.$$

Beweis: Sei $\|\cdot\|$ eine Norm auf $\mathbb{R}^n$ und $\|\cdot\|_\infty$ die Maximumsnorm. Es genügt zu zeigen, daß Konstanten $a, b > 0$ existieren mit

$$\|x\| \leq a\|x\|_\infty \,, \quad \|x\|_\infty \leq b\|x\|$$

für alle $x \in \mathbb{R}^n$. Die erste Abschätzung ergibt sich folgendermaßen:

$$\begin{aligned}
\|x\| &= |x_1 e_1 + x_2 e_2 + \ldots + x_n e_n| \\
&\leq \|x_1 e_1\| + \ldots + \|x_n e_n\| = |x_1|\,\|e_1\| + \ldots + |x_n|\,\|e_n\| \\
&\leq (\|e_1\| + \ldots + \|e_n\|)\,\|x\|_\infty = a\|x\|_\infty\,,
\end{aligned}$$

mit $a = \|e_1\| + \ \ldots + \|e_n\|$.

Die zweite Abschätzung beweist man durch Widerspruch: Angenommen, es gäbe keine solche Konstante $b > 0$. Dann könnte man für alle $k \in \mathbb{N}$ ein $x_k \in \mathbb{R}^n$ finden mit

$$\|x_k\|_\infty > k\|x_k\|\,.$$

Setze $y_k := \frac{x_k}{\|x_k\|_\infty}$. Für die Folge $\{y_k\}_{k=1}^\infty$ gelten

$$\|y_k\| = \|\,\frac{x_k}{\|x_k\|_\infty}\,\| = \frac{1}{\|x_k\|_\infty}\,\|x_k\| < \frac{1}{k}$$

und

$$\|y_k\|_\infty = \left\|\,\frac{x_k}{\|x_k\|_\infty}\,\right\|_\infty = 1\,.$$

Nach dem Satz von Bolzano–Weierstraß hat $\{y_k\}_{k=1}^\infty$ somit eine Teilfolge $\{z_j\}_{j=1}^\infty$, $\ z_j = y_{kj}$, die bezüglich der Maximumsnorm konvergiert. Sei $z \in \mathbb{R}^n$ der Grenzwert. Dann gilt

$$\lim_{k\to\infty} \|z_k - z\|_\infty = 0\,,$$

also, wegen $\|z_k\|_\infty = 1$,

$$1 = \lim_{k\to\infty} \|z_k\|_\infty = \lim_{k\to\infty} \|z_k - z + z\|_\infty \leq \|z\|_\infty + \lim_{k\to\infty} \|z_k - z\|_\infty = |z|$$

also $z \neq 0$.
Andererseits gilt $\|z_j\| = \|y_{k_j}\| < \frac{1}{k_j} \leq \frac{1}{j}$, also

$$\begin{aligned}
\|z\| &= \|z - z_k + z_k\| = \lim_{k\to\infty} \|z - z_k + z_k\| \\
&\leq \lim_{k\to\infty} \|z - z_k\| + \lim_{k\to\infty} \|z_k\| \\
&\leq a \lim_{k\to\infty} \|z - z_k\|_\infty + \lim_{k\to\infty} \frac{1}{k} = 0\,,
\end{aligned}$$

also $z = 0$. Widerspruch! ■

Wenn zwei Normen die Ungleichungen des eben bewiesenen Satzes erfüllen, sagt man, sie seien **äquivalent.** Auf dem $\mathbb{R}^n$ sind also alle Normen äquivalent. Hieraus ergibt sich unmittelbar, daß eine Folge, die bezüglich einer Norm gegen a konvergiert, auch bezüglich jeder anderen Norm gegen a konvergiert. Folglich hängt der Konvergenzbegriff nicht von der verwendeten Norm ab. Außerdem ergibt sich sofort , daß das obenstehende Lemma und der Satz nicht nur für die Maximumsnorm, sondern für alle Normen gelten.

Lemma: (Cauchysches Konvergenzkriterium) Sei $\|\cdot\|$ eine Norm auf $\mathbb{R}^n$. Eine Folge $\{x_k\}_{k=1}^{\infty}$, $x_k \in \mathbb{R}^n$, ist konvergent, genau dann wenn zu jedem $\varepsilon > 0$ ein $k_0 \in \mathbb{N}$ existiert mit

$$\|x_k - x_\ell\| < \varepsilon$$

für alle $k, \ell \geq k_0$.

Beweis: $\{x_k\}_{k=1}^{\infty}$ ist eine Cauchyfolge im $\mathbb{R}^n$, genau dann wenn jede der Komponentenfolgen $\{x_k^{(i)}\}_{k=1}^{\infty}$ eine Cauchyfolge im $\mathbb{R}^1$ ist. Denn es existieren Konstanten $a, b > 0$ mit

$$a|x_k^{(i)} - x_\ell^{(i)}| \leq \|x_k - x_\ell\| \leq b\Big(|x_k^{(1)} - x_\ell^{(1)}| + \ldots + |x_k^{(n)} - x_\ell^{(n)}|\Big), \quad i = 1, \ldots, n.$$

Also folgt die Aussage des Lemmas, aus dem Cauchyschen Konvergenzkriterium im $\mathbb{R}^1$.

■

Unendliche Reihen: Unter einer unendlichen Reihe $\sum_{k=1}^{\infty} x_k$, $x_k \in \mathbb{R}^n$, versteht man die Folge $\{s_\ell\}_{\ell \in \mathbb{N}}$ der Partialsummen $s_\ell = \sum_{k=1}^{\ell} x_k$. Falls $s = \lim_{\ell \to \infty} s_\ell$ existiert, heißt s die Summe der Reihe: $s = \sum_{k=1}^{\infty} x_k$. Eine Reihe heißt absolut konvergent, wenn

$$\sum_{k=1}^{\infty} \|x_k\|$$

konvergiert. Aus

$$\|\sum_{k=\ell}^{m} x_m\| \leq \sum_{k=\ell}^{m} \|x_k\|$$

und dem Cauchyschen Konvergenzkriterium folgt, daß eine absolut konvergente Reihe auch konvergiert. Die Umkehrung gilt im allgemeinen nicht. Eine Reihe ist genau dann absolut konvergent, wenn jede ihrer Komponentenreihen absolut konvergiert. Hieraus folgt, daß eine absolut konvergente Reihe bei jeder Umordnung gegen dieselbe Summe konvergiert, da dies für die Komponentenreihen richtig ist.

1 b.) Topologie des $\mathbb{R}^n$

Sei $\|\cdot\|$ eine Norm auf $\mathbb{R}^n$.

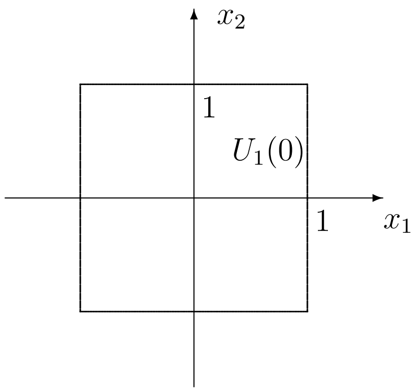
Definition: Sei $a \in \mathbb{R}^n$, $\varepsilon > 0$. Die Menge

$$U_\varepsilon(a) = \{x \in \mathbb{R}^n | \, \|x - a\| < \varepsilon\}$$

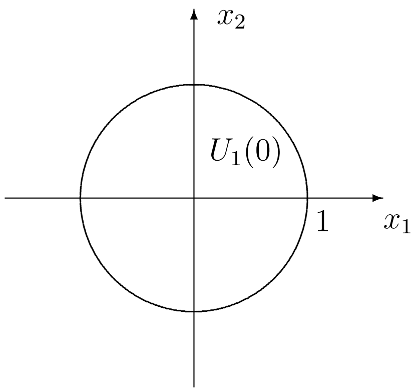
heißt offene ε–Umgebung von a bezüglich der Norm $\|\cdot\|$, oder Kugel um a mit Radius ε. Eine Teilmenge U von $\mathbb{R}^n$ heißt Umgebung von $a \in \mathbb{R}^n$, wenn U eine ε–Umgebung von a enthält.
Die Menge $U_1(0) = \{x \in \mathbb{R}^n | \, \|x\| < 1\}$ heißt offene Einheitskugel bezüglich der Norm $\|\cdot\|$.

Im $\mathbb{R}^2$ kann man sich die Form der „Einheitskugel“ für verschiedene Normen veranschaulichen:

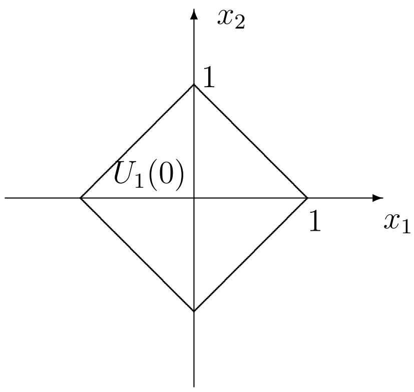
Maximumsnorm: $\|\cdot\|_\infty$



Euklidsche Norm: $|\cdot|$



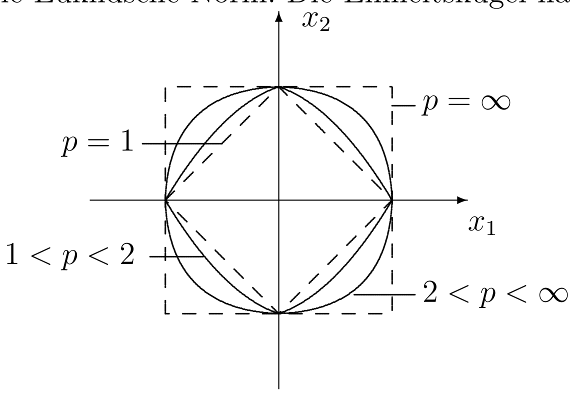
Auch $\|x\|_1 := |x_1| + |x_2| + \ldots + |x_n|$ ist eine Norm. Der Beweis wird dem Leser überlassen. Für diese Norm hat die Einheitskugel folgende Form



Allgemein wird für jedes $p \geq 1$ eine Norm definiert durch

$$\|x\|_p := (|x_1|^p + \ldots + |x_n|^p)^{1/p} .$$

Für $p = 2$ ergibt sich die Euklidsche Norm. Die Einheitskugel hat folgende Form



Der Umgebungsbegriff hängt nicht von der verwendeten Norm ab. Denn seien $\|\cdot\|_1$, $\|\cdot\|_2$ Normen auf $\mathbb{R}^n$, sei $a \in \mathbb{R}^n$. Aus der Äquivalenz der Normen

$$c_1\|x\|_1 \leq \|x\|_2 \leq c_2\|x\|_1$$

folgt, daß jede ε–Umgebung $U_\varepsilon(a) = \{x \big| \, \|x-a\|_1 < \varepsilon\}$ bezüglich $\|\cdot\|_1$ die $c_1\varepsilon$–Umgebung

$$V_{c_1\varepsilon}(a) = \Big\{x \Big| \, \|x-a\|_2 < c_1\varepsilon\Big\}$$

bezüglich $\|\cdot\|_2$ enthält. Also ist $U_\varepsilon(a)$ Umgebung von a bezüglich der $\|\cdot\|_2$ Norm, und somit ist jede Umgebung von a bezüglich $\|\cdot\|_1$ auch Umgebung bezüglich $\|\cdot\|_2$ und umgekehrt.

Definition: Sei $M \subseteq \mathbb{R}^n$. Ein Punkt $x \in \mathbb{R}^n$ heißt innerer Punkt von M , wenn es eine Umgebung von x gibt, die in M enthalten ist, d.h. wenn M Umgebung von x ist.

$x \in \mathbb{R}^n$ heißt Häufungspunkt von M , wenn in jeder Umgebung von x ein Punkt von M liegt, der von x verschieden ist.

$x \in \mathbb{R}^n$ heißt Randpunkt von M, wenn in jeder Umgebung von x mindestens ein Punkt von M und ein Punkt des Komplementes $\mathbb{R}^n \backslash M$ liegt.

Eine Menge heißt offen, wenn sie nur aus inneren Punkten besteht. Eine Menge heißt abgeschlossen, wenn sie alle ihre Häufungspunkte enthält.

Wie in $\mathbb{R}^1$ beweist man:
Das Komplement einer offenen Menge ist abgeschlossen, das Komplement einer abgeschlossenen Menge ist offen. Die Vereinigung eines beliebigen Systems offener Mengen ist offen, der Durchschnitt von endlich vielen offenen Mengen ist offen. Der Durchschnitt eines beliebigen Systems abgeschlossener Mengen ist abgeschlossen, die Vereinigung von endlich vielen abgeschlossenen Mengen ist abgeschlossen.

Sei $M \subseteq \mathbb{R}^n$ beschränkt und $\|\cdot\|$ eine Norm auf $\mathbb{R}^n$. Als Durchmesser von M bezeichnet man die Zahl

$$\delta(M) := \sup_{y,x \in M} \|y - x\| .$$

Satz: Sei $\{A_k\}_{k=1}^{\infty}$ eine Folge von abgeschlossenen nichtleeren Mengen $A_k \subseteq \mathbb{R}^n$ mit $A_{k+1} \subseteq A_k$ und

$$\lim_{k \to \infty} \delta(A_k) = 0 .$$

Dann existiert $x \in \mathbb{R}^n$ mit

$$\bigcap_{k=1}^{\infty} A_k = \{x\} .$$

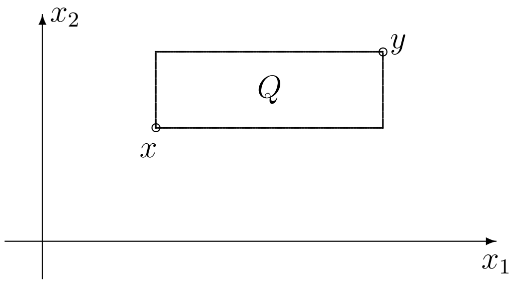
Beweis: Wähle $x_k \in A_k$. Die Folge $\{x_k\}_{k=1}^{\infty}$ ist dann eine Cauchyfolge, weil $x_{k+\ell} \subseteq A_{k+\ell} \subseteq A_k$, also $\|x_k - x_{k+\ell}\| \leq \delta(A_k) \to 0$.
Der Grenzwert von $\{x_k\}_{k=1}^{\infty}$ sei x. Es gilt $x \in \bigcap_{k=1}^{\infty} A_k$. Denn wenn $k \in \mathbb{N}$ existieren würde mit $x \notin A_k$, dann würde eine Umgebung $U_\varepsilon(x)$ existieren mit $U_\varepsilon(x) \cap A_k = \emptyset$, also auch $U_\varepsilon(x) \cap A_\ell = \emptyset$ für alle $\ell \geq k$, wegen $A_\ell \subseteq A_k$, also $\|x - x_\ell\| \geq \varepsilon$ für alle $\ell \geq k$, im Widerspruch zur Annahme daß x Grenzwert von $\{x_k\}_{k=1}^{\infty}$ sei.

Sei $y \in \bigcap_{k=1}^{\infty} A_k$. Dann gilt $\|y - x\| \leq \delta(A_k)$ für alle k, also $\|y - x\| = \lim_{k\to\infty} \|y - x\| \leq \lim_{k\to\infty} \delta(A_k) = 0$, also $x = y$. Somit gilt

$$\bigcap_{k=1}^{\infty} A_k = \{x\} .$$

$\blacksquare$

Kompakte Mengen. Kompakte Mengen werden im $\mathbb{R}^n$ wie im $\mathbb{R}^1$ definiert und es gelten entsprechende Aussagen. Um dies zu zeigen, müssen noch zwei weitere Begriffe eingeführt werden:
Seien $x, y \in \mathbb{R}^n$. Dann bezeichnet man die Menge $\{z \in \mathbb{R}^n \,\big|\, x_i \leq z_i \leq y_i\,,\ \ i = 1, \dots, n\}$ als abgeschlossenen Quader. Ist $y_1 - x_1 = y_2 - x_2 = \ldots = y_n - x_n = a \geq 0$, dann bezeichnet man diese Menge als Würfel mit Kantenlänge a .



Sei $M \subseteq \mathbb{R}^n$. Ein Mengensystem $\mathcal{U}$ aus offenen Mengen im $\mathbb{R}^n$ mit $M \subseteq \bigcap_{U \in \mathcal{U}} U$ heißt offene Überdeckung von M .

Satz: Sei $M \subseteq \mathbb{R}^n$. Die folgenden drei Aussagen sind äquivalent:

(1) M ist beschränkt und abgeschlossen

(2) Sei $\mathcal{U}$ eine offene Überdeckung von M . Dann gibt es endlich viele Mengen $U_i \in \mathcal{U}\,,\ \ i = 1, \dots, m$ mit $M \subseteq \bigcup_{i=1}^m U_i$

(3) Jede unendliche Teilmenge von M besitzt einen Häufungspunkt in M .

Beweis: (1) $\Rightarrow$ (2): Angenommen, es gäbe eine offene Überdeckung $\mathcal{U}$, so daß (2) nicht richtig ist. Da M beschränkt ist, kann M in einen abgeschlossenen Würfel W eingeschlossen werden. Man unterteile diesen Würfel nun in 2^n Würfel mit der halben Kantenlänge. Nach Annahme gilt für wenigstens einen dieser kleineren Würfel, daß sein Durchschnitt mit M nicht durch endlich viele Mengen aus $\mathcal{U}$ überdeckt werden kann. W_1 sei einer dieser Würfel. Nun unterteile man W_1 und konstruiere analog W_2 . Dies gibt eine Folge $\{W_k\}_{k=1}^\infty$ von abgeschlossenen Würfeln mit

1.) $W \supseteq W_1 \supseteq W_2 \supseteq \dots$

2.) $\lim\limits_{k \to \infty} \delta(W_k) = 0$

3.) $M \cap W_k$ kann nicht durch endlich viele Mengen aus $\mathcal{U}$ überdeckt werden.

Die Folge $\{M \cap W_k\}_{k=1}^\infty$ erfüllt die Voraussetzungen des vorangehenden Satzes, also existiert $x \in \mathbb{R}^n$ mit

$$x \in \bigcap_{k=1}^{\infty} (M \cap W_k)\,.$$

Wegen $x \in M$ gibt es $U \in \mathcal{U}$ mit $x \in U$. U ist offen, enthält also eine ε–Umgebung von x . Wegen $\lim_{k\to\infty} \delta(W_k) = 0$ und $x \in W_k$ ist daher $M \cap W_k \subseteq W_k \subseteq U$ für alle genügend großen k , also genügt zur Überdeckung von $M \cap W_k$ bereits eine einzige Menge $U \in \mathcal{U}$, im Widerspruch zu 3.). Also muß (2) richtig sein.

(2) $\Rightarrow$ (3): Sei $A \subseteq M$ eine Menge, die keine Häufungspunkte in M besitzt. Dann ist kein Punkt von M Häufungspunkt von A , also gibt es zu jedem $x \in M$ eine offene Umgebung, die außer möglicherweise x selbst keinen Punkt aus A enthält. Die Menge dieser Umgebungen bildet eine offene Überdeckung von M , also genügen endlich viele derartige Umgebungen zur Überdeckung von M . Da jede dieser Umgebungen höchstens einen Punkt aus A enthält, muß A endlich sein. Eine unendliche Teilmenge A muß also einen Häufungspunkt in M haben.

(3) $\Rightarrow$ (1): Wäre M nicht beschränkt, dann gäbe es zu jedem $k \in \mathbb{N}$ ein $x_k \in M$ mit

$$\|x_k\| \geq k\,.$$

Sei A die Menge dieser Punkte. A ist unendliche Teilmenge von M , hat aber keinen Häufungspunkt. Denn wäre y Häufungspunkt, dann müßte es unendlich viele $x \in A$ geben mit

$$\|x - y\| < 1\,, \quad \text{also} \quad \|x\| = \|x - y + y\| \leq 1 + \|y\|\,,$$

im Widerspruch zur Konstruktion von A . Also muß M beschränkt sein.

Wäre M nicht abgeschlossen, dann gäbe es einen Häufungspunkt $x \in \mathbb{R}^n$ von M , der nicht zu M gehört, und es gäbe zu jedem $k \in \mathbb{N}$ ein $x_k \in M$ mit $\|x_k - x\| < \frac{1}{k}$. Die Folge $\{x_k\}_{k=1}^{\infty}$ konvergiert gegen x , hat also x als einzigen Häufungspunkt, also müßte x nach Voraussetzung zu M gehören, im Widerspruch zur Annahme. ■

Definition: Eine Menge $M \subseteq \mathbb{R}^n$ heißt kompakt, wenn sie eine der drei Eigenschaften (und damit alle drei Eigenschaften) des vorangehenden Satzes hat.

Wie im $\mathbb{R}^1$ beweist man:

Satz: Eine Menge $M \subseteq \mathbb{R}^n$ ist kompakt, genau dann wenn jede Folge in M eine konvergente Teilfolge besitzt mit Grenzwert in M .

Mengen mit dieser Eigenschaft heißen **folgenkompakt**. Im $\mathbb{R}^n$ sind also Kompaktheit und Folgenkompaktheit äquivalente Begriffe. Klar ist auch:

Satz: Jede beschränkte unendliche Teilmenge $M \subseteq \mathbb{R}^n$ hat mindestens einen Häufungspunkt.

1 c.) Stetige Abbildungen vom $\mathbb{R}^n$ in den $\mathbb{R}^m$

Sei $D \subseteq \mathbb{R}^n$ eine Menge. Ich betrachte nun Abbildungen $f : D \to \mathbb{R}^m$. Derartige Abbildungen heißen „Funktionen von n Variablen“.

Für $x \in D$ seien $f_1(x), f_2(x), \ldots, f_m(x)$ die Komponenten von $f(x) \in \mathbb{R}^m$. Hierdurch werden Abbildungen

$$f_i : D \to \mathbb{R}\,, \quad i = 1, \ldots, m$$

definiert. Umgekehrt seien m Abbildungen $f_1, \ldots, f_m : D \to \mathbb{R}$ gegeben. Dann wird durch

$$x \mapsto f(x) := \Big(f_1(x), \ldots, f_m(x)\Big)$$

eine Abbildung

$$f : D \to R^m$$

definiert. Jede Abbildung $f : D \to \mathbb{R}^m$, $D \subseteq \mathbb{R}^n$, wird also beschrieben durch m – Abbildungsgleichungen

$$\left.\begin{array}{rcl} y_1 & = & f_1(x_1, \ldots, x_n) \\ \vdots & & \\ y_m & = & f_m(x_1, \ldots, x_n) \end{array}\right. , \quad x \in D\,.$$

Beispiele:

1.) Sei $f : \mathbb{R}^n \to \mathbb{R}^m$ und es gelte für alle $x, y \in \mathbb{R}^n$ und $c \in \mathbb{R}$

$$\begin{aligned} f(x+y) &= f(x) + f(y) \\ f(cx) &= cf(x)\,. \end{aligned}$$

Dann heißt f linear. In der linearen Algebra wird gezeigt, daß zu jeder linearen Abbildung $f : \mathbb{R}^n \to \mathbb{R}^m$ eine eindeutige Matrix

$$A = \begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & & \\ a_{m1} & \ldots & a_{mn} \end{pmatrix}, \quad a_{ij} \in \mathbb{R}\,,$$

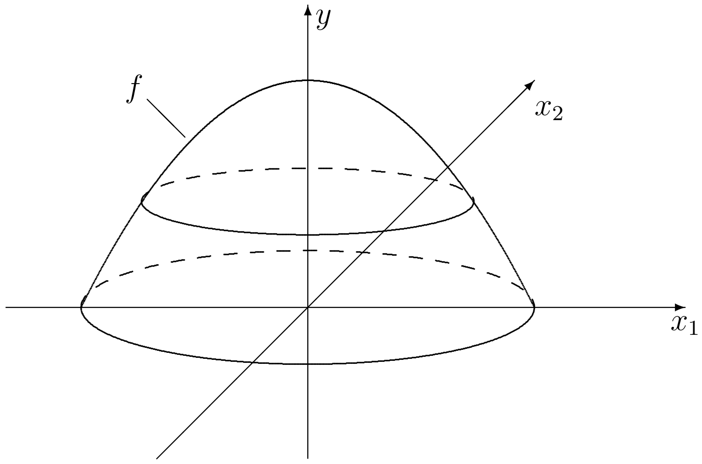
existiert mit

$$f(x) = Ax = \begin{pmatrix} a_{11}x_1 & + \ldots + & a_{1n}x_n \\ \vdots & & \\ a_{m1}x_1 & + \ldots + & a_{mn}x_n \end{pmatrix}.$$

2.) Sei $n = 2$, $m = 1$, $D = \{x \in \mathbb{R}^2 \,\big|\, |x| < 1\}$. Sei

$$f : D \to \mathbb{R},$$

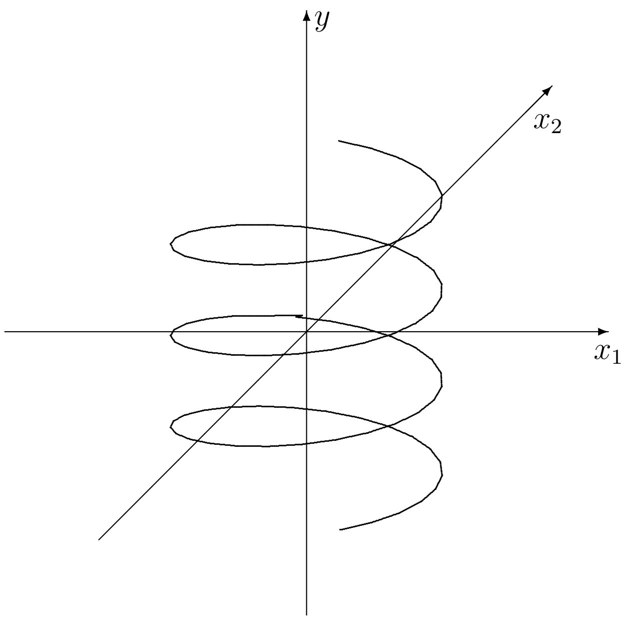
$$f(x) = \sqrt{1 - x_1^2 - x_2^2}$$



„oberer Teil der Einheitsphäre“.

3.) Jede Abbildung $f : \mathbb{R} \to \mathbb{R}^m$ heißt Weg im $\mathbb{R}^m$. Zum Beispiel sei für $t \in \mathbb{R}$

$$t \mapsto f(t) = \begin{pmatrix} f_1(t) \\ f_2(t) \\ f_3(t) \end{pmatrix} := \begin{pmatrix} \cos t \\ \sin t \\ t \end{pmatrix}$$



„Spiralfeder“

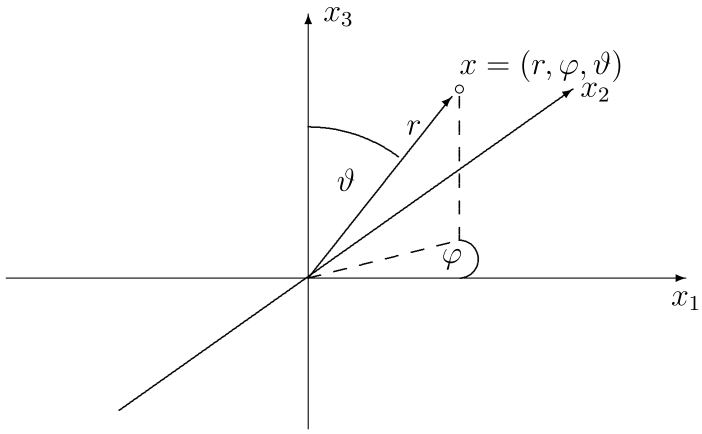
4.) (Polarkoordinaten). Sei

$$D = \left\{ (r, \varphi, \psi) \in \mathbb{R}^3 \,\middle|\, 0 < r\,,\ \ 0 \le \varphi < 2\pi\,,\ \ 0 < \vartheta < \pi \right\},$$

und sei $f : D \to \mathbb{R}^3$,

$$f(r, \varphi, \psi) = \begin{pmatrix} r \cos\varphi \sin\vartheta \\ r \sin\varphi \sin\vartheta \\ r \cos\vartheta \end{pmatrix}.$$

Der Wertebereich dieser Abbildung ist $\mathbb{R}^3$ ohne die x_3–Achse.



Definition: Sei $D \subseteq \mathbb{R}^n$. Eine Abbildung $f : D \to \mathbb{R}^m$ heißt stetig in $a \in D$, wenn es zu jedem $\varepsilon > 0$ ein $\delta > 0$ gibt mit

$$\|f(x) - f(a)\| < \varepsilon$$

für alle $x \in D$ mit $\|x - a\| < \delta$.

Beachte, daß die beiden Normen im $\mathbb{R}^n$ und im $\mathbb{R}^m$ mit demselben Symbol $\|\cdot\|$ bezeichnet wurden. Es kommt bei dieser Definition nicht darauf an, welche Norm verwendet wird. Fast alle Sätze über reelle stetige Funktionen übertragen sich auf stetige Funktionen vom $\mathbb{R}^n$ in den $\mathbb{R}^m$ mit denselben Beweisen. Ich will einige Dinge kurz wiederholen.

Satz: f ist genau dann stetig an a, wenn es zu jeder Umgebung V des Punktes $f(a)$ eine Umgebung U von a gibt mit $f(U \cap D) \subseteq V$.

Satz: f ist genau dann stetig in $a \in D$, wenn für jede Folge $\{x_k\}_{k=1}^\infty$, $x_k \in D$, mit $\lim_{k\to\infty} x_k = a$ gilt

$$\lim_{k\to\infty} f(x_k) = f(a).$$

Definition: Sei $f : D \to \mathbb{R}^n$ und sei a Häufungspunkt von D. (Es muß nicht notwendig $a \in D$ sein.) Man schreibt

$$\lim_{x\to a} f(x) = b\,,$$

wenn zu jedem $\varepsilon > 0$ ein $\delta > 0$ existiert mit

$$\|f(x) - b\| < \varepsilon$$

für alle $x \in D\backslash\{a\}$ mit $\|x - a\| < \delta$.

Beispiel: $f : \mathbb{R}^2 \to \mathbb{R}$ sei definiert durch

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

Diese Funktion ist für $(x,y) \neq 0$ stetig, im Punkt $(x,y) = 0$ aber nicht stetig. Denn es gilt

$$f(x,0) = f(0,y) = 0\,,$$

also ist f auf den Geraden $x = 0$ und $y = 0$ identisch Null. Auf der Diagonalen $x = y$ gilt aber

$$f(x,y) = \frac{2x^2}{2x^2} = 1\,.$$

Für die beiden Folgen $\{z_k\}_{k=1}^\infty$ mit $z_k = (\frac{1}{k}, 0)$ und $\{\tilde{z}_k\}_{k=1}^\infty$ mit $\tilde{z}_k = (\frac{1}{k}, \frac{1}{k})$ gilt also $\lim_{k\to\infty} z_k = \lim_{k\to\infty} \tilde{z}_k = 0$, aber

$$\lim_{k\to\infty} f(z_k) = 0 = f(0)\,, \quad \lim_{k\to\infty} f(\tilde{z}_k) = 1\,.$$

Weil diese Grenzwerte der beiden Bildfolgen nicht übereinstimmen, besitzt f keinen Grenzwert in 0, und ist somit nicht stetig, und kann auch nicht durch eine andere Festlegung des Wertes $f(0)$ zu einer stetigen Funktion gemacht werden. Jedoch sind die Funktionen

$$x \mapsto f(x,y)\,, \quad y \mapsto f(x,y)$$

stetig, also ist f stetig in jeder Variablen.

Satz: Sei $f : D \to \mathbb{R}^m$, $D \subseteq \mathbb{R}^n$, $a \in D$. f ist stetig in a, genau dann wenn alle Komponentenfunktionen $f_1, \ldots, f_m : D \to \mathbb{R}^1$ in a stetig sind.

Beweis: Sei $\{x_k\}_{k=1}^\infty$ eine Folge mit $x_k \in D$ und $\lim_{k\to\infty} x_k = a$. Es gilt $\lim_{k\to\infty} f(x_k) = f(a)$ genau dann wenn $\lim_{k\to\infty} f_i(x_k) = f_i(a)$ gilt für $i = 1, \ldots, m$. Hieraus folgt die

Behauptung. ■

Definition: $f : D \to \mathbb{R}^m$ heißt stetig, wenn f in jedem Punkt $x \in D$ stetig ist.

Sei $D \subseteq \mathbb{R}^n$. Eine Teilmenge D' von D heißt relativ offen bezüglich D wenn eine offene Menge $O \subseteq \mathbb{R}^n$ existiert mit $D' = O \cap D$.

Lemma: $O \subseteq D$ ist genau dann relativ offen bezüglich D, wenn zu jedem $x \in O$ eine Umgebung U von x existiert mit $U \cap D \subseteq O$.

Beweis: Sei O relativ offen mit $x \in O$. Dann existiert eine offene Menge O' mit $O = O' \cap D$. Dann ist O' die gesuchte Umgebung. Umgekehrt existiere zu jedem $x \in O$ eine Umgebung $U(x)$ mit $U(x) \cap D \subseteq O$. Weil jede Umgebung eine offene Umgebung enthält, kann angenommen werden, daß $U(x)$ offen ist. Dann ist

$$O \subseteq D \cap \bigcup_{x \in O} U(x)$$

und

$$D \cap \bigcup_{x \in O} U(x) = \bigcup_{x \in O} \Big(D \cap U(x)\Big) \subseteq O\,,$$

also $O = D \cap \bigcup_{x \in O} U(x)$. Da $\bigcup_{x \in O} U(x)$ offen ist, folgt die Behauptung. ■

Satz: Sei $D \subseteq \mathbb{R}^n$. $f : D \to \mathbb{R}^m$ ist stetig, genau dann wenn das Urbild jeder offenen Teilmenge von $\mathbb{R}^m$ unter f relativ offen ist bezüglich D.

Beweis: Sei f stetig, $O \subseteq \mathbb{R}^m$ offen und $x \in f^{-1}(O)$. Dann ist $f(x) \in O$, also ist O Umgebung von $f(x)$, also existiert eine Umgebung V von x mit $f(V \cap D) \subseteq O$, also gilt $V \cap D \subseteq f^{-1}(O)$, also ist $f^{-1}(O)$ relativ offen bezüglich D.

Sei umgekehrt das Urbild jeder offenen Menge relativ offen. Sei $x \in D$. Sei U eine offene Umgebung von $f(x)$. Dann ist $f^{-1}(U)$ relativ offen, also existiert eine offene Menge $O \subseteq \mathbb{R}^n$ mit $f^{-1}(U) = O \cap D$. Wegen $x \in f^{-1}(U) \subseteq O$ ist O Umgebung von x, also ist f stetig wegen

$$f(O \cap D) = f\Big(f^{-1}(U)\Big) \subseteq U\,.$$

■

Satz: (i) Sei $D \subseteq \mathbb{R}^n$, $f : D \to \mathbb{R}^m$, $g : D \to \mathbb{R}^m$ seien stetig. Dann sind auch die Abbildungen

$$f + g\,,\ \ cf\ \ (c \in \mathbb{R})$$

stetig.

(ii) Seien $f : D \to \mathbb{R}$, $g : D \to \mathbb{R}$ stetig. Dann ist auch $f \cdot g$ stetig. $\frac{f}{g}$ ist stetig in allen Punkten, in denen g nicht verschwindet.

(iii) Seien $f : D \to \mathbb{R}^m$, $\varphi : D \to \mathbb{R}$ stetig. Dann ist auch φf stetig.

Beweis: klar!

Satz: Seien $D_1 \subseteq \mathbb{R}^n$, $f : D_1 \to \mathbb{R}^p$, $D_2 \subseteq \mathbb{R}^p$, $g : D_2 \to \mathbb{R}^m$ stetige Abbildungen, so daß $g \circ f$ existiert. Dann ist auch $g \circ f$ stetig.

Beweis: klar!

Definition: Sei $D \subseteq \mathbb{R}^n$. $f : D \to \mathbb{R}^m$ heißt gleichmäßig stetig, wenn zu jedem $\varepsilon > 0$ ein $\delta > 0$ existiert mit

$$\|f(x) - f(y)\| < \varepsilon$$

für alle $x, y \in D$ mit $\|x - y\| < \delta$.

Satz: Sei $f : D \to \mathbb{R}^m$ stetig und D kompakt. Dann ist f gleichmäßig stetig.

Satz: Sei $D \subseteq \mathbb{R}^n$ kompakt und f stetig. Dann ist $f(D)$ kompakt.

Folgerung: Sei $D \subseteq \mathbb{R}^n$ kompakt und $f : D \to \mathbb{R}$ stetig. Dann nimmt f das Maximum und Minimum an.

Definition: Sei $M \subseteq \mathbb{R}^n$. M heißt zusammenhängend, wenn gilt:
Sei $U_1, U_2 \subseteq M$ relativ offen mit $U_1 \cap U_2 = \emptyset$ und $M = U_1 \cup U_2$. Dann muß entweder $M = U_1$ und $U_2 = \emptyset$ oder $M = U_2$ und $U_1 = \emptyset$ gelten.

Beispiel: Jedes Intervall in $\mathbb{R}$ ist zusammenhängend.

Satz: Sei $D \subseteq \mathbb{R}^n$ zusammenhängend und $f : D \to \mathbb{R}^m$ stetig. Dann ist $f(D)$ zusammenhängend.

Beweis: Seien $U_1, U_2 \subseteq f(D)$ relativ offen mit $U_1 \cap U_2 = \emptyset$ und $f(D) = U_1 \cup U_2$. Da f stetig ist, sind dann $f^{-1}(U_1), f^{-1}(U_2)$ relativ offen mit $f^{-1}(U_1) \cap f^{-1}(U_2) = \emptyset$ und $f^{-1}(U_1) \cup f^{-1}(U_2) = D$. Da D zusammenhängend ist, ist $f^{-1}(U_1)$ oder $f^{-1}(U_2) = \emptyset$, also ist U_1 oder $U_2 = \emptyset$, und hieraus folgt die Behauptung. ∎

Definition: Sei $[a, b] \subseteq \mathbb{R}$ ein Intervall, $\gamma : [a, b] \to \mathbb{R}^m$ stetig. Dann heißt γ ein Weg.

Definition: Eine Menge $M \subseteq \mathbb{R}^n$ heißt wegzusammenhängend, wenn zwei beliebige Punkte von M durch einen in M verlaufenden Weg stetig miteinander verbunden werden können, d.h. wenn es zu $x, y \in M$ ein Intervall $[a,b]$ und eine stetige Abbildung $\gamma : [a,b] \to M$ gibt mit $\gamma(a) = x\,,\;\; \gamma(b) = y\,.$

$\gamma(a)$ heißt Anfangspunkt, $\gamma(b)$ Endpunkt des Weges.

Satz: Sei $D \subseteq \mathbb{R}^n$ wegzusammenhängend, $f : D \to \mathbb{R}^m$ sei stetig. Dann ist $f(D)$ wegzusammmenhängend.

Beweis: Sei $a, b \in f(D)$, und seien $x \in f^{-1}(a)\,,\;\; y \in f^{-1}(b)$. Dann existiert ein Weg γ , der x in D mit y verbindet. $f \circ \gamma$ ist dann ein Weg, der a mit b in $f(D)$ verbindet. ∎

Satz: Sei $M \subseteq \mathbb{R}^n$ wegzusammenhängend. Dann ist M zusammenhängend.

Beweis: Angemommen, M sei nicht zusammenhängend. Dann existieren bezüglich M relativ offene Mengen $U_1 \neq \emptyset\,,\;\; U_2 \neq \emptyset$ mit $U_1 \cap U_2 = \emptyset$ und $U_1 \cup U_2 = M$. Wähle $x \in U_1\,,\;\; y \in U_2$, und wähle einen Weg $\gamma : [a,b] \to M$, der x mit y innerhalb M verbindet. Seien

$$\begin{aligned} V_1 &= \gamma\Big([a,b]\Big) \cap U_1 \\ V_2 &= \gamma\Big([a,b]\Big) \cap U_2\,. \end{aligned}$$

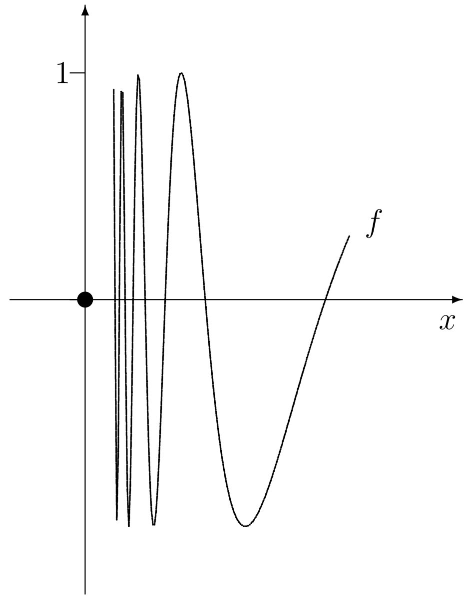
V_1 und V_2 sind relativ offen bezüglich $\gamma([a,b])$. Außerdem gilt $V_1 \neq \emptyset\,,\;\; V_2 \neq \emptyset\,,\;\; V_1 \cap V_2 = \emptyset\,,\;\; V_1 \cup V_2 = \gamma([a,b])$, also ist $\gamma([a,b])$ nicht zusammenhängend.

Andererseits ist $[a,b]$ zusammenhängend und γ stetig, also $\gamma([a,b])$ zusammenhängend. Dies ist ein Widerspruch, also muß M zusammenhängend sein. ∎

Beispiel: Betrachte die Abbildung $f : \mathbb{R}_0^+ \to \mathbb{R}$,

$$f(x) = \begin{cases} \sin\dfrac{1}{x}\,, & x > 0 \\ 0\,, & x = 0\,. \end{cases}$$

Dann ist die Bildmenge $M = \{(x, f(x)) \,\big|\, x \in \mathbb{R}_0^+\} \subseteq \mathbb{R}^2$ zusammenhängend, aber nicht wegzusammenhängend.



Der Beweis, daß M nicht wegezusammenhängend ist, bleibt dem Leser als Übungsaufgabe überlassen.

Ich zeige, daß M zusammenhängend ist. Angenommen, M sei nicht zusammenhängend. Seien $U_1, U_2 \subseteq M$ relativ offene Mengen mit $U_1 \neq \emptyset$, $U_2 \neq \emptyset$, $U_1 \cap U_2 = \emptyset$ und $U_1 \cup U_2 = M$. Die Menge $M' = \{(x, f(x)) \,\big|\, x \in \mathbb{R}^+\} \subseteq M$ ist zusammenhängend als Bild von $\mathbb{R}^+$ unter der stetigen Abbildung

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

Also muß entweder $U_1 \cap M' = \emptyset$ oder $U_2 \cap M' = \emptyset$ sein. O.B.d.A. sei $U_1 \cap M' = \emptyset$. Dann gilt $U_2 = M'$ und $U_1 = \{(0,0)\}$. Diese Menge U_1 ist aber nicht relativ offen bezüglich M, denn sonst würde eine offene Menge $O \subseteq \mathbb{R}^2$ existieren mit $U_1 = M \cap O$. O wäre Umgebung von $(0,0)$, würde also noch eine ε–Umgebung von $(0,0)$ enthalten, also würde O außer $(0,0)$ auch noch andere Punkte von M enthalten, also $U_1 \neq M \cap O$. Dies ist ein Widerspruch, also ist M zusammenhängend. ∎

Dieses Beispiel zeigt, daß die Umkehrung des letzten Satzes nicht gilt.

Sei $D \subseteq \mathbb{R}^n$, $f : D \to \mathbb{R}^m$ sei stetig und injektiv. Die Umkehrabbildung $f^{-1} : f(D) \to \mathbb{R}^n$ braucht nicht stetig zu sein. Es gilt aber:

Satz: Ist f eine injektive, stetige Abbildung mit kompaktem Definitionsbereich, dann ist auch f^{-1} stetig.

Beweis: wie im $\mathbb{R}^1$!

Definition: Sei $D \subseteq \mathbb{R}^n$, sei $f : D \to W \subseteq \mathbb{R}^m$ bijektiv, und seien f und $f^{-1} : W \to D$ stetig. Dann heißt f Homöomorphismus von D auf W.

1 d.) Gleichmäßige Konvergenz

Sei $\|\cdot\|$ eine Norm auf $\mathbb{R}^m$.

Definition: Sei D eine Menge, $f : D \to \mathbb{R}^m$ sei beschränkt. Dann heißt

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

die Supremumsnorm von f.

Daß $\|\cdot\|_\infty$ eine Norm ist, beweist man wie für reellwertige Abbildungen. (Siehe das Skriptum Analysis I, Abschnitt 9 c.)

Die Menge $B(D, \mathbb{R}^m)$ aller beschränkten Abbildungen von D nach $\mathbb{R}^m$ ist ein reeller Vektorraum. Mit $\|\cdot\|_\infty$ wird $B(D, \mathbb{R}^m)$ zu einem normierten Raum. Die Norm in $B(D, \mathbb{R}^m)$ hängt natürlich davon ab, welche Norm in $\mathbb{R}^m$ gewählt wird. Da aber in $\mathbb{R}^m$ alle Normen äquivalent sind, gilt dies auch für die hiermit definierten Supremumsnormen auf $B(D, \mathbb{R}^m)$. Seien $\|\cdot\|^{(1)}$, $\|\cdot\|^{(2)}$ Normen auf $\mathbb{R}^m$, und $\|\cdot\|_\infty^{(1)}$, $\|\cdot\|_\infty^{(2)}$ die zugehörigen Normen auf $B(D, \mathbb{R}^m)$. Dann existieren Konstanten $a, b > 0$ mit

$$\begin{aligned} a\|x\|^{(2)} &\leq \|x\|^{(1)} \leq b\|x\|^{(2)}, \quad \text{für alle} \quad x \in \mathbb{R}^m \\ a\|f\|_\infty^{(2)} &\leq \|f\|_\infty^{(1)} \leq b\|f\|_\infty^{(2)}, \quad \text{für alle} \quad f \in B(D, \mathbb{R}^m). \end{aligned}$$

Somit hängt die folgende Definition nicht von der auf $\mathbb{R}^m$ gewählten Norm ab:

Definition: Sei D eine Menge, $\{f_k\}_{k=1}^\infty$ sei eine Folge von Abbildungen $f_k \in B(D, \mathbb{R}^m)$. $\{f_k\}_{k=1}^\infty$ heißt gleichmäßig konvergent, wenn $f \in B(D, \mathbb{R}^m)$ existiert mit

$$\lim_{k \to \infty} \|f_k - f\|_\infty = 0.$$

Satz: (Cauchysches Konvergenzkriterium) $\{f_k\}_{k=1}^\infty$ konvergiert genau dann gleichmäßig, wenn zu jedem $\varepsilon > 0$ ein $k_0 \in \mathbb{N}$ existiert mit

$$\|f_k - f_\ell\|_\infty < \varepsilon$$

für alle $k, \ell \geq k_0$.

Also ist $B(D, \mathbb{R}^m)$ vollständig, d.h. $B(D, \mathbb{R}^m)$ ist ein Banachraum. Man beweist diesen

Satz wie für reellwertige Abbildungen. (Siehe Skriptum Analysis I, Abschnitt 9 c).

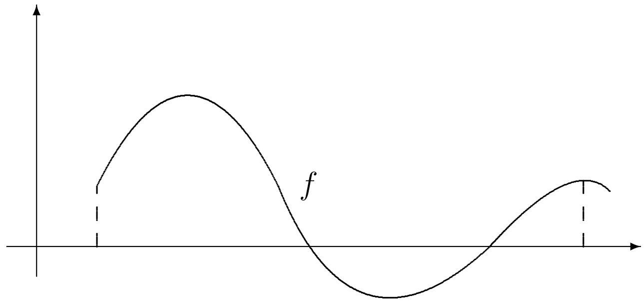
Satz: Sei $D \subseteq \mathbb{R}^n$, sei $\{f_k\}_{k=1}^\infty$ eine Folge stetiger Funktionen $f_k \in B(D, \mathbb{R}^m)$. $\{f_k\}_{k=1}^\infty$ konveriere gleichmäßig gegen f. Dann ist f stetig.

Beweis: wie für reellwertige Abbildungen.

Sei $D \subseteq \mathbb{R}^n$, $C(D, \mathbb{R}^m)$ sei der Raum der beschränkten stetigen Abbildungen. Aus diesem Satz folgt, daß auch $C(D, \mathbb{R}^m)$ ein Banachraum ist.

2 Das Riemannsche Integral

Für eine möglichst große Klasse von reellen Funktionen möchte man den Inhalt der Fläche bestimmen, die begrenzt ist durch den Graphen der gegebenen Funktion und der Abszissenachse.



Für „komplizierte“ Funktionen ist es allerdings schwer zu sagen, was dieser Flächeninhalt sein soll. Als Beispiel betrachte man die Dirichlet–Funktion

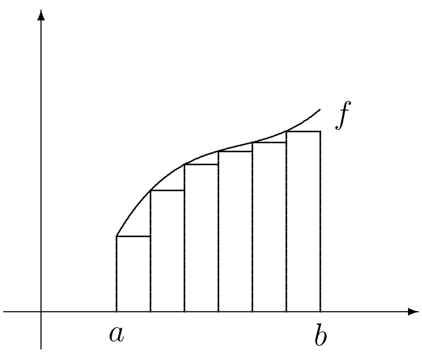
$$f(x) = \begin{cases} 1\,, & x \in \mathbb{Q} \\ 0\,, & x \in \mathbb{R}\backslash\mathbb{Q}\,. \end{cases}$$

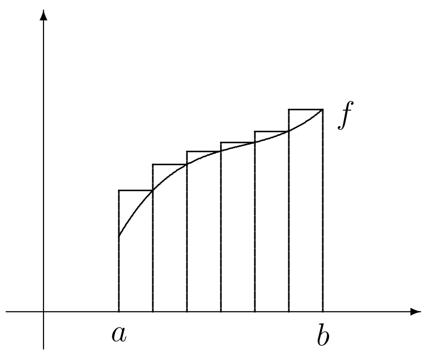
Die Aufgabe besteht also darin, Mengen, die „begrenzt“ werden durch Funktionsgraphen, eine Zahl zuzuordnen, die man das Integral der entsprechenden Funktion nennt, und die Eigenschaft hat, wie man sie intuitiv vom Flächeninhalt erwartet. Es wird sich zeigen, daß dies nicht für alle Funktionen möglich ist.

In diesem Kapitel werden wir das Riemannsche Integral für reelle Funktionen besprechen. Man kann auch Riemannsche Integrale für Funktionen von n Veränderlichen definieren.

2 a.) Definition des Riemannschen Integrals für Funktionen einer Variablen

Sei $-\infty < a < b < \infty$, und sei $f : [a, b] \to \mathbb{R}$. Zur Berechnung des Inhalts der Fläche unter dem Graphen von f ist es naheliegend, diese Fläche durch Rechtecke auszuschöpfen:





Bei Verfeinerung der Unterteilung wird der Flächeninhalt der Rechtecke in anschaulichem Sinn gegen den Flächeninhalt der Fläche unter dem Graphen von f konvergieren. Man kann die Fläche unter dem Graphen von f auch durch Rechtecke überdecken. Auch in diesem Fall konvergiert der Flächeninhalt der Rechtecke in anschaulichem Sinn gegen den Flächeninhalt der Fläche unter dem Graphen von f, wenn man die Unterteilung in Rechtecke verfeinert.

Man erwartet also, daß der Flächeninhalt der „ausschöpfenden Rechtecksfläche“ und der Flächeninhalt der „überdeckenden Rechtecksfläche“ bei Verfeinerung der Unterteilung gegen dieselbe Zahl konvergieren. Diese Zahl wird man als Flächeninhalt der Fläche unter dem Graphen von f bezeichnen.

Klar ist aber, daß diese Flächeninhalte der überdeckenden und der ausschöpfenden Rechtecksflächen nicht für alle f bei Verfeinerung der Unterteilung gegen dieselbe Zahl konvergieren werden. Ein Beispiel dafür ist wieder die Dirichletfunktion.

Diejenigen Funktionen f, für die die Flächeninhalte der überdeckenden Rechtecksfläche und der ausschöpfenden Rechtecksfläche gegen dieselbe Zahl konvergieren, heißen „Riemann–integrierbar“, und diese Zahl heißt „Riemann–Integral“ von f über dem Intervall $[a, b]$. Die anderen Funktionen heißen „nicht Riemann–integrierbar“. Dieses „Programm“ wird nun durchgeführt.

Definition: Sei $-\infty < a < b < \infty$. Unter einer Partition P des Intervalls $[a, b]$ versteht man eine endliche Menge $\{x_0, \ldots, x_n\} \subseteq \mathbb{R}$ mit

$$a = x_0 < x_1 < \ldots \quad < x_{n-1} < x_n = b\,.$$

Zur Abkürzung sei $\Delta x_i = x_i - x_{i-1} \quad (i = 1, \ldots, n)$.

Sei $f : [a, b] \to \mathbb{R}$ eine beschränkte reelle Funktion und $P = \{x_0, \ldots, x_n\}$ eine Partition von $[a, b]$. Sei

$$\begin{aligned} M_i &= \sup \{ f(x) \,\big|\, x_{i-1} \le x \le x_i \} \\ m_i &= \inf \{ f(x) \,\big|\, x_{i-1} \le x \le x_i \} \end{aligned}, \qquad i = 1, \ldots, n,$$

und sei

$$\begin{aligned} U(P, f) &= \sum_{i=1}^{n} M_i \Delta x_i \\ L(P, f) &= \sum_{i=1}^{n} m_i \Delta x_i\,. \end{aligned}$$

Definition: Sei

$$\begin{aligned}\overline{\int_a^b} f dx &= \inf \left\{U(P,f) \,\middle|\, P \quad \text{Partition von} \quad [a,b]\right\}\\ \underline{\int_a^b} f dx &= \sup \left\{L(P,f) \,\middle|\, P \quad \text{Partition von} \quad [a,b]\right\}.\end{aligned}$$

Die Zahlen $\overline{\int_a^b} f dx$ und $\underline{\int_a^b} f dx$ heißen oberes und unteres Riemannintegral von f. Wenn das obere und untere Riemannintegral übereinstimmen, heißt f Riemann–integrierbar, und der gemeinsame Wert des oberen und unteren Riemannintegrals wird mit

$$\int_a^b f dx \quad \text{oder} \quad \int_a^b f(x) dx$$

bezeichnet. Diese Zahl heißt Riemannintegral von f. Die Menge der beschränkten, Riemann–integrierbaren Funktionen auf dem Intervall $[a,b]$ wird mit $\mathcal{R}([a,b])$ bezeichnet.

Nach Voraussetzung ist f beschränkt, also existieren Zahlen m, M mit

$$m \leq f(x) \leq M$$

für alle $x \in [a,b]$. Hieraus folgt $m \leq m_i \leq M_i \leq M$, also

$$\begin{aligned} m(b-a) &= \sum_{i=1}^n m \Delta x_i \leq \sum_{i=1}^n m_i \Delta x_i = L(P,f)\\ &\leq \sum_{i=1}^n M_i \Delta x_i = U(P,f) \leq \sum_{i=1}^n M \Delta x_i = M(b-a)\,,\end{aligned}$$

also existieren das Infimum der Menge

$$\left\{U(P,f) \,\middle|\, P \quad \text{Partition von} \quad [a,b]\right\}$$

und das Supremum der Menge

$$\left\{L(P,f) \,\middle|\, P \quad \text{Partition von} \quad [a,b]\right\},$$

also ist die voranstehende Definition sinnvoll.

2 b.) Kriterien für Riemann–integrierbare Funktionen

Um mit dem Begriff des Riemannintegrals sinnvoll arbeiten zu können, müssen einfache Kriterien dafür gefunden werden, daß eine gegebene Funktion Riemann–integrierbar ist.

Im Folgenden werden solche Kriterien hergeleitet.

Definition: Seien P, P^* Partitionen von $[a,b]$. P^* heißt Verfeinerung von P, wenn $P \subseteq P^*$ gilt. P^* heißt gemeinsame Verfeinerung der Partitionen P_1 und P_2, wenn $P^* = P_1 \cup P_2$.

Satz: Sei $f : [a,b] \to \mathbb{R}$ eine beschränkte Funktion, und sei P^* eine Verfeinerung der Partition P von $[a,b]$. Dann gilt

$$\begin{aligned} L(P,f) &\leq L(P^*,f) \\ U(P^*,f) &\leq U(P,f)\,. \end{aligned}$$

Beweis: Sei $P = \{x_0, \ldots, x_n\}$, und es werde zunächst angenommen, daß P^* genau einen Punkt x^* mehr enthält als P. Dann gibt es $x_{j-1}, x_j \in P$ mit $x_{j-1} < x^* < x_j$. Seien

$$\begin{aligned} w_1 &= \inf\left\{ f(x) \,\middle|\, x_{j-1} \leq x \leq x^* \right\} \\ w_2 &= \inf\left\{ f(x) \,\middle|\, x^* \leq x \leq x_j \right\}. \\ m_i &= \inf\left\{ f(x) \,\middle|\, x_{i-1} \leq x \leq x_i \right\}, \quad i = 1, \ldots, n\,. \end{aligned}$$

Dann gilt

$$\begin{aligned} L(P,f) &= \sum_{i=1}^{n} m_i \Delta x_i = \sum_{i=1}^{j-1} m_i \Delta x_i \\ &\quad + m_j(x^* - x_{j-1} + x_j - x^*) + \sum_{i=j+1}^{n} m_i \Delta x_i \\ &\leq \sum_{i=1}^{j-1} m_i \Delta x_i + w_1(x^* - x_{j-1}) + w_2(x_j - x^*) + \sum_{i=j+1}^{n} m_i \Delta x_i \\ &= L(P^*, f)\,. \end{aligned}$$

Wenn P^* eine Verfeinerung von P ist, die k Punkte mehr enthält als P, genügt es diese Überlegungen k mal zu wiederholen. (Vollständige Induktion!)
Die zweite Ungleichung des Satzes beweist man ebenso. ■

Satz: Sei $f : [a,b] \to \mathbb{R}$ beschränkt. Dann gilt

$$\underline{\int_a^b} f dx \leq \overline{\int_a^b} f dx$$

Beweis: Seien P_1 und P_2 Partitionen, und sei P^* die gemeinsame Verfeinerung. Nach Definition gilt

$$L(P^*, f) \leq U(P^*, f)\,,$$

und aus dem vorangehenden Satz folgt

$$L(P_1, f) \leq L(P^*, f) \leq U(P^*, f) \leq U(P_2, f)\,,$$

also

$$L(P_1, f) \leq U(P_2, f)$$

für alle Partitionen P_1, P_2 von $[a, b]$. Somit ist $U(P_2, f)$ eine obere Schranke der Menge

$$\Big\{L(P, f) \Big| P \quad \text{Partition von} \quad [a, b]\Big\}\,.$$

Somit ist $U(P_2, f)$ nicht kleiner als das Supremum dieser Menge, also

$$\underline{\int_a^b} f dx \leq U(P_2, f)\,.$$

Aus dieser Ungleichung folgt nun, daß $\underline{\int_a^b} f dx$ eine untere Schranke der Menge

$$\Big\{U(P, f) \Big| P \quad \text{Partition von} \quad [a, b]\Big\}$$

ist, also ist $\underline{\int_a^b} f dx$ nicht größer als das Infimum dieser Menge, somit

$$\underline{\int_a^b} f dx \leq \overline{\int_a^b} f dx\,.$$

■

Satz: Sei $f : [a, b] \to \mathbb{R}$ beschränkt. Es gilt $f \in \mathcal{R}([a, b])$ dann und nur dann wenn zu jedem $\varepsilon > 0$ eine Partition P existiert mit

$$U(P, f) - L(P, f) < \varepsilon\,.$$

Beweis: „$\Longleftarrow$" Zu jedem $\varepsilon > 0$ existiere eine Partition P mit $U(P, f) - L(P, f) < \varepsilon$. Da für jede Partition P^* gilt

$$L(P^*, f) \leq \underline{\int_a^b} f dx \leq \overline{\int_a^b} f dx \leq U(P^*, f)\,,$$

folgt

$$0 \leq \overline{\int_a^b} f dx - \underline{\int_a^b} f dx \leq U(P, f) - L(P, f) < \varepsilon\,,$$

also

$$0 \leq \overline{\int_a^b} f dx - \underline{\int_a^b} f dx < \varepsilon$$

für jedes $\varepsilon > 0$, somit

$$\underline{\int_a^b} f dx = \overline{\int_a^b} f dx\,,$$

also $f \in \mathcal{R}([a,b])$.

„$\Longrightarrow$“ Sei $f \in \mathcal{R}([a,b])$. Nach Definition von Infimum und Supremum gibt es zu jedem $\varepsilon > 0$ Partitionen P_1 und P_2 mit

$$\begin{aligned}\int_a^b f dx &= \overline{\int_a^b} f dx \le U(P_1, f) \le \overline{\int_a^b} f dx + \frac{\varepsilon}{2}\\ \int_a^b f dx &= \underline{\int_a^b} f dx \ge L(P_2, f) \ge \underline{\int_a^b} f dx - \frac{\varepsilon}{2}\,.\end{aligned}$$

Sei P die gemeinsame Verfeinerung von P_1 und P_2. Dann folgt

$$\int_a^b f dx - \frac{\varepsilon}{2} \le L(P,f) \le \int_a^b f dx \le U(P,f) \le \int_a^b f dx + \frac{\varepsilon}{2}$$

also

$$U(P,f) - L(P,f) < \varepsilon\,.$$

■

Aus diesem Satz folgt, daß $\mathcal{R}([a,b])$ die Klasse $C([a,b])$ enthält. Denn es gilt:

Satz: Sei $f : [a,b] \to \mathbb{R}$ stetig. Dann ist $f \in \mathcal{R}([a,b])$. Außerdem gilt: Zu jedem $\varepsilon > 0$ gibt es $\delta > 0$ mit

$$\Big| \sum_{i=1}^n f(t_i)\Delta x_i - \int_a^b f dx \Big| < \varepsilon$$

für jede Partition $P = \{x_0, \ldots, x_n\}$ von $[a,b]$ mit $\max_{i=1,\ldots,n} \Delta x_i < \delta$, und für jede Wahl von Punkten $t_1, \ldots, t_n$ mit $t_i \in [x_{i-1}, x_i]$.

$$\sum_{i=1}^n f(t_i)\Delta x_i$$

heißt Riemannsche Summe.

Beweis: Sei $\varepsilon > 0$. Setze

$$\eta = \frac{\varepsilon}{(b-a)}\,.$$

Da f stetig ist auf dem kompakten Intervall $[a,b]$, ist f beschränkt und auch gleichmäßig stetig. Also existiert $\delta > 0$, so daß

$$|f(x) - f(t)| < \eta \qquad (*)$$

für alle $x, t \in [a,b]$ mit $|x-t| < \delta$. Man wähle nun eine Partition $P = \{x_0, \ldots, x_n\}$ von $[a,b]$ mit $\max_{i=1,\ldots,n} \Delta x_i < \delta$. Dann folgt aus $(*)$ für alle $x, t \in [x_{i-1}, x_i]$

$$f(x) - f(t) < \eta\,,$$

also

$$\begin{aligned} M_1 - m_1 &= \sup_{x_{i-1} \le x \le x_i} f(x) - \inf_{x_{i-1} \le t \le x_i} f(t) \\ &= \max_{x_{i-1} \le x \le x_i} f(x) - \min_{x_{i-1} \le t \le x_i} f(t) \\ &= f(x_0) - f(t_0) < \eta\,, \end{aligned}$$

für geeignete $x_0, t_0 \in [x_{i-1}, x_i]$. Also folgt

$$\begin{aligned} U(P,f) - L(P,f) &= \textstyle\sum_{i=1}^n (M_i - m_i)\Delta x_i < \eta \sum_{i=1}^n \Delta x_i \\ &= \eta(b-a) = \varepsilon\,. \end{aligned} \qquad \binom{*}{*}$$

Nach dem vorangehenden Satz ist also $f \in \mathcal{R}([a,b])$. Wegen

$$\begin{aligned} L(P,f) &\le \sum_{i=1}^n f(t_i)\Delta x_i \le U(P,f) \\ L(P,f) &\le \int_a^b f dx \le U(P,f) \end{aligned}$$

folgt aus $\binom{*}{*}$ auch

$$\Big| \int_a^b f dx - \sum_{i=1}^n f(t_i)\Delta x_i \Big| < \varepsilon\,.$$

■

Auch die Klasse der monotonen Funktionen gehört zu $\mathcal{R}([a,b])$:

Satz: Sei $f : [a,b] \to \mathbb{R}$ monoton. Dann ist $f \in \mathcal{R}([a,b])$.

Beweis: Sei f monoton wachsend. f ist beschränkt wegen $f(a) \le f(x) \le f(b)$ für alle $x \in [a,b]$. Sei $\varepsilon > 0$. Zu beliebigem $n \in \mathbb{N}$ setze

$$x_i = a + \frac{b-a}{n}\, i\,, \quad i = 0, 1, \ldots, n\,.$$

$P = \{x_0, \ldots, x_n\}$ ist eine Partition von $[a,b]$, und wegen der Monotonie von f gilt

$$\begin{aligned} m_i &= \inf\Big\{f(x) \Big| x_{i-1} \le x \le x_i\Big\} = f(x_{i-1}) \\ M_1 &= \sup\Big\{f(x) \Big| x_{i-1} \le x \le x_i\Big\} = f(x_i)\,, \end{aligned}$$

also

$$
\begin{aligned}
U(P,f) - L(P,f) &= \sum_{i=1}^{n} (M_i - m_i)\Delta x_i \\
= \quad \sum_{i=1}^{n} \Big(f(x_i) - f(x_{i-1}\Big) \frac{b-a}{n} &= \Big(f(b) - f(a) \Big) \frac{b-a}{n} < \varepsilon\,,
\end{aligned}
$$

wenn $n \in \mathbb{N}$ hinreichend groß gewählt ist. Also ist $f \in \mathcal{R}([a,b])$. Für monoton fallendes f verläuft der Beweis genauso. ■

2 c.) Einfache Eigenschaften des Riemann-Integrals

Satz: (i) Für $f_1, f_2 \in \mathcal{R}([a,b])$ gilt $f_1 + f_2 \in \mathcal{R}([a,b])$ und

$$
\int_a^b (f_1 + f_2)dx = \int_a^b f_1 dx + \int_a^b f_2 dx\,.
$$

Für $f \in \mathcal{R}([a,b])$ und $c \in \mathbb{R}$ gilt $cf \in \mathcal{R}[(a,b)]$ und

$$
\int_a^b cf dx = c \int_a^b f dx\,.
$$

Also ist $\mathcal{R}([a,b])$ ein Vektorraum.

(ii) Wenn $f_1, f_2 \in \mathcal{R}([a,b])$ und $f_1(x) \leq f_2(x)$ ist für alle $x \in [a,b]$, dann folgt

$$
\int_a^b f_1 dx \leq \int_a^b f_2 dx
$$

(iii) Wenn $f \in \mathcal{R}([a,b])$ und wenn $a < c < b$, dann

$$
f\big|_{[a,c]} \in \mathcal{R}\Big([a,c]\Big)\,, \quad f\big|_{[c,b]} \in \mathcal{R}\Big([c,b]\Big)\,,
$$

und

$$
\int_a^c f dx + \int_c^b f dx = \int_a^b f dx\,.
$$

(iv) Wenn $f \in \mathcal{R}([a,b)]$ und $|f(x)| \leq M$, dann ist

$$
\Big| \int_a^b f dx \Big| \leq M(b-a)\,.
$$

Beweis: (i) Sei $f = f_1 + f_2$, und sei P eine Partition von $[a,b]$. Dann gilt

$$
\begin{aligned}
\inf_{x_{i-1} \leq x \leq x_i} f(x) &= \inf_{x_{i-1} \leq x \leq x_i} \Big(f_1(x) + f_2(x) \Big) \\
&\geq \inf_{x_{i-1} \leq x \leq x_i} f_1(x) + \inf_{x_{i-1} \leq x \leq x_i} f_2(x)\,,
\end{aligned}
$$

und

$$\begin{aligned}\sup_{x_{i-1}\le x\le x_i} f(x) &= \sup_{x_{i-1}\le x\le x_i} \Big(f_1(x)+f_2(x)\Big)\\ &\le \sup_{x_{i-1}\le x\le x_i} f_1(x) + \sup_{x_{i-1}\le x\le x_i} f_2(x)\,,\end{aligned}$$

also

$$\begin{aligned}L(P,f_1)+L(P,f_2) &\le L(P,f)\le U(P,f)\\ &\le U(P,f_1)+U(P,f_2)\,.\end{aligned} \tag{+}$$

Zu $\varepsilon > 0$ wähle Partitionen P_1 und P_2 mit

$$U(P_j,f_j)-L(P_j,f_j)<\varepsilon\,,\quad j=1,2\,.$$

Sei P die gemeinsame Verfeinerung von P_1 und P_2 . Dann folgt

$$U(P,f_j)-L(P,f_j)<\varepsilon\,,\quad j=1,2\,, \tag{++}$$

und somit, wegen (+) ,

$$\begin{aligned}U(P,f)-L(P,f) \le U(P,f_1)+U(P,f_2)\\ -L(P,f_1)-L(P,f_2)<2\varepsilon\,.\end{aligned}$$

Da $\varepsilon > 0$ beliebig war, folgt hieraus $f = f_1 + f_2 \in \mathcal{R}([a,b])$.
Aus (++) folgt auch

$$U(P,f_j) < \int_a^b f_j dx + \varepsilon$$

und somit, wegen (+) ,

$$\int_a^b f dx \le U(P,f) \le \int_a^b f_1 dx + \int_a^b f_2 dx + 2\varepsilon\,.$$

Da ε beliebig war, ergibt dies

$$\int_a^b f dx \le \int_a^b f_1 dx + \int_a^b f_2 dx\,.$$

Ebenso folgt

$$\int_a^b f dx \ge \int_a^b f_1 dx + \int_a^b f_2 dx\,,$$

also

$$\int_a^b f dx = \int_a^b f_1 dx + \int_a^b f_2 dx\,.$$

Die Behauptungen (ii) – (iv) werden ganz ähnlich bewiesen. Die Beweise bleiben dem Leser überlassen. ■

Satz: Seien $-\infty < m < M < \infty$ und $f \in \mathcal{R}([a,b])$ mit $f : [a,b] \to [m,M]$. Sei $\Phi : [m,M] \to \mathbb{R}$ stetig, und sei $h = \Phi \circ f$. Dann ist $h \in \mathcal{R}([a,b])$.

Beweis: Sei $\varepsilon > 0$. Da Φ gleichmäßig stetig ist auf $[m,M]$, gibt es ein $\delta > 0$ mit $\delta < \varepsilon$ und mit $|\Phi(s) - \Phi(t)| < \varepsilon$ für alle $s,t \in [m,M]$ mit $|s-t| \le \delta$. Weil $f \in \mathcal{R}[(a,b)]$ ist gibt es außerdem eine Partition $P = \{x_0, \ldots, x_n\}$ von $[a,b]$, so daß

$$U(P,f) - L(P,f) < \delta^2 . \qquad (*)$$

Seien

$$\begin{aligned} M_i &= \sup_{x_{i-1} \le x \le x_i} f(x), & m_i &= \inf_{x_{i-1} \le x \le x_i} f(x) \\ M_i^* &= \sup_{x_{i-1} \le x \le x_i} h(x), & m_i^* &= \inf_{x_{i-1} \le x \le x_i} h(x) . \end{aligned}$$

Setze

$$\begin{aligned} A &= \left\{ i \,\middle|\, i \in \mathbb{N},\ 1 \le i \le n,\ M_i - m_i < \delta \right\} \\ B &= \{1, \ldots, n\} \setminus A . \end{aligned}$$

Wenn $i \in A$ ist, folgt für alle x, y mit $x_{i-1} \le x, y \le x_i$

$$|h(x) - h(y)| = \Big| \Phi\Big(f(x)\Big) - \Phi\Big(f(y)\Big) \Big| < \varepsilon$$

wegen $|f(x) - f(y)| \le M_i - m_i < \delta$. Dies liefert

$$M_i^* - m_i^* \le \varepsilon .$$

Wenn $i \in B$ ist, gilt

$$M_i^* - m_i^* \le 2K ,$$

mit $K = \sup_{m \le t \le M} |\Phi(t)|$. Aus $(*)$ folgt außerdem

$$\delta \sum_{i \in B} \Delta x_i \le \sum_{i \in B} (M_i - m_i) \Delta x_i \le \sum_{i=1}^{n} (M_i - m_i) \Delta x_i = U(P,f) - L(P,f) < \delta^2 ,$$

also

$$\sum_{i \in B} \Delta x_i < \delta .$$

Somit folgt

$$\begin{aligned} U(P,h) - L(P,h) &= \sum_{i\in A}(M_i^* - m_i^*)\Delta x_i + \sum_{i\in B}(M_i^* - m_i^*)\Delta x_i \\ &\le \delta \sum_{i\in A} \Delta x_i + 2K \sum_{i\in B} \Delta x_i \\ &\le \delta(b-a) + 2K\delta \le \varepsilon(b-a+2K)\,, \end{aligned}$$

wegen $\delta < \varepsilon$. Da ε beliebig gewählt war, folgt hieraus $h \in \mathcal{R}([a,b])$. ■

Folgerung: Seien $f, g \in \mathcal{R}([a,b])$. Dann gilt

a.) $fg \in \mathcal{R}([a,b])$

b.) $|f| \in \mathcal{R}([a,b])$ und $\left| \int_a^b f dx \right| \le \int_a^b |f|\, dx$.

Beweis: Mit $\Phi(t) = t^2$ zeigt der vorangehende Satz, daß $f^2 = \Phi \circ f \in \mathcal{R}([a,b])$ ist. Wegen

$$fg = \frac{1}{4}\left[(f+g)^2 - (f-g)^2\right]$$

folgt also auch $fg \in \mathcal{R}([a,b])$.
Mit $\Phi(t) = |t|$ folgt aus dem vorangehenden Satz, daß $|f| = \Phi \circ f \in \mathcal{R}([a,b])$. Wähle $c = \pm 1$, so daß

$$c \int_a^b f dx \ge 0\,.$$

Dann ergibt sich

$$\left| \int_a^b f dx \right| = c \int_a^b f dx = \int_a^b c f dx \le \int_a^b |f| dx\,,$$

wegen $cf(x) \le |f(x)|$ für alle $x \in [a,b]$. ■

2 d.) Hauptsatz der Differential– und Integralrechnung

Sei $-\infty < a < b < \infty$ und $f \in \mathcal{R}([a,b])$. Man setzt

$$\int_b^a f dx = -\int_a^b f dx\,.$$

Es gilt dann

$$\int_u^v f dx + \int_v^w f dx = \int_u^w f dx$$

wenn u, v, w beliebige Punkte von $[a,b]$ sind.

Satz: (Mittelwertsatz der Integralrechnung). Sei $f : [a, b] \to \mathbb{R}$ stetig. Dann gibt es eine Stelle c mit $a \leq c \leq b$, so daß gilt

$$\int_a^b f dx = f(c)(b-a) .$$

Beweis: f ist auf einem kompakten Intervall definiert und stetig, also integrierbar. Da das Integral monoton ist, gilt

$$\begin{aligned} (b-a) \min_{x\in[a,b]} f(x) &= \int_a^b \min_{y\in[a,b]} f(y)dx \leq \int_a^b f(x)dx \\ &\leq \int_a^b \max_{y\in[a,b]} f(y)dx = \max_{x\in[a,b]} f(x)(b-a) . \end{aligned}$$

Da f das Minimum und das Maximum auf $[a, b]$ annimmt, existiert nach dem Zwischenwertsatz ein $c \in [a, b]$ mit

$$f(c) = \frac{1}{b-a} \int_a^b f(x)dx .$$

■

Satz: Sei $f \in \mathcal{R}([a, b])$. Dann ist

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

eine stetige Funktion $F : [a, b] \to \mathbb{R}$.

Beweis: Es existiert M mit $|f(x)| \leq M$ für alle $x \in [a, b]$. Also gilt für $x, x_0 \in [a, b]$ mit $x_0 < x$

$$\begin{aligned} \Big| F(x) - F(x_0) \Big| &= \Big| \int_a^x f(t)dt - \int_a^{x_0} f(t)dt \Big| \\ &= \Big| \int_{x_0}^x f(t)dt \Big| \leq M(x - x_0) . \end{aligned}$$

Hieraus folgt die Stetigkeit von F auf $[a, b]$. ■

Satz: Sei $f : [a, b] \to \mathbb{R}$ stetig. Dann ist die durch

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

definierte Funktion differenzierbar, und es gilt

$$F' = f .$$

Die Funktion F ist also Stammfunktion von f.

Beweis: Sei $x_0 \in [a,b]$. Dann gilt nach dem Mittelwertsatz der Integralrechnung

$$\begin{aligned}
\lim_{\substack{x\to x_0\\ x\neq x_0}} \frac{F(x)-F(x_0)}{x-x_0} &= \lim_{\substack{x\to x_0\\ x\neq x_0}} \frac{1}{x-x_0}\Big[\int_a^x f(t)dt - \int_a^{x_0} f(t)dt\Big]\\
&= \lim_{\substack{x\to x_0\\ x\neq x_0}} \frac{1}{x-x_0}\int_{x_0}^x f(t)dt = \lim_{\substack{x\to x_0\\ x\neq x_0}} \frac{1}{x-x_0} f(y)(x-x_0)\\
&= \lim_{y\to x_0} f(y) = f(x_0)\,,
\end{aligned}$$

für geeeignetes y zwischen x_0 und x. ∎

Hauptsatz der Differential und Integralrechnung: Ist f Stammfunktion der stetigen Funktion $f : [a,b] \to \mathbb{R}$, dann gilt

$$\int_a^b f(t)dt = F(b) - F(a) = F(x)\Big|_a^b .$$

Beweis: Die Funtkion $x \mapsto \int_a^b f(t)dt$ ist eine Stammfunktion zu f, also existiert eine Konstante c mit

$$F(x) = \int_a^x f(t)dt + c$$

für alle $x \in [a,b]$. Hieraus folgt $c = F(a)$, also $F(b) - F(a) = \int_a^b f(t)dt$. ∎

Dieser Satz ist das wichtigste Hilfsmittel zur Berechnung von Integralen:

Beispiele:

1.) $$\int_a^b x^c dx = \frac{1}{c+1}\, x^{c+1}\Big|_a^b ,$$

falls $0 < a < b$, $c \in \mathbb{R}$, $c \neq -1$. Für $c < -1$ gilt

$$\lim_{m\to\infty} \int_a^m x^c dx = \lim_{m\to\infty} \frac{1}{c+1}\, m^{c+1} - \frac{1}{c+1}\, a^{c+1} = -\frac{1}{c+1}\, a^{c+1} .$$

Daher definiert man

$$\int_a^\infty x^c dx = \lim_{m\to\infty} \int_a^m x^c dx = -\frac{1}{c+1}\, a^{c+1} .$$

Das Integral $\int_a^\infty x^c dx$ heißt uneigentliches Riemannintegral, und man sagt, für $c < -1$ sei x^c im uneigentlichen Sinn Riemann–integrierbar über $[a,\infty)$ mit $a > 0$. Insbesondere ergibt sich

$$\int_1^\infty x^{-2} dx = 1 .$$

2.) $$\int_a^b \frac{1}{x}\, dx = \ln b - \ln a\,,$$

für $0 < a < b < \infty$. $\frac{1}{x}$ ist nicht integrierbar über $[1,\infty)$.

3.)
$$\int_a^b \frac{1}{\sqrt{1-x^2}}\,dx = \arcsin b - \arcsin a$$
für $-1 < a < b < 1$.
Man setzt
$$\begin{aligned}\int_{-1}^1 \frac{1}{\sqrt{1-x^2}}\,dx &= \lim_{\substack{b\to 1\\ b<1}} \lim_{\substack{a\to -1\\ a>-1}} \int_a^b \frac{1}{\sqrt{1-x^2}}\,dx \\ &= \lim_{b\to 1} \arcsin b - \lim_{a\to -1} \arcsin a = \frac{\pi}{2} - \Big(-\frac{\pi}{2}\Big) = \pi\,.\end{aligned}$$
Die Funktion $\frac{1}{\sqrt{1-x^2}}$ ist im uneigentlichen Sinn Riemann–integrierbar über dem Intervall $[-1,1]$.

Satz: (Substitutionsregel) f sei stetig, $g : [a,b] \to \mathbb{R}$ sei stetig differenzierbar, und die Hintereinanderausführung $f \circ g$ existiere. Dann gilt
$$\int_a^b f\Big(g(t)\Big)g'(t)dt = \int_{g(a)}^{g(b)} f(x)dx\,.$$

Beweis: Es existiert eine Stammfunktion F von f, und da F denselben Definitionsbereich wie f hat, existiert die Hintereinanderausführung $F \circ g$. Nach der Kettenregel gilt
$$(F\circ g)' = (F'\circ g)g' = (f\circ g)g'\,,$$
also
$$F\Big(g(b)\Big) - F\Big(g(a)\Big) = \int_a^b f\Big(g(t)\Big)g'(t)dt\,.$$
Wegen
$$F\Big(g(b)\Big) - F\Big(g(a)\Big) = \int_{g(a)}^{g(b)} f(x)dx$$
folgt die Behauptung. ■

Bemerkung: Falls g^{-1} existiert, kann die Substitutionsregel auch in der Form
$$\int_a^b f(x)dx = \int_{g^{-1}(a)}^{g^{-1}(b)} f\Big(g(t)\Big)g'(t)dt$$
geschrieben werden.

Beispiel: $\displaystyle\int_0^1 \sqrt{1-x^2}\,dx$ soll berechnet werden. Mit der Substitution $x = x(t) = \cos t$ folgt wegen der Umkehrbarkeit von cos im Intervall $[0,\frac{\pi}{2}]$:

$$
\begin{aligned}
\int_0^1 \sqrt{1-x^2}\,dx &= \int_{x^{-1}(0)}^{x^{-1}(1)} \sqrt{1-x(t)^2}\,\frac{dx(t)}{dx}\,dt \\
&= \int_{\frac{\pi}{2}}^{0} \sqrt{1-\cos^2 t}\,(-\sin t)dt = \int_0^{\pi/2} (\sin t)^2 dt \\
&= \int_0^{\pi/2} \Big(\frac{1}{2} - \frac{1}{2}\cos 2t\Big)dt = \frac{\pi}{4} - \frac{1}{4}\sin 2t\Big|_0^{\pi/2} = \frac{\pi}{4}\,.
\end{aligned}
$$

Satz: (Produktintegration) $f : [a,b] \to \mathbb{R}$ sei stetig, F sei eine Stammfunktion von f und $g : [a,b] \to \mathbb{R}$ sei stetig differenzierbar. Dann gilt

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

Beweis: Es gilt $(F \cdot g)' = F' \cdot g + F \cdot g' = f \cdot g + F \cdot g'$, also

$$
F(x)g(x)\Big|_a^b = \int_a^b f(x)g(x)dx + \int_a^b F(x)g'(x)dx\,.
$$

■

Beispiel:

$$
\begin{aligned}
\int_0^\pi \sin^2 x dx &= -\cos x \sin x\Big|_0^\pi + \int_0^\pi \cos^2 x dx \\
&= -\cos x \sin x\Big|_0^\pi + \int_0^\pi (1-\sin^2 x)dx\,,
\end{aligned}
$$

also

$$
\int_0^\pi \sin^2 x dx = \frac{1}{2}\int_0^\pi 1 dx = \frac{\pi}{2}\,.
$$

3 Das Lebesguesche Integral

Sei $f : [0,1] \to \mathbb{R}$ definiert durch

$$f(x) = \begin{cases} \dfrac{1}{\sqrt{x}}\,, & 0 < x \le 1 \\ 0\,, & x = 0\,. \end{cases}$$

Diese Funktion ist nicht Riemann–integrierbar über dem Intervall $[0,1]$, weil sie unbeschränkt ist, und somit das obere Riemann–Integral nicht existiert. Eine Folge $\{f_n\}_{n=1}^{\infty}$ Riemann–integrierbarer Funktionen $f_n : [0,1] \to \mathbb{R}$ erhält man aus dieser Funktion durch Abschneiden:

$$f_n(x) = \begin{cases} 0\,, & 0 \le x < \dfrac{1}{n} \\ \dfrac{1}{\sqrt{x}}\,, & \dfrac{1}{n} \le x \le 1\,. \end{cases}$$

Es gilt $f_n \in \mathcal{R}([0,1])$ für alle n und

$$f_n(x) \nearrow f(x)\,, \quad n \to \infty\,,$$

für alle $x \in [0,1]$. Außerdem gilt

$$\int_0^1 f_n dx = \int_{\frac{1}{n}}^1 \frac{1}{\sqrt{x}}\, dx = \Big[2\sqrt{x}\Big]_{\frac{1}{n}}^1 = 2 - \frac{2}{\sqrt{n}} \nearrow 2\,, \quad n \to \infty\,.$$

Obwohl also die Folge $\{f_n\}_{n=1}^{\infty}$ von Riemann–integrierbaren Funktionen monoton wächst und gegen die Grenzfunktion f konvergiert, und die Folge der Integrale $\{\int_0^1 f_n dx\}_{n=1}^{\infty}$ konvergiert, ist die Grenzfunktion nicht Riemann–integrierbar. Dies ist ein erheblicher Nachteil des Riemann–Integrals.

Es liegt aber nahe, den Integralbegriff auszudehnen und auch der Funktion f ein Integral zuzuordnen durch die Definition

$$\int_0^1 f(x)dx := \lim_{n\to\infty} \int_0^1 f_n(x)dx\,.$$

f ist die Grenzfunktion einer monoton wachsenden Folge von Riemann–integrierbaren Funktionen mit konvergenter Integralfolge; f ist dann selbst wieder integrierbar und das Integral von f ist gleich dem Grenzwert der Integralfolge $\{f_n\}_{n=1}^{\infty}$. Dies ist die Grundidee bei der Einführung eines leistungsfähigeren Integralbegriffes, des Lebesgue–Integrals. Dieses Lebesgue–Integral soll nun definiert und untersucht werden. Ich werde dabei von Anfang an die Integration von Funktionen mehrerer Variabler untersuchen.

3 a.) Nullmengen

Definition: Für $i = 1, \ldots, n$ seien I_i Intervalle der Form $(-\infty, \infty)$, $(-\infty, b_i)$, $[a_i, \infty)$, $[a_i, b_i)$. Die Menge

$$Q = \prod_{i=1}^{n} I_i \subseteq \mathbb{R}^n$$

heißt n–dimensionaler, halboffener Quader. Der Quader heißt unbeschränkt, wenn wenigstens eines der Intervalle unbeschränkt ist, andernfalls heißt der Quader beschränkt. Für beschränkte Quader $Q = \prod_{i=1}^{n}[a_i, b_i)$ sei

$$|Q| = \prod_{i=1}^{n}(b_i - a_i)$$

das „Maß“ von Q.

Definition: Die Menge $N \subseteq \mathbb{R}^n$ heißt Nullmenge oder Menge vom Maß 0, wenn es zu jedem $\varepsilon > 0$ höchstens abzählbar viele n–dimensionale, beschränkte Quader $Q_1, Q_2, \ldots$ gibt mit

$$N \subseteq \bigcup_{j=1}^{\infty} Qj\,, \quad \sum_{j=1}^{\infty} |Q_j| < \varepsilon\,.$$

Satz: (i) Jede Teilmenge einer Nullmenge ist eine Nullmenge.

(ii) Die Vereinigung höchstens abzählbar vieler Nullmengen ist wieder eine Nullmenge.

(iii) Abzählbare Teilmengen von $\mathbb{R}^n$ sind Nullmengen.

Beweis: Der Beweis der Aussage (i) ist klar.

(ii) Seien $N_1, N_2, N_3, \ldots$ Nullmengen und sei $\varepsilon > 0$. Dann gibt es Quader $\{Q_{kj}\}_{k,j=1,2,\ldots}$ mit

$$N_k \subseteq \bigcup_{j=1}^{\infty} Q_{kj}\,, \quad \sum_{j=1}^{\infty} |Q_{kj}| \leq \frac{1}{2^k}\,\varepsilon\,.$$

Für die abzählbare Menge $\{Q_{kj}\}_{k,j=1,2,\ldots}$ gilt also

$$\bigcup_{k=1}^{\infty} N_k \subseteq \bigcup_{k,j=1}^{\infty} Q_{kj}\,, \quad \sum_{j,k=1}^{\infty} |Q_{kj}| = \sum_{k=1}^{\infty} \sum_{j=1}^{\infty} |Q_{kj}| \leq \sum_{k=1}^{\infty} \frac{1}{2^k}\,\varepsilon = \varepsilon\,.$$

(Man beachte, daß nach dem großen Umordnungssatz die Reihenfolge der Summation in $\sum_{k,j=1}^{\infty} |Q_{k,j}|$ beliebig ist.) Somit ist $\bigcup_{k=1}^{\infty} N_k$ eine Nullmenge.

(iii) Es ist klar, daß die leere Menge eine Nullmenge ist. Sei $N = \{x\}$ eine einpunktige Menge im $\mathbb{R}^n$. Sei $x = (x_1, \ldots, x_n)$, $\varepsilon > 0$, und sei

$$Q = \Big[x_1 - \frac{\varepsilon}{4}, x_1 + \frac{\varepsilon}{4}\Big) \times \Big[x_2 - \frac{1}{2}, x_2 + \frac{1}{2}\Big) \times \ldots \times \Big[x_n - \frac{1}{2}, x_n + \frac{1}{2}\Big)\,.$$

Dann gilt $N \subseteq Q$ und

$$|Q| = \frac{\varepsilon}{2} \cdot 1 \cdot \ldots \cdot 1 = \frac{\varepsilon}{2} < \varepsilon\,.$$

Also ist N eine Nullmenge. Nach (ii) ist dann jede abzählbar Menge eine Nullmenge. ■

Beispiele: 1.) $\mathbb{Q}$ ist eine Nullmenge von $\mathbb{R}$, weil $\mathbb{Q}$ abzählbar ist.

2.) Sei N der Graph einer stetigen Funktion $f : \mathbb{R}^{n-1} \to \mathbb{R}$, d.h.

$$N = \left\{ x \in \mathbb{R}^n \,\middle|\, x = (y, f(y))\,, y \in \mathbb{R}^{n-1} \right\}.$$

Dann ist N Nullmenge.

Zum Beweis genügt es, f auf einem Quader $Q \subseteq \mathbb{R}^{n-1}$ mit der Kantenlänge 1 zu betrachten und zu zeigen, daß der Graph von $f|_Q$ eine Nullmenge ist. Denn $\mathbb{R}^{n-1}$ kann von abzählbar vielen solchen Quadern überdeckt werden. f ist auf der kompakten Menge $\overline{Q}$ und damit auch auf Q gleichmäßig stetig, also gibt es zu jedem $\varepsilon > 0$ ein $\delta > 0$ mit $|f(y) - f(z)| < \varepsilon$ für $y, z \in Q$ mit $\|y - z\|_\infty < \delta$. Wähle $k \in \mathbb{N}$, $\frac{1}{k} < \delta$, und zerlege Q in k^{n-1} Teilquader mit der Kantenlänge $\frac{1}{k}$. Dann ist der Graph von f auf jedem dieser Quader enthalten in einem Quader mit dem Volumen $(\frac{1}{k})^{n-1}\varepsilon$. Der Graph von f über Q ist in der Vereinigungsmenge dieser k^{n-1} Quader enthalten. Für die Summe der Volumen dieser Quader gilt

$$k^{n-1} \left(\frac{1}{k}\right)^{n-1} \varepsilon = \varepsilon\,,$$

also ist N eine Nullmenge.

3.) Hieraus folgt, daß alle Hyperebenen im $\mathbb{R}^n$ Nullmengen sind.

4.) Die Cantor–Menge: Jede reelle Zahl a kann dargestellt werden durch eine triadische Entwicklung:

$$a = \sum_{k=-\ell}^{\infty} a_k 3^{-k}\,,$$

wobei $a_k = 0, 1$ oder 2 ist. Die Zahlen im Intervall $[0, 1)$ sind genau diejenigen Zahlen mit einer Entwicklung der Form

$$a = \sum_{k=1}^{\infty} a_k 3^{-k}\,.$$

Die Cantor–Menge C besteht aus allen Zahlen in $[0, 1)$, deren triadische Entwicklung nur die Ziffern 0 und 2 enthält:

$$C = \{0, a_1 a_2 a_3 \ldots \,\big|\, a_k = 0 \quad \text{oder} \quad 2\}\,.$$

Geometrisch kann diese Menge folgendermaßen dargestellt werden: Alle Zahlen, deren triadische Entwicklung mit 0 oder 2 beginnt, liegen in den schraffierten Bereichen:

Alle Zahlen, deren triadische Entwicklung in den ersten beiden Stellen nur 0 oder 2 enthält, liegen in den folgenden Bereichen:

Man fahre so fort und nehme jeweils das mittlere Drittel der verbleibenden Intervalle heraus.

In jedem Schritt besteht die schraffierte Menge aus endlich vielen halboffenen „Quadern“ . Die Summe der Volumen dieser Quader ist gleich der Gesamtlänge der schraffierten Menge. In jedem Schritt verringert sich diese Gesamtlänge auf $\frac{2}{3}$ der Länge im vorangehenden Schritt. Im k–ten Schritt ist diese Länge also $(\frac{2}{3})^k$. Wegen $(\frac{2}{3})^k \to 0$ für $k \to \infty$, und da C in jedem Schritt in der schraffierten Menge enthalten ist, folgt daß C eine Nullmenge ist.
C ist aber überabzählbar. Denn folgendermaßen kann eine bijektive Abbildung f von C auf das Intervall $[0, 1)$ konstruiert werden: Setze

$$f\Big(\sum_{k=1}^{\infty} a_k 3^{-k}\Big) = \sum_{k=1}^{\infty} \Big(\frac{1}{2}\, a_k\Big) 2^{-k} ,$$

wobei rechts eine Binärentwicklung steht.

Definition: Sei $D \subseteq \mathbb{R}^n$. Man sagt, eine Eigenschaft gelte für fast alle $x \in D$ oder kürzer, fast überall auf D , wenn es eine Nullmenge N gibt, so daß die Eigenschaft für alle $x \in D \backslash N$ gilt.

3 b.) Treppenfunktionen

Definition: Unter einer Zerlegung Z des halboffenen Quaders $Q \subseteq \mathbb{R}^n$ in Unterquader versteht man eine endliche Menge von n–dimensionalen, halboffenen Quadern

$$Z = \{Q_1, \ldots, Q_m\}$$

mit

$$Q_k \cap Q_j = \emptyset\,, \quad k \neq j\,, \quad \text{und} \quad \bigcup_{j=1}^{m} Q_j = Q\,.$$

Definition: Sei Q ein beschränkter halboffener Quader. Die Funktion $\varphi : Q \to \mathbb{R}$ heißt Treppenfunktion auf Q, wenn es eine Zerlegung $Z = \{Q_1, \ldots, Q_m\}$ von Q gibt, so daß φ auf jedem Unterquader Q_j einen konstanten Wert c_j hat. Das Integral von φ über Q ist definiert durch

$$\int_Q \varphi dx = \int_Q \varphi(x) dx = \sum_{j=1}^{m} c_j |Q_j| \,. \qquad (*)$$

Zur Definition des Integrals von φ können verschiedene Zerlegungen von Q gewählt werden. Zunächst muß gezeigt werden, daß die Definition des Integrals sinnvoll ist, also nicht von der Zerlegung abhängt. Sei $Z' = \{Q'_1, \ldots, Q'_\ell\}$ eine andere Zerlegung von Q, so daß die Treppenfunktion φ auf jedem Q'_j einen konstanten Wert c'_j hat. Falls $Q_i \cap Q'_j \neq \emptyset$ ist, folgt $c_i = c'_j$. Wegen $|Q_i \cap Q'_j| = 0$ für $Q_i \cap Q'_j = \emptyset$, gilt also in jedem Fall

$$c_i |Q_i \cap Q'_j| = c'_j |Q_i \cap Q'_j| \,.$$

Wegen

$$\bigcup_{i=1}^{m} (Q_i \cap Q'_j) = (\bigcup_{i=1}^{m} Q_i) \cap Q'_j = Q \cap Q'_j = Q'_j$$

folgt

$$\sum_{i=1}^{\infty} |Q_i \cap Q'_j| = \Big| \bigcup_{i=1}^{m} (Q_k \cap Q'_j) \Big| = |Q'_j| \,,$$

und in ähnlicher Weise

$$\sum_{j=1}^{\ell} |Q_i \cap Q'_j| = |Q_i| \,.$$

Also ergibt sich

$$\begin{aligned}
\sum_{i=1}^{m} c_i |Q_i| &= \sum_{i=1}^{m} c_i \sum_{j=1}^{\ell} |Q_i \cap Q'_j| \\
&= \sum_{j=1}^{\ell} \sum_{i=1}^{m} c_i |Q_i \cap Q'_j| = \sum_{j=1}^{\ell} \sum_{i=1}^{m} c'_j |Q_i \cap Q'_j| \\
&= \sum_{j=1}^{\ell} c'_j \sum_{i=1}^{m} |Q_i \cap Q'_j| = \sum_{j=1}^{\ell} c'_j |Q'_j| \,.
\end{aligned}$$

Die Summe auf der rechten Seite von $(*)$ ist also unabhängig von der gewählten Zerlegung von Q, und somit ist die Definition von $\int_Q \varphi dx$ sinnvoll.

Definition: Sei $M \subseteq \mathbb{R}^n$ ein beschränkter oder unbeschränkter, n–dimensionaler Quader. Die Funktion $\varphi : M \to \mathbb{R}$ heißt Treppenfunktion auf M, wenn es einen n–dimensionalen, beschränkten, halboffenen Quader $Q \subseteq \overline{M}$ gibt und wenn φ von M zu einer auch mit φ bezeichneten Funktion auf $\overline{M}$ fortgesetzt werden kann, so daß $\varphi|_Q$ eine Treppenfunktion auf Q ist, und so daß $\varphi|_{M\setminus Q} = 0$ gilt. Das Integral von φ über M ist definiert durch

$$\int_M \varphi dx = \int_Q \varphi|_Q dx\,.$$

$T(M)$ sei die Menge aller Treppenfunktionen auf M.

Die Wahl des Quaders Q in dieser Definition ist im allgemeinen nicht eindeutig bestimmt. Trotzdem ist die Definition des Integrals von φ über M sinnvoll, weil der Wert von $\int_Q \varphi|_Q dx$ nicht von der Wahl von Q abhängt. Denn seien Q und Q' zwei halboffene Quader mit $\varphi|_{M\setminus Q} = \varphi|_{M\setminus Q'} = 0$. Jeder Punkt x mit $\varphi(x) \neq 0$ gehört dann sowohl zu Q als auch zu Q', folglich ist auch $P = Q \cap Q'$ ein halboffener Quader mit $\varphi|_{M\setminus P} = 0$. Es ist P ein Unterquader sowohl von Q als auch von Q', und man sieht sofort, daß $\int_Q \varphi|_Q dx = \int_P \varphi|_P dx$ und $\int_{Q'} \varphi|_{Q'} dx = \int_P \varphi|_P dx$, also auch

$$\int_Q \varphi|_Q dx = \int_{Q'} \varphi|_{Q'} dx$$

gilt.

Lemma a.) Seien φ und ψ Treppenfunktionen auf M und sei $c \in \mathbb{R}$. Dann ist auch $(\varphi + \psi) \in T(M)$ und $c\varphi \in T(M)$ mit

(i) $\displaystyle\int_M \varphi + \psi dx = \int_M \varphi dx + \int_M \psi dx$

(ii) $\displaystyle\int_M c\varphi dx = c \int_M \varphi dx\,.$

Also ist $T(M)$ ein Vektorraum und das Integral ist eine lineare Abbildung auf $T(M)$ mit Werten in $\mathbb{R}$.

b.) Sind $\varphi, \psi \in T(M)$ mit $\varphi(x) \le \psi(x)$ für alle $x \in M$, dann folgt

$$\int_M \varphi dx \le \int_M \psi dx\,.$$

Beweis: a.) Es gibt halboffene beschränkte Quader $P, P' \subseteq \overline{M}$, so daß $\varphi|_P$ und $\psi|_{P'}$ Treppenfunktionen sind mit $\varphi|_{M\setminus P} = 0$, $\psi|_{M\setminus P'} = 0$. Sei $Q \subseteq \overline{M}$ ein beschränkter halboffener Quader mit $P, P' \subseteq Q$. φ und ψ sind dann auch Treppenfunktionen auf Q. Seien $\{Q_1, \ldots, Q_m\}$ und $\{Q_1', \ldots, Q_\ell'\}$ Zerlegungen von Q, so daß φ auf Q_i den konstanten Wert c_i und ψ auf Q_j' den konstanten Wert c_j' hat. Dann ist $\{Q_i \cap Q_j' \mid i = 1, \ldots, m\,;\, j = 1, \ldots, \ell\}$ eine Zerlegung von Q. Falls der beschränkte, halboffene Quader $Q_i \cap Q_j' \neq \emptyset$ ist, hat φ den konstanten Wert c_i und ψ den konstanten Wert c'_j auf diesem Quader, also hat $\varphi+\psi$ den konstanten Wert $c_i+c'_j$ auf dem Quader. Somit ist $\varphi+\psi$ eine Treppenfunktion auf Q mit $(\varphi+\psi)|_{M\setminus Q} = 0$.

Weil $|Q_i \cap Q'_j| = 0$ ist falls $Q_i \cap Q'_j = \emptyset$ gilt, folgt

$$\begin{aligned}
\int_M \varphi + \psi dx = \int_Q \varphi + \psi dx &= \sum_{i=1}^m \sum_{j=1}^\ell (c_i + c'_j)|Q_i \cap Q'_j| \\
&= \sum_{i=1}^m c_i \sum_{j=1}^\ell |Q \cap Q'_j| + \sum_{j=1}^\ell c'_j \sum_{i=1}^m |Q_i \cap Q'_j| \\
&= \sum_{i=1}^m c_i |Q_i| + \sum_{j=1}^\ell c_j |Q'_j| = \int_Q \varphi dx + \int_Q \psi dx \\
&= \int_M \varphi dx + \int_M \psi dx\,.
\end{aligned}$$

Dies beweist (i). Der Beweis, daß $c\varphi \in T(M)$ ist und daß (ii) gilt, ist klar.

b.) Falls $\varphi \le \psi$ gilt und $Q_i \cap Q'_j \neq \emptyset$ ist, folgt $c_i \le c'_j$ für die konstanten Werte von φ und ψ auf $Q_i \cap Q'_j$. Also ergibt sich

$$\begin{aligned}
\int_M \varphi dx &= \sum_{i=1}^m \sum_{j=1}^\ell c_i |Q_i \cap Q'_j| \le \sum_{i=1}^m \sum_{j=1}^\ell c'_j |Q_i \cap Q'_j| \\
&= \int_M \psi dx\,.
\end{aligned}$$

∎

Definition: Sei $A \subseteq \mathbb{R}^n$ eine Menge. Die durch

$$\chi_A(x) = \begin{cases} 1\,, & x \in A \\ 0\,, & x \in \mathbb{R}^n \setminus A \end{cases}$$

definierte Funktion $\chi_A : \mathbb{R}^n \to \mathbb{R}$ heißt charakteristische Funktion der Menge A.

Jede Treppenfunktion kann in der Form

$$\varphi(x) = \sum_{j=1}^m c_j \chi_{Q_j}(x)$$

mit geeigneten Quadern $Q_1, \ldots, Q_m \subseteq \mathbb{R}^n$ geschrieben werden.

3 c.) Definition des Lebesgueschen Integrals

Definition: Sei $D \subseteq \mathbb{R}^n$. Eine Folge $\{\varphi_m\}_{m=1}^\infty$ von Funktionen $\varphi_m : D \to \mathbb{R}$ heißt wachsend (fallend), wenn die Folge $\{\varphi_m(x)\}_{m=1}^\infty$ monoton wachsend (fallend) ist für alle $x \in D$.

Definition: Sei $M \subseteq \mathbb{R}^n$ ein n–dimensionaler Quader. Die Menge $\mathcal{L}^+(M)$ besteht aus allen Funktionen $f : M \to \mathbb{R}$ mit folgender Eigenschaft: Es gibt eine wachsende Folge $\{\varphi_m\}_{m=1}^\infty$ von Treppenfunktionen $\varphi_m \in T(M)$, die fast überall auf M gegen f strebt, und deren Integralfolge $\{\int_M \varphi_m dx\}_{m=1}^\infty$ konvergiert.

Nach dem oben bewiesenen Lemma ist die Folge der Integrale $\{\int_M \varphi_m dx\}_{m=1}^\infty$ monoton wachsend, wenn $\{\varphi_m\}_{m=1}^\infty$ monoton wachsend ist. Also ist die Folge $\{\int_M \varphi_m dx\}_{m=1}^\infty$ konvergent genau dann wenn sie beschränkt ist.

Jedem $f \in \mathcal{L}^+(M)$ möchte man nun ein Integral zuordnen durch die Definition

$$\int_M f dx = \lim_{m\to\infty} \int_M \varphi_m dx \,,$$

wobei φ_m eine solche wachsende Folge von Treppenfunktionen ist, die fast überall gegen f konvergiert. Es gibt aber verschiedene solche Folgen $\{\varphi_m\}_{m=1}^\infty$. Daher mußt zuerst gezeigt werden, daß der Grenzwert auf der rechten Seite von der speziell gewählten Folge unabhängig ist. Hierzu benötigt man zwei Hilfssätze:

Hilfssatz: Sei $Q \subseteq \mathbb{R}^n$ ein beschränkter halboffener Quader, sei $\{\varphi_m\}_{m=1}^\infty$ eine monoton fallende Folge von Treppenfunktionen $\varphi_m : Q \to \mathbb{R}$ mit $\varphi_m(x) \geq 0$ für alle $m \in \mathbb{N}$ und alle $x \in Q$ und mit $\lim_{m\to\infty} \varphi_m(x) = 0$ für fast alle $x \in Q$. Dann gilt:

$$\lim_{m\to\infty} \int_Q \varphi_m dx = 0 \,.$$

Beweis: **I.)** Nach Definition gibt es zur Treppenfunktion φ_m eine Zerlegung von Q in endlich viele halboffene Quader $Q_1, \ldots, Q_t$, so daß φ_m konstant ist auf jeder der offenen Mengen

$$\overset{\circ}{Q}_1, \ldots, \overset{\circ}{Q}_t \,.$$

Sei V_m die Vereinigung dieser Mengen. V_m ist eine offene Menge, und das Komplement $\overline{Q} \backslash V_m$ ist gleich der Menge $N_m = \bigcup_{j=1}^t \partial Q_j$, wobei ∂Q_j die Menge der Randpunkte von Q_j ist. Die Seiten eines Quaders sind Teilmenge einer Hyperebene im $\mathbb{R}^n$, und damit Nullmengen. Also ist ∂Q_j eine Nullmenge, also auch N_m, und somit auch

$$N' = \bigcup_{m=1}^\infty N_m \,.$$

Nach Definition gibt es eine Nullmenge N'' mit $\varphi_m(x) \searrow 0$ für $m \to \infty$ und für alle $x \in Q \backslash N''$. Es sei nun

$$N = N' \cup N''.$$

Dies ist eine Nullmenge.

Sei $\varepsilon > 0$. Dann gibt es abzählbar viele beschränkte Quader $W_1, W_2, \ldots$ mit

$$N \subseteq \bigcup_{\ell=1}^{\infty} W_\ell \,, \quad \sum_{\ell=1}^{\infty} |W_\ell| < \varepsilon \,.$$

In dem man W_ℓ etwas vergrößert, kann man abzählbar viele halboffene Quader $W_1', W_2', \ldots$ finden mit $W_\ell \subseteq \overset{\circ}{W'}_\ell$ und mit

$$N \subseteq \bigcup_{\ell=1}^{\infty} \overset{\circ}{W'}_\ell \,, \quad \sum_{\ell=1}^{\infty} |W_\ell'| < \varepsilon \,. \tag{$*$}$$

II.) Sei nun $x \in \overline{Q} \backslash N = Q \backslash N$. Dann ist $x \notin N''$, also gilt $\varphi_m(x) \searrow 0$ für $m \to \infty$, und somit gibt es $k(x) \in \mathbb{N}$ mit $\varphi_{k(x)}(x) < \varepsilon$.

Es ist $x \notin N'$, also $x \notin N_{k(x)}$, also gehört x zur Menge $V_{k(x)}$, die aus endlich vielen offenen Quadern besteht, auf denen φ jeweils konstant ist. Somit gehört x zu einem dieser Quader, der mit $\overset{\circ}{P}(x)$ bezeichnet werde. Nach Konstruktion ist $\overset{\circ}{P}(x)$ das Innere eines halboffenen Quaders $P(x)$, auf dem φ konstant ist. Weil $\{\varphi_m\}_{m=1}^{\infty}$ monoton fallend ist, folgt also

$$\varphi_m(y) \le \varphi_{k(x)}(y) = \varphi_{k(x)}(x) < \varepsilon \tag{$**$}$$

für alle $y \in P(x)$ und alle $m \ge k(x)$.

Es ist $\{\overset{\circ}{P}(x)\}_{x \in \overline{Q} \backslash N} \cup \{\overset{\circ}{W'}_1, \overset{\circ}{W'}_2, \ldots\}$ eine offene Überdeckung der kompakten Menge $\overline{Q}$, also gibt es endlich viele Quader $\overset{\circ}{P}(x_1), \ldots, \overset{\circ}{P}(x_s)$ und $\overset{\circ}{W'}_{\ell_1}, \ldots, \overset{\circ}{W'}_{\ell_2}$, die schon eine Überdeckung von $\overline{Q}$ bilden. Die Menge der halboffenen Quader $\{P(x_1), \ldots, P(x_s),$ $W_{\ell_1}', \ldots, W_{\ell_r}'\}$ überdeckt dann erst recht Q. Sei $k_0 = \max\{k(x_1), \ldots, k(x_s)\}$. Aus $(**)$ folgt dann

$$\varphi_m(y) < \varepsilon \tag{$+$}$$

für alle $y \in [P(x_1) \cup \ldots \cup P(x_s)]$ und alle $m \ge k_0$. Für alle $y \in Q$ gilt

$$\varphi_m(y) \le \varphi_1(y) \le \max_{z \in Q} \varphi_1(z) =: K \,.$$

Wegen $(Q \backslash \bigcup_{j=1}^r W_{\ell_j}') \subseteq \bigcup_{i=1}^s P(x_i)$ ergibt sich aus $(+)$ also

$$\varphi_m(y) \le \begin{cases} \varepsilon \,, & y \in Q \backslash \bigcup_{j=1}^r W_{\ell_j}' \\ K \,, & y \in Q \cap \bigcup_{j=1}^r W_{\ell_j}' \,, \end{cases}$$

für $m \geq k_0$. Somit folgt mit den charakteristischen Funktionen von Q und von W'_{ℓ_j}, daß

$$\varphi_m(y) \leq \varepsilon \chi_Q(y) + \sum_{j=1}^{r} K \chi_{W'_{\ell_j}}(y)$$

gilt für alle $y \in Q$ und $m \geq k_0$, also

$$\begin{aligned}
\int_Q \varphi_m dy \leq \varepsilon \int_Q \psi_Q dy + \sum_{j=1}^{r} K \int_Q \chi_{W'_{\ell_j}} dy \\
&= \varepsilon |Q| + \sum_{j=1}^{r} K |Q \cap W'_{\ell_j}| \\
&\leq \varepsilon |Q| + K \sum_{j=1}^{r} |W'_{\ell_j}| \leq \varepsilon |Q| + K \sum_{\ell=1}^{\infty} |W'_\ell| \leq \varepsilon(|Q| + K)\,.
\end{aligned}$$

Da $\varepsilon > 0$ beliebig gewählt war, folgt $\lim_{n\to\infty} \int_Q \varphi_m dx = 0$. ∎

Hilfssatz: Seien $M \subseteq \mathbb{R}^n$ ein n–dimensionaler Quader, seien $f, g \in \mathcal{L}^+(M)$ mit $f(x) \leq g(x)$ für fast alle $x \in M$. Seien $\{\varphi_m\}_{m=1}^{\infty}$, $\{\psi_m\}_{m=1}^{\infty}$ wachsende Folgen von Treppenfunktionen $\varphi_m, \psi_m : M \to \mathbb{R}$ mit beschränkten Integralfolgen. $\{\varphi_m\}_{m=1}^{\infty}$ konvergiere fast überall gegen f, und $\{\psi_m\}_{m=1}^{\infty}$ konvergiere fast überall gegen g. Dann gilt

$$\lim_{m\to\infty} \int_M \varphi_m dx \leq \lim_{m\to\infty} \int_M \psi_m dx\,.$$

Beweis: Für festes k ist die Folge $\{\varphi_k - \psi_m\}_{m=1}^{\infty}$ monoton fallend und strebt fast überall auf M gegen den Grenzwert $\varphi_k - g$, wobei für fast alle $x \in M$ gilt

$$\varphi_k(x) - g(x) \leq f(x) - g(x) \leq 0\,.$$

Sei

$$(\varphi_k - \psi_m)^+(x) = \begin{cases} (\varphi_k - \psi_m)(x)\,, & \text{falls} \quad (\varphi_k - \psi_m)(x) \geq 0\,, \\ 0\,, & \text{sonst}\,. \end{cases}$$

Somit ist auch die Folge $\{(\varphi_k - \psi_m)^+\}_{m=1}^{\infty}$ von nichtnegativen Treppenfunktionen monoton fallend und konvergiert fast überall gegen 0. Außerhalb eines beschränkten halboffenen Quaders Q verschwindet $(\varphi_k - \psi_1)^+$. Wegen $0 \leq (\varphi_k - \psi_m)^+ \leq (\varphi_k - \psi_1)^+$ verschwinden dann auch alle $(\varphi_k - \psi_m)^+$ außerhalb von Q. Also sind die Voraussetzungen des vorangehenden Hilfssatzes erfüllt, und es folgt

$$\begin{aligned}
\int_M \varphi_k dx - \lim_{m\to\infty} \int_M \psi_m dx = \lim_{m\to\infty} \Big[\int_M \varphi_k dx - \int_M \psi_m dx \Big] \\
&= \lim_{m\to\infty} \int_M (\varphi_k - \psi_m) dx \leq \lim_{m\to\infty} \int_Q (\varphi_k - \psi_m)^+ dx = 0\,.
\end{aligned}$$

Diese Abschätzung gilt für jedes k , also folgt

$$\lim_{k\to\infty}\int_M \varphi_k dx \le \lim_{m\to\infty}\int_M \psi_m dx\,.$$

Damit ist der Satz bewiesen. ∎

Nun kann das Lebesgue-Integral für Funktionen aus $\mathcal{L}^+(M)$ definiert werden:

Definition: Sei M ein n–dimensionaler Quader, $f \in \mathcal{L}^+(M)$, und $\{\varphi_m\}_{m=1}^\infty$ eine wachsende Folge von Treppenfunktionen $\varphi_m : M \to \mathbb{R}$ mit beschränkter Integralfolge, die fast überall gegen f konvergiert. Dann ist das Lebesguesche Integral von f definiert durch

$$\int_M f dx = \lim_{m\to\infty}\int_M \varphi_m dx\,.$$

Der Grenzwert auf der rechten Seite dieser Gleichung hängt nicht von der gewählten Folge von Treppenfunktionen ab. Denn seien $\{\varphi_m\}_{m=1}^\infty$ und $\{\psi_m\}_{m=1}^\infty$ wachsende Folgen von Treppenfunktionen $\varphi_m,\, \psi_m \in T(M)$ mit $\lim_{m\to\infty}\varphi_m(x) = f(x)$ fast überall, $\lim_{m\to\infty}\psi_m(x) = f(x)$ fast überall, und mit beschränkten Integralfolgen. Aus dem obigen Hilfssatz folgt dann

$$\lim_{m\to\infty}\int_M \varphi_m dx = \lim_{m\to\infty}\int_M \psi_m dx.$$

Beispiele: **1.)** Jede stetige beschränkte Funktion $f : M \to \mathbb{R}$ auf einem beschränkten Quader $M \subseteq \mathbb{R}^n$ gehört zu $\mathcal{L}^+(M)$. Zum Beweis muß eine wachsende Folge $\{\varphi_m\}_{m=1}^\infty$, $\varphi_m \in T(M)$, konstruiert werden, die fast überall gegen f konvergiert, und deren Integralfolge beschränkt ist. Hierzu wähle man eine Folge $\{Z_m\}_{m=1}^\infty$ von Zerlegungen von M mit folgenden Eigenschaften:

(i) Z_{m+1} sei eine Verfeinerung von Z_m , d.h. zu jedem Quader $Q \in Z_{m+1}$ gibt es einen Quader $P \in Z_m$ mit $Q \subseteq P$.

(ii) Es sei $\lim_{k\to\infty}\sup_{Q\in Z_k}\delta(Q) = 0$, wobei $\delta(Q)$ der Durchmesser von Q sei.

Nun definiere man $\varphi_m \in T(M)$ durch

$$\varphi_m(x) = \sum_{Q\in Z_m}\Big(\inf_{y\in Q} f(y)\Big)\chi_Q(x)\,.$$

Dann ist die Folge $\{\varphi_m(x)\}_{m=1}^\infty$ wachsend für alle $x \in M$ mit $\lim_{m\to\infty}\varphi_m(x) = f(x)$. Zum Beweis der Monotonie beachte man, daß zu jedem x Quader $Q \in Z_{m+1}, P \in Z_m$ existieren mit $x \in Q \subseteq P$, also

$$\varphi_{m+1}(x) = \inf_{y\in Q} f(y) \geq \inf_{y\in P} f(y) = \varphi_m(x)\,.$$

Die Folge der Integrale ist beschränkt wegen

$$\int_M \varphi_m dx \leq \int_M \Big(\sup_{y\in M} f(y)\Big) dx = |Q| \sup_{y\in M} f(y)\,.$$

Somit ist $f \in \mathcal{L}^+(M)$.

2.) Die Dirichletfunktion $f : [0,1] \to \mathbb{R}$,

$$f(x) = \begin{cases} 1\,, & x \in \mathbb{Q} \cap [0,1] \\ 0\,, & \text{sonst} \end{cases}$$

gehört zu $\mathcal{L}^+([0,1])$. Weil $[0,1] \cap \mathbb{Q}$ eine Nullmenge ist, ist $\{\varphi_m\}_{m=1}^\infty$ mit

$$\varphi_m \equiv 0$$

eine wachsende Folge von Treppenfunktionen, die fast überall gegen f konvergiert, und deren Integralfolge beschränkt ist. Für das Integral von f gilt

$$\int_0^1 f dx = \lim_{m\to\infty} \int_0^1 \varphi_m dx = 0.$$

3.) Für $n, k \in \mathbb{N}$ mit $1 \leq k \leq n-1$ sei

$$I_{nk} = \Big[\frac{k}{n} - \frac{1}{4(n-1)}\, 2^{-n}, \frac{k}{n} + \frac{1}{4(n-1)}\, 2^{-n}\Big] \subseteq [0,1]$$

ein Intervall. Die Menge

$$I = \bigcup_{n=1}^{\infty} \bigcup_{k=1}^{n-1} I_{nk} \subseteq [0,1]$$

besteht aus abzählbar vielen nichtleeren Intervallen mit positivem Maß, ist also keine Nullmenge. Für Ihr „Maß“ gilt

$$|I| \leq \sum_{n=1}^{\infty} \sum_{k=1}^{n-1} |I_{nk}| \leq \sum_{n=1}^{\infty} \sum_{k=1}^{n-1} \frac{1}{2(n-1)}\, 2^{-n} = \frac{1}{2}\,,$$

also ist auch ihr Komplement $[0,1]\backslash I$ keine Nullmenge.
Die charakteristische Funktion $f : [0,1] \to \mathbb{R}$ von I,

$$f(x) = \begin{cases} 1\,, & x \in I \\ 0\,, & x \in [0,1]\backslash I \end{cases}$$

gehört zu $\mathcal{L}^+([0,1])$. Denn definiert man $\varphi_m \in T([0,1])$ durch

$$\varphi_m(x) = \begin{cases} 1\,, & x \in \bigcup_{n=1}^{m} \bigcup_{k=1}^{n-1} I_{nk} \\ 0\,, & \text{sonst} \end{cases} \,,$$

dann ist $\{\varphi_m\}_{m=1}^{\infty}$ eine wachsende Folge von Treppenfunktionen (siehe das unten folgende Bild) mit $\lim_{m\to\infty} \varphi_m(x) = f(x)$ für alle $x \in [0,1]$ und mit beschränkter Integralfolge:

$$\int_0^1 \varphi_m dx \le \int_0^1 1\, dx = 1\,.$$

Jedoch ist $-f \notin \mathcal{L}^+([0,1])$. Zum Beweis sei angenommen, daß eine wachsende Folge $\{\varphi_m\}_{m=1}^{\infty}$ mit $\varphi_m \in T([0,1])$ existiere, die fast überall gegen $-f$ konvergiere. Für eine beliebig gewählte Funktion φ_s aus dieser Folge sei $[a,b)$ mit $a < b$ ein Intervall, auf dem φ_s konstant ist. Dann gibt es eine rationale Zahl $\frac{k}{n} \in [a,b)$ mit $1 \le k \le n-1$, also gehört mindestens die Hälfte des Intervalls I_{nk} zu $[a,b)$, folglich ist

$$I_{nk} \cap [a,b)$$

keine Nullmenge. Hieraus folgt, daß der konstante Wert von φ_s auf $[a,b)$ kleiner oder gleich -1 sein muß, weil sonst für alle $m \ge s$ und alle $x \in I_{nk} \cap [a,b)$

$$\varphi_m(x) \ge \varphi_s(x) > -1 = -f(x)\,,$$

also $\lim_{m\to\infty} \varphi_m(x) \ne -f(x)$ gelten würde. Weil $I_{nk} \cap [a,b)$ keine Nullmenge ist, widerspricht dies der Voraussetzung.

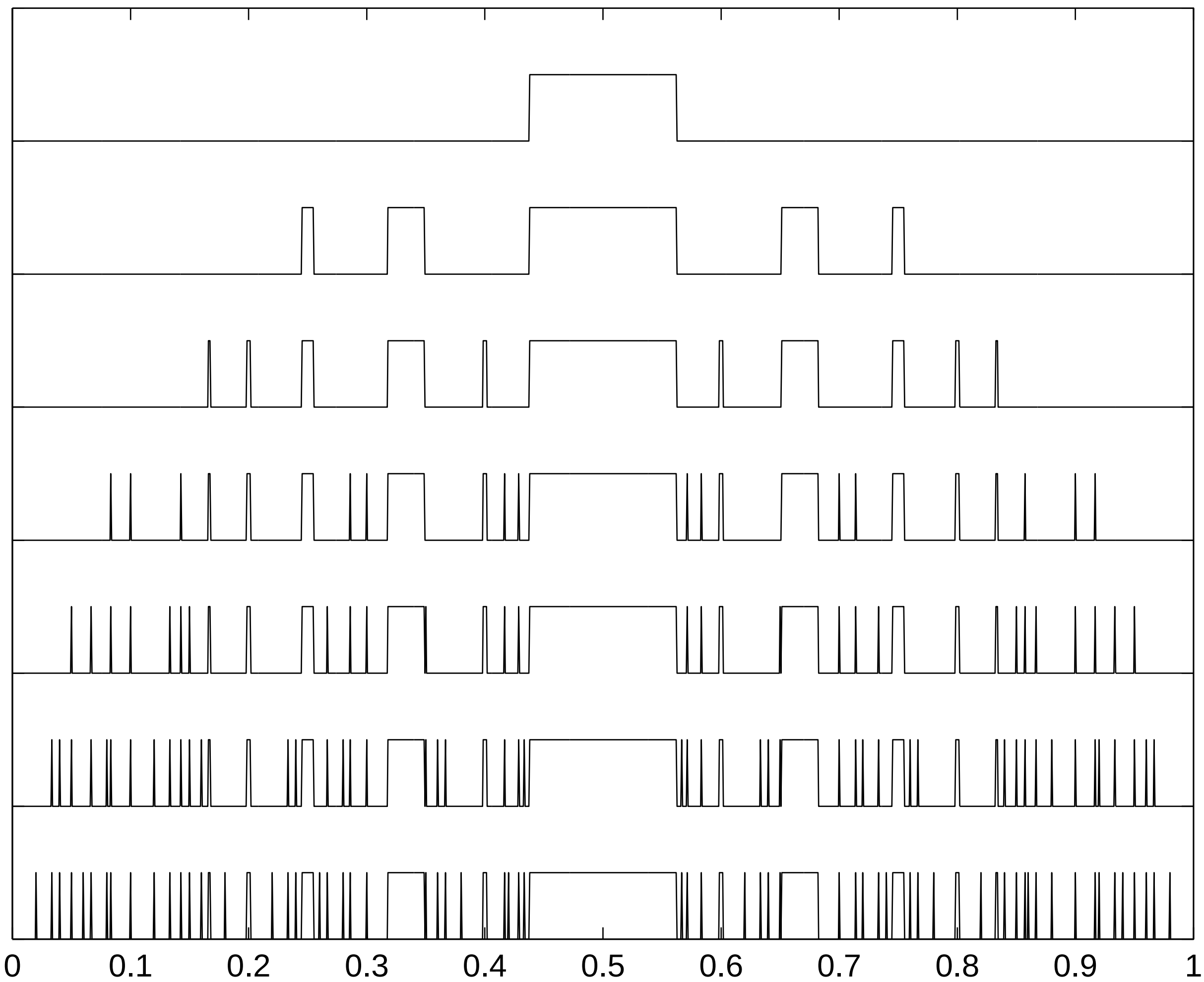
Weil $[a,b)$ ein beliebig gewähltes Intervall war, auf dem φ_s konstant ist, folgt daß

$$\varphi_s(x) \le -1$$

gilt für alle $x \in [0,1]$, und weil φ_s ein beliebiges Element von $\{\varphi\}_{m=1}^{\infty}$ war, gilt $\varphi_m \le -1$ für alle $m \in \mathbb{N}$. Für $x \in [0,1]\backslash I$ folgt somit

$$\lim_{m\to\infty} \varphi_m(x) \le -1 < -f(x) = 0\,,$$

also konvergiert $\{\varphi_m\}_{m=1}^{\infty}$ nicht fast überall gegen $-f$, weil $[0,1]\backslash I$ keine Nullmenge ist. Dies beweist, daß $-f \notin \mathcal{L}^+([0,1])$.



Graph der Treppenfunktionen φ_2, φ_4, φ_6, φ_{13}, φ_{20}, φ_{30} und φ_{50} aus Beispiel 3.

Das letzte Beispiel zeigt, daß $\mathcal{L}^+(M)$ kein Vektorraum ist, weil sonst mit f auch $-f$ zu $\mathcal{L}^+(M)$ gehören müßte. Natürlich sollte aber mit f auch die Funktion $-f$ integrierbar sein. Also muß $\mathcal{L}^+(M)$ in geeigneter Weise zu einem Vektorraum erweitert werden. Dazu benutzt man folgendes Resultat für $\mathcal{L}^+(M)$:

Satz: Wenn $f, g \in \mathcal{L}^+(M)$ und $c \geq 0$, dann sind auch $f + g \in \mathcal{L}^+(M)$, $cf \in \mathcal{L}^+(M)$ und

$$\int_M (f+g)dx = \int_M f dx + \int_M g dx, \quad \int_M cf dx = c \int_M f dx \, .$$

Beweis: Es seien $\{\varphi_m\}_{n=1}^\infty, \{\psi_m\}_{m=1}^\infty$ wachsende Folgen von Treppenfunktionen mit $\lim_{m\to\infty} \varphi_m(x) = f(x)$ fast überall, $\lim_{m\to\infty} \psi_m(x) = g(x)$ fast überall, und mit beschränkten Integralfolgen. Dann ist $\{\varphi_m + \psi_m\}_{m=1}^\infty$ eine wachsende Folge von Treppenfunktionen, die fast überall gegen $f+g$ konvergiert, und

$$\begin{aligned}\int_M (f+g)dx &= \lim_{m\to\infty}\int_M (\varphi_m+\psi_m)dx \\ &= \lim_{m\to\infty}\int_M \varphi_m dx + \lim_{m\to\infty}\int_M \psi_m dx = \int_M f dx + \int_M g dx\,.\end{aligned}$$

Ebenso folgt dies für cf. ■

Nun kann $\mathcal{L}^+(M)$ zu einem Vektorraum $\mathcal{L}(M)$ erweitert werden:

Definition: Sei M ein n-dimensionaler Quader. $\mathcal{L}(M)$ sei die Menge aller Funktionen $f=g-h$ mit $g,h\in\mathcal{L}^+(M)$. Für $f\in\mathcal{L}(M)$ definiert man das Lebesgue Integral durch

$$\int_M f dx = \int_M g dx - \int_M h dx\,.$$

Diese Definition ist sinnvoll, weil das Integral von f nicht von der Darstellung von f abhängt. Denn seien $f=g_1-h_1$ und $f=g_2-h_2$ mit $g_1,h_1,g_2,h_2\in\mathcal{L}^+(M)$. Dann folgt $g_1-h_1=g_2-h_2$, also $g_1+h_2=g_2+h_1\in\mathcal{L}^+(M)$ und

$$\int_M g_1 dx + \int_M h_2 dx = \int_M (g_1+h_2)dx = \int_M (g_2+h_1)dx = \int_M g_2 dx + \int_M h_1 dx\,,$$

folglich

$$\int_M g_1 dx - \int_M h_1 dx = \int_M g_2 dx - \int_M h_2 dx\,,$$

und somit ist $\int_M f dx$ unabhängig von der gewählten Zerlegung von f.

Definition: Sei $A\subseteq\mathbb{R}^n$ eine beliebige Menge. $\mathcal{L}(A)$ sei die Menge aller Funktionen $f:A\to\mathbb{R}$, für die die Fortsetzung $\tilde f:\mathbb{R}^n\to\mathbb{R}$,

$$\tilde f(x) = \begin{cases} f(x)\,, & x\in A \\ 0\,, & x\in\mathbb{R}^n\setminus A \end{cases}$$

zu $\mathcal{L}(\mathbb{R}^n)$ gehört. Die Funktionen $f\in\mathcal{L}(A)$ heißen Lebesgue-integrierbar über A mit Lebesgue-Integral

$$\int_A f dx = \int_{\mathbb{R}^n}\tilde f dx\,.$$

Daß $\mathcal{L}(A)$ ein Vektorraum ist, wird im nächsten Abschnitt bewiesen.

3 d.) Einfache Eigenschaften des Lebesgue–Integrals

Satz: Sei $A \subseteq \mathbb{R}^n$ eine beliebige Menge, sei $f_1 \in \mathcal{L}(A)$ und sei fast überall $f_2 = f_1$. Dann ist auch $f_2 \in \mathcal{L}(A)$ und

$$\int_A f_2 dx = \int_A f_1 dx\,.$$

Beweis: Zunächst werde angenommen, daß $A = M$ sei mit einem n–dimensionalen Quader M, und daß $f_1 \in \mathcal{L}^+(M)$ sei. Dann existiert eine wachsende Folge von Treppenfunktionen $\{\varphi_m\}_{m=1}^\infty$, die fast überall gegen f_1 konvergiert, und deren Integralfolge beschränkt ist. Dann konvergiert diese Folge auch fast überall gegen f_2, also $f_2 \in \mathcal{L}^+(M)$ und

$$\int_M f_2 dx = \lim_{m\to\infty} \int_M \varphi_m dx = \int_M f_1 dx\,.$$

Um die Aussage für $f_1 \in \mathcal{L}(A)$ mit beliebigem A zu beweisen, genügt es bereits, sie für $f_1 \in \mathcal{L}(\mathbb{R}^n)$ zu zeigen. Sei also $f_1 = g_1 - h_1 \in \mathcal{L}(\mathbb{R}^n)$ mit $g_1, h_1 \in \mathcal{L}^+(\mathbb{R}^n)$. Weil $h_2 = h_1 + (f_1 - f_2)$ fast überall mit h_1 übereinstimmt, ist $h_2 \in \mathcal{L}^+(\mathbb{R}^n)$, und weil $\mathbb{R}^n$ ein Quader ist, folgt aus dem soeben bewiesenen Resultat, daß

$$\int_{\mathbb{R}^n} h_2 dx = \int_{\mathbb{R}^n} h_1 dx\,.$$

Wegen $f_2 = g_1 - h_2$ ist dann auch $f_2 \in \mathcal{L}(\mathbb{R}^n)$ mit

$$\int_{\mathbb{R}^n} f_2 dx = \int_{\mathbb{R}^n} g_1 dx - \int_{\mathbb{R}^n} h_2 dx = \int_{\mathbb{R}^n} g_1 dx - \int_{\mathbb{R}^n} h_1 dx = \int_{\mathbb{R}^n} f_1 dx\,.$$

■

Satz: a.) Für $f, g \in \mathcal{L}(A)$ und $c \in \mathbb{R}$ gilt $(f+g) \in \mathcal{L}(A)$ und $cf \in \mathcal{L}(A)$ mit

(i) $\displaystyle\int_A (f+g)dx = \int_A f dx + \int_A g dx\,,$

(ii) $\displaystyle\int_A cf dx = c\int_A f dx\,.$

Also ist $\mathcal{L}(A)$ ein Vektorraum und das Integral ist eine lineare Abbildung auf $\mathcal{L}(A)$ mit Werten in $\mathbb{R}$.

b.) Sind $f, g \in \mathcal{L}$ mit $f(x) \geq g(x)$ fast überall in A, dann folgt

$$\int_A f dx \geq \int_A g dx\,.$$

Beweis: Es genügt diesen Satz für $A = \mathbb{R}^n$ zu beweisen.

a.) Seien $f = f_1 - f_2$, $g = g_1 - g_2$ mit $f_1, f_2, g_1, g_2 \in \mathcal{L}^+(\mathbb{R}^n)$. Dann sind $f_1 + g_1, f_2 + g_2 \in$

$\mathcal{L}^+(\mathbb{R}^n)$. Wegen $f+g=(f_1+g_1)-(f_2+g_2)$ folgt also $f+g=\mathcal{L}(\mathbb{R}^n)$ und

$$\begin{aligned}
\int_{\mathbb{R}^n}(f+g)dx &= \int_{\mathbb{R}^n}(f_1+g_1)dx-\int_{\mathbb{R}^n}(f_2-g_2)dx\\
&= \int_{\mathbb{R}^n} f_1dx+\int_{\mathbb{R}^n} g_1dx-\Big[\int_{\mathbb{R}^n} f_2dx+\int_{\mathbb{R}^n} g_2dx\Big]\\
&= \Big[\int_{\mathbb{R}^n} f_1dx-\int_{\mathbb{R}^n} f_2dx\Big]+\Big[\int_{\mathbb{R}^n} g_1dx-\int_{\mathbb{R}^n} g_2dx\Big]\\
&= \int_{\mathbb{R}^n} fdx+\int_{\mathbb{R}^n} gdx\,.
\end{aligned}$$

Daß $cf=c(f_1-f_2)\in\mathcal{L}(\mathbb{R}^n)$ ist mit $\int_{\mathbb{R}^n} cf dx = c\int_{\mathbb{R}^n} f dx$ sieht man, wenn man cf in der Form

$$\begin{aligned}
cf &= cf_1-cf_2\,, && \text{falls} \quad c\geq 0\\
cf &= (-c)f_2-(-c)f_1\,, && \text{falls} \quad c<0
\end{aligned}$$

schreibt, und wenn man benützt, daß $dg\in\mathcal{L}^+(\mathbb{R}^n)$ ist für $d\geq 0$ und $g\in\mathcal{L}^+(\mathbb{R}^n)$ mit $\int_{\mathbb{R}^n} dgdx = d\int_{\mathbb{R}^n} gdx$. ■

b.) Sei $f\geq 0$ fast überall und $f=f_1-f_2$ mit $f_1,f_2\in\mathcal{L}^+(\mathbb{R}^n)$. Dann ist $f_1\geq f_2$ fast überall, und es existieren wachsende Folgen $\{\varphi_m\}_{m=1}^\infty$, $\{\psi_m\}_{m=1}^\infty$ von Treppenfunktionen $\varphi_m,\psi_m\in T(\mathbb{R}^n)$ mit beschränkter Integralfolge, die fast überall gegen f_1 beziehungsweise f_2 konvergieren. Nach dem früher bewiesenen Hilfssatz folgt

$$\int_{\mathbb{R}^n} f_1dx=\lim_{m\to\infty}\int_{\mathbb{R}^n}\varphi_m dx\geq\lim_{m\to\infty}\int_{\mathbb{R}^n}\psi_m dx=\int_{\mathbb{R}^n} f_2dx\,,$$

also

$$\int_{\mathbb{R}^n} fdx=\int_{\mathbb{R}^n} f_1dx-\int_{\mathbb{R}^n} f_2dx\geq 0\,.$$

Falls $f,g\in\mathcal{L}(\mathbb{R}^n)$ ist mit $f\geq g$ fast überall, folgt $f-g\in\mathcal{L}(\mathbb{R}^n)$ und $f-g\geq 0$ fast überall, also

$$\int_{\mathbb{R}^n} fdx-\int_{\mathbb{R}^n} gdx=\int_{\mathbb{R}^n}(f-g)dx\geq 0\,.$$

■

Satz: Seien $f,g\in\mathcal{L}(A)$. Dann gehören auch $\max(f,g)$, $\min(f,g)$, f^+, f^- und $|f|$ zu $\mathcal{L}(A)$ und es gilt

$$\Big|\int_A fdx\Big|\leq\int_A|f|dx\,.$$

Hierbei sei

$$f^+(x) \;=\; \begin{cases} f(x)\,, & \text{falls} \quad f(x) \geq 0 \\ 0\,, & \text{falls} \quad f(x) < 0 \end{cases}$$

$$f^-(x) \;=\; \begin{cases} 0\,, & \text{falls} \quad f(x) \geq 0 \\ -f(x)\,, & \text{falls} \quad f(x) < 0\,. \end{cases}$$

Beweis: Es genügt wieder, diesen Satz für $A = \mathbb{R}^n$ zu beweisen. Erst wird gezeigt, daß $|f| \in \mathcal{L}(\mathbb{R}^n)$ ist. Sei $f = f_1 - f_2$ mit $f_1, f_2 \in \mathcal{L}^+(\mathbb{R}^n)$. Wegen $|f| = \max(f_1, f_2) - \min(f_1, f_2)$ genügt es zu zeigen, daß $\max(f_1, f_2) \in \mathcal{L}^+(\mathbb{R}^n)$ und $\min(f_1, f_2) \in \mathcal{L}^+(\mathbb{R}^n)$ sind.

Seien $\{\varphi_m\}_{m=1}^\infty$, $\{\psi_m\}_{m=1}^\infty$ wachsende Folgen von Treppenfunktionen mit beschränkten Integralfolgen, die fast überall gegen f_1 beziehungsweise f_2 konvergieren. Dann sind $\{\max(\varphi_m, \psi_m)\}_{m=1}^\infty$, $\{\min(\varphi_m, \psi_m)\}_{m=1}^\infty$ wachsende Folgen von Treppenfunktionen mit

$$\begin{aligned} \max(\varphi_m, \psi_m) &\nearrow \max(f_1, f_2)\,, \\ \min(\varphi_m, \psi_m) &\nearrow \min(f_1, f_2) \end{aligned}$$

fast überall. Sei $h = \min(\varphi_1, \psi_1)$. Dies ist eine Treppenfunktion, also auch $\max(\varphi_m - h, \psi_m - h) = \max(\varphi_m, \psi_m) - h$, und wegen $\varphi_m - h \geq 0$, $\psi_m - h \geq 0$ folgt

$$\begin{aligned} \int_{\mathbb{R}^n} \max(\varphi_m, \psi_m) dx &= \int_{\mathbb{R}^n} \Big[\max(\varphi_m - h, \psi_m - h) + h\Big] dx \\ &\leq \int_{\mathbb{R}^n} \Big[(\varphi_m - h) + (\psi_m - h) + h\Big] dx \\ &= \int_{\mathbb{R}^n} \varphi_m dx + \int_{\mathbb{R}^n} \psi_m dx - \int_{\mathbb{R}^n} h dx \\ &\leq \int_{\mathbb{R}^n} f_1 dx + \int_{\mathbb{R}^n} f_2 dx - \int_{\mathbb{R}^n} h dx\,. \end{aligned}$$

Folglich ist die Folge der Integral $\{\int_{\mathbb{R}^n} \max(\varphi_m, \psi_m) dx\}_{m=1}^\infty$ beschränkt, und somit folgt $\max(f_1, f_2) \in \mathcal{L}^+(\mathbb{R}^n)$. Ebenso folgt dies für $\min(f_1, f_2)$.

Also ist $|f| \in \mathcal{L}(\mathbb{R}^n)$. Wegen

$$\begin{aligned} f^+ &= \frac{1}{2}(|f| + f)\,, \;\; f^- = \frac{1}{2}(|f| - f)\,, \\ \max(f, g) &= \frac{1}{2}(f + g + |f - g|)\,, \;\; \min(f, g) = \frac{1}{2}(f + g - |f - g|) \end{aligned}$$

folgen die anderen Aussagen hieraus und aus dem zuletzt bewiesenen Satz.

Wegen $f \leq |f|$, $-f \leq |f|$ folgt

$$\int_A f dx \leq \int_A |f| dx\,, \;\; -\int_A f dx = \int_A (-f) dx \leq \int_A |f| dx\,,$$

also
$$\Big|\int_A f dx\Big| \leq \int_A |f| dx\,.$$

∎

4 Lebesgue–Integration und Grenzübergang. Der Satz von Fubini

4 a.) Vorbereitende Resultate und Satz von Beppo Levi für Funktionen aus $\mathcal{L}^+$

Im Gegensatz zum Riemannschen Integral hat man für das Lebesguesche Integral starke Grenzwertsätze. Diese Grenzwertsätze sind eine der wichtigsten Rechtfertigungen für die Einführung des Lebesgueschen Integrals. In diesem Abschnitt werden einige vorbereitende Resultate zum Beweis dieser Sätze hergeleitet.

Es seien $Q_1, \ldots, Q_m \subseteq \mathbb{R}^n$ beschränkte, paarweise disjunkte Quader und $F = Q_1 \cup \ldots \cup Q_m$ die Vereinigungsmenge. Das Maß von F wird definiert durch

$$|F| = \sum_{k=1}^{m} |Q_k| \,.$$

Dieses Maß hängt nicht von der Zerlegung von F in disjunkte Quader ab.

Lemma: Sei $K > 0$ und für jedes $m \in \mathbb{N}$ sei $F_m \subseteq \mathbb{R}^n$ Vereinigung von endlich vielen beschränkten, paarweise disjunkten Quadern mit $|F_m| \leq K$ und mit $F_m \subseteq F_{m+1}$. Dann ist $F = \bigcup_{m=1}^{\infty} F_m$ Vereinigung von abzählbar vielen beschränkten, paarweise disjunkten Quadern $Q_1, Q_2, \ldots$ mit

$$\sum_{k=1}^{\infty} |Q_k| \leq K.$$

Beweis: Es sei $F_0 = \emptyset$. Für k, $m \in \mathbb{N}$ mit $k < m$ gilt dann wegen $F_k \subseteq F_{m-1}$

$$(F_m \setminus F_{m-1}) \cap (F_k \setminus F_{k-1}) \subseteq (F_m \setminus F_{m-1}) \cap F_k = \emptyset$$

und

$$\begin{aligned}(F_m \setminus F_{m-1}) &\cup \ldots \cup (F_3 \setminus F_2) \cup (F_2 \setminus F_1) \cup (F_1 \setminus F_0) \\ &= (F_m \setminus F_{m-1}) \cup \ldots \cup (F_3 \setminus F_2) \cup F_2 = \ldots = F_m \,,\end{aligned}$$

also kann F in der Form

$$F = \bigcup_{m=1}^{\infty} F_m = \bigcup_{m=1}^{\infty} (F_m \setminus F_{m-1})$$

als Vereinigung abzählbar vieler disjunkter Mengen geschrieben werden. Man sieht schnell, daß jede der Mengen $F_m \setminus F_{m-1}$ in endlich viele beschränkte, paarweise disjunkte Quader zerlegt werden kann, also ist F die Vereinigung abzählbar vieler paarweise disjunkter Quader

$$F = \bigcup_{k=1}^{\infty} Q_k \,,$$

wobei mit geeigneten Zahlen $s_1 = 1 < s_2 < s_3 < \ldots$ gilt

$$F_m \setminus F_{m-1} = \bigcup_{k=s_m}^{s_{m+1}-1} Q_k \,.$$

Folglich ist $F_m = \bigcup_{k=1}^{s_{m+1}-1} Q_k$, also ergibt sich für jedes m

$$\sum_{k=1}^{s_{m+1}-1} |Q_k| = |F_m| \leq K,$$

und somit auch $\sum_{k=1}^{\infty} |Q_k| \leq K$, weil die Folge $\{s_{m+1} - 1\}_{m=1}^{\infty}$ unbeschränkt wächst.

■

Hilfssatz: Sei $M \subseteq \mathbb{R}^n$ ein n–dimensionaler Quader und sei $\{\varphi_m\}_{m=1}^{\infty}$ eine wachsende Folge von Treppenfunktionen aus $T(M)$. Wenn die Folge $\{\int_M \varphi_m dx\}_{m=1}^{\infty}$ beschränkt ist, dann konvergiert $\{\varphi_m\}_{m=1}^{\infty}$ fast überall gegen eine Funktion aus $\mathcal{L}^+(M)$.

Beweis: Sei N die Menge aller $x \in M$, für die $\lim_{m\to\infty} \varphi_m(x)$ nicht existiert, für die also die Folge $\{\varphi_m(x)\}_{m=1}^{\infty}$ unbeschränkt ist. Es genügt zu zeigen, daß N eine Nullmenge ist. Denn dann ist die Funktion $f : M \to \mathbb{R}$,

$$f(x) = \begin{cases} \lim\limits_{m\to\infty} \varphi_m(x) \,, & x \in M \setminus N \\ 0 \,, & x \in N \,, \end{cases}$$

nach Definition von $\mathcal{L}^+(M)$ in diesem Raum enthalten.

Zum Beweis, daß N eine Nullmenge ist, sei o.B.d.A. $\varphi_m \geq 0$ für alle m . Sonst betrachte man die Folge $\{\varphi_m - \varphi_1\}_{m=1}^{\infty}$, die ebenfalls auf N divergiert und auf $M \setminus N$ konvergiert. Zu $\varepsilon > 0$ definiere man die Menge

$$F_{\varepsilon,m} = \{x \in M \mid \varphi_m(x) \geq K/\varepsilon\} \,,$$

mit der Konstanten

$$K = \sup_{m\in\mathbb{N}} \int_M \varphi_m dx \,.$$

Es gilt

$$N \subseteq F_\varepsilon := \bigcup_{m=1}^{\infty} F_{\varepsilon,m} \,.$$

Denn für $x \in N$ wächst die Folge $\{\varphi_m(x)\}_{m=1}^\infty$ unbeschränkt, also existiert m_0 mit $\varphi_{m_0}(x) \geq K/\varepsilon$, und somit ist $x \in F_{\varepsilon,m_0} \subseteq F_\varepsilon$.

Weil φ_m eine Treppenfunktion ist, besteht jede der Mengen $F_{\varepsilon,m}$ aus endlich vielen beschränkten, paarweise disjunkten Quadern, und weil $\{\varphi_m\}_{m=1}^\infty$ wachsend ist, gilt $F_{\varepsilon,m} \subseteq F_{\varepsilon,m+1}$. Wegen $\varphi_m \geq 0$ auf M und $\varphi_m(x) \geq K/\varepsilon$ für $x \in F_{\varepsilon,m}$ ergibt sich auch

$$\frac{K}{\varepsilon}\,|F_{\varepsilon,m}| \leq \int_{F_{\varepsilon,m}} \varphi_m(x)\,dx \leq \int_M \varphi_m(x)\,dx \leq K\,,$$

also $|F_{\varepsilon,m}| \leq \varepsilon$. Nach dem vorangehenden Lemma ist also F_ε Vereinigung von abzählbar vielen Quadern $Q_1, Q_2, \ldots$ mit $\sum_{k=1}^\infty |Q_k| \leq \varepsilon$. Weil $N \subseteq F_\varepsilon = \bigcup_{k=1}^\infty Q_k$ gilt und weil ε beliebig gewählt war, bedeutet dies, daß daß N eine Nullmenge ist. ■

Dieser Satz kann von Treppenfunktionen auf wachsende Folgen von Funktionen aus $\mathcal{L}^+$ verallgemeinert werden:

Satz: von Beppo Levi für Funktionen aus $\mathcal{L}^+$: Sei $M \subseteq \mathbb{R}^n$ ein n–dimensionaler Quader, und sei $\{f_m\}_{m=1}^\infty$ eine wachsende Folge von Funktionen aus $\mathcal{L}^+(M)$ mit beschränkter Integralfolge $\{\int_M f_m dx\}_{m=1}^\infty$. Dann konvergiert $\{f_m\}_{m=1}^\infty$ fast überall gegen eine Funktion $f \in \mathcal{L}^+(M)$ und

$$\lim_{m\to\infty} \int_M f_m dx = \int_M f dx\,.$$

Beweis: Für jedes $m \in \mathbb{N}$ sei $\{\varphi_{mj}\}_{j=1}^\infty$ eine wachsende Folge von Treppenfunktionen aus $T(M)$ und N_m eine Nullmenge mit $\lim_{j\to\infty} \varphi_{mj}(x) = f(x)$ für alle $x \in M \backslash N_m$. Dann gilt

$$\lim_{j\to\infty} \int_M \varphi_{mj} dx = \int_M f_m dx\,.$$

Setze

$$\varphi_m(x) := \max\{\varphi_{kj}(x) \,\big|\, j,k = 1,\, \ldots, m\}$$

für alle $x \in M$. Dann ist $\{\varphi_m\}_{m=1}^\infty$ eine wachsende Folge von Treppenfunktionen. Da $\varphi_{kj}(x) \leq f_k(x)$ gilt für alle $x \in M\backslash N_k$ und alle j, folgt

$$\varphi_m(x) = \max\{\varphi_{kj}(x) \,\big|\, j,k = 1,\, \ldots, m\} \leq \max\{f_k(x) \,\big|\, k = 1,\, \ldots, m\} = f_m(x) \qquad (*)$$

für alle $x \in M\backslash N$, mit der Nullmenge $N = \bigcup_{k=1}^\infty N_k$. Wegen der Monotonie des Lebesgue–Integrals resultiert hieraus

$$\int_M \varphi_m dx \leq \int_M f_m dx \leq \sup_{m\in\mathbb{N}} \int_M f_m dx < \infty\,,$$

also liefert der Hilfssatz die Existenz einer Funktion $f \in \mathcal{L}^+(M)$ und einer Nullmenge N' mit $\lim_{m\to\infty} \varphi_m(x) = f(x)$ für alle $x \in M \backslash N'$, und mit

$$\int_M f dx = \lim_{m\to\infty} \int_M \varphi_m dx\,. \tag{+}$$

Hieraus folgert man

$$f_k(x) = \lim_{m\to\infty} \varphi_{km}(x) \le \lim_{m\to\infty} \varphi_m(x) = f(x)$$

für alle $x \in M\backslash(N' \cup N_k)$, und somit ergibt $(*)$, daß die Ungleichungen

$$\varphi_m(x) \le f_m(x) \le f(x)$$

erfüllt sind für alle $x \in M\backslash N''$ und für alle $m \in \mathbb{N}$, mit der Nullmenge $N'' = N \cup \bigcup_{n=1}^\infty (N' \cup N_m) = N' \cup N$. Hieraus folgt

$$\lim_{m\to\infty} f_m(x) = f(x)$$

für fast alle $x \in M$ und

$$\int_M \varphi_m dx \le \int_M f_m dx \le \int_M f dx\,,$$

also

$$\lim_{m\to\infty} \int_M f_m dx = \int_M f dx\,,$$

wegen $(+)$. ■

Um diesen Satz zum allgemeinen Satz von Beppo Levi für Funktionen aus $\mathcal{L}$ zu verallgemeinern, benötigt man ein weiteres Hilfsresultat:

Hilfssatz: Sei $M \subseteq \mathbb{R}^n$ ein n–dimensionaler Quader und sei $f \in \mathcal{L}(M)$. Dann gibt es zu jedem $\varepsilon > 0$ Funktionen $g, h \in \mathcal{L}^+(M)$ mit $f = g - h, \;\; h \ge 0$ und

$$0 \le \int_M h dx < \varepsilon\,.$$

Beweis: Sei $f = g_1 - h_1$ mit $g_1, h_1 \in \mathcal{L}^+(M)$, sei $\varepsilon > 0$ und sei $\{\varphi_m\}_{m=1}^\infty$ eine wachsende Folge von Treppenfunktionen, die fast überall gegen h_1 konvergiert und deren Integralfolge beschränkt ist. Dann gibt es $k \in \mathbb{N}$ mit

$$0 \le \int_M h_1 dx - \int_M \varphi_k dx < \varepsilon\,.$$

Aus der Ungleichung $\varphi_k(x) \le h_1(x)$, die für fast alle x erfüllt ist, folgt daß $(h_1 - \varphi_k)^- = 0$ ist fast überall in M, mit der im letzten Kapitel eingeführten Bezeichnung $(h_1 - \varphi_k)^-$ für den negativen Anteil von $(h_1-\varphi_k)$. Folglich ist $(h_1-\varphi_k)^- \in \mathcal{L}^+(M)$ und $\int_M (h_1-\varphi_k)^- dx = 0$. Die gesuchte Zerlegung von f erhält man nun in der Form

$$f = g - h$$

mit

$$g = g_1 - \varphi_k + (h_1 - \varphi_k)^- \,, \quad h = h_1 - \varphi_k + (h_1 - \varphi_k)^- \,.$$

Denn $-\varphi_k$ ist eine Treppenfunktion und gehört somit zu $\mathcal{L}^+(M)$. Wegen $g_1 \in \mathcal{L}^+(M)$ und $(h_1 - \varphi_k)^- \in \mathcal{L}^+(M)$ folgt also

$$g = g_1 - \varphi_k + (h_1 - \varphi_k)^- = g_1 + (-\varphi_k) + (h_1 - \varphi_k)^- \in \mathcal{L}^+(M) \,,$$

und ebenso $h \in \mathcal{L}^+(M)$. Außerdem ist

$$h(x) = (h_1 - \varphi_k)(x) + (h_1 - \varphi_k)^-(x) = (h_1 - \varphi_k)^+(x) \geq 0$$

für alle $x \in M$ und

$$0 \leq \int_M h dx = \int_M (h_1 - \varphi_k) dx + \int_M (h_1 - \varphi_k)^- dx = \int_M h_1 dx - \int_M \varphi_k dx < \varepsilon \,.$$

■

4 b.) Grenzübergang bei monotoner und dominierter Konvergenz

Satz von Beppo Levi über monotone Konvergenz: Sei $A \subseteq \mathbb{R}^n$ und sei $\{f_m\}_{m=1}^\infty$ eine wachsende Folge von Funktionen aus $\mathcal{L}(A)$ mit beschränkter Integralfolge. Dann konvergiert $\{f_m\}_{m=1}^\infty$ fast überall gegen eine Funktion $f \in \mathcal{L}(A)$ und

$$\int_A f dx = \lim_{m\to\infty} \int_A f_m dx \,.$$

Beweis: Es genügt den Satz für $A = \mathbb{R}^n$ zu beweisen. O.B.d.A. kann angenommen werden, daß $f_m \geq 0$ gilt für alle m , weil man sonst $\{f_m - f_1\}_{m=1}^\infty$ betrachten kann. Sei $f_0 \equiv 0$. Dann kann man die Folge $\{f_m\}_{m=1}^\infty$ als Reihe

$$\sum_{m=1}^{\infty} (f_m - f_{m-1})$$

mit nichtnegativen Gliedern schreiben. Nach dem obenstehenden Hilfssatz gibt es zu jedem m Funktionen $g_m, h_m \in \mathcal{L}^+(\mathbb{R}^n)$ mit

$$f_m - f_{m-1} = g_m - h_m \,, \quad h_m \geq 0 \,, \quad \int_{\mathbb{R}^n} h_m dx \leq 2^{-m} \,.$$

Seien

$$s_m = \sum_{j=1}^{m} g_j \,, \quad t_m = \sum_{j=1}^{m} h_j \,.$$

Dann gilt $s_m, t_m \in \mathcal{L}^+(\mathbb{R}^n)$, und wegen $g_m = (f_m - f_{m-1}) + h_m \geq 0$ sind die Folgen $\{s_m\}_{m=1}^{\infty}$ und $\{t_m\}_{m=1}^{\infty}$ wachsend. Außerdem gilt

$$f_m = \sum_{j=1}^{m} (f_j - f_{j-1}) = \sum_{j=1}^{m} (g_j - h_j) = s_m - t_m \,.$$

Die Folge $\{\int_{\mathbb{R}^n} t_m dx\}_{m=1}^{\infty}$ ist beschränkt wegen

$$\int_M t_m dx = \sum_{j=1}^{\infty} \int_M h_j dx \leq \sum_{j=1}^{m} 2^{-j} \leq 1 \,,$$

also ist auch $\{\int_{\mathbb{R}^n} s_m dx\}_{m=1}^{\infty}$ beschränkt wegen

$$\int_{\mathbb{R}^n} s_m dx = \int_{\mathbb{R}^n} f_m dx + \int_M t_m dx \,.$$

Nach dem Satz von Beppo Levi für Funktionen aus $\mathcal{L}^+$ konvergieren $\{s_m\}_{m=1}^{\infty}$ und $\{t_m\}_{m=1}^{\infty}$ fast überall gegen Funktionen $s, t \in \mathcal{L}^+(\mathbb{R}^n)$ mit

$$\int_M s dx = \lim_{n\to\infty} \int_M s_m dx \,, \qquad \int_M t dx = \lim_{m\to\infty} \int_M t_m dx \,.$$

Sei nun

$$f = s - t \,.$$

Dann ist $f \in \mathcal{L}(\mathbb{R}^n)$ und

$$\lim_{m\to\infty} f_m(x) = \lim_{m\to\infty} \Big[s_m(x) - t_m(x) \Big] = s(x) - t(x) = f(x)$$

für fast alle x, sowie

$$\lim_{m\to\infty} \int f_m dx = \lim_{m\to\infty} \int_M (s_m - t_m) dx = \int_M s dx - \int_M t dx = \int_M f dx \,.$$

$\blacksquare$

Folgerung: Sei $A \subseteq \mathbb{R}^n$. Für jede Funktion $f \in \mathcal{L}(A)$ gilt $f(x) = 0$ für fast alle $x \in A$, genau dann wenn $\int_A |f| dx = 0$ ist.

Beweis: Es wurde schon gezeigt, daß $\int_A |f| dx = \int_A 0 dx = 0$ ist, wenn f fast überall mit der Nullfunktion übereinstimmt. Zum Beweis der Umkehrung sei $\int_A |f| dx = 0$. Die Folge $\{m|f|\}_{m=1}^{\infty}$ ist wachsend und besteht aus Funktionen von $\mathcal{L}(A)$. Wegen

$$\int_A m|f| dx = m \int_A |f| dx = 0$$

ist die Integralfolge beschränkt, also existiert der Limes

$$\lim_{m\to\infty} m|f(x)| = |f(x)| \lim_{m\to\infty} m$$

für fast alle $x \in A$, was nur möglich ist wenn $f(x) = 0$ ist für fast alle $x \in A$. ∎

Folgerung: Sei $-\infty < a < b < \infty$ und sei f Riemann–integrierbar über $[a, b]$. Dann ist $f \in \mathcal{L}^+([a, b])$ und

$$L - \int_a^b f dx = R - \int_a^b f dx \,,$$

wobei $L - \int_A^b$ das Lebesgue–Integral und $R - \int_a^b$ das Riemann–Integral bedeutet.

Beweis: Sei $f \in \mathcal{R}([a, b])$. Zu jeder Partition $P = \{x_0, x_1, \ldots, x_\ell\}$ von $[a, b]$ definiere man Treppenfunktionen $x \mapsto t(x, P)$ und $x \mapsto s(x, P)$ durch

$$t(x, P) = \sum_{i=1}^{\ell} c_i' \chi_{[x_{i-1}, x_i)}(x) \,, \quad s(x, P) = \sum_{i=1}^{\ell} c_i \chi_{[x_{i-1}, x_i)}(x)$$

mit den Konstanten

$$c_i' = \sup_{y \in [x_{i-1}, x_i)} f(y) \,, \quad c_i = \inf_{y \in [x_{i-1}, x_i)} f(y) \,.$$

Mit der Bezeichnung aus Kapitel 2 ist dann

$$\int_a^b t(x, P) dx = U(P, f) \,, \ \int_a^b s(x, P) dx = L(P, f) \,.$$

Wähle nun eine Folge von Partitionen $\{P_m\}_{m=1}^{\infty}$ mit

$$\lim_{m\to\infty} L(P_m, f) = \lim_{m\to\infty} \int_a^b s(x, P_m) dx = \underline{\int_a^b} f dx = R - \int_a^b f dx \,,$$

und setze $P_m' = \bigcup_{k=1}^m P_k$. Dann ist die Partition P_m' Verfeinerung von P_k für $k = 1, \ldots, m$, also gilt

$$\int_a^b s(x, P_m') dx \geq \int_a^b s(x, P_k) dx \,, \quad k = 1, \ldots, m \,,$$

also folgt $\lim_{m\to\infty} \int_a^b s(x, P_m') dx = R - \int_a^b f dx$. Außerdem ist P_{m+1}' eine Verfeinerung von P_m', also folgt nach Definition der Treppenfunktionen

$$s(x, P_{m+1}') \geq s(x, P_m') \,,$$

also ist $\{s(\cdot, P_m')\}_{m=1}^{\infty}$ eine wachsende Folge von Treppenfunktionen mit beschränkter Integralfolge. Aus dem Satz von Beppo Levi schließt man nun, daß diese Treppenfunktionsfolge fast überall gegen eine Funktion $g \in \mathcal{L}^+([a, b])$ konvergiert mit

$$L - \int_a^b g dx = \lim_{m\to\infty} \int_a^b s(x, P_m') dx = R - \int_a^b f dx \,.$$

Weil für alle x und m nach Definition der Treppenfunktion $s(x,P'_m) \leq f(x)$ gilt, folgt $g \leq f$ fast überall.

Auf dieselbe Weise konstruiert man eine fallende Folge von Treppenfunktionen $\{t(\cdot,P''_m)\}_{m=1}^{\infty}$, die nach dem Satz von Beppo Levi (der auch für fallende Funktionenfolgen gilt) fast überall gegen eine Funktion $h \in \mathcal{L}([a,b])$ konvergiert mit

$$L-\int_a^b h dx = \lim_{m\to\infty}\int_a^b t(x,P''_m)dx = R-\int_a^b f dx$$

und mit $f \leq h$ fast überall.
Weil fast überall $g \leq f \leq h$ gilt, resultiert somit

$$\begin{aligned} L-\int_a^b |h-g|dx &= L-\int_a^b (h-g)dx = (L-\int_a^b hdx)-(L-\int_a^b gdx) \\ &= (R-\int_a^b fdx)-(R-\int_a^b fdx) = 0\,, \end{aligned}$$

also ergibt die voranstehende Folgerung $h = g$ fast überall, und somit $g = f = h$ fast überall, also $f \in \mathcal{L}^+([a,b])$ und

$$L-\int_a^b fdx = L-\int_a^b g\,dx = R-\int_a^b fdx\,.$$

■

Beispiel 1: Für jedes $s < -1$ und $a > 0$ is $f(x) = x^s$ integrierbar über $[a,\infty)$. Denn für $m \in \mathbb{N}$ sei

$$f_m(x) = \begin{cases} x^s\,, & a \leq x < a+m \\ 0\,, & a+m \leq x\,. \end{cases}$$

Als stetige Funktion auf dem beschränkten Intervall $[a,a+m)$ gehört $f_m|_{[a,a+m)}$ zum Raum $\mathcal{L}^+([a,a+m))$, und ist damit integrierbar. Hieraus folgt sofort, daß f_m über $[a,\infty)$ integrierbar ist mit

$$\int_a^\infty f_m dx = \int_a^{a+m} f_m dx = \mathcal{R}-\int_a^{a+m} f_m dx = \frac{1}{s+1}\Big[(a+m)^{s+1}-a^{s+1}\Big] \nearrow -\frac{a^{s+1}}{s+1}\,.$$

Da die Folge $\{f_m\}_{m=1}^{\infty}$ wachsend und die Integralfolge beschränkt ist, liefert der Satz von Beppo Levi, daß eine Nullmenge $N \subseteq [a,\infty)$ und eine Funktion $g \in \mathcal{L}([a,\infty))$ existieren mit $\lim_{m\to\infty} f_m(x) = g(x)$ für alle $x \in [a,\infty)\backslash N$. Weil aber $\lim_{m\to\infty} f_m(x) = f(x)$ ist für alle $x \in [a,\infty)$, folgt $g(x) = f(x)$ für alle $x \in [a,\infty)\backslash N$, also $f \in \mathcal{L}([a,\infty))$ und

$$\int_a^\infty x^s dx = \int_a^\infty g\,dx = \lim_{m\to\infty}\int_a^\infty f_m dx = -\frac{a^{s+1}}{s+1}\,.$$

Beispiel 2: Ebenso sieht man, daß für alle $a, b, c, s \in \mathbb{R}$ mit $a < c < b$ und $s > -1$ die durch

$$|x-c|_0^s = \begin{cases} |x-c|^s\,, & \text{für} \quad x \neq c \\ 0\,, & \text{für} \quad x = c \end{cases}$$

definierte Funktion $x \mapsto |x-c|_0^s : (a,b) \to \mathbb{R}$ integrierbar ist über (a,b) mit dem Integral

$$\begin{aligned} \int_a^b |x-c|_0^s dx &= \int_c^b (x-c)^s dx + \int_a^c (c-x)^s dx \\ &= \frac{1}{s+1}\,(x-c)^{s+1}\Big|_c^b - \frac{1}{s+1}\,(c-x)^{s+1}\Big|_a^c = \frac{(b-c)^{s+1} + (c-a)^{s+1}}{s+1}\,. \end{aligned}$$

Beispiel 3: Sei $\alpha > 2$. Für $m \in \mathbb{N}$ sei $f_m : (0,1) \to \mathbb{R}$ definiert durch

$$f_m(x) = \sum_{k=2}^{m} \sum_{\ell=1}^{k-1} \frac{1}{k^\alpha}\,|x - \frac{\ell}{k}|_0^{-1/2}\,.$$

Nach dem vorangehenden Beispiel ist die (natürlich fast überall mit $\frac{1}{\sqrt{x-\ell/k}}$ übereinstimmende) Funktion $|x-\ell/k|_0^{-1/2}$ für $1 \leq \ell \leq k-1$ integrierbar über dem Intervall $(0,1)$ mit

$$\int_0^1 |x-\ell/k|_0^{-1/2} dx = 2\Big((1-\ell/k)^{1/2} + (\ell/k)^{1/2}\Big) \leq 4\,.$$

Als endliche Summe integrierbarer Funktionen ist somit auch f_m integrierbar mit

$$\int_0^1 f_m dx = \sum_{k=2}^{m} \sum_{\ell=1}^{k-1} \frac{1}{k^\alpha} \int_0^1 |x-\ell/k|_0^{-1/2} dx \leq \sum_{k=2}^{m} \frac{4(k-1)}{k^\alpha} \leq 4 \sum_{k=2}^{\infty} \frac{1}{k^{\alpha-1}} < \infty\,.$$

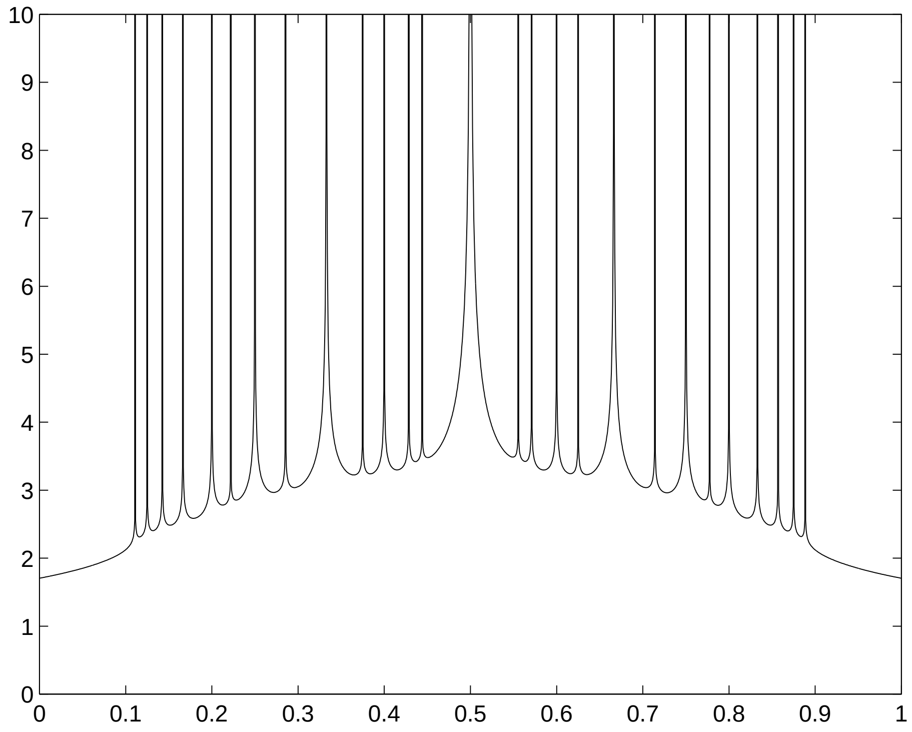
Zur wachsenden Folge $\{f_m\}_{m=1}^\infty$ von Funktionen $f_m \in \mathcal{L}((0,1))$ ist also die Integralfolge $\{\int_0^1 f_m dx\}_{m=1}^\infty$ beschränkt. Damit liefert der Satz von Beppo Levi, daß die Reihe

$$f(x) = \sum_{k=2}^{\infty} \sum_{\ell=1}^{k-1} \frac{1}{k^\alpha}\,|x - \frac{\ell}{k}\,|_0^{-1/2}$$

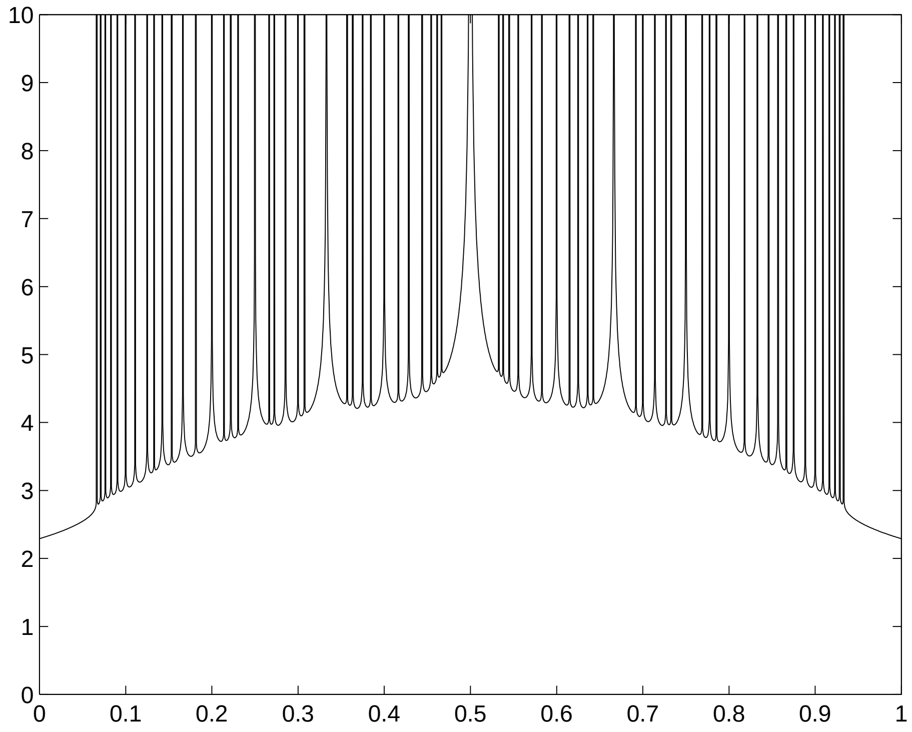
für fast alle $x \in (0,1)$ konvergiert, und daß die durch diese Reihe definierte Funktion f zu $\mathcal{L}((0,1))$ gehört mit

$$\int_0^1 f dx = \lim_{m\to\infty} \int_0^1 f_m dx \leq 4 \sum_{k=2}^{\infty} \frac{1}{k^{\alpha-1}}\,.$$

Jedoch besitzt f unendlich viele Singularitäten, die im Intervall $(0,1)$ „dicht“ liegen. Für $\alpha = 2.1$ sind die Graphen der Funktionen f_9 und f_{15} in der Abbildung auf der folgenden Seite dargestellt. Dieses Beispiel zeigt, daß es sehr ungewöhnliche Funktionen gibt, die Lebesgue–integrierbar sind.



Graph der Funktion f_9



Graph der Funktion f_{15}

Satz von Lebesgue über dominierte Konvergenz: Sei $A \subseteq \mathbb{R}^n$ und sei $\{f_m\}_{m=1}^{\infty}$ eine Folge von Funktionen $f_m \in \mathcal{L}(A)$, die fast überall gegen eine Funktion $f : A \to \mathbb{R}$

konvergiert. $g \in \mathcal{L}(A)$ sei eine Dominante dieser Folge, d.h. für alle $m \in \mathbb{N}$ gelte

$$|f_m| \leq g$$

fast überall in A. Dann ist $f \in \mathcal{L}(A)$ und

$$\lim_{m\to\infty} \int_A f_m dx = \int_A f dx\,.$$

Beweis: N sei die Nullmenge aller $x \in A$, für die $\{f_m(x)\}_{m=1}^\infty$ nicht gegen $f(x)$ konvergiert. Sei

$$g_m(x) = \begin{cases} \inf\{f_k(x) \mid k \geq m\}\,, & \text{für} \quad x \in A\backslash N \\ 0\,, & \text{für} \quad x \in N\,, \end{cases}$$

$$h_m(x) = \begin{cases} \sup\{f_k(x) \mid k \geq m\}\,, & \text{für} \quad x \in A\backslash N \\ 0\,, & \text{für} \quad x \in N\,. \end{cases}$$

Dann ist $\{g_m\}_{m=1}^\infty$ wachsend, $\{h_m\}_{m=1}^\infty$ fallend und $\lim_{m\to\infty} g_m(x) = \lim_{m\to\infty} h_m(x) = f(x)$ für alle $x \in A\backslash N$. Es soll nun der Satz von Beppo Levi auf die Folgen $\{g_m\}_{m=1}^\infty$ und $\{h_m\}_{m=1}^\infty$ angewandt werden. Dazu muß gezeigt werden, daß $g_m, h_m \in \mathcal{L}(A)$ ist, und daß die Integralfolgen $\{\int_A g_m dx\}_{m=1}^\infty$, $\{\int_A h_m dx\}_{m=1}^\infty$ beschränkt sind.

Zum Beweis sei

$$u_{mk}(x) = \min\{f_m(x), \ldots, f_{m+k}(x)\}\,.$$

Wegen $f_\ell \in \mathcal{L}(A)$ ist auch $u_{mk} \in \mathcal{L}(A)$. Die Funktionenfolge $\{u_{mk}\}_{k=1}^\infty$ ist fallend und konvergiert auf der Menge $A\backslash N$ gegen g_m. Schließlich ist auch

$$|u_{mk}(x)| \leq \max\{|f_m(x)|, \ldots, |f_{m+k}(x)|\} \leq g(x)$$

für fast alle x, also

$$|\int_A u_{mk} dx| \leq \int_A |u_{mk}| dx \leq \int_A g dx\,.$$

Somit sind die Voraussetzungen des Satzes von Beppo Levi erfüllt. Nach diesem Satz existiert eine Funktion $G_m \in \mathcal{L}(A)$ und eine Nullmenge N' mit $\lim_{k\to\infty} u_{mk}(x) = G_m(x)$ für alle $x \in A\backslash N'$. Da $\lim_{k\to\infty} u_{mk}(x) = g_m(x)$ gilt für alle $x \in A\backslash N$, stimmt $g_m(x)$ für alle x, die nicht zur Nullmenge $N \cup N'$ gehören, mit $G_m(x)$ überein, also ist auch $g_m \in \mathcal{L}(A)$. Ebenso folgt $h_m \in \mathcal{L}(A)$.

Daß die Folgen $\{\int_A g_m dx\}_{m=1}^\infty$ und $\{\int_A h_m dx\}_{m=1}^\infty$ beschränkt sind, folgt sofort aus

$$\int_A g_1 dx \leq \int_A g_m dx \leq \int_A h_m dx \leq \int_A h_1 dx\,.$$

Somit kann der Satz von Beppo Levi auf die wachsende Folge $\{g_m\}_{m=1}^{\infty}$ und auf die fallende Folge $\{h_m\}_{m=1}^{\infty}$ angewandt werden, und mit derselben Schlussweise wie eben erhält man für die gemeinsame Grenzfunktion f dieser Folgen, daß $f \in \mathcal{L}(A)$ und

$$\lim_{m\to\infty} \int_A g_m dx = \int_A f dx = \lim_{n\to\infty} \int_A h_m dx$$

gilt. Wegen $g_m \leq f_m \leq h_m$ fast überall folgt

$$\int_A g_m dx \leq \int_A f_m dx \leq \int_A h_m dx\,,$$

also

$$\lim_{m\to\infty} \int_A f_m dx = \int_A f dx\,.$$

■

4 c.) Parameterabhängige Integrale

Der Wert des Integrales

$$\int_A f(x,t)dx$$

ist vom Parameter t abhängig, über den nicht integriert wird. In diesem Abschnitt werden die Stetigkeits– und Differenzierbarkeitseigenschaften der durch dieses parameterabhängige Integral definierten Funktion

$$F(t) = \int_A f(x,t)dx$$

studiert.

Satz: Sei $A \subseteq \mathbb{R}^n$, sei $D \subseteq \mathbb{R}^p$ offen und

$$f : A \times D \to \mathbb{R}$$

bei jedem festen $t \in D$ Lebesgue–integrierbar bezüglich x sowie für fast alle $x \in A$ stetig bezüglich t. Wenn f eine Lebensgue–integrierbare Dominante $\Phi : A \to \mathbb{R}$ besitzt, d.h. wenn für alle $(x,t) \in A \times D$ die Ungleichung

$$|f(x,t)| \leq \Phi(x)$$

gilt, dann ist die durch

$$F(t) = \int_A f(x,t)dx$$

definierte Funktion $F : D \to \mathbb{R}$ stetig.

Beweis: Sei $\{t_k\}_{k=1}^{\infty}$ eine Folge mit $t_k \in D$ und mit $\lim_{k\to\infty} t_k = t \in D$. Sei $\varphi_k : A \to \mathbb{R}$ definiert durch

$$\varphi_k(x) = f(x, t_k)\,.$$

Dann konvergiert die Folge $\{\varphi_k\}_{k=1}^{\infty}$ punktweise gegen die durch $\varphi(x) = f(x,t)$ definierte Funktion $\varphi : A \to \mathbb{R}$, und hat die Dominante Φ. Also gilt nach dem Satz von Lebesgue

$$\lim_{k\to\infty} F(t_k) = \lim_{k\to\infty} \int_A f(x,t_k)dx = \int_A \lim_{k\to\infty} f(x,t_k)dx = F(t)\,.$$

Also ist F stetig. ∎

Als nächstes soll die Differenzierbarkeit untersucht werden.

Satz: Sei $A \subseteq \mathbb{R}^n$, sei $D \subseteq \mathbb{R}$ offen und sei $f : A \times D \to \mathbb{R}$ bei jedem festen $t \in D$ Lebesgue–integrierbar. Für fast alle $x \in A$ sei die Funktion $t \mapsto f(x,t)$ differenzierbar und ihre Ableitung $\frac{\partial f}{\partial t}(x,t)$ besitze eine Lebesgue–integrierbare Dominante Φ, d.h. es gelte

$$\Big|\frac{\partial f}{\partial t}(x,t)\Big| \le \Phi(x)$$

für alle $t \in D$ und fast alle $x \in A$. Dann ist für jedes $t \in D$ die Funktion

$$x \mapsto \frac{\partial f}{\partial t}(x,t)$$

Lebesgue–integrierbar, die durch

$$F(t) = \int_A f(x,t)dx$$

definierte Funktion F ist in D differenzierbar und hat die Ableitung

$$F'(t) = \int_A \frac{\partial f}{\partial t}(x,t)dx\,.$$

Beweis: Sei $\{t_k\}_{k=1}^{\infty}$ eine Folge mit $t_k \in D$ mit $\lim_{k\to\infty} t_k = t \in D$ und mit $t_k \neq t$ für alle $k \in \mathbb{N}$. Betrachte

$$\psi_k(x) = \frac{f(x,t_k) - f(x,t)}{t_k - t}\,.$$

ψ_k ist Lebesgue–integrierbar, und für fast alle $x \in A$ gilt $\lim_{k\to\infty}\psi_k(x) = \frac{\partial f}{\partial t}(x,t)$. Nun wird gezeigt, daß Φ Dominante ist für die Funktionen ψ_k. Aus dem Lebesgueschen Satz folgt dann $(x \mapsto \frac{\partial f}{\partial t}(x,t)) \in \mathcal{L}(A)$ und

$$\begin{aligned} F'(t) &= \lim_{k\to\infty} \int_A \frac{f(x,t_k) - f(x,t)}{t_k - t}dx = \lim_{k\to\infty}\int_A \psi_k(x)dx \\ &= \int_A \lim_{k\to\infty}\psi_k(x)dx = \int_A \frac{\partial f}{\partial t}(x,t)dx\,. \end{aligned}$$

Daß Φ Dominante ist, folgt aus dem Mittelwertsatz:

$$|\psi_k(x)| = \Big| \frac{f(x,t_k) - f(x,t)}{t_k - t} \Big| = \Big| \frac{\partial f}{\partial t}(x, \tau_k) \Big| \le \Phi(x)\,,$$

für geeignetes τ_k zwischen t und t_k. ■

Beispiel: Sei $A = (0,1)$, $D = (0,1)$, $f : (0,1) \times (0,1) \to \mathbb{R}$ definiert durch

$$f(x,t) = \frac{1}{\sqrt{x} + t}\,.$$

Dann ist $x \mapsto f(x,t)$ Lebesgue–integrierbar für jedes $t \in (0,1)$. Für die Ableitung gilt

$$\frac{\partial f}{\partial t} = \frac{-1}{(\sqrt{x} + t)^2}\,.$$

Für jedes $\varepsilon > 0$ gilt

$$\Big|\frac{\partial f}{\partial t}(x,t)\Big| = \frac{1}{(\sqrt{x} + t)^2} \le \frac{1}{(\sqrt{x} + \varepsilon)^2} = \Phi(x)$$

falls $(x,t) \in (0,1) \times (\varepsilon, 1)$ ist; Φ ist Lebesgue–integrierbar, also folgt für $t \in (\varepsilon, 1)$ mit $F(t) = \int_0^1 f(x,t)dx$:

$$\frac{dF}{dt}(t) = \int_0^1 \frac{\partial f}{\partial t}(x,t)dx = -\int_0^1 \frac{1}{(\sqrt{x} + t)^2}\,dx\,.$$

Da dies für alle $\varepsilon > 0$ gilt, ist F differenzierbar auf $(0,1)$.

4 d.) Satz von Fubini

Der Satz von Fubini erlaubt die Berechnung n–dimensionaler Integrale auf die iterierte Berechnung eindimensionaler Integrale zurückzuführen. Dies ist von großer Bedeutung, weil man zur Berechnung eindimensionaler Integrale viele Regeln und Hilfsmittel zur Verfügung hat.

In diesem Abschnitt sei $n \in \mathbb{N}$ mit $n \ge 2$. Elemente von $\mathbb{R}^{n-1}$ werden mit x' bezeichnet, Elemente von $\mathbb{R}^n$ mit x. Für $A \subseteq \mathbb{R}^n$ und $z \in \mathbb{R}$ sei

$$A(z) = \{x' \in \mathbb{R}^{n-1} \;\Big|\; (x', z) \in A\}$$

ein $(n-1)$–dimensionaler Querschnitt von A.

Sei χ_Q die charakteristische Funktion zum n–dimensionalen, beschränkten Quader $Q =$

$\Pi_{i=1}^n I_i$. Für $z \in I_n$ ist dann $x' \mapsto \chi_Q(x', z)$ die charakteristische Funktion zum $(n-1)$–dimensionalen Quader $Q' = \Pi_{i=1}^{n-1} I_i$, also gilt $\int_{\mathbb{R}^{n-1}} \chi_Q(x', z)\, dx' = |Q'|$, während $\chi_Q(x'z) = 0$ ist für $z \notin I_N$. Somit ist

$$z \mapsto \int_{\mathbb{R}^{n-1}} \chi_Q(x', z)\, dx' = |Q'|\; \chi_{I_n}(z) : \mathbb{R} \to \mathbb{R}$$

eine Treppenfunktion und

$$\int_{-\infty}^{\infty} \int_{\mathbb{R}^{n-1}} \chi_Q(x', z)\, dx' \;=\; \int_{-\infty}^{\infty} |Q'|\; \chi_{I_n}(z)\, dz = |Q'|\, |I_n| = |Q| = \int_{\mathbb{R}^n} \chi_Q(x)\, dx\, .$$

Wegen der Linearität des Integrals ist damit für jede Treppenfunktion

$$\varphi(x) = \sum_{i=1}^{m} c_i \chi_{Q_i}(x)$$

mit n–dimensionalen Quadern $Q_1, \ldots, Q_m$ auch die durch

$$\psi(z) = \int_{\mathbb{R}^{n-1}} \varphi(x', z)\, dx' = \sum_{i=1}^{m} c_i \int_{\mathbb{R}^{n-1}} \chi_{Q_i}(x', z)\, dx' = \sum_{i=1}^{m} c_i |Q_i'|\; \chi_{I_n^{(i)}}(z)$$

definierte Funktion $\psi : \mathbb{R} \to \mathbb{R}$ eine Treppenfunktion mit

$$\begin{aligned} \int_{-\infty}^{\infty} \int_{\mathbb{R}^{n-1}} \varphi(x', z)\, dx' dz \;&=\; \int_{-\infty}^{\infty} \psi(z)\, dz = \sum_{i=1}^{m} c_i \int_{-\infty}^{\infty} \int_{\mathbb{R}^{n-1}} \chi_{Q_i}(x', z)\, dx' dz \\ &=\; \sum_{i=1}^{m} c_i \int_{\mathbb{R}^n} \chi_{Q_i}(x)\, dx = \int_{\mathbb{R}^n} \varphi(x)\, dx\, . \end{aligned}$$

Die Berechnung des n–dimensionalen Integrals von φ kann also auf die Berechnung eines $(n-1)$–dimensionalen und eines eindimensionalen Integrals reduziert werden, und durch n–fache Wiederholung dieses Schlusses sogar auf die iterierte Berechnung von n eindimensionalen Integralen zurückgeführt werden. Der Satz von Fubini sagt aus, daß dies für jede Lebesgue–integrierbare Funktion möglich ist.
Zum Beweis des Satzes von Fubini benötigen wir zwei Hilfssätze:

Hilfssatz: Sei $\{\varphi_m\}_{m=1}^{\infty}$ eine wachsende Folge von Treppenfunktionen $\varphi_m \in T(\mathbb{R}^n)$ mit beschränkter Integralfolge, und sei N die Menge aller $x \in \mathbb{R}^n$, für die die Folge $\{\varphi_m(x)\}_{m=1}^{\infty}$ divergiert. Sei $f : \mathbb{R}^n \to \mathbb{R}$ definiert durch

$$f(x) = \begin{cases} \lim\limits_{m\to\infty} \varphi_m(x)\, , & x \in \mathbb{R}^n \backslash N \\ 0\, , & x \in N\, . \end{cases}$$

Dann gilt:

(i) N ist eine Nullmenge im $\mathbb{R}^n$, und es gilt $f \in \mathcal{L}^+(\mathbb{R}^n)$ mit

$$\int_{\mathbb{R}^n} f(x)dx = \lim_{m\to\infty} \int_{\mathbb{R}^n} \varphi_m(x)dx\,.$$

(ii) Es gibt eine Nullmenge $N_1 \subseteq \mathbb{R}$, so daß für alle $z \in \mathbb{R}\backslash N_1$ die Menge $N(z)$ eine Nullmenge im $\mathbb{R}^{n-1}$ ist, die Funktion $x' \mapsto f(x',z) : \mathbb{R}^{n-1} \to \mathbb{R}$ zu $\mathcal{L}(\mathbb{R}^{n-1})$ gehört und die Funktion $F : \mathbb{R} \to \mathbb{R}$,

$$F(z) = \begin{cases} \displaystyle\int_{\mathbb{R}^{n-1}} f(x',z)dx'\,, & z \in \mathbb{R}\backslash N_1 \\ \hfill 0\,, & z \in N_1\,, \end{cases}$$

in $\mathcal{L}(\mathbb{R})$ liegt mit

$$\int_{\mathbb{R}^n} f(x)dx = \int_{-\infty}^{\infty} F(z)dz = \int_{-\infty}^{\infty}\int_{\mathbb{R}^{n-1}} f(x',z)dx'dz\,.$$

Beweis: Die Aussage (i) ist der Satz von Beppo Levi und wurde schon bewiesen. Zum Beweis von (ii) sei $\psi_m : \mathbb{R} \to \mathbb{R}$ definiert durch

$$\psi_m(z) = \int_{\mathbb{R}^{n-1}} \varphi_m(x',z)dx'\,.$$

Dann ist $\{\psi_m\}_{m=1}^{\infty}$ eine wachsende Folge von Treppenfunktionen mit

$$\int_{-\infty}^{\infty} \psi_m(z)dz = \int_{-\infty}^{\infty}\int_{\mathbb{R}^{n-1}} \varphi_m(x',z)dx'dz = \int_{\mathbb{R}^n} \varphi_m(x)dx \le \int_{\mathbb{R}^n} f dx\,.$$

Nach dem Satz von Beppo Levi gibt es also eine Nullmenge $N_1 \subseteq \mathbb{R}$, so daß $\{\psi_m(z)\}_{m=1}^{\infty}$ für alle $z \in \mathbb{R}\backslash N_1$ konvergiert. Für die durch

$$F(z) = \begin{cases} \displaystyle\lim_{m\to\infty} \psi_m(z)\,, & z \in \mathbb{R}\backslash N_1 \\ \hfill 0\,, & z \in N_1 \end{cases}$$

definierte Funktion gilt $F \in \mathcal{L}(\mathbb{R})$ und

$$\begin{aligned} \int_{-\infty}^{\infty} F(z)dz &= \lim_{n\to\infty}\int_{-\infty}^{\infty} \psi_m(z)dz = \lim_{m\to\infty}\int_{-\infty}^{\infty}\int_{\mathbb{R}^{n-1}} \varphi_m(x',z)dx'dz \\ &= \lim_{m\to\infty}\int_{\mathbb{R}^n} \varphi_m(x)dx = \int_{\mathbb{R}^n} f(x)dx\,. \end{aligned} \tag{$*$}$$

Für $z \in \mathbb{R}\backslash N_1$ ist die Treppenfunktionsfolge $\{x' \mapsto \varphi_m(x',z)\}_{m=1}^{\infty}$ wachsend und hat wegen

$$\lim_{m\to\infty}\int_{\mathbb{R}^{n-1}} \varphi_m(x',z)dx' = \lim_{m\to\infty} \psi_m(z) = F(z) < \infty$$

eine konvergente Integralfolge. Wieder nach dem Satz von Beppo Levi gibt es daher eine Nullmenge $N_z \subseteq \mathbb{R}^{n-1}$, so daß die Folge $\{\varphi_m(x', z)\}_{m=1}^{\infty}$ für alle $x' \in \mathbb{R}^{n-1} \backslash N_z$ konvergiert. Nach Definition der Menge N und der Funktion f bedeutet dies für $x' \notin N_z$, daß $(x', z) \notin N$, folglich $x' \notin N(z)$ und somit $N(z) \subseteq N_z$ gilt, und daß

$$\lim_{m \to \infty} \varphi_m(x', z) = f(x', z)$$

ist. Damit liefert der Satz von Beppo Levi außerdem, daß $(x' \mapsto f(x', z)) \in \mathcal{L}(\mathbb{R}^{n-1})$ ist für $z \in \mathbb{R} \backslash N_1$ mit

$$\int_{\mathbb{R}^{n-1}} f(x'z)\, dx' = \int_{\mathbb{R}^{n-1}} \lim_{m \to \infty} \varphi_m(x', z)\, dx' = \lim_{m \to \infty} \int_{\mathbb{R}^{n-1}} \varphi_m(x', z)\, dx' = F(z)\, .$$

Als Teilmenge der Nullmenge N_z ist also auch $N(z)$ eine Nullmenge in $\mathbb{R}^{n-1}$. Wegen $(*)$ folgt schließlich auch

$$\int_{-\infty}^{\infty} \int_{\mathbb{R}^{n-1}} f(x', z)\, dx' dz = \int_{-\infty}^{\infty} F(z)\, dz = \int_{\mathbb{R}^n} f(x)\, dx\, .$$

■

Mit diesem Ergebnis kann der zweite Hilfssatz bewiesen werden.:

Hilfssatz: Sei $N \subseteq \mathbb{R}^n$ eine Nullmenge. Dann ist $N(z)$ für fast alle $z \in \mathbb{R}$ eine Nullmenge in $\mathbb{R}^{n-1}$.

Beweis: Sei $\varepsilon > 0$ und seien $Q_1, Q_2, \ldots$ abzählbar viele Quader mit $N \subseteq \bigcup_{\ell=1}^{\infty} Q_\ell$ und mit $\sum_{\ell=1}^{\infty} |Q_\ell| < \varepsilon$. Für jedes $z \in \mathbb{R}$ ist dann $Q_\ell(z)$ ein $(n-1)$–dimensionaler Quader und es gilt

$$N(z) \subseteq \Big[\bigcup_{\ell=1}^{\infty} Q_\ell\Big](z) = \bigcup_{\ell=1}^{\infty} Q_\ell(z)\, .$$

Mit den charakteristischen Funktionen χ_{Q_ℓ} wird durch

$$\varphi_m(x) = \sum_{\ell=1}^{m} \chi_{Q_\ell}(x)$$

eine Treppenfunktion auf $\mathbb{R}^n$ definiert. Die Folge $\{\varphi_m\}_{m=1}^{\infty}$ ist wachsend und die Integralfolge ist beschränkt wegen

$$\int_{\mathbb{R}^n} \varphi_m(x) dx = \sum_{\ell=1}^{m} \int_{\mathbb{R}^n} \chi_{Q_\ell}(x) dx = \sum_{\ell=1}^{m} |Q_\ell| < \varepsilon\, .$$

Sei $K \subseteq \mathbb{R}^n$ die Menge aller x, für die $\{\varphi_m(x)\}_{m=1}^{\infty}$ divergiert, und sei

$$f_\varepsilon(x) = \begin{cases} \lim\limits_{m \to \infty} \varphi_m(x)\, , & x \in \mathbb{R}^n \backslash K \\ 0\, , & x \in K\, . \end{cases}$$

Nach dem vorangehenden Hilfssatz ist K eine Nullmenge im $\mathbb{R}^n$, für fast alle $z \in \mathbb{R}^n$ ist $K(z)$ eine Nullmenge im $\mathbb{R}^{n-1}$, die Funktion f_ε gehört zu $\mathcal{L}(\mathbb{R}^n)$, für fast alle z gehört $x' \mapsto f_\varepsilon(x', z)$ zu $\mathcal{L}(\mathbb{R}^{n-1})$ und die durch

$$g_\varepsilon(z) := \int_{\mathbb{R}^{n-1}} f_\varepsilon(x', z)dx' = \lim_{m\to\infty} \int_{\mathbb{R}^{n-1}} \varphi_m(x', z)dx'$$

definierte Funktion gehört zu $\mathcal{L}(\mathbb{R})$ mit

$$\int_{\mathbb{R}^n} f_\varepsilon(x)dx = \int_{-\infty}^{\infty} g_\varepsilon(z)\, dz = \lim_{m\to\infty} \int_{\mathbb{R}^n} \varphi_m(x)dx = \sum_{\ell=1}^{\infty} |Q_\ell| < \varepsilon\,.$$

Wegen

$$\int_{\mathbb{R}^{n-1}} \varphi_m(x', z)dx' = \sum_{\ell=1}^{m} \int_{\mathbb{R}^{n-1}} \chi_{Q_\ell(z)}(x')dx' = \sum_{\ell=1}^{m} |Q_\ell(z)|$$

ist also $\{Q_\ell(z)\}_{\ell=1}^{\infty}$ für fast alle $z \in \mathbb{R}$ eine Überdeckung von $N(z)$ mit

$$\sum_{\ell=1}^{\infty} |Q_\ell(z)| = \lim_{m\to\infty} \int_{\mathbb{R}^{n-1}} \varphi_m(x', z)dx' = g_\varepsilon(z) < \infty\,.$$

Setzt man nun

$$h_k(z) = \min\{g_1(z), \dots, g_{1/k}(z)\},$$

dann genügt es zum Beweis des Hilfssatzes zu zeigen, daß für alle z außerhalb einer Nullmenge $K_1 \subseteq \mathbb{R}$

$$\lim_{k\to\infty} h_k(z) = 0 \tag{$*$}$$

gilt, weil dann für alle $z \in \mathbb{R}\backslash K_1$ und alle $\delta > 0$ Zahlen $k, l \in \mathbb{N}$ mit $1 \leq \ell \leq k$ existieren, so daß

$$\sum_{\ell=1}^{\infty} |Q_\ell(z)| = g_{1/\ell}(z) = h_k(z) < \delta$$

gilt. Also ist $N(z)$ für $z \in \mathbb{R}\backslash K_1$ eine Nullmenge.

Zum Beweis von $(*)$ beachte man, daß $h_k \in \mathcal{L}(\mathbb{R})$ ist und daß $h_k(z) \geq 0$ gilt für alle $z \in \mathbb{R}$, weil $g_\varepsilon \in \mathcal{L}(\mathbb{R})$ ist für alle $\varepsilon > 0$ und weil $g_\varepsilon(z) \geq 0$ gilt. Nach Definition ist die Folge $\{h_k\}_{k=1}^{\infty}$ fallend, also existiert $\lim_{k\to\infty} h_k(z) \geq 0$ für alle $z \in \mathbb{R}$, und weil

$$0 \leq \int_{-\infty}^{\infty} h_k(z)dz \leq \int_{-\infty}^{\infty} g_{1/k}(z)dz = \int_{\mathbb{R}^n} f_{1/k}(x)dx < \frac{1}{k}$$

gilt, folgt aus dem Satz von Beppo Levi, angewandt auf fallende Folgen, daß

$$\int_{-\infty}^{\infty} |\lim_{k\to\infty} h_k(z)|dz = \int_{-\infty}^{\infty} \lim_{k\to\infty} h_k(z)dz = \lim_{k\to\infty} \int_{-\infty}^{\infty} h_k(z)dz = 0$$

hieraus. Hieraus folgt $(*)$. ■

Satz von Fubini: Seien $A \subseteq \mathbb{R}^n$ und $f \in \mathcal{L}(A)$. Dann exisitert eine Nullmenge $L \subseteq \mathbb{R}$, so daß die Funktion

$$\Big(x' \mapsto f(x', z)\Big) : A(z) \to \mathbb{R}$$

für alle $z \in \mathbb{R} \backslash L$ zu $\mathcal{L}(A(z))$ gehört, und die durch

$$F(z) = \begin{cases} \displaystyle\int_{A(z)} f(x', z) dx', & z \in \mathbb{R} \backslash L \\ 0, & z \in L \end{cases}$$

erklärte Funktion $F : \mathbb{R} \to \mathbb{R}$ zu $\mathcal{L}(\mathbb{R})$ gehört mit

$$\int_A f(x) dx = \int_{-\infty}^{\infty} F(z) dz = \int_{-\infty}^{\infty} \int_{A(z)} f(x', z) dx' dz .$$

Beweis: Es genügt, diesen Satz für $A = \mathbb{R}^n$ zu beweisen, weil nach Definition eine Funktion zu $\mathcal{L}(A)$ gehört, genau dann wenn die durch 0 von A auf $\mathbb{R}^n$ fortgesetzte Funktion zu $\mathcal{L}(\mathbb{R}^n)$ gehört. Weiterhin genügt es, diesen Satz für $f \in \mathcal{L}^+(\mathbb{R}^n)$ zu beweisen, weil jede Funktion aus $\mathcal{L}(\mathbb{R}^n)$ als Differenz zweier Funktionen aus $\mathcal{L}^+(\mathbb{R}^n)$ geschrieben werden kann.

Sei also $f \in \mathcal{L}^+(\mathbb{R}^n)$. Dann gibt es eine wachsende Folge $\{\varphi_m\}_{m=1}^{\infty}$ von Treppenfunktionen auf $\mathbb{R}^n$ und eine Nullmenge N im $\mathbb{R}^n$ mit $\lim_{m\to\infty} \varphi_m(x) = f(x)$ für alle $x \in \mathbb{R}^n \backslash N$ und mit

$$\lim_{m\to\infty} \int_{\mathbb{R}^n} \varphi_m(x) dx = \int_{\mathbb{R}^n} f(x) dx .$$

Sei K die Menge aller $x \in \mathbb{R}^n$, für die $\{\varphi_m(x)\}_{m=1}^{\infty}$ divergiert und sei

$$\tilde{f}(x) = \begin{cases} \lim\limits_{m\to\infty} \varphi_m(x), & x \in \mathbb{R}^n \backslash K \\ 0, & x \in K . \end{cases}$$

K ist eine Teilmenge der Nullmenge N, also ist K selbst eine Nullmenge von $\mathbb{R}^n$, und $\tilde{f}$ stimmt außerhalb von N mit f überein. Nach dem zweiten Hilfssatz gibt es eine Nullmenge K_1 von $\mathbb{R}$, so daß $N(z)$ eine Nullmenge von $\mathbb{R}^{n-1}$ ist für alle $z \in \mathbb{R} \backslash K_1$. Für alle diese z stimmt also die Funktion $x' \mapsto \tilde{f}(x', z)$ fast überall in $\mathbb{R}^{n-1}$ mit der Funktion $x' \mapsto f(x', z)$ überein. Nach dem ersten Hilfssatz gehört $x' \mapsto \tilde{f}(x', z)$ zu $\mathcal{L}(\mathbb{R}^{n-1})$ für alle z außerhalb einer Nullmenge N_1, also gehört $x' \mapsto f(x', z)$ zu $\mathcal{L}(\mathbb{R}^{n-1})$ mit

$$\int_{\mathbb{R}^{n-1}} f(x', z) dx' = \int_{\mathbb{R}^{n-1}} \tilde{f}(x', z) dx'$$

für alle z außerhalb der Nullmenge $L = K_1 \cup N_1$ von $\mathbb{R}$. Die übrigen Behauptungen des Satzes von Fubini folgen nun unmittelbar aus dem ersten Hilfssatz. ■

Beispiel: Seien $a, b, c > 0$,

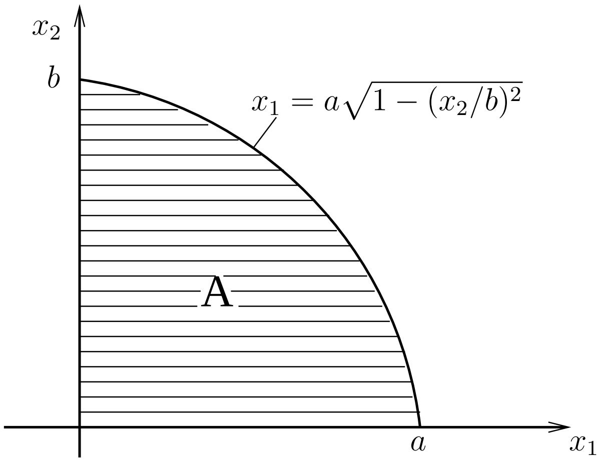
$$A = \{x = (x_1, x_2) \in \mathbb{R}^2 \;\Big|\; \frac{x_1^2}{a^2} + \frac{x_2^2}{b^2} \le 1,\ x_1 \ge 0,\ x_2 \ge 0\}$$

und

$$f(x_1, x_2) = c\sqrt{1 - \frac{x_1^2}{a^2} - \frac{x_2^2}{b^2}}\,.$$

Berechne

$$\int_A f(x)\, dx = \int_A c\sqrt{1 - \frac{x_1^2}{a^2} - \frac{x_2^2}{b^2}}\, d(x_1, x_2)\,.$$



Mit der Fortsetzung

$$\tilde{f}(x) = \begin{cases} f(x)\,, & x \in A \\ 0\,, & x \in \mathbb{R}^2 \setminus A\,, \end{cases}$$

gilt nach dem Satz von Fubini

$$\begin{aligned}
\int_A f(x)\, dx &= \int_{\mathbb{R}^2} \tilde{f}(x)\, dx = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \tilde{f}(x_1, x_2)\, dx_1\, dx_2 \\
&= \int_0^b \int_0^{a\sqrt{1-(x_2/b)^2}} c\sqrt{1 - \frac{x_1^2}{a^2} - \frac{x_2^2}{b^2}}\, dx_1\, dx_2 \\
&= \int_0^b \int_0^{a\sqrt{1-(x_2/b)^2}} c\sqrt{1 - \frac{x_2^2}{b^2}} \sqrt{1 - \frac{x_1^2}{a^2(1 - x_2^2/b^2)}}\, dx_1\, dx_2\,.
\end{aligned}$$

Mit der Substitution $y(x_1) = \frac{x_1}{a\sqrt{1-x_2^2/b^2}}$ ergibt sich

$$
\begin{aligned}
\int_A f(x)\,dx &= \int_0^b \int_0^1 c\sqrt{1-\frac{x_2^2}{b^2}}\sqrt{1-y^2}\,\frac{dx_1(y)}{dy}\,dy\,dx_2 \\
&= \int_0^b ac\Big(1-\frac{x_2^2}{b^2}\Big)\,dx_2 \int_0^1 \sqrt{1-y^2}\,dy \\
&= ac\frac{\pi}{4}\Big[x_2-\frac{x_2^3}{3b^2}\Big]_0^b = \frac{\pi}{6}abc\,.
\end{aligned}
$$

Dies ist das Volumen des achten Teiles eines Ellipsoides mit den Hauptachsen a, b und c.

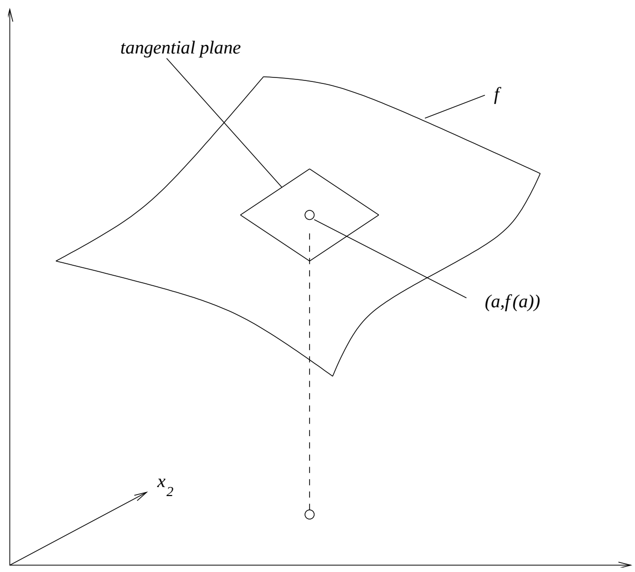
5 Differentialrechnung im $\mathbb{R}^n$

5 a.) Definition der Ableitung

Es soll der Begriff der Ableitung von reellen Funktionen auf Funktionen von n Veränderlichen verallgemeinert werden. Die Grundidee dabei ist folgende: Seien $U \subseteq \mathbb{R}^n$ offen, $f : U \to \mathbb{R}^m$ und $a \in U$. Unter der Ableitung von f an der Stelle a versteht man diejenige lineare Abbildung, die f in einer Umgebung des Punktes a „am besten approximiert“.
Ich betrachte als Beispiel Abbildungen der Form $f : \mathbb{R}^2 \to \mathbb{R}$. Der Graph dieser Abbildung ist eine Fläche im $\mathbb{R}^3$. Der Graph einer linearen Abbildung $T : \mathbb{R}^2 \to \mathbb{R}$ ist eine durch den Nullpunkt gehende Ebene. Man bestimme nun diejenige lineare Abbildung T, deren Graph parallel ist zu der Tangentialebene an den Graphen von f im Punkt $(a, f(a))$. Dann heißt T „Ableitung von f im Punkt a “, und die Funktion

$$x \mapsto f(a) + T(x - a)$$

approximiert f in einer Umgebung von a.



Diese auf der Anschauung beruhende Idee wird in der folgenden Definition mathematisch genau gefasst:

Definition: Sei $U \subseteq \mathbb{R}^n$ offen. $f : U \to \mathbb{R}^m$ heißt differenzierbar an der Stelle $a \in U$, wenn es eine lineare Abbildung $T : \mathbb{R}^n \to \mathbb{R}^m$ und eine an der Stelle a stetige Abbildung $r : U \to \mathbb{R}^m$ mit $r(a) = 0$ gibt derart, daß für alle $x \in U$ gilt

$$f(x) = f(a) + T(x - a) + r(x)\|x - a\| .$$

Um zu prüfen, ob f an der Stelle $a \in D$ differenzierbar ist muß man zuerst eine geeignete lineare Abbildung T finden und dann prüfen, ob für

$$r(x) = \frac{f(x) - f(a) - T(x-a)}{\|x-a\|}$$

$\lim_{x \to a} r(x) = 0$ gilt. Ich werde später angeben, wie man T findet. Es kann höchstens ein solches T geben. Denn es gilt:

Lemma: T ist eindeutig bestimmt.

Beweis: Seien T_1, T_2 lineare Abbildungen mit

$$\begin{aligned} f(x) &= f(a) + T_1(x-a) + r_1(x)\|x-a\| \\ f(x) &= f(a) + T_2(x-a) + r_2(x)\|x-a\|\,. \end{aligned} \quad , \quad \lim_{x \to a} r_1(x) = \lim_{x \to a} r_2(x) = 0$$

Dann folgt

$$(T_1 - T_2)(x-a) = \Big(r_2(x) - r_1(x)\Big)\|x-a\|\,.$$

Sei $h \in \mathbb{R}^n$. Für alle hinreichend kleinen $t > 0$ gilt dann $x = a + th\, eU$, und es folgt

$$(T_1 - T_2)(th) = t(T_1 - T_2)(h) = \Big(r_2(a+th) - r_1(a+th)\Big)\|th\|\,,$$

also

$$(T_1 - T_2)(h) = \lim_{t \to 0} (T_1 - T_2)(h) = \lim_{t \to 0} \Big(r_2(a+th) - r_1(a+th)\Big)\|h\| = 0\,,$$

also $T_1 = T_2$, da h beliebig gewählt war. ■

Definition: Sei $U \subseteq \mathbb{R}^n$ offen und $f : U \to \mathbb{R}^m$ sei in $a \in U$ differenzierbar. Dann heißt die eindeutig bestimmte lineare Abbildung $T : \mathbb{R}^n \to \mathbb{R}^m$ mit

$$f(x) = f(a) + T(x-a) + r(x)\|x-a\|\,, \;\; \lim_{x \to a} r(x) = 0\,,$$

die Ableitung von f an der Stelle a. Bezeichnung: $T = f'(a)$.

Bei linearen Abbildungen läßt man häufig die Klammern um das Argument weg und schreibt $T(h) = Th = f'(a)h$.

Ist f reellwertig, dann ist $f'(a)$ eine lineare Abbildung $f'(a) : \mathbb{R}^m \to \mathbb{R}$. Solche lineare Abbildungen nennt man auch Linearformen. In diesem Fall bezeichnet man $f'(a)$ auch als Gradienten von f und schreibt $\text{grad} f(a) := f'(a)$. Aus der linearen Algebra weiß man, daß jede Linearform auf $\mathbb{R}^n$ mit Hilfe des Skalarproduktes in eindeutiger Weise durch einen Vektor im $\mathbb{R}^n$ dargestellt werden kann: Es gibt ein eindeutig bestimmtes $y \in \mathbb{R}^n$, so daß für alle $h \in \mathbb{R}^n$ gilt

$$\mathrm{grad} f(a)h = y \cdot h \,.$$

Oft wird auch der Vektor y mit $\mathrm{grad} f(a)$ bezeichnet. Man muß sich aber im klaren darüber sein, daß dies unpräzise ist, weil man für zwei verschiedene Dinge dieselbe Bezeichnung benutzt. Der Vektor $\mathrm{grad} f(a)$ zeigt in die Richtung des größten Zuwachses der Abbildung $f'(a)$, und damit auch in die Richtung des größten Zuwachses der Abbildung f an der Stelle a, weil $f'(a)$ die Abbildung f in einer Umgebung von a approximiert.
Die Tangentialhyperebene der Abbildung $f : U \to \mathbb{R}$ im Punkt $a \in U$ wird definiert durch

$$\Big\{ (x, z) \in \mathbb{R}^{n+1} \,\Big|\, z = f(a) + \Big[f'(a) \Big] (x - a) \,, \;\; x \in \mathbb{R}^n \Big\} \,.$$

Der Vektor $(-\mathrm{grad} f(a), 1) \in \mathbb{R}^{n+1}$ steht senkrecht auf dieser Hyperebene. Denn für zwei Vektoren (x_1, z_1) und (x_2, z_2) aus dieser Hyperebene gilt

$$\begin{aligned} (x_2, z_2) - (x_1, z_1) &= \Big(x_2 - x_1, \Big[f'(a) \Big] (x_2 - a) - \Big[f'(a) \Big] (x_1 - a) \Big) \\ &= \Big(x_2 - x_1, \Big[f'(a) \Big] (x_2 - x_1) \Big) \,, \end{aligned}$$

also

$$\begin{aligned} &\Big(-\mathrm{grad} f'(a), 1 \Big) \cdot \Big[(x_2, z_2) - (x_1, z_1) \Big] \\ &= \Big(-\mathrm{grad} f'(a) \Big) \cdot (x_2 - x_1) + \Big[f'(a) \Big] (x_2 - x_1) \\ &= -\Big(\mathrm{grad} f'(a) \Big) \cdot (x_2 - x_1) + \Big(\mathrm{grad} f'(a) \Big) \cdot (x_2 - x_1) = 0 \,. \end{aligned}$$

Ist insbesondere $U \subseteq \mathbb{R}$ und $f : U \to \mathbb{R}$ eine reelle Funktion, dann ist die lineare Abbildung $T = f'(a) : \mathbb{R} \to \mathbb{R}$ gegeben durch

$$Th = \frac{df}{dx}(a)h, \;\; h \in \mathbb{R} \,,$$

mit der klassischen Ableitung $\frac{df}{dx}(a) \in \mathbb{R}$ von f an der Stelle a.

Es gilt:

Lemma: Die Abbildung $f : U \to \mathbb{R}^m$, $U \subseteq \mathbb{R}^n$, ist differenzierbar in $a \in U$, genau dann wenn jede der Komponentenfunktionen $f_1, \ldots, f_m : U \to \mathbb{R}$ differenzierbar ist in $a \in U$. Es gilt dann

$$f_j'(a) = \Big(f'(a) \Big)_j \,, \quad j = 1, \ldots, m \,.$$

Beweis: Wenn $f'(a)$ existiert, ist $(f'(a))_j : \mathbb{R}^n \to \mathbb{R}$ linear, und es gilt

$$\lim_{\substack{h\to 0\\ h\neq 0}} \frac{f_j(h+a) - f_j(a) - (f'(a))_j h}{\|h\|} = 0\,,$$

also ist $(f'(a))_j = f_j'(a)$. Ist umgekehrt $f_j'(a)$ die Ableitung von f_j für $j = 1, \ldots, m$, dann wird durch

$$Th = \begin{pmatrix} f_1'(a)h \\ \vdots \\ f_m'(a)h \end{pmatrix} : \mathbb{R}^n \to \mathbb{R}^m$$

eine lineare Abbildung definiert für die gilt

$$\lim_{\substack{h\to 0\\ h\neq 0}} \frac{f(a+h) - f(a) - Th}{\|h\|} = 0\,,$$

also ist $T = f'(a)$. ∎

Um $f'(a)v$ für $v \in \mathbb{R}^n$ zu bestimmen, setze man $x = a + tv$ mir $t \in \mathbb{R}\,,\ t \neq 0$. Es gilt dann

$$f(a+tv) = f(a) + f'(a)(tv) + r(tv+a)|t|\,\|v\|\,,$$

also

$$f'(a)v = \frac{f(a+tv) - f(a)}{t} - r(tv+a)\,\frac{|t|}{t}\,\|v\|\,,$$

somit

$$f'(a)v = \lim_{\substack{t\to 0\\ t\neq 0}} f'(a)v = \lim_{\substack{t\to 0\\ t\neq 0}} \frac{f(a+tv) - f(a)}{t}\,.$$

Der rechtsstehende Grenzwert heißt Richtungsableitung von f an der Stelle a in Richtung von $v \in \mathbb{R}^n$. Für die Richtungsableitung benutze ich die Bezeichnung

$$D_v f(a) := \lim_{\substack{t\to 0\\ t\neq 0}} \frac{f(a+tv) - f(a)}{t}\,.$$

Zur Bestimmung der linearen Abbildung $f'(a)$ genügt es, die Richtungsableitungen $D_{v_i} f(a)$ für eine Basis $v_1, \ldots, v_n$ von $\mathbb{R}^n$ zu berechnen, weil man jedes $v \in \mathbb{R}^n$ als Linearkombination $v = \sum_{i=1}^n \alpha_i v_i$ schreiben kann mit $\alpha_i \in \mathbb{R}^n$, also

$$f'(a)v = f'(a)(\sum_{i=1}^n \alpha_i v_i) = \sum_{i=1}^n \alpha_i f'(a) v_i.$$

Es ist naheliegend, als Basis die Standardbasis $e_1, \ldots, e_n$ zu wählen. Die dabei benötigte Richtungsableitung $D_{e_i} f(a)$ nennt man i–te partielle Ableitung. Partielle Ableitungen bezeichnet man durch

$$\frac{\partial f}{\partial x_i}\,,\ \ D_i f\,,\ \ f_{x_i}\,,\ \ f_{x_i}'\,,$$

manchmal auch durch $f_{|i}$ oder f_i . Hierbei können aber Verwechslungen auftreten. Es gilt:

$$\frac{\partial f}{\partial x_i}(a) = \lim_{\substack{x_i \to a_i \\ x_i \neq a_i}} \frac{f(a_1, \ldots, x_i, \ldots, a_n) - f(a_1, \ldots, a_i, \ldots, a_n)}{x_i - a_i},$$

d.h.

$$\frac{\partial f_j}{\partial x_i}(a) \quad = \quad \lim_{\substack{x_i \to a_i \\ x_i \neq a_i}} \frac{f_j(a_1, \ldots, x_i, \ldots, a_n) - f_j(a_1, \ldots, a_i, \ldots, a_n)}{x_i - a_i};$$

$$i = 1, \ldots, n; \quad j = 1, \ldots, m$$

Zur Bestimmung der partiellen Ableitungen von f an der Stelle a genügt somit die Differentialrechnung einer reellen Variablen.

Sei $U \subseteq \mathbb{R}^n$ eine offene Menge und sei $f : U \to \mathbb{R}^m$ differenzierbar an der Stelle $a \in U$. Zur Bestimmung von $f'(a)$ geht man nun folgendermaßen vor: Weil f im Punkt a differenzierbar ist, existieren alle partiellen Ableitungen $D_i f(a) = \frac{\partial f}{\partial x_i}(a)$ an der Stelle a . Für beliebiges $h \in \mathbb{R}^n$ gilt $h = \sum_{i=1}^n h_i e_i$, $h_i \in \mathbb{R}$, also

$$f'(a)h = f'(a)(\sum_{i=1}^n h_i e_i) = \sum_{i=1}^n \Big(f'(a)e_i\Big)h_i = \sum_{i=1}^n D_i f(a) h_i$$

oder, in der üblichen Matrizenschreibweise,

$$f'(a)h = \begin{pmatrix} [f'(a)h]_1 \\ \vdots \\ [f'(a)h]_m \end{pmatrix} = \begin{pmatrix} D_1 f_1(a) & \ldots & D_n f_1(a) \\ \vdots & & \\ D_1 f_m(a) & \ldots & D_n f_m(a) \end{pmatrix} \begin{pmatrix} h_1 \\ \vdots \\ h_n \end{pmatrix}.$$

Also ist

$$\begin{pmatrix} D_1 f_1(a) & \ldots & D_n f_1(a) \\ \vdots & & \\ D_1 f_m(a) & \ldots & D_n f_m(a) \end{pmatrix}$$

die zu den Standardbasen $e_1, \ldots, e_n$ in $\mathbb{R}^n$ und $e_1, \ldots, e_m$ in $\mathbb{R}^m$ gehörende Darstellung von $f'(a)$ also $m \times n$ Matrix. Diese Matrix heißt Jacobi–Matrix von f an der Stelle a .

Um zu prüfen ob f an der Stelle a differenzierbar ist, prüft man zuerst, ob alle partiellen Ableitungen $\frac{\partial f_j}{\partial x_i}(a)$ existieren. Dies ist eine notwendige Bedingung für die Differenzierbarkeit. Wenn alle partiellen Ableitungen existieren braucht aber f nicht differenzierbar zu sein. Daher muß man mit der Matrix

$$T = \Big(\frac{\partial f_j}{\partial x_i}(a)\Big)_{\substack{i=1, \ldots n \\ j=1, \ldots, m}}$$

prüfen, ob

$$\lim_{\substack{h\to 0\\ h\neq 0}} \frac{f(a+h)-f(a)-Th}{\|h\|} = 0\,,$$

gilt. Falls dies richtig ist, ist $f'(a) := T$ die Ableitung von f an der Stelle a.

5 b.) Beispiele

1.) Sei $f : \mathbb{R}^2 \to \mathbb{R}^2\,,\quad x = (x_1, x_2) \mapsto f(x) = (f_1(x_1, x_2),\ f_2(x_1, x_2))$, definiert durch

$$\begin{aligned} f_1(x_1,x_2) &= x_1^2 - x_2^2 \\ f_2(x_1,x_2) &= 2x_1x_2\,. \end{aligned}$$

Falls f an der Stelle $a = (a_1, a_2) \in \mathbb{R}^2$ differenzierbar ist, muß gelten

$$f'(a) = \begin{pmatrix} 2a_1 & -2a_2 \\ 2a_2 & 2a_1 \end{pmatrix}.$$

Sei $r(x) = \Big(r_1(x), r_2(x)\Big) = \dfrac{f(x) - f(a) - f'(a)(x-a)}{\|x-a\|}\,.$

Dann gilt

$$\begin{aligned} r_1(x) &= \frac{x_1^2 - x_2^2 - a_1^2 + a_2^2 - 2a_1(x_1-a_1) + 2a_2(x_2-a_2)}{\|x-a\|} = \frac{(x_1-a_1)^2 - (x_2-a_2)^2}{\|x-a\|} \\ r_2(x) &= \frac{2x_1x_2 - 2a_1a_2 - 2a_2(x_1-a_1) - 2a_1(x_2-a_2)}{\|x-a\|} = \frac{2(x_1-a_1)(x_2-a_2)}{\|x-a\|}\,. \end{aligned}$$

Mit der Maximumsnorm ergibt sich

$$\begin{aligned} |r_1(x)| &\le 2\|x-a\| \\ |r_2(x)| &\le 2\|x-a\|\,, \end{aligned}$$

also

$$\lim_{x\to a} \|r(x)\| \le \lim_{x\to a} 2\|x-a\| = 0\,,$$

also ist f an der Stelle a, und weil a beliebig war, in ganz $\mathbb{R}^2$ differenzierbar.

2.) Sei $A : \mathbb{R}^n \to \mathbb{R}^m$ linear, und sei $f : \mathbb{R}^n \to \mathbb{R}^m$ deffiniert durch

$$f(x) = Ax + c\,,\quad c \in \mathbb{R}^m\,.$$

Dann ist f in ganz $\mathbb{R}^n$ differenzierbar, und es gilt $f'(a) = A$. Denn

$$\frac{f(a+h) - f(a) - Ah}{\|h\|} = \frac{A(a+h) + c - Aa - c - Ah}{\|h\|} = 0\,.$$

3.) $f : \mathbb{R}^2 \to \mathbb{R}$ sei definiert durch

$$f(x_1, x_2) = \begin{cases} 0\,, & \text{für} \quad (x_1, x_2) = 0 \\ \dfrac{x_1 x_2}{\sqrt{x_1^2 + x_2^2}}\,, & \text{für} \quad (x_1, x_2) \neq 0\,. \end{cases}$$

f ist an der Stelle $a = 0$ nicht differenzierbar, aber die partiellen Ableitungen existieren im Nullpunkt und es gilt

$$\frac{\partial f}{\partial x_1}(0) = 0\,, \quad \frac{\partial f}{\partial x_2}(0) = 0\,.$$

Wäre also f in 0 differenzierbar, müßte

$$\mathrm{grad} f(0) = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

sein. Es gilt aber für

$$r(h) = \frac{f(h) - f(0)}{|h|} = \frac{h_1 h_2}{\sqrt{h_1^2 + h_2^2}\sqrt{h_1^2 + h_2^2}} = \frac{h_1 h_2}{h_1^2 + h_2^2}$$

auf der Diagonalen $h = (h_1, h_1)$

$$\lim_{h \to 0} |r(h)| = \lim_{h \to 0} \frac{h_1^2}{2h_1^2} = \frac{1}{2} \neq 0\,.$$

5 c.) Einfache Eigenschaften und Rechenregeln für differenzierbare Abbildungen

Zur Vorbereitung benötige ich ein Resultat über lineare Abbildungen:

Lemma: Sei $A : \mathbb{R}^n \to \mathbb{R}^m$ linear. Dann ist A stetig, und es existiert eine nicht negative Konstante, die mit $\|A\|$ bezeichnet wird, so daß $\|Ax\| \leq \|A\|\,\|x\|$ gilt für alle $x \in \mathbb{R}^n$.

Beweis: Es existiert eine $m \times n$ Matrix $(a_{ij})_{\substack{i=1,\ldots,m \\ j=1,\ldots,n}}$, so daß für $y = Ax$ gilt

$$\begin{aligned} y_1 &= a_{11}x_1 + \ldots + a_{1n}x_n \\ &\vdots \\ y_m &= a_{m1}x_1 + \ldots + a_{mn}x_n\,. \end{aligned}$$

Jede dieser Abbildungsgleichungen definiert eine stetige Abbildung von $\mathbb{R}^n$ nach $\mathbb{R}$, also ist A stetig.

Sei $E = \{x \in \mathbb{R}^n \,\big|\, \|x\| \leq 1\}$. Dies ist eine kompakte Menge. Also existiert

$$\|A\| := \sup_{x \in E} \|Ax\|\,,$$

da A stetig ist. Für alle $x \in \mathbb{R}^n$ gilt nun

$$\|Ax\| = \left\|A\Big(\|x\|\frac{x}{\|x\|}\Big)\right\| = \left\|\,\|x\|A\Big(\frac{x}{\|x\|}\Big)\right\| = \|x\| \left\|A\Big(\frac{x}{\|x\|}\Big)\right\| \le \|A\|\,\|x\|\,,$$

wegen $\frac{x}{\|x\|} \in E$. ■

Definition: Die Zahl $\|A\| := \sup_{\|x\|\le 1} \|Ax\|$ heißt Norm der linearen Abbildung A.

Lemma: Es gilt für alle $x \in \mathbb{R}^n$, $c \in \mathbb{R}$ und für alle linearen Abbildungen $A, B : \mathbb{R}^n \to \mathbb{R}^m$

(i) $\|A\| \ge 0$,
$A = 0 \iff \|A\| = 0$,

(ii) $\|cA\| = |c|\,\|A\|$,

(iii) $\|A + B\| \le \|A\| + \|B\|$,

(iv) $\|Ax\| \le \|A\|\,\|x\|$.

Beweis: (ii) ist klar, (iv) wurde schon gezeigt. Zum Beweis von (iii) beachte, daß

$$\begin{aligned} \|(A+B)x\| &= \|Ax + Bx\| \le \|Ax\| + \|Bx\| \\ &\le \|A\|\,\|x\| + \|B\|\,\|x\| = (\|A\| + \|B\|)\|x\| \end{aligned}$$

gilt. Hieraus folgt

$$\|A+B\| = \sup_{\|x\|\le 1} \|(A+B)x\| \le \sup_{\|x\|\le 1} (\|A\| + \|B\|)\|x\| = \|A\| + \|B\|\,.$$

Zum Beweis von (i) sei $\|A\| = 0$. Für alle $x \in \mathbb{R}^n$ folgt dann aus (iv), daß $0 \le \|Ax\| \le \|A\|\,\|x\| = 0$ gilt. Somit ist $A = 0$. Die anderen Aussagen von (i) sind klar. ■

Die Menge $L(\mathbb{R}^n, \mathbb{R}^m)$ der linearen Abbildungen von $\mathbb{R}^n$ nach $\mathbb{R}^m$ bildet einen Vektorraum, und dieses Lemma zeigt, daß $\|A\|$ wirklich die Eigenschaften einer Norm besitzt. Also wird $L(\mathbb{R}^n, \mathbb{R}^m)$ mit dieser Norm zu einem normierten Raum.

Wir studieren nun wieder differenzierbare Abbildungen.

Seien $U \subseteq \mathbb{R}^n$ offen, $f : U \to \mathbb{R}^m$, $a \in U$. Wenn für f alle Richtungsableitungen im Punkt a existieren, braucht f doch nicht stetig zu sein. Ein Beispiel ist

$$f : \mathbb{R}^2 \to \mathbb{R}\,,$$

$$f(x_1, x_2) = \begin{cases} 0\,, & (x_1, x_2) = 0 \\ \dfrac{x_1 x_2^2}{x_1^2 + x_2^6}\,, & (x_1, x_2) \ne 0\,. \end{cases}$$

Die Richtungsableitungen existieren alle im Nullpunkt, weil für $v = (v_1, v_2) \in \mathbb{R}^2$, $v \neq 0$, gilt

$$D_v f(0) = \lim_{\substack{t\to 0\\ t\neq 0}} \frac{f(tv) - f(0)}{t} = \begin{cases} \lim\limits_{t\to 0} \dfrac{v_1 v_2^2}{v_1^2 + t^4 v_2^6} = \dfrac{v_2^2}{v_1}\,, & v_1 \neq 0 \\ 0\,, & v_1 = 0\,. \end{cases}$$

Aber für

$$h = (h_1, \sqrt{h_1})\,, \quad h_1 > 0\,,$$

gilt

$$f(h) = \frac{h_1^2}{h_1^2 + h_1^3} = \frac{1}{1 + h_1} \to 1 \neq f(0)\,,$$

für $h_1 \to 0$.

Es gilt aber

Satz: Sei $U \subseteq \mathbb{R}^n$, und $f : U \to \mathbb{R}^m$ sei an der Stelle $a \in U$ differenzierbar. Dann existiert $c > 0$, so daß für alle x aus einer Umgebung von a gilt

$$\|f(x) - f(a)\| \leq c\|x - a\|\,.$$

Insbesondere ist f in a stetig.

Beweis: Es gilt

$$f(x) = f(a) + f'(a)(x - a) + r(x)\|x - a\|\,,$$

also

$$\|f(x) - f(a)\| \leq \|f'(a)\|\,\|x - a\| + \|r(x)\|\,\|x - a\|\,.$$

Wegen $\lim_{x\to a} r(x) = 0$ folgt

$$\|f(x) - f(a)\| \leq c\|x - a\|\,,$$

also

$$\lim_{x\to a} \|f(x) - f(a)\| = 0\,.$$

∎

Satz: Sei $U \subseteq \mathbb{R}^n$ offen, $f : U \to \mathbb{R}^m$, $g : U \to \mathbb{R}^m$ seien beide an der Stelle $a \in U$ differenzierbar. Dann sind auch $f + g$ und cf $(c \in \mathbb{R})$ an der Stelle a differenzierbar und es gilt

$$\begin{aligned} (f + g)'(a) &= f'(a) + g'(a) \\ (cf)'(a) &= cf'(a)\,. \end{aligned}$$

Beweis: Es gilt

$$\begin{aligned} f(a+h) &= f(a) + f'(a)h + r_1(a+h)\|h\|\,, \quad \lim_{h\to 0} r_1(a+h) = 0 \\ g(a+h) &= g(a) + g'(a)h + r_2(a+h)\|h\|\,, \quad \lim_{h\to 0} r_2(a+h) = 0\,. \end{aligned}$$

Also folgt

$$(f+g)(a+h) = (f+g)(a) + \Big(f'(a) + g'(a)\Big)h + (r_1 + r_2)(a+h)\|h\|\,.$$

Hieraus resultiert $(f+g)'(a) = f'(a) + g'(a)$. Die andere Aussage ergibt sich ebenso. ■

Satz (Produktregel): Die Funktionen $f : U \to \mathbb{R}$ und $g : U \to \mathbb{R}$ seien beide an der Stelle $a \in U$ differenzierbar. Dann ist auch $f \cdot g$ an der Stelle a differenzierbar mit

$$(f \cdot g)'(a)h = f(a)g'(a)h + g(a)f'(a)h\,.$$

Beweis:

$$\begin{aligned} (f \cdot g)(a+h) = &\Big(f(a) + f'(a)h + r_1(a+h)\|h\|\Big) \\ &\cdot \Big(g(a) + g'(a)h + r_2(a+h)\|h\|\Big) \\ = \;\; &(f \cdot g)(a) + f(a)g'(a)h + g(a)f'(a)h + r(a+h)\|h\|\,, \end{aligned}$$

mit

$$\begin{aligned} r(a+h)\|h\| &= f'(a)h\, g'(a)h + \Big(g(a) + g'(a)h\Big)r_1(a+h)\|h\| \\ &\quad + \Big(f(a) + f'(a)h\Big)r_2(a+h)\|h\| + r_1(a+h)r_2(a+h)\|h\|^2\,. \end{aligned}$$

Es gilt

$$\begin{aligned} \lim_{h\to 0} \|r(a+h)\| \;\leq\; &\lim_{h\to 0} \frac{1}{\|h\|}\Big[\|f'(a)\|\,\|g'(a)\|\,\|h\|^2 \\ &+ \Big(|g(a)| + \|g'(a)\|\,\|h\|\Big)|r_1(a+h)|\,\|h\| \\ &+ \Big(|f(a)| + \|f'(a)\|\,\|h\|\Big)|r_2(a+h)|\,\|h\| \\ &+ |r_1(a+h)|\,|r_2(a+h)|\,\|h\|^2\Big] = 0\,. \end{aligned}$$

■

Bemerkung: Natürlich kann man die Produktregel auch in der Form

$$\mathrm{grad}(fg)(a) = f(a)\,\mathrm{grad}g(a) + g(a)\,\mathrm{grad}f(a)$$

schreiben.

Satz (Kettenregel): Sei $V \subseteq \mathbb{R}^n$ offen, $g : V \to \mathbb{R}^m$ sei an der Stelle $b \in V$ differenzierbar. Sei $U \subseteq \mathbb{R}^p$ offen, $f : U \to V$ sei an der Stelle $a \in U$ differenzierbar, und es sei $b = f(a)$. Dann ist $g \circ f$ an der Stelle $a \in U$ differenzierbar, und es gilt

$$(g \circ f)'(a) = g'\Big(f(a)\Big) \circ f'(a)\,.$$

Beweis: Zur Abkürzung seien

$$T_2 = g'(b)\,, \quad T_1 = f'(a)\,,$$

und für $h \in \mathbb{R}^p$, $\|h\|$ genügend klein, sei

$$R(h) = (g \circ f)(a+h) - (g \circ f)(a) - T_2T_1h\,.$$

Es muß gezeigt werden, daß

$$\lim_{\substack{h\to 0\\ h\neq 0}} \frac{\|R(h)\|}{\|h\|} = 0$$

ist. Es gilt

$$\begin{aligned} f(x) - f(a) - T_1(x-a) &= r_1(x-a)\|x-a\|\,, \quad \lim_{x\to 0} r_1(x) = 0\\ g(y) - g(b) - T_2(y-b) &= r_2(y-b)\|y-b\|\,, \quad \lim_{y\to 0} r_2(y) = 0\,. \end{aligned}$$

Somit folgt

$$\begin{aligned} R(h) &= g\Big(f(a+h)\Big) - g\Big(f(a)\Big) - T_2\Big(f(a+h) - f(a)\Big)\\ &\quad + T_2\Big(f(a+h) - f(a) - T_1h\Big)\\ &= r_2\Big(f(a+h) - f(a)\Big)\|f(a+h) - f(a)\| + T_2\Big(r_1(h)\|h\|\Big)\,, \end{aligned}$$

d.h.

$$\begin{aligned} \lim_{h\to 0} \frac{\|Rh\|}{\|h\|} &\le \lim_{h\to 0}\Big[\frac{1}{\|h\|}\,\|r_2\Big(f(a+h) - f(a)\Big)\|\;\|f(a+h) - f(a)\|\Big]\\ &\quad + \lim_{h\to 0} \|T_2\Big(r_1(h)\Big)\|\Big]\,. \end{aligned}$$

Wegen der Stetigkeit von T_2 folgt $\lim_{h\to 0} T_2(r_1(h)) = 0$. Wegen $\|f(a+h) - f(a)\| \le c\|h\|$ ergibt sich

$$\lim_{h\to 0} \frac{\|Rh\|}{\|h\|} \le c \lim_{h\to 0} \|r_2\Big(f(a+h) - f(a)\Big)\| = 0\,.$$

Damit ist der Satz bewiesen. ∎

Für die Jacobi–Matrizen von $f : U \subseteq \mathbb{R}^p \to \mathbb{R}^n\,,\quad g : V \subseteq \mathbb{R}^n \to \mathbb{R}^m$ und $h = g \circ f : U \subseteq \mathbb{R}^p \to \mathbb{R}^m$ ergibt sich also

$$\begin{pmatrix} \frac{\partial h_1}{\partial x_1} & \cdots & \frac{\partial h_1}{\partial x_p} \\ \vdots & & \\ \frac{\partial h_m}{\partial x_1} & \cdots & \frac{\partial h_m}{\partial x_p} \end{pmatrix} = \begin{pmatrix} \frac{\partial g_1}{\partial y_1} & \cdots & \frac{\partial g_1}{\partial y_n} \\ \vdots & & \\ \frac{\partial g_m}{\partial y_1} & \cdots & \frac{\partial g_m}{\partial y_n} \end{pmatrix} \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_p} \\ \vdots & & \\ \frac{\partial f_n}{\partial x_1} & \cdots & \frac{\partial f_n}{\partial x_p} \end{pmatrix},$$

wobei die partiellen Ableitungen von h und f an der Stelle a, von g an der Stelle $b = f(a)$ zu bilden sind.

Es ergibt sich also

$$\frac{\partial h_j}{\partial x_i}(a) = \sum_{k=1}^{n} \frac{\partial g_j}{\partial y_k}(b)\, \frac{\partial f_k}{\partial x_i}(a)\,, \quad i = 1, \ldots, p \quad j = 1, \ldots, m\,.$$

Folgerung: Sei $U \subseteq \mathbb{R}^n$ offen, $f : U \to \mathbb{R}$ sei in $a \in U$ differenzierbar, und es gelte $f(a) \neq 0$. Dann gilt

$$\operatorname{grad} \frac{1}{f}(a) = (\frac{1}{f})'(a) = -\frac{1}{f(a)^2} \operatorname{grad} f(a)\,.$$

Beweis: Betrachte die Abbildung $g : \mathbb{R}\backslash\{0\} \to \mathbb{R}\,,\ g(x) := \frac{1}{x}$. Dann gilt

$$\frac{1}{f} = g \circ f\,,$$

also

$$\operatorname{grad} \frac{1}{f}(a) = \Big(\frac{1}{f}\Big)'(a) = g'\Big(f(a)\Big) f'(a) = -\frac{1}{f(a)^2} \operatorname{grad} f(a)\,.$$

■

Man kann die Ableitung der Umkehrabbildung einer bijektiven Abbildung $f : U \to V$, $U, V \subseteq \mathbb{R}^n$ offen, mit der Kettenregel berechnen.

Denn sei $g : V \to U$, die Umkehrabbildung zu f, sei f an der Stelle $a \in U$ und g an der Stelle $b = f(a) \in V$ differenzierbar. Dann gilt

$$g \circ f = \mathrm{id}_U\,,$$

also

$$(g \circ f)'(a) = g'\Big(f(a)\Big) \circ f'(a) = \mathrm{id}_{\mathbb{R}^n}\,,$$

folglich

$$g'\Big(f(a)\Big) = \Big(f'(a)\Big)^{-1},$$

oder

$$g'(b) = \Big[f'\Big(g(b)\Big)\Big]^{-1}.$$

Wenn man voraussetzt, daß die Umkehrabbildung der linearen Abbildung $f'(a)$ existiert, genügt es sogar vorauszusetzen, daß g stetig sei. Nach einem Satz der linearen Algebra existiert die Umkehrabbildung von $f'(a)$, wenn die Determinante $\det f'(a)$ der $f'(a)$ repräsentierenden $n \times n$–Matrix von Null verschieden ist. Man nennt $\det f'(a)$ Jacobi–Determinante.

Satz: Sei $U \subseteq \mathbb{R}^n$ offen, $f : U \to \mathbb{R}^n$ sei umkehrbar, an der Stelle a differenzierbar, und die Jacobi–Determinante $\det f'(a)$ sei von Null verschieden. Sei $f(U)$ offen und die Umkehrabbildung $g : f(U) \to \mathbb{R}^n$ sei an der Stelle $b = f(a)$ stetig. Dann ist g an der Stelle b differenzierbar, und es gilt

$$g'(b) = \Big(f'(a)\Big)^{-1}.$$

Beweis: Zunächst zeige ich: Es gibt eine Umgebung V von b und eine Konstante $c > 0$ mit

$$\frac{\|g(y) - g(b)\|}{\|y - b\|} \leq c\,,$$

für alle $y \in V \cap f(U)$.

Es gilt

$$f(x) - f(a) = f'(a)(x - a) \quad + \quad r(x)\|x - a\|\,, \quad \lim_{x \to a} r(x) = 0\,.$$

Also folgt mit der umgekehrten Dreiecksungleichung

$$\begin{aligned}
\frac{\|g(y) - g(b)\|}{\|y - b\|} &= \frac{\|g(y) - g(b)\|}{\|f(g(y)) - f(g(b))\|} \\
= &\quad \frac{\|g(y) - g(b)\|}{\|f'(a)(g(y) - g(b)) \quad + \quad r(g(y))\|g(y) - g(b)\|\,\|} \\
\leq &\quad \frac{\|(f'(a))^{-1}f'(a)\,(g(y) - g(b))\|}{\|f'(a)(g(y) - g(b))\| - \|r(g(y))\|\,\|(f'(a))^{-1}f'(a)(g(y) - g(b))\|} \\
\leq &\quad \frac{\|(f'(a))^{-1}\|\,\|f'(a)(g(y) - g(b))\|}{\|f'(a)(g(y) - g(b))\|(1 - \|r(g(y))\|\,\|(f'(a))^{-1}\|)} \\
= &\quad \frac{\|(f'(a))^{-1}\|}{1 - \|r(g(y))\|\,\|f'(a)^{-1}\|}\,.
\end{aligned}$$

Wegen $\lim_{y \to b} r(g(y)) = 0$ folgt die Behauptung. Hierbei wird die Stetigkeit von g benutzt. Nun ergibt sich der Satz folgendermaßen:

Es muß gezeigt werden, daß

$$\lim_{\substack{y \to b \\ y \neq b}} \frac{g(y) - g(b) - f'(a)^{-1}(y-b)}{\|y-b\|} = 0$$

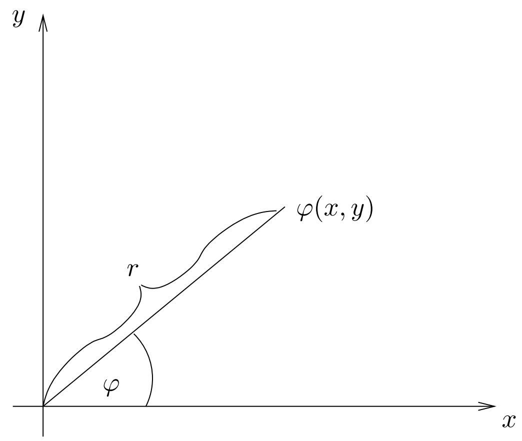
ist. Es gilt

$$\begin{aligned}
&\lim_{\substack{y \to b \\ y \neq b}} \frac{g(y) - g(b) - f'(a)^{-1}(y-b)}{\|y-b\|} \\
&\quad = \lim_{\substack{y \to b \\ y \neq b}} \frac{g(y) - g(b) - f'(a)^{-1}(f(g(y)) - f(g(b)))}{\|y-b\|} \\
&\quad = \lim_{\substack{y \to b \\ y \neq b}} \frac{g(y) - g(b) - f'(a)^{-1}\Big(f'(a)(g(y) - g(b)) + r(g(y))\|g(y) - g(b)\|\Big)}{\|y-b\|} \\
&\quad = \lim_{\substack{y \to b \\ y \neq b}} f'(a)^{-1}\Big(r\Big(g(y)\Big)\Big) \frac{\|g(y) - g(b)\|}{\|y-b\|} = 0\,.
\end{aligned}$$

■

Beispiel (Polarkoordinatenabbildung): Seien $\varepsilon > 0$ und $c_2 > c_1 > 0$, und für $c_1 \leq r \leq c_2$, $0 \leq \varphi \leq 2\pi - \varepsilon$ sei

$$\begin{aligned}
x &= f_1(r, \varphi) = r\cos\varphi \\
y &= f_2(r, \varphi) = r\sin\varphi\,.
\end{aligned}$$



Diese Abbildung ist injektiv, differenzierbar, die Jacobi–Determinante ist von Null verschieden, und die Umkehrabbildung ist stetig, weil f auf einer kompakten Menge definiert ist. Ohne die Umkehrabbildung bestimmen zu müssen, kann die Ableitung der Umkehrabbildung bestimmt werden. Im Punkt $(x, y) = f(r, \varphi)$ gilt

$$[f^{-1}]'(x, y) = f'(r, \varphi)^{-1} = \begin{pmatrix} \cos\varphi & -r\sin\varphi \\ \sin\varphi & r\cos\varphi \end{pmatrix}^{-1}$$

$$= \begin{pmatrix} \cos\varphi & \sin\varphi \\ -\frac{1}{r}\sin\varphi & \frac{1}{r}\cos\varphi \end{pmatrix} = \begin{pmatrix} \frac{x}{\sqrt{x^2+y^2}} & \frac{y}{\sqrt{x^2+y^2}} \\ \frac{-y}{x^2+y^2} & \frac{x}{x^2+y^2} \end{pmatrix}.$$

5 d.) Mittelwertsatz

Der Mittelwertsatz für reelle Funktionen kann auf *reellwertige* Funktionen verallgemeinert werden.

Satz: Sei $U \subseteq \mathbb{R}^n$ offen, $f : U \to \mathbb{R}$ sei differenzierbar, und die Verbindungsstrecke der beiden Punkte $a, b \in U$ sei ganz in U enthalten. Dann gibt es einen Punkt c auf dieser Verbindungsstrecke mit

$$f(b) - f(a) = f'(c)(b - a).$$

Beweis: Definiere die Abbildung $\gamma : [0, 1] \to U$ durch $t \mapsto \gamma(t) := a + t(b - a)$. Hierdurch wird $[0, 1]$ auf die Verbindungstrecke von a und b abgebildet. γ ist differenzierbar mit

$$\gamma'(t) = b - a.$$

Auf die differenzierbare Funktion $F : [0, 1] \to \mathbb{R}$,

$$F = f \circ \gamma,$$

wende man den Mittelwertsatz für reelle Funktionen an. Es folgt mit geeignetem $\vartheta \in (0, 1)$

$$f(a) - f(b) = F(1) - F(0) = F'(\vartheta) = f'\Big(\gamma(\vartheta)\Big)\gamma'(\vartheta) = f'(c)(b - a),$$

mit $c = \gamma(\vartheta)$. ■

Natürlich kann man den Mittelwertsatz auch folgendermaßen formulieren: Zu $x, x+h \in U$ gibt es $\vartheta, 0 < \vartheta < 1$, mit

$$f(x + h) - f(x) = f'(x + \vartheta h)h.$$

Schrankensatz: Sei $U \subseteq \mathbb{R}^n$ offen, $f : U \to \mathbb{R}^m$ sei differenzierbar, und die Verbindungsstrecke der beiden Punkte $x, x + h \in U$ sei in U enthalten. Dann gibt es eine Zahl ϑ mit $0 < \vartheta < 1$, so dass

$$\|f(x + h) - f(x)\| \leq \|f'(x + \vartheta h)\| \, \|h\|$$

gilt.

Beweis: Um diesen Satz zu beweisen, benötigen wir folgendes Resultat, das wir nicht beweisen:

Satz (von Hahn-Banach): Sei $\|\cdot\|$ eine Norm auf $\mathbb{R}^m$. Dann gibt es zu jedem $u \in \mathbb{R}^m$ eine lineare Abbildung $A_u : \mathbb{R}^m \to \mathbb{R}$, so dass $\|A_u\| = 1$ und $A_u(u) = \|u\|$ gilt.

Beispiel: Es sei $u \in \mathbb{R}^m$, $u \neq 0$. Für die Euklidische Norm $\|\cdot\| = |\cdot|$ definiere man A_u durch

$$A_u(v) = \frac{u}{|u|} \cdot v; \quad v \in \mathbb{R}^m.$$

Dann gilt $A_u(u) = \frac{u}{|u|} \cdot u = |u|$ und

$$\begin{aligned} 1 = \left|\frac{u}{|u|}\right| &= \frac{1}{|u|} A_u(u) \leq \frac{1}{|u|} \|A_u\|\,|u| = \|A_u\| \\ &= \sup_{|v|\leq 1} |A_u(v)| = \sup_{|v|\leq 1} \left|\frac{u}{|u|} \cdot v\right| \leq \sup_{|v|\leq 1} \frac{|u| \cdot |v|}{|u|} = 1. \end{aligned}$$

Also gilt $\|A_u\| = 1$.

Beweis des Schrankensatzes: Zu $f(x+h) - f(x) \in \mathbb{R}^m$ sei $A : \mathbb{R}^m \to \mathbb{R}$ eine lineare Abbildung mit $\|A\| = 1$ und $A\big(f(x+h) - f(x)\big) = \|f(x+h) - f(x)\|$. Als lineare Abbildung ist A differenzierbar mit Ableitung $A'(y) = A$ für alle $y \in \mathbb{R}^m$. Wendet man den Mittelwertsatz auf die Abbildung $F = A \circ f : U \to \mathbb{R}$ an und verwendet die Kettenregel, dann erhält man mit einer Zahl $0 < \vartheta < 1$, dass

$$\begin{aligned} \|f(x+h) - f(x)\| &= A\big(f(x+h) - f(x)\big) \\ &= A\big(f(x+h)\big) - A\big(f(x)\big) = F(x+h) - F(x) = F'(x+\vartheta h)h \\ &= Af'(x+\vartheta h)h \leq \|A\|\,\|f'(x+\vartheta h)\|\,\|h\| = \|f'(x+\vartheta h)\|\,\|h\|. \end{aligned}$$

■

Satz: Sei $U \subseteq \mathbb{R}^n$ offen und wegzusammenhängend. $f : U \to \mathbb{R}^m$ sei differenzierbar. Dann gilt: f ist konstant, genau dann wenn $f'(x) = 0$ ist für alle $x \in U$.

Zum Beweis benützen wir folgendes

Lemma: Sei $U \subseteq \mathbb{R}^n$ offen und wegzusammenhängend, und seien $a, b \in U$. Dann können a, b durch einen ganz in U verlaufenden Streckenzug mit den ‚Eckpunkten'

$$a_0 = a\,, \;\; a_1, \,\ldots, a_{k-1}, \;\; a_k = b$$

verbunden werden.

Dieses Lemma beweise ich nicht. Man findet einen Beweis im Buch von Barner–Flohr, Analysis II, S. 56.

Beweis des Satzes: Falls f konstant, ist $f' = 0$. Zum Beweis der Umkehrung sei $f'(x) = 0$ für alle $x \in U$. Es genügt, die Behauptung für Funktionen $f : U \to \mathbb{R}$ zu beweisen, weil man im allgemeinen Fall die Komponentenfunktionen $f_1, \ldots, f_m$ von f betrachten kann. Sei also f reellwertig.

Seien $a, b \in U$. Man verbinde diese Punkte durch einen Streckenzug in U mit den angegebenen Eckpunkten, und wende den Mittelwertsatz auf jede der Strecken mit den Endpunkten a_j, a_{j+1} an, $j = 0, 1, \ldots, k-1$. Es folgt

$$f(a_{j+1}) = f(a_j) + f'(c)(a_{j+1} - a_j) = f(a_j)\,,$$

also

$$f(b) = f(a)\,.$$

■

Wenn f differenzierbar ist, existieren alle partiellen Ableitungen. Wenn die partiellen Ableitungen existieren, braucht f aber nicht differenzierbar zu sein. Es gilt jedoch:

Satz: Sei $U \in \mathbb{R}^n$ offen. Wenn die Funktion $f : U \to \mathbb{R}^m$ sämtliche partiellen Ableitungen $\frac{\partial f_j}{\partial x_i}$, $i = 1, \ldots, n$, $j = 1, \ldots, m$, besitzt, und diese an der Stelle $a \in U$ stetig sind, dann ist f an der Stelle a differenzierbar.

Beweis: Es genügt zu zeigen, daß jede der Komponentenfunktionen $f_1, \ldots, f_m$ differenzierbar ist. Also kann man annehmen, daß $f : U \to \mathbb{R}$ gilt. Es ist zu zeigen, daß

$$\lim_{\substack{h \to 0 \\ h \neq 0}} \frac{f(a+h) - f(a) - Th}{\|h\|_\infty} = 0$$

ist mit

$$T := \left(\frac{\partial f}{\partial x_1}(a), \ldots, \frac{\partial f}{\partial x_n}(a) \right).$$

Für $h \in \mathbb{R}^n$ setze

$$\begin{aligned} a_0 &= a\,, \\ a_1 &= a_0 + h_1 e_1 \\ a_2 &= a_1 + h_2 e_2 \\ &\vdots \\ a + h = a_n &= a_{n-1} + h_n e_n\,, \end{aligned}$$

wobei $e_1, \ldots, e_n \in \mathbb{R}^n$ die kanonische Basis sei. Es gilt dann

$$\begin{aligned} & f(a+h) - f(a) \\ = \quad & \Big(f(a+h) - f(a_{n-1})\Big) + \Big(f(a_{n-1}) - f(a_{n-2})\Big) + \ldots + \Big(f(a_1) - f(a)\Big) . \qquad (*) \end{aligned}$$

Läuft x auf der Verbindungsstrecke zwischen a_{j-1} und a_j, dann variiert nur die Komponente x_j von x. Da die Abbildung $x_j \mapsto f(x_1, \ldots, x_j, \ldots, x_n)$ nach Voraussetzung differenzierbar ist, kann der Mittelwertsatz auf jeden Summanden in der Formel $(*)$ angewendet werden. c_j sei der Zwischenpunkt auf der Verbindungsstrecke von a_{j-1} und a_j. Dann gilt

$$f(a+h) - f(a) = \sum_{j=1}^{n} \Big(f(a_j) - f(a_{j-1})\Big) = \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}(c_j) h_j .$$

Also folgt

$$\begin{aligned} f(a+h) - f(a) - Th &= \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}(c_j) h_j - Th \\ &= \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}(c_j) h_j - \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}(a) h_j = \sum_{j=1}^{n} \Big(\frac{\partial f}{\partial x_j}(c_j) - \frac{\partial f}{\partial x_j}(a)\Big) h_j , \end{aligned}$$

somit

$$|f(a+h) - f(a) - Th| \leq \|h\|_\infty \sum_{j=1}^{n} \Big| \frac{\partial f}{\partial x_j}(c_j) - \frac{\partial f}{\partial x_j}(a) \Big| .$$

Wegen $\|c_j - a\|_\infty \leq \|h\|_\infty$ folgt die Behauptung aus der Stetigkeit aller partiellen Ableitungen $\frac{\partial f}{\partial x_j}$, in a. ∎

Dieser Satz liefert eine einfache hinreichende Bedingung für die Differenzierbarkeit einer Abbildung

$$f : U \to \mathbb{R}^m , \;\; U \subseteq \mathbb{R}^n .$$

Beispiel: $f : \mathbb{R}^n \backslash \{0\} \to \mathbb{R}$ sei definiert durch

$$f(x) = (x_1^2 + \ldots + x_n^2)^s , \quad s \in \mathbb{R} .$$

Diese Abbildung ist überall differenzierbar. Denn die partiellen Ableitungen

$$\frac{\partial f}{\partial x_j} = s \cdot (x_1^2 + \ldots + x_n^2)^{s-1} 2x_j$$

sind stetig.

5 e.) Stetig differenzierbare Abbildungen

Sei $U \subseteq \mathbb{R}^n$ offen, $f : U \to \mathbb{R}^m$ sei in allen Punkten $x \in U$ differenzierbar. Dann wird durch

$$x \mapsto f'(x) : U \to L(\mathbb{R}^n, \mathbb{R}^m)$$

eine Abbildung von U in die Menge der linearen Abbildungen von $\mathbb{R}^n$ nach $\mathbb{R}^m$ definiert. Wendet man die lineare Abbildung $f'(x)$ auf einen beliebigen Vektor $h \in \mathbb{R}^n$ an, dann erhält man einen Vektor in $\mathbb{R}^m$:

$$f'(x,h) := f'(x)h \in \mathbb{R}^m \,.$$

Also kann man f' auch als Abbildung von $U \times \mathbb{R}^n$ nach $\mathbb{R}^m$ auffassen:

$$(x,h) \mapsto f'(x,h) : U \times \mathbb{R}^n \to \mathbb{R}^m \,.$$

f' ist bezüglich des zweiten Arguments linear. Welche Auffassung man verwendet, ist eine Frage der Zweckmäßigkeit.

Weil $L(\mathbb{R}^n, \mathbb{R}^m)$ mit der am Anfang dieses Abschnittes eingeführten Norm ein normierter Raum ist, ist f' bei beiden Auffassungen eine Abbildung zwischen normierten Räumen. Für Abbildungen zwischen normierten Räumen ist der Begriff der Stetigkeit definiert, und man kann daher untersuchen, ob f' bei einer der beiden verschiedenen Auffassungen eine stetige Abbildung ist. Das folgende Lemma zeigt, daß es bei der Untersuchung der Stetigkeit nicht darauf ankommt, welche Auffassung man zu Grunde legt:

Lemma: Sei $U \subseteq \mathbb{R}^n$ eine offene Menge. Genau dann ist $f' : U \times \mathbb{R}^n \to \mathbb{R}^m$ stetig, wenn $f' : U \to L(\mathbb{R}^n, \mathbb{R}^m)$ stetig ist.

Beweis: Auf $\mathbb{R}^n$ und auf $U \times \mathbb{R}^n \subseteq \mathbb{R}^n \times \mathbb{R}^n$ verwende ich die Maximimsnorm. Sei $f' : U \times \mathbb{R}^n \to \mathbb{R}^m$ stetig und sei $a \in U$. Wähle $c > 0$ mit $K = \{x \in \mathbb{R}^n \,\big|\, \|x-a\|_\infty \leq c\} \subseteq U$. Weil f' auf der kompakten Menge

$$K \times \{h \in \mathbb{R}^n \,\big|\, \|h\|_\infty \leq 1\}$$

gleichmäßig stetig ist, existiert zu jedem $\varepsilon > 0$ ein $\delta > 0$ mit

$$\|f'(x,h) - f'(a,h)\| \leq \varepsilon$$

für alle $x, h \in \mathbb{R}^n$ mit $\|x-a\|_\infty < \delta$ und $\|h\|_\infty \leq 1$, weil dann $\|(x,h)-(a,h)\|_\infty = \|(x-a,h)\|_\infty = \|x-a\|_\infty < \delta$ gilt. Also folgt für diese x und für die Norm $\|f'(x) - f'(a)\|$

der linearen Abbildung $f'(x) - f'(a) \in L(\mathbb{R}^n, \mathbb{R}^m)$:

$$\begin{aligned} \|f'(x) - f'(a)\| &= \sup_{\|h\|\le 1} \left\| \Big[f'(x) - f'(a) \Big] h \right\| \\ &= \sup_{\|h\|\le 1} \|f'(x)h - f'(a)h\| = \sup_{\|h\|\le 1} \|f'(x,h) - f'(a,h)\| \le \varepsilon\,. \end{aligned}$$

Dies bedeutet, daß $f' : U \to L(\mathbb{R}^n, \mathbb{R}^m)$ in a stetig ist. Weil a beliebig gewählt war, ist diese Abbildung stetig.

Sei umgekehrt $f' : U \to L(\mathbb{R}^n, \mathbb{R}^m)$ stetig und sei $(a, h) \in U \times \mathbb{R}^n$. Zu jedem $\varepsilon > 0$ gibt es dann eine Zahl $\delta > 0$, die kleiner oder gleich $\min(\varepsilon, 1)$ gewählt werden kann, so daß $\|f'(x) - f'(a)\| \le \varepsilon$ gilt für alle $x \in U$ mit $\|x - a\|_\infty < \delta$. Für $(x, h_1) \in U \times \mathbb{R}^n$ mit $\|(x, h_1) - (a, h)\|_\infty < \delta$ folgt dann

$$\begin{aligned} \|f'(x,h_1) - f'(a,h)\| &= \|f'(x)h_1 - f'(a)h\| \\ &= \Big\| \Big[f'(x) - f'(a) \Big] h_1 - f'(a)\,(h_1 - h) \Big\| \\ &\le \|f'(x) - f'(a)\|\,\|h_1\| + \|f'(a)\|\,\|h_1 - h\| \\ &\le \varepsilon(\|h\|_\infty + \|h_1 - h\|_\infty) + \|f'(a)\|\delta \le \varepsilon\Big(\|h\|_\infty + 1 + \|f'(a)\|\Big)\,, \end{aligned}$$

wegen $\|x - a\|_\infty$, $\|h_1 - h\|_\infty < \delta \le \min(\varepsilon, 1)$. Weil $\|h\|_\infty + 1 + \|f'(a)\|$ unabhängig von (x, h_1) ist, folgt hieraus die Stetigkeit der Abbildung $f' : U \times \mathbb{R}^n \to \mathbb{R}^m$ in (a, h). Da dieser Punkt beliebig gewählt war, ist diese Abbildung stetig. ∎

Definition: (i) Sei $U \subseteq \mathbb{R}^n$ offen und sei $f : U \to \mathbb{R}^m$ differenzierbar. Ist $f' : U \times \mathbb{R}^n \to \mathbb{R}^m$ beziehungsweise $f' : U \to L(\mathbb{R}^n, \mathbb{R}^m)$ stetig, dann heißt f stetig differenzierbar.

(ii) Seien $U, V \subseteq \mathbb{R}^n$ offen und sei $f : U \to V$ stetig differenzierbar und umkehrbar. Ist die Umkehrabbildung $f^{-1} : V \to U$ ebenfalls stetig differenzierbar, dann heißt f Diffeomorphismus.

Der folgende Satz gibt ein handhabbares Kriterium, mit dem man nachprüfen kann, ob eine Abbildung stetig differenzierbar ist:

Satz: Sei $U \subseteq \mathbb{R}^n$ offen. Die Abbildung $f : U \to \mathbb{R}^m$ ist stetig differenzierbar, genau dann wenn alle partiellen Ableitungen $\frac{\partial}{\partial x_i} f_j$ in U existieren und stetig sind.

Beweis: Die Abbildung $f' : U \times \mathbb{R}^n \to \mathbb{R}^m$ ist stetig, genau dann wenn jede der Komponentenfunktionen

$$(x, h) \mapsto f_j'(x, h) = \sum_{i=1}^{n} \frac{\partial}{\partial x_i} f_j(x) h_i \qquad (*)$$

stetig ist. Wählt man für h den Einheitsbasisvektor e_i , dann folgt aus der stetigen Differenzierbarkeit von f , daß die partielle Ableitung

$$x \mapsto \frac{\partial}{\partial x_i} f_j(x) = f_j'(x, e_i) : U \to \mathbb{R}$$

stetig ist. Wenn umgekehrt alle partiellen Ableitungen von f in U existieren und stetig sind, dann ist f in U differenzierbar und die j–te Komponente der Ableitung ist gegeben durch die rechte Seite von $(*)$. Man sieht sofort, daß diese rechte Seite eine stetige Funktion von (x, h) ist. ■

5 f.) Höhere Ableitungen, Taylorsche Formel

Die Ableitung von $f : U \to \mathbb{R}^m$ ist eine Abbildung $f : U \to L(\mathbb{R}^n, \mathbb{R}^m)$. Die Ableitung von f' wird man als zweite Ableitung f'' von f bezeichnen. Also ist die zweite Ableitung $f''(x)$ von f an der Stelle x eine lineare Abbildung von $\mathbb{R}^n$ in den Raum der linearen Abbildungen $L(\mathbb{R}^n, \mathbb{R}^m)$:

$$f'' : U \to L\Big(\mathbb{R}^n, L(\mathbb{R}^n, \mathbb{R}^m)\Big) .$$

Es ist möglich, die zweite Ableitung von f so zu definieren, weil $L(\mathbb{R}^n, \mathbb{R}^m)$ ein normierter Raum ist (sogar ein Banachraum). Denn man kann die Definition der Ableitung einer Funktion von $\mathbb{R}^n$ nach $\mathbb{R}^m$ ohne Änderung auf Funktionen zwischen allgemeinen normierten Räumen übertragen. Jedoch will ich die zweite Ableitung weniger abstrakt aber in äquivalenter Weise folgendermaßen definieren:

Definition: Sei $U \subseteq \mathbb{R}^n$ eine offene Menge und sei $f : U \to \mathbb{R}^m$ differenzierbar in U . Die Funktion f heißt zweimal differenzierbar in einem Punkt $x \in U$, wenn zu jeden festen $h \in \mathbb{R}^n$ die durch

$$g_h(x) = f'(x, h) = f'(x)h$$

definierte Funktion $g_h : U \to \mathbb{R}^m$ in x differenzierbar ist. Als zweite Ableitung von f im Punkt x bezeichnet man die durch

$$f''(x, h, k) = g_h'(x)k$$

definierte Funktion $(h, k) \mapsto f''(x, h, k) : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^m$. Ist f in jedem Punkt von U zweimal differenzierbar, dann gilt $f'' : U \times \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^m$.

Für jedes $x \in U$ ist

$$(h, k) \mapsto f''(x, h, k) : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^m$$

eine bilineare Abbildung, d.h. eine Abbildung, die in beiden Variablen linear ist. Denn da $g_{h_1+h_2}(x) = f'(x)(h_1 + h_2) = f'(x)h_1 + f'(x)h_2 = g_{h_1}(x) + g_{h_2}(x)$ gilt, folgt

$$\begin{aligned}
&f''(x, h_1 + h_2, k_1 + k_2) = g'_{h_1+h_2}(x)(k_1 + k_2) \\
&= \Big[g_{h_1}(x) + g_{h_2}(x)\Big]'(k_1 + k_2) = g'_{h_1}(x)(k_1 + k_2) + g'_{h_2}(x)(k_1 + k_2) \\
&= f''(x, h_1, k_1) + f''(x, h_1, k_2) + f''(x, h_2, k_1) + f''(x, h_2, k_2)\,.
\end{aligned}$$

Ebenso folgt

$$\begin{aligned}
f''(x, ch, k) &= cf''(h, k)\,, \\
f''(x, h, ck) &= cf''(h, k)\,.
\end{aligned}$$

Seien $h = (h_1, \ldots, h_n)$ und $k = (k_1, \ldots, k_n)$. Dann gilt

$$f''(x, h, k) = g'_h(x)k = \sum_{j=1}^{n} \frac{\partial}{\partial x_j}\, g_h(x)k_j\,.$$

Wegen

$$g_h(x) = f'(x)h = \sum_{i=1}^{n} \frac{\partial}{\partial x_i}\, f(x)h_i$$

folgt also

$$\begin{aligned}
f''(x, h, k) = g'_h(x)k &= \sum_{j=1}^{n} \frac{\partial}{\partial x_j}\Big(\sum_{i=1}^{n} \frac{\partial}{\partial x_i}\, f(x)h_i\Big)k_j \\
&= \sum_{j=1}^{n}\sum_{i=1}^{n} \frac{\partial}{\partial x_j}\,\frac{\partial}{\partial x_i}\, f(x)h_ik_j\,.
\end{aligned}$$

Hierbei sieht man, daß die zweiten partiellen Ableitungen $\frac{\partial}{\partial x_j}\frac{\partial}{\partial x_i} f(x)$ alle existieren, indem man für h und k die Standardbasisvektoren e_i und e_j wählt. Es gilt

$$\frac{\partial^2 f}{\partial x_j \partial x_i}\,(x) = \frac{\partial}{\partial x_j}\,\frac{\partial}{\partial x_i}\, f(x) = \begin{pmatrix} \frac{\partial}{\partial x_j}\,\frac{\partial}{\partial x_i}\, f_1(x) \\ \vdots \\ \frac{\partial}{\partial x_j}\,\frac{\partial}{\partial x_i}\, f_m(x) \end{pmatrix} \in \mathbb{R}^m\,.$$

Man setzt auch

$$\frac{\partial^2 f}{\partial x_j^2}\,(x) := \frac{\partial^2 f}{\partial x_j \partial x_j}\,(x)\,.$$

Für reellwertiges $f : U \to \mathbb{R}$ erhält man in Matrizenschreibweise

$$\begin{aligned} f''(x,h,k) &= (k_1, \ldots, k_n) \begin{pmatrix} \frac{\partial^2 f}{\partial x_1 \partial x_1}(x) & \ldots & \frac{\partial^2 f}{\partial x_1 \partial x_n}(x) \\ \vdots & & \\ \frac{\partial^2 f}{\partial x_n \partial x_1}(x) & \ldots & \frac{\partial^2 f}{\partial x_n \partial x_n}(x) \end{pmatrix} \begin{pmatrix} h_1 \\ \vdots \\ h_n \end{pmatrix} \\ &= k \cdot Hh\,, \end{aligned}$$

wobei man

$$H = \Big(\frac{\partial^2 f}{\partial x_i \partial x_j}\Big)_{i,j=1,\ldots,n}$$

als die Hessesche Matrix bezeichnet. Für beliebiges $f : U \to \mathbb{R}^m$ erhält man

$$\Big[f''(x,h,k)\Big]_j = k \cdot H_j h\,,$$

wobei H_j die Hessesche Matrix der j–ten Komponentenfunktion f_j ist. Insbesondere folgt hieraus

$$(f'')_j(x,h,k) = (f_j)''(x,h,k)\,,$$

d.h. die j–te Komponente von f'' ist die zweite Ableitung der Komponentenfunktion f_j. Falls $f : U \to \mathbb{R}^m$ differenzierbar ist, und f in $a \in U$ zweimal differenzierbar ist, dann ist H beziehungsweise H_j eine symmetrische Matrix, d. h. es gilt

$$\frac{\partial^2 f_j}{\partial x_i \partial x_k}(a) = \frac{\partial^2 f_j}{\partial x_k \partial x_i}(a)\,.$$

Dies ergibt sich aus dem folgenden Satz. Man beachte aber, daß alle zweiten partiellen Ableitungen in a existieren können, ohne daß f diese Voraussetzungen erfüllt. Dann braucht H nicht symmetrisch zu sein.

Satz von H.A. Schwartz: Sei $U \subseteq \mathbb{R}^n$ offen, $f : U \to \mathbb{R}^m$ sei differenzierbar und in einem Punkt $x \in U$ zweimal differenzierbar. Dann gilt für alle $h, k \in \mathbb{R}^n$

$$f''(x,h,k) = f''(x,k,h)\,.$$

(Die bilineare Abbildung $(h,k) \mapsto f''(x,h,k) : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^m$ ist symmetrisch.)

Beweis: Die Bilinearform $(h,k) \mapsto f''(x,h,k)$ ist symmetrisch, genau dann wenn jede ihrer Komponenten $(h,k) \mapsto (f'')_j(x,h,k) = (f_j)''(x,h,k)$ symmetrisch ist. Es genügt also, die Symmetrie für die Komponentenfunktionen f_j zu beweisen, wobei ich den Index j weglasse und voraussetze, daß $f : U \to \mathbb{R}$ gilt. Zum Beweis des Satzes zeige ich, daß für alle $h, k \in \mathbb{R}^n$

$$\lim_{\substack{s\to 0\\ s>0}} \frac{f(x+sh+sk)-f(x+sh)-f(x+sk)+f(x)}{s^2} = f''(x,h,k) \qquad (*)$$

gilt. Hieraus folgt die Behauptung, weil sich die linke Seite bei Vertauschen von h und k nicht ändert.

$f''(x,h,k)$ ist die Ableitung der Funktion $x \mapsto f'(x,h)$. Also gilt

$$f'(x+k,h)-f'(x,h) = f''(x,h,k)+R_x(h,k)\|k\|$$

mit

$$\lim_{k\to 0} R_x(h,k) = 0\,.$$

$R_x(h,k)$ ist linear bezüglich h, weil $f'(x+k,h)$, $f'(x,h)$ und $f''(x,h,k)$ linear in h sind, und es existiert eine von h und k abhängige Zahl ϑ mit $0<\vartheta<1$, so daß

$$\begin{aligned} &f(x+h+k)-f(x+h)-f(x+k)+f(x)\\ &\quad = f''(x,h,k)+R_x(h,\vartheta h+k)\|\vartheta h+k\|-R_x(h,\vartheta h)\|\vartheta h\| \end{aligned} \qquad (**)$$

gilt. Zum Beweis dieser Gleichung betrachte man die Hilfsfunktion

$$F:[0,1]\to\mathbb{R}\,,\;\; F(t):=f(x+th+k)-f(x+th)\,.$$

Wegen

$$F'(t)=f'(x+th+k)h-f'(x+th)h=f'(x+th+k,h)-f'(x+th,h)$$

und wegen

$$F(1)-F(0)=f(x+h+k)-f(x+h)-f(x+k)+f(x)$$

folgt nach dem Mittelwertsatz

$$F(1)-F(0)=F'(\vartheta)\,,$$

mit geeignetem $\vartheta, 0<\vartheta<1$, also

$$\begin{aligned} &f(x+h+k)-f(x+h)-f(x+k)+f(x)\\ &= f'(x+\vartheta h+k,h)-f'(x+\vartheta h,h)\\ &= \Big(f'(x+\vartheta h+k,h)-f'(x,h)\Big)-\Big(f'(x+\vartheta h,h)-f'(x,h)\Big)\,. \end{aligned}$$

Mit

$$\begin{aligned} f'(x+\vartheta h+k,h)-f'(x,h) &= f''(x,h,\vartheta h+k)+R_x(h,\vartheta h+k)\|\vartheta h+k\|\\ f'(x+\vartheta h,h)-f'(x,h) &= f''(x,h,\vartheta h)+R_x(h,\vartheta h)\|\vartheta h\| \end{aligned}$$

und mit

$$f''(x,h,\vartheta h+k) - f''(x,h,\vartheta h) = f''(x,h,k)$$

folgt $(**)$.

Sei $s > 0$. Ersetzt man in $(**)$ den Vektor k durch sk und den Vektor h durch sh, dann kann man auf der rechten Seite wegen der Bilinearität oder Linearität der Terme den Faktor s^2 herausziehen und erhält

$$\begin{aligned} & f(x+sh+sk) - f(x+sh) - f(x+sk) + f(x) \\ = \;& s^2\Big[f''(x,h,k) + R_x\Big(h, s(\vartheta h+k)\Big)\|\vartheta h+k\| - R_x(h,s\vartheta h)\|\vartheta h\|\Big]\,. \end{aligned}$$

Wegen

$$\lim_{s\to 0} R_x\Big(h, s(\vartheta h+k)\Big) = 0\,, \quad \lim_{s\to 0} R_x(h,s\vartheta h) = 0$$

folgt $(*)$.

Beispiel:

$$f : \mathbb{R}^2 \to \mathbb{R}\,,\;\; f(x_1,x_2) = x_1^2 x_2 + x_1 + x_2^3\,.$$

Die partiellen Ableitungen jeder Ordnung existieren und sind stetig, also ist f zweimal differenzierbar. Es gilt

$$\begin{aligned} \mathrm{grad} f &= \begin{pmatrix} \dfrac{\partial f}{\partial x_1} \\ \dfrac{\partial f}{\partial x_2} \end{pmatrix} = \begin{pmatrix} 2x_1x_2+1 \\ \\ x_1^2+3x_2^2 \end{pmatrix} \\ f''(x) := H &= \begin{pmatrix} \dfrac{\partial^2 f}{\partial x_1^2} & \dfrac{\partial^2 f}{\partial x_2 \partial x_1} \\ \dfrac{\partial^2 f}{\partial x_1 \partial x_2} & \dfrac{\partial^2 f}{\partial x_2^2} \end{pmatrix} = \begin{pmatrix} 2x_2 & 2x_1 \\ \\ 2x_1 & 6x_2 \end{pmatrix}\,. \end{aligned}$$

Höhere Ableitungen: Höhere Ableitungen definiert man *induktiv.* Die p–te Ableitung von $f : U \to \mathbb{R}^m$ ist eine Abbildung

$$f^{(p)} : U \times \underbrace{\mathbb{R}^n \times \ldots \times \mathbb{R}^n}_{p \text{ Faktoren}} \to \mathbb{R}^m\,,$$

die man folgendermaßen aus $f^{(p-1)}$ erhält: Sind $x \in U$ und $h^{(1)}, \ldots, h^{(p)} \in \mathbb{R}^n$, dann ist $f^{(p)}$ definiert durch

$$f^{(p)}(x,h^{(1)},\ldots,h^{(p)}) := \Big[y \mapsto f^{(p-1)}(y,h^{(1)},\ldots,h^{(p-1)})\Big]'\Big|_{y=x}(h^{(p)})\,.$$

$f^{(p)}$ ist linear in den letzten p Argumenten und ist total symmetrisch: Für $1 \leq i \leq j \leq p$ gilt

$$f^{(p)}(x, \ldots, h^{(i)}, \ldots, h^{(j)}, \ldots) = f^{(p)}(x, \ldots, h^{(j)}, \ldots, h^{(i)}, \ldots)\,.$$

Ist $f^{(p)}$ stetig, dann heißt f p–mal stetig differenzierbar. Wenn $f^{(p)}$ für alle $p \in \mathbb{N}$ existiert, heißt f unendlich oft differenzierbar. Wie für f'' sieht man, daß

$$f^{(p)}(x, h^{(1)}, \ldots, h^{(p)}) = \sum_{i_1=1}^{n} \ldots \sum_{i_p=1}^{n} \frac{\partial^p f}{\partial x_{i_1} \ldots \partial x_{i_p}}(x) h_{i_1}^{(1)} \ldots h_{i_p}^{(p)}$$

gilt.

Satz (Taylorformel): Sei $U \subseteq \mathbb{R}^n$ eine offene Menge $f : U \to \mathbb{R}$ sei $(p+1)$–mal differenzierbar und die Verbindungsstrecke der beiden Punkte x und $x+h$ gehöre zu U. Dann existiert ϑ, $0 < \vartheta < 1$, mit

$$f(x+h) = f(x) + f'(x,h) + \frac{1}{2!} f''(x,h,h) + \ldots + \frac{1}{p!} f^{(p)}(x, \underbrace{h, \ldots, h}_{p \text{ mal}}) + R_p(x,h)\,,$$

wobei

$$R_p(x,h) = \frac{1}{(p+1)!} f^{(p+1)}(x + \vartheta h, \underbrace{h, \ldots, h}_{p+1 \text{ mal}})$$

sei.

Beweis: Sei $\gamma : [0,1] \to U$, $\quad \gamma(t) = x + th$. Auf $F = f \circ \gamma : [0,1] \to \mathbb{R}$ wende man den Taylorschen Satz für reelle Funktionen an:

$$F(1) = \sum_{j=0}^{p} \frac{F^{(j)}(0)}{j!} + \frac{1}{(p+1)!} F^{(p+1)}(\vartheta)\,.$$

Wegen

$$\begin{aligned}
F'(t) &= f'\Big(\gamma(t)\Big)\gamma'(t) = f'\Big(\gamma(t), \gamma'(t)\Big) = f'\Big(\gamma(t), h\Big) \\
F''(t) &= f''\Big(\gamma(t), h, \gamma'(t)\Big) = f''\Big(\gamma(t), h, h\Big) \\
&\vdots \\
F^{(p+1)}(t) &= f^{(p+1)}\Big(\gamma(t), \underbrace{h, \ldots, h}_{(p+1) \text{ mal}}\Big)
\end{aligned}$$

folgt hieraus die Behauptung. ∎

Man kann die Taylorformel auch in folgender Form schreiben:

$$\begin{aligned} f(x+h) &= \sum_{j=0}^{p} \frac{1}{j!}\Big[\sum_{i_1=1}^{n} \cdots \sum_{i_j=1}^{n} \frac{\partial^j f(x)}{\partial x_{i_1} \ldots \partial x_{i_j}}\, h_{i_1} \ldots h_{i_j}\Big] \\ &\quad + \frac{1}{(p+1)!} \sum_{i_1=1}^{n} \cdots \sum_{i_{p+1}=1}^{n} \frac{\partial^{p+1} f(x+\vartheta h)}{\partial x_{i_1} \ldots \partial x_{i_{p+1}}}\, h_{i_1} \ldots h_{i_{p+1}}\,. \end{aligned}$$

Zur Abkürzung führt man die folgenden Bezeichnungen ein: Für $\alpha = (\alpha_1, \ldots, \alpha_n) \in \mathbb{N}_0^n$ und für $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$ sei

$$\begin{aligned} |\alpha| &:= \alpha_1 + \ldots + \alpha_n \\ \alpha! &:= \alpha_1! \ldots \alpha_n! \\ x^\alpha &:= x_1^{\alpha_1} \ldots x_n^{\alpha_n}\,, \\ D^\alpha f(a) &:= \frac{\partial^{|\alpha|} f}{\partial^{\alpha_1} x_1 \ldots \partial^{\alpha_n} x_n}\,(a)\,. \end{aligned}$$

Man bezeichnet $\alpha \in \mathbb{N}_0^n$ als Multiindex und $|\alpha|$ als Länge von α. Bei vorgegebenem Multiindex α mit $|\alpha| = j$ gibt es in

$$\sum_{i_1=1}^{n} \cdots \sum_{i_j=1}^{n} \frac{\partial^j f(x)}{\partial x_{i_1} \ldots \partial x_{i_j}}\, h_{i_1} \ldots h_{i_j}$$

$\frac{j!}{\alpha_1! \ldots \alpha_n!}$ Glieder, die aus dem Glied $D^\alpha f(x) h^\alpha$ durch Vertauschen der Reihenfolge, in der die Ableitungen gebildet werden, entstehen. Also folgt

$$\begin{aligned} f(x+h) &= \sum_{j=0}^{p} \sum_{|\alpha|=j} \frac{1}{\alpha!}\, D^\alpha f(x) h^\alpha + \sum_{|\alpha|=p+1} \frac{1}{\alpha!}\, D^\alpha f(x+\vartheta h) h^\alpha \\ &= \sum_{|\alpha| \le p} \frac{1}{\alpha!}\, D^\alpha f(x) h^\alpha + \sum_{|\alpha|=p+1} \frac{1}{\alpha!}\, D^\alpha f(x+\vartheta h) h^\alpha\,. \end{aligned}$$

6 Lokale Extrema, Sätze von der inversen und der impliziten Funktion.

In diesem Kapitel werden die Ergebnisse von Kapitel 5 angewandt.

6 a.) Lokale Extrema

Definition: Sei $U \subseteq \mathbb{R}^n$ offen, sei $f : U \to \mathbb{R}$ differenzierbar und sei $a \in U$. Gilt $\mathrm{grad} f(a) = 0$, dann heißt a kritischer Punkt von f.

Satz: Sei $U \subseteq \mathbb{R}^n$ offen und sei $f : U \to \mathbb{R}$ differenzierbar. Ist a lokale Extremalstelle von f, dann ist a kritischer Punkt von f.

Beweis: O.B.d.A. habe f in a ein Maximum. Dann gibt es eine Umgebung V von a mit $f(x) \le f(a)$ für alle $x \in V$. Sei $h \in \mathbb{R}^n$. Wähle $\delta > 0$ so klein, daß $a + th \in V$ ist für alle $t \in \mathbb{R}$ mit $|t| \le \delta$. Sei $F : [-\delta, \delta] \to \mathbb{R}$ definiert durch

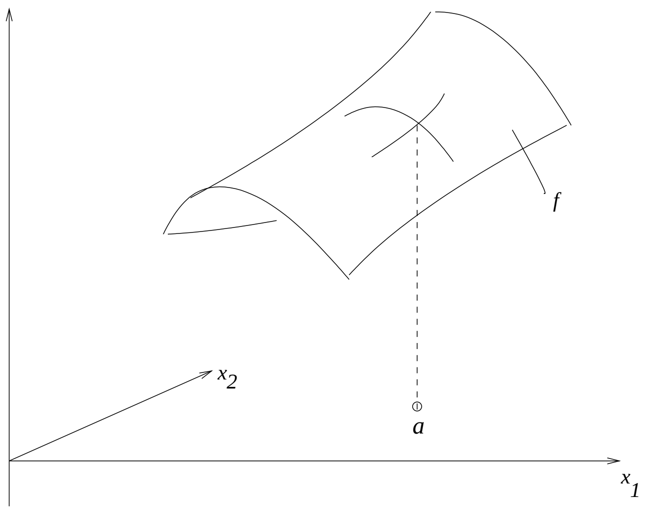
$$F(t) := f(a + th)\,.$$

Dann hat F ein Maximum in $t = 0$, also folgt

$$0 = F'(0) = f'(a)h\,.$$

Weil dies für alle $h \in \mathbb{R}^n$ gilt, resultiert $f'(a) = 0$. ■

Dies ist eine notwendige Bedingung an f für eine lokale Extremalstelle, aber keine hinreichende. Zum Beispiel ist der Sattelpunkt in der folgenden Skizze zwar ein kritischer Punkt, aber keine lokale Extremalstelle:



Wie für reelle Funktionen kann man mit Hilfe der zweiten Ableitung hinreichende Bedingungen erhalten. Hierzu benötigt man Resultate über quadratische Formen, die ich hier ohne Beweis angebe:

Vorbemerkung über quadratische Formen: 1.) Sei $Q : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ eine bilineare Abbildung. Dann heißt die Abbildung $h \mapsto Q(h,h) : \mathbb{R}^n \to \mathbb{R}$ quadratische Form. Man teilt quadratische Formen folgendermaßen ein: Sei

$$\begin{array}{llll}
Q(h,h) > 0 & \text{für alle} \quad h \neq 0 \ : & \text{Dann heißt} \quad Q & \text{positiv definit} \\
Q(h,h) \geq 0 & \text{für alle} \quad h \neq 0 \ : & \text{Dann heißt} \quad Q & \text{positiv semidefinit} \\
Q(h,h) < 0 & \text{für alle} \quad h \neq 0 \ : & \text{Dann heißt} \quad Q & \text{negativ definit} \\
Q(h,h) \leq 0 & \text{für alle} \quad h \neq 0 \ : & \text{Dann heißt} \quad Q & \text{negativ semidefinit.}
\end{array}$$

Q heißt indefinit, wenn Q sowohl positive wie negative Werte annimmt.

2.) Eine quadratische Form kann man immer in der Form

$$Q(h,h) = \sum_{i,j=1}^{n} c_{ij} h_i h_j = h \cdot Ch$$

darstellen mit einer symmetrischen Koeffizientenmatrix

$$C = \begin{pmatrix} c_{11} & \dots & c_{1n} \\ \vdots & & \\ c_{n1} & \dots & c_{nn} \end{pmatrix}, \quad c_{ij} = c_{ji}\,.$$

3.) Ein Kriterium dafür, daß Q positiv definit ist, ist

$$c_{11} > 0\,, \ \det \begin{pmatrix} c_{11} & c_{12} \\ c_{21}, & c_{22} \end{pmatrix} > 0\,, \quad \det \begin{pmatrix} c_{11} & c_{12} & c_{13} \\ c_{21} & c_{22} & c_{23} \\ c_{31} & c_{32} & c_{33} \end{pmatrix} > 0\,, \ \dots\,, \det(c_{ij})_{i,j=1,\dots,n} > 0\,.$$

4.) Für eine in $a \in U$ zweimal differenzierbare Abbildung $f : U \to \mathbb{R}$ ist $h \to f''(a,h,h)$ eine quadratische Form. Wegen

$$f''(a,h,h) = \sum_{i,j=1}^{n} \frac{\partial^2}{\partial x_i \partial x_j} f(a) h_i h_j$$

ist die Koeffizientenmatrix dieser quadratischen Form die Hessesche Matrix

$$H = \Big(\frac{\partial^2}{\partial x_i \partial x_j} f(a) \Big)_{i,j=1,\dots,n}.$$

Mit diesen Definitionen und Resultaten für quadratische Formen kann ein hinreichendes Kriterium für Extremalstellen formuliert werden:

Satz: Sei $U \subseteq \mathbb{R}^n$ offen, $f : U \to \mathbb{R}$ sei zweimal stetig differenzierbar und sei $a \in U$ kritischer Punkt von f . Ist dann die quadratische Form $f''(a, h, h)$

(i) positiv definit, so ist a Minimalstelle von f

(ii) negativ definit, so ist a Maximalstelle von f

(iii) indefinit, so ist a keine Extremalstelle von f .

Beweis: Aus der Taylorformel ergibt sich

$$f(x) = f(a) + f'(a, x-a) + \frac{1}{2} f''\Big(a + \vartheta(x-a), x-a, x-a\Big),$$

mit geeignetem $0 < \vartheta < 1$, also, wegen $f'(a) = 0$,

$$\begin{aligned} f(x) &= f(a) + \frac{1}{2} f''\Big(a + \vartheta(x-a), x-a, x-a\Big) \\ &= f(a) + \frac{1}{2} f''(a, x-a, x-a) + R(x, x-a, x-a), \end{aligned} \qquad (*)$$

mit

$$R(x, h, k) = \frac{1}{2} f''\Big(a + \vartheta(x-a), h, k\Big) - \frac{1}{2} f''(a, h, k\Big).$$

Es ist $\mathbb{R}(a, h, k) = 0$, und es gilt sogar, daß zu jedem $\varepsilon > 0$ ein $\delta > 0$ existiert mit

$$|R(x, h, h)| \le \varepsilon \|h\|^2, \qquad (**)$$

für alle $x \in U$ mit $\|x - a\| < \delta$ und alle $h \in \mathbb{R}^n$. Zum Beweis wähle man $r > 0$ so klein, daß die abgeschlossene Kugel $K_r(a) = \{x \in \mathbb{R}^n \,\big|\, \|x - a\| \le r\}$ ganz zu U gehört. Nach Voraussetzung ist $(x, h) \mapsto f''(x, h, h) : U \times \mathbb{R}^n \to \mathbb{R}$ stetig, also ist diese Funktion auf der abgeschlossenen und beschränkten, folglich kompakten Teilmenge

$$K_r(a) \times \{h \in \mathbb{R}^n \,\big|\, \|h\| = 1\}$$

sogar gleichmäßig stetig. Dies bedeutet, daß zu $\varepsilon > 0$ eine Zahl $\delta > 0$ existiert mit

$$\frac{1}{2} \Big| f''(x, h, h) - f''(a, h, h) \Big| < \varepsilon$$

für alle $h \in \mathbb{R}^n$ mit $\|h\| = 1$ und alle $x \in U$ mit $\|x - a\| < \delta$. Weil für $0 < \vartheta < 1$ auch $\|a + \vartheta(x-a) - a\| = \vartheta\|x - a\| < \delta$ gilt, folgt

$$\begin{aligned} |R(x, h, h)| &= \|h\|^2 \Big| R\Big(x, \frac{h}{\|h\|}, \frac{h}{\|h\|}\Big) \Big| \\ &= \|h\|^2 \frac{1}{2} \Big| f''\Big(a + \vartheta(x-a), \frac{h}{\|h\|}, \frac{h}{\|h\|}\Big) - f''\Big(a, \frac{h}{\|h\|}, \frac{h}{\|h\|}\Big) \Big| < \varepsilon \|h\|^2. \end{aligned}$$

Dies beweist $(**)$.

Sei nun $h \mapsto f''(a,h,h)$ eine positiv definite quadratische Form. Dann gilt $f''(a,h,h) > 0$ für alle $h \in \mathbb{R}^n$ mit $h \neq 0$, und da die stetige Abbildung $h \mapsto f''(a,h,h) : \mathbb{R}^n \to \mathbb{R}$ auf der abgeschlossenen und beschränkten, also kompakten Menge $\{h \in \mathbb{R}^n \,\big|\, \|h\| = 1\}$ ihr Minimum an einer Stelle h_0 annimmt, folgt für alle $h \in \mathbb{R}^n$

$$f''(a,h,h) = \|h\|^2 f''\Big(a, \frac{h}{\|h\|}, \frac{h}{\|h\|}\Big) \geq \|h\|^2 \min_{\|\eta\|=1} f''(a,\eta,\eta) = c\|h\|^2 ,$$

mit

$$c = \min_{\|\eta\|=1} f''(a,\eta,\eta) = f''(a,h_0,h_0) > 0 .$$

Wählt man nun $\varepsilon = c/4$, dann folgt hieraus und aus $(*), (**)$, daß $\delta > 0$ existiert mit

$$\begin{aligned} f(x) - f(a) &= \frac{1}{2} f''(a, x-a, x-a) + R(x, x-a, x-a) \\ &\geq \frac{c}{2} \|x-a\|^2 - \frac{c}{4} \|x-a\|^2 = \frac{c}{4} \|x-a\|^2 \geq 0 , \end{aligned}$$

für alle x mit $\|x-a\| < \delta$, also ist a ein lokales Minimum.

Entsprechend beweist man, daß bei negativ definitem $f''(a,h,h)$ ein lokales Maximum vorliegt. Wenn $f''(a,h,h)$ indefinit ist, gibt es $h_0 \in \mathbb{R}^n$, $k_0 \in \mathbb{R}^n$ mit $\|h_0\| = \|k_0\| = 1$ und mit

$$f''(a,h_0,h_0) > 0 , \quad f''(a,k_0,k_0) < 0 .$$

Hieraus folgt, daß auf der Geraden mit Richtungsvektor h_0 bzw. k_0 für genügend kleines $\|x-a\|, x \neq a$ die Differenz $f(x) - f(a)$ positiv bzw. negativ ist. Dies beweist man wie oben. Also ist a kein lokales Extremum. ∎

Beispiel: Sei $f : \mathbb{R}^2 \to \mathbb{R}$ definiert durch $f(x,y) = 6xy - 3y^2 - 2x^3$. Für jeden kritischen Punkt (x,y) gilt

$$f'(x,y) = \begin{pmatrix} \dfrac{\partial f}{\partial x}(x,y) \\[2ex] \dfrac{\partial f}{\partial y}(x,y) \end{pmatrix} = \begin{pmatrix} 6y - 6x^2 \\ 6x - 6y \end{pmatrix} = 0 .$$

Hieraus können die kritischen Punkte bestimmt werden. Man erhält für die kritischen Punkte $(x,y) = (0,0)$ und $(x,y) = (1,1)$.

Um festzustellen, ob diese Punkte Extremalpunkte sind, muß die Hessesche Matrix

$$H(x,y) = \begin{pmatrix} -12x & 6 \\ 6 & -6 \end{pmatrix}$$

von f an den kritischen Punkten untersucht werden. Die durch die Matrix

$$H(0,0) = \begin{pmatrix} 0 & 6 \\ 6 & -6 \end{pmatrix}$$

definierte quadratische Form $f''(0,0,h,h)$ ist indefinit. Denn es gilt für $h = (1,1)$

$$f''(0,0,h,h) = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \cdot \begin{pmatrix} 0 & 6 \\ 6 & -6 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \cdot \begin{pmatrix} 6 \\ 0 \end{pmatrix} = 6$$

und für $h = (0,1)$

$$f''(0,0,h,h) = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \cdot \begin{pmatrix} 0 & 6 \\ 6 & -6 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \cdot \begin{pmatrix} 6 \\ -6 \end{pmatrix} = -6\,,$$

also ist $(0,0)$ keine Extremalstelle. Dagegen ist die durch die Matrix

$$H(1,1) = \begin{pmatrix} -12 & 6 \\ 6 & -6 \end{pmatrix}$$

definierte quadratische Form $f(1,1,h,h)$ negativ definit. Denn nach dem oben angegebenen Kriterium ist die Matrix $-H(1,1)$ positiv definit wegen $12 > 0$ und

$$\det \begin{pmatrix} 12 & -6 \\ -6 & 6 \end{pmatrix} = 72 - 36 > 0\,.$$

Somit ist $H(1,1)$ negativ definit und $(1,1)$ ein lokales Maximum.

6 b.) Lokale Umkehrbarkeit von Abbildungen

Früher wurde gezeigt, daß wenn f invertierbar und in einem Punkt a differenzierbar ist, wenn außerdem $\det f'(a) \neq 0$ gilt und die Inverse g in $b = f(a)$ stetig ist, dann ist g in b differenzierbar. Man kann sich fragen, ob aus $\det f'(a) \neq 0$ bereits folgt, daß f in einer Umgebung von a invertierbar ist. Das folgende Beispiel zeigt, daß dies im Allgemeinen nicht richtig ist.

Gegenbeispiel: Sei $f : (-1,1) \to \mathbb{R}$ definiert durch

$$f(x) = \begin{cases} x + 3x^2 \sin \dfrac{1}{x}\,, & x \neq 0 \\ 0\,, & x = 0\,. \end{cases}$$

f ist für alle $|x| < 1$ differenzierbar, $f'(x)$ ist beschränkt, $f'(0) = 1$, aber f ist in keiner Umgebung des Nullpunktes invertierbar.

Jedoch ist in diesem Beispiel f' nicht stetig im Nullpunkt, weil der Grenzwert von

$$f'(x) = 1 + 6x \sin \frac{1}{x} - 3 \cos \frac{1}{x}$$

an der Stelle 0 nicht existiert. Setzt man auch noch die Stetigkeit der Ableitung voraus, dann kann man folgern, daß eine lokale Inverse existiert:

Satz: Sei $U \subseteq \mathbb{R}^n$ offen und $a \in U$, $f : U \to \mathbb{R}^n$ sei stetig differenzierbar, und es sei $\det f'(a) \neq 0$. Sei $b = f(a)$. Dann existieren offene Mengen $V, W \subseteq \mathbb{R}^n$ mit $a \in V$, $b \in W$, so daß $f : V \to W$ bijektiv ist, und so daß die Inverse $g : W \to V$ stetig differenzierbar ist. (Natürlich gilt dann $g'(y) = [f'(g(y))]^{-1}$.)

Beweis: Setze $A := f'(a)$ und $\lambda := \frac{1}{4\|A^{-1}\|}$. Die Inverse A^{-1} existiert, weil nach Voraussetzung $\det A \neq 0$ ist. Da nach Vorasusetzung $f' : U \mapsto L(\mathbb{R}^n, \mathbb{R}^n)$ stetig ist, existiert eine offene Kugel V um a mit

$$\|f'(x) - A\| < 2\lambda$$

für alle $x \in V$.

1.) Zunächst soll gezeigt werden: Für beliebige $x, x + h \in V$ gilt

$$\|f(x+h) - f(x) - Ah\| \leq \frac{1}{2} \|Ah\|, \qquad (*)$$

also, wegen der umgekehrten Dreiecksungleichung,

$$\begin{aligned} \|f(x+h) - f(x)\| &\geq \|Ah\| - \|f(x+h) - f(x) - Ah\| \geq \frac{1}{2}\|Ah\| \\ &= 2\lambda\|A^{-1}\| \, \|Ah\| \geq 2\lambda\|A^{-1}Ah\| = 2\lambda\|h\|, \end{aligned} \qquad (**)$$

woraus dann folgt, daß f in V injektiv ist.

Hierzu definiere man $F : [0, 1] \to \mathbb{R}^n$ durch

$$F(t) = f(x + th) - tAh.$$

Da V eine Kugel ist, gehört mit x und $x + h$ auch die Verbindungsstrecke $\{x + th \,\big|\, 0 \leq t \leq 1\}$ zu V, und es gilt

$$\begin{aligned} \|F'(t)\| &= \|f'(x+th)h - Ah\| \leq \|f'(x+th) - A\| \, \|h\| \leq 2\lambda\|h\| \\ &= 2\lambda\|A^{-1}Ah\| \leq 2\lambda\|A^{-1}\| \, \|Ah\| = \frac{1}{2} \|Ah\|. \end{aligned}$$

Aus dem Schrankensatz folgt nun $\|F(1) - F(0)\| \leq \frac{1}{2} \|Ah\|$, also $(*)$.

Somit existiert die Inverse $g : W \to V$ mit $W = f(V)$.

2.) Es ist zu zeigen, daß W offen ist. Sei $x_0 \in V$ und sei K_r eine offene Kugel mit Mittelpunkt x_0 und Radius $r > 0$, so daß $\overline{K}_r \subseteq V$ ist. Es soll gezeigt werden, daß $f(K_r)$ eine offene Kugel um $f(x_0)$ mit Radius λr enthält.

Hierzu wähle $y \in \mathbb{R}^n$ mit $\|y - f(x_0)\| < \lambda r$. Es wird das Urbild von y unter f in K_r konstruiert. Wähle $x^* \in \overline{K}_r$ mit

$$\|y - f(x^*)\| = \min_{x \in \overline{K}_r} \|y - f(x)\| .$$

Es wird sich zeigen, daß x^* das Urbild ist. Zunächst muß gezeigt werden, daß ein solches x^* existiert. Hierzu beachte man, daß die durch $\phi(x) := \|y - f(x)\|$ definierte Funktion $\phi : \overline{K}_r \to \mathbb{R}$ stetig ist. Denn es gilt für $z \in \overline{K}_r$

$$\begin{aligned}\lim_{x \to z} |\phi(x) - \phi(z)| &= \lim_{x \to z} \Big| \|y - f(x)\| - \|y - f(z)\| \Big| \\ &\leq \lim_{x \to z} \Big\| \Big(y - f(x)\Big) - \Big(y - f(z)\Big) \Big\| \leq \lim_{x \to z} \|f(x) - f(z)\| = 0 ,\end{aligned}$$

wobei wieder die umgekehrte Dreiecksungleichung verwendet wurde. Also nimmt ϕ auf der kompakten Menge $\overline{K}_r$ das Minimum in mindestens einem Punkt x^* an. Es soll nun gezeigt werden, daß $\|y - f(x^*)\| = \phi(x^*) = 0$ ist. Hierzu zeigt man, daß ein $\tilde{x} \in K_r$ existieren würde mit $\|y - f(\tilde{x})\| < \|y - f(x^*)\|$, falls $\|y - f(x^*)\| \neq 0$ wäre. Dies ist ein Widerspruch zur Definition von x^*.

Sei $h = A^{-1}(y - f(x^*))$. Für hinreichend kleines $t \in (0,1)$ ist

$$\tilde{x} = x^* + th \in V .$$

Es gilt nun nach $(*)$ wegen $x^*, x^* + th \in V$:

$$\begin{aligned}&\|f(x^* + th) - y\| = \|f(x^* + th) - f(x^*) - Ath + f(x^*) - y + Ath\| \\ &\leq \|f(x^* + th) - f(x^*) - Ath\| + \|f(x^*) - y + Ath\| \\ &\leq \frac{1}{2} \|Ath\| + \|f(x^*) - y + Ath\| \\ &= \frac{1}{2} \Big\|A\Big[tA^{-1}\Big(y - f(x^*)\Big)\Big]\Big\| + \Big\|f(x^*) - y + A\Big[tA^{-1}\Big(y - f(x^*)\Big)\Big]\Big\| \\ &= \frac{1}{2} t\|y - f(x^*)\| + \Big\|(1-t)\Big(f(x^*) - y\Big)\Big\| = \Big(1 - \frac{t}{2}\Big)\|y - f(x^*)\| < \|y - f(x^*)\| ,\end{aligned}$$

falls $y \neq f(x^*)$. Weil diese Ungleichung für alle $t \in (0,1)$ mit $x^* + th \in V$ gilt, bleibt nur noch zu zeigen, daß $\tilde{x} = x^* + th \in \overline{K}_r$ ist für alle hinreichend kleines t. Dann ist der Widerspruch konstruiert. Hierzu genügt es zu zeigen, daß x^* nicht auf dem Rand der Kugel K_r liegt. Für einen Randpunkt x von K_r gilt $\|x - x_0\| = r$. Aus $(**)$ folgt somit

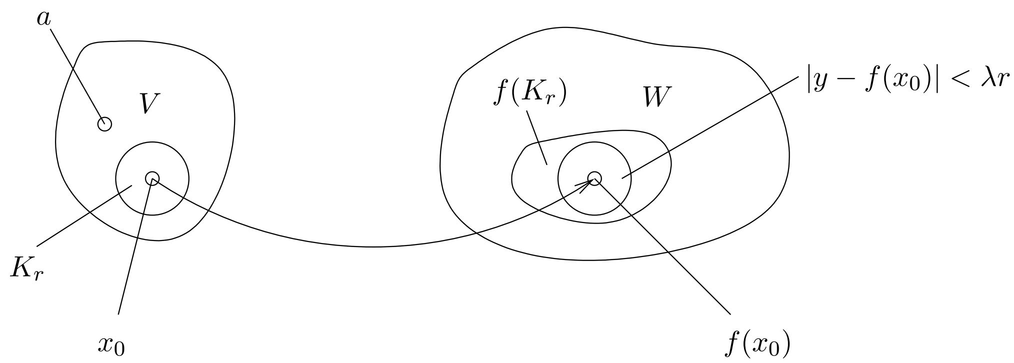
$$2\lambda r \leq \|f(x) - f(x_0)\| \leq \|y - f(x)\| + \|y - f(x_0)\| < \phi(x) + \lambda r ,$$

also

$$\phi(x_0) = \|y - f(x_0)\| < \lambda r < \phi(x)\,,$$

so daß ϕ in keinem Randpunkt das Minimum annehmen kann. Also ist x^* innerer Punkt von K_r .

Damit ist bewiesen, daß $y = f(x^*)$ gilt, somit gehören alle $y \in \mathbb{R}^n$ mit $\|y - f(x_0)\| < \lambda r$ zu $f(V)$, also ist $W = f(V)$ offen, weil $x_0 \in V$ beliebig gewählt war.



3.) Es bleibt zu zeigen, daß die Inverse $g : W \to V$ stetig differenzierbar ist. Da f stetig differenzierbar ist, sind die partiellen Ableitungen $\frac{\partial f_i}{\partial x_j} : V \to \mathbb{R}$ stetig. Da $\det f'(x) = \det(\frac{\partial f_i}{\partial x_j})_{i,j=1,\ldots,n}$ eine Summe aus Produkten der $\frac{\partial f_i}{\partial x_j}(x)$ ist, ist $x \mapsto \det f'(x) : V \to \mathbb{R}$ stetig, also gibt es eine Umgebung von a in der $\det f'(x) \neq 0$ ist, in der also $f'(x)$ invertierbar ist. Man verkleinere die Umgebung V soweit, daß $f'(x)$ invertierbar ist für alle $x \in V$. Nach einem Satz aus Abschnitt 5 c.) folgt dann, daß die Inverse $g : W \to V$ in jedem Punkt von W differenzierbar ist, wenn sie stetig ist. Die Stetigkeit von g folgt unmittelbar aus $(**)$. Denn es gilt für $y, y + k \in W$:

$$\|k\| = \|y + k - y\| = \|f\Big(g(y+k)\Big) - f\Big(g(y)\Big)\| \geq 2\lambda \|g(y+k) - g(y)\|\,,$$

also

$$\lim_{k\to 0} \|g(y+k) - g(y)\| \leq \frac{1}{2\lambda} \lim_{k\to 0} \|k\| = 0\,.$$

Somit ist g stetig, also differenzierbar mit $g'(y) = [f'(g(y))]^{-1}$.

Aus dieser Formel folgt auch, daß g' stetig ist. Denn in Abschnitt 5 e.) wurde gezeigt, daß g' stetig ist, wenn die Elemente der Matrix $g'(y)$, also die partiellen Ableitungen von g , stetige Funktionen von y sind. Weil $g'(y)$ die Inverse der Matrix $f'(g(y))$ ist, werden die Elemente von $g'(y)$ aus den Elementen von $f'(g(y))$ durch Bildung von Determinanten und Quotienten berechnet (Cramersche Regel!), also sind die Elemente von $g'(y)$ stetige Funktionen der Elemente der Matrix $f'(g(y))$, die selber wieder stetige Funktionen von y

sind, weil f' und g stetig sind. Also ist g stetig differenzierbar. ■

Beispiel: $f : \mathbb{R}^3 \to \mathbb{R}^3$ sei erklärt durch

$$\begin{aligned} f_1(x_1, x_2, x_3) &= x_1 + x_2 + x_3 \\ f_2(x_1, x_2, x_3) &= x_2x_3 + x_3x_1 + x_1x_2 \\ f_3(x_1, x_2, x_3) &= x_1x_2x_3 \,. \end{aligned}$$

Weil die partiellen Ableitungen alle existieren und stetig sind, ist f stetig differenzierbar mit

$$f'(x) = \begin{pmatrix} 1 & 1 & 1 \\ x_3 + x_2 & x_3 + x_1 & x_2 + x_1 \\ x_2x_3 & x_1x_3 & x_1x_2 \end{pmatrix},$$

also

$$\begin{aligned} \det f'(x) &= \begin{vmatrix} 1 & 0 & 0 \\ x_3 + x_2 & x_1 - x_2 & x_1 - x_3 \\ x_2x_3 & (x_1 - x_2)x_3 & (x_1 - x_3)x_2 \end{vmatrix} \\ &= (x_1 - x_2)(x_1 - x_3)x_2 - (x_1 - x_2)(x_1 - x_3)x_3 \\ &= (x_1 - x_2)(x_1 - x_3)(x_2 - x_3)\,. \end{aligned}$$

Sei also $b = f(a)$ mit $(a_1 - a_2)(a_1 - a_3)(a_2 - a_3) \neq 0$. Dann gilt es Umgebungen V von a und W von b, so daß das Gleichungssystem

$$\begin{aligned} y_1 &= x_1 + x_2 + x_3 \\ y_2 &= x_2x_3 + x_3x_1 + x_1x_2 \\ y_3 &= x_1x_2x_3 \end{aligned}$$

zu jedem $y \in W$ eine eindeutige Lösung $x \in V$ hat.

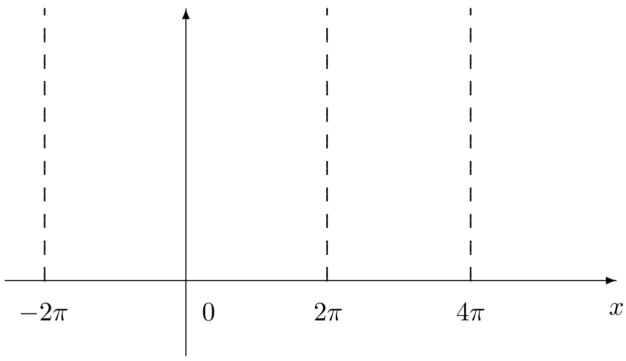
Man beachte aber, daß aus der lokalen Invertierbarkeit nicht die globale folgt. Man sieht dies an folgendem Beispiel: Sei $f : \{(x, y) \in \mathbb{R}^2 \,\big|\, y > 0\} \to \mathbb{R}^2$ definiert durch

$$\begin{aligned} f_1(x, y) &= y \cos x \\ f_2(x, y) &= y \sin x\,. \end{aligned}$$

f ist stetig differenzierbar mit

$$\det f'(x, y) = \begin{vmatrix} -y \sin x & \cos x \\ y \cos x & \sin x \end{vmatrix} = -y \sin^2 x - y \cos^2 x = -y \neq 0$$

für alle (x, y) aus dem Definitionsbereich. Also ist f in jedem Punkt lokal invertierbar, jedoch nicht global. Denn sei $b = f(a)$ mit $a = (a_1, a_2)$. Dann gilt auch $b = f(a_1 + 2\pi m, a_2)$, $\quad m \in \mathbb{Z}$, weil f bezüglich der x–Koordinate 2π–periodisch ist.



6 c.) Implizite Funktionen

Es sei eine Abbildung $f : \mathbb{R}^{n+m} \to \mathbb{R}^n$ gegeben mit den Komponenten f_j, also $f = (f_1, \ldots, f_n)$, und es sei $y = (y_1, \ldots, y_m) \in \mathbb{R}^m$ gegeben. Es liegt nahe zu fragen, ob $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$ so bestimmt werden kann, daß

$$\begin{aligned} f_1(x_1, \ldots, x_n, y_1, \ldots, y_m) &= 0 \\ \vdots \\ f_n(x_1, \ldots, x_n, y_1 \ldots, y_m) &= 0 \end{aligned}$$

gilt. Dies sind n Gleichungen zur Bestimmung von n Unbekannten $x_1, \ldots, x_n$. Zunächst betrachte man den Fall, daß $f = A : R^{n+m} \to \mathbb{R}^n$ eine lineare Abbildung ist,

$$A(x,y) = \begin{pmatrix} A_1(x,y) \\ \vdots \\ A_n(x,y) \end{pmatrix} = \begin{pmatrix} a_{11}x_1 + \ldots + a_{1n}x_n + b_{11}y_1 + \cdots + b_{1m}y_m \\ \vdots \\ a_{n1}x_1 + \ldots + a_{nn}x_n + b_{n1}y_1 + \ldots + b_{nm}y_m \end{pmatrix}.$$

A habe folgende Eigenschaft:

$$A(h,0) = 0 \implies h = 0\,.$$

Dies ist genau dann erfüllt, wenn die Matrix

$$\begin{pmatrix} a_{11} \ldots a_{1n} \\ \\ \vdots \\ \\ a_{n1} \ldots a_{nn} \end{pmatrix} = \begin{pmatrix} \frac{\partial A_1}{\partial x_1} \cdots \frac{\partial A_1}{\partial x_n} \\ \vdots \\ \frac{\partial A_n}{\partial x_1} \cdots \frac{\partial A_n}{\partial x_n} \end{pmatrix}$$

invertierbar ist, also genau dann wenn

$$\det\left(\frac{\partial A_j}{\partial x_i}\right)_{\substack{i=1,\,\dots,n\\ j=1,\,\dots,n}} \neq 0$$

ist. Unter dieser Bedingung ist

$$h \mapsto Ch := A(h,0) : \mathbb{R}^n \to \mathbb{R}^n$$

eine invertierbare lineare Abbildung, folglich hat das Gleichungssystem

$$A(h,k) = A(h,0) + A(0,k) = Ch + A(0,k) = 0$$

für jedes $k \in \mathbb{R}^m$ die eindeutig bestimmte Lösung

$$h = \varphi(k) := -C^{-1}A(0,k)\,.$$

Für $\varphi : \mathbb{R}^m \to \mathbb{R}^n$ gilt

$$A(\varphi(k),k) = 0\,,$$

für alle $k \in \mathbb{R}^m$. Man sagt, φ sei durch diese Gleichung implizit gegeben. Der Satz über implizit gegebene Funktionen betrifft dieselbe Situation für stetig differenzierbare Abbildungen f, die nicht notwendig linear sein müssen:

Satz (über implizite Funktionen): Sei $D \subseteq \mathbb{R}^{n+m}$ offen, $f : D \to \mathbb{R}^n$ sei stetig differenzierbar. Sei $(a,b) \in \mathbb{R}^n \times \mathbb{R}^m$ mit $f(a,b) = 0$ und mit

$$\det\begin{pmatrix} \frac{\partial f_1}{\partial x_1}(a,b) & \dots & \frac{\partial f_1}{\partial x_n}(a,b) \\ \vdots & & \\ \frac{\partial f_n}{\partial x_1}(a,b) & \dots & \frac{\partial f_n}{\partial x_n}(a,b) \end{pmatrix} \neq 0\,. \qquad (*)$$

Dann gibt es eine Umgebung $U \subseteq \mathbb{R}^m$ von b und eine eindeutig bestimmte stetig differenzierbare Abbildung $\varphi : U \to \mathbb{R}^n$ mit $\varphi(b) = a$ und mit

$$f\Big(\varphi(y), y\Big) = 0$$

für alle $y \in U$.

Bemerkung: Sei $A = f'(a,b) : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$. Die Bedingung $(*)$ ist äquivalent zur Bedingung

$$A(h,0) = 0 \Longrightarrow h = 0\,.$$

Beweis des Satzes: Betrachte die Abbildung $F : D \to \mathbb{R}^{n+m}$,

$$(x, y) \mapsto F(x, y) := \Big(f(x, y), y\Big) \in \mathbb{R}^{n+m} .$$

Es gilt $F(a, b) = (0, b)$. Es soll gezeigt werden, daß F die Voraussetzungen des Satzes über lokale Umkehrbarkeit erfüllt. Aus diesem Satz folgt dann, daß es Umgebungen $V \subseteq \mathbb{R}^{n+m}$ von (a, b) und $W \subseteq \mathbb{R}^{n+m}$ von $(0, b)$ gibt, so daß $F : V \to W$ bijektiv ist und eine stetig differenzierbare Inverse $F^{-1} : W \to V$ besitzt. Die Inverse ist von der Form

$$F^{-1}(z, w) = \Big(\phi(z, w), w\Big) ,$$

mit einer stetig differenzierbaren Funktion $\phi : W \to \mathbb{R}^n$. Aus W und ϕ erhält man die gesuchte Umgebung $U \subseteq \mathbb{R}^m$ von b und die gesuchte Funktion $\varphi : U \to \mathbb{R}^n$ durch die Definitionen

$$U = \{w \in \mathbb{R}^m \mid (0, w) \in W\}$$

und

$$\varphi(w) := \phi(0, w) , \quad w \in U .$$

Denn wegen $(0, b) \in W$ ist U eine Umgebung von b in $\mathbb{R}^m$, und für alle $w \in U$ gilt

$$(0, w) = F\Big(F^{-1}(0, w)\Big) = F\Big(\phi(0, w), w\Big) = F\Big(\varphi(w), w\Big) = \Big(f\Big(\varphi(w), w\Big), w\Big) ,$$

also

$$f\Big(\varphi(w), w\Big) = 0 .$$

Also genügt es, die Voraussetzungen des Satzes über lokale Umkehrbarkeit nachzuprüfen. Weil f nach Voraussetzung stetig differenzierbar ist, folgt aus der Definition von F sofort, daß alle partiellen Ableitungen von F existieren und stetig sind. Also ist F stetig differenzierbar. Der Satz über lokale Umkehrbarkeit kann somit angewandt werden, wenn $F'[a, b]$ invertierbar ist. Mit $A = f'(a, b)$ gilt für $(h, k) \in \mathbb{R}^{n+m}$

$$(h, k) \mapsto F'[a, b](h, k) = \Big(A(h, k), k\Big) \in \mathbb{R}^{n+m} . \tag{$*$}$$

Denn da f differenzierbar ist, folgt

$$f(a + h, b + k) = f(a, b) + A(h, k) + r(h, k)\|(h, k)\| ,$$

mit $\lim_{(h,k)\to 0} r(h, k) = 0$. Also gilt

$$\begin{aligned} F(a + h, b + k) &= \Big(f(a + h, b + k) ,\ \ b + k\Big) \\ &= \Big(f(a, b), b\Big) + \Big(A(h, k), k\Big) + \Big(r(h, k)\|(h, k)\|, 0\Big) \\ &= F(a, b) + \Big(A(h, k), k\Big) + \Big(r(h, k), 0\Big)\|(h, k)\| , \end{aligned}$$

mit $\lim_{(h,k)\to 0}\Big(r(h,k),0\Big) = 0$. Dies beweis $(*)$. Hieraus folgt, daß $F'[a,b]$ invertierbar ist. Denn aus

$$F'[a,b](h,k) = \Big(A(h,k),k\Big) = (0,0)$$

resultiert $k=0$, also $A(h,0)=0$, somit $h=0$. Also besteht der Nullraum der linearen Abbildung $F'[a,b]$ nur aus der Menge $\{0\}$, also ist die Abbildung invertierbar. ■

Man kann auch die Ableitung der Funktion φ berechnen. Nach der Kettenregel gilt für die Ableitung $\frac{d}{dy}f(\varphi(y),y)$ der Funktion $y \mapsto f(\varphi(y),y)$:

$$\begin{aligned} 0 = \frac{d}{dy} f\Big(\varphi(y),y\Big) &= \Big(\frac{\partial}{\partial x} f\Big(\varphi(y),y\Big), \frac{\partial}{\partial y} f\Big(\varphi(y),y\Big)\Big) \begin{pmatrix} \varphi'(y) \\ I_{m\times m} \end{pmatrix} \\ &= \frac{\partial}{\partial x} f\Big(\varphi(y),y\Big) \circ \varphi'(y) + \frac{\partial}{\partial y} f\Big(\varphi(y),y\Big) \end{aligned}$$

mit der Einheitsmatrix $I_{m\times m}$ auf $\mathbb{R}^m$. Hieraus folgt

$$\varphi'(y) = -\Big[\frac{\partial}{\partial x} f\Big(\varphi(y),y\Big)\Big]^{-1} \circ \frac{\partial}{\partial y} f\Big(\varphi(y),y\Big),$$

mit

$$\begin{aligned} \frac{\partial}{\partial x} f(x,y) &= \Big(\frac{\partial f_j}{\partial x_i}(x,y)\Big)_{j=1,\dots,n;\, i=1,\dots,n} \\ \frac{\partial}{\partial y} f(x,y) &= \Big(\frac{\partial f_j}{\partial y_i}(x,y)\Big)_{j=1,\dots,n;\, i=1,\dots,m} \end{aligned}$$

Beispiele 1.) Sei eine Gleichung

$$f(x_1, \dots, x_n) = 0$$

gegeben mit stetig differenzierbarem $f : \mathbb{R}^n \to \mathbb{R}$. Zu gegebenen $x_1, \dots, x_{n-1}$ ist x_n gesucht, so daß diese Gleichung erfüllt ist. Angenommen, es existiere $a = (a_1, \dots, a_n) \in \mathbb{R}^n$ mit

$$f(a_1, \dots, a_n) = 0$$

und mit

$$\frac{\partial f}{\partial x_n}(a_1, \dots, a_n) \neq 0.$$

Dann existiert eine Umgebung $U \subseteq \mathbb{R}^{n-1}$ von $(a_1, \dots, a_{n-1})$, so daß zu jedem $(x_1, \dots, x_{n-1}) \in U$ ein eindeutiges $x_n = \varphi(x_1, \dots, x_{n-1})$ aus einer Umgebung von a_n existiert mit

$$f(x_1 \dots, x_{n-1}, x_n) = 0.$$

Für die Ableitung von φ gilt

$$\begin{aligned}
\mathrm{grad}\varphi(x_1,\,\ldots,x_{n-1}) &= \frac{-1}{\frac{\partial}{\partial x_n} f(x_1,\,\ldots,x_n)}\,\mathrm{grad}_{n-1} f(x_1,\,\ldots,x_n) \\
&= \frac{-1}{\frac{\partial}{\partial x_n} f} \begin{pmatrix} \dfrac{\partial}{\partial x_1} f \\ \vdots \\ \dfrac{\partial}{\partial x_{n-1}} f \end{pmatrix} .
\end{aligned}$$

2.) $F : \mathbb{R}^3 \to \mathbb{R}^2$ sei definiert durch

$$\begin{aligned}
f_1(x,y,z) &= 3x^2 + xy - z - 3 \\
f_2(x,y,z) &= 2xz + y^3 + xy\,.
\end{aligned}$$

Es gilt $f(1,0,0) = 0$. Zu gegebenen $z \in \mathbb{R}$ aus einer Umgebung des Nullpunktes ist $(x,y) \in \mathbb{R}^2$ in einer Umgebung von $(1,0)$ gesucht so daß $f(x,y,z) = 0$ gilt. Es ist

$$\begin{pmatrix} \dfrac{\partial f_1}{\partial x} & \dfrac{\partial f_1}{\partial y} \\ \dfrac{\partial f_2}{\partial x} & \dfrac{\partial f_2}{\partial y} \end{pmatrix} = \begin{pmatrix} 6x+y & x \\ 2z+y & 3y^2+x \end{pmatrix},$$

also

$$\begin{vmatrix} \dfrac{\partial f_1}{\partial x}(1,0,0) & \dfrac{\partial f_1}{\partial y}(1,0,0) \\ \dfrac{\partial f_2}{\partial x}(1,0,0) & \dfrac{\partial f_2}{\partial y}(1,0,0) \end{vmatrix} = \begin{vmatrix} 6 & 1 \\ 0 & 1 \end{vmatrix} = 6 \neq 0\,,$$

also kann eine genügend kleine Zahl $\delta > 0$ und eine Funktion $\varphi : (-\delta,\delta) \to \mathbb{R}^2$ gefunden werden mit $f(\varphi_1(z),\varphi_2(z),z) = 0$ für alle z mit $|z| < \delta$. Es gilt

$$\begin{aligned}
\varphi'(z) &= -\begin{pmatrix} 6x+y & x \\ 2z+y & 3y^2+x \end{pmatrix}^{-1} \begin{pmatrix} -1 \\ 2x \end{pmatrix} \\
&= \frac{-1}{(6x+y)(3y^2+x) - x(2z+y)} \begin{pmatrix} 3y^2+x & -x \\ -(2z+y) & 6x+y \end{pmatrix} \begin{pmatrix} -1 \\ 2x \end{pmatrix} \\
&= \frac{-1}{(6x+y)(3y^2+x) - x(2z+y)} \begin{pmatrix} -3y^2 - x - 2x^2 \\ +(2z+y) + 12x^2 + 2xy \end{pmatrix},
\end{aligned}$$

mit $x = \varphi_1(z)$ und mit $y = \varphi_2(z)$. Insbesondere gilt

$$\varphi'(0) = -\frac{1}{6} \begin{pmatrix} -3 \\ \\ 12 \end{pmatrix} = \begin{pmatrix} \frac{1}{2} \\ -2 \end{pmatrix} ,$$

wegen $\varphi(0) = (1, 0)$, also $\varphi_1(0) = 1, \;\; \varphi_2(0) = 0$.