Analysis II

Gunther H. Peichl

Skriptum zur Vorlesung im WS 2011 / 2012

Institut für Mathematik und Wissenschaftliches Rechnen
Karl–Franzens–Universität Graz

KAPITEL VII

Integralrechnung

Historisch wurzelt der Integralbegriff in der Ermittelung von Flächeninhalten. Aber auch physikalische Problemstellungen führen auf Probleme der Integralrechnung: Etwa die Berechnung des zurückgelegten Weges während einer Zeitspanne $[a, b]$, wenn man die Geschwindigkeit v als Funktion der Zeit kennt. Besonders einfach ist die Lösung dieses Problems, wenn die Geschwindigkeit auf jedem Intervall $[t_{i-1}, t_i]$, $t_0 = a < t_1 < \cdots < t_{i-1} < t_i < \cdots < t_n = b$, einen konstanten Wert v_i, $i = 1, \dots, n$ annimmt. Da der bei konstanter Geschwindigkeit v_i während der Zeitspanne $t_i - t_{i-1}$ zurückgelegte Weg durch $v_i(t_i - t_{i-1})$ gegeben ist, ergibt sich für den gesamten Weg der Ausdruck $\sum_{i=1}^{n} v_i(t_i - t_{i-1})$. Dies ist ein Beispiel eines besonders einfachen Integrals. Selbstverständlich genügt es nicht, sich auf einen derart engen Integralbegriff zu beschränken. Wir zeigen im folgenden, wie man das Integral von stückweise konstanten Funktionen wie oben auf eine für die Praxis ausreichend reichhaltige Klasse von Funktionen systematisch ausdehnen kann.

Notation: In diesem Kapitel bezeichnet I stets das abgeschlossene, beschränkte Intervall $[a, b] \subset \mathbb{R}$. Diese Voraussetzung wird im Folgenden daher nicht mehr gesondert ausgewiesen. Wir erinnern an die Norm von $f \in \mathcal{B}(I, \mathbb{K})$

$$||f|| \colon = \sup\{|f(x)| \colon x \in I\},$$

das ist die **Norm** von f. Die gleichmäßige Konvergenz einer Folge von Funktionen (f_n) gegen f werden wir durch $f_n \rightrightarrows f$ andeuten. Mit $\mathbb{K}$ bezeichnen wir den Körper der reellen oder komplexen Zahlen.

1. Treppenfunktionen, Regelfunktionen

Wir präzisieren vorerst die Klasse der stückweise konstanten Funktionen, welchen auf anschauliche Weise ein Integralwert zugeordnet werden kann.

Definition VII-1.1. *Es sei $f \colon I \to \mathbb{K}$ und $n \in \mathbb{N}$.*

i) $\mathcal{Z} = \{x_0, x_1, \dots, x_n\}$ *heißt* ***Zerlegung*** *von I* $\underset{Def}{\Leftrightarrow}$ $a = x_0 < x_1 < \cdots < x_n = b$.

   $|\mathcal{Z}| := \max\{x_i - x_{i-1} : x_i \in \mathcal{Z}, i = 1, \dots, n\}$ *heißt* ***Feinheitsmaß*** *der Zerlegung $\mathcal{Z}$.*

ii) *f heißt* ***Treppenfunktion*** $\underset{Def}{\Leftrightarrow}$ *es gibt eine Zerlegung $\mathcal{Z} = \{x_0, x_1, \dots, x_n\}$ von I und Zahlen c_i, $i = 1, \dots, n$, derart, daß $f|_{(x_{i-1}, x_i)} = c_i$,*

iii) $\mathcal{T}(I, \mathbb{K}) := \{f \colon I \to \mathbb{K}$ *ist eine Treppenfunktion*$\}$.

Eine Treppenfunktion nimmt nur endlich viele Werte an, somit gilt trivialerweise

$$\mathcal{T}(I,\mathbb{K}) \subset \mathcal{B}(I,\mathbb{K}).$$

Da f auf I definiert ist, sind natürlich auch die Funktionswerte von $f(x_i)$, $i = 0,\ldots,n$, festgelegt. Sie sind aber im Folgenden ohne Bedeutung. Verschiedene Zerlegungen $\mathcal{Z}$ von I können zur gleichen Treppenfunktion f führen, z.B. wenn man noch weitere Teilpunkte in eine Zerlegung einfügt. Dies ist zweckmäßig etwa beim Nachweis, daß für $f, g \in \mathcal{T}(I,\mathbb{K})$ auch $f + g \in \mathcal{T}(I,\mathbb{K})$ gilt. Bezeichnet man mit $\mathcal{Z}_f$ bzw. $\mathcal{Z}_g$ den Treppenfunktionen f bzw. g zugrundeliegenden Zerlegungen von I, und mit $\mathcal{Z}$ jene Zerlegung, in der sämtliche Teilpunkte von $\mathcal{Z}_f$ und $\mathcal{Z}_g$ vorkommen, d.h. $\mathcal{Z}$ ist eine **Verfeinerung** von $\mathcal{Z}_f$ und $\mathcal{Z}_g$, dann sind sowohl f als auch g und damit auch $f + g$ (aber auch fg) auf den Teilintervallen von $\mathcal{Z}$ konstant. Da natürlich $\lambda f \in \mathcal{T}(I,\mathbb{K})$ für $f \in \mathcal{T}(I,\mathbb{K})$ und $\lambda \in \mathbb{K}$ gilt, ergibt sich, daß $\mathcal{T}(I,\mathbb{K})$ ein (komplexer) Vektorraum ist. Im nächsten Abschnitt werden wir zeigen, daß man auf $\mathcal{T}(I,\mathbb{K})$ in naheliegender Weise einen Integralbegriff erklären kann. Um zu verstehen, auf welche Funktionen sich dieser Integralbegriff erweitern läßt, ist es notwendig zu untersuchen, welche Funktionen als *gleichmäßige* Grenzwerte von Treppenfunktionen darstellbar sind. Dies trifft beispielsweise für stetige Funktionen zu.

THEOREM VII-1.2. *Es sei $f \in C(I,\mathbb{K})$. Dann gibt es eine Folge $(t_n) \subset \mathcal{T}(I,\mathbb{K})$, die gleichmäßig gegen f konvergiert:* $\lim_{n\to\infty} \|f - t_n\| = 0$.

BEWEIS. Wegen der Kompaktheit von I ist f gleichmäßig stetig, d.h.

$$(*) \qquad \forall \varepsilon > 0 \exists \delta = \delta(\varepsilon) \forall x, y \in I\colon |x - y| < \delta \Rightarrow |f(x) - f(y)| < \varepsilon.$$

Es sei $\varepsilon > 0$ und δ durch $(*)$ gegeben. Wir wählen eine Zerlegung $\mathcal{Z} = \{x_0,\ldots,x_n\}$ mit Feinheitsmaß $|\mathcal{Z}| \le \delta$ und definieren eine Treppenfunktion t_ε, indem wir $\xi_i \in [x_{i-1}, x_i)$ beliebig wählen und

$$t_\varepsilon(x) = \begin{cases} f(\xi_i), & x \in [x_{i-1}, x_i), \quad i = 1,\ldots,n, \\ f(b), & x = b \end{cases}$$

festsetzen. Da jedes $x \in I$ in genau einem der Intervalle $[x_{i-1}, x_i)$, $i = 1,\ldots,n$, liegt (oder $x = b$ gilt), folgt mit $|\xi_i - x| < \delta$ aus $(*)$ für $x \ne b$

$$|f(x) - t_\varepsilon(x)| = |f(x) - f(\xi_i)| < \varepsilon,$$

(für $x = b$ erhält man $|f(b) - t_\varepsilon(b)| = 0$) also

$$\|f - t_\varepsilon\| \le \varepsilon.$$

Läßt man nun ε eine Nullfolge durchlaufen, etwa $(\frac{1}{n})$, erhält man eine Folge $(t_n) \subset \mathcal{T}(I,\mathbb{K})$ mit

$$\|f - t_n\| \le \frac{1}{n},$$

d.h. (t_n) konvergiert gleichmäßig gegen f. □

Auf die Kompaktheit von I kann nicht verzichtet werden: es ist nicht möglich, die stetige Funktion $x \mapsto \frac{1}{x}$ auf $(0,1)$ gleichmäßig durch eine Folge von Treppenfunktionen zu approximieren.
Der Beweis zeigt, wie man vorgehen muß, wenn man zu einer gegebenen *stetigen* Funktion f eine approximierende Folge von Treppenfunktionen (t_n) konstruieren soll. Aus der Definition VII-1.1 geht hervor, daß Treppenfunktionen alle sinnvollen links- und rechtsseitigen Grenzwerte besitzen. Diese Eigenschaft überträgt sich auch auf die gleichmäßigen Grenzwerte von Treppenfunktionen.

Theorem VII-1.3. *Die Abbildung $f\colon I \to \mathbb{K}$ werde gleichmäßig durch eine Folge $(t_n) \subset \mathcal{T}(I,\mathbb{K})$ approximiert. Dann besitzt f an jeder Stelle – wo dies möglich ist – sowohl den rechtsseitigen als auch den linksseitigen Grenzwert.*

Beweis. Wir zeigen nur die Existenz des rechtsseitigen Grenzwertes $f(x^+)$ für $x \in [a,b)$ und überlassen den entsprechenden Nachweis für $f(x^-)$ dem interessierten Leser. Dazu betrachten wir eine beliebige Folge (x_n) mit $x_n \geq \xi$ und $\lim_{n\to\infty} x_n = \xi$. Es genügt zu zeigen, daß $\bigl(f(x_n)\bigr)$ eine Cauchyfolge ist. Wir wählen $\varepsilon > 0$. Nach Voraussetzung gibt es einen Index N_ε und eine Treppenfunktion t_{N_ε} mit $\|f - t_{N_\varepsilon}\| < \varepsilon$. Somit gilt

$$|f(x) - t_{N_\varepsilon}(x)| < \varepsilon, \qquad \text{für alle } x \in I.$$

Die Treppenfunktion t_{N_ε} besitzt den rechtsseitigen Grenzwert in $x = \xi$. Folglich existiert $\delta = \delta(\varepsilon,\xi) > 0$ mit

$$|t_{N_\varepsilon}(x) - t_{N_\varepsilon}(y)| < \varepsilon, \qquad \text{für alle } x, y \in (\xi, \xi+\delta).$$

Insgesamt ergibt sich daher für solche x, y

$$(*) \qquad |f(x) - f(y)| \leq |f(x) - t_{N_\varepsilon}(x)| + |t_{N_\varepsilon}(x) - t_{N_\varepsilon}(y)| + |t_{N_\varepsilon}(y) - f(y)| < 3\varepsilon.$$

Da (x_n) gegen ξ konvergiert, gibt es einen Index $\tilde{N}(\varepsilon)$, sodaß $x_n \in (\xi, \xi+\delta)$ für alle $n \geq \tilde{N}(\varepsilon)$ zutrifft. Für n, $m \geq \tilde{N}(\varepsilon)$ gilt dann wegen $(*)$

$$|f(x_n) - f(x_m)| \leq 3\varepsilon.$$

Somit ist $\bigl(f(x_n)\bigr)$ eine Cauchyfolge und besitzt wegen der Vollständigkeit von $\mathbb{K}$ einen Grenzwert L. Man überzeuge sich, daß L unabhängig von der Wahl der approximierenden Folge (x_n) ist. Es folgt $L = f(\xi^+)$. □

Definition VII-1.4. *Es sei $f\colon I \to \mathbb{K}$.*

i) *f heißt **Regelfunktion** $\underset{Def}{\Leftrightarrow} \exists (t_n) \subset \mathcal{T}(I,\mathbb{K})\colon \lim_{n\to\infty} t_n = f$, glm.*

ii) *$\mathcal{R}(I,\mathbb{K}) := \{f\colon I \to \mathbb{K}, f$ ist Regelfunktion $\}$.*

Satz VII-1.3 zeigt, daß eine Regelfunktion nur Unstetigkeiten 1. Art aufweisen kann. Diese für Regelfunktionen notwendige Bedingung ist auch hinreichend.

Theorem VII-1.5. *Besitzt $f\colon I \to \mathbb{K}$ überall den links- und rechtsseitigen Grenzwert (wo dies möglich ist), dann ist f eine Regelfunktion.*

Beweis. Es sei $\varepsilon > 0$ gegeben. Für alle $\xi \in [a,b]$ gibt es nach Voraussetzung eine Umgebung $K(\xi,\delta_\xi)$, $\delta_\xi > 0$, mit folgender Eigenschaft:

$$(*) \qquad \forall x,y \in I\colon \big(x,y \in (\xi,\xi+\delta_\xi) \vee x,y \in (\xi-\delta_\xi,\xi)\big) \Rightarrow |f(x)-f(y)| < \varepsilon.$$

Die Familie $\{K(\xi,\delta_\xi)\colon \xi \in I\}$ bildet eine offene Überdeckung von I. Da I kompakt ist, genügen bereits endlich viele dieser Umgebungen, etwa $K(\xi_i,\delta_i)$, $i = 1,\dots,k$, um I zu überdecken. Die Punkte ξ_i und die Endpunkte der Intervalle $K(\xi_i,\delta_i)$ denken wir uns der Größe nach geordnet und erhalten dadurch eine Zerlegung $\mathcal{Z}$:

$$a = x_0 < x_1 < \cdots < x_n = b.$$

Aus jedem offenen Intervall (x_{i-1},x_i) wählen wir einen Punkt z_i und definieren $t_\varepsilon \in \mathcal{T}(I,\mathbb{K})$ durch

$$t_\varepsilon(x) = \begin{cases} f(z_i), & x \in (x_{i-1},x_i), \qquad i = 1,\dots,n, \\ f(x_i), & x = x_i \qquad\qquad\quad\; i = 0,\dots,n. \end{cases}$$

Für alle $x \in I$ gilt dann

$$|f(x) - t_\varepsilon(x)| < \varepsilon.$$

Dies ist trivialerweise richtig für $x = x_j$. Zu jedem anderen x gibt es genau einen Index j und mindestens einen Index $1 \le \ell \le k$ mit $x \in (x_{j-1},x_j) \subset K(\xi_\ell,\delta_\ell)$. Dies ist trivial, falls x_{j-1} oder x_j mit einem der Mittelpunkte ξ_ℓ übereinstimmt. Im Fall $x_{j-1} = \xi_\ell - \delta_\ell$ oder $x_j = \xi_\ell + \delta_\ell$ folgt zwangsläufig $\xi_\ell \ge x_j$ bzw. $\xi_\ell \le x_{j-1}$. Andere Fälle kann man analog behandeln, problematisch ist lediglich die Situation $x_{j-1} = \xi_r + \delta_r$ und $x_j = \xi_s - \delta_s$ mit $1 \le r,s \le k$. Dann muß es ein weiteres Intervall $K(\xi_\ell,\delta_\ell)$ mit $(x_{j-1},x_j) \subset K(\xi_\ell,\delta_\ell)$ geben, da andernfalls $I \subset \bigcup_{\nu=1}^k K(\xi_\nu,\delta_\nu)$ verletzt wäre.
Da somit (x_{j-1},x_j) entweder links oder rechts von ξ_ℓ liegt, folgt aus $(*)$

$$|f(x) - t_\varepsilon(x)| = |f(x) - f(z_j)| < \varepsilon.$$

Läßt man ε eine Nullfolge durchlaufen, erhält man eine Folge (t_n) mit $\lim_{n\to\infty} t_n = f$, glm. □

Kombiniert man die Sätze VII-1.5 und VII-1.3, ergibt sich folgende Charakterisierung von Regelfunktionen.

Theorem VII-1.6. *Es sei $f\colon I \to \mathbb{K}$. Äquivalent sind*

(1) *f ist eine Regelfunktion,*
(2) *f besitzt überall den links- und rechtsseitigen Grenzwert (wo dies möglich ist).*

Ist $f\colon [a,b] \to \mathbb{R}$ monoton, dann existiert für alle $x \in (a,b]$ der rechtsseitige Grenzwert, für alle $x \in [a,b)$ der linksseitige Grenzwert und es gilt für monoton fallende Funktionen

$$f(x^+) = \inf_{s>x} f(s), \qquad \text{bzw. } f(x^-) = \sup_{s<x} f(s).$$

Monotone Funktionen sind also Regelfunktionen.

Sind f, g Regelfunktionen, folgt aus Satz VII-1.6, daß auch $f + g$ bzw. λf für $\lambda \in \mathbb{K}$ Regelfunktionen sind. $\mathcal{R}(I, \mathbb{K})$ ist somit ein Vektorraum. Es gilt

$$\mathcal{R}(I, \mathbb{K}) \subset \mathcal{B}(I, \mathbb{K}).$$

Theorem VII-1.7. *Jede Regelfunktion* $f \colon I = [a, b] \to \mathbb{K}$ *besitzt höchstens abzählbar viele Unstetigkeiten.*

Beweis. Es sei $(t_n) \subset \mathcal{T}(I, \mathbb{K})$ eine approximierende Folge von Treppenfunktionen, d.h. es gilt $f = \lim_{n\to\infty} t_n$ glm. auf I. Jede Treppenfunktion t_n ist stetig, ausgenommen auf einer höchstens endlichen Menge $A_n \subset I$. Somit ist ist $A := \bigcup_{n=1}^{\infty} A_n$ abzählbar. Da jede Treppenfunktion t_n auf $I \setminus A$ stetig ist, folgt aus Satz IV-3.6, daß auch f auf $I \setminus A$ stetig ist. $\square$

Abschließend zeigen wir, daß $\mathcal{R}(I, \mathbb{K})$ gegenüber der Bildung gleichmäßiger Grenzwerte abgeschlossen ist:

Theorem VII-1.8. *Es sei* $(f_n)_{n\geq 1} \subset \mathcal{R}(I, \mathbb{K})$ *und* $f_n \rightrightarrows f$. *Dann ist* f *eine Regelfunktion.*

Beweis. Es sei $\varepsilon > 0$ gegeben. Zu jedem f_n existiert $t_{n_\varepsilon} \in \mathcal{T}(I, \mathbb{K})$ mit

$$\|f_n - t_{n_\varepsilon}\| < \frac{\varepsilon}{2}.$$

Zu $\varepsilon > 0$ gibt es nach Voraussetzung einen Index $N(\varepsilon)$ mit $\|f - f_n\| < \frac{\varepsilon}{2}$ für alle $n \geq N(\varepsilon)$. Somit folgt für alle $n \geq N(\varepsilon)$

$$\|f - t_{n_\varepsilon}\| \leq \|f - f_n\| + \|f_n - t_{n_\varepsilon}\| < \varepsilon.$$

Somit kann auch f durch eine Folge von Treppenfunktionen gleichmäßig approximiert werden und ist daher selbst eine Regelfunktion. $\square$

2. Das Cauchy Integral

Es sei $t \in \mathcal{T}(I, \mathbb{K})$ und $\mathcal{Z} = \{x_0, x_1, \ldots, x_n\}$ eine zu t passende Zerlegung von I. Ferner bezeichnen wir mit c_i den Wert von t auf (x_{i-1}, x_i), $i = 1, \ldots, n$. Wir definieren die Abbildung $\mathcal{I} \colon \mathcal{T}(I, \mathbb{K}) \to \mathbb{C}$ durch

$$\mathcal{I}(t) = \sum_{k=1}^{n} c_k (x_k - x_{k-1}).$$

Der Wert von $\mathcal{I}(t)$ könnte möglicherweise nicht nur von t, sondern auch von der jeweiligen Zerlegung $\mathcal{Z}$ abhängen. Tatsächlich ist dies aber nicht der Fall: Man überzeugt sich davon, indem man sich vorerst überlegt, daß sich $\mathcal{I}(t)$ nicht ändert, wenn man einen Teilpunkt in $\mathcal{Z}$ einfügt, oder einen (redundanten) Teilpunkt von $\mathcal{Z}$, in dem t stetig ist, wegläßt. Es seien nun t eine Treppenfunktion und $\mathcal{Z}_i$, $i = 1, 2$, zwei nach Definition VII-1.1 zugeordnete Zerlegungen von I. Wir bilden die Zerlegung $\mathcal{Z}_{12}$, in welcher sämtliche Teilpunkte von $\mathcal{Z}_1$ und $\mathcal{Z}_2$ nach der Größe geordnet auftreten. Ferner sei $\mathcal{I}_i(t)$ der Wert der Summe, wenn man der Berechnung die Zerlegung $\mathcal{Z}_i$ zugrundelegt, $i = 1, 2, 12$. Man kann nun der Zerlegung $\mathcal{Z}_1$ bzw. $\mathcal{Z}_2$ schrittweise einen

Teilpunkt hinzufügen, ohne daß sich der Wert von $\mathcal{I}_i(t)$, $i = 1, 2$ verändert. Nach endlich vielen Schritten ist man schließlich bei der Zerlegung $\mathcal{Z}_{12}$ angelangt und erhält

$$\mathcal{I}_1(t) = \mathcal{I}_{12}(t) = \mathcal{I}_2(t).$$

Somit ist folgende Definition sinnvoll:

DEFINITION VII-2.1. *Es sei $t \in \mathcal{T}(I, \mathbb{K})$ und $\mathcal{Z} = \{x_0, \dots, x_n\}$ eine Zerlegung von I. Ferner seien $c_i \in \mathbb{K}$, $i = 1, \dots, n$, die Werte von t auf (x_{i-1}, x_i).*

$$\mathcal{I}(t) := \sum_{k=1}^{n} c_k(x_k - x_{k-1})$$

*heißt **(bestimmtes) Integral** von t. Man schreibt auch*

$$\mathcal{I}(t) = \int_a^b t = \int_a^b t(x)dx$$

*und nennt t **Integrand**, a die* **untere** *und b die* **obere** ***Integrationsgrenze**.*

BEISPIEL VII-2.2. Wir betrachten die Abbildung (für festes $n \in \mathbb{N}$)

$$t(x) = \frac{1}{n}[nx], \quad x \in [0, 1].$$

(vgl. Beispiel IV-3.2). t nimmt auf $[\frac{k}{n}, \frac{k+1}{n})$ den Wert $c_k = \frac{k}{n}$ an, $k = 0, \dots, n-1$. Eine zu t passende Zerlegung von $[0, 1]$ ist die äquidistante Einteilung

$$0 < \frac{1}{n} < \dots < \frac{k}{n} < \dots < \frac{n-1}{n} < 1,$$

d.h. $x_i = \frac{i}{n}$, $i = 0, \dots, n$, und $c_i = \frac{i-1}{n}$, $i = 1, \dots, n$. Somit folgt

$$\mathcal{I}(t) = \sum_{i=1}^{n} \frac{i-1}{n} \cdot \frac{1}{n} = \frac{1}{n^2} \sum_{i=1}^{n} (i-1) = \frac{1}{n^2} \sum_{i=0}^{n-1} i = \frac{n(n-1)}{n^2 \cdot 2} = \frac{1}{2} \frac{n-1}{n}.$$

Wir zeigen nun, daß $\mathcal{I} \colon \mathcal{T}(I, \mathbb{K}) \to \mathbb{K}$ eine lineare Abbildung ist:

THEOREM VII-2.3. *Für alle $f, g \in \mathcal{T}(I, \mathbb{K})$, $I = [a, b]$, und für alle $\lambda \in \mathbb{K}$ gilt*
1. *$\mathcal{I}(f + g) = \mathcal{I}(f) + \mathcal{I}(g)$,*
2. *$\mathcal{I}(\lambda f) = \lambda \mathcal{I}(f)$.*

BEWEIS. Für den Beweis von (1) bildet man die Zerlegung $\mathcal{Z}$, in der alle Teilpunkte der zu f und zu g gehörenden Zerlegungen auftreten. Dann sind sowohl f als auch g konstant auf den Teilintervallen von $\mathcal{Z}$. Die Behauptung (1) ist nun evident. Die Behauptung (2) folgt unmittelbar aus der Definition von $\mathcal{I}$. □

THEOREM VII-2.4. *1. Es sei $f \in \mathcal{T}(I, \mathbb{R})$ und $f(x) \geq 0$ für alle $x \in I$. Dann ist $\mathcal{I}(f) \geq 0$ (d.h. $\mathcal{I}$ ist eine **positive** lineare Abbildung).*
*2. Es seien $f, g \in \mathcal{T}(I, \mathbb{R})$ und $f(x) \geq g(x)$ für alle $x \in I$. Dann ist $\mathcal{I}(f) \geq \mathcal{I}(g)$ (d.h. $\mathcal{I}$ ist eine **monotone** lineare Abbildung).*

3. Für alle $f \in \mathcal{T}(I,\mathbb{K})$ gilt

$$|\mathcal{I}(f)| \leq \mathcal{I}(|f|) \qquad \textit{und} \qquad |\mathcal{I}(f)| \leq \|f\|(b-a).$$

Somit folgt $\mathcal{I} \in \mathcal{L}(\mathcal{T}(I,\mathbb{K}),\mathbb{K})$.

Beweis. Die Aussage 1. liest man unmittelbar aus der Definition ab und 2. folgt aus 1.
ad 3. Es sei $\mathcal{Z} = \{x_0, \ldots, x_n\}$ eine zu f gehörige Zerlegung von I und c_i der Wert von f auf (x_{i-1}, x_i). Aus

$$|c_i| \leq \sup\{|f(x)| \colon x \in I\} = \|f\|, \quad i = 1, \ldots, n,$$

folgt

$$|\mathcal{I}(f)| \leq \sum_{i=1}^{n} |c_i|(x_i - x_{i-1}) \leq \|f\|(b-a).$$

□

Neben der Additivität bezüglich des Integranden hat das Integral auch eine Additivitätseigenschaft bezüglich des Integrationsintervalls: Dazu bemerken wir, daß die Einschränkung einer Treppenfunktion $f \in \mathcal{T}(I,\mathbb{K})$ auf ein beliebiges Teilintervall $J \subset I$ wieder eine Treppenfunktion ist, die wir der Einfachheit halber wieder mit f bezeichnen:

Theorem VII-2.5. *Es sei $f \in \mathcal{T}(I,\mathbb{K})$ und $c \in I$. Dann gilt*

$$\int_a^c f(x)dx + \int_c^b f(x)dx = \int_a^b f(x)dx.$$

Beweis. Die Behauptung ergibt sich unmittelbar indem man den Punkt c in die Zerlegung aufnimmt, welche für die Berechnung von $\int_a^b f(x)dx$ verwendet wird. □

Die Beschränktheit von $\mathcal{I}$ ermöglicht es, die Definition des Integrals von $\mathcal{T}(I,\mathbb{K})$ auf $\mathcal{R}(I,\mathbb{K})$ zu erweitern: Dazu betrachten wir eine Regelfunktion f und eine Folge von Treppenfunktionen (t_n), welche gleichmäßig gegen f konvergiert. Es folgt

$$|\mathcal{I}(t_n) - \mathcal{I}(t_m)| \leq (b-a)\|t_n - t_m\|.$$

Wegen der Gleichmäßigkeit der Konvergenz von (t_n) ist für alle $x \in I$ die Folge $(t_n(x))$ eine (gleichmäßige) Cauchy Folge, d.h. es existiert ein $N(\varepsilon) \in \mathbb{N}$ derart, daß für alle $n, m \geq N(\varepsilon)$

$$\|t_n - t_m\| < \frac{\varepsilon}{b-a}$$

zutrifft. Dies zeigt für $n, m \geq N(\varepsilon)$

$$|\mathcal{I}(t_n) - \mathcal{I}(t_m)| < \varepsilon,$$

d.h. $(\mathcal{I}(t_n))_{n\geq 1}$ ist eine Cauchy Folge in $\mathbb{K}$ und somit konvergent. Es liegt nahe

$$\mathcal{I}(f) := \lim_{n\to\infty} \mathcal{I}(t_n)$$

zu setzen. Diese Definition ist sinnvoll, soferne gezeigt werden kann, daß $\lim_{n\to\infty} \mathcal{I}(t_n)$ unabhängig ist von der jeweiligen Folge (t_n), welche gegen f konvergiert: Es gelte also für zwei Folgen von Treppenfunktionen $t_n \rightrightarrows f$ und $s_n \rightrightarrows f$. Zu zeigen ist: $\lim_{n\to\infty} \mathcal{I}(t_n) = \lim_{n\to\infty} \mathcal{I}(s_n)$ bzw. $\lim_{n\to\infty} \mathcal{I}(t_n - s_n) = 0$. Letzteres ist aber eine unmittelbare Folge der Beschränktheit von $\mathcal{I}$

$$|\mathcal{I}(t_n - s_n)| \leq (b-a)\|t_n - s_n\|$$

und des Faktums $\lim_{n\to\infty} \|t_n - s_n\| \leq \lim_{n\to\infty} \|t_n - f\| + \lim_{n\to\infty} \|f - s_n\| = 0$.
Es ist daher sinnvoll, folgende Definition zu vereinbaren:

Definition VII-2.6. *Es sei $f \in \mathcal{R}(I,\mathbb{K})$ und $(t_n) \subset \mathcal{T}(I,\mathbb{K})$ mit $t_n \rightrightarrows h$.*

$$\mathcal{I}(f) := \lim_{n\to\infty} \mathcal{I}(t_n) \text{ heißt } \textbf{Integral der Regelfunktion } f.$$

Für das Integral $\mathcal{I}(f)$ schreibt man auch

$$\mathcal{I}(f) = \int_a^b f = \int_a^b f(x)dx.$$

Wegen $\mathcal{T}(I,\mathbb{K}) \subset \mathcal{R}(I,\mathbb{K})$ kann man für $f \in \mathcal{T}(I,\mathbb{K})$ die konstante Folge $t_n = f$, $n \in \mathbb{N}$, zur Approximation verwenden. Da der Grenzwert $\mathcal{I}(f)$ unabhängig von der approximierenden Folge von Treppenfunktionen ist, ergibt sich, daß das Integral gemäß Definition VII-2.6 auf $\mathcal{T}(I,\mathbb{K})$ mit dem Integral gemäß Definition VII-2.1 übereinstimmt. Durch Definition VII-2.6 wird also der Integralbegriff aus Definition VII-2.1 von $\mathcal{T}(I,\mathbb{K})$ auf $\mathcal{R}(I,\mathbb{K})$ fortgesetzt. Satz VII-1.6 kann nun als notwendige *und* hinreichende Integrabilitätsbedingung betrachtet werden: Die Klasse der **integrierbaren Funktionen** ist der Vektorraum der Funktionen, für welche überall (soferne sinnvoll) rechts- und linksseitige Grenzwerte existieren.
Wir veranschaulichen die Handhabung der Definition an einigen Beispielen:

Beispiel VII-2.7. 1) Wir betrachten $f(x) = x$, $x \in [0,1]$. Da f stetig ist, ist f eine Regelfunktion. Als approximierende Folge von Treppenfunktionen können wir $t_n(x) = \frac{1}{n}[nx]$, $n \in \mathbb{N}$, verwenden (vgl. IV-3.2). Nach Beispiel VII-2.2 gilt

$$\mathcal{I}(t_n) = \frac{1}{2}(1 - \frac{1}{n}),$$

und daher

$$\mathcal{I}(f) = \int_0^1 f(x)dx = \lim_{n\to\infty} \mathcal{I}(t_n) = \frac{1}{2}.$$

2) Für $f(x) = e^x$, $x \in [a,b]$, zeigen wir nun

$$\mathcal{I}(f) = \int_a^b f(x)dx = e^b - e^a.$$

Auch in diesem Falle können wir eine äquidistante Zerlegung $\mathcal{Z}_n$ von I verwenden:

$$a = x_0 < x_1 < \cdots < x_n = b, \quad x_i = \frac{b-a}{n} i + a, \quad i = 0, \ldots, n.$$

Wir setzen

$$t_n(x) = \begin{cases} e^{x_{i-1}}, & x \in [x_{i-1}, x_i), \quad i = 1, \ldots, n, \\ e^b, & x = b. \end{cases}$$

Auf Grund des Beweises von Satz VII-1.2 wissen wir zwar bereits, daß wegen $\lim_{n\to\infty} |\mathcal{Z}_n| = 0$ auch $t_n \rightrightarrows f$ gelten muß. Wir wollen uns von der gleichmäßigen Konvergenz der Folge (t_n) jedoch direkt überzeugen: Jedes $x \neq b$ liegt in einem eindeutig bestimmten Intervall $[x_{i-1}, x_i)$. Es folgt also mit Hilfe des Mittelwertsatzes die Existenz von $\xi_i \in (x_{i-1}, x_i)$ mit

$$|f(x) - t_n(x)| = |e^x - e^{x_{i-1}}| \le e^{\xi_i} |x - x_{i-1}| \le e^b |x - x_{i-1}| \le e^b |\mathcal{Z}_n|.$$

Für $x = b$ ist diese Abschätzung trivialerweise efüllt. Daraus folgt

$$\|f - t_n\| \le e^b |\mathcal{Z}_n|.$$

Als nächstes berechnen wir $\mathcal{I}(t_n)$:

$$\begin{aligned} \mathcal{I}(t_n) &= \sum_{k=1}^n \exp(a + \frac{k-1}{n}(b-a)) \frac{b-a}{n} \\ &= \frac{b-a}{n} e^a \sum_{k=1}^n \exp(\frac{k-1}{n}(b-a)) = \frac{b-a}{n} e^a \sum_{k=1}^n (\exp(\frac{b-a}{n}))^{k-1} \\ &= \frac{b-a}{n} e^a \frac{\exp(b-a) - 1}{\exp \frac{b-a}{n} - 1} = \frac{\frac{b-a}{n}}{\exp \frac{b-a}{n} - 1} (e^b - e^a) \xrightarrow[n\to\infty]{} e^b - e^a, \end{aligned}$$

wobei $\lim_{z\to 0} \frac{z}{e^z - 1} = 1$ verwendet wurde.

THEOREM VII-2.8. *Für alle $f, g \in \mathcal{R}(I, \mathbb{K})$ und alle $\lambda \in \mathbb{K}$ gilt*

(1) $\mathcal{I}(f + g) = \mathcal{I}(f) + \mathcal{I}(g)$,
(2) $\mathcal{I}(\lambda f) = \lambda \mathcal{I}(f)$.

BEWEIS. Es seien $(t_n), (s_n) \subset \mathcal{T}(I, \mathbb{K})$ mit

$$t_n \rightrightarrows f, \qquad s_n \rightrightarrows g.$$

Dann gilt $t_n + s_n \in \mathcal{T}(I, \mathbb{K})$ und

$$t_n + s_n \rightrightarrows f + g.$$

Auf $\mathcal{T}(I, \mathbb{K})$ ist $\mathcal{I}$ linear, also gilt

$$\mathcal{I}(t_n + s_n) = \mathcal{I}(t_n) + \mathcal{I}(s_n).$$

Durch Grenzübergang erhält man

$$\mathcal{I}(f + g) = \lim_{n\to\infty} \mathcal{I}(t_n + s_n) = \lim_{n\to\infty} \mathcal{I}(t_n) + \lim_{n\to\infty} \mathcal{I}(s_n) = \mathcal{I}(f) + \mathcal{I}(g).$$

Analog ergibt sich (2). $\square$

THEOREM VII-2.9. *Für alle* $f \in \mathcal{R}(I, \mathbb{K})$ *gilt*

$$|\mathcal{I}(f)| \leq \mathcal{I}(|f|) \qquad \textit{und} \qquad |\mathcal{I}(f)| \leq (b-a)\|f\|.$$

BEWEIS. Bezeichnet man mit (t_n) eine Folge von Treppenfunktionen, welche gleichmäßig gegen f konvergiert, folgt die Behauptung aus der Abschätzung

$$|\mathcal{I}(t_n)| \leq (b-a)||t_n|| \leq (b-a)(||t_n - f|| + ||f||).$$

□

Die Sätze VII-2.8, VII-2.9 zeigen, daß die Erweiterung von $\mathcal{I}$ eine *lineare* und *stetige* Abbildung $\mathcal{R}(I, \mathbb{K}) \to \mathbb{K}$ darstellt. Wir weisen nun nach, daß auch die Positivität erhalten bleibt.

THEOREM VII-2.10. *1) Es sei* $f \in \mathcal{R}(I, \mathbb{R})$ *und* $f(x) \geq 0$ *für alle* $x \in I$. *Dann gilt* $\mathcal{I}(f) \geq 0$.
2) Es sei $f \in C(I, \mathbb{R})$ *und* $f(x) \geq 0$ *für alle* $x \in I$. *Ist* $f(x_0) > 0$ *für ein* $x_0 \in I$, *dann ist* $\mathcal{I}(f) > 0$.

BEWEIS. 1) Es sei $(t_n) \in \mathcal{T}(I, \mathbb{R})$ eine f approximierende Folge von Treppenfunktionen, d.h. $\lim_{n\to\infty} \|f - t_n\| = 0$, und $\varepsilon > 0$. Es gibt also ein $N(\varepsilon) \in \mathbb{N}$, sodaß für alle $n \geq N(\varepsilon)$ und für alle $x \in I$

$$t_n(x) \geq f(x) - \varepsilon \geq -\varepsilon$$

zutrifft. Nach Satz VII-2.4-2) gilt

$$\mathcal{I}(t_n) \geq -\varepsilon(b-a)$$

und daher auch $\mathcal{I}(f) = \lim_{n\to\infty} \mathcal{I}(t_n) \geq -\varepsilon(b-a)$. Da $\varepsilon > 0$ beliebig gewählt war, gilt die Behauptung.
2) O.B.d.A können wir $x_0 \in (a, b)$ annehmen. Es sei $\delta > 0$ so gewählt, daß für alle $x \in (x_0 - \delta, x_0 + \delta) \subset I$,

$$f(x) > \frac{1}{2} f(x_0) > 0.$$

Wir definieren nun die Treppenfunktion

$$t(x) = \begin{cases} \frac{1}{2} f(x_0), & x \in (x_0 - \delta, x_0 + \delta), \\ 0, & \text{sonst}, \end{cases}$$

sodaß $f(x) \geq t(x)$ für alle $x \in I$ gilt. Wegen 1) folgt

$$\mathcal{I}(f) \geq \mathcal{I}(t) = \frac{1}{2} f(x_0) 2\delta > 0.$$

□

Eine unmittelbar Folge der Linearität und Positivität ist wieder die Monotonie:

Korollar VII-2.11. Es seien $f, g \in \mathcal{R}(I, \mathbb{R})$ und $f(x) \geq g(x)$ für alle $x \in I$. Dann gilt

$$\mathfrak{I}(f) \geq \mathfrak{I}(g).$$

Sind darüber hinaus f und g stetig und gilt an mindestens einer Stelle $x_0 \in I$ $f(x_0) > g(x_0)$, dann gilt

$$\mathfrak{I}(f) > \mathfrak{I}(g).$$

Bemerkung VII-2.12. Geht man den beschriebenen Fortsetzungsprozeß von $\mathfrak{I}$ und den Beweis der Sätze VII-2.8 und VII-2.10 noch einmal durch, erkennt man, daß weder die spezielle Bedeutung von $\mathfrak{I}$ – das Integral von Treppenfunktionen – noch spezielle Eigenschaften von $\mathbb{K}$ verwendet wurden. Die Beweise beruhen lediglich auf den *strukturellen* Eigenschaften Linearität und Beschränkheit (also Stetigkeit) und der Vollständigkeit $\mathbb{K}$. Die Konstruktion des Cauchy Integrals kann daher auf Funktionen $f\colon I \to X$, X ein Banachraum übertragen werden.

Theorem VII-2.13. *Es sei $f \in \mathcal{R}(I, \mathbb{K})$ und $c \in I$. Dann gilt*

$$\int_a^b f(x)dx = \int_a^c f(x)dx + \int_c^b f(x)dx.$$

Beweis. Zunächst überlegen wir, daß die Restriktion einer Regelfunktion auf ein Teilintervall von I wieder eine Regelfunktion ist. Dies folgt unmittelbar aus der Charakterisierung in Satz VII-1.6. Es sei nun $(t_n) \subset \mathcal{T}(I, \mathbb{K})$ eine approximierende Folge von Treppenfunktionen. Für jedes $n \in \mathbb{N}$ gilt nach Satz VII-2.5

$$\int_a^b t_n(x)dx = \int_a^c t_n(x)dx + \int_c^b t_n(x)dx.$$

Da natürlich $(t_n|_{[a,c]})$ gleichmäßig gegen $f|_{[a,c]}$ konvergiert (und analog für $[c, b]$), ergibt sich die Behauptung aus der Definition des Integrals. □

Es ist zweckmäßig, auf die Voraussetzung $a \leq b$ zu verzichten und für $a > b$ zu definieren

$$\int_a^b f(x)dx := -\int_b^a f(x)dx.$$

Es seien nun a, b und c beliebige reelle Zahlen und $f\colon [\min(a, b, c), \max(a, b, c)] \to \mathbb{K}$ eine Regelfunktion. Mit Hilfe dieser erweiterten Integraldefinition gilt kann man sich von der Gültigkeit folgender Gleichheit überzeugen

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

Theorem VII-2.14. *[2. Mittelwertsatz der Integralrechnung] Es sei* $f \in C(I, \mathbb{R})$, $g \in \mathcal{R}(I, \mathbb{R})$ *und* $g(x) \geq 0$ *für alle* $x \in I$. *Dann gibt es eine Zwischenstelle* $\xi \in I$ *mit*

$$\int_a^b f(x)g(x)dx = f(\xi) \int_a^b g(x)dx.$$

Beweis. Wegen der Kompaktheit von I nimmt f nach Korollar III-4.3 Maximum und Minimum an, d.h. es existieren

$$m = \min\{f(x) \colon x \in I\} \text{ und } M = \max\{f(x) \colon x \in I\}.$$

Somit gilt für alle $x \in I$

$$mg(x) \leq f(x)g(x) \leq Mg(x),$$

und wegen der Monotonie und Linearität des Integrals auch

$$m \int_a^b g(x)\, dx \leq \int_a^b f(x)g(x)\, dx \leq M \int_a^b g(x)\, dx.$$

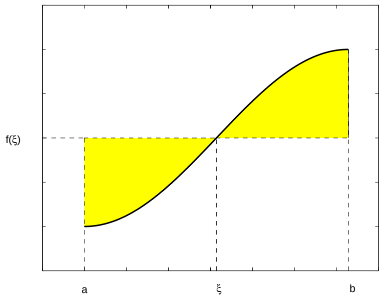
Falls $\int_a^b g(x)\, dx = 0$, muß auch $\int_a^b f(x)g(x)\, dx = 0$ gelten und die Behauptung des Satzes ist richtig. Es sei nun $\int_a^b g(x)\, dx \neq 0$. Dann folgt $\kappa = \frac{\int_a^b f(x)g(x)dx}{\int_a^b g(x)dx} \in [m, M]$. Aus dem Zwischenwertsatz IV-1.3 folgt nun die Existenz einer Zwischenstelle $\xi \in I$ mit $\kappa = f(\xi)$. ☐

Eine eingehende Analyse würde zeigen, daß ξ sogar in (a, b) gewählt werden kann. Für $g \equiv 1$ ergibt sich als Spezialfall:

Korollar VII-2.15 (1. Mittelwertsatz der Integralrechnung). Es sei $f \in C(I, \mathbb{R})$. Dann gibt es ein $\xi \in I$ mit

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

Deutet man $\int_a^b f(x)dx$ (mit $f \geq 0$ auf I) als die Fläche zwischen dem Graph von f und der x-Achse, dann ist nach dem 1. Mittelwertsatz diese gleich dem Flächeninhalt eines Rechteckes mit den Seitenlängen $b - a$ und $f(\xi)$:



3. Stammfunktion

Viele Probleme in Naturwissenschaft und Technik führen auf die Aufgabe, den Differentiationsprozeß umzukehren, d.h. zu einer gegebenen Funktion $f\colon I \to \mathbb{K}$ eine Funktion $F\colon I \to \mathbb{K}$ so zu bestimmen, daß $F' = f$ gilt.

Definition VII-3.1. *Es sei $f\colon I \to \mathbb{K}$. $F\colon I \to \mathbb{K}$ heißt* ***Stammfunktion*** *von f* $\underset{Def}{\Leftrightarrow}$

(1) *F ist stetig.*
(2) *Es gibt eine höchstens abzählbare Teilmenge $A \subset I$ so, daß F auf $I \setminus A$ differenzierbar ist mit*

$$F'(x) = f(x), \quad x \in I \setminus A.$$

Beispiel VII-3.2. Zu jeder Ableitung gibt es eine Stammfunktion (wir lassen in der folgenden Tabelle die jeweiligen Definitionsbereiche weg):

Funktion f	Stammfunktion F	Funktion f	Stammfunktion F
$f(x)$	$F(x)$	$f(x)$	$F(x)$
$x^c \quad (c \neq -1)$	$\frac{1}{c+1}x^{c+1}$	$\frac{1}{\cos^2 x}$	$\tan x$
$\frac{1}{x}$	$\ln x$	$\frac{1}{\sin^2 x}$	$-\cot x$
$a^x \quad (a \neq 1)$	$\frac{1}{\ln a}a^x$	$\frac{1}{\sqrt{1-x^2}}$	$\arcsin x$
$\exp ax \quad (a \neq 0)$	$\frac{1}{a} \cdot \exp ax$	$\frac{1}{1+x^2}$	$\arctan x$
$\exp ix$	$-i \cdot \exp ix$	$\frac{f'(x)}{f(x)}$	$\ln \lvert f(x) \rvert$
$\cos x$	$\sin x$	$\frac{f'(x)}{1+(f(x))^2}$	$\arctan f(x)$
$\sin x$	$-\cos x$	$\sum_{k=0}^{\infty} a_k x^k$	$\sum_{k=0}^{\infty} \frac{a_k}{k+1} x^{k+1}$

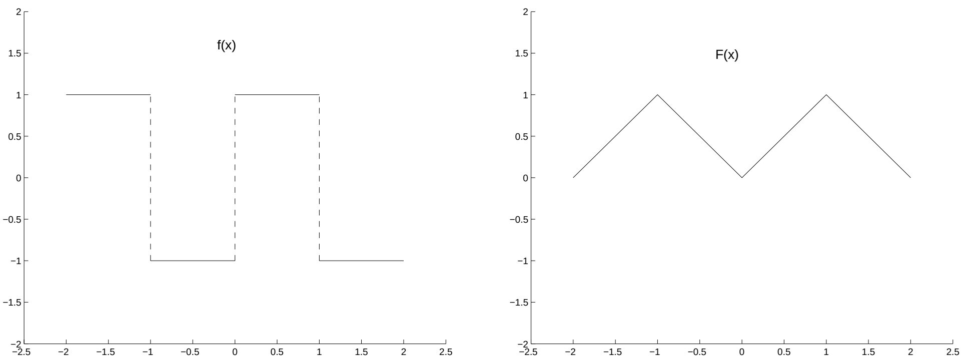
Die meisten Lehrbücher verlangen für eine Stammfunktion F die Differenzierbarkeit und Identität $F' = f$ auf *ganz* I. Bereits einfache Anwendungen in der Technik erfordern aber einen allgemeineren und flexibleren Begriff der Stammfunktion.

Beispiel VII-3.3. Es sei $f\colon \mathbb{R} \to \mathbb{R}$ gegeben durch

$$f(x) = \begin{cases} 1, & 2n \leq x < 2n+1, \quad n \in \mathbb{Z}, \\ -1 & \text{sonst.} \end{cases}$$

Eine Stammfunktion von f ist gegeben durch

$$F(x) = \begin{cases} x - 2n, & 2n \leq x < 2n+1, \\ -x + 2n + 2 & 2n+1 \leq x < 2n+2 \end{cases} n \in \mathbb{Z}.$$



Wegen der Linearität der Differentiation ist klar, daß mit zwei Stammfunktionen F und G für f bzw.g auch $\alpha F + \beta G$, $\alpha, \beta \in \mathbb{K}$, Stammfunktionen von $\alpha f + \beta g$ sind (dabei wird auch verwendet, daß die Vereinigung von zwei abzählbaren Mengen abzählbar ist. Trivialerweise ist mit F natürlich auch $F + \alpha$, $\alpha \in \mathbb{K}$, eine Stammfunktion von f. Wir zeigen nun, daß sich umgekehrt je zwei Stammfunktionen von f höchstens um eine additive Konstante unterscheiden können. Dieses Resultat ist eine unmittelbare Konsequenz von Korallar V-4.7 falls $A = \emptyset$. Der allgemeinere Fall folgt aus

Theorem VII-3.4. *Es sei $f \in C(I, \mathbb{K})$. Ferner gebe es eine abzählbare Teilmenge $A \subset I$ und eine Konstante $L > 0$ derart, daß gilt*

(1) *f ist rechtsseitig differenzierbar auf $I \setminus A$,*
(2) $\forall x \in I \setminus A\colon |f'_+(x)| \leq L$.

Für alle $x, y \in I$ gilt dann

$$|f(x) - f(y)| \leq L|x - y|,$$

d.h. f ist Lipschitz-stetig.

Beweis. Es sei $x < y$. Für beliebiges $\varepsilon > 0$ definieren wir $F_\varepsilon\colon [x, y] \to \mathbb{R}$,

$$F_\varepsilon(\xi) := |f(\xi) - f(x)| - (L + \varepsilon)(\xi - x), \quad \xi \in [x, y],$$

und zeigen $F_\varepsilon(y) \leq 0$. Daraus folgt für $\varepsilon \to 0$ die Behauptung. Angenommen es wäre $F_\varepsilon(y) > 0$. Da $F_\varepsilon(A)$ abzählbar, $[0, F_\varepsilon(y)]$ aber überabzählbar ist, muß es ein $\gamma \in \mathbb{R}$ geben mit

$$0 = F_\varepsilon(x) < \gamma < F_\varepsilon(y) \text{ und } \gamma \notin F_\varepsilon(A).$$

Nach dem Zwischenwertsatz IV-1.3 gibt es eine Zwischenstelle $\xi_\gamma \in (x,y)$ mit $F_\varepsilon(\xi_\gamma) = \gamma$. Somit ist $F_\varepsilon^{-1}(\{\gamma\}) \neq \emptyset$. Wegen der Stetigkeit von F_ε ist $F_\varepsilon^{-1}(\{\gamma\}) \cap [x,y]$ abgeschlossen und besitzt ein maximales Element ξ^*. Für alle $\xi \in (\xi^*, y)$ gilt daher

$$F_\varepsilon(\xi) > \gamma.$$

Dann gilt einerseits

$$(*) \qquad \varphi(\xi) := \frac{F_\varepsilon(\xi) - F_\varepsilon(\xi^*)}{\xi - \xi^*} > 0 \text{ für alle } \xi \in (\xi^*, y],$$

andererseits wegen der Definition von F_ε

$$\begin{aligned}
\varphi(\xi) &= (\xi - \xi^*)^{-1}\big[|f(\xi) - f(x)| - |f(\xi^*) - f(x)| - (L+\varepsilon)(\xi - x - (\xi^* - x))\big] \\
&= (\xi - \xi^*)^{-1}(|f(\xi) - f(x)| - |f(\xi^*) - f(x)|) - (L+\varepsilon) \\
&\leq \frac{|f(\xi) - f(\xi^*)|}{\xi - \xi^*} - L - \varepsilon.
\end{aligned}$$

Aus $\gamma \notin F_\varepsilon(A)$ folgt $\xi^* \notin A$. f besitzt also in ξ^* eine rechtsseitige Ableitung. Wegen $|f'_+(\xi^*)| \leq L$ gibt es also eine rechtsseitige Umgebung $(\xi^*, \xi^* + \delta)$, $\delta > 0$, so, daß

$$\frac{|f(\xi) - f(\xi^*)|}{\xi - \xi^*} \leq L + \varepsilon/2 \quad \text{für alle } \xi \in (\xi^*, \xi^* + \delta)$$

zutrifft. Dies hat jedoch

$$\varphi(\xi) \leq -\frac{\varepsilon}{2} < 0$$

für alle $\xi \in (\xi^*, \xi^* + \delta)$ zur Folge, ein Widerspruch zu $(*)$. □

Korollar VII-3.5. Es sei $f \in C(I, \mathbb{K})$, und $A \subset I$ abzählbar. Auf $I \setminus A$ sei f rechtsseitig differenzierbar mit $f'_+(x) = 0$, $x \in I \setminus A$. Dann ist f konstant.

Beweis. Setze $L = 0$ in Satz VII-3.4. □

Für Stammfunktionen bedeutet Korollar VII-3.5:

Theorem VII-3.6. *Es seien $f\colon I \to \mathbb{K}$ und $F, G \in C(I, \mathbb{K})$ Stammfunktionen von f. Dann ist $F - G$ konstant.*

4. Der Hauptsatz der Differential- und Integralrechnung

In diesem Abschnitt zeigen wir, daß die Integration in gewisser Weise die Umkehroperation zur Differentiation ist.

THEOREM VII-4.1. *Es sei $f \in \mathcal{R}(I,\mathbb{K})$ und $F\colon I \to \mathbb{K}$ die Abbildung*

$$x \to F(x) = \int_a^x f(t)dt, \quad x \in I.$$

Dann folgt:

(1) *F ist Lipschitz stetig auf I,*
(2) *F ist auf $[a,b)$ rechtsseitig differenzierbar und auf $(a,b]$ linksseitig differenzierbar und es gilt*

$$\forall x \in [a,b)\colon F'_+(x) = f(x^+),$$
$$\forall x \in (a,b]\colon F'_-(x) = f(x^-).$$

BEWEIS. 1) Da $f|_{[a,x]}$ eine Regelfunktion ist, ist F sinnvoll definiert. Aus Satz VII-2.13 ergibt sich für $x, y \in I$, $x < y$, die Abschätzung

$$|F(x) - F(y)| = |\int_a^x f(t)dt - \int_a^y f(t)dt| = |\int_a^x f(t)dt - \int_a^x f(t)dt - \int_x^y f(t)dt|$$
$$= |\int_x^y f(t)dt| \le \|f\| \cdot |y - x|.$$

2) Wir führen den Beweis nur für die rechtsseitige Ableitung an $x \in [a,b)$. Für $h > 0$ (h so klein, daß $x + h \in I$) ergibt sich für den rechtsseitigen Differenzenquotienten von F

$$\frac{1}{h}\big(F(x+h) - F(x)\big) = \frac{1}{h}\big(\int_a^{x+h} f(t)dt - \int_a^x f(t)\,dt\big) = \frac{1}{h}\int_x^{x+h} f(t)\,dt.$$

Zu $\varepsilon > 0$ wählen wir nun $\delta > 0$ so, daß $|f(t) - f(x^+)| < \varepsilon$ für alle $t \in (x, x+\delta)$ zutrifft. Damit folgt für $0 < h < \delta$ mit Satz VII-2.9

$$\big|\frac{1}{h}\big(F(x+h) - F(x)\big) - f(x^+)\big| = \big|\frac{1}{h}\int_x^{x+h} f(t)\,dt - f(x^+)\big| = \frac{1}{h}\big|\int_x^{x+h} \big(f(t) - f(x^+)\big)\,dt\big|$$
$$\le \frac{1}{h}\int_x^{x+h} |f(t) - f(x^+)|dt \le \frac{1}{h}\int_x^{x+h} \sup_{t\in(x,x+h)} |f(t) - f(x^+)|dt < \varepsilon.$$

$\square$

THEOREM VII-4.2 (Hauptsatz der Differential- und Integralrechnung).

1) Für jedes $f \in \mathcal{R}(I, \mathbb{K})$, $I = [a, b]$, ist die Abbildung $f\colon I \to \mathbb{K}$,

$$F(x) = \int_a^x f(t)\,dt, \quad x \in I,$$

eine Stammfunktion von f. Insbesondere ist F in den Stetigkeitsstellen von f differenzierbar und es gilt dort

$$F'(x) = f(x).$$

2) Mit einer beliebigen Stammfunktion Φ von f gilt

$$\int_a^b f(t)\,dt = \Phi(b) - \Phi(a).$$

Für die Differenz $\Phi(b) - \Phi(a)$ schreibt man auch $\Phi|_a^b$.

Beweis. 1) Nach Satz VII-4.1 ist F auf I stetig und in allen Stetigkeitsstellen von f gilt

$$F'_+(x) = f(x^+) = f(x) = f(x^-) = F'_-(x),$$

also

$$F'(x) = f(x),$$

d.h. F ist dort differenzierbar (in den Randpunkten von I natürlich nur rechts- bzw. linksseitig). Da nach Satz VII-1.7 f nur abzählbar viele Unstetigkeitsstellen besitzt, ist F eine Stammfunktion von f im Sinne von Definition VII-3.1.
2) Die Behauptung ist trivial für F. Für jede weitere Stammfunktion gilt nach Satz VII-3.6 $F(x) - \Phi(x) = c$, $x \in I$, für ein geeignetes $c \in \mathbb{K}$. Es folgt somit

$$\int_a^b f(t)\,dt = F(b) - F(a) = (\Phi(b) + c) - (\Phi(a) + c) = \Phi(b) - \Phi(a).$$

□

Korollar VII-4.3. Für jedes $F \in C^1(I, \mathbb{K})$, gilt für alle $x \in I$

$$\int_a^x F'(t)\,dt = F(x) - F(a).$$

Bemerkung VII-4.4. 1) Der Hauptsatz beinhaltet die theoretisch sehr zufriedenstellende Erkenntnis, daß zumindest jede integrierbare Funktion (Regelfunktion) eine Stammfunktion besitzt. Diese ist in der Form eines Integrals mit fester unterer und variabler oberer Grenze gegeben. Dieser Zusammenhang zwischen Integration und Differentiation ist umso bemerkenswerter, wenn man sich die höchst unterschiedlichen Ausgangspositionen für den Begriff der Stammfunktion und des Integrals vor Augen hält.

2) Für die Menge aller Stammfunktionen einer Regelfunktion f,

$$\{x \mapsto \Phi(x) = \int\limits_a^x f(t)\,dt + c\colon c \in \mathbb{K}\}$$

ist auch das Symbol

$$\int f(x)\,dx \text{ oder } \int f\,dx \text{ oder } \int f$$

gebräuchlich, welches man **unbestimmtes Integral** von f nennt. Der Einfachheit halber wird dasselbe Symbol auch zur Bezeichnung irgendeiner Stammfunktion benützt, wie etwa in

$$\int \sin x\,dx = -\cos x, \qquad \int e^{\alpha x}\,dx = \frac{1}{\alpha}e^{\alpha x}, \quad \alpha \neq 0.$$

Diese Bezeichnung führt zu keinen Mißverständnissen, wenn man berücksichtigt, daß mit $\int f(x)\,dx = F$ auch $\int f(x)\,dx = F + c$, $c \in \mathbb{K}$, gilt. Hat man auf einem Intervall I die Beziehungen $\int f\,dx = F$ und $\int f\,dx = G$ gefunden, so ist es nicht zulässig, auf $F = G$ zu schließen. Vielmehr gilt dann $F - G = c$ für ein $c \in \mathbb{K}$.
3) Die Ableitung einer differenzierbaren Funktion $F\colon I \to \mathbb{K}$ muß keine Regelfunktion sein. Als Beispiel betrachte man $F\colon [-1,1] \to \mathbb{R}$, gegeben durch

$$F(x) = \begin{cases} x^2 \sin\frac{1}{x}, & x \neq 0, \\ 0, & x = 0. \end{cases}$$

F ist überall differenzierbar, aber in $x = 0$ besitzt F' weder einen rechts-, noch eine linksseitigen Grenzwert. F' ist also keine Regelfunktion auf I. Insbesondere kann daher eine Stammfunktion von F' *nicht* durch eine Integration (nach Cauchy) gefunden werden.
4) Mit dem Hauptsatz läßt sich die oft mühsame Berechnung eines unbestimmten Integrals über die Definition zurückführen auf das Auffinden einer Stammfunktion. Der Hauptsatz ermöglicht es auch, die Produktregel und die Kettenregel aus der Differentialrechnung in häufig benutzte Integrationsregeln umzusetzen. Um die Beziehung $\int u'\,dx = u$ zu haben, formulieren wir diese Regeln für *stetig* differenzierbare Funktionen:

4.1. Partielle Integration.

Theorem VII-4.5. *Es sei $f \in C(I,\mathbb{K})$, $g \in C^1(I,\mathbb{K})$ und F eine Stammfunktion von f. Dann gilt*

$$\int\limits_a^b f(x)g(x)\,dx = F(x)g(x)\Big|_a^b - \int\limits_a^b F(x)g'(x)\,dx \qquad \textit{(partielle Integration).}$$

Beweis. Es gilt $F \in C^1(I, \mathbb{K})$ und somit $Fg \in C^1(I, \mathbb{K})$. Nach der Produktregel gilt

$$(Fg)' = F'g + Fg' = fg + Fg'.$$

Somit ist Fg eine Stammfunktion von $fg + Fg'$ und es folgt

$$(Fg)(b) - (Fg)(a) = \int\limits_a^b f(x)g(x)\,dx + \int\limits_a^b F(x)g'(x)\,dx.$$

Dies ist äquivalent zur Behauptung. □

Häufig wird die Regel für die partielle Integration in der einprägsameren Form

$$\int\limits_a^b u'v\,dx = uv|_a^b - \int\limits_a^b uv'\,dx$$

bzw. als unbestimmtes Integral

$$\int u'v\,dx = uv - \int uv'\,dx$$

formuliert.

Beispiel VII-4.6. 1) $\int x^\alpha \ln x\,dx = \frac{x^{\alpha+1}}{(\alpha+1)^2}((\alpha+1)\ln x - 1)$, $\alpha \neq -1$.
Wir definieren

$$f(x) = x^\alpha \qquad g(x) = \ln x$$

(bzw. $u'(x) = x^\alpha$, $v = \ln x$, für $x > 0$). Eine Stammfunktion von f ist gegeben durch $F(x) = \frac{x^{\alpha+1}}{\alpha+1}$. Wir erhalten somit

$$\begin{aligned}\int x^\alpha \ln x\,dx &= \frac{x^{\alpha+1}}{\alpha+1}\ln x - \int \frac{x^{\alpha+1}}{\alpha+1}\cdot\frac{1}{x}\,dx\\ &= \frac{x^{\alpha+1}}{\alpha+1}\ln x - \frac{x^{\alpha+1}}{(\alpha+1)^2}.\end{aligned}$$

2) $\int \sqrt{1-x^2}\,dx = \frac{1}{2}(x\sqrt{1-x^2} + \arcsin x)$, $x \in [-1, 1]$.
Für $|x| \leq 1 - \varepsilon$, $0 < \varepsilon < 1$ setzen wir

$$f(x) = 1, \qquad g(x) = \sqrt{1-x^2}$$

(bzw. $u'(x) = 1$, $v(x) = \sqrt{1-x^2}$), und erhalten

$$\begin{aligned}\int 1\cdot\sqrt{1-x^2}\,dx &= x\sqrt{1-x^2} - \int x\cdot\frac{(-x)}{\sqrt{1-x^2}}\,dx\\ &= x\sqrt{1-x^2} - \int \frac{1-x^2-1}{\sqrt{1-x^2}}\,dx\\ &= x\sqrt{1-x^2} + \int \frac{dx}{\sqrt{1-x^2}} - \int \sqrt{1-x^2}\,dx,\end{aligned}$$

d.h.

$$2\int \sqrt{1-x^2}\,dx = x\sqrt{1-x^2} + \int \frac{dx}{\sqrt{1-x^2}}.$$

Nach Beispiel VII-3.2 ist $\arcsin x$ eine Stammfunktion von $\frac{1}{\sqrt{1-x^2}}$ auf $(-1,1)$. Daraus folgt die behauptete Formel zunächst auf $(-1,1)$. Da $f\colon x \mapsto \sqrt{1-x^2}$ stetig auf $[-1,1]$ ist, besitzt f eine Stammfunktion auf I. Die Abbildung $x \mapsto \frac{1}{2}(x\sqrt{1-x^2} + \arcsin x)$ ist auf $[-1,1]$ stetig und stellt daher auch auf $[-1,1]$ eine Stammfunktion dar.

Als Anwendung der partiellen Integration beweisen wir folgende nützliche Variante des Satzes von Taylor:

Theorem VII-4.7. *Es sei $f \in C^{n+1}(I,\mathbb{K})$ und $x_0 \in I$. Dann gilt für alle $x \in I$*

$$f(x) = \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x-x_0)^k + \int_{x_0}^{x} \frac{(x-t)^n}{n!} f^{(n+1)}(t)\,dt.$$

Beweis. Wir führen den Beweis mit vollständiger Induktion. Für $f \in C^1(I,\mathbb{K})$ ergibt der Hauptsatz VII-4.2 die Identität

$$f(x) = f(x_0) + \int_{x_0}^{x} f'(t)\,dt, \quad x \in I.$$

Dies ist gerade die Behauptung für $n = 0$. Die Behauptung gelte nun für alle C^n-Funktionen und es sei $f \in C^{n+1}(I,\mathbb{K})$. Insbesondere gilt also $f \in C^n(I,\mathbb{K})$ und daher auch

$$f(x) = \sum_{k=0}^{n-1} \frac{f^{(k)}(x_0)}{k!}(x-x_0)^k + \int_{x_0}^{x} \frac{(x-t)^{n-1}}{(n-1)!} f^{(n)}(t)\,dt, \quad x \in I.$$

$f^{(n)} \in C^1(I,\mathbb{K})$, somit können wir partiell integrieren und erhalten

$$\begin{aligned} f(x) &= \sum_{k=0}^{n-1} \frac{f^{(k)}(x_0)}{k!}(x-x_0)^k - \frac{(x-t)^n}{n!} f^{(n)}(t)\Big|_{x_0}^{x} + \int_{x_0}^{x} \frac{(x-t)^n}{n!} f^{(n+1)}(t)\,dt \\ &= \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x-x_0)^k + \int_{x_0}^{x} \frac{(x-t)^n}{n!} f^{(n+1)}(t)\,dt. \end{aligned}$$

□

4.2. Integration durch Substitution. Die Kettenregel der Differentialrechnung ermöglicht die Integration durch Substitution. Diese Technik kann in zwei unterschiedlichen Situationen eingesetzt werden:

Fall 1: Der Integrand besitzt die Form $f(g(t))g'(t)$: Als Beispiel betrachte man den Integranden
$$h(t) = (\sin^3 t + e^{\sin t}) \cos t,$$
welcher offensichtlich die Ableitung von
$$H(t) = \frac{1}{4} \sin^4 t + e^{\sin t}$$
ist. H ist also Stammfunktion von h. Der folgende Satz beschreibt die allgemeine Situation.

THEOREM VII-4.8 (1. Substitutionsregel). *Es sei $f \in C(I, \mathbb{K})$, $g \in C^1(J, \mathbb{R})$, $I = [a, b]$, $J = [\alpha, \beta]$ und $g(J) \subset I$. Dann gilt für alle $x, y \in J$*
$$\int\limits_x^y f(g(t))g'(t)\, dt = \int\limits_{g(x)}^{g(y)} f(u) du$$
und
$$\int f(g(t))g'(t)\, dt = \int f(u) du\Big|_{u=g(t)}.$$
Dabei bedeutet $\int f(u)du\big|_{u=g(t)}$ die Auswertung einer Stammfunktion von f an der Stelle $g(t)$.

BEWEIS. Da f stetig ist, besitzt f eine Stammfunktion $F \in C^1(I, \mathbb{K})$. Wegen $g(J) \subset \operatorname{def} F$ ist die Komposition $F \circ g$ möglich. Wir zeigen, daß $F \circ g$ eine Stammfunktion von $(f \circ g)g'$ ist: Nach der Kettenregel V-2.2 gilt für alle $t \in J$
$$(F \circ g)'(t) = F'(g(t))g'(t) = f(g(t))g'(t)$$
(die letzte Gleichheit gilt überall, weil F eine Stammfunktion von f ist und f stetig ist). Damit folgt
$$\int\limits_x^y f(g(t))g'(t)\, dt = (F \circ g)(y) - (F \circ g)(x) \; = F(g(y)) - F(g(x)) \overset{\text{S.VII-4.2}}{=} \int\limits_{g(x)}^{g(y)} f(u) du.$$
$\square$

BEISPIEL VII-4.9. 1) $\mathcal{I} = \int (\cos t + \cos^3 t)\, dt$.
Um den Satz anwenden zu können, schreiben wir das Integral in der Form
$$\mathcal{I}_t = \int (1 + \cos^2 t) \cos t\, dt = \int (2 - \sin^2 t) \cos t\, dt.$$
1. Schritt: *formale* Substitution: $g(t) = u$, $g'(t)\, dt = du$,
hier: $g(t) = \sin t = u$, $\cos t\, dt = du$,
$$\mathcal{I}_u = \int (2 - u^2) du.$$

2. Schritt: unbestimmte Integration nach u:

$$\mathcal{I}_u = 2u - \frac{1}{3}u^3.$$

3. Schritt: Rücksubstitution $u = g(t)$:

$$\mathcal{I}_t = 2\sin t - \frac{1}{3}\sin^3 t.$$

Diese formale Vorgangsweise spiegelt die Anwendung der Substitutionsregel folgendermaßen wieder: Es ist $f(g(t))g'(t) = (2-\sin^2 t)\cos t$ mit $g(t) = \sin t$ und $f(u) = 2-u^2$. Gemäß der Substitutionsregel hat man die Stammfunktion von f an der Stelle $u = \sin t$ auszuwerten.
2)

$$\mathcal{I} = \int_0^2 e^{t^2} t\,dt = \frac{1}{2}\int_0^2 e^{t^2} 2t\,dt.$$

1. Schritt: *formale* Substitution $t^2 = u$, $2t\,dt = du$.
2. Schritt: Transformation der Grenzen: $t = 0 \Rightarrow u = 0$, $t = 2 \Rightarrow u = 4$.
3. Schritt: Integration: $\mathcal{I} = \int_0^4 e^u du = e^4 - e^0 = e^4 - 1$.

Fall 2: Einsetzen einer beliebigen Substitutionsfunktion g. Dies ist der eigentliche Anwendungsbereich der Substitutionsregel. Man führt ein Integral $\int f(x)\,dx$ durch geschickte Wahl einer Substitutionsfunktion g in die Form

$$\int f(g(t))g'(t)\,dt,$$

über, in der Hoffnung, daß letzteres Integral einfacher als das Ausgangsintegral ist.

Theorem VII-4.10 (2. Substitutionsregel). *Es sei $f \in C(I, \mathbb{K})$, $g \in C^1(J, \mathbb{R})$, $J = [\alpha, \beta]$ und $g(J) \subset I$. Ferner sei g injektiv. Dann gilt für alle $u, v \in g(J)$*

$$\int_u^v f(x)\,dx = \int_{g^{-1}(u)}^{g^{-1}(v)} f(g(t))g'(t)\,dt$$

bzw.

$$\int f(x)\,dx = \int f(g(t))g'(t)\,dt\Big|_{t=g^{-1}(x)}.$$

Beweis. Wie im Beweis von Satz VII-4.8 sieht man, daß $F \circ g$ eine Stammfunktion von $(f \circ g)g'$ ist, falls F eine Stammfunktion von f ist. Für die linke Seite erhält man daher

$$\int_u^v f(x)\,dx = F(v) - F(u),$$

für die rechte

$$\int\limits_{g^{-1}(u)}^{g^{-1}(v)} f(g(t))g'(t)\,dt = (F \circ g)(g^{-1}(v)) - (F \circ g)(g^{-1}(u)) = F(v) - F(u).$$

$\square$

Beispiel VII-4.11. 1) $\int \frac{dx}{\sin x} = \ln|\tan\frac{x}{2}|$.
Ist der Integrand eine rationale Funktion in sin und cos, $R(\sin x, \cos x)$, kann man immer – soferne sich nicht andere, einfachere Methoden anbieten – folgende Substitution ansetzen:

$$x = g(u) = 2\arctan u \quad \text{bzw} \quad u = \tan\frac{x}{2}.$$

Ersetzt man in den Identitäten

$$\cos 2z = \frac{1-\tan^2 z}{1+\tan^2 z}, \quad \sin 2z = \frac{2\tan z}{1+\tan^2 z}$$

z durch $\arctan u$, erhält man

$$\cos(g(u)) = \frac{1-u^2}{1+u^2}, \quad \sin(g(u)) = \frac{2u}{1+u^2}.$$

Wegen Satz VII-4.10 gilt die Identität

$$\int R(\sin x, \cos x)\,dx = \int R(\sin(g(u)), \cos g(u))g'(u)du\big|_{u=g^{-1}(x)}$$
$$= \int R\Big(\frac{2u}{1+u^2}, \frac{1-u^2}{1+u^2}\Big)\frac{2}{1+u^2}du\big|_{u=\tan\frac{x}{2}}.$$

Umgelegt auf das konkrete Beispiel bedeutet dies

$$\int \frac{dx}{\sin x} = \int \frac{1+u^2}{2u}\frac{2}{1+u^2}du = \int \frac{du}{u} = \ln|u|\big|_{u=\tan\frac{x}{2}} = \ln|\tan\frac{x}{2}|.$$

2) $\mathcal{I} = \int\limits_{-1}^{1} \sqrt{1-x^2}\,dx = \frac{\pi}{2}$.
Dieses Integral stellt somit eine erste Verbindung zwischen der Zahl π und dem Einheitskreis her: Interpretiert man das Integral als Flächeninhalt, ergibt sich, daß $\frac{\pi}{2}$ – die kleinste positive Nullstelle des Kosinus – den Flächeninhalt eines Halbkreises mit Radius 1 angibt.
1. Schritt: Wahl der Substitutionsfunktion. Zu $f(x) = \sqrt{1-x^2}$ kann man, um die Wurzel zu eliminieren, als Substitutionsfunktion g wählen:

$$g(t) = \cos t, \quad t \in [0, \pi].$$

Man beachte, daß g auf $[0, \pi]$ streng monoton ist und $g([0, \pi]) = [-1, 1]$ gilt. Natürlich ist g stetig differenzierbar.
2. Schritt: formale Substitution: Auf $[0, \pi]$ gilt

$$f(g(t)) = \sqrt{1-\cos^2 t} = \sin t, \quad g'(t) = -\sin t.$$

Wir erhalten also das unbestimmte Integral

$$\int \sqrt{1-x^2}\,dx = -\int \sin^2 t\,dt.$$

3. Schritt: Transformation der Grenzen: $x=-1 \Rightarrow t=\arccos(-1)=\pi$, $x=1 \Rightarrow t = \arccos 1 = 0$
4. Schritt: Integration durch Auswertung der Stammfunktion von $f(g(t))g'(t)$ in $t=0$ und $t=\pi$.
Zur Bestimmung der Stammfunktion $F\circ g$ von $t \mapsto -\sin^2 t$ verwenden wir die Umformung

$$\sin^2 t = \frac{1}{2} - \frac{1}{2}\cos 2t,$$

und erhalten somit unmittelbar

$$(F\circ g)(t) = -\frac{1}{2}t + \frac{1}{4}\sin 2t.$$

Nach Satz VII-4.11 erhalten wir

$$\int_{-1}^{1} \sqrt{1-x^2}\,dx = (F\circ g)(t)\Big|_{t=\pi}^{t=0} = \frac{1}{2}\big(-t\big|_{\pi}^{0} + \frac{1}{2}\sin 2t\big|_{\pi}^{0}\big) = \frac{\pi}{2}.$$

4.3. Partialbruchzerlegung. Ausgangspunkt ist folgende **kanonische Produktdarstellung** für Polynome, welche wir ohne Beweis mitteilen.

Theorem VII-4.12. *Jedes Polynom p mit $\operatorname{grad} p = n \geq 1$ läßt sich mit Hilfe seiner paarweise verschiedenen Nullstellen $z_i \in \mathbb{C}$, $i=1,\ldots,m$, in der Form*

$$p(z) = a_n \prod_{i=1}^{m} (z-z_i)^{\nu_i}$$

darstellen, wobei $a_n \in \mathbb{C}$, $a_n \neq 0$. Die Zahlen $\nu_i \in \mathbb{N}$, $i=1,\ldots,m$, sind eindeutig bestimmt und erfüllen

$$\sum_{i=1}^{m} \nu_i = n.$$

*Man nennt ν_i die **Vielfachheit (Multiplizität)** der Nullstelle z_i.*

Beweis. Übung. □

Sind alle Koeffizienten des Polynoms $p(z) = \sum_{k=0}^{n} a_k z^k$ reell, folgt wegen $a_k = \bar{a}_k$, $k=0,\ldots,n$,

$$p(\bar{z}) = \sum_{k=0}^{n} a_k (\bar{z})^k = \sum_{k=0}^{n} a_k \overline{z^k} = \sum_{k=0}^{n} \overline{a_k z^k} = \overline{\sum_{k=0}^{n} a_k z^k} = \overline{p(z)}.$$

Mit $\xi \in \mathbb{C}$ ist somit auch $\bar{\xi} \in \mathbb{C}$ eine Nullstelle von p. Besitzt ξ die Vielfachheit μ, d.h. gilt

$$p(z) = (z-\xi)^\mu q(z), \quad q(\xi) \neq 0, \quad z \in \mathbb{C},$$

dann folgt

$$\overline{p(z)} = (\bar{z}-\bar{\xi})^\mu \overline{q(z)} = p(\bar{z}).$$

Schreibt man für $\bar{z}$ wieder z, erhält man

$$p(z) = (z-\bar{\xi})^\mu \overline{q(\bar{z})} \equiv (z-\bar{\xi})^\mu r(z),$$

also ist $\bar{\xi}$ eine Nullstelle von p (dies wissen wir bereits) mit der Vielfachheit $\mu' \geq \mu$. Aus $r(\bar{\xi}) = \overline{q(\xi)} \neq 0$ folgt $\mu' = \mu$, also stimmen die Vielfachheiten der Nullstellen ξ und $\bar{\xi}$ überein.

Wir betrachten nun den Fall $\xi = \alpha + i\beta$, $\beta \neq 0$. Für jedes *reelle* x gilt dann

$$\begin{aligned}(x-\xi)(x-\bar{\xi}) &= (x-(\alpha+i\beta))(x-(\alpha-i\beta)) = (x-\alpha)^2+\beta^2\\ &= x^2+Ax+B,\end{aligned}$$

mit

$$A = -2\alpha, \quad B = \alpha^2+\beta^2.$$

Das quadratische Polynom x^2+Ax+B besitzt also keine *reellen* Nullstellen. Zusammen mit Satz VII-4.12 erhalten wir die *reelle kanonische Produktdarstellung* von p:

Theorem VII-4.13. *Es sei* $p(z) = \sum\limits_{k=0}^{n} a_k z^k$ *ein Polynom mit* reellen *Koeffizienten* a_k, $k = 0, \dots, n$. *Dann gilt*

(1) *Komplexe Nullstellen* ξ *mit* $\operatorname{Im}\xi \neq 0$ *treten in konjugierten Paaren auf: Mit* ξ *ist auch* $\bar{\xi}$ *eine Nullstelle gleicher Multiplizität wie* ξ.

(2) *Sind* $x_1, \dots, x_r$ *alle verschiedenen reellen Nullstellen von* p *mit der Multiplizität* ρ_i, $i = 1, \dots, r$, *gilt die Darstellung*

$$p(x) = a_n \prod_{i=1}^{r}(x-x_i)^{\rho_i} \prod_{j=1}^{s}(x^2+A_jx+B_j)^{\sigma_j}, \quad x \in \mathbb{R}.$$

Die Polynome $x^2+A_jx+B_j$, $A_j, B_j \in \mathbb{R}$, $j = 1, \dots, s$, *sind paarweise verschieden und besitzen keine reellen Nullstellen. Die natürlichen Zahlen* ρ_i, $i = 1, \dots, r$, *und* σ_j, $j = 1, \dots, s$, *erfüllen*

$$\sum_{i=1}^{r}\rho_i + 2\sum_{j=1}^{s}\sigma_j = n.$$

Theorem VII-4.14 (Partialbruchzerlegung). *Es sei* $r = \frac{p}{q}$ *eine rationale Funktion mit* $\operatorname{grad} p < \operatorname{grad} q = n$. *Das Nennerpolynom habe die kanonische Produktdarstellung*

$$q(z) = a_n \prod_{j=1}^{m}(z-z_j)^{\nu_j}, \quad z \in \mathbb{C},$$

*mit $z_j \neq z_k$ für $j \neq k$. Dann besitzt r eine eindeutig bestimmte Summendarstellung (**Partialbruchzerlegung**) der Form*

$$r(z) = \sum_{i=1}^{m} \sum_{j=1}^{\nu_i} \frac{a_{ij}}{(z - z_i)^j}, \quad z \in \mathbb{C} \setminus \{z_1, \dots, z_m\},$$

mit komplexen Zahlen a_{ij}.

Beweis. 1. Existenz: Wir führen den Beweis durch Induktion nach $\operatorname{grad} q = n$. Im Fall $n = 1$ gilt $\operatorname{grad} p < 1$, d.h. p ist eine Konstante $c \in \mathbb{C}$. Es gilt also

$$r(z) = \frac{c}{a_1(z - z_1)} = \frac{a_{11}}{z - z_1} \quad \text{mit} \quad a_{11} := \frac{c}{a_1}.$$

Induktionsschritt: Es sei $n > 1$. Wir nehmen an, der Satz sei für alle rationalen Funktionen $\frac{P}{Q}$ mit $\operatorname{grad} P < \operatorname{grad} Q \leq n - 1$ bewiesen. Wir schreiben q in der Form

$$q(z) = (z - z_1)^{\nu_1} s(z) \quad \text{mit} \quad s(z) = a_n \prod_{j=2}^{m} (z - z_j)^{\nu_j}.$$

Wegen $z_1 \neq z_i$, $i = 2, \dots, n$, ist $s(z_1) \neq 0$. Für $a \in \mathbb{C}$ berechnen wir

$$\frac{p(z)}{q(z)} - \frac{a}{(z - z_1)^{\nu_1}} = \frac{p(z) - as(z)}{(z - z_1)^{\nu_1} s(z)}.$$

Wählt man speziell $a = \frac{p(z_1)}{s(z_1)}$, gilt $p(z_1) - as(z_1) = 0$. Gilt sogar

$$p(z) - as(z) = 0 \quad \text{für } \textit{alle } z \in \mathbb{C},$$

folgt

$$r(z) = \frac{a}{(z - z_1)^{\nu_1}}$$

und wir sind fertig. Ist jedoch $p - as$ nicht das Nullpolynom, kann man nach Satz VII-4.12 den Linearfaktor $z - z_1$ abspalten, d.h. es gibt ein Polynom P mit

$$\begin{gathered} p(z) - as(z) = (z - z_1) P(z) \\ \operatorname{grad} P \leq \max\{\operatorname{grad} p, \operatorname{grad} s\} - 1 \leq n - 2. \end{gathered}$$

Somit folgt

$$r(z) = \frac{a}{(z - z_1)^{\nu_1}} + \frac{P(z)}{Q(z)},$$

mit

$$Q(z) = (z - z_1)^{\nu_1 - 1} s(z) = (z - z_1)^{\nu_1 - 1} a_n \prod_{j=2}^{m} (z - z_j)^{\nu_j},$$

also

$$\operatorname{grad} Q = n - 1.$$

Nach Induktionsvoraussetzung besitzt $\frac{P}{Q}$ eine Partialbruchzerlegung

$$\frac{P(z)}{Q(z)} = \sum_{j=1}^{\nu_1 - 1} \frac{a_{1j}}{(z-z_1)^j} + \sum_{i=2}^{m} \sum_{j=1}^{\nu_i} \frac{a_{ij}}{(z-z_i)^j},$$

und daher mit $a_{1\nu_1} = a$

$$r(z) = \sum_{i=1}^{m} \sum_{j=1}^{\nu_i} \frac{a_{ij}}{(z-z_i)^j}.$$

2. Eindeutigkeit. Es genügt zu zeigen, daß sämtliche Koeffizienten der Zerlegung von $r_0(z) \equiv 0$ verschwinden. Angenommen es ist

$$(*) \qquad r_0(z) = \sum_{i=1}^{m} \sum_{j=1}^{\nu_i} \frac{a_{ij}}{(z-z_i)^j} = 0, \quad z \in \mathbb{C} \setminus \{z_1, \ldots, z_m\}.$$

Wir betrachten eine feste Nullstelle z_{i_0} und wählen $1 \leq j_0 \leq \nu_{i_0}$ so, daß $a_{i_0 j} = 0$ für $j > j_0$ zutrifft. Multipliziert man $(*)$ mit $(z - z_{i_0})^{j_0}$, erhält man

$$0 = \left(\sum_{\substack{i=1 \\ i \neq i_0}}^{\nu_i} \sum_{j=1}^{\nu_i} \frac{a_{ij}}{(z-z_i)^j} \right) (z - z_{i_0})^{j_0} + \sum_{j=1}^{j_0} (z - z_{i_0})^{j_0 - j} a_{i_0 j}.$$

Durch Grenzübergang $z \to z_{i_0}$ erhält man $a_{i_0 j_0} = 0$. Schrittweise ergibt sich auf diese Weise für sämtliche Koeffizienten $a_{ij} = 0$. □

Für die praktische Berechnung der Partialbruchzerlegung gibt es mehrere Möglichkeiten:

1. **Koeffizientenvergleich:** Multipliziert man die Darstellung von $r = \frac{p}{q}$ mit q, erhält man

$$\begin{aligned} p(z) &= \sum_{i=1}^{m} \left(\sum_{j=1}^{\nu_i} \frac{a_{ij}}{(z-z_i)^j} \right) a_n \prod_{k=1}^{m} (z-z_k)^{\nu_k} \\ &= a_n \sum_{i=1}^{m} \sum_{j=1}^{\nu_i} a_{ij} (z-z_i)^{\nu_i - j} \prod_{\substack{k=1 \\ k \neq i}}^{m} (z-z_k)^{\nu_k}, \quad z \in \mathbb{C}. \end{aligned}$$

Der Identitätssatz IV-4.12 ist nun die Grundlage für den Koeffizientenvergleich, bei dem man die Koeffizienten gleicher Potenzen von z^k in p und im Polynom auf der rechten Seite gleichsetzt. Dies führt auf ein lineares Gleichungssystem für die Unbekannten a_{ij}, welches wegen der Existenz- und Eindeutigkeitsaussage in Satz VII-4.14 genau eine Lösung besitzen *muß*.

2. **Substitutionsmethode:** Ein anderes lineares Gleichungssystem für a_{ij} läßt sich gewinnen, wenn man für z nacheinander n verschiedene und möglichst zweckmäßig gewählte Werte $\xi_1, \ldots, \xi_n$ einsetzt.

3. **Grenzwertmethode:** Die „höchsten“ Koeffizienten $a_{i\nu_i}$ erhält man besonders einfach, indem man die Darstellung für $r(z)$ mit $(z-z_i)^{\nu_i}$ multipliziert und den Grenzwert $z \to z_i$ betrachtet (in der Praxis bedeutet dies natürlich, daß man nach der Multiplikation den gekürzten Ausdruck in $z = z_i$ auswertet). Hat man $a_{i\nu_i}$ berechnet, kann man den Term $\frac{a_{i\nu_i}}{(z-z_i)^{\nu_i}}$ auf die linke Seite bringen und das Verfahren mit dem nächstniedrigeren Koeffizienten $a_{i\nu_i-1}$ wiederholen.

Beispiel VII-4.15. 1) $r(z) = \frac{z+1}{z(z-1)(z-i)(z+i)}$.
Da sämtliche Nullstellen des Nennerpolynoms Vielfachheit 1 haben, führt folgender Ansatz zum Ziel ($a_{11} = A$, $a_{21} = B$, $a_{31} = C$, $a_{41} = D$):

$$(*) \qquad r(z) = \frac{A}{z} + \frac{B}{z-1} + \frac{C}{z-i} + \frac{D}{z+i}.$$

a) Koeffizientenvergleich:
Wir multiplizieren $(*)$ mit $z(z-1)(z-i)(z+i)$ und erhalten die Identität

$$\begin{aligned}
z+1 &= A(z-1)(z^2+1) + Bz(z^2+1) + Cz(z-1)(z+i) + Dz(z-1)(z-i) \\
&= A(z^3-z^2+z-1) + B(z^3+z) + C(z^3+(i-1)z^2-iz) \\
&\quad + D(z^3-(1+i)z^2+iz) \\
&= (A+B+C+D)z^3 + (-A+(i-1)C-(1+i)D)z^2 \\
&\quad + (A+B-iC+iD)z - A.
\end{aligned}$$

Vergleicht man die Koeffizienten gleicher Potenzen, erhält man das Gleichungssystem

$$\begin{array}{rrrrcl}
A & +B & +C & +D & = & 0 \\
-A & & +(i-1)C & -(1+i)D & = & 0 \\
A & +B & -iC & +iD & = & 1 \\
-A & & & & = & 1\ .
\end{array}$$

b) Substitutionsmethode:
Setzt man nacheinander in $(*)$ etwa $z = -1, -2, 2, 3$ ein, erhält man das Gleichungssystem

$$\begin{array}{rrcrrrr}
z = & -1\colon & 0 & = & -A & -\frac{1}{2}B & -\frac{1}{1+i}C & +\frac{1}{-1+i}D \\
z = & -2\colon & \frac{-1}{30} & = & -\frac{1}{2}A & -\frac{1}{3}B & -\frac{1}{2+i}C & +\frac{1}{-2+i}D \\
z = & 2\colon & \frac{3}{10} & = & \frac{1}{2}A & +B & +\frac{1}{2-i}C & +\frac{1}{2+i}D \\
z = & 3\colon & \frac{4}{60} & = & \frac{1}{3}A & +\frac{1}{2}B & +\frac{1}{3-i}C & +\frac{1}{3+i}D\ .
\end{array}$$

c) Grenzwertmethode:
Wir multiplizieren $(*)$ mit z. Dies ergibt

$$\frac{z+1}{(z-1)(z-i)(z+i)} = A + z\left(\frac{B}{z-1} + \frac{C}{z-i} + \frac{D}{z+i}\right).$$

Für $z = 0$ erhält man $A = -1$. Multipliziert man mit $z-1$, folgt

$$\frac{z+1}{z(z-i)(z+i)} = B + (z-1)\left(-\frac{1}{z} + \frac{C}{z-i} + \frac{D}{z+i}\right),$$

für $z = 1$ ergibt sich $B = 1$. Ganz entsprechend findet man $C = \frac{i}{2}$, $D = -\frac{i}{2}$ und damit die Partialbruchzerlegung

$$\frac{z+1}{z(z-1)(z-i)(z+i)} = -\frac{1}{z} + \frac{1}{z-1} + \frac{1}{2}\frac{i}{z-i} - \frac{1}{2}\frac{i}{z+i}.$$

2) $r(z) = \frac{z^2+1}{z(z-1)^2}$.
1 ist eine zweifache Nullstelle des Nennerpolynoms, somit ist ($a_{11} = A$, $a_{21} = B$, $a_{22} = C$)

$$r(z) = \frac{A}{z} + \frac{B}{z-1} + \frac{C}{(z-1)^2}$$

anzusetzen. Multipliziert man mit z und wertet in $z = 0$ aus, erhält man $A = 1$. Multiplikation mit $(z-1)^2$ liefert

$$\frac{z^2+1}{z} = \frac{(z-1)^2}{z} + B(z-1) + C.$$

Für $z = 1$ erhält man also $C = 2$. Setzt man nun in

$$r(z) = \frac{z^2+1}{z(z-1)^2} = \frac{1}{z} + \frac{B}{z-1} + \frac{2}{(z-1)^2}$$

$z = 2$ ein, erhält man

$$\frac{5}{2} = \frac{1}{2} + B + 2,$$

also $B = 0$. Die gesuchte Partialbruchzerlegung lautet somit

$$r(z) = \frac{1}{z} + \frac{2}{(z-1)^2}.$$

Das Beispiel zeigt, daß es vorteilhaft sein kann, verschiedene Methoden zur Berechnung der Partialbruchzerlegungen zu kombinieren.

Ist die rationale Funktion reell, d.h. sind alle Koeffizienten der Polynome p und q reell, ist es z.B. für Anwendungen in der Integralrechnung oft zweckmäßig, eine *reelle* Partialbruchzerlegung zur Verfügung zu haben.

Theorem VII-4.16. *Es sei $r = \frac{p}{q}$ eine rationale Funktion mit* $\operatorname{grad} p < \operatorname{grad} q = n$. *Sämtliche Koeffizienten der Polynome p und q seien reell. Das Nennerpolynom besitze die reelle kanonische Produktdarstellung*

$$q(x) = a_n \prod_{i=1}^{r}(x-x_i)^{\rho_i} \prod_{j=1}^{s}(x^2 + A_j x + B_j)^{\sigma_j}, \quad x \in \mathbb{R}$$

(vgl. Satz VII-4.13). Dann besitzt r eine eindeutig bestimmte Partialbruchzerlegung der Form

$$r(x) = \sum_{i=1}^{r} \sum_{j=1}^{\rho_i} \frac{a_{ij}}{(x - x_i)^j} + \sum_{i=1}^{s} \sum_{j=1}^{\sigma_i} \frac{\alpha_{ij} x + \beta_{ij}}{(x^2 + A_i x + B_i)^j}, \quad x \in \mathbb{R},$$

wobei die Koeffizienten a_{ij}, α_{ij}, β_{ij} reelle Zahlen sind.

Beispiel VII-4.17. $r(x) = \frac{x+1}{x(x-1)(x^2+x+1)}, x \in \mathbb{R} \setminus \{0, 1\}$.
Die einzigen reellen Nullstellen von $q(x) = x(x-1)(x^2 + x + 1)$, $x = 0$ und $x = 1$, besitzen Multiplizität 1. Somit lautet der Ansatz für die reelle Partialbruchzerlegung

$$r(x) = \frac{A}{x} + \frac{B}{x-1} + \frac{Cx + D}{x^2 + x + 1}.$$

Die Koeffizienten A und B erhält man leicht mit Hilfe der Grenzwertmethode. Man findet

$$A = -1, \quad B = \frac{2}{3}.$$

Es ist also

$$\frac{x+1}{x(x-1)(x^2+x+1)} = -\frac{1}{x} + \frac{2}{3}\frac{1}{x-1} + \frac{Cx + D}{x^2 + x + 1}.$$

Setzt man die Werte $x = -1$ und $x = 2$ ein, erhält man das Gleichungssystem

$$0 = 1 - \frac{1}{3} - C + D$$
$$\frac{3}{14} = -\frac{1}{2} + \frac{2}{3} + \frac{2}{7}C + \frac{1}{7}D,$$

woraus $C = \frac{1}{3}$ und $D = -\frac{1}{3}$ folgt. Insgesamt gilt also die Darstellung

$$f(x) = -\frac{1}{x} + \frac{2}{3}\frac{1}{x-1} + \frac{1}{3}\frac{x-1}{x^2 + x + 1}.$$

Die reelle bzw. komplexe Partialbruchzerlegung ermöglicht die systematische, geschlossene Integration rationaler Funktionen $r(x) = \frac{p(x)}{q(x)}$. Im Falle $\operatorname{grad} p \geq \operatorname{grad} q$ ist vorerst der polynomiale Anteil in r durch Division von p durch q abzuspalten. Wir demonstrieren die Methode an einigen einfachen Beispielen:

Beispiel VII-4.18. 1) Mit Hilfe von Beispiel VII-4.15-1) finden wir

$$\int \frac{(x+1)\,dx}{x(x-1)(x^2+1)} = -\int \frac{dx}{x} + \int \frac{dx}{x-1} + \frac{i}{2} \int \left(\frac{1}{x-i} - \frac{1}{x+i} \right) dx$$
$$= -\ln|x| + \ln|x-1| + \frac{i}{2} \int \frac{2i}{x^2+1}\,dx$$
$$= -\ln|x| + \ln|x-1| - \arctan x.$$

(Wir haben $\frac{1}{x-i} - \frac{1}{x+i}$ durch $\frac{2i}{x^2+1}$ ersetzt, da wir den natürlichen Logarithmus vorerst nur für *reelle* Argumente definiert haben).

2) Aus Beispiel VII-4.15-2) folgt

$$\int \frac{x^2+1}{x(x-1)^2}\,dx = \int \frac{dx}{x} + 2\int \frac{dx}{(x-1)^2} = \ln|x| - \frac{2}{x-1}.$$

3) Wir benützen Beispiel VII-4.17 in

$$\int \frac{x+1}{x(x-1)(x^2+x+1)}\,dx = -\int \frac{dx}{x} + \frac{2}{3}\int \frac{dx}{x-1} + \frac{1}{3}\int \frac{x-1}{x^2+x+1}\,dx.$$

Lediglich das letzte Integral erfordert etwas Aufwand:

$$\begin{aligned}\int \frac{x-1}{x^2+x+1}\,dx &= \frac{1}{2}\int \frac{2x+1}{x^2+x+1}\,dx - \frac{3}{2}\int \frac{dx}{x^2+x+1}\\ &= \frac{1}{2}\ln(x^2+x+1) - \frac{3}{2}\int \frac{dx}{x^2+x+1}\end{aligned}$$

(man beachte $x^2+x+1>0$ für $x \in \mathbb{R}$). Den Integranden im letzten Integral formt man um zu

$$x^2+x+1 = (x+\frac{1}{2})^2 + \frac{3}{4} = \frac{3}{4}\big((\frac{2}{\sqrt{3}}(x+\frac{1}{2}))^2+1\big).$$

Substituiert man $g(x) = \frac{2}{\sqrt{3}}(x+\frac{1}{2})$, erhält man

$$\int \frac{dx}{x^2+x+1} = \frac{2}{\sqrt{3}}\arctan(\frac{2}{\sqrt{3}}(x+\frac{1}{2})).$$

Insgesamt ergibt sich also

$$\begin{aligned}\int \frac{x+1}{x(x-1)(x^2+x+1)}\,dx = &-\ln|x| + \frac{2}{3}\ln|x-1| + \frac{1}{6}\ln(x^2+x+1)\\ &-\sqrt{3}\arctan(\frac{2}{\sqrt{3}}(x+\frac{1}{2})).\end{aligned}$$

Bemerkung VII-4.19. Im Prinzip ist es nicht notwendig, für die Integration einer *reellen* rationalen Funktion, die *reelle* Partialbruchzerlegung aus Satz VII-4.16 zu verwenden. Man kann natürlich auch die komplexe Partialbruchzerlegung zur Integration heranziehen und erhält

$$\begin{aligned}\int r(x)\,dx &= \sum_{k=1}^{m}\sum_{j=1}^{\nu_k}\int \frac{a_{kj}}{(x-z_k)^j}\,dx\\ &= \sum_{k=1}^{m}\left[\int \frac{a_{k1}}{x-z_k}\,dx - \sum_{j=2}^{\nu_k}\frac{a_{kj}}{j-1}(x-z_k)^{-j+1}\right].\end{aligned}$$

Problematisch ist nur die Berechnung der Stammfunktion von $(x-z_k)^{-1}$ für $z_k \in \mathbb{C}\backslash\mathbb{R}$, da die Logarithmen nur für reelle Argumente definiert wurden. Es sei $z_k = \alpha_k + i\beta_k$, $\beta_k \neq 0$. Mit der Umformung

$$\frac{1}{x-z_k} = \frac{x-\alpha_k+i\beta_k}{(x-\alpha_k)^2+\beta_k^2}$$

erhalten wir

$$\begin{aligned}\int \frac{dx}{x-z_k} &= \int \frac{x-\alpha_k}{(x-\alpha_k)^2+\beta_k^2}\,dx + i\beta_k \int \frac{dx}{(x-\alpha_k)^2+\beta_k^2} \\ &= \frac{1}{2}\int \frac{2(x-\alpha_k)}{(x-\alpha_k)^2+\beta_k^2}\,dx + i\beta_k \int \frac{dx}{(x-\alpha_k)^2+\beta_k^2} \\ &= \frac{1}{2}\ln((x-\alpha_k)^2+\beta_k^2) + i\arctan\frac{x-\alpha_k}{\beta_k}.\end{aligned}$$

Der gesamte Ausdruck für $\int r(x)\,dx$ ist natürlich wieder reell. Allgemein sind jedoch die *reellen* Gleichungssysteme, auf welche die reelle Partialbruchzerlegung führt, händisch leichter zu lösen, als Gleichungssysteme mit komplexen Koeffizienten.

5. Integration und Grenzübergang

In diesem Abschnitt untersuchen wir die Vertauschbarkeit von Limes und Integral. Diese ist im allgemeinen *nicht* gegeben:

Beispiel VII-5.1. $f_n(x) = \begin{cases} n & x \in [\frac{1}{2n}, \frac{1}{n}] \\ 0 & x \in [0,1] \setminus [\frac{1}{2n}, \frac{1}{n}] \end{cases}.$

Es gilt $\lim f_n(x) = f(x) = 0$ für alle $x \in [0,1]$ und $\int_0^1 f_n(x)\,dx = \frac{1}{2}$, $n \in \mathbb{N}$, und $\int_0^1 f(x)\,dx = 0$. Man beachte, daß in diesem Beispiel die Folge (f_n) nur punktweise gegen f konvergiert und darüber hinaus $\{\|f_n\| \colon n \in \mathbb{N}\}$ unbeschränkt ist.

In Hinblick auf die Definition des Cauchy Integrals ist es jedoch nicht weiter überraschend, daß die Integrale einer gleichmäßig konvergenten Folge von Regelfunktionen gegen das Integral der Grenzfunktion konvergieren (diese Eigenschaft ist gewissermaßen von Haus aus in das Integral eingebaut):

Theorem VII-5.2. *Es sei $(f_n)_{n\geq 1} \subset \mathcal{R}(I,\mathbb{K})$, $f\colon I \to \mathbb{K}$, und $f_n \rightrightarrows f$. Dann ist f eine Regelfunktion, die Folge der Integrale $(\mathcal{I}(f_n))$ ist konvergent und es gilt*

$$\lim_{n\to\infty} \int_a^b f_n(x)\,dx = \int_a^b \lim_{n\to\infty} f_n(x)\,dx = \int_a^b f(x)\,dx.$$

Beweis. Wegen Satz VII-1.8 gilt $f \in \mathcal{R}(I,\mathbb{K})$. Aus der Beschränktheit des Integrals Satz VII-2.9 folgt

$$|\mathcal{I}(f_n) - \mathcal{I}(f)| = |\mathcal{I}(f_n - f)| \leq (b-a)\|f_n - f\|.$$

□

Korollar VII-5.3. Es sei $(f_n) \subset \mathcal{R}(I,\mathbb{K})$ und die Reihe $\sum_{n=0}^{\infty} f_n$ gleichmäßig konvergent. Dann gilt

$$\sum_{n=0}^{\infty} \int_a^b f_n(x)\,dx = \int_a^b \sum_{n=0}^{\infty} f_n(x)\,dx.$$

Dieses Ergebnis eröffnet eine weitere Möglichkeit, Potenzreihenentwicklungen zu berechnen:

Beispiel VII-5.4. Aus der Binomialreihe (Beispiel VI-7.13) ergibt sich für $|t| < 1$ die Reihenentwicklung

$$(1-t^2)^{-\frac{1}{2}} = \sum_{k=0}^{\infty} (-1)^k \binom{-\frac{1}{2}}{k} t^{2k}.$$

Auf kompakten Teilintervallen von $(-1,1)$ ist die Potenzreihe gleichmäßig konvergent. Auf $[0,x]$ bzw. $[x,0]$, $|x|<1$, kann nach Korollar VII-5.3 die Reihe gliedweise integriert werden: Es gilt somit

$$\int_0^x \frac{dt}{\sqrt{1-t^2}} = \sum_{k=0}^{\infty} (-1)^k \binom{-\frac{1}{2}}{k} \int_0^x t^{2k}\,dt = \sum_{k=0}^{\infty} (-1)^k \binom{-\frac{1}{2}}{k} \frac{x^{2k+1}}{2k+1}.$$

Die Stammfunktion auf der linken Seite ist $\arcsin x$. Die rechte Seite ergibt mit $(k \geq 1)$

$$\binom{-\frac{1}{2}}{k} = \frac{1}{k!} \prod_{j=0}^{k-1} \left(-\frac{1}{2} - j\right) = \frac{1}{k!}(-1)^k \frac{1}{2^k} \prod_{j=0}^{k-1} (1+2j)$$

die Reihenentwicklung

$$\arcsin x = x + \sum_{k=1}^{\infty} \frac{1}{k!} \frac{1}{2^k} \frac{1}{2k+1} \prod_{j=0}^{k-1} (1+2j) x^{2k+1},$$

die zumindest für $|x| < 1$ konvergiert. Tatsächlich läßt sich zeigen, daß diese Darstellung sogar für $|x| \leq 1$ gültig ist.

Unser nächstes Ziel ist es, Satz VII-5.2 erheblich zu verallgemeinern. Dazu benötigen wir zwei technische Resultate. Im folgenden wollen wir unter einer **Elementarmenge** eine endliche Vereinigung von disjunkten, beschränkten Intervallen verstehen. Ist $E \subset \mathbb{R}$ eine Elementarmenge, d.h. ist

$$E = \bigcup_{i=1}^{n} J_i,$$

$J_i \subset \mathbb{R}$ beschränktes Intervall, $i = 1, \ldots, n$, und bezeichnet $\lambda(J)$ die Länge eines Intervalls J, dann nennen wir

$$\lambda(E) = \sum_{i=1}^{n} \lambda(J_i)$$

Länge von E. Wir übergehen hier den Nachweis, daß $\lambda(E)$ unabhängig ist von der speziellen Darstellung $\bigcup_{i=1}^{n} J_i$ von E. Ohne Beweis zitieren wir folgendes Resultat:

Lemma VII-5.5. *Es sei $(E_n)_{n\geq 1}$ eine Folge von Elementarmengen mit $E_n \subset I = [a, b]$, $n \in \mathbb{N}$, und es existiere $c > 0$ so, daß*

$$\forall n \in \mathbb{N}\colon \lambda(E_n) \geq c > 0.$$

Dann gibt es ein Element $x_0 \in I$, das zu unendlich vielen Elementarmengen gehört.

Lemma VII-5.6. *Es sei $t \in \mathcal{T}(I, \mathbb{K})$, $I = [a, b]$. Gilt für ein $\varepsilon > 0$*

$$\left|\int_a^b t(x)\,dx\right| \geq \varepsilon,$$

dann ist die Menge

$$E = \{x \in I\colon |t(x)| \geq \frac{\varepsilon}{2(b-a)}\}$$

eine Elementarmenge, für deren Länge die Abschätzung

$$\lambda(E) \geq \frac{\varepsilon}{2\|t\|}$$

zutrifft.

Beweis. Es sei $\mathcal{Z}\colon a = x_0 < \cdots < x_i < \cdots < x_n = b$ eine zu t gehörende Zerlegung von I und es sei c_i der Wert von t auf $J_i = (x_{i-1}, x_i)$. Dann ist klar, daß E eine endliche Vereinigung von disjunkten Intervallen ist (es können nur Intervalle des Typs J_i oder $[x_i, x_i]$ auftreten). Es gilt

$$\int_a^b t(x)\,dx = \sum_{i=1}^n c_i(x_i - x_{i-1}) = \sum_{i=1}^n c_i \lambda(J_i).$$

Wir spalten nun die Summe folgendermaßen auf:

$$\sum_{i=1}^n c_i\lambda(J_i) = \sum_{|c_i| \geq \frac{\varepsilon}{2(b-a)}} c_i\lambda(J_i) + \sum_{|c_i| < \frac{\varepsilon}{2(b-a)}} c_i\lambda(J_i).$$

Dies ergibt die Abschätzung

$$\begin{aligned}\varepsilon \leq \left|\int_a^b t(x)\,dx\right| &\leq \sum_{|c_i| \geq \frac{\varepsilon}{2(b-a)}} |c_i|\lambda(J_i) + \sum_{|c_i| < \frac{\varepsilon}{2(b-a)}} |c_i|\lambda(J_i) \\ &\leq \|t\| \cdot \lambda(E) + \frac{\varepsilon}{2(b-a)} \cdot (b-a)\end{aligned}$$

und somit (man beachte $t \neq 0$ und daher auch $\|t\| > 0$)

$$\lambda(E) \geq \frac{\varepsilon}{2\|t\|}.$$

□

Theorem VII-5.7. *Es sei $(t_n)_{n\geq 1} \subset \mathcal{T}(I, \mathbb{K})$ mit*

i) $\lim_{n\to\infty} t_n(x) = 0$ *für alle* $x \in I$ *(t_n konvergiert punktweise gegen* 0*),*
ii) $\exists c > 0\,\forall n \in \mathbb{N}\colon \|t_n\| \le c$ *((t_n) ist beschränkt).*

Dann gilt

$$\lim_{n\to\infty} \int_a^b t_n(x)\,dx = 0.$$

Beweis. Angenommen, die Folge $(\mathcal{I}(t_n))$ wäre keine Nullfolge, dann gäbe es ein $\varepsilon_0 > 0$ und eine Teilfolge $(\mathcal{I}(t_{\varphi(n)}))$ mit $|\mathcal{I}(t_{\varphi(n)})| \ge \varepsilon_0$, $n \in \mathbb{N}$. Wir bezeichnen der Einfachheit halber nun diese Teilfolge wieder mit $\mathcal{I}(t_n)$ und gehen daher von der Annahme aus:

$$\forall n \in \mathbb{N}\colon |\mathcal{I}(t_n)| \ge \varepsilon_0.$$

Nach Lemma VII-5.6 besitzt für jedes t_n die Elementarmenge

$$E_n = \{x \in I\colon |t_n(x)| \ge \frac{\varepsilon_0}{2(b-a)}\},$$

die Länge

$$\lambda(E_n) \ge \frac{\varepsilon_0}{2\|t_n\|}.$$

Wegen $\|t_n\| \le c$, $n \in \mathbb{N}$, gilt daher

$$\lambda(E_n) \ge \frac{\varepsilon_0}{2c}, \quad n \in \mathbb{N}.$$

Nach Lemma VII-5.5 gibt es daher eine Stelle $x_0 \in I$, welche in unendlich vielen Elementarmengen E_n liegt. Für unendlich viele $n \in \mathbb{N}$ gilt also

$$|t_n(x_0)| \ge \frac{\varepsilon_0}{2(b-a)},$$

im Widerspruch zur Voraussetzung $\lim_{n\to\infty} t_n(x_0) = 0$. □

Theorem VII-5.8 (Arzela-Osgood). *Die Folge $(f_n) \subset \mathcal{R}(I,\mathbb{K})$ konvergiere punktweise gegen die Regelfunktion f. Ferner sei $\{\|f_n\|\colon n \in \mathbb{N}\}$ beschränkt. Dann gilt*

$$\lim_{n\to\infty} \int_a^b f_n(x)\,dx = \int_a^b f(x)\,dx = \int_a^b \lim_{n\to\infty} f_n(x)\,dx.$$

Wir bemerken, daß aus der *punktweisen* Konvergenz von f_n gegen eine Funktion f nicht gefolgert werden kann, daß f eine Regelfunktion ist. Es ist daher notwendig, $f \in \mathcal{R}(I,\mathbb{K})$ vorauszusetzen.

Beweis. Da f nach Voraussetzung eine Regelfunktion ist, gilt auch $f_n - f \in \mathcal{R}(I,\mathbb{K})$. Da $(f - f_n)$ eine Nullfolge und $(\|f - f_n\|)$ beschränkt ist, genügt es wegen der Linearität des Integrals den Satz für den Spezialfall

$$(*) \qquad \lim_{n\to\infty} f_n(x) = 0, \quad x \in I,$$

zu beweisen. Zu jeder Regelfunktion f_n wählen wir – wie im Beweis von Satz VII-1.8 – eine Treppenfunktion $t_n \in \mathcal{T}(I, \mathbb{K})$ mit

$$\|t_n - f_n\| \leq \frac{1}{n}, \quad n \in \mathbb{N}.$$

Daraus folgt die Abschätzung

$$\|t_n\| \leq \|f_n\| + \|f_n - t_n\| < \|f_n\| + \frac{1}{n},$$

aus welcher wir die Beschränktheit von $(\|t_n\|)$ ablesen. Insbesondere folgt aber auch für jedes $x \in I$

$$|t_n(x)| \leq |f_n(x) - t_n(x)| + |f_n(x)| \leq \frac{1}{n} + |f_n(x)|.$$

Somit konvergiert auch die Folge von Treppenfunktionen (t_n) punktweise gegen die Nullfunktion. Aus Satz VII-5.7 folgt daher

$$\lim_{n\to\infty} \mathcal{I}(t_n) = 0.$$

Zusammen mit der Beschränktheit des Integrals ergibt sich daraus die Behauptung

$$\begin{aligned}|\mathcal{I}(f_n)| &\leq |\mathcal{I}(t_n)| + |\mathcal{I}(t_n - f_n)| \leq |\mathcal{I}(t_n)| + (b-a)\|t_n - f_n\| \\ &< |\mathcal{I}(t_n)| + \frac{1}{n}(b-a).\end{aligned}$$

□

Wir formulieren den Satz von Arzela-Osgood auch für Funktionenreihen:

Korollar VII-5.9. Es sei $(f_n) \subset \mathcal{R}(I, \mathbb{K})$. Die Reihe $\sum_{k=1}^{\infty} f_k$ konvergiere punktweise gegen eine Regelfunktion. Ferner sei die Folge der Partialsummen gleichmäßig beschränkt, d.h.

$$\exists c \geq 0 \,\forall x \in I \,\forall n \in \mathbb{N} \colon |\sum_{k=1}^{n} f_k(x)| \leq c.$$

Dann gilt

$$\sum_{k=0}^{\infty} \left(\int_a^b f_k(x)\, dx \right) = \int_a^b \left(\sum_{k=0}^{\infty} f_k(x) \right) dx.$$

6. Parameterabhängige Integrale

Wir gehen von einer Funktion $f \colon I \times D \to \mathbb{K}$, $D \subset \mathbb{K}$, mit folgender Eigenschaft aus:

$$\forall t \in D \colon x \mapsto f(x,t) \text{ ist eine Regelfunktion.}$$

(Die Abbildung $x \mapsto f(x,t)$ bezeichnen wir mit $f(\cdot,t)$). Man bezeichnet t in diesem Zusammenhang oft als **Parameter**. Definiert man

$$F : \begin{cases} D \to \mathbb{K} \\ t \mapsto \int\limits_a^b f(x,t)\,dx, \end{cases}$$

sagt man, F sei durch ein **parameterabhängiges Integral** gegeben. Wir untersuchen nun die Eigenschaften von F.

Theorem VII-6.1. *$f \in \mathcal{B}(I \times D, \mathbb{K})$, $D \subset \mathbb{K}$, habe folgende Eigenschaften:*

i) $\forall t \in D\colon f(\cdot,t) \in \mathcal{R}(I,\mathbb{K})$,
ii) $\forall x \in I\colon f(x,\cdot) \in C(D,\mathbb{K})$.

Dann ist die Abbildung $F\colon D \to \mathbb{K}$,

$$F(t) = \int\limits_a^b f(x,t)\,dx$$

stetig auf D.

Beweis. Es sei $t_0 \in D$ ein Häufungspunkt von D und $(t_n) \subset D$ konvergiere gegen t_0. Wir betrachten nun die Folge von Regelfunktionen (g_n),

$$g_n(x) = f(x,t_n), \quad x \in I, n \in \mathbb{N}.$$

Aus ii) folgt für alle $x \in I$

$$\lim_{n\to\infty} g_n(x) = \lim_{n\to\infty} f(x,t_n) = f(x,t_0).$$

Es gilt $f(\cdot,t_0) \in \mathcal{R}(I,\mathbb{K})$ und wegen $f \in \mathcal{B}(I \times D,\mathbb{K})$ ist die Folge $(\|g_n\|)$, d.h. $(\|f(\cdot,t_n)\|)$ beschränkt. Nach dem Satz von Arzela-Osgood VII-5.8 folgt daher

$$\begin{aligned}\lim_{n\to\infty} F(t_n) = \lim_{n\to\infty} \int\limits_a^b f(x,t_n)\,dx &= \lim_{n\to\infty} \int\limits_a^b g_n(x)\,dx = \int\limits_a^b \lim_{n\to\infty} g_n(x)\,dx \\ &= \int\limits_a^b f(x,t_0)\,dx = F(t_0).\end{aligned}$$

□

Der Beweis zeigt, daß die Forderung der globalen Beschränktheit von f abgeschwächt werden kann zu

$$\forall t \in D\,\exists U \in \mathcal{U}(t)\colon f \text{ ist beschränkt auf } I \times (U \cap D).$$

Beispiel VII-6.2. 1) Es sei $I = [a,b]$, $0 < a < b$, $D = \mathbb{R}$, und $f(x,t) = x^t$. Aus dem Satz folgt, daß

$$F(t) = \begin{cases} \frac{b^{t+1}-a^{t+1}}{t+1}, & t \neq -1, \\ \ln \frac{b}{a}, & t = -1, \end{cases}$$

auf beliebigen kompakten Intervallen von $\mathbb{R}$ und damit auf $\mathbb{R}$ selbst stetig ist. Insbesondere folgt

$$\lim_{t \to -1} \frac{b^{t+1} - a^{t+1}}{t+1} = F(-1) = \ln \frac{b}{a}.$$

2) Es sei $I = [a,b] \subset \mathbb{R}$, $D = (0,\infty)$ und $f(x,t) = t^x$. Mit Hilfe von Satz VII-6.1 schließen wir auf die Stetigkeit von

$$F(t) = \begin{cases} \frac{1}{\ln t}(t^b - t^a), & t \in (0,\infty) \setminus \{1\}, \\ b - a, & t = 1, \end{cases}$$

und insbesondere $F(1) = \int\limits_a^b 1\,dx = b - a = \lim_{t\to 1} \frac{1}{\ln t}(t^b - t^a)$.

Das folgende Resultat benötigt den Begriff der partiellen Ableitungen, der im Abschnitt über die Differenzierbarkeit von Funktionen in mehreren Veränderlichen diskutiert wird.

Theorem VII-6.3 (Vertauschung von Integration und Differentiation). *Es sei $f\colon I \times D \to \mathbb{K}$, und $D \subset \mathbb{K}$. f besitze folgende Eigenschaften:*

i) $\forall t \in D\colon f(\cdot,t) \in \mathcal{R}(I,\mathbb{K})$,
ii) $\forall x \in I\colon$ *existiere die partielle Ableitung* $\frac{\partial f}{\partial t}\colon I \times D \to \mathbb{K}$,
iii) $\forall t \in D\colon \frac{\partial f}{\partial t}(\cdot,t) \in \mathcal{R}(I,\mathbb{K})$,
iv) $\frac{\partial f}{\partial t} \in \mathcal{B}(I \times D,\mathbb{K})$, *d.h.* $\exists c \geq 0\,\forall (x,t) \in I \times D\colon \left|\frac{\partial f}{\partial t}(x,t)\right| \leq c$.

Dann ist die Abbildung $F\colon D \to \mathbb{K}$, definiert durch

$$F(t) = \int\limits_a^b f(x,t)\,dx,$$

auf D differenzierbar und es gilt

$$F'(t) = \int\limits_a^b \frac{\partial f}{\partial t}(x,t)\,dx.$$

Beweis. Es sei $t \in D$ und $(t_n) \subset D$ mit $\lim_{n\to\infty} t_n = t$ und $t_n \neq t$, $n \in \mathbb{N}$. Wir betrachten den Differenzenquotient von F:

$$\frac{F(t_n) - F(t)}{t_n - t} = \int\limits_a^b \frac{f(x,t_n) - f(x,t)}{t_n - t}\,dx =: \int\limits_a^b g_n(x)\,dx, \quad n \in \mathbb{N}.$$

Wegen i) gilt $(g_n) \subset \mathcal{R}(I, \mathbb{K})$ und aus ii) folgt für alle $x \in I$

$$\lim_{n\to\infty} g_n(x) = \frac{\partial f}{\partial t}(x,t).$$

Nach iii) ist $\frac{\partial f}{\partial t}$, also der punktweise Grenzwert der Folge (g_n), eine Regelfunktion. Wir zeigen nun die Beschränktheit von $(\|g_n\|)$: Der Mittelwertsatz V-4.6 zeigt

$$g_n(x) = \frac{f(x,t_n) - f(x,t)}{t_n - t} = \frac{\partial f}{\partial t}(x, t^*(x)), \quad t^*(x) \in (\min\{t, t_n\}, \max\{t, t_n\})$$

und iv) ergibt

$$\forall x \in I\, \forall n \in \mathbb{N}\colon |g_n(x)| \le c.$$

Aus dem Satz von Arzela-Osgood VII-5.8 folgt somit

$$\begin{aligned} F'(t) &= \lim_{n\to\infty} \frac{F(t_n) - F(t)}{t_n - t} = \lim_{n\to\infty} \int_a^b g_n(x)\, dx \\ &= \int_a^b \lim_{n\to\infty} g_n(x)\, dx = \int_a^b \frac{\partial f}{\partial t}(x,t)\, dx. \end{aligned}$$

$\square$

Man beachte, daß in Satz VII-6.3 nicht die Beschränktheit von f, sondern jene von $\frac{\partial f}{\partial t}$ gefordert wird.

Beispiel VII-6.4 (Gaußsches Fehlerintegral). Wir betrachten das parameterabhängige Integral

$$F(t) = \int_0^1 \frac{e^{-(1+x^2)t^2}}{1+x^2}\, dx$$

für $t \in \mathbb{R}$. Wir setzen also $I = [0,1]$ und

$$f := \begin{cases} [0,1] \times \mathbb{R} \to \mathbb{R}, \\ (x,t) \mapsto \frac{e^{-(1+x^2)t^2}}{1+x^2}. \end{cases}$$

Wir verifizieren die Voraussetzungen von Satz VII-6.3: Für jedes $t \in \mathbb{R}$ ist die Abbildung $x \mapsto f(x,t)$ stetig, also eine Regelfunktion und für jedes $x \in [0,1]$ ist $t \mapsto f(x,t)$ differenzierbar. Es gilt

$$\frac{\partial f}{\partial t}(x,t) = -2te^{-(1+x^2)t^2},$$

somit ist $\frac{\partial f}{\partial t}(\cdot, t) \in \mathcal{R}(I, \mathbb{R})$ für alle $t \in \mathbb{R}$. Ferner gilt für alle $(x,t) \in I \times \mathbb{R}$

$$\Big|\frac{\partial f}{\partial t}(x,t)\Big| = 2|t|e^{-(1+x^2)t^2} \le \begin{cases} 2|t| \le 2 & |t| \le 1, x \in I, \\ \frac{2|t|}{(1+x^2)t^2} \le \frac{2}{|t|} \le 2 & |t| \ge 1, x \in I, \end{cases}$$

also ist $\frac{\partial f}{\partial t} \in \mathcal{B}(I \times \mathbb{R}, \mathbb{R})$. Nach Satz VII-6.3 ist F auf $\mathbb{R}$ differenzierbar und es gilt

$$F'(t) = -2e^{-t^2} \int_0^1 e^{-x^2t^2} t \, dx.$$

Substituiert man für festes t im Integral $g(x) = xt$, also $g(0) = 0$ und $g(1) = t$, erhält man mit Hilfe von Satz VII-4.8

$$F'(t) = -2e^{-t^2} \int_0^t e^{-u^2} du.$$

Beschränken wir uns vorübergehend auf ein kompaktes t-Intervall, folgt wegen der Stetigkeit von $u \mapsto e^{-u^2}$ aus dem Hauptsatz VII-4.2 die Beziehung

$$\frac{d}{dt}(\int_0^t e^{-u^2} du)^2 = 2 \int_0^t e^{-u^2} du \cdot e^{-t^2}.$$

Wir erhalten also

$$F'(t) = -\frac{d}{dt}\big(\int_0^t e^{-u^2} du\big)^2,$$

d.h.

$$F(t) = -\big(\int_0^t e^{-u^2} du\big)^2 + c, \quad t \in \mathbb{R}.$$

Die Konstante c ist festgelegt durch

$$c = F(0) = \int_0^1 \frac{dx}{1+x^2} = \arctan x\Big|_0^1 = \frac{\pi}{4}.$$

Wir erhalten daraus für alle $t \in \mathbb{R}$

$$\left(\int_0^t e^{-u^2} du \right)^2 = \frac{\pi}{4} - F(t).$$

Wir betrachten nun eine beliebige Folge (t_n) mit $\lim_{n\to\infty} t_n = \infty$. Für alle $n \in \mathbb{N}$ und $x \in I$ gilt dann

$$\lim_{n\to\infty} f(x, t_n) = 0$$

und

$$|f(x, t_n)| \leq 1.$$

Aus dem Satz von Arzela-Osgood folgt daher

$$\lim_{n\to\infty} F(t_n) = \lim_{n\to\infty} \int_0^1 f(x,t_n)\,dx = \int_0^1 \lim_{n\to\infty} f(x,t_n)\,dx = 0,$$

und folglich

$$\lim_{t\to\infty} \int_0^t e^{-u^2} du = \frac{\sqrt{\pi}}{2}.$$

Vorausgreifend bemerken wir, daß man den Grenzwert auf der linken Seite mit dem Symbol $\int_0^\infty e^{-u^2} du$ bezeichnet. Somit wurde gezeigt

$$\int_0^\infty e^{-u^2} du = \frac{\sqrt{\pi}}{2}.$$

Schließlich betrachten wir noch die Integration einer durch ein parameterabhängiges Integral definierten Funktion F. Unter den Voraussetzungen von Satz VII-6.1 ist F stetig und somit auf kompakten Teilintervallen von D integrierbar. Es sei etwa $[\alpha,\beta] \subset D$, dann gilt

$$\int_\alpha^\beta F(t)\,dt = \int_\alpha^\beta \Big[\int_a^b f(x,t)\,dx\Big]\,dt.$$

Die Frage liegt nahe, ob die Gleichung

$$\int_\alpha^\beta \Big[\int_a^b f(x,t)\,dx\Big]\,dt = \int_a^b \Big[\int_\alpha^\beta f(x,t)\,dt\Big]\,dx$$

gilt, d.h. ob es zulässig ist, die Integrationsreihenfolge zu vertauschen. Da die Variablen x und t bei dieser Problemstellung gleichberechtigt auftreten, fordern wir, daß $f(\cdot,t) \in C(I,\mathbb{K})$ für alle $t \in D$ ist.

Theorem VII-6.5 (Vertauschung der Integrationsreihenfolge). *Es sei $f \in \mathcal{B}([a,b]\times[\alpha,\beta],\mathbb{K})$ und es gelte*

i) $\forall x \in [a,b]\colon f(x,\cdot) \in C([\alpha,\beta],\mathbb{K})$,
ii) $\forall t \in [\alpha,\beta]\colon f(\cdot,t) \in C([a,b],\mathbb{K})$.

Dann gilt

$$\int_\alpha^\beta \Big[\int_a^b f(x,t)\,dx\Big]\,dt = \int_a^b \Big[\int_\alpha^\beta f(x,t)\,dt\Big]\,dx.$$

Es ist klar, daß $f \in C([a,b]\times[\alpha,\beta],\mathbb{K})$ eine hinreichende Bedingung für i) und ii) darstellt.

Beweis. Nach Satz VII-6.1 ist die Abbildung $F\colon [\alpha,\beta] \to \mathbb{K}$

$$F(t) = \int_a^b f(x,t)\,dx$$

stetig auf $[\alpha,\beta]$ und somit integrierbar. Wir definieren die Abbildung $H\colon [\alpha,\beta] \to \mathbb{K}$ durch

$$H(y) = \int_\alpha^y F(t)\,dt, \quad y \in [\alpha,\beta].$$

Nach dem Hauptsatz VII-4.2 gilt für alle $y \in [\alpha,\beta]$

$$H'(y) = F(y) = \int_a^b f(x,y)\,dx.$$

Wir untersuchen nun die Abbildung $g\colon [a,b] \times [\alpha,\beta] \to \mathbb{K}$, welche gegeben ist durch

$$g(x,y) = \int_\alpha^y f(x,t)\,dt,$$

und verifizieren die Voraussetzungen von Satz VII-6.3. Für jedes $y \in [\alpha,\beta]$ ist $g(\cdot,y) \in C([a,b],\mathbb{K})$ nach Satz VII-6.1 (mit y als Parameter). Für jedes $x \in [a,b]$ ist $g(x,\cdot)$ nach dem Hauptsatz VII-4.2 wegen der Stetigkeit von $f(x,\cdot)$ stetig differenzierbar auf $[\alpha,\beta]$ und es gilt für alle $y \in [\alpha,\beta]$, $x \in [a,b]$

$$\frac{\partial g}{\partial y}(x,y) = f(x,y).$$

Es ist also auch $\frac{\partial g}{\partial y}(\cdot,y) \in C([a,b],\mathbb{K})$ und $\frac{\partial g}{\partial y} \in \mathcal{B}([a,b]\times[\alpha,\beta],\mathbb{K})$. Nach Satz VII-6.3 ist die Abbildung $G\colon [\alpha,\beta] \to \mathbb{K}$, definiert durch

$$G(y) = \int_a^b g(x,y)\,dx,$$

auf $[\alpha,\beta]$ differenzierbar mit

$$G'(y) = \int_a^b \frac{\partial g}{\partial y}(x,y)\,dx = \int_a^b f(x,y)\,dx.$$

Somit folgt für alle $y \in [\alpha,\beta]$

$$H'(y) = G'(y).$$

Es gibt daher eine Konstante $c \in \mathbb{K}$ mit

$$H(y) = G(y) + c, \quad y \in [\alpha,\beta].$$

Wegen $H(\alpha) = G(\alpha) = 0$ ($g(x,\alpha) = 0$ für alle $x \in [a,b]$) folgt sogar $H = G$, d.h. es ist

$$\int_\alpha^y \Big[\int_a^b f(x,t)\,dx\Big]\,dt = \int_a^b \Big[\int_\alpha^y f(x,t)\,dt\Big]\,dx, \quad y \in [\alpha,\beta].$$

Für $y = \beta$ ergibt sich die Behauptung. □

7. Uneigentliche Integrale

Die bisher entwickelte Integrationstheorie bezog sich auf Funktionen, deren Definitionsbereich ein *kompaktes* Intervall in $\mathbb{R}$ ist. Dies hat zwei starke Einschränkungen der Anwendbarkeit der Theorie zur Folge: Einerseits muß der Integrationsbereich beschränkt sein, andererseits sind integrierbare Funktionen notwendigerweise beschränkt. Wir zeigen nun, wie man Funktionen, die diesen Anforderungen nicht genügen, unter bestimmten Umständen einen sinnvollen Wert des Integrals zuordnen kann.

Definition VII-7.1. *Es sei $I = [a,b] \subset \bar{\mathbb{R}}$, $-\infty \le a < b \le \infty$ und $f\colon I \cap \mathbb{R} \to \mathbb{K}$. Wir definieren folgende* ***uneigentliche Integrale*** $\int_a^b f(x)\,dx$*:*

(1) *Es sei $a \in \mathbb{R}$ und $f|_{[a,\beta]} \in \mathcal{R}([a,\beta],\mathbb{K})$ für alle $\beta \in [a,b)$. Existiert $\lim_{\beta\uparrow b} \int_a^\beta f(x)\,dx$, so definiert man*

$$\int_a^b f(x)\,dx := \lim_{\beta\uparrow b} \int_a^\beta f(x)\,dx.$$

(2) *Eine analoge Definition gilt, falls $b \in \mathbb{R}$ für $\alpha \downarrow a$.*
(3) *Es sei $-\infty \le a < b \le \infty$. Existieren für ein $c \in (a,b)$ die uneigentlichen Integrale $\int_a^c f(x)\,dx$ und $\int_c^b f(x)\,dx$, definieren wir*

$$\int_a^b f(x)\,dx := \int_a^c f(x)\,dx + \int_c^b f(x)\,dx.$$

(4) *Das uneigentliche Integral $\int_a^b f(x)\,dx$ heißt* ***divergent****, wenn der entsprechende Grenzwert nicht existiert.*

Die Unabhängigkeit des uneigentlichen Integrals in (3) von der Wahl der Zwischenstelle wird später behandelt. Ist $I = [a,b] \subset \mathbb{R}$ und $f \in \mathcal{R}(I,\mathbb{K})$, so stimmt das uneigentliche Integral mit dem Integral gemäß Definition VII-2.1 überein. Wegen der stetigen Abhängigkeit etwa von der oberen Grenze gilt nämlich

$$\int_a^b f(x)\,dx = \lim_{\beta \uparrow b} \int_a^\beta f(x)\,dx.$$

(Beweis: Übung).
Diese Definition erfaßt die für die Praxis relevanten Typen von uneigentlichen Integralen:

Beispiel VII-7.2. 1) $\int_1^\infty x^{-s}\,dx$.
(Typ 1: $a \in \mathbb{R}$, $b = \infty$, unbeschränkter Integrationsbereich). Der Integrand ist stetig auf $[1, \beta]$ für $\beta \in [1, \infty)$. Man erhält

$$\int_1^\beta x^{-s}\,dx = \begin{cases} \frac{1}{s-1}(1-\beta^{1-s}), & s \neq 1, \\ \ln \beta, & s = 1. \end{cases}$$

Somit existiert $\lim_{\beta\to\infty} \int_1^\beta x^{-s}\,dx$ genau dann, wenn $s > 1$ gilt. Sein Wert ist

$$\int_1^\infty x^{-s}dx = \lim_{\beta\to\infty} \int_1^\beta x^{-s}ds = \frac{1}{s-1}.$$

2) $\int_0^1 x^{-s}dx$.
(Typ 2: a, $b \in \mathbb{R}$, Integrand unbeschränkt am Rand des Integrationsbereiches). Der Integrand ist stetig auf $[\alpha, 1]$ für alle $\alpha \in (0, 1]$. Man erhält

$$\int_\alpha^1 x^{-s}dx = \begin{cases} \frac{1}{1-s}(1-\alpha^{1-s}), & s \neq 1, \\ -\ln \alpha, & s = 1. \end{cases}$$

Das uneigentliche Integral $\int_0^1 x^{-s}dx$ existiert also genau für $s < 1$. Es hat den Wert

$$\int_0^1 x^{-s}dx = \lim_{\alpha\downarrow 0} \int_\alpha^1 x^{-s}dx = \frac{1}{1-s}.$$

3) $\int_{-\infty}^\infty \frac{dx}{1+x^2} = \pi$.

(Typ 3: $a = -\infty, b = \infty, c = 0$). Wie vorhin untersuchen wir die beiden uneigentlichen Integrale

$$\int\limits_0^\infty \frac{dx}{1+x^2} = \lim_{\beta\to\infty} \int\limits_0^\beta \frac{dx}{1+x^2} = \lim_{\beta\to\infty} \arctan\beta = \frac{\pi}{2}$$

$$\int\limits_{-\infty}^0 \frac{dx}{1+x^2} = \lim_{\alpha\to-\infty} \int\limits_\alpha^0 \frac{dx}{1+x^2} = -\lim_{\alpha\to-\infty} \arctan\alpha = \frac{\pi}{2}.$$

Beide zusammen ergeben die Behauptung. Aus Symmetriegründen könnte man sich in diesem Beispiel die Berechnung von $\int\limits_{-\infty}^0 \frac{dx}{1+x^2}$ ersparen.

4) $\int\limits_{-1}^1 \frac{dx}{\sqrt{|x|}} = 4.$
(Typ 4: a, $b \in \mathbb{R}$, $a < b$, $c = 0$, Integrand unbeschränkt im Inneren des Integrationsbereiches). Nach Beispiel 2) existieren die beiden uneigentlichen Integrale

$$\int\limits_0^1 \frac{dx}{\sqrt{x}} = \int\limits_{-1}^0 \frac{dx}{\sqrt{|x|}} = 2.$$

Daraus folgt die Behauptung.

Wir weisen darauf hin, daß bei uneigentlichen Integralen vom Typ 3) und 4) die Grenzwerte in den beiden uneigentlichen Integralen $\int\limits_a^c f(x)\,dx$ und $\int\limits_c^b f(x)\,dx$ voneinander *unabhängig* durchzuführen sind. Durch eine geeignete Koppelung der beiden Grenzwerte kann man manchmal die Existenz des Grenzwertes erzwingen: Als Beispiel betrachte man das uneigentliche Integral

$$\int\limits_a^b \frac{1}{x}\,dx, \quad a < 0 < b.$$

Dieses existiert nicht, da die beiden uneigentlichen Integrale $\int_a^0 \frac{1}{x}\,dx$ und $\int_0^b \frac{1}{x}\,dx$ nicht existieren. Koppelt man jedoch die beiden Grenzwerte $\lim_{\beta\uparrow 0} \int\limits_a^\beta \frac{1}{x}\,dx$ und $\lim_{\alpha\downarrow 0} \int\limits_\alpha^b \frac{1}{x}\,dx$ in der Form

$$\lim_{\varepsilon\downarrow 0}\left(\int\limits_a^{-\varepsilon} \frac{1}{x}\,dx + \int\limits_\varepsilon^b \frac{1}{x}\,dx\right) = \lim_{\varepsilon\downarrow 0}(\ln|-\varepsilon| - \ln|a| + \ln b - \ln\varepsilon) = \ln\frac{b}{|a|},$$

existiert der Grenzwert. Man nennt diesen *speziellen* Grenzwert **Cauchyscher Hauptwert** des divergenten uneigentlichen Integrals und schreibt

$$(C)\int_a^b \frac{1}{x}\,dx = \ln\frac{b}{|a|}.$$

Existiert das uneigentliche Integral $\int\limits_a^\infty f(x)\,dx$ für ein $a \in \mathbb{R}$, dann existiert auch das uneigentliche Integral $\int_\alpha^\infty f(x)\,dx$ für alle $\alpha \geq a$ und sein Wert ist

$$\int_\alpha^\infty f(x)\,dx = \int_a^\infty f(x)\,dx - \int_a^\alpha f(x)\,dx.$$

So selbstverständlich dies erscheinen mag, bedarf es trotzdem einer Rechtfertigung. Die Existenz von $\int_a^\infty f(x)\,dx$ bedeutet

$$(*) \qquad \forall\varepsilon > 0\,\exists\xi \geq a\,\forall\beta \geq \xi\colon \Big|\int_a^\infty f(x)\,dx - \int_a^\beta f(x)\,dx\Big| < \varepsilon.$$

Es sei nun $\alpha \geq a$ und $\varepsilon > 0$ beliebig und ξ entsprechend $(*)$ gewählt (O.B.d.A. können wir $\xi \geq \alpha$ annehmen). Für $\beta \geq \alpha$ gilt dann

$$\Big|\Big(\int_a^\infty f(x)\,dx - \int_a^\alpha f(x)\,dx\Big) - \int_\alpha^\beta f(x)\,dx\Big| = \Big|\int_a^\infty f(x)\,dx - \int_a^\beta f(x)\,dx\Big| < \varepsilon,$$

d.h.

$$\lim_{\beta\to\infty}\int_\alpha^\beta f(x)\,dx = \int_a^\infty f(x)\,dx - \int_a^\alpha f(x)\,dx.$$

Auch uneigentliche Integrale sind also additiv in Bezug auf das Integrationsintervall und verhalten sich in dieser Hinsicht wie das Cauchy Integral. Insbesondere gilt auch für alle $\alpha \leq \beta$

$$\int_\alpha^\beta f(x)\,dx = \int_\alpha^\infty f(x)\,dx - \int_\beta^\infty f(x)\,dx$$

soferne die uneigentlichen Integrale existieren. Tatsächlich gilt diese Formel für alle $\alpha, \beta \geq a$.

Ähnlich der Situation bei Reihen ist es meist nicht möglich, uneigentliche Integrale zu berechnen, sodaß man sich mit der bloßen Existenz (man sagt auch Konvergenz) der uneigentlichen Integrale bescheiden muß. Hiefür gibt es verschiedene Kriterien, welche wir nur für den Typ $\int_a^\infty f(x)\,dx$ formulieren. Die Modifikation für die anderen Möglichkeiten sind evident. Universell einsetzbar ist das Cauchy-Kriterium:

Theorem VII-7.3 (Cauchy-Kriterium). *Es sei* $f\colon [a,\infty) \to \mathbb{K}$ *und* $f|_{[a,\beta]} \in \mathcal{R}([a,\beta],\mathbb{K})$ *für alle* $\beta \geq a$. *Das uneigentliche Integral* $\int_a^\infty f(x)\,dx$ *existiert genau dann, wenn gilt:*

$$\forall \varepsilon > 0\, \exists \xi \geq a\, \forall \alpha, \beta \geq \xi\colon \Big|\int\limits_\alpha^\beta f(x)\,dx\Big| < \varepsilon.$$

Beweis. Man betrachte $\lim_{\beta\to\infty} F(\beta)$ für $F(\beta) = \int\limits_a^\beta f(x)\,dx$, $\beta \geq a$. □

Beispiel VII-7.4. $\int_0^\infty \frac{\sin x}{x}\,dx$.
Wir definieren die Funktion $f\colon [0,\infty) \to \mathbb{R}$ durch

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

und bemerken, daß f auf $[0,\infty)$ stetig ist. Insbesondere ist f integrierbar auf $[0,a]$ und auf $[a,\beta]$ für beliebige $0 < a < \beta < \infty$. Spalten wir daher $\int_0^\beta f(x)\,dx$ auf in

$$\int\limits_0^\beta f(x)\,dx = \int\limits_0^a f(x)\,dx + \int\limits_a^\beta f(x)\,dx,$$

genügt es, die Existenz von $\lim_{\beta\to\infty} \int\limits_a^\beta f(x)\,dx$ nachzuweisen. Dazu formen wir vorerst das Integral mit partieller Integration um:

$$\int\limits_a^\beta \frac{1}{x}\sin x\,dx = -\frac{1}{x}\cos x\Big|_a^\beta - \int\limits_a^\beta \frac{1}{x^2}\cos x\,dx.$$

Der erste Term auf der rechten Seite besitzt einen Grenzwert für $\beta \to \infty$. Auf das Integral wenden wir das Cauchy-Kriterium an und erhalten für $a \leq u < v$

$$\Big|\int\limits_u^v \frac{\cos x}{x^2}\,dx\Big| \leq \int\limits_u^v \frac{dx}{x^2} = \frac{1}{u} - \frac{1}{v} < \frac{1}{u}.$$

Wählt man $\varepsilon > 0$ und setzt $\xi(\varepsilon) = \frac{1}{\varepsilon}$, erhält man für $\xi(\varepsilon) < u < v$

$$\Big|\int\limits_u^v \frac{\cos x}{x^2}\,dx\Big| < \varepsilon,$$

und damit die Konvergenz von $\int_a^\infty \frac{\cos x}{x^2}\,dx$. Somit existiert auch $\int_0^\infty \frac{\sin x}{x}\,dx$.

Definition VII-7.5. *Es sei* $f\colon [a,\infty) \to \mathbb{K}$ *und* $f|_{[a,\beta]} \in \mathcal{R}([a,\beta],\mathbb{K})$ *für alle* $\beta \in [a,\infty)$. *Das uneigentliche Integral* $\int\limits_a^\infty f(x)\,dx$ *heißt* ***absolut konvergent****, wenn* $\int_a^\infty |f(x)|\,dx$ *existiert.*

Aus der absoluten Konvergenz eines uneigentlichen Integrals folgt dessen Konvergenz (Existenz). Dies ergibt sich mit Hilfe des Cauchy-Kriteriums aus der Ungleichung $(u < v)$

$$\Big|\int\limits_u^v f(x)\,dx\Big| \leq \int\limits_u^v |f(x)|\,dx.$$

Die Umkehrung ist – wie bei Reihen – nicht richtig: Aus der bloßen Existenz eines uneigentlichen Integrals folgt nicht dessen absolute Konvergenz. Als Beispiel betrachten wir das konvergente Integral $\int_0^\infty \frac{\sin x}{x}\,dx$. Die Abbildung $F\colon [0,\infty) \to \mathbb{R}$ sei gegeben durch

$$F(\beta) = \int\limits_0^\beta \Big|\frac{\sin x}{x}\Big|\,dx.$$

Für $\beta = n\pi$ erhalten wir

$$F(n\pi) = \int\limits_0^{n\pi} \Big|\frac{\sin x}{x}\Big|\,dx = \sum_{i=1}^n \int\limits_{(i-1)\pi}^{i\pi} |\sin x|\frac{1}{x}\,dx \geq \sum_{i=1}^n \frac{1}{i\pi} \int\limits_{(i-1)\pi}^{i\pi} |\sin x|\,dx = \frac{2}{\pi}\sum_{i=1}^n \frac{1}{i}.$$

Wegen der Divergenz der harmonischen Reihe kann also $\lim_{\beta\to\infty} F(\beta)$ nicht existieren.

Theorem VII-7.6 (Vergleichskriterium). *Es sei* $f\colon [a,\infty) \to \mathbb{R}$, $g\colon ([a,\infty) \to \mathbb{R}_+)$ *und* $f|_{[a,\beta]} \in \mathcal{R}([a,\beta],\mathbb{R})$, $g|_{[a,\beta]} \in \mathcal{R}([a,\beta],\mathbb{R}_+)$ *für alle* $\beta \geq a$. *Gilt* $|f(t)| \leq g(t)$ *für alle* $t \in [a,\infty)$ *und existiert* $\int_a^\infty g(t)\,dt$, *dann ist* $\int_a^\infty f(t)\,dt$ *absolut konvergent.*
Umgekehrt: Gilt $f(t) \geq g(t) \geq 0$ *für alle* t *(notwendigerweise ist* $\operatorname{bild} f \subset \mathbb{R}$*) und ist* $\int_a^\infty g(t)\,dt$ *divergent, dann divergiert auch das uneigentliche Integral* $\int_a^\infty f(t)\,dt$.

Beweis. Die erste Behauptung ergibt sich aus der Abschätzung

$$\Big|\int\limits_u^v f(x)\,dx\Big| \leq \int\limits_u^v |f(x)|\,dx \leq \int\limits_u^v g(x)\,dx$$

für $a \leq u < v$ und dem Cauchy-Kriterium VII-7.3. Die zweite Behauptung folgt aus

$$\int\limits_a^\beta g(x)\,dx \leq \int\limits_a^\beta f(x)\,dx$$

für $\beta \in [a,\infty)$ und der Unbeschränktheit von $\{\int\limits_a^\beta g(x)\,dx\colon \beta \in [a,\infty)\}$. $\square$

KOROLLAR VII-7.7. Es sei $f\colon [a,\infty) \to \mathbb{R}$, $a > 0$, und $f|_{[a,\beta]} \in \mathcal{R}([a,\beta],\mathbb{R})$ für alle $\beta \in [a,\infty)$. Gilt

$$|f(x)| \leq \frac{1}{x^s}, \quad \text{für ein } s > 1$$

für alle $x \geq x_0 \geq a$, dann ist $\int_a^\infty f(x)\,dx$ absolut konvergent. Ist $\operatorname{bild} f \subset \mathbb{R}$ und gilt

$$0 < \frac{1}{x^s} \leq f(x), \quad \text{für ein } s \leq 1$$

für alle $x \geq x_0 \geq a$, dann ist $\int_a^\infty f(x)\,dx$ divergent.

Die manchmal recht mühsamen Abschätzungen bei der Anwendung des Vergleichskriteriums kann man u.U. auch umgehen:

THEOREM VII-7.8 (Grenzwertkriterium). *Es seien $f, g\colon [a,\infty), \mathbb{R}$ und $f|_{[a,\beta]}, g|_{[a,\beta]} \in \mathcal{R}([a,\beta]),\mathbb{R})$ für alle $\beta \geq a$. Darüber hinaus gelte $f(x) \geq 0$ und $g(x) > 0$ für alle $x \in [a,\infty)$. Es existiere*

$$\rho := \lim_{x\to\infty} \frac{f(x)}{g(x)} \in \bar{\mathbb{R}}.$$

i) *Gilt $\rho \in (0,\infty)$, dann sind $\int_a^\infty f(x)\,dx$ und $\int_a^\infty g(x)\,dx$ beide konvergent oder beide divergent.*
ii) *Gilt $\rho = 0$ und existiert $\int_a^\infty g(x)\,dx$, dann existiert auch $\int_a^\infty f(x)\,dx$.*
ii) *Gilt $\rho = \infty$ und divergiert $\int_a^\infty g(x)\,dx$, dann divergiert auch $\int_a^\infty f(x)\,dx$.*

BEWEIS. i) Es gibt ein $\xi \geq a$ so, daß für alle $x \geq \xi$ die Abschätzung

$$\frac{\rho}{2} g(x) \leq f(x) \leq \frac{3\rho}{2} g(x)$$

gilt. Die Behauptung folgt nun aus dem Vergleichskriterium.
ii) Es gibt ein $\xi \geq a$ so, daß für alle $x \geq \xi$ die Abschätzung

$$0 \leq f(x) \leq g(x)$$

zutrifft.
iii) In diesem Falle gilt für alle hinreichend großen x die Abschätzung

$$1 \leq g(x) \leq f(x).$$

□

Ersetzt man im Grenzwertkriterium f durch $|f|$, erhält man natürlich ein Kriterium für die absolute Konvergenz des uneigentlichen Integrals. Wir betonen noch einmal, daß analoge Kriterien auch für die übrigen Typen von uneigentlichen Integralen gelten.

BEISPIEL VII-7.9 (Eulersches Γ-Integral).

$$\Gamma(s) = \int_0^\infty x^{s-1} e^{-x}\,dx, \quad s > 0.$$

Da beide Grenzen kritisch sind, spalten wir das Integral auf in

$$\Gamma(s) = \int_0^1 x^{s-1} e^{-x}\, dx + \int_1^\infty x^{s-1} e^{-x}\, dx.$$

Für das erste Integral verwenden wir die Vergleichsfunktion $g(x) = x^{s-1}$. Nach Beispiel VII-7.2 existiert $\int_0^1 x^{s-1}\, dx$ genau für $s > 0$. Mit $f(x) = x^{s-1} e^{-x}$ ergibt sich

$$\rho = \lim_{x \downarrow 0} \frac{f(x)}{g(x)} = \lim_{x \downarrow 0} e^{-x} = 1.$$

Nach dem Grenzwertkriterium existiert $\int_0^1 x^{s-1} e^{-x}\, dx$ genau für $s > 0$. Für das zweite Integral verwenden wir die Vergleichsfunktion $g(x) = e^{-\frac{x}{2}}$ (um das mögliche Anwachsen von $x \mapsto x^{s-1}$ zu kompensieren). Wir erhalten nun

$$\rho = \lim_{x \to \infty} \frac{f(x)}{g(x)} = \lim_{x \to \infty} \frac{x^{s-1} e^{-x}}{e^{-x/2}} = \lim_{x \to \infty} x^{s-1} e^{-\frac{x}{2}} = 0.$$

Aus der Konvergenz von $\int_1^\infty e^{-\frac{x}{2}}\, dx$ folgt daher mit Satz VII-7.8 die Existenz von $\int_1^\infty x^{s-1} e^{-x}\, dx$ (sogar für alle $s \in \mathbb{R}$). Somit existiert $\Gamma(s)$ genau für $s > 0$. Als Übung beweise man

$$\forall n \in \mathbb{N} \colon \Gamma(n) = \int_0^\infty x^{n-1} e^{-x}\, dx = (n-1)!.$$

Mit Hilfe uneigentlicher Integrale läßt sich in manchen Fällen bequem die Konvergenz unendlicher Reihen nachweisen:

Theorem VII-7.10 (Integralkriterium). *Es sei $f \colon [1, \infty) \to \mathbb{R}$ monoton fallend und $f(x) \geq 0$ für alle $x \in [1, \infty)$. Die unendliche Reihe*

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

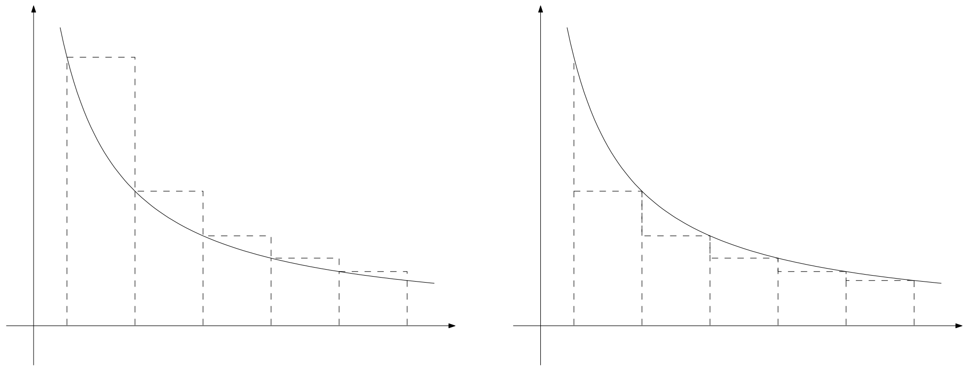
konvergiert genau dann, wenn das uneigentliche Integral

$$\int_1^\infty f(x)\, dx$$

existiert.

Beweis. Da f monoton fällt, ist $f|_{[1,\beta]} \in \mathcal{R}([1, \beta], \mathbb{R})$ für alle $\beta \geq 1$. Ferner gilt für alle $k \in \mathbb{N}$ die Abschätzung

$$f(k+1) \leq \int_k^{k+1} f(x)\, dx \leq f(k).$$



Durch Addition folgt

$$\sum_{k=2}^{n} f(k) \leq \int_{1}^{n} f(x)\,dx \leq \sum_{k=1}^{n-1} f(k), \quad n \in \mathbb{N}.$$

Bezeichnet S_n die n-te Partialsumme der Reihe $\sum_{k=1}^{\infty} f(k)$ erhält man

$$\begin{aligned} &S_n - f(1) \leq \int_{1}^{n} f(x)\,dx \leq S_{n-1}, \\ &S_n \leq S_{n+1}, \quad n \in \mathbb{N}. \end{aligned}$$

Existiert $\int\limits_{1}^{\infty} f(x)\,dx$, ergibt sich aus

$$S_n \leq f(1) + \int_{1}^{n} f(x)\,dx \leq f(1) + \int_{1}^{\infty} f(x)\,dx$$

die Konvergenz der Folge (S_n) aus dem Monotoniekriterium. Divergiert hingegen $\int_1^{\infty} f(x)\,dx$, ist wegen

$$\int_{1}^{n} f(x)\,dx \leq S_{n-1}$$

auch die Reihe $\sum_{k=1}^{\infty} f(k)$ divergent. □

Beispiel VII-7.11. $\sum_{n=2}^{\infty} \frac{1}{n(\ln n)^s}$ konvergiert für $s > 1$ und divergiert für $s \leq 1$.
Die Abbildung $x \to f(x) = \frac{1}{x}(\ln x)^{-s}$ erfüllt für $x \geq 2$ die Voraussetzungen des Integralkriteriums. Substitutiert man $g(x) = \ln x$, erhält man

$$\int_{2}^{n} \frac{dx}{x(\ln x)^s} = \int_{\ln 2}^{\ln n} \frac{dt}{t^s} = \frac{1}{1-s}[(\ln n)^{1-s} - (\ln 2)^{1-s}], \quad s \neq 1.$$

Für $s > 1$ folgt die Existenz des uneigentlichen Integrals

$$\int\limits_2^\infty \frac{dx}{x(lnx)^s},$$

für $s < 1$ dessen Divergenz. Für $s = 1$ ergibt sich wegen

$$\int\limits_3^n \frac{dx}{x \ln x} = \int\limits_{\ln 3}^{\ln n} \frac{dt}{t} = \ln(\ln n) - \ln(\ln 3)$$

ebenfalls die Divergenz des uneigentlichen Integrals. Dies zeigt die Behauptung.

Da uneigentliche Integrale durch einen Grenzübergang definiert sind, bleiben Linearität, Positivität und Monotonie erhalten. Der Betrag des Integrals kann aber nicht mehr gegen die Intervallänge und $\|f\|$ abgeschätzt werden. Dies wurde jedoch wesentlich im Beweis des Satzes von Arzela-Osgood verwendet. Wir zeigen nun, daß dieser Satz für uneigentliche Integrale nicht mehr gilt:

Beispiel VII-7.12. Es sei $(f_n) \subset C(\mathbb{R})$ gegeben durch

$$f_n(x) = \frac{n}{n^2 + x^2}, \quad x \in \mathbb{R}, n \in \mathbb{N}.$$

Es gilt $\lim_{n\to\infty} f_n = 0$ gleichmäßig auf $\mathbb{R}$. Für jedes f_n existiert das uneigentliche Integral

$$\int\limits_0^\infty \frac{n}{n^2 + x^2}\, dx = \int\limits_0^\infty \frac{dt}{1 + t^2} = \frac{\pi}{2}.$$

Es gilt also *nicht* $\lim_{n\to\infty} \int\limits_0^\infty f_n(x)\, dx = \int\limits_0^\infty \lim_{n\to\infty} f_n(x)\, dx$.

8. Parameterabhängige uneigentliche Integrale

Wie bei Funktionenreihen spielt bei parameterabhängigen uneigentlichen Integralen die Gleichmäßigkeit der Konvergenz eine wesentliche Rolle:

Definition VII-8.1. *Es sei $f\colon I \times D \to \mathbb{K}$, $I = [a, \infty)$, $D \subset \mathbb{R}$. Für alle $\beta \in [a, \infty)$ und für alle $t \in D$ sei $f|_{[a,\beta]}(\cdot, t) \in \mathcal{R}([a, \beta], \mathbb{K})$.*

$\int_a^\infty f(x, t)\, dx$ heißt ***gleichmäßig konvergent*** *(auf D) $\underset{Def}{\Leftrightarrow}$*

$$\forall \varepsilon > 0\, \exists \xi \in [a, \infty)\, \forall \alpha, \beta \geq \xi\, \forall t \in D\colon |\int\limits_\alpha^\beta f(x, t)\, dx| < \varepsilon.$$

Eine offensichtlich hinreichende Bedingung für gleichmäßige Konvergenz des uneigentlichen Integrals ist die Existenz einer bezüglich t gleichmäßigen Majorante $g\colon [a, \infty) \to \mathbb{R}$. Gilt also für alle $x \in [a, \infty)$ und für alle $t \in D$

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

und existiert $\int_a^\infty g(x)\,dx$, dann ist $\int_a^\infty f(x,t)\,dx$ gleichmäßig konvergent. Dies folgt unmittelbar aus der Abschätzung ($\alpha \le \beta$)

$$|\int_\alpha^\beta f(x,t)\,dx| \le \int_\alpha^\beta g(x)\,dx.$$

Beispiel VII-8.2. $\int_a^\infty \frac{1}{x}\exp((-t+i)x)\,dx$, $a > 0$, $t \ge 0$. Für $t \ge c > 0$ ergibt sich aus $|\frac{1}{x}\exp((-t+i)x)| \le \frac{1}{c}\frac{1}{x^2}$ die gleichmäßige Konvergenz auf $[c,\infty)$. Wir dehnen nun die Gleichmäßigkeit der Konvergenz auf $t \ge 0$ aus. Durch partielle Integration erhält man für $a \le u \le v$

$$\int_u^v \frac{1}{x}e^{(-t+i)x}\,dx = \frac{1}{x}\cdot\frac{1}{-t+i}e^{(-t+i)x}\Big|_u^v + \int_u^v \frac{1}{x^2}\frac{1}{-t+i}e^{(-t+i)x}\,dx.$$

Wegen $|\frac{1}{-t+i}| \le 1$ und $|e^{(-t+i)x}| = e^{-tx} \le 1$ für $t \ge 0$, ergibt sich die Abschätzung

$$\Big|\int_u^v \frac{1}{x}e^{(-t+i)x}\,dx\Big| \le \frac{1}{u} + \frac{1}{v} + \int_u^v \frac{1}{x^2}\,dx = \frac{2}{u}$$

und damit die gleichmäßige Konvergenz für $t \ge 0$. Durch Zerlegung in Real- und Imaginärteil folgt dann die gleichmäßige Konvergenz der reellen Integrale

$$\int_a^\infty \frac{\cos x}{x}e^{-tx}\,dx \quad \text{und} \quad \int_a^\infty \frac{\sin x}{x}e^{-tx}\,dx, \quad a > 0,$$

auf $t \ge 0$. Da $x \mapsto \frac{\sin x}{x}$ stetig nach 0 fortgesetzt werden kann, ist sogar

$$\int_0^\infty \frac{\sin x}{x}e^{-tx}\,dx$$

gleichmäßig konvergent auf $[0,\infty)$.

Theorem VII-8.3. *Für $f \in \mathcal{B}(I \times D, \mathbb{K})$, $I = [a,\infty)$, $D = [c,d] \subset \mathbb{R}$, gelte*

i) *$\int_a^\infty f(x,t)\,dx$ sei gleichmäßig konvergent auf D,*
ii) *$\forall x \in I\colon f(x,\cdot) \in C(D,\mathbb{K})$.*

Dann ist die Abbildung $F\colon D \to \mathbb{K}$, definiert durch

$$F(t) = \int_a^\infty f(x,t)\,dx,$$

stetig auf D.

Beweis. Es sei $t_0 \in D$. Für $\alpha \in [a, \infty)$ betrachten wir

$$\begin{aligned} |F(t) - F(t_0)| &= \Big|\int_a^\infty f(x,t)\,dx - \int_a^\infty f(x,t_0)\,dx\Big| \\ &\le \Big|\int_a^\alpha f(x,t)\,dx - \int_a^\alpha f(x,t_0)\,dx\Big| + \Big|\int_\alpha^\infty f(x,t)\,dx\Big| + \Big|\int_\alpha^\infty f(x,t_0)\,dx\Big|. \end{aligned}$$

Wegen der gleichmäßigen Konvergenz von $\int_a^\infty f(x,t)\,dx$ für $t \in D$ gibt es zu $\varepsilon > 0$ ein $\xi \ge a$ derart, daß für alle $\xi \le \alpha < \infty$ und für alle $t \in D$

$$\Big|\int_\alpha^\infty f(x,t)\,dx\Big| < \varepsilon.$$

Wir fixieren nun ein derartiges α. Auf $[a, \alpha]$ erfüllt f die Voraussetzungen von Satz VII-6.1. Also gibt es ein $\delta > 0$ so, daß $|t - t_0| < \delta$ und $t \in D$

$$\Big|\int_a^\alpha f(x,t)\,dx - \int_a^\alpha f(x,t_0)\,dx\Big| < \varepsilon$$

nach sich zieht. Insgesamt erhält man daher für $|t - t_0| < \delta$ und $t \in D$

$$|F(t) - F(t_0)| < 3\varepsilon.$$

□

Theorem VII-8.4. *Für $f \in \mathcal{B}(I \times D, \mathbb{K})$, $I = [a, \infty)$, $D = [\alpha, \beta] \subset \mathbb{R}$, gelte*

i) *$\int_a^\infty f(x,t)\,dx$ sei gleichmäßig konvergent auf D,*
ii) *$\forall t \in D\colon f(\cdot, t) \in C(I, \mathbb{K})$,*
iii) *$\forall x \in I\colon f(x, \cdot) \in C(D, \mathbb{K})$.*

Dann konvergiert das uneigentliche Integral $\int_a^\infty [\int_\alpha^\beta f(x,t)\,dt]\,dx$ und es gilt

$$\int_\alpha^\beta \Big[\int_a^\infty f(x,t)\,dx\Big]\,dt = \int_a^\infty \Big[\int_\alpha^\beta f(x,t)\,dt\Big]\,dx.$$

Beweis. Nach Satz VII-8.3 ist $t \mapsto \int_a^\infty f(x,t)\,dx$ stetig auf D. Somit existiert das iterierte Integral $\int_\alpha^\beta [\int_a^\infty f(x,t)\,dx]\,dt$. Es sei $\varepsilon > 0$ beliebig gewählt. Wegen der gleichmäßigen Konvergenz von $\int_a^\infty f(x,t)\,dx$ gibt es ein $\xi \ge a$, so daß für alle $u \ge \xi$

und $t \in D$ die Abschätzung

$$\Big|\int_a^\infty f(x,t)\,dx - \int_a^u f(x,t)\,dx\Big| < \varepsilon$$

zutrifft. Damit ergibt sich die Ungleichung

$$\begin{aligned}&\Big|\int_\alpha^\beta\Big[\int_a^\infty f(x,t)\,dx\Big]\,dt - \int_\alpha^\beta\Big[\int_a^u f(x,t)\,dx\Big]\,dt\Big| \\ &\le \int_\alpha^\beta \Big|\int_a^\infty f(x,t)\,dx - \int_a^u f(x,t)\,dx\Big|\,dt \le \varepsilon(\beta-\alpha).\end{aligned}$$

Auf Grund der Voraussetzungen an f kann man in $\int_\alpha^\beta[\int_a^u f(x,t)\,dx]\,dt$ nach Satz VII-6.5 die Integrationsreihenfolge vertauschen und erhält für alle $u \ge \xi$

$$\Big|\int_\alpha^\beta\Big[\int_a^\infty f(x,t)\,dx\Big]\,dt - \int_a^u\Big[\int_\alpha^\beta f(x,t)\,dt\Big]\,dx\Big| \le \varepsilon(\beta-\alpha).$$

Daraus folgt die Behauptung. □

Theorem VII-8.5. *Für $f\colon I\times D \to \mathbb{K}$, $I=[a,\infty)$, $D=[\alpha,\beta]\subset\mathbb{R}$, gelte:*

i) *Es existiere* $\int_a^\infty f(x,\alpha)\,dx$,
ii) *Es existiere* $\frac{\partial f}{\partial t}\colon I\times D\to\mathbb{K}$,
iii) $\frac{\partial f}{\partial t}\in\mathcal{B}(I\times D,\mathbb{K})$,
iv) $\forall t\in D\colon \frac{\partial f}{\partial t}(\cdot,t)\in C(I,\mathbb{K})$,
v) $\forall x\in I\colon \frac{\partial f}{\partial t}(x,\cdot)\in C(D,\mathbb{K})$,
vi) $\int_a^\infty \frac{\partial f}{\partial t}(x,t)\,dx$ *sei gleichmäßig konvergent auf* D.

Dann konvergiert

$$F(t) := \int_a^\infty f(x,t)\,dx$$

für alle $t\in D$, F ist differenzierbar auf D und es gilt

$$F'(t) = \int_a^\infty \frac{\partial f}{\partial t}(x,t)\,dx.$$

Beweis. $\frac{\partial f}{\partial t}$ erfüllt die Voraussetzungen von Satz VII-8.4 auf $I\times[\alpha,y]$ für alle $y\in D$. Somit konvergiert für alle $y\in D$ das uneigentliche Integral $\int_a^\infty[\int_\alpha^y \frac{\partial f}{\partial t}(x,t)\,dt]\,dx$

und es gilt

$$(*) \qquad \int_\alpha^y \Big[\int_a^\infty \frac{\partial f}{\partial t}(x,t)\,dx\Big]\,dt = \int_a^\infty \Big[\int_\alpha^y \frac{\partial f}{\partial t}(x,t)\,dt\Big]\,dx.$$

Für festes $x \in I$ ist $\frac{\partial f}{\partial t}(x,\cdot) \in C(D,\mathbb{K})$. Nach dem Hauptsatz VII-4.2 folgt daher für alle $y \in D$

$$\int_\alpha^y \frac{\partial f}{\partial t}(x,t)\,dt = f(x,y) - f(x,\alpha).$$

Wegen der Konvergenz von $\int_a^\infty f(x,\alpha)\,dx$ ergibt sich damit aus $(*)$ die Konvergenz von $\int_a^\infty f(x,y)\,dx$ und

$$F(y) := \int_a^\infty f(x,y)\,dx = \int_a^\infty f(x,\alpha)\,dx + \int_\alpha^y \Big[\int_a^\infty \frac{\partial f}{\partial t}(x,t)\,dx\Big]\,dt, \quad y \in D.$$

$\frac{\partial f}{\partial t}$ erfüllt auch die Voraussetzungen von Satz VII-8.3. Also ist $t \mapsto \int_a^\infty \frac{\partial f}{\partial t}(x,t)\,dx$ stetig auf D. Der Hauptsatz VII-4.2 sichert daher die Differenzierbarkeit von F auf D und

$$F'(t) = \int_a^\infty \frac{\partial f}{\partial t}(x,t)\,dx.$$

□

Beispiel VII-8.6. 1) Wir betrachten das komplexe uneigentliche Integral

$$\int_0^\infty e^{(-1+it)x}\,dx = -\frac{1}{-1+it}, \quad t \in \mathbb{R}.$$

Dies folgt aus

$$\int_0^\beta e^{(-1+it)x}\,dx = \frac{e^{-\beta}e^{it\beta}}{-1+it} - \frac{1}{-1+it}.$$

Durch Aufspalten in Real- und Imaginärteil ergibt sich

$$\int_0^\infty e^{(-1+it)x}\,dx = \int_0^\infty e^{-x}\cos tx\;dx + i\int_0^\infty e^{-x}\sin tx\;dx = \frac{1+it}{1+t^2},$$

und damit die reellen uneigentlichen Integrale

$$(*)\qquad \left.\begin{aligned} \int_0^\infty e^{-x} \cos tx \; dx &= \frac{1}{1+t^2} \\ \int_0^\infty e^{-x} \sin tx \; dx &= \frac{t}{1+t^2} \end{aligned}\right\}, \quad t \in \mathbb{R}.$$

Aus $|e^{(-1+it)x}| \leq e^{-x}$, $(t,x) \in [0,\infty) \times \mathbb{R}$, folgt die gleichmäßige Konvergenz der uneigentlichen Integrale. Wir schränken nun t auf $D = [0,1]$ ein. Die Abbildung $f\colon [0,\infty) \times D \to \mathbb{R}$, $f(x,t) = e^{-x} \cos tx$, erfüllt die Voraussetzungen von Satz VII-8.4. Somit existiert $\int_0^\infty [\int_0^1 e^{-x} \cos tx \, dt] \, dx$ und es gilt

$$\int_0^1 \Big[\int_0^\infty e^{-x} \cos tx \; dx\Big] \, dt = \int_0^\infty \Big[\int_0^1 e^{-x} \cos tx \; dt\Big] \, dx.$$

Auf der linken Seite erhält man mit $(*)$

$$\int_0^1 \frac{dt}{1+t^2} = \arctan 1 - \arctan 0 = \frac{\pi}{4},$$

auf der rechten Seite kann man das innere Integral ausführen:

$$\int_0^\infty \Big[\int_0^1 e^{-x} \cos tx \, dt\Big] \, dx = \int_0^\infty e^{-x} \frac{\sin x}{x} \, dx.$$

(Man beachte, daß der Integrand stetig nach $x = 0$ fortgesetzt werden kann). Insgesamt wurde also gezeigt

$$\int_0^\infty e^{-x} \frac{\sin x}{x} \, dx = \frac{\pi}{4}.$$

2) Für $t \geq 0$ definieren wir die Abbildung F durch

$$F(t) = \int_0^\infty e^{-tx} \frac{\sin x}{x} \, dx.$$

Die Existenz von $F(0)$ wurde in Beispiel VII-7.4 nachgewiesen, für $t > 0$ kann man wie in Beispiel 1) argumentieren. Dort wurde auch

$$F(1) = \frac{\pi}{4}$$

gezeigt. Wir setzen $f\colon [0,\infty) \times [0,1] \to \mathbb{R}$ fest durch

$$f(x,t) = e^{-tx} \frac{\sin x}{x},$$

und erhalten

$$\frac{\partial f}{\partial t}(x,t) = -e^{-tx}\sin x.$$

Für jedes $t > 0$ ergibt eine einfache Rechnung wie in 1)($*$) das uneigentliche Integral

$$\int\limits_0^\infty \frac{\partial f}{\partial t}(x,t)\,dx = -\int\limits_0^\infty e^{-tx}\sin x\,dx = \frac{-1}{1+t^2},$$

welches gleichmäßig für $t \in [\alpha, 1]$, $\alpha \in (0,1]$ konvergiert. Somit erfüllt f sämtliche Voraussetzungen von Satz VII-8.5 auf $[0,\infty) \times [\alpha, \beta]$ für alle $\alpha \in (0,1]$. Daher ist F auf $(0,1]$ differenzierbar und es gilt

$$F'(t) = -\int\limits_0^\infty e^{-tx}\sin x\,dx = -\frac{1}{1+t^2}.$$

Folglich ist mit einer Konstanten $c \in \mathbb{R}$

$$F(t) = -\arctan t + c.$$

Der Wert von c ergibt sich aus

$$F(1) = \frac{\pi}{4} = -\arctan 1 + c$$

zu

$$c = \frac{\pi}{2}.$$

Für alle $t \in (0,1]$ gilt daher

$$F(t) = \frac{\pi}{2} - \arctan t.$$

Nach Satz VII-8.3 ist F stetig auf $[0,1]$, soferne $\int_0^\infty f(x,t)\,dx$ gleichmäßig auf $[0,1]$ konvergiert. Nehmen wir dies vorerst an, kann man auf $F(0) = \frac{\pi}{2}$ schließen. Also gilt

$$\int\limits_0^\infty \frac{\sin x}{x}\,dx = \frac{\pi}{2}.$$

Für den Nachweis der gleichmäßigen Konvergenz von $\int_0^\infty f(x,t)\,dx$ für $t \in [0,1]$ bemerken wir, daß es genügt, die uneigentlichen Integrale $\int_1^\infty f(x,t)\,dx$ zu betrachten. Integrieren wir partiell und beachten

$$\int e^{-tx}\sin x\,dx = -\frac{1}{1+t^2}[t\sin x + \cos x]e^{-tx},$$

erhalten wir

$$\begin{aligned}\int_1^\beta e^{-tx}\frac{\sin x}{x}\,dx &= -\frac{1}{x}\frac{e^{-tx}}{1+t^2}[t\sin x+\cos x]\Big|_1^\beta - \frac{1}{1+t^2}\int_1^\beta \frac{1}{x^2}e^{-tx}[t\sin x+\cos x]\,dx\\ &= -\frac{1}{\beta}\frac{e^{-t\beta}}{1+t^2}[t\sin\beta+\cos\beta] + \frac{e^{-t}}{1+t^2}[t\sin 1+\cos 1]\\ &\quad - \frac{1}{1+t^2}\int_1^\beta \frac{1}{x^2}e^{-tx}[t\sin x+\cos x]\,dx.\end{aligned}$$

Der erste Term auf der rechten Seite kann gleichmäßig in $t \in [0,1]$ durch $\frac{2}{\beta}$ abgeschätzt werden. Ebenso kann das Integral durch $2\int_1^\beta \frac{1}{x^2}\,dx$ beschränkt werden. Daraus folgt die gleichmäßige Konvergenz von $\int_1^\infty e^{-tx}\frac{\sin x}{x}\,dx$ für $t \in [0,1]$.

KAPITEL VIII

Differenzierbare Funktionen

1. Differenzierbarkeit

Die Stetigkeit einer Abbildung f an einer Stelle x_0 bedeutet, daß man lokal, dh. in einer Umgebung U von x_0, die Funktionswerte $f(x)$ durch $f(x_0)$ approximieren kann. Natürlich bringt dies i.A. nur eine sehr grobe Näherung an den tatsächlichen Funktionsverlauf. Bedeutend besser kann man den Graph von f durch eine affine Funktion $t_1(x) = f(x_0) + A_{x_0}(x - x_0)$ approximieren. Die Forderung, der Approximationsfehler $r(x) = f(x) - t_1(x)$ möge in einer Umgebung von x_0 klein sein, liefert jedoch allein keine Bedingung für A_{x_0}. Es ist also notwendig, die Anforderungen an die Approximationsgüte zu erhöhen. Eine Präzisierung dieser Idee führt auf das Konzept der Differenzierbarkeit einer Funktion.
Im Folgenden bezeichnen X, Y jeweils normierte Räume mit Normen $||\cdot||_X$ bzw. $||\cdot||_Y$. Wenn die verwendete Norm aus dem Zusammenhang ersichtlich ist, wird auf die Unterscheidung der Normen durch einen Index verzichtet.

Definition VIII-1.1. *Es sei $f\colon U \to Y$, $U \subset X$ offen und $x_0 \in U$.*
1.) f heißt ***FRECHET-differenzierbar*** *an der Stelle $x_0 \in U$, wenn es eine stetige, lineare Abbildung $A_{x_0} \in \mathcal{L}(X,Y)$, eine Kugel $K(0,\delta) \subset X$ und eine Abbildung $r\colon K(0,\delta) \to Y$ gibt mit der Eigenschaft*

i) $\forall h \in K(0,\delta)\colon f(x_0 + h) = f(x_0) + A_{x_0}(h) + r(h)$,

ii) $\lim_{h \to x_0} \frac{r(h)}{||h||} = 0$.

2.) Die lineare Abbildung A_{x_0} heißt ***(FRECHET) Ableitung*** *von f an der Stelle x_0. Man schreibt $f'(x_0) = \frac{df}{dx}(x_0) = Df(x_0) = A_{x_0}$*
3.) f heißt (FRECHET) differenzierbar auf U, wenn f an jeder Stelle $x \in U$ differenzierbar ist. Die Abbildung

$$f'\colon \begin{cases} U \to \mathcal{L}(X,Y) \\ x \mapsto f'(x) \end{cases}$$

heißt Ableitung von f. Anstelle von f' schreibt man auch Df.

Wegen i) muß $r(0) = 0$ gelten. Wir zeigen zuerst, daß diese Definition sinnvoll ist, d.h. daß durch die beiden Bedingungen i) und ii) die Abbildung A_{x_0} und der Korrekturterm r eindeutig festgelegt werden: Angenommen es gäbe $T_i \in \mathcal{L}(X,Y)$ und $r_i\colon K(0,\delta) \to Y$, $i = 1, 2$, mit

$$f(x_0 + h) = f(x_0) + T_1(h) + r_1(h) = f(x_0) + T_2(h) + r_2(h).$$

Dann gilt

$$T_1(h) + r_1(h) = T_2(h) + r_2(h)$$

für alle $h \in K(0,\delta)$. Ersetzt man in dieser Beziehung h durch th mit $0 < t \leq 1$ und dividiert durch t, findet man

$$T_1(h) + \frac{1}{t} r_1(th) = T_2(h) + \frac{1}{t} r_2(th).$$

Beachtet man $\frac{1}{t}||r_i(th)|| = ||h||\,\mathrm{sign}\ t\ \frac{||r_i(th)||}{||th||}$, folgt mit $\lim_{t\to 0} \frac{||r_i(th)||}{||th||} = 0$, $i = 1, 2$,

$$T_1 h = T_2 h$$

für alle $h \in K(0,\delta)$. Wegen der Linearität von T_i folgt daraus $T_1 = T_2$. Die Gleichheit von r_1 und r_2 ist dann eine unmittelbare Folge von $T_1 = T_2$.

Beispiel VIII-1.2. Wir betrachten die affine Abbildung $f\colon X \to Y$, $f(x) = Ax + b$, mit $b \in Y$ und $A \in \mathcal{L}(X,Y)$. Aus der Linearität von A folgt

$$f(x+h) = A(x+h) + b = f(x) + Ah.$$

Setzt man $r(h) = 0$ für alle $h \in X$, folgt die Differenzierbarkeit von f und

$$f'(x) = A$$

für alle $x \in X$.

Differenzierbarkeit ist eine stärkere Eigenschaft einer Funktion als Stetigkeit:

Theorem VIII-1.3. *Ist eine Abbildung $f\colon U \to Y$ differenzierbar in $x_0 \in U$, dann ist f stetig in x_0.*

Beweis. Die Behauptung ergibt sich unmittelbar aus der Darstellung

$$f(x_0 + h) = f(x_0) + f'(x_0)(h) + r(h)$$

und der Stetigkeit von $f'(x_0)$ und r. □

Bemerkung VIII-1.4. 1.) Die Frechet Ableitung wird auch **totale Ableitung** oder **(totales) Differential** der Funktion f genannt.
2.) Gleichwertig ist folgende Charakterisierung der Differenzierbarkeit von f in $x_0 \in U$: es gibt eine stetige, lineare Abbildung $A_{x_0} \in \mathcal{L}(X,Y)$ mit

$$\lim_{h\to 0} \frac{f(x_0+h) - f(x_0) - A_{x_0}(h)}{||h||_X} = 0.$$

Dieser Grenzwert kann auch geschrieben werden als

$$\lim_{x\to x_0} \frac{f(x) - f(x_0) - A_{x_0}(x - x_0)}{||x - x_0||_X} = 0.$$

3.) Sind die Dimensionen der Räume X und Y endlich, genügt es, nur die Linearität von A_{x_0} zu fordern, da eine lineare Abbildung zwischen endlich dimensionalen Räumen automatisch stetig ist.
4.) Aus Satz III-3.11 folgt, daß eine Abbildung $f\colon U \to \mathbb{K}^n$, $f = (f_1, \ldots, f_n)$, in $x_0 \in U$

genau dann differenzierbar ist, wenn jede Komponentenfunktion f_i, $i = 1, \dots, n$, in x_0 differenzierbar ist. Daher kann man sich bei vielen Untersuchungen auf skalare Funktionen beschränken.
5.) Man beachte, daß in Definition VIII-1.1 die Differenzierarkeit nur in einem inneren Punkt des Definitionsbereiches von f betrachtet wird. Dies wird erzwungen durch die Annahme, daß der Definitionsbereich offen ist. Ein Grund für diese Einschränkung besteht darin, daß dadurch Randpunkte ausgeschlossen werden und somit Annäherungen an x_0 aus beliebiger Richtung möglich sind. Das folgende Beispiel zeigt, daß die Eindeutigkeit der Ableitung verloren gehen kann, wenn der Definitionsbereich nicht offen ist:

Beispiel VIII-1.5. Es sei $f\colon \mathbb{R} \times \{0\} \to \mathbb{R}$ gegeben durch $(x, 0) \mapsto 2x$. Eine lineare Abbildung von $\mathbb{R}^2$ nach $\mathbb{R}$ wird durch eine Matrix (α, β) dargestellt. Für die Untersuchung der Differenzierbarkeit von f in $(0, 0)$ betrachten wir daher den Grenzwert

$$\lim_{||(x,0)||\to 0} \frac{1}{||(x,0)||}\Big(f(x,0) - f(0,0) - (\alpha,\beta)\begin{pmatrix} x \\ 0 \end{pmatrix}\Big) \tag{5}$$

$$= \lim_{|x|\to 0} \frac{1}{|x|}(2x - \alpha x) = (2-\alpha) \lim_{|x|\to 0} \operatorname{sign} x \tag{6}$$

(In $\mathbb{R}^2$ können wir z.B. die euklidische Norm verwenden). Durch die Forderung, daß der Grenzwert existiert und gleich Null ist, wird offenbar β nicht festgelegt. Jede durch $(2, \beta)$ beschriebene lineare Abbildung käme als Ableitung von f in $(0, 0)$ in Frage.

Wir betrachten nun den wichtigen Spezialfall reeller Funktionen $f\colon I \to \mathbb{R}$. Das Intervall I sei offen. Wenn f in $x_0 \in I$ differenzierbar ist, dann gibt es eine lineare Abbildung $A_{x_0} \in \mathcal{L}(\mathbb{R}, \mathbb{R})$ mit den Eigenschaften

$$f(x_0 + h) = f(x_0) + A_{x_0}(h) + r(h),$$
$$\lim_{h\to 0} \frac{r(h)}{|h|} = 0.$$

Jede lineare Abbildung $A_{x_0} \in \mathcal{L}(\mathbb{R}, \mathbb{R})$ wird durch eine reelle Zahl $\alpha \in \mathbb{R}$ dargestellt und es gilt $A_{x_0}(h) = \alpha h$ mit $\alpha = A_{x_0}(1)$, und umgekehrt. Diese Überlegung erlaubt es, die Ableitung von f in x_0 mit der reellen Zahl α zu identifizieren. Ist f in $x_0 \in I$ differenzierbar, dann gibt es also $f'(x_0) \in \mathbb{R}$ mit

$$f(x_0 + h) = f(x_0) + f'(x_0)h + r(h),$$
$$\lim_{h\to 0} \frac{r(h)}{h} = 0.$$

Die Ableitung $f'(x_0)$ ist daher gegeben durch den Grenzwert

$$f'(x_0) = \lim_{h\to 0} \frac{1}{h}(f(x_0 + h) - f(x_0)). \tag{7}$$

Existiert umgekehrt der Grenzwert (7) und definiert man $r\colon K(0,\delta) \to \mathbb{R}$ durch

$$f(x_0 + h) = f(x_0) + f'(x_0)h + r(h)$$

für $h \in K(0,\delta)$ (δ so, daß $K(x_0,\delta) \subset I$), erhält man

$$\lim_{h\to 0} \frac{r(h)}{h} = \lim_{h\to 0} \frac{1}{h}(f(x_0+h) - f(x_0)) - f'(x_0) = 0,$$

d.h. f ist differenzierbar in x_0. Für skalare Funktionen in einer Veränderlichen kann die Differenzierbarkeit auch in Häufungspunkten des Definitionsbereiches untersucht werden (vgl. Definition V-1.1).

Abschließend demonstrieren wir den Einsatz von Definition VIII-1.1 für den Nachweis der Differenzierbarkeit einer Funktion in mehreren Veränderlichen.

Beispiel VIII-1.6. Es sei $f\colon \mathbb{R}^3 \to \mathbb{R}$ gegeben durch $f(x,y,z) = zx^2 + ye^z$. Definition VIII-1.1 verlangt, den Zuwachs der Funktionswerte linear zu approximieren. Wir betrachten daher

$$\begin{aligned} f(x+h, y+k, z+l) - f(x,y,z) &= (z+l)(x+h)^2 + (y+k)e^{z+l} - zx^2 - ye^z \\ &= 2xzh + ke^z + l(x^2 + ye^z) + r(h,k,l) \end{aligned}$$

mit

$$r(h,k,l) = zh^2 + 2xlh + lh^2 + ke^z(e^l - 1) + ye^z(e^l - l - 1).$$

Diese Aufspaltung berücksichtigt

$$\begin{aligned} ke^{z+l} &= ke^z + k\mathcal{O}(l), \\ ye^{z+l} - ye^z &= lye^z + \mathcal{O}(l^2). \end{aligned}$$

Wir schätzen nun den Korrekturterm $r(h,k,l)$ unter Verwendung der Maximum Norm $||(h,k,l)||_\infty = \max(|h|,|k|,|l|)$ ab. Als Abkürzung schreiben wir vorübergehend $\delta = (h,k,l)$.

$$\begin{aligned} |r(h,k,l)| \le (|z| + 2|x|)||\delta||_\infty^2 + ||\delta||_\infty^3 + ||\delta||_\infty^2\, e^z |\frac{1}{l}(e^l - 1)| \\ + |y|\, e^z ||\delta||_\infty^2 \sum_{k=2}^{\infty} \frac{|l|^{k-2}}{k!}. \end{aligned}$$

Da $\frac{1}{l}(e^l - 1)$ beschränkt ist (vgl. Lemma V-3.1) und $\sum_{k=2}^{\infty} \frac{|l|^{k-2}}{k!} \le e^{|l|}$ ergibt sich

$$\lim_{(h,k,l)\to(0,0,0)} \frac{|r(h,k,l)|}{||(h,k,l)||_\infty} = 0.$$

Somit ist f auf $\mathbb{R}^3$ differenzierbar und es gilt

$$f'(x,y,z) = (2xz, e^z, x^2 + ye^z).$$

Der Nachweis der Differenzierbarkeit für Funktionen in mehreren Veränderlichen kann noch etwas vereinfacht werden, da der Kandidat für $f'(x_0)$ a priori bekannt ist. Wir kommen darauf in Abschnitt 3 zurück.

2. Rechenregeln für differenzierbare Funktionen

In diesem Abschnitt entwickeln wir einige Regeln, die es erlauben, die Ableitung komplizierterer Funktionen mit Hilfe der bekannten Ableitung einfacherer Funktionen zu berechnen.

Theorem VIII-2.1. *Es seien $f,\, g\colon U \to Y$, $U \subset X$ offen, differenzierbar in $x_0 \in U$ und $\lambda \in \mathbb{K}$. Dann sind auch $f+g$ und λf an der Stelle x_0 differenzierbar und es gilt*

$$\begin{aligned}(f+g)'(x_0) &= f'(x_0) + g'(x_0,)\\ (\lambda f)'(x_0) &= \lambda f'(x_0).\end{aligned}$$

Die Menge der in x_0 differenzierbaren Funktionen bildet somit einen Vektorraum über $\mathbb{K}$.

Beweis. Der Beweis ist eine unmittelbare Folgerung aus den Darstellungen

$$\begin{aligned}f(x_0+h) &= f(x_0) + f'(x_0)(h) + r_f(h),\\ g(x_0+h) &= g(x_0) + g'(x_0)(h) + r_g(h)\end{aligned} \tag{8}$$

mit geeigneten Funktionen r_f und r_g, die sich aus der Differenzierbarkeit von f und g ergeben. □

Für Funktionen mit Werten in $\mathbb{K}$ ist folgende Regel nützlich.

Theorem VIII-2.2. *Es seien $f,\, g\colon U \to \mathbb{K}$, $U \subset X$ offen, differenzierbar in $x_0 \in U$. Dann sind fg und, falls $f(x_0) \neq 0$ ist, auch $\frac{1}{f}$ an der Stelle x_0 differenzierbar und es gilt*

$$\begin{aligned}(fg)'(x_0) &= g(x_0)f'(x_0) + f(x_0)g'(x_0) && \text{Produktregel}\\ (\tfrac{1}{f})'(x_0) &= -\tfrac{1}{f^2(x_0)}f'(x_0) && \text{Reziprokregel}\\ (\tfrac{g}{f})'(x_0) &= \tfrac{1}{f^2(x_0)}(f(x_0)g'(x_0) - g(x_0)f'(x_0)) && \text{Quotientenregel}\end{aligned}$$

Beweis. Mit Hilfe der Darstellungen (8) folgt

$$(fg)(x_0+h) = (fg)(x_0) + g(x_0)f'(x_0)(h) + f(x_0)g'(x_0)(h) + R(h)$$

mit

$$R(h) = (f'(x_0)(h) + r_f(h))(g'(x_0)(h) + r_g(h)) + f(x_0)r_g(h) + g(x_0)r_f(h). \tag{9}$$

Wegen $f'(x_0) \in \mathcal{L}(X,\mathbb{K})$ ist $f'(x_0)$ beschränkt, vgl. Satz III-2.16, d.h. es gilt

$$|f'(x_0)(h)| \le M_f ||h||_X, \quad x \in X$$

mit einer geeigneten Konstanten $M_f \geq 0$ und analog für $g'(x_0)(h)$. Somit kann man $R(h)$ abschätzen durch

$$|R(h)| \leq (M_f + \frac{|r_f(h)|}{||h||_X})(M_g + \frac{|r_g(h)|}{||h||_X})||h||_X^2 + |f(x_0)|\,|r_g(h)| + |g(x_0)|\,|r_f(h)|,$$

was $\lim_{h\to 0} \frac{R(h)}{||h||_X} = 0$ zur Folge hat. Mit der Beobachtung $g(x_0)f'(x_0) + f(x_0)g'(x_0) \in \mathcal{L}(X, \mathbb{K})$ ist die Produktregel bewiesen. Für den Nachweis der Reziprokregel betrachten wir

$$\begin{aligned}
(\frac{1}{f})(x_0 + h) - (\frac{1}{f})(x_0) &= -\frac{f(x_0 + h) - f(x_0)}{f(x_0 + h)f(x_0)} \\
&= -\frac{1}{f(x_0 + h)f(x_0)}(f'(x_0)(h) + r_f(h)) \\
&= -\frac{1}{f(x_0)^2}f'(x_0)(h) + R(h)
\end{aligned}$$

mit

$$R(h) = (\frac{1}{f(x_0)} - \frac{1}{f(x_0 + h)})\frac{1}{f(x_0)}(f'(x_0)(h) + r_f(h)).$$

Aus der Abschätzung

$$|R(h)| \leq |\frac{1}{f(x_0)} - \frac{1}{f(x_0 + h)}|\,|\frac{1}{f(x_0)}|(M_f||h||_X + |r_f(h)|)$$

und der Stetigkeit von f in x_0 folgt

$$\lim_{h\to 0} \frac{1}{||h||}R(h) = 0.$$

Somit ist die Reziprokregel gezeigt. Die Quotientenregel ergibt sich nun durch Kombination der Produkt- und der Reziprokregel. □

Besonders vielfältige Anwendungen findet die Regel für die Ableitung der Verkettung differenzierbarer Funktionen.

THEOREM VIII-2.3 (Kettenregel). *Es seien X, Y, Z normierte Räume, $f\colon U \to Y$, $U \subset X$ offen, $g\colon V \to Z$, $V \subset Y$ offen und $f(U) \subset V$. Sind f in $x_0 \in U$ und g in $f(x_0) \in V$ differenzierbar, dann ist $g \circ f$ in x_0 differenzierbar und es gilt*

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

BEWEIS. Wir greifen wieder auf die lokalen Darstellungen

$$\begin{aligned}
f(x_0 + h) &= f(x_0) + f'(x_0)(h) + r_f(h), \\
g(f(x_0) + k) &= g(f(x_0)) + g'(f(x_0))(k) + r_g(k)
\end{aligned}$$

zurück, welche in $K_X(x_0, \delta)$ beziehungsweise in $K_Y(f(x_0), \delta)$ mit einem geeigneten $\delta > 0$ gelten. Die Komposition von g nach f läßt sich dann schreiben in der Form

$$(g \circ f)(x_0 + h) = g(f(x_0 + h)) = g(f(x_0) + f'(x_0)(h) + r_f(h)).$$

Für h hinreichend klein liegt $k = f'(x_0)(h) + r_f(h)$ in $K_Y(0,\delta)$ und somit gilt

$$\begin{aligned}(g\circ f)(x_0+h) &= (g\circ f)(x_0) + g'(f(x_0))\big(f'(x_0)(h)+r_f(h)\big) + r_g\big(f'(x_0)(h)+r_f(h)\big)\\ &= (g\circ f)(x_0) + \big(g'(f(x_0))\circ f'(x_0)\big)(h) + R(h)\end{aligned}$$

mit

$$R(h) = g'(f(x_0))(r_f(h)) + r_g\big(f'(x_0)(h)+r_f(h)\big).$$

Wegen $g'(f(x_0))\circ f'(x_0) \in \mathcal{L}(X,Z)$ genügt es $\lim_{h\to 0}\frac{1}{||h||}R(h) = 0$ zu zeigen. Dies folgt aus der Beschränktheit der linearen Abbildungen $f'(x_0)$ und $g'(f(x_0))$, d.h. aus den Abschätzungen

$$\begin{aligned}\|f'(x_0)(h)\|_Y &\le M_1\|h\|_X, \qquad h\in X\\ \|g'(f(x_0))(y)\|_Z &\le M_2\|y\|_Y, \qquad y\in Y,\end{aligned}$$

mit passenden Konstanten M_1, $M_2 \ge 0$, kombiniert mit

$$||R(h)||_Z \le M_2||r_f(h)||_Y + \frac{||r_g(f'(x_0)(h)+r_f(h))||_Z}{||f'(x_0)(h)+r_f(h)||Y}(M_1||h||_X + ||r_f(h)||_Y)$$

und den Eigenschaften von r_f und r_g. □

Beispiel VIII-2.4. Es sei $f\colon U\to Y$ injektiv und es seien f in x_0 und f^{-1} in $y_0 = f(x_0)$ differenzierbar. Dann ist nach der Kettenregel $f^{-1}\circ f$ in x_0 differenzierbar und es gilt

$$(f^{-1}\circ f)' = (f^{-1})'(f(x_0))\circ f'(x_0).$$

Wegen $f^{-1}\circ f = \mathrm{id}_X$ folgt mit Beispiel VIII-1.2

$$(f^{-1})'(f(x_0))\circ f'(x_0) = \mathrm{id}_X$$

Ist nun $f'(x_0)$ invertierbar, erhält man

$$(f^{-1})'(f(x_0)) = (f'(x_0))^{-1}.$$

Die Schwachstelle dieser Argumentation ist natürlich, daß die Differenzierbarkeit der Umkehrfunktion bereits gesichert sein muß. Dieser Mangel wird im Folgenden behoben.

Theorem VIII-2.5. *Es sei $f\colon U\to Y$, $U\subset X$ offen, differenzierbar in $x_0\in U$. Ferner sei f injektiv, $V = f(U)$ offen und die Umkehrabbildung $f^{-1}\colon V\to U$ sei stetig an der Stelle $y_0 = f(x_0)$. Besitzt die Ableitung $f'(x_0)\in\mathcal{L}(X,Y)$ eine stetige Inverse, dann ist auch f^{-1} an der Stelle y_0 differenzierbar und es gilt*

$$(f^{-1})'(f(x_0)) = \big(f'(x_0)\big)^{-1}.$$

Beweis. O.B.d.A. können wir $x_0 = 0$ und $y_0 = f(x_0) = 0$ annehmen (ersetze $x\to f(x)$ durch die Abbildung $x\to f(x_0+x) - f(x_0)$. Man überzeuge sich davon, daß auch diese Abbildung die Voraussetzungen des Satzes erfüllt). Da V offen ist, gilt

$K_Y(0,\delta) \subset V$ für $\delta > 0$ hinreichend klein. Wegen der Differenzierbarkeit von f in 0 gilt

$$(*) \qquad f(x) = f'(0)(x) + r_f(x)$$

für alle $x \in K_X(0,\delta_X)$. Da f^{-1} in 0 stetig und V offen ist, gibt es eine Kugel $K_Y(0,\delta_Y) \subset V$ so, daß $f^{-1}(y) \in K_X(0,\delta_X)$ für alle $y \in K_Y(0,\delta_Y)$ zutrifft. Somit gilt die Entwicklung $(*)$ für alle $x = f^{-1}(y)$ mit $y \in K_Y(0,\delta_Y)$ und daher auch

$$y = f'(0)(f^{-1}(y)) + r_f(f^{-1}(y)),$$

für alle $y \in K_Y(0,\delta_Y)$. Nach Voraussetzung besitzt $f'(0)$ eine stetige Inverse, somit folgt

$$(10) \qquad \begin{aligned} f^{-1}(y) &= f'(0)^{-1}(y) - f'(0)^{-1}\big(r_f(f^{-1}(y))\big) \\ &= f'(0)^{-1}(y) + R(y) \end{aligned}$$

mit

$$R(y) = -f'(0)^{-1}\big(r_f(f^{-1}(y))\big).$$

Wegen $f'(0)^{-1} \in \mathcal{L}(Y,X)$ genügt es $\lim_{y\to 0} \frac{1}{||y||} R(y) = 0$ zu zeigen. Dabei verwenden wir wieder die Beschränktheit von $f'(0)^{-1}$,

$$\|f'(0)^{-1}(y)\|_X \le M\|y\|_Y, \qquad y \in Y$$

mit einer geeigneten Konstanten $M \ge 0$. In der Abschätzung

$$\frac{||R(y)||}{||y||} \le M\,\frac{||r_f(f^{-1}(y))||}{||f^{-1}(y)||} \cdot \frac{||f^{-1}(y)||}{||y||} \qquad y \neq 0$$

bemerken wir, daß wegen $f^{-1}(0) = 0$ aus der Injektivität von f auch $f^{-1}(y) \neq 0$ für $y \neq 0$ folgt. Wegen der Stetigkeit von f^{-1} in $y_0 = 0$ folgt weiters $\lim_{y\to 0} f^{-1}(y) = 0$ und somit

$$(11) \qquad \lim_{y\to 0} \frac{||r_f(f^{-1}(y))||}{||f^{-1}(y)||} = 0.$$

Es genügt daher zu zeigen, daß $\frac{||f^{-1}(y)||}{||y||}$ beschränkt ist. Dazu schätzen wir das Wachstum von $||f^{-1}(y)||$ mit Hilfe von (10) folgendermaßen ab

$$||f^{-1}(y)|| \le M\,||y|| + M\frac{||r_f(f^{-1}(y))||}{||f^{-1}(y)||}\,||f^{-1}(y)||$$

Für y hinreichend klein erhält man schließlich mit (11)

$$||f^{-1}(y)|| \le \frac{M}{1 - M\frac{||r_f(f^{-1}(y))||}{||f^{-1}(y)||}}\,||y|| \le 2M\,||y||.$$

$\square$

Bemerkung VIII-2.6. 1.) Im Allgemeinen muß die Existenz *und* Stetigkeit der Inversen von $f'(x_0)$ gefordert werden. Sind X und Y jedoch endlich dimensionale Räume, ist $f'(x_0)^{-1}$ nach Satz III-3.13 stetig (falls die Inverse existiert).
2.) Die Umkehrfunktion einer streng monotonen Funktion $f\colon I \to \mathbb{R}$, $I \subset \mathbb{R}$ ein Intervall, ist automatisch stetig. Ein vergleichbares Resultat existiert für Funktionen in mehreren Veränderlichen nicht. Für reelle Funktionen können daher die Voraussetzungen in Satz VIII-2.5 erheblich abgeschwächt werden:

Theorem VIII-2.7. *Es sei I ein Intervall und $f\colon I \to \mathbb{R}$ streng monoton. Ist f in $x_0 \in I$ differenzierbar mit $f'(x_0) \neq 0$, dann ist f^{-1} in $y_0 = f(x_0)$ differenzierbar und es gilt*

$$(f^{-1})'(f(x_0)) = \frac{1}{f'(x_0)}.$$

Beweis. Die einfachen Modifikationen des Beweises von Satz VIII-2.5 seien dem Leser überlassen. □

3. Richtungsableitungen, partielle Ableitungen, Jakobi-Matrizen

Bei der totalen Ableitung einer Funktion an der Stelle x_0 werden die Funktionswerte von f in einer *vollen* Umgebung von x_0 mit $f(x_0)$ verglichen. Wesentlich schwächer ist das Konzept der Richtungsableitung, bei welchem f nur auf einem durch x_0 verlaufenden Geradensegment ausgewertet wird.

Definition VIII-3.1. *Es sei $f\colon U \to Y$, $U \subset X$ offen, $x_0 \in U$ und $h \in X$. Die Funktion f besitzt in x_0 die* ***Richtungsableitung in Richtung (längs) h****, wenn der Grenzwert*

$$\lim_{t\to 0} \frac{1}{t}(f(x_0 + th) - f(x_0)) \tag{12}$$

existiert. Übliche Bezeichnungen für die Richtungsableitung von f in x_0 längs h sind $\frac{\partial f}{\partial h}(x_0)$ oder $\partial_h f(x_0)$ oder $f'(x_0; h)$.

Da U offen ist, gibt es $\delta > 0$ so, daß $\{x_0 + th\colon |t| < \delta\} \subset U$. Definiert man die Funktion $\phi_h\colon (-\delta, \delta) \to Y$ durch $\phi_h(t) = f(x_0 + th)$, dann bedeutet die Existenz des Grenzwertes (12) die übliche Differenzierbarkeit von ϕ_h in $t = 0$ und es gilt

$$\phi_h'(0) = \frac{\partial f}{\partial h}(x_0)$$

Ersetzt man in (12) die Richtung h durch λh, $\lambda \in \mathbb{R}$, erhält man

$$f'(x_0; \lambda h) = \lim_{t\to 0} \frac{1}{t}(f(x_0 + t\lambda h) - f(x_0)) = \lambda \lim_{t\to 0} \frac{1}{\lambda t}(f(x_0 + t\lambda h) - f(x_0)) = \lambda f'(x_0; h),$$

die Richtungsableitung ist somit homogen bezüglich der Richtung h. Aus diesem Grunde wird ein Richtungsvektor oft normiert, d.h. $||h|| = 1$ gesetzt.

Beispiel VIII-3.2. Die Abbildung $f\colon \mathbb{R}^2 \to \mathbb{R}$ sei gegeben durch

$$f(x,y) = \begin{cases} \frac{xy^2}{x^2+y^4} & (x,y) \neq (0,0), \\ 0 & (x,y) = (0,0). \end{cases}$$

Um die Richtungsableitung von f in $(0,0)$ in Richtung $h = (\xi, \eta)$ zu berechnen, betrachten wir den Differenzenquotienten

$$\begin{aligned} \frac{1}{t}(f(x_0+th) - f(x_0)) &= \frac{1}{t}(f(th) - f(0,0)) \\ &= \frac{1}{t} f(t\xi, t\eta) = \frac{\xi\eta^2}{\xi^2 + t^2\eta^4} \underset{t\to 0}{\longrightarrow} \begin{cases} \frac{\eta^2}{\xi}, & \xi \neq 0, \\ 0, & \xi = 0. \end{cases} \end{aligned}$$

Die Abbildung f besitzt also in $(0,0)$ in jede Richtung h eine Richtungsableitung. Wir erinnern daran, daß f in $(0,0)$ jedoch nicht stetig ist. Somit kann f in $(0,0)$ nicht differenzierbar sein. Dies äußert sich auch darin, daß $h \to f'(0,0)h$ nicht linear ist.

Theorem VIII-3.3. *Ist $f\colon U \to Y$, $U \subset X$ offen, in $x_0 \in U$ differenzierbar, dann ist f in x_0 längs jeder Richtung h differenzierbar und es gilt*

$$\partial_h f(x_0) = f'(x_0)(h).$$

Beweis. Da f in x_0 differenzierbar ist, gilt

$$f(x_0+th) - f(x_0) = tf'(x_0)(h) + r(th),$$

also

$$\frac{1}{t}(f(x_0+th) - f(x_0)) = f'(x_0)(h) + \frac{r(th)}{t||h||}||h||.$$

Die Behauptung folgt nun aus $\lim_{t\to 0} \frac{r(th)}{t||h||} = \lim_{t\to 0} \frac{r(th)}{||th||} \operatorname{sign} t = 0.$ □

Das Beispiel VIII-3.2 zeigt, daß die Umkehrung dieses Satzes falsch ist.
Wir betrachten nun den Spezialfall $X = \mathbb{K}^m$ versehen mit der Standardbasis $(e_1, \dots, e_m)$.

Definition VIII-3.4. *Es sei $f\colon U \to Y$, $U \subset \mathbb{K}^m$ offen, $x_0 \in U$ und $h \in \mathbb{K}^m$. Die Richtungsableitung von f in x_0 in Richtung des j-ten Standardbasisvektors e_j heißt* **j**-*te* ***partielle Ableitung*** *1. Ordnung von f in x_0. Anstelle von $\frac{\partial f}{\partial e_j}(x_0)$ schreibt man $\frac{\partial f}{\partial x_j}(x_0)$ bzw. $\partial_j f(x_0)$ bzw. $f_{x_j}(x_0)$, $j = 1, \dots, m$.*

Wird $x_0 = (\xi_1^0, \dots, \xi_m^0)$ gesetzt, ist die partielle Ableitung nach der j-ten Koordinate an der Stelle x_0 durch den Grenzwert des Differenzenquotienten bestimmt

$$\begin{aligned} \frac{1}{t}(f(x_0+te_j) - f(x_0)) = \frac{1}{t}(&f(\xi_1^0, \dots, \xi_{j-1}^0, \xi_j^0 + t, \xi_{j+1}^0, \dots, \xi_m^0) \\ &- f(\xi_1^0, \dots, \xi_i^0, \dots, \xi_m^0)). \end{aligned}$$

Dies zeigt, daß man die partielle Ableitung nach der j-ten Koordinate an der Stelle x_0 erhält, indem man alle Veränderlichen mit Ausnahme von ξ_j als Konstante betrachtet und die nunmehr nur noch von ξ_j abhängige Funktion in gewohnter Weise differenziert.

Ist f an jeder Stelle $x \in U$ partiell nach ξ_j differenzierbar, sagt man, f sei auf U partiell nach ξ_j differenzierbar und nennt die Abbildung

$$\frac{\partial f}{\partial \xi_j}\colon \begin{cases} U \to \mathbb{R} \\ x \mapsto \frac{\partial f}{\partial \xi_j}(x) \end{cases}$$

die partielle Ableitung von f nach ξ_j auf U.
Die Funktion $f\colon U \to \mathbb{K}^n$, $U \subset \mathbb{K}^m$ offen, $f = (f_1, \dots, f_n)$ sei differenzierbar in $x_0 \in U$. Sowohl in $\mathbb{K}^n$, als auch in $\mathbb{K}^m$ sei die Standardbasis zugrundegelegt. Dann wird die Ableitung $f'(x_0)$ durch eine $n \times m$- Matrix dargestellt, in deren j-ten Spalte der Vektor

$$f'(x_0)(e_j) = \begin{pmatrix} f_1'(x_0)(e_j) \\ \vdots \\ f_n'(x_0)(e_j) \end{pmatrix} = \begin{pmatrix} \frac{\partial f_1}{\partial \xi_j}(x_0) \\ \vdots \\ \frac{\partial f_m}{\partial \xi_j}(x_0). \end{pmatrix}$$

steht. Die Matrixdarstellung von $f'(x_0)$ bezüglich der Standardbasen in $\mathbb{K}^m$ und $\mathbb{K}^n$ ist daher durch die **Jacobimatrix** (**Funktionalmatrix**) gegeben:

$$\begin{pmatrix} \frac{\partial f_1}{\partial \xi_1}(x_0) & \dots & \frac{\partial f_1}{\partial \xi_m}(x_0) \\ \vdots & & \vdots \\ \frac{\partial f_n}{\partial \xi_1}(x_0) & \dots & \frac{\partial f_n}{\partial \xi_m}(x_0) \end{pmatrix}. \tag{13}$$

Da wir in $\mathbb{K}^m$ und $\mathbb{K}^n$ stets die Standardbasen verwenden, werden wir $f'(x_0)$ und die zugehörige Jacobimatrix identifizieren und daher gleich bezeichnen. Wir schreiben $f'(x_0)h$ anstelle von $f'(x_0)(h)$, wenn wir $f'(x_0)$ als Matrix auffassen.

Als Kandidat für $f'(x_0)$ kommt also nur die Jacobimatrix in Frage. Man kann daher beim Nachweis der Differenzierbarkeit einer Funktion $f\colon U \to \mathbb{K}^n$ die gesuchte lineare Abbildung durch die Jacobimatrix ersetzen. Die Differenzierbarkeit von f in x_0 ist daher gleichwertig mit

$$\lim_{h \to 0} \frac{1}{||h||}\Big(f_k(x_0 + h) - f_k(x_0) - \sum_{j=1}^{m} \frac{\partial f_k}{\partial \xi_j}(x_0) h_j\Big) \tag{14}$$

$k = 1, \dots, n$, $h = (h_1, \dots, h_m)$. Beispiel VIII-3.2 zeigt jedoch, daß die Existenz der partiellen Ableitungen allein über die Differenzierbarkeit von f nichts aussagt. Noch deutlicher kommt dies in folgendem Beispiel zum Ausdruck.

Beispiel VIII-3.5. Es sei $f\colon \mathbb{R}^2 \to \mathbb{R}$ gegeben durch $f(t, 0) = f(0, t) = 1$, $t \in \mathbb{R}$ und sonst vollkommen beliebig. Dann existieren die partiellen Ableitungen 1. Ordnung in $(0, 0)$ und es ist $\frac{\partial f}{\partial x}(0, 0) = \frac{\partial f}{\partial y}(0, 0) = 0$. Sonst kann über f keine weitere Aussage gemacht werden.

Es sei $f\colon U \to \mathbb{R}$, $U \subset \mathbb{R}^m$ offen, differenzierbar in $x_0 \in U$. In Bezug auf die Standardbasis im $\mathbb{R}^m$ ist die Jacobimatrix von f gegeben durch die $1 \times m$-Matrix

$$(\frac{\partial f}{\partial \xi_1}(x_0), \dots, \frac{\partial f}{\partial \xi_m}(x_0)).$$

Somit gilt für $h = (\sigma_1, \dots, \sigma_m)$

$$f'(x_0)(h) = \sum_{i=1}^{m} \frac{\partial f}{\partial \xi_j}(x_0)\sigma_i.$$

Oft ist es zweckmäßig, die partiellen Ableitungen erster Ordnung von f zu einem Vektor zusammenzufassen:

Definition VIII-3.6. *$f\colon U \to \mathbb{R}$, $U \subset \mathbb{R}^m$ offen, besitze in x_0 sämtliche partielle Ableitungen $\frac{\partial f}{\partial \xi_i}$, $i = 1, \dots, m$. Der Spaltenvektor*

$$\operatorname{grad} f(x_0) := \left(\frac{\partial f}{\partial \xi_1}(x_0), \dots, \frac{\partial f}{\partial \xi_m}(x_0)\right)^T \in \mathbb{R}^m$$

heißt ***Gradient*** *von f an der Stelle x_0. Ebenfalls gebräuchlich ist die Schreibweise: $\nabla f(x_0)$, eine veraltete Bezeichung ist „Nabla“.*

Wegen Satz VIII-3.3 kann man daher die Richtungsableitung von f in x_0 längs einer Richtung h mit Hilfe des Gradienten als inneres Produkt schreiben

$$(*) \qquad \partial_h f(x_0) = \langle \operatorname{grad} f(x_0), h \rangle.$$

Aus der Ungleichung von Cauchy-Schwarz folgt dann

$$|\partial_h f(x_0)| \le ||\operatorname{grad} f(x_0)||_2 ||h||_2 = ||\operatorname{grad} f(x_0)||_2.$$

Die Änderungsrate von f in x_0 entlang jeder Richtung h ist somit durch die (euklidische) Norm des Gradienten von f in x_0 begrenzt. Diese Schranke ist sogar scharf: dies ist trivial für $\operatorname{grad} f(x_0) = 0$, im Falle $\operatorname{grad} f(x_0) \neq 0$ setzen wir in $(*)$ die Richtung $h^* = \frac{\operatorname{grad} f(x_0)}{||\operatorname{grad} f(x_0)||}$ ein und erhalten

$$\partial_{h^*} f(x_0) = ||\operatorname{grad} f(x_0)||.$$

Lemma VIII-3.7. *Es sei $f\colon U \to \mathbb{R}$, $U \subset \mathbb{R}^m$ offen, differenzierbar in $x_0 \in U$. Dann gibt der Gradient von f in x_0 die Richtung des stärksten Anstieges der Funktionswerte von f an.*

Analog folgt, daß $-\operatorname{grad} f(x_0)$ in die Richtung des stärksten Gefälles der Funktionswerte weist.

Wir kehren noch einmal zur Kettenregel zurück und betrachten die Situation $f\colon U \to \mathbb{K}^p$, $U \subset \mathbb{K}^m$ offen, $g\colon V \to \mathbb{K}^n$, $V \subset \mathbb{K}^p$ offen, $f(U) \subset V$, f sei in $x_0 \in U$ und g in $f(x_0) \in V$ differenzierbar. Nach der Kettenregel ist $h = g \circ f\colon U \to \mathbb{K}^n$ in x_0 differenzierbar und es gilt

$$(15) \qquad h'(x_0) = (g \circ f)'(x_0) = g'(f(x_0)) \circ f'(x_0).$$

Der Verknüpfung der linearen Abbildungen auf der rechten Seite entspricht das Produkt der jeweiligen Jacobimatrizen. Konkret bedeutet (15) in diesem Fall also

(16)
$$\begin{pmatrix} \frac{\partial h_1}{\partial \xi_1}(x_0) & \dots & \frac{\partial h_1}{\partial \xi_m}(x_0) \\ \vdots & & \vdots \\ \frac{\partial h_n}{\partial \xi_1}(x_0) & \dots & \frac{\partial h_n}{\partial \xi_m}(x_0) \end{pmatrix} = \begin{pmatrix} \frac{\partial g_1}{\partial \eta_1}(f(x_0)) & \dots & \frac{\partial g_1}{\partial \eta_p}(f(x_0)) \\ \vdots & & \vdots \\ \frac{\partial g_n}{\partial \eta_1}(f(x_0)) & \dots & \frac{\partial g_n}{\partial \eta_p}(f(x_0)) \end{pmatrix} \begin{pmatrix} \frac{\partial f_1}{\partial \xi_1}(x_0) & \dots & \frac{\partial f_1}{\partial \xi_m}(x_0) \\ \vdots & & \vdots \\ \frac{\partial f_p}{\partial \xi_1}(x_0) & \dots & \frac{\partial f_p}{\partial \xi_m}(x_0) \end{pmatrix}$$

Wir diskutieren nun zwei häufige Anwendungen der Kettenregel. Zuerst diskutieren wir die Ableitung einer Funktion g entlang einer differenzierbaren Kurve γ. Darunter verstehen wir vorerst eine differenzierbare Abbildung $\gamma\colon I \to \mathbb{R}^p$.

Korollar VIII-3.8. Es sei $I \subset \mathbb{R}$ ein offenes Intervall, $\gamma\colon I \to \mathbb{R}^p$ differenzierbar auf I, $g\colon V \to \mathbb{R}$, $\gamma(I) \subset V \subset \mathbb{R}^p$, V offen, differenzierbar auf V. Dann ist die reelle Funktion $\Phi := g \circ \gamma\colon I \to \mathbb{R}$ auf I differenzierbar und es gilt für alle $t \in I$

$$\begin{aligned} \Phi'(t) = g'(\gamma(t)) \circ \gamma'(t) &= \sum_{j=1}^{p} \frac{\partial g}{\partial \xi_j}(\gamma(t))\, \frac{d\gamma_j}{dt}(t) \\ &= \langle \operatorname{grad} g(\gamma(t)), \gamma'(t) \rangle \end{aligned}$$

mit $\gamma(t) = (\gamma_1(t), \dots, \gamma_p(t))$.

Beweis. Satz VIII-2.3. □

Beispiel VIII-3.9. Wir betrachten die Abbildungen

$$\gamma\colon \begin{cases} \mathbb{R} \to \mathbb{R}^2 \\ t \mapsto (t^2+1, t) \end{cases} \qquad g\colon \begin{cases} \mathbb{R}^2 \to \mathbb{R} \\ (\xi, \eta) \mapsto \xi - \eta^2. \end{cases}$$

Man überzeuge sich davon, daß $\Phi = g \circ \gamma$ durch

$$\Phi(t) \equiv 1, \quad t \in \mathbb{R},$$

gegeben ist. Also gilt $\Phi'(t) = 0$, $t \in \mathbb{R}$. Für die Anwendung der Kettenregel benötigen wir

$$\gamma'(t) = \begin{pmatrix} 2t \\ 1 \end{pmatrix} \in \mathbb{R}^{2\times 1}, \quad g'(\xi, \eta) = (1 \quad -2\eta) \in \mathbb{R}^{1\times 2}.$$

Nach der Kettenregel gilt

$$\Phi'(t) = g'(\gamma(t)) \circ \gamma'(t) = (1 \quad -2t) \begin{pmatrix} 2t \\ 1 \end{pmatrix} = 0.$$

Korollar VIII-3.10. Es sei $f\colon U \to \mathbb{R}^p$, $U \subset \mathbb{R}^m$ offen, differenzierbar auf U und $g\colon V \to \mathbb{R}$, $V \subset \mathbb{R}^p$ offen, differenzierbar auf V und $f(U) \subset V$. Dann ist die reellwertige Abbildung $h := g \circ f\colon U \to \mathbb{R}$ differenzierbar auf U und es gilt für alle

$x \in U$

$$\begin{aligned} h'(x) &= g'(f(x)) \circ f'(x) \\ &= \left(\frac{\partial g}{\partial \eta_1}(f(x)) \dots \frac{\partial g}{\partial \eta_p}(f(x))\right) \begin{pmatrix} \frac{\partial f_1}{\partial \xi_1}(x) & \dots & \frac{\partial f_1}{\partial \xi_m}(x) \\ \vdots & & \vdots \\ \frac{\partial f_p}{\partial \xi_1}(x) & \dots & \frac{\partial f_p}{\partial \xi_m}(x) \end{pmatrix}. \end{aligned}$$

Insbesondere gilt

$$(*) \qquad \frac{\partial h}{\partial \xi_i}(x) = \sum_{j=1}^{p} \frac{\partial g}{\partial \eta_j}(f(x)) \frac{\partial f_j}{\partial \xi_i}(x), \quad i = 1, \dots, m.$$

Beweis.

$$\frac{\partial h}{\partial \xi_i}(x) = (h'(x))(e_i) = (g'(f(x)))(f'(x)e_i).$$

□

Der typische Anwendungsbereich dieses Korollars umfaßt Koordinatentransformationen bei reellwertigen Funktionen mehrerer Veränderlicher:

Beispiel VIII-3.11. Es sei $g\colon V \to \mathbb{R}$, $V \subset \mathbb{R}^p$. Oft ist es zweckmäßig, anstelle der cartesischen Koordinaten $(\eta_1, \dots, \eta_p)$ eines Punktes $x \in \mathbb{R}^p$ andere, problemangepaßte Koordinaten $(\xi_1, \dots, \xi_p)$ zu verwenden. Der Zusammenhang zwischen $(\eta_1, \dots, \eta_p)$ und $(\xi_1, \dots, \xi_p)$ wird durch eine zumindest injektive Abbildung $f\colon U \to \mathbb{R}^p$, $U \subset \mathbb{R}^p$, $f(\xi_1, \dots, \xi_p) = (\eta_1(\xi_1, \dots, \xi_p), \dots, \eta_p(\xi_1, \dots, \xi_p))$, beschrieben. In den neuen Variablen $(\xi_1, \dots, \xi_p)$ wird die Abbildung g dann dargestellt durch $h = g \circ f$. Unter den Voraussetzungen von Korollar VIII-3.10 gilt die Ableitungsregel $(*)$,welche häufig in folgender einprägsamen Form geschrieben wird

$$(\dagger) \qquad \frac{\partial h}{\partial \xi_i}(x) = \sum_{j=1}^{p} \frac{\partial g}{\partial \eta_j}(f(x)) \frac{\partial \eta_j}{\partial \xi_i}(x), \quad i = 1, \dots, m,$$

in der die generische unabhängige Variable von g und die Komponentenfunktionen von f gleich bezeichnet werden.

Zur Illustration betrachten wir die Abbildung $g(\eta_1, \eta_2) = \eta_1^2 + \eta_2^2$ in ebenen Polarkoordinaten $(r, \theta) \in (0, \infty) \times (-\pi, \pi]$,

$$\begin{aligned} \eta_1 &= r\cos\theta, \\ \eta_2 &= r\sin\theta. \end{aligned}$$

Der Übergang zu Polarkoordinaten wird beschrieben durch die Abbildung $((\xi_1, \xi_2) \sim (r, \theta))$

$$f\colon \begin{cases} (0, \infty) \times (-\pi, \pi] \to \mathbb{R}^2 \setminus \{(0,0)\}, \\ (r, \theta) \mapsto (\eta_1, \eta_2) = (r\cos\theta, r\sin\theta) \end{cases}$$

und $h(r, \theta) = (g \circ f)(r, \theta) = r^2$. Somit gilt $\frac{\partial h}{\partial r}(r, \theta) = 2r$, $\frac{\partial h}{\partial \theta}(r, \theta) = 0$. Zu demselben Ergebnis gelangen wir auch, wenn wir die Kettenregel in der Form $(*)$ anwenden. Dazu berücksichtigen wir, daß die Abbildung f auf der offenen Menge $(0,\infty)\times(-\pi,\pi)$ differenzierbar ist, und berechnen zuerst die benötigten partiellen Ableitungen für $(r,\theta)\in(0,\infty)\times(-\pi,\pi)$

$$\frac{\partial g}{\partial\eta_1}(\eta_1,\eta_2)=2\eta_1,\qquad \frac{\partial\eta_1}{\partial r}(r,\theta)=\cos\theta,\qquad \frac{\partial\eta_1}{\partial\theta}(r,\theta)=-r\sin\theta,$$
$$\frac{\partial g}{\partial\eta_2}(\eta_1,\eta_2)=2\eta_2,\qquad \frac{\partial\eta_1}{\partial r}(r,\theta)=\sin\theta,\qquad \frac{\partial\eta_2}{\partial\theta}(r,\theta)=r\cos\theta.$$

Aus (†) folgt dann

$$\begin{aligned}\frac{\partial h}{\partial r}(r,\theta)&=\frac{\partial g}{\partial\eta_1}(r\cos\theta,r\sin\theta)\frac{\partial\eta_1}{\partial r}(r,\theta)+\frac{\partial g}{\partial\eta_2}(r\cos\theta,r\sin\theta)\frac{\partial\eta_2}{\partial r}(r,\theta)\\&=2r\cos\theta\cdot\cos\theta+2r\sin\theta\cdot\sin\theta=2r\end{aligned}$$

$$\begin{aligned}\frac{\partial h}{\partial\theta}(r,\theta)&=\frac{\partial g}{\partial\eta_1}(r\cos\theta,r\sin\theta)\frac{\partial\eta_1}{\partial\theta}(r,\theta)+\frac{\partial g}{\partial\eta_2}(r\cos\theta,r\sin\theta)\frac{\partial\eta_2}{\partial\theta}(r,\theta)\\&=2r\cos\theta(-r\sin\theta)+2r\sin\theta(r\cos\theta)=0.\end{aligned}$$

Dieselbe Rechnung wird übersichtlicher in der Matrixschreibweise der Kettenregel dargestellt:

$$\begin{aligned}h'(r,\theta)&=\begin{pmatrix}\frac{\partial h}{\partial r}(r,\theta) & \frac{\partial h}{\partial\theta}(r,\theta)\end{pmatrix}=g'(r\cos\theta,r\sin\theta)\circ f'(r,\theta)\\&=(2r\cos\theta\quad 2r\sin\theta)\begin{pmatrix}\cos\theta & -r\sin\theta\\ \sin\theta & r\cos\theta\end{pmatrix}=(2r\quad 0).\end{aligned}$$

Wir bemerken, daß in Korollar VIII-3.10 die Existenz der partiellen Ableitungen von h und die Darstellung der partiellen Ableitungen bereits unter der schwächeren Voraussetzung bewiesen werden kann, daß die Komponenten von f lediglich alle partiellen Ableitungen auf U besitzen.

4. Differenzierbarkeit komplexer Funktionen

Wir gehen nun etwas näher auf die Differenzierbarkeit komplexer Funktionen $f\colon U\to\mathbb{C}$, $U\subset\mathbb{C}$ offen, ein. Nach Definition VIII-1.1 ist f differenzierbar, wenn es eine (stetige) lineare Abbildung $T\in\mathcal{L}(\mathbb{C},\mathbb{C})$ und eine Funktion $r\colon K(0,\delta)\to\mathbb{C}$ gibt, sodaß

$$(*)\qquad \begin{aligned}&f(z_0+h)=f(z_0)+T(h)+r(h),\qquad h\in K(0,\delta),\\&\lim_{h\to0}\frac{r(h)}{h}=0\end{aligned}$$

gilt. Jede lineare Abbildung $T\colon\mathbb{C}\to\mathbb{C}$ kann dargestellt werden als komplexe Multiplikation mit $a=T(1)$, d.h.

$$T(z)=az.$$

Da $\mathbb{C}$ isomorph zu $\mathbb{R}^2$ ist, kann man f auch als Abbildung $F\colon V \to \mathbb{R}^2$ auffassen, wobei V bestimmt ist durch die Forderung

$$(x,y) \in V \Leftrightarrow z = x + iy \in U.$$

F heißt reelle Interpretation von f und ist definiert durch

$$F(x,y) = (u(x,y), v(x,y)), \qquad (x,y) \in V$$

genau dann, wenn $z = x + iy \in U$ und

$$f(z) = u(z) + iv(z)$$

gilt (etwas unpräzise unterscheiden wir dabei in der Notation nicht zwischen $u(x,y)$ und $u(x+iy)$ und analog für v). Ist $a = \alpha + i\beta$ und $z = x + iy$ dann ist die reelle Interpretation von $z \to az = \alpha x - \beta y + i(\beta x + \alpha y))$ gegeben durch

$$(x,y) \to \begin{pmatrix} \alpha & -\beta \\ \beta & \alpha \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}.$$

Man kann daher die Differenzierbarkeitsbedingung $(*)$ übersetzen in eine Bedingung an die reelle Interpretation F von f. Mit $h = \sigma_1 + i\sigma_2$ und $r(h) = r_1(h) + ir_2(h)$ ergibt sich

$$(\dagger) \qquad \begin{gathered} F(x_0 + \sigma_1, y_0 + \sigma_2) = F(x_0, y_0) + \begin{pmatrix} \alpha & -\beta \\ \beta & \alpha \end{pmatrix} \begin{pmatrix} \sigma_1 \\ \sigma_2 \end{pmatrix} + \begin{pmatrix} r_1(\sigma_1,\sigma_2) \\ r_2(\sigma_1,\sigma_2) \end{pmatrix} \\ \lim_{||(\sigma_1,\sigma_2)||_2 \to 0} \frac{r_1(\sigma_1,\sigma_2)}{||(\sigma_1,\sigma_2)||_2} = 0, \quad \text{und} \quad \lim_{||(\sigma_1,\sigma_2)||_2 \to 0} \frac{r_2(\sigma_1,\sigma_2)}{||(\sigma_1,\sigma_2)||_2} = 0 \end{gathered}$$

(man beachte $|h| = ||(\sigma_1,\sigma_2)||_2$). Somit ist $F\colon V \to \mathbb{R}^2$ Frechet-differenzierbar und es gilt

$$F'(x_0,y_0) = \begin{pmatrix} \alpha & -\beta \\ \beta & \alpha \end{pmatrix}.$$

Andererseits ist $F'(x_0,y_0)$ auch gegeben durch die Jacobimatrix von F, d.h.

$$F'(x_0,y_0) = \begin{pmatrix} u_x(x_0,y_0) & u_y(x_0,y_0) \\ v_x(x_0,y_0) & v_y(x_0,y_0) \end{pmatrix}.$$

Wegen der Eindeutigkeit der Ableitung muß

$$(\ddagger) \qquad \begin{aligned} u_x(x_0,y_0) &= v_y(x_0,y_0) \\ u_y(x_0,y_0) &= -v_x(x_0,y_0) \end{aligned}$$

gelten.
Ist umgekehrt $F\colon V \to \mathbb{R}^2$, $V \subset \mathbb{R}^2$ offen, differenzierbar in $(x_0,y_0) \in V$ und gelten die beiden Gleichungen $(\ddagger)$, dann kann man

$$F'(x_0,y_0)(h) = \begin{pmatrix} u_x(x_0,y_0) & -v_x(x_0,y_0) \\ v_x(x_0,y_0) & u_x(x_0,y_0) \end{pmatrix} \begin{pmatrix} \sigma_1 \\ \sigma_2 \end{pmatrix}$$

interpretieren als $(u_x(x_0,y_0)+iv_x(x_0,y_0))(\sigma_1+i\sigma_2)$ und (†) führt auf die Differenzierbarkeitsbedingung $(*)$ für die komplexe Funktion $z\to f(z)=u(z)+iv(z)$. Somit gilt

Theorem VIII-4.1. *Die komplexe Funktion $f\colon U\to\mathbb{C}$, $f(z)=u(z)+iv(z)$ ist in $z_0=x_0+iy_0$ genau dann differenzierbar, wenn ihre reelle Interpretation $F\colon V\to\mathbb{R}^2$, $F(x,y)=(u(x,y),v(x,y))$, in (x_0,y_0) differenzierbar ist und die* ***Cauchy-Riemannschen Gleichungen***

$$\text{(CR)}\qquad u_x=v_y,\qquad u_y=-v_x$$

in (x_0,y_0) erfüllt sind. Dabei ist $V\subset\mathbb{R}^2$ bestimmt durch die Forderung, $(x,y)\in V\Leftrightarrow x+iy\in U$. Ferner gilt dann

$$\begin{aligned}f'(z_0)&=u_x(x_0,y_0)+iv_x(x_0,y_0),\\ |f'(z_0)|^2&=\det F'(x_0,y_0).\end{aligned}$$

Beispiel VIII-4.2. Wir untersuchen die Differenzierbarkeit der Abbildung $f(z)=\operatorname{Re}z$, $z\in\mathbb{C}$. Mit $z=x+iy$ ist die reelle Interpretation von f gegeben durch

$$F(x,y)=\begin{pmatrix}x\\0\end{pmatrix}=\begin{pmatrix}1&0\\0&0\end{pmatrix}\begin{pmatrix}x\\y\end{pmatrix}.$$

Als lineare Abbildung ist F nach Beispiel VIII-1.2 differenzierbar und es gilt

$$F'(x_0,y_0)=\begin{pmatrix}1&0\\0&0\end{pmatrix}.$$

Da die Cauchy-Riemannschen Gleichungen für kein $(x,y)\in\mathbb{R}^2$ erfüllt sind, ist die Abbildung $z\to\operatorname{Re}z$ nirgendwo (komplex) differenzierbar. Es ist überraschend, daß es im Komplexen so leicht ist, eine Funktion zu finden, welche an keiner Stelle differenzierbar ist. Dies liegt daran, daß komplexe Differenzierbarkeit eine sehr starke Eigenschaft einer Funktion ist: man kann nämlich zeigen, daß die Differenzierbarkeit von f auf U bereits die Existenz aller höheren Ableitungen von f auf U nach sich zieht, d.h. $C^1(U)=C^\infty(U)$! Im Gegensatz dazu ist es sehr mühsam, eine *reelle* Funktion zu konstruieren, die an keiner Stelle eines (offenen) Intervalles differenzierbar ist.

4.1. Höhere partielle Ableitungen. Wir schieben die Diskussion höherer Ableitungen von Abbildungen zwischen normierten Räumen etwas auf und betrachten vorerst nur Funktionen $f\colon U\to\mathbb{R}$, $U\subset\mathbb{R}^m$ offen. Ist f an jeder Stelle $x\in U$ partiell nach ξ_i differenzierbar, sagt man, f sei auf U partiell nach ξ_i differenzierbar und nennt die Abbildung

$$f_{\xi_i}\colon\begin{cases}U\to\mathbb{R}\\ x\mapsto f_{\xi_i}(x)\end{cases}$$

die partielle Ableitung von f nach ξ_i auf U. Anstelle von f_{ξ_i} schreibt man auch $\frac{\partial f}{\partial\xi_i}$ bzw. D_if. Existiert f_{ξ_i} (auf U), kann man versuchen, die partielle Ableitung f_{ξ_i} an einer Stelle $x_0 \in U$ partiell nach ξ_j zu differenzieren, $j = 1, \ldots, m$. Wir erhalten auf diese Weise die zweite partielle Ableitung von f nach ξ_i und ξ_j an der Stelle x_0,

$$f_{\xi_i\xi_j}(\xi_1^0, \ldots, \xi_m^0) =: \frac{\partial^2 f}{\partial\xi_j\partial\xi_i}(\xi_1^0, \ldots, \xi_m^0).$$

Auf analoge Weise definiert man partielle Ableitungen höherer Ordnung.
Wir illustrieren diese Begriffe an einem einfachen Beispiel:

Beispiel VIII-4.3. Es sei $f\colon \mathbb{R}^2 \to \mathbb{R}$ gegeben durch $f(\xi, \eta) = \xi\eta^3 + \xi^2\eta^2$. Man verifiziert

$$\begin{aligned}
\frac{\partial f}{\partial \xi}(\xi,\eta) &= \eta^3 + 2\xi\eta^2, & \frac{\partial f}{\partial \eta}(\xi,\eta) &= 3\xi\eta^2 + 2\xi^2\eta,\\
\frac{\partial^2 f}{\partial \xi^2}(\xi,\eta) &= 2\eta^2, & \frac{\partial^2 f}{\partial \eta\partial\xi}(\xi,\eta) &= 3\eta^2 + 4\xi\eta,\\
\frac{\partial^2 f}{\partial \eta^2}(\xi,\eta) &= 6\xi\eta + 2\xi^2, & \frac{\partial^2 f}{\partial \xi\partial\eta}(\xi,\eta) &= 3\eta^2 + 4\xi\eta,\\
\frac{\partial^3 f}{\partial \xi^3}(\xi,\eta) &= 0, & \frac{\partial^3 f}{\partial \xi\partial\eta\partial\xi}(\xi,\eta) &= 4\eta,\\
\frac{\partial^3 f}{\partial \xi\partial\eta^2}(\xi,\eta) &= 6\eta + 4\xi, & \frac{\partial^3 f}{\partial \xi^2\partial\eta}(\xi,\eta) &= 4\eta,\\
&\text{etc.}
\end{aligned}$$

Es fällt auf, daß die gemischten Ableitungen $\frac{\partial^2 f}{\partial\eta\partial\xi}$, $\frac{\partial^2 f}{\partial\xi\partial\eta}$ bzw. $\frac{\partial^3 f}{\partial\xi\partial\eta\partial\xi}$, $\frac{\partial^3 f}{\partial\xi^2\partial\eta}$ gleich sind. Dies ist kein Zufall, wie der nächste Satz zeigt:

Theorem VIII-4.4 (H.A. Schwarz). *Es sei $f\colon U \to \mathbb{R}$, $U \subset \mathbb{R}^m$ offen und $x_0 \in U$. f besitze auf U die partiellen Ableitungen $\frac{\partial f}{\partial\xi_k}$, $\frac{\partial f}{\partial\xi_l}$ und die gemischte Ableitung $\frac{\partial^2 f}{\partial\xi_k\partial\xi_\ell}$,$k \neq \ell$. Ist die gemischte partielle Ableitung $\frac{\partial^2 f}{\partial\xi_k\partial\xi_\ell}$ stetig in x_0, dann existiert auch $\frac{\partial^2 f}{\partial\xi_\ell\partial\xi_k}$ und es gilt*

$$\frac{\partial^2 f}{\partial\xi_\ell\partial\xi_k}(\xi_1^0, \ldots, \xi_m^0) = \frac{\partial^2 f}{\partial\xi_k\partial\xi_\ell}(\xi_1^0, \ldots, \xi_m^0).$$

Beweis. O.B.d.A. können wir von $\xi_k = \xi_1$ und $\xi_\ell = \xi_2$ ausgehen (anderenfalls ordnet man die Variablen um). Da die übrigen Variablen als Konstante behandelt werden, genügt es, den Satz für eine Abbildung $f\colon U \to \mathbb{R}$, $U \subset \mathbb{R}^2$, zu zeigen, für welche die partiellen Ableitungen erster Ordnung und die gemischte Ableitung $\frac{\partial^2 f}{\partial\xi\partial\eta}$ auf U existieren. Ferner sei $\frac{\partial^2 f}{\partial\xi\partial\eta}$ stetig in $x_0 = (\xi_0, \eta_0)$. Zu $\varepsilon > 0$ gibt es daher eine Kugel $K_\infty(0, \delta)$ so, daß für alle $y = (h, k) \in K_\infty(0, \delta)$

$$\begin{aligned}
&x_0 + y \in U,\\
(\dagger)\qquad &|\frac{\partial^2 f}{\partial\xi\partial\eta}(\xi_0 + h, \eta_0 + k) - \frac{\partial^2 f}{\partial\xi\partial\eta}(\xi_0, \eta_0)| < \varepsilon
\end{aligned}$$

gelten. Für $y = (h,k) \in K_\infty(0,\delta)$ betrachten wir nun den Ausdruck

$$F(h,k) = f(\xi_0 + h, \eta_0 + k) - f(\xi_0, \eta_0 + k) - f(\xi_0 + h, \eta_0) + f(\xi_0, \eta_0).$$

und zeigen

$$F(h,k) = \frac{\partial^2 f}{\partial\xi\partial\eta}(\tilde{\xi},\tilde{\eta})hk$$

für ein $(\tilde{\xi},\tilde{\eta}) \in K_\infty(x_0,\delta)$. Man beachte $F(h,k) = 0$, falls $hk = 0$. Wir können daher $hk \neq 0$ annehmen. Wir bezeichnen mit $I(a,b)$ das Intervall mit den Randpunkten a und b. Die Abbildung

$$\varphi(\eta) = f(\xi_0 + h, \eta) - f(\xi_0, \eta), \quad \eta \in I(\eta_0, \eta_0 + k)$$

ist auf $I(\eta_0, \eta_0 + k)$ differenzierbar. Aus dem Mittelwertsatz folgt daher die Existenz einer Stelle $\tilde{\eta}$ zwischen η_0 und $\eta_0 + k$ mit

$$\begin{aligned} F(h,k) &= \varphi(\eta_0 + k) - \varphi(\eta_0) = \varphi'(\tilde{\eta})k \\ &= \left[\frac{\partial f}{\partial\eta}(\xi_0 + h, \tilde{\eta}) - \frac{\partial f}{\partial\eta}(\xi_0, \tilde{\eta})\right] k. \end{aligned}$$

Nach Voraussetzung existiert $\frac{\partial^2 f}{\partial\xi\partial\eta} = \frac{\partial}{\partial\xi}(\frac{\partial f}{\partial\eta})$ auf U. Wendet man daher noch einmal den Mittelwertsatz auf $\frac{\partial f}{\partial\eta}(\xi_0 + h, \tilde{\eta}) - \frac{\partial f}{\partial\eta}(\xi_0, \tilde{\eta})$ an, ergibt sich für eine Stelle $\tilde{\xi}$ zwischen ξ_0 und $\xi_0 + h$

$$F(h,k) = \frac{\partial^2 f}{\partial\xi\partial\eta}(\tilde{\xi},\tilde{\eta})hk.$$

Wegen $(\tilde{\xi},\tilde{\eta}) \in K_\infty(x_0,\delta)$ und (†) folgt

$$(*) \qquad |\frac{1}{hk}F(h,k) - \frac{\partial^2 f}{\partial\xi\partial\eta}(\xi_0,\eta_0)| = |\frac{\partial^2 f}{\partial\xi\partial\eta}(\tilde{\xi},\tilde{\eta}) - \frac{\partial^2 f}{\partial\xi\partial\eta}(\xi_0,\eta_0)| < \varepsilon.$$

Nun ist

$$\frac{1}{hk}F(h,k) = \frac{1}{k}\Big[\frac{1}{h}\big(f(\xi_0 + h, \eta_0 + k) - f(\xi_0, \eta_0 + k)\big) - \frac{1}{h}\big(f(\xi_0 + h, \eta_0) - f(\xi_0, \eta_0)\big)\Big],$$

und somit

$$\frac{1}{k}\lim_{h\to 0}\frac{1}{h}F(h,k) = \frac{1}{k}(\frac{\partial f}{\partial\xi}(\xi_0, \eta_0 + k) - \frac{\partial f}{\partial\xi}(\xi_0,\eta_0)).$$

Führt man daher den Grenzübergang $h \to 0$ in $(*)$ durch, erhält man daher

$$|\frac{1}{k}(\frac{\partial f}{\partial\xi}(\xi_0, \eta_0 + k) - \frac{\partial f}{\partial\xi}(\xi_0,\eta_0)) - \frac{\partial^2 f}{\partial\xi\partial\eta}(\xi_0,\eta_0)| \leq \varepsilon$$

für $|k| < \delta$. Dies zeigt die Existenz von $\frac{\partial^2 f}{\partial\eta\partial\xi}(\xi_0,\eta_0)$ und

$$\frac{\partial^2 f}{\partial\eta\partial\xi}(\xi_0,\eta_0) = \frac{\partial^2 f}{\partial\xi\partial\eta}(\xi_0,\eta_0).$$

□

5. Mittelwertsatz

5.1. Lokale Extrema. Nach dem Satz von Weierstraß (Korollar III-4.2) nimmt eine stetige, reellwertige Funktion f auf einer kompakten Menge K Maximum und Minimum an. Im Fall des Maximums bedeutet dies: $\exists x_0 \in K\, \forall x \in K\colon f(x) \leq f(x_0)$. Dieser Satz bringt eine *globale* Eigenschaft stetiger Funktionen zum Ausdruck, da sämtliche Funktionswerte zum Vergleich zugelassen sind. Er gibt aber keinerlei Hinweis darauf, *wo* ein globales Extremum liegt. Bei differenzierbaren Funktionen ist es jedoch möglich, aus dem *lokalen* Verhalten der Funktion, d.h. dem Verhalten in einer *Umgebung* einer Stelle x_0, auf das Vorliegen eines Extremums (relativ zur Umgebung) in x_0 zu schließen.

Definition VIII-5.1. *Es sei $f\colon U \to \mathbb{R}$, $U \subset X$ und $x_0 \in U$.*

i) *f besitzt in x_0 ein* ***lokales Maximum*** *(**Minimum**) $\underset{Def}{\Leftrightarrow}$ es gibt eine Umgebung V von x_0 mit der Eigenschaft*

$$\forall x \in V \cap U\colon f(x) \leq f(x_0) \quad (f(x) \geq f(x_0)).$$

ii) *f besitzt in $x_0 \in U$ ein* ***lokales Extremum*** $\underset{Def}{\Leftrightarrow}$ *f besitzt in x_0 ein lokales Maximum oder lokales Minimum.*

Wir formulieren vorerst eine **notwendige** Bedingung für das Vorliegen eines lokalen Extremums.

Theorem VIII-5.2 (Notwendige Optimalitätsbedingung 1. Ordnung). *Es sei $f\colon U \to \mathbb{R}$, $U \subset X$ offen. Ist f differenzierbar in x_0 und besitzt f in x_0 ein lokales Extremum, dann gilt $f'(x_0) = 0$.*

Beweis. O.B.d.A. besitze f in x_0 ein lokales Maximum (anderenfalls betrachte man $-f$). Nach Definition VIII-5.1 gibt es ein $\delta > 0$ mit $f(x_0+h) - f(x_0) \geq 0$ für $h \in K(0,\delta)$. Für ein festes $h \in X$ und $t \in (0,1)$ erhalten wir aus der Differenzierbarkeit von f in x_0

$$f(x_0 + th) - f(x_0) = tf'(x_0)(h) + r(th) \geq 0.$$

Dividiert man diese Ungleichung durch t und führt anschließend den Grenzübergang $t \to 0^+$ durch, erhält man $f'(x_0)(h) \geq 0$. Da h beliebig gewählt war, bedeutet dies $f'(x_0) = 0$. ☐

Bemerkung VIII-5.3. (1) Die notwendige Bedingung für das Auftreten eines lokalen Extremums gilt nur in *inneren* Punkten des Definitionsbereiches von f (wir haben dies erzwungen, indem wir als Definitionsbereich eine offene Menge wählten). Der Satz gilt nicht in den Randpunkten von D: Als Beispiel betrachte man $f = \mathrm{id}\,|_{[0,1]}$. f besitzt ein lokales Minimum in $x = 0$ und ein lokales Maximum in $x = 1$, aber es ist $f'(0) = f'(1) = 1$.
(2) Die Bedingung $f'(x_0) = 0$ ist nicht hinreichend: Für $f(x) = x^3$, $x \in [-1,1]$ gilt $f'(0) = 0$, aber f besitzt in $x = 0$ kein lokales Extremum.
(3) Es gibt auch innere lokale Extrema, die durch Satz VIII-5.2 nicht erfaßt werden:

$f(x) = |x|$, $x \in [-1, 1]$, besitzt in $x = 0$ ein lokales Minimum und ist an dieser Stelle nicht differenzierbar.
(4) Als Kandidaten für lokale Extremstellen einer Abbildung $f : D \to \mathbb{R}$ kommen also in Frage

a) die Randpunkte von D,
b) die Nullstellen von f',
c) jene Stellen, in denen f nicht differenzierbar ist.

5.2. Der Mittelwertsatz für Funktionen in mehreren Veränderlichen. Der Mittelwertsatz läßt sich mühelos auf *reellwertige* Funktionen in mehreren Veränderlichen übertragen.

Theorem VIII-5.4. *Es sei X ein normierter Raum, $f : U \to \mathbb{R}$, $U \subset X$ offen, differenzierbar, x_0 und $x_0 + h$ seien Punkte, die mitsamt ihrer Verbindungsstrecke in U liegen. Dann gibt es ein $\vartheta \in (0, 1)$, so daß gilt*

$$f(x_0 + h) - f(x_0) = f'(x_0 + \vartheta h)(h).$$

Beweis. Wir parametrisieren Verbindungsstrecke zwischen x_0 und $x_0 + h$ durch $\gamma : [0, 1] \to X$, $\gamma(t) = x_0 + th$. Diese Abbildung ist differenzierbar und hat die konstante Ableitung

$$\gamma'(t) = h.$$

Die Abbildung $\varphi : [0, 1] \to \mathbb{R}$, $\varphi(t) = (f \circ \gamma)(t)$ ist nach der Kettenregel differenzierbar. Somit kann der gewöhnliche Mittelwertsatz V-4.6 angewendet werden: Es gibt $\vartheta \in (0, 1)$ mit

$$f(x_0 + h) - f(x_0) = \varphi(1) - \varphi(0) = \varphi'(\vartheta).$$

Andererseits ergibt die Kettenregel für die Ableitung von φ

$$\varphi'(\vartheta) = f'(\gamma(\vartheta))\gamma'(\vartheta) = f'(x_0 + \theta h)h.$$

□

Bemerkung VIII-5.5. Der Mittelwertsatz gilt allerdings nicht in dieser Form (als Gleichung) für vektorwertige Funktionen. Dies liegt daran, daß die Zwischenstellen $x_0 + \vartheta h$ für die einzelnen Komponentenfunktionen i.a. nicht gleich sind: Es sei $f : \mathbb{R} \to \mathbb{R}^2$ die Abbildung $t \mapsto (t - t^2, t - t^3)$. Es ist also $f(1) - f(0) = (0, 0)$. Die Ableitung ist gegeben durch $f'(t) = (1 - 2t, 1 - 3t^2)$. Da $f'(t) \neq 0$ für alle $t \in \mathbb{R}$ gilt, kann die Gleichung $f(1) - f(0) = f'(t)(1 - 0)$ für keinen Wert von t zutreffen. Einen teilweisen Ersatz bietet folgende Ungleichung.

Theorem VIII-5.6. *Es seien X und Y normierte Räume.*
1.) E sei $f : [a, b] \to Y$, stetig und differenzierbar auf (a, b). Ferner sei f' beschränkt, d.h. es gelte mit einer Konstanten $M \geq 0$ die Abschätzung

$$||f'(t)s|| \leq M|s|, \qquad \text{für } t \in (a, b) \text{ und } s \in \mathbb{R}.$$

Dann gilt für alle t, $s \in [a, b]$

$$||f(t) - f(s)|| \leq M|t - s|.$$

2.) Es sei $f\colon U \to Y$, $U \subset X$ offen. Die Verbindungsstrecke $[x,y]$ zweier Elemente $x,\ y \in U$ liege in U. Ferner sei f auf $[x,y]$ differenzierbar und es gelte mit einer Konstanten $M \geq 0$

$$\|f'(z)h\| \leq M\|h\| \qquad \textit{für alle } z \in [x,y] \textit{ und } h \in X.$$

Dann folgt

$$||f(y) - f(x)||_Y \leq M||y-x||_X.$$

Beweis. ad 1.) Wegen der Stetigkeit von f auf $[a,b]$ genügt es, die Ungleichung

$$||f(t) - f(s)|| \leq M|t-s|$$

für alle $a < s < t < b$ zu zeigen. Dazu betrachten wir für beliebiges $\varepsilon > 0$

$$S_\varepsilon = \{\alpha \in [s,t] \colon ||f(\beta) - f(s)|| \leq (M+\varepsilon)(\beta - s), \quad \beta \in [s,\alpha]\}.$$

Wegen $s \in S_\varepsilon$ ist S_ε nicht leer. Da S_ε beschränkt ist, existiert $\alpha^* = \sup S_\varepsilon$. Die Menge S_ε ist auch abgeschlossen: betrachten wir eine Folge $(\alpha_n) \subset S_\varepsilon$, die gegen $\bar\alpha \in [s,t]$ konvergiert. Ist $\bar\alpha = s$, gilt $\bar\alpha \in S_\varepsilon$. Es sei $s < \bar\alpha \leq t$. Dann gibt es für jedes $\alpha < \bar\alpha$ ein Folgenglied α_n mit $\alpha < \alpha_n$. Dies hat $\alpha \in S_\varepsilon$ zur Folge. Somit gilt

$$||f(\beta) - f(s)|| \leq (M+\varepsilon)(\beta - s)$$

für alle $\beta \in [s,\bar\alpha)$ und wegen der Stetigkeit von f auch auf $[s,\bar\alpha]$.
Wegen der Abgeschlossenheit S_ε erhält man $\alpha^* = \max S_\varepsilon$. Angenommen, es wäre $\alpha^* < t$. Da f in α^* differenzierbar ist, gäbe es zu dem gewählten $\varepsilon > 0$ ein $\delta = \delta(\varepsilon, \alpha^*)$ mit der Eigenschaft

$$||f(\beta) - f(\alpha^*) - f'(\alpha^*)(\beta - \alpha^*)|| \leq \varepsilon|\beta - \alpha^*|$$

für $|\beta - \alpha^*| < \delta$. Somit folgt für $\alpha^* \leq \beta < \alpha^* + \delta$

$$\begin{aligned}||f(\beta) - f(\alpha^*)|| &\leq ||f(\beta) - f(\alpha^*) - f'(\alpha^*)(\beta - \alpha^*)|| + ||f'(\alpha^*)(\beta - \alpha^*)|| \\ &\leq \varepsilon(\beta - \alpha^*) + ||f'(\alpha^*)||(\beta - \alpha^*) \leq (M+\varepsilon)(\beta - \alpha^*)\end{aligned}$$

und daher

$$\begin{aligned}||f(\beta) - f(s)|| &\leq ||f(\beta) - f(\alpha^*)|| + ||f(\alpha^*) - f(s)|| \\ &\leq (M+\varepsilon)(\beta - \alpha^* + \alpha^* - s) = (M+\varepsilon)(\beta - s).\end{aligned}$$

Dies hat $[s, \alpha^* + \delta) \subset S_\varepsilon$ zur Folge und somit muß $\alpha^* = t$ gelten. Folglich gilt

$$||f(t) - f(s)|| \leq (M+\varepsilon)|t-s|$$

für alle $a < s < t < b$ und $\varepsilon > 0$. Führt man nun den Grenzübergang $\varepsilon \to 0^+$ und anschließend $s \to a^+$ und $t \to b^-$ durch, erhält man die gesuchte Ungleichung.
ad 2.) Wir betrachten die Funktion $g(t) = f(x + t(y-x))$ als Verkettung $g = f \circ \varphi$, mit $\varphi(t) = x + t(y-x)$, $t \in [0,1]$. Dann folgt $g(1) - g(0) = f(y) - f(x)$ und g ist differenzierbar mit $g'(t) = f'(x + t(y-x))(y-x)$. Somit erhalten wir aus Teil 1

$$||f(y) - f(x)|| \leq (\sup_{z \in [x,y]} ||f'(z)||)||y-x||.$$

□

Definition VIII-5.7. *Es sei $(X, ||\cdot||)$ ein normierter Raum und $-\infty < a \leq b < \infty$.*
1.) Jede stetige Abbildung $\gamma : [a, b] \to X$ heißt ***(parametrisierte) Kurve*** *(stetiger Weg), mit Anfangspunkt $\gamma(a)$ und Endpunkt $\gamma(b)$. Das Bild $\gamma([a, b])$ heißt* ***Spur*** *der Kurve.*
2.) Eine Teilmenge M von X heißt ***wegzusammenhängend****, wenn es zu jedem Punktepaar $(x, y) \in M \times M$ eine Kurve mit Anfangspunkt x und Endpunkt y gibt, deren Spur zur Gänze in M liegt.*
3.) Für je zwei Punkte x, $y \in X$ ist $\gamma\colon [0, 1] \to X$, $\gamma(t) = x + t(y - x)$ eine stetige Parametrisierung eines Geradensegmentes mit Anfangspunkt x und Endpunkt y. Dieses wird auch mit $[x, y]$ bezeichnet. Eine Teilmenge $M \subset X$ heißt ***konvex*** *, wenn mit je zwei Punkten x, $y \in M$ auch die Verbindungsstrecke $[x, y]$ zu M gehört*

Somit ist unter den Voraussetzungen von Satzes VIII-5.6 f Lipschitz stetig, falls U konvex ist.
Ist M offen, kann man mit einem Kompaktheitsargument zeigen, daß in der Definition des Wegzusammenhangs als parametrisierte Kurve sogar ein polygonaler Weg gewählt werden kann, d.h. es gibt Punkte x_i, $i = 0, \ldots, n$, $x_0 = x$, $x_n = y$, so, daß $[x_{i-1}, x_i] \subset M$ für $i = 1, \ldots, n$ gilt.

Korollar VIII-5.8. Die offene Menge $U \subset X$ sei konvex und $f\colon U \to Y$ differenzierbar. Ist $f'(x) = 0$ für alle $x \in U$, dann ist f konstant.

Theorem VIII-5.9. *$U \subset X$ sei offen und wegzusammenhängend und f, $g\colon U \to Y$ differenzierbar.*

(1) *Ist $f' = 0$ auf U, dann ist f konstant.*
(2) *Ist $f' = g'$ auf U, dann gibt es eine Konstante $c \in Y$ mit $f = g + c$. Ist $f(x_0) = g(x_0)$ für mindestens ein $x_0 \in U$, gilt $f = g$.*

Beweis. Es seien x, $y \in U$ beliebig gewählt. Da U wegzusammenhängend ist, kann man x und y durch einen polygonalen Weg verbinden, d.h. es gibt Punkte x_i, $i = 0, \ldots, n$, $x_0 = x$, $x_n = y$, so, daß jedes Segment $[x_{i-1}, x_i]$ in U liegt. Als Folge von Satz VIII-5.6 ist f auf $[x_{i-1}, x_i]$, $1 \leq i \leq n$, konstant. Somit folgt $f(x) = f(x_0) = f(x_1) = \cdots = f(x_n) = f(y)$. Also ist f konstant. □

Bemerkung VIII-5.10. 1.) Die einzigen zusammenhängenden Teilmengen in $\mathbb{R}$ sind Intervalle. Das Beispiel $f \equiv -1$ auf $[0, 1]$ und $f \equiv 1$ auf $[2, 3)$ zeigt, daß der Zusammenhang des Definitionsbereiches notwendig ist.

6. Differenzierbarkeit von Funktionen in mehreren Veränderlichen

Aus dem Mittelwertsatz ergibt sich sich auch ein sehr praktisches Kriterium für die Differenzierbarkeit von Funktionen in mehreren Veränderlichen.

Theorem VIII-6.1. *Die Abbildung $f\colon U \to \mathbb{R}^n$, $U \subset \mathbb{R}^m$ offen, besitze auf U sämtliche partiellen Ableitungen $\frac{\partial f_j}{\partial \xi_i}(x)$, $j = 1, \ldots, n$, $i = 1, \ldots, m$. Sind diese stetig an der Stelle $x_0 \in U$, dann ist f in x_0 differenzierbar.*

Beweis. Es genügt, den Satz für reellwertige Funktionen $f\colon U \to \mathbb{R}$ zu beweisen. Da als Ableitung nur die lineare Abbildung mit der Matrixdarstellung $f'(x_0) = \left(\frac{\partial f}{\partial \xi_1}(x_0) \dots \frac{\partial f}{\partial \xi_m}(x_0)\right) \in \mathbb{R}^{1\times m}$ in Frage kommt, müssen wir

$$(*) \qquad \lim_{h\to 0} \frac{f(x_0+h) - f(x_0) - f'(x_0)h}{\|h\|} = 0$$

nachweisen. Wir verbinden x_0 und x_0+h, $h = (\sigma_1, \dots, \sigma_m)$, durch einen achsenparallelen Weg durch die Punkte x_i,

$$x_i = x_{i-1} + \sigma_i e_i, \quad i = 1, \dots, m,$$

also $x_i = x_0 + \sum_{j=1}^{i} \sigma_j e_j$ und $x_m = x_0 + h$. Für h hinreichend klein liegt dieser Weg in U. Man kann daher $f(x_0+h) - f(x_0)$ in der Form

$$f(x_0+h) - f(x_0) = \sum_{i=1}^{m} \big(f(x_i) - f(x_{i-1})\big).$$

schreiben. Es sei $x_{i-1} = (\xi_1^{i-1}, \dots, \xi_m^{i-1})$, $i = 1, \dots, m$. Aus

$$\begin{aligned} f(x_i) - f(x_{i-1}) &= f(x_{i-1} + \sigma_i e_i) - f(x_{i-1}) \\ &= f(\xi_1^{i-1}, \dots, \xi_i^{i-1} + \sigma_i, \dots, \xi_m^{i-1}) - f(\xi_1^{i-1}, \dots, \xi_i^{i-1}, \dots, \xi_m^{i-1}) \end{aligned}$$

folgt mit Hilfe des Mittelwertsatzes V-4.6

$$f(x_i) - f(x_{i-1}) = \frac{\partial f}{\partial \xi_i}(\xi_1^{i-1}, \dots, \xi_i^{i-1} + \theta_i\sigma_i, \dots, \xi_m^{i-1})\sigma_i$$

mit einem geeigneten $\theta_i \in (0,1)$. Bezeichnen wir die Zwischenstellen $(\xi_1^{i-1}, \dots, \xi_i^{i-1} + \theta_i\sigma_i, \dots, \xi_m^{i-1})$ mit $x_i^*(h)$, ergibt sich schließlich

$$f(x_0+h) - f(x_0) - f'(x_0)h = \sum_{i=1}^{m} \left(\frac{\partial f}{\partial \xi_i}(x_i^*(h)) - \frac{\partial f}{\partial \xi_i}(x_0)\right)\sigma_i,$$

und weiter

$$|f(x_0+h) - f(x_0) - f'(x_0)h| \le \sum_{i=1}^{m} \Big|\frac{\partial f}{\partial \xi_i}(x_i^*(h)) - \frac{\partial f}{\partial \xi_i}(x_0)\Big| \|h\|_\infty.$$

Wegen $x_i^*(h) = x_{i-1} + \theta_i\sigma_i e_i$ folgt

$$\|x_i^*(h) - x_0\|_\infty = \|x_{i-1} - x_0 + \theta_i\sigma_i e_i\|_\infty = \|\sum_{j=1}^{i-1} \sigma_j e_j + \theta_i\sigma_i e_i\|_\infty \le \|h\|_\infty,$$

d.h. $\lim_{h\to 0} x_i^*(h) = x_0$, $i = 1, \dots, m$. Wegen der Stetigkeit der partiellen Ableitungen ergibt sich nun die Behauptung, wenn wir in $(*)$ die Norm $\|\cdot\|_\infty$ verwenden. □

Definition VIII-6.2. *$f\colon U \to \mathbb{R}^n$, $U \subset \mathbb{R}^m$ offen, sei differenzierbar und $x_0 \in U$.*

*f heißt **stetig differenzierbar** (an der Stelle x_0) $\underset{Def}{\Leftrightarrow}$*

$f'\colon U \to \mathcal{L}(\mathbb{R}^m, \mathbb{R}^n)$, $x \mapsto f'(x)$, ist stetig (an der Stelle x_0).

Wodurch ist die Stetigkeit einer Abbildung $F\colon U \to \mathcal{L}(\mathbb{R}^m, \mathbb{R}^n)$ charakterisiert? Ohne Beschränkung der Allgemeinheit genügt es, den Fall $n = 1$ zu betrachten.

THEOREM VIII-6.3. *Es sei $F = (f_1, \dots, f_m)\colon U \to \mathcal{L}(\mathbb{R}^m, \mathbb{R})$, $U \subset \mathbb{R}^m$ und $f_i\colon U \to \mathbb{R}$. Die Abbildung F ist in $x_0 \in U$ stetig genau dann, wenn f_i, $i = 1, \dots, m$, in x_0 stetig ist.*

BEWEIS. Verwendet man in $\mathbb{R}^m$ die Norm $\|x\| = \max\{|\xi_i|, 1 \le m\}$, ergibt sich für die lineare Abbildung $F(x) = (f_1(x), \dots, f_m(x)$ die Norm $\|F(x)\|_{\mathcal{L}(\mathbb{R}^m,\mathbb{R})} = \sum_{i=1}^m |f_i(x)|$. Es sei F stetig in x_0. Dann gibt es zu $\varepsilon > 0$ eine δ-Umgebung $K(x_0, \delta)$, von x_0 mit

$$\|F(x) - F(x_0)\|_{\mathcal{L}(\mathbb{R}^m,\mathbb{R})} = \sum_{i=1}^m |f_i(x) - f_i(x_0)| < \varepsilon$$

für $x \in K(x_0, \delta)$. Dies hat die Stetigkeit der Komponenentenfunktionen f_i in x_0 zur Folge. Umgekehrt seien nun die Funktionen f_i, $i = 1 \dots, m$, stetig in x_0. Zu $\varepsilon > 0$ wählt man $\delta > 0$ so, daß

$$|f_i(x) - f_i(x_0)| < \frac{\varepsilon}{m}, \qquad 1 \le i \le m$$

für alle $x \in K(x_0, \delta)$ gilt. Für $h = (\sigma_1, \dots, \sigma_m)$ und $x \in K(x_0, \delta)$ folgt dann die Abschätzung

$$\begin{aligned} |F(x)h - F(x_0)h| &= |\sum_{i=1}^m (f_i(x)\sigma_i - f_i(x_0)\sigma_i)| \\ &\le \sum_{i=1}^m |(f_i(x) - f_i(x_0))|\,|\sigma_i| \le \sum_{i=1}^m |(f_i(x) - f_i(x_0))|\,\|h\| < \varepsilon\|h\|, \end{aligned}$$

also

$$\|F(x) - F(x_0)\|_{\mathcal{L}(\mathbb{R}^m,\mathbb{R})} < \varepsilon.$$

Somit ist F stetig in x_0. □

Nach Satz VIII-6.1 ist die Stetigkeit von f' somit gleichbedeutend mit der Stetigkeit der Elemente der Jacobimatrix:

THEOREM VIII-6.4. *Es sei $f\colon U \to \mathbb{R}^n$, $U \subset \mathbb{R}^m$ offen. Äquivalent sind*

i) *f ist stetig differenzierbar,*
ii) *alle partiellen Ableitungen $\frac{\partial f_i}{\partial x_j}$, $i = 1 \dots, n$, $j = 1 \dots, m$ sind stetig auf U.*

BEISPIEL VIII-6.5. $f\colon \mathbb{R}^2 \to \mathbb{R}$ sei gegeben durch $f(\xi, \eta) = e^{\xi\eta}$. Die Existenz der partiellen Ableitungen $\frac{\partial}{\partial\xi} f(\xi, \eta) = \eta e^{\xi\eta}$, $\frac{\partial}{\partial\eta} f(\xi, \eta) = \xi e^{\xi\eta}$ ist trivial, deren Stetigkeit als Abbildung von $\mathbb{R}^2 \to \mathbb{R}$ leicht zu argumentieren. Daher ist f auf $\mathbb{R}^2$ stetig differenzierbar. Man versuche zum Vergleich den Nachweis der Differenzierbarkeit an einer Stelle (ξ_0, η_0) an Hand der Definition.

Man kann zeigen, daß auch höhere Ableitungen auf eine ähnliche Weise charakterisiert werden können. Wir vereinbaren vorerst:

Definition VIII-6.6. *Es sei $U \subset \mathbb{R}^m$ offen und $k \in \mathbb{N}_0$.*

$$C^k(U) := \{f\colon U \to \mathbb{R}\colon \text{ alle partiellen Ableitungen von } f \text{ der Ordnung} \leq k \text{ existieren und sind auf } U \text{ stetig}\}.$$

6.1. Höhere Ableitungen. In diesem Abschnitt gehen wir etwas näher auf höhere Ableitungen ein und beschränken uns vorerst auf die Ableitungen zweiter Ordnung. Es sei $f\colon U \to Y$, $U \subset X$ offen, differenzierbar. Dann existiert die Ableitung $f'\colon U \to \mathcal{L}(X, Y)$. Ist diese Abbildung $x_0 \in U$ differenzierbar erhält man die zweite Ableitung $f''(x_0) \in \mathcal{L}(X, \mathcal{L}(X, Y))$. Betrachten wir also vorerst Abbildungen vom Typ $F \in \mathcal{L}(X, \mathcal{L}(Y, Z))$. Dann existiert eine Konstante $M \geq 0$ derart, daß

$$\|F(x)\|_{\mathcal{L}(X,Y)} \leq M\|x\|_X$$

für alle $x \in X$ gilt. Für jedes $x \in X$ ist $F(x)$ eine stetige lineare Abbildung von Y nach Z. Es sei $F(x)y$ das Bild von $y \in Y$ unter $F(x)$. Wegen der Definition der Norm in $\mathcal{L}(X, Y)$ folgt dann

$$\|F(x)y\|_Z \leq M\|x\|_X\|y\|_Y.$$

Umgekehrt folgt aus obiger Abschätzung für eine Abbildung $F \in L(X, L(Y, Z))$, daß sogar $F \in \mathcal{L}(X, \mathcal{L}(Y, Z))$ zutrifft ($L(X, Y)$ bezeichnet den Vektorraum der linearen, nicht notwendig stetigen Abbildungen von X nach Y). Lineare Abbildungen von diesem Typ sind eng mit bilinearen Abbildungen verknüpft.

Definition VIII-6.7. *1.) Es seien X, Y und Z normierte Räume. Eine Abbildung $b\colon X \times Y \to Z$ heißt* ***bilinear****, wenn*

(1) *$x \to b(x, y)$ für jedes $y \in Y$ linear und*
(2) *$y \to b(x, y)$ für jedes $x \in X$ linear ist.*

2.) Eine bilineare Abbildung $b\colon X \times Y \to Z$ heißt beschränkt, wenn es eine Konstante $M \geq 0$ gibt, so daß

$$\|b(x, y)\|_Z \leq M\|x\|_X\|y\|_Y$$

für alle $(x, y) \in X \times Y$ gilt.
3.) Eine bilineare Abbildung $b\colon X \times X \to Z$ heißt symmetrisch, wenn für alle x, $y \in X$

$$b(x, y) = b(y, x)$$

zutrifft.
4.) Mit $\mathcal{L}_2(X \times Y, Z)$ bezeichnet man den Vektorraum der bilinearen, beschränkten Abbildungen von $X \times Y$ nach Z.

Beispiel VIII-6.8. 1.) Es sei $A \in \mathbb{R}^{n\times m}$ und $b(x, y) = y^T Ax$ für $(x, y) \in \mathbb{R}^m \times \mathbb{R}^n$. Eine einfache Rechnung zeigt, daß b eine beschränkte, bilineare Abbildung von $\mathbb{R}^m \times \mathbb{R}^n$ nach $\mathbb{R}$ ist.
2.) Setzt man $X = Y = \mathbb{R}^m$ und $Z = \mathbb{R}$, dann definiert

$$b(x, y) = \langle x, y\rangle$$

eine beschränkte, bilineare Abbildung von $\mathbb{R}^m \times \mathbb{R}^m$ nach $\mathbb{R}$.

Man kann sich leicht davon überzeugen, daß

$$\|b\|_{\mathcal{L}_2(X\times Y,Z)} = \sup_{x\neq 0, y\neq 0} \frac{\|b(x,y\|_Z}{\|x\|_X\|y\|_Y}$$

eine Norm in $\mathcal{L}_2(X\times Y,Z)$ definiert.
Es sei nun $b\in\mathcal{L}_2(X\times Y,Z)$. Für jedes $x\in X$ betrachtet man die Abbildung $\lambda_x\colon Y\to Z$

$$\lambda_x(y) = b(x,y).$$

Dann gilt für α, $\beta\in\mathbb{R}$ und y, $\tilde{y}\in Y$

$$\lambda_x(\alpha y+\beta\tilde{y}) = b(x,\alpha y+\beta\tilde{y}) = \alpha b(x,y)+\beta b(x,\tilde{y}) = \alpha\lambda_x(y)+\beta\lambda_x(\tilde{y}),$$

d.h. λ_x ist linear. Für α, $\beta\in\mathbb{R}$ und x, $\tilde{x}\in X$ erhält man für jedes $y\in Y$

$$\lambda_{\alpha x+\beta\tilde{x}}(y) = b(\alpha x+\beta\tilde{x},y) = \alpha b(x,y)+\beta b(\tilde{x},y) = \alpha\lambda_x(y)+\beta\lambda_{\tilde{x}}(y).$$

Somit ist auch die Abbildung $\lambda\colon X\to L(X,Y)$, $x\to\lambda_x$ linear, also $\lambda\in L(X,L(Y,Z))$. Wegen der Abschätzung

$$\|\lambda_x(y)\|_Z \leq M\|x\|_X\|y\|_Y$$

gilt sogar $\lambda\in\mathcal{L}(X,\mathcal{L}(Y,Z))$.
Umgekehrt sei $F\in\mathcal{L}(X,\mathcal{L}(Y,Z))$ und $b_F\colon X\times Y\to Z$ definiert durch

$$b_F(x,y) = F(x)y.$$

Eine einfache Rechnung zeigt, daß b_F bilinear und wegen

$$\|b_F(x,y)\|_Z = \|F(x)y\|_Z \leq M\|x\|_X\|y\|_Y$$

sogar beschränkt ist. Man kann daher die Räume $\mathcal{L}(X,\mathcal{L}(Y,Z))$ und $\mathcal{L}_2(X\times Y,Z)$ identifizieren. Dazu betrachte man die Abbildung

$$\sigma\colon\mathcal{L}(X,\mathcal{L}(Y,Z))\to\mathcal{L}_2(X\times Y,Z), \qquad F\to b_F$$

Gilt $\sigma(F)=\sigma(\tilde{F})$, also $b_F=b_{\tilde{F}}$ folgt $F(x)y=\tilde{F}(x)y$ für alle $x\in X$ und $y\in Y$. Dies hat insbesondere die Gleichheit der linearen Abbildungen $F(x)$ und $\tilde{F}(x)$ für alle $x\in X$ und somit $F=\tilde{F}$ zur Folge. Also ist σ injektiv. In den vorangehenden Betrachtungen wurde gezeigt, daß jedes $b\in\mathcal{L}_2(X\times Y,Z)$ eine Abbildung $\lambda\in\mathcal{L}(X,\mathcal{L}(Y,Z))$ definiert. Für diese Abbildung gilt

$$\sigma(\lambda)(x,y) = \lambda_x(y) = b(x,y), \qquad (x,y)\in X\times Y.$$

Dies zeigt die Surjektivität und damit auch die Bijektivität von σ. Betrachtet man $\mathcal{L}(X,\mathcal{L}(Y,Z))$ und $\mathcal{L}_2(X\times Y,Z)$ als normierte Räume, dann ist σ sogar eine Isometrie, das heißt es gilt

$$\|\sigma(F)\|_{\mathcal{L}_2(X\times Y,Z)} = \|F\|_{\mathcal{L}_2(X\times Y,Z)}.$$

Der einfache Beweis sei dem Leser überlassen.

Wir betrachten nun den Spezialfall $X = Y = \mathbb{R}^m$ und $Z = \mathbb{R}$. Dann ist $\mathcal{L}(\mathbb{R}^m, (\mathbb{R}^m, \mathbb{R}))$ isomorph zu $\mathcal{L}_2(\mathbb{R}^m \times \mathbb{R}^m, \mathbb{R})$. Legt man dem $\mathbb{R}^m$ die kanonische Basis zugrunde, erhält man für eine Bilinearform $b \in \mathcal{L}(\mathbb{R}^m, (\mathbb{R}^m, \mathbb{R})$ und $x = \sum_{i=1}^m \xi_i e_i$, $y = \sum_{i=1}^m \eta_i e_i$

$$b(x, y) = b(\sum_{j=1}^m \xi_j e_j, \sum_{i=1}^m \eta_i e_i) = \sum_{i,j=1}^m b(e_j, e_i)\xi_j \eta_i$$

Definiert man die Matrix $B \in \mathbb{R}^{m\times m}$ durch $B_{ij} = b(e_j, e_i)$, läßt sich dies schreiben in der Form

$$b(x, y) = \sum_{i,j=1}^m B_{ij}\xi_j \eta_i = \eta^T B\xi.$$

Umgekehrt wurde in Beispiel VIII-6.8 gezeigt, daß jede Matrix $B \in \mathbb{R}^{m\times m}$ eine Bilinearform $b \in \mathcal{L}_2(\mathbb{R}^m \times \mathbb{R}^m, \mathbb{R})$ und somit auch eine Abbildung $F \in \mathcal{L}(\mathbb{R}^m, (\mathbb{R}^m, \mathbb{R}))$ definiert. Es gilt dann

$$F(e_i)e_j = b(e_j, e_i).$$

Wir wenden nun diese Überlegungen auf die Charakterisierung der zweiten Ableitung einer Funktion $f : U \to Y$ an. Ist f differenzierbar, kann man $f' : U \to \mathcal{L}(X, Y)$ betrachten und die Differenzierbarkeit von f' untersuchen. Gemäß der Definition VIII-1.1 ist f' in $x_0 \in U$ differenzierbar, wenn es eine Abbildung $F \in \mathcal{L}(X, \mathcal{L}(X, Y))$ gibt mit der Eigenschaft

$$\lim_{h\to 0} \frac{1}{\|h\|} \|f'(x_0 + h) - f'(x_0) - Fh\|_{\mathcal{L}(X,Y)} = 0.$$

Dies ist gleichwertig mit

$$\lim_{h\to 0} \frac{1}{\|h\|} \sup_{k\neq 0} \frac{1}{\|k\|} \|f'(x_0 + h)k - f'(x_0)k - (Fh)k\|_Y = 0. \tag{17}$$

Identifiziert man F mit der bilinearen Abbildung $b \in \mathcal{L}_2(X \times X, Y)$, dann ist (17) gleichwertig mit

$$\lim_{h\to 0} \frac{1}{\|h\|} \sup_{k\neq 0} \frac{1}{\|k\|} \|f'(x_0 + h)k - f'(x_0)k - b(h, k)\|_Y = 0.$$

Anstelle von $b(h, k)$ schreibt man üblicherweise $f''(x_0)(h, k)$. Wir kehren nun wieder zum Spezialfall $X = \mathbb{R}^m$ und $Y = \mathbb{R}$ zurück, also $f : U \to \mathbb{R}$, $U \subset \mathbb{R}^m$. Es sei $f \in C^2(U)$. Dann ist f auf U stetig differenzierbar und es gilt

$$f'(x)k = \sum_{i=1}^m \partial_i f(x)k_i$$

Der Nachweis der Existenz der zweiten Ableitung von f in x_0 erfordert daher eine Matrix $\mathcal{H}(x_0) \in \mathbb{R}^{m\times m}$ mit

$$\lim_{h\to 0} \frac{1}{\|h\|} \sup_{k\neq 0} \frac{1}{\|k\|} |\sum_{i=1}^m \partial_i f(x_0 + h)k_i - \sum_{i=1}^m \partial_i f(x_0)k_i - k^T \mathcal{H}(x_0)h| = 0.$$

Wegen $f \in C^2(U,\mathbb{R})$ liegen die partiellen Ableitungen $\partial_i f$ in $C^1(U,\mathbb{R})$ und sind daher stetig differenzierbar. Für h so klein, daß $[x_0, x_0+h] \subset U$, ergibt der Mittelwertsatz

$$\partial_i f(x_0+h) - \partial_i f(x_0) = (\partial_i f)'(x_i^*(h))h = \sum_{j=1}^m \frac{\partial}{\partial \xi_j}\partial_i f((x_i^*(h))h_j$$

mit einer Zwischenstelle $x_i^*(h) \in [x_0, x_0+h]$. Dies hat

$$\sum_{i=1}^m \partial_i(f(x_0+h) - \partial_i f(x_0))k_i = \sum_{i=1}^m\sum_{j=1}^m \frac{\partial^2 f}{\partial \xi_j \partial \xi_i}(x_i^*(h))h_j k_i.$$

zur Folge. Setzt man

$$\mathcal{H}(x_0)_{ij} = \frac{\partial^2 f}{\partial \xi_j \partial \xi_i}(x_0)$$

ergibt sich für

$$\sum_{i=1}^m \partial_i(f(x_0+h) - \partial_i f(x_0))k_i - k^T\mathcal{H}(x_0)h = \sum_{i,j=1}^m (\frac{\partial^2 f}{\partial \xi_j \partial \xi_i}(x_i^*(h)) - \frac{\partial^2 f}{\partial \xi_j \partial \xi_i}(x_0))h_j k_i$$

und somit

$$\begin{aligned}|\sum_{i=1}^m \partial_i(f(x_0+h) - \partial_i f(x_0))k_i - k^T\mathcal{H}(x_0)h| \\ \le \sum_{i,j=1}^m |\frac{\partial^2 f}{\partial \xi_j \partial \xi_i}(x_i^*(h)) - \frac{\partial^2 f}{\partial \xi_j \partial \xi_i}(x_0)|\,\|k\|_\infty \|h\|_\infty.\end{aligned}$$

Daraus folgt

$$\begin{aligned}\frac{1}{\|h\|}\sup_{k\neq 0}\frac{1}{\|k\|}\|f'(x_0+h)k - f'(x_0)k - k^T\mathcal{H}(x_0)h\| \\ \le \sum_{i,j=1}^m |\frac{\partial^2 f}{\partial \xi_j \partial \xi_i}(x_i^*(h)) - \frac{\partial^2 f}{\partial \xi_j \partial \xi_i}(x_0)| \underset{h\to 0}{\to} 0.\end{aligned}$$

Der letzte Schluß ist durch die Stetigkeit von $\frac{\partial^2 f}{\partial \xi_j \partial \xi_i}$ in x_0 und $\lim_{h\to 0} x_i^*(h) = x_0$ gerechtfertigt. Die zweite Ableitung von $f\colon U \to \mathbb{R}$ in x_0 wird also durch die Hesse Matrix $\mathcal{H}(x_0)$ dargestellt.

Umgekehrt habe nun $f\colon U \to \mathbb{R}$ auf U eine stetige 2. Ableitung. Wir zeigen, daß dann die partiellen Ableitungen 2.Ordnung auf U existieren und stetig sind. Identifiziert man die 2. Ableitung mit der Bilinearform $b \in \mathcal{L}_2(\mathbb{R}^m \times \mathbb{R}^m, \mathbb{R})$ folgt für $x_0 \in U$ aus der Voraussetzung

$$\lim_{h\to 0}\frac{1}{\|h\|}\sup_{k\neq 0}\frac{1}{\|k\|}|f'(x_0+h)k - f'(x_0)k - b(h,k)| = 0$$

Wählt man $h = te_i$ und $k = e_j$ ergibt sich

$$0 = \lim_{t\to 0} \frac{1}{|t|} |(\partial_j f(x_0 + te_i) - \partial_j f(x_0)) - tb(e_i, e_j)|$$

$$= \lim_{t\to 0} |\frac{1}{t}(\partial_j f(x_0 + te_i) - \partial_j f(x_0)) - b(e_i, e_j)|$$

Somit existiert $\frac{\partial^2 f}{\partial\xi_i \partial\xi_j}(x_0)$ und es gilt $\frac{\partial^2 f}{\partial\xi_i \partial\xi_j}(x_0) = b(e_i, e_j)$. Die Stetigkeit der zweiten Ableitung in $x_0 \in U$ bedeutet, daß zu $\varepsilon > 0$ ein $\delta > 0$ gefunden werden kann, daß für alle $x \in U$ mit $\|x - x_0\| < \delta$

$$\|f''(x) - f''(x_0)\|_{\mathcal{L}_2(\mathbb{R}^m \times \mathbb{R}^m, \mathbb{R})} = \sup_{h\neq 0, k\neq 0} \frac{|(f''(x) - f''(x_0))(h, k)|}{\|h\| \, \|k\|} < \varepsilon$$

zutrifft.Wählt man $h = e_i$, $k = e_j$ und berücksichtigt $f''(x_0)(e_i, e_j) = b(e_i, e_j) = \frac{\partial^2 f}{\partial\xi_i \partial\xi_j}(x_0)$ kann man auf

$$|\frac{\partial^2 f}{\partial\xi_i \partial\xi_j}(x) - \frac{\partial^2 f}{\partial\xi_i \partial\xi_j}(x_0)| < \varepsilon$$

schließen falls $\|x - x_0\| < \delta$, d.h. die partiellen Ableitungen $\frac{\partial^2 f}{\partial\xi_i \partial\xi_j}$ sind stetig in x.

7. Der Satz von Taylor für Funktionen in mehreren Veränderlichen

Es sei $f \in C^{n+1}(u, \mathbb{R})$, $U \subset \mathbb{R}^n$ offen, $x_0 \in U$ und $h = (\sigma_1, \ldots, \sigma_m)$ so, daß die Verbindungsstrecke von x_0 und $x_0 + h$ in U liegt. Man kann dann den Satz von Taylor für Funktionen in mehreren Veränderlichen zurückführen auf den Satz von Taylor für Funktionen in einer Veränderlichen: dazu parametrisiert man die Gerade durch die Punkte x_0 und $x_0 + h$ durch $\varphi\colon (-\delta, 1 + \delta) \to \mathbb{R}^m$, $\varphi(t) = x_0 + th$ und setzt

$$F = f \circ \varphi\colon (-\delta, 1 + \delta) \to \mathbb{R}$$

für ein hinreichend kleines $\delta > 0$. Die Kettenregel sichert dann die Differenzierbarkeit von F und es gilt

$$F'(t) = f'(\varphi(t)) \circ \varphi'(t) = \sum_{i=1}^{m} (\partial_i f)(\varphi(t))\sigma_i.$$

Für $\frac{\partial f}{\partial x_i}$ wurde die abkürzende Schreibweise $\partial_i f$ verwendet. Setzt man $g_1 = \sum_{i=1}^m \sigma_i \partial_i f$, dann ist $g_1 \in C^n(U, \mathbb{R})$ und $F' = g_1 \circ \varphi$. Die Funktion g_1 besitzt somit zumindest stetige partielle Ableitungen 1. Ordnung und ist nach Satz VIII-6.1 differenzierbar. Somit existiert F'' und es gilt

$$F''(t) = g_1'(\varphi(t)) \circ \varphi'(t) = \sum_{i=1}^{m} (\partial_i g_1)(\varphi(t))\sigma_i$$

$$= \sum_{i=1}^{m} \sigma_i \partial_i (\sum_{j=1}^{m} \sigma_j \partial_j f)(\varphi(t)) = \sum_{i=1}^{m} \sum_{j=1}^{m} \sigma_i \sigma_j (\partial_i \partial_j f)(\varphi(t)) \equiv (g_2 \circ \varphi)(t),$$

mit

$$g_2 = \sum_{i=1}^{m} \sum_{j=1}^{m} \sigma_i \sigma_j (\partial_i \partial_j f).$$

Es folgt $g_2 \in C^{n-1}(U, \mathbb{R})$ und wie vorhin schließt man auf die Existenz von F'''. Um die höheren Ableitungen von F übersichtlich anschreiben zu können, betrachten wir die lineare Abbildung

$$D_h \colon \begin{cases} C^k(U, \mathbb{R}) \to C^{k-1}(U, \mathbb{R}), & k = 1, \dots, n+1 \\ D_h u = \sum_{i=1}^m \sigma_i \partial_i u \end{cases}$$

Dann kann man F' schreiben als

$$F' = (D_h f) \circ \varphi.$$

Wir zeigen induktiv

$$(*) \qquad F^{(k)} = (D_h^k f) \circ \varphi.$$

Der Induktionsschritt verläuft folgendermaßen: Es gelte $(*)$. Da in D_h^k nur partielle Ableitungen k-ter Ordnung auftreten, gilt $D_h^k f \in C^{n+1-k}(U, \mathbb{R})$. Insbesondere ist D_h^k nach Satz VIII-6.1 differenzierbar und somit existiert $F^{(k+1)}$. Die Kettenregel ergibt

$$\begin{aligned} F^{(k+1)}(t) &= \frac{d}{dt}((D_h^k f)(\varphi(t)) \\ &= (\sum_{i=1}^{m} \sigma_i \partial_i D_h^k f)(\varphi(t)) = (D_h (D_h^k f)(\varphi(t)) \\ &= (D_h^{k+1} f)(\varphi(t)). \end{aligned}$$

F ist demnach $n+1$-mal differenzierbar und es folgt mit dem Satz von Taylor VI-7.3

$$F(1) = \sum_{k=0}^{n} \frac{1}{k!} F^{(k)}(0) + \frac{1}{(n+1)!} F^{(n+1)}(\vartheta),$$

also

$$f(x_0 + h) = \sum_{k=0}^{n} \frac{1}{k!} (D_h^k f)(x_0) + \frac{1}{(n+1)!} (D_h^{n+1} f)(x_0 + \vartheta h),$$

für ein $\vartheta \in (0,1)$. In jedem der Terme

$$D_h^k f = (\sum_{i=1}^{m} \sigma_i \partial_i)^k f$$

kommen alle m^k partiellen Ableitungen der Ordnung k vor. Für $f \in C^{n+1}(U, \mathbb{R})$ sind nach dem Satz von Schwarz, Satz VIII-4.4, alle Ableitungen gleich, in welchen beispielsweise α_1-mal nach x_1, …, α_m-mal nach x_m abgeleitet wird. Mit Hilfe von

Multiindices kann man diese Terme zusammenfassen. Ein Multiindex der Dimension m ist ein m-Tupel $\alpha = (\alpha_1, \dots, \alpha_m) \in \mathbb{N}_0^m$. Folgende Notation ist üblich:

$$\begin{aligned}
|\alpha| &= \sum_{i=1}^{m} \alpha_i, \\
\alpha! &= \prod_{i=1}^{m} \alpha_i!, \\
h^\alpha &= \prod_{i=1}^{m} \sigma_i^{\alpha_i}, \qquad \text{für } h = (\sigma_1, \dots, \sigma_m) \\
\partial_\alpha &= \frac{\partial^{|\alpha|} f}{\partial x_1^{\alpha_1} \dots \partial x_m^{\alpha_m}}.
\end{aligned}$$

Beispiel VIII-7.1. Es sei $m = 3$ und $\alpha = (2, 0, 1)$. Dann ist $|\alpha| = 3$, $\alpha! = 2!0!1! = 2$, $h^\alpha = \sigma_1^2 \sigma_2^0 \sigma_3^1 = \sigma_1^2 \sigma_3$ und $\partial_\alpha f = \partial_{(2,0,1)} f = \frac{\partial^3}{\partial x_1^2 \partial x_3}$.

Lemma VIII-7.2. *Es gilt*

$$D_h^k f = (\sum_{i=1}^{m} \sigma_i \partial_i)^k f = \sum_{|\alpha|=k} \frac{k!}{\alpha!} h^\alpha \partial_\alpha f.$$

Beweis. Die einzelnen Summanden von $D_h^k f$ haben die Form

$$\sigma_{i_1} \cdots \sigma_{i_k} \frac{\partial^k}{\partial x_{i_1} \dots \partial x_{i_k}},$$

$1 \leq i_j \leq m$, $j = 1, \dots, k$. Wegen des Satzes von Schwarz VIII-4.4 ergeben einige dieser Summanden denselben Wert, beispielsweise $h^\alpha \partial_\alpha f$. Dabei ist $\alpha = (\alpha_1, \dots, \alpha_m)$ ein Multiindex mit $|\alpha| = k$. Da es $k!$ Anordnungen von k partiellen Ableitungen gibt, die Reihenfolge der partiellen Ableitungen $\frac{\partial^{\alpha_j}}{\partial x_{i_j}^{\alpha_j}}$ unerheblich ist, folgt, daß es $\frac{k!}{\alpha!}$ solcher Summanden gibt. □

Bisher wurde somit gezeigt:

Theorem VIII-7.3 (Satz von Taylor). *Es sei $f \in C^{n+1}(U, \mathbb{R})$, $U \subset \mathbb{R}^m$ offen, $x_0 \in U$ und $h = (\sigma_1, \dots, \sigma_m)$ so, daß die Verbindungstrecke von x_0 und $x_0 + h$ in U liegt. Dann gibt es ein $\vartheta \in (0, 1)$ so, daß*

$$\begin{aligned}
f(x_0 + h) &= \sum_{k=0}^{n} \frac{1}{k!} (D_h^k f)(x_0) + \frac{1}{(n+1)!} (D_h^{n+1} f)(x_0 + \vartheta h) \\
&= \sum_{|\alpha| \leq n} \frac{h^\alpha}{\alpha!} (\partial_\alpha f)(x_0) + \sum_{|\alpha|=n+1} \frac{h^\alpha}{\alpha!} (\partial_\alpha f)(x_0 + \vartheta h).
\end{aligned}$$

Beispiel VIII-7.4. Für $m = 2$ und $f \in C^3(U, \mathbb{R})$ erhalten wir

$$\begin{aligned} f(x_0 + h) &= f(x_0) + \sigma_1 f_x(x_0) + \sigma_2 f_y(x_0) \\ &\quad + \frac{1}{2!}(\sigma_1^2 f_{xx}(x_0) + 2\sigma_1\sigma_2 f_{xy}(x_0) + \sigma_2^2 f_{yy}(x_0)) + R_2(h) \\ R_2(h) &= \frac{1}{3!}(\sigma_1^3 f_{xxx}(x_0 + \vartheta h) + 3\sigma_1^2\sigma_2 f_{xxy}(x_0 + \vartheta h) \\ &\quad + 3\sigma_1\sigma_2^2 f_{xyy}(x_0 + \vartheta h) + \sigma_2^3 f_{yyy}(x_0 + \vartheta h)). \end{aligned}$$

Man überzeuge sich davon, daß die Verwendung von Multiindices auf dasselbe Resultat führt. Es treten folgende Multiindices auf

$$\begin{array}{lllll} k = 0: & (0,0) \\ k = 1: & (1,0), & (0,1) \\ k = 2: & (2,0), & (1,1), & (0,2) \\ k = 3: & (3,0), & (2,1), & (1,2), & (0,3). \end{array}$$

Bemerkung VIII-7.5. Es wurde eingangs gezeigt, daß für eine C^{n+1}-Funktion f die Funktionen $D_h^k f$, $k = 1, \ldots, n$, differenzierbar sind. Wir geben nun eine neue Interpretation von $D_h^k f$. Wegen

$$f'(x)h = \sum_{i=1}^{m} \sigma_i (\partial_i f)(x)$$

kann man

$$f'(x)h = (D_h f)(x)$$

schreiben. Differenziert man $x \to f'(x)h$ noch einmal (mit Inkrement h) erhält man

$$(f'(x)h)'h = (D_h^2 f)(x).$$

Anstelle von $(f'(x)h)'h$ schreibt man üblicherweise $f''(x)(h, h)$. Induktiv fortfahrend überzeugt man sich von der Beziehung

$$f^{(k)}(x)(h, \ldots, h) = (D_h^k f)(x).$$

Somit läßt sich der Satz von Taylor auch in der vertrauten Form anschreiben

$$f(x_0 + h) = \sum_{k=0}^{n} \frac{1}{k!} f^{(k)}(x_0)(\underbrace{h, \ldots, h}_{k \text{ fach}}) + \frac{1}{(n+1)!} f^{(n+1)}(x_0 + \vartheta h)(\underbrace{h, \ldots, h}_{(n+1) \text{ fach}}).$$

Definition VIII-7.6. *Es sei $f : U \to \mathbb{R}$, $U \subset \mathbb{R}^m$ offen, zweimal stetig differenzierbar. Die Matrix*

$$\mathcal{H}_f(x_0) = \begin{pmatrix} \frac{\partial^2 f}{\partial \xi_1^2}(x_0) & \ldots & \frac{\partial^2 f}{\partial \xi_m \partial \xi_1}(x_0) \\ \vdots & & \vdots \\ \frac{\partial^2 f}{\partial \xi_1 \partial \xi_m}(x_0) & \ldots & \frac{\partial^2 f}{\partial \xi_m^2}(x_0) \end{pmatrix} \in \mathbb{R}^{m \times m}$$

*heißt **Hesse Matrix** von f an der Stelle x_0.*

Für eine C^2-Funktion ist die Hesse-Matrix nach dem Satz von Schwarz symmetrisch. Mit ihrer Hilfe kann man den Term 2. Ordnung in der Taylorentwicklung darstellen als

$$\frac{1}{2}(D_h^2 f)(x_0) = \frac{1}{2}h^T \mathcal{H}_f(x_0)h.$$

Die Taylorentwicklung einer C^2-Funktion f kann man daher auch folgendermaßen anschreiben:

$$f(x_0+h) = f(x_0) + \langle \operatorname{grad} f(x_0), h\rangle + \frac{1}{2}h^T \mathcal{H}_f(x_0+\vartheta h)h.$$

Theorem VIII-7.7. *Es sei $f \in C^2(U)$, $U \subset \mathbb{R}^m$ offen, $x_0 \in U$ und $K(x_0,\varepsilon) \subset U$. Dann gibt es eine Abbildung $\rho\colon K(0,\varepsilon) \to \mathbb{R}$ so, daß für alle $h \in K(0,\varepsilon)$ die Darstellung gilt:*

$$f(x_0+h) = f(x_0) + f'(x_0)h + \frac{1}{2}h^T\, \mathcal{H}_f(x_0)h + \|h\|^2\rho(h)$$

mit $\lim_{h\to 0}\rho(h) = 0$.

Beweis. Aus dem Satz von Taylor folgt mit $h = (\sigma_1, \dots, \sigma_m)$

$$f(x_0+h) = f(x_0) + f'(x_0)h + \frac{1}{2}\sum_{i,k=1}^{m} \frac{\partial^2 f(x_0)}{\partial\xi_i\partial\xi_k}\sigma_i\sigma_k + r(h)$$

wobei

$$r(h) = \frac{1}{2}\sum_{i,k=1}^{m}\left(\frac{\partial^2 f}{\partial\xi_i\partial\xi_k}(x_0+\theta h) - \frac{\partial^2 f(x_0)}{\partial\xi_i\partial\xi_k}\right)\sigma_i\sigma_k.$$

Die Behauptung folgt aus der Abschätzung

$$|r(h)| \le \frac{1}{2}||h||_\infty^2 \sum_{i,k=1}^{m}\Big|\frac{\partial^2 f}{\partial\xi_i\partial\xi_k}(x_0+\theta h) - \frac{\partial^2 f}{\partial\xi_i\partial\xi_k}(x_0)\Big| \equiv ||h||_\infty^2\,\rho(h).$$

und der Stetigkeit der zweiten partiellen Ableitungen. □

8. Lokale Extrema für Funktionen in mehreren Veränderlichen

Wir haben bereits in Satz VIII-5.2 gesehen, daß eine notwendige Bedingung für das Auftreten eines lokalen Extremums einer differenzierbaren Funktion f an einer inneren Stelle x_0 des Definitionsbereiches $f'(x_0) = 0$ ist. Man nennt die Nullstellen der 1. Ableitung auch **kritische Stellen**. Ist $\operatorname{def} f \subset \mathbb{R}^m$, dann ist $x_0 \in \operatorname{def} f$ genau dann kritisch, wenn $\operatorname{grad} f(x_0) = 0$ ist, also x_0 das nichtlinearen Gleichungssystem

$$f_{x_i}(x) = 0, \quad i = 1, \dots, m$$

löst.

Wie im skalaren Fall läßt sich die Natur eines lokalen Extremums unter bestimmten Voraussetzungen am quadratischen Term in der Taylorformel ablesen. Dazu benötigen wir einige Eigenschaften quadratischer Formen. Wir erinnern: Es sei $A = (a_{ij}) \in \mathbb{R}^{m\times m}$, die Abbildung $Q_A\colon \mathbb{R}^m \to \mathbb{R}$,

$$Q_A(x) = x^T Ax = \sum_{i,j=1}^{m} a_{ij}\xi_i\xi_j$$

heißt **quadratische Form**. Offenbar gilt

$$(*) \qquad Q_A(\lambda x) = \lambda^2 Q_A(x)$$

für alle $\lambda \in \mathbb{R}$. Gilt für alle $x \neq 0$

$$Q_A(x) > 0 \quad (\text{bzw. } Q_A(x) < 0),$$

nennt man A **positiv definit** (bzw. **negativ definit**). Gilt für alle $x \in \mathbb{R}^m$

$$Q_A(x) \geq 0 \quad (\text{bzw. } Q_A(x) \leq 0),$$

heißt A **positiv semidefinit** (bzw. **negativ semidefinit**). Nimmt Q_A sowohl positive wie negative Werte an, nennt man A **indefinit**.
Ohne Beweis zitieren wir folgendes Definitheitskriterium.

THEOREM VIII-8.1 (Hurwitz-Kriterium). *$A \in \mathbb{R}^{m\times m}$ sei symmetrisch, d.h. $A = A^T$. Wir bezeichnen mit δ_k den k-ten Hauptminor von A, d.h.*

$$\delta_k = \det\begin{pmatrix} a_{11} & \dots & a_{1k} \\ \vdots & & \vdots \\ a_{k1} & \dots & a_{kk} \end{pmatrix}, \quad k = 1, \dots, m.$$

Dann gilt

a) *A ist positiv definit $\Leftrightarrow \delta_k > 0$, $k = 1, \dots, m$.*
b) *A ist negativ definit $\Leftrightarrow$ die Vorzeichen von δ_k alternieren in folgender Weise: $(-1)^k\delta_k > 0$, $k = 1, \dots, m$.*
c) *A ist indefinit, falls $\delta_k \neq 0$, $k = 1, \dots, m$, und weder a) noch b) zutrifft.*

KOROLLAR VIII-8.2. Es sei $A = \begin{pmatrix} a & b \\ b & c \end{pmatrix} = A^T$, $\det A = ac - b^2$. Dann gilt

a) A ist positiv definit $\Leftrightarrow a > 0, \det A > 0$.
b) A ist negativ definit $\Leftrightarrow a < 0, \det A > 0$.
c) A ist indefinit $\Leftrightarrow \det A < 0$.

LEMMA VIII-8.3. *Für $A \in \mathbb{R}^{m\times m}$ sind folgende Aussagen äquivalent:*

(1) *A ist positiv (negativ) definit.*
(2) *Es gibt ein $\alpha > 0$ derart, daß für alle $x \in \mathbb{R}^m$ gilt:*

$$Q_A(x) \geq \alpha\|x\|^2 \quad (Q_A(x) \leq -\alpha\|x\|^2)$$

BEWEIS. Wegen $Q_{-A}(x) = -Q_A(x)$ ist A genau dann negativ definit, wenn $-A$ positiv definit ist. Es genügt daher, positiv definite Matrizen zu betrachten.
(2) $\Rightarrow$ (1) trivial.

(1) $\Rightarrow$ (2) Es sei A positiv definit. $Q_A\colon \mathbb{R}^m \to \mathbb{R}$ ist stetig (Übung) und nimmt daher auf der kompakten Menge $S(0,1) := \{x \in \mathbb{R}^m\colon \|x\| = 1\}$ ein Minimum an, d.h. es gibt $x_0 \in S(0,1)$ so, daß für alle $x \in S(0,1)$

$$Q_A(x) \geq Q_A(x_0) = \alpha > 0$$

erfüllt ist (die strikte Ungleichung gilt, da A positiv definit ist). Für beliebige $x \neq 0$ bilden wir den Vektor $z_x = \frac{x}{\|x\|} \in S(0,1)$. Es folgt

$$\alpha \leq Q_A(z_x) = Q_A(\frac{x}{\|x\|}) = \frac{1}{\|x\|^2} Q_A(x)$$

d.h.

$$\alpha\|x\|^2 \leq Q_A(x), \quad x \neq 0.$$

Für $x = 0$ ist die Ungleichung trivialerweise erfüllt. □

Bemerkung VIII-8.4. Eine elegante Beweisvariante von Lemma VIII-8.3 für $A = A^T$ verwendet, daß

$$\sqrt{Q_A(x)} = \sqrt{\langle x, Ax\rangle}$$

eine Norm auf $\mathbb{R}^m$ darstellt. Die Behauptung folgt dann unmittelbar aus der Äquivalenz aller Normen in Räumen endlicher Dimension.

Theorem VIII-8.5 (Hinreichende Optimalitätsbedingung). *Es sei $f \in C^2(U)$, $U \subset \mathbb{R}^m$ offen, und $x_0 \in U$ ein kritischer Punkt von f, d.h. $f'(x_0) = 0$. Dann gilt:*

i) $\mathcal{H}_f(x_0)$ *positiv definit* $\Rightarrow$ f *besitzt in* x_0 *ein striktes lokales Minimum.*
ii) $\mathcal{H}_f(x_0)$ *negativ definit* $\Rightarrow$ f *besitzt in* x_0 *ein striktes lokales Maximum.*
iii) $\mathcal{H}_f(x_0)$ *indefinit* $\Rightarrow$ f *besitzt in* x_0 *kein lokales Extremum.*

Kritische Punkte von f in denen die Hesse-Matrix indefinit ist, nennt man auch **Sattelpunkte**.

Beweis. ad i) Wir bezeichnen mit Q die zur Matrix $\frac{1}{2}\mathcal{H}_f(x_0)$ gehörige quadratische Form. Da x_0 ein kritischer Punkt von f ist, folgt nach Satz VIII-7.7

$$f(x_0 + h) - f(x_0) = Q(h) + \|h\|^2\rho(h)$$

für alle $h \in K(0, \varepsilon)$, ε hinreichend klein, mit

$$\lim_{h\to 0} \rho(h) = 0. \tag{1}$$

Mit der Eigenschaft $(*)$ quadratischer Formen schließen wir für $h \neq 0$ auf

$$\frac{f(x_0 + h) - f(x_0)}{\|h\|^2} = Q(\frac{h}{\|h\|}) + \rho(h). \tag{2}$$

$\mathcal{H}_f(x_0)$ ist positiv definit, nach Lemma VIII-8.3 gibt es also ein $\alpha > 0$ mit

$$Q(x) \geq \alpha > 0 \quad \text{für alle } x \in S(0,1).$$

Wegen (1) gibt es ein $0 < \delta < \varepsilon$ so, daß für alle $h \in K(0,\delta)$

$$|\rho(h)| < \frac{\alpha}{2}$$

erfüllt ist. Für $h \in K(0,\delta)$ ergibt sich daher

$$\frac{f(x_0+h)-f(x_0)}{\|h\|^2} \geq \alpha + \rho(h) \geq \alpha - |\rho(h)| \geq \alpha - \frac{\alpha}{2} = \frac{\alpha}{2},$$

d.h.

$$f(x_0+h) - f(x_0) \geq \frac{\alpha}{2}\|h\|^2.$$

Insbesondere gilt also für alle $x \in K(x_0,\delta)$

$$f(x) \geq f(x_0),$$

f besitzt somit in x_0 ein striktes lokales Minimum.
ii) (Übung).
iii) Es sei $\mathcal{H}_f(x_0)$ indefinit. Es gibt also von 0 verschiedene Vektoren h_+ und h_- mit

$$Q(h_+) > 0 \quad \text{und} \quad Q(h_-) < 0.$$

Ersetzen wir in (2) h durch λh_+, bzw. λh_-, $\lambda \neq 0$, ergibt sich

$$\begin{aligned}\frac{f(x_0+\lambda h_\pm)-f(x_0)}{\lambda^2\|h_\pm\|^2} &= Q(\frac{\lambda h_\pm}{\|\lambda h_\pm\|}) + \rho(\lambda h_\pm)\\ &= \frac{\lambda^2}{\|\lambda h_\pm\|^2}Q(h_\pm) + \rho(\lambda h_\pm) = \frac{1}{\|h_\pm\|^2}Q(h_\pm) + \rho(\lambda h_\pm).\end{aligned}$$

Wählen wir nun $0 < \delta$ so klein, daß für alle $|\lambda| < \delta$

$$x_0 + \lambda h_\pm \in U,$$

und

$$|\rho(\lambda h_\pm)| \leq \frac{1}{2}\frac{1}{\|h_\pm\|^2}|Q(h_\pm)|,$$

erfüllt sind, dann erhalten wir für alle $0 < |\lambda| < \delta$

$$\begin{aligned}\frac{f(x_0+\lambda h_+)-f(x_0)}{\lambda^2\|h_+\|} &\geq \frac{1}{2}\frac{1}{\|h_+\|^2}Q(h_+) > 0,\\ \frac{f(x_0+\lambda h_-)-f(x_0)}{\lambda^2\|h_+\|} &\leq -\frac{1}{2}\frac{1}{\|h_-\|^2}|Q(h_-)| < 0,\end{aligned}$$

also

$$f(x_0+\lambda h_+) \geq f(x_0) \quad \text{und} \quad f(x_0+\lambda h_-) \leq f(x_0).$$

f kann also in x_0 kein lokales Extremum besitzen. □

Im speziellen Fall von Funktionen in zwei Veränderlichen ist die Anwendung der hinreichenden Bedingungen aus Satz VIII-8.5 besonders einfach.

Theorem VIII-8.6. *Es sei $f \in C^2(D)$, $D \subset \mathbb{R}^2$ offen, und $x_0 \in D$ sei ein kritischer Punkt von f. Ferner sei*

$$\delta = \det \mathcal{H}_f(x_0) = f_{\xi\xi}(x_0)f_{\eta\eta}(x_0) - f_{\xi\eta}(x_0)^2.$$

Dann gilt

$$\begin{array}{rccl} i) & f_{\xi\xi}(x_0) > 0, \delta > 0 & \Rightarrow & f \textit{ besitzt in } x_0 \textit{ ein striktes lokales Minimum,} \\ ii) & f_{\xi\xi}(x_0) < 0, \delta > 0 & \Rightarrow & f \textit{ besitzt in } x_0 \textit{ ein striktes lokales Maximum,} \\ iii) & \delta < 0 & \Rightarrow & f \textit{ besitzt in } x_0 \textit{ einen Sattelpunkt.} \end{array}$$

Beispiel VIII-8.7. $f\colon \mathbb{R}^2 \to \mathbb{R}$, $f(\xi,\eta) = \xi^3 - 6\xi\eta + 3\eta^2 - 24\xi + 4$. Es ist $f'(\xi,\eta) = (3\xi^2 - 6\eta - 24, \quad -6\xi + 6\eta)$. Die Gleichung $f'(\xi,\eta) = (0 \quad 0)$ ergibt die kritischen Punkte $(4,4)$, $(-2,-2)$. Die Hesse Matrix ist gegeben durch

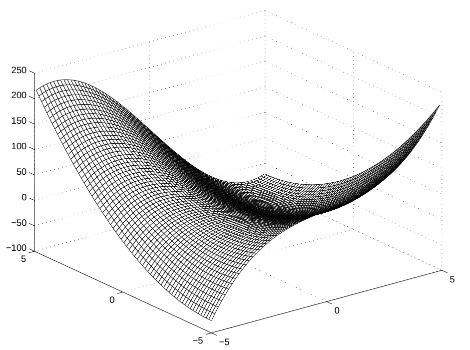
$$\mathcal{H}_f(\xi,\eta) = \begin{pmatrix} f_{\xi\xi} & f_{\xi\eta} \\ f_{\xi\eta} & f_{\eta\eta} \end{pmatrix} = \begin{pmatrix} 6\xi & -6 \\ -6 & 6 \end{pmatrix},$$

also

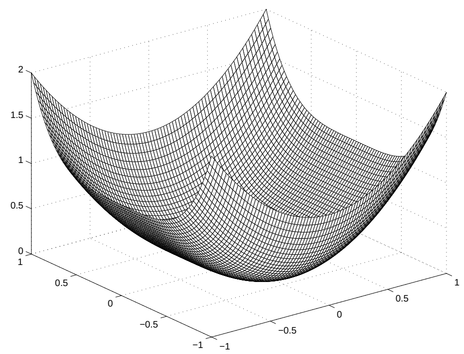
$$\mathcal{H}_f(4,4) = \begin{pmatrix} 24 & -6 \\ -6 & 6 \end{pmatrix}, \quad \mathcal{H}_f(-2,-2) = \begin{pmatrix} -12 & -6 \\ -6 & 6 \end{pmatrix}.$$

Satz VIII-8.6 zeigt, daß f in $(4,4)$ ein striktes lokales Minimum und in $(-2,-2)$ einen Sattelpunkt besitzt.

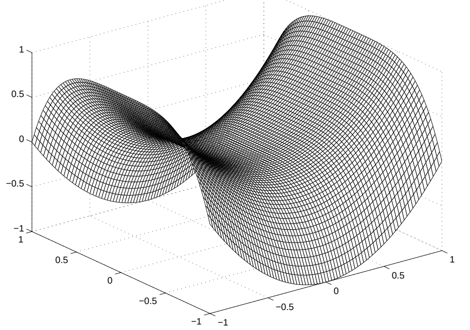
Leider werden mit den Sätzen VIII-8.5, VIII-8.6 nicht alle Möglichkeiten erfaßt. Dies liegt daran, daß nicht jede quadratische Form entweder positiv definit oder negativ definit oder indefinit ist. Zur Illustration betrachten wir folgendes Beispiel: Die Abbildungen



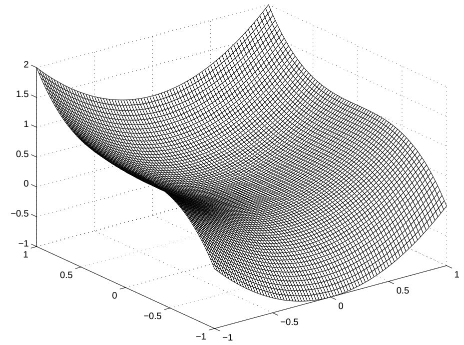
$f(\xi,\eta) = \xi^3 - 6\xi\eta + 3\eta^2 - 24\xi + 4$



$f(\xi,\eta) = \xi^2 + \eta^4$



$g(\xi,\eta) = \xi^2 - \eta^4$



$h(\xi,\eta) = \xi^2 + \eta^3$

$$\begin{aligned} f(\xi,\eta) &= \xi^2 + \eta^4, \\ g(\xi,\eta) &= \xi^2 - \eta^4, \\ h(\xi,\eta) &= \xi^2 + \eta^3, \end{aligned}$$

haben alle in $(0,0)$ einen kritischen Punkt. In jedem Falle ist die Hesse-Matrix in $(0,0)$ durch

$$\mathcal{H}(0,0) = \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix}$$

gegeben. $\mathcal{H}(0,0)$ ist also positiv semidefinit. In $(0,0)$ besitzt jedoch f ein (globales) Minimum, g einen Sattelpunkt, während h an der Stelle weder einen Sattelpunkt noch ein lokales Extremum aufweist. Bei der Untersuchung derartiger kritischer Stellen, an denen die Sätze VIII-8.5, VIII-8.6 nicht anwendbar sind, ist man auf ad hoc Methoden angewiesen.

9. Lokale Umkehrbarkeit, implizite Funktionen

9.1. Fixpunktsatz von Banach. Wir wenden uns nun einer Satzgruppe zu, die die Lösbarkeit von nichtlinearen Gleichungen $f(x) = y$ zum Inhalt hat. Ein wesentliches Beweismittel ist folgendes Fixpunktprinzip.

Theorem VIII-9.1 (Fixpunktsatz von Banach). *Es sei X ein Banach Raum, $U \subset X$, $f\colon U \to X$ eine Abbildung und M eine abgeschlossene Teilmenge von U mit folgenden Eigenschaften:*
i) $\exists \ell \in [0,1) \forall x,y \in M\colon \|f(x) - f(y)\| \le \ell \|x - y\|$,
*d.h. f ist eine **Kontraktion** auf M.*
ii) $f(M) \subset M$*, d.h. M ist **invariant** unter f.*
Dann gibt es in M genau ein Element x^* mit $f(x^*) = x^*$.

Elemente x mit der Eigenschaft $f(x) = x$ nennt man **Fixpunkte** von f. Satz VIII-9.1 sichert also unter bestimmten Vorausssetzungen die Existenz eines eindeutig bestimmten Fixpunktes von f.

Beweis. Wir wählen ein *beliebiges* Element x_0 in M und definieren rekursiv die Folge (x_n) durch

$$x_{n+1} = f(x_n), \quad n \in \mathbb{N}_0.$$

Wegen der Invarianz von M folgt aus $x_0 \in M$, daß auch $x_1 = f(x_0)$ in M liegt. Eine einfache Induktion zeigt dann $(x_n)_{n\ge 1} \subset M$.
Wir zeigen nun, daß (x_n) eine Cauchy Folge ist: dazu verifiziere man vorerst die Abschätzung (Übung):

$$\|f(x_n) - f(x_{n-1})\| \le \ell^n \|f(x_0) - x_0\|, \quad n \in \mathbb{N}.$$

Daraus folgt mit Hilfe eines "Teleskoparguments" für alle $n \in \mathbb{N}$ und $p \in \mathbb{N}$ die Ungleichung

$$(*) \qquad \|f(x_{n+p}) - f(x_n)\| \le \sum_{i=0}^{p-1} \|f(x_{n+i+1}) - f(x_{n+i})\| \le \sum_{i=0}^{p-1} \ell^{n+i+1} \|f(x_0) - x_0\| < \ell^{n+1} \frac{\|f(x_0) - x_0\|}{1-\ell}.$$

Wegen $\ell < 1$ ist (x_n) eine Cauchy Folge. X ist vollständig, also ist (x_n) konvergent. Somit gibt es ein Element $x^* \in X$ mit

$$\lim_{n\to\infty} x_n = x^*.$$

Aus $(x_n) \subset M$ und der Abgeschlossenheit von M folgt mit Satz III-1.7 $x^* \in M$. Die Stetigkeit von f ergibt

$$x^* = \lim_{n\to\infty} x_{n+1} = \lim_{n\to\infty} f(x_n) = f(\lim_{n\to\infty} x_n) = f(x^*),$$

x^* ist also ein Fixpunkt von f.
Die Eindeutigkeit des Fixpunktes ist eine Folge der Kontraktionseigenschaft von f. Angenommen es gäbe einen weiteren Fixpunkt y^* in M, dann folgt

$$\|x^* - y^*\| = \|f(x^*) - f(y^*)\| \le \ell \|x^* - y^*\|,$$

also

$$(1-\ell)\|x^* - y^*\| \le 0.$$

Dies kann nur für $x^* = y^*$ gelten. □

Bemerkung VIII-9.2. 1) Eine Schwierigkeit bei der Anwendung des Banachschen Fixpunktsatzes besteht oft darin, eine geeignete invariante Teilmenge M zu finden, auf der man die Kontraktionseigenschaft von f nachweisen kann. Die Lipschitzkonstante ℓ kann dabei nicht durch 1 ersetzt werden. Es genügt also nicht, die etwas schwächere Ungleichung

$$\forall x, y \in M\colon \|f(x) - f(y)\| \le \|x - y\|$$

nachzuweisen.
2) Der Beweis von Satz VIII-9.1 ist konstruktiv: der Fixpunkt von f kann berechnet werden, indem man für einen beliebigen Startwert $x_0 \in M$ die Folge der iterierten Bilder $(f^n(x_0)) = (x_n)$ berechnet. Man erhält auch eine Abschätzung des Fehlers, der entsteht, wenn man die Iteration nach dem n-ten Schritt abbricht. Führt man nämlich in $(*)$ den Grenzübergang $p \to \infty$ durch (und ersetzt n durch $n-1$), ergibt sich

$$\|x^* - x_n\| = \lim_{p\to\infty} \|x_{n+p} - x_n\| = \lim_{p\to\infty} \|f(x_{n-1+p}) - f(x_{n-1})\| \le \frac{\ell^n}{1-\ell} \|f(x_0) - x_0\|.$$

Beispiel VIII-9.3. Wir demonstrieren die Anwendung des Fixpunktsatzes auf die Abbildung $f(x) = \frac{1}{2} + x - \frac{x^2}{6}$, $x \geq 0$. Offensichtlich ist $x^* = \sqrt{3}$ der gesuchte Fixpunkt. Wir wählen $M = [\frac{3}{2}, 2]$. Wegen $0 < \frac{1}{3} \leq f'(x) \leq \frac{1}{2}$, $x \in M$, ist f auf M streng monoton steigend. Aus $f(\frac{3}{2}) = \frac{13}{8} > \frac{3}{2}$ und $f(2) = \frac{11}{6} < 2$ folgt somit die Invarianz von M. Als Lipschitzkonstante kann man $L = \frac{1}{2}$ wählen. Nach dem Fixpunktsatz von Banach existiert in M genau ein Fixpunkt, welcher mit Hilfe der Iteration $x_{n+1} = f(x_n)$, $x_o \in M$, berechnet werden kann. Wie groß muß n gewählt werden, damit $|x^* - x_n| < 10^{-5}$ gewährleistet ist? Mit Hilfe der vorausgehenden Bemerkung findet man für $x_0 = \frac{3}{2}$

$$|x^* - x_n| \leq (\frac{1}{2})^n \frac{\frac{1}{2} - \frac{x_0^2}{6}}{1 - \frac{1}{2}} = (\frac{1}{2})^{n+2}.$$

Eine einfache Rechnung zeigt die Gültigkeit der Ungleichung $(\frac{1}{2})^{n+2} < 10^{-5}$ für $n \geq 15$. Führt man die Rechnung durch, findet man $x_{15} = 1.73205$ und $|x_{15} - \sqrt{3}| \leq 7 \cdot 10^{-7}$.

9.2. Lokale Invertierbarkeit. Im Folgenden greifen wir auf den Begriff der Norm einer Matrix $M \in \mathbb{K}^{m\times m}$ zurück: verwenden wir in $\mathbb{K}^m$ die Norm $||\cdot||_p$, dann ist die zugehörige Matrixnorm festgelegt durch $||M||_p = \inf\{c \geq 0\colon ||Mx||_p \leq c||x||_p, x \in \mathbb{R}^m\} = \sup\{||Mx||_p\colon ||x||_p \leq 1\}$ Verwenden wir im $\mathbb{K}^m$ beispielsweise die Norm $||x||_\infty = \max\{|\xi_i|\colon 1 \leq i \leq m\}$, $x = (\xi_1, \ldots, \xi_i)$, dann ist ergibt sich für die zugehörige Matrix Norm

$$||M||_\infty = \max_{1\leq i\leq m} \sum_{j=1}^m |m_{ij}|.$$

Theorem VIII-9.4. *Es bezeichne $||\cdot||$ eine beliebige Matrixnorm.*
1.) Gilt $||I - A|| < 1$, dann ist A regulär.
2.) Ist $A \in \mathbb{K}^{m\times m}$ regulär und $||B - A|| < \frac{1}{||A^{-1}||}$, dann ist B regulär. Für $||B - A|| < \frac{1}{2||A^{-1}||}$ gilt dann $||B^{-1}|| < 2||A^{-1}||$. Somit ist die Gruppe der regulären Matrizen $GL_m(\mathbb{K})$ offen in $\mathbb{K}^{m\times m}$.
3.) Die Abbildung $inv\colon GL_m(\mathbb{K}) \to \mathcal{L}(\mathbb{K}^m, \mathbb{K}^m)$, $inv(A) = A^{-1}$ ist stetig differenzierbar mit

$$inv'(A)\delta A = -A^{-1}\,\delta A\, A^{-1}, \quad \delta A \in \mathbb{K}^{m\times m}.$$

Beweis. ad 1.) Es genügt die Injektivität von A nachzuweisen. Die Annahme $Ax = 0$ für ein $x \neq 0$ führt dann auf

$$||x|| \leq ||x - Ax|| + ||Ax|| \leq ||I - A||\,||x||,$$

also auf den Widerspruch $||I - A|| \geq 1$.
ad 2.) Wir betrachten

$$||I - A^{-1}B|| = ||A^{-1}(A - B)|| \leq ||A^{-1}||\,||A - B|| < 1.$$

Also ist $A^{-1}B$ regulär und daher auch $B = A(A^{-1}B)$. Aus $x = A^{-1}y$ folgern wir

$$\|x\| \leq \|A^{-1}\|\,\|y\|$$

und somit

$$\|Ax\| \leq \frac{1}{\|A^{-1}\|}\|x\|.$$

Für jede reguläre Matrix $||B||$ folgt aus der Ungleichung

$$||Bx|| \geq \alpha||x||, \qquad x \in \mathbb{K}^m$$

für ein $\alpha > 0$ die Abschätzung

$$||B^{-1}|| \leq \frac{1}{\alpha}.$$

Für $||B - A|| < \frac{1}{2||A^{-1}||}$ ergeben sich somit folgende Abschätzungen

$$\begin{aligned} ||Bx|| &\geq ||(B - A)x + Ax|| \geq ||Ax|| - ||(B - A)x|| \\ &\geq \frac{1}{||A^{-1}||}||x|| - ||B - A||\,||x|| \geq \frac{1}{2||A^{-1}||}||x||. \end{aligned}$$

ad 3.) Wir zeigen *inv* ist stetig in $A \in GL_m(\mathbb{K})$. Es sei $\varepsilon > 0$ beliebig gewählt und $0 < \delta < \min\{\frac{1}{2||A^{-1}||}, \frac{\varepsilon}{2||A^{-1}||^2}\}$. Falls $||B - A|| < \delta$ ist B regulär und es folgt

$$\begin{aligned} ||inv(A) - inv(B)|| &= ||A^{-1} - B^{-1}|| = ||B^{-1}(A - B)A^{-1}|| \\ &\leq ||A^{-1}||\,||B^{-1}||\,||A - B|| \leq 2||A^{-1}||^2\delta < \varepsilon. \end{aligned}$$

Die Differenzierbarkeit von *inv* sieht man folgendermaßen ein: Es sei $A \in GL_m(\mathbb{K})$ und $||\delta A|| < \frac{1}{2||A^{-1}||}$. Dann ist $(A + \delta A) \in GL_m(\mathbb{K})$ und

$$\begin{aligned} inv(A + \delta A) - inv(A) &= (A + \delta A)^{-1} - A^{-1} = -A^{-1}\delta A(A + \delta A)^{-1} \\ &= -A^{-1}\delta A\, A^{-1} + r(\delta A) \end{aligned}$$

mit

$$r(\delta A) = -A^{-1}\delta A((A + \delta A)^{-1} - A^{-1})$$

Aus der Abschätzung

$$||r(\delta A)|| \leq ||A^{-1}||\,||\delta A||\,||(A + \delta A)^{-1} - A^{-1}||$$

zusammen mit der Stetigkeit von *inv* folgt

$$\lim_{\delta A \to 0} \frac{r(\delta A)}{||\delta A||} = 0.$$

Wegen der Linearität und Stetigkeit von $\delta A \to -A^{-1}\delta A\, A^{-1}$ ist *inv* differenzierbar und es gilt

$$inv'(A)(\delta A) = -A^{-1}\delta A\, A^{-1}.$$

Die Stetigkeit der Ableitung $inv'\colon GL_m(\mathbb{K}) \to \mathcal{L}(\mathbb{K}^{m\times m}, \mathbb{K}^{m\times m})$ betrachten wir für A, $B \in GL_m(\mathbb{K})$ und $\delta A \in \mathbb{K}^{m\times m}$

$$\begin{aligned}||inv'(A)\delta A - inv'(B)\delta A|| &= ||-A^{-1}\delta A\, A^{-1} + B^{-1}\delta A\, B^{-1}|| \\ &\leq ||-(A^{-1}-B^{-1})\delta A\, A^{-1}|| + ||B^{-1}\delta A\,(A^{-1}-B^{-1})|| \\ &\leq ||A^{-1}-B^{-1}||(||A^{-1}|| + ||B^{-1}||)||\delta A||\end{aligned}$$

also

$$||inv'(A) - inv'(B)|| \leq (||A^{-1}|| + ||B^{-1}||)||inv(A) - inv(B)||.$$

Die Stetigkeit von inv' folgt nun aus der Stetigkeit von inv. □

Theorem VIII-9.5. *Es sei $f\colon U \to \mathbb{R}^m$, $U \subset \mathbb{R}^m$ offen, stetig differenzierbar. An der Stelle $x_0 \in U$ sei $f'(x_0)\colon \mathbb{R}^m \to \mathbb{R}^m$ invertierbar. Dann gibt es eine offene Umgebung $O \in \mathcal{U}(x_0)$ mit folgenden Eigenschaften:*

i) *$f|_O$ ist injektiv,*
ii) *$f(O)$ ist offen,*
iii) *die Umkehrabbildung g von $f|_O$ ist stetig differenzierbar und es gilt für alle $y \in f(O)$*

$$g'(y) = (f'(g(y)))^{-1}.$$

Beweis. Wir unterteilen den Beweis in mehrere Schritte.
1) Vereinfachung des Problems: Indem wir an Stelle von f die Abbildung $x \mapsto f(x_0 + x) - f(x_0)$ betrachten, können wir o.B.d.A. von $x_0 = 0$ und $f(0) = 0$ ausgehen. Nach Voraussetzung ist also $f'(0)$ regulär. Ersetzen wir f durch $h = f'(0)^{-1} \circ f$, folgt aus der Kettenregel

$$h'(x) = f'(0)^{-1} \circ f'(x), \text{ also } h'(0) = id_{\mathbb{R}^m}.$$

Wir setzen deshalb o.B.d.A. von nun an voraus, daß

$$f(0) = 0 \text{ und } f'(0) = id_{\mathbb{R}^m}$$

gilt.
2) Fixpunktargument: Wir zeigen, daß jedes y aus einer geeigneten Umgebung von 0 genau ein Urbild besitzt, indem wir Fixpunkte der Abbildung $\phi_y\colon U \to \mathbb{R}^m$

$$\phi_y(x) = x - f(x) + y$$

suchen. Es gilt nämlich

$$\phi_y(x) = x \Leftrightarrow y = f(x).$$

Wir greifen dabei auf den Fixpunktsatz von Banach VIII-9.1 zurück.
3) Invarianz von $\bar{K}(0, 2r)$ für geeignete $r > 0$: Wegen $f'(0) = id_{\mathbb{R}^m}$ und der Stetigkeit von f' an der Stelle $x_0 = 0$ gibt es ein $r > 0$, sodaß

$$(*) \qquad \|x\| \leq 2r \Rightarrow \|id_{\mathbb{R}^m} - f'(x)\| \leq \frac{1}{2}$$

gilt. Setzen wir $h(x) = x - f(x)$, folgt $h(0) = 0$ und $h'(x) = id_{\mathbb{R}^m} - f'(x)$. Somit gilt $||h'(x)|| \leq \frac{1}{2}$ für $||x|| \leq 2r$. Mit Hilfe des Schrankensatzes VIII-5.6 ergibt sich daraus für $x \in \bar{K}(0, 2r)$

$$||x - f(x)|| = ||h(x) - h(0)|| = \sup_{z \in \bar{K}(0,2r)} ||h'(z)|| \, ||x|| \leq \frac{1}{2}||x|| \leq r.$$

Dies ergibt für $||y|| < r$ und $||x|| \leq 2r$

$$||\phi_y(x)|| \leq ||x - f(x)|| + ||y|| < r + r = 2r,$$

d.h. $\phi_y(\bar{K}(0, 2r)) \subset K(0, 2r)$ für alle $y \in K(0, r)$.
4) ϕ_y ist für jedes $y \in K(0, r)$ eine Kontraktion auf $\bar{K}(0, 2r)$: Dies ist wiederum eine unmittelbare Folge des Schrankensatzes.

$$(\dagger) \qquad ||\phi_y(x) - \phi_y(\tilde{x})|| = ||h(x) - h(\tilde{x})|| \leq \sup_{z \in \bar{K}(0,2r)} ||h'(z)|| \, ||x - \tilde{x}|| \leq \frac{1}{2}||x - \tilde{x}||.$$

5) Beweis von i) und ii): Da $\bar{K}(0, 2r)$ abgeschlossen ist, garantiert der Fixpunktsatz von Banach für alle $y \in K(0, r)$ die Existenz eines *eindeutigen* Fixpunktes von ϕ_y, der sogar in $K(0, 2r)$ liegt.
Mit anderen Worten: Jedes $y \in K(0, r)$ besitzt ein eindeutiges Urbild in $K(0, 2r)$. Setzen wir $O = f^{-1}(K(0, r)) \cap K(0, 2r)$, dann ist $f|_O \colon O \to K(0, r)$ ist eine Bijektion. Die die Menge O ist offen: wegen der Stetigkeit von f ist $f^{-1}(K(0, r))$ relativ offen in U und da U selber offen ist, folgt daß $f^{-1}(K(0, r))$ und daher auch $f^{-1}(K(0, r)) \cap K(0, 2r)$ offen ist.
6) Differenzierbarkeit von $g := (f|_O)^{-1}$: In Hinsicht auf Satz VIII-2.5 ist die Stetigkeit von g nachzuweisen. Für x und $\tilde{x} \in O$ gilt

$$f(x) - f(\tilde{x}) = x - \tilde{x} - (h(x) - h(\tilde{x})).$$

Aus $(\dagger)$ folgt daher

$$\|f(x) - f(\tilde{x})\| \geq \|x - \tilde{x}\| - \|h(x) - h(\tilde{x})\| \geq \frac{1}{2}\|x - \tilde{x}\|.$$

Schreibt man $y = f(x)$, $\tilde{y} = f(\tilde{x})$ und $x = g(y)$, $\tilde{x} = g(\tilde{y})$, dann gilt $y, \tilde{y} \in K(0, r)$ und

$$\|g(y) - g(\tilde{y})\| \leq 2\|y - \tilde{y}\|,$$

d.h. g ist Lipschitz stetig auf $K(0, r)$. Für $y \in K(0, r)$ liegt $g(y)$ in $K(0, 2r)$ und somit ist wegen $(*)$ und Satz VIII-9.4 $f'(g(y))$ invertierbar. Aus Satz VIII-2.5 folgt nun die Differenzierbarkeit von g und

$$g'(y) = (f'(g(y)))^{-1}, \quad y \in K(0, r).$$

7) Stetige Differenzierbarkeit von g: Dies ergibt sich aus der Beobachtung, daß g' sich folgendermaßen als Verkettung stetiger Funktionen darstellen läßt

$$g' = inv \circ f' \circ g.$$

$\square$

KOROLLAR VIII-9.6. Es sei $f\colon U \to \mathbb{R}^m$, $U \subset \mathbb{R}^m$ offen, stetig differenzierbar. Ist $f'(x)$ für alle $x \in U$ regulär, dann ist $f(U)$ offen.

BEWEIS. Jeder Punkt $x \in U$ besitzt nach dem Umkehrsatz eine Umgebung O_x, so daß $f(O_x)$ offen ist. Dann ist auch $f(U) = \cup_{x\in U} f(O_x)$ offen. $\square$

DEFINITION VIII-9.7. *Es seien U und V offene Teilmengen von $\mathbb{R}^m$. Eine Abbildung $f\colon U \to V$ heißt C^k-**Diffeomorphismus**, wenn f bijektiv ist und sowohl f, als auch f^{-1} von der Klasse C^k für ein $k \in \mathbb{N}$ sind.*

KOROLLAR VIII-9.8. Es sei $f\colon U \to \mathbb{R}^m$, $U \subset \mathbb{R}^m$ offen, stetig differenzierbar. Ist $f'(x)$ für alle $x \in U$ regulär und ist f injektiv, dann ist f ein C^1-Diffeomorphismus von U auf die offene Menge $f(U)$.

BEWEIS. $f(U)$ ist wegen Korollar VIII-9.6 offen. Nach Voraussetzung existiert die Umkehrabbildung $f^{-1}\colon f(U) \to U$. Nach dem Umkehrsatz existiert für jedes $x \in U$ eine offene Umgebung O_x, so daß die Umkehrabbildung g der Einschränkung von f auf O_x stetig differenzierbar ist. Da f^{-1} auf $f(O_x)$ mit g übereinstimmt folgt die Behauptung. $\square$

BEMERKUNG VIII-9.9. Die Voraussetzung der Regularität von $f'(x)$ für *alle* $x \in U$ sichert nur die *lokale* Injektivität von f. Folgendes Beispiel, zeigt daß globale Injektivität nicht gefolgert werden kann. Es sei $f\colon \mathbb{R}^2 \to \mathbb{R}^2$

$$f(\xi,\eta) = \begin{pmatrix} e^\xi \cos\eta \\ e^\xi \sin\eta \end{pmatrix}.$$

Dann ist $\det(f'(\xi,\eta)) = e^{2\xi} > 0$ für alle $(\xi,\eta) \in \mathbb{R}^2$. somit ist f an jeder Stelle (ξ,η) auf einer geeigneten Umgebung injektiv, aber nicht injektiv auf $\mathbb{R}^2$.

9.3. Implizite Funktionen. Wir wenden uns nun der Frage zu, ob und unter welchen Umständen z.B. eine Gleichung in zwei Unbekannten

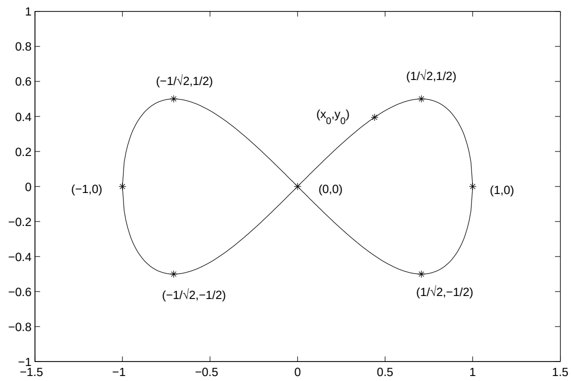
$$f(x,y) = 0 \tag{$*$}$$

nach y oder nach x aufgelöst werden kann. Also, ob es Funktionen $\varphi\colon I \to \mathbb{R}$ oder $\psi\colon J \to \mathbb{R}$ gibt mit der Eigenschaft

$$\begin{aligned} \forall x \in I\colon y = \varphi(x) &\Leftrightarrow f(x,\varphi(x)) = 0, \\ \forall y \in J\colon x = \psi(y) &\Leftrightarrow f(\psi(y),y) = 0. \end{aligned}$$

Man sagt auch, die Funktionen φ oder ψ seien **implizit definiert** durch die Gleichung $(*)$ und die Gleichung $(*)$ wird durch φ nach y, bzw. durch ψ nach x aufgelöst.
Zur Illustration der auftretenden Probleme betrachten wir das Beispiel $f(x,y) = x^2(1-x^2) - y^2$. Die Gleichung $f(x,y) = 0$ besitzt offenbar die Lösungen

$$\begin{aligned} y &= \pm\sqrt{x^2(1-x^2)}, \quad |x| \le 1, \\ x &= \pm\sqrt{\frac{1}{2} \pm \sqrt{\frac{1}{4} - y^2}}, \quad |y| \le \frac{1}{2}. \end{aligned}$$



Die Menge der Nullstellen von f wird also durch eine Lemniskate beschrieben. Ohne Zusatzforderung an φ oder ψ gibt es offenbar unendlich viele Möglichkeiten, Lösungen von $(*)$ zu einer Funktion zusammenzufügen. Wir fordern daher, daß φ oder ψ zumindest stetig sind. In einer hinreichend kleinen Umgebung einer Lösung (x_0, y_0) gibt es für alle x genau ein $y = \varphi(x)$ mit $f(x, \varphi(x)) = 0$ und für alle y genau ein $x = \psi(y)$ mit $f(\psi(y), y) = 0$, die Gleichung $(*)$ ist also nach x und nach y auflösbar. Wir bemerken: $f_x(x_0, y_0) \neq 0$ und $f_y(x_0, y_0) \neq 0$. Anders ist die Situation in $(0, 0)$: in jeder (noch so kleinen) Umgebung hat die Gleichung $(*)$ für jedes $x \neq 0$, aber auch für jedes $y \neq 0$ mehrere Lösungen und kann daher weder nach y noch nach x aufgelöst werden. Eine einfache Rechnung zeigt $f'(0, 0) = (0, 0)$. In jeder Umgebung der Lösungen $(\pm 1, 0)$ kann $(*)$ nicht nach y aufgelöst werden, da es zu $x \neq \pm 1$ jeweils zwei unterschiedliche Lösungen für y gibt. Es gibt aber eine Auflösung nach x, z.B.

$$\psi(y) = \sqrt{\frac{1}{2} + \sqrt{\frac{1}{4} - y^2}}, \quad |y| \leq \frac{1}{2}$$

in $(1, 0)$, bzw. durch $-\psi$ in $(-1, 0)$. An diesen Stellen gilt $f_x(\pm 1, 0) \neq 0$ und $f_y(\pm 1, 0) = 0$. In einer hinreichend kleinen Umgebung der Stellen $(\pm\frac{1}{\sqrt{2}}, \pm\frac{1}{2})$ gibt es nur eine Auflösung nach y. Man beachte $f_x(\pm\frac{1}{\sqrt{2}}, \pm\frac{1}{2}) = 0$ und $f_y(\pm\frac{1}{\sqrt{2}}, \pm\frac{1}{2}) \neq 0$.
Dies legt die Vermutung nahe, daß eine Gleichung $f(x, y) = 0$ in der Nähe einer Nullstelle (x_0, y_0) eine Auflösung $y = \varphi(x)$ bzw. $x = \psi(y)$ besitzt, wenn $f_y(x_0, y_0) \neq 0$ bzw. $f_x(x_0, y_0) \neq 0$ ist.
Allgemeiner betrachten wir die Lösbarkeit eines Systems von n nichtlinearen Gleichungen in m Unbekannten $m > n$. Der Umkehrsatz VIII-9.5 behandelt den Fall $n = m$. Lösen des nichtlinearen Gleichungssystems bedeutet hier, daß wir n der Variablen, wir bezeichnen sie mit $\xi_1, \ldots, \xi_n$, in eindeutiger Weise durch die übrigen Variablen, wir nennen sie $u_1, \ldots, u_q$, $q = m - n$, ausdrücken wollen. Wir untersuchen also folgendes

Gleichungssystem in den Unbekannten $(\xi_1, \dots, \xi_n, u_1, \dots, u_q)$:

$$\begin{aligned} f_1(\xi_1, \dots, \xi_n, u_1, \dots, u_q) &= 0 \\ &\vdots \\ f_n(\xi_1, \dots, \xi_n, u_1, \dots, u_q) &= 0 \end{aligned} .$$

Ziel ist es, n Funktionen φ_i, $i = 1, \dots, n$, so zu finden, daß

$$f_i(\varphi_1(u_1, \dots, u_q), \dots, \varphi_n(u_1, \dots, u_q), u_1, \dots, u_q) = 0, \quad i = 1, \dots, n,$$

auf einer offenen Teilmenge $U \subset \mathbb{R}^q$ erfüllt ist.
Wir verwenden folgende Bezeichnungen

$$\begin{aligned} &x = (\xi_1, \dots, \xi_n), && x_0 = (\xi_1^0, \dots, \xi_n^0), \\ &u = (u_1, \dots, u_q), && u_0 = (u_1^0, \dots, u_q^0), \\ &f\colon D \to \mathbb{R}^n, D \subset \mathbb{R}^n \times \mathbb{R}^q, && f = (f_1, \dots, f_n). \end{aligned}$$

Es ist bequem, in diesem Zusammenhang den Begriff der partiellen Ableitung zu erweitern. Wir bezeichnen mit $\partial_x f(x_0, u_0) \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^n)$, $\partial_u f(x_0, u_0) \in \mathcal{L}(\mathbb{R}^q, \mathbb{R}^n)$ die linearen Abbildungen

$$\begin{aligned} \partial_x f(x_0, u_0)h &:= f'(x_0, u_0)(h, 0), && (h, 0) \in \mathbb{R}^n \times \mathbb{R}^q, \\ \partial_u f(x_0, u_0)k &:= f'(x_0, u_0)(0, k), && (0, k) \in \mathbb{R}^n \times \mathbb{R}^q. \end{aligned}$$

Man überzeuge sich davon, daß $\partial_x f(x_0, u_0)$ ($\partial_u f(x_0, u_0)$) die Ableitung im Sinne der Definition VIII-1.1 der *partiellen* Funktion $x \mapsto f(x, u_0)$ ($u \mapsto f(x_0, u)$) ist. Mit dieser Bezeichnung gilt

$$f'(x_0, u_0)(h, k) = \partial_x f(x_0, u_0)h + \partial_u f(x_0, u_0)k.$$

Theorem VIII-9.10 (Satz über implizite Funktionen). *Es sei $D \subset \mathbb{R}^n \times \mathbb{R}^q$ offen und $f\colon D \to \mathbb{R}^n$ stetig differenzierbar. An einer Stelle $(x_0, u_0) \in D$ gelte $f(x_0, u_0) = 0$ und es sei $\partial_x f(x_0, u_0)$ invertierbar. Dann gibt es eine Umgebung $U \in \mathcal{U}(u_0)$, eine Umgebung $V \in \mathcal{U}(x_0)$ und eine eindeutig bestimmte, stetig differenzierbare Abbildung $\varphi\colon U \to V$ mit*

$$\begin{aligned} \varphi(u_0) &= x_0 \\ \forall u \in U\colon f(\varphi(u), u) &= 0. \end{aligned}$$

Ferner gilt $\varphi'(u) = -(\partial_x f(\varphi(u), u))^{-1} \circ \partial_u f(\varphi(u), u)$. (Die Invertierbarkeit von $\partial_x f(x_0, u_0)$ ist gleichbedeutend mit

$$\det \begin{pmatrix} \frac{\partial f_1}{\partial \xi_1}(x_0, u_0) & \dots & \frac{\partial f_1}{\partial \xi_n}(x_0, u_0) \\ \vdots & & \vdots \\ \frac{\partial f_n}{\partial \xi_1}(x_0, u_0) & \dots & \frac{\partial f_n}{\partial \xi_n}(x_0, u_0) \end{pmatrix} \neq 0).$$

Beweis. Um den Satz von der Umkehrfunktion anwenden zu können, definieren wir die Abbildung

$$F := \begin{cases} D \to \mathbb{R}^n \times \mathbb{R}^q \\ (x,u) \mapsto (f(x,u),u) \end{cases} .$$

Ihre Ableitung ist für $(x,u) \in D$ und $(h,k) \in \mathbb{R}^n \times \mathbb{R}^q$ gegeben durch

$$F'(x,u)(h,k) = (\partial_x f(x,u)h + \partial_u f(x,u)k, k),$$

bzw. in Matrixschreibweise

$$(*) \qquad F'(x,u)(h,k) = \begin{pmatrix} \partial_x f(x,u) & \partial_u f(x,u) \\ 0 & id_{\mathbb{R}^q} \end{pmatrix} \begin{pmatrix} h \\ k \end{pmatrix}.$$

Aus der Stetigkeit der Elemente der Jakobi Matrix folgt die stetige Differenzierbarkeit von F. Nach Voraussetzung ist $\partial_x f(x_0,u_0)$ regulär und daher auch $F'(x_0,u_0)$. F erfüllt also alle Voraussetzungen des Umkehrsatzes VIII-9.5. Es gibt daher eine offene Umgebung $W \in \mathcal{U}((x_0,u_0))$ so, daß $F(W)$ offen und $F|_W$ ein C^1-Diffeomorphismus von W auf $F(W)$ ist. Es sei $G\colon F(W) \to W$, $G = (F|_W)^{-1}$. Dann gilt

$$(**) \qquad G'(F(x,u)) = (F'(x,u))^{-1}, \qquad (x,u) \in W.$$

G besitzt dieselbe Struktur wie F, d.h. es gibt eine stetig differenzierbare Funktion $g\colon F(W) \to \mathbb{R}^n$ mit

$$G(y,v) = (g(y,v),v), \qquad (y,v) \in F(W).$$

Es sei $(x,u) \in W$ eine Nullstelle von f. Es folgt $(0,u) \in F(W)$ und

$$(\dagger) \qquad \begin{aligned} f(x,u) = 0 \Leftrightarrow F(x,u) = (0,u) &\Leftrightarrow G(0,u) = (x,u) \\ &\Leftrightarrow g(0,u) = x. \end{aligned}$$

Setzt man umgekehrt $x = g(0,u)$ für jedes $u \in \mathbb{R}^q$ mit $(0,u) \in F(W)$, dann liegt $(x,u) \in W$ und $(\dagger)$ zeigt, daß (x,u) eine Nullstelle von f ist. Da W offen ist, gibt es Umgebungen $U \in \mathcal{U}(u_0)$ und $V \in \mathcal{U}(x_0)$ mit $V \times U \subset W$. Da auch $F(W)$ offen ist und $(0,u_0) \in F(W)$, kann man U so wählen, daß $\{0\} \times U \subset F(W)$ gilt. Wegen der Stetigkeit von g kann man weiters U so einrichten, daß $g(0,u) \in V$ für $u \in U$ gilt. Definieren wir nun die Abbildung

$$\varphi\colon U \to V, \quad \varphi(u) = g(0,u),$$

dann ist φ stetig differenzierbar auf U, es gilt $\varphi(u_0) = x_0$ und für alle $u \in U$ gilt wegen $(\dagger)$ die Äquivalenz

$$g(0,u) = \varphi(u) \Leftrightarrow f(\varphi(u),u) = 0.$$

Aus der Beziehung $(*)$ liest man ab, daß $F'(x,u)$ genau dann invertierbar ist, wenn $\partial_x f(x,u)$ invertierbar ist. Wegen $(**)$ existiert also $(\partial_x f(x,u))^{-1}$ auf W. Für die Berechnung der Ableitung von φ definieren wir die Abbildung $H\colon U \to \mathbb{R}^n \times \mathbb{R}^q$ durch

$$H(u) := (\varphi(u),u)(= (g(0,u),u) = G(0,u)) \in W.$$

Dann ist H auf U stetig differenzierbar mit

$$H'(u)k = (\varphi'(u)k, k), \quad k \in \mathbb{R}^q.$$

Ausgehend von der Identität $0 \equiv f(\varphi(u), u) = (f \circ H)(u)$ findet man mit Hilfe der Kettenregel

$$\begin{aligned} 0 &= ((f \circ H)'(u))k = (f'(H(u)) \circ H'(u))k = f'(\varphi(u), u)(\varphi'(u)k, k) \\ &= \partial_x f(\varphi(u), u)\varphi'(u)k + \partial_u f(\varphi(u), u)k. \end{aligned}$$

Wegen der Invertierbarkeit von $\partial_x f(x, u)$ für $(x, u) \in W$ erhält man schließlich

$$\varphi'(u)k = -(\partial_x f(\varphi(u), u))^{-1}\partial_u f(\varphi(u), u)k.$$

Schließlich zeigen wir noch die Eindeutigkeit der lösenden Funktion φ: Angenommen es gäbe eine weitere Funktion $\tilde{\varphi}\colon \tilde{U} \to V$, $\tilde{U} \in \mathcal{U}(u_0)$ mit $\tilde{\varphi}(u_0) = x_0$ und

$$f(\tilde{\varphi}(u), u) = 0, \qquad u \in \tilde{U}.$$

Da $(\tilde{\varphi}(u), u)$ die eindeutige Lösung der Gleichung $F(x, u) = 0$ in $U \times V$ ist, stimmt $\tilde{\varphi}(u)$ mit $\varphi(u)$ auf $U \cap \tilde{U} \in \mathcal{U}(u_0)$ überein. □

Speziell erhalten wir für eine Gleichung mit m Unbekannten:

Korollar VIII-9.11. Es sei $f\colon U \to \mathbb{R}$, $U \subset \mathbb{R}^m$ offen, stetig differenzierbar und $x_0 = (\xi_1^0, \ldots, \xi_m^0)$ eine Lösung von $f(x) = 0$ mit $\frac{\partial f(x_0)}{\partial \xi_m} \neq 0$. Dann gibt es Umgebungen U von $x_0' = (\xi_1^0, \ldots, \xi_{m-1}^0)$, V von ξ_m^0 und eine stetig differenzierbare Funktion $\varphi\colon U \to V$ mit $f(x', \varphi(x')) = 0$, $x' \in U$ und

$$\varphi'(x') = -\frac{1}{f_{\xi_m}(x)}(f_{\xi_1}(x), \ldots, f_{\xi_{m-1}}(x)), \quad x = (x', \varphi(x')) \in U \times V.$$

Als Anwendung betrachten wir die Niveaumenge einer stetig differenzierbaren Funktion (der Einfachheit halber) in zwei Veränderlichen zum Wert c, also

$$N_f(c) = \{x \in \mathbb{R}^2\colon f(x) = c\}.$$

Für alle $x \in N_f(c)$ sei $f'(x) \neq 0$. An einer Stelle $x_0 = (\xi_0, \eta_0) \in N_f(c)$ sei beispielsweise $f_\eta(x_0) \neq 0$. Dann gibt es Umgebungen $U \in \mathfrak{U}(\xi_0)$, $V \in \mathfrak{U}(\eta_0)$ und eine stetig differenzierbare Funktion $\varphi\colon U \to V$ mit

$$f(\xi, \varphi(\xi)) = c, \qquad \xi \in U.$$

Differenziert man diese Identität, erhält man

$$(*) \qquad f_\xi(\xi, \varphi(\xi)) + f_\eta(\xi, \varphi(\xi))\varphi'(\xi) = 0, \qquad \xi \in U.$$

Setzt man $\gamma\colon U \to \mathbb{R}^2$, $\gamma(x) = (x, \varphi(x))$, stellt γ eine stetig differenzierbare Parametrisierung von $N_f(c)$ in der Umgebung $U \times V$ von x_0 dar. Der Tangentialvektor an diese Kurve ist durch $\gamma'(\xi) = (1, \varphi'(\xi))$, $\xi \in U$, gegeben. In dieser Sprechweise ist $(*)$ äquivalent zu

$$\langle \operatorname{grad} f(\gamma(\xi)), \gamma'(\xi) \rangle = 0,$$

d.h. in jedem Punkt der Niveaumenge von f ist der Gradient von f orthogonal zum Tangentialvektor an die Niveaumenge in diesem Punkt. Dies stimmt mit dem früheren Befund überein, daß der Gradient von f stets in die Richtung des stärksten Anstiegs der Funktionswerte von f weist.

10. Extrema mit Nebenbedingungen

Bisher haben wir nur kritische Punkte aus dem *Inneren* des Definitionsbereiches von f zu untersucht (wir haben dies erzwungen, indem wir $\operatorname{def} f$ offen voraussetzten). Der Funktionswert $f(x_0)$ kann daher mit den Werten von f an jeder benachbarten Stelle aus einer *vollen* Umgebung von x_0 verglichen werden. Dies ermöglicht den effizienten Einsatz von Methoden der Differentialrechnung. Oft treten lokale Extrema jedoch am Rand von $\operatorname{def} f$ auf. Suchen wir etwa die lokalen Extrema von f auf $\bar{K}(0,1)$, beschränkt man sich vorerst auf das Innere $K(0,1)$ und kann die Methoden von Abschnitt 7 anwenden. Als nächstes ist zu untersuchen, ob auf dem Rand von $\bar{K}(0,1)$, d.h. im Falle $\operatorname{def} f \subset \mathbb{R}^2$ auf $S(0,1) = \{(\xi,\eta)\colon \xi^2+\eta^2 = 1\}$, lokale Extrema liegen. Man sucht also in diesem Falle lokale Extrema von f unter der *Gleichungs*nebenbedingung $g(\xi,\eta) = \xi^2+\eta^2-1 = 0$. Anschließend muß man prüfen, ob diese Stellen auch lokale Extrema bzgl. $\bar{K}(0,1)$ sind.

Beispiel VIII-10.1. Wir untersuchen das Problem

$$\text{Maximiere: } f(\xi,\eta) = \xi\eta$$

unter der

$$\text{Nebenbedingung: } g(\xi,\eta) = \xi^2+\eta^2-8 = 0.$$

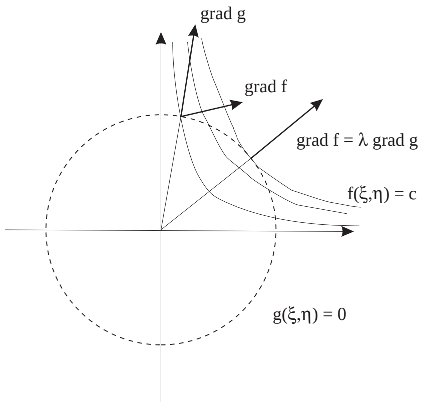
In diesem Zusammenhang nennt man die zu minimierende (maximierende) Funktion f oft **Kostenfunktional**. Wir geben eine geometrische, heuristische Lösung an. Dazu beachten wir, daß die Niveaulinien von f, d.h. die Kurven, auf denen f einen konstanten Wert annimmt, Hyperbeln sind. Weiters gilt

$$\begin{aligned} f'(\xi,\eta) &= (\eta \quad \xi), \\ g'(\xi,\eta) &= (2\xi \quad 2\eta). \end{aligned}$$

Die Niveaulinie von f zum Wert $c = 1$ schneidet den Kreis (im 1. Quadranten) in zwei Punkten. Betrachtet man die Niveaulinien in Richtung des Gradienten von f, wird der zugehörige Funktionswert größer. Auch diese Niveaulinien schneiden den Kreis solange an zwei Stellen, bis die Niveaulinie den Kreis nur mehr berührt. Dies ist in $(2,2)$ der Fall. Da weitere Niveaulinien in Richtung des Gradienten den Kreis nicht mehr schneiden, ist $(2,2)$ jener Punkt des Kreises, in welchem f maximalen Wert hat.

Man erkennt also: die Lösung muß an einer Stelle auftreten, an der $\operatorname{grad} f$ und $\operatorname{grad} g$ linear abhängig sind. Dies führt auf das Gleichungssystem

$$\begin{cases} \operatorname{grad} f + \lambda \operatorname{grad} g = 0, \\ g(\xi,\eta) = 0, \end{cases}$$



mit den beiden Lösungen $(2,2)$ und $(-2,-2)$. Die lineare Abhängigkeit von f' und g' in Beispiel VIII-10.1 ist kein Zufall, sondern problemimmanent.
In vielen Fällen ist es möglich, die Gleichungsnebenbedingung nach einer Variablen aufzulösen. Angenommen wir interessieren uns für die lokalen Extrema von $f\colon \mathbb{R}^3 \to \mathbb{R}$ unter der Nebenbedingung $g(\xi,\eta,\zeta) = 0 = h(\xi,\eta) - \zeta$. Nehmen wir ferner an, h und f seien stetig differenzierbar. Das Problem

$$\text{Maximiere: } f(\xi,\eta,\zeta)$$

unter der

$$\text{Nebenbedingung: } g(\xi,\eta,\zeta) = 0$$

ist daher gleichwertig mit dem *nicht restringierten* Problem

$$\text{Maximiere } F(\xi,\eta) = f(\xi,\eta,h(\xi,\eta)).$$

Setzt man $H\colon \mathbb{R}^2 \to \mathbb{R}^3$, $(\xi,\eta) \mapsto (\xi,\eta,h(\xi,\eta))$, gilt $F = f \circ H$. Nach Satz **??** ergeben sich die kritischen Punkte von F aus $(x = (\xi,\eta))$

$$\begin{aligned}
0 = F'(x) &= f'(H(x)) \circ H'(x) \\
&= \begin{pmatrix} \frac{\partial f}{\partial \xi}(H(x)) & \frac{\partial f}{\partial \eta}(H(x)) & \frac{\partial f}{\partial \zeta}(H(x)) \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ \frac{\partial h}{\partial \xi}(x) & \frac{\partial h}{\partial \eta}(x) \end{pmatrix} \\
&= \begin{pmatrix} \frac{\partial f}{\partial \xi}(H(x)) + \frac{\partial f}{\partial \zeta}(H(x))\frac{\partial h}{\partial \xi}(x), & \frac{\partial f}{\partial \eta}(H(x)) + \frac{\partial f}{\partial \zeta}(H(x))\frac{\partial h}{\partial \eta}(x) \end{pmatrix}.
\end{aligned}$$

Dies kann mit $\partial_x = (\partial_\xi \quad \partial_\eta)$ kompakter in der Form

$$(*) \qquad \partial_x f(H(x)) + \frac{\partial f}{\partial \zeta}(H(x))\partial_x h(x) = 0$$

geschrieben werden. Berücksichtigen wir noch

$$g'(\xi,\eta,\zeta) = (\partial_x h(\xi,\eta) \quad -1),$$

und ergänzen wir $(*)$ durch die triviale Gleichung

$$-\frac{\partial f}{\partial \zeta}(H(x)) = \frac{\partial g}{\partial \zeta}\frac{\partial f}{\partial \zeta}(H(x)),$$

also

$$\frac{\partial f}{\partial \zeta}(H(x)) + \frac{\partial f}{\partial \zeta}(H(x))\frac{\partial g}{\partial \zeta} = 0,$$

können wir $(*)$ in der Form

$$f'(H(x)) + \frac{\partial f}{\partial \zeta}(H(x))g'(x) = 0$$

schreiben. Setzen wir $\lambda = \frac{\partial f}{\partial \zeta}(H(x))$, erhalten wir die kritischen Punkte von F aus dem Gleichungssystem

$$\begin{aligned} f'(H(x)) + \lambda g'(x) &= 0, \\ g(x) &= 0. \end{aligned}$$

Diese Vorgangsweise ermöglicht es, ein Optimierungsproblem mit einer Gleichungsnebenbedingung auf ein nicht restringiertes Problem zurückzuführen. Allerdings tritt nun eine neue, zusätzliche Variable λ auf, die ebenfalls berechnet werden muß. Wir zeigen nun, daß dieses Verfahren unter bestimmten Voraussetzungen auch dann funktioniert, wenn es nicht möglich ist, die Gleichungsnebenbedingung explizit nach einer Variablen aufzulösen.

Theorem VIII-10.2 (Lagrangesche Multiplikatorregel). *Die Abbildungen $f\colon D \to \mathbb{R}$, $g\colon D \to \mathbb{R}^n$, $D \subset \mathbb{R}^m$, $n < m$, seien stetig differenzierbar. f besitze in x_0 ein lokales Extremum unter der Nebenbedingung $g(x) = 0$. Die Jacobimatrix von g an der Stelle x_0 habe Rang n (d.h. die Vektoren $\operatorname{grad} g_i(x_0)$, $i = 1, \dots, n$, sind linear unabhängig). Dann existieren n Zahlen $\lambda_1, \dots, \lambda_n \in \mathbb{R}$ (**Lagrange Multiplikatoren**) derart, daß*

$$f'(x_0) + \sum_{i=1}^{n} \lambda_i g_i'(x_0) = 0,$$

also das Gleichungssystem

$$\begin{array}{ccccccccccc} \frac{\partial f(x_0)}{\partial \xi_1} & + & \lambda_1 \frac{\partial g_1(x_0)}{\partial \xi_1} & + & \lambda_2 \frac{\partial g_2(x_0)}{\partial \xi_1} & + & \dots & + & \lambda_n \frac{\partial g_n(x_0)}{\partial \xi_1} & = & 0 \\ & & & & & & & & & \vdots & \\ \frac{\partial f(x_0)}{\partial \xi_m} & + & \lambda_1 \frac{\partial g_1(x_0)}{\partial \xi_m} & + & \lambda_2 \frac{\partial g_2(x_0)}{\partial \xi_m} & + & \dots & + & \lambda_n \frac{\partial g_n(x_0)}{\partial \xi_m} & = & 0 \end{array}$$

erfüllt ist.

Beweis. 1. Schritt: Elimination von n Variablen in g.
Die Voraussetzung Rang $g'(x_0) = n$ bedeutet, daß n Spalten von $g'(x_0)$ linear unabhängig sind. O.B.d.A. können wir annehmen, daß dies die ersten n Spalten sind.

Wir setzen nun

$$\begin{aligned} x &= (\xi_1, \dots, \xi_m), & x_0 &= (\xi_1^0, \dots, \xi_m^0),\\ y &= (\xi_1, \dots, \xi_n), & y_0 &= (\xi_1^0, \dots, \xi_n^0),\\ z &= (\xi_{n+1}, \dots, \xi_m), & z_0 &= (\xi_{n+1}^0, \dots, \xi_m^0). \end{aligned}$$

An Stelle von $f(x)$, $g(x)$ schreiben wir $f(y,z)$, $g(y,z)$. Da f in $x_0 = (y_0, z_0)$ ein lokales Extremum unter der Nebenbedingung $g(x) = g(y,z) = 0$ besitzt, gilt

$$g(y_0, z_0) = 0.$$

Die ersten n Spalten von $g'(x_0)$ sind linear unabhängig, dies ist gleichwertig mit der Invertierbarkeit von $\partial_y g(y_0, z_0)$. Dem Satz über implizite Funktionen VIII-9.10 entnehmen wir nun, daß $g(y,z)$ in einer Umgebung von z_0 stetig differenzierbar nach y aufgelöst werden kann, d.h. es gibt eine Umgebung $U \in \mathcal{U}(z_0)$, $U \subset \mathbb{R}^{m-n}$, und eine stetig differenzierbare Funktion $\varphi\colon U \to \mathbb{R}^n$ mit der Eigenschaft

$$\begin{aligned} \forall z \in U\colon g(\varphi(z), z) &= 0,\\ \varphi(z_0) &= y_0. \end{aligned}$$

Ferner gilt für alle $z \in U$

$$\varphi'(z) = -(\partial_y g(\varphi(z), z))^{-1} \partial_z g(\varphi(z), z).$$

2. Schritt: Ableitung des nicht restringierten Kostenfunktionals.
Für alle $z \in U$ gilt $(\varphi(z), z) \in D$. Daher können wir auch eine Abbildung F auf U definieren durch

$$F := \begin{cases} U \to \mathbb{R} \\ z \mapsto f(\varphi(z), z) \end{cases}.$$

Es ist klar, daß f an der Stelle $x_0 = (y_0, z_0)$ genau dann ein lokales Extremum unter der Nebenbedinung $g(y,z) = 0$ besitzt, wenn F an der Stelle z_0 ein lokales Extremum *ohne* Nebenbedinungen besitzt. Nach Satz VIII-5.2 muß also

$$F'(z_0) = 0$$

gelten. Setzen wir

$$H : \begin{cases} U \to \mathbb{R}^n \times \mathbb{R}^{m-n} \\ z \mapsto (\varphi(z), z), \end{cases}$$

ist H stetig differenzierbar auf U und es gilt für $h \in \mathbb{R}^{m-n}$

$$H'(z)h = (\varphi'(z)h, h).$$

Wegen $F = f \circ H$ (beachte $H(U) \subset D$) erhalten wir mit Hilfe der Kettenregel für $h \in \mathbb{R}^{m-n}$

$$\begin{aligned} F'(z)h &= f'(H(z)) \circ H'(z)h \\ &= f'(H(z))(\varphi'(z)h, h) \\ &= f'(H(z))((\varphi'(z)h, 0) + (0, h)) \\ &= \partial_y f(H(z))\varphi'(z)h + \partial_z f(H(z))h, \end{aligned}$$

d.h. z_0 genügt der Gleichung

$$(*) \qquad \partial_y f(H(z_0))\varphi'(z_0) + \partial_z f(H(z_0)) = 0.$$

3. Schritt: Konstruktion des Lagrange Multiplikators.
Wir eliminieren $\varphi'(z_0)$ in $(*)$, indem wir

$$\varphi'(z_0) = -(\partial_y g(y_0, z_0))^{-1}\partial_z g(y_0, z_0)$$

einsetzen und erhalten mit $H(z_0) = (y_0, z_0)$

$$(**) \qquad -\partial_y f(y_0, z_0)(\partial_y g(y_0, z_0))^{-1}\partial_z g(y_0, z_0) + \partial_z f(y_0, z_0) = 0.$$

Beachten wir

$$\partial_y f(y_0, z_0) \colon \mathbb{R}^n \to \mathbb{R}, \quad (\partial_y g(y_0, z_0))^{-1} \colon \mathbb{R}^n \to \mathbb{R}^n,$$

also

$$\Lambda := -\partial_y f(y_0, z_0)(\partial_y g(y_0, z_0))^{-1} \colon \mathbb{R}^n \to \mathbb{R},$$

läßt sich $(**)$ mit Hilfe von Λ in der Form

$$\partial_z f(y_0, z_0) + \Lambda \partial_z g(y_0, z_0) = 0$$

schreiben. Aus der Definition von Λ ergibt sich weiters die Beziehung

$$\partial_y f(y_0, z_0) + \Lambda \partial_y g(y_0, z_0) = 0.$$

Die Matrixdarstellung von Λ ist eine $1 \times n$- Matrix $(\lambda_1 \ldots \lambda_n)$. Die beiden letzten Gleichungen können kombiniert werden zu

$$f'(y_0, z_0) + \Lambda g'(y_0, z_0) = f'(x_0) + \sum_{i=1}^{n} \lambda_i g_i'(x_0) = 0.$$

□

Bei der Lösung der Aufgabe, die Stellen lokaler Extrema von f unter der Nebenbedingung $g(x) = 0$ zu bestimmen, kann man routinemäßig so vorgehen, daß man zuerst die sogenannte **Lagrangefunktion** $\mathcal{L} \colon \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}$,

$$\mathcal{L}(x, \lambda) = f(x) + \sum_{i=1}^{n} \lambda_i g_i(x)$$

bildet und dann die kritischen Stellen von $\mathcal{L}$ berechnet,

$$\begin{aligned}\partial_x \mathcal{L}(x,\lambda) &= f'(x) + \sum_{i=1}^{n} \lambda_i g_i'(x) = 0,\\ \partial_\lambda \mathcal{L}(x,\lambda) &= g(x) = 0.\end{aligned}$$

Jede Lösung (x_0, λ_0) dieses Gleichungssystems mit der Eigenschaft Rang $g'(x_0) = n$ kommt als Stelle eines lokalen Extremums von f unter der Nebenbedingung g in Frage. Ob dies tatsächlich der Fall ist, bedarf einer eigenen Untersuchung. Ebenso sind jene Stellen, welche von der Lagrangeschen Multiplikatorregel nicht erfaßt werden, also Stellen mit $g(x) = 0$ und Rang $g'(x) < n$, gesondert zu behandeln.

Bemerkung VIII-10.3. Ist die Restriktionsmenge $R = \{x \in \mathbb{R}^m \colon g(x) = 0\}$ kompakt, so besitzt $f|_R$ ein globales Maximum und Minimum, mit anderen Worten, f hat unter der Nebenbedingung $g(x) = 0$ ein globales und damit auch lokales Maximum und Minimum. Der größte bzw. kleinste der Werte $f(x)$, die man mit Hilfe des vorhin beschriebenen Verfahrens gefunden hat, ist dann das gesuchte Maximum bzw. Minimum von f unter der Nebenbedingung $g(x) = 0$.

Beispiel VIII-10.4.

$$\begin{aligned}\text{Maximiere: } & f(\xi,\eta,\zeta) = \xi + \eta + \zeta,\\ \text{Nebenbedingung: } & \zeta = 1 - 7\xi^2 - 3\eta^2.\end{aligned}$$

Eliminationsmethode:
Wir lösen das gleichwertige Problem, in dem die Nebenbedingung eliminiert wurde:

$$\text{Maximiere } F(\xi,\eta) = (\xi + \eta) + (1 - 7\xi^2 - 3\eta^2).$$

Die notwendige Optimialitätsbedingung lautet

$$F'(\xi,\eta) = 0 = (1 - 14\xi \quad 1 - 6\eta),$$

d.h.

$$(\xi^0, \eta^0) = (\frac{1}{14}, \frac{1}{6}).$$

Aus der Nebenbedingung erhalten wir $\zeta^0 = \frac{37}{42}$. Die Hesse Matrix von F

$$\mathcal{H}_F(\frac{1}{14}, \frac{1}{6}) = \begin{pmatrix} -14 & 0 \\ 0 & -6 \end{pmatrix}$$

ist negativ definit, somit besitzt F an der Stelle $(\frac{1}{14}, \frac{1}{6})$ und damit f an der Stelle $(\frac{1}{14}, \frac{1}{6}, \frac{37}{42})$ ein lokales Maximum. Wegen $\lim_{\|(\xi,\eta)\|\to\infty} F(\xi,\eta) = -\infty$ wird an dieser Stelle sogar das globale Maximum angenommen.
Lagrangesche Multiplikatormethode:
Wir bilden die Lagrange Funktion $(g(\xi,\eta,\zeta) = 1 - 7\xi^2 - 3\eta^2 - \zeta)$

$$\mathcal{L}(\xi,\eta,\zeta,\lambda) = \xi + \eta + \zeta + \lambda(1 - 7\xi^2 - 3\eta^2 - \zeta)$$

und bestimmen die Nullstellen von $\mathcal{L}'$:

$$\frac{\partial \mathcal{L}}{\partial \xi} = 1 - \lambda \cdot 14\xi = 0$$
$$\frac{\partial \mathcal{L}}{\partial \eta} = 1 - \lambda \cdot 6\eta = 0$$
$$\frac{\partial \mathcal{L}}{\partial \zeta} = 1 - \lambda = 0$$
$$\frac{\partial \mathcal{L}}{\partial \lambda} = 1 - 7\xi^2 - 3\eta^2 - \zeta = 0.$$

Als Lösung ergibt sich $\lambda = 1$ und $(\xi, \eta, \zeta) = (\frac{1}{14}, \frac{1}{6}, \frac{37}{42})$. Wegen

$$g'(\xi, \eta, \zeta) = (-14\xi \quad -6\eta \quad -1) \neq 0$$

kommt nur die Stelle $(\frac{1}{14}, \frac{1}{6}, \frac{37}{42})$ für ein lokales Extremum in Frage. Da g' überall regulär ist, kommen alle lokalen Extrema unter den Nullstellen von $\mathcal{L}'$ vor. Es gibt allerdings nur eine Nullstelle von $\mathcal{L}'$, somit muß an dieser Stelle ein globales Extremum vorliegen. Durch Vergleich mit dem Funktionswert an einer beliebigen zulässigen Stelle erkennt man, daß in $(\frac{1}{14}, \frac{1}{6}, \frac{37}{42})$ das globale Maximum angenommen wird.

Beispiel VIII-10.5.

$$\begin{aligned} \text{Maximiere } & f(\xi, \eta, \zeta) = \zeta, \\ \text{Nebenbedingung: } & g_1(\xi, \eta, \zeta) = \xi^2 + \eta^2 + \zeta^2 - 1 = 0, \\ & g_2(\xi, \eta, \zeta) = 3\xi\eta - 4\zeta = 0. \end{aligned}$$

Wir bilden die Lagrange Funktion

$$\mathcal{L}(\xi, \eta, \zeta, \lambda_1, \lambda_2) = \zeta + \lambda_1(\xi^2 + \eta^2 + \zeta^2 - 1) + \lambda_2(3\xi\eta - 4\zeta).$$

Die Nullstellen von $\mathcal{L}'$ ergeben sich aus

$$\frac{\partial \mathcal{L}}{\partial \xi} = 2\lambda_1\xi + 3\lambda_2\eta = 0$$
$$\frac{\partial \mathcal{L}}{\partial \eta} = 2\lambda_1\eta + 3\lambda_2\xi = 0$$
$$\frac{\partial \mathcal{L}}{\partial \zeta} = 1 + 2\lambda_1\zeta - 4\lambda_2 = 0$$
$$\frac{\partial \mathcal{L}}{\partial \lambda_1} = \xi^2 + \eta^2 + \zeta^2 - 1 = 0$$
$$\frac{\partial \mathcal{L}}{\partial \lambda_2} = 3\xi\eta - 4\zeta = 0.$$

Wir überlegen zuerst, daß $\lambda_2 = 0$ nicht möglich ist. Die 3. Gleichung zeigt, daß beide Lagrangemultiplikatoren nicht gleichzeitig verschwinden können. Wäre nun $\lambda_2 = 0$, ergäbe sich aus den ersten beiden Gleichungen ($\lambda_1 \neq 0!$) $\xi = \eta = 0$ und aus der letzten Gleichung $\zeta = 0$. Dies widerspricht der vorletzten Gleichung. Auf ähnliche Weise folgt aus den ersten beiden Gleichungen $\xi \neq 0$, $\eta \neq 0$ und damit

$$\frac{2\lambda_1}{3\lambda_2} = -\frac{\eta}{\xi} = -\frac{\xi}{\eta},$$

also

$$\xi^2 = \eta^2,$$

bzw.

$$\xi = \pm\eta.$$

Aus der letzten Gleichung folgt

$$\zeta = \pm\frac{3}{4}\xi^2,$$

die vierte Gleichung ergibt

$$2\xi^2 + \frac{9}{16}\xi^4 - 1 = 0,$$

also $\xi^2 = \frac{4}{9}$ (bzw. $\xi^2 = -4$, dies führt auf keine reelle Lösung).
Wir erhalten also

$$\xi = \pm\frac{2}{3},$$

und damit die kritischen Stellen von $\mathcal{L}$,

$$(\frac{2}{3}, \frac{2}{3}, \frac{1}{3}), \quad (\frac{2}{3}, -\frac{2}{3}, -\frac{1}{3}), \quad (-\frac{2}{3}, -\frac{2}{3}, \frac{1}{3}), \quad (-\frac{2}{3}, \frac{2}{3}, -\frac{1}{3}).$$

Die Restriktionsmenge R ist eine abgeschlossene Teilmenge von $S(0,1)$ und somit kompakt. f nimmt auf R das Maximum an. Dies hat den Wert $\frac{1}{3}$ und tritt an den Stellen $(\frac{2}{3}, \frac{2}{3}, \frac{1}{3})$ und $(-\frac{2}{3}, -\frac{2}{3}, \frac{1}{3})$ auf.

Ob an einer kritischen Stelle der Lagrangefunktion tatsächlich ein lokales Extremum vorliegt, kann man aus deren zweiter Ableitung ablesen. Wir nennen einen kritischen Punkt (x_0, λ_0) der Lagrangefunktion **nichtentartet** , wenn die sogenannte **geränderte Hessesche Matrix** $\mathcal{G}$ von $\mathcal{L}$ in (x_0, λ_0) regulär ist, d.h. wenn

$$\mathcal{G}(x_0, \lambda_0) = \begin{pmatrix} \mathcal{H}_{\mathcal{L}}(x_0) & g'(x_0)^T \\ g'(x_0) & 0 \end{pmatrix}$$

vollen Rang hat. Wir weisen darauf hin, daß im Folgenden die Lagrangefunktion stets für den festen Lagrange Multiplikator $\lambda = \lambda_0$ ausgewertet wird. Insbesondere bedeutet $\mathcal{H}_{\mathcal{L}}(x_0)$ die Hesse Matrix der Abbildung $x \to \mathcal{L}(x, \lambda_0)$ an der Stelle $x = x_0$.

Theorem VIII-10.6. *Die Abbildungen $f\colon D \to \mathbb{R}$ und $g\colon D \to \mathbb{R}^n$, $D \subset \mathbb{R}^m$ offen, $n < m$,seien stetig differenzierbar und (x_0, λ_0) sei ein nichtentarteter kritischer Punkt der Lagrangefunktion. Dann besitzt f in x_0 ein lokales Minimum unter der Nebenbedingung $g(x) = 0$ genau dann, wenn $h^T \mathcal{H}_{\mathcal{L}}(x_0) h \geq 0$ für alle $h \in ker\, g'(x_0)$ gilt.*

Für den Beweis dieses Kriteriums verweisen wir auf die einschlägige Literatur Wir demonstrieren die Anwendung dieses Kriteriums an Hand der Beispiele VIII-10.4 und VIII-10.5. Es wurde allerdings bereits gezeigt, daß es gelegentlich einfachere Argumente gibt, um die kritischen Stellen der Lagrangefunktion zu untersuchen.

Beispiel VIII-10.7 (Fortsetzung von Beipiel VIII-10.4). Es wurde bereits gezeigt, daß die Lagrangefunktion

$$\mathcal{L}(\xi, \eta, \zeta, \lambda) = \xi + \eta + \zeta + \lambda(1 - 7\xi^2 - 3\eta^2 - \zeta)$$

nur in $(x_0, \lambda_0) = (\frac{1}{14}, \frac{1}{6}, \frac{37}{42}, 1)$ eine kritische Stelle hat. Wertet man die Jacobi Matrix von g in x_0 aus, erhält man

$$J_g(x_0) = (-1, -1, -1).$$

Wir benötigen noch die Hesse Matrix von $\mathcal{L}(x, 1)$ in $x = x_0$. Dazu berechnen wir

$$\begin{array}{llll}
\mathcal{L}_\xi = 1 - 14\lambda\xi, & \mathcal{L}_{\xi\xi} = -14\lambda, & \mathcal{L}_{\xi\eta} = 0, & \mathcal{L}_{\xi\zeta} = 0, \\
\mathcal{L}_\eta = 1 - 6\lambda\eta, & \mathcal{L}_{\eta\xi} = 0, & \mathcal{L}_{\eta\eta} = -6\lambda, & \mathcal{L}_{\eta\zeta} = 0, \\
\mathcal{L}_\zeta = 1 - \lambda, & \mathcal{L}_{\zeta\xi} = 0, & \mathcal{L}_{\zeta\eta} = 0, & \mathcal{L}_{\zeta\zeta} = 0.
\end{array}$$

Als nächstes bauen wir die geränderte Hesse Matrix in $(x_0, \lambda_0) = (x_0, 1)$ auf:

$$\mathcal{G}(x_0, 1) = \begin{pmatrix} \mathcal{H}_{\mathcal{L}}(x_0) & J_g(x_0)^T \\ J_g(x_0) & 0 \end{pmatrix} = \begin{pmatrix} -14 & 0 & 0 & -1 \\ 0 & -6 & 0 & -1 \\ 0 & 0 & 0 & -1 \\ -1 & -1 & -1 & 0 \end{pmatrix}.$$

Man verifiziert leicht, daß die Matrix $\mathcal{G}(x_0, 1)$ regulär ist. Die Hesse Matrix $\mathcal{H}_{\mathcal{L}}(x_0)$ ist eine Diagonalmatrix mit den Eigenwerten $\{-14, -6, 0\}$ und daher negativ semidefinit. Nach Satz VIII-10.6 nimmt f in x_0 ein lokales Maximum an. Es wurde bereits im Beispiel VIII-10.4 gezeigt, daß in x_0 sogar das globale Maximum angenommen wird.

Beispiel VIII-10.8 (Fortsetzung von Beipiel VIII-10.5). Wir verifizieren mit Hilfe von Satz VIII-10.6, daß an der Stelle $x_0 = (\frac{2}{3}, \frac{2}{3}, \frac{1}{3})$ ein lokales Maximum angenommen wird. Eine einfache Rechnung ergibt die Werte der zugehörigen Lagrange Multiplikatoren

$$\lambda_1 = -\frac{3}{10}, \quad \lambda_2 = \frac{1}{5}.$$

Weiters ist

$$Dg(\xi, \eta, \zeta) = \begin{pmatrix} 2\xi & 2\eta & 2\zeta \\ 3\eta & 3\xi & -4 \end{pmatrix}$$

und somit

$$Dg(x_0) = \begin{pmatrix} \frac{4}{3} & \frac{4}{3} & \frac{2}{3} \\ 2 & 2 & -4 \end{pmatrix}.$$

Der Kern von $Dg(x_0)$ ist gegeben durch

$$\ker Dg(x_0) = \{ \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} t \colon t \in \mathbb{R} \}$$

Die Hesse Matrix von $(\xi, \eta, \zeta) \mapsto \mathcal{L}(\xi, \eta, \zeta, \lambda_1, \lambda_2)$ ist gegeben durch

$$\mathcal{H}_{\mathcal{L}}(x_0) = \begin{pmatrix} 2\lambda_1 & 3\lambda_2 & 0 \\ 3\lambda_2 & 2\lambda_1 & 0 \\ 0 & 0 & 2\lambda_1 \end{pmatrix} = \begin{pmatrix} -\frac{3}{5} & \frac{3}{5} & 0 \\ \frac{3}{5} & -\frac{3}{5} & 0 \\ 0 & 0 & -\frac{3}{5} \end{pmatrix},$$

die geränderte Hesse Matrix ist somit

$$\mathcal{G}(x_0, \lambda_1, \lambda_2) = \begin{pmatrix} -\frac{3}{5} & \frac{3}{5} & 0 & \frac{4}{3} & 2 \\ \frac{3}{5} & -\frac{3}{5} & 0 & \frac{4}{3} & 2 \\ 0 & 0 & -\frac{3}{5} & \frac{2}{3} & -4 \\ \frac{4}{3} & \frac{4}{3} & \frac{2}{3} & 0 & 0 \\ 2 & 2 & -4 & 0 & 0 \end{pmatrix}.$$

Man verifiziere, daß $\mathcal{G}(x_0, \lambda_1, \lambda_2)$ regulär ist. Da der Kern von $Dg(x_0)$ Dimension 1 hat, genügt es, $h^T \mathcal{H}_{\mathcal{L}}(x_0) h \geq 0$ für einen Basisvektor von $\ker Dg(x_0)$ zu zeigen, etwa für $h = (-1, 1, 0)^T$. Wegen

$$\begin{pmatrix} -1 & 1 & 0 \end{pmatrix} \begin{pmatrix} -\frac{3}{5} & \frac{3}{5} & 0 \\ \frac{3}{5} & -\frac{3}{5} & 0 \\ 0 & 0 & -\frac{3}{5} \end{pmatrix} \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} = -\frac{12}{5} < 0$$

liegt in x_0 ein lokales Maximum vor. Da $Dg(\xi, \eta, \zeta)$ für jede Wahl von (ξ, η, ζ) vollen Rang hat, kann es, abgesehen von $(-\frac{2}{3}, -\frac{2}{3}, \frac{1}{3})$, keine weiteren Stellen geben, an denen f ein lokales Maximum annimmt.

KAPITEL IX

Integralrechnung im $\mathbb{R}^n$

Notation: Im Folgenden bezeichnen wir mit $f_n \rightrightarrows f$ die gleichmäßige Konvergenz der Funktionen f_n gegen die Grenzfunktion f.

1. Halbstetige Funktionen

Definition IX-1.1. (1) *Eine Abbildung $f\colon \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ heißt im Punkt $x_0 \in \mathbb{R}^n$ **von unten halbstetig**, wenn zu jedem $c \in \mathbb{R}$ mit $c < f(x_0)$ eine Umgebung $U \in \mathcal{U}(x_0)$ mit $f(U) \subset (c, \infty]$ existiert.*

(2) *Eine Abbildung $f\colon \mathbb{R}^n \to \mathbb{R} \cup \{-\infty\}$ heißt im Punkt $x_0 \in \mathbb{R}^n$ **von oben halbstetig**, wenn zu jedem $c \in \mathbb{R}$ mit $c > f(x_0)$ eine Umgebung $U \in \mathcal{U}(x_0)$ mit $f(U) \subset [-\infty, c)$ existiert.*

*Die Abbildung f heißt **von unten (oben) halbstetig**, wenn sie an jeder Stelle $x_0 \in \mathbb{R}^n$ von unten (oben) halbstetig ist.*

Der Einfachheit halber betrachten wir nur Funktionen, welche auf $\mathbb{R}^n$ definiert sind. Die Erweiterung dieses Konzeptes auf Funktionen $f\colon D \to \mathbb{R}$, $D \subset \mathbb{R}^n$, sei dem Leser überlassen. Offenbar ist eine Abbildung f in x_0 von oben halbstetig genau dann, wenn $-f$ in x_0 von unten halbstetig ist.
Schreibt man c in der Form $f(x_0) - \varepsilon$, dann ist die Halbstetigkeit von unten von f in x_0 gleichwertig mit

$$(*) \qquad \forall \varepsilon > 0 \, \exists \delta > 0 \, \forall x \in \mathbb{R}^n\colon \|x - x_0\| < \delta \Rightarrow f(x_0) - f(x) < \varepsilon.$$

Wäre f stetig in x_0, dann wäre in $(*)$ die Folgerung $f(x_0) - \varepsilon < f(x) < f(x_0) + \varepsilon$. Halbstetige Funktionen nützen also nur eine der beiden Ungleichungen. Somit ist eine Funktion in x_0 genau dann stetig, wenn sie in x_0 von unten und von oben halbstetig ist.

Lemma IX-1.2. *1.) Eine Abbildung $f\colon \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ ist von unten halbstetig genau dann, wenn $f^{-1}((c, \infty])$ für alle $c \in \mathbb{R}$ offen ist.*
2.)Eine Abbildung $f\colon \mathbb{R}^n \to \mathbb{R} \cup \{-\infty\}$ ist von oben halbstetig genau dann, wenn $f^{-1}([-\infty, c))$ für alle $c \in \mathbb{R}$ offen ist.

Beweis. Es genügt die erste Behauptung zu zeigen: "$\Rightarrow$" Es sei $c \in \mathbb{R}$ und $x \in f^{-1}((c, \infty])$, also $c < f(x)$. Da f in x von unten halbstetig ist, gibt es eine Umgebung $U \in \mathcal{U}(x)$ mit $f(U) \subset (c, \infty]$, also $U \subset f^{-1}((c, \infty])$. Somit ist x ein innerer Punkt von $f^{-1}((c, \infty])$.

"$\Leftarrow$"Es sei $x \in \mathbb{R}^n$ und $c < f(x)$, also $x \in f^{-1}((c,\infty])$. Da $f^{-1}((c,\infty])$ offen ist, existiert eine Umgebung $U \in \mathcal{U}(x)$ mit $U \subset f^{-1}((c,\infty])$, also $f(U) \subset (c,\infty]$. □

Beispiel IX-1.3. 1) Die **charakteristische Funktion** χ_A einer Teilmenge $A \subset \mathbb{R}^n$ ist gegeben durch

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

Es gilt: A ist offen genau dann, wenn χ_A von unten halbstetig ist. Dies folgt aus

$$\chi_A^{-1}((c,\infty]) = \begin{cases} \mathbb{R}^n, & c < 0, \\ A, & 0 \le c < 1, \\ \emptyset, & c \ge 1. \end{cases}$$

2) Ersetzt man A durch $\complement A$ und berücksichtigt $\chi_{\complement A} = 1 - \chi_A$ findet man: A ist abgeschlossen genau dann, wenn χ_A von oben halbstetig ist.
3) Die Abbildung $f\colon \mathbb{R} \to \mathbb{R}$ definiert durch $f(x) = 0$ für $x \le 0$ und $f(x) = 1$ für $x > 0$ ist in $x = 0$ von unten halbstetig. Für $c < 0 = f(0)$ erfüllt nämlich sogar jede Umgebung U von $x = 0$ die Bedingung $f(U) \subset (c,\infty]$. Ein anderer Beweis ergibt sich aus der Beobachtung $f = \chi_{(0,\infty)}$.

Auch die Halbstetigkeit einer Funktion kann man mittels Folgen charakterisieren:

Lemma IX-1.4. *Eine Funktion $f\colon \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ ist in $x_0 \in \mathbb{R}^n$ von unten halbstetig genau dann, wenn*

$$\liminf_{k\to\infty} f(x_k) \ge f(x_0)$$

für jede nach x_0 konvergierende Folge (x_k) gilt.

Beweis. "$\Rightarrow$" Es sei f in x_0 halbstetig von unten, also $f^{-1}((c,\infty])$ offen in $\mathbb{R}^n$ für alle $c \in \mathbb{R}$ mit $c < f(x_0)$, und (x_k) eine Folge mit Grenzwert x_0. Wählt man $c = f(x_0) - \varepsilon$, $\varepsilon > 0$ beliebig, gibt es somit eine Kugel $K(x_0, \delta(\epsilon))$, welche in $f^{-1}((c,\infty])$ enthalten ist. Wegen der Konvergenz der Folge (x_k) gilt ferner $x_k \in K(x_0, \delta(\epsilon))$ für $k \ge N(\delta(\epsilon))$ mit einem geeigneten $N(\delta(\epsilon))$. Somit folgt $f(x_k) > c = f(x_0) - \varepsilon$, $k \ge N(\delta(\epsilon))$, d.h.

$$\liminf_{k\to\infty} f(x_k) \ge f(x_0) - \varepsilon.$$

Da $\varepsilon > 0$ beliebig gewählt war, ergibt sich die Behauptung.
"$\Leftarrow$" Umgekehrt sei f in $x_0 \in \mathbb{R}^n$ nicht von unten halbstetig. Dann gibt es ein $c < f(x_0)$ derart, daß jede Umgebung $U \in \mathcal{U}(x_0)$ einen Punkt x_U enthält mit $f(x_U) \le c$. Also existiert eine Folge (x_k) so, daß $\lim_{k\to\infty} = x_0$ und $f(x_k) \le c$, $k \in \mathbb{N}$. Dies hat $\liminf_{k\to\infty} f(x_k) \le c < f(x_0)$ zur Folge. □

Lemma IX-1.5. *Eine von unten halbstetige Funktion nimmt auf einer kompakten Menge das Minimum an.*

Beweis. Es sei f von unten halbstetig, $K \subset \mathbb{R}^n$ kompakt und $(x_k) \subset K$ eine Folge mit $\lim_{k\to\infty} f(x_k) = \inf_{x\in K} f \in \bar{\mathbb{R}}$. Wegen der Kompaktheit von K gibt es $x_0 \in K$ und eine Teilfolge $(x_{\varphi_k}) \subset (x_k)$ mit $\lim_{k\to\infty} x_{\varphi_k} = x_0$. Mit Lemma IX-1.4 folgt $\inf_{x\in K} f(x) = \lim_{k\to\infty} f(x_{\varphi_k}) \geq f(x_0) > -\infty$. Dies zeigt $f(x_0) = \min_{x\in K} f(x)$ und $\min_{x\in K} f(x) > -\infty$. □

Definition IX-1.6. *Es sei $f_\alpha \colon \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$, $\alpha \in I$, I eine nichtleere Indexmenge, eine Familie von Funktionen. Man nennt*

$$f = \sup_{\alpha\in I} f_\alpha \colon \begin{cases} \mathbb{R}^n \to \mathbb{R} \cup \{\infty\} \\ x \mapsto \sup_{\alpha\in I} f_\alpha(x) \end{cases}$$

*die **obere Einhüllende (Enveloppe)** der Familie $(f_\alpha)_{\alpha\in I}$. .*

Analog definiert man für $f_\alpha \colon \mathbb{R}^n \to \mathbb{R} \cup \{-\infty\}$ die untere Einhüllende durch $f = \inf_{\alpha\in I}$.

Lemma IX-1.7. *Es sei $f_\alpha \colon \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$, $\alpha \in I$, eine Familie von in x_0 von unten halbstetigen Funktionen. Dann ist auch die obere Einhüllende von unten halbstetig in x_0.*

Beweis. Es sei $f = \sup_{\alpha\in I} f_\alpha$ und $c < f(x_0)$. Dann gibt es einen Index $\alpha \in I$ mit $c < f_\alpha(x_0) \leq f(x_0)$. Da f_α in x_0 halbstetig von unten ist, existiert eine Umgebung $U \in \mathcal{U}(x_0)$ mit $f_\alpha(x) > c$, somit auch $f(x) > c$, für alle $x \in U$. □

Beispiel IX-1.8. Es sei $f_n \colon [0,1] \to \mathbb{R}$, $n \in \mathbb{N}$, definiert durch

$$f_n(x) = \begin{cases} nx, & x \in [0, \frac{1}{n}], \\ 1, & x \in (\frac{1}{n}, 1]. \end{cases}$$

Die obere Einhüllende f ist gegeben durch

$$f_n(x) = \begin{cases} 0, & x = 0, \\ 1, & x \in (0, 1]. \end{cases}$$

Sie ist nicht stetig, nach Lemma IX-1.7 aber zumindest halbstetig von unten. Außerdem gilt $f_n \leq f_{n+1}$ und $f = \lim_{n\to\infty} f_n$.

Definition IX-1.9. *Es sei $f \colon \mathbb{R}^n \to \mathbb{R}$. Man nennt*

$$\operatorname{supp} f = \overline{\{x \in \mathbb{R}^n \colon f(x) \neq 0\}}$$

*den **Träger (Support)** von f. Weiters definieren wir*

$$C_c(\mathbb{R}^n) := \{f \in C(\mathbb{R}^n, \mathbb{R}) \colon \operatorname{supp} f \text{ ist kompakt } \}.$$

Man nimmt den Abschluß in der Definition des Trägers einer Funktion f, um sicherzustellen, daß es zu jedem $x \notin \operatorname{supp} f$ eine Kugel $K(x, \varepsilon)$ mit $K(x, \varepsilon) \subset \complement \operatorname{supp} f$ gibt. Somit gilt dann $f \equiv 0$ auf $K(x, \varepsilon)$. Isolierte Nullstellen und Häufungspunkte von

Nullstellen einer Funktion liegen in ihrem Träger. Bei der Analyse der qualitativen Eigenschaften einer Funktion kann man sich also ohne Beschränkung der Allgemeinheit auf deren Träger beschränken.

Einen wesentlichen Baustein bei der Entwicklung eines mehrdimensionalen Integrals bilden jene Funktionen, welche durch monotone Grenzwerte stetiger Funktionen mit kompaktem Träger darstellbar sind:

Definition IX-1.10. *Für eine Folge* $(f_k) \subset C_c(\mathbb{R}^n)$ *vereinbaren wir die Schreibweise*

$$f_k \uparrow f \Leftrightarrow \forall x \in \mathbb{R}^n \, \forall k \in \mathbb{N} f_k(x) \leq f_{k+1}(x) \wedge \lim_{k\to\infty} f_k(x) = f(x),$$
$$f_k \downarrow f \Leftrightarrow \forall x \in \mathbb{R}^n \, \forall k \in \mathbb{N} f_{k+1}(x) \leq f_k(x) \wedge \lim_{k\to\infty} f_k(x) = f(x).$$

Wir definieren ferner

$$\mathcal{H}_U(\mathbb{R}^n) = \{f \colon \mathbb{R}^n \to \mathbb{R} \cup \{\infty\} | \exists (f_k) \subset C_c(\mathbb{R}^n), f_k \uparrow f\},$$
$$\mathcal{H}_O(\mathbb{R}^n) = \{f \colon \mathbb{R}^n \to \mathbb{R} \cup \{-\infty\} | \exists (f_k) \subset C_c(\mathbb{R}^n), f_k \downarrow f\}.$$

Wenn keine Verwechslung möglich ist, verwenden wir die einfachere Schreibweise $\mathcal{H}_U$, bzw. $\mathcal{H}_O$. Die Funktionen in $\mathcal{H}_U$ sind die obere Einhüllende einer Folge stetiger Funktionen und daher von unten halbstetig, jene in $\mathcal{H}_O$ von oben halbstetig. Die Mengen $\mathcal{H}_U$ und $\mathcal{H}_O$ sind jedoch keine Teilräume. Eine alternative Charakterisierung ist folgende.

Theorem IX-1.11. *Eine Funktion* $f \colon \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ *gehört genau dann zu* $\mathcal{H}_U$, *wenn folgende Bedingungen erfüllt sind:*

(1) *f ist halbstetig von unten.*
(2) *Es gibt eine kompakte Teilmenge $K \subset \mathbb{R}^n$ mit*

$$f(x) \geq 0, \quad x \in \mathbb{R}^n \setminus K.$$

Beweis. "$\Rightarrow$" Es sei $f \in \mathcal{H}_U$ und $(f_k) \subset C_c(\mathbb{R}^n)$, $f_k \uparrow f$. Wähle $K = \operatorname{supp} f_1$.
"$\Leftarrow$" Die Abbildung f besitze die Eigenschaften (1) und (2). Nach Lemma IX-1.5 nimmt f auf K das Minimum an. Somit existiert $M \in \mathbb{Q}$, $M \geq 0$, mit

$$\forall x \in \mathbb{R}^n \colon f(x) \geq -M.$$

Wir betrachten nun die Menge

$$\mathcal{J} = \{(a, \varepsilon, c) \in \mathbb{Q}^n \times \mathbb{Q} \times \mathbb{Q} \colon \varepsilon > 0, c \geq -M, \forall x \in K(a,\varepsilon) \colon f(x) > c\}.$$

Offensichtlich ist $\mathcal{J}$ abzählbar. Für $j \in \mathcal{J}$ sei $g_j \in C_c(\mathbb{R}^n)$ eine Funktion mit folgenden Eigenschaften:

(1) $g_j(x) = c, \quad x \in K(a, \varepsilon/2)$,
(2) $g_j(x) \leq c, \quad x \in K(a, \varepsilon)$,
(3) $g_j(x) = -M, \quad x \in K \setminus K(a, \varepsilon)$,
(4) $g_j(x) \leq 0, \quad x \in \mathbb{R}^n \setminus (K \cup K(a, \varepsilon))$.

Es folgt $g_j \le f$, $j \in \mathcal{J}$ und somit $\sup_{j\in\mathcal{J}} g_j \le f$. Wir zeigen $\sup_{j\in\mathcal{J}} g_j = f$: Angenommen es gäbe $x_0 \in \mathbb{R}^n$, $c_0 \in \mathbb{Q}$, $c_0 \ge -M$ mit $\sup_{j\in\mathcal{J}} g_j(x_0) < c_0 < f(x_0)$. Da f in x_0 von unten halbstetig ist, gibt es eine Umgebung $K(x_0, r_0)$ derart, daß

$$(*) \qquad f(y) > c_0, \qquad y \in K(x_0, r_0).$$

Da $\mathbb{Q}^n$ in $\mathbb{R}^n$ dicht liegt (Beweis: übung), existiert eine Kugel $K(a_0, \varepsilon_0) \subset K(x_0, r_0)$ und $x_0 \in K(a_0, \varepsilon_0/2)$. Natürlich gilt $(*)$ auch auf $K(a_0, \varepsilon_0)$. Zu $j_0 = (a_0, \varepsilon_0, c_0)$ kann man eine Funktion g_{j_0} mit den Eigenschaften (1) - (4) konstruieren. Daraus ergibt sich der Widerspruch $g_{j_0}(x_0) = c_0 > \sup_{j\in\mathcal{J}} g_j(x_0)$.
Es sei nun $j_1, j_2, \ldots$ eine Aufzählung von $\mathcal{J}$. Setzt man $f_k = \max\{g_{j_1}, \ldots, g_{j_k}\} \in C_c(\mathbb{R}^n)$, gilt $f_k \le f_{k+1}$ und $\lim_{k\to\infty} f_k = f$, d.h. $f \in \mathcal{H}_U$. □

Als Konsequenz dieses Satzes halten wir fest

$$\mathcal{H}_U \cap \mathcal{H}_O = C_c(\mathbb{R}^n).$$

Bemerkung IX-1.12. Es sei $f \in \mathcal{H}_U$ und $f \ge 0$. Dann gibt es eine Folge $(f_k) \subset C_c(\mathbb{R}^n)$ mit $f_k \uparrow f$ und $f_k \ge 0$, $k \in \mathbb{N}$ (man ersetze nötigenfalls f_k durch $\max(0, f_k)$). Setzt man $\varphi_k = f_k - f_{k-1}$, $k \ge 2$, $\varphi_1 = f_1$, dann ist $\varphi_k \ge 0$ und $\varphi_k \in C_c(\mathbb{R}^n)$, $k \in \mathbb{N}$ und es gilt $f = \sum_{k=1}^{\infty} \varphi_k$.

2. Das Integral für halbstetige Funktionen

Unter einem achsenparallelen Quader versteht man die Menge $Q = X_{i=1}^n [a_i, b_i]$, $-\infty < a_i \le b_i < \infty$, $i = 1, \ldots, n$. Gilt $a_i = a$ und $b_i = b$, $i = 1, \ldots, n$ spricht man von einem Würfel. Durch Translation und Rotation eines achsenparallelen Quaders erhält man einen Quader in allgemeiner Lage. Setzt man $V(Q) = \prod_{i=1}^n (b_i - a_i)$, stimmt $V(Q)$ für $n = 2$ und $n = 3$ mit der Fläche eines Rechtecks bzw. mit dem üblichen Volumen eines Quaders im $\mathbb{R}^3$ überein. Es ist daher sinnvoll, die Größe $V(Q)$ als Volumen des Quaders Q zu bezeichnen. Wir werden später sehen, daß sich diese Vereinbarung aus der allgemeinen Definition des Volumens kompakter Körper zwanglos ergibt.

Definition IX-2.1. *Es sei $Q \subset \mathbb{R}^n$ ein Quader.*

(1) *Eine Menge von Quadern $\{Q_1, \ldots, Q_k\}$ heißt* ***Zerlegung (Partition)*** *von Q genau dann, wenn*
   (a) $Q_i \subset Q, \qquad i = 1, \ldots, k.$
   (b) $\bigcup_{i=1}^k Q_i = Q, \qquad Q_i^\circ \cap Q_j^\circ = \emptyset, \quad i \ne j$
(2) *Eine Abbildung $t\colon Q \to \mathbb{R}$ heißt* ***Treppenfunktion*** *wenn es eine Zerlegung $\{Q_1, \ldots, Q_k\}$ von Q gibt, so daß $f|_{Q_i^\circ}$ jeweils konstant ist, $i = 1, \ldots, k$.*
(3) $\mathcal{T}(Q, \mathbb{R}) := \{t\colon Q \to \mathbb{R}, t$ *ist eine Treppenfunktion*$\}$.

Eine Zerlegung in Teilwürfel entspricht einer äquidistanten Partition in Definition VII-VII-1.1. Die Verfeinerung einer Zerlegung kann wie im Fall $n = 1$ erklärt werden. Der Leser überzeuge sich davon, daß $\mathcal{T}(Q, \mathbb{R})$ versehen mit den natürlichen algebraischen Operationen ein Vektorraum ist.

Definition IX-2.2. *Es sei $t \in \mathcal{T}(Q, \mathbb{R})$ und $\{Q_1, \dots, Q_k\}$ eine passende Zerlegung des (nicht notwendig achsenparallelen) Quaders Q in Teilquader Q_i mit Volumen $V(Q_i)$ und $f|_{Q_i^\circ} = c_i$, $i = 1, \dots, k$. Das Integral von t ist definiert durch*

$$\int_Q t(x)\,dx = \sum_{i=1}^{k} c_i V(Q_i).$$

Wie im 1-dimensionalen Fall kann man sich davon überzeugen, daß das Integral einer Treppenfunktion unabhängig ist von der Zerlegung, welche für die Darstellung der Treppenfunktion verwendet wird.
Folgende Eigenschaften des Integrals ergeben sich unmittelbar aus der Definition.

Lemma IX-2.3. *Es seien f, $g \in \mathcal{T}(Q, \mathbb{R})$ und α, $\beta \in \mathbb{R}$ und $V(Q)$ das Volumen des Quaders Q. Ferner sei $A \in \mathbb{R}^{n \times n}$ eine orthogonale Matrix, $b \in \mathbb{R}^n$ und $\varphi\colon \tilde{Q} \to Q$ die lineare Abbildung $\varphi(x) = Ax + b$, $x \in \mathbb{R}^n$ des Quaders $\tilde{Q} = \varphi^{-1}(Q)$ auf Q. Dann gilt:*

1) $\alpha \int_Q f(x)\,dx + \beta \int_Q g(x)\,dx = \int_Q (\alpha f + \beta g)(x)\,dx,$ *Linearität*
2) $f \geq g \Rightarrow \int_Q f(x)\,dx \geq \int_Q g(x)\,dx,$ *Monotonie*
3) $|\int_Q f(x)\,dx| \leq \|f\|_\infty V(Q)$ *Beschränktheit*
4) $\int_{\tilde{Q}} (f \circ \varphi)(x)\,dx = \int_Q f(y)\,dy$ *Bewegungsinvarianz.*

Der Nachweis der Bewegungsinvarianz des Integrals ergibt sich aus der Beobachtung, daß φ und somit φ^{-1} eine Rotation, gefolgt von einer Translation beschreiben. Somit ist $\tilde{Q}$ ein zu Q kongruenter Quader. Eine Zerlegung $\{Q_1, \dots, Q_m\}$ von Q in Teilquader wird daher abgebildet auf eine Zerlegung des Quaders $\tilde{Q}$ in Teilquader $\{\varphi(Q_1), \dots, \varphi(Q_m)\}$ mit $V(\varphi(Q_i)) = V(Q_i)$, $i = 1, \dots, m$. Einer Treppenfunktion f auf Q entspricht daher die Treppenfunktion $f \circ \varphi$ auf $\tilde{Q}$.
Wir erweitern nun das Integral auf stetige Funktionen mit kompaktem Träger. Dazu benötigen wir folgendes Approximationsresultat, vgl. Satz VII-VII-1.2.

Theorem IX-2.4. *Es sei $f \in C_c(\mathbb{R}^n)$. Dann gibt es eine Folge von Treppenfunktionen (t_k), welche gleichmäßig gegen f konvergiert.*

Beweis. Der Träger von f sei in einem Quader Q enthalten. Es genügt, folgende Behauptung zu beweisen:

$$\forall \varepsilon > 0\, \exists t_\varepsilon \in \mathcal{T}(Q, \mathbb{R})\colon \|f - t_\varepsilon\|_\infty < \varepsilon.$$

Nach Satz III-5.4 ist f auf Q gleichmäßig stetig. Es sei $\varepsilon > 0$ beliebig gewählt. Dann gibt es ein $\delta > 0$, so daß

$$\forall x, y \in Q\colon \|x - y\| < \delta \Rightarrow |f(x) - f(y)| < \varepsilon.$$

Ferner sei $\{Q_1, \ldots, Q_k\}$ eine Zerlegung von Q in Teilquader derart, daß $\operatorname{diam} Q_i < \delta$, $i = 1, \ldots, k$. In jedem Teilquader Q_i wählen wir einen Punkt ξ_i und definieren die Treppenfunktion t_ε durch

$$t_\varepsilon(x) = \begin{cases} f(\xi_i), & x \in Q_i^\circ, \quad i = 1, \ldots, k, \\ f(\xi_{j^*}), & x \in Q_j \cap Q_l \text{ und } j^* \leq j < l, \\ 0 & \text{sonst.} \end{cases}$$

Der Index j^* ist z.B. charakerisiert durch die Bedingung

$$j^* = \min\{1 \leq j \leq n \colon \exists (l \in \{1, \ldots, n\})(j < l \wedge x \in Q_j \cap Q_l)\}$$

Diese Konstruktion stellt sicher, daß $|f(x) - t_\varepsilon(x)| < \varepsilon$ für alle $x \in Q$ gilt. □

Es sei nun $f \in C_c(\mathbb{R}^n)$ und (t_k) eine Folge von Treppenfunktionen, welche nach Satz IX-2.4 gleichmäßig gegen f konvergiert. Die Linearität und Beschränktheit des Integrals ergibt

$$\left| \int_Q t_k(x)\, dx - \int_Q t_l(x)\, dx \right| = \left| \int_Q (t_k - t_l)(x)\, dx \right| \leq \|t_k - t_l\|_\infty V(Q).$$

Somit ist $(\int_Q t_k(x)\, dx)_{k \geq 1}$ eine Cauchy Folge in $\mathbb{R}$ und daher konvergent. Wir zeigen, daß ihr Grenzwert von der Folge (t_k), welche zur Approximation von f verwendet wird, nicht abhängt. Es sei also (s_k) eine weitere Folge von Treppenfunktionen mit $s_k \rightrightarrows f$. Die Beschränktheit des Integrals ergibt wie vorhin

$$\left| \int_Q (s_k(x) - t_k(x))\, dx \right| \leq \|s_k - t_k\|_\infty V(Q),$$

also $\lim_{k \to \infty} \int_Q (s_k(x) - t_k(x))\, dx = 0$. Daraus folgt

$$\lim_{k \to \infty} \int_Q s_k(x)\, dx = \lim_{k \to \infty} \int_Q t_k(x)\, dx,$$

(die Existenz der Limiten wurde ja bereits nachgewiesen). Somit ist folgende Definition sinnvoll.

Definition IX-2.5. *Es sei $f \in C_c(\mathbb{R}^n)$, $\operatorname{supp} f \subset Q$, und (t_k) eine Folge von Treppenfunktionen, welche gleichmäßig gegen f konvergiert. Das Integral von f ist definiert durch*

$$\int_{\mathbb{R}^n} f(x)\, dx = \lim_{k \to \infty} \int_Q t_k(x)\, dx.$$

Für $n = 1$ stimmt dieses Integral mit dem Cauchy Integral überein. Die Eigenschaften des Integrals für Treppenfunktionen bleiben bei der Bildung des Grenzwertes erhalten.

Lemma IX-2.6. *Es seien f, $g \in C_c(\mathbb{R}^n)$ und α, $\beta \in \mathbb{R}$ und* $\operatorname{supp} f \subset Q$. *Ferner sei $A \in \mathbb{R}^{n\times n}$ orthogonal, $b \in \mathbb{R}^n$ und $\varphi\colon \mathbb{R}^n \to \mathbb{R}^n$ definiert durch $\varphi(x) = Ax + b$. Dann gilt:*

1) $\alpha \int_{\mathbb{R}^n} f(x)\,dx + \beta \int_{\mathbb{R}^n} g(x)\,dx = \int_{\mathbb{R}^n} (\alpha f + \beta g)(x)\,dx,$ *Linearität*
2) $f \geq g \Rightarrow \int_{\mathbb{R}^n} f(x)\,dx \geq \int_{\mathbb{R}^n} g(x)\,dx,$ *Monotonie*
3) $|\int_{\mathbb{R}^n} f(x)\,dx| \leq \int_{\mathbb{R}^n} |f(x)|\,dx \leq \|f\|_\infty V(Q)$ *Beschränktheit*
4) $\int_{\mathbb{R}^n} (f \circ \varphi)(x)\,dx = \int_{\mathbb{R}^n} f(y)\,dy$ *Bewegungsinvarianz.*

Beweis. Wir zeigen nur die Bewegungsinvarianz. Es sei $\tilde{Q} = \varphi^{-1}(Q)$. Dann ist $f \circ \varphi \in C_c(\mathbb{R}^n)$ und $\operatorname{supp} f \circ \varphi \subset \tilde{Q}$. Für eine Folge $(t_k) \subset \mathcal{T}(Q, \mathbb{R})$ mit $t_k \rightrightarrows f$ ergibt sich somit $t_k \circ \varphi \in \mathcal{T}(\tilde{Q}, \mathbb{R})$ und $t_k \circ \varphi \rightrightarrows f \circ \varphi$. Aus der Definition des Integrals und Lemma IX-2.3 folgt daher

$$\int_{\mathbb{R}^n} (f \circ \varphi)(x)\,dx = \lim_{k\to\infty} \int_{\tilde{Q}} (t_k \circ \varphi)(x)\,dx = \lim_{k\to\infty} \int_Q t_k(y)\,dy = \int_{\mathbb{R}^n} f(y)\,dy.$$

□

Als nächstes erweitern wir das Integral auf die Familien $\mathcal{H}_\mathcal{U}$ und $\mathcal{H}_\mathcal{O}$ halbstetiger Funktionen.

Definition IX-2.7. *Es sei $f \in \mathcal{H}_U$ und $(f_k) \subset C_c(\mathbb{R}^n)$ eine Funktionenfolge mit $f_k \uparrow f$. Das Integral von f ist definiert durch*

$$\int_{\mathbb{R}^n} f(x)\,dx = \lim_{k\to\infty} \int_{\mathbb{R}^n} f_k(x)\,dx \in \mathbb{R} \cup \{\infty\}.$$

Für $f \in \mathcal{H}_O$ setzen wir

$$\int_{\mathbb{R}^n} f(x)\,dx = -\int_{\mathbb{R}^n} (-f)(x)\,dx \in \mathbb{R} \cup \{-\infty\}.$$

Da die Folge $(\int_{\mathbb{R}^n} f_k(x)\,dx)$ monoton wächst, vgl. Lemma IX-2.6-2), existiert der Grenzwert in $\bar{\mathbb{R}}$. Wir zeigen nun, daß dieser Grenzwert unabhängig ist von der speziellen Folge $(f_k) \subset C_c(\mathbb{R}^n)$, die zur Approximation von f verwendet wird.

Lemma IX-2.8. *Es seien $f \in \mathcal{H}_U$ und (f_k), $(g_k) \subset C_c(\mathbb{R}^n)$ Folgen mit $f_k \uparrow f$ und $g_k \uparrow f$. Dann gilt*

$$\lim_{k\to\infty} \int_{\mathbb{R}^n} f_k(x)\,dx = \lim_{k\to\infty} \int_{\mathbb{R}^n} g_k(x)\,dx.$$

Beweis. Wir zeigen: Für festes $k \in \mathbb{N}$ gilt

$$\int_{\mathbb{R}^n} f_k(x)\,dx \leq \lim_{l\to\infty} \int_{\mathbb{R}^n} g_l(x)\,dx.$$

Aus Symmetriegründen ergibt sich daraus die Behauptung. Wir betrachten die Funktionen $h_l = \min(f_k, g_l)$. Wegen $f_k \le f$ und $g_l \uparrow f$ folgt $h_l \uparrow f_k$. Die Funktionen h_l sind stetig und haben kompakten Träger

$$\operatorname{supp} h_l \subset \operatorname{supp} h_1 \cup \operatorname{supp} f_k, \quad l \in \mathbb{N}.$$

Dies sieht man folgendermaßen ein: Wählt man $x \in \complement \operatorname{supp} h_1 \cap \complement \operatorname{supp} f_k$, existiert eine Umgebung $U \in \mathcal{U}(x)$ mit $f_k = h_1 \equiv 0$ auf U. Wegen $h_1 \le h_l \le f_k$ folgt daraus $h_l \equiv 0$ auf U und somit $x \notin \operatorname{supp} h_l$. Nach dem Satz von Dini **??** ist die Konvergenz daher gleichmäßig. Es sei nun Q ein Quader mit $\operatorname{supp} h_l \subset Q$. Mit Hilfe von Lemma IX-2.6 erhält man

$$\left| \int_{\mathbb{R}^n} (f_k(x) - h_l(x))\, dx \right| \le \|f_k - h_l\|_\infty V(Q),$$

also

$$\int_{\mathbb{R}^n} f_k(x)\, dx = \lim_{l\to\infty} \int_{\mathbb{R}^n} h_l(x)\, dx \le \lim_{l\to\infty} \int_{\mathbb{R}^n} g_l(x)\, dx,$$

wobei wir $h_l \le g_l$ und wiederum Lemma IX-2.6 verwendet haben. □

Dieses Lemma zeigt, daß die Definition des Integrals für Funktionen in $\mathcal{H}_U$ sinnvoll ist. Setzt man insbesondere eine Funktion $f \in C_c(\mathbb{R}^n)$ ein, kann man $f_k = f$ wählen, so daß die Integrale gemäß Definition IX-2.5 und Definition IX-2.7 übereinstimmen. Da die Funktionen in $\mathcal{H}_U$ nicht beschränkt sein müssen, ihr Träger nicht unbedingt kompakt ist, macht die Beschränktheit des Integrals keinen Sinn mehr. Die übrigen Eigenschaften des Integrals können jedoch übertragen werden.

Lemma IX-2.9. *Es seien f, $g \in \mathcal{H}_U$ und α, $\beta \ge 0$. Ferner sei $A \in \mathbb{R}^{n\times n}$ orthogonal, $b \in \mathbb{R}^n$ und $\varphi\colon \mathbb{R}^n \to \mathbb{R}^n$ definiert durch $\varphi(x) = Ax + b$. Dann gehören auch die Funktionen $\alpha f + \beta g$, sowie $f \circ \varphi$ zu $\mathcal{H}_U$ und es gilt*

1) $\alpha \int_{\mathbb{R}^n} f(x)\, dx + \beta \int_{\mathbb{R}^n} g(x)\, dx = \int_{\mathbb{R}^n} (\alpha f + \beta g)(x)\, dx,$
2) $f \ge g \Rightarrow \int_{\mathbb{R}^n} f(x)\, dx \ge \int_{\mathbb{R}^n} g(x)\, dx,$ *Monotonie*
3) $\int_{\mathbb{R}^n} (f \circ \varphi)(x)\, dx = \int_{\mathbb{R}^n} f(y)\, dy$ *Bewegungsinvarianz.*

Beweis. Es seien (f_k), $(g_k) \subset C_c(\mathbb{R}^n)$ Funktionenfolgen mit $f_k \uparrow f$ und $g_k \uparrow g$. Dann gilt auch $\alpha f_k + \beta g_k \uparrow \alpha f + \beta g$ für α, $\beta \ge 0$ und somit $\alpha f + \beta g \in \mathcal{H}_U$. Aus der Definition des Integrals in $\mathcal{H}_U$ folgt dann

$$\begin{aligned}\int_{\mathbb{R}^n} (\alpha f + \beta g)(x)\, dx &= \lim_{k\to\infty} \int_{\mathbb{R}^n} (\alpha f_k + \beta g_k)(x)\, dx \\ &= \alpha \lim_{k\to\infty} \int_{\mathbb{R}^n} f_k(x)\, dx + \beta \int_{\mathbb{R}^n} g_k(x)\, dx = \alpha \int_{\mathbb{R}^n} f(x)\, dx + \beta \int_{\mathbb{R}^n} g(x)\, dx.\end{aligned}$$

Die restlichen Aussagen beweist man analog. □

Bemerkung IX-2.10. Das Integral in $\mathcal{H}_U$ ist *nicht* linear, da $\mathcal{H}_U$ kein Vektorraum ist und daher 1) nur für α, $\beta \geq 0$ gilt. Allerdings gilt für $\alpha \in \mathbb{R}$ und $f \in \mathcal{H}_U$

$$\alpha \int_{\mathbb{R}^n} f(x)\,dx = \int_{\mathbb{R}^n} (\alpha f)(x)\,dx.$$

Für $\alpha < 0$ folgt nämlich $\alpha f \in \mathcal{H}_O$ und

$$\int_{\mathbb{R}^n} (\alpha f)(x)\,dx = -\int_{\mathbb{R}^n} -(\alpha f)(x)\,dx = -\int_{\mathbb{R}^n} (-\alpha) f(x)\,dx = -(-\alpha)\int_{\mathbb{R}^n} f(x)\,dx.$$

Die erste Gleichheit ergibt sich aus der Definition des Integrals in $\mathcal{H}_O$, die letzte folgt aus 1).

Lemma IX-2.11. *Es sei* $(f_k) \subset \mathcal{H}_U$ *und* $f_k \geq 0$, $k \in \mathbb{N}_0$, *und* $f = \sum_{k=0}^{\infty} f_k$. *Dann ist* $f \in \mathcal{H}_U$ *und es gilt*

$$\int_{\mathbb{R}^n} f(x)\,dx = \sum_{k=0}^{\infty} \int_{\mathbb{R}^n} f_k(x)\,dx.$$

Beweis. Wegen Bemerkung IX-1.12 gibt es Funktionen $\varphi_{kl} \in C_c(\mathbb{R}^n)$, $\varphi_{kl} \geq 0$, k, $l \in \mathbb{N}_0$ mit $f_k = \sum_{l=0}^{\infty} \varphi_{kl}$. Somit gilt $f = \sum_{k=0}^{\infty} \sum_{l=0}^{\infty} \varphi_{kl}$. Für jedes $x \in \mathbb{R}^n$ ist die iterierte Reihe (und somit auch die entsprechende Doppelreihe) (absolut) konvergent oder divergent. Dann ist jede Anordnung der Doppelreihe in eine einfache Reihe konvergent (mit gleichem Grenzwert) oder divergent. Somit gilt insbesondere

$$f = \sum_{k=0}^{\infty} \sum_{l=0}^{k} \varphi_{l,k-l} = \sum_{k=0}^{\infty} h_k$$

mit

$$h_k = \sum_{l=0}^{k} \varphi_{l,k-l} \in C_c(\mathbb{R}^n), \quad h_k \geq 0, k \in \mathbb{N}_0.$$

Für die Partialsummen $s_m = \sum_{k=0}^{m} h_k$ folgt dann $s_m \in C_c(\mathbb{R}^n)$, $s_m \uparrow f$, $m \in \mathbb{N}_0$, und somit $f \in \mathcal{H}_U$. Aus der Abschätzung

$$s_m = \sum_{k=0}^{m} h_k \leq \sum_{k=0}^{m} f_k,$$

der Definition des Integrals in $\mathcal{H}_U$ und Lemma IX-2.9 ergibt sich daher

$$\begin{aligned}\int_{\mathbb{R}^n} f(x)\,dx &= \lim_{m\to\infty} \int_{\mathbb{R}^n} s_m(x)\,dx \leq \lim_{m\to\infty} \int_{\mathbb{R}^n} \sum_{k=0}^{m} f_k(x)\,dx \\ &= \lim_{m\to\infty} \sum_{k=0}^{m} \int_{\mathbb{R}^n} f_k(x)\,dx = \sum_{k=0}^{\infty} \int_{\mathbb{R}^n} f_k(x)\,dx\end{aligned}$$

also

$$\int_{\mathbb{R}^n} f(x)\,dx \leq \sum_{k=0}^{\infty} \int_{\mathbb{R}^n} f_k(x)\,dx.$$

Die Abschätzung

$$\sum_{k=0}^{m} f_k \leq f, \quad m \in \mathbb{N}_0$$

ergibt wieder mit Lemma IX-2.9

$$\sum_{k=0}^{m} \int_{\mathbb{R}^n} f_k(x)\,dx = \int_{\mathbb{R}^n} \sum_{k=0}^{m} f_k(x)\,dx \leq \int_{\mathbb{R}^n} f(x)\,dx$$

also auch

$$\sum_{k=0}^{\infty} \int_{\mathbb{R}^n} f_k(x)\,dx \leq \int_{\mathbb{R}^n} f(x)\,dx$$

□

Wir kommen nun zu einem zentralen Satz der Integrationstheorie.

THEOREM IX-2.12 (Fubini). *Es seien $f \in \mathcal{H}_U(\mathbb{R}^n)$, $(i_1, \ldots, i_n)$ eine Permutation von $(1, \ldots, n)$ und $1 \leq k < n$. Für $x = (x_1, \ldots, x_n)$ setzen wir $\xi = (x_{i_1}, \ldots, x_{i_k}) \in \mathbb{R}^k$, $\eta = (x_{i_{k+1}}, \ldots, x_{i_n}) \in \mathbb{R}^{n-k}$ und*

$$\tilde{f}(\xi, \eta) = \tilde{f}(x_{i_1}, \ldots, x_{i_k}, x_{i_{k+1}}, \ldots, x_{i_n}) = f(x_1, \ldots, x_n).$$

Dann gilt: Für jedes feste $\eta \in \mathbb{R}^{n-k}$ gehört die Funktion

$$\xi \mapsto \tilde{f}(\xi, \eta)$$

zu $\mathcal{H}_U(\mathbb{R}^k)$. Es existiert das Integral

$$F(\eta) := \int_{\mathbb{R}^k} \tilde{f}(\xi, \eta)\,d\xi = \int_{\mathbb{R}^k} f(x_1, \ldots, x_n)\,dx_{i_1} \ldots dx_{i_k},$$

es ist $F \in \mathcal{H}_U(\mathbb{R}^{n-k})$ und es gilt

$$(\dagger) \qquad \int_{\mathbb{R}^{n-k}} F(\eta)\,d\eta = \int_{\mathbb{R}^{n-k}} [\int_{\mathbb{R}^k} \tilde{f}(\xi, \eta)\,d\xi]\,d\eta = \int_{\mathbb{R}^n} f(x)\,dx.$$

BEWEIS. Wir gehen vorerst von der identischen Permutation aus. Es sei also $\xi = (x_1, \ldots, x_k)$ und $\eta = (x_{k+1}, \ldots, x_n)$. Wir zeigen, daß

$$(\ddagger) \qquad \int_{\mathbb{R}^{n-k}} F(\eta)\,d\eta = \int_{\mathbb{R}^{n-k}} [\int_{\mathbb{R}^k} f(\xi, \eta)\,d\xi]\,d\eta = \int_{\mathbb{R}^n} f(x)\,dx$$

für jede Funktion $f \in \mathcal{H}_U(\mathbb{R}^n)$ gilt. Wir führen den Beweis in mehreren Schritten.
Schritt 1: Es sei χ_Q die charakteristische Funktion eines achsenparallelen Quaders

$Q = X_{i=1}^m [a_i, b_i]$. Zerlegt man Q in $Q = Xi = 1^k [a_i, b_i] \times X_{j=k+1}^m [a_j, b_j] \equiv Q' \times Q''$ folgt $\chi_Q(\xi, \eta) = \chi_{Q'}(\xi)\chi_{Q''}(\eta)$. Somit folgt

$$\begin{aligned}\int_{\mathbb{R}^{n-k}} [\int_{\mathbb{R}^k} \chi_Q(\xi, \eta)\, d\xi]\, d\eta &= \int_{\mathbb{R}^{n-k}} \chi_{Q''}(\eta)[\int_{\mathbb{R}^k} \chi_{Q'}(\xi)\, d\xi]\, d\eta \\ &= V_k(Q')V_{n-k}(Q'') = V_n(Q) = \int_{\mathbb{R}^n} \chi_Q(x)\, dx.\end{aligned}$$

Jede Treppenfunktion $t \in \mathcal{T}(Q, \mathbb{R})$ läßt sich mit einer geeigneten Zerlegung von Q in Quader $\{Q_1, \dots, Q_l\}$ darstellen in der Form

$$t = \sum_{i=1}^{l} c_i \chi_{Q_i},$$

wobei $c_i \in \mathbb{R}$, $i = 1, \dots, l$.
Wegen der Linearität des Integrals ergibt sich daher

$$(*) \qquad \int_{\mathbb{R}^{n-k}} [\int_{\mathbb{R}^k} t(\xi, \eta)\, d\xi]\, d\eta = \int_{\mathbb{R}^n} t(x)\, dx,$$

für alle $t \in \mathcal{T}(Q, \mathbb{R})$.
Schritt 2: Wir zeigen nun die Gültigkeit des Satzes für $f \in C_c(\mathbb{R}^n)$. Wir schließen den Träger von f in einen achsenparallelen Quader Q ein und konstruieren nach Satz IX-2.4 zu einer Folge von Zerlegungen $\{Q_{1l}, \dots, Q_{m_l l}\}$ von Q, $l \in \mathbb{N}$, Treppenfunktionen $t_l = \sum_{i=1}^{m_l} c_{il}\chi_{Q_{il}}$ mit $t_l \rightrightarrows f$. Wie vorhin setzen wir $Q = Q' \times Q'' \subset \mathbb{R}^k \times \mathbb{R}^{n-k}$ und $Q_{il} = Q'_{il} \times Q''_{il} \subset Q' \times Q''$. Für $\eta \in \mathbb{R}^{n-k}$ betrachten wir die Schnittfunktionen t_l^η, $f^\eta \colon \mathbb{R}^k \to \mathbb{R}$

$$t_l^\eta(\xi) = t_l(\xi, \eta) \qquad \text{und} \qquad f^\eta(\xi) = f(\xi, \eta), \quad \xi \in Q'.$$

Auch die Funktionen $t_l^\eta = \sum_{i=1}^{m_l} c_{il}\chi_{Q''_{il}}(\eta)\chi_{Q'_{il}}$, $l \in \mathbb{N}$, sind Treppenfunktionen auf Q' und es gilt $t_l^\eta \rightrightarrows f^\eta \in C_c(\mathbb{R}^k)$ (sogar gleichmäßig in $\eta \in Q''$). Dann sind auch die Funktionen

$$T_l(\eta) = \int_{\mathbb{R}^k} t_l^\eta(\xi)\, d\xi = \sum_{i=1}^{m_l} c_{il} V(Q'_{il})\chi_{Q''_{il}}(\eta) = \int_{\mathbb{R}^k} t_l(\xi, \eta)\, d\xi$$

Treppenfunktionen auf Q''. Wegen der Definition des Integrals in $C_c(\mathbb{R}^k)$ folgt daher

$$F(\eta) = \int_{\mathbb{R}^k} f^\eta(\xi)\, d\xi = \lim_{l\to\infty} \int_{\mathbb{R}^k} t_l^\eta(\xi)\, d\xi = \lim_{l\to\infty} T_l(\eta)$$

Dieser Grenzwert existiert sogar gleichmäßig in η. Um dies einzusehen, betrachten wir vorerst

$$|T_l(\eta) - T_k(\eta)| = |\int_{\mathbb{R}^k} (t_l^\eta(\xi) - t_k^\eta(\xi))\, d\xi| \le \|t_l^\eta - t_k^\eta\|_\infty V(Q').$$

Wegen $t_l^\eta \rightrightarrows f^\eta$ ist die Folge $(T_l(\eta))$ eine in η gleichmäßige Cauchy Folge. Bildet man den Grenzwert $k \to \infty$ erhält man $T_l \rightrightarrows F$. Aus der analogen Abschätzung

$$|F(\eta) - F(\tilde{\eta})| = |\int_{\mathbb{R}^k} (f^\eta(\xi) - f^{\tilde{\eta}}(\xi))\, d\xi| \le \max_{\xi\in\mathbb{R}^k} |f(\xi,\eta) - f(\xi,\tilde{\eta})| V(Q')$$

erkennt man wegen der gleichmäßigen Stetigkeit von f, daß auch F stetig (sogar gleichmäßig stetig) ist. Somit liegt F in $C(\mathbb{R}^{n-k})$ und wegen $\operatorname{supp} F \subset Q''$ auch in $C_c(\mathbb{R}^{n-k})$. Wegen der Definition des Integrals in $C_c(\mathbb{R}^{n-k}$ und $T_l \rightrightarrows F$ erhält man daher

$$(**)\qquad \begin{aligned} \int_{\mathbb{R}^{n-k}} F(\eta)\, d\eta &= \lim_{l\to\infty} \int_{\mathbb{R}^{n-k}} T_l(\eta)\, d\eta = \lim_{l\to\infty} \int_{\mathbb{R}^{n-k}} [\int_{\mathbb{R}^k} t_l(\xi,\eta)\, d\xi]\, d\eta \\ &\overset{(*)}{=} \lim_{l\to\infty} \int_{\mathbb{R}^n} t_l(x)\, dx = \int_{\mathbb{R}^n} f(x)\, dx. \end{aligned}$$

Die mit "$(*)$" markierte Gleichheit folgt aus Schritt 1.
Schritt 3: Es sei nun $f \in \mathcal{H}_U$ und $(f_l) \subset C_c(\mathbb{R}^n)$ eine Folge von Funktionen mit $f_l \uparrow f$. Definiert man für festes $\eta \in \mathbb{R}^{n-k}$ wie vorhin die Funktionen f_l^η und f^η, ergibt sich $f_l^\eta \in C_c(\mathbb{R}^k, \mathbb{R})$ und $f_l^\eta \uparrow f^\eta$. Somit gilt $f^\eta \in \mathcal{H}_U(\mathbb{R}^k)$ und es existiert das Integral

$$F(\eta) = \int_{\mathbb{R}^k} f^\eta(\xi)\, d\xi = \int_{\mathbb{R}^k} f(\xi,\eta)\, d\xi.$$

Setzt man wie in Schritt 2

$$F_l(\eta) = \int_{\mathbb{R}^k} f_l^\eta(\xi)\, d\xi = \int_{\mathbb{R}^k} f_l(\xi,\eta)\, d\xi,$$

folgt aus der Definition des Integrals in $\mathcal{H}_U(\mathbb{R}^k)$

$$F(\eta) = \lim_{l\to\infty} \int_{\mathbb{R}^k} f_l^\eta(\xi)\, d\xi = \lim_{l\to\infty} F_l(\eta).$$

Es sei nun Q_l ein achsenparalleler Quader mit $\operatorname{supp} f_l \subset Q_l = Q_l' \times Q_l''$. Dann folgert man aus der Abschätzung

$$|F_l(\eta) - F_l(\tilde{\eta})| = |\int_{\mathbb{R}^k} (f_l^\eta(\xi) - f_l^{\tilde{\eta}}(\xi))\, d\xi| \le \max_{\xi\in\mathbb{R}^k} |f_l(\xi,\eta) - f_l(\xi,\tilde{\eta})| V(Q_l')$$

wie vorhin $F_l \in C_c(\mathbb{R}^{n-k})$. Da aber auch $F_l \le F_{l+1}$ gilt, folgt $F \in \mathcal{H}_U(\mathbb{R}^{n-k})$. Eine zweifache Anwendung der Definition des Integrals in $\mathcal{H}_U$ ergibt nun

$$\begin{aligned} \int_{\mathbb{R}^{n-k}} F(\eta)\, d\eta &= \lim_{l\to\infty} \int_{\mathbb{R}^{n-k}} F_l(\eta)\, d\eta = \lim_{l\to\infty} \int_{\mathbb{R}^{n-k}} [\int_{\mathbb{R}^k} f_l(\xi,\eta)\, d\xi]\, d\eta \\ &\overset{(**)}{=} \lim_{l\to\infty} \int_{\mathbb{R}^n} f_l(x)\, dx = \int_{\mathbb{R}^n} f(x)\, dx. \end{aligned}$$

Die mit "$(**)$" markierte Gleichheit wurde in Schritt 2 bewiesen.
Schritt 4: Es sei nun $\tilde{x} = (x_{i_1}, \dots, x_{i_n})$ eine beliebige Permutation der Variablen $(x_1, \dots, x_n)$, $\xi = (x_{i_1}, \dots, x_{i_k})$ und $\eta = (x_{i_{k+1}}, \dots, x_{i_n})$. Die orthogonale Matrix

$A \in \mathbb{R}^{n\times n}$ entstehe aus der Einheitsmatrix durch der Permutation $(i_1,\dots,i_n)$ entsprechende Vertauschungen der Spalten. Dann ist $\tilde{x} = Ax$ und $\tilde{f} = f \circ A^{-1}$. Aus Lemma IX-2.3 folgt $\tilde{f} \in \mathcal{H}_U(\mathbb{R}^n)$. Aus dem bisher Gezeigten erhält man

$$\int_{\mathbb{R}^{n-k}} [\int_{\mathbb{R}^k} \tilde{f}(\xi,\eta)\,d\xi]\,d\eta = \int_{\mathbb{R}^n} \tilde{f}(\tilde{x})\,d\tilde{x} = \int_{\mathbb{R}^n} \tilde{f} \circ A^{-1}(\tilde{x})\,d\tilde{x} = \int_{\mathbb{R}^n} f(x)\,dx.$$

Für das letzte Gleichheitszeichen verwende man wieder Lemma IX-2.3. □

Der Satz von Fubini beinhaltet zwei Aussagen: zum einen zeigt er auf, daß man ein n-dimensionales Integral zurückführen kann auf **iterierte Integrale** kleinerer Dimension. Eine wiederholte Anwendung dieses Argumentes reduziert die Berechnung eines n-dimensionalen Integrales auf die Auswertung von n eindimensionalen Integralen, nämlich

$$\int_{\mathbb{R}^n} f(x)\,dx = \int_{\mathbb{R}} \dots [\int_{\mathbb{R}} f(x_1,\dots,x_n)\,dx_1]\dots dx_n.$$

Zum anderen besagt der Satz, daß die Gruppierung der Integrationsvariablen und somit die Reihenfolge der iterierten Integrale belanglos ist. Im Fall $n = 2$ ergibt die spezielle Wahl der Permutationen $(i_1,i_2) = (1,2)$ bzw. $(i_1,i_2) = (2,1)$

$$\int_{\mathbb{R}^2} f(\xi,\eta)\,d\xi d\eta = \int_{\mathbb{R}} [\int_{\mathbb{R}} f(\xi,\eta)\,d\xi]d\eta = \int_{\mathbb{R}} [\int_{\mathbb{R}} f(\xi,\eta)\,d\eta]d\xi.$$

Korollar IX-2.13. Es seien $\varphi \in \mathcal{H}_U(\mathbb{R}^k,\mathbb{R}_+)$ und $\psi \in \mathcal{H}_U(\mathbb{R}^{n-k},\mathbb{R}_+)$. Dann liegt die Abbildung $f\colon (\xi,\eta) \to f(\xi,\eta) = \varphi(\xi)\psi(\eta)$ in $\mathcal{H}_U(\mathbb{R}^n,\mathbb{R}_+)$ und es gilt mit $x = (\xi,\eta)$

$$\int_{\mathbb{R}^n} f(x)\,dx = \int_{\mathbb{R}^k} \varphi(\xi)\,d\xi \int_{\mathbb{R}^{n-k}} \psi(\eta)\,d\eta.$$

Ein weiteres nützliches Werkzeug in der Integrationstheorie ist die Variablentransformation. Wir betrachten vorerst einen einfachen Spezialfall.

Theorem IX-2.14. *Es seien $f \in \mathcal{H}_U$, $A \in \mathbb{R}^{n\times n}$ regulär und $b \in \mathbb{R}^n$. Dann liegt die Abbildung $x \mapsto f(Ax+b)$ in $\mathcal{H}_U$ und es gilt*

$$\int_{\mathbb{R}^n} f(Ax+b)\,dx = \frac{1}{|\det A|} \int_{\mathbb{R}^n} f(x)\,dx.$$

Der Beweis dieses Satzes stützt sich auf die Singulärwertzerlegung einer Matrix.

Theorem IX-2.15. *Zu jeder Matrix $A \in \mathbb{R}^{n\times m}$ mit Rang $r \le \min(m,n)$ gibt es orthogonale Matrizen $U \in \mathbb{R}^{n\times n}$ und $V \in \mathbb{R}^{m\times m}$ und eine Diagonalmatrix $D \in \mathbb{R}^{n\times m}$, $D = diag\,(\sigma_1,\dots,\sigma_r,0,\dots,0)$, $\sigma_i > 0$, $i = 1,\dots,r$, so daß*

$$A = UDV.$$

Beweis des Satzes IX-2.14. Es sei $(f_k) \subset C_c(\mathbb{R}^n)$ eine Folge von Funktionen mit $f_k \uparrow f$. Da der Träger von $x \mapsto f_k(Ax+b)$ kompakt ist und $f_k(Ax+b) \uparrow f(Ax+b)$ für alle $x \in \mathbb{R}^n$ gilt, folgt $x \mapsto f(Ax+b) \in \mathcal{H}_U$. Es sei $A = UDV$ die Singulärwertzerlegung von A mit einer Diagonalmatrix $D = \text{diag}\,(\sigma_1,\dots,\sigma_n)$, $\sigma_i > 0$,

$i = 1, \ldots, n$. Wegen der Invarianz des Integrals unter Translationen und orthogonalen Transformationen folgt

$$\int_{\mathbb{R}^n} f_k(Ax + b)\, dx = \int_{\mathbb{R}^n} f(UDVx)\, dx = \int_{\mathbb{R}^n} f_k(UDx)\, dx.$$

Mit Hilfe des Satzes von Fubini schreiben wir das letzte Integral als n-fach iteriertes Integral und substituieren in jedem der eindimensionalen (Cauchy-)Integrale $y_i = \sigma_i x_i$, $i = 1, \ldots, n$. Auf diese Weise erhalten wir

$$\begin{aligned}\int_{\mathbb{R}^n} f_k(UDx)\, dx &= \int_{\mathbb{R}} \ldots [\int_{\mathbb{R}} f_k(U(\sigma_1 x_1, \ldots, \sigma_n x_n)^T) dx_1] \ldots dx_n \\ &= \frac{1}{\sigma_1 \cdot \ldots \cdot \sigma_n} \int_{\mathbb{R}} \ldots [\int_{\mathbb{R}} f_k(U(y_1, \ldots, y_n)^T) dy_1] \ldots dy_n \\ &= \frac{1}{\det D} \int_{\mathbb{R}^n} f_k(Uy)\, dy = \frac{1}{\det D} \int_{\mathbb{R}^n} f_k(y)\, dy.\end{aligned}$$

In den letzten beiden Schritten wurde der Satz von Fubini und anschließend wieder Lemma IX-2.6 verwendet. Da der Betrag der Determinante einer orthogonalen Matrix gleich 1 ist, folgt

$$|\det A| = |\det U||\det D||\det V| = \det D > 0.$$

Eine neuerliche Anwendung des Satzes von Fubini und der Grenzübergang $k \to \infty$ beenden den Beweis. $\square$

In vielen Anwendungen will man eine Funktion f nicht über $\mathbb{R}^n$, sondern über eine kompakte Teilmenge K integrieren. Dies kann man in die bisher entwickelte Theorie einbauen, wenn man zusätzlich $f\chi_K \in \mathcal{H}_O$ fordert (wir erinnern daran, daß die charakteristische Funktion einer kompakten Teilmenge in $\mathcal{H}_O$ liegt, vgl. Satz IX-1.11 und Beispiel IX-1.3). Wir vereinbaren folgende Schreibweise.

Definition IX-2.16. *Es sei $K \subset \mathbb{R}^n$ kompakt und $f\colon \mathbb{R}^n \to \mathbb{R} \cup \{-\infty\}$ derart, daß $f\chi_K \in \mathcal{H}_O$. Wir setzen*

$$\int_K f(x)\, dx := \int_{\mathbb{R}^n} f(x)\chi_K(x)\, dx.$$

Ist K das Intervall $[a, b]$ in $\mathbb{R}$, schreiben wir auch

$$\int_K f(x)\, dx = \int_a^b f(x)\, dx.$$

3. Berechnung elementarer Volumina

In diesem Abschnitt wenden wir die Integralrechnung auf die Berechnung der Volumina von elementaren geometrischen Körpern an.

DEFINITION IX-3.1. *Das* ***Volumen*** $V(K)$ *einer kompakten Teilmenge* K *in* $\mathbb{R}^n$ *ist definiert durch*

$$V(K) = \int_{\mathbb{R}^n} \chi_K(x)\,dx.$$

Im Fall $n = 2$ *nennt man* $V(K)$ ***Flächeninhalt*** *von* K*, im Fall* $n = 1$ *ergibt* $V(K)$ *die* ***Länge*** *von* K*.*

Da es zu einer kompakte Menge K stets einen Würfel $W = [-M, M]^n$ mit $K \subset W$ gibt, folgt aus $\chi_K \leq \chi_W$, daß eine kompakte Menge endliches Volumen besitzt. Gelegentlich verwenden wir die Schreibweise $V_n(K)$, um anzudeuten, daß das Volumen von K in $\mathbb{R}^n$ gemeint ist. Wir ziehen zuerst eine Folgerung aus dem Satz von Fubini.

THEOREM IX-3.2. *Es sei* $1 \leq k < n$ *und* $K_1 \subset \mathbb{R}^k$*,* $K_2 \subset \mathbb{R}^{n-k}$ *seien kompakte Teilmengen. Dann gilt*

$$V_n(K_1 \times K_2) = V_k(K_1)V_{n-k}(K_2).$$

BEWEIS. Schreibt man $x = (x_1, \ldots, x_n)$, $\xi = (x_1, \ldots, x_k)$ und $\eta = (x_{k+1}, \ldots, x_n)$, gilt

$$\chi_{K_1 \times K_2}(\xi, \eta) = \chi_{K_1}(\xi)\chi_{K_2}(\eta).$$

Die Behauptung folgt nun aus Korollar IX-2.13. □

BEISPIEL IX-3.3. 1) Intervall, $K = [a, b]$, $V(K) = \int_a^b dx = b - a$.
2) Parallelepiped $K = X_{i=1}^n [a_i, b_i]$, $V(K) = \prod_{i=1}^n (b_i - a_i)$.
3) gerader Zylinder mit Basis B und Höhe h, $K = B \times [0, h]$, $V(K) = V(B)h$.

THEOREM IX-3.4. *Es sei* $\varphi\colon \mathbb{R}^n \to \mathbb{R}^n$ *die affine Abbildung* $x \mapsto Ax + b$ *mit* $A \in \mathbb{R}^{n\times n}$ *und* $b \in \mathbb{R}^n$*. Dann gilt für jede kompakte Teilmenge* $K \subset \mathbb{R}^n$

$$V(\varphi(K)) = |\det A| V(K).$$

Insbesondere ist das Volumen unter längentreuen Abbildungen invariant (diese werden durch orthogonale Matrizen beschrieben).

BEWEIS. Ist $\det A \neq 0$, folgt die Behauptung aus

$$\chi_{\varphi(K)} = \chi_K \circ \varphi^{-1}$$

und Satz IX-2.14

$$\begin{aligned} V(\varphi(K)) &= \int_{\mathbb{R}^n} \chi_{\varphi(K)}(y)\,dy = \int_{\mathbb{R}^n} \chi_K(A^{-1}y - A^{-1}b)\,dy \\ &= |\frac{1}{\det(A^{-1})}| \int_{\mathbb{R}^n} \chi_K(x)\,dx = |\det A| V(K). \end{aligned}$$

Falls aber $\det A = 0$, ist $\varphi(K)$ in einer Hyperebene enthalten, die man nach einer orthogonalen Koordinatentransformation als $\{(x_1, \ldots, x_n) \in \mathbb{R}^n : x_n = 0\}$ annehmen darf. Somit ist $\varphi(K)$ ein Zylinder mit der Höhe Null und hat somit nach dem vorigen Beispiel das Volumen null. □

Theorem IX-3.5. *Es sei $K \subset \mathbb{R}^n$ kompakt, $r \geq 0$ und $rK = \{rx \colon x \in K\}$. Dann gilt*

$$V(rK) = r^n V(K).$$

Beweis. Für $r = 0$ ist die Behauptung klar. Für $r > 0$ überzeuge man sich von der Gültigkeit der Beziehung

$$\chi_{rK}(x) = \chi_K(\frac{1}{r}x), \quad x \in \mathbb{R}^n.$$

Mit Satz IX-2.14 folgt

$$V(rK) = \int_{\mathbb{R}^n} \chi_{rK}(x)\,dx = \int_{\mathbb{R}^n} \chi_K(\frac{x}{r})\,dx = r^n \int_{\mathbb{R}^n} \chi_K(x)\,dx = r^n V(K).$$

□

Das Volumen von Körpern, die in einer Koordinatenrichtung durch den Graph einer Funktion begrenzt werden, erhält man wie im eindimensionalen Fall.

Theorem IX-3.6. *Es sei $K \subset \mathbb{R}^n$ kompakt, $f \in C_c(\mathbb{R}^n)$ und $f \geq 0$ auf K. Das Volumen des Körpers*

$$K_f = \{(x,t) \colon x \in K, 0 \leq t \leq f(x)\} \subset \mathbb{R}^{n+1}$$

ist gegeben durch

$$V_{n+1}(K_f) = \int_K f(x)\,dx.$$

Beweis. Mit Hilfe des Satzes von Fubini findet man wegen $\chi_{K_f}(x,t) = \chi_{[0,f(x)]}(t)\chi_K(x)$ mit $z = (x,t) \in \mathbb{R}^{n+1}$

$$\begin{aligned} V_{n+1}(K_f) &= \int_{\mathbb{R}^{n+1}} \chi_{K_f}(z)\,dz = \int_{\mathbb{R}^n} [\int_{\mathbb{R}} \chi_{[0,f(x)]}(t)\chi_K(x)\,dt]\,dx \\ &= \int_{\mathbb{R}^n} [\int_0^{f(x)} 1\,dt]\chi_K(x)\,dx = \int_{\mathbb{R}^n} f(x)\chi_K(x)\,dx = \int_K f(x)\,dx. \end{aligned}$$

□

Theorem IX-3.7 (Cavalierisches Prinzip). *Es sei $K \subset \mathbb{R}^n$ kompakt. Für $t \in \mathbb{R}$ bezeichne K_t die $(n-1)$-dimensionale Schnittmenge*

$$K_t = \{x' \in \mathbb{R}^{n-1} \colon (x',t) \in K\},$$

(K_t kann auch leer sein). Dann gilt

$$V_n(K) = \int_{\mathbb{R}} V_{n-1}(K_t)\,dt.$$

Beweis. Die charakteristische Funktion von K_t erfüllt für $x' \in \mathbb{R}^{n-1}$

$$\chi_{K_t}(x') = \chi_K(x',t),$$

und somit

$$V_{n-1}(K_t) = \int_{\mathbb{R}^{n-1}} \chi_K(x',t)\,dx'.$$

Die Behauptung ergibt sich nun aus dem Satz von Fubini

$$V_n(K) = \int_{\mathbb{R}^n} \chi_K(x)\,dx = \int_{\mathbb{R}} [\int_{\mathbb{R}^{n-1}} \chi_K(x',t)\,dx']\,dt = \int_{\mathbb{R}} V_{n-1}(K_t)\,dt.$$

$\square$

Das klassische Prinzip von Cavalieri ist nun eine einfache Folgerung.

Korollar IX-3.8. Es seien K, $L \subset \mathbb{R}^n$ kompakt. Gilt $V_{n-1}(K_t) = V_{n-1}(L_t)$ für alle $t \in \mathbb{R}$, dann haben K und L gleiches Volumen.

Beispiel IX-3.9. Es sei $B \subset \mathbb{R}^{n-1}$ kompakt und $h > 0$. Wir definieren einen Kegel $K_h(B)$ mit Basis B und Höhe h durch

$$K_h(B) := \{((1-\lambda)\xi, \lambda h) \in \mathbb{R}^{n-1} \times \mathbb{R} \colon \xi \in B, 0 \leq \lambda \leq 1\}.$$

Für die Schnittmengen

$$K_h(B)_t = \{x' \in \mathbb{R}^{n-1} \colon (x',t) \in K_h(B)\},$$

$t \in \mathbb{R}$, gilt

$$K_h(B)_t = \begin{cases} (1-\frac{t}{h})B, & t \in [0,h], \\ \emptyset, & \text{sonst.} \end{cases}.$$

Mit den Sätzen IX-3.7 und IX-3.5 folgt

$$\begin{aligned} V_n(K_h(B)) &= \int_0^h V_{n-1}(K_h(B)_t)\,dt = \int_0^h V_{n-1}((1-\frac{t}{h})B)\,dt \\ &= V_{n-1}(B)\int_0^h (1-\frac{t}{h})^{n-1}\,dt = \frac{h}{n}V_{n-1}(B). \end{aligned}$$

Beispiel IX-3.10. Es seien $a_1, \dots, a_n$ Vektoren im $\mathbb{R}^n$. Das Volumen des von diesen Vektoren aufgespannten Simplex

$$S(a_1,\dots,a_n) := \{x = \sum_{i=1}^n \lambda_i a_i \colon \sum_{i=1}^n \lambda_i = 1, \lambda_i \geq 0, i = 1,\dots,n\}$$

ist gegeben durch

$$V(S(a_1,\dots,a_n)) = \frac{1}{n!}|\det(a_1,\dots,a_n)|.$$

Wir betrachten zuerst den Einheitssimplex $S(e_1,\dots,e_n)$, der von den Vektoren e_j der kanonischen Basis aufgespannt wird. Mittels vollständiger Induktion zeigen wir

$$V(S(e_1,\dots,e_n)) = \frac{1}{n!}.$$

Für $n = 1$ ist $S(e_1) = [0,1]$, also $V(S(e_1)) = 1$. Allgemein ist $S(e_1,\dots,e_n)$ ein Kegel mit der Höhe 1 über der Basis $S(e_1,\dots,e_{n-1})$. Nach Induktionsvoraussetzung gilt

$$V_{n-1}(S(e_1,\dots,e_{n-1})) = \frac{1}{(n-1)!}.$$

Der Induktionsschritt folgt nun aus Beispiel IX-3.9

$$V_n(K_1(S(e_1,\dots,e_{n-1}))) = \frac{1}{n}V_{n-1}(S(e_1,\dots,e_{n-1})) = \frac{1}{n!}.$$

Da $x \in S(a_1,\dots,a_n)$ genau dann gilt, wenn

$$x = \sum_{i=1}^n \lambda_i a_i = \sum_{i=1}^n \lambda_i(a_1,\dots,a_n)e_i = A\sum_{i=1}^n \lambda_i e_i$$

also

$$S(a_1,\dots,a_n) = AS(e_1,\dots,e_n),$$

mit $A = (a_1 \dots a_n)$, ergibt Satz IX-2.14 die Behauptung.

Beispiel IX-3.11. Das Volumen der n-dimensionalen Kugel $\bar{K}(0,r) = \{x \in \mathbb{R}^n \colon \|x\|_2 \le r\}$.
Nach Satz IX-3.5 gilt

$$V_n(\bar{K}(0,r)) = r^n V_n(\bar{K}(0,1)).$$

Gebräuchlich ist auch die Bezeichnung $\omega_n = V_n(\bar{K}(0,1))$. Für $n = 1$ ist $K(0,1) = [-1,1]$, also $\omega_1 = 2$. Für $n > 1$ führen wir die Berechnung von ω_n mit dem Cavalierischen Prinzip auf ω_{n-1} zurück. Für die Schnittmengen gilt

$$\bar{K}_n(0,1)_t = \begin{cases} \bar{K}_{n-1}(0,\sqrt{1-t^2}), & |t| \le 1, \\ \emptyset, & \text{sonst.} \end{cases}$$

Daraus folgt mit Satz IX-3.5

$$(*) \qquad \omega_n = \int_{-1}^1 V_{n-1}(\bar{K}_{n-1}(0,\sqrt{1-t^2}))\,dt = w_{n-1}\int_{-1}^1 (1-t^2)^{\frac{n-1}{2}}\,dt.$$

Substituiert man im letzten Integral $t = \cos x$, erhält man

$$c_n := \int_{-1}^1 (1-t^2)^{\frac{n-1}{2}}\,dt = 2\int_0^{\frac{\pi}{2}} \sin^n x\,dx.$$

Setzt man $\omega_0 = 1$, gilt $(*)$ für alle $n \ge 1$. Mit partieller Integration und Induktion nach n zeigt man für $k \ge 1$

$$c_{2k} = \pi\prod_{i=1}^k \frac{2i-1}{2i}, \qquad c_{2k+1} = 2\prod_{i=1}^k \frac{2i}{2i+1}.$$

Insbesondere gilt

$$c_k c_{k-1} = \frac{2\pi}{k}, \qquad k \in \mathbb{N},$$

und somit wegen $(*)$

$$\omega_n = \omega_{n-1}c_n = \omega_{n-2}c_n c_{n-1} = \frac{2\pi}{n}\omega_{n-2}.$$

Eine einfache Induktion ergibt schließlich für $k \geq 1$

$$\omega_{2k} = \frac{\pi^k}{k!}, \qquad \text{und } \omega_{2k+1} = \frac{2^{k+1}}{1 \cdot 3 \cdot \ldots \cdot (2k+1)} \pi^k.$$

Bemerkung IX-3.12. Der Flächeninhalt des Einheitskreises beträgt demnach π. Hier zeigt sich zum ersten Mal die enge Beziehung zwischen der Zahl π und gewissen Kenngrößen des Kreises. Wir erinnern daran, daß π als kleinste, positive Nullstelle des Kosinus eingeführt wurde.

Beispiel IX-3.13. Volumen eines Rotationskörpers.
Es sei $[a, b] \subset \mathbb{R}$, $f \in C([a, b], \mathbb{R})$, $f \geq 0$ und

$$K = \{(x, y, t) \in \mathbb{R}^2 \times [a, b] \colon x^2 + y^2 \leq f(t)^2\}$$

Dann gilt

$$V(K) = \pi \int_a^b f(t)^2 \, dt.$$

Die Schnittmengen K_t sind nämlich gegeben durch

$$K_t = \begin{cases} \{(x, y) \in \mathbb{R}^2 \colon x^2 + y^2 \leq f(t)^2\}, & t \in [a, b], \\ \emptyset, & \text{sonst.} \end{cases}$$

Nach dem vorangehenden Beispiel gilt $V_2(K_t) = \pi f(t)^2$. Das Volumen des Rotationskörpers ergibt sich nun aus dem Cavalierischen Prinzip.

4. Das Lebesgue Integral

Wir führen nun die letzte Erweiterung des Integrals durch. Dazu benötigen wir den Begriff des Ober- und des Unterintegrals.

Definition IX-4.1. *Für eine beliebige Funktion $f \colon \mathbb{R}^n \to \bar{\mathbb{R}}$ definieren wir das* ***Oberintegral*** *durch*

$$\int^* f(x) \, dx = \inf\{\int_{\mathbb{R}^n} \varphi(x) \, dx \colon \varphi \in \mathcal{H}_U, \varphi \geq f\} \in \bar{\mathbb{R}},$$

und das ***Unterintegral*** *durch*

$$\int_* f(x) \, dx = \sup\{\int_{\mathbb{R}^n} \psi(x) \, dx \colon \psi \in \mathcal{H}_O, \psi \leq f\} \in \bar{\mathbb{R}},$$

Die Funktion $\varphi \equiv \infty$ gehört zu $\mathcal{H}_U$, die Funktion $\psi \equiv -\infty$ gehört zu $\mathcal{H}_O$. Daher sind die Mengen, über welche das Supremum bzw. das Infimum gebildet wird, nicht leer. Wir betonen, daß das Oberintegral bzw. das Unterintegral für beliebige Funktionen gebildet werden kann. Eine unmittelbare Folgerung aus der Definition ist die Beziehung

$$\int_* f(x) \, dx = -\int^* (-f)(x) \, dx.$$

Wir können uns im Folgenden also ohne Beschränkung der Allgemeinheit auf die Diskussion des Oberintegrals beschränken.

Lemma IX-4.2. *1) Für jedes $f\colon \mathbb{R}^n \to \bar{\mathbb{R}}$ gilt*

$$\int_* f(x)\,dx \le \int^* f(x)\,dx.$$

2)Für jedes $f \in \mathcal{H}_U$ oder $f \in \mathcal{H}_O$ gilt

$$\int_* f(x)\,dx = \int^* f(x)\,dx = \int_{\mathbb{R}^n} f(x)\,dx.$$

Beweis. 1) Es seien $\varphi \in \mathcal{H}_U$ und $\psi \in \mathcal{H}_O$ Funktionen mit $\psi \le f \le \varphi$. Daraus folgt $\varphi - \psi \ge 0$ und $\varphi - \psi \in \mathcal{H}_U$, da $-\psi \in \mathcal{H}_U$. Mit Lemma IX-2.9 schließt man

$$\int_{\mathbb{R}^n} (\varphi - \psi)(x)\,dx = \int_{\mathbb{R}^n} \varphi(x)\,dx + \int_{\mathbb{R}^n} (-\psi)(x)\,dx \ge 0,$$

also

$$\int_{\mathbb{R}^n} \varphi(x)\,dx \ge -\int_{\mathbb{R}^n} (-\psi)(x)\,dx = \int_{\mathbb{R}^n} \psi(x)\,dx.$$

Man achte auf die unterschiedliche Definition der Integrale für φ und ψ.
2) Es sei $f \in \mathcal{H}_U$. Die Gleichheit der Integrale

$$\int^* f(x)\,dx = \int_{\mathbb{R}^n} f(x)\,dx$$

ergibt sich aus der Definition. Um die Gleichheit

$$\int_* f(x)\,dx = \int_{\mathbb{R}^n} f(x)\,dx$$

einzusehen, wähle man eine Folge $(\psi_k) \subset C_c(\mathbb{R}^n)$ mit $\psi_k \uparrow f$. Aus der Definition IX-2.7 des Integrals von f folgt

$$(*) \qquad \int_{\mathbb{R}^n} f(x)\,dx = \sup_k \int_{\mathbb{R}^n} \psi_k(x)\,dx.$$

Da die Funktionen ψ_k auch in $\mathcal{H}_O$ liegen, erhält man aus der Definition IX-4.1 andererseits

$$\sup_k \int_{\mathbb{R}^n} \psi_k(x)\,dx \le \int_* f(x)\,dx \le \int^* f(x)\,dx = \int_{\mathbb{R}^n} f(x)\,dx.$$

Zusammen mit $(*)$ ergibt dies die Behauptung. □

Lemma IX-4.3. *1) Für f, $g\colon \mathbb{R}^n \to \bar{\mathbb{R}}$ mit $f \le g$ gilt*

$$\int^* f(x)\,dx \le \int^* g(x)\,dx, \qquad \int_* f(x)\,dx \le \int_* g(x)\,dx$$

2) Für $f\colon \mathbb{R}^n \to \bar{\mathbb{R}}$ und $\lambda \ge 0$ gilt

$$\int^* (\lambda f)(x)\,dx = \lambda \int^* f(x)\,dx, \qquad \int_* (\lambda f)(x)\,dx = \lambda \int_* f(x)\,dx.$$

3) Für $f,\, g\colon \mathbb{R}^n \to \bar{\mathbb{R}}$ gilt

$$\int^* (f(x)+g(x))\,dx \le \int^* f(x)\,dx + \int^* g(x)\,dx,$$
$$\int_* (f(x)+g(x))\,dx \ge \int_* f(x)\,dx + \int_* g(x)\,dx,$$

(wir nehmen an, daß f und g nicht gleichzeitig die Werte ∞ und $-\infty$ annehmen).

Beweis. 1) Übung.
2) Für $\lambda = 0$ ist die Aussage trivial (wir treffen die Vereinbarung $0\cdot\infty = 0$). Es sei also $\lambda > 0$. Zu beliebig gewähltem $\varepsilon > 0$ gibt es nach der Definition des Infimums eine Abbildung $\varphi \in \mathcal{H}_U$ mit $\varphi \ge f$ und

$$\int_{\mathbb{R}^n} \varphi(x)\,dx < \int^* f(x)\,dx + \frac{\varepsilon}{\lambda}.$$

Wegen $\lambda\varphi \in \mathcal{H}_U$ und $\lambda\varphi \ge \lambda f$ folgt aus der Definition des Oberintegrals

$$\int^* (\lambda f)(x)\,dx \le \int_{\mathbb{R}^n} (\lambda\varphi)(x)\,dx = \lambda \int_{\mathbb{R}^n} \varphi(x)\,dx < \lambda \int^* f(x)\,dx + \varepsilon,$$

also

$$(*)\qquad \int^* (\lambda f)(x)\,dx \le \lambda \int^* f(x)\,dx.$$

Ersetzt man f in $(*)$ durch λf und gleichzeitig λ durch $\frac{1}{\lambda}$, erhält man

$$\int^* f(x)\,dx \le \frac{1}{\lambda} \int^* (\lambda f)(x)\,dx.$$

Zusammen mit $(*)$ ergibt dies die Behauptung.
3) Zu beliebigem $\varepsilon > 0$ gibt es Funktionen $\varphi,\, \psi \in \mathcal{H}_U$ mit $\varphi \ge f$, $\psi \ge g$ und

$$\int^* f(x)\,dx + \frac{\varepsilon}{2} > \int_{\mathbb{R}^n} \varphi(x)\,dx, \qquad \int^* g(x)\,dx + \frac{\varepsilon}{2} > \int_{\mathbb{R}^n} \psi(x)\,dx.$$

Wegen $\varphi + \psi \in \mathcal{H}_U$ und $\varphi + \psi \ge f + g$ folgt

$$\int^* f(x)\,dx + \int^* g(x)\,dx + \varepsilon > \int_{\mathbb{R}^n} (\varphi(x)+\psi(x))\,dx \ge \int^* (f(x)+g(x))\,dx$$

Dies zeigt die Subadditivität des Oberintegrals. Die Ungleichung für das Unterintegral ist nun eine einfache Konsequenz:

$$\int_* (f(x)+g(x))\,dx = -\int^* (-f(x)-g(x))\,dx \ge -\int^* (-f)(x)\,dx - \int^* (-g)(x)\,dx$$
$$= \int_* f(x)\,dx + \int_* g(x)\,dx$$

$\square$

Theorem IX-4.4. *Für beliebige Funktionen $f_k\colon \mathbb{R}^n \to [0,\infty]$, $k \in \mathbb{N}$, gilt*

$$\int^* \sum_{k=1}^{\infty} f_k(x)\,dx \le \sum_{k=1}^{\infty} \int^* f_k(x)\,dx.$$

Beweis. Zu $\varepsilon > 0$ gibt es für jedes $k \in \mathbb{N}$ eine Funktion $\varphi_k \in \mathcal{H}_U$ mit $0 \le f_k \le \varphi_k$ und

$$\int_{\mathbb{R}^n} \varphi_k(x)\,dx < \int^* f_k(x)\,dx + \frac{\varepsilon}{2^k}.$$

Nach Lemma IX-2.11 gilt $\varphi = \sum_{k=1}^{\infty} \varphi_k \in \mathcal{H}_U$ und

$$\int_{\mathbb{R}^n} \varphi(x)\,dx = \sum_{k=1}^{\infty} \int_{\mathbb{R}^n} \varphi_k(x)\,dx \le \sum_{k=1}^{\infty} \int^* f_k(x)\,dx + \varepsilon.$$

Wegen

$$\sum_{k=1}^{\infty} f_k \le \sum_{k=1}^{\infty} \varphi_k = \varphi$$

folgt aus der Definition des Oberintegrals

$$\int^* \sum_{k=1}^{\infty} f_k(x)\,dx \le \int_{\mathbb{R}^n} \varphi(x)\,dx$$

also

$$\int^* \sum_{k=1}^{\infty} f_k(x)\,dx \le \sum_{k=1}^{\infty} \int^* f_k(x)\,dx + \varepsilon.$$

Dies ist gleichwertig mit der Behauptung, da $\varepsilon > 0$ beliebig klein gewählt werden kann. □

Definition IX-4.5. *Eine Abbildung $f\colon \mathbb{R}^n \to \bar{\mathbb{R}}$ mit*

$$-\infty < \int_* f(x)\,dx = \int^* f(x)\,dx < \infty.$$

heißt ***Lebesgue-integrierbar****. Den gemeinsamen Wert des Ober- und Unterintegrals nennt man* ***Lebesgue-Integral*** *von f, $\int_{\mathbb{R}^n} f(x)\,dx$. Die Menge der Lebesgue-integrierbaren Funktionen $f\colon \mathbb{R}^n \to \mathbb{R}$ bezeichnen wir mit $\mathcal{L}^1(\mathbb{R}^n)$.*

Die Funktionen in $\mathcal{H}_U$ sind nach Lemma IX-4.2 somit genau dann Lebesgue-integrierbar, wenn ihr Integral gemäß Definition IX-2.7 endlich ist. Insbesondere sind somit stetige Funktionen mit kompaktem Träger Lebesgue-integrierbar. Lemma IX-4.2 zeigt auch auf, daß der Wert des Lebesgue-Integrals dieser Funktionen mit dem früheren Integral übereinstimmt. Aus diesem Grunde werden wir im Folgenden Lebesgue-integrierbare Funktionen kurz integrierbar bezeichnen.
Als nächstes leiten wir eine äquivalente Charakterisierung der Integrierbarkeit her.

THEOREM IX-4.6. *1) Eine Abbildung $f\colon \mathbb{R}^n \to \bar{\mathbb{R}}$ ist integrierbar genau dann, wenn*

$$\forall \varepsilon > 0\, \exists g \in C_c(\mathbb{R}^n)\colon \int^* |f(x) - g(x)|\, dx < \varepsilon.$$

2) Es sei $(\varphi_k) \subset C_c(\mathbb{R}^n)$ und $f\colon \mathbb{R}^n \to \bar{\mathbb{R}}$ derart, daß

$$\lim_{k\to\infty} \int^* |f(x) - \varphi_k(x)|\, dx = 0.$$

Dann ist f integrierbar und es gilt

$$\int_{\mathbb{R}^n} f(x)\, dx = \lim_{k\to\infty} \int_{\mathbb{R}^n} \varphi_k(x)\, dx.$$

BEWEIS. 1) Es sei f integrierbar, also $\int^* f(x)\, dx = \int_* f(x)\, dx \in \mathbb{R}$. Zu beliebig gewähltem $\varepsilon > 0$ gibt es daher Funktionen $\varphi \in \mathcal{H}_U$ und $\psi \in \mathcal{H}_O$ mit $\psi \le f \le \varphi$, so daß

$$-\infty < \int_{\mathbb{R}^n} \psi(x)\, dx \le \int_{\mathbb{R}^n} f(x)\, dx \le \int_{\mathbb{R}^n} \varphi(x)\, dx < \infty$$

und

$$\int_{\mathbb{R}^n} (\varphi - \psi)(x)\, dx < \frac{\varepsilon}{2}.$$

Bei der Herleitung der letzten Ungleichung verwendet man $-\psi \in \mathcal{H}_U$, somit $\varphi - \psi \in \mathcal{H}_U$ und $-\int_{\mathbb{R}^n} \psi(x)\, dx = \int_{\mathbb{R}^n} (-\psi)(x)\, dx$ (man achte darauf, daß auf der linken Seite das Integral in $\mathcal{H}_O$, auf der rechten Seite das Integral in $\mathcal{H}_U$ zur Anwendung kommt). Aus der Ungleichung

$$\varphi - f \le \varphi - \psi$$

folgert man

$$\int^* |\varphi(x) - f(x)|\, dx = \int^* (\varphi(x) - f(x))\, dx < \frac{\varepsilon}{2}.$$

Weiters existiert nach der Definition des Integrals für Funktionen in $\mathcal{H}_U$ eine Abbildung $g \in C_c(\mathbb{R}^n)$ mit $g \le \varphi$ und

$$\int_{\mathbb{R}^n} |\varphi(x) - g(x)|\, dx < \frac{\varepsilon}{2}.$$

Insgesamt erhält man

$$\int^* |f(x) - g(x)|\, dx \le \int^* |f(x) - \varphi(x)|\, dx + \int^* |\varphi(x) - g(x)|\, dx < \varepsilon.$$

Umgekehrt existiere zu jedem $\varepsilon > 0$ eine Funktion $g \in C_c(\mathbb{R}^n)$ mit $\int^* |f(x) - g(x)|\, dx < \varepsilon$. Nach der Definition des Oberintegrals existiert $h \in \mathcal{H}_U$ mit

$$|f - g| \le h \qquad \text{und} \qquad \int_{\mathbb{R}^n} h(x)\, dx < \varepsilon.$$

Es folgt

$$g - h \le f \le g + h$$

(wegen der Beschränktheit von g sind die Funktionen $g \pm h\colon \mathbb{R}^n \to \bar{\mathbb{R}}$ definiert), $g - h \in \mathcal{H}_O$, $g + h \in \mathcal{H}_U$ und somit

$$\int_{\mathbb{R}^n} (g+h)(x)\,dx - \int_{\mathbb{R}^n} (g-h)(x)\,dx = 2\int_{\mathbb{R}^n} h(x)\,dx < 2\varepsilon.$$

Dies zeigt die Integrierbarkeit von f.
2) Die Integrierbarkeit von f ist eine Folge von 1). Es sei nun $\varepsilon > 0$ und $N(\varepsilon) \in \mathbb{N}$ so groß, daß $\int^* |f(x) - \varphi_k(x)|\,dx < \varepsilon$ für $k \geq N(\varepsilon)$ gilt. Wie im Beweis des ersten Teiles schließt man auf die Existenz von $h_k \in \mathcal{H}_U$ mit

$$|f - \varphi_k| \leq h_k \qquad \text{und} \qquad \int_{\mathbb{R}^n} h_k(x)\,dx < \varepsilon.$$

Aus $\varphi_k - h_k \leq f \leq \varphi_k + h_k$ folgt

$$\int_{\mathbb{R}^n} (\varphi_k - h_k)(x)\,dx \leq \int_{\mathbb{R}^n} f(x)\,dx \leq \int_{\mathbb{R}^n} (\varphi_k + h_k)(x)\,dx,$$

und daraus wegen Lemma IX-2.9

$$|\int_{\mathbb{R}^n} f(x)\,dx - \int_{\mathbb{R}^n} \varphi_k(x)\,dx| \leq \int_{\mathbb{R}^n} h_k(x)\,dx < \varepsilon$$

für $k \geq N(\varepsilon)$. □

Für das nächste Resultat ist folgende Notation zweckmäßig. Es sei $f\colon \mathbb{R}^n \to \bar{\mathbb{R}}$. Wir vereinbaren

$$f_+(x) := \max\{f(x), 0\}, \qquad f_-(x) := -\min\{f(x), 0\}, \quad x \in \mathbb{R}.$$

Offensichtlich gelten die Beziehungen

$$f = f_+ - f_- \qquad \text{und} \qquad |f| = f_+ + f_-.$$

Theorem IX-4.7. *Ist $f\colon \mathbb{R}^n \to \bar{\mathbb{R}}$ integrierbar, dann sind auch $|f|$, f_+ und f_- integrierbar und es gilt*

$$(*) \qquad |\int_{\mathbb{R}^n} f(x)\,dx| \leq \int_{\mathbb{R}^n} |f(x)|\,dx.$$

Sind umgekehrt f_+ und f_- integrierbar, dann ist auch f integrierbar.

Beweis. Ist f integrierbar, dann gibt es nach Satz IX-4.6 zu jedem $\varepsilon > 0$ eine Funktion $\varphi \in C_c(\mathbb{R}^n)$ mit $\int^* |f(x) - \varphi(x)|\,dx < \varepsilon$. Die Funktionen φ_+, φ_- und $|\varphi|$ liegen ebenfalls in $C_c(\mathbb{R}^n)$. Unter Berücksichtigung von

$$f_+(x) - \varphi_+(x) = \begin{cases} f(x) - \varphi(x), & f(x) \geq 0, \varphi(x) \geq 0, \\ f(x), & f(x) \geq 0, \varphi(x) < 0, \\ 0, & f(x) < 0, \varphi(x) < 0, \\ -\varphi(x), & f(x) < 0, \varphi(x) \geq 0, \end{cases}$$

(analoge für $f_-(x) - \varphi_-(x)$) findet man

$$|f_\pm - \varphi_\pm| \leq |f - \varphi| \qquad \text{und} \qquad \big||f| - |\varphi|\big| \leq |f - \varphi|.$$

Mit Lemma IX-4.3 folgt

$$\int^* |f_\pm(x) - \varphi_\pm(x)|\,dx < \varepsilon \qquad \text{und} \qquad \int^* \big||f(x)| - |\varphi(x)|\big|\,dx < \varepsilon,$$

was gleichwertig mit der Integrierbarkeit von $f_\pm$ und $|f|$ ist. Die Ungleichung $(*)$ ist eine Konsequenz der Monotonie des Ober- und des Unterintegrals und der Ungleichungen $-|f| \le f \le |f|$.
Sind umgekehrt die Funktionen f_+ und f_- integrierbar, dann gibt es nach Satz IX-4.6 Funktionen φ und $\psi \in C_c(\mathbb{R}^n)$ mit

$$\int_{\mathbb{R}^n} |f_+(x) - \varphi(x)|\,dx < \frac{\varepsilon}{2} \qquad \text{und} \qquad \int_{\mathbb{R}^n} |f_-(x) - \psi(x)|\,dx < \frac{\varepsilon}{2}.$$

Dies hat

$$\int^* |f(x) - \varphi(x) + \psi(x)|\,dx \le \int^* |f_+(x) - \varphi(x)|\,dx + \int^* |f_-(x) - \psi(x)|\,dx < \varepsilon$$

zur Folge. Wegen $\varphi - \psi \in C_c(\mathbb{R}^n)$ bedeutet dies die Integrierbarkeit von f. $\square$

Wir können nun zeigen, daß $\mathcal{L}^1(\mathbb{R}^n)$ ein Vektorraum ist. Die Addition von Funktionen in $\mathcal{L}^1(\mathbb{R}^n)$ ist definiert, da wir die Funktionswerte $\pm\infty$ für Elemente von $\mathcal{L}^1(\mathbb{R}^n)$ ausgeschlossen haben.

Theorem IX-4.8. *Es seien f, $g \in \mathcal{L}^1(\mathbb{R}^n)$ und $\lambda \in \mathbb{R}$. Dann sind auch $f + g$, $\lambda f \in \mathcal{L}^1(\mathbb{R}^n)$ und es gilt*

(1) $\int_{\mathbb{R}^n} (f+g)(x)\,dx = \int_{\mathbb{R}^n} f(x)\,dx + \int_{\mathbb{R}^n} g(x)\,dx,$
(2) $\int_{\mathbb{R}^n} (\lambda f)(x)\,dx = \lambda \int_{\mathbb{R}^n} f(x)\,dx,$
(3) *aus* $f \le g$ *folgt* $\int_{\mathbb{R}^n} f(x)\,dx \le \int_{\mathbb{R}^n} g(x)\,dx.$

Beweis. Wegen der Integrierbarkeit von f und g gibt es nach Satz IX-4.6-1) Folgen (φ_k), $(\psi_k) \subset C_c(\mathbb{R}^n)$ mit

$$\int^* |f(x) - \varphi_k(x)|\,dx \underset{k\to\infty}{\to} 0 \quad \text{und} \quad \int^* |g(x) - \psi_k(x)|\,dx \underset{k\to\infty}{\to} 0.$$

Mit Lemma IX-4.3 folgt dann

$$\int^* |\lambda f(x) - \lambda\varphi_k(x)|\,dx = \int^* |\lambda||f(x) - \varphi_k(x)|\,dx = |\lambda| \int^* |f(x) - \varphi_k(x)|\,dx \underset{k\to\infty}{\to} 0,$$

und

$$\int^* |(f+g)(x) - (\varphi_k + \psi_k)(x)|\,dx \le \int^* |f(x) - \varphi_k(x)|\,dx + \int^* |g(x) - \psi_k(x)|\,dx \underset{k\to\infty}{\to} 0.$$

Nach Satz IX-4.6-2) sind daher die Funktionen λf und $f + g$ integrierbar. Die Gültigkeit der Regeln (1) und (2) ergibt sich aus deren Gültigkeit in $C_c(\mathbb{R}^n)$. Die Aussage (3) ist ein Spezialfall von Lemma IX-4.3-1). $\square$

Theorem IX-4.9. *Es seien f, $g \in \mathcal{L}^1(\mathbb{R}^n)$. Ist g beschränkt, dann gilt $fg \in \mathcal{L}^1(\mathbb{R}^n)$*

Beweis. Da g beschränkt ist, existiert eine Konstante $M > 0$ mit $|g(x)| \le M$, $x \in \mathbb{R}^n$. Nach Satz IX-4.6-1) gibt es zu jedem $\varepsilon > 0$ Funktionen φ, $\psi \in C_c(\mathbb{R}^n)$ mit

$$\int^* |f(x) - \varphi(x)|\, dx < \frac{\varepsilon}{2M} \quad \text{und} \quad \int^* |g(x) - \psi(x)|\, dx < \frac{\varepsilon}{2(\|\varphi\|_\infty + 1)}.$$

Aus der Ungleichung

$$|fg - \varphi\psi| \le |f - \varphi||g| + |\varphi||g - \psi| \le |f - \varphi|M + \|\varphi\|_\infty |g - \psi|$$

folgt mit Lemma IX-4.3

$$\int^* |(fg)(x) - (\varphi\psi)(x)|\, dx \le M \int^* |f(x) - \varphi(x)|\, dx + \|\varphi\|_\infty \int^* |g(x) - \psi(x)|\, dx < \varepsilon.$$

Gemäß Satz IX-4.6 ist fg integrierbar. Da die Werte von fg in $\mathbb{R}$ liegen, gilt $fg \in \mathcal{L}(\mathbb{R}^n)$. $\square$

Definition IX-4.10. *1) Eine Teilmenge $M \subset \mathbb{R}^n$ heißt* ***integrierbar****, wenn ihre charakteristische Funktion χ_M integrierbar ist.*
2) Ist $M \subset \mathbb{R}^n$ integrierbar, nennt man

$$\lambda(M) := \int_{\mathbb{R}^n} \chi_M(x)\, dx$$

Lebesgue Maß *(oder* ***Volumen****) von M.*

Da die charakteristische Funktion einer kompakten Teilmenge des $\mathbb{R}^n$ in $\mathcal{H}_O$ liegt, sind nach Lemma IX-4.2-2 kompakte Mengen integrierbar und das Lebesgue Maß $\lambda(M)$ gemäß Definition IX-4.10 und das Volumen $V(M)$ gemäß Definition IX-3.1 stimmen überein.

Theorem IX-4.11. *Eine offene oder abgeschlossene Teilmenge $M \subset \mathbb{R}^n$ ist integrierbar genau dann, wenn $\int^* \chi_M(x)\, dx < \infty$.*

Beweis. Die charakteristischen Funktionen offener bzw. abgeschlossener Mengen liegen in $\mathcal{H}_U$ bzw. $\mathcal{H}_O$. Die Behauptung folgt nun aus Lemma IX-4.2 und Definition IX-4.10. $\square$

Theorem IX-4.12. *Es seien A, $B \subset \mathbb{R}^n$ integrierbare Mengen. Dann sind auch die Mengen $A \cap B$, $A \cup B$ und $A \setminus B$ integrierbar und es gilt*

$$\lambda(A \cup B) = \lambda(A) + \lambda(B) - \lambda(A \cap B),$$
$$\lambda(A \setminus B) = \lambda(A) - \lambda(A \cap B).$$

Beweis. Für charakteristische Funktionen gelten die Beziehungen

$$\chi_{A\cap B} = \chi_A \chi_B,$$
$$\chi_{A\cup B} = \chi_A + \chi_B - \chi_{A\cap B},$$
$$\chi_{A\setminus B} = \chi_A - \chi_{A\cap B}.$$

Da χ_A und χ_B integrierbar sind, folgt mit Satz IX-4.9 die Integrierbarkeit von $\chi_{A\cap B}$. Die Integrierbarkeit von $\chi_{A\cup B}$ und $\chi_{A\setminus B}$ ist dann eine Folge von Satz IX-4.8. $\square$

Wir wollen uns nun von der Einschränkung befreien, daß integrierbare Funktionen a priori auf $\mathbb{R}^n$ definiert sind.

Definition IX-4.13. *Es sei $M \subset \mathbb{R}^n$ eine* beliebige *Teilmenge. Eine Funktion $f\colon M \to \mathbb{R} \cup \{\pm\infty\}$ heißt **integrierbar über** M, falls die trivial fortgesetzte Funktion $\tilde{f}\colon \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$*

$$\tilde{f}(x) := \begin{cases} f(x), & x \in M, \\ 0, & \textit{sonst}, \end{cases}$$

integrierbar ist. Wir setzen dann

$$\int_M f(x)\,dx := \int_{\mathbb{R}^n} \tilde{f}(x)\,dx.$$

Beispiel IX-4.14. 1) Es sei $f\colon K \to \mathbb{R}$ stetig und $K \subset \mathbb{R}^n$ kompakt. Dann ist f über K integrierbar.

Beweis. Es sei $\tilde{f}\colon \mathbb{R}^n \to \mathbb{R}$ die trivial fortgesetzte Funktion. Ist $f \geq 0$, dann ist $\tilde{f}$ von oben halbstetig. Sie gehört somit zu $\mathcal{H}_O$ und ist daher integrierbar. Im allgemeinen Fall schreiben wir

$$\tilde{f} = (\tilde{f} + c\chi_K) - c\chi_K,$$

$c \in \mathbb{R}_+$ so groß, daß $f + c\chi_K \geq 0$ gilt. Wegen $\tilde{f} + c\chi_K = \widetilde{f + c}$ ist $\tilde{f} + c\chi_K$ integrierbar. Die Integrierbarkeit von $\tilde{f}$ ergibt sich nun aus der Integrierbarkeit von χ_K. □

2) Es sei $U \subset \mathbb{R}^n$ eine beschränkte, offene Menge. Dann ist jede beschränkte, stetige Funktion $f\colon U \to \mathbb{R}$ integrierbar über U.
3) Es sei $f \in \mathcal{L}^1(\mathbb{R}^n)$ und $M \subset \mathbb{R}^n$ integrierbar. Dann ist $f|_M$ integrierbar über M. Dies folgt aus der Beobachtung $\widetilde{f|_M} = f\chi_M$ und Satz IX-4.9.

5. Nullmengen

Definition IX-5.1. *1) Eine Teilmenge $M \subset \mathbb{R}^n$ heißt **Nullmenge**, wenn sie integrierbar ist und Lebesgue Maß null hat.*
*2) Ein Prädikat $E(x)$ gilt **fast überall**, wenn*

$$\{x \in \mathbb{R}^n \colon \neg E(x)\}$$

eine Nullmenge ist.

Wegen

$$0 \leq \int_* \chi_M(x)\,dx \leq \int^* \chi_M(x)\,dx$$

ist $M \subset \mathbb{R}^n$ eine Nullmenge genau dann, wenn

$$\int^* \chi_M(x)\,dx = 0.$$

Beispiel IX-5.2. Jede einelementige Menge $M = \{x_0\} \subset \mathbb{R}^n$ ist eine Nullmenge, da M in einen Würfel W_ε beliebig kleiner Seitenlänge $\varepsilon > 0$ eingeschlossen werden kann. Aus $\chi_M \leq \chi_{W_\varepsilon}$ folgt

$$\int^* \chi_M(x)\,dx \leq \int^* \chi_{W_\varepsilon}(x)\,dx = \varepsilon^n.$$

Theorem IX-5.3. *1) Jede Teilmenge einer Nullmenge ist eine Nullmenge.*
2) Eine abzählbare Vereinigung von Nullmengen ergibt wieder eine Nullmenge.

Beweis. Der Beweis der Aussage 1) ist trivial, der Beweis von 2) ergibt sich aus der Beobachtung

$$\chi_M \leq \sum_{k=1}^{\infty} \chi_{M_k}$$

für $M = \cup_{=1}^{\infty} M_k$. Gilt $\lambda(M_k) = 0$ für alle $k \in \mathbb{N}$, ergibt Satz IX-4.4

$$\int^* \chi_M(x)\,dx \leq \int^* \sum_{k=1}^{\infty} \chi_{M_k}(x)\,dx \leq \sum_{k=1}^{\infty} \int^* \chi_{M_k}(x)\,dx = 0.$$

□

Korollar IX-5.4. Das Intervall $[0,1]$ ist überabzählbar.

Beweis. Wäre $[0,1]$ abzählbar, dann wäre $[0,1]$ eine Nullmenge im Widerspruch zu $V([0,1]) = 1$. □

Beispiel IX-5.5. Es sei $K \subset \mathbb{R}^{n-1}$ kompakt und $f\colon K \to \mathbb{R}$ stetig. Dann ist der Graph $G(f)$ von f eine Nullmenge in $\mathbb{R}^n$.
Um dies einzusehen, beachte man, daß der Graph von f wegen der Kompaktheit von K und der Stetigkeit von f kompakt ist. Somit ist $\chi_{G(f)}$ integrierbar. Wegen

$$G(f) = \{(x', t) \in \mathbb{R}^{n-1} \times \mathbb{R}\colon x' \in K, t = f(x')\}.$$

gilt $\chi_{G(f)}(x', t) = \chi_K(x')\chi_{\{f(x')\}}(t)$. Die Behauptung ergibt sich nun aus dem Satz von Fubini und Beispiel IX-5.2

$$\begin{aligned}\int^* \chi_{G(f)}(x', t)\,dx'dt &= \int_{\mathbb{R}^n} \chi_{G(f)}(x', t)\,dx'dt \\ &= \int_{\mathbb{R}^{n-1}} [\int_{\mathbb{R}} \chi_{\{f(x')\}}(t)\,dt]\chi_K(x')\,dx' = 0.\end{aligned}$$

Als Folgerung halten wir fest, daß beispielsweise der Rand eines Kreises oder der Rand einer Kugel Nullmengen im $\mathbb{R}^2$ bzw. $\mathbb{R}^3$ sind.

Beispiel IX-5.6. Jede Hyperebene $H \subset \mathbb{R}^n$ ist eine Nullmenge. Man kann nämlich H auffassen als abzählbare Vereinigung kompakter Quader, von denen eine Seite die Länge null hat (Details Übung).

Nullmengen haben folgende bemerkenswerte Eigenschaft.

LEMMA IX-5.7. *Es sei* $N \subset \mathbb{R}^n$ *eine Nullmenge und* $u_N\colon \mathbb{R}^n \to [0,\infty]$ *gegeben durch*

$$u_N(x) = \begin{cases} 0, & x \in \mathbb{R}^n \setminus N, \\ \infty, & x \in N. \end{cases}$$

Dann ist u_N *integrierbar und es gilt*

$$\int_{\mathbb{R}^n} u_N(x)\,dx = 0.$$

BEWEIS. Man überzeuge sich von der Gültigkeit von $u_N = \sum_{k=1}^{\infty} \chi_N$. Aus Satz IX-4.4 folgt daher

$$\int^* u_N\,dx \le \sum_{k=1}^{\infty} \int^* \chi_N\,dx = 0.$$

Andererseits gilt $0 \le \int_* u_N\,dx \le \int^* u_N\,dx \le 0$. Dies zeigt die Behauptung. □

Mit Hilfe dieses Resultates können wir eine fundamentale Eigenschaft integrierbarer Funktionen zeigen.

THEOREM IX-5.8. *Stimmen zwei Funktionen* $f,\ g\colon \mathbb{R}^n \to \bar{\mathbb{R}}$ *fast überall überein und ist* f *integrierbar, dann ist auch* g *integrierbar und es gilt*

$$\int_{\mathbb{R}^n} f(x)\,dx = \int_{\mathbb{R}^n} g(x)\,dx.$$

BEWEIS. Es sei $(\varphi_k) \subset C_c(\mathbb{R}^n)$ eine Folge von Funktionen mit

$$\lim_{k\to\infty} \int^* |f(x) - \varphi_k(x)|\,dx = 0,$$

vgl. Satz IX-4.6. Ferner seien

$$N = \{x \in \mathbb{R}^n \colon f(x) \neq g(x)\}$$

und u_N wie in Lemma IX-5.7. Dann gilt

$$|g - \varphi_k| \le |f - \varphi_k| + u_N.$$

Da N eine Nullmenge ist, ergibt sich

$$\begin{aligned} \int^* |g(x) - \varphi_k(x)|\,dx &\le \int^* |f(x) - \varphi_k(x)|\,dx + \int^* u_N(x)\,dx \\ &= \int^* |f(x) - \varphi_k(x)|\,dx \underset{k\to\infty}{\to} 0. \end{aligned}$$

Bezieht man sich wieder auf Satz IX-4.6, folgt die Integrierbarkeit von g und

$$\int_{\mathbb{R}^n} g(x)\,dx = \lim_{k\to\infty} \int_{\mathbb{R}^n} \varphi_k(x)\,dx = \int_{\mathbb{R}^n} f(x)\,dx.$$

□

Satz IX-5.8 zeigt, daß man die Werte einer integrierbaren Funktion auf einer Nullmenge abändern kann, ohne den Wert des Integrals zu verändern. Wir zeigen nun, daß eine integrierbare Funktion die Werte $\pm\infty$ höchstens auf einer Nullmenge annehmen kann.

THEOREM IX-5.9. *Es sei $f\colon \mathbb{R}^n \to \bar{\mathbb{R}}$ und es gelte*

$$M := \int^* |f(x)|\, dx < \infty.$$

Dann ist die Menge

$$N = \{x \in \mathbb{R}^n \colon |f(x)| = \infty\}$$

eine Nullmenge.

BEWEIS. Für alle $\varepsilon > 0$ gilt

$$\chi_N \le \varepsilon |f|$$

und somit

$$\int^* \chi_N(x)\, dx \le \varepsilon \int^* |f(x)|\, dx \le \varepsilon M.$$

Dies ist nur möglich, falls $\int^* \chi_N(x)\, dx = 0$, also N eine Nullmenge ist. □

Als Konsequenz aus den letzten beiden Sätzen erkennen wir, daß man jede integrierbare Funktion durch eine fast überall gleiche Funktion, die allerdings nur endliche Werte annimmt, ersetzen kann, ohne den Wert des Integrals zu verändern. Die Beschränkung auf endliche Funktionswerte für Funktionen in $\mathcal{L}^1(\mathbb{R}^n)$ stellt demnach keine Beschränkung der Allgemeinheit dar.

THEOREM IX-5.10. *Für $f\colon \mathbb{R}^n \to \bar{\mathbb{R}}$ gilt*

$$\int^* |f(x)|\, dx = 0$$

genau dann, wenn $f = 0$ fast überall gilt.

BEWEIS. Sei $A = \{x \in \mathbb{R}^n \colon f(x) \neq 0\}$. Ist A eine Nullmenge und u_A definiert wie in Lemma IX-5.7, dann folgt aus

$$|f| \le u_A$$

mit Lemma IX-4.3 und Lemma IX-5.7

$$\int^* |f(x)|\, dx \le \int^* u_A(x)\, dx = 0$$

Es sei nun umgekehrt $\int^* |f(x)| = 0$. Für $F(x) = \sum_{k=1}^\infty |f(x)|$, $x \in \mathbb{R}^n$, gilt $F = u_A$ und somit $\chi_A \le u_A$. Mit Satz IX-4.4 erhält man

$$\int^* \chi_A\, dx \le \int^* u_A\, dx = \int^* \sum_{k=1}^\infty |f(x)|\, dx \le \sum_{k=1}^\infty \int^* |f(x)|\, dx = 0$$

Somit ist $\chi(A)$ integrierbar und es gilt $\lambda(A) = 0$. □

THEOREM IX-5.11. *Eine Teilmenge $N \subset \mathbb{R}^n$ ist genau dann eine Nullmenge, wenn es zu jedem $\varepsilon > 0$ eine offene Menge mit $N \subset U$ und $\lambda(U) < \varepsilon$ gibt.*

BEWEIS. Aus $N \subset U$ folgt $\lambda(N) \leq \lambda(U) < \varepsilon$ und somit $\lambda(N) = 0$. Es sei umgekehrt N eine Nullmenge und $\alpha > 1$. Dann ist

$$\int^* \alpha\chi_N \, dx = 0$$

und zu $\varepsilon > 0$ existiert $\varphi \in \mathcal{H}_U$ mit $\alpha\chi_N \leq \varphi$ und

$$\int_{\mathbb{R}^n} \varphi \, dx < \varepsilon.$$

Wegen $\varphi \geq \alpha\chi_N > 1$ auf N und da φ halbstetig von unten ist, folgt aus Lemma IX-1.2, daß

$$U = \{x \in \mathbb{R}^n \colon \varphi(x) > 1\}$$

offen ist. Ferner gilt $N \subset U$ und $\chi_U \leq \varphi$, was

$$\lambda(U) = \int \chi_U \, dx \leq \int_{\mathbb{R}^n} \varphi \, dx < \varepsilon$$

zur Folge hat. □

Eine weitere Charakterisierung von Nullmengen beruht auf folgender Eigenschaft offener Mengen.

LEMMA IX-5.12. *Jede offene Teilmenge $U \subset \mathbb{R}^n$ ist die Vereinigung von abzählbar vielen kompakten Quadern, deren Inneres disjunkt ist.*

BEWEIS. Für jedes $k \in \mathbb{N}_0$ und $m = (m_1, \ldots, m_n) \in \mathbb{Z}^n$ betrachten wir die Quader

$$Q_{km} = \{(x_1, \ldots, x_n) \colon x_i \in [\frac{m_i}{2^k}, \frac{m_i + 1}{2^k}]\}$$

und setzen

$$Q_k = \bigcup_{\substack{m \in \mathbb{Z}^n \\ Q_{km} \subset U}} Q_{km}, \quad k \in \mathbb{N}_0.$$

Dann gilt

$$U = \cup_k Q_k.$$

Die Inklusion $\cup_k Q_k \subset U$ ist klar. Da U offen ist, gibt es für jedes $x = (x_1, \ldots, x_n) \in U$ eine quaderförmige Kugel $K_\infty(x, r)$, die in U enthalten ist. Es sei $k \in \mathbb{N}_0$ so gewählt, daß

$$\frac{1}{2^k} < r.$$

Setzt man $m_i = [2^k x_i]$, $i = 1, \ldots, n$, also

$$\frac{m_i}{2^k} \leq x_i < \frac{m_i + 1}{2^k},$$

erhält man

$$x_i - r < x_i - \frac{1}{2^k} < \frac{m_i}{2^k} \leq x_i < \frac{m_i+1}{2^k} \leq x_i + \frac{1}{2^k} < x_i + r.$$

Dies zeigt $[\frac{m_i}{2^k}, \frac{m_i+1}{2^k}] \subset (x_i - r, x_i + r)$, $i = 1, \ldots, n$ und daher

$$Q_{km} \subset K_\infty(x, r) \subset U.$$

□

THEOREM IX-5.13. *Eine Teilmenge $N \subset \mathbb{R}^n$ ist genau dann eine Nullmenge, wenn es zu jedem $\varepsilon > 0$ eine abzählbare Familie von Quadern $(Q_i)_{i=1}^\infty$ mit paarweise disjunktem Inneren gibt mit der Eigenschaft*

$$N \subset \cup_{i=1}^\infty Q_i \quad \textit{und} \quad \sum_{i=1}^\infty \lambda(Q_i) < \varepsilon$$

BEWEIS. Die Bedingung ist hinreichend. Aus $N \subset \cup_{i=1}^\infty Q_i$ folgt nämlich $\chi_N \leq \sum_{i=1}^\infty \chi_{Q_i}$ und daraus mit Satz IX-4.4

$$\int^* \chi_N \leq \int^* \sum_{i=1}^\infty \chi_{Q_i} \leq \sum_{i=1}^\infty \int^* \chi_{Q_i} = \sum_{i=1}^\infty \lambda(Q_i) < \varepsilon,$$

also $\lambda(N) = 0$. Es sei nun umgekehrt N eine Nullmenge. Dann gibt es nach Satz IX-5.11 für jedes $\varepsilon > 0$ eine offene Menge U mit $N \subset U$ und $\lambda(U) < \varepsilon$. Wegen Lemma IX-5.12 kann man U darstellen als abzählbare Vereinigung von Quadern mit disjunktem Inneren, $U = \cup_{k=1}^\infty Q_k$. Wegen $Q_i^\circ \cap Q_j^\circ = \emptyset$ ist $Q_i \cap Q_j = \emptyset$ oder es liegt $Q_i \cap Q_j$ in einer $n-1$-dimensionalen Hyperebene und es gilt $\lambda(Q_i \cap Q_j) = 0$ für $i \neq j$. Mit Hilfe von Satz IX-4.12 folgert man

$$\sum_{i=1}^k \lambda(Q_i) = \lambda(\cup_{i=1}^k Q_i) \leq \lambda(U) < \varepsilon$$

für alle $k \in \mathbb{N}$ und somit $\sum_{i=1}^\infty \lambda(Q_i) < \varepsilon$. □

Wir untersuchen nun, wie sich Nullmengen unter Lipschitz stetigen Abbildungen verhalten.

THEOREM IX-5.14. *Es sei $N \subset \mathbb{R}^n$ eine Nullmenge und $f\colon N \to \mathbb{R}^n$ Lipschitz stetig mit Lipschitz Konstante $\alpha \geq 0$. Dann ist $f(N)$ eine Nullmenge.*

BEWEIS. Die Behauptung ist trivial für $\alpha = 0$. Es sei also $\alpha > 0$ und $N \subset \mathbb{R}^n$ eine Nullmenge. Dann existieren nach dem vorangehenden Satz abzählbar viele Quader Q_i, $i \in \mathbb{N}$, mit disjunktem Inneren, so daß

$$N \subset \cup_{i=1}^\infty Q_i \quad \text{und} \quad \sum_{i=1}^\infty \lambda(Q_i) < \varepsilon.$$

Wegen $N = \cup_{i=1}^\infty (Q_i \cap N)$ und $f(N) = \cup_{i=1}^\infty f(Q_i \cap N)$ betrachten wir vorerst $f(Q \cap N)$ für einen beliebigen Quader Q mit $Q \cap N \neq \emptyset$. Wir können den Quader auffassen als abgeschlossene Kugel $\bar{K}_\infty(a,r)$ für ein geeignetes $a \in \mathbb{R}^n$ und $r > 0$. Für ein festes $x_0 \in Q \cap N$ erhält man für jedes $x \in Q \cap N$

$$\begin{aligned}||f(x) - f(x_0)||_\infty &\le \alpha ||x - x_0||_\infty \\ &\le \alpha(||x-a||_\infty + ||x_0 - a||_\infty) \le 2\alpha r.\end{aligned}$$

Diese Ungleichung zeigt

$$f(N \cap Q) \subset \bar{K}_\infty(f(x_0), 2\alpha r)(\equiv \tilde{Q})$$

und daher auch

$$\lambda(\bar{K}_\infty(f(x_0), 2\alpha r)) = (2\alpha 2r)^n = (2\alpha)^n \lambda(Q).$$

Es gibt also Quader $\tilde{Q}_i$ mit $f(Q_i \cap N) \subset \tilde{Q}_i$, $i \in \mathbb{N}$. Dies hat

$$f(N) = \cup_{i=1}^\infty f(Q_i \cap N) \subset \cup_{i=1}^\infty \tilde{Q}_i$$

und

$$\sum_{i=1}^\infty \tilde{\lambda}(Q_i) \le (2\lambda)^n \sum_{i=1}^\infty \lambda(Q_i) < (2\lambda)^n \varepsilon.$$

zur Folge. Somit ist $\lambda(f(N)) = 0$. □

Korollar IX-5.15. Es sei $U \subset \mathbb{R}^n$ offen, $f\colon U \to \mathbb{R}^n$ stetig differenzierbar und $N \subset U$ eine Nullmenge. Dann ist $f(N)$ eine Nullmenge.

Beweis. Wir stellen U gemäß Lemma IX-5.12 als abzählbare Vereinigung kompakter Quader dar, $U = \cup_{i=1}^\infty Q_i$. Dann ist $f(N) = \cup_{i=1}^\infty f(N \cap Q_i)$. Für jeden Quader Q_i und $x, y \in Q_i$ ergibt der Mittelwertsatz VII-VIII-5.6 (beachte: Q_i ist konvex)

$$||f(x) - f(y)||_\infty \le \max_{z \in Q_i} ||f'(z)||\, ||x - y||,$$

d.h. $f|_{Q_i}$ ist Lipschitz stetig für alle $i \in \mathbb{N}$. Da $N \cap Q_i$ eine Nullmenge ist, folgt aus den Sätzen IX-5.14 und IX-5.3, daß auch $f(N \cap Q_i)$ und $f(N)$ Nullmengen sind. □

Man beachte, daß eine stetig differenzierbare Abbildung auf einer offenen Menge nicht notwendig Lipschitz ist.

6. Konvergenzsätze

Die guten Konvergenzeigenschaften des Lebesgueintegrales gehören zu den wichtigsten Gründen für dessen weite Verbreitung.

Theorem IX-6.1 (Satz von der monotonen Konvergenz, B.Levi). *Es sei $f_k\colon \mathbb{R}^n \to \bar{\mathbb{R}}$, $k \in \mathbb{N}$, eine Folge integrierbarer Funktionen. Für alle $k \in \mathbb{N}$ gelte $f_k(x) \le f_{k+1}(x)$ für fast alle $x \in \mathbb{R}$. Es sei*

$$\int_{\mathbb{R}^n} f_k(x)\, dx \le M < \infty,$$

für ein $M \in \mathbb{R}$. Dann ist $f = \lim_{k\to\infty} f_k$ integrierbar und

$$\int_{\mathbb{R}^n} f(x)\, dx = \lim_{k\to\infty} \int_{\mathbb{R}^n} f_k(x)\, dx.$$

Beweis. Da nach Satz IX-5.9 jede Funktion f_k die Werte $\pm\infty$ nur auf einer Nullmenge N_k annehmen kann und $N = \cup_k N_k$ ebenfalls eine Nullmenge ist, kann man nach Satz IX-5.8 ohne Beschränkung der Allgemeinheit $f_k\colon \mathbb{R}^n \to \mathbb{R}$ annehmen. Ebenso können wir nach einer weiteren Modifikation auf einer Nullmenge davon ausgehen, daß $f_k(x) \le f_{k+1}(x)$ für alle $k \in \mathbb{N}$ und für alle $x \in \mathbb{R}$ gilt.
Die Voraussetzungen sichern die Existenz des Grenzwertes $\alpha = \lim_{k\to\infty} \int f_k\,dx$.Wegen der Monotonie der Folge (f_k) existiert der punktweise Grenzwert $f(x) = \lim_{k\to\infty} f_k(x)$, $x \in \mathbb{R}^n$. Die Darstellung

$$\begin{aligned}|(f - f_m)(x)| &= \lim_{k\to\infty} (f_k(x) - f_m(x))\\ &= \lim_{k\to\infty} \sum_{i=m+1}^{k} (f_i(x) - f_{i-1}(x)) = \sum_{i=m+1}^{\infty} (f_i(x) - f_{i-1}(x)),\end{aligned}$$

$m \in \mathbb{N}$, (die Differenz $f(x) - f_m(x)$ ist wegen $f_m(x) \in \mathbb{R}$ auch sinnvoll falls $|f(x)| = \infty$), zusammen mit $f_i - f_{i-1} \ge 0$, $i > m$, und Satz IX-4.4 ergeben

$$\begin{aligned}\int^* |f - f_m|\,dx = \int^* \sum_{i=m+1}^{\infty} (f_i - f_{i-1})\,dx &\le \sum_{i=m+1}^{\infty} \int^* (f_i - f_{i-1})\,dx\\ &= \lim_{k\to\infty} \int^* f_k\,dx - \int^* f_m\,dx = \alpha - \int^* f_m\,dx.\end{aligned}$$

Für $\varepsilon > 0$ gilt daher

$$\int^* |f - f_m|\,dx \le \alpha - \int^* f_m\,dx < \varepsilon$$

für m hinreichend groß. Da f_m integrierbar ist, gibt es eine Funktion $\varphi_m \in C_c(\mathbb{R}^n)$ mit

$$\int^* |f_m - \varphi_m|\,dx < \frac{\varepsilon}{2}$$

und folglich

$$\int^* |f - \varphi_m|\,dx < \varepsilon,$$

d.h. f ist integrierbar und $\int f\,dx = \lim_{m\to\infty} \int_{\mathbb{R}^n} \varphi_m\,dx$. Die Behauptung ergibt sich nun aus der Abschätzung

$$\begin{aligned}|\alpha - \int \varphi_m\,dx| &\le |\alpha - \int f_m\,dx| + |\int f_m\,dx - \int \varphi_m\,dx|\\ &\le |\alpha - \int f_m\,dx| + \int |f_m - \varphi_m|\,dx < \varepsilon.\end{aligned}$$

$\square$

Wir betrachten nun einen nützlichen Spezialfall:

Korollar IX-6.2. Es seien f_k, $k \in \mathbb{N}$, und $g\colon \mathbb{R}^n \to \bar{\mathbb{R}}$ integrierbare Funktionen.
1.) Ist $f_k \leq g$ fast überall, $k \in \mathbb{N}$, dann ist $\sup_k f_k$ integrierbar und

$$\sup_k \int f_k \, dx \leq \int \sup_k f_k \, dx.$$

2.) Ist $f \geq g$ fast überall, $k \in \mathbb{N}$, dann ist $\inf_k f_k$ integrierbar und

$$\inf_k \int f_k \, dx \geq \int \inf_k f_k \, dx.$$

(Wir erinnern daran, daß $\sup_k f_k$, bzw. $\inf_k f_k$ die obere, respektive untere Einhüllende der Funktionen (f_k) bezeichnet.

Beweis. Wir modifizieren vorerst die Funktionen auf einer Nullmenge derart, daß f_k, $g\colon \mathbb{R}^n \to \mathbb{R}$ und $f_k(x) \leq g(x)$ für alle $x \in \mathbb{R}^n$ und $k \in \mathbb{N}$ gilt. Wir betrachten die Funktionen

$$g_l = \max(f_1, \ldots, f_l).$$

Es folgt

$$(*) \qquad g_l \leq g_{l+1} \leq g, \quad l \in \mathbb{N}.$$

Somit existiert für alle $x \in \mathbb{R}^n$ der Grenzwert $h(x) = \lim_{l\to\infty} g_l(x)$. Wir zeigen

$$h = \sup_k f_k$$

Aus

$$f_k \leq g_k \leq h$$

folgt

$$\sup_k f_k \leq h.$$

Angenommen es wäre $\sup_k f_k(x_0) < h(x_0) - \varepsilon$ (man beachte $h(x_0) \leq g(x_0) \in \mathbb{R}$) für ein $x_0 \in \mathbb{R}^n$ und $\varepsilon > 0$. Dann wäre $f_k(x_0) < h(x_0) - \varepsilon$ für alle $k \in \mathbb{N}$ und daher auch $g_l(x_0) < h(x_0) - \varepsilon$ für alle $l \in \mathbb{N}$, im Widerspruch zur Definition von h. Aus $(*)$ erhält man die Beschränktheit der Integrale

$$\int g_l \, dx \leq \int g \, dx, \quad l \in \mathbb{N}.$$

Nach dem Satz von der monotonen Konvergenz ist daher h und somit auch $\sup_k f_k$ integrierbar. Dann gilt für jedes $k \in \mathbb{N}$

$$\int f_k \, dx \leq \int \sup_j f_j \, dx,$$

und somit auch

$$\sup_k \int f_k \, dx \leq \int \sup_k f_k \, dx.$$

Ersetzt man f_k und g durch $-f_k$ und $-g$, ergibt sich die zweite Behauptung. $\square$

Das nächste Resultat ist eine weitgehende Verallgemeinerung des Satzes von Arzela-Osgood.

THEOREM IX-6.3 (H.Lebesgue, majorisierte Konvergenz). *Es sei* $f_k\colon \mathbb{R}^n \to \bar{\mathbb{R}}$*,* $k \in \mathbb{N}$*, eine Folge integrierbarer Funktionen, welche fast überall punktweise gegen eine Funktion* $f\colon \mathbb{R}^n \to \bar{\mathbb{R}}$ *konvergiert. Gibt es eine integrierbare Funktion* $F\colon \mathbb{R}^n \to [0,\infty]$*, so daß für alle* $k \in \mathbb{N}$

$$|f_k(x)| \leq F(x),$$

für fast alle $x \in \mathbb{R}^n$ *gilt, dann ist* f *integrierbar und*

$$\int_{\mathbb{R}^n} f(x)\,dx = \lim_{k\to\infty} \int_{\mathbb{R}^n} f_k(x)\,dx.$$

BEWEIS. Nach einer Modifikation der Funktionen auf einer Nullmenge kann man $f_k\colon \mathbb{R}^n \to \mathbb{R}$ und $F\colon \mathbb{R}^n \to [0,\infty)$ annehmen, und die Gültigkeit der Majorisierung $f_k(x) \leq F(x)$ für alle $x \in \mathbb{R}^n$ und $k \in \mathbb{N}$ voraussetzen. Für $i \in \mathbb{N}$ setzen wir

$$g_i = \sup_{k\geq i} f_k.$$

Nach Korollar IX-6.2 sind die Funktionen g_i integrierbar und

$$(\dagger) \qquad \sup_{k\geq i} \int f_k\,dx \leq \int \sup_{k\geq i} f_k\,dx, \quad i \in \mathbb{N}.$$

Berücksichtigt man $-F \leq g_{i+1} \leq g_i$, $i \in \mathbb{N}$, sichert der Satz von der monotonen Konvergenz die Integrierbarkeit von

$$\limsup_{k\to\infty} f_k = \lim_{i\to\infty} g_i$$

(die Grenzwerte sind punktweise zu verstehen) und

$$\int \limsup_{k\to\infty} f_k\,dx = \lim_{i\to\infty} \int g_i\,dx = \lim_{i\to\infty} \int \sup_{k\geq i} f_k\,dx.$$

Zusammen mit (†) erhält man daher

$$\limsup_{k\to\infty} \int f_k\,dx = \lim_{i\to\infty} \sup_{k\geq i} \int f_k\,dx \leq \lim_{i\to\infty} \int \sup_{k\geq i} f_k\,dx = \int \limsup_{k\to\infty} f_k\,dx.$$

Ersetzt man g_i durch $g_i = \inf_{k\geq i} f_k$ und berücksichtigt $g_i \leq g_{i+1} \leq F$, $i \in \mathbb{N}$, erhält man auf analoge Weise die Integrierbarkeit von $\liminf_{k\to\infty} f_k$ und

$$\liminf_{k\to\infty} \int f_k\,dx \geq \int \liminf_{k\to\infty} f_k\,dx.$$

Da die Funktionenfolge (f_k) punktweise gegen f konvergiert, folgt

$$\liminf_{k\to\infty} f_k(x) = \limsup_{k\to\infty} f_k(x) = \lim_{k\to\infty} f_k(x) = f(x), \quad x \in \mathbb{R}^n.$$

Somit ist f integrierbar und es gilt

$$\limsup_{k\to\infty} \int f_k\,dx \leq \int f\,dx \leq \liminf_{k\to\infty} \int f_k\,dx,$$

also

$$\lim_{k\to\infty}\int f_k\,dx = \int f\,dx.$$

□

Als Anwendung untersuchen wir die Abhängigkeit des Integrals vom Integrationsbereich.

THEOREM IX-6.4. *Es sei $a_1 \subset A_2 \subset \ldots$ eine aufsteigende Mengenfolge und $A = \cup_{k=1}^{\infty} A_k$. Die Abbildung $f\colon A \to \mathbb{R}$ sei integrierbar über alle A_k, $k \in \mathbb{N}$. Dann sind äquivalent:*
1.) Die folge der Integrale $\int_{A_k} |f|\,dx$ ist beschränkt.
2.) f ist über A integrierbar.
Es gilt

$$\int_A f\,dx = \lim_{k\to\infty}\int_{A_k} f\,dx.$$

BEWEIS. 2.) $\Rightarrow$ 1.) ist eine unmittelbare Folge von $|f|\chi_{A_k} \le |f|\chi_A$ und der Monotonie des Integrals.
1.) $\Rightarrow$ 2.) Es sei $f_k = f\chi_{A_k}$ und $\tilde f$ die triviale Fortsetzung von f außerhalb von A. Wir zeigen vorerst, daß $|\tilde f|$ integrierbar ist: wegen $\chi_{A_k} \uparrow \chi_A$ gilt auch $|f_k| \uparrow |\tilde f|$. Nach Voraussetzungist die Folge der Integrale $\int_{\mathbb{R}^n} |f_k|\,dx = \int_{A_k} |f|\,dx$ beschränkt. Aus dem Satz von der monotonen Konvergenz ergibt sich nun die Integrierbarkeit von $|\tilde f|$. Da $|f_k| \le |\tilde f|$ für alle $k \in \mathbb{N}$ gilt und f_k punktweise gegen $\tilde f$ konvergiert, folgt aus dem Satz von der majorisierten Konvergenz die Integrierbarkeit von $\tilde f$, d.h. f ist über A integrierbar. Ferner gilt

$$\int \tilde f\,dx = \int_A d\,dx = \lim_{k\to\infty}\int f_k\,dx = \lim_{k\to\infty}\int_{A_k} f\,dx.$$

□

7. Die L^p-Räume

DEFINITION IX-7.1. *Für jede reelle Zahl $p \ge 1$ und für jede Funktion $f\colon \mathbb{R}^n \to \bar{\mathbb{R}}$ setzen wir*

$$||f||_{L^p} = (\int^* |f|^p\,dx)^{1/p} \in [0,\infty].$$

Wir zeigen, daß $||\cdot||_{L^p}$ fast alle Eigenschaften einer Norm hat.

THEOREM IX-7.2 (Hölder Ungleichung). *Es seien p, $q > 1$ konjugierte Exponenten, also $\frac{1}{p} + \frac{1}{q} = 1$, und f, $g\colon \mathbb{R}^n \to \bar{\mathbb{R}}$. Dann gilt*

$$||fg||_{L^1} \le ||f||_{L^p}||g||_{L^q}.$$

Beweis. Die Ungleichung ist trivial falls $||f||_{L^p}||g||_{L^q} = \infty$. Ist beispielseise $||f||_{L^p} = 0$, dann ist nach Satz IX-5.10 $f = 0$ fast überall, also $||fg||_{L^1} = 0$. Ohne Beschränkung der Allgemeinheit können wir also $||f||_{L^p}|, ||g||_{L^q}| \in (0,\infty)$ annehmen. Für die Funktionen

$$\varphi = \frac{|f|^p}{||f||^p_{L^p}}, \qquad \text{und} \qquad \psi = \frac{|g|^q}{||g||^q_{L^q}}$$

folgt dann

$$(*) \qquad \int^* \varphi\, dx = 1, \qquad \text{und} \qquad \int^* \psi\, dx = 1.$$

Nach der Ungleichung zwischen dem arithmetischen und geometrischen Mittel Korrollar **??** gilt für alle $a \geq 0$ und $b \geq 0$

$$a^{1/p}b^{1/q} \leq \frac{a}{p} + \frac{b}{q},$$

also

$$\frac{|f|}{||f||_{L^p}}\frac{|g|}{||g||_{L^q}} = \varphi^{1/p}\psi^{1/q} \leq \frac{1}{p}\varphi\frac{1}{q}\psi.$$

Bildet man nun das Oberintegral,folgt aus $(*)$

$$\frac{1}{||f||_{L^p}}\frac{1}{||g||_{L^q}}\int^* |fg|\, dx \leq 1,$$

d.h.

$$\int^* |fg|\, dx \leq ||f||_{L^p}||g||_{L^q}.$$

□

Man beachte die Analogie zum Beweis der Hölderschen Ungleichung für endliche Summen.

Theorem IX-7.3 (Minkowski Ungleichung). *Für alle Funktionen f, $g\colon \mathbb{R}^n \to \bar{\mathbb{R}}$ und $p \geq 1$ gilt*

$$||f+g||_{L^p} \leq ||f||_{L^p} + ||g||_{L^p}.$$

(Wir nehmen an, daß f und g nicht gleichzeitig die Werte ∞ und $-\infty$ annehmen).

Beweis. Für $p = 1$ ergibt sich die Ungleichung aus Lemma IX-4.3. Für $p > 1$ verläuft der Beweis vollkommen analog zum Beweis der Minkowski Ungleichung für endliche Summen, sofern man die Summation durch das Oberintegral ersetzt. □

Definition IX-7.4. *1.) Eine Funktion $f\colon \mathbb{R}^n \to \mathbb{R}$ heißt* ***lokal integrierbar****, wenn f über jede kompakte Menge $K \subset \mathbb{R}^n$ integrierbar ist. Wir schreiben*

$$\mathcal{L}^1_{loc}(\mathbb{R}^n) = \{f\colon \mathbb{R}^n \to \mathbb{R},\ f \textit{ ist lokal integrierbar}\}.$$

2.) Für alle $p \geq 1$ setzt man

$$\mathcal{L}^p(\mathbb{R}^n) = \{f \in \mathcal{L}^1_{loc}(\mathbb{R}^n)\colon ||f||_{L^p} < \infty\}.$$

Bemerkung IX-7.5. 1.) Stetige Funktionen sind lokal integrierbar.
2.) Die lokale Integrierbarkeit kann auch dadurch charakterisiert werden, daß es für alle $x \in \mathbb{R}^n$ eine offene Umgebung U_x gibt, so daß $f|_{U_x}$ integrierbar ist.
3.) Die Funktionen in $\mathcal{L}^p(\mathbb{R}^n)$ werden durch zwei unterschiedliche Bedingungen charakterisiert: $f \in \mathcal{L}^1_{loc}(\mathbb{R}^n)$ verlangt, daß für jede kompakte Menge K das Ober- und Unterintegral von $f|_K$ übereinstimmen und der gemeinsame Wert endlich ist. Die Bedingung $||f||_{L^p} < \infty$ fordert, daß das Oberintegral von $|f|^p$ endlich ist. Es wird später gezeigt, daß aus beiden Bedingungen folgt, daß $|f|^p$ sogar integrierbar ist. Es ist jedoch nicht möglich, $\mathcal{L}^p(\mathbb{R}^n)$ allein durch die Forderung der Integrierbarkeit von $|f|^p$ zu charakterisieren, da dies lediglich eine Bedingung für den Betrag der Funktion darstellt. Eine weitere Bedingung an f selbst, wie zum Beispiel die lokale Integrierbarkeit ist notwendig.

Wir geben nun eine einfache hinreichende Bedingung für $||f||_{L^p} < \infty$ an.

Theorem IX-7.6. *Es sei $f \in \mathcal{L}^1(\mathbb{R}^n)$ beschränkt. Dann ist auch $|f|^p \in \mathcal{L}^1(\mathbb{R}^n)$ für alle $p \geq 1$.*

Beweis. Da f integrierbar ist, ist auch $|f|$ integrierbar. Somit kann man $f \geq 0$ annehmen. Ohne Beschränkung der Allgemeinheit sei auch $f \geq 1$. Zu jedem $\varepsilon > 0$ existiert $\varphi \in C_c(\mathbb{R}^n)$ mit
$$\int^* |f - \varphi|\, dx < \varepsilon.$$
Setzt man $\psi = \min(\varphi^+, 1) \in C_c(\mathbb{R}^n)$ und berücksichtigt
$$|f - \psi| \leq |f - \varphi|$$
folgt
$$\int^* |f - \psi|\, dx < \varepsilon.$$
Wendet man den Mittelwertsatz auf die Funktion $x \to x^p$, $x \in [0,1]$ an, erhält man die Ungleichung
$$|a^p - b^p| \leq p|a - b|$$
für a, $b \in [0,1]$ und somit
$$|f^p - \psi^p| \leq p|f - \psi|,$$
also
$$\int^* |f^p - \psi^p|\, dx \leq p \int^* |f - \psi|\, dx < p\varepsilon.$$
Dies zeigt die Integrierbarkeit von f^p. □

$\mathcal{L}^1(\mathbb{R}^n)$ wurde bereits in Definition IX-4.5 eingeführt. Wir zeigen nun, daß die beiden Definitionen konsistent sind.

Theorem IX-7.7. *Es sei $f\colon \mathbb{R}^n \to \mathbb{R}$. Äquivalent sind folgende Aussagen.*
1.) f ist integrierbar.
2.) $f \in \mathcal{L}^1_{loc}(\mathbb{R}^n)$ und $||f||_{L^1} < \infty$.

Beweis. 1.)$\Rightarrow$ 2.) folgt aus Satz IX-4.7 und Satz IX-4.9.
2.)$\Rightarrow$ 1.) Wir betrachten die folge der kompakten Kugeln $A_k = \bar{K}(0,k)$, $k \in \mathbb{N}$. Dann ist $\mathbb{R}^n = \cup_k A_k$, $|f|_{A_k}$ ist integrierbar und

$$\int_{A_k} |f|\,dx \leq \int_{\mathbb{R}^n} |f|\,dx < \infty.$$

Die Integrierbarkeit von f folgt nun aus Satz IX-6.4. $\square$

Theorem IX-7.8. *Es sei $p \geq 1$ und $f \in \mathcal{L}^p(\mathbb{R}^n)$. Dann ist $|f|^p \in \mathcal{L}^1(\mathbb{R}^n)$.*

Beweis. Da mit f auch $|f|$ in $\mathcal{L}^1_{loc}(\mathbb{R}^n)$ liegt, kann man ohne Beschränkung der Allgemeinheit $f \geq 0$ voraussetzen. Für $k \in \mathbb{N}$ sei

$$f_k(x) = \begin{cases} \min(f(x),k), & ||x|| \leq k, \\ 0, & ||x|| > k. \end{cases}$$

Da $f \pm k$ lokal integrierbar ist und $\min(f,k) = \frac{1}{2}(f + k - |f - k|)$, folgt $\min(f,k) \in \mathcal{L}^1_{loc}(\mathbb{R}^n)$, somit $f_k \in \mathcal{L}^1(\mathbb{R}^n)$. Wegen der Beschränktheit von f_k^p ergibt Satz IX-7.6 die Integrierbarkeit von f_k^p und daher

$$\int f_k^p\,dx \leq \int^* f^p\,dx.$$

Wegen $f_k^p \uparrow f^p$ ist f^p nach dem Satz von der monotonen Konvergenz integrierbar. $\square$

Sind f, $g \in \mathcal{L}^1_{loc}(\mathbb{R}^n)$ und ist $K \subset \mathbb{R}^n$ kompakt, dann sind $f|_K$. $(\alpha f)|_K$ für $\alpha \in \mathbb{R}$ und $g|_K$ und somit auch $f|_K + g|_K = (f+g)|_K$ integrierbar. Die Menge $\mathcal{L}^1_{loc}(\mathbb{R}^n)$ ist daher ein Vektorraum. Die Minkowksi Ungleichung sichert $(f+g) \in \mathcal{L}^p(\mathbb{R}^n)$ für f und $g \in \mathcal{L}^p(\mathbb{R}^n)$. Da dann auch αf in $\mathcal{L}^p(\mathbb{R}^n)$ liegt, ist $\mathcal{L}^p(\mathbb{R}^n)$ ein Vektorraum für alle $p \geq 1$. Die Einschränkung von $||\cdot||_{L^p}$ auf $\mathcal{L}^p(\mathbb{R}^n)$ hat offenbar folgende Eigenschaften:

(1) $||f||_{L^p} \geq 0$ und $f = 0 \Rightarrow ||f||_{L^p} = 0$.
(2) $||\alpha f||_{L^p} = |\alpha|||f||_{L^p}$ für alle $\alpha \in \mathbb{R}$.
(3) $||f+g||_{L^p} \leq ||f||_{L^p} + ||g||_{L^p}$.

Allerdings folgt aus $||f||_{L^p} = 0$ nicht $f = 0$, sondern nur $f = 0$ fast überall. Somit definiert $||\cdot||_{L^p}$ keine Norm auf $\mathcal{L}^p(\mathbb{R}^n)$. Man nennt eine Funktion auf einem Vektorraum mit den Eigenschaften 1) - 3) eine **Halbnorm** . Die Menge

$$\mathcal{N} = \{f \in \mathcal{L}^p(\mathbb{R}^n) \colon f = 0 \text{ fast überall}\}$$

ist ein Unterraum von $\mathcal{L}^p(\mathbb{R}^n)$.

Definition IX-7.9. *Für $p \geq 1$ setzen wir*

$$L^p(\mathbb{R}^n) = \mathcal{L}^p(\mathbb{R}^n)/\mathcal{N}.$$

Die Elemente $[f]$ von $L^p(\mathbb{R}^n)$ sind Äquivalenzklassen von Funktionen in $\mathcal{L}^p(\mathbb{R}^n)$. Eine Funktion $g \in \mathcal{L}^p(\mathbb{R}^n)$ liegt genau dann in $[f]$, wenn $f - g \in \mathcal{N}$, d.h. wenn $f = g$

fast überall gilt. Da das Lebesgue Integral eine lineare Abbildung von $\mathcal{L}^p(\mathbb{R}^n)$ nach $\mathbb{R}$ darstellt, induziert es eine lineare Abbildung auf $L^p(\mathbb{R}^n)$, $[f] \to \int [f]\,dx$ durch

$$\int [f]\,dx = \int h\,dx$$

für einen beliebigen Repräsentanten $h \in [f]$. Wegen Satz IX-5.8 ist $\int [f]\,dx$ tatsächlich unabhängig von der Wahl von $h \in [f]$. Aus diesem Grunde ist auch

$$||[f]||_{L^p} = ||h||_{L^p} \quad \text{für ein} \quad h \in [f]$$

wohldefiniert und hat die Eigenschaften einer Halbnorm. Allerdings folgt nun aus $||[f]||_{L^p} = 0$, daß ein (und somit jeder) Repräsentant von $[f]$ fast überall Null ist, d.h. $[f] = \mathcal{N}$. Da $\mathcal{N}$ das Nullelement in $L^p(\mathbb{R}^n)$ darstellt, definiert $||\cdot||_{L^p}$ für alle $p \geq 1$ eine Norm auf $L^p(\mathbb{R}^n)$. Somit ist $(L^p(\mathbb{R}^n), ||\cdot||_{L^p})$ ein normierter Raum. Eine Folge von Äquivalenzklassen $([f_k])$ konvergiert genau dann in $(L^p(\mathbb{R}^n), ||\cdot||_{L^p})$ gegen $[f]$, wenn

$$||[f_k] - [f]||_{L^p} = ||[f_k - f]||_{L^p} = (\int^* |f_k - f|^p\,dx)^{1/p} \underset{k\to\infty}{\to} 0.$$

Vereinbart man nun, daß eine Folge von Funktionen $(f_k) \subset \mathcal{L}^p(\mathbb{R}^n)$ genau dann gegen $f \in \mathcal{L}^p(\mathbb{R}^n)$ konvergiert, wenn

$$\lim_{k\to\infty} (\int^* |f_k - f|^p\,dx)^{1/p} = 0$$

folgt

$$[f_k] \to [f] \quad \text{in } L^p(\mathbb{R}^n) \Leftrightarrow f_k \to f \quad \text{in } \mathcal{L}^p(\mathbb{R}^n).$$

Bemerkung IX-7.10. Meist verzichtet man auf die Unterscheidung zwischen der Äquivalenzklasse $[f]$ und dem Repräsentanten f und schreibt

$$\int [f]\,dx = \int f\,dx.$$

Wir greifen nun noch einmal die Höldersche Ungleichung auf.

Theorem IX-7.11 (Höldersche Ungleichung). *Es sei $f \in \mathcal{L}^p(\mathbb{R}^n)$, $g \in \mathcal{L}^q(\mathbb{R}^n)$, $1 < p, q < \infty$ und $\frac{1}{p} + \frac{1}{q} = 1$. Dann ist $fg \in \mathcal{L}^1(\mathbb{R}^n)$ und es gilt*

$$|\int fg\,dx| \leq ||f||_{L^p} ||g||_{L^q}.$$

Beweis. Wegen Satz IX-7.2 ist nur $fg \in \mathcal{L}^1(\mathbb{R}^n)$ zu zeigen. Wegen $fg = (f^+ - f^-)(g^+ - g^-)$ kann man $f \geq 0$ und $g \geq 0$ annehmen. Wegen Satz IX-7.8 sind f^p und g^q integrierbar. Definiert man für $k \in \mathbb{N}$ die Funktionen f_k und g_k wie im Beweis von Satz IX-7.8, folgt aus Satz IX-4.9 die Integrierbarkeit von $f_k g_k$. Ferner gilt $f_k g_k \uparrow fg$ und

$$\int f_k g_k\,dx \leq \int^* fg\,dx \leq ||f||_{L^p}\,||g||_{L^q}$$

(Satz IX-7.2). Eine erneute Anwendung des Satzes von der monotonen Konvergenz ergibt $fg \in \mathcal{L}^1(\mathbb{R}^n)$. $\square$

Der Satz von der majorisierten Konvergenz gilt auch in $\mathcal{L}^p(\mathbb{R}^n)$.

LEMMA IX-7.12. *Es sei $p \geq 1$ und $(f_k) \subset \mathcal{L}^p(\mathbb{R}^n)$ eine Funktionenfolge, welche fast überall gegen eine Funktion $f\colon \mathbb{R}^n \to \mathbb{R}$ konvergiert. Ferner sei $F\colon \mathbb{R}^n \to [0,\infty]$ mit $||F||_{L^p} < \infty$ und*

$$|f_k| \leq F \quad \textit{für alle } k \in \mathbb{N}.$$

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

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

BEWEIS. Für jede kompakte Teilmenge $K \subset \mathbb{R}^n$ gilt $f_i\chi_K \in \mathcal{L}^1(\mathbb{R}^n)$ da $\mathcal{L}^p(\mathbb{R}^n) \subset \mathcal{L}^1_{loc}(\mathbb{R}^n)$. Ferner gilt

$$f_i\chi_K \to f\chi_K \quad \text{fast überall}$$

und

$$|f_i\chi_K| \leq F\chi_K, \quad i \in \mathbb{N}.$$

Da χ_K integrierbar (K ist kompakt) und beschränkt ist, folgt aus Satz IX-7.6 $\chi_K \in \mathcal{L}^p(\mathbb{R}^n)$. Somit ergibt die Höldersche Ungleichung

$$\int^* F\chi_K \, dx \leq ||F||_{L^p}\lambda(K)^{1/q} < \infty.$$

Daher kann man auf die Folge $(f_i\chi_K)$ den Satz von der majorisierten Konvergenz anwenden und auf $f\chi_K \in \mathcal{L}^1(\mathbb{R}^n)$ schließen. Es folgt $f \in \mathcal{L}^1_{loc}(\mathbb{R}^n)$. Nach Abänderung auf einer Nullmenge gilt außerdem

$$|f| \leq F$$

und somit

$$\int^* |f|^p \, dx \leq \int^* F^p \, dx.$$

Dies zeigt $f \in \mathcal{L}^p(\mathbb{R}^n)$. Aus der Abschätzung

$$|f_k - f|^p \leq (|f_k| + |f|)^p \leq 2^p F^p$$

und $\lim_{k\to\infty} |f_k - f|^p = 0$ fast überall erhält man schließlich wieder mit Hilfe des Satzes von der majorisierten Konvergenz

$$\lim_{k\to\infty} \int |f_k - f|^p \, dx = 0.$$

□

Wir senden uns nun der Vollständigkeit der Räume $(L^p(\mathbb{R}^n), ||\cdot||_{L^p})$ zu. Ausgangspunkt ist folgendes Resultat für Reihen, welches zum ersten Mal einen Zusammenhang zwischen der Konvergenz in $\mathcal{L}^p(\mathbb{R}^n)$ und der punktweisen Konvergenz aufzeigt.

Lemma IX-7.13. *Es sei $p \geq 1$ und $g_k \in \mathcal{L}^p(\mathbb{R}^n)$ eine Folge von Funktionen mit*

$$M\colon = \sum_{k=1}^{\infty} ||g_k||_{L^p} < \infty.$$

Dann konvergiert die Reihe $\sum_{k=1}^{\infty} g_k$ fast überall absolut gegen eine Funktion $g \in \mathcal{L}^p(\mathbb{R}^n)$ und es gilt

$$\lim_{k\to\infty} ||g - \sum_{i=1}^{k} g_i||_{L^p} = 0.$$

Beweis. Wir setzen

$$G_k = \sum_{i=1}^{k} |g_i| \qquad \text{und} \qquad G = \sum_{i=1}^{\infty} |g_i|.$$

Dann ist $G_k \in \mathcal{L}^p(\mathbb{R}^n)$ und somit $G_k^p \in \mathcal{L}^1(\mathbb{R}^n)$. Die Minkowski Ungleichung ergibt für $k \in \mathbb{N}$

$$\int G_k^p \, dx = ||G_k||_{L^p}^p = ||\sum_{i=1}^{k} |g_i|||_{L^p} \leq \sum_{i=1}^{k} ||g_i||_{L^p} \leq M^p.$$

Wegen $G_k^p \uparrow G^p$ folgt aus dem Satz von der monotonen Konvergenz $G^p \in \mathcal{L}^1(\mathbb{R}^n)$. Nach Satz IX-5.9 nimmt G^p und damit auch G den Wert ∞ höchstens auf einer Nullmenge N an, d.h. $G(x) = \sum_{k=1}^{\infty} |g_k(x)| < \infty$ für alle $x \in \mathbb{R}^n \setminus N$. Somit konvergiert die Reihe absolut auf $\mathbb{R}^n \setminus N$. Auf N setzen wir g_i, $i \in \mathbb{N}$, und g gleich Null. Dann gilt für alle $x \in \mathbb{R}^n$

$$|\sum_{i=1}^{k} g_i(x)| \leq G_k(x) \leq G(x).$$

Da $G^p \in \mathcal{L}^1(\mathbb{R}^n)$ und somit $G \in \mathcal{L}^p(\mathbb{R}^n)$ liegt, folgt aus Lemma IX-7.12

$$\lim_{k\to\infty} ||g - \sum_{i=1}^{k} g_i||_{L^p} = 0.$$

□

Theorem IX-7.14. *Es sei $p \geq 1$ und $(f_k) \subset \mathcal{L}^p(\mathbb{R}^n)$ eine Cauchy Folge in $\mathcal{L}^p(\mathbb{R}^n)$. Dann gilt*
1.) Es existiert eine Teilfolge $(f_{\varphi(k)}) \subset (f_k)$ welche fast überall gegen eine Grenzfunktion $f\colon \mathbb{R}^n \to \mathbb{R}$ konvergiert.
2.) Die Grenzfunktion f liegt in $\mathcal{L}^p(\mathbb{R}^n)$ und es gilt

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

Beweis. Da (f_k) eine Cauchy Folge in $\mathcal{L}^p(\mathbb{R}^n)$ ist, gibt es für alle $k \in \mathbb{N}$ Indices $\varphi(k) \in \mathbb{N}$ mit $\varphi(k+1) > \varphi(k)$ und

$$||f_{\varphi(k)} - f_{\varphi(k+1)}||_{L^p} < 2^{-k}, \qquad f_{\varphi(0)} = 0.$$

Wegen Lemma IX-7.13 konvergiert die Reihe $\sum_{k\to\infty}(f_{\varphi(k)} - f_{\varphi(k+1)})$ fast überall gegen eine Grenzfunktion $f \in \mathcal{L}^p(\mathbb{R}^n)$. Insbesondere gilt

$$f = \lim_{k\to\infty} \sum_{k\to k} (f_{\varphi(i)} - f_{\varphi(i+1)}) = \lim_{k\to\infty} f_{\varphi(k)}$$

fast überall und

$$\lim_{k\to\infty} ||f_{\varphi(k)} - f||_{L^p} = 0.$$

Da (f_k) eine Cauchy Folge ist, folgt daraus

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

□

KOROLLAR IX-7.15. Die normierten Räume $(L^p(\mathbb{R}^n), ||\cdot||_{L^p})$ sind für $p \geq 1$ vollständig.

KOROLLAR IX-7.16. Es seien $p \geq 1$, f, $f_k \in \mathcal{L}^p(\mathbb{R}^n)$, $k \in \mathbb{N}$, mit

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

Dann gibt es eine Teilfolge $(f_{\varphi(k)}) \subset (f_k)$ welche fast überall gegen f konvergiert.

BEWEIS. Die Folge (f_k) ist eine Cauchy Folge in $\mathcal{L}^p(\mathbb{R}^n)$. Somit gibt es eine Teilfolge $(f_{\varphi(k)})$, welche fast überall gegen eine Funktion $g \in \mathcal{L}^p(\mathbb{R}^n)$ konvergiert. Außerdem gilt

$$\lim_{k\to\infty} ||f_k - g||_{L^p} = 0.$$

Es folgt

$$||f - g||_{L^p} \leq ||f - f_k||_{L^p} + ||f_k - g||_{L^p} \underset{k\to\infty}{\to} 0,$$

also $f = g$ fast überall. Dann konvergiert $(f_{\varphi(k)})$ auch fast überall gegen f. □

8. Die zentralen Sätze der Integralrechnung

8.1. Der Satz von Fubini.

THEOREM IX-8.1 (Fubini). *Es sei $f\colon \mathbb{R}^k \times \mathbb{R}^m \to \mathbb{R}$ integrierbar. Dann gibt es eine Nullmenge $N \subset \mathbb{R}^m$, so daß für jedes feste $\eta \in \mathbb{R}^m \setminus N$ die Funktion*

$$\begin{aligned} &\mathbb{R}^k \to \mathbb{R} \\ &\xi \mapsto f(\xi, \eta) \end{aligned}$$

integrierbar ist. Setzt man

$$F(\eta) = \begin{cases} \int_{\mathbb{R}^k} f(\xi,\eta)\, d\xi, & \eta \in \mathbb{R}^m \setminus N, \\ 0, & \eta \in N, \end{cases}$$

dann ist die Funktion $F\colon \mathbb{R}^m \to \mathbb{R}$ integrierbar und es gilt mit $x = (\xi, \eta)$

$$\int_{\mathbb{R}^{k+m}} f(x)\, dx = \int_{\mathbb{R}^m} F(\eta)\, d\eta = \int_{\mathbb{R}^m} [\int_{\mathbb{R}^k} f(\xi,\eta)\, d\xi]\, d\eta.$$

Bemerkung IX-8.2. Da die Vertauschung der Variablen $(\xi,\eta)\to(\eta,\xi)$ ein Automorphismus des $\mathbb{R}^{k+m}$ mit Determinante ± 1 ist, folgt mit Satz IX-2.14, der offenbar auch für integrierbare Funktionen gilt,

$$\int_{\mathbb{R}^m}[\int_{\mathbb{R}^k} f(\xi,\eta)\,d\xi]\,d\eta=\int_{\mathbb{R}^k}[\int_{\mathbb{R}^m} f(\xi,\eta)\,d\eta]\,d\xi.$$

In vielen Anwendungen ist eine stetige Funktion über eine kompakte Teilmenge $K\subset\mathbb{R}^n$ zu integrieren. Wir zeigen nun, wie man mit Hilfe des Satzes von Fubini dieses Problem auf die Berechnung von n iterierten Ingeralen in einer Veränderlichen zurückführen kann. Es sei also $f\in C(K,\mathbb{R})$. Nach Definition IX-4.13 gilt

$$\int_K f(x)\,dx=\int_{\mathbb{R}^n}\tilde f(x)\chi_K(x)\,dx,$$

wobei $\tilde f$ die triviale Fortsetzung von f durch Null bezeichnet. Nach Beispiel IX-4.14 und Satz IX-4.9 ist $\tilde f\chi_K$ integrierbar. Partitionieren wir $x\in\mathbb{R}^n$ in der Form $x=(x',t)$, $x'\in\mathbb{R}^{n-1}$, $t\in\mathbb{R}$, folgt aus dem Satz von Fubini (wir können sogar die schwächere Version Satz IX-2.12 verwenden) die Integrierbarkeit von $t\mapsto\tilde f(x',t)\chi_K(x',t)$ für alle $x'\in\mathbb{R}^{n-1}$. Für jedes $x'\in\mathbb{R}^{n-1}$ betrachten wir die Schnittmengen

$$K_{x'}=\{t\in\mathbb{R}\colon (x',t)\in K\}.$$

Die Abbildung $t\mapsto\chi_K(x',t)$ stimmt also mit $\chi_{K_{x'}}$ überein. Nach dem Satz von Fubini gilt ferner

$$(*)\qquad \begin{aligned}\int_K f(x)\,dx&=\int_{\mathbb{R}^{n-1}}[\int_{\mathbb{R}}\tilde f(x',t)\chi_K(x',t)\,dt]\,dx'\\&=\int_{\mathbb{R}^{n-1}}[\int_{K_{x'}}\tilde f(x',t)\,dt]\,dx'.\end{aligned}$$

Setzt man

$$B'=\{x'\in\mathbb{R}^{n-1}\colon K_{x'}\neq\emptyset\},$$

läßt sich $(*)$ schreiben in der suggestiven Form

$$(\dagger)\qquad \int_K f(x)\,dx=\int_{B'}[\int_{K_{x'}} f(x',t)\,dt]\,dx'.$$

Beispiel IX-8.3. Wir berechnen das Volumen des Körpers

$$K=\{(x,y,z)\in\mathbb{R}^3\colon x\geq 0, y\geq 0, z\geq 0, x+y+z\leq\sqrt{2}, x^2+y^2\leq 1\}.$$

Wir setzen im ersten Reduktionsschritt

$$B'=\{(x,y)\in\mathbb{R}^2\colon x\geq 0, y\geq 0, x^2+y^2\leq 1\}$$

und für $(x,y)\in B'$

$$K_{(x,y)}=[0,\sqrt{2}-x-y].$$

Wir erhalten mit (†)

$$V(K) = \int_{\mathbb{R}^3} \chi_K(x,y,z)\,dxdydz = \int_{B'} [\int_0^{\sqrt{2}-x-y} dz]\,dxdy = \int_{B'} (\sqrt{2}-x-y)\,dxdy.$$

Auf das letzte Integral wenden wir wieder (†) an und setzen im zweiten Reduktionsschritt

$$\begin{aligned} B'' &= [0,1], \\ K_x &= [0, \sqrt{1-x^2}], \quad x \in B''. \end{aligned}$$

Dies ergibt

$$\begin{aligned} V(K) &= \int_{B'} (\sqrt{2}-x-y)\,dxdy = \int_{x=0}^1 [\int_{y=0}^{\sqrt{1-x^2}} (\sqrt{2}-x-y)\,dy]\,dx \\ &= \int_{x=0}^1 [\sqrt{2}y - xy - \frac{1}{2}y^2]|_{y=0}^{\sqrt{1-x^2}} dx = \int_0^1 [(\sqrt{2}-x)\sqrt{1-x^2} - \frac{1}{2}(1-x^2)]\,dx \\ &= \frac{\pi}{2\sqrt{2}} - \frac{2}{3}. \end{aligned}$$

Die tragende Voraussetzung des Satzes von Fubini ist die Integrierbarkeit der Abbildung $f\colon \mathbb{R}^k \times \mathbb{R}^m \to \mathbb{R}$, welche manchmal nicht leicht zu verifizieren ist. In solchen Fällen kann der Satz von Tonelli hilfreich sein.

Theorem IX-8.4 (Tonelli). *Es sei $f\colon \mathbb{R}^k \times \mathbb{R}^m \to \mathbb{R}$. Ist für fast alle $\eta \in \mathbb{R}^m$ die Abbildung*

$$\xi \mapsto |f(\xi,\eta)|$$

integrierbar und existiert das iterierte Integral

$$\int_{\mathbb{R}^m} [\int_{\mathbb{R}^k} |f(\xi,\eta)|\,d\xi]\,d\eta < \infty,$$

dann ist $f\colon \mathbb{R}^k \times \mathbb{R}^m \to \mathbb{R}$ integrierbar.

8.2. Konvergenzsätze. Die guten Konvergenzeigenschaften des Lebesgueintegrales gehören zu den wichtigsten Gründen für dessen weite Verbreitung. Wir zitieren zwei zentrale Ergebnisse.

Theorem IX-8.5 (B.Levi, monotone Konvergenz). *Es sei $f_k\colon \mathbb{R}^n \to \mathbb{R}$, $k \in \mathbb{N}$, eine Folge integrierbarer Funktionen mit $f_k \le f_{k+1}$ für alle $k \in \mathbb{N}$. Gilt*

$$M := \lim_{k\to\infty} \int_{\mathbb{R}^n} f_k(x)\,dx < M,$$

dann ist $f = \lim_{k\to\infty} f_k$ integrierbar und es ist

$$\int_{\mathbb{R}^n} f(x)\,dx = \lim_{k\to\infty} \int_{\mathbb{R}^n} f_k(x)\,dx.$$

Theorem IX-8.6 (H.Lebesgue, majorisierte Konvergenz). *Es sei* $f_k\colon \mathbb{R}^n \to \bar{\mathbb{R}}$, $k \in \mathbb{N}$, *eine Folge integrierbarer Funktionen, welche fast überall punktweise gegen eine Funktion* $f\colon \mathbb{R}^n \to \bar{\mathbb{R}}$ *konvergiert. Gibt es eine integrierbare Funktion* $F\colon \mathbb{R}^n \to [0,\infty]$, *so daß für alle* $k \in \mathbb{N}$

$$|f_k(x)| \le F(x),$$

für fast alle $x \in \mathbb{R}^n$ *gilt, dann ist* f *integrierbar und*

$$\int_{\mathbb{R}^n} f(x)\,dx = \lim_{k\to\infty} \int_{\mathbb{R}^n} f_k(x)\,dx.$$

Wir können nun mühelos zeigen, daß Regelfunktionen Lebesgue-integrierbar sind und ihr Lebesgue Integral mit dem Cauchy Integral übereinstimmt. Es sei also $f \in \mathcal{R}(I,\mathbb{R})$, $I = [a,b]$, eine Regelfunktion. Dann gibt es eine Folge von Treppenfunktionen $(t_k) \subset \mathcal{T}(I,\mathbb{R})$ derart, daß $t_k \rightrightarrows f$. Da f beschränkt ist, folgt aus der gleichmäßigen Konvergenz der Folge (t_k) die Existenz einer positiven Konstanten M mit

$$\forall k \in \mathbb{N} \forall x \in I\colon |t_k(x)| \le M.$$

Da I kompakt ist, ist die Funktion $F(x) = M$, $x \in I$, und $F(x) = 0$, $x \in \mathbb{R} \setminus I$, integrierbar. Nach dem Satz von Lebesgue ist daher $\tilde{f}$ integrierbar, also f integrierbar über I und es gilt

$$\int_I f(x)\,dx = \int_{\mathbb{R}^n} \tilde{f}(x)\,dx = \lim_{k\to\infty} \int_{\mathbb{R}^n} \tilde{t}_k(x)\,dx = \lim_{k\to\infty} \int_I t_k(x)\,dx.$$

Auf der rechten Seite steht das Lebesgue Integral von f, die linke Seite definiert gerade das Cauchy Integral von f.
Als weitere Anwendung des Satzes von der majorisierten Konvergenz betrachten wir parameterabhängige Integrale.

Theorem IX-8.7. *Es sei* $f\colon \mathbb{R}^n \times D \to \mathbb{R}$, $D \subset \mathbb{R}^k$ *und* $t_0 \in D$. *Es gelte*

(1) $\forall t \in D\colon f(\cdot,t) \in \mathcal{L}^1(\mathbb{R}^n).$,
(2) *für fast alle* $x \in \mathbb{R}^n$ *sei* $t \mapsto f(x,t)$ *stetig in* t_0,
(3) *es existiert eine integrierbare Funktion* $F\colon \mathbb{R}^n \to [0,\infty]$, *so daß für alle* $t \in D$

$$|f(x,t)| \le F(x)$$

für fast alle $x \in \mathbb{R}^n$ *gilt.*

Dann ist die Abbildung

$$g\colon \begin{cases} I \to \mathbb{R}, \\ t \mapsto \int_{\mathbb{R}^n} f(x,t)\,dx \end{cases}$$

stetig in $t = t_0$.

Beweis. Vgl. Satz VII-VII-6.1. □

Theorem IX-8.8. *Es sei* $I \subset \mathbb{R}$ *ein offenes Intervall und* $f\colon \mathbb{R}^n \times I \to \mathbb{R}$. *Ferner gelte*

(1) $\forall t \in I\colon f(\cdot, t) \in \mathcal{L}^1(\mathbb{R}^n)$,
(2) *für fast alle* $x \in \mathbb{R}^n$ *ist* $t \mapsto f(x,t)$ *differenzierbar auf* I,
(3) *es gibt eine integrierbare Funktion* $F\colon \mathbb{R}^n \to [0,\infty)$, *so daß für alle* $t \in I$

$$|\frac{\partial f}{\partial t}(x,t)| \leq F(x)$$

für fast alle $x \in \mathbb{R}^n$ *gilt.*

Dann ist die Abbildung

$$g\colon \begin{cases} D \to \mathbb{R}, \\ t \mapsto \int_{\mathbb{R}^n} f(x,t)\,dx \end{cases}$$

differenzierbar auf I *und es gilt*

$$g'(t) = \int_{\mathbb{R}^n} \frac{\partial f}{\partial t}(x,t)\,dx.$$

Beweis. Vgl. Satz VII-VII-6.3. □

9. Die Transformationsformel

Definition IX-9.1. *Es seien* U, $V \subset \mathbb{R}^n$ *offene Mengen. Ein Homöomorphismus* $\phi\colon U \to V$ *heißt* ***Diffeomorphismus*** , *wenn* ϕ *und* ϕ^{-1} *stetig differenzierbar sind.*

Theorem IX-9.2 (Transformationssatz). *Es seien* U, $V \subset \mathbb{R}^n$ *offene Mengen und* $\phi\colon U \to V$ *ein Diffeomorphismus. Eine Abbildung* $f\colon V \to \mathbb{R}$ *ist genau dann integrierbar, wenn* $(f \circ \phi)|\det D\phi|$ *über* U *integrierbar ist. Es gilt dann*

$$\int_{\phi(U)} f(y)\,dy = \int_U f(\phi(x))|\det D\phi(x)|\,dx.$$

Korollar IX-9.3 (Ebene Polarkoordinaten). Es sei

$$\phi\colon \begin{cases} [0,\infty) \times [0,2\pi] \to \mathbb{R}^2 \\ (r,\varphi) \to (r\cos\varphi, r\sin\varphi). \end{cases}$$

Eine Abbildung $f\colon \mathbb{R}^2 \to \mathbb{R}$ ist integrierbar genau dann, wenn die Abbildung

$$(r,\varphi) \to r f(\phi(r,\varphi))$$

über $[0,\infty) \times [0,2\pi]$ integrierbar ist. Es gilt dann

$$\int_{\mathbb{R}^2} f(x,y)\,dxdy = \int_0^{2\pi}\int_0^{\infty} f(r\cos\varphi, r\sin\varphi)r\,drd\varphi.$$

Beweis. Wir setzen in Satz IX-9.2

$$U = \{(r,\varphi)\colon r > 0, 0 < \varphi < 2\pi\}$$

und

$$V = \phi(U) = \mathbb{R}^2 \setminus \{(x,0)\colon x \geq 0\}.$$

Dann ist ϕ ein Diffeomorphismus von U auf V. Wegen $|\det D\phi(r,\varphi)| = r$ folgt die Behauptung aus dem Transformationssatz, da ∂U und ∂V Nullmengen sind. □

Beispiel IX-9.4. Wir demonstrieren eine elegante Berechnung des Gauß'schen Integrals $\int_0^\infty e^{-x^2}\,dx = \frac{\sqrt{\pi}}{2}$

Beweis. Wir zeigen zuerst die Existenz des Integrals. Wir gehen aus von den Treppenfunktionen

$$t_n(x) = \begin{cases} 1, & x \in [0,2), \\ \frac{1}{k^2}, & x \in [k, k+1), \quad 2 \le k \le n-1, \\ \frac{1}{n^2}, & x \in [n, n+1] \end{cases}$$

$n \ge 2$. Treppenfunktionen sind integrierbar und es gilt für alle $n \ge 2$

$$\int_0^\infty t_n(x) = 1 + \sum_{i=1}^n \frac{1}{i^2} \le 1 + \sum_{i=1}^\infty \frac{1}{i^2}.$$

Wir definieren nun für $n \ge 2$ die Funktionen

$$f_n(x) = \begin{cases} e^{-x^2}, & x \in [0, n+1], \\ 0, & x > n+1. \end{cases}$$

Diese Funktionen sind integrierbar, vgl. Beispiel IX-4.14, es gilt $f_n \le t_n$ und somit

$$\int_0^{n+1} f_n(x)\,dx \le \int_0^{n+1} t_n(x)\,dx \le 1 + \sum_{i=1}^\infty \frac{1}{i^2}$$

für alle $n \ge 2$. Berücksichtigt man noch $f_n \uparrow f$ mit $f(x) = e^{-x^2}$, $x \ge 0$, folgt aus dem Satz von B. Levi die Integrierbarkeit von f, d.h

$$\int_0^\infty e^{-x^2}\,dx < \infty.$$

Dann existiert aber auch das iterierte Integral

$$\int_0^\infty [\int_0^\infty e^{-x^2-y^2}\,dx]\,dy = (\int_0^\infty e^{-x^2}\,dx)^2.$$

Nach dem Satz IX-8.4 von Tonelli ist die Abbildung $h\colon [0,\infty) \times [0,\infty) \to \mathbb{R}$, $(x,y) \mapsto e^{-x^2+-y^2}$ integrierbar. Kombiniert man daher den Satz von Fubini und das Korollar IX-9.3, ergibt sich

$$\begin{aligned} (\int_0^\infty e^{-x^2}\,dx)^2 &= \int_0^\infty \int_0^\infty e^{-x^2-y^2}\,dxdy = \int_0^\infty \int_0^{\frac{\pi}{2}} e^{-r^2} r\,dr d\varphi \\ &= \int_0^\infty e^{-r^2} r\,dr \int_0^{\frac{\pi}{2}} d\varphi = \frac{1}{2}\frac{\pi}{2}. \end{aligned}$$

$\square$