Infinitesimalrechnung II[1]

Prof. J. Frehse

http://www.iam.uni-bonn.de/AngewandteAnalysis

Bei diesem Skript handelt es sich
um die korrigierte Version
einer Mitschrift einiger Studierender.
Ihnen sei Lob, Preis und Dank !!

8. Februar 2000

[1]geTEXed von der „3. Reihe“

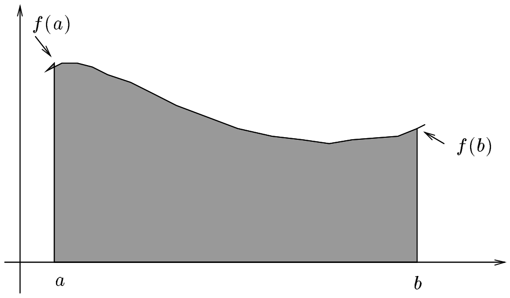
Inhaltsverzeichnis

Kapitel 1

Integralrechnung

1.1 Das eindimensionale Riemannsche Integral

Es sei $f : [a,b] \to \mathbb{R}$, $a, b \in \mathbb{R}$, $a < b$, eine vorgegebene Funktion, deren Graph durch folgende Skizze veranschaulicht wird:



Das Problem, den Flächeninhalt des schraffierten Gebietes zu bestimmen beziehungsweise überhaupt zu definieren, hat historisch zur Entwicklung der Integralrechnung und insbesondere zum Begriff der Riemannschen Integrals geführt.
Es ist einleuchtend, daß man zur Definition des Flächeninhaltes des schraffierten Gebietes dasselbe durch schmale, kleine Rechtecke approximiert, deren Flächeninhalte aufsummiert und anschließend einen Grenzübergang durchführt. Diesen heuristischen Vorgang wollen wir im folgenden exakt durchführen.

Definition:
Es sei $a, b \in \mathbb{R}$ und $a < b$. Eine Zerlegung $\mathfrak{Z}$ *des Intervalls $\overline{I} = [a,b]$ ist ein $(m+1)$-Tupel von reellen Zahlen $x_0, \ldots, x_m$ mit*

$$a = x_0 < x_1 < x_2 < \cdots < x_{m-1} < x_m = b.$$

Wir schreiben auch $\mathfrak{Z} = [x_0, \ldots, x_m]$ und sprechen von einer Zerlegung $\mathfrak{Z}$ bezüglich der Teilungspunkte $x_0, \ldots, x_m$. Die Intervalle $I_i = [x_{i-1}, x_i]$, $i = 1, \ldots, m$, heißen die Teilintervalle

der Zerlegung.
Schließlich heißt die Zahl

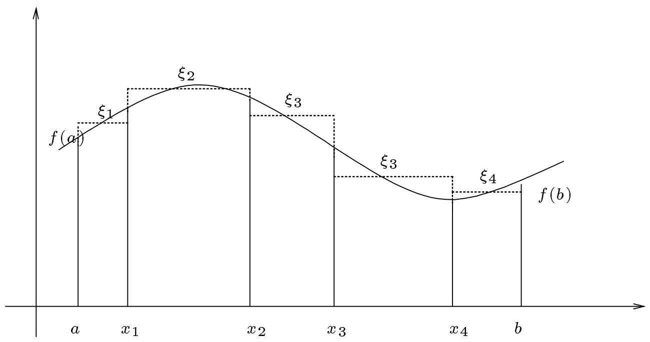
$$|\mathfrak{Z}| = \max_{i=1,\dots,m} (x_i - x_{i-1})$$

die Feinheit *der Zerlegung.*

Definition:
Es sei $\mathfrak{Z} = [x_0, \dots, x_m]$ *eine Zerlegung des Intervalls* $[a, b]$ *und aus jedem der Teilintervalle* $[x_{i-1}, x_i]$ *seien Zahlen* ξ_i *gewählt. Ferner sei eine Funktion* $f\colon [a, b] \to \mathbb{R}$ *gegeben. Dann heißt*

$$\sum(f, \xi_i, \mathfrak{Z}) = \sum_{i=1}^{m} f(\xi_i)(x_i - x_{i-1})$$

die Riemannsche Summe *von* f *zur Zerlegung* $\mathfrak{Z}$ *und den Zwischenpunkten* ξ_i.



Definition:
Gegeben sei die Funktion $f\colon [a, b] \to \mathbb{R}$. f *heißt über* $\overline{I} = [a, b]$ Riemann-integrierbar, *wenn es eine Zahl* A *mit der folgenden Eigenschaft gibt:*
Zu jedem $\varepsilon > 0$ *gibt es ein* $\delta > 0$, *so daß für jede Zerlegung* $\mathfrak{Z}$ *von* $\overline{I}$ *mit* $|\mathfrak{Z}| < \delta$ *und jede Wahl von Zwischenpunkten* ξ_i *zu* $\mathfrak{Z}$ *die Ungleichung*

$$\left|\sum(f, \xi_i, \mathfrak{Z}) - A\right| < \varepsilon$$

gilt. Die Zahl A *nennt man das* Riemannsche Integral *von* f *über* $[a, b]$*; man schreibt dafür* $\int\limits_a^b f\, dx$.

Der in Epsilontik geschulte Leser zieht daraus die Folgerung, daß jede Folge von Riemannschen Summen einer Riemann-integrierbaren Funktion gegen das Riemann-Integral der Funktion konvergiert, wenn die Feinheit der Zerlegungen gegen Null geht.

Unser Ziel ist es zunächst, zu beweisen, daß stetige Funktionen auf einem abgeschlossenen Intervall Riemann-integrierbar sind. Hierzu benötigen wir einige Vorbereitungen.

Definition:
Die Funktion $f\colon [a,b] \to \mathbb{R}$ sei beschränkt, und es sei $\mathfrak{Z} = \mathfrak{Z}(x_0, \dots, x_m)$ eine Zerlegung von $[a,b]$. Die reelle Zahl

$$U_{\mathfrak{Z}} = \sum_{i=1}^{m} \inf\Big\{ f(\xi) \Big| \xi \in [x_{i-1}, x_i] \Big\} (x_i - x_{i-1})$$

heißt Riemannsche Untersumme *der Funktion f bezüglich der Zerlegung $\mathfrak{Z}$, die reelle Zahl*

$$O_{\mathfrak{Z}} = \sum_{i=1}^{m} \sup\Big\{ f(\xi) \Big| \xi \in [x_{i-1}, x_i] \Big\} (x_i - x_{i-1})$$

Riemannsche Obersumme *der Funktion f bezüglich der Zerlegung $\mathfrak{Z}$.*

Es ist klar, daß

$$U_{\mathfrak{Z}} \leq \sum (f, \xi_k, \mathfrak{Z}) \leq O_{\mathfrak{Z}}.$$

Definition:
Die Funktion $f\colon [a,b] \to \mathbb{R}$ sei beschränkt. Dann heißt die reelle Zahl

$$\underline{J} = \sup\Big\{ U_{\mathfrak{Z}} \Big| \ \mathfrak{Z} \quad \textit{ist Zerlegung von} \quad [a,b] \Big\}$$

das über $[a,b]$ erstreckte (Riemannsche) Unterintegral *der Funktion f und die reelle Zahl*

$$\overline{J} = \inf\Big\{ O_{\mathfrak{Z}} \Big| \ \mathfrak{Z} \quad \textit{ist Zerlegung von} \quad [a,b] \Big\}$$

das über $[a,b]$ erstreckte (Riemannsche) Oberintegral *der Funktion f.*

Aus der Annahme der Beschränktheit von f folgt, daß $\underline{J}$ und $\overline{J}$ existieren, d.h. reelle Zahlen sind. Aus $O_{\mathfrak{Z}} \geq U_{\mathfrak{Z}}$ folgt $\overline{J} \geq \underline{J}$.

Es gilt der

Satz 1.1.
Die Funktion $f\colon [a,b] \to \mathbb{R}$ sei beschränkt. Mit den vorangegangenen Bezeichnungen ist f genau dann über $[a,b]$ integrierbar, wenn $\overline{J} = \underline{J}$ ist.
Beweis:

$\Rightarrow$ Die Funktion $f\colon [a,b] \to \mathbb{R}$ sei integrierbar. Es gibt dann eine Zahl A, so daß für jedes vorgegebene $\varepsilon > 0$ und jede Zerlegung $\mathfrak{Z}$ von $[a,b]$ mit Zwischenpunkten ξ_k die Ungleichung

$$\Big| A - \sum (f, \xi_k, \mathfrak{Z}) \Big| < \varepsilon$$

stattfindet, sofern die Feinheit der Zerlegung kleiner als $\delta = \delta(\varepsilon)$ ist. Auflösen des Betrages liefert

$$-\varepsilon < A - \sum_{i=1}^{m} f(\xi_i)(x_i - x_{i-1}) < \varepsilon.$$

Geht man in den einzelnen Summanden zum Supremum bzw. Infimum über, so erhält man

$$-\varepsilon \leq A - \sup\Big\{f(\xi), \xi \in [x_i - x_{i-1}]\Big\}(x_i, x_{i-1}) \leq \varepsilon$$

und

$$-\varepsilon \leq A - \inf\Big\{f(\xi), \xi \in [x_i - x_{i-1}]\Big\}(x_i, x_{i-1}) \leq \varepsilon.$$

Daraus folgt nach Definition der Riemannschen Ober- bzw. Untersumme

$$|A - O_{\mathfrak{Z}}| \leq \varepsilon \quad \text{und} \quad |A - U_{\mathfrak{Z}}| \leq \varepsilon$$

und somit nach der Dreiecksungleichung

$$|O_{\mathfrak{Z}} - U_{\mathfrak{Z}}| \leq 2\varepsilon \quad \text{und} \quad 0 \leq O_{\mathfrak{Z}} - U_{\mathfrak{Z}} \leq 2\varepsilon.$$

Aus den Ungleichungen $O_{\mathfrak{Z}} \geq \overline{J}$ und $U_{\mathfrak{Z}} \leq \underline{J}$ folgt $0 \leq \overline{J} - \underline{J} \leq 2\varepsilon$ für alle $\varepsilon > 0$. Da $\overline{J}$ und $\underline{J}$ nicht von ε abhängen, ergibt sich $\overline{J} = \underline{J}$.

$\Leftarrow$ Es gelte $\overline{J} = \underline{J} = J$.
Den trivialen Fall $f = const.$ schließen wir aus.
Wir zeigen zunächst: Ist $\varepsilon > 0$ vorgegeben, so gibt es eine Zerlegung $\mathfrak{Z}_1$ von $[a, b]$, so daß

$$O_{\mathfrak{Z}_1} - U_{\mathfrak{Z}_1} < \frac{\varepsilon}{3}. \tag{1.1}$$

Dies sieht man so:
Da $\underline{J}$ bzw. $\overline{J}$ nach Definition das Infimum bzw. Supremum der Riemannschen Unter- bzw. Obersummen bedeutet, gibt es Zerlegungen $\mathfrak{U}$ und $\mathfrak{L}$ von $[a, b]$ mit

$$J - U_{\mathfrak{U}} < \frac{\varepsilon}{6} \quad \text{und} \quad O_{\mathfrak{L}} - J < \frac{\varepsilon}{6}. \tag{1.2}$$

(Andernfalls wäre die Extremumseigenschaft von J verletzt.)
Aus den Zerlegungen $\mathfrak{U}$ und $\mathfrak{L}$ bilden wir durch „Zusammenlegen“ eine neue Zerlegung, die wir mit $\mathfrak{UL}$ bezeichnen, indem wir die Mengen der Teilpunkte von $\mathfrak{U}$ und $\mathfrak{L}$ vereinigen und die Elemente der Vereinigungsmenge der Größe nach anordnen und numerieren.
Es gilt $U_{\mathfrak{U}} \leq U_{\mathfrak{UL}}$ und $O_{\mathfrak{UL}} \leq O_{\mathfrak{L}}$.

Dies sieht man folgendermaßen ein: Verfeinert man die Zerlegung $\mathfrak{U} = [x_0, \dots, x_m]$ zu einer Zerlegung $\mathfrak{U}' = [x_0, \dots, x_{k-1}, \zeta, x_k, \dots, x_m]$, so gilt

$$\begin{aligned}
U_{\mathfrak{U}} &= \sum_{\substack{i=1\\i\neq k}}^{m} \inf\Big\{ f(\xi) \Big| \xi \in [x_{i-1}, x_i] \Big\} (x_i - x_{i-1}) \\
&\quad + \inf\Big\{ f(\xi) \Big| \xi \in [x_{k-1}, x_k] \Big\} (x_k - x_{k-1}) \\
&\leq \sum_{\substack{i=1\\i\neq k}}^{m} \inf\Big\{ f(\xi) \Big| \xi \in [x_{i-1}, x_i] \Big\} (x_i - x_{i-1}) \\
&\quad + \inf\Big\{ f(\xi) \Big| \xi \in [x_{k-1}, \zeta] \Big\} (\zeta - x_{k-1}) \\
&\quad + \inf\Big\{ f(\xi) \Big| \xi \in [\zeta, x_k] \Big\} (x_k - \zeta) \\
&= U_{\mathfrak{U}'}
\end{aligned}$$

Hierbei wurde benutzt, daß

$$\begin{aligned}
\inf\Big\{ f(\xi) \Big| \xi \in [x_{k-1}, x_k] \Big\} &\leq \inf\Big\{ f(\xi) \Big| \xi \in [\zeta, x_k] \Big\} \\
\text{beziehungsweise} &\leq \inf\Big\{ f(\xi) \Big| \xi \in [x_{k-1}, \zeta] \Big\}
\end{aligned}$$

ist.

Da $\mathfrak{UL}$ eine Verfeinerung der Zerlegung $\mathfrak{U}$ ist (d.h. die Teilpunkte von $\mathfrak{U}$ enthält), bekommt man mit dem beschriebenen Verfahren sukzessive durch jeweiliges Hinzufügen eines Teilungspunktes schließlich die Ungleichung

$$U_{\mathfrak{U}} \leq U_{\mathfrak{UL}} \tag{1.3}$$

und entsprechend

$$O_{\mathfrak{UL}} \leq O_{\mathfrak{L}}. \tag{1.4}$$

Die Zerlegung $\mathfrak{UL}$ ist unsere gesuchte Zerlegung $\mathfrak{Z}_1$, denn es gilt wegen (1.2)

$$O_{\mathfrak{L}} - U_{\mathfrak{U}} < \frac{\varepsilon}{6} + \frac{\varepsilon}{6} = \frac{\varepsilon}{3}$$

und wegen (1.3) und (1.4)

$$O_{\mathfrak{UL}} - U_{\mathfrak{UL}} < \frac{\varepsilon}{3}.$$

Die Zerlegung $\mathfrak{Z}_1$ mit der Eigenschaft (1.1) ist damit konstruiert. Es sei p die Anzahl ihrer Teilpunkte und

$$\sigma = \sup\Big\{ f(\xi) \Big| \xi \in [a, b] \Big\} - \inf\Big\{ f(\xi) \Big| \xi \in [a, b] \Big\}.$$

Da f nicht die konstante Funktion ist, gilt $\sigma \neq 0$. Wir setzen nun $\delta = \frac{\varepsilon}{3p\sigma}$ und behaupten, daß dieses δ die Eigenschaft hat, daß für alle Zerlegungen $\mathfrak{Z}$ von $[a, b]$ mit Feinheit kleiner δ die Beziehung

$$O_{\mathfrak{Z}} - U_{\mathfrak{Z}} < \varepsilon \tag{1.5}$$

gilt. Es sei also $\mathfrak{Z}$ eine Zerlegung mit $|\mathfrak{Z}| < \delta$. Wir bilden eine dritte Zerlegung $\mathfrak{Z}'$ durch Übereinanderlegen von $\mathfrak{Z}$ und $\mathfrak{Z}_1$, d.h. $\mathfrak{Z}'$ entsteht aus $\mathfrak{Z}$ und $\mathfrak{Z}_1$ durch Vereinigung der Mengen der Teilungspunkte. Analog zu unserer Betrachtung über $U_{\mathfrak{U}}$ und $U_{\mathfrak{UL}}$ erhält man auch hier wieder

$$U_{\mathfrak{Z}_1} \leq U_{\mathfrak{Z}'}, \quad O_{\mathfrak{Z}_1} \geq O_{\mathfrak{Z}'},$$

daher also wegen (1.1)

$$O_{\mathfrak{Z}'} - U_{\mathfrak{Z}'} < \frac{\varepsilon}{3}. \tag{1.6}$$

Wegen der Beziehung

$$O_{\mathfrak{Z}} - U_{\mathfrak{Z}} = (O_{\mathfrak{Z}} - O_{\mathfrak{Z}'}) + (O_{\mathfrak{Z}'} - U_{\mathfrak{Z}'}) + (U_{\mathfrak{Z}'} - U_{\mathfrak{Z}})$$

und wegen (1.6) ist es zum Nachweis von (1.5) nur erforderlich, zu zeigen, daß die (nicht negativen) Differenzen

$$O_{\mathfrak{Z}} - O_{\mathfrak{Z}'} \quad \text{und} \quad U_{\mathfrak{Z}'} - U_{\mathfrak{Z}}$$

kleiner oder gleich $\frac{\varepsilon}{3}$ sind. O.B.d.A. zeigen wir nun, daß

$$O_{\mathfrak{Z}} - O_{\mathfrak{Z}'} < \frac{\varepsilon}{3}$$

ausfallen muß. Hierzu bedenken wir Folgendes: Die Teilungsintervalle der Zerlegung $\mathfrak{Z}$ seien $[x_i, x_{i-1}]$, $i = 1, \ldots, m$. Die Teilungsintervalle der Verfeinerung $\mathfrak{Z}\mathfrak{Z}_1$ müssen in diesen Intervallen $[x_i, x_{i-1}]$ liegen; wir bezeichnen die in $[x_i, x_{i-1}]$ liegenden Intervalle mit $[z_{i,j}, z_{i,j-1}]$, $\quad j = 1, \ldots, m_i$. Hierbei ist $x_{i-1} = z_{i,0} < z_{i,1} < \cdots < z_{i,m_j} = x_i$. Mit diesen Bezeichnungen ist

$$O_{\mathfrak{Z}} = \sum_{i=1}^{m} \sup\Big\{ f(\xi) \Big| \xi \in [x_i, x_{i-1}] \Big\} (x_i - x_{i-1}),$$

$$O_{\mathfrak{Z}'} = \sum_{i=1}^{m} \sum_{j=1}^{m_i} \sup\Big\{ f(\zeta) \Big| \zeta \in [z_{i,j}, z_{i,j-1}] \Big\} (z_{i,j} - z_{i,j-1}).$$

Es gilt

$$\Big| \sup\Big\{ f(\xi) \Big| \xi \in [x_i, x_{i-1}] \Big\} - \sup\Big\{ f(\zeta) \Big| \zeta \in [z_{i,j}, z_{i,j-1}] \Big\} \Big| < \sigma$$

und damit

$$\begin{aligned}
&\left|\sup\left\{f(\xi)\middle|\xi\in[x_i,x_{i-1}]\right\}(x_i-x_{i-1})-\sum_{j=1}^{m_i}\sup\left\{f(\zeta)\middle|\zeta\in[z_{i,j},z_{i,j-1}]\right\}(z_{i,j},z_{i,j-1})\right|\\
\leq\ &\sum_{j=1}^{m_i}\left|\sup\left\{f(\xi)\middle|\xi\in[x_i,x_{i-1}]\right\}-\sup\left\{f(\zeta)\middle|\zeta\in[z_{i,j},z_{i,j-1}]\right\}\right|(z_{i,j},z_{i,j-1})\\
\leq\ &\sigma\sum_{j=1}^{m_i}(z_{i,j}-z_{i,j-1})=\sigma(x_i-x_{i-1})\\
<\ &\frac{\sigma\varepsilon}{3p\sigma}=\frac{\varepsilon}{3p},
\end{aligned}$$

da die Feinheit von $\mathfrak{Z}$ kleiner als $\frac{\varepsilon}{3p\sigma}$ sein sollte. Da $\mathfrak{Z}_1$ p Teilungspunkte enthält, muß bei der Abschätzung der Differenz $O_{\mathfrak{Z}}-O_{\mathfrak{Z}'}$ die eben durchgeführte Abschätzung höchstens p-mal durchgeführt werden, d.h. bis auf p Ausnahmefälle ist $m_i=1$ und damit

$$\begin{aligned}
&\sup\left\{f(\xi)\middle|\xi\in[x_i,x_{i-1}]\right\}(x_i-x_{i-1})\\
=\ &\sum_{j=1}^{m_i}\sup\left\{f(\zeta)\middle|\zeta\in[z_{i,j},z_{i,j-1}]\right\}(z_{i,j},z_{i,j-1})
\end{aligned}$$

bis auf höchstens p Ausnahmefälle für die Indizes i.
Damit ergibt sich

$$O_{\mathfrak{Z}}-O_{\mathfrak{Z}'}<p\frac{\varepsilon}{3p}=\frac{\varepsilon}{3}$$

und die gewünschte Abschätzung ist bewiesen. Der Beweis ist damit beendet. □

Satz 1.2. (Riemannsches Integrabilitätskriterium)
Es sei $f\colon[a,b]\to\mathbb{R}$ eine beschränkte Funktion. f ist genau dann Riemann-integrierbar über $[a,b]$, wenn es zu jedem $\varepsilon>0$ eine Zerlegung $\mathfrak{Z}$ von $[a,b]$ gibt, so daß

$$O_{\mathfrak{Z}}-U_{\mathfrak{Z}}<\varepsilon.$$

(Hierbei bedeutet $O_{\mathfrak{Z}}$ bzw. $U_{\mathfrak{Z}}$ die Riemannsche Ober- bzw. Untersumme von f über $[a,b]$ bzgl. $\mathfrak{Z}$.)
Die Bedeutung des Satzes liegt darin, daß man zum Nachweis der Integrabitität von f nur eine *spezielle* Zerlegung $\mathfrak{Z}$ mit $O_{\mathfrak{Z}}-U_{\mathfrak{Z}}<\varepsilon$ konstruieren muß.
Beweis:

⇒ Sei f integrierbar. Dann ist wegen Satz 1.1. $\overline{J}=\underline{J}$, und die in der Richtung ⇐ des Beweises von Satz 1.1. konstruierte Zerlegung $\mathfrak{Z}_1$ leistet das Gewünschte, d.h. es gilt (Formel (1.1))

$$O_{\mathfrak{Z}_1}-U_{\mathfrak{Z}_1}<\frac{\varepsilon}{3}<\varepsilon.$$

$\Leftarrow$ Wir haben die Integrierbarkeit von f zu zeigen. Nach Satz 1.1. genügt es, zu zeigen, daß $\underline{J} = \overline{J}$ ist.
Es sei $\varepsilon > 0$ vorgegeben. Nach Voraussetzung gibt es eine Zerlegung $\mathfrak{Z}$ von $[a,b]$ mit $O_{\mathfrak{Z}} - U_{\mathfrak{Z}} < \varepsilon$. Da $O_{\mathfrak{Z}} \geq \overline{J}$ und $U_{\mathfrak{Z}} \leq \underline{J}$, folgt $\overline{J} - \underline{J} < \varepsilon$. Da $\overline{J}$ und $\underline{J}$ von ε unabhängig sind, kann man nun ε gegen Null gehen lassen, und es folgt $\overline{J} - \underline{J} \leq 0$ und da nach Definition $\underline{J} \leq \overline{J}$ ist, erhalten wir die Gleicheit $\overline{J} = \underline{J}$. $\square$

Mit Hilfe von Satz 1.2. erhalten wir nun

Satz 1.3.
Es sei $f\colon [a,b] \to \mathbb{R}$ stetig. Dann ist f Riemann-integrierbar über $[a,b]$.
Beweis:
Es sei $\varepsilon > 0$ vorgegeben. Da stetige Funktionen auf abgeschlossenen Intervallen gleichmäßig stetig sind, gibt es ein $\delta > 0$ mit

$$|f(x) - f(y)| < \frac{\varepsilon}{b-a}$$

für alle $x, y \in [a,b]$ mit $|x-y| < \delta$. Es sei nun $\mathfrak{Z} = \mathfrak{Z}(x_0, \ldots, x_m)$ eine beliebige Zerlegung des Intervalls $[a,b]$ mit der Feinheit $|\mathfrak{Z}| < \delta$. Es gilt dann

$$\begin{aligned} O_{\mathfrak{Z}} - U_{\mathfrak{Z}} &= \sum_{i=1}^{m} \Bigg(\sup\Big\{ f(\xi) \Big| \xi \in [x_i, x_{i-1}] \Big\} \\ &\qquad - \inf\Big\{ f(\zeta) \Big| \zeta \in [x_i, x_{i-1}] \Big\} \Bigg) (x_i - x_{i-1}) \\ &< \frac{\varepsilon}{b-a} \sum_{i=1}^{m} (x_i - x_{i-1}) \\ &= \frac{\varepsilon}{b-a}(b-a) = \varepsilon, \end{aligned}$$

d.h. $O_{\mathfrak{Z}} - U_{\mathfrak{Z}} < \varepsilon$, und nach Satz 1.2. ist f daher Riemann integrierbar. $\square$

Man kann allgemeiner zeigen, daß stückweise stetige, beschränkte Funktionen Riemann-integrierbar sind.
Eine Funktion $f\colon [a,b] \to \mathbb{R}$ nennen wir *stückweise stetig*, wenn eine Zerlegung $\mathfrak{Z}$ des Intervalls existiert, so daß f auf den endlich vielen offenen Teilintervallen der Zerlegung stetig ist. Die Frage nach der Berechnung konkreter Riemannscher Integrale stellen wir zunächst zurück und beweisen im nächsten Kapitel einige Eigenschaften des Riemannschen Integrals.

1.2 Elementare Integrationsregeln

Satz 2.1.
Es sei $f\colon [a,b] \to \mathbb{R}$ über $\overline{I} = [a,b]$ Riemann-integrierbar. Dann ist f auch über jedes abgeschlos-

sene Teilintervall $\overline{I}_1 \subset \overline{I}$ *Riemann-integrierbar.*

Beweis:

Es sei $\varepsilon > 0$ vorgegeben. Wir zeigen, daß es eine Zerlegung $\mathfrak{Z}$ von $\overline{I}_1$ gibt, so daß $O_{\mathfrak{Z}} - U_{\mathfrak{Z}} < \varepsilon$ ausfällt. ($O_{\mathfrak{Z}}$ bzw. $U_{\mathfrak{Z}}$ ist wieder die Riemannsche Obersumme bzw. Untersumme von f bzgl. $\mathfrak{Z}$ und $\overline{I}_1$). Aus dem Riemannschen Integrabilitätskriterium folgt dann die behauptete Integrierbarkeit.

Da nach Voraussetzung f über $\overline{I}$ Riemann-integrierbar ist, gibt es eine Zerlegung $\tilde{\mathfrak{Z}}$ von $\overline{I}$ mit $O_{\tilde{\mathfrak{Z}}} - U_{\tilde{\mathfrak{Z}}} < \varepsilon$. Die in $\overline{I}$ liegenden Punkte von $\tilde{\mathfrak{Z}}$ bilden zusammen mit den Randpunkten von $\overline{I}_1$ eine Zerlegung $\mathfrak{Z}$ von $\overline{I}_1$, für die erst recht die Ungleichung

$$O_{\mathfrak{Z}} - U_{\mathfrak{Z}} < \varepsilon$$

gilt. □

Satz 2.2.

Es sei $f\colon [a,b] \to \mathbb{R}$ *über* $\overline{I} = [a,b]$ *Riemann-integrierbar und* $a < c < b$. *Dann gilt*

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

Beweis:

Nach Definition des Riemannschen Integrals gibt es zu jedem $\varepsilon > 0$ eine Zahl δ, so daß für alle Zerlegungen $\mathfrak{Z}$ von $[a,b]$, $\mathfrak{Z}_1$ von $[a,c]$ und $\mathfrak{Z}_2$ von $[c,b]$ mit Feinheit $|\mathfrak{Z}| < \delta$ die Ungleichungen

$$\left| \int_a^b f\,dx - \sum(f, \xi_i, \mathfrak{Z}) \right| < \varepsilon,$$

$$\left| \int_a^c f\,dx - \sum(f, \xi_i, \mathfrak{Z}_1) \right| < \varepsilon,$$

$$\left| \int_c^b f\,dx - \sum(f, \xi_i, \mathfrak{Z}_2) \right| < \varepsilon$$

stattfinden. Insbesondere *gibt es auch* zwei Zerlegungen $\mathfrak{Z}_1$ und $\mathfrak{Z}_2$ mit dieser Eigenschaft. Aus $\mathfrak{Z}_1$ und $\mathfrak{Z}_2$ bilden wir durch Zusammenlegen eine Zerlegung $\mathfrak{Z}$ von $[a,b]$ und wählen die zugehörigen Zwischenpunkte so, daß $\sum(f, \xi_i, \mathfrak{Z}) = \sum(f, \xi_i, \mathfrak{Z}_1) + \sum(f, \xi_i, \mathfrak{Z}_2)$. Wir erhalten dann unter Benutzung der Dreiecksungleichung

$$\left| \int_a^b f\,dx - \sum(f, \xi_i, \mathfrak{Z}) \right| < 3\varepsilon.$$

Durch Grenzübergang für $\varepsilon \to 0$ ergibt sich die Behauptung. □

Satz 2.3. (Linearität des Riemannschen Integrals)
Es seien $f_1, f_2 \colon [a,b] \to \mathbb{R}$ über $\overline{I} = [a,b]$ Riemann-integrierbar. Dann ist auch $f_1 + f_2$ und αf_1 für $\alpha \in \mathbb{R}$ Riemann-integrierbar, und es gilt

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

und

$$\int_a^b \alpha f_1\, dx = \alpha \int_a^b f_1\, dx.$$

Beweis:
Sei $\varepsilon > 0$ vorgegeben. Nach Definition der Riemannschen Integrals gibt es ein δ, so daß für alle Zerlegungen $\mathfrak{Z}$ von $\overline{I}$ mit $|\mathfrak{Z}| < \delta$ die Ungleichungen

$$\left| \int_a^b f_1\, dx - \sum(f_1, \xi_i, \mathfrak{Z}) \right| < \frac{\varepsilon}{2},$$

$$\left| \int_a^b f_2\, dx - \sum(f_2, \xi_i, \mathfrak{Z}) \right| < \frac{\varepsilon}{2}$$

gelten. Wir setzen $A = \int_a^b f_1\, dx + \int_a^b f_2\, dx$ und erhalten unter Benutzung der Linearität der Riemannschen Summe

$$\begin{aligned} \left| A - \sum(f_1, \xi_i, \mathfrak{Z}) \right| &= \left| \int_a^b f_1\, dx - \sum(f_1, \xi_i, \mathfrak{Z}) + \int_a^b f_2\, dx - \sum(f_2, \xi_i, \mathfrak{Z}) \right| \\ &< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon \end{aligned}$$

nach der Dreiecksungleichung.
Es gibt also eine Zahl A, so daß für alle Zerlegungen $\mathfrak{Z}$ mit Feinheit $|\mathfrak{Z}| < \delta = \delta(\varepsilon)$ die Ungleichung

$$\left| A - \sum(f_1 + f_2, \xi_i, \mathfrak{Z}) \right| < \varepsilon$$

gilt. Daher ist $f_1 + f_2$ Riemann-integrierbar, und es gilt

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

Um die Gleichheit

$$\int_a^b \alpha f_1 \, dx = \alpha \int_a^b f_1 \, dx$$

zu beweisen, nutzt man ebenfalls die Linearität der Riemannschen Summe aus. □

Satz 2.4. (Monotonie des Riemannschen Integrals)
Es seien $f_1, f_2 \colon [a, b] \to \mathbb{R}$ über $\overline{I} = [a, b]$ Riemann-integrierbar, und es sei $f_1(x) \leq f_2(x)$ für alle $x \in [a, b]$. Dann gilt

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

Der Beweis wird auch hier wieder auf die Riemannschen Summen zurückgeführt, worauf wir verzichten wollen.

Aus Satz 2.4. folgt

Satz 2.4'.
Die Funktion $f \colon [a, b] \to \mathbb{R}$ sei beschränkt und Riemann-integrierbar über $[a, b]$. Dann gilt

$$\left| \int_a^b f(x) \, dx \right| \leq (b - a) \sup\Big\{ f(\xi) \Big| \xi \in [a, b] \Big\}.$$

Beweis:
Sei $S := \sup\Big\{ f(\xi) \Big| \xi \in [a, b] \Big\}$. Dann ist $-S \leq f(x) \leq S$ für alle $x \in [a, b]$, und wegen Satz 2.4 gilt dann

$$-S \int_a^b 1 \, dx \leq \int_a^b f(x) \, dx \leq S \int_a^b 1 \, dx.$$

Die zu $\int_a^b 1 \, dx$ gehörigen Riemannschen Summen haben alle den Wert $b - a$, daher gilt auch $\int_a^b 1 \, dx = b - a$. Daraus folgt

$$-S(b - a) \leq \int_a^b f(x) \, dx \leq S(b - a)$$

und somit die Behauptung. □

1.3 Der Hauptsatz der Differential- und Integralrechnung

Es sei $\overline{I}$ das abgeschlossene Intervall $[a,b]$ und f eine über $\overline{I}$ integrierbare Funktion. Für $x_1, x_2 \in \overline{I}$ ist dann, falls $x_1 < x_2$ ist,

$$\int_{x_1}^{x_2} f(x)\,dx$$

wohldefiniert. Es ist nützlich, diese Definition auch auf die Fälle $x_1 = x_2$ oder $x_1 > x_2$ folgendermaßen auszudehnen:

$$\int_{x_1}^{x_1} f(x)\,dx = 0, \qquad \int_{x_1}^{x_2} f(x)\,dx = -\int_{x_2}^{x_1} f(x)\,dx.$$

Man überzeugt sich sofort von der Gültigkeit der in 1.2 aufgestellten Regeln auch für den erweiterten Integralbegriff; insbesondere ist

$$\int_{x_1}^{x_2} f(x)\,dx + \int_{x_2}^{x_3} f(x)\,dx = \int_{x_1}^{x_3} f(x)\,dx,$$

ganz gleich, wie die Punkte x_1, x_2, x_3 in $\overline{I}$ gewählt werden.
Die Funktion f ist auch über Teilintervalle der Form $[a,x]$, $x \in \overline{I}$, integrierbar; wir definieren also mit Hilfe von f eine neue Funktion F durch

$$F(x) = \int_a^x f(\xi)\,d\xi\ , \quad x \in \overline{I}.$$

Der Zusammenhang zwischen F und f soll nun untersucht werden.Es gilt der

Satz 3.1. (Hauptsatz der Differential und Integralrechnung)
Es sei $f\colon [a,b] \to \mathbb{R}$ stetig. Dann ist die durch

$$F(x) = \int_a^x f(\xi)\,d\xi$$

erklärte Funktion $F\colon [a,b] \to \mathbb{R}$ in $]a,b[$ differenzierbar, und es gilt

$$\frac{d}{dx}F(x_0) = f(x_0) \text{ für } x_0 \in]a,b[.$$

Beweis:
Es gilt für $x, x_0 \in]a,b[,\ x \neq x_0$:

$$F(x) - F(x_0) = \int_a^x f(\xi)\,d\xi - \int_a^{x_0} f(\xi)\,d\xi = \int_{x_0}^x f(\xi)\,d\xi$$

und damit

$$\left|\frac{F(x)-F(x_0)}{x-x_0}-f(x_0)\right| = \left|\frac{1}{x-x_0}\int\limits_{x_0}^{x}\Big(f(\xi)-f(x_0)\Big)\,d\xi\right| . \tag{1.7}$$

Daraus folgt mit Satz 2.4'

$$\left|\frac{F(x)-F(x_0)}{x-x_0}-f(x_0)\right| \le \sup\Big\{|f(\xi)-f(x_0)| \;\Big|\; \xi\in[x_0,x]\Big\} \tag{1.8}$$

(Falls $x_0 > x$, setzen wir $[x_0, x] = [x, x_0]$).
Wegen der Stetigkeit von f geht die rechte Seite für $x \to x_0$ von (1.8) gegen Null, das heißt

$$\frac{F(x)-F(x_0)}{x-x_0} \to f(x_0) \quad \text{für} \quad x \to x_0,\ x \neq x_0$$

Der Satz ist damit bewiesen. □

Aus (1.7) folgt wegen der Beschränktheit von f für beliebige Riemann-integrierbare Funktionen: *Das Riemannsche Integral hängt stetig von den Grenzen ab.*

Die Beziehungen zwischen Integration und Differentiation gestatten es, Riemannsche Integrale explizit zu berechnen. Hierzu dienen die beiden folgenden Definitionen sowie Satz 3.2.

Definition:
Eine Stammfunktion *zur Funktion f ist eine differenzierbare Funktion F, für die $F' = f$ ist.*

Definition:
Eine Funktion F heißt unbestimmtes Integral *der integrierbaren Funktion f, wenn für je zwei Punkte $x_1, x_2 \in \overline{I}$*

$$\int\limits_{x_1}^{x_2} f(\xi)\,d\xi = F(x_2) - F(x_1)$$

ist.

Satz 3.2.

(a) Sind F_1 und F_2 Stammfunktonen von $f\colon [a,b] \to \mathbb{R}$, so ist $F_1 - F_2$ konstant. Mit F ist auch $F + C$ für $C \in \mathbb{R}$ Stammfunktion.

(b) Ist F Stammfunktion der integrierbaren Funktion f, so gilt

$$\int\limits_{x_1}^{x_2} f(\xi)\,d\xi = F(x_2) - F(x_1).$$

Beweis:

(a) Aus $F_1' = F_2' = f$ folgt $\frac{d}{dx}(F_1 - F_2) = 0$ und daher $F_1 - F_2 = konstant$. Aus $F' = f$ folgt $(F + C)' = f$.

(b) Da $\int_a^x f(\xi)\, d\xi$ als Funktion von x Stammfunktion zu f ist, folgt nach (a) $F(x) - \int_a^x f(\xi)\, d\xi = C$ (konstant) für $x \in [a, b]$. Daraus folgt

$$F(x_2) - \int_a^{x_2} f(\xi)\, d\xi - F(x_1) + \int_a^{x_1} f(\xi)\, d\xi = 0,$$

also

$$F(x_2) - F(x_2) = \int_a^{x_2} f(\xi)\, d\xi - \int_a^{x_1} f(\xi)\, d\xi = \int_{x_1}^{x_2} f(\xi)\, d\xi.$$

□

Da für jede stetige Funktion $f\colon [a, b] \to \mathbb{R}$ das Integral $\int_a^x f(x)\, dx$ definiert und differenzierbar ist mit $F' = f$ folgt:

Satz 3.3.
Jede stetige Funktion $f\colon [a, b] \to \mathbb{R}$ besitzt eine Stammfunktion. □

Stammfunktionen elementarer Funktionen:
In vielen Fällen lassen sich die Stammfunktionen „elementarer“ Funktionen explizit durch elementare Funktionen ausdrücken. Wir geben hier einige Beispiele. Die Richtigkeit kann durch Differentiation überprüft werden.

Funktion f	Stammfunktion F	Definitionsbereich
$x^k, \quad k \in \mathbb{R}^+$	$\frac{1}{k+1}x^{k+1}$	$x \in \mathbb{R}$
$x^k, \quad k = -2, -3, \dots$	$\frac{1}{k+1}x^{k+1}$	$x \neq 0$
x^{-1}	$\ln \lvert x \rvert$	$x \neq 0$
$\frac{1}{1+x^2}$	$\arctan x$	$x \in \mathbb{R}$
$\frac{1}{\sqrt{1-x^2}}$	$\arcsin x$	$\lvert x \rvert < 1$
e^x	e^x	$x \in \mathbb{R}$
$\sin x$	$-\cos x$	$x \in \mathbb{R}$
$\cos x$	$\sin x$	$x \in \mathbb{R}$
$\frac{1}{\sin^2 x}$	$-\cot x$	$x \neq k\pi, \ k \in \mathbb{Z}$
$\frac{1}{\cos^2 x}$	$\tan x$	$x \neq (k + \frac{1}{2})\pi, \ k \in \mathbb{Z}$

Ebenso kann man Stammfunktionen für Potenzreihen angeben:

Satz 3.4.
Eine Stammfunktion der Potenzreihe $\sum\limits_{k=0}^{\infty} a_k(x - x_0)^k$ *im Konvergenzintervall* I *ist die Reihe* $\sum\limits_{k=0}^{\infty} \frac{a_k}{k+1}(x - x_0)^{k+1}$, *die ebenfalls das Konvergenzintervall* I *hat.*

Die Stammfunktionen von einigen weiteren Funktionen werden in den folgenden Kapiteln angegeben.

1.4 Partielle Integration

Die im Folgenden bewiesene Regel von der partiellen Integration ist in der Analysis von großer Bedeutung und wird sehr oft angewandt:

Satz 4.1.
Die Funktion $f, g \colon]a, b[\to \mathbb{R}$ *seien in* $]a, b[$ *differenzierbar und es sei* F *eine Stammfunktion von* $g'f$. *Dann ist* $fg - F$ *eine Stammfunktion von* $f'g$. *Sind* f, g, f' *und* g' *zusätzlich stetige*

Funktionen von $[a,b]$ nach $\mathbb{R}$, so gilt

$$\int_a^b f'(x)g(x)\,dx = f(b)g(b) - f(a)g(a) - \int_a^b g'(x)f(x)\,dx.$$

Beweis:
Es ist

$$\frac{d}{dx}(fg - F) = f'g + g'f - g'f = f'g,$$

das heißt $fg - F$ ist in der Tat Stammfunktion zu $f'g$.
Zum Beweis der zweiten Behauptung sei $x_1 < x_2,\ x_1, x_2 \in]a,b[$. Die durch

$$F(x) = \int_{x_1}^{x} g'(\xi)f(\xi)\,d\xi$$

definierte Funktion $F\colon]a,b[\to \mathbb{R}$ ist Stammfunktion zu $g'f$, daher gilt

$$\frac{d}{dx}\left((fg)(x) - \int_{x_1}^{x} g'(\xi)f(\xi)d\xi\right) = (f'g)(x)$$

und

$$(fg)(x_2) - \int_{x_1}^{x_2} g'(\xi)f(\xi)\,d\xi - (fg)(x_1) + \int_{x_1}^{x_1} g'(\xi)f(\xi)\,d\xi = \int_{x_1}^{x_2} (f'g)(\xi)\,d\xi.$$

Daraus folgt

$$\int_{x_1}^{x_2} f'g(\xi)\,d\xi = (fg)(x_2) - (fg)(x_1) - \int_{x_1}^{x_2} g'(\xi)f(\xi)\,d\xi.$$

Wegen der vorausgesetzten Stetigkeit von g' und f' sowie f und g darf man in der letzten Gleichung den Grenzübergang $x_1 \to a, \quad x_2 \to b$ durchführen. Daraus ergibt sich

$$\int_a^b f'g(\xi)\,d\xi = (fg)(b) - (fg)(a) - \int_a^b g'(\xi)f(\xi)\,d\xi.$$

$\square$

Der Leser möge meinen, daß man doch auch gleich hätte $x_1 = a$ und $x_2 = b$ setzen können. Dies ist jedoch nicht möglich, da $f'(a)$ in unserem Beweis nicht den Differentialquotient von f an der Stelle a, sondern die Fortsetzung $f'\colon]a,b[\to \mathbb{R}$ auf ganz $[a,b]$ bedeutet. Dies wurde in der Voraussetzung formuliert: „*f' sei eine stetige Funktion von $[a,b]$ nach $\mathbb{R}$*“.

Anwendungen:

1. Integration von $\ln x$ (für $x > 0$). Wir setzen $f'(x) = 1$ sowie $g(x) = \ln x$ und erhalten

$$\begin{aligned}\int_{x_0}^{x} \ln x\, dx &= x \ln x - x_0 \ln x_0 - \int_{x_0}^{x} x \cdot \frac{1}{x}\, dx \quad (x_0, x > 0)\\ &= x \ln x - x_0 \ln x_0 - x + x_0\\ &= x \ln x - x + const.\end{aligned}$$

2. Zur Integration von $x \sin x$ setzt man

$$\begin{aligned}f'(x) &= \sin x,\\ f(x) &= -\cos x,\\ g(x) &= x,\\ g'(x) &= 1\end{aligned}$$

und erhält

$$\begin{aligned}\int_{x_0}^{x} x \sin x\, dx &= -x \cos x + x_0 \cos x_0 + \int_{x_0}^{x} \cos x\, dx\\ &= -x \cos x + x_0 \cos x_0 + x \sin x + x_0 \sin x_0.\end{aligned}$$

3. Berechne $\int_{x_0}^{x} \sin^2 x\, dx$. Wir setzen

$$\begin{aligned}f'(x) &= \sin x,\\ f(x) &= -\cos x,\\ g(x) &= \sin x,\\ g'(x) &= \cos x\end{aligned}$$

und erhalten

$$\begin{aligned}\int_{x_0}^{x} \sin^2 x\, dx &= -\sin x \cos x + \sin x_0 \cos x_0 + \int_{x_0}^{x} \cos^2 x\, dx\\ &= -\sin x \cos x + \sin x_0 \cos x_0 + \int_{x_0}^{x} dx - \int_{x_0}^{x} \sin^2 dx,\\ &\text{da} \quad \sin^2 x + \cos^2 x = 1.\end{aligned}$$

Also ergibt sich

$$\int_{x_0}^{x} \sin^2 x\, dx = \frac{1}{2}(-\sin x \cos x + \sin x_0 \cos x_0 + x - x_0).$$

1.5 Substitutionsregel

Satz 5.1.
Es seien $f\colon [a,b] \to \mathbb{R}$ *stetig und* $\varphi\colon [\alpha,\beta] \to \mathbb{R}$ *stetig differenzierbar, und es gelte* $\varphi([\alpha,\beta]) \subset [a,b]$*. Dann gilt*

$$\int\limits_{\varphi(\alpha)}^{\varphi(\beta)} f(x)\,dx = \int\limits_{\alpha}^{\beta} f(\varphi(u))\varphi'(u)\,du.$$

Anmerkung:
φ heißt auf $[\alpha,\beta]$ stetig differenzierbar, wenn $\varphi\colon [\alpha,\beta] \to \mathbb{R}$ stetig ist, φ in $]\alpha,\beta[$ differenzierbar ist und die Ableitung $\varphi'\colon]\alpha,\beta[\to \mathbb{R}$ eine auf $[\alpha,\beta]$ stetige Fortsetzung besitzt, die ebenfalls mit φ' bezeichnet wird.

Beweis:
Es sei F eine Stammfunktion von f. Dann ist in $]\alpha,\beta[$

$$(F\circ\varphi)' = (F'\circ\varphi)\varphi' = (f\circ\varphi)\varphi'. \tag{1.9}$$

Daraus folgt

$$\begin{aligned}\int\limits_{\alpha}^{\beta} \Big[(f\circ\varphi)\varphi'\Big] &= \int\limits_{\alpha}^{\beta} (F\circ\varphi)'(x)\,dx \qquad (1.10)\\ &= (F\circ\varphi)(\beta) - (F\circ\varphi)(\alpha)\\ &= F(\varphi(\beta)) - F(\varphi(\alpha))\end{aligned}$$

und

$$\int\limits_{\alpha}^{\beta} f(\varphi(u))\varphi'(u)\,du = F(\varphi(\beta)) - F(\varphi(\alpha)).$$

Anmerkung:
Die Folgerung nach (1.9) ist nicht trivial, man muß $x_1, x_2 \in]\alpha,\beta[,\ x_1 < x_2$, wählen, und die Formel

$$\int\limits_{x_1}^{x_2} \Big[(f\circ\varphi)\varphi'\Big](u)\,du = F(\varphi(x_2)) - F(\varphi(x_1))$$

anwenden. Der Grenzübergang $x_1 \to \alpha,\ x_2 \to \beta$ ergibt dann Formel (1.10). □

In vielen Fällen läßt sich eine vorgelegte Funktion in der Form $f(\varphi(u))\varphi'(u)$ schreiben; um sie zu integrieren, braucht man dann aufgrund des obigen Satzes nur noch eine Stammfunktion von f zu finden.

Beispiele:

1. $\int\limits_\alpha^\beta (a+bu)^n\,du \quad (\text{mit } b \neq 0)$: Mit

$$x = \varphi(u) = a + bu, \quad f(x) = x^n, \quad \varphi'(u) = b$$

ist

$$\begin{aligned} \int\limits_\alpha^\beta (a+bu)^n\,du &= \frac{1}{b}\int\limits_\alpha^\beta (a+bu)^n b\,du \\ &= \frac{1}{b}\int\limits_{\varphi(\alpha)}^{\varphi(\beta)} x^n\,dx \\ &= \frac{1}{b(n+1)}\Big[(a+b\beta)^{n+1} - (a+b\alpha)^{n+1}\Big]. \end{aligned}$$

2. Es sei $g(u) > 0$ und stetig differenzierbar für $u \in [\alpha, \beta]$. Wir berechnen $\int\limits_\alpha^\beta \frac{g'(u)}{g(u)}\,du$: Mit

$$\begin{aligned} x = \varphi(u) &= g(u), \\ \varphi'(u) &= g'(u) \end{aligned}$$

und

$$f(x) = \frac{1}{x}$$

ergibt sich

$$\int\limits_\alpha^\beta \frac{g'(u)}{g(u)}\,du = \int\limits_\alpha^\beta \frac{dx}{x} = \ln g(\beta) - \ln g(\alpha) = \ln\frac{g(\beta)}{g(\alpha)}.$$

3. $\int\limits_\alpha^\beta \frac{1}{u}\ln u\,du \quad (0 < \alpha < \beta)$: Mit

$$\begin{aligned} x = \varphi(u) &= \ln u \\ \varphi'(u) &= \frac{1}{u} \end{aligned}$$

und

$$f(x) = x$$

erhalten wir

$$\int\limits_{\alpha}^{\beta} \frac{1}{u} \ln u \, du = \int\limits_{\ln \alpha}^{\ln \beta} x \, dx = \frac{1}{2}[\ln^2(\beta) - \ln^2(\alpha)].$$

4. Gelegentlich muß man das vorgelegte Integral erst geschickt umformen, um die Substitutionsregel anwenden zu können. Zum Beispiel ist

$$\begin{aligned} \frac{1}{\sin u} &= \frac{\sin u}{\sin^2 u} = \frac{\sin u}{1 - \cos^2 u} \\ &= \frac{\sin u}{(1 - \cos u)(1 + \cos u)} \\ &= \frac{1}{2} \cdot \frac{\sin u}{1 - \cos u} + \frac{1}{2} \cdot \frac{\sin u}{1 + \cos u} \end{aligned}$$

Ist also $0 < \alpha < \beta < \pi$, so ist

$$\int\limits_{\alpha}^{\beta} \frac{du}{\sin u} = \frac{1}{2} \int\limits_{\alpha}^{\beta} \frac{\sin u \, du}{1 - \cos u} + \frac{1}{2} \int\limits_{\alpha}^{\beta} \frac{\sin u \, du}{1 + \cos u}.$$

Im ersten Integral rechts setzen wir

$$\begin{aligned} x = \varphi(u) &= 1 - \cos u, \\ \varphi'(u) &= \sin u, \end{aligned}$$

$$f(x) = \frac{1}{x}$$

und erhalten

$$\frac{1}{2} \int\limits_{\alpha}^{\beta} \frac{\sin u \, du}{1 - \cos u} = \frac{1}{2} \int\limits_{1-\cos \alpha}^{1-\cos \beta} \frac{dx}{x} = \frac{1}{2} \ln \frac{1 - \cos \beta}{1 - \cos \alpha}.$$

Im zweiten Integral auf der rechten Seite substituiert man

$$\begin{aligned} x = \varphi(u) &= 1 + \cos u, \\ \varphi'(u) &= -\sin u, \end{aligned}$$

$$f(x) = -\frac{1}{x},$$

also:

$$\frac{1}{2} \int\limits_{\alpha}^{\beta} \frac{\sin u \, du}{1 + \cos u} = -\frac{1}{2} \int\limits_{1+\cos \alpha}^{1+\cos \beta} \frac{dx}{x} = -\frac{1}{2} \ln \frac{1 + \cos \beta}{1 + \cos \alpha}.$$

Insgesamt ergibt sich

$$\int_{\alpha}^{\beta} \frac{du}{\sin u} = -\frac{1}{2} \ln \frac{(1+\cos\alpha)(1-\cos\beta)}{(1+\cos\beta)(1-\cos\alpha)}.$$

Als Stammfunktion erhält man etwa:

$$F(u) = \frac{1}{2} \ln \frac{1-\cos u}{1+\cos u}.$$

1.6 Die Integration der rationalen Funktion

In diesem Paragraphen wollen wir zeigen, daß sich die Stammfunktionen rationaler Funktionen durch „elementare analytische Funktionen“ ausdrücken lassen (d.h. durch rationale Funktionen, e-Funktionen, arctan sowie ln).
Zunächst bemerken wir, daß jede rationale Funktion f sich in der Form

$$f(x) = g_0(x) + \frac{h(x)}{g(x)}$$

darstellen läßt, wobei g_0, h und g Polynome sind und der Grad von h kleiner als der von g ist. Dies folgt z.B. aus dem Euklidischen Algorithmus, siehe z.B. [3, Seite 73]. Da wir schon wissen, wie man g_0 integriert, brauchen wir also nur Ausdrücke $\frac{h(x)}{g(x)}$ zu betrachten, bei denen der Grad von h kleiner als der von g ist. Aus dem Fundamentalsatz der Algebra folgt, daß man jedes Polynom g mit reellen Koeffizienten in der folgenden Weise schreiben kann:

$$g(x) = \prod_{j=1}^{m} (x-a_j)^{r_j} \prod_{k=1}^{n} Q_k^{s_k}(x) \tag{1.11}$$

mit quadratischen Polynomen Q_k ohne reelle Nullstellen und Zahlen $a_j \in \mathbb{R}$ und $r_j, s_k \in \mathbb{N}$.
Mit Hilfe der Zerlegung (1.11) läßt sich die Existenz einer *Partialbruchzerlegung* zeigen. Es gilt die Darstellung

$$\frac{h(x)}{g(x)} = \sum_{j=1}^{m} \sum_{\nu=1}^{r_j} \frac{c_{\nu j}}{(x-a_j)^{\nu}} + \sum_{k=1}^{n} \sum_{\lambda=1}^{s_k} \frac{L_{\lambda k}(x)}{{Q_k}^{\lambda}(x)}$$

mit linearen Funktionen $L_{\lambda k}$ und $c_{\nu j} \in \mathbb{R}$.
Wir werden dies nicht beweisen, aber bei Beispielen wird uns die Richtigkeit einleuchten. Aufgrund der obigen Partialbruchzerlegung genügt es zur Angabe einer Stammfunktion von $\frac{h}{g}$, daß man die Stammfunktion für

$$\frac{c_{\nu j}}{(x-a_j)^{\nu}} \quad \text{sowie} \quad \frac{L_{\lambda k}(x)}{{Q_k}^{\lambda}(x)} \quad \text{angibt.}$$

Zunächst werden wir jedoch die Partialbruchzerlegung einfacher rationaler Funktionen angeben.

Partialbruchzerlegung von $r(x) = \frac{x^2+1}{(x-1)^2(x+1)}$. Nach der oben angegebenen Zerlegung muß gelten

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

Daraus folgt:

$$\begin{aligned} r(x) &= \frac{A(x^2-1) + B(x+1) + C(x-1)^2}{(x-1)^2(x+1)} \\ &= \frac{x^2(A+C) + x(B-2C) + (-A+B+C)}{(x-1)^2(x+1)} \end{aligned}$$

Durch Koeffizientenvergleich ergibt sich

$$\begin{aligned} A + C &= 1 \\ B - 2C &= 0 \\ -A + B + C &= 1. \end{aligned}$$

Daraus folgt

$$\begin{aligned} A + C &= 1 \\ -A + 3C &= 1, \end{aligned}$$

also

$$C = \frac{1}{2}, \qquad B = 1, \qquad A = \frac{1}{2}.$$

Damit kann man r schreiben als

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

Probe:

$$\begin{aligned} \frac{1}{2(x-1)} + \frac{1}{(x-1)^2} + \frac{1}{2(x+1)} &= \frac{x^2 - 1 + 2x + 2 + (x-1)^2}{2(x-1)^2(x+1)} \\ &= \frac{2x^2+2}{2(x-1)^2(x+1)} \\ &= \frac{x^2+1}{(x-1)^2(x+1)} \\ &= r(x). \end{aligned}$$

Zweites Beispiel:

$$r(x) = \frac{2x^4 + 2x^2 - 5x + 1}{x(x^2 + x + 1)^2}$$

Aus dem Satz von der Existenz der Partialbruchzerlegung folgt, daß man r wie folgt zerlegen kann:

$$r(x) = \frac{A}{x} + \frac{B_1 x + C_1}{x^2 + x + 1} + \frac{B_2 x + C_2}{(x^2 + x + 1)^2}.$$

Es sei angemerkt, daß $x^2 + x + 1$ keine reellen Nullstellen besitzt. Durch Koeffizientenvergleich ergibt sich

$$\begin{aligned} A + B_1 &= 2 \\ 2A + B_1 + C_1 &= 0 \\ 3A + B_1 + C_1 + B_2 &= 2 \\ 2A + C_1 + C_2 &= -5 \\ A &= 1. \end{aligned}$$

Daraus folgt

$$A = 1, \qquad B_1 = 1, \qquad C_1 = -3, \qquad B_2 = 1, \qquad C_2 = -4.$$

Aus diesen Überlegungen folgt, daß wir zur Integration rationaler Funktionen bloß die Stammfunktionen F der folgenden Funktionen f angeben müssen ($x \neq x_0$):

1. $f(x) = \frac{1}{x - x_0}$, $\quad F(x) = \ln|x - x_0|$.

2. $f(x) = \frac{1}{(x-x_0)^k}$, $\quad (k \geq 2) \quad F(x) = \frac{-1}{(k-1)(x-x_0)^{k-1}}$.

3. $f(x) = \frac{1}{x^2+ax+b}$, $4b - a^2 > 0$ (Die Bedingung $4b - a^2 > 0$ ist äquivalent dazu, daß $x^2 + ax + b$ keine reellen Nullstellen hat.)
   In diesem Falle ist
   $$F(x) = \frac{1}{\sigma} \arctan\left[\frac{1}{\sigma}\left(x + \frac{a}{2}\right)\right] \quad \text{mit} \quad \sigma = \sqrt{b - \frac{a^2}{4}}.$$
   Man kann dies durch Differentiation nachweisen:
   $$\begin{aligned} F'(x) &= \frac{1}{\sigma^2} \cdot \frac{1}{\left[\frac{1}{\sigma}\left(x + \frac{a}{2}\right)\right]^2 + 1} = \frac{1}{x^2 + ax + \frac{a^2}{4} + \sigma^2} \\ &= \frac{1}{x^2 + ax + b}. \end{aligned}$$

4. $f(x) = \frac{1}{(x^2+ax+b)^k}$, $k \geq 1$. Durch die Substitution $y = \frac{1}{\sigma}(x + \frac{a}{2}), \sigma = \sqrt{b - \frac{a^2}{4}}$ erhält man

$$\begin{aligned} f(x) &= \frac{1}{\left[\left(x+\frac{a}{2}\right)^2 + b - \frac{a^2}{4}\right]^k} \\ &= \frac{1}{\sigma^{2k}\left[\frac{1}{\sigma^2}\left(x+\frac{a}{2}\right)^2 + 1\right]^k} \\ &= \frac{1}{\sigma^{2k}\left(y^2+1\right)^k}, \end{aligned}$$

das heißt, es genügt eine Stammfunktion zu $\frac{1}{(y^2+1)^k}$ zu finden. Es gilt

$$\begin{aligned} \frac{d}{dy}\frac{y}{(1+y^2)^k} &= \frac{1}{(1+y^2)^k} - \frac{2k\,y^2}{(1+y^2)^{k+1}} \\ &= \frac{1-2k}{(1+y^2)^k} + \frac{2k}{(1+y^2)^{k+1}}. \end{aligned}$$

Daraus folgt

$$\frac{1}{(1+y^2)^{k+1}} = \frac{1}{2k}\frac{d}{dy}\frac{y}{(1+y^2)^k} + \frac{2k-1}{2k}\frac{1}{(1+y^2)^k}.$$

Bezeichnet nun $F_k(y)$ eine Stammfunktion von $\frac{1}{(1+y^2)^{k+1}}$ an der Stelle y, so folgt aus der letzten Gleichung

$$F_{k+1}(y) = \frac{1}{2k}\frac{y}{(1+y^2)^k} + \frac{2k-1}{2k}F_k \; + C.$$

Da $F_1(y) = \arctan y$ bekannt ist, hat man damit eine Rekursionsformel gefunden, aus der man F_k berechnen kann.

5. $f(x) = \frac{x}{(x^2+ax+b)^k}$, $k \geq 1$. (Der Nenner habe keine Nullstellen).
Zur Bestimmung einer Stammfunktion schreiben wir f in der Form

$$f(x) = \frac{1}{2}\frac{2x+a}{(x^2+ax+b)^k} - \frac{a}{2}\frac{1}{(x^2+ax+b)^k}.$$

Der zweite Summand rechts wurde unter 4. behandelt, eine Stammfunktion des ersten Summanden ist

$$\frac{1}{2}\cdot\frac{1}{1-k}\cdot\frac{1}{(x^2+ax+b)^{k-1}} \quad \text{für} \quad k \geq 2$$

und

$$\frac{1}{2}\ln(x^2 + ax + b) \quad \text{für} \quad k = 1.$$

Damit ist bewiesen, daß die rationalen Funktionen elementar integrierbar sind. □

1.7 Uneigentliche Riemannsche Integrale

Aus verschiedenen Gründen möchte man den in den vorangegangenen Kapiteln eingeführten Integralbegriff erweitern und auch auf *nicht kompakte Intervalle*, etwa $[0, \infty[$, als Integrationsintervall zulassen. Dies führt auf den Begriff des uneigentlichen Integrals. Sei J ein beliebiges (endliches oder unendliches, abgeschlossenes, halboffenes oder offenes) Intervall mit den Endpunkten a, b. Die Funkion $f(x)$ sei auf J integrierbar. Das Symbol $\int_a^b f(x)\,dx$ nennen wir dann ein *uneigentliches Integral.* Das ist bis jetzt eine reine Bezeichnung, ohne wesentlichen Inhalt. Den Inhalt bekommt der Begriff erst, wenn wir jetzt erklären, wann wir ein solches uneigentliches Integral *konvergent* (nicht konvergent = divergent) nennen.

Definition:
Das uneigentliche Integral $\int_a^b f(\xi)\,d\xi$ *heißt* konvergent, *wenn für ein* $\gamma \in J$ *das Integral* $\int_x^\gamma f(\xi)\,d\xi$ *als Funktion von* x *beim Grenzübergang* $x \to a$ $(x \in J)$ *und* $\int_\gamma^y f(\xi)\,d\xi$ *als Funktion von* y *beim Grenzübergang* $y \to b$ $(y \in J)$ *konvergieren. Wir nennen dann die Summe dieser beiden Grenzwerte den Wert von* $\int_a^b f(\xi)\,d\xi$ *und schreiben*

$$\int_a^b f(\xi)\,d\xi = \lim_{x\to a} \int_x^\gamma f(\xi)\,d\xi + \lim_{x\to b} \int_\gamma^y f(\xi)\,d\xi.$$

Konvergenzbegriff und Wert eines uneigentlichen Integrals sind unabhängig von der Wahl von $\gamma \in J$, wie man sich leicht überlegt.

Beispiele:

1.

$$\int_0^\infty e^{-\xi}\,d\xi$$

ist an der oberen Grenze uneigentlich. Wegen $e^{-\xi} \to 0$ für $\xi \to \infty$ folgt aber

$$\int_0^\infty e^{-\xi}\,d\xi = -e^{-\xi}\Big|_0^\infty = 1.$$

2.

$$\int_{-\infty}^{\infty} \frac{d\xi}{1+\xi^2} = \arctan \xi|_{-\infty}^{\infty} = \frac{\pi}{2} - (-\frac{\pi}{2}) = \pi.$$

3.

$$\int_{1}^{\infty} \frac{d\xi}{\xi^\alpha} = \begin{cases} \left.\frac{\xi^{-\alpha+1}}{-\alpha+1}\right|_1^\infty & \text{für } \alpha \neq 1, \quad = \frac{1}{\alpha-1} \text{ für } \alpha > 1 \\ \log \xi|_1^\infty & \text{für } \alpha = 1. \end{cases}$$

Für $\alpha < 1$ liegt Divergenz dieses Integrals vor, und zwar die Divergenz gegen $+\infty$.

4.

$$\int_{0}^{1} \frac{d\xi}{\xi^\alpha} = \begin{cases} \left.\frac{\xi^{-\alpha+1}}{-\alpha+1}\right|_1^0 & \text{für } \alpha \neq 1, \quad = \frac{1}{1-\alpha} \text{ für } \alpha < 1 \\ \log \xi|_0^1 & \text{für } \alpha = 1. \end{cases}$$

Im übrigen läßt sich dieses Integral mit Hilfe der Substitution $\xi = \frac{1}{\eta}$ auf das vorher behandelte zurückführen.

5.

$$\int_{0}^{1} \frac{d\xi}{\sqrt{1-\xi^2}} = \arcsin \xi|_0^1 = \frac{\pi}{2}.$$

Im Folgenden geben wir einige Konvergenzkriterien für uneigentliche Integrale an.

Aus dem Cauchyschen Hauptkriterium folgt:

Satz 7.1.
Das an der oberen Grenze uneigentliche Integral $\int_a^b f(\xi)\, d\xi$ konvergiert genau dann, wenn zu jedem $\varepsilon > 0$ eine Umgebung U von b existiert, so daß gilt:

$$\int_{x}^{y} f(\xi)\, d\xi < \varepsilon \quad \text{für alle } J \text{ angehörenden } x, y \in U.$$

Die folgenden Kriterien beziehen sich auf die absolute Konvergenz uneigentlicher Integrale. $\int_a^b f(\xi)\,d\xi$ heißt *absolut konvergent*, wenn $\int_a^b |f(\xi)|\,d\xi$ konvergiert. Es gilt zunächst der diese Bezeichnung rechtfertigende

Satz 7.2.
Aus der absoluten Konvergenz eines uneigentlichen Integrals folgt dessen Konvergenz.
Beweis:
Der Beweis ergibt sich unmittelbar aus

$$\left|\int_x^y f(\xi)\,d\xi\right| \leq \left|\int_x^y |f(\xi)|\,d\xi\right| \qquad (x, y \in J).$$

□

Ein konvergentes Integral $\int_a^b g(\xi)\,d\xi$ mit $g(\xi) \geq 0$ heißt eine *Majorante* für $\int_a^b f(\xi)\,d\xi$, wenn

$$|f(\xi)| \leq g(\xi) \quad \text{für alle} \quad \xi \in J$$

ist. Aus Satz 7.2 ergibt sich mittels Satz 2.4. ohne weiteres der

Satz 7.3.
Jedes uneigentliche Integral $\int_a^b f(\xi)\,d\xi$, das eine Majorante $\int_a^b g(\xi)\,d\xi$ besitzt, ist absolut konvergent, und es gilt dann

$$\left|\int_a^b f(\xi)\,d\xi\right| \leq \int_a^b g(\xi)\,d\xi, \textit{ falls } a < b.$$

Beispiele:

1. Die Funktion

$$Ei(x) = \int_{-\infty}^{x} \frac{e^\xi}{\xi}\,d\xi \quad (x < 0) \tag{1.12}$$

wird Exponentialintegral genannt. Daß dieses uneigentliche Integral konvergiert, zeigt man leicht durch Angabe einer Majorante. Zu jedem $x < 0$ existiert bekanntlich M, so daß

$$|\xi|e^\xi \leq M \quad \text{für} \quad \xi \leq x$$

ist.
Daher ist $\int\limits_{-\infty}^{x} \frac{M}{\xi^2} d\xi$ Majorante für (1.12). Die Substitution $e^\xi = \eta$ führt uns, wenn wir noch $e^x = y$ setzen, zum Integrallogarithmus

$$Li(y) = \int_0^y \frac{d\eta}{\log\eta} \quad (0 < y < 1).$$

2. Das Dirichlet-Integral

$$J = \int_0^\infty \frac{\sin\xi}{\xi}\,d\xi. \tag{1.13}$$

Der Nachweis der Konvergenz erfolgt nach dem Cauchyschen Hauptkriterium.
Für $0 < x < y$ ist

$$\int_x^y \frac{\sin\xi}{\xi}\,d\xi = \left[\frac{\cos\xi}{\xi}\right]_x^y - \int_x^y \frac{\cos\xi}{\xi^2}\,d\xi \ ,$$

$$\left|\int_x^y \frac{\sin\xi}{\xi}\,d\xi\right| \leq \frac{1}{x} + \frac{1}{y} + \int_x^y \frac{d\xi}{\xi^2} = \frac{2}{x},$$

und die rechte Seite $\frac{2}{x}$ strebt gegen 0 für $x \to \infty$.
Das Dirichletintegral ist übrigens nicht absolut konvergent.

3. Ein weiteres interessantes uneigentliches Integral definiert die sogenannte Γ-*Funktion (Gamma-Funktion).*
   Man überlegt sich mit Hilfe der bisherigen Sätze leicht, daß für jedes $x > 0$ das Integral
   $$\Gamma(x) = \int_0^\infty e^{-\xi}\xi^{x-1}\,d\xi$$
   konvergiert.
   Mit Hilfe partieller Integration erhält man die Funktionalgleichung
   $$\Gamma(x+1) = x\Gamma(x) \quad \text{für alle} \quad x > 0.$$
   Da $\Gamma(1) = 1$, folgt durch vollständige Induktion
   $$\Gamma(n+1) = n!\,,\; n \in \mathbb{N},$$
   d.h. die Γ-Funktion setzt die zunächst nur auf den natürlichen Zahlen erklärte Funktion $n!$ auf die positiven reellen Zahlen fort.

Satz 7.4. (Integralkriterium zum Nachweis der Konvergenz unendlicher Reihen)
Sei f Riemann-integrierbar, $f \geq 0$, $\int_1^\infty f(\xi)\,d\xi$ existiere, f sei monoton fallend und es gelte $f(n) \geq |a_n|$, dann ist die unendliche Reihe $\sum_{n=1}^\infty a_n$ konvergent.

Mit diesem Kriterium kann man z.B. die Konvergenz der Reihe $\sum_{n=1}^\infty \frac{1}{n^2}$ nachweisen.

1.8 Weitere Sätze zum Riemannschen Integral

Satz 8.1. (Der Mittelwertsatz der Integralrechnung)
Es seien $f, g\colon [a,b] \to \mathbb{R}$ beschränkt und Riemann-integrierbar, sowie $g \geq 0$ in $[a,b]$ und $m := \inf\{f(\xi) | \xi \in [a,b]\}$, $M := \sup\{f(\xi) | \xi \in [a,b]\}$. Dann gibt es ein $\mu \in [m,M]$ mit
$$\int_a^b f(x)g(x)\,dx = \mu \int_a^b g(x)\,dx.$$
Ist f stetig, so gibt es ein $\xi^* \in [a,b]$ mit
$$\int_a^b f(x)g(x)\,dx = f(\xi^*) \int_a^b g(x)\,dx.$$

Beweis:
Es gilt

$$mg(x) \leq f(x)g(x) \leq Mg(x) \quad \text{für alle } x \in [a,b].$$

Aus der Monotonie des Riemannschen Integrals folgt

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

Setze $\varphi(t) = t\int_a^b g(x)\,dx$, $\varphi\colon [m,M] \to \mathbb{R}$.
Dann nimmt φ jeden Wert zwischen $\varphi(m)$ und $\varphi(M)$ an, d.h. es gibt insbesondere ein $\mu \in [m,M]$ mit

$$\varphi(\mu) = \int_a^b f(x)g(x)\,dx.$$

Dies beweist die erste Behauptung.
Ist nun f stetig, so gibt es nach dem Zwischenwertsatz ein $\xi^* \in [a,b]$ mit $f(\xi^*) = \mu$. Daraus folgt die zweite Behauptung. □

Satz 8.2. (Die Integraldarstellung des Restgliedes in der Taylorschen Formel)
Die reellwertige Funktion f sei in einem Intervall, dem die Punkte x und $x+h$ angehören, $(n+1)$-mal stetig differenzierbar. Dann gilt:

$$f(x+h) = \sum_{k=0}^{n} \frac{f^{(k)}(x)}{k!}h^k + \frac{1}{n!}\int_0^h (h-\xi)^n f^{(n+1)}(x+\xi)\,d\xi.$$

Beweis: (Vollständige Induktion)
Für $n = 0$ lautet die obige Formel:

$$f(x+h) = f(x) + \int_0^h f'(x+\xi)\,d\xi$$

Dies ist richtig wegen des Hauptsatzes der Differential- und Integralrechnung.
Schluß von $n-1$ auf n:

Mittels partieller Integration erhält man

$$\begin{aligned}\int\limits_0^h (h-\xi)^{n-1} f^{(n)}(x+\xi)\,d\xi &= -\left[\frac{(h-\xi)^n}{n} f^{(n)}(x+\xi)\right]_0^h \\ &\quad + \frac{1}{n}\int\limits_0^h (h-\xi)^n f^{(n+1)}(x+\xi)\,d\xi \\ &= \frac{f^{(n)}(x)}{n} h^n + \frac{1}{n}\int\limits_0^h (h-\xi)^n f^{(n+1)}(x+\xi)\,d\xi.\end{aligned}$$

Gilt die Formel in der Behauptung des Satzes für $n-1$, so leitet man sie sich aus der letzten Gleichung für n her. □

Satz 8.3. (Integration unendlicher Reihen und Folgen)
Die Riemann-integrierbaren Funktionen $f_n\colon [a,b] \to \mathbb{R}$ mögen für $n \to \infty$ gleichmäßig gegen die Riemann-integrierbare Funktion $f\colon [a,b] \to \mathbb{R}$ konvergieren. Dann gilt:

$$\int\limits_a^b f_n(x)\,dx \longrightarrow \int\limits_a^b f(x)\,dx \qquad (n \to \infty).$$

Beweis:
Wegen der gleichmäßigen Konvergenz $f_n \to f$ gibt es zu jedem $\varepsilon > 0$ ein $n_0 \in \mathbb{N}$, so daß für alle $n \geq n_0$, $n \in \mathbb{N}, x \in [a,b]$ die folgende Ungleichung gilt:

$$|f_n(x) - f(x)| < \frac{\varepsilon}{2(b-a)}.$$

Aus der Monotonie des Riemannschen Integrals folgt

$$\left|\int\limits_a^b (f_n(x) - f(x))\;dx\right| \leq \int\limits_a^b \frac{\varepsilon}{2(b-a)}\,dx = \frac{\varepsilon}{2}$$

und

$$\left|\int\limits_a^b f_n(x)\,dx - \int\limits_a^b f(x)\,dx\right| \leq \varepsilon, \qquad n \geq n_0.$$

Die Behauptung ist damit bewiesen. □

Entsprechend beweist man, daß gleichmäßig gegen eine Riemann-integrable Funktion konvergente unendliche Reihen von Funktionen gliedweise integriert werden dürfen, d.h es gilt

$$\int\limits_a^b \sum_{k=1}^{\infty} f_k\,dx = \sum_{k=1}^{\infty} \int\limits_a^b f_k\,dx,$$

wenn $\sum\limits_{k=1}^{\infty} f_k$ gleichmäßig auf $[a, b]$ gegen eine integrable Funktion konvergiert und die f_k Riemann-integrierbar sind.

1.9 Numerische Integration

Häufig lassen sich die Stammfunktionen vorgegebener Funktionen nicht durch bekannte Funktionen ausdrücken, so daß man sich mit der *näherungsweisen* Berechnung von

$$\int_a^b f(x)\,dx$$

begnügen muß.
Die Riemannschen Summen bieten uns ein einfaches, aber bereits sehr brauchbares Mittel zur numerischen Approximation bestimmter Integrale

$$\int_a^b f(x)\,dx \quad (a < b).$$

Wir wählen die Zerlegung des Intervalls $[a, b]$ äquidistant (etwa $n-1$ Teilungspunkte) und setzen im Folgenden zur Abkürzung

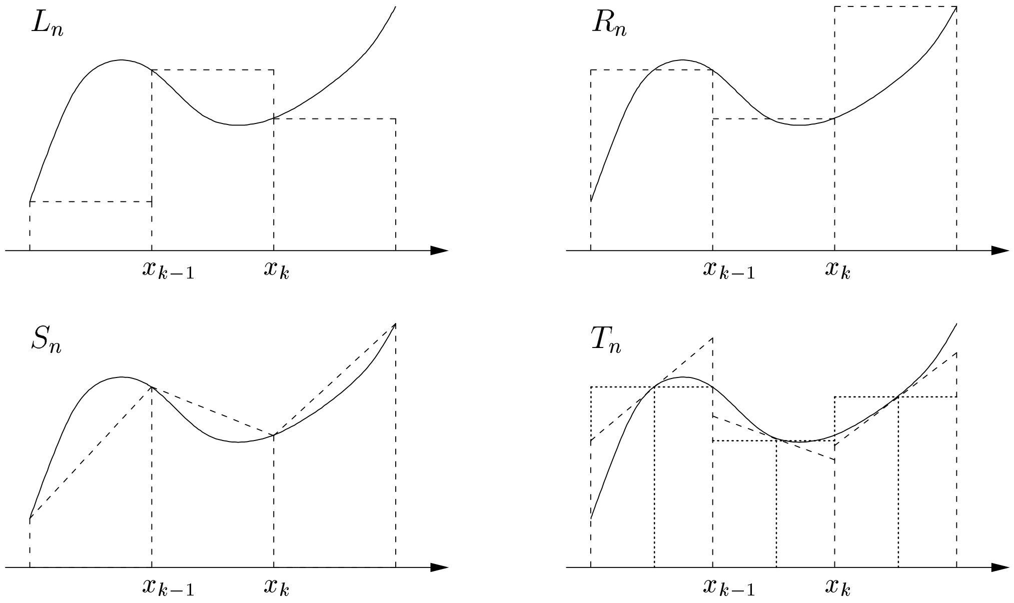
$$f\left(a + \frac{t}{n}(b-a)\right) =: f_t.$$

Von den vielen Möglichkeiten der Bildung Riemannschen Summen seien nur drei notiert:

$$\begin{aligned} L_n &= \frac{b-a}{n}\left(f_0 + f_1 + \cdots + f_{n-1}\right), \\ R_n &= \frac{b-a}{n}\left(f_1 + f_2 + \cdots + f_n\right), \\ T_n &= \frac{b-a}{n}\left(f_{\frac{1}{2}} + f_{\frac{3}{2}} + \cdots + f_{n-\frac{1}{2}}\right). \end{aligned}$$

Außerdem definieren wir

$$S_n := \frac{1}{2}\left(L_n + R_n\right) = \frac{b-a}{n}\left(\frac{1}{2}f_0 + f_1 + f_2 + \cdots + f_{n-1} + \frac{1}{2}f_n\right).$$



L_n, R_n, S_n und T_n streben geben $\int_a^b f(x)\,dx$ für $n \to \infty$. Die ersten beiden dieser Aussagen heißen Rechtecks-, die letzten beiden Trapezregeln. Das Integral wird durch Rechtecks- bzw. Trapezsummen approximiert (siehe Skizze).
Die Trapeze werden durch Sehnen (bei S_n) bzw. Tangenten (bei T_n) nach oben begrenzt; man spricht daher von Sehnen- bzw. Trapezregel.
Bei der Intetration differenzierbarer Funktionen liefert die *Simpsonsche Regel* genauere Resultate.
So wie bei den Rechtecks- und Trapezregeln die Funktion $f(x)$ durch stückweise lineare Funktionen approximiert wurde, kann man zur Approximation stückweise auch Polynome höheren Grades nehmen und bessere Resultate erwarten. In einem Teilintervall $[\alpha, \beta]$ von $[a, b]$ wählen wir als approximierendes Polynom $p(x)$ ein solches höchstens zweiten Grades, das an den Stellen $\alpha, \beta, \frac{\alpha+\beta}{2}$ dieselben Werte wie $f(x)$ hat. Ein solches p ist

$$p(x) = f(\alpha) + f(\alpha, \beta)(x - \alpha) + f\left(\alpha, \beta, \frac{\alpha+\beta}{2}\right)(x-\alpha)(x-\beta)$$

mit

$$\begin{aligned} f(\alpha, \beta) &= \frac{f(\beta) - f(\alpha)}{\beta - \alpha}, \\ f\left(\alpha, \beta, \frac{\alpha+\beta}{2}\right) &= \frac{2}{(\beta-\alpha)^2}\left(f(\beta) - 2f\left(\frac{\alpha+\beta}{2}\right) + f(\alpha)\right). \end{aligned}$$

Wegen

$$\begin{aligned}
\int_\alpha^\beta (x-\alpha)\,dx &= \frac{1}{2}(x-\alpha)^2\Big|_\alpha^\beta = \frac{1}{2}(\beta-\alpha)^2, \\
\int_\alpha^\beta (x-\alpha)(x-\beta)\,dx &= \frac{1}{2}\left[(x-\alpha)^2(x-\beta)\right]_\alpha^\beta - \frac{1}{2}\int_\alpha^\beta (x-\alpha)^2\,dx \\
&= -\frac{1}{2}\left[\frac{1}{3}(x-\alpha)^3\right]_\alpha^\beta = -\frac{1}{6}(\beta-\alpha)^3,
\end{aligned}$$

$$\begin{aligned}
f(\alpha,\beta) &= \frac{f(\beta)-f(\alpha)}{\beta-\alpha}, \\
f\left(\alpha,\beta,\frac{\alpha+\beta}{2}\right) &= \frac{2}{(\beta-\alpha)^2}\left(f(\beta)-2f\left(\frac{\alpha+\beta}{2}\right)+f(\alpha)\right)
\end{aligned}$$

gilt die „Keplersche Faßregel“

$$\int_\alpha^\beta p(x)\,dx = \frac{\beta-\alpha}{6}\left[f(\alpha)+4f\left(\frac{\alpha+\beta}{2}\right)+f(\beta)\right].$$

Man kann sich überlegen (vgl. [2, Seite 54]), daß

$$\int_\alpha^\beta f(x)\,dx = \frac{\beta-\alpha}{6}\left[f(\alpha)+4f\left(\frac{\alpha+\beta}{2}\right)+f(\beta)\right] - \frac{f^{(4)}(\eta)}{2880}(\beta-\alpha)^5$$

für geeignetes $\eta \in [\alpha,\beta]$ gilt.

$$\Rightarrow \quad \left|\int_\alpha^\beta \Big(f(x)-p(x)\Big)\,dx\right| \le \frac{M}{2880}(\beta-\alpha)^5 \text{ mit } M = \max\left\{\left|f^{(4)}(x)\right| \,\middle|\, x\in[\alpha,\beta]\right\},$$

wenn f fünfmal stetig differenzierbar ist. Will man nun $\int_a^b f(x)\,dx$ berechnen, so zerlegt man das Intervall in kleine äquidistante Teilintervalle $[x_l, x_{l+1}]$ und wendet auf diese die erhaltene Simpsonsche Formel an und summiert die Teilintegrale $\int_{x_l}^{x_{l+1}} f(x)\,dx$. Dies ergibt eine verhältnismäßig genaue Approximation des Integrals $\int_a^b f(x)\,dx$.

Kapitel 2

Differential- und Integralrechnung mehrerer Veränderlicher

2.10 Der n-dimensionale Raum

Es sei n eine natürliche Zahl. Unter dem n-dimensionalen reellen Zahlenraum (in Zeichen: $\mathbb{R}^n$) wollen wir die Menge aller geordneten n-Tupel $(\xi_1, \dots, \xi_n)$ von reellen Zahlen verstehen:

$$\mathbb{R}^n = \Big\{(\xi_1, \dots, \xi_n)\Big|\xi_i \in \mathbb{R} \quad \text{für } i = 1, \dots, n\Big\}.$$

Ein Element des $\mathbb{R}^n$ nennen wir auch *Punkt* und bezeichnen es abkürzend durch lateinische Buchstaben, z.B. $x = (\xi_1, \dots, \xi_n)$.
Auf der Menge $\mathbb{R}^n$ läßt sich die algebraische Struktur eines Vektorraums über dem Körper $\mathbb{R}$ (kurz: eines reellen Vektorraums) einführen:
Zu zwei Elementen $x = (\xi_1, \dots, \xi_n)$ und $y = (\eta_1, \dots, \eta_n)$ werde als Summe definiert:

$$x + y = (\xi_1 + \eta_1, \dots, \xi_n + \eta_n) \in \mathbb{R}^n.$$

Zu einer reellen Zahl α und einem Element $x = (\xi_1, \dots, \xi_n) \in \mathbb{R}^n$ werde als Produkt definiert:

$$\alpha x = (\alpha\xi_1, \dots, \alpha\xi_n) \in \mathbb{R}^n.$$

Mit Hilfe der Additionsaxiome für den reellen Zahlkörper prüft man leicht nach, daß $\mathbb{R}^n$ unter der eben eingeführten Addition eine kommutative Gruppe bildet. Das neutrale Element ist das n-Tupel $(0, \dots, 0)$ der „Nullvektor “ oder „Nullpunkt“ des $\mathbb{R}^n$, den wir der Einfachheit halber als 0 bezeichnen, sofern Mißverständnisse nicht zu befürchten sind. – Ebenso verifiziert man unter Hinzuziehung der Multiplikations- und Distributivitäts-Axiome von $\mathbb{R}$ folgende Regeln

$$\begin{aligned}
(\alpha + \beta)x &= \alpha x + \beta x, \\
\alpha(x + y) &= \alpha x + \alpha y, \\
\alpha(\beta x) &= (\alpha\beta)x, \\
1 \cdot x &= x
\end{aligned}$$

für alle $x, y \in \mathbb{R}^n$, $\alpha, \beta \in \mathbb{R}$. Steht bei einer Betrachtung die Vektorraumstruktur des $\mathbb{R}^n$ im Vordergrund, so wird man die Elemente des $\mathbb{R}^n$ als *Vektoren* bezeichnen.

Definition:
Sind $x = (\xi_1, \ldots, \xi_n)$ und $y = (\eta_1, \ldots, \eta_n)$ Vektoren aus $\mathbb{R}^n$, so wird die reelle Zahl

$$\sum_{i=1}^{n} \xi_i \eta_i$$

das Skalarprodukt *von x und y genannt und mit $x \cdot y$ bezeichnet*

Mit Hilfe der Körperaxiome von $\mathbb{R}$ erkennt man sofort die Richtigkeit von

Satz 10.1.
Das Skalarprodukt genügt folgenden Regeln:

$$\begin{array}{lrcl}
(a) & x \cdot y & = & y \cdot x, \\
(b) & (x_1 + x_2) \cdot y & = & x_1 \cdot y + x_2 \cdot y, \\
(c) & (\alpha x) \cdot y & = & \alpha(x \cdot y), \\
(d) & x \cdot x \geq 0, & & x \cdot x = 0 \Leftrightarrow x = 0
\end{array}$$

für alle $x, x_1, x_2, y \in \mathbb{R}^n$, $\alpha \in \mathbb{R}$.

Für $x \cdot x$ schreiben wir auch x^2. Mit Hilfe des Skalarproduktes definieren wir eine *Norm* genannte Abbildung des $\mathbb{R}^n$ nach $\mathbb{R}$, indem wir jedem $x \in \mathbb{R}^n$ als Norm die Zahl $\|x\| := \sqrt{x^2}$ zuordnen.

Satz 10.2.
Die Norm hat die folgenden Eigenschaften:

$$\begin{array}{lrcl}
(1) & \|x\| \geq 0, & & \|x\| = 0 \Leftrightarrow x = 0, \\
(2) & \|\alpha x\| & = & |\alpha| \cdot \|x\|, \\
(3) & \|x + y\| & \leq & \|x\| + \|y\|
\end{array}$$

für alle $x, y \in \mathbb{R}^n$, $\alpha \in \mathbb{R}$.

Regel (1) ist die Übersetzung der Regel (d) für das Skalarprodukt, Regel (2) folgt sofort aus Regel (c). Um Regel (3) zu verifizieren, beweisen wir zuerst den

Satz 10.3. (Schwarzsche Ungleichung)
Für zwei Vektoren $x, y \in \mathbb{R}^n$ gilt stets:

$$(x \cdot y)^2 \leq x^2 \cdot y^2.$$

Das Gleichheitszeichen gilt hierbei genau dann, wenn x und y linear abhängig sind.

Beweis:

Ist $y = 0$, so hat man

$$x \cdot 0 = x(0+0) = x \cdot 0 + x \cdot 0,$$

also $x \cdot 0 = 0$. In diesem Fall verschwinden beide Seiten der behaupteten Ungleichung.

Ist $y \neq 0$, so gilt wegen (1) auch $\|y\| \neq 0$. Wenn wir noch $y^2 = \|y\|^2$ bedenken, können wir für beliebiges $t \in \mathbb{R}$ schreiben:

$$0 \leq (x+ty)^2 = \left(\frac{xy}{\|y\|} + t\|y\|\right)^2 + x^2 - \frac{(xy)^2}{\|y\|^2}.$$

Nun kann man t so wählen, daß $\frac{xy}{\|y\|} + t\|y\| = 0$ ist. Dann erhält man

$$0 \leq x^2 - \frac{(xy)^2}{y^2}$$

und daraus die Behauptung. Sind x und y linear unabhängig, so verschwindet $x + ty$ für kein t, also ist stets $0 < (x+ty)^2$, und damit wird die behauptete Ungleichung streng. Sind x und y linear abhängig, so gibt es wegen $y \neq 0$ ein t_0 so, daß $x + t_0 y = 0$. Dann ist aber

$$\frac{xy}{\|y\|} + t_0\|y\| = -t_0\|y\| + t_0\|y\| = 0$$

und wir bekommen das Gleichheitszeichen in der Schwarzschen Ungleichung. $\square$

Folgerung:

$|x \cdot y| \leq \|x\| \cdot \|y\|$ für $x, y \in \mathbb{R}^n$.

Das folgt durch Wurzelziehen aus der Schwarzschen Ungleichung.

Zum Nachweis von (3) in Satz 10.2 betrachten wir:

$$\begin{aligned}
\|x+y\|^2 &= (x+y)^2 \\
&= x^2 + 2x \cdot y + y^2 \quad \text{nach (a) und (b)} \\
&\leq x^2 + 2|x \cdot y| + y^2 \\
&\leq \|x\|^2 + 2\|x\| \cdot \|y\| + \|y\|^2 \quad \text{nach der Folgerung} \\
&= (\|x\| + \|y\|)^2.
\end{aligned}$$

Durch Wurzelziehen erhält man (3).

Ist auf dem $\mathbb{R}^n$ eine reellwertige Funktion gegeben, die den Regeln (1) bis (3) genügt, so nennen wir diese Funktion eine *Norm* und sprechen von einem *normierten* reellen Vektorraum.

Die oben mittels des Skalarproduktes definierte Funktion wollen wir die *euklidische* Norm nennen.
Man kann dem $\mathbb{R}^n$ auch andere Normen aufprägen. Bei späteren Untersuchungen werden wir oft für $x = (x_1, \dots, x_n)$ setzen:

$$|x| = \max_{i=1,\dots,n} |x_i|.$$

Man verifiziert leicht (1) und (2), (3) folgt so:

$$|x+y| = \max_i |x_i + y_i| \leq \max_i (|x_i| + |y_i|) \leq \max_i |x_i| + \max_i |y_i| = |x| + |y|.$$

Mit Hilfe der euklidischen Norm wollen wir jetzt den *euklidischen Abstand (Distanz)* zweier Punkte des $\mathbb{R}^n$ definieren, indem wir für $x, y \in \mathbb{R}^n$ setzen:

$$\operatorname{dist}(x,y) := ||y - x|| .$$

Satz 10.4.
Die Distanz hat folgende Eigenschaften:

$$\operatorname{dist}(x,y) \geq 0, \quad \operatorname{dist}(x,y) = 0 \Leftrightarrow x = y \tag{2.1}$$
$$\operatorname{dist}(x,y) = \operatorname{dist}(y,x), \tag{2.2}$$
$$\operatorname{dist}(x,z) \leq \operatorname{dist}(x,y) + \operatorname{dist}(y,z) \tag{2.3}$$

für alle $x, y, z \in \mathbb{R}^n$.

Beweis:
(2.1) ist die Übersetzung der Regel (1) für die Norm;

(2.2) folgt aus (2) für $\alpha = -1$;

(2.3) folgt aus (3) so:

$$\operatorname{dist}(x,z) = \|z - x\| = \|z - y + y - x\| \leq \|z - y\| + \|y - x\| = \operatorname{dist}(y,z) + \operatorname{dist}(x,y).$$

Ist zu einer beliebigen Menge $X = \{x, y, \dots\}$ eine Funktion gegeben, die jedem Paar (x, y) von Elementen von X eine reelle Zahl $\operatorname{dist}(x, y)$ zuordnet, und genügt diese Funktion den Regeln (2.1) bis (2.3), so sagt man, sie sei eine *Metrik* auf X und man nennt X einen *metrischen Raum.*

2.11 Topologie des $\mathbb{R}^n$

Definition:
Es sei $x_0 \in \mathbb{R}^n$ *und* ε *eine positive reelle Zahl. Eine* ε-Umgebung *von* x_0 *ist eine Menge der Gestalt*

$$U_\varepsilon(x_0) = \{x \in \mathbb{R}^n : |x - x_0| < \varepsilon\}.$$

Dabei ist $|x| = \max_{i=1\dots n} |x_i|$.

$U_\varepsilon(x_0)$ ist also ein n-dimensionaler Würfel mit Mittelpunkt x_0, Kantenlänge 2ε und zu den Koordinatenachsen parallelen Kanten. Offenbar gilt stets $x_0 \in U_\varepsilon(x_0)$, und für $0 < \varepsilon_1 \leq \varepsilon_2$ ist $U_{\varepsilon_1}(x_0) \subset U_{\varepsilon_2}(x_1)$.

Definition:
Es sei $M \subset \mathbb{R}^n$. Ein Punkt $x_0 \in \mathbb{R}^n$ heißt innerer Punkt *von M, wenn es ein $\varepsilon > 0$ gibt, so daß $U_\varepsilon(x_0) \subset M$.*

Definition:
Unter einer Umgebung *eines Punktes $x \in \mathbb{R}^n$ verstehen wir eine Punktmenge, die x als inneren Punkt enthält.*
Eine Menge $U \subset \mathbb{R}^n$ ist also genau dann eine Umgebung von x, wenn sie eine ε-Umgebung von x enthält. Insbesondere sind ε-Umgebungen von x auch Umgebungen von x. Wir werden Umgebungen von x oft mit $U(x)$ bezeichnen.

Definition:
Es sei $M \subset \mathbb{R}^n$. M heißt offen *oder ein* Bereich, *wenn jeder Punt von M innerer Punkt von M ist.*

Beispiel:
Es seien $a_1, \dots, a_n, b_1, \dots, b_n, \in \mathbb{R}$ und $a_k < b_k$ für $k = 1, \dots, n$. Dann ist

$$Q = \Big\{(x_1, \dots, x_n) \in \mathbb{R}^n \ : \ a_k < x_k < b_k \ ; \ k = 1, \dots, n\Big\}$$

eine offene Menge (offener Quader):
Für $\vec{x}_0 = ({x_1}^{(0)}, \dots, {x_n}^{(0)}) \in Q$ setzen wir

$$\varepsilon = \min_{k=1,\dots,n} \{({x_k}^{(0)} - a_k), (b_k - {x_k}^{(0)})\}.$$

Dann ist offenbar $U_\varepsilon(\vec{x}_0) \subset Q$.

Satz 11.1.
Das System der offenen Mengen hat die folgenden Eigenschaften:

(1) $\varnothing$ und $\mathbb{R}^n$ sind offen,

(2) die Vereinigung beliebig vieler offener Mengen ist offen,

(3) der Durchschnitt endlich vieler offener Mengen ist offen.

Beweis:

(1) ist trivial.

(2) sieht man so: Sind U_i, wobei i eine beliebige Indexmenge J durchläuft, offene Mengen, und ist $\vec{x}_0 \in \bigcup_{i\in J} U_i$, so gehört $\vec{x}_0$ mindestens einem U_{i_0} an. Dann gibt es $\varepsilon > 0$, so daß $U_\varepsilon(\vec{x}_0) \subset U_{i_0} \subset \bigcup_{i\in J} U_i$.

(3) ergibt sich auf folgende Weise: Sind $U_1, \dots, U_i$ offen und ist $\vec{x}_0 \in \bigcap_{\nu=1}^{i} U_\nu$. Dann ist $\varepsilon = \min_{\nu=1,\dots,i} \varepsilon > 0$ und $U_\varepsilon(\vec{x}_0) \subset U_{\varepsilon_\nu}(x_o) \subset U_\nu$ für $\nu = 1, \dots, i$, also $U_\varepsilon(\vec{x}_0) \subset \bigcap_{\nu=1}^{i} U_\nu$.

□

Der Durchschnitt beliebig vieler offener Mengen ist nicht notwendig offen: $M_k = \{\vec{x} : |\vec{x}| < \frac{1}{k}\}$ ist für $k = 1, 2, 3, \dots$ ein offener Quader, $\bigcap_{k=1}^{\infty} M_k = \{0\}$ ist offenbar keine offene Menge. Hat man auf einer beliebigen Menge X ein System von Teilmengen mit den Eigenschaften (1)-(3) aus Satz 11.1 (dabei ersetze man in (1) die Menge $\mathbb{R}^n$ durch X), so nennt man das System eine *Topologie* auf X und und die einzelnen Mengen des Systems die „offenen Mengen" der Topologie. x zusammen mir der Topologie auf X heißt *topologischer Raum.* Auf einem beliebigen metrischen Raum kann auch wörtlich wie auf dem $\mathbb{R}^n$ eine Topologie definieren.

Definition:
Eine Teilmenge $M \subset \mathbb{R}^n$ heißt abgeschlossen, *wenn ihr Komplement offen ist.*

Satz 11.2.
Das System der abgeschlossenen Mengen hat die folgenden Eigenschaften:

(1) $\varnothing$ und $\mathbb{R}^n$ sind abgeschlossen,

(2) der Durchschnitt beliebig vieler abgeschlossener Mengen ist abgeschlossen,

(3) die Vereinigung endlich vieler abgeschlossener Mengen ist abgeschlossen.

Der Beweis ergibt sich sofort aus Satz 11.1 und allgemeinen mengentheoretischen Sätzen.

Definition:
Ist $M \subset \mathbb{R}^n$, so heißt ein Punkt $x \in \mathbb{R}^n$ Randpunkt *von M, wenn in jeder Umgebung von x sowohl Punkte aus M als auch Punkte aus dem Komplement von M liegen.*
Mit ∂M bezeichnen wir die Menge aller Randpunkte von M.

Definition:
Die Menge $\overline{M} = M \cup \partial M$ heißt die abgeschlossene Hülle *M oder auch die Abschließung von M.*

Die Bezeichnung „abgeschlossenen Hülle“ wird durch Satz 11.4 gerechtfertigt.
Zunächst gilt:

Satz 11.3.
Sei $M \subset \mathbb{R}^n$. Dann gilt $\overline{M} = \overline{(\overline{M})}$.
Beweis:
Aus der Definition folgt $\overline{(\overline{M})} = \overline{M} \cup \partial\bar{M} \supset \overline{M}$; es bleibt zu zeigen, daß $\overline{M} \supset \partial\overline{M}$. Hierzu sei $x \in \partial\overline{M}$. Angenommen, es wäre $x \notin \overline{M}$, d.h. also $x \in \complement\overline{M}$. Wegen $x \in \partial\overline{M}$ gibt es zu jeder offenen Umgebung U von x ein $y \in \overline{M}$ mit $y \in U$.
In jeder offenen Umgebung von y muß ein Punkt $\tilde{y} \in M$ liegen (entweder $\tilde{y} = y$, wenn $y \in M$ oder, falls $y \in \partial M$, aufgrund der Definition der Begriffs „Randpunkt“.)
Da U offen ist, ist U auch Umgebung von y, und wir haben die Situation, daß in jeder offenen Umgebung von $x \in \complement\overline{M} \subset \complement M$ ein Punkt $\tilde{y} \in M$ liegt. Daraus folgt aber, daß $x \in \partial M$ gelten muß, im Widerspruch zu $x \in \complement\overline{M}$. Also ist $x \in \overline{M} und \partial\overline{M} \subset \overline{M}$. $\square$

Satz 11.4.
Es sei $M \subset \mathbb{R}^n$. Dann ist $\overline{M}$ abgeschlossen. Überdies ist $\overline{M}$ die kleinste abgeschlossene in $\mathbb{R}^n$ liegende Obermenge von M.
Beweis:
Es sei $x \in \complement\overline{M}$. Zum Nachweis der Abgeschlossenheit von M muß gezeigt werden, daß $\complement\overline{M}$ offen ist. Wir müssen also zeigen, daß eine in $\complement\overline{M}$ liegende Umgebung von x existiert. Wenn dies nicht der Fall wäre, so enthielte jede Umgebung von x mindestens einen Punkt von $\overline{M}$, d.h., der Punkt $x \in \complement\bar{M}$ wäre Randpunkt von $\bar{M}$ Dies bedeutet $x \in \partial\overline{M} \subset \overline{(\overline{M})}$. Aus Satz 11.3. folgt dann $x \in \overline{M}$ im Widerspruch zu $x \in \complement\overline{M}$.
Zum Nachweis der zweiten Behauptung ist zu zeigen: Wenn N eine abgeschlossene Teilmenge von $\mathbb{R}^n$ ist, die M enthält, so gilt $N \supset \overline{M}$.
Da $N \supset M$, genügt es, die Inklusion $N \supset \partial M$ nachzuweisen. Hierzu wiederum genügt der Nachweis der Inklusion $N \supset \{x \in \partial M | x \notin M\}$. Sei daher $x \in \partial M,\ x \notin M$. Angenommen, es wäre $n \notin N$. Wegen der Abgeschlossenheit von N ist $\complement N$ offen, d.h. es gibt eine Umgebung U von x, die in $\complement N$ ligt. Da $x \in \partial N$, liegt in jeder Umgebung von x ein Punkt von M, d.h. das obige $U \subset \complement N$ enthält mindestens einen Punkt von M. Da $M \subset N$ ist, ergibt sich ein Widerspruch. $\square$

Den zweiten Teil von Satz 11.4. kann man auch folgendermaßen ausdrücken:
$\overline{M}$ ist der Durchschnitt aller abgeschlossenen Obermengen von M.

Definition:
Die Menge aller inneren Punkte einer Menge $M \subset \mathbb{R}^n$ wird als „offener Kern $\overset{\circ}{M}$ von M“ oder als das „Innere von M“ bezeichnet.

Satz 11.5.
Es sei $M \subset \mathbb{R}^n$. Dann ist die Menge $\overset{\circ}{M}$ offen. Ferner ist $\overset{\circ}{M}$ die größte offene Teilmenge von M.
Beweis:
Es sei $x \in \overset{\circ}{M}$, d.h. x ist innerer Punkt von M. Dann gibt es eine offene Umgebung U von x mit $U \subset M$.
Da U offen ist, gibt es zu jedem $y \in U$ eine Umgebung V_y von y mit $V_y \subset U \subset M$, d.h. es gilt auch $y \in \overset{\circ}{M}$ und $U \subset \overset{\circ}{M}$. Daraus folgt die Offenheit von $\overset{\circ}{M}$. □

Zum Nachweis der zweiten Behauptung muß gezeigt werden, daß aus $W \subset M$, W offen, die Inklusion $W \subset \overset{\circ}{M}$ folgt. Dies folgt unmittelbar aus der Offenheit von W. □

Wir notieren noch

Sätzchen 11.6.
Es sei $M \subset \mathbb{R}^n$. Dann gilt

$$\overset{\circ}{\left(\overset{\circ}{M}\right)} = \overset{\circ}{M} \, .$$

Beweis:
Nach Satz 11.5. ist $\overset{\circ}{\overset{\circ}{M}}$ die größte offene Menge, die in $\overset{\circ}{M}$ enthalten ist. Da $\overset{\circ}{M}$ offen ist, ist dies die Menge $\overset{\circ}{M}$ selbst. □

Eine weitere wichtige „topologische Erkenntnis" ist

Satz 11.7. (Hausdorffsches Trennungsaxiom)
Es seien x_1 und x_2 verschiedene Punkte des $\mathbb{R}^n$. Dann gibt es Umgebungen $U(x_1)$ und $U(x_2)$, so daß $U(x_1) \cap U(x_2) = \varnothing$.
Beweis:
Man setze $\varepsilon := \frac{1}{4}|x_2 - x_1| > 0$ und $U(x_k) = U_\varepsilon(x_k)$ für $k = 1, 2$. Gäbe es ein $x \in U_\varepsilon(x_1) \cap U_\varepsilon(x_2)$, so folgte nach der Dreicksungleichung der Widerspruch $4\varepsilon = |x_1 - x_2| \leq |x_2 - x| + |x - x_1| \leq \varepsilon + \varepsilon = 2\varepsilon$. □
Folgerung:
Eine aus einem Punkt bestehende Menge x_0 ist abgeschlossen.
Zu jedem $x \in \complement\{x_0\}$ gibt es nämlich nach Satz 11.7. eine ε-Umgebung $U_\varepsilon(x)$, die x_0 nicht enthält, die also ganz in $\complement\{x_0\}$ liegt. Also ist $\complement\{x_0\}$ offen.

Definition:
Es sei $M \subset \mathbb{R}^n$. Ein Punkt $x_0 \in \mathbb{R}^n$ heißt Häufungspunkt *von M, wenn in jeder Umgebung von x_0 unendlich viele Punkte von M liegen.*

Es genügt, hierbei die offenen Umgebungen von x_0 zu betrachten.

Satz 11.8.
Eine Teilmenge $M \subset \mathbb{R}^n$ ist genau dann abgeschlossen, wenn sie alle ihre Häufungspunkte enthält.
Beweis:

$\Rightarrow$ Es sei M abgeschlossen und x ein Häufungspunkt von M. Läge x in $\complement M$, so gibt es wegen der Offenheit von $\complement M$ eine Umgebung U von x mit $U \subset \complement M$. U kann daher nicht unendlich viele Punkte von M enthalten, und x ist kein Häufungspunkte von M. Dies ist ein Widerspruch.

$\Leftarrow$ M enthalte alle seine Häufungspunkte. Wir müssen zeigen, daß $\complement M$ offen ist. Hierzu sei $x \in \complement M$. Dann gibt es eine offene Umgebung U von x, so daß $U \cap M$ nur endlich viele Elemente enthält. Andernfalls enthielte jede Umgebung von x unendlich viele Elemente aus M und x wäre Häufungspunkt von M, also $x \in M$. Die Menge $U \cap M$ ist daher abgeschlossen, ihr Komplement $\complement(U \cap M) = \complement U \cup \complement M$ offen. Da U offen ist, ist auch $U \cap (\complement U \cup \complement M)$ offen, und wegen der Identität

$$U \cap (\complement U \cup \complement M) = (U \cap \complement U) \cup (U \cap \complement M) = \varnothing \cup (U \cap \complement M) = U \cap \complement M$$

ist die Menge $U \cap \complement M$ offen. Da $x \in U \cap \complement M$, ist $U \cap \complement M$ eine Umgebung von x, die in $\complement M$ liegt. M ist daher abgeschlossen. $\square$

2.11.1 Kompakte Mengen

Definition:
Eine Teilmenge $M \subset \mathbb{R}^n$ heißt beschränkt, *wenn ein $C \in \mathbb{R}$ existiert mit $|x| \leq C$ für alle $x \in M$.*

Definition:
Eine Teilmenge $M \subset \mathbb{R}^n$ heißt kompakt, *wenn M abgeschlossen und beschränkt ist.*

Kompakte Mengen haben eine wichtige Eigenschaft, welche bei Beweisen häufig verwendet wird. Es ist dies die sogenannte *Überdeckungseigenschaft*:

Definition:
Ein System $\{U_i,\ i \in J\}$ von Mengen $U \subset \mathbb{R}^n$ mit einer Indexmenge J ist eine Überdeckung *der Menge $M \subset \mathbb{R}^n$, wenn zu jedem $x \in M$ ein $i \in J$ existiert mit $x \in U_i$.*
Das System $\{U_i,\ i \in J\}$ heißt offene Überdeckung *von M, wenn die U_i zusätzlich offene Mengen sind.*

Definition:
Eine Menge $M \subset \mathbb{R}^n$ hat die Überdeckungseigenschaft, *wenn jede offene Überdeckung* $\{U_i,\ i \in J\}$ *von* M *eine endliche Teilüberdeckung besitzt, d.h. es gibt eine endliche Teilmenge* $J_0 \subset J$, *so daß* $\{U_i,\ i \in J_0\}$ *immer noch eine Überdeckung von* M *ist.*

Satz 11.9. (Satz von Heine-Borel)
Eine Menge $M \subset \mathbb{R}^n$ *ist genau dann kompakt, wenn* M *die Überdeckungseigenschaft besitzt.*
Beweis:
Der Beweis kann nachgelesen werden in [4, Seite 30] oder [6].

2.12 Punktfolgen und Konvergenz

Ist jeder natürlichen Zahl $i = 1, 2, 3, \ldots$ ein Punkt $a_i \in \mathbb{R}^n$ zugeordnet, so bilden die $\vec{a_i}$ $, i = 1, 2, 3, \ldots$ eine *Punktfolge* im $\mathbb{R}^n$. Präzise ausgedrückt:

Eine *Punktfolge* $(\vec{a_i})_{i=1}^{\infty}$ oder $(\vec{a_i})_{i\in\mathbb{N}}$ oder $(\vec{a_1}, \vec{a_2}, \ldots)$ ist eine Abbildung der natürlichen Zahlen in den $\mathbb{R}^n$. Das Bild der Zahl $i \in \mathbb{N}$ ist $\vec{a_i} \in \mathbb{R}^n$. Jedes einzelne $\vec{a_i}$, $i \in \mathbb{N}$, heißt *Glied* der Folge.

Definition:
Es sei $(\vec{a_i})_{i=1}^{\infty}$ *eine Punktfolge in* $\mathbb{R}^n$. *Ein Punkt* $\vec{a}$ *heißt* Häufungspunkt *von* $(\vec{a_i})$, *wenn in jeder Umgebung von* $\vec{a_i}$ *unendlich viele Glieder der Folge* $(\vec{a_i})$ *liegen.*

Definition:
Eine Punktfolge $(\vec{a_i})$ *im* $\mathbb{R}^n$ *heißt* beschränkt, *wenn die Punktmenge* $\{\vec{a_i} | i \in \mathbb{N}\}$ *beschränkt ist.*

Definition:
Eine Punktfolge $(\vec{a_i})_{i\in\mathbb{N}}$ *im* $\mathbb{R}^n$ konvergiert gegen den Punkt $\vec{a} \in \mathbb{R}^n$, *in Zeichen:*

$$\vec{a_i} \to \vec{a} \ \ (i \to \infty) \qquad \textit{oder} \qquad \lim_{i\to\infty} \vec{a_i} = \vec{a},$$

wenn in jeder Umgebung von $\vec{a}$ *fast alle Glieder der Folge* $(\vec{a_i})_{i\in\mathbb{N}}$ *liegen.*

Äquivalente Definition:
Eine Punktfolge $(\vec{a_i})_{i\in\mathbb{N}}$ konvergiert gegen $\vec{a} \in \mathbb{R}^n$, *wenn zu jedem* $\varepsilon > 0$ *ein* $i_0 \in \mathbb{N}$ *existiert, so daß für alle* $i \geq i_0$ *die Ungleichung*

$$|\vec{a} - \vec{a_i}| < \varepsilon$$

stattfindet.

Definition:
Die Punktfolge $(\vec{a}_i)_{i\in\mathbb{N}}$ *heißt* konvergent, *wenn ein* $\vec{a}\in\mathbb{R}^n$ *mit*

$$\vec{a}_i \to \vec{a} \qquad (i\to\infty)$$

existiert.

Aus der Definition folgt leicht der

Satz 12.1.
Eine Punktfolge $(\vec{a}_i)$ *im* $\mathbb{R}^n$ *konvergiert genau dann gegen einen Punkt* $\vec{a}\in\mathbb{R}^n$, *wenn die Komponenten* $\alpha_i^{(k)}$ *von* $\vec{a}_i = (\alpha_i^{(1)},\dots,\alpha_i^{(n)})$ *für* $i\to\infty$ *gegen die Komponenten* $\alpha^{(k)}$ *von* $\vec{a}=(\alpha^{(1)},\dots,\alpha^{(n)})$ *konvergieren.*

Hieraus folgen leicht die Sätze 12.2 und 12.3:

Satz 12.2.
Es seien $(\vec{x}_i)_{i\in\mathbb{N}}$ *und* $(\vec{y}_i)_{i\in\mathbb{N}}$ *Punktfolgen im* $\mathbb{R}^n$ *mit*

$$\vec{x}_i\to\vec{x},\quad \vec{y}_i\to\vec{y}\quad (i\to\infty),\ \vec{x},\vec{y}\in\mathbb{R}^n.$$

Dann gilt:

$$\begin{aligned}\lim_{i\to\infty}(\vec{x}_i+\vec{y}_i) &= \vec{x}+\vec{y},\\ \lim_{i\to\infty}\vec{x}_i\cdot\vec{y}_i &= \vec{x}\cdot\vec{y},\\ \lim_{i\to\infty}\alpha_i\vec{x}_i &= \alpha\vec{x}_i \qquad \textit{für}\quad \alpha_i\to\alpha\ \ (i\to\infty),\quad \alpha_i,\alpha\in\mathbb{R}.\end{aligned}$$

Satz 12.3. (Cauchy-Kriterium)
Es sei $(\vec{a}_i)_{i\in\mathbb{N}}$ *eine Punktfolge im* $\mathbb{R}^n$. *Die Folge* $(\vec{a}_i)_{i\in\mathbb{N}}$ *konvergiert genau dann, wenn zu jedem* $\varepsilon>0$ *ein* $m_0\in\mathbb{N}$ *existiert, so daß für alle* $k,m\in\mathbb{N}$ *mit* $k,m\geq m_0$ *die Ungleichung*

$$|\vec{a_k}-\vec{a_m}|<\varepsilon$$

stattfindet.

Satz 12.4. (Satz von Bolzano-Weierstraß)
Jede beschränkte Punktfolge $(\vec{a}_i)_{i\in\mathbb{N}}$ *im* $\mathbb{R}^n$ *hat mindestens einen Häufungspunkt.*
Beweis:
Aus der Voraussetzung folgt, daß die Folgen der Komponenten $(\alpha_i^{(k)})_{i\in\mathbb{N}}, k=1\dots,n$, von $\vec{a}_i=(\alpha_i^{(1)},\dots,\alpha_i^{(n)})$ beschränkt sind. Es gibt daher eine Teilfolge $(\varphi_1(i))_{i\in\mathbb{N}}, \varphi(i)\in\mathbb{N}$,

so daß die Folge $(\alpha^{(1)}_{\varphi_1(i)})_{i\in\mathbb{N}}$ konvergiert. Dies folgt aus dem „eindimensionalen“ Satz von Bolzano-Weierstraß, angewandt auf $(\alpha^{(1)}_i)_{i\in\mathbb{N}}$. Erneute Anwendung dieses Satzes auf die Folge $(\alpha^{(2)}_{\varphi_1(i)})_{i\in\mathbb{N}}$ ergibt die Existenz einer Teilfolge $(\varphi_2(i))_{i\in\mathbb{N}}$ von $(\varphi_1(i))_{i\in\mathbb{N}}$, so daß $(\alpha^{(2)}_{\varphi_2(i)})_{i\in\mathbb{N}}$ konvergiert und $(\alpha^{(1)}_{\varphi_2(i)})_{i\in\mathbb{N}}$ als Teilfolge von $(\alpha^{(1)}_{\varphi_1(i)})_{i\in\mathbb{N}}$ konvergiert. n-malige Auswahl von Teilfolgen nach diesem Vorbild ergibt die Existenz einer Teilfolge $(\varphi_n(i))_{i\in\mathbb{N}}$, so daß $(\alpha^{(k)}_{\varphi_n(i)})_{i\in\mathbb{N}}$ für $k = 1,\dots,n$ konvergiert. Dies bedeutet aber, daß $(\vec{a}_{\varphi_n(i)})_{i\in\mathbb{N}}$ konvergiert, und der Satz ist bewiesen. □

Satz 12.5.
Es sei a ein Häufungspunkt der Punktfolge $(a_i)_{i\in\mathbb{N}}$. Dann gibt es eine Teilfolge $(a_{\varphi(i)})_{i\in\mathbb{N}}$, die gegen a konvergiert.

Beweis:
Der Beweis geschieht wie im Fall reeller Zahlenfolgen. □
Der Satz von Bolzano-Weierstraß kann damit auch so formuliert werden:

Jede beschränkte Folge besitzt eine konvergente Teilfolge.

2.13 Funktion von n Veränderlichen, Stetigkeit

Naive Definition:
Es seien $m,n \in \mathbb{N}$ und $M \subset \mathbb{R}^n$, $N \subset \mathbb{R}^m$. Eine Funktion von M nach N, in Zeichen $f\colon M \to N$, ist eine Zuordnungsvorschrift, die jedem Element x aus M genau ein Element $f(x)$ aus N zuordnet.
M heißt Definitionsbereich *von f oder* Urbildmenge*, N heißt* Bildbereich *von f.*
In dem vorliegenden Fall $M \subset \mathbb{R}^n$, $N \subset \mathbb{R}^m$ heißt f auch „Funktion von n Veränderlichen“ oder „Funktion mehrerer Variabler“. Ist $m = 1$, spricht man auch von einer *reellen Funktion*, ist $m \geq 2$, so spricht man auch von einer *Vektorfunktion*.

Definition:
Sei $M \subset \mathbb{R}^n$ und $x \in M$. Eine Funktion $f\colon M \to \mathbb{R}^n$ heißt stetig im Punkte *x, wenn zu jeder Umgebung U von $f(x)$ eine Umgebung V von x existiert, so daß $\{f(y)|y \in V \cap M\} \subset U$.*
Eine äquvalante Definition mit Epsilontik lautet folgendermaßen:
Eine Funktion $f\colon M \to \mathbb{R}^m$ heißt *stetig im Punkte x*, wenn zu jedem $\varepsilon > 0$ ein $\delta > 0$ existiert, so daß für alle $y \in M$ mit $|x - y| < \delta$ die Ungleichung $|f(x) - f(y)| < \varepsilon$ stattfindet.
Wie im Fall einer Variablen beweist man

Satz 13.1.
Es seien $n,m \in \mathbb{N}$, $x_0 \in M \subset \mathbb{R}^n$ und $f\colon M \to \mathbb{R}^m$ stetig in x_0. Dann gibt es eine Umgebung U von x_0, in der f beschränkt ist, d.h. es gibt ein $C \in \mathbb{R}$, so daß $|f(x)| \leq C$ für $x \in U \cap M$.

Auch der folgende Satz ist im n-dimensionalen gültig:

Satz 13.2.
Es seien $n, m \in \mathbb{N}$, $x_0 \in M \subset \mathbb{R}^n$ und die Funktionen $f, g\colon M \to \mathbb{R}^m$ seien stetig in x_0. Dann ist auch $f+g$, αf mit $\alpha \in \mathbb{R}$, sowie $f \cdot g$ (Skalarprodukt!) in x_0 stetig. Ist $m = 1$ und $g(x_0) \neq 0$, so ist auch $\frac{f}{g}$ in einer Umgebung von x_0 erklärt und in x_0 stetig.
Beweis:
Wir beweisen nur die letzte Behauptung.
Ist $g(x_0) \neq 0$, so gibt es eine Umgebung $U(x_0)$, in der gilt

$$\Big|g(x)\Big| > \beta > 0 \quad \text{mit} \quad \beta = \frac{1}{2}\Big|g(x_0)\Big| \quad \text{für alle} \quad x \in U(x_0).$$

Wegen der Stetigkeit von g gibt es nämlich zu $\varepsilon = \beta$ eine Umgebung $U(x_0)$, so daß

$$\Big|g(x) - g(x_0)\Big| < \beta = \varepsilon.$$

Daraus folgt

$$\Big|g(x_0)\Big| - \Big|g(x)\Big| < \beta = \frac{|g(x_0)|}{2}$$

und

$$\Big|g(x)\Big| > \frac{|g(x_0|}{2} = \beta.$$

$\frac{f}{g}$ ist also in $U(x_0)$ erklärt.
Es gilt

$$\begin{aligned}
\left|\frac{f(x)}{g(x)} - \frac{f(x_0)}{g(x_0)}\right| &= \frac{|f(x)g(x_0) - g(x)f(x_0)|}{|g(x)||g(x_0)|} \\
&\leq \frac{1}{\beta^2}\Big|f(x)g(x_0) - g(x)f(x_0)\Big| \\
&= \frac{1}{\beta^2}\Big|f(x)(g(x_0) - g(x)) + g(x)(f(x) - f(x_0))\Big| \\
&\leq \frac{1}{\beta^2}\Big|f(x)\Big| \cdot \Big|g(x_0) - g(x)\Big| + \frac{1}{\beta^2}\Big|g(x)\Big| \cdot \Big|f(x) - f(x_0)\Big|.
\end{aligned}$$

Nach Satz 13.1. gibt es eine Umgebung $V(x_0)$ und ein $C \in \mathbb{R}$ mit $|f(x)| \leq C$, $|g(x)| \leq C$ für $x \in V(x_o)$. Mit der obigen Abschätzung folgt

$$\left|\frac{f(x)}{g(x)} - \frac{f(x_0)}{g(x_0)}\right| \leq \frac{C}{\beta^2}\left(\Big|g(x_0) - g(x)\Big| + \Big|f(x) - f(x_0)\Big|\right)$$

für $x \in U(x_0) \cap V(x_0) =: W$.
Wegen der Stetigkeit von f und g gibt es zu vorgegebenem $\varepsilon > 0$ eine Umgebung T von x_0 mit

$$\begin{aligned}
\Big|g(x_0) - g(x)\Big| &< \frac{\beta^2\varepsilon}{2C} \quad \text{und} \\
\Big|f(x_0) - f(x)\Big| &< \frac{\beta^2\varepsilon}{2C} \quad \text{für} \quad x \in T.
\end{aligned}$$

Daraus folgt

$$\left| \frac{f(x)}{g(x)} - \frac{f(x_0)}{g(x_0)} \right| < \varepsilon \quad \text{für} \quad x \in T \cap W,$$

d.h. zu jedem vorgegebenen $\varepsilon > 0$ gibt es eine Umgebung von x_0 mit

$$\left| \frac{f(x)}{g(x)} - \frac{f(x_0)}{g(x_0)} \right| < \varepsilon, \quad \text{d.h. } \tfrac{f}{g} \text{ ist in } U(x_0) \text{ stetig.}$$

□

Der folgende Satz besagt, daß die „Hintereinanderausführung“ stetiger Funktionen stetig ist:

Satz 13.3.
Es seien $k, m, n \in \mathbb{N}$, $M \subset \mathbb{R}^n$, $g\colon M \to \mathbb{R}^m$ stetig und $g(M) \subset N \subset \mathbb{R}^m$, sowie $f\colon N \to \mathbb{R}^k$ stetig. Dann ist die zusammengesetzte Funktion $f \circ g\colon M \to \mathbb{R}^k$, die definiert ist durch $(f \circ g)(x) = f(g(x))$, $x \in M$, ebenfalls stetig.

Der Beweis geschieht wie im eindimensionalen Fall mit Hilfe des Folgenkriteriums:

Satz 13.4. (Folgenkriterium)
Eine Funktion $f\colon M \to \mathbb{R}^m$, $M \subset \mathbb{R}^n$, ist genau dann im Punkt $x_0 \in M$ stetig, wenn für jede Folge $(x_i)_{i\in\mathbb{N}}$, $x_i \in \mathbb{R}^n$, mit $x_i \to x_0$ $(i \to \infty)$ die Konvergenz $f(x_i) \to f(x_0)$ stattfindet.

Der Beweis verläuft fast wörtlich wie im Fall einer Variablen.

Beispiele stetiger Funktionen:
Man überlegt sich leicht, daß lineare Funktionen stetig sind. Eine lineare Funktion $l : \mathbb{R}^n \to \mathbb{R}^m$ ist durch

$$l(x) = \mathbf{A}x = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix}$$

mit einer reellen $m \times n$-Matrix $\mathbf{A}$ erklärt.
Aufgrund von Satz 13.2. sind Produkte von linearen Funktionen und Summen hiervon stetig auf $\mathbb{R}^n$. Dies führt zu dem Satz, daß Polynome in n Variablen (oder „Unbestimmten“) stetig sind. Polynome – erst recht rationale Funktionen $\frac{p(x)}{q(x)}$ – sind nur sinnvoll einzuführen als reell- (oder komplex-) wertige Abbildungen $p : \mathbb{R}^m \to \mathbb{R}$.
Eine kompakte Schreibweise für Polynome mit n Veränderlichen arbeitet mit Multi-Indizes: Es sei $\alpha = (\alpha_1, \dots, \alpha_n)$ mit $\alpha_i \in \{0\} \cup \mathbb{N}$. Ausnahmsweise sei $|\alpha| = \sum_{i=1}^{n} \alpha_i$. Für $x = (\xi_1, \dots, \xi_n) \in \mathbb{R}^n$

sei

$$x^\alpha = \prod_{i=1}^{n} {\xi_i}^{\alpha_i}.$$

Ein Polynom p m-ten Grades in n Veränderlichen läßt sich darstellen durch

$$p(x) = \sum_{|\alpha|\le m} a_\alpha x^\alpha$$

mit Koeffizienten $a_\alpha \in \mathbb{R}$, die durch Multi-Indizes indiziert sind. $\sum\limits_{|\alpha|\le m}$ bedeutet die Summation über alle Multi-Indizes α mit $|\alpha| \le m$.
Beispiel: $p(x) = \xi_1\xi_2 + 1$

$$\begin{array}{rccccccccccc} p(x) = & 0\cdot{\xi_1}^2 & + & 0\cdot{\xi_2}^2 & + & 1\cdot\xi_1\xi_2 & + & 0\cdot\xi_1 & + & 0\cdot\xi_2 & + & 1 \\ & \uparrow & & \uparrow & & \uparrow & & \uparrow & & \uparrow & & \uparrow \\ & a_{(2,0)} & & a_{(0,2)} & & a_{(1,1)} & & a_{(1,0)} & & a_{(0,1)} & & a_{(0,0)} \end{array}$$

Sind p und q Polynome in n Veränderlichen, so läßt sich eine rationale Funktion $\frac{p}{q}$ definieren durch

$$\frac{p}{q}(x) := \frac{p(x)}{q(x)} \quad \text{für} \quad x \in \mathbb{R}^n \setminus \Big\{y \Big| q(y) = 0\Big\}.$$

Ähnlich wie im eindimensionalen Fall kennt man auch hier den Begriff der *gleichmäßigen Stetigkeit*:

Definition:
Seien $m, n \in \mathbb{N}$ und $M \subset \mathbb{R}^n$. Eine Funktion $f\colon M \to \mathbb{R}^n$ heißt gleichmäßig stetig *auf M, wenn zu jedem $\varepsilon > 0$ ein $\delta > 0$ existiert, so daß für alle $x, y \in M$ mit $|x - y| < \delta$ die Ungleichung $|f(x) - f(y)| < \varepsilon$ stattfindet.*

Es gilt

Satz 13.5.
Es sei M eine kompakte Teilmenge des $\mathbb{R}^n$, $n \in \mathbb{N}$, und $f\colon M \to \mathbb{R}^m$, $m \in \mathbb{N}$, stetig. Dann ist f auf M gleichmäßig stetig.
Beweis:
Man kann dies mit Hilfe des Satzes von Heine-Borel zeigen. Wir schlagen einen anderen Weg ein. – Angenommen, die Aussage wäre falsch. Dann gibt es ein $\varepsilon > 0$, so daß zu jedem $\delta > 0$ ein Pärchen $x_\delta, y_\delta \in M$ existiert mit den Eigenschaften

$$\Big|x_\delta - y_\delta\Big| < \delta, \qquad \Big|f(x_\delta) - f(y_\delta)\Big| \ge \varepsilon.$$

Wähle $\delta = \frac{1}{n}$, $n \in \mathbb{N}$. Aus dem Satz von Bolzano-Weierstraß folgt die Existenz eines $x \in M$ und einer Teilfolge $(\varphi(i))_{i \in \mathbb{N}}$, so daß

$$x_{\frac{1}{\varphi(i)}} \to x \quad (i \to \infty).$$

Außerdem gilt mit $\delta = \frac{1}{\varphi(i)}$

$$\left| x_{\frac{1}{\varphi(i)}} - y_{\frac{1}{\varphi(i)}} \right| < \frac{1}{\varphi(i)} \to 0 \quad (i \to \infty),$$

d.h. es gilt auch

$$y_{\frac{1}{\varphi(i)}} \to x.$$

Daraus folgt

$$\left| f\left(x_{\frac{1}{\varphi(i)}}\right) - f\left(y_{\frac{1}{\varphi(i)}}\right) \right| \to \left| f(x) - f(x) \right| = 0 \quad (i \to \infty).$$

Andererseits muß

$$\left| f\left(x_{\frac{1}{\varphi(i)}}\right) - f\left(y_{\frac{1}{\varphi(i)}}\right) \right| \geq \varepsilon$$

gelten. Dies ist ein Widerspruch, und der Satz ist bewiesen. □

Eine weitere Eigenschaft stetiger Funktionen ist

Satz 13.6.
Es sei $n \in \mathbb{N}$ und $M \subset \mathbb{R}^n$ kompakt sowie $f\colon M \to \mathbb{R}$ stetig. Dann nimmt die Funktion f ihr Maximum und ihr Minimum auf M an, d.h. es existieren $y, z \in M$ so daß für alle $x \in M$ die Ungleichungen $f(y) \leq f(x) \leq f(z)$ gelten.
Man schreibt auch

$$\begin{aligned} f(y) &= \min_M f, \\ f(z) &= \max_M f. \end{aligned}$$

Der Beweis geschieht wie im eindimensionalen Fall unter Benutzung des Satzes von Bolzano-Weierstraß.
Wie im Fall reeller Funktionen einer Variablen kann man punktweise konvergente Folgen von stetigen Funktionen betrachten und sich fragen, unter welchen Bedingungen die Grenzfunktion stetig ist:

Definition:
Eine Folge $(f_i)_{i \in \mathbb{N}}$ von Funktionen $f_i\colon M \to \mathbb{R}^m$, $M \subset \mathbb{R}^n$, $m, n \in \mathbb{N}$ heißt punktweise

konvergent *gegen die Funktion* $f\colon M \to \mathbb{R}^m$, *in Zeichen* $f_i \to f \quad (i \to \infty)$, *wenn für jedes* $x \in M$ $f_i(x) \to f(x) \quad (i \to \infty)$.
Die Folge $(f_i)_{i\in\mathbb{N}}$ *heißt (auf* M*)* gleichmäßig konvergent *gegen* f, *wenn zu jedem* $\varepsilon > 0$ *ein* $i_0 \in \mathbb{N}$ *existiert, so daß* $|f_i(x) - f(x)| < \varepsilon$ *für alle* $i \in \mathbb{N},\ i \geq i_0$ *und alle* $x \in M$.

Satz 13.7.
Es seien $m, n \in \mathbb{N},\ M \subset \mathbb{R}^n$ *und* $f_i\colon M \to \mathbb{R}^n,\ i \in \mathbb{N}$ *stetig. Die Folge* $(f_i)_{i\in\mathbb{N}}$ *konvergiere auf* M *gleichmäßig gegen* $f\colon M \to \mathbb{R}^n$. *Dann ist* f *auf* M *stetig.*

Der Beweis geschieht wie im eindimensionalen Fall. □

2.14 Partielle Ableitungen

Es sei M eine offene Teilmenge des $\mathbb{R}^n$ und $f\colon M \to \mathbb{R}$ mit $n \in \mathbb{N}$. Ist $\vec{x}_0 = \left(\xi_1^{(0)}, \ldots, \xi_n^{(0)}\right) \in M$, so gibt es eine ε-Umgebung $U_\varepsilon(\vec{x}_0)$ von $\vec{x}_0$ mit $U_\varepsilon(\vec{x}_0) \subset M$. Die Funktionen $g_i, i = 1, \ldots, n$, die definiert sind durch

$$\begin{aligned}
g_1(\xi) &= f(\xi, \xi_2^{(0)}, \xi_3^{(0)}, \ldots, \xi_{n-1}^{(0)}, \xi_n^{(0)}) \\
g_2(\xi) &= f(\xi_1^{(0)}, \xi, \xi_3^{(0)}, \ldots, \xi_{n-1}^{(0)}, \xi_n^{(0)}) \\
&\ \vdots \\
g_n(\xi) &= f(\xi_1^{(0)}, \xi_2^{(0)}, \xi_3^{(0)}, \ldots, \xi_{n-1}^{(0)}, \xi)
\end{aligned}$$

sind dann zumindest auf ε-Umgebungen $V_\varepsilon(\xi_i^{(0)}) \subset \mathbb{R}$ von $\xi_i^{(0)}$ erklärt:

$$g_i\colon V_\varepsilon(\xi_i^{(0)}) \to \mathbb{R} \qquad , i = 1, \ldots, n.$$

Es kann vorkommen, daß die Funktionen g_i an der Stelle $\xi_i^{(0)}$ differenzierbar sind.

In diesem Fall sagen wir, daß die Funktion f *an der Stelle* $\vec{x}_0$ *eine partielle Ableitung erster Ordnung in der* i*-ten Komponente besitzt* oder *einmal partiell differenzierbar nach der* i*-ten Komponente ist.*

Wir schreiben

$$g_i'(\xi_i^{(0)}) = \frac{\partial f}{\partial \xi_i}(\vec{x}_0)$$

oder, ohne Rücksicht auf die Nennung der Variablen,

$$g_i'(\xi_i^{(0)}) = \frac{\partial f}{\partial x_i}(\vec{x}_0)$$

oder

$$g_i'(\xi_i^{(0)}) = f_{x_i}(\vec{x}_0).$$

Hat man die partielle Differenzierbarkeit nach allen Komponenten, so setzt man

$$\nabla f(\vec{x}_0) := \frac{\partial f(\vec{x}_0)}{\partial \vec{x}} = \left(\frac{\partial f}{\partial x_1}(\vec{x}_0), \dots, \frac{\partial f}{\partial x_n}(\vec{x}_0) \right),$$

$\nabla f(\vec{x}_0) =$ „Nabla f an der Stelle $\vec{x}_0$“ oder „Gradient von f an der Stelle $\vec{x}_0$“ $= \mathrm{grad} f(\vec{x}_0)$. Wenn $f = (f_1, \dots, f_m)^T$ eine Vektorfunktion vom M nach $\mathbb{R}^m$ ist und die Funktionen $f_i \colon M \to \mathbb{R}$ an der Stelle $\vec{x}_0 \in M$ einmal partiell differenzierbar sind, so bezeichnet man die Matrix

$$\frac{\partial f}{\partial \vec{x}}(x_0) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(\vec{x}_0) & \cdots & \frac{\partial f_1}{\partial x_n}(\vec{x}_0) \\ \vdots & & \vdots \\ \frac{\partial f_m}{\partial x_1}(\vec{x}_0) & \cdots & \frac{\partial f_1}{\partial x_m}(\vec{x}_0) \end{pmatrix}$$

als *Funktionalmatrix* oder *Jacobische (Matrix)* von f.

Ist $f \colon M \to \mathbb{R}^n$ partiell differenzierbar, so ist die Funktionalmatrix $\frac{\partial f}{\partial \vec{x}}(\vec{x}_0)$ quadratisch und man kann die Determinante von $\frac{\partial f}{\partial \vec{x}_0}(\vec{x}_0)$, in Zeichen $\det \frac{\partial f}{\partial \vec{x}_0}(\vec{x}_0)$, die sogenannte *Funktionaldeterminate* von f an der Stelle $\vec{x}_0$ bilden.

Schließlich erwähnen wir noch, daß man

$$\sum_{i=1}^{n} \frac{\partial f_i}{\partial x_i}(x_0)$$

als *Divergenz* von f an der Stelle $\vec{x}_0$, in Zeichen $\mathrm{div} f(\vec{x}_0)$, bezeichnet.

Existieren die ersten Ableitungen von f in allen Punkten $\vec{x}_0$ von M, so sagt man, daß f in M einmal partiell differenzierbar ist. Ist $f \colon M \to \mathbb{R}$, so definieren die partiellen Ableitungen von f dann eine Abbildung $\nabla f \colon M \to \mathbb{R}^n$. Ist ∇f partiell differenzierbar, so kann man partielle Ableitung zweiter Ordnung

$$\frac{\partial^2 f}{\partial x_i \partial x_k}$$

und entsprechend höherer Ordnung bilden.

Beispiele:

1. $f(\vec{x}) = \sin \xi_1 \xi_2, \quad \vec{x} = (\xi_1, \xi_2), \quad f \colon \mathbb{R}^2 \to \mathbb{R},$
   $\nabla f(\vec{x}) = (\xi_2 \cos(\xi_1 \xi_2), \xi_1 \cos(\xi_1 \xi_2))$

2. $f(\vec{x}) = \sum\limits_{i=1}^{n} a_i \xi_i, \quad \vec{x} = (\xi_1, \dots, \xi_n)^T,$
   $\nabla f(\vec{x}) = (a_1, \dots, a_n).$

3. $f(\vec{x}) = A\vec{x} + \vec{b} = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix} + \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}$,
   $f\colon \mathbb{R}^n \to \mathbb{R}^m, \quad \vec{x} = (\xi_1, \dots, \xi_n)^T$,
   $\frac{\partial f}{\partial \vec{x}} = A$.

Für die partiellen Ableitungen einer Funktion gelten z.T. ähnliche Sätze wie für die gewöhnliche Ableitung einer Funktion einer Variablen.

Satz 14.1.
Die Funktionen $f, g\colon M \to \mathbb{R}^m$, mit $M \subset \mathbb{R}^n$ offen, seien an der Stelle $\vec{x}_0 \in M$ einmal partiell differenzierbar. Dann sind auch $f + g$, αf mit $\alpha \in \mathbb{R}$ und $f \cdot g$(Skalarprodukt) an der Stelle $\vec{x}_0$ einmal partiell differenzierbar, und es gilt:

$$\begin{aligned} \frac{\partial(f+g)}{\partial x_i}(\vec{x}_0) &= \frac{\partial f}{\partial x_i}(\vec{x}_0) + \frac{\partial g}{\partial x_i}(\vec{x}_0), \\ \frac{\partial(\alpha f)}{\partial x_i}(\vec{x}_0) &= \alpha \frac{\partial f}{\partial x_i}(\vec{x}_0), \\ \frac{\partial(f \cdot g)}{\partial x_i}(\vec{x}_0) &= \left(\frac{\partial f}{\partial x_i} \cdot g\right)(\vec{x}_0) + \left(f \cdot \frac{\partial g}{\partial x_i}\right)(\vec{x}_0). \end{aligned}$$

Beweis:
Der Beweis wird auf den entsprechenden eindimensionalen Fall zurückgeführt.

Interessanterweise folgt aus der partiellen Differenzierbarkeit einer Funktion nicht deren Stetigkeit.

Beispiel:
Sei

$$f(\xi_1, \xi_2) = \begin{cases} \frac{\xi_1 \xi_2}{\xi_1^2 + \xi_2^2} & \text{für } \xi_1^2 + \xi_2^2 \neq 0, \\ 0 & \text{für } \xi_1 = \xi_2 = 0. \end{cases}$$

Dann besitzt f überall auf $\mathbb{R}^2$ partielle Ableitungen, ist aber im Nullpunkt unstetig, denn es gilt

$$\left(\frac{1}{j}, \frac{1}{j}\right) \to 0 \qquad (j \to \infty),$$

aber $f\left(\frac{1}{j}, \frac{1}{j}\right) = \frac{1}{2} \neq 0 = f(0,0)$.

Setzt man jedoch voraus, daß die ersten partiellen Ableitungen von $f\colon M \to \mathbb{R}$, $M \subset \mathbb{R}^n$ offen, in M existieren und ∇f auf M beschränkt ist, so folgert man leicht aus dem Mittelwertsatz die Lipschitzstetigkeit von f.

Satz 14.2. (Kettenregel)
Es sei $M \subset \mathbb{R}^n$ offen und $g\colon M \to \mathbb{R}^m$ stetig und einmal partiell differenzierbar. Ferner sei N eine offene Teilmenge des $\mathbb{R}^m$ mit $N \supset g(M)$ und $f\colon N \to \mathbb{R}$ besitze in $\mathbb{N}$ stetige erste partielle Ableitungen. Dann ist die durch $(f \circ g)(\vec{x}) = f(g(\vec{x}))$ definierte Funktion $(f \circ g) : M \to \mathbb{R}$ einmal partiell differenzierbar, und es gilt:

$$(\nabla(f \circ g))(\vec{x}) = ((\nabla f)(g(\vec{x})))\frac{\partial g}{\partial \vec{x}}(\vec{x})$$

Die letzte Gleichung ist so zu verstehen: Der Zeilenvektor $(\nabla(f \circ g))(\vec{x})$ ist gleich dem Matrizenprodukt aus dem Zeilenvektor $(\nabla f)(g(\vec{x}))$ mit der Matrix

$$\begin{pmatrix} \frac{\partial g_1}{x_1}(\vec{x}) & \cdots & \frac{\partial g_1}{\partial x_n}(\vec{x}) \\ \vdots & & \vdots \\ \frac{\partial g_m}{x_1}(\vec{x}) & \cdots & \frac{\partial g_m}{\partial x_n}(\vec{x}) \end{pmatrix}.$$

Beweis:
Es sei e_i der i-te Einheitsvektor, d.h. der Vektor des $\mathbb{R}^n$, der in der i-ten Komponente eine 1 stehen hat und dessen restliche Komponenten verschwinden. Wir müssen zeigen, daß:

$$\frac{f(g(\vec{x} + h \cdot e_i)) - f(g(\vec{x}))}{h} \to \sum_{k=1}^{m} \left(\left(\frac{\partial g_k}{\partial x_i} \right)(\vec{x}) \right) \left(\frac{\partial f}{\partial x_k} \right)(g(\vec{x}))$$

konvergiert für $h \to 0, \quad h \neq 0$. Es folgt dann

$$\left(\tfrac{f(g(\vec{x}+h\cdot e_1))-f(g(\vec{x}))}{h}, \cdots, \tfrac{f(g(\vec{x}+h\cdot e_n))-f(g(\vec{x}))}{h} \right) \to \nabla(f \circ g)(\vec{x}) = ((\nabla f)(g(x))\tfrac{\partial g}{\partial \vec{x}}(\vec{x}).$$

Es gilt

$$\begin{aligned} &\frac{f(g(\vec{x} + he_i)) - f(g(\vec{x}))}{h} = \\ &= \frac{f(g_1(\vec{x} + he_i), \cdots, g_m(\vec{x} + he_i)) - f(g_1(\vec{x} + he_i), \cdots, g_{m-1}(\vec{x} + he_i), g_m(\vec{x}))}{h} + \\ &+ \frac{f(g_1(\vec{x} + he_i), \cdots, g_{m-1}(\vec{x} + he_i), g_m(\vec{x})) - f(g_1(\vec{x} + he_i), \cdots, g_{m-1}(\vec{x}), g_m(\vec{x}))}{h} + \\ &\vdots \\ &+ \frac{f(g_1(\vec{x} + he_i), g_2(\vec{x}), \cdots, g_m(\vec{x})) - f(g_1(\vec{x}), \cdots, g_m(\vec{x}))}{h}. \end{aligned}$$

Wir wenden auf die einzelnen Brüche den Mittelwertsatz an und erhalten

$$\frac{f(g(\vec{x} + he_i)) - f(g(\vec{x}))}{h} = \sum_{k=1}^{n} \left(\frac{\partial f}{\partial x_k} g_*^{(k)} \right) \frac{g_k(\vec{x} + he_j) - g_k(\vec{x})}{h}.$$

Die Vektoren $g_*^{(k)}$ hängen von $\vec{x}$ und h ab, ihre j-ten Komponenten liegen zwischen $g_j(\vec{x}+he_i)$ und $g_j(\vec{x})$, z.B. ist

$$g_*^{(m)} = (g_1(\vec{x}+he_i), \dots, g_{m-1}(\vec{x}+he_i), \gamma_*)$$

mit einer Zahl γ_* zwischen $g_m(\vec{x}+he_i)$ und $g_m(\vec{x})$.
Wegen der Stetigkeit von g konvergiert $g_*^{(k)} \to g(\vec{x})$ für $h \to 0$. Wegen der Stetigkeit von $\frac{\partial f}{\partial x_k}$ konvergiert daher

$$\frac{\partial f}{\partial x_k}(g_*^{(k)}) \to \frac{\partial f}{\partial x_k}(g(\vec{x})) \qquad (h \to 0)$$

und daher

$$\frac{f(g(\vec{x}+he_i)) - f(g(\vec{x}))}{h} \to \sum_{k=1}^{n} \left(\left(\frac{\partial f}{\partial x_k} \right) (g(\vec{x})) \right) \left(\frac{\partial g_k}{\partial x_i} \right) (\vec{x}).$$

□

Ist f in Satz 14.2 eine Abbildung von N nach $\mathbb{R}^s$, so kann man die Kettenregel folgendermaßen formulieren:

Satz 14.2'. *(Vorraussetzungen analog zu Satz 14.2.)*

$$\frac{\partial(f \circ g)}{\partial x}(x) = \left(\frac{\partial f}{\partial x}\right)(g(x))\left(\frac{\partial g}{\partial x}(x)\right),$$

d.h die Funktionalmatrix von $f \circ g$ an der Stelle x, also die Matrix

$$\begin{pmatrix} \frac{\partial(f_1 \circ g)}{\partial x_1}(x) & \cdots & \frac{\partial(f_1 \circ g)}{\partial x_n}(x) \\ \vdots & & \vdots \\ \frac{\partial(f_s \circ g)}{\partial x_1}(x) & \cdots & \frac{\partial(f_s \circ g)}{\partial x_n}(x) \end{pmatrix}$$

ist gleich dem Matrixenprodukt der beiden Matrizen

$$\begin{pmatrix} \frac{\partial f_1}{\partial x_1}(g(x)) & \cdots & \frac{\partial f_1}{\partial x_m}(g(x)) \\ \vdots & & \vdots \\ \frac{\partial f_s}{\partial x_1}(g(x)) & \cdots & \frac{\partial f_s}{\partial x_m}(g(x)) \end{pmatrix}$$

und

$$\begin{pmatrix} \frac{\partial g_1}{\partial x_1}(x) & \cdots & \frac{\partial g_1}{\partial x_n}(x) \\ \vdots & & \vdots \\ \frac{\partial g_m}{\partial x_1}(x) & \cdots & \frac{\partial g_m}{\partial x_n}(x) \end{pmatrix}.$$

Wir erwähnen noch eine Anwendung der Kettenregel, die man manchmal benötigt:

Satz:
Es sei $M \subset \mathbb{R}^n$ offen und $f\colon M \to \mathbb{R}$ einmal stetig partiell differenzierbar. Ist dann $\vec{x} \in M$ und $\vec{h} \in \mathbb{R}^n$, so ist durch $g(t) := f(\vec{x} + t\vec{h})$ eine Funktion $g\colon \mathbb{R} \to \mathbb{R}$ für alle t aus einer Umgebung der $0 \in \mathbb{R}$ definiert und nach dem letzten Satz bei $t = 0$ differenzierbar, und es gilt

$$\frac{dg}{dt}(0) = \vec{h} \cdot \nabla f(\vec{x}).$$

Man bezeichnet $\vec{h} \cdot \nabla f(\vec{x})$ auch als *Ableitung von f in Richtung von h.* („Richtungsableitung“ – allerdings verwendet man diesen Namen noch in einem allgemeineren Zusammenhang).

Eine weitere wichtige Eigenschaft partieller Ableitungen ist ihre Vertauschbarkeit.

Beispiel:

$$\begin{aligned}\frac{\partial}{\partial\xi_1}\left(\frac{\partial}{\partial\xi_2}\sin(\xi_1\xi_2)\right) &= \frac{\partial}{\partial\xi_1}\left(\xi_1\cos(\xi_1\xi_2)\right)\\ &= \cos(\xi_1\xi_2)-\xi_2\xi_1\sin(\xi_1\xi_2)\\ &= \frac{\partial}{\partial\xi_2}\left(\xi_2\cos(\xi_1\xi_2)\right)\\ &= \frac{\partial}{\partial\xi_2}\left(\frac{\partial}{\partial\xi_1}\sin(\xi_1\xi_2)\right).\end{aligned}$$

Satz 14.3.
Es sei $M\subset\mathbb{R}^n$ offen und die Funktion $f\colon M\to\mathbb{R}$ möge in M erste und zweite stetige partielle Ableitungen haben. Dann gilt in M die Gleichheit

$$\frac{\partial^2 f}{\partial x_i\partial x_k}=\frac{\partial^2 f}{\partial x_k\partial x_i}.$$

Entsprechend kann man die Vertauschbarkeit höherer Ableitungen behaupten:

$$\left(\prod_{i=1}^N\frac{\partial^{\alpha(i)}}{\partial x_{\beta(i)}}\right)f=\left(\prod_{i=1}^N\frac{\partial^{\alpha(\pi(i))}}{\partial x_{\beta(\pi(i))}}\right)f.$$

Hierbei ist π eine Permutation der Zahlen von 1 bis N, $\alpha(i)\in\mathbb{N}$, $\beta(i)\in\{1,\dots,n\}$.

Beweis:
Es sei $x\in M$ und e_j, $j=1,\dots,n$ die Einheitsvektoren im $\mathbb{R}^n$. Da M offen ist, gibt es ein $h_0>0$, so daß für alle $h\in\mathbb{R}$ mit $|h|\le h_0$ die Punkte $x+he_i$ und $x+he_i+he_j$, $i,j=1,\dots n$, in M liegen. Für $|h|\le h_0$ sind dann die ersten und zweiten Differenzenquotienten von f an der Stelle x erklärt:

$$w(x)=D_j^h f(x)=\frac{f(x+he_j)-f(x)}{h},\quad D_k^hD_j^hf(x)=D_k^hw(x).$$

Man rechnet leicht nach, daß

$$D_j^hD_k^hf(x)=D_k^hD_j^hf(x),$$

und zum Nachweis des Satzes genügt es, zu zeigen, daß für $j,k=1,\dots,n$

$$D_j^hD_k^hf(x)\to\frac{\partial}{\partial x_j}\frac{\partial}{\partial x_k}f(x)\qquad(h\to 0).$$

Dies geschieht folgendermaßen:
Es gelten die Gleichungen

$$\begin{aligned}D_k^hf(x) &= \frac{1}{h}\int_0^1\frac{\partial f}{\partial x_k}(x+te_k)\,dt,\\ D_k^hf(x+he_j) &= \frac{1}{h}\int_0^1(\frac{\partial f}{\partial x_k}(x+he_j+te_k)\,dt.\end{aligned}$$

Daraus folgt

$$\begin{aligned} D_j^h D_k^h f(x) &= \frac{1}{h^2}\int\limits_0^1 \left(\frac{\partial f}{\partial x_k}(x+he_j+te_k)-\frac{\partial f}{\partial x_k}(x+te_k)\right)dt \\ &= \frac{1}{h}\int\limits_0^1 \left(\frac{\frac{\partial f}{\partial x_k}(x+he_j+te_k)-\frac{\partial f}{\partial x_k}(x+te_k)}{h}-\frac{\partial^2 f(x)}{\partial x_j \partial x_k}\right)dt \\ &\quad +\frac{\partial^2 f(x)}{\partial x_j \partial x_k}. \end{aligned}$$

Aus dem Mittelwertsatz folgt

$$\frac{1}{h}\left(\frac{\partial f}{\partial x_k}(x+he_j+te_k)-\frac{\partial f}{\partial x_k}(x+te_k)\right)=\frac{\partial^2}{\partial x_j \partial x_k}f(x+\theta h e_j+te_k)$$

mit $0<\theta<1$. Aus der gleichmäßigen Stetigkeit von $\frac{\partial^2 f}{\partial x_j \partial x_k}$ in einer Umgebung von x folgt für $0<|h|<\delta$, daß der Integrand im letzten Integral zum Betrag noch kleiner als $\varepsilon>0$ sein muß. Daraus folgt

$$D_j^h D_k^h f(x)\to\frac{\partial^2 f(x)}{\partial x_j \partial x_k}\qquad (h\to 0),$$

und der Satz ist bewiesen. □

2.15 Mittelwertsatz und Taylorscher Satz. Totale Differenzierbarkeit

Satz 15.1. (Mittelwertsatz)
Es sei $M\subset\mathbb{R}^n$ offen und $f\colon M\to\mathbb{R}$ einmal stetig partiell differenzierbar. Es seien $\vec{x},\vec{y}\in M$ und die $\vec{x}$ und $\vec{y}$ verbindende Strecke $\{t\vec{x}+(1-t)\vec{y}|0\le t\le 1\}$ gehöre zu M. Dann gibt es eine Zahl $\vartheta\in]0,1[$ mit

$$f(\vec{x})=f(\vec{y})+(\vec{x}-\vec{y})\cdot\nabla f\left(\vec{y}+\vartheta(\vec{x}-\vec{y})\right).$$

Beweis:
Die durch $g(t)=f(\vec{y}+t(\vec{x}-\vec{y}))$ definierte Funktion $g\colon[0,1]\to\mathbb{R}$ ist stetig und in $]0,1[$ differenzierbar. Nach dem „eindimensionalen“ Mittelwertsatz ist $g(1)=g(0)+g'(\vartheta)$ mit einem $\vartheta\in]0,1[$. Daraus folgt mit der Kettenregel

$$f(\vec{x})=f(\vec{y})+(\vec{x}-\vec{y})\cdot\nabla f\left(\vec{y}+\vartheta(\vec{x}-\vec{y})\right).$$

□

Folgerung:
Die Punktmenge $M\subset\mathbb{R}^n$ sei offen und habe die Eigenschaft, daß je zwei Punkte $\vec{x}$ und $\vec{y}$

aus M durch einen endlichen Polygonzug verbunden werden können, d.h. es existieren Punkte $\vec{x}_i \in M\,,\ i = 1,\dots,N$, mit $\vec{x}_1 = \vec{x}, \vec{x}_N = y$, so daß die Strecken $\{t\vec{x}_i + (1-t)\vec{x}_{i+1} | t \in [0,1]\}$ für $i = 1,\dots,N-1$ alle in M liegen. (Man sagt dann, M sei zusammenhängend). Die ersten partiellen Ableitungen der Funktion $f\colon M \to \mathbb{R}$ mögen verschwinden. Dann ist f konstant auf M.

Der Mittelwertsatz für Funktionen in n Veränderlichen ist die einfachste Form des entsprechenden Taylorschen Satzes. Hierzu bedienen wir uns der folgenden Schreibweise:

Es sei $\vec{h} \in \mathbb{R}^n$ und $f\colon M \to \mathbb{R}$, M offen, $\vec{x} \in M\,,\ f$ sei n-mal partiell differenzierbar. Wir setzen

$$(\vec{h} \cdot \nabla) f(x) = \vec{h} \cdot \nabla f(x)$$

und bei entsprechenden Differenzierbarkeitsvoraussetzungen

$$(\vec{h} \cdot \nabla)^{k+1} f(x) = (\vec{h} \cdot \nabla)((h \cdot \nabla)^k f)(x), \quad k+1 \leq N.$$

Es ist dann formal (siehe die Multi-Index-Schreibweise bei Polynomen in n Veränderlichen)

$$(\vec{h} \cdot \nabla)^k = \sum_{|\mu|=k} \vec{h}^\mu \nabla^\mu$$

mit

$$\nabla^\mu = \prod_{i=1}^{n} \left(\frac{\partial}{\partial x_i}\right)^{\mu_i}, \qquad \mu = (\mu_1, \dots \mu_n).$$

Setzt man $g(t) = f(\vec{x} + t\vec{h})$, so gilt

$$\frac{d^k g}{dt^k}(t) = (\vec{h} \cdot \nabla)^k f(\vec{x} + t\vec{h}).$$

Aus dem Taylorschen Satz, angewandt auf g, folgt

Satz 15.2. (Taylorsche Formel)
Es sei $M \subset \mathbb{R}^n$ offen, $\vec{x} \in M$ und $\vec{h} \in \mathbb{R}^n$ sowie $\{\vec{x} + t\vec{h} | t \in [0,1]\} \subset M$. Die Funktion $f\colon M \to \mathbb{R}$ sei $(N+1)$-mal in M stetig partiell differenzierbar. Dann gilt

$$f(\vec{x} + \vec{h}) = \sum_{t=0}^{N} \frac{1}{k!} (\vec{h} \cdot \nabla)^k f(\vec{x}) + \frac{1}{(N+1)!} \left(\vec{h} \cdot\ \nabla\right)^{N+1} f(\vec{x} + \vartheta \vec{h})$$

mit einer Zahl $\vartheta \in]0,1[$.

Man kann die Taylorsche Formel als Darstellung der Funktion f an der Stelle $\vec{x} + \vec{h}$ durch ein Polynom N-ten Grades in $\vec{h}$ + einem Restglied auffassen.

Für die Approximation der Funktion f an der Stelle $\vec{x} + \vec{h}$ durch eine *lineare Funktion* gilt der

Satz 15.3.
Es sei $M \subset \mathbb{R}^n$ *offen,* $\vec{x}, \vec{x}_0 \in M$, $\{t\vec{x} + (1-t)\vec{x}_0 | t \in [0,1]\} \subset M$ *sowie* $f\colon M \to \mathbb{R}$ *einmal stetig partiell differenzierbar. Dann gilt*

$$\frac{f(\vec{x}) - f(\vec{x}_0) - (\nabla f(x_0)) \cdot (\vec{x} - \vec{x}_0)}{|\vec{x} - \vec{x}_0|} \to 0$$

für $\vec{x} \to \vec{x}_0, \vec{x} \neq \vec{x}_0$.

Ist $f\colon M \to \mathbb{R}^m$ einmal stetig partiell differenzierbar, d.h. die Komponenten f sind stetig partiell differenzierbar, so ist $\nabla f(\vec{x}_0)$ als Funktionalmatrix von f an der Stelle $\vec{x}_0$ aufzufassen.

Beweis:
Nach dem Mittelwertsatz ist für $f\colon M \to \mathbb{R}$

$$\frac{f(\vec{x}) - f(\vec{x}_0) - (\nabla f(x_0)) \cdot (\vec{x} - \vec{x}_0)}{|\vec{x} - \vec{x}_0|} =$$

$$= [\nabla f(\vec{x}_0 + \vartheta(\vec{x} - \vec{x}_0)) - \nabla f(\vec{x}_0)] \frac{(\vec{x} - \vec{x}_0)}{|\vec{x} - \vec{x}_0|}, \quad \vartheta \in]0,1[.$$

Der letzte Ausdruck konvergiert für $\vec{x} \to \vec{x}_0$, $\vec{x} \neq x_0$, gegen Null, da $\left|\frac{(\vec{x} - \vec{x}_0)}{|\vec{x} - \vec{x}_0|}\right| = 1$ und mit $\vec{x}_0 + \vartheta(\vec{x} - \vec{x}_0) \to \vec{x}_0$ nach Voraussetzung auch

$$\nabla f(\vec{x}_0 + \vartheta(\vec{x} - \vec{x}_0)) \to \nabla f(\vec{x}_0)$$

konvergiert. □

Ist f eine *Vektor*funktion, so ist die eben gemachte Schlußweise für jede Komponente einzeln durchzuführen.
Man kann Satz 15.3. auch so formulieren: Unter den dortigen Voraussetzungen ($f\colon M \to \mathbb{R}^m$) gibt es eine lineare Funktion $l\colon \mathbb{R}^n \to R^m$ die sich in der Form $l(\vec{z}) = \mathbf{A}\vec{z}$ mit einer $m \times n$-Matrix $\mathbf{A}$ darstellen läßt, so daß gilt

$$\frac{f(\vec{x}) - f(\vec{x}_0) - l(\vec{x} - \vec{x}_0)}{|\vec{x} - \vec{x}_0|} \to 0 \quad (\vec{x} \to \vec{x}_0, \vec{x} \neq \vec{x}_0). \tag{2.4}$$

In der Tat hat die Matrix $\mathbf{A} = \nabla f(\vec{x}_0)$ diese Eigenschaft. Formel (2.4) kann auch zur Definition der Differenzierbarkeit von f verwendet werden. Man sagt dann, wenn eine Funktion l mit den obigen Eigenschaften existiert, daß die Funktion f *total differenzierbar* sei. Nach Satz 15.3. ist eine Funktion f total differenzierbar, wenn ∇f stetig ist.
Im Anschluß an die Taylorsche Formel erwähnen wir noch *Mehrfachpotenzreihen.* Eine Reihe der Form

$$P(\vec{x}) = \sum_{v_1, \dots, v_n = 0}^{\infty} a_{v_1 v_2 \dots v_n} (\xi_1 - \xi_{01})^{v_1} (\xi_2 - \xi_{02})^{v_2} \cdots (\xi_n - \xi_{0n})^{v_n}$$

mit rellen Zahlkoeffizienten $a_{v_1 v_2 \dots v_n}$ und Variablen $\vec{x} = (\xi_1, \dots, \xi_n)$ heißt eine n-fache (oder Mehrfach-) Potenzreihe in $\vec{x}$ mit Entwicklungspunkt $\vec{x}_0 = (\xi_{01}, \dots, \xi_{0n})$. (Genaugenommen muß $\sum_{v_1,\dots,v_n=0}^{\infty} \beta_{v_1 v_2 \dots v_n}$, $\beta_{v_1 v_2 \dots v_n} \in \mathbb{R}$ als Folge von Partialsummen erklärt werden, wobei die n-Tupel $(v_1, v_2, \dots, v_n)$ in einer geeigneten Reihenfolge durch natürliche Zahlen numeriert werden und die Partialsummen konsistent mit dieser Numerierung gebildet werden.). Setzen wir voraus, daß die obige Reihe absolut konvergiert, so ist ihr Wert unabhängig von der Weise, wie wir die n-Tupel $(v_1, v_2, \dots, v_n)$ bei der Summation durchlaufen. Wie im eindimensionalen Fall gilt der

Satz 15.4.
Wenn P für ein $\vec{x} = (\xi_{11}, \xi_{12}, \dots, \xi_{1n}) \in \mathbb{R}^n$ absolut konvergiert, so konvergiert P auch für alle $\vec{x} = (\xi_1, \xi_2, \dots, \xi_n) \in \mathbb{R}^n$ mit

$$|\xi_i - \xi_{0i}| < |\xi_{1i} - \xi_{0i}|, \quad i = 1, \dots, n.$$

Der offene Kern der Menge aller $\vec{x} \in \mathbb{R}^n$, für die P absolut konvergiert, heißt *Konvergenzgebiet* von P.
Es gilt der

Satz 15.5.
Die durch die Potenzreihe P dargestellte Funktion ist in ihrem Konvergenzgebiet beliebig oft stetig partiell differenzierbar. Es gilt

$$a_{v_1 v_2 \dots v_n} = \frac{1}{v_1! \cdots v_n!} \frac{\partial^{v_1 + \cdots + v_n}}{(\partial x_1)^{v_1} \cdots (\partial x_n)^{v_n}} P(\vec{x}_0).$$

2.16 Extrema bei Funktionen mehrerer Veränderlicher

Satz 16.1.
Es sei $M \subset \mathbb{R}^n$ offen und $f\colon M \to \mathbb{R}$ einmal partiell differenzierbar. An der Stelle $\vec{x}_0 \in M$ befinde sich ein Maximum oder Minimum von f auf M, d.h.

$$f(\vec{x}_0) \geq f(\vec{x}), \qquad \vec{x} \in M$$

bzw.

$$f(\vec{x}_0) \leq f(\vec{x}), \qquad \vec{x} \in M.$$

Dann gilt

$$\nabla f(\vec{x}_0) = 0.$$

Beweis:
Sei $g_i(\xi) = f(\xi_1^{(0)}, \dots, \xi_{i-1}^{(0)}, \xi, \xi_{i+1}^{(0)}, \dots, \xi_n^{(0)})$ mit $\vec{x}_0 = (\xi_1^{(0)}, \dots, \xi_n^{(0)})$.
Für alle ξ aus einer Umgebung $U(\xi_i^{(0)})$ von $\xi_i^{(0)}$ in $\mathbb{R}$ gilt

$$g_i(\xi_i^{(0)}) \geq g_i(\xi)$$

bzw.

$$g_i(\xi_i^{(0)}) \leq g_i(\xi),$$

denn f hat an der Stelle $\vec{x}_0$ ein Extremum und

$$\left(\xi_1^{(0)}, \dots, \xi_{i-1}^{(0)}, \xi, \xi_{i+1}^{(0)}, \dots, \xi_n^{(0)}\right) \in M \text{für } \xi \in U(\xi_i^{(0)})$$

Daraus folgt $g_i'(\xi_i^{(0)}) = 0$ und

$$\frac{\partial f}{\partial x_i}(\vec{x}_0) = 0.$$

$\square$

Beispiel: $f(\vec{x}) = \left(\sum\limits_{i,k=1}^{n} a_{i\,k}\xi_i\xi_k + b_i\xi_i\right) + c.$
f habe an der Stelle $\vec{x}_0 = (\xi_1^{(0)}, \dots, \xi_n^{(0)})$ ein Extremum (d.h. Maximum oder Minimum). Dann gilt wegen

$$\begin{aligned}
\frac{\partial f}{\partial \xi_j}(\vec{x}_0) &= 2a_{j\,j}\xi_j^{(0)} + \sum_{\substack{i=1\\ i\neq j}}^{n} a_{i\,j}\xi_i^{(0)} + \sum_{\substack{k=1\\ k\neq j}}^{n} a_{j\,k}\xi_k^{(0)} + b_j \\
&= (a_{1\,j}, \cdots, a_{n\,j}) \cdot \vec{x}_0 + (a_{j\,1}, \cdots, a_{j\,n}) \cdot \vec{x}_0 + b_j
\end{aligned}$$

und

$$\nabla f(\vec{x}_0) = (A + A^T)\vec{x}_0 + \vec{b}$$

die Beziehung

$$(A + A^T)\vec{x}_0 + \vec{b} = 0,$$

und dies liefert eine notwendige Bedingung für das Extremum.

Eine hinreichende Bedingung für ein Maximum einer partiell differenzierbaren Funktion liefert

Satz 16.2.
Es sei $U \subset \mathbb{R}^n$ eine offene ε-Umgebung des Punktes $\vec{x}_0 \in \mathbb{R}^n$ und $f\colon U \to \mathbb{R}$ zweimal stetig partiell differenzierbar. Die Matrix $\left(\frac{\partial^2 f}{\partial x_i \partial x_k}\right)_{i,k=1}^n$ der zweiten Ableitungen von f sei in U negativ semidefinit, *d.h. für alle $\vec{h} = (h_1, \dots, h_n) \in \mathbb{R}^n$ gelte*

$$\sum_{i,k=1}^n \frac{\partial^2 f}{\partial x_i \partial x_k} h_i h_k \le 0 \quad \textit{in } U.$$

Außerdem gelte im Punkt $\vec{x}_0 \in U$ die Gleichung $\nabla f(\vec{x}_0) = 0$. Dann nimmt f im Punkt $\vec{x}_0$ in U ein Maximum an.

Beweis:
Aus der Taylorschen Formel folgt, falls $\vec{x}_0 + \vec{h} \in U$, daß

$$f(\vec{x}_0 + \vec{h}) = f(\vec{x}_0) + (\vec{h} \cdot \nabla) f(\vec{x}_0) + \frac{1}{2}(\vec{h} \cdot \nabla)^2 f(\vec{x}_0 + \theta \vec{h}), \theta \in]0,1[.$$

Wegen $\nabla f(\vec{x}_0) = 0$ folgt $(\vec{h} \cdot \nabla) f(\vec{x}_0) = 0$ und wegen

$$\frac{1}{2}(\vec{h} \cdot \nabla)^2 f = \frac{1}{2} \sum_{i,k=1}^n \frac{\partial^2 f}{\partial x_i \partial x_k} h_i h_k \le 0$$

folgt

$$f(\vec{x}_0 + \vec{h}) \le f(\vec{x}_0),$$

d.h. f hat an der Stelle $\vec{x}_0$ tatsächlich ein Maximum.

Entsprechend hat f an der Stelle $\vec{x}_0$ ein Minimum, wenn $\nabla f(\vec{x}_0) = 0$ gilt und $\frac{\partial^2 f}{\partial x_i \partial x_k}$ in U positiv semidefinit ist.
Die Eindeutigkeit des Extremums erhält man, wenn man die Voraussetzung der Semidefinitheit durch die der Definitheit ersetzt.

2.16.1 Extrema mit Nebenbedingungen

Wir betrachten Optimierungsprobleme mit Gleichungsnebenbedingungen: Seien $f\colon \mathbb{R}^n \to \mathbb{R}$, $g\colon \mathbb{R}^n \to \mathbb{R}$, $M = \{x | g(x) = 0\}$.
Gesucht ist $x^* \in M$, so daß

$$f(x^*) \le f(x) \quad \forall x \in M.$$

Kürzer schreibt man:

$$f(x) = \min!, \qquad\qquad g(x) = 0. \tag{2.5}$$

(Analog betrachtet man Maximumaufgaben.)

Ähnlich wie im unrestringierten Fall

$$f(x) = \min!, \qquad x \in \mathbb{R}^n,$$

läßt sich eine zu $\nabla f(z) = 0$, z Maximal- bzw. Minimalstelle, notwendige Bedingung angeben. Dies leistet die sogenannte „Theorie der Lagrangemultiplikatoren", die wir zunächst im linearen Fall

$$g(x) = Ax - b, \qquad A \in M(n \times n, \mathbb{R})$$

untersuchen.
Wir schreiben

$$g = (g_1, \ldots, g_m), \qquad g_i(x) = a^{(i)} \cdot x - \beta_i, \qquad i = 1, \ldots, m \qquad a^{(i)} \in \mathbb{R}^n, \beta_i \in \mathbb{R}.$$

Satz 16.3.
Sei $f \in C^1(\mathbb{R}^n)$, $a^{(i)} \in \mathbb{R}^n$, $\beta_i \in \mathbb{R}$ *und* x^* *Lösung von*

$$f(x) = \min!, \qquad a^{(i)} \cdot x - \beta_i = 0, \qquad i = 1, \ldots, m. \tag{2.6}$$

Dann gibt es reelle Zahlen λ_i, *genannt* „Lagrangemultiplikatoren", *so daß*

$$\nabla f(x^*) + \sum_{i=1}^{m} \lambda_i a^{(i)} = 0. \tag{2.7}$$

Anmerkung:
(2.7) ergibt n Gleichungen in $\mathbb{R}$ für die $n + m$ Unbekannten

$$x_i^*, i = 1, \ldots, n, \qquad \lambda_i, i = 1, \ldots, m.$$

Die Nebenbedingung

$$g_i(x^*), i = 1, \ldots, m,$$

liefert die „fehlenden" m Gleichungen.

Beweis von Satz 16.3:
Sei $h \in \mathbb{R}^n$, $h \perp a^{(i)}$.
Dann ist $g_i(x + th) = g_i(x)$ und $x^* + th \in M$, d.h. $x^* + th$ ist ein zulässiges Vergleichselement. Da x^* Minimalstelle ist, gilt $f(x^*) \leq f(x^* + th) =: \varphi(t)$. Mit diesen Bezeichnungen erhalten wir also $\varphi(0) \leq \varphi(t)$, d.h. $\varphi'(0) = 0$, da φ bei 0 ein Minimum besitzt. Nach der Kettenregel gilt dann $h \cdot \nabla f(x^*) = 0 \quad \forall h \perp a^{(i)}$. Nach Sätzen der linearen Algebra folgt schließlich $\nabla f(x^*) \in \operatorname{span}\{a^{(1)}, \ldots, a^{(m)}\}$. $\square$

Im Fall *allgemeiner nichtlinearer Nebenbedingungen* gilt der Satz

Satz 16.4.
Sei $f \in C^1$, $g \in C^1(\mathbb{R}^n, \mathbb{R}^m)$ und x^* (lokale) Lösung von

$$f(x) = \min!, \qquad g(x) = 0.$$

Es gelte die „Regularitätsbedingung“, *nämlich die lineare Unabhängigkeit der Vektoren $\nabla g_i(x^*)$, $i = 1, \dots, m$. Dann existieren $\lambda_i \in \mathbb{R}$ mit*

$$\nabla f(x^*) + \sum_{i=1}^{m} \lambda_i \nabla g_i(x^*) = 0. \tag{2.8}$$

Zum Beweis:
Man kann (2.8) ähnlich wie (2.7) beweisen, indem die Gleichung

$$h \cdot \nabla f(x^*) = 0 \quad \forall h \text{ mit } h \perp g_i(x^*)$$

gezeigt wird. Dieses wiederum geschieht durch Konstruktion zulässiger Vergleichsvektoren

$$x^* + th + o(t) \in M,$$

was jedoch relativ aufwendig ist und den sogenannten Satz über implizite Funktionen verwendet (siehe z.B. [5, Band 3]). In der Optimierung werden wesentlich elegantere Beweise geliefert, die auf einer Konvergenzanalyse der Methode der *Straffunktionale* (Penalty-method) beruhen. In der Methode der Straffunktionale wird das Problem

$$f(x) = \min!, \quad g(x) = 0$$

durch ein unrestringiertes Problem

$$f(x) + \varepsilon^{-1}|g(x)|^2 = \min!, \quad x \in \mathbb{R}^n,$$

approximiert ($\varepsilon \to 0$). („Die Verletzung von $g(x) = 0$ wird durch den großen Fehler ε^{-1} bestraft“). Aus der notwendigen Bedingung für die Lösung x_ε folgt

$$\nabla f(x_\varepsilon) + \sum_{i=1}^{m} \lambda_i^\varepsilon \nabla g_i(x_\varepsilon) = 0, \quad \lambda_i^\varepsilon = 2\varepsilon^{-1} g_i(x_\varepsilon),$$

und man konstruiert sich λ_i als Häufungspunkt der λ^ε.

Wir beweisen hier zur Illustration nur den Fall der Nebenbedingung $g(x) = 1 - |x|^2$.

Beweis für einen Spezialfall:
Der Vektor $\frac{x^*+th}{|x^*+th|}$ erfüllt die Nebenbedingung, denn er hat Betrag 1 ($|t|$ klein, $h \neq 0$). Es gilt wegen der Minimalität von x^*:

$$0 = \frac{d}{dt} f\left(\frac{x^* + th}{|x^* + th|}\right)\Bigg|_{t=0} = [h - x^*(h \cdot x^*)] \cdot \nabla f(x^*),$$

da $\frac{d}{dt}\left(\frac{x^*+th}{|x^*+th|}\right) = \frac{h}{|x^*+th|} - \frac{x^*+th}{|x^*+th|^3}(h \cdot (x^* + th))$

Über lineare Algebra folgt dann

$$\nabla f(x^*) = \lambda x^*.$$

2.16.2 Optimierungsaufgaben mit Ungleichungsnebenbedingungen

Sei $f \in C^1(\mathbb{R}^n, \mathbb{R})$ und $g \in C^1(\mathbb{R}^n, \mathbb{R}^m)$. Wir betrachten die Aufgabe

(P) Gesucht ist $x^* \in \mathbb{R}^n$ mit $g(x^*) \geq 0$ und $f(x^*) \geq f(x)\ \forall x$ mit $g(x) \geq 0$.

($g(x^*) \geq 0$ bedeutet $g_i(x^*) \geq 0,\ i = 1, \ldots m$.)

Wir schreiben kürzer: x^* löst

$$f(x) = \max!\,, \qquad g(x) \geq 0$$

Ähnlich wie im Fall von Gleichungsnebenbedingungen gibt es auch hier eine Multiplikatorentheorie, die „KUHN-TUCKER-THEORIE".

Es gilt der folgende

Satz:
Sei $f \in C^1(\mathbb{R}^n, \mathbb{R})$, $g \in C^1(\mathbb{R}^n, \mathbb{R}^m)$. x^* sei eine Lösung von (P). Es gelte die Regularitätsbedingung: *Die Vektoren $\nabla g_i(x^*)$, $i = 1, \ldots, m$, seien linear unabhängig. Dann gelten die* Kuhn-Tucker-Bedingungen *mit geeigneten Zahlen* („Multiplikatoren") λ_i, $i = 1, \ldots m$.

1. $g(x^*) \geq 0,\ \lambda_i \geq 0,\ i = 1, ..., m$. „*Zulässigkeit*"
2. $\lambda_i g_i(x^*) = 0,\ i = 1, \ldots m$, „*Gleichgewichtssatz*"
3. $\nabla_x L(x^*, \lambda) = 0$ mit $L(x, \lambda) = f(x) + \sum_{i=1}^m \lambda_i g_i(x)$

L heißt hierbei *Lagrange-Funktion.*

Beispiel:

$$-x_1^2 - x_1 x_2 - x_2^2 - x_1 = \max! \qquad x_1 + x_2 \geq 2$$
$$L(x, \lambda) = -x_1^2 - x_1 x_2 - x_2^2 - x_1 + \lambda(x_1 + x_2 - 2)$$

3. Kuhn-Tucker-Bed. $\Longrightarrow$ (wir schreiben x statt x^*)

$$-2x_1 - x_2 - 1 + \lambda = 0$$
$$-2x_2 - x_1 + \lambda = 0$$

Erster Fall: $x_1 + x_2 > 2 \quad \Longrightarrow \quad$ (wg. 2. KUHN-TUCKER-Bed.) $\lambda = 0$.
Die Lösung des Gleichungssystems für x ergibt dann $x_1 = -\frac{2}{3}$, $x_2 = \frac{1}{3}$. Dies verletzt aber Bed. 1. Der erste Fall kann daher nicht eintreten. Zweiter Fall: $\lambda > 0 \Longrightarrow$ (wg. Bed. 2) $x_1 + x_2 = 2$. Die Lösung des Gleichungssystems für x und λ ergibt $\lambda = \frac{7}{2}$ und $x_1 = \frac{1}{2}$, $x_2 = \frac{3}{2}$. Dies erfüllt tatsächlich alle 3 KUHN-TUCKER-Bed. und der zweite Fall führt zur Lösung.

2.17 Der Satz über implizite Funktionen

Die folgende Fragestellung ist sehr wichtig:
Es sei $g\colon \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^m$ eine Funktion. Ist es möglich, „die Gleichung $g(x,y) = 0, x \in \mathbb{R}^n$, nach $y \in \mathbb{R}^m$ aufzulösen“? Mathematisch präzise lautet die Fragestellung: Gibt es eine Funktion $F\colon \mathbb{R}^n \to \mathbb{R}^m$ mit

$$g(x, F(x)) = 0 \text{ für alle } x \in \mathbb{R}^n \text{ oder wenigstens } x \in U \subset \mathbb{R}^n?$$

Eine Auskunft hierüber gibt der

Satz 17.1. (Satz über implizite Funktionen)
Es sei U eine Umgebung des Nullvektors $0 \in \mathbb{R}^n$ sowie V eine Umgebung des Nullvektors $0 \in \mathbb{R}^m$. Die Funktion $g : U \times V \to \mathbb{R}^m$ sei stetig und sei bezüglich der Variablen $y \in V$ einmal partiell differenzierbar. Die $m \times m$-Matrix $\frac{\partial g}{\partial y}(x,y)$ der partiellen Ableitungen von g bezüglich $y \in \mathbb{R}^m, x \in U \subset \mathbb{R}^n, y \in V \subset \mathbb{R}^m$, sei eine stetige Funktion bezüglich aller Variablen, und es gelte

$$\det\left(\frac{\partial g}{\partial y}(0,0)\right) \neq 0$$

sowie $g(0,0) = 0$.
Dann gibt es eine ϑ-Umgebung U_ϑ von $0 \in \mathbb{R}^n$ sowie eine Funktion $F\colon U_\vartheta \to V$, so daß gilt

$$g(x, F(x)) = 0.$$

Die Funktion F hat noch gewisse Eigenschaften, die ebenfalls als Teil des Satzes über implizite Funktionen gelten, die wir jedoch in den anschließenden Sätzen festhalten.
Beweis:
Da $\det\frac{\partial g}{\partial y}(0,0) \neq 0$, existiert $\Gamma := (\frac{\partial g}{\partial y}(0,0))^{-1}$. Für jedes $x \in U$ definieren wir die Abbildung $T_x\colon V \to \mathbb{R}^m$ durch $T_x(y) = y - \Gamma\, g(x,y)$. Aus dem Mittelwertsatz und der Stetigkeit von $\frac{\partial g}{\partial y}$ bei $(0,0) \in \mathbb{R}^n \times \mathbb{R}^m$ folgt, daß zu vorgegebenem $\varepsilon > 0$ ein $\delta > 0$ existiert mit

$$\left| g(x,y) - g(x,z) - \frac{\partial g}{\partial y}(0,0)(y-z) \right| < \varepsilon |y - z|$$

für alle $x \in U_\delta(0) \subset \mathbb{R}^n$ und $y, z \in V_\delta(0) \subset \mathbb{R}^m$.
Unter Benutzung der Matrix $\Gamma = (\frac{\partial g}{\partial y}(0,0))^{-1}$ folgt

$$\begin{aligned} |y - z - \Gamma(g(x,y) - g(x,z))| &= \left| \Gamma\left(\frac{\partial g}{\partial y}(0,0)(y-z) - g(x,y) + g(x,z)\right)\right| \\ &< C \left| g(x,y) - g(x,z) - \frac{\partial g}{\partial y}(0,0)(y-z) \right| \\ &< \varepsilon C |y - z|, \end{aligned}$$

wobei man für C z.B. die Summe der Absolutbeträge der Koeffizienten von Γ nehmen kann. Wählt man $\varepsilon = \frac{q}{C}$, so erkennen wir, daß zu vorgegebenem q ein $\delta > 0$ existiert, so daß für alle $x \in U_\delta(0) \subset \mathbb{R}^n$, $y, z \in V_\delta(0) \subset \mathbb{R}^m$ die Ungleichung

$$T_x(y) - T_x(z) = |y - z - \Gamma(g(x,y) - g(x,z))| < q|y-z|$$

folgt. Außerdem gilt wegen der Stetigkeit von g bei vorgegebenen $\varepsilon > 0$

$$|T_x(0)| = |\Gamma(g(x,0) - g(0,0))| < \varepsilon$$

für $x \in U_\vartheta(0)$ bei genügend kleinem $\vartheta > 0, \vartheta \leq \delta$. Insbesondere kann man $\varepsilon = \delta(1-q)$ setzen. Aus Satz 17.3. folgt die Existenz eines *Fixpunktes* $y \in V_\delta(0) \subset V$ von T_x, d.h. es gilt

$$T_x(y) = y \quad \text{oder} \quad y - \Gamma g(x,y) = y \quad \text{oder} \quad \Gamma g(x,y) = 0.$$

Wir setzen dieses y gleich $F(x)$, und es gilt

$$g(x, F(x)) = 0\,, \; x \in U_\delta(0).$$

Der Satz ist damit bewiesen. □

Der Beweis von Satz 17.1. beruhte wesentlich auf dem *Prinzip der kontrahierenden Abbildung*, deren Standardformulierung in dem folgenden Satz gegeben wird:

Satz 17.2.

Es sei $T\colon \mathbb{R}^m \to R^m$ stetig und es existiere ein $q \in]0,1[$ mit

$$|T(y) - T(z)| \leq q|y - z|$$

für alle $y, z \in \mathbb{R}^m$. Dann besitzt T einen eindeutigen Fixpunkt y_0, d.h. es gilt $T(y_0) = y_0$. Außerdem konvergiert für jedes $z_0 \in \mathbb{R}^m$ die Folge $(T^N(z_0))_{N\in\mathbb{N}}$ gegen y_0.

Beweis:

Es sei $z_0 \in \mathbb{R}^m$ beliebig. Wir betrachten die Folge $(T^N(z_0))_{N\in\mathbb{N}}$. Wir erhalten durch wiederholtes Einsetzen der Voraussetzung

$$\begin{aligned} \left|T^{N+1}(z_0) - T^N(z_0)\right| &\leq q\left|T^N(z_0) - T^{N-1}(z_0)\right| \\ &\leq q^N\left|T(z_0) - z_0\right|. \end{aligned}$$

Daraus folgt für $N > M \geq 1$, $N, M \in \mathbb{N}$

$$\begin{aligned} |T^N(z_0) - T^M(z_0)| &\leq \sum_{i=M}^{N-1} |T^{i+1}(z_0) - T^i(z_0)| \\ &\quad \text{nach der Dreiecksungleichung} \\ &\leq \sum_{i=M}^{N-1} q^i |T(z_0) - z_0| \\ &= |T(z_0) - z_0| \sum_{i=M}^{N-1} q^i. \end{aligned}$$

Da wegen der Voraussetzung $0 < q < 1$ die Reihe $\sum_{i=1}^{\infty} q^i$ konvergiert, kann wegen des Cauchyschen Konvergenzkriteriums der Ausdruck $\sum_{i=M}^{N-1} q^i$ unter jede Schranke ε gedrückt werden, wenn man M und N genügend groß wählt. Daraus folgt die Existenz einer Zahl N_0, so daß für $N, M \geq N_0$ die Ungleichung

$$|T^N(z_0) - T^M(z_0)| \leq |T(z_0) - z_0| \sum_{i=M}^{N-1} q^i$$

gilt. Dies bedeutet, daß die Folge $(T^N(z_0))_{N\in\mathbb{N}}$ eine Cauchyfolge ist und gegen einen Vektor y_0 konvergieren muß.Wegen der Stetigkeit von T konvergiert

$$T(T^N(z_0)) \to T(y_0) \qquad (N \to \infty).$$

Andererseits konvergiert $T^{N+1}(z_0) \to y_0$.
Daraus folgt $y_0 = T(y_0)$ und somit die Existenz eines Fixpunktes y_0 von T sowie die Konvergenzaussage über $T^N(z_0)$.
Die Eindeutigkeit folgt ebenfalls in einfacher Weise: Sind y_0 und $\tilde{y}_0$ Fixpunkte von T, so gilt

$$|y_0 - \tilde{y}_0| = |T(y_0) - T(\tilde{y}_0)| \leq q|y_0 - \tilde{y}_0|.$$

Da nach Voraussetzung $q < 1$ ist, folgt $y_0 = \tilde{y}_o$. Der Satz ist damit bewiesen. □

Zum Nachweis von Satz 17.1. benötigten wir eine Variante von Satz 17.2, nämlich

Satz 17.3.
Es sei T eine stetige Abbindlung einer Nullumgebung $V \subset \mathbb{R}^m$ nach $\mathbb{R}^m$, und zu jedem $q \in]0,1[$ existiere ein $\delta > 0$, so daß $V_\delta(0) \subset V$ und

$$|T(y) - T(z)| < q|y - z| \quad \textit{sowie} \quad |T(0)| \leq \delta(1 - q)$$

für alle $y, z \in V_\delta(0)$. Dann gibt es einen Fixpunkt *$y_0 \in V_\delta(0)$ von T.*
Beweis:
Sei $q \in]0,1[$ vorgegeben. Es gibt dann ein $\delta_0 > 0$ mit $V_{\delta_0}(0) \subset V$ und $|T(y) - T(z)| \leq q|y - z|$ für $y, z \in V_{\delta_0}(0)$. Wie beim Beweis von Satz 17.2. bilden wir die Iterationsfolge $(T^N(0))_{N\in\mathbb{N}}$. Es gilt dann (vgl. den obigen Beweis)

$$\begin{aligned} |T^N(0)| &\leq \sum_{i=0}^{N-1} |T^{i+1}(0) - T^i(0)| \\ &\leq \sum_{i=0}^{N-1} q^i |T(0) - 0| \\ &\leq |T(0)| \frac{1}{1-q} \\ &\leq \delta. \end{aligned}$$

Daraus folgt, daß die Vektoren $T^N(0)$ alle in $V_\delta(0)$ liegen, $T^N(0)$ also immer definiert ist. Die Existenz eines Fixpunktes folgt nun wie in Satz 17.2. $\square$

Das Resultat über die Eindeutigkeit des Fixpunktes *kontrahierender* Abbildungen (also Abbildungen mit der Eigenschaft $|T(y) - T(z) \leq q|y - z|,\ 0 < q < 1)$ ergibt ein entsprechendes Resultat bei impliziten Funktionen:

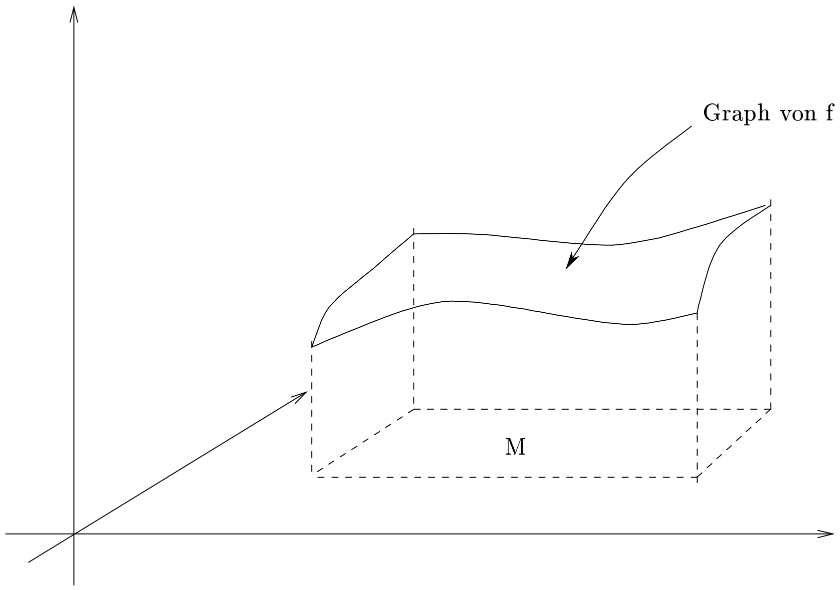
Satz 17.4.
Es mögen die Voraussetzungen von Satz 17.1. gelten. Dann gibt es eine ϑ-Umgebung U_ϑ des Nullvektors $0 \in \mathbb{R}^n$ und eine δ-Umgebung V_δ des Nullvektors $0 \in \mathbb{R}^m$, in der die Funktion F aus Satz 17.1. eindeutig ist, d.h. aus

$$g(x,y) = 0\ ,\ x \in U_\vartheta(0)\ ,\ y \in V_\delta(0) \quad \textit{folgt } y = F(x).$$

Außerdem läßt sich beweisen(siehe z.B.[5, Band 2, Nr. 134]) , daß die Funktion F in Satz 17.1. k-mal stetig partiell differenzierbar ist, wenn die Funktion g k-mal stetig partiell differenzierbar ist.

2.18 Das n-dimensionale Riemannsche Intergral

Analog zum eindimensionalen Fall wird man zum Begriff des n-dimensionalen Riemannschen Integerals einer Funktion $f : \mathbb{R}^n \to \mathbb{R}$ über einer Punktmenge M geführt, wenn man den „Rauminhalt“ der „$(n+1)$-dimensionalen“ Punktmenge bestimmen will, die durch M und den Graphen von f begrenzt wird:



Zur Definition des Rauminhalts der Punktmenge, die durch das Rechteck M und den Graphen der Funktion f (sowie den „Seitenwänden“ –siehe Skizze) begrenzt wird, geht man ähnlich vor wie im eindimensionalen Fall.

Zunächst bereiten wir die Definition des Riemannschen Integrals über ein n-dimensionales Intervall

$$I = [a,b] = \{ x = (\xi_1, \dots , \xi_n) \in \mathbb{R}^n \,|\, \alpha_i \leq \xi_1 \leq \beta_i \}$$

mit $a = (\alpha_1, \dots , \alpha_n) \in \mathbb{R}^n$ und $b = (\beta_1, \dots , \beta_n) \in \mathbb{R}^n$ vor.

Definition:
Der Inhalt des n-dimensionalen Intervalls *I=[a,b] ist die Zahl*

$$|I| = \prod_{i=1}^{n} (\beta_i - \alpha_i).$$

Definition:
Es seien $I_1, \dots , I_N$ n-dimensionale abgeschlossene Intervalle. Das System $\mathfrak{Z} = \{I_1, \dots , I_N\}$ heißt Zerlegung *des n-dimensionalen Intervalls I, wenn die offenen Kerne von $I_1, \dots , I_N$ paarweise disjunkt sind und*

$$\bigcup_{i=1}^{N} I_i = I$$

gilt. Die Intervalle $I_i \in \mathfrak{Z}$ heißen Teilintervalle der Zerlegung. *Die größte Kantenlänge der Intervalle I_i, also*

$$\beta_j^{(i)} - \alpha_j^{(i)}, \qquad j = 1, \dots , n, \qquad i = 1, \dots , N \qquad , I_i = \left[a^{(i)}, b^{(i)} \right]$$

mit $a^{(i)} = (\alpha_1^{(i)}, \dots , \alpha^{(i)})$, $b^{(i)} = (\beta_1^{(i)}, \dots , \beta_n^{(i)})$ heißt die Feinheit $|\mathfrak{Z}|$ *von $\mathfrak{Z}$.*

Die Zerlegung $\mathfrak{Z}^*$ von I heißt eine Verfeinerung *von $\mathfrak{Z}$, wenn jedes Teilungsintervall von $\mathfrak{Z}^*$ ganz in einem Teilungsintervall vom $\mathfrak{Z}$ enthalten ist. (Es gilt dann $|\mathfrak{Z}^*| \leq |\mathfrak{Z}|$).*

Es seien $\mathfrak{Z}_1$ und $\mathfrak{Z}_2$ zwei Zerlegungen desselben Intervalls I, etwa

$$\mathfrak{Z}_1 = \left\{ I_1^{(1)}, \dots , I_N^{(1)} \right\} \qquad , \mathfrak{Z}_2 = \left\{ I_1^{(2)}, \dots , I_N^{(2)} \right\}.$$

Die nichtleeren Punktmengen unter den Durchschnitten $I_i^{(1)} \cap I_k^{(2)}$ sind ebenfalls Intervalle, und diese bilden wieder eine Zerlegung von I, die man Superposition $\mathfrak{Z}_1\ \mathfrak{Z}_2$ *von $\mathfrak{Z}_1$ und $\mathfrak{Z}_2$ nennt. Sie ist die gröbste gemeinsame Verfeinerung von $\mathfrak{Z}_1$ und $\mathfrak{Z}_2$.*

Es ist anschaulich einleuchtend, daß

$$|I| = \sum_{i=1}^{N} |I_i|$$

gilt, wenn $\mathfrak{Z} = \{I_1, \dots, I_N\}$ eine Zerlegung von I ist.

Beweis:
Der Beweis kann z.B. nachgelesen werden in [2, Seite 106].

Definition:
Es sei I ein n-dimensionales Intervall und $\mathfrak{Z} = \{I_1, \dots, I_N\}$ eine Zerlegung von I. Ferner seien Punkte $\xi_i \in I_1$ $i = 1, \dots .N$ (sogenannte Teilungspunkte *zur Zerlegung $\mathfrak{Z}$) gegeben. Die* Riemannsche Summe *der Funktion $f \colon I \to \mathbb{R}$ ist die Zahl*

$$\sum(f, \xi_i, \mathfrak{Z}) = \sum_{i=1}^{N} f(\xi_i)\, |I_i|\,.$$

Ist f beschränkt, so lassen sich die Riemannsche Obersumme $O_{\mathfrak{Z}}$ *und die* Riemannsche Untersumme $U_{\mathfrak{Z}}$ *definieren:*

$$O_{\mathfrak{Z}} = \sum_{i=1}^{N} \sup\{f(\xi)\,|\,\xi \in I_i\}\,|I_i|\,,$$
$$U_{\mathfrak{Z}} = \sum_{i=1}^{N} \inf\{f(\xi)\,|\,\xi \in I_i\}\,|I_i|\,.$$

Definition:
Es sei I ein n-dimensionales Intervall. Die Funktion $f \colon I \to \mathbb{R}$ heißt Riemann-integrierbar *über I, wenn eine Zahl A mit der folgenden Eigenschaft existiert:*

$$\forall \varepsilon > 0\, \exists \delta > 0 \quad \textit{mit}\ \forall \mathfrak{Z}, |\mathfrak{Z}| < \delta \ \textit{mit Zwischenpunkten}\ \xi_i : \ \left|\sum(f, \xi_i, \mathfrak{Z}) - A\right| < \varepsilon.$$

*Die Zahl A heißt dann das n-*dimensionale Riemannsche Integral *von f über I, in Zeichen*

$$A = \int_I f(x)\, dx.$$

Definition:
Es sei I ein n-dimensionales Intervall und die Funktion $f \colon I \to \mathbb{R}$ gegeben. Das Riemannsche Unterintegral $\underline{J}$ *bzw.* Oberintegral $\overline{J}$ *ist definiert durch*

$$\underline{J} = \sup\{U_{\mathfrak{Z}}\,|\,\mathfrak{Z} \quad \textit{ist Zerlegung von}\ I\}$$

bzw.

$$\overline{J} = \inf\{O_{\mathfrak{Z}}\,|\,\mathfrak{Z} \quad \textit{ist Zerlegung von}\ I\}\,.$$

Ähnlich wie im Fall des eindimensionalen Riemannschen Integrals beweist man

Satz 18.1.
Es sei I ein n-dimensionales Intervall und $f\colon I \to \mathbb{R}$ eine beschränkte Funktion. Die folgenden Aussagen sind äquivalent:

1. *f ist Riemann-integrierbar über I.*
2. *Das Riemannsche Oberintegral von f ist gleich dem Riemannschen Unterintegral.*
3. *Riemannsches Integrabilitätskriterium:*
   Zu jedem $\varepsilon > 0$ gibt es eine Zerlegung $\mathfrak{Z}$ von I, so daß
   $$O_{\mathfrak{Z}} - U_{\mathfrak{Z}} < \varepsilon.$$
   (Hierbei bedeutet $O_{\mathfrak{Z}}$ bzw. $U_{\mathfrak{Z}}$ die Riemannsche Obersumme bzw. Untersumme bzgl. der Zerlegung $\mathfrak{Z}$.)

Die Bedeutung des Riemannschen Integrabilitätskriterium liegt wieder darin, daß man zum Nachweis der Riemann-Integrierbarkeit die Bedingung

$$O_{\mathfrak{Z}} - U_{\mathfrak{Z}}$$

nur für spezielle Zerlegungen nachzuweisen braucht.

Mit Hilfe dieses Kriteriums ergibt sich wie im eindimensionalen Fall der

Satz 18.2.
Es sei I ein n-dimensionales abgeschlosenes Intervall und $f\colon I \to \mathbb{R}$ stetig. Dann ist f Riemann-integrierbar über I.

Wir wollen nun die Definition des Riemann-Integrals auf den Fall ausdehnen, daß das Integrationsgebiet kein n-dimensionales Intervall, sondern eine beliebige beschränkte Punktmenge M ist.

Definition:
Es sei $f\colon M \to \mathbb{R}$ eine Funktion und T ein n-dimensionales Intervall mit $T \supset M$. Wir setzen f zu einer Funktion $\tilde{f}\colon T \to \mathbb{R}$ fort durch

$$\tilde{f}(x) = \begin{cases} f(x) & \text{für } x \in M, \\ 0 & \text{für } x \in T \setminus M. \end{cases}$$

Die Funktion $f\colon M \to \mathbb{R}$ *heißt* Riemann-integrierbar über M, *wenn* $\tilde{f}$ *Riemann-integrierbar über* T *ist.*

Wir setzen

$$\int\limits_M f(x)\,dx := \int\limits_T \tilde{f}(x)\,dx.$$

Eine Funktion $f\colon M \to \mathbb{R}$ muß nicht notwendig Riemann-integrierbar über M sein, auch wenn f stetig auf M oder sogar konstant ist.

Wir notieren einige Regeln:

Satz 18.3.(Linearität)
Es sei M *eine beschränkte Punktmenge des* $\mathbb{R}^n$, *und* f_1, f_2 *seien Riemann-integrierbar über* M. *Dann ist für* $\alpha, \beta \in \mathbb{R}$ *auch* $\alpha f_1 + \beta f_2$ *über* M *Riemann-integrierbar, und es gilt*

$$\int\limits_M (\alpha f_1 + \beta f_2)(x)\,dx = \alpha \int\limits_M f_1(x)\,dx + \beta \int\limits_M f_2(x)\,dx.$$

Der Beweis wird wie im eindimensionalen Fall auf die Riemannschen Summen zurückgespielt. Ähnliches gilt für die folgenden Sätze:

Satz 18.4. (Monotonie)
Es sei M *eine beschränkte Punktmenge des* $\mathbb{R}^n$, *und* f_1, f_2 *seien Riemann-integrierbar über* M. *Überdies sei*

$$f_1(x) \le f_2(x) \quad \text{für alle } x \in M.$$

Dann gilt

$$\int\limits_M f_1(x)\,dx \le \int\limits_M f_2(x)\,dx.$$

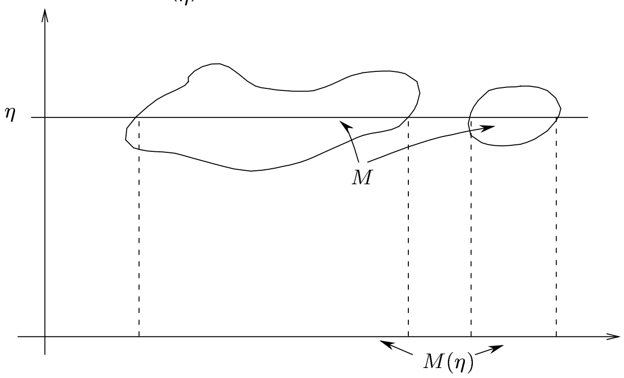
Satz 18.5.
Es seien M_1, M_2 *beschränkte Teilmengen des* $\mathbb{R}^n$, *und die Funktion* $f\colon M_1 \cup M_2 \to \mathbb{R}$ *sei sowohl über* M_1 *als auch über* M_2 *Riemann-integrierbar. Dann ist auch* f *über* $M_1 \cup M_2$ *Riemann-integrierbar.*

2.19 Mehrfache Integrale

In diesem Kapitel wird eine Methode angegeben, wie man die explizite Berechnung von mehrdimensionalen Riemannschen Integralen gegebenenfalls durch mehrfache eindimensionale

Integrale erreicht. Zur Erläuterung beschränken wir uns zunächst auf zweidimensionale Integrale.

Definition:
Es sei $M \subset \mathbb{R}^2$ beschränkt. Jedem $\eta \in \mathbb{R}$ ordnen wir die Menge $M(\eta)$ derjenigen ξ zu, für die $\binom{\xi}{\eta} \in M$ ist. Die Menge der η, für die $M(\eta)$ nicht leer ist, nennen wir $\mathfrak{O}$.



Wenn nun für jedes $\eta \in \mathfrak{O}$ das eindimensionales Integral $F(\eta) = \int\limits_{M(\eta)} f(\xi,\eta)\,d\xi$ existiert und das Integral $\int\limits_{\mathfrak{O}} F(\eta)d\eta$ existiert, so nennt man letzteres das Doppelintegral

$$\iint\limits_{M} f(\xi,\eta)\,d\xi\,d\eta$$

von f über M. Wenn speziell

$$M = \left\{ \begin{pmatrix} \xi \\ \eta \end{pmatrix} \middle| \varphi(\eta) \leq \xi \leq \psi(\eta),\ \eta \in [\alpha_2, \beta_2] \right\}$$

mit reellen Funktionen $\varphi, \psi \colon \mathbb{R} \to \mathbb{R}$ und Zahlen $\alpha_2, \beta_2 \in \mathbb{R}$, so schreibt man auch

$$\int\limits_{\alpha_2}^{\beta_2} \int\limits_{\varphi(\eta)}^{\psi(\eta)} f(\xi,\eta)\,d\xi\,d\eta,$$

insbesondere im Fall $\psi(\eta) = \beta_1$, $\varphi(\eta) = \alpha_1$ für $\eta \in [\alpha_2, \beta_2]$

$$\int\limits_{\alpha_2}^{\beta_2} \int\limits_{\alpha_1}^{\beta_1} f(\xi,\eta)\,d\xi\,d\eta.$$

Es erhebt sich die Frage, ob die Gleichung

$$\iint\limits_{M} f(\xi,\eta)\,d\xi\,d\eta = \int\limits_{M} f(x)\,dx$$

gilt.
Hierzu gilt

Satz 19.1.
Es sei M eine beschränkte Punktmenge, und die Funktion $f\colon M \to \mathbb{R}$ sei beschränkt und Riemann-integrierbar. Außerdem existiere das Doppelintegral $\iint\limits_M f(\xi,\eta)d\xi d\eta$, d.h. für jedes $\eta \in \mathfrak{D}$ existere $F(\eta) = \int_{M(\eta)} f(\xi,\eta)d\xi$ und es existere $\int_{\mathfrak{D}} F(\eta)d\eta$, jeweils im Riemannschen Sinne. Dann gilt

$$\int\limits_M f(x)\,dx = \iint\limits_M f(\xi,\eta)\,d\xi\,d\eta.$$

Beweis:
Es sei

$$I = \left\{ \begin{pmatrix} \xi \\ \eta \end{pmatrix} \middle| \alpha \le \xi \le \beta, \alpha \le \eta \le \beta \right\}$$

ein Quadrat, welches M enthalte. Wir setzen

$$\bar{f}(x) = \begin{cases} f(x) & \text{für } x \in M, \\ 0 & \text{für } x \in I - M. \end{cases}$$

Dann gilt

$$\begin{aligned} \int\limits_M f\,dx &= \int\limits_I \bar{f}\,dx \quad \text{und} \\ \int\limits_\alpha^\beta \int\limits_\alpha^\beta \bar{f}(\xi,\eta)\,d\xi\,d\eta &= \iint\limits_I \bar{f}(\xi,\eta)\,d\xi\,d\eta, \end{aligned}$$

und wir brauchen nur die Gleichheit

$$\int\limits_I \bar{f}\,dx = \int\limits_\alpha^\beta \int\limits_\alpha^\beta \bar{f}(\xi,\eta)\,d\xi\,d\eta$$

zu beweisen.
Wir bilden hierzu die eindimensionalen Intervalle

$$I_j = \left[\alpha + \frac{j-1}{N}(\beta-\alpha) \middle| \alpha + \frac{j}{N}(\beta-\alpha) \right], \quad j = 1, \dots, N,$$

die das Intervall $[\alpha, \beta]$ äquidistant unterteilen, und die zweidimensionalen Intervalle

$$I_{jk} = \left\{ \begin{pmatrix} \xi \\ \eta \end{pmatrix}, \xi \in I_j, \eta \in I_k \right\}.$$

Die Menge $\{I_{jk} | j, k = 1, \ldots, N\}$ bildet dann eine Zerlegung $\mathfrak{Z}_N$ von I.
Es sei $\overline{S_j}(\eta) = \sup\{\bar{f}(\xi, \eta) | \xi \in I_j\}$ und $\overline{S_{jk}} = \sup\{\bar{f}(\xi, \eta) | \binom{\xi}{\eta} \in I_{jk}\}$.
Die Summe

$$\frac{\beta - \alpha}{N} \sum_{j=1}^{N} \bar{S}_j(\eta) \quad \text{ist eine Obersumme für} \quad \int_{\alpha}^{\beta} \bar{f}(\xi, \eta)\, d\xi,$$

die Summe

$$\left(\frac{\beta - \alpha}{N}\right)^2 \sum_{j,k=1}^{N} \bar{S}_{jk} \quad \text{eine Obersumme für} \quad \int_{I} \bar{f}(x)\, dx.$$

Für die zugehörigen Riemannschen Integrale gilt daher

$$\int_{\alpha}^{\beta} \bar{f}(\xi, \eta) d\xi \leq \frac{\beta - \alpha}{N} \sum_{j=1}^{N} \bar{S}_j(\eta) \leq \frac{\beta - \alpha}{N} \sum_{j=1}^{N} \bar{S}_{jk} \quad \text{für} \quad \eta \in I_k.$$

Integriert man beide Seiten der Ungleichung bezüglich η, so erhält man

$$\int_{I_k} \left(\int_{\alpha}^{\beta} \bar{f}(\xi, \eta)\, d\xi \right) d\eta \leq \left(\frac{\beta - \alpha}{N}\right)^2 \sum_{j=1}^{N} \bar{S}_{jk}$$

und nach der Summation von $k = 1$ bis N

$$\int_{\alpha}^{\beta} \int_{\alpha}^{\beta} \bar{f}(\xi, \eta)\, d\xi\, d\eta \leq \left(\frac{\beta - \alpha}{N}\right)^2 \sum_{j=1}^{N} \bar{S}_{jk} = O_{\mathfrak{Z}_k}.$$

Da nach den allgemeinen Sätzen

$$O_{\mathfrak{Z}_N} \to \int \bar{f}\, dx \quad (N \to \infty)$$

konvergiert, folgt

$$\int_{\alpha}^{\beta} \int_{\alpha}^{\beta} \bar{f}(\xi, \eta)\, d\xi\, d\eta \leq \int_{I} \bar{f}\, dx,$$

und eine entsprechende Überlegung über Untersummen ergibt

$$\int_{\alpha}^{\beta} \int_{\alpha}^{\beta} \bar{f}(\xi, \eta)\, d\xi\, d\eta \geq \int_{I} \bar{f}\, dx.$$

Der Satz ist damit bewiesen. $\square$

Analog zu $\iint_M f(\xi, \eta)\, d\xi\, d\eta$ kann man auch das Doppelintegral $\iint_M f(\xi, \eta)\, d\eta\, d\xi$ definieren, das entsteht, wenn man zuerst die Integration bzgl. η und dann bzgl. ξ ausführt. Wenn beide Doppelintegrale existieren und außerdem das zweidimensionale Riemannsche Integral $\int\limits_M f\,dx$ existiert, so müssen wegen Satz 19.1 auch die Doppelintegrale $\int_M f(\xi,\eta)\,d\xi\,d\eta$ und $\int_M \int f(\xi,\eta)\,d\eta\,d\xi$ gleich sein, d.h. es gilt

Satz 19.2.
Es mögen die Voraussetzungen von Satz 19.1. gelten, außerdem existiere

$$\iint\limits_M f(\xi,\eta)\,d\eta\,d\xi.$$

Dann ist die Reihenfolge der Integration vertauschbar.

Entsprechend zum zweidimensionalen Fall läßt sich ein n-faches Integral definieren. Es gilt dann ein n-dimensionales Analogon zu Satz 19.1. und 19.2.:

Satz 19.3.
Es sei $M \subset \mathbb{R}^n$ eine beschränkte Punktmenge und die Funktion $f\colon M \to \mathbb{R}$ sei beschränkt und Riemann-integrierbar. Dann ist jedes n-fache Integral von f über M, falls es existiert, gleich dem n-dimensionalen Riemannschen Integral $\int\limits_M f\,dx$. Das n-fache Integral ist also in diesem Sinne von der Reihenfolge der Integration unabhängig.

Anwendungen der Sätze 19.1. – 19.3. finden sich in Kapitel 21.

2.20 Vertauschbarkeit von Differentiation und Integration; Differentiation von Integralen mit variablen Grenzen

Die Fragestellung in der Überschrift ist naheliegend.

Es gilt hierzu

Satz 20.1.
Es sei $a < b, \alpha < \beta$ $, \alpha, \beta, a, b \in \mathbb{R}$. Es sei f eine auf dem Rechteck $[a,b]\times]\alpha,\beta[$ definierte stetige Funktion mit Werten in $\mathbb{R}$. Für jedes $\xi \in [a,b]$ sei $f(\xi,\cdot)\colon]\alpha,\beta[\to \mathbb{R}$ überall in $]\alpha,\beta[$ differenzierbar, und die durch die partielle Ableitung $\frac{\partial}{\partial x} f(\xi,x)$ definierte Funktion sei stetig auf $[a,b]\times]\alpha,\beta[$.

Dann ist die Funktion $F\colon]\alpha,\beta[\to \mathbb{R}$, die definiert ist durch

$$F(x) = \int\limits_a^b f(\xi,x)\,d\xi,$$

differenzierbar, und es gilt

$$\frac{d}{dx}F(x) = \int_a^b \frac{\partial}{\partial x} f(\xi, x)\, d\xi.$$

Beweis:
Es sei $x \in]\alpha, \beta[$. Für genügend kleines $h \neq 0$, $h \in \mathbb{R}$, ist $F(x+h)$ definiert, und es gilt

$$\frac{F(x+h) - F(x)}{h} = \int_a^b \frac{f(\xi, x+h) - f(\xi, x)}{h}\, d\xi.$$

Aus dem Mittelwertsatz folgt

$$\left| \frac{f(\xi, x+h) - f(\xi, x)}{h} - \frac{\partial f(\xi, x)}{\partial x} \right| = \left| \frac{\partial f(\xi, x+\theta h)}{\partial x} - \frac{\partial f(\xi, x)}{\partial x} \right|.$$

Da $\frac{\partial f}{\partial x}$ gleichmäßig stetig in $[a, b] \times U(x)$ ist — $U(x)$ ist hierbei eine in $]\alpha, \beta[$ liegende abgeschlossene Umgebung von x — , folgt, daß zu vorgegebenem ε ein δ existiert, so daß für alle $h \neq 0$, $h \in \mathbb{R}$, mit $|h| < \delta$ die Ungleichung

$$\left| \frac{f(\xi, x+h) - f(\xi, x)}{h} - \frac{\partial f(\xi, x)}{\partial x} \right| < \frac{\varepsilon}{b-a}$$

gilt. Daraus folgt für $h \neq 0$, $h \in \mathbb{R}$, mit $|h| < \delta$

$$\begin{aligned} \left| \frac{F(x+h) - F(x)}{h} - \int_a^b \frac{\partial}{\partial x} f(\xi, x)\, d\xi \right| &\leq \left| \int_a^b \left[\frac{f(\xi, x+h) - f(\xi, x)}{h} - \frac{\partial}{\partial x} f(\xi, x) \right] d\xi \right| \\ &\leq (b-a)\frac{\varepsilon}{b-a} = \varepsilon. \end{aligned}$$

Daraus folgt schließlich die Differenzierbarkeit von F mit

$$F'(x) = \int_a^b \frac{\partial}{\partial x} f(\xi, x)\, d\xi.$$

□

Die Formel in diesem Satz gilt nur bei festen Differentiationsgrenzen. Im Fall variabler Grenzen gilt der folgende

Satz 20.2.
Es sei $a < b$, $\alpha < \beta$, $a, b, \alpha, \beta \in \mathbb{R}$ und $f\colon [a,b]\times]\alpha,\beta[\to \mathbb{R}$ stetig und bzgl. der Variablen aus $]\alpha,\beta[$ stetig partiell differenzierbar. Außerdem sei $\psi\colon]\alpha,\beta[\to [a,b]$ und $\phi\colon]\alpha,\beta[\to [a,b]$ stetig differenzierbar. Dann ist die durch

$$F(x) = \int\limits_{\phi(x)}^{\psi(x)} f(\xi, x)\, d\xi$$

definierte Funktion $F\colon]\alpha,\beta[\to \mathbb{R}$ differenzierbar, und es gilt

$$F'(x) = \int\limits_{\phi(x)}^{\psi(x)} \frac{\partial}{\partial x} f(\xi, x)\, d\xi + \psi'(x) f(\psi(x), x) - \phi'(x) f(\phi(x), x).$$

Beweis:
Wir definieren $G\colon]\alpha,\beta[\times\phi(]\alpha,\beta[) \times \psi(]\alpha,\beta[) \to \mathbb{R}$ durch

$$G(x,y,z) = \int\limits_{y}^{z} f(\xi, x)\, d\xi.$$

Es gilt dann

$$F(x) = \int\limits_{\phi(x)}^{\psi(x)} f(\xi, x)\, d\xi = G(x, \phi(x), \psi(x)).$$

Die Funktion G ist in allen Variablen stetig partiell differenzierbar. Nach der Kettenregel gilt daher

$$\begin{aligned}
\frac{d}{dx} F(x) &= \Big((\nabla G)\big(x, \phi(x), \psi(x)\big) \Big) \cdot \frac{d}{dx} \begin{pmatrix} x \\ \phi(x) \\ \psi(x) \end{pmatrix} \\
&= \left(\int\limits_{\phi(x)}^{\psi(x)} \frac{\partial}{\partial x} f(\xi, x)\, d\xi \,,\, -f(\phi(x), x) \,,\, f(\psi(x), x) \right) \cdot \begin{pmatrix} 1 \\ \phi'(x) \\ \psi'(x) \end{pmatrix} \\
&= \int\limits_{\phi(x)}^{\psi(x)} \frac{\partial}{\partial x} f(\xi, x)\, d\xi + f(\psi(x), x)\psi'(x) - f(\phi(x), x)\phi'(x).
\end{aligned}$$

$\square$

2.21 Der Riemannsche Inhalt beschränkter Punktmengen

Es sei $M \subset \mathbb{R}^n$ beschränkt. Es ist naheliegend, das n-dimensionale Riemannsche Integral $\int\limits_M 1\,dx$, kürzer $\int\limits_M dx$, als (n-dimensionales) *Volumen* von M (auch n-dimensionaler *Riemannscher Inhalt* von M oder *Riemannsches Maß* von M) zu definieren. Ist $T \supset M$ ein n-dimensionales Intervall und die charakteristische Funktion χ_M von M definiert durch

$$\chi_M(x) = \begin{cases} 1 & \text{für } x \in M, \\ 0 & \text{für } x \in \mathbb{R}^n - M, \end{cases}$$

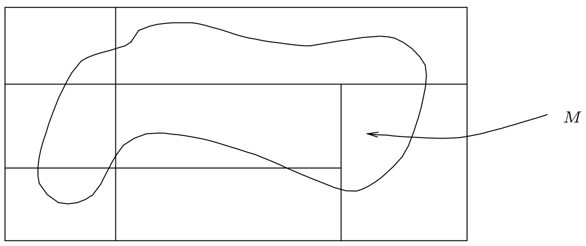
so kann man auch schreiben

$$\int\limits_M 1\,dx = \int\limits_M dx = \int\limits_T \chi_M(x) := |M|.$$

Wenn M „teuflisch" ist, braucht $\int\limits_M dx$ nicht zu existieren.

Definition:
Wenn $\int\limits_M dx$ existert, so heißt M Riemann-meßbar*.*

Ob $\int\limits_M dx$ existiert, kann mit Hilfe von Satz 18.1. ausgedrückt werden, wenn das Ober- und Unterintegral von χ_M verwendet wird.
Im vorliegenden Fall bezeichnet man das Oberintegral der Funktion $\chi_M \colon T \to \mathbb{R}$ auch als das *äußere Riemannsche Maß* von M, das Unterintegral von $\chi_M \colon T \to \mathbb{R}$ als das *innere Riemannsche Maß* von M. Man mache sich klar, wie die Ober- und Untersummen von χ_M aussehen:



Die Menge M wird durch die Quader einer Zerlegung überdeckt. Die Untersumme zu χ_M ist gleich dem Inhalt aller Quader der Zerlegung, die ganz in M liegen, die Obersumme zu χ_M ist gleich dem Inhalt aller Quader der Zerlegung, die mindestens einen Punkt mir N gemeinsam haben. Wegen Satz 18.1. gilt

Satz 21.1.
Eine beschränkte Menge M ist genau dann Riemann-meßbar, wenn ihr äußeres und inneres Riemann-Maß übereinstimmen.

Ein weiteres wichtiges Kriterium für die Riemann-Meßbarkeit von Punktmengen ist

Satz 21.2.
Eine Menge M ist genau dann Riemann-meßbar, wenn der Rand ∂M das Riemann-Maß Null hat.
Beweis:
Für den Beweis verweisen wir auf [2, Seite 124].

Beispiel:

1. Es sei M beschränkte Teilmenge der $\mathbb{R}^n$ und H eine Hyperebene des $\mathbb{R}^n$. Dann hat $M \cap H$ das n-dimensionale Riemannsche Maß Null.

2. Es sei $g\colon \mathbb{R}^n \to \mathbb{R}^n$ stetig differenzierbar und $M \subset \mathbb{R}^n$ habe das n-dimensionale Riemannsche Maß Null. Dann hat $g(M)$ das n-dimensionale Riemannsche Maß Null.
   Beweis: siehe [2, Seite 124].

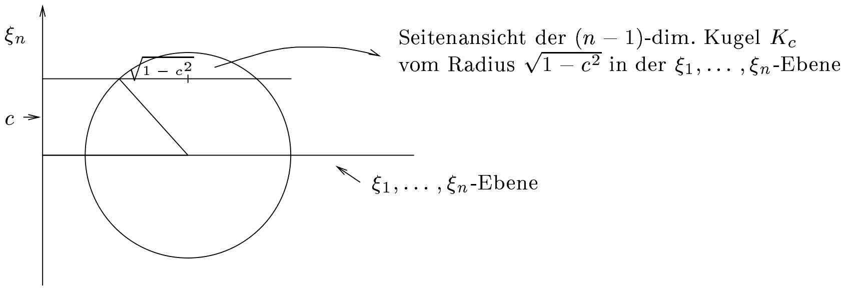
Folgerung:
Die $(n-1)$-dimensionale Einheitssphäre $\{x \in \mathbb{R}^n \,|\, ||x|| = 1\}$ hat das n-dimensionale Riemannsche Maß Null.

Folgerung:
Die n-dimensionale Einheitskugel ist Riemann-meßbar.

Als Beispiel wollen wir den n-dimensionalen Riemannschen Inhalt der n-dimensionalen Einheitskugel berechnen. Hierzu benutzen wir eine Folgerung aus dem Satz über die Zurückführung des n-dimensionalen Riemannschen Integrals auf Mehrfachintegrale, die als *„Prinzip von Cavalieri“* bezeichnet wird:

Satz 21.3.
Es sei $M \subset \mathbb{R}^n$ Riemann-meßbar und M liege ganz zwischen den beiden Hyperebenen $\{x = (\xi_1, \dots, \xi_n) \in \mathbb{R}^n | \xi_n = a\}$ und $\{x = (\xi_1, \dots, \xi_n) \in \mathbb{R}^n | \xi_n = b\}$. Für jedes $c \in [a, b]$ besitze der Durchschnitt von M mit der Hyperebene $\{x = (\xi_1, \dots, \xi_n) \in \mathbb{R}^n | \xi_n = c\}$ den $(n-1)$-dimensionalen Inhalt $v(c)$. Es existiere $\int_a^b v(c)\, dc$. Dann gilt für den n-dimensionalen Inhalt von M $|M| = \int_a^b v(c)\, dc$.

Wir wenden dies auf die Berechnung des n-dimensionalen Riemann-Inhalts der n-dimensionalen Einheitskugel an.



Wie wir später bei der Besprechung der Substitutionsregel für n-dimensionale Integrale sehen werden, ist der k-dimensionale Inhalt einer k-dimensionalen Kugel vom Radius R gleich $R^k S_k$, wenn S_k der Inhalt der k-dimensionalen Einheitskugel ist. Auf unser Problem angewandt, bedeutet dies, daß in Satz 21.3.

$$v(c) = (\sqrt{1-c^2})^{n-1} S_{n-1}$$

und

$$S_n = \int_{-1}^{1} (\sqrt{1-c^2})^{n-1} S_{n-1}\, dc.$$

Nach der Substitution $c = \sin\xi$ folgt

$$S_n = S_{n-1} \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \cos^n(\xi)\, d\xi$$

und mit Hilfe partieller Integration

$$\frac{S_n}{S_{n-1}} = \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \cos^n(\xi)\, d\xi = \frac{n-1}{n} \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \cos^{n-2}(\xi)\, d\xi = \frac{n-1}{n} \frac{S_{n-2}}{S_{n-3}}$$

oder

$$S_n = \frac{n-1}{n} \frac{S_{n-1} S_{n-2}}{S_{n-3}}.$$

Wir beachten, daß $S_1 = 2$ gilt. Daraus folgt

$$S_2 = S_1 \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \cos^2(\xi)\, d\xi = 2 \cdot \frac{\pi}{2} = \pi.$$

Für S_3 ergibt sich

$$S_3 = \frac{3-1}{3} \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \cos^2(\xi)\, d\xi = \pi \cdot \frac{2}{3} \cdot 2 = \frac{4\pi}{3}.$$

$$S_4 = \frac{4-1}{4}\frac{S_3 S_2}{S_1} = \frac{3}{4} \cdot \frac{4\pi}{3} \cdot \pi \cdot \frac{1}{2} = \frac{\pi^2}{2}.$$

Der 4-dimensionale Inhalt der vierdimensionalen Einheitskugel ist also $\frac{\pi^2}{2}$. Mit der obigen Rekursionsformel läßt sich leicht S_5, S_6 etc. berechnen.

Kapitel 3

Gewöhnliche Differentialgleichungen

3.1 Einleitung und Beispiele

Die Gleichung

$$F(t, y(t), y'(t), \ldots, y^{(m)}(t)) = 0 \ , \ t \in (\alpha, \beta)$$

mit einem Intervall $(\alpha, \beta) \subset \mathbb{R}$ und Funktionen

$$F\colon (\alpha, \beta) \times \mathbb{R}^{m+1} \to \mathbb{R}, y\colon (\alpha, \beta) \to \mathbb{R}$$

heißt allgemeine, implitize, (skalare) *gewöhnliche Differentialgleichung* m-ter Ordnung.

Entsprechend redet man von einem *System gewöhnlicher Differentialgleichungen*, wenn $y\colon (\alpha, \beta) \to \mathbb{R}^n$ und $F\colon (\alpha, \beta) \times \mathbb{R}^{n(m+1)} \to \mathbb{R}^n$.

Im Gegensatz hierzu nennt man die Gleichung

$$G(x, u(x), \nabla u(x), \ldots, \nabla^m u(x)) = 0 \ , \ x \in \Omega \subset \mathbb{R}^d,$$

eine *partielle Differentialgleichung.*

Gewöhnliche Differentialgleichungen und Systeme *m-ter Ordnung* lassen sich durch Erhöhung der Zahl der Variablen auf ein *System erster Ordnung* zurückführen.

Beispiel:
$F(., y, y', y'') = 0$. Setze $y_1 := y, y_1' := y_2$.
Damit erhalten wir das folgende System erster Ordnung:

$$\begin{pmatrix} y_1' - y_2 \\ F(., y_1, y_1', y_2') \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Hat die gewöhnliche DGL die Gstalt $y' = f(.,y)$, allgemeiner $y^{(m)} = f(.,y,\dots,y^{(m-1)})$, so spricht man von einer *expliziten gewöhnlichen Differentialgleichung.*

Beispiel:

1. Lösungen der skalaren Differentialgleichung $y' = 0$ ist $y = c\ ,\ c \in \mathbb{R}$ beliebig.

2. Wir betrachten eine weitere skalare Differentialgleichung: $y' = y$. Lösungen sind $y = ce^t\ ,\ c \in \mathbb{R}$ beliebig.

3. $y' = \mathbf{A}y$, wobei $\mathbf{A}$ eine $n \times n$-Matrix mit konstanten Koeffizienten ist. Lösungen sind $y(t) = ce^{\mathbf{A}t}\ ,\ c \in \mathbb{R}$ beliebig.
   Dabei ist $e^{\mathbf{B}}$ für eine Matrix $\mathbf{B}$ über die Reihenentwicklung erklärt: $e^{\mathbf{B}} = I + \mathbf{B} + \frac{1}{2}\mathbf{B}^2 + \frac{1}{3!}\mathbf{B}^3 + \cdots$.

Beobachtung:
Die Lösung ist nicht eindeutig, im allgemeinen hat eine Differentialgleichung eine *Lösungsschar.* Man hat einen Parameter c frei, der es erlaubt, die gesuchte Funktion y an einer Stelle t_0 *vorzuschreiben.* Man spricht dann von einem *Anfangswertproblem* (AWP)

Nicht eindeutig lösbares AWP:
$y' = \sqrt{y}\ ,\ y(0) = 0$. Lösungen sind $y(t) = \frac{1}{4}t^2$ und $y(t) \equiv 0$.

Singularität in endlicher Zeit:
$y' = y^2\ ,\ y(0) = 1$. Die Lösung dieses Anfangswertproblems ist $y(t) = \frac{1}{1-t}$, welche für $t \to 1$ eine Singularität entwickelt.

3.2 Elementare Lösungsmethoden für gewöhnliche Differentialgleichungen

3.2.1 Variation der Konstanten

Wir betrachten die skalare Differentialgleichung

$$y' = a(.)y + b(.)$$

und wählen den Ansatz

$$y(t) = c(t)w(t) \quad \text{mit} \qquad w' = aw \quad \Rightarrow \quad w = \text{const.} \cdot e^{\left(\int_{t_0}^t a(s)\,ds\right)}.$$

Nach der Produktregel erhalten wir, wenn wir unser speziell gewähltes y in die Differentialgleichung einsetzen,

$$y' = c'w + w'c = c'w + acw = c'w + ay.$$

Damit y eine Lösung ist, muß also gelten

$$c'w = b.$$

Für c erhalten wir damit

$$c = \int_{t_0}^{t} w^{-1} b \, ds,$$

und eine partiluläre Lösung ist dann

$$y(t) = e^{(\int_{t_0}^{t} a(s)\, ds)} \int_{t_0}^{t} e^{(-\int_{t_0}^{\xi} a(s)\, ds)} b(\xi)\, d\xi.$$

Die *allgemeine* Lösung von $y' = ay + b$ ergibt sich als Summe einer partiklären Lösung der inhomogenen Gleichung $y' = ay + b$ und der allgemeinen Lösung der homogenen Gleichung $y' = ay$.

3.2.2 Bernoullische Differentialgleichung

Wir betrachten Differentialgleichungen der Form

$$y' = ay^{\alpha} + b.$$

Setze $y = z^{\frac{1}{1-\alpha}}$. Eine einfache Rechnung ergibt

$$z' = a + bz.$$

Daraus ergibt sich z mit Hilfe der „Variation der Konstanten“.

3.2.3 Riccatische Differentialgleichung

$$y' = ay^2 + by + c$$

Diese Gleichung läßt sich i.a. *nicht* durch „elementare Integrationen“ lösen, d.h. eine Lösungsdarstellung unter Verwendung der elementaren Funktionen und Integrationen und Umkehrfunktionen gibt es nicht.
Die Gleichung $y' = ay^2 + by + c$ läßt sich durch Substitution $y = -a(\log z)' = (-a)\frac{z'}{z}$ auf die lineare, homogene Gleichung 2. Ordnung

$$-z'' = z'a' + bz$$

transformieren.

3.2.4 Differentialgleichungen mit getrennten Veränderlichen

$$y' = f(t)g(y).$$

Sei

$$G' = \frac{1}{g}. \quad \Rightarrow \quad \frac{d}{dt}(G(y)) = f(t), \quad G(y) = c + \int_{t_0}^{t} f(s)\,ds.$$

y ergibt sich durch Umkehren von G.

Beispiel:

$$\begin{aligned} & y' = t \cdot e^y \\ \Rightarrow \quad & (e^{-y})' = -\left(\frac{t^2}{2}\right)' \\ \Rightarrow \quad & e^{-y} = -\frac{t^2}{2} + c \\ \Rightarrow \quad & y = -\log\left(c - \frac{t^2}{2}\right). \end{aligned}$$

(Für $c > 0$ ergibt sich in endlicher Zeit eine Sigularität.)

3.2.5 Homogene Differentialgleichung

$$y' = f\left(\frac{y}{t}\right).$$

Setze

$$z = \frac{y}{t}.$$

$$\Rightarrow \quad z' = \frac{1}{t}(f(z) - z),$$

d.h. es liegt eine Differentialgleichung mit getrennten Veränderlichen vor.

3.3 Der Existenz- und Eindeutigkeitssatz von Picard-Lindelöf

Sei $f\colon [t_0, T] \times \mathbb{R}^n \to \mathbb{R}^n$ stetig und $y_0 \in \mathbb{R}^n$.
Wir betrachten das Anfangswertproblem $y' = f(.,y)$ in $]t_0, t_0 + \delta[\ ,\ y(t_0) = y_0$.
Das Anfangswertproblem ist äquivalent zur Integralgleichung

$$y(t) = y_0 + \int_{t_0}^{t} f(s, y(s))\,ds. \tag{3.1}$$

Es sei $V = C([t_0, t_0+\delta]; \mathbb{R}^n)$ (= Raum der stetigen Funktionen mit Werten in $\mathbb{R}^n$). Die Gleichung (3.1) läßt sich als *Fixpunktgleichung* $\mathfrak{A}y = y$ mit dem Operator $\mathfrak{A}\colon V \to V$, definiert durch

$$(\mathfrak{A}z)(t) = y_0 + \int_{t_0}^{t} f(s, z(s))\, ds$$

schreiben. (Ist f nur auf $[t_0, T] \times U(y_0)$ definiert, ist der Definitionsbereich von $\mathfrak{A}$ entsprechend kleiner.)

Um die Existenz eines Fixpunktes von $\mathfrak{A}$ und damit die Lösbarkeit von (3.1) zu gewährleisten, benötigen wir den *Banachschen Fixpunktsatz* .

Definition:
Ein Banachraum *ist ein vollständiger, normierter linearer Raum, in unserem Fall der obige Raum V mit der Norm*

$$||z|| = \sup_{t \in [t_0, t_0+\delta]} |z(t)|.$$

Banachscher Fixpunktsatz
Es sei $\mathfrak{A}\colon V \to V$ eine kontrahierende Abbildung eines Banachraumes V in sich, d.h. es gibt ein $q < 1$ mit

$$||\mathfrak{A}x - \mathfrak{A}y|| \leq q\, ||x - y||\, \forall x, y \in V.$$

Dann besitzt $\mathfrak{A}$ einen eindeutigen Fixpunkt.

Beweis:
Die Eindeutigkeit ist klar:
Seien x und y Fixpunkte von $\mathfrak{A}$, d.h. $\mathfrak{A}x = x, \quad \mathfrak{A}y = y$. Dann gilt nach Voraussetzung $||x - y|| = ||\mathfrak{A}x - \mathfrak{A}y|| \leq q\, ||x - y||$. Da $q < 1$ ist, folgt $||x - y|| = 0$, also $x = y$.
Für den Beweis der Existenz wählt man ein beliebiges Startelement x_0 und beweist die Konvergenz $\mathfrak{A}^m x_0 \to x^*$. Wegen

$$x^* \leftarrow \mathfrak{A}^m x_0 = \mathfrak{A}(\mathfrak{A}^{m-1} x_0) \to \mathfrak{A}(\lim_{m\to\infty} \mathfrak{A}^{m-1} x_0) = \mathfrak{A}x^*$$

(beachte, daß aus der Kontraktionsbedingung die Stetigkeit von $\mathfrak{A}$ folgt), ergibt sich $\mathfrak{A}x^* = x^*$.
(Erinnerung: $x_j \to x \quad \Leftrightarrow \quad ||x_j - x|| \to 0\ (j \to \infty)$).
Nachweis der Konvergenz von $\mathfrak{A}^m x_0$: Man zeigt, daß $(\mathfrak{A}^m x_0,\ m \in \mathbb{N})$ eine *Cauchyfolge* ist. Aus der Vollständigkeit von V folgt dann die Existenz des Limes.

$$\begin{aligned} \left|\left|\mathfrak{A}^m x_0 - \mathfrak{A}^k x_0\right|\right| &\leq \left|\left|\mathfrak{A}^m x_0 - \mathfrak{A}^{m-1} x_0\right|\right| + \left|\left|\mathfrak{A}^{m-1} x_0 - \mathfrak{A}^{m-2} x_0\right|\right| + \cdots + \left|\left|\mathfrak{A}^{k+1} x_0 - \mathfrak{A}^k x_0\right|\right| \\ &\leq q^{m-1}\, ||\mathfrak{A}x_0 - x_0|| + q^{m-2}\, ||\mathfrak{A}x_0 - x_0|| + \cdots + q^k\, ||\mathfrak{A}x_0 - x_0|| \quad (m \geq k). \end{aligned}$$

Dies führt auf den *Reihenrest* einer konvergenten geometrischen Reihe.

$$\Rightarrow \qquad \left|\left|\mathfrak{A}^m x_0 - \mathfrak{A}^k x_0\right|\right| \leq \varepsilon \qquad \text{für} \quad m \geq k \geq k_0(\varepsilon).$$

□

Der Satz von Banach gilt auch unter der schwächeren Voraussetzung, daß die Abbildung $\mathfrak{A}$ eine „Kugel" $K = \{x \in V \,|\, ||x_0 - x|| \leq R\}$ in sich abbildet und auf dieser kontrahierend ist. Es gibt dann einen Fixpunkt y^*, der in K liegt („lokale Version des Banachschen Fixpunktsatzes").

Anwendung des Banachschen Fixpunktsatzes auf Anfangswertprobleme

$$y' = f(.,y) \ , \quad y(t_0) = y_0.$$

Hierbei sei $y_0 \in \mathbb{R}^n$, $f \in C([t_0,T] \times \mathbb{R}^n; \mathbb{R}^n)$. Außerdem genüge f einer *Lipschitzbedingung* in einer Umgebung von y_0, d.h. es existiere ein $L \in \mathbb{R}$ mit $|f(t,\xi) - f(t,\eta)| \leq L|\xi - \eta| \quad \forall \xi, \eta \in U(y_0)$. Man sagt auch „lokale Lipschitzbedingung", da $\xi, \eta \in U(y_0)$, und spricht von globaler Lipschitzbedingung, wenn obige Ungleichung für alle ξ, η gilt.

Umformung des AWP in eine Integralgleichung:

$$y(t) = y_0 + \int_{t_0}^{t} f(s, y(s))\, ds.$$

Mit $(\mathfrak{A}z)(t) = y_0 + \int_{t_0}^{t} f(s, z(s))\, ds$ schreibt sich das AWP als *Fixpunktproblem* für die Abbildung $\mathfrak{A} \colon C([t_0, T'); \mathbb{R}^n) \to C([t_0, T'); \mathbb{R}^n)$. U.U. werden wir $T' = t_0 + \delta < T$ wählen.

Lemma:
Es gelte die obige Lipschitzbedingung für $f \in C$. $\Rightarrow \quad \exists \delta > 0$ *und ein* $q \in]0,1[$, *so daß* $||\mathfrak{A}z - \mathfrak{A}w|| \leq q\, ||z - w|| \quad \forall z, w \in K \subset C([t_0, t_0 + \delta])$, $\quad K = \{z \in C([t_0, t_0 + \delta]) | z(t) \in U_\varepsilon(y_0)\}$, $\quad U_\varepsilon(y_0) \subset U(y_0)$. *Hierbei ist* $||z|| = \max\limits_{t \in [t_0, t_0 + \delta]} |z(t)|$.

Beweis:

$$\begin{aligned} |\mathfrak{A}z(t) - \mathfrak{A}w(t)| &= \left| \int_{t_0}^{t} (f(s, z(s)) - f(s, w(s)))\, ds \right| \\ &\leq \left| \int_{t_0}^{t} L|z(s) - w(s)|\, ds \right| \qquad \text{nach der Lipschitzbedingung} \\ &\leq L\delta\, ||z - w||\,. \end{aligned}$$

Mit $\delta = \frac{q}{L}$ und Übergang zum Supremum bzgl. t ergibt sich die Kontraktionsbedingung. □

Lemma:
Es gelten dieselben Voraussetzungen wie oben. Dann existiert $\delta > 0$, so daß die obige Abbildung $\mathfrak{A}$ die Kugel $K = \{z \in C([t_0, t_0 + \delta]) | z(t) \in U_\varepsilon(y_0)\}$ in sich abbildet.

Beweis:

$$|\mathfrak{A}z(t) - y_0| = \left| \int_{t_0}^{t} f(s, z(s))\, ds \right| \leq \delta \sup_{[t_0, t_0+\delta] \times U_\varepsilon} f < \varepsilon \qquad \text{für} \quad \delta = \delta(\varepsilon, f).$$

□

Damit sind die Voraussetzungen der lokalen Version des Banachschen Fixpunktsatzes gesichert und es gilt der

Existenz- und Eindeutigkeitssatz (nach Picard-Lindelöf)
Seien f und y_0 wie oben.
Dann existiert ein $\delta > 0$, so daß das AWP

$$y' = f(., y) \;, \quad y(t_0) = y_0,$$

im Intervall $[t_0, t_0 + \delta]$ eindeutig lösbar ist.

3.4 Fortsetzbarkeit von Lösungen

Der Existenz- und Eideutigkeitssatz für Anfangswertprobleme liefert nur die Lösbarkeit in einem *kleinen* Zeitintervall. Bei vielen Problemen — z.B. der Planetenbewegung — wünscht man sich die globale Lösbarkeit, d.h. die Lösbarkeit in dem unendlichen Zeitintervall, so daß keine Singularität („Katastrophe") nach endlicher Zeit eintrifft. – Man könnte versuchen, folgendermaßen vorzugehen: Man löse das Anfangswertproblem zur Differentialgleichung $y' = f(., y)$ mit der Anfangsbedingung $y(t_0) = y_0$ und erhält nach dem Satz von Picard-Lindelöf ein $\delta_0 > 0$, so daß das AWP in $[t_0, t_0 + \delta_0]$ gelöst wird. Sodann löse man das neue AWP mit der Anfangsbedingung $y_{neu}(t_0 + \delta_0) = y_{alt}(t_0 + \delta_0)$ und erhält eine Lösung der Differentialgleichung in $[t_0 + \delta_0, t_0 + \delta_0 + \delta_1]$. Man „stückelt" die beiden Lösungen y_{alt} und y_{neu} zusammen und erhält eine Lösung auf $[t_0, t_0 + \delta_0 + \delta_1]$. Durch Wiederholung dieser Methode erhält man eine Lösung auf $[t_0, t_0 + \delta_0 + \delta_1 + \delta_2 + \cdots + \delta_N]$. Leider konvergiert $\sum_{j=0}^{\infty} \delta_j$ zumeist *nicht* gegen unendlich, so daß man mit dieser Methode im allgemeinen keine globale Lösung erhält. Eine Ausnahme liegt vor, wenn man $\delta_j = \delta > 0 \;, \quad j = 1, 2, \ldots,$ wählen kann. Dies ist möglich, wenn eine *globale Lipschitzbedingung* vorausgesetzt wird, d.h. es gilt $|f(t, \xi) - f(t, \eta)| \leq L|\xi - \eta|$ für alle $\xi, \eta \in \mathbb{R}^n \;, \; t \geq t_0$. Die globale Lipschitzbedingung ist leider sehr einschränkend und impliziert lineares Wachstum von $f(t,\eta)$ bezüglich η. Wenigstens gilt die globale Lipschitzbedingung im linearen Fall $f(t,\eta) = \mathbf{A}(t)\eta + b(t)$ mit einer beschränkten $n \times n$-Matrizenfunktion $\mathbf{A}$.

Satz:

Sei $\mathbf{A}(t) = (a_{ik}(t))_{i,k=1}^n$ und $b(t) \in \mathbb{R}^n$ stetig bzgl. t und $\mathbf{A}(t)$ für $t \to \infty$ beschränkt. Dann gibt es eine Lösung des Anfangswertproblems

$$y' = \mathbf{A}(.)y + b(.) \qquad , \quad y(t_0) = y_0, \qquad \textit{in} \quad [t_0, \infty[.$$

3.5 Lineare Differentialgleichungssysteme

Sei $\mathbf{A}(t) = (a_{ik}(t))_{i,k=1}^n$ eine stetige Matrizenfunktion, $b \in C([t_0,T] \times \mathbb{R}^n)$.
Wir betrachten das „homogene System“

$$y' = \mathbf{A}(.)y.$$

Es ist klar, daß die Gesamtheit der Lösungen des homogenen Systems einen linearen Raum bildet. Es gilt sogar der

Satz:

Die Dimension des Lösungsraumes des homogenen Systems ist n.

Beweis:

Durch die Wahl von n linear unabhängigen Anfangsvektoren konstruiert man sich nach dem Existenz- und Exideutigkeitssatz mindestens n linear unabhängige Lösungen des homogenen Systems. Gäbe es mehr als n linear unabhängige Lösungen, so könnte man diese an der Stelle t_0 nichttrivial zu Null kombinieren. Diese Linearkombination von Lösungen müßte aber wegen des Eindeutigkeitssatz gleich der Lösung „identisch Null“ sein, d.h. es kann keine $n+1$ linear unabhängigen Lösungen geben.

Es ist klar, daß sich die allgemeine Lösung des inhomogenen Systems als Summe einer partikulären Lösung des inhomogenen Systems und der allgemeinen Lösung des homogenen Systems schreiben läßt:

Verfügt man über n linear unabhängige Lösungen des homogenen Systems, etwa $\varphi^{(1)}, \ldots, \varphi^{(n)}$, so lassen sich diese zu einer Matrix $\Phi = (\varphi^{(1)}, \ldots, \varphi^{(n)})$, der sogenannten *Wronski*-Matrix zusammenfassen. Jede Lösung des homogenen Systems hat dann die Gestalt $\Phi(t) \cdot c \;,\;\; c =$ konst. Vektor. Es gilt $\dot{\Phi} = \mathbf{A}\Phi$. Eine partikuläre Lösung des inhomogenen Systems erhält man mit Hilfe der Variation der Konstanten. Ansatz: $y(t) = \Phi(t)c(t) \quad \Rightarrow$

$$\dot{y} = \dot{\Phi}c + \Phi\dot{c} = \mathbf{A}\Phi c + \Phi\dot{c} = \mathbf{A}y + \Phi\dot{c}$$

Aus der Forderung $\Phi\dot{c} = b$ folgt $c(t) = \int\limits_{t_0}^{t} \Phi^{-1}(s)b(s)\,ds$.

Kapitel 4

Grundbegriffe der Vektoranalysis

4.1 Kurven im $\mathbb{R}^n$

Definition:
Es sei $[a,b] \subset \mathbb{R}^1$, $a < b$, *und* $p\colon [a,b] \to \mathbb{R}^n$ *eine stetige Funktion. Dann nennt man die Menge* $K = \{x \in \mathbb{R}^n \mid x = p(t),\ t \in [a,b]\}$ *ein* Kurvenstück *im* $\mathbb{R}^n$.

1. **Beispiel:**

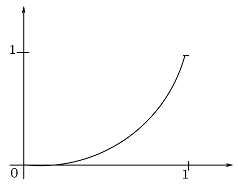
$$[a,b] = [0,1], \quad p(t) = \begin{pmatrix} \sin 2\pi t \\ \cos 2\pi t \end{pmatrix}.$$

Das Kurvenstück K ist dann der Rand des Einheitskreises.
Verschiedene Funktionen p können zu demselben Kurvenstück führen, z.B. ergibt

$$q(t) = \begin{pmatrix} \sin 4\pi t \\ \cos 4\pi t \end{pmatrix}$$

dasselbe Kurvenstück K.

2. **Beispiel:** $[a,b] = [0,1], \quad p(t) = \binom{t}{t^2}$.



Für mathematische Zwecke (etwa zur Definition des Kurvenintegrales) ist es wichtig, einen dem Begriff des Kurvenstückes ähnlichen Begriff (den der *Kurve*) zu haben, bei dem jedoch die intuitive dynamische Vorstellung in einem mathematisch präzisierten Sinne erhalten bleibt, die sich ergibt, wenn man die Variable t von a nach b „laufen" läßt und untersucht, wie sich die

Werte $p(t)$ verhalten.
„Läuft" man mit der Variablen t von 0 nach 1, so durchläuft

$$p(t) = \begin{pmatrix} \sin 2\pi t \\ \cos 2\pi t \end{pmatrix}$$

im Uhrzeigersinn den Einheitskreis, man beginnt mit $p(0) = \binom{0}{1}$ und endet dort. Betrachtet man stattdessen

$$q(t) = \begin{pmatrix} \sin 4\pi t \\ \cos 4\pi t \end{pmatrix},$$

so durchläuft man zweimal den Einheitskreis.
Schließlich durchläuft man mit der durch

$$\tilde{p}(t) = \begin{pmatrix} -\sin 2\pi t \\ \cos 2\pi t \end{pmatrix}$$

definierten Funktion $\tilde{p}$ den Einheitskreis entgegen dem Uhrzeigersinn.
Setzt man

$$\begin{aligned} p(t) &= \begin{pmatrix} t \\ t^2 \end{pmatrix}, \ t \in [0,1], \\ \tilde{p}(t) &= \begin{pmatrix} 2t \\ 4t^2 \end{pmatrix}, \ t \in \left[0, \frac{1}{2}\right], \\ \tilde{p}(t) &= \begin{pmatrix} 1 \\ 1 \end{pmatrix} \ \text{für } t \in \left[\frac{1}{2}, 1\right], \end{aligned}$$

so ergeben p und $\tilde{p}$ dasselbe Kurvenstück. Läuft man jedoch mit t von 0 bis 1, so hat man – betrachtet man $\tilde{p}$ – nach dem Punkt $\frac{1}{2}$ bereits das ganze Kurvenstück durchlaufen und ruht sich anschließend im Punkt $\binom{1}{1}$ aus.

Um einen Begriff zu haben, der diesen Effekt (mehrfaches Durchlaufen des Kurvenstückes, Durchlaufen des Einheitskreises mit oder entgegen dem Uhrzeigersinn, Verweilen in einem Punkt, Hin- und Zurücklaufen etc., etc.) mathematisch festhält, wählen wir folgende

Definition:
Eine (orientierte) Kurve *im* $\mathbb{R}^n$ *ist eine Äquivalenzklasse von stetigen Funktionen* $p\colon [a,b] \to \mathbb{R}^n$. *Zwei Funktionen* $p\colon [a,b] \to \mathbb{R}^n$ *und* $q\colon [c,d] \to \mathbb{R}^n$ *heißen hierbei äquivalent, wenn es eine stetige, streng monoton steigende Funktion* $\varphi\colon [a,b] \to [c,d]$ *gibt mit* $\varphi(a) = c$, $\varphi(b) = d$, *so daß* $q(\varphi(t)) = p(t)$ *für* $t \in [a,b]$ *gilt.*
Ist $p\colon [a,b] \to \mathbb{R}^n$ *ein Repräsentant einer solchen Klasse, so nennt man* p *eine* Parameterdarstellung *der Kurve und* $[a,b]$ *das zugehörige* Parameterintervall. *Den Punkt* $p(a)$ *nennt man Anfangspunkt, den Punkt* $p(b)$ *Endpunkt. Die Menge* $\{p(t) \mid t \in [a,b]\}$ *heißt der* Wertebereich *oder* Wertevorrat *der Kurve.*

Beispiel:
Seien
$$[a,b]=[0,\frac{\pi}{4}],\quad p(t)=\begin{pmatrix}\sin t\\ \cos t\end{pmatrix}\quad \text{und}\quad [c,d]=[0,1],\quad q(t)=\begin{pmatrix}\frac{1}{\sqrt{1+t^2}}\\ \frac{1}{\sqrt{1+t^2}}\end{pmatrix}.$$
Die beiden Funktionen $p\colon \left[0,\frac{\pi}{4}\right]\to\mathbb{R}^2$ und $q\colon[0,1]\to\mathbb{R}^2$ stellen dieselbe Kurve dar, als Funktion φ in der Definition kann man $\varphi(t)=\tan t$ nehmen.
Hingegen werden durch
$$p(t)=\begin{pmatrix}\sin 2\pi t\\ \cos 2\pi t\end{pmatrix},\quad \tilde{p}(t)=\begin{pmatrix}-\sin 2\pi t\\ \cos 2\pi t\end{pmatrix},\quad g(t)=\begin{pmatrix}\sin 4\pi t\\ \cos 4\pi t\end{pmatrix}$$
verschiedene Kurven definiert.

Ältere Bücher verwenden die Begriffe Kurve und Kurvenstück synonym. Statt „orientierte Kurve“ sagt man meistens „Kurve“.

Eine Kurve heißt *geschlossen*, wenn Anfangs- und Endpunkt übereinstimmen.

Der angegebene Kurvenbegriff ist noch sehr allgemein und schließt u.a. mathematische Objekte ein, die von der anschaulichen Vorstellung einer Kurve erheblich abweichen. Z.B. konstruierte der Mathematiker Peano eine Kurve, deren Wertebereich in einem Quadrat dicht liegt. So etwas vermeidet man, wenn man den Begriff der *Jordankurve* einführt.

Definition:
Eine Jordankurve *ist eine Kurve, die eine Parameterdarstellung besitzt, die eine topologische Abbildung (d.h. umkehrbar eindeutige, in beiden Richtungen stetige surjektive Abbildung) des Parameterintervalls auf den Wertebereich der Kurve definiert. Dies gilt dann für alle Parameterdarstellungen der Kurve.*

Eine geschlossene Jordankurve ist eine geschlossene Kurve mit der Eigenschaft:
Es gibt eine Parameterdarstellung $p\colon[a,b]\to\mathbb{R}^n$, so daß p das Parameterintervall $[a,b[$ umkehrbar eindeutig auf den Wertebereich W der Kurve abbildet und die Umkehrabbildungen $p^{-1}\colon W\to[a,b[$ und $p^{-1}\colon W\to]a,b]$ stetig sind.

In der Ebene gilt der bekannte

Jordansche Kurvensatz:
Es sei Γ ein geschlossenes Jordansches Kurvenstück im $\mathbb{R}^2$. Dann unterteilt Γ den $\mathbb{R}^2$ in zwei offene, zusammenhängende Punktmengen mit Rand Γ, d.h. es existieren offene, zusammenhängende Mengen $\Omega_1,\Omega_2\subset\mathbb{R}^2$ mit

i) $\Omega_1\cap\Omega_2=\varnothing$, *ii)* $\partial\Omega_1 = \partial\Omega_2 = \Gamma$,

iii) $\Omega_1 \cup \Gamma \cup \Omega_2 = \mathbb{R}^2$.

Der Satz ist anschaulich evident, der Beweis jedoch schwierig.

Von der Anwendung her stellt sich die Frage, was man unter der Länge einer Kurve versteht und wie man sie eventuell berechnet.
Es sei Γ eine Kurve mit Parameterdarstellung $p\colon [a,b] \to \mathbb{R}^n$. Es sei $a = t_0 < t_1 < \ldots < t_N = b$ eine Zerlegung $\mathfrak{Z}$ des Parameterintervalls. Unter einem Γ eingeschriebenen Polygonzug bzgl. $\mathfrak{Z}$ verstehen wir den Polygonzug mit den Ecken $p(t_i)$, $i = 0, \ldots, N$, d.h. die Menge

$$\mathfrak{p}_{\mathfrak{Z}} = \left\{ \tau p(t_i) + (1-\tau) p(t_{i-1}) \;\middle|\; 0 \le \tau \le 1,\, i = 1, \ldots, N \right\}.$$

Unter der *Länge des Polygonzugs* $\mathfrak{p}_{\mathfrak{Z}}$ verstehen wir die Zahl

$$|\mathfrak{p}_{\mathfrak{Z}}| = \sum_{i=1}^{N} |p(t_i) - p(t_{i-1})|.$$

Hierbei ist

$$|a| = \left(\sum_{j=1}^{n} |a_j|^2 \right)^{\frac{1}{2}}, \; a = (a_1, \ldots, a_n).$$

Definition:
Eine Kurve Γ heißt rektifizierbar*, wenn das Supremum S der Längen aller einbeschriebenen Polygonzüge beschränkt ist.*
Man nennt $|\Gamma| := S$ die Länge *der Kurve. In Zeichen:*
Mit der Parameterdarstellung $p\colon [a,b] \to \mathbb{R}^n$ ist

$$|\Gamma| = \sup\left\{ |\mathfrak{p}_{\mathfrak{Z}}| \;\middle|\; \mathfrak{Z} \text{ ist eine Zerlegung von } [a,b],\ \mathfrak{p}_{\mathfrak{Z}} \text{ siehe oben} \right\}.$$

Es ist leicht zu sehen, daß diese Definition von der Parameterdarstellung unabhängig ist und der intuitiven Vorstellung des Begriffs der Länge gerecht wird.
Man kann für (stetig) differenzierbare Kurven eine Formel für die Länge angeben. Eine Kurve heißt hierbei (stetig) differenzierbar, wenn sie eine (stetig) differenzierbare Parameterdarstellung besitzt.

Satz:
Es sei Γ eine stetig differenzierbare Kurve mit Parameterdarstellung $p\colon [a,b] \to \mathbb{R}^n$, p stetig

differenzierbar. Dann ist Γ rektifizierbar , und es gilt:

$$|\Gamma| = \int_a^b |\dot{p}(t)|\, dt.$$

Hierbei ist $\dot{p} = \frac{\partial p}{\partial t}$.

Beweis:

Wegen der Stetigkeit von p gilt nach Definition des Riemannschen Integrals bei vorgegebenem ε für alle Zerlegungen des Parameterintervalls $[a,b]$ mit Feinheit $< \delta(\varepsilon)$ die Ungleichung

$$\Big|\int_a^b |\dot{p}(t)|\, dt \;-\; \sum_{i=1}^N |\dot{p}(t_i)|(t_i - t_{i-1})\Big| < \varepsilon,$$

wenn $\mathfrak{Z} = \{t_0, \dots, t_N\}$, $a = t_0 < t_1 < \dots < t_n = b$.

Nach Definition der Kurvenlänge gibt es eine Zerlegung $\mathfrak{Z}_0 = \{\tau_0, \dots, \tau_M\}$ von $[a,b]$ mit

$$0 \le |\Gamma| - \sum_{i=1}^M |p(\tau_i) - p(\tau_{i-1})| < \varepsilon.$$

Es sei $\mathfrak{Z}' = \{\tilde{\tau}_0, \dots, \tilde{\tau}_S\}$ eine gemeinsame Verfeinerung von $\mathfrak{Z}_0$ und einer Zerlegung mit Feinheit $\delta(\varepsilon)$. Es gilt dann erst recht

$$0 \le |\Gamma| - \sum_{i=1}^S |p(\tilde{\tau}_i) - p(\tilde{\tau}_{i-1})| < \varepsilon.$$

Es gilt nach dem Mittelwertsatz

$$\begin{aligned} |p(\tilde{\tau}_{i+1}) - p(\tilde{\tau}_i)| &= \left(\sum_{j=0}^n |p_j(\tilde{\tau}_{i+1}) - p_j(\tilde{\tau}_i)|^2\right)^{\frac{1}{2}} \\ &= (\tilde{\tau}_{i+1} - \tilde{\tau}_i)\left(\sum_{j=1}^n |\dot{p}_j(\tilde{\tau}_i^{(j)})|^2\right)^{\frac{1}{2}} \end{aligned}$$

mit Zahlen $\tilde{\tau}_i^{(j)} \in [\tilde{\tau}_i, \tilde{\tau}_{i+1}]$. Wählt man $\delta_0 < \delta(\varepsilon)$ genügend klein, so gilt wegen der Stetigkeit von $|\dot{p}|$

$$|p(\tilde{\tau}_{i+1}) - p(\tilde{\tau}_i)| = (\tilde{\tau}_{i+1} - \tilde{\tau}_i)\left(\sum_{j=1}^n |\dot{p}_j(\tilde{\tau}_i)|^2\right)^{\frac{1}{2}} + E_{ij}(\tilde{\tau}_{i+1} - \tilde{\tau}_i),$$

mit $|E_{ij}| < \varepsilon$.

Daraus folgt

$$\begin{aligned} \left|\sum_{i=0}^{S-1} |p(\tilde{\tau}_{i+1}) - p(\tilde{\tau}_i)| - \sum_{i=0}^{S-1} (\tilde{\tau}_{i+1} - \tilde{\tau}_i)\, |\dot{p}(\tilde{\tau}_i)|\right| &< \varepsilon \sum_{i=0}^{S-1} (\tilde{\tau}_{i+1} - \tilde{\tau}_i) \\ &= \varepsilon(b-a). \end{aligned}$$

Zusammen mit der beim Beginn des Beweises aufgestellten Ungleichung gilt damit

$$\Big|\int\limits_a^b |\dot p(t)|\,dt - \sum_{i=0}^{S-1} |p(\tilde\tau_{i+1}) - p(\tilde\tau_i)|\Big| < \varepsilon(b-a) + \varepsilon$$

und somit

$$\Big|\int\limits_a^b |\dot p(t)|\,dt - |\Gamma|\Big| < \varepsilon + \varepsilon(b-a) + \varepsilon.$$

Der Grenzübergang für $\varepsilon \to 0$ ergibt die Behauptung. □

Beispiel:
Eine Ellipse E wird gegeben durch die Parameterdarstellung

$$p\colon [0,2\pi] \to \mathbb{R}^2, \quad p(t) = \begin{pmatrix} a\sin t \\ b\cos t \end{pmatrix}.$$

Es gilt

$$\begin{aligned} \dot p(t) &= \begin{pmatrix} a\cos t \\ -b\sin t \end{pmatrix}, \\ |\dot p(t)| &= \sqrt{a^2\cos^2 t + b^2 \sin^2 t}, \\ |E| &= \int\limits_0^{2\pi} \sqrt{a^2\cos^2 t + b^2\sin^2 t}\,dt. \end{aligned}$$

Im Fall des Kreises $a = b = r$ ergibt sich für den Kreisumfang $E_{a=b=r} = 2\pi r$. Falls $a \neq b$ ist, läßt sich das Integral nicht mehr elementar auswerten, nach der Substitution $\sin t = u$ wird man auf ein sogenanntes *elliptisches Integral* geführt.

Definition:
Eine rektifizierbare Jordansche Kurve nennt man einen Weg. *Allgemeiner definieren die meisten Autoren einen* Weg *als Kurve, die sich durch Aneinanderfügen endlich vieler Jordanscher Kurven ergibt. Dem wollen wir uns auch hier anschließen.*
Mathematisch bedeutet dies: Ist Γ eine Kurve mit der Parameterdarstellung $p\colon [a,b] \to \mathbb{R}^n$, so nennt man Γ einen Weg, wenn Zahlen $t_0, t_1, \ldots, t_N$ mit $a = t_0 < t1 < \ldots < t_N = b$ existieren, so daß die durch die Restriktion von p auf $[t_i, t_{i+1}]$ erklärten Kurven rektifizierbare jordansche Kurven sind, d.h. topologische Abbildungen von $[t_i, t_{i+1}]$ auf $\{p(t) \mid t \in [t_i, t_{i+1}]\}$ sind.
Ein Weg heißt geschlossen, *wenn Anfangspunkt und Endpunkt übereinstimmen.*

Definition:
Eine Punktmenge M heißt (weg-)zusammenhängend, *wenn sich je zwei Punkte $P_1, P_2 \in M$ durch ein in M liegendes Wegstück verbinden lassen, d.h. es gibt einen Weg Γ mit Parameterdarstellung $p\colon [a,b] \to \mathbb{R}^n$, so daß $p(a) = P_1, \quad p(b) = P_2, \quad p(t) \in M$ für $t \in [a,b]$.*

Beispiel:
$\{x \in \mathbb{R}^n \mid x = (x_1, \dots, x_n),\, x_i \neq 0$ für alle $i = 1, \dots, n\}$ ist nicht zusammenhängend.
$\{x \in \mathbb{R}^n \mid |x| = 1\}$ ist zusammenhängend, ebenso $\mathbb{R}^n$.

Nimmt man aus einem Kreis den Mittelpunkt oder einen kleineren konzentrischen Kreis weg, so ist die neue Punktmenge zwar zusammenhängend, hat jedoch anschaulich ein Loch. Diese anschauliche Tatsache läßt sich mit Hilfe des Begriffs des *einfachen Zusammenhangs* einer Punktmenge mathematisch ausdrücken, was im folgenden erläutert wird.
Es seien $I = [\alpha, \beta] \subset \mathbb{R}$ und $[a, b] \subset \mathbb{R}$ Intervalle und $\phi\colon I \times [a, b] \to \mathbb{R}^n$ stetig. Für jedes $s \in I$ definiere $\phi(s, \cdot)\colon [a, b] \to \mathbb{R}^n$ eine Kurve im $\mathbb{R}^n$. Man sagt dann, daß ϕ eine *stetige Deformationsschar* von Kurven definiert und die durch $\phi(\alpha, \cdot)\colon [a, b] \to \mathbb{R}^n$ definierte Kurve in die durch $\phi(\beta, \cdot)\colon [a, b] \to \mathbb{R}^n$ definierte Kurve stetig deformiert wird.

Man sagt nun, eine Kurve Γ lasse sich *auf einen Punkt zusammenziehen*, wenn es eine stetige Deformationsschar gibt, die Γ in eine Kurve mit einpunktigem Wertebereich stetig deformiert. Mit diesen Begriffen ergibt sich die folgende

Definition:
Eine Punktmenge M heißt einfach (weg-)zusammenhängend, *wenn M zusammenhängend ist und sich jeder geschlossene Weg in M auf einen Punkt zusammenziehen läßt, wobei die Kurven der zugehörigen Deformationsschar Wege sein sollen.*

Beispiele:

$\mathbb{R}^n$	ist einfach zusammenhängend.
$\left\{x \in \mathbb{R}^n \,\middle|\, \lvert x\rvert \leq 1\right\}$	ist einfach zusammenhängend.
$\left\{x \in \mathbb{R}^2 \,\middle|\, \lvert x\rvert = 1\right\}$	ist nicht einfach zusammenhängend.
$\left\{x \in \mathbb{R}^3 \,\middle|\, \lvert x\rvert = 1\right\}$	ist einfach zusammenhängend.

4.2 Kurvenintegrale

Für viele physikalische Anwendungen ist es wichtig, das Integral über eine Funktion längs einer Kurve zu bilden. Will man z.B. eine Ladung in einem elektromagnetischen Feld von einem Punkt zu einem anderen längs einer Kurve verschieben, so muß man (evtl. negative) Arbeit aufwenden. Um diese Arbeit zu berechnen, approximiert man die Kurve durch einen Polygonzug und summiert Arbeitsanteile, die sich für die Verschiebung der Ladung längs einer Strecke ergeben. Es ist dies das Skalarprodukt aus dem der Strecke zugeordneten Vektor und dem dort wirkenden

Kraftvektor. Im Grenzübergang erhält man dann ein „Kurvenintegral".
Dies soll im Folgenden mathematisch präzisiert werden. Es sei Γ ein Weg im $\mathbb{R}^n$ mit Parameterdarstellung $p\colon [a,b] \to \mathbb{R}^n$ und $f\colon \mathbb{R}^n \to \mathbb{R}^n$ stetig. Es sei $\mathfrak{Z} = \{t_0, t_1, \dots, t_N\}$, $a = t_0 < t_1 < \dots < t_N = b$ eine Zerlegung des Parameterintervalls. Dann verstehen wir unter einer Riemannschen Summe von f über Γ (bzgl. p und $\mathfrak{Z}$ und bzgl. der Zwischenpunkte $\xi_i \in [t_i, t_{i+1}]$) die Zahl

$$\sum(f, \xi_i, \mathfrak{Z}) = \sum_{i=0}^{N-1} f\left(p(\xi_i)\right) \cdot \left(p(t_{i+1}) - p(t_i)\right).$$

Man beachte, daß $f\left(p(\xi_i)\right) \cdot \left(p(t_{i+1}) - p(t_i)\right)$ das *Skalarprodukt* der beiden Vektoren $f(p(\xi_i))$ und $(p(t_{i+1}) - p(t_i))$ ist.
Es gilt der

Satz 2.1.
Es sei $f\colon \mathbb{R}^n \to \mathbb{R}^n$ stetig und Γ ein Weg im $\mathbb{R}^n$ mit Parameterdarstellung $p\colon [a,b]$. Dann gibt es eine Zahl A mit der Eigenschaft:
Zu jedem $\varepsilon > 0$ gibt es ein $\delta > 0$, so daß für alle Zerlegungen $\mathfrak{Z} = \{t_0, \dots, t_N\}$ von $[a,b]$ mit Feinheit $|\mathfrak{Z}| < \delta$ und alle Zwischenpunkte $\xi_i \in [t_i, t_{i+1}]$ die Ungleichung

$$\left|\sum (f, \xi_i, \mathfrak{Z}) - A\right| < \varepsilon$$

gilt.
Man nennt A das *Kurvenintegral* von f über Γ und schreibt

$$A = \int_\Gamma f\, d\vec{s}.$$

Man überlegt sich leicht, daß die Definition unabhängig von der Parameterdarstellung ist.
Der folgende Satz liefert eine Möglichkeit zur Berechnung von Kurvenintegralen.

Satz 2.2.
Zusätzlich zu den Voraussetzungen von Satz 2.1 möge der Weg Γ eine stetig differenzierbare Parameterdarstellung $p\colon [a,b] \to \mathbb{R}^n$ besitzen. Dann gilt

$$\int_\Gamma f\, d\vec{s} = \int_a^b f((p(t)) \cdot \dot{p}(t)\, dt.$$

Der Beweis läuft ähnlich wie beim Nachweis der entsprechenden Formel für die Kurvenlänge.

Für den Fall, daß f eine Stammfunktion besitzt, d.h. es gilt $f = \nabla g$ mit einer stetig differenzierbaren Funktion $g\colon \mathbb{R}^n \to \mathbb{R}$, läßt sich der Weg des Kurvenintegrals in Satz 2.2 aufgrund der Substitutionsregel sofort angeben:

$$\begin{aligned}\int\limits_\Gamma f\,d\vec{s} &= \int\limits_a^b ((p(t))\cdot \dot{p}(t)\,dt \\ &= \int\limits_a^b \nabla(g\circ p)(t)\,dt \\ &= g\,(p(b)) - g\,(p(a))\,.\end{aligned}$$

Die bisherigen Überlegungen gelten selbstverständlich auch, wenn f und g nur in einer offenen Punktmenge $\Omega \subset \mathbb{R}^n$ erklärt sind und Γ in Ω verläuft.
Aus der obigen Formel

$$\int\limits_\Gamma f\,d\vec{s} = g(p(b)) - g(p(a))$$

im Falle $f = \nabla g$ erkennt man, daß der Wert des Kurvenintegrals nur vom Anfangs- und Endpunkt der Kurve, nicht jedoch von der verbindenden Kurve abhängt. Man nennt diese Eigenschaft die *„Wegunabhängigkeit des Kurvenintegrals“*.
Im allgemeinen sind Kurvenintegrale nicht wegunabhängig. Diese Frage ist sehr wichtig im Zusammenhang mit den Erhaltungssätzen der Physik.

Satz 2.3.
Das Kurvenintegral $\int\limits_\Gamma f d\vec{s}$ ist in Ω genau dann wegunabhängig, wenn $\int\limits_W f d\vec{s} = 0$ gilt für jeden geschlossenen Weg in Ω. Hierbei ist Ω eine offene, zusammenhängende Punktmenge des $\mathbb{R}^n$ und $f\colon \Omega \to \mathbb{R}^n$ stetig.

Beweis:

(a) Es sei $\int\limits_\Gamma f\,d\vec{s}$ in Ω wegunabhängig und Γ ein geschlossener Weg in Ω mit Anfangspunkt $=$ $a =$ Endpunkt. Da Ω offen ist, gibt es einen in Ω liegenden Kreis mit Radius $< \varepsilon$, dessen Rand $\tilde{W}$ den Punkt a enthält. $\tilde{W}$ definiert einen geschlossenen Weg W mit der Länge $|W| = 2\pi\varepsilon$.
Wegen der Wegunabhängigkeit gilt

$$\int\limits_\Gamma f\,d\vec{s} = \int\limits_W f\,d\vec{s},$$

da sowohl Γ als auch W den Punkt a mit a verbindet. Es gilt

$$\left|\int\limits_\Gamma f\,d\vec{s}\right| = \left|\int\limits_W f\,d\vec{s}\right| \le K\cdot 2\pi\varepsilon,$$

da $|f| \le K$ in $U(a)$, denn f ist stetig. Grenzübergang $\varepsilon \to 0$ ergibt

$$\int\limits_{\Gamma} f\,d\vec{s} = 0.$$

(b) Es gelte

$$\int\limits_{\Gamma} f\,d\vec{s} = 0$$

für jeden geschlossenen Weg Γ in Ω. Die Wege W_1 und W_2 mögen beide a und b verbinden (mit a als Anfangs- und b als Endpunkt). Es sei $-W_2$ der in entgegengesetzter Richtung durchlaufene Weg W_2, d.h. ist $p_1\colon [\alpha,\beta] \to \mathbb{R}^n$ eine Parameterdarstellung von W_1, dann ist $-W_2$ der durch die Parameterdarstellung $q\colon [\alpha,\beta] \to \mathbb{R}^n$
mit $q(t) = p_2(\alpha+\beta-t)$ definierte Weg. Es gilt

$$\int\limits_{W_2} f\,d\vec{s} = -\int\limits_{-W_2} f\,d\vec{s}.$$

Es sei W der durch Aneinandersetzen von W_1 und $-W_2$ entstehende Weg, d.h. W hat die Parameterdarstellung $g\colon [\alpha, 2\beta-\alpha] \to \mathbb{R}^n$ mit

$$g(t) = \begin{cases} p_1(t) & \text{für } t \in [\alpha,\beta], \\ q(t-\beta+\alpha) & \text{für } t \in [\beta, 2\beta-\alpha]. \end{cases}$$

Es gilt

$$\begin{aligned} \int\limits_{W} f\,d\vec{s} &= \int\limits_{W_1} f\,d\vec{s} + \int\limits_{-W_2} f\,d\vec{s} \\ &= \int\limits_{W_1} f\,d\vec{s} - \int\limits_{W_2} f\,d\vec{s}. \end{aligned}$$

Da W geschlossen ist, folgt

$$\int\limits_{W} f\,d\vec{s} = 0$$

und damit

$$\int\limits_{W_1} f\,d\vec{s} = \int\limits_{W_2} f\,d\vec{s},$$

also die behauptete Wegunabhängigkeit. $\square$

Es gilt auch die Umkehrung zu dem Satz, daß die Existenz einer Stammfunktion die Wegunabhängigkeit eines Kurvenintegrals impliziert.

Satz 2.4.
Es sei Ω eine beschränkte, offene, zusammenhängende Punktmenge des $\mathbb{R}^n$ und $f\colon \Omega \to \mathbb{R}^n$ stetig. Das Kurvenintegral $\int\limits_\Gamma f d\vec{s}$ sei wegunabhängig. Dann besitzt f eine Stammfunktion $g\colon \Omega \to \mathbb{R}$, d.h. es gilt: $f = \nabla g$ in Ω. Eine Darstellung von g ist

$$g(x) = \int\limits_{x_0}^{x} f\, d\vec{s}.$$

Hierbei ist $\int\limits_{x_0}^{x}$ das Kurvenintegral längs eines x_0 mit x verbindenden Weges.

Beweis:
Siehe [2, S.100, Satz 45].
Man definiert g durch die obige Darstellung und analysiert das Grenzverhalten der Differenzenquotienten.
Für den Fall, daß die Funktion f in Satz 2.4 stetige erste partielle Ableitungen hat, ist die Funktion g zweimal stetig partiell differenzierbar und wegen der Vertauschbarkeit der Ableitungen gilt

$$\frac{\partial^2 g}{\partial x_i \partial x_k} = \frac{\partial^2 g}{\partial x_k \partial x_i}$$

in Ω. Mit $\nabla g = f = (f_1, \dots, f_n)$ folgt

$$\frac{\partial}{\partial x_i} f_k = \frac{\partial}{\partial x_k} f_i,$$

und es erhebt sich die Frage, ob letztere Bedingung, die auch *Integrabilitätsbedingung* heißt, alleine hinreicht, die Wegunabhängigkeit des Kurvenintegrals zu sichern. Ein Gegenbeispiel zeigt, daß dies nicht immer der Fall ist: Man setze

$$\Omega = \left\{ x = (\xi, \eta) \in \mathbb{R}^2 \mid \xi^2 + \eta^2 > 0 \right\}$$

und

$$f(\xi, \eta) = \left(\frac{-\eta}{\xi^2 + \eta^2}, \frac{\xi}{\xi^2 + \eta^2} \right)^T.$$

Wählt man jedoch für Γ die Kurve mit der Parameterdarstellung

$$p\colon [0, 2\pi] \to \mathbb{R}^2, \quad p(t) = \begin{pmatrix} \cos t \\ \sin t \end{pmatrix},$$

so gilt

$$\begin{aligned}
\int_\Gamma f\,d\vec{s} &= \int_0^{2\pi} f(p(t))\cdot \dot{p}(t)\,dt \\
&= \int_0^{2\pi} (-\sin t, \cos t)\cdot(-\sin t, \cos t)\,dt \\
&= \int_0^{2\pi} (\sin^2 t + \cos^2 t)\,dt \\
&= 2\pi \neq 0.
\end{aligned}$$

D.h., das Kurvenintegral über den geschlossenen Weg Γ verschwindet nicht. Der Grund dafür, daß dies nicht gilt, liegt in der Tatsache, daß Ω in diesem Beispiel nicht einfach zusammenhängend ist. Es gilt nämlich der

Satz 2.5.
Es sei Γ eine beschränkte, offene und einfach zusammenhängende Punktmenge des $\mathbb{R}^n$ und $f\colon \Omega \to \mathbb{R}^n$ stetig differenzierbar. Notwendig und hinreichend für die Wegunabhängigkeit des Kurvenintegrals $\int_\Gamma f\,d\vec{s}$ ist das Erfülltsein der Integrabilitätsbedingungen

$$\frac{\partial f_i}{\partial x_k} = \frac{\partial f_k}{\partial x_i}.$$

in Ω, $\quad i,\ k = 1,\dots,n,\ f = (f_1,\dots,f_n)$.

Beweis:
Siehe [2, S.104, Satz 47]

Index

Literaturverzeichnis

[1] F. Erwe: Differential- und Integralrechnung I. BI, Mannheim, 1962.

[2] F. Erwe: Differential- und Integralrechnung II. BI, Mannheim, 1962.

[3] H. Grauert, I. Lieb: Differential- und Integralrechnung I. Springer-Verlag, Berlin, Heidelberg, 1967.

[4] H. Grauert, W. Fischer: Differential- und Integralrechnung II. Springer-Verlag, Berlin, Heidelberg, 1968.

[5] H. v. Mangoldt, K. Knopp, Einführung in die höhere Mathematik, Band I, II, III. Zwölfte Auflage. Stuttgart, 1963.

[6] jedes andere *anständige* Anfängerlehrbuch