Vorlesung über Analysis II

Philipps--Universität Marburg

Wintersemester 2020/21

Pablo Ramacher

Inhalt

[Vorlesung über Analysis II 1](#_Toc150088732)

[Kapitel 4: Das Riemann-Integral 3](#_Toc150088733)

[4.1 Zum Begriff des Volumens 3](#_Toc150088734)

[4.2 Das Jordan'sche Volumen 4](#_Toc150088735)

[4.3 Integration beschränkter Funktionen 16](#_Toc150088736)

[4.4 Charakterisierung Riemann-integrierbarer Funktionen 23](#_Toc150088737)

[4.5 Das Riemann'sche Integral in \R 30](#_Toc150088738)

[Das Wallis'sche Produkt 34](#_Toc150088739)

[4.6 Der Satz von Fubini und seine Anwendungen 38](#_Toc150088740)

[4.7 Uneigentliche Riemann-Integrale auf \R 47](#_Toc150088741)

[Kapitel 5: Differentialrechnung für Abbildungen mehrerer Variablen 52](#_Toc150088742)

[5.1 Die Ableitung einer Abbildung 52](#_Toc150088743)

Kapitel 4: Das Riemann-Integral

 ''Rückschauend glaube ich, es war schwieriger,

die Probleme zu erkennen, als sie zu lösen..."

Charles Darwin (Brief an Charles Lyell vom 30. September 1859)

## 4.1 Zum Begriff des Volumens

Im Folgenden soll der Frage nachgegangen werden, inwiefern für jede Teilmenge A\subset\R^n ein n-dimensionales „Volumen“ definiert werden kann, welches folgenden Forderungen genügt:

1. Jeder Menge A\subset\R^n ist eine Zahl \mu(A) zugeordnet, die nicht negativ oder unendlich ist und das n-dimensionale Volumen heißt;
2. Die Funktion \mu: \P(\R^n) \rightarrow \R\_+ \cup\{\infty\} ist translationsinvariant, d.h. für jedes x\_0 \in \R^n gilt \mu(A+x\_0)=\mu(A);
3. Sind A\_1, A\_2,... disjunkte Mengen, so gilt \mu(\cup\_{i=1}^\infty A\_i)= \sum\_{i=1}^\infty \mu(A\_i) (\sigma-Additivität);
4. Ist W=[a\_1,b\_1] \times \dots \times [a\_n,b\_n] ein Quader, so gilt \mu(W)=(b\_1-a\_1) \cdot...\cdot (b\_n-a\_n).

Hierbei bezeichnet \P(X) die Potenzmenge der Menge X, also die Menge aller Teilmengen von X, inklusive der leeren Menge und der Menge selbst. Wie der folgende Satz zeigt, gibt es jedoch keine Funktion \mu: \P(\R^n)\rightarrow \R\_+\cup\{\infty\}, die diese Anforderungen erfüllt:

<satz> Satz 4.1.1

Es existiert keine Funktion \mu: \P(\R^n)\rightarrow \R\_+\cup\{\infty\}, die

den Bedingungen (1)-(4) genügt.

</satz>

<beweis>

Der Einfachheit halber zeigen wir das Ergebnis nur für n=1. Wir nehmen zum Zwecke des Widerspruchs an, dass eine solche Funktion \mu: \P(\R)\rightarrow [0,\infty] existiert und betrachten im Intervall [0,1] die Äquivalenzrelation

x \sim y \Leftrightarrow x-y\in\Q.

Sei A \subset [0,1] eine Teilmenge, die aus jeder Äquivalenzklasse genau eine Zahl enthält und setze

B = \bigcup\_{r \in [-1,1] \cap \Q} (A+r).

Zuerst bemerken wir, dass für r\_1, r\_2 \in \Q

r\_1 \neq r\_2 \implies (A+r\_1) \cap (A+r\_2)=\emptyset.

Denn wäre a+r\_1 = a' + r\_2 für gewisse a, a' \in A, dann folgte a-a'=r\_1-r\_2\in\Q, d.h. a \sim a'. Aber nach Voraussetzung sollte A aus jeder Klasse genau einen Repräsentanten enthalten. Insgesamt ist B eine Vereinigung von abzählbar vielen disjunkten Mengen, so dass die triviale Inklusion [0,1] \subset B \subset [-1,2] zusammen mit den Eigenschaften (3) und (4) impliziert, dass

1 \leq \sum\_{r\in [-1,1]\cap \Q} \mu(A+r) \leq 3.

Weil \mu translationsinvariant ist nach Bedingung (2), muss jedoch auch

1 \leq \sum{r\in [-1,1]\cap \Q} \mu(A) \leq 3

gelten. Die Abschätzung nach oben bedeutet aber \mu(A)=0, was mit der nach unten nicht vereinbar ist.

</beweis>

Der einzige Ausweg besteht darin, ein n-dimensionales Volumen nicht für alle Teilmengen A\subset\R^n, sondern nur für bestimmte Teilmengen zu definieren. Es gibt zwei verschiedene Wege, dies zu tun:

1. Das Jordan'sche Volumen: Dieses Volumen ist nur für gewisse beschränkte Teilmengen definiert. Historisch gesehen ist dies die erste erfolgreiche Präzisierung des Volumenbegriffs. Vom Standpunkt der Forschung ist sie zwar überholt, aber ausgesprochen praktikabel und deswegen als Einstieg geeignet.
2. Das Lebesgue-Maß: Es wurde 1902 von Henri Lebesgue entwickelt und kann als Ausgangspunkt der modernen Maßtheorie gelten, auf der die ganze Wahrscheinlichkeitsrechnung fußt. Die Klasse der Lebesgue-messbaren Mengen ist sehr viel größer als die der Jordan-messbaren.

## 4.2 Das Jordan'sche Volumen

In diesem Abschnitt werden wir die Mengen beschreiben, die später als

zulässige Integrationsbereiche vorkommen können.

<definition> Definition 4.2.1

Wir definieren das Volumen eines Quaders W=[a\_1,b\_1]\times \dots\times [a\_n,b\_n] als

 \text{vol} (W):=(b\_1-a\_1)\cdot...\cdot (b\_n-a\_n).

</definition>

<definition> Definition 4.2.2

Eine Zerlegung des n-dimensionalen Quaders W ist ein n-Tupel P=(P\_1,..., P\_n) aus n endlichen Folgen P\_i=\{ a\_i=t^i\_0<t^i\_1<... < t^i\_{k(i)} = b\_i\} unterschiedlicher Länge.

Die Koordinaten t^i\_j werden Stützstellen der Zerlegung genannt. Durch

W\_{\alpha\_1,..., \alpha\_n} := [t^1\_{\alpha\_1}, t^1\_{\alpha\_1+1}]

\times \dots \times [t^n\_{\alpha\_n}, t^n\_{\alpha\_n+1}],

\alpha\_1= 0,..., k(1)-1, ..., \alpha\_n=0,..., k(n)-1,

werden dann k(1) \cdot ... \cdot k(n) kleinere Quader definiert und es gilt

1. $W = \bigcup\_{\alpha\_1=0}^{k(1)-1} ...

\bigcup\_{\alpha\_n=0}^{k(n)-1} W\_{\alpha\_1,..., \alpha\_n} $,

1. $\sum\_{\alpha\_1,..., \alpha\_n} \text{vol}

(W\_{\alpha\_1,..., \alpha\_n}) = \sum\_{\alpha\_1,..., \alpha\_n}

(t^1\_{\alpha\_1+1} - t^1\_{\alpha\_1})... (t^n\_{\alpha\_n+1} - t^n\_{\alpha\_n})

= (b\_1-a\_1) ... (b\_n-a\_n) = \text{vol}(W) $.

</definition>

<beispiel> Beispiel 4.2.3

Wir betrachten einen 2-dimensionalen Quader mit k(1)=5 und k(2)=3, also

P\_1= (a\_1=t^1\_0, t^1\_1,..., b\_1=t^1\_5) und

P\_2= (a\_2=t^2\_0, t^2\_1,..., b\_2=t^2\_3).

Damit haben wir eine Zerlegung des großen Quaders in 15 kleine Quader.

</beispiel>

<definition> Definition 4.2.4

Die Zerlegung P=(P\_1,..., P\_n) des Quaders W heißt feiner als die Zerlegung P'=(P'\_1,..., P'\_n) oder Verfeinerung von P', falls P'\_i\subset P\_i für alle i=1,...,n gilt.

</definition>

Offenbar gilt mit der Notation W\_\alpha:=W\_{\alpha\_1,...,\alpha\_n}

1. Ist P' eine Verfeinerung von P, so ist jeder Quader W'\_{\beta} der Zerlegung P' in einem Quader W\_{\alpha} der Zerlegung P enthalten.
2. Jeder Quader W\_{\alpha} der Zerlegung P ist die Summe gewisser Quader W'\_{\beta} der Zerlegung P' und \text{vol} W\_{\alpha} = \sum\_{W'\_{\beta}\subset W\_\alpha}\text{vol} W'\_{\beta}.

Wir bemerken, dass diese und alle weiter auftretenden Summen endlich sind, weil eine Zerlegung nur endlich viele Teilwürfel hat. Sei nun A\subset \R^n eine beschränkte Menge A\subset \R^n und wähle zunächst einen Quader W mit A\subset W. Zu jeder Zerlegung P von W betrachten wir jeweils die Obersumme zu P

\overline{S} (A;P) := \sum\_{W\_\alpha \cap \overline{A} \neq \emptyset} \text{vol} W\_{\alpha}

sowie die Untersumme zu P

\underline{S}(A;P) := \sum\_{W\_\alpha \subset Int A} vol W\_{\alpha}.

Offenbar gilt \underline{S}(A;P)\leq \overline{S}(A;P).

<lemma> Lemma 4.2.5

Ist P' eine Verfeinerung der Zerlegung P, so gilt

\underline{S}(A;P) \leq \underline{S}(A;P') \leq \overline{S}(A;P') \leq \overline{S}(A;P).

</lemma>

<beweis>

Es ist

\underline{S}(A;P) = \sum\_{W\_\alpha\subset \text{Int} A}\text{vol} W\_{\alpha}

 = \sum\_{W\_\alpha\subset \text{Int} A}

\left[ \sum\_{W'\_\beta\subset W\_\alpha} \text{vol} W'\_{\beta} \right]

\leq \sum\_{W'\_\beta\subset\text{Int} A} \text{vol} W'\_{\beta} = \underline{S}(A;P').

Ein entsprechendes Argument für die Obersummen liefert die Behauptung.

</beweis>

<korollar> Korollar 4.2.6

Sind P und Q zwei Zerlegungen von W, so gilt

\underline{S}(A;P)\leq \overline{S}(A;Q).

<\korollar>

<beweis>

Wir wählen eine Zerlegung P', die sowohl eine Verfeinerung von P als

auch von Q ist und wenden das vorherige Lemma an. Damit folgt

\underline{S}(A;P)\leq\underline{S}(A;P')\leq

\overline{S}(A;P')\leq\overline{S}(A;Q). \qedhere

</beweis>

<definition> Definition 4.2.7 Jordan'sches Volumen

Sei A \subset \R^n eine beschränkte Menge.

1. Das innere Jordan'sche Volumen von A ist durch

\underline{\text{vol}} A := \sup\_{P} \underline{S}(A;P) gegeben.

1. Das äußere Jordan'sche Volumen von A ist durch

\overline{\text{vol}} A := \inf\_{P} \overline{S}(A;P) gegeben.

1. Die Menge A heißt Jordan-messbar, falls

\underline{\text{vol}} A = \overline{\text{vol}} A gilt. In diesem Fall

wird diese Zahl Jordan-Volumen genannt und \text{vol} A geschrieben.

</definition>

Man beachte, dass Korollar 4.2.6 impliziert, dass

0 \leq \underline{\text{vol}} A \leq \overline{\text{vol}} A < \infty gilt.

<beispiel> Beispiel 4.2.8

Die Menge A= \Q \cap [0,1] ist nicht Jordan-messbar, denn für jede

Zerlegung 0 = t\_0 < t\_1 < ... < t\_k = 1 des Intervalls [0,1] gilt

\underline{S}(A;P)=0, da A keine inneren Punkte hat, und

\overline{S}(A;P')=1, da \overline {A}=[0,1].

</beispiel>

Folgende Eigenschaften sind sofort einsichtig:

<lemma> Lemma 4.2.9

1. Ist A\subset \R^n eine beschränkte Menge mit \overline{\text{vol}} A=0, so ist A Jordan-messbar und es gilt \text{vol} A=0. Eine solche Menge wird (Jordan)-Nullmenge genannt.
2. Jede Teilmenge einer Nullmenge ist wieder eine Nullmenge.
3. Sind A,B beschränkte Jordan-messbare Mengen mit A\subset B, so gilt \text{vol}A \leq \text{vol}B.
4. Ist A Jordan-messbar und x\_0\in\R^n, so ist auch die verschobene Menge x\_0+A Jordan-messbar und es gilt \text{vol} A=\text{vol}(x\_0+A) (Translations-Invarianz des Jordan-Volumens).

</lemma>

Wir erinnern im Folgenden an einige Eigenschaften des topologischen Randes einer Menge aus der Analysis I.

<lemma> Lemma 4.2.10

Sei (X,d) ein metrischer Raum, A\subset X und \partial(A):=\overline{A}\cap\overline{X-A} der Rand der Menge A.

1. \overline{A}=\text{Int}(A)\stackrel{.}{\cup}{\partial}(A) (disjunkte Vereinigung),
2. {\partial}(\text{Int}(A))\subset {\partial}(A),
3. {\partial}(\overline{A})\supset{\partial}(A),
4. Ist C\subset X eine beliebige zusammenhängende Menge mit C\cap A, C\cap(X-A) \neq \emptyset, so folgt C\cap{\partial}(A)\neq\emptyset.

</lemma>

Wir zeigen als nächstes, dass der Rand einer Menge eng mit der Differenz zwischen äußerem und innerem Jordan'schen Volumen verbunden ist:

<satz> Satz 4.2.11

Für jede beschränkte Menge A\subset \R^n gilt

\overline{\text{vol}} A - \underline{\text{vol}}A = \overline{\text{vol}}{\partial}(A).

</satz>

<beweis>

Sei W ein Quader, der A enthält, und P eine beliebige Unterteilung

von W. Wir schreiben wie zuvor W\_{\alpha\_1, ...,\alpha\_n} =: W\_\alpha.

Es folgt

\overline{S}(A;P)- \underline{S}(A;P) =

\sum\_{W\_\alpha\cap\overline{A}\neq\emptyset} \text{vol} W\_{\alpha}

 - \sum\_{W\_\alpha\subset \text{Int} A} \text{vol} W\_{\alpha}

 = \sum\_{ W\_\alpha \cap \overline{A} \neq \emptyset \hat W\_\alpha \not \subset \text{Int} A }} \text{vol} W\_{\alpha}.

Sei W\_\alpha ein Quader mit

W\_\alpha\cap\overline{A}\neq\emptyset und W\_\alpha\not\subset \text{Int} A.

Wir zeigen, dass diese beiden Bedingungen zu W\_\alpha \cap {\partial}(A) \neq \emptyset äquivalent sind. Angenommen, es gelten die beiden Bedingungen. Nach Eigenschaft (1) des Randes muss dann W\_\alpha \cap \text{Int}{A} \neq \emptyset oder W\_\alpha \cap \partial (A) \neq \emptyset oder beides sein.

Im zweiten Fall ist nichts mehr zu zeigen, im ersten folgt W\_\alpha\cap A\neq\emptyset und aus W\_\alpha\not\subset \text{Int} A sofort W\_\alpha\cap (\R^n-\overline{A})\neq\emptyset, also erst recht W\_\alpha\cap (\R^n-A)\neq\emptyset;

weil Quader zusammenhängend sind, liefert Eigenschaft (4) mit C=W\_\alpha auch in diesem Fall

W\_\alpha\cap{\partial}(A)\neq\emptyset.

Gelte umgekehrt W\_\alpha\cap{\partial}(A)\neq\emptyset. Nach Eigenschaft (1) folgt dann sofort, dass W\_\alpha\cap\overline{A}\neq\emptyset und W\_\alpha\not\subset \text{Int} A, was den Beweis der Äquivalenz der verschiedenen Bedingungen beendet. Insgesamt können wir die Summe also umschreiben und erhalten unter Verwendung von {\partial}(A)=\overline{{\partial}(A)}:

\overline{S}(A;P)- \underline{S}(A;P) =

\sum\_{W\_\alpha\cap{\partial}(A)\neq\emptyset}\text{vol} W\_{\alpha}

 = \overline{S}({\partial}(A);P).

Weil dies für jede beliebige Unterteilung gilt, folgt die Behauptung.

</beweis>

<bemerkung> Bemerkung 4.2.12

Um zu zeigen, dass eine beschränkte Menge A Jordan-messbar ist, genügt es zu zeigen, dass für eine beliebige Folge von Partitionen P\_k \limsup \underline S(A,P\_k)= \liminf \overline S(A,P\_k) erfüllt ist.

</bemerkung>

<lemma> Lemma 4.2.13

Sind A und B (beschränkte) Nullmengen, so ist auch A\cup B eine Nullmenge.

</lemma>

<beweis>

Wir wählen einen Quader W, der A\cup B enthält.

Weil A und B Nullmengen sind, existiert zu jedem \epsilon>0 eine Zerlegung P\_{\epsilon} von W mit

\overline{S}(A;P\_\epsilon) = \sum\_{W\_\alpha\cap\overline{A}\neq\emptyset}

\text{vol} W\_{\alpha} < \epsilon,

\overline{S}(B;P\_\epsilon) = \sum\_{W\_\alpha\cap\overline{A}\neq\emptyset}

\text{vol} W\_{\alpha} < \epsilon.

Dabei sind die W\_{\alpha} Quader der Zerlegung P\_{\epsilon} von W.

Wegen W\_\alpha \cap \overline{A\cup B} = (W\_\alpha\cap \overline{A}) \cup

(W\_\alpha\cap \overline{B}) folgt desweiteren

W\_\alpha \cap \overline{A\cup B} \neq \emptyset \Leftrightarrow

W\_\alpha\cap \overline{A}\neq\emptyset

\text{ oder } W\_\alpha\cap \overline{B}\neq\emptyset.

Damit ist

\overline{S}(A\cup B;P\_\epsilonilon) = \sum\_{W\_\alpha\cap\overline{A\cup B}\neq\emptyset} \text{vol} W\_\alpha \leq

\sum\_{W\_\alpha\cap\overline{A}\neq\emptyset} \text{vol} W\_{\alpha} +

\sum\_{W\_\alpha\cap\overline{B}\neq\emptyset} \text{vol} W\_{\alpha} < 2\epsilon.

Da \epsilon beliebig war, folgt insgesamt \overline{\text{vol}}( A\cup B) = 0.

</beweis>

Wir bezeichnen im Folgenden mit {\mathcal{J}}(\R^n) \subset \P(\R^n) die Familie aller

beschränkten Jordan-messbaren Mengen.

<definition> Definition 4.2.14

Eine Familie \D \subset \P(\R^n) heißt Mengenring, falls für beliebige Mengen A,B\in\D auch A\cup B, A\cap B und A - B in \D liegen.

</definition>

<satz> Satz 4.2.15

Die Familie der Jordan-messbaren Mengen {\mathcal{J}}(\R^n) bildet einen Mengenring.

</satz>

<beweis>

Seien A,B\in {\mathcal{J}}(\R^n).

Wir zeigen zunächst {A\cup B\in {\mathcal{J}}(\R^n)}. Hierzu bemerken wir, dass in einem metrischen Raum allgemein die Inklusion

{\partial}(A\cup B)\subset {\partial}(A)\cup{\partial}(B)

gilt. So ist nach Definition des Randes

\partial(A\cup B) = \overline{[A\cup B]}\cap \overline{[X - (A\cup B)]} =

[\overline{A}\cup\overline{B}] \cap [ \overline{(X-A)\cap (X-B)}]

\subset [\overline{A} \cap \overline{(X-A)}]\cup [\overline{B} \cap \overline{(X-B)}]

= \partial(A)\cup \partial(B).

Weil A,B \in {\mathcal{J}}(\R^n) gilt, sind {\partial}(A), {\partial}(B) Nullmengen, so dass nach vorigem Lemma auch {\partial}(A)\cup {\partial}(B) eine Nullmenge ist. Dann ist aber trivialerweise auch die Teilmenge {\partial}(A\cup B) eine Nullmenge und wir erhalten A\cup B\in{\mathcal{J}}(\R^n).

Als nächstes zeigen wir A-B \in {\mathcal{J}}(\R^n). Dazu bemerken wir, dass allgemein \partial(A-B) \subset {\partial}(A)\cup{\partial}(B) sowie A-B=A\cap (X-B) gilt.

Tatsächlich ist \partial(A - B) = \overline{[A- B]} \cap \overline{[X - (A- B)]} = \overline{[A\cap (X-B)]} \cap \overline{[X-(A\cap(X-B))]}

\subset \overline{A}\cap \overline{(X-B)} \cap \overline{[(X-A)\cup B]}.

Dabei wurde im zweiten Schritt die allgemeine Identität X-(C\cap D)= (X-C)\cup (X-D) benutzt. Nun vertauschen aber Abschluß und Vereinigung und wir erhalten insgesamt das Ergebnis

\partial(A - B)\subset \overline{A}\cap \overline{(X-B)} \cap

[\overline{(X-A)}\cup \overline{B}]

\subset [\overline{A} \cap \overline{(X-A)}] \cup

[\overline{B} \cap \overline{(X-B)}]={\partial}(A)\cup{\partial}(B).

Eine Argumentation wie im vorherigen Fall liefert dann das Gewünschte.

Schließlich zeigen wir {A\cap B\in {\mathcal{J}}(\R^n)}. Dies folgt jedoch aus den beiden anderen Teilergebnissen, da für einen Quader mit A\cup B\subset W gilt A\cap B = W- [(W-A)\cup(W-B)].

</beweis>

<bemerkung> Bemerkung 4.2.16

Da Komplemente beschränkter Mengen im allgemeinen nicht wieder beschränkt sind, kann {\mathcal{J}}(\R^n) keine Mengenalgebra bzw. Boole'sche Algebra bilden, da hierfür mit jedem A auch dessen Komplement in {\mathcal{J}}(\R^n) liegen müßte.

</bemerkung>

<satz> Satz 4.2.17

Für A,B\in {\mathcal{J}}(\R^n) gilt

\text{vol} (A\cup B) \leq \text{vol}A +\text{vol} B.

Gilt \text{Int}(A\cap B)=\emptyset, so steht hier sogar Gleichheit.

</satz>

<beweis>

Sei W ein Quader, der A\cup B enthält, und P eine beliebige Zerlegung von W. Es ist

\overline{S}(A\cup B;P) =

\sum\_{W\_\alpha\cap\overline{A\cup B}\neq\emptyset} \text{vol} W\_{\alpha}

 \leq \sum\_{W\_\alpha\cap\overline{A} \neq \emptyset} \text{vol} W\_{\alpha}

+ \sum\_{W\_\alpha\cap\overline{B}\neq\emptyset} \text{vol} W\_{\alpha}

 = \overline{S}(A;P) + \overline{S}(B;P).

Durch Übergang zum Infimum, erst auf der rechten, dann auf der linken Seite, entsteht \overline{\text{vol}}(A\cup B) \leq \overline{\text{vol}} A +\overline{\text{vol}} B.

Da all diese Mengen Jordan-messbar sind, ist dies gleichbedeutend mit

\text{vol} (A\cup B) \leq \text{vol}A +\text{vol} B.

Gilt nun \text{Int}(A\cap B)=\emptyset, so folgt wegen \text{Int}(A\cap B)=\text{Int}A\cap\text{Int}B, dass \text{Int}A und \text{Int}B keine gemeinsamen Punkte haben; ein Quader, der in \text{Int}A\cup\text{Int}B liegt, ist also vollständig in einem der beiden Inneren enthalten. Dies bedeutet

\underline{S}(A;P) + \underline{S}(B;P)

 = \sum\_{W\_\alpha\subset\text{Int}A} \text{vol} W\_{\alpha}

+ \sum\_{W\_\alpha\subset\text{Int}B} \text{vol} W\_{\alpha}

= \sum\_{W\_\alpha\subset\text{Int}A\cup\text{Int}B} \text{vol} W\_{\alpha}

\leq \sum\_{W\_\alpha\subset\text{Int}(A\cup B)} \text{vol} W\_{\alpha}

= \underline{S}(A\cup B;P),

wobei wir die Inklusion \text{Int}A\cup\text{Int}B\subset \text{Int}(A\cup B) berücksichtigten. Durch Übergang zum Supremum und Ausnutzen der Jordan-Messbarkeit erhält man nun wie eben \text{vol} (A\cup B) \geq \text{vol}A +\text{vol} B und damit die umgekehrte Ungleichung.

</beweis>

Wir zeigen im folgenden, dass das zu Anfang definierte Volumen eines Quaders mit seinem Jordan-Volumen übereinstimmt und somit dieser

Volumenbegriff in sich konsistent ist.

<satz> Satz 4.2.18

Jeder Quader W=[a\_1,b\_1] \times \dots \times [a\_n,b\_n] ist Jordan-messbar und sein

Jordan-Volumen ist \text{vol} (W)=(b\_1-a\_1) \cdot ... \cdot (b\_n-a\_n).

</satz>

<beweis>

Wir wählen einen größeren Quader \tilde{W}, der W enthält, und eine Zerlegung P\_\epsilon=(P\_1,..., P\_n) von \tilde{W} mit der Eigenschaft, dass in P\_i die Folge von Stützstellen a\_i-\epsilon<a\_i+\epsilon<b\_i-\epsilon<b\_i+\epsilon vorkommt.

Von allen Quadern dieser Zerlegung liegt nur [a\_1+\epsilon,b\_1-\epsilon]\times \dots\times [a\_n+\epsilon,b\_n-\epsilon] in \text{Int} W. Damit ist \underline{S}(W;P\_\epsilon) = (b\_1-a\_1-2\epsilon)\times \dots \times (b\_n-a\_n-2\epsilon).

Weiter schneiden nur die 3^n Quader vom Typ [t\_1,s\_1]\times \dots\times [t\_n,s\_n] mit

t\_i = a\_i-\epsilon, a\_i+\epsilon \text{ bzw. } b\_i-\epsilon,

s\_i = a\_i+\epsilon, b\_i-\epsilon \text{ bzw. } b\_i+\epsilon

den Quader W; die Summe ihrer Volumina ist gleich dem Volumen des Quaders

[a\_1-\epsilon,b\_1+\epsilon]\times \dots\times [a\_n-\epsilon,b\_n+\epsilon],

so dass

\overline{S}(W;P\_\epsilon) = (b\_1-a\_1+2\epsilon)\times \dots \times (b\_n-a\_n+2\epsilon).

Lassen wir \epsilon\rightarrow 0 gehen, so folgt

\prod\_{i=1}^n (b\_i-a\_i) \leq \sup\_P \underline{S}(W;P)

\leq \inf\_P \overline{S}(W;P) \leq \prod\_{i=1}^n (b\_i-a\_i),

was sofort \overline{\text{vol}} W=\underline{\text{vol}} W = \prod\_{i=1}^n (b\_i-a\_i) impliziert.

</beweis>

<satz> Satz 4.2.19 [ Regularität des Jordan-Volumens]

Zu jeder Jordan-messbaren Menge A und jedem \epsilon>0 existieren eine kompakte Menge K und eine offene Menge U mit K\subset A\subset U, K,U\in{\mathcal{J}}(R^n), \text{vol} U-\epsilon\leq \text{vol}A\leq \text{vol} K+\epsilon.

</satz>

<beweis>

Wegen \text{vol}A=\underline{\text{vol}} A existiert eine Zerlegung P mit

 \text{vol} A-\epsilon < \sum\_{W\_\alpha\subset\text{Int} A}\text{vol} W\_\alpha

und wir setzen K:=\bigcup\_{W\_\alpha\subset\text{Int} A} W. Diese Menge ist als Vereinigung endlich vieler kompakter Quader kompakt. Weil sich Quader einer Zerlegung niemals in inneren Punkten schneiden, gilt zudem mit Satz 4.2.17

\text{vol} K = \sum\_{W\_\alpha\subset\text{Int} A}\text{vol} W\_\alpha,

also \text{vol} A\leq \text{vol} K+\epsilon. Ähnlich konstruieren wir die offene Menge U. So existiert wegen

\text{vol}A=\overline{\text{vol}} A eine Zerlegung P' mit \sum\_{W\_\alpha\cap\overline{A}\neq\emptyset} \text{vol} W\_\alpha < \text{vol} A+\epsilon.

Damit gilt

A\subset \bigcup\_{W\_\alpha\cap\overline{A}\neq\emptyset} W\_\alpha =: \tilde{U}.

Es gibt nun zwei Arten von Punkten aus A. Solche x \in A, die im Inneren eines Quaders W\_\alpha mit W\_\alpha\cap\overline{A}\neq\emptyset liegen, kommen nur in diesem einen offenen Quader vor und liegen natürlich in \text{Int}\tilde{U}. Daneben kommen noch Punkte x\in A vor, die in {\partial} W\_\alpha für ein W\_\alpha sind, also auf einer Quaderkante oder Quaderecke liegen. Auf jeden Fall liegt damit x nicht nur in diesem Quader, sondern auch noch in einem weiteren Nachbarwürfel (bei Kanten) bzw. in drei Nachbarwürfeln (bei Ecken). Nimmt man diese Quader zusammen, insgesamt also zwei bzw. vier, so ist x darin ein innerer Punkt, liegt also wieder in \text{Int}\tilde{U}.

Insgesamt ist somit A\subset \text{Int}\tilde{U}=:U. Zudem ist U Jordan-messbar, weil es sich von \tilde{U} nur um eine Nullmenge unterscheidet. Damit ist alles bewiesen.

</beweis>

Für endlich viele disjunkte Jordan-messbare Mengen wissen wir bereits, dass das Volumen der Vereinigung die Summe der einzelnen Volumina ist. Wie der folgende Satz zeigt, gilt eine entsprechende Aussage auch für abzählbare Vereinigungen.

<satz> Satz 4.2.20 [Eingeschränkte \sigma-Additivität des Jordan-Volumens]

Sind A\_1,A\_2, A\_3 ... Jordan-messbar, paarweise disjunkt und ist auch A:=\bigcup\_{i=1}^\infty A\_i Jordan-messbar, so gilt \text{vol} A = \text{vol} \bigcup\_{i=1}^\infty A\_i = \sum\_{i=1}^\infty \text{vol}A\_i.

</satz>

<bemerkung>

Uneingeschränkte \sigma-Additivität würde vorliegen, wenn die abzählbare Vereinigung messbarer Mengen automatisch wieder messbar ist. Beim Lebesgue-Maß ist dies der Fall.

</bemerkung>

<beweis>

Der Beweis stützt sich auf die Regularität des Jordan-Volumens. Ist \epsilon>0 gegeben, so existieren offene Mengen U\_i\in{\mathcal{J}}(\R^n) mit A\_i\subset U\_i und \text{vol} U\_i< \text{vol} A\_i + \epsilon/2^i. Weil zudem A Jordan-messbar ist, existiert eine kompakte Menge K\in{\mathcal{J}}(\R^n) mit K\subset A und \text{vol} A<\text{vol} K+\epsilon. Insgesamt haben wir die Inklusionen

K \subset A = \bigcup\_{i=1}^\infty A\_i \subset \bigcup\_{i=1}^\infty U\_i,

so dass die Mengen U\_i eine \Überdeckung des Kompaktums K bilden.

Nach der Überdeckungseigenschaft von Heine-Borel existieren dann bereits endlich viele offene Mengen U\_i dergestalt, dass K\subset U\_{i\_1}\cup...\cup U\_{i\_N}. Es folgt mit Satz 4.2.17

\text{vol} K \leq \text{vol} (U\_{i\_1}\cup...\cup U\_{i\_N})

\leq \text{vol} U\_{i\_1}+...+ \text{vol} U\_{i\_N} \leq \sum\_{i=1}^\infty \text{vol} U\_i

\leq \sum\_{i=1}^\infty \text{vol} A\_i +\sum\_{i=1}^\infty \frac{\epsilon}{2^i}

= \sum\_{i=1}^\infty \text{vol} A\_i +\epsilon.

Damit gilt

\text{vol} A \leq \text{vol} K +\epsilon\leq \sum\_{i=1}^\infty \text{vol} A\_i + 2\epsilon

für alle \epsilon>0, also im Limes \epsilon\rightarrow 0

<eq1>

\text{vol} A \leq \sum\_{i=1}^\infty \text{vol} A\_i .

</eq1>

Andererseits haben wir für beliebiges N\in\N die Inklusion A\_1\cup...\cup A\_N\subset A, also mit Satz 4.2.17 \text{vol} A \geq \text{vol} (A\_1\cup...\cup A\_N) = \sum\_{i=1}^N \text{vol} A\_i,

da die A\_i disjunkt sind. Die Folge auf der rechten Seite ist monoton und beschränkt, mithin konvergent und es gilt im Limes N\rightarrow \infty

<eq2>

\text{vol} A \geq \sum\_{i=1}^\infty \text{vol} A\_i .

</eq2>

Die Gleichungen \eqref{eq:1} und \eqref{eq:2} ergeben zusammen dann das gewünschte Ergebnis.

</beweis>

## 4.3 Integration beschränkter Funktionen

In diesem Abschnitt soll eine geeignete Klasse von Funktionen eingeführt werden, die über Jordan-messbaren Mengen integriert werden können. Wie bei den Jordan-messbare Mengen ist auch hier die Beschränktheit wesentlich.

<definition>

Sei W\subset\R^n ein Quader und f: W\rightarrow \R eine beschränkte Funktion. Ist P eine Zerlegung von W mit den Quadern W\_\alpha:=W\_{\alpha\_1,...,\alpha\_n}, so definieren wir

\overline{S}(f,W;P) := \sum\_{W\_\alpha} \sup f\big|\_{W\_\alpha}\cdot \text{vol} W\_{\alpha},

\underline{S}(f,W;P) := \sum\_{W\_\alpha} \inf f\big|\_{W\_\alpha}\cdot \text{vol} W\_{\alpha}.

 </definition>

Offenbar gilt \underline{S}(f,W;P) \leq \overline{S}(f,W;P).

<lemma>

Ist P' eine Verfeinerung von P, so gilt \underline{S}(f,W;P) \leq \underline{S}(f,W;P') \leq \overline{S}(f,W;P') \leq \overline{S}(f,W;P).

</lemma>

<beweis>

Ist W'\_{\beta} ein Quader von P', so ist dieser in einem Quader W\_\alpha von P enthalten. Also ist

\underline{S}(f,W;P')& =& \sum\_{W'\_\beta} \inf f\big|\_{W'\_\beta}\text{vol}

W'\_\beta = \sum\_{W\_\alpha} \left[ \sum\_{W'\_\beta\subset W\_\alpha}

\inf f\big|\_{W'\_\beta} \text{vol} W'\_\beta\right]\\

&\geq & \sum\_{W\_\alpha} \left[ \sum\_{W'\_\beta\subset W\_\alpha}

\inf f\big|\_{W\_\alpha} \text{vol} W'\_\beta\right] = \

\sum\_{W\_\alpha} \inf f\big|\_{W\_\alpha} \left[ \sum\_{W'\_\beta\subset W\_\alpha}

\text{vol} W'\_\beta\right]\\

&=& \sum\_{W\_\alpha} \inf f\big|\_{W\_\alpha} \text{vol} W\_{\alpha} =

\underline{S}(f,W;P).

Analog zeigt man die zweite Ungleichung.

</beweis>

<definition>[\bf Oberes und unteres Riemann-Integral]

Sei W\subset\R^n ein Quader und f: W\rightarrow \R eine beschränkte Funktion. Bezeichnet P eine Zerlegung von W, so definieren wir das untere Riemann-Integral von f über W durch

 \sup\_{P} \underline{S}(f,W;P) =

\int^u\_W f

und das obere Riemann-Integral von f über W durch

 \inf\_{P} \overline{S}(f,W;P) =

\int^o\_W f.

Die Funktion f wird

Riemann-integrierbar genannt, falls \int^u\_W f= \int^o\_W f gilt.

In diesem Fall schreiben wir für diese Zahl

\int\_W f = \int\_W f(x) dx.

</definition>

<satz>\label{int-stetig}

Ist f:W\rightarrow \R eine beschränkte, Riemann-integrierbare Funktion und h:\R \rightarrow \R

stetig, so ist auch h\circ f: W\rightarrow \R Riemann-integrierbar.

</satz>

<beweis>

Es gelte f(W)\subset [-m,m]. Weil h stetig ist, ist h auf dem Kompaktum [-m,m] nach dem Satz von Heine sogar gleichmäßig stetig, d.h.für jedes \epsilon>0 existiert ein \delta mit |h(t)-h(s)|<\epsilon für alle t,s\in [-m,m] mit |t-s|<\delta. Dabei kann ohne Einschränkung angenommen werden, dass \delta\leq \epsilon. Sei weiterhin k:=\sup\_{t\in [-m,m]} |h(t)|. Weil f Riemann-integrierbar ist, existiert eine Zerlegung P von W mit

\underline{S}(f,W;P) > \int\_W f - \frac{\delta^2}{2},

\overline{S}(f,W;P) < \int\_W f + \frac{\delta^2}{2}.

Dann ist

<eq:sumdiff>

\overline{S}(f,W;P) - \underline{S}(f,W;P) =\

\sum\_{W^\*\in P} [\sup f\big|\_{W^\*} - \inf f\big|\_{W^\*}]\text{vol} W^\* < \delta^2.

</eq:sumdiff>

Wir unterteilen jetzt P in zwei Familien von Quadern gemäß

P\_1 := \{ W^\* \in P : \sup f\big|\_{W^\*} - \inf f\big|\_{W^\*} < \delta \},

P\_2 := \{ W^\* \in P : \sup f\big|\_{W^\*} - \inf f\big|\_{W^\*} \geq \delta\}.

Aufgrund von \eqref{eq:sumdiff} ist

\delta \sum\_{W^\*\in P\_2} \text{vol} W^\*

\leq \sum\_{W^\*\in P\_2} [\sup f \big |\_{W^\*} - \inf f \big|\_{W^\*}] \text{vol} W^\*

\leq \sum\_{W^\*\in P} [\sup f\big|\_{W^\*} - \inf f\big|\_{W^\*}]\text{vol} W^\* < \delta^2

und es ergibt sich

\sum\_{W^\*\in P\_2} \text{vol} W^\* < \delta.

Wir betrachten nun Quader W^\* \in P\_1. Für solche gilt \sup f\big|\_{W^\*} - \inf

f\big|\_{W^\*} < \delta. Weil trivialerweise f(W^\*)\subset [\inf f\big|\_{W^\*} , \sup f\big|\_{W^\*}] gilt, ist f(W^\*) in einem Intervall der Länge \leq \delta enthalten; damit gilt für alle s,t \in f(W^\*) wegen der gleichmäßigen Stetigkeit von h, dass |h(t)-h(s)|<\epsilon. Hieraus folgert man

\sup (h\circ f) \big|\_{W^\*} - \inf (h\circ f) \big|\_{W^\*} < \epsilon \forall W^\*\in P\_1 .

Schätzen wir die Ober- und Untersumme für h\circ f ab, so berechnet man

\overline{S}(h\circ f,W;P) & - & \underline{S}(h \circ f,W;P)

= \sum\_{W^\*\in P\_1} [\sup (h\circ f)\big|\_{W^\*} - \inf (h\circ f)\big|\_{W^\*}] \text{vol} W^\*

 + \sum\_{W^\*\in P\_2} [\sup (h\circ f)\big|\_{W^\*} - \inf (h\circ f) \big|\_{W^\*}]\text{vol} W^\*

\leq \epsilon \sum\_{W^\*\in P\_1} \text{vol} W^\* + 2k \delta .

Dabei ist \sum\_{W^\*\in P\_1} \text{vol} W^\* \leq \text{vol} W, während \delta so gewählt war, dass \delta \leq \epsilon. Damit ergibt sich insgesamt

\overline{S}(h\circ f,W;P) - \underline{S}(h \circ f,W;P)

\leq \epsilon (\text{vol} W+2k). </beweis>

<satz> Satz 4.3.5

Seien f,g:W\rightarrow \R beschränkt und Riemann-integrierbar. Dann gilt:

(1) f+g ist Riemann-integrierbar und \int\_W (f+g)(x) dx = \int\_W f(x) dx+\int\_W g(x) dx;

(2) f\cdot g ist Riemann-integrierbar;

(3) |f|:W\rightarrow [0,\infty [ ist Riemann-integrierbar und |\int\_W f |\leq \int\_W |f|;

(4) Gilt f\leq g, so gilt auch \int\_W f \leq \int\_W g;

(5) c\cdot f ist für jede Konstante c\in \R

Riemann-integrierbar und \int\_W c\cdot f = c\int\_W f.

</satz>

<beweis>

(1) Für eine beliebige Zerlegung P von W gilt

\underline{S}(f+g,W;P) = \sum\_{W^\*\in P} [\inf (f+g)\big|\_{W^\*}]\text{vol} W^\*

\geq \sum\_{W^\*\in P} [\inf f\big|\_{W^\*}+ \inf g\big|\_{W^\*}]\text{vol} W^\*\\

\geq \underline{S}(f,W;P)+ \underline{S}(g,W;P).

Ebenso zeigt man

\overline{S}(f+g,W;P) \leq \overline{S}(f,W;P)+ \overline{S}(g,W;P),

so dass insgesamt für die oberen und unteren Riemann-Integrale die Abschätzungen

\int\_W^u f+\int\_W^u g \leq \int\_W^u f+g \leq \int\_W^o f+g \leq \int\_W^o f+\int\_W^o g

gelten. Weil f und g Riemann-integrierbar sind, folgt \int\_W^u f+g = \int\_W^o f+g. Also ist f+g ebenfalls Riemann-integrierbar, und dies ist gleich \int\_W f+\int\_W g.

(2) Wegen f \cdot g = \frac{1}{4} [(f+g)^2 - (f-g)^2] folgt (2) unmittelbar aus (1) zusammen mit Satz \ref{int-stetig} mit h(t)=t^2.

(3) Die Intergrierbarkeit folgt aus Satz \ref{int-stetig}, mit h(t)=|t|. Die Abschätzung ist elementar. Gleiches gilt für (4).

(5) Hier wendet man Satz \ref{int-stetig} auf h(t)=ct an.

</beweis>

<definition> Definition 4.3.6

Sei A\subset W eine Jordan-messbare Menge. Wir definieren die charakteristische Funktion \chi\_A von A durch

\chi\_A (x) = \left\{\ba{ll} 1 & x\in A\ 0 & x\notin A. \ea\right.

</definition>

<satz> Satz 4.3.7

Die charakteristische Funktion einer Jordan-messbaren Menge A\subset W ist Riemann-integrierbar und es gilt \int\_W \chi\_A = \text{vol} A.

</satz>

<beweis>

Für eine Zerlegung P von W ist

\overline{S}(\chi\_A,W;P) = \sum\_{W^\*\in P} \sup \chi\_A\big|\_{W^\*} \text{vol} W^\*

= \sum\_{W^\*\cap A \neq\emptyset}\text{vol} W^\*

\leq \sum\_{W^\*\cap \overline{A} \neq\emptyset} \text{vol}W^\* = \overline{S}(A;P)

und ebenso

\underline{S}(\chi\_A,W;P) = \sum\_{W^\*\in P} \inf \chi\_A\big|\_{W^\*} \text{vol} W^\*

= \sum\_{W^\*\subset A }\text{vol} W^\*

\geq \sum\_{W^\*\subset \text{Int} A} \text{vol}W^\* = \underline{S}(A;P).

Insgesamt gelten bei beliebiger Zerlegung die Abschätzungen

\underline{S}(A;P)\leq \underline{S}(\chi\_A,W;P) \leq \overline{S}(\chi\_A,W;P)\leq \overline{S}(A;P).

Bildet man in der ersten Abschätzung beiderseits das Supremum über alle Zerlegungen, so folgt

\sup\_P \underline{S}(A;P)\leq \sup\_P \underline{S}(\chi\_A,W;P)

und hiermit \underline{\text{vol}} A\leq \int\_W^u\chi\_A.

Bildet man in der letzten Abschätzung ebenso das Infimum über alle Zerlegungen, so findet man insgesamt

\underline{\text{vol}} A\leq \int\_W^u\chi\_A \leq \int\_W^o\chi\_A \leq \overline{\text{vol}} A.

Weil A nach Voraussetzung Jordan-messbar war, sind alle diese Größen gleich und wir erhalten die Behauptung.

</beweis>

<definition> Definition 4.3.8

Sind f: W \rightarrow \R Riemann-integrierbar und A\subset W Jordan-messbar, so ist \chi\_A \cdot f ebenfalls Riemann-integrierbar. Wir definieren dann das Integral von f über A gemäß

\int\_A f(x) dx := \int\_W (f\cdot \chi\_A ) dx.

</definition>

Wie man leicht nachprüft, übertragen sich die Rechenregeln aus Satz 4.3.5 sinngemäß auf dieses Integral ( Übungsaufgabe). Der folgende Satz besagt, dass für eine gleichmäßig konvergente Funktionenfolge Limes und Integration vertauscht werden können.

<satz> Satz 4.3.9 Vertauschung von Grenzwert und Integration

Sind f\_n: W\rightarrow \R Riemann-integrierbare Funktionen, die gleichmäßig gegen f konvergieren, so ist auch f Riemann-integrierbar und es gilt

\lim\_{n\rightarrow \infty} \int\_W f\_n(x) dx = \int\_W f(x) dx.

</satz>

<beweis>

Sei \epsilon>0 gegeben. Nach Voraussetzung existiert ein N mit

f(x)-\epsilon \leq f\_n(x) \leq f(x)+\epsilon \forall n\geq N, x\in W.

Bringt man in jeder Ungleichung das \epsilon auf die andere Seite, so folgt für jede Zerlegung P von W und jedes n\geq N

\overline{S}(f,W;P)\leq \overline{S}(f\_n,W;P)+\epsilon\text{vol} W,

\underline{S}(f\_n,W;P) - \epsilon\text{vol} W\leq \underline{S}(f,W;P).

Weil jedes f\_n Riemann-integrierbar ist, ergibt sich

\int^o\_W f\leq \int\_W f\_n+\epsilon\text{vol} W,

\int f\_n - \epsilon\text{vol} W\leq\int\_W^u f, also

0 \leq \int^o\_W f - \int^u\_W f \leq 2 \epsilon\text{vol} W.

Damit ist f Riemann-integrierbar und es gilt für alle n \geq N die Abschätzung

\int\_W f\_n -\epsilon\text{vol} W\leq \int\_W f\leq \int\_W f\_n +\epsilon\text{vol} W,

also \bigg|\int\_W f - \int\_W f\_n\bigg| \leq \epsilon \text{vol} W .

Läßt man nun n gegen \infty gehen, folgt die Behauptung.

</beweis>

<korollar> Korollar 4.3.10

Sind f\_n: W \rightarrow \R Riemann-integrierbar und ist die Funktionenreihe f(x)=\sum\_{i=1}^\infty f\_n(x) gleichmäßig konvergent, so ist f Riemann-integrierbar und darf gliedweise integriert werden, so dass

\int\_W f(x) dx = \sum\_{i=1}^\infty\int\_W f\_i(x) dx.

<\korollar>

Wir beenden diesen Abschnitt mit einigen Bemerkungen zu den Skalierungseigenschaften des Volumens. Ist A\subset \R^n eine beliebige Teilmenge und \lambda>0, so setzen wir

\lambda\cdot A := \{ \lambda a | a\in A\}.

Ist insbesondere W=[a\_1,b\_1]\times \dots\times [a\_n,b\_n] ein Quader, so gilt \lambda\cdot W

= [\lambda a\_1,\lambda b\_1]\times \dots\times [\lambda a\_n,\lambda b\_n]

und damit \text{vol} (\lambda\cdot W) = \lambda^n\text{vol} W.

Wir zeigen nun, dass diese Formel für jede Jordan-messbare Menge gültig bleibt.

<lemma> Lemma 4.3.11

Für jede Jordan-messbare Menge A\subset W\subset\R^n ist auch \lambda\cdot A Jordan-messbar und es gilt: \text{vol} (\lambda\cdot A)=\lambda^n\text{vol} A.

</lemma>

<beweis>

Für jede Zerlegung P von W mit Quadern W^\*\in P ist \lambda P eine Zerlegung von \lambda\cdot W mit Quadern \lambda W^\*. Dann aber gilt

\lambda^n\underline{S}(A;P) = \sum\_{W^\* \subset \text{Int} A} \lambda^n \cdot \text{vol} W^\* = \underline{S}(\lambda A;\lambda P),

\lambda^n\overline{S}(A;P) = \overline{S}(\lambda A;\lambda P).

Damit ergibt sich

\underline{\text{vol}} (\lambda\cdot A) = \sup\_P \underline{S}(\lambda A;\lambda P)

= \lambda^n\sup\_P \underline{S}(A;P) = \lambda^n\text{vol} A.

Mit einem ähnlichen Argument zeigt man \overline{\text{vol}} (\lambda\cdot A) = \lambda^n\text{vol} A und die Behauptung folgt.

</beweis>

## 4.4 Charakterisierung Riemann-integrierbarer Funktionen

In diesem Abschnitt sollen Riemann-integrierbare Funktionen charakterisiert werden. Hierzu benötigen wir einerseits die Lebesgue-Zahl einer Überdeckung, sowie andererseits den Begriff der Schwankung einer Funktion in einem Punkt.

<proposition> Proposition 4.4.1

Sei A\subset X eine kompakte Teilmenge des metrischen Raums (X,d) und

A\subset \bigcup\_{i\in I} U\_i eine offene \Überdeckung von A. Es gibt dann eine Zahl \kappa>0 mit folgender Eigenschaft: ist B\subset A eine Teilmenge von A mit \mathrm{diam} B:=\sup\{d(x,y) | x,y\in B\}<\kappa, so ist B vollständig in einer der Mengen U\_i enthalten. Die Zahl \kappa heißt Lebesgue-Zahl der \Überdeckung \bigcup\_{i \in I} U\_i.

</proposition>

<beweis> Übungsaufgabe </beweis>

Um den Begriff der Schwankung einer Funktion in einem Punkt einzuführen, betrachten wir einen metrischen Raum X und f:X\rightarrow \R eine beschränkte Funktion. Ist x\_0\in X, so betrachten wir zu gegebenem \delta >0

M(f,x\_0,\delta) := \sup\{ f(x) | x\in K (x\_0,\delta)\},

m(f,x\_0,\delta) := \inf\{ f(x) | x\in K (x\_0,\delta)\}.

Diese Zahlen existieren aufgrund der Beschränktheit von f.

Ist \delta\_1<\delta\_2, also K (x\_0,\delta\_1)\subset K (x\_0,\delta\_2), dann ist M(f,x\_0,\delta\_1)\leq M(f,x\_0,\delta\_2) und m(f,x\_0,\delta\_1)\geq m(f,x\_0,\delta\_2). Zusammen bedeutet dies

M(f,x\_0,\delta\_1) - m(f,x\_0,\delta\_1) \leq M(f,x\_0,\delta\_2) - m(f,x\_0,\delta\_2).

Die Funktion \delta \mapsto M(f,x\_0,\delta) - m(f,x\_0,\delta) ist also monoton sowie beschränkt und es existiert der Grenzwert

\lim\_{\delta\rightarrow 0^+} M(f,x\_0,\delta) - m(f,x\_0,\delta) =: o(f, x\_0) \geq 0.

<definition> Definition 4.4.2

Die Zahl o(f, x\_0) heißt die Schwankung oder Variation von f im Punkt x\_0.

</definition>

<beispiel> Beispiel 4.4.3

Die Funktion f:\R\rightarrow \R, f(0)=0, f(x)=\sin(1/x) für x\neq 0 ist in null unstetig und hat dort die Schwankung o(f,0)=2.

</beispiel>

<bemerkung> Bemerkung 4.4.4

* Ein verwandter Begriff, ist der der Funktion beschränkter Schwankung. Damit wird gemessen, wie oft eine Funktion auf einem gegebenen Intervall hin- und herschwankt.
* Der Begriff der Schwankung in einem Punkt hat über die im nächsten Abschnitt gegebene Anwendung noch weitere Bedeutung in der Analysis wie zum Beispiel bei der Definition sogenannter schwacher Ableitungen.

\end{itemize}

</bemerkung>

<beispiel> Beispiel 4.4.5

Hat f in x\_0 eine Sprungstelle wie etwa f(x)= x für x\geq 0, f(x)=x-1 für x<0, so ist die Schwankung genau die Höhe der Sprungstelle, denn für jedes \delta ist M(f,0,\delta) = \delta,

m(f,0,\delta) = -1 -\delta,

also M(f,0,\delta) - m(f,0,\delta) = 1+2\delta \stackrel{\delta\rightarrow 0}{\longrightarrow} 1.

</beispiel>

<satz> Satz 4.4.6

Sei X ein metrischer Raum. Für jede beschränkte Funktion f: X\rightarrow \R gilt:

(1) f ist genau dann in x\_0\in X stetig, falls o(f,x\_0)=0;

(2) Die Menge B\_\epsilon =\{ x\in X : o(f,x)\geq \epsilon\} ist in X abgeschlossen.

</satz>

<beweis>

(1) Ist f in x\_0 stetig und \epsilon>0 gegeben, so existiert ein \delta>0 mit K (x\_0,\delta)\subset f^{-1}(f(x\_0)-\epsilon,f(x\_0)+\epsilon). Also ist

M(f,x\_0,\delta)< f(x\_0)+\epsilon,

m(f,x\_0,\delta)> f(x\_0)-\epsilon,

also

M(f,x\_0,\delta) - m(f,x\_0,\delta)< 2\epsilon.

Aus der Monotonie der linken Seite folgt o(f,x\_0)=0. Gilt umgekehrt o(f,x\_0)=0, so existiert zu jedem \epsilon>0 ein \delta>0 mit M(f,x\_0,\delta) - m(f,x\_0,\delta) <\epsilon, also

\sup\_{x\in K (x\_0,\delta)} f(x)- \inf\_{x\in K (x\_0,\delta)} f(x) \leq \epsilon.

Man überlegt sich aber leicht, dass dies K (x\_0,\delta)\subset f^{-1}(K(f(x\_0),\epsilon)) impliziert, so dass f stetig in x\_0 stetig ist.

(2) Wir zeigen, dass X-B\_\epsilon offen ist. Für x\in X- B\_\epsilon gilt o(f,x)<\epsilon, so dass ein \delta>0 mit M(f,x,\delta) - m(f,x,\delta)<\epsilon existiert. Ist nun y\in K (x,\delta), so gibt es ein \delta\_y mit K (y,\delta\_y)\subset K (x,\delta). Damit ist insbesondere

M(f,y,\delta\_y)\leq M(f,x,\delta),

m(f,y,\delta\_y)\geq m(f,x,\delta

und damit

M(f,y,\delta\_y) - m(f,y,\delta\_y) < \epsilon.

Dann ist aber o(f,y) < \epsilon für alle y \in K (x,\delta), was bedeutet, dass die Kugel K (x,\delta) in

X-B\_\epsilon liegt.

</beweis>

Zur Vorbereitung für die Charakterisierung Riemann-integrierbarer Funktionen benötigen wir folgende

<definition> Definition 4.4.7

Eine {beliebige} Teilmenge A\subset\R^n hat das Lebesgue-Maß null, falls zu jeder Zahl \epsilon>0 eine Folge von Quadern W\_n existiert mit

A\subset \bigcup\_{i=1}^\infty \text{Int} W\_i

Und \sum\_{i=1}^\infty \text{vol} W\_i < \epsilon.

In diesem Fall schreibt man \mu\_L(A)=0.

</definition>

<bemerkung> Bemerkung 4.4.8

(1) Obwohl Mengen vom Lebesgue-Nullmengen "`klein"' sind, brauchen sie nicht beschränkt zu sein, im Gegensatz zu Jordan-Nullmengen.

(2) Statt mit Zerlegungen eines großen Quaders arbeitet man hier mit Ausschöpfungen durch beliebige Quader, wodurch der Begriff flexibler wird.

(3) Jede Jordan-messbare Menge vom Volumen null ist auch eine Menge vom Lebesgue-Maß null.

</bemerkung>

<satz> Satz 4.4.9

(1) Haben Mengen A\_n\subset \R^n das Lebesgue-Maß null, so hat auch die Vereinigung A=\bigcup\_{i=1}^\infty A\_i das Lebesgue-Maß null.

(2) Jede abzählbare Teilmenge A\subset\R^n hat das Lebesgue-Maß null.

(3) Eine kompakte Menge A hat genau dann das Lebesgue-Maß null, falls \text{vol} A=0 gilt.

</satz>

<beweis>

(1) Ist \epsilon>0, so existieren wegen \mu\_L(A\_i)=0 Quader W^1\_i, W^2\_i,... mit

A\_i\subset \bigcup\_{j=1}^\infty \text{Int} W^j\_i und

\sum\_{j=1}^\infty \text{vol} W^j\_i < \frac{\epsilon}{2^i}.

Dann gilt

A = \bigcup\_{i=1}^\infty A\_i \subset \bigcup\_{i=1}^\infty \bigcup\_{j=1}^\infty \text{Int} W^j\_i und

\sum\_{i=1}^\infty \sum\_{j=1}^\infty \text{vol} W^j\_i

< \sum\_{i=1}^\infty \frac{\epsilon}{2^i} = \epsilon,

was gleichbedeutend mit \mu\_L(A)=0 ist. Insbesondere folgt (2) sofort aus (1), denn die Mengen A\_i=\{a\_i\} haben trivialerweise das Lebesgue-Maß null.

(3) Ist A kompakt vom Lebesgue-Maß null, so existieren für jedes \epsilon>0 Quader W\_i mit

A\subset \bigcup\_{i=1}^\infty \text{Int} W\_i und \sum\_{i=1}^\infty \text{vol} W\_i < \epsilon.

Weil aber A kompakt ist, überdecken bereits endlich viele Quader W\_{i\_1},..., W\_{i\_k} die Menge A. Dann gilt für eine geeignete Partition P'

\overline{\text{vol}} A = \inf\_P \overline{S} (A,P)

= \inf\_P \sum\_{W^\* \in P, W^\* \cap \bar A\not= \emptyset} \text{vol} W^\* \leq \text{vol} \bigcup\_{W^\* \in P', W^\* \cap \bar A\not=\emptyset} W^\*

\leq \text{vol} W\_{i\_1}+... + \text{vol} W\_{i\_k}

\leq \sum\_{i=1}^\infty \text{vol} W\_i < \epsilon.

Damit ist \overline{\text{vol}} A =0. Ist umgekehrt A Jordan-messbar mit \text{vol} A=0, dann gilt immer \mu\_L(A)=0 infolge obiger Bemerkung.

</beweis>

Der folgende Satz gibt eine Charakterisierung für Riemann-integrierbare Funktionen und zeigt, dass die Theorie des Riemannschen Integrals von der Lebesgue-Theorie nicht unabhängig ist.

<satz> Satz 4.4.10

Eine beschränkte Funktion f: W\rightarrow \R ist genau dann Riemann-integrierbar, falls die Menge ihrer Unstetigkeitsstellen das Lebesgue-Maß null hat. Insbesondere ist jede beschränkte Funktion f: W\rightarrow \R mit höchstens abzählbar vielen Unstetigkeitsstellen Riemann-integrierbar.

</satz>

<beispiel> Beispiel 4.4.11

Sei W=[0,1] und f: W\rightarrow \R definiert durch f(0)=1, f(x)=1-1/n auf ]1/(n+1), 1/n]. Nach dem vorherigen Satz ist diese Funktion Riemann-integrierbar. Es gilt weiterhin

\int\_{[0,1]} f = \sum\_{n=1}^\infty \left[1-\frac{1}{n}\right] \left[\frac{1}{n} - \frac{1}{n+1}\right]

= \sum\_{n=1}^\infty \frac{n-1}{n^2(n+1)} < \sum\_{n=1}^\infty \frac{1}{n^2}

= \zeta(2) = \frac{\pi^2}{6}.

</beispiel>

<beweis> [Beweis von Satz 4.4.10]

Sei B\subset W die Menge der Unstetigkeitsstellen von f und M=\max\_{x\in W} |f(x)|.

(1) Angenommen, \mu\_L(B)=0. Wir können also zu \epsilon >0 eine Folge von Quadern W\_i wählen mit B\subset \bigcup\_{i=1}^\infty \text{Int} W\_i \text{ und } \sum\_{i=1}^\infty \text{vol} W\_i

< \epsilon. Ist f in x stetig, x\in W - B, so hat f dort die Schwankung null und es existiert ein von x abhängiges \delta\_x>0 mit

\sup\_{y\in K (x,\delta\_x)} f(y) - \inf\_{y\in K (x,\delta\_x)} f(y) < \epsilon.

Die Familie \{ \text{Int} W\_i,K (x,\delta\_x)\} mit x\in W-B ist eine offene Überdeckung von W. Sei \kappa die Lebesgue-Zahl dieser Überdeckung. Wir wählen eine Zerlegung P von W, die so fein ist, dass für alle Teilwürfel \mathrm{diam} W^\*<\kappa gilt. Dann zerfällt P in die Partitionen

P\_1 = \{ W^\*\in P | W^\*\text{ liegt in einer der Mengen } \text{Int} W\_i\},

P\_2 = \{ W^\*\in P | W^\*\notin P\_1\}.

Ist nun W^\* ein Quader, der nicht zu P\_1 gehört, also nicht vollständig in einer der Mengen \text{Int} W\_i enthalten ist, so existiert nach Eigenschaft der Lebesgue-Zahl eine Kugel K (x,\delta\_x) mit x\in W-B, in der W^\* liegt. Für P\_1 gilt zudem die Abschätzung

\sum\_{W^\*\in P\_1}\text{vol} W^\*\leq \sum\_{i=1}^\infty\text{vol} W\_i < \epsilon.

Daraus folgt \overline{S}(f,W;P)-\underline{S}(f,W;P)

=\sum\_{W^\*\in P\_1} \underbrace{\left[ \sup f\big|\_{W^\*}

 - \inf f\big|\_{W^\*}\right]}\_{\leq 2M} \text{vol} W^\*

+ \sum\_{W^\*\in P\_2}\underbrace{\left[ \sup f\big|\_{W^\*}

- \inf f\big|\_{W^\*}\right]}\_{\leq \epsilon} \text{vol} W^\*\\

 \leq 2M \underbrace{\sum\_{W^\*\in P\_1}\text{vol} W^\*}\_{\leq \epsilon}

+ \epsilon \underbrace{\sum\_{W^\*\in P\_2} \text{vol} W^\*}\_{\leq \text{vol} W }

\leq \epsilon(2M+\text{vol} W).

Zu jedem \epsilon existiert demnach eine Zerlegung P mit

\overline{S}(f,W;P)-\underline{S}(f,W;P)\leq \epsilon\cdotconst.

Damit ist \int^o\_W f = \int^u\_W f , also f Riemann-integrierbar.

(2) Sei umgekehrt f Riemann-integrierbar und betrachte die nach Satz 4.4.6 abgeschlossenen Mengen B\_\epsilon = \{ x\in W | o(f,x) \geq \epsilon\}.

Weil f genau dann stetig in x ist, falls o(f,x)=0 ist, gilt offenbar

B = \bigcup\_{n\in\N} B\_{1/n}.

Um also \mu\_L(B)=0 zu zeigen, genügt es nach Satz 4.4.9 \mu\_L(B\_{1/n})=0 für jedes n zu zeigen. Wir fixieren ein n\in\N und ein \epsilon>0. Da f Riemann-integrierbar ist, können wir eine Zerlegung P von W mit

\overline{S}(f,W;P)-\underline{S}(f,W;P) < \frac{\epsilon}{n}

finden. Sei

P\_1 = \{ W^\*\in P | W^\*\cap B\_{1/n}\neq\emptyset\}.

Wir zeigen zunächst, dass \sum\_{W^\*\in P\_1}\text{vol} W^\*<\epsilon. In der Tat, ist W^\*\in P\_1, so gilt nach Definition der Schwankung und weil W^\*\cap B\_{1/n}\neq\emptyset, dass

\sup f\big|\_{W^\*}-\inf f\big|\_{W^\*} \geq \frac{1}{n}.

Dies impliziert

\frac{1}{n}\sum\_{W^\*\in P\_1}\text{vol} W^\*

\leq \sum\_{W^\*\in P\_1} \left[\sup f\big|\_{W^\*}-\inf f\big|\_{W^\*}\right] \text{vol} W^\*

\leq \sum\_{W^\*\in P}\left[\sup f\big|\_{W^\*}-\inf f\big|\_{W^\*}\right]\text{vol} W^\*

= \overline{S}(f,W;P)-\underline{S}(f,W;P) < \frac{\epsilon}{n}.

Aufgrund von B\_{1/n} \subset \bigcup\_{W^\*\in P\_1} W^\*

folgt mit \overline{ B\_{1/n}} = B\_{1/n}

\overline{\text{vol}} B\_{1/n} = \inf\_{P'} \sum\_{W^\* \in P', W^\* \cap \overline{B\_{1/n}} \not=\emptyset} \text{vol} W^\* \leq \sum\_{W^\*\in P\_1}\text{vol} W^\* < \epsilon.

Damit ist \overline{\text{vol}} B\_{1/n}=0 und erst recht \mu\_L (B\_{1/n})=0.

</beweis>

Man beachte im obigen Beweis, dass die Charakterisierung von Stetigkeit mittels der Schwankung über die Betrachtung von \sup f - \inf f erfolgt und aufgrund dessen mit \overline S(f,W;P) - \underline S(f,W;P) in Verbindung gebracht werden kann.

<korollar> Korollar 4.4.12

Jede monotone Funktion f: [a,b]\rightarrow \R ist Riemann-integrierbar.

<\korollar>

<beweis>

Eine auf einem kompakten intervall [a,b] definierte monotone Funktion ist dort notwendig beschränkt. Da jede monotone Funktion auf einem kompakten Intervall höchstens abzählbarviele Unstetigkeitsstellen hat (Übungsaufgabe!), folgt die Behauptung.

</beweis>

<satz> Satz 4.4.13 Erster Mittelwertsatz der Integralrechnung

Ist f: W\rightarrow \R stetig, g: W\rightarrow \R Riemann-integrierbar und g\geq 0, dann existiert ein Punkt x\_0\in W mit

\int\_W f(x)g(x) dx = f(x\_0)\int\_W g(x) dx.

Ist insbesondere g=\chi\_A die charakteristische Funktion einer Jordan-messbaren Menge, so existiert ein Punkt x\_0\in W mit

\int\_A f(x) dx = f(x\_0)\text{vol} A.

</satz>

<beweis>

Angenommen, es gilt \int\_W g(x) dx=0. Wegen g\geq 0 folgt dann mit m:=\min\{f(x) | x\in W\} und M=\max\{f(x) | x\in W\} m g(x)\leq f(x)g(x)\leq Mg(x), also

m \int\_W g(x)dx \leq \int\_W f(x)g(x) dx\leq M\int\_W g(x) dx.

Beachte hierbei, dass f als stetige Funktion auf W sein Minimum und Maximum annimmt. Nach Voraussetzung sind aber die linke und die rechte Seite null, es folgt also \int\_W f(x)g(x) \d x=0 und die Behauptung ist trivial. Sei andererseits \int\_W g(x) dx>0. Wieder gilt

m \int\_W g(x) dx \leq \int\_W f(x)g(x) dx\leq M\int\_W g(x) dx, also

m\leq \frac{\int\_W f(x) g(x) dx}{\int\_W g(x) dx } \leq M.

Damit existieren zwei Punkte x\_1,x\_2 \in W derart, dass f(x\_1) \leq \lambda \leq f(x\_2), wobei \lambda den in der obigen Ungleichung auftretenden Quotienten bezeichnet. Desweiteren hat f als stetige Funktion die Darboux-Eigenschaft \footnote{Die Darboux-Eigenschaft besagt, dass zwischen zwei Punkten a,c mit f(a)<\lambda, f(c)>\lambda ein b existiert mit f(b)=\lambda.}

und somit existiert ein x\_0\in W mit

\lambda = \frac{\int\_W f(x) g(x) dx}{\int\_W g(x) dx } = f(x\_0)

und wir erhalten die Behauptung.

</beweis>

## 4.5 Das Riemann'sche Integral in \R

Ist f: [a,b]\rightarrow \R eine Riemann-integrierbare Funktion, so schreiben wir

\int\_a^b f(x) dx := \int\_{[a,b]} f(x) dx

für das Integral über den eindimensionalen Quader [a,b]. Dabei soll immer a\leq b sein, so dass etwa ein Integral mit vertauschten Integrationsgrenzen keinen Sinn macht. Wir untersuchen als nächstes die Funktion, die entsteht, wenn man eine der Intervallgrenzen als variabel auffaßt.

<satz> Satz 4.5.1

Zu einer Riemann-integrierbaren Funktion f: [a,b]\rightarrow \R definieren wir F(x):=\int\_a^x f(t) dt. Dann gilt:

(1) F: [a,b]\rightarrow \R ist gleichmäßig stetig.

(2) Ist f in x\_0\in [a,b] stetig, so ist F dort differenzierbar und es gilt F'(x\_0)=f(x\_0).

</satz>

<beweis>

Für M:=\sup\_{t\in[a,b]} f(t) gilt (y<x) |F(x)-F(y)| = \bigg|\int\_a^x f(t) dt - \int\_a^y f(t) dt\bigg|

= \bigg| \int\_y^x f(t) dt\bigg| \leq \int\_y^x |f(t)| dt\leq M(x-y).

Diese Abschätzung besagt gerade, dass F auf [a,b] gleichmäßig stetig ist.

Sei nun f in x\_0 stetig, \epsilon>0 und \delta>0 dergestalt, dass |f(t)-f(x\_0)|<\epsilon für alle t mit |t-x\_0|<\delta. In diesem Fall gilt auf dem Intervall [x\_0,t] (analoges Argument auf dem Intervall [t,x\_0])

\bigg| \frac{F(t)-F(x\_0)}{t-x\_0} - f(x\_0) \bigg|

= \bigg| \frac{\int\_a^t f(u) du - \int\_a^{x\_0} f(u) du}{t-x\_0}

-\frac{1}{t-x\_0}\int\_{x\_0}^t f(x\_0) du \bigg|

= \bigg|\frac{1}{t-x\_0}\int\_{x\_0}^t [f(u)-f(x\_0)]du \bigg|

< \frac{1}{t-x\_0} \epsilon (t-x\_0) = \epsilon.

Folglich existiert F'(x\_0) und ist gleich f(x\_0).

</beweis>

Der folgende Satz stellt den Zusammenhang zwischen der Differential- und der Integralrechnung in einer Variablen her.

<satz> Satz 4.5.2 Hauptsatz der Differential- und Integralrechnung

Ist f: [a,b]\rightarrow \R Riemann-integrierbar und existiert eine differenzierbare Funktion F mit F'(x)=f(x), so gilt

\int\_a^b f(t) dt = F(b) - F(a).

</satz>

<beweis>

Wir zerlegen das gegebene Intervall mit einer Partition P gemäß a=x\_0<x\_1<... <x\_k=b und wenden auf F\big|\_{[x\_{i-1},x\_i]} den Mittelwertsatz der Differentialrechnung an. Es existieren demnach Zahlen t\_i\in ]x\_{i-1},x\_i[ mit

F(x\_i)-F(x\_{i-1}) = F'(t\_i)(x\_i-x\_{i-1}) = f(t\_i)(x\_i-x\_{i-1}).

Damit folgt

\overline{S}(f,P)= \sum\_{i=1}^k \sup f\big|\_{[x\_{i-1},x\_i]}(x\_i-x\_{i-1})

 \geq \sum\_{i=1}^k \frac{F(x\_i)-F(x\_{i-1})}{x\_i-x\_{i\_1}} (x\_i-x\_{i-1}) = F(b) - F(a)

und analog

\underline{S}(f,P)\leq F(b) - F(a).

Dann ist aber für jede Zerlegung

\underline{S}(f,P)\leq F(b) - F(a)\leq \overline{S}(f,P),

also \int^u f=\sup \underline{S}(f,P)\leq F(b) - F(a)\leq \inf\overline{S}(f,P)= \int^o f.

Da f Riemann-integrierbar ist, sind Ober- und Unterintegral gleich und damit gleich F(b)-F(a).

</beweis>

<definition> Definition 4.5.3

Ist f: [a,b]\rightarrow \R eine beliebige Funktion, so heißt jede differenzierbare Funktion F: [a,b]\rightarrow \R mit F'=f eine Stammfunktion von f.

</definition>

Falls f eine Stammfunktion besitzt, so ist es am einfachsten, das Integral über eine Stammfunktion F, und nicht über Riemann'sche Summen zu berechnen, also als

\int\_a^b f(t) dt = F(b)-F(a).

<lemma> Lemma 4.5.4

(1) Jede stetige Funktion f besitzt eine Stammfunktion.

(2) Sind F, G Stammfunktionen der gleichen Riemann-integrierbaren Funktion f: [a,b]\rightarrow \R, so gilt F(x)=G(x)+c für eine Konstante c und alle x\in [a,b].

</lemma>

<beweis>

Die erste Aussage folgt sofort aus Satz 4.5.1, die zweite folgt aus der Tatsache, dass nach dem Mittelwertsatz der Differentialrechnung eine differenzierbare Funktion H(=F-G) mit H'=0 konstant ist.

</beweis>

Die Bestimmung von Stammfunktionen stellt ein tiefliegendes Problem dar. Wie wir später anhand der sogenannten elliptischen Funktionen sehen werden, können Funktionen Stammfunktionen besitzen, die jedoch keine Darstellung mittels elementarer Funktionen haben. Tatsächlich ist das Problem systematischer Natur. So besitzt nicht jede Riemann-integrierbare Funktion eine Stammfunktion und nicht jede Funktion, die eine Stammfunktion besitzt, ist Riemann-integrierbar.

<beispiel> Beispiel 4.5.5

Nicht jede Riemann-integrierbare Funktion besitzt eine Stammfunktion. Ist nämlich F auf [a,b] differenzierbar, so hat F' dort die Darboux-Eigenschaft -- auch wenn F' nicht stetig ist (Übungsaufgabe). Damit gilt: Jede Riemann-integrierbare Funktion, die eine Stammfunktion hat, besitzt die Darboux-Eigenschaft. Es gibt aber Riemann-integrierbare Funktionen, die diese Eigenschaft nicht haben. Zum Beispiel ist dies für

f(t) = \left\{\ba{ll} 0 & t\leq 1/2\ 1 & t>1/2 \ea\right.

der Fall.

</beispiel>

<beispiel> Beispiel 4.5.6

Nicht jede Funktion, die eine Stammfunktion besitzt, ist Riemann-integrierbar. Zum Beispiel ist g(x)=x^{3/2}\sin(1/x) auf [0,1] eine anständige differenzierbare Funktion, aber ihre Ableitung

g'(x) = \frac{3}{2}\sqrt{x}\sin\frac{1}{x} -\frac{1}{\sqrt{x}}\cos \frac{1}{x}

ist auf [0,1] unbeschränkt, also auf gar keinen Fall Riemann-integrierbar.

Wie man desweiteren in der Lebesgue-Theorie zeigt, ist die Ableitung einer Funktion genau dann Lebesgue-integrierbar, falls sie absolut stetig ist (das ist stärker als stetig). Ein analoger Satz scheint für in der Riemannsche Integrationstheorie jedoch nicht bekannt zu sein.

</beispiel>

<definition> Definition 4.5.7

Sei f eine Riemann-integrierbare Funktion, die eine Stammfunktion besitzt. Unter dem unbestimmten Integral \int f(t)dt versteht man die Familie aller Stammfunktionen von f. In diesem Sinne gilt:

(1) \int x^n dx = \frac{x^{n+1}}{n+1}+c (n\neq -1)

(2) \int \frac{1}{x} dx= \ln x +c, x>0

(3) \int\sin x dx = -\cos x + c

(4) \int\sin^2 x dx = \frac{1}{2} x - \frac 12 \sin x \cos x

(5) \int\cos x dx = \sin x + c

(6) \int\tan x dx = -\ln \cos x + c, x\in [0,\pi/2]

(7) \int e^x dx = e^x + c

(8) \int \ln x dx = x\ln x - x + c

(9) \int \frac{1}{1+x^2} dx = \arctan x + c

(10) \int \frac{1}{1-x^2} dx = \left\{\ba{ll} \mathrm{arctanh} x + c

= \frac{1}{2}\ln\frac{1+x}{1-x} + c & |x|<1

\mathrm{arccoth} x +c = \frac{1}{2}\ln\frac{x+1}{x-1} + c

 |x|>1 \ea\right.

(11) \int \sqrt{1-x^2} dx = \frac{1}{2}(x \sqrt{1-x^2} +\arcsin x) (|x|<1)

(12) \int \frac{1}{\sqrt{1-x^2}} dx = \arcsin x (|x|<1)

(13) \int \frac{1}{\sqrt{1+x^2}} dx = \mathrm{arsinh} x +c = \ln(x+\sqrt{1+x^2}) +c'

(14) \int \frac{1}{\sqrt{x^2-1}} dx = \mathrm{arcosh} x +c = \ln(x+\sqrt{x^2-1}) +c'

</definition>

Aus den Rechenregeln für das Differenzieren ergeben sich folgende Rechenregeln für Integrale:

<satz> Satz 4.5.8

(1) Partielle Integration: Sind f und g differenzierbar auf [a,b], so gilt

\int\_a^b f(x)g'(x) dx = f(x)g(x)\big|\_a^b - \int\_a^b f'(x) g(x) dx .

(2) Substitutionsregel: Ist g: [a,b]\rightarrow \R Riemann-integrierbar mit Stammfunktion G, f: [\alpha,\beta]\rightarrow [a,b] stetig differenzierbar und f(\alpha)=a, f(\beta)=b, so hat (g\circ f)f': [\alpha,\beta]\rightarrow \R als Stammfunktion G\circ f und es gilt

\int\_\alpha^\beta (g\circ f)(x)f'(x) dx = (G\circ f) (\beta) - (G\circ f)(\alpha)

= G(b)-G(a) = \int\_a ^b g(y) dy.

</satz>

<beweis>

Folgt direkt aus der Produkt- bzw. Kettenregel.

</beweis>

Für unbestimmte Integrale gelten entsprechende Rechenregeln.

<bemerkung> Bemerkung 4.5.9

Die Substitutionsregel wendet man intuitiv wie folgt an. Ersetze die Variable y durch y=f(x) und schreibe die Integrationsgrenzen als a=f(\alpha),b=f(\beta). Dann ist dy = f'(x)dx und es gilt

\int\_a^b \underbrace{g(y)}\_{=g(f(x))} \underbrace{dy}\_{= f'(x)dx} = \int\_\alpha^\beta g(f(x)) f'(x)dx.

</bemerkung>

### Das Wallis'sche Produkt

Im folgenden werden wir eine Rekursionsformel für I\_n :=\int \sin^n x dx herleiten. So berechnet man mittels partieller Integration für n \geq 2

I\_n &=& -\int \sin^{n-1}(x)\cos' x dx {=} - \sin^{n-1}(x)\cos x + (n-1)\int \sin^{n-2}(x)\cos^2 (x) dx

= - \sin^{n-1}(x)\cos x + (n-1)\int \sin^{n-2}(x) (1-\sin^2 (x)) dx

= - \sin^{n-1}(x)\cos x - (n-1)I\_n + (n-1)I\_{n-2}.

Also gilt I\_0=x, I\_1=-\cos x und demnach I\_n = - \frac{1}{n}\sin^{n-1}( x)\cos x +\frac{n-1}{n}I\_{n-2}.

Wir berechnen nun den Wert von I\_n bei Integration über das Intervall [0,\pi/2] jeweils für geraden und ungeraden Index. Wegen \sin 0=0, \cos \pi/2=0 folgt

I\_n(0) = \frac{n-1}{n}I\_{n-2}(0), I\_n(\pi/2) = \frac{n-1}{n} I\_{n-2}(\pi/2).

Zudem ist I\_0(0)=0, I\_1(0)=-1, I\_0(\pi/2)=\pi/2, I\_1(\pi/2)=0, so dass insgesamt

l\_{2n}(0) = \frac{2n-1}{2n}I\_{2n-2}(0) = \frac{2n-1}{2n}\cdot\frac{2n-3}{2n-2}\cdot

... \cdot\frac{1}{2} \cdot I\_{0}(0) = 0,

I\_{2n}(\pi/2)& =& \frac{2n-1}{2n}I\_{2n-2}(\pi/2)

= \frac{2n-1}{2n}\cdot\frac{2n-3}{2n-2}\cdot ... \cdot\frac{1}{2} \cdot \frac{\pi}{2}.

Ebenso beweist man

I\_{2n+1}(\pi/2) = 0,

I\_{2n+1}(0) = \frac{2n}{2n+1}\cdot\frac{2n-2}{2n-1}\cdot... \cdot (-1).

Daraus folgt

<eq:3>

\int\_0^{\pi/2} \sin^{2n}x dx = I\_{2n}(\pi/2)- I\_{2n}(0)

= \frac{1\cdot 3\cdot ...\cdot (2n-1)} {2\cdot 4\cdot...\cdot (2n)}

\cdot\frac{\pi}{2},

\int\_0^{\pi/2} \sin^{2n+1}x dx& =& I\_{2n+1}(\pi/2)- I\_{2n+1}(0)

= \frac{2\cdot 4\cdot ...\cdot (2n)}{3\cdot 5\cdot...\cdot (2n+1)}.

</eq:3>

Mit Kenntnis dieser beiden Integrale wollen wir folgendes Ergebnis herleiten, das zuerst von John Wallis \footnote{ Britischer Mathematiker (1616-1703), der ab 1649 Professor an der Universität Oxford war. Berufen wurde er für seine Leistungen beim Entziffern geheimer Botschaften.} bewiesen wurde:

<proposition> Proposition 4.5.10 Wallis-Formel, 1656

\pi = \lim\_{n\rightarrow \infty} \frac{1}{n}

\left[\frac{2\cdot 4\cdot...\cdot (2n)} {1\cdot 3\cdot...\cdot (2n-1)} \right]^2.

</proposition>

<beweis>

Mit den Integrationsformeln \eqref{eq:3} erhalten wir zunächst

\frac{\pi}{2} = \left[\frac{2\cdot 4\cdot...\cdot (2n)} {1\cdot 3\cdot...\cdot (2n-1)} \right]^2 \frac{1}{2n+1} \frac{\int\_0^{\pi/2} \sin^{2n}x dx}{\int\_0^{\pi/2} \sin^{2n+1}x dx}

= \frac{1}{n}\left[\frac{2\cdot 4\cdot...\cdot (2n)}{1\cdot 3\cdot...\cdot (2n-1)} \right]^2 \underbrace{\frac{n}{2n+1}}\_{\rightarrow 1/2} \cdot

\frac{\int\_0^{\pi/2} \sin^{2n}x dx}{\int\_0^{\pi/2} \sin^{2n+1}x dx}.

Wenn wir zeigen können, dass das Verhältnis der Integrale für n\rightarrow \infty gegen 1 konvergiert, so folgt die Behauptung.

Für 0\leq x\leq \pi/2 haben wir jedoch 0\leq\sin^{2n+1} x\leq\sin^{2n}x\leq\sin^{2n-1}x,

also nach Integration

0 \leq \int\_{0}^{\pi/2}\sin^{2n+1}x dx\leq \int\_0^{\pi/2}\sin^{2n}xdx\leq\int\_0^{\pi/2}\sin^{2n-1}x dx

und weiter nach Bilden des Verhältnisses

1 \leq \frac{\int\_0^{\pi/2}\sin^{2n}xdx}{\int\_{0}^{\pi/2}\sin^{2n+1}x dx}

\leq \frac{\int\_0^{\pi/2}\sin^{2n-1}xdx}{\int\_{0}^{\pi/2}\sin^{2n+1}x dx}

= \frac{\frac{2\cdot 4\cdot ...\cdot (2n-2)}{3\cdot 5\cdot...\cdot (2n-1)}}

{\frac{2\cdot 4\cdot ...\cdot (2n)}{3\cdot 5\cdot...\cdot (2n+1)}}

= \frac{2n+1}{2n}= 1+\frac{1}{2n}.

Damit ist die Wallis-Formel bewiesen.

</beweis>

Weil man den in der Wallis-Formel auftretenden Quotienten noch gemäß

\frac{2\cdot 4\cdot...\cdot (2n)}{1\cdot 3\cdot...\cdot (2n-1)}

 = \frac{(2^n n!)2\cdot 4\cdot...\cdot (2n)}{(2n)!} = \frac{(2^n n!)^2}{(2n)!}

umformen kann, kann besagte Formel äquivalent auch geschrieben werden als

\sqrt{\pi} = \lim\_{n\rightarrow \infty} \frac{(n!)^2 2^{2n}}{\sqrt{n}(2n)!}.

<beispiel> Beispiel 4.5.11

Ein weiteres Integral, welches sich mittels Substitution auf die Integrale \eqref{eq:3} zurückführen läßt, ist \int\_0^1 (1-x^2)^n dx. Wir substitutieren x=\cos t, dx=-\sin tdt mit Integrationsgrenzen x=0 \leftrightarrow t=\pi/2, x=1\leftrightarrow t=0, so dass

\int\_0^1 (1-x^2)^ndx = -\int\_{\pi/2}^0 (1-\cos^2 t)^n\sin t dt

= \int\_0^{\pi/2}\sin^{2n+1}t dt = \frac{2\cdot 4 \cdot... (2n)}{3\cdot 5\cdot...(2n+1)}.

</beispiel>

<beispiel> Beispiel 4.5.12

Auch das Integral \int\_0^{\pi/2} (1-k^2\sin^2 x)^{-1/2}dx mit |k|<1 lässt sich nach Reihenentwicklung auf die Integrale \eqref{eq:3} zurückführen. So lautet die in der Analysis I besprochene Binomialreihe

\frac{1}{\sqrt{1+y}} = 1-\frac{1}{2}y+\frac{1\cdot 3}{2\cdot 4}y^2

-\frac{1\cdot 3\cdot 5}{2\cdot 4\cdot 6}y^3\pm....

Für y=-k^2\sin^2 x erhalten wir somit

\frac{1}{\sqrt{1-k^2\sin^2 x}} = 1+ \frac{1}{2}k^2\sin^2 x +\frac{1\cdot 3}{2\cdot 4}k^4\sin^4x+

 \frac{1\cdot 3\cdot 5}{2\cdot 4\cdot 6}k^6\sin^6x+...

Nach dem Weierstraß-Kriterium ist diese Reihe wegen |\sin x|\leq 1 und |k|< 1 gleichmäßig konvergent, alle Terme sind polynomial, also klarerweise Riemann-integrierbar. Damit ist Korollar 4.3.10 anwendbar und es darf gliedweise integriert werden, so dass

\int\_0^{\pi/2} \frac{1}{\sqrt{1-k^2\sin^2 x}} dx

= \int\_0^{\pi/2} dx+ \frac{k^2}{2}\int\_0^{\pi/2}\sin^2 x dx

+ \frac{3k^4}{2\cdot 4} \int\_0^{\pi/2} \sin^4x dx

+ \frac{3\cdot 5\cdot k^6}{2\cdot 4\cdot 6} \int\_0^{\pi/2}\sin^6x dx+...

= \frac{\pi}{2}\left( 1+\left[\frac{1}{2}\right]^2k^2 + \left[\frac{1\cdot 3}{2\cdot 4}\right]^2k^4

+ \left[\frac{1\cdot 3\cdot 5}{2\cdot 4\cdot 6}\right]^2k^6 +...\right)

</beispiel>

<satz> Satz 4.5.13 Zweiter Mittelwertsatz der Integralrechnung

Sei f: [a,b]\rightarrow \R stetig, g: [a,b]\rightarrow \R monoton wachsend und stetig differenzierbar. Dann existiert ein Punkt \xi\in [a,b] mit

\int\_a^b f(x)g(x) dx = g(a)\int\_a^\xi f(x)dx+ g(b)\int\_\xi^b f(x) dx.

</satz>

<beweis>

Sei F die Stammfunktion der stetigen Funktion f. Weil g monoton wachsend und differenzierbar ist, gilt g'\geq 0. Da F als Stammfunktion stetig und g' als stetige Funktion Riemann-integrierbar ist, kann der erste Mittelwertsatz auf F(x)g'(x) angewandt werden. Demnach existiert ein \xi\in [a,b] mit

\int\_a^b F(x)g'(x) dx = F(\xi)\int\_a^b g'(x) dx.

Hiermit folgt aber

\int\_a^b f(x)g(x)dx = \int\_a^b F'(x)g(x) dx = F(x)g(x)\big|\_a^b - \int\_a^b F(x)g'(x) dx

= F(b)g(b)-F(a)g(a)-F(\xi)\int\_a^b g'(x) dx

= F(b)g(b)-F(a)g(a)-F(\xi) (g(b)-g(a))

= g(a)(F(\xi)-F(a)) + g(b)(F(b)-F(\xi))

= g(a)\int\_a^\xi f(x) dx+ g(b)\int\_{\xi}^b f(x) dx

und wir erhalten die Behauptung.

</beweis>

<bemerkung> Bemerkung 4.5.14

Der zweite Mittelwertsatz gilt auch, wenn f nur integrierbar und g nur monoton ist. Der Beweis ist dann allerdings erheblich schwieriger, vgl.Mangoldt-Knopp, Einführung in die höhere Mathematik III, Abschnitt 44.

</bemerkung>

<beispiel> Beispiel 4.5.15

Wir zeigen, dass für 0<a<b die Beziehung

\lim\_{n\rightarrow \infty} \int\_a^b \frac{\sin nx}{x} dx = 0 gilt.

Sei dazu f(x)=\sin nx, g(x)=-1/x, wobei n nicht ganzzahlig,

sondern nur positiv sein muss. Nach zweitem Mittelwertsatz gilt für ein \xi\in [a,b]

-\int\_a^b \frac{\sin nx}{x} dx = - \frac{1}{a}\int\_a^\xi \sin nx dx - \frac{1}{b}\int\_\xi^b \sin nx dx.

Dies impliziert

\bigg|\int\_a^b \frac{\sin nx}{x} dx \bigg|

\leq \frac{1}{an}\big|\cos n\xi -\cos n a\big|+ \frac{1}{bn}\big|\cos nb - \cos n\xi \big|

\leq \frac{2}{an}+\frac{2}{bn},

so dass sich im Limes n\rightarrow \infty die Behauptung ergibt.

</beispiel>

## 4.6 Der Satz von Fubini und seine Anwendungen

Sind W\_1\subset\R^n und W\_2\subset \R^m Quader, so ist auch W\_1 \times W\_2 ein Quader, und zwar in \R^{n+m}. Der Satz von Fubini besagt, dass es meistens möglich ist, das Integral einer Funktion f: W\_1\times W\_2\rightarrow \R sukzessive, also erst über W\_1 und anschließend über W\_2 zu berechnen. Dieser Satz ist bei der praktischen Berechnung mehr-dimensionaler Integrale von großer Bedeutung.

<satz> Satz 4.6.1 Satz von Fubini

Sei f: W\_1\times W\_2\rightarrow \R eine beschränkte und Riemann-integrierbare Funktion. Für jedes x\in W\_1 betrachten wir die Funktion g\_x: W\_2\rightarrow \R, g\_x(y)= f(x,y) und setzen

\mathcal{U}(x) := \int^u\_{W\_2} g\_x(y) dy = \int^u\_{W\_2} f(x,y) dy,

\mathcal{O}(x) := \int^o\_{W\_2} g\_x(y) dy = \int^o\_{W\_2} f(x,y) dy.

Dann sind \mathcal{U}, \mathcal{O}: W\_1 \rightarrow \R Riemann-integrierbar und

\int\limits\_{W\_1\times W\_2} f(x,y)dxdy

= \int\_{W\_1}\mathcal{U}(x) dx = \int\_{W\_1} \left[\int^u\_{W\_2} f(x,y) dy\right] dx

= \int\_{W\_1}\mathcal{O}(x) dx = \int\_{W\_1} \left[\int^o\_{W\_2} f(x,y) dy\right] dx.

</satz>

<beweis>

Sei P=(P\_1,...,P\_n) eine Zerlegung von W\_1 mit den Quadern W,

P'=(P'\_1,...,P'\_m) eine Zerlegung von W\_2 mit den Quadern W'.

Durch Zusammmensetzen erhält man eine Zerlegung Q von W\_1 \times W\_2 mit den Quadern W \times W'. Wir haben zunächst

\underline{S}(f,W\_1\times W\_2;Q)

= \sum\_{W\in P, W'\in P'} \inf f\big|\_{W\times W'}\text{vol}W\text{vol} W'.

Ist aber x \in W, so gilt \inf f \big|\_{W\times W'}\leq \inf g\_x \big|\_{W'}. Daraus folgt für festes W und x\in W

\sum\_{W'\in P'}\inf f\big|\_{W\times W'}\text{vol}W'

\leq \sum\_{W'\in P'}\inf g\_x\big|\_{W'}\text{vol}W'

= \underline{S}(g\_x, W\_2;P') \leq \int^u\_{W\_2} g\_x = \mathcal{U}(x).

Weil dies für alle x \in W zutrifft, gilt auch

\sum\_{W'\in P'}\inf f\big|\_{W\times W'}\text{vol}W' \leq \inf\_{x\in W}\mathcal{U}(x).

Damit folgt für die ursprüngliche Untersumme

\underline{S}(f,W\_1\times W\_2;Q)

\leq \sum\_{W\in P} \inf\_{x\in W}\mathcal{U}(x) \text{vol} W \leq \underline{S}(\mathcal{U},W\_1;P).

Analog zeigt man \overline{S}(f,W\_1\times W\_2;Q) \geq \overline{S}(\mathcal{O},W\_1;P).

Weil natürlich \mathcal{U}\leq\mathcal{O} gilt, haben wir damit insgesamt

\underline{S}(f,W\_1\times W\_2;Q) \leq \underline{S}(\mathcal{U},W\_1;P)

\leq \left\{\ba{c} \overline{S}(\mathcal{U},W\_1;P)

\underline{S}(\mathcal{O},W\_1;P) \ea\right\} \leq \overline{S}(\mathcal{O},W\_1;P)

\leq \overline{S}(f,W\_1\times W\_2;Q).

Da f auf W\_1\times W\_2 Riemann-integrierbar ist, gilt

\int\limits\_{W\_1\times W\_2} f = \sup\_Q \underline{S}(f,W\_1\times W\_2;Q)

= \inf\_Q \overline{S}(f,W\_1\times W\_2;Q),

woraus folgt, dass

\sup\_P \underline{S}(\mathcal{U},W\_1;P) = \inf\_P \overline{S}(\mathcal{U},W\_1;P)

und

\sup\_P \underline{S}(\mathcal{O},W\_1;P) = \inf\_P \overline{S}(\mathcal{O},W\_1;P).

Damit sind \mathcal{U}, \mathcal{O}: W\_1\rightarrow \R Riemann-integrierbar und es gilt

\int\_{W\_1} \mathcal{U} = \int\limits\_{W\_1\times W\_2} f = \int\_{W\_1} \mathcal{O}.

</beweis>

<bemerkung> Bemerkung 4.6.2

Ist f stetig, so sind auch alle Funktionen g\_x(y) stetig und demzufolge Riemann-integrierbar, und die Formel des Satzes von Fubini vereinfacht sich zu

\int\limits\_{W\_1\times W\_2} f =\int\_{W\_1}\left[\int\_{W\_2} f(x,y) dy\right] dx.

Nach n-facher Anwendung auf einen Quader W=[a\_1,b\_1]\times \dots [a\_n,b\_n] folgt etwa für eine stetige Funktion f: W\rightarrow \R

\int\_W f(x) dx = \int\_{a\_1}^{b\_1} \left[ ...

\left[ \int\_{a\_n}^{b\_n} f(x\_1,..., x\_n) dx\_n\right]dx\_{n-1}

...\right] dx\_1.

</bemerkung>

<beispiel> Beispiel 4.6.3

Der Satz von Fubini muss mit oberen und unteren Integralen formuliert werden, weil g\_x(y)=f(x,y) bei festem x als Funktion auf W\_2 nicht integrierbar sein muss, auch wenn f auf W\_1 \times W\_2 es ist. Betrachte so die Funktion

f: [0,1]^2\rightarrow \R, f(x,y)=\left\{\ba{ll}

1 & \text{ falls x irrational,}

1 & \text{ falls x rational, y irrational,}

1- 1/q & \text{ falls x=p/q, y rational. } \ea\right.

Diese Funktion ist unstetig in [0,1]\times (\Q\cap [0,1]), was als abzählbare Vereinigung von Lebesgue-Nullmengen der Form [0,1]\times \{ r\} wieder eine Lebesgue-Nullmenge ist. Damit ist f Riemann-integrierbar. Aber es ist g\_x=1 für x irrational, also integrierbar, dagegen ist für x=p/q die Einschränkung

g\_x(y)=\left\{

1 \text{ y irrational}

1-1/q \text{ y rational}.

in keinem Punkte stetig, also auch nicht Riemann-integrierbar.

Trotzdem existiert die Funktion

\mathcal{U}(x) = \int^u\_{[0,1]} g\_x(y) dy =

1 & \text{ falls x irrational}

1-1/q & \text{ falls x rational}

Deren Unstetigkeitstellen sind genau die rationalen Zahlen ("Ubungsaufgabe), also abzählbar. Als Gesamtwert des Integrals erhält man

\int\_{[0,1]^2} f = \int\_{[0,1]} \mathcal{U}(x) dx = \sup\_P\sum\_{W\in P} \max \mathcal{U}\big|\_W \text{vol} W = \sup\_P\sum\_{W\in P} 1\cdot \text{vol} W = 1.

</beispiel>

Wir besprechen nun eine Reihe von wichtigen Anwendungsbeispielen des Satzes von Fubini.

<beispiel> Beispiel 4.6.4 Flächen unter Graphen

Sei f: [a,b] \rightarrow \R eine Riemann-integrierbare Funktion, f \geq 0.

Wir betrachten das Gebiet A = \{ (x,y) | a \leq x\leq b, 0\leq y\leq f(x)\}, von welchem wir annehmen, dass es in W\_1\times W\_2 für geeignete Quader W\_1, W\_2 \subset \R enthalten ist.

Nach dem Satz von Fubini gilt dann

\text{vol} A = \int\_{W\_1 \times W\_2} \chi\_A = \int\_{W\_1} \left[\int\_{W\_2} \chi\_A(x,y) dy\right]dx

= \int\_a^b\left[ \int\_0^{f(x)} 1\cdot dy \right] dx = \int\_a^b f(x) dx.

Dies begründet die vertraute Formel zur Berechnung der Fläche unter einem Graphen. Der Sachverhalt gilt aber noch viel allgemeiner. Sei so A\subset \R^n Jordan-messbar und

f: A\rightarrow [0,\infty[ Riemann-integrierbar.

Wir betrachten das Gebiet Z = \{ (x,y)\in \R^n\times \R | x\in A \text{ und } 0\leq y\leq f(x)\}

mit Querschnitt A, welches oben von dem Graphen von f abgeschlossen wird. Mit dem gleichen Beweis gilt dann \text{vol} Z = \int\_A f(x) dx.

</beispiel>

<beispiel> Beispiel 4.6.5 Das Volumen von Kugeln

Wir berechnen mit Hilfe des Satzes von Fubini eine Rekursionsformel für das Volumen der n-dimensionalen Kugel D^n(1)\subset\R^n vom Radius 1. Zunächst versehen wir \R^n=\R^{n-2}\times \R^2 mit den Koordinaten (x,y,z\_1,...,z\_{n-2}) und berechnen

\text{vol} D^n(1) = \int\limits\_{D^n(1)}1 = \int\limits\_{x^2+y^2\leq 1}

\left[ \int\limits\_{\sum z\_i^2\leq 1- x^2-y^2}1\cdot dz\_1... dz\_{n-2} \right]dxdy

= \int\limits\_{x^2+y^2\leq 1} \text{vol} (D^{n-2} (\sqrt{1-x^2-y^2})) dxdy

= \text{vol} D^{n-2}(1) \int\limits\_{x^2+y^2\leq 1} (1-x^2-y^2)^{\frac{n-2}{2}} dxdy

=: \text{vol} D^{n-2}(1)\cdot {\mathcal{J}}\_{\frac{n-2}{2}}.

Dabei haben wir im letzten Schritt die Skalierungseigenschaft für Jordan-messbare Mengen, siehe Lemma 4.3.11, verwendet. Um eine Rekursionsformel zu erhalten, benötigen wir den Wert des zweidimensionalen Integrals. Nach Fubini ist

{\mathcal{J}}\_{\frac{n-2}{2}} = \int\limits\_{x^2+y^2\leq 1} (1-x^2-y^2)^{\frac{n-2}{2}} dxdy

= \int\_{-1}^1 \left[\int\limits\_{-\sqrt{1-x^2 }}^{\sqrt{1-x^2 }} (1-x^2-y^2)^{\frac{n-2}{2}}dy \right]dx.

Das innere Integral ist von einer Gestalt, die sich mit einer einfachen Substitution vereinfachen läßt. Sei so c>0 eine beliebige Konstante und betrachte die Substitution: t = \sqrt{c}\cos\phi dt

= -\sqrt{c}\sin \phi d\phi ,

 t=-\sqrt{c}\leftrightarrow \phi =\pi,

 t=\sqrt{c}\leftrightarrow \phi =0. Dann ist

\int\_{-\sqrt{c}}^{\sqrt{c}} (c-t^2)^{\frac{n-2}{2}}dt

= \int^{\pi}\_0 (c-c\cos^2\phi )^{\frac{n-2}{2}}\sqrt{c}\sin\phi d\phi

 = c^{\frac{n-1}{2}}\int\_0^\pi \sin^{n-1}\phi d\phi

= 2c^{\frac{n-1}{2}}\int\_0^{\pi/2} \sin^{n-1}\phi d\phi .

Der Wert des letzten Integrals läßt sich mittels der Wallis-Formel berechnen. Einsetzen ergibt nun mit c=1-x^2

{\mathcal{J}}\_{\frac{n-2}{2}}

= 2 \int\_{-1}^1 \left[(1-x^2)^{\frac{n-1}{2}} \int\_0^{\pi/2} \sin^{n-1}\phi d\phi \right] dx

= 2\int\_0^{\pi/2} \sin^{n-1}\phi d\phi \int\_{-1}^1 (1-x^2)^{\frac{n-1}{2}} dx

= 2\int\_0^{\pi/2} \sin^{n-1}\phi d\phi \int\_0^\pi\sin^n\theta d\theta

= 4 \int\_0^{\pi/2} \sin^{n-1}\phi d\phi \int\_0^{\pi/2}\sin^n\theta d\theta.

Ist nun etwa n=2k+1 ungerade, so gilt

\int\_0^{\pi/2}\sin^{n-1}\phi d\phi = \int\_0^{\pi/2}\sin^{2k}\phi d\phi

= \frac{1\cdot 3\cdot...\cdot (2k-1)}{2\cdot 4\cdot...\cdot (2k)}\frac{\pi}{2},

\int\_0^{\pi/2}\sin^{n}\theta d\theta = \int\_0^{\pi/2}\sin^{2k+1}\theta d\theta

= \frac{2\cdot 4\cdot...\cdot (2k)}{1\cdot 3\cdot...\cdot (2k+1)}

und damit {\mathcal{J}}\_{\frac{n-2}{2}} = \frac{2\pi}{n}.

Man rechnet nach, dass dieser Ausdruck auch für n=2k gerade richtig bleibt. Insgesamt erhalten wir

\text{vol} D^n(1) = \frac{2\pi}{n} \text{vol} D^{n-2}(1).

Da nach der Skalierungseigenschaft des Volumens aber \text{vol} D^n(R)=R^n\text{vol} D^n(1) gilt, kann man dies alternativ schreiben als

\text{vol} D^n(R) = \frac{2\pi R^2}{n} \text{vol} D^{n-2}(R).

Hat man \text{vol}D^1(R) und \text{vol}D^2(R) bestimmt, so kann man daraus alle anderen Kugelvolumina rekursiv berechnen. Das erste ist die Intervalllänge, \text{vol}D^1(R)=2R, und für D^2(R) erhält man bekanntlich mit

\sin^2\phi=(1-\cos 2 \phi)/2 \text{vol} D^2(R)

= \int\limits\_{-R}^R\left[ \int\limits\_{-\sqrt{R^2-x^2}}^{\sqrt{R^2-x^2}}dy\right] dx

= 2 \int\_{-R}^R \sqrt{R^2-x^2} dx

\stackrel{x=R\sin\phi }{=} 2\int\_{-\pi/2}^{\pi/2} R^2\cos^2\phi d\phi

= 4R^2\int\_0^{\pi/2} \frac{1+\cos 2\phi }{2} d\phi

= 4R^2\left[\frac{\phi }{2} + \frac{\sin 2\phi }{4}\right]\_0^{\pi/2}

= \pi R^2.

Daraus erhält man nun die expliziten Formeln

\text{vol} D^n(R) = \left\{\ba{ll} \frac{\pi^{\frac{n}{2}} R^n}{(n/2)!} \text{ falls n gerade}

\frac{2^{\frac{n+1}{2}}\pi^{\frac{n-1}{2}}R^n}{1\cdot 3\cdot 5\cdot...\cdot

n} \text{ falls n ungerade.}

So ist beispielsweise

D^3(R) = \frac{4\pi}{3}R^3, D^4(R) = \frac{\pi^2}{2}R^4, D^5(R) = \frac{8\pi^2}{15}R^5.

</beispiel>

<beispiel> Beispiel 4.6.6 Volumen eines Rotationskörpers

Sei f: [a,b]\rightarrow \R eine nicht-negative, Riemann-integrierbare Funktion. Wir betrachten in der (x,y)-Ebene den Graphen von f und drehen diese Kurve in \R^3 um die x-Achse. Die Menge V, die von der so entstehenden Fäche eingeschlossen wird, heißt der von f als Leitkurve erzeugte Rotationskörper.

Ein Punkt (x,y,z) liegt genau dann in V, falls a\leq x \leq b und y^2+z^2\leq f^2(x), da der Schnitt von V mit der Ebene x=x\_0 aus dem Kreis mit Radius f(x\_0) besteht.

Der Satz von Fubini liefert dann

\text{vol} V = \int \chi\_V dx dy dz = \int\_a^b \left[\int\_{y^2+z^2\leq f^2(x)} dydz \right]dx

= \int\_a^b\text{vol} D^2(f(x)) dx = \int\_a^b \pi f^2(x)dx.

Intuitiv ist das Volumen des Rotationskörper somit genau die infinitesimale Summe der Flächen aller seiner Querschnitte.

</beispiel>

Motivierend für die Entwicklung des Riemann-Integrales war ursprünglich

<proposition> Proposition 4.6.7 Cavalieri'sches Prinzip

Seien A,B \subset\R^3 Jordan-messbar. Für c \in\R betrachten wir die Mengen

 A\_c = \{ (x,y) \in \R^2: (x,y,c) \in A \} und

B\_c = \{ (x,y) \in \R^2: (x,y,c) \in B \} .

Für jede reelle Zahl c seien A\_c und B\_c Jordan-messbar und es gelte \text{vol} A\_c = \text{vol} B\_c. Dann ist sogar \text{vol} A = \text{vol}B.

</proposition>

<beweis>

Wir betrachten Intervalle X,Y,Z in x-, y- und z-Richtung derart, dass A und B im Quader W=X\times Y\times Z liegen. Nach Voraussetzung sind \chi\_A, \chi\_B, \chi\_{A\_c},\chi\_{B\_c} Riemann-integrierbar. Wir betrachten nun f=\chi\_A als Abbildung W\_2\times W\_1\rightarrow \R mit W\_1=Z\subset\R, W\_2=X\times Y\subset \R^2 sowie entsprechend g\_c=\chi\_{A\_c}: W\_2\rightarrow \R. Weil diese Funktion Riemann-integrierbar ist, liefert der Satz von Fubini in der vereinfachten Form

\text{vol} A =\int\limits\_{W\_1\times W\_2} \chi\_A

= \int\limits\_{W\_1} \left[ \int\limits\_{W\_2} \chi\_{A\_z}dxdy\right]dz

= \int\limits\_{W\_1} \text{vol} A\_z dz.

Nach Voraussetzung ist aber \text{vol} A\_c=\text{vol} B\_c und damit

\int\limits\_{W\_1} \text{vol} A\_z dz = \int\limits\_{W\_1} \text{vol} B\_z dz.

Anwendung des Satzes von Fubini liefert, dass dieses Integral gleich \text{vol} B ist, was den Beweis beendet.

</beweis>

<beispiel> Beispiel 4.6.8 Masse, Schwerpunkt und Trägheitsmoment

Sei V\subset \R^3 eine Jordan-messbarer Körper in \R^3 bzw. F\subset\R^2 eine Jordan-messbare Fläche in \R^2. Letztere idealisiert einen starren Körper vernachlässigbarer Dicke.

Mit x bezeichnen wir den Punkt mit Koordinaten (x\_1,x\_2,x\_3) bzw. (x\_1,x\_2), dx ist je nach Kontext dx\_1dx\_2dx\_3 bzw. dx\_1dx\_2.

Sind \rho(x) bzw.~\mu(x) die Massendichten pro Volumen- bzw. Flächeneinheit von V bzw. F, so sind deren Gesamtmassen

m\_V = \int\_V \rho(x)dx,

m\_F = \int\_F \mu(x) dx.

Die Koordinaten x\_i^V, x\_i^F des Schwerpunkts G sind dann

x^V\_i = \frac{1}{m\_V}\int\_V x\_i(x) \rho(x)dx (i=1,2,3),

x^F\_i = \frac{1}{m\_F}\int\_F x\_i(x)\mu(x) dx (i=1,2).

Als Beispiel bestimmen wir die Koordinaten des Schwerpunkts einer zweidimensionalen Platte, die durch die Kurven y^2=ax und x=a bei gegebenem a>0 begrenzt wird und die konstante Massendichte \mu hat. Aus Symmetriegründen hat der Schwerpunkt die Ordinate Null, y^F=0, und wegen \mu=konst ist die Masse proportional zur Fläche. Diese ist das Doppelte der Fläche unter dem Graphen von y=\sqrt{ax}, also m\_S = 2\mu \int\_0^a \sqrt{ax} dx = 2\mu\sqrt{a} \int\_0^a x^{1/2}dx

= 2\mu\sqrt{a} \frac{2}{3}x^{3/2}\big|\_0^a = \frac{4}{3}\mu a^2.

Damit ergibt sich für die x-Koordinate des Schwerpunkts

m\_s\cdot x^S = \int\_0^a 2\mu \sqrt{ax}\cdot x dx

= 2\mu\sqrt{a}\frac{2}{5}\big|\_0^a = \frac{4}{5}\mu a^3,

also x^S = \frac{3}{5}a.

Wir besprechen nun das Trägheitsmoment. Sind n Massenpunkte X\_i der Masse m\_i gegeben, so ist das Trägheitsmoment dieses Systems bezüglich einer Achse A definiert als J\_A:= \sum\_{i=1}^n m\_i d^2(X\_i,A), wobei d der euklidische Abstand ist.

Verallgemeinert dies auf kontinuierliche Massenverteilungen, so bedeutet dies

J\_{A}^V = \int\_V \rho(x) d^2(A,x) dx,

J^{A}\_S = \int\_S \mu(x) d^2(A,x) dx.

 Als Beispiel bestimmen wir das Trägheitsmoment einer Stange der Länge l und konstanter Massendichte \mu bezüglich einer Achse, die senkrecht durch eines der Enden der Stange verläuft. Dann ist

J = \int\_0^l \mu x^2dx = \mu\left[\frac{x^3}{3}\right]^l\_0 = \mu l^3/3.

Eine wesentliche Vereinfachung zur Berechnung von Trägheitsmomenten liefert der Satz von Huygens.

<satz> Satz 4.6.9 Satz von Huygens

Sei V ein Jordan-messbarer Körper in \R^3 mit Schwerpunkt G und Masse m, A eine Achse durch G und A' eine beliebige Parallele zu A. Dann gilt

J\_{A'}^V = J^V\_A + md^2(A,A').

</satz>

<beweis>

Wir drehen unser Koordinatensystem so, dass A,A' parallel zur x\_3-Achse sind, und verschieben es derart, dass A' sogar mit der x\_3-Achse zusammenfällt. Damit ist für einen Punkt x=(x\_1,x\_2,x\_3)\in V

d^2(A,x) = (x\_1-x\_1^V)^2+ (x\_2-x\_2^V)^2,

d^2(A',x) = x\_1^2+x\_2^2,

d^2(A,A') = d^2 = (x\_1^V)^2+ (x\_2^V)^2.

Hieraus ergibt sich dann

J^V\_{A} = \int\_V\rho(x)d^2(A,x)dx

= \int\_V \rho(x)\left[(x\_1-x\_1^V)^2+ (x\_2-x\_2^V)^2\right] dx

= \int\_V \rho(x)\left[x\_1^2+x\_2^2\right] dx

+ \left[(x\_1^V)^2+(x\_2^V)^2\right] \int\_V \rho(x) dx

- 2 x\_1^V\int\_V \rho(x) x\_1 dx - 2 x\_2^V\int\_V \rho(x) x\_2 dx

= \int\_V \rho(x) d^2(A',x) dx

+ d^2(A,A')\cdot m - 2x\_1^V(m\cdot x\_1^V) - 2x\_2^V(m\cdot x\_2^V)

= J^V\_{A'}+ md^2(A,A') -2 md^2(A,A') = J^V\_{A'}- md^2(A,A').

</beweis>

</beispiel>

## 4.7 Uneigentliche Riemann-Integrale auf \R

Ist f: [a,\infty[\rightarrow \R eine auf jedem endlichen Intervall [a,x] beschränkte, Riemann-integrierbare Funktion, so können wir \int\_a^x f(t) dt bilden und uns fragen, ob ihr Limes x\rightarrow \infty existiert.

<definition> Definition 4.7.1 Uneigentliches Integral

Angenommen, f:[a,\infty[\rightarrow \R ist eine auf jedem endlichen Intervall [a,x] beschränkte, Riemann-integrierbare Funktion. Existiert der Grenzwert \lim\_{x\rightarrow \infty}\int\_a^x f(t) dt, so sagt man, dass f auf [a,\infty[ uneigentlich Riemann-integrierbar ist und setzt

\int\_a^\infty f(t) dt := \lim\_{x\rightarrow \infty}\int\_a^x f(t) dt.

Analog definiert man \int\_{-\infty}^b f(t) dt für eine Funktion

f: ]-\infty, b]\rightarrow \R, die auf endlichen Intervallen beschränkt und Riemann-integrierbar ist.

Für eine Funktion f: ]-\infty,\infty[\rightarrow \R mit diesen Eigenschaften wird weiterhin vereinbart, dass

\int\_{-\infty}^\infty f(t) dt := \int\_{-\infty}^0 f(t) dt + \int\_{0}^\infty f(t) dt.

</definition>

Obige Definition bedeutet im letzteren Fall, dass die Grenzwerte gegen +\infty bzw. -\infty unabhängig voneinander existieren müssen, es ist also nicht zulässig, \int\_{-\infty}^\infty f(t) dt mit

\lim\_{c\rightarrow \infty} \int\_{-c}^c f(t) dt

gleichzusetzen. Ein solcher unbestimmer Integralbegriff wäre denkbar, ist aber nicht Gegenstand dieser Vorlesung. Für Integrale dieser drei Typen gelten die üblichen Rechenregeln, sofern alle auftretenden Ausdrücke sinnvoll sind. So ist etwa

\int\_a^\infty (f(t)+g(t)) dt = \int\_a^\infty f(t) dt+ \int\_a^\infty g(t) dt.

<beispiel> Beispiel 4.7.2

Das unbestimmte Integral \int\_0^\infty \sin t dt existiert nicht, weil \int\_0^x\sin t dt = -\cos x+1 ist und \lim\_{x\rightarrow \infty} \cos x nicht existiert.

</beispiel>

<beispiel> Beispiel 4.7.3

Das unbestimmte Integral \int\_0^\infty \frac{dt}{(1+t^2)^n} existiert (n>1) und hat den Wert

\frac{1\cdot 3\cdot...\cdot (2n-3)}{2\cdot 4\cdot ...\cdot (2n-2)}\frac{\pi}{2}.

Substituiert man nämlich t=\cot \phi , dt=-(1+\cot^2\phi d\phi ), t=0

\leftrightarrow \phi =\pi/2, t=x\leftrightarrow \phi =\mathrm{arccot} x,

so gilt

\int\_0^x \frac{dt}{(1+t^2)^n} = \int\_{\mathrm{arccot} x}^{\pi/2} \frac{d\phi }{(1+\cot^2\phi )^{n-1}} = \int\_{\mathrm{arccot} x}^{\pi/2}\sin^{2n-2}\phi d\phi .

Gilt x\rightarrow \infty, so strebt \mathrm{arccot} gegen 0. Folglich ist

 \lim\_{x\rightarrow \infty}\int\_0^x \frac{dt}{(1+t^2)^n}

= \int\_0^{\pi/2} \sin^{2n-2}\phi d\phi = \frac{1\cdot 3\cdot...\cdot (2n-3)}

{2\cdot 4\cdot ...\cdot (2n-2)}\frac{\pi}{2}.

</beispiel>

<beispiel> Beispiel 4.7.4 Poisson-Integral

Wir zeigen als nächstes \int\_0^\infty e^{-x^2} dx =\frac{\sqrt{\pi}}{2} und betrachten hierzu die Kreisscheibe D^2(R)\subset\R^2 und das Integral

F(R) = \int\_{D^2(R)}e^{-x^2-y^2} dxdy.

Wir wollen zeigen, dass F als Funktion von R differenzierbar ist. So ist für ein beliebiges \epsilon>0

F(R+\epsilon) - F(R) = \int\limits\_{D^2(R +\epsilon) - D^2(R)} e^{-x^2-y^2} dxdy.

Wenden wir den ersten Mittelwertsatz der Integralrechnung an, so wissen wir, dass ein Punkt (x\_0(\epsilon), y\_0(\epsilon))\in D^2(R +\epsilon) - D^2(R) existiert mit F(R+\epsilon) - F(R)

= e^{-x\_0^2(\epsilon) - y\_0^2(\epsilon)} \text{vol} (D^2(R +\epsilon) - D^2(R))

 = e^{-x\_0^2(\epsilon) - y\_0^2(\epsilon)} \pi [(R+\epsilon)^2-R^2].

Damit folgt

\frac{F(R+\epsilon) - F(R)}{\epsilon} = e^{-x\_0^2(\epsilon) - y\_0^2(\epsilon)} \pi(2R+\epsilon).

Für \epsilon>0 muss die Folge der Punkte (x\_0(\epsilon), y\_0(\epsilon)) nicht konvergent sein; es ist aber auf jeden Fall x\_0^2(\epsilon) + y\_0^2(\epsilon)\rightarrow R^2. Damit gilt

\lim\_{\epsilon\rightarrow 0} \frac{F(R+\epsilon) - F(R)}{\epsilon} = e^{-R^2} 2\pi R.

Damit ist F wie behauptet differenzierbar und es gilt F'(R)=e^{-R^2} 2\pi R. Nach dem Hauptsatz der Differential- und Integralrechnung ist dann F(R) = F(R) - F(0) = \int\_0^R F'(t) dt

= \int\_0^R \pi e^{-t^2}2t dt = \pi (-e^{-t^2})\big|\_0^R = \pi(1-e^{-R^2}).

Nun ist aber D^2(R) \subset [-R,R]^2 \subset D^2(\sqrt{2}R), also

F(R) \leq \int\_{[-R,R]^2}e^{-x^2-y^2} dxdy \leq F(\sqrt{2}R)

Der Satz von Fubini liefert für das mittlere Integral

\int\_{[-R,R]^2}e^{-x^2-y^2} dxdy = \left[\int\_{-R}^R e^{-t^2}dt\right]^2

= 4 \left[\int\_{0}^R e^{-t^2}dt\right]^2,

so dass insgesamt

\pi(1-e^{-R^2}) \leq 4 \left[\int\_{0}^R e^{-t^2}dt\right]^2

\leq \pi(1-e^{-2R^2}).

Läßt man nun R gegen unendlich gehen, so folgt die Behauptung.

</beispiel>

<satz> Satz 4.7.5 Satz von Cauchy-McLaurin

Ist f: [a,\infty[ \to \R^+ stetig, monoton fallend und positiv, dann existiert \int\_a^\infty f(t) dt genau dann, falls die Reihe \sum\_{n=1}^\infty f(a+n) konvergiert.

</satz>

<beweis>

Der Satz beruht auf einer simplen Abschätzung. Für x\in [a+k, a+k+1] gilt nach Voraussetzung f(a+k+1) \leq f(x) \leq f(a+k), also nach Integration über [a+k,a+k+1]

f(a+k+1) \leq \int\_{a+k}^{a+k+1}f(x) dx \leq f(a+k).

Dies jedoch impliziert

\sum\_{n=1}^{k+1} f(a+n) \leq \int\_{a}^{a+k+1}f(x) dx \leq \sum\_{n=0}^{k} f(a+n).

Weil alle Zahlen positiv sind, folgt hieraus die Behauptung.

</beweis>

<bemerkung> Bemerkung 4.7.6

Diesen Satz kann zum Beweis der Existenz des uneigentlichen Integrals bei feststehender Konvergenz der Reihe angewandt werden, sowie umgekehrt.

</bemerkung>

<beispiel> Beispiel 4.7.7

Wir wir bereits sahen, ist \zeta(s)=\sum\_{n=1}^\infty 1/ n^s für s>1 konvergent. Tatsächlich ist für s>1

\int\_1^x \frac{1}{t^s} dt = \frac{-1}{s-1} \left[ \frac{1}{x^{s-1}} -1\right],

also

\int\_1^\infty \frac{1}{t^s} dt = \frac{1}{s-1}.

Nun folgt die Konvergenz der Reihe aus dem vorherigen Satz.

</beispiel>

Ein weiterer Typ uneigentlicher Integrale ist derjenige, wo bis zu einer Stelle integriert werden soll, in der der Integrand gar nicht definiert ist.

<definition> Definition 4.7.8

Sei f: ]a,b[\rightarrow \R auf dem offenen Intervall ]a,b[ definiert und auf jedem abgeschlossenen Intervall [\alpha,\beta] \subset ]a,b[ Riemann-integrierbar, also insbesondere \int\_\alpha^\beta f(t) dt existent.

Wir sagen, dass f auf ]a,b[ uneigentlich integrierbar ist, falls für ein x\_0 \in ]a,b[ die Grenzwerte

\lim\_{\alpha\rightarrow a^+}\int\_\alpha^{x\_0} f(t) dt,

\lim\_{\beta\rightarrow b^-}\int\_{x\_0}^\beta f(t) dt

existieren und schreiben dann

\int\_a^b f(t) dt = \lim\_{\alpha\rightarrow a^+}\int\_\alpha^{x\_0} f(t)+

\lim\_{\beta\rightarrow b^-}\int\_{x\_0}^\beta f(t) dt.

</definition>

In obiger Definition ist x\_0 ein beliebiger Punkt in ]a,b[ und man prüft leicht nach, dass die Definition nicht von x\_0 abhängt. Ist a=-\infty oder b=+\infty, so stimmt dieser Begriff mit dem zuvor definierten überein.

Abschließend besprechen wir im folgenden noch die für die Analysis und Zahlentheorie wichtige Gamma-Funktion.

<beispiel> Beispiel 4.7.9 Gamma-Funktion

Wir betrachten für x>0 die Funktion \Gamma(x) = \int\_0^\infty t^{x-1} e^{-t} dt,

 die Integral-Darstellung der Gamma-Funktion nach Euler (1730). Um die Existenz von \Gamma(x) einzusehen, bemerken wir zunächst, dass

0 < t^{x-1} e^{-t} < t^{x-1} für t>0.

Damit ist

0 < \int\_\alpha^1 t^{x-1}e^{-t} dt \leq \int\_\alpha^1 t^{x-1} dt

= \frac{1}{x} t^x\bigg|\_\alpha^1 = \frac{1}{x} -\frac{\alpha^x}{x}.

Wegen x>0 gilt also für \alpha >0

\int\_\alpha^1 t^{x-1}e^{-t} dr < \frac{1}{x},

und weil bei \alpha\rightarrow 0 das Integral

\int\_\alpha^1 t^{x-1}e^{-t} dt wächst und durch 1/x beschränkt ist, muss demzufolge \lim\_{\alpha\rightarrow 0^+} \int\_\alpha^1 t^{x-1}e^{-t} dt existieren.

Damit existiert \int\_0^1 t^{x-1} e^{-t} dt. Analog ist t^{x-1}e^{-t}=\frac{t^{x+1}}{e^t}\frac{1}{t^2} und \lim\_{t\rightarrow \infty} \frac{t^{x+1}}{e^t}=0. Damit existiert ein M>0 mit

t^{x-1} e^{-t} \leq \frac{1}{t^2} \text{ für alle } t\geq M

und es folgt

\int\_M^\beta t^{x-1}e^{-t} dt \leq \int\_M^\beta \frac{1}{t^2} dt

= - \frac{1}{t}\bigg|\_M^\beta \leq \frac{1}{M}.

Weil \int\_M^\beta t^{x-1} e^{-t} dt wachsend in \beta ist, folgt, dass \lim\_{\beta\rightarrow \infty} \int\_M^\infty t^{x-1} e^{-t} dt existiert. Damit existiert auch

\int\_1^\infty t^{x-1} e^{-t} dt = \int\_1^M t^{x-1} e^{-t} dt + \int\_M^\infty t^{x-1}e^{-t} dt.

Damit ist gezeigt, dass t^{x-1} e^{-t} auf (0,\infty) uneigentlich integrierbar ist und \Gamma(x) für x>0 existiert. Weiterhin haben wir vermöge der Substitution \sqrt t =u

\Gamma\left(\frac{1}{2}\right) = \int\_0^\infty \frac{e^{-t}}{\sqrt{t}} dt

= \int\_0^\infty \frac{e^{-u^2}}{u}2u du = 2 \int\_0^\infty e^{-u^2} du.

Also folgt aus dem Wert des Poisson-Integrals \Gamma(1/2)=\sqrt{\pi}. Weiterhin haben wir für x>1

\Gamma (x) = \int\_0^\infty t^{x-1} e^{-t} dt = -\int\_0^\infty t^{x-1}\left( e^{-t}\right)' dt

= - t^{x-1} e^{-t}\bigg|\_0^\infty +\int\_0^\infty (x-1)t^{x-2}e^{-t} dt

= 0+ (x-1)\int\_0^\infty t^{x-2}e^{-t} dt,

also \Gamma(x)=(x-1)\Gamma(x-1) für x>1.

Letztlich gilt \Gamma(1) = \int\_0^\infty e^{-t} dt = -e^{-t}\bigg|\_0^\infty = 1.

Damit folgt \Gamma(n) = (n-1)\Gamma(n-1) = (n-1)(n-2)\Gamma(n-2) = ... = (n-1)!.

So war Eulers Motivation für die Einführung der Gamma-Funktion tatsächlich auch gewesen, eine Funktion zu finden, die eine Fortsetzung der Fakultät darstellt. Desweiteren läßt sich das Volumen der Kugel D^n(R) durch die \Gamma-Funktion gemäß

<eq:4>

\text{vol} D^n(R) = \frac{R^n \pi^{n/2}}{\Gamma(n/2+1)}

</eq:4>

ausdrücken. Wir beweisen dies induktiv. Tatsächlich gilt für n=1,2

\text{vol} D^1(R) = 2R,

\frac{R^1 \pi^{1/2}}{\Gamma(1+1/2)} = \frac{R\pi^{1/2}}{\frac{1}{2}\Gamma (1/2)} = 2R

\text{vol} D^2(R) = \pi R^2,

\frac{R^2\pi}{\Gamma(1+1)} = \frac{R^2\pi}{1!} = R^2\pi.

Gelte jetzt \eqref{eq:4} für n-2. Wir beweisen dann die Gleichung für n unter Verwendung der Rekursionsformel für das Volumen von D^n. So berechnet man

\text{vol} D^n(R) = \frac{2\pi}{n}R^2\text{vol} D^{n-2}(R)

= \frac{2\pi}{n} R^2\cdot \frac{R^{n-2}\pi^{\frac{n-2}{2}}}{\Gamma \left( \frac{n-2}{2}+1\right)}

= \frac{R^n \pi^{n/2}}{\frac{n}{2}\Gamma\left(\frac{n}{2}\right)}

= \frac{R^n\pi^{n/2}}{\Gamma\left( \frac{n}{2} +1 \right)},

und \eqref{eq:4} folgt. Damit erhalten wir insbesondere die Stirlingsche Formel

\Gamma(x) \simeq \sqrt{2}\pi x^{x-1/2}e^{-x}.

Die Gamma-Funktion kann, ähnlich wie die Riemannsche Zeta-Funktion, auf komplexe Werte x\in \C fortgesetzt werden. Sie spielt eine wichtige Rolle in der Theorie der speziellen Funktionen. Eine weitere, von Gauß 1813 gefundene Darstellung für die Gamma-Funktion lautet ,C.F.}

\Gamma(x) = \lim\_{n\rightarrow \infty} \frac{n! x^x}{x(x+1)... (x+n)} .

</beispiel>

Kapitel 5: Differentialrechnung für Abbildungen mehrerer Variablen

## 5.1 Die Ableitung einer Abbildung

Wie wir bereits bei der Differentialrechnung einer Variablen sahen, stellt die Ableitung einer Funktion eine lineare Näherung derselben dar. Ein entsprechender Ableitungsbegriff soll nun im folgenden auch für Funktionen mehrerer Veränderlichen eingeführt werden. Dies führt dazu, die Ableitung als lineare Abbildung im Sinne der linearen Algebra zu verstehen.

<definition> Definition 5.1.1

Sei f: U\rightarrow \R^m eine Abbildung, U\subset \R^n eine offene Teilmenge. Dann heißt f in x\_0 \in U differenzierbar, falls eine lineare Abbildung L: \R^n\rightarrow \R^m derart existiert, dass

<eq:6>

\lim\_{h\rightarrow 0, h\in\R^n} \frac{\| f(x\_0+h) - f(x\_0) - L(h)\|}{\| h\|} = 0.

</eq:6>

Die lineare Abbildung L heißt die Ableitung von f in x\_0 und man schreibt für dieselbe L := Df(x\_0).

</definition>

Der folgende Satz besagt, dass die Ableitung eindeutig ist und Differenzierbarkeit Stetigkeit impliziert.

<satz> Satz 5.1.2

(1) Ist f: U\rightarrow \R^m in x\_0\in U differenzierbar, so ist die lineare Abbildung L:\R^n\rightarrow \R^m durch die Bedingung \eqref{eq:6} eindeutig bestimmt.

(2) Ist f in x\_0\in U differenzierbar, dann ist f in x\_0 stetig.

</satz>

<beweis>

Seien L,L': \R^n\rightarrow \R^m zwei lineare Abbildungen, die die Bedingung \eqref{eq:6} erfüllen. Dann gilt zunächst

0 \leq \lim\_{h\rightarrow 0} \frac{\| L(h)-L'(h)\| }{\| h\|}

\leq \lim\_{h\rightarrow 0} \frac{\| f(x\_0+h) - f(x\_0) - L'(h)\|}{\| h\|}

+ \lim\_{h\rightarrow 0} \frac{\| L(h) - f(x\_0+h) + f(x\_0)\|}{\| h\|} = 0.

Daraus folgt \lim\_{h\rightarrow 0} \frac{\| L(h)-L'(h)\| }{\| h\|} =0.

Wir betrachten weiter h= t x mit t\in\R und 0\not=x\in\R^n. Es folgt

0 = \lim\_{t\rightarrow 0} \frac{\| L(t x)-L'(t x)\| }{\|t x\|} = \frac{\| L(x)-L'(x)\| }{\|x \|}.

Also gilt L(x)=L'(x) für alle x\in\R^n. Ist weiterhin f: U\rightarrow \R^m in x\_0 differenzierbar und x\_k\rightarrow x\_0, so haben wir

\|f(x\_k) - f(x\_0)\|

\leq \frac{\|f(x\_k) - f(x\_0) - L(x\_k- x\_0)) \|}{\|x\_k-x\_0\|}\cdot \|x\_k-x\_0\| +\|L(x\_k - x\_0)\|,

woraus sofort \lim\_{k\rightarrow \infty} f(x\_k)=f(x\_0) folgt.

</beweis>

Einen weiteren, schwächeren Ableitungsbegriff stellt der Begriff der Richtungsableitung dar. Der Zusammenhang zwischen beiden Begriffen wird im folgenden genauer untersucht werden.

<definition> Definition 5.1.3

Sei f: U\rightarrow \R^m eine Abbbildung. Wir sagen, dass f im Punkt x\_0 in Richtung 0\not=\vec{a}\in\R^n abgeleitet werden kann, falls

\lim\_{t\rightarrow 0} \frac{f(x\_0+ t\vec{a}) - f(x\_0)}{t} =: \nabla\_{\vec{a}}f(x\_0)

existiert. Die rechte Seite heißt Richtungsableitung oder Ableitung von f in Richtung \vec{a} an der Stelle x\_0.

</definition>

Zunächst zeigen wir, dass aus der Differenzierbarkeit die Existenz aller Richtungsableitungen folgt:

<satz> Satz 5.1.4

Ist f: U\rightarrow \R^m in x\_0 \in \R^n differenzierbar, so besitzt f in x\_0 in jeder Richtung \vec{a} die Richtungsableitung und es gilt:

\nabla\_{\vec{a}} f(x\_0) = D f(x\_0) (\vec{a}) = L(\vec{a}).

</satz>

<beweis>

Sei f in x\_0 differenzierbar mit Ableitung L=Df(x\_0): \R^n\rightarrow \R^m. Wir betrachten alsdann den Vektor h=t\vec{a}\in\R^n. Für ihn gilt nach Voraussetzung

\lim\_{t\rightarrow 0} \frac{\|f(x\_0+ t\vec{a}) - f(x\_0) - L(t\vec{a})\|}{\|t\vec{a}\|} = 0,

also

\lim\_{t\rightarrow 0} \left\| \frac{f(x\_0+t\vec{a}) - f(x\_0)}{t } - L(\vec{a})\right\| = 0.

Dies beweist sowohl die Existenz der Richtungsableitung als auch die behauptete Formel.

</beweis>

Wir besprechen nun ein Beispiel einer Abbildung, die in einem Punkt alle Richtungsableitungen besitzt, aber dort nicht stetig, also erst recht nicht differenzierbar ist.

<beispiel> Beispiel 5.1.5

Sei f: \R^2\rightarrow \R definiert durch

f(x,y)=\left\{

\frac{xy^2}{x^2+y^4} \text{ falls } (x,y)\neq (0,0)

0 \text{ falls } (x,y)=(0,0).

Ist \vec{a}=(a\_1,a\_2) gegeben, so gilt

\nabla\_{\vec{a}} f(0) = \lim\_{t\rightarrow 0} \frac{f(ta\_1, ta\_2)}{t}

= \lim\_{t\rightarrow 0} \frac{t^3 a\_1a\_2^2}{t(t^2a\_1^2+t^4a\_2^4)}

= \left\{

0 falls a\_1=0

\frac{a\_2^2}{a\_1} falls a\_1\neq 0.

Dagegen ist f in (0,0) nicht stetig, also erst recht nicht differenzierbar, weil für die beiden Nullfolgen (0,1/n) und (1/n^2,1/n) gilt

\lim\_{n\rightarrow \infty} f(0,1/n) = 0,

\lim\_{n\rightarrow \infty} f(1/n^2,1/n)

 = \lim\_{n\rightarrow \infty} \frac{1/n^4}{1/n^4+ 1/n^4} = \frac{1}{2}.

Insbesondere existiert der Grenzwert \lim\limits\_{(x,y) \rightarrow (0,0)} f(x,y) nicht. Die folgenden Bilder stellen den Graphen der Funktion f dar.

</beispiel>

<beispiel> Beispiel 5.1.6 Lineare Abbildungen

Wir überlegen uns, dass eine lineare Abbildung L: \R^n\rightarrow \R^m in jedem Punkt differenzierbar ist und die Ableitung DL(x\_0)=L besitzt. In der Tat,

\frac{\| L(x\_0+h) - L(x\_0) - L(h)\|}{\| h\|}

= \frac{\| L(x\_0) +L( h) - L(x\_0) - L(h)\|}{\| h\|} = 0.

</beispiel>

<beispiel> Beispiel 5.1.7

Besitzt f in x\_0 die Richtungsableitungen \nabla\_{\vec{a}} f(x\_0),

\nabla\_{\vec{b}} f(x\_0), so muss \nabla\_{\vec{a}+\vec{b}} f(x\_0) nicht unbedingt existieren. Betrachte etwa

g: \R^2\rightarrow \R, g(x,y)=

\left\{ \frac{xy}{x^2+y^2} falls (x,y)\neq (0,0)

 0 falls (x,y)=(0,0)

sowie die beiden Vektoren \vec{a}=(1,0), \vec{b}=(0,1). Dann gilt

\nabla\_{\vec{a}} g(0,0) = \lim\_{t\rightarrow 0} \frac{g(t,0)}{t} = 0,

\nabla\_{\vec{b}} g(0,0) = \lim\_{t\rightarrow 0} \frac{g(0,t)}{t} = 0,

während jedoch

\nabla\_{\vec{a}+\vec{b}} g(0,0) = \lim\_{t\rightarrow 0} \frac{g(t,t)}{t}

= \lim\_{t\rightarrow 0}\frac{1}{2t}

nicht existiert.

</beispiel>

Eine wichtige Klasse von Abbildungen sind sogenannte multilineare Abbildungen. Sie interessieren uns hier unter dem Gesichtspunkt der Differenzierbarkeit.

<definition> Definition 5.1.8

Es seien V\_1,..., V\_k und W endlich-dimensionale Vektorräume. Eine Abbildung f: V\_1\times V\_2 ...\times V\_k\rightarrow W heißt multilinear, falls sie linear in jeder Komponente ist, d.h.falls für jedes 1\leq i\leq k gilt:

(1) f(x\_1,..., x\_i+y\_i, ..., x\_k) = f(x\_1,..., x\_i,..., x\_k)+ f(x\_1,...,y\_i,..., x\_k)

(2) f(x\_1,..., \lambda\cdot x\_i, ..., x\_k) = \lambda f(x\_1,..., x\_i, ..., x\_k)

</definition>

<lemma> Lemma 5.1.9 Fundamentalabschätzung für multilineare Abbildungen]

Für jede multilineare Abbildung f: V\_1\times ...\times V\_k\rightarrow W existiert eine Konstante A\_f mit

\|f(x\_1,..., x\_k)\| \leq A\_f \|x\_1\|\cdots \|x\_k\| \forall x\_i\in V\_i.

</lemma>

<beweis>

Jede multilineare Abbildung f ist überall stetig, insbesondere im Nullpunkt. Damit existiert ein \delta>0 derart, dass aus \|x\_1\|\leq \delta,..., \|x\_k\| \leq \delta die Beziehung \|f(x\_1,..., x\_k)\|\leq1 folgt. Dann erhält man jedoch für beliebige x\_1, ..., x\_k\neq 0

\|f(x\_1,..., x\_k)\|

 = \frac{\|x\_1\|}{\delta} ... \frac{\|x\_k\|}{\delta} \cdot

 \| f\Big( \frac{\delta x\_1}{\|x\_1\|},..., \frac{\delta x\_k}{\|x\_k\|}\Big)\|

\leq \frac{1}{\delta^k} \|x\_1\|\cdots \|x\_k\| .

</beweis>

<lemma> Lemma 5.1.10

Ist f: V\_1\times ...\times V\_k\rightarrow W eine multilineare Abbildung, so ist

Df(x\_1,..., x\_k)\in {\mathrm{End}} (V\_1\times ...\times V\_k, W) gegeben durch

Df(x\_1,..., x\_k) (a\_1,..., a\_k) = \sum\_{i=1}^k f(x\_1,..., x\_{i-1}, a\_i, x\_{i+1},..., x\_k).

</lemma>

<beweis>

Wir betrachten die Größe

F := f(x\_1+a\_1,..., x\_k+a\_k) - f(x\_1,..., x\_k) - \sum\_{i=1}^k f(x\_1, ..., a\_i,..., x\_k)

= \sum\_{i\_1< i\_2} f(x\_1,..., a\_{i\_1},..., a\_{i\_2}, \dots, x\_k)

+\sum\_{i\_1< i\_2<i\_3} f(x\_1,..., a\_{i\_1},..., a\_{i\_2},..., a\_{i\_3},... x\_k)+...

+ f(a\_1,..., a\_k).

Die Fundamentalabschätzung liefert mit der Konstanten A\_f

\|F\|\leq A\_f\left[\sum\_{i\_1< i\_2} \|x\_1\|\cdots \|a\_{i\_1}\| \cdots \|a\_{i\_2}\|\cdots \|x\_k\|

+ \sum\_{i\_1< i\_2<i\_3}... + ... + \|a\_1\|\cdots\|a\_k\| \right].

Die Ungleichung zwischen geometrischem und quadratischem Mittel für n nicht negative Zahlen b\_1,..., b\_n besagt, dass \sqrt[n]{b\_1\cdots b\_n} \leq \sqrt{\frac{b\_1^2+... + b\_n^2}{n}},

also b\_1\cdots b\_n\leq c\_n (b\_1^2+... + b\_n^2)^{n/2}. Angewandt auf die Zahlen b\_i= \|a\_i\| erhalten wir mit a\_1 statt a\_{i\_1} usw.

\|a\_1\|\cdot \|a\_2\|\leq c\_2 \left[\|a\_1\|^2+\|a\_2\|^2\right]^1

\leq c\_2 \left[\sum\_{i=1}^k \|a\_i\|^2\right]^1,

\|a\_1\|\cdot \|a\_2\|\cdot \|a\_3\|\leq c\_3 \left[\|a\_1\|^2+\|a\_2\|^2 +\|a\_3\|^2\right]^{3/2}

\leq c\_3 \left[\sum\_{i=1}^k \|a\_i\|^2\right]^{3/2} \text{ usw.}

Da die Zahlen \|x\_i\| konstant sind, spielen sie bei der Abschätzung keine Rolle. Zudem kann k\geq 2 vorausgesetzt werden, da für lineare Abbildungen das Ergebnis bereits bekannt ist. Eine Abschätzung ergibt nun

\frac{\|F\|}{\sqrt{\sum\_{i=1}^k \|a\_i\|^2}}

\leq A\_f\left[C\_2 \left[\sum\_{i=1}^k \|a\_i\|^2\right]^{1/2}

+ C\_3 \left[\sum\_{i=1}^k \|a\_i\|^2\right]^{1}+...

+ C\_k \left[\sum\_{i=1}^k \|a\_i\|^2\right]^{(k-1)/2} \right].

Dabei sind die Koeffizienten C\_i angepaßt worden, so dass etwa

C\_2 = c\_2 \sum\_{i\_1< i\_2} \|x\_1\|\cdots \|\hat{x}\_{i\_1}\| \cdots \|\hat{x}\_{i\_2}\|\cdots \|x\_k\|

Wichtig ist nur, dass diese Konstanten nicht mehr von den a\_i abhängen. Lässt man nun (a\_1, ..., a\_k) gegen Null gehen, so folgt

\lim\_{(a\_1,..., a\_k)\rightarrow (0,...,0)} \frac{\|F\|}{\sqrt{\sum\_{i=1}^k \|a\_i\|^2}} = 0.

</beweis>

<beispiel> Beispiel 5.1.11

Jede Bilinearform f: V\times V\rightarrow \K, f(x,y)=\langle x,y\rangle ist bilinear, also gilt

Df(x,y)(a\_1,a\_2) = \langle a\_1, y\rangle + \langle x, a\_2\rangle.

</beispiel>

<beispiel> Beispiel 5.1.12

Das Vektorprodukt f:\R^3\times \R^3\rightarrow \R^3 ist bilinear. Demzufolge gilt

Df(x,y)(a\_1,a\_2) = a\_1\times y+ x\times a\_2.

</beispiel>

<beispiel> Beispiel 5.1.13

Die Determinante \det: V\times \dots \times V\rightarrow \K, wobei die Anzahl der Faktoren gleich \dim V=n ist, ist multilinear, weswegen

D \det (x\_1,..., x\_n)(a\_1,..., a\_n)

= \det (a\_1, x\_2,..., x\_n)+ \det (x\_1, a\_2,..., x\_n)+... + \det(x\_1,..., x\_{n-1}, a\_n).

</beispiel>

Angenommen, f: U\rightarrow \R^n ist an der Stelle x\_0 differenzierbar. Die Ableitung Df(a): \R^n\rightarrow \R^m ist eine lineare Abbildung, die sich bezüglich zweier Basen durch eine (n\times m)-Matrix M[Df(a)] gemäß Df(a)\left ( b \right )

= M[Df(a)]\cdot \left (\begin{matrix} b\_1\ \vdots\ b\_n\end{matrix}\right ), b \in \R^n,

darstellen läßt. Die Einträge der darstellenden Matrix sind dabei durch die Bilder der Basisvektoren gegeben. Später werden wir zwischen der linearen Abbildung und der sie darstellenden Matrix nicht mehr unterscheiden, jedoch ist die Ableitung als abstrakte Abbildung wohldefiniert, während die Matrix von der Wahl von Basen abhängt. Im folgenden Satz stellen wir Rechenregeln für Ableitungen und Richtungsableitungen zusammen.

<satz> Satz 5.1.14

Gegeben seien Funktionen f,g: U\rightarrow \R^m, wobei U\subset \R^n eine offene Teilmenge bezeichnet.

(1) Existieren \nabla\_{\vec{a}}f(x\_0) und \nabla\_{\vec{a}}g(x\_0), so existiert auch \nabla\_{\vec{a}}(f+g)(x\_0) und es gilt

\nabla\_{\vec{a}}f(x\_0) + \nabla\_{\vec{a}}g(x\_0) = \nabla\_{\vec{a}}(f+g)(x\_0).

(2) Sind f und g in x\_0 differenzierbar, so ist auch f+g in x\_0 differenzierbar und

D(f+g)(x\_0) = Df(x\_0)+ Dg(x\_0).

(3) Produktregel. Sei h: U\rightarrow \R eine Funktion. Existieren die Richtungsableitungen bzw. die Ableitung von h und f, so gilt dies auch für hf und \nabla\_{\vec{a}} (hf)(x\_0)

= \nabla\_{\vec{a}} h(x\_0)\cdot f(x\_0) + h(x\_0)\nabla\_{\vec{a}} f(x\_0), D(hf)(x\_0)

= Dh(x\_0) f(x\_0) + h(x\_0)Df(x\_0).

(4)Quotientenregel. Sei m=1. Sind f,g in x\_0 differenzierbar und g(x\_0)\neq 0, so ist f/g in einer Umgebung von x\_0 definiert, in x\_0 differenzierbar und

D\left[\frac{f}{g}\right] (x\_0) = \frac{g(x\_0)Df(x\_0) - f(x\_0)Dg(x\_0)}{g(x\_0)^2}.

(5) Kettenregel. Ist f: U\rightarrow V\subset \R^m in x\_0 differenzierbar, g: V \rightarrow \R^p in f(x\_0) differenzierbar, so ist auch g\circ f: U\rightarrow \R^p in x\_0 differenzierbar und

D(g\circ f) (x\_0) = Dg (f(x\_0))\circ Df(x\_0).

(6) Die Abbildung f=(f\_1,f\_2): U\rightarrow \R^{m\_1}\times \R^{m\_2} ist genau dann in x\_0 differenzierbar, falls f\_1 und f\_2 dort differenzierbar sind und in diesem Falle gilt

Df(x\_0) = (Df\_1(x\_0), Df\_2(x\_0)).

Für die darstellenden Matrizen M[Df\_i(x\_0)]: \R^n\rightarrow \R^{m\_i} und M[Df(x\_0)]: \R^n\rightarrow \R^{m\_1+m\_2} bedeutet dies

M[Df(x\_0)] = \begin{bmatrix} M[Df\_1(x\_0)] \ M[Df\_2(x\_0)]\end{bmatrix} .

</satz>

<beweis>

Die Eigenschaften (1)-(3) und (6) sind trivial. Wir zeigen die Kettenregel (5): Sei y\_0=f(x\_0), L=Df(x\_0) und L'=Dg(y\_0). Wir betrachten zunächst

\phi (x) = f(x)-f(x\_0) - L(x-x\_0),

\psi(y) = g(y) - g(y\_0) - L'(y-y\_0),

\rho(x) = (g\circ f)(x) - (g\circ f)(x\_0) - (L'\circ L)(x-x\_0).

Die Differenzierbarkeit von f und g in x\_0 bzw. y\_0 bedeutet

\lim\_{x\rightarrow x\_0} \frac{\|\phi (x)\|}{\|x-x\_0\|} = 0,

\lim\_{y\rightarrow y\_0} \frac{\|\psi(y)\|}{\|y-y\_0\|} = 0 .

Drückt man \rho(x) durch die anderen Größen aus, so erhält man

\rho (x) &= & g (f(x)) - g(y\_0) - L'(L(x-x\_0))

= g (f(x)) - g(y\_0) -L'(f(x)-f(x\_0) -\phi (x))

= \left[ g (f(x)) - g(y\_0) -L' (f(x)-f(x\_0))\right] + L'(\phi (x))

 = \psi(f(x))+L'(\phi (x)).

Wir zeigen als nächstes \lim\_{x\rightarrow x\_0} \|\rho (x)\| / \|x-x\_0\| = 0. Dafür ist es nach der obigen Umformung ausreichend, die Gleichheiten

\lim\_{x\rightarrow x\_0} \frac{\| \psi(f(x))\|}{\| x-x\_0 \|} = 0,

\lim\_{x\rightarrow x\_0} \frac{\| L'(\phi (x))\|}{ \| x-x\_0\| } = 0

zu beweisen. Aber aus der Fundamentalabschätzung für multilineare Abbildungen wissen wir, dass eine Konstante A' existiert mit \|L'(\phi (x))\|\leq A' \|\phi (x)\|, und dies impliziert

0 \leq \lim\_{x\rightarrow x\_0} \frac{\| L'(\phi (x))\| }{\|x-x\_0\|}

\leq A' \lim\_{x\rightarrow x\_0} \frac{\|\phi (x)\|}{\|x-x\_0\| } = 0.

Der zweite Limes ist etwas schwieriger zu behandeln. Ist \epsilon>0, so existiert wegen \lim\limits\_{y\rightarrow y\_0} \frac{\|\psi(y)\|}{\|y-y\_0\|}=0 ein \delta>0 mit

\| \psi(f(x))\| < \epsilon \| f(x) - y\_0 \| \text{ für } \| f(x) -y\_0\| < \delta.

Nun ist als differenzierbare Funktion f in x\_0 stetig, so dass ein

\delta'>0 mit \|f(x)-y\_0\|<\delta für alle \|x-x\_0\|< \delta' existiert.

Damit gilt für alle x mit \|x-x\_0\|

< \delta' \| \psi(f(x))\| \leq \epsilon \| f(x) - y\_0 \|

= \epsilon \|\phi (x) +L(x-x\_0)\| \leq \epsilon \|\phi (x)\|+ \epsilon A \|x-x\_0\|,

wobei A die zu L gehörende Konstante in der Fundamentalabschätzung ist. Teilt man durch \|x-x\_0\|, so ergibt sich

0 \leq \frac{\|\psi(f(x))\|}{\|x-x\_0\|} \leq \epsilon \frac{\|\phi (x)\|}{\|x-x\_0\|} + \epsilon A

und damit

0 \leq \limsup\_{x\rightarrow x\_0} \frac{\|\psi(f(x))\|}{\|x-x\_0\|} \leq 0 + \epsilon A.

Da \epsilon beliebig klein werden kann, folgt die Behauptung. Wir zeigen als letztes die Quotientenregel (4). Die Funktion

f/g: \R^n\rightarrow \R ist die Verknüpfung von \Psi:\R^n\rightarrow \R^2,

\Psi(x)=(f(x), g(x)) und \Phi:\R^2\rightarrow \R, \Phi(x,y)=x/y.

Aber nach (6) ist D\Psi=(Df, Dg), und D\Phi(x,y)(a\_1,a\_2) = \frac{ya\_1 - xa\_2}{y^2}.

In der Tat,

\frac{| \Phi(x+a\_1,y+a\_2) - \Phi(x,y) - \frac{ya\_1 - xa\_2}{y^2} |}{\sqrt{a\_1^2+ a\_2^2}}

= \frac{| \frac{x+a\_1}{y+a\_2} - \frac{x}{y} - \frac{ya\_1 - xa\_2}{y^2}|}{\sqrt{a\_1^2+ a\_2^2}}

= \frac{| -ya\_1a\_2+xa\_2^2| }{y^2|y+a\_2|\sqrt{a\_1^2+ a\_2^2}}

= \frac{ |a\_2|\cdot |xa\_2 - y a\_1|}{y^2|y+a\_2|\sqrt{a\_1^2+ a\_2^2}}

\leq |a\_2| \frac{\sqrt{x^2+ y^2}\sqrt{a\_1^2+ a\_2^2}}{y^2|y+a\_2|\sqrt{a\_1^2+ a\_2^2}} \stackrel{(a\_1,a\_2) \rightarrow (0,0)}{\longrightarrow} 0.

Nun liefert die Kettenregel

D\left[ \frac{f}{g}\right](x\_0) = D(\Phi\circ \Psi)(x\_0)

= D\Phi (\Psi(x\_0))\circ D\Psi(x\_0) = D\Phi (\Psi(x\_0))\circ (Df(x\_0), Dg(x\_0))

= \frac{g(x\_0) Df(x\_0) - f(x\_0) Dg(x\_0)}{g(x\_0)^2}

und wir erhalten die Behauptung.

</beweis>

Ist L: \R^n\rightarrow \R eine lineare Abbildung und \langle \cdot , \cdot \rangle das euklidische Skalarprodukt auf \R^n, so ist aus der linearen Algebra bekannt, dass es genau einen Vektor v\_L\in\R^n gibt, welcher

L(x) = \langle x, v\_L\rangle \forall x\in\R^n

erfüllt. Dies führt zu folgender

<definition> Definition 5.1.15

Sei U offen in \R^n und f: U\rightarrow \R differenzierbar in x\_0 mit Ableitung Df(x\_0): \R^n\rightarrow \R. Dann ist der Gradient von f am Punkt x\_0 als derjenige Vektor

\grad f(x\_0)\in\R^n gegeben, welcher durch

\langle \grad f(x\_0),\vec{a}\rangle = Df(x\_0)(\vec{a}) = \nabla\_{\vec{a}} f(x\_0)

eindeutig bestimmt ist.

</definition>

Der Gradient hat eine wichtige geometrische Interpretation. In diesem Zusammenhang führen wir den Begriff der Niveau-Menge M\_c einer Funktion f: U\rightarrow \R ein, unter der man die Menge

M\_c := \{ x\in U | f(x)=c \} , c \in \R,

derjenigen Punkte, wo sie einen konstanten Wert annimmt, versteht.

<lemma> Lemma 5.1.16

(1) Der Gradient \grad f(x\_0) gibt diejenige Richtung an,

in der die Funktion im Punkte x\_0 am schnellsten wächst.

(2) Ist \gamma: ]a,b[\rightarrow M\_c eine differenzierbare Kurve (Dieser Begriff wird im folgenden noch präzisiert werden.) in einer Niveaufläche, so gilt \gamma'(t) \perp \grad f(\gamma(t)).

</lemma>

<beweis>

(1) Die Richtung steilstens Anstiegs ist genau die Richtung mit maximaler Richtungsableitung. Ist \|\vec{a}\|=1, so gilt nach der Cauchy-Schwarz-Ungleichung

\nabla\_{\vec{a}} f(x\_0) = \langle \grad f(x\_0), \vec{a}\rangle \leq \| \grad f(x\_0)\| ,

wobei Gleichheit genau dann vorliegt, falls

\vec{a}=\frac{\grad f(x\_0)}{\|\grad f(x\_0) \|}.

Damit wird die Richtungsableitung in der Tat in Richtung der Gradientenlinien maximal.

(2) Die Bilder der Kurve \gamma(t) sind Punkte in U, auf denen f konstant ist, mithin f(\gamma(t))=c gilt, und wir betrachten die Verknüpfung f\circ\gamma: ]a,b[ \stackrel{\gamma}{\longrightarrow} U\stackrel{f}{\longrightarrow} \R. Nach der Kettenregel ist nun

0 = D(f\circ \gamma)(t) = Df(\gamma(t))\circ \gamma'(t) = Df(\gamma(t)) ( \gamma'(t))

=\eklm{\grad f(\gamma(t)), \gamma ' (t) }.

Dies bedeutet genau \gamma'(t) \perp \grad f(\gamma(t)).

</beweis>

<beispiel>

Die Abbildung f(x,y)=x^2+3y^2 hat als Niveaulinien genau Ellipsen, welche man sinnvollerweise als Flächen z=f(x,y) darstellt. Um die Gradientenvektorfelder besser darstellen zu können, wurden die Niveaulinien in die Eben projiziert.

Abbildung: Fläche f(x,y)=x^2+3y^2 mit Niveaulinien und Gradientenvektorfeldern

</beispiel>

## 5.2 Partielle Ableitungen und die Matrix-Darstellung der Ableitung

Sei U\subset\R^n eine offene Teilmenge und f: U\rightarrow \R eine Funktion. Wir bezeichnen mit \mklm{e\_1,..., e\_n} die Standardbasis von \R^n.

<definition> Definition 5.2.1

Die i-te partielle Ableitung von f an der Stelle a ist die Richtungsableitung in Richtung e\_i, in Zeichen \frac{\partial f}{\partial x\_i}(a) := \nabla\_{e\_i}f (a) = \lim\_{t\rightarrow 0} \frac{f(a+te\_i) - f(a)}{t}.

</definition>

<beispiel> Beispiel 5.2.2

In der Situation des soeben betrachteten Beispiels f: \R^2\rightarrow \R, f(x,y)=x^2+3y^2, ergibt sich

\frac{\partial f}{\partial x} (x\_0,y\_0) = \lim\_{t\rightarrow 0}\frac{(x\_0+t)^2 + 3y\_0^2 - x\_0^2-3y\_0^2}{t} = 2x\_0,

\frac{\partial f}{\partial y} (x\_0,y\_0) = ... = 6y\_0.

</beispiel>

<lemma> Lemma 5.2.3

Ist f: U\rightarrow \R an der Stelle a\in U differenzierbar, dann ist Df(a): \R^n\rightarrow \R eine lineare Abbildung, deren Matrix M[Df(a)] in der Basis e\_1,..., e\_n gegeben ist durch

M[Df(a)] = \left( \frac{\partial f}{\partial x\_1}(a),...,\frac{\partial f}{\partial x\_n}(a) \right).

Es gilt also

Df(a)\begin{bmatrix} b\_1 \vdots b\_n \end{bmatrix} =

\left(\frac{\partial f}{\partial x\_1}(a),..., \frac{\partial f}{\partial x\_n}(a) \right)

 \begin{bmatrix} b\_1\ \vdots\\ b\_n\end{bmatrix}

= \sum\_{i=1}^n \frac{\partial f}{\partial x\_i}(a)\cdot b\_i .

</lemma>

<beweis>

Wenn f an der Stelle a differenzierbar ist, so gilt \nabla\_{\vec{b}}f(a)=Df(a)(\vec{b}) für alle \vec{b}, also

Df(a)(e\_i) = \nabla\_{e\_i} f(a) = \frac{\partial f}{\partial x\_i} (a).

Da die darstellende Matrix M[L] einer linearen Abbildung L: \R^n\rightarrow \R in der Basis e\_1,..., e\_n aber genau durch M[L] = (L(e\_1),..., L(e\_n)) gegeben ist, folgt die Behauptung.

</beweis>

<bemerkung> Bemerkung 5.2.4

Die zweite Formel dieses Lemmas kann man auch wie folgt interpretieren.

Für Abbildungen f: U\rightarrow \R ist

\grad f(x) = \left(\frac{\partial f}{\partial x\_1},..., \frac{\partial f}{\partial x\_n}\right)

= \sum\_{i=1}^n \frac{\partial f}{\partial x\_i} (x)e\_i.

</bemerkung>

Das folgende Kriterium ist ausgesprochen nützlich, um dieDifferenzierbarkeit einer Funktion zu beweisen.

<satz> Satz 5.2.5 Kriterium für Differenzierbarkeit

Sei f: U\rightarrow \R eine Abbildung. Existieren alle partiellen Ableitungen \frac{\partial f}{\partial x\_1},...,\frac{\partial f}{\partial x\_n} auf U und sind sie dort stetig, so ist f auf ganz U differenzierbar.

</satz>

<beweis>

Aus der Existenz der partiellen Ableitungen ergibt sich als natürlicher Kandidat für die Ableitung Df(a)(h)=\sum\frac{\partial f}{\partial x\_i}(a)h\_i, wobei h=(h\_1,..., h\_n)\in\R^n.

In der Tat,betrachte für a\in U und hinreichend kleines |h|die Differenz

f(a+h)-f(a)& =& f(a\_1+h\_1,..., a\_n+h\_n)-f(a\_1,..., a\_n)

= f(a\_1+h\_1,..., a\_n+h\_n) - f(a\_1+h\_1,..., a\_{n-1}+h\_{n-1}, a\_n)

+ f(a\_1+h\_1,..., a\_{n-1}+h\_{n-1}, a\_n) - f(a\_1+h\_1,..., a\_{n-2}+h\_{n-2}, a\_{n-1}, a\_n)

+ ... + f(a\_1+h\_1,a\_2,..., a\_n) -f(a\_1,..., a\_n).

Bei festem a betrachten wir nun die Funktion g\_1(x):= f(x,a\_2,..., a\_n), x\in\R. Diese ist nach Voraussetzung stetig differenzierbar, nach dem ersten Mittelwertsatz der Differentialrechnung existiert also ein \xi\_1\in [a\_1,a\_1+h\_1] mit

f(a\_1+h\_1,a\_2,..., a\_n) -f(a\_1,..., a\_n) = h\_1 g\_1' (\xi\_1)

= h\_1 \frac{\partial f}{\partial x\_1}(\xi\_1, a\_2,..., a\_n).

Ebenso können wir g\_2(x):= f(a\_1+h\_1,x,a\_3,..., a\_n) betrachten und erhalten ein \xi\_2\in [a\_2,a\_2+h\_2] mit

f(a\_1+h\_1,a\_2+h\_2,a\_3,..., a\_n) -f(a\_1+h\_1,a\_2,a\_3,..., a\_n)

= h\_2 \frac{\partial f}{\partial x\_2}(a\_1+h\_1,\xi\_2, a\_3,..., a\_n).

So verfahren wir weiter für i=3,...,n. Wir kürzen die hier auftretenden Punkte mit

p\_i := (a\_1+h\_1,..., a\_{i-1}+h\_{i-1},\xi\_i, a\_{i+1},..., a\_n)

ab, so dass die eingangs betrachtete Differenz f(a+h)-f(a) geschrieben werden kann als

f(a+h)-f(a) = \sum\_{i=1}^n h\_i\frac{\partial f}{\partial x\_i}(p\_i).

Nun können wir den Differentialquotienten betrachten und die Cauchy-Schwarz-Ungleichung anwenden, nach welcher

\frac{|f(a+h)-f(a) - \sum\frac{\partial f}{\partial x\_i}(a)h\_i |}{\|h \|}

= \frac{ \big|\sum \left[\frac{\partial f}{\partial x\_i}(p\_i)

– \frac{\partial f}{\partial x\_i}(a) \right] h\_i\big|}{\|h \|}

 \leq \sqrt{\sum \left[\frac{\partial f}{\partial x\_i}(p\_i)

 – \frac{\partial f}{\partial x\_i}(a) \right]^2 }.

Doch für h\rightarrow 0 gilt p\_i\rightarrow a, so dass infolge der Stetigkeit der partiellen Ableitungen auch \frac{\partial f}{\partial x\_i}(p\_i) \rightarrow \frac{\partial f}{\partial x\_i}(a) gelten muss. Damit hat der Differentialquotient für h\rightarrow 0 den Grenzwert 0.

</beweis>

<bemerkung> Bemerkung 5.2.6

Für den Beweis wurde nicht nur der Punkt a\in U, sondern eine kleine offene Umgebung um diesen Punkt benötigt. Es ist also falsch, aus der Stetigkeit der partiellen Ableitungen in einem Punkt die Differenzierbarkeit in diesem Punkt zu schließen.

</bemerkung>

<beispiel> Beispiel 5.2.7

Wir zeigen an einem Beispiel, dass die partiellen Ableitungen einer differenzierbaren Funktion nicht notwendigerweise stetig sein müssen. Betrachte so die Funktion f: \R^2\rightarrow \R,

f(x,y) = \left\{ \ba{ll} (x^2+y^2)\sin\frac{1}{\sqrt{x^2+y^2}} falls x^2+y^2>0

0 falls (x,y)=(0,0).\ea\right.

Dies ist die Fläche, die durch Drehung um die z-Achse des rechten Teils des Graphen der

Funktion z= x^2\sin(1/x) entsteht.

<abbildung>

Graph der Funktion z= x^2\sin(1/x). Die Funktion wird von der Normalparabel und ihrer Spiegelung an der x-Achse beschränkt. Sie oszilliert zwischen den beiden Parabeln hin und her, wobei die Oszillationen in Richtung des Nullpunktes immer schneller und schneller werden.

<\abbildung>

Wie man nun nachrechnet, sind die partiellen Ableitungen \partial f / \partial x, \partial f/ \partial y im Punkt (0,0) unstetig. So ist etwa \frac{\partial f}{\partial x}

= 2x \left[\sin\frac{1}{\sqrt{x^2+y^2}} - \frac{1}{2\sqrt{x^2+y^2}} \cos \frac{1}{\sqrt{x^2+y^2}}\right].

Also ist \frac{\partial f}{\partial x}(0,y\_0)=0, aber

\frac{\partial f}{\partial x}(x\_0,0)=2x\_0\sin\frac{1}{|x\_0|} \pm \cos \frac{1}{|x\_0|}

 hat keinen Grenzwert für x\_0\rightarrow 0.

</beispiel>

Ist f: U\rightarrow \R eine Funktion, für die alle partiellen Ableitungen existieren, so sind diese ebenfalls Funktionen von U nach \R, und es macht Sinn, nach deren partiellen Ableitungen zu fragen. Auf diese Weise erhält man partielle Ableitungen beliebiger Ordnungen, und man vereinbart für den Fall der Existenz die Schreibweise

\frac{\partial}{\partial x\_i}\left[\frac{\partial f}{\partial x\_j}\right]

=: \frac{\partial^2 f}{\partial x\_i\partial x\_j},

\frac{\partial}{\partial x\_i}\left[\frac{\partial f}{\partial x\_i}\right]

=: \frac{\partial^2 f}{\partial x\_i^2}.

Ist nun \alpha=(\alpha\_1,...,\alpha\_n) ein Multi-Index mit \alpha\_k\in\N, so setzen wir |\alpha|=\sum \alpha\_k und schreiben \frac{\partial^{|\alpha|} f }{\partial x^\alpha} = \frac{\partial^{|\alpha| }f}{(\partial x\_1)^{\alpha\_1}... (\partial x\_n)^{\alpha\_n}}.

<definition> Definition 5.2.8

Eine Funktion f: U\rightarrow \R, U\subset\R^n heißt der Klasse C^k, in Zeichen f\in C^k(U), falls für jeden Multi-Index \alpha=(\alpha\_1,...,\alpha\_n) mit |\alpha|\leq k die partiellen Ableitungen \frac{\partial^{|\alpha|} f }{\partial x^\alpha} auf U existieren und stetig sind.

</definition>

Eine Funktion ist also genau der Klasse C^1, falls alle partiellen Ableitungen existieren und stetig sind; der Klasse C^2, falls alle ersten und zweiten partiellen Ableitungen existieren und stetig sind usw. Existieren alle zweiten partiellen Ableitungen, so hat man immer zwei Optionen bei der Wahl der Differentiationsreihenfolge, etwa:

\frac{\partial^2 f}{\partial x\_1 \partial x\_2}

und

\frac{\partial^2 f}{\partial x\_2 \partial x\_1}.

Der folgende Satz von Schwarz besagt, dass die Reihenfolge der zweiten partiellen

Ableitungen irrelevant ist, falls f\in C^2(U).

<satz> Satz 5.2.9 H.A.Schwarz, 1873

Für jede Funktion f: U\rightarrow \R der Klasse C^2(U) gilt

\frac{\partial^2 f}{\partial x\_i \partial x\_j} = \frac{\partial^2 f}{\partial x\_j \partial x\_i}

für alle i,j=1,..., n.

</satz>

<beweis>

Angenommen, es gelte in einem Punkte a \in U

\frac{\partial^2 f}{\partial x\_1 \partial x\_2}(a) - \frac{\partial^2 f}{\partial x\_2 \partial x\_1}(a)

> 0.

Aufgrund der Stetigkeit der zweiten partiellen Ableitungen existiert dann ein n-dimensionaler

Quader W= \prod\_{i=1}^n [a\_i, b\_i] \subset U, auf dem immer noch

\frac{\partial^2 f}{\partial x\_1 \partial x\_2} - \frac{\partial^2 f}{\partial x\_2 \partial x\_1}>0

gilt. Insbesondere wäre dann

\int\_W \frac{\partial^2 f}{\partial x\_1 \partial x\_2} - \frac{\partial^2 f}{\partial x\_2 \partial x\_1}

 > 0.

Zudem sichert die Stetigkeit die Riemann-Integrierbarkeit der zweiten partiellen Ableitungen. Nach dem Satz von Fubini und dem Hauptsatz der Differential- und Integralrechnung ist dann aber einerseits

\int\limits\_{[a\_1,b\_1]\times [a\_2,b\_2]} \frac{\partial^2 f}{\partial x\_1 \partial x\_2} dx\_1 dx\_2

= \int\_{a\_2}^{b\_2}\int\_{a\_1}^{b\_1} \frac{\partial^2 f}{\partial x\_1 \partial x\_2} dx\_1 dx\_2

= \int\_{a\_2}^{b\_2} \left[ \frac{\partial f}{ \partial x\_2} (b\_1,x\_2,..., x\_n)

- \frac{\partial f}{ \partial x\_2} (a\_1,x\_2,..., x\_n) \right] dx\_2

= f (b\_1,x\_2,..., x\_n)\big|\_{a\_2}^{b\_2} - f (a\_1,x\_2,..., x\_n)\big|\_{a\_2}^{b\_2}

und andererseits

\int\limits\_{[a\_1,b\_1]\times [a\_2,b\_2]} \frac{\partial^2 f}{\partial x\_2 \partial x\_1} dx\_1 dx\_2

= \int\_{a\_1}^{b\_1}\int\_{a\_2}^{b\_2} \frac{\partial^2 f}{\partial x\_2 \partial x\_1} dx\_2 dx\_1

= \int\_{a\_1}^{b\_1} \left[ \frac{\partial f}{ \partial x\_1} (x\_1,b\_2,..., x\_n)

- \frac{\partial f}{ \partial x\_1} (x\_1,a\_2,..., x\_n) \right] dx\_1

= f(x\_1,b\_2,..., x\_n)\big|\_{a\_1}^{b\_1} - f (x\_1,a\_2,..., x\_n)\big|\_{a\_1}^{b\_1}.

Diese Größen sind identisch, also sind es auch die ursprünglichen Integrale, und damit kann ihre Differenz nicht strikt größer null sein, Widerspruch.

</beweis>

<bemerkung> Bemerkung 5.2.10

Es gibt mehrere moderne Varianten dieses Satzes, die die Vertauschbarkeit der partiellen Ableitungen auch unter gewissen schwächeren Voraussetzungen garantieren, siehe z.B. Young, 1909.

</bemerkung>

Wir betrachten nun wieder Abbildungen der Gestalt f=(f\_1,..., f\_m): U\subset \R^n\rightarrow \R^m. Nach Satz 5.1.14 (6) gilt dann in einem Punkt a, in dem f differenzierbar ist, dass

\label{eq:7m}

M[Df(a)] = \begin{bmatrix} M[Df\_1(a)] \vdots\ M[Df\_m(a)]\end{bmatrix}

<satz> Satz 5.2.11

Ist f=(f\_1,..., f\_m): U\subset \R^n\rightarrow \R^m in a\in U differenzierbar, so hat die Ableitung Df(a): \R^n\rightarrow \R^m in den Standardbasen von \R^n und \R^m die Matrixdarstellung

M[Df(a)] = \begin{bmatrix}

\frac{\partial f\_1}{\partial x\_1} (a) &\cdots &\frac{\partial f\_1}{\partial x\_n} (a)\\

\vdots & \ddots & \vdots\\

\frac{\partial f\_m}{\partial x\_1} (a) &\cdots &\frac{\partial f\_m}{\partial x\_n} (a)

\end{bmatrix}

</satz>

<beweis>

Die Behauptung ist eine unmittelbare Konsequenz von \eqref{eq:7m} und Lemma 5.2.3.

</beweis>

<korollar> Korollar 5.2.12

Sind g\_1,..., g\_m: U\subset \R^n\rightarrow \R sowie f: \R^m\rightarrow \R differenzierbar, so gilt für F(x)=f(g\_1(x),..., g\_m(x)) als Abbildung von U nach \R

\frac{\partial F}{\partial x\_i}(x)

= \sum\_{j=1}^m \frac{\partial f}{\partial y\_j} \big(g\_1(x),..., g\_m(x)\big)\frac{\partial g\_j}{\partial x\_i}(x).

Sind dabei sowohl die Funktionen g\_1,..., g\_m als auch f der Klasse C^k, so ist auch F der Klasse C^k.

<\korollar>

<beweis>

Sei \Phi: U\rightarrow \R^m die Abbildung \Phi(x)=(g\_1(x),..., g\_m(x)), so dass F=f\circ \Phi gilt. Nach der Kettenregel ist DF(x) = Df(\Phi(x))\circ D\Phi(x). Gehen wir hier zu den Matrizen über, so bedeutet dies einfach

M[DF(x)] = M[Df(\Phi(x))] \cdot M[D\Phi(x)].

Aber nach vorigem Satz ist

M[DF(x)]= \left( \frac{\partial F}{\partial x\_1}(x), ... , \frac{\partial F}{\partial x\_n}(x) \right),

M[Df(\Phi(x))] = \left( \frac{\partial f}{\partial y\_1}(\Phi(x)),...,\frac{\partial f}{\partial y\_n}(\Phi(x))\right),

M[D\Phi(x)] = \begin{bmatrix}

\frac{\partial g\_1}{\partial x\_1}(x) &\cdots &\frac{\partial g\_1}{\partial x\_n}(x) \\

\vdots & \ddots & \vdots\\

\frac{\partial g\_m}{\partial x\_1}(x) &\cdots &\frac{\partial g\_m}{\partial x\_n} (x)

\end{bmatrix} .

Explizites Ausmultiplizieren und Vegleichen der Einträge an jeweils einer gegebenen Stelle liefert dann die Behauptung.

</beweis>

Wir diskutieren als Anwendung sogenannte parameterabhängige Integrale.

<beispiel> Beispiel 5.2.13

Sei f: [a,b]\times [c,d]\rightarrow \R stetig und existiere die partielle Ableitung \frac{\partial f}{\partial y}, die ebenfalls stetig sei. Wir betrachten

F(y) = \int\_a^b f(x,y) dx

und zeigen, dass F differenzierbar ist mit Ableitung

F'(y) = \int\_a^b \frac{\partial f}{\partial y}(x,y) dx.

In der Tat, da \partial f/ \partial y eine Stammfunktion besitzt, gilt für y \in [c,d]

\int\_c^y \frac{\partial f}{\partial y}(x,\xi) d\xi = f(x,y)-f(x,c), also

F(y) = \int\_a^b \left[\int\_c^y \frac{\partial f}{\partial y}(x,\xi) d\xi+f(x,c) \right] dx.

Nach dem Satz von Fubini ist weiter

F(y+h)-F(y) = \int\_a^b \left[\int\_y^{y+h} \frac{\partial f}{\partial y}(x,\xi) d\xi \right] dx

= \int\_y^{y+h}\left[\int\_a^b \frac{\partial f}{\partial y}(x,\xi) d x\right] d\xi

und damit folgt mit Satz 4.5.1 (2)

\lim\_{h\rightarrow 0} \frac{F(y+h)-F(y)}{h} = \int\_a^b \frac{\partial f}{\partial y} (x,y) dx.

</beispiel>

<beispiel> Beispiel 5.2.14

Sei f: [a,b]\times [c,d]\rightarrow \R eine stetige Funktion. Betrachte

F(x,y) = \int\_a^x f(t,y) dt .

Wie oben sieht man dann, dass

\frac{\partial F}{\partial x}(x,y) = f(x,y),

\frac{\partial F}{\partial y}(x,y) = \int\_a^x\frac{\partial f}{\partial y}(t,y) dt.

</beispiel>

<beispiel> Beispiel 5.2.15

Seien f: [a,b]\times [c,d]\rightarrow \R und g: [c,d]\rightarrow [a,b] differenzierbar und

G(y) = \int\_a^{g(y)}f(t,y) dt = F(g(y),y)

mit F wie im vorigen Beispiel. Dann gilt nach Korollar 5.2.12

G'(y) = \frac{\partial F}{\partial x}\frac{\partial g(y)}{\partial y} + \frac{\partial F}{\partial y}

= f(g(y),y) g'(y) +\int\_{a}^{g(y)} \frac{\partial f}{\partial y} (t,y) dt.

</beispiel>

In Kapitel 3 hatten wir die Differenzierbarkeit von Funktionen einer komplexen Variablen besprochen. Eine solche Funktion kann auch als Funktion zweier reeller Variablen aufgefaßt werden und es macht Sinn, nach den partiellen Ableitungen und deren Beziehung zur komplexen Differenzierbarkeit zu fragen. Wir betrachten so \C\simeq\R^2 versehen mit den Koordinaten z=x+iy, sowie eine offene Teilmenge U\subset \C und eine holomorphe Funktion f: U\rightarrow \C.

<satz> Satz 5.2.16 Cauchy-Riemann-Differentialgleichungen

Ist f: U\rightarrow \C holomorph, so ist f als Abbildung U\subset\R^2\rightarrow \R^2 differenzierbar und mit der Bezeichnung f(z)=f(x+iy)=u(x,y)+iv(x,y) gilt

\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y},

\frac{\partial u}{\partial y} = -\frac{\partial v}{\partial x} .

</satz>

<beweis>

Ist f=u+iv im Punkt z\_0=x\_0+iy\_0 \C-differenzierbar, so gilt für ein a\in\C

\label{eq:cr}

\lim\_{h\rightarrow 0}\frac{f(z\_0+h) - f(z\_0)}{h} = a, also

\lim\_{h\rightarrow 0}\frac{f(z\_0+h) - f(z\_0)- ah }{h} = 0.

Mit a=a\_1+i a\_2 und h=h\_1+ih\_2, ergibt sich nun ah= a\_1 h\_1 - a\_2 h\_2+i(h\_1 a\_2+h\_2 a\_1). Unter der Identifikation \C\simeq \R^2 definiert somit die komplexe Multiplikation mit a\in\C eine reell-lineare Abbildung A: \R^2\rightarrow \R^2, deren darstellende Matriz durch

M[A] = \begin{bmatrix} a\_1 & -a\_2 \\

 a\_2 & a\_1

\end{bmatrix}

gegeben ist. Aus \eqref{eq:cr} folgt dann

\lim\_{(h\_1,h\_2)\rightarrow 0}\frac{\| f(z\_0+h) - f(z\_0) - A(h)\|}{\|h\|} = 0.

Damit ist f in z\_0 differenzierbar mit Ableitung A und

M[A] = \begin{bmatrix} a\_1 & -a\_2\ a\_2 & a\_1 \end{bmatrix}

= \begin{bmatrix}

\frac{\partial u}{\partial x} & \frac{\partial u}{\partial y}\\

\frac{\partial v}{\partial x} & \frac{\partial v}{\partial y}

\end{bmatrix}.

Aus der Gestalt der Matrix A folgen die Cauchy-Riemann-Differentialgleichungen.

</beweis>

<bemerkung> Real- und Imaginärteil u,v erfüllen sogar die Laplace-Gleichung

\frac{\partial^2 u}{\partial x^2}+\frac{\partial^2 u}{\partial y^2}

= \frac{\partial^2 v}{\partial x^2}+\frac{\partial^2 v}{\partial y^2} = 0.

So ergibt sich aus den Cauchy-Riemann-Differentialgleichungen zunächst durch Einsetzen

\frac{\partial^2 u}{\partial x^2}+\frac{\partial^2 u}{\partial y^2}

= \frac{\partial}{\partial x} \left[\frac{\partial u}{\partial x}\right]

+ \frac{\partial}{\partial y} \left[\frac{\partial u}{\partial y}\right]

= \frac{\partial}{\partial x} \left[\frac{\partial v}{\partial y}\right]

+ \frac{\partial}{\partial y} \left[-\frac{\partial v}{\partial x}\right].

Um hieraus auf die Laplace-Gleichung zu schließen, benötigen wir noch die Stetigkeit der partiellen Ableitungen. Wie in der Funktionentheorie gezeigt wird, ist diese jedoch gegeben und die gewünschte Gleichung folgt aus der Vertauschbarkeit der partiellen Ableitungen.

</bemerkung>

## 5.3 Taylor-Entwicklung und Extrema einer Funktion mehrerer Variablen

Wir wollen nun eine Näherungsformel für Funktionen mehrerer Variablen besprechen. Als Bezeichnungen vereinbaren wir für einen Multiindex \alpha=(\alpha\_1,...,\alpha\_n)\in \N^n

|\alpha| := \alpha\_1+...+\alpha\_n

\alpha! := (\alpha\_1!)\cdot...\cdot (\alpha\_n!).

Für y=(y\_1,...,y\_n)\in\R^n sei zudem y^\alpha= y\_1^{\alpha\_1}\cdot...\cdot y\_n^{\alpha\_n}.

Die mehrdimensionale Taylorformel reduziert sich im Kern auf die eindimensionale, angewandt in einer gewissen ausgewählten Richtung. Sie ist von beweistechnischer Relevanz, wird aber selten zur Berechnung von Reihenentwicklungen verwendet. Deswegen werden wir der Frage nach der Konvergenz der mehrdimensionalen Taylorreihe nicht weiter nachgehen, sondern verweisen im Bedarfsfall auf die Literatur.

<satz> Satz 5.3.1 Taylorformel

Sei U\subset \R^n eine offene Teilmenge, f\in C^k(U) und x,y\in U zwei Punkte derart, dass die Strecke [x,y]:=\{ tx+(1-t)y: t \in [0,1]\} vollständig in U liegt. Dann existiert ein Punkt \xi\in [x,y] mit

f(x) = \sum\_{|\alpha|\leq k-1} \frac{1}{\alpha!} \frac{\partial^{|\alpha| } f }{\partial x^\alpha}(y)(x-y)^\alpha

+ \sum\_{|\alpha|=k} \frac{1}{\alpha!} \frac{\partial^{|\alpha|} f }{\partial x^\alpha}(\xi)(x-y)^\alpha .

</satz>

<beweis>

Wir betrachten die Funktion g: [0,1]\rightarrow \R, g(t)=f(y+t(x-y)). Diese ist von der Klasse C^k und nach der eindimensionalen Taylorformel gilt

f(x) = g(1) = \sum\_{\beta=0}^{k-1} \frac{1}{\beta!} \frac{d^\beta g}{dt^\beta}(0) + \frac{1}{k!}\frac{d^k g}{dt}(t\_0)

für ein t\_0 \in [0,1]. Aber mit Korollar 5.2.12 ist

\frac{d^0 g}{dt^0} (0)& =& g(0) = f(y),

\frac{d g}{dt} (0)& =& \sum\_{i=1}^n \frac{\partial f}{\partial x\_i}(y) (x\_i-y\_i)

 = \sum\_{|\alpha|=1}\frac{\partial^\alpha f}{\partial x^\alpha}(y) (x-y)^\alpha,

\frac{d^2 g}{dt^2} (0) =\sum\_{i,j=1}^n \frac{\partial^2 f}{\partial x\_i\partial x\_j}(y) (x\_i-y\_i)(x\_j-y\_j)

= \sum\_{|\alpha|=2}\frac{\partial^\alpha f}{\partial x^\alpha}(y) (x-y)^\alpha\frac{2!}{\alpha!}.

Dies liegt daran, dass in der Summe über alle i und j sowohl

\partial / \partial x\_i \partial x\_j als auch \partial / \partial x\_j\partial x\_i vorkommt, in der

Summe über |\alpha|=2 dagegen nur einmal, so dass ein Korrekturfaktor

nötig ist. Insgesamt erhält man

\frac{d^j g}{dt^j}(0) = \sum\_{|\alpha|=j}

\frac{j!}{\alpha!}\frac{\partial^{|\alpha|} f}{\partial x^\alpha} (y) (x-y)^\alpha.

und mit \xi:=y+t\_0(x-y) folgt die Behauptung.

</beweis>

Zur Veranschaulichung schreiben wir diese Formel für den Fall n=2 explizit aus. Dabei verwenden wir die Notation x=(x\_1,x\_2) und setzen h:=x-y, lassen bei den höheren partiellen Ableitungen von f den Punkt y weg und kürzen die partiellen Ableitungen durch

\frac{\partial}{\partial x\_i} =: \partial\_i

ab. Dann ist für n=2

f(x\_1,x\_2) = f(y\_1,y\_2)+ \partial\_1 f(y\_1,y\_2) h\_1+\partial\_2 f h\_2

+ \frac{1}{2}\partial\_{11}f h\_1^2+ \frac{1}{2}\partial\_{22}f h\_2^2+ \partial\_{12}f h\_1 h\_2

+ \frac{1}{6}\partial\_{111}f h\_1^3 + \frac{1}{6}\partial\_{222}f h\_2^3

+ \frac{1}{2}\partial\_{112}f h\_1h\_2^2+ \frac{1}{2}\partial\_{122}f h\_1^2h\_2+ ...

<beispiel> Beispiel 5.3.2 Tangentialhyperebene

Sei f eine reellwertige, stetig differenzierbare Funktion auf U\subset\R^n. Dann stellt der Graph F=\{ (x,z)\in\R^{n+1} : f(x)=z\} von f eine Fläche in \R^{n+1} dar.

Das Taylorpolynom erster Ordnung

z = f(y)+\sum\_{j=1}^n \partial\_{j}f(y)(x\_j-y\_j) = f(y)+Df(y)(x-y)

definiert eine affine Hyperebene in \R^{n+1} im Flächenpunkt (y,f(y)), welche als Tangentialhyperebene bezeichnet wird. Im Fall n=1 führt dies auf die bekannte Gleichung für die Tangente an einen Graphen, z=f(y)+ f'(y)(x-y).

</beispiel>

Wir wollen nun die Extremalstellen von Funktionen mehrerer Variablen studieren.

<definition> Definition 5.3.3

Sei U\subset\R^n offen, f: U\rightarrow \R differenzierbar.

(1) x\_0 \in U heißt lokales Minimum von f, falls ein \epsilon>0 existiert, so dass f(x)\geq f(x\_0) für alle x mit |x-x\_0|<\epsilon gilt.

(2) x\_0\in U heißt lokales Maximum von f, falls ein \epsilon>0 existiert, so dass f(x)\leq f(x\_0) für alle x mit |x-x\_0|<\epsilon gilt.

(3) x\_0\in U heißt kritischer oder stationärer Punkt von f, falls Df(x\_0)=0 gilt. In jedem kritischen Punkt definiert man eine symmetrische Bilinearform, die Hessesche Bilinearform

Hf(x\_0): \R^n\times \R^n\rightarrow \R durch

Hf(x\_0)(\vec{a},\vec{b}) := \sum\_{i,j=1}^n \frac{\partial^2 f}{\partial x\_i \partial x\_j}a\_i b\_j,

falls die zweiten partiellen Ableitungen existieren.

</definition>

<beispiel> Beispiel 5.3.4 Schmiegquadrik

Auch das Taylorpolynom zweiter Ordnung hat noch eine geometrische Bedeutung für Abbildungen f: U\subset\R^n\rightarrow \R der Klasse C^2. An einem kritischen Punkt lässt es sich nämlich mittels der Hesseschen Bilinearform schreiben als

z = f(y)+ Df(y)(x-y) + \frac{1}{2} Hf(y)(x-y,x-y)

und definiert eine Quadrik in \R^{n+1}, welche die Schmiegquadrik von f im Punkt Flächenpunkt (y,f(y)) genannt wird. Im Fall n=2 lehrt die lineare Algebra, dass die Schmiegquadrik durch eine geeignete Koordinatentransformation auf eine der folgenden Normalformen gebracht werden kann:

* z=x^2+y^2 (elliptisches Paraboloid)
* z=x^2-y^2 (hyperbolisches Paraboloid)
* z=x^2 (parabolischer Zylinder)

</beispiel>

In diesem Zusammenhang erinnern wir nun an folgende Definitionen aus der linearen Algebra.

<definition> Definition 5.3.5

Eine symmetrische Bilinearform A: \R^n\times \R^n\rightarrow \R,

A(\vec{a},\vec{b}) = \sum\_{i,j=1}^n A\_{ij}a\_i b\_j heißt

(1) positiv semidefinit, falls A(\vec{a},\vec{a})\geq 0 für alle \vec{a}\in\R^n,

(2) positiv definit, falls A(\vec{a},\vec{a}) >0 für alle 0\neq \vec{a}\in\R^n; für n=2 ist das äquivalent zu A\_{11}>0 und \det A>0,

(3) negativ semidefinit, falls A(\vec{a},\vec{a})\leq 0 für alle \vec{a}\in\R^n,

(4) negativ definit, falls A(\vec{a},\vec{a}) <0 für alle 0\neq \vec{a}\in\R^n; für n=2 ist das äquivalent zu A\_{11}<0 und \det A>0.

</definition>

<satz> Satz 5.3.6

Sei f\in C^3(U).

(1) Ist x\_0\in U ein lokales Minimum bzw. Maximum, so ist x\_0 ein kritischer Punkt von f und Hf(x\_0) ist positiv bzw. negativ semidefinit.

(2) Ist x\_0\in U ein kritischer Punkt von f und Hf(x\_0) positiv bzw. negativ definit, so ist x\_0 ein lokales Minimum bzw. Maximum von f.

</satz>

<beispiel> Beispiel 5.3.7

Wir betrachten die Funktion f(x,y)=\epsilon\_1 x^2+\epsilon\_2 y^2 auf ganz \R^2. Wegen

Df(x\_0,y\_0) = \big(2\epsilon\_1x\_0,2\epsilon\_2 y\_0 \big) ist (0,0) ein kritischer Punkt mit Hessescher Bilinearform Hf(0,0) = \begin{bmatrix} 2\epsilon\_1 & 0\ 0 & 2\epsilon\_2 \end{bmatrix}.

Demnach gilt

(1) Für \epsilon\_1=\epsilon\_2=1 ist Hf(0,0) positiv definit, und es liegt ein lokales Minimum vor;

(2) Für \epsilon\_1=\epsilon\_2=-1 ist Hf(0,0) negativ definit, und es liegt ein lokales Maximum vor;

(3) Für \epsilon\_1+\epsilon\_2=0 ist (0,0) zwar kritisch, aber kein Extremum. Es liegt ein sogenannter Sattelpunkt vor.

</Aufzählung>

</beispiel>

Die genauere Klassifizierung der Sattelpunkte ist Gegenstand der Morse-Theorie.

Bevor wir zum Beweis des obigen Satzes schreiten, bemerken wir kurz folgenden Sachverhalt der linearen Algebra, der einfach einzusehen ist: Ist \lambda >0 der kleinste Eigenwert der Bilinearform A, insbesondere also A positiv definit, so gilt A(x,x)\geq \lambda \|x\|^2 für alle Vektoren, und Gleichheit tritt ein, falls x ein Eigenvektor von A ist.

<beweis> Beweis von Satz 5.3.6

(1) Sei x\_0\in U ein lokales Minimum von f und \vec{a}\in\R^n ein beliebiger Vektor.

Die Funktion g: \R\rightarrow \R, g(t)=f(x\_0+t\vec{a}) hat dann ein lokales Minimum bei t=0, d.h.~g'(0)=0, g''(0)\geq 0. Ausgedrückt durch f ist dies gleichbedeutend mit

0 = \sum\_1^n \frac{\partial f}{\partial x\_i} (x\_0)a\_i,

0 \leq \sum\_{i,j} \frac{\partial^2 f}{\partial x\_i\partial x\_j} (x\_0)a\_ia\_j

= Hf(x\_0)(\vec{a},\vec{a}).

 Weil diese Identitäten für einen beliebigen Vektor \vec{a} gelten, muss \frac{\partial f}{\partial x\_i} (x\_0)=0 für alle i sein. Also ist x\_0 ein kritischer Punkt und Hf(x\_0) ist positiv semidefinit. Analog behandelt man den Fall eines lokalen Maximums.

(2) Sei x\_0 ein kritischer Punkt und \epsilon>0 derart, dass K (x\_0,\epsilon)\subset U gilt. Für y \in K (x\_0,\epsilon) gilt nach Satz 5.3.1, weil f der Klasse C^3 ist, für ein \xi \in [x\_0,y] f(y) = f(x\_0)+ 0+ \frac{1}{2}\sum\_{i,j=1}^n

\frac{\partial^2 f}{\partial x\_i\partial x\_j} (x\_0) ((x\_0)\_i -y\_i)((x\_0)\_j -y\_j)

+\sum\_{|\alpha|=3} \frac{1}{\alpha!}\frac{\partial^\alpha f}{\partial x^\alpha}(\xi)(x\_0-y)^\alpha\\

= f(x\_0)+ Hf(x\_0)(x\_0-y,x\_0-y) +\sum\_{|\alpha|=3}

\frac{1}{\alpha!}\frac{\partial^\alpha f}{\partial x^\alpha} (\xi)(x\_0-y)^\alpha.

Angenommen, Hf(x\_0) ist positiv definit mit kleinstem Eigenwert \lambda>0, so dass

Hf(x\_0)(x\_0-y,x\_0-y) \geq \lambda\|x\_0-y\|^2.

Für die ursprüngliche Gleichung ergibt sich dann

f(y) - f(x\_0) \geq \lambda\|x\_0-y\|^2 + \sum\_{|\alpha|=3}

\frac{1}{\alpha!}\frac{\partial^{|\alpha|} f}{\partial x^\alpha} (\xi)(x\_0-y)^\alpha.

Wir möchten zeigen, dass die rechte Seite positiv ist. Ist der zweite Summand null, so ist dies offenbar der Fall. Im gegenteiligen Fall können wir ein \mu wählen mit

\mu < \epsilon,

0 < \mu < \lambda \left[\sum\_{|\alpha|=3} \frac{1}{\alpha!}

\sup\_{z\in K (x\_0,\epsilon)} \bigg|\frac{\partial^{|\alpha|} f}{\partial x^\alpha} (z)\bigg|

\right]^{-1} .

Ist nun 0< \|x\_0-y\|<\mu und |\alpha|=3, so gilt \frac{|(x\_0-y)^\alpha|}{\|x\_0-y \|^2}\leq\mu.

Tatsächlich, ist etwa \alpha=(0,\ldots,0,1,0,\ldots,0,1,0,\ldots,0,1,0\ldots, 0) mit Einsen

an den Stellen i,j,k, so folgern wir unter Verwendung von ab/(a^2+b^2) \leq 1/2

\frac{|(x\_0-y)^\alpha|}{\|x\_0-y \|^2}

= \frac{|(x\_0)\_i - y\_i|\cdot|(x\_0)\_j - y\_j|\cdot |(x\_0)\_k - y\_k|}{ ((x\_0)\_1 - y\_1)^2+\ldots + ((x\_0)\_1 - y\_1)^2}

\leq |(x\_0)\_i - y\_i|\cdot \frac{|(x\_0)\_j - y\_j|\cdot |(x\_0)\_k - y\_k|}{((x\_0)\_j - y\_j)^2+ ((x\_0)\_k - y\_k)^2}

\leq \frac{1}{2} |(x\_0)\_i - y\_i|

\leq \frac{1}{2} \|x\_0-y\| \leq \|x\_0-y\| < \mu.

Ebenso zeigt man diese Ungleichung, falls \alpha nur zwei Einträge (=1 und 2) oder nur einen Eintrag (=3) hat. Insgesamt gilt für jedes y\in K (x\_0,\mu)

\frac{f(y)-f(x\_0)}{\|x\_0-y\|^2}

\geq \lambda+ \sum\_{|\alpha|=3} \frac{1}{\alpha!}\frac{\partial^{|\alpha|} f}{\partial x^\alpha}

(\xi)\frac{(x\_0-y)^\alpha}{\|x\_0-y\|^2 }

\geq \lambda - \sum\_{|\alpha|=3} \frac{1}{\alpha!} \sup\_{z\in K (x\_0,\epsilon)} \bigg|\frac{\partial^{|\alpha|} f}{\partial x^\alpha} (z)\bigg|

\frac{\big |(x\_0-y)^\alpha\big |}{\|x\_0-y\|^2 }\\

\geq \lambda - \sum\_{|\alpha|=3} \frac{1}{\alpha!} \sup\_{z\in K (x\_0,\epsilon)} \bigg|\frac{\partial^{|\alpha|} f}{\partial x^\alpha} (z)\bigg|\mu > 0

nach Wahl von \mu. Damit ist gezeigt, dass f(x\_0) ein echtes lokales Minimum von f ist.

</beweis>

Wir beenden diesen Abschnitt mit einem Beispiel.

<beispiel> Beispiel 5.3.8

Sei V ein reeller Vektorraum, \langle \cdot ,\cdot \rangle ein positiv definites Skalarprodukt auf V und A ein Endomorphismus V\rightarrow V, der bezüglich des Skalarprodukts symmetrisch ist, also \langle Ax,y\rangle=\langle x,Ay\rangle erfüllt. Wir betrachten die Funktion

f: V-\{0\}\rightarrow \R, f(x) = \frac{\langle Ax,x\rangle}{\langle x,x\rangle}.

und fragen nach den kritischen Punkte von f. Nenner und Zähler sind Multilinearformen, so dass nach den bekannten Rechenregeln

Df(x)(v) = \frac{(\langle v, Ax\rangle + \langle x, Av\rangle)\langle x, x\rangle

- 2\langle Ax,x\rangle \langle x,v\rangle}{\langle x,x\rangle^2}

 = 2 \frac{\langle v, Ax\rangle\langle x, x\rangle

- \langle Ax,x\rangle \langle x,v\rangle}{\langle x,x\rangle^2}.

Damit ist Df(x)(v)=0 gleichbedeutend mit

\langle v, \langle x, x\rangle Ax - \langle Ax,x\rangle x\rangle=0

für alle v\in V.

Weil das Skalarprodukt nicht entartet ist, muss alsdann \langle x, x\rangle Ax = \langle Ax,x\rangle x gelten. Ax ist also proportional zu x und x ein Eigenvektor von A. Umgekehrt sieht man, dass ein Vektor x mit Ax=\lambda x in der Tat Df(x)=0 impliziert. Ob wirklich ein Extremum vorliegt, hängt dann entscheidend von den Eigenwerten von A ab. Sind diese zum Beispiel alle positiv und \lambda der kleinste, so gilt \langle Ax,x\rangle\geq \lambda \langle x,x\rangle, also liegt in einem Eigenvektor zu \lambda ein globales Minimum vor.

</beispiel>

<bemerkung> Bemerkung 5.3.9

Wir haben in diesem Abschnitt Kriterien zur Extremwertbestimmung für Funktionen kennengelernt, die auf offenen Mengen definiert sind. Auf abgeschlossenen Menge kann obige Methode nicht angewandt werden. Hierzu betrachten wir als Beispiel die Funktion

f: \R^2\rightarrow \R, f(x,y)=-(x^2+y^2)^2 +x^2-y^2.

Die kritischen Punkte ergeben sich aus den Bedingungen

\frac{\partial f}{\partial x} = -2x(2x^2+2y^2-1) = 0,

\frac{\partial f}{\partial y} = -2y(2x^2+2y^2+1) = 0.

Weil der zweite Faktor von \partial f/\partial y immer mindestens 1 ist, liefert die zweite Bedingung notwendigerweise y=0. Aus der ersten Bedingung folgt dann x=0 oder x=\pm 1/\sqrt{2}. Damit haben wir insgesamt drei kritische Punkte, nämlich (0,0) und (\pm 1/\sqrt{2},0).

Die Hessesche Form lautet

\begin{bmatrix} -12 x^2-4y^2+2 & -8xy\ -8xy & -4x^2-12y^2-2\end{bmatrix},

also gilt in den kritischen Punkten

Hf(0,0) = \begin{bmatrix} 2 & 0 \ 0 & -2 \end{bmatrix},

Hf(\pm 1/\sqrt{2},0) = \begin{bmatrix} -4 & 0 \ 0 & -4 \end{bmatrix}.

Damit ist (0,0) ein Sattelpunkt, (\pm 1/\sqrt{2},0) sind lokale Maxima mit Funktionswert f(\pm 1/\sqrt{2},0)=1/4. Kritische Punkte, die einem lokalen Minimum entsprächen, gibt es nicht.

Wir schränken nun f auf \bar{D}(1)=\{(x,y)\in\R^2 : x^2+y^2\leq 1\} ein. Diese Menge ist kompakt, die stetige Funktion f muss auf ihr also ihr Minimum und ihr Maximum annehmen, obwohl wir insbesondere für ein Minimum gar keinen Kandidaten haben. Letzteres muss demzufolge auf dem Rand, wo die Methode der kritischen Punkte nicht anwendbar ist, liegen. Die lokalen Extremwerte hat die obige Rechnung nur auf der offenen Kugel

K (0,1)=\text{Int} \bar{D}(1) geliefert. Den Rand {\partial} \bar{D}(1)=S^1

muss man getrennt beachten. Die Einschränkung von f auf S^1 ist

aber f\big|\_{s^1} = -1 +x^2-y^2 = -2y^2, y\in [-1,1],

mithin nur von y abhängig und maximal für y=0 mit Funktionswert f(\pm 1, 0)=0 und minimal für y=\pm 1 mit Wert f(0, \pm 1)=-2. Damit liegt das globale Maximum von f\big|\_{\bar{D}(1)} in den beiden kritischen Punkten (\pm 1/\sqrt{2},0), das globale Minimum auf dem Rand bei (0, \pm 1). In Abschnitt 5.5 werden wir eine Methode kennenlernen, mit der der Rand systematisch behandelt werden kann.

</bemerkung>

## 5.4 Der Satz über die Umkehrabbildung

Für Funktionen einer Variablen ist folgender Sachverhalt bekannt.

<satz> Satz 5.4.1

Ist U \subset \R offen und zusammenhängend, f: U\rightarrow \R der Klasse C^k und überall f'\neq 0, dann ist f: U\rightarrow f(U) bijektiv, also global invertierbar, und f^{-1} ebenfalls der Klasse C^k. Insbesondere ist f(U) ebenfalls offen und zusammenhängend.

</satz>

Das folgende Beispiel zeigt, dass für Abbildungen f: \R^n\rightarrow \R^n diese Aussage nicht mehr richtig ist.

<beispiel> Beispiel 5.4.2

Sei f: \R^2\rightarrow \R^2 definiert durch f(x,y)= (e^x\cos y, e^x\sin y). Ihr Differential erfüllt

Df(x,y) = \begin{bmatrix} e^x \cos(y) & - e^x \sin(y)

e^x \sin(y) & e^x \cos(y) \end{bmatrix},

\det Df(x,y) = e^{2x} > 0.

Damit ist Df überall invertierbar, aber wegen f(x,y+2k\pi)=f(x,y) ist f nicht global invertierbar.

</beispiel>

Unser Ziel ist es zu zeigen, dass aus \det Df \neq 0 zumindest lokal die Existenz einer Umkehrabbildung folgt.

<lemma> Lemma 5.4.3

Seien W ein Quader und U eine offene Menge derart, dass W\subset U\subset\R^n. Ist f=(f\_1,\ldots, f\_n): U\rightarrow \R^n eine Abbildung der Klasse C^1, für die eine Konstante M existiert mit

\bigg| \frac{\partial f\_i}{\partial x\_j} (a)\bigg| \leq M \forall 1\leq i,j\leq n, \forall a\in W,

so gilt

\| f(a) - f(b)\| \leq n^2 M \|a-b\| \forall a,b\in W.

</lemma>

<beweis>

Ähnlich wie beim Beweis von Satz 5.2.5 schreiben wir für die i-te Komponente f\_i die Differenz

f\_i(a)-f\_i(b)

= \sum\_{j=1}^n [f\_i(a\_1,\ldots, a\_j,b\_{j+1},\ldots, b\_n) - f\_i(a\_1,\ldots, a\_{j-1},b\_{j},\ldots, b\_n)]

und fassen den Summanden zum Index j als Funktion von a\_j auf. Der Mittelwertsatz der Differentialrechnung liefert dann Zahlen \xi\_j\in [a\_j,b\_j] mit

f\_i(a)-f\_i(b)

= \sum\_{j=1}^n (a\_j -b\_j) \frac{\partial f\_i}{\partial x\_j}(a\_1,\ldots, a\_{j-1},\xi\_j,b\_{j+1},\ldots, b\_n).

Die Punkte (a\_1,\ldots, a\_{j-1},\xi\_j,b\_{j+1},\ldots, b\_n) liegen wieder in W, so dass

|f\_i(a)-f\_i(b)| \leq \sum\_{j=1}^n M |a\_j -b\_j| \leq nM\|a-b\|.

Für die Gesamtnorm ergibt sich nun die Abschätzung

\|f(a)-f(b)\|

= \sqrt{ \sum |f\_i(a)-f\_i(b)|^2 } \leq \sqrt{n^2M^2\|a-b\|^2 n } = \n^{3/2} M \|a-b\|.

</beweis>

<satz> Satz 5.4.4 Satz über die Umkehrabbildung

Sei U\in\R^n offen, f: U\rightarrow \R^n stetig differenzierbar, und gelte \det Df(a) \neq 0 für ein a\in U. Für b:=f(a) existieren dann offene Mengen V\_a\subset U und V\_b\subset\R^n derart, dass

* V\_b=f(V\_a),
* f\_{|V\_a}: V\_a\rightarrow V\_b ist bijektiv,
* f^{-1}: V\_b\rightarrow V\_a ist stetig differenzierbar.

Ist sogar f\in C^k(U), k\geq 1, dann ist auch f^{-1}: V\_b\rightarrow V\_a der Klasse C^k auf V\_b.

</satz>

<beweis>

Der Beweis ist in mehrere Schritte untergliedert. Zunächst bemerken wir, dass mit L=Df(a) in diesem Punkt

D(L^{-1}\circ f)= DL^{-1} \circ Df = L^{-1}\circ L={\mathrm{Id}}\_{\R^n}

ist. Wir können also annehmen, dass

Df(a)={\mathrm{Id}}\_{\R^n}, mithin \partial f\_i(a)/\partial x\_j =\delta\_{ij}

gilt. Denn ist L^{-1} \circ f lokal umkehrbar, so auch f=L \circ ( L^{-1} \circ f), da L eine Bijektion ist. Wir wählen als nächstes einen Quader K \subset U, der

a \in\text{Int} K,

\det Df |\_K \neq 0,

\left|\frac{\partial f\_i}{\partial x\_j}(x) - \frac{\partial f\_i}{\partial x\_j}(a) \right| < \frac{1}{2n^2} \forall x\in K

erfüllt. Für die neue Funktion g(x):=f(x)-x gilt für alle x\in K

\left|\frac{\partial g\_i}{\partial x\_j}(x) \right|

= \left|\frac{\partial f\_i}{\partial x\_j}(x) - \delta\_{ij}\right| < \frac{1}{2n^2},

so dass das vorherige Lemma angewandt werden kann. Man erhält

\|f(x\_1) - x\_1 - f(x\_2)+ x\_2\| \leq \frac{1}{2}\|x\_1 - x\_2\| \forall x\_1, x\_2\in K.

Mit der Dreiecksungleichung folgt

\|x\_1 - x\_2\| -\|f(x\_1) - f(x\_2) \| \leq \|f(x\_1) - x\_1 - [f(x\_2) - x\_2]\| \leq \frac{1}{2}\|x\_1 - x\_2\|

und damit \tag{\*}

\|x\_1 - x\_2\| \leq 2 \|f(x\_1) - f(x\_2) \| \forall x\_1, x\_2\in K.

Daraus folgt bereits, dass f auf K injektiv ist.

Insbesondere impliziert a \notin {\partial} K, dass f(a)\notin f( {\partial} K). Jedoch muss f(\stackrel{\circ}K) nicht notwendig offen sein; stattdessen ist f(\partial K) kompakt. Also hat f(a) von diesem Kompaktum einen echt positiven Abstand

0 < \delta := \inf\_{x\in {\partial} K} \|f(a)-f(x)\|.

Wir setzen jetzt V\_b:= \{ y\in\R^n : \|y-f(a) \|<\delta / 2 \}.

Für y\in V\_b und x\in{\partial} K gilt \|y-f(a)\|<\delta/2,

\|f(x)-f(a)\| \geq \delta,

also \tag{\*\*}

\|y-f(a)\| < \|y -f(x)\| \forall y\in V\_b, x\in{\partial} K.

Wir wollen nun zeigen, dass zu jedem y\in V\_b genau ein x\in \text{Int} K existiert mit f(x)=y. Zu einem solchen y\in V\_b betrachten wir

h\_y: K\rightarrow \R, h\_y(x)=\|y-f(x)\|^2 = \sum (y\_i-f\_i(x))^2.

Weil K kompakt ist, nimmt h\_y darauf ihr (absolutes) Minimum an. Aber (\*\*) impliziert für x \in \partial K, dass h\_y(a)< h\_y (x). Also kann h\_y das Minimum nicht auf \partial K annehmen, sondern muss es in einem Punkt x\_0\in\text{Int} K tun. An dieser Stelle gilt nach Satz 5.3.6

\frac{\partial h\_y}{\partial x\_j}(x\_0) = 0

= 2\sum\_{i=1}^n (y\_i-f\_i(x\_0)) \frac{\partial f\_i}{\partial x\_j} (x\_0).

Da x\_0\in K, gilt \det \left[\frac{\partial f\_i}{\partial x\_j}(x\_0)\right]\neq 0, weswegen y\_i - f\_i(x\_0)=0. Also ist y=f(x\_0). Auf V\_a :=\text{Int} K\cap f^{-1}(V\_b) ist somit f\big|\_{V\_a}: V\_a\rightarrow V\_b bijektiv, und beide Mengen V\_a, V\_b sind offen. Wir schreiben nun (\*) mit x\_i=f^{-1}(y\_i)\in V\_a\subset K, i=1,2, um und erhalten

\| f^{-1}(y\_1)- f^{-1}(y\_2) \| \leq 2\|y\_1-y\_2\| \forall y\_1,y\_2\in V\_b.

Daraus folgt, dass \big(f\_{|V\_a} \big)^{-1}: V\_b\rightarrow V\_a stetig ist. Es verbleibt zu zeigen, dass diese Abbildung stetig differenzierbar bzw. der Klasse C^k ist, falls auch f\in C^k(U) ist.

Wir betrachten x\in V\_a, y=f(x)\in V\_b und setzen L\_x=Df(x). Wir behaupten, dass f^{-1} in y differenzierbar ist und Df^{-1}(y)=L\_x^{-1}=[Df(x)]^{-1} gilt. Tatsächlich, für x\_1\in V\_a existiert eine Funktion \phi: \R^n \rightarrow \R^n derart, dass

f(x\_1) = f(x)+ L\_x(x\_1 - x) +\phi (x\_1-x) und

\lim\_{x\_1\rightarrow x} \frac{\|\phi (x\_1-x) \|}{\| x\_1- x\|} = 0.

Umgeformt lautet dies L^{-1}\_x (f(x\_1) - f(x)) = x\_1 - x + L^{-1}\_x (\phi (x\_1-x)) und es ergibt sich mit y\_1 = f(x\_1)

L^{-1}\_x (y\_1- y) = f^{-1}(y\_1)- f^{-1}(y) +L^{-1}\_x\big(\phi (f^{-1}(y\_1) - f^{-1} (y))\big).

Für die Differenzierbarkeit ist es also ausreichend,

\lim\_{y\_1\rightarrow y} \frac{\| L^{-1}\_x\big(\phi (f^{-1}(y\_1) - f^{-1} (y))\big)\|}{\|y\_1 - y\|} = 0

zu beweisen. Wendet man die Fundamentalabschätzung auf L\_x^{-1} mit der Konstanten A an, so ist

\|L^{-1}\_x\big(\phi (f^{-1}(y\_1) - f^{-1} (y))\big) \| \leq A\cdot \| \phi (f^{-1}(y\_1) - f^{-1}(y)) \|.

Nun gilt wegen der Stetigkeit von f^{-1}

\lim\_{y\_1\rightarrow y} \frac{\| \phi (f^{-1}(y\_1) - f^{-1}(y))\|}{\| y\_1-y\|}

= \lim\_{x\_1\rightarrow x} \frac{\| \phi ( x\_1 - x)\|}{\| f(x\_1) - f(x)\|} \\

= \lim\_{x\_1\rightarrow x} \underbrace{\frac{\| \phi ( x\_1 - x)\|}{\|x\_1 - x\|}}\_{\rightarrow 0} \cdot

\underbrace{\frac{\|x\_1 - x\|}{\| f(x\_1) - f(x)\|}}\_{\leq 2 \text{ nach }(\*)} = 0.

Demnach ist f^{-1} differenzierbar und Df^{-1}(f(x))=[Df(x)]^{-1}.

Sei nun f=(f\_1,\ldots, f\_n): V\_a\rightarrow V\_b, mit Umkehrfunktion f^{-1}=(u\_1,\ldots, u\_n): V\_b\rightarrow V\_a, also f\_i (u\_1(y),\ldots, u\_n(y))=y\_i. Mittels Differenzieren folgt

\sum\_{k=1}^n \frac{\partial f\_i}{\partial x\_k}(f^{-1} (y))\frac{\partial u\_k}{\partial y\_j}(y)

= \delta\_{ij}, also

\left[\frac{\partial u\_k}{\partial y\_j} (y)\right]

= \left[\frac{\partial f\_i}{\partial x\_k}(f^{-1} (y))\right]^{-1}.

Damit ist jedes \frac{\partial u\_k}{\partial y\_j} (y) eine rationale Funktion q\_{kj} in den partiellen Ableitungen \frac{\partial f\_i}{\partial x\_k}(f^{-1} (y)). Ist dann f der Klasse C^k, so ist es auch f^{-1}, weil f^{-1} stetig ist.

</beweis>

<definition> Definition 5.4.5

* Eine bijektive Abbildung f: U\rightarrow V der Klasse C^k, deren Umkehrfunktion auch der Ordnung C^k ist, heißt C^k-Diffeomorphismus.
* Ein Punkt x\in U heißt regulärer Punkt einer C^1-Funktion f: U\rightarrow V, falls Df(x) invertierbar ist.

</definition>

Der Satz über die Umkehrfunktion besagt also qualitativ folgendes: Ist f: U\rightarrow V der Klasse C^k und x ein regulärer Punkt von f, so existiert eine hinreichend kleine Umgebung um x, auf der f ein C^k-Diffeomorphismus ist.

## 5.5 Der Satz über implizite Funktionen und Extrema mit Nebenbedingungen

Oft ist man in der Situation, dass eine Funktion nur implizit durch eine Gleichung gegeben ist, die man gerne nach dieser auflösen bzw. freistellen würde. Obwohl eine explizite Formel meist unerreichbar ist, kann man dennoch die Bedingungen angeben, unter denen eine solche Auflösung existiert.

<beispiel> Beispiel 5.5.1

Für die Menge der Punkte (x,y)\in\R^2 mit x^2+y^2-1=0 suchen wir eine Funktion \phi (x): \R\rightarrow \R mit x^2+\phi ^2(x)-1=0.

Es gibt aber nicht eine solche Funktion, sondern zwei, nämlich \phi (x)=\pm \sqrt{1-x^2}. Zudem ist sie nicht überall definiert, sondern nur auf [0,1].

</beispiel>

<beispiel> Beispiel 5.5.2

Wir betrachten die Funktion f:\R^n\rightarrow \R, f(x\_1,\ldots, x\_n)=x\_1^2 + \ldots x\_{n-1}^2. Die Menge aller Punkte x, die f(x)=0 erfüllen, ist nicht leer, denn sie besteht aus der x\_n-Achse. Aber es gibt keine Funktion \phi : \R^{n-1}\rightarrow \R, welche die Identität f(x\_1,\ldots,x\_{n-1},\phi)=0 erfüllt, da es grundsätzlich nicht möglich ist, die letzte Koordinate als Funktion der anderen auszudrücken.

</beispiel>

<beispiel> Beispiel 5.5.3

Betrachte die Nullstellen der Gleichung f(x,y)=x-y^3. Zwar wird f(x,y)=0 durch \phi (x)=\sqrt[3]{x} aufgelöst, doch im Gegensatz zur ursprünglichen Funktion ist diese bei 0 nicht differenzierbar.

</beispiel>

<satz> Satz 5.5.4 Satz über implizite Funktionen

Sei U\subset \R^n\times \R^m offen, f: U\rightarrow \R^m stetig differenzierbar, (a,b) \in U \subset \R^n\times \R^m ein Punkt, in dem f(a,b)=0 und

\det \begin{bmatrix} \frac{\partial f\_1}{\partial y\_1} & \cdots & \frac{\partial f\_1}{\partial y\_m}\\

\vdots & \ddots & \vdots\\

\frac{\partial f\_m}{\partial y\_1} & \cdots & \frac{\partial f\_m}{\partial y\_m} \end{bmatrix} (a,b)

\neq 0

gilt. Dann existiert eine offene zusammenhängende Menge A\subset\R^n, die a enthält, sowie eine Abbildung g: A\rightarrow \R^m mit

1. g(a)=b, f(x, g(x))=0 \forall x\in A {\bf (Auflösung bzw. Freistellung nach y)},
2. g ist eindeutig,
3. ist f\in C^k(U), dann ist auch g\in C^k(A).

</satz>

<beweis>

Wir definieren eine Abbildung F: U\rightarrow \R^{n+m} durch F(x,y)=(x,f(x,y)). Diese erfüllt

\det DF(a,b) = \det \begin{bmatrix} 1 & & & & & \\

 & \ddots & & &0 & \\

& & 1 & & & \\

& & & \frac{\partial f\_1}{\partial y\_1} & \cdots & \frac{\partial f\_1}{\partial y\_m}\ & 0 & & \vdots & \ddots & \vdots\\

& & & \frac{\partial f\_m}{\partial y\_1} & \cdots & \frac{\partial f\_m}{\partial y\_m}\\ \end{bmatrix}(a,b) \neq 0.

Nach dem Satz über die Umkehrfunktion existiert eine offene Menge \tilde{W} \subset\R^{n+m}, die F(a,b)=(a,0) enthält sowie eine offene Menge \tilde{V}\subset \R^{n+m}, die (a,b) enthält derart, dass F\_{|\tilde{V}}: \tilde{V}\rightarrow \tilde{W} bijektiv und F^{-1} der gleichen Klasse wie F ist. Wähle nun eine offene Menge A\subset\R^n um a und ein \epsilon>0 derart, dass W:=A\times ]-\epsilon,\epsilon[^m\subset\tilde{W}. Dies ist möglich, weil (a,0)\in\tilde{W}. Sei V:= F^{-1}(W).

Durch weiteres Verkleinern von A erreichen wir zudem, dass eine Umgebung B um b existiert mit A\times B\subset \tilde V. Nach Konstruktion ist F\_{|V}: V\rightarrow W bijektiv und F^{-1} der gleichen Klasse wie f. Wir schreiben F^{-1}: W\rightarrow V als F^{-1}=(F\_1,F\_2) und überlegen uns, dass F\_2 fast die gesuchte Funktion g ist. So gilt

F F^{-1}(x,y)=(x,y), also \big( F\_1(x,y), f(F\_1(x,y), F\_2(x,y))\big) = (x,y).

Damit folgt \label{eq:8}

F\_1(x,y) = x \text{ und } f(x, F\_2(x,y)) =y \forall (x,y)\in W.

Weiterhin gilt F(a,b)=(a,0), also

(a,b)=F^{-1}(a,0)=(a, F\_2(a,0)). Demzufolge ist \label{eq:9}

F\_2(a,0) = b.

Setzt man g(x):=F\_2(x,0), so ist g damit auch auf a\in A\subset\R^n definiert und wegen \eqref{eq:8} und \eqref{eq:9} gilt g(a) = b, f(x, g(x)) = 0 \forall x \in A.

Wenn F^{-1} der Klasse C^k ist, so gilt dies auch für g. Es verbleibt noch, die Eindeutigkeit von g zu zeigen. Angenommen, \tilde{g}: A\rightarrow \R^m wäre eine zweite Funktion mit \tilde{g}(a)=b, f(x, \tilde{g}(x))=0. Dann gälte

F(x,g(x)) = (x, f(x,g(x))) = (x,0),

F(x,\tilde{g}(x)) = (x,0).

Weil F auf A\times B injektiv ist, muss dann aber g(x)=\tilde{g}(x) gelten.

</beweis>

<beispiel> Beispiel 5.5.5

Betrachte f: \R^2\rightarrow \R, f(x,y)=x^2+y^2-1 und die Gleichung 0=f(x,y). Es ist \frac{\partial f}{\partial y}=2y, die Ableitung also invertierbar in den Punkten (a,b)\neq (\pm 1, 0). In Punkten (a,b) mit b>0 ist y=\sqrt{1-x^2}, für b<0 dagegen y=-\sqrt{1-x^2}. In den Punkten (\pm1 ,0) ist der Satz nicht anwendbar.

</beispiel>

Angenommen, es gelten die Voraussetzungen des Satzes über implizite Funktionen, und es ist g=(g\_1,\ldots, g\_m): A\subset\R^n\rightarrow \R^m die gefundenene lokale Auflösung. Aus f(x,g(x))=0\in\R^m folgt dann durch Ableiten für jede Komponente f\_\alpha, \alpha=1,\ldots, m

\frac{\partial f\_\alpha}{\partial x\_i}(x,g(x))+\sum\_{j=1}^m \frac{\partial f\_\alpha}{\partial y\_j} (x, g(x))\frac{\partial g\_j(x)}{\partial x\_i} = 0,

also

\begin{bmatrix} \frac{\partial g\_j(x)}{\partial x\_i} \end{bmatrix}

= - \begin{bmatrix}\frac{\partial f\_\alpha}{\partial y\_j} (x, g(x))\end{bmatrix}^{-1}

\begin{bmatrix} \frac{\partial f\_\alpha}{\partial x\_i}(x,g(x))\end{bmatrix}.

Die Existenz der hier auftauchenden Inversmatrix ist dabei durch die Voraussetzung des Satzes über implizite Funktionen gewährleistet. Qualitativ besagt diese Formel, dass alle partiellen Ableitungen der unbekannten Funktion g als rationale Ausdrücke der Einträge der rechten Seite berechnet werden können -- worin allerdings g auftaucht. Man bekommt also ein Differentialgleichungssystem für die partiellen Ableitungen von g. Wir erläutern diesen Sachverhalt an einem Beispiel.

<beispiel> Beispiel 5.5.6

Betrachte die Funktion f(x,y)=x^2+2xy-y^2-a^2. Da \partial\_y f=2x-2y, ist der Satz über implizite Funktionen anwendbar in allen Punkten mit x\neq y. Nun ist f(x,y)=0 äquivalent zu (x+y)^2-2y^2=a^2, die Ausnahmepunkte entsprechen also Punkten mit 2x^2=a^2. Weiter lautet f(x,g(x))=0 explizit x^2+2xg(x)-g^2(x) = a^2 und Ableiten ergibt x+g+(x-g)g'=0, also

g' = \frac{-x-g}{x-g} = -\frac{(x+g)^2}{x^2-g^2} = -\frac{(x+g)^2}{a^2-2xg}.

</beispiel>

Das folgende Ergebnis ist eine Variante des Satzes über implizite Funktionen.

<satz> Satz 5.5.7

Sei U\subset \R^n offen, f: U\rightarrow \R^p eine Funktion der Klasse C^k und p\leq n. Sei a\in U und gelte f(a)=0, {\mathrm{rk} } Df(a)=p. Dann existiert eine Umgebung U\_a\subset U um a sowie ein C^k-Diffeomorphismus h\_a: U\_a\rightarrow h\_a(U\_a)\subset\R^n mit

(1) h\_a (a)=0,

(2) f\circ h^{-1}\_a (x\_1,\ldots, x\_n)=(x\_{n-p+1},\ldots, x\_n).

</satz>

<beweis>

Ist {\mathrm{rk} }Df(a)=p, so existieren p Spalten i\_1,\ldots, i\_p mit

\det \begin{bmatrix}

\frac{\partial f\_1}{\partial x\_{i\_1}}(a) & \cdots & \frac{\partial f\_1}{\partial x\_{i\_p}}(a)\\

\vdots & \ddots & \vdots\\

\frac{\partial f\_p}{\partial x\_{i\_1}}(a) & \cdots & \frac{\partial f\_p}{\partial x\_{i\_p}}(a)

\end{bmatrix} \neq 0.

Wir betrachten zunächst den Fall, dass es die letzten p Indizes sind, es sich also um die Ableitungen nach x\_{n-p+1},\ldots, x\_n handelt. Wir fassen f als Abbildung

f: \R^{n-p}\times \R^p\rightarrow \R^p auf und betrachten wie eben

F: \R^{n-p}\times \R^p\rightarrow \R^{n-p}\times \R^p,

F(x\_1,\ldots, x\_{n-p}, x\_{n-p+1},\ldots, x\_n) =(x\_1,\ldots, x\_{n-p}, f\_1(x), \ldots, f\_p(x)).

Es ist \det DF(a)\neq 0, F(a)=(a\_1,\ldots, a\_{n-p}, 0,\ldots,0). Für H(x)=F(x)- (a\_1,\ldots, a\_{n-p}, 0,\ldots,0) gilt also H: U\rightarrow \R^{n-p}\times \R^p, H(a)=0, \det dH(a)\neq 0. Nach dem Satz über die Umkehrfunktion existieren offene Umgebungen a\in U\_a\subset\R^n, 0\in V\_a\subset\R^n derart, dass h\_a:=H\_{|U\_a}: U\_a\rightarrow V\_a ein Diffeomorphismus ist.

Schreibt man alles aus, so sieht man, dass dies genau die gesuchte Funktion ist. Den allgemeinen Fall behandelt man dadurch, dass man eine lineare Abbildung L wählt, die die Koordinaten entsprechend umsortiert. Auch hier verzichten wir auf die Details.

</beweis>

Wir wenden uns nun der Extremwertbestimmung mit Nebenbedingungen zu und illustrieren die Fragestellung an zwei Beispielen.

<beispiel> Beispiel 5.5.8

Gesucht ist das Minimum der Funktion f(x)=x\_1+x\_2+\ldots+x\_n unter der Bedingung, dass x\_1 \cdot x\_2\cdot …\cdot x\_n=1 und x\_i\geq 0 gilt. Diese Aufgabe hängt mit der Ungleichung

\sqrt[n]{x\_1\cdots x\_n} \leq \frac{x\_1+\ldots + x\_n}{n}

zwischen geometrischem und arithmetischem Mittel zusammen. Denn aus x\_1\cdot x\_2 \cdot … \cdot x\_n = 1 folgt aus ihr \sum x\_i\geq n und Gleichheit tritt genau dann auf, falls alle x\_i gleich sind. Damit hat f bei (1,1,\ldots,1) ein Minimum.

</beispiel>

<beispiel> Beispiel 5.5.9

Wir betrachten ein n-Eck, dessen Ecken P\_1,\ldots, P\_n auf dem Einheitskreis liegen. Bezeichnet man die Innenwinkel mit x\_i, so ist die Fläche des n-Ecks gegeben durch f(x)=\frac{1}{2}\sum\_{i=1}^n \sin x\_i

\footnote{Die Fläche eines jeden Teildreiecks ist die Hälfte des Produkts von Basis und Höhe: dreht man eine Kante auf die positive x-Achse, so sieht man, dass die Basis 1 ist und die Höhe gerade 1\cdot \sin x\_i.}.

Wir suchen die Extrema von f unter den Nebenbedingungen x\_i\in [0,\pi], \sum x\_i=2\pi, also das eingeschriebene n-Eck mit größtem Flächeninhalt.

\includegraphics[width=5cm]{bilder/extrema-mit-NB.eps}

Abbildung, die ein in den Einheitskreis eingeschriebenes Fünfeck zeigt, das über den Kreismittelpunkt in Dreiecke unterteilt ist.

</beispiel>

Zur allgemeinen Behandlung des Problems seien x=(x\_1,\ldots, x\_n)\in\R^n, y=(y\_1,\ldots, y\_m)\in\R^m Koordinaten in \R^n bzw. \R^m und f: U\subset \R^{n+m}\rightarrow \R, g: U\rightarrow \R^m gewisse Funktionen. Es sind die Extrema von f(x,y) unter den m Nebenbedingungen g\_i(x,y)=0, i=1,\ldots,m gesucht. Die Grundidee ist nun, g nach y aufzulösen.

<satz> Satz 5.5.10

Sei U eine offene Umgebung von (a,b)\in\R^{n+m}, f: U\rightarrow \R und

g: U\rightarrow \R^m stetig differenzierbar und gelte

\det \begin{bmatrix} \frac{\partial g\_1}{\partial y\_1} & \cdots & \frac{\partial g\_1}{\partial y\_m}\\

\vdots & \ddots & \vdots\\

\frac{\partial g\_m}{\partial y\_1} & \cdots & \frac{\partial g\_m}{\partial y\_m}

\end{bmatrix} (a,b) \neq 0 .

Hat die Funktion f(x,y) unter der Nebenbedingung g(x,y)=0 an der Stelle (a,b) ein lokales Extremum, so gibt es ein \lambda\_0\in\R^m derart, dass die Funktion

H(x,y,\lambda) := f(x,y)+\langle \lambda , g(x,y)\rangle = f(x,y)+\sum\_{j=1}^m \lambda\_j g\_j(x,y)

an der Stelle (a,b,\lambda\_0) einen stationären Punkt hat.

</satz>

<beweis>

Nach dem Satz über implizite Funktionen hat g=0 in einer Umgebung von (a,b) eine eindeutige Auflösung y=h(x)=(h\_1(x),\ldots, h\_m(x)) mit h(a)=b und g(x,h(x))=0. Also hat k(x):= f(x, h(x)) an der Stelle a ein Extremum (ohne Nebenbedingungen), und es ist für alle i

\label{eq:10}

\frac{\partial k}{\partial x\_i} = \frac{\partial f}{\partial x\_i}+

\sum\_{j=1}^m \frac{\partial f}{\partial y\_j}\frac{\partial h\_j}{\partial x\_i} = 0.

Zudem gilt für das Differential von h

\begin{bmatrix}\frac{\partial h\_j(x)}{\partial x\_i}\end{bmatrix}

= -\begin{bmatrix}\frac{\partial g\_\alpha}{\partial y\_j} (x, h(x))\end{bmatrix}^{-1}

\begin{bmatrix}\frac{\partial g\_\alpha}{\partial x\_i}(x,h(x))\end{bmatrix}.

Eingesetzt in \eqref{eq:10} und als vektorielle Identität geschrieben bedeutet dies

\begin{bmatrix} \frac{\partial f}{\partial x\_i}\end{bmatrix}

- \begin{bmatrix} \frac{\partial f}{\partial y\_j} \end{bmatrix}^t

 \begin{bmatrix}\frac{\partial g\_\alpha}{\partial y\_j} (x, h(x))\end{bmatrix}^{-1}

\begin{bmatrix}\frac{\partial g\_\alpha}{\partial x\_i}(x,h(x))\end{bmatrix} = 0.

Setzt man nun

\lambda\_0 = (\lambda\_0^1,\ldots, \lambda^m\_0 )

= -\begin{bmatrix} \frac{\partial f}{\partial y\_j}\end{bmatrix}^t

\begin{bmatrix}\frac{\partial g\_\alpha}{\partial y\_j} (a,b)\end{bmatrix}^{-1},

 so gilt an der Stelle (a,b,\lambda\_0)

\begin{bmatrix} \frac{\partial f}{\partial x\_i}\end{bmatrix}

+\lambda\_0 \cdot \begin{bmatrix}\frac{\partial g\_\alpha}{\partial x\_i}(a,b)\end{bmatrix}\ = 0

und

\begin{bmatrix} \frac{\partial f}{\partial y\_j}\end{bmatrix}

+\lambda\_0 \cdot \begin{bmatrix}\frac{\partial g\_\alpha}{\partial y\_j} (a,b)\end{bmatrix} = 0.

Dies sind die partiellen Ableitungen von H nach den Variablen x und y. Weil zudem \partial H/ \partial \lambda\_k = g\_k=0 ist, folgt, dass H dort einen stationären Punkt hat und damit die Behauptung.

</beweis>

Man nennt \lambda\_1,\ldots, \lambda\_m die Lagrange'schen Multiplikatoren. An der Stelle (a,b,\lambda\_0) sind nach obigem die Gleichungen

\frac{\partial H}{\partial x\_i}& =& \frac{\partial f}{\partial x\_i}

+\sum\_{j=1}^m \lambda\_j \frac{\partial g\_j}{\partial x\_i} = 0, i=1,\ldots,n,\\

\frac{\partial H}{\partial y\_k}& =& \frac{\partial f}{\partial y\_k}+

\sum\_{j=1}^m \lambda\_j \frac{\partial g\_j}{\partial y\_k} = 0, k=1,\ldots,m,\\

\frac{\partial H}{\partial \lambda\_k}& =& g\_k = 0, k=1,\ldots,m,

notwendige Bedingungen für das Auftreten eines lokalen Extremums von f in (a,b) unter den Nebenbediungen g=0, sofern die Matrix [\partial g/\partial y] dort invertierbar ist. In der Praxis setzt man die Funktion H an mittels Hinzuzählen der Terme, die verschwinden sollen, und bestimmt aus den partiellen Ableitungen nach x notwendige Bedingungen für ein Extremum, die man dann weiter studiert.

Wir illustrieren die Anwendung der Lagrange'schen Multiplikatoren an den eingangs gestellten Aufgaben:

Fortsetzung von Beispiel 5.5.8

Wir suchen das Minimum von f(x)=x\_1+x\_2+\ldots+x\_n unter der Bedingung g(x)= x\_1\cdot\ldots\cdot x\_n-1=0. Demnach ist

H(x,\lambda) = x\_1+x\_2+\ldots+x\_n+\lambda (x\_1\cdot\ldots\cdot x\_n-1)

mit partiellen Ableitungen

\frac{\partial H}{\partial x\_i} = 1+ \lambda\prod\_{j=1, j\neq i}^n x\_j = 1+\frac{\lambda}{x\_i} = 0.

Also ist x\_i=-1/\lambda eine notwendige Bedingung für ein Extremum. Da alle x\_i gleich sind, folgt x\_i=1.

Fortsetzung von Beispiel 5.5.9

Für das Problem des größten n-Ecks im Einheitskreis ist H(x,\lambda) = \sum\_{i=1}^n \sin x\_i+\lambda\left[-2\pi+ \sum\_{i=1}^n x\_i \right]

mit partiellen Ableitungen

\frac{\partial H}{\partial x\_k} = \cos x\_k+\lambda = 0,

also \cos x\_k=const. Es folgt x\_k= 2\pi/n, das regelmäßige n-Eck ist also der Kandidat für den maximalen Flächeninhalt. Man überlegt sich leicht, dass ein echtes Maximum vorliegt.

## 5.6 Der Transformationssatz für Integrale

Wir kommen nun zur mehrdimensionalen Variante der Substitutionsformel, deren Beweis jedoch deutlich schwieriger ist.

<satz> Satz 5.6.1 Transformationsformel

Seien U,V\in\R^n offene Mengen, g: U\rightarrow V bijektiv, stetig differenzierbar und \det Dg (x)\neq 0 \forall x\in U. Ist zudem A\in\mathcal{J}(U) Jordan-messbar und f: g(A)\rightarrow \R Riemann-integrierbar, so gilt

\int\_{g(A)} f(x) dx = \int\_A f(g(y))|\det Dg(y)| dy.

</satz>

<beweis>

1. Schritt: Wir überlegen uns, dass es ausreicht,

\label{eq:11}

\int\_{g(A)} f(x) dx \leq \int\_A f(g(y))|\det Dg(y)| dy

zu beweisen. Wenn dies nämlich gilt, so kann man anstatt A\subset U, g: U\rightarrow V und f: g(A)\rightarrow \R die Objekte g(A)\subset U, g^{-1}: V\rightarrow U und

\tilde{f}=(f\circ g) |\det Dg|: A=g^{-1}g(A)\rightarrow \R

betrachten und hat für diese

\int\_{g^{-1}g(A)} \tilde{f}(y) dy \leq \int\_{g(A)} \tilde{f}(g^{-1}(x)) |\det D g^{-1}(x)| dx.

Dann ist aber

\int\_A f(g(y))|\det Dg(y)| dy \leq \int\_{g(A)} f(x) |\det Dg(x)|\cdot |\det D g^{-1}(x)| dx

= \int\_{g(A)} f(x) dx .

2. Schritt: Wir zeigen als nächstes, dass es sogar ausreicht, \eqref{eq:11} für die konstante Funktion f=c zu beweisen. Denn sei f beliebig und P eine Zerlegung eines Quaders \tilde{W} um g(A). Desweiteren gelte \eqref{eq:11} für die konstante Funktion. Dann ist

\underline{S}(\chi\_{g(A)} f, W;P)

= \sum\_{W\in P} \inf \chi\_{g(A)} f\big|\_W \text{vol} W

= \sum\_{W\in P} \int\_W \inf \chi\_{g(A)} f\big|\_W

\leq \sum\_{W\in P} \int\_{g^{-1}(W)} \inf [\chi\_{g(A)} f\big|\_W] (g(x)) |\det Dg(x)| dx

\leq \sum\_{W\in P} \int\_{g^{-1}(W)}\chi\_{g(A)}(g(x))f(g(x))|\det Dg(x)| dx

= \sum\_{W\in P} \int\_{g^{-1}(W)}\chi\_{g(A)}(x)f(g(x))|\det Dg(x)| dx.

 Weil g nach Voraussetzung bijektiv ist, kann dies durch das Integral über ganz U abgeschätzt werden, so dass

\underline{S}(\chi\_{g(A)} f, W;P) \leq\ \int\_U\chi\_{g(A)}(x)f(g(x)) |\det Dg(x)| dx

= \int\_A f(g(x))|\det Dg(x)| dx.

Aus den ersten beiden Schritte ergibt sich somit, dass es genügt, \label{eq:12}

\text{vol} g(A) \leq \int\_A |\det Dg(y)| dy

zu beweisen.

3. Schritt: Wir zeigen, dass es ausreicht, \eqref{eq:12} für einen Quader A=W zu wissen. Trifft dies zu, so betrachten wir für ein beliebiges A\in {\mathcal{J}}(U) endlich viele disjunkte Quader W\_i \in Int A, i=1,\ldots, k. Für diese ist

\text{vol} g\big(\cup\_i W\_i\big) \leq \int\_{\cup\_i W\_i} |\det Dg(x)| dx \leq \int\_A |\det Dg(x)| dx,

weil diese Ungleichung nach Voraussetzung ja für jeden einzelnen Quader W\_i gilt. Dann ist aber auch

\sup\_{W\_i, W\_i\in\text{Int} A} \text{vol} g\big(\cup\_i W\_i\big) \leq\ \int\_A |\det Dg(x)| dx,

und die linke Seite ist nach Regularität des Jordan-Volumens genau \text{vol} g(A). Wir müssen also \eqref{eq:12} nur für Quader in U beweisen.

4. Schritt: Wir zeigen \eqref{eq:12} zunächst für Quader und eine lineare Funktion g:\R^n\rightarrow \R^n. In diesem Fall gilt sogar

\text{vol} g(W) = |\det g|\text{vol} W = \int\_W |\det Dg(y)| dy.

Um dies einzusehen, verwendet man einen Fakt der linearen Algebra, demzufolge jede lineare Abbildung sich als Verknüpfung von Elementartransformationen t\_i schreiben läßt. Diese Elementartransformationen bestehen aus der Multiplikation eines einzelnen Vektors mit einer Konstanten, der Addition zweier Vektoren sowie der Vertauschung zweier Vektoren. Für diese prüft man die Aussage leicht nach. Dann aber gilt für eine Verknüpfung g=t\_1\dots t\_k

\text{vol} g(W) = \text{vol} t\_1 t\_2\ldots t\_k (W) = \det t\_1 \text{vol} t\_2\ldots t\_k (W)

= \ldots = \det t\_1 \det t\_2\ldots\det t\_k \text{vol} W.

5. Schritt: Wir zeigen nun allgemein \eqref{eq:12}. Sei \delta>0 gegeben. Wir überdecken den Quader W\subset U mit einer gewissen Anzahl k\_\delta von Quadern K\_i, i=1,\ldots, k\_\delta, der Kantenlänge \delta, so dass

W \subset \bigcup\_{i=1}^{k\_\delta} K\_i\subset U.

Für x,y\in K\_i gilt nach der Taylorformel für alle \alpha=1,\ldots,n und gewisse t\_\alpha\in K\_i

g\_\alpha(x) - g\_\alpha(y) = \sum\_{j=1}^n \frac{\partial g\_\alpha}{\partial x\_j} (t\_\alpha)(x\_j-y\_j),

also

|g\_\alpha(x) - g\_\alpha(y)| \leq \delta \| Dg(t\_\alpha)\| \leq \delta \sup\_{t\in K\_i}\| Dg(t)\| .

Dabei bezeichnet \| \cdot \| die Maximumsnorm, so dass \| (a\_{ij})\|=\max\sum |a\_{ij}|. Insgesamt liegt g(K\_i) in einem Quader der Kantenlänge \delta \sup\_{t\in K\_i}\| Dg(t)\|^n, was

\text{vol} g(K\_i)\leq \delta^n \sup \| Dg(t)\|^n

nach sich zieht. Sei m\_i\in K\_i der Mittelpunkt des Quaders K\_i. Dann ist L=[Dg(m\_i)]^{-1} eine lineare Abbildung und wir betrachten die Verknüpfung h=L\circ g. Nach dem 4. Schritt und der Abschätzung angewandt auf h statt g gilt

\text{vol} g(K\_i)|\det L| = \text{vol} h(K\_i) = \leq \delta^n \sup\_{t\in K\_i}\| Dh(t)\|^n

= \delta^n \sup\_{t\in K\_i}\| DL\circ Dg(t)\|^n .

Aber auf der linken Seite ist |\det L|=|\det [Dg(m\_i)]^{-1} |, so dass man nach Multiplikation der Gleichung mit |\det L|:

\text{vol} g(K\_i) \leq |\det Dg(m\_i)| \delta^n \sup\_{t\in K\_i}\| DL\circ Dg(t)\|^n

= |\det Dg(m\_i)|\text{vol} K\_i \sup\_{t\in K\_i}\| Dg(m\_i)^{-1}\circ Dg(t)\|^n

erhält. Damit folgt

\text{vol} g(W) \leq \sum\_{i=1}^{k\_\delta} \text{vol} g(K\_i)

\leq \sum |\det Dg(m\_i)| \text{vol} K\_i \sup\_{t\in K\_i}\| Dg(m\_i)^{-1}\circ Dg(t)\|^n

\leq \max\_{i=1,\ldots, k\_\delta} \sup\_{t\in K\_i} \|Dg(m\_i)^{-1}\circ Dg(t)\|^n

\sum\_{i=1}^{k\_\delta} |\det Dg(m\_i)| \text{vol} K\_i.

Da m\_i\in K\_i, gilt dann für \delta\rightarrow 0 aufgrund der Definition des Riemann-integrals

\sum\_{i=1}^{k\_\delta} |\det Dg(m\_i)| \text{vol} K\_i \longrightarrow \int\_W|\det Dg(x)| dx.

Damit verbleibt,

\lim\_{\delta\rightarrow 0} \max\_{i=1,\ldots, k\_\delta} \sup\_{t\in K\_i}\| Dg(m\_i)^{-1}\circ Dg(t)\|^n

zu zeigen. Sei dazu 0<M=: \min\_{x\in W} \|Dg(x)\|. Weil W kompakt ist, ist die stetige Funktion x\mapsto \|Dg(x)\| sogar gleichmäßig stetig. Zu gegebenem \epsilon>0 existiert also ein \delta>0 mit

\|Dg(x) - Dg(y)\| < \epsilon \forall x,y\in K\_i.

Dann aber ist

\|Dg(x) Dg(y)^{-1} -{\mathrm{Id}}\| < \frac{\epsilon}{\|Dg(y)\|} \leq \frac{\epsilon}{M},

also

\max\_{i=1,\ldots, k\_\delta} |\sup\_{t\in K\_i}\| Dg(m\_i)^{-1}\circ Dg(t)\|^n -1|

\leq \frac{\epsilon^n}{M^n}

für hinreichend kleines \delta>0.

</beweis>

Das wichtige an dieser Formel sind die Anwendungen.

<beispiel> Beispiel 5.6.2 Polarkoordinaten

Sei f: D^2(R)\rightarrow \R eine Riemann-integrierbare Funktion.

Wir betrachten die Abbildung

g: (0,R)\times (0, 2\pi) \rightarrow K (0,R)-\{(x,0): x\in[ 0,R[\}, g(r,\theta)=(r\cos\theta,r\sin\theta),

die der Transformation von Polarkoordinaten auf kartesische Koordinaten entspricht. Sie ist ein Diffeomorphismus mit

Dg(r,\theta)

= \begin{bmatrix} \cos\theta & \sin\theta\

-r\sin\theta & r\cos\theta\end{bmatrix}

Und |\det Dg| = r.

Bei der Integration über D^2(R) spielen die weggelassenen Randpunkte keine Rolle, denn sie bilden eine Jordan-Nullmenge und liefern keinen Beitrag zum Integral. Damit gilt explizit

\int\_{D^2(R)}f(x,y) dxdy = \int\_0^R\int\_0^{2\pi} f(r\cos\theta,r\sin\theta) r drd\theta.

Zum Beispiel erhält man so ganz leicht wieder das Volumen der zweidimensionalen Kugel, denn nach Fubini ist

\text{vol} D^2(R) = \int\_{D^2(R)}1 dxdy

= \int\_0^Rr dr \int\_0^{2\pi} d\theta = 2\pi\cdot \frac{1}{2}r^2\bigg|^R\_0

 =\pi R^2.

In diesem Zusammenhang sei noch die Leibniz'sche Sektorformel erwähnt. Wir betrachten einen in Polarkoordinaten gegebenen Graphen einer Funktion r=f(\theta), \theta\in[\alpha,\beta]. Der von dem Graphen und den beiden Geraden \theta=\alpha, \theta=\beta eingeschlossene Sektor

S = \{ (r\cos\theta,r\sin\theta)\in\R^2: \alpha\leq\theta\leq\beta, 0\leq r\leq f(\theta)\}

hat dann nach dem Transformationssatz den Flächeninhalt

\text{vol} S = \int\_\alpha^\beta \int\_0^{f(\theta)} r dr d\theta

= \frac{1}{2}\int\_\alpha^\beta f^2(\theta) d\theta.

</beispiel>

# Kapitel 6: Kurven im dreidimensionalen Raum

## 6.1 Grundbegriffe der Kurventheorie

Im 19. Jahrhundert ging man lange davon aus, dass unter einer Kurve im \R^3 das Bild einer stetigen Abbildung f: [0,1] \to A \subset \R^3 verstanden werden kann. Injektivität zu fordern, würde Kurven mit Selbstschnitten ausschließen. Wie jedoch Peano zeigte, gilt folgender

<satz> Satz 6.1.1 [Peano 1890]

Es existiert eine stetige, surjektive Abbildung des Intervalls auf das Quadrat.

</satz>

<beweis>

Die in Frage kommende Abbildung f: [0,1] \to [0,1] \times [0,1] wird konstruiert als Grenzwert einer gleichmäßig konvergenten Folge stetiger Abbildungen f\_i. Im ersten Schritt wird das Intervall [0,1] in neun Teilintervalle gleicher Länge zerlegt und eine analoge Zerlegung des Quadrates vorgenommen. Die Abbildung f\_1 bildet nun jedes der Teilintervalle [(i-1)/9, i/9] auf die Diagonale des i-ten Quaders stetig und bijektiv ab, wobei die Diagonalen jeweils wie angedeutet miteinander verbunden werden.

Die Abbildung f\_2 entsteht auf die gleiche Weise, wofür jetzt die Unterteilung in jedem der Intervalle [(i-1)/9, i/9] und in jedem der Teilquadrate nochmals durchgeführt wird. Man erhält 81 Intervalle und Quadrate und f\_2 bildet die kleinen Intervalle stetig und bijektiv auf die Diagonalen der kleinen Quadrate so ab, dass f\_2 ([(i-1)/9, i/9]) im i-ten Quadrat der ersten Zerlegungsstufe liegt. Nach Konstruktion gilt dann f\_1(i/9) = f\_2(i/9).

Auf diese Weise entsteht eine Folge von stetigen Abbildungen mit

* \| f\_{n} (t) - f\_{n+1} (t)\| \le \sqrt{2}/ 3^{n} für alle t \in [0,1];
* f\_n (i / 3^{n+1})=f\_{n+1} (i/3^{n+1})= f\_{n+2} (i/3^{n+1}) für alle n \ge 1 und 0 \le i \le 3^{n+1}.

Beachte hierbei, dass \sqrt{2}/3^n die Länge der Diagonalen in der n-ten Zerlegung ist. Aufgrund der Abschätzung (1) gilt

| f\_n (t) -f\_m (t)\| \leq \sqrt{2} (3^{-n}+ 3^{-n-1}+ \dots )

\leq \frac{\sqrt{2}}{3^n} \sum \_{k=0}^\infty 3^{-k} = \frac{1}{\sqrt{2}\cdot 3^{n-1}}

für alle t \in [0,1] und m \ge n. Damit bilden die \mklm{f\_n} eine Cauchy-Folge und konvergieren gleichmäßig gegen eine stetige Abbildung f:[0,1] \to [0,1] \times [0,1]. Nach Konstruktion liegen jedoch alle Punkte der Gestalt (\frac{k}{3^n}, \frac{l}{3^m}) im Bild der Abbildung f. Das Intervall [0,1] ist kompakt und damit ist die Bildmenge f([0,1]) gleichfalls eine kompakte Teilmenge des Quadrates. Die beschriebene, auf jeden Fall im Bild von f liegende Menge ist dicht im Quadrat [0,1] \times [0,1] und somit folgt f([0,1])=[0,1] \times [0,1], d.h. f ist stetig und surjektiv.

</beweis>

Dieses Beispiel zeigt, dass die "Dimension"' eines topologischen Raumes unter stetigen Abbildungen wachsen kann. Eine Präzisierung dieser Größe erfolgt in der topologischen Dimensionstheorie.

<definition> Definition 6.1.2

Eine parametrisierte Kurve ist eine C^1-Abbildung \gamma : [a, b] \rightarrow \R^3 eines Intervalls nach \R^3, deren Ableitung nirgends verschwindet.

</definition>

Um die Länge einer Kurve zu definieren, wählen wir auf der Kurve endlich viele Punkte P\_0 = \gamma(t\_0) , P\_1 = \gamma(t\_1), P\_2 = \gamma(t\_2), \ldots, P\_k = \gamma(t\_k) , welche einer wachsenden Parameterfolge a = t\_0 < t\_1 < t\_2 < \ldots < t\_k = b entsprechen. Ersetzen wir die Kurve durch den durch diese Punkte verlaufenden Polygonzug \mathcal{P} und messen wir dessen Länge, so erhalten wir

L (\mathcal{P}) = \sum\_{i=1}^k |P\_i - P\_{i-1}| =

\sum\_{i=1}^k

\Big| \frac{\gamma(t\_i) - \gamma(t\_{i-1})}{t\_i - t\_{i-1}}\Big| \cdot \big( t\_i - t\_{i-1} \big).

Die Länge L(\gamma) der Kurve wird als das Supremum der Längen aller dieser die Kurve approximierenden Polygonzüge definiert,

L(\gamma) := \mathrm{sup} \big( L (\mathcal{P}) \big) .

Sie muss für differenzierbare Kurven nicht notwendigerweise endlich sein.

<proposition> Proposition 6.1.3

Ist \gamma:[a,b]\rightarrow \R^3 der Klasse C^1, so gilt L(\gamma)<\infty und

\label{eq:kurvenlaenge}

L(\gamma) = \int\_a^b \| \gamma'(t)\| dt .

</proposition>

<beweis>

Da die stetige Funktion \|\gamma'(t)\| auf [a,b] ihr Maximum und Minimum annimmt, existiert ein \alpha>0 mit \|\gamma'(t)\|\leq \alpha. Aufgrund der Mittelwertungleichung folgt

\|\gamma(t)- \gamma(t')\| \leq \alpha |t-t'| \text{ für alle } t,t'\in [a,b]

und wir erhalten sofort für eine beliebige Zerlegung die Abschätzung

L(\mathcal{P}) \leq \sum\_{i=1}^k \alpha \big( t\_i - t\_{i-1} \big) = \alpha (b-a) < \infty.

Damit hat \gamma endliche Länge. Die Gleichung \eqref{eq:kurvenlaenge} folgt aus den Mittelwertsätzen der Differentialrechnung. Insbesondere liefert das Kriterium für Riemann-Integrierbarkeit, dass das Integral existiert.

</beweis>

Der Vektor \gamma'(t) wird als der Tangential- oder Geschwindigkeitsvektor von \gamma bezeichnet. Wegen \gamma'(t)\neq 0 ist bei festem Anfangspunkt die Längenfunktion

L(t) := \int\_{a}^{t} \| \gamma'(x)\| dx

eine streng monoton wachsende Funktion und insbesondere injektiv. Ein Kurvenpunkt wird eindeutig durch denjenigen Parameter festgelegt, der die Länge des Kurvenstückes vom Anfangspunkt zu diesem betrachteten Punkt angibt. Mit anderen Worten, invertieren wir die Längenfunktion, so wird

\tilde \gamma =\gamma \circ L^{-1} : [0 , L(\gamma)] \rightarrow \R^3

eine Parametrisierung der Kurve. Der Ableitungsvektor bezogen auf diesen Längenparameter s hat die Länge eins, denn nach dem Satz über die Ableitung der Umkehrabbildung und Satz 4.5.1 ist

 \|d\tilde \gamma /ds\| =\norm{\gamma'(L^{-1}(s)) \cdot \frac d{ds} L^{-1}(s)}

= \norm{ \frac{\gamma'(L^{-1}(s))}{\norm{\gamma'(L^{-1}(s))}}}=1.

Eine derartige Parametrisierung nennt man die natürliche Parametrisierung der Kurve. Man sagt auch, die Kurve sei auf Bogenlänge parametrisiert. Wir treffen folgende Vereinbarung:

* Ist die Kurve in einer beliebigen Parametrisierung \gamma(t) gegeben, so bezeichnen wir mit \dot{\gamma}(t) etc. die Ableitung nach dem Parameter t;
* liegt jedoch die Kurve in ihrer natürlichen Parametrisierung vor, so sei \gamma'(s) etc. die Ableitung nach dem Längenparameter der Kurve.

<beispiel> Beispiel 6.1.4

Wir betrachten die Schraubenlinie in der Parametrisierung \gamma(t) = \big(a\cdot \cos t, a\cdot \sin t , b\cdot t\big) t \in [0, \infty).

Wegen ||\dot{\gamma}(t)|| = \sqrt{a^2 + b^2} ist die Längenfunktion gegeben durch L(t) = \sqrt{a^2 + b^2}\cdot t. Der natürliche Parameter ist s = L(t) = \sqrt{a^2 + b^2} \cdot t und daraus entsteht die natürliche Parametrisierung, indem man den Parameter t durch den Parameter s ersetzt:

\gamma(s) = \left(

a \cdot\cos \frac{s}{\sqrt{a^2+b^2}},

a \cdot \sin \frac{s}{\sqrt{a^2+b^2}},

\frac{b\cdot s}{\sqrt{a^2+b^2}}

\right).

Man sieht, dass der Geschwindigkeitsvektor dieser Kurve in der Tat immer Länge 1 hat.

</beispiel>

<beispiel> Beispiel 6.1.5

Die Ellipse \frac{x^2}{a^2}+\frac{y^2}{b^2}=1, a,b>0 kann parametrisiert werden durch

\gamma: [0,2\pi]\rightarrow \R^2, \gamma(t) = (a\cos t, b\sin t)

und hat damit die Längenfunktion

L(t) = \int\_0^t \sqrt{a^2\sin^2 x+b^2\cos^2 x} dx.

Dieses Integral kann nicht geschlossen ausgewertet werden, es gehört zur Gattung der elliptischen Integrale, so dass es nicht möglich ist, die Ellipse explizit auf Bogenlänge zu parametrisieren.

 </beispiel>

Zur Beschreibung der Geometrie einer Kurve ordnen wir ihr in jedem Punkt drei Vektoren zu.

<definition> Definition 6.1.6

Ist \gamma(s) eine Kurve in natürlicher Parametrisierung, so nennt man den Vektor \vec{t}(s) := \gamma'(s) den Einheitstangentialvektor an die Kurve.

</definition>

Die Krümmung einer Kurve in einem Punkt ist ein Maß für die Winkeländerung der Tangentialvektoren pro Längeneinheit. Dies führt zu folgender

<definition> Definition 6.1.7

Sei \mathcal{C} \subset \R^3 eine Kurve und p \in \mathcal{C} ein Kurvenpunkt. Die Krümmung \kappa(p) der Kurve \mathcal{C} im Punkte p ist der Grenzwert

\kappa(p) = \lim\_{q\rightarrow p}\frac{\phi}{\widehat{pq}} ,

wobei \phi der Winkel zwischen den Tangentialvektoren in den Punkten p,q und \widehat{pq} die Länge des Kurvenstückes zwischen den Punkten p und q ist.

</definition>

<satz> Satz 6.1.8

In der natürlichen Parametrisierung der Kurve gilt \kappa(s) = ||\gamma''(s)||.

</satz>

<beweis>

Setzen wir p = \gamma (s) und q = \gamma (s+h), so gilt \widehat{pq} = h in der natürlichen Parametrisierung der Kurve und der Winkel \phi berechnet sich unter Verwendung der Formel für den Winkel in einem gleichschenkligen Dreieck durch

\sin(\phi/2) = ||\vec{t}(s+h) - \vec{t}(s)||/2 .

Damit erhalten wir mit l'Hopital

\lim\_{q\rightarrow p}\frac{\phi}{\widehat{pq}} =

\lim\_{h\rightarrow 0}\frac{\phi/2} {\sin(\phi/2)}

\cdot \lim\_{h\rightarrow 0}\frac{2\cdot \sin(\phi/2)}{h} =

\lim\_{h\rightarrow 0} \frac{|| \vec{t}(s+h) - \vec{t}(s)||}{h} = ||\gamma''(s)|| .

</beweis>

In einer beliebigen Parametrisierung der Kurve zeigt eine leichte Umrechnung, dass (Übungsaufgabe!)

\kappa(t) = \frac{||\dot{\gamma}(t) \times \ddot{\gamma}(t)||}{||\dot{\gamma}(t)||^3} .

<definition> Definition 6.1.9

Der Hauptnormalenvektor \vec{h} einer Kurve in Punkten mit nicht-verschwindender Krümmung ist die normalisierte Ableitung des Einheitstangentialvektors,

\vec{h} (s):= \frac{1}{\kappa (s)} {\vec{t^{'}}} (s)= \frac{{\gamma}^{''}(s)}{\| {\gamma}^{''} (s)\|} .

Das Vektorprodukt \vec{b} (s):= \vec{t} (s) \times \vec{h} (s) nennt man den Binormalenvektor der Kurve. Die drei wegen \norm {\vec{t}(s)}^2=1 paarweise orthogonalen Vektoren \vec{t} (s), \vec{h} (s) und \vec{b} (s) der Länge eins bilden das Frenet-Reper der Kurve.

Die von dem Tangentialvektor \vec{t} und dem Hauptnormalenvektor \vec{h} aufgespannte Ebene nennt man die Schmiegebene der Kurve in einem Punkt. Der Binormalenvektor \vec{b} ist zu dieser Ebene orthogonal.

</definition>

<definition> Definition 6.1.10

Die Windung oder Torsion einer Kurve \mathcal{C} \subset \R^3 der Klasse C^3 in Punkten nicht verschwindender Krümmung ist das Skalarprodukt

\tau (s) = - \Big\langle \frac{d \vec{b} (s)}{ds}, \vec{h} (s) \Big\rangle .

</definition>

Die Strukturgleichungen einer Kurve im euklidischen Raum sind die Frenet-Formeln, welche die Ableitungen des Frenet-Repers durch eben dieses Reper ausdrücken. Sie sind der Inhalt des folgenden Satzes.

<satz> Satz 6.1.11 Fundamentalsatz der Kurventheorie

In der natürlichen Parametrisierung einer Kurve der Klasse C^3 nicht verschwindender Krümmung gilt

\frac{d}{ds} \begin{bmatrix} \vec{t} (s) \\ \vec{h} (s) \\ \vec{b} (s) \end{bmatrix}

 = \begin{bmatrix}

0 & \kappa (s) & 0 \\ - \kappa (s) & 0 & \tau (s) \\ 0 & - \tau (s) & 0 \end{bmatrix}

\cdot \begin{bmatrix} \vec{t} (s) \\ \vec{h} (s) \\ \vec{b} (s) \end{bmatrix} .

Sind \gamma, \gamma^\* :[0,L] \to \R^3 zwei Kurven gleicher Krümmung und gleicher Windung in der natürlichen Parametrisierung, so existiert eine euklidische Bewegung A: \R^3 \to \R^3 mit \gamma^\* (s) = A \circ \gamma (s). Sind eine positive Funktion \kappa (s) und eine beliebige Funktion \tau (s) auf einem Intervall [0,L] gegeben, dann existiert eine Kurve, deren Krümmung und Windung gleich den vorgegebenen Funktionen sind.

</satz>

<beweis>

Nach Definition des Hauptnormalenvektors gilt \vec{t}' (s) = \kappa (s)\cdot \vec{h} (s).

Wegen \| \vec{h} (s) \| \equiv 1 ist die Ableitung \vec{h}' (s) orthogonal zu \vec{h} (s), kann also als Linearkombination von \vec{t} (s) und \vec{b} (s) dargestellt werden. Aus

\Big\langle \frac{d \vec{h} (s)}{ds}, \vec{t} (s) \Big\rangle

= \frac{d}{ds} \Big\langle \vec{h} (s), \vec{t} (s) \Big \rangle

- \Big\langle \vec{h} (s) , \frac{d \vec{t} (s)}{ds} \Big\rangle

= 0- \Big\langle \vec{h}(s), \kappa (s)\cdot \vec{h} (s) \Big\rangle

= - \kappa(s)

sowie der entsprechenden Gleichung

\Big\langle \frac{d \vec{h} (s)}{ds}, \vec{b} (s) \Big\rangle =

\frac{d}{ds} \Big\langle \vec{h} (s), \vec{b} (s) \Big\rangle

- \Big\langle \vec{h} (s) , \frac{d \vec{b} (s)}{ds} \Big\rangle

= \tau(s)

ergibt sich die zweite der Frenet-Formeln. Dann aber gilt

\frac{d \vec{b} (s)}{ds}

= \frac{d \vec{t} (s)}{ds} \times \vec{h} (s)+\vec{t} \times \frac{d \vec{h} (s)}{ds}

= - \tau (s)\cdot \vec{h} (s).

Den zweiten Teil des Satzes wird in der Theorie der gewöhnlichen Differentialgleichungen bewiesen.

</beweis>

<bemerkung> Bemerkung 6.1.12

1. Der obige Satz besagt, dass eine Kurve bis auf euklidische Bewegungen eindeutig durch ihre Krümmung und ihre Torsion bestimmt ist.
2. Die dritte Frenet-Formel ergibt |\tau (s)|=\| \vec{b}' (s)\|. Daher kann der Absolutbetrag

| \tau (p)| der Windung einer Kurve in einem Punkt p als der Grenzwert

| \tau (p)| = \lim\limits\_{q \to p} \frac{\psi}{\widehat{pq}}

erhalten werden, wobei \psi der Winkel zwischen den Schmiegebenen der

Kurve in den Punkten p,q ist (siehe Beweis von Satz 6.1.8). Damit ist der Absolutbetrag der Windung ein Maß dafür, wie stark sich die Kurve aus ihrer Schmiegebene herauswindet.

</bemerkung>

# Kapitel 7: Einführung in die Theorie der gewöhnlichen Differentialgleichungen

## 7.1 Beispiele für gewöhnliche Differentialgleichungen

Die Theorie der Differentialgleichungen ist für die Betrachtung naturwissenschaftlicher Vor\-gänge von zentraler Bedeutung. Sie beschreibt, wie sich Systeme in der Zeit verändern und mit ihr sind einige der tiefsten Probleme der Mathematik verbunden.

Die Aufgabenstellung der Theorie der gewöhnlichen Differentialgleichungen läßt sich wie folgt umschreiben. Gegeben seien

* eine Abbildung F: \R^n\times \ldots \R^n\times \R = \R^{kn+1} \longrightarrow\R^n,
* eine Zahl t\_0\in\R,
* k Punkte a\_0,\ldots, a\_{k-1}\in\R^n.

Gesucht ist dann eine Abbildung x:I\rightarrow \R^n der Klasse C^k auf einer Umgebung

I := ]t\_0-\epsilon,t\_0+\epsilon[ von t\_0 derart, dass

* x^{(k)}(t) = F(x(t),\dot x(t), \ddot x (t),\ldots, x^{(k-1)}(t),t) für alle t\in I, (DGL k-ter Ordnung)
* x(t\_0)=a\_0, \dot x (t\_0)=a\_1,\ldots,x^{(k-1)}(t\_0)=a\_{k-1}. (Anfangsbedingungen)

Dabei heißt die Zahl k\in \N die Ordnung der Differentialgleichung. Gewöhnlich heißt eine Differentialgleichung dann, falls sie nur von einer reellen Variablen abhängt; anderenfalls spricht man von partiellen Differentialgleichungen. Ist F eine multilineare Funktion, so redet man von einer linearen Differentialgleichung. Dies bedeutet, dass x(t) und seine Ableitungen etwa nicht in Potenzen oder dergleichen auftreten. Zur Einführung in das Gebiet der Differentialgleichungen besprechen wir im folgenden eine Reihe von verschiedenen Beispielen.

<beispiel> Beispiel 7.1.1 Newton'sches Bewegungsgesetz

Das Newton'sche Bewegungsgesetz besagt, dass die Bahn oder Trajektorie eines Punktteilchens der Masse m im Raum durch eine Abbildung x:\R\rightarrow \R^3 beschrieben wird, welche einer Differentialgleichung der Gestalt

m \ddot x (t) = F(x(t), \dot x(t), t)

genügt. Dabei beschreibt F: \R^3\times \R^3\times \R\rightarrow \R^3 die Kräfte, die auf das Punktteilchen wirken. So ist etwa

* \ddot x(t)= (0,0,-g) die Differentialgleichung des freien Falls in x\_3-Richtung, wobei g die Gravitationskonstante ist;
* m \ddot x(t)= -c\_W A \dot x (t)\norm{\dot x(t)} +(0,0,-gm)

die Differentialgleichung des freien Falls mit Luftwiderstand bei Windstille, wobei c\_W der (dimensionslose) Widerstandsbeiwert und A der maximale Querschnitt des fallenden Körpers ist;

* m \ddot x(t)=-\omega^2 x(t) die Differentialgleichung für eine harmonische Schwingung.

</beispiel>

<beispiel> Beispiel 7.1.2 Exponentielles Wachstum

Ein exponentielles Wachstums- bzw. Zerfallsverhalten ist typisch für Populationsprobleme, den radioaktiven Zerfall oder die Abnahme der Temperatur eines Gegenstandes. Derartige Prozesse werden wie folgt modelliert:

Sei N(t) eine Menge von Individuen oder Objekten, die zum Zeitpunkt t=0 den Wert N\_0\neq 0 hat. Wir nehmen an, dass sich N(t) in gleichen Zeitabständen um die gleiche Proportion ändert. Dies bedeutet mathematisch, dass \frac{N(t)}{N(0)} = \frac{N(h+t)}{N(h)}, d.h.

\frac{N(t)}{N(0)}\cdot \frac{N(h)}{N(0)} = \frac{N(h+t)}{N(0)}.

Setzt man E(t):= N(t)/N\_0 mit N\_0:=N(0), so erhält man die Funktionalgleichung,

E(t)\cdot E(h) = E(t+h) \forall t,h\in\R.

Aus der Funktionalgleichung kann man dann eine Differentialgleichung für N gemäß

\lim\_{h\rightarrow 0} \frac{E(t+h)-E(t)}{h} = E(t)\lim\_{h\rightarrow 0} \frac{E(h)-1}{h}, also

\dot N(t) = \dot E(0)N(t) =: c N(t)

herleiten. Man kann die Differentialgleichung auch ohne die Funktionalgleichung einsehen. In einem bestimmten Zeitabschnitt \Delta t=(t+h)-t=h ändert sich N um \Delta N = N(t+h)-N(t), und wir setzen voraus, dass \Delta N proportional der noch vorhandenen Menge N(t) mal dem Zeitabschnitt \Delta t ist, so dass \Delta N \sim N \Delta t, \frac{N(t+h)-N(t)}{h} = \lambda N(t).

Dabei ist, wenn N monoton fallend bzw. wachsend sein soll, \lambda<0 bzw.>0 zu wählen. Die Konstante \lambda heißt Zerfallskonstante. Im Limes h\rightarrow 0 ergibt sich damit wieder die Differentialgleichung \dot N(t)= \lambda N(t), welche man durch Integration direkt lösen kann gemäß

\lambda = \frac{\dot N(t)}{ N(t)} = \left(\ln N \right)', wobei

\ln N = \lambda t + c\_0, N(t) = C\_0 e^{\lambda t}.

Bei t=0 ist N(t)=C\_0, so dass wir C\_0=N\_0 als die Ausgangsmenge des Stoffes interpretieren können. Wir erhalten also das Gesetz des exponentiellen Wachstums bzw. Zerfalls:

N(t) = N\_0 e^{\lambda t}

Bei einem echten Zerfallsprozess definiert man die Halbwertszeit \tau als diejenige Zeit t, zu der die Hälfte des ursprünglich zerfallsfähigen Materials zerfallen ist; sie ist also die Lösung der Gleichung

N\_0/ 2 = N\_0e^{\lambda\tau}, \text{ also } \tau = \frac{\ln 2}{|\lambda|}.

Nimmt man dagegen \lambda>0, so ergibt sich ein exponentielles Wachstum. Dies ist etwa ein gutes Modell für die Vermehrung von Bakterien bei unbeschränktem Nahrungsangebot.

</beispiel>

<beispiel> Beispiel 7.1.3 logistisches Wachstum

In manchen Situationen ist unbeschränktes exponentielles Wachstum nicht realistisch, etwa wenn Bakterien auf einer begrenzten Nährlösung gezüchtet werden, eine sich mit dem Wachstum verbrauchende Ressource, oder es um den Absatz eines Produkts geht, da ein beschränkter Markt irgendwann gesättigt ist. Sei deshalb N(t)\geq 0 eine nicht weiter spezifizierte, von oben durch M>0 begrenzte Menge. Man setzt dann voraus, dass \dot N(t) proportional zum aktuellen Bestand N(t) und der noch vorhandenen Kapazität M-N(t) ist, also

\dot N(t)= \lambda N(t) (M-N(t))

gilt. Wieder ist eine Auflösung mittels Integration möglich, denn unter Benutzung einer Partialbruchzerlegung folgt

\lambda = \frac{\dot N(t)}{N(t) (M-N(t))} =

\frac{\dot N(t)}{M} \left[ \frac{1}{N(t)}+\frac{1}{M-N(t)}\right]

= \frac{1}{M}\left[ (\ln N)' - (\ln (M-N(t)))'\right],

so dass man M \lambda t+ \ln c = \ln N(t) - \ln (M-N(t)) = \ln\frac{N(t)}{M-N(t)}

für ein c>0 erhält. Exponentiation und Auflösung nach N führen dann auf

ce^{M\lambda t} = \frac{N(t)}{M-N(t)},

ce^{-M\lambda t} = \frac{M-N(t)}{N(t)} = \frac{M}{N(t)} - 1

und man erhält die Gesamtlösung

N(t) = \frac{M}{ce^{-M\lambda t} +1} mit c= \frac{M}{N\_0} -1.

Der Graph einer solchen logistischen Funktion ist typischerweise von folgender Gestalt.

<Abbildung>

Graph der Funktion N(t)=\frac{3}{1+5e^{-2t}}.

Der Graph steigt erst nur schwach an, und wird dann immer steiler. Dann flacht er aber wieder ab und nähert sich asymptotisch einem Wert an. Er hat eine ähnliche Form wie der Graph des Arcus Tangens.

</Abbildung>

Wie man leicht zeigt, hat er bei t= \ln c / \lambda M einen Wendepunkt.

</beispiel>

Wir diskutieren nun den Fall einer homogenen linearen Differentialgleichung zweiter Ordnung und betrachten hierzu die allgemeine homogene lineare Differentialgleichung mit konstanten Koeffizienten

\label{eq:15}

a \ddot{x}(t)+ b\dot{x}(t)+cx(t) = 0, a,b,c\in \R, a\neq 0.

Offenbar gilt:

1. x(t)=0 ist eine Lösung von \eqref{eq:15},
2. Ist x(t) eine Lösung, dann ist auch k\cdot x(t) für jedes k\in\R eine Lösung,
3. Sind x\_1(t) und x\_2(t) Lösungen, dann ist auch x\_1(t)+x\_2(t) eine Lösung.

Der Lösungsraum von \eqref{eq:15} ist demnach ein Vektorraum.

Wir machen nun für eine mögliche Lösung den Ansatz x(t)=e^{rt}. Wegen \dot{x}(t)=re^{rt} und \ddot{x}(t)=r^2e^{rt} ist \eqref{eq:15} dann äquivalent zu ar^2 e^{rt}+ br e^{rt}+ ce^{rt}=0.

Damit ist x(t)=e^{rt} genau dann eine Lösung, falls r eine Nullstelle der sogenannten charakteristischen Gleichung ar^2+br+c=0 ist. Wir wissen, dass die Lösung dieser Gleichung entscheidend von ihrer Diskrimante \Delta:= b^2-4ac abhängt, man erhält also folgende Fallunterscheidung:

#### 1. Fall: \Delta = b^2-4ac \geq 0}

In diesem Fall hat die charakteristische Gleichung mindestens eine reelle Nullstelle, welche wir mit r\_1 bezeichnen. Dann ist x\_1(t)=e^{r\_1t} eine Lösung von \eqref{eq:15}. Ist y(t) irgendeine Lösung, dann kann man y(t)=h(t)e^{r\_1t} schreiben und erhält

\dot{y} = e^{r\_1t}[\dot{h}+ h r\_1],

\ddot{y} = e^{r\_1t}[\ddot{h}+ 2\dot{h}r\_1 + h r\_1^2].

Also ist y(t) genau dann eine Lösung, falls

0 = a\ddot{y} + b\dot{y}+ cy

= a e^{r\_1t}[\ddot{h}+ 2\dot{h}r\_1 + h r\_1^2] + b e^{r\_1t}[\dot{h}+ h r\_1] + c he^{r\_1t}

= e^{r\_1t}[ a(\ddot{h}+ 2\dot{h}r\_1 + h r\_1^2)+ b(\dot{h}+ h r\_1)+ch]

= e^{r\_1t}[ a\ddot{h}+ 2a \dot{h}r\_1+ b\dot{h}+0\cdot h]

= e^{r\_1t}[ a \ddot{h}+ \dot{h} (2ar\_1+b)].

Dies bedeutet genau, dass a \ddot{h}+ \dot{h} (2ar\_1+b)=0 gelten muß, und wir erhalten eine Differentialgleichung für h, welche leicht zu lösen ist. Ist r\_2 die zweite reelle Nullstelle der charakteristischen Gleichung, so gilt stets r\_1+r\_2=-b/a, also 2 ar\_1+ b = a (r\_1-r\_2), so dass die Differentialgleichung für h wegen a\neq 0 äquivalent zu \ddot{h}+ \dot{h} (r\_1-r\_2)=0 ist.

Angenommen, \Delta=b^2-4ac = 0. In diesem Fall hat die charakteristische Gleichung eine doppelte Nullstelle und es ist r\_1=r\_2, \ddot{h}=0. Damit existieren Konstanten

k\_1,k\_2\in\R mit h(t)=k\_1+k\_2 t. Der Lösungsraum von \eqref{eq:15} ist in diesem Fall zweidimensional und die allgemeine Lösung hat die Gestalt \tag{\Delta=0}

y(t) = (k\_1+k\_2 t)e^{r\_1t}.

Gelte nun \Delta=b^2-4ac > 0. In diesem Fall sind r\_1 \neq r\_2 beide reell und es gilt \dot{h}= k e^{(r\_2-r\_1)t}, so dass h = \frac{k}{r\_2-r\_1} e^{(r\_2-r\_1)t} + k\_1. Setzt man dieses Ergebnis ist den ursprünglichen Ansatz für y ein, so erhält man y= k\_1e^{r\_1t}+ \frac{k}{r\_2-r\_1}e^{r\_2t}=: k\_1e^{r\_1t}+ k\_2 e^{r\_2 t}. Die allgemeine Lösung hat demzufolge die Gestalt \tag{\Delta>0}

y(t) = k\_1e^{r\_1t}+ k\_2 e^{r\_2 t}.

#### 2. Fall: \Delta=b^2-4ac < 0

In diesem Fall hat die charakteristische Gleichung zwei zueinander komplex konjugierte Nullstellen r\_{1,2}=\alpha\pm i\beta, r\_1-r\_2=2i\beta und h erfüllt \ddot{h}=-2i\beta \dot{h}. Nun hat man zwei Möglichkeiten. So prüft man einerseits nach, dass e^{\alpha t}\sin (\beta t) und

e^{\alpha t}\cos(\beta t) zwei Lösungen sind, und argumentiert, dass der Lösungsraum maximal zweidimensional sein kann. Demzufolge ist die allgemeine Lösung gemäß \tag{\Delta <0}

y(t) = k\_1 e^{\alpha t}\sin (\beta t)+ k\_2 e^{\alpha t}\cos (\beta t)

darstellbar. Andererseits können wir stattdessen auch einfach die Gleichung für h lösen. Sie lautet nach erster Integration \dot{h}= k\_1 e^{-2i\beta t } und nach zweiter Integration

h= k\_1 e^{-2i\beta t } +k\_2.

Damit ist

y(t) = h(t)e^{(\alpha+i\beta)t} = e^{\alpha t}[k\_1 e^{-i\beta t } + k\_2e^{i\beta t }]

= e^{\alpha t}[ (k\_1+k\_2)\cos(\beta t)+ i(k\_2-k\_1) \sin(\beta t)] .

Damit die Funktion insgesamt reell ist, fordert man k\_{1,2}= c\_1\pm i c\_2, was auf das vorherige Ergebnis führt. Lösungen, die eine Basis des Lösungsraums bilden, nennt man auch Fundamentallösungen. Eine einzelne Lösung x(t) im Lösungsraum wird durch die Vorgabe von x(0) und \dot x(0) ausgewählt. Insgesamt erhalten wir somit folgenden

<satz> Satz 7.1.4

Die reellen Lösungen der allgemeinen homogenen linearen Differentialgleichung zweiter Ordnung mit konstanten Koeffizienten

a \ddot{x}(t)+ b\dot{x}(t)+cx(t) = 0, a,b,c\in \R, a\neq 0,

bilden einen zweidimensionalen reellen Vektorraum V, dessen Basis von den Nullstellen r\_1,r\_2 sowie der Diskriminante \Delta=b^2-4ac der charakteristischen Gleichung ar^2+br+c = 0 abhängt. Es gilt:

1. Für \Delta>0 sind r\_1, r\_2 reell und die Funktionen e^{r\_1 t}, e^{r\_2 t} bilden eine Basis von V;
2. Für \Delta=0 ist r\_1=r\_2\in\R und die Funktionen e^{rt}, t e^{rt} bilden eine Basis von V;
3. Für \Delta<0 sind r\_1, r\_2 komplex konjugiert, r\_{1,2}=\alpha\pm i\beta, und

e^{\alpha t}\sin (\beta t), e^{\alpha t}\cos (\beta t) sind eine Basis von V.

</satz>

Eine Anwendung des obigen Satzes ist folgendes Beispiel.

<beispiel> Satz 7.1.5 Der gedämpfte Oszillator

Aus der Mechanik ist bekannt, dass die Differentialgleichung für einen Oszillator mit Masse m=1 und linearisierter Dämpfung \ddot{x} +2\gamma \dot{x}+ \omega^2x = 0, \gamma, \omega \in \R, lautet. Dabei ist etwa x(t) die Auslenkung aus der Ruhelage eines Pendels und \dot x(t) seine Winkelgeschwindigkeit. Für \gamma=0 ist dies die Gleichung einer harmonischen Schwingung der Frequenz \omega, für \gamma>0 bzw. \gamma<0 wird die Schwingung gedämpft bzw. angetrieben.

Wir betrachten zunächst den Fall \gamma=0. Die charakteristische Gleichung hat dann die Nullstellen \pm i\omega und \eqref{eq:15} die Fundamentallösungen e^{\pm i\omega t} bzw.

\cos (\omega t),\sin(\omega t). Alle Lösungen sind periodisch und stellen geschlossene Trajektorien dar. Zur Veranschaulichung derselben, bedient man sich des sogenannten Phasenraumes \P, welcher im vorliegenden Fall durch die Ebene mit Koordinaten (x(t), \dot x(t)) gegeben ist. Die Vorgabe der Anfangsbedingungen x(0)=x\_0 und \dot x(0)=v\_0 entspricht dann genau einem Punkt im Phasenraum, durch den eine eindeutige Lösung der Differentialgleichung geht.

Ist etwa (1,0)\in\P=\R^2 mit \omega=1 gegeben, also x(0)=1, \dot x(0)=0, so lautet die Lösung (x(t),\dot x(t))= (\cos t, -\sin t) und man erhält folgendes Bild im Phasenraum.

<Abbildung>

Harmonischer Oszillator mit zwei typischen Lösungen (\omega=1, \gamma=0).

Man sieht ein Koordinatensystem, in dem viele kleine Vektoren eingezeichnet sind, wie bei einem Kraftfeld in der Physik. Diese folgen konzentrischen Kreisen, und sind unabhängig von ihrer Position stets gleich lang. Es sind zwei Trajektorien eingezeichnet, die je einen der konzentrischen Kreise verfolgen.

</Abbildung>

Gilt \gamma\neq 0, so sind die Nullstellen der charakteristischen Gleichung

\lambda\_{1,2} = \left\{

-\gamma \pm i\sqrt{\omega^2-\gamma^2} \text{ für }\gamma^2<\omega^2 (\Delta<0)

-\gamma \pm \sqrt{\gamma^2-\omega^2} \text{ für } \gamma^2 >\omega^2 (\Delta > 0)

-\gamma \text{ (doppelt) } \text{ für }\gamma^2 =\omega^2 (\Delta=0).

\right.

Ist \Delta<0, so lauten die reellen Fundamentallösungen

e^{-\gamma t}\cos \sqrt{\omega^2-\gamma^2 }t,

e^{-\gamma t} \sin \sqrt{\omega^2-\gamma^2 }t.

Gilt \Delta>0, so haben die Nullstellen \lambda\_{1,2} dasselbe Vorzeichen wie \gamma. Die reellen Fundamentallösungen sind in diesem Fall durch

e^{-\gamma t +t\sqrt{\gamma^2-\omega^2 }},

e^{-\gamma t -t\sqrt{\gamma^2-\omega^2 }}

gegeben. Im Grenzfall {\Delta=0} ist \gamma^2=\omega^2. Dann sind die Fundamentallösungen durch e^{-\gamma t} und te^{-\gamma t} gegeben.

Für \gamma>0 laufen beide Fundamentallösungen für t\rightarrow \infty in x\_0=0 hinein. In diesem Fall liegt eine echte Dämpfung vor.

<Abbildung>

\gamma>0: Oszillator mit Lösungen durch (-1,0), (1,1) und (1,-1).

Die Lösungen nähern sich zunächst der zweiten Winkelhalbierenden, und dann immer mehr dem Ursprung des Koordinatensystems. Die Pfeile auf den Lösungen zeigen in den Ursprung hinein. Die Lösung für (1, -1) bewegt sich komplett entlang der zweiten Winkelhalbierenden.

</Abbildung>

Gilt jedoch {\gamma=-1<0}, so entfernt sich jede Anfangskonfiguration vom Nullpunkt.

<Abbildung>

\gamma<0: Oszillator mit Lösungen durch (-1,0), (1,1), (1,-1).

Die Lösungen bewegen sich ungefähr parallel zur ersten Winkelhalbierenden in entgegengesetzte Richtung des Ursprungs. Die Pfeile auf den Lösungen zeigen vom Ursprung weg.

</Abbildung>

Offensichtlich liegen für \gamma=0 und \gamma<0 zwei unterschiedliche Formen stabilen Verhaltens vor.

</beispiel>

## 7.2 Die Reduktion einer Differentialgleichung k-ter Ordnung auf eine Differentialgleichung erster Ordnung

Gegeben sei eine Abbildung F: \R^n\times \ldots\times \R^n\times \R\rightarrow \R^n und

F^\*: \R^{kn}\times \R\rightarrow \R^{kn} durch

F^\*(y\_0,\ldots, y\_{k-1},t) = (y\_1,\ldots, y\_{k-1}, F(y\_0,\ldots, y\_{k-1},t)), y\_i \in \R^n,

gegeben.

<lemma> Lemma 7.2.1

Ist x: (t\_0-\epsilon, t\_0+\epsilon)\rightarrow \R^n eine Lösung der Differentialgleichung k-ter Ordnung

\label{eq:16}

\frac{d^k x}{d t^k}(t) = F \bigg(x(t), \frac{dx}{dt}(t), \ldots, \frac{d^{k-1}x}{dt^{k-1}} (t) , t\bigg)

mit den Anfangsbedingungen

x(t\_0)=a\_0, dx/dt (t\_0)=a\_1,\ldots, d^{k-1} x/ dt^{k-1} (t\_0)=a\_{k-1},

so ist y: (t\_0-\epsilon, t\_0+\epsilon) \rightarrow \R^{kn}, y(t)=(y\_0(t), \dots y\_{k-1}(t)) =\big(x(t),\frac{dx}{dt}(t),\ldots, \frac{d^{k-1}x}{dt^{k-1}} (t) \big)

eine Lösung der Differentialgleichung erster Ordnung

\label{eq:17}

\frac{dy}{dt} = F^\*(y(t),t)

mit den Anfangsbedingungen y(t\_0)=(a\_0,\ldots,a\_{k-1}) und umgekehrt.

</lemma>

<beweis>

Offenbar ist \dot y (t) = F^\* (y(t), t) äquivalent zu den Gleichungen

\dot y\_0 (t) = y\_1 (t), \dot y\_1 (t) = y\_2 (t), \ldots,

\dot y\_{k-2} (t) = y\_{k-1} (t), \dot y\_{k-1} (t) = F(y\_0 (t), \ldots, y\_{k-1} (t), t),

woraus sich unmittelbar die Behauptung ergibt.

</beweis>

<beispiel> Beispiel 7.2.2

Die Differentialgleichung für den Oszillator lautet \ddot x(t) +2\gamma\dot x(t)+\omega^2 x(t)=0. Es ist n=1, k=2 und die Abbildung F(x,y,t) ist F(x,y,t) = -2\gamma y -\omega^2 x. Dann ist

F^\* : \R^2\times \R \rightarrow \R^2, F^\* (y\_0, y\_1,t) = (y\_1,-\omega^2 y\_0-2\gamma y\_1)

und die Gleichung \ddot x(t) = -2\gamma\dot x(t)-\omega^2 x(t) ist äquivalent zu

\frac{d}{dt} \begin{bmatrix} y\_0\ y\_1 \end{bmatrix}

= \begin{bmatrix} 0 & 1\ -\omega^2 & -2\gamma \end{bmatrix}

\begin{bmatrix} y\_0\ y\_1 \end{bmatrix},

mithin zu \dot y\_0 = y\_1, \dot y\_1 = -\omega^2 y\_0 - 2\gamma y\_1.

</beispiel>

Infolge obigen Lemmas kann die Untersuchung von Differentialgleichungen in vielen Fällen auf das Studium von Differentialgleichungen erster Ordnung zurückgeführt werden. Ein Spezialfall von Lemma 7.2.1 liegt vor, falls F nicht von t abhängt. Dies führt zu folgender

<definition> Definition 7.2.3

Eine Differentialgleichung heißt autonom, falls sie von der Form \label{eq:18}

\frac{dx}{dt} (t)= F(x(t))

ist mit F : U\subset \R^n \rightarrow \R^n unabhängig von t. In diesem Fall können wir

F = (F^1, F^2,\ldots, F^n) als Vektorfeld \vec{F} auf U \subset \R^n interpretieren. Die Lösungen von \eqref{eq:18} werden Integralkurven des Vektorfeldes \vec{F} genannt.

</definition>

Autonome Differentialgleichungen lassen eine einfache geometrische Bedeutung zu. So ist eine Lösung von \eqref{eq:18} durch eine Abbildung

x : (t\_0 - \epsilon, t\_0 + a) \rightarrow U \subset \R^n

gegeben, welche als Kurve in U verstanden werden kann, deren Tangentialvektor in jedem Punkt durch das Vektorfeld \vec{F} vorgegeben ist.

<beispiel> Beispiel 7.2.4

Dem Vektorfeld \vec{F} (x,y) =( - y ,x) entspricht die Differentialgleichung

\frac{d}{dt} \begin{bmatrix} x(t) \ y(t) \end{bmatrix} = \begin{bmatrix} -y(t) \ x(t) \end{bmatrix}

= \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} x(t) \ y(t) \end{bmatrix}.

Die dazugehörige Differentialgleichung zweiter Ordnung ist die Schwingungsgleichung \frac{d^2 x (t)}{dt^2} = -x (t).

</beispiel>

## 7.3 Elementare Integrationsmethoden

Die Bestimmung der Lösungen einer gegebenen Differentialgleichung stellt im allgemeinen ein schwieriges Problem dar. Im folgenden besprechen wir einige elementare Methoden zur Lösung von Differentialgleichungen.

### 7.3.1 Trennung der Variablen

Seien U\_1 und U\_2 offene Mengen in \R. Die Methode der Trennung der Variablen ist anwendbar auf Differentialgleichungen auf \R vom Typ \label{eq:19}

\frac{dx (t)}{dt} = f (t) g(x(t)), x(t\_0) = x\_0,

wobei f : U\_1\subset \R \rightarrow \R, g : U\_2\subset \R \rightarrow \R stetige Funktionen und t\_0 \in U\_1, x\_0 \in U\_2, sind. Es können nun folgende beiden Fälle eintreten.

#### 1. Fall

Gilt g(x\_0) = 0, so ist x(t) \equiv x\_0 eine Lösung des Problems \eqref{eq:19}, welche auf der gesamten Menge U\_1 \subset \R definiert ist.

#### 2. Fall

Sei g(x\_0) \neq 0. Dann existiert eine Zahl \epsilon > 0 mit |x-x\_0| < \epsilon \Longrightarrow g(x) \neq 0. Auf (x\_0 - \epsilon, x\_0 + \epsilon) ist also \frac{1}{g(x)} definiert.

Sei G(x) eine Stammfunktion von \frac{1}{g(x)} auf (x\_0 - \epsilon, x\_0 + \epsilon), welche eine streng monotone Funktion G : (x\_0 - \epsilon, x\_0 + \epsilon) \rightarrow (G(x\_0) - \epsilon\_1, G (x\_0) + \epsilon\_2) darstellt.

<satz> Satz 7.3.1

Die einzige in einer Umgebung von t\_0 \in U\_1 definierte Lösung x(t) des Problems \eqref{eq:19} ist durch

x(t) = G^{-1} \left(G(x\_0) + \int^t\_{t\_0} f(\mu) d\mu)\right)

gegeben, vorausgesetzt, dass g(x\_0) \neq 0. Diese Lösung ist auf der Menge

U\_1 \cap \{t : -\epsilon\_1 < \int^t\_{t\_0} f(\mu) d\mu) < \epsilon\_2\}

definiert.

</satz>

<beweis>

Angenommen, x(t) ist Lösung von \frac{dx(t)}{dt} = f(t) g(x(t)), x (t\_0) = x\_0 und g(x\_0) \neq 0. Für t nahe an t\_0 können wir dann durch g(x (t)) teilen und erhalten \frac{dx(t)/dt}{g(x(t))} = f(t).

Integrieren wir von t\_0 bis t, so folgt mit dem Hauptsatz der Differential- und Integralrechnung

G(x(t)) - G(x(t\_0)) = \int^t\_{t\_0} f(\mu) d\mu,

 also x(t) = G^{-1} (G (x\_0) + \int^t\_{t\_0} f(\mu) d\mu). Beachte dabei, dass G aufgrund der Monotonie injektiv ist und somit die inverse Abbildung G^{-1} existiert.

Damit ist die Existenz und Eindeutigkeit gezeigt.

 </beweis>

<bemerkung> Bemerkung 7.3.2

Im ersten Fall kann man nicht beweisen, dass x(t) \equiv x\_0 die einzige Lösung ist. Betrachte zum Beispiel f(t) \equiv 1, t\_0 = 0, g(x) = 3 \sqrt[3]{x^2}, x\_0 = 0. Dann ist x(t)\equiv 0 Lösung der Gleichung \frac{dx (t)}{dt} = 3 \sqrt[3]{x(t)^2} mit x(0) = 0. Jedoch stellt x^\*(t) = t^3 eine weitere Lösung dar. Tatsächlich erfüllt diese Funktion auch x^\* (0) = 0 und

 \frac{dx^\* (t)}{dt} = 3t^2,

3 \sqrt[3]{(x^\*(t))^2} = 3 \sqrt[3]{t^6} = 3t^2,

so dass dx^\* (t)/dt= 3 \sqrt[3]{x^\*(t)^2}.

</bemerkung>

<beispiel> Beispiel 7.3.3

Wir wollen die Differentialgleichung \frac{dx}{dt} = \frac{x^2-1}{2} mit x(0) = x\_0 lösen.

1. Fall

Es gilt x\_0 = \pm 1. Dann ist x(t) \equiv x\_0 = \pm 1 eine Lösung

2. Fall

Es ist x\_0 \neq \pm 1. Dann folgt aus 2 \frac{\dot x}{x^2-1}=1

durch Integration 2 \int^t\_0 \frac{\dot x (\mu)}{x^2 (\mu)-1} d\mu = t.

Da jedoch \int \frac{dy}{y^2-1} = \frac{1}{2} \ln \big| \frac{y-1}{y+1}\big|,

folgt \ln \bigg|\frac{x(t)-1}{x(t)+1}\bigg| - \ln \bigg| \frac{x\_0 -1}{x\_0 +1}\bigg| = t

und damit \big|\frac{x(t)-1}{x(t)+1}\big| = \big|\frac{x\_0 -1}{x\_0 +1}\big| e^t.

Weil e^t nirgends veschwindet, schließt man jetzt

\frac{x(t)-1}{x(t) +1} = \frac{x\_0 -1}{x\_0 +1} e^t.

Damit folgt x(t) - 1 = x(t) \frac{x\_0 -1}{x\_0 +1} e^t + \frac{x\_0 -1}{x\_0 +1} e^t,

x(t) \bigg[1 - \frac{x\_0 -1}{x\_0 +1} e^t \bigg] = 1 + \frac{x\_0 -1}{x\_0 +1} e^t.

Somit ist x(t) = \frac{ 1+ \frac{x\_0 -1}{x+ 1} e^t}{1- \frac{x\_0 -1}{x+ + 1} e^t}

die Lösung des Problems \dot x = \frac{x^2 -1}{2}, x(0) = x\_0 für x\_0 \neq \pm 1.

Wir diskutieren diese Lösung ausführlicher:

1. Ist x\_0 < -1, so gilt 1 > \frac{x\_0 + 1}{x\_0-1} > 0 und die Lösung ist nur für t\in \big( \ln \big(\frac{x\_0 + 1}{x\_0-1}\big), \infty\big) definiert
2. Ist -1 < x\_0 < 1, so folgt \frac{x\_0 + 1}{x\_0-1} < 0 und die Lösung ist für alle Parameterwerte t definiert.
3. Ist 1 < x\_0, so gilt \frac{x\_0 + 1}{x\_0-1} > 1 und die Lösung ist für

t\in (-\infty, \ln \big(\frac{x\_0 + 1}{x\_0-1}\big)\big) definiert.

Insgesamt ergibt sich für die Graphen der Lösungen des Problems

\dot x = x^2 = \frac{x^2 -1}{2}, x(0) = x\_0 folgendes Bild.

<Abbildung>

Man sieht ein Koordinatensystem mit einem Vektorfeld, das horizontal (also entlang der t Achse) verschiebungssymmetrisch ist. Die Vektoren zeigen im Allgemeinen in positive t und x Richtung. Näher der t-Achse ist der x-Anteil deutlich kleiner, und der t-Anteil etwas größer.

Die Lösungen verlaufen für den 1.Fall in zwei parallelen Linien oberhalb und unterhalb der t-Achse.

Für (1) in der obigen Auflistung verläuft die Lösung unterhalb der t-Achse zunächst steil in Richtung der Achse und flacht dann schnell ab. Die Lösung existiert nicht für alle Zeiten, weshalb die Trajektorie plötzlich endet.

Für (2) verläuft die Lösung entlang der t-Achse.

Für (3) verläuft die Lösung oberhalb der t-Achse zunächst nah an der t-Achse und langsam von ihr weg, und steigt dann exponentiell an. Die Lösung existiert nicht für alle Zeiten, weshalb die Trajektorie plötzlich endet.

Im Punkt (t,x) ist ein Vektor mit Steigung (x^2-1)/2 eingezeichnet, dessen Länge keine weitere Bedeutung hat. Weil die Differentialgleichung autonom ist, ist das Vektorfeld t-unabhängig, also translationsinvariant in t-Richtung.

</Abbildung>

</beispiel>

<beispiel> Beispiel 7.3.4

Wir betrachten die Differentialgleichung \dot x = \frac{x}{t^2+1} x(0) = x\_0. Wieder sind zwei Fälle zu unterscheiden.

#### 1. Fall

Gilt x\_0 = 0, so ist x(t) \equiv 0 eine Lösung.

#### 2. Fall

Für x\_0=c\neq 0 formen wir die Differentialgleichung gemäß \frac{\dot x}{x} = \frac{1}{1+t^2} um und erhalten als Lösung \ln |x(t)| = \arctan t + \ln c, also x(t) = c e^{\arctan t}, was in der Tat x\_0 = c erfüllt. Damit ist x(t) = x\_0 e^{\arctan t} eine Lösung, die asymptotisch gegen x\_0 e^{\pi/2} tendiert

</beispiel>

### 7.3.2 Ähnlichkeitsdifferentialgleichungen

Eine Klasse von Differentialgleichungen, auf welche die Methode der Trennung der Variablen angewandt werden kann, sind sogenannte Ähnlichkeitsdifferentialgleichungen. Eine solche Differentialgleichung ist eine Gleichung vom Typ

\frac{dx }{dt} (t) = f\bigg( \frac{x (t)}{t}\bigg).

Mit der Substitution u(t) := \frac{x (t)}{t} ergibt sich

\frac{du}{dt} = \frac{t \frac{dx}{dt} - x(t)}{t^2} = \frac{t f(u(t))-x(t)}{t^2}

= \frac{1}{t} (f(u(t)) - u(t))

 und wir erhalten folgende

<proposition> Proposition 7.3.5

Die Ähnlichkeitsdifferentialgleichung \dot x = f\big(\frac{x}{t}\big) geht durch die Substitution u(t) = \frac{x(t)}{t} in die Gleichung \frac{du (t)}{dt} = \frac{1}{t} \cdot (f(u) - u) über, welche mittels der Methode der Trennung der Variablen gelöst werden kann.

</proposition>

<beispiel> Beispiel 7.3.6

Betrachte das Problem \dot x = 1 + \frac{x}{t}, x(1) = x\_0.

Dann liefert die Substitution u(t) = \frac{x(t)}{t} die Ableitung

\frac{du}{dt} = \frac{t \dot x -x}{t^2} = \frac{t (1+ \frac{x}{t})-x}{t^2} = \frac{1}{t}.

Also ist u(t) = \ln t + C und damit \frac{x(t)}{t} = \ln t + C, x(t) = t ( \ln t + x\_0) eine Lösung des Problems. Diese Lösungen sind für alle t\in (0,\infty) definiert.

</beispiel>

<beispiel> Beispiel 7.3.7

Wir betrachten die Differentialgleichung \dot x = 1 + \frac{x}{t} + ( \frac{x}{t})^2, t>0.

Mit der Substitution u = \frac{x}{t} ist dann \dot u = \frac{t \dot x-x}{t^2} = \frac{1+u^2}{t}, also \frac{\dot u}{1+u^2} = \frac{1}{t}, woraus sich

\arctan u(t) = \ln t + C,

u(t) = \tan (\ln t + C),

x(t) = t\cdot \tan (\ln t + C)

ergibt.

</beispiel>

### 7.3.3 Lineare Differentialgleichungen erster Ordnung und die Methode der Variation der Konstanten

Eine lineare Differentialgleichung erster Ordnung ist eine Differentialgleichung der Form

\frac{dx}{dt} = p(t) x(t) + q(t), x(t\_0) = x\_0.

Zur Lösung derselben verfahren wir wie folgt. Wir lösen zunächst die homogene Gleichung \frac{dx}{dt} = p(t) x(t) mittels der Methode der Trennung der Variablen und erhalten zunächst als Lösungsschar x(t) = C \cdot \exp\left[\int^t\_{t\_0} p(\mu) d\mu\right].

Jetzt fassen wir C als Funktion von t auf und setzen

x(t) = C(t) \exp\left[\int^t\_{t\_0} p(\mu) d\mu\right]

als Lösung von \dot x = p(t) x(t) + q(t) an. Dann folgt

\dot C (t) \exp\left[{\int^t\_{t\_0} p (\mu) d\mu}\right]

+ C(t) \exp\left[{\int^t\_{t\_0} p(\mu) d\mu}\right] p(t)

= p(t) C(t) \exp\left[{\int^t\_{t\_0} p (\mu) d\mu}\right] + q(t),

also \dot C (t) = q(t) / \exp\left[{\int^t\_{t\_0} p (\mu) d\mu}\right]

und somit

C(t) = \int^t\_{t\_0} q(s) \exp\left[{-\int^s\_{t\_0} p (\mu) d\mu}\right] ds .

Damit ist die Lösung des Problems gegeben durch

x(t) = \left\{ \int^t\_{t\_0} q(s) \exp\left[{-\int^s\_{t\_0} p (\mu) d\mu} \right] ds

+ x\_0\right\} \exp\left[{\int^t\_{t\_0} p (\mu) d\mu}\right].

<beispiel> Beispiel 7.3.8

Sei \dot x = t x + t e^{t^2}, x(0) = x\_0. Hier gilt p(t) = t, q(t) = t e^{t^2}. Für die homogene Gleichung gilt \dot x = t x \Longrightarrow \frac{\dot x}{x} = t, also

\ln x = \frac{1}{2} t^2 + C^\*, x(t) = C e^{1/2 t^2}. Mit dem Ansatz x(t) = C(t) e^{1/2 t^2} lautet die Bedingung \dot c (t) e^{ t^2/2} + C(t) e^{t^2/2} t = t C(t) e^{t^2/2} + t e^{t^2},

\dot c (t) e^{ t^2/2} = t e^{t^2},

c(t) = e^{t^2/2} + D.

Also ist die Lösung

x(t) = \big( e^{1/2 t^2} + D\big) e^{1/2 t^2} .

Mit x(0) = x\_0 folgt x\_0 = 1 + D. Somit ist die Lösung des Problems \dot x = t x + t e^{t^2},

x(0) = x\_0 gegeben durch

x(t) = e^{t^2} + (x\_0 - 1) e^{1/2 t^2}.

Alle Lösungen sind in diesem Fall auf t \in (-\infty, \infty) definiert.

</beispiel>

<beispiel> Beispiel 7.3.9

Wir betrachten die Differentialgleichung \cos t \cdot \frac{dx}{dt} = -2 \sin t x(t) + 2 \sin t, x(0) = x\_0 mit der Bedingung x\_0 \neq \frac{2k+1}{2} \pi.

Die homogene Gleichung \dot x = -2 \frac{\sin t}{\cos t} x, beziehungsweise

\frac{\dot x}{x} = -2 \frac{\sin t}{\cost}

lösen wir durch den Ansatz \ln x = 2 \ln \cos t + \ln C, was auf x(t) = C \cos^2 (t) führt. Variation der Konstanten ergibt nun x(t) = C(t) \cos^2 (t), was

\cos t (\dot C \cos^2 (t) - 2C \cos t \sin t) = -2C \sin t \cos^2 t + 2 \sin t

 zur Folge hat, so dass \dot C(t) = 2\sin t / \cos^3 t. Damit ist

 C(t) = \int \frac{2 \sin t}{\cos^2 t} dt = \frac{1}{\cos^2 t} + D, also

 x(t) = \big( \frac{1}{\cos^2 t} + D\big) \cos^2 t = 1 + D \cos^2 t,

und die Lösung des Problems insgesamt gegeben durch

x(t) = 1 + (x\_0 - 1) \cos^2 t.

In der folgenden Grafik ist nur das Vektorfeld eingezeichnet, da das numerische Verfahren zur Berechnung der Lösung äußerst instabil instabil ist und auf keine periodische Lösung führt.

</beispiel>

### 7.3.4 Bernoulli'sche Differentialgleichungen

Ein weitere wichtige Klasse von Differentialgleichungen sind Bernoulli'sche Differentialgleichungen. Dies sind Gleichungen vom Typ \dot{x}(t) = p(t) x(t) + q (t) (x(t))^n, wobei n eine reelle Zahl ist. Eine solche Gleichung wird durch die Substitution z(t) = (x(t))^{1-n} behandelt. Man erhält dann

\dot z = (1-n)x^{-n} \dot x = (1-n)x^{-n} (p x + q x^n)

= (1-n)(p x^{1-n} + q),

was auf die lineare Differentialgleichung

\dot{z}(t) = (1-n) p(t) z(t) + (1-n) q(t)

führt. Man löst jetzt diese lineare Differentialgleichung nach z(t) und erhält daraus als Lösung

x(t) = [z(t)]^{\frac{1}{1-n}}.

<beispiel> Beispiel 7.3.10

 Sei \dot x = -x + t \sqrt{x}, also n= \frac{1}{2}. Mit dem Ansatz z(t) = \sqrt{x(t)} ergibt sich

 \dot z = \frac{\dot x}{2 \sqrt{x}} = - \frac{x+ t \sqrt{x}}{2 \sqrt{x}} = - \frac{1}{2} \sqrt{x} + \frac{t}{2}

= - \frac{1}{2} z + \frac{t}{2}.

Wir lösen wieder zuerst die homogene Gleichung \dot z = - \frac{1}{2} z und erhalten

 \ln z = - \frac{1}{2} t + C^\* und damit z(t) = C e^{-\frac{1}{2} t}.

Variation der Konstanten führt nun auf

\dot C e^{-\frac{1}{2} t} - \frac{1}{2} C e^{-\frac{1}{2} t} = - \frac{1}{2} C e^{-\frac{1}{2}t} + \frac{t}{2}, \dot C = \frac{t}{2} e^{\frac{1}{2} t}.

Nach partieller Integration folgt C(t) = t e^{\frac{1}{2} t} - 2 e^{\frac{1}{2} t} + D. Damit ist

z(t) = (t e^{\frac{1}{2} t} - 2 e^{\frac{1}{2} t} + D) e^{-\frac{1}{2} t} = (t - 2 + D e^{-\frac{1}{2} t})

und folglich x(t) = z^2(t) = (t-2 + D e^{-\frac{1}{2} t})^2 die gesuchte Lösung des Problems.

</beispiel>

### 7.3.5 Exakte Differentialgleichungen

Wir betrachten als nächstes in \R^2 die Differentialgleichung P(t,x)+ Q(t,x) \dot{x} = 0 . Sie heißt vollständige oder exakte Differentialgleichung. Existiert eine Funktion G(t,x), die

P(t,x) = \frac{\partial G}{\partial t},

Q(t,x) = \frac{\partial G}{\partial x}

erfüllt, so sind die Funktionen G(t,x) = const implizite Lösungen der ursprünglichen Gleichung. Im Allgemeinen ist dies natürlich nicht der Fall; betrachtet man jedoch für ein f(t,x) die äquivalente Differentialgleichung

(f \cdot P)(t,x)+(f \cdot Q)(t,x) \dot{x} = 0,

so besagt der Satz von Frobenius, der erst im Rahmen der Analysis 3 bewiesen werden kann, dass es immer möglich ist, eine solche Funktion f zu finden, für die eine Funktion G mit

f(t,x) P(t,x) = \frac{\partial G}{\partial t},

f(t,x) Q(t,x) = \frac{\partial G}{\partial x}

existiert. Die Lösungskurven sind dann wieder implizit durch G(t,x)= const gegeben. Diese Funktion f wird integrierender Faktor genannt. Allerdings liefert der Satz von Frobenius keinen Algorithmus, wie der integrierende Faktor gefunden werden kann. In einfachen Fällen kann dieser jedoch direkt ersehen werden. Können wir beispielsweise nur von den Variablen t und x abhängende Funktionen F(t) und G(x) mit

\frac{\partial P{(t,x)}}{\partial x} - \frac{\partial Q{(t,x)}} {\partial t} = Q (t,x)F(t) - P(t,x)G(x)

finden, dann ist f(t,x)= \exp (\int F(t)dt) \exp(\int G(x)dx) ein integrierender Faktor.

<beispiel> Beispiel 7.3.11

Wir betrachten die Differentialgleichung (2t^2+3tx-4t)\dot{x}+(3x-2tx-3x^2) = 0 .

Dann können wir F(t)= 2/t und G(x)=- 5/x wählen und erhalten als integrierenden Faktor f=t^2 x^{-5}.

Die Lösungen der Differentialgleichung sind unter Verwendung dieses integrierenden Faktors die durch die Gleichung t^3x^{-4} - \frac{1}{2}t^4 x^{-4} -t^3x^{-3} \equiv const beschriebenen Kurven.

<Abbildung>

Die Kurven t^3-\frac{1}{2}t^2-t^3 x = c x^4 für verschiedene Werte von c

Die Kurven schneiden die x-Achse alle im Wert x = 0,5 und verlaufen nahe der x-Achse ähnlich. Je größer oder kleiner der t-Wert wird, desto weiter entfernen sich die Lösungen voneinander. Einige Lösungen steigen sehr stark an, andere sind eher flach. Alle Kurven sind linksgekrümmt und haben für negative t Werte etwas kleinere x-Werte als für positive t Werte gleichen Betrags.

</Abbildung> </beispiel>

# Kapitel 8: Lineare Differentialgleichungen

Ziel dieses Kapitels ist es, die Lösung der allgemeinen inhomogenen, linearen Differentialgleichung k-ter Ordnung

x^{(k)}(t)+ a\_1(t) x^{(k-1)}(t)+ \ldots + a\_{k-1}(t)\dot x(t)+ a\_k(t) x(t) = f(t), a\_1,\ldots, a\_k\in\R,

mit variablen Koeffizienten zu bestimmen. Zu deren Behandlung benötigen wir einige Vorbereitungen.

## 8.1 Die Exponentialfunktion für Matrizen

Es bezeichne End(\R^n) die Menge aller linearen Abbildungen L: \R^n\rightarrow \R^n. Die Operatornorm, auf End(\R^n) ist dann definiert durch

\|L\| :=\sup\_{0\neq x\in\R^n} \frac{\|Lx\|}{\|x\|}.

Infolge der Linearität von L ist

\|L\| =\sup\_{\|x\|=1} \frac{\|Lx\|}{\|x\|} = \max\_{\|x\|=1} \|Lx\|,

wobei wir berücksichtigten, dass die Einheitsphäre kompakt ist und somit das Supremum auf ihr angenommen wird. Wir beweisen als nächstes einige elementare Eigenschaften der Operatornorm.

<satz> Satz 8.1.1

 Für alle L,L'\in{\mathrm{End}}(\R^n) gilt

(1) \| Lx\|\leq \|L\|\cdot \|x\| \forall x\in\R^n,

(2) \|\lambda L\| = |\lambda|\cdot \|L\| \forall \lambda\in\R,

(3) \|L+L'\|\leq \|L\|+\|L'\|, \hfill (Subadditivität)

(4) \|LL'\|\leq \|L\|\cdot\|L'\|, \hfill (Submultiplikativität)

(5) Mit d(L,L'):=\|L-L'\| ist ({\mathrm{End}}(\R^n),d) ein vollständiger metrischer Raum.

</satz>

<beweis>

Die Eigenschaften (1)-(3) sind trivial. Für (4) wenden wir (1) an und erhalten

\|LL'\| = \sup\_{0\neq x\in\R^n} \frac{\|LL'x\|}{\|x\|}

\leq \sup\_{0\neq x\in\R^n} \frac{\|L\|\cdot \|L'x\|}{\|x\|} = \|L\|\cdot \|L'\|.

Um (5) einzusehen, betrachten wir eine Cauchy-Folge L\_n\in {\mathrm{End}}(\R^n). Dann existiert zu gegebenem \epsilon>0 eine Zahl N mit

\|L\_n-L\_m\| < \epsilon \forall m,n\geq N.

Für jedes x\in\R^n gilt demzufolge

\label{eq:20}

\|L\_n x - L\_m x\| \leq \|L\_n - L\_m\|\cdot \|x\| \leq \epsilon\|x\|,

so dass L\_n x eine Cauchy-Folge in \R^n ist. Aufgrund der Vollständigkeit von \R^n existiert der Grenzwert \lim\_{n\rightarrow \infty} L\_nx=: L(x). Damit ist für jedes x\in\R^n ein Element L(x) definiert, welches wiederum eindeutig eine lineare Abbildung L definiert, da offenbar

L(x+y) = \lim\_{n\rightarrow \infty} L\_n(x+y)

= \lim\_{n\rightarrow \infty} L\_n(x)+\lim\_{n\rightarrow \infty} L\_n(y) = Lx+Ly.

Bildet man nun in \eqref{eq:20} dem Limes m\rightarrow \infty, so folgt

\|L\_n x - Lx\| \leq \epsilon\|x\|, also

\sup\_{0\neq x\in\R^n} \frac{\|L\_n x - Lx\|}{\|x\|} \leq \epsilon

und damit \|L\_n-L\|\leq\epsilon für jedes \epsilon>0.

</beweis>

Sei nun L\in{\mathrm{End}}(\R^n) und betrachte die abgebrochene Exponentialreihe

e\_k(L) := {\mathrm{Id}} + L+\frac{L^2}{2!}+\cdots+\frac{L^k}{k!} \in {\mathrm{End}}(\R^n)

Der vorherige Satz liefert dann für k\leq m die Abschätzung

\| e\_m(L)- e\_k(L)\| \leq \frac{\|L\|^{k+1}}{(k+1)!}+\cdots+\frac{\|L\|^{m}}{m!}.

Die Konvergenz der Exponentialreihe impliziert nun, dass ein N existiert mit

\| e\_m(L)- e\_k(L)\| \leq \epsilon\forall k,m\geq N.

Damit ist e\_k(L) eine Cauchy-Folge in {\mathrm{End}}(\R^n) und wegen der Vollständigkeit dieses Raums existiert somit der Grenzwert

e^L := \lim\_{k\rightarrow \infty} e\_k(L)

= \lim\_{k\rightarrow \infty} \left[{\mathrm{Id}} + L+\frac{L^2}{2!}+\ldots

+\frac{L^k}{k!}\right] \in {\mathrm{End}}(\R^n) .

Dieses Element heißt das Exponential von L.

<beispiel> Beispiel 8.1.2

Das Exponential linearer Abbildungen erfüllt nicht die Funktionalgleichung der Exponentialfunktion e^Ae^B\neq e^{A+B}. Betrachte etwa die Matrizen A=\begin{bmatrix}1&0\ 0& 2\end{bmatrix}, B=\begin{bmatrix} 0 & 0\ 1 & 0\end{bmatrix}. Dann gilt wegen B^2=0

e^B = \begin{bmatrix}1&0\ 0& 1\end{bmatrix} + \begin{bmatrix} 0 & 0\ 1 & 0\end{bmatrix}

= \begin{bmatrix} 1 & 0\ 1 & 1\end{bmatrix},

e^A = \begin{bmatrix} e & 0\ 0 & e^2\end{bmatrix}.

Damit ist einerseits

e^A e^B = \begin{bmatrix} e & 0\ e & e^2\end{bmatrix}

und andererseits

A+B = \begin{bmatrix} 1 & 0\ 1 & 2\end{bmatrix},

(A+B)^2 = \begin{bmatrix} 1 & 0\ 3 & 4\end{bmatrix},

(A+B)^3 = \begin{bmatrix} 1 & 0\ 7 & 8\end{bmatrix}.

Also hat e^{A+B} in der linken unteren Ecke als Eintrag

\left[e^{A+B}\right]\_{21} = 1+\frac{3}{2!}+\frac{7}{3!}+\ldots \geq 3,

welcher nicht gleich e sein kann.

</beispiel>

<satz> Satz 8.1.3

Das Exponential hat folgende Eigenschaften:

(1) e^0={\mathrm{Id}};

(2) sind A,B\in End(\R^n) vertauschbar, gilt also AB=BA, so ist e^{A+B}=e^A e^B = e^B e^A;

(3) e^A ist für jedes A\in{\mathrm{End}}(\R^n) invertierbar und [e^A]^{-1}=e^{-A};

(4) \det e^A = e^{\tr A};

(5) e^{BAB^{-1}} = Be^AB^{-1} für jedes invertierbare Element B\in End(\R^n).

</satz>

<beweis>

(1) ist klar. Um (2) einzusehen, betrachten wir für m\in\N

\mathcal{R}\_m = e\_{2m}(A+B)- e\_m(A)e\_m(B) =

\sum\_{j=0}^{2m}\frac{(A+B)^j}{j!} –

\left[\sum\_{k=0}^m \frac{A^k}{k!}\right] \left[\sum\_{l=0}^m \frac{B^l}{l!}\right].

Weil A und B kommutieren, können wir die binomischen Formeln anwenden und erhalten

\mathcal{R}\_m = \sum\_{j=0}^{2m}\sum\_{k=0}^j \frac{1}{j!}

\frac{j!}{k!(j-k)!} A^k B^{j-k} - \sum\_{k,l=0}^m \frac{1}{k! l!}A^kB^l

 = \sum\_{\substack{k+l\leq 2m\ k>m \text{ oder } l>m }} \frac{1}{k!}A^k\frac{1}{l!}B^l .

Damit folgt für die Norm

\|\mathcal{R}\_m\|

\leq \sum\_{\substack{ k>m \text{ oder }\ l>m} } \frac{1}{k!}\|A\|^k\frac{1}{l!}\|B\|^l

\leq \left[\sum\_{k>m} \frac{\|A\|^k}{k!}\right]\cdot \left[\sum\_{l\geq 0} \frac{\|B\|^l}{l!}\right]

+ \left[\sum\_{k\geq 0} \frac{\|A\|^k}{k!}\right]\cdot\left[\sum\_{l>m} \frac{\|B\|^l}{l!}\right]

\leq e^{\|B\|} \cdot \sum\_{k>m} \frac{\|A\|^k}{k!}+ e^{\|A\|}\cdot \sum\_{l>m} \frac{\|B\|^l}{l!}.

Mit m\rightarrow \infty folgt alsdann \|\mathcal{R}\_m\|\rightarrow 0 und somit

0 = \lim\_{m\rightarrow \infty} \mathcal{R}\_m

= \lim\_{m\rightarrow \infty} e\_{2m}(A+B) - \lim\_{m\rightarrow \infty} e\_{m}(A)e\_{m}(B).

(3) folgt nun direkt aus (2), da [A,-A]= 0 und somit

\mathrm{Id} = e^0= e^{A-A} = e^A e^{-A}=e^{-A }e^A

gilt. Die Eigenschaft (5) ergibt sich nun unmittelbar aus

e^{BAB^{-1}} = \lim\_{k\rightarrow \infty} \left[ {\mathrm{Id}} + \frac{BAB^{-1}}{1!}

+\frac{(BAB^{-1})^2}{2!} + \cdots + \frac{(BAB^{-1})^k}{k!}\right]\\

= B\cdot \lim\_{k\rightarrow \infty} \left[ {\mathrm{Id}} + \frac{A}{1!}

+\frac{A^2}{2!}+ \cdots + \frac{A^k}{k!}\right] B^{-1}

= B e^A B^{-1}.

Wir wenden uns nun (4) zu und betrachten einen Basiswechsel B, welcher A als Endomorphismus auf \C^n auf obere Dreiecksgestalt BAB^{-1} transformiert. Bezeichnen wir die Diagonaleinträge mit \lambda\_1,\ldots, \lambda\_n, so ist einerseits

\det (e^A)& =& \det(B e^A B^{-1}) = \det(e^{BAB^{-1}})

= \det \exp\begin{bmatrix}

\lambda\_ 1 & \* & \cdots & \* \\

0 & \lambda\_2 & \* & \vdots \\

0 & \cdots & 0 & \lambda\_n \end{bmatrix}

= \det \begin{bmatrix} e^{\lambda\_ 1 }& \*& \cdots& \* \\

 0& e^{\lambda\_2} & \* & \vdots\\

0 & \cdots & 0 & e^{\lambda\_n} \end{bmatrix}

 = e^{\sum \lambda\_i}

und andererseits

e^{\tr A} = e^{\tr BAB^{-1}}

= \exp \tr \begin{bmatrix} e^{\lambda\_ 1 }& \*& \cdots& \* \\

 0& e^{\lambda\_2} & \* & \vdots\\

0 & \cdots & 0 & e^{\lambda\_n} \end{bmatrix}

= e^{\sum \lambda\_i}.

</beweis>

<definition> Definition 8.1.4

Es sei

\GL^+(n,\R) := \{B\in{\mathrm{End}}(\R^n): \det B>0\}.

</definition>

\GL^+(n,\R) ist eine offene Untergruppe von {\mathrm{End}}(\R^n) und nach obigem ist e: {\mathrm{End}}(\R^n)\rightarrow \GL^+(n,\R).

<lemma> Lemma 8.1.5

Das Differential der Abbildung e: {\mathrm{End}}(\R^n)\rightarrow \GL^+(n,\R) im

Nullpunkt ist De(0)={\mathrm{Id}}: {\mathrm{End}}(\R^n)\rightarrow {\mathrm{End}}(\R^n).

</lemma>

<beweis>

Man berechnet [De(0)](B) = \nabla\_B e(0) = \lim\_{t\rightarrow 0} \frac{e^{tB} - {\mathrm{Id}}}{t}

= \lim\_{t\rightarrow 0} \frac{tB+ t^2B^2/2+t^3(\ldots)}{t} = B.

</beweis>

Aus dem Satz über die Umkehrabbildung folgt nun unmittelbar folgendes

<korollar> Korollar 8.1.6

Es existieren offene Teilmengen 0\in U \subset End(\R^n) und {\mathrm{Id}}\in V\subset \GL^+(n,\R) derart, dass e: U\rightarrow V ein Diffeomorphismus ist.

<\korollar>

\qed

<bemerkung> Bemerkung 8.1.7

Es ist falsch, dass e global ein Diffeomorphismus ist. Tatsächlich ist die Exponentialabbildung von Matrizen nicht surjektiv. Zum Beispiel liegt

B=\begin{bmatrix} -1 & 0 \ 1 & -1 \end{bmatrix}

nicht im Bild von e, denn eine Matrix A mit e^A=B müsste spurfrei sein, also die Gestalt

A=\begin{bmatrix} a & b \ c & -a \end{bmatrix} haben,

und man überprüft, dass A^2=-\det A\cdot {\mathrm{Id}} gilt. Damit ist

e^A= C\_1\cdot {\mathrm{Id}}+ C\_2\cdot A. Die Konstante C\_2 kann nicht Null sein, denn e^A ist kein Vielfaches der Identität. Dann aber folgt b=0 und a=0. Demzufolge ist A nilpotent und kann nicht B als Exponential haben.

</bemerkung>

## 8.2. Homogene lineare Differentialgleichungssysteme erster Ordnung mit konstanten Koeffizienten

Wir beginnen nun mit der systematischen Behandlung von linearen Differentialgleichungen. Hierzu ordnen wir jedem Paar (A,x\_0) bestehend aus einer linearen Abbildung

A\in{\mathrm{End}}(\R^n) und einem Vektor x\_0\in\R^n das System linearer Differentialgleichungen erster Ordnung

\label{eq:21}

\dot{x} = Ax, x = x(t) = (x\_1(t),\ldots, x\_n(t))

mit der Anfangsbedingung x(t\_0)=x\_0 zu. Gewisse Lösungen der Differentialgleichung kann man unmittelbar angeben. Ist etwa v ein Eigenvektor von A zum Eigenwert \lambda, so ist x(t)=e^{\lambda t}v eine Lösung von \dot x=Ax. Allerdings erfüllt sie nicht die geforderte Anfangsbedingung. Nimmt man diese hinzu, so ist die Lösung eindeutig und lässt sich mit Hilfe der Exponentialabbildung von Matrizen beschreiben. Tatsächlich gilt folgender

<satz> Satz 8.2.1

Die einzige Lösung des Anfangswertproblems \eqref{eq:21} ist

\phi : \R\rightarrow \R^n, \phi (t) = e^{(t-t\_0)A} \cdot x\_0.

</satz>

<beweis>

Wir zeigen zunächst, dass \phi (t) eine Lösung ist. So ist

\phi (t\_0) = e^0\cdot x\_0 = {\mathrm{Id}}(x\_0) = x\_0.

Andererseits berechnet sich die Ableitung von \phi gemäß Lemma 8.1.5 zu

\dot{\phi }(t) = \frac{d}{dt} e^{(t-t\_0)A} \cdot x\_0

= \lim\_{h\rightarrow 0} \frac{e^{(t+h-t\_0)A} - e^{(t-t\_0)A}}{h} (x\_0)

= \lim\_{h\rightarrow 0} \left[ \frac{e^{hA} - {\mathrm{Id}}}{h}\right] e^{(t-t\_0)A} x\_0

= A e^{(t-t\_0)A} x\_0 = A\phi (t).

Für den Nachweis der Eindeutigkeit sei \psi(t) irgendeine Lösung und betrachte e^{-(t-t\_0)A} \cdot \psi(t). Eine analoge Rechnung zeigt dann

\frac{d}{dt} [e^{-(t-t\_0)A} \cdot \psi(t)] = -Ae^{-(t-t\_0)A}\psi(t)+ e^{-(t-t\_0)A} \dot\psi(t)

= -Ae^{-(t-t\_0)A}\psi(t)+e^{-(t-t\_0)A} A\psi(t) = 0,

weil A mit e^{-(t-t\_0)A} kommutiert. Damit ist e^{-(t-t\_0)A} \cdot \psi(t)=c konstant, und der Wert der Konstanten berechnet sich durch Betrachten des Anfangswerts zu \psi(t\_0)=x\_0=c. Insgesamt ist also \psi(t)= e^{(t-t\_0)A} x\_0 und die Lösung eindeutig.

</beweis>

Im Allgemeinen ist für eine Matrix A deren Exponential nur schwer zu berechnen, so dass Aussagen über das Verhalten der Lösung auf diese Weise kaum zu erzielen sind. Im Folgenden soll stattdessen ein Algorithmus zur allgemeinen Bestimmung von e^{A} vorgestellt werden.

#### 1.Schritt: Berechnung der Jordan'schen Normalform

Seien \lambda\_1,\ldots, \lambda\_k die reellen, \mu\_1=a\_1\pm ib\_1, \ldots, \mu\_l=a\_l \pm i b\_l die komplexen Eigenwerte \footnote{Beachte, dass mit \mu auch \bar \mu ein Eigenwert einer reellen Matrix ist.} von A entsprechend ihrer Vielfachheit und \alpha\_1,\ldots, \alpha\_k beziehungsweise \beta\_1,\ldots,\beta\_l die Größen ihrer Jordan-Kästchen. Nach der allgemeinen Jordan-Theorie existiert dann eine Basis von \R^n aus Eigenvektoren und Hauptvektoren, in der A die Jordan'sche Normalform

A = \begin{bmatrix} D(\lambda\_1) & & & & & \\

& \ddots & & & & \\

& & D(\lambda\_k) & & & \\

& & & D(\mu\_1) & & \\

& & & & \ddots & \\

& & & & & D(\mu\_{l}) \end{bmatrix}

hat, wobei die einzelnen Jordan-Blöcke die Gestalt

D(\lambda\_i) = \begin{bmatrix} \lambda\_i & & & & \\

1 & \lambda\_i & &0 & \\

& 1 & \ddots & & \\

0 & & 1 & \lambda\_i & \end{bmatrix},

D(\mu\_j) = \begin{bmatrix} a\_j & b\_j & & & & & \\

-b\_j & a\_j & & & & & \\

1 & 0 & a\_j& b\_j& & & \\

0 & 1 & -b\_j & a\_j & & & \\

 & & 1& 0& \ddots& & \\

 & & 0 & 1& & & \\

 & & & & & a\_j & b\_j\\

 & & & & \ddots & -b\_j& a\_j \end{bmatrix}

haben. Dann gilt in dieser Basis

\label{eq:22}

e^{tA} = \begin{bmatrix} e^{tD(\lambda\_1)} & & & & & \\

& \ddots & & & & \\

& & e^{tD(\lambda\_k)} & & & \\

& & & e^{tD(\mu\_1)} & & \\

& & & & \ddots & \\

& & & & & e^{tD(\mu\_{l})}

\end{bmatrix},

wodurch das Problem der Berechnung von e^{tA} auf die Berechnung des Exponentials von Jordan-Blöcken reduziert.

#### 2.Schritt: Berechnung von e^{tD(\lambda\_i)}

Nach obigem zerfällt D(\lambda\_i) in einen nilpotenten und einen zur Identität proportionalen Anteil gemäß

D(\lambda\_i) = \lambda\_i\cdot {\mathrm{Id}} +

\begin{bmatrix} 0 & & & & \\

1 & 0 & &0 & \\

& 1 & \ddots & & \\

0 & & 1 & 0 & \end{bmatrix}

=: \lambda\_i\cdot {\mathrm{Id}} + N\_i.

Weil diese Matrizen vertauschen, gilt e^{t D(\lambda\_i)} = e^{t\lambda\_i} e^{tN\_i}.

Für die nilpotente Matrix N\_i gilt N\_i^{\alpha\_i}=0. Die Exponentialreihe bricht also ab und man erhält

e^{tN\_i} = \begin{bmatrix} 1 & & & & \\

t & 1 & & 0 & \\

\frac{t^2}{2} & t & 1 & & \\

\vdots & \ddots & \ddots& \ddots & \\

\frac{t^{\alpha\_i-1}}{(\alpha\_i-1)!} & \cdots & \frac{t^2}{2} & t & 1\\

\end{bmatrix}.

Für den ersten Typ von Jordan-Blöcken läßt sich e^{t D(\lambda\_i)} also vollständig berechnen.

#### 3.Schritt: Berechnung von e^{tD(\mu\_j)}

Wir vereinbaren für die in D(\mu\_j) auftretenden Blöcke die Abkürzungen

E = \begin{bmatrix} 1 & 0 \ 0 & 1\end{bmatrix},

\theta\_j = \begin{bmatrix} a\_j & b\_j \ -b\_j & a\_j\end{bmatrix},

F = \begin{bmatrix} 0 & 1 \ -1 & 0\end{bmatrix}.

Auch hier zerlegen wir D(\mu\_j) in seinen nilpotenten Anteil und ein Vielfaches der Identität und erhalten

D(\mu\_j) = \begin{bmatrix} \theta\_j & & & \\

E & \theta\_j & 0& \\

 & \ddots & \ddots & \\

0 & & E & \theta\_j \end{bmatrix}

= \begin{bmatrix} \theta\_j & & & \\

0 & \theta\_j & 0& \\

 & \ddots & \ddots & \\

0 & & 0 & \theta\_j \end{bmatrix}

+ \begin{bmatrix} 0 & & & \\

E & 0 & 0& \\

 & \ddots & \ddots & \\

0 & & E & 0 \end{bmatrix} =: \Theta\_j+ M\_j.

Man überlegt sich alsdann, dass auch diese beiden Matrizen miteinander vertauschen. Zerlegt man \Theta\_j nochmals innerhalb jedes Blocks gemäß

\theta\_j = a\_j\cdot E+ \begin{bmatrix} 0 & b\_j \ -b\_j & 0\end{bmatrix}

=: a\_j\cdot E + b\_j\cdot F,

so folgt e^{t\theta\_j} = e^{t a\_j} e^{t b\_j F}. Zur weiteren Berechnung beachten wir, dass für alle k\in\N

F^{4k} = E, F^{4k+1} = F, F^{4k+2} = -E,

F^{4k+3} = -F,

so dass

e^{tb\_j F} = \left[1-\frac{(t b\_j)^2}{2}+\frac{(t b\_j)^4}{4!}\mp\ldots \right]

\cdot E+ \left[tb\_j-\frac{(tb\_j)^3}{3!}+\frac{(tb\_j)^5}{5!}\mp\ldots\right]\cdot F

= \begin{bmatrix} \cos (tb\_j) & \sin(tb\_j) \ -\sin(tb\_j) & \cos(t b\_j) \end{bmatrix}.

Insgesamt ist somit

e^{tD(\mu\_j)} = e^{t\Theta\_j}e^{tM\_j}

= e^{t a\_j} \begin{bmatrix}

\cos(tb\_j) & \sin(tb\_j) & & & 0 \\

-\sin(tb\_j) & \cos(tb\_j) & & & \\

& & \ddots & & \\

& & & \cos(tb\_j)& \sin(tb\_j) \\

 0& & & -\sin(tb\_j) & \cos(tb\_j) \end{bmatrix}

\cdot \begin{bmatrix} E & & & & \\

tE & E & & 0 & \\

\frac{t^2}{2}E & tE & E & & \\

\vdots & \ddots & \ddots& \ddots & \\

\frac{t^{\beta\_i-1}}{(\beta\_i-1)!} E& \cdots & \frac{t^2}{2}E & tE & E \end{bmatrix}.

Damit ist die Berechnung von e^{tD(\mu\_j)} und damit e^{tA} im Wesentlichen abgeschlossen. Wir besprechen im Folgenden einige qualitative Folgerungen des vorangehenden Satzes.

<korollar> Korollar 8.2.2

Sei A: \R^n\rightarrow \R^n eine lineare Abbildung mit n verschiedenen reellen Eigenwerten

\lambda\_1,\ldots,\lambda\_n und n zugehörigen Eigenvektoren v\_1,\ldots, v\_n. Die allgemeine Lösung der Differentialgleichung \dot x(t)=Ax(t) ist dann gegeben durch

x(t) = c\_1 e^{t\lambda\_1}v\_1+\ldots + c\_n e^{t\lambda\_n}v\_n.

<\korollar>

\qed

<beispiel> Beispiel 8.2.3

Die Matrix A=\begin{bmatrix}3 & 2 \ 1 & 4 \end{bmatrix} definiert das System von Differentialgleichungen

\label{eq:23}

\dot x\_1 (t) = 3 x\_1(t)+ 2 x\_2(t),

\dot x\_2(t) = x\_1(t)+4 x\_2(t).

Das charakteristische Polynom faktorisiert über \R als

\det (A-\lambda )

= \det \begin{bmatrix}3-\lambda & 2 \ 1 & 4-\lambda \end{bmatrix}

= \lambda^2 -7\lambda + 10 = (\lambda-5)(\lambda -2).

Die Eigenwerte von A lauten also \lambda\_1=5, \lambda\_2=2 und die zugehörigen Eigenvektoren sind die Lösungen v\_1, v\_2 von (A-\lambda\_i E)v\_i=0, welche sich zu

(A-5E)v\_1 = 0, v\_1 = \begin{bmatrix}1\\1\end{bmatrix} und

(A-2E)v\_2 = 0, v\_2 = \begin{bmatrix}2\\-1\end{bmatrix}

berechnen. Mit dem Basiswechsel P=(v\_1,v\_2)^{-1}

=\begin{bmatrix}1 & 2 \\1 & -1\end{bmatrix}^{-1}

=-\frac{1}{3}\begin{bmatrix}-1 & -2\\-1& 1\end{bmatrix}

ist

PAP^{-1} = A' = \begin{bmatrix}5 & 0 \\0 & 2\end{bmatrix}

diagonal und die allgemeine Lösung von \eqref{eq:23} lautet

\phi (t) = \begin{bmatrix}x\_1(t)\ x\_2(t)\end{bmatrix}

= \sum\_{i=1}^2 c\_i e^{t\lambda\_i} v\_i

= c\_1 e^{5t}\begin{bmatrix}1\\1\end{bmatrix}

+ c\_2 e^{2t}\begin{bmatrix}2\\-1\end{bmatrix}.

</beispiel>

<korollar> 8.2.4

Sei A: \R^{2n}\rightarrow \R^{2n} linear mit 2n verschiedenen, komplexen, paarweise konjugierten Eigenwerten \mu\_j=a\_j\pm i b\_j, j=1,\dots,n. Seien weiter 2n Vektoren u\_i, v\_i, i=1,\ldots, n, gegeben mit

(1) Jeder der Unterräume U\_i:=\mathrm{Span} ( u\_i, v\_i) ist invariant unter A,

(2) A wirkt auf U\_i durch Multiplikation mit \begin{bmatrix} a\_i & b\_i\ - b\_i & a\_i \end{bmatrix}.

Die allgemeine Lösung der Differentialgleichung \dot x(t)=Ax(t) ist dann gegeben durch

\phi (t) = \sum\_{i=1}^n c\_i e^{t a\_i} [\cos (tb\_i) u\_i - \sin (tb\_i) v\_i]

+ \sum\_{i=1}^n d\_i e^{t a\_i} [\sin (tb\_i) u\_i + \cos (tb\_i) v\_i]\\

= \sum\_{i=1}^n e^{t a\_i}

\bigg[ [c\_i\cos (tb\_i)+ d\_i \sin (tb\_i)]u\_i + [d\_i \cos (tb\_i) - c\_i \sin (tb\_i)] v\_i\bigg].

<\korollar>

\qed

<bemerkung> Bemerkung 8.2.5

In den Anwendungen bestimmt man die Vektoren u\_i, v\_i aus den komplexen Eigenvektoren mit Hilfe eines Basiswechsels, der dem gleicht, der beim Übergang von der komplexen zur reellen Jordan'schen Normalform verwendet wird. Alternativ kann man auch vollständig mit der komplexen Jordan'schen Normalform arbeiten und verwendet zum Schluss folgende Beobachtung: Ist \phi (t) eine komplexwertige Lösung der ursprünglichen Differentialgleichung, dann sind auch ihr Real- und Imaginärteil Lösungen derselben, was ebenfalls zu einer reellen Lösungsbasis führt.

</bemerkung>

<beispiel> Beispiel 8.2.6

Die Matrix A=\begin{bmatrix}1 & 1 \ -1 & 1 \end{bmatrix} definiert das System von Differentialgleichungen

\dot x\_1 (t) = x\_1(t)+ x\_2(t),

\dot x\_2(t) = -x\_1(t)+ x\_2(t).

Das charakteristische Polynom faktorisiert über \C als

\det (A-\lambda E) = \det \begin{bmatrix}1-\lambda & 1 \ -1 & 1-\lambda \end{bmatrix}

= \lambda^2 -2\lambda + 2 = (\lambda-1-i)(\lambda -1+i).

Die komplexen Eigenwerte von A sind somit 1\pm i, a=1, b=1, so dass die Matrix A bereits in reeller Jordan'scher Normalform ist. Als Vektoren u, v kann man deshalb die Standardbasis e\_1,e\_2 verwenden und die allgemeine Lösung lautet nach obigem Korollar

\phi (t) = ce^t(\cos t u - \sin t v) + de^t (\sin t u+\cos t v)

= ce^t\begin{bmatrix}\cos t\ -\sin t\end{bmatrix}+ de^t \begin{bmatrix}\sin t\ \cos t\end{bmatrix}.

Alternativ kann man nach obiger Bemerkung die beiden komplexen Eigenvektoren berechnen, welche

v\_1=\begin{bmatrix}1\ i\end{bmatrix}, v\_2=\begin{bmatrix}I \\ 1\end{bmatrix}

lauten. Damit gelangt man zu den komplexen Fundamentallösungen

\phi ^\C (t) = c\_1 e^{(1+i)t}\begin{bmatrix}1\ i\end{bmatrix}

+ c\_2e^{(1-i)t}\begin{bmatrix}i\ 1\end{bmatrix}.

Real- und Imaginärteile sind ebenfalls Lösungen und man erhält als reelle Lösung

\phi (t) = d\_1\begin{bmatrix}e^t\cos t \ -e^{t}\sin t\end{bmatrix}

+ d\_2 \begin{bmatrix}e^t\sin t \ e^{t}\cos t\end{bmatrix},

welche mit der zuvor gefundenen übereinstimmt.

</beispiel>

## 8.3 Die Lösung einer homogenen, linearen Differentialgleichung mit konstanten Koeffizienten

Wir kommen nun zur Bestimmung der allgemeinen Lösung einer homogenen, linearen Differentialgleichung k-ter Ordnung

\label{eq:24}

x^{(k)}(t)+ a\_1 x^{(k-1)}(t)+ \ldots + a\_{k-1}\dot x(t)+ a\_k x(t) = 0, a\_1,\ldots, a\_k\in\R,

durch Umformung in ein äquivalentes System von Differentialgleichungen erster Ordnung und Anwendung der soeben erarbeiteten Lösungstheorie für derartige Systeme.

Als charakteristische Gleichung dieser Differentialgleichung bezeichnet man die Polynomgleichung

\label{eq:25}

\chi(\lambda) = \lambda^k+ a\_1\lambda^{k-1}+\ldots + a\_{k-1}\lambda+ a\_k = 0.

Seien \lambda\_1,\ldots,\lambda\_i die reellen Lösungen der charakteristischen Gleichung mit Vielfachheiten m\_1,\ldots, m\_i, sowie

\mu\_1=a\_1+i b\_1,

\ldots, \mu\_j=a\_j+ib\_j,

\bar{\mu}\_1=a\_1-i b\_1, \ldots,

\bar{\mu}\_j=a\_j-ib\_j

die komplexen Lösungen mit Vielfachheiten n\_1,\ldots, n\_j.

<satz> Satz 8.3.1

Die allgemeine Lösung der Differentialgleichung \eqref{eq:24} ist von der Gestalt

x(t)& =& p\_1(t)e^{t\lambda\_1}+\cdots+p\_i(t)e^{t\lambda\_i}

+e^{ta\_1}[q\_1(t)\cos (b\_1 t)+ r\_1(t)\sin(b\_1t)]+ \cdots

+ e^{ta\_j}[q\_j(t)\cos (b\_j t)+ r\_j(t)\sin(b\_jt)],

wobei p\_1,\ldots,p\_i Polynome vom Grad echt kleiner als m\_1,\ldots, m\_i und q\_1, r\_1,\ldots, q\_j,r\_j Polynome vom Grad echt kleiner als n\_1,\ldots, n\_j sind.

</satz>

<beispiel> Beispiel 8.3.2

Die Differentialgleichung x^{(3)}(t)-x^{(2)}(t)+ \dot x (t)-x(t)=0 hat die charakteristische Gleichung

\chi(\lambda) = \lambda^3-\lambda^2+\lambda-1 = (\lambda-1)(\lambda-i)(\lambda+i).

Damit ist die allgemeine Lösung der Gestalt x(t) = ae^t+b\cos t+c\sin t.

</beispiel>

<beispiel> Beispiel 8.3.3

Die Differentialgleichung x^{(4)}(t)+2x^{(2)}(t)+ x(t)=0 mit der charakteristischen Gleichung

\chi(\lambda) = \lambda^4+2\lambda^2+1 = (\lambda^2+1)^2

hat i und -i als doppelte Eigenwerte. Damit ist die allgemeine Lösung x(t) = (at+b) \cos t+ (ct+d)\sin t.

</beispiel>

<beweis>[Beweis von Satz 8.3.1]

Die Differentialgleichung \eqref{eq:24} ist äquivalent zu dem Differentialgleichungssystem

\dot y\_0=y\_1, \dot y\_1=y\_2,\

\ldots, \dot y\_{k-2}=y\_{k-1},

\dot y\_{k-1}=-a\_1 y\_{k-1}-\ldots - a\_{k-1}y\_1 - a\_k y\_0.

In Matrixform geschrieben lautet dieses

\frac{d}{dt} \begin{bmatrix}y\_0\ y\_1\ y\_2 \\\vdots \ y\_{k-1}\end{bmatrix}

= \begin{bmatrix}0 & 1 & 0 & \ldots & 0\\

0 & 0 & 1 & & 0\\

 & & & \ddots & \\

 & & & & 1\\

-a\_k &-a\_{k-1} & & & -a\_1 \end{bmatrix}

\begin{bmatrix} y\_0 \\ y\_1 \\ y\_2 \\ \vdots \\ y\_{k-1} \end{bmatrix},

wobei die Matrix fortan mit A bezeichnet werde. Das charakteristische Polynom ist nun

\chi\_A(\lambda) = \det (\lambda E - A)

= \det \begin{bmatrix}\lambda & -1 & 0 & \ldots & 0 \\

0 & \lambda & -1 & & 0\\

 & & & \ddots & \\

 & & & & -1\\

a\_k &a\_{k-1} & & & a\_1+\lambda \end{bmatrix}

Wir entwickeln diese Determinante nach der letzten Zeile, wobei alternierend die Vorzeichen + und - zu wählen sind. Damit folgt

\pm \chi(\lambda) = a\_k (-1)^{k-1}- a\_{k-1} \lambda(-1)^{k-2}+

\ldots \pm a\_2\lambda^{k-2}(-1) \mp (a\_1+\lambda)\lambda^{k-1}\\

= a\_k+ a\_{k-1} \lambda+ a\_{k-2}\lambda^2 +\ldots

+ a\_2\lambda^{k-2} +a\_1\lambda^{k-1}+\lambda^k,

in Übereinstimmung mit \eqref{eq:25}. Man führt alsdann A in seine Jordan'sche Normalform über und berechnet wie im vorigen Abschnitt beschrieben in einer Basis aus Eigen- und Hauptvektoren das Exponential e^{tA}. Die allgemeine Lösung ist dann e^{tA} mal einer beliebigen Linearkombination der Eigen- und Hauptvektoren, und die uns interessierende Lösung x(t) die erste Komponente des so entstehenden Vektors. Die Aussage über die Polynome in t erhält man durch direkten Vergleich über die Berechnung von e^{tD(\lambda\_i)} und e^{tD(\mu\_j)}.

</beweis>

## 8.4 Homogene lineare Differentialgleichungen mit veränderlichen Koeffizienten

Als nächstes wenden wir uns der Untersuchung von linearen Differentialgleichungen mit veränderlichen Koeffizienten zu. Sei so zu jedem t\in (a,b) ein Endomorphismus A(t): \R^n\rightarrow \R^n gegeben derart, dass die Zuordnung t \longmapsto A(t) bezüglich der durch die Operatornorm definierten Metrik stetig ist. Wir betrachten dann das homogene lineare Differentialgleichungssystem

\label{eq:26}

\dot x(t) = A(t) x(t), x(t)\in\R^n.

Im kommenden Kapitel werden wir folgende Aussagen beweisen.

- Jede maximale Lösung von \eqref{eq:26} ist auf ganz ]a,b[ definiert.

- Sind x\_1(t), x\_2(t): ]a,b[\rightarrow \R^n zwei Lösungen von \eqref{eq:26} mit x\_1(t\_0)=x\_2(t\_0) für ein t\_0\in ]a,b[, so folgt x\_1(t) =x\_2(t) auf ganz ]a,b[; insbesondere ist jede Lösung x(t), die an einer Stelle t\_0 verschwindet, identisch null.

- Zu jedem t\_0\in ]a,b[ und x\_0\in\R^n existiert eine Lösung mit x(t\_0)=x\_0.

Damit ist die Menge V=\{ x: ]a,b[\rightarrow \R^n, x \text{ löst } \eqref{eq:26} \} ein n-dimensionaler Vektorraum.

<definition> Definition 8.4.1

Ein Fundamentalsystem von \eqref{eq:26} ist eine Basis \phi \_1,\ldots, \phi \_n des Lösungsraumes V. Eigenschaft (2) impliziert, dass für jeden einzelnen Wert von t die Fundamentallösungen \phi \_1(t),\ldots, \phi \_n(t) eine Basis des \R^n bilden. Ist \phi \_1,\ldots, \phi \_n ein Fundamentalsystem, so nennt man die Funktion W: ]a,b[\rightarrow \R W(t) := \det \big(\phi \_1(t),\ldots, \phi \_n(t) \big) die Wronski-Determinante des Fundamentalsystems. Nach dem soeben gesagten ist W(t)\neq 0 für alle t\in ]a,b[.

</definition>

<satz> Satz 8.4.2 Satz von Liouville

Für jedes fest gewählte t\_0\in]a,b[ gilt

W(t) = W(t\_0) \exp\left[\int\_{t\_0}^t \tr A(s) ds\right].

</satz>

<beweis>

Sei A(t)=\big[a\_{ij}(t)\big]\_{i,j=1}^n, \phi \_j(t)=(\phi ^1\_j(t),\ldots,\phi ^n\_j(t) ).

Dann ergibt Ableitung der Wronski-Determinante

W(t) = \det \big[\phi ^i\_j(t)\big]

= \det\begin{bmatrix}\phi ^1\_1(t)&\ldots & \phi ^1\_n(t)\\

\vdots & \ddots& \vdots\\

\phi ^n\_1(t)& \ldots & \phi ^n\_n(t)\end{bmatrix}

nach den Zeilen

\dot W(t) = \det \begin{bmatrix}\dot \phi ^1\_1(t)&\ldots & \dot\phi ^1\_n(t)\\

\vdots & \ddots& \vdots\\

\phi ^n\_1(t)& \ldots & \phi ^n\_n(t)\end{bmatrix} +\ldots +

\det \begin{bmatrix} \phi ^1\_1(t)&\ldots & \phi ^1\_n(t)\\

\vdots & \ddots& \vdots\\

\dot\phi ^n\_1(t)& \ldots & \dot\phi ^n\_n(t)\end{bmatrix}.

Weil jedes \phi \_j(t) eine Lösung der ursprünglichen Differentialgleichung ist, gilt \dot \phi ^i\_j = \sum\_k a\_{ik}\phi ^k\_j. Einsetzen liefert dann

\dot W(t) = \det \begin{bmatrix}\sum\_k a\_{1k}\phi ^k\_1(t)&\cdots & \sum\_k a\_{1k}\phi ^k\_n(t)\\

\vdots & \ddots& \vdots\\

\phi ^n\_1(t)& \ldots & \phi ^n\_n(t)\end{bmatrix}

+\ldots + \det \begin{bmatrix} \phi ^1\_1(t)&\ldots & \phi ^1\_n(t)\\

\vdots & \ddots& \vdots\\

\sum\_k a\_{nk}\phi ^k\_1(t)& \ldots & \sum\_k a\_{nk}\phi ^k\_n(t)\end{bmatrix} .

Nun ändert sich die Determinante einer Matrix nicht unter elementaren Transformationen, so dass

\dot W(t) = \det \begin{bmatrix}a\_{11}\phi ^1\_1(t)&\ldots & a\_{11}\phi ^1\_n(t)\\

\vdots & \ddots& \vdots\\

\phi ^n\_1(t)& \ldots & \phi ^n\_n(t)\end{bmatrix} +\cdots +

\det \begin{bmatrix} \phi ^1\_1(t)&\ldots & \phi ^1\_n(t)\\

\vdots & \ddots& \vdots\\

a\_{nn}\phi ^n\_1& \ldots & a\_{nn}\phi ^n\_n\end{bmatrix}

= (a\_{11}+a\_{22}+\cdots +a\_{nn})

\det \begin{bmatrix} \phi ^1\_1(t)&\ldots & \phi ^1\_n(t)\\

\vdots & \ddots& \vdots\\

\phi ^n\_1(t)& \ldots & \phi ^n\_n(t)\end{bmatrix} = \tr A(t) \cdot W(t).

Weil W(t)\neq 0 für alle t ist, impliziert diese Differentialgleichung sofort

W(t)=c \exp\left[\int\_{t\_0}^t \tr A(s) ds\right].

Der Wert der Konstanten c folgt dann durch Auswerten bei t=t\_0.

</beweis>

Wir wenden jetzt das bisher gesagte auf die allgemeine lineare homogene Differentialgleichung k-ter Ordnung

\label{eq:27}

x^{(k)}(t)+ a\_1(t)x^{(k-1)}(t)+\ldots + a\_{k-1}\dot x(t)+a\_k(t) x(t) = 0

an, wobei a\_i(t)\in C^0(]a,b[) stetige Funktionen sind. Wir wissen bereits, dass diese Gleichung äquivalent ist zum System

\frac{d}{dt}\begin{bmatrix}y\_0\ y\_1\ y\_2 \\\vdots \ y\_{k-1}\end{bmatrix}

= \begin{bmatrix}0 & 1 & 0 & \ldots & 0\\

0 & 0 & 1 & & 0\\

 & & & \ddots & \\

 & & & & 1\\

-a\_k(t) &-a\_{k-1}(t) & & & -a\_1(t)

\end{bmatrix}

\begin{bmatrix}y\_0\ y\_1\ y\_2 \\\vdots \ y\_{k-1}\end{bmatrix} ,

wobei die entsprechende Matrix A stetige Funktionen als Einträge hat. Die Gleichung ist damit vom Typ \eqref{eq:26} und der entsprechende Lösungsraum k-dimensional. Insbesondere werden wir unter einem Fundamentalsystem der Differentialgleichung \eqref{eq:27} eine Basis x\_1(t),\ldots, x\_k(t) dieses k-dimensionalen Vektorraums verstehen, und als Wronski-Determinante der ursprünglichen Gleichung die Funktion

W(t) = \det\begin{bmatrix}x\_1(t) & \ldots & x\_k(t)\\

\dot x\_1(t) & \ldots & \dot x\_k(t)\\

\vdots & \ddots & \vdots\\

x\_1^{(k-1)}(t)& \ldots & x\_k^{(k-1)}(t) \end{bmatrix}

bezeichnen. Nach dem Satz von Liouville gilt in diesem Fall wegen \tr A=-a\_1

W(t) = W(t\_0)\exp \left[-\int\_{t\_0}^{t}a\_1(s) ds\right].

<beispiel> Beispiel 8.4.3

Wir betrachten auf ]0,\infty[ die Differentialgleichung

\ddot x(t)- \frac{2}{t}\dot x(t) + \bigg[1+\frac{2}{t^2} \bigg] x(t) = 0.

Eine allgemeine Methode zur Berechung der Lösungen existiert nicht, aber wir wissen, dass der Lösungsraum 2-dimensional ist. Nun verifiziert man, dass x\_1(t) = t\cos t, x\_2(t) = t \sin t

Lösungen der betrachteten Differentialgleichung sind. Tatsächlich ist

\dot x\_1 = \cos t - t\sin t,

\ddot x\_1 = -2\sin t - t\cos t,

also

\ddot x\_1- \frac{2}{t}\dot x\_1 + \bigg[1+\frac{2}{t^2} \bigg] x\_1

= -2\sin t -t\cos t -\frac{2}{t}(\cos t - t\sin t)+ \bigg[1+\frac{2}{t^2} \bigg] t\cos t = 0.

Folglich ist jede Lösung der Differentialgleichung auf ]0,\infty[ eine Linearkombination dieser beiden Lösungen. Wir berechnen noch dieWronski-Determinante, einmal direkt aus der Kenntnis der Lösungen,

W(t) = \det \begin{bmatrix}t\cos t & t\sin t \\

\cos t -t\sin t & \sin t + t\cos t\end{bmatrix} = t^2,

und einmal mit Hilfe des Satzes von Liouville ohne Kenntnis der Lösungen für t\_0=1 gemäß

-\int\_1^t a\_1(s) ds = -\int\_1^t \left[-\frac{2}{s}\right] ds = 2 [\ln s]\_1^t = 2\ln t.

Damit ist

W(t) = W(1) e^{2\ln t} = W(1) t^2.

Beachte, dass die Differentialgleichung linear ist, so dass jedes Vielfache einer Lösung wieder eine Lösung ist. Demnach kann W(t) nur bis auf eine multiplikative Konstante bestimmt werden.

</beispiel>

<beispiel> Beispiel 8.4.4 Jacobi'sche Differentialgleichung

Wir betrachten die elliptischen Integrale

E(t) = \int\_0^{\pi/2} \sqrt{1-t^2\sin^2 x} dx,

F(t) = \int\_0^{\pi/2} \frac{1}{\sqrt{1-t^2\sin^2 x}} dx,

vergleiche Beispiel 6.1.5. Wie sich zeigen lässt, ist E(t) eine Lösung der Differentialgleichung

\ddot x(t) + \frac{\dot x}{t} +\frac{x}{1-t^2} = 0 (Jacobi'sche Differentialgleichung)

auf ]0,1[. Nach den vorherigen Betrachtungen ist der Lösungsraum insgesamt zweidimensional. Ist somit G(t) eine weitere Lösung, so impliziert der Satz von Liouville, dass

\det \begin{bmatrix} E(t)& G(t)\ \dot E(t) & \dot G(t) \end{bmatrix}

= C \exp\left[ -\int \frac{1}{s} ds\right] = Ce^{-\ln t} = \frac{C}{t}.

Wir normieren G(t) so, dass C=1 gilt. Dann erfüllt G(t) die Differentialgleichung erster Ordnung E(t)\dot G(t)- G(t)\dot E(t) = \frac{1}{t}. Löst man nun diese Gleichung mittels Variation der Konstanten, so führt der Ansatz G(t)=c(t)E(t) auf die Differentialgleichung \dot c(t)=1/t E^2(t). Demzufolge ist

G(t) = E(t)\int \frac{1}{tE^2(t)} dt.

Damit ist auch die zweite Fundamentallösung bestimmt, wobei eine geschlossene Formel für G(t) ebenso unerreichbar ist wie für E(t).

</beispiel>

## 8.5 Inhomogene lineare Differentialgleichungen mit veränderlichen Koeffizienten

Abschließend behandeln wir schließlich noch inhomogene, lineare Differentialgleichungen mit veränderlichen Koeffizienten. Seien so a\_1(t),\ldots, a\_k(t), f(t) stetige Funktionen auf einem offenen Intervall ]a,b[ und betrachte auf diesem Intervall die allgemeine inhomogene Differentialgleichung k-ter Ordnung

\label{eq:28}

x^{(k)}(t)+ a\_1(t)x^{(k-1)}(t)+\ldots + a\_k(t)x(t) = f(t),

sowie die entsprechende homogene Differentialgleichung

\label{eq:29}

x^{(k)}(t)+ a\_1(t)x^{(k-1)}(t)+\ldots + a\_k(t)x(t) = 0.

Wir wissen bereits, dass der Lösungsraum von \eqref{eq:29} ein k-dimensionaler, reeller Vektorraum ist. Für das Lösen des inhomogenen Problems ist folgender Satz von entscheidender Bedeutung.

<satz> Satz 8.5.1 Überlagerungssatz

(1) Sind x\_1(t), x\_2(t) Lösungen der inhomogenen Gleichung \eqref{eq:28}, so ist ihre Differenz x\_1(t)-x\_2(t) eine Lösung der homogenen Gleichung \eqref{eq:29}.

(2) Die allgemeine Lösung der inhomogenen Gleichung \eqref{eq:28} entsteht aus der Summe der allgemeinen Lösung der homogenen Gleichung \eqref{eq:29} und einer beliebigen speziellen Lösung der inhomogenen Gleichung.

</satz>

<beweis>

(1) Weil die Gleichung \eqref{eq:28} linear ist, folgt aus

x\_1^{(k)}(t)+\ldots + a\_k(t)x\_1(t) = f(t),

x\_2^{(k)}(t)+\ldots + a\_k(t)x\_2(t) = f(t),

durch Subtraktion sofort

(x\_1-x\_2)^{(k)}(t)+\ldots + a\_k(t)(x\_1(t)-x\_2(t)) = 0.

(2) Es durchlaufe x\_1(t) nun alle Lösungen der inhomogenen Gleichung \eqref{eq:28}. Ziehen wir von der Gesamtheit all dieser Lösungen, also von der allgemeinen Lösung x\_1(t), eine beliebige spezielle Lösung x\_2(t) der inhomogenen Gleichung ab, so erhalten wir die allgemeine Lösung z(t) der zugehörigen homogenen Gleichung. Denn ist \tilde{z}(t) irgendeine beliebige Lösung der homogenen Gleichung, so ist \tilde{z}(t)+ x\_2(t) eine Lösung der inhomogenen Gleichung, also unter den Funktionen x\_1(t) enthalten. Damit ist auch \tilde{z}(t) unter den Funktionen z(t)=x\_1(t)-x\_2(t) enthalten.

</beweis>

Das Lösungsverfahren für die inhomogene Differentialgleichung \eqref{eq:28} lautet somit wie folgt.

1. Bestimme ein Fundamentalsystem \phi \_1(t),\ldots, \phi \_k(t) der homogenen Gleichung \eqref{eq:29}.

2. Bestimme mittels Variation der Konstanten eine spezielle Lösung

x\_2(t) = c\_1(t)\phi \_1(t)+\ldots+ c\_k(t)\phi \_k(t)

der inhomogenen Differentialgleichung \eqref{eq:28}.

3. Die allgemeine Lösung der inhomogenen Differentialgleichung \eqref{eq:28} lautet dann

x\_1(t) = [c\_1(t)+ d\_1]\phi \_1(t)+\ldots+ [c\_k(t)+d\_k]\phi \_k(t), d\_1,\ldots, d\_k\in\R.

Kapitel 9: Allgemeine Existenz- und Eindeutigkeitsaussagen für Differentialgleichungen

## 9.1 Der Satz von Picard-Lindelöf

Wir wenden uns in diesem Kapitel der Herleitung von hinreichenden Voraussetzungen für die Existenz und Eindeutigkeit der Lösung einer Differentialgleichung zu. Dies ist der Inhalt des Satzes von Picard-Lindelöf. Sei im folgenden U \subset \R^n \times \R eine offene Teilmenge und F : U \rightarrow \R^n eine Abbildung.

<definition> Definition 9.1.1

F erfüllt die Lipschitz-Bedingung bezüglich der x-Variablen auf einer Teilmenge A\subset U, falls eine Konstante K = K(A) mit

\|F(x\_1, t) - F(x\_2, t)\| \leq K \|x\_1 - x\_2\|

für alle Punkte (x\_1,t), (x\_2, t) \in A existiert. Man sagt auch, dass F auf A Lipschitz-stetig ist. Die Zahl K heißt Lipschitz-Konstante.

</definition>

Ist F Lipschitz-stetig, so muss also bei beliebigem, festem t die Sekantensteigung \Delta F/\Delta x durch K beschränkt sein.

<Abbildung>

Ein Bild des Graphs einer Funktion. Auf der x-Achse sind x\_1 und x\_2 eingezeichnet. Die Strecke zwischen x\_1 und x\_2 hat die Länge \Delta x. Außerdem sind die Schnittpunkte der Vertikalen durch x\_1 und x\_2 mit dem Graphen von F markiert. Der vertikale Unterschied zwischen diesen beiden Punkten ist \Delta F. Das Steigungsdreieck ist eingezeichnet.

</Abbildung>

Dieser Begriff spielt in der Theorie der gewöhnlichen Differentialgleichungen eine zentrale Rolle. Er gestattet es, den Banach'schen Fixpunktsatz auf einen geeigneten Funktionenraum anzuwenden, um die Existenz und Eindeutigkeit der Lösung von Differentialgleichungen zu beweisen.

<satz> Satz 9.1.2

Ist F : U \rightarrow \R^n der Klasse C^1 und W \times [a,b] \subset U, W \subset \R^n ein Quader, so erfüllt F die Lipschitz-Bedingung auf W\times [a,b] bezüglich der Variablen x\in W.

</satz>

<beweis>

Sind x\_1, x\_2 \in W und t\in [a,b], so folgt mittels des Mittelwertsatzes F^i (x\_1, t) - F^i (x\_2, t) = \sum^n\_{j=1} \frac{\partial F^i}{\partial x^j} (x^\*\_i, t)(x^j\_1 - x^j\_2),

 wobei x^\*\_i\in W ein Punkt auf der Strecke [x\_1, x\_2] \subset W ist. Daraus ergibt sich

\|F (x\_1, t) & - & F(x\_2, t)\|^2 = \sum^n\_{i=1} |F^i (x\_1, t) - F^i (x\_2,t)|^2

\leq \sum^n\_{i=1} \sum^n\_{j=1} \big| \frac{\partial F^i}{\partial x^j}

(x^\*\_i, t)\big|^2 |x^j\_1 - x^j\_2|^2

\leq \underbrace{n\cdot \max\limits\_{x^\* \in W\atop t\in [a,b]}

\max\limits\_{1\leq i \leq n\atop 1\leq j \leq n}

 \Big|\frac{\partial F^i}{\partial x^j} (x^\*\_i, t)\Big|} {:= K} \sum^n\_{j=1} |x^j\_1 -x^j\_2|^2

= K \|x\_1 - x\_2\|^2.

</beweis>

<beispiel> Beispiel 9.1.3

Die Funktion f: \R\rightarrow \R, f(x)=|x| ist Lipschitz-stetig mit Lipschitz-Konstante K=1, denn \big| |x| - |y| \big| \leq |x-y|. Aber natürlich ist sie nicht stetig differenzierbar.

</beispiel>

<beispiel> Beispiel 9.1.4

Die Funktion f(x) = \sqrt{|x|} auf [-1,1] ist stetig, aber nicht Lipschitz-stetig, da f infolge von

\sup\limits\_{x\in [-1,1]} \frac{|f(x)-f(0)|}{|x-0|} = \sup\limits\_{x\in [-1,1]} \frac{1}{\sqrt{|x|}} = \infty keine beschränkte Sekantensteigung hat.

</beispiel>

<satz> Satz 9.1.5 Satz von Picard-Lindelöf -- vereinfachte Version

Sei (x\_0, t\_0)\in U \subset \R^n\times \R, F : U \rightarrow \R^n stetig und a,b > 0 positive Zahlen mit \mathscr{B}:=\overline{K (x\_0, b)}\times [t\_0 - a, t\_0 +a]\subset U.

Angenommen, F ist Lipschitz-stetig auf \mathscr{B} bezüglich der x-Variablen mit der Lipschitz-Konstanten K und setze

M = \sup \big\{ |F(x,t)| : (x,t) \in \mathscr{B}\big\}.

Sei weiterhin \alpha = \min (a, \frac{b}{M}, \frac{1}{2 K}). Dann hat \dot x = F (x,t), x(t\_0) = x\_0 genau eine Lösung

\phi \_{x\_0} : [t\_0 - \alpha, t\_0 +\alpha] \longrightarrow \overline{K (x\_0, b)} \subset \R^n.

</satz>

<beweis>

Wir betrachten den Raum X := C\big( [t\_0-\alpha, t\_0 + \alpha];

\overline{K (x\_0, b})\big) aller stetigen Abbildungen von [t\_0 - \alpha, t\_0 + \alpha] nach \overline{K (x\_0, b)}, versehen mit der Supremumsmetrik

 d(\phi \_1, \phi \_2) = \sup\limits\_{|t-t\_0| \leq \alpha} \|\phi \_1 (t) - \phi \_2 (t)\|.

Wie in der Analysis I gezeigt wurde, ist (X,d) ein vollständiger metrischer Raum. Sei desweiteren

Y := \{\phi \in X : \phi (t\_0) = x\_0\}.

Y\subset X ist ein abgeschlossener Teilraum, so dass auch (Y,d\_Y) ein vollständiger metrischer Raum ist. Wir definieren nun eine Abbildung A : Y \rightarrow Y durch die Integralgleichung

(A \phi )(t) := x\_0 + \int^t\_{t\_0} F(\phi (s), s) ds.

Tatsächlich ist A \phi \in Y für \phi \in Y. So ist (A\phi )(t\_0) = x\_0. Demnach bleibt zu zeigen, dass A\phi das Intervall [t\_0 -\alpha, t\_0 + \alpha] stetig nach \overline{K (x\_0, b)} abbildet. Aber

\|(A \phi )(t)- x\_0\| = \big\| \int^t\_{t\_0} F(\phi (s),s)ds\big\|

\leq M(t-t\_0) \leq M\cdot \alpha \leq M\cdot \frac{b}{M} \leq b

 und wir erhalten das Gewünschte, da offensichtlich A\phi(t) stetig in t ist. Als nächstes überlegt man sich, dass für alle \phi , \psi \in Y

 d(A \phi , A \psi) \leq \frac{1}{2} d(\phi , \psi)

 gilt. Tatsächlich impliziert die Lipschitz-Stetigkeit von F

\|(A \phi )(t) - (A \psi)(t)\| = \big\| \int^t\_{t\_0} F(\phi (s), s) - F(\psi (s), s) ds \big\|

\leq \int^t\_{t\_0} \|F(\phi (s), s) - F(\psi (s), s)\big\| ds

\leq \int^t\_{t\_0} K \| \phi (s) - \psi (s)\| ds

\leq K \cdot \sup \{ \| \phi (s) - \psi (s)\| : |s-t\_0| \leq \alpha \} \cdot \alpha .

Damit folgt wie behauptet

d (A \phi , A \psi) \leq K\cdot \alpha d(\phi , \psi)

\leq K \cdot \frac{1}{2K} d(\phi , \psi) \leq \frac{1}{2} d(\phi , \psi).

Damit ist A eine kontrahierende Abbildung und aus dem Banach'schen Fixpunktsatz der Analysis 1 ergibt sich jetzt, dass es genau eine Abbildung

\phi ^\* : [t\_0-\alpha, t\_0 + \alpha]\rightarrow \overline{K (x\_0, b)} in Y mit

\label{eq:30}

\phi ^\* (t) = (A \phi ^\*)(t) = x\_0 + \int^t\_{t\_0} F (\phi ^\* (t), s) ds

gibt. \phi ^\* ist zunächst nur stetig, aber aus \eqref{eq:30} folgt, dass \phi ^\* der Klasse C^1 ist und

\frac{d\phi ^\*}{dt} = F (\phi ^\* (t), t), \phi ^\* (t\_0) = x\_0

erfüllt. Also ist \phi ^\* eine Lösung des Problems.

Angenommen, es läge eine weitere Lösung \tilde \phi :

[t\_0 - \alpha, t\_0 + \alpha] \rightarrow \overline{K (x\_0, b)} des Problems

\frac{dx}{dt} = F(x,t) x(t\_0) = x\_0 vor. Integration von t\_0 bis t führt dann auf

\tilde \phi (t)-\tilde\phi (t\_0) = \int^t\_{t\_0}

\frac{d\tilde\phi }{dt} (s) ds = \int^t\_{t\_0} F(\tilde \phi(s), s) ds,

so dass abermals

\tilde\phi (t) = x\_0 + \int^t\_{t\_0} F(\tilde \phi (s), s) ds gilt. Also ist auch \tilde \phi ein Fixpunkt von A. Da A genau einen Fixpunkt besitzt, muss demzufolge \phi^\*=\tilde \phi gelten.

</beweis>

<bemerkung> Bemerkung 9.1.6

Der Fixpunktsatz von Banach wurde von ihm im Jahre 1920 bei der Untersuchung von Integralgleichungen aufgestellt. Er war Teil seiner Dissertation \emph{Sur les op\'erations dans les ensembles abstraits et leur application aux \'equations int\'egrales}. Hierbei zeigte er nicht nur die Existenz und Eindeutigkeit des Fixpunktes, sondern gab auch ein numerisches Verfahren zur näherungsweisen Berechnung desselben an.

</bemerkung>

<beispiel> Beispiel 9.5.7

Fehlende Lipschitz-Stetigkeit kann den Verlust der Eindeutigkeit der Lösung zur Folge haben kann. Betrachte so die Schar kubischer Parabeln

x(t) = (t-c)^3.

Differentiation ergibt \dot x = 3(t-c)^2 und Elimination des Parameters c führt auf die Differentialgleichung

\dot x(t) = 3 x^{2/3}(t).

Durch jeden Punkt der t-Achse gehen unendlich viele Lösungen; durch den Koordinatenursprung etwa x=t^3, x=0 und ferner alle C^2-Kurven

x(t) = \left\{\ba{ll} 0 & 0\leq x\leq a\ (t-a)^3 & x\geq a \ea\right. \text{a>0 beliebig},

die aus einem Stück der t-Achse und der anschließenden Hälfte der kubischen Parabel bestehen. Dies liegt daran, dass f(x,t)=3x^{2/3} in null keine beschränkte Ableitung hat und somit die Lipschitz-Bedingung bezüglich x nicht erfüllt.

</beispiel>

Es sei schließlich noch ohne Beweis bemerkt, dass der Wert der Zahl \alpha noch verbessert werden kann.

<satz> Satz 9.5.8 Picard-Lindelöf -- endgültige Fassung

Sei (x\_0, t\_0)\in U \subset \R^n \times \R, F : U \rightarrow \R^n stetig und seien positive Zahlen a, b > 0 mit \mathscr{B}:=\overline{K (x\_0, b)}\times [t\_0 - a, t\_0 + a] \subset U gegeben. Des Weiteren nehmen wir an, dass F auf \mathscr{B} Lipschitz-stetig bezüglich der x-Variablen mit der Konstanten K ist und setzen

M := \sup \{|F(x, t)| : (x,t)\in \mathscr{B}\},

sowie \alpha: = \min \big( a, \frac{b}{M}\big).

Dann hat das Problem \dot x = F (x,t), x(t\_0) = x\_0 genau eine Lösung

 \phi \_{x\_0} : [t\_0 - \alpha, t\_0 + \alpha] \longrightarrow \overline{K (x\_0, b)}\subset \R^n,

welche auf dem Intervall [t\_0 - \alpha, t\_0 + \alpha] definiert ist und Werte in K (x\_0, b) annimmt.

</satz>

\qed