

Analysis II

OSWALD RIEMENSCHNEIDER

Hamburg, 23. Dezember 2004

Vorbemerkung

Es handelt sich bei dem vorliegenden Text um die Fortführung meines Manuskripts *Analysis I*. Verweise auf Kapitel 1 bis 12 sind demnach dort zu suchen; ebenso beziehen sich alle mit einem „label“ [ab] versehenen Literaturangaben auf das Verzeichnis in Band 1. Weitere Literatur wird in der Regel im Text direkt angegeben. Es seien hier zwei weitere Werke neueren Erscheinungsdatums erwähnt, die für den Leserkreis von Interesse sein könnten. Sie verdeutlichen exemplarisch die Spannweite der an deutschen Universitäten möglichen Konzepte in der Anfängerausbildung:

- H. Amann, J. Escher: *Analysis I, II*. Birkhäuser: Basel–Boston–Berlin 1998/1999.
- E. Behrends: *Analysis*, Band 1. Ein Lernbuch für den sanften Wechsel von der Schule zur Uni. Vieweg & Sohn: Braunschweig–Wiesbaden 2003.

In die Literaturliste zu Band 1 haben keine Werke über die *Geschichte der Mathematik* Einlaß gefunden mit Ausnahme von [38], worin zumindest die historische Entwicklung besondere Berücksichtigung findet. Ähnliche Zielsetzungen hat auch das Buch

- H. Schröder: *Wege zur Analysis – genetisch - geometrisch - konstruktiv*. Springer: Berlin–Heidelberg–New York 2001.

Wir geben hier noch zwei Bände aus der Überfülle der bestehenden mathemathikhistorischen Literatur für die Zeitspanne 1700-1900 an; es kann in diesem Zusammenhang nicht oft genug betont werden, daß zum Studium moderner Forschungsergebnisse die (zumindest passive) Beherrschung des Englischen und nach Möglichkeit auch des Französischen eine unerläßliche Voraussetzung bildet. Zum Publizieren eigener Ergebnisse muß noch die aktive Handhabung des Englischen hinzukommen.

- Jean Dieudonné, ed.: *Abrégé d'histoire des mathématiques*.
- Morris Kline: *Mathematical thoughts from ancient to modern times*.

Denjenigen, die durch meinen Aufbau angeregt oder aus eigenem Antrieb sich für die Grundlegung der reellen Zahlen durch Axiome und ihre konkrete Konstruktion interessieren, sei ein neuerer Artikel empfohlen, der im *internet* leicht aufzufinden ist.

- N. A'Campo, A natural construction for the real numbers. arXiv:math. GN/0301015.

Für wertvolle Hinweise bei der Korrektur des hiermit vorliegenden zweiten Teils danke ich erneut Stephan Tolksdorf und auch Lilian Matthiesen. Es wird so bald wie möglich ein separater Index erstellt und angefügt werden.

Hamburg, 23. 12. 2004
Oswald Riemenschneider

Inhalt

Vorbemerkung	xxxiii
Inhalt	xxxv
Index (Stand: fehlt)	xxxvii

Teil II: Grundlagen der Analysis (Fortsetzung)

13 Intervalle und stetige Funktionen	225
Anhang: Stetige Abbildungen zwischen topologischen Räumen	240
14 Differenzierbare Funktionen	243
Anhang: Weitere Charakterisierungen der reellen Zahlen	291
15 Kompakte Räume und stetige Funktionen	293
Anhang: Weitere Charakterisierungen der reellen Zahlen	304
16 Der Hauptsatz der Differential- und Integralrechnung	307
Anhang: Tabelle der wichtigsten Stammfunktionen	337
17 Differenzierbare Abbildungen, Kettenregel und Taylor-Formel	341
Anhang: Die Äquivalenz von Normen auf endlich-dimensionalen Vektorräumen und Anwendungen	363
18 Das Darboux-, Riemann- und Lebesgue-Integral	369
Anhang: Die letzten 5 der 55 Charakterisierungen der reellen Zahlen	392
19 Die schwingende Saite und Fourier-Reihen	395
Anhang: Skalarprodukt- und Hilbert-Räume	410
20 Gewöhnliche Differentialgleichungen	425
Anhang: Der Existenzsatz von Peano	444
21 Gewöhnliche Differentialgleichungen höherer Ordnung und Differentialgleichungssysteme	449
Anhang: Die Erhaltungssätze der klassischen Mechanik und die Keplerschen Gesetze	458
22 Lineare Differentialgleichungen und -Gleichungssysteme	467
23 Das Newton-Verfahren und der Banachsche Fixpunktsatz	485
Anhang: Die Stetigkeit der allgemeinen Lösung von gewöhnlichen Differentialgleichungen ...	494

13 Intervalle und stetige Funktionen

In diesem Kapitel geht es im wesentlichen um den *Zwischenwertsatz* und seine logische Verknüpfung mit dem Begriff des *Zusammenhangs*. Wir werden dabei nicht nur wichtige Aussagen über die reellen Zahlen kennenlernen, sondern erneut auch deren Äquivalenz zum Axiom der Vollständigkeit. Auch wenn das DEDEKINDSche *Schnittaxiom* heutzutage keine entscheidende Rolle mehr spielt, findet es seinen angemessenen Platz in diesem Gedankenkreis. Es liegt in der Natur der Sache, daß wir neue, über die Konvergenz von Folgen hinausgehende (topologische) Begriffsbildungen einführen müssen. Hierzu gehört das Studium von *offenen* und *abgeschlossenen* Mengen, die wir schon früher kurz gestreift haben (Kapitel 7). Vor allem aber müssen wir über *zusammenhängende* Mengen sprechen und ihre Bedeutung für die möglichen Werte von reellwertigen stetigen Funktionen. Hierbei spielt eine entscheidende Rolle, daß der Körper der reellen Zahlen charakterisiert werden kann als der einzige angeordnete Körper, in dem die Intervalle - und nur diese - zusammenhängende Mengen sind. Ein weiteres grundlegendes Resultat schließt sich an: Stetige reellwertige Funktionen auf beschränkten abgeschlossenen Intervallen $[a, b]$ nehmen ihr Maximum und ihr Minimum an. Einige Beispiele und Anwendungen auf klassische Funktionen sowie zahlreiche neue „Vollständigkeitsaxiome“ beschließen das Kapitel.

Zuerst müssen wir den zentralen Begriff der *offenen* Mengen in einem angeordneten Körper \mathbb{K} rekapitulieren. (Zu der allgemeineren Situation von \mathbb{K} -metrischen bzw. *topologischen Räumen* X siehe Kapitel 7 und dessen Anhang).

Definition. Eine Menge $U \subset \mathbb{K}$ heißt *offen*, wenn es zu jedem $x_0 \in U$ ein $\varepsilon > 0$ gibt, so daß

$$U_\varepsilon(x_0) := \{x \in \mathbb{K} : |x - x_0| < \varepsilon\} \subset U.$$

Eine Menge $A \subset \mathbb{K}$ heißt *abgeschlossen*, wenn ihr Komplement $\mathbb{K} \setminus A$ offen ist.

Beispiele. 1. \mathbb{K} selbst und die leere Menge sind trivialerweise sowohl offen als auch abgeschlossen. Es ist ein Ausdruck der Vollständigkeit der reellen Zahlen, wie wir weiter unten einsehen werden, daß im Fall $\mathbb{K} = \mathbb{R}$ dies auch die *einzigsten* gleichzeitig offenen und abgeschlossenen Mengen sind. Analoges gilt auch für $\mathbb{K} = \mathbb{C}$ und noch allgemeiner für jeden normierten reellen oder komplexen Vektorraum.

2. Die sogenannten „offenen“ Intervalle in einem angeordneten Körper sind tatsächlich offene Mengen in diesem Sinne. Dies ist eine unmittelbare Konsequenz aus der Dreiecksungleichung. Entsprechendes läßt sich für die „abgeschlossenen“ Intervalle aussagen.

3. In \mathbb{Q} ist die Menge

$$\{x \in \mathbb{Q} : x^2 < 2\}$$

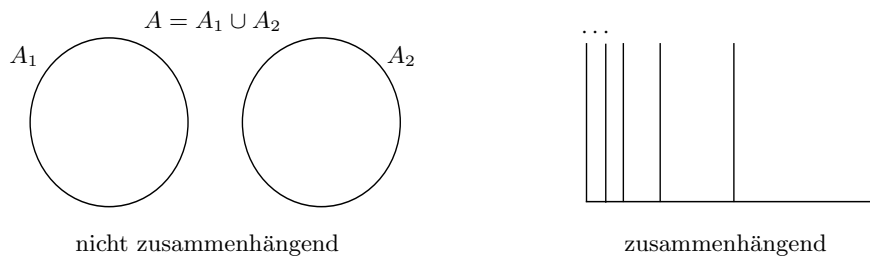
sowohl offen als auch abgeschlossen, da ihr Komplement Vereinigung der beiden offenen Mengen

$$U = \{x \in \mathbb{Q} : x > 0 \text{ und } x^2 > 2\} \quad \text{und} \quad -U$$

ist. Der Beweis sei dem Leser überlassen.

Definition. Eine Teilmenge $A \subset X$ in einem \mathbb{K} -metrischen (oder sogar topologischen) Raum X heißt *zusammenhängend*, wenn für je zwei offene Mengen $U_0, U_1 \subset X$ mit $A_0 := A \cap U_0, A_1 := A \cap U_1$ gilt: Ist $A = A_0 \cup A_1$ und $A_0 \cap A_1 = \emptyset$, so ist notwendig schon eine der Mengen A_0, A_1 leer. Dies ist offensichtlich gleichbedeutend damit, daß die andere der beiden Mengen gleich A ist.

Beispiel. Die leere Menge sowie *einpunktige* Mengen sind stets zusammenhängend. Dagegen braucht X nicht zusammenhängend zu sein, wie man sich an einfachen Beispielen klar machen kann. Zusammenhängende Mengen können eine recht komplizierte Struktur aufweisen (siehe die folgende Skizze und das Beispiel am Ende de Anhangs).



Figur 13.1

Bemerkung. Versieht man A mit der von X induzierten *Relativtopologie* (siehe den Anhang zu Kapitel 7), so bedeutet die vorige Definition, daß die Menge A nicht als disjunkte Vereinigung von nichtleeren, in A offenen Mengen realisiert werden kann. Oder in anderen Worten (da Komplemente von offenen Mengen abgeschlossen heißen): die (in der Relativtopologie von A) einzigen zugleich offenen als auch abgeschlossenen Teilmengen von A sind A selbst und die leere Menge.

Wir werden im folgenden zeigen, daß der Körper der reellen Zahlen den einzigen angeordneten Körper darstellt, für den *jedes Intervall* zusammenhängend ist. Grundlegend dafür ist der Zusammenhang des *abgeschlossenen Einheitsintervalls* $I = [0, 1]$.

Satz 13.1 *Der angeordnete Körper \mathbb{K} erfülle das Supremumsaxiom. Dann ist das abgeschlossene Einheitsintervall I zusammenhängend.*

Beweis. Es seien U_0, U_1 offene Mengen in \mathbb{K} mit $I_0 \cup I_1 = I = [0, 1]$ und $I_0 \cap I_1 = \emptyset$, wobei $I_j := I \cap U_j$, $j = 0, 1$. Ohne Einschränkung sei $0 \in I_0$. Wir setzen dann $A := \{a \in I : [0, a] \subset I_0\}$. Wegen $0 \in A$ ist A nicht leer und wegen $A \subset I$ nach oben beschränkt. Folglich existiert das Supremum α von A und damit eine Folge von Elementen $a_j \in A$, so daß $a_j \rightarrow \alpha$. Aufgrund der Definition von A ist weiter

$$[0, \alpha) \subset \bigcup_{j \in \mathbb{N}} [0, a_j] \subset I_0.$$

Wäre nun $\alpha \notin I_0$, also $\alpha \in I_1 \subset U_1$, so wären wegen der Offenheit von U_1 auch fast alle $a_j \in U_1 \cap I = I_1$, was wegen $a_j \in I_0$ für alle $j \in \mathbb{N}$ nicht sein kann. Also ist sogar $[0, \alpha] \subset I_0 \subset U_0$, und wegen der Offenheit von U_0 muß $\alpha = 1$ sein, denn sonst gäbe es ein positives ε , so daß $[0, \alpha + \varepsilon] \subset I_0$ im Gegensatz zur Definition von α . Somit ist $I_0 = I$ und $I_1 = \emptyset$. \square

Bemerkung. Der Beweis funktioniert wortwörtlich auch für beschränkte abgeschlossene Intervalle $[a, b]$. Wir werden aber weiter unten zeigen, wie man aus dem Zusammenhang des Einheitsintervalls sogar auf den Zusammenhang *aller* - also nicht nur der abgeschlossenen und beschränkten - Intervalle schließen kann.

Die Bedeutung des Zusammenhangs eines abgeschlossenen Intervalls I kommt erst durch Untersuchung von *stetigen Funktionen* auf I voll zum Tragen. Man macht sich sofort klar, daß die Stetigkeit einer Funktion $f : D \rightarrow \mathbb{K}$ durch den Begriff der Offenheit von Mengen charakterisiert werden kann.

Lemma 13.2 *Eine Funktion $f : D \rightarrow \mathbb{K}$, $D \subset \mathbb{K}$, ist genau dann stetig, wenn für jede offene Menge $U_2 \subset \mathbb{K}$ eine offene Menge $U_1 \subset \mathbb{K}$ existiert, so daß für das Urbild $f^{-1}(U_2) = D \cap U_1$ gilt.*

Den (sehr einfachen) *Beweis* lassen wir hier aus, da wir ihn weiter unten im Anhang in noch viel allgemeinerem Rahmen erbringen werden. \square

Bemerkung. Ist insbesondere $D = \mathbb{K}$, so bedeutet dieser Satz: Eine Abbildung $f : \mathbb{K} \rightarrow \mathbb{K}$ ist genau dann stetig, wenn Urbilder offener Mengen offen sind.

Kommen wir nun zur Untersuchung von *Zusammenhangsfragen* zurück. Wir gewinnen aus der gerade gewonnenen Einsicht sofort ein hinreichendes Kriterium, mit dem man aus zusammenhängenden Teilmengen mittels stetiger Funktionen neue gewinnen kann.

Lemma 13.3 *Ist $D \subset \mathbb{K}$ zusammenhängend und $f : D \rightarrow \mathbb{K}$ stetig, so ist auch das Bild $f(D)$ zusammenhängend.*

Beweis. Wir nehmen an, $E = f(D)$ sei nicht zusammenhängend. Dann gibt es offene Mengen $V_0, V_1 \subset \mathbb{K}$, so daß für $E_j := E \cap V_j$, $j = 0, 1$, gilt: $E_j \neq \emptyset$, $E = E_0 \cup E_1$, $E_0 \cap E_1 = \emptyset$. Seien U_j gemäß Lemma 2 offene Mengen in \mathbb{K} mit $f^{-1}(V_j) =: D_j = D \cap U_j$. Dann ist offensichtlich D die disjunkte Vereinigung der nichtleeren (relativ offenen) Mengen D_j . Also war auch D nicht zusammenhängend. \square

Definition. Man nennt einen angeordneten Körper *total unzusammenhängend*, wenn außer der leeren Menge nur die einpunktigen Teilmengen zusammenhängend sind.

Wir haben oben gesehen, daß der Körper der reellen Zahlen nichttriviale zusammenhängende Teilmengen besitzt, also nicht total unzusammenhängend ist. Diese Eigenschaft hat sofort weitere weitreichende Konsequenzen.

Satz 13.4 *Ist der angeordnete Körper \mathbb{K} nicht total unzusammenhängend, so ist jedes Intervall zusammenhängend. Insbesondere ist der Körper \mathbb{K} selbst zusammenhängend.*

Beweis. Wir nehmen zunächst an, daß es ein nichttriviales zusammenhängendes abgeschlossenes Intervall $I = [a, b]$ gibt. Für $c \leq d$ ist die affine Abbildung

$$\alpha(x) := c + \frac{d-c}{b-a}(x-a), \quad x \in \mathbb{K},$$

stetig und bildet I auf das Intervall $[c, d]$ ab. Wegen des vorstehenden Lemmas ist also auch dieses zusammenhängend. Hat man ein beliebiges Intervall J vorgegeben, so schließe man wie folgt: Ist J nicht zusammenhängend, so schreibe man J als disjunkte Vereinigung zweier nichtleerer Mengen J_0, J_1 , die Durchschnitte von J mit offenen Mengen in \mathbb{K} sind. Wähle nun $c \in J_0$ und $d \in J_1$ und nehme ohne Einschränkung $c < d$ an. Aus der Annahme über J folgt dann aber sofort, daß das Intervall $[c, d]$ nicht zusammenhängend ist, was im krassen Widerspruch zum ersten Teil unseres Beweises steht.

Wir müssen also nur noch unsere erste Annahme rechtfertigen. Es sei $Z \subset \mathbb{K}$ eine nichttriviale zusammenhängende Teilmenge, und es seien a, b Elemente in Z mit $a < b$. Wir setzen $I = [a, b]$. Gäbe es ein $c \in I$, das nicht in Z enthalten ist, so wäre $Z = \{x \in Z : x < c\} \cup \{x \in Z : x > c\}$ im Gegensatz zur Voraussetzung des Zusammenhangs von Z ; somit ist $I \subset Z$. Wäre I nicht zusammenhängend, so würde eine stetige Funktion f auf I existieren, die genau die Werte 0 und 1 annimmt: Läßt sich nämlich I nichttrivial als disjunkte Vereinigung $I_0 \cup I_1$ schreiben mit $I_j = I \cap U_j$, U_j offen in \mathbb{K} , so ist die Funktion

$$f(x) := \begin{cases} 0, & x \in I_0 \\ 1, & x \in I_1 \end{cases}$$

stetig auf I . Dann ist aber die durch $f(x) := f(a)$, $x \leq a$, $x \in Z$, $f(x) := f(b)$, $x \geq b$, $x \in Z$, definierte Funktion stetig auf Z und somit Z nicht zusammenhängend, da die Bildmenge $f(Z) = \{0, 1\}$ nicht zusammenhängend ist. \square

Wir wenden uns nun der Formulierung des außerordentlich wichtigen *Zwischenwertsatzes* zu, der oft auch als *Satz von Bolzano* bezeichnet wird.

Satz 13.5 *Das Intervall $I = [a, b] \subset \mathbb{K}$ sei zusammenhängend. Dann nimmt jede stetige Funktion $f : I \rightarrow \mathbb{K}$ auf I jeden Wert zwischen $f(a)$ und $f(b)$ an.*

Beweis. Es sei ohne Einschränkung $f(a) < f(b)$. Da das Bild $f(I)$ zusammenhängend ist, muß nach dem Beweis von Satz 4 auch das Intervall $[f(a), f(b)]$ in $f(I)$ enthalten sein. (Siehe auch Folgerung 14). Dies ist aber gerade unsere Behauptung. \square

Bemerkung. Aus der Gültigkeit des Zwischenwertsatzes folgt umgekehrt auch wieder, daß jedes Intervall $I = [a, b]$ und damit auch \mathbb{K} zusammenhängend sein muß. Anderenfalls wähle wie oben eine stetige Funktion $f : I \rightarrow \mathbb{K}$, die genau die Werte 0 und 1 annimmt. Nach evtl. Verkleinerung von I dürfen wir weiter ohne Einschränkung voraussetzen, daß $f(a) = 0$, $f(b) = 1$. Trotzdem nimmt f den Zwischenwert $1/2$ nicht an. \square

Wir haben oben eingesehen, daß das Einheitsintervall $[0, 1]$ und damit auch jedes abgeschlossene Intervall $[a, b]$ im Körper der reellen Zahlen zusammenhängend ist. Somit haben wir den klassischen Zwischenwertsatz bewiesen:

Satz 13.6 (Zwischenwertsatz) *Es sei $I \subset \mathbb{R}$ ein nicht leeres Intervall, und $f : I \rightarrow \mathbb{R}$ sei eine stetige Funktion. Seien ferner $a, b \in I$ beliebig vorgegeben, so nimmt f jeden Wert zwischen $f(a)$ und $f(b)$ an (mindestens) einer Stelle zwischen a und b an.*

Wegen seiner grundlegenden Bedeutung wollen wir noch einen *direkten Beweis* des Zwischenwertsatzes einfügen, der auf dem Supremumsaxiom beruht. Wir können dazu ohne Einschränkung annehmen, daß $a < b$ und $f(a) < f(b)$ ist. Wählt man dann $f(a) < \eta < f(b)$ beliebig, so müssen wir die Existenz einer Zahl ξ mit $a < \xi < b$ nachweisen, für die $f(\xi) = \eta$ gilt. Man kann dann zu der Funktion $f - \eta$ übergehen und damit einsehen, daß man nur den folgenden Spezialfall des Zwischenwertsatzes zu beweisen braucht. \square

Folgerung 13.7 *Ist $I = [a, b] \subset \mathbb{R}$ und $f : I \rightarrow \mathbb{R}$ eine stetige Funktion mit $f(a) < 0 < f(b)$, so besitzt f im Inneren von I eine Nullstelle.*

Beweis. Das Supremum $\xi := \sup \{x \in I : f(x) < 0\}$ existiert und ist kleiner gleich b . Mit einer Folge $x_j \nearrow \xi$ und der Stetigkeit von f ergibt sich sofort $f(\xi) = f(\lim_j x_j) \leq 0$. Wäre $f(\xi) < 0$, so müßte auch $\xi < b$ sein, und wegen der Stetigkeit von f würde es Stellen ξ' echt zwischen ξ und b geben, an denen f negativ wäre. Dies widerspricht aber der Definition von ξ . \square

Bemerkung. Der Zwischenwertsatz gibt uns eine vierte Methode an die Hand, die Existenz von Wurzeln in \mathbb{R} herzuleiten. Sind nämlich eine reelle Zahl $a > 0$ und $k \in \mathbb{N}^*$ vorgegeben, so ist die Funktion $f(x) = x^k - a$ stetig auf \mathbb{R} . Ihr Funktionswert an der Stelle 0 ist negativ; für hinreichend große natürliche Zahl n ist jedoch $f(n)$ positiv wegen des Archimedischen Axioms. Also besitzt f eine Nullstelle in dem Intervall $[0, n]$.

Beispiel. Selbstverständlich kann der Wert η mehrfach, ja sogar unendlich oft angenommen werden, ohne daß die Funktion f konstant ist. Z. B. nimmt die Funktion $f(x) = x \sin 1/x$, $x \neq 0$, $f(0) = 0$ in jedem Intervall $[-a, a]$ den Wert $\eta = 0$ unendlich oft an. Dies impliziert übrigens auch, daß diese Funktion im Nullpunkt nicht analytisch ist (Identitätssatz für Potenzreihen).

Als nächstes wollen wir einige Anwendungen des Zwischenwertsatzes diskutieren. Wir erinnern daran, daß wir im vorigen Kapitel den Nachweis für die Tatsache schuldig geblieben sind, daß die Exponentialfunktion im Reellen alle positiven Werte annimmt. Dies können wir jetzt schnell nachholen. Es sei also $y_0 > 0$ vorgegeben. Wegen

$$\lim_{x \rightarrow -\infty} e^x = 0, \quad \lim_{x \rightarrow \infty} e^x = \infty$$

gibt es dann Zahlen $a < b$ mit $e^a < y_0 < e^b$. Nach dem Satz von Bolzano wird also der Wert y_0 zwischen a und b angenommen. Fast wörtlich kann man entsprechend für die Wurzelfunktionen auf $[0, \infty)$ argumentieren.

Daß die Wurzelfunktionen, der natürliche Logarithmus und viele andere klassische Funktionen *stetig* sind, ergibt sich aus einem allgemeinen Resultat über Umkehrfunktionen, das wir im folgenden vorbereiten wollen. Es sei also $I \subset \mathbb{R}$ ein Intervall und $f : I \rightarrow \mathbb{R}$ eine stetige Funktion. Wir fragen uns zunächst, unter welchen Voraussetzungen f das Intervall I bijektiv auf sein Bild $f(I)$ abbildet, wann also f injektiv ist. Die nicht überraschende Antwort liefert der folgende

Satz 13.8 *Die stetige Funktion $f : I \rightarrow \mathbb{R}$ ist genau dann injektiv, wenn sie streng monoton ist.*

Wir führen den *Beweis* auf das folgende Lemma zurück.

Lemma 13.9 *Ist $f : I \rightarrow \mathbb{R}$ stetig und injektiv, und sind x_0, x_1, x_2 Punkte in I mit $x_0 < x_1 < x_2$, so folgt aus $f(x_0) < f(x_1)$ stets $f(x_1) < f(x_2)$.*

Beweis (Lemma). Angenommen, es wäre $f(x_1) > f(x_2)$. Dann würde jeder Wert η zwischen dem Maximum von $f(x_0)$ und $f(x_2)$ und $f(x_1)$ sowohl zwischen x_0 und x_1 als auch zwischen x_1 und x_2 angenommen, was der Injektivität von f widerspräche. \square

Beweis (Satz). Jedes Intervall I ist (siehe Satz 16) die abzählbare Vereinigung von einer aufsteigenden Folge von kompakten Intervallen $I_j := [a_j, b_j] \subset I_{j+1} \subset I$:

$$I = \bigcup_{j=0}^{\infty} I_j.$$

Wir können zusätzlich ohne Einschränkung annehmen, daß I (und jedes I_j) mindestens zwei Punkte enthält. Es genügt dann zu zeigen, daß die Einschränkung von f auf jedes Intervall I_j streng monoton ist; denn aus $I_0 \subset I_1 \subset I_2 \subset \dots$ folgt, daß f auf jedem Intervall I_j monoton wächst (bzw. fällt), wenn dies auf I_0 richtig ist, und daraus folgt sofort die strenge Monotonie auf ganz I .

Mit anderen Worten: Wir können uns auf den Fall beschränken, daß $I = [a, b]$ mit $a < b$ gilt. Wegen der Injektivität von f ist ferner $f(a) \neq f(b)$, und wir nehmen ohne Einschränkung an, daß $f(a) < f(b)$. Wir werden zeigen, daß dann f streng monoton wächst. Es seien also $a \leq x_1 < x_2 \leq b$ vorgegeben. Wäre nun $f(x_1) > f(x_2)$, so müßte nach dem voranstehenden Lemma (angewandt auf $-f$) auch $f(x_2) \geq f(b)$ folgen. Ganz entsprechend schließt man auf $f(a) \geq f(x_1)$. Insgesamt wäre also $f(a) > f(b)$ im Gegensatz zu unserer Annahme. Also ist stets $f(x_1) \leq f(x_2)$ und damit wegen der Injektivität von f auch $f(x_1) < f(x_2)$. \square

Da nach Lemma 3 das Bild $J := f(I) \subset \mathbb{R}$ zusammenhängend ist, folgt aus dem später noch zu beweisenden Satz 16, daß J sogar ein Intervall ist. Wir können daher einen großen Teil unserer bisherigen Überlegungen in dem folgenden Satz zusammenfassen.

Satz 13.10 *Die Funktion $f : I \rightarrow \mathbb{R}$, $I \subset \mathbb{R}$ ein Intervall, sei stetig und injektiv. Dann ist f streng monoton wachsend oder fallend. Der Bildbereich $J := f(I)$ ist ebenfalls ein Intervall, und die Umkehrfunktion $f^{-1} : J \rightarrow \mathbb{R}$ von f ist streng monoton wachsend oder fallend und stetig.*

Beweis. Es bleibt nur zu zeigen, daß die Umkehrfunktion $g := f^{-1}$ stetig ist. Hierzu benutzen wir das Folgenkriterium. Es sei also (y_j) eine konvergente Folge in J mit Grenzwert η . Setzen wir $x_j := g(y_j)$ und $\xi := g(\eta)$, so ist zu zeigen, daß die Folge der x_j gegen ξ konvergiert. Nehmen wir an, daß dies nicht der Fall sei. Dann gibt es ein $\varepsilon_0 > 0$, so daß die Abschätzung $|x_j - \xi| < \varepsilon_0$ für unendlich viele j verletzt ist. Durch Übergang zu einer Teilfolge der y_j können wir daher annehmen, daß für alle j die Ungleichung $|x_j - \xi| \geq \varepsilon_0$ besteht. Nach erneutem Übergang zu einer Teilfolge können wir sogar erreichen, daß für alle j entweder

$$(*) \quad x_j \geq \xi + \varepsilon_0$$

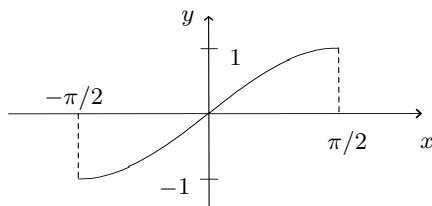
oder für alle j die Ungleichung $x_j \leq \xi - \varepsilon_0$ gilt. Es sei ohne Einschränkung der erste Fall gegeben. Aufgrund der Wahl der Folge (x_j) ist nun $y_j = f(x_j)$ und $\eta = f(\xi)$. Wegen der strengen Monotonie der Funktion f ist unter der Voraussetzung $(*)$ aber

$$y_j = f(x_j) \geq f(\xi + \varepsilon_0),$$

woraus sich sofort ein Widerspruch ergibt:

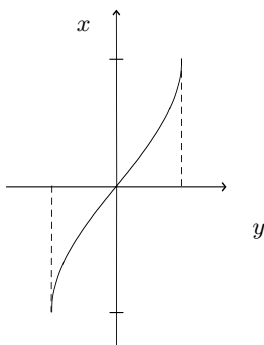
$$\eta = \lim_{j \rightarrow \infty} y_j = \lim_{j \rightarrow \infty} f(x_j) \geq f(\xi + \varepsilon_0) > f(\xi) = \eta. \quad \square$$

Beispiel. Wir wenden dieses Resultat auf die Umkehrfunktion des Sinus an. Im Intervall $[-\pi/2, \pi/2]$ erfüllt die Funktion $\sin x$ die Voraussetzungen des gerade bewiesenen Satzes (siehe auch das nachfolgende Kapitel):



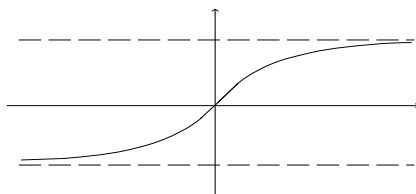
Figur 13.2

Damit besitzt der Sinus eine auf dem Intervall $[-1, 1]$ definierte stetige und streng monoton wachsende Umkehrfunktion, die als *Arcussinus-Funktion*¹⁷ bezeichnet wird, in Symbolen: $x = \arcsin y$.



Figur 13.3

Ganz entsprechend sieht man ein, daß der Tangens auf dem (offenen) Intervall $(-\pi/2, \pi/2)$ eine Umkehrfunktion $\arctan: \mathbb{R} \rightarrow (-\pi/2, \pi/2)$ besitzt.



Figur 13.4

¹⁷arcus (lat.) ist der *Bogen*; also ist x der Bogen, der zu dem Sinuswert y gehört.

Wir schließen als Zwischenspiel einige Bemerkungen zum DEDEKINDSchen *Schnittaxiom* an. Wir sagen, daß ein angeordneter Körper \mathbb{K} diesem Axiom genügt, wenn es zu nichtleeren Mengen $A \leq B$ in \mathbb{K} mit $\mathbb{K} = A \cup B$ stets ein Element $c \in \mathbb{K}$ gibt mit $A \leq \{c\} \leq B$. Die nächsten beiden Sätze liefern die Einsicht, daß alle in diesem Kapitel betrachteten Eigenschaften äquivalent zu den Axiomen (I) bis (VII) sind.

Satz 13.11 *Ist der angeordnete Körper \mathbb{K} zusammenhängend, so genügt er auch dem Dedekindschen Schnittaxiom.*

Beweis (Satz 9). Es seien A, B nichtleere Teilmengen von \mathbb{K} mit $A \cup B = \mathbb{K}$ und $A \leq B$, d. h. $a \leq b$ für alle $a \in A, b \in B$. Wären A und B offen, so müßte $A \cap B = \emptyset$ gelten (anderenfalls gäbe es Elemente $a \in A, b \in B$ mit $a > b$), was dem Zusammenhang von \mathbb{K} widerspräche. Also ist eine der beiden Mengen nicht offen; sei diese ohne Einschränkung A . Dann gibt es einen *nicht inneren* Punkt $c \in A$, für den also kein offenes Intervall mit Mittelpunkt c in A enthalten ist. Da alle kleineren Werte als c aber zu A gehören müssen, ist dann notwendig $[c, c + \varepsilon) \cap B \neq \emptyset$ für alle $\varepsilon > 0$. Daraus folgt sofort $A \leq \{c\} \leq B$. \square

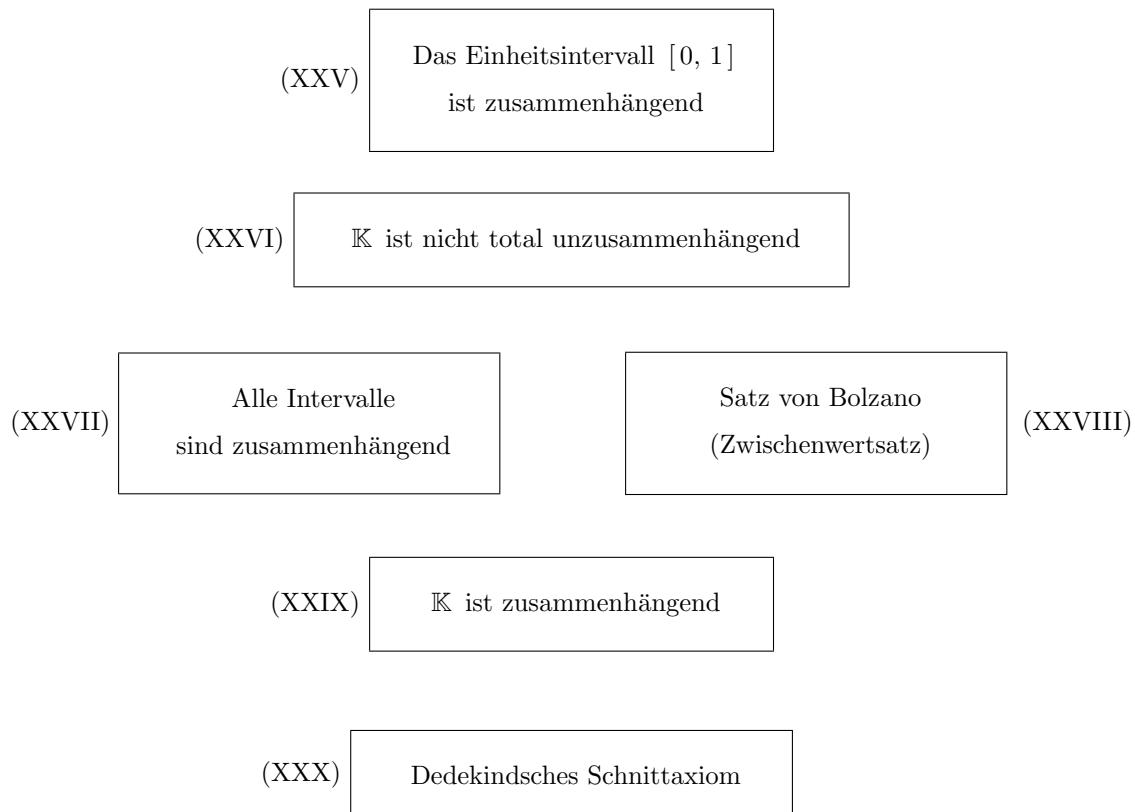
Satz 13.12 *Genügt der angeordnete Körper \mathbb{K} dem Dedekindschen Schnittaxiom, so erfüllt er auch das Prinzip der monotonen Konvergenz.*

Beweis. Es sei $a_j \in \mathbb{K}$ eine monoton aufsteigende, nach oben beschränkte Folge. Man setzt dann

$$A := \{x \in \mathbb{K} : x \leq a_j \text{ für ein } j\}, \quad B := \{x \in \mathbb{K} : x > a_j \text{ für alle } j\}.$$

Sei c der durch $A \leq B$ bestimmte Schnitt. Nach Voraussetzung ist $a_j \leq c$ für alle $j \in \mathbb{N}$. Wird nun $\varepsilon > 0$ beliebig vorgegeben, so gibt es ein j_0 , so daß $a_{j_0} > c - \varepsilon$ (denn sonst wäre $c - \varepsilon/2 \in B$). Da die Folge a_j monoton aufsteigt, ist sie gegen c konvergent. \square

Wir fassen unsere Ergebnisse jetzt wieder in Form von Axiomen(systemen) zusammen, die jedes für sich die reellen Zahlen charakterisieren.



Wir notieren noch eine weitere Charakterisierung der reellen Zahlen, die die allgemeine Theorie der \mathbb{K} -metrischen Räume X benutzt. Ein solcher heie \mathbb{K} -wegweise zusammenhngend, wenn es zu je zwei Punkten $a, b \in X$ eine stetige Abbildung $\gamma : I \rightarrow X$, $I := [0, 1] \subset \mathbb{K}$, gibt mit $\gamma(0) = a$, $\gamma(1) = b$ (siehe den Anhang zur allgemeinen Definition der Stetigkeit). Wie im Beweis von Lemma 3 sieht man, da der Nichtzusammenhang eines wegzusammenhngenden Raumes den Nichtzusammenhang des abgeschlossenen Einheitsintervalls impliziert. Somit ergibt sich aus Axiom (XXV) die folgende Aussage:

(XXXI)
 \mathbb{K} -wegweise zusammenhngende
 \mathbb{K} -metrische Rume
sind zusammenhngend

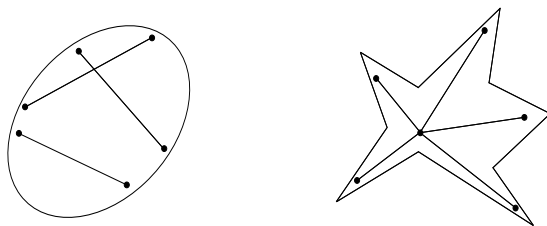
Umgekehrt sind aber offensichtlich alle Intervalle in diesem Sinne wegzusammenhngende \mathbb{K} -metrische Rume und damit (XXV) und (XXXI) sogar quivalent. \square

Wir knnen hieraus schon jetzt eine zu Beginn des Kapitels erwhnte allgemeine Konsequenz ziehen.

Satz 13.13 *Jeder normierte reelle oder komplexe Vektorraum ist zusammenhngend. Insbesondere ist der Krper der komplexen Zahlen zusammenhngend.*

Beweis. Mit je zwei Punkten a, b ist auch deren reelle Verbindungsstrecke $(1-t)a + tb$, $t \in [0, 1]$, in dem Vektorraum enthalten. Also sind normierte reelle oder komplexe Vektorrume stets \mathbb{R} -zusammenhngend, da die Abbildung $t \mapsto (1-t)a + tb$ stetig ist. Nach Axiom (XXXI) sind sie dann auch zusammenhngend. \square

Bemerkung. Das vorige Argument kann man sogar auf *konvexe Teilmengen* K in einem normierten Vektorraum anwenden, wobei eine Teilmenge $K \subset V$ konvex heit, wenn mit zwei Punkten $a, b \in K$ auch ihre Verbindungsstrecke in K liegt. Offene oder abgeschlossene *Kugeln* im euklidischen Vektorraum erfllen diese Bedingung (siehe auch den Anhang zu Kapitel 17). Noch allgemeiner kann man *sternfrmige* Mengen betrachten.



Figur 13.5

In total unzusammenhngenden angeordneten Krpern sind genau die einpunktigen Mengen die (nichtleeren) zusammenhngenden Mengen. Fr den Krper der reellen Zahlen wissen wir bisher nur, da alle Intervalle zusammenhngend sind. Es gilt aber auch die Umkehrung. Diese Charakterisierung der reellen Intervalle benutzt ein wichtiges Nebenresultat aus dem Beweis von Satz 4, das wir deshalb hier gesondert notieren, zumal wir es weiter oben schon beim Beweis des *Zwischenwertsatzes* benutzt haben.

Folgerung 13.14 *Ist $Z \subset \mathbb{K}$ zusammenhngend, \mathbb{K} ein angeordneter Krper, so ist fr alle $a, b \in Z$, $a < b$, das Intervall $I = [a, b] \subset Z$.*

Diese letzte Eigenschaft von zusammenhängenden Mengen in angeordneten Körpern ist auch für sich interessant. Wir sammeln einige dazu äquivalente Bedingungen. Der einfache Beweis sei dem Leser überlassen.

Satz 13.15 *Es sei \mathbb{K} ein angeordneter Körper und Z eine Teilmenge. Dann sind die folgenden Aussagen äquivalent :*

i) Für alle $a, b \in Z$, $a \leq b$, ist $I = [a, b] \subset Z$;

ii) $Z = \bigcup_{\substack{a, b \in Z \\ a < b}} [a, b]$;

iii) für alle $c \in Z$ ist $Z = \bigcup_{\substack{a, b \in Z \\ a \leq c \leq b}} [a, b]$;

iv) es gibt ein $c \in Z$, so daß $Z = \bigcup_{\substack{a, b \in Z \\ a \leq c \leq b}} [a, b]$.

Die oben angekündigte und auch schon verwendete Charakterisierung der Intervalle in \mathbb{R} kann jetzt in die folgende Form gegossen werden.

Satz 13.16 *Es sei J eine nichtleere Menge reeller Zahlen. Dann sind die folgenden Aussagen äquivalent :*

a) J ist ein Intervall ;

b) J ist zusammenhängend ;

c) J ist wegweise zusammenhängend ;

d) J erfüllt eine der äquivalenten Bedingungen aus Satz 15 ;

e) J ist abzählbare Vereinigung einer aufsteigenden Folge von Intervallen $I_j = [a_j, b_j]$;

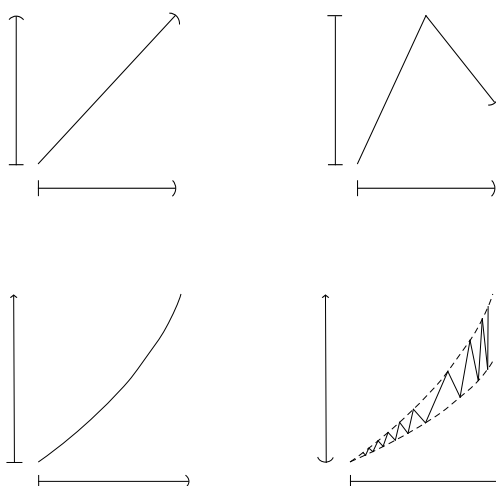
f) J ist abzählbare Vereinigung von Intervallen $I_j = [a_j, b_j]$ mit nichtleerem Durchschnitt $\bigcap I_j$.

Beweis. Die Implikationen a) \implies b) \implies d) haben wir oben schon begründet. Völlig elementar sind die Folgerungen a) \implies e) \implies f) \implies d). Ist d) erfüllt, also z. B. die Aussage i) aus Satz 15, so ist J offensichtlich wegweise zusammenhängend und damit wegen Axiom XXXI zusammenhängend. Es bleibt noch d) \implies a). Hierzu verwendet man z. B. die Aussage iv) in Satz 15 und bildet $\beta := \sup J$ im eigentlichen oder uneigentlichen Sinne. Je nachdem, ob β in J enthalten ist oder nicht, ist dann

$$J \cap [c, \infty) = [c, \beta] \quad \text{bzw.} \quad = [c, \beta).$$

Entsprechend behandelt man die untere Grenze. □

Aus unseren bisherigen Überlegungen ergibt sich, da Bilder zusammenhängender Mengen unter stetigen Abbildungen wieder zusammenhängend sind, daß unter einer stetigen Funktion $f : D \rightarrow \mathbb{R}$, $D \subset \mathbb{R}$, jedes Intervall $I \subset D$ auf ein Intervall $J \subset \mathbb{R}$ abgebildet wird. Man kann sich aber leicht mit verschiedenen Beispielen klar machen, daß J alle „Intervalltypen“ annehmen kann, sofern das Intervall I nicht abgeschlossen und beschränkt ist. Ist z. B. $I = \mathbb{R}$, so ist $J = \mathbb{R}_+^*$ für $f = \exp$, $J = [-1, 1]$ für $f = \sin$, $J = (-\pi/2, \pi/2)$ für $f = \tan$, $J = [0, \pi/2)$ für $f = |\tan|$, $J = \mathbb{R}^*$ für $f(x) := |x \sin x|$. Weitere Beispiele kann man den folgenden Skizzen entnehmen.



Figur 13.6

Ist dagegen das Intervall I abgeschlossen und beschränkt, oder wie wir später sagen werden: ist I *kompakt*, so kann das Bildintervall unter einer stetigen Funktion nur wieder von demselben Typ sein. Dies ist der Inhalt des folgenden wichtigen Satzes, den man als weitreichende Verallgemeinerung von Satz 5 ansehen kann.

Wir beginnen mit dem Beweis von

Satz 13.17 *Erfüllt der angeordnete Körper \mathbb{K} Axiom (IV), so nimmt jede stetige Funktion $f : I := [a, b] \rightarrow \mathbb{K}$ ihr Supremum auf I an, d. h.: es gibt (mindestens) eine Stelle $c \in I$ mit maximalem Funktionswert: $f(x) \leq f(c)$ für alle $x \in I$.*

Bemerkung. Auch wenn es nicht korrekt ist, wird der im vorstehenden Satz benannte Tatbestand auch allgemein dadurch ausgedrückt, daß man sagt, die Funktion f nehme ihr *Maximum* auf I an. Wir werden dennoch diese Sprechweise gelegentlich verwenden.

Beweis. Mit Axiom (IV) sind auch die Axiome (V), (XXI) und (XXII) erfüllt. Es sei nun $K := \sup f(I) \in \mathbb{K}$, wenn die Menge $f(I)$ nach oben beschränkt ist, wobei wir (XXI) verwenden. Im anderen Fall setzen wir $K = \infty$. Es gibt dann im ersten Fall wegen (XXII), im zweiten Fall wegen des Archimedischen Axioms eine Folge $x_j \in I$ mit $\lim_{j \rightarrow \infty} f(x_j) = K$. Wegen (V) können wir nach Übergang zu einer Teilfolge annehmen, daß der Grenzwert c der Folge (x_j) existiert, der wegen der Monotonieeigenschaft des Grenzwertes in I liegt. Da f (folgen-) stetig ist, ergibt sich sofort

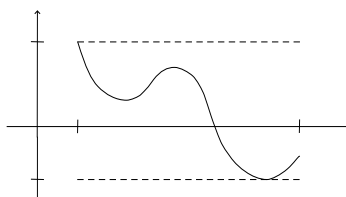
$$K = \lim_{j \rightarrow \infty} f(x_j) = f(c).$$

Insbesondere ist K endlich, und die Funktion f nimmt an der Stelle c ihr Supremum an. □

Folgerung 13.18 *Unter der obigen Voraussetzung nimmt die Funktion f auch ihr Infimum an, und der Bildbereich $f(I)$ von f ist das kompakte Intervall $[\min_{x \in I} f(x), \max_{x \in I} f(x)]$.*

Beweis. Es seien $c, d \in I$ Stellen, an denen $-f$ bzw. f ihr Maximum annehmen; dann gilt $f(c) \leq f(x) \leq f(d)$ für alle $x \in I$. Wegen des *Zwischenwertsatzes* werden auch alle Werte $y \in [f(c), f(d)]$ von der Funktion f angenommen. □

Bemerkungen. 1. In gewissem Sinne ist der vorstehende Satz eine der tiefstliegenden Aussagen der Analysis. Sein Beweis ist zwar elementar, aber nicht trivial, da er mehrere grundlegende Eigenschaften der reellen Zahlen verwendet. Man kann seinen Inhalt etwa in dem folgenden Bild veranschaulichen.



Figur 13.7

2. Aus der Aussage der Folgerung ergibt sich umgekehrt auch wieder sowohl der Zwischenwertsatz als auch Satz 17.

3. Die stetigen Funktionen auf einem Intervall I bilden offensichtlich einen \mathbb{R} -Vektorraum (und sogar eine kommutative \mathbb{R} -Algebra mit Eins), den man stets mit $C^0(I, \mathbb{R})$ bezeichnet. Ist das Intervall I kompakt, so ist dieser Vektorraum sogar im Raum der beschränkten Funktionen enthalten. Somit besitzt er dann zusammen mit der Supremumsnorm die Struktur eines vollständig normierten Vektorraums (siehe Folgerung 12.31).

Möchte man nur die Annahme des Supremums von $f : I \rightarrow \mathbb{K}$ garantieren (nicht aber die des Infimums), so zeigt eine einfache Modifikation des obigen Beweises, wobei man statt des Folgenkriteriums die ε - δ -Variante heranziehen muß, daß die Funktion f auch nur die Hälfte der Eigenschaften einer stetigen Funktion zu besitzen braucht.

Definition. Eine Funktion $f : D \rightarrow \mathbb{K}$, $D \subset \mathbb{K}$, heißt im Punkte $a \in D$ *halbstetig nach oben* (oder *nach oben halbstetig*), wenn es zu jedem $\varepsilon > 0$ ein $\delta = \delta(\varepsilon) > 0$ gibt, so daß

$$f(x) \leq f(a) + \varepsilon \quad \text{für } x \in D \quad \text{mit } |x - a| \leq \delta.$$

Dies ist gleichbedeutend mit der Forderung: Zu jedem $L \in \mathbb{K}$ mit $f(a) < L$ gibt es ein $\delta = \delta(L) > 0$, so daß

$$f(x) < L \quad \text{für } x \in D \quad \text{mit } |x - a| \leq \delta.$$

Sie heißt halbstetig *nach unten*, wenn $-f$ halbstetig nach oben ist.

Bemerkung. Eine Funktion ist genau dann stetig in a , wenn sie dort halbstetig nach oben und nach unten ist.

Satz 13.19 *Erfüllt der angeordnete Körper \mathbb{K} Axiom (IV), so nimmt jede nach oben halbstetige Funktion $f : I := [a, b] \rightarrow \mathbb{K}$ ihr Supremum auf I an.*

Beweis. Es sei wie oben $K := \sup f(I) \in \mathbb{K} \cup \{\infty\}$ und (x_j) eine (ohne Einschränkung gegen $c \in I$ konvergente) Folge mit $\lim_j f(x_j) = K$. Nach Voraussetzung gibt es zu $L_0 := f(c) + 1$ ein $\delta_0 = \delta(L_0) > 0$, so daß $f(x) < L_0$ für $x \in I$ mit $|x - c| \leq \delta_0$. Dies gilt dann insbesondere für fast alle x_j , so daß notwendigerweise $K \leq L_0$ und insbesondere $K \in \mathbb{K}$. Es bleibt zu zeigen, daß $f(c) = K$. Wäre nämlich $f(c) < K$, so könnte man ein L finden mit $f(c) < L < K$ und für fast alle j wäre $x_j \leq L < K$ im Widerspruch zur Definition von K . \square

Da umgekehrt die Folgerung 18 den Zwischenwertsatz nach sich zieht, gewinnen wir die Äquivalenz aller bisherigen Axiome zu den beiden folgenden.

(XXXII) (Nach oben Halb-) Stetige Funktionen auf $[a, b]$ nehmen ihr Supremum an

(XXXIII) Stetige Bilder von kompakten Intervallen
sind kompakte Intervalle

Es sei noch ein weitere Satz angeführt, der ebenfalls zum Vollständigkeitsaxiom äquivalent ist.

Satz 13.20 *Jede monoton aufsteigende Funktion $f : I \rightarrow I$ des Einheitsintervalls $I := [0, 1] \subset \mathbb{R}$ besitzt einen Fixpunkt.*

Beweis. Wir führen einen Beweis für die Existenz eines Fixpunktes mit Hilfe des Prinzips der Intervallschachtelung. Dazu setze man $a_0 = 0$, $b_0 = 1$, so daß wegen der wachsenden Monotonie von $f : [0, 1] \rightarrow [0, 1]$ gilt:

$$a_0 \leq f(a_0) \leq f(b_0) \leq b_0.$$

Wir konstruieren nun durch sukzessives Halbieren Intervalle

$$I_j = [a_j, b_j] \subset I_0 = [a_0, b_0] \quad \text{mit} \quad a_j \leq f(a_j) \leq f(b_j) \leq b_j$$

und (automatisch) $b_j - a_j = (1/2)^j$. Ist für beliebiges j nämlich I_j schon konstruiert, so sei $m_j = \frac{1}{2}(a_j + b_j)$, und man setze

$$\begin{cases} a_{j+1} = a_j, & b_{j+1} = m_j, & \text{falls } f(m_j) \leq m_j, \\ a_{j+1} = m_j, & b_{j+1} = b_j, & \text{falls } f(m_j) > m_j. \end{cases}$$

Es sei $x_0 \in \mathbb{R}$ das wohlbestimmte Element in $\bigcap [a_j, b_j]$. Wegen $a_j \leq x_0 \leq b_j$ ist dann $a_j \leq f(a_j) \leq f(x_0) \leq f(b_j) \leq b_j$; also liegt auch $f(x_0)$ in dem Durchschnitt aller dieser Intervalle, so daß notwendigerweise

$$f(x_0) = x_0.$$

Wir geben noch einen *zweiten* Beweis mit Hilfe des Supremumsaxioms. Man setze

$$A = \{x \in [0, 1] : x \leq f(x)\}.$$

Wegen $0 \leq f(0)$ ist A nicht leer, und somit existiert $x_0 = \sup A \in [0, 1]$. Wir behaupten, daß $f(x_0) = x_0$. Wäre nämlich $x_0 < f(x_0)$, so würde mit der Monotonie von f die Ungleichung $f(x_0) \leq f(f(x_0))$ folgen, also $f(x_0) \in A$, was der Definition von x_0 als kleinster oberer Schranke von A widerspräche. Andererseits gibt es im Fall $f(x_0) < x_0$ Punkte $x \in A$ mit $f(x_0) < x \leq x_0$, womit sich ebenfalls ein Widerspruch einstellt:

$$f(x) \leq f(x_0) < x \leq f(x).$$

□

Bemerkung. Da in beiden Beweisen wesentlich benutzt wurde, daß der Körper \mathbb{R} archimedisch und vollständig ist, kann man vermuten, daß die Aussage des Satzes wieder *charakteristisch* für \mathbb{R} ist, d. h. in einem beliebigen angeordneten Körper \mathbb{K} *nur* dann gilt, wenn dieser archimedisch und vollständig, also zu \mathbb{R} isomorph ist. Dies ist tatsächlich richtig; wir zeigen sogar mehr:

Satz 13.21 *Ist der angeordnete Körper \mathbb{K} nicht archimedisch oder nicht vollständig, so gibt es stetige, streng monoton wachsende Funktionen $f : [0, 1] \rightarrow [0, 1]$ ohne Fixpunkte.*

Als Folgerung hieraus gewinnt man noch einmal die Einsicht, daß der Zwischenwertsatz nur im Körper der reellen Zahlen gültig ist.

Folgerung 13.22 *In solchen Körpern gilt der Zwischenwertsatz nicht, d. h. auch der Zwischenwertsatz ist charakteristisch für \mathbb{R} .*

Beweis (Folgerung). Bilde $g(x) := f(x) - x$ mit einer Funktion f wie im Satz; insbesondere ist $g(x) \neq 0$ für alle $x \in [0, 1]$, aber $g(0) = f(0) \geq 0$ und $g(1) = f(1) - 1 \leq 0$. \square

Beweis (Satz). Es sei \mathbb{K} zunächst als nicht archimedisch angeordneter Körper vorausgesetzt. Dann gibt es *unendlich kleine* positive Elemente ε in \mathbb{K} , also Elemente $\varepsilon > 0$ mit $\varepsilon \leq 1/n$ für alle $n \in \mathbb{N}^*$. Wir definieren mit einem solchen ε die Mengen

$$I_0 = \bigcup_{n \in \mathbb{N}} \{x \in [0, 1] : x \leq n\varepsilon\}, \quad I_1 = \{x \in [0, 1] : x > n\varepsilon \text{ für alle } n \in \mathbb{N}\},$$

und überlassen dem Leser den Nachweis der folgenden Behauptungen:

1. $0 \in I_0, 1 \in I_1$;
2. $I = [0, 1] = I_0 \cup I_1$;
3. $I_0 < I_1$, d. h. für alle $x_0 \in I_0$ und $x_1 \in I_1$ gilt $x_0 < x_1$; insbesondere ist $I_0 \cap I_1 = \emptyset$;
4. für $x_j \in I_j$ ist $(x_j - \varepsilon, x_j + \varepsilon) \cap I \subset I_j, j = 1, 2$.

Die Konstruktion der gesuchten Funktion ist nun denkbar einfach. Man setzt

$$f(x) = \begin{cases} x + \varepsilon/3, & x \in I_0 \\ x - \varepsilon/3, & x \in I_1 \end{cases}$$

und begründet leicht die Stetigkeit von f mit der Aussage 4. Die strenge Monotonie von f braucht nur nachgewiesen zu werden für Punktepaare (x_0, x_1) mit $x_j \in I_j$. Wäre $f(x_1) \leq f(x_0)$ für ein solches Paar, so erhielte man

$$x_1 - \frac{\varepsilon}{3} \leq x_0 + \frac{\varepsilon}{3}, \text{ also } x_0 < x_1 \leq x_0 + 2\frac{\varepsilon}{3} < x_0 + \varepsilon,$$

also, wieder nach 4., den Widerspruch $x_1 \in (x_0, x_0 + \varepsilon) \cap I \subset I_0$. Schließlich ist $f(0) = \varepsilon/3 > 0$ und $f(1) = 1 - \varepsilon/3 < 1$, also auch $0 < f(x) < 1$ für alle $x \in I$. Augenscheinlich besitzt f keine Fixpunkte. Wir können also nunmehr annehmen, daß \mathbb{K} archimedisch angeordnet, aber nicht vollständig ist. Dann ist (ohne Einschränkung) $\mathbb{K} \subset \mathbb{R}$, aber $\mathbb{K} \neq \mathbb{R}$; somit existiert ein $x_0 \in [0, 1]_{\mathbb{R}}$ mit $x_0 \notin [0, 1]_{\mathbb{K}} = [0, 1]_{\mathbb{R}} \cap \mathbb{K}$. Da $\mathbb{Q} \subset \mathbb{K} \subset \mathbb{R}$ und \mathbb{Q} in \mathbb{R} dicht liegt, gibt es eine aufsteigende Folge $x'_0 = 0 < x'_1 < x'_2 < \dots$ von Elementen $x'_j \in \mathbb{K}$ mit $\lim x'_j = x_0$ (als Grenzwert in \mathbb{R}). Entsprechend hat man $x''_0 = 1 > x''_1 > \dots$ in \mathbb{K} mit $\lim x''_j = x_0$. Man erklärt nun f als stückweise affine Funktion auf den Intervallen $[x'_j, x'_{j+1}]_{\mathbb{K}}$ bzw. $[x''_{j+1}, x''_j]_{\mathbb{K}}$, und zwar so, daß

$$f(x'_j) = x'_{j+1}, f(x'_{j+1}) = x'_{j+2} \quad \text{und} \quad f(x''_j) = x''_{j+1}, f(x''_{j+1}) = x''_{j+2}.$$

Man sieht leicht, daß hierdurch eine stetige, streng monoton wachsende Funktion $f : [0, 1]_{\mathbb{K}} \rightarrow [0, 1]_{\mathbb{K}}$ ohne Fixpunkt definiert wird. \square

Wir notieren damit als 34. Axiom:

(XXXIV)	<p>Jede monoton aufsteigende Funktion</p> $f : I \rightarrow I, I := [0, 1] \subset \mathbb{K},$ <p>besitzt einen Fixpunkt.</p>
---------	---

Zu guter Letzt soll noch kurz auf die Frage eingegangen werden, welchen Einfluß die *Monotonie* einer reellen Funktion $f : I \rightarrow \mathbb{R}, I \subset \mathbb{R}$, auf die Menge ihrer *Unstetigkeitsstellen* hat. In der Tat

kann diese Menge *dicht* im Definitionsbereich liegen.

Beispiel. Es sei $\mathbb{Q} = \{r_n : n \in \mathbb{N}\}$ eine Abzählung der rationalen Zahlen; man definiere

$$f(x) := \sum_{r_n < x} 2^{-n}, \quad x \in \mathbb{R}.$$

f ist dann eine auf \mathbb{R} wohldefinierte Funktion, denn für jedes $x \in \mathbb{R}$ ist die definierende Reihe absolut summierbar, weil sie die obere Schranke

$$\sum_{n=0}^{\infty} 2^{-n} = 2$$

besitzt. Ist $x_1 < x_2$, so liegen rationale Zahlen echt zwischen diesen reellen Werten, und somit enthält die Reihe für $f(x_2)$ mehr (positive) Summanden als die Reihe für $f(x_1)$. Damit ist f streng monoton wachsend. Wir behaupten, daß f in jedem Punkt von $\mathbb{R} \setminus \mathbb{Q}$ stetig und in jedem Punkt von \mathbb{Q} unstetig ist.

Ist nämlich $a \in \mathbb{Q}$, also $a = r_N$ für ein $N \in \mathbb{N}$, so ist für jedes $x \in \mathbb{R}$, welches größer ist als a ,

$$f(a) + 2^{-N} = \sum_{r_n < r_N} 2^{-n} + 2^{-N} = \sum_{r_n \leq r_N} 2^{-n} \leq \sum_{r_n < x} 2^{-n} = f(x).$$

Damit ist $f(x) - f(a) \geq 2^{-N}$, $x > a$, und folglich ist f in a *nicht* stetig. In jedem $a \in \mathbb{R} \setminus \mathbb{Q}$ ist f jedoch stetig. Dazu reicht es zu zeigen, daß

$$\lim_{x \searrow a} f(x) = f(a) \quad \text{und} \quad \lim_{x \nearrow a} f(x) = f(a).$$

Wir beweisen die erste Aussage; die zweite läßt sich ganz analog herleiten. Wegen der strengen Monotonie von f ist $f(x) > f(a)$ für alle $x > a$, und damit genügt zu zeigen:

Zu jedem $\varepsilon > 0$ existiert ein $\delta > 0$, so daß für alle $x > a$ mit $x - a < \delta$ folgt: $f(x) < f(a) + \varepsilon$.

Nun gibt es zu $\varepsilon > 0$ ein N , so daß $\sum_{n=N}^{\infty} 2^{-n} < \varepsilon$. Da $a \notin \mathbb{Q}$, ist $\delta := \min_{n < N} \{ |a - r_n| \} > 0$ und folglich $|a - r_n| \geq \delta$ für alle $n < N$. Für alle $x \in \mathbb{R}$ mit $0 < x - a < \delta$ gilt dann aber

$$f(x) - f(a) = \sum_{r_n < x} 2^{-n} - \sum_{r_n < a} 2^{-n} = \sum_{a \leq r_n < x} 2^{-n} \leq \sum_{a \leq r_n < a + \delta} 2^{-n} \leq \sum_{n=N}^{\infty} 2^{-n} < \varepsilon.$$

Dieses Beispiel ist insofern „optimal“, als die Mächtigkeit der Menge der Unstetigkeitsstellen die der rationalen Zahlen nicht übersteigen kann. Dies ist der Inhalt des folgenden Satzes, der in der Integrationstheorie noch Verwendung finden wird.

Satz 13.23 *Es sei $f : I \rightarrow \mathbb{R}$, $I \subset \mathbb{R}$ ein nichttriviales Intervall, eine monoton wachsende Funktion. Dann ist die Menge der Unstetigkeitsstellen von f höchstens abzählbar.*

Beweis. Wir können die beiden eventuell vorhandenen Endpunkte des Intervalls I getrost außer Acht lassen und betrachten im folgenden nur Stellen ξ im Inneren von I . Für eine solche setzen wir

$$S := \sup_{x < \xi} f(x).$$

Wegen $f(x) \leq f(\xi)$ für alle $x < \xi$ existiert dieses Supremum und ist kleiner oder gleich $f(\xi)$. Ist nun (x_j) eine von links gegen ξ konvergente Folge, so gibt es zu vorgegebenem $\varepsilon > 0$ ein $x < \xi$ mit $f(x) > S - \varepsilon$ und ein $N \in \mathbb{N}$ mit $x_j \geq x$ für alle $j > N$. Für diese j ist dann

$$S - \varepsilon < f(x) \leq f(x_j) \leq S < S + \varepsilon,$$

also $\lim_{j \rightarrow \infty} f(x_j) = S$. Mit anderen Worten: Es existiert für alle betrachteten ξ der linksseitige Grenzwert

$$f^-(\xi) := \lim_{x \nearrow \xi} f(x) = S \leq f(\xi).$$

Genauso überzeugt man sich von der Existenz des rechtsseitigen Grenzwertes

$$f^+(\xi) := \lim_{x \searrow \xi} f(x) \geq f(\xi).$$

Die Funktion f ist nun an der Stelle ξ genau dann stetig, wenn diese beiden Grenzwerte übereinstimmen. Ist also f an der Stelle ξ unstetig, so ist $f^-(\xi) < f^+(\xi)$ und wir können eine rationale Zahl r_ξ auswählen, die echt zwischen diesen beiden Grenzwerten liegt. Ist $\xi' > \xi$ eine weitere Unstetigkeitsstelle, so ist - wieder wegen der Monotonie -

$$\lim_{x \searrow \xi} f(x) \leq \lim_{x \nearrow \xi'} f(x).$$

Infolgedessen ist die rationale Zahl $r_{\xi'}$ von r_ξ verschieden, und die Abbildung $\xi \mapsto r_\xi \in \mathbb{Q}$ ist injektiv. Also ist die Menge der Unstetigkeitsstellen gleichmächtig zu einer Teilmenge der abzählbar unendlichen Menge \mathbb{Q} und damit höchstens abzählbar (siehe Satz 3.9). \square

Anhang: Stetige Abbildungen zwischen topologischen Räumen

Wir stellen in diesem Anhang die wichtigsten Definitionen und Tatsachen über *stetige Abbildungen* zwischen metrischen und topologischen Räumen zusammen. Wir geben ohne Umschweife die folgende

Definition. Eine Abbildung $f : X \rightarrow Y$ zwischen topologischen Räumen X und Y heißt *stetig* in einem Punkt $a \in X$, wenn es zu jeder Umgebung $V \subset Y$ von $b = f(a)$ eine Umgebung $U \subset X$ von a gibt mit $f(U) \subset V$. Eine Abbildung $f : X \rightarrow Y$ heißt *stetig* (schlechthin), wenn sie an jeder Stelle $a \in X$ stetig ist. Ist $A \subset X$ eine Teilmenge und $f : A \rightarrow Y$ gegeben, so heißt f *stetig* (auf A), wenn f bzgl. der von X auf A induzierten *Relativtopologie* stetig ist.

Bemerkungen. 1. Anstelle der *vollen* Umgebungsfilter braucht man bei der obigen Definition selbstverständlich nur *Umgebungsbasen* von $b = f(a)$ und a zu testen.

2. Sind X und Y metrische Räume (evtl. sogar mit Metriken in unterschiedlichen angeordneten Körpern), so ist wegen 1. die obige Definition gleichbedeutend mit: Zu jeder Kugel $B(b, \varepsilon)$ gibt es eine Kugel $B(a, \delta)$ mit

$$f(B(a, \delta)) \subset B(b, \varepsilon).$$

3. Ist die Abbildung f nur auf einer Teilmenge $A \subset X$ mit $a \in A$ erklärt, so definiert man in Übereinstimmung mit der allgemeinen Definition: $f : A \rightarrow Y$ heißt stetig in a , wenn es zu jeder Kugel $B(b, \varepsilon) \subset Y$ eine Kugel $B(a, \delta) \subset X$ gibt mit

$$f(B(a, \delta) \cap A) \subset B(b, \varepsilon).$$

3. Sind X und Y normierte Vektorräume, so bedeutet dies: Zu jedem $\varepsilon > 0$ gibt es ein $\delta > 0$, so daß aus $\|x - a\| < \delta$, $x \in A$, folgt, daß $\|f(x) - f(a)\| < \varepsilon$. Insbesondere stimmt die obige Definition im Falle $A \subset X = Y = \mathbb{K}$ mit unser früher gegebenen überein.

4. Die identische Abbildung $\text{id} : X \rightarrow X$ ist stets stetig. Ist $f : X \rightarrow Y$ stetig in a und $g : Y \rightarrow Z$ stetig in $b = f(a)$, so ist $g \circ f : X \rightarrow Z$ stetig in a .

Die folgende Aussage wurde schon zitiert und verwendet; ihr Beweis wird nun endlich nachgetragen.

Satz 13.24 *Eine Abbildung $f : X \rightarrow Y$ zwischen topologischen Räumen ist genau dann stetig, wenn Urbilder offener Mengen unter f offen sind.*

Beweis. Es sei f stetig und $V \subset Y$ offen. Ist dann $a \in f^{-1}(V)$ und $b = f(a)$, so ist V eine Umgebung von b . Folglich gibt es eine Umgebung U von a mit $f(U) \subset V$, also $U \subset f^{-1}(V)$. Somit ist das Urbild von V Umgebung jedes seiner Punkte und folglich eine offene Menge in X . Ist umgekehrt diese Bedingung erfüllt und V eine Umgebung eines Punktes $b = f(a) \in Y$, so gibt es per definitionem eine *offene* Umgebung $V' \subset V$ von b . Nach Voraussetzung ist dann $U := f^{-1}(V')$ eine (offene) Umgebung von a mit $f(U) \subset V' \subset V$. Somit ist f an jeder Stelle $a \in X$ stetig. \square

Das folgende Beispiel ist vollständig elementar, erhält aber aufgrund seiner häufigen Verwendung den Rang eines Lemmas.

Lemma 13.25 *Ist $(V, \|\cdot\|)$ ein normierter \mathbb{K} -Vektorraum, so ist die Normabbildung*

$$\|\cdot\| : V \longrightarrow \mathbb{K}_0$$

von V in den Bewertungskörper \mathbb{K}_0 stetig.

Beweis. Es sei $a \in V$ fest gewählt, und $x \in V$ sei beliebig. Aus der Dreiecksungleichung folgt dann $\|x\| \leq \|x - a\| + \|a\|$ und damit $\|x\| - \|a\| \leq \|x - a\|$. Durch Vertauschen der Rollen von x und a erhält man hieraus die auch sonst nützliche Ungleichung

$$|\|x\| - \|a\|| \leq \|x - a\|,$$

aus der unmittelbar die Stetigkeit der Normfunktion folgt. \square

Ist auf der anderen Seite $Y = V$ ein normierter Vektorraum über dem bewerteten Körper \mathbb{K} , so lassen sich Abbildungen von X nach V auch addieren und mit Funktionen $X \rightarrow \mathbb{K}$ multiplizieren. Fast wörtlich wie oben beweist man den

Satz 13.26 Sind $f, g : X \rightarrow V$ und $c : X \rightarrow \mathbb{K}$ stetig in $a \in X$, so auch $f + g$ und cf .

Den Zusammenhang zwischen stetigen und folgenstetigen Abbildungen kann man jetzt wie folgt ausdrücken.

Lemma 13.27 Ist $f : X \rightarrow Y$ stetig in $a \in X$, so auch folgenstetig:

$$\lim_{j \rightarrow \infty} f(x_j) = f(a) \quad \text{für} \quad \lim_{j \rightarrow \infty} x_j = a.$$

Ist umgekehrt f folgenstetig in a und besitzt a eine abzählbare Umgebungsbasis, so ist f stetig in a .

Zum Schluß sollen noch einige kurze Bemerkungen zum allgemeinen Begriff des Zusammenhangs von topologischen Räumen angefügt werden.

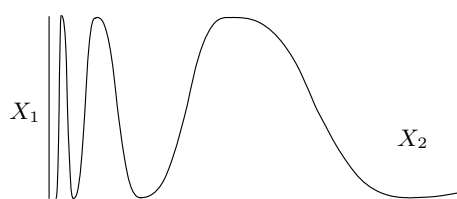
Definition. Einen topologischen Raum X nennt man *zusammenhängend*, wenn er nicht als disjunkte Vereinigung von zwei nicht leeren offenen Mengen geschrieben werden kann. Eine Teilmenge $A \subset X$ heißt *zusammenhängend*, wenn sie in der Relativtopologie zusammenhängend ist. Ein topologischer Raum X heißt *wegweise zusammenhängend*, wenn er im obigen Sinne \mathbb{R} -wegweise zusammenhängend ist, wenn es also zu je zwei Punkten $a, b \in X$ eine stetige Abbildung $\gamma : I \rightarrow X$ des abgeschlossenen reellen Einheitsintervalls $I = [0, 1] \subset \mathbb{R}$ gibt mit $\gamma(0) = a$, $\gamma(1) = b$. Man nennt eine solche Abbildung auch einen (stetigen) *Weg* oder eine *Kurve* in X mit *Anfangspunkt* a und *Endpunkt* b und die Bildmenge $\gamma(I) \subset X$ die *Spur* des Weges $\gamma : I \rightarrow X$.

Man überzeugt sich sofort aufgrund unserer oben unter restriktiveren Voraussetzungen gemachten Überlegungen, daß generell die folgenden Aussagen richtig sind.

Lemma 13.28 Stetige Bilder von zusammenhängenden Mengen, insbesondere Spuren von stetigen Wegen, sind zusammenhängend.

Lemma 13.29 Ist der topologische Raum X wegzusammenhängend, so auch zusammenhängend.

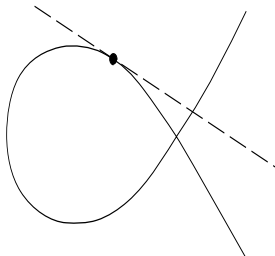
Die Umkehrung des vorigen Lemmas ist jedoch nicht richtig. Es existieren durchaus zusammenhängende topologische (sogar metrische) Räume, die nicht wegzusammenhängend sind. Ein Beispiel ist die mit der Relativmetrik versehene Teilmenge X der euklidischen Ebene \mathbb{R}^2 , die die Vereinigung der Strecke $X_1 := \{(x, y) \in \mathbb{R}^2 : x = 0, -1 \leq y \leq 1\}$ mit dem Graphen der Funktion $\sin(1/x)$ auf \mathbb{R}_+^* , also der Menge $X_2 := \{(x, y) \in \mathbb{R}^2 : x > 0, y = \sin(1/x)\}$ ist (natürlich kann man hier auch die Sinusfunktion durch die weiter oben in Kapitel 12 definierte stückweise lineare Funktion g ersetzen; siehe Figur 12.3). Es ist eine leichte Übungsaufgabe einzusehen, daß dieser Raum zusammenhängend ist (im Wesentlichen liegt dies daran, daß jede Umgebung eines Punktes in X_1 auch die Menge X_2 treffen muß). Selbstverständlich sind auch die Mengen X_1 und X_2 mit der jeweiligen Relativtopologie für sich wegzusammenhängend; ihre Vereinigung kann es jedoch nicht sein: Kein Punkt in X_2 kann durch einen stetigen Weg mit einem Punkt in X_1 verbunden werden. Anschaulich ist dies einleuchtend, da ein solcher Weg „unendlich lang“ sein müßte. Am einfachsten sieht man dies mit einem *Kompaktheitsargument* ein (siehe das übernächste Kapitel). Man kann aber auch einen direkten Beweis führen, den zu finden wir dem Leser überlassen.



Figur 13.8

14 Differenzierbare Funktionen

Der Begriff der *Differenzierbarkeit* hat zu tun mit (momentanen) *Wachstumsraten* (wie z. B. *Geschwindigkeiten* oder *Beschleunigungen*), aber auch mit *linearer Approximation*. Historisch erwachsen ist dieser Begriff aus dem Bemühen, für allgemeinere als ebene algebraische Kurven C *Tangenten* zu finden, also Geraden, die die Kurve an einer festen Stelle (x_0, y_0) berühren:



Figur 14.1

Gemäß unserer Philosophie wollen wir diese Begriffsbildung von vornherein etwas allgemeiner fassen. Wir betrachten einen bewerteten Körper \mathbb{K} mit Bewertungskörper \mathbb{K}_0 ; ferner sei V ein normierter \mathbb{K} -Vektorraum (mit Wertekörper \mathbb{K}_0). In den uns am meisten interessierenden Fällen ist $\mathbb{K}_0 = \mathbb{R}$ und $\mathbb{K} = \mathbb{R}$ oder \mathbb{C} .

Definition. Es sei $D \subset \mathbb{K}$ eine nichtleere Teilmenge, $a \in D$ sei ein Häufungspunkt von D , und $f : D \rightarrow V$ sei eine Abbildung. Dann heißt f *differenzierbar* an der Stelle a , wenn der Grenzwert

$$f'(a) := \lim_{\substack{x \rightarrow a \\ x \neq a}} \frac{f(x) - f(a)}{x - a}$$

existiert. $f'(a)$ heißt dann die (erste) *Ableitung* von f an der Stelle a . f heißt *differenzierbar auf D* , wenn f an jeder Stelle a differenzierbar (und damit auch jeder Punkt $a \in D$ Häufungspunkt von D) ist. Die Abbildung $f' : D \rightarrow V$ heißt dann die *erste Ableitungsfunktion* oder auch kurz die *erste Ableitung* der Abbildung f . Man schreibt auch suggestiver (nach LEIBNIZ)

$$df/dx \quad \text{bzw.} \quad \frac{df}{dx}$$

anstelle von f' und entsprechend für $f'(a)$ auch

$$\frac{df}{dx}(a) \quad \text{oder} \quad \left. \frac{df}{dx} \right|_{x=a}.$$

Es ist üblich, den Bruch

$$\frac{\Delta f}{\Delta x} := \frac{f(x) - f(a)}{x - a}, \quad x \neq a,$$

als einen *Differenzen-Quotienten* und entsprechend

$$\frac{df}{dx}(a)$$

als *Differential-Quotienten* an der Stelle a zu bezeichnen. Es ist manchmal ratsam und wirkungsvoll, den Differential-Quotienten an einer beliebigen, aber festen Stelle $x \in D$ zu betrachten und dann den Differenzen-Quotienten in der folgenden Form zu schreiben:

$$\frac{df}{dx}(x) = \lim_{\substack{h \rightarrow 0 \\ h \neq 0}} \frac{f(x+h) - f(x)}{h}.$$

Warnung und Bemerkung. Trotz dieser Bezeichnung darf der Differentialquotient df/dx nicht als *Quotient* der Grössen df und dx , die für sich ohnehin keinen Sinn ergeben, aufgefaßt werden, auch wenn wir den Ausdruck oft *formal* wie einen Quotienten behandeln dürfen, wie wir noch sehen werden. Es ist allerdings möglich, im Rahmen der *nonstandard Analysis* auch die *infinitesimalen* Grössen df und dx mathematisch exakt einzuführen und ihren Quotienten zu bilden, womit die Leibnizsche Schreibweise, ungeachtet ihrer genialen Praktikabilität, eine endgültige Rechtfertigung erfährt.

Beispiel. Ist $f(x) \equiv c \in V$, so gilt $f'(x) \equiv 0$. Ebenso leicht bestimmt man für $V = \mathbb{K}$ und $f = \text{id} : \mathbb{K} \rightarrow \mathbb{K}$ die Ableitung zu $f' = 1$.

Bemerkung. Zur Erinnerung rekapitulieren wir die *exakte* Definition des Grenzwertes $f'(a)$: Zu jedem $\varepsilon > 0$ gibt es ein $\delta > 0$, so daß für alle $x \in D$ mit $0 < |x - a| < \delta$ die Ungleichung

$$\left\| f'(a) - \frac{f(x) - f(a)}{x - a} \right\| < \varepsilon$$

besteht. Für Körper \mathbb{K} mit der Bedingung (*) ist die Existenz des Grenzwertes, wie wir früher schon gezeigt haben, äquivalent zur Existenz *aller* Grenzwerte

$$\lim_{j \rightarrow \infty} \frac{f(x_j) - f(a)}{x_j - a} \text{ für beliebige Folgen } x_j \in D, x_j \neq a, \text{ und } \lim x_j = a,$$

die dann automatisch alle übereinstimmen müssen.

Man macht sich die Bedeutung des Differentialquotienten am einfachsten klar in der speziellen Situation, daß $V = \mathbb{K}$ und $\mathbb{K} = \mathbb{R}$. Hier ist der Differenzenquotient natürlich nichts anderes als die *Steigung* einer Geraden in \mathbb{R}^2 , nämlich der *Sekante* durch die Punkte $(a, f(a))$ und $(x, f(x))$, und der Differentialquotient ist die „Grenzsteigung“ bei dem Grenzübergang $x \rightarrow a$, die wir als Steigung der *Tangente* an den Graphen von f an der Stelle $(a, f(a))$ interpretieren werden (siehe Figur 12.4 und Satz 1). – Wir geben sogleich eine entsprechende Definition in unserer allgemeinen Situation.

Definition und Bemerkung. Es sei $f : D \rightarrow V$ im Punkte $a \in D \subset \mathbb{K}$ differenzierbar. Dann wird durch

$$\mathbb{K} \ni t \mapsto (t, f(a) + (t - a)f'(a)) \in \mathbb{K} \times V$$

eine ausgezeichnete Gerade in $\mathbb{K} \times V$ durch den Punkt $(a, f(a))$ beschrieben, die aus den *Sekanten*

$$\mathbb{K} \ni t \mapsto \left(t, f(a) + (t - a) \frac{f(x_j) - f(a)}{x_j - a} \right) \in \mathbb{K} \times V$$

durch den Grenzübergang $D \ni x_j \rightarrow a$ entsteht. Sie heißt die *Tangente* an den Graphen von f an der Stelle $(a, f(a))$. Manchmal bezeichnet man, wenn auch nicht völlig korrekt, die affine Funktion $t \mapsto f(a) + (t - a)f'(a)$ als die *Tangentengleichung* von f an der Stelle a . Aus dem nächsten Satz geht hervor (Satz 1. iii) und Folgerung 2), daß die Tangente unter allen Geraden in $\mathbb{K} \times V$, die durch den Punkt $(a, f(a))$ gehen, dadurch eindeutig charakterisiert ist, daß sie den Graphen der Abbildung f an der Stelle $(a, f(a))$ „am besten“ (und das heißt: „besser als linear“) approximiert. Eine andere, *dynamische* Interpretation geben wir im Anschluß an den Beweis von Folgerung 2.

Wir fügen im ersten Satz noch eine zweite äquivalente Formulierung ein, da mit ihrer Hilfe alle Standard-Aussagen über Differenzierbarkeit von Summe und Produkt von Funktionen etc. bequem auf die entsprechenden Aussagen über stetige Funktionen zurückgeführt werden können. Außerdem bildet sie die Grundlage für die Übertragung des Differenzierbarkeitsbegriffes auf Funktionen in mehreren Veränderlichen.

Satz 14.1 *Es seien $a \in D \subset \mathbb{K}$ und $f : D \rightarrow V$ gegeben. Dann sind die folgenden Aussagen äquivalent :*

i) f ist in a differenzierbar, d. h. es existiert

$$f'(a) := \lim_{\substack{x \rightarrow a \\ x \neq a}} \frac{f(x) - f(a)}{x - a}.$$

ii) Es gibt eine in a stetige Funktion $\psi : D \rightarrow V$ mit

$$f(x) = f(a) + (x - a)\psi(x), \quad x \in D.$$

iii) Es gibt eine Konstante $c \in V$, so daß für die durch

$$f(x) = f(a) + (x - a)c + \varphi(x), \quad x \in D$$

definierte Funktion $\varphi : D \rightarrow V$ gilt:

$$\lim_{\substack{x \rightarrow a \\ x \neq a}} \left\| \frac{\varphi(x)}{x - a} \right\| = 0.$$

Die Funktion ψ und die Konstante c sind eindeutig bestimmt, und es gilt:

$$f'(a) = \psi(a) = c.$$

φ ist insbesondere stetig in a mit $\varphi(a) = 0$.

Beweis. i) \implies ii). Die Funktion ψ ist außerhalb von a notwendig von der Gestalt

$$\psi(x) = \frac{f(x) - f(a)}{x - a}.$$

Daß diese stetig, und dann sogar eindeutig, in den Punkt a hinein fortgesetzt werden kann, ist äquivalent zu der Existenz des Differentialquotienten. Insbesondere muß man $\psi(a) = f'(a)$ setzen.

ii) \implies iii). Man setze $\varphi(x) := (x - a)(\psi(x) - \psi(a))$ und $c := \psi(a)$. Dann ist iii) erfüllt wegen

$$\frac{\varphi(x)}{x - a} = \psi(x) - \psi(a) \xrightarrow{x \rightarrow a} 0.$$

Gäbe es eine weitere solche Darstellung mit einer Konstanten \tilde{c} und der Funktion $\tilde{\varphi}(x)$, so wäre wegen

$$\|c - \tilde{c}\| = \left\| \frac{\varphi(x) - \tilde{\varphi}(x)}{x - a} \right\| \xrightarrow{x \rightarrow a} 0$$

notwendig $\tilde{c} = c$ und damit auch $\tilde{\varphi} = \varphi$.

iii) \implies i). Aus

$$\left\| \frac{f(x) - f(a)}{x - a} - c \right\| = \left\| \frac{\varphi(x)}{x - a} \right\| \xrightarrow{x \rightarrow a} 0$$

folgt die Existenz von

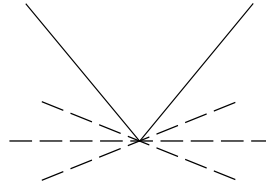
$$f'(a) = \lim_{\substack{x \rightarrow a \\ x \neq a}} \frac{f(x) - f(a)}{x - a} = c.$$

Die Zusätze, außer der letzten Bemerkung, wurden schon weiter vorne mitbewiesen. Ist nun $\varepsilon > 0$ vorgegeben, so folgt aufgrund der Grenzwertaussage in iii) die Existenz eines $\delta > 0$, so daß $|\varphi(x)| < \varepsilon|x - a|$ für alle $x \in D$ mit $|x - a| < \delta$. Dies bleibt natürlich richtig, wenn man δ durch $\min(\delta, 1)$ ersetzt. Für diese x ist dann $|\varphi(x)| < \varepsilon$. \square

Eine unmittelbare Konsequenz aus der Stetigkeit von ψ oder φ ist die

Folgerung 14.2 Ist die Funktion $f : D \rightarrow V$ differenzierbar in a , so ist sie dort auch stetig.

Warnung. Die Umkehrung dieser Aussage ist natürlich nicht richtig, wie schon das einfache Beispiel $f(x) = |x|$, $D = \mathbb{R}$, $a = 0$, zeigt.



Figur 14.2

Um die eingangs gemachten Aussagen über die bestmögliche Approximation des Graphen einer differenzierbaren Funktion durch ihre Tangente zu präzisieren, benötigen wir eine Definition.

Definition und Bemerkung. Es seien $f, g : D \rightarrow V$ Funktionen, $a \in D$ sei ein Häufungspunkt von D . Wir sagen, die Funktionen f und g berühren sich von mindestens erster Ordnung im Punkte a , wenn $f(a) = g(a)$ und

$$\lim_{\substack{x \rightarrow a \\ x \neq a}} \frac{f(x) - g(x)}{x - a} = 0.$$

Dies bedeutet nicht nur, daß die Differenz $f - g$ für $x \rightarrow a$ gegen Null geht (so daß die Bedingung $f(a) = g(a)$ überflüssig ist, sofern nur eine der beiden Funktionen an der Stelle a stetig, insbesondere also dort differenzierbar ist), sondern sogar, daß diese Differenz „schneller als linear“ gegen Null geht. Denn ausgeschrieben lautet unsere Definition insbesondere: Zu jedem $\varepsilon > 0$ existiert ein $\delta > 0$, so daß für alle $x \in D$ mit $|x - a| < \delta$ gilt:

$$\|f(x) - g(x)\| \leq \varepsilon |x - a|.$$

Es ist unmittelbar klar, daß die Relation „Berühren von mindestens erster Ordnung in a “ eine Äquivalenzrelation ist. - Ferner ist eine unmittelbare Konsequenz aus der Definition der Differenzierbarkeit:

Satz 14.3 Sind die beiden Funktionen f, g differenzierbar im Punkte a , so berühren sie sich genau dann von mindestens erster Ordnung im Punkte a , wenn

$$f(a) = g(a) \quad \text{und} \quad f'(a) = g'(a).$$

Bei vorgegebener Funktion f ist die einzige affine Funktion, die diesen beiden Bedingungen genügt, die durch die Tangentengleichung gegebene. - Wir können somit guten Gewissens festhalten:

Folgerung 14.4 Ist die Funktion $f : D \rightarrow V$ im Punkte $a \in D$ differenzierbar, so wird f an der Stelle a von der Tangentengleichung $x \mapsto f(a) + f'(a)(x - a)$ (und unter allen affinen Funktionen nur von dieser) von mindestens erster Ordnung berührt.

Beispiel. Ist f eine Abbildung von $D \subset \mathbb{K}$ nach $V = \mathbb{K}^n$, so läßt sich f schreiben in der Form $f = (f_1, \dots, f_n)$ mit den Komponentenfunktionen f_j , $j = 1, \dots, n$. Ist speziell $D = I \subset \mathbb{R}$ ein Intervall, aufgefaßt als ein Zeitintervall, und $f : I \rightarrow \mathbb{R}^n$, so kann man diese auffassen als die Beschreibung der zeitlichen Bewegung eines Masseteilchens im Raum (unter dem Einfluß von äußeren Kräften). (Man spricht auch von einem (parametrisierten) Weg oder einer Kurve, wenn f zumindest stetig ist). Aus diesem Grunde benutzt man dann auch oft das Symbol t für die Koordinate in D (für „tempus“ lat. Zeit) und (x_1, \dots, x_n) für die Koordinaten in \mathbb{R}^n (oder $(z_1, \dots, z_n) \in \mathbb{C}^n$). Aus dem Kriterium ii) von Satz 1 ergibt sich unmittelbar die folgende Aussage.

Lemma 14.5 Die Abbildung $f = (f_1, \dots, f_n) : D \rightarrow \mathbb{K}^n$, $D \subset \mathbb{K}$, ist genau dann in $t_0 \in D$ differenzierbar, wenn alle Komponentenfunktionen $f_j : D \rightarrow \mathbb{K}$ in t_0 differenzierbar sind, und es gilt

$$f'(t_0) = (f'_1(t_0), \dots, f'_n(t_0)).$$

Wir wollen diese physikalische Vorstellung einer Abbildung $\mathbb{R} \supset D \ni t \mapsto f(t) \in V$ als die zeitliche Bewegung eines Masseteilchens im Raum V auch weiterhin in Anspruch nehmen und in dieser Situation insbesondere auch die *Ableitung* entsprechend interpretieren. Die einfachste solche Bewegung ist *geradlinig* und *gleichförmig*:

$$\mathbb{R} \ni t \mapsto \alpha(t) := x_0 + (t - t_0)v_0$$

mit einem festen *Zeitpunkt* t_0 , zu dem sich das Masseteilchen an der Stelle $x_0 \in V$ befindet. Zur Zeit $t_1 > t_0$ befindet es sich an der Stelle $\alpha(t_1)$ und hat daher in dem Zeitintervall der Länge $\Delta t := t_1 - t_0$ die Strecke

$$\Delta \alpha := \alpha(t_1) - \alpha(t_0) = (t_1 - t_0)v_0 = \Delta t \cdot v_0$$

zurückgelegt. Der Vektor $v_0 \in V$ stellt somit im landläufigen Sinne die (konstante) *Geschwindigkeit* des Masseteilchens dar. Selbstverständlich ist, wie man unmittelbar nachrechnet,

$$v_0 = \alpha'(t) \quad \text{zu jeder Zeit } t \in \mathbb{R}.$$

Im allgemeinen Fall einer Bewegung $t \mapsto f(t) \in V$ stellen wir uns einen festen Zeitpunkt t_0 und einen variablen Zeitpunkt t vor. Damit das Masseteilchen von dem Ort $f(t_0)$ nach $f(t)$ in der Zeit von t_0 nach t in einer geradlinigen und gleichförmigen Bewegung gelangen kann, müßte es, in $f(t_0)$ startend, mit der Geschwindigkeit

$$v_t := \frac{f(t) - f(t_0)}{t - t_0}$$

fliegen. Durch immer feinere Unterteilung des Graphen von f in affine Teilstücke gelangt man dann zwangsläufig zu der Vorstellung der *momentanen* Geschwindigkeit des Masseteilchens zur Zeit t_0 in Form des Grenzwertes

$$\lim_{\substack{t \rightarrow t_0 \\ t \neq t_0}} \frac{f(t) - f(t_0)}{t - t_0},$$

also der ersten Ableitung $f'(t_0)$ der *Bewegungsgleichung* f zur Zeit $t = t_0$.

Bemerkung. Physikalisch zu interpretieren ist diese Größe auch als Geschwindigkeitsvektor einer geradlinigen und gleichförmigen Bewegung $t \mapsto f(t_0) + (t - t_0)f'(t_0)$, die das Teilchen vom Zeitpunkt t_0 ab durchlaufen würde, wenn exakt zu diesem Zeitpunkt sämtliche Kräfte „abgeschaltet“ werden könnten. Man denke hierbei zum Beispiel an einen Hammerwerfer im Moment des „Loslassens“ des Hammers. Ist speziell $V = \mathbb{R}^n$, so bestehen die Komponenten des momentanen *Geschwindigkeitsvektors* gerade aus den Geschwindigkeiten der Projektionen der Bewegung in die einzelnen Koordinatenrichtungen. Wir kommen auf ähnliche Fragen später im Zusammenhang mit der elementaren Kurventheorie zurück.

Die *momentane absolute Geschwindigkeit* ist selbstverständlich der Betrag $\|f'(t_0)\|$. Wir notieren:

Folgerung 14.6 Ist $f : D \rightarrow V$ differenzierbar in $a \in D$, so gilt

$$\lim_{\substack{x \rightarrow a \\ x \neq a}} \left\| \frac{f(x) - f(a)}{x - a} \right\| = \|f'(a)\|.$$

Beweis. Nach der Dreiecksungleichung ist

$$\left| \left\| \frac{f(x) - f(a)}{x - a} \right\| - \|f'(a)\| \right| \leq \left\| \frac{f(x) - f(a)}{x - a} - f'(a) \right\|. \quad \square$$

Beispiel. Ein außerordentlich nützliches, insbesondere für die Einführung der *trigonometrischen Funktionen* bedeutsames Beispiel ist die Abbildung $\mathbb{R} \ni t \mapsto e^{it} \in \mathbb{C}$. Es ist

$$|e^{it}|^2 = e^{it} \overline{e^{it}} = e^{it} e^{-it} = e^{it-it} = e^0 = 1.$$

Also handelt es sich um die Bewegung eines Massenpunktes auf dem Einheitskreis $S^1 = \{z \in \mathbb{C} : |z| = 1\} \subset \mathbb{C} \cong \mathbb{R}^2$, der sich zur Zeit $t = 0$ an der Stelle $e^0 = 1 = (1, 0)$ befindet. Sein Geschwindigkeitsvektor zur Zeit t_0 ist

$$\lim_{t \rightarrow t_0} \frac{e^{it} - e^{it_0}}{t - t_0} = i e^{it_0} \lim_{t \rightarrow t_0} \frac{e^{i(t-t_0)} - 1}{i(t-t_0)} = i e^{it_0}$$

nach Satz 12.11. Schreibt man für den Realteil von e^{it} kurz $c(t)$ und entsprechend $s(t)$ für den Imaginärteil, so ist $e^{it} = c(t) + i s(t)$, wofür man besonders in der angelsächsischen Literatur auch $\text{cis}(t)$ schreibt. Unsere Ableitungsformel besagt dann

$$c'(t) + i s'(t) = \text{cis}'(t) = i \text{cis}(t) = -s(t) + i c(t)$$

und damit

$$s'(t) = c(t), \quad c'(t) = -s(t).$$

Der Geschwindigkeitsvektor $\text{cis}'(t) = (-s(t), c(t))$ steht damit stets senkrecht auf dem *Ortsvektor* $\text{cis}(t) = (c(t), s(t))$ und besitzt die Länge 1. Also bewegt sich das Masseteilchen mit konstanter (absoluter) Geschwindigkeit 1 auf dem Einheitskreis, so daß der in der Zeit t zurückgelegte *Bogen* nicht nur proportional zu, sondern sogar gleich t sein muß. Dies werden wir später noch beweisen. Wir werden ferner weiter unten noch zeigen, daß durch die beiden vorstehenden *Differentialgleichungen* und die *Anfangsbedingungen* $s(0) = 0$, $c(0) = 1$ die Funktionen $c(t)$ und $s(t)$ eindeutig festgelegt sind. Es handelt sich dabei selbstverständlich um die trigonometrischen Funktionen $c(t) = \cos t$ und $s(t) = \sin t$.

Wir müssen als nächstes die üblichen „algebraischen“ Resultate herleiten. Je nach Geschmack kann man sich dabei auf irgendeine der äquivalenten Definitionen stützen. Wir verwenden wegen der besonderen Eleganz der Beweisführung die Charakterisierung ii).

Satz 14.7 *Es seien $f, g : D \rightarrow V$, $c : D \rightarrow \mathbb{K}$ differenzierbar in a . Dann sind dort auch die Funktionen $f + g$, cf differenzierbar, und es gilt*

$$\frac{d(f+g)}{dx}(a) = \frac{df}{dx}(a) + \frac{dg}{dx}(a)$$

$$\frac{d(cf)}{dx}(a) = c(a) \frac{df}{dx}(a) + \frac{dc}{dx}(a) f(a) \quad (\text{Produktregel}).$$

Ferner gilt die Quotientenregel: Ist $c(a) \neq 0$, so ist $\frac{1}{c} f$ nahe a definiert und in a differenzierbar mit

$$\left(\frac{1}{c} f\right)'(a) = \frac{c(a) f'(a) - c'(a) f(a)}{c(a)^2}.$$

Einen Teil der vorstehenden Resultate kann man wie folgt interpretieren.

Folgerung 14.8 *Die in a differenzierbaren Funktionen aus $\text{Abb}(D, V)$ bilden einen \mathbb{K} -Vektorraum, für den die Zuordnung $f \mapsto \frac{df}{dx}(a)$ eine \mathbb{K} -lineare Abbildung ist.*

Bemerkungen und Definition. Insbesondere trägt die Menge $\text{Diff}(D, V)$ der auf ganz D differenzierbaren Abbildungen die kanonische Struktur eines \mathbb{K} -Untervektorraums von $\mathcal{C}^0(D, V)$, dem \mathbb{K} -Vektorraum der *stetigen* Abbildungen von D nach V . (Man beachte, daß die Ableitung einer *konstanten* Funktion identisch verschwindet). Wegen der Produktregel ist $\text{Diff}(D, V)$ sogar ein $\text{Diff}(D, \mathbb{K})$ -Modul. Für jedes Element $f \in \text{Diff}(D, V)$ ist die Ableitung $f' \in \text{Abb}(D, V)$. Die Zuordnung

$$\frac{d}{dx} : \begin{cases} \text{Diff}(D, V) \longrightarrow \text{Abb}(D, V) \\ f \longmapsto f' \end{cases}$$

ist \mathbb{K} -linear. Setzt man voraus, daß die Ableitung sogar *stetig* ist: $f' \in \mathcal{C}^0(D, V)$, so nennt man die Funktion f (*einmal*) *stetig differenzierbar*. Aufgrund der Rechenregeln für die Ableitungen und für stetige Funktionen folgt unmittelbar, daß die Menge der stetig differenzierbaren Funktionen wiederum einen Vektorraum bilden, der stets mit $\mathcal{C}^1(D, V)$ bezeichnet wird. Der Ableitungsoperator kann dann als lineare Abbildung $\mathcal{C}^1(D, V) \rightarrow \mathcal{C}^0(D, V)$ aufgefaßt werden.

Beweis von Satz 3. Beginnen wir mit der Additivität: Aus

$$\begin{aligned} f(x) &= f(a) + (x-a)\psi(x) \\ g(x) &= g(a) + (x-a)\gamma(x) \end{aligned}$$

folgt durch Addition der beiden Zeilen

$$(f+g)(x) = (f+g)(a) + (x-a)(\psi(x) + \gamma(x)),$$

und mit ψ und γ ist auch deren Summe in a stetig. Folglich ist $f+g$ differenzierbar in a mit

$$(f+g)'(a) = \psi(a) + \gamma(a) = f'(a) + g'(a).$$

Der Beweis für die *Produktregel* ist nur unwesentlich schwerer: Wir schreiben statt c wie in der Formulierung des Satzes wieder g und multiplizieren den obigen Ansatz aus. Es ergibt sich diesmal (wobei wir die Reihenfolge beachten sollten):

$$(gf)(x) = (gf)(a) + (x-a)[g(a)\psi(x) + \gamma(x)f(a) + (x-a)\gamma(x)\psi(x)].$$

Die in der eckigen Klammer stehende Funktion $\Psi(x)$ ist in a stetig. Also ist auch das Produkt differenzierbar, und es gilt

$$(gf)'(a) = \Psi(a) = g(a)\psi(a) + \gamma(a)f(a) = g(a)f'(a) + g'(a)f(a).$$

Somit bleibt noch die Quotientenregel zu beweisen. Wegen der Produktregel braucht man die Quotientenregel aber nur für die Funktion c herzuleiten, für die wir wieder g schreiben: Ist nämlich die Differenzierbarkeit von $1/g$ in a gezeigt mit

$$(+)\quad \left(\frac{1}{g}\right)'(a) = -\frac{g'(a)}{g(a)^2},$$

so folgt aus der Produktregel (wir lassen das Argument a fort):

$$\left(\frac{1}{g}f\right)' = \left(\frac{1}{g}\right)'f + \frac{1}{g}f' = -\frac{g'}{g^2}f + \frac{1}{g}f' = \frac{gf' - g'f}{g^2}.$$

Kommen wir also zu (+). Da g in a insbesondere stetig ist, folgt aus $g(a) \neq 0$, daß $1/g$ in einer Umgebung von a definiert und stetig ist. Genauer gibt es ein $\varepsilon_0 > 0$, s. d. $|g(x)| \geq \varepsilon_0$ für alle x in einer Umgebung von a gilt. Hieraus folgt

$$\lim_{\substack{x \rightarrow a \\ x \neq a}} \frac{1}{x-a} \left(\frac{1}{g(x)} - \frac{1}{g(a)} \right) = - \lim_{\substack{x \rightarrow a \\ x \neq a}} \frac{g(x) - g(a)}{g(x)g(a)(x-a)} = - \lim_{\substack{x \rightarrow a \\ x \neq a}} \frac{g(x) - g(a)}{x-a} \lim_{\substack{x \rightarrow a \\ x \neq a}} \frac{1}{g(x)g(a)},$$

sofern die beiden Grenzwerte auf der rechten Seite existieren. Diese sind aber nach Voraussetzung existent und gleich $g'(a)$ bzw. $g(a)^{-2}$. \square

Aus diesem Satz gewinnt man unmittelbar die Erkenntnis, daß *polynomiale Abbildungen* mit Koeffizienten in dem \mathbb{K} -Vektorraum V differenzierbare Abbildungen $\mathbb{K} \rightarrow V$ sind. Normalerweise läßt man hier nur Koeffizienten in \mathbb{K} zu, betrachtet also nur *Polynome*, was aber für den Beweis gänzlich unerheblich ist. Man beachte aber, daß wir die Skalarenmultiplikation im Vektorraum V im Gegensatz zum üblichen Gebrauch erneut von (links und) *rechts* schreiben, um für Polynome nicht gezwungen zu sein, die bekannten Formeln umzuschreiben.

Satz 14.9 *Polynomiale Abbildungen* $P(x) = a_0x^n + \dots + a_n$, $a_j \in V$, sind als Abbildungen $\mathbb{K} \rightarrow V$ differenzierbar mit

$$P'(x) = n a_0 x^{n-1} + \dots + a_{n-1}.$$

Beweis. Da konstante Funktionen eine identisch verschwindende Ableitung besitzen, brauchen wir nach dem vorigen Satz nur die Differenzierbarkeit der Monome $P_n(x) = x^n$ nachzuweisen mit $P'_n(x) = n x^{n-1}$. Dies folgt aber sofort durch vollständige Induktion nach $n \geq 1$ mittels der Produktregel, wobei der Induktionsanfang $P'_1(x) = 1$ völlig elementar ist. Natürlich kann man auch für jedes n direkt argumentieren: Für beliebiges $a \in \mathbb{K}$ gilt

$$\frac{x^n - a^n}{x - a} = x^{n-1} + a x^{n-2} + \dots + a^{n-1}, \quad x \neq a.$$

Da die rechte Seite stetig in a ist, ergibt sich

$$\lim_{\substack{x \rightarrow a \\ x \neq a}} \frac{x^n - a^n}{x - a} = n a^{n-1}. \quad \square$$

Wesentlich interessanter ist die Frage, ob auch *Potenzreihen* in ihrem (offenen) Konvergenzradius differenzierbare Funktionen darstellen. Dies ist in der Tat der Fall, wenn wir (sinnvollerweise) voraussetzen, daß der Bewertungskörper \mathbb{K}_0 der Körper \mathbb{R} der reellen Zahlen ist.

Satz 14.10 *Die Potenzreihe*

$$P(x) = \sum_{n=0}^{\infty} a_n x^n, \quad a_n \in V,$$

besitze den positiven Konvergenzradius $R \in \mathbb{R}_+^* \cup \{\infty\}$. Dann ist P für alle $x \in \mathbb{K}$ mit $|x| < R$ differenzierbar, und es gilt

$$P'(x) = \sum_{n=1}^{\infty} n a_n x^{n-1} = \sum_{n=0}^{\infty} (n+1) a_{n+1} x^n.$$

Mit anderen Worten: Man erhält die Ableitung P' durch (formales) „Differenzieren unter dem Summenzeichen“. Die formal abgeleitete Reihe besitzt sogar den gleichen Konvergenzradius wie die ursprüngliche.

Beweis. Wegen Folgerung 11.11 besitzen die gegebene Potenzreihe und die Potenzreihe

$$P_1(x) = \sum_{n=1}^{\infty} n a_n x^{n-1}$$

den gleichen Konvergenzradius R .

Es sei nun $a \in B(0, R) \subset \mathbb{K}$ beliebig, aber fest gewählt und r eine Zahl mit $|a| < r < R$. Wir setzen ferner

$$f_n(x) := a_n (x^{n-1} + a x^{n-2} + \dots + a^{n-1}).$$

Für $|x| \leq r$ gilt dann

$$\|f_n(x)\| \leq n \|a_n\| r^{n-1}.$$

Da die Reihe $P_1(r)$ absolut konvergent ist, ist die Reihe

$$\sum_{n=1}^{\infty} f_n(x) \quad \text{für } |x| \leq r$$

gleichmäßig konvergent, stellt also insbesondere eine stetige Funktion dar. Es gilt nun für $x \neq a$:

$$\frac{P(x) - P(a)}{x - a} = \sum_{n=0}^{\infty} a_n \frac{x^n - a^n}{x - a} = \sum_{n=1}^{\infty} f_n(x),$$

und somit existiert für $x \rightarrow a$ der Grenzwert

$$P'(a) = \lim_{\substack{x \rightarrow a \\ x \neq a}} \frac{P(x) - P(a)}{x - a} = \lim_{\substack{x \rightarrow a \\ x \neq a}} \sum_{n=1}^{\infty} f_n(x) = \sum_{n=1}^{\infty} \lim_{\substack{x \rightarrow a \\ x \neq a}} f_n(x) = \sum_{n=1}^{\infty} f_n(a) = \sum_{n=1}^{\infty} n a_n a^{n-1} . \quad \square$$

Beispiele. 1. Die Exponentialreihe $\exp x$ ist überall auf \mathbb{C} konvergent und *reproduziert* sich selbst bei Differentiation, wie wir schon wissen:

$$\exp' x = \sum_{n=1}^{\infty} \frac{n}{n!} x^{n-1} = \sum_{n=1}^{\infty} \frac{1}{(n-1)!} x^{n-1} = \exp x .$$

Sie ist eindeutig festgelegt durch ihre Funktionalgleichung und die Ableitung $f'(0) = 1$.

2. Die Reihe

$$L(x) := \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n} x^n$$

besitzt den Konvergenzradius 1 (man verwende hierzu das Quotientenkriterium). Für ihre Ableitung ergibt sich mit Hilfe der geometrischen Reihe

$$L'(x) = \sum_{n=1}^{\infty} (-1)^{n-1} x^{n-1} = \sum_{n=0}^{\infty} (-x)^n = \frac{1}{1+x} .$$

Es ist $L(x) = \ln(x+1)$ (siehe Lemma 12.20).

3. Die auf ganz \mathbb{C} konvergente *Sinus-Reihe*

$$\sin z := \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!} z^{2n+1}$$

besitzt als Ableitung die *Cosinus-Reihe*

$$\cos z = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!} z^{2n} .$$

Definition. Eine Funktion $f : D \rightarrow V$, $D \subset \mathbb{K}$, heißt *n-mal differenzierbar*, $n \geq 2$, wenn sie alle höheren Ableitungen $f^{(1)} := f'$, $f'' := f^{(2)} := (f')'$, \dots , $f^{(n)} := (f^{(n-1)})'$ bis zur Ordnung n besitzt. Sie heißt *n-mal stetig differenzierbar*, wenn sie n -mal differenzierbar und $f^{(n)}$ stetig ist (es sind dann $f, f', \dots, f^{(n-1)}$ automatisch stetig). Es ist unmittelbar einsichtig, daß die Menge der n -mal stetig differenzierbaren Abbildungen $D \rightarrow V$ einen \mathbb{K} -Vektorraum bildet. In Übereinstimmung mit früheren Notationen bezeichnen wir diesen mit $\mathcal{C}^n(D, V)$. f heißt *beliebig oft differenzierbar*, wenn alle Ableitungen $f^{[n]}$, $n \in \mathbb{N}$, existieren (dabei setzt man stets $f^{(0)} := f$). Der entsprechende \mathbb{K} -Vektorraum wird mit $\mathcal{C}^\infty(D, V)$ bezeichnet. Aus der Produktregel folgt übrigens unmittelbar durch Induktion, daß für n -mal (stetig) differenzierbare Funktionen f, g auch das Produkt $f \cdot g$ die Eigenschaft besitzt, n -mal stetig differenzierbar zu sein, und daß die *LEIBNIZ-Formel*

$$(f \cdot g)^{(n)} = \sum_{k=0}^n \binom{n}{k} f^{(k)} g^{(n-k)}$$

gilt. Damit ist die Menge $\mathcal{C}^n(I, \mathbb{R})$ der n -mal stetig differenzierbaren Funktionen $f : I \rightarrow \mathbb{R}$ auf dem nicht leeren Intervall $I \subset \mathbb{R}$ nicht nur ein \mathbb{R} -Vektorraum, sondern sogar eine *kommutative Algebra* (mit der Funktion $f(x) = 1$ als Einselement). Als einfaches Beispiel bestimmt man leicht $f^{(n)}(x)$, $n = 1, 2, 1000$, für die Funktion $f(x) := x^2 e^x$.

Mit diesen Bezeichnungen können wir sofort weiter notieren:

Folgerung 14.11 *Potenzreihen sind beliebig oft differenzierbare Funktionen in ihrem Konvergenzkreis.*

Bemerkung. Es gibt nach WEIERSTRASS reellwertige Funktionen in einer reellen Veränderlichen, die überall stetig, aber nirgends differenzierbar sind. Ein solches Beispiel ist die Funktion G aus Kapitel 12, deren Verlauf in Figur 12.11 skizziert wurde. Durch Integration konstruiert man dann auch differenzierbare Funktionen mit stetiger, aber nirgends differenzierbarer Ableitung. Wir müssen diese eigentümliche Eigenschaft von G noch nachweisen. Da die Funktion $G : \mathbb{R} \rightarrow \mathbb{R}$ periodisch mit Periode 1 ist, brauchen wir nur zu zeigen, daß sie an keiner Stelle $a \in [0, 1]$ differenzierbar ist. Hierzu benutzen wir die Dualbruchentwicklung

$$a = \sum_{k=1}^{\infty} \frac{a_k}{2^k} \quad \text{mit den Abschnitten} \quad A_n := \sum_{k=1}^n \frac{a_k}{2^k} .$$

Es sei ferner $B_n := A_n + 1/2^n$. Dann ist $A_n \leq a \leq B_n$, und die Folgen (A_n) und (B_n) konvergieren gegen a . Nun sind für $k \geq n$ die Zahlen $2^k A_n$ und $2^k B_n$ ganz und folglich die Funktionswerte von $g = g_0$ an diesen Stellen gleich Null. Für $k < n$ ist dagegen

$$g(2^k A_n) = \begin{cases} \sum_{j=k+1}^n \frac{a_j}{2^{j-k}} = g(2^k B_n) - \frac{1}{2^{n-k}} & \text{bzw.} \\ 1 - \sum_{j=k+1}^n \frac{a_j}{2^{j-k}} = g(2^k B_n) + \frac{1}{2^{n-k}} & , \end{cases}$$

je nachdem, ob $a_{k+1} = 0$ oder $= 1$. Also ist für $k < n$

$$\delta_k := \frac{g(2^k A_n) - g(2^k B_n)}{2^{k-n}} = -1 ,$$

wenn $a_{k+1} = 0$, und $= 1$ sonst, und folglich

$$\lim_{n \rightarrow \infty} \frac{G(A_n) - G(B_n)}{A_n - B_n} = \lim_{n \rightarrow \infty} \sum_{k=1}^n \frac{g(2^k A_n) - g(2^k B_n)}{2^{k-n}} = \sum_{k=1}^{\infty} \delta_k$$

wegen $|\delta_k| = 1$ nicht konvergent. Dies widerspricht aber dem folgenden allgemeinen Satz.

Satz 14.12 *Es sei a innerer Punkt des Intervalls I , und die Funktion $f : I \rightarrow \mathbb{R}$ sei in a differenzierbar. Dann gilt*

$$\lim_{j \rightarrow \infty} \frac{f(x'_j) - f(x''_j)}{x'_j - x''_j} = f'(a)$$

für alle Folgen $x'_j, x''_j \in I$ mit $x'_j < a < x''_j$ und $\lim_j x'_j = \lim_j x''_j = a$.

Beweis. Man kann annehmen, daß $f'(a) = 0$ ist; denn anderenfalls gehe man zu der Funktion

$$g(x) := f(x) - (x - a) f'(a)$$

über. Die Behauptung folgt dann sofort aus der Beziehung

$$\frac{f(x'_j) - f(x''_j)}{x'_j - x''_j} = \frac{f(x'_j) - f(a)}{x'_j - a} \frac{x'_j - a}{x'_j - x''_j} + \frac{f(a) - f(x''_j)}{a - x''_j} \frac{a - x''_j}{x'_j - x''_j}$$

und der Tatsache, daß aufgrund der Voraussetzung die Folgen

$$\frac{x'_j - a}{x'_j - x''_j} \quad \text{und} \quad \frac{a - x''_j}{x'_j - x''_j}$$

dem Betrage nach nach oben durch 1 beschränkt sind. □

Dagegen ist im *Komplexen* die Differenzierbarkeit eine weit stärkere Bedingung. Für Funktionen $f : U \rightarrow \mathbb{C}$, $U \subset \mathbb{C}$, schreibt man selbstverständlich für die komplexe Variable im Allgemeinen nicht x , sondern $z \in U$ (das Symbol x ist dem *Realteil* von z vorbehalten). In diesem Fall sollten wir für den Differentialquotienten $f'(a)$ im Leibnizschen Sinne besser

$$\frac{df}{dz}(a)$$

schreiben. Man nennt diese Ableitung daher auch manchmal die *komplexe* Ableitung von f im Punkte a .

Man zeigt in der Funktionentheorie den folgenden tiefliegenden

Satz 14.13 *Es sei $U \subset \mathbb{C}$ eine offene Menge und $f : U \rightarrow \mathbb{C}$ (komplex) differenzierbar. Dann gilt:*

- i) *Die Ableitung f' ist differenzierbar, also automatisch stetig.*
- ii) *f ist sogar beliebig oft differenzierbar.*
- iii) *f ist in jedem Punkt $z_0 \in U$ im größten Kreis $B(z_0, R)$, der noch in U enthalten ist, in eine Potenzreihe entwickelbar.*

Als nächstes wollen wir der Frage nachgehen, wie sich Differenzierbarkeit bei *Komposition* von Funktionen verhält. Dies macht in unserem bisherigen Aufbau nur einen Sinn, wenn wir als Bildraum nur $V = \mathbb{K}$ zulassen (den allgemeineren Fall verschieben wir auf später, wenn wir auch Funktionen auf Teilmengen von Vektorräumen betrachten). Etwas allgemeiner betrachten wir Funktionen $f : D \rightarrow \mathbb{K}$, $D \subset \mathbb{K}$, und $g : E \rightarrow V$, $E \subset \mathbb{K}$, mit $f(D) \subset E$. Dann ist auch $g \circ f : D \rightarrow V$ erklärt, und es ist sinnvoll, nach der Differenzierbarkeit dieser Komposition zu fragen.

Satz 14.14 (Kettenregel) *Es seien f und g wie vorstehend gegeben. Es seien weiter f in a und g in $b = f(a)$ differenzierbar. Dann ist $g \circ f$ in a differenzierbar, und es gilt*

$$(g \circ f)'(a) = f'(a)g'(f(a)).$$

Beweis. Wir schreiben $f(x) = f(a) + (x-a)\psi(x)$, $g(y) = g(b) + (y-b)\gamma(y)$, $b = f(a)$, $y = f(x)$. Dann folgt

$$\begin{aligned} (g \circ f)(x) &= (g \circ f)(a) + (f(x) - f(a))\gamma(f(a) + (x-a)\psi(x)) \\ &= (g \circ f)(a) + (x-a)\psi(x)\gamma(f(a) + (x-a)\psi(x)) \\ &= (g \circ f)(a) + (x-a)\Psi(x) \end{aligned}$$

mit der in a stetigen Funktion $\Psi(x) = \psi(x)\gamma(f(a) + (x-a)\psi(x))$. Wegen $\Psi(a) = \psi(a)\gamma(f(a)) = f'(a)g'(f(a))$ folgt die Behauptung. \square

Beispiele. 1. Die Funktion $h(x) = a^x$ ist die Zusammensetzung von $h(y) = e^y$ mit $y = f(x) = (\ln a)x$. Wegen $(d/dy)e^y = e^y$ und $(d/dx)((\ln a)x) = \ln a$ ist dann

$$\frac{d}{dx} a^x = e^y|_{y=(\ln a)x} \cdot (\ln a) = (\ln a) a^x.$$

2.
$$\frac{d}{dx} (\sin e^x) = e^x \cos e^x.$$

Es ist ebenso von allgemeinem Interesse, wie sich *Umkehrfunktionen* bezüglich Differentiation verhalten. Wir setzen im folgenden voraus, daß $f : D \rightarrow \mathbb{K}$, $D \subset \mathbb{K}$, *injektiv* ist, was z. B. im Fall eines angeordneten Körpers \mathbb{K} richtig ist, wenn f streng monoton wächst oder fällt. Dann existiert die Umkehrabbildung $g : E := f(D) \rightarrow \mathbb{K}$. Wir setzen ferner voraus, daß für einen (Häufungs-) Punkt a von D die Umkehrfunktion g in dem Bildpunkt $b := f(a) \in E$ *stetig* ist. Dies ist z. B. automatisch

erfüllt, wenn $f : I := [a, b] \rightarrow \mathbb{R}$, $I \subset \mathbb{R}$, stetig, insbesondere also differenzierbar ist. Damit g an der Stelle b differenzierbar ist, muß notwendigerweise $f'(a) \neq 0$ sein, denn aus der Kettenregel folgt

$$g'(b) f'(a) = (g \circ f)'(a) = (\text{id})'(a) = 1.$$

Dies reicht als Bedingung aber auch aus, wie der folgende Satz besagt.

Satz 14.15 (Differenzierbarkeit der Umkehrfunktion) *Unter den vorstehenden Voraussetzungen sei f an einer Stelle $a \in D$ differenzierbar mit $f'(a) \neq 0$. Dann ist der Bildpunkt $b = f(a)$ ein Häufungspunkt von $E = f(D)$, und die Umkehrfunktion $g : E \rightarrow \mathbb{K}$ ist an der Stelle $b = f(a)$ differenzierbar mit*

$$g'(b) = \frac{1}{f'(a)}.$$

Beweis. Da $a \in D$ ein Häufungspunkt von D ist, gibt es nichttriviale, gegen a konvergente Folgen (x_j) , $x_j \in D \setminus \{a\}$. Da f in a differenzierbar und damit auch stetig ist, konvergiert die Bildfolge $f(x_j) \in E$ gegen $b = f(a)$. Da weiter nach Voraussetzung f injektiv ist, ist $f(x_j) \neq b$ für alle j . Also ist b ein Häufungspunkt von E . Wir schreiben nun $y = f(x) = f(a) + (x - a)\psi(x)$, wobei $\psi(x)$ in a stetig ist mit $\psi(a) = f'(a) \neq 0$. Wegen der Injektivität von f ist ψ überall von Null verschieden, und wir erhalten für alle $y \in E$ mit $b = f(a)$:

$$g(y) - g(b) = x - a = \frac{y - b}{\psi(g(y))}.$$

Da $\psi(g(y))$ nirgends Null und nach der Stetigkeitsvoraussetzung an g in $b \in E$ stetig ist, folgt die Behauptung aus $\psi(g(b)) = \psi(a) = f'(a)$. \square

Bemerkung. Für komplex differenzierbare Funktionen $f : G \rightarrow \mathbb{C}$, $G \subset \mathbb{C}$, ist die Situation wieder entschieden komfortabler. Hier reicht schon die Bedingung $f'(a) \neq 0$, daß die Funktion f in einer offenen Umgebung $U \subset G$ von a bijektiv, die Bildmenge $f(U)$ offen und die Umkehrabbildung $g : f(U) \rightarrow U$ komplex differenzierbar, insbesondere also stetig ist.

Beispiele. 1. Die Funktion $y \mapsto e^y$ ist injektiv auf \mathbb{R} mit dem Bild $(0, \infty)$; für die Umkehrfunktion $x \mapsto \ln x$ erhalten wir daraus mit Satz 15:

$$(\ln x)' = \frac{1}{\frac{de^y}{dy} \Big|_{y=\ln x}} = \frac{1}{e^{\ln x}} = \frac{1}{x}.$$

Eine entsprechende Formel gilt auch - mit den notwendigen Vorsichtsmaßnahmen - im Komplexen (siehe weiter unten).

2. Die Funktion $y \mapsto y^2$ ist auf dem reellen Intervall $[0, \infty)$ streng monoton wachsend und stetig mit Bild $[0, \infty)$. Da ihre Ableitung außerhalb des Nullpunkts nicht verschwindet, ist die Umkehrfunktion $x \mapsto \sqrt{x}$ differenzierbar auf dem offenen Intervall $(0, \infty)$ mit

$$(\sqrt{x})' = \frac{1}{\frac{dy^2}{dy} \Big|_{y=\sqrt{x}}} = \frac{1}{2\sqrt{x}}.$$

Dies Ergebnis läßt sich auch anders gewinnen und beträchtlich verallgemeinern. Man beachte, daß $\sqrt{x} = x^{1/2}$ und daß wir für positive x und beliebige $r \in \mathbb{R}$ die Potenz x^r durch $\exp(r \ln x)$ definieren konnten. Somit können wir auch die Kettenregel zusammen mit Beispiel 1. in Anspruch nehmen und finden

$$(x^r)' = (\exp(r \ln x))' = \frac{r}{x} \exp(r \ln x) = r \exp((r-1) \ln x) = r x^{r-1}, \quad x > 0, r \in \mathbb{R}.$$

3. Die Funktion $y \mapsto \sin y$ ist auf dem reellen Intervall $[-\pi/2, \pi/2]$ streng monoton wachsend und stetig mit Bild $[-1, 1]$. Die stetige Umkehrabbildung $[-1, 1] \rightarrow [-\pi/2, \pi/2]$ wird, wie wir im vorigen Kapitel schon angemerkt haben, als $\arcsin x$ bezeichnet. Diese Arcusfunktion ist, da die Ableitung des Sinus nur an den Stellen $\pm\pi/2$ verschwindet, auf dem offenen Intervall $(-1, 1)$ differenzierbar mit

$$\begin{aligned} (\arcsin x)' &= \frac{1}{\frac{d \sin y}{dy} \Big|_{y=\arcsin x}} = \frac{1}{\cos(\arcsin x)} = \frac{1}{\sqrt{1 - \sin^2(\arcsin x)}} \\ &= \frac{1}{\sqrt{1 - x^2}}. \end{aligned}$$

Das positive Vorzeichen der Quadratwurzel ergibt sich aus der Tatsache, daß der Arcus-Sinus streng monoton *wächst* (siehe die Ausführungen am Ende des Kapitels).

Weitere Beispiele werden wir weiter unten angeben.

Bemerkung. Unter den Bedingungen des Satzes über die Differenzierbarkeit der Umkehrfunktion sieht man sofort durch vollständige Induktion, daß die Umkehrfunktion $g := f^{-1}$ genauso oft differenzierbar ist wie f selbst. Schreibt man nämlich Inv für die Funktion $\text{Inv}(t) = 1/t$, $t \in \mathbb{R}^*$, so lautet die Ableitungsregel für g einfach $g' = \text{Inv} \circ f' \circ f^{-1}$. Da Inv beliebig oft differenzierbar ist, folgt unsere Behauptung aus der *Kettenregel*. Es ergibt sich damit z. B.

$$g'' = (\text{Inv} \circ f' \circ f^{-1})' = (\text{Inv}' \circ f' \circ f^{-1}) \cdot (f'' \circ f^{-1}) \cdot (f^{-1})'$$

und folglich $(f^{-1})'' = -\frac{1}{[f' \circ f^{-1}]^3} f'' \circ f^{-1}$.

In den folgenden Abschnitten beweisen wir die grundlegenden Sätze der Theorie der differenzierbaren reellwertigen Funktionen auf Intervallen in \mathbb{R} , darunter den *Satz von Rolle*, den *1. und 2. Mittelwertsatz* und die *Taylor-Formel*. Als Konsequenz erhalten wir eine Begründung für das oben schon mehrfach verwendete Argument, daß eine differenzierbare Funktion mit positiver Ableitung auf einem Intervall dort notwendig streng monoton wachsend ist. Ferner sammeln wir noch einige Aussagen über *konvexe* Funktionen. Auch wenn wir die Sätze für beliebige angeordnete Körper formulieren, so werden wir doch im Nachhinein feststellen, daß sie allesamt *nur* für den Körper der reellen Zahlen gültig sind. Wir können damit unseren charakterisierenden Axiomen des reellen Zahlkörpers noch mehr als eine Handvoll weitere hinzufügen.

Es sei also im folgenden \mathbb{K} ein beliebiger angeordneter Körper, und $f : [a, b] \rightarrow \mathbb{K}$ sei eine Funktion auf einem nichttrivialen abgeschlossenen Intervall $[a, b] \subset \mathbb{K}$, $a < b$. Wir sagen für den Rest dieses Paragraphen, f erfülle die Bedingung $(*)_n$, $n \in \mathbb{N}$ fest, wenn f auf dem *abgeschlossenen* Intervall $[a, b]$ mindestens n -mal stetig differenzierbar und die (nach Voraussetzung stetige) n -te Ableitung $f^{(n)}$ noch auf dem *offenen* Intervall (a, b) differenzierbar ist. Dies ist natürlich insbesondere erfüllt, wenn die Funktion f auf einem Intervall I $n+1$ -mal differenzierbar oder sogar stetig differenzierbar ist und a, b in I enthalten sind.

Satz 14.16 *Die folgenden Eigenschaften sind äquivalent.*

- o) \mathbb{K} ist archimedisch und vollständig, also isomorph zu \mathbb{R} .
- i) *Es gilt der Satz von Rolle: Ist $f : [a, b] \rightarrow \mathbb{K}$ eine Funktion mit $(*)_0$ und $f(a) = f(b)$, so gibt es ein $\xi \in (a, b)$ mit $f'(\xi) = 0$.*
- ii) *Es gilt der verallgemeinerte (oder 2.) Mittelwertsatz: Sind $f, g : [a, b] \rightarrow \mathbb{K}$ zwei Funktionen mit $(*)_0$ und ist $g'(x) \neq 0$ für alle $x \in (a, b)$, so ist $g(a) \neq g(b)$, und es gibt ein $\xi \in (a, b)$ mit*

$$\frac{f(b) - f(a)}{g(b) - g(a)} = \frac{f'(\xi)}{g'(\xi)}.$$

- iii) *Es gilt der Mittelwertsatz: Erfüllt $f : [a, b] \rightarrow \mathbb{K}$ die Bedingung $(*)_0$, so existiert ein $\xi \in (a, b)$ mit*

$$\frac{f(b) - f(a)}{b - a} = f'(\xi).$$

- iv) *Es gilt die Taylor-Formel: Erfüllt $f : [a, b] \rightarrow \mathbb{K}$ die Bedingung $(*)_n$, so gibt es für jedes $x \in (a, b]$ ein ξ mit $a < \xi < x$, so daß*

$$f(x) = f(a) + \frac{f'(a)}{1!}(x-a) + \dots + \frac{f^{(n)}(a)}{n!}(x-a)^n + R_n(x),$$

wobei

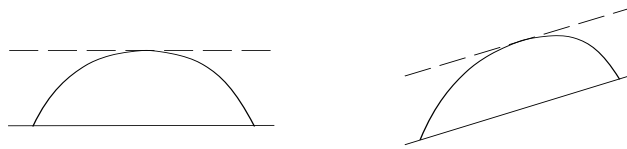
$$R_n(x) = \frac{f^{(n+1)}(\xi)}{(n+1)!}(x-a)^{n+1}$$

das sogenannte Restglied (in der Form von Lagrange) ist.

- v) *Erfüllt f die Bedingung $(*)_n$ und ist $f^{(n+1)} = 0$ auf (a, b) , so ist f ein Polynom höchstens n -ten Grades.*

- vi) *Erfüllt f die Bedingung $(*)_0$ und ist $f' = 0$ auf (a, b) , so ist f konstant.*

Bevor wir den Beweis dieses grundlegenden Satzes durchführen, illustrieren wir den Satz von Rolle und den ersten Mittelwertsatz an einer Zeichnung.



Figur 14.3

Bemerkung. Wir werden die außerordentlich wichtige Interpretation von iv) nach später verschieben, zumal schon die ersten beiden Aussagen und die letzte weittragende Anwendungen zulassen. Der Leser kann sogar, wenn ihm dies aufgrund seiner Vorkenntnisse sinnvoll erscheint, den Beweis von iv) und v) zunächst auslassen und erst nach der eben angekündigten Interpretation in Angriff nehmen. Es sollte klar sein, daß man die Aussage vi), die ebenfalls mannigfache Anwendungen erlaubt, auch direkt aus dem 1. Mittelwertsatz ableiten kann. Der Vollständigkeit halber werden wir dies weiter unten ausführen. Der Schritt von vi) nach o) beruht auf unseren Überlegungen zu dem Begriff des Zusammenhangs in angeordneten Körpern; der Beweis kann bei der ersten Lektüre übergangen werden.

Beweis. o) \implies i). Nach Voraussetzung ist die Funktion f stetig auf $[a, b] \subset \mathbb{R}$, nimmt also ihr Maximum und Minimum an. Wir können annehmen, daß eine der Extremalstellen ξ im offenen Intervall (a, b) liegt, denn sonst ist $f(x) \equiv f(a)$ konstant und $f' = 0$. Es liege ohne Einschränkung bei ξ ein Maximum vor. Dann ist der Differentialquotient

$$\frac{f(x) - f(\xi)}{x - \xi} \begin{cases} \geq 0, & x < \xi, \\ \leq 0, & x > \xi, \end{cases}$$

und somit

$$f'(\xi) = \lim_{x \rightarrow \xi} \frac{f(x) - f(\xi)}{x - \xi} = 0.$$

- i) \implies ii). Betrachte die Funktion

$$F(x) := (f(x) - f(a))(g(b) - g(a)) - (g(x) - g(a))(f(b) - f(a)).$$

F erfüllt $(*)_0$, und es ist $F(a) = F(b) = 0$. Somit existiert ein $\xi \in (a, b)$ mit $F'(\xi) = 0$, also

$$f'(\xi)(g(b) - g(a)) = g'(\xi)(f(b) - f(a)).$$

Wäre $g(a) = g(b)$, so müßte es ein $\eta \in (a, b)$ geben mit $g'(\eta) = 0$ im Widerspruch zu unserer Voraussetzung an g .

ii) \implies iii). Setze $g(x) = x$.

iii) \implies i). Trivial.

i) \implies iv). Zu jedem $a < x \leq b$ gibt es eine eindeutig bestimmte Zahl $\rho(x) \in \mathbb{K}$ mit

$$f(x) = \sum_{j=0}^n \frac{f^{(j)}(a)}{j!} (x-a)^j + \rho(x) \frac{(x-a)^{n+1}}{(n+1)!}.$$

Setze nun bei festem x mit $a < x \leq b$:

$$F(t) := f(x) - \sum_{j=0}^n \frac{f^{(j)}(t)}{j!} (x-t)^j - \rho(x) \frac{(x-t)^{n+1}}{(n+1)!}.$$

Nach Voraussetzung an f ist F stetig für $t \in [a, x]$ und differenzierbar auf (a, x) . Ferner ist $F(x) = 0$ und $F(a) = 0$ nach Definition von $\rho(x)$. Somit existiert ein ξ zwischen a und x , s. d. $F'(\xi) = 0$. Nun ist

$$\begin{aligned} -F'(t) &= f'(t) + \sum_{j=1}^n \left\{ \frac{f^{(j+1)}(t)}{j!} (x-t)^j - \frac{f^{(j)}(t)}{(j-1)!} (x-t)^{j-1} \right\} - \rho(x) \frac{(x-t)^n}{n!} \\ &= \frac{f^{(n+1)}(t)}{n!} (x-t)^n - \rho(x) \frac{(x-t)^n}{n!}. \end{aligned}$$

Wegen $a < \xi < x$ folgt $\rho(x) = f^{(n+1)}(\xi)$.

iv) \implies v). Nach Voraussetzung ist $R_{n+1}(x) = 0$ für alle $x > a$. Somit gilt

$$f(x) = \sum_{j=0}^n \frac{f^{(j)}(a)}{j!} (x-a)^j, \quad x \in [a, b].$$

v) \implies vi). Polynome vom Grad ≤ 0 sind Konstanten.

vi) \implies o). Ist $\mathbb{K} \not\cong \mathbb{R}$, so besitzt \mathbb{K} Intervalle $[a, b]$, die nicht zusammenhängen (siehe Kapitel 12). Somit gibt es *lokal konstante* Funktionen, die aber nicht konstant sind. \square

Wir haben in dem Beweis von o) nach i) einen Zwischenschritt benutzt, der es aufgrund seiner Anwendungsfähigkeit verdient, gesondert hervorgehoben zu werden.

Definition. Eine Stelle $\xi \in D$ heißt eine *Extremalstelle* einer Funktion $f: D \rightarrow \mathbb{K}$, wenn f in ξ einen maximalen bzw. minimalen Wert annimmt. Sie heißt eine *lokale Extremalstelle*, wenn es ein $\varepsilon > 0$ gibt, so daß ξ eine Extremalstelle von f eingeschränkt auf $D \cap \{x \in \mathbb{K} : |x - \xi| < \varepsilon\}$ ist. Im ersten Fall spricht man auch von einer *absoluten Extremalstelle*. Eine absolute oder lokale Extremalstelle ξ heißt *isoliert*, wenn der extreme Wert (in einer Umgebung) nur an der Stelle ξ angenommen wird, wenn also z. B. (im Falle eines absoluten Maximums) $f(x) < f(\xi)$ für alle $x \neq \xi$ gilt.

Aus dem Beweis des vorigen Satzes ergibt sich das nachstehende außerordentlich wichtige *notwendige* Kriterium für eine lokale Extremalstelle.

Lemma 14.17 *Es sei ξ eine lokale Extremalstelle von $f : D \rightarrow \mathbb{K}$, f sei in $\xi \in D$ differenzierbar, wobei ξ sowohl Häufungspunkt von $D^+ = \{x \in D : x > \xi\}$ als auch von $D^- = \{x \in D : x < \xi\}$ sei. Dann gilt*

$$f'(\xi) = 0.$$

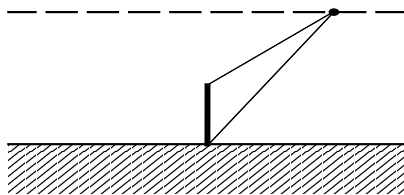
Bemerkungen. 1. Ist $D = I$ ein Intervall, so sind die Voraussetzungen an den Punkt ξ auf jeden Fall erfüllt, wenn er kein *Endpunkt* des Intervalls ist. Wenn dies doch eintritt, so braucht die Konklusion des Satzes nicht zu gelten. Ist z. B. $I = [a, b] \subset \mathbb{R}$ und $f : I \rightarrow \mathbb{R}$ stetig, so nimmt f auf I zwar ein Maximum an, sagen wir bei $\xi \in I$. Dies kann aber einer der Endpunkte sein, wie das Beispiel $f(x) = x$, $\xi = b$, zeigt, und in diesem Fall ist tatsächlich $f'(\xi) \neq 0$.

2. Die Bedingung ist zwar notwendig, i. A. aber nicht *hinreichend*. Das einfachste Gegenbeispiel wird gegeben durch die Funktion $f(x) = x^3$ auf dem Intervall $[-1, 1] \subset \mathbb{R}$. Diese ist streng monoton wachsend und nimmt ihre Extremwerte an den Stellen ± 1 an. Trotzdem ist $f'(0) = 0$. - Wir kommen hierauf später noch einmal zurück.

An dieser Stelle sollen einige Beispiele die Bedeutung von Lemma 17 beleuchten.

Beispiele. 1. Welches Rechteck in der Ebene mit gegebenem Umfang $2L$ besitzt den größten Flächeninhalt? Bezeichnen wir mit x die Länge einer der Seiten, so ist notwendig die andere von der Länge $L - x$ und der Flächeninhalt $F = F(x) = x(L - x)$. Ferner ist $0 \leq x \leq L$ und damit am Rande des Definitionsintervalls $F(0) = F(L) = 0$. Somit muß es eine Extremalstelle im Inneren dieses Intervalls geben. Wegen $F'(x) = L - 2x$ ist die einzige Stelle dieser Art $x = L/2$, und da hier $F(L/2) = L^2/4$ positiv ist, ist dies tatsächlich das absolute Maximum. Mit anderen Worten: Von allen Rechtecken mit gegebenem Umfang besitzt das Quadrat den größten Flächeninhalt. - Viel schwerer ist übrigens die Frage zu beantworten, welche allgemeine Figur gegebenen Umfangs in der Ebene den größten Flächeninhalt besitzt (*isoperimetrisches Problem* oder auch als *Problem der Dido* bekannt); selbstverständlich ist die einzige Lösung die Kreisscheibe mit dem korrekten Umfang.

2. Wir behandeln noch ein etwas anspruchsvolleres Beispiel: Ein Schiff fahre parallel zur Küste im Abstand A . An welcher Stelle sieht man vom Schiff aus eine senkrecht ins Meer ragende Mole der Länge $L < A$ unter maximalem Winkel? Was kann man in den Fällen $L \geq A$ aussagen? Wie verändert sich das Ergebnis, wenn man stattdessen in einem Flugzeug über dem Schiff in der Höhe H fliegt?



Figur 14.4

Es bezeichne x den horizontalen Abstand des Schiffsorts von dem Fußpunkt der Mole. Wenn wir uns (zunächst) auf *positive* x beschränken, so ist der Blickwinkel BW (absolut gemessen) auszudrücken in der Gestalt¹⁸:

$$\text{BW}(x) = \arctan \frac{A}{x} - \arctan \frac{A-L}{x}.$$

Diese Formel ist auch korrekt, wenn $A \leq L$. Ist x negativ, so ist dieser Ausdruck stets negativ, also als *orientierter* Winkel zu interpretieren. Um die „Katastrophen“ zu verstehen, die bei $x \rightarrow 0$ und $A \leq L$ geschehen, ist es überaus sinnvoll, dies als korrekte Definition zu akzeptieren. Allerdings muß man dann

¹⁸Zur Arcus Tangens-Funktion siehe weiter unten.

im negativen Bereich nach *Minima* von BW suchen (was aber wegen $BW(-x) = -BW(x)$ gerade die Negativen der Maxima im positiven Bereich sind).

Kehren wir nun zu dem Fall $x > 0$ und $A > L$ zurück. Da die Arcus Tangens-Funktion für $x \rightarrow \infty$ nach $\pi/2$ strebt und im Nullpunkt stetig mit dem Wert 0 ist (siehe die Ausführungen weiter unten), ergeben sich, was anschaulich ohnehin klar ist, die uneigentlichen Grenzwerte

$$\lim_{x \rightarrow 0} BW(x) = 0, \quad \lim_{x \rightarrow \infty} BW(x) = 0.$$

Somit ist die Funktion BW nach ganz \mathbb{R} stetig fortsetzbar, und da sie auf \mathbb{R}_+^* positiv ist, muß sie wegen der zweiten Grenzwertaussage dort ein Maximum annehmen. Wenn es nun nur genau eine Stelle in diesem Intervall gibt, an der die Ableitung der Funktion verschwindet, so muß dort das (absolute) Maximum vorliegen. Dies ist aber der Fall, da (wegen $\arctan'(x) = (1+x^2)^{-1}$)

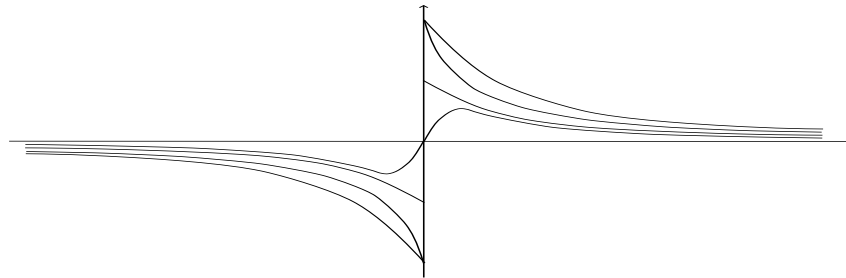
$$BW'(x) = \frac{A-L}{x^2 + (A-L)^2} - \frac{A}{x^2 + A^2}$$

nur an den Stellen mit $x^2 = A(A-L)$ verschwindet, und davon gibt es nur eine im positiven Bereich, nämlich $x_{\max} = \sqrt{A(A-L)}$.

Geht nun A gegen L , so konvergiert x_{\max} gegen Null, und dies ist durchaus der richtige Wert. Denn bei diesem Kurs schrammt der Pott haarscharf an der Katastrophe vorbei. Der Blickwinkel wächst stetig von (fast) Null auf $\pi/2$, und wenn der Kapitän sich nach der Beihnahe-Katastrophe von seinem Schrecken erholt hat, muß er sich umdrehen, um die Mole im Winkel von $\pi/2$ (und Gott sei Dank Schiff und Mole intakt) zu sehen. Dreht er sich nicht um, so sieht er die Mole selbstverständlich vor seinem geistigen Auge mit dem orientierten Winkel $-\pi/2$.

Ist schließlich $A < L$, so kommt es tatsächlich zum Zusammenstoß. Hierbei wächst der Blickwinkel von (fast) Null auf π , und wenn der Kapitän nach der Katastrophe noch in der Lage ist, mathematisch zu denken, so sieht er die (etwas demolierte) Mole nunmehr unter dem (orientierten) Winkel $-\pi$.

Bemerkung. Man stellt bei den obigen Betrachtungen schnell fest, daß es im Wesentlichen nur auf das Verhältnis $\ell := L/A$ ankommt, wir also ohne Einschränkung $A = 1$ setzen und die Molenlänge ℓ als variabel ansehen können. Dann hängt die Funktion BW auch noch von dem Parameter ℓ ab, und das folgende Schaubild gibt alle Möglichkeiten zweifelsfrei wider.

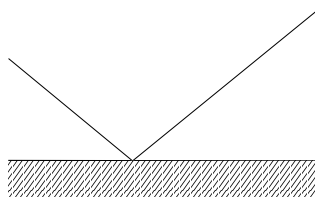


Figur 14.5

3. Es folgen noch zwei Beispiele aus der Mathematischen Physik. In der Hamiltonschen Mechanik beweist man z. B. den Satz, daß ein (System von) Masseteilchen unter geeigneten Voraussetzungen von allen denkbaren Bewegungsbahnen diejenige auswählt, auf denen das *Wirkungsintegral* stationär, also extremal (maximal oder minimal) wird.¹⁹ In der Regel sucht die Natur (jedenfalls lokal) die Bewegung mit *kleinstem* Wirkungsintegral aus. Deswegen formulierte MAUPERTUIS dies auch als das *Gesetz der*

¹⁹Siehe z. B. meinen Vorlesungstext *Differentialgeometrie II*.

kleinsten Wirkung; seine genaue Formulierung stammt von EULER und LAGRANGE. Im einfachsten Fall einer Bewegung in einem homogenen Medium ohne Einwirkung äußerer Kräfte, also bei konstanter Geschwindigkeit, impliziert dieses Gesetz, daß ein Masseteilchen auf dem kürzesten (genauer: einem stationären) Wege zu seinem Ziel gelangen muß oder auch: auf dem schnellsten Wege. In dieser Form stammt das Extremalprinzip von FERMAT. Es führt leicht zu dem mathematisch ableitbaren Satz, daß im (mit einem homogenen Medium angefüllten oder total leeren) euklidischen Raum sich das Licht auf einer Geraden bewegt (siehe auch Kapitel 21), entsprechend in einem gekrümmten Raum auf einer sogenannten *Geodätischen*. Nun kann aber auch in einem homogenen Medium evtl. das Licht nicht nur direkt von einer Lichtquelle zu einem Beobachter gelangen, sondern z. B. indirekt durch Spiegelung an einer glatten Fläche.



Figur 14.6

Verlangen wir auch hier, daß das Licht (immer mit der gleichen Geschwindigkeit c sich bewegend), von allen über den Rand des reflektierenden Mediums verlaufenden (stückweise geraden) Wegen den zeitlich kürzesten auswählt, so benötigt es, von dem Punkt $(0, b_0)$ über $(x, 0)$ nach (a_1, b_1) gelangend, die Zeit L/c mit der Weglänge

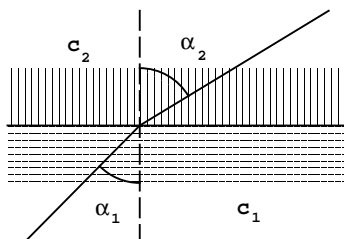
$$L = \sqrt{x^2 + b_0^2} + \sqrt{(a_1 - x)^2 + b_1^2}.$$

Nach Differentiation und Bestimmung des stationären Punktes, den wir mit x_0 bezeichnen wollen, findet man für den *Reflektionspunkt* die Beziehung

$$\frac{x_0}{\sqrt{x_0^2 + b_0^2}} = \frac{a_1 - x_0}{\sqrt{(a_1 - x_0)^2 + b_1^2}}.$$

Mit anderen Worten: Die von dem Lichtstrahl zurückgelegten Strecken bilden die Hypothenusen von zwei kongruenten rechtwinkligen Dreiecken. Somit haben wir das bekannte *Reflektionsgesetz* nachgewiesen: Das Licht bewegt sich bei Reflektion so, daß der Einfallswinkel gleich dem Ausfallwinkel ist.

Ganz entsprechend kann man das *Brechungsgesetz* herleiten. Hier wechselt beim Übergang von einem homogenen Medium zu einem anderen die Lichtgeschwindigkeit von c_1 nach c_2 .



Figur 14.7

Eine analog zu der eben ausgeführten Bestimmung der benötigten Zeit liefert die Funktion

$$T = \frac{\sqrt{x^2 + b_0^2}}{c_1} + \frac{\sqrt{(a_1 - x)^2 + b_1^2}}{c_2}.$$

Die Berechnung des stationären Punktes x_0 verläuft fast genauso wie oben und liefert das SCHNELLI-
USSCHE *Brechungsgesetz*:

$$\frac{c_1}{c_2} = \frac{\frac{\sqrt{x_0^2 + b_0^2}}{x_0}}{\frac{\sqrt{(a_1 - x_0)^2 + b_1^2}}{a_1 - x_0}} = \frac{\sin \alpha_1}{\sin \alpha_2},$$

wobei die Eintritts- und Austrittswinkel α_1, α_2 relativ zu der *Normalen* auf die Reflektionsgrenze bestimmt werden müssen.

Wir wollen die vorstehenden Untersuchungen zur Bestimmung von Extrema noch etwas weiter treiben. Zur Untersuchung des genauen Verlaufs von Funktionen, der aus der Schule unter dem Begriff der „Kurvendiskussion“ geläufig sein sollte, sind (*differentielle*) Kriterien für (strenge) Monotonie und *Konvexität* von großer Nützlichkeit. Wir formulieren die Aussagen für angeordnete Körper \mathbb{K} ; sie sind aber nur für die reellen Zahlen erfüllt, wie man leicht feststellen kann (siehe den Anhang).

Satz 14.18 *Ist $f : I \rightarrow \mathbb{K}$ auf einem nichttrivialen Intervall $I \subset \mathbb{K}$ differenzierbar und ist $f' \geq 0$, so ist f monoton wachsend, sofern der Körper \mathbb{K} dem Mittelwertsatz genügt (also isomorph zu \mathbb{R} ist). Gilt sogar $f'(x) > 0$ für alle $x \in I$, so wächst f streng monoton.*

Beweis. Wegen des 1. Mittelwertsatzes ist $f(x_2) - f(x_1) = f'(\xi)(x_2 - x_1)$ für alle $x_1 < x_2$ in I mit einem geeigneten Element ξ zwischen x_1 und x_2 . Daraus ergibt sich sofort die (strenge) Monotonie von f , wenn $f' \geq 0$ (bzw. $f' > 0$). Umgekehrt sind bei einer monoton wachsenden Funktion f alle Differenzenquotienten notwendig nicht negativ. \square

Bemerkung. Um sicher zu sein, daß sogar $f(a) < f(b)$ ist, brauchen wir nur zu wissen, daß an jeder Stelle x *echt* zwischen a und b die Ableitung der Funktion f positiv ist.

Hieraus gewinnt man sofort die folgende Aussage.

Folgerung 14.19 *Es sei $f : I \rightarrow \mathbb{R}$ eine differenzierbare Funktion und $\xi \in I$ eine Stelle mit $f'(x) \geq 0$ für alle $x < \xi$ und $f'(x) \leq 0$ für alle $x > \xi$. Dann liegt bei ξ ein absolutes Maximum der Funktion f . Sind die vorausgesetzten Ungleichungen überdies strikt, so ist das Maximum sogar isoliert.*

Diese Folgerung impliziert eine hinreichende Bedingung dafür, daß das notwendige Kriterium in Lemma 17 auch tatsächlich hinreichend ist.

Satz 14.20 *Es sei $f : I = (a, b) \rightarrow \mathbb{R}$ differenzierbar mit $f(\xi) = 0$ an einer Stelle ξ mit $a < \xi < b$. Ist dann f sogar zweimal differenzierbar an der Stelle ξ mit $f''(\xi) > 0$ bzw. < 0 , so besitzt f an der Stelle ξ ein lokales isoliertes Minimum (bzw. Maximum).*

Beweis. Es ist nach Voraussetzung

$$\lim_{x \rightarrow \xi} \frac{f'(x)}{x - \xi} = \lim_{x \rightarrow \xi} \frac{f'(x) - f'(\xi)}{x - \xi} = f''(\xi) > 0.$$

Infolgedessen existiert ein $\delta > 0$, so daß $f'(x)/(x - \xi) > 0$ für alle $x \in I$ mit $0 < |x - \xi| \leq \delta$. Damit ist $f'(\xi) < 0$ für $x < \xi$ und $f'(\xi) > 0$ für $x > \xi$, und wir können den vorstehenden Satz anwenden. \square

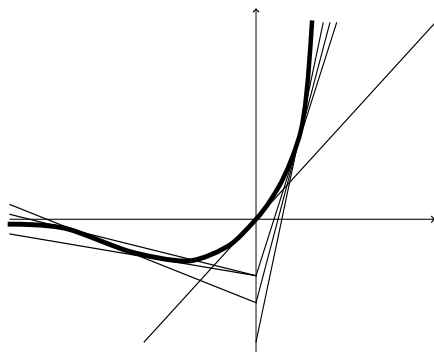
Bemerkung. Man beachte, daß auch die Bedingung $f''(\xi) > 0$ (bzw. < 0) wiederum nur hinreichend, nicht aber notwendig ist, da z. B. die Funktion $f(x) := x^4$ an der Stelle 0 ein globales isoliertes

Minimum besitzt, obwohl $f'(0) = f''(0) = 0$.

Zur Illustration der vorstehenden Überlegungen fügen wir ein Beispiel ein.

Beispiel. Wieviele Tangenten an den Graphen der durch $f(x) = xe^x$ definierten Funktion gehen durch einen vorgegebenen Punkt $(0, Y) \in \mathbb{R}^2$?

Der Verlauf der Funktion f und die Antwort auf die obige Frage sieht wie folgt aus.



Figur 14.8

In der Tat: Die Tangente an den Graphen von f im Punkte $(\xi, \eta) = (\xi, f(\xi))$ besitzt die Gleichung

$$y = f(\xi) + (t - \xi)f'(\xi),$$

in unserem konkreten Fall also

$$y = \xi e^\xi + (\xi + 1)e^\xi(t - \xi).$$

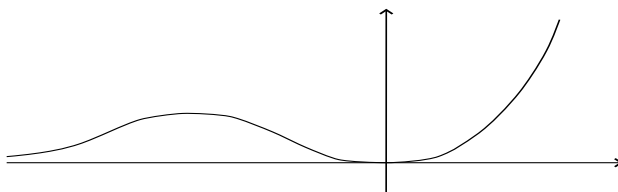
Diese soll durch den Punkt $(0, Y)$ gehen. Wir suchen also zu gegebenem $Y \in \mathbb{R}$ die Anzahl der Lösungen der Gleichung

$$Y = -g(\xi), \quad g(\xi) := \xi^2 e^\xi.$$

Da die Funktion g nicht negativ ist, gibt es mit Sicherheit für $Y > 0$ keine Lösung. Auf der anderen Seite ist (wir ersetzen ξ durch x):

$$\lim_{x \rightarrow -\infty} g(x) = 0, \quad g(0) = 0, \quad \lim_{x \rightarrow \infty} g(x) = \infty.$$

Damit gibt es stets mindestens eine Tangente, wenn $Y < 0$. Eine genauere Untersuchung des Verlaufs der Funktion g ergibt das folgende qualitative Bild:



Figur 14.9

Eine exakte Bestimmung der Stelle des lokalen Maximums von g auf der negativen Achse liefert dann das endgültige Ergebnis:

Intervall	Anzahl der Tangenten
$Y > 0$	0
$Y = 0$	1
$-4e^{-2} < Y < 0$	3
$Y = -4e^{-2}$	2
$Y < -4e^{-2}$	1

Als nächstes wollen wir als Konsequenz aus dem Mittelwertsatz den Beweis des *Zwischenwertsatzes für Ableitungen* einfügen. Selbstverständlich braucht die Ableitung f' einer differenzierbaren Funktion $f: I = [a, b] \rightarrow \mathbb{R}$ nicht stetig zu sein. Ein einfaches Beispiel ist die differenzierbare Funktion $f(x) := x^2 \sin(1/x)$, $x \neq 0$, die durch $f(0) = 0$ stetig und sogar differenzierbar fortgesetzt werden kann; es ist $f'(0) = 0$. Außerhalb des Nullpunkts berechnet sich die Ableitung leicht zu

$$f'(x) = 2x \sin \frac{1}{x} - \cos \frac{1}{x}, \quad x \neq 0.$$

Der erste Summand auf der rechten Seite geht nun zwar gegen Null für $x \rightarrow 0$, aber der zweite besitzt keinen Grenzwert. Folglich ist die Ableitung von f an der Stelle 0 nicht stetig. Dennoch gilt für solche Ableitungen immer der *Zwischenwertsatz*; d. h. mit anderen Worten, daß nicht stetige Ableitungsfunktionen nicht völlig willkürlich vorgegeben werden können.

Lemma 14.21 *Ist $I \subset \mathbb{R}$ ein nichttriviales Intervall und $f: I \rightarrow \mathbb{R}$ eine differenzierbare Funktion, so gibt es für alle $a, b \in I$ mit $a < b$ und alle η zwischen $f'(a)$ und $f'(b)$ ein $\xi \in I$ mit $a < \xi < b$, so daß*

$$f'(\xi) = \eta.$$

Beweis. Wir definieren Funktionen $\varphi_a, \varphi_b: I \rightarrow \mathbb{R}$ durch

$$\varphi_a(x) := \begin{cases} f'(a) & , \quad x = a, \\ \frac{f(a) - f(x)}{a - x} & , \quad a < x \leq b, \end{cases} \quad \text{und} \quad \varphi_b(x) := \begin{cases} \frac{f(b) - f(x)}{b - x} & , \quad a \leq x < b, \\ f'(b) & , \quad x = b. \end{cases}$$

Dies sind stetige Funktionen, deren Bilder also nach dem Zwischenwertsatz kompakte Intervalle sind. Nun ist aber

$$\varphi_a(b) = \varphi_b(a) = \frac{f(b) - f(a)}{b - a}$$

und damit der Durchschnitt der beiden Bildintervalle nicht leer, womit auch deren Vereinigung ein kompaktes Intervall J ist. Ferner liegen $f'(a) = \varphi_a(a)$ und $f'(b) = \varphi_b(b)$ in J , also auch η . Ist dann ohne Beschränkung der Allgemeinheit $\eta \in \varphi_a(I)$, so existiert ein x mit $a < x \leq b$, so daß

$$\eta = \frac{f(a) - f(x)}{a - x}.$$

Nach dem Mittelwertsatz gibt es dann aber ein ξ , das echt zwischen a und x und damit echt zwischen a und b liegt, mit $f'(\xi) = \eta$. \square

Auch *theoretische* Anwendungen von Teil v) von Satz 16 sind Legion. Wir beginnen mit einer Herleitung von v) aus einer oft sehr nützlichen Abschätzung, die sich wiederum sofort aus dem ersten Mittelwertsatz ergibt, so daß wir auf einen Beweis verzichten können.

Lemma 14.22 Die reelle Funktion $f : I = [a, b] \rightarrow \mathbb{R}$ erfülle $(*)_0$, und die erste Ableitung von f sei auf (a, b) beschränkt: $m \leq f'(x) \leq M$, $a < x < b$. Dann gilt für alle x_1, x_2 mit $a \leq x_1 < x_2 \leq b$, daß

$$m(x_2 - x_1) \leq f(x_2) - f(x_1) \leq M(x_2 - x_1).$$

Bemerkungen. 1. Die Aussage v) in Satz 16 ergibt sich hieraus selbstverständlich mit $M = m = 0$.

2. Faßt man die Beschränktheitsaussage von Lemma 22 zusammen in

$$|f'(x)| \leq C, \quad a < x < b,$$

so gewinnt man die folgende Aussage über die mögliche Abweichung zweier Funktionswerte:

$$|f(x_1) - f(x_2)| \leq C|x_2 - x_1| \quad \text{für alle } x_1, x_2 \in I.$$

Man bezeichnet diese Eigenschaft als LIPSCHITZ-*Stetigkeit* von f auf dem Intervall I . Sie verdeutlicht erneut die Stetigkeit einer differenzierbaren Funktion, liefert aber, unter den eben gemachten stärkeren Voraussetzungen, erheblich mehr, nämlich z. B. die *gleichmäßige Stetigkeit* auf dem Intervall I . (Siehe Kapitel 15).

Aus der Aussage v) können wir beispielhaft eine neue Charakterisierung der Exponentialfunktion herausfiltern.

Lemma 14.23 Es sei $f : I \rightarrow \mathbb{R}$ eine differenzierbare Funktion auf einem (nichttrivialen) Intervall $I \subset \mathbb{R}$ mit $f'(x) = cf(x)$, $x \in I$, $c \in \mathbb{R}$ fest. Dann gilt für jeden Punkt $x_0 \in I$:

$$f(x) = Ae^{cx}, \quad A = f(x_0)e^{-cx_0}.$$

Beweis. Betrachte mit der im Satz definierten Konstanten A die Funktion

$$g(x) = f(x)e^{-cx} - A.$$

Dann ist

$$g'(x) = f'(x)e^{-cx} - cf(x)e^{-cx} \equiv 0,$$

also $g(x) = \text{konstant}$ und $g(x) \equiv 0$ wegen $g(x_0) = 0$. □

Eine weitere Anwendung betrifft die Charakterisierung der trigonometrischen Funktionen \cos und \sin , die wir schon am Anfang des Kapitels erwähnt hatten. Den Verlauf dieser Funktionen skizzieren wir im Anschluß an diese Überlegungen.

Satz 14.24 Es gibt eindeutig bestimmte differenzierbare Funktionen $c, s : \mathbb{R} \rightarrow \mathbb{R}$, die den folgenden Bedingungen genügen :

$$c' = -s, \quad s' = c, \quad c(0) = 1, \quad s(0) = 0.$$

Für diese Funktionen ist notwendig

$$c(t) + is(t) = e^{it},$$

also insbesondere $c^2 + s^2 = 1$. Sie besitzen die in ganz \mathbb{R} konvergente Potenzreihenentwicklung

$$c(t) = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!} t^{2n}, \quad s(t) = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!} t^{2n+1}.$$

Beweis. Wir setzen $\text{cis}(t) := c(t) + is(t)$ und $f(t) := \text{cis}(t)e^{-it}$. Nach Voraussetzung und der Produktregel ist die Funktion f differenzierbar mit

$$f'(t) = (-s(t) + ic(t))e^{-it} + (c(t) + is(t))(-ie^{-it}) = 0.$$

Damit sind sowohl der Real- als auch der Imaginärteil von f differenzierbar, und beide besitzen eine identisch verschwindende Ableitung. Also ist auch $f'(t) = \text{const.} = f'(0) = 1$ und folglich $\text{cis}(t) = e^{it}$, also $c(t) = \text{Re } e^{it}$, $s(t) = \text{Im } e^{it}$. Damit ist $c^2 + s^2 = 1$, was man auch aus der Produktregel herleiten kann:

$$(c^2 + s^2)' = 2(cc' + ss') = 2(-cs + sc) = 0.$$

Die Potenzreihenentwicklungen ergeben sich aus der Entwicklung der Exponentialfunktion. □

Bemerkung. Es ist also notwendig

$$c(t) = \cos t, \quad s(t) = \sin t, \quad t \in \mathbb{R}.$$

Ferner gilt

$$c(t) = \text{Re } e^{it} = \frac{e^{it} + \overline{e^{it}}}{2} = \frac{e^{it} + e^{-it}}{2} \quad \text{und} \quad s(t) = \text{Im } e^{it} = \frac{e^{it} - e^{-it}}{2i}.$$

Diese Formeln sind nur Spezialfälle der für alle $z \in \mathbb{C}$ gültigen EULERSchen *Formeln*

$$\cos z = \frac{e^{iz} + e^{-iz}}{2} \quad \text{und} \quad \sin z = \frac{e^{iz} - e^{-iz}}{2i},$$

aus denen man durch Addition die fundamentale Beziehung

$$\cos z + i \sin z = e^{iz}$$

ableitet. Man rechnet ebenso leicht nach, daß auch für komplexe Werte z die Beziehung

$$\sin^2 z + \cos^2 z = 1$$

besteht.

Aus den Eulerschen Formeln und dem Additionstheorem der Exponentialfunktion kann man auch die *Additionstheoreme* der trigonometrischen Funktionen gewinnen.

Satz 14.25 Für alle $z, w \in \mathbb{C}$ bestehen die Relationen :

$$\sin(z+w) = \sin z \cos w + \cos z \sin w, \quad \cos(z+w) = \cos z \cos w - \sin z \sin w.$$

Insbesondere gelten die Verdoppelungsformeln

$$\sin 2z = 2 \sin z \cos z, \quad \cos 2z = \cos^2 z - \sin^2 z = 1 - 2 \sin^2 z.$$

Den leichten *Beweis* überlassen wir den Leser/inne/n. Für reelle Zahlen ist der Beweis noch einfacher:

$$c(x+h) + i s(x+h) = \text{cis}(x+h) = e^{i(x+h)} = e^{ix} e^{ih} = (c(x) + i s(x))(c(h) + i s(h)). \quad \square$$

Man kann übrigens auch die trigonometrischen Funktionen ähnlich wie die Exponentialfunktion durch ihre Additionstheoreme und die Vorgabe der Ableitung im Nullpunkt charakterisieren. Beachtet man, daß

$$\lim_{z \rightarrow 0} \frac{\sin z}{z} = \lim_{z \rightarrow 0} \frac{\sin z - \sin 0}{z} = \sin'0 = \cos 0 = 1$$

und

$$\lim_{z \rightarrow 0} \frac{\cos z - 1}{z} = \lim_{z \rightarrow 0} \frac{\cos z - \cos 0}{z} = \cos'0 = -\sin 0 = 0,$$

so hat man schon die erste Implikationsrichtung in dem folgenden Satz eingesehen.

Satz 14.26 Es gibt genau zwei Funktionen $c, s: \mathbb{R} \rightarrow \mathbb{R}$, die den Additionstheoremen

$$c(x+h) = c(x)c(h) - s(x)s(h), \quad s(x+h) = s(x)c(h) + c(x)s(h)$$

genügen, und für die

$$\lim_{h \rightarrow 0} \frac{c(h) - 1}{h} = 0, \quad \lim_{h \rightarrow 0} \frac{s(h)}{h} = 1$$

gilt.

Beweis. Wir zeigen, daß bei Richtigkeit der Additionstheoreme und den Zusatzbedingungen die Funktionen c und s überall differenzierbar sind und gemäß Satz 24 die „richtigen“ Ableitungen besitzen:

$$c'(x) = \lim_{h \rightarrow 0} \frac{c(x+h) - c(x)}{h} = \lim_{h \rightarrow 0} \frac{c(x)(c(h) - 1) - s(x)s(h)}{h} = -s(x),$$

und entsprechend für die Ableitung von $s(x)$. Insbesondere sind c und s stetig im Nullpunkt, so daß sich aus den vorgegebenen Grenzwerten notwendig auch die richtigen Anfangsbedingungen ergeben:

$$c(0) - 1 = \lim_{h \rightarrow 0} (c(h) - 1) = 0 \quad \text{und} \quad s(0) = \lim_{h \rightarrow 0} s(h) = 0. \quad \square$$

Man kann die trigonometrischen Funktionen auch allein für sich durch das Erfüllen einer Differentialgleichung charakterisieren. Diese ist notwendig von zweiter Ordnung; es handelt sich dabei um die für die Physik zentrale *Schwingungsgleichung*.

Satz 14.27 Die Differentialgleichung $y'' + y = 0$ besitzt (auf ganz \mathbb{R}) genau eine Lösung $y = c(x)$ mit $c(0) = 1$ und $c'(0) = 0$ und genau eine Lösung $y = s(x)$ mit $s(0) = 0$ und $s'(0) = 1$. Selbstverständlich ist $s(x) = \sin x$ und $c(x) = \cos x$.

Beweis. Nach Voraussetzung ist eine beliebige Lösung c zweimal differenzierbar mit $c'' = -c$. Infolgedessen ist c sogar beliebig oft differenzierbar, wie man durch Differenzieren dieser Gleichung und vollständige Induktion einsieht. Wir setzen $s(x) := -c'(x)$. Dann erfüllt das Paar (c, s) das System von Differentialgleichungen erster Ordnung

$$c' = -s, \quad s' = -c'' = c$$

und besitzt die Anfangsbedingungen $c(0) = 1$, $s(0) = c'(0) = 0$. Nach Satz 23 ist notwendig $c = \cos$ und $s = \sin$. \square

Bemerkung. Mit einer leichten Verfeinerung der Argumente kann man sofort beweisen, daß der gesamte Lösungsraum der Schwingungsgleichung $y'' + y = 0$ aus allen Linearkombinationen $A \cos x + B \sin x$ besteht. Insbesondere ist er ein Vektorraum der Dimension 2 mit der Basis $\cos x, \sin x$. Dies ist ein Spezialfall der Lösungstheorie von linearen Differentialgleichungen n -ter Ordnung mit konstanten Koeffizienten, die wir später in Kapitel 22 in voller Allgemeinheit entwickeln werden. Wir werden dort unter anderem zeigen, daß die Lösungen solcher Gleichungen in Potenzreihen entwickelt werden können, also analytisch sind, und eindeutig bestimmt sind durch ihre Werte und Ableitungen bis zur $n - 1$ -ten Ordnung an einer festen Stelle. Man kann somit zum Lösen solcher Gleichungen stets einen Potenzreihen-Ansatz machen, der in unserem obigen konkreten Beispiel sofort auf die Cosinus- bzw. Sinus-Reihe führt.

Wir schieben jetzt noch einige weitere Untersuchungen zum Verlauf der *trigonometrischen Funktionen* im Reellen - und ihrer Interpretation als *Kreisfunktionen* - und der Exponentialfunktion im Komplexen ein. Gemäß unserer Einführung der trigonometrischen Funktionen als Reihen dürfen wir zunächst noch kein Schulwissen verwenden. Immerhin wissen wir aber schon, daß für reelles t die Werte $\cos t$ und $\sin t$ die Koordinaten des Punktes e^{it} auf dem Einheitskreis sind. Um den Anschluß an unser Schulwissen zu gewinnen, muß es uns gelingen, die reelle Zahl t als geeignetes Maß für den

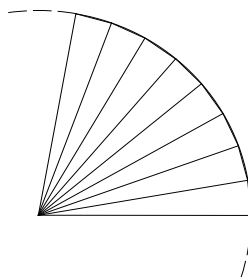
Winkel zwischen dem Vektor $(\cos t, \sin t)$ und der reellen Achse zu identifizieren. Physikalisch ist dies absolut klar: Die Bewegung $t \mapsto e^{it}$ verläuft auf dem Einheitskreis S^1 mit konstanter absoluter Geschwindigkeit; zur Zeit $t > 0$ hat also ein Masseiteilchen, wenn es zur Zeit $t = 0$ im Punkte $e^0 = 1$ gestartet ist, einen Weg der Länge t auf einem Kreisbogen zurückgelegt. Wir werden dies später noch einmal allgemein im Rahmen der Längenbestimmung von Kurven nachprüfen. Wir können aber auch direkt diesen Gedankengang einsichtig machen. Stellen wir uns also wieder die Bewegung eines Masseiteilchens nach dem obigen Bewegungsgesetz zwischen der Zeit 0 und t vor. Unterteilen wir dann das Zeitintervall in n gleiche Teile, so durchläuft das Masseiteilchen nacheinander die Punkte

$$\zeta_j := \exp\left(j \frac{it}{n}\right), \quad j = 0, \dots, n.$$

Offensichtlich ist aber $\zeta_j - \zeta_{j-1}$ dem Betrage nach gleich $|e^{it/n} - 1|$, also unabhängig von j . Damit legt das Teilchen, wenn man seinen Weg durch den entsprechenden Polygonzug durch die Punkte ζ_j ersetzt, eine Gesamtstrecke von

$$n \left| e^{it/n} - 1 \right| = t \left| \frac{e^{it/n} - 1}{it/n} \right|$$

zurück. Die tatsächlich zurückgelegte Strecke sollte der Grenzwert sein bei $n \rightarrow \infty$. Dieser ist aber, da dann it/n nach 0 geht, gleich t , wie nicht anders zu erwarten war. Wir werden also im folgenden die reelle Zahl t als *Bogenmaß* des Winkels interpretieren, der von den Zahlen $z = e^{it}$ und $1 = e^{i0}$ eingeschlossen wird, wobei wir zulassen, daß dieser Winkel *mehrdeutig* ist, wir also mehrfach um den Einheitskreis herumlaufen, bevor wir an der Stelle z ankommen.



Figur 14.10

Wir zeigen nun als erstes, daß der Cosinus im Intervall $[0, 2]$ eine Nullstelle besitzen muß. Dies folgt aus dem Zwischenwertsatz und $\cos 0 = 1$, wenn wir nachweisen können, daß $\cos 2 < 0$. Das letztere ist aber eine Konsequenz aus der Potenzreihenentwicklung des Cosinus und der Fehlerabschätzung im Leibniz-Kriterium: Die ersten beiden Terme in der Entwicklung des Cosinus um Null an der Stelle 2 ergeben den Wert

$$1 - \frac{2^2}{2!} = -1,$$

und der Fehler ist kleiner als der Absolutbetrag des nächsten Gliedes, also $2^4/4! = 2/3 < 1$. Wegen des Identitätssatzes für Potenzreihen besitzt der Cosinus im Reellen sogar eine kleinste positive Nullstelle, die man mit $\pi/2$ bezeichnet²⁰. Somit ist $\cos x > 0$ für alle $x \in [0, \pi/2)$, und wegen $\sin' = \cos$ ist der Sinus *streng monoton wachsend* auf dem *abgeschlossenen* Intervall $[0, \pi/2]$. Damit ist der Sinus außerdem positiv auf dem halboffenen Intervall $(0, \pi/2]$ und besitzt den Wert 1 an der Stelle $\pi/2$. Aus der Beziehung $\cos' = -\sin$ folgt weiterhin, daß der Cosinus in dem betrachteten abgeschlossenen Intervall *streng monoton fallend* ist. Aus den Potenzreihenentwicklungen liest man selbstverständlich unmittelbar ab, daß der Sinus eine *ungerade* und der Cosinus eine *gerade* Funktion ist:

$$\sin(-z) = -\sin z, \quad \cos(-z) = \cos z.$$

²⁰Wir werden später nachweisen müssen, daß diese Zahl tatsächlich ein Viertel des Umfangs des Einheitskreises ausmacht.

Außerdem impliziert das Additionstheorem für den Sinus und den Cosinus die Relationen

$$\sin(z + \pi/2) = \cos z, \quad \cos(z + \pi/2) = -\sin z.$$

Damit ist der Verlauf beider Funktionen auf ganz \mathbb{R} vollständig bestimmt durch den Verlauf von einer von beiden auf dem Intervall $[0, \pi/2]$. Insbesondere ergibt sich

$$\sin(z + \pi) = -\sin z, \quad \cos(z + \pi) = -\cos z$$

und damit die *Periodizität* der Funktionen:

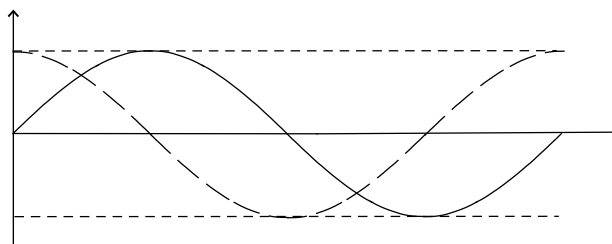
$$\sin(z + 2\pi) = \sin z, \quad \cos(z + 2\pi) = \cos z.$$

Da aber die Sinusfunktion im Reellen (und auch, wie wir unten zeigen werden, im Komplexen) genau die Nullstellen $0, \pm\pi, \pm2\pi, \dots$ besitzt, kann sie keine kleineren Perioden als 2π besitzen. Mit anderen Worten: Es gilt für eine Zahl p genau dann $\sin(z + p) = \sin z$ für alle $z \in \mathbb{C}$, wenn $p = 2n\pi$, $n \in \mathbb{Z}$. Die entsprechende Aussage ist auch für den Cosinus richtig.

Als Konsequenz aus den Werten der trigonometrischen Funktionen an den Vielfachen von π erhält man weiter die berühmten Eulerschen Formeln

$$e^{2\pi i} = 1, \quad e^{\pi i} = -1.$$

Der Verlauf der Funktionen sieht damit im Intervall $[0, 2\pi]$ in etwa wie folgt aus (der Graph des Cosinus ist „gestrichelt“ gezeichnet):



Figur 14.11

Als spezielle Werte hat man

$$\sin \pi/4 = \sin(\pi/2 - \pi/4) = \cos(-\pi/4) = \cos \pi/4, \quad \text{also} \quad \sin \pi/4 = \cos \pi/4 = \frac{1}{2}\sqrt{2}.$$

Die Werte an den Stellen $\pi/6$ und $\pi/12$ kann man rein algebraisch wie folgt bestimmen. Mit $\sin 2x = 2 \sin x \cos x$ und $\cos 2x = \cos^2 x - \sin^2 x = 1 - 2 \sin^2 x$ gewinnt man die Gleichung dritten Grades

$$\begin{aligned} \sin 3x &= \sin(2x + x) = \sin 2x \cos x + \cos 2x \sin x = 2 \sin x \cos^2 x + \sin x (1 - 2 \sin^2 x) \\ &= 2 \sin x (1 - \sin^2 x) + \sin x (1 - 2 \sin^2 x) = 3 \sin x - 4 \sin^3 x, \end{aligned}$$

die man genauso einfach auch aus den Eulerschen Formeln herleiten kann²¹. Setzt man hier $x = \pi/3$, so findet man bei anschließender Verwendung der Verdoppelungsformeln schnell alle gesuchten Werte. Aus reiner Freude am Rechnen gehen wir noch ein wenig anders vor. Wir setzen zuerst $S := \sin \pi/6$ und erhalten mit der obigen Gleichung mit $x = \pi/6$ wegen $\sin \pi/2 = 1$ die Beziehung

$$4S^3 - 3S + 1 = 0.$$

²¹Mit dieser Gleichung zeigt man übrigens im Rahmen der Galois-Theorie, daß die Dreiteilung des allgemeinen Winkels mit Zirkel und Lineal unmöglich ist.

Eine offensichtliche Lösung dieser Gleichung ist $S = 1/2$. Nun ist

$$4S^3 - 3S + 1 = (S - 1/2)(4S^2 + 2S - 2),$$

und das rechts stehenden quadratische Polynom hat die Nullstellen $1/2$ und -1 . Da aber $\sin \pi = -1$ nicht für S in Frage kommt, ist notwendig $\sin \pi/6 = S = 1/2$ und folglich $\cos \pi/6 = \sqrt{1 - (1/2)^2} = \sqrt{3}/2$. Weiter ist

$$\sin(x + \pi/4) = \sin((x - \pi/4) + \pi/2) = \cos(x - \pi/4),$$

und folglich wegen $1/4 + 1/12 = 1/3$, $1/4 - 1/12 = 1/6$ auch $\sin \pi/3 = \cos \pi/6 = \sqrt{3}/2$ und $\cos \pi/3 = 1/2$. Schließlich erhält man mit einer der Verdoppelungsformeln die Gleichung

$$1 - 2 \sin^2 \pi/12 = \cos \pi/6 = \frac{\sqrt{3}}{2},$$

aus der unmittelbar

$$\sin \pi/12 = \frac{\sqrt{2 - \sqrt{3}}}{2} = \frac{\sqrt{6} - \sqrt{2}}{4}, \quad \cos \pi/12 = \frac{\sqrt{2 + \sqrt{3}}}{2} = \frac{\sqrt{6} + \sqrt{2}}{4}$$

kommt.

Der *Tangens* eines Bogens oder Winkels ist definiert als das Verhältnis von Sinus zu Cosinus (und der *Cotangens* als dessen Kehrwert):

$$\tan x := \frac{\sin x}{\cos x}, \quad \cot x := \frac{\cos x}{\sin x}.$$

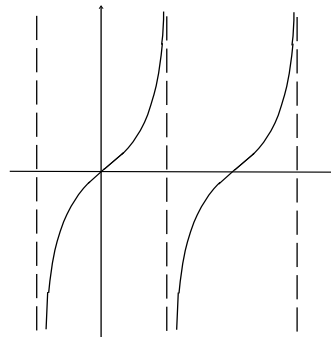
Diese Funktionen sind im Reellen nur erklärt an den Stellen, an denen der Cosinus bzw. der Sinus nicht verschwinden, also im Falle des Tangens für $x \neq \pi/2 \pm n\pi$. Man überzeugt sich leicht davon, daß beide Funktionen sogar die Periode π besitzen. Da der Tangens eine ungerade Funktion ist, ist sein Verlauf vollständig bestimmt durch seine Werte auf dem Intervall $[0, \pi/2)$. Die Ableitung bestimmt sich nach der Quotientenregel sofort zu

$$\tan' x = \frac{\cos x \cos x + \sin x \sin x}{\cos^2 x} = \frac{1}{\cos^2 x} = 1 + \tan^2 x,$$

ist insbesondere also überall positiv. Damit ist der Tangens im Reellen eine streng monoton wachsende Funktion; und es gilt:

$$\lim_{x \rightarrow \pi/2} \tan x = \infty, \quad \lim_{x \rightarrow -\pi/2} \tan x = -\infty.$$

Somit ist sein Bild ganz \mathbb{R} , und sein Verlauf sieht wie folgt aus:



Figur 14.12

Die Umkehrabbildung

$$\arctan : \mathbb{R} \longrightarrow (-\pi/2, \pi/2)$$

wird wiederum als *Arcus-Tangens* bezeichnet. Aus der Gleichung

$$\tan' x = 1 + \tan^2 x .$$

und dem Satz über die Differentiation der Umkehrfunktion gewinnen wir ebenso leicht wie im Falle des Arcussinus:

$$\frac{d \arctan x}{dx} = \frac{1}{1 + x^2} .$$

Man beachte, daß auch hier wie beim Logarithmus die Ableitung eine *rationale Funktion* ist.

Die letzte Gleichung kann sofort dazu benutzt werden, um die Potenzreihenentwicklung des \arctan durch Integration aus der geometrischen Reihe mit $q := -x^2$ zu gewinnen (siehe auch Kapitel 12). Dagegen kann man die Potenzreihenentwicklung für den Tangens selbst nicht so einfach berechnen. Man beschreitet hierzu im Allgemeinen den folgenden Weg, der über die sogenannten *BERNOULLI-Zahlen* führt. Zunächst betrachtet man die Funktion

$$f(z) := \frac{e^z - 1}{z}, \quad z \neq 0,$$

die sich offensichtlich nach 0 durch $f(0) := 1$ stetig fortsetzen läßt. Aus der Potenzreihenentwicklung der Exponentialfunktion gewinnt man unmittelbar die für alle $z \in \mathbb{C}$ gültige Darstellung

$$f(z) = \sum_{k=0}^{\infty} \frac{z^k}{(k+1)!} .$$

Die Funktion $g := 1/f$ läßt sich dann ebenfalls um den Nullpunkt entwickeln; man setzt diese Reihe meist in der Form

$$g(z) = \frac{z}{e^z - 1} = \sum_{k=0}^{\infty} \frac{B_k}{k!} z^k$$

mit den Bernoulli-Zahlen B_k an. Mit den geläufigen Rekursionsformeln ergeben sich dann sofort die Beziehungen

$$B_0 = 1, \quad \sum_{j=0}^n \frac{B_{n-j}}{(n-j)!} \frac{1}{(j+1)!} = 0, \quad n \geq 1,$$

aus denen man die Bernoulli-Zahlen rekursiv berechnen kann. Man findet z. B.:

$$B_1 = -\frac{1}{2}, \quad B_2 = \frac{1}{6}, \quad B_3 = 0, \quad B_4 = -\frac{1}{30}, \quad B_6 = \frac{1}{42}, \quad B_8 = -\frac{1}{30}, \quad B_{10} = \frac{5}{66}, \quad B_{12} = -\frac{691}{2730}$$

und $B_{14} = 7/6$. Selbstverständlich sind alle B_k rationale Zahlen, und die B_k mit *ungeradem* Index $k \geq 3$ müssen alle verschwinden. Die letzte Behauptung folgt sofort aus der Beziehung

$$g(z) - B_1 z = B_0 + \sum_{k=2}^{\infty} \frac{B_k}{k!} z^k = \frac{z}{e^z - 1} + \frac{z}{2} = \frac{z}{2} \left(\frac{e^{z/2} + e^{-z/2}}{e^{z/2} - e^{-z/2}} \right) = \frac{z \cosh(z/2)}{2 \sinh(z/2)} .$$

Ersetzt man in dieser Formel z durch $2z$, so gewinnt man die Beziehung

$$z \cosh z = \sinh z \cdot \sum_{k=0}^{\infty} \frac{B_{2k}}{(2k)!} (2z)^k$$

und hieraus durch Einsetzen der bekannten Potenzreihenentwicklungen für \cosh und \sinh

$$z \sum_{k=0}^{\infty} \frac{z^{2k}}{(2k)!} = \left(\sum_{k=0}^{\infty} \frac{z^{2k+1}}{(2k+1)!} \right) \left(\sum_{k=0}^{\infty} \frac{B_{2k}}{(2k)!} (2z)^k \right) .$$

Ein analoger Ansatz

$$z \cos z = \sin z \cdot G(z)$$

mit der notwendig *geraden* Funktion

$$G(z) = \sum_{k=0}^{\infty} (-1)^k \frac{B_{2k}^*}{(2k)!} (2z)^{2k}$$

führt dann zwangsläufig zu dem gleichen Rekursionssystem wie für die Bernoulli-Zahlen. Es ist somit $B_{2k}^* = B_{2k}$ und weiter

$$z \cot z = z \frac{\cos z}{\sin z} = \sum_{k=0}^{\infty} (-1)^k \frac{2^{2k} B_{2k}}{(2k)!} z^{2k}.$$

Mit der trigonometrischen Umformung

$$z \tan z = z \frac{\sin^2 z}{\sin z \cos z} = z \left(\frac{\cos z}{\sin z} - 2 \frac{\cos^2 z - \sin^2 z}{2 \sin z \cos z} \right) = z (\cot z - 2 \cot 2z)$$

erhält man schließlich die gewünschte Potenzreihenentwicklung des Tangens in der Form

$$\tan z = \sum_{n=1}^{\infty} (-1)^{n-1} \frac{2^{2n} (2^{2n} - 1) B_{2n}}{(2n)!} z^{2n-1}.$$

An dieser Stelle ist es angebracht, noch einige Bemerkungen über den Verlauf der Exponentialfunktion im Komplexen zu machen. Da der reelle Cosinus alle Werte zwischen -1 und 1 annimmt und der Sinus mit einem Wert auch sein Negatives, durchläuft die *Kurve*

$$\mathbb{R} \ni t \mapsto (\cos t, \sin t) \in \mathbb{C}$$

alle Punkte des Einheitskreises

$$S^1 := \{ z = x + iy \in \mathbb{C} : |z| = 1 \}.$$

Zu jeder komplexen Zahl $w \neq 0$ gibt es daher (nicht eindeutig bestimmte) reelle Zahlen t mit

$$\frac{w}{|w|} = e^{it} = \cos t + i \sin t,$$

d. h.

$$w = |w| e^{it}.$$

Man nennt dies die Darstellung von w in *Polarkoordinaten*. Selbstverständlich ist hierin der „Winkel“ t (gemessen im Bogenmaß) nur bis auf additive Vielfache der Periode 2π eindeutig bestimmt.

Man erkennt hiermit sofort, daß die komplexe Exponentialfunktion jeden komplexen Wert außer Null tatsächlich annimmt. Ist nämlich $w \neq 0$ vorgegeben und $w = |w| \exp(it)$ die Polarkoordinatendarstellung, so wähle man einen reellen Logarithmus r von $|w|$, also $e^r = |w|$, und bilde

$$z = r + i(t + 2n\pi), \quad n \in \mathbb{Z}.$$

Dann ist wegen der oben angegebenen Eulerschen Formel, wie gewünscht,

$$e^z = e^r e^{it} = |w| e^{it} = w.$$

Insbesondere ist aber die Exponentialfunktion im Komplexen nicht injektiv. Die Urbilder eines gegebenen Wertes $w \neq 0$ sind genau die eben angegebenen abzählbar unendlich vielen Werte z . - Dies liegt an dem folgenden

Lemma 14.28 *Es gilt genau dann $e^{z_1} = e^{z_2}$, wenn $z_2 - z_1 \in 2\pi i \mathbb{Z}$.*

Beweis. Die Bedingung $e^{z_1} = e^{z_2}$ ist wegen der Funktionalgleichung der Exponentialfunktion gleichbedeutend mit $e^{z_1 - z_2} = 1$. Wir brauchen daher nur die $z \in \mathbb{C}$ zu bestimmen mit $e^z = 1$. Schreiben wir $z = x + iy$, so ist diese Bedingung äquivalent zu

$$e^x(\cos y + i \sin y) = 1, \quad \text{also } e^x \cos y = 1, \quad e^x \sin y = 0.$$

Die zweite Forderung führt sofort auf $y \in \pi\mathbb{Z}$. Von diesen sind aber die Werte $\pm\pi, \pm 3\pi$ etc. nicht zulässig, da dort der Cosinus negativ ist. Somit verbleiben die Werte $y \in 2\pi\mathbb{Z}$; an diesen Stellen ist $\cos y = 1$, und der einzig verbleibende mögliche x -Wert ist $x = 0$. \square

Aufgrund dieser Überlegungen ist die komplexe Exponentialfunktion auf jedem Streifen $\{z \in \mathbb{C} : a < \operatorname{Im} z < a + 2\pi\}$ injektiv und läßt in \mathbb{C} nur die Werte auf der Halbgeraden $\{r e^{ia} : r \geq 0\}$ aus. Die zugehörigen Umkehrfunktionen nennt man auch *Zweige* des Logarithmus und bezeichnet sie einheitlich mit $\log z$. Speziell für $a = -\pi$ erhält man einen Zweig, der auf \mathbb{C}^* ohne die negative reelle Halbachse erklärt ist und auf der positiven reellen Achse mit dem reellen Logarithmus übereinstimmt. Man nennt diesen oft auch den *Hauptzweig* des Logarithmus und schreibt hierfür manchmal auch $\operatorname{Log} x$:

$$\log : \mathbb{C} \setminus \{r e^{ia} : r \geq 0\} \longrightarrow \{z \in \mathbb{C} : a < \operatorname{Im} z < a + 2\pi\}.$$

Für die (komplexe) Ableitung aller Zweige des Logarithmus folgt wie im reellen Fall

$$\frac{d \log z}{dz} = \frac{1}{z}.$$

Bemerkung. Mit den Eulerschen Formeln kann man leicht auch die Nullstellen des komplexen Sinus und Cosinus berechnen. Es stellt sich heraus, daß beide Funktionen nur die schon bekannten *reellen* Nullstellen besitzen. Somit ist der Tangens auch im Komplexen außer an den Stellen $\pm\pi/2, \pm 3\pi/2$ etc. definiert. Für seine Ableitung gewinnt man die gleichen Formeln wie im Reellen:

$$\tan' z = \frac{1}{\cos^2 z} = 1 + \tan^2 z.$$

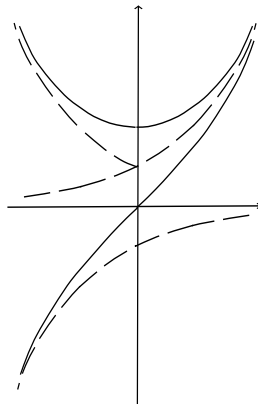
Weitere wichtige Funktionen, die einen gesonderten Namen erhalten, sind die sogenannten *Hyperbelfunktionen*. Im Grunde sind sie jedoch nur Bestandteile der (reellen) Exponentialfunktion, genauer: der *gerade* und der *ungerade* Anteil derselben. In Formeln also:

$$\cosh x := \frac{e^x + e^{-x}}{2}, \quad \sinh x := \frac{e^x - e^{-x}}{2}.$$

Man nennt sie auch den *hyperbolischen Cosinus* bzw. *Sinus* oder auch lat. *cosinus hyperbolicus* bzw. *sinus hyperbolicus*. Selbstverständlich gilt

$$\cosh x = \cosh(-x) = \sum_{n=0}^{\infty} \frac{x^{2n}}{(2n)!} \quad \text{und} \quad \sinh x = -\sinh(-x) = \sum_{n=0}^{\infty} \frac{x^{2n+1}}{(2n+1)!}.$$

Der Graph des Cosinus hyperbolicus wird auch als *Kettenlinie* bezeichnet, da er (unter idealisierenden Voraussetzungen) die Form einer hängenden Kette unter dem Einfluß der Schwerkraft modelliert. Der Verlauf der beiden Funktionen kann der folgenden Skizze entnommen werden (zum Vergleich sind Teile der Graphen der Funktionen $(1/2)e^{\pm x}$ und $-1/2e^{-x}$ gestrichelt eingezeichnet).



Figur 14.13

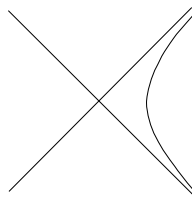
Die Bezeichnung der Funktionen als Hyperbel-Funktionen beruht auf der leicht zu beweisenden Tatsache, daß die Kurve

$$\mathbb{R} \ni t \mapsto (x, y) := (\cosh t, \sinh t)$$

gerade die Punkte des rechten „Astes“ der *Hyperbel*

$$x^2 - y^2 = 1$$

durchläuft.



Figur 14.14

Selbstverständlich setzt man die Hyperbelfunktionen durch ihre Potenzreihenentwicklungen ins Komplexe fort. Es zeigt sich dann, daß sie eigentlich nichts anderes sind als die komplexen Cosinus- und Sinus-Funktionen, in denen man nur Urbild- und/oder Bildraum um 90° „gedreht“ hat:

$$\cosh(iz) = \cos z, \quad \sinh(iz) = i \sin z.$$

Man rechnet leicht die folgenden *Additionstheoreme* nach:

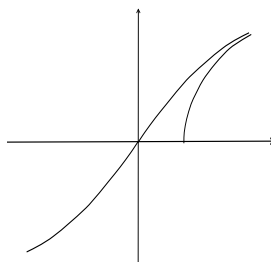
$$\sinh(z+w) = \sinh z \cosh w + \cosh z \sinh w, \quad \cosh(z+w) = \cosh z \cosh w + \sinh z \sinh w.$$

Ferner definiert man den hyperbolischen Tangens und Cotangens durch

$$\tanh z = \frac{\sinh z}{\cosh z}, \quad \coth z = \frac{\cosh z}{\sinh z},$$

wobei im letzteren Fall $z \neq 0$ vorausgesetzt werden muß.

Kehrt man wieder ins Reelle zurück, so sieht man, daß \sinh auf ganz \mathbb{R} , \cosh auf dem Intervall $x \geq 1$, \tanh wiederum auf ganz \mathbb{R} und \coth auf \mathbb{R}^* Umkehrfunktionen besitzen; man bezeichnet diese als *Area-Funktionen* arsinh , arcosh etc.. Eine anschauliche Begründung mit Hilfe von *Flächeninhalten* (= lat. *area*) werden wir später geben. Die ersten beiden Funktionen verlaufen wie folgt:



Figur 14.15

Wie der Cosinus und der Sinus lassen sich auch die Hyperbelfunktionen durch ihre Ableitungen und die entsprechende Anfangsbedingung charakterisieren. Es gilt insbesondere

$$\cosh' z = \sinh z, \quad \sinh z = \cosh' z$$

und

$$\tanh' z = \frac{1}{\cosh^2 z} = 1 - \tanh^2 z, \quad \coth' z = \frac{-1}{\sinh^2 z} = 1 - \cosh^2 z.$$

Hiermit berechnet man die Ableitungen der Areafunktionen im Reellen zu

$$\operatorname{arcosh}' x = \frac{1}{\sqrt{x^2 - 1}}, \quad \operatorname{arsinh}' x = \frac{1}{\sqrt{x^2 + 1}}, \quad \operatorname{artanh}' x = \frac{1}{1 - x^2}, \quad \operatorname{arcoth}' x = \frac{1}{x^2 - 1}.$$

Wir geben noch einige Beispiele zur Differentiation an.

Beispiele.

$$\begin{aligned} \frac{d}{dx} \ln(\ln x) &= \frac{1}{x \ln x}, & x > 1, \\ \frac{d}{dx} (1 + x^2)^{\sin x} &= (1 + x^2)^{\sin x} \left(\frac{2x \sin x}{1 + x^2} + \ln(1 + x^2) \cdot \cos x \right), & x \in \mathbb{R}, \\ \frac{d}{dx} \left(\frac{1+x}{1-x} \right)^{x^2} &= \left(\frac{1+x}{1-x} \right)^{x^2} 2x \left(\ln \frac{1+x}{1-x} + \frac{x}{1-x^2} \right), & |x| < 1, \\ \frac{d}{dx} \frac{\sqrt{x} \sin x}{\ln x} &= \frac{(2x \cos x + \sin x) \ln x - 2 \sin x}{2\sqrt{x} (\ln x)^2}, & x > 0, x \neq 1, \\ \frac{d}{dx} \sqrt{e^{\sin \sqrt{x}}} &= \frac{\cos \sqrt{x}}{4\sqrt{x}} \sqrt{e^{\sin \sqrt{x}}}, & x > 0. \end{aligned}$$

Im dritten Fall ist z. B.

$$\left(\frac{1+x}{1-x} \right)^{x^2} = \exp \left(x^2 \ln \frac{1+x}{1-x} \right),$$

woraus das Ergebnis unter mehrfacher Benutzung der Produkt- und Kettenregel abgelesen werden kann, wenn man noch

$$\left(\ln \frac{1+x}{1-x} \right)' = \frac{(1-x) + (1+x)}{(1-x)^2 \frac{1+x}{1-x}} = \frac{2}{1-x^2}$$

berücksichtigt.

Nach diesem Abstecher in den Bereich der klassischen transzendenten Funktionen kehren wir ein weiteres Mal zu theoretischen Überlegungen zurück.

Man gelangt unter Verwendung des verallgemeinerten Mittelwertsatzes zu einem Resultat, das ein merkwürdiges „Zwitterdasein“ führt. Zum einen scheint es eine eher zufällige Beobachtung zu sein, zum anderen aber führt es zu tiefliegenden Erkenntnissen über Grenzwerte von Funktionen. Man nennt die Formeln auch die HOSPITAL-schen *Regeln* (auch wenn der Marquis de l'Hospital- oder DE L'HÔPITAL - wahrscheinlich das „copyright“ für sie einem der Bernoullis „abgekauft“ hat).

Satz 14.29 *Es seien $f, g : (a, b) \rightarrow \mathbb{R}$ differenzierbare Funktionen auf dem Intervall $(a, b) \subset \overline{\mathbb{R}} := \mathbb{R} \cup \{\infty, -\infty\}$, d. h. $-\infty \leq a < b \leq \infty$, es sei $g'(x) \neq 0$ für alle $x \in (a, b)$, und es gelte entweder $\lim_{x \searrow a} f(x) = \lim_{x \searrow a} g(x) = 0$ oder $\lim_{x \searrow a} |g(x)| = \infty$. Existiert dann*

$$A := \lim_{x \searrow a} \frac{f'(x)}{g'(x)}$$

im eigentlichen oder uneigentlichen Sinne (d. h. $A \in \overline{\mathbb{R}}$), so existiert auch (im eigentlichen oder uneigentlichen Sinne) der Grenzwert

$$\lim_{x \searrow a} \frac{f(x)}{g(x)} = A.$$

Eine entsprechende Aussage gilt auch für den Grenzübergang $x \nearrow b$.

Bemerkung. Man beachte, daß man im ersten Fall auch für $a \in \mathbb{R}$ wegen $\lim_{x \searrow a} g(x) = 0$ die einfachen algebraischen Grenzwertsätze nicht anwenden darf.

Beweis. Wir behandeln nur einige der möglichen Fälle²². Der einfachste ist der, daß $a \in \mathbb{R}$ und $\lim_{x \rightarrow a+} f(x) = \lim_{x \rightarrow a+} g(x) = 0$. Dann gibt es für beliebiges $x \in (a, b)$ ein ξ mit $a < \xi < x$, s. d.

$$\frac{f(x)}{g(x)} = \frac{f(x) - f(a)}{g(x) - g(a)} = \frac{f'(\xi)}{g'(\xi)},$$

wobei wir f und g nach a durch $f(a) = g(a) = 0$ stetig fortgesetzt haben. Mit x geht dann auch ξ gegen a , womit die Behauptung aus dem Folgenkriterium folgt.

Den Beweis unter der Voraussetzung $\lim_{x \rightarrow a} |g(x)| = \infty$ führen wir nur für $a, A \in \mathbb{R}$ und überlassen dem Leser die Mühe, sich die Einzelheiten für $a, A = \pm \infty$ selbst zu erarbeiten. Es sei

$$\lim_{x \rightarrow a} \frac{f'(x)}{g'(x)} = A \in \mathbb{R}.$$

Dann gibt es zu jedem $\varepsilon > 0$ ein $\delta_1 > 0$, s. d.

$$\left| \frac{f'(\xi)}{g'(\xi)} - A \right| < \frac{\varepsilon}{2}$$

für alle ξ mit $0 < \xi - a < \delta_1$. Liegen x und y mit $x < y$ in diesem Bereich, so ist nach dem verallgemeinerten Mittelwertsatz auch

$$\left| \frac{f(x) - f(y)}{g(x) - g(y)} - A \right| < \frac{\varepsilon}{2}.$$

Nun ist (bei festem y) nach Voraussetzung

$$\frac{f(x)}{g(x)} = \frac{f(x) - f(y)}{g(x) - g(y)} \frac{1 - \frac{g(y)}{g(x)}}{1 - \frac{f(y)}{f(x)}} \xrightarrow{x \rightarrow a} \frac{f(x) - f(y)}{g(x) - g(y)}.$$

²²Einen vollständigen Beweis findet man z. B. in [9], Satz 50.1.

Hält man also y fest und wählt man $0 < \delta < y - a < \delta_1$ geeignet, so ist

$$\left| \frac{f(x)}{g(x)} - \frac{f(y)}{g(y)} \right| < \frac{\varepsilon}{2}$$

für alle x mit $0 < x - a < \delta$. Somit ist für diese x

$$\left| \frac{f(x)}{g(x)} - A \right| < \varepsilon. \quad \square$$

Mit diesen Regeln lassen sich unglaublich viele interessante Grenzwerte bestimmen. Zuerst leiten wir in den folgenden Beispielen zwei der uns schon bekannten Grenzwertformeln erneut her.

Beispiele. 1. $\lim_{x \rightarrow 0} \frac{e^x - 1}{x} = \lim_{x \rightarrow 0} \frac{e^x}{1} = 1.$

2. $\lim_{x \rightarrow 0} \frac{\sin x}{x} = \lim_{x \rightarrow 0} \frac{\cos x}{1} = 1.$

3. $\lim_{x \rightarrow 0} \frac{\tan x}{x} = \lim_{x \rightarrow 0} \frac{1 + \tan^2 x}{1} = 1.$

4. $\lim_{x \rightarrow \infty} \frac{e^{ax}}{x} = \lim_{x \rightarrow \infty} \frac{a e^x}{1} = \infty$ für positives a .

Hieraus ergibt sich sofort per vollständiger Induktion:

5. $\lim_{x \rightarrow \infty} \frac{e^x}{x^n} = \lim_{x \rightarrow \infty} \frac{e^x}{n x^{n-1}} = \infty.$

Mit anderen Worten: Die *Exponentialfunktion* wächst stärker als jedes *Polynom* für $x \rightarrow \infty$.

Wegen $e^{ax}/x^b = (e^{(a/b)x}/x)^b$ bei positiven a, b impliziert 4. eine noch stärkere Aussage:

6. $\lim_{x \rightarrow \infty} \frac{e^{ax}}{x^b} = \infty$ für alle $a, b > 0$.

In Worten: Jede noch so kleine (positive) Potenz von e^x geht für $x \rightarrow \infty$ schneller gegen Unendlich als jede noch so große Potenz von x .

Hieraus folgt unmittelbar

7. $\lim_{x \rightarrow \infty} P(x) e^{-ax} = 0$ für jedes Polynom $P(x)$ und jedes $a > 0$.

8. $\lim_{x \rightarrow \infty} \frac{\ln x}{x^a} = \lim_{x \rightarrow \infty} \frac{1/x}{a x^{a-1}} = \lim_{x \rightarrow \infty} \frac{1}{a x^a} = 0$ für jedes positive a .

Hieraus ergibt sich wie beim Übergang von 4. zu 6.:

9. $\lim_{x \rightarrow \infty} \frac{(\ln x)^b}{x^a} = 0$ für alle positiven a, b .

In Worten: Jede noch so große Potenz von $\ln x$ geht für $x \rightarrow \infty$ langsamer gegen Unendlich als jede noch so kleine (positive) Potenz von x .

Den Grenzübergang $x \searrow 0$ kann man durch die Substitution $y = 1/x$ auf den Grenzübergang $y \rightarrow \infty$ zurückführen. Hiermit gewinnt man z. B. aus 9. die Aussage:

10. $\lim_{x \rightarrow 0^+} x^a (\ln x)^n = \lim_{y \rightarrow \infty} \frac{(-\ln y)^n}{y^a} = 0$ für $a > 0$ und $n \in \mathbb{N}$.

$$11. \lim_{x \rightarrow 0^+} x^x = \lim_{x \rightarrow 0^+} \exp(x \ln x) = \exp\left(\lim_{x \rightarrow 0^+} x \ln x\right) = e^0 = 1.$$

In den folgenden beiden Beispielen muß man die Regeln sogar zweimal anwenden.

$$12. \lim_{x \rightarrow 0} \frac{1 - \cos \frac{x}{2}}{1 - \cos x} = \lim_{x \rightarrow 0} \frac{\frac{1}{2} \sin \frac{x}{2}}{\frac{1}{2} \sin x} = \lim_{x \rightarrow 0} \frac{\frac{1}{4} \cos \frac{x}{2}}{\cos x} = \frac{1}{4}.$$

$$13. \lim_{x \rightarrow 0} \left(\frac{1}{\sin x} - \frac{1}{x} \right) = \lim_{x \rightarrow 0} \frac{x - \sin x}{x \sin x} = \lim_{x \rightarrow 0} \frac{1 - \cos x}{x \cos x + \sin x} = \lim_{x \rightarrow 0} \frac{\sin x}{x(-\sin x) + 2 \cos x} = 0.$$

Zum Amusement des Lesers hier noch einige weitere Beispiele, die dem Buch [124] entnommen sind.

$$14. \lim_{x \rightarrow 1^-} \frac{\pi/2 - \arcsin x}{\sqrt{1-x}} = \lim_{x \rightarrow 1^-} \frac{\frac{1}{\sqrt{1-x^2}}}{1} = \lim_{x \rightarrow 1^-} \frac{2}{\sqrt{1+x}} = \sqrt{2}.$$

$$15. \lim_{x \rightarrow \infty} \frac{\ln(x^m + 1)}{\ln x^n} = \lim_{x \rightarrow \infty} \frac{m x^{m-1}}{x^m + 1} \frac{x^n}{n x^{n-1}} = \lim_{x \rightarrow \infty} \frac{m x^m}{n(x^m + 1)} = \frac{m}{n}, \quad m, n \in \mathbb{N}^*.$$

$$16. \lim_{x \rightarrow 1^-} \sin(\pi x) \ln(1-x) = \lim_{x \rightarrow 1^-} \frac{\ln(1-x)}{\frac{1}{\sin(\pi x)}} = \lim_{x \rightarrow 1^-} \frac{\sin^2(\pi x)}{(1-x)\pi \cos(\pi x)} = \frac{-1}{\pi} \lim_{x \rightarrow 1^-} \frac{\sin^2(\pi x)}{1-x} =$$

$$\frac{-1}{\pi} \lim_{x \rightarrow 1^-} \frac{2\pi \sin(\pi x) \cos(\pi x)}{-1} = 0.$$

$$17. \lim_{x \rightarrow 0} \frac{\ln(1 + \arctan x)}{x} = \lim_{x \rightarrow 0} \frac{1}{1 + \arctan x} \frac{1}{1 + x^2} = 1.$$

Daraus folgt unmittelbar

$$\lim_{x \rightarrow 0} (1 + \arctan x)^{1/x} = \lim_{x \rightarrow 0} \exp\left(\frac{1}{x} \ln(1 + \arctan x)\right) = e^1 = e.$$

$$18. \lim_{x \rightarrow 0} \left(\frac{1}{x} - \frac{1}{\ln(x+1)} \right) = \lim_{x \rightarrow 0} \frac{\ln(x+1) - x}{x \ln(x+1)} = \lim_{x \rightarrow 0} \frac{\frac{1}{x+1} - 1}{\ln(x+1) + \frac{x}{x+1}} =$$

$$\lim_{x \rightarrow 0} \frac{-x}{x + (x+1) \ln(x+1)} = \lim_{x \rightarrow 0} \frac{-1}{2 + \ln(x+1)} = -\frac{1}{2}.$$

Nach diesem Feuerwerk der verschiedensten Grenzwerte von Funktionen wenden wir uns kurz dem Begriff der *Konvexität* zu.

Definition. Eine Funktion $f : I \rightarrow \mathbb{K}$, I ein beliebiges Intervall in \mathbb{K} , heißt *konvex*, wenn für alle $a, b \in I$, $a < b$, gilt:

$$f(x) \leq f(a) + (x-a) \frac{f(b) - f(a)}{b-a}, \quad x \in [a, b].$$

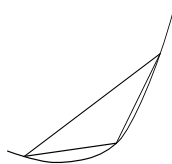
Jedes solche x schreibt sich auch in der Form $x = ta + (1-t)b$, $t \in [0, 1]$, so daß die Konvexitätsbedingung sich auch wie folgt umschreiben läßt:

$$f(ta + (1-t)b) \leq f(a) + (ta + (1-t)b - a) \frac{f(b) - f(a)}{b-a} = tf(a) + (1-t)f(b), \quad t \in [0, 1].$$

Hierin ist $t = (b-x)/(b-a)$, so daß die Konvexitätsbedingung auch noch gelesen werden kann in der Form

$$f(x) \leq \frac{b-x}{b-a} f(a) + \frac{x-a}{b-a} f(b), \quad a < x < b.$$

Mit diesen Vorüberlegungen ist es ein Leichtes, das folgende Lemma herzuleiten. Sein Inhalt wird in dem anschließenden Bild verdeutlicht.



Figur 14.16

Lemma 14.30 Für eine Funktion $f : I \rightarrow \mathbb{K}$ sind die folgenden Aussagen äquivalent :

- i) f ist konvex ;
 ii) für jedes Paar $a < b$ in I und alle x zwischen a und b gilt

$$\frac{f(x) - f(a)}{x - a} \leq \frac{f(b) - f(x)}{b - x} ;$$

- iii) für jedes Paar $a < b$ in I und alle x zwischen a und b gilt

$$\frac{f(x) - f(a)}{x - a} \leq \frac{f(b) - f(a)}{b - a} \leq \frac{f(b) - f(x)}{b - x} .$$

Beweis. Aus der Definition der Konvexität ergibt sich unmittelbar die Ungleichung

$$f(x) - f(x_1) \leq (x - x_1) \frac{f(x_2) - f(x_1)}{x_2 - x_1} ,$$

also die Hälfte der Behauptung in iii). Die andere Hälfte folgt ganz entsprechend. Offensichtlich impliziert iii) die Aussage ii). Ebenso leicht gelangt man von ii) nach i). \square

Ist nun f sogar differenzierbar auf dem Intervall I , so erhält man hieraus leicht ein *differentielles Kriterium* für Konvexität. Wir setzen im folgenden Satz wieder voraus, daß im Körper \mathbb{K} der Mittelwertsatz gilt.

Satz 14.31 Eine differenzierbare Funktion $f : I \rightarrow \mathbb{K}$ ist genau dann konvex, wenn die Ableitung f' monoton wachsend ist.

Beweis. a). Es seien $a < b$ zwei Punkte in I . Läßt man dann in der Formel iii) des vorigen Satzes x gegen a bzw. b gehen, so ergibt sich

$$f'(a) \leq \frac{f(b) - f(a)}{b - a} \leq f'(b) .$$

b). Es seien nun a, b, x drei Punkte in dem Intervall I mit $a < x < b$. Nach dem Mittelwertsatz gibt es dann Punkte $\xi_1 \in (a, x)$ und $\xi_2 \in (x, b)$ mit

$$\frac{f(x) - f(a)}{x - a} = f'(\xi_1) \quad \text{und} \quad \frac{f(b) - f(x)}{b - x} = f'(\xi_2) .$$

Da $\xi_1 < \xi_2$ und f' monoton wächst, hat man für jedes solche Tripel von Punkten

$$\frac{f(x) - f(a)}{x - a} \leq \frac{f(b) - f(x)}{b - x} ,$$

woraus nach ii) in Lemma 30 die Konvexität von f folgt. \square

Eine unmittelbare Konsequenz hieraus ist die

Folgerung 14.32 Ist $f : I \rightarrow \mathbb{K}$ auf einem nichttrivialen Intervall $I \subset \mathbb{K}$ zweimal differenzierbar und ist $f'' \geq 0$, so ist f konvex, sofern \mathbb{K} dem Mittelwertsatz genügt (also isomorph zu \mathbb{R} ist).

Beweis. Wegen des Mittelwertsatzes ist die Ableitung f' monoton wachsend. \square

Bemerkungen. 1. Selbstverständlich braucht eine konvexe Funktion nicht differenzierbar zu sein, wie das Beispiel $f(x) = |x|$ zeigt.

2. *Strenge* Monotonie von f' bzw. Positivität der zweiten Ableitung implizieren natürlich *strenge* Konvexität.

3. Eine Funktion f heißt (streng) *konkav*, wenn die Funktion $-f$ (streng) konvex ist.

Wir wollen die Bedeutung auch dieses Konzeptes durch den Beweis einiger wichtiger Ungleichungen hervorheben. Zunächst bemerken wir, daß eine der möglichen Definitionen der Konvexität einer Funktion wesentlich verallgemeinert werden kann.

Lemma 14.33 (Jensen'sche Ungleichung) Ist $f : I \rightarrow \mathbb{R}$ eine konvexe Funktion, so gilt für je n Elemente $x_j \in I$ und n nicht negative reelle Zahlen λ_j mit $\lambda_1 + \dots + \lambda_n = 1$ die folgende Ungleichung:

$$f(\lambda_1 x_1 + \dots + \lambda_n x_n) \leq \lambda_1 f(x_1) + \dots + \lambda_n f(x_n).$$

Ist f streng konvex, so gilt hier sogar die strikte Ungleichung, außer wenn $x_1 = \dots = x_n$.

Beweis. Wir führen den Beweis durch vollständige Induktion nach $n \geq 1$, wobei für $n = 1$ nichts zu zeigen ist. Sind uns dann $n + 1$ Punkte und Zahlen gegeben, die ohne Einschränkung alle positiv seien, so setzen wir $\lambda := \lambda_1 + \dots + \lambda_n$ und

$$x := \frac{\lambda_1}{\lambda} x_1 + \dots + \frac{\lambda_n}{\lambda} x_n.$$

Man sieht sofort, daß auch $x \in I$. (Das ist nichts anderes als die „Konvexität“ aller Intervalle). Dann ist aber auch

$$f\left(\sum_{j=1}^n \lambda_j x_j + \lambda_{n+1} x_{n+1}\right) = f(\lambda x + \lambda_{n+1} x_{n+1}) \leq \lambda f(x) + \lambda_{n+1} f(x_{n+1}) \leq \sum_{j=1}^{n+1} \lambda_j f(x_j).$$

Gilt jedoch Gleichheit für streng konvexes f , so ergibt sich aus der letzten Beziehung aber $x = x_{n+1}$ und per Induktion auch $x_1 = \dots = x_n = x = x_{n+1}$. \square

Da der natürliche Logarithmus wegen $(\ln x)'' = -1/x^2 < 0$ streng konkav ist, erhält man aus der Jensenschen Ungleichung die folgende Abschätzung:

$$\ln(\lambda_1 x_1 + \dots + \lambda_n x_n) \geq \lambda_1 \ln(x_1) + \dots + \lambda_n \ln(x_n).$$

Durch „Exponieren“ folgt hieraus eine Ungleichung, die eine weitreichende Verallgemeinerung der Ungleichung zwischen dem arithmetischen und dem geometrischen Mittel darstellt. Sie wird aus naheliegenden Gründen die *Ungleichung zwischen dem gewichteten arithmetischen und dem gewichteten geometrischen Mittel* genannt.

Satz 14.34 Sind x_1, \dots, x_n beliebige positive reelle Zahlen und $\lambda_1, \dots, \lambda_n$ beliebige nicht negative reelle Zahlen mit $\lambda_1 + \dots + \lambda_n = 1$, so gilt:

$$x_1^{\lambda_1} \cdot \dots \cdot x_n^{\lambda_n} \leq \lambda_1 x_1 + \dots + \lambda_n x_n.$$

Das Gleichheitszeichen steht nur dann, wenn $x_1 = \dots = x_n$.

Zum Abschluß dieses Abschnitts sollen noch Beweise der HÖLDERSchen und der MINKOWSKISchen Ungleichung angegeben werden, die häufig Anwendung in der höheren Analysis finden.

Satz 14.35 Für reelle Zahlen $p, q > 1$ mit $\frac{1}{p} + \frac{1}{q} = 1$ und beliebige Vektoren $z, w \in \mathbb{C}^n$ gilt :

$$\sum_{j=1}^n |z_j w_j| \leq \|z\|_p \cdot \|w\|_q,$$

wobei für einen Vektor $z = (z_1, \dots, z_n) \in \mathbb{C}^n$ sein p -Norm erklärt ist durch

$$\|z\|_p := \left(\sum_{j=1}^n |z_j|^p \right)^{1/p}.$$

Beweis. Wir können ohne Beschränkung der Allgemeinheit annehmen, daß die beiden beteiligten Vektoren von Null verschieden sind. Dann ist nach der Ungleichung zwischen dem gewichteten arithmetischen und geometrischen Mittel

$$\frac{|z_j w_j|}{\|z\|_p \|w\|_q} \leq \frac{1}{p} \frac{|z_j|^p}{\|z\|_p^p} + \frac{1}{q} \frac{|w_j|^q}{\|w\|_q^q},$$

woraus schon die Behauptung durch Aufsummieren folgt. \square

Bemerkung. Offensichtlich geht für $p = q = 2$ die Höldersche in die CAUCHY - SCHWARZsche Ungleichung über.

Die Minkowskische Ungleichung ist nichts anderes als die Dreiecksungleichung für die p -Norm. Sie ist eine Konsequenz aus der Hölderschen Ungleichung.

Satz 14.36 Für $p \geq 1$ und alle $z, w \in \mathbb{C}^n$ gilt :

$$\|z + w\|_p \leq \|z\|_p + \|w\|_p.$$

Beweis. Für $p = 1$ ist nichts zu zeigen. Es sei also $p > 1$ und $q > 1$ so gewählt, daß $1/p + 1/q = 1$. Mit $s_j := |z_j + w_j|^{p-1}$ ist $|z_j + w_j|^p = s_j |z_j + w_j| \leq |z_j| s_j + |w_j| s_j$, so daß die Höldersche Ungleichung zu

$$(*) \quad \|z + w\|_p^p \leq \|z\|_p \|s\|_q + \|w\|_p \|s\|_q$$

führt. Wegen $q = p/(p-1)$ ist aber $s_j^q := |z_j + w_j|^p$, und damit ist

$$\|s\|_q = \left(\sum_{j=1}^n |s_j|^q \right)^{1/q} = \left(\sum_{j=1}^n |z_j + w_j|^p \right)^{1/q} = \|z + w\|_p^{p-1},$$

woraus mit (*) die Behauptung folgt. \square

Wir haben oben schon erwähnt, daß die Dreiteilung des allgemeinen Winkels mit Zirkel und Lineal nicht möglich ist. Dies schreckt jedoch Laien nicht davon ab, immer wieder neue Konstruktionen zu ersinnen, die das Problem zwar nicht lösen, aber im besten Falle doch sehr gut approximieren. Eine solche Lösung hat kein Geringerer als OSKAR PERRON in seinem Artikel *Die Winkeldreiteilung des Schneidermeisters Kopf*, Sitzungsberichte der Bayerischen Akademie der Wissenschaften, Jahrgang 1929, 341–343, vorgestellt. (Siehe auch vom selben Autor *Eine neue Winkeldreiteilung des Schneidermeisters Kopf*, *ibid.*, Jahrgang 1933, 439–445). Wir zitieren die ersten Absätze:

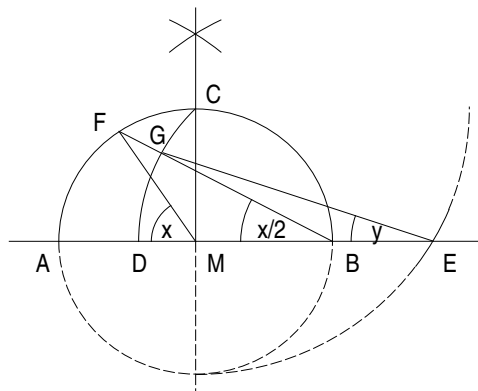
„Herr Eugen Kopf, Schneidermeister in Ludwigshafen a. Rh., behauptet, daß eine von ihm ersonnene Konstruktion mit Lineal und Zirkel die genaue Dreiteilung des Winkels leiste. Die Behauptung ist natürlich falsch, und was Herr Kopf zu ihrer Begründung anführt, steht auf keinem höheren Niveau

als man es sonst bei Winkelteilern und Kreisquadrirern gewöhnt ist. Um so überraschender ist es, daß mit der Konstruktion selbst ihr Entdecker doch eine wirklich glückliche Hand bewiesen hat; denn durch rechnerische Nachprüfung fand ich, daß sie im Verhältnis zu ihrer Einfachheit eine ganz erstaunlich gute Näherung liefert. Der Fehler ist für spitze Winkel im ungünstigsten Fall nur $8' 12''$, liegt also unter der Genauigkeitsgrenze der besten handlichen Zeichnung.

Die Konstruktion findet sich nicht in den einschlägigen Büchern von Enriques und Vahlen; die dort mitgeteilten Konstruktionen sind zum Teil viel ungenauer und nicht einfacher, zum Teil nur wenig genauer und dann bedeutend komplizierter, so daß die Kopfsche Konstruktion entschieden hübscher ist. Man darf aus diesem Umstand wohl mit ziemlicher Wahrscheinlichkeit schließen, daß die Konstruktion neu ist. Aber selbst, wenn sie irgendwo in einem vergessenen alten Schmöcker stehen sollte, was man bei derlei Dingen nie wissen kann, ist sie immerhin wert, der Vergessenheit entrissen zu werden.

Die Konstruktion verläuft, wenn man sie von den unnötigen Linien befreit, folgendermaßen (in der Figur sind gar keine Hilfslinien unterdrückt, nicht einmal solche, die lediglich zur Konstruktion eines rechten Winkels dienen):

Man zeichne einen Halbkreis und den zu seinem Durchmesser AB senkrechten Radius MC . Durch C zeichne man den Kreisbogen CD mit dem Mittelpunkt B . Schließlich markiere man auf der Verlängerung von AB den Punkt E so, daß $CE = AB$ ist. Ist nun $\sphericalangle AMF = x$ der gegebene spitze Winkel, so bringe man die Gerade FB mit dem Kreisbogen CD zum Schnitt in G . Verbindet man dann G mit E , so ist $\sphericalangle AEG = y$ die Kopfsche Näherung für den dritten Teil von x .



Figur 14.17

Wir wollen die Behauptungen Perrons nachvollziehen, wobei wir uns seinem Gedankengang fast wörtlich anschließen.

Mit Hilfe der Zeichnung und mit dem Sinussatz erhält man sofort

$$\frac{\sin(x/2 - y)}{\sin y} = \frac{\sqrt{3} - 1}{\sqrt{2}},$$

woraus sich mit dem Additionstheorem für den Sinus sofort

$$\cot y = \frac{\sqrt{3} - 1 + \sqrt{2} \cos(x/2)}{\sqrt{2} \sin(x/2)}$$

und damit

$$y = \arctan \frac{\sqrt{2} \sin(x/2)}{\sqrt{3} - 1 + \sqrt{2} \cos(x/2)}$$

ergibt. Der absolute Fehler ist dann

$$f(x) = \frac{x}{3} - y.$$

Man stellt sofort fest, daß $f(0) = 0$ und (durch Inspektion der Zeichnung) auch $f(\pi/2) = 0$ (oder durch Nachrechnen: Für $x = \pi/2$ ist

$$\tan y = \frac{\sqrt{2} \sin(\pi/4)}{\sqrt{3} - 1 + \sqrt{2} \cos(\pi/4)} = \frac{1}{\sqrt{3}}, \quad \text{da} \quad \sin(\pi/4) = \cos(\pi/4) = \frac{1}{2}\sqrt{2},$$

und damit $y = \pi/6$, da $\sin(\pi/6) = 1/2$, $\cos(\pi/6) = \sqrt{3}/2$).

Durch Differenzieren dieser Beziehung findet man unmittelbar (indem man noch mit $\sqrt{3} + 1$ erweitert):

$$f'(x) = -\frac{1}{3} - \frac{\sqrt{3} + 1 + \sqrt{2} \cos(x/2)}{4(\sqrt{3} + \sqrt{2} \cos(x/2))}.$$

Durch nochmaliges Differenzieren kommt

$$f''(x) = \frac{-\sin(x/2)}{4\sqrt{2}(\sqrt{3} + \sqrt{2} \cos(x/2))^2} < 0, \quad 0 < x \leq \pi/2;$$

so daß die Fehlerfunktion streng konkav und insbesondere im Inneren des Definitionsintervalls stets positiv, d. h. der Kopfsche Winkel stets zu klein ist. Des weiteren muß die Ableitung der Fehlerfunktion zuerst streng monoton steigen und dann nach Annahme ihres Maximums wieder streng monoton fallen. Die Extremalstelle ξ berechnet sich aus der Ableitungsformel sofort implizit zu

$$\cos \frac{\xi}{2} = \frac{3 - \sqrt{3}}{\sqrt{2}}, \quad \text{also} \quad \cos \xi = 11 - 6\sqrt{3} \quad \text{und} \quad \sin \frac{\xi}{2} = \sqrt{3\sqrt{3} - 5}.$$

Der maximale Fehler kann damit völlig explizit angegeben werden in der Form

$$M := f_{\max} = \frac{1}{3} \arccos(11 - 6\sqrt{3}) - \arctan \sqrt{\frac{3\sqrt{3} - 5}{2}}.$$

Mit einer sehr guten trigonometrischen Tafel findet man die entsprechenden Winkel zu

$$\xi = 52^\circ 34' 37'' \quad \text{und} \quad \eta = 17^\circ 23' 20,6'',$$

so daß der maximale Fehler in etwa $0^\circ 8' 12''$ beträgt. Nimmt man zusätzlich an, daß der vorgegebene Winkel größer als 45° ist, was man ja immer durch Verdoppelung und anschließende Halbierung erreichen kann, so ist übrigens der relative Fehler in grober Abschätzung maximal in etwa gleich

$$\frac{8,2}{60 \cdot 15} = 0,91111 \dots \%$$

Die letzte Argumentation zur Abschätzung des maximalen Fehlers ist nicht völlig befriedigend, auch wenn sie ohne große Mühe bei etwas genauerer Schlußweise zu einer tatsächlichen oberen Schranke für den Fehler führt. Bei weniger Mühe gewinnt man eine gröbere, nämlich ca. doppelt so große) Abschätzung leicht wie folgt: Da der Arcuscosinus eine konkave Funktion ist, liegt er im Intervall von $1/2$ bis $\sqrt{3}/2$ oberhalb der Sekante

$$y = \frac{\sqrt{3} + 9}{24} \pi - \frac{\sqrt{3} + 1}{12} \pi x.$$

Nun ist offensichtlich $1/2 < 11 - 6\sqrt{3} < \sqrt{3}/2$. Hiermit berechnet man

$$\arccos \xi \geq \pi \frac{23 - 9\sqrt{3}}{24}$$

und damit

$$M \leq \frac{\pi}{2} \frac{136 - 78\sqrt{3}}{12 \cdot 24}.$$

Der letzte Ausdruck ist mit Sicherheit kleiner oder gleich $0^\circ 16' 53''$.

Zum Abschluß dieses Kapitels wollen wir uns mit der Interpretation und Ausgestaltung der *Taylor-Formel* beschäftigen. Ist die Funktion f an einer Stelle a differenzierbar bis zur Ordnung n , so kann man ihr das sogenannte *Taylor-Polynom* n -ter Ordnung an der Stelle a zuordnen:

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

Man rechnet sofort nach, daß dieses Polynom die Eigenschaft besitzt, daß

$$\left. \frac{d^k}{dx^k} T_{f,a}^n(x) \right|_{x=a} = \left. \frac{d^k}{dx^k} f(x) \right|_{x=a}, \quad k = 0, \dots, n.$$

Selbstverständlich ist das Taylor-Polynom durch diese Bedingungen auch eindeutig bestimmt.

Differenzierbarkeit bedeutet per definitionem gerade, daß der Graph der Funktion f von dem Graphen der Tangente *berührt*, also besser als linear approximiert wird. Wir können nun erwarten, daß das Taylor-Polynom n -ter Ordnung den Graphen der Funktion nahe der Stelle a besser als von n -ter Ordnung berührt. Dies ist in der Tat der Fall, zumindest, wenn die Funktion f nicht nur eine Funktion mit $(*)_0$ ist, sondern sogar n -mal *stetig* differenzierbar ist.

Satz 14.37 *Ist $f : I \rightarrow \mathbb{R}$ differenzierbar bis zur Ordnung n auf dem nicht trivialen Intervall $I \subset \mathbb{R}$, und ist die n -te Ableitung $f^{(n)}$ im Punkte $a \in I$ noch stetig, so gilt*

$$f(x) - T_{f,a}^n(x) = \varphi(x) (x-a)^n$$

mit einer Funktion $\varphi : I \rightarrow \mathbb{R}$, die im Punkte a stetig ist und dort verschwindet. Das Taylor-Polynom ist das einzige Polynom höchstens n -ten Grades, das diese Eigenschaft besitzt.

Beweis. Nach der Taylor-Formel mit Restglied nach Lagrange ist

$$f(x) - T_{f,a}^{n-1}(x) = \frac{f^n(\xi)}{n!} (x-a)^n,$$

wobei $\xi = \xi(x)$ zwischen a und x liegt. Man hat daher eine wie im Satz geforderte Darstellung, wenn man

$$\varphi(x) := \frac{f^n(\xi) - f^n(a)}{n!}$$

setzt. Da mit x auch ξ gegen a geht, ist die Funktion φ tatsächlich stetig an der Stelle a , wo sie nach Definition auch verschwindet. Die Eindeutigkeitsaussage erhält man durch Bildung der Differenz aus zwei solchen Darstellungen als Korollar zu dem folgenden Lemma. \square

Lemma 14.38 *Gilt für ein Polynom P höchstens n -ten Grades, daß $P(x) = \varphi(x)x^n$ in einer Umgebung von 0 mit $\lim_{x \rightarrow 0} \varphi(x) = 0$, so ist P das Nullpolynom.*

Beweis. Es sei $P(x) = a_0 + a_1x + \dots + a_nx^n$ und für ein $k \in \mathbb{N}$ mit $k \leq n$ schon nachgewiesen, daß $a_0 = \dots = a_{k-1} = 0$. Dann gilt für $x \neq 0$:

$$\frac{P(x)}{x^k} = a_k + a_{k+1}x + \dots + a_nx^{n-k} = \varphi(x)x^{n-k}.$$

Da die rechte Seite mit $x \rightarrow 0$ immer noch gegen Null geht, muß auch $a_k = 0$ sein. \square

Bemerkung. Das Restglied in der Taylor-Formel kann in ganz unterschiedlicher Weise ausgedrückt werden. Die *Integral-Form*

$$R_{n+1}(x) = \frac{1}{n!} \int_a^x (x-t)^n f^{(n+1)}(t) dt$$

werden wir später noch herleiten. Aus dieser folgt übrigens die Lagrange-Form unter Berücksichtigung des Mittelwertsatzes der Integralrechnung.

Wir wollen hier noch eine weitere Form des Restgliedes angeben, das auf SCHLÖMILCH zurückgeht. Es enthält als Spezialfälle die Lagrange-Form ($p = n + 1$) als auch die sogenannte *Cauchy-Form* ($p = 1$).

Satz 14.39 Erfüllt die Funktion f die Bedingung des Taylorschen Satzes, so gibt es zu jedem $p \in \mathbb{N}^*$ und jedem $x \in I$ ein $\vartheta \in (0, 1)$, so daß

$$f(x) - T_{f,a}^n(x) = R_{n+1}(x) = \frac{f^{(n+1)}(a + \vartheta(x-a))}{n!p} (1-\vartheta)^{n+1-p} (x-a)^{n+1}$$

ist.

Beweis. Die angegebene Formel ist offensichtlich im Punkte $x = a$ richtig. Wir können uns daher auf den Fall $a < x$ beschränken und betrachten wie im Beweis von Satz 16 bei fixiertem x die Funktionen $F(t)$ mit $t \in [a, x]$. Die Funktion $g(t) := (x-t)^p$ besitzt in dem fraglichen Intervall die Ableitung $g'(t) = -p(x-t)^{p-1}$ und ist damit auf jeden Fall von Null verschieden für $a < t < x$. Somit können wir den zweiten Mittelwertsatz auf die Funktionen F und g anwenden. Leichte Rechnungen ergeben $F(x) = g(x) = 0$ und

$$F(a) = f(x) - \sum_{j=0}^n \frac{f^{(j)}(a)}{j!} (x-a)^j, \quad g(a) = (x-a)^p.$$

Ferner berechnet man wie früher $F'(t) = -f^{(n+1)}(t)(x-t)^n/n!$. Somit gibt es nach dem zweiten Mittelwertsatz ein ξ zwischen a und x , so daß

$$\frac{f(x) - \sum_{j=0}^n \frac{f^{(j)}(a)}{j!} (x-a)^j}{(x-a)^p} = \frac{f^{(n+1)}(\xi)}{n!} \frac{(x-\xi)^n}{p(x-\xi)^{p-1}}.$$

Schreibt man ξ in der Form $a + \vartheta(x-a)$ mit einem geeigneten $\vartheta \in (0, 1)$, so ist $x - \xi = (1 - \vartheta)(x - a)$, und die Behauptung ergibt sich sofort aus der vorstehenden Identität. \square

Als letzte Anwendung in diesem Kapitel charakterisieren wir schließlich die sogenannten *analytischen* Funktionen. Hierbei ist $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$.

Definition. Es sei $U \subset \mathbb{K}$ eine offene Menge, $f : U \rightarrow \mathbb{K}$ eine Funktion und $x_0 \in U$. f heißt in x_0 *analytisch*, wenn es eine Umgebung $B(x_0, r) \subset U$ von x_0 und eine in $B(x_0, r)$ konvergente Potenzreihe

$$\sum_{n=0}^{\infty} a_n (x - x_0)^n$$

gibt, die dort gegen f konvergiert. Sie heißt *analytisch* schlechthin, wenn sie in jedem Punkt $x_0 \in U$ analytisch ist.

Bemerkung. Die Potenzreihenentwicklung einer analytischen Funktion f um einen Punkt x_0 ist *eindeutig bestimmt*. Wir haben weiter oben ja schon gesehen, daß - wenn wir ohne Einschränkung $x_0 = 0$ annehmen - aus $f(x) = \sum_{n=0}^{\infty} a_n x^n$ die Entwicklung

$$f'(x) = \sum_{n=0}^{\infty} (n+1) a_{n+1} x^n$$

folgt, aus der man durch vollständige Induktion für jedes natürliche k die Formel

$$(\times)_k \quad f^{(k)}(x) = \sum_{n=0}^{\infty} \frac{(n+k)!}{n!} a_{n+k} x^n$$

gewinnt. Setzt man hier speziell $x = 0$ ein, so hat man $f^{(k)}(0) = k! a_k$, also

$$a_k = \frac{1}{k!} f^{(k)}(0).$$

Die Potenzreihenentwicklung an der Stelle $x_0 = 0$ und damit auch an jeder beliebigen Stelle x_0 ist daher notwendigerweise die sogenannte *Taylor-Entwicklung*

$$f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(x_0)}{n!} (x - x_0)^n.$$

Wegen der nicht zu unterschätzenden grundlegenden Bedeutung sei diese Aussage als Satz formuliert.

Satz 14.40 *Ist f in a analytisch, so ist f in einer Umgebung von a beliebig oft differenzierbar, und die Potenzreihenentwicklung von f in a stimmt mit der Taylorentwicklung überein :*

$$f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(a)}{n!} (x - a)^n.$$

Aus dem Satz von der Taylor-Formel leitet man unmittelbar das folgende hinreichende Kriterium dafür ab, daß eine beliebig oft differenzierbare reelle Funktion analytisch ist. Am Ende des Paragraphen werden wir daraus eine weitere hinreichende Bedingung ableiten, deren Notwendigkeit man aber erst durch den Übergang ins Komplexe einsehen kann.

Satz 14.41 *Eine C^∞ -Funktion $f : I := (-\delta, \delta) \rightarrow \mathbb{R}$ ist genau dann in I durch ihre Taylor-Reihe darstellbar, wenn die Folge der Reste $R_{n+1}(x)$ in der Taylor-Formel für alle $x \in I$ gegen Null konvergiert.*

Beispiel. Von EULER wurden 1755 für $y = f(x) = \arctan x$ die folgenden Formeln bewiesen:

$$f^{(n)}(x) = (n-1)! \sin\left(n\left(y + \frac{\pi}{2}\right)\right) \cos^n y, \quad n \geq 1.$$

In der Tat: Für $n = 1$ ist einerseits

$$y' = \frac{1}{1+x^2} = \frac{1}{1+\tan^2 y} = \cos^2 y$$

und andererseits

$$\sin(y + \pi/2) = \cos y,$$

womit der Induktionsanfang erledigt ist. Ist die Formel

$$y^{(n)} = (n-1)! \sin\left(n\left(y + \frac{\pi}{2}\right)\right) \cos^n y$$

nun für ein $n \geq 1$ schon gezeigt, so erhält man durch eine weitere Differentiation

$$\begin{aligned} y^{(n+1)} &= n! y' \cos^{n-1} y [\cos(n(y + \pi/2)) \cos y - \sin y \sin(n(y + \pi/2))] \\ &= n! \cos^2 y \cos^{n-1} y \cos((n+1)y + n(\pi/2)) \\ &= n! \cos^{n+1} y \sin((n+1)(y + \pi/2)). \end{aligned}$$

Aus den Eulerschen Formeln folgt nun speziell mit $x_0 = 0$ auch $y_0 = 0$ und damit

$$f^{(n)}(0) = (n-1)! \sin(n\pi/2) = \begin{cases} 0 & \text{falls } n \text{ gerade,} \\ (-1)^{(n-1)/2} (n-1)! & \text{falls } n \text{ ungerade.} \end{cases}$$

Damit erhalten wir für die Taylor-Reihe des \arctan die (uns schon bekannte) Gestalt

$$\sum_{n=0}^{\infty} \frac{(-1)^n}{2n+1} x^{2n+1}.$$

Um ohne Integration und den Abelschen Grenzwertsatz zu beweisen, daß diese Reihe für $|x| \leq 1$ gegen den Arcustangens konvergiert, müssen wir noch das Restglied R_{n+1} für diese Werte von x abschätzen. Nach den Eulerschen Formeln ist aber offensichtlich $|f^{(n+1)}|_{\infty} \leq n!$. Das Restglied in der Form von Lagrange kann dann für alle $x \in \mathbb{R}$ abgeschätzt werden durch

$$|R_{n+1}(x)| \leq \frac{n!}{(n+1)!} |x|^{n+1} = \frac{|x|^{n+1}}{n+1}.$$

Hieraus folgt die Behauptung.

Wir wollen als nächstes einsehen, daß die Eigenschaft einer Funktion, analytisch zu sein, in Wahrheit eine *offene* Eigenschaft ist, d. h. daß sie mit x_0 auch für alle x_1 in einer Umgebung von x_0 gilt. Dies ist aber offensichtlich gerade der Inhalt der Folgerung 12.19, die wir hier noch einmal gesondert aufführen.

Satz 14.42 *Es sei $f(x) = \sum_{n=0}^{\infty} a_n x^n$ konvergent in dem Kreis $B(0, R)$. Ist dann $a \in B(0, R)$ und ordnet man formal um :*

$$\sum_{n=0}^{\infty} a_n x^n = \sum_{n=0}^{\infty} a_n ((x-a) + a)^n = \sum_{n=0}^{\infty} a_n \sum_{k=0}^n \binom{n}{k} a^{n-k} (x-a)^k = \sum_{k=0}^{\infty} b_k (x-a)^k$$

mit

$$b_k = \sum_{n=k}^{\infty} \binom{n}{k} a_n a^{n-k},$$

so existieren alle b_k , und die Reihe $\sum_{k=0}^{\infty} b_k (x-a)^k$ konvergiert in dem Kreis $B(a, R - |a|)$ gegen die ursprüngliche Reihe (wobei $r := R - |a|$ gleich ∞ zu setzen ist, wenn auch R unendlich ist).

Bemerkungen. 1. Aus diesem Umordnungssatz ergibt sich noch einmal auf andere Weise das früher schon hergeleitete Resultat über die *Differenzierbarkeit* von Potenzreihen in ihrem Konvergenzkreis. Subtrahiert man von der Entwicklung um den Punkt a den Wert $f(a) = b_0$, so erhält man für $0 < |x-a| < r$ die konvergente Entwicklung

$$f(x) - f(a) = (x-a) \sum_{k=1}^{\infty} b_k (x-a)^{k-1}.$$

Wegen der Stetigkeit von Potenzreihen in ihrem Konvergenzkreis impliziert dies die Existenz von

$$f'(a) = \lim_{x \rightarrow a} \frac{f(x) - f(a)}{x-a} = b_1 = \sum_{n=1}^{\infty} n a_n a^{n-1}. \quad \square$$

2. Der vorstehende Umordnungssatz drückt auf der anderen Seite nichts anderes aus als die Taylorentwicklung der durch die Potenzreihe gegebenen Funktion f an der Stelle a . Denn nach den Formeln $(\times)_k$ ist

$$\begin{aligned} \sum_{k=0}^{\infty} \frac{f^{(k)}(a)}{k!} (x-a)^k &= \sum_{k=0}^{\infty} \left(\frac{1}{k!} \sum_{n=0}^{\infty} \frac{(n+k)!}{n!} a_{n+k} a^n \right) (x-a)^k \\ &= \sum_{k=0}^{\infty} \left(\sum_{n=0}^{\infty} \frac{(n+k)!}{n! k!} a_{n+k} a^n \right) (x-a)^k \\ &= \sum_{k=0}^{\infty} \left(\sum_{n=0}^{\infty} \binom{n+k}{k} a_{n+k} a^n \right) (x-a)^k \\ &= \sum_{k=0}^{\infty} \left(\sum_{n=k}^{\infty} \binom{n}{k} a_n a^{n-k} \right) (x-a)^k. \end{aligned}$$

Satz 41 wirft nun die Frage auf, unter welchen einfacheren Bedingungen eine beliebig oft differenzierbare Funktion $f : U \rightarrow \mathbb{K}$ in einer evtl. kleineren Umgebung von $a \in U$ durch ihre *formale Taylorreihe*

$$T_{f,a}(x) := \sum_{n=0}^{\infty} \frac{f^{(n)}(a)}{n!} (x - a)^n$$

an der Stelle $a \in U$ dargestellt wird, wann also gilt, daß

$$T_{f,a}(x) = f(x) ?$$

Bemerkungen. 1. Die Zuordnung $f \mapsto T_{f,a}$ kann man auffassen als einen \mathbb{K} -Algebra-Homomorphismus von der Algebra der sogenannten *Keime* von C^∞ -Funktionen in a in die Algebra der *formalen* Potenzreihen in a :

$$\mathcal{C}_{\mathbb{K},a} \xrightarrow{T} \mathbb{K}\{\{x - a\}\}.$$

Nach einem Satz von E. Borel²³ ist die Abbildung surjektiv, d. h. zu vorgegebenen $a_n \in \mathbb{K}$ gibt es eine C^∞ -Funktion f (in ganz \mathbb{K}) mit

$$f^{(n)}(0) = a_n.$$

Wählt man hier speziell $a_n = (n!)^2$, so sieht man, daß die Taylor-Reihe (außer in 0) nicht zu konvergieren braucht.

2. Die Abbildung $f \mapsto T_{f,a}$ ist nicht injektiv, d. h. $\ker T$ ist nicht das Null-Ideal. Man hat tatsächlich „viele“ Funktionen $f \in \ker T$, also Funktionen mit $f^{(n)}(0) = 0$ für alle $n \in \mathbb{N}$, insbesondere nichttriviale Funktionen (man nennt diese dann „flach“ in 0). Ein typisches Beispiel ist

$$f(x) = \begin{cases} e^{-1/x^2}, & x \neq 0 \\ 0, & x = 0. \end{cases}$$

3. Betrachtet man nur die *analytischen* Funktionskeime $\mathcal{O}_{\mathbb{K},a} \subset \mathcal{C}_{\mathbb{K},a}^\infty$, so ist $T|_{\mathcal{O}_{\mathbb{K},a}}$ ein \mathbb{K} -Algebra-Isomorphismus von $\mathcal{O}_{\mathbb{K},a}$ auf die Algebra der *konvergenten* Potenzreihen $\mathbb{K}\{x - a\}$.

Satz 14.43 *Es sei $f : (-\delta_0, \delta_0) \rightarrow \mathbb{R}$ eine C^∞ -Funktion. Genau dann ist f um 0 in eine konvergente Potenzreihe entwickelbar, wenn es ein δ mit $0 < \delta \leq \delta_0$ und eine Konstante $M > 0$ gibt, so daß*

$$|f^{(n)}(x)| \leq Mn! \delta^{-n}, \quad |x| < \delta.$$

Bemerkung. Die Notwendigkeit dieser Bedingung kann man erst über den Weg ins *Komplexe* verstehen (siehe den Teil dieses Manuskripts über (elementare) Funktionentheorie).

Beweis (daß die angegebene Bedingung hinreichend ist). Es sei $0 < \delta_1 < \delta$ und $|x| \leq \delta_1$. Dann gilt für das n -te Restglied $R_n(x)$ der Taylor-Entwicklung von f um 0:

$$|R_n(x)| = \frac{|f^{(n+1)}(\xi)|}{(n+1)!} |x|^{n+1} \leq M \left(\frac{\delta}{\delta_1} \right)^{n+1},$$

und die rechte Seite konvergiert gegen Null für $n \rightarrow \infty$. □

Zum Abschluß dieses Kapitels wollen wir uns noch kurz mit dem Konvergenzverhalten von Folgen *differenzierbarer* Funktionen beschäftigen. Offensichtlich reicht es nicht aus, wenn die Folge der differenzierbaren Funktionen f_n *gleichmäßig* konvergent ist, um auf die Differenzierbarkeit der Grenzfunktion $f = \lim f_n$ schließen zu können.

²³Für einen Beweis siehe z. B. die Note auf p. 99 von M. GOLUBITSKY, V. GUILLEMIN, *Stable Mappings and Their Singularities*, Springer: New York, Heidelberg, Berlin 1973.

Beispiel. Man mache aus der Betragsfunktion $f(x) = |x|$ durch Abänderung auf dem Intervall $|x| \leq 1/n$, $n \geq 1$ durch $f_n(x) := nx^2/2 + 1/2n$ eine auf ganz \mathbb{R} stetig differenzierbare Funktion f_n . Die Folge f_n konvergiert gleichmäßig auf \mathbb{R} gegen die Funktion f , die aber an der Stelle 0 nicht differenzierbar ist.

Man kann aber auf die Differenzierbarkeit der Grenzfunktion (im Wesentlichen) schon dann schließen, wenn die *Folge der Ableitungen gleichmäßig konvergiert*. Einen sehr einfachen Beweis werden wir im übernächsten Kapitel unter der stärkeren Voraussetzung der *stetigen Differenzierbarkeit* der Funktionen mit Hilfe des Hauptsatzes der Differential- und Integralrechnung gewinnen.

Satz 14.44 *Es sei $f_n : [a, b] \rightarrow \mathbb{R}$ eine Folge differenzierbarer Funktionen, die an einer einzigen Stelle $x_0 \in [a, b]$ konvergiert, und deren Ableitungsfolge (f'_n) auf $[a, b]$ gleichmäßig konvergiert. Dann ist auch die Folge (f_n) gleichmäßig konvergent, und ihre Grenzfunktion $f = \lim f_n$ ist differenzierbar mit*

$$f'(x) = \lim_{n \rightarrow \infty} f'_n(x), \quad x \in [a, b].$$

Bemerkungen. 1. Auf die Konvergenz an einer einzigen Stelle $x_0 \in [a, b]$ kann man selbstverständlich nicht verzichten, wie das Beispiel $f_n \equiv n$ zeigt.

2. Auch wenn die Funktionenfolge gleichmäßig gegen eine *differenzierbare* Grenzfunktion konvergiert, braucht die Folge der Ableitungen nicht konvergent zu sein. Ein Beispiel hierfür ist die Folge

$$f_n(x) = \frac{\sin n^2 x}{n}.$$

3. Unter den Voraussetzungen des vorigen Beispiels kann die Folge der Ableitungen sogar konvergent sein, aber nicht gegen die Ableitung der Grenzfunktion. So konvergiert die Folge

$$f_n(x) := x - \frac{x^n}{n}$$

auf dem Einheitsintervall $[0, 1]$ gleichmäßig gegen die Grenzfunktion $f(x) = x$. Die Folge der Ableitungen $f'(x) = 1 - x^{n-1}$ konvergiert punktweise in diesem Intervall, aber nicht im Punkte $x_0 = 1$ gegen $f'(1) = 1$.

4. Die Funktionenfolge $f_n(x) = nx(1-x)^n$ ist an den Stellen $x = 0, 1$ trivialerweise und an den Stellen $0 < x < 1$ wegen $\lim_{n \rightarrow \infty} na^n = 0$ für $|a| < 1$ gegen 0 konvergent. Die Grenzfunktion $f \equiv 0$ ist also stetig (und sogar beliebig oft differenzierbar). Dennoch konvergiert die Folge (f_n) *nicht gleichmäßig* gegen f . Dies sieht man z. B. wie folgt: Man betrachte zu jedem $n \in \mathbb{N}$ die Stelle $x_n := 1/(n+1) \in I$. Man berechnet dann sofort

$$f_n(x_n) = \frac{n}{n+1} \left(1 - \frac{1}{n+1}\right)^n = \left(1 - \frac{1}{n+1}\right)^{n+1},$$

und die rechte Seite konvergiert gegen e^{-1} . Setzt man dann $\varepsilon_0 := (2e)^{-1}$, so ist $f_n(x_n) \geq \varepsilon_0$ für fast alle n . Damit kann die Folge (f_n) auf I nicht gleichmäßig gegen Null konvergieren. Die Punkte x_n sind, wovon man sich sofort überzeugt, die Stellen, an denen die Funktion f_n jeweils ihr Maximum annimmt, welches nicht nach Null konvergiert. Dennoch konvergiert die Folge (f_n) gegen Null, da die Folge (x_n) gegen den rechten Endpunkt des Intervalls konvergiert. Das hier auftretende Phänomen wird aus naheliegenden Gründen als *wandernder Buckel* bezeichnet.

Bevor wir mit dem *Beweis* von Satz 44 beginnen können, müssen wir unser früheres Ergebnis zur Stetigkeit von Grenzfunktionen (Satz 12.14) verallgemeinern.

Satz 14.45 *Es sei $f_j : D \rightarrow \mathbb{R}$ eine gleichmäßig konvergente Folge, $D \subset X$, und $a \in X$ sei ein Häufungspunkt von D . Existiert dann*

$$\lim_{x \rightarrow a} f_j(x) \quad \text{für alle } j,$$

so existieren auch die Grenzwerte

$$\lim_{j \rightarrow \infty} \lim_{x \rightarrow a} f_j(x) = \lim_{x \rightarrow a} \lim_{j \rightarrow \infty} f_j(x).$$

Beweis. Wir setzen $f(x) := \lim_{j \rightarrow \infty} f_j(x)$ und $\alpha_j := \lim_{x \rightarrow a} f_j(x)$. Nach dem Cauchy-Kriterium gibt es zu vorgegebenem $\varepsilon > 0$ ein $N = N(\varepsilon)$, s. d. für alle $j, k \geq N$ und alle $x \in X$ gilt:

$$|f_j(x) - f_k(x)| < \varepsilon.$$

Bei festem j, k kann man in dieser Ungleichung x gegen a gehen lassen, so daß

$$|\alpha_j - \alpha_k| \leq \varepsilon \quad \text{für alle } j, k \geq N(\varepsilon).$$

Damit ist $(\alpha_j)_{j \in \mathbb{N}}$ eine Cauchy-Folge in \mathbb{R} , womit der Nachweis der Existenz des Grenzwertes

$$\lim_{j \rightarrow \infty} \lim_{x \rightarrow a} f_j(x) = \lim_{j \rightarrow \infty} \alpha_j =: \alpha$$

erbracht ist. Es bleibt noch zu zeigen, daß

$$\lim_{x \rightarrow a} f(x) = \alpha.$$

Dazu betrachten wir die Abschätzung

$$|f(x) - \alpha| \leq |f(x) - f_j(x)| + |f_j(x) - \alpha_j| + |\alpha_j - \alpha|.$$

Bei gegebenem $\varepsilon > 0$ können wir N so groß wählen, daß

$$|f(x) - f_N(x)| \leq \frac{\varepsilon}{3} \quad \text{für alle } x \in D$$

und $|\alpha_N - \alpha| < \frac{\varepsilon}{3}$. Nach Definition gibt es ferner ein $\delta = \delta(\varepsilon) > 0$, so daß

$$|f_N(x) - \alpha_N| < \frac{\varepsilon}{3} \quad \text{für alle } x \in B(a, \delta) \cap D.$$

Somit ist

$$|f(x) - \alpha| < \varepsilon \quad \text{für alle } x \in B(a, \delta) \cap D$$

und folglich

$$\lim_{x \rightarrow a} f(x) = \alpha. \quad \square$$

Wir wenden diesen Satz an, um die scharfe Version des Satzes über die Differenzierbarkeit von Grenzfunktionen herzuleiten.

Beweis (von Satz 44). Wir ersetzen die Funktionen f_j durch $f_j - f_j(x_0)$, so daß also ohne Einschränkung $f_j(x_0) = 0$ angenommen werden kann. Für die gleichmäßige Konvergenz der Funktionenfolge (f_j) bemühen wir das Cauchy-Kriterium und den 1. Mittelwertsatz: Zu jedem $x \in I$ gibt es ein ξ zwischen x_0 und x mit

$$|f_j(x) - f_k(x)| = |f'_j(\xi) - f'_k(\xi)| \cdot |x - a| \leq \sup |f'_j(x) - f'_k(x)| \cdot |x - a| \leq \varepsilon(b - a)$$

für alle $j, k \geq N = N(\varepsilon)$. Somit existiert $f = \text{Lim } f_j$. Selbstverständlich ist f stetig auf I . Es bleibt zu zeigen, daß f an jeder Stelle $\xi \in I$ differenzierbar ist. Wir betrachten dazu auf $D := I \setminus \{\xi\}$ die Funktion

$$g_j(x) = \frac{f_j(x) - f_j(\xi)}{x - \xi}.$$

Es folgt erneut aus dem 1. Mittelwertsatz und dem Cauchy-Kriterium, daß die Folge g_j auf D gleichmäßig konvergiert. Selbstverständlich ist

$$\lim_{j \rightarrow \infty} g_j(x) = \frac{f(x) - f(\xi)}{x - \xi} =: g(x).$$

Nun existieren aber nach Voraussetzung die Grenzwerte

$$\lim_{x \rightarrow \xi} g_j(x) = \lim_{x \rightarrow \xi} \frac{f_j(x) - f_j(\xi)}{x - \xi} = f'_j(\xi),$$

so daß aus dem zuvor bewiesenen Satz alles Weitere folgt:

$$\lim_{x \rightarrow \xi} \frac{f(x) - f(\xi)}{x - \xi} = \lim_{x \rightarrow \xi} \lim_{j \rightarrow \infty} g_j(x) = \lim_{j \rightarrow \infty} \lim_{x \rightarrow \xi} g_j(x) = \lim_{j \rightarrow \infty} f'_j(\xi). \quad \square$$

Anhang: Weitere Charakterisierungen der reellen Zahlen

Wir fassen in diesem Anhang kurz die Aussagen zusammen, die wir im vorangehenden Kapitel als äquivalent zu der Vollständigkeit des angeordneten Körpers \mathbb{K} eingesehen haben. Es handelt sich dabei also erneut um Axiome der reellen Zahlen.

(XXXV) Satz von Rolle

(XXXVI) Verallgemeinerter Mittelwertsatz

(XXXVII) Mittelwertsatz

(XXXVIII) Taylor-Formel mit Restglied nach Schlömilch

(XXXIX) Taylor-Formel mit Restglied nach Lagrange

(XL) Taylor-Formel mit Restglied nach Cauchy

(XLI) Charakterisierung von
polynomialen Funktionen
durch $f^{(n+1)} \equiv 0$

(XLII) $f' \equiv 0 \implies f = \text{const.}$

(XLIII) $f' \geq 0 \implies f$ monoton wachsend

(XLIV) $f'' \geq 0 \implies f$ konvex

Die Äquivalenz von (XXXV), (XXXVI), (XXXVII), (XLI) und (XLII) mit $\mathbb{K} \cong \mathbb{R}$ haben wir schon in Satz 16 bewiesen. Aus dem verallgemeinerten Mittelwertsatz folgt die Taylor-Formel mit dem Schlömilchschen Restglied, die (XXXIX) und (XL) impliziert. Die letzten beiden Aussagen haben aber die Charakterisierung (XLI) der Polynome zur Folge. (XLIII) folgt aus dem Mittelwertsatz, impliziert aber seinerseits wieder (XLII), da sich mit $f' = 0$ auch $f' \geq 0$ und $(-f)' \geq 0$ und daraus die wachsende Monotonie von f und $-f$ ergibt. Also muß f konstant sein. Auch (XLIV) gewinnt man auf bekannte Weise aus Eigenschaften des reellen Zahlkörpers. Ist umgekehrt (XLIV) erfüllt und f eine differenzierbare Funktion mit $f' = 0$, so ist f auch zweimal differenzierbar mit $f'' = 0$. Nach (XLIII) muß dann sowohl f als auch $-f$ konvex sein; also ist f eine affin-lineare Funktion: $f(x) = dx + c$. Da aber die erste Ableitung verschwindet, ist notwendig $d = 0$ und damit $f = c$ konstant. Somit ist wiederum notwendigerweise (XLII) erfüllt. \square

Bemerkung. Verlangt man die Richtigkeit z. B. des Zwischenwertsatzes oder des Mittelwertsatzes nur für *Polynome* anstelle von stetigen oder differenzierbaren Funktionen, so stößt man auf die weit größere Klasse der von Emil Artin und Otto Schreier eingeführten und intensiv studierten *reell abgeschlossenen Körper*.

15 Kompakte Räume und stetige Funktionen

Beschränkte, abgeschlossene Intervalle $[a, b] \subset \mathbb{R}$ werden, wie wir im vorletzten Kapitel schon angemerkt haben, auch als *kompakt* bezeichnet. *Stetige* Funktionen verhalten sich besonders angenehm auf solchen Intervallen; sie sind nicht nur *beschränkt* und nehmen ihr *Maximum* und *Minimum* an, sondern lassen sich auch sehr gut durch *Treppenfunktionen* approximieren, was die Grundlage für die Riemannsche Integrationstheorie darstellt. Dies liegt an einer *Überdeckungseigenschaft*, die man auch im allgemeineren Rahmen von normierten Vektorräumen (und sogar topologischen Räumen) studieren kann und sogar muß, wie wir später noch sehen werden. All dies soll in diesem Paragraphen untersucht werden. Wir werden wieder (scheinbar) in beliebigen angeordneten Körpern arbeiten, um bei den Beweisen deutlich herauszuheben, welches Axiom der reellen Zahlen wir jeweils benutzen. Im Anhang werden wir jedoch sehen, daß diese zentralen Sätze der Analysis (wenn wir uns gegebenenfalls auf archimedisch angeordnete Körper oder solche mit nichttrivialen Nullfolgen beschränken) tatsächlich *nur* im Körper der reellen Zahlen gelten.

Als erstes wollen wir in diesem Kapitel die eingangs angesprochene *Kompaktheitsaussage* exakt formulieren und beweisen. Wir werden anschließend an mehreren Folgerungen (insbesondere durch erneuten Beweis von Satz 13.17) demonstrieren, welche weitreichenden Konsequenzen diese Eigenschaft nach sich zieht. - Wir benötigen zuerst einige neue Begriffe.

Definition. Eine (*offene*) *Überdeckung* einer Teilmenge $M \subset \mathbb{K}$ ist eine Familie $(U_\iota)_{\iota \in J}$ von *offenen* Mengen $U_\iota \subset \mathbb{K}$ mit

$$M \subset \bigcup_{\iota \in J} U_\iota .$$

Eine Menge M heißt *kompakt*, wenn man aus *jeder* offenen Überdeckung $(U_\iota)_{\iota \in J}$ von M eine *endliche Teilüberdeckung* auswählen kann, d. h.: Es gibt zu jeder solchen Überdeckung mit Indexmenge J eine *endliche* Teilmenge $J_0 \subset J$, so daß

$$M \subset \bigcup_{\iota \in J_0} U_\iota .$$

Kompakte Mengen unterliegen starken Einschränkungen. Es gilt z. B. das folgende

Lemma 15.1 *Kompakte Mengen M in angeordneten Körpern \mathbb{K} sind beschränkt und abgeschlossen.*

Beweis. Es ist $M \subset \mathbb{K} = \bigcup_{r \in \mathbb{K}_+} (-r, r)$. Ist M kompakt, so reichen endlich viele solcher Intervalle mit Radien r_j , $j = 1, \dots, n$, zur Überdeckung aus. Setzt man $R := \max r_j$, so ist $M \subset (-R, R)$. Somit ist M beschränkt.

Um zu beweisen, daß M abgeschlossen ist, muß man per definitionem zeigen, daß das *Komplement* $\mathbb{K} \setminus M$ offen ist. Es sei also c ein Element des Komplements. Für jedes $m \in M$ ist dann $\varepsilon_m := |m - c| > 0$ und damit das System der

$$U_m := \{x \in \mathbb{K} : |x - m| < \varepsilon_m/2\}$$

eine offene Überdeckung von M . Da M kompakt ist, gibt es endlich viele Punkte m_j , $j = 1, \dots, n$, in M , so daß die endlich vielen offenen Mengen $U_j := U_{m_j}$ das Kompaktum M überdecken. Ist $\varepsilon := (1/2) \min \varepsilon_j$, $\varepsilon_j := \varepsilon_{m_j}$, so ist $U := \{x \in \mathbb{K} : |x - c| < \varepsilon\} \subset \mathbb{K} \setminus M$, d. h. das Komplement ist offen. \square

Bemerkung. Man kann leicht zu jedem (halb)offenen (oder unbeschränkten) Intervall in einem *archimedisch* angeordneten Körper \mathbb{K} *abzählbare* offene Überdeckungen $(U_\iota)_{\iota \in \mathbb{N}}$ angeben, aus denen man keine endliche Teilüberdeckung auswählen kann. Z. B. ist

$$(0, 1] = \bigcup_{n=0}^{\infty} (1/n, 1] .$$

Die obigen Konzepte und Ergebnisse lassen sich im abstrakteren Rahmen einführen und herleiten.

Definition. Ein \mathbb{K}_0 -metrischer Raum X heißt *kompakt*, wenn jede offene Überdeckung von X eine endliche Teilüberdeckung besitzt. Dieser Begriff ist auch für beliebige *topologische Räume* X sinnvoll. Man verwendet ihn dann aber nur, wenn X zusätzlich *hausdorffsch* ist. Eine Teilmenge $M \subset X$ wird *kompakt* genannt, wenn M in der *Relativtopologie* kompakt ist. Dies ist offensichtlich genau die oben gegebene Definition im Fall $X = \mathbb{K}$.

Man erhält sofort mutatis mutandis den folgenden Satz.

Satz 15.2 *Kompakte Mengen M in topologischen (hausdorffschen) Räumen X sind notwendig abgeschlossen. Ist X sogar ein \mathbb{K}_0 -metrischer Raum, so sind kompakte Mengen auch beschränkt (und folgenabgeschlossen, wenn \mathbb{K}_0 nichttriviale Nullfolgen besitzt).*

Wir gehen im folgenden genauso wie bei den zusammenhängenden Mengen vor, indem wir zeigen, daß das abgeschlossene Einheitsintervall in \mathbb{R} tatsächlich kompakt ist und stetige Bilder von kompakten Mengen wieder kompakt sind.

Satz 15.3 *Erfüllt der angeordnete Körper \mathbb{K} das (schwache) Intervallschachtelungsprinzip, so ist das Einheitsintervall $I = [0, 1]$ kompakt.*

Beweis. Es sei $\bigcup_{\iota \in J} U_\iota \supset I$ eine offene Überdeckung, aus der man keine endliche Teilüberdeckung auswählen kann. Dies bleibt dann richtig für mindestens eines der beiden Intervalle $[0, \frac{1}{2}]$ und $[\frac{1}{2}, 1]$. So fortfahrend findet man eine Intervallschachtelung $I = I_0 \supset I_1 \supset I_2 \supset \dots$, so daß kein I_j von endlich vielen U_ι überdeckt wird. Es sei nun $x_0 \in \bigcap_{j=0}^{\infty} I_j$. Dann gibt es ein $\iota_0 \in J$ mit $x_0 \in U_{\iota_0}$ und damit auch ein j_0 mit $I_{j_0} \subset U_{\iota_0}$. Widerspruch! \square

Bemerkung. Der gleiche Beweis funktioniert auch für beliebige Intervalle $[a, b]$. Man kann diesen Fall (wobei man $a < b$ voraussetzen darf, da einpunktige Intervalle offensichtlich immer kompakt sind) aber auch auf den ersten zurückführen, indem man das folgende Lemma anwendet.

Lemma 15.4 *Ist $M \subset \mathbb{K}$ kompakt und $f : M \rightarrow \mathbb{K}$ eine stetige Funktion, so ist das Bild $f(M)$ kompakt.*

Beweis. Ist (V_ι) eine offene Überdeckung von $f(M)$, so existieren offene Mengen U_ι in \mathbb{K} , so daß $f^{-1}(f(M) \cap V_\iota) = M \cap U_\iota$. Also ist, wie man sofort einsieht, (U_ι) eine offene Überdeckung von M , aus der man ein durch J_0 indiziertes *endliches* Teilsystem auswählen kann, das M schon überdeckt. Es folgt

$$f(M) \subset f\left(\bigcup_{\iota \in J_0} f^{-1}(V_\iota)\right) \subset \bigcup_{\iota \in J_0} f(f^{-1}(V_\iota)) \subset \bigcup_{\iota \in J_0} V_\iota \quad \square$$

Folgerung 15.5 *Ist ein einziges nichttriviales beschränktes abgeschlossenes Intervall I in dem angeordneten Körper \mathbb{K} kompakt, so auch jedes andere solche Intervall.*

Denn jedes abgeschlossene beschränkte Intervall ist stetiges Bild von I . \square

Bemerkungen. 1. Ein bewerteter Körper \mathbb{K} heißt *lokalkompakt*, wenn die Null in \mathbb{K} eine *kompakte Umgebung* besitzt, also eine *kompakte* Menge K , in der eine *offene* Menge U liegt mit $0 \in U \subset K$. Die obigen Aussagen implizieren, daß \mathbb{R} lokalkompakt ist (es ist $0 \in (-1, 1) \subset [-1, 1]$). Wir werden weiter unten einsehen, daß \mathbb{R} der einzige lokalkompakte Körper unter den *angeordneten* Körpern ist. Man beachte aber, daß es eine Klassifikationsliste aller *bewerteten* lokalkompakten Körper gibt, die

nicht nur den reellen Zahlkörper enthält.

2. Aus der vorigen Folgerung kann man Satz 13.17 erneut (in seiner verschärften Fassung in Form von Folgerung 13.18) gewinnen. Denn ist \mathbb{K} der reelle Zahlkörper, so ist jedes Intervall $I = [a, b]$ sowohl kompakt nach Satz 3 als auch zusammenhängend nach den Ergebnissen in Kapitel 13. Also ist das Bild $f(I)$ unter einer stetigen Funktion $f : [a, b] \rightarrow \mathbb{K}$ kompakt und zusammenhängend, also insbesondere ein beschränktes Intervall. Da kompakte Intervalle in archimedisch angeordneten Körpern aber weder halboffen noch offen sein können, wie wir oben angemerkt haben, ist notwendig $f(I) = [c, d]$. \square

Aus der Lokalkompaktheit eines angeordneten Körpers kann man eine nützliche „Gleichmäßigkeitsaussage“ ableiten, die Auswirkungen auf das Verhalten von stetigen Funktionen auf kompakten Intervallen hat.

Definition. Wir sagen, in dem angeordneten Körper gelte das *Lebesguesche Lemma*, wenn es zu jedem Intervall $I = [a, b] \subset \mathbb{K}$ und jeder offenen Überdeckung

$$I \subset \bigcup_{\iota \in J} U_\iota$$

von I eine Zahl $\delta > 0$ gibt (die auch *Lebesguesche Zahl der Überdeckung* genannt wird), s. d. es zu je zwei Elementen $x_1, x_2 \in [a, b]$ mit $|x_1 - x_2| \leq \delta$ ein $\iota \in J$ gibt mit $x_1, x_2 \in U_\iota$.

Satz 15.6 *Gibt es in dem angeordneten Körper \mathbb{K} nichttriviale kompakte Intervalle, so erfüllt er die Aussage des Lebesgueschen Lemmas.*

Beweis. Das Intervall $I = [a, b]$ werde von den offenen Mengen U_ι , $\iota \in J$, überdeckt. Jedes U_ι ist Vereinigung von Intervallen $\{|x - x_0| < \delta_0\}$, $\delta_0 := \delta(x_0)$, mit

$$\{|x - x_0| < 2\delta_0\} \subset U_\iota.$$

Da das Intervall $[a, b]$ kompakt ist, wird es von endlich vielen solchen Intervallen überdeckt:

$$[a, b] \subset \bigcup_{j=1}^n \{|x - x_j| < \delta_j\}, \quad \{|x - x_j| < 2\delta_j\} \subset U_{\iota_j}.$$

Es sei nun $\delta := \min(\delta_1, \dots, \delta_n)$ und $x \in [a, b]$, also $|x - x_j| < \delta_j$ für ein j , und x' ein weiteres Element mit $|x - x'| \leq \delta$. Dann folgt $|x' - x_j| \leq |x' - x| + |x - x_j| < \delta + \delta_j \leq 2\delta_j$ und insbesondere $x, x' \in U_{\iota_j}$. \square

Das Lebesguesche Lemma hat in *archimedisch* angeordneten Körpern zur Folge, daß stetige Funktionen $f : I = [a, b] \rightarrow \mathbb{K}$ sogar *gleichmäßig* stetig sind. Dieser Sachverhalt ist (in einer äquivalenten Formulierung) die Grundlage der RIEMANNschen *Integrationstheorie* für stetige Funktionen auf kompakten reellen Intervallen.

Satz 15.7 *Ist \mathbb{K} ein archimedisch angeordneter Körper, der dem Lebesgueschen Lemma genügt, so gibt es für jede stetige Funktion $f : [a, b] \rightarrow \mathbb{K}$ und zu jedem $\varepsilon > 0$ eine endliche (sogar äquidistante) Unterteilung $a_0 = a < a_1 < \dots < a_n = b$ von $[a, b]$, s. d. für alle x_j, ξ_j mit $a_{j-1} \leq x_j, \xi_j \leq a_j$ gilt:*

$$|f(x_j) - f(\xi_j)| < \varepsilon.$$

Beweis. Es sei $\varepsilon > 0$ vorgegeben. Dann gibt es zu jedem $x_0 \in [a, b]$ ein $\delta_0 = \delta(x_0)$, s. d. aus $|x - x_0| < \delta_0$, $x \in [a, b]$ die Ungleichung $|f(x) - f(x_0)| < \varepsilon/2$ folgt. Es sei $\delta > 0$ die Lebesguesche Zahl zu der Überdeckung

$$\bigcup_{x_0 \in [a, b]} \{|x - x_0| < \delta_0\},$$

und $n \in \mathbb{N}$ sei so gewählt, daß

$$\frac{1}{n}(b-a) \leq \frac{\delta}{2}.$$

Für die Zerlegung $a_j = a + \frac{j}{n}(b-a)$, $j = 0, 1, \dots, n$, ist die Behauptung dann erfüllt. \square

Hieraus folgt unter den gegebenen Voraussetzungen die *gleichmäßige* Stetigkeit von f auf I . Wir geben hierzu gleich eine allgemeine

Definition. Es sei $f : X \rightarrow Y$ eine stetige Abbildung zwischen metrischen Räumen. Man nennt f *gleichmäßig stetig*, wenn es zu jedem $\varepsilon > 0$ ein $\delta > 0$ gibt, so daß für $x, x' \in X$ mit $d_X(x, x') < \delta$ stets $d_Y(f(x), f(x')) < \varepsilon$ folgt.

Bemerkung. Setzt man speziell $x' = a$ für einen festen Punkt $a \in X$, so sieht man sofort, daß jede gleichmäßig stetige Funktion insbesondere an jeder Stelle a stetig ist. Die Umkehrung gilt nicht, wie man an dem Beispiel der Funktion $f(x) := 1/x$ auf dem Intervall $(0, 1]$ sieht.

Lemma 15.8 *Gibt es zu einer stetigen Funktion $f : I = [a, b] \rightarrow \mathbb{K}$ für jedes $\varepsilon > 0$ eine endliche Unterteilung $a_0 = a < a_1 < \dots < a_n = b$ von $[a, b]$, s. d. für alle x_j, ξ_j mit $a_{j-1} \leq x_j, \xi_j \leq a_j$ gilt :*

$$|f(x_j) - f(\xi_j)| < \varepsilon,$$

so ist f gleichmäßig stetig. Die Umkehrung gilt auf jeden Fall, wenn \mathbb{K} archimedisch angeordnet ist.

Beweis. Wähle zu vorgegebenem $\varepsilon > 0$ die Unterteilung so, daß $|f(x_j) - f(\xi_j)| < \varepsilon/2$. Ist dann $\delta := \min(a_j - a_{j-1})$, so sind je zwei Elemente $x, x' \in I$ mit $|x - x'| < \delta$ in demselben Intervall $I_j := [a_{j-1}, a_j]$ oder in benachbarten Intervallen I_j und I_{j+1} enthalten. Es folgt auf jeden Fall

$$|f(x) - f(x')| \leq |f(x) - f(a_j)| + |f(x') - f(a_j)| < \varepsilon.$$

Ist umgekehrt f gleichmäßig stetig, so wähle man zu vorgegebenem $\varepsilon > 0$ erst das zugehörige δ und anschließend $n \in \mathbb{N}$ so groß, daß $(b-a)/n \leq \delta$. Für die Zerlegung $a_j = a + j(b-a)/n$, $j = 0, 1, \dots, n$, ist die Behauptung erfüllt. \square

Das vorstehende Lemma können wir noch in einer anderen Weise formulieren, die auf die Bedürfnisse der *Integrationstheorie* zugeschnitten ist.

Definition. Eine Funktion $\varphi : I = [a, b] \rightarrow \mathbb{K}$ heißt eine *Treppenfunktion*, wenn es eine endliche Unterteilung $a_0 = a < a_1 < \dots < a_n = b$ von I gibt, so daß φ eingeschränkt auf die offenen Intervalle (a_{j-1}, a_j) , $j = 1, \dots, n$, konstant ist. (Über die Werte von φ an den Stellen a_j setzen wir nichts voraus).

Bemerkung. Treppenfunktionen kann man auf elementare Weise ein *Integral* zuordnen. Das Ziel der Integrationstheorie ist es, diesen Integralbegriff auf größere Funktionsklassen auszudehnen. Daß dies für *stetige* Funktionen gelingt, liegt an dem folgenden Satz.

Lemma 15.9 *Unter der Voraussetzung des vorigen Lemmas läßt sich die stetige Funktion $f : I \rightarrow \mathbb{K}$ (in der Supremumsnorm) beliebig genau durch eine Treppenfunktion approximieren.*

Beweis. Wird zu vorgegebenem $\varepsilon > 0$ eine endliche Unterteilung $a_0 = a < a_1 < \dots < a_n = b$ von I gewählt, so daß

$$|f(x_j) - f(\xi_j)| < \varepsilon,$$

für alle x_j, ξ_j mit $a_{j-1} \leq x_j, \xi_j \leq a_j$, so wird z. B. durch $\varphi(a_j) := f(a_j)$, $\varphi(x) := f(\xi_j)$ für $x \in (a_{j-1}, a_j)$ mit jeweils fest gewähltem ξ_j in diesem Intervall eine Treppenfunktion φ definiert,

für die $|\varphi(x) - f(x)| \leq \varepsilon$ für alle $x \in I$ gilt. \square

Wir fassen noch einmal zusammen: *Ist $I = [a, b] \subset \mathbb{R}$ ein abgeschlossenes reelles Intervall, so ist dieses kompakt, es gilt das Lebesguesche Lemma für I , und stetige Funktionen $f : I \rightarrow \mathbb{R}$ sind sogar gleichmäßig stetig und lassen sich beliebig genau durch Treppenfunktionen approximieren.*

Bevor wir noch einige Folgerungen herleiten und (im Anhang) weitere äquivalente Charakterisierungen für den Körper der reellen Zahlen angeben, halten wir hier kurz inne, um uns zu fragen, welche der vorigen Resultate verallgemeinerungsfähig sind. Wir überlassen der Leserin und dem Leser die (leichte) Mühe, sich an Hand des obigen Beweises von der Richtigkeit der folgenden Aussage zu überzeugen.

Satz 15.10 *Ist $f : X \rightarrow Y$ eine stetige Abbildung zwischen topologischen Hausdorff-Räumen und ist $M \subset X$ kompakt, so auch das Bild $f(M) \subset Y$.*

Ist $f : X \rightarrow \mathbb{R}$ eine stetige reellwertige Funktion auf dem kompakten Raum X , so ist die Bildmenge $f(X) \subset \mathbb{R}$ kompakt, also beschränkt und abgeschlossen; sie nimmt somit ihr Supremum an. Ist X (zudem) zusammenhängend, so auch die Bildmenge, die dann nach früheren Sätzen ein (kompaktes) Intervall sein muß. Wir erhalten somit als eine wesentliche Verschärfung früherer Sätze über stetige Funktionen auf kompakten Intervallen, insbesondere als Verallgemeinerung des *Zwischenwertsatzes*:

Satz 15.11 *Stetige reellwertige Funktionen f auf kompakten Räumen X nehmen ihr Maximum und Minimum an. Ist X zusammenhängend, so wird jeder Wert zwischen zwei beliebig gewählten Funktionswerten f angenommen.*

Wir beweisen außerdem noch einmal *direkt* den folgenden Satz, den wir weiter oben in einem Spezialfall durch Kombination zweier Ergebnisse gefunden haben.

Satz 15.12 *Ist $f : X \rightarrow Y$ eine stetige Abbildung zwischen metrischen Räumen, von denen X kompakt ist, so ist f gleichmäßig stetig.*

Beweis. Es werde $\varepsilon > 0$ fest gewählt. Dann gibt es zu jedem $a \in X$ ein $\delta = \delta(a)$, so daß $d_Y(f(x), f(a)) < \varepsilon/2$, wenn $d_X(x, a) < \delta$. Nach Voraussetzung besitzt die offene Überdeckung

$$(U_a)_{a \in X}, \quad U_a := B(a, \delta(a)/2)$$

eine endliche Teilüberdeckung. Wird diese durch die Punkte a_1, \dots, a_n indiziert: $U_j := U_{a_j}$, und schreiben wir $\delta_j := \delta(a_j)/2$, so ist mit $\delta := \min \delta_j$, $x \in U_j$ und $d_X(x, x') < \delta$ offensichtlich auch $x' \in U_j$, und damit folgt für solche x, x' sofort

$$d_Y(f(x), f(x')) \leq d_Y(f(x), f(a_j)) + d_Y(f(x'), f(a_j)) < 2(\varepsilon/2) = \varepsilon. \quad \square$$

Die Frage nach der Angabe *aller* kompakten Teilmengen eines angeordneten Körpers läßt sich nicht so leicht beantworten wie im Falle der zusammenhängenden Mengen. - Wir geben als erstes ein

Beispiel. Ist X ein hausdorffscher topologischer Raum und (x_j) eine *konvergente* Folge in X mit Grenzwert a , so ist die Menge $M := \{x_j : j \in \mathbb{N}\} \cup \{a\}$ kompakt. Den einfachen Beweis überlassen wir dem Leser.

Bevor wir ein wenig genauer auf die kompakten Teilmengen von \mathbb{R} eingehen, leiten wir ein hinreichendes Kriterium in topologischen Räumen her (notwendige Bedingungen haben wir schon in Satz 2 kennengelernt).

Lemma 15.13 *Ist X ein Hausdorff-Raum, und sind A und K eine abgeschlossene bzw. kompakte Teilmenge von X mit $A \subset K$, so ist auch A kompakt.*

Beweis. Ist $(U_{i \in J})$ eine offene Überdeckung von A , so ist $K \subset (X \setminus A) \cup \bigcup_{i \in J} U_i$ eine offene Überdeckung von K . Wählt man aus dieser eine endliche Teilüberdeckung von K aus und läßt gegebenenfalls, wenn vorkommend, die Menge $X \setminus A$ fort, so kommt zu einer endlichen Teilüberdeckung von A , die aus Elementen der ursprünglichen Überdeckung besteht. \square

Folgerung 15.14 *In einem \mathbb{K}_0 -metrischen Raum (X, d) sind die folgenden Aussagen äquivalent :*

- i) *alle abgeschlossenen Kugeln $\overline{B}(a, R)$ sind kompakt ;*
- ii) *die kompakten Teilmengen von X sind genau die beschränkten und abgeschlossenen Mengen.*

Beweis. Ist ii) erfüllt, so müssen die beschränkten und abgeschlossenen Kugeln $\overline{B}(a, R)$ kompakt sein. Ist andererseits i) erfüllt und K eine beschränkte und abgeschlossene Teilmenge von X , so ist sie insbesondere in einer offenen Kugel $B(a, R)$ und damit auch in der abgeschlossenen Kugel $\overline{B}(a, R)$ enthalten. Nach dem vorigen Lemma ist K kompakt. Umgekehrt ist jede kompakte Teilmenge in X notwendig beschränkt und abgeschlossen. \square

In \mathbb{R} ist die Aussage i) der vorigen Folgerung wegen der Kompaktheit aller Intervalle $[-R, R]$ erfüllt. Es ergibt sich damit sofort:

Satz 15.15 (Heine - Borel) *Eine Teilmenge in \mathbb{R} ist genau dann kompakt, wenn sie beschränkt und abgeschlossen ist.*

Beispiel. Eine sehr interessante kompakte Menge hat CANTOR konstruiert. Man setze für ein beliebiges kompaktes Intervall $I = [a, b]$, $a < b$:

$$I_0 := [a, a + (b - a)/3], \quad I_2 := [a + 2(b - a)/3, b]$$

und induktiv für $j_1, \dots, j_k \in \{0, 2\}$:

$$I_{j_1, \dots, j_k} := (I_{j_1, \dots, j_{k-1}})_{j_k}.$$

Dann ist bei festem k

$$I^{(k)} := \bigcup I_{j_1, \dots, j_k}$$

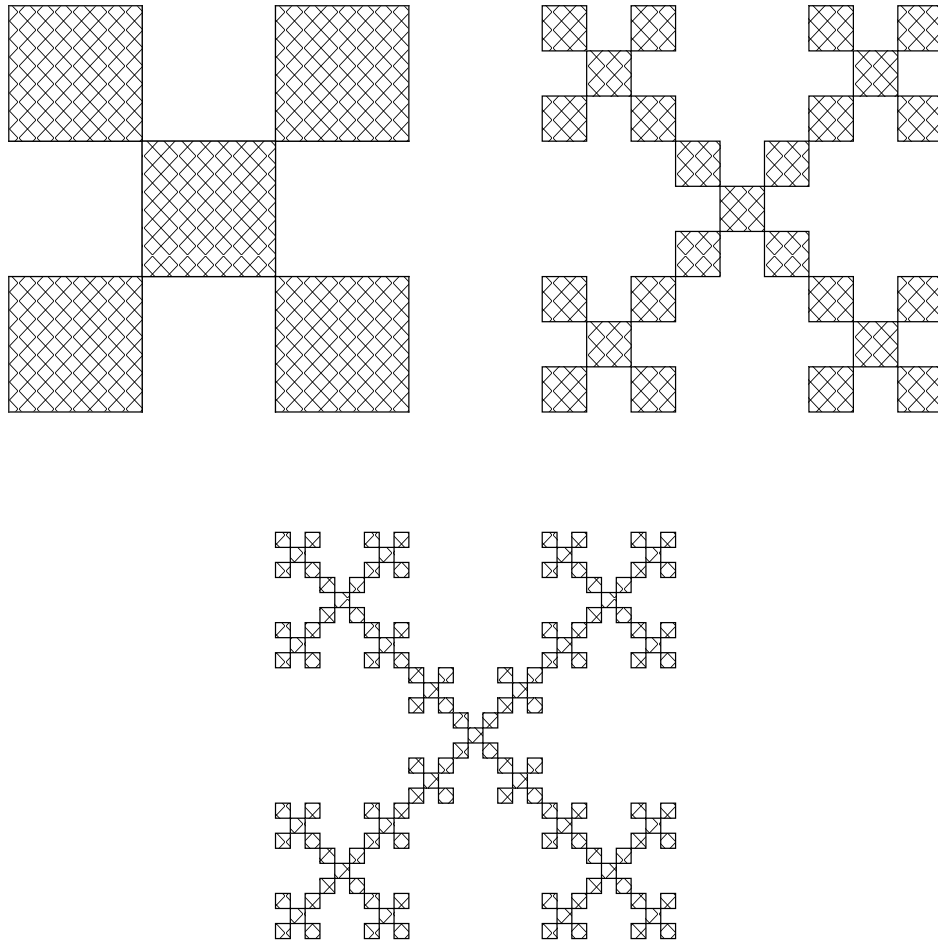
Vereinigung von endlich vielen abgeschlossenen Teilintervallen von I , also insbesondere eine kompakte Menge, und es gilt $I \supset I^{(0)} \supset I^{(1)} \supset \dots$. Der Durchschnitt

$$C := \bigcap_{k=0}^{\infty} I^{(k)}$$

ist dann ebenfalls abgeschlossen und beschränkt, also kompakt. Man nennt C das CANTORSche Diskontinuum oder auch die 1/3-Menge. C hat die folgenden merkwürdigen Eigenschaften:

1. C hat dieselbe Mächtigkeit wie \mathbb{R} , ist also insbesondere überabzählbar;
2. C liegt nirgends dicht in I , d. h. $C \cap J$ ist niemals dicht in J für beliebige nichttriviale Intervalle $J \subset I$;
3. C ist total unzusammenhängend;
4. jeder Punkt von C ist Häufungspunkt von C ;
5. C ist eine LEBESGUESche Nullmenge, d. h. zu jedem $\varepsilon > 0$ gibt es abzählbar viele Intervalle $I_j^\varepsilon = [a_j^\varepsilon, b_j^\varepsilon]$ mit $C \subset \bigcup_j I_j^\varepsilon$ und $\sum_j (b_j^\varepsilon - a_j^\varepsilon) \leq \varepsilon$.

Hier ist eine Bildfolge zum Entstehungsprozeß eines *zweidimensionalen Analogons* zur Cantor-Menge (siehe [11]), die sowohl die weitere Bezeichnung als CANTORSche *Wischmenge* beleuchtet als auch eine wesentliche Eigenschaft solcher *Fraktale* verdeutlicht, nämlich ihre *Selbstähnlichkeit*:



Figur 15.1

Tatsächlich besteht die Cantor-Menge C genau aus den Zahlen der Gestalt

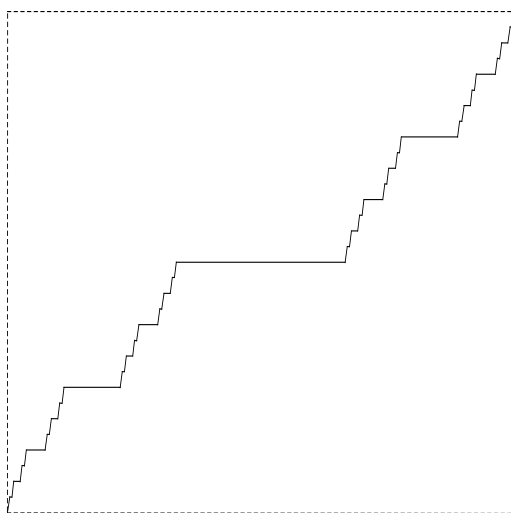
$$x = \sum_{n=1}^{\infty} \frac{a_n}{3^n} \quad \text{mit} \quad a_n \in \{0, 2\}.$$

Die durch

$$C \ni x \mapsto \varphi(x) := \sum_{n=1}^{\infty} \frac{a_n}{2^{n+1}}$$

definierte CANTOR-Funktion $\varphi : C \rightarrow I := [0, 1]$ ist *monoton wachsend*, *stetig* und *surjektiv*. Sie besitzt eine *stetige Fortsetzung* $\Theta : I \rightarrow I$, die auf jedem der offenen Intervalle, in die $I \setminus C$ zerfällt, konstant ist.

Ein (natürlich zum Scheitern verurteilter) Versuch, die Cantor-Funktion *graphisch* darzustellen, möge den Leser erfreuen und ihn zumindest für kurze Zeit den hohen Abstraktheitsgrad des vorliegenden Kapitels vergessen machen.



Figur 15.2

Wie ist nun die Situation im \mathbb{R}^n (mit der üblichen euklidischen Topologie)? Eine kurze Analyse des Beweises von Satz 3 zeigt, daß auch die folgende Aussage richtig ist:

Satz 15.16 Sind $I_1, \dots, I_n \subset \mathbb{R}$ kompakte Intervalle, so ist auch der Quader $Q := I_1 \times \dots \times I_n$ kompakt in \mathbb{R}^n . \square

Dieselbe Beweisidee führt ohne große Schwierigkeiten sogar zu einem noch viel allgemeineren Satz.

Satz 15.17 Sind X und Y kompakte topologische Räume, so ist auch das kartesische Produkt $X \times Y$ mit der Produkttopologie ein kompakter Raum. \square

Nach Folgerung 14 ist damit eine Menge $K \subset \mathbb{R}^n$ genau dann kompakt, wenn sie abgeschlossen ist und in einem abgeschlossenen Quader, insbesondere also z. B. in einem Würfel mit Mittelpunkt 0 enthalten ist. Nun ist aber jede (euklidische) Kugel mit Mittelpunkt 0 in einem solchen Würfel enthalten. - Wir können also festhalten:

Satz 15.18 (Heine - Borel) Eine Teilmenge $K \subset \mathbb{R}^n$ ist (in der üblichen Topologie) genau dann kompakt, wenn sie in der euklidischen Norm beschränkt und abgeschlossen ist. \square

Bemerkung. Da auf \mathbb{R}^n alle Normen äquivalent sind (siehe den Anhang zu Kapitel 17), kann man in dem vorigen Satz die euklidische Norm auf \mathbb{R}^n auch durch jede andere ersetzen.

Warnung. Trotzdem darf man die Begriffe *kompakt* einerseits und *beschränkt und abgeschlossen* andererseits nicht gleichsetzen. In einem normierten \mathbb{R} -Vektorraum V ist nämlich die Bedingung i) in der obigen Folgerung 14 nur dann erfüllt, wenn V endlich-dimensional ist. (Siehe hierzu den Anhang zu Kapitel 17).

Der Satz von Heine-Borel liefert ein sehr einfaches Kriterium, um Kompaktheit zu testen. Oft werden Teilmengen $A \subset \mathbb{R}^n$ beschrieben in der Form

$$A = \{x \in \mathbb{R}^n : f_1(x) = \dots = f_r(x) = 0, \quad g_1(x) \geq 0, \dots, g_s(x) \geq 0\}$$

mit stetigen Funktionen $f_1, \dots, f_r, g_1, \dots, g_s : \mathbb{R}^n \rightarrow \mathbb{R}$. Aus dem Folgenkriterium folgt sofort, daß A in \mathbb{R}^n abgeschlossen ist. Kann man noch die Beschränktheit von A nachweisen, so ist A kompakt.

Beispiele. 1. Offensichtlich ist jedes endliche Produkt $I_1 \times \cdots \times I_n \subset \mathbb{R}^n$ von *kompakten* Intervallen $I_j = [a_j, b_j] \subset \mathbb{R}$ von dieser Bauart.

2. Für jedes $r \geq 0$ ist $\overline{B}_r = \overline{B}(r, 0) = \{x \in \mathbb{R}^n : r^2 - \sum_{j=1}^n x_j^2 \geq 0\}$.

3. Entsprechend ist $S_r = \partial \overline{B}_r = \{x \in \mathbb{R}^n : r^2 - \sum_{j=1}^n x_j^2 = 0\}$.

Zum Abschluß dieses Kapitels wollen wir noch einige Eigenschaften kompakter Mengen und weitere Kompaktheitskonzepte in \mathbb{K}_0 -metrischen Räumen X diskutieren. Wir setzen überall in diesem Abschnitt voraus, daß der angeordnete Körper \mathbb{K}_0 der Bedingung (*) genügt, d. h. nichttriviale Nullfolgen besitzt.

Völlig banal ist die folgende Bemerkung, die auch in beliebigen topologischen Räumen gültig ist und nur der Vollständigkeit halber hier aufgeführt sei.

Satz 15.19 Sind $A_1, \dots, A_n \subset X$ kompakte Teilmengen, so auch $A = \bigcup_{j=1}^n A_j$.

Des weiteren haben wir:

Satz 15.20 Ist $A \subset X$ eine kompakte Teilmenge, so ist A (als metrischer Raum) vollständig und beschränkt. Zudem ist $A \subset X$ (folgen-) abgeschlossen.

Beweis. Wir werden anschließend in Satz 21 zeigen, daß bei der General-Voraussetzung an \mathbb{K}_0 jeder kompakte \mathbb{K}_0 -metrische Raum *vollständig* ist. Setzen wir dies hier voraus, so ergibt sich sofort die (Folgen-) Abgeschlossenheit: Ist $x_j \in A$, $\lim x_j = x \in X$, so ist die Folge (x_j) eine Cauchy-Folge in X und damit auch in A . Da A vollständig ist, muß x in A liegen.

Nach Voraussetzung besitzt \mathbb{K}_0 auch monoton aufsteigende, (nach oben) unbeschränkte Folgen $0 < N_0 < N_1 < N_2 < \cdots$. Es sei $a \in X$ fest und $U_j = B(a, N_j)$; dann gilt $U_0 \subset U_1 \subset \cdots$ und

$$A \subset X = \bigcup_{j=0}^{\infty} U_j.$$

Wegen der Kompaktheit von A existiert ein ℓ mit $A \subset U_\ell = B(a, N_\ell)$. □

Definition und Bemerkung. Ein \mathbb{K}_0 -metrischer Raum X heißt *präkompakt* oder auch *total beschränkt*, wenn er für jedes $\varepsilon > 0$ eine *endliche* Überdeckung mit Kugeln vom Radius ε gestattet. Dies ist offensichtlich äquivalent dazu, daß es zu jedem $\varepsilon > 0$ eine *endliche* Teilmenge $E \subset X$ gibt, so daß für alle $x \in X$ gilt: $d_X(x, E) := \min_{y \in E} d_X(x, y) < \varepsilon$. Man beachte, daß im Gegensatz zur *Kompaktheit* der Begriff der *Präkompaktheit* kein rein topologischer Begriff ist.

Wir leiten jetzt den zentralen Charakterisierungssatz für kompakte metrische Räume her.

Satz 15.21 Es sei (X, d) ein \mathbb{K}_0 -metrischer Raum, wobei \mathbb{K}_0 nichttriviale Nullfolgen besitze. Dann sind äquivalent :

- o) X ist kompakt ;
- i) X ist abzählbar kompakt, d. h. jede abzählbare offene Überdeckung von X besitzt eine endliche Teilüberdeckung ;
- ii) X ist folgenkompakt, d. h. jede Folge in X besitzt eine konvergente Teilfolge ;
- iii) X ist vollständig und präkompakt, d. h. zu jedem $\varepsilon > 0$ gibt es endliche viele Punkte x_1, \dots, x_N ,
s. d. $X = \bigcup_{j=1}^N B(x_j, \varepsilon)$.

Unter diesen Bedingungen gilt zusätzlich :

- iv) X besitzt eine abzählbare dichte Teilmenge ;
 v) X besitzt abzählbare Topologie, d. h. es gibt abzählbar viele offene Kugeln, B_0, B_1, \dots in X mit $U = \bigcup_{j: B_j \subset U} B_j$ für alle offenen Mengen U .

Beweis. o) \implies i) ist völlig banal und gilt ganz allgemein für beliebige topologische Räume.

i) \implies ii) Es sei (x_j) eine beliebige Folge. Wir müssen zeigen, daß sie einen Häufungspunkt besitzt (denn nach Voraussetzung an \mathbb{K}_0 gibt es dann auch eine konvergente Teilfolge). Setze also

$$R_n = \{x_j : j \geq n\}, \quad A_n := \overline{R_n}.$$

Ist $x \in \bigcap_{n=0}^{\infty} A_n$, so sieht man sofort, daß x ein Häufungspunkt der Folge ist. Es werde also angenommen,

daß $\bigcap_{n=0}^{\infty} A_n = \emptyset$. Dann ist $X = \bigcup_{n=0}^{\infty} U_n$, $U_n = X \setminus A_n$. Nach Voraussetzung gibt es eine endliche Teilmenge $N \subset \mathbb{N}$ mit $X = \bigcup_{n \in N} U_n$ und damit $\bigcap_{n \in N} A_n = \emptyset$. Dies kann aber nicht sein wegen

$$\bigcap_{n \in N} R_n \neq \emptyset.$$

ii) \implies iii) Sei (x_j) eine Cauchy-Folge. Nach ii) besitzt diese eine konvergente Teilfolge und ist damit konvergent. X ist also vollständig. Angenommen, X sei nicht präkompakt. Dann gibt es ein $\varepsilon > 0$, s. d. $X \setminus \bigcup_{j=1}^N B(x_j, \varepsilon) \neq \emptyset$ für beliebige endliche Punktfolgen $\{x_1, \dots, x_N\}$. Damit konstruiert man

eine Folge x_0, x_1, x_2, \dots , s. d. $x_{n+1} \notin \bigcup_{j=0}^n B(x_j, \varepsilon)$. Dies gilt dann auch für jede Teilfolge, so daß wir annehmen können, daß der Grenzwert $x = \lim x_j$ existiert. Insbesondere ist (x_j) eine Cauchy-Folge, s. d. $d_X(x_j, x_k) < \varepsilon$ für alle $j, k \geq N = N(\varepsilon)$. Dann folgt aber $x_{N+1} \in B(x_N, \varepsilon)$. Widerspruch!

iii) \implies o) Wir nehmen an, daß iii), aber nicht o) erfüllt sei, s. d. also eine offene Überdeckung $\bigcup_{\iota \in I} U_\iota$ von X existiert, die keine endliche Teilüberdeckung von X enthält. Es sei ε_n eine (ohne Einschränkung) streng monoton fallende Nullfolge in \mathbb{K}_0 und $\alpha_n := \varepsilon_n - \varepsilon_{n+1} > 0$, $n \in \mathbb{N}$. Wir konstruieren induktiv Kugeln $B_n = B_n(x_n, \alpha_n)$ mit $B_0 \cap B_1 \neq \emptyset$, $B_1 \cap B_2 \neq \emptyset$ etc., s. d. kein B_n von endlich vielen U_ι überdeckt wird. Es seien nämlich B_0, \dots, B_{n-1} schon vorhanden. Nach Voraussetzung ist

$$B_{n-1} \subset X = \bigcup_{j \in J_n} B(x_{n,j}, \alpha_n)$$

mit einer endlichen Menge J_n und damit für eine geeignete Teilmenge $J'_n \subset J_n$:

$$B_{n-1} \subset \bigcup_{j \in J'_n} B(x_{n,j}, \alpha_n), \quad B_{n-1} \cap B(x_{n,j}, \alpha_n) \neq \emptyset.$$

Ließe sich jede der Kugeln $B(x_{n,j}, \alpha_n)$ durch endlich viele U_ι überdecken, so auch B_{n-1} . Damit muß es eine unter ihnen geben, die auch die gewünschte Eigenschaft besitzt. Wir nennen sie $B_n = B(x_n, \alpha_n)$. Wir betrachten nun die Folge (x_n) . Aus $B_n \cap B_{n+1} \neq \emptyset$ ergibt sich

$$d_X(x_n, x_{n+1}) \leq \alpha_n + \alpha_{n+1}$$

und daraus durch Induktion für $m > n \geq 1$

$$\begin{aligned} d_X(x_n, x_m) &\leq d_X(x_n, x_{n+1}) + \cdots + d_X(x_{m-1}, x_m) \\ &\leq (\alpha_n + \alpha_{n+1}) + (\alpha_{n+1} + \alpha_{n+2}) + \cdots + (\alpha_{m-1} + \alpha_m) \\ &\leq 2 \sum_{j=n}^m \alpha_j = 2(\varepsilon_n - \varepsilon_{m+1}) < 2\varepsilon_n. \end{aligned}$$

Ist nun $\varepsilon > 0$ beliebig vorgegeben und $\varepsilon_n \leq \varepsilon/2$ für alle $n \geq N = N(\varepsilon)$, so ist für alle $m > n \geq N$:

$$d_X(x_n, x_m) < \varepsilon.$$

Folglich ist (x_n) eine Cauchy-Folge, die nach Voraussetzung konvergiert: Es existiert $a = \lim_{n \rightarrow \infty} x_n$. Da a in einer der offenen Mengen U_ι liegt, gibt es ein $\alpha > 0$ und einen Index ι_0 mit

$$B(a, \alpha) \subset U_{\iota_0}.$$

Für geeignetes großes N ist $d_X(a, x_N) < \alpha/2$ und $\alpha_N \leq \varepsilon_N < \alpha/2$. Aus der Dreiecksungleichung folgt dann

$$B_N = B(x_N, \alpha_N) \subset B(a, \alpha) \subset U_{\iota_0},$$

und dies ist ein Widerspruch zur Konstruktion von B_N !

iii) \implies iv) Es sei ε_k eine (monoton fallende) Nullfolge, und $\{x_0^{(k)}, \dots, x_{N_k}^{(k)}\} =: M_k$ sei die zu ε_k gehörende Punktmenge. Dann ist die Menge $M := \bigcup_{k=0}^{\infty} M_k$ abzählbar, und wegen $X = \bigcup_{x \in M_k} B(x, \varepsilon_k)$ ist M dicht in X .

iv) \implies v) Es sei $M \subset X$ abzählbar und dicht in X , und die $\varepsilon_k > 0$ seien wie oben gewählt. Dann ist das System der offenen Kugeln

$$B(x, \varepsilon_k), \quad x \in M, \quad k \in \mathbb{N},$$

abzählbar. Es sei nun U offen und $x' \in U$. Es genügt dann zu zeigen, daß es ein $B(x, \varepsilon_k)$ gibt mit $x' \in B(x, \varepsilon_k) \subset U$. Nun ist auf jeden Fall $B(x', \varepsilon) \subset U$. Wähle dann ein k mit $\varepsilon_k \leq \varepsilon/2$ und ein $x \in M$ mit $d_X(x, x') < \varepsilon_k$. Bei dieser Wahl ist

$$x' \in B(x, \varepsilon_k) \subset B(x', \varepsilon) \subset U. \quad \square$$

Anhang: Weitere Charakterisierungen der reellen Zahlen

Wir wollen noch zeigen, daß alle wichtigen Sätze des vorliegenden Paragraphen äquivalent zu den uns schon bekannten 44 Axiomen der reellen Zahlen sind. Wir sind damit in der Lage, weitere 6 Charakterisierungen beizusteuern.

(XLV)

Stetige Funktionen auf $[a, b]$ sind durch eine natürliche Zahl n nach oben beschränkt
--

(XLVI)

\mathbb{K} erfüllt (*) & Stetige Funktionen auf $[a, b]$ sind beschränkt

(XLVII)

$[0, 1]$ ist kompakt

(XLVIII)

\mathbb{K} ist archimedisch & Lebesguesches Lemma

(XLIX)

\mathbb{K} ist archimedisch & Stetige Funktionen auf $[a, b]$ sind gleichmäßig stetig
--

(L)

Stetige Funktionen auf $[a, b]$ können durch Treppenfunktionen beliebig genau approximiert werden

Beweis. Wir haben gesehen, daß das Intervallschachtelungs-Axiom (IV) die Aussagen (XLV) und (XLVI) impliziert.

Um von (XLVI) nach (XLVII) zu gelangen, brauchen wir nur einzusehen, daß lokalkompakte angeordnete Körper archimedisch sein müssen. Oder anders gewendet: Ist der angeordnete Körper \mathbb{K} nicht archimedisch, so gibt es nicht kompakte Intervalle $[0, K]$. In diesem Fall gibt es „unendlich kleine“

positive Elemente in \mathbb{K} , also Elemente $\varepsilon > 0$ mit $\varepsilon < 1/n$ für alle $n \in \mathbb{N}^*$, und damit auch „unendlich große“ Elemente K , d. h. $K > n$ für alle $n \in \mathbb{N}$. Die Mengen

$$U_m := \{x \in \mathbb{K} : x = m \text{ oder } |x - m| \text{ unendlich klein}\}, \quad m \in \mathbb{N}^*,$$

sind offen (und paarweise disjunkt), da mit ε und ε' auch $\varepsilon + \varepsilon'$ unendlich klein ist: Ist nämlich $x \in U_m$ und $|x - x'| < \varepsilon$ für ein unendlich kleines positives ε , so ist auch $x' \in U_m$. Wir behaupten, daß die (notwendig offene) Vereinigung der U_m aber auch *abgeschlossen* ist, also die Offenheit des Komplementes

$$U_0 := \mathbb{K} \setminus \bigcup_{m \in \mathbb{N}^*} U_m.$$

Ist nämlich $x \in U_0$ und $\varepsilon > 0$ unendlich klein, so sieht man wie oben, daß das gesamte Intervall $(x - \varepsilon, x + \varepsilon)$ in U_0 enthalten sein muß.

Für jedes unendlich große $K > 0$ ist somit $\bigcup_{m \in \mathbb{N}} U_m$ eine offene Überdeckung von $[0, K]$, aus der man kein einziges Element fortlassen darf.

Genauso sieht man, daß auch (XLVII) nur in archimedisch angeordneten Körpern richtig sein kann. Ist nämlich \mathbb{K} nicht archimedisch, so definiert man zu den eben konstruierten Mengen eine Funktion f durch

$$f(x) := \begin{cases} m, & x \in U_m, \\ 0, & \text{sonst.} \end{cases}$$

Dann ist f stetig, nimmt aber auf dem Intervall $[0, K]$ kein Maximum an. Somit haben wir auch die Implikation (XLVII) \implies (XLVIII) gezeigt.

Für den Schluss von (XLVIII) nach (XLIX) genügt die (sogar tiefere, aber nicht schwere) Einsicht, daß aus der Voraussetzung das Archimedische Axiom folgt: Für positives $a \in \mathbb{K}$ ist die Funktion $f(x) = ax$ stetig auf dem Intervall $[0, 1]$; folglich ist $a = f(1) \leq n$ mit geeignetem $n \in \mathbb{N}$.

Um von (XLIX) nach (L) zu kommen, braucht man sich nur zu vergewissern, daß die Approximationsaussage das Archimedische Axiom nach sich zieht. Nach Voraussetzung gibt es insbesondere für die stetige Funktion $f(x) = x$ auf $[0, 1]$ und zu unendlich kleinem $\varepsilon > 0$ eine Unterteilung $a_0 = 0 < a_1 < \dots < a_n = 1$ und Elemente $c_j \in \mathbb{K}$ mit

$$|x - c_j| < \varepsilon, \quad a_{j-1} \leq x \leq a_j.$$

Hieraus folgt sofort, daß alle a_j und c_j unendlich klein sein müßten, was wegen $a_n = 1$ nicht der Fall sein kann.

Es bleibt damit nur noch zu zeigen, daß man von (L) zu einem der früheren Axiome, z. B. zu dem Prinzip der monotonen Konvergenz (II), zurückschließen kann. Wir nehmen dazu an, daß der Körper \mathbb{K} die Bedingung (*) erfüllt, nicht aber (II), und konstruieren dann eine stetige Funktion f auf einem Intervall $[0, K]$, die nicht beschränkt ist. Da (II) nicht gilt, gibt es eine (ohne Einschränkung streng) monoton wachsende Folge $a_0 < a_1 < \dots$, die nach oben beschränkt ist, aber nicht konvergiert. Wir definieren

$$I_0 = \{x \in \mathbb{K} : x \leq a_0\}, \quad I_j = \{x \in \mathbb{K} : a_{j-1} \leq x \leq a_j\}, \quad j \geq 1,$$

und $A = \bigcup_{j=0}^{\infty} I_j$. Offensichtlich ist A auch Vereinigung der offenen Intervalle $(-\infty, a_{j+1})$ und damit selbst eine *offene* Menge. Wegen des nicht existierenden Limes der Folge $(a_j)_{j \in \mathbb{N}}$ ist aber leicht einzusehen, daß die Menge A in \mathbb{K} auch *folgenabgeschlossen* ist.

Es sei nämlich (b_k) eine Folge in A mit Grenzwert $b \in \mathbb{K}$. Dann gibt es zwei Möglichkeiten: Entweder gibt es ein j , so daß $b_k \leq a_j$ für fast alle k . Dann ist auch $b \leq a_j$ und damit $b \in A$. Oder es gibt zu jedem j unendlich viele k mit $b_k > a_j$. Hieraus folgt dann $b \geq a_j$ für alle j . Wähle nun zu vorgegebenem $\varepsilon > 0$ ein k , so daß $b_k > b - \varepsilon$, und ein j_0 mit $b_k \leq a_{j_0}$. Dann hat man für alle $j \geq j_0$:

$$0 \leq b - a_j \leq b - a_{j_0} \leq b - b_k < \varepsilon$$

im Gegensatz zu der Voraussetzung, daß die Folge (a_j) nicht konvergiert.

Wegen der Bedingung (*) ist die Menge A sogar *abgeschlossen* und damit

$$\mathbb{K} \setminus A = \{x \in \mathbb{K} : x > a_j \text{ für alle } j\}$$

offen (und nach Voraussetzung nicht leer). Es sei weiter eine (ebenfalls nach (*) existierende) Folge $c_j \in \mathbb{K}$ mit $c_0 = 0 < c_1 < c_2 < \dots$ und $\lim_{j \rightarrow \infty} c_j = \infty$ gegeben. Man definiere nun $f : \mathbb{K} \rightarrow \mathbb{K}$ durch

$$f(x) := \begin{cases} 0, & x \in I_0 \\ c_{j-1} + \frac{x - a_{j-1}}{a_j - a_{j-1}} (c_j - c_{j-1}), & x \in I_j, \quad j \geq 1 \\ 0, & x \notin A. \end{cases}$$

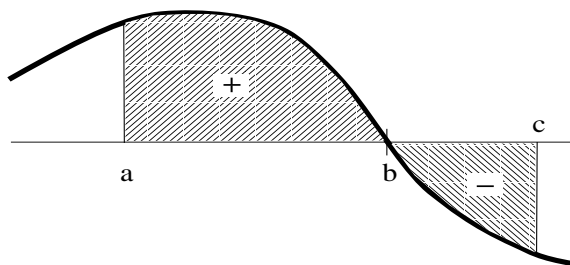
Offensichtlich ist f stetig, aber unbeschränkt auf jedem Intervall $[0, K]$ mit $K \in \mathbb{K} \setminus A$. □

Bemerkung. Man findet das Axiom (XLV) z. B. bei Steiner [163]. Für die anderen Axiome (XLVI), (XLVII), (XLVIII), (XLIX) und (L) sind mir keine Literaturzitate bekannt außer meiner eigenen Arbeit [164].

16 Der Hauptsatz der Differential- und Integralrechnung

Wir formalisieren zuerst die Eigenschaften, die wir gern von einem Integral erfüllt wissen möchten, und zeigen erst später, daß ein solches Integral tatsächlich für eine große Klasse von Funktionen (die die *Treppenfunktionen* und *stetigen* Funktionen umfaßt), existiert und selbstverständlich eindeutig bestimmt ist. Alle Erweiterungen des Integralbegriffs (*Riemann- und Darboux-Integral*, *Lebesgues-Integral*) dehnen die Klasse der integrierbaren Funktionen weiter aus. Dabei ist die Riemannsches Erweiterung nicht weitreichend genug, insbesondere für Anwendungen der Integrationstheorie bei Fourier-Transformationen. Deshalb werden wir uns im anschließenden Kapitel nicht allzulänglich mit diesem Begriff auseinandersetzen und eine kurze Einführung in die Lebesguesche Integrationstheorie geben. Den vollen Ausbau der Integrationstheorie bringen wir erst im Zusammenhang mit der Behandlung des Integrals in mehreren Dimensionen (Teil IV).

Es sei I ein beliebiges Intervall in \mathbb{R} und $f : I \rightarrow \mathbb{R}$ eine zunächst völlig willkürliche (evtl. beschränkte) Funktion. Wir wollen f einen *orientierten* (mit Vorzeichen versehenen) *Flächeninhalt* der Fläche zwischen dem Graphen der Funktion und der reellen Achse für alle x zwischen zwei Grenzen $a, b \in I$, $a < b$, zuordnen.



Figur 16.1

Ein *Integral* von f ist dann eine Abbildung, die jedem Paar $a, b \in I$, $a < b$, eine reelle Zahl zuordnet, die wir mit

$$\mathcal{I}_a^b(f)$$

bezeichnen wollen, und die die folgenden Eigenschaften besitzen soll:

1. $\mathcal{I}_a^b(f) + \mathcal{I}_b^c(f) = \mathcal{I}_a^c(f)$,
2. $m \leq f(x) \leq M$, $a \leq x \leq b \implies m(b-a) \leq \mathcal{I}_a^b(f) \leq M(b-a)$.

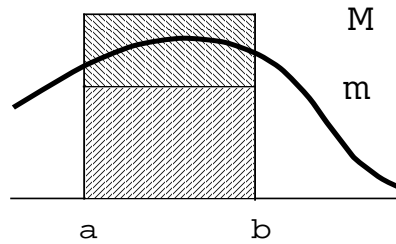
Will man sich bei der Bedingung 1. („Additivität“ des Integrals) von der lästigen Voraussetzung $a < b$ befreien, so hat man selbstverständlich

$$\mathcal{I}_a^a(f) := 0, \quad \mathcal{I}_b^a(f) := -\mathcal{I}_a^b(f)$$

zu setzen. Hat f dann ein einheitliches Vorzeichen auf $[a, b]$, so folgt aus 2., daß das Vorzeichen von $\mathcal{I}_a^b(f)$ positiv oder negativ ist, je nachdem, ob man das in Rede stehende Flächenstück links oder rechts liegend umrundet, wenn man zunächst auf der x -Achse von a in Richtung b startet (und die x - und y -Achse in der üblichen Weise orientiert sind). Die Aussage 2. kann man in dieser allgemeineren Situation noch abschwächen zu

$$2'. \text{ Aus } |f(x)| \leq M \text{ für } x \text{ zwischen } a \text{ und } b \text{ folgt } |\mathcal{I}_a^b(f)| \leq M|b-a|.$$

Diese reicht für viele Anwendungen schon aus.



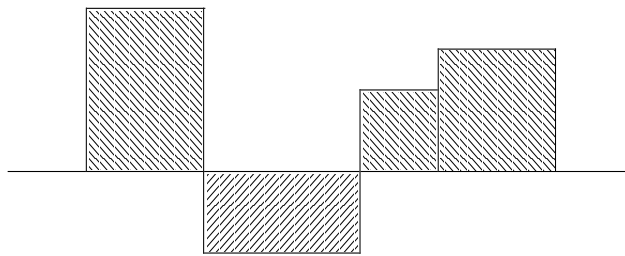
Figur 16.2

Es ist klar, daß diese Eigenschaften das Integral schon für *konstante* Funktionen festlegt: Für $f(x) = \alpha$ kann man $m = M = \alpha$ wählen und erhält notwendigerweise

$$\mathcal{I}_a^b(\alpha) = \alpha(b - a),$$

in Übereinstimmung mit unserer Vorstellung eines „orientierten“ Flächeninhaltes. Aufgrund der Eigenschaft 2. sieht man außerdem, daß auch das Integral für *Treppenfunktionen* φ schon festgelegt ist. Hierbei heißt φ auf einem *kompakten* Intervall $[a, b]$ eine Treppenfunktion, falls eine sogenannte *Zerlegung* $\mathcal{Z} : a = x_0 < x_1 < \dots < x_n = b$ existiert, s. d.

$$\varphi|_{(x_{j-1}, x_j)} = \text{const.}$$



Figur 16.3

Eine Funktion φ auf einem beliebigen Intervall I heißt eine Treppenfunktion, wenn es ein kompaktes Intervall $[a, b] \subset I$ gibt, so daß $\varphi|_{[a, b]}$ eine Treppenfunktion im vorigen Sinne und $\varphi|_{I \setminus [a, b]}$ identisch Null ist. Wir nennen dann auch \mathcal{Z} eine *Zerlegung zur Treppenfunktion* φ . Selbstverständlich sind solche Zerlegungen einer Treppenfunktion a priori nicht eindeutig zugeordnet; man kann z. B. eine Zerlegung weiter „unterteilen“, was bei zahlreichen Überlegungen auch erforderlich ist. Wir nennen deshalb eine Zerlegung $\mathcal{Z}' : a' = x'_0 < \dots < x'_{n'} = b'$ *feiner* als \mathcal{Z} , in Zeichen: $\mathcal{Z}' < \mathcal{Z}$, falls es eine Teilfolge $0 \leq i_0 < i_1 < \dots < i_{n-1} < i_n \leq n'$ gibt mit $x'_{i_j} = x_j$, $j = 0, \dots, n$. Insbesondere ist dann notwendig $[a, b] \subset [a', b']$.

Unmittelbar klar ist, daß es zu je zwei Zerlegungen $\mathcal{Z}_1, \mathcal{Z}_2$ eine *gemeinsame* Verfeinerung \mathcal{Z}' gibt: $\mathcal{Z}' < \mathcal{Z}_1, \mathcal{Z}_2$. (Man braucht nur die Vereinigung der beiden zugehörigen Punktfolgen zu bilden und deren Punkte linear anzuordnen). Sind nun Treppenfunktionen φ_j gegeben bzgl. \mathcal{Z}_j , $j = 1, 2$, so sind beide auch definiert zu einer gemeinsamen Verfeinerung \mathcal{Z}' , und bezüglich der letzteren ist auch $\varphi_1 + \lambda\varphi_2$, $\lambda \in \mathbb{R}$, eine Treppenfunktion. - Zusammenfassend ergibt dies einen Beweis für das folgende Lemma.

Lemma 16.1 Die Menge $\mathcal{T}(I)$ der Treppenfunktionen auf einem beliebigen Intervall I bildet bzgl. der üblichen Addition und Multiplikation mit reellen Konstanten einen (falls I nicht nur aus einem Punkt besteht, unendlich-dimensionalen) \mathbb{R} -Untervektorraum des Vektorraums aller Funktionen auf I .

Bemerkung. Man kann $\mathcal{T}(I)$ auch auf ganz andere Weise beschreiben. Wir erinnern daran, daß wir mit χ_A die *charakteristische Funktion* einer Teilmenge $A \subset X$ bezeichnen wollen: $\chi_A(x) = 1$, wenn $x \in A$, und $\chi_A(x) = 0$ für $x \in X \setminus A$. - Wir überlassen dem Leser den leichten Nachweis des folgenden Lemmas.

Lemma 16.2 $\mathcal{T}(I)$ ist das lineare Erzeugnis der Menge der charakteristischen Funktionen

$$\chi_{[a,b]}, \quad a, b \in I, \quad a \leq b,$$

im Vektorraum aller Funktionen auf I .

Das Integral einer Treppenfunktion läßt sich ohne weiteres Nachdenken elementar geometrisch einführen (siehe Figur 3). Da die Einschränkung einer Treppenfunktion auf ein kompaktes Intervall selbst wieder eine Treppenfunktion ist, können wir dabei von vornherein annehmen, daß $I = [a, b]$ und φ bzgl. der Zerlegung \mathcal{Z} mit $x_0 = a, \dots, x_n = b$ gegeben ist. Wir setzen dann $I_j := (x_{j-1}, x_j)$, $c_j := \varphi(x)$, $x \in I_j$, und

$$\mathcal{I}_a^b(\varphi) = \sum_{j=1}^n c_j (x_j - x_{j-1}).$$

Lemma 16.3 $\mathcal{I}_a^b(\varphi)$ ist unabhängig von der gewählten definierenden Zerlegung \mathcal{Z} . Hierdurch wird ein Integral auf dem Raum aller Treppenfunktionen erklärt.

Beweis. Es ist nur zu zeigen, daß die Definition sich nicht bei Verfeinerung der gegebenen Zerlegung verändert. Mit Bezeichnungen, die sich von selbst verstehen, ist nämlich

$$\begin{aligned} \sum_{k=1}^{n'} c'_k (x'_k - x'_{k-1}) &= \sum_{j=1}^n \sum_{k=i_{j-1}+1}^{i_j} c'_k (x'_k - x'_{k-1}) = \sum_{j=1}^n c_j \sum_{k=i_{j-1}+1}^{i_j} (x'_k - x'_{k-1}) \\ &= \sum_{j=1}^n c_j (x'_{i_j} - x'_{i_{j-1}}) = \sum_{j=1}^n c_j (x_j - x_{j-1}). \end{aligned}$$

Der Nachweis der axiomatischen Integral-Eigenschaften ist mit dieser notwendigen Vorbemerkung eine leichte Übung. \square

Die obige Definition ist in der Tat das *einzige mögliche* Integral auf der Menge der Treppenfunktion. Dies folgt unmittelbar aus dem folgenden Satz zusammen mit dem Ergebnis über das Integral konstanter Funktionen. Er begründet zudem, warum das Integral einer Treppenfunktion φ überhaupt nicht von ihren Werten an den Stellen a, b und $\overline{I_j} \cap \overline{I_{j+1}}$ abhängt.

Satz 16.4 Ist $f : I \rightarrow \mathbb{R}$ eine beschränkte Funktion und $\mathcal{I}_a^b(f)$ ein Integral, so ist die Funktion

$$F(x) = \mathcal{I}_a^x(f), \quad x \in I,$$

Lipschitz-stetig, also insbesondere stetig.

Beweis. Es sei $|f(x)| \leq K$, und $x_0, x_1 \in I$ seien beliebig vorgegeben. Dann gilt

$$|F(x_0) - F(x_1)| = |\mathcal{I}_a^{x_0}(f) - \mathcal{I}_a^{x_1}(f)| = |\mathcal{I}_{x_1}^{x_0}(f)| \leq K |x_0 - x_1|. \quad \square$$

Folgerung 16.5 Es gilt

$$\lim_{\delta \rightarrow 0} \mathcal{I}_b^{c-\delta}(f) = \mathcal{I}_b^c(f)$$

für beschränkte Funktionen f .

Bemerkung. In dem vorstehenden Satz benötigt man lediglich eine etwas *schwächere* Voraussetzung. Es genügt vorauszusetzen, daß die Funktion f nur *lokal beschränkt* ist: Zu jedem Punkt $x_0 \in I$ gibt es eine Konstante $K > 0$ und eine $\delta > 0$, so daß $|f(x)| \leq K$ für alle $x \in I \cap (x_0 - \delta, x_0 + \delta)$. Dies ist gleichbedeutend damit, daß f auf jedem kompakten Intervall $[a, b] \subset I$ beschränkt ist. Dann ist das *unbestimmte Integral* $x \mapsto \mathcal{I}_a^x(f)$ in Abhängigkeit von x eine stetige Funktion.

Bemerkung und Definition. Wir werden im folgenden weiterhin \mathcal{I}_a^b für ein *abstraktes* Integral schreiben, also für eine irgendwie gegebene Zuordnung, die nur den Axiomen 1. und 2. genügt. Wenn wir jedoch schon eingesehen haben, daß es zu einer Funktion f bzw. einer ganzen Funktionenklasse nur ein solches Integral gibt, so benutzen wir auch das auf LEIBNIZ zurückgehende *Integral-Symbol* eines stilisierten Summenzeichens. (Zur Berechtigung dieses Zeichens siehe mehr im Zusammenhang mit unseren Ausführungen zu Riemannschen Summen und zum Riemann-Integral im folgenden Kapitel). Insbesondere bedeutet also

$$\int_a^b \varphi(x) dx := \sum_{j=1}^n c_j (x_j - x_{j-1})$$

für eine Treppenfunktion φ zur Zerlegung \mathcal{Z} .

Das Integral auf dem *Raum der Treppenfunktionen* erfüllt weitere wichtige *Linearitäts-Eigenschaften*, die wir wie folgt zusammenfassen.

Satz 16.6 *Das Integral $\int : \mathcal{T}(I) \rightarrow \mathbb{R}$ ist ein lineares monotonen Funktional, d. h.*

- i) $\int_a^b (\lambda \varphi(x) + \mu \psi(x)) dx = \lambda \int_a^b \varphi(x) dx + \mu \int_a^b \psi(x) dx, \quad \lambda, \mu \in \mathbb{R};$
 ii) *aus $\varphi \leq \psi$ auf $[a, b]$ folgt $\int_a^b \varphi(x) dx \leq \int_a^b \psi(x) dx$.*

Beweis. Der Beweis des ersten Teils ist „straightforward“. Der zweite folgt unmittelbar aus der Linearität und Axiom 2: Offensichtlich ist ii) gleichbedeutend damit, daß mit $\varphi \geq 0$ auf $[a, b]$ auch $\int_a^b \varphi(x) dx \geq 0$ ist. □

Definition und Bemerkung. Wir werden im folgenden ein Integral mit der Eigenschaft i) ein *lineares Integral* nennen. Die Monotonie ist dann eine automatische Konsequenz, wie aus dem vorstehenden Beweis hervorgeht. Man beachte aber, daß in dem üblichen Aufbau der Integrationstheorie i. A. abstrakte lineare Funktionale auf Funktionenräumen untersucht werden, die nicht notwendig unseren Integral-Axiomen 1. und 2. unterworfen sind. In diesem Fall muß man dann die Monotonie des Funktional gesondert fordern bzw. gegebenenfalls nachweisen.

An *Stetigkeitsstellen* von f besitzt das *unbestimmte Integral* $x \mapsto \mathcal{I}_a^x(f)$ weit stärkere Eigenschaften als die Folgerung 5 besagt.

Satz 16.7 *Es sei \mathcal{I}_a^b ein Integral von f . Dann ist die Funktion*

$$F(x) = \mathcal{I}_a^x(f), \quad x \in I, \quad a \in I \text{ fest},$$

differenzierbar an Stellen x , in denen f stetig ist.

Beweis. Wir betrachten ohne Einschränkung nur *positive* h in dem Differenzenquotienten

$$\frac{F(x_0 + h) - F(x_0)}{h} = \frac{1}{h} \mathcal{I}_{x_0}^{x_0+h}(f).$$

Da f in x_0 stetig ist, gibt es zu jedem $\varepsilon > 0$ ein $\delta > 0$, so daß $h < \delta$ die Abschätzung

$$f(x_0) - \varepsilon \leq f(x_0 + h) \leq f(x_0) + \varepsilon$$

impliziert. Es folgt

$$\left| \frac{F(x_0 + h) - F(x_0)}{h} - f(x_0) \right| = \frac{1}{h} \left| \mathcal{I}_{x_0}^{x_0+h}(f) - f(x_0)h \right| \leq \varepsilon.$$

Für $h < 0$ argumentiert man genauso. □

Dieser Satz ist der wahre Grund für das Bestehen des *Hauptsatzes der Differential- und Integralrechnung*. Man nennt $F : I \rightarrow \mathbb{R}$ eine *Stammfunktion* der Funktion f , wenn F differenzierbar mit $F' = f$ ist. Zwei Stammfunktionen derselben Funktion unterscheiden sich auf einem Intervall natürlich höchstens um eine additive Konstante (Satz 14.16 vi)), und mit F ist auch $F + \text{const.}$ eine Stammfunktion von f .

Eine mögliche Formulierung eines Teils des Hauptsatzes ist tatsächlich die aus dem vorstehenden Satz resultierende

Folgerung 16.8 *Ist $f : I \rightarrow \mathbb{R}$ stetig, so gibt es höchstens einen Integralbegriff; denn dann ist für fest gewähltes, aber beliebiges $a \in I$ notwendig*

$$F(x) = \mathcal{I}_a^x(f)$$

die eindeutig bestimmte Stammfunktion F von f mit $F(a) = 0$.

Beweis. Da Stammfunktionen bis auf additive Konstanten eindeutig bestimmt sind, ergibt sich die Behauptung sofort aus $\mathcal{I}_a^a(f) = 0$. □

Satz 16.9 *Es gibt genau ein Integral auf dem Vektorraum*

$$S(I) = \{ f : I \rightarrow \mathbb{R}, f \text{ stetig, } f \text{ besitzt eine Stammfunktion} \}.$$

Dieses ist linear und monoton. Es ist gegeben durch

$$\mathcal{I}_a^b(f) = F \Big|_a^b := F(b) - F(a),$$

F eine Stammfunktion von f .

Bemerkungen. 1. Wir werden anschließend zeigen, daß die Bedingung des Vorhandenseins einer Stammfunktion überflüssig, also tatsächlich

$$S(I) = \mathcal{C}(I)$$

ist. Somit bleiben alle Aussagen sogar für *stetige* Funktionen gültig. Dies ist dann eine *Existenzaussage* und als solche der zweite und wichtigere Teil des Hauptsatzes der Differential- und Integralrechnung.

2. Unsere Tabelle für (stetige) Ableitungen liefert nun sofort eine Riesenauswahl von Integralen, z. B. ist

$$\int_a^b \cos x \, dx = \sin x \Big|_a^b = \sin b - \sin a$$

$$\int_a^b e^x \, dx = e^x \Big|_a^b = e^b - e^a,$$

u. v. a. m. (Siehe auch den Anhang zu dem vorliegenden Kapitel).

Beweis des Satzes. Wegen Folgerung 8 muß $\mathcal{I}_a^b(f)$ von der angegebenen Gestalt sein. Da sich Stammfunktionen nur um additive Konstanten unterscheiden, ist

$$\mathcal{I}_a^b(f) = F(b) - F(a)$$

unabhängig von F definiert. Die Bedingung 2 ist mit dieser Definition automatisch erfüllt. Die Linearität folgt sofort aus der Tatsache, daß mit $F'_1 = f_1$ und $F'_2 = f_2$ auch $(F_1 + \lambda F_2)' = f_1 + \lambda f_2$ ist. Wegen der Linearität ist $\int f dx \leq \int g dx$ für $f \leq g$ gleichbedeutend mit $\int f dx \geq 0$ für $f \geq 0$. Wegen $F' \geq f \geq 0$ muß in diesem Fall aber F monoton wachsen, woraus die Behauptung folgt. Die 2. Bedingung folgt aus dem Mittelwertsatz:

$$\frac{F(d) - F(c)}{d - c} = F'(\xi) = f(\xi) \quad \text{und} \quad m \leq f(\xi) \leq M, \quad c \leq \xi \leq d. \quad \square$$

Die Hauptidee der Integrationstheorie besteht darin, Grenzfunktionen von geeigneten konvergenten Folgen „elementar integrierbarer“ Funktionen zu betrachten und versuchsweise

$$\int \lim_{j \rightarrow \infty} f_j dx := \lim_{j \rightarrow \infty} \int f_j dx$$

zu setzen. Dies geht i. a. nur gut, wenn man den Begriff der Konvergenz nicht zu schwach wählt. Wir wollen zeigen, daß sich jedenfalls bei *gleichmäßiger* Konvergenz *jeder* Integralbegriff vernünftig verhält.

Satz 16.10 *Es sei $\mathcal{I}_a^b(f_j)$ ein lineares Integral für die Funktionen $f_j : I \rightarrow \mathbb{R}$, und es gelte $f_j \implies f$ gleichmäßig auf jedem kompakten Intervall $[a, b] \subset I$. Dann ist*

$$\mathcal{I}_a^b(f) := \lim_{j \rightarrow \infty} \mathcal{I}_a^b(f_j)$$

ein Integral für f , welches überdies von der speziellen Auswahl der Folge (f_j) unabhängig ist. Insbesondere erhält man das ursprüngliche Integral zurück, sofern f schon integrierbar war.

Beweis. a) Zeigen wir zuerst die Eindeutigkeitsaussage. Es gelte $\lim f_j = \lim f_j^* = f$ gleichmäßig auf $[a, b]$. Dann ist

$$|f_j^*(x) - f_j(x)| < \varepsilon / (b - a)$$

für alle $j \geq N(\varepsilon)$ und alle $x \in [a, b]$. Dies impliziert (wegen der Voraussetzung der Linearität des Integrals)

$$|\mathcal{I}_a^b(f_j^*) - \mathcal{I}_a^b(f_j)| = |\mathcal{I}_a^b(f_j^* - f_j)| \leq \varepsilon(b - a)/(b - a) = \varepsilon$$

für alle $j \geq N(\varepsilon)$ und damit $\lim_{j \rightarrow \infty} \mathcal{I}_a^b(f_j^*) = \lim_{j \rightarrow \infty} \mathcal{I}_a^b(f_j)$, falls einer der beiden Grenzwerte existiert.

b) Wie unter a) beweist man die Existenz des Grenzwertes $\lim_{j \rightarrow \infty} \mathcal{I}_a^b(f_j) =: \mathcal{I}_a^b(f)$ mit Hilfe des Cauchy-Kriteriums. Ist f schon integrierbar, so folgt aus Teil a) mit der Folge $f_j^* = f$, daß die rechte Seite mit dem ursprünglichen Integral übereinstimmt.

Es bleibt nur noch zu zeigen, daß $\mathcal{I}_a^b(f)$ die Eigenschaften 1. und 2. eines Integrals besitzt. Dies folgt für 1. aus der Tatsache, daß eine die Funktion f gleichmäßig auf dem Intervall $[a, c]$ approximierende Folge integrierbarer Funktionen auch auf den Teilintervallen $[a, b]$, $[b, c]$ gleichmäßig approximiert. Es ist somit

$$\mathcal{I}_a^c(f) = \lim_{j \rightarrow \infty} \mathcal{I}_a^c(f_j) = \lim_{j \rightarrow \infty} \mathcal{I}_a^b(f_j) + \lim_{j \rightarrow \infty} \mathcal{I}_b^c(f_j) = \mathcal{I}_a^b(f) + \mathcal{I}_b^c(f).$$

2. folgt aus der Tatsache, daß sich aus $m \leq f(x) \leq M$, $x \in [a, b]$, und der gleichmäßigen Konvergenz der Folge f_j auf $[a, b]$ gegen f sofort $m - \varepsilon \leq f_j(x) \leq M + \varepsilon$ für alle $j \geq N(\varepsilon)$ und $x \in [a, b]$ und damit

$$(m - \varepsilon)(b - a) \leq \mathcal{I}_a^b(f_j) \leq (M + \varepsilon)(b - a)$$

ergibt. Daraus folgen die entsprechenden Abschätzungen für den Grenzwert bei $j \rightarrow \infty$, und da diese für alle $\varepsilon > 0$ richtig sein müssen, gilt auch

$$m(b - a) \leq \mathcal{I}_a^b(f) \leq M(b - a). \quad \square$$

Definition. Wir nennen dieses neue Integral das von dem vorgegebenen Integral induzierte (abstrakte) *Regel-Integral* auf dem Raum $\text{Reg-}\mathcal{F}$ der abstrakten *Regelfunktionen* und schreiben dafür manchmal auch $\text{Reg-}\mathcal{I}_a^b(f)$.

Bemerkung. Ist \mathcal{I}_a^b ein lineares Integral auf einem Vektorraum von „integrierbaren“ Funktionen, so bilden auch die gleichmäßigen Limites einen Vektorraum, auf dem das induzierte Regelintegral selbst wieder linear ist. Es handelt sich bei diesem Prozeß also um die Fortsetzung eines linearen *Operators* von einem normierten Vektorraum auf einen gewissen Abschluß. Wir werden diesen Vorgang im nächsten Kapitel noch einmal genauer im Zusammenhang mit dem Riemann-Integral studieren und dann durch einen geeigneten Kunstgriff zur Einführung des Lebesgue-Integrals heranziehen. Man beachte allerdings, daß wir auf diese Weise nicht immer zu einer größeren Klasse von integrierbaren Funktionen gelangen. Zum Beispiel besitzen die *stetigen* Funktionen ein (eindeutig bestimmtes) Integral (siehe unten); bei dem *Regel-Abschluß* verlassen wir aber den Raum der stetigen Funktionen überhaupt nicht.

Die Situation ist hiervon aber gänzlich verschieden bei dem Raum der *Treppenfunktionen*. Aufgrund des vorigen Satzes geben wir die folgende

Definition. Eine Funktion $f : I \rightarrow \mathbb{R}$ heißt eine *Regelfunktion* (schlechthin), wenn sie sich auf jedem kompakten Intervall $[a, b] \subset I$ beliebig genau in der Supremums-Norm durch Treppenfunktionen approximieren läßt: Für alle $\varepsilon > 0$ gibt es eine Treppenfunktion $\varphi : I \rightarrow \mathbb{R}$, so daß

$$\|f - \varphi\|_{[a,b]} < \varepsilon.$$

Das aufgrund des vorigen Satzes existierende lineare (und damit monotone) Integral auf dem Vektorraum der Regelfunktionen heißt das *Regel-Integral* (schlechthin).

Auch wenn es fast wie eine Tautologie anmutet, notieren wir dieses Ergebnis gesondert.

Folgerung 16.11 *Für jede Regelfunktion f existiert das (Regel-) Integral. Ist f auf dem kompakten Intervall $[a, b]$ gleichmäßiger Limes einer Folge φ_j von Treppenfunktionen, so ist*

$$\int_a^b f(x) dx = \lim_{j \rightarrow \infty} \int_a^b \varphi_j(x) dx.$$

Bemerkung. Wir haben in dem vorigen Kapitel nachgewiesen, daß *stetige* Funktionen diese Approximations-Eigenschaft besitzen. Dies wird aus dem anschließend zu beweisenden Charakterisierungssatz für Regelfunktionen erneut folgen. Somit besitzen alle stetigen Funktionen auf kompakten Intervallen das Regel-Integral. Allerdings geben uns die reinen Existenzbeweise keine approximierende Folge von Treppenfunktionen an die Hand, so daß wir i. A. nicht in der Lage sein werden, das Regel-Integral auch auszurechnen. Wir notieren deshalb hier einen weiteren, sehr präziseren Sachverhalt, der die in Rede stehende Eigenschaft für stetige Funktionen nicht nur ein drittes Mal impliziert, sondern uns auch konkrete Möglichkeiten zur Berechnung von Integralen stetiger Funktionen eröffnet.

Satz 16.12 *Es sei $f : [a, b] \rightarrow \mathbb{R}$ eine stetige Funktion. Dann gibt es zu jedem $\varepsilon > 0$ ein $\delta > 0$, so daß sich für alle Zerlegungen $\mathcal{Z} : a = x_0 < x_1 < \dots < x_n = b$ mit Feinheit*

$$\delta(\mathcal{Z}) := \max_{j=1, \dots, n} (x_j - x_{j-1}) \leq \delta$$

jede Treppenfunktion φ der Bauart

$$\varphi(x) = \begin{cases} f(x_j), & x = x_j, \\ f(\xi_j), & x \in (x_{j-1}, x_j), \quad \xi_j \in [x_{j-1}, x_j] \text{ fest,} \end{cases}$$

von der Funktion f in der Maximumnorm auf $[a, b]$ höchstens um ε unterscheidet.

Beweis. Offensichtlich ist die Supremumsnorm der Differenz $f - \varphi$ höchstens gleich der nur von f und der zu φ gehörenden Zerlegung \mathcal{Z} abhängenden Zahl

$$\mu = \max_{j=1, \dots, n} (M_j - m_j),$$

wobei M_j das Maximum von f auf dem Intervall $I_j := [x_j, x_{j-1}]$ bezeichnet und m_j entsprechend das Minimum. Wir zeigen durch Widerspruch, daß $\mu < \varepsilon$ bei vorgegebenem ε , sofern die Feinheit der unterliegenden Zerlegung hinreichend klein ist. Angenommen, dies sei nicht der Fall. Dann gibt es ein $\varepsilon_0 > 0$ und zu jedem $\delta = 1/n$, $n \in \mathbb{N}^*$, eine Zerlegung \mathcal{Z}_n der Feinheit $\leq 1/n$, so daß $\mu_n := \mu(f, \mathcal{Z}_n) \geq \varepsilon_0$. Da f auf dem kompakten Intervall I_j ihr Maximum und Minimum annimmt, existieren folglich Stellen $\xi_n, \xi'_n \in [a, b]$, so daß

$$f(\xi_n) - f(\xi'_n) \geq \varepsilon_0 \quad \text{und} \quad |\xi_n - \xi'_n| \leq 1/n.$$

Durch Übergang zu einer Teilfolge der ξ_n können wir annehmen, daß

$$\xi := \lim_{n \rightarrow \infty} \xi_n \in [a, b]$$

existiert. Wegen der zweiten Bedingung ist dann auch $\lim_{n \rightarrow \infty} \xi'_n = \xi$, und mit der Stetigkeit der Funktion f erhalten wir den offensichtlichen Unsinn

$$\varepsilon_0 \leq \lim_{n \rightarrow \infty} (f(\xi_n) - f(\xi'_n)) = f(\xi) - f(\xi) = 0. \quad \square$$

Bemerkung und Definition. Das Regel-Integral einer stetigen Funktion f wird somit durch Integrale von Treppenfunktionen der obigen Gestalt, also durch Summen der Form

$$S(f, \mathcal{Z}, (\xi_j)) := \sum_{j=1}^n f(\xi_j)(x_j - x_{j-1}),$$

approximiert. Man nennt diese auch *Riemannsche Summen* von f zur Zerlegung \mathcal{Z} und den *Stützstellen* ξ_j . Sie spielen die zentrale Rolle beim Aufbau des Integrals nach RIEMANN. Schreibt man jetzt noch $\Delta x := x_j - x_{j-1}$ und für „sehr kleine“ Differenzen wie Leibniz symbolisch einfach dx , so ist es von der Form der Riemannschen Summen bis zu Leibniz' Schreibweise des Integrals nicht mehr weit:

$$\int f(x) dx.$$

Die vorige Bemerkung kann noch wesentlich genauer gefaßt werden.

Folgerung 16.13 *Es sei f eine stetige Funktion auf dem kompakten Intervall $[a, b]$, und \mathcal{Z}_n sei eine Folge von Zerlegungen mit $\lim_{n \rightarrow \infty} \delta(\mathcal{Z}_n) = 0$. Dann konvergiert jede beliebige Folge*

$$S_n := S(f, \mathcal{Z}_n, (\xi_{n,j}))$$

von Riemannschen Summen gegen das Integral von f .

Beispiele. 1. Man wird sehr oft \mathcal{Z}_n als die *äquidistante* Zerlegung des Intervalls $[a, b]$ in n gleichlange Unterintervalle wählen: $x_j = a + j(b-a)/n$, $j = 0, \dots, n$. (Dies ist allerdings nicht immer die beste Wahl; siehe Beispiel 4). Hier gilt auf jeden Fall $\delta(\mathcal{Z}_n) = (b-a)/n \rightarrow 0$. Wir betrachten als Beispiel die Funktion $f(x) = x^m$, $m \in \mathbb{N}$, auf dem Intervall $[0, a]$, und wählen als Stützstellen zu der äquidistanten Zerlegung \mathcal{Z}_n jeweils den rechten Endpunkt. Dann wird

$$S_n = \sum_{j=1}^n \left(\frac{ja}{n}\right)^m \frac{a}{n} = \frac{a^{m+1}}{n^{m+1}} \sum_{j=1}^n j^m.$$

Nun ist (Übungsaufgabe!) die Summe $\sum_{j=1}^n j^m$ ein Polynom in n vom Grad $m+1$ mit höchstem Koeffizienten $1/(m+1)$ und damit

$$\int_0^a x^m = \lim_{n \rightarrow \infty} \frac{a^{m+1}}{n^{m+1}} \left\{ \frac{n^{m+1}}{m+1} + \text{niedrigere Terme in } n \right\} = \frac{a^{m+1}}{m+1},$$

was nicht anders zu erwarten war und auch von ARCHIMEDES schon auf ähnliche Weise hergeleitet wurde.

2. Man kann auf die gleiche Weise die Integrale von Sinus und Cosinus berechnen. Der Einfachheit halber beschränken wir uns auf ein Intervall der Form $[0, a]$. Für $z \in \mathbb{C}$ definieren wir zunächst die Funktion F durch

$$F(z) := \sum_{k=0}^n \cos kz + i \sum_{k=0}^n \sin kz = \sum_{k=0}^n e^{ikz}.$$

Die rechte Seite ist als Partialsumme der geometrischen Reihe in $q := \exp(iz)$ für $q \neq 1$, also $z \notin 2\pi\mathbb{Z}$, gleich

$$\begin{aligned} \frac{q^{n+1} - 1}{q - 1} &= \frac{q^{(n+1)/2} (q^{(n+1)/2} - q^{-(n+1)/2})}{q^{1/2} (q^{1/2} - q^{-1/2})} = e^{inz/2} \frac{\sin(n+1)z/2}{\sin z/2} \\ &= \left(\cos \frac{nz}{2} + i \sin \frac{nz}{2} \right) \frac{\sin(n+1)z/2}{\sin z/2}. \end{aligned}$$

Daraus folgen unmittelbar die auch für sich interessanten Identitäten:

$$\begin{aligned} \sum_{k=0}^n \cos kz &= \frac{F(z) + F(-z)}{2} = \cos \frac{nz}{2} \cdot \frac{\sin(n+1)z/2}{\sin z/2}, \\ \sum_{k=0}^n \sin kz &= \frac{F(z) - F(-z)}{2i} = \sin \frac{nz}{2} \cdot \frac{\sin(n+1)z/2}{\sin z/2}. \end{aligned}$$

Zur Berechnung des Integrals über den Cosinus mittels Riemannscher Summen nehmen wir die Folge der äquidistanten Zerlegungen \mathcal{Z}_n des Intervalls $[0, a]$ in n Abschnitte und als Folge der Stützstellen jeweils den linken Endpunkt. Die entsprechende Riemannsche Summe $S_n = S(\cos, \mathcal{Z}_n, (z_j)_{j=0, \dots, n-1})$ ist dann gleich

$$S_n = \left(\sum_{k=0}^{n-1} \cos k \frac{a}{n} \right) \cdot \frac{a}{n} = 2 \cos \frac{(n-1)a}{2n} \cdot \sin a/2 \cdot \frac{a/2n}{\sin a/2n}.$$

Wegen $\lim_{x \rightarrow 0} (\sin x)/x = 1$ konvergiert die rechte Seite gegen $2 \cos a/2 \sin a/2 = \sin a$.

Im Falle des Sinus als Integranden nehme man die gleiche Zerlegungsfolge, aber die rechten Endpunkte als Stützstellen. Hierdurch kommt man zu der Folge

$$S_n = \left(\sum_{k=1}^n \sin k \frac{a}{n} \right) \cdot \frac{a}{n} = 2 \sin \frac{(n+1)a}{2n} \cdot \sin a/2 \cdot \frac{a/2n}{\sin a/2n}.$$

mit dem Grenzwert $2 \sin^2 a/2 = 1 - \cos a$.

3. Für natürliche Zahlen $p > q$ ist

$$\sum_{n=kq}^{kp-1} \frac{1}{n} = \frac{p-q}{k(p-q)} \left(\frac{1}{q} + \frac{1}{q+(1/k)} + \dots + \frac{1}{p-(1/k)} \right)$$

eine Riemannsche Summe der Funktion $f(x) := 1/x$ auf dem Intervall $[q, p]$ bzgl. der äquidistanten Zerlegung mit $k(p-q)$ „Teilintervallen“ mit dem jeweiligen linken Endpunkt als Stützstelle. (Da f

monoton fällt, ist dies sogar eine Riemannsche „Obersumme“). Folglich ist der Grenzwert dieser Summe für $k \rightarrow \infty$ gleich

$$\int_q^p \frac{dx}{x} = \ln x \Big|_{x=q}^p = \ln p - \ln q = \ln \frac{p}{q}.$$

4. Zur Berechnung des Integrals $\int_1^a \frac{dx}{x}$, $a > 1$, unterteilt man das Intervall $[1, a]$ am besten in *geometrischer* Progression: $x_0 = 1 = \alpha_n^0$, $x_1 = \alpha_n$, $x_2 = \alpha_n^2, \dots, x_n = \alpha_n^n = a$ mit $\alpha_n := \sqrt[n]{a} > 1$. Dann ist $x_{j+1} - x_j = \alpha_n^j (\alpha_n - 1) \leq x_n - x_{n-1}$, also die Feinheit der Zerlegung gleich $\alpha_n^{n-1} (\alpha_n - 1)$, und diese geht mit $n \rightarrow \infty$ gegen Null. Wählt man jeweils den linken Endpunkt der Teilintervalle als „Stützstellen“, so findet man die folgenden Riemanschen Summen:

$$S_n = \sum_{j=0}^{n-1} \frac{\alpha_n^j (\alpha_n - 1)}{\alpha_n^j} = n (\sqrt[n]{a} - 1).$$

Nun ist (bekanntlich)

$$\lim_{n \rightarrow \infty} n (\sqrt[n]{a} - 1) = \lim_{n \rightarrow \infty} \frac{a^{1/n} - 1}{1/n} = \lim_{x \rightarrow 0} \frac{a^x - 1}{x} = \frac{d}{dx} e^{x \ln a} \Big|_{x=0} = \ln a.$$

Da die Folge der Feinheiten der Zerlegungen \mathcal{Z}_n gegen Null geht und der Integrand $1/x$ in dem betrachteten Bereich (definiert und) stetig, also insbesondere Riemann-integrierbar ist, ist dieser Grenzwert der Folge (S_n) gleich dem gesuchten Integral:

$$\int_1^a \frac{dx}{x} = \ln a.$$

Bei unseren Untersuchungen zum Regel-Integral erhebt sich aus theoretischer Sicht die Frage, ob man einer Funktion leicht ansehen kann, daß sie eine Regelfunktion ist. Mit anderen Worten: Lassen sich die Regelfunktionen durch einfache *intrinsische* Eigenschaften charakterisieren? Da man jedes Intervall I durch *kompakte* Intervalle ausschöpfen kann, genügt die Charakterisierung auf solchen Intervallen. Es gilt tatsächlich der folgende

Satz 16.14 Eine Funktion $f : I = [a, b] \rightarrow \mathbb{R}$ ist genau dann eine Regelfunktion, wenn sie

- i) für alle $x_0 \in (a, b]$ einen linksseitigen,
- ii) für alle $x_0 \in [a, b)$ einen rechtsseitigen

Grenzwert besitzt.

Schärfer als oben können wir hiermit (erneut) konstatieren:

Folgerung 16.15 Stetige und monotone Funktionen sind Regelfunktionen und besitzen deshalb ein Integral.

Beweis (des Satzes). a) Sei f eine Regelfunktion. Wir müssen z. B. nachweisen, daß der rechtsseitige Limes für alle $x_0 \in [a, b)$ existiert. Sei also zu $\varepsilon > 0$ die Treppenfunktion φ so gewählt, daß

$$\|f - \varphi\|_{[a,b]} < \varepsilon/2.$$

Ferner sei (x_0, t) ein (offenes !) Intervall, auf dem φ konstant sei. Dann gilt für alle $x, x' > x_0$ und nahe bei x_0 wegen $\varphi(x) = \varphi(x')$, daß

$$|f(x) - f(x')| \leq |f(x) - \varphi(x)| + |\varphi(x') - f(x')| < \varepsilon.$$

Mit dem Cauchy-Kriterium folgt dann sofort die Existenz des rechtsseitigen Grenzwertes

$$\lim_{x \searrow x_0} f(x).$$

b) Besitzt umgekehrt f rechts- und linksseitige Grenzwerte, aber keine ε_0 -approximierende Treppenfunktion für ein $\varepsilon_0 > 0$, so konstruiert man durch Intervallhalbierung eine Intervallschachtelung $I = I_0 \supset I_1 \supset I_2 \supset \dots$, so daß $f|_{I_n}$ auch keine ε_0 -approximierende Treppenfunktion besitzt. Ist aber $\xi \in \cap I_n$ und c_ℓ bzw. c_r der links- bzw. rechtsseitige Grenzwert von f in ξ , so gibt es ein $\delta > 0$, so daß

$$\begin{aligned} |f(x) - c_\ell| &< \varepsilon_0 \quad \text{für } x \in [\xi - \delta, \xi), \\ |f(x) - c_r| &< \varepsilon_0 \quad \text{für } x \in (\xi, \xi + \delta]. \end{aligned}$$

Für hinreichend großes n ist aber $I_n = [a_n, b_n] \subset [\xi - \delta, \xi + \delta]$, und die durch

$$\varphi(x) = \begin{cases} c_\ell, & x \in [a_n, \xi) \\ f(\xi), & x = \xi \\ c_r, & x \in (\xi, b_n] \end{cases}$$

definierte Treppenfunktion auf I_n wäre doch eine ε_0 -Approximation von $f|_{I_n}$. Widerspruch! \square

Aus der Definition der Regelfunktionen folgt unter anderem, daß sie wie die zu ihnen gehörenden monotonen Funktionen nur abzählbar viele Unstetigkeitsstellen besitzen können.

Folgerung 16.16 *Jede Regelfunktion (insbesondere jede monotone Funktion) ist mit Ausnahme höchstens abzählbar vieler Stellen stetig.*

Beweis. Die Gesamtheit aller Unstetigkeitsstellen einer approximierenden Folge von Treppenfunktionen ist höchstens abzählbar. Außerhalb dieser Menge sind alle Funktionen stetig und, da sie gleichmäßig konvergieren, ist auch die Grenzfunktion an diesen Stellen stetig. \square

Für das Integral auf dem Vektorraum der *Regelfunktionen* konstatiert man leicht (wir haben weiter oben schon eine viel allgemeinere Aussage formuliert und sie wegen ihrer Einfachheit nicht eines expliziten Beweises für würdig befunden):

Satz 16.17 *Das Integral auf dem Vektorraum der Regelfunktionen ist linear und monoton:*

$$\begin{aligned} \int_a^b (\alpha f(x) + \beta g(x)) dx &= \int_a^b \alpha f(x) dx + \int_a^b \beta g(x) dx, \\ \int_a^b f(x) dx &\leq \int_a^b g(x) dx, \quad \text{falls } f \leq g \text{ auf } [a, b]. \end{aligned}$$

Insbesondere ist dies alles richtig für den Untervektorraum der *stetigen* Funktionen auf I , und in Verbindung mit Folgerung 8 schließen wir unmittelbar:

Folgerung 16.18 *Stetige Funktionen auf (beliebigen) Intervallen in \mathbb{R} besitzen Stammfunktionen.*

Bevor wir uns der reichen Ernte widmen, die uns dieser Satz eröffnet, geben wir noch einen *alternativen Beweis* für die Existenz von Stammfunktionen stetiger Funktionen, der allerdings auf dem WEIERSTRASSSchen *Approximationssatz* beruht und daher an dieser Stelle noch nicht vollständig ist. (Der fehlende Beweis wird am Ende von Kapitel 19 nachgeliefert). Es sei also $f : I \rightarrow \mathbb{R}$ stetig. Wir wählen dann einen festen Punkt $x_0 \in I$ und schreiben I als aufsteigende Vereinigung kompakter Intervalle, die x_0 enthalten. Da es auf jedem solchen Intervall höchstens eine Stammfunktion von f gibt, die in x_0 verschwindet, können wir uns auf die Existenz von Stammfunktionen stetiger Funktionen

auf *kompakten* Intervallen I konzentrieren. Nach dem Weierstraßschen Approximationssatz gibt es nun eine Folge P_n von Polynomen, die auf I gleichmäßig gegen f konvergiert. Die P_n besitzen aber Stammfunktionen Q_n , die nach Satz 14.44 auf I gleichmäßig gegen eine (stetig) differenzierbare Grenzfunktion F konvergieren mit $F' = \lim_n Q'_n = \lim_n P_n = f$. \square

Damit können wir tatsächlich, wie am Anfang behauptet, alle Integrale über stetige Funktionen im Prinzip mittels von Stammfunktionen berechnen. Aus diesem Grunde schreibt man für eine Stammfunktion F von f oftmals auch

$$\int f(x) dx$$

und nennt dieses Symbol ein *unbestimmtes Integral* von f . Der Grund hierfür liegt natürlich in der Tatsache begründet, daß

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

Um die Mehrdeutigkeit von Stammfunktionen anzudeuten, schreibt man manchmal auch

$$\int f(x) dx + C.$$

Es ist üblich, den Zusammenhang zwischen Integralen und Stammfunktionen bei stetigen Funktionen als den *Hauptsatz der Differential- und Integralrechnung* zu bezeichnen. Wir wollen ihn zur Sicherheit noch einmal formulieren.

Satz 16.19 (Hauptsatz der Differential- und Integralrechnung) *Es sei f eine stetige Funktion auf dem kompakten Intervall $I = [a, b]$. Dann existiert die Funktion*

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

und definiert eine Stammfunktion von f . Umgekehrt gilt für jede Stammfunktion F von f , daß

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

Aufgrund des Hauptsatzes können Sätze der Differentialrechnung in solche der Integrationstheorie auf die einfachste Weise „umgemünzt“ werden. Als Beispiele beweisen wir den Satz über die *partielle Integration* und die *Substitutionsregel*. Der erste ist schlicht eine Umformulierung der *Produktregel*.

Satz 16.20 (Partielle Integration) *Sind u, v stetig differenzierbare Funktionen, so gilt*

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

Beweis. Der Hauptsatz, angewandt auf die stetige Funktion $(uv)'$, ergibt

$$uv = \int (uv)' dx = \int (u'v + uv') dx = \int u'v dx + \int uv' dx. \quad \square$$

Beispiele. 1. Setzt man $u(x) = e^x$, $v(x) = x$, so erhält man

$$\int x e^x dx = x e^x - e^x = (x - 1) e^x,$$

was man auch durch Differenzieren sofort verifiziert, und speziell also

$$\int_0^1 x e^x dx = (x - 1) e^x \Big|_0^1 = e^0 = 1.$$

Ganz entsprechend findet man *Rekursionsformeln* für die Integrale

$$\int x^n e^x dx .$$

2. Etwas komplizierter ist die Situation, wenn man $u'(x) = 1$, also $u(x) = x$ und $v(x) = \sqrt{1-x^2}$ im Intervall $(-1, 1)$ setzt. Dann folgt

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

und somit

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

Aus Stetigkeitsgründen gilt diese Formel auch noch in den Randpunkten $x = \pm 1$. Ganz entsprechend leitet man die folgenden Beziehungen ab:

$$\begin{aligned} \int \sqrt{1+x^2} dx &= \frac{1}{2} \left(x \sqrt{1+x^2} + \operatorname{Arsinh} x \right) , \quad x \in \mathbb{R} , \\ \int \sqrt{x^2-1} dx &= \frac{1}{2} \left(x \sqrt{x^2-1} - \operatorname{Arcosh} x \right) , \quad x \in [1, \infty) . \end{aligned}$$

3. Wir leiten eine Rekursionsformel für die Integrale $I_k := \int \sin^k x dx$ ab, mit deren Hilfe wir die *Produkt-Darstellung* von $\pi/2$ durch WALLIS endlich begründen können. Für $k \geq 2$ setzen wir $u' = \sin$ und $v = \sin^{k-1}$, womit wir sofort

$$\begin{aligned} I_k &= -\cos x \sin^{k-1} x + (k-1) \int \cos^2 x \sin^{k-2} x dx \\ &= -\cos x \sin^{k-1} x + (k-1) \int (1 - \sin^2 x) \sin^{k-2} x dx \\ &= -\cos x \sin^{k-1} x + (k-1) I_{k-2} - (k-1) I_k \end{aligned}$$

erhalten. Diese Gleichung kann man nach I_k auflösen:

$$I_k = -\frac{1}{k} \cos x \sin^{k-1} x + \frac{k-1}{k} I_{k-2} .$$

Da $I_0 = \int \sin^0 x dx = x$ und $\int \sin x dx = -\cos x$, kann man damit I_k rekursiv für alle natürlichen Zahlen k berechnen.

Es sei nun A_k das *bestimmte* Integral $\int_0^{\pi/2} \sin^k x dx$. Es ist $A_0 = \pi/2$ und $A_1 = 1$; nach der zuvor hergeleiteten Rekursionsformel gilt dann für alle $k \geq 2$ die Beziehung $k A_k = (k-1) A_{k-2}$, so daß wir schließlich alle Werte der Integrale I_k angeben können:

$$A_{2k} = \frac{(2k-1)(2k-3) \cdots 3 \cdot 1}{(2k)(2k-2) \cdots 4 \cdot 2} \frac{\pi}{2} , \quad A_{2k+1} = \frac{(2k)(2k-2) \cdots 4 \cdot 2}{(2k+1)(2k-1) \cdots 5 \cdot 3} .$$

Weiter ist $A_{2k+2} \leq A_{2k+1} \leq A_{2k}$ wegen $\sin^{2k+2} x \leq \sin^{2k+1} x \leq \sin^{2k} x$ auf dem Intervall $[0, \pi/2]$. Mit

$$\lim_{k \rightarrow \infty} \frac{A_{2k+2}}{A_{2k}} = \lim_{k \rightarrow \infty} \frac{2k+1}{2k+2} = 1 \quad \text{ist auch} \quad \lim_{k \rightarrow \infty} \frac{A_{2k+1}}{A_{2k}} = 1 ,$$

so daß schließlich aus

$$\frac{A_{2k+1}}{A_{2k}} = \frac{(2k) \cdot (2k) \cdot \dots \cdot 4 \cdot 4 \cdot 2 \cdot 2}{(2k+1) \cdot (2k-1) \cdot \dots \cdot 3 \cdot 1} \frac{2}{\pi}$$

die schon früher erwähnte WALLISSche *Produktformel*

$$\frac{\pi}{2} = \prod_{k=1}^{\infty} \frac{4k^2}{4k^2 - 1}$$

folgt.

Die *Substitutionsformel* ist entsprechend eine Umschreibung der *Kettenregel*.

Satz 16.21 (Substitutionsformel) *Es sei $f : D \rightarrow \mathbb{R}$ eine stetige Funktion mit Stammfunktion F , und die Funktion $\varphi : [a, b] \rightarrow D$ sei stetig differenzierbar. Dann ist $F \circ \varphi$ eine Stammfunktion von $(f \circ \varphi) \varphi'$, und es gilt somit*

$$\int_a^b f(\varphi(x)) \varphi'(x) dx = \int_{\varphi(a)}^{\varphi(b)} f(t) dt.$$

Beweis. Nach der Kettenregel ist $(F \circ \varphi)' = (F' \circ \varphi) \varphi' = (f \circ \varphi) \varphi'$. Dies ist der Inhalt der ersten Behauptung. Mit dem Hauptsatz folgt

$$\begin{aligned} \int_a^b f(\varphi(x)) \varphi'(x) dx &= \int_a^b (F \circ \varphi)'(x) dx = (F \circ \varphi)(b) - (F \circ \varphi)(a) = F(\varphi(b)) - F(\varphi(a)) \\ &= \int_{\varphi(a)}^{\varphi(b)} f(t) dt, \end{aligned}$$

was zu beweisen war. □

Bemerkung. Man „substituiert“ also in dem rechts stehenden Integral $t := \varphi(x)$. Man muß dann aber auch das „Differential“ dt substituieren durch $dt = \varphi'(x) dx$, was man sich formal aus der Leibnizschen Schreibweise

$$\frac{dt}{dx} = \frac{d\varphi(x)}{dx} = \varphi'(x)$$

ableiten kann. Das Substituieren der Integrationsgrenzen muß man dagegen von links nach rechts vornehmen; selbstverständlich ist die entsprechende Regel ebenfalls völlig natürlich, denn wenn x das Intervall von a nach b durchläuft, läuft $t = \varphi(x)$ von $\varphi(a)$ nach $\varphi(b)$.

Beispiele. 1. Ist $f(x) \neq 0$ in dem betrachteten Intervall, so ist

$$\int \frac{f'(x)}{f(x)} dx = \ln |f(x)|,$$

wie man unmittelbar durch Differentiation verifiziert. Man gewinnt dasselbe Resultat auch aus dem Integral über $1/t$ mit der Substitution $t = f(x)$. Als Spezialfall hat man z. B.

$$\int \cot x dx = \int \frac{\cos x}{\sin x} dx = \ln |\sin x|.$$

2. Ganz entsprechend erhält man die Formeln

$$\int f(x) f'(x) dx = \frac{1}{2} f^2(x), \quad \int \frac{f'(x)}{\sqrt{f(x)}} = 2\sqrt{f(x)},$$

wobei im zweiten Fall $f(x) > 0$ vorausgesetzt werden muß.

3. Ist f eine *ungerade* Funktion, also $f(-t) = -f(t)$, so liefert die Substitution $t := -s$, $dt = -ds$, die Beziehung

$$\int_{-a}^0 f(t) dt = - \int_a^0 f(-s) ds = - \int_0^a f(s) ds$$

und damit

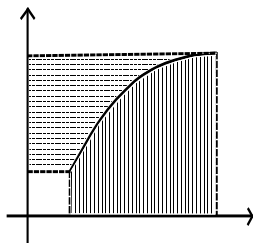
$$\int_{-a}^a f(t) dt = \int_{-a}^0 f(t) dt + \int_0^a f(t) dt = 0$$

für alle $a > 0$.

4. Ist $f : \mathbb{R} \rightarrow \mathbb{R}$ *periodisch* mit Periode p , so folgt mit der Substitution $t := x + p$ sofort, daß das Integral von f über ein Intervall $[0, a]$ mit dem über $[p, a + p]$ übereinstimmt, und dies wiederum impliziert

$$\int_a^{a+p} f(x) dx = \int_0^a f(x) dx + \int_0^p f(x) dx + \int_p^{a+p} f(x) dx = \int_0^p f(x) dx .$$

Es gibt auch eine offensichtliche Verbindung zwischen dem Integral einer Funktion und dem ihrer Umkehrfunktion, die sich in dem folgenden Bild manifestiert:



Figur 16.4

Satz 16.22 Die Funktion $f : [a, b] \rightarrow \mathbb{R}$ sei stetig differenzierbar und streng monoton wachsend, und $g : [f(a), f(b)] \rightarrow \mathbb{R}$ sei ihre Umkehrfunktion. Dann gilt

$$\int_a^b f(x) dx + \int_{f(a)}^{f(b)} g(y) dy = b f(b) - a f(a) .$$

Beweis. Es ist verlockend, vielversprechend und in der Tat erfolgreich, auf das zweite Integral die Substitutionsregel mit $y = f(x)$ anzuwenden. Damit vergewissert man sich der Formel auf einen Schlag:

$$\begin{aligned} \int_a^b f(x) dx + \int_{f(a)}^{f(b)} g(y) dy &= \int_a^b (f(x) + x f'(x)) dx = \int_a^b (x f(x))' dx \\ &= x f(x) \Big|_a^b = b f(b) - a f(a) , \end{aligned}$$

was zu beweisen war. □

Beispiele. Anwendungen der Formel sind „straightforward“. Es ist z. B. mit $f(x) = \sqrt{x}$, $g(y) = y^2$:

$$\int_a^b \sqrt{x} dx + \int_{\sqrt{a}}^{\sqrt{b}} y^2 dy = b \sqrt{b} - a \sqrt{a}$$

und

$$\int_{\sqrt{a}}^{\sqrt{b}} y^2 dy = \frac{1}{3} y^3 \Big|_{\sqrt{a}}^{\sqrt{b}} ,$$

und damit insgesamt

$$\int_a^b \sqrt{x} dx = \frac{2}{3} (b\sqrt{b} - a\sqrt{a}) \quad \text{oder} \quad \int_a^b x^{1/2} dx = \frac{b^{3/2} - a^{3/2}}{3/2}.$$

Ebenso einfach findet man z. B. $\int_0^1 \arcsin x dx = \pi/2 - 1$.

Man interessiert sich insbesondere für *Klassen* von Funktionen, die sich *elementar* integrieren lassen, d. h. die Stammfunktionen besitzen, die Quotienten sind von Linearkombinationen aus Potenzfunktionen, der Exponentialfunktion und den trigonometrischen Funktionen und deren Umkehrfunktionen. Die letzteren pflegt man auch als *elementare transzendente Funktionen* zu bezeichnen. - Aus dem Satz über die *Partialbruch-Zerlegung* (Satz 6.22) folgt unmittelbar

Satz 16.23 *Die rationalen Funktionen sind elementar integrierbar.*

Beweis. Man braucht nur Stammfunktionen für die in dem o. a. Satz auftretenden *Partial-Brüche* anzugeben. Für die Linearfaktoren und ihre Potenzen ist uns das Ergebnis schon wohlbekannt:

$$\int \frac{dx}{x-a} = \ln|x-a|, \quad \int \frac{dx}{(x-a)^m} = -\frac{1}{m-1} \frac{1}{(x-a)^{m-1}}, \quad m \geq 2.$$

Für die quadratischen Anteile $x^2 + bx + c$ ohne reelle Nullstellen, also mit $4c - b^2 > 0$, und ihre Potenzen ersetzt man zunächst den linearen Zähler $Bx + C$ im Falle $B \neq 0$ durch $Bx + (Bb/2) + (C - (Bb/2))$ und gewinnt damit unter Heranziehung der Substitutionsregel die Formeln

$$\int \frac{Bx + C}{(x^2 + bx + c)^m} dx = \left(C - \frac{Bb}{2}\right) \int \frac{dx}{(x^2 + bx + c)^m} + \begin{cases} \frac{B}{2} \ln(x^2 + bx + c), & m = 1, \\ \frac{-B}{2(m-1)(x^2 + bx + c)^{m-1}}, & m \geq 2. \end{cases}$$

Diese Formeln sind auch bei $B = 0$ noch korrekt, auch wenn sie uns dann nicht weiterhelfen. Es verbleiben somit noch die entsprechenden Integrale mit *konstantem* Zähler $C = 1$. Hier verifiziert man schnell für $m \geq 2$ die folgende Rekursionsformel:

$$\int \frac{dx}{(x^2 + bx + c)^m} = \frac{2x + b}{(m-1)(4c - b^2)(x^2 + bx + c)^{m-1}} + \frac{2(2m-3)}{(m-1)(4c - b^2)} \int \frac{dx}{(x^2 + bx + c)^{m-1}},$$

und mit

$$\int \frac{dx}{x^2 + bx + c} = \frac{2}{\sqrt{4c - b^2}} \arctan \frac{2x + b}{\sqrt{4c - b^2}}$$

sind alle Fälle abgehandelt. □

Bemerkung. Wir wollen kurz aufzeigen, wie man die soeben formulierte Rekursionsformel auch durch geschickte partielle Integration herleiten kann. Wir schreiben $Q = Q(x)$ für das quadratische Polynom $x^2 + bx + c$ und bemerken, daß

$$\frac{1}{Q^{m-1}} = \frac{x^2 + bx + c}{Q^m} = \frac{x}{2} \frac{2x + b}{Q^m} + \frac{\frac{b}{2}x + c}{Q^m}.$$

Mit partieller Integration kann man eine Stammfunktion des ersten Ausdrucks auf der rechten Seite schreiben in der Form

$$\frac{-x}{2(m-1)Q^{m-1}}$$

plus eine Stammfunktion von

$$\frac{1}{2(m-1)} \frac{1}{Q^{m-1}},$$

die man auf die linke Seite bringen kann. Es verbleibt auf der rechten Seite noch der zweite Ausdruck, der uns schon zu der gewünschten Rekursionsformel führt, wenn $b = 0$. Ist jedoch $b \neq 0$, so formt man ihn um zu

$$\frac{b}{4} \frac{2x + b}{Q^m} + \frac{1}{4} \frac{4c - b^2}{Q^m},$$

woraus ebenfalls die erwünschte Rekursionsformel folgt.

Beispiele. 1. Das Integral $\int (e^x - 1)/(e^x + 1) dx$ wird durch die Substitution $t := e^x$, $dt = t dx$ übergeführt in das Integral mit dem rationalen Integranden

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

Somit ist

$$\int \frac{e^x - 1}{e^x + 1} dx = (2 \ln(t + 1) - \ln t) \Big|_{t=e^x} = 2 \ln(e^x + 1) - x.$$

2. Entsprechend wird mit der Substitution $t := \sqrt{x}$, $dt = dx/2t$:

$$\begin{aligned} \int \frac{x - \sqrt{x}}{x + \sqrt{x}} dx &= \int \frac{2t(t^2 - t)}{t^2 + t} dt = \int \frac{2t(t - 1)}{t + 1} dt = 2 \int \left(t - 2 + \frac{2}{t + 1} \right) dt \\ &= (t^2 - 4t + 4 \ln(t + 1)) \Big|_{t=\sqrt{x}} = x - 4\sqrt{x} + 4 \ln(\sqrt{x} + 1). \end{aligned}$$

3. Das letzte Beispiel läßt sich zu einer Methode zur Behandlung einer ganzen Klasse von Integralen ausbauen. Wir bezeichnen im Folgenden mit $R = R(s, t)$ eine *rationale Funktion in den zwei reellen Variablen* s und t , also den Quotienten zweier Polynome in s, t . Dann lassen sich Integrale der Form

$$\int R(x, \sqrt[n]{ax + b}) dx, \quad a \neq 0,$$

stets auf ein Integral mit einem rationalen Integranden (in einer Veränderlichen) zurückführen, und zwar durch die offensichtliche Substitution

$$t := \sqrt[n]{ax + b}, \quad x = \frac{t^n - b}{a}, \quad dx = \frac{n t^{n-1} dx}{a}.$$

4. Eine weitere Klasse dieser Art sind Integrale der Form

$$\int R(\sin x, \cos x) dx,$$

die sich stets mit der Substitution

$$t := \tan \frac{x}{2}$$

in ein *rationales Integral* umformen lassen. Es ist nämlich

$$\frac{dt}{dx} = \frac{1/2}{1 + \tan^2(x/2)}, \quad \text{also} \quad dx = \frac{2 dt}{1 + t^2}$$

und wegen $\frac{1}{1 + t^2} = \cos^2(x/2)$ auch

$$\cos x = \cos 2 \frac{x}{2} = \frac{1 - t^2}{1 + t^2}, \quad \sin x = \sin 2 \frac{x}{2} = \frac{2t}{1 + t^2}.$$

Man sollte aber, bevor man die vorgeschlagene Substitution gedankenlos anwendet, erst prüfen, ob nicht eine näherliegende Substitution wie z. B. $t := \tan x$ schneller zum Ziel führt.

Weitere Beispiele findet man in umfangreicheren Lehrbüchern wie z. B. [9] oder [14]. Eine schöne Methode zur Bestimmung von Integralen führt über das Komplexe: Einzelheiten zum *Residuensatz* und seinen Anwendungen gehören zu jedem Standardtext der *Funktionentheorie* (siehe z. B. [82]).

Wir wollen an dieser Stelle des Kapitels noch einige theoretische Anwendungen des Hauptsatzes ergänzend anfügen. Hierzu gehören ein (einfacherer) Beweis von Satz 14.44 über die *Differenzierbarkeit von Grenzfunktionen* in einer spezielleren Situation und die Darstellung des *Restglieds in der Taylor-Formel in Integralform*, die auf dem *Mittelwertsatz der Integralrechnung* beruht.

Wir setzen voraus, daß die Funktionen $f_j : I = [a, b] \rightarrow \mathbb{R}$ nicht nur differenzierbar, sondern sogar *stetig differenzierbar* sind. Ist dann die Folge (f_j) an einer einzigen Stelle $x_0 \in I$ konvergent und die Folge der Ableitungen (f'_j) gleichmäßig konvergent auf I , so ist die Grenzfunktion $f := \lim_j f_j$ nach Satz 14.44 differenzierbar mit

$$f' = \left(\lim_{j \rightarrow \infty} f_j \right)' = \lim_{j \rightarrow \infty} f'_j .$$

Der ganz rechts stehende Limes ist aber unter der zusätzlichen Voraussetzung ebenfalls stetig, so daß die Grenzfunktion f sogar *stetig differenzierbar* ist.

Wir wollen für diese letztgenannte Konsequenz einen einfachen Beweis geben und setzen zu diesem Zwecke

$$f^* := \text{Lim}_{j \rightarrow \infty} f'_j .$$

Nach Voraussetzung ist die Funktion f^* auf I stetig und besitzt daher die Stammfunktion

$$F := f(x_0) + \int_{x_0}^x f^*(t) dt .$$

Insbesondere ist $F(x_0) = f(x_0)$ und $F' = f^* = \lim_j f'_j$. Es braucht daher nur noch gezeigt zu werden, daß die Folge der f_j selbst gleichmäßig auf I gegen F konvergiert. Nach dem Hauptsatz der Differential- und Integralrechnung ist aber

$$f_j(x) := f_j(x_0) + \int_{x_0}^x f'_j(t) dt ,$$

und durch Subtraktion von der obigen Definition von F ergibt sich

$$F(x) - f_j(x) := f(x_0) - f_j(x_0) + \int_{x_0}^x (f^*(t) - f'_j(t)) dt .$$

Hieraus folgt die Behauptung sofort unter Verwendung der Standardabschätzung von Integralen. \square

Satz 16.24 (Integralform des Restglieds) Für eine $n+1$ -mal stetig differenzierbare Funktion $f : I \rightarrow \mathbb{R}$ kann man das Restglied R_{n+1} bei der Taylor-Entwicklung an einer Stelle $a \in I$ schreiben in der Form

$$R_{n+1}(x) = \frac{1}{n!} \int_a^x (x-t)^n f^{(n+1)}(t) dt .$$

Beweis. Wir führen vollständige Induktion nach n . Der Induktionsanfang $n=0$ ist nichts anderes als die schon mehrfach verwendete Version des Hauptsatzes der Differential- und Integralrechnung:

$$f(x) = f(a) + \int_a^x f'(t) dt .$$

Ist die Aussage für eine Zahl n schon bewiesen, die Funktion f aber sogar $n+2$ -mal stetig differen-

zierbar, so erhält man mit partieller Integration sofort

$$\begin{aligned} R_{n+1}(x) &= \frac{1}{n!} \int_a^x (x-t)^n f^{(n+1)}(t) dt = - \int_a^x f^{(n+1)}(t) \frac{d}{dt} \left(\frac{(x-t)^{n+1}}{(n+1)!} \right) \\ &= - f^{(n+1)}(t) \frac{(x-t)^{n+1}}{(n+1)!} \Big|_{t=a}^x + \int_a^x \frac{(x-t)^{n+1}}{(n+1)!} f^{(n+2)}(t) dt \\ &= \frac{f^{(n+1)}(a)}{(n+1)!} (x-a)^{n+1} + \frac{1}{(n+1)!} \int_a^x (x-t)^{n+1} f^{(n+2)}(t) dt, \end{aligned}$$

woraus die Behauptung folgt. \square

Beispiele. Wir berechnen erneut $\ln 2$ und $\pi/2$, diesmal ohne Verwendung des *Abelschen Grenzwertsatzes* (siehe Kapitel 12). Dazu genügt der Nachweis der folgenden Aussage:

Die Reihenentwicklung

$$\ln(1+x) = \sum_{k=0}^{\infty} (-1)^k \frac{x^{k+1}}{k+1}$$

ist gleichmäßig auf jedem Intervall der Form $[-1+\delta, 1]$, $\delta > 0$.

Beweis. Es gilt

$$\frac{1}{1+t} = \sum_{k=0}^{\infty} (-1)^k t^k + (-1)^{n+1} \frac{t^{n+1}}{1+t}$$

und damit

$$\ln(1+x) = \int_0^x \frac{dt}{1+t} = \sum_{k=0}^n (-1)^k \frac{x^{k+1}}{k+1} - (-1)^n \int_0^x \frac{t^{n+1} dt}{1+t}.$$

Kürzt man die Summe auf der rechten Seite mit $s_n(x)$ ab und schreibt man $I := [-1+\delta, 1]$, so folgt

$$\sup_{x \in I} |\ln(1+x) - s_n(x)| \leq \sup_{x \in I} \left| \int_0^x \frac{t^{n+1}}{1+t} dt \right| \leq \frac{1}{\delta} \left| \int_0^x t^{n+1} dt \right| \leq \frac{1}{\delta(n+2)} \xrightarrow{n \rightarrow \infty} 0. \quad \square$$

Mit einem ähnlichen Argument erhält man wegen

$$\sup_{|x| \leq 1} \left| \int_0^x \frac{t^{2(n+2)}}{1+t^2} dt \right| \leq \frac{1}{2n+3} \xrightarrow{n \rightarrow \infty} 0$$

auch die *gleichmäßige Konvergenz* der Reihenentwicklung

$$\arctan x = \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k+1}}{2k+1}$$

auf dem kompakten Intervall $[-1, 1]$.

Für die letzte Anwendung benötigen wir noch den *Mittelwertsatz* der Integralrechnung, der auch per se interessant ist und zu zahlreichen weiteren Anwendungen Anlaß gibt.

Satz 16.25 (Mittelwertsatz der Integralrechnung) *Es seien $f, \varphi : I = [a, b] \rightarrow \mathbb{R}$ stetige Funktionen und $\varphi \geq 0$. Dann existiert ein $\xi \in I$ mit*

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

Im Spezialfall $\varphi \equiv 1$ hat man insbesondere

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

für ein $\xi \in I$.

Beweis. Es sei m das Minimum und M das Maximum von f auf I . Somit gilt $m\varphi \leq f\varphi \leq M\varphi$, und wegen der Monotonie des Integrals ist

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

Also existiert eine Zahl μ mit $m \leq \mu \leq M$, so daß

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

Nach dem Zwischenwertsatz ist aber $\mu = f(\xi)$ für ein $\xi \in I$. □

Bemerkung. Aus der Integralform des Restglieds gewinnt man erneut die LAGRANGESche Version unter Verwendung des Mittelwertsatzes der Integralrechnung (dies ist die oben angekündigte) Anwendung:

$$\begin{aligned} R_{n+1}(x) &= \frac{1}{n!} \int_a^x (x-t)^n f^{(n+1)}(t) dt = f^{(n+1)}(\xi) \int_a^x \frac{(x-t)^n}{n!} dt \\ &= -f^{(n+1)}(\xi) \frac{(x-t)^{n+1}}{(n+1)!} \Big|_{t=a}^x = \frac{f^{(n+1)}(\xi)}{(n+1)!} (x-a)^{n+1} . \end{aligned}$$

Trotz der unglaublichen Fülle von Integralen, die wir mit dem Hauptsatz bestimmen können, entziehen sich schon ganz einfache Funktionen dem „Regelansatz“. So sollte z. B. die Funktion $f(x) := \sin 1/x$ auf jedem Intervall $[0, A]$, unabhängig von dem Funktionswert an der Stelle 0, ein Integral besitzen, da wegen der Beschränktheit von f der Grenzwert

$$\lim_{\varepsilon \searrow 0} \int_{\varepsilon}^A \sin \frac{1}{x} dx$$

existiert (siehe Folgerung 27 weiter unten). Die Funktion f ist aber auf keinen Fall in den Nullpunkt hinein als Regel-Funktion fortsetzbar !

Mögliche Auswege aus diesem Lemma sind die Bereitstellung eines leistungsfähigeren Integrals (im vorigen Beispiel reicht schon das RIEMANN-Integral; siehe nächstes Kapitel) oder, wie wir in unserer Argumentation schon haben anklingen lassen, die Verwendung von sogenannten *uneigentlichen Integralen*. Wir wollen hier einige Anmerkungen zu dem letzteren Konzept im Rahmen einer allgemeinen Integrationstheorie im Sinne unserer Axiome 1 und 2 machen. Wir betrachten also wieder eine Funktion f oder eine Funktionenklasse auf einem Intervall $I \subset \mathbb{R}$ und verlangen die Existenz eines Integrals $\mathcal{I}_a^b(f)$ für jedes kompakte Intervall $[a, b] \subset I$. Beim *uneigentlichen Integral* stellen wir uns die Frage, wie sich diese Integrale bei Annäherung von a an das Infimum von I bzw. von b an das Supremum des Intervalls verhalten. Durch Zerlegung des Intervalls können wir uns auf die Annäherung an eine der beiden Grenzen beschränken; es sei dies im folgenden die rechte und damit ohne Einschränkung $I = [a, B)$ mit $B \in \mathbb{R} \cup \{\infty\}$. Wir schreiben dann

$$\mathcal{I}_a^{B-}(f) := \lim_{b \nearrow B} \mathcal{I}_a^b(f)$$

oder bei Benutzung des üblichen Integralzeichens

$$\int_a^{B-} f(x) dx := \lim_{b \nearrow B} \int_a^b f(x) dx ,$$

sofern der rechts stehende Grenzwert existiert. Dies ist auf jeden Fall richtig, wenn $B \in \mathbb{R}$ und f nahe B lokal beschränkt ist (siehe Folgerung 27). Ist insbesondere $B \in I$ und f in B lokal beschränkt, so existiert der Grenzwert und stimmt mit dem Integral $\mathcal{I}_a^B(f)$ in der alten Bedeutung überein (Folgerung 5).

Definition. Der Grenzwert $\mathcal{I}_a^{B-}(f) := \lim_{b \nearrow B} \mathcal{I}_a^b(f)$, $B \in \mathbb{R} \cup \{\infty\}$, heißt ein *an der oberen Grenze B uneigentliches Integral*. Ist B endlich, so reservieren manche Autoren diese Bezeichnung nur für die Situation, daß f bei Annäherung an B von links *nicht beschränkt* ist (denn im anderen Fall existiert der Grenzwert immer). Entsprechend definiert man an der unteren Grenze und an beiden Grenzen uneigentliche Integrale. Man sagt dann auch, das entsprechende uneigentliche Integral *existiere* oder *konvergiere*. Im anderen Fall spricht man auch von einem *divergierenden uneigentlichen Integral*.

Bemerkung. Da nach der zuvor gemachten Bemerkung keine Verwechslungen zu befürchten sind, schreibt man oft für ein an der oberen Grenze B uneigentliches Integral auch

$$\int_a^B f(x) dx \quad \text{anstelle von} \quad \int_a^{B-} f(x) dx .$$

Diese Konvention ist insbesondere angebracht, wenn $B = \infty$. Entsprechende Notationen werden für die anderen Typen von uneigentlichen Integralen verwendet.

Bemerkung. Die Existenz solcher uneigentlicher Integrale ist offensichtlich unabhängig von der Auswahl des Punktes a . Man beachte, daß man im Falle eines beidseitig uneigentlichen Integrals die beiden Grenzübergänge *unabhängig voneinander* durchführen muß:

$$\mathcal{I}_{A+}^{B-}(f) := \lim_{a \searrow A} \mathcal{I}_a^{x_0}(f) + \lim_{b \nearrow B} \mathcal{I}_{x_0}^b(f) ,$$

was man im Falle $A, B \in \mathbb{R}$ auch schreiben kann in der Form

$$\mathcal{I}_{A+}^{B-}(f) := \lim_{\varepsilon \searrow 0} \mathcal{I}_{A+\varepsilon}^{x_0}(f) + \lim_{\varepsilon \searrow 0} \mathcal{I}_{x_0}^{B-\varepsilon}(f) ,$$

und diese Formulierung impliziert insbesondere die Gleichheit

$$\mathcal{I}_{A+}^{B-}(f) := \lim_{\varepsilon \searrow 0} \mathcal{I}_{A+\varepsilon}^{B-\varepsilon}(f) ,$$

die jedoch im Allgemeinen keinesfalls gleichbedeutend mit der Konvergenz des uneigentlichen Integrals $\mathcal{I}_{A+}^{B-}(f)$ ist.

Beispiele. 1. Da $\int_0^R \sin x dx = 1 - \cos R$ für $R \rightarrow \infty$ nicht konvergiert, existiert keines der beiden uneigentlichen Integrale

$$\int_0^\infty \sin x dx , \quad \int_{-\infty}^\infty \sin x dx .$$

Der sogenannte *Hauptwert*

$$\lim_{R \rightarrow \infty} \int_{-R}^R \sin x dx$$

ist aber gleich Null.

2. Das Integral $\int_0^1 \frac{dx}{x^s}$, $s > 0$, ist uneigentlich bei 0. Es gilt

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

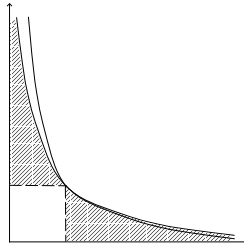
Somit konvergiert das uneigentliche Integral gegen $1/(1-s)$ für $s < 1$, und es divergiert für $s \geq 1$.

3. Wie in 2. rechnet man nach, daß das uneigentliche Integral

$$\int_1^{\infty} \frac{dx}{x^s} = \lim_{R \rightarrow \infty} \int_1^R \frac{dx}{x^s}$$

konvergent für $s > 1$ gegen $1/(s-1)$ und divergent für $0 < s \leq 1$ ist.

4. Insbesondere existiert niemals $\int_0^{\infty} \frac{dx}{x^s}$ für $s > 0$.



Figur 16.5

5. Das Integral $\int_{-1}^1 \frac{dx}{\sqrt{1-x^2}}$ ist an beiden Endpunkten uneigentlich. Es berechnet sich einfach als das Doppelte von

$$\int_0^1 \frac{dx}{\sqrt{1-x^2}} = \lim_{\varepsilon \searrow 0} \int_0^{1-\varepsilon} \frac{dx}{\sqrt{1-x^2}} = \lim_{\varepsilon \searrow 0} \left(\arcsin x \Big|_0^{1-\varepsilon} \right) = \lim_{\varepsilon \searrow 0} \arcsin(1-\varepsilon) = \frac{\pi}{2}.$$

Hiermit kann man auch das *eigentliche* Integral $\int_{-1}^1 \sqrt{1-x^2} dx$ berechnen, also den Flächeninhalt des oberhalb der x -Achse liegenden Teils des Einheitskreises um den Nullpunkt:

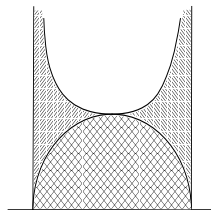
$$\int_{-1}^1 \sqrt{1-x^2} dx = \lim_{a \nearrow 1} \int_{-a}^a \sqrt{1-x^2} dx = \lim_{a \nearrow 1} \left(\frac{a}{2} \sqrt{1-a^2} + \frac{1}{2} \int_{-a}^a \frac{dx}{\sqrt{1-x^2}} \right) = \frac{\pi}{2}.$$

Selbstverständlich kann man das schneller erledigen mit den Substitutionen $x := \sin t$ und $x := \cos t$:

$$2 \int_{-1}^1 \sqrt{1-x^2} dx = \int_{-\pi/2}^{\pi/2} \cos^2 t dt + \int_0^{\pi} \sin^2 t dt.$$

Nun ist die Funktion $\cos^2 t$ periodisch mit der Periode π , so daß sich schließlich, wie nicht anders erwartet, die erwünschte Formel einstellt:

$$\int_{-1}^1 \sqrt{1-x^2} dx = \frac{1}{2} \int_0^{\pi} (\cos^2 t + \sin^2 t) dt = \frac{\pi}{2}.$$



Figur 16.6

6. Das uneigentliche Integral

$$\int_{-\infty}^{\infty} \frac{dx}{1+x^2}$$

berechnet man wie oben zu

$$2 \lim_{R \rightarrow \infty} \left(\arctan x \Big|_0^R \right) = 2 \lim_{R \rightarrow \infty} \arctan R = 2 \frac{\pi}{2} = \pi.$$

7. Das Integral $\int_0^1 \frac{\arcsin x}{\sqrt{1-x^2}} dx$ ist uneigentlich bei $x = 1$. Die Substitution $t := \arcsin x$, $dt = dx/\sqrt{1-x^2}$ macht hieraus

$$\int_0^1 \frac{\arcsin x}{\sqrt{1-x^2}} dx = \lim_{B \rightarrow \pi/2} \int_0^B t dt = \int_0^{\pi/2} t dt = \frac{\pi^2}{8}.$$

Durch Substitution kann ein uneigentliches Integral also durchaus in eigentliches umgewandelt werden.

8. Das Integral $\int_0^{\infty} \frac{\ln t}{t^2} dt$ konvergiert gegen 1, wie man sofort mit der Substitution $t := e^s$ sieht.

9. Für $0 < a < b$ ergibt sich mit der Substitution $t := (x-a)/(b-a)$, daß das an beiden Grenzen uneigentliche Integral

$$\int_a^b \frac{dx}{\sqrt{(x-a)(b-x)}} = \int_0^1 \frac{dt}{\sqrt{t(1-t)}},$$

also unabhängig von a und b ist, sofern es konvergiert. Bei dem rechts stehenden Integral bietet sich die erneute Substitution $t = \sin^2 u$ mit $dt = 2 \sin 2u du$ an. Damit berechnet man leicht

$$\int_0^1 \frac{dt}{\sqrt{t(1-t)}} = \int_0^{\pi/2} \frac{2 \sin 2u du}{\sin u \cos u} = 2 \int_0^{\pi/2} du = \pi.$$

10. Noch raffinierter ist die Berechnung des an beiden Endpunkten uneigentlichen Integrals

$$J := \int_0^{\pi/2} \ln(\sin x) dx.$$

Wir setzen zuerst $x := 2t$, woraus sich

$$J = 2 \int_0^{\pi/4} \ln(\sin 2t) dt = \frac{\pi}{2} \ln 2 + 2 \int_0^{\pi/4} \ln(\sin t) dt + 2 \int_0^{\pi/4} \ln(\cos t) dt$$

ergibt. Das ganz rechts stehende Integral wird durch $u := \pi/2 - t$ zu

$$\int_{\pi/4}^{\pi/2} \ln(\sin u) du,$$

so daß man nach erneuter Umbenennung der Variablen zu der Gleichung

$$J = \frac{\pi}{2} \ln 2 + 2J, \quad \text{also zu } J = -\frac{\pi}{2} \ln 2$$

gelangt.

Wir überlassen es dem Leser, sich selbst die *algebraischen* Regeln für uneigentliche Integrale zu erarbeiten. Wir stellen nur noch einige allgemeine *Konvergenzkriterien* auf. Man hat viele Sätze über uneigentliche Integrale, die nicht nur formal wie bei unendlichen Reihen aussehen. Z. B. gilt das *CAUCHY-Kriterium* in diesem Kontext in der folgenden Form.

Satz 16.26 Das bei $B \in \mathbb{R} \cup \{\infty\}$ uneigentliche Integral $\int_a^{B-} f(x) dx$ konvergiert genau dann, wenn es zu jedem $\varepsilon > 0$ ein β_0 mit $a \leq \beta_0 < B$ gibt, so daß für alle β, β' mit $\beta_0 \leq \beta < \beta' < B$ gilt:

$$\left| \int_{\beta}^{\beta'} f(x) dx \right| < \varepsilon.$$

Beispiel. Durch partielle Integration findet man für $0 < a < b$ die Beziehung

$$\int_a^b \sqrt{x} \cos x^2 dx = \int_a^b \frac{1}{2\sqrt{x}} (2x \cos x^2) dx = \frac{\sin x^2}{2\sqrt{x}} \Big|_{x=a}^b - \int_a^b \left(\frac{1}{2\sqrt{x}} \right)' \sin x^2 dx.$$

Wähle nun zu vorgegebenem $\varepsilon > 0$ die Zahl $R := \varepsilon^{-2}$. Dann ist für alle $R < a < b$ wegen der Tatsache, daß die unter dem zweiten Integralzeichen stehende Ableitung durchweg negativ ist,

$$\begin{aligned} \left| \int_a^b \sqrt{x} \cos x^2 dx \right| &\leq \frac{1}{2\sqrt{a}} + \frac{1}{2\sqrt{b}} + \int_a^b |\sin x^2| \left| \left(\frac{1}{2\sqrt{x}} \right)' \right| dx \\ &\leq \frac{1}{2\sqrt{a}} + \frac{1}{2\sqrt{b}} + \left| \int_a^b \left(\frac{1}{2\sqrt{x}} \right)' dx \right| = \frac{1}{\sqrt{a}} < \varepsilon. \end{aligned}$$

Also ist das uneigentliche Integral $\int_0^{\infty} \sqrt{x} \cos x^2 dx$ konvergent. Dieses Beispiel zeigt im Übrigen, daß man die Analogie zwischen dem Verhalten von uneigentlichen Integralen und unendlichen Reihen nicht zu weit treiben darf. In einer konvergenten Reihe muß zum Beispiel die Folge a_n der Reihenglieder gegen Null gehen. In dem vorliegenden Beispiel geht aber der Integrand keineswegs bei $x \rightarrow \infty$ gegen Null; ganz im Gegenteil ist der Integrand sogar auf keinem Teilintervall $[a, \infty)$ beschränkt.

Aus dem Cauchy-Kriterium folgt nun die schon früher erwähnte Tatsache, daß man an sich einer *endlichen* Stelle nicht um die Konvergenz uneigentlicher Integrale kümmern braucht, sofern der Integrand dort lokal beschränkt ist. In der Tat werden wir im nachfolgenden Kapitel zeigen, daß im Rahmen der RIEMANNschen Integrationstheorie das Problem in dieser Situation völlig obsolet ist.

Folgerung 16.27 Ist die Funktion $f : [a, B) \rightarrow \mathbb{R}$, $B \in \mathbb{R}$, beschränkt nahe B , und existiert das Integral $\int_a^{\beta} f(x) dx$ für alle $a \leq \beta < B$, so konvergiert das uneigentliche Integral $\int_a^{B-} f(x) dx$.

Beweis. Wir können ohne Einschränkung annehmen, daß f dem Betrage nach auf dem gesamten Intervall $[a, B)$ kleiner oder gleich M ist. Nach Satz 4 ist dann das Integral als Funktion der oberen Grenze LIPSCHITZ-stetig mit Lipschitz-Konstante M :

$$\left| \int_a^{\beta'} f(x) dx - \int_a^{\beta} f(x) dx \right| = \left| \int_{\beta}^{\beta'} f(x) dx \right| \leq M |\beta' - \beta|.$$

Der Rest ergibt sich aus dem Cauchy-Kriterium. □

Eine direkte Folgerung aus dem Prinzip der monotonen Konvergenz ist das *Monotoniekriterium*.

Satz 16.28 (Monotoniekriterium) Die Funktion $f : [a, B) \rightarrow \mathbb{R}$ sei nicht negativ und auf allen kompakten Intervallen $[a, b] \subset [a, B)$ integrierbar. Genau dann konvergiert das uneigentliche Integral $\int_a^{B-} f(x) dx$, wenn die Menge der Integralwerte $\mathcal{I}_a^b(f)$, $a \leq b < B$, nach oben beschränkt ist.

Man kann in diesem Zusammenhang auch über *absolut konvergente* uneigentliche Integrale und das *Majorantenkriterium* reden. Wir setzen für diesen Teil des Kapitels voraus, daß sowohl die Funktion $f : I = [a, B) \rightarrow \mathbb{R}$, $B \in \mathbb{R} \cup \{\infty\}$, als auch ihr Absolutbetrag $|f|$ auf jedem kompakten Teilintervall $[a, b] \subset I$ ein Integral besitze, das zudem als *monoton* vorausgesetzt sei.

Satz 16.29 (Absolute Konvergenz für uneigentliche Integrale) Existiert unter den zuvor gemachten Voraussetzungen das uneigentliche Integral $\mathcal{I}_a^{B-}(|f|)$, so konvergiert auch das uneigentliche Integral $\mathcal{I}_a^{B-}(f)$, und es ist

$$|\mathcal{I}_a^{B-}(f)| \leq \mathcal{I}_a^{B-}(|f|).$$

Beweis. Dies ist eine unmittelbare Folgerung aus dem Cauchy-Kriterium und der wegen $-|f| \leq f \leq |f|$ nach Voraussetzung gültigen Abschätzung

$$|\mathcal{I}_a^{\beta}(f)| \leq \mathcal{I}_a^{\beta}(|f|). \quad \square$$

Ist zusätzlich zu den obigen Voraussetzungen $g \geq 0$ eine weitere „integrierbare“ Funktion, so ergibt sich hieraus ohne weitere Mühe das *Majorantenkriterium* in der folgenden Fassung.

Satz 16.30 (Majorantenkriterium für uneigentliche Integrale) Ist $|f| \leq g$ und existiert das uneigentliche Integral $\mathcal{I}_a^{B-}(g)$, so ist $\mathcal{I}_a^{B-}(f)$ absolut konvergent.

Beispiel. Es ist $\sin x < x$ für $x > 0$ und folglich

$$0 < \frac{\sin 1/x}{x} < \frac{1}{x^2} \quad \text{für } x \geq 1.$$

Da das uneigentliche Integral über $1/x^2$ existiert, konvergiert auch

$$\int_1^{\infty} \frac{\sin(1/x)}{x} dx.$$

Wie im Falle der Reihen kann man hieraus ganz entsprechend ein *Minorantenkriterium* gewinnen.

Folgerung 16.31 (Minorantenkriterium für uneigentliche Integrale) Ist $f(x) \geq g(x) \geq 0$ für alle $a \leq x < B$ und divergiert das Integral $\mathcal{I}_a^{B-}(g)$, so auch $\mathcal{I}_a^{B-}(f)$.

Durch gleichzeitige Anwendung des Monotoniekriteriums und des Satzes über absolute Konvergenz kann man das folgende kleine hinreichende Kriterium für Integranden in Produktgestalt herleiten.

Lemma 16.32 Es seien $f, g : [a, B) \rightarrow \mathbb{R}$ stetig, f sei beschränkt, und das uneigentliche Integral

$$\int_a^{B-} g(x) dx$$

sei absolut konvergent. Dann konvergiert auch

$$\int_a^{B-} f(x)g(x) dx$$

absolut.

Beweis. Nach dem Monotoniekriterium reicht zu zeigen, daß die Funktion

$$x \mapsto \int_a^x |f(t)g(t)| dt, \quad x \in [a, B),$$

beschränkt ist. Nun ist aber $G(x) := \int_a^x |g(t)| \leq M$ und $|f(x)| \leq K$ für alle $a \leq x < B$, so daß KM eine geeignete Schranke ist. \square

Bemerkung. In dem vorigen Lemma kann die Voraussetzung der *absoluten* Konvergenz des Integrals über die Funktion g nicht abgeschwächt werden. Wir werden unten sehen, daß das Integral $\int_0^\infty (\sin x)/x dx$ tatsächlich konvergiert. Es ist aber nicht absolut konvergent, da das Monotoniekriterium verletzt ist:

$$\int_0^{n\pi} \left| \frac{\sin x}{x} \right| dx = \sum_{k=1}^n \int_{(k-1)\pi}^{k\pi} \left| \frac{\sin x}{x} \right| dx \geq \sum_{k=1}^n \frac{1}{k\pi} \int_{(k-1)\pi}^{k\pi} |\sin x| dx = \frac{2}{\pi} \sum_{k=1}^n \frac{1}{k}.$$

Ebenso leicht sieht man, daß auch das Integral $\int_0^\infty (\sin^2 x)/x dx$ divergiert.

Von den vergleichsweise einfach zu beweisenden Konvergenzkriterien notieren wir noch das folgende.

Satz 16.33 (Grenzwertkriterium) *Sind die Funktionen f, g durchweg positiv auf dem Intervall $[a, B)$ und existiert der Grenzwert $\lim_{x \rightarrow B^-} \frac{f(x)}{g(x)} =: \alpha \in \mathbb{R}_+$, so ist mit $\mathcal{I}_a^{B^-}(g)$ auch $\mathcal{I}_a^{B^-}(f)$ konvergent.*

Beweis. Nach Voraussetzung gibt es zu $\varepsilon = 1$ ein $\beta < B$, so daß für alle x mit $\beta \leq x < B$ gilt:

$$\left| \frac{f(x)}{g(x)} - \alpha \right| < 1.$$

Somit ist für diese x auch $0 < f(x) \leq (\alpha + 1)g(x)$, woraus die Behauptung unter Verwendung des Majorantenkriteriums folgt. \square

Bemerkung. Ist im obigen Satz der Grenzwert α positiv, so ist natürlich das uneigentliche Integral $\mathcal{I}_a^{B^-}(f)$ *genau dann* konvergent, wenn dies auf $\mathcal{I}_a^{B^-}(g)$ zutrifft. Der Beweis zeigt übrigens, daß man die Voraussetzung über die Existenz des Grenzwertes α noch abschwächen kann.

Wir unterbrechen unsere theoretischen Überlegungen an dieser Stelle zu einigen kurzen Bemerkungen über die Γ -Funktion. Nach EULER definiert man

$$\Gamma(x) := \int_0^\infty e^{-t} t^{x-1} dt$$

und zeigt den folgenden

Satz 16.34 1. $\Gamma(x)$ konvergiert für alle $x > 0$;

2. für alle $x > 0$ gilt $\Gamma(x+1) = x\Gamma(x)$;

3. für alle $n \in \mathbb{N}^*$ ist $\Gamma(n) = (n-1)!$.

Bemerkung. Die Γ -Funktion ist somit eine (wie wir später zeigen werden, sogar beliebig oft differenzierbare) „Interpolation“ der nur auf den natürlichen Zahlen definierten Fakultät.

Beweis des Satzes. An der unteren Grenze ist das Γ -Integral nur für $x < 1$ uneigentlich. Nun konvergieren die Integrale $\int_0^1 t^{-s} dt$ für $s < 1$. Nach dem Grenzwertkriterium konvergiert daher auch das Integral

$$\int_0^1 e^{-t} t^{x-1} dt \quad \text{für } 1-x < 1, \quad \text{also } x > 0.$$

Die Konvergenz an der oberen Grenze wird von dem Majorantenkriterium bestätigt: Es ist bei festem $x > 0$ ab einer Stelle t_0 wegen $\lim_{t \rightarrow \infty} e^{-t} t^{x+1} = 0$ stets

$$0 < e^{-t} t^{x-1} < \frac{1}{t^2}.$$

Die Funktionalgleichung der Γ -Funktion erhält man durch partielle Integration:

$$\begin{aligned}\Gamma(x+1) &= \int_0^\infty e^{-t} t^x dt = -e^{-t} t^x \Big|_{t=0}^\infty + \int_0^\infty e^{-t} x t^{x-1} dt \\ &= x \int_0^\infty e^{-t} t^{x-1} dt = x \Gamma(x).\end{aligned}$$

Die dritte Formel folgt hieraus durch vollständige Induktion nach n und dem Wert

$$\Gamma(1) = \int_0^\infty e^{-t} dt = 1. \quad \square$$

Auch die tiefer liegenden Konvergenzsätze über Reihen lassen sich auf die Situation von (uneigentlichen) Integralen übertragen. Wir formulieren und beweisen als Beispiel das Analogon zu dem DIRICHLET-Kriterium.

Satz 16.35 Ist f stetig und g stetig differenzierbar auf dem Intervall $[a, B)$, $B \in \mathbb{R} \cup \{\infty\}$, ist

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

in $[a, B)$ beschränkt, und strebt g monoton gegen 0 für $t \rightarrow B-$, so existiert das uneigentliche Integral

$$\int_a^{B-} f(t) g(t) dt.$$

Beispiel. Das Integral $\int_0^\infty \frac{\sin t}{t} dt$ ist tatsächlich nur an der oberen Grenze uneigentlich, da der Integrand sich stetig durch 1 in den Nullpunkt hinein fortsetzen läßt. Wir können uns also z. B. auf das Intervall $I := [\pi/2, \infty)$ beschränken. Die Funktion $f(t) := \sin t$ ist stetig auf I , und $F(x) := \int_{\pi/2}^x f(t) dt = -\cos x$ ist dort beschränkt. Ferner ist $g(t) := 1/t$ stetig differenzierbar und monoton fallend auf I , so daß das anfangs genannte Integral wegen des Dirichlet-Kriteriums konvergiert. Sein Wert ist übrigens $\pi/2$. Wir werden am Ende des Kapitels zeigen, daß dieses Integral nicht absolut konvergiert.

Beweis (des Dirichlet-Kriteriums). Für beliebige $[\alpha, \beta] \subset [a, b)$ findet man vermöge partieller Integration, die im Fall der unendlichen Reihen der Abelschen partiellen Summation entspricht,

$$\int_\alpha^\beta f(x) g(x) dx = F g \Big|_\alpha^\beta - \int_\alpha^\beta F(x) g'(x) dx.$$

Aufgrund der Voraussetzungen geht der „ausintegrierte“ Anteil gegen Null für $\beta \rightarrow b$. Es braucht also nur noch gezeigt zu werden, daß das uneigentliche Integral über $F g'$ existiert. Hierzu benutzen wir das Cauchy-Kriterium, wobei wir noch annehmen dürfen, daß g monoton wächst und damit wegen der Stetigkeit von g' überall $g' \geq 0$ gilt. Ist nun M eine Schranke zu F , so gibt es zu vorgegebenem $\varepsilon > 0$ ein $\beta_0 < \beta$, so daß für alle $x \in [\beta_0, b)$ gilt: $g(x) < \varepsilon/2M$. Dann ist für alle Intervalle $[\alpha, \beta] \subset [\beta_0, b)$

$$\left| \int_\alpha^\beta F(x) g'(x) dx \right| \leq \int_\alpha^\beta |F(x)| g'(x) dx \leq M g \Big|_\alpha^\beta \leq \varepsilon. \quad \square$$

Man kann das vorstehende Resultat auch verwenden zum Nachweis der folgenden Aussage:

Folgerung 16.36 Ist $f : [a, \infty) \rightarrow \mathbb{R}$ stetig differenzierbar mit monoton wachsender Ableitung f' und $\lim_{x \rightarrow \infty} f'(x) = \infty$, so konvergiert das uneigentliche Integral

$$\int_a^\infty \sin f(x) dx .$$

Beweis. Wegen der Grenzwertaussage für die Ableitung von f ist ab f' ab einer geeigneten Stelle positiv und damit f von da ab streng monoton wachsend. Ohne Einschränkung der Allgemeinheit können wir also annehmen, daß dies schon auf dem Definitionsintervall $[a, \infty)$ der Fall ist. Wegen $\lim_{x \rightarrow \infty} f(x) = \infty$ besitzt f eine (streng monotone) Umkehrfunktion $g : [f(a), \infty)$. Wir substituieren $x = f(t)$, also $t = g(x)$ und $dt = \frac{dx}{f'(g(x))}$ und gewinnen daraus

$$\int_0^\infty \sin f(t) dt = \int_{f(a)}^\infty \frac{\sin x}{f'(g(x))} dx .$$

Das umgeformte Integral ist nach dem vorher bewiesenen Kriterium konvergent. \square

Beispiel und Bemerkung. Insbesondere existieren gemäß dieser Folgerung die sogenannten FRESNELSchen Integrale $\int_0^\infty \sin x^2 dx$ und $\int_0^\infty \cos x^2 dx$. Man beweist in der Funktionentheorie, daß

$$\int_0^\infty \sin x^2 dx = \int_0^\infty \cos x^2 dx = \sqrt{\frac{\pi}{8}} .$$

Mit der gleichen Methode wie beim Dirichlet-Kriterium gelangt man auch beim Abelschen Kriterium zum Ziel, weshalb wir auf einen Beweis verzichten.

Satz 16.37 Es sei f stetig und g stetig differenzierbar auf dem Intervall $[a, B)$, $B \in \mathbb{R} \cup \{\infty\}$, es existiere das uneigentliche Integral

$$\int_a^{B-} f(x) dx ,$$

und g sei monoton und beschränkt. Dann existiert auch das uneigentliche Integral

$$\int_a^{B-} f(x) g(x) dx .$$

Es gibt ferner noch einen wunderbaren *Vergleichssatz* zwischen Reihen und uneigentlichen Integralen, den man in beiden Richtungen verwenden kann.

Satz 16.38 Die Funktion $f : [0, \infty) \rightarrow \mathbb{R}$ sei positiv und monoton fallend (und damit über jedes kompakte Intervall integrierbar im Sinne des Regel-Integrals). Dann gilt: Das uneigentliche Integral

$$\int_0^\infty f(x) dx$$

konvergiert genau dann, wenn die Reihe

$$\sum_{n=0}^{\infty} f(n)$$

konvergiert.

Beweis. Im Intervall $n \leq x \leq n+1$ gilt nach Voraussetzung $f(n) \geq f(x) \geq f(n+1) \geq 0$, so daß

$$f(n+1) \leq \int_n^{n+1} f(x) dx \leq f(n)$$

ist und

$$\sum_{n=1}^{N+1} f(n) \leq \int_0^{N+1} f(x) dx \leq \sum_{n=0}^N f(n)$$

für alle $N \in \mathbb{N}$. Wegen $f(n) \geq 0$ ist aber die Reihe $\sum_{n=0}^{\infty} f(n)$ genau dann konvergent, wenn die Folge ihrer Partialsummen beschränkt ist. Nach dem Vorhergehenden ist dies wiederum äquivalent zur Beschränktheit der monoton aufsteigenden Folge der Integrale $\int_0^{N+1} f(x) dx$, also zur Konvergenz des uneigentlichen Integrals $\int_0^{\infty} f(x) dx$. \square

Beispiel. Wir haben oben gesehen, daß das uneigentliche Integral $\int_1^{\infty} \frac{dx}{x^s}$ für $s > 1$ existiert. Da der Integrand die Voraussetzungen des vorigen Satzes erfüllt, kann man hieraus erneut schließen, daß für alle $s > 1$ die RIEMANNSCHE ζ -Funktion

$$\zeta(s) := \sum_{n=1}^{\infty} \frac{1}{n^s}$$

konvergiert.

Wir beenden dieses Kapitel mit einigen Resultaten über das Verhalten von uneigentlichen Integralen bei *konvergenten Funktionenfolgen*. Bei *gleichmäßiger Konvergenz* herrschen die alten Verhältnisse, sofern das betrachtete Intervall *endlich* ist. Dies besagt der folgende Satz.

Satz 16.39 Die Funktionenfolge $f_j : [a, B) \rightarrow \mathbb{R}$, $B \in \mathbb{R}$, sei gleichmäßig konvergent gegen eine Grenzfunktion f . Existieren die uneigentlichen Integrale

$$\int_a^{B-} f_j(x) dx, \quad \text{so auch} \quad \int_a^{B-} f(x) dx,$$

und es gilt

$$\lim_{j \rightarrow \infty} \int_a^{B-} f_j(x) dx = \int_a^{B-} \lim_{j \rightarrow \infty} f_j(x) dx = \int_a^{B-} f(x) dx.$$

Beweis. Wir beweisen zuerst mit dem Cauchy-Kriterium, daß die Folge $I_j := \int_a^{B-} f_j(x) dx$ der uneigentlichen Integrale konvergiert: Es existiert zu vorgegebenem $\varepsilon > 0$ ein $N \in \mathbb{N}$, so daß für alle $j, k \geq N$ und alle $x \in [a, b)$ die Abschätzung

$$|f_j(x) - f_k(x)| < \frac{\varepsilon}{b-a}$$

besteht. Für diese j, k und alle $\beta < b$ hat man dann weiter die Abschätzung

$$\left| \int_a^{\beta} f_j(x) dx - \int_a^{\beta} f_k(x) dx \right| \leq \int_a^{\beta} |f_j(x) - f_k(x)| dx \leq \varepsilon,$$

woraus sich unmittelbar

$$|I_j - I_k| = \left| \lim_{\beta \nearrow b} \int_a^{\beta} (f_j(x) - f_k(x)) dx \right| = \lim_{\beta \nearrow b} \left| \int_a^{\beta} (f_j(x) - f_k(x)) dx \right| \leq \varepsilon$$

ergibt. Also existiert

$$A := \lim_{j \rightarrow \infty} \int_a^{b-} f_j(x) dx .$$

Wegen der gleichmäßigen Konvergenz der Funktionenfolge (f_j) auf $[a, b)$ ist die Grenzfunktion f auf jedem kompakten Intervall $[a, \beta] \subset [a, b)$ integrierbar. Es ist noch zu zeigen, daß

$$\int_a^{b-} f(x) dx = \lim_{\beta \nearrow b} \int_a^\beta f(x) dx = A .$$

Nun gibt es zu beliebigem $\varepsilon > 0$ ein $N \in \mathbb{N}$ mit

$$\left| \int_a^{b-} f_N(x) - A dx \right| < \frac{\varepsilon}{3} \quad \text{und} \quad |f_N(x) - f(x)| < \frac{\varepsilon}{3(b-a)} \quad \text{für alle } x \in [a, b) ,$$

da die Folge der f_j gleichmäßig auf $[a, b)$ gegen f konvergiert. Ferner gibt es ein $\beta_0 < b$, so daß

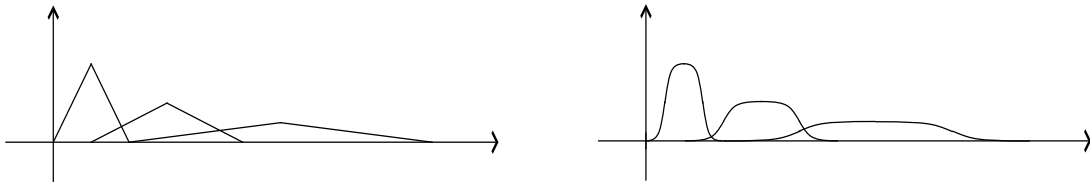
$$\left| \int_a^\beta f_N(x) dx - \int_a^{\beta_0} f_N(x) dx \right| \leq \frac{\varepsilon}{3} \quad \text{für alle } \beta_0 \leq \beta < b .$$

Dann gilt für alle diese β :

$$\begin{aligned} \left| \int_a^\beta f(x) dx - A \right| &\leq \left| \int_a^\beta f(x) dx - \int_a^\beta f_N(x) dx \right| + \left| \int_a^\beta f_N(x) dx - \int_a^{\beta_0} f_N(x) dx \right| \\ &\quad + \left| \int_a^{\beta_0} f_N(x) dx - A \right| \\ &\leq \int_a^\beta |f_N(x) dx - f(x)| dx + \frac{2}{3} \varepsilon \leq \varepsilon , \end{aligned}$$

und die Behauptung ist bewiesen. \square

Bemerkung und Beispiel. Die vorstehende Aussage ist i. A. nicht richtig für *unbeschränkte* Intervalle: Es gibt stetige und sogar C^∞ -Funktionen $g: \mathbb{R} \rightarrow \mathbb{R}$, die außerhalb des Intervalls $(-1, 1)$ verschwinden, auf diesem positiv sind und (damit) ein positives Integral $\int_{-1}^1 g(x) dx =: A > 0$ besitzen. Eine solche stetige Funktion ist $g(x) = 1 - |x|$, $|x| < 1$, $g(x) = 0$ sonst, oder im differenzierbaren Fall die durch Null fortgesetzte Funktion $g(x) := \exp\left(\frac{-1}{1-x^2}\right)$, $|x| < 1$. Betrachte nun die Funktionenfolge $f_0(x) := g(x-1)$, $f_n(x) := 2^{-n} f_0(2^{-n}(x-n))$. Damit ist $f_n(x) = 0$ für $x < n$, so daß die Folge (f_n) auf \mathbb{R} punktweise gegen 0 konvergiert. (Die folgenden Skizzen sind überhöht).



Figur 16.7

Die Konvergenz ist sogar gleichmäßig, da $\sup f_n = 2^{-n} \sup f_0$. Mit der Substitution $t := 2^{-n}(x-n)$ rechnet man aber sofort nach, daß

$$\int_0^\infty f_n(x) dx = A \quad \text{und damit} \quad \lim_{n \rightarrow \infty} \int_0^\infty f_n(x) dx = A > 0 = \int_0^\infty \lim_{n \rightarrow \infty} f_n(x) dx .$$

Man kann auch auf $(0, \infty)$ z. B. die (bei $n > 0$ nirgends verschwindenden) Funktionen $f_n(x) := n e^{-n/(2x^2)}/x^3$ betrachten. Mit dem Hauptsatz der Differential- und Integralrechnung findet man sehr leicht

$$\int_0^{\infty} f_n(x) dx = e^{-n/(2x^2)} \Big|_{x=0}^{\infty} = 1.$$

Andererseits besitzt die Funktion f_n den maximalen Wert $3\sqrt{3}e^{-3/2}/\sqrt{n}$, und dieser geht gegen Null. Also ist die Funktionenfolge gleichmäßig auf $(0, \infty)$ gegen Null konvergent.

Anhang: Tabelle der wichtigsten Stammfunktionen

Zur Bequemlichkeit des Lesers stellen wir hier eine Tabelle von einigen der bisher gewonnenen Stammfunktionen zusammen.

$$\begin{aligned}\int x^n dx &= \frac{x^{n+1}}{n+1}, & n \in \mathbb{Z} \setminus \{-1\}, \\ \int x^\alpha dx &= \frac{x^{\alpha+1}}{\alpha+1}, & \alpha \in \mathbb{R} \setminus \{-1\}, \\ \int a^x dx &= \frac{a^x}{\ln a}, & x \in \mathbb{R}, a > 0, \\ \int \frac{dx}{x} &= \ln |x|, & x \in \mathbb{R}^* = \mathbb{R} \setminus \{0\}, \\ \int \ln x dx &= x(\ln x - 1), & x > 0, \\ \int \sin x dx &= -\cos x, & x \in \mathbb{R}, \\ \int \cos x dx &= \sin x, & x \in \mathbb{R}, \\ \int \sinh x dx &= \cosh x, & x \in \mathbb{R}, \\ \int \cosh x dx &= \sinh x, & x \in \mathbb{R}, \\ \int \frac{dx}{\cos^2 x} &= \tan x, & x \neq \pi/2 + \pi\mathbb{Z}, \\ \int \frac{dx}{\sin^2 x} &= -\cot x, & x \notin \pi\mathbb{Z}, \\ \int \frac{dx}{\cosh^2 x} &= \tanh x, & x \in \mathbb{R}, \\ \int \frac{dx}{\sinh^2 x} &= -\coth x, & x \in \mathbb{R}^*, \\ \int \tan x dx &= -\ln |\cos x|, \\ \int \cot x dx &= \ln |\sin x|, \\ \int \tanh x dx &= \ln \cosh x, \\ \int \coth x dx &= \ln |\sinh x|, \\ \int \frac{dx}{1+x^2} &= \arctan x, & x \in \mathbb{R}, \\ \int \frac{dx}{1-x^2} &= \operatorname{Artanh} x, & |x| < 1, \\ \int \frac{dx}{1-x^2} &= \frac{1}{2} \ln \left| \frac{1+x}{1-x} \right|, & |x| \neq 1,\end{aligned}$$

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

$$\int \sqrt{1+x^2} dx = \frac{1}{2} (x \sqrt{1+x^2} + \operatorname{Arsinh} x), \quad x \in \mathbb{R},$$

$$\int \sqrt{x^2-1} dx = \frac{1}{2} (x \sqrt{x^2-1} - \operatorname{Arcosh} x), \quad x \geq 1,$$

$$\int \frac{dx}{\sqrt{1-x^2}} = \arcsin x, \quad |x| < 1,$$

$$\int \frac{dx}{\sqrt{x^2-1}} = \operatorname{Arcosh} x = \ln(x + \sqrt{x^2-1}), \quad x > 1,$$

$$\int \frac{dx}{\sqrt{x^2-1}} = -\operatorname{Arcosh}(-x), \quad x < -1,$$

$$\int \frac{dx}{\sqrt{1+x^2}} = \operatorname{Arsinh} x = \ln(x + \sqrt{1+x^2}), \quad x \in \mathbb{R},$$

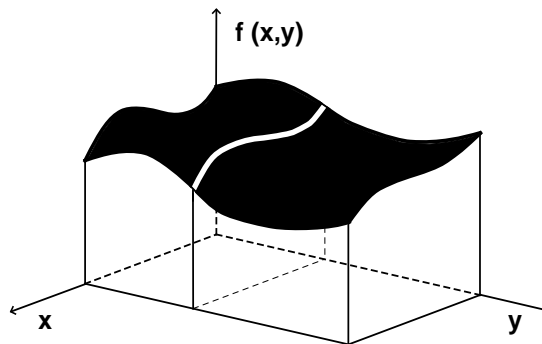
$$\int \arcsin x dx = x \arcsin x + \sqrt{1-x^2},$$

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

17 Differenzierbare Abbildungen, Kettenregel und Taylor-Formel

Stetige Funktionen f in einer reellen Veränderlichen nehmen ihre *Extrema* auf *kompakten* Intervallen $I = [a, b]$ an. Ist f überdies differenzierbar, so liefert das Verschwinden der Ableitung f' eine *notwendige* Bedingung für die Stellen, an denen das Extremum angenommen wird. Wir haben in Kapitel 15 gesehen, daß die erste Aussage auch für reellwertige Funktionen auf *metrischen* (oder sogar *topologischen*) Räumen X richtig bleibt, wenn man das kompakte Intervall I durch eine *kompakte Menge* $K \subset X$ ersetzt. Um die entsprechenden *differentiellen Kriterien* herleiten zu können, müssen wir uns überlegen, in welchem *generellen* Rahmen man über das Konzept der *Differenzierbarkeit* für Abbildungen $f : X \rightarrow Y$ verfügen kann. Verfolgt man das Konzept des *Differentialquotienten*, so hat man wie in Kapitel 15 nur die Wahl, f als Abbildung von (einer Teilmenge eines *bewerteten Körpers*) \mathbb{K} in einen *normierten* \mathbb{K} -Vektorraum W zuzulassen. Verfolgt man jedoch das Konzept der *linearen Approximation*, so kann der Definitionsbereich U der Abbildung f auch eine (geeignete) Teilmenge eines weiteren normierten \mathbb{K} -Vektorraumes V sein, wie wir anschließend genauer erläutern werden.

Es sei also im folgenden \mathbb{K} ein bewerteter Körper (vorzugsweise $\mathbb{K} = \mathbb{R}, \mathbb{C}$) mit Bewertungskörper \mathbb{K}_0 , und V, W seien normierte \mathbb{K} -Vektorräume; ferner sei $f : U \rightarrow W, U \subset V$, eine Abbildung, wobei zunächst über U nichts weiter gesagt sei. Der uns am meisten interessierende Fall ist der, daß U eine *offene* Teilmenge von $V := \mathbb{R}^n$ und $W = \mathbb{R}^m$, speziell also $m = 1$ und damit f eine Funktion in n reellen Veränderlichen ist:



Figur 17.1

Als erstes Konzept führen wir nun durch eine geeignete „Reduzierung“ der Parameter die sogenannten *Richtungsableitungen* ein. Es sei $a \in U$ und $v \in V$ ein Vektor. Wir betrachten die parametrisierte Gerade

$$\mathbb{K} \ni t \longrightarrow \gamma(t) := a + tv$$

und nennen die Menge U in a bzgl. der „Richtung“ v *zulässig*, wenn $0 \in \mathbb{K}$ ein Häufungspunkt der Menge $\gamma^{-1}(U \setminus \{a\})$ ist. Man sieht, daß dann $v \neq 0$ sein muß und die Bedingung auch automatisch für alle Vielfachen $\lambda v, \lambda \neq 0$, erfüllt ist (somit also nur von der *nichtorientierten Richtung* von v abhängt oder, anders ausgedrückt: nur von der v zugeordneten *nichtparametrisierten* Geraden). Ist U *offen*, so ist U in jedem Punkt bzgl. jeder Richtung zulässig.

Definition. Unter der obigen Voraussetzung heißt

$$\partial_v f(a) := \lim_{t \rightarrow 0} \frac{f(\gamma(t)) - f(a)}{t} = \left. \frac{d}{dt} (f \circ \gamma(t)) \right|_{t=0}$$

die *Richtungsableitung* von f in Richtung v an der Stelle a , sofern dieser Grenzwert existiert.

Bemerkung. Aus der Kettenregel des Kapitels 14 folgt sofort: Existiert $\partial_v f(a)$, so auch $\partial_{\lambda v} f(a)$ für alle $\lambda \in \mathbb{K}$, und es ist

$$\partial_{\lambda v} f(a) = \lambda \partial_v f(a), \quad \lambda \neq 0.$$

Selbstverständlich gilt immer $\partial_{0v} f(a) = 0$. Ist speziell $U \subset \mathbb{R}^n$, so kann man insbesondere v gleich dem j -ten Einheitsvektor e_j wählen. Man schreibt dann

$$\frac{\partial f}{\partial x_j}(a) := \partial_j f(a) := D^j f(a) = \partial_{e_j} f(a)$$

und nennt diesen Ausdruck die *partielle Ableitung* von f an der Stelle a nach der Variablen x_j , $j = 1, \dots, n$. Andere Bezeichnungen für die partiellen Ableitungen sind auch $f_{x_j}(a)$ oder $f_{,j}(a)$.

Wir stellen jetzt die folgende Frage: *Unter welcher Voraussetzung an f in a folgt aus der Existenz der partiellen Ableitungen $\partial_1 f(a), \dots, \partial_n f(a)$ die Existenz aller Richtungsableitungen mit*

$$\partial_v f(a) = \sum_{j=1}^n v_j \frac{\partial f}{\partial x_j}(a), \quad v = (v_1, \dots, v_n) \in V = \mathbb{R}^n ?$$

Dieses sicherlich erwünschte *lineare Verhalten der Richtungsableitungen* hat tatsächlich mit *linearer Approximierbarkeit* der Abbildung f bei a zu tun. Wir verallgemeinern dazu unsere frühere Definition für Funktionen. Es seien dazu V , W und $a \in U \subset V$ wie oben gegeben, U sei hierbei stets eine *offene* Menge.

Definition. $f : U \rightarrow W$ heißt in a *total differenzierbar*, wenn f dort (stetig) linear approximierbar ist, d. h. wenn es eine stetige Linearform $L : V \rightarrow W$ gibt mit

$$f(a + v) = f(a) + L(v) + \varphi(v), \quad v \in V \text{ nahe } 0,$$

wobei

$$\lim_{v \rightarrow 0} \frac{\|\varphi(v)\|_W}{\|v\|_V} = 0.$$

Bemerkungen. 1. Eine *Linearform* $L : V \rightarrow W$ ist genau dann stetig, wenn es eine Konstante C gibt mit $\|L(v)\|_W \leq C \|v\|_V$ (siehe Satz 1). Dies ist immer richtig, wenn V endlich-dimensional ist.

2. Ist f total differenzierbar in a , so ist insbesondere φ stetig in 0 und damit f stetig in a .

3. Die Linearform $L : V \rightarrow W$ ist eindeutig bestimmt. Denn sonst würde die Differenz zweier solcher Formen eine Linearform φ sein mit

$$\lim_{v \rightarrow 0} \frac{\|\varphi(v)\|_W}{\|v\|_V} = 0.$$

Wäre $\varphi(v) \neq 0$ für ein $v \neq 0$, so wäre

$$\frac{\|\varphi(\lambda v)\|_W}{\|\lambda v\|_V} = \frac{\|\varphi(v)\|_W}{\|v\|_V} \neq 0$$

und damit der Grenzwert der linken Seite für $\lambda \rightarrow 0$ nicht gleich Null.

Man schreibt oft auch $(Df)(a)$ oder $D_a f$ für L und nennt $(Df)(a)$ das *totale Differential* von f an der Stelle a .

Bevor wir weiter gehen im Studium differenzierbarer Abbildungen, wollen wir erst die Bemerkung 1 begründen.

Satz 17.1 *Es seien V und W normierte Vektorräume, $F : V \rightarrow W$ sei eine lineare Abbildung. Dann sind die folgenden Aussagen äquivalent :*

- i) F ist stetig in 0 ;
- ii) F ist stetig ;
- iii) F ist Lipschitz-stetig, d. h. es gibt eine Zahl $L > 0$, s. d. $\|F(v) - F(w)\| \leq L\|v - w\|$ für alle $v, w \in V$;
- iv) F ist gleichmäßig stetig ;
- v) F ist beschränkt auf der Einheitskugel $\{v \in V : \|v\| \leq 1\}$;
- vi) es gibt eine Zahl $C > 0$, s. d. $\|F(v)\| \leq C\|v\|$ für alle $v \in V$.

Beweis. Es gilt mit Sicherheit (mit $L = C$) vi) \implies iii) \implies iv) \implies ii) \implies i). Wir brauchen daher nur noch i) \implies v) \implies vi) zu zeigen. Ist $\varepsilon = 1$, so gibt es unter der Voraussetzung i) ein $\delta > 0$, so daß aus $\|v\| \leq \delta$ folgt: $\|F(v)\| \leq 1$. Ist nun $\|v\| \leq 1$, so ist $\|\delta \cdot v\| = |\delta| \|v\| \leq \delta$ und damit

$$\|F(v)\| = \left\| F\left(\frac{1}{\delta} \delta v\right) \right\| = \frac{1}{\delta} \|F(\delta v)\| \leq \frac{1}{\delta} =: C.$$

Ist andererseits diese Ungleichung für alle v in der Einheitskugel erfüllt, so ist vi) mit der gleichen Konstanten richtig für alle $v \in V$, denn für $v \neq 0$ ist

$$\frac{1}{\|v\|} F(v) = F\left(\frac{v}{\|v\|}\right) \leq C. \quad \square$$

Bemerkungen und Definition. 1. Obwohl lineare Abbildungen *niemals* im *eigentlichen Sinne* beschränkt sind, d. h. niemals $\|F(v)\| \leq C$ für alle $v \in V$ gilt, außer wenn $F = 0$ ist, nennt man lineare Abbildungen mit einer der obigen äquivalenten Eigenschaften i) bis vi) doch *beschränkt linear*. Dies ist also immer zu lesen als linear *und* beschränkt auf der Einheitskugel. Man nennt C wie oben eine *Schranke* für F .

2. Sind V und W endlich-dimensional, so sind alle Aussagen für *jede* lineare Abbildung erfüllt. In diesem Falle stimmen also die Begriffe *Linearität* und *stetige Linearität* überein. Für unendlich-dimensionale Vektorräume ist dies jedoch nicht der Fall. (Siehe hierzu ein Beispiel im Anhang zu diesem Kapitel).

3. Es ist übrigens klar, daß der Graph einer total differenzierbaren Abbildung gut *linear approximiert* werden kann: Ist $f : U \rightarrow W$ in $a \in U \subset V$ total differenzierbar und $L : V \rightarrow W$ das zugehörige totale Differential, so approximiert der affine Raum

$$\{(x, y) \in U \times W : y = f(a) + L(x - a)\}$$

den Graphen

$$\{(x, y) \in U \times W : y = f(x)\}$$

„besser als linear“. Wir nennen ihn den *Tangentenraum* an den Graphen von f an der Stelle a .

Wir wollen als nächstes den Zusammenhang zwischen totaler Differenzierbarkeit und der Existenz von partiellen Ableitungen bzw. allen Richtungsableitungen aufklären. Eine Richtung ist sehr einfach.

Lemma 17.2 *Ist f in a total differenzierbar, so existieren alle Richtungsableitungen, und es ist*

$$\partial_v f(a) = L(v).$$

Ist insbesondere $U \subset V = \mathbb{R}^n$, so gilt

$$\partial_v f(a) = L(v) = \sum_{j=1}^n v_j \frac{\partial f}{\partial x_j}(a)$$

für $v = (v_1, \dots, v_n)$.

Beweis. Es gilt

$$\lim_{t \rightarrow 0} \frac{f(a + tv) - f(a)}{t} = \lim_{t \rightarrow 0} \frac{L(tv) + \varphi(tv)}{t} = L(v) + \lim_{t \rightarrow 0} \frac{\varphi(tv)}{t} = L(v),$$

da

$$\left\| \frac{\varphi(tv)}{t} \right\|_W = \frac{\|\varphi(tv)\|_W}{\|tv\|_V} \|v\|_V \xrightarrow{t \rightarrow 0} 0 \quad \text{für } v \neq 0. \quad \square$$

Kurz gesagt, impliziert also *totale Differenzierbarkeit* in einem Punkt die *partielle Differenzierbarkeit*. Die Umkehrung gilt nicht, wie man sich leicht an geeigneten Beispielen klar machen kann. Man kann sogar Beispiele konstruieren, bei denen *alle* Richtungsableitungen in einem Punkt existieren und trotzdem die Funktion dort nicht total differenzierbar ist.

Beispiele. 1. Die Funktion in zwei Veränderlichen

$$f(x, y) := \begin{cases} \frac{xy}{x^2 + y^2} & \text{für } (x, y) \neq (0, 0), \\ 0 & \text{für } (x, y) = (0, 0) \end{cases}$$

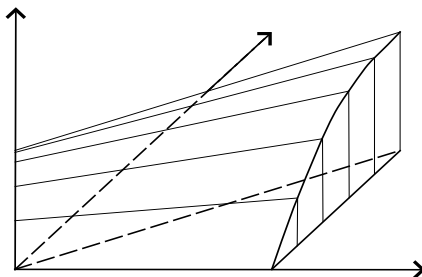
ist offensichtlich an allen Stellen $(x, y) \in \mathbb{R}^2 \setminus \{(0, 0)\}$ *stetig partiell differenzierbar* mit

$$\frac{\partial f}{\partial x}(x, y) = \frac{y(y^2 - x^2)}{(x^2 + y^2)^2} \quad \text{und} \quad \frac{\partial f}{\partial y}(x, y) = \frac{x(x^2 - y^2)}{(x^2 + y^2)^2},$$

und auch im Nullpunkt existieren beide partiellen Ableitungen:

$$\frac{\partial f}{\partial x}(0, 0) = \lim_{x \rightarrow 0} \frac{f(x, 0) - f(0, 0)}{x} = 0, \quad \frac{\partial f}{\partial y}(0, 0) = 0.$$

Der Graph dieser Funktion sieht (ausschnittsweise) in etwa wie folgt aus:



Figur 17.2

Die Funktion f selbst ist aber im Nullpunkt nicht einmal stetig, kann insbesondere also dort auch nicht total differenzierbar sein: Schränkt man die Funktion nämlich auf die Gerade $y = cx$ ein, so ist dort $f(x, y) = c/(1 + c^2)$, und der Grenzwert für $x \rightarrow 0$ ist nur dann gleich 0, wenn $c = 0$.

Die Funktion f besitzt auch *keine Richtungsableitungen* im Nullpunkt (außer denen in Richtung der Koordinatenachsen): Mit einem beliebigen Vektor $(v_1, v_2) \in \mathbb{R}^2$ mit $v_1 v_2 \neq 0$ ist

$$\partial_{(v_1, v_2)} f(0, 0) = \lim_{t \rightarrow 0} \frac{f(tv_1, tv_2) - f(0, 0)}{t} = \lim_{t \rightarrow 0} \frac{f(v_1, v_2)}{t} = \pm \infty.$$

Mit einem ähnlichen Argument sieht man ein, daß die partiellen Ableitungen $\partial_1 f$, $\partial_2 f$ im Nullpunkt nicht stetig sind.

2. Die Funktion

$$f(x, y) := \begin{cases} \frac{x^2 y}{x^2 + y^2} & \text{für } (x, y) \neq (0, 0), \\ 0 & \text{für } (x, y) = (0, 0) \end{cases}$$

besitzt im Nullpunkt sogar alle *Richtungsableitungen*. Für $(v_1, v_2) \in \mathbb{R}^2$ ist wegen $f(tv_1, tv_2) = t f(v_1, v_2)$:

$$\partial_{(v_1, v_2)} f(0, 0) = f(v_1, v_2),$$

insbesondere also $\partial_1 f(0, 0) = \partial_2 f(0, 0) = 0$. Wäre aber f in 0 total differenzierbar, so müßte für das totale Differential $Df_0 = 0$ gelten, was für alle Vektoren $v = (v_1, v_2) \in \mathbb{R}^2$ den offensichtlichen Unsinn $f(v_1, v_2) = (Df)_0(v) = 0$ nach sich ziehen würde. Wie im vorigen Beispiel rechnet man leicht die partiellen Ableitungen außerhalb des Nullpunkts aus und stellt fest, daß diese im Nullpunkt nicht stetig sind.

Eine *hinreichende* Bedingung für totale Differenzierbarkeit kann man nun leicht für Abbildungen $f: U \rightarrow \mathbb{R}^m$, $U \subset \mathbb{R}^n$ eine offene Menge, angeben. Man beachte, daß in diesem Fall f als ein m -tupel von Funktionen aufgefaßt werden kann: $f = {}^t(f_1, \dots, f_m)$, und $(Df)(a)$ identifiziert sich, wenn man Vektoren in V und W als Spaltenvektoren auffaßt, mit der *Funktionalmatrix* oder *Jacobi-Matrix* von f an der Stelle a :

$$(Df)(a) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(a) & \cdots & \frac{\partial f_1}{\partial x_n}(a) \\ \frac{\partial f_2}{\partial x_1}(a) & \cdots & \frac{\partial f_2}{\partial x_n}(a) \\ \vdots & & \vdots \\ \frac{\partial f_m}{\partial x_1}(a) & \cdots & \frac{\partial f_m}{\partial x_n}(a) \end{pmatrix} = \left(\frac{\partial f_j}{\partial x_k}(a) \right)_{\substack{j=1, \dots, m \\ k=1, \dots, n}}.$$

Es ist klar, daß f genau dann *stetig partiell differenzierbar* ist, wenn dies für die Funktionen f_1, \dots, f_m gilt. Eine entsprechende Aussage ist auch für *totale Differenzierbarkeit* richtig; dies folgt unmittelbar aus der *Äquivalenz der Normen* auf \mathbb{R}^m (siehe Anhang 1 zu diesem Kapitel). Man kann sich daher, wie auch bei vielen weiteren Überlegungen, auf den Fall $m = 1$ beschränken.

In sämtlichen „pathologischen“ Beispielen waren die partiellen Ableitungen an den fraglichen Stellen *unstetig*. Schließt man diese Möglichkeit aus, so ist die Welt wieder in bester Ordnung.

Satz 17.3 *Es sei $f: Q \rightarrow \mathbb{R}$, $Q = (-a, a)^n$, $a > 0$, in jedem Punkt von Q partiell differenzierbar, und die partiellen Ableitungen $\partial f / \partial x_j$, $j = 1, \dots, n$, seien stetig im Ursprung. Dann ist f in 0 total differenzierbar.*

Folgerung 17.4 *Ist $f: U \rightarrow \mathbb{R}$, U eine offene Menge in \mathbb{R}^n , stetig partiell differenzierbar, so ist f in jedem Punkt $a \in U$ total differenzierbar.*

Beweis (Satz 3). Es sei $x = (x_1, \dots, x_n) \in Q$ ein fest gewählter Punkt. Dann sind auch die Punkte $x^{(n)} := x$, $x^{(n-1)} = (x_1, \dots, x_{n-1}, 0)$, $x^{(n-2)} = (x_1, \dots, x_{n-2}, 0, 0), \dots, x^{(0)} = (0, \dots, 0) = 0 \in Q$. Betrachte die Funktionen

$$g_j(t) := f(x_1, \dots, x_{j-1}, t, 0, \dots, 0), \quad |t| < a.$$

Diese Funktionen sind differenzierbar für $|t| < a$ mit

$$\frac{dg_j}{dt}(t) = \frac{\partial f}{\partial x_j}(x_1, \dots, x_{j-1}, t, 0, \dots, 0).$$

Aus dem 1. Mittelwertsatz der Differentialrechnung folgt dann mit einem geeigneten Wert t_j zwischen 0 und x_j , daß

$$f(x^{(j)}) - f(x^{(j-1)}) = g_j(x_j) - g_j(0) = g'_j(t_j)x_j = \frac{\partial f}{\partial x_j}(\xi_j)x_j$$

mit $\xi_j = \xi_j(x) = (x_1, \dots, x_{j-1}, t_j, 0, \dots, 0)$. Setze nun

$$\psi_j(x) = \frac{\partial f}{\partial x_j}(\xi_j(x));$$

dann gilt

$$f(x) - f(0) = \sum_{j=1}^n (f(x^{(j)}) - f(x^{(j-1)})) = \sum_{j=1}^n \psi_j(x)x_j,$$

und die Funktionen ψ_j sind in 0 stetig, da mit x gegen 0 auch $\xi_j(x)$ gegen 0 geht. Setzt man $c_j := \psi_j(0)$ und $\varphi(x) := \sum_{j=1}^n (\psi_j(x) - \psi_j(0))x_j$, so ist

$$f(x) = f(0) + \sum_{j=1}^n c_j x_j + \varphi(x),$$

und $\varphi(x)/\|x\|_\infty$ geht gegen Null für $x \rightarrow 0$. □

Offensichtlich haben wir am Ende des Beweises noch den folgenden Zusatz gezeigt:

Satz 17.5 Eine Funktion $f : U \rightarrow \mathbb{R}$ ist in $a \in U \subset \mathbb{R}^n$ genau dann total differenzierbar, wenn es (i. a. nicht eindeutig bestimmte) Funktionen $\psi_j : U \rightarrow \mathbb{R}$, $j = 1, \dots, n$, gibt, die in a stetig sind, und für die gilt:

$$f(x) = f(a) + \sum_{j=1}^n \psi_j(x)(x_j - a_j).$$

Die f in a zugeordnete totale Ableitung $(Df)_a$ ist dann gegeben durch

$$(Df)_a(v) = \sum_{j=1}^n \psi_j(a)v_j, \quad v = (v_1, \dots, v_n) \in \mathbb{R}^n.$$

Insbesondere ist

$$\frac{\partial f}{\partial x_j} = \psi_j(a).$$

Bemerkung. Im Komplexen ist Differenzierbarkeit, wie wir früher schon angemerkt haben, eine viel stärkere Eigenschaft als im Reellen. Dies zeigt sich auch in mehreren Veränderlichen: Ist $f : D_1 \times \dots \times D_n \rightarrow \mathbb{C}$ partiell komplex differenzierbar, $D_j \subset \mathbb{C}$ eine offene Kreisscheibe, so sind

die partiellen Ableitungen $\partial f / \partial z_j$, $j = 1, \dots, n$, automatisch stetig, und f ist an jeder Stelle $a \in D$ total (komplex) differenzierbar. Es gilt sogar, daß f in jedem Punkt *analytisch* (oder *holomorph*) ist, d. h. um jeden Punkt in eine konvergente Potenzreihe in n komplexen Veränderlichen entwickelt werden kann.

Wir kommen nun zu den *algebraischen* Eigenschaften des Begriffs der totalen und partiellen Differenzierbarkeit.

Satz 17.6 *Es seien die Abbildungen $f, f_1, f_2 : U \rightarrow W$ total differenzierbar in a , und das gleiche gelte für $c : U \rightarrow \mathbb{K}$. Dann sind auch die Abbildungen*

$$f_1 + f_2, \quad cf$$

total differenzierbar in a , und es gilt

$$D_a(f_1 + f_2) = D_a f_1 + D_a f_2, \quad D_a(cf) = c(a) \cdot D_a f + D_a c \cdot f(a).$$

Insbesondere gilt

$$\begin{aligned} \frac{\partial (f_1 + f_2)}{\partial x_j}(a) &= \frac{\partial f_1}{\partial x_j}(a) + \frac{\partial f_2}{\partial x_j}(a), \\ \frac{\partial (cf)}{\partial x_j}(a) &= c(a) \frac{\partial f}{\partial x_j}(a) + \frac{\partial c}{\partial x_j}(a) f(a). \end{aligned}$$

Die letzten Gleichungen bestehen natürlich auch für den Fall, daß die betreffenden Funktionen in a nur nach x_j partiell differenzierbar sind.

Der *Beweis* verläuft völlig analog zu dem Fall einer Veränderlichen und sei deshalb dem Leser überlassen. \square

Von großer Wichtigkeit ist die Verallgemeinerung der *Kettenregel* auf mehrere Veränderliche.

Satz 17.7 (Kettenregel) *Es seien f und g verknüpfbar, f sei in a und g in $b = f(a)$ total differenzierbar. Dann ist $g \circ f$ in a total differenzierbar, und es gilt*

$$D_a(g \circ f) = D_b g \circ D_a f,$$

wobei auf der rechten Seite die Verknüpfung von linearen Abbildungen steht.

Beweis. Wir setzen

$$\begin{aligned} f(a + v) &= f(a) + L(v) + \varphi(v), \\ g(b + w) &= g(b) + M(w) + \psi(w) \end{aligned}$$

mit stetigen Linearformen L und M und

$$\lim_{v \rightarrow 0} \frac{\|\varphi(v)\|}{\|v\|} = 0, \quad \lim_{w \rightarrow 0} \frac{\|\psi(w)\|}{\|w\|} = 0.$$

Daraus ergibt sich für $F = g \circ f$:

$$\begin{aligned} F(a + v) &= g(f(a + v)) = g(b + L(v) + \varphi(v)) \\ &= g(b) + M(L(v) + \varphi(v)) + \psi(L(v) + \varphi(v)) \\ &= F(a) + (M \circ L)(v) + \Phi(v) \end{aligned}$$

mit

$$\Phi(v) = M(\varphi(v)) + \psi(L(v) + \varphi(v)).$$

Hieraus folgt wegen der Stetigkeit von M :

$$\lim_{v \rightarrow 0} \frac{\|M(\varphi(v))\|}{\|v\|} = \lim_{v \rightarrow 0} \left\| M \left(\frac{\varphi(v)}{\|v\|} \right) \right\| = 0.$$

Für den zweiten Ausdruck setzen wir $w := L(v) + \varphi(v)$. Es ist klar, daß $w \rightarrow 0$ mit $v \rightarrow 0$. Damit folgt

$$\lim_{v \rightarrow 0} \frac{\|\psi(L(v) + \varphi(v))\|}{\|v\|} = \lim_{w \rightarrow 0} \frac{\|\psi(w)\|}{\|w\|} \lim_{v \rightarrow 0} \frac{\|L(v) + \psi(v)\|}{\|v\|} = 0.$$

(Man beachte, daß $\psi(L(v) + \varphi(v)) = 0$, falls $w = L(v) + \varphi(v) = 0$.) □

Folgerung 17.8 In dem vorigen Satz seien $V \cong \mathbb{R}^n$ und $W \cong \mathbb{R}^m$ endlich-dimensional, es sei $f = {}^t(f_1, \dots, f_m)$ eine Abbildung $V \rightarrow W$, $f_j = f_j(x_1, \dots, x_n)$, und $g = g(y_1, \dots, y_m)$ sei eine Funktion. Dann gilt

$$\frac{\partial g(f_1(x_1, \dots, x_n), \dots, f_m(x_1, \dots, x_n))}{\partial x_k} = \sum_{j=1}^m \frac{\partial g}{\partial y_j}(f(x)) \cdot \frac{\partial f_j}{\partial x_k}(x).$$

Wir wollen noch eine *notwendige* Bedingung dafür ableiten, daß eine gegebene Abbildung f , die in a total differenzierbar ist, in $b = f(a)$ eine total differenzierbare Umkehrabbildung g besitzt. In diesem Fall muß mit $F = g \circ f = \text{id}$ gelten:

$$F(a) + (M \circ L)(v) + \Phi(v) = \text{id}(a + v) = a + v, \quad a = F(a),$$

also

$$(M \circ L - \text{id})(v) + \Phi(v) \equiv 0$$

und damit $M \circ L = \text{id}_V$. Entsprechend folgt $L \circ M = \text{id}_W$, so daß insbesondere $L : V \rightarrow W$ ein linearer *Isomorphismus* sein muß. Ist speziell $V = \mathbb{R}^n$, $W = \mathbb{R}^m$, so muß damit $m = n$ und die Jacobi-Matrix

$$(Df)_a := \left(\frac{\partial f_j}{\partial x_k}(a) \right)_{1 \leq j, k \leq n}$$

invertierbar sein. Wir werden später zeigen (Kapitel 25), daß diese Bedingung auch (lokal) hinreichend ist.

Satz 17.9 (Umkehrsatz) Es sei $f = {}^t(f_1, \dots, f_n) : U \rightarrow \mathbb{R}^n$, U offen in \mathbb{R}^n , stetig partiell differenzierbar, und

$$(Df)_a \in M(n \times n, \mathbb{R})$$

sei an einer Stelle $a \in U$ invertierbar. Dann gibt es eine offene Umgebung $U_0 \subset U$ von a und eine offene Umgebung U_1 von $b = f(a)$, so daß

$$f|_{U_0} : U_0 \longrightarrow U_1$$

bijektiv und die Umkehrabbildung

$$(f|_{U_0})^{-1} : U_1 \longrightarrow U_0$$

stetig differenzierbar ist.

Wir wollen uns jetzt mit *höheren Ableitungen* und der *Taylorformel* in mehreren Veränderlichen beschäftigen. Wir setzen

$$\frac{\partial^2 f}{\partial x_j \partial x_k} := \frac{\partial}{\partial x_j} \left(\frac{\partial}{\partial x_k} \right),$$

sofern diese Ableitung (an einer Stelle) existiert, und schreiben dafür auch

$$\partial_{jk}f \quad \text{oder} \quad D^j D^k f \quad \text{oder} \quad f_{x_j x_k} \quad \text{oder} \quad f_{,jk}.$$

Man kann einfache Beispiele dafür angeben, daß zweite partielle Ableitungen *nicht zu vertauschen brauchen*: I. A. ist $\partial_{jk}f \neq \partial_{kj}f$.

Beispiel. Die Funktion $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ sei definiert durch

$$f(x, y) := \begin{cases} xy \frac{x^2 - y^2}{x^2 + y^2} & \text{für } (x, y) \neq (0, 0), \\ 0 & \text{für } (x, y) = (0, 0). \end{cases}$$

Dann ist f überall zweimal partiell differenzierbar; es gilt aber

$$\frac{\partial^2 f}{\partial x \partial y}(0, 0) = 1 \neq -1 = \frac{\partial^2 f}{\partial y \partial x}(0, 0).$$

Unter recht einfachen Voraussetzungen tritt dieses unangenehme Phänomen aber nicht auf.

Satz 17.10 *Es sei $U \subset \mathbb{R}^n$ offen, $f : U \rightarrow \mathbb{R}$ sei stetig partiell differenzierbar, und es existiere $\partial^2 f / \partial x_j \partial x_k$. Ferner sei diese zweite Ableitung an einer Stelle $a \in U$ stetig. Dann existiert auch $\partial^2 f / \partial x_k \partial x_j$ an der Stelle a , und es gilt*

$$\frac{\partial^2 f}{\partial x_k \partial x_j}(a) = \frac{\partial^2 f}{\partial x_j \partial x_k}(a).$$

Beweis. Ohne Einschränkung sei $n = 2$ und $(j, k) = (1, 2)$; ferner sei $a = 0$ und $Q = [-\delta, \delta] \subset U$. Betrachte die Funktion (wir schreiben $(x_1, x_2) = (x, y)$):

$$g(x, y) = f(x, y) - f(x, 0), \quad (x, y) \in Q.$$

Dann ist nach dem 1. Mittelwertsatz

$$\begin{aligned} F(x, y) &:= f(x, y) - f(x, 0) - f(0, y) + f(0, 0) = g(x, y) - g(0, y) = x \frac{\partial g}{\partial x}(\xi, y) \\ &= x \left(\frac{\partial f}{\partial x}(\xi, y) - \frac{\partial f}{\partial x}(\xi, 0) \right) = xy \frac{\partial^2 f}{\partial y \partial x}(\xi, \eta) \end{aligned}$$

mit $\xi = \xi_{x,y}$ zwischen 0 und x und $\eta = \eta_{x,y}$ zwischen 0 und y . Nach Voraussetzung gibt es zu $\varepsilon > 0$ ein $\delta > 0$ (evtl. kleiner als das schon gewählte) mit

$$\left| \frac{\partial^2 f}{\partial y \partial x}(x, y) - \frac{\partial^2 f}{\partial y \partial x}(0, 0) \right| < \varepsilon$$

für $|x| < \delta$, $|y| < \delta$. Mit der obigen Gleichung ist dann

$$\left| \frac{F(x, y)}{xy} - \frac{\partial^2 f}{\partial y \partial x}(0, 0) \right| < \varepsilon, \quad |x| < \delta, \quad |y| < \delta.$$

Nun ist

$$\frac{F(x, y)}{xy} = \frac{1}{x} \left\{ \frac{f(x, y) - f(x, 0)}{y} - \frac{f(0, y) - f(0, 0)}{y} \right\},$$

und dieser Ausdruck geht mit $y \rightarrow 0$ gegen $\frac{1}{x} \left(\frac{\partial f}{\partial y}(x, 0) - \frac{\partial f}{\partial y}(0, 0) \right)$. Also ist für alle hinreichend kleinen x :

$$\left| \frac{(\partial f / \partial y)(x, 0) - (\partial f / \partial y)(0, 0)}{x} - \frac{\partial^2 f}{\partial y \partial x}(0, 0) \right| \leq \varepsilon$$

und damit

$$\frac{\partial^2 f}{\partial x \partial y}(0, 0) = \frac{\partial^2 f}{\partial y \partial x}(0, 0). \quad \square$$

Definition. Sei $U \subset \mathbb{R}^n$ offen. Eine Funktion $f : U \rightarrow \mathbb{R}$ heißt k -mal stetig differenzierbar, wenn alle Ableitungen

$$\partial_{i_1} \cdots \partial_{i_k} f, \quad i_1, \dots, i_k \in \{1, \dots, n\},$$

in U existieren und stetig sind. Wir schreiben dann $f \in \mathcal{C}^k(U)$. Nach dem vorigen Satz ist dann jede Ableitung $\partial_{i_1} \cdots \partial_{i_k} f$ unabhängig von der Reihenfolge der Differentiationen; m. a. W.: Es gilt stets

$$\partial_{i_1} \cdots \partial_{i_k} f = D^{\alpha_1} D^{\alpha_2} \cdots D^{\alpha_n} f,$$

wobei α_j die Anzahl der i_k bezeichnet, die gleich j sind, $j = 1, \dots, n$. Ferner setzt man

$$\mathcal{C}^\infty(U) = \bigcap_{k=0}^{\infty} \mathcal{C}^k(U).$$

Funktionen in diesem Vektorraum heißen *unendlich oft partiell differenzierbar*.

Man kann leicht die *Taylor-Formel* auf mehrere Veränderliche übertragen, wobei letztendlich nur Schreibarbeit zu bewältigen ist. Man betrachtet $f \in \mathcal{C}^{k+1}(U)$ und Punkte $a, x \in U$, deren *Verbindungsstrecke* in U verläuft. Dies ist zum Beispiel der Fall, wenn U *konvex*, also z. B. eine Kugel ist, oder wenn U *sternförmig* bzgl. a ist. Durch Betrachtung von $F(t) := f(a + th)$, $h = x - a$, $t \in [0, 1]$, kann man das ganze Problem auf den Fall einer einzigen Variablen t zurückführen. Wir setzen im folgenden

$$\alpha! := \alpha_1! \cdots \alpha_n!, \quad |\alpha| := \alpha_1 + \cdots + \alpha_n \quad \text{und} \quad D^\alpha f := \partial_1^{\alpha_1} \cdots \partial_n^{\alpha_n} f \quad \text{für} \quad \alpha := (\alpha_1, \dots, \alpha_n) \in \mathbb{N}^n,$$

und

$$h^\alpha = h_1^{\alpha_1} \cdots h_n^{\alpha_n} \quad \text{für} \quad h = (h_1, \dots, h_n) \in \mathbb{R}^n.$$

Mit diesen Bezeichnungen gilt der folgende

Satz 17.11 *Zu jedem $x = a + h$ mit $a + th \in U$, $t \in [0, 1]$, existiert ein $\vartheta \in [0, 1]$ mit*

$$f(x) = \sum_{|\alpha| \leq k} \frac{D^\alpha f(a)}{\alpha!} h^\alpha + \sum_{|\alpha| = k+1} \frac{D^\alpha f(a + \vartheta h)}{\alpha!} h^\alpha.$$

Insbesondere ist

$$f(x) = \sum_{|\alpha| \leq k+1} \frac{D^\alpha f(a)}{\alpha!} h^\alpha + o(\|h\|^{k+1}).$$

Hierbei steht $o(\|h\|^{k+1})$ für eine Funktion $\varphi(h)$ mit

$$\lim_{h \rightarrow 0} \frac{\varphi(h)}{\|h\|^{k+1}} = 0.$$

Beweis. a) Wir betrachten die Funktion

$$F(t) = f(a + th), \quad t \in [0, 1],$$

und zeigen, daß sie $(k+1)$ -mal stetig differenzierbar ist mit

$$\begin{aligned} \frac{d^\ell F}{dt^\ell}(t) &= \sum_{j_1, \dots, j_\ell=1}^n \partial_{j_\ell} \cdots \partial_{j_1} f(a + th) h_{j_1} \cdots h_{j_\ell}, \\ &= \sum_{|\alpha|=\ell} \frac{\ell!}{\alpha!} D^\alpha f(a + th) h^\alpha, \quad 0 \leq \ell \leq k+1. \end{aligned}$$

Die erste Formel erhält man leicht durch Induktion nach $\ell \geq 1$, wobei der Fall $\ell = 1$ gerade die Kettenregel ist:

$$\frac{dF}{dt} = \frac{d}{dt} f(a_1 + th_1, \dots, a_n + th_n) = \sum_{j=1}^n \partial_j f(a + th) h_j.$$

Ist die Formel nun für $\ell \geq 1$ schon bewiesen, so ergibt sich wiederum mit der Kettenregel

$$\begin{aligned} \frac{d^{\ell+1}F}{dt^{\ell+1}}(t) &= \frac{d}{dt} \left(\sum_{j_1, \dots, j_\ell=1}^n \partial_{j_\ell} \dots \partial_{j_1} f(a + th) \right) h_{j_1} \dots h_{j_\ell} \\ &= \sum_{j=1}^n \left(\sum_{j_1, \dots, j_\ell=1}^n \partial_j \partial_{j_\ell} \dots \partial_{j_1} f(a + th) h_{j_1} \dots h_{j_\ell} \right) h_j \\ &= \sum_{j_1, \dots, j_{\ell+1}=1}^n \partial_{j_{\ell+1}} \dots \partial_{j_1} f(a + th) h_{j_1} \dots h_{j_{\ell+1}}. \end{aligned}$$

Kommt nun unter den Indizes (j_1, \dots, j_ℓ) der Index 1 genau α_1 -mal vor, der Index 2 genau α_2 -mal etc., so ist wegen des vorigen Satzes

$$\partial_{j_\ell} \dots \partial_{j_1} f(a + th) h_{j_1} \dots h_{j_\ell} = D^\alpha f(a + th) h^\alpha.$$

Es gibt aber genau $\ell!/\alpha_1! \dots \alpha_n!$ ℓ -tupel (j_1, \dots, j_ℓ) von Zahlen $1 \leq j_\lambda \leq n$, bei denen die Zahl ν genau ν -mal vorkommt.

b) Wegen der Taylor-Formel in einer Veränderlichen gibt es ein $\vartheta \in (0, 1)$ mit

$$f(a + h) = F(1) = \sum_{\ell=0}^k \frac{F^{(\ell)}(0)}{\ell!} + \frac{F^{(\ell+1)}(\vartheta)}{(\ell+1)!} = \sum_{|\alpha| \leq k} \frac{D^\alpha f(a)}{\alpha!} h^\alpha + \sum_{|\alpha|=k+1} \frac{D^\alpha f(a + \vartheta h)}{\alpha!} h^\alpha.$$

c) Nach dem Vorstehenden ist

$$f(a + h) = \sum_{|\alpha| \leq k+1} \frac{D^\alpha f(a)}{\alpha!} h^\alpha + \sum_{|\alpha|=k+1} r_\alpha(h) h^\alpha$$

mit

$$r_\alpha(h) = \frac{1}{\alpha!} (D^\alpha f(a + \vartheta h) - D^\alpha f(a)).$$

Wegen der Stetigkeit von $D^\alpha f$ an der Stelle a für $|\alpha| \leq k+1$ ist $\lim_{h \rightarrow 0} r_\alpha(h) = 0$. Nun ist in der Maximumnorm von \mathbb{R}^n offensichtlich für $|\alpha| = k+1$:

$$\frac{|h^\alpha|}{\|h\|_\infty^{k+1}} = \frac{|h_1|^{\alpha_1} \dots |h_n|^{\alpha_n}}{\|h\|_\infty^{\alpha_1} \dots \|h\|_\infty^{\alpha_n}} \leq 1.$$

Somit ergibt sich

$$\lim_{h \rightarrow 0} \frac{\varphi(h)}{\|h\|_\infty^{k+1}} = 0, \quad \text{also} \quad \varphi = o(\|h\|_\infty^{k+1})$$

mit

$$\varphi(h) = \sum_{|\alpha|=k+1} r_\alpha(h) h^\alpha. \quad \square$$

Für Anwendungen beschränken wir uns auf den Fall $k = 1$. Es ist klar, daß in $a \in U$ höchstens dann ein (lokales) Extremum vorliegen kann, wenn alle partiellen Ableitungen $f_{,j}$ in a verschwinden,

denn dann muß ja auch die Funktion $f(a + te_j)$ ein lokales Extremum bei $t = 0$ besitzen.

Definition. Ist $f : U \rightarrow \mathbb{R}$ in $a \in U$ partiell differenzierbar, so heißt der Vektor

$$\text{grad}_a f = \left(\frac{\partial f}{\partial x_1}(a), \dots, \frac{\partial f}{\partial x_n}(a) \right)$$

der *Gradient* der Funktion f an der Stelle a . Manchmal schreiben wir auch $\text{grad } f(a)$ oder Ähnliches.

Bemerkung. Weiter unten werden wir den Gradienten einer Funktion auch *geometrisch* deuten.

Hier können wir zunächst anmerken:

Lemma 17.12 *Eine notwendige Bedingung für ein lokales Extremum einer partiell differenzierbaren Funktion $f : U \rightarrow \mathbb{R}$ an der Stelle a in der offenen Menge U ist das Verschwinden des Gradienten :*

$$\text{grad}_a f = 0.$$

Für Funktionen in einer Veränderlichen konnten wir mit Hilfe der Taylor-Formel hinreichende Bedingungen dafür angeben, wann an einer Stelle mit verschwindender Ableitung tatsächlich ein lokales Minimum oder Maximum vorliegt. Auch für Funktionen in mehreren Veränderlichen gibt es solche Kriterien, die die entsprechenden Bedingungen in einer Veränderlichen verallgemeinern. Nach dem Vorigen hat man nämlich an Stellen a mit verschwindendem Gradienten für kleine h die Darstellung

$$f(a + h) = f(a) + \frac{1}{2} \sum_{j,k=1}^n \frac{\partial^2 f}{\partial x_j \partial x_k}(a) h_j h_k + o(\|h\|^2).$$

Definition. Die (symmetrische) quadratische Matrix

$$H_a = \left(\frac{\partial^2 f}{\partial x_j \partial x_k}(a) \right)_{j,k=1,\dots,n}$$

heißt die *HESSESche Matrix* von f an der Stelle a . Die zugehörige *quadratische Form*

$$h \longmapsto H(h) := \langle h, H_a h \rangle = \sum_{j,k=1}^n \frac{\partial^2 f}{\partial x_j \partial x_k}(a) h_j h_k$$

heißt die *Hessesche Form* von f an der Stelle a .

Definition. Es sei $S = (a_{jk}) \in M(n \times n, \mathbb{R})$ *symmetrisch*, d. h. $a_{kj} = a_{jk}$ für alle $j, k = 1, \dots, n$. Die zugehörige *quadratische Form* $q(h) = q_S(h) := \sum a_{j,k} h_j h_k$ heißt

- a) *positiv definit*, wenn $q(h) > 0$, $h \neq 0$;
- b) *positiv semidefinit*, wenn $q(h) \geq 0$, $h \in \mathbb{R}^n$;
- c) *indefinit*, wenn es h' und h'' gibt mit $q(h') > 0 > q(h'')$.

Ein nützliches Kriterium ist das folgende:

Satz 17.13 *Für die quadratische Form $q = q_S$ sind die folgenden Aussagen äquivalent :*

- i) $q = q_S$ ist *positiv (semi-) definit* ;
- ii) die *Eigenwerte der Matrix S sind (sämtlich reell und) positiv (nicht negativ)* ;
- iii) alle *Hauptunterdeterminanten der symmetrischen Matrix S sind positiv (nicht negativ)*.

Zum *Beweis* siehe das Ende dieses Kapitels im Anschluß an den Beweis des Satzes über die *Hauptachsentransformation*. \square

Bemerkung. Das Kriterium iii) wird oft HURWITZ zugeschrieben, findet sich aber schon in einer früheren Arbeit von JACOBI.

Satz 17.14 *Es sei $f \in C^2(U)$, $a \in U$, und $\text{grad}_a f = 0$. Dann gilt :*

- Ist die Hessesche Form $H = H_a$ positiv definit, so besitzt f in a ein (lokales) isoliertes Minimum.*
- Ist H negativ definit, so besitzt f in a ein (lokales) isoliertes Maximum.*
- Ist H indefinit, so besitzt f in a kein lokales Extremum.*

Beweis. Selbstverständlich folgt b) aus a) durch Übergang zu $-f$. Wir zeigen also a). Im folgenden bezeichnet $\|\cdot\|$ die euklidische Norm von \mathbb{R}^n . Nach Voraussetzung gilt

$$h(a+h) = f(a) + \frac{1}{2} \langle h, Hh \rangle + \varphi(h) \quad \text{mit} \quad \varphi(h) = o(\|h\|^2).$$

Somit gibt es zu jedem $\varepsilon > 0$ ein $\delta > 0$, so daß aus $\|h\| < \delta$ die Ungleichung $|\varphi(h)| < \varepsilon \|h\|^2$ folgt. Ist nun H positiv definit, so existiert

$$\alpha := \inf \{ \langle h, Hh \rangle : h \in S^{n-1} \} > 0.$$

Für $h \neq 0$ sei $\lambda := 1/\|h\|$ und damit $g := \lambda h \in S^{n-1}$. Dann ist

$$\langle h, Hh \rangle = \frac{1}{\lambda^2} \langle g, Hg \rangle \geq \frac{1}{\lambda^2} \alpha = \alpha \|h\|^2.$$

(Diese Ungleichung gilt natürlich auch für $h = 0$). Wähle jetzt $\delta = \delta(\alpha/4)$, so daß

$$|\varphi(h)| < \frac{\alpha}{4} \|h\|^2 \quad \text{für} \quad \|h\| < \delta.$$

Da andererseits $\frac{1}{2} \langle h, Hh \rangle \geq \frac{\alpha}{2} \|h\|^2$, ergibt sich

$$f(a+h) \geq f(a) + \frac{\alpha}{4} \|h\|^2 > f(a) \quad \text{für} \quad 0 < \|h\| \leq \delta.$$

c) Für indefinites H sei $\xi \in \mathbb{R}^n \setminus \{0\}$ ein Vektor mit $\langle \xi, H\xi \rangle = \alpha > 0$. Für kleine reelle Zahlen t ist dann

$$f(a+t\xi) = f(a) + \frac{t^2}{2} \langle \xi, H\xi \rangle + \varphi(t\xi) = f(a) + \frac{\alpha}{2} t^2 + \varphi(t\xi).$$

Ist t hinreichend klein, so ist $|\varphi(t\xi)| \leq \frac{\alpha}{4} t^2$ und folglich $f(a+t\xi) > f(a)$. Entsprechend erhält man für $\eta \in \mathbb{R}^n$ mit $\langle \eta, H\eta \rangle < 0$, daß $f(a+t\eta) < f(a)$ für kleine t . \square

Bemerkung. Ist H nur *semidefinit*, so kann man keine Aussagen über f nahe a mit $\text{grad}_a f = 0$ machen. Es ist z. B.

$$H = \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix}$$

positiv semidefinit für die Funktionen

$$f(x, y) = x^2, \quad x^2 + y^3, \quad x^2 + y^4$$

an der Stelle $a = 0 = (0, 0) \in \mathbb{R}^2$. Im ersten Fall hat man ein (nicht isoliertes) Minimum in a , im zweiten Fall kein Extremum und im dritten Fall ein isoliertes Minimum.

Wir kommen nun zu der oben schon angekündigten *geometrischen Deutung* des Gradienten einer Funktion f in n reellen Veränderlichen. Als erstes notieren wir:

Satz 17.15 Die Funktion $f : U \rightarrow \mathbb{R}$, $U \subset \mathbb{R}^n$ offen, sei in dem Punkt $a \in U$ stetig partiell differenzierbar. Dann ist der Gradient $\text{grad}_a f$ der eindeutig bestimmte Vektor $\gamma \in \mathbb{R}^n$, so daß für alle Vektoren $v \in \mathbb{R}^n$ die Richtungsableitung $\partial_v f(a)$ von f in a geschrieben werden kann in der Gestalt

$$\partial_v f(a) = \langle \gamma, v \rangle .$$

Beweis. Eine Richtung haben wir schon gezeigt: Ist f in a total differenzierbar, so besitzt f dort alle Richtungsableitungen, und es ist

$$\partial_v f(a) = \sum_{j=1}^n \frac{\partial f}{\partial x_j}(a) v_j = \langle \text{grad}_a f, v \rangle \quad \text{für alle } v = (v_1, \dots, v_n) \in \mathbb{R}^n .$$

Ist umgekehrt $\partial_v f(a) = \langle \gamma, v \rangle$ für alle $v \in \mathbb{R}^n$, so gilt speziell mit $v := e_j$: $\partial_j f(a) = \langle \gamma, e_j \rangle = \gamma_j$, also $\gamma = \text{grad}_a f$. \square

Als nächstes müssen wir ganz allgemein *orientierte Winkel* in *euklidischen Vektorräumen* einführen²⁴. Wir erinnern dazu zunächst daran, daß wegen der Cauchy–Schwarzschen Ungleichung in euklidischen Vektorräumen²⁵ V mit Skalarprodukt $\langle \cdot, \cdot \rangle$ für zwei von Null verschiedene Vektoren $v, w \in V$ die Zahl

$$c(v, w) := \frac{\langle v, w \rangle}{\|v\| \|w\|}$$

dem Betrage nach stets ≤ 1 ist. Ist $A \in \text{End}(V)$ eine lineare Abbildung, die das Skalarprodukt invariant läßt, also eine *orthogonale Abbildung* oder *euklidische Drehung um den Nullpunkt*, so bleibt auch $c(v, w)$ invariant: $c(Av, Aw) = c(v, w)$. Da offensichtlich

$$c(v, w) = c\left(\frac{v}{\|v\|}, \frac{w}{\|w\|}\right),$$

können wir ferner annehmen, daß v und w die Länge 1 besitzen. Durch eine geeignete euklidische Bewegung A läßt sich offensichtlich weiter erreichen, daß $v = e_1$ ist und w in der von e_1 und e_2 aufgespannten Ebene liegt. Ist nun φ , $0 \leq \varphi \leq \pi$, der (orientierte) Winkel zwischen der e_1 -Achse und w , also $w = (\cos \varphi) e_1 + (\sin \varphi) e_2$, so ist $\langle v, w \rangle = \cos \varphi$ wegen $\langle e_j, e_k \rangle = \delta_{jk}$.

Lemma 17.16 *Durch*

$$\cos \varphi := \frac{\langle v, w \rangle}{\|v\| \|w\|}, \quad 0 \leq \varphi \leq \pi,$$

wird je zwei von Null verschiedenen Vektoren v, w eindeutig ihr orientierter Zwischenwinkel φ zugeordnet. Es gilt

$$\langle v, w \rangle = (\cos \varphi) \|v\| \|w\| .$$

Bezeichnen wir nun mit $\varphi := \varphi_v$ diesen Winkel zwischen v und dem Gradienten γ von f in a , $0 \leq \varphi \leq \pi$, so ist nach den obigen Darlegungen

$$\langle \gamma, v \rangle = (\cos \varphi) \|v\| \|\gamma\|$$

und damit $\partial_v f(a)$ am größten, wenn $\cos \varphi = 1$ ist und folglich v und $\text{grad}_a f$ in dieselbe Richtung zeigen.

Folgerung 17.17 Die Richtung des Gradienten $\text{grad}_a f$ ist die Richtung des steilsten Anstiegs der Funktion f im Punkte a .

²⁴Zu diesen Bemerkungen vergleiche auch die ausführlicheren Darlegungen im Anhang zu Kapitel 19.

²⁵Siehe loc. cit.

Man kann dieses Ergebnis auch noch anders deuten. Auf einer sogenannten *Niveau-Menge*

$$N_f(c) := f^{-1}(c) = \{x \in U : f(x) = c\}$$

steigt die Funktion f überhaupt nicht an. Ist speziell $n = 2$, so sind diese Niveau-Mengen im Allgemeinen *Kurven*, die man auch als *Höhenlinien* bezeichnet. Die Funktion f läßt sich dann aus ihrem *Höhenportrait*, also einer Landkarte mit Höhenlinien, rekonstruieren.

Nun weiß jeder Bergwanderer, daß der *steilste Anstieg* auf den Berg *senkrecht zu den Höhenlinien* verläuft. Wir können diese alltägliche Erfahrung mathematisch völlig exakt formulieren und beweisen.

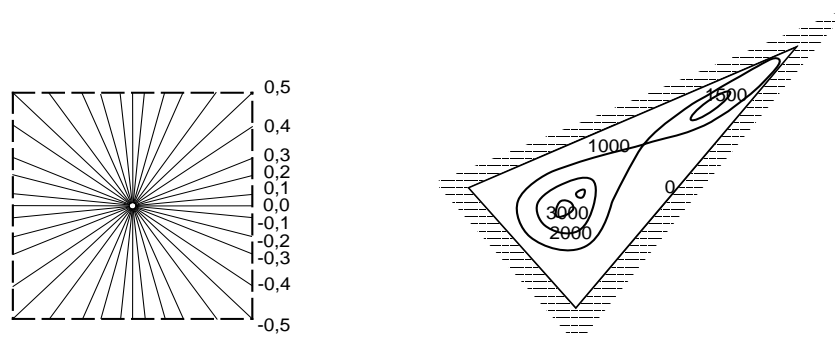
Lemma 17.18 *Es sei $f : U \rightarrow \mathbb{R}$ eine in dem Punkt $a \in U \subset \mathbb{R}^n$ total differenzierbare Funktion, und $\alpha : I = (-\varepsilon, \varepsilon) \rightarrow U$ sei eine differenzierbare Kurve mit $\alpha(0) = a$, deren Bild $\alpha(I)$ ganz in der Höhenlinie $N_f(f(a))$ liegt. Dann gilt*

$$\langle \text{grad}_a f, \alpha'(0) \rangle = 0.$$

Beweis. Die Bedingung $\alpha(t) \in N_f(f(a))$ ist gleichbedeutend damit, daß die Zusammensetzung $f \circ \alpha$ identisch $f(a)$ auf dem Intervall I ist. Der Rest ist nur noch eine direkte Anwendung der Kettenregel:

$$\langle \text{grad}_a f, \alpha'(0) \rangle = \sum_{j=1}^n \frac{\partial f}{\partial x_j}(a) \alpha'_j(0) = \left. \frac{d}{dt} (f \circ \alpha) \right|_{t=0} = 0. \quad \square$$

Anbei geben wir zwei Beispiele, wobei das erste der Figur 2 entspricht²⁶.



Figur 17.3

Zum Abschluß dieses Kapitels wollen wir uns mit einer Verfeinerung des notwendigen Kriteriums für lokale Extrema von Funktionen in mehreren Veränderlichen beschäftigen, dem sogenannten Kriterium für ein (*lokales*) *Extremum mit Nebenbedingung*. Wir behandeln hier nur den einfachsten Fall und stellen das allgemeine Resultat zurück, bis wir es im Rahmen der Theorie der *Untermannigfaltigkeiten* von \mathbb{R}^n (siehe Kapitel 25) in seine natürliche Form bringen und vollständig erklären können.

Ein Beispiel dieser Art haben wir schon früher behandelt: *Man suche das Rechteck größten Flächeninhalts mit vorgegebenem Umfang $2L$* . Mit anderen Worten: Man maximiere die Funktion $F(x, y) := xy$ auf $\mathbb{R}_+^* \times \mathbb{R}_+^*$ unter der *Nebenbedingung* $f(x, y) = 2L = \text{const.}$, wobei $f(x, y) := x + y$. Die Lösung des *speziellen* Problems ist sehr einfach: Man *löst die Gleichung $f(x, y) = 2L$ nach einer der Variablen auf*, z. B. also $y = 2L - x$, und setzt in die Funktion F ein. Dies liefert eine Funktion in einer einzigen Variablen, deren Maxima und Minima man mit der Differentialrechnung der Funktionen in einer Variablen zu bestimmen versuchen kann.

Im Allgemeinfall haben wir es also mit zwei (stetig differenzierbaren) Funktionen f, F auf einer offenen Menge $U \subset \mathbb{R}^n$ zu tun.

²⁶In dem zweiten sehr schematisierten Beispiel werden einige Leser vielleicht die höchste Erhebung in Spanien erkennen.

Definition. Ein Punkt $a \in U$ heißt ein *lokales Maximum* von F unter der Nebenbedingung $f = 0$, wenn es eine Umgebung V von a in U gibt, so daß $F(x) \leq F(a)$ für alle $x \in V$ mit der zusätzlichen Bedingung $x \in N_f(0)$, d. h. $f(x) = 0$, gilt. Analog erklärt man die Begriffe des *lokalen Minimums* und zusammenfassend des *lokalen Extremums* von F an der Stelle a unter der Nebenbedingung $f = 0$.

Das Problem der Bestimmung einer notwendigen Bedingung für die Existenz eines solchen Extremums in einem Punkt a läßt sich auf den „absoluten“ Fall der Bestimmung von lokalen Extrema einer Funktion in $n - 1$ Variablen zurückführen, wenn man die Gleichung $f(x_1, \dots, x_n) = 0$ nach einer der Variablen x_1, \dots, x_n nahe a auflösen kann, wenn es also z. B. offene Mengen $V' \subset \mathbb{R}^{n-1}$, $V'' \subset \mathbb{R}$ mit

$$a \in V' \times V'' \subset U$$

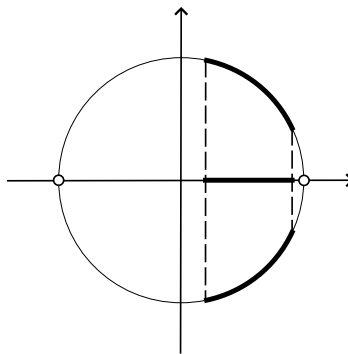
und eine Funktion $\varphi : V'' \rightarrow V'$ gibt, so daß

$$\{x \in V' \times V'' : f(x) = 0\} = \{x = (x', x'') \in \mathbb{R}^{n-1} \times \mathbb{R} : x'' = \varphi(x')\},$$

wenn also die Nullstellenmenge von f in der Produktumgebung $V' \times V''$ von a der Graph der Funktion φ ist. Denn dann besitzt F genau dann ein lokales Extremum in $a := (a', a'') \in \mathbb{R}^{n-1} \times \mathbb{R}$ unter der Nebenbedingung $f = 0$, wenn die zusammengesetzte Funktion $F(x_1, \dots, x_{n-1}, \varphi(x_1, \dots, x_{n-1}))$ in $n - 1$ Veränderlichen ein lokales Extremum (ohne Nebenbedingung) in dem Punkt $a' \in V' \subset \mathbb{R}^{n-1}$ annimmt.

Erstaunlicherweise braucht man unter einer Zusatzvoraussetzung weder nach der Auflösung der impliziten Gleichung $f = 0$ zu suchen noch diese, wenn man sie kennt, einzusetzen. Der Schlüssel des Geheimnisses ist der leicht aus dem Umkehrsatz deduzierbare Satz über implizite Funktionen (siehe Kapitel 25), der besagt, daß man die Funktion f lokal bei einem Punkt $a \in N_f(0)$ immer dann nach x_j auflösen kann, wenn $\partial f / \partial x_j(a) \neq 0$.

Beispiel. Wir betrachten die Funktion $f(x_1, \dots, x_n) := x_1^2 + \dots + x_n^2 - 1$. Das Nullstellengebilde dieser Funktion ist die Einheitskugel S^{n-1} , die sich an allen Stellen $x' := (x_1, \dots, x_{n-1}) \in \mathbb{R}^{n-1}$ mit $x_1^2 + \dots + x_{n-1}^2 < 1$ als Graph der (differenzierbaren) Funktion $\varphi(x') := \sqrt{1 - (x_1^2 + \dots + x_{n-1}^2)}$ oder der Funktion $-\varphi$ darstellen läßt, und offensichtlich ist eine Auflösung von $f = 0$ nach x_n nur an diesen Stellen möglich. Andererseits verschwindet die Ableitung $\partial f / \partial x_n = 2x_n$ auf S^{n-1} genau in den Punkten $(x_1, \dots, x_{n-1}, 0)$ mit $x_1^2 + \dots + x_{n-1}^2 = 1$.



Figur 17.4

Eine (lokale) Auflösung der Gleichung $f = 0$ in der Nähe von $a \in U$ mit $f(a) = 0$ ist also immer dann möglich, wenn $\text{grad}_a f \neq 0$. Unter dieser Voraussetzung erhalten wir aber auch eine notwendige Bedingung für die lokale Existenz von Extrema unter der Nebenbedingung $f = 0$, ohne tatsächlich die Auflösung vornehmen zu müssen.

Satz 17.19 Es seien $F, f : U \rightarrow \mathbb{R}$ stetig partiell differenzierbare Funktionen, und $a \in U$ sei eine Stelle mit $\text{grad}_a f \neq 0$. Besitzt dann F ein lokales Extremum in a unter der Nebenbedingung $f = 0$, so ist notwendig

$$\text{grad}_a F = \lambda \text{grad}_a f$$

mit einer Konstanten $\lambda \in \mathbb{R}$.

Definition. Man nennt die (a priori unbekannt) Zahl λ einen LAGRANGE-Multiplikator.

Beispiel. Wir kommen noch einmal auf das Beispiel $F(x, y) = xy$ und $f(x, y) = x + y - 2L$ zurück. Hier ist $\text{grad}_a f = (1, 1) \neq 0$ an allen Stellen des Definitionsbereichs. Die Lagrange-Bedingung für ein lokales Extremum der Funktion F unter der Nebenbedingung $f = 0$ liefert dann die Bedingung

$$(a_2, a_1) = \text{grad}_a F = \lambda \text{grad}_a f = (\lambda, \lambda),$$

und obwohl wir den Lagrange-Faktor λ nicht kennen, folgt hieraus schon $(a_2, a_1) = (\lambda, \lambda)$, also $a_1 = \lambda = a_2$, und aus der Nebenbedingung ergibt sich das uns schon bekannte Resultat $a_1 = a_2 = L$.

Beweis von Satz 19. Wir können ohne Einschränkung annehmen, daß $\partial f / \partial x_n$ an der Stelle $a = (a', a'')$ von Null verschieden ist und die Gleichung $f(x_1, \dots, x_n) = 0$ nahe a nach $x_n = \varphi(x_1, \dots, x_{n-1})$ aufgelöst werden kann. Dann ist $f(x_1, \dots, x_{n-1}, \varphi(x_1, \dots, x_{n-1})) = 0$ in einer Umgebung des Punktes $a' \in \mathbb{R}^{n-1}$, und aus der Kettenregel folgt durch Differentiation nach x_j , $j = 1, \dots, n-1$:

$$\frac{\partial f}{\partial x_j}(a) + \frac{\partial f}{\partial x_n}(a) \frac{\partial \varphi}{\partial x_j}(a') = 0.$$

Auf der anderen Seite besitzt die zusammengesetzte Funktion

$$F(x_1, \dots, x_{n-1}, \varphi(x_1, \dots, x_{n-1}))$$

ein lokales Extremum in a' . Damit muß der Gradient dieser Funktion in a' verschwinden, und dies liefert erneut mit der Kettenregel die Gleichungen

$$\frac{\partial F}{\partial x_j}(a) + \frac{\partial F}{\partial x_n}(a) \frac{\partial \varphi}{\partial x_j}(a') = 0, \quad j = 1, \dots, n-1.$$

Mit $\lambda := \frac{\partial F}{\partial x_n}(a) \left(\frac{\partial \varphi}{\partial x_n}(a') \right)^{-1}$ folgt aus diesen Gleichungen

$$\frac{\partial F}{\partial x_j}(a) = \lambda \frac{\partial f}{\partial x_j}(a), \quad j = 1, \dots, n,$$

also $\text{grad}_a F = \lambda \text{grad}_a f$. □

Bemerkung. Unter der Bedingung $\text{grad}_a f \neq 0$ ist die Menge $N := N_f(f(a))$ tatsächlich eine $(n-1)$ -dimensionale Untermannigfaltigkeit von \mathbb{R}^n (zu diesem Begriff siehe Kapitel 25). Insbesondere besitzt sie einen Tangentialraum $T_{N,a}$, der als Vektorraum erzeugt wird von den partiellen Ableitungen

$$\frac{\partial \varphi}{\partial x_1}(a'), \dots, \frac{\partial \varphi}{\partial x_{n-1}}(a'),$$

wenn $\varphi(x')$, $x' \in U' \subset \mathbb{R}^{n-1}$, eine Auflösung der Niveaumenge N z. B. nach der Variablen x_n bezeichnet. Die Bezeichnung Tangentialraum verdient sich $T_{N,a}$ zu Recht aufgrund der Tatsache, daß er genau aus allen Tangentialvektoren $\alpha'(0)$ von Kurven $\alpha : (-\varepsilon, \varepsilon) \rightarrow N$ mit $\alpha(0) = a$ besteht. (Zu allen diesen Behauptungen vgl. loc. cit.). Hiermit kann man Lemma 18 umformulieren zu

Lemma 17.20 *Es sei $f : U \rightarrow \mathbb{R}$ eine in dem Punkt $a \in U \subset \mathbb{R}^n$ total differenzierbare Funktion mit $\text{grad}_a f \neq 0$. Dann steht der Gradient $\text{grad}_a f$ senkrecht auf dem Tangentialraum $T_{N,a}$ der Höhenlinie $N = N_f(f(a))$.*

Als besonderen Höhepunkt dieses letzten Abschnitts deduzieren wir einen zentralen Satz der linearen Algebra: den *Satz über die Hauptachsentransformation symmetrischer Matrizen* bzw. *quadratischer Formen*, samt einiger nützlicher Folgerungen, die wir im Zusammenhang mit der Bestimmung lokaler Extrema von Funktionen in mehreren Veränderlichen schon erwähnt hatten.

Wir erinnern daran, daß eine quadratische Form auf \mathbb{R}^n eine Funktion des Typs

$$q(x) := \sum_{j,k=1}^n a_{jk} x_j x_k$$

ist. Nach Vertauschung der Indizes j, k und Anwendung des Kommutativ-Gesetzes erhält man auch

$$q(x) = \sum_{j,k=1}^n a_{kj} x_j x_k = \frac{1}{2} \sum_{j,k=1}^n (a_{jk} + a_{kj}) x_j x_k = \sum_{j,k=1}^n \alpha_{jk} x_j x_k$$

mit $\alpha_{jk} := (1/2)(a_{jk} + a_{kj}) = \alpha_{kj}$. Wir können daher und werden von vornherein annehmen, daß die Matrix $S := (a_{jk})$ *symmetrisch* ist. Wir schreiben von nun an

$$q(x) = q_S(x) := \langle x, Sx \rangle.$$

Ist umgekehrt $S \in M(n \times n, \mathbb{R})$ symmetrisch, so ist

$$\langle x, y \rangle_S := {}^t x S y$$

eine *symmetrische Bilinearform* auf \mathbb{R}^n und

$$q(x) = q_S(x) = \langle x, x \rangle_S$$

die zugehörige quadratische Form. Mit dieser Beziehung kann man allgemein quadratische Formen q auf reellen Vektorräumen mit Hilfe von symmetrischen Bilinearformen b erklären durch $q(x) = b(x, x)$, wobei umgekehrt die quadratische Form q ihrerseits die sie definierende symmetrische Bilinearform b bestimmt durch die sogenannte *Polarisationsformel*:

$$b(x, y) = \frac{1}{2} (q(x + y) - q(x) - q(y)).$$

Es ist nicht schwer, den *Gradienten* einer solchen quadratischen Form auszurechnen. Man bekommt sofort

$$\frac{\partial q}{\partial x_\ell} = \sum_{j,k=1}^n a_{jk} (\delta_{j\ell} x_k + x_j \delta_{k\ell}) = \sum_{k=1}^n a_{\ell k} x_k + \sum_{j=1}^n a_{j\ell} x_j = 2 \sum_{j=1}^n a_{j\ell} x_j$$

und damit die einfache Formel

$$\text{grad } q_S = 2 S x.$$

In dem Spezialfall $S := E_n$ ist q_S nichts anderes als das Quadrat der euklidischen Norm: $q_{E_n}(x) = \|x\|_2^2$. Nun ist aber die Einheitskugel S^{n-1} kompakt, so daß für jede symmetrische Matrix S die quadratische Form q_S auf S^{n-1} z. B. ein Maximum annimmt. Mit anderen Worten: Es existiert ein (globales) Maximum der Funktion q_S unter der Nebenbedingung $q_{E_n} - 1 = 0$. Somit gibt es nach Satz 18 eine Stelle $v_1 \in S^{n-1}$, so daß mit einem geeigneten $\lambda \in \mathbb{R}$ gilt:

$$S v_1 = \frac{1}{2} \text{grad } q_S(v_1) = \frac{\lambda}{2} \text{grad } q_{E_n}(v_1) = \lambda E_n v_1 = \lambda v_1.$$

Wir haben somit unter Verwendung *analytischer Methoden* bewiesen:

Satz 17.21 Jede reelle symmetrische Matrix $S = (a_{jk})$ besitzt einen Eigenvektor mit reellem Eigenwert λ .

Wählt man in der Situation wie zuvor eine Hyperebene H durch den Ursprung, die senkrecht steht auf dem Eigenvektor v_1 , so liefert ein ähnliches Argument die Existenz eines weiteren Eigenvektors $v_2 \in S^{n-1} \cap H$ mit reellem Eigenwert. Induktiv so fortfahrend, beweist man den

Satz 17.22 (Hauptachsentransformation) Zu jeder reellen symmetrischen Matrix $S = (a_{jk}) \in M(n \times n, \mathbb{R})$ gibt es ein Orthonormalsystem v_1, \dots, v_n von Eigenvektoren mit reellen Eigenwerten.

Selbstverständlich ist unter der Voraussetzung des vorigen Satzes das System der Vektoren v_1, \dots, v_n eine Basis von \mathbb{R}^n . Bezeichnet Q die Matrix mit den Koordinaten der Vektoren v_j relativ zu der Standardbasis e_1, \dots, e_n , so ist ${}^t Q Q = E_n$, da die Spaltenvektoren ein Orthonormalsystem bilden. Q ist somit eine *orthogonale* Matrix, also eine Matrix, die das euklidische Skalarprodukt invariant läßt²⁷:

$$\langle Qx, Qy \rangle = {}^t(Qx)(Qy) = {}^t x ({}^t Q Q) y = {}^t x y = \langle x, y \rangle.$$

Mit den üblichen Argumenten der Linearen Algebra folgt hieraus, daß die Matrix ${}^t Q S Q$ *Diagonalgestalt* besitzt: ${}^t Q S Q = \text{diag}(\lambda_1, \dots, \lambda_n)$, wobei die $\lambda_1, \dots, \lambda_n$ die *Eigenwerte* der Matrix S bezeichnet. Satz 22 kann also auch wie folgt umformuliert werden:

Satz 17.23 (Hauptachsentransformation: Zweite Fassung) Zu jeder reellen symmetrischen Matrix $S = (a_{jk}) \in M(n \times n, \mathbb{R})$ gibt es eine orthogonale Matrix $Q \in M(n \times n, \mathbb{R})$ und reelle Zahlen $\lambda_1, \dots, \lambda_n$ mit ${}^t Q S Q = \text{diag}(\lambda_1, \dots, \lambda_n)$.

Mit einer anschließenden geeigneten *Homothetie* $e_j \mapsto \alpha_j e_j$ mit $\alpha_j > 0$ läßt sich dann jede quadratische Form über den reellen Zahlen in die Gestalt $\langle x, Sx \rangle$ bringen, wobei die Matrix $S = S_{p,p'}$ die *Normalform*

$$S_{p,p'} = \text{diag}(\underbrace{1, \dots, 1}_p, \underbrace{-1, \dots, -1}_{p'}, 0, \dots, 0)$$

besitzt. Durch eine *lineare* Transformation kann also jede quadratische Form in die Gestalt

$$q(x) = \sum_{j=1}^p x_j^2 - \sum_{j=1}^{p'} x_{p+j}^2$$

gebracht werden.

Bemerkung. Die Eigenwerte einer Matrix $S \in M(n \times n, \mathbb{R})$ sind *Invarianten* des S zugeordneten *linearen Endomorphismus*; oder mit anderen Worten: Die Matrizen S und $Q^{-1} S Q$, $Q \in \text{GL}(n, \mathbb{R})$, besitzen das gleiche charakteristische Polynom und damit die gleichen Eigenwerte. Die symmetrische Matrix einer quadratischen Form transformiert sich jedoch auf andere Weise; es wird nämlich S bei Basiswechsel zu der symmetrischen Matrix ${}^t Q S Q$, $Q \in \text{GL}(n, \mathbb{R})$. Dies hat zur Folge, daß die soeben formulierte Aussage im Allgemeinen ihre Gültigkeit verliert (es sei denn, daß man nur *orthogonale* Transformationsmatrizen Q zuläßt). Immerhin besagt aber der SYLVESTERsche *Trägheitssatz*, daß die Anzahlen p, p' der positiven bzw. negativen Eigenwerte einer quadratischen Form *eindeutig* zugeordnet sind:

Satz 17.24 Es sei q eine quadratische Form und S die darstellende symmetrische Matrix bzgl. einer Basis. Dann sind die folgenden Zahlen Invarianten von q (also unabhängig von der Basis):

$$\begin{aligned} p &= \text{Anzahl der positiven Eigenwerte von } S, \\ p' &= \text{Anzahl der negativen Eigenwerte von } S, \\ r &= \text{Rang von } S. \end{aligned}$$

²⁷Die Gesamtheit der reellen orthogonalen $n \times n$ -Matrizen bildet eine Gruppe, die üblicherweise mit $O(n)$ bzw. genauer mit $O(n, \mathbb{R})$ bezeichnet wird.

Beweis. Siehe z. B. FISCHER [42], *Lineare Algebra*, Satz 5.7.4. □

Man nennt dann auch r den *Rang*, p den *Index* und $p - p'$ die *Signatur* von q . Hat q den Rang r , den Index p und die Signatur s , so besitzt q eine beschreibende Matrix der Gestalt

$$\left(\begin{array}{cc|cc|cc} 1 & 0 & & & & & \\ & \ddots & & & & & \\ 0 & 1 & & & & & \\ \hline & & -1 & 0 & & & \\ & & & \ddots & & & \\ & & 0 & -1 & & & \\ \hline & & & & 0 & 0 & \\ & & & & & \ddots & \\ & & & & 0 & 0 & \end{array} \right) \left. \begin{array}{l} \vphantom{\left(\right.} \\ \vphantom{\left. \right)} \\ \vphantom{\left(\right.} \\ \vphantom{\left. \right)} \\ \vphantom{\left(\right.} \\ \vphantom{\left. \right)} \\ \vphantom{\left(\right.} \\ \vphantom{\left. \right)} \end{array} \right\} \begin{array}{l} p \\ p' = r - p \\ n - r \end{array}$$

und es ist $s = p - (r - p) = 2p - r$.

Ebenso ist der Wert $\det S$ der Determinante von S der quadratischen Form q_S nicht eindeutig zugeordnet. Wegen

$$\det({}^t Q S Q) = (\det Q)^2 \cdot \det S$$

ist aber das *Vorzeichen* von $\det S$ eindeutig bestimmt, sofern nur $\det S \neq 0$, d. h. wenn die Form q *nicht ausgeartet* ist. Wir sagen kurz, die quadratische Form q *habe positive Determinante*, wenn $\det S > 0$ für eine und damit für alle darstellenden Matrizen S , d. h. wenn $r = n$ und $p' \equiv 0 \pmod{2}$.

Man kann nun einfache Kriterien dafür angeben, daß eine quadratische Form q *positiv definit* ist.

Satz 17.25 *Es sei q eine quadratische Form auf dem (euklidischen) Vektorraum V . Dann sind die folgenden Aussagen äquivalent :*

- i) q ist positiv definit ;
- ii) es gibt eine (Orthonormal-) Basis (v_1, \dots, v_n) von V und positive Zahlen $\lambda_j > 0$, s. d.

$$q \left(\sum_{j=1}^n a_j v_j \right) = \sum_{j=1}^n \lambda_j a_j^2 ;$$

- iii) es gibt eine (Orthogonal-) Basis (u_1, \dots, u_n) von V , so daß

$$q \left(\sum_{j=1}^n a_j u_j \right) = \sum_{j=1}^n a_j^2 .$$

D. h. mit anderen Worten : Es sei S eine symmetrische Matrix. Dann sind äquivalent :

- i)' ${}^t x S x > 0$ für alle $x \neq 0$;
- ii)' es existiert $Q \in O(n)$, s. d. ${}^t Q S Q = \text{diag}(\lambda_1, \dots, \lambda_n)$, $\lambda_j > 0$, $j = 1, \dots, n$;
- iii)' es existiert $Q' \in GL(n, \mathbb{R})$, s. d. ${}^t Q' S Q' = E_n$, d. h. es existiert $T \in GL(n, \mathbb{R})$, so daß $S = {}^t T T$.

Beweis. i) \implies ii). Wegen des Satzes über die Hauptachsentransformation kann man S diagonalisieren: es gibt $Q \in O(n)$, so daß

$${}^t Q S Q = \text{diag}(\lambda_1, \dots, \lambda_n).$$

Mit $x \neq 0$ ist auch $Qx \neq 0$, also

$$0 < {}^t(Qx) S (Qx) = {}^t x {}^t Q S Q x = \sum_{j=1}^n \lambda_j x_j^2.$$

Daraus folgt aber $\lambda_j > 0$ für $j = 1, \dots, n$ durch Einsetzen von $x = e_j$.

ii) \implies iii). Man ersetze Q durch

$$Q' := Q \cdot \text{diag}\left(\frac{1}{\sqrt{\lambda_1}}, \dots, \frac{1}{\sqrt{\lambda_n}}\right).$$

iii) \implies i) Es sei $x \neq 0$ und $y = Q'^{-1}x$. Dann folgt $y \neq 0$ und

$${}^t x S x = {}^t y {}^t Q' S Q' y = {}^t y E_n y = \sum_{j=1}^n y_j^2 > 0. \quad \square$$

Das HURWITZ-Kriterium für positive Definitheit läßt sich nun wie folgt formulieren:

Satz 17.26 *Es sei q eine quadratische Form auf dem endlich-dimensionalen \mathbb{R} -Vektorraum V der Dimension n . Dann sind die folgenden Aussagen äquivalent:*

- i) q ist positiv definit;
- ii) für jeden linearen Unterraum $U \subset V$ hat $q|_U$ positive Determinante;
- iii) es gibt eine vollständige Fahne von Untervektorräumen

$$\{0\} = U_0 \subset U_1 \subset \dots \subset U_{n-1} \subset U_n = V,$$

d. h. eine ineinandergeschachtelte Folge von Untervektorräumen $U_j \subset V$ mit $\dim U_j = j$, so daß $q|_{U_j}$ positive Determinante besitzt für $j = 1, \dots, n$.

Die klassische Form dieses Satzes ist nur eine Umformulierung.

Folgerung 17.27 *Es sei $S \in M(n \times n, \mathbb{R})$ eine symmetrische Matrix. Dann ist*

$$q_S(x) = {}^t x S x$$

positiv definit genau dann, wenn alle Hauptunterdeterminanten

$$S_j := \det \begin{pmatrix} a_{11} & \cdots & a_{1j} \\ \vdots & & \vdots \\ a_{j1} & \cdots & a_{jj} \end{pmatrix}, \quad j = 1, \dots, n,$$

der Matrix S positiv sind.

Beweis von Satz 26. i) \implies ii). $q|_U$ ist nach Definition auch positiv definit. Man braucht also ii) nur für $U = V$ zu beweisen. Wegen Satz 25 ist aber E_n eine beschreibende Matrix für q in einer geeigneten Basis; somit besitzt q positive Determinante.

ii) \implies iii) ist trivial.

iii) \implies i) wird per Induktion nach n bewiesen: Für $n = 1$ ist die Aussage trivialerweise richtig. Sei

sie also für Vektorräume der Dimension $n - 1$ schon als richtig erkannt. Dann können wir sie auch auf $q|_{U_{n-1}}$ anwenden und finden eine Basis (u_1, \dots, u_{n-1}) von U_{n-1} , so daß

$$q(a_1 u_1 + \dots + a_{n-1} u_{n-1}) = a_1^2 + \dots + a_{n-1}^2.$$

In dieser Basis gilt dann für die zugehörige Bilinearform b :

$$b(u_j, u_k) = \delta_{jk}, \quad 1 \leq j, k \leq n - 1.$$

Es sei nun $(u_1, \dots, u_{n-1}, v_n)$ eine Basis von V , und man setze

$$u_n = v_n - b(u_1, v_n) u_1 - \dots - b(u_{n-1}, v_n) u_{n-1}.$$

Dann ist auch (u_1, \dots, u_n) eine Basis von V , und es gilt für $j \leq n - 1$:

$$b(u_j, u_n) = b(u_j, v_n) - \sum_{k=1}^{n-1} b(u_k, v_n) b(u_j, u_k) = b(u_j, v_n) - \sum_{k=1}^{n-1} \delta_{jk} b(u_k, v_n) = 0.$$

Infolgedessen hat b in dieser Basis Diagonalgestalt:

$$\text{diag}(1, \dots, 1, b(u_n, u_n)).$$

Da q nach Voraussetzung positive Determinante besitzt, muß auch $b(u_n, u_n) > 0$ gelten, und folglich ist q positiv definit. \square

Anhang: Die Äquivalenz von Normen auf endlich - dimensionalen Vektorräumen und Anwendungen

Man nennt zwei Normen $\| \cdot \|$ und $\| \cdot \| ^*$ auf einem Vektorraum V *äquivalent*, wenn sie dieselbe *Topologie* erzeugen. - Man sieht sofort, daß folgendes richtig ist:

Lemma 17.28 Die Normen $\| \cdot \|$ und $\| \cdot \| ^*$ auf dem Vektorraum V sind genau dann äquivalent, wenn es Konstanten $C_1, C_2 > 0$ gibt, so daß für alle $v \in V$ gilt :

$$C_1 \| v \| ^* \leq \| v \| \leq C_2 \| v \| ^* .$$

Beweis. Wenn die Ungleichungen erfüllt sind, sind die beiden Normen offensichtlich äquivalent. Umgekehrt ist die offene Kugel $B_1^*(0)$ bzgl. der „gesternnten“ Norm notwendigerweise auch offen in der Norm $\| \cdot \|$. Infolgedessen gibt es eine positive reelle Zahl r , so daß $B_r(0) \subset B_1^*(0)$; mit anderen Worten: Ist $\| v \| < r$, so ist $\| v \| ^* < 1$. Für $v \neq 0$ ist aber

$$\frac{rv}{2\|v\|} \in B_r(0) \subset B_1^*(0) \quad \text{und somit} \quad \left\| \frac{rv}{2\|v\|} \right\| ^* < 1 ,$$

also

$$C \| v \| ^* \leq \| v \| \quad \text{mit} \quad C := \frac{r}{2} .$$

Die andere Ungleichung gewinnt man durch Vertauschung der Rollen der beiden Normen. □

Bemerkung. Es ist eine leichte Übungsaufgabe zu zeigen, daß Äquivalenz von Normen tatsächlich eine *Äquivalenzrelation* ist. Äquivalenz von Normen impliziert natürlich die Äquivalenz der induzierten Metriken. Bei Normen ist diese Eigenschaft jedoch viel stärker als bei Metriken.

Beispiel. Im \mathbb{R}^n gilt offensichtlich

$$\| v \|_\infty \leq \| v \|_2 \leq \sqrt{n} \| v \|_\infty .$$

Also sind die *euklidische* und die *Maximum-Norm* äquivalent.

Dies ist kein Zufall, denn es gilt der viel allgemeinere Satz, daß alle Normen auf *endlich-dimensionalen* Vektorräumen äquivalent sind. Wir werden weiter unten an Beispielen sehen, daß dies im *unendlich-dimensionalen* Fall nicht richtig ist.

Satz 17.29 Je zwei Normen auf einem endlich-dimensionalen reellen Vektorraum V sind äquivalent. Damit sind topologische Begriffe wie Offenheit, Abgeschlossenheit, Beschränktheit und Kompaktheit von Mengen und Konvergenz, Cauchy-Bedingung, Häufungspunkt etc. bei Folgen *unabhängig von der gewählten Norm*. Insbesondere ist V bzgl. jeder Norm vollständig.

Beweis. Ist $V \cong \mathbb{R}^n$, so induziert jede Norm auf V eine Norm auf \mathbb{R}^n und umgekehrt. Wir können uns daher im folgenden auf den Fall $V = \mathbb{R}^n$ beschränken und werden beweisen, daß jede beliebige Norm $\| \cdot \|$ auf \mathbb{R}^n zu der *Maximumnorm*

$$v \longmapsto \| v \|_\infty := \max_{j=1, \dots, n} |a_j| , \quad v = \sum_{j=1}^n a_j v_j ,$$

äquivalent ist. Hierbei bezeichnet (v_1, \dots, v_n) eine beliebige Basis von $V \cong \mathbb{R}^n$.

Eine Richtung ist sehr einfach. Es gilt nämlich

$$\| v \| = \left\| \sum_{j=1}^n a_j v_j \right\| \leq \sum_{j=1}^n |a_j| \| v_j \| \leq (n \cdot \max_{j=1, \dots, n} \| v_j \|) \| v \|_\infty =: C \| v \|_\infty .$$

Man hätte auch aus der Cauchy–Schwarzschen Ungleichung eine Abschätzung

$$\|v\| \leq C_1 \|v\|_2 \quad \text{mit} \quad C_1 := \sqrt{\sum_{j=1}^n \|v_j\|}$$

bekommen können.

Für die umgekehrte Richtung geben wir *drei* Beweise, von denen nur die ersten beiden den *Kompaktheitsbegriff* benötigen.

Wir beachten als erstes, daß aus der Dreiecksungleichung nach unten sofort

$$|\|v\| - \|v_0\|| \leq \|v - v_0\| \leq C \|v - v_0\|_\infty$$

folgt. Somit ist *jede* Norm $\|\cdot\| : V \rightarrow \mathbb{R}$ eine *stetige* Funktion auf V sowohl bzgl. der von $\|\cdot\|$ selbst als auch der von $\|\cdot\|_\infty$ induzierten Topologie.

Es ist nun der *Rand* $\partial Q = \{v \in V : \|v\|_\infty = 1\}$ des Einheitswürfels $Q = \{v \in V : \|v\|_\infty \leq 1\} \cong I \times \cdots \times I \subset \mathbb{R}^n$, $I = [-1, 1]$, eine *kompakte* Menge, auf der jede stetige Funktion ihr Maximum und Minimum annimmt. Wegen der obigen Überlegung gibt es also ein $v_0 \in \partial Q$, so daß

$$\|v\| \geq \|v_0\| > 0 \quad \text{für alle} \quad v \in \partial Q.$$

Ist nun $v \in V$ beliebig, aber von Null verschieden, so besitzt λv , $\lambda = 1/\|v\|_\infty$, die Eigenschaft, daß $\|\lambda v\|_\infty = |\lambda| \|v\|_\infty = 1$. Es folgt

$$\|v\| = \frac{1}{\lambda} \|\lambda v\| \geq \tilde{C} \|v\|_\infty \quad \text{mit} \quad \tilde{C} := \|v_0\| > 0.$$

2. Hier noch eine Beweisvariante: Es sei ∂Q wie zuvor definiert. Wir nehmen an, daß $\gamma := \inf\{\|v\| : v \in \partial Q\}$ *nicht* positiv sei. Dann gibt es eine Folge $v_j \in \partial Q$ mit $\lim \|v_j\| = 0$. Da für $\partial Q \subset \bar{Q} = \{v : \|v\|_\infty \leq 1\}$ der Satz von Bolzano und Weierstraß gilt, können wir ohne Einschränkung annehmen, daß die Folge v_j in der Maximumnorm gegen $a \in \partial Q$ konvergiert: $\|v_j - a\|_\infty \xrightarrow{j \rightarrow \infty} 0$. Da die Norm $\|\cdot\|$ stetig bezüglich $\|\cdot\|_\infty$ ist, wie wir oben gesehen haben, folgt

$$\|v_j\| \xrightarrow{j \rightarrow \infty} \|a\| \neq 0 \quad \text{wegen} \quad a \neq 0.$$

Widerspruch !

3. Für den dritten *sehr elementaren* Beweis führen wir Induktion nach n , wobei wir für $n - 1$ schon annehmen dürfen, daß auf jedem Vektorraum W mit $\dim W = n - 1$ alle Normen äquivalent sind und W deshalb vollständig ist (der Fall $n = 1$ ist trivial). Angenommen, es gäbe keine Konstante $C > 0$ mit

$$\|v\|_\infty \leq C \|v\|.$$

Dann muß es eine Folge $v_j \in V$ geben, so daß

$$\|v_j\|_\infty > j \|v_j\|.$$

Durch Multiplikation mit geeigneten Konstanten $\lambda_j \neq 0$ können wir erreichen, daß für alle $j \in \mathbb{N}$ die Vektoren v_j in der Supremumsnorm normiert sind: $\|v_j\|_\infty = 1$, und damit

$$(*) \quad \|v_j\| < \frac{1}{j}, \quad j \in \mathbb{N}^*,$$

gilt. Wir können die Normierungseigenschaft nach evtl. Multiplikation mit dem Faktor -1 dahingehend verschärfen, daß es ein $m \in \{1, \dots, n\}$ und eine aufsteigende unendliche Folge $\{j_k\}$ natürlicher Zahlen gibt, so daß in den Darstellungen

$$v_{j_k} = x_1^{(k)} w_1 + \cdots + x_n^{(k)} w_n$$

bzgl. einer Basis w_1, \dots, w_n von V stets $x_m^{(k)} = 1$ gilt. Es ist damit

$$v'_k = v_{j_k} - w_m \in W := \{v \in V : x_m = 0\}$$

und

$$(**) \quad \|v_{j_k}\| < \frac{1}{j_k} \quad \text{mit} \quad j_1 < j_2 < \dots \longrightarrow \infty.$$

Aus diesem Grunde ist die Folge (v'_k) in W eine Cauchyfolge bzgl. der Einschränkung der Norm $\|\cdot\|$ auf W :

$$\|v'_k - v'_\ell\| = \|(v_{j_k} - w_m) - (v_{j_\ell} - w_m)\| = \|v_{j_k} - v_{j_\ell}\| \leq \frac{1}{j_k} + \frac{1}{j_\ell} \leq \frac{2}{j_k} < \varepsilon$$

für $\ell \geq k \geq N(\varepsilon)$. Nach Induktionsvoraussetzung gibt es ein $v' \in W$ mit

$$\lim_{k \rightarrow \infty} v'_k = v'$$

bzgl. dieser Norm. Es folgt

$$\lim_{k \rightarrow \infty} v_{j_k} = v' + w_m$$

in V bzgl. $\|\cdot\|$. Dies widerspricht aber (**), da von daher

$$\|\lim_{k \rightarrow \infty} v_{j_k}\| = 0$$

sein müßte, andererseits aber offensichtlich $\|v' + w_m\| \geq 1$ gilt. □

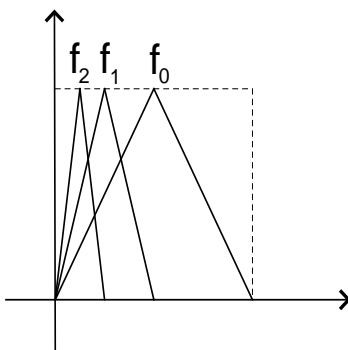
Beispiele. 1. Auf dem (unendlich-dimensionalen) Vektorraum $\mathcal{C}^0(I)$, $I = [a, b]$, $a < b$, wird durch

$$\|f\|_{L^1} := \int_a^b |f(x)| dx$$

eine Norm erklärt. Wir zeigen tatsächlich in Kapitel 18, daß man auf dem Raum *aller* Funktionen auf I eine „Pseudo-Halbnorm“ definieren kann, die bei stetigen Funktionen mit der eben gegebenen Definition übereinstimmt. Selbstverständlich kann der Norm-Wert für eine *stetige* Funktion niemals unendlich sein, und ist $f \neq 0$, so ist der Absolutbetrag $|f| \geq 0$ und auf einem ganzen nichttrivialen Teilintervall von I echt grösser als eine feste positive Konstante. Nach allgemeinen Gesetzen der Integrationstheorie (siehe Kapitel 18) ist dann $\|f\|_{L^1} > 0$ für jede stetige Funktion $f \neq 0$. Selbstverständlich gilt auch die Abschätzung:

$$\|f\|_{L^1} \leq (b - a) \|f\|_\infty.$$

Diese impliziert insbesondere, daß das Integral der Grenzfunktion einer gleichmäßig konvergenten Folge von stetigen Funktionen gleich dem Grenzwert der entsprechenden Integralfolge ist. Eine Abschätzung in der umgekehrten Richtung kann es jedoch nicht geben. Eine solche würde zur Konsequenz haben, daß jede Folge f_j nichtnegativer stetiger Funktionen auf I , deren Folge der Integrale gegen Null geht, gleichmäßig gegen Null konvergieren müßte. Man kann aber leicht Funktionenfolgen angeben, die der ersten Bedingung genügen, jedoch *konstante Supremumsnorm* besitzen. Durch geeignete Modifikation findet man Beispiele von solchen Funktionsfolgen, die sogar *unbeschränkt* in der Supremumsnorm sind.



Figur 17.5

2. Der (unendlich-dimensionale) Vektorraum $\mathcal{C}^1(I)$, $I = [a, b]$, $a < b$, besitzt neben der Supremumsnorm $\|\cdot\|_\infty$ zum Beispiel auch die folgende Norm (es sei dem Leser überlassen, die Norm-Axiome nachzuprüfen):

$$\|f\| := \|f'\|_\infty (b - a) + |f(a)|, \quad f \in \mathcal{C}^1(I).$$

Aufgrund des Hauptsatzes der Differential- und Integralrechnung ist

$$f(x) = f(a) + \int_a^x f'(t) dt \quad \text{und damit} \quad \|f\|_\infty \leq \|f\|.$$

Wäre aber umgekehrt $\|f\| \leq C \|f\|_\infty$, so wäre der Grenzwert f jeder gleichmäßig konvergenten Folge (f_j) von stetig differenzierbaren Funktionen f_j auf I automatisch wieder stetig differenzierbar. Denn dann wäre sowohl die Folge $(f_j(a))$ konvergent als auch die Folge der Ableitungen f'_j gleichmäßig konvergent auf I gegen eine (stetige) Grenzfunktion g , und dies hätte nach Satz 14.44 die stetige Differenzierbarkeit von f zur Folge. Wir wissen aber aufgrund von Beispielen aus dem gleichen Kapitel, daß diese Behauptung nicht zutreffend ist.

3. Die Abbildung

$$D : \begin{cases} \mathcal{C}^1([0, 1]) & \longrightarrow & \mathbb{R} \\ f & \longmapsto & f'(0) \end{cases}$$

ist linear, aber wenn \mathcal{C}^1 mit der *Supremumsnorm* versehen wird, *nicht* stetig. Z. B. bilden die Funktionen $f_j(x) = (\sin j^2 x)/j$, $j \geq 1$, eine Nullfolge, aber die Folge $(f'_j(0))_{j \in \mathbb{N}^*}$ divergiert.

Als erste Anwendung des Satzes über die Äquivalenz von Normen auf endlich-dimensionalen Vektorräumen beweisen wir den Satz von TSCHEBYSCHEF über *Bestapproximation* in Banachräumen. Man beachte, daß ein entsprechendes Ergebnis in *Skalarprodukträumen* viel einfacher zu beweisen ist, da man stets von einem Vektor das *Lot* auf einen endlich-dimensionalen Unterraum fallen kann. Dies ist letztendlich das ganze Geheimnis der Theorie der *Fourier-Reihen*, auf die wir später noch zu sprechen kommen (siehe Kapitel 19 samt Anhang). Ohne den Begriff des *Senkrechtstehens* ist aber ein so schneller Zugang nicht möglich. In der Tat haben wir es hier mit einem reinen *Existenzproblem* zu tun, das sogar im Allgemeinen nicht mehr *eindeutig* lösbar ist.

Satz 17.30 *Es sei V ein normierter Vektorraum, $U \subset V$ ein endlich-dimensionaler Untervektorraum, und ferner sei $v \in V$. Dann gibt es eine (i. A. nicht eindeutig bestimmte) Bestapproximation $u_0 \in U$ von v , d. h. einen Vektor $u_0 \in U$ mit*

$$\|v - u_0\| = \inf_{u \in U} \|v - u\|.$$

Beispiel und Bemerkung. In \mathbb{R}^2 mit der Maximumnorm besteht die Menge aller Bestapproximation des Vektors $v = (1, 1)$ in dem von $u_0 = (1, 0)$ erzeugten Untervektorraum U , wie man sofort nachrechnet, aus den Vektoren λu_0 mit $0 \leq \lambda \leq 2$. Man zeigt leicht, daß dieses Phänomen allgemeingültig im folgenden Sinne ist: Die Gesamtheit aller Bestapproximationen eines Vektors v bildet stets eine konvexe Teilmenge von U .

Beweis von Satz 30. Es sei ohne Einschränkung $v \notin U$ und $\alpha := \inf_{u \in U} \|v - u\|$. Dann gibt es eine Folge $u_j \in U$ mit $\alpha = \lim_{j \rightarrow \infty} \|v - u_j\|$. Wegen

$$\|u_j - u_0\| \leq \|u_j - v\| + \|u_0 - v\|$$

ist die Folge (u_j) in U beschränkt und damit wegen der Äquivalenz aller Normen auf U ohne Einschränkung (nach Auswahl einer Teilfolge gemäß Bolzano–Weierstraß) konvergent: Es existiert $\lim u_j =: \tilde{u}$. Wegen der Stetigkeit der Norm ist dann aber für alle $u \in U$:

$$\|v - \tilde{u}\| = \lim_{j \rightarrow \infty} \|v - u_j\| = \alpha \leq \|v - u\|. \quad \square$$

Wir formulieren eine erste Anwendung des Satzes von Tschebyschef, um seine starke Kraft an einem sehr einfachen Beispiel zu demonstrieren. Man beachte, daß hier absolut nichts mehr zu beweisen ist, wenn man sich klar gemacht hat, welche Räume die Rolle von V und U spielen sollen.

Folgerung 17.31 *Es sei $I \subset \mathbb{R}$ ein Intervall, $f : I \rightarrow \mathbb{R}$ sei eine stetige Funktion, und die Zahl $n \in \mathbb{N}$ sei vorgegeben. Dann gibt es ein Polynom P_0 vom Grad $\leq n$, so daß*

$$\|f - P_0\|_\infty \leq \|f - P\|_\infty$$

für alle $P \in \mathbb{R}[x]$ mit $\deg P \leq n$.

Beispiel. Die TSCHEBYSCHEF–Polynome

$$T_0(x) := 1, \quad T_n(x) := \frac{1}{2^{n-1}} \cos(n \arccos x), \quad n \in \mathbb{N}^*,$$

sind normiert und vom Grad n . Dies folgt aus der trigonometrischen Formel

$$\cos(n+1)\vartheta + \cos(n-1)\vartheta = 2 \cos n\vartheta \cos \vartheta$$

und der daraus resultierenden Beziehung

$$\cos n\vartheta = 2^{n-1} \cos^n \vartheta + \sum_{k=0}^{n-1} a_{k,n} \cos^k \vartheta.$$

Diese Polynome lösen ein noch allgemeineres Approximationsproblem: *Unter allen reellen Polynomen mit Grad n und höchstem Koeffizienten 1 approximiert T_n auf dem Intervall $[0, 1]$ das Nullpolynom am besten.* Man beachte, daß wir hier keine Bestapproximation in einem Untervektorraum suchen. Immerhin bilden die zur Konkurrenz zugelassenen Polynome eine konvexe Teilmenge in dem endlich-dimensionalen Unterraum der Polynome vom Grad $\leq n$, die das Nullpolynom nicht enthält.

Als weitere und weit wichtigere Anwendung beweisen wir den schon zitierten Satz über die Charakterisierung endlich-dimensionaler Vektorräume:

Satz 17.32 *Es sei V ein normierter Vektorraum mit Norm $\|\cdot\|$. Dann gilt: die abgeschlossene Einheitskugel*

$$\{v \in V : \|v\| \leq 1\}$$

ist genau dann kompakt, wenn $\dim V < \infty$.

Beweis. Ist die Kugel $B := \{ \|v\| \leq 1 \}$ kompakt, so wird sie insbesondere überdeckt von endlich vielen Kugeln

$$B_j := B(u_j, 1/2), \quad j = 1, \dots, n,$$

vom Radius $1/2$. Wir setzen $U = \text{span}(u_1, \dots, u_n)$ und werden zeigen, daß $V = U$ ist. Wäre dies nämlich nicht der Fall, so gäbe es ein $v \in V \setminus U$ mit Bestapproximation $\tilde{u} \in U$:

$$0 < \alpha = \|v - \tilde{u}\| \leq \|v - u\| \quad \text{für alle } u \in U.$$

Setze

$$\tilde{v} := \frac{v - \tilde{u}}{\|v - \tilde{u}\|},$$

so daß also $\|\tilde{v}\| = 1$. Nach Voraussetzung existiert ein j , so daß $\|\tilde{v} - u_j\| \leq 1/2$. Mit

$$v = \tilde{u} + \|v - \tilde{u}\| \tilde{v} = \tilde{u} + \|v - \tilde{u}\| u_j + \|v - \tilde{u}\| (\tilde{v} - u_j)$$

und $\tilde{u} + \|v - \tilde{u}\| u_j \in U$ folgt

$$\alpha \leq \|v - \tilde{u}\| \cdot \|\tilde{v} - u_j\| \leq \frac{1}{2} \alpha.$$

Widerspruch! □

Zum Abschluß dieses Anhangs machen wir noch einige Bemerkungen über *stetige lineare Operatoren*. Wir wissen, daß lineare Abbildungen $A : V \rightarrow W$ normierter Vektorräume stets stetig sind, wenn V und W *endliche* Dimension besitzen. Im unendlich-dimensionalen Fall ist dies nicht immer richtig (siehe das obige Beispiel). Man bezeichnet nun generell mit $L(V, W)$ den Vektorraum der *stetigen* linearen Abbildungen der normierten Vektorräume V, W . Man nennt lineare Abbildungen $V \rightarrow W$ in der Physik auch *lineare Operatoren* und Elemente in $L(V, W)$ *stetige* oder *beschränkte Operatoren*. Nach Voraussetzung ist für $A \in L(V, W)$ die Zahl

$$\|A\| := \sup\{ \|Av\| : v \in V, \|v\| \leq 1 \}$$

endlich. Sie heißt die *Operatornorm* von A . Man kann $\|A\|$ geometrisch als den größten „Dehnungskoeffizienten“ der Abbildung A interpretieren.

Satz 17.33 1. Die Operatornorm ist tatsächlich eine Norm auf $L(V, W)$.

2. Es gilt $\|Av\| \leq \|A\| \|v\|$ für alle $v \in V$.

3. Ist zusätzlich $B \in L(U, V)$, so gilt

$$\|AB\| \leq \|A\| \cdot \|B\|.$$

Die *Beweise* sind nicht schwer und seien deshalb dem Leser überlassen. □

Beispiel. Sind speziell $V = \mathbb{R}^n$ und $W = \mathbb{R}^m$ jeweils mit der *Maximumnorm* versehen, und identifiziert man $L(V, W)$ mit $\text{Hom}(\mathbb{R}^n, \mathbb{R}^m) \cong M(n \times m, \mathbb{R}) \cong \mathbb{R}^{n \times m}$, repräsentiert man also die (stetigen) Operatoren durch die darstellenden Matrizen bzgl. der Standardbasen, so berechnet man leicht

$$\|A\| = \max_j \left(\sum_{k=1}^m |a_{jk}| \right).$$

Insbesondere gilt

$$\|A\|_\infty := \max_{j,k} |a_{jk}| \leq \|A\| \leq n \|A\|_\infty.$$

Man nennt in diesem Fall die Operatornorm $\|A\|$ aus ersichtlichen Gründen auch die *Zeilensummen-Norm*.

18 Das Darboux-, Riemann- und Lebesgue-Integral

Unsere Forderungen an eine „integrierbare“ Funktion, die wir bisher gestellt haben, scheinen doch zu einschränkend gewesen zu sein, zumal wir nicht einmal der Funktion $f(x) := \sin(1/x)$, $x \neq 0$, $f(0) = 0$, ohne Verwendung eines weiteren Grenzwertbegriffes, dem *uneigentlichen Regel-Integral*, einen vernünftigen Integralwert zuordnen konnten. Abhilfe schafft hier der Begriff des RIEMANN-*Integrals*, also der Approximation des Integrals durch RIEMANNsche *Summen*, dem wir schon indirekt im letzten Kapitel bei der Integration *stetiger* Funktionen begegnet sind. Wir wollen diesen Begriff aber auch hier zunächst nicht direkt einführen, sondern über einen scheinbaren Umweg, der von DARBOUX stammt. Wir können dabei sehr viel deutlicher sehen, worin die Verallgemeinerung bzgl. des Regel-Integrals liegt und wie wir zu dem noch wesentlich weiter gehenden Konzept des LEBESGUE-*Integrals* gelangen.

Unser „Fehler“ bei der Einführung des Regelintegrals beruht offenbar darin, daß wir einen zu starken Grenzwertbegriff für Folgen integrierbarer Funktionen, nämlich den der *gleichmäßigen Konvergenz* (auf kompakten Intervallen), verwendet haben. Aber tatsächlich wollen wir ja nicht die zu integrierende Funktion beliebig genau approximieren, sondern „nur“ ihr Integral. Wir betrachten dazu das Beispiel 4 nach Satz 44 in Kapitel 14 der Funktionenfolge $f_n(x) := nx(1-x)^n \geq 0$ auf dem Einheitsintervall $[0, 1]$. Diese Folge ist *punktweise*, aber nicht gleichmäßig konvergent gegen die Grenzfunktion $f_0 \equiv 0$. Man berechnet aber sofort

$$I_n := \int_0^1 f_n(x) dx = \int_0^1 n(1-x)x^n dx = n \left(\frac{x^{n+1}}{n+1} - \frac{x^{n+2}}{n+2} \right) \Big|_0^1 = \frac{n}{(n+1)(n+2)},$$

so daß $\lim_{n \rightarrow \infty} I_n = 0 = I_0$.

Zur Erläuterung des DARBOUXschen Ansatzes nehmen wir gleich wieder eine abstraktere Haltung ein. Wir starten wieder mit einem allgemeinen *linearen Integral* $\mathcal{I}_a^b(f)$ auf einem Vektorraum \mathcal{F} von Funktionen $f: I \rightarrow \mathbb{R}$ und versuchen, einen größeren als den „Regelabschluß“ zu konstruieren. Da wir zunächst das Integral über ein festes kompaktes Intervall $[a, b]$ zu erklären haben, nehmen wir an, daß $\varphi \leq f \leq \psi$ auf $[a, b]$ für eine beliebige Funktion $f: I \rightarrow \mathbb{R}$ und Funktionen $\varphi, \psi \in \mathcal{F}$ gelte; dann sollte, da wir nach Möglichkeit das Integral zu einem *linearen* und damit auch zu einem *monotonen* Funktional fortsetzen wollen, die Ungleichungskette

$$\mathcal{I}_a^b(\varphi) \leq \mathcal{I}_a^b(f) \leq \mathcal{I}_a^b(\psi)$$

bestehen. Es ist daher naheliegend, die folgende Definition zu geben.

Definition und Bemerkung. Man nennt

$$\overline{\mathcal{I}}_a^b(f) := \inf \{ \mathcal{I}_a^b(\psi) : f \leq \psi \text{ auf } [a, b], \psi \in \mathcal{F} \}$$

das DARBOUXsche *Oberintegral* von f . Es existiert stets im uneigentlichen Sinne, wobei $\overline{\mathcal{I}}_a^b(f) = \infty$ gesetzt wird, wenn die rechts stehende definierende Menge leer ist, wenn es also kein $\psi \in \mathcal{F}$ gibt mit $f \leq \psi$ auf $[a, b]$. Entsprechend definiert man das *Unterintegral*

$$\underline{\mathcal{I}}_a^b(f) := \sup \{ \mathcal{I}_a^b(\varphi) : \varphi \leq f \text{ auf } [a, b], \varphi \in \mathcal{F} \}.$$

Bei Verwendung des Integralzeichens schreibt man auch

$$\int_a^{\overline{b}} f(x) dx \text{ oder } \int_a^{*b} f(x) dx \quad \text{und} \quad \int_a^b f(x) dx \text{ oder } \int_{*a}^b f(x) dx$$

oder Ähnliches.

Die Funktion f heißt DARBOUX-*integrierbar* (bzgl. \mathcal{I}), wenn das Ober- und das Unterintegral (reell sind und) übereinstimmen:

$$\overline{\mathcal{I}}_a^b(f) = \underline{\mathcal{I}}_a^b(f) \in \mathbb{R}.$$

Der gemeinsame Wert heißt selbstverständlich das DARBOUX-Integral von f und wird mit $\text{Darb-}\mathcal{I}_a^b(f)$ oder kürzer $D\text{-}\mathcal{I}_a^b(f)$ oder wieder einfach mit $\mathcal{I}_a^b(f)$ bezeichnet.

Man kann sich intensiv mit den Eigenschaften der Ober- und Unterintegrale beschäftigen. Wir wollen uns aber damit nicht lange aufhalten, sondern sogleich die Darboux-integrierbaren Funktionen in Augenschein nehmen. Die zentrale Charakterisierung ist die folgende einfache Bemerkung.

Lemma 18.1 *Eine Funktion $f : I \rightarrow \mathbb{R}$ ist genau dann DARBOUX-integrierbar bzgl. \mathcal{I} , wenn es zu jedem $\varepsilon > 0$ und zu je zwei Elementen $a < b$ in I Funktionen $\varphi, \psi \in \mathcal{F}$ gibt, so daß*

$$\varphi \leq f \leq \psi \text{ auf } [a, b] \text{ und } \mathcal{I}_a^b(\psi - \varphi) \leq \varepsilon.$$

Sind $\varphi_j, \psi_j \in \mathcal{F}$ zwei Folgen mit $\varphi_j \leq f \leq \psi_j$ auf $[a, b]$ und $\lim_{j \rightarrow \infty} \mathcal{I}_a^b(\psi_j - \varphi_j) = 0$, so ist

$$\text{Darb-}\mathcal{I}_a^b(f) = \lim_{j \rightarrow \infty} \mathcal{I}_a^b(\varphi_j) = \lim_{j \rightarrow \infty} \mathcal{I}_a^b(\psi_j).$$

Beweis. Ist f Darboux-integrierbar und ist J sein Integral auf $[a, b]$, so gibt es zu vorgegebenem $\varepsilon > 0$ eine Funktion $\varphi \in \mathcal{F}$ mit $\varphi \leq f$ auf $[a, b]$ und $J - \mathcal{I}_a^b(\varphi) < \varepsilon/2$. Entsprechend findet man eine Funktion $\psi \in \mathcal{F}$ mit $f \leq \psi$ auf $[a, b]$ und $\mathcal{I}_a^b(\psi) - J < \varepsilon/2$. Zusammen ergibt sich unter Einbeziehung der Linearität des zugrunde liegenden Integrals das gewünschte Kriterium. Die Umkehrung und die letzte Behauptung folgen ebenso einfach. \square

Wir bezeichnen im folgenden mit $\text{Darb-}\mathcal{F}$ den *Darboux-schen Abschluß* von \mathcal{F} , also die Menge aller Funktionen, die Darboux-integrierbar sind. Selbstverständlich ist $\mathcal{F} \subset \text{Darb-}\mathcal{F}$, und die Funktionen in \mathcal{F} „behalten“ ihre Integrale. Eine unmittelbare Konsequenz aus dem vorstehenden Lemma ist die

Folgerung 18.2 *Darb-}\mathcal{F} ist ein Vektorraum, und das Darboux-Integral ist ein lineares Funktional.*

Beweis. Aus Ungleichungen $\varphi_j \leq f_j \leq \psi_j$ und $\mathcal{I}_a^b(\psi_j - \varphi_j) \leq \varepsilon_j$, $j = 1, 2$, folgt sofort $\varphi_1 + \varphi_2 \leq f_1 + f_2 \leq \psi_1 + \psi_2$ und $\mathcal{I}_a^b((\psi_1 + \psi_2) - (\varphi_1 + \varphi_2)) = \mathcal{I}_a^b(\psi_1 - \varphi_1) + \mathcal{I}_a^b(\psi_2 - \varphi_2) \leq \varepsilon_1 + \varepsilon_2$, so daß mit f_1 und f_2 auch ihre Summe in dem Darboux-Abschluß liegt. Noch einfacher ist der Nachweis, daß mit f auch λf , $\lambda \in \mathbb{R}$, Darboux-integrierbar ist. Also ist $\text{Darb-}\mathcal{F}$ ein \mathbb{R} -Vektorraum. Die Linearität des Darboux-Integrals ergibt sich aus den Grenzwertsätzen für konvergente Folgen: Nach Definition gilt für jede Folge $\varphi_j \leq f \leq \psi_j$ mit $\lim_{j \rightarrow \infty} \mathcal{I}_a^b(\psi_j - \varphi_j) = 0$, daß $\mathcal{I}_a^b(f) := \text{Darb-}\mathcal{I}_a^b(f) = \lim_{j \rightarrow \infty} \mathcal{I}_a^b(f_j)$. Hieraus gewinnt man sofort die Einsicht, daß das Darboux-„Integral“ zumindest ein *lineares* Funktional ist. \square

Es bleibt noch zu zeigen, daß das Darboux-Integral tatsächlich ein *Integral* in unserem abstrakten Sinne ist und das Regel-Integral fortsetzt. Dies können wir zumindest unter einer sehr schwachen weiteren Voraussetzung herleiten. Wir erinnern daran, daß wir schon früher den *Regel-Abschluß* $\text{Reg-}\mathcal{F}$ von \mathcal{F} eingeführt hatten.

Satz 18.3 *Der Vektorraum \mathcal{F} enthalte zu je zwei Elementen $a, b \in I$ mit $a < b$ eine Funktion χ_{ab} , die auf dem kompakten Intervall $[a, b]$ identisch gleich 1 ist. Dann ist $\mathcal{F} \subset \text{Reg-}\mathcal{F} \subset \text{Darb-}\mathcal{F}$ eine Kette von (Unter-) Vektorräumen. Das Darboux-sche Integral auf $\text{Darb-}\mathcal{F}$ ist ein lineares Integral, welches das Regel-Integral fortsetzt.*

Beweis. Es ist zuerst zu zeigen, daß der Raum $\text{Reg-}\mathcal{F}$ der Regel-Funktionen in $\text{Darb-}\mathcal{F}$ enthalten ist. Es sei also $f \in \text{Reg-}\mathcal{F}$; dann gibt es zu jedem kompakten Intervall $[a, b] \subset I$ und jedem $\varepsilon > 0$ eine Funktion $\chi \in \mathcal{F}$, so daß $|f - \chi|_{[a, b]} < \varepsilon/2$, also $-\varepsilon/2 < f(x) - \chi(x) < \varepsilon/2$ für alle $x \in [a, b]$. Nach Voraussetzung ist $\Phi := (\varepsilon/2)\chi_{ab} \in \mathcal{F}$ und folglich

$$\varphi(x) := \chi(x) - \Phi(x) \leq f(x) \leq \chi(x) + \Phi(x) =: \psi(x) \text{ für alle } x \in [a, b].$$

Wegen $\psi(x) - \varphi(x) = 2\Phi(x)$ ist aber $\mathcal{I}_a^b(\psi - \varphi) = \mathcal{I}_a^b(2\Phi) = \varepsilon(b - a)$, womit auch die Inklusion $\text{Reg-}\mathcal{F} \subset \text{Darb-}\mathcal{F}$ gezeigt ist, von der man sofort einsieht, daß sie eine Inklusion von Vektorräumen ist.

Ist f eine Regelfunktion, so zeigt der erste Teil des Beweises, daß wir f beliebig genau in der *Supremumsnorm* von $[a, b]$ durch zwei Funktionenfolgen φ_j und $\psi_j \in \mathcal{F}$ approximieren können mit $\varphi_j \leq f \leq \psi_j$ und $\mathcal{I}_a^b(\psi_j - \varphi_j) \rightarrow 0$. Es folgt für $f \in \text{Reg-}\mathcal{F}$, daß

$$\text{Reg-}\mathcal{I}_a^b(f) = \lim_{j \rightarrow \infty} \mathcal{I}_a^b(\varphi_j) = \text{Darb-}\mathcal{I}_a^b(f).$$

Wir müssen weiter nachweisen, daß das Darboux-Integral tatsächlich unseren Integral-Axiomen genügt. Das erste Axiom gewinnt man wie folgt: Für beliebige $a < b < c$ und $\varphi_j \leq f \leq \psi_j$ auf $[a, c]$ mit $\mathcal{I}_a^c(\psi_j - \varphi_j) < \varepsilon_j$ ist auch $\varphi_j \leq f \leq \psi_j$ auf $[a, b]$ und

$$\mathcal{I}_a^b(\psi_j - \varphi_j) \leq \mathcal{I}_a^b(\psi_j - \varphi_j) + \mathcal{I}_b^c(\psi_j - \varphi_j) = \mathcal{I}_a^c(\psi_j - \varphi_j) < \varepsilon_j$$

und entsprechend für das Intervall $[b, c]$. Damit sind die Folgen φ_j und ψ_j approximierende Folgen für f im Sinne des Darboux-Integrals auf jedem der beiden Teilintervalle, so daß wir als Konsequenz das erste Axiom bestätigen können:

$$\mathcal{I}_a^c(f) = \lim_{j \rightarrow \infty} \mathcal{I}_a^c(\varphi_j) = \lim_{j \rightarrow \infty} (\mathcal{I}_a^b(\varphi_j) + \mathcal{I}_b^c(\varphi_j)) = \lim_{j \rightarrow \infty} \mathcal{I}_a^b(\varphi_j) + \lim_{j \rightarrow \infty} \mathcal{I}_b^c(\varphi_j) = \mathcal{I}_a^b(f) + \mathcal{I}_b^c(f).$$

Ist schließlich f Darboux-integrierbar und $f \geq m$ auf $[a, b]$, also auch $f(x) \geq m \chi_{ab}(x)$, $x \in [a, b]$, so ist $f - m \chi_{ab} \geq 0$ auf $[a, b]$, und wegen der Linearität des Darboux-Integrals brauchen wir noch zu zeigen, daß aus $f \geq 0$ auf $[a, b]$ auch $\text{Darb-}\mathcal{I}_a^b(f) \geq 0$ folgt. Dies ist aber klar, da in diesem Fall das Integral $\mathcal{I}_a^b(0) = 0$ in der Menge der Integrale liegt, deren Supremum das Unterintegral und damit das Integral von f liefert. \square

Bemerkung und Definition. Die zusätzliche Voraussetzung an den Funktionenraum \mathcal{F} ist zum Beispiel dann erfüllt, wenn er die konstanten Funktionen oder den Raum $\mathcal{T}(I)$ der Treppenfunktionen auf I enthält. Insbesondere ist der obige Satz also anwendbar im Fall, daß $\mathcal{F} = \mathcal{T}(I)$ ist. In dieser Situation sprechen wir von dem Raum $\text{Darb}(I)$ der DARBOUX-integrierbaren Funktionen auf I (schlechthin). Wir haben also stets die Inklusion von Vektorräumen

$$\mathcal{T}(I) \subset \text{Reg}(I) \subset \text{Darb}(I).$$

In der Tat ist der Raum $\text{Darb}(I)$ *echt größer* als der Raum der Regelfunktionen. Z. B. ist das anfangs erwähnte Beispiel $\sin(1/x)$ mit beliebig fest gewähltem Wert bei 0 eine Darboux-integrierbare Funktion. Dies ist eine Konsequenz aus dem folgenden viel allgemeineren Lemma. Wir erinnern daran, daß eine Menge $D \subset I$ *diskret* ist, wenn ihr Durchschnitt $D \cap [a, b]$ mit einem kompakten Intervall $[a, b] \subset I$ stets nur endlich viele Punkte besitzt.

Lemma 18.4 *Ist $f : I \rightarrow \mathbb{R}$ eine (lokal) beschränkte Funktion, die außerhalb einer diskreten Menge D Darboux-integrierbar ist, so ist f selbst Darboux-integrierbar.*

Beweis. Wir brauchen nur zu zeigen, daß das Darboux-Integral von f über jedem kompakten Intervall $[a, b]$ existiert. Da $D \cap [a, b]$ endlich ist, können wir durch eine endliche Unterteilung erreichen, daß nur jeweils höchstens einer der Endpunkte der Teilintervalle zu D gehört. Sei dies ohne Einschränkung der rechte, und das Teilintervall selbst gleich $[a, b]$. Da die Funktion f nahe bei b beschränkt ist, gibt es eine positive reelle Zahl M und ein $\beta < b$, so daß $|f(x)| < M$ für alle x mit $\beta \leq x \leq b$. Wähle dann zu vorgegebenem $\varepsilon > 0$ ohne Einschränkung β so nahe an b , daß $2M(b - \beta) < \varepsilon/2$. Nach Voraussetzung ist f auf dem Intervall $[a, \beta]$ Darboux-integrierbar. Folglich gibt es Treppenfunktionen φ, ψ auf diesem Intervall, so daß dort $\varphi \leq f \leq \psi$ und $\int_a^\beta (\psi - \varphi) < \frac{\varepsilon}{2}$ gilt. Man setze nun diese beiden Treppenfunktionen auf $(\beta, b]$ durch $-M$ bzw. M fort. Dies liefert Treppenfunktionen auf dem ganzen Intervall $[a, b]$, die die Funktion f „einhüllen“ und deren Integraldifferenz kleiner als ε ist. Nach der Definition des Darbouxschen Integrals ist die Funktion f integrierbar. Die letzte Behauptung ist wegen der Standardabschätzung für Integrale trivial: Liegt β so nahe bei b , daß f auf $[\beta, b]$ dem Betrage nach durch M beschränkt ist, so folgt

$$\left| \int_a^b f(x) dx - \int_a^\beta f(x) dx \right| = \left| \int_\beta^b f(x) dx \right| \leq (b - \beta) M. \quad \square$$

Bemerkung. Im Rahmen der LEBESGUE-Theorie kann man eine *intrinsische* Charakterisierung der Darboux-integrierbaren Funktionen gewinnen: Eine Funktion $f : I \rightarrow \mathbb{R}$ ist genau dann Darboux-integrierbar, wenn die Menge ihrer Unstetigkeitsstellen eine LEBESGUE-Nullmenge ist. Zur Definition dieses Konzeptes siehe weiter unten; z. B. sind höchstens abzählbar unendliche Mengen Nullmengen im Lebesgueschen Sinne.

Beispiel. Hiernach ist die nirgends stetige DIRICHLETSche Funktion $\chi_{\mathbb{Q} \cap [0,1]}$ nicht Darboux-integrierbar. Dies kann man aber auch direkt ohne tiefliegende Hilfsmittel einsehen, denn es ist offensichtlich das Unterintegral dieser Funktion gleich 0 und ihr Oberintegral gleich 1.

Ist $I = [a, b]$ ein *kompaktes* Intervall, so ist $f : [a, b] \rightarrow \mathbb{R}$ genau dann eine Treppenfunktion, wenn $f|_{[a,c]}$ und $f|_{[c,b]}$ Treppenfunktionen sind für ein c (oder auch für alle c) mit $a < c < b$. Man macht sich schnell klar, daß eine entsprechende Aussage auch für Regelfunktionen gültig ist, und für Darboux-integrierbare Funktionen folgt sie aus der einfachen Tatsache, daß man zwei Treppenfunktionen auf den Intervallen $[a, b]$ und $[b, c]$ durch geeignete Festlegung des Funktionswertes bei b zu einer Treppenfunktion auf $[a, c]$ „zusammenfügen“ kann, ohne die Additivität des Integrals zu beeinträchtigen.

Lemma 18.5 Für eine Funktion $f : [a, c] \rightarrow \mathbb{R}$, $a < c$, sind die folgenden Aussagen äquivalent :

- i) f ist Darboux-integrierbar ;
- ii) für alle b mit $a < b < c$ sind $f|_{[a,b]}$ und $f|_{[b,c]}$ Darboux-integrierbar ;
- iii) es gibt ein b mit $a < b < c$, so daß $f|_{[a,b]}$ und $f|_{[b,c]}$ Darboux-integrierbar sind.

Beweis. Man zeigt i) \implies ii) \implies iii) \implies i). □

Wir wollen weiter zeigen, daß mit f notwendig auch $|f|$ Darboux-integrierbar ist. Diese Eigenschaft erfüllt auch sowohl der Vektorraum der Treppenfunktionen als auch derjenige der stetigen Funktionen. Allgemein gilt die folgende Charakterisierung solcher *Funktionenverbände*.

Satz 18.6 Es sei \mathcal{F} ein Vektorraum von Funktionen $f : I \rightarrow \mathbb{R}$. Dann sind die folgenden Aussagen äquivalent :

- i) Aus $f \in \mathcal{F}$ folgt stets $|f| \in \mathcal{F}$;
- ii) aus $f \in \mathcal{F}$ folgt stets $f^+ := \max(f, 0) \in \mathcal{F}$;
- iii) aus $f, g \in \mathcal{F}$ folgt stets $\max(f, g) \in \mathcal{F}$.

Beweis. Die Schlußkette i) \implies iii) \implies ii) \implies i) verläuft über die offensichtlichen Identitäten

$$2 \max(f, g) = f + g + |f - g|, \quad f^+ = \max(f, 0), \quad f^- = \max(-f, 0), \quad |f| = f^+ + f^-. \quad \square$$

Satz 18.7 Der Vektorraum $\text{Darb}(I)$ der Darboux-integrierbaren Funktionen auf einem Intervall I ist ein Funktionenverband.

Beweis. Nach dem vorstehenden Satz brauchen wir nur nachzuweisen, daß mit f auch f^+ Darboux-integrierbar ist; ferner können wir annehmen, daß $I = [a, b]$. Sind dann φ, ψ Treppenfunktionen mit $\varphi \leq f \leq \psi$, so ist auch $\varphi^+ \leq f^+ \leq \psi^+$. An Stellen, an denen ψ negativ ist, verschwinden ψ^+ und φ^+ , so daß dort $\psi^+ - \varphi^+ = 0 \leq \psi - \varphi$, und an den anderen Stellen ist ebenfalls $\psi^+ - \varphi^+ \leq \psi^+ - \varphi = \psi - \varphi$. Insbesondere gilt

$$\int_a^b (\psi^+(x) - \varphi^+(x)) dx \leq \int_a^b (\psi(x) - \varphi(x)) dx. \quad \square$$

Satz 18.8 Es seien $f, g : [a, b] \rightarrow \mathbb{R}$ Darboux-integrierbare Funktionen, und p sei eine positive natürliche Zahl. Dann sind auch $|f|^p$ und $f \cdot g$ Darboux-integrierbar.

Beweis. 1. Nach dem zuvor bewiesenen Satz dürfen wir voraussetzen, daß $f = |f| \geq 0$. Da Darboux-integrierbare Funktionen notwendig lokal beschränkt sind, können wir ferner annehmen, daß auf dem kompakten Intervall $[a, b] \subset I$ die Funktion f nach oben durch 1 beschränkt ist. Wählt man nun zu vorgegebenem $\varepsilon > 0$ Treppenfunktionen φ und ψ mit $\varphi \leq f \leq \psi$ und $\int_a^b (\psi - \varphi) dx < \varepsilon/p$, so kann man ohne Einschränkung $\varphi \geq 0$ und $\psi \leq 1$ verlangen. Für die Treppenfunktionen φ^p und ψ^p ist dann auch $\varphi^p \leq f^p \leq \psi^p$. Nun hat man für $0 \leq \alpha < \beta \leq 1$ die Ungleichung

$$\frac{\beta^p - \alpha^p}{\beta - \alpha} = \beta^{p-1} + \beta^{p-2}\alpha + \dots + \alpha^{p-1} \leq p,$$

die man auch vermöge des Mittelwertsatzes aus der Beziehung $\frac{dx^p}{dx} = px^{p-1}$ ableiten kann, so daß

$$\int_a^b (\psi^p(x) - \varphi^p(x)) dx \leq p \int_a^b (\psi(x) - \varphi(x)) dx < \varepsilon.$$

2. Es ist

$$f \cdot g = \frac{1}{4} [(f + g)^2 - (f - g)^2].$$

Hieraus folgt die Behauptung aus dem ersten Teil mit $p = 2$. □

Warnung. Selbstverständlich ist im Allgemeinen

$$\int_a^b f(x) \cdot g(x) dx \neq \left(\int_a^b f(x) dx \right) \cdot \left(\int_a^b g(x) dx \right).$$

Bemerkung. Auch für das LEBESGUE-Integral $\int_{\mathbb{R}} f(x) dx$ hat man notwendig die Integrierbarkeit von $|f|$. Dies hat einen gewissen Nachteil gegenüber dem uneigentlichen RIEMANN-Integral über \mathbb{R} zur Konsequenz. So existiert $\int_{-\infty}^{\infty} \frac{\sin x}{x} dx$ im uneigentlichen Riemannschen Sinne, aber nicht im Lebesgueschen, da sonst auch das Integral über den Betrag des Integranden im Lebesgueschen wie auch im uneigentlichen Riemannschen Sinne existieren müßte. Das Lebesgue-Integral läßt im Übrigen so viele Funktionen zur Integration zu, daß die direkte Übertragung der zweiten Aussage über das Produkt von zwei integrierbaren Funktionen falsch wird.

Die *Substitutionsformel* oder *Transformationsformel* können wir für das Darboux-Integral in der folgenden Fassung herleiten.

Satz 18.9 Ist $\alpha : J \rightarrow I$ eine stetig differenzierbare Abbildung zwischen reellen Intervallen mit $\alpha' \geq 0$, so ist mit $f : I \rightarrow \mathbb{R}$ auch auch $(f \circ \alpha) \cdot \alpha' : J \rightarrow \mathbb{R}$ Darboux-integrierbar, und es gilt für alle $a, b \in J$:

$$\int_a^b f(\alpha(t)) \cdot \alpha'(t) dt = \int_{\alpha(a)}^{\alpha(b)} f(x) dx.$$

Beweis. Die Aussage folgt offensichtlich für Treppenfunktionen auf I wegen der Linearität des Integrals aus der Substitutionsregel für stetige Funktionen. Ist nun f Darboux-integrierbar, so ist für alle $a, b \in J$ mit $a < b$ das Bildintervall gleich $[\alpha(a), \alpha(b)]$ (nach Voraussetzung ist α monoton steigend). Nach Voraussetzung gibt es Folgen von Treppenfunktionen φ_j, ψ_j mit $\varphi_j \leq f \leq \psi_j$ auf dem Intervall $K := [\alpha(a), \alpha(b)]$, und es konvergiert $\int_K (\psi_j - \varphi_j) dx \rightarrow 0$, so daß

$$\int_{\alpha(a)}^{\alpha(b)} f(x) dx = \lim_{j \rightarrow \infty} \int_{\alpha(a)}^{\alpha(b)} \varphi_j(x) dx.$$

Nun sind die Funktionen $(\varphi_j \circ \alpha) \cdot \alpha'$, $(\psi_j \circ \alpha) \cdot \alpha'$ auf jeden Fall Regelfunktionen und damit Darboux-integrierbar auf $[a, b]$. Ferner ist dort $\varphi_j \circ \alpha \leq f \circ \alpha \leq \psi_j \circ \alpha$ und folglich, aufgrund der Voraussetzung $\alpha' \geq 0$, auch $(\varphi_j \circ \alpha) \cdot \alpha' \leq (f \circ \alpha) \cdot \alpha' \leq (\psi_j \circ \alpha) \cdot \alpha'$. Wegen

$$\int_a^b ((\psi_j \circ \alpha) \cdot \alpha' - (\varphi_j \circ \alpha) \cdot \alpha') dx = \int_a^b (((\psi_j - \varphi_j) \circ \alpha) \cdot \alpha') dx = \int_{\alpha(a)}^{\alpha(b)} (\psi_j - \varphi_j) dx \longrightarrow 0$$

ist dann f Darboux-integrierbar auf $[a, b]$, und es folgt

$$\int_a^b (f \circ \alpha)(x) \cdot \alpha'(x) dx = \lim_{j \rightarrow \infty} \int_a^b (\varphi_j \circ \alpha)(x) \cdot \alpha'(x) dx = \lim_{j \rightarrow \infty} \int_{\alpha(a)}^{\alpha(b)} \varphi_j(x) dx = \int_{\alpha(a)}^{\alpha(b)} f(x) dx. \quad \square$$

Hieraus leitet man unmittelbar ab:

Folgerung 18.10 *Ist $\alpha : J \rightarrow I$ bijektiv mit einer stetig differenzierbaren Umkehrabbildung, so ist die Funktion $f : I \rightarrow \mathbb{R}$ genau dann Darboux-integrierbar, wenn $f \circ \alpha : J \rightarrow \mathbb{R}$ Darboux-integrierbar ist.*

Beweis. Man braucht nur eine Richtung zu zeigen: Mit f ist auch $f \circ \alpha$ Darboux-integrierbar; denn die Umkehrung folgt hieraus durch Anwendung auf die Umkehrabbildung von α . Ist nun α , wie vorausgesetzt, ein *Diffeomorphismus*, so ist an allen Stellen $\alpha' \neq 0$ und folglich α' überall positiv oder überall negativ. Insbesondere ist $1/\alpha'$ eine wohldefinierte stetige Funktion auf J , also insbesondere Darboux-integrierbar. Nach dem ersten Teil ist, wenn wir ohne Einschränkung $\alpha' > 0$ annehmen, $(f \circ \alpha) \cdot \alpha'$ auf J Darboux-integrierbar. Da das Produkt von Darboux-integrierbaren Funktionen wieder Darboux-integrierbar ist, folgt hieraus durch Multiplikation mit $1/\alpha'$ die Behauptung. \square

Im folgenden Abschnitt werden wir nachweisen, daß der Zugang von DARBOUX mit dem von RIEMANN identisch ist. Dazu müssen wir den Begriff des RIEMANN-Integrals erst einmal sorgfältig formulieren. Der Leser wird bemerken, daß die abstrakte Definition wesentlich komplizierter ausfällt als die des Darboux-Integrals. Wir widerstehen der Versuchung, die oben bewiesenen Resultate über das Darboux-Integral noch einmal für das Riemann-Integral zu formulieren.

Definition. Eine Funktion $f : I \rightarrow \mathbb{R}$ heißt RIEMANN-integrierbar, wenn es zu jedem kompakten Intervall $[a, b] \subset I$ eine Zahl R_a^b gibt derart, daß zu jedem $\varepsilon > 0$ ein $\delta > 0$ existiert mit

$$|R_a^b - S(f)| < \varepsilon$$

für alle Riemanschen Summen $S(f) = S(f, \mathcal{Z}, (\xi_j))$ zu $f|_{[a,b]}$ bzgl. Zerlegungen \mathcal{Z} des Intervalls $[a, b]$ der Feinheit $\delta(\mathcal{Z}) \leq \delta$.

Bemerkung. Aus der Definition geht unmittelbar hervor, daß die Zahlen R_a^b eindeutig bestimmt sind. Man nennt R_a^b natürlich das RIEMANN-Integral von f über $[a, b]$ und schreibt wie üblich

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

Wenn wir wissen, daß eine Funktion Riemann-integrierbar ist, so brauchen wir zur Berechnung des Riemann-Integrals nur eine geeignete Folge \mathcal{Z}_n von Zerlegungen des Intervalls $[a, b]$ und zu jeder Zerlegung \mathcal{Z}_n geeignete Stützstellen (ξ_{nj}) auszuwählen. Konvergiert dann die Folge $\delta(\mathcal{Z}_n)$ der Feinheiten gegen Null, so ist automatisch

$$\int_a^b f(x) dx = \lim_{n \rightarrow \infty} S(f|_{[a,b]}, \mathcal{Z}_n, (\xi_{nj})).$$

Beispiel. Wir haben im vorigen Kapitel nachgewiesen, daß die *stetigen* Funktionen Riemann-integrierbar sind und ihre Riemann-Integrale mit den Regel-Integralen übereinstimmen. Selbstverständlich ist dies auch für die *Treppenfunktionen* richtig.

Man kann alle relevanten Eigenschaften des Riemann-Integrals direkt ableiten, also insbesondere, daß die Menge $\text{Riem}(I)$ der Riemann-integrierbaren Funktionen einen Vektorraum bildet und das Integral über jedem kompakten Intervall $[a, b] \subset I$ ein Integral in unserer abstrakten Definition und zusätzlich *linear* ist, also insbesondere ein *monotones* Funktional darstellt. Wir können uns aber diesen Nachweis ersparen, wenn wir zeigen - was wir anschließend ausführen werden:

Satz 18.11 *Es gilt $\text{Darb}(I) = \text{Riem}(I)$.*

Der *Beweis* erfolgt in mehreren Schritten. Im ersten benötigen wir, daß Riemann-integrierbare Funktionen stets lokal beschränkt sind. (Dies ist der Grund, warum in zahlreichen Texten das Riemann-Integral von vornherein nur für beschränkte Funktionen auf kompakten Intervallen entwickelt wird).

Lemma 18.12 *Eine Riemann-integrierbare Funktion $f : I \rightarrow \mathbb{R}$ ist auf jedem kompakten Intervall $[a, b] \subset I$ beschränkt.*

Beweis (durch Widerspruch). Nehmen wir an, daß f auf einem Intervall $[a, b]$ z. B. nicht nach oben beschränkt ist, also $\sup_{x \in [a, b]} f(x) = \infty$ gilt. Wir wählen dann eine Folge $x_j \in [a, b]$ mit $\lim_{j \rightarrow \infty} f(x_j) = \infty$. Nach Übergang zu einer Teilfolge können wir nach Investition des Satzes von Bolzano-Weierstraß annehmen, daß die Folge (x_j) gegen einen Wert $\xi \in [a, b]$ konvergiert. Wir können des Weiteren ohne Beschränkung der Allgemeinheit annehmen, daß $\xi > a$ ist und daß die Folge der x_j *streng monoton aufsteigend* gegen ξ geht. Wir betrachten nun die Folge \mathcal{Z}_n der n -fachen äquidistanten Zerlegungen von $[a, b]$. Für \mathcal{Z}_n bezeichne J_n dasjenige eindeutig bestimmte Teilintervall, in dem ξ liegt, aber nicht als linker Endpunkt. Wir wählen dann zur Zerlegung \mathcal{Z}_n eine beliebige Riemann-Summe S'_n für alle Teilintervalle außer für das Teilintervall J_n . Nach Konstruktion von ξ gibt es dann ein $x_j \in J_n$, so daß

$$f(x_j) \frac{b-a}{n} \geq n - S'_n.$$

Dann ist

$$S_n := f(x_j) \frac{b-a}{n} + S'_n \geq n$$

eine Folge Riemannscher Summen von f auf dem Intervall $[a, b]$, die nach oben unbeschränkt ist. Dies widerspricht aber der Riemann-Integrierbarkeit von f . \square

Wir können nun den Beweis von Satz 11 in Angriff nehmen und beginnen mit der Inklusion $\text{Riem}(I) \subset \text{Darb}(I)$. In diesem Fall können wir uns offensichtlich auf die Situation zurückziehen, daß wir von vornherein $I = [a, b]$ haben. Eine Riemann-integrierbare Funktion f ist dann auf I und somit auch auf jedem Teilintervall von I bzgl. einer Zerlegung \mathcal{Z} beschränkt. Es bezeichne M_1, \dots, M_r die jeweiligen Maxima und entsprechend m_1, \dots, m_r die Minima. Werden dann bei vorgegebenem $\varepsilon > 0$ die Stützstellen ξ_1, \dots, ξ_r so gewählt, daß $M_j - f(\xi_j) < \varepsilon/4(b-a)$, so unterscheidet sich die zugeordnete Riemannsche Summe von dem Integral der durch die M_j definierten Treppenfunktion $\psi \geq f$ um höchstes $\varepsilon/4$. Ganz entsprechend findet man die Treppenfunktion $\varphi \leq f$, deren Integral sich ebenfalls von einer geeigneten Riemannschen Summe höchstens um $\varepsilon/4$ unterscheidet. Ist aber die Feinheit der benutzten Zerlegung \mathcal{Z} klein genug, so unterscheidet sich das Riemann-Integral von f auf $[a, b]$ von jeder Riemannschen Summe ebenfalls um höchstens $\varepsilon/4$. Somit gilt $\int_a^b (\psi(x) - \varphi(x)) dx < \varepsilon$ und f ist als Darboux-integrierbar entlarvt.

Es bleibt noch, die Inklusion $\text{Darb}(I) \subset \text{Riem}(I)$ zu begründen. Ist f Darboux-integrierbar, so gibt es zu dem Intervall $[a, b]$ und zu $\varepsilon > 0$ Treppenfunktionen $\varphi \leq f \leq \psi$ auf $[a, b]$ mit Integraldifferenz höchstens gleich $\varepsilon/2$. Wir können ohne Einschränkung annehmen, daß φ und ψ bzgl. der gleichen Zerlegung $a = t_0 < t_1 < \dots < t_{m-1} < t_m = b$ definiert sind. Wir setzen dann

$$\delta = \delta(\varepsilon) := \frac{\varepsilon}{8mM},$$

wobei M eine positive obere Schranke von $|f|_{[a,b]}$ bezeichnet, und zeigen, daß jede Riemann-Summe $S = S(f, \mathcal{Z}, (\xi_k))$ zu einer Zerlegung $\mathcal{Z} : a = x_0 < x_1 < \dots < x_{n-1} < x_n = b$ mit Feinheit $\delta(\mathcal{Z}) \leq \delta$ das Darboux-Integral von f über $[a, b]$ bis auf ε genau approximiert. Es sei Φ die Treppenfunktion mit $\Phi(x) = f(\xi_k)$ für $x_{k-1} < x < x_k$ und $\Phi(x_k) = f(x_k)$; sie besitzt das Integral S und erfüllt die Bedingung $\varphi(x) - 2M \leq \Phi(x) \leq \psi(x) + 2M$. Ist ferner das Intervall $[x_{k-1}, x_k]$ in einem der Intervalle (t_{k-1}, t_k) enthalten, so ist dort sogar $\varphi(x) \leq f(x) \leq \psi(x)$. Auf der Vereinigung dieser Intervalle definieren wir die Treppenfunktion Ψ durch Null, und sonst sei $\Psi(x) = 2M$ gesetzt. Dann ist für alle $x \in [a, b]$

$$\varphi(x) - \Psi(x) \leq \Phi(x) \leq \psi(x) + \Psi(x).$$

Nun gibt es höchstens $2m$ Intervalle $[x_{k-1}, x_k]$, auf denen Ψ von Null verschieden ist. Folglich ist

$$\int_a^b \Psi(x) dx \leq (2M) \cdot (2m\delta) \leq \frac{\varepsilon}{2}.$$

Daraus folgt

$$\int_a^b \varphi(x) dx - \frac{\varepsilon}{2} \leq S = \int_a^b \Phi(x) dx \leq \int_a^b \psi(x) dx + \frac{\varepsilon}{2}.$$

Nach Wahl von φ und ψ ist aber

$$\int_a^b f(x) dx - \int_a^b \varphi(x) dx \leq \int_a^b (\psi(x) - \varphi(x)) dx \leq \frac{\varepsilon}{2}$$

und entsprechend

$$\int_a^b \psi(x) dx - \int_a^b f(x) dx \leq \frac{\varepsilon}{2}.$$

Zusammen mit den Ungleichungen davor gewinnt man die gesuchte Ungleichung

$$\left| \int_a^b f(x) dx - S \right| \leq \varepsilon. \quad \square$$

Wir fragen uns sogar noch etwas allgemeiner: Wann ist für einen vorgegebenen Funktionenraum \mathcal{F} von integrierbaren Funktionen auf dem Intervall I der Darboux-Abschluß gleich dem Raum der Riemann-integrierbaren Funktionen? Dazu muß der Funktionenraum \mathcal{F} einige notwendige Bedingungen erfüllen, die wir zunächst herleiten wollen. Aus den Inklusionen $\mathcal{F} \subset \text{Darb-}\mathcal{F} \subset \text{Riem}(I)$ folgt, daß die Funktionen in \mathcal{F} notwendig Riemann-integrierbar sein müssen. Gilt umgekehrt $\text{Riem}(I) \subset \text{Darb-}\mathcal{F}$, so müssen insbesondere die Treppenfunktionen im Darboux-Abschluß von \mathcal{F} liegen. Dies führt zu der folgenden Bedingung:

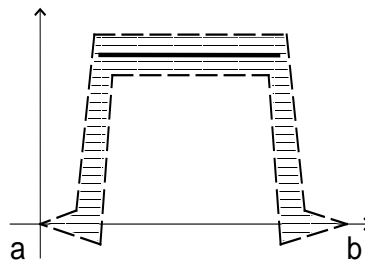
(+) Zu jedem kompakten Intervall $I' \subset [a, b]$ und zu jedem $\varepsilon > 0$ gibt es Funktionen $\varphi, \psi \in \mathcal{F}$, so daß $\varphi(x) \leq 1 \leq \psi(x)$, $x \in I'$, $\varphi(x) \leq 0 \leq \psi(x)$, $x \in [a, b] \setminus I'$, und

$$\int_a^b (\psi(x) - \varphi(x)) dx < \varepsilon.$$

Lemma 18.13 *Der Funktionenraum \mathcal{F} erfüllt die Bedingung (+) genau dann, wenn $\mathcal{T}(I) \subset \text{Darb-}\mathcal{F}$.*

Den leichten *Beweis* überlassen wir dem Leser. □

Bemerkung. Man kann die Bedingung (+) für viele Klassen von Riemann-integrierbaren, insbesondere von stetigen Funktionen nachprüfen. Sie ist nach dem vorstehenden Lemma schon dann erfüllt, wenn die Treppenfunktionen im Raum \mathcal{F} enthalten sind. Sie ist aber auch erfüllt für den Raum $\mathcal{F} = \mathcal{C}^0(I)$ der stetigen Funktionen auf I und den Unterraum $\mathcal{C}_{00}^0(I) \subset \mathcal{C}^0(I)$ der stetigen Funktionen mit kompaktem Träger, also der stetigen Funktionen $f : I \rightarrow \mathbb{R}$, für die es ein kompaktes Intervall $[a, b] \subset I$ gibt mit $f|_{I \setminus [a, b]} \equiv 0$. Der Grund hierfür sollte aus der folgenden Zeichnung deutlich werden. Man kann diese Räume auch noch verkleinern durch die Forderung, daß die Funktionen in \mathcal{F} eine gewisse Differenzierbarkeitsordnung besitzen oder sogar beliebig oft differenzierbar sein sollen.



Figur 18.1

Wir können nun beweisen:

Satz 18.14 *Der Darboux-Abschluß des Funktionenraumes \mathcal{F} ist genau dann der Raum der Riemann-integrierbaren Funktionen auf I , wenn $\mathcal{F} \subset \text{Riem}(I)$ die Bedingung (+) erfüllt.*

Beweis. Die Notwendigkeit der Bedingungen haben wir schon erläutert. Ist die Bedingung (+) erfüllt, so ist nach dem vorstehenden Lemma der Raum der Treppenfunktionen enthalten in dem Darboux-Abschluß bzgl. \mathcal{F} . Also ist auch der Raum der Darboux-integrierbaren Funktionen enthalten in dem Darboux-Abschluß von \mathcal{F} , der selbst wieder aufgrund der ersten Voraussetzung in $\text{Riem}(I)$ enthalten ist. Wegen Satz 11 gilt überall Gleichheit. \square

Es sollte dem Leser nicht verborgen geblieben sein, daß der mehrfach in diesem Kapitel gebrauchte Begriff des *Abschlusses* eine topologische Bedeutung hat, die wir anschließend noch etwas genauer beleuchten werden. Insbesondere überlassen wir ihm den einfachen Nachweis der folgenden Aussagen, die zur Folge haben, daß wir sowohl den Regelabschluß als auch den Darboux-Abschluß nur einmal durchführen können.

Satz 18.15 *Eine Funktion $f : I \rightarrow \mathbb{R}$ ist genau dann eine Regelfunktion, wenn sie sich auf jedem kompakten Intervall beliebig genau in der Supremumsnorm durch Regelfunktionen approximieren läßt. Sie liegt genau dann im Darboux-Abschluß von \mathcal{F} , wenn es zu jedem kompakten Intervall $[a, b] \subset I$ und jedem $\varepsilon > 0$ Funktionen $\varphi, \psi \in \text{Darb-}\mathcal{F}$ gibt mit $\varphi \leq f \leq \psi$ auf $[a, b]$ und*

$$\int_a^b (\psi(x) - \varphi(x)) dx < \varepsilon. \quad \square$$

Wir wollen nun den Aufbau des (verallgemeinerten) Darboux-Integrals noch etwas umformulieren. Dazu beginnen wir wieder mit einem abstrakten linearen Integral \mathcal{I}_a^b auf einem Vektorraum \mathcal{F} von Funktionen $f : I \rightarrow \mathbb{R}$, setzen aber diesmal zusätzlich voraus, daß mit $f \in \mathcal{F}$ auch $|f|$ in \mathcal{F} enthalten ist, also ein Vektorverband vorliegt, und I ein *kompaktes* Intervall $[a, b]$ ist. Wir setzen für eine beliebige Funktion $f : [a, b] \rightarrow \mathbb{R}$:

$$\|f\|_1^{\mathcal{F}} := \inf \left\{ \int_a^b \varphi(x) dx : |f| \leq \varphi, \varphi \in \mathcal{F} \right\}.$$

Mit anderen Worten: $\|f\|_1^{\mathcal{F}}$ ist per definitionem nichts anderes als das \mathcal{F} -Oberintegral der Funktion $|f|$.

Lemma 18.16 *Die Zuordnung*

$$\begin{cases} \text{Abb}(I, \mathbb{R}) \longrightarrow \mathbb{R}_+ \cup \{\infty\} \\ f \longmapsto \|f\|_1^{\mathcal{F}} \end{cases}$$

definiert auf dem Vektorraum $\text{Abb}(I, \mathbb{R})$ der reell-wertigen Funktionen auf I eine Pseudo-Halbnorm.

Bemerkung und Definition. Unter einer *Pseudo-Halbnorm* auf einem Vektorraum V wollen wir eine Abbildung $\|\cdot\| : V \rightarrow \mathbb{R}_+ \cup \{\infty\}$ verstehen, die alle Axiome einer *Pseudonorm* erfüllt außer, daß $\|f\| = 0$ nur für $f = 0$ gelten darf. Selbstverständlich soll aber $\|0\| = 0$ sein.

Beweis. Es ist, wegen $0 \in \mathcal{F}$, auch $\|0\|_1^{\mathcal{F}} = 0$ und damit $\|0f\|_1^{\mathcal{F}} = 0$ für alle $f \in \text{Abb}(I, \mathbb{R})$. Wir können daher im folgenden für die Betrachtung der Vielfachen λf stets $\lambda \neq 0$ voraussetzen. Dann ist aber

$$\{\psi \in \mathcal{F} : |\lambda f| \leq \psi\} = \{\lambda \varphi : \varphi \in \mathcal{F} \text{ und } |f| \leq \varphi\}$$

und folglich

$$\|\lambda f\|_1^{\mathcal{F}} := \inf \left\{ \lambda \int_a^b \varphi(x) dx : |f| \leq \varphi, \varphi \in \mathcal{F} \right\}.$$

Nun macht man sich aber sofort klar, daß für eine Menge A nicht negativer reeller Zahlen und $\lambda > 0$ die einfache Beziehung $\inf(\lambda A) = \lambda \inf A$ besteht. Infolgedessen ist

$$\|\lambda f\|_1^{\mathcal{F}} := \lambda \inf \left\{ \int_a^b \varphi(x) dx : |f| \leq \varphi, \varphi \in \mathcal{F} \right\} = \lambda \|f\|_1^{\mathcal{F}}.$$

Zum Nachweis der Dreiecksungleichung braucht man nur zu beachten, daß aus $|f_1| \leq \varphi_1$ und $|f_2| \leq \varphi_2$ die Ungleichung $|f_1 + f_2| \leq \varphi_1 + \varphi_2$ folgt. Also ist

$$\{\varphi_1 + \varphi_2 : |f_1| \leq \varphi_1, |f_2| \leq \varphi_2, \varphi_1, \varphi_2 \in \mathcal{F}\} \subset \{\psi \in \mathcal{F} : |f_1 + f_2| \leq \psi\}.$$

Damit ist die in Rede stehende Pseudo-Halbnorm von $f_1 + f_2$ höchstens so groß wie das Infimum über die Menge aller Integrale von Funktionen aus der links stehenden Menge. Diese Menge A von Integralen ist aber eine Summe $A_1 + A_2$ der entsprechenden Mengen von Integralen für die Funktionen f_1 bzw. f_2 , und wegen der leicht zu verifizierenden Beziehung

$$\inf(A_1 + A_2) = \inf A_1 + \inf A_2$$

ergibt sich

$$\|f_1 + f_2\|_1^{\mathcal{F}} \leq \|f_1\|_1^{\mathcal{F}} + \|f_2\|_1^{\mathcal{F}}. \quad \square$$

Bemerkung. Auch bei *Pseudo-Halbnormen* hat man einen Konvergenzbegriff, der im Wesentlichen alle Merkmale wie auf dem uns schon vertrauten Terrain der (Pseudo-) Normen aufweist. Allerdings geht eine ganz entscheidende Eigenschaft verloren: Die *Eindeutigkeit* von Grenzwerten. Dies ist allerdings ebenso wenig dramatisch wie das Auftreten des Norm-Wertes ∞ im Falle von Pseudonormen, den wir durch Einschränkung auf den Unterraum der *beschränkten* Elemente wegdiskutieren konnten. Im Falle von *Halb-Normen* muß man dagegen zu *Klassen* von Elementen übergehen, deren Differenz die Halbnorm 0 besitzt. Man könnte dies alles leicht im abstrakten Rahmen abhandeln; bei der LEBESGUESchen Integrationstheorie wird uns dieses Phänomen noch einmal konkret begegnen.

Die entscheidende Beobachtung ist nun, daß der Raum der Darb- \mathcal{F} -integrierbaren Funktionen nichts anderes ist als der *Abschluß* von \mathcal{F} im Raum $\text{Abb}(I, \mathbb{R})$ bzgl. dieser Pseudo-Halbnorm.

Lemma 18.17 *Eine Funktion $f \in \text{Abb}(I, \mathbb{R})$ liegt genau dann in Darb- \mathcal{F} , wenn es zu jedem $\varepsilon > 0$ eine Funktion $\varphi \in \mathcal{F}$ gibt, so daß*

$$\|f - \varphi\|_1^{\mathcal{F}} \leq \varepsilon.$$

Ist $\varphi_j \in \mathcal{F}$ eine Folge mit

$$\lim_{j \rightarrow \infty} \|f - \varphi_j\|_1^{\mathcal{F}} = 0,$$

so ist

$$\int_a^b f(x) dx = \lim_{j \rightarrow \infty} \int_a^b \varphi_j(x) dx.$$

Beweis. Nur der erste Teil bedarf der Begründung. Nach Definition ist $f \in \text{Darb-}\mathcal{F}$, wenn es zu $\varepsilon > 0$ Funktionen $\varphi, \psi \in \mathcal{F}$ gibt mit $\varphi \leq f \leq \psi$ und $\int_a^b (\psi - \varphi) dx \leq \varepsilon$. Dann ist aber $0 \leq f - \varphi \leq \psi - \varphi$, $\psi - \varphi \in \mathcal{F}$, und folglich

$$\|f - \varphi\|_1^{\mathcal{F}} \leq \int_a^b (\psi(x) - \varphi(x)) dx \leq \varepsilon.$$

Ist umgekehrt

$$\|f - \varphi_1\|_1^{\mathcal{F}} \leq \frac{\varepsilon}{3},$$

so gibt es eine Funktion $\varphi_2 \in \mathcal{F}$ mit $|f - \varphi_1| \leq \varphi_2$ und $\int_a^b \varphi_2(x) dx \leq \varepsilon/2$. Damit gilt

$$\varphi := \varphi_1 - \varphi_2 \leq f \leq \varphi_1 + \varphi_2 =: \psi$$

und

$$\int_a^b (\psi(x) - \varphi(x)) dx = 2 \int_a^b \varphi_2(x) dx \leq \varepsilon. \quad \square$$

Folgerung 18.18 *Ist \mathcal{F} ein Vektorverband, so ist mit f auch $|f|$ Darboux-integrierbar bzgl. \mathcal{F} .*

Beweis. Es gilt $\| |f| - |\varphi| \|_1^{\mathcal{F}} \leq \|f - \varphi\|_1^{\mathcal{F}}$. □

Wir können nun leicht einsehen, daß die Erweiterung des Integrals von Treppenfunktionen auf sowohl die Regelfunktionen als auch die Darboux-integrierbaren Funktionen dem gleichen Prinzip folgt. Wir werden dieses Prinzip in abstrakter Form herausarbeiten und dann anschließend bei der Einführung des LEBESGUESchen *Integrals* erneut zum Einsatz bringen. Im Fall der Regelfunktionen verfügen wir über den Konvergenzbegriff der *gleichmäßigen* Konvergenz auf dem kompakten Intervall $I = [a, b]$, der von der Supremumsnorm $\|\varphi\|_\infty$ induziert wird. Für das lineare Funktional des Integrals auf dem Raum \mathcal{F} der integrierbaren Funktionen φ gilt dann

$$\left| \int_a^b \varphi(x) dx \right| \leq \|\varphi\|_\infty (b - a).$$

Beim Darboux-Integral dagegen hat man die Pseudo-Halbnorm $\|\varphi\|_1^{\mathcal{F}}$. Ist \mathcal{F} ein Vektorverband, so hat man wegen $-\varphi \leq |\varphi| \leq \varphi$ sofort auch die noch einfachere Beziehung

$$\left| \int_a^b \varphi(x) dx \right| \leq \int_a^b |\varphi(x)| dx = \|\varphi\|_1^{\mathcal{F}}.$$

Da außerdem $\int_a^b |\varphi(x)| dx \leq \|\varphi\|_\infty (b - a)$ ist, impliziert die gleichmäßige Konvergenz die Konvergenz im Darboux'schen Sinne, d. h. wir sehen noch einmal in diesem Fall, daß $\text{Reg-}\mathcal{F} \subset \text{Darb-}\mathcal{F}$.

In beiden Fällen geht es also insbesondere darum, ein lineares Funktional von einem Unterraum eines Vektorraums mit einer Pseudo-Halbnorm auf seinen Abschluß fortzusetzen, wobei das Funktional seinem Betrage nach ebenfalls durch die Pseudo-Halbnorm abgeschätzt werden kann. Wir wollen die zugrundeliegenden Begriffe nunmehr herausarbeiten.

Es stellt sich zunächst die Frage, ob *jede* lineare Abbildung von einem Untervektorraum V_0 eines Vektorraumes V (nach einem Vektorraum W) *eindeutig* fortsetzbar ist zu einer linearen Abbildung $V_1 \rightarrow W$, wobei $V_1 \supset V_0$ ein beliebig vorgegebener Untervektorraum von V ist. Nach dem *allgemeinen*

Basisergänzungssatz ist eine Fortsetzung immer möglich, wobei jedoch die Fortsetzung im Falle $V_1 \neq V_0$ niemals eindeutig bestimmt ist. Es muß also hier noch mehr im Spiel sein. Tatsächlich ist in unserer Situation V ein (pseudo-halb-) normierter Vektorraum, und V_1 ist der Abschluß von V_0 in V bzgl. der von der Pseudo-Halbnorm induzierten Topologie. Ferner ist $W = \mathbb{R}$ ein vollständig normierter \mathbb{R} -Vektorraum. Überdies ist, was wegen der Eigenschaft 2 des Integrals unmittelbar klar ist, \mathcal{I} eine stetige Linearform, und was wir tatsächlich suchen, ist eine stetige Fortsetzung nach $V_1 = \overline{V_0}$. - Wir müssen daher zunächst klären, wann eine Linearform stetig ist. Wie wir wissen, ist dies im Endlich-dimensionalen immer richtig. Der Beweis von Satz 17.1 überträgt sich auch auf den Fall von pseudo-halbnormierten Vektorräumen.

Satz 18.19 *Es seien V und W pseudo-halbnormierte Vektorräume, $F : V \rightarrow W$ sei eine lineare Abbildung. Dann sind die folgenden Aussagen äquivalent :*

- i) F ist stetig in 0;
- ii) F ist stetig ;
- iii) F ist Lipschitz-stetig, d. h. es gibt eine Zahl $L > 0$, s. d. $\|F(v) - F(w)\| \leq L \|v - w\|$ für alle $v, w \in V$;
- iv) F ist gleichmäßig stetig ;
- v) F ist beschränkt auf der Einheitskugel $\{v \in V : \|v\| \leq 1\}$;
- vi) es gibt eine Zahl $C > 0$, s. d. $\|F(v)\| \leq C \|v\|$ für alle $v \in V$.

Wir benötigen jetzt noch zwei einfache Sätze.

Satz 18.20 *Es sei V_0 ein Untervektorraum eines pseudo-halbnormierten Vektorraums V . Dann ist auch der Abschluß $V_1 = \overline{V_0}$ von V_0 in V ein Untervektorraum.*

Beweis. Es ist $v \in V_1$ genau dann, wenn $v = \lim v_j$ für eine Folge $v_j \in V_0$. Aus den algebraischen Rechenregeln folgt die Behauptung. \square

Satz 18.21 *Es seien V ein pseudo-halbnormierter und W ein vollständig normierter Vektorraum, $V_0 \subset V$ sei ein Untervektorraum, und V_1 sei der Abschluß von V_0 in V . Dann läßt sich jede stetige lineare Abbildung $F_0 : V_0 \rightarrow W$ eindeutig zu einer stetigen linearen Abbildung $F_1 : V_1 \rightarrow W$ fortsetzen. Ist C eine Schranke für F_0 , so auch für F_1 .*

Beweis. Es sei $v \in V_1 = \overline{V_0}$, also $v = \lim v_j$, $v_j \in V_0$. Ist C eine Schranke für F_0 , so ist für $\varepsilon > 0$ mit $\delta = \delta(\varepsilon/C)$

$$\|F_0(v_j) - F_0(v_k)\| = \|F_0(v_j - v_k)\| < \frac{\varepsilon}{C} C = \varepsilon, \quad \|v_j - v_k\| < \delta.$$

Also ist die Folge $(F_0(v_j))_{j \in \mathbb{N}}$ eine Cauchy-Folge in W und damit konvergent. Sei w der Grenzwert; wir setzen dann

$$F_1(v) := w$$

und müssen zeigen, daß $F_1(v)$ unabhängig von der Auswahl der Folge (v_j) ist und F_1 alle weiteren behaupteten Eigenschaften besitzt.

Ist z. B. (v'_j) eine weitere Folge mit $\lim v'_j = v$ und $w' = \lim F_0(v'_j)$, so wird wegen

$$\|w - w'\| \leq \|w - F_0(v_j)\| + \|F_0(v_j) - F_0(v'_j)\| + \|F_0(v'_j) - w'\|$$

und $\|F_0(v_j) - F_0(v'_j)\| \leq C \|v_j - v'_j\|$ die rechte Seite beliebig klein, und damit ist $w = w'$. Infolgedessen ist auch $F_1 = F_0$ auf V_0 .

Die *Linearität* von F_1 folgt einfach aus den Konvergenzsätzen: $v = \lim v_j$ und $v' = \lim v'_j$ implizieren

$$\alpha v + \beta v' = \lim(\alpha v_j + \beta v'_j)$$

und damit

$$\begin{aligned} F_1(\alpha v + \beta v') &= \lim F_0(\alpha v_j + \beta v'_j) = \lim(\alpha F_0(v_j) + \beta F_0(v'_j)) \\ &= \alpha \lim F_0(v_j) + \beta \lim F_0(v'_j) = \alpha F_1(v) + \beta F_1(v'). \end{aligned}$$

Schließlich ist mit $\|F_0(v_j)\| \leq C \|v_j\|$, $v = \lim v_j$, auch

$$\|F_1(v)\| = \|\lim F_0(v_j)\| \leq \lim C \|v_j\| \leq C \|v\|. \quad \square$$

Damit können wir sowohl das Regel- als auch das Darboux-, also das Riemann-Integral unter dem einheitlichen Aspekt des Abschlusses des Raums der Treppenfunktionen unter einer geeigneten (Pseudohalb-) Norm verstehen.

Um zu begründen, warum wir uns mit dem Riemann-Integral nicht zufrieden geben können, kehren wir noch einmal zu der Dirichlet-Funktion $\chi := \chi_{\mathbb{Q} \cap [0,1]}$ zurück. So problematisch es erscheinen mag, sie überhaupt als Funktion zu akzeptieren, ist sie doch von einer sehr einfachen Bauart, wenn wir den Prozeß der punktweisen Konvergenz von Funktionen mit berücksichtigen. Ist nämlich $Q \cap [0, 1] = \{a_0, a_1, a_2, \dots\}$ eine Abzählung und $A_n := \{a_j : j = 0, \dots, n\}$, so bilden die $\chi_n := \chi_{A_n}$ eine monoton aufsteigende Folge von Riemann-integrierbaren Funktionen mit

$$\int_0^1 \chi_n(x) dx = 0,$$

die punktweise gegen χ konvergiert. Nun gilt tatsächlich der folgende Satz, den wir nicht beweisen wollen, da uns die *LEBESGUE-Theorie* ein stärkeres Mittel an die Hand geben wird.

Satz 18.22 *Es seien $f, f_j : [a, b] \rightarrow \mathbb{R}$ Riemann-integrierbare Funktionen, und die Folge f_j sei punktweise monoton aufsteigend konvergent gegen f . Dann gilt*

$$\lim_{j \rightarrow \infty} \int_a^b f_j(x) dx = \int_a^b \lim_{j \rightarrow \infty} f_j(x) dx = \int_a^b f(x) dx.$$

Warnung. Die entsprechende Aussage ist falsch, wenn man nur punktweise Konvergenz (und nicht monotone Konvergenz) voraussetzt.

Unter den vorigen Voraussetzungen ist die Folge $I_j := \int_a^b f_j(x) dx$ der Integrale selbst monoton aufsteigend und daher schon konvergent, *wenn sie nach oben beschränkt ist*. Wir würden daher eigentlich erwarten, daß der *monotone Limes* $f = \lim_j f_j$ unter dieser leichten Zusatzvoraussetzung automatisch integrierbar ist. Das Lebesgue-Integral erfüllt uns diesen Wunsch.

Zur Einführung des Lebesgue-Integrals folgen wir dem von KÖNIGSBERGER gewählten Weg, der auf STONE zurückgeht. Wir betrachten hierbei grundsätzlich Funktionen, die auf ganz \mathbb{R} definiert sind: $f : \mathbb{R} \rightarrow \mathbb{R}$. Ist f a priori nur auf einer Teilmenge $A \subset \mathbb{R}$ definiert, so setzen wir sie stets *trivial* zu einer Funktion f^A nach ganz \mathbb{R} fort:

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

Um solche Funktionen überhaupt „einfangen“ zu können, muß man den Raum der Treppenfunktionen in gewisser Weise ersetzen durch den Raum der verallgemeinerten Treppenfunktionen mit höchstens abzählbar vielen Treppen.

Definition. Eine *Hüllreihe* Φ zu einer Funktion $f : \mathbb{R} \rightarrow \mathbb{R}$ ist eine unendliche Reihe

$$\Phi = \sum_{j=0}^{\infty} c_j \chi_j, \quad \chi_j := \chi_{I_j},$$

mit (evtl. auch zu einpunktigen Mengen entarteten) kompakten²⁸ Intervallen $I_j := [a_j, b_j]$ und reellen Zahlen $c_j \geq 0$, so daß

$$|f(x)| \leq \Phi(x) := \sum_{j:x \in I_j} c_j$$

für alle $x \in \mathbb{R}$. Der *Inhalt* $I(\Phi)$ der Hüllreihe Φ wird gegeben durch

$$I(\Phi) := \sum_{j=0}^{\infty} c_j (b_j - a_j) \in \mathbb{R}_+ \cup \{\infty\}.$$

Bemerkung und Definition. Jede Funktion $f : \mathbb{R} \rightarrow \mathbb{R}$ besitzt eine Hüllreihe, nämlich $\Phi_\infty := \sum_j j \chi_{[-j,j]}$. Damit existiert die Zahl

$$\|f\|_{L^1} = \inf \{ I(\Phi) : \Phi \text{ ist Hüllreihe zu } f \} \in \mathbb{R}_+ \cup \{\infty\}.$$

Man nennt sie die L^1 -Norm von f und schreibt meist auch kürzer $\|f\|_1$.

Diese Bezeichnung ist nicht voll gerechtfertigt. Es gilt nur, was wir dem Leser zur Überprüfung lassen, das folgende

Lemma 18.23 $\|f\|_{L^1}$ ist eine Pseudo-Halbnorm auf dem Vektorraum aller Funktionen $f : \mathbb{R} \rightarrow \mathbb{R}$.

Die Beziehung dieser Halbnorm zu der das Riemann-Integral beherrschenden Halbnorm ist schnell aufgeklärt.

Lemma 18.24 Es sei $I := [a, b]$ ein kompaktes Intervall und $f : I \rightarrow \mathbb{R}$ eine Funktion. Dann gilt

$$\|f^I\|_{L^1} \leq \|f\|_1^T.$$

Ist f eine Treppenfunktion auf $[a, b]$, so besteht hier sogar Gleichheit, und es ist

$$\|f^I\|_{L^1} = \int_a^b |f(x)| dx \leq \|f\|_\infty (b - a).$$

Beweis. Man kann jede Treppenfunktion mit nicht negativen Funktionswerten als eine Hüllreihe auffassen, deren Inhalt mit dem Integral der Treppenfunktion übereinstimmt. \square

Definition und Bemerkung. Eine Funktion $f : \mathbb{R} \rightarrow \mathbb{R}$ heißt *Lebesgue-integrierbar*, wenn es eine Folge von Treppenfunktionen φ_j gibt, so daß

$$\|f - \varphi_j\|_{L^1} \longrightarrow 0.$$

Da $\|\cdot\|_{L^1}$ eine Pseudo-Halbnorm ist, bildet die Menge der Lebesgue-integrierbaren Funktionen einen Vektorraum.

Wegen des vorstehenden Lemmas hat man die

²⁸Man kann hier genauso gut auch beschränkte offene Intervalle verwenden. Zu den Einzelheiten siehe die mehrdimensionale Integrationstheorie.

Folgerung 18.25 Ist $f : I = [a, b] \rightarrow \mathbb{R}$ Riemann-integrierbar, so ist die Fortsetzung f^I Lebesgue-integrierbar.

Zur Einführung des LEBESGUE-Integrals benötigen wir das folgende

Lemma 18.26 Ist f Lebesgue-integrierbar mit approximierender Folge φ_j von Treppenfunktionen, so ist die Folge

$$I_j := \int_{\mathbb{R}} \varphi_j(x) dx := \int_{a_j}^{b_j} \varphi_j(x) dx, \quad \varphi_j(x) = 0 \text{ außerhalb } [a_j, b_j],$$

der Integrale konvergent. Der Grenzwert der Folge I_j ist unabhängig von der Auswahl der Folge φ_j .

Beweis. Daß die Folge I_j der Integrale eine Cauchy-Folge ist und ihr Grenzwert nicht von der Auswahl der φ_j abhängt, ergibt sich unmittelbar aus der folgenden Ungleichung für Treppenfunktionen φ, ψ :

$$\left| \int \varphi dx - \int \psi dx \right| \leq \int |\varphi - \psi| dx = \|\varphi - \psi\|_1 \leq \|\varphi - f\|_1 + \|f - \psi\|_1. \quad \square$$

Definition und Bemerkung. Man setzt unter der vorigen Voraussetzung

$$\int_{\mathbb{R}} f(x) dx := \lim_{j \rightarrow \infty} \int \varphi_j(x) dx$$

und nennt diese Zahl das *Lebesgue-Integral* von f . Ist $f : [a, b] \rightarrow \mathbb{R}$ Riemann-integrierbar, so ist die Fortsetzung $f^I : \mathbb{R} \rightarrow \mathbb{R}$ Lebesgue-integrierbar mit

$$\int_{\mathbb{R}} f^I(x) dx = \int_a^b f(x) dx.$$

Selbstverständlich ist das Lebesgue-Integral ein lineares und monotonies Funktional auf dem Raum aller Lebesgue-integrierbaren Funktionen.

Weiter hat man die Verbandseigenschaft des Vektorraums der Lebesgue-integrierbaren Funktionen.

Satz 18.27 Mit f ist auch $|f|$ Lebesgue-integrierbar, und es gilt

$$\left| \int_{\mathbb{R}} f(x) dx \right| \leq \int_{\mathbb{R}} |f(x)| dx = \|f\|_1.$$

Beweis. Es sei φ_j eine Folge von Treppenfunktionen, die in der L^1 -Pseudo-Halbnorm gegen f konvergiert. Wegen $||f| - |\varphi_j|| \leq |f - \varphi_j|$ ist dann auch

$$\| |f| - |\varphi_j| \|_1 \leq \|f - \varphi_j\|_1.$$

Infolgedessen ist $|f|$ integrierbar mit

$$\left| \int_{\mathbb{R}} f(x) dx \right| = \left| \lim_{j \rightarrow \infty} \int_{\mathbb{R}} \varphi_j(x) dx \right| \leq \lim_{j \rightarrow \infty} \int_{\mathbb{R}} |\varphi_j(x)| dx = \int_{\mathbb{R}} |f(x)| dx.$$

Um die Gleichung auf der rechten Seite einzusehen, beachten wir die aus der Dreiecksungleichung für die L^1 -Pseudo-Halbnorm entspringenden Ungleichungen

$$\|f\|_1 - \|f - \varphi_j\|_1 \leq \|\varphi_j\|_1 \leq \|f\|_1 + \|f - \varphi_j\|_1.$$

Wegen

$$\|\varphi_j\|_1 = \int_{\mathbb{R}} |\varphi_j(x)| dx$$

implizieren diese, wie gewünscht, die Ungleichungen

$$\|f\|_1 \leq \int_{\mathbb{R}} |f(x)| dx \leq \|f\|_1. \quad \square$$

Im Gegensatz zu den Verhältnissen bei dem Darboux-Integral kann man für Produkte integrierbarer Funktionen nur zeigen:

Satz 18.28 Sind f, g Lebesgue-integrierbar, und ist g beschränkt, so ist auch das Produkt $f \cdot g$ Lebesgue-integrierbar.

Beweis. Es sei M eine positive obere Schranke für g . Wir wählen dann zu vorgegebenem $\varepsilon > 0$ eine Treppenfunktion φ mit $\|f - \varphi\|_1 \leq \varepsilon/2M$ und eine Treppenfunktion ψ mit $\|g - \psi\|_1 \leq \varepsilon/2\tilde{M}$, wobei \tilde{M} eine positive obere Schranke von $|\varphi|$ bezeichnet. Aus der Ungleichung $|fg - \varphi\psi| \leq |f - \varphi||g| + |\varphi||g - \psi|$ erschließt man

$$\|fg - \varphi\psi\|_1 \leq M\|f - \varphi\|_1 + \tilde{M}\|g - \psi\|_1 \leq \varepsilon. \quad \square$$

Ist $f: \mathbb{R} \rightarrow \mathbb{R}$ eine beliebige Funktion und $A \subset \mathbb{R}$ eine Teilmenge, so setzt man

$$\int_A f(x) dx := \int_{\mathbb{R}} f(x) \chi_A(x) dx,$$

sofern das rechts stehende Integral existiert. Dies ist aufgrund von Satz 27 immer dann richtig, wenn f Lebesgue-integrierbar und $A = [a, b]$ ein kompaktes Intervall ist. Es gilt dann selbstverständlich auch

$$\int_{[a,b]} f(x) dx + \int_{[b,c]} f(x) dx = \int_{[a,c]} f(x) dx.$$

Mit anderen Worten:

Folgerung 18.29 Das Lebesgue-Integral auf dem Raum der Lebesgue-integrierbaren Funktionen $f: \mathbb{R} \rightarrow \mathbb{R}$ ist tatsächlich ein Integral im Sinne von Kapitel 16.

Daß wir es sowohl in der Riemann- als auch in der Lebesgue-Theorie nur mit *Halbnormen* zu tun haben, also Grenzwerte nicht eindeutig bestimmt sind, äußert sich in der Tatsache, daß man Funktionen auf „nicht zu großen“ Teilmengen abändern kann, ohne den Wert des Integrals zu beeinflussen. Im Falle des Riemann-Integrals sind dies *endliche* Mengen. Im Falle des Lebesgue-Integrals können solche Mengen sogar *dicht* in \mathbb{R} liegen. Um dieses Phänomen genauer unter die Lupe nehmen zu können, notieren wir als Folgerung aus der Definition (man setze alle Treppenfunktionen $\varphi_j = 0$):

Folgerung 18.30 Funktionen f mit verschwindender L^1 -Halbnorm $\|f\|_{L^1} = 0$ sind Lebesgue-integrierbar mit verschwindendem Integral. Ist die Differenz zweier Funktionen von diesem Typ, so sind beide integrierbar oder beide nicht integrierbar, und im ersten Fall stimmen ihre Integrale überein.

Beispiel. Die charakteristische Funktion $\chi_{\mathbb{Q}}$ der rationalen Zahlen und damit auch die Dirichlet-Funktion $\chi_{\mathbb{Q} \cap [0,1]}$ sind Lebesgue-integrierbar (aber nicht Riemann-integrierbar).

Dies folgt sofort aus dem weiter unten formulierten Satz von Beppo Levi über die monotone Konvergenz, läßt sich aber auch unmittelbar mit der soeben notierten Folgerung begründen. - Wir notieren gleich etwas allgemeiner:

Lemma 18.31 Ist $A \subset \mathbb{R}$ eine höchstens abzählbare Teilmenge, so gilt

$$\|\chi_A\|_{L^1} = 0.$$

Beweis. Wir numerieren die Menge A durch: $A = \{x_0, x_1, \dots\}$, und wählen zu vorgegebenem $\varepsilon > 0$ kompakte Intervalle $I_j = [a_j, b_j]$ mit $x_j \in I_j$ und $(b_j - a_j) = \varepsilon/2^{j+1}$. Dann ist

$$A \subset \bigcup_{j=1}^{\infty} I_j \quad \text{und} \quad \chi_A \leq \Phi := \sum_{j=0}^{\infty} \chi_{I_j}.$$

Wegen $I(\Phi) = \varepsilon$ ist folglich $\|\chi_A\|_{L^1} \leq \varepsilon$. □

Der Beweis dieses Lemmas überträgt sich unmittelbar auf eine allgemeinere Situation.

Definition. Eine Teilmenge $A \subset \mathbb{R}$ heißt eine **LEBESGUE-Nullmenge** oder eine Menge vom *Lebesgue-Maß Null*, wenn es zu jedem $\varepsilon > 0$ abzählbar unendlich viele kompakte Intervalle $I_j = [a_j, b_j]$ gibt mit

$$A \subset \bigcup_{j=0}^{\infty} I_j \quad \text{und} \quad \sum_{j=0}^{\infty} (b_j - a_j) \leq \varepsilon.$$

Beispiele sind abzählbare Mengen, aber auch die **CANTOR-Menge**.

Das vorige Lemma ist also auch in der folgenden, sogar schärferen Form gültig:

Lemma 18.32 *Eine Teilmenge $A \subset \mathbb{R}$ ist genau dann eine Lebesguesche Nullmenge, wenn*

$$\|\chi_A\|_{L^1} = 0.$$

Ist $f : \mathbb{R} \rightarrow \mathbb{R}$ Lebesgue-integrierbar und A eine Menge vom Maß Null, so beweist man sofort, daß auch $\|f \chi_A\|_{L^1} = 0$ und damit

$$\int_A f(x) dx = 0.$$

Man kann sogar beweisen, daß jede beliebige Funktion über Nullmengen integriert werden können und das Integral Null besitzen und somit integrierbare Funktionen auf Nullmengen beliebig abgeändert werden können, ohne daß das Integral affektiert wird.

Wir kommen auf diese Fragen später noch einmal zurück, und stellen jetzt nur noch die wichtigsten Konvergenzeigenschaften des Lebesgue-Integrals zusammen. Die *Beweise* erbringen wir in *Analysis 3* gleich für die *mehrdimensionale* Integration. Die erste ist der nach **BEPPLO LEVI** benannte *Satz von der monotonen Konvergenz*.

Satz 18.33 *Es sei f_j eine monoton aufsteigende Folge integrierbarer Funktionen $f_j : \mathbb{R} \rightarrow \mathbb{R}$ derart, daß die Folge der Integrale*

$$\int_{\mathbb{R}} f_j(x) dx$$

nach oben beschränkt ist. Dann gibt es eine Teilmenge $A \subset \mathbb{R}$ vom Lebesgue-Maß 0, so daß die Funktionenfolge f_j auf $\mathbb{R} \setminus A$ punktweise nach oben beschränkt ist, also insbesondere dort eine Grenzfunktion f besitzt, und setzt man f irgendwie (z. B. durch $f(x) := 0, x \in A$) zu einer Funktion auf ganz \mathbb{R} fort, so ist f integrierbar mit

$$\int_{\mathbb{R}} f(x) dx = \lim_{j \rightarrow \infty} \int_{\mathbb{R}} f_j(x) dx.$$

Die Haupteigenschaft des Lebesgue-Integrals in dieser Hinsicht ist der von **LEBESGUE** selbst stammende *Satz über die majorisierte Konvergenz*.

Satz 18.34 Es sei f_j eine Folge integrierbarer Funktionen $f_j: \mathbb{R} \rightarrow \mathbb{R}$ derart, die fast überall, d. h. außerhalb einer Menge A vom Lebesgue-Maß 0, punktweise gegen eine Grenzfunktion f konvergiere. Ferner gebe es eine Funktion F mit $\|F\|_1 < \infty$, also z. B. eine integrierbare Funktion F , so daß fast überall $|f_j| \leq F$ gelte. Setzt man f irgendwie zu einer Funktion auf ganz \mathbb{R} fort, so ist f integrierbar mit

$$\int_{\mathbb{R}} f(x) dx = \lim_{j \rightarrow \infty} \int_{\mathbb{R}} f_j(x) dx .$$

Als Anwendung des Lebesgueschen Konvergenzsatzes beweisen wir noch die außerordentlich wichtigen Vertauschungssätze von Grenzwerten bei Integralen. Wir formulieren und beweisen einen dieser Sätze ohne Umschweife für Funktionen in n Veränderlichen, die von weiteren q Parametern abhängen dürfen, indem wir die Existenz des mehrdimensionalen Lebesgue-Integrals und die Richtigkeit des analogen Konvergenzsatzes voraussetzen. Diese fehlenden Beweisstücke werden später nachgeliefert.

Satz 18.35 Es sei eine Teilmenge $X \times T \subset \mathbb{R}^q \times \mathbb{R}^n$ und eine Funktion $f: X \times T \rightarrow \mathbb{R}$ gegeben, so daß für alle $x \in X$ die Integrale

$$F(x) = \int_T f(x, t) dt$$

existieren.

a) Ist zusätzlich $f(x, t)$ bei festem $t \in T$ eine stetige Funktion in x und ferner $\Phi(t)$, $t \in T$, eine integrierbare Funktion mit $|f(x, t)| \leq \Phi(t)$ für alle $(x, t) \in X \times T$, so ist auch $F(x)$ stetig auf X , d. h. es gilt

$$\lim_{\substack{x \rightarrow a \\ x \in X}} \int_T f(x, t) dt = \int_T \lim_{\substack{x \rightarrow a \\ x \in X}} f(x, t) dt .$$

b) Ist X offen in \mathbb{R}^n , f stetig partiell differenzierbar nach x_j bei festem t , und gilt

$$\left| \frac{\partial f}{\partial x_j}(x, t) \right| \leq \Phi(t), \quad (x, t) \in X \times T,$$

mit einer integrierbaren Funktion $\Phi(t)$, so ist F nach x_j stetig partiell differenzierbar, und es gilt

$$\frac{\partial}{\partial x_j} \int_T f(x, t) dt = \int_T \frac{\partial f}{\partial x_j}(x, t) dt .$$

Beweis. a) Es gelte $\lim_{k \rightarrow \infty} x_k = a$; man setzt dann $f_k(t) := f(x_k, t)$ und hat punktweise Konvergenz $\lim_{k \rightarrow \infty} f_k(t) = f(a, t)$. Ferner ist $|f_k(t)| \leq \Phi(t)$. Nach dem Satz von Lebesgue folgt

$$F(a) = \int_T f(a, t) dt = \int_T \lim_{k \rightarrow \infty} f_k(t) dt = \lim_{k \rightarrow \infty} \int_T f_k(t) dt = \lim_{k \rightarrow \infty} \int_T f(x_k, t) dt .$$

b) Wir setzen mit einer Nullfolge $h_k \neq 0$:

$$g_k(t) := \frac{f(a + h_k e_j, t) - f(a, t)}{h_k} .$$

Aus dem Mittelwertsatz folgt

$$|g_k(t)| \leq \Phi(t)$$

und damit die Existenz von

$$\begin{aligned} \frac{\partial F}{\partial x_j}(a) &= \lim_{k \rightarrow \infty} \frac{1}{h_k} \left(\int_T f(a + h_k e_j, t) dt - \int_T f(a, t) dt \right) = \lim_{k \rightarrow \infty} \int_T \frac{f(a + h_k e_j, t) - f(a, t)}{h_k} dt \\ &= \int_T \lim_{k \rightarrow \infty} \frac{f(a + h_k e_j, t) - f(a, t)}{h_k} dt = \int_T \frac{\partial f}{\partial x_j}(a, t) dt , \end{aligned}$$

was zu beweisen war. □

Bemerkung. Die Konklusion des Teils a) kann schon dann getroffen werden, wenn $T \subset \mathbb{R}$ ein *kompaktes Intervall* und f eine *stetige Funktion* auf $X \times T$ ist. Denn da die Aussage *lokal* bzgl. $a \in X$ ist, können wir ohne Einschränkung X als kompakte Teilmenge annehmen, so daß f auf $X \times T$ sogar *beschränkt* ist. Somit ist $\Phi := M \chi_T$ mit geeigneter Konstanten M eine Majorante wie im Satz. Entsprechende Bedingungen findet man auch für die Differenzierbarkeitsaussage. In diesen Fällen existieren alle Integrale auch im Riemannsches Sinne.

Die vorstehende Bemerkung findet eine unmittelbare Anwendung auf sogenannte *Doppelintegrale* oder noch allgemeiner auf *n-fach iterierte Integrale stetiger Funktionen*, die man als Ausgangspunkt in die *mehrdimensionale Integrationstheorie* verwenden kann, auch wenn wir später einem anderen Weg folgen werden.

Wir betrachten der Einfachheit halber im folgenden eine *stetige Funktion* in *zwei Veränderlichen* x, y auf einem *Produkt von kompakten Intervallen*:

$$f : [a, b] \times [c, d] \rightarrow \mathbb{R} .$$

Nach Satz 35 und der soeben festgehaltenen Bemerkung ist dann die Funktion

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

stetig auf $[c, d]$ und kann also ihrerseits integriert werden. Man bezeichnet

$$\int_c^d F(y) dy := \int_c^d \left(\int_a^b f(x, y) dx \right) dy$$

als ein *Doppelintegral*.

Unter den soeben gemachten Voraussetzungen existiert aber auch das Doppelintegral bei *Vertauschung* der Reihenfolge der Integrationen:

$$\int_a^b \left(\int_c^d f(x, y) dy \right) dx .$$

Für die *Integrationstheorie von stetigen Funktionen in mehreren Veränderlichen* ist zweifelsohne unerläßlich, daß sich unabhängig von der Reihenfolge der eindimensionalen Integrationen stets derselbe Wert einstellt.

Satz 18.36 *Ist $f : [a, b] \times [c, d] \rightarrow \mathbb{R}$ eine stetige Funktion, so gilt*

$$\int_c^d \left(\int_a^b f(x, y) dx \right) dy = \int_a^b \left(\int_c^d f(x, y) dy \right) dx .$$

Beweis. Man definiere

$$\Phi(x) := \int_c^d \left(\int_a^x f(\xi, y) d\xi \right) dy .$$

Es ist $\Phi(a) = 0$, und wegen Satz 35 ist Φ differenzierbar mit

$$\Phi'(x) = \int_c^d \frac{d}{dx} \left(\int_a^x f(\xi, y) d\xi \right) dy = \int_c^d f(x, y) dy .$$

Also ist mit dem Hauptsatz der Differential- und Integralrechnung

$$\int_c^d \left(\int_a^b f(x, y) dx \right) dy = \Phi(b) = \Phi(b) - \Phi(a) = \int_a^b \Phi'(x) dx = \int_a^b \left(\int_c^d f(x, y) dy \right) dx . \quad \square$$

Wir wollen dieses Kapitel beschließen mit einigen Anmerkungen zu dem *Hauptsatz der Differential- und Integralrechnung*. Dieser beantwortet in der üblichen Fassung *partiell* eine Frage, die wir bisher noch gar nicht gestellt hatten, die aber - auch historisch - eine der zentralen Motivationen für die Konstruktion des Integrals war: *Unter welchen Voraussetzungen kann man eine differenzierbare Funktion f aus ihrer Ableitung f' rekonstruieren?*

Angesichts der früher schon notierten Tatsache, daß Ableitungen von differenzierbaren Funktionen nicht notwendig sehr „angenehme“ Eigenschaften besitzen, aber auch nicht völlig willkürliche Funktionen sind, ist zu vermuten, daß die *präzise* Antwort auf diese Frage interessant sein sollte, wenn auch nicht einfach zu gewinnen. Immerhin wissen wir schon aufgrund des *Hauptsatzes*, daß das (Regel-) Integral diese Aufgabe löst, wenn f' *stetig* ist, denn dann ist ja

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

Es ist nicht schwer zu sehen, daß diese Aussage sinngemäß immer noch gilt, wenn f' zumindest eine *Regel-Funktion* ist. Allerdings kann i. A. nicht einmal das umfassendere *Riemann-Integral* das Problem erledigen, da es differenzierbare Funktionen $f : [0, 1] \rightarrow \mathbb{R}$ mit beschränkter, aber nicht Riemann-integrierbarer Ableitung f' gibt. VITO VOLTERRA konstruierte 1881 sogar ein Beispiel dieser Art, bei dem f' auf einer Teilmenge von positivem Maß unstetig ist. In seiner Thesis 1902 verwies LEBESGUE hierauf als Motivation zu seiner Integrationstheorie, die tatsächlich Differentiation und Integration wieder als inverse Operatoren in einer großen Klasse von Funktionen etabliert²⁹. Tatsächlich kann man als erste teilweise Antwort den folgenden Satz festhalten.

Satz 18.37 *Ist $f : [a, b] \rightarrow \mathbb{R}$ differenzierbar und besitzt die Ableitung f' eine Lebesgue-integrierbare Majorante, so ist sie selbst Lebesgue-integrierbar. Es gilt dann der Hauptsatz der Differential- und Integralrechnung :*

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

Einen *Beweis* für diese Behauptungen findet man in [11], Band 2, Satz 8.6. Wir begnügen uns hier mit dem Fall, daß die Ableitung *beschränkt* ist. Wir betrachten für hinreichend große $k \in \mathbb{N}$ die Funktionenfolge $f_k : [a, b] \rightarrow \mathbb{R}$ mit

$$f_k(x) := \begin{cases} \frac{f(x + 1/k) - f(x)}{1/k} & \text{für } x \in [a, b - 1/k] \\ 0 & \text{sonst.} \end{cases}$$

Dies sind *Regelfunktionen*; ihre Folge konvergiert *punktweise* auf $[a, b)$ gegen die Ableitungsfunktion f' . Aufgrund des Mittelwertsatzes der Differentialrechnung ist die Folge gleichmäßig beschränkt durch eine Schranke M zu $|f'|$. Folglich ist f' Lebesgue-integrierbar mit

$$\int_a^b f'(x) dx = \lim_{k \rightarrow \infty} \int_a^b f_k(x) dx.$$

Die Integrale auf der rechten Seite kann man aber weiter auswerten:

$$\int_a^b f_k(x) dx = k \left(\int_{a+1/k}^b f(x) dx - \int_a^{b-1/k} f(x) dx \right) = k \left(\int_{b-1/k}^b f(x) dx - \int_a^{a+1/k} f(x) dx \right).$$

Da f stetig ist, folgt aus dem klassischen Hauptsatz, daß

²⁹Als Literatur für die im folgenden teilweise nur zitierten, aber nicht bewiesenen Ergebnisse seien zusätzlich zu unserer Literaturliste genannt:

J. J. Benedetto, *Real Variable and Integration*. B. G. Teubner: Stuttgart 1976.

E. Hewitt, K. Stromberg, *Real and abstract analysis*. Springer: Berlin-Heidelberg-New York 1965.

D. E. Varberg, On absolutely continuous functions. Amer. Math. Monthly **72**, 831-841 (1965).

K. Floret, *Maß- und Integrationstheorie*. B. G. Teubner: Stuttgart 1981.

$$\int_a^b f'(x) dx = f(b) - f(a) . \quad \square$$

Bemerkung und Beispiel. Man darf allerdings nicht annehmen, daß die Ableitung einer Funktion stets Lebesgue-integrierbar oder gar beschränkt ist. Die Funktion

$$f(x) := \begin{cases} x^2 \sin \frac{1}{x^2} & \text{für } 0 < x \leq 1, \\ 0 & \text{für } x = 0, \end{cases}$$

besitzt tatsächlich eine nicht Lebesgue-integrierbare Ableitung.

Dagegen existiert die Ableitung der Funktion

$$f(x) := \begin{cases} x^{3/2} \sin \frac{1}{x} & \text{für } 0 < x \leq 1, \\ 0 & \text{für } x = 0, \end{cases}$$

überall auf $[0, 1]$, ist *unbeschränkt*, aber trotzdem Lebesgue-integrierbar.

Bemerkung und Beispiel. In Satz 37 darf man, wie es im Rahmen der Lebesgueschen Integrationstheorie nahezuliegen scheint, die Existenz der Ableitung f' nicht ersetzen durch die schwächere Voraussetzung, daß diese nur *fast überall* vorhanden ist. Wir erinnern daran (siehe Kapitel 15), daß die auf der CANTOR-Menge $C \subset I = [0, 1]$ definierte CANTOR-Funktion eine *monoton wachsende, stetige und surjektive* Fortsetzung $\Theta : I \rightarrow I$ besitzt, die auf jedem der abzählbar vielen disjunkten offenen Intervalle, in die $I \setminus C$ zerfällt, konstant ist. Folglich ist Θ fast überall differenzierbar mit $\Theta' = 0$ und daher

$$\int_0^1 \Theta'(x) dx = 0 < 1 = \Theta(1) - \Theta(0) .$$

Immerhin steht dieses Beispiel noch im Einklang mit dem LEBESGUESchen *Differenzierbarkeitssatz*.

Satz 18.38 *Ist die Funktion $f : [a, b] \rightarrow \mathbb{R}$ monoton wachsend, so existiert die Ableitung f' fast überall. Sie ist Lebesgue-integrierbar, und es gilt die Abschätzung*

$$\int_a^b f'(x) dx \leq f(b) - f(a) .$$

Definition und Bemerkung. Eine Funktion $f : [a, b] \rightarrow \mathbb{R}$ heißt von *beschränkter Variation*, wenn

$$\sup v(\mathcal{Z}) < \infty ,$$

wobei für eine Zerlegung $\mathcal{Z} : a = x_0 < x_1 < \dots < x_n = b$ die *Variation* $v(\mathcal{Z})$ definiert wird durch

$$v(\mathcal{Z}) := \sum_{j=1}^n |f(x_j) - f(x_{j-1})| .$$

Selbstverständlich ist eine *monotone* Funktion von beschränkter Variation.

Darüberhinaus gilt der JORDANSche *Zerlegungssatz*:

Satz 18.39 *Eine Funktion $f : [a, b] \rightarrow \mathbb{R}$ ist genau dann von beschränkter Variation, wenn sie Differenz zweier monoton wachsender Funktionen ist.*

Folgerung 18.40 *Funktionen $f : [a, b] \rightarrow \mathbb{R}$ von beschränkter Variation sind Lebesgue-integrierbar.*

Es ist nun zwingend erforderlich, nach Eigenschaften von f selbst zu fragen, die die Lebesgue-Integrierbarkeit von f' und damit die Gültigkeit des Hauptsatzes nach sich ziehen. Hierbei ist die ungemein *glättende* Eigenschaft des Integrals zu berücksichtigen

Definition und Bemerkungen. Eine Funktion $f : [a, b] \rightarrow \mathbb{R}$ heißt *absolut stetig*, falls für alle $\varepsilon > 0$ ein $\delta > 0$ existiert, so daß für jede endliche Menge von paarweise disjunkten Intervallen $[a_j, b_j]$, $j = 1, \dots, n$, in $[a, b]$ mit

$$\sum_{j=1}^n (b_j - a_j) \leq \delta$$

gilt:

$$\sum_{j=1}^n |f(b_j) - f(a_j)| \leq \varepsilon.$$

Absolut stetige Funktionen sind automatisch (gleichmäßig) stetig und von beschränkter Variation. Aus dem Mittelwertsatz der Differentialrechnung folgt unmittelbar, daß differenzierbare Funktionen mit beschränkter Ableitung absolut stetig sind. Es gibt aber stetige Funktionen, die keine beschränkte Variation besitzen.

Aufgrund der vorstehenden Bemerkungen ist das folgende Resultat, das ebenfalls auf VITALI und LEBESGUE zurückgeht, vielleicht keine Überraschung mehr.

Satz 18.41 *Es sei $f : [a, b] \rightarrow \mathbb{R}$ differenzierbar. Genau dann ist f' Lebesgue-integrierbar, wenn f absolut stetig ist. In diesem Fall gilt der Hauptsatz der Differential- und Integralrechnung :*

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

Ohne Beweis. □

In der Tat gilt viel mehr:

Satz 18.42 *Ist $f : [a, b] \rightarrow \mathbb{R}$ absolut stetig, so ist f fast überall differenzierbar, die Ableitung f' ist Lebesgue-integrierbar, und es gilt der Hauptsatz der Differential- und Integralrechnung :*

$$f(x) - f(a) = \int_a^x f'(t) dt, \quad x \in [a, b].$$

Ohne Beweis. □

Es bleibt aber noch, eine weitere Frage zu beantworten: Es sei $g : [a, b] \rightarrow \mathbb{R}$ Lebesgue-integrierbar und

$$G(x) := \int_a^x g(t) dt, \quad x \in [a, b].$$

An wievielen Stellen in $[a, b]$ ist die Funktion G differenzierbar? - Die (naheliegende) Antwort wird durch den folgenden Satz gegeben:

Satz 18.43 *Ist $g : [a, b] \rightarrow \mathbb{R}$ Lebesgue-integrierbar und*

$$G(x) := \int_a^x g(t) dt, \quad x \in [a, b],$$

so ist G absolut stetig und fast überall differenzierbar mit $G' = g$.

Ohne Beweis. □

Man kann nun die *absolut stetigen* Funktionen $f : [a, b] \rightarrow \mathbb{R}$ nach BANACH und ZARIZKI auf die folgenden äquivalenten Weisen charakterisieren, wobei wir f eine *Nullfunktion* nennen, wenn sie jede Lebesgue-Nullmenge $N \subset [a, b]$ in eine Lebesgue-Nullmenge überführt:

Satz 18.44 Für Funktionen $f : [a, b] \rightarrow \mathbb{R}$ sind äquivalent :

- i) f ist absolut stetig ;
- ii) f ist stetig, von beschränkter Variation und eine Nullfunktion ;
- iii) f ist eine stetige Nullfunktion und die Ableitung f' existiert fast überall und ist Lebesgue-integrierbar.

Unter einer und damit jeder dieser Voraussetzungen gilt der Hauptsatz der Differential- und Integralrechnung :

$$f(x) - f(a) = \int_a^x f'(t) dt, \quad x \in [a, b].$$

Ohne Beweis. □

Faßt man alle bisherigen Aussagen zusammen, so sieht man, daß *genau die absolut stetigen Funktionen dem Hauptsatz der Differential- und Integralrechnung genügen.*

Beispiel und Bemerkung. Es mag überraschen, daß es stetige Funktionen von beschränkter Variation geben soll, die keine Nullfunktionen sind. Wir kennen aber schon ein Beispiel, nämlich die Cantorsche Funktion $\Theta : [0, 1] \rightarrow [0, 1]$. Sie ist nicht nur stetig, sondern auch monoton wachsend und damit von beschränkter Variation. Da sie außerhalb der Cantor-Menge C nur abzählbar viele Werte annimmt, ist das Bild von $[0, 1] \setminus C$ unter Θ eine Nullmenge. Infolgedessen besitzt das Bild der Nullmenge C unter Θ das Maß 1.

Wir zitieren schließlich noch die *ultimative* Fassung der *Substitutionsformel*.

Satz 18.45 Ist $\alpha : J \rightarrow I$ eine monoton wachsende, absolut stetige (und damit fast überall differenzierbare) Abbildung zwischen reellen Intervallen, so ist mit $f : I \rightarrow \mathbb{R}$ auch $(f \circ \alpha) \cdot \alpha' : J \rightarrow \mathbb{R}$ (fast überall definiert und) Lebesgue-integrierbar, und es gilt für alle $a, b \in J$:

$$\int_a^b f(\alpha(t)) \cdot \alpha'(t) dt = \int_{\alpha(a)}^{\alpha(b)} f(x) dx.$$

Bemerkung. Es gibt noch weitere Integralbegriffe wie die von DENJOY, PERRON und seit kürzerem auch von KURZWEIL und HENSTOCK. Wir können in diesem grundlegenden Band nicht auf diese speziellen Integrale eingehen. Den besonders physikalisch interessierten Lesern sei aber vor allem eine andere Verallgemeinerung, nämlich die von STIELTJES, ans Herz gelegt.

Anhang: Die letzten 5 der 55 Charakterisierungen des reellen Zahlkörpers

Der definitiv letzte Kreis unserer Axiome beschäftigt sich mit der (Darboux'schen) *Integrationstheorie* und dem *Hauptsatz der Differential- und Integralrechnung*. Wir zeigen die Äquivalenz der folgenden Aussagen zu den früheren. Es ist stets f eine stetige Funktion $f : [a, b] \rightarrow \mathbb{K}$. Der Hauptsatz hat hierbei die folgende Formulierung: *Jede stetige Funktion $f : [a, b] \rightarrow \mathbb{K}$ besitzt eine Stammfunktion, und je zwei Stammfunktionen unterscheiden sich additiv um eine Konstante.*

$$(LI) \quad \begin{array}{c} \mathbb{K} \text{ erfüllt } (*) \\ \& \\ \text{Für stetiges } f \text{ existiert das} \\ \text{Oberintegral } \int_a^{\bar{b}} f(x) dx \end{array}$$

$$(LII) \quad \begin{array}{c} \mathbb{K} \text{ erfüllt } (*) \\ \& \\ \text{Für stetiges } f \text{ existiert das} \\ \text{Darboux'sche Integral } \int_a^b f(x) dx \end{array}$$

$$(LIII) \quad \begin{array}{c} \text{Für stetiges } f \geq 0 \text{ existiert eine Stammfunktion,} \\ \text{und jede Stammfunktion ist monoton wachsend} \end{array}$$

$$(LIV) \quad \begin{array}{c} \text{Hauptsatz der Differential- und Integralrechnung} \end{array}$$

Beweis. Aus dem Axiom (XLVI) folgt, daß es zu jeder stetigen Funktion $f : [a, b] \rightarrow \mathbb{K}$ Treppenfunktionen $\varphi, \psi : [a, b] \rightarrow \mathbb{K}$ mit $\varphi \leq f \leq \psi$ gibt. Damit ist die Menge

$$\left\{ \int_a^b \psi(x) dx : f \leq \psi \text{ Treppenfunktion} \right\}$$

nicht leer und nach unten durch

$$\int_a^b \varphi(x) dx$$

beschränkt. Also existiert wegen (XXI) das Oberintegral

$$\int_a^{\bar{b}} f(x) dx = \inf \left\{ \int_a^b \psi(x) dx : f \leq \psi, \psi \text{ Treppenfunktion} \right\}$$

und dann auch das Unterintegral wegen

$$\int_a^b f(x) dx = - \int_a^{\bar{b}} (-f)(x) dx .$$

Ist umgekehrt (LI) oder (LII) erfüllt, so ist jede stetige Funktion $f : [a, b] \rightarrow \mathbb{K}$ notwendig (nach oben) beschränkt, also das Axiom (XLVI) gültig und folglich $\mathbb{K} \cong \mathbb{R}$. Insbesondere gibt es wegen (L) zu jeder stetigen Funktion $f : [a, b] \rightarrow \mathbb{K}$ und jedem $\varepsilon > 0$ Treppenfunktionen φ, ψ mit $\varphi \leq f \leq \psi$ und $\sup_{[a,b]}(\psi - \varphi) \leq \frac{\varepsilon}{b-a}$. Somit ist

$$0 \leq \int_a^b \psi(x) dx - \int_a^b \varphi(x) dx \leq \varepsilon ,$$

und es ergibt sich mit (XXI) sofort

$$\int_a^{\bar{b}} f(x) dx = \int_a^b f(x) dx ,$$

also (noch einmal) die Implikation (LI) \implies (LII).

Von (LII) nach (LIII) schließt man wie folgt: Nach Voraussetzung existiert für alle $x \in [a, b]$ das Darboux'sche Integral

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

Mit dem üblichen Beweis ergibt sich die Differenzierbarkeit von F mit $F' = f$. Da das Integral additiv ist, gilt für $x_1 < x_2$:

$$F(x_2) - F(x_1) = \int_{x_1}^{x_2} f(x) dx = \int_a^{x_2} f(x) dx - \int_a^{x_1} f(x) dx \geq 0 ,$$

da $\varphi : [x_1, x_2] \rightarrow \mathbb{K}$ mit $\varphi(x) = 0$ eine Treppenfunktion mit $\varphi \leq f|_{[x_1, x_2]}$ ist. Ist G eine weitere Stammfunktion, so ist $(G - F)' = G' - F' = f - f = 0$ und wegen (XLII) $G - F = \text{const.}$ Insbesondere ist auch G monoton wachsend.

(LIII) \implies (LIV) Jede stetige Funktion f schreibt sich in der Form $f = f^+ - f^-$ mit den stetigen Funktionen

$$f^+ = \max(f, 0) , \quad f^- = -\min(f, 0) .$$

Da f^+ und f^- nichtnegativ sind, besitzen sie Stammfunktionen F^+ und F^- , und $F := F^+ - F^-$ ist eine Stammfunktion von f . Sei G eine weitere Stammfunktion von f , so ist $H := G - F$ eine Stammfunktion von 0. Also ist $H' = 0 \geq 0$ und $(-H)' = 0 \geq 0$. Nach (LIII) müssen dann sowohl H als auch $-H$ monoton aufsteigend sein, woraus sofort $H = \text{const.}$ folgt.

Ist schließlich (LIV) erfüllt, so folgt aus $f' = 0 = 0'$, daß $f = \text{const.}$ Dies ist aber gerade die Aussage (XLII). \square

Bemerkung. In [171] wurde Axiom (LII) in der Form

(LII)'	\mathbb{K} erfüllt (*) & Für stetiges f existiert das Riemannsche Integral $\int_a^b f(x) dx$
--------	--

formuliert, wobei aber in der Argumentation das Riemann-Integral mit dem Darboux-Integral gleichgesetzt wurde. Bei Verwendung der Definitionen des *vorliegenden* Textes ist die Äquivalenz mit diesem Axiom (LII)' daher nicht nachgewiesen, zumal unser Beweis für die Gleichheit der beiden Integral-Begriffe ganz erheblich von Eigenschaften der reellen Zahlen Gebrauch macht. Auf jeden Fall geht der Beweis (man benötigt nur Lemma 12) aber durch, wenn man den Körper \mathbb{K} in den Körper der reellen Zahlen einbetten kann, d. h. wenn \mathbb{K} *archimedisch* ist.

$$(LV) \quad \begin{array}{c} \mathbb{K} \text{ ist archimedisch} \\ \& \\ \text{Für stetiges } f \text{ existiert das} \\ \text{Riemannsches Integral } \int_a^b f(x) dx \end{array}$$

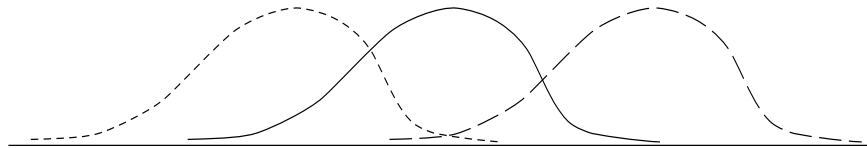
Ich weiß im Übrigen nicht, ob man bei den beiden den Hauptsatz betreffenden Axiomen die jeweiligen Zusätze fortlassen darf, also nur zu verlangen braucht, daß jede stetige Funktion f auf Intervallen $[a, b] \subset \mathbb{K}$ eine Stammfunktion besitzt.

19 Die schwingende Saite und Fourier-Reihen

Wir wollen zunächst über die *Wellengleichung* sprechen. Dazu stellen wir uns eine unendlich lange Saite (oder besser ein Seil) vor. Man kennt dann das Phänomen, daß sich *Wellen* mit konstanter Geschwindigkeit entlang des Seiles ausbreiten können (entsprechend Schall in Luft, Licht im Vakuum, etc.)

Es bezeichne zur Zeit $t = 0$ die Funktion $F(x)$ die Auslenkung an der Stelle $x \in \mathbb{R}$ in y -Richtung. Sei weiter die Ausbreitungsgeschwindigkeit (in positiver x -Richtung) von der Größe c . Dann wird der gesamte raumzeitliche Vorgang beschrieben durch die Funktion

$$y = \varphi(x, t) = F(x - ct).$$



Figur 19.1

Als Grundprinzip der Physik sollte gelten, daß alle dynamischen Vorgänge durch partielle Differentialgleichungen beschrieben werden. Nehmen wir also der Einfachheit halber an, daß F (stetig) differenzierbar ist. Dann gilt:

$$\begin{aligned} \frac{\partial \varphi}{\partial x} &= F'(x - ct) \frac{\partial(x - ct)}{\partial x} = F'(x - ct) \\ \frac{\partial \varphi}{\partial t} &= F'(x - ct) \frac{\partial(x - ct)}{\partial t} = -cF'(x - ct). \end{aligned}$$

Wir haben es also mit der partiellen Differentialgleichung

$$\frac{\partial \varphi}{\partial t} + c \frac{\partial \varphi}{\partial x} = 0$$

zu tun. Nun lehrt die Erfahrung, daß Wellen auch in negativer Richtung mit gleichem Geschwindigkeitsbetrag laufen können. Also sollten wir durch die Substitution $c \mapsto -c$ wieder eine Lösung erhalten, was aber nicht der Fall ist. Da wir alle Möglichkeiten durch ein *einziges* Gesetz erhalten wollen, differenzieren wir noch einmal und erhalten in beiden Fällen, wenn über die Lösung *hinreichend gute* Differenzierbarkeits-Voraussetzungen (nämlich zweimalige stetige Differenzierbarkeit) gemacht werden:

$$(*) \quad \frac{\partial^2 \varphi}{\partial t^2} = -c \frac{\partial}{\partial t} \frac{\partial \varphi}{\partial x} = -c \frac{\partial}{\partial x} \frac{\partial \varphi}{\partial t} = c^2 \frac{\partial^2 \varphi}{\partial x^2}.$$

Dies ist die *eindimensionale Wellengleichung*. In höheren Dimensionen hat man entsprechend die *Wellengleichung*

$$\frac{\partial^2 \varphi}{\partial t^2} = c^2 \Delta \varphi$$

mit dem *Laplace-Operator*

$$\Delta = \frac{\partial^2}{\partial x_1^2} + \dots + \frac{\partial^2}{\partial x_n^2}.$$

(Dies ist natürlich keine vernünftige physikalische Ableitung der Wellengleichung. Selbstverständlich ist auch der von uns vermittelte Eindruck falsch, daß die Welle *wandert*. Tatsächlich bewegen sich nur die einzelnen Teilchen *harmonisch* um eine Ruhelage).

Die Vorstellung, daß eine momentane Störung der Ruhelage sich bei Vorliegen von (*) wie eine Welle ausbreiten wird, drückt sich gut in den beiden folgenden Sätzen aus:

Satz 19.1 Jede zweimal stetig differenzierbare Lösung von (*) ist von der Gestalt

$$\varphi(x, t) = F(x - ct) + G(x + ct), \quad F, G \in \mathcal{C}^2(\mathbb{R}).$$

Satz 19.2 Jede Lösung ist vollständig bestimmt durch die Ausgangslage $\varphi(x, 0)$ und die Anfangsgeschwindigkeit $\frac{\partial \varphi}{\partial t}(x, 0)$.

Beweis (von Satz 1). Wir setzen zur Vereinfachung $c = 1$ und

$$\psi(\xi, \tau) = \varphi(\xi + \tau, \xi - \tau),$$

führen also die Substitution $x = \xi + \tau, t = \xi - \tau$ durch. Dann impliziert die Kettenregel:

$$\frac{\partial \psi}{\partial \xi} = \frac{\partial \varphi}{\partial x} \frac{\partial(\xi + \tau)}{\partial \xi} + \frac{\partial \varphi}{\partial t} \frac{\partial(\xi - \tau)}{\partial \xi} = \frac{\partial \varphi}{\partial x} + \frac{\partial \varphi}{\partial t}$$

und damit

$$\frac{\partial^2 \psi}{\partial \tau \partial \xi} = \frac{\partial^2 \varphi}{\partial x^2} - \frac{\partial^2 \varphi}{\partial t \partial x} + \frac{\partial^2 \varphi}{\partial x \partial t} - \frac{\partial^2 \varphi}{\partial t^2} = 0.$$

Dies bedeutet aber, daß ψ Summe zweier Funktionen in jeweils einer der Veränderlichen sein muß:

$$\psi = \tilde{F}(\tau) + \tilde{G}(\xi).$$

Macht man jetzt die obige Substitution rückgängig, so ergibt sich

$$\xi = \frac{1}{2}(x + t), \quad \tau = \frac{1}{2}(x - t) \quad \text{und damit} \quad \varphi(x, t) = F(x - t) + G(x + t). \quad \square$$

Beweis (von Satz 2). Aus $\varphi(x, t) = F(x - t) + G(x + t)$ ergibt sich $\varphi(x, 0) = F(x) + G(x)$ und

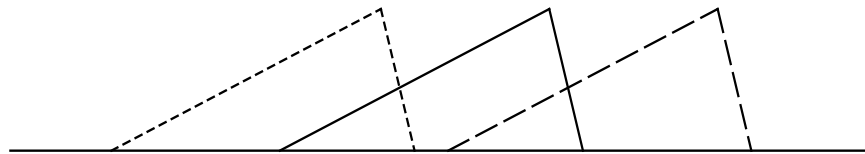
$$\frac{\partial \varphi}{\partial t}(x, t) = -F'(x - t) + G'(x + t),$$

also $\frac{\partial \varphi}{\partial t}(x, 0) = -F'(x) + G'(x)$. Damit sind $F + G$ und $-F + G$ (vermittels Integration bis auf additive Konstanten) durch die Anfangsbedingungen bestimmt, also auch F und G und damit φ . \square

Bemerkung. Man kann sich ohne Schwierigkeiten Wellenpakete vorstellen, die keine differenzierbare Gestalt haben. Diese können in einem strengen Sinne natürlich niemals Lösung einer Differentialgleichung wie die Wellengleichung sein. Die Theorie der *Distributionen* stellt aber den Begriff von *verallgemeinerten Lösungen* bereit, mit dem dann *alle* Funktionen der Gestalt

$$\varphi(x, t) = F(x - ct) + G(x + ct), \quad F, G \text{ stetig}$$

als Lösungen angesehen werden dürfen.

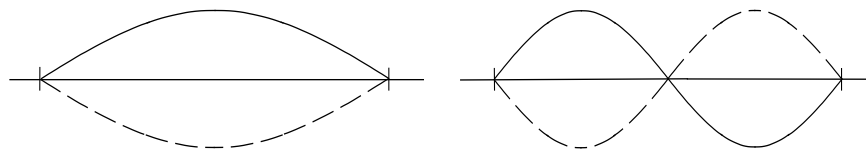


Figur 19.2

Spezielle Lösungen der eindimensionalen Wellengleichung sind offensichtlich die Funktionen $\varphi(x, t) = A \cos(\omega(x - ct))$, $\omega > 0$. Für festgehaltenes $x = x_0$ stellen sie harmonische Schwingungen mit der Frequenz $c\omega$ dar. Also ist *jede* Frequenz zulässig, wie wir auch aus Erfahrung wissen: Wir hören Töne beliebiger Höhe (mit Einschränkungen, die nicht prinzipieller Natur sind, sondern nur von unserer - leider auch altersabhängigen - Physiologie bestimmt werden), das Vakuum kann Licht aller Wellenlängen transportieren usw.

Dies alles ändert sich *dramatisch*, wenn man die Saite (das Seil, die Luftsäule) „einspannt“ wie z. B. bei einem *Musikinstrument* (Klavier, Geige, Flöte) ! Mathematisch bedeutet dies, daß wir jetzt *Randbedingungen* zu erfüllen haben, wodurch die Lösungen tatsächlich drastisch reduziert werden. Bei geeigneter Normierung können wir natürlich annehmen, daß diese Bedingungen von der folgenden Gestalt³⁰ sind:

$$\varphi(0, t) = \varphi(\pi, t) = 0 \quad \text{für alle } t \in \mathbb{R} .$$



Figur 19.3

Wir setzen dann wieder ohne Einschränkung $c = 1$ und versuchen unser Glück zunächst mit einem *Separationsansatz*

$$y = \varphi(x, t) = v(x)w(t), \quad v(0) = v(\pi) = 0 .$$

Lösungen von dieser Gestalt nennt man aus naheliegenden Gründen *synchrone Schwingungen* oder *stehende Wellen*. Aus dem Ansatz folgt wegen

$$\frac{\partial^2 y}{\partial t^2} = v(x)\ddot{w}(t), \quad \frac{\partial^2 y}{\partial x^2} = v''(x)w(t)$$

durch Differenzieren nach x bzw. t sofort die Beziehung (wir benutzen die NEWTONSche *Konvention*, Ableitungen nach der Zeit mit einem Punkt statt einem Strich zu kennzeichnen)

$$\frac{v''}{v} = \frac{\ddot{w}}{w} = \lambda$$

mit einer reellen Konstanten λ . Also erfüllen v und w die *Eigenwertgleichungen* für den Operator D^2 :

$$v'' + \lambda v = 0, \quad \ddot{w} + \lambda w = 0,$$

die nach früheren Überlegungen die Lösungen

$$v(x) = A \cos(\omega x - B), \quad w(t) = a \cos(\omega t - b)$$

mit der Konstanten $\omega = \sqrt{\lambda}$ besitzen. Dabei muß ω reell, d. h. λ positiv sein, denn anderenfalls „explodieren“ die Lösungen, was wegen der Randbedingungen nicht sein kann.

Wir geben für die letzte Feststellung noch ein physikalisches Argument: $v(x)$ ist nicht konstant, da sonst $v(x) \equiv 0$ und $y(x, t) \equiv 0$ wäre, was die ruhende Saite beschreibt und wegen absoluter Langweiligkeit nicht weiter verfolgt zu werden braucht. Deswegen ist $v'(x)$ nicht identisch Null, so daß mit partieller Integration folgt:

$$\lambda \int_0^\pi v^2 dx = - \int_0^\pi v v'' dx = - v v' \Big|_0^\pi + \int_0^\pi (v')^2 dx = \int_0^\pi (v')^2 dx > 0 .$$

³⁰Für einige Blasinstrumente sind diese Bedingungen jedoch nicht die richtigen, wenn π die Länge des Instruments angibt.

λ ist aber noch weit stärkeren Einschränkungen unterworfen: Mit

$$0 = v(0) = A \cos(\sqrt{\lambda} \cdot 0 - B) = A \cos B$$

folgt, wenn man ohne Einschränkung $B = \pi/2$ wählt,

$$v(x) = A \cos\left(\sqrt{\lambda} \cdot x - \frac{\pi}{2}\right) = A \sin(\sqrt{\lambda} x).$$

Dann ist aber $0 = v(\pi) = A \sin(\sqrt{\lambda} \pi)$, also $\sqrt{\lambda} \cdot \pi = n\pi$, $n \in \mathbb{N}^*$, so daß nur die $\lambda = \lambda_n$ mit

$$\lambda_n = n^2$$

möglich sind. Insgesamt erhalten wir damit durch den Separationsansatz Lösungen der Form:

$$u_n(x, t) = \sin nx \cdot (a_n \cos nt + b_n \sin nt).$$

Diese repräsentieren die sogenannten *Grundschnwingungen* oder (akustisch) *Grundtöne*.

Nun ist die Wellengleichung *linear*. Dies impliziert, daß Lösungen addiert oder mit Konstanten multipliziert werden können, d. h. daß die Lösungen einen \mathbb{R} -Vektorraum bilden (*Überlagerung von Wellen*). Man kann sogar vermuten, daß jede Lösung als *unendliche konvergente Reihe* in den Grundschnwingungen

$$(*) \quad u(x, t) = \sum_{n=1}^{\infty} u_n(x, t) = \sum_{n=1}^{\infty} \sin nx \cdot (A_n \cos nt + B_n \sin nt)$$

dargestellt werden kann. Setzen wir zunächst voraus, daß wir eine solche Reihe vorliegen haben und daß ihre Konvergenz (und evtl. die von formal abgeleiteten Reihen) gleichmäßig ist. Wir können also differenzieren und $t = 0$ setzen:

$$\begin{aligned} g(x) &= u(x, 0) = \sum_{n=1}^{\infty} A_n \sin nx \\ h(x) &= \left. \frac{\partial u}{\partial t}(x, t) \right|_{t=0} = \sum_{n=1}^{\infty} \sin nx \cdot (-nA_n \sin nt + nB_n \cos nt) \Big|_{t=0} \\ &= \sum_{n=1}^{\infty} nB_n \sin nx. \end{aligned}$$

Wir wissen schon von früher, daß die Lösung $u(x, t)$ durch die Anfangsbedingungen g und h eindeutig bestimmt ist. Damit eine Lösung in der Form $(*)$ darstellbar ist, brauchen wir also (für zweimal differenzierbare Funktionen) gleichmäßig konvergente Entwicklungen der Form

$$g(x) = \sum_{n=1}^{\infty} A_n \sin nx, \quad h(x) = \sum_{n=1}^{\infty} nB_n \sin nx$$

(diese sind natürlich nur möglich, falls $g(0) = g(\pi) = 0$ und entsprechend für h).

Wenn solche *trigonometrische* Reihen auf $[0, \pi]$ konvergieren, dann auch auf ganz \mathbb{R} , und die dargestellten Funktionen g und h sind ungerade:

$$g(-x) = -g(x), \quad h(-x) = -h(x).$$

Da eine beliebige Funktion diese Bedingungen nicht erfüllt, wird man von vornherein allgemeinere trigonometrische Reihen betrachten, wie z. B.

$$(**) \quad f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos nx + b_n \sin nx)$$

(wobei die Koeffizienten a_n , b_n andere als früher sind). Wenn eine solche Reihe überall konvergiert, so stellt sie automatisch eine 2π -periodische Funktion dar: $f(x+2\pi) = f(x)$ für alle $f \in \mathbb{R}$, so daß wir nur ihre Eigenschaften auf dem Intervall $[0, 2\pi]$ zu studieren brauchen (von anderen Autoren wird das Intervall $[-\pi, \pi]$ bevorzugt, was aber keinen Unterschied macht).

Als erstes zeigen wir, daß eine gegebene Funktion f (unter milden Voraussetzungen) ihre *Fourier-Koeffizienten* a_n , b_n eindeutig bestimmt. Es seien nämlich $f(x) \cos mx$ und $f(x) \sin mx$ für alle m Riemann-integrierbar, und die in Frage stehenden Reihen seien gliedweise integrierbar; dann folgt mit den *Orthogonalitäts-Relationen* für die Sinus- und Cosinusfunktionen (die wir weiter unten im Komplexen ableiten werden):

$$\begin{aligned} \int_0^{2\pi} f(x) \cos mx \, dx &= \int_0^{2\pi} \frac{a_0}{2} \cos mx \, dx + \sum_{n=1}^{\infty} a_n \int_0^{2\pi} \cos nx \cos mx \, dx \\ &\quad + \sum_{n=1}^{\infty} b_n \int_0^{2\pi} \sin nx \cos mx \, dx \\ &= \frac{a_0}{2} \cdot 2\pi \delta_{0m} + a_n \pi \delta_{mn} = a_m \pi. \end{aligned}$$

(Diese einheitliche Formel ist der Grund dafür, daß der „nullte“ Koeffizient mit $a_0/2$ angesetzt wurde). Ebenso erhält man eine Formel für die Koeffizienten b_m . Zusammenfassend hat man also notwendigerweise

$$a_m = \frac{1}{\pi} \int_0^{2\pi} f(x) \cos mx \, dx, \quad b_m = \frac{1}{\pi} \int_0^{2\pi} f(x) \sin mx \, dx.$$

Die f vermöge dieser *Fourier-Koeffizienten* zugeordnete trigonometrische Reihe (**) heißt die *Fourier-Reihe* von f . Fragen, die sich sofort stellen, sind: Wo konvergent diese Reihe, wo konvergiert sie gegen f , wo ist sie gleichmäßig konvergent?

Es gilt z. B. der folgende Satz, dessen Beweis wir später nachtragen. (Man beachte, daß unter den Voraussetzungen des Satzes die Fourier-Koeffizienten existieren, da wegen der Stetigkeit von f alle Funktionen $f(x) \cos mx$ etc. Riemann-integrierbar sind).

Satz 19.3 *Es sei die Funktion f stetig und stückweise stetig differenzierbar. Dann konvergiert ihre Fourier-Reihe gleichmäßig gegen f .*

Kehren wir wieder zurück zu dem Problem der schwingenden Saite. Wir haben dabei vorgegebene Anfangsbedingungen

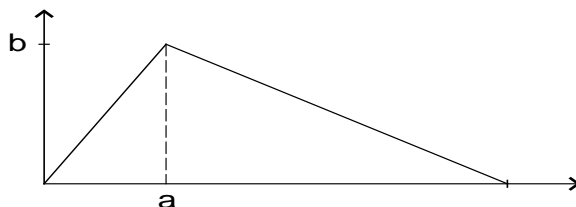
$$\begin{aligned} g(x) &\quad \text{zweimal stetig diffenzierbar auf } [0, \pi], \quad g(0) = g(\pi) = 0, \\ h(x) &\quad \text{einmal stetig diffenzierbar auf } [0, \pi], \quad h(0) = h(\pi) = 0. \end{aligned}$$

Setze g nach \mathbb{R} durch $g(-x) = -g(x)$ auf $[-\pi, 0]$ fort zu einer 2π -periodischen Funktion und verfähre ebenso mit h . Nach dem obigen Satz konvergieren die Fourier-Reihen dieser Fortsetzungen dann gleichmäßig gegen g bzw. h . Da die Fortsetzungen nach Konstruktion ungerade sind, sind die Fourier-Reihen reine *Sinus-Reihen*; z. B. gilt für die Koeffizienten a_n von g :

$$\begin{aligned} \pi a_n &= \int_0^{2\pi} g(x) \cos nx \, dx = \int_0^{\pi} g(x) \cos nx \, dx + \int_{-\pi}^0 g(x) \cos nx \, dx \\ &= \int_0^{\pi} g(x) \cos nx \, dx - \int_0^{\pi} g(-x) \cos n(-x) d(-x) \\ &= \int_0^{\pi} g(x) \cos nx \, dx - \int_0^{\pi} g(x) \cos nx \, dx = 0. \end{aligned}$$

Einsetzen in (***) ergibt dann eine gleichmäßige konvergente Reihe, die allen Bedingungen genügt.

Hinweis. Man kann offensichtlich so auch die *gezupfte Saite* behandeln: $g(x)$ ist hierbei nicht zweimal stetig differenzierbar, aber zumindest *stückweise* von dieser Güte. Das Resultat erfüllt dann die Wellengleichung nicht unmittelbar, aber doch in einem früher schon angedeuteten verallgemeinerten Sinne, der jetzt mit Hilfe der Fourier-Reihen leicht zu interpretieren ist.



Figur 19.4

Eine *Lösung* für den oben gezeichneten Fall ist gegeben durch

$$\varphi(x, t) = \frac{2b}{a(\pi - a)} \sum_{n=1}^{\infty} \left(\frac{\sin na}{n^2} \sin nx \right) \cos nt, \quad 0 < a < \pi.$$

Das Herumschleppen von cos und sin in der Fourier-Theorie ist etwas mühsam. Man kann sich das Leben vereinfachen, wenn man im *Komplexen* arbeitet. Wir erinnern daher zunächst an die EULERSchen Formeln:

$$\begin{aligned} a_n \cos nx + b_n \sin nx &= a_n \frac{e^{inx} + e^{-inx}}{2} + b_n \frac{e^{inx} - e^{-inx}}{2i} \\ &= \frac{1}{2} (a_n - ib_n) e^{inx} + \frac{1}{2} (a_n + ib_n) e^{-inx}. \end{aligned}$$

Umgekehrt ist (mit komplexen A_n, A_{-n}):

$$\begin{aligned} A_n e^{inx} + A_{-n} e^{-inx} &= A_n (\cos nx + i \sin nx) + A_{-n} (\cos nx - i \sin nx) \\ &= (A_n + A_{-n}) \cos nx + i(A_n - A_{-n}) \sin nx. \end{aligned}$$

Wir betrachten daher im folgenden *komplexwertige* Funktionen

$$f: \mathbb{R} \rightarrow \mathbb{C}, \quad f(x + 2\pi) = f(x).$$

Setzt man $f = u + iv$, wobei $u = \operatorname{Re} f$, $v = \operatorname{Im} f$, so folgt natürlich auch

$$u, v: \mathbb{R} \rightarrow \mathbb{R}, \quad u(x + 2\pi) = u(x), \quad v(x + 2\pi) = v(x).$$

Wir können also allgemeiner nach einer Entwicklung

$$f(x) = \sum_{n=-\infty}^{\infty} c_n e^{inx}, \quad c_n \in \mathbb{C},$$

fragen (die natürlich äquivalent ist zu reellen Fourier-Entwicklungen von $u(x)$ und $v(x)$, $f = u + iv$).

Um die komplexen Fourier-Koeffizienten c_n bestimmen zu können, benötigen wir das (RIEMANN-)Integral für *komplexwertige* Funktionen $f = u + iv$:

$$\int_a^b f(x) dx := \int_a^b u(x) dx + i \int_a^b v(x) dx,$$

d. h.

$$\operatorname{Re} \int_a^b f(x) dx := \int_a^b \operatorname{Re} f(x) dx, \quad \operatorname{Im} \int_a^b f(x) dx := \int_a^b \operatorname{Im} f(x) dx.$$

Wir berechnen (für $n \neq 0$):

$$\begin{aligned} \int e^{inx} dx &= \int (\cos nx + i \sin nx) dx = \int \cos nx dx + i \int \sin nx dx \\ &= \frac{1}{n} \sin nx - \frac{i}{n} \cos nx = \frac{1}{in} (\cos nx + i \sin nx) = \frac{1}{in} e^{inx} \end{aligned}$$

und finden also erneut die gleichen Formeln wie im Reellen. Daraus folgt sofort für $m, n \in \mathbb{Z}$:

$$\int_0^{2\pi} e^{inx} \overline{e^{imx}} dx = \int_0^{2\pi} e^{inx} e^{-imx} dx = \int_0^{2\pi} e^{i(n-m)x} dx = \begin{cases} 2\pi, & m = n, \\ 0, & m \neq n. \end{cases}$$

Man bekommt auf diese Weise sofort (bei geeigneten Integrier- und Summierbarkeits-Bedingungen):

$$\frac{1}{2\pi} \int_0^{2\pi} f(x) \overline{e^{imx}} dx = \frac{1}{2\pi} \int_0^{2\pi} f(x) e^{-imx} dx = \frac{1}{2\pi} \int_0^{2\pi} \sum_{n=-\infty}^{\infty} c_n e^{i(n-m)x} dx = c_m.$$

Wir führen jetzt auf dem Raum der auf $[0, 2\pi]$ Riemann-integrierbaren komplexwertigen Funktionen f das übliche *Skalarprodukt* (bis auf einen Normierungsfaktor) ein vermöge:

$$\langle f, g \rangle = \frac{1}{2\pi} \int_0^{2\pi} \overline{f(x)} g(x) dx.$$

Dann gelten auf diesem Raum die Axiome eines *hermiteschen* Skalarproduktes³¹:

$$\langle \lambda f + \mu g, h \rangle = \bar{\lambda} \langle f, h \rangle + \bar{\mu} \langle g, h \rangle$$

$$\langle f, \lambda g + \mu h \rangle = \lambda \langle f, g \rangle + \mu \langle f, h \rangle$$

$$\langle f, g \rangle = \overline{\langle g, f \rangle}.$$

Insbesondere folgt $\overline{\langle f, f \rangle} = \langle f, f \rangle$, d. h. $\langle f, f \rangle \in \mathbb{R}$. Genauer ist

$$2\pi \langle f, f \rangle = \int_0^{2\pi} f \bar{f} dx = \int_0^{2\pi} |f(x)|^2 dx \geq 0.$$

Ist f stetig, so folgt aus $\langle f, f \rangle = 0$ sogar, daß $f = 0$. (I. a. ist f jedoch *nicht* $\equiv 0$). Wir setzen

$$\|f\|_2 := \sqrt{\langle f, f \rangle} = \left(\frac{1}{2\pi} \int_0^{2\pi} |f|^2 dx \right)^{1/2}.$$

Der normierte Vektorraum $(\mathcal{C}^0([0, 2\pi]), \|\cdot\|_2)$ ist übrigens *nicht vollständig*. (Man vervollständigt ihn zum Raum der sogenannten L^2 -Funktionen wie man den rationalen Zahlkörper \mathbb{Q} zum Körper der reellen Zahlen vervollständigt).

Betrachte nun die Funktionen

$$e_k : \mathbb{R} \longrightarrow \mathbb{C}, \quad e_k(x) := e^{ikx}.$$

³¹Siehe auch den Anhang zu diesem Kapitel.

Diese sind 2π -periodisch, stetig, und erfüllen bzgl. des oben definierten Skalarproduktes

$$\langle e_k, e_\ell \rangle = \delta_{k\ell},$$

bilden also ein sogenanntes *Orthonormal-System*. Sie bilden jedoch *keine Basis* des Vektor-Raumes

$$\{ f \in \mathcal{C}^0(\mathbb{R}, \mathbb{C}) : f \text{ ist } 2\pi\text{-periodisch} \}.$$

(Dies würde ansonsten bedeuten, daß sich jede solche Funktion *endlich* aus den e_k linear kombinieren ließe, d. h. daß jede solche Funktion ein *trigonometrisches Polynom* wäre).

Definition. Eine Folge $f_n : [0, 2\pi] \rightarrow \mathbb{C}$ von Funktionen *konvergiert im quadratischen Mittel* gegen $f : [0, 2\pi] \rightarrow \mathbb{C}$, falls

$$\lim_{n \rightarrow \infty} \|f_n - f\|_2 = 0, \quad \text{d. h.} \quad \lim_{n \rightarrow \infty} \int_0^{2\pi} |f(x) - f_n(x)|^2 dx = 0.$$

Bemerkung. Ist die Folge f_n *gleichmäßig konvergent* gegen f , dann auch im *quadratischen Mittel*.

Warnung. Die Umkehrung dieser Aussage ist nicht richtig; ja nicht einmal *punktweise Konvergenz* kann aus der Konvergenz im quadratischen Mittel geschlossen werden.

Satz 19.4 Sei $f : \mathbb{R} \rightarrow \mathbb{C}$ *Riemann-integrierbar über $[0, 2\pi]$ und 2π -periodisch*. Seien ferner $c_k = \langle e_k, f \rangle$ die *Fourier-Koeffizienten von f* und

$$f_n = \sum_{k=-n}^n c_k e^{ikx}$$

die *Partialsummen der Fourier-Reihe*. Dann gilt:

$$\|f - f_n\|_2 \longrightarrow 0,$$

d. h. die *Fourier-Reihe von f konvergiert im quadratischen Mittel gegen f* .

Der *Beweis* dieses Satzes erfolgt in mehreren Schritten.

Lemma 19.5
$$\|f - f_n\|_2^2 = \|f\|_2^2 - \sum_{k=-n}^n |c_k|^2.$$

Folgerung 19.6 *Es gilt stets die BESSELSche Ungleichung*

$$\sum_{k=-\infty}^{\infty} |c_k|^2 \leq \|f\|_2^2.$$

Folgerung 19.7 *Es gilt $\lim_{n \rightarrow \infty} \|f - f_n\|_2 = 0$ genau dann, wenn die Vollständigkeitsrelation (oder BESSELSche Gleichung)*

$$\|f\|_2^2 = \sum_{n=-\infty}^{\infty} |c_n|^2$$

erfüllt ist.

Beweis (von Lemma 5). Es ist per definitionem

$$\begin{aligned}\langle f, f_n \rangle &= \sum_{k=-n}^n c_k \langle f, e_k \rangle = \sum_{k=-n}^n c_k \overline{c_k} = \sum_{k=-n}^n |c_k|^2, \\ \langle f_n, f_n \rangle &= \sum_{k, \ell=-n}^n c_\ell \overline{c_k} \delta_{\ell k} = \sum_{k=-n}^n |c_k|^2.\end{aligned}$$

Also ergibt sich, wie gewünscht,

$$\begin{aligned}\|f - f_n\|_2^2 &= \langle f - f_n, f - f_n \rangle = \|f\|_2^2 - \langle f_n, f \rangle - \langle f, f_n \rangle + \|f_n\|_2^2 \\ &= \|f\|_2^2 - \sum_{k=-n}^n |c_k|^2.\end{aligned}$$

□

Lemma 19.8 *Reelle 2π -periodische Treppenfunktionen erfüllen die Vollständigkeitsrelation.*

Beweis. Ohne Einschränkung sei

$$f(x) = \begin{cases} 1, & 0 \leq x < a, \\ 0, & a \leq x < 2\pi. \end{cases}$$

(Alle anderen Treppenfunktionen entstehen, außer an den endlich vielen Sprungstellen, durch endliche Linearkombinationen von Translaten dieser Funktion). Es ergibt sich durch eine einfache Rechnung $c_0 = a/2\pi$ und

$$\begin{aligned}\sum_{k=-\infty}^{\infty} |c_k|^2 &= \frac{a^2}{4\pi^2} + \sum_{k=1}^{\infty} \frac{1 - \cos ka}{\pi^2 k^2} \\ &= \frac{a^2}{4\pi^2} + \frac{1}{\pi^2} \sum_{k=1}^{\infty} \frac{1}{k^2} - \frac{1}{\pi^2} \sum_{k=1}^{\infty} \frac{\cos ka}{k^2} \\ &= \frac{a^2}{4\pi^2} + \frac{1}{\pi^2} \frac{\pi^2}{6} - \frac{1}{\pi^2} \left(\frac{(\pi - a)^2}{4} - \frac{\pi^2}{12} \right) \\ &= \frac{a}{2\pi} = \frac{1}{2\pi} \int_0^{2\pi} f^2(x) dx = \|f\|_2^2.\end{aligned}$$

□

Nun zum *Beweis* von Satz 4 durch Ableitung der *Vollständigkeitsrelation*. Wir setzen

$$f = u + iv, \quad c_k = \langle e_k, f \rangle, \quad a_k = \langle e_k, u \rangle, \quad b_k = \langle e_k, v \rangle,$$

so daß $c_k = a_k + ib_k$. Dann folgt

$$\int |f|^2 dx = \int |u|^2 dx + \int |v|^2 dx,$$

und, da u reell, ist, ergibt sich per definitionem

$$\langle e_k, u \rangle = \langle u, e_{-k} \rangle = \overline{\langle e_{-k}, u \rangle}, \quad \text{also} \quad a_k = \overline{a_{-k}} \quad \text{und entsprechend} \quad b_k = \overline{b_{-k}}.$$

Daraus folgt wegen

$$|c_k|^2 = |a_k|^2 + |b_k|^2 + i(b_k \overline{a_k} - a_k \overline{b_k})$$

sofort

$$\sum_{k=-\infty}^{\infty} |c_k|^2 = \sum_{k=-\infty}^{\infty} |a_k|^2 + \sum_{k=-\infty}^{\infty} |b_k|^2.$$

Also können wir ohne Einschränkung f als reell (und dem Betrage nach kleiner oder gleich 1) annehmen. Wegen der Definition der Riemann-Integrierbarkeit findet man dann zu jedem $\varepsilon > 0$ Treppenfunktionen φ, ψ , so daß $-1 \leq \varphi \leq f \leq \psi \leq 1$ und

$$\int_0^{2\pi} (\psi - \varphi) dx \leq \frac{\pi}{4} \varepsilon^2.$$

Setzt man noch $g := f - \varphi$, so folgt

$$|g|^2 = |f - \varphi|^2 \leq (\psi - \varphi)^2 \leq 2(\psi - \varphi),$$

da $\psi + \varphi \leq 2$, und damit

$$\|g\|_2^2 \leq \frac{1}{2\pi} \frac{\pi}{2} \varepsilon^2 = \frac{\varepsilon^2}{4}.$$

Sind nun f_n, φ_n, g_n jeweils die n -ten Partialsummen der Fourier-Reihen von f, φ bzw. g , so muß $g_n = f_n - \varphi_n$ gelten. Nach den Lemmata 5 und 8 ist aber

$$\|g - g_n\|_2^2 \leq \|g\|_2^2 \leq \frac{\varepsilon^2}{4}, \quad \|\varphi - \varphi_n\|_2 \leq \frac{\varepsilon}{2} \quad \text{für alle } n \geq N = N(\varepsilon)$$

und damit schließlich

$$\|f - f_n\|_2 = \|(g + \varphi) - (g_n + \varphi_n)\|_2 \leq \|g - g_n\|_2 + \|\varphi - \varphi_n\|_2 \leq 2 \frac{\varepsilon}{2} = \varepsilon$$

für alle $n \geq N$. □

Wir tragen jetzt noch den *Beweis* von Satz 3 nach. Nach Voraussetzung gibt es eine Unterteilung $\mathcal{Z} : 0 = t_0 < t_1 < \dots < t_r = 2\pi$ des Grundintervalls und stetige Funktionen $\varphi_j : [t_{j-1}, t_j] \rightarrow \mathbb{R}$ mit $f'(x) = \varphi_j(x)$ für $x \in (t_{j-1}, t_j)$. Es werde ferner φ auf dem Intervall $[0, 2\pi)$ erklärt durch $\varphi|_{[t_{j-1}, t_j)} := \varphi_j$ und dann als 2π -periodische Funktion fortgesetzt nach ganz \mathbb{R} . Sind γ_k die Fourier-Koeffizienten von φ , so ist nach der BESSELSchen *Ungleichung*:

$$\sum_{k=-\infty}^{\infty} |\gamma_k| \leq \|\varphi\|_2^2 < \infty.$$

Man kann nun durch *partielle Integration* die Fourier-Koeffizienten c_k von f für $k \neq 0$ aus denjenigen von φ berechnen. Wir zeigen weiter unten, daß

$$(+)\quad c_k = \frac{i\gamma_k}{k}.$$

Aus der Ungleichung zwischen dem geometrischen und arithmetischen Mittel ergibt sich hieraus unmittelbar

$$|c_k| \leq \frac{1}{2} \left(\frac{1}{|k|^2} + |\gamma_k|^2 \right).$$

Somit ist die Reihe

$$\sum_{k=-\infty}^{\infty} c_k$$

absolut konvergent, so daß die Fourier-Reihe $\sum c_k e^{ikx}$ von f absolut und gleichmäßig gegen eine (stetige) Funktion g konvergiert. Damit konvergiert die Fourier-Reihe im quadratischen Mittel sowohl gegen f als auch gegen g . Also ist $\|f - g\|_2 = 0$ und somit, da f und g stetig sind, $f = g$.

Es bleibt noch (+) zu begründen. Es ist

$$\begin{aligned} 2\pi c_k &= \int_0^{2\pi} f(x) e^{-ikx} dx = \sum_{j=1}^r \int_{t_{j-1}}^{t_j} f(x) e^{-ikx} dx \\ &= \sum_{j=1}^r \left(f(x) \frac{e^{-ikx}}{-ik} \Big|_{t_{j-1}}^{t_j} - \int_{t_{j-1}}^{t_j} \varphi_j(x) \frac{e^{-ikx}}{-ik} dx \right) \\ &= f(x) \frac{e^{-ikx}}{-ik} \Big|_0^{2\pi} - \int_0^{2\pi} \varphi(x) \frac{e^{-ikx}}{-ik} dx = 2\pi \frac{\gamma_k}{ik}. \quad \square \end{aligned}$$

Bemerkung. In gewisser Weise spiegelt die „Schnelligkeit“ der Konvergenz der Reihe $\sum |c_k|$ der Fourier-Koeffizienten einer Funktion f die „Differenzierbarkeitsordnung“ von f wider. Denn bei einer stetig differenzierbaren Funktion ist nach dem Beweis des vorangehenden Satzes nicht nur die eben erwähnte Reihe konvergent, sondern sogar

$$\sum k |c_k|.$$

Ähnliche Aussagen spielen in der Theorie der FOURIER-Integrale eine ganz entscheidende Rolle.

Es sei noch angemerkt, daß im Allgemeinfall nicht einmal *punktweise Konvergenz* der Fourier-Reihe gegen ihre Ausgangsfunktion stattfindet. Hat man aber eine *Regelfunktion* f vorliegen, so existieren natürlich die Fourier-Koeffizienten und damit auch die f zugeordnete Fourier-Reihe, und es gilt der folgende Satz.

Satz 19.9 (Dirichlet) Für eine 2π -periodische Regelfunktion $f : \mathbb{R} \rightarrow \mathbb{R}$ konvergiert die Fourier-Reihe von f an jeder Stelle $x \in \mathbb{R}$, an der f sowohl eine linksseitige als auch eine rechtsseitige Ableitung besitzt, punktweise gegen den Mittelwert der Funktionswerte :

$$\frac{1}{2} (f(x+) + f(x-)),$$

insbesondere also gegen $f(x)$, wenn f in x stetig ist.

Zum Beweis siehe z. B. [11], Band 1, Satz 16.3. □

Bemerkungen. 1. Aufgrund der Konvergenz der Fourier-Reihen im quadratischen Mittel ist der natürliche Definitionsbereich der Fourier-Theorie der Vektorraum L^2 der im Lebesgueschen Sinne *quadratintegrierbaren* periodischen Funktionen. Wir werden auf diesen erst in den Kapiteln über die mehrdimensionale Integration eingehen können. Der vorstehende Satz bleibt auch dann für eine Stelle x noch richtig, wenn neben den beiden Ableitungen auch die beiden Grenzwerte $f(x+)$ und $f(x-)$ als existent vorausgesetzt werden (siehe [9], Teil 2, Satz 136.4).

2. Wenn man den Begriff der *Konvergenz* einer Reihe abschwächt, so erhält man noch weit stärkere Aussagen. Man nennt eine Reihe $\sum_{k=0}^{\infty} a_k$ C-summierbar zum Wert s , wenn $\lim_{n \rightarrow \infty} \sigma_n = s$ gilt für die Folge

$$\sigma_n := \frac{s_0 + \cdots + s_n}{n+1}$$

der *arithmetischen Mittel* der Folge (s_n) der Partialsummen $s_n = a_0 + \cdots + a_n$. Man schreibt dann auch

$$C - \sum_{k=0}^{\infty} a_k := s.$$

Das Symbol „C“ steht für den Erfinder dieses Begriffes, ERNESTO CESARÒ. Ist die Reihe *konvergent* gegen den Grenzwert s , so ist sie auch C-summierbar zum gleichen Wert. Die Umkehrung braucht

selbstverständlich nicht zu gelten, wie das divergente Beispiel der Reihe $\sum_{k=0}^{\infty} (-1)^k$ mit

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

zeigt. - Es gilt nun der folgende Satz (zum *Beweis* siehe loc. cit., Satz 139.3).

Satz 19.10 (Fejér) *Existieren für eine 2π -periodische L^2 -Funktion f an einer Stelle x der links- und der rechtsseitige Grenzwert $f(x+)$ und $f(x-)$, so ist ihre Fourier-Reihe an dieser Stelle C -summierbar zum Wert*

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

Als Anwendung von Satz 3 geben wir einen Beweis des folgenden wohlbekannten Satzes, den wir schon einmal in Kapitel 14 in einer Beweisvariante herangezogen haben.

Satz 19.11 (Weierstraßscher Approximationssatz) *Jede stetige Funktion $f : [a, b] \rightarrow \mathbb{R}$ läßt sich beliebig genau in der Supremumsnorm durch Polynome approximieren, ist also der gleichmäßige Limes einer Folge von Polynomen.*

Beweis. Es sei ohne Einschränkung $[a, b] = [0, 1]$. Jede stetige Funktion $f : [0, 1] \rightarrow \mathbb{R}$ läßt sich gleichmäßig durch *Treppenfunktionen* und damit auch, wie man sich sofort klar macht, durch eine *stetige, stückweise lineare* Funktion approximieren. Diese läßt sich aber sofort fortsetzen zu einer *2π -periodischen stetigen, stückweise linearen*, also insbesondere stückweise differenzierbaren Funktion auf ganz \mathbb{R} . Nach Satz 3 ist dann f durch Partialsummen f_k ihrer Fourier-Entwicklung approximierbar. Die f_k ihrerseits sind (endliche) *trigonometrische Summen*, die man auf ganz \mathbb{R} in Potenzreihen entwickeln kann. Durch einfaches Zusammensetzen dieser Puzzle-Stückchen gelangt man zu der Behauptung. \square

Bemerkung. Will man die Güte der Approximation verbessern, so hat man i. A. den *Grad* des approximierenden Polynoms zu erhöhen. Dabei *ändern sich i. A. auch die Koeffizienten in niedrigeren Graden*. Wäre dies nämlich nicht der Fall, so hätte dieser Satz zur Folge, daß jede Funktion der gegebenen Art *analytisch* wäre, was bekanntlich definitiv falsch ist. Es sind eben gerade die Potenzreihenentwicklungen analytischer Funktionen, bei denen wir zur Verbesserung der Approximation jeweils nur *höhere Terme anhängen müssen*.

Der obige Beweis des Weierstraßschen Approximationssatzes besitzt den Nachteil, daß man die approximierende Polynomfolge nicht explizit angeben kann. Wir skizzieren daher einen *weiteren Beweis*, der die *BERNSTEIN-Polynome* benutzt. Wir folgen dabei dem Buch von Storch und Wiebe [17]. Ohne Einschränkung sei wieder $I = [0, 1]$. Nach dem *schwachen Gesetz der großen Zahlen* (siehe z. B. loc. cit., p.185) ist für jedes $x \in [0, 1]$, jedes $a > 0$ und jede natürliche Zahl $n \in \mathbb{N}^*$

$$\sum_{k \in \mathbb{N}, \left| \frac{k}{n} - x \right| \geq a} \binom{n}{k} x^k (1-x)^{n-k} \leq \frac{x(1-x)}{a^2 n} \leq \frac{1}{4a^2 n}.$$

Betrachte nun zu f die Bernstein-Polynome

$$B_n(x) := \sum_{k=0}^n f\left(\frac{k}{n}\right) \binom{n}{k} x^k (1-x)^{n-k}.$$

Wegen der gleichmäßigen Stetigkeit von f auf I gibt es zu jedem $\varepsilon > 0$ ein $\delta > 0$, so daß aus $x, y \in [0, 1]$, $|x - y| \leq \delta$ die Ungleichung $|f(x) - f(y)| \leq \varepsilon/2$ folgt. Beachtet man noch $|f(x)| \leq M$ und

$$\sum_{k=0}^n \binom{n}{k} x^k (1-x)^{n-k} = 1,$$

so ergibt sich mit $n \geq M/\delta^2 \varepsilon$ und den obigen Abschätzungen sofort die Behauptung:

$$\begin{aligned} |f(x) - B_n(x)| &= \left| \sum_{k=0}^n \left(f(x) - f\left(\frac{k}{n}\right) \right) \binom{n}{k} x^k (1-x)^{n-k} \right| \\ &\leq \sum_{k \in \mathbb{N}, \left| \frac{k}{n} - x \right| < \delta} \left(f(x) - f\left(\frac{k}{n}\right) \right) \binom{n}{k} x^k (1-x)^{n-k} \\ &\quad + \sum_{k \in \mathbb{N}, \left| \frac{k}{n} - x \right| \geq \delta} \left(f(x) - f\left(\frac{k}{n}\right) \right) \binom{n}{k} x^k (1-x)^{n-k} \\ &\leq \frac{\varepsilon}{2} + 2M \frac{1}{4\delta^2 n} \leq 2 \frac{\varepsilon}{2} = \varepsilon. \quad \square \end{aligned}$$

Es gibt noch eine wesentliche Verschärfung des Weierstraßschen Approximationssatzes für stetige Funktionen auf kompakten metrischen und sogar topologischen Räumen, die in der Literatur als *Approximationssatz von STONE* und *WEIERSTRASS* bezeichnet wird.

Satz 19.12 *Es sei X ein kompakter topologischer Raum und D eine Teilmenge der Algebra $\mathcal{C}^0(X)$ der stetigen Funktionen auf X mit den folgenden Eigenschaften :*

- Für alle $x \in X$ gibt es ein $f_x \in D$ mit $f_x(x) \neq 0$;
- für alle $x, y \in X$, $x \neq y$, gibt es ein $f \in D$ mit $f(x) \neq f(y)$ (man sagt: D trennt die Punkte von X).

Dann liegt die von D erzeugte Unter algebra $A(D)$ dicht in $\mathcal{C}^0(X)$ bzgl. der Supremumsnorm; d. h. jede stetige Funktion auf X ist gleichmäßiger Limes einer Folge von Polynomen

$$\sum_{0 \leq j_1, \dots, j_r \leq n} a_{j_1, \dots, j_r} p_1^{j_1} \cdot \dots \cdot p_r^{j_r}, \quad p_1, \dots, p_r \in D.$$

Bemerkungen. 1. Wählt man für $I = [0, 1]$ die Menge $D := \{1, x\}$, so erhält man aus dem obigen Satz den Approximationssatz von Weierstraß erneut als Korollar.

2. Entsprechend erhält man mit $D = \{1, x_1, \dots, x_n\}$, x_j die j -te Koordinatenfunktion auf \mathbb{R}^n , die entsprechende Verallgemeinerung auf den \mathbb{R}^n :

Folgerung 19.13 *Fede stetige Funktion f auf einem Kompaktum $K \subset \mathbb{R}^n$ läßt sich beliebig genau in der Supremumsnorm durch Polynome in n Veränderlichen approximieren.*

Zum *Beweis* des Satzes von Stone–Weierstraß benötigt man erstaunlicherweise nur einen ganz konkreten *Spezialfall* des Weierstraßschen Approximationssatzes.

Lemma 19.14 *Es gibt eine Folge von Polynomen $p_n : \mathbb{R} \rightarrow \mathbb{R}$ mit $p_n(0) = 0$, die auf $[0, 1]$ gleichmäßig gegen $f(x) := \sqrt{x}$ konvergiert.*

Beweis des Lemmas. Man setze induktiv $p_0(x) := 0$ und $p_{n+1}(x) := p_n(x) + \frac{1}{2}(x - p_n^2(x))$.

Es ist dann

$$\sqrt{x} - p_{n+1}(x) = (\sqrt{x} - p_n(x)) \left(1 - \frac{1}{2}(\sqrt{x} + p_n(x)) \right),$$

woraus sich per Induktion

$$\frac{nx}{2 + n\sqrt{x}} \leq p_n(x) \leq \sqrt{x}, \quad \text{d. h.} \quad 0 \leq \sqrt{x} - p_n(x) \leq \frac{2\sqrt{x}}{2 + n\sqrt{x}}$$

auf $[0, 1]$ ergibt. Hieraus folgt

$$\sup_{0 \leq x \leq 1} |\sqrt{x} - p_n(x)| \leq \sup_{0 \leq x \leq 1} \frac{2\sqrt{x}}{2 + n\sqrt{x}} \leq \frac{2}{n}$$

und damit die gleichmäßige Konvergenz. $p_n(0) = 0$ gilt nach Konstruktion. \square

Folgerung 19.15 Auf jedem Intervall $[-a, a]$ läßt sich die Betragsfunktion $h(x) := |x|$ beliebig genau durch Polynome q_n approximieren.

Beweis. Setze $q_n(x) := a p_n\left(\frac{x^2}{a^2}\right)$. \square

Bemerkung. Das Iterationsverfahren im Beweis von Lemma 14 beruht auf einem Spezialfall des BANACHSchen Fixpunktsatzes (siehe Kapitel 23). Die Idee trägt wesentlich weiter (siehe loc. cit.).

Wir benötigen schließlich noch das folgende

Lemma 19.16 Es sei A eine abgeschlossene Unteralgebra von $\mathcal{C}^0(X)$, X ein kompakter topologischer Raum. Dann enthält A mit f und g auch $|f|$, $\max(f, g)$, $\min(f, g)$.

Beweis. Nach unseren früheren Überlegungen zu *Funktionenverbänden* in Kapitel 18 brauchen wir nur zu zeigen, daß mit f auch $|f|$ in A enthalten ist. Es sei $a = \sup_{x \in X} |f(x)|$. Nach der obigen Folgerung gibt es zu jedem $\varepsilon > 0$ ein Polynom p_ε mit $p_\varepsilon(0) = 0$, so daß

$$\sup_{x \in X} ||f(x)| - p_\varepsilon(f(x))| \leq \varepsilon.$$

Da mit f auch $p_\varepsilon(f)$ zu A gehört (p_ε besitzt nach Konstruktion keinen konstanten Term), liegt in jeder ε -Umgebung von $|f|$ eine Funktion aus A . Also ist $|f| \in \overline{A} = A$. \square

Es dürfte andererseits dem aufmerksamen Leser bewußt sein, daß für jede Unteralgebra $A \subset \mathcal{C}^0(X)$ auch der Abschluß \overline{A} eine Unteralgebra der Algebra der stetigen Funktionen auf X bildet. - Wir können nunmehr den Satz von Stone-Weierstraß beweisen:

Es ist zu zeigen: Zu jeder beliebigen Funktion $f \in \mathcal{C}^0(X)$ und jedem $\varepsilon > 0$ gibt es eine Funktion $g_\varepsilon \in A(D)$ mit $\sup |f(x) - g_\varepsilon(x)| \leq \varepsilon$. Der *Beweis* erfolgt in 3 Schritten:

1. Zu je zwei Punkten $x_1 \neq x_2$ gibt es ein $h \in A(D)$ mit $h(x_j) = f(x_j)$, $j = 1, 2$.

Seien $f_1, f_2 \in D$ so gewählt, daß $f_j(x_j) \neq 0$, $j = 1, 2$. Dann liegen die Funktionen

$$F_j := \frac{1}{f_j(x_j)} f_j \quad \text{in } A(D)$$

und erfüllen $F_j(x_j) = 1$, $j = 1, 2$, und

$$h_1 := F_1 + F_2 - F_1 F_2$$

liegt in $A(D)$ und erfüllt $h_1(x_1) = h_1(x_2) = 1$. Ferner existiert ein $h_2 \in D$, s. d. $h_2(x_1) \neq h_2(x_2)$. Wähle

$$h(x) := \frac{f(x_2) - f(x_1)}{h_2(x_2) - h_2(x_1)} h_2(x) - \frac{f(x_1)h_2(x_2) - f(x_2)h_2(x_1)}{h_2(x_2) - h_2(x_1)} h_1(x);$$

dann liegt h in $A(D)$ und erfüllt die behaupteten Identitäten $h(x_j) = f(x_j)$, $j = 1, 2$. Auch für $x_1 = x_2$ hat man eine solche Funktion, nämlich z. B.

$$h(x) = \frac{f(x_1)}{f_1(x_1)} f_1(x).$$

2. Zu jedem $\varepsilon > 0$ und jedem $x_0 \in X$ gibt es eine Funktion $h_0 \in \overline{A(D)}$ mit $h_0(x_0) = f(x_0)$ und $h_0(x) \leq f(x) + \varepsilon$ für alle $x \in X$.

Zu jedem $x_1 \in X$ gibt es nämlich ein $g_1 \in A(D)$ mit $g_1(x_0) = f(x_0)$, $g_1(x_1) = f(x_1)$ nach dem 1. Schritt. Für jedes $x_1 \in X$ sei $U = U(x_1)$ eine Umgebung mit $g_1(x) \leq f(x) + \varepsilon$ für alle $x \in U$. Wegen der Kompaktheit von X kann man eine endliche Überdeckung auswählen, und definiert für die entsprechenden $g_1^{(1)}, \dots, g_1^{(L)}$:

$$h_0 := \min_{\lambda} g_1^{(\lambda)}.$$

Nach den vorigen Bemerkungen ist $h_0 \in \overline{A(D)}$ und besitzt die gewünschten Eigenschaften.

3. Seien schließlich $f \in \mathcal{C}^0(X)$ und $\varepsilon > 0$ vorgegeben. Zu festem $x_0 \in X$ sei $h_0 \in A(D)$ wie im zweiten Schritt konstruiert, und $W = W(x_0)$ sei eine Umgebung von x_0 mit

$$h_0(x) > f(x) - \varepsilon, \quad x \in W(x_0).$$

Mit dem gleichen Kompaktheitsargument wie im 2. Schritt geht man zu einem geeigneten endlichen Maximum dieser Funktionen h_0 über, das wieder zu $\overline{A(D)}$ gehört. Wegen 2. ergibt sich sofort

$$\sup_{x \in X} |h_0(x) - f(x)| \leq \varepsilon.$$

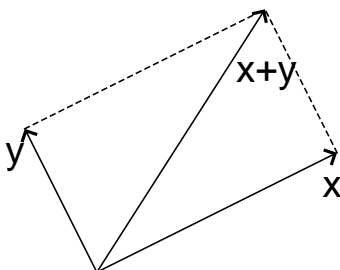
□

Anhang: Skalarprodukt- und Hilbert-Räume

In unserer Anschauung sind die Ebene und der uns umgebende Raum nicht nur mit einer linearen Struktur versehen; wir haben auch eine Vorstellung von Begriffen wie *Längen* von Strecken und dem *Senkrechtstehen* (*Orthogonalität*) von Strecken oder Geraden. Überdies reden wir auch von *Winkeln* zwischen sich schneidenden Geraden (obwohl dieser Begriff, wie wir sehen werden, aus den beiden anderen abgeleitet werden kann). Alle diese Begriffsbildungen gehören in den Bereich der *euklidischen Geometrie*, die wir hier noch einmal kurz streifen möchten, obwohl wir vieles hiervon früher schon benutzt haben.

Orthogonalität und Längenmessung sind in der euklidischen Ebene eng miteinander durch den *Satz des PYTHAGORAS* verknüpft: Für je zwei Vektoren $x, y \in \mathbb{R}^2$ mit $x \perp y$ gilt:

$$\|x + y\|^2 = \|x\|^2 + \|y\|^2.$$



Figur 19.5

Aber auch umgekehrt kann man mit elementarer Geometrie sofort zeigen, daß aus der Gültigkeit des Satzes von Pythagoras die *Orthogonalität* von x und y folgt.

Setzen wir nun allgemein im \mathbb{R}^n voraus, daß wir eine *Längenmessung*

$$\|\cdot\| : \mathbb{R}^n \rightarrow \mathbb{R}_+$$

besitzen und einen symmetrischen *Orthogonalitätsbegriff* \perp , so können wir leicht die übliche Form der *euklidischen Norm* und des *euklidischen Skalarproduktes* aus wenigen (unmittelbar einleuchtenden) Axiomen ableiten. – Wir fordern:

1. Es gibt eine *orthonormierte Basis* (e_1, \dots, e_n) , d. h. eine Basis mit $\|e_j\| = 1$ für alle j und $e_j \perp e_k$ für alle $j \neq k$.
2. Zwei Vektoren x und y stehen genau dann senkrecht aufeinander, wenn

$$\|x + y\|^2 = \|x\|^2 + \|y\|^2.$$

3. Gilt $x \perp y_1$ und $x \perp y_2$, so ist auch $x \perp y$ für alle $y \in \text{span}(y_1, y_2)$.
4. $\|ax\| = |a| \cdot \|x\|$ für alle $a \in \mathbb{R}$, $x \in \mathbb{R}^n$.

Schreibt man nun $x = x_1e_1 + \dots + x_n e_n$, so folgt aus 1. und 3. die Relation $x_n e_n \perp x_1 e_1 + \dots + x_{n-1} e_{n-1}$ und damit

$$\|x\|^2 = \|x_1 e_1 + \dots + x_{n-1} e_{n-1}\|^2 + \|x_n e_n\|^2.$$

Per Induktion ergibt sich daraus sofort

$$\|x\|^2 = x_1^2 + \dots + x_n^2,$$

und, da Längen nicht negativ sein sollen,

$$\|x\| = \left(\sum_{j=1}^n x_j^2 \right)^{1/2}.$$

Weiter kann man, ausgehend von dem Axiom 2, die *Orthogonalität* im euklidischen (Standard-) Raum \mathbb{R}^n mit Hilfe einer *Bilinearform* analytisch fassen. Es gilt per definitionem für $x, y \in \mathbb{R}^n$, daß $x \perp y$ genau dann, wenn $\|x + y\|^2 = \|x\|^2 + \|y\|^2$, und dies ist seinerseits äquivalent zu

$$\sum_{j=1}^n (x_j + y_j)^2 = \sum_{j=1}^n x_j^2 + \sum_{j=1}^n y_j^2 \iff \langle x, y \rangle := x_1 y_1 + \cdots + x_n y_n = 0.$$

Wir haben also die kanonische symmetrische Bilinearform $\langle x, y \rangle = x_1 y_1 + \cdots + x_n y_n$ auf \mathbb{R}^n zur Verfügung, so daß

$$\|x\| = \langle x, x \rangle^{1/2} \quad \text{und} \quad x \perp y \iff \langle x, y \rangle = 0.$$

Dieses Beispiel verallgemeinern wir sofort zu der folgenden generellen

Definition. Es sei V ein \mathbb{R} -Vektorraum. Ein *euklidisches Skalarprodukt* auf V ist eine symmetrische Bilinearform

$$\langle \cdot, \cdot \rangle : V \times V \longrightarrow \mathbb{R}$$

mit $\langle v, v \rangle > 0$ für $v \neq 0$ (also eine sogenannte *positiv-definite Form*). Das Paar $(V, \langle \cdot, \cdot \rangle)$ heißt ein *euklidischer Vektorraum*.

Ein *Beispiel* für einen euklidischen Vektorraum ist neben \mathbb{R}^n mit dem obigen Standard-Skalarprodukt $\langle x, y \rangle = \sum_{j=1}^n x_j y_j$ der Raum $\mathcal{C}^0(I)$ der stetigen Funktionen $f : I \rightarrow \mathbb{R}$ auf einem Intervall $I \subset \mathbb{R}$ mit

$$\langle f, g \rangle := \int_I f(x) g(x) dx.$$

Die *Positiv-Definitheit* des Skalarprodukts ist natürlich notwendig, um überhaupt eine Norm definieren zu können: Der Ausdruck

$$\|v\| = \langle v, v \rangle^{1/2}$$

erfüllt Bedingung i) einer Norm, aber auch ii): Aus $\|av\|^2 = \langle av, av \rangle = a^2 \langle v, v \rangle$ folgt $\|av\| = |a| \cdot \|v\|$. Die Dreiecksungleichung iii) ist eine direkte Folgerung aus der (allgemeinen) *Cauchy-Schwarzschen Ungleichung*, die wir jetzt beweisen wollen.

Satz 19.17 (Cauchy - Schwarzsche Ungleichung) *Es sei V ein euklidischer Vektorraum. Dann gilt stets*

$$|\langle v, w \rangle| \leq \|v\| \cdot \|w\|,$$

und Gleichheit besteht hier genau dann, wenn v und w linear abhängig sind.

Beweis. Seien v, w abhängig, also ohne Einschränkung $w = av$. Dann ist

$$|\langle v, w \rangle| = |\langle v, av \rangle| = |a| \cdot \|v\|^2 = \|v\| \cdot \|w\|.$$

Seien umgekehrt v, w linear unabhängig, so daß also insbesondere $v - aw \neq 0$ für alle $a \in \mathbb{R}$ und $w \neq 0$. Dann ergibt sich

$$0 < \langle v - aw, v - aw \rangle = \|v\|^2 - 2a \langle v, w \rangle + |a|^2 \|w\|^2.$$

Wählt man hier speziell

$$a = \frac{\langle v, w \rangle}{\|w\|^2},$$

so folgt sofort

$$0 < \|v\|^2 - \frac{\langle v, w \rangle^2}{\|w\|^2}$$

und damit die Behauptung

$$|\langle v, w \rangle| < \|v\| \cdot \|w\|. \quad \square$$

Folgerung 19.18 Ist $\langle \cdot, \cdot \rangle$ ein euklidisches Skalarprodukt auf dem Vektorraum V , so wird durch $\|v\| := \sqrt{\langle v, v \rangle}$ eine Norm auf V definiert.

Beweis. Es fehlt nur noch der Nachweis der *Dreiecksungleichung*, die sich einfach aus der Cauchy-Schwarzschen Ungleichung ergibt:

$$\begin{aligned} \|v + w\|^2 &= \langle v + w, v + w \rangle = \|v\|^2 + 2\langle v, w \rangle + \|w\|^2 \\ &\leq \|v\|^2 + 2\|v\| \cdot \|w\| + \|w\|^2 \\ &= (\|v\| + \|w\|)^2. \end{aligned} \quad \square$$

Definition. $\|v\| := \langle v, v \rangle^{1/2}$ heißt die dem Skalarprodukt $\langle \cdot, \cdot \rangle$ zugeordnete Norm.

In einem euklidischen Raum gilt insbesondere die Formel

$$\|v + w\|^2 = \|v\|^2 + \|w\|^2 + 2\langle v, w \rangle,$$

die man auch als (allgemeinen) *Satz des PYTHAGORAS* bezeichnet. Somit ist der klassische Pythagoras

$$\|v + w\|^2 = \|v\|^2 + \|w\|^2$$

genau dann erfüllt, wenn $\langle v, w \rangle = 0$ ist. Deswegen definiert man auch in einem beliebigen euklidischen Vektorraum das *Senkrechtstehen* oder die *Orthogonalität* zweier Vektoren v, w durch

$$v \perp w \iff \langle v, w \rangle = 0 \iff \|v + w\|^2 = \|v\|^2 + \|w\|^2.$$

Man kann im euklidischen \mathbb{R}^n sogar *Winkel* zwischen Vektoren v, w definieren³². Das geht natürlich nur, wenn $v \neq 0, w \neq 0$. Der Winkel sollte unabhängig von der Länge der Vektoren sein. Wir normieren sie deshalb durch

$$v_0 = \frac{v}{\|v\|}, \quad w_0 = \frac{w}{\|w\|}$$

auf die Länge 1 und versuchen eine Zerlegung

$$w_0 = w_1 + w_2, \quad w_1 \in \text{span}(v_0), \quad w_2 \perp v_0,$$

was tatsächlich mit $w_1 := \langle w_0, v_0 \rangle v_0, w_2 := w_0 - w_1$ gelingt:

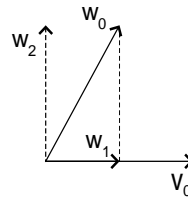
$$\langle w_2, v_0 \rangle = \langle w_0, v_0 \rangle - \langle w_1, v_0 \rangle = \langle w_0, v_0 \rangle - \langle w_0, v_0 \rangle \langle v_0, v_0 \rangle = 0.$$

Wir haben hierbei nichts anderes getan, als den Vektor w_0 auf v_0 *orthogonal* zu projizieren (siehe die nachfolgende Figur). Ferner ist $\|w_1\|^2 + \|w_2\|^2 = \|w_1 + w_2\|^2 = \|w_0\|^2 = 1$, so daß also aufgrund der Zeichnung der Winkel ϑ durch

$$\cos \vartheta = \pm \|w_1\| = \langle w_0, v_0 \rangle = \frac{\langle w, v \rangle}{\|w\| \cdot \|v\|}$$

bestimmt ist. Nach dieser Definition sind zwei Vektoren genau dann orthogonal zueinander, wenn der Cosinus ihres Zwischenwinkels gleich Null, der Zwischenwinkel selbst also gleich $\pi/2$ ist, wie es unserer Anschauung entspricht.

³²Siehe auch Lemma 17.16.



Figur 19.6

Definition und Bemerkung. Auf exakt die gleiche Weise kann man auch in *beliebigen* euklidischen Vektorräumen immer einen Winkel ϑ definieren, was wiederum an der *Cauchy-Schwarzschen Ungleichung* liegt:

$$\cos \vartheta = \frac{\langle v, w \rangle}{\|v\| \cdot \|w\|}, \quad 0 \leq \vartheta \leq \pi.$$

Als weitere Konsequenz aus der Cauchy-Schwarzschen Ungleichung notieren wir noch die

Folgerung 19.19 *In einem euklidischen Raum ist das Skalarprodukt $V \times V \ni (v, w) \mapsto \langle v, w \rangle \in \mathbb{R}_+$ eine stetige Funktion.*

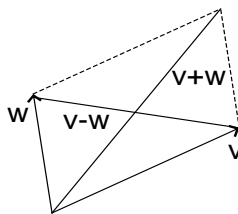
Beweis. Es gilt $\langle v, w \rangle - \langle v_n, w_n \rangle = \langle v - v_n, w \rangle + \langle v_n, w - w_n \rangle$ und damit

$$|\langle v, w \rangle - \langle v_n, w_n \rangle| \leq \|v - v_n\| \|w\| + \|v_n\| \|w - w_n\|. \quad \square$$

In euklidischen Vektorräumen ist die Norm weiteren Bedingungen unterworfen. So hat man die folgende *geometrische* Beziehung, die man einfach mit dem Satz von Pythagoras in der oben allgemein formulierten Version in einer Zeile ausrechnet.

Lemma 19.20 *Es sei V ein euklidischer Vektorraum. Dann gilt die Parallelogramm-Gleichung :*

$$\|v + w\|^2 + \|v - w\|^2 = 2(\|v\|^2 + \|w\|^2).$$



Figur 19.7

Außerdem erhält man hieraus leicht die sogenannte *Polarisationsformel*:

Lemma 19.21 *Es sei V ein euklidischer Vektorraum. Dann gilt*

$$\langle v, w \rangle = \frac{1}{2} (\|v + w\|^2 - \|v\|^2 - \|w\|^2) = \frac{1}{4} (\|v + w\|^2 - \|v - w\|^2).$$

D. h. also: Im euklidischen Fall ist das Skalarprodukt $\langle \cdot, \cdot \rangle$ schon durch die Norm $\|\cdot\|$ bestimmt. – In Bezug auf die Umkehrung gilt nun der folgende Satz.

Satz 19.22 *Eine Norm $\|\cdot\|$ kommt genau dann von einem Skalarprodukt her, wenn sie der Parallelogramm-Gleichung genügt.*

Beweis. Nach einer der Polarisationsformeln muß für das gesuchte Skalarprodukt die Formel

$$(*) \quad \langle v, w \rangle := \frac{1}{4} (\|v + w\|^2 - \|v - w\|^2)$$

gelten. Wir bemerken sofort, daß die durch $\langle \cdot, \cdot \rangle$ auf einem beliebigen normierten Vektorraum V definierte Form *symmetrisch* ist und die Eigenschaften

$$\langle v, v \rangle = \|v\|^2, \quad \langle 0, w \rangle = 0$$

besitzt.

Wir brauchen daher nur noch die Linearität in dem ersten Faktor nachzuweisen, sofern die gegebene Norm der *Parallelogramm-Gleichung* genügt. Dazu setzen wir in diese die Vektoren $x = v_1 \pm \frac{1}{2} w$ und $y = v_2 \pm \frac{1}{2} w$ ein und erhalten

$$\|v_1 + v_2 \pm w\|^2 + \|v_1 - v_2\|^2 = 2\|v_1 \pm \frac{1}{2} w\|^2 + 2\|v_2 \pm \frac{1}{2} w\|^2.$$

Subtraktion dieser beiden Gleichungen liefert dann die Formel

$$(+)$$

$$\langle v_1 + v_2, w \rangle = 2\langle v_1, \frac{1}{2} w \rangle + 2\langle v_2, \frac{1}{2} w \rangle.$$

Setzt man hierin $v_1 = v, v_2 = 0$, so ergibt sich

$$\langle v, w \rangle = 2\langle v, \frac{1}{2} w \rangle$$

für alle v und w , was, in die rechte Seite von (+) eingesetzt, zumindest die *Additivität* ergibt:

$$\langle v_1 + v_2, w \rangle = \langle v_1, w \rangle + \langle v_2, w \rangle.$$

Hieraus folgt nun induktiv für alle $n \in \mathbb{N}$, daß

$$\langle nv, w \rangle = n\langle v, w \rangle,$$

und diese Formel bleibt wegen $0 = \langle 0, w \rangle = \langle nv + (-n)v, w \rangle = \langle nv, w \rangle + \langle (-n)v, w \rangle$ auch für alle *ganzen* Zahlen $n \in \mathbb{Z}$ richtig. Genauso schließt man dann

$$\langle av, w \rangle = a\langle v, w \rangle \quad \text{für alle } a \in \mathbb{Q}.$$

Da \mathbb{Q} in dem Körper der reellen Zahlen dicht liegt, braucht man jetzt nur noch zu zeigen, daß bei festen $v, w \in V$ die reellen Funktionen

$$\varphi(a) := \langle av, w \rangle, \quad \psi(a) := a\langle v, w \rangle, \quad a \in \mathbb{R},$$

stetig sind. Dies ist für ψ natürlich trivialerweise richtig. Um die Stetigkeit von φ einzusehen, genügt es wegen der Definition von $\langle \cdot, \cdot \rangle$, die Stetigkeit der Funktion $\|av + w\|^2$ und damit auch die der Funktion $\nu(a) = \|av + w\|$ bei fest vorgegebenen v, w zu beweisen. Nun ist in einem normierten Vektorraum stets die Ungleichung

$$\left| \|x\| - \|y\| \right| \leq \|x - y\|$$

erfüllt. Damit ergibt sich sofort für alle $a, b \in \mathbb{R}$:

$$|\nu(a) - \nu(b)| = \left| \|av + w\| - \|bv + w\| \right| \leq \|(a - b)v\| = |a - b| \cdot \|v\|,$$

also sogar die gleichmäßige Stetigkeit von ν . □

Bemerkung. Alternativ kann man im letzten Teil des vorigen Beweises nachprüfen, daß der Beweis für die Cauchy-Schwarzsche Ungleichung zumindest mit rationalen Koeffizienten und dann durch Grenzübergang auch *allgemein* in unserer Situation durchgeht. Wir haben nämlich für alle $v, w \in V, a \in \mathbb{Q}$:

$$0 \leq \|av - w\|^2 = \langle av - w, av - w \rangle = a^2 \|v\|^2 - 2a \langle v, w \rangle + \|w\|^2.$$

Da die rechte Seite stetig in a ist, folgt diese Ungleichung für alle $a \in \mathbb{R}$, und man kann dann wie oben (siehe auch [42]) weiterschließen, daß

$$|\langle v, w \rangle| \leq \|v\| \cdot \|w\|.$$

Ist nun $a = \lim_j a_j$ eine beliebige reelle Zahl, $a_j \in \mathbb{Q}$, so gilt für beliebige $v, w \in V$:

$$\begin{aligned} |\langle av, w \rangle - a_j \langle v, w \rangle| &= |\langle av, w \rangle - \langle a_j v, w \rangle| = |\langle (a - a_j)v, w \rangle| \\ &\leq \|(a - a_j)v\| \cdot \|w\| = |a - a_j| \cdot \|v\| \|w\| \xrightarrow{j \rightarrow \infty} 0 \end{aligned}$$

und damit

$$a \langle v, w \rangle = \left(\lim_{j \rightarrow \infty} a_j \right) \langle v, w \rangle = \lim_{j \rightarrow \infty} (a_j \langle v, w \rangle) = \langle av, w \rangle. \quad \square$$

Folgerung 19.23 Die Maximum-Norm kommt nicht von einem Skalarprodukt her.

Beweis. Man wähle z. B. $v = (1, 0), w = (0, 1)$. □

Der \mathbb{R}^n mit euklidischem Skalarprodukt $\langle \cdot, \cdot \rangle$ besitzt die *Orthonormalbasis* (e_1, \dots, e_n) :

$$\langle e_j, e_k \rangle = \delta_{jk}.$$

Wir wollen zeigen, daß jeder (endlich-dimensionale) euklidische Vektorraum solche Basen besitzt. Genauer gilt:

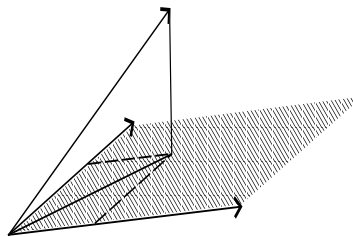
Satz 19.24 (Schmidtsches Orthonormalisierungs - Verfahren) *Es sei V ein euklidischer Vektorraum mit Skalarprodukt $\langle \cdot, \cdot \rangle$ und Basis (u_1, \dots, u_n) . Dann gibt es eine Orthonormalbasis (v_1, \dots, v_n) mit*

$$\text{span}(u_1, \dots, u_j) = \text{span}(v_1, \dots, v_j), \quad j = 1, \dots, n.$$

Beweis. Setze $U_j := \text{span}(u_1, \dots, u_j), j = 1, \dots, n$. Im Falle $j = 1$ braucht man nur $v_1 = u_1 / \|u_1\|$ zu wählen. Sei die Aussage für $j < n$ schon bewiesen. Dann ist

$$u_{j+1} \notin U_j = \text{span}(v_1, \dots, v_j).$$

Es sei nun \tilde{u}_{j+1} die *orthogonale Projektion* von u_{j+1} nach U_j : $\tilde{u}_{j+1} := \sum_{k=1}^j \langle u_{j+1}, v_k \rangle v_k$.



Figur 19.8

Dann steht

$$\tilde{v}_{j+1} = u_{j+1} - \tilde{u}_{j+1}$$

(tatsächlich) senkrecht auf U_j :

$$\begin{aligned} \langle \tilde{v}_{j+1}, v_\ell \rangle &= \langle u_{j+1}, v_\ell \rangle - \sum_{k=1}^j \langle u_{j+1}, v_k \rangle \langle v_k, v_\ell \rangle \\ &= \langle u_{j+1}, v_\ell \rangle - \langle u_{j+1}, v_\ell \rangle = 0, \quad \ell = 1, \dots, j. \end{aligned}$$

Es ist ferner $\tilde{v}_{j+1} \neq 0$, da sonst $u_{j+1} = \tilde{u}_{j+1} \in \text{span}(v_1, \dots, v_j) = U_j$ wäre. Also ist

$$(v_1, \dots, v_{j+1}) \quad \text{mit} \quad v_{j+1} = \frac{\tilde{v}_{j+1}}{\|\tilde{v}_{j+1}\|}$$

orthonormiert, und

$$\begin{aligned} \text{span}(v_1, \dots, v_{j+1}) &= \text{span}(v_1, \dots, v_j, \tilde{v}_{j+1}) = \text{span}(v_1, \dots, v_j, u_{j+1}) \\ &= \text{span}(u_1, \dots, u_j, u_{j+1}) = U_{j+1}. \end{aligned}$$

Für $j = n$ ergibt sich $V = \text{span}(v_1, \dots, v_n)$, und wegen $\dim V = n$ muß (v_1, \dots, v_n) eine Basis von V bilden. \square

Das letzte Argument ist bei unendlicher Dimension nicht anwendbar. Deswegen ist eine andere Aussage nützlich, die aber auch für endliche Dimension von Interesse ist.

Satz 19.25 *Es sei $(v_\iota)_{\iota \in I}$ ein orthonormales System von Vektoren $v_\iota \neq 0$. Dann ist $(v_\iota)_{\iota \in I}$ linear unabhängig.*

Beweis. Die Voraussetzung bedeutet natürlich $v_\iota \perp v_\kappa$ für alle $\iota \neq \kappa$ in I und $\|v_\iota\| = 1$ für alle $\iota \in I$. Ist nun für eine endliche Teilmenge $I_0 \subset I$

$$\sum_{\iota \in I_0} a_\iota v_\iota = 0,$$

so multipliziert man mit einem beliebigen v_κ , $\kappa \in I_0$, und erhält

$$a_\kappa = a_\kappa \|v_\kappa\|^2 = \sum_{\iota \in I_0} a_\iota \langle v_\iota, v_\kappa \rangle = \left\langle \sum_{\iota \in I_0} a_\iota v_\iota, v_\kappa \right\rangle = 0. \quad \square$$

Bei komplexen Vektorräumen muß man notgedrungen die „euklidischen“ Axiome modifizieren. Z. B. hat die naheliegende Form

$$\langle z, w \rangle = \sum_{j=1}^n z_j w_j$$

auf \mathbb{C}^n isotrope Vektoren (für $n \geq 2$), ist also ausgeartet (das Produkt von ${}^t(1, i, 0, \dots, 0)$ mit sich selbst ist gleich Null). Außerdem ist diese Form (auch schon für $n = 1$) nicht positiv-definit. Aber nur unter dieser Bedingung können wir jedem $z \in \mathbb{C}^n$ eine „vernünftige“ (sprich: *reelle*) Länge zuordnen. Dieses Malheur passiert übrigens für *jede* komplexe Bilinearform, denn es ist ja stets für alle $v \in V$:

$$\langle iv, iv \rangle = i^2 \langle v, v \rangle = -\langle v, v \rangle,$$

und nur einer der beiden Ausdrücke $\langle v, v \rangle$ und $\langle iv, iv \rangle$ kann positiv sein.

Man muß sich daher von der Linearität in beiden Argumenten trennen und definiert in leichter Modifikation:

Definition. Eine *Sesquilinearform* auf einem komplexen Vektorraum V ist eine Abbildung

$$\langle \cdot, \cdot \rangle : V \times V \rightarrow \mathbb{C},$$

die linear im 2. und antilinear (bzgl. Konjugation) im 1. Argument ist, d. h.:

$$\begin{aligned} \langle v, aw_1 + bw_2 \rangle &= a \langle v, w_1 \rangle + b \langle v, w_2 \rangle, \\ \langle av_1 + bv_2, w \rangle &= \bar{a} \langle v_1, w \rangle + \bar{b} \langle v_2, w \rangle. \end{aligned}$$

Hinweis. Manchmal wird in der Literatur auch die Antilinearität im zweiten Argument und die Linearität im ersten Argument verlangt. „Sesqui“ heißt in etwa „anderthalbfach“.

Beispiele für solche Vektorräume mit Sesquilinearform sind:

1. \mathbb{C}^n mit $\langle z, w \rangle = \sum_{j=1}^n \bar{z}_j w_j$.
2. $\mathcal{C}^0(I, \mathbb{C})$ mit $\langle f, g \rangle = \int_I \overline{f(x)} g(x) dx$.
3. $\ell_{\mathbb{C}}^2 = \{(a_j)_{j \in \mathbb{N}} : a_j \in \mathbb{C}, \sum_{j=0}^{\infty} |a_j|^2 < \infty\}$ mit $\langle (a_j), (b_j) \rangle = \sum_{j=0}^{\infty} \bar{a}_j b_j$.

Konzeptionell kann man die Antilinearität im 1. Argument noch eleganter fassen: Ist V ein \mathbb{C} -Vektorraum, so bezeichne \bar{V} als *Menge* V selbst. Die *Addition* sei ebenfalls genau wie in V erklärt, die *Multiplikation* mit Skalaren $\lambda \in \mathbb{C}$ aber durch

$$\mathbb{C} \times \bar{V} \ni (\lambda, v) \mapsto \bar{\lambda} v \in \bar{V}.$$

\bar{V} ist dann ebenfalls ein \mathbb{C} -Vektorraum, und eine Sesquilinearform auf V ist nichts anderes als eine Bilinearform

$$\bar{V} \times V \longrightarrow \mathbb{C}.$$

So betrachtet ist es klar, daß wir auch die Symmetrie aus der Definition von euklidischen Skalarprodukten im komplexen Fall nicht einfach übernehmen können.

Definition. Eine Sesquilinearform heißt eine *hermitesche Form*, wenn

$$\langle w, v \rangle = \overline{\langle v, w \rangle} \quad \text{für alle } v, w \in V.$$

Diese Bedingung ist jedenfalls verträglich mit den Linearitätsbedingungen und wird von unseren Standardformen erfüllt. Man beachte, daß man bei Voranstellung dieses Axioms auf die Antilinearitäts-Bedingung im 1. Argument vollständig verzichten kann.

Lemma 19.26 *Eine Abbildung $\langle \cdot, \cdot \rangle : V \times V \rightarrow \mathbb{C}$ ist genau dann hermitesch, wenn*

- i) $\langle w, v \rangle = \overline{\langle v, w \rangle}$,
- ii) $\langle v, aw_1 + bw_2 \rangle = a \langle v, w_1 \rangle + b \langle v, w_2 \rangle$.

Hermitesche Formen werden relativ zu Basen per definitionem durch *hermitesche Matrizen* beschrieben:

$$\langle v, w \rangle = {}^t \bar{v} H w, \quad {}^t H = \bar{H},$$

wobei stets $\bar{H} = (\overline{h_{jk}})$ für $H = (h_{jk})$ gesetzt wird und $\bar{v} \in \mathbb{C}^n$ natürlich den Vektor ${}^t(\bar{z}_1, \dots, \bar{z}_n)$ bezeichnet, wenn $v = {}^t(z_1, \dots, z_n)$. Hermitesche Matrizen transformieren sich demgemäß nach der Formel

$$A = {}^t \bar{S} B S.$$

Für hermitesche Formen ist das Produkt von v mit sich selbst stets reell. Es ist daher sinnvoll, auch für solche Formen den Begriff der *Positiv-Definitheit* einzuführen:

$$\langle v, v \rangle > 0, \quad v \neq 0.$$

Definition. Ein komplexer Vektorraum V heißt *unitär*, wenn er mit einer positiv-definiten hermiteschen Form $\langle \cdot, \cdot \rangle$ versehen ist.

Die meisten Aussagen über euklidische Vektorräume übertragen sich jetzt fast wörtlich auf unitäre Vektorräume. Z. B. hat man:

1. $\|v\| := \langle v, v \rangle^{1/2}$ ist eine Norm auf V .

2. Es gilt die Cauchy-Schwarzsche Ungleichung:

$$|\langle v, w \rangle| \leq \|v\| \cdot \|w\|.$$

3. Orthonormalisierung kann im Endlich-Dimensionalen wie über \mathbb{R} durchgeführt werden.

4. Der *Satz von Pythagoras* muß in unitären Vektorräumen leicht modifiziert werden. Er lautet:

$$\|v + w\|^2 = \|v\|^2 + \|w\|^2 + 2 \operatorname{Re} \langle v, w \rangle.$$

5. Die *Parallelogramm-Gleichung* ist wie im euklidischen Fall gültig (und folgt auch hier schlicht und einfach durch Addition des Pythagoras für $v + w$ und $v - w$):

$$\|v + w\|^2 + \|v - w\|^2 = 2(\|v\|^2 + \|w\|^2).$$

Hieraus ergibt sich sofort:

$$4 \operatorname{Re} \langle v, w \rangle = \|v + w\|^2 - \|v - w\|^2.$$

5. Erinnern wir uns, daß im euklidischen Fall das Skalarprodukt durch die Norm bestimmt ist, so verwundert uns vielleicht diese und die vorvorletzte Formel, die *nur den Realteil* von $\langle v, w \rangle$ durch die Norm auszudrücken scheint. Nun ist aber im unitären Fall $\langle w, v \rangle = \overline{\langle v, w \rangle}$ und damit

$$\operatorname{Im} \langle v, w \rangle = -\operatorname{Im} \overline{\langle v, w \rangle} = -\operatorname{Im} \langle w, v \rangle = \operatorname{Re} i \langle w, v \rangle,$$

so daß wir leicht die folgende Formel erhalten:

$$\begin{aligned} 4 \langle v, w \rangle &= 4 \operatorname{Re} \langle v, w \rangle + 4i \operatorname{Im} \langle v, w \rangle \\ &= 4 \operatorname{Re} \langle v, w \rangle + 4i \operatorname{Re} i \langle w, v \rangle \\ &= 4 \operatorname{Re} \langle v, w \rangle + 4i \operatorname{Re} \langle -iw, v \rangle \\ &= \|v + w\|^2 - \|v - w\|^2 + i \|v - iw\|^2 - i \|v + iw\|^2, \end{aligned}$$

die man wie im Reellen die *Polarisationsformel* nennt.

Definition. Wir sprechen im folgenden allgemein von einem *Skalarprodukt-Raum*, wenn er euklidisch (und dann über \mathbb{R}) bzw. unitär (und dann über \mathbb{C} erklärt) ist.

Wir wollen uns als nächstes dem Problem der *Bestapproximation* in solchen Skalarprodukt-Räumen zuwenden. Im Gegensatz zu lediglich nur normierten Räumen ist die Situation hier wesentlich durchsichtiger. Es gilt der folgende Satz, den wir der Bequemlichkeit halber nur für den euklidischen Fall notieren.

Satz 19.27 *Es sei U ein Untervektorraum des euklidischen Vektorraums V , $v_0 \in V$ sei fest vorgegeben. Es sei ferner $u_0 \in U$. Dann gilt :*

- i) u_0 ist eine Bestapproximation von v_0 bzgl. U genau dann, wenn $v_0 - u_0 \perp u$ für alle $u \in U$;
 ii) wenn eine solche Bestapproximation u_0 existiert, so ist sie eindeutig bestimmt;
 iii) ist $U \subset V$ endlich-dimensional und $\{u_1, \dots, u_n\}$ eine orthonormale Basis von U , so ist

$$u_0 := \sum_{j=1}^n \langle v_0, u_j \rangle u_j$$

eine Bestapproximation von v_0 bzgl. U .

Beweis. Ist $u \in U$ ein beliebiger Vektor, so ist $v_0 - u = (v_0 - u_0) + (u_0 - u)$ und

$$(*) \quad \|v_0 - u\|^2 = \|v_0 - u_0\|^2 + 2\langle v_0 - u_0, u_0 - u \rangle + \|u_0 - u\|^2.$$

Setze nun voraus, daß $v_0 - u_0 \perp w$ für alle $w \in U$ und daß ferner $u \in U$, $u \neq u_0$, fest gewählt sei. Da dann auch $u_0 - u \in U$ ist, so folgt

$$\|v_0 - u\|^2 = \|v_0 - u_0\|^2 + \|u_0 - u\|^2 > \|v_0 - u_0\|^2.$$

Dies beschließt eine Richtung in i) und (unter Voraussetzung der Umkehrrichtung, die gleich gezeigt wird) auch die Eindeutigkeitsaussage in ii).

Sei also umgekehrt $\|v_0 - u\| \geq \|v_0 - u_0\|$ für alle $u \in U$ vorausgesetzt; dann folgt aus (*) die Ungleichung

$$2\langle v_0 - u_0, u_0 - u \rangle + \|u_0 - u\|^2 \geq 0 \quad \text{für alle } u \in U.$$

Da jeder Vektor in U von der Gestalt $u_0 - u$, $u \in U$, ist, so ergibt sich auch

$$(**) \quad 2\langle v_0 - u_0, \tilde{u} \rangle + \|\tilde{u}\|^2 \geq 0 \quad \text{für alle } \tilde{u} \in U.$$

Ersetzt man in (**) den Vektor \tilde{u} durch $-\tilde{u}$, so folgt

$$|\langle v_0 - u_0, \tilde{u} \rangle| \leq \frac{1}{2} \|\tilde{u}\|^2 \quad \text{für alle } \tilde{u} \in U$$

und damit für beliebiges $a \in \mathbb{R}^*$ und $u \in U$:

$$|\langle v_0 - u_0, u \rangle| = \frac{1}{|a|} |\langle v_0 - u_0, au \rangle| \leq \frac{1}{2|a|} \|au\|^2 \leq \frac{|a|}{2} \|u\|^2.$$

Läßt man hierin a bei festem u gegen 0 gehen, so ergibt sich die gewünschte Relation $\langle v_0 - u_0, u \rangle = 0$ für alle $u \in U$.

Will man in dem vorigen Argument den Grenzübergang vermeiden, so setze man in (**) bei festem $u \in U$, $u \neq 0$, den Vektor

$$\tilde{u} = -\frac{\langle v_0 - u_0, u \rangle}{\|u\|^2} u$$

ein. Es ergibt sich unmittelbar die Ungleichung

$$-\frac{|\langle v_0 - u_0, u \rangle|^2}{\|u\|^2} \geq 0,$$

aus der ebenfalls die Behauptung folgt.

Wegen i) und ii) brauchen wir für iii) nur zu zeigen, daß $v_0 - u_0$ auf jedem Vektor u_j , $j = 1, \dots, n$, senkrecht steht. Dies ist aber gerade die Aussage des Schmidtschen Orthogonalisierungs-Verfahrens. \square

Diesen Satz kann man ganz erheblich verallgemeinern. Wie wir in dem Abschnitt über Bestapproximation in *normierten Vektorräumen* schon gesehen haben, spielt *Konvexität* hier eine wesentliche Rolle. (Zu weiteren Informationen siehe z. B. HEUSER [157]). Wir begnügen uns an dieser Stelle mit dem Nachweis des folgenden Satzes.

Satz 19.28 *Es sei V ein Skalarprodukt-Raum und $K \subset V$ eine nicht leere konvexe Menge, die als metrischer Unterraum von V vollständig sei. Dann besitzt jedes Element v_0 eine eindeutig bestimmte Bestapproximation bzgl. K .*

Bemerkung. Lineare Unterräume sind automatisch konvex und, wenn sie endlich-dimensional sind, wegen der Äquivalenz aller Normen auch vollständig. Von daher ist Satz 28 eine weitreichende Verallgemeinerung von Satz 27.

Beweis (von Satz 28). Es sei $\alpha := \min_{u \in K} \|v_0 - u\|$ und $u_j \in U$ eine Folge mit

$$\alpha = \lim_{j \rightarrow \infty} \|v_0 - u_j\|.$$

In der Parallelogramm-Gleichung $\|v + w\|^2 + \|v - w\|^2 = 2(\|v\|^2 + \|w\|^2)$ setzen wir jetzt $v := v_0 - u_j$, $w := v_0 - u_k$. Damit erhält man

$$\begin{aligned} \|u_j - u_k\|^2 &= 2\|v_0 - u_j\|^2 + 2\|v_0 - u_k\|^2 - 4\left\|v_0 - \frac{u_j + u_k}{2}\right\|^2 \\ &\leq 2\|v_0 - u_j\|^2 + 2\|v_0 - u_k\|^2 - 4\alpha \xrightarrow{j,k \rightarrow \infty} 0, \end{aligned}$$

da wegen der Konvexität von K das Element $(1/2)(u_j - u_k)$ ebenfalls in K liegt. $(u_j)_{j \in \mathbb{N}}$ ist folglich eine Cauchy-Folge in K , die nach Voraussetzung gegen einen Grenzwert $u_0 \in K$ konvergiert. Wegen der Stetigkeit der Metrik ist dann $\|v_0 - u_0\| = \alpha$, also u_0 eine Bestapproximation. Ist \tilde{u}_0 eine weitere solche, so kann man für die Folge u_j im obigen Beweis die Folge $u_0, \tilde{u}_0, u_0, \tilde{u}_0, \dots$ nehmen. Sie ist dann ebenfalls eine Cauchy-Folge, was nur möglich ist, wenn $\tilde{u}_0 = u_0$. \square

Aufgrund des zuvor bewiesenen Satzes sind Bedingungen von Interesse, wann ein Unterraum eines normierten oder sogar nur metrischen Raumes vollständig ist. Eine einfache Antwort gibt

Satz 19.29 *Es sei (X, d_X) ein metrischer Raum und $Y \subset X$ ein Unterraum.*

- i) *Ist Y vollständig, so ist Y abgeschlossen in X .*
- ii) *Ist X vollständig und ist Y abgeschlossen in X , so ist auch Y vollständig.*

Der *Beweis* ist denkbar einfach.

i) Ist (w_j) eine Folge in Y , die gegen ein Element $w \in X$ konvergiert, so ist sie eine Cauchy-Folge sowohl in X als auch in Y und muß deshalb gegen ein Element in Y konvergieren, das notwendig gleich w ist.

ii) Ist (w_j) eine Cauchy-Folge in Y , so ist sie auch eine solche in X und damit gegen ein Element $v \in X$ konvergent. Da Y abgeschlossen in X ist, muß v ein Element von Y sein. \square

Wir erinnern jetzt daran, daß ein Skalarprodukt-Raum ein **HILBERT-Raum** genannt wird, wenn er als normierter Raum bzgl. der kanonischen Norm vollständig ist. Wir können damit notieren:

Folgerung 19.30 *Abgeschlossene Teilmengen eines Hilbert-Raumes V sind vollständig; insbesondere sind abgeschlossene lineare Unterräume $U \subset V$ wieder Hilbert-Räume mit dem induzierten Skalarprodukt. Zu jedem $v_0 \in V$ gibt es dann eine eindeutig bestimmte Bestapproximation bzgl. U . Dies gilt im Besonderen für alle endlich-dimensionalen linearen Unterräume in V .*

Beispiel. Wir haben im Anhang zu Kapitel 8 die Hilbert-Räume der quadrat-summierbaren Folgen $\ell_{\mathbb{R}}^2$ und $\ell_{\mathbb{C}}^2$ als Vervollständigung der Unterräume der Folgen „mit endlichem Träger“ konstruiert. Da diese Unterräume nicht selbst schon vollständig sind, sind sie auch nicht abgeschlossen in dem jeweiligen Hilbertschen Folgenraum. Aus Gründen, die sehr bald einsichtig werden, wollen wir diese Beispiele noch

etwas verallgemeinern. Wir schreiben \mathbb{K} für einen der beiden Körper \mathbb{R} oder \mathbb{C} und betrachten zu einer beliebigen Menge T den Vektorraum aller Funktionen $\varphi : T \rightarrow \mathbb{K}$ mit $\varphi(t) \neq 0$ für höchstens abzählbar viele $t \in T$ und

$$\sum_{t \in T} |\varphi(t)|^2 < \infty,$$

wobei selbstverständlich nur über die $t \in T$ zu summieren ist, für die $\varphi(t) \neq 0$. Wir schreiben $\ell^2(T)$ oder genauer $\ell_{\mathbb{K}}^2(T)$ für diesen Vektorraum, zusammen mit dem (wohldefinierten) Skalarprodukt

$$\langle \varphi, \psi \rangle := \sum_{t \in T} \overline{\varphi(t)} \psi(t).$$

Im Spezialfall $T = \{1, \dots, n\}$ schreiben wir auch $\ell_{\mathbb{K}}^2(n)$; dies ist natürlich nichts anderes als der euklidische \mathbb{R}^n bzw. unitäre \mathbb{C}^n . Im Fall $T = \mathbb{N}$ ist $\ell_{\mathbb{K}}^2(\infty) := \ell_{\mathbb{K}}^2(\mathbb{N}) = \ell_{\mathbb{K}}^2$. - Wir überlassen dem Leser den einfachen Nachweis des folgenden Satzes.

Lemma 19.31 Die Skalarprodukt-Räume $\ell_{\mathbb{K}}^2(T)$ sind vollständig, also Hilbert-Räume.

Wir werden im folgenden unter anderem zeigen, daß jeder Hilbert-Raum über \mathbb{K} normisomorph zu einem der Räume $\ell_{\mathbb{K}}^2(T)$ bei geeignetem T ist. Man beachte, daß wegen der Polarisationsformeln ein linearer Isomorphismus zwischen Skalarprodukt-Räumen, der die Normen erhält, automatisch auch die Skalarprodukte invariant läßt.

Definition. Ein indiziertes System $(u_s)_{s \in S}$ in einem Skalarprodukt-Raum V heißt ein *Orthonormal-System*, wenn $\langle u_s, u_{s'} \rangle = \delta_{ss'}$ für alle $s, s' \in S$. Ein solches System heißt *vollständig*, wenn es nicht (orthonormal) vergrößert werden kann, d. h. wenn aus $u \perp u_s$ für alle $s \in S$ stets $u = 0$ folgt. Für ein *Orthogonal-System* fordert man lediglich, daß $\langle u_s, u_{s'} \rangle = 0$ für alle $s \neq s'$ in S . Orthonormal-Systeme sind linear unabhängig (siehe Satz 25). Der Einfachheit halber werden wir im folgenden die Menge S mit dem indizierten System $(u_s)_{s \in S}$ identifizieren, d. h. $S \subset V$ annehmen.

Eine einfache Folgerung aus dem ZORNschen Lemma ist die Tatsache, daß jeder nichttriviale Skalarprodukt-Raum $V \neq \{0\}$ vollständige Orthonormal-Systeme besitzt. Es gilt sogar wesentlich mehr.

Satz 19.32 Jedes Orthonormal-System in einem nichttrivialen Skalarprodukt-Raum $V \neq \{0\}$ läßt sich vervollständigen. Insbesondere existieren vollständige Orthonormal-Systeme in V .

Beweis. Es sei $S_0 \subset V$ ein gegebenes Orthonormal-System. Betrachte dann die Menge aller Orthonormal-Systeme S , die S_0 umfassen. Diese ist bzgl. der Inklusionsrelation induktiv geordnet und besitzt daher ein maximales Element. Für die zweite Aussage wählt man ein Element $v_0 \in V$ mit $\|v_0\| = 1$ und setzt $S_0 := \{v_0\}$. \square

Beispiel. Wir betrachten im Raum der stetigen 2π -periodischen Funktionen $f : \mathbb{R} \rightarrow \mathbb{K}$ die Funktionen

$$\frac{1}{\sqrt{2\pi}}, \frac{\cos x}{\sqrt{\pi}}, \frac{\sin x}{\sqrt{\pi}}, \frac{\cos 2x}{\sqrt{\pi}}, \frac{\sin 2x}{\sqrt{\pi}}, \dots$$

(im Falle $\mathbb{K} = \mathbb{R}$) bzw. (im Falle $\mathbb{K} = \mathbb{C}$)

$$\frac{e^{inx}}{\sqrt{2\pi}}, n \in \mathbb{Z}.$$

Wir haben im Hauptteil dieses Kapitels nachgerechnet, daß diese Systeme orthonormal sind. Sie sind aber auch vollständig in den zu Grunde liegenden Vektorräumen mit dem üblichen L^2 -Skalarprodukt. Ist nämlich z. B. $f : \mathbb{R} \rightarrow \mathbb{C}$ stetig und 2π -periodisch mit

$$\langle f, e_n \rangle = \int_0^{2\pi} f(x) e^{-inx} dx = 0$$

für alle $n \in \mathbb{Z}$, so bedeutet dies, daß alle *Fourier-Koeffizienten* c_n von f verschwinden. Damit konvergiert die Nullfolge im quadratischen Mittel gegen f , d. h. es ist

$$\int_0^{2\pi} |f(x)|^2 dx = 0,$$

woraus wegen der Stetigkeit von f sofort $f = 0$ folgt.

Bemerkung. (Orthonormal-) Systeme von *speziellen Polynomen* und *Funktionen* spielen in der *Mathematischen Physik* eine große Rolle. So führen die *HERMITESchen Polynome*

$$H_n(x) := (-1)^n e^{x^2} \frac{d^n}{dx^n} e^{-x^2}, \quad n \in \mathbb{N},$$

zu dem (vollständigen) Orthonormal-System der *HERMITESchen Funktionen*

$$\frac{1}{\sqrt{2^n n! \sqrt{\pi}}} e^{-x^2/2} H_n(x), \quad n \in \mathbb{N},$$

in $L^2(\mathbb{R})$. Entsprechend hat man die *LEGENDRE-Polynome* und *-Funktionen* für $L^2([-1, 1])$ und die *LAGUERRE-Polynome* und *-Funktionen* für $L^2(\mathbb{R}_+)$. Sie alle erfüllen wichtige Differentialgleichungen der Physik. Auch die *TSCHEBYSCHEFF-Polynome* bilden (nach geeigneter Normierung) ein vollständiges Orthonormal-System in dem Raum $C^0((-1, 1))$, wenn dieser mit einem „gewichteten“ Skalarprodukt versehen wird:

$$\langle f, g \rangle := \int_{-1}^1 \overline{f(x)} g(x) \frac{1}{\sqrt{1-x^2}} dx.$$

Die klassische Fourier-Theorie legt es nun nahe, ganz allgemein *Fourier-Analysis* in Skalarprodukt-Räumen V relativ zu vorgegebenen (vollständigen) Orthonormal-Systemen $S \subset V$ zu betreiben. Wir ordnen jedem $v \in V$ die *Fourier-Koeffizienten* $c_u := \langle v, u \rangle$ zu, wobei $u \in S$, und fragen ganz naiv, ob - oder wann - der Vektor v die *Fourier-Entwicklung*

$$(\times) \quad v = \sum_{u \in S} c_u u$$

besitzt. Dies ist insbesondere dadurch gerechtfertigt, daß für jede *endliche* Teilmenge $S_0 \subset S$ die endliche Summe

$$\sum_{u \in S_0} c_u u$$

die *Bestapproximation* von v bzgl. des durch S_0 in V erzeugten endlich-dimensionalen Untervektorraums ist.

Nun scheint aber die Summe auf der rechten Seite von (\times) keinen rechten Sinn zu ergeben außer, wenn die Menge S höchstens abzählbar ist. Dies braucht selbstverständlich nicht der Fall zu sein; glücklicherweise wird hier stets nur über eine höchstens abzählbare Teilmenge $S' \subset S$ summiert. - Dies ist der Inhalt des übernächsten Satzes. Zuvor notieren wir noch die verallgemeinerte *BESSELsche Gleichung* und *Ungleichung*.

Lemma 19.33 *Es sei S_0 ein endliches Orthonormal-System in dem Skalarprodukt-Raum V . Dann gilt für jedes $v \in V$ die Gleichung*

$$\left\| v - \sum_{u \in S_0} c_u u \right\|^2 = \|v\|^2 - \sum_{u \in S_0} |\langle v, u \rangle|^2$$

und damit auch die *Ungleichung*

$$\sum_{u \in S_0} |\langle v, u \rangle|^2 \leq \|v\|^2.$$

Beweis. Man rechnet direkt nach, daß für beliebige Koeffizienten γ_u die Gleichung

$$\left\| v - \sum_{u \in S_0} \gamma_u u \right\|^2 = \|v\|^2 - \sum_{u \in S_0} |\langle v, u \rangle|^2 + \sum_{u \in S} |c_u - \gamma_u|^2$$

gilt. Hieraus folgt die erste Behauptung mit $\gamma_u = c_u$, und die zweite ist eine unmittelbare Konsequenz aus der ersten. \square

Satz 19.34 *Unter den weiter oben gemachten Voraussetzungen ist bei vorgegebenem $v \in V$ der Fourier-Koeffizient $c_u := \langle v, u \rangle$ für höchstens abzählbar viele $u \in S$ von Null verschieden. Summiert man nur über diese, so gilt auch hier stets die verallgemeinerte Besselsche Ungleichung in der Form*

$$\sum_{u \in S} |\langle v, u \rangle|^2 \leq \|v\|^2.$$

Beweis. Bei vorgegebenem $\varepsilon > 0$ seien u_1, \dots, u_n Elemente in S mit $|\langle v, u_j \rangle| \geq \varepsilon$ für $j = 1, \dots, n$. Nach der Besselschen Ungleichung in Lemma 33 ist dann $n\varepsilon^2 \leq \|v\|^2$ und damit $n \leq \varepsilon^{-2}\|v\|^2$. Es können daher höchstens endlich viele solcher u_j bei gegebenem ε vorhanden sein. Läßt man jetzt noch ε eine Nullfolge durchlaufen, so folgt die Behauptung. Die Besselsche Ungleichung gewinnt man durch Grenzübergang aus der entsprechenden Formel bei endlichen Summen. \square

Der alles entscheidende Satz ist nun der folgende:

Satz 19.35 *Ist S ein beliebiges Orthonormal-System in dem Skalarprodukt-Raum V , so konvergiert die Fourier-Reihe*

$$\sum_{u \in S} \langle v, u \rangle u$$

von v (in beliebiger Anordnung der höchstens abzählbar vielen von Null verschiedenen Summanden) genau dann gegen v , falls die Parsevalsche Gleichung

$$\sum_{u \in S_0} |\langle v, u \rangle|^2 = \|v\|^2$$

besteht. In einem Hilbert-Raum V ist diese Beziehung genau dann für alle $v \in V$ gültig, wenn das System S vollständig ist.

Der *Beweis* ist nach den bisherigen Ausführungen nicht mehr schwer und kann dem Leser zur Übung überlassen werden. (Siehe z. B. HEUSER [157], Satz 24.1, Satz 24.3). \square

Definition. Aus naheliegenden Gründen nennt man daher ein *vollständiges Orthonormalsystem* in einem *Hilbert-Raum* auch eine *Orthonormal-Basis*. Man beachte jedoch, daß dieser Basis-Begriff im *analytischen* Sinne zu verstehen ist, nicht im Sinne der Linearen Algebra.

In allen „echten“ Beispielen, die uns bisher begegnet sind, ist die Kardinalität der Orthonormal-Basen die von \mathbb{N} . Dies hat mit der *Separabilität* der unterliegenden metrischen Räume zu tun. Wir erinnern an die

Definition und Beispiel. Ein metrischer Raum (X, d_X) heißt *separabel*, wenn er eine abzählbare dichte Teilmenge besitzt. Dies ist z. B. erfüllt für \mathbb{R}^n bzgl. irgendeiner Norm, da \mathbb{Q}^n eine solche Teilmenge ist.

Man beweist nun leicht den folgenden

Satz 19.36 *Jeder separable Skalarprodukt-Raum besitzt eine höchstens abzählbare Orthonormal-Basis.*

Beweis. Siehe loc. cit., Satz 24.2. □

Zum Schluß skizzieren wir den Beweis für die *Klassifikation der Hilbert-Räume*.

Satz 19.37 *Jeder Hilbert-Raum V ist als Skalarprodukt-Raum isomorph zu einem Raum $\ell_{\mathbb{K}}^2(T)$. Hierbei ist T eine (ansonsten beliebige) Menge von der gleichen Kardinalität wie ein vollständiges Orthonormal-System S von V .*

Beweis. Wir schreiben $S = \{u_t : t \in T\}$ und haben damit für jedes $v \in V$ die Entwicklungen

$$v = \sum_{t \in T} \langle v, u_t \rangle u_t \quad \text{und} \quad \|v\|^2 = \sum_{t \in T} |\langle v, u_t \rangle|^2.$$

Nach Satz 34 sind in diesen beiden Summen höchstens abzählbar viele Fourier-Koeffizienten von Null verschieden. Die $v \in V$ zugeordnete Funktion $\varphi_v(t) := \langle v, u_t \rangle$ von T mit Werten in \mathbb{K} liegt damit in $\ell_{\mathbb{K}}^2(T)$, so daß die durch $v \mapsto \varphi_v$ definierte Abbildung $\Phi : V \rightarrow \ell_{\mathbb{K}}^2(T)$ normerhaltend und trivialerweise linear ist. Wegen der Normerhaltung ist sie automatisch injektiv, so daß nur noch die *Surjektivität* nachzuprüfen ist. Es sei also ohne Einschränkung $T' \subset T$ so gewählt, daß für $\varphi \in \ell_{\mathbb{K}}^2(T)$ die Werte $\varphi(t) = 0$ sind für $t \in T \setminus T'$ und $T' \simeq \mathbb{N}$. Wir schreiben dann j für die Elemente in T' und u_j , $j \in \mathbb{N}$, für die entsprechende Teilmenge von S . Ferner sei $c_j := \varphi(j)$ und

$$v_n := \sum_{j=0}^n c_j u_j.$$

Da die u_j ein Orthonormal-System bilden, gewinnt man aus dem Satz von Pythagoras für $m \leq n$ die Abschätzung

$$\|s_n - s_{m-1}\|^2 = \sum_{j=m}^n |c_j|^2 \leq \varepsilon^2,$$

sofern nur $m \geq N = N(\varepsilon)$, da nach Voraussetzung die Reihe $\sum_{j=0}^{\infty} |c_j|^2$ konvergiert. Also ist

$$\sum_{j=0}^{\infty} c_j u_j$$

eine Cauchy-Reihe, die wegen der vorausgesetzten Vollständigkeit von V gegen ein Element v konvergent ist. Wegen der Stetigkeit des Skalarprodukts folgt dann schließlich

$$\langle v, u_k \rangle = \left\langle \sum_{j=0}^{\infty} c_j u_j, u_k \right\rangle = \sum_{j=0}^{\infty} c_j \langle u_j, u_k \rangle = c_k,$$

also die Surjektivität des Homomorphismus Φ . □

Als Folgerung erhält man im Falle $T = \mathbb{N}$ den folgenden

Satz 19.38 *Alle (unendlich-dimensionalen) separablen Hilbert-Räume über \mathbb{K} sind isomorph zu $\ell_{\mathbb{K}}^2$.*

20 Gewöhnliche Differentialgleichungen

Viele Fragen aus der *Geometrie* und der *Mechanik*, wie auch aus fast allen Bereichen der Theoretischen Physik, führen auf *gewöhnliche* und sogar *partielle Differentialgleichungen*. Wie man solche Differentialgleichungen im einfachsten Fall aufstellt und löst, wollen wir an zwei Beispielen demonstrieren.

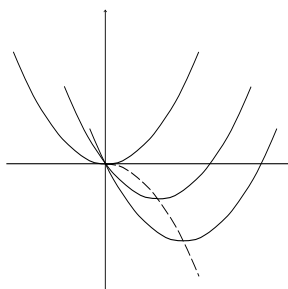
Beispiel 1. Man verschiebe die Parabel $y = x^2$ in \mathbb{R}^2 so, daß sie weiterhin durch den Nullpunkt verläuft. Zu der hierdurch entstehenden „Kurvenschar“ bestimme man die Schar der *orthogonalen* Kurven.

Die Gesamtheit *aller* verschobenen Parabeln wird beschrieben durch

$$y - b = (x - a)^2, \quad (a, b) \in \mathbb{R}^2.$$

Damit der Nullpunkt auf einer dieser Parabeln liegt, ist notwendig und hinreichend, daß $a^2 = -b$. Die zu betrachtende „Schar“ von Parabeln ist daher gegeben durch $P_a : y = x(x - 2a)$, $a \in \mathbb{R}$. Durch einen vorgegebenen Punkt $(x_0, y_0) \in \mathbb{R}^2$ geht genau eine Kurve dieser Schar, wenn $x_0 \neq 0$, und zwar P_a mit

$$a = \frac{x_0^2 - y_0}{2x_0}.$$



Figur 20.1

Wir werden uns daher im folgenden auf die Betrachtung der Menge $G := \mathbb{R}^* \times \mathbb{R} = \{(x_0, y_0) \in \mathbb{R}^2 : x_0 \neq 0\}$ beschränken (müssen). Da die Ableitung von P_a nach x an der Stelle x_0 gleich $2(x_0 - a) = x_0 + y_0/x_0$ ist, bestimmt die Schar von Parabeln auf G das „Richtungsfeld“ $P(x, y) = x + y/x$. Man sieht sofort, daß P einer lokalen Lipschitz-Bedingung genügt³³ (oder schließt dies daraus, daß die Funktion P nach y stetig partiell differenzierbar ist). Also gibt es durch jeden Punkt $(x_0, y_0) \in G$ genau eine Lösung der Differentialgleichung

$$y' = P(x, y) = \frac{x^2 + y}{x},$$

und da die betrachteten Parabeln tatsächlich Lösungen sind, können wir festhalten:

Die Schar der (Graphen der) P_a (durchschnitten mit G) ist identisch mit der Gesamtheit der Lösungsgraphen der Differentialgleichung $y' = P(x, y)$.

Die Aufgabe verlangt die Aufstellung des zu P *orthogonalen* Richtungsfeldes. Nun ist aber $(-\beta, \alpha)$ orthogonal zu (α, β) im euklidischen \mathbb{R}^2 . Somit wird die Differentialgleichung für die „Orthogonalschar“:

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

³³Zu dieser und den beiden folgenden Behauptungen siehe weiter unten.

auf der Menge $G' := \{(x, y) \in \mathbb{R}^2 : x \neq 0, x^2 + y \neq 0\}$. Ist nun $\varphi(x)$ eine (lokale) Lösung dieser Differentialgleichung, also $(x^2 + \varphi(x))\varphi'(x) = -x$, so ist, wie man sofort nachrechnet,

$$\frac{d}{dx} \left(e^{2\varphi(x)} \left(\varphi(x) + x^2 - \frac{1}{2} \right) \right) = 0.$$

Also ist notwendig

$$\varphi(x) + x^2 - \frac{1}{2} = C e^{-2\varphi(x)},$$

woraus man durch Elimination leicht eine Gleichung für die *Umkehrfunktionen* $x = \psi(y)$ der Lösungen findet:

$$x = \pm \sqrt{\frac{1}{2} - y + C e^{-2y}}.$$

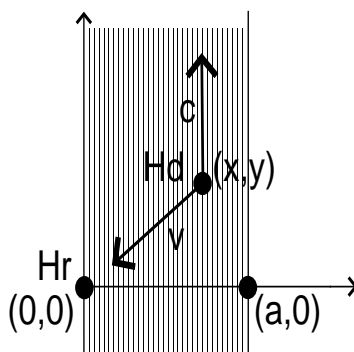
Bemerkung. Betrachtet man von vornherein die Differentialgleichung für die Umkehrfunktionen, also $x'(y) + x + y/x = 0$, so ist diese vom BERNOULLI-Typ³⁴. In der Literatur findet man hierzu reiches Material; insbesondere liefert die Theorie dieses Typs ebenfalls die eben angegebenen Lösungen.

Beispiel 2. Man bestimme die sogenannte *Hundekurve*: Ein Fluß fließt in einem von parallelen Geraden begrenzten Kanalbecken der Breite a mit der konstanten Geschwindigkeit c in Richtung der Uferbefestigung. Ein (selbstverständlich punktförmiger) Hund befindet sich an einem Ufer, sein (ebenfalls punktförmiges) Herrchen steht senkrecht gegenüber am anderen Ufer und pfeift. Daraufhin springt der Hund ins Wasser und schwimmt, relativ zum Wasser, mit konstantem Geschwindigkeitsbetrag $\|v\|$, die „Schnauze“ (macht irgendwie keinen Sinn bei einem punktförmigen Hund, oder doch?) stets auf das Herrchen gerichtet, zu ihm zurück. Welche Kurve legt der Hund zurück? Unter welcher Voraussetzung kann der Hund - nach ausgiebigem Ausschütteln des Felles - sein Herrchen schwanzwedelnd wieder begrüßen?

In geeigneten Koordinaten (siehe Zeichnung) ist die Eigengeschwindigkeit des Hundes wegen der Bedingung, daß seine Schwimmrichtung stets zu seinem Herrchen, der sich im Ursprung befindet, zeigt, gleich

$$v = \frac{v_0}{\sqrt{x^2 + y^2}} (-x, -y) \quad \text{mit} \quad v_0 := \|v\|.$$

Die Strömungsgeschwindigkeit ist dann in der Form $(0, c)$ anzusetzen.



Figur 20.2

³⁴Siehe weiter unten.

Die *Bewegungsgleichungen* des Hundes sind somit (siehe auch den Anhang zu diesem Kapitel):

$$\dot{x}(t) := \frac{dx}{dt} = -\frac{v_0 x}{\sqrt{x^2 + y^2}}, \quad \dot{y}(t) := \frac{dy}{dt} = c - \frac{v_0 y}{\sqrt{x^2 + y^2}}.$$

Offensichtlich ist gerade der Standort $(0, 0)$ des Herrchens eine Stelle, wo die rechte Seite nicht definiert ist. Unsere Aufgabe besteht also darin, Lösungen zu finden, die in endlicher Zeit gegen den Ursprung konvergieren.

Aus den Bewegungsgleichungen folgt sofort, daß die zeitliche Ableitung der x -Koordinate überall negativ ist und somit $x = x(t)$ streng monoton fällt. Also ist auch die verstrichene Zeit t eine (eindeutig bestimmte) Funktion des Ortes: $t = t(x)$, und durch Einsetzen erhalten wir die Gleichung der *Bahnkurve* des Hundes als $y(t(x))$, die wegen der Kettenregel der Differentialgleichung

$$\frac{dy}{dx} = \frac{v_0 y - c\sqrt{x^2 + y^2}}{v_0 x} = \frac{y}{x} - \frac{c}{v_0} \sqrt{1 + \left(\frac{y}{x}\right)^2}$$

genügt. Dies ist eine sogenannte *homogene Differentialgleichung*, die man bekanntlich mit dem Ansatz $z = y/x$, $x > 0$, löst³⁵. Mit $y' = z'x + z$ gewinnt man die Differentialgleichung

$$\frac{z'}{\sqrt{1 + z^2}} = -\lambda \cdot \frac{1}{x}, \quad \lambda := \frac{c}{v_0},$$

die man ohne Umschweife integriert: $\operatorname{Arsinh} z = -\lambda \ln \frac{x}{C}$. Die Anfangsbedingung $z = 0$ für $x = a$ liefert weiter $C = a$ und folglich die Bahngleichung

$$y = xz = x \sinh\left(\lambda \ln \frac{a}{x}\right) = \frac{x}{2} \left(\left(\frac{a}{x}\right)^\lambda - \left(\frac{x}{a}\right)^\lambda \right).$$

Diese Lösungen konvergieren mit $x \searrow 0$ gegen Null für $\lambda < 1$, gegen $a/2$ für $\lambda = 1$ und gegen ∞ für $\lambda > 1$. Der arme Hund hat also höchstens dann eine Chance, sein unbewegtes Herrchen wiederzutreffen, wenn $\lambda < 1$. Um nachzuprüfen, ob er ihn tatsächlich schwanzwedelnd wieder begrüßen kann, müssen wir noch die Zeit ausrechnen, die er benötigt, um am anderen Ufer anzukommen. Diese können wir ohne Schwierigkeiten aus der ersten Bewegungsgleichung bestimmen. Wegen $\cosh^2 - \sinh^2 \equiv 1$ erhält man aus dieser nämlich

$$\dot{x}(t) = \frac{-v_0}{\cosh(\lambda \ln(a/x))},$$

woraus man durch Integration für die benötigte Gesamtzeit T die Formel

$$2v_0 T = \int_0^a \left(\frac{a^\lambda}{x^\lambda} + \frac{x^\lambda}{a^\lambda} \right) dx$$

gewinnt. Bei $\lambda \neq 1$ ist der Wert dieses uneigentlichen Integrals gleich

$$\lim_{x \searrow 0} \left(\frac{x^{1-\lambda}}{(1-\lambda)a^{-\lambda}} + \frac{x^{1+\lambda}}{(1+\lambda)a^\lambda} \right) \Big|_x^a,$$

also ∞ für $\lambda > 1$ und

$$2v_0 T = \frac{2a}{1-\lambda^2} \quad \text{für } \lambda < 1.$$

Hieran sieht man, daß auch für $\lambda < 1$, aber nahe bei 1, der Hund eine sehr lange Zeit braucht, sein Herrchen zu erreichen. Für $\lambda = 1$ ist die Bahn zwar endlich, nämlich ein Stück der Parabel

³⁵Siehe weiter unten.

$y = (a^2 - x^2)/2a$, aber dennoch kommt der Hund niemals an dem Uferpunkt $(0, a/2)$ an.

Bemerkung. Eine bekannte Variante dieses Problems sucht die Bahn des Hundes unter der Voraussetzung, daß die Strömungsgeschwindigkeit des Flusses nicht konstant c_0 , sondern von der Form

$$c(x) = 4c_0 \frac{x(a-x)}{a^2}, \quad 0 \leq x \leq a,$$

ist, d. h. insbesondere, daß der Fluß am Ufer sich gar nicht bewegt und seine größte Strömungsgeschwindigkeit c_0 genau in der Mitte hat.

Bezeichnet v_0 wie in der Lösung zuvor den konstanten Geschwindigkeitsbetrag des Hundes, so lauten jetzt die Bewegungsgleichungen

$$\dot{x} := \frac{dx}{dt} = -\frac{v_0 x}{\sqrt{x^2 + y^2}}, \quad \dot{y} := \frac{dy}{dt} = 4c_0 \frac{x(a-x)}{a^2} - \frac{v_0 y}{\sqrt{x^2 + y^2}},$$

aus denen man sofort die Bahngleichung in der Form

$$\frac{dy}{dx} = \frac{y}{x} - 2\lambda(a-x)\sqrt{x^2 + y^2}, \quad 0 < x \leq a, \quad \lambda := \frac{2c_0}{v_0 a^2},$$

gewinnt. Mit dem Ansatz $z = y/x$ folgt hieraus wieder wie oben eine Differentialgleichung mit getrennten Variablen für z :

$$\frac{z'}{\sqrt{1+z^2}} = -2\lambda(a-x).$$

Durch einfaches „Hochintegrieren“ und unter Beachtung von $z = 0$ für $x = a$ ergibt sich hieraus

$$\ln(z + \sqrt{1+z^2}) = \lambda(x-a)^2$$

und damit

$$y + \sqrt{x^2 + y^2} = x e^{\lambda(x-a)^2}.$$

Bringt man y auf die rechte Seite und quadriert, so hebt sich der Term y^2 heraus, und man kann die Gleichung nach y auflösen. Dabei erhält man die folgende einfache Formel für die gesuchte Lösung:

$$y = x \sinh[\lambda(x-a)^2].$$

Diese Beispiele sollten genügen, um den Begriff einer (*expliziten*) *gewöhnlichen Differentialgleichung erster Ordnung* und des zugehörigen *Lösungsbegriffs* herauszuarbeiten. Wir haben stets die folgende Situation: $G \subset \mathbb{R}^2$ ist ein *Gebiet*, d. h. eine offene zusammenhängende Teilmenge von \mathbb{R}^2 , und $f : G \rightarrow \mathbb{R}$ ist eine stetige Funktion. Unter einer *expliziten gewöhnlichen Differentialgleichung 1. Ordnung* verstehen wir eine Gleichung der folgenden Bauart:

$$(*) \quad y' = f(x, y).$$

Eine Lösung dieser Differentialgleichung besteht in einer differenzierbaren Funktion $\varphi : I \rightarrow \mathbb{R}$ mit

a) $\Gamma_\varphi = \{(x, y) : y = \varphi(x), x \in I\} \subset G,$

b) $\varphi'(x) = f(x, \varphi(x))$ für alle $x \in I.$

Eine *implizite* Differentialgleichung erster Ordnung ist von der Form

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

mit einer stetigen Funktion F auf einem geeigneten Gebiet $G \subset \mathbb{R}^3.$

Beispiel 3. In dem Spezialfall $f(x, y) = f(x)$, bei dem also die Funktion f nicht von der Variablen y explizit abhängt, sind Lösungen gerade Funktionen φ mit $\varphi'(x) = f(x)$, also *Stammfunktionen* von f . Somit werden die Lösungen gegeben durch

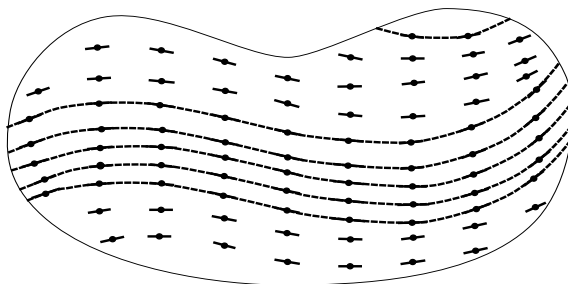
$$\varphi(x) = \int_{x_0}^x f(t) dt + y_0.$$

Bemerkungen. 1. Aus der Bedingung b) folgt, daß jede Lösung automatisch eine *stetig differenzierbare* Funktion auf dem Intervall I ist. Ist die Funktion f sogar stetig partiell differenzierbar, so sind Lösungen aufgrund der Kettenregel sogar zweimal stetig differenzierbar, etc.

2. Eine *geometrische Interpretation von Lösungen* einer Differentialgleichung gewinnt man wie folgt: Durch

$$G \ni (x_0, y_0) \mapsto \{y = y_0 + f(x_0, y_0) \cdot (x - x_0)\} =: T_{(x_0, y_0)}$$

wird jedem Punkt $(x_0, y_0) \in G$ eine Gerade durch (x_0, y_0) mit Steigung $f(x_0, y_0)$ zugeordnet. Insgesamt erhält man also durch die Funktion f ein sogenanntes „Richtungsfeld“ auf dem Gebiet G wie in der folgenden Zeichnung:



Figur 20.3

Lösungen $y = \varphi(x)$ der Differentialgleichung sind dann differenzierbare Funktionen $\varphi : I \rightarrow \mathbb{R}$ mit $(x_0, \varphi(x_0)) \in G$ für alle $x_0 \in I$, deren Graph an jeder Stelle (x_0, y_0) gerade die Tangente $T_{(x_0, y_0)}$ besitzt, wobei $y_0 := \varphi(x_0)$.

3. Eine andere geometrische Interpretation kann man wie folgt gewinnen. Man schreibt die Differentialgleichung in der Form

$$\frac{dy}{dx} = f(x, y)$$

und schreibt sie formal um zu

$$f(x, y) dx - dy = 0.$$

Will man jetzt die Variablen x und y wirklich gleichberechtigt behandeln, so macht man den allgemeineren Ansatz

$$f(x, y) dx + g(x, y) dy = 0$$

mit stetigen Funktionen f, g auf einem Gebiet $G \subset \mathbb{R}^2$. Als *Lösung* sollte man jetzt *differenzierbare Kurven* $\gamma(t) := (\alpha(t), \beta(t))$, $t \in I$, mit $\gamma(t) \in G$ ansehen, die der Gleichung

$$f(\gamma(t)) \alpha'(t) + g(\gamma(t)) \beta'(t) = 0$$

genügen. Bei der ursprünglichen Gleichung sind gerade die zweiten Komponenten von nach x parametrisierten Kurven $\gamma(x) = (x, \varphi(x))$ die Lösungen im klassischen Sinne. Eine schöne Interpretation der Lösungskurven erhält man, wenn die vorgegebene Gleichung *exakt* ist, d. h. wenn es eine (stetig)

partiell differenzierbare Funktion F auf G gibt mit $F_x = f$, $F_y = g$.³⁶ Ist dann nämlich $\gamma(t)$ eine Lösung, so gewinnt man aus der Kettenregel die Identität

$$\begin{aligned} \frac{d}{dt} F(\alpha(t), \beta(t)) &= F_x(\alpha(t), \beta(t)) \alpha'(t) + F_y(\alpha(t), \beta(t)) \beta'(t) \\ &= f(\alpha(t), \beta(t)) \alpha'(t) + g(\alpha(t), \beta(t)) \beta'(t) = 0. \end{aligned}$$

Somit ist F auf der Lösungskurve *konstant*, und umgekehrt ist jede Kurve mit dieser Eigenschaft auch Lösung. Mit anderen Worten: Ist die vorgegebene Differentialgleichung exakt mit Stammfunktion F , so sind die Lösungen gerade die *Niveaulinien* von F .

Es erhebt sich damit die Frage, wann eine Differentialgleichung in der obigen verallgemeinerten Form exakt ist. Setzt man zusätzlich voraus, daß es eine Stammfunktion F gibt, die zweimal stetig differenzierbar ist, so findet man die folgende *notwendige* Bedingung:

$$g_x = (F_y)_x = (F_x)_y = f_y.$$

Wir werden später sehen, daß in *einfach zusammenhängenden* Gebieten G die Bedingung $g_x = f_y$ auch hinreichend ist für die Existenz einer Stammfunktion. Da die Klasse der einfach zusammenhängenden Gebiete die *Kreise* umfaßt (und sogar *konvexe* und noch allgemeiner *sternförmige* Gebiete), lassen sich unter dieser Bedingung jedenfalls *lokal* immer Stammfunktionen finden. Wenn die Bedingung nicht erfüllt ist, so kann man versuchen, sie durch einen kleinen Kunstgriff zu erzwingen. Unser Lösungsbegriff ändert sich ja offensichtlich nicht, wenn wir die beiden Funktionen f und g mit einer festen Funktion λ multiplizieren, die auf G nirgends verschwindet. Man nennt dann λ einen *EULERSchen Multiplikator* oder *integrierenden Faktor*, wenn die multiplizierte Gleichung exakt wird. Dazu ist unter geeigneten Differenzierbarkeitsbedingungen notwendig, daß die unbekannte Funktion λ die *partielle Differentialgleichung*

$$g \lambda_x + g_x \lambda = f \lambda_y + f_y \lambda$$

erfüllt. Eine solche Gleichung zu lösen, ist i. A. wesentlich schwerer, als die Lösungen der gegebenen gewöhnlichen Differentialgleichung zu finden. Allerdings braucht man ja nur eine einzige Lösung zu kennen, die evtl. durch gezieltes Raten aufgespiürt werden kann. Genau dies haben wir bei der Auffindung der Orthogonalschar in unserem ersten Beispiel zu Beginn des Kapitels getan.

Es mag überraschen, daß unter den obigen sehr allgemeinen Voraussetzungen *lokal* stets Lösungen von expliziten Differentialgleichungen 1. Ordnung (zu vorgegebenem *Anfangswert*) existieren.

Satz 20.1 (Peanoscher Existenzsatz) *Ist wie oben $f : G \rightarrow \mathbb{R}$ stetig und $(a, c) \in G$, so gibt es ein offenes Intervall I mit $a \in I$ und eine Lösung $\varphi : I \rightarrow \mathbb{R}$ der Differentialgleichung $y' = f(x, y)$ mit $\varphi(a) = c$.*

Wir werden diesen Satz im Anhang zu dem vorliegenden Kapitel mit zwei verschiedenen Methoden beweisen. Die Beweise lassen sich leicht auch auf *Systeme von Differentialgleichungen erster Ordnung* übertragen.

In unserer Diskussion des Beispiels 1 spielte die *Eindeutigkeit* von Lösungen eine ganz entscheidende Rolle. Diese ist nun aber unter diesen schwachen Voraussetzungen an die Funktion f nicht gewährleistet.

Beispiel 4. Für beliebige Zahlen $a \leq b$ wird durch

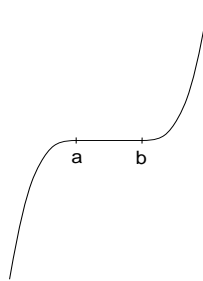
$$\varphi_{ab}(x) := \begin{cases} \sqrt{x-b}^3, & x \geq b \\ 0, & a \leq x \leq b \\ -\sqrt{a-x}^3, & x \leq a \end{cases}$$

³⁶Dies bedeutet in der Theorie der Differentialformen, die wir später entwickeln, daß $dF = f dx + g dy$. Man nennt dann auch die rechts stehende Form exakt und F eine *Stammfunktion*.

eine zweiparametrische Schar von *differenzierbaren* Funktionen $\varphi_{ab} : \mathbb{R} \rightarrow \mathbb{R}$ definiert, die der Differentialgleichung

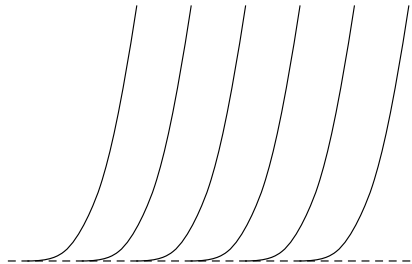
$$y' = \frac{3}{2} y^{1/3}$$

genügen.



Figur 20.4

Man kann leicht einsehen, daß dies in der Tat alle Lösungen der gegebenen Differentialgleichung auf $G := \mathbb{R} \times \mathbb{R}$ sind. Für Anfangswerte $y_0 = 0$ sind damit Lösungen nicht eindeutig bestimmt. Dies ändert sich schlagartig, wenn man sich z. B. auf das Teilgebiet $G_+ := \mathbb{R} \times \mathbb{R}_+^*$ beschränkt. Der Grund hierfür besteht in der Tatsache, daß die Funktion $f(x, y) := y^{1/3}$ auf G nicht nach y stetig partiell differenzierbar ist, wohl aber auf G_+ . Die Gesamtheit der Lösungskurven sieht hier wie folgt aus.



Figur 20.5

Beispiel 5 und Bemerkungen. Das obige Phänomen der „Nichteindeutigkeit“ von Lösungen kann generell auch bei der sogenannten Differentialgleichung erster Ordnung mit *getrennten Variablen* auftreten:

$$y' = f(x)g(y).$$

Hierbei sind $f : I \rightarrow \mathbb{R}$ und $g : J \rightarrow \mathbb{R}$ stetige Funktionen auf offenen Intervallen $I, J \subset \mathbb{R}$. Ist nämlich y_0 eine Nullstelle von g , so ist immer $\varphi(x) = y_0$ eine Lösung der Differentialgleichung auf ganz I . Es kann aber durchaus noch weitere Lösungen geben, die an einer Stelle x_0 den Wert y_0 annehmen. Man beschränkt sich daher bei der Untersuchung dieser Differentialgleichung auf Intervalle J , auf denen g keine Nullstellen besitzt. Wegen der Stetigkeit der Funktion g können wir dann weiter annehmen, daß sie auf ganz J *positiv* ist (ein negatives Vorzeichen können wir bei f aufnehmen). Zu vorgegebenen $(x_0, y_0) \in I \times J$ gibt es dann Stammfunktionen F von f auf I und G von $1/g$ auf J mit $F(x_0) = 0$ und $G(y_0) = 0$. Zudem ist G streng monoton wachsend auf J und besitzt daher eine streng monoton wachsende stetig differenzierbare Umkehrfunktion $H : G(J) \rightarrow J$ auf dem offenen Intervall $G(J) \subset \mathbb{R}$ mit $H(0) = y_0$. Ist nun $\varphi : I' \rightarrow J$, $x_0 \in I' \subset I$, eine Lösung der Differentialgleichung mit $\varphi(x_0) = y_0$, so folgt wegen $\varphi'(x) = f(x)g(\varphi(x))$ unmittelbar

$$\int_{x_0}^x \frac{\varphi'(t)}{g(\varphi(t))} dt = \int_{x_0}^x f(t) dt$$

und hieraus mit der Substitution $u := \varphi(x)$ im linken Integral:

$$\int_{y_0}^{\varphi(x)} \frac{du}{g(u)} = \int_{x_0}^x f(t) dt .$$

Dies bedeutet aber

$$G(\varphi(x)) = F(x) .$$

Wählt man jetzt noch I' so klein, daß $F(I') \subset G(J)$, so ist notwendig

$$\varphi(x) = H(F(x)) .$$

Insbesondere ist die Funktion φ auf I' *eindeutig* bestimmt. Man rechnet nun einfach nach, daß die Funktion $\varphi := H \circ F$ tatsächlich die Differentialgleichung zu dem vorgegebenen Anfangswert löst: Es ist

$$\varphi(x_0) = H(F(x_0)) = H(0) = y_0 ,$$

und wegen $G(\varphi(x)) = F(x)$ erhält man mit der Kettenregel $G'(\varphi(x))\varphi'(x) = F'(x) = f(x)$ und folglich

$$\varphi'(x) = f(x)g(\varphi(x)) , \quad x \in I' .$$

Wir haben damit den folgenden Satz bewiesen (F und G seien wie zuvor gewählt):

Satz 20.2 *Es sei $y' = f(x)g(y)$ eine Differentialgleichung mit getrennten Variablen auf $I \times J$ und $(x_0, y_0) \in I \times J$. Ist dann das Intervall I' mit $x_0 \in I' \subset I$ (maximal) so gewählt, daß $F(I') \subset G(J)$, so existiert genau eine Lösung φ der gegebenen Differentialgleichung auf I' mit $\varphi(x_0) = y_0$.*

Bemerkung. Rein „formal“ besteht die eben geschilderte Lösungsmethode in der Umformung der Gleichung

$$\frac{dy}{dx} = f(x)g(y)$$

in

$$\frac{dy}{g(y)} = f(x) dx$$

mit anschließender Integration und Auflösung nach y . Hierbei sollte man aber nicht die evtl. notwendige Beschränkung des Existenzintervalls der Lösungen, wie sie durch den vorstehenden Satz explizit aufgezeigt werden, außer Acht lassen. Man sieht übrigens unmittelbar, daß wir es hier mit einer *exakten* Differentialgleichung zu tun haben, deren Stammfunktion die Gestalt $F(x) - G(y)$ besitzt.

Wir demonstrieren das soeben erläuterte Verfahren an dem folgenden

Beispiel 6. Die Differentialgleichung $y' = y^2$ ist vom Typ der getrennten Variablen mit $I = J = \mathbb{R}$. Da die rechte Seite stetig differenzierbar nach y ist, gilt sogar der Eindeutigkeitssatz, was wir hier schon voraussetzen wollen. Also gibt es zu $(0, c) \in \mathbb{R}^2$ eine eindeutig bestimmte Lösung φ_c in einer Umgebung von 0 mit $\varphi_c(0) = c$. Für $c = 0$ ist die identisch verschwindende Funktion eine Lösung und damit $\varphi_0 = 0$. Keine andere Lösung kann jemals 0 werden. Ist also $c > 0$, so können wir uns von vornherein auf das Gebiet $\mathbb{R} \times \mathbb{R}_+^*$ beschränken. Mit den obigen Bezeichnungen ist dann

$$F(x) = x \quad \text{und} \quad G(y) = \frac{1}{c} - \frac{1}{y} .$$

Es gilt $G(\mathbb{R}_+^*) = (-\infty, 1/c) =: I'$, so daß genau dieses Intervall *maximal* ist mit $F(I') \subset G(\mathbb{R}_+^*)$. Für die Lösung φ der Differentialgleichung gilt auf diesem Intervall nach dem vorstehenden Satz

$$\frac{1}{c} - \frac{1}{\varphi(x)} = x$$

und damit

$$\varphi(x) = \frac{c}{1 - cx}, \quad x < \frac{1}{c}.$$

Entsprechend findet man für $c < 0$ die Lösungen

$$\varphi(x) = \frac{c}{1 - cx}, \quad x > \frac{1}{c}.$$

Man sieht also an diesem Beispiel, daß die Lösungen nicht auf dem vollen Intervall $I = \mathbb{R}$ zu existieren brauchen, obwohl der Definitionsbereich der rechten Seite, also $f(x, y) = y^2$ ganz $\mathbb{R} \times \mathbb{R}$ ist.

Bevor wir uns den drängenden *theoretischen* Fragen zuwenden, behandeln wir einige weitere einfache Klassen von Differentialgleichungen und elementare Lösungsmethoden in Beispielen.

Beispiel 7. Die (*inhomogene*) *lineare Differentialgleichung* ist von der Form

$$y' = A(x)y + B(x) \quad \text{mit stetigen Funktionen } A, B: I \rightarrow \mathbb{R}.$$

Die homogene Gleichung ($B = 0$) ist vom Typ der getrennten Variablen und läßt sich sofort lösen. Wählt man einen Punkt $a \in I$ fest, so werden die Lösungen gegeben durch

$$\varphi(x) = c \exp\left(\int_a^x A(t) dt\right), \quad c = \varphi(a).$$

Insbesondere bildet die *Gesamtheit* der Lösungen einen eindimensionalen Vektorraum mit einer *Basislösung* φ_0 , die nirgends auf dem Definitionsintervall I verschwindet. Den *inhomogenen* Fall führt man auf den homogenen zurück durch die sogenannte *Methode der variablen Koeffizienten*. Man braucht wie in der Theorie der Systeme von linearen Gleichungen wegen der Linearität des Problems nur eine einzige spezielle Lösung ψ des inhomogenen Problems zu finden (alle anderen Lösungen sind dann Summe von ψ mit der allgemeinen Lösung $c\varphi_0$, $c \in \mathbb{R}$, des homogenen Problems), und diese gewinnt man mit dem Ansatz

$$\psi(x) := c(x)\varphi_0(x).$$

Wegen

$$c'\varphi_0 + cA\varphi_0 = c'\varphi_0 + c\varphi_0' = \psi' = A\psi + B = cA\varphi_0 + B$$

ist ψ genau dann eine Lösung, wenn

$$c'(x) = \frac{B(x)}{\varphi_0(x)}.$$

Hieraus gewinnt man die gesuchte Funktion $c(x)$ durch Integration. - Wir behandeln hierzu das konkrete

Beispiel 8. $y' = 2xy + x^3$. Die homogene Gleichung $y' = 2xy$ hat die fundamentale Lösung

$$\varphi_0(x) = e^{x^2},$$

und mit den vorstehenden Überlegungen findet man die folgende allgemeine Lösung (wir betrachten z. B. die Stelle $a = 0$):

$$\psi(x) = e^{x^2} \left(c + \int_0^x t^3 e^{-t^2} dt \right).$$

Substituiert man in dem Integral noch $s = t^2$, so erhält man

$$\int_0^x t^3 e^{-t^2} dt = \frac{1}{2} \int_0^{x^2} s e^{-s} ds = \frac{1}{2} \left(1 - e^{-x^2} (1 + x^2) \right)$$

und damit die einfachere Darstellung

$$\psi(x) = \left(c + \frac{1}{2}\right) e^{x^2} - \frac{1}{2} (x^2 + 1) .$$

Bemerkung. Die inhomogene lineare Differentialgleichung ist i. A. nicht exakt, besitzt aber stets einen Eulerschen Multiplikator. Schreiben wir die Differentialgleichung in der Form $dy - (Ay + B)dx$, so muß ein Eulerscher Multiplikator λ der Bedingung $\lambda_x + [\lambda(Ay + B)]_y = 0$ genügen, also $\lambda_x + \lambda_y(Ay + B) + A\lambda = 0$. Man kann versuchen, hierfür eine Lösung zu finden, die nicht von y abhängt, was offensichtlich möglich ist:

$$\lambda(x, y) := \exp(-\alpha(x)) ,$$

wobei α eine Stammfunktion von A bezeichnet. Eine Stammfunktion der mit λ multiplizierten Gleichung ist dann notwendig von der Form $F(x, y) = y e^{-\alpha(x)} - \beta(x)$, wobei die Funktion f leicht durch Integration zu bestimmen ist, da $\beta' = \lambda B = B e^{-\alpha}$ sein muß. Für die Lösungen der Differentialgleichung, also die Niveau-Linien von F , findet man so Gleichungen, die man unmittelbar nach y auflösen kann:

$$y = e^{\alpha(x)} (C + \beta(x)) .$$

Beispiel 9. Unter einer *homogenen* Differentialgleichung³⁷ versteht man eine solche vom Typ

$$y' = f\left(\frac{y}{x}\right) .$$

Hierbei ist $f : J \rightarrow \mathbb{R}$ stetig auf einem (offenen) Intervall J und $G := \{(x, y) \in \mathbb{R}^2 : x \neq 0, \frac{y}{x} \in J\}$. Differentialgleichungen dieses Typs lassen sich auf solche mit getrennten Variablen durch den Ansatz $z := y/x$ zurückführen. Aus der Produktregel folgt nämlich $y' = xz' + z$, woraus man sofort schließt:

Satz 20.3 *Es sei $I \subset \mathbb{R}^*$ ein offenes Intervall, es seien $a \in I$, $(a, c) \in G$, und $\varphi : I \rightarrow \mathbb{R}$ sei eine Funktion. Genau dann ist φ Lösung der homogenen Differentialgleichung $y' = f(y/x)$ mit Anfangsbedingung $\varphi(a) = c$, wenn die durch*

$$\psi(x) = \frac{\varphi(x)}{x}$$

definierte Funktion $\psi : I \rightarrow \mathbb{R}$ Lösung der Differentialgleichung

$$z' = \frac{f(z) - z}{x}$$

mit der Anfangsbedingung $\psi(a) = c/a$ ist.

Als konkretes *Beispiel 10* hierzu betrachten wir $y' = 1 + \left(\frac{y}{x}\right) + \left(\frac{y}{x}\right)^2$. Hierbei ist also $f(z) = 1 + z + z^2$; wir setzen ferner $J := \mathbb{R}_+^*$. Der vorstehende Satz führt das Problem zurück auf die Lösung der Differentialgleichung mit getrennten Variablen:

$$z' = \frac{1 + z^2}{x} .$$

Dies führt auf die Stammfunktionen

$$F(x) = \int_a^x \frac{dt}{t} = \ln \frac{x}{a} \quad \text{und} \quad G(z) = \int_{c/a}^z \frac{dt}{1+t^2} = \arctan z - \arctan \frac{c}{a}$$

und zu den Lösungen

$$z = \psi(x) = \tan \left(\ln \frac{x}{a} + \arctan \frac{c}{a} \right) .$$

³⁷Dieser Begriff darf nicht verwechselt werden mit dem der homogenen linearen Differentialgleichung.

Die ursprüngliche Gleichung besitzt daher die Lösungen

$$y = x \psi(x) = x \tan \left(\ln \frac{x}{a} + \arctan \frac{c}{a} \right).$$

Bemerkung und Beispiel 11. Man kann unter der Voraussetzung des vorigen Satzes natürlich nicht erwarten, daß man die notwendige Auflösung der Gleichung nach einer der Variablen explizit vornehmen kann. Wir betrachten dazu die Differentialgleichung

$$y' = \frac{x + y}{x + 2y}$$

auf $G := \mathbb{R}^2 \setminus \{(x, y) \in \mathbb{R}^2 : x + 2y = 0\}$. Nach dem allgemeinen Existenz- und Eindeigkeitssatz besitzt diese Differentialgleichung lokal eindeutig bestimmte Lösungen durch jeden Punkt $(a, c) \in G$. Sie läßt sich mit dem üblichen Ansatz $z = y/x$ für *homogene Differentialgleichungen* leicht lösen. Dieser ergibt die Gleichung

$$xz' = \frac{1+z}{1+2z} - z = \frac{1-2z^2}{1+2z}, \quad \text{und wegen} \quad \frac{1+2z}{1-2z^2} = \frac{A}{1-\sqrt{2}z} - \frac{B}{1+\sqrt{2}z}$$

mit

$$A = \frac{\sqrt{2}+1}{2}, \quad B = \frac{\sqrt{2}-1}{2}$$

liefert dies unmittelbar die Lösungen in implizierter Form:

$$\ln|x| + \frac{A}{\sqrt{2}} \ln|1-\sqrt{2}z| + \frac{B}{\sqrt{2}} \ln|1+\sqrt{2}z| = C.$$

Nun ist $A+B = \sqrt{2}$, so daß man $\ln|x|$ auf die beiden anderen Summanden „aufteilen“ kann. Dadurch kann man z wieder durch y ersetzen und erhält nach Multiplikation mit $2\sqrt{2}$ und mit einer neuen Konstanten C :

$$(\sqrt{2}+1) \ln|x-\sqrt{2}y| + (\sqrt{2}-1) \ln|x+\sqrt{2}y| = C.$$

Durch „Exponieren“ gewinnt man daraus schließlich die einfache implizite Gestalt

$$|x-\sqrt{2}y|^{\sqrt{2}+1} \cdot |x+\sqrt{2}y|^{\sqrt{2}-1} = C',$$

wobei man zunächst $C' > 0$ voraussetzen muß. Durch $C' = 0$ findet man als mögliche Lösungen noch die Geraden $x-\sqrt{2}y=0$ und $x+\sqrt{2}y=0$, die wir bei unseren bisherigen Überlegungen ausschließen mußten. Man rechnet aber schnell nach, daß sie tatsächlich die Differentialgleichung lösen.

Bemerkungen. 1. Die Lösungsgesamtheit sieht also so ähnlich aus wie die (verzerrte) Schar aller Hyperbeln zu zwei festen Asymptoten (mit Einschluß derselben). Die einzig wirklich auszuschließende Gerade $x+2y=0$ ist der Ort der Punkte, wo die obigen Lösungskurven Tangenten parallel zur y -Achse besitzen.

2. Man kann das gleiche Ergebnis auch recht schnell mit Hilfe eines integrierenden Faktors gewinnen. In der Tat ist die Differentialgleichung

$$(x+y)dx - (x+2y)dy = 0$$

nicht exakt. Man rechnet aber nach, daß die Gleichung

$$\frac{x+y}{x^2-2y^2}dx - \frac{x+2y}{x^2-2y^2}dy = 0$$

eine Stammfunktion (bis auf einen Faktor) vom Typ

$$F(x, y) = (\sqrt{2} + 1) \ln |x - \sqrt{2}y| + (\sqrt{2} - 1) \ln |x + \sqrt{2}y|$$

besitzt.

3. Man versteht die Form des integrierenden Faktors besser, wenn man die *lineare Transformation* $u = x + \sqrt{2}y$, $v = x - \sqrt{2}y$ vornimmt. Diese ist äquivalent zu

$$x = \frac{u + v}{2}, \quad y = \frac{u - v}{2\sqrt{2}}.$$

Die obige Differentialform transformiert sich damit zu

$$\frac{1}{2\sqrt{2}} \left((\sqrt{2} + 1) u dv + (\sqrt{2} - 1) v du \right),$$

und diese besitzt offenbar den Eulerschen Multiplikator $1/uv$, also $1/(x^2 - 2y^2)$.

Beispiel 12. Die *Methode der Variablentransformation* besteht darin, eine Differentialgleichung $y' = f(x, y)$ auf einem *Produktgebiet* $G = I \times J$ durch bijektive, umkehrbar stetig differenzierbare Abbildungen $g: I^* \rightarrow I$, $h: J^* \rightarrow J$ in eine Differentialgleichung auf $G^* := I^* \times J^*$ umzuformen. Durch $F(u, v) := (g(u), h(v))$ wird eine bijektive Abbildung $F: G^* \rightarrow G$ gestiftet, mit deren Hilfe man sofort den folgenden Satz beweisen kann.

Satz 20.4 Ist $\varphi: I \rightarrow J$ eine Lösung der Differentialgleichung $y' = f(x, y)$, so ist $\psi := h^{-1} \circ \varphi \circ g: I^* \rightarrow J^*$ eine Lösung der Differentialgleichung

$$v' = \frac{g'(u)}{h'(v)} f(g(u), h(v)),$$

und umgekehrt.

Die Nützlichkeit dieser Transformation erläutern wir an dem *Beispiel 13* der BERNOULLISCHEN *Differentialgleichung* $y' = A(x)y + B(x)y^\alpha$, wobei $A, B: I \rightarrow \mathbb{R}$ stetig sind, α eine beliebige reelle Konstante bezeichnet und $J := \mathbb{R}_+^*$ gesetzt sei. Wir können weiter $\alpha \neq 1$ annehmen, da es sich sonst um eine homogene lineare Differentialgleichung handelt. Man setzt nun $x = g(u) := u$ und $y = h(v) := v^{1/1-\alpha}$ für $u \in I^* := I$ bzw. $v \in J^* := J$; damit transformiert sich die Bernoullische Differentialgleichung in

$$v' = \frac{1 - \alpha}{v^{1/(1-\alpha)-1}} \left(A(x) v^{1/1-\alpha} + B(x) v^{\alpha/1-\alpha} \right),$$

also in die inhomogene lineare Differentialgleichung $v' = (1 - \alpha)(A(x)v + B(x))$.

In den obigen Beispielen der linearen Differentialgleichung oder z. B. bei $y' = y^2$ haben wir nicht nur lokale Existenz, sondern auch *Eindeutigkeit* von Lösungen vorgefunden. Wir wollen dieses Verhalten noch ein wenig genauer studieren.

Definition. Wir sagen, die Differentialgleichung $y' = f(x, y)$ genügt dem *Eindeutigkeitssatz*, wenn für zwei Lösungen $\varphi: I \rightarrow \mathbb{R}$ und $\tilde{\varphi}: \tilde{I} \rightarrow \mathbb{R}$ gilt: Ist $\varphi(a) = \tilde{\varphi}(a)$ für ein $a \in I \cap \tilde{I}$, so ist $\varphi = \tilde{\varphi}$ auf $I \cap \tilde{I}$.

Wie wir an einem Beispiel gesehen haben, reicht die *Stetigkeit* von f nicht aus, um den *Eindeutigkeitssatz* für die Differentialgleichung $y' = f(x, y)$ zu garantieren. Eine hinreichende Bedingung ist durch das (lokale) Erfülltsein einer Lipschitz-Bedingung gegeben. Eine schwächere hinreichende

Bedingung werden wir weiter unten noch kennenlernen.

Definition. Es sei $G \subset \mathbb{R}^2$ offen und $f : G \rightarrow \mathbb{R}$ eine Funktion. Man sagt, f genüge (bzgl. der zweiten Variablen y) einer *LIPSCHITZ-Bedingung* auf G , wenn es eine Konstante $L > 0$ gibt, so daß für alle $(x, y), (x, \tilde{y}) \in G$ die Abschätzung

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

besteht. Sie genügt *lokal* in G einer Lipschitz-Bedingung, wenn es zu jedem Punkt in G eine Umgebung $U \subset G$ gibt, so daß f auf U einer Lipschitz-Bedingung genügt.

Bemerkungen. Aus dem Mittelwertsatz folgt unmittelbar, daß eine Funktion f mit stetiger partieller Ableitung f_y lokal stets einer Lipschitz-Bedingung genügt.

Satz 20.5 *Erfüllt die stetige Funktion $f : G \rightarrow \mathbb{R}$ eine lokale LIPSCHITZ-Bedingung auf G , so besteht für die Differentialgleichung $y' = f(x, y)$ der Eindeigkeitsatz.*

Beweis. Ohne Einschränkung seien die beiden Lösungen φ und ψ auf dem *gleichen* Intervall I definiert. Nach Voraussetzung ist die Menge

$$A := \{a \in I : \varphi(a) = \psi(a)\} \subset I$$

nicht leer. Wir zeigen anschließend, daß die Menge A in I sowohl (relativ) abgeschlossen als auch offen ist. Dies reicht aus, um, wie erwünscht, auf $A = I$ zu schließen. Denn dann hat man eine *disjunkte* Zerlegung $I = (I \setminus A) \cup A$ von I in die (relativ) offenen Teilmengen A und $I \setminus A$, was aber aufgrund des Zusammenhangs des reellen Intervalls I und wegen $A \neq \emptyset$ nur dann möglich ist, wenn $I \setminus A$ leer, also $A = I$ ist. Die *Abgeschlossenheit* von A folgt aus der Stetigkeit der Funktionen φ, ψ : Ist nämlich $a_j \in A$ eine Folge, die gegen eine Zahl $a \in I$ konvergiert, so ist

$$\varphi(a) = \lim_j \varphi(a_j) = \lim_j \psi(a_j) = \psi(a),$$

d. h. $a \in A$.

Für die Offenheit von A benötigen wir die Lipschitz-Bedingung. Es sei dazu $a \in A$, $c := \varphi(a) = \psi(a)$ und $\varepsilon > 0$ und $r > 0$ seien so klein gewählt, daß f auf dem Rechteck $Q := \{(x, y) : |x - a| \leq \varepsilon, |y - c| \leq r\} \subset G$ der Lipschitz-Bedingung mit einer Konstanten L genügt. Wir können ferner ohne Beschränkung der Allgemeinheit annehmen, daß $\varepsilon L < 1$ ist. Wir setzen dann z. B. $J := \{x \in I : 0 \leq x - a \leq \varepsilon\}$. Ist ε hinreichend klein, so ist $J \subset I$ ein kompaktes Intervall und die Einschränkung der Graphen von φ und ψ auf J sind ganz in Q enthalten. Insbesondere existiert

$$S := \sup_{x \in J} |\varphi(x) - \psi(x)| \leq 2r.$$

Es genügt zu zeigen, daß $S = 0$; denn dann ist $\varphi = \psi$ auf J , und mit dem entsprechenden Argument für Punkte links von a erhält man $I \cap \{x \in \mathbb{R} : |x - a| < \varepsilon\} \subset A$.

Dies folgt aber unter den gemachten Voraussetzungen aus der Standard-Abschätzung für Integrale, wenn man die Differentialgleichung in eine *Integralgleichung* umschreibt. Aus dem Hauptsatz der Differential- und Integralrechnung gewinnt man unmittelbar die Einsicht, daß die Bedingungen $\varphi'(x) = f(x, \varphi(x))$, $x \in I$, und $\varphi(a) = c$ äquivalent sind zu

$$(*) \quad \varphi(x) = c + \int_a^x f(t, \varphi(t)) dt.$$

Für $x \in J$ ist dann

$$|\varphi(x) - \psi(x)| = \left| \int_a^x (f(t, \varphi(t)) - f(t, \psi(t))) dt \right| \leq \varepsilon L S.$$

Durch Übergang zum Supremum auf der linken Seite folgt hieraus $0 \leq S \leq \varepsilon L S$, was aufgrund von $\varepsilon L < 1$ aber nur für $S = 0$ möglich ist. \square

Man kann die lokale LIPSCHITZ-Bedingung noch wesentlich abschwächen, ohne die (lokale) Eindeutigkeitsaussage zu verlieren. Es gilt z. B. der folgende

Satz 20.6 *Ist die Funktion $f : I_\varepsilon(a) \times I_r(c) \rightarrow \mathbb{R}$ stetig und genügt sie der Bedingung*

$$|x - a| \cdot |f(x, y_1) - f(x, y_2)| \leq |y_1 - y_2| \quad \text{für alle } (x, y_1), (x, y_2) \in I_\varepsilon(a) \times I_r(c),$$

so besitzt die Differentialgleichung $y' = f(x, y)$ höchstens eine Lösung φ lokal um a mit $\varphi(a) = c$.

Beweis. Es seien φ und ψ zwei Lösungen auf einem kompakten Intervall I , die an der Stelle $a \in I$ übereinstimmen. Wir setzen

$$\lambda(x) := \frac{\varphi(x) - \psi(x)}{x - a}, \quad x \neq a,$$

und zeigen als erstes, daß diese Funktion durch $\lambda(a) := 0$ stetig nach a fortgesetzt werden kann. Dieses folgt unmittelbar wegen

$$\lambda(x) := \frac{\varphi(x) - \varphi(a)}{x - a} - \frac{\psi(x) - \psi(a)}{x - a} \longrightarrow \varphi'(a) - \psi'(a) = f(a, \varphi(a)) - f(a, \psi(a)) = 0.$$

Es sei ferner $M := \sup_{x \in I} |\lambda(x)|$. Es ist dann zu zeigen, daß $M = 0$. Nehmen wir also $M > 0$ an, und es sei $\gamma \in I$ ein Punkt, so daß $|\lambda(\gamma)| = M$. Insbesondere ist dann $\gamma \neq a$ und

$$\left| \int_a^\gamma \left| \frac{\varphi(t) - \psi(t)}{t - a} \right| dt \right| < M |\gamma - a|$$

wegen $\lambda(a) = 0$. Dies führt aufgrund unserer Voraussetzung zu einem Widerspruch. Es müßte dann nämlich gelten:

$$\begin{aligned} M |\gamma - a| &= |\lambda(\gamma)| |\gamma - a| = |\varphi(\gamma) - \psi(\gamma)| = \left| \int_a^\gamma (f(t, \varphi(t)) - f(t, \psi(t))) dt \right| \\ &\leq \left| \int_a^\gamma |f(t, \varphi(t)) - f(t, \psi(t))| dt \right| \leq \left| \int_a^\gamma \left| \frac{\varphi(t) - \psi(t)}{t - a} \right| dt \right| \\ &< M |\gamma - a|. \end{aligned}$$

Widerspruch! \square

Bemerkung. Der obige Beweis stammt von OSKAR PERRON, *Eine hinreichende Bedingung für die Unität der Lösung von Differentialgleichungen erster Ordnung*, Math. Zeitschrift **28** (1928), pp. 216–219. Als erster hat aber MITIO NAGUMO diesen Satz (in etwas schwächerer Form) formuliert und bewiesen in einer Arbeit mit exakt dem gleichen Titel, erschienen in Japanese Journal of Mathematics **3**, (1926), 107–112. Perron hat in loc. cit. des weiteren durch ein Beispiel begründet, daß man auf der rechten Seite der vorausgesetzten Ungleichung keinen Faktor > 1 hinzufügen darf. Er betrachtet zu positivem ε die Funktion $f(x, y)$ mit

$$f(x, y) := \begin{cases} (1 + \varepsilon) \frac{y}{x}, & 0 < y < x^{1+\varepsilon}, \\ (1 + \varepsilon) x^\varepsilon, & y \geq x^{1+\varepsilon}, \\ 0, & y \leq 0. \end{cases}$$

Sie ist stetig auf $\mathbb{R}_+ \times \mathbb{R}$ und erfüllt die Bedingung

$$x |f(x, y_1) - f(x, y_2)| \leq (1 + \varepsilon) |y_1 - y_2|.$$

Dennoch gibt es unendlich viele im Nullpunkt verschwindende Lösungen, nämlich $y = Cx^{1+\varepsilon}$, $0 \leq C \leq 1$.

Die Formel (*) ist auch der Ausgangspunkt für ein explizites *Iterationsverfahren*, mit dem man gewöhnliche Differentialgleichungen lokal lösen kann, sofern sie (lokal) einer Lipschitz-Bedingung genügen. Wir beginnen mit einem einfachen Lemma, das den „Operator“-Gedanken schon in den Vordergrund stellt.

Lemma 20.7 Zu $(a, c) \in G$ seien $\varepsilon > 0$ und $r > 0$ so gewählt, daß $Q := \{(x, y) \in \mathbb{R}^2 : |x - a| \leq \varepsilon, |y - c| \leq r\} \subset G$. Die stetige Funktion $f : G \rightarrow \mathbb{R}$ sei auf Q durch die Konstante M dem Betrage nach nach oben beschränkt. Wird dann ε so klein gewählt, daß $\varepsilon \leq r/M$, so ordnet der Operator („Picard-Operator“)

$$\varphi \longmapsto \psi := \mathcal{P}(\varphi) \quad \text{mit} \quad \psi(x) := c + \int_a^x f(t, \varphi(t)) dt$$

jeder stetigen Funktion $\varphi : I := [a - \varepsilon, a + \varepsilon] \rightarrow [c - r, c + r] =: J$ mit $\varphi(a) = c$ eine stetige Funktion $\psi : I \rightarrow J$ mit $\psi(a) = c$ zu.

Beweis. Nach allgemeinen Sätzen der Integrationstheorie ist die Funktion ψ wohldefiniert und stetig, und sie erfüllt die Anfangsbedingung $\psi(a) = c$. Die verbleibende Behauptung ist erneut eine Konsequenz der Standardabschätzung für Integrale:

$$|\psi(x) - c| = \left| \int_a^x f(t, \varphi(t)) dt \right| \leq \varepsilon M \leq r. \quad \square$$

Satz 20.8 (Iterationsverfahren von Picard und Lindelöf) Unter den Voraussetzungen des vorigen Lemmas seien zu $(a, c) \in G$ die Konstanten $\varepsilon > 0$ und $r > 0$ zusätzlich so gewählt, daß die Funktion $f : G \rightarrow \mathbb{R}$ auf $Q = I \times J$ eine Lipschitz-Bedingung mit der Lipschitz-Konstanten $L > 0$ erfüllt. Ist dann $\varphi_0 \equiv c$ und definiert man induktiv

$$\varphi_{k+1} := \mathcal{P}(\varphi_k),$$

so konvergiert die Folge (φ_k) auf I gleichmäßig gegen eine Grenzfunktion φ , die die Differentialgleichung $y' = f(x, y)$ mit Anfangsbedingung $\varphi(a) = c$ löst.

Folgerung 20.9 Erfüllt $f : G \rightarrow \mathbb{R}$ eine lokale Lipschitz-Bedingung, so besitzt die Differentialgleichung $y' = f(x, y)$ lokal (eindeutig bestimmte) Lösungen.

Beweis des Satzes. Es genügt zu zeigen, daß für alle $x \in I$ die Ungleichungen

$$(+) \quad |\varphi_{k+1}(x) - \varphi_k(x)| \leq M L^k \frac{|x - a|^{k+1}}{(k + 1)!}$$

bestehen. Denn dann besitzt die Teleskopreihe $\sum_{k=0}^{\infty} (\varphi_{k+1} - \varphi_k)$ auf dem Intervall I die konvergente Majorante $(M/L) \sum_{k=0}^{\infty} (\varepsilon L)^{k+1} / (k + 1)!$, so daß

$$\varphi := \lim \varphi_k = c + \sum_{k=0}^{\infty} (\varphi_{k+1} - \varphi_k)$$

als gleichmäßiger Limes stetiger Funktionen stetig auf I ist. Wegen

$$|f(x, \varphi(x)) - f(x, \varphi_k(x))| \leq L |\varphi(x) - \varphi_k(x)|$$

konvergiert auch die Folge $f(x, \varphi_k(x))$ auf I gleichmäßig gegen $f(x, \varphi(x))$. Damit ist die Vertauschung des Grenzprozesses in der folgenden Zeile gerechtfertigt:

$$\varphi(x) - c = \lim_k (\varphi_{k+1}(x) - c) = \lim_k \int_a^x f(t, \varphi_k(t)) dt = \int_a^x \lim_k f(t, \varphi_k(t)) dt = \int_a^x f(t, \varphi(t)) dt.$$

Es bleibt also noch der Nachweis von (+). Für $k = 0$ ist dies einfach die Standard-Abschätzung für Integrale. Ist die Aussage für ein $k - 1 \geq 0$ schon bewiesen, so folgt genauso einfach

$$\begin{aligned} |\varphi_{k+1}(x) - \varphi_k(x)| &= \left| \int_a^x (f(t, \varphi_k(t)) - f(t, \varphi_{k-1}(t))) dt \right| \\ &\leq \left| \int_a^x L |\varphi_k(t) - \varphi_{k-1}(t)| dt \right| \leq \frac{ML^k}{k!} \left| \int_a^x |t - a|^k dt \right| = ML^k \frac{|x - a|^{k+1}}{(k+1)!}, \end{aligned}$$

womit der Beweis vollendet ist. \square

Bemerkung. Der sogenannte Satz von Picard und Lindelöf stammt eigentlich schon von LIPSCHITZ 1876 und wurde sogar noch weitere 50 Jahre früher von CAUCHY unter der Voraussetzung bewiesen, daß die Funktion f stetige partielle Ableitungen besitzt. Die Beiträge von PICARD 1893 und später von LINDELÖF beschränken sich im Wesentlichen auf die Bestimmung eines möglichst großen Definitivintervalls I der Lösung.

Wir betrachten zu dem Picard-Lindelöf-Verfahren noch ein Beispiel, dessen Lösungen wir allerdings schon kennen.

Beispiel. Wir wollen das Anfangswert-Problem $y' = 2xy$, $\varphi(0) = c$, lösen. Da die Funktion $f(x, y) = 2xy$ stetig partiell nach y differenzierbar ist, erfüllt sie lokal eine Lipschitz-Bedingung. Die gesuchte Lösung läßt sich dann (in einer a priori kleinen geeigneten Umgebung des Nullpunktes) iterativ aufbauen: Es ist

$$\begin{aligned} \varphi_0(x) &= c, \quad \varphi_1(x) = c + c \int_0^x 2t dt = c(1 + x^2), \\ \varphi_2(x) &= c + 2c \int_0^x (t + t^3) dt = c \left(1 + x^2 + \frac{x^4}{2} \right). \end{aligned}$$

Durch vollständige Induktion gewinnt man offensichtlich

$$\varphi_k(x) = c \left(1 + \frac{x^2}{1!} + \dots + \frac{x^{2k}}{k!} \right)$$

und damit

$$\varphi(x) = \lim_{k \rightarrow \infty} \varphi_k(x) = ce^{x^2}.$$

Die Lösung ist damit sogar auf ganz \mathbb{R} definiert.

Bemerkung. In dem vorigen Beispiel sind die Lösungen sogar *reell-analytisch*. Dies ist z. B. immer der Fall, wenn wir lineare Differentialgleichungen mit reell-analytischen (also insbesondere konstanten) „Koeffizienten“ behandeln wollen. In einer solchen Situation kann man einen *Potenzreihen-Ansatz* machen und versuchen, die Lösung durch *Koeffizienten-Vergleich* zu ermitteln. In dem vorigen Beispiel führt dies mit dem Ansatz

$$\varphi(x) := \sum_{k=0}^{\infty} a_k x^k$$

zu der Gleichung

$$\sum_{k=0}^{\infty} (k+1) a_{k+1} x^k = \varphi'(x) = 2x\varphi(x) := \sum_{k=1}^{\infty} 2a_{k-1} x^k,$$

woraus durch Koeffizienten-Vergleich $a_1 = a_3 = a_5 = \dots = 0$ und $2a_2 = 2a_0$, $4a_4 = 2a_2$, $6a_6 = 2a_4$ etc. folgt. Mit $c := a_0$ ergibt sich daraus unmittelbar $a_2 = c$, $a_4 = c/2$, $a_6 = c/6$ und per vollständiger Induktion $a_{2\ell} = c/\ell!$.

Wir wollen uns zum Abschluß dieses Kapitels noch ein wenig mit der sogenannten „allgemeinen“ Lösung einer Differentialgleichung mit lokaler Lipschitz-Bedingung beschäftigen. Wir betrachten zu fest vorgegebenem Punkt $(a, c) \in G$ die Gesamtheit $\mathcal{I}_{(a,c)}$ aller Intervalle $I \subset \mathbb{R}$ mit $a \in I$, zu denen es eine Lösung $\varphi: I \rightarrow \mathbb{R}$ der Differentialgleichung $y' = f(x, y)$ mit $\varphi(a) = c$ gibt. Nach dem (lokalen) Existenzsatz ist $\mathcal{I}_{(a,c)} \neq \emptyset$, und

$$I_{(a,c)} := \bigcup_{I \in \mathcal{I}_{(a,c)}} I$$

ist nach unseren früheren Sätzen über die Charakterisierung von reellen Intervallen tatsächlich wieder ein Intervall. Wegen des Eindeutigkeitsatzes fügen sich alle Lösungen auf den Intervallen $I \in \mathcal{I}_{(a,c)}$ zu einer Lösung $\varphi_{(a,c)}$ auf $I_{(a,c)}$ zusammen. Wir nennen sie die *maximale* Lösung mit Anfangs-Bedingung (a, c) , da jede andere solche Lösung eine Einschränkung von dieser auf ein evtl. kleineres Intervall I mit $a \in I$ ist.

Satz 20.10 *Das Definitionsintervall $I_{(a,c)}$ der maximalen Lösung $\varphi_{(a,c)}$ ist offen. Die Lösung verläßt „nach beiden Seiten“ jedes Kompaktum $K \subset G$.*

Bemerkung und Definition. Man umschreibt den Sachverhalt des vorigen Satzes mit der Phrase, daß die maximale Lösung „von Rand zu Rand“ in G verläuft. Wir schreiben im folgenden auch $I_{(a,c)} := (\alpha_-(a, c), \alpha_+(a, c))$ und nennen die Elemente $\alpha_+(a, c) \in \mathbb{R} \cup \{\infty\}$ und $\alpha_-(a, c) \in \mathbb{R} \cup \{-\infty\}$ (bei Interpretation der Variablen x als *Zeit*) die *positive* bzw. *negative Entweichzeit*.

Beweis des Satzes. Es sei $\alpha_+ := \sup I_{(a,c)}$. Wir zeigen dann als erstes, daß der Graph von $\varphi_{(a,c)}$ auf dem Intervall $[a, \alpha_+)$ in keinem Kompaktum $K \subset G$ enthalten sein kann, womit die zweite Aussage („nach rechts“) bewiesen ist. Wäre dies doch für ein Kompaktum K der Fall, so besäße der Graph einen Häufungspunkt der Form $(\alpha_+, \gamma) \in K \subset G$. Es werde nun das kompakte Rechteck $Q = I \times J \subset G$ mit Mittelpunkt (α_+, γ) so gewählt, daß dort $\varepsilon \leq r/M$ gilt (siehe Lemma 7) und für f eine Lipschitz-Bedingung mit Lipschitz-Konstanten L erfüllt ist. Es sei nun (a', γ') ein Punkt des Graphen mit $|a' - \alpha_+| < \varepsilon/2$, $|\gamma' - \gamma| < r/2$. Dann ist der Quader Q' mit den „halbierten“ Kantenlängen ε bzw. r und Mittelpunkt (a', γ') in Q enthalten, und wegen $\varepsilon/2 \leq r/2M$ gibt es eine Lösung der Differentialgleichung mit Anfangsbedingung (a', γ') , die auf dem Intervall $[a' - \varepsilon/2, a' + \varepsilon/2]$ definiert ist. Nun ist aber $a' + \varepsilon/2 > \alpha_+$, so daß im Widerspruch zur Voraussetzung die Lösung $\varphi_{(a,c)}$ doch nach rechts über α_+ fortgesetzt werden kann.

Hieraus folgt sofort auch der erste Teil. Wäre nämlich $\alpha_+ \in I_{(a,c)}$, so wäre $\varphi_{(a,c)}$ auf dem kompakten Intervall $[a, \alpha_+]$ erklärt und damit der Graph von $\varphi_{(a,c)}$ auf diesem Intervall eine kompakte Teilmenge in G . \square

Wir können nunmehr die „allgemeine Lösung“ $\varphi_{(a,c)}$ als eine Funktion auf der Menge

$$\tilde{G} := \bigcup_{(a,c) \in G} I_{(a,c)} \times \{(a, c)\} \subset \mathbb{R} \times G \subset \mathbb{R}^3$$

auffassen. Der Deutlichkeit halber schreiben wir (ξ, η) anstelle von (a, c) und

$$\Phi(x; \xi, \eta) := \varphi_{(\xi, \eta)}(x), \quad (x, \xi, \eta) \in \tilde{G}.$$

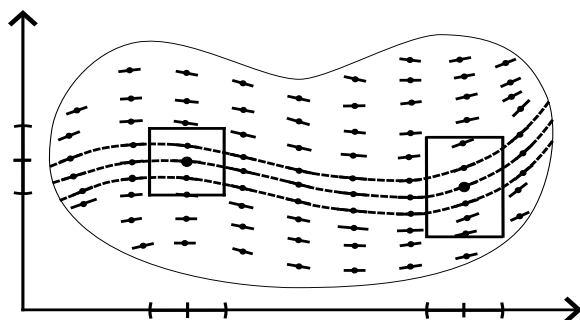
Diese Funktion hat offensichtlich die folgenden Eigenschaften:

- o) $\Phi(\xi; \xi, \eta) := \eta$,
- i) $\Phi(\xi'; \xi, \eta) := \eta' \implies \Phi(\xi; \xi', \eta') := \eta$,
- ii) $\frac{\partial}{\partial x} \Phi(x; \xi, \eta) = f(x, \Phi(x; \xi, \eta))$.

Überraschenderweise gilt nun der folgende tiefliegende Satz, den wir erst im Anhang zu Kapitel 23 beweisen können, da wir dabei Methoden benutzen werden, die uns jetzt noch nicht zur Verfügung stehen.

Satz 20.11 Die allgemeine Lösung $\Phi(x; \xi, \eta)$ einer Differentialgleichung $y' = f(x, y)$ mit lokaler Lipschitz-Bedingung ist eine stetige Funktion auf der offenen Menge \tilde{G} .

Wir sollten diesen Satz aber zumindest *interpretieren*. Er besagt, daß es zu jedem $(\xi_0, \eta_0) \in G$, jedem x_0 im Definitionsbereich der maximalen Lösung durch (ξ_0, η_0) und jedem $\varepsilon > 0$ ein $\delta > 0$ gibt, so daß jede Lösung mit Anfangsbedingung $|\xi - \xi_0| < \delta$, $|\eta - \eta_0| < \delta$ noch ganz über dem Intervall $(x_0 - \delta, x_0 + \delta)$ definiert ist und dort nur Werte in dem Intervall $(y_0 - \varepsilon, y_0 + \varepsilon)$ annimmt, wobei $y_0 = \Phi(x_0; \xi_0, \eta_0)$.



Figur 20.6

Bemerkung. Zusammen mit ii) impliziert dieser Satz, daß die allgemeine Lösung $\Phi(x; \xi, \eta)$ auf \tilde{G} sogar stetig partiell nach der Variablen x differenzierbar ist. Setzt man zusätzlich voraus, daß die Funktion f sogar stetig partiell nach y differenzierbar ist, so ist die allgemeine Lösung auch stetig partiell nach den Variablen ξ und η differenzierbar.

Wir geben noch ein *Beispiel* für eine allgemeine Lösung. Betrachten wir dazu die inhomogene lineare Differentialgleichung

$$y' + \frac{2x}{1+x^2} y = \frac{2x^2}{1+x^2}.$$

Die Funktion $\varphi_0(x) := (1+x^2)^{-1}$ ist eine erzeugende Lösung der entsprechenden homogenen linearen Differentialgleichung. Der Ansatz $c\varphi_0$ der *variablen Koeffizienten* liefert die Gleichung $c'(x) = 2x^2$, also $c(x) = (2/3)x^3$. Die Gesamtheit der Lösungen kann man somit schreiben in der Form

$$\varphi(x) = \frac{2}{3} \frac{c + x^3}{1 + x^2}.$$

Hieraus läßt sich c leicht so bestimmen, daß der Wert η an der Stelle ξ angenommen wird. Durch Einsetzen gewinnt man dann die auf ganz \mathbb{R}^3 definierte allgemeine Lösung

$$\Phi(x; \xi, \eta) = \frac{1 + \xi^2}{1 + x^2} \left(\eta + \frac{2}{3} \frac{x^3 - \xi^3}{1 + \xi^2} \right),$$

und diese Funktion ist stetig partiell nach allen drei Variablen x, ξ, η differenzierbar (also auch total differenzierbar und (damit) stetig).

Bemerkung. Bei Vorhandensein einer globalen Lipschitz-Bedingung hat man sogar eine quantitative Abschätzung für das Abweichen zweier Lösungen voneinander, wenn ihre Anfangsbedingungen nahe

beieinander liegen. Allerdings ist diese Beziehung exponentiell bzgl. der Variablen x , so daß die Größe δ in dem vorigen Satz bei vorgegebenem ε und wachsendem x sehr schnell gegen Null geht. Damit kann man schon nach verhältnismäßig kurzer „Zeitdifferenz“ $x - x_0$ bei vorgegebener Approximationsgüte ε zu Werten von δ gelangen, die unterhalb jeder Meßgenauigkeit oder der vorhandenen Präzision der Information liegen. Dies mag erklären, warum langfristige *Wettervorhersage* ein so schwieriges Geschäft ist. Solche Überlegungen schließen den sogenannten *Schmetterlingseffekt* zumindest nicht aus, wenn dieser auch meiner Meinung nach hierdurch nicht logisch begründet werden kann. Im Übrigen ist deterministisch-chaotisches Verhalten, wie es sich z. B. im Modell des *LORENZ-Attraktors* manifestiert, ein wesentlich komplizierteres und erkenntnis-theoretisch bedeutsameres Phänomen.

Bemerkung. Die Kenntnis der Lösung $\varphi_{(a,c)}$ liefert uns, wenn wir von der soeben dargestellten Problematik absehen, nicht nur die prinzipielle Möglichkeit, das Verhalten des vorgelegten Systems mathematisch exakt für alle Zukunft zu bestimmen, sondern auch seine Entwicklung vor dem Zeitpunkt a . Dieses Vermögen zur *Rekonstruktion der Vergangenheit* geht allerdings auch verloren, wenn der lokale Eindeutigkeitssatz verletzt ist. Dies kann man sehr anschaulich an einem System demonstrieren, dessen Lösungen ähnlich wie in Figur 5 aussehen, aber an der y -Achse gespiegelt sind. Qualitativ erhält man ein solches Lösungssystem für die Differentialgleichung $y' = -\sqrt{y}$, wenn man sich auf Werte $y \geq 0$ beschränkt. Diese Differentialgleichung beschreibt zum Beispiel (bei geeigneter Wahl der Konstanten) die Wasserstandshöhe $y \geq 0$ in einem zylindrischen Becher, wenn das Wasser durch ein (kleines) kreisförmiges Loch im Boden abfließen kann (siehe z. B. [165], p. 272). Trifft man das System im Zustand $y = 0$ an, so hilft einem nur noch detektivische Kleinarbeit, aber kein noch so grandioses mathematisches Wissen, wenn man herausfinden möchte, zu welchem exakten Zeitpunkt sich der Behälter entleert hatte.

Anhang: Der Existenzsatz von Peano

Wir wollen in diesem Anhang den Beweis für den Peanoschen Existenzsatz (Satz 20.1) nachtragen. Wir können uns dabei auf eines der beiden von a begrenzten Halb-Intervalle beschränken, also z. B. das rechte. Es seien also im folgenden $b > a$ und $r > 0$ in Abhängigkeit von $(a, c) \in G$ so klein gewählt, daß $Q := \{(x, y) \in \mathbb{R}^2 : a \leq x \leq b, |y - c| \leq r\} \subset G$ und f dem Betrage nach auf Q nach oben durch eine Schranke M begrenzt ist, so daß $M(b - a) \leq r$ gilt.

In den beiden Versionen des Beweises, die wir hier präsentieren, konstruiert man gleichmäßig konvergente Folgen von „näherungsweise“ Lösungen, von deren Grenzfunktion man anschließend nachweisen muß, daß sie die vorgegebene Differentialgleichung löst. Hinter diesem Schluß steckt ein allgemeines Prinzip, das wir vorweg formulieren wollen. Dazu führen wir noch Symbole ein für Räume von Funktionen, mit denen wir in diesem Anhang ständig arbeiten werden: Wir setzen $I := [a, b]$ und

$$\mathcal{F} := \{\varphi \in \mathcal{C}^0(I, \mathbb{R}) : \varphi(a) = c \text{ und } |\varphi(x) - c| \leq r, x \in I\}.$$

Dies ist natürlich nichts anderes als eine (abgeschlossene) Teilmenge der abgeschlossenen Kugel um die Funktion $\varphi_0 \equiv c$ in $\mathcal{C}^0(I, \mathbb{R})$, versehen mit der Supremumsnorm.

Das nachfolgende Lemma wird meistens nur für den Fall $\psi_n = \varphi_n$ formuliert. In der präsentierten allgemeineren Fassung läßt es sich aber mit Gewinn zu einem äußerst eleganten Schluß verwenden.

Lemma 20.12 *Es seien $\varphi_k, \psi_k \in \mathcal{F}$ gleichmäßig konvergente Folgen mit der gleichen Grenzfunktion $\varphi := \lim \varphi_k = \lim \psi_k$, und $\varepsilon_k > 0$ sei eine Nullfolge, so daß*

$$\left| \varphi_k(x) - c - \int_a^x f(t, \psi_k(t)) dt \right| \leq \varepsilon_k, \quad x \in I.$$

Dann ist φ eine Lösung der Differentialgleichung $y' = f(x, y)$.

Beweis. Nach Voraussetzung existiert

$$\varphi(x) = \lim_{k \rightarrow \infty} \varphi_k(x) = c + \lim_{k \rightarrow \infty} \int_a^x f(t, \psi_k(t)) dt.$$

Da f auf dem kompakten Rechteck Q gleichmäßig stetig ist, konvergiert die Folge der Funktionen $f(x, \psi_k(x))$ gleichmäßig auf I : Zu vorgegebenem $\varepsilon > 0$ existiert ein $\delta > 0$ mit $|f(x, y_1) - f(x, y_2)| < \varepsilon$ für alle $(x, y_1), (x, y_2) \in Q$ mit $|y_1 - y_2| < \delta$, und zu δ gibt es ein $N \in \mathbb{N}$, so daß für alle $k, \ell \geq N$ und alle $x \in I$ gilt: $|\psi_k(x) - \psi_\ell(x)| < \delta$. Somit ist das Cauchy-Kriterium für gleichmäßige Konvergenz der Funktionenfolge $f(x, \psi_k(x))$ erfüllt, und wir dürfen in der vorstehenden Formel das Limeszeichen mit dem Integralzeichen vertauschen:

$$\lim_{k \rightarrow \infty} \int_a^x f(t, \psi_k(t)) dt = \int_a^x \lim_{k \rightarrow \infty} f(t, \psi_k(t)) dt = \int_a^x f(t, \lim_{k \rightarrow \infty} \psi_k(t)) dt = \int_a^x f(t, \varphi(t)) dt. \quad \square$$

Wir können solche „Näherungslösungen“ φ_k z. B. als Teilfolge der Folge von sogenannten EULERSchen Polygonzügen ψ_n gewinnen. Um deren Eigenschaften bequem herleiten zu können, geben wir noch die folgende

Definition. Eine stetige Funktion $\varphi : I = [a, b] \rightarrow \mathbb{R}$ gehöre zu \mathcal{F}' , wenn sie stückweise stetig differenzierbar ist, der Anfangsbedingung $\varphi(a) = c$ genügt, und für ihre (evtl. rechtsseitige und linksseitige) Ableitung die Abschätzung

$$|\varphi'(x)| \leq M, \quad M(b - a) \leq r,$$

erfüllt ist.

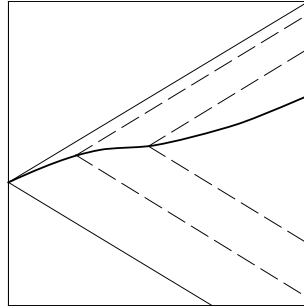
Als kleine Übungsaufgabe überlassen wir dem Leser den Nachweis des folgenden Lemmas.

Lemma 20.13 *Für alle $x_1 < x_2$ in I und alle $\varphi \in \mathcal{F}'$ gilt :*

$$(*) \quad |\varphi(x_2) - \varphi(x_1)| \leq M(x_2 - x_1).$$

Speziell für $x_1 := a$, $x_2 := x$ impliziert dies $\mathcal{F}' \subset \mathcal{F}$.

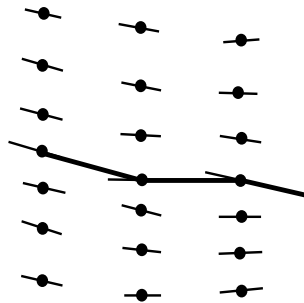
Die folgende Skizze dient zur Illustration des Lemmas:



Figur 20.7

Sind nun f, I, Q, M, r etc. wie oben gegeben, so definiert man den n -ten EULERSchen Polygonzug ψ_n auf $I = I_1 \cup I_2 \cup \dots \cup I_n$, $I_\nu := [a_{\nu-1} := a + (\nu-1)(b-a)/n, a + \nu(b-a)/n =: a_\nu]$ induktiv nach ν durch

- o) $\psi_n(a) := c$,
- i) $\psi_n(x) := \psi_n(a_{\nu-1}) + f(a_{\nu-1}, \psi_n(a_{\nu-1}))(x - a_{\nu-1})$, $x \in I_\nu$.



Figur 20.8

Man sieht aufgrund des vorstehenden Lemmas, daß die Funktionen ψ_n tatsächlich auf ganz I wohldefiniert und somit in \mathcal{F}' enthalten sind. Damit sind die ersten beiden Aussagen des folgenden Satzes schon bewiesen.

Satz 20.14 Die Polygonzüge ψ_n besitzen die folgenden Eigenschaften :

- i) Die ψ_n sind wohldefiniert, und ihr Graph liegt in Q ;
- ii) es ist $|\psi_n(x_1) - \psi_n(x_2)| \leq M|x_1 - x_2|$ für alle $x_1, x_2 \in I$;
- iii) zu jedem $\varepsilon > 0$ gibt es ein $N \in \mathbb{N}$, so daß für alle $n \geq N$ und alle $x \in I$ gilt :

$$|\psi'_n(x) - f(x, \psi_n(x))| \leq \varepsilon.$$

Beweis. Nur noch die Aussage iii) ist zu begründen. Aufgrund der Definition der ψ_n ist sie äquivalent zu

$$|f(a_\nu, \psi_n(a_\nu)) - f(x, \psi_n(x))| \leq \varepsilon$$

für $a_\nu \leq x \leq a_{\nu+1}$ und alle hinreichend großen n . Dies ist aber eine unmittelbare Konsequenz aus der gleichmäßigen Stetigkeit der Funktion f auf Q : Zu jedem $\varepsilon > 0$ gibt es ein $\delta > 0$, so daß $|f(x_1, y_1) - f(x_2, y_2)| \leq \varepsilon$, sofern nur die Punkte (x_1, y_1) und $(x_2, y_2) \in Q$ in der Maximumnorm höchstens die Entfernung δ besitzen. Unterteilt man aber das Intervall I fein genug, wählt man also die Anzahl der Teilintervalle N groß genug, so ist dies für die x -Komponente der Punkte $(a_\nu, \psi_n(a_\nu))$, $(x, \psi_n(x))$, $x \in I_\nu$, schon für alle $n \geq N$ erfüllt. Wegen der Eigenschaft ii) liegen aber auch die y -Komponenten nahe genug beieinander, wenn man evtl. N noch vergrößert. \square

Wir zeigen als nächstes, daß die Eigenschaft iii) in dem vorigen Satz eine Integralabschätzung wie in Lemma 12 impliziert. Durch Integration und Anwendung des Hauptsatzes der Differential- und Integralrechnung erhalten wir für jedes $x \in I_\nu$ die Ungleichung

$$\begin{aligned} \left| \psi_n(x) - \psi_n(a_{\nu-1}) - \int_{a_{\nu-1}}^x f(t, \psi_n(t)) dt \right| &= \left| \int_{a_{\nu-1}}^x (\psi'_n(t) - f(t, \psi_n(t))) dt \right| \\ &\leq \int_{a_{\nu-1}}^x |\psi'_n(t) - f(t, \psi_n(t))| dt \leq \varepsilon (x - a_{\nu-1}). \end{aligned}$$

Wendet man diese Ungleichung speziell auf die oberen Grenzen der links von I_ν liegenden Teilintervalle an und addiert alles auf, so erhält man unmittelbar

$$\left| \psi_n(x) - c - \int_a^x f(t, \psi_n(t)) dt \right| \leq \varepsilon ((x - a_{\nu-1}) + (a_{\nu-1} - a_{\nu-2}) + \cdots + (a_1 - a_0)) \leq \varepsilon (b - a).$$

Was uns jetzt lediglich noch zum Beweis des Peanoschen Existenzsatzes fehlt, ist die *gleichmäßige Konvergenz* der Folge der Eulerschen Polygonzüge. Eine solche Aussage darf aber nicht erwartet werden, da man eine entsprechende „Vermutung“ leicht an Gegenbeispielen widerlegen kann.

Abhilfe schafft aber die Erkenntnis, daß es uns reichen würde, wenn eine *Teilfolge* der ψ_k gleichmäßig konvergiert. Dies ist tatsächlich aufgrund eines sehr allgemeinen Satzes immer richtig.

Definition. Eine Familie $\{f_\iota\}_{\iota \in J}$ von stetigen Abbildungen $f_\iota : X \rightarrow Y$ zwischen metrischen Räumen X, Y heißt *gleichgradig stetig*, wenn es zu jedem $\varepsilon > 0$ ein $\delta = \delta(\varepsilon) > 0$ gibt, so daß

$$d_Y(f_\iota(x_1), f_\iota(x_2)) < \varepsilon \quad \text{für alle } x_1, x_2 \in X \quad \text{mit } d_X(x_1, x_2) < \delta \quad \text{und alle } \iota \in J.$$

Bemerkung. Die vorige Definition impliziert, daß notwendigerweise alle Abbildungen f_ι *gleichmäßig stetig* sind.

Aufgrund der Aussage (*) in Lemma 13 ist unmittelbar klar, daß die Familie \mathcal{F}' gleichgradig stetig ist. Sie ist wegen der Eigenschaft $\mathcal{F}' \subset \mathcal{F}$ offensichtlich auch gleichmäßig (also auch punktweise) *beschränkt*, und zwar durch $|c| + r =: C$. - Somit sind alle Schwierigkeiten behoben, wenn der folgende Satz richtig ist.

Satz 20.15 (Arzelà - Ascoli) *Es sei $f_\iota : I = [a, b] \rightarrow \mathbb{R}^m$, $\iota \in J$, eine Familie von Funktionen, die gleichgradig stetig und punktweise beschränkt ist. Dann gibt es zu jeder Folge in dieser Familie eine Teilfolge, die gleichmäßig auf I konvergiert.*

Beweis. Es sei ohne Einschränkung von vornherein $J = \mathbb{N}$ und f_n eine gleichgradig stetige und punktweise beschränkte Folge von Funktionen. Wir wählen weiter eine abzählbare dichte Teilmenge $M = \{a_0, a_1, a_2, \dots\} \subset I$, also z. B. $M := \mathbb{Q} \cap I$ mit einer geeigneten Abzählung $\mathbb{N} \cong M$. Nach

dem Satz von Bolzano–Weierstraß kann man eine Teilfolge $f_{0,n}$ der Folge f_n auswählen, die an der Stelle a_0 konvergiert, sodann aus $f_{0,n}$ eine Teilfolge $f_{1,n}$, die an der Stelle a_1 konvergiert etc. Auf diese Weise erhält man induktiv für alle $k \in \mathbb{N}$ eine Teilfolge $f_{k,n}$, die an der Stelle a_k konvergiert. Die „Diagonal-Folge“ $f_{k,k}$ ist dann eine Teilfolge von f_n , und da diese bei festem ℓ für $k \geq \ell$ eine Teilfolge von $(f_{\ell,n})_{n \in \mathbb{N}}$ ist, konvergiert sie an jeder Stelle a_ℓ .

Wir können also ohne Einschränkung der Allgemeinheit annehmen, daß die ursprüngliche Folge der f_n selbst schon auf der Menge M punktweise konvergiert, und zeigen, daß die Folge dann automatisch auf I gleichmäßig konvergent ist. Wegen der Kompaktheit von I genügt dazu der Nachweis der *lokal gleichmäßigen* Konvergenz: Zu jedem $x_0 \in I$ und jedem $\varepsilon > 0$ gibt es ein $\delta > 0$ und ein $N \in \mathbb{N}$, so daß $\|f_j(x) - f_k(x)\| < \varepsilon$ für alle $j \geq k \geq N$ und alle $x \in I$ mit $|x - x_0| < \delta$.

Die letzte Aussage folgt aber aus der Voraussetzung der *gleichgradigen Stetigkeit*. Setzen wir in der Definition 2 $\delta := \delta(\varepsilon/3)$, so ist also $\|f_n(x) - f_n(x')\| < \varepsilon/3$ für alle $n \in \mathbb{N}$ und alle $x, x' \in I \cap I_\delta(x_0)$. Da die Menge M dicht in I liegt, existiert des Weiteren ein $a_\ell \in I \cap I_\delta(x_0)$, und nach Konstruktion gibt es ein $N \in \mathbb{N}$, so daß $\|f_j(a_\ell) - f_k(a_\ell)\| < \varepsilon/3$ für alle $j \geq k \geq N$. Zusammen ergibt sich aus der Dreiecksungleichung für alle $j \geq k \geq N$ und alle $x \in I \cap I_\delta(x_0)$ die gewünschte Abschätzung

$$\|f_j(x) - f_k(x)\| \leq \|f_j(x) - f_j(a_\ell)\| + \|f_j(a_\ell) - f_k(a_\ell)\| + \|f_k(a_\ell) - f_k(x)\| < \varepsilon. \quad \square$$

Bemerkung. Der Satz von ARZELÀ–ASCOLI hat offensichtlich das Flair des Satzes von Bolzano und Weierstraß. Dies ist nicht zufällig so. Man kann nämlich die *relativ-kompakten*, d. h. nach Abschluß kompakten Teilmengen H im Raum der stetigen Abbildungen $f: X \rightarrow Y$ von einem lokal-kompakten topologischen Raum X in einen normierten Raum Y , versehen mit der sogenannten kompakt-offenen Topologie, gerade charakterisieren durch die gleichgradige Stetigkeit der Familie H und die Forderung, daß für alle $x \in X$ die Menge $\{h(x) : h \in H\}$ relativ kompakt in Y liegt. Hierbei können die Ansprüche an Y sogar noch abgeschwächt werden³⁸.

Wir beschließen diesen Anhang wie angekündigt mit einer zweiten, wesentlich eleganteren Beweisvariante des Existenzsatzes von Peano, die ohne die Eulerschen Polygonzüge auskommt. Dazu setzen wir für $\alpha := (b - a)/n$, $n \in \mathbb{N}^*$,

$$\varphi_\alpha(t) := \begin{cases} c & \text{für } t \leq a, \\ c + \int_a^t f(s, \varphi_\alpha(s - \alpha)) ds & \text{für } a \leq t \leq b. \end{cases}$$

Hierdurch ist die Funktion φ_α tatsächlich wohldefiniert; denn für $0 \leq t - a \leq \alpha$ ist $\varphi_\alpha(t - \alpha) = c$, und damit existiert das Integral auf der rechten Seite. Es ist daher nur durch vollständige Induktion nach ν sicherzustellen, daß aus $|\varphi_\alpha - c| \leq r$ für $(\nu - 1)\alpha \leq t - a \leq \nu\alpha$ dieselbe Abschätzung auf dem Intervall $\nu\alpha \leq t - a \leq (\nu + 1)\alpha$ folgt. Dies ist aber unmittelbar klar wegen

$$|\varphi_\alpha(t) - c| \leq \int_a^t |f(s, \varphi_\alpha(s - \alpha))| ds \leq M(b - a) \leq r.$$

Weiter ist φ_α stückweise differenzierbar mit $|\varphi'_\alpha| \leq M$ auf den obigen Teilintervallen, woraus sich sofort

$$|\varphi_\alpha(t_1) - \varphi_\alpha(t_2)| \leq M|t_1 - t_2| \quad \text{für } t_1, t_2 \in I$$

ergibt.

Somit ist die Folge $\varphi_\alpha = \varphi_{\alpha_n}$ auf I punktweise (sogar gleichmäßig) beschränkt (durch $|c| + r$) und wegen der letzten Ungleichung auch gleichgradig stetig. Man kann daher nach dem Satz von ARZELÀ - ASCOLI eine auf I gleichmäßig konvergente Teilfolge $\varphi_k = \varphi_{\alpha_{n_k}}$ auswählen. Ist φ der (stetige) Grenzwert, so konvergiert wegen

$$|\varphi_k(s - \alpha_{n_k}) - \varphi(s)| \leq |\varphi_k(s - \alpha_{n_k}) - \varphi_k(s)| + |\varphi_k(s) - \varphi(s)| \leq M\alpha_{n_k} + |\varphi_k(s) - \varphi(s)|$$

³⁸Siehe z. B. [39], pp. 167 ff.

auch die Folge $\psi_k(s) := \varphi_k(s - \alpha_{n_k})$ gleichmäßig gegen φ , so daß wir das obigen Lemma 12 direkt anwenden können. \square

21 Gewöhnliche Differentialgleichungen höherer Ordnung und Differentialgleichungssysteme

Wir beschäftigen uns in diesem Kapitel mit *gewöhnlichen Differentialgleichungen höherer Ordnung* und *Systemen von Differentialgleichungen*. Die Behandlung solcher Differentialgleichungstypen wird schon allein durch die Erfordernisse der klassischen Physik erzwungen. Bewegt sich z. B. ein Masseteilchen der Masse m in einem Bereich G des dreidimensionalen Raumes unter dem Einfluß einer Kraft F , so erfüllt sein Bahnvektor $x = x(t)$ die NEWTONschen *Bewegungsgleichungen*

$$m \frac{d^2 x}{dt^2} = F(x(t)).$$

Mathematisch exakt formuliert, handelt es sich hierbei um ein *System von drei gewöhnlichen Differentialgleichungen zweiter Ordnung*. Bezeichnet man nämlich die Koordinaten in \mathbb{R}^3 mit $x = (x_1, x_2, x_3)$, so ist eine Kraft F nichts anderes als eine Abbildung $F : G \rightarrow \mathbb{R}^3$, und sind F_1, F_2, F_3 ihre Komponentenfunktionen, so besteht eine Lösung der obigen „Differentialgleichung“ aus einem System $x(t) = (x_1(t), x_2(t), x_3(t))$ von drei *differenzierbaren* Funktionen auf einem Intervall $I = [a, b]$, so daß $x(t) \in G$ für alle $t \in I$ und

$$m \frac{d^2 x_i(t)}{dt^2} = F(x_1(t), x_2(t), x_3(t)), \quad t \in [a, b], \quad i = 1, 2, 3.$$

Noch etwas *geometrischer* gewendet suchen wir also *differenzierbare Kurven* $\alpha : I \rightarrow G$ mit

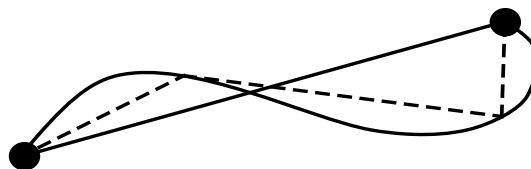
$$m \frac{d^2 \alpha(t)}{dt^2} = F(\alpha(t)), \quad t \in I.$$

Bemerkung. Wir werden weiter unten sehen, daß sich solche Systeme nach Einführung neuer Variabler stets in Systeme von Differentialgleichungen erster Ordnung umwandeln lassen.

Die Notwendigkeit, Systeme von Differentialgleichungen zu untersuchen, kommt aber auch aus der einfachsten Grundaufgabe der *Variationsrechnung*. Es handelt sich hierbei um eine Verallgemeinerung der Aufgabe, ein Extremum einer Funktion zu finden, wie z. B. in der folgenden Situation: Man betrachte die Menge \mathcal{F} aller *stetig differenzierbaren Kurven* $\alpha : I \rightarrow \mathbb{R}^n$ auf dem Einheitsintervall $I = [0, 1]$ mit fest gewähltem *Anfangspunkt* x_0 und festem *Endpunkt* x_1 und bestimme unter diesen eine Kurve α_0 , die das *Funktional*

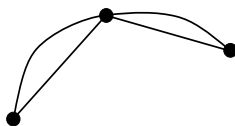
$$L(\alpha) := \int_0^1 \|\alpha'(t)\|_2 dt$$

minimiert, für die also $L(\alpha_0) \leq L(\alpha)$ für alle $\alpha \in \mathcal{F}$. Die Größe $L(\alpha)$ ist natürlich nichts anderes als die (euklidische) *Länge* des Kurvenbogens $\alpha(I)$, so daß wir also nach der *kürzesten Verbindung* zwischen zwei Punkten im euklidischen Raum (auf allen stetig differenzierbaren Wegen) fragen. Selbstverständlich besitzt dieses Problem eine (eindeutig bestimmte) Lösung, nämlich die *geradlinige Strecke* zwischen x_0 und x_1 , denn wegen der Dreiecksungleichung ist dies offensichtlich sogar die Lösung des Problems in der größeren Klasse der *rektifizierbaren* Kurven.



Figur 21.1

Bemerkung. Man muß sich allerdings vor der Annahme hüten, daß ein solches Problem stets eine Lösung besitzt. Darauf hat z. B. HILBERT mit dem folgenden Beispiel hingewiesen: In der vorigen Situation setze man zusätzlich voraus, daß die Lösungskurve α auf dem Weg von x_0 nach x_1 stets noch durch einen weiteren Punkt x_2 verläuft, der nicht auf der Verbindungsgeraden von x_0 nach x_1 liegt. Dann gibt es natürlich eine eindeutig bestimmte Lösung in der Menge der rektifizierbaren Kurven, nämlich den Streckenzug von x_0 über x_2 nach x_1 . Diese ist jedoch nicht *differenzierbar*, wohl aber beliebig genau bzgl. der Länge durch stetig differenzierbare Kurven approximieren. Somit kann das Problem keine *stetig differenzierbare* Lösung besitzen.



Figur 21.2

Wir wollen nun zeigen, daß unter sehr präzisen Voraussetzungen eine mögliche Minimal-Lösung einem Differentialgleichungssystem genügen muß. Dies ist aber, wie im Falle der Bestimmung von Extrema einer Funktion, nur eine *notwendige* Bedingung, und es ist somit stets nachzuprüfen, ob das Differentialgleichungssystem eine eindeutige Lösung besitzt und diese tatsächlich unser ursprüngliches Problem löst. Insbesondere ist a priori nicht ausgeschlossen, daß unser Verfahren eine *Maximal*-Lösung ergibt.

Satz 21.1 (Euler - Lagrange - Gleichungen der Variationsrechnung) *Es sei $\tilde{G} \subset \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}$ ein Gebiet und $L = L(q_1, \dots, q_n, p_1, \dots, p_n, t)$ eine zweimal stetig partiell differenzierbare Funktion auf \tilde{G} . Es sei $I := [a, b]$ ein fest vorgegebenes kompaktes Intervall, x_0, x_1 seien Punkte in \mathbb{R}^n , und \mathcal{F} sei die Menge aller zweimal stetig differenzierbaren Kurven $\alpha : I \rightarrow \mathbb{R}^n$ mit $\alpha(0) = x_0$, $\alpha(1) = x_1$ und $(\alpha(t), \alpha'(t), t) \in \tilde{G}$ für alle $t \in I$. Dann ist die Kurve $\alpha \in \mathcal{F}$ extremal bzgl. des Funktionals*

$$W(\alpha) := \int_a^b L(\alpha(t), \alpha'(t), t) dt,$$

wenn α den folgenden Differentialgleichungen genügt:

$$\frac{d}{dt} \left(\frac{\partial L}{\partial p_j}(\alpha(t), \alpha'(t), t) \right) = \frac{\partial L}{\partial q_j}(\alpha(t), \alpha'(t), t), \quad j = 1, \dots, n.$$

Bemerkungen. 1. Selbstverständlich kann die Menge \mathcal{F} bei beliebig vorgegebenen I, \tilde{G}, x_0, x_1 leer sein. Dies ist jedoch nicht der Fall, wenn $G \subset \mathbb{R}^n$ ein Gebiet und \tilde{G} von der Gestalt $G \times \mathbb{R}^n \times I$ bzw. $G \times (\mathbb{R}^n \setminus \{0\}) \times I$ ist. Im zweiten Fall bedeutet dies, daß es stets von x_0 nach x_1 eine durch das vorgegebene Intervall I parametrisierte stetig differenzierbare Kurve $\alpha : I \rightarrow G$ gibt, deren Ableitung nirgends verschwindet. Solche (parametrisierten) Kurven nennt man auch *glatt*.

2. Wir schreiben die EULER-LAGRANGE-Differentialgleichungen oft auch in der abgekürzten Form

$$\frac{d}{dt} \left(\frac{\partial L}{\partial p_j} \right) = \frac{\partial L}{\partial q_j}, \quad j = 1, \dots, n.$$

Wegen der Voraussetzung, daß die sogenannte LAGRANGE-Funktion L zweimal stetig partiell differenzierbar ist, ist diese Gleichung nach der Kettenregel ein (implizites) „stetiges“ System von Differentialgleichungen zweiter Ordnung in den unbekanntenen Komponentenfunktionen von α .

3. Man nennt in der Physik $W = W(\alpha)$ auch die *Wirkung* des Systems³⁹ auf die Kurve oder *Bahn*

³⁹Dort hat die Lagrange-Funktion die physikalische Dimension einer *Energie*. Größen der Dimension Energie mal Zeit heißen stets *Wirkung*.

α . Man spricht daher auch (fälschlicher Weise) von dem „Prinzip der kleinsten Wirkung“. Korrekt muß es natürlich das Prinzip der *stationären* Wirkung heißen.

Bevor wir den (einfachen) Beweis des vorstehenden zentralen Satzes durchführen, beleuchten wir seinen Inhalt an dem weiter oben schon erwähnten *Beispiel* der kürzesten Verbindung zwischen zwei Punkten in \mathbb{R}^n . Die LAGRANGE-Funktion lautet hier

$$L = L(p_1, \dots, p_n) := \|p\|_2, \quad p = (p_1, \dots, p_n) \in \mathbb{R}^n;$$

sie ist also unabhängig von den Variablen q_j und t . Damit sie die geforderte Differenzierbarkeitseigenschaft besitzt, muß man $\tilde{G} = \mathbb{R}^n \times (\mathbb{R}^n \setminus \{0\}) \times I$ setzen. Wir suchen damit die Lösungen des Problems nur unter allen *glatten* parametrisierten Kurven $\alpha : I \rightarrow \mathbb{R}^n$. Die Euler-Lagrangeschen Differentialgleichungen lauten dann:

$$\frac{d}{dt} \left(\frac{p_j}{\|p\|_2} \right) = \frac{d}{dt} \left(\frac{\partial L}{\partial p_j} \right) = \frac{\partial L}{\partial q_j} = 0.$$

Also ist

$$\frac{p_j}{\|p\|_2} = c_j$$

konstant; für die gesuchte glatt parametrisierte Kurve $\alpha : I \rightarrow \mathbb{R}^n$ muß also notwendig - neben der Anfangs- und Endbedingung - die Gleichung

$$\frac{\alpha'_j(t)}{\|\alpha'(t)\|} = c_j \quad \text{oder zusammengefaßt} \quad \frac{\alpha'(t)}{\|\alpha'(t)\|} = c$$

für einen festen Vektor $c \in \mathbb{R}^n$ erfüllt sein. Wir führen jetzt für die Kurve α durch

$$s = \sigma(t) = \int_a^t \|\alpha'(\tau)\| d\tau$$

eine neue Parametrisierung ein: Die Funktion σ ist differenzierbar mit Ableitung $\sigma'(t) = \|\alpha'(t)\| > 0$ und damit streng monoton wachsend. Sie bildet das Intervall $[a, b]$ bijektiv auf das Intervall $[0, \lambda]$ ab, wobei λ gerade die *Kurvenlänge* bedeutet. Damit beschreibt auch

$$\tilde{\alpha}(s) := \alpha(\sigma^{-1}(s)), \quad s \in [0, \lambda],$$

den gleichen „Kurvenbogen“ wie α , und aus der Kettenregel folgt

$$\tilde{\alpha}'(s) := \frac{\alpha'(\sigma^{-1}(s))}{\|\alpha'(\sigma^{-1}(s))\|} = c$$

und damit notwendigerweise unter Berücksichtigung der Randbedingungen die erwartete *Streckengleichung*

$$\tilde{\alpha}(s) := cs + x_0 \quad \text{mit} \quad c = \frac{x_1 - x_0}{\lambda}, \quad s \in [0, \lambda].$$

Bemerkung. Parametrisierungen von Kurven sind in weitem Maße willkürlich. Hat man es aber mit einer *glatten* Kurve zu tun, so kann man stets ihre *Bogenlänge* als „den natürlichen“ Parameter einführen. Genau dies haben wir in dem Beispiel oben getan. Mehr über dieses Sachverhalt findet man in einem späteren Kapitel über Kurven.

Beweis von Satz 1. Es sei ohne Beschränkung der Allgemeinheit $\alpha \in \mathcal{F}$ eine *minimale* Lösung. Es seien weiter $g_1, \dots, g_n : I \rightarrow \mathbb{R}$ beliebige zweimal stetig differenzierbare Funktionen mit $g_j(a) = g_j(b) = 0$, und $\alpha_s : I \rightarrow \mathbb{R}^n$ sei für beliebiges $s \in \mathbb{R}$ definiert durch $\alpha_s(t) := \alpha(t) + sg(t)$ mit $g := (g_1, \dots, g_n) : I \rightarrow \mathbb{R}^n$. Dann existiert ein $\varepsilon > 0$, so daß für alle $|s| < \varepsilon$ und alle $t \in I$ der

Vektor $(\alpha_s(t), \alpha'(t) + s g'(t), t)$ in \tilde{G} enthalten ist. Denn sonst existierte eine Folge von Zahlen $s_j \rightarrow 0$ und $t_j \in I$ mit $(\alpha(t_j) + s_j g(t_j), \alpha'(t_j) + s_j g'(t_j), t_j) \notin \tilde{G}$. Nach Übergang zu einer Teilfolge können wir weiter annehmen, daß die Folge (t_j) gegen ein Element τ in I konvergiert. Da das Komplement von \tilde{G} abgeschlossen ist, liegt dann auch der Grenzwert der obigen Folge im Komplement, also nicht in \tilde{G} . Andererseits ist dieser Grenzwert aber gleich $(\alpha(\tau), \alpha'(\tau), \tau) \in \tilde{G}$. Widerspruch!

Wir können daher für hinreichend kleine s die Funktion

$$\tilde{W}(s) := W(\alpha_s)$$

bilden. Nach Voraussetzung ist $\tilde{W}(s) = W(\alpha_s) \geq \tilde{W}(\alpha) = \tilde{W}(0)$. Nun ist $\tilde{W}(s)$ ein parameterabhängiges Integral mit dem (stetigen) Integranden

$$L(\alpha_s(t), \alpha'_s(t), t),$$

der nach der Kettenregel stetig partiell nach s differenzierbar ist mit der Ableitung

$$\frac{\partial L(\alpha_s(t), \alpha'_s(t), t)}{\partial s} = \sum_{j=1}^n \frac{\partial L}{\partial q_j}(\alpha_s(t), \alpha'_s(t), t) g_j(t) + \sum_{j=1}^n \frac{\partial L}{\partial p_j}(\alpha_s(t), \alpha'_s(t), t) g'_j(t).$$

Somit ist nach dem Satz über die Differenzierbarkeit solcher Integrale (siehe Satz 18.35 mit anschließender Bemerkung):

$$\sum_{j=1}^n \int_a^b \left(\frac{\partial L}{\partial q_j}(\alpha(t), \alpha'(t), t) g_j(t) + \frac{\partial L}{\partial p_j}(\alpha(t), \alpha'(t), t) g'_j(t) \right) dt = \left. \frac{d\tilde{W}}{ds} \right|_{s=0} = 0.$$

Mit partieller Integration erhält man aber

$$\int_a^b \frac{\partial L}{\partial p_j}(\alpha(t), \alpha'(t), t) g'_j(t) dt = \left. \frac{\partial L}{\partial p_j}(\alpha(t), \alpha'(t), t) g_j(t) \right|_a^b - \int_a^b \frac{d}{dt} \left(\frac{\partial L}{\partial p_j}(\alpha(t), \alpha'(t), t) \right) g_j(t) dt,$$

so daß diese Bedingung wegen $g_j(a) = g_j(b) = 0$ umformuliert werden kann in

$$\sum_{j=1}^n \int_a^b \left(\frac{\partial L}{\partial q_j}(\alpha(t), \alpha'(t), t) - \frac{d}{dt} \left(\frac{\partial L}{\partial p_j}(\alpha(t), \alpha'(t), t) \right) \right) g_j(t) dt = 0.$$

Wählt man hier alle g_k , $k \neq j$, bei festem j identisch gleich Null, so impliziert dies die Gleichungen

$$\int_a^b \left(\frac{\partial L}{\partial q_j}(\alpha(t), \alpha'(t), t) - \frac{d}{dt} \left(\frac{\partial L}{\partial p_j}(\alpha(t), \alpha'(t), t) \right) \right) g_j(t) dt = 0$$

für alle $j = 1, \dots, n$ und beliebige Funktionen g_j . Dann muß aber der Integrand

$$\frac{\partial L}{\partial q_j}(\alpha(t), \alpha'(t), t) - \frac{d}{dt} \left(\frac{\partial L}{\partial p_j}(\alpha(t), \alpha'(t), t) \right)$$

gleich Null sein. □

Bemerkung. Im letzten Schritt des vorigen Beweises haben wir den folgenden Schluß verwendet, den wir gesondert notieren.

Lemma 21.2 *Es sei $f : I := [a, b] \rightarrow \mathbb{R}$ eine stetige Funktion mit*

$$\int_a^b f(t) g(t) dt = 0$$

für alle zweimal stetig differenzierbaren Funktionen $g : I := [a, b] \rightarrow \mathbb{R}$ mit $g(a) = g(b) = 0$. Dann ist f identisch Null.

Beweis. Angenommen, f sei an einer Stelle $c \in I$ von Null verschieden, also ohne Einschränkung der Allgemeinheit positiv. Wegen der Stetigkeit von f können wir dann annehmen, daß es ein $\varepsilon > 0$ und ein $\delta > 0$ gibt, so daß $a + \delta \leq c \leq b - \delta$ und $f(x) \geq \varepsilon$ für alle $x \in J$ mit $J := \{x \in \mathbb{R} : |x - c| \leq \delta\}$. Es sei nun $g : \mathbb{R} \rightarrow \mathbb{R}$ eine nicht negative, zweimal stetig differenzierbare Funktion mit $g(x) = 0$ für $|x - c| \geq \delta$ und $g(x) = \varepsilon$ für $|x - c| \leq \delta/2$.⁴⁰ Dann ist

$$\int_a^b f(t)g(t) dt = \int_{c-\delta}^{c+\delta} f(t)g(t) dt \geq \int_{c-\delta/2}^{c+\delta/2} f(t)g(t) dt \geq \delta\varepsilon^2 > 0. \quad \square$$

Wir führen als nächstes den Begriff des *Systems von gewöhnlichen Differentialgleichungen erster Ordnung* präzise ein und übertragen unsere früheren Existenz- und Eindeutigkeitsresultate auf diese neue Situation.

Definition. Es sei $G \subset \mathbb{R}^{n+1} = \mathbb{R} \times \mathbb{R}^n$ ein Gebiet (mit Variablen x, y_1, \dots, y_n , wobei wir den Vektor ${}^t(y_1, \dots, y_n) \in \mathbb{R}^n$ oft mit Y oder y bezeichnen, und $F : G \rightarrow \mathbb{R}^n$ sei eine stetige Abbildung. Dann heißt

$$(*) \quad Y' = F(x, Y)$$

ein *System von (expliziten) gewöhnlichen Differentialgleichungen erster Ordnung*. Eine *Lösung* dieses Systems ist eine differenzierbare Abbildung $\Phi : I \rightarrow \mathbb{R}^n$ auf einem Intervall $I \subset \mathbb{R}$ mit den folgenden Eigenschaften:

- i) $(x, \Phi(x)) \in G$ für alle $x \in I$.
- ii) $\Phi'(x) = F(x, \Phi(x))$ für alle $x \in I$.

Bezeichnet man die *Komponentenfunktionen* von F mit f_1, \dots, f_n und die der Lösung Φ mit $\varphi_1, \dots, \varphi_n$, so schreibt sich das obige System explizit in der Form

$$\begin{aligned} y'_1 &= f_1(x, y_1, \dots, y_n) \\ y'_2 &= f_2(x, y_1, \dots, y_n) \\ &\vdots \\ y'_n &= f_n(x, y_1, \dots, y_n), \end{aligned}$$

und die Lösungsbedingung ii) lautet vollständig ausgeschrieben:

$$\begin{aligned} \varphi'_1 &= f_1(x, \varphi_1(x), \dots, \varphi_n(x)) \\ \varphi'_2 &= f_2(x, \varphi_1(x), \dots, \varphi_n(x)) \\ &\vdots \\ \varphi'_n &= f_n(x, \varphi_1(x), \dots, \varphi_n(x)). \end{aligned}$$

Wir sagen, das System genüge einer *LIPSCHITZ-Bedingung*, wenn es eine Konstante $L \geq 0$ gibt, so daß

$$\|F(x, Y) - F(x, \tilde{Y})\| \leq L\|Y - \tilde{Y}\| \quad \text{für alle } (x, Y), (x, \tilde{Y}) \in G.$$

Hierbei ist $\|\cdot\|$ irgendeine Norm auf \mathbb{R}^n .

Bemerkung. Da alle Normen auf \mathbb{R}^n äquivalent sind, ist die letzte Bedingung in der Tat *unabhängig* von der gewählten Norm. Allerdings ist zu beachten, daß die *Lipschitz-Konstante* L sich bei Übergang zu einer anderen Norm im Allgemeinen ändert.

Der *Existenz- und Eindeutigkeitsatz* für gewöhnliche Differentialgleichungen erster Ordnung mit lokaler Lipschitz-Bedingung bleibt (samt Beweis) wortwörtlich für Systeme richtig. Insbesondere führt auch hier das Picard-Lindelöfsche Iterationsverfahren zum Ziel.

⁴⁰Es existieren sogar C^∞ -Funktionen mit diesen Eigenschaften.

Satz 21.3 Das System $Y' = F(x, Y)$ erfülle auf $G \subset \mathbb{R}^{n+1}$ lokal eine Lipschitz-Bedingung. Dann gibt es zu vorgegebener Anfangsbedingung $(a, c) \in G$, $c = (c_1, \dots, c_n)$ lokal um a eine eindeutig bestimmte Lösung $\Phi : I \rightarrow \mathbb{R}^n$ mit $\Phi(a) = c$.

Beweis. Man schreibt das Differentialgleichungssystem wieder um in ein Integralgleichungssystem

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

wobei man unter dem Integral der „vektorwertigen“ Funktion $F(t, \Phi(t))$ den Vektor der Integrale über die Koeffizientenfunktionen zu verstehen hat:

$$\int_a^x F(t, \Phi(t)) dt := \left(\int_a^x f_1(t, \Phi(t)) dt, \dots, \int_a^x f_n(t, \Phi(t)) dt \right).$$

Geht man den ursprünglichen Beweis Punkt für Punkt durch, so stellt man leicht fest, daß er einzig und allein neben Standard-Aussagen über stetige Funktionen, die auch in mehreren Veränderlichen gültig sind, die geläufige *Abschätzung für Integrale* benutzt. Wir benötigen also nur noch eine entsprechende Aussage für die soeben eingeführten Vektoren von Integralen bzgl. einer geeigneten Norm in \mathbb{R}^n . Diese findet man in dem anschließenden Lemma. \square

Lemma 21.4 Es sei $\Phi : I = [a, b] \rightarrow \mathbb{R}^n$ eine stetige Kurve. Dann gilt, wobei $\|\cdot\|_2$ die euklidische Norm in \mathbb{R}^n bezeichnet:

$$\left\| \int_a^b \Phi(t) dt \right\|_2 \leq \int_a^b \|\Phi(t)\|_2 dt.$$

Beweis: Sei $J := \int_a^b \Phi(t) dt$ und $L := \|J\|_2 > 0$. (Bei $L = 0$ ist die Abschätzung trivialerweise richtig). Dann gilt wegen der Cauchy-Schwarzschen Ungleichung:

$$\begin{aligned} L^2 &= \langle J, J \rangle = \left\langle \int_a^b \Phi(t) dt, J \right\rangle = \int_a^b \langle \Phi(t), J \rangle dt \\ &\leq \int_a^b \|\Phi(t)\|_2 \|J\|_2 dt = L \int_a^b \|\Phi(t)\|_2 dt, \end{aligned}$$

also $L \leq \int_a^b \|\Phi(t)\|_2 dt$. \square

Bemerkung. Wir beweisen den Existenz- und Eindeutigkeitssatz für Systeme in Kapitel 23 noch einmal vollständig mit Hilfe des BANACHSchen Fixpunktsatzes.

Bevor wir uns der (einfachen) Beziehung zwischen Differentialgleichungen höherer Ordnung und Systemen erster Ordnung zuwenden, wollen wir noch eine wichtige Abschätzung einschieben, die den Mittelwertsatz von differenzierbaren Funktionen in einer Veränderlichen auf differenzierbare Abbildungen verallgemeinert. Diese Abschätzung wird noch vielfältige Anwendungen finden. In dem gegenwärtigen Zusammenhang gewährt sie uns die Einsicht, daß Systeme von Differentialgleichungen einer lokalen Lipschitz-Bedingung genügen, wenn die rechte Seite F stetig partiell nach den Variablen y_1, \dots, y_n differenzierbar ist.

Wir formulieren dazu den klassischen Mittelwertsatz etwas um.

Satz 21.5 (Mittelwertsatz) Es sei $f : I \rightarrow \mathbb{R}$ einmal stetig differenzierbar, und es seien $x \in I$ und $\xi \in \mathbb{R}$ so beschaffen, daß $x + \xi \in I$. Dann gilt:

$$f(x + \xi) - f(x) = \left(\int_0^1 f'(x + t\xi) dt \right) \xi.$$

Beweis. Mit dem Hauptsatz der Differential- und Integralrechnung wird

$$f(x + \xi) - f(x) = \int_x^{x+\xi} f'(u) du .$$

Wir substituieren $u = x + t\xi$ als Funktion in $t \in [0, 1]$ bei festem x und ξ und erhalten die Behauptung aus der Substitutionsformel. \square

Wir definieren jetzt auch für eine „matrixwertige stetige Kurve“ auf $I = [a, b]$, also für eine stetige Abbildung

$$A : I \longrightarrow M(m \times n, \mathbb{R}) = \mathbb{R}^{mn} , \quad A(t) = (a_{jk}(t)) , \quad t \in I ,$$

das Integral als die reelle $m \times n$ -Matrix

$$\int_a^b A(t) dt := \left(\int_a^b a_{jk}(t) dt \right)_{j,k} \in M(m \times n, \mathbb{R})$$

und gewinnen sofort die folgende Verallgemeinerung:

Satz 21.6 (Mittelwertsatz für stetig differenzierbare Abbildungen) *Es sei $U \subset \mathbb{R}^n$ offen, $F : U \rightarrow \mathbb{R}^m$ stetig differenzierbar, und $x \in U$, $\xi \in \mathbb{R}^n$ seien so beschaffen, daß die Strecke $\{x + t\xi : t \in [0, 1]\}$ ganz in U verläuft. Dann gilt*

$$F(x + \xi) - F(x) = \left(\int_0^1 DF(x + t\xi) dt \right) \xi \quad \text{mit der Funktionalmatrix } DF = \frac{\partial (f_1, \dots, f_m)}{\partial (x_1, \dots, x_n)} .$$

Beweis. Dies ist eine unmittelbare Anwendung der Kettenregel, angewendet auf die Funktionen

$$g_j(t) := f_j(x + t\xi) , \quad t \in [0, 1] , \quad j = 1, \dots, m ,$$

wobei die f_j die Komponentenfunktionen von F bezeichnen. Es ist nämlich

$$g'_j(t) = \sum_{k=1}^n \frac{\partial f_j}{\partial x_k}(x + t\xi) \xi_k$$

und damit für alle $j = 1, \dots, m$:

$$f_j(x + \xi) - f_j(x) = g_j(1) - g_j(0) = \int_0^1 g'_j(t) dt = \sum_{k=1}^n \left(\int_0^1 \frac{\partial f_j}{\partial x_k}(x + t\xi) dt \right) \cdot \xi_k . \quad \square$$

Als Folgerung erhalten wir hieraus den auch oft so genannten *Schrankensatz*.

Folgerung 21.7 *Es sei $F : U \rightarrow \mathbb{R}^m$ stetig differenzierbar, $K \subset U$ sei kompakt und konvex, und es werde*

$$\|DF\|_K := \sup_{x \in K} \|DF(x)\|$$

mit der Operatornorm $\|\cdot\|$ für Matrizen definiert. Dann gilt für alle $x, \tilde{x} \in K$:

$$\|F(\tilde{x}) - F(x)\|_2 \leq \|DF\|_K \|\tilde{x} - x\|_2 .$$

Beweis. Mit $\tilde{x} = x + \xi$ folgt bzgl. der euklidischen Normen in \mathbb{R}^n bzw. \mathbb{R}^m :

$$\|F(x + \xi) - F(x)\| = \left\| \int_0^1 DF(x + t\xi) \cdot \xi dt \right\| \leq \int_0^1 \|DF(x + t\xi) \cdot \xi\| dt \leq \|DF\|_K \cdot \|\xi\| . \quad \square$$

Bemerkung. Mit einer völlig analogen Argumentation erhält man den Schrankensatz auch bzgl. der *Supremumsnorm* in der folgenden Gestalt:

$$\|F(\tilde{x}) - F(x)\|_\infty \leq \|DF\|_K \|\tilde{x} - x\|_\infty,$$

wobei hier aber $\|DF\|_K$ stehen muß für das Supremum aller Operatornormen $\|DF(x)\|$, $x \in K$, in Bezug auf die *Supremumsnormen* in \mathbb{R}^n und \mathbb{R}^m . Diese ist bekanntlich die sogenannte *Zeilensummen-norm* (siehe den Anhang zu Kapitel 17), so daß wir hier die *Schranke* ganz explizit schreiben können als

$$\sup_{x \in K} \left(\max_{j=1, \dots, m} \sum_{k=1}^n \left| \frac{\partial f_j}{\partial x_k}(x) \right| \right).$$

In dieser Version findet der Schrankensatz oft Verwendung in numerischen Kalkulationen.

Aus den vorigen Überlegungen folgt weiter die früher schon erwähnte hinreichende Bedingung für lokale Lipschitz-Bedingung.

Satz 21.8 *Es sei $F : G \rightarrow \mathbb{R}^n$, $G \subset \mathbb{R} \times \mathbb{R}^n$, stetig partiell differenzierbar nach y_1, \dots, y_n . Dann genügt F lokal einer Lipschitz-Bedingung.*

Beweis. Sei $(a, c) \in G$ ein fest gewählter Punkt. Dann existiert eine Zahl $r > 0$, so daß der Zylinder $Z := \{(x, Y) \in \mathbb{R} \times \mathbb{R}^n : |x - a| \leq r, \|Y - c\| \leq r\}$ in G enthalten ist. Da Z kompakt ist, existiert

$$M := \sup_{(x, Y) \in Z} \left\{ \left\| \frac{\partial F}{\partial Y}(x, Y) \right\| \right\},$$

wobei $\left\| \frac{\partial F}{\partial Y}(x, Y) \right\|$ die Operatornorm der $n \times n$ -Matrix

$$\frac{\partial F}{\partial Y} := \frac{\partial (f_1, \dots, f_n)}{\partial (y_1, \dots, y_n)}$$

bezeichnet. Der Mittelwertsatz (bei festem x und variablem Y) impliziert die Behauptung. \square

Kommen wir nun schließlich noch zu Differentialgleichungen n -ter Ordnung.

Definition. Es sei $f : G \rightarrow \mathbb{R}$ eine stetige Funktion auf dem Gebiet $G \subset \mathbb{R} \times \mathbb{R}^n$. Unter einer (expliziten) gewöhnlichen Differentialgleichung n -ter Ordnung verstehen wir eine Gleichung der Form

$$(**) \quad y^{(n)} = f(x, y, y', \dots, y^{(n-1)}).$$

Eine *Lösung* einer solchen Differentialgleichung ist eine n -mal differenzierbare reellwertige Funktion $\varphi : I \rightarrow \mathbb{R}$ mit

- a) $(x, \varphi(x), \varphi'(x), \dots, \varphi^{(n-1)}(x)) \in G$ für alle $x \in I$.
- b) $\varphi^{(n)}(x) = f(x, \varphi(x), \varphi'(x), \dots, \varphi^{(n-1)}(x))$ für alle $x \in I$.

Man kann ohne Schwierigkeiten eine solche Differentialgleichung *höherer Ordnung* „umwandeln“ in ein System 1. Ordnung. Man führt dazu neue Variable y_1, \dots, y_n für die sukzessiven Ableitungen der gesuchten Lösung ein und ordnet also der Differentialgleichung

$$y^{(n)} = f(x, y, y', \dots, y^{(n-1)})$$

das System

$$(***) \quad \begin{cases} y_1' &= y_2 \\ y_2' &= y_3 \\ &\vdots \\ y_{n-1}' &= y_n \\ y_n' &= f(x, y_1, y_2, \dots, y_n) \end{cases}$$

zu, oder anders gesagt: der Funktion $f : G \rightarrow \mathbb{R}$ wird die Abbildung $F = {}^t(f_1, \dots, f_n) : G \rightarrow \mathbb{R}^n$ zugeordnet mit $f_1(x, Y) := y_2, \dots, f_{n-1}(x, Y) := y_n, f_n(x, Y) := f(x, Y)$.

Daß dieses System tatsächlich „äquivalent“ zu der ursprünglichen Differentialgleichung n -ter Ordnung ist, besagt das folgende, leicht einzusehende

Lemma 21.9 *Jeder Lösung $\varphi : I \rightarrow \mathbb{R}$ der Differentialgleichung $(**)$ wird durch*

$$\Phi := {}^t(\varphi, \varphi', \dots, \varphi^{(n-1)})$$

*eine Lösung des Systems $(***)$ zugeordnet.*

*Ist umgekehrt $\Phi = {}^t(\varphi_1, \dots, \varphi_n)$ eine Lösung des zugeordneten Systems, so ist $\varphi := \varphi_1$ Lösung der ursprünglichen Differentialgleichung $(**)$.*

Beweis. Im zweiten Fall ist

$$\varphi_1^{(n)}(x) = \varphi_2^{(n-1)}(x) = \dots = \varphi_n'(x) = f(x, \varphi_1(x), \varphi_1'(x), \dots, \varphi_1^{(n-1)}(x)).$$

Der erste Teil folgt ebenso schnell. □

Es sollte dem Leser ohne weitere Erläuterungen klar sein, was es besagen soll, daß die rechte Seite f einer gewöhnlichen Differentialgleichung n -ter Ordnung einer (lokalen) Lipschitz-Bedingung genügt. Wir überlassen ihm/ihr den Nachweis des folgenden Lemmas, wobei man am günstigsten die *Maximum-Norm* auf \mathbb{R}^n verwendet.

Lemma 21.10 *Die Differentialgleichung $y^{(n)} = f(x, y, y', \dots, y^{(n-1)})$ genügt genau dann einer lokalen Lipschitz-Bedingung, wenn das zugeordnete System $Y' = F(x, Y)$ einer solchen genügt.*

Damit können wir den Existenz- und Eindeigkeitssatz für Systeme gewöhnlicher Differentialgleichungen erster Ordnung ohne Nachdenken auf den Fall von gewöhnlichen Differentialgleichungen n -ter Ordnung übertragen.

Satz 21.11 *Die Funktion $f : G \rightarrow \mathbb{R}, G \subset \mathbb{R} \times \mathbb{R}^n$, genüge lokal einer Lipschitz-Bedingung. Dann gibt es zu vorgegebenen Anfangsbedingungen $(a, c) = (a, c_1, \dots, c_n)$ lokal um a genau eine Lösung $\varphi : I \rightarrow \mathbb{R}$ der Differentialgleichung $y^{(n)} = f(x, y, y', \dots, y^{(n-1)})$ mit*

$$\varphi^{(j)}(a) = c_{j-1}, \quad j = 1, \dots, n.$$

Bemerkung. Selbstverständlich kann man ebenso auch *Systeme* von gewöhnlichen Differentialgleichungen *höherer Ordnung* in Systeme erster Ordnung umwandeln. Für die klassische Mechanik relevant sind insbesondere Systeme *zweiter Ordnung* (siehe den Anhang zu diesem Kapitel). Wie uns aus der Physik vertraut ist, sind Lösungen $\Phi = {}^t(\varphi_1, \dots, \varphi_N)$ solcher Systeme (bei Vorliegen einer lokalen Lipschitz-Bedingung) eindeutig festgelegt durch den „Anfangsort“ $\Phi(a)$ des Systems und seine „Anfangsgeschwindigkeit“ $\Phi'(a)$. Bei n Masseteilchen im dreidimensionalen Raum ist die Anzahl N der Gleichungen zweiter Ordnung gleich $3n$ und damit die Anzahl der eine Bahn bestimmenden „Anfangsparameter“ gleich $2N = 6n$.

Anhang: Die Erhaltungssätze der klassischen Mechanik und die Keplerschen Gesetze

Wir denken uns n *Teilchen* (Massenpunkte) gegeben mit den Massen m_j , $j = 1, \dots, n$, die sich in einem (abgeschlossenen) physikalischen System unter dem Einfluß des von ihrer gegenseitigen Lage abhängigen *Potentials*

$$U = U(r_1, \dots, r_n) = U(x_1, y_1, z_1, \dots, x_n, y_n, z_n)$$

bewegen mögen. Wir schreiben für den Ortsvektor des j -ten Teilchens

$$r_j = r_j(t) = (x_j(t), y_j(t), z_j(t)) \in \mathbb{R}^3$$

und für seinen Geschwindigkeitsvektor

$$v_j = v_j(t) = \frac{dr_j}{dt} = \dot{r}_j.$$

Die NEWTONSchen *Bewegungsgleichungen* schreiben wir leicht abgekürzt in der Form

$$m_j \frac{d^2 r_j}{dt^2} = -\text{grad}_{r_j} U(r_1, \dots, r_n) := -\left(\frac{\partial U}{\partial x_j}, \frac{\partial U}{\partial y_j}, \frac{\partial U}{\partial z_j} \right) =: -\frac{\partial U}{\partial r_j}.$$

Wir wollen diese Gleichungen im LAGRANGE-*Formalismus* umformulieren. Dazu setzen wir

$$T := \sum_{j=1}^n \frac{m_j}{2} \|v_j\|^2, \quad L := T - U.$$

T ist dabei die gesamte *kinetische Energie* des Systems. L heißt auch die *Lagrange-Funktion* des Systems; sie wird aufgefaßt als eine Funktion in den $6n$ Variablen $r_1, \dots, r_n, v_1, \dots, v_n$. Die Gleichungen

$$\frac{\partial L}{\partial \dot{x}_j} = m_j \dot{x}_j, \quad \frac{\partial L}{\partial \dot{y}_j} = m_j \dot{y}_j, \quad \frac{\partial L}{\partial \dot{z}_j} = m_j \dot{z}_j$$

fassen wir zusammen zu

$$\frac{\partial L}{\partial v_j} = m_j v_j.$$

Die Bewegungsgleichungen lauten dann ausgeschrieben

$$\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{x}_j} \right) = m_j \ddot{x}_j = -\frac{\partial U}{\partial x_j} = \frac{\partial L}{\partial x_j} \quad \text{etc.}$$

In der Lagrangeschen Fassung bekommen diese also die Gestalt⁴¹

$$(*) \quad \frac{d}{dt} \left(\frac{\partial L}{\partial v_j} \right) = \frac{\partial L}{\partial r_j}.$$

Diese Gleichungen sollten bei *konservativen* Systemen invariant unter *Zeitverschiebungen* und bei Annahme von *Homogenität* und *Isotropie des Raumes* auch invariant unter Translationen und Drehungen des Raumes sein. Wir wollen zeigen, daß aus diesen Annahmen wichtige *Erhaltungssätze* der klassischen Mechanik folgen. Dahinter steht ein allgemeiner, von E. Noether bewiesener *Erhaltungssatz* bei der Aktion von *Liegruppen*.

⁴¹Lösungen der Newtonschen Bewegungsgleichungen sind also *stationäre Kurven* bzgl. des mit der Lagrange-Funktion gebildeten Wirkungsintegrals. (Siehe Satz 1).

1. Die Homogenität in der Zeit impliziert den *Energieerhaltungssatz*. Hierbei soll die Forderung der Homogenität bedeuten, daß das Potential U *nicht explizit* von der Zeit abhängt. Wir definieren wie üblich die (totale) *Energie* E des Systems durch

$$E := T + U = 2T - L.$$

Die zeitliche Variation auf einer Bahn des Systems ergibt sich dann zu

$$\begin{aligned} \frac{dL}{dt} &= \sum_j \frac{\partial L}{\partial \dot{x}_j} \frac{d\dot{x}_j}{dt} + \dots + \sum_j \frac{\partial L}{\partial \dot{x}_j} \frac{d\dot{x}_j}{dt} + \dots = \sum_j \left\langle \frac{\partial L}{\partial v_j}, \frac{dv_j}{dt} \right\rangle + \sum_j \left\langle \frac{\partial L}{\partial r_j}, \frac{dr_j}{dt} \right\rangle \\ &= \sum_j \left\langle \frac{\partial L}{\partial v_j}, \frac{dv_j}{dt} \right\rangle + \sum_j \left\langle \frac{d}{dt} \left(\frac{\partial L}{\partial v_j} \right), \frac{dr_j}{dt} \right\rangle = \frac{d}{dt} \sum_j \left\langle \frac{\partial L}{\partial v_j}, v_j \right\rangle \\ &= \frac{d}{dt} \sum_j \langle m_j v_j, v_j \rangle. \end{aligned}$$

Daraus folgt sofort $E = -L + 2T = -L + \sum_j m_j \|v_j\|^2 = \text{const.}$

2. Bei *Homogenität des Raumes* wird das Bewegungsgesetz invariant unter *Translationen* sein; wir fordern daher für das Potential die Bedingung

$$U(r_1, \dots, r_n) = U(r_1 + a, \dots, r_n + a)$$

für jeden beliebigen Vektor $a \in \mathbb{R}^3$. Hieraus kann man den *Impulserhaltungssatz* ableiten. Dazu wählen wir r_1, \dots, r_n und a fest und bilden $\tilde{U}(t) = U(r_1 + ta, \dots, r_n + ta) = \text{const.}$ Durch Differentiation an der Stelle $t = 0$ folgt

$$\left\langle \sum_{j=1}^n \frac{\partial U}{\partial r_j}, a \right\rangle = \sum_{j=1}^n \left\langle \frac{\partial U}{\partial r_j}, a \right\rangle = \left. \frac{d\tilde{U}}{dt} \right|_{t=0} = 0.$$

Da a beliebig war, erhält man

$$\sum_{j=1}^n \frac{\partial U}{\partial r_j} = 0.$$

Nun ist aber

$$\frac{d}{dt} \left(\sum_j \frac{\partial L}{\partial v_j} \right) = \sum_j \frac{d}{dt} \left(\frac{\partial L}{\partial v_j} \right) = \sum_j \frac{\partial L}{\partial r_j} = - \sum_j \frac{\partial U}{\partial r_j} = 0$$

und damit

$$\sum_{i=j}^n m_j v_j = \sum_{j=1}^n \frac{\partial L}{\partial v_j} = \text{const.}$$

3. Die *Isotropie* des Raumes bedeutet schließlich, daß $U(r_1, \dots, r_n) = U(Ar_1, \dots, Ar_n)$ für alle (speziellen) *orthogonalen* Matrizen A gilt. Wir wählen also z. B. die *Drehmatrix*

$$A = A_\alpha := \begin{pmatrix} \cos \alpha & -\sin \alpha & 0 \\ \sin \alpha & \cos \alpha & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

und bilden $\tilde{U} = \tilde{U}(\alpha) = U(x_1 \cos \alpha - y_1 \sin \alpha, x_1 \sin \alpha + y_1 \cos \alpha, z_1, \dots) = \text{const.}$ Hieraus folgt durch Differentiation nach α :

$$\begin{aligned} 0 = \frac{d\tilde{U}}{d\alpha} &= \sum_j \frac{\partial U}{\partial x_j} (x_1 \cos \alpha - \dots) \cdot [-x_j \sin \alpha - y_j \cos \alpha] \\ &\quad + \sum_j \frac{\partial U}{\partial y_j} (x_1 \cos \alpha - \dots) \cdot [x_j \cos \alpha - y_j \sin \alpha]. \end{aligned}$$

Läßt man hier α gegen Null gehen, so erhält man

$$\sum_j \left(y_j \frac{\partial U}{\partial x_j} - x_j \frac{\partial U}{\partial y_j} \right) = 0.$$

Nun ist der (*Gesamt-*) *Drehimpuls* des Systems definiert durch

$$\sum_{j=1}^n r_j \times m_j v_j,$$

wobei das *Kreuzprodukt* für zwei Vektoren $a := {}^t(a_1, a_2, a_3)$, $b = {}^t(b_1, b_2, b_3) \in \mathbb{R}^3$ formal definiert werden kann als

$$a \times b = \begin{vmatrix} a_1 & b_1 & e_1 \\ a_2 & b_2 & e_2 \\ a_3 & b_3 & e_3 \end{vmatrix} := {}^t(a_2 b_3 - a_3 b_2, a_3 b_1 - a_1 b_3, a_1 b_2 - a_2 b_1).^{42}$$

Erhaltung des Drehimpulses bedeutet natürlich die Erhaltung jeder einzelnen *Komponente* dieses Vektors. Per definitionem ist z. B. die dritte Komponente gegeben durch

$$\sum_j m_j (x_j \dot{y}_j - y_j \dot{x}_j),$$

so daß Differenzieren nach der Zeit wie gewünscht

$$\left(\sum_j m_j (x_j \dot{y}_j - y_j \dot{x}_j) \right)' = \sum_j m_j (x_j \ddot{y}_j - y_j \ddot{x}_j) = \sum_j \left(y_j \frac{\partial U}{\partial x_j} - x_j \frac{\partial U}{\partial y_j} \right) = 0$$

ergibt.

Mit diesen Vorkenntnissen wollen wir jetzt die berühmten KEPLERSchen *Gesetze* aus den NEWTONSchen *Bewegungsgleichungen* in einem Schwerfeld ableiten.

Ein Massenteilchen (mit Masse m) bewege sich in einem *Zentralfeld* mit Potential U (z. B. in dem Gravitationsfeld einer anderen Masse $M \gg m$, die im Nullpunkt ruhend angenommen wird, also z. B. ein Planet im Gravitationsfeld der Sonne oder das System Erde–Mond). Zentralfeld bedeutet, daß das Potential nur von dem *Abstand* r ⁴³ des Teilchens vom Ursprung abhängt: $U(x) = U(r)$, $r = \|x\|$. Setzen wir

$$x = x(t) = (x_1(t), x_2(t), x_3(t))$$

für den Ortsvektor des Teilchens, so lauten die Newtonschen Bewegungsgleichungen

$$m \ddot{x}(t) = m \frac{d^2 x}{dt^2}(t) = -\text{grad}_x U(x(t)),$$

oder ausgeschrieben:

$$m \ddot{x}_1 = -\frac{\partial U}{\partial x_1}(x_1, x_2, x_3), \quad m \ddot{x}_2 = -\frac{\partial U}{\partial x_2}(x_1, x_2, x_3), \quad m \ddot{x}_3 = -\frac{\partial U}{\partial x_3}(x_1, x_2, x_3).$$

Dies ist ein System von drei Differentialgleichungen zweiter Ordnung und entspricht daher einem System von sechs Gleichungen erster Ordnung. Wir wissen dann, daß, falls U hinreichend „gute“ Eigenschaften hat, sechs Größen (z. B. die Orts- und Geschwindigkeitskomponenten des Teilchens zu

⁴²Bekanntlich steht das Kreuzprodukt senkrecht auf a und b und seine Länge ist gleich dem Flächeninhalt des von a und b aufgespannten Parallelogramms. Insbesondere ist stets $a \times a = 0$.

⁴³Im Gegensatz zu den vorhergehenden Überlegungen bezeichnet r jetzt keinen Ortsvektor, sondern seine Länge.

einer Anfangszeit t_0) die Bewegung vollständig bestimmen.

Die *Erhaltungssätze der klassischen Mechanik* (siehe oben) liefern die Möglichkeit, die Anzahl der Gleichungen zu reduzieren und sie evtl. völlig zu lösen. Da die Gleichungen nach Voraussetzung invariant unter der Gruppe $SO(3, \mathbb{R})$ sind, gilt der *Drehimpulserhaltungssatz*:

$$x(t) \times m \dot{x}(t) = c = \text{const.}$$

Wir zeigen dies noch einmal direkt: Für zwei differenzierbare vektorwertige Funktionen $u(t), v(t)$ gilt nach Definition des Kreuzproduktes stets

$$(u(t) \times v(t))' = \dot{u}(t) \times v(t) + u(t) \times \dot{v}(t)$$

und damit

$$(x \times m \dot{x})' = \dot{x} \times m \dot{x} + x \times m \ddot{x} = -x \times \text{grad}_x U = -\frac{1}{r} \frac{dU}{dr} x \times x = 0,$$

da nach Voraussetzung

$$\text{grad}_x U = \frac{dU}{dr} \text{grad}_x r \quad \text{und} \quad \text{grad}_x r = \frac{1}{r} x.$$

Wir unterscheiden nun zwei Fälle:

1. $c = 0$; dann müssen $x(t)$ und $\dot{x}(t)$ für alle t linear abhängig sein. Setzen wir $x(t) \neq 0$ für alle t voraus, so heißt dies

$$\dot{x}(t) = a(t) \cdot x(t), \quad a: \mathbb{R} \rightarrow \mathbb{R} \text{ stetig,}$$

d. h. $\dot{x}_i(t) = a(t) x_i(t)$ und damit

$$x(t) = \exp\left(\int_{t_0}^t a(\tau) d\tau\right) x(t_0).$$

Insbesondere bewegt sich der Massenpunkt auf „Kollisionskurs“, d. h. auf einer Geraden durch den Nullpunkt.

2. $c \neq 0$. Dann steht $x(t)$ stets senkrecht auf c . Wegen der Symmetrie des Problems können wir $c = \lambda e_3$ annehmen, so daß die Bewegung in der (x_1, x_2) -Ebene stattfindet. Wir führen in dieser Ebene zweckmäßigerweise Polarkoordinaten ein:

$$x = x_1 = r \cos \varphi, \quad y = x_2 = r \sin \varphi,$$

wobei stets $r \neq 0$ sein muß (wegen $c \neq 0$). Es ist dann

$$\begin{aligned} \lambda &= m(x \dot{y} - y \dot{x}) \\ &= m r (\cos \varphi (\dot{r} \sin \varphi + r \dot{\varphi} \cos \varphi) - \sin \varphi (\dot{r} \cos \varphi - r \dot{\varphi} \sin \varphi)) \\ &= m r^2 \dot{\varphi}. \end{aligned}$$

Nun ist das Flächenelement ΔF in Polarkoordinaten gegeben⁴⁴ durch

$$\Delta F = \frac{1}{2} r(t) \cdot (r(t) \Delta \varphi) = \frac{1}{2} r^2(t) \cdot \dot{\varphi}(t) \Delta t,$$

so daß also die von $x(t)$ zwischen den Zeiten t_0 und t_1 überstrichene Fläche gleich dem Integral

$$F_{t_0, t_1} = \int_{t_0}^{t_1} \frac{1}{2} r^2(t) \dot{\varphi}(t) dt = \frac{\lambda}{2m} (t_1 - t_0)$$

⁴⁴Mathematisch exakt begründen können wir dies erst, wenn wir die *Transformationsformel* für Doppelintegrale zur Verfügung haben.

ist. Wir haben somit den folgenden Satz abgeleitet, der in jedem beliebigen Zentralfeld gültig ist:

2. Keplersches Gesetz. *Der Ortsvektor eines Planeten überstreicht in gleichen Zeiten gleiche Flächen.*

Zur Ableitung der *Bahngleichung* benutzen wir den *Energieerhaltungssatz*: $E = T + U = \text{const.}$ In unserem konkreten Fall ist

$$\begin{aligned} T &= \frac{m}{2} (\dot{x}^2(t) + \dot{y}^2(t)) \\ &= \frac{m}{2} [(\dot{r} \cos \varphi - r \dot{\varphi} \sin \varphi)^2 + (\dot{r} \sin \varphi + r \dot{\varphi} \cos \varphi)^2] \\ &= \frac{m}{2} (\dot{r}^2 + r^2 \dot{\varphi}^2) . \end{aligned}$$

Zusammen mit der *Drehimpulserhaltung* ergibt sich hieraus

$$E = \frac{m}{2} \left(\dot{r}^2 + \frac{\lambda^2 r^2}{m^2 r^4} \right) + U(r) = \frac{m}{2} \dot{r}^2 + \frac{\lambda^2}{2mr^2} + U(r)$$

und damit

$$\dot{r}(t) = \frac{dr}{dt} = \sqrt{\frac{2}{m} (E - U(r)) - \frac{\lambda^2}{m^2 r^2}} .$$

Dies ist eine Differentialgleichung mit getrennten Variablen, aus der sich $r = r(t)$ (jedenfalls prinzipiell) bestimmen läßt. Aus der Gleichung $\lambda = m r^2 \dot{\varphi}$ gewinnt man dann auch $\varphi = \varphi(t)$, also den gesamten Bewegungsvorgang des Massenpunktes in Polarkoordinaten.

Man kann versuchen, aus $r = r(t)$ und $\varphi = \varphi(t)$ die Zeit t zu eliminieren. Man würde damit (evtl.) r als Funktion von φ , $r = r(\varphi)$ bekommen, d. h. die *Bahngleichung* (ohne zu wissen, wie sich der Massenpunkt zeitlich auf dieser Bahn bewegt). Es ist in unserem Fall einfacher, φ als Funktion von r zu bestimmen:

$$\frac{d\varphi}{dr} = \frac{d\varphi}{dt} \cdot \left(\frac{dr}{dt} \right)^{-1} = \frac{\lambda}{r^2 \sqrt{2m(E - U(r)) - \lambda^2 r^{-2}}} .$$

In dem Spezialfall der *Massenanziehung* (Planet um Sonne, Mond oder Satellit um Erde, etc.) ist das Potential von einer ganz speziellen Gestalt:

$$U(r) = -\frac{\alpha}{r} , \quad r = \|x\| , \quad \alpha > 0 .$$

Dabei ist $\alpha = \gamma \cdot m M$ mit einer universellen Konstanten γ (Gravitations-Konstante). Damit ergibt sich

$$\begin{aligned} \varphi = \varphi(r) &= \lambda \int \frac{dr}{r^2 \sqrt{2m \left(E + \frac{\alpha}{r} \right) - \frac{\lambda^2}{r^2}}} \\ &= \lambda \int \frac{dr}{r^2 \sqrt{\left(2mE + \frac{\alpha^2 m^2}{\lambda^2} \right) - \left(\frac{\lambda}{r} - \frac{\alpha m}{\lambda} \right)^2}} , \end{aligned}$$

und mit der Substitution

$$\rho = \frac{\lambda}{r} - \frac{\alpha m}{\lambda} , \quad d\rho = -\frac{\lambda}{r^2} dr , \quad C = \sqrt{2mE + \frac{\alpha^2 m^2}{\lambda^2}}$$

folgt

$$\varphi = - \int \frac{d\rho}{\sqrt{C^2 - \rho^2}} = \arccos \frac{\rho}{C} .$$

Durch Umkehrung dieser Gleichung kommt dann schließlich

$$\frac{1}{C} \left(\frac{\lambda}{r} - \frac{\alpha m}{\lambda} \right) = \cos \varphi,$$

und damit wird die *Bahngleichung* bestimmt zu

$$\frac{p}{r} = 1 + e \cos \varphi \quad \text{mit} \quad p = \frac{\lambda^2}{m\alpha}, \quad e = \sqrt{1 + \frac{2E\lambda^2}{m\alpha^2}}.$$

Dies ist die tatsächlich die Gleichung einer *Ellipse*, *Parabel* oder *Hyperbel*, je nachdem, ob die *Exzentrizität* $e < 1$, $e = 1$ oder $e > 1$, d. h. die Energie $E < 0$, $E = 0$ oder $E > 0$ ist; denn aus $p = r + er \cos \varphi = \sqrt{x^2 + y^2} + ex$ folgt unmittelbar durch Quadrieren

$$x^2 + y^2 = (p - ex)^2 = p^2 - 2pex + e^2 x^2,$$

und dies ist für

$e = 1$ eine *Parabel*: $y^2 = p(p - 2x)$;

$e < 1$ eine *Ellipse* mit *Mittelpunkt*

$$\left(-\frac{pe}{1 - e^2}, 0 \right)$$

und *Halbachsen*

$$a = \frac{p}{1 - e^2}, \quad b = \frac{p}{\sqrt{1 - e^2}},$$

insbesondere ein *Kreis* im Falle $e = 0$, da Multiplikation der Gleichung

$$(1 - e^2)x^2 + y^2 + 2epx = p^2$$

mit $(1 - e^2)/p^2$ die Gleichung

$$\left(\frac{1 - e^2}{p} x + e \right)^2 + \left(\frac{\sqrt{1 - e^2}}{p} y \right)^2 = 1$$

nach sich zieht;

$e > 1$ eine *Hyperbel*, da entsprechend wie oben folgt:

$$\left(\frac{1 - e^2}{p} x + e \right)^2 - \left(\frac{\sqrt{e^2 - 1}}{p} y \right)^2 = 1.$$

Da Planetenbahnen beschränkt sind, kommt für sie nur der Fall der Ellipse⁴⁵ in Frage, den wir noch etwas näher untersuchen wollen. Hier gilt für den Abstand $2f$ der beiden Brennpunkte (bei $a > b$) die Beziehung $f^2 = a^2 - b^2$, und somit ist

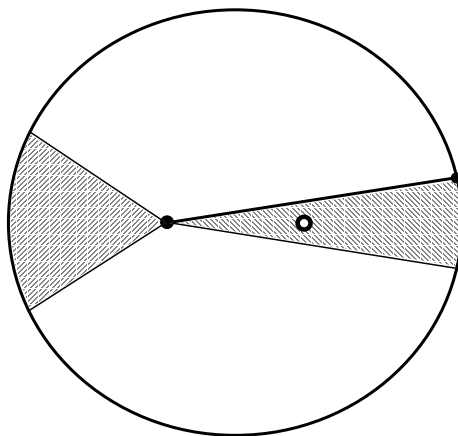
$$f^2 = p^2 \frac{1 - (1 - e^2)}{(1 - e^2)^2} = \left(\frac{pe}{1 - e^2} \right)^2.$$

Zusammenfassend ergibt sich:

1. Keplersches Gesetz. *Planeten bewegen sich auf Ellipsen, in deren einem Brennpunkt die Sonne steht.*

Das erste und zweite Keplersche Gesetz werden in der folgenden Skizze illustriert.

⁴⁵Die anderen Bahnformen werden z. B. durch *Kometen* realisiert.



Figur 21.3

Wir wollen noch eine kürzere Ableitung des 1. Keplerschen Gesetzes geben, wobei wir der Einfachheit wegen alle Konstanten gleich 1 setzen. Aus dem Bewegungsgesetz

$$\ddot{x} = -\frac{x}{\|x\|^3}$$

folgt die zeitliche Konstanz des Drehimpulsvektors:

$$J = x \times \dot{x} = \text{const.}$$

In dem speziellen Fall der Massenanziehung gibt es aber noch ein weiteres *Integral*, nämlich den sogenannten *Achsenvektor* oder *Lenzschen Vektor*

$$A = J \times \dot{x} + \frac{x}{\|x\|}.$$

(Eine theoretische Begründung für die Existenz dieser Invarianten wird bei LANDAU - LIFSCHITZ, Mechanik, Paragraph 50, Aufgabe gegeben). Wir prüfen die Behauptung einfach nach: Es ist

$$\dot{A} = J \times \ddot{x} + \frac{\dot{x}}{\|x\|} - \frac{\langle x, \dot{x} \rangle}{\|x\|^3} x$$

und

$$J \times \ddot{x} = -(x \times \dot{x}) \times \frac{x}{\|x\|^3} = \frac{\langle \dot{x}, x \rangle}{\|x\|^3} x - \frac{\langle x, x \rangle}{\|x\|^3} \dot{x}.$$

Aus der zeitlichen Konstanz des Vektors A folgt nun unmittelbar, daß sich ein Massenpunkt im Schwerefeld auf einer Bahn bewegen muß, die durch eine *quadratische Gleichung* gegeben wird. Es ist

$$\langle A, x \rangle = \langle J \times \dot{x}, x \rangle + r = \langle \dot{x} \times x, J \rangle + r = -\|J\|^2 + \|x\|,$$

woraus durch Auflösen nach $\|x\|$ und anschließendes Quadrieren die vorstehende Behauptung folgt. Da die Bahn zudem in einer Ebene verläuft, stellt sie eine *ebene Quadrik* dar, von denen in der *Analytischen Geometrie* nachgewiesen wird, daß sie genau die (evtl. ausgearteten) *Kegelschnitte* darstellen. Dies kann man auch ganz konkret nachrechnen. Man wählt die $x = x_1$ -Richtung so, daß sie mit der Richtung des Achsenvektors A übereinstimmt (daher rührt auch sein Name). Dann ist

$$\|A\| r \cdot \cos \varphi = \langle A, x \rangle = -\|J\|^2 + r \quad \text{und somit} \quad r = \frac{\|J\|^2}{1 - \|A\| \cos \varphi}.$$

Um das dritte Keplersche Gesetz ableiten zu können, benötigen wir nur das zweite Gesetz und die konkreten Parameter aus dem ersten Gesetz. Hiernach ist der Flächeninhalt der bei einem vollen Umlauf überstrichenen Fläche gleich

$$\pi a b = \pi \sqrt{p} a^{3/2}$$

und

$$\sqrt{p} = \frac{\lambda}{m} \cdot \frac{1}{\sqrt{\gamma M}} .$$

Somit gilt nach dem 2. Gesetz für die *Umlaufszeit* T :

$$T = \frac{2m}{\lambda} \pi \frac{\lambda}{m} \frac{1}{\sqrt{\gamma M}} a^{3/2} ,$$

und folglich ist

$$\frac{T^2}{a^3} = \frac{4\pi^2}{\gamma M}$$

eine Konstante, die nicht von dem einzelnen Planeten abhängt. Dies ist der Inhalt des dritten Keplerschen Gesetzes:

3. Keplersches Gesetz. *Das Verhältnis der Quadrate der Umlaufzeiten zu den Kuben der großen Halbachsen der Umlaufbahnen ist für alle Planeten konstant.*

Bemerkung. NEWTON hat in der Tat zuerst das Massenanziehungsgesetz aus den Keplerschen Gesetzen abgeleitet. Die entsprechenden Überlegungen wollen wir an dieser Stelle einfügen. Nach dem ersten Gesetz verläuft die Bewegung insbesondere in einer Ebene; wir nehmen an, daß dies die (x, y) -Ebene sei, in deren Ursprung die Sonne stehe. Weiter seien (r, φ) *Polarkoordinaten*.

Wir müssen zunächst den Beschleunigungsvektor in diesen Koordinaten ausrechnen. Durch zweimalige Differentiation der Gleichungen $x = r \cos \varphi$, $y = r \sin \varphi$ erhält man sofort

$$\begin{aligned} \ddot{x} &= (\ddot{r} - r\dot{\varphi}^2) \cos \varphi - (2\dot{r}\dot{\varphi} + r\ddot{\varphi}) \sin \varphi \\ \ddot{y} &= (\ddot{r} - r\dot{\varphi}^2) \sin \varphi + (2\dot{r}\dot{\varphi} + r\ddot{\varphi}) \cos \varphi . \end{aligned}$$

Dies impliziert (in selbstverständlicher Schreibweise)

$$(\ddot{r})_r = \ddot{r} - r\dot{\varphi}^2 , \quad (\ddot{r})_\varphi = 2\dot{r}\dot{\varphi} + r\ddot{\varphi} .$$

Das zweite Keplersche Gesetz („Flächensatz“) besagt nun aber, daß $r^2\dot{\varphi}$ *konstant* (gleich F) ist. Differentiation dieses Gesetzes führt somit zu

$$(\ddot{r})_\varphi = 0 ,$$

oder, in anderen Worten, zu der Erkenntnis, daß *die Beschleunigung stets die Richtung der Verbindungslinie Sonne-Planet besitzt*. Es liegt also eine sogenannte *Zentralbeschleunigung* vor, und wir können festhalten, daß *das Vorhandensein einer Zentralbeschleunigung mit dem Flächensatz äquivalent ist*.

Schreiben wir wieder die Ellipsengleichung in der Form

$$\frac{1}{r} = \frac{1}{k} (1 + \varepsilon \cos \varphi) , \quad k = a(\varepsilon^2 - 1) ,$$

so gewinnt man durch Differentiation und mit Hilfe des Flächensatzes

$$\dot{r} = \frac{\varepsilon}{k} \sin \varphi \cdot r^2 \dot{\varphi} = \frac{\varepsilon F}{k} \sin \varphi \quad \text{und} \quad \ddot{r} = \frac{\varepsilon F}{k} \cos \varphi \cdot \dot{\varphi} = \frac{\varepsilon F^2}{k r^2} \cos \varphi ,$$

und damit wird der Radialanteil des Beschleunigungsvektors gleich

$$\ddot{r} - r \dot{\varphi}^2 = \frac{F^2}{r^2} \left(\frac{\varepsilon}{k} \cos \varphi - \frac{1}{r} \right) = - \frac{F^2}{k r^2} .$$

Infolgedessen erhalten wir bei jedem einzelnen Planeten die Zentralbeschleunigung $-F^2/k r^2$.

Die Größe F^2/k hängt zunächst von jedem einzelnen Planeten ab. Man rechnet aber leicht nach, daß

$$\frac{F^2}{k} = \frac{4 \pi^2 a^3}{T^2} ,$$

und nach dem 3. Keplerschen Gesetz ist dies die (universelle) Konstante γM .

Zum Schluß wollen wir noch begründen, wodurch die Annahme des Ruhens der Sonne im Ursprung gerechtfertigt ist. Wir behandeln dazu ansatzweise das allgemeine *Zweikörperproblem* von zwei Massen m_1 und m_2 , die sich unter Einfluß des Potentials

$$U = U(r_1, r_2)$$

bewegen. Ferner möge das Potential U nur vom Abstand beider Massenpunkte abhängen:

$$U = U(r) , \quad r = \|r_1 - r_2\| .$$

Die Lagrange-Funktion des Systems lautet

$$L = \frac{m_1}{2} \|v_1\|^2 + \frac{m_2}{2} \|v_2\|^2 - U(r) .$$

Der Impulserhaltungssatz liefert für den *Schwerpunkt*

$$s = \frac{m_1 r_1 + m_2 r_2}{(m_1 + m_2)}$$

das Verschwinden der zweiten Ableitung; also ist \dot{s} konstant, und wir können s als Ursprung eines *Inertialsystems* wählen (das sich per definitionem geradlinig und gleichförmig im Raum bewegt).

Führen wir jetzt noch den Abstandsvektor $r = r_1 - r_2$ ein, so folgt mit $m_1 r_1 + m_2 r_2 = 0$ und $v = v_1 - v_2$ sofort

$$r_1 = \frac{m_2}{m_1 + m_2} r , \quad r_2 = - \frac{m_1}{m_1 + m_2} r .$$

Einsetzen in die Lagrange-Funktion ergibt dann

$$(+) \quad L = \frac{m}{2} \|v\|^2 - U(r)$$

mit der sogenannten *reduzierten* Masse

$$m = \frac{m_1 m_2}{m_1 + m_2} .$$

Ist nun $m_2 \gg m_1$, so ist $r_1 \sim r$, $r_2 \sim 0$, $v_1 \sim v$, $v_2 \sim 0$ und $m_1 \sim m$. Wir können daher den schwereren Körper als im Ursprung ruhend annehmen und Lösungen $r = r(t)$ der zu (+) gehörenden Bewegungsgleichungen als Bahngleichungen des leichteren Körpers interpretieren.

22 Lineare Differentialgleichungen und -Gleichungssysteme

Es sei $I \subset \mathbb{R}$ ein (beliebiges) Intervall, und $a_{jk}, b_j : I \rightarrow \mathbb{R}$ seien stetige Funktionen, $1 \leq j, k \leq n$. Ein System linearer *Differentialgleichungen erster Ordnung* ist von der Form

$$(*) \quad Y' = A(x)Y + b(x),$$

wobei $A(x)$ die Matrix

$$\begin{pmatrix} a_{11}(x) & \cdots & a_{1n}(x) \\ \vdots & & \vdots \\ a_{n1}(x) & \cdots & a_{nn}(x) \end{pmatrix}$$

und $b(x)$ den Spaltenvektor

$$\begin{pmatrix} b_1(x) \\ \vdots \\ b_n(x) \end{pmatrix}$$

bezeichnet. Ausgeschrieben kann dieses System auch in der Form

$$y'_j = \sum_{k=1}^n a_{jk}(x) y_k + b_j(x), \quad j = 1, \dots, n,$$

notiert werden. Eine *Lösung* von $(*)$ besteht dann aus einem n -tupel von differenzierbaren Funktionen

$$\varphi(x) = \begin{pmatrix} \varphi_1(x) \\ \vdots \\ \varphi_n(x) \end{pmatrix} \quad \text{mit} \quad \varphi'(x) := \begin{pmatrix} \varphi'_1(x) \\ \vdots \\ \varphi'_n(x) \end{pmatrix} = A(x) \cdot \varphi(x) + b(x).$$

In Hinblick auf die zentralen Existenz- und Eindeutigkeitsätze für Lösungen von Systemen von Differentialgleichungen ist, wie wir wissen, das Bestehen einer lokalen *LIPSCHITZ-Bedingung* erforderlich. Wir können in der vorliegenden speziellen Situation sogar noch mehr beweisen. Man wähle ein beliebiges Kompaktum $J \subset I$ und setze

$$F(x, y) = A(x)y + b(x), \quad L = \sup_{x \in J} \{ \|A(x)\| \} < \infty.$$

Hierbei wird für eine beliebige quadratische Matrix $A = (a_{j,k})$ die *Operator-Norm* $\|A\|$ wie üblich bzgl. einer festgelegten Norm in \mathbb{R}^n definiert durch

$$\|A\| = \sup_{\|y\| \leq 1} \|Ay\|.$$

Aus der Tatsache der Äquivalenz aller Normen auf $M(n \times n, \mathbb{R}) \simeq \mathbb{R}^{n^2}$ folgt, daß jede Operatornorm $\|A\|$ stetig in den Einträgen der Matrix A und folglich L tatsächlich endlich ist. Speziell für die *Zeilensummen-Norm*, die der Supremumsnorm auf \mathbb{R}^n zugeordnet ist, hat man die konkrete Abschätzung

$$\|A\|_\infty := \max_{j,k} |a_{jk}| \leq \|A\|.$$

Damit ergibt sich für beliebige $y, \tilde{y} \in \mathbb{R}^n$:

$$\|F(x, y) - F(x, \tilde{y})\| = \|A(x)(y - \tilde{y})\| \leq L \cdot \|y - \tilde{y}\|, \quad x \in J.$$

Mit anderen Worten: *Die Funktion $F(x, y)$ ist Lipschitz-stetig bzgl. y auf $J \times \mathbb{R}^n$.*

Insbesondere ist die Lipschitz-Bedingung *lokal* erfüllt, und dies zieht unmittelbar die lokale Existenz und die Eindeutigkeit der Lösungen nach sich. Wir zeigen die Eindeutigkeitsaussage anschließend noch einmal für lineare Differentialgleichungssysteme und beweisen, daß Lösungen stets auf ganz I existieren.

Satz 22.1 Es seien $\varphi, \psi : I \rightarrow \mathbb{R}^n$ zwei Lösungen des Systems (*), die an einer Stelle in dem Intervall I übereinstimmen. Dann ist $\varphi = \psi$.

Beweis. Wegen der Stetigkeit der Lösungen genügt es zu zeigen, daß solche lokal um einen Punkt $a \in I$ übereinstimmen, wenn dies nur an der Stelle a richtig ist. Man setzt

$$K_\delta := \sup_{\substack{x \in I \\ |x-a| < \delta}} \{ \|\varphi(x) - \psi(x)\| \},$$

und erhält wegen

$$\begin{aligned} \|\varphi(x) - \psi(x)\| &= \left\| \int_a^x (f(t, \varphi(t)) - f(t, \psi(t))) dt \right\| = \left\| \int_a^x A(t) (\varphi(t) - \psi(t)) dt \right\| \\ &\leq L \left| \int_a^x \|\varphi(t) - \psi(t)\| dt \right| \leq K_\delta \cdot L \cdot |x - a| \end{aligned}$$

für alle $x \in I$ mit $|x - a| < \delta$ durch Supremumbildung sofort

$$K_\delta \leq K_\delta L \delta.$$

Wählt man schließlich $\delta > 0$ so klein, daß $2\delta L < 1$, so folgt $K_\delta = 0$. \square

Wir können jetzt den entscheidenden *globalen Existenzsatz* für Lösungen formulieren:

Satz 22.2 Zu jedem $a \in I$ und jedem $c \in \mathbb{R}^n$ gibt es (genau) eine globale Lösung $\varphi : I \rightarrow \mathbb{R}^n$ mit $\varphi(a) = c$.

Beweis. I ist die Vereinigung aller kompakten Intervall $J \subset I$, die a enthalten. Wegen des Eindeutigkeitsatzes für Lösungen genügt es dann, nur diese *kompakten* Teilintervalle zu betrachten. Aus diesem Grunde können wir von vornherein annehmen, daß das Intervall I selbst schon kompakt ist und wir deshalb wie oben eine *globale* Lipschitz-Konstante L auf ganz $I \times \mathbb{R}^n$ zur Verfügung haben.

Wir konstruieren nun eine Lösung gemäß dem *Iterationsverfahren* von PICARD und LINDELÖF. Setzen wir also

$$\varphi_0(x) = c, \quad \varphi_1(x) = c + \int_a^x f(t, \varphi_0(t)) dt$$

und allgemein

$$\varphi_{k+1}(x) = c + \int_a^x f(t, \varphi_k(t)) dt,$$

und definieren wir dann noch

$$K := \sup \{ \|\varphi_1(x) - \varphi_0(x)\| \},$$

so erhalten wir wie im Beweis des Eindeutigkeitsatzes

$$\|\varphi_2(x) - \varphi_1(x)\| = \left\| \int_a^x A(t) (\varphi_1(t) - \varphi_0(t)) dt \right\| \leq K \cdot L \cdot |x - a|$$

und per Induktion

$$\|\varphi_{k+1}(x) - \varphi_k(x)\| \leq K \frac{L^k |x - a|^k}{k!},$$

so daß der Limes

$$\varphi(x) = \lim_{k \rightarrow \infty} \varphi_k(x)$$

gleichmäßig auf I existiert. Wegen

$$\varphi(x) = \lim_{k \rightarrow \infty} \varphi_{k+1}(x) = c + \int_a^x \lim_{k \rightarrow \infty} f(\xi, \varphi_k(\xi)) d\xi = c + \int_a^x f(\xi, \varphi(\xi)) d\xi$$

ist $\varphi'(x) = f(x, \varphi(x))$ und $\varphi(a) = c$ und damit φ die gesuchte Lösung. \square

Bemerkung und Beispiel. Unter geeigneten Voraussetzungen kann man die *globale* Existenz von Lösungen auch für nichtlineare Differentialgleichungen begründen. Ist z. B. die Differentialgleichung

$$y' = x |\sin xy|$$

vorgelegt, so besitzt diese zu vorgegebenem Anfangswert bei $a \in \mathbb{R}$ genau eine auf ganz \mathbb{R} definierte Lösung. Die Lösungen $y = \varphi(x)$ mit $\varphi(a) \neq 0$ verschwinden nirgends, da $\varphi_0(x) \equiv 0$ offensichtlich die Differentialgleichung löst. Der Grund für diese Behauptungen liegt in der Abschätzung

$$\begin{aligned} \left| |\sin xy_1| - |\sin xy_2| \right| &\leq |\sin xy_1 - \sin xy_2| = 2 \left| \cos x \frac{y_1 + y_2}{2} \right| \left| \sin x \frac{y_1 - y_2}{2} \right| \\ &\leq 2 \left| \sin x \frac{y_1 - y_2}{2} \right| \leq |x| |y_1 - y_2|. \end{aligned}$$

Somit liegt wie bei linearen Differentialgleichungen auf $I \times \mathbb{R}$ eine *globale* Lipschitz-Bedingung vor, wenn $I \subset \mathbb{R}$ ein *kompaktes* Intervall ist, und daraus folgt wie oben die Tatsache der Existenz von Lösungen auf ganz \mathbb{R} und ihre eindeutige Bestimmtheit durch den Anfangswert.

Definition. Ist $b = 0$, so nennt man das System $Y' = A(x)Y$ wie im Falle von linearen Gleichungssystemen *homogen*. Der Lösungsraum ist dann ein Untervektorraum des reellen Vektorraums der n -tupel stetig differenzierbarer Funktionen auf I .

Satz 22.3 *Der Lösungsraum des homogenen Systems $Y' = A(x)Y$ ist ein n -dimensionaler Vektorraum. Für jeden Punkt $a \in I$ ist die Abbildung, die einer Lösung φ ihren Wert $\varphi(a)$ zuordnet, ein Vektorraum-Isomorphismus.*

Beweis. Nach dem vorigen Satz existieren Lösungen ψ_1, \dots, ψ_n auf I mit

$$\psi_j(a) = e_j = {}^t(0, \dots, 0, \overbrace{1}^j, 0, \dots, 0), \quad j = 1, \dots, n,$$

für einen fest gewählten Punkt $a \in I$. Ist nun φ eine beliebige Lösung, so setzt man $\varphi(a) = c = \sum c_j e_j = \sum c_j \psi_j(a)$ und bildet

$$\psi := \varphi - \sum_{j=1}^n c_j \psi_j.$$

Daraus ergibt sich unmittelbar

$$\psi' = \left(\varphi - \sum_{j=1}^n c_j \psi_j \right)' = \varphi' - \sum_{j=1}^n c_j \psi_j' = A \left(\varphi - \sum_{j=1}^n c_j \psi_j \right) = A\psi.$$

Also ist ψ eine Lösung, die in a verschwindet. Infolgedessen muß nach dem Eindeutigkeitsatz ψ mit der trivialen Lösung $\varphi_0 = 0$ übereinstimmen, also $\varphi = \sum_{j=1}^n c_j \psi_j$ gelten.

Damit ist der Lösungsraum als ein \mathbb{R} -Vektorraum der Dimension $\leq n$ erkannt. Ist aber $\sum a_j \psi_j = 0$, so auch $0 = \sum a_j \psi_j(a) = \sum a_j e_j$ und folglich $a_j = 0$ für alle j . \square

Sind uns allgemein n Lösungen

$$\psi_j = \begin{pmatrix} \psi_{1j} \\ \vdots \\ \psi_{nj} \end{pmatrix}, \quad j = 1, \dots, n,$$

des homogenen Systems gegeben, so bilden wir die $n \times n$ -Matrix

$$\Psi = \begin{pmatrix} \psi_{11} & \cdots & \psi_{1n} \\ \vdots & & \vdots \\ \psi_{n1} & \cdots & \psi_{nn} \end{pmatrix},$$

für die wir

$$\Psi' = \begin{pmatrix} \psi'_{11} & \cdots & \psi'_{1n} \\ \vdots & & \vdots \\ \psi'_{n1} & \cdots & \psi'_{nn} \end{pmatrix}$$

setzen und sofort die Beziehung

$$\Psi' = A \Psi$$

erhalten.

Definition. Eine $n \times n$ -Matrix $\Psi = \Psi(x)$ mit $\Psi'(x) = A(x)\Psi(x)$, $x \in I$, heißt ein *Lösungsfundamentalsystem* des homogenen Systems $Y' = A(x)Y$, falls die Spalten von Ψ eine Basis des Lösungsraumes des homogenen Systems bilden.

Diese Bedingung ist nach der obigen Diskussion äquivalent dazu, daß die Spalten von Ψ mindestens an einer Stelle a linear unabhängige Vektoren in \mathbb{R}^n sind; oder anders ausgedrückt: es reicht, daß

$$\det \Psi(a) \neq 0 \quad \text{für ein } a \in I.$$

Satz 22.4 *Es sei Ψ eine $n \times n$ -Matrix mit Einträgen in $C^1(I, \mathbb{R})$, die der Differentialgleichung $\Psi' = A\Psi$ genügt. Dann sind äquivalent:*

- i) Ψ ist ein Lösungsfundamentalsystem der Gleichung $Y' = A(x)Y$;
- ii) es gibt ein $a \in I$, so daß $\det \Psi(a) \neq 0$;
- iii) für alle $x \in I$ ist $\det \Psi(x) \neq 0$.

Kennt man ein Fundamentalsystem des homogenen Gleichungssystems, so kann man auch jedes zugehörige *inhomogene* System ohne größere Schwierigkeiten lösen. Der folgende Satz ist aus der Theorie der linearen Gleichungssysteme wohlvertraut und wird wortwörtlich wie dort bewiesen.

Satz 22.5 *Sei ψ_0 eine spezielle Lösung Des Systems $Y' = A(x)Y + b(x)$. Dann ist jede andere Lösung von der Form $\psi_0 + \psi$, wobei ψ Lösung des homogenen Gleichungssystems ist, und jede(s) Funktion(ensystem) der Gestalt $\psi_0 + \psi$ löst das inhomogene System.*

Haben wir also das homogene System vollständig gelöst, so brauchen wir nur noch eine einzige Lösung des inhomogenen Systems zu finden. Dies geschieht, wie im Falle einer einzigen linearen Differentialgleichung, mit der sogenannten *Variation der Konstanten* (was eigentlich eine *contradictio in adjecto* ist):

Wir machen mit einer differenzierbaren Abbildung $u : I \rightarrow \mathbb{R}^n$ den Ansatz

$$\psi_0(x) = \Psi(x) \cdot u(x) = \begin{pmatrix} \psi_{11}(x) & \cdots & \psi_{1n}(x) \\ \vdots & & \vdots \\ \psi_{n1}(x) & \cdots & \psi_{nn}(x) \end{pmatrix} \begin{pmatrix} u_1(x) \\ \vdots \\ u_n(x) \end{pmatrix}$$

und versuchen, die „variierende Konstanten“ u so zu bestimmen, daß ψ_0 das inhomogene System $Y' = A(x)Y + b(x)$ löst. Dies ist äquivalent zu

$$A(\Psi u) + b = A\psi_0 + b = \psi_0' = \Psi u' + \Psi' u = \Psi u' + (A\Psi)u,$$

also zu $\Psi u' = b$. Nun ist die Matrix $\Psi(x)$ nach Satz 4 an jeder Stelle $x \in I$ invertierbar. Damit ist notwendig und hinreichend, daß $u(x)$ die Gleichung

$$u'(x) = \Psi^{-1}(x) \cdot b(x)$$

erfüllt. Mit Hilfe der *CRAMERSchen Regel* erkennt man aber sofort, daß die inverse Matrix ebenfalls stetig differenzierbare Einträge besitzt, so daß man diese Gleichung einfach hochintegrieren kann. – Zur Illustration geben wir ein

Beispiel. Für das System

$$\begin{cases} y_1' = -y_2, \\ y_2' = y_1 + x \end{cases}$$

haben wir

$$\Psi(x) = \begin{pmatrix} \cos x & -\sin x \\ \sin x & \cos x \end{pmatrix}$$

mit

$$\det \Psi \equiv 1 \quad \text{und} \quad \Psi^{-1}(x) = \begin{pmatrix} \cos x & \sin x \\ -\sin x & \cos x \end{pmatrix}.$$

Somit folgt

$$u'(x) = \begin{pmatrix} u_1'(x) \\ u_2'(x) \end{pmatrix} = \begin{pmatrix} \cos x & \sin x \\ -\sin x & \cos x \end{pmatrix} \begin{pmatrix} 0 \\ x \end{pmatrix} = \begin{pmatrix} x \sin x \\ x \cos x \end{pmatrix}$$

und

$$u(x) = \begin{pmatrix} \sin x - x \cos x \\ \cos x + x \sin x \end{pmatrix}.$$

Dies impliziert sofort

$$\psi_0(x) = \Psi(x) u(x) = \begin{pmatrix} -x \\ 1 \end{pmatrix},$$

so daß die allgemeine Lösung die Gestalt

$$\psi(x) = \begin{pmatrix} a \cos x - b \sin x - x \\ a \sin x + b \cos x + 1 \end{pmatrix}, \quad a, b \in \mathbb{R} \text{ beliebig,}$$

hat.

Die vorstehenden Ergebnisse lassen sich ohne weiteres übertragen auf lineare Differentialgleichungen n -ter Ordnung:

$$y^{(n)} + a_{n-1}(x)y^{(n-1)} + \dots + a_0(x)y = b(x).$$

Ordnet man nämlich einer Lösung φ der vorliegenden Gleichung wie im vorangegangenen Kapitel das n -tupel ${}^t(\varphi, \varphi', \dots, \varphi^{(n-1)})$ zu, so erfüllt dieses ein (spezielles) System von gewöhnlichen Differentialgleichungen 1. Ordnung, woraus man sofort schließt:

Satz 22.6 Der Lösungsraum einer gewöhnlichen linearen Differentialgleichung n -ter Ordnung ist ein reeller n -dimensionaler affiner Raum (bzw. ein Vektorraum derselben Dimension im homogenen Fall). n Lösungen $\varphi_1, \dots, \varphi_n$ der homogenen Gleichung spannen den Lösungsraum genau dann auf, wenn es ein $a \in I$ gibt, so daß die sogenannte Wronski-Determinante

$$W(a) = \det \begin{pmatrix} \varphi_1(a) & \cdots & \varphi_n(a) \\ \varphi_1'(a) & \cdots & \varphi_n'(a) \\ \vdots & & \vdots \\ \varphi_1^{(n-1)}(a) & \cdots & \varphi_n^{(n-1)}(a) \end{pmatrix} \neq 0$$

ist. Partikuläre Lösungen der inhomogenen Gleichung erhält man durch Variation der Konstanten.

Beispiel. Die Differentialgleichung $y'' - \frac{1}{2x}y' + \frac{1}{2x^2}y = 0$, $I = \mathbb{R}_{>0}$, wird gelöst durch die Funktionen $\varphi_1(x) = x$, $\varphi_2(x) = \sqrt{x}$. Hier ist

$$W(x) = \det \begin{pmatrix} x & \sqrt{x} \\ 1 & \frac{1}{2\sqrt{x}} \end{pmatrix} = \frac{\sqrt{x}}{2} - \sqrt{x} = -\frac{1}{2}\sqrt{x} \neq 0.$$

Die allgemeine Lösung ist also von der Form

$$ax + b\sqrt{x}, \quad x \in \mathbb{R}_{>0}.$$

Bemerkung. Alle speziellen Funktionen in einer reellen Veränderlichen, die in der mathematischen Physik auftreten, sind Lösungen von homogenen linearen Differentialgleichungen zweiter Ordnung (z. B. der LEGENDRE-, HERMITE- oder LAGUERRE-Gleichung), also allgemein von Gleichungen des Typs

$$y'' + a(x)y' + b(x)y = 0.$$

Wir wollen noch ein Verfahren besprechen, wie man in diesem Fall aus einer gegebenen Lösung eine zweite, unabhängige gewinnen kann: Sei φ eine Lösung, die nirgends verschwindet. Dann setzt man $\psi = \varphi \cdot u$ und rechnet aus, daß ψ genau dann die obige Differentialgleichung löst, wenn

$$u'' + \left(2\frac{\varphi'}{\varphi} + a\right)u' = 0$$

erfüllt ist. Dies ist nun eine lineare Differentialgleichung 1. Ordnung für die unbekannte Funktion u' , die man leicht lösen kann. Man nennt dieses Verfahren deshalb auch *Reduktion der Ordnung*.

Beispiel. Wir betrachten die Legendresche Differentialgleichung für $n = 1$:

$$y'' - \frac{2x}{1-x^2}y' + \frac{2}{1-x^2}y = 0, \quad |x| < 1.$$

Eine Lösung ist offensichtlich $\varphi(x) = x$. Diese verschwindet aber in 0, so daß das obige Verfahren z. B. nur in dem Bereich $0 < x < 1$ sinnvoll ist. Mit dem Ansatz $\psi(x) := x \cdot u(x)$ ergibt sich die Differentialgleichung

$$u'' + \left(2\frac{1}{x} - \frac{2x}{1-x^2}\right)u' = 0,$$

aus der man u' zu

$$\begin{aligned} u'(x) &= \exp(-2 \log x - \log(1-x^2)) \\ &= \frac{1}{x^2(1-x^2)} = \frac{1}{x^2} + \frac{1}{2} \left(\frac{1}{1-x} + \frac{1}{1+x} \right) \end{aligned}$$

berechnet. Damit ist

$$u(x) = -\frac{1}{x} + \frac{1}{2} \log \frac{1+x}{1-x},$$

so daß die zweite Lösung durch

$$\psi(x) = \frac{x}{2} \log \frac{1+x}{1-x} - 1$$

gegeben wird, von der man im nachhinein aber feststellt, daß sie tatsächlich für alle x mit $|x| < 1$ erklärt ist und die Differentialgleichung löst.

Hier noch ein Beispiel aus der Theorie der gewöhnlichen Differentialgleichungen zweiter Ordnung.

Satz 22.7 Für die Lösungen einer homogenen linearen Differentialgleichung zweiter Ordnung mit stetigen Koeffizienten auf einem Intervall I ist folgendes richtig :

- a) Jede von Null verschiedene Lösung hat nur einfache Nullstellen, und die Menge ihrer Nullstellen besitzt keinen Häufungspunkt in I .
- b) Ist φ_1, φ_2 ein Fundamentalsystem von Lösungen, so liegt zwischen je zwei Nullstellen von φ_1 eine Nullstelle von φ_2 .

Beweis. a) Lösungen von linearen Differentialgleichungen zweiter Ordnung existieren auf dem gesamten Definitionsintervall I und sind eindeutig bestimmt durch ihren Wert und ihre Ableitung an einer Stelle $a \in I$. An einer mehrfachen Nullstelle a verschwinden diese beiden Werte. Ist die Differentialgleichung homogen, so ist die identisch verschwindende Funktion eine Lösung und damit die einzige mit dieser Eigenschaft. Häufen sich Nullstellen einer Lösung an der Stelle a , so auch die Nullstellen ihrer Ableitung wegen des Satzes von Rolle. Da die Lösung stetig differenzierbar ist, sind die Voraussetzungen des ersten Teiles in a erfüllt, und folglich ist die Lösung identisch Null.

b) Zwei Lösungen φ_1, φ_2 bilden genau dann ein Fundamentalsystem, wenn die Wronski–Determinante

$$W(x) := \det \begin{pmatrix} \varphi_1(x) & \varphi_2(x) \\ \varphi_1'(x) & \varphi_2'(x) \end{pmatrix}$$

an allen Stellen $x \in I$ von Null verschieden ist. Ist insbesondere $\varphi_1(a) = \varphi_1(b) = 0$, $a < b$, so ist φ_2 an beiden Stellen a, b von Null verschieden. Wäre nun φ_2 auch im Inneren des Intervalls $J := [a, b]$ von Null verschieden, so würde die Funktion $\psi := \varphi_1/\varphi_2$ dort existieren und differenzierbar sein. Wegen der Voraussetzung wäre dann $\psi(a) = \psi(b) = 0$ und also nach dem Satz von Rolle

$$\psi'(\xi) = \frac{\varphi_2(\xi)\varphi_1'(\xi) - \varphi_1(\xi)\varphi_2'(\xi)}{\varphi_2^2(\xi)}$$

gleich Null an einer Stelle ξ zwischen a und b im Widerspruch zur Eigenschaft von $W(x)$. \square

Wir wollen uns von nun an in diesem Kapitel mit dem speziellen Fall beschäftigen, daß die Koeffizienten in einer linearen Differentialgleichung n -ter Ordnung bzw. in einem System von linearen Differentialgleichungen erster Ordnung *Konstanten* sind. Hierbei können wir wesentliche Aussagen der linearen Algebra, wie z. B. *Normalformensätze*, direkt zur Lösung der Gleichung(en) verwenden. Da diese jedoch ihre einfachste Form im *Komplexen* annehmen, ist es sinnvoll, sich auch mit komplexwertigen Differentialgleichungen und Lösungen zu beschäftigen. Wir werden uns dabei meistens mit dem *homogenen* Fall begnügen, da auch in dieser speziellen Situation i. A. für inhomogene Systeme und Gleichungen der Ansatz der variablen Koeffizienten das Mittel der Wahl ist.

Es sei also wie bisher $I \subset \mathbb{R}$ ein (beliebiges) Intervall, aber die gegebenen Funktionen dürfen jetzt auch \mathbb{C} -wertig sein: $a_{jk}, b_j : I \rightarrow \mathbb{C}$ seien stetig, $1 \leq j, k \leq n$. Ein *komplexwertiges* System linearer Differentialgleichungen erster Ordnung ist dann von der gleichen Form wie in der reellen Situation:

$$Y' = A(x)Y + b(x),$$

und *Lösungen* bestehen aus n -tupeln von differenzierbaren Funktionen $\varphi_j : I \rightarrow \mathbb{C}$,

$$\varphi(x) = \begin{pmatrix} \varphi_1(x) \\ \vdots \\ \varphi_n(x) \end{pmatrix} \quad \text{mit} \quad \varphi'(x) := \begin{pmatrix} \varphi_1'(x) \\ \vdots \\ \varphi_n'(x) \end{pmatrix} = A(x) \cdot \varphi(x) + b(x).$$

Genau wie im reellen Fall beweist man den globalen Existenz- und Eindeigkeitssatz und vollständige Analoga zu den Sätzen 4 und 6.

Satz 22.8 *Der Lösungsraum eines Systems von n gewöhnlichen komplexen linearen Differentialgleichungen erster Ordnung ist ein komplexer n -dimensionaler affiner Raum (bzw. ein Vektorraum derselben Dimension im homogenen Fall). n Lösungen $\varphi_1, \dots, \varphi_n$ der homogenen Gleichung spannen den Lösungsraum genau dann auf, wenn es ein $a \in I$ gibt, so daß die Anfangswerte $\varphi_1(a), \dots, \varphi_n(a) \in \mathbb{C}$ über \mathbb{C} unabhängig sind. Dies ist dann auch an allen anderen Stellen in I richtig. Partikuläre Lösungen der inhomogenen Gleichung erhält man durch Variation der Konstanten.*

Ist nun $Y' = A(x)Y + b(x)$ wieder ein *reelles* System, so kann man dennoch auch von einer *komplexen* Lösung $\varphi : I \rightarrow \mathbb{C}$ sprechen. Das nächste Lemma, das wegen seiner Einfachheit keines echten Beweises bedarf, gibt erschöpfenden Aufschluß über den Zusammenhang zwischen reellen und komplexen Lösungen homogener Systeme.

Satz 22.9 *Für zwei reelle Lösungen φ_1, φ_2 des reellen Systems $Y' = A(x)Y$ ist $\varphi_1 + i\varphi_2$ eine komplexe Lösung. Umgekehrt sind mit jeder komplexen Lösung $\varphi : I \rightarrow \mathbb{C}$ die Funktionen $\operatorname{Re} \varphi : I \rightarrow \mathbb{R}$ und $\operatorname{Im} \varphi : I \rightarrow \mathbb{R}$ reelle Lösungen.*

Damit erhebt sich aber ein Problem: Hat man ein *komplexes Fundamentalsystem* $\varphi_1, \dots, \varphi_n$ eines homogenen reellen Systems gefunden, so geben diese Anlaß zu den $2n$ reellen Lösungen $\operatorname{Re} \varphi_1, \operatorname{Im} \varphi_1, \dots, \operatorname{Re} \varphi_n, \operatorname{Im} \varphi_n : I \rightarrow \mathbb{R}$, und von diesen sind mit Sicherheit n Stück überflüssig. Tatsächlich gilt stets die Aussage, daß diese $2n$ Lösungen ein reelles Fundamentalsystem von Lösungen enthalten.

Lemma 22.10 *Ist $\varphi_1, \dots, \varphi_n : I \rightarrow \mathbb{R}$ ein komplexes Fundamentalsystem des homogenen reellen Systems $Y' = A(x)Y$, so lassen sich aus den $2n$ Funktionen $\operatorname{Re} \varphi_1, \operatorname{Im} \varphi_1, \dots, \operatorname{Re} \varphi_n, \operatorname{Im} \varphi_n : I \rightarrow \mathbb{R}$ stets n über \mathbb{R} linear unabhängige, also ein reelles Fundamentalsystem, auswählen.*

Beweis. Zu fest vorgegebenem $a \in I$ und vorgegebenem j zwischen 1 und n gibt es komplexe Zahlen a_{jk} , so daß

$$\sum_{k=1}^n a_{jk} \varphi_k(a) = e_j.$$

Bildet man die entsprechende Gleichung der Realteile, so gewinnt man

$$\sum_{k=1}^n ((\operatorname{Re} a_{jk}) \operatorname{Re} \varphi_k(a) - (\operatorname{Im} a_{jk}) \operatorname{Im} \varphi_k(a)) = e_j.$$

Damit erzeugen die Vektoren $\operatorname{Re} \varphi_k(a), \operatorname{Im} \varphi_k(a)$ den Vektorraum \mathbb{R}^{2n} , so daß man aus ihnen n reell linear unabhängige auswählen kann. Die entsprechenden Funktionen bilden dann automatisch ein reelles Lösungs-Fundamentalsystem. \square

Bemerkung. Sind die komplexen Lösungen $\varphi_1, \dots, \varphi_n$ so beschaffen, daß an einer Stelle $a \in I$ die Werte $\varphi_j(a) = e_j$ angenommen werden, so verschwinden die Imaginärteile automatisch identisch auf

I und die $\varphi_1, \dots, \varphi_n$ sind sowohl ein komplexes als auch ein reelles Lösungs–Fundamentalsystem.

Wir kommen nun zu den Differentialgleichungen mit konstanten Koeffizienten und behandeln zuerst eine einzige (homogene) Differentialgleichung n -ter Ordnung:

$$y^{(n)} + a_{n-1}y^{(n-1)} + \dots + a_0y = 0 .$$

Als Definitionsintervall können wir jetzt selbstverständlich $I = \mathbb{R}$ wählen. Wesentlich ist für dieses Problem das Polynom

$$P(T) = T^n + a_{n-1}T^{n-1} + \dots + a_0 .$$

Da dieses i. a. ein gutes Zerfällungsverhalten nur über dem algebraisch abgeschlossenen Körper \mathbb{C} besitzt, setzen wir von vornherein voraus, daß auch die Koeffizienten komplexe Zahlen sein dürfen:

$$a_0, \dots, a_{n-1} \in \mathbb{C} ,$$

und daß wir wie weiter oben erläutert auch *komplexe Lösungen* zulassen, d. h. Funktionen

$$\varphi : I \rightarrow \mathbb{C} , \quad \varphi = g + ih , \quad g, h : I \rightarrow \mathbb{R} ,$$

also $g = \operatorname{Re} \varphi$, $h = \operatorname{Im} \varphi$, so daß mit

$$\varphi'(x) = g'(x) + ih'(x)$$

gilt:

$$\varphi^{(n)} + a_{n-1}\varphi^{(n-1)} + \dots + a_0\varphi = 0 .$$

Eine leichte Übertragung der obigen Überlegungen zeigt dann, daß die Lösungen in diesem Fall einen n -dimensionalen \mathbb{C} -Vektorraum bilden. Wenn ferner a_0, \dots, a_{n-1} *reelle* Zahlen sind und $\varphi = g + ih$ eine *komplexe* Lösung bezeichnet, so ist

$$0 = P(D)\varphi = P(D)g + iP(D)h , \quad \text{also} \quad P(D)g = P(D)h = 0 ,$$

d. h. mit φ sind auch Real- und Imaginärteil Lösungen. Insbesondere führen, wie gerade zuvor dargelegt wurde, *reelle* Anfangsbedingungen für ein *komplexes* Fundamentalsystem

$$\varphi_1, \dots, \varphi_n , \quad \varphi_j = g_j + ih_j , \quad j = 1, \dots, n ,$$

automatisch zu dem *reellen* Fundamentalsystem g_1, \dots, g_n .

Zur Vereinfachung der Schreibweise setzen wir mit dem Polynom P im folgenden

$$P(D) := D^n + a_{n-1}D^{n-1} + \dots + a_0 ,$$

wobei $D = d/dx$ den *Ableitungsoperator* bezeichnet, und schreiben die in Frage stehende Differentialgleichung in der Form

$$P(D)y = 0 .$$

Wir nennen $P(D)$ auch einen *linearen Differentialoperator mit konstanten Koeffizienten*.

Es ist nun naheliegend, das Polynom $P(T)$ zu faktorisieren:

$$P(T) = (T - \lambda_1)^{k_1} \cdot \dots \cdot (T - \lambda_r)^{k_r} ,$$

wobei $\lambda_1, \dots, \lambda_r$ die paarweise verschiedenen *komplexen* Nullstellen von $P(T)$ bezeichnen. Es gilt dann

$$P(D) = (D - \lambda_1)^{k_1} \cdot \dots \cdot (D - \lambda_r)^{k_r} .$$

Satz 22.11 V sei ein (nicht notwendig endlich-dimensionaler) komplexer Vektorraum, es sei $\alpha \in \text{End}(V)$, $P \in \mathbb{C}[x]$ ein nicht verschwindendes Polynom, und W sei der Kern des Endomorphismus $P(\alpha)$. Ferner sei $P = P_1^{k_1} \cdot \dots \cdot P_r^{k_r}$ die Zerlegung von P in paarweise verschiedene Linearfaktoren $P_j = P_j(T) := T - \lambda_j$, und es sei $W_j = \ker P_j(\alpha)^{k_j}$, $j = 1, \dots, r$. Dann gilt:

- (i) $W = W_1 \oplus \dots \oplus W_r$;
- (ii) W_j ist α -invariant für alle $j = 1, \dots, r$.

Bemerkung. Dieser Satz bleibt auch im Reellen richtig, wenn man den Begriff „Linearfaktoren“ durch *Primfaktoren* ersetzt, wobei die Primfaktoren eines reellen Polynoms die Linearfaktoren mit reellem λ oder die quadratischen Faktoren mit nichtreellen, konjugiert-komplexen Nullstellen sind. Der Satz wird manchmal auch als *Satz über die Primär-Zerlegung* bezeichnet⁴⁶.

Für die *Anwendung* dieses Satzes auf unser Problem setzen wir

$$V = C^\infty(I, \mathbb{C}), \quad \alpha = D : V \rightarrow V,$$

$$W = \ker P(D) = \{ \varphi \in C^\infty(I, \mathbb{C}) : P(D)\varphi = 0 \};$$

W ist also gerade der Raum der C^∞ -Lösungen unseres Systems. Da dieser aber, wie wir sehen werden, die erwartete Dimension n hat, haben wir damit schon *alle* Lösungen gefunden und insbesondere gezeigt, daß diese beliebig oft differenzierbar sein müssen (was aber auch schon unmittelbar aus der Definition folgt). Mit Hilfe des früheren Lipschitz-Argumentes kann man übrigens *a priori* beweisen, daß die Lösungen *analytisch* sein müssen. Gleiches ergibt sich aus den obigen Argumenten, wenn wir den Raum der C^∞ -Funktionen durch den der reell-analytischen, also C^ω -Funktionen ersetzen. Nach dem obigen Satz gilt dann

$$W = \bigoplus_{j=1}^r \{ \varphi \in C^\infty(I, \mathbb{C}) : (D - \lambda_j)^{k_j} \varphi = 0 \},$$

(wobei es nicht schwer ist, dies auch ohne Verwendung von *Linearer Algebra* direkt zu zeigen; siehe z. B. FORSTER, *Analysis 2*). Also ist nur der Lösungsraum für den speziellen Gleichungstyp

$$(D - \lambda)^k \varphi = 0, \quad \lambda \in \mathbb{C},$$

zu studieren. Ist $k = 1$, so bedeutet dies einfach

$$D\varphi = \lambda\varphi, \quad \text{so daß} \quad \ker(D - \lambda) = \{ c e^{\lambda x} : c \in \mathbb{C} \}.$$

Ist nun $f \in C^\infty(I, \mathbb{C})$ beliebig, so gilt

$$(D - \lambda)f = Df - \lambda f = e^{\lambda x} e^{-\lambda x} Df - \lambda e^{\lambda x} e^{-\lambda x} f = e^{\lambda x} D(e^{-\lambda x} f),$$

und hieraus folgt durch vollständige Induktion $(D - \lambda)^k f = e^{\lambda x} D^k(e^{-\lambda x} f)$ wegen

$$(D - \lambda)^{k+1} f = (D - \lambda)[e^{\lambda x} D^k(e^{-\lambda x} f)] = e^{\lambda x} D[e^{-\lambda x} (e^{\lambda x} D^k(e^{-\lambda x} f))] = e^{\lambda x} D^{k+1}(e^{-\lambda x} f).$$

Die letzte Beziehung impliziert die Äquivalenz $(D - \lambda)^k f = 0 \iff D^k(e^{-\lambda x} f) = 0$, und die letzte Gleichung ist gleichbedeutend damit, daß $e^{-\lambda x} f$ ein Polynom eines Grades $\leq k - 1$, die Lösung f also von der Gestalt

$$f(x) = e^{\lambda x} (c_0 + c_1 x + \dots + c_{k-1} x^{k-1}), \quad c_0, \dots, c_{k-1} \in \mathbb{C} \text{ beliebig},$$

ist. Also ist insbesondere der Lösungsraum dieser speziellen Gleichung k -dimensional.

Im ursprünglichen Problem ist aber $\sum k_j = n$, und die Summe der einzelnen Lösungsräume ist direkt, so daß wir insgesamt n über \mathbb{C} unabhängige Lösungen konstruiert haben. - Wir haben insbesondere bewiesen:

⁴⁶Einen Beweis findet man bei Hoffmann-Kunze [50], pp 219 ff., und für den Fall, daß V endlich-dimensional und F das charakteristische Polynom von α ist, so daß $W = V$ wegen des Satzes von CAYLEY-HAMILTON, in [41].

Satz 22.12 Die Funktionen $x^\ell e^{\lambda_j x}$, $\ell = 0, \dots, k_j - 1$, $j = 1, \dots, r$, formen ein komplexes Lösungsfundamentalsystem der Differentialgleichung $P(D)y = 0$, wobei $P(T) = (T - \lambda_1)^{k_1} \dots (T - \lambda_r)^{k_r}$ die Linearfaktorzerlegung des Polynoms $P(T)$ ist.

Beispiel. $y^{(iv)} + 8y'' + 16y = 0$. Hier ist das zugrundeliegende Polynom

$$P(T) = T^4 + 8T^2 + 16 = (T^2 + 4)^2 = (T - 2i)^2(T + 2i)^2$$

mit den komplexen Nullstellen $\lambda_{1,2} = \pm 2i$ der Vielfachheit 2. Es ergeben sich die vier fundamentalen komplexen Lösungen

$$\begin{cases} e^{2ix} = \cos 2x + i \sin 2x, & xe^{2ix} = x \cos 2x + ix \sin 2x, \\ e^{-2ix} = \cos 2x - i \sin 2x, & xe^{-2ix} = x \cos 2x - ix \sin 2x, \end{cases}$$

deren acht Real- und Imaginärteile reelle Lösungen der Gleichung sind. Offensichtlich gibt es aber unter diesen nur vier paarweise verschiedene, nämlich

$$\cos 2x, x \cos 2x, \sin 2x, x \sin 2x,$$

und diese bilden (automatisch) ein reelles Fundamentalsystem.

Beispiel (Gedämpfte und ungedämpfte Schwingungen). Wir behandeln die gewöhnliche lineare Differentialgleichung zweiter Ordnung

$$\frac{d^2 y}{dx^2} + 2\mu \frac{dy}{dx} + \omega_0^2 y = 0, \quad \mu \geq 0.$$

Für $\mu = 0$ beschreibt diese Gleichung eine *ungedämpfte Schwingung* mit den fundamentalen Lösungen

$$\begin{cases} y_1 = \cos \omega_0 x, \\ y_2 = \sin \omega_0 x. \end{cases}$$

Ist $\mu > 0$ (*gedämpfte Schwingung*), so hat man den Operator

$$P(D) = D^2 + 2\mu D + \omega_0^2,$$

wobei

$$P(T) = T^2 + 2\mu T + \omega_0^2 = (T - \lambda_1)(T - \lambda_2)$$

mit

$$\lambda_{1,2} = -\mu \pm \sqrt{\mu^2 - \omega_0^2}.$$

Selbstverständlich sind jetzt drei Fälle zu unterscheiden:

a) $\mu^2 - \omega_0^2 < 0$, d. h. $0 < \mu < \omega_0$. Hier ist

$$\lambda_{1,2} = -\mu \pm i\omega, \quad \omega = \sqrt{\omega_0^2 - \mu^2}.$$

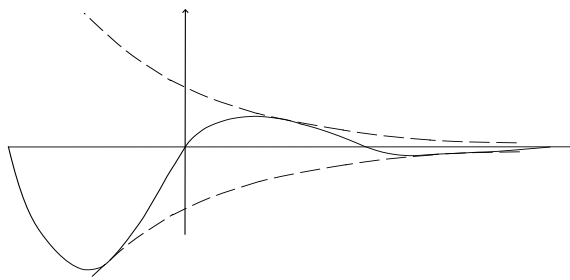
Also findet man die beiden komplexen Lösungen

$$e^{\lambda_1 x} = e^{-\mu x} e^{i\omega x}, \quad e^{\lambda_2 x} = e^{-\mu x} e^{-i\omega x}$$

und damit die (linear unabhängigen) reellen Lösungen

$$e^{-\mu x} \cos(\omega x), \quad e^{-\mu x} \sin(\omega x),$$

die mit wachsendem $x \rightarrow \infty$ gegen Null konvergieren.



Figur 22.1

b) $\mu^2 - \omega_0^2 = 0$, d. h. $\mu = \omega_0$. Hier hat man die doppelte Nullstelle $-\mu$ und damit die beiden Fundamentallösungen

$$e^{-\mu x}, \quad x e^{-\mu x}.$$

Im Gegensatz zu der schwachen Dämpfung im Fall a) liegt hier der Grenzfall zur starken Dämpfung vor, bei dem nach evtl. einem Ausschlag keine weiteren *Schwingungen* mehr stattfinden.

c) Im letzten Fall hat man $\mu^2 - \omega_0^2 > 0$, also die beiden verschiedenen reellen Nullstellen $\lambda_{1,2} = -\mu \pm \sqrt{\mu^2 - \omega_0^2} < 0$. Wir setzen $\mu_{1,2} := -\lambda_{1,2} > 0$ und finden die beiden unabhängigen Fundamentallösungen

$$e^{-\mu_1 x}, \quad e^{-\mu_2 x}.$$

Die inhomogene Gleichung $P(D)y = b(x)$ ist natürlich mit der im letzten Kapitel beschriebenen allgemeinen Methode der variablen Koeffizienten zu lösen, wenn die Fundamentallösungen für die homogene Gleichung $P(D)y = 0$ bekannt sind. Bei *speziellen* Inhomogenitäten b kann man aber auch andere Methoden anwenden. Wir betrachten z. B. die folgende Situation.

Satz 22.13 Die Differentialgleichung

$$P(D)y = e^{\mu x}, \quad P(\mu) \neq 0,$$

besitzt die partikuläre Lösung

$$\psi(x) := \frac{1}{P(\mu)} e^{\mu x}.$$

Beweis. Aus $D e^{\mu x} = \mu e^{\mu x}$ folgt $D^k e^{\mu x} = \mu^k e^{\mu x}$ und damit sowohl $P(D) e^{\mu x} = P(\mu) e^{\mu x}$ als auch

$$P(D) \left(\frac{1}{P(\mu)} e^{\mu x} \right) = \frac{1}{P(\mu)} P(D) e^{\mu x} = \frac{P(\mu)}{P(\mu)} e^{\mu x} = e^{\mu x}. \quad \square$$

Beispielsweise kann man hiermit sehr einfach das Problem der *erzwungenen Schwingung* (im Nichtresonanzfall) behandeln.

Wir besprechen als nächstes den Fall der *Differentialgleichungssysteme mit konstanten Koeffizienten*, aus dem man umgekehrt die vorstehenden Ergebnisse über Differentialgleichungen höherer Ordnung noch einmal auf andere Weise ableiten kann.

Es sei dazu A eine komplexe $n \times n$ -Matrix. Wir betrachten gemäß der allgemeinen Theorie nur (mit einer Ausnahme analog zu Satz 13) *homogene* Systeme von linearen Differentialgleichungen erster Ordnung mit konstanten Koeffizienten:

$$Y' = AY$$

und suchen natürlich ein komplexes (falls A reell ist, auch ein reelles) Fundamentalsystem von Lösungen. Dies ist ein schwierigeres Problem als das zuvor behandelte.

Zerfällt die Matrix A entlang der Hauptdiagonalen in kleinere „Kästchen“, so zerfällt auch das vorgelegte Differentialgleichungssystem in mehrere unabhängige Systeme, die man entsprechend auch unabhängig voneinander lösen kann. Im Extremfall besitzt A Diagonalgestalt; dann besteht das System aus n völlig unabhängigen Differentialgleichungen erster Ordnung.

Es ist somit naheliegend zu versuchen, A in eine geeignete einfachere Gestalt zu transformieren. Daher untersuchen wir als erstes das allgemeine *Transformationsverhalten* solcher Differentialgleichungssysteme. Sei dazu S eine invertierbare Matrix und $B := S^{-1} A S$. Dann gilt:

$$\varphi' = A \varphi \quad \text{genau dann, wenn} \quad \psi' = B \psi, \quad \text{wobei} \quad \psi = S^{-1} \varphi.$$

Denn: $\psi' = (S^{-1} \varphi)' = S^{-1} \varphi' = S^{-1} A \varphi = B \psi$ und umgekehrt.

Wir wissen aus der *Linearen Algebra*, daß eine quadratische Matrix zu einer oberen Dreiecksmatrix ähnlich ist, wenn ihr charakteristisches Polynom vollständig in Linearfaktoren zerfällt. Da dies im Komplexen stets erfüllt ist, gibt es also eine invertierbare komplexe Matrix S , so daß $B = S^{-1} A S$ obere Dreiecksgestalt besitzt. Nach dem eben Gesagten genügt es dann, das System $Y' = B Y$ zu lösen, das ausgeschrieben von der folgenden Form ist:

$$\left. \begin{array}{l} y_1' = b_{11} y_1 + b_{12} y_2 + \cdots + b_{1n} y_n \\ y_2' = \phantom{b_{11} y_1} + b_{22} y_2 + \cdots + b_{2n} y_n \\ \vdots \\ y_n' = \phantom{b_{11} y_1} \phantom{b_{22} y_2} + b_{nn} y_n \end{array} \right\}.$$

Ein solches System kann aber sukzessive von unten nach oben aufgelöst werden, da es sich bei jedem Lösungsschritt nur um eine einzige (inhomogene) lineare Differentialgleichung handelt. Man findet auf diese Weise auch geeignete komplexe Lösungen ψ_1, \dots, ψ_n mit gegebenen Anfangsbedingungen, z. B.

$$\psi_j(0) = e_j, \quad j = 1, \dots, n,$$

also ein spezielles Fundamentalsystem. Ist nun A reell, so geht man genauso vor. Für jede komplexe Lösung $\psi = g + ih$ sind dann automatisch g und h reelle Lösungen. Hat man n Lösungen mit den obigen Anfangsbedingungen gewählt: $\psi_j(0) = e_j$, so ist notwendig $g_j(0) = e_j$, woraus man wie schon weiter oben schließt, daß die Funktionen g_1, \dots, g_n ein reelles Fundamentalsystem bilden (und die h_j alle identisch verschwinden). Also ist das ganze Problem sowohl im Komplexen als auch im Reellen im Prinzip mit dem Verfahren der Trigonalisierung zu behandeln.

Dies ist aber unbefriedigend, da wenig konzeptionell. Eine konzeptionellere Methode gewinnt man durch die Verwendung der *Exponentialfunktion für Matrizen*.

Satz 22.14 *Es sei $A \in M(n \times n, \mathbb{C})$. Dann konvergiert die Reihe*

$$\exp A = \sum_{k=0}^{\infty} \frac{1}{k!} A^k,$$

und zwar sogar absolut, in dem Sinne, daß in der Folge der Partialsummen dieser Reihe alle Einträge absolut konvergente Folgen bilden. Vertauschen überdies zwei Matrizen A, B , so ist

$$\exp(A + B) = \exp A \cdot \exp B.$$

Beweis. Die Zeilensummen-Norm $\|A\|$ für Matrizen (siehe den Anhang zu Kapitel 17) erfüllt die Dreiecksungleichung, und aus der Abschätzung $\|AB\| \leq \|A\| \cdot \|B\|$ folgt

$$\left\| \sum_{k=\ell}^m \frac{1}{k!} A^k \right\|_{\infty} \leq \left\| \sum_{k=\ell}^m \frac{1}{k!} A^k \right\| \leq \sum_{k=\ell}^m \frac{1}{k!} \|A\|^k.$$

Mit dem Majorantenkriterium ergibt sich daraus unmittelbar die absolute Konvergenz aller Eintragsfolgen. Die zweite Aussage beweisen wir nach einer kurzen Zwischenüberlegung. \square

Ist $A \in M(n \times n, \mathbb{C})$, so ist $\exp(xA)$ für alle $x \in \mathbb{C}$ eine Matrix, deren Einträge überall auf \mathbb{C} konvergente Potenzreihen darstellen. Man kann also $\exp(xA)$ differenzieren:

$$\exp(xA)' = \sum_{k=0}^{\infty} \frac{1}{k!} (x^k \cdot A^k)' = \sum_{k=1}^{\infty} \frac{k}{k!} x^{k-1} A^k = A \cdot \exp(xA).$$

Somit sind die Spalten von $\exp(xA)$, $x \in \mathbb{R}$, Lösungen des Systems $Y' = AY$ auf \mathbb{R} , und man kann unmittelbar einsehen, daß sie ein Fundamentalsystem bilden, denn offensichtlich ist

$$\exp(0 \cdot A) = \exp(0) = E_n.$$

Übrigens gilt allgemein wegen $A(-A) = (-A)A$ und Satz 1, daß

$$\exp(xA) \exp(-xA) = \exp(xA - xA) = \exp 0 = E_n,$$

so daß die Matrix $\exp(xA)$ an jeder Stelle $x \in \mathbb{R}$ invertierbar ist.

Wir können mit der vorstehenden Betrachtung den Beweis für die *Funktionalgleichung der matrixwertigen Exponentialfunktion* auf äußerst elegante Weise erledigen, also für die Gleichung

$$\exp(A+B) = \exp A \cdot \exp B, \quad \text{wenn } AB = BA.$$

Man beweist leicht induktiv:

$$A^k B = A^{k-1} (AB) = A^{k-1} (BA) = (A^{k-1} B) A = (B A^{k-1}) A = B A^k,$$

woraus sofort $\exp(xA) \cdot B = B \cdot \exp(xA)$ folgt. Man definiere die Matrix

$$\Phi(x) := \exp(x(A+B)) - \exp(xA) \exp(xB).$$

Dann berechnet man die Ableitung zu

$$\begin{aligned} \Phi'(x) &= (A+B) \exp(x(A+B)) - A \exp(xA) \exp(xB) - \exp(xA) B \exp(xB) \\ &= (A+B) \exp(x(A+B)) - A \exp(xA) \exp(xB) - B \exp(xA) \exp(xB) \\ &= (A+B) \Phi(x). \end{aligned}$$

Also haben wir n Spaltenlösungen der Gleichung $Y' = (A+B)Y$ vorliegen. Nun ist aber

$$\Phi(0) = \exp 0 - \exp 0 \cdot \exp 0 = E_n - E_n \cdot E_n = 0,$$

so daß wegen des Identitätssatzes für Lösungen von Systemen mit Lipschitzbedingung alle Spalten Null sein müssen, d. h. $\Phi(x) \equiv 0$. \square

Um ehrlich zu sein, liefert auch Satz 14 i. A. kein *effektives* Verfahren, um ein Fundamentalsystem von Lösungen zu bestimmen. Es ergibt sich aber eine Vereinfachung und eine wichtige theoretische Einsicht, wenn wir die Theorie der *Jordanschen Normalform* für Matrizen heranziehen. Wir benötigen sogar nur, daß A eine obere Dreiecksmatrix ist, für deren Zerlegung

$$A = D + N$$

in den *Diagonalanteil* D und den *nilpotenten Anteil* $N = A - D$ die Vertauschungsrelation $DN = ND$ erfüllt ist. Dies ist auf jeden Fall gegeben für ein einzelnes *Jordan-Kästchen* zu einem festen Eigenwert λ . Wir schreiben genauer

$$D = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$$

mit den Eigenwerten $\lambda_1, \dots, \lambda_n$ von A . N ist nilpotent von einer Ordnung $\leq n$. Da beide Matrizen nach Voraussetzung vertauschen sollen, erhalten wir

$$\exp(xA) = \exp(xD + xN) = \exp(xD) \cdot \exp(xN),$$

wobei offensichtlich

$$\exp(xD) = \begin{pmatrix} e^{\lambda_1 x} & & 0 \\ & \ddots & \\ 0 & & e^{\lambda_n x} \end{pmatrix}$$

eine Diagonalmatrix ist und

$$\exp(xN) = \sum_{k=0}^{n-1} \frac{1}{k!} x^k N^k$$

eine obere Dreiecksmatrix darstellt, die in der Hauptdiagonalen nur Einsen enthält, in der ersten oberen „Nebendiagonale“ nur Polynome höchstens erster Ordnung in x , in der zweiten nur solche höchstens zweiter Ordnung, etc. Man kann im Übrigen alles noch weiter signifikant vereinfachen, wenn man die volle Kraft der Jordanschen Normalform benutzt.

Wir illustrieren die bisherigen Darlegungen an einem *Beispiel*:

$$\begin{cases} y_1' = -y_2 \\ y_2' = y_1, \end{cases}$$

also für das Gleichungssystem $Y' = AY$, wobei

$$A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

das charakteristische Polynom $\chi_A(\lambda) = \lambda^2 + 1$ mit den Nullstellen $\lambda_{1,2} = \pm i$ besitzt. Nun ist

$$D := \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix} = S^{-1} A S, \quad \text{wenn man } S = \begin{pmatrix} 1/\sqrt{2} & -i/\sqrt{2} \\ -i/\sqrt{2} & 1/\sqrt{2} \end{pmatrix}$$

wählt. Das zu D gehörige Gleichungssystem lautet

$$Z' = D Z, \quad \text{also } \begin{cases} z_1' = i z_1, \\ z_2' = -i z_2, \end{cases}$$

dessen Fundamentallösung einfach durch

$$\Psi(x) = \begin{pmatrix} e^{ix} & 0 \\ 0 & e^{-ix} \end{pmatrix}$$

gegeben ist. Daraus berechnet sich ein Fundamentalsystem für die ursprüngliche Gleichung zu

$$\Phi(x) = S \Psi(x) = \begin{pmatrix} 1/\sqrt{2} & -i/\sqrt{2} \\ -i/\sqrt{2} & 1/\sqrt{2} \end{pmatrix} \begin{pmatrix} e^{ix} & 0 \\ 0 & e^{-ix} \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} e^{ix} & -ie^{-ix} \\ -ie^{ix} & e^{-ix} \end{pmatrix}.$$

Geht man jetzt noch zu den Real- und Imaginärteilen über, so sieht man, daß nach neuer Normierung die Spalten des Realteils

$$\begin{pmatrix} \cos x \\ \sin x \end{pmatrix} \quad \text{und} \quad \begin{pmatrix} -\sin x \\ \cos x \end{pmatrix}$$

als reelles Fundamentalsystem überleben.

Wir geben noch ein weiteres, anspruchsvolleres *Beispiel*, nämlich das homogene System $Y' = AY$ mit der Matrix

$$A := \begin{pmatrix} 2 & 0 & 1 \\ 0 & 2 & 0 \\ 0 & 1 & 3 \end{pmatrix}.$$

Sie besitzt den Eigenvektor $v_1 := e_1$ zum Eigenwert 2 und den Eigenvektor $v_3 := e_1 + e_3$ zum Eigenwert 3. Der Eigenwert 2 hat die Vielfachheit 2; man rechnet leicht nach, daß der Eigenraum zu diesem Eigenwert nur die Dimension 1 besitzt. Somit muß die Jordansche Normalform von A wie folgt aussehen:

$$J = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix},$$

und der Vektor $v_2 := -e_2 + e_3$ wird unter A tatsächlich abgebildet auf $e_1 - 2e_2 + 2e_3 = v_1 + 2v_2$. Der Basiswechsel von der Standard-Basis in die Basis v_1, v_2, v_3 wird also geliefert von der Matrix

$$C := \begin{pmatrix} 1 & 0 & 1 \\ 0 & -1 & 0 \\ 0 & 1 & 1 \end{pmatrix} \quad \text{mit der Inversen} \quad C^{-1} = \begin{pmatrix} 1 & -1 & -1 \\ 0 & -1 & 0 \\ 0 & 1 & 1 \end{pmatrix},$$

denn der inverse Basiswechsel wird beschrieben durch $e_1 = v_1$, $e_2 = v_3 - v_1 - v_2$, $e_3 = v_3 - v_1$. Man rechnet leicht nach, daß in der Tat $C^{-1}AC = J$. Die Differentialgleichung $Y' = JY$ hat nun, wie oben ausgeführt und wie man auch leicht nachrechnen kann, das Lösungsfundamentalsystem

$$\psi_1(x) = \begin{pmatrix} e^{2x} \\ 0 \\ 0 \end{pmatrix}, \quad \psi_2(x) = \begin{pmatrix} x e^{2x} \\ e^{2x} \\ 0 \end{pmatrix}, \quad \psi_3(x) = \begin{pmatrix} 0 \\ 0 \\ e^{3x} \end{pmatrix},$$

so daß die ursprüngliche homogene Gleichung $Y' = AY$ gelöst wird durch die vektorwertigen Funktionen $C\varphi_j$, $j = 1, 2, 3$, also durch

$$\varphi_1(x) = \begin{pmatrix} e^{2x} \\ 0 \\ 0 \end{pmatrix}, \quad \varphi_2(x) = \begin{pmatrix} x e^{2x} \\ -e^{2x} \\ e^{2x} \end{pmatrix}, \quad \varphi_3(x) = \begin{pmatrix} e^{3x} \\ 0 \\ e^{3x} \end{pmatrix}.$$

Wir fügen noch ein einfaches Resultat an, das man bei speziellen Inhomogenitäten von Differentialgleichungssystemen mit konstanten Koeffizienten manchmal mit Gewinn verwenden kann.

Satz 22.15 *Es sei $v \in \mathbb{C}^n$ ein Eigenvektor der Matrix $A \in M(n \times n, \mathbb{C})$ zum Eigenwert $\lambda \in \mathbb{C}$. Dann besitzt die inhomogene Differentialgleichung*

$$Y' = AY + e^{\omega x} v$$

eine spezielle komplexwertige Lösung der Gestalt

$$Y = \frac{1}{\omega - \lambda} e^{\omega x} v, \quad \omega \neq \lambda, \quad \text{bzw.} \quad Y = x e^{\omega x} v, \quad \omega = \lambda.$$

Beweis. Mit $Y = \beta e^{\omega x} v$ ist $Y' = \beta \omega e^{\omega x} v$ und $AY + e^{\omega x} v = (\lambda \beta + 1) e^{\omega x} v$. Damit wird die Differentialgleichung genau dann durch Y gelöst, wenn $\lambda \beta + 1 = \omega \beta$, d. h. genau dann, wenn $\beta = (\omega - \lambda)^{-1}$. Im Fall $\lambda = \omega$ sieht man genauso schnell, daß die angegebene Funktion $Y = \beta x e^{\omega x} v$ genau für $\beta = 1$ eine Lösung ist. \square

Beispiel. Man löse das inhomogene Differentialgleichungssystem

$$Y' = \begin{pmatrix} 2 & 0 & 1 \\ 0 & 2 & 0 \\ 0 & 1 & 3 \end{pmatrix} Y + e^{2x} \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix}.$$

Hier ist der Vektor $v := 2e_1 + e_3 = {}^t(2, 0, 1)$ kein Eigenvektor der gegebenen Matrix A (die Eigenvektoren haben wir weiter oben bestimmt), aber offenbar Summe des Eigenvektors $v_1 := e_1$ zum Eigenwert 2 und des Eigenvektors $v_3 := e_1 + e_3$ zum Eigenwert 3. Infolgedessen kann man wegen der Linearität des Problems eine *partikuläre* Lösung der vorgegebenen Differentialgleichung durch *Addition* zweier vektorwertiger Funktionen wie im Satz davor mit $\omega = \lambda = 2$ bzw. $2 = \omega < \lambda = 3$ gewinnen. Das präzise Aufschreiben dieser Lösung liefert keine weitere tiefere Einsicht und wird deshalb unterlassen.

Wir geben noch ein weiteres *Anwendungsbeispiel*. Der durch einen Transportmechanismus bedingte Austausch zwischen zwei Behältern mit Salzen in homogener Lösung läßt sich durch das Gleichungssystem

$$\begin{aligned} \dot{m}_1 &= -k_1 m_1 + k_2 m_2, \\ \dot{m}_2 &= k_1 m_1 - k_2 m_2, \end{aligned}$$

beschreiben. Dabei bezeichnen m_1, m_2 die zur Zeit t in den beiden Behältern befindlichen Salzmenngen, und die positiven Zahlen k_1, k_2 sind konstante Übergangsraten. Man berechne die Lösungen dieses Systems unter den Anfangsbedingungen $m_1(0) = M_1, m_2(0) = M_2$ und untersuche ihr Langzeitverhalten.

Durch Addition der beiden Gleichungen erhält man $(m_1 + m_2)' = 0$ und damit $m_1 + m_2 = \text{const.} = M_1 + M_2 =: M$. Einsetzen in die erste Gleichung liefert dann

$$\dot{m}_1 = -k m_1 + k_2 M \quad \text{mit} \quad k := k_1 + k_2$$

mit der speziellen konstanten Lösung $(k_2/k) M$. Die gesuchte Lösung mit der geforderten Anfangsbedingung ist dann

$$m_1(t) = \frac{k_1 M_1 - k_2 M_2}{k_1 + k_2} e^{-(k_1+k_2)t} + \frac{k_2(M_1 + M_2)}{k_1 + k_2},$$

woraus sich mit $m_2 = M - m_1$ sofort auch

$$m_2(t) = \frac{k_2 M_2 - k_1 M_1}{k_1 + k_2} e^{-(k_1+k_2)t} + \frac{k_1(M_1 + M_2)}{k_1 + k_2}$$

ergibt. Für $t \rightarrow \infty$ geht m_1 gegen $\frac{k_2(M_1 + M_2)}{k_1 + k_2}$ und m_2 gegen $\frac{k_1(M_1 + M_2)}{k_1 + k_2}$.

Wir wollen an einem weiteren Beispiel demonstrieren, wie die Theorie der linearen Differentialgleichungssysteme auch zu *theoretischen* Einsichten genutzt werden kann.

Satz 22.16 *Es sei A eine reelle $n \times n$ -Matrix. Genau dann ist A schiefsymmetrisch, d. h. ${}^t A = -A$, wenn für jede Lösung $\varphi : \mathbb{R} \rightarrow \mathbb{R}^n$ der Differentialgleichung $Y' = AY$ die euklidische Norm $\|\varphi(x)\| = \text{konstant}$ ist.*

Beweis. a) Wir zeigen gleich noch allgemeiner: Ist A schief-symmetrisch, so ist das Skalarprodukt $\langle \varphi_1(x), \varphi_2(x) \rangle$ für zwei beliebige Lösungen stets konstant. Der Beweis ist denkbar einfach:

$$\begin{aligned} \langle \varphi_1, \varphi_2 \rangle' &= \langle \varphi_1', \varphi_2 \rangle + \langle \varphi_1, \varphi_2' \rangle = \langle A \varphi_1, \varphi_2 \rangle + \langle \varphi_1, A \varphi_2 \rangle \\ &= {}^t(A \varphi_1) \varphi_2 + {}^t \varphi_1 A \varphi_2 = {}^t \varphi_1 ({}^t A + A) \varphi_2 = 0. \end{aligned}$$

b) Ist umgekehrt die euklidische Norm für jede Lösung konstant, so bleibt wegen der *Polarisierungsformel* $2 \langle \varphi_1, \varphi_2 \rangle = \|\varphi_1 + \varphi_2\|^2 - \|\varphi_1\|^2 - \|\varphi_2\|^2$ auch das Skalarprodukt von je zwei Lösungen konstant. Damit ist notwendig ${}^t \varphi_1 ({}^t A + A) \varphi_2 = 0$ für je zwei Lösungen. Wählt man nun zu beliebigen $j, k = 1, \dots, n$ die beiden Lösungen so, daß $\varphi_1(0) = e_j$, $\varphi_2(0) = e_k$, so impliziert diese Bedingung das Verschwinden des Eintrags der Matrix ${}^t A + A$ an der Stelle j, k . \square

Zum Abschluß dieses Kapitels wollen wir noch kurz auf homogene lineare Differentialgleichungen höherer Ordnung mit konstanten Koeffizienten und ihre Beziehung zu entsprechenden Systemen eingehen. Der Differentialgleichung

$$y^{(n)} + a_{n-1}(x)y^{(n-1)} + \dots + a_0(x)y = P(D)y = 0$$

entspricht offenbar das System

$$Y' = \Lambda Y,$$

wobei Λ die sogenannte *Begleitmatrix*

$$\Lambda := \begin{pmatrix} 0 & 0 & \dots & 0 & -a_0 \\ 1 & 0 & \dots & 0 & -a_1 \\ 0 & 1 & \dots & 0 & -a_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & -a_{n-1} \end{pmatrix}$$

ist. Hier sind die Koeffizienten a_0, \dots, a_{n-1} nach dem VIETA-schen *Wurzelsatz* (bis auf ein geeignetes Vorzeichen) die *elementar-symmetrischen* Funktionen in den Nullstellen des Polynoms P .

Es stellt sich somit natürlich die Frage nach der *JORDAN-Normalform* einer solchen Matrix Λ . Wegen unserer früheren Überlegungen sollte es nicht verwundern, daß diese zerfällt in eine direkte Summe von Jordanschen Normalformen, die zu den Potenzen der Linearfaktoren gehören. Es genügt daher, den Spezialfall des Polynoms $P(T) = (T - \lambda)^n$ zu behandeln. Es ist eine wohlbekanntes Tatsache⁴⁷, daß die dem Polynom $(T - \lambda)^n$ zugehörige Begleitmatrix eine *JORDAN-Matrix* der Form

$$\begin{pmatrix} \lambda & 0 & 0 & \dots & 0 \\ 1 & \lambda & 0 & \dots & 0 \\ 0 & 1 & \lambda & \dots & 0 \\ \vdots & \vdots & & \ddots & \vdots \\ 0 & 0 & \dots & 1 & \lambda \end{pmatrix}$$

ist. Somit kann man Satz 12 auch direkt aus den Ergebnissen über Systeme mit konstanten Koeffizienten herleiten.

⁴⁷Hierzu siehe z. B. LORENZ, *Lineare Algebra II*, p. 161.

23 Das Newton - Verfahren und der Banachsche Fixpunktsatz

Viele Existenzaussagen der Mathematik lassen sich auf Lösungen von *Iterationsproblemen* zurückführen. Hierbei hat man eine angenäherte Lösung zu einem Problem und eine genaue *Iterationsvorschrift*, also einen *Algorithmus*, der aus der ersten Näherung eine bessere macht, so daß die Folge der Iterationen gegen die gesuchte Lösung konvergiert. Das Standardbeispiel eines solchen Algorithmus ist das *NEWTON-Verfahren*, das wir zur Bestimmung k -ter Wurzeln reeller Zahlen schon früh in diesem Kurs kennengelernt haben. Das Verfahren besitzt aber ein wesentlich größeres Anwendungsfeld, wenn man die entscheidenden Eigenschaften der Wurzelfunktionen in abstrakter Form herausarbeitet. - Man kommt so zu dem folgenden Resultat.

Satz 23.1 (Newton - Verfahren) *Es sei $f : I := [a, b] \rightarrow \mathbb{R}$ eine zweimal differenzierbare konvexe Funktion, so daß also $f'' \geq 0$, mit $f(a) < 0 < f(b)$. Dann gibt es genau eine Nullstelle von f in I , also ein $\xi \in I$ mit*

$$f(\xi) = 0.$$

Es ist ferner $f'(x) > 0$ für alle $x > \xi$, und die Folge

$$x_{n+1} := x_n - \frac{f(x_n)}{f'(x_n)}, \quad f(x_0) \geq 0$$

ist wohldefiniert und konvergiert monoton fallend gegen ξ . Ist sogar $f'(x) \geq C > 0$ und $f''(x) \leq K$ auf I , so gilt

$$|x_{n+1} - x_n| \leq |\xi - x_n| \leq \frac{K}{2C} |x_n - x_{n-1}|^2.$$

Bemerkungen. 1. Die im Satz angegebene Iterationsfolge konvergiert daher unter den letzten Voraussetzungen wie im Falle der Wurzelapproximation *quadratisch* und folglich *fehlerkorrigierend* gegen die gesuchte Nullstelle.

2. Analoge Aussagen wie in Satz 1 erhält man auch unter den folgenden Voraussetzungen:

- f ist konvex auf I , es ist $f(a) > 0 > f(b)$, und der Startwert x_0 erfüllt die Bedingung $f(x_0) \geq 0$;
- f ist *konkav* auf I , es ist $f(a) > 0 > f(b)$, und der Startwert x_0 erfüllt die Bedingung $f(x_0) \leq 0$;
- f ist *konkav* auf I , es ist $f(a) < 0 < f(b)$, und der Startwert x_0 erfüllt die Bedingung $f(x_0) \leq 0$.

3. Ist f in I zweimal stetig differenzierbar, hat f'' dort keine Nullstelle und ist $f(a)f(b) < 0$, so liegt eine dieser vier angegebenen Situationen vor. Der Startwert x_0 ist dann so zu wählen, daß $f(x_0)f''(x_0) \geq 0$. Ist also f'' in einer Umgebung der gesuchten Stelle ξ von Null verschieden, so führt das Newton-Verfahren stets zum Ziel, falls der Startwert x_0 nur nahe genug bei ξ gewählt wird. Liegt x_0 noch nicht auf der richtigen Seite von ξ , so tut dies auf jeden Fall x_1 , wovon man sich leicht überzeugt.

4. *Regula falsi.* Wir haben den Zwischenwertsatz und damit die Existenz von Nullstellen unter anderem mit der Methode der *Intervallhalbierung* bewiesen. Unter Umständen führt die folgende Methode schneller zum Ziel: Für stetiges $f : I := [a, b] \rightarrow \mathbb{R}$ mit $f(a) < 0 < f(b)$ bringt man die Sekante durch die Punkte $(a, f(a))$ und $(b, f(b))$ zum Schnitt mit der x -Achse. Da die Sekante die Gleichung

$$s(x) = \frac{f(b) - f(a)}{b - a} (x - a) + f(a)$$

besitzt, liefert dies den Wert

$$a - \frac{b - a}{f(b) - f(a)} f(a) = \frac{a f(b) - b f(a)}{f(b) - f(a)} = b - \frac{a - b}{f(a) - f(b)} f(b).$$

Man rechnet den Funktionswert von f an dieser Stelle aus und wählt, falls man nicht zufällig eine Nullstelle von f getroffen hat, eines der beiden von diesem Wert begrenzten Teilintervalle von I aus, auf das die ursprüngliche Voraussetzungen zutreffen. Auf diese Weise gewinnt man eine *Iterationsvorschrift*, die immer gegen eine Nullstelle von f konvergiert, wovon sich der Leser überzeugen möge.

Man benutzt oft die regula falsi zur Bestimmung eines „guten“ Startwerts für das Newton-Verfahren. Es gibt auch weitere Algorithmen, die das Newton-Verfahren und die regula falsi mischen.

Beispiel. Es sei $k \geq 2$ und $a \in \mathbb{R}_+^*$. Das Newton-Verfahren ist im Intervall $x \geq 0$ auf die Funktion $f(x) = x^k$ anwendbar und liefert die Iterationsfolge

$$x_{n+1} = \frac{1}{k} \left[(k-1)x_n + \frac{a}{x_n^{k-1}} \right]$$

mit

$$\lim x_n = \sqrt[k]{a}.$$

Man kann solche Folgen auch im Komplexen betrachten, z. B. für $k = 3$, $a = 1$:

$$3z_{n+1} = 2z_n + \frac{1}{z_n^2},$$

und betritt damit das „Wunderland“ der Iteration von *rationalen Funktionen* auf \mathbb{C} , einem der bevorzugten Tummelplätze der Propagandisten für die *Fraktale*.

Beweis (des Newton-Verfahrens). Nach Voraussetzung ist die Ableitung f' im gesamten Intervall I monoton wachsend. Ist $x_0 \in I$ eine Stelle, an der f ihr Minimum annimmt, so ist $x_0 < b$, und ist $a < x_0$, so ist notwendig $f'(x_0) = 0$. Also ist f im Intervall $[a, x_0]$ monoton fallend und kann dort keine Nullstelle besitzen. Auf dem Intervall $[x_0, b]$ ist f monoton steigend und nimmt nach dem Zwischenwertsatz dort mindestens eine Nullstelle an. Es sei ξ die kleinste unter diesen. (Da f stetig ist, existiert eine kleinste Nullstelle). Nach dem Mittelwertsatz gibt es ein $\xi_1 \in [x_0, \xi]$, so daß

$$f'(\xi_1) = \frac{f(\xi) - f(x_0)}{\xi - x_0} = \frac{-f(x_0)}{\xi - x_0} > 0.$$

Damit ist auch $f'(x) > 0$ für alle $x \geq \xi_1$ und folglich f im Intervall $[\xi_1, b]$ streng monoton wachsend. Wegen $\xi \in [\xi_1, b]$ kann f keine weiteren Nullstellen besitzen.

Es sei nun $x_0 \in I$ eine Stelle mit $f(x_0) \geq 0$, also insbesondere $x_0 \geq \xi$. Man beweist dann durch vollständige Induktion, daß die durch

$$x_{n+1} := x_n - \frac{f(x_n)}{f'(x_n)}$$

induktiv definierte Folge die folgenden Eigenschaften hat: $f(x_n) \geq 0$ und $\xi \leq x_{n+1} \leq x_n$. Insbesondere ist sie dann wegen $f'(x_n) \geq f'(\xi) > 0$ wohldefiniert, und aus der ersten Behauptung folgt damit schon die erwartete Monotonieaussage $x_{n+1} \leq x_n$. Es genügt somit zu zeigen, daß $f(x_{n+1}) \geq 0$, denn dann ist auch notwendig $x_{n+1} \geq \xi$, da anderenfalls ein Widerspruch zum Zwischenwertsatz deduziert werden könnte.

Hierzu definiert man die Hilfsfunktion $\varphi_n(x) := f(x) - f(x_n) - f'(x_n)(x - x_n)$. Wegen der Monotonie von f' gilt

$$\varphi_n'(x) := f'(x) - f'(x_n) \leq 0 \quad \text{für } x \leq x_n.$$

Da $\varphi_n(x_n) = 0$ ist, ist $\varphi_n(x) \geq 0$ für $x \leq x_n$, also insbesondere

$$0 \leq \varphi_n(x_{n+1}) = f(x_{n+1}) - f(x_n) - f'(x_n)(x_{n+1} - x_n) = f(x_{n+1}).$$

(Wir haben hier nur noch einmal indirekt die uns schon bekannte Tatsache ausgenutzt, daß für die konvexe Funktion f alle Funktionswerte oberhalb jeder Tangente, insbesondere der Tangente an der Stelle x_{n+1} , liegen).

Wir haben also gezeigt, daß die Folge x_n monoton fällt und nach unten durch ξ beschränkt ist. Also existiert der Grenzwert $x_\infty = \lim_{n \rightarrow \infty} x_n$. Aus dem Bildungsgesetz der Folge ergibt sich daraus aber

$$x_\infty = x_\infty - \frac{f(x_\infty)}{f'(x_\infty)}$$

und damit $f(x_\infty) = 0$. Wegen der Eindeutigkeit der Nullstelle ist schließlich $x_\infty = \xi$.

Es fehlt noch der Nachweis der letzten Aussage. Nach Voraussetzung ist $f(x) \geq f(\xi) \geq C$ für alle $x \geq \xi$. Daraus folgt $f(x) \geq C(x - \xi)$ für alle $x \geq \xi$ und damit

$$|x_{n+1} - x_n| \leq |\xi - x_n| \leq \frac{f(x_n)}{C}.$$

Um den Wert $f(x_n)$ abzuschätzen, betrachten wir die Hilfsfunktion $\psi_n(x) := \varphi_{n-1}(x) - (K/2)(x - x_{n-1})^2$. Differenzieren liefert

$$\psi'_n(x) = f'(x) - f'(x_{n-1}) - K(x - x_{n-1}) \quad \text{und} \quad \psi''_n(x) = f''(x) - K \leq 0 \quad \text{für alle } x \in [\xi, b].$$

Die Funktion ψ'_n ist also im Intervall $[\xi, b]$ monoton fallend, und wegen $\psi_n(x_{n-1}) = 0$ ist $\psi'_n(x) \geq 0$ für $x \in [\xi, x_{n-1}]$. Da schließlich auch $\psi_n(x_{n-1}) = 0$ ist, folgt weiter $\psi_n(x) \leq 0$ für $x \in [\xi, x_{n-1}]$, also insbesondere $\psi_n(x_n) = 0$ und damit

$$f(x_n) \leq \frac{K}{2}(x_n - x_{n-1})^2.$$

Man hätte die letzte Abschätzung übrigens auch sofort aus der Taylorformel zweiter Ordnung von f an der Stelle x_{n-1} durch Einsetzen von $x = x_n$ erhalten können. \square

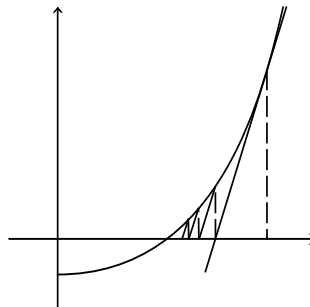
Bemerkungen. 1. Eine andere konkrete Variante des Newtonschen Iterationsverfahrens werden wir weiter unten aus dem Banachschen Fixpunktsatz ableiten.

2. Man kann sich leicht davon überzeugen, daß unter den obigen Voraussetzungen auch das sogenannte *vereinfachte Newton-Verfahren* stets gegen eine Nullstelle der Funktion f konvergiert. Es handelt sich hierbei um die induktiv definierte Folge

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_0)},$$

wozu man nur die Ableitung von f an dem Startwert x_0 auszurechnen braucht.

Eine Skizze hierzu sieht in etwa wie folgt aus:



Figur 23.1

Ein weiteres Iterationsproblem haben wir im Kapitel über die Existenz und Eindeutigkeit der Lösungen von Differentialgleichungssystemen kennengelernt. Das Verfahren von PICARD und LINDELÖF wurde später von BANACH zu einer schlagkräftigen allgemeinen Methode vervollkommen. Wir wollen zu Beginn dieses Kapitels den BANACHSchen *Fixpunktsatz* in seiner einfachsten Form beweisen und danach seine Nützlichkeit demonstrieren, indem wir mit seiner Hilfe erneut den Existenz- und Eindeutigkeitsatz für Systeme von gewöhnlichen Differentialgleichungen erster Ordnung herleiten. Will man allerdings ein größtmögliches Existenzintervall für Lösungen wie im Verfahren von Picard und Lindelöf erzielen, so muß man eine verbesserte Version bereitstellen. Wir bieten dafür dem Leser den Fixpunktsatz von WEISSINGER an mit einer weiteren Anwendung, dem noch ausstehenden Nachweis der Stetigkeit der allgemeinen Lösung des Differentialgleichungssystems $Y' = F(x, Y)$ (Satz 20.11), den wir in dem Anhang vorstellen. Der Banachsche Fixpunktsatz findet eine erneute Verwendung im nachfolgenden ersten Kapitel des Teils 3 beim Beweis des *Satzes über implizite Funktionen*.

Satz 23.2 (Banachscher Fixpunktsatz) *Es sei V ein Banach-Raum mit Norm $\|\cdot\|$, $A \subset V$ sei eine abgeschlossene Teilmenge, und es sei*

$$\alpha : A \longrightarrow V$$

eine Abbildung, für die gilt

$$(*) \quad \|\alpha(v) - \alpha(v')\| \leq C \|v - v'\| \quad \text{mit einer Konstanten } 0 < C < 1$$

für alle $v, v' \in A$. Ist dann zusätzlich

$$(**) \quad \alpha(A) \subset A,$$

so besitzt die Einschränkung von α auf A genau einen Fixpunkt v_∞ , also eine Stelle $v_\infty \in A$ mit $\alpha(v_\infty) = v_\infty$. Für jeden Startwert $v_0 \in A$ konvergiert die durch $v_{j+1} := \alpha(v_j)$ iterativ definierte Folge gegen v_∞ .

Ist $A = \bar{B} := \bar{B}(0, r)$ eine abgeschlossene Kugel um den Ursprung mit Radius r , so ist die Bedingung $(**)$ unter der Voraussetzung $(*)$ zum Beispiel dann erfüllt, wenn

$$\|\alpha(0)\| \leq (1 - C)r.$$

Gilt neben $(*)$ sogar die stärkere Abschätzung

$$\|\alpha(0)\| \leq \vartheta(1 - C)r$$

mit einer Zahl $0 < \vartheta < 1$, so bildet α die offene Kugel $B = B(0, r)$ in sich ab, und auch der Fixpunkt v_∞ liegt in B .

Definition und Bemerkung. Eine Abbildung α mit der Eigenschaft $(*)$ wie oben nennt man auch eine *Kontraktion*. Kontraktionen sind automatisch (gleichmäßig) *stetige* Abbildungen.

Beweis. Die Eindeutigkeitsaussage ergibt sich aus der Kontraktionseigenschaft: Sind v_∞ und w_∞ Fixpunkte von α , so ist

$$\|v_\infty - w_\infty\| = \|\alpha(v_\infty) - \alpha(w_\infty)\| \leq C \|v_\infty - w_\infty\|$$

und damit $v_\infty = w_\infty$ wegen $C < 1$. Die letzte Aussage ist ebenfalls denkbar einfach einzusehen: Für $\|v\| \leq r$ ist

$$\|\alpha(v)\| \leq \|\alpha(v) - \alpha(0)\| + \|\alpha(0)\| \leq C \|v - 0\| + (1 - C)r \leq r.$$

Es bleibt noch, die Konvergenz der Iterationsfolge (v_j) in V zu zeigen. Denn wegen der Abgeschlossenheit von A ist der Grenzwert v_∞ dann in A enthalten und ein Fixpunkt von α wegen

$$\alpha(v_\infty) = \alpha\left(\lim_{j \rightarrow \infty} v_j\right) = \lim_{j \rightarrow \infty} \alpha(v_j) = \lim_{j \rightarrow \infty} v_{j+1} = v_\infty.$$

Da V nach Voraussetzung ein Banach-Raum ist, müssen wir also nur noch nachweisen, daß (v_j) eine Cauchy-Folge ist. Nun ist aber leicht per vollständiger Induktion einzusehen, daß

$$\|v_{j+1} - v_j\| \leq C^j \|v_1 - v_0\| \quad \text{für alle } j \in \mathbb{N},$$

und damit gilt für alle $k \geq j$:

$$\|v_{k+1} - v_j\| \leq (C^j + C^{j+1} + \dots + C^k) \|v_1 - v_0\|.$$

Aus der Konvergenz der geometrischen Reihe für $0 < C < 1$ folgt die Behauptung.

Es bleibt schließlich noch, die letzte Aussage zu begründen. Wie im vorstehenden Beweis sieht man, daß für die Iterationsfolge mit $v_0 = 0$ gilt:

$$\|v_{k+1} - v_j\| \leq \vartheta (C^j + C^{j+1} + \dots + C^k) r (1 - C).$$

Insbesondere ist

$$\|v_{k+1}\| \leq \vartheta (C^0 + C^{j+1} + \dots + C^k) r (1 - C) \leq \vartheta r$$

und damit $v_\infty \in \overline{B}(0, \vartheta r) \subset B(0, r)$. □

Beispiele. 1. Wir behandeln noch einmal den Existenz- und Eindeutigkeitsatz für Systeme von gewöhnlichen Differentialgleichungen erster Ordnung $Y' = F(x, Y)$ mit einer stetigen Funktion $F : G \rightarrow \mathbb{R}^n$, $G \subset \mathbb{R} \times \mathbb{R}^n$ ein Gebiet. Ist nämlich $(a, c) \in G$, so wählen wir ein Intervall $I := [a, b] \subset \mathbb{R}$ und eine abgeschlossene Kugel $\overline{B}(c, r) \subset \mathbb{R}^n$ so, daß $Z := I \times \overline{B}(c, r) \subset G$ und $(b - a)M \leq r$, wobei M eine obere Schranke von $\|F\|$ auf dem Zylinder Z ist. Dann ist

$$V := \mathcal{C}^0(I, \mathbb{R}^n), \quad \text{versehen mit der Supremumsnorm,}$$

wegen des Weierstraßschen Konvergenzsatzes ein Banach-Raum, und die Menge der stetigen Funktionen

$$\overline{B} := \mathcal{C}^0(I, \overline{B}(c, r))$$

ist nichts anderes als die abgeschlossene Kugel in V mit Radius r und Mittelpunkt φ_0 mit $\varphi_0(x) \cong c$. Offensichtlich ist die Teilmenge

$$A := \{\varphi \in \overline{B} : \varphi(a) = c\}$$

eine abgeschlossene Teilmenge von V . Wegen der Standard-Abschätzung für Integrale hat der Picard-Operator $\alpha : A \rightarrow V$, der definiert wird durch

$$(\alpha(\varphi))(x) = c + \int_a^x F(t, \varphi(t)) dt,$$

die Eigenschaft, daß $\alpha(A) \subset A$. Setzen wir weiter voraus, daß α lokal einer Lipschitz-Bedingung genügt, so können wir $b > a$ und $r > 0$ so klein wählen, daß auf dem entsprechenden Zylinder Z eine Lipschitz-Konstante L zu F existiert und $(b - a)L =: C < 1$ ist. Dann ist α sogar eine Kontraktion, und der Banachsche Fixpunktsatz garantiert die Existenz genau einer Abbildung $\varphi \in A$ mit $\alpha(\varphi) = \varphi$, also einer Lösung φ des Differentialgleichungssystems $Y' = F(x, Y)$ auf I mit $\varphi(a) = c$.

2. Man kann auch *quadratische lineare Gleichungssysteme* iterativ lösen. Wir können ein solches stets in der Form

$$x_j - \sum_{k=1}^n c_{jk} x_k = b_j, \quad j = 1, \dots, n,$$

schreiben und die Abbildung

$$\mathbb{R}^n \ni x \longmapsto \alpha(x) := \left(\sum_{k=1}^n c_{1k} x_k + b_1, \dots, \sum_{k=1}^n c_{nk} x_k + b_n \right) \in \mathbb{R}^n$$

betrachten. Lösungen des Systems entsprechen dann Lösungen des Fixpunktproblems $\alpha(x) = x$. Es sei nun $\|\cdot\|$ irgendeine Norm auf \mathbb{R}^n und $\|C\|$ die zugehörige *Operatornorm* für die Matrix $C := (c_{jk})$. Dann gilt:

Lemma 23.3 Ist $\|C\| < 1$, so konvergiert die Folge $x_n := \alpha^n(x_0)$ für einen beliebigen Anfangswert $x_0 \in \mathbb{R}^n$ gegen eine Lösung des gegebenen linearen Gleichungssystems $x - Cx = b$.

Beweis. Es sei $A := V = \mathbb{R}^n$ mit der gegebenen Norm. Nach Voraussetzung ist für alle $x, \bar{x} \in \mathbb{R}^n$:

$$\|\alpha(x) - \alpha(\bar{x})\| = \|Cx - C\bar{x}\| \leq \|C\| \cdot \|x - \bar{x}\|. \quad \square$$

3. Das NEWTON-Verfahren läßt sich stets in einer kleinen Umgebung der gesuchten Nullstelle ξ ohne jegliche Konvexitätseigenschaften anwenden, wenn dort die Ableitung der Funktion f nirgends verschwindet. Wir setzen also voraus, daß die Funktion f auf dem Intervall $I := [a, b]$ zweimal stetig differenzierbar sei mit $f'(x) \neq 0$. Dann ist $f(\xi) = 0$ genau dann, wenn ξ Fixpunkt der Funktion

$$g(x) := x - \frac{f(x)}{f'(x)}$$

ist. Wegen

$$g'(x) = \frac{f(x)f''(x)}{(f'(x))^2}, \quad \text{also insbesondere } g'(\xi) = 0,$$

gibt es eine positive Zahl $L < 1$ und ein Intervall $J = [\xi - \delta, \xi + \delta] \subset I$ mit $\delta > 0$, so daß $|g'(x)| \leq L$ für alle $x \in J$. Dann wird J von g in sich abgebildet wegen

$$|g(x) - \xi| = |g(x) - g(\xi)| = |g'(\eta)| |x - \xi| \leq L|x - \xi|.$$

Ferner ist g kontrahierend, so daß die Newton-Folge $x_{n+1} := g(x_n)$ gegen einen Fixpunkt von g und damit eine Nullstelle von f konvergiert.

4. Die Idee des Beweises von Lemma 19.14, das essentiell für die Herleitung des Satzes von STONE-WEIERSTRASS war, läßt eine weitgehende Verallgemeinerung zu. Man betrachte eine nichtleere Teilmenge X und die Algebra $B(X)$ der beschränkten reellwertigen Funktionen auf X , die offensichtlich bzgl. der Supremumsnorm eine BANACH-Algebra bildet. Es gilt dann der folgende

Satz 23.4 Ist $B_0 \subset B(X)$ eine abgeschlossene Unteralgebra, die die 1 enthält, so enthält B_0 mit jeder Funktion $f \geq 0$ auch ihre Wurzel \sqrt{f} und damit für alle $f, g \in B_0$ auch $|f| = \sqrt{f^2}$, $\max(f, g)$ und $\min(f, g)$.

Beweisskizze. Man betrachtet zunächst die Folge $f_n := f + 1/n$ und beweist ähnlich wie in Lemma 19.14, daß die Folge $\sqrt{f_n}$ in $B(X)$ gegen \sqrt{f} konvergiert. Wegen $f_n \in B_0$ für $f \in B_0$ können wir daher annehmen, daß f nicht nur positiv ist, sondern auch nach unten durch eine positive Konstante α beschränkt ist: $f(x) \geq \alpha > 0$, $x \in X$. Da f auch nach oben beschränkt ist, erreicht man durch Multiplikation mit einer geeigneten positiven Konstanten β , daß für βf gilt: $\|1 - \beta f\| < 1$. Nun liegt genau dann die Wurzel von f in B_0 , wenn dies für βf gilt, so daß wir von vornherein $\|1 - f\| < 1$ voraussetzen dürfen. Unter dieser Voraussetzung versuchen wir, f in der Form $(1 - h)^2$ mit $h \in B_0$ zu schreiben, was zu $\sqrt{f} = 1 - h \in B_0$ führen würde. Schreibt man noch $g := 1 - f$, so ist das Problem äquivalent zu dem Fixpunktproblem

$$h = \frac{1}{2}(g + h^2), \quad \|g\| < 1.$$

Dieses besitzt in der Tat eine Lösung in $B_0 \cup \overline{B}_r(0)$, wobei r zwischen $\|g\|$ und 1 liegt und $\overline{B}_r(0) \subset B(X)$ die abgeschlossene Kugel vom Radius r bezeichnet. Man rechnet nämlich einfach nach, daß der Operator

$$h \mapsto \frac{1}{2}(g + h^2)$$

die Kugel $\overline{B}_r(0)$ in sich abbildet und kontrahierend mit Kontraktionskonstante $r < 1$ ist. Nach Definition bildet der Operator dann auch die abgeschlossene Menge $B_0 \cup \overline{B}_r(0)$ in sich ab, woraus mit

dem Banachschen Fixpunktsatz die Behauptung folgt. \square

Wir machen nun, wie schon oben angekündigt, einen weiteren Schritt vorwärts und formulieren einen Satz, der den BANACHSchen *Fixpunktsatz* als Spezialfall umfaßt, aber von größerer Kraft ist. Wir werden z. B. mit seiner Hilfe den *Existenzsatz* aus der Theorie der Systeme gewöhnlicher Differentialgleichungen noch einmal in voller Schärfe ableiten.

Satz 23.5 (Weissinger) *Es sei X ein vollständiger metrischer Raum, $A \subset X$ sei eine abgeschlossene Teilmenge, ε_n seien positive reelle Zahlen mit*

$$\sum_{n=1}^{\infty} \varepsilon_n < \infty,$$

und $\alpha : A \rightarrow A$ sei eine (notwendig stetige) Abbildung mit

$$d(\alpha^n(x), \alpha^n(x')) \leq \varepsilon_n d(x, x')$$

für alle $n \in \mathbb{N}^*$ und alle $x, x' \in A$. Dann besitzt α genau einen Fixpunkt $x_\infty \in A$:

$$\alpha(x_\infty) = x_\infty,$$

und für alle $x_0 \in A$ und alle $n \in \mathbb{N}^*$ gilt:

$$d(\alpha^n(x_0), x_\infty) \leq \left(\sum_{j=n}^{\infty} \varepsilon_j \right) d(\alpha(x_0), x_0),$$

d. h. insbesondere: Die Iterationsfolge $x_n := \alpha^n(x_0)$ konvergiert gegen x_∞ .

Beweis. Nach Definition ist $x_n = \alpha^n(x_0) = \alpha(\alpha^{n-1}(x_0)) = \alpha(x_{n-1})$. Es folgt

$$d(x_{n+1}, x_n) = d(\alpha(x_n), \alpha(x_{n-1})) = d(\alpha^n(x_1), \alpha^n(x_0)) \leq \varepsilon_n d(x_1, x_0),$$

und für $k \geq 1$ ist

$$d(x_{n+k}, x_n) \leq d(x_{n+k}, x_{n+k-1}) + \cdots + d(x_{n+1}, x_n) \leq (\varepsilon_n + \cdots + \varepsilon_{n+k-1}) d(x_1, x_0).$$

Ist also zu vorgegebenem $\varepsilon > 0$ die Zahl $N = N(\varepsilon)$ hinreichend groß, so gilt

$$(\varepsilon_n + \cdots + \varepsilon_{n+k-1}) d(x_1, x_0) < \varepsilon$$

für alle $n \geq N$ und alle $k \geq 1$. Also ist die Folge (x_n) eine Cauchy-Folge, die somit gegen einen Grenzwert $x_\infty \in X$ konvergiert. Da alle $x_n \in A$ sind und A abgeschlossen ist, ist notwendig $x_\infty \in A$. Weiter ist für $n \geq N$, wie man sofort aus der Dreiecksungleichung sieht:

$$d(x_n, x_\infty) = \lim_{k \rightarrow \infty} d(x_n, x_{n+k}) \leq \left(\sum_{j=n}^{\infty} \varepsilon_j \right) d(x_1, x_0).$$

(Hinter dieser Aussage steht die Tatsache, daß $d : X \times X \rightarrow \mathbb{R}$ eine *stetige* Abbildung ist bei festgehaltenem zweiten Argument bzgl. der von d induzierten Topologie auf X . Man erinnere sich an die entsprechende Aussage bzgl. *normierter Vektorräume*).

Wegen $d(\alpha(x), \alpha(x_0)) \leq \varepsilon_1 d(x, x_0)$ ist α eine stetige Abbildung. Infolgedessen ist selbstredend

$$\alpha(x_\infty) = \alpha\left(\lim_{n \rightarrow \infty} x_n\right) = \lim_{n \rightarrow \infty} \alpha(x_n) = \lim_{n \rightarrow \infty} x_{n+1} = x_\infty,$$

x_∞ also ein Fixpunkt von α .

Es bleibt zu zeigen, daß es *nur einen* solchen Fixpunkt geben kann: Sei \bar{x}_∞ ein weiterer, so ist nämlich

$$0 \leq d(x_\infty, \bar{x}_\infty) = d(\alpha^n(x_\infty), \alpha^n(\bar{x}_\infty)) \leq \varepsilon_n d(x_\infty, \bar{x}_\infty)$$

für alle $n \in \mathbb{N}$ und damit wegen $\lim_{n \rightarrow \infty} \varepsilon_n = 0$ auch $d(x_\infty, \bar{x}_\infty) = 0$, d. h. $x_\infty = \bar{x}_\infty$. \square

Bemerkung. Der BANACHsche Fixpunktsatz ist tatsächlich ein Spezialfall dieses Resultates. Aus (*) (siehe Satz 1) folgt nämlich durch Induktion sofort

$$\|\alpha^n(x) - \alpha^n(x')\| \leq C^n \|x - x'\|,$$

wobei die (geometrische) Reihe $\sum_{n=0}^{\infty} C^n$ wegen der Voraussetzung $0 < C < 1$ konvergent ist.

Aus dem Weissingerschen Fixpunktsatz folgt in der Tat erneut der Satz 20.8 in der genauen Fassung für Differentialgleichungssysteme:

Satz 23.6 (Iterationsverfahren von Picard und Lindelöf) *Es sei $F : G \rightarrow \mathbb{R}^n$ eine stetige Abbildung auf dem Gebiet $G \subset \mathbb{R} \times \mathbb{R}^n$. Zu $(a, c) \in G$ seien $\varepsilon > 0$ und $r > 0$ so gewählt, daß $Q := \{(x, y) \in \mathbb{R}^{n+1} : |x - a| \leq \varepsilon, \|y - c\| \leq r\} \subset G$. Die Abbildung F sei auf Q durch die Konstante M dem Betrage nach nach oben beschränkt und erfülle eine Lipschitz-Bedingung mit der Konstanten $L > 0$. Wird dann ε so klein gewählt, daß $\varepsilon \leq r/M$, so ordnet der „Picard-Operator“*

$$\varphi \longmapsto \psi := \mathcal{P}(\varphi) \quad \text{mit} \quad \psi(x) := c + \int_a^x F(t, \varphi(t)) dt$$

jeder stetigen Funktion $\varphi : I := [a - \varepsilon, a + \varepsilon] \rightarrow \bar{B}(c, r)$ mit $\varphi(a) = c$ eine stetige Funktion $\psi : I \rightarrow \bar{B}(c, r)$ mit $\psi(a) = c$ zu.

Ist ferner $\varphi_0 : I \rightarrow \bar{B}(c, r)$ irgendeine stetige Funktion mit $\varphi_0(a) = c$ und definiert man induktiv

$$\varphi_{k+1} := \mathcal{P}(\varphi_k),$$

so konvergiert die Folge (φ_k) auf I gleichmäßig gegen eine Grenzfunktion φ_∞ , die die Differentialgleichung $Y' = F(x, Y)$ mit der Anfangsbedingung $\varphi_\infty(a) = c$ löst.

Beweis. Wie in dem Beispiel nach dem Banachschen Fixpunktsatz sei

$$V := \mathcal{C}^0(I, \mathbb{R}^n), \quad \text{versehen mit der Supremumsnorm}$$

und

$$\bar{B} := \mathcal{C}^0(I, \bar{B}(c, r)), \quad A := \{\varphi \in \bar{B} : \varphi(a) = c\}.$$

Wegen der Standard-Abschätzung für Integrale bildet der Picard-Operator die Menge A in sich ab. Es braucht daher nur noch die zentrale Abschätzung für die Iterierten in dem Satz von Weissinger nachgeprüft zu werden. Seien dazu also $\varphi_0, \psi_0 \in A$ und φ_k, ψ_k ihre k -fachen Iterierten. Wie im Beweis von Satz 20.8 zeigt man durch Induktion, daß

$$(+)$$

$$\|\varphi_k(x) - \psi_k(x)\| \leq L^k \frac{|x - a|^k}{k!} \|\varphi_0 - \psi_0\|_\infty,$$

so daß man schließlich die Abschätzung

$$\|\varphi_k - \psi_k\|_\infty \leq \frac{(L\varepsilon)^k}{k!} \|\varphi_0 - \psi_0\|_\infty$$

erhält, die nach dem Satz von Weissinger auf jeden Fall hinreichend ist. \square

Bemerkung. Aus den allgemeinen Abschätzungen des Weissingerschen Fixpunktsatzes gewinnt man unmittelbar aus dem vorstehenden Beweis ganz vorzügliche Fehlerabschätzungen für das Iterationsverfahren von Picard und Lindelöf. Es ist

$$\|\varphi_\infty - \varphi_k\| \leq \left(\sum_{j=k}^{\infty} \frac{(L\varepsilon)^j}{j!} \right) \|\varphi_1 - \varphi_0\|,$$

und wenn man speziell $\varphi_0 \equiv c$ wählt,

$$\|\varphi_\infty - \varphi_k\| \leq \varepsilon M \left(\sum_{j=k}^{\infty} \frac{(L\varepsilon)^j}{j!} \right).$$

Bemerkung. Man kann die präzise Form des Iterationssatzes von Picard und Lindelöf dennoch aus dem Banachschen Fixpunktsatz ableiten, wenn man statt der Supremumsnorm eine andere „gewichtete“ Supremumsnorm verwendet. Einen Hinweis, wie man diese einzuführen hat, gewinnt man aus der vorstehenden Bemerkung. Siehe hierzu HEUSER [157], p. 70.

Analysiert man den Beweis des Weissingerschen Resultats noch etwas genauer, so stellt man schnell fest, daß sich die Voraussetzungen an die *stetige* Abbildung $\alpha : A \rightarrow A$ wesentlich abschwächen lassen. Dies führt zu dem folgenden Resultat.

Folgerung 23.7 *Es sei X ein vollständiger metrischer Raum, $A \subset X$ sei eine abgeschlossene Teilmenge, ε_n seien positive reelle Zahlen mit*

$$\sum_{n=1}^{\infty} \varepsilon_n < \infty,$$

und $\alpha : A \rightarrow A$ sei eine stetige Abbildung mit

$$d(\alpha^n(x), \alpha^n(x')) \leq \varepsilon_n C(x, x')$$

für alle n und alle $x, x' \in A$ mit einer nur von dem Paar $x, x' \in A$ abhängenden Konstanten $0 < C(x, x') < \infty$. Dann besitzt α genau einen Fixpunkt $x_\infty \in A$, und es gelten die folgenden Abschätzungen :

$$d(\alpha^n(x_0), x_\infty) \leq \left(\sum_{j=n}^{\infty} \varepsilon_j \right) C(\alpha(x_0), x_0) \quad \square$$

Anhang: Die Stetigkeit der allgemeinen Lösung von gewöhnlichen Differentialgleichungen

In diesem Anhang wollen wir u. a. der Frage nachgehen, wie weit sich zwei Lösungen einer gewöhnlichen Differentialgleichung voneinander entfernen können, falls ihre Anfangsbedingungen nahe beieinander liegen. Hierzu kann man nutzbringend ein Resultat heranziehen, das nach dem schwedischen Mathematiker und Ingenieur THOMAS HAKON GRONWALL benannt ist. Als tiefliegende Anwendung leiten wir mit Hilfe des Weissauerschen Fixpunktsatzes die im Titel genannte *Stetigkeit der allgemeinen Lösung von gewöhnlichen Differentialgleichungen* ab, die wir schon in Kapitel 20 diskutiert haben.

Lemma 23.8 (Gronwall) *Die Funktion $u : I \rightarrow \mathbb{R}$ sei stetig auf dem Intervall $I = [a, b)$, $a < b \leq \infty$, und es gelte die Abschätzung*

$$0 \leq u(x) \leq A + B \int_a^x u(\xi) d\xi \quad \text{für alle } x \in I$$

mit positiven Konstanten A, B . Dann ist notwendig

$$0 \leq u(x) \leq A e^{B(x-a)}, \quad x \in I.$$

Beweis. Da das Intervall I die Vereinigung aller *kompakten* Intervalle $[a, \beta]$, $\beta < b$, ist, genügt es, die behauptete Abschätzung nur für solche Intervalle zu beweisen. Wir nehmen daher an, daß I selbst das kompakte Intervall $[a, b]$, $a < b < \infty$, ist. Dann ist u auf I nach oben durch eine positive Konstante M beschränkt, und aus der vorausgesetzten Ungleichung ergibt sich

$$0 \leq u(x) \leq A + MB(x-a), \quad a \leq x \leq b.$$

Durch vollständige Induktion gewinnt man hieraus leicht unter Verwendung derselben Ungleichung:

$$0 \leq u(x) \leq A \sum_{j=0}^k \frac{B^j (x-a)^j}{j!} + \frac{MB^{k+1} (x-a)^{k+1}}{(k+1)!} \quad \text{für alle } k \in \mathbb{N} \text{ und } x \in [a, b].$$

Mit $k \rightarrow \infty$ folgt die Behauptung. □

Wir gewinnen aus Lemma 8 eine Abschätzung über das Maß des „Auseinanderdriftens“ zweier Lösungen eines Differentialgleichungssystems.

Satz 23.9 *Die Abbildung $F : G \rightarrow \mathbb{R}^n$, $G \subset \mathbb{R} \times \mathbb{R}^n$, erfülle eine globale Lipschitz-Bedingung mit der Lipschitz-Konstanten L . Dann gilt für zwei Lösungen $\Phi, \tilde{\Phi} : I \rightarrow \mathbb{R}^n$ des Differentialgleichungssystems $Y' = F(x, Y)$ bei festem a in dem gemeinsamen Definitionsintervall I :*

$$\|\Phi(x) - \tilde{\Phi}(x)\| \leq \|\Phi(a) - \tilde{\Phi}(a)\| e^{L|x-a|}, \quad x \in I.$$

Beweis. Die übliche Umwandlung der Differentialgleichung in eine Integralgleichung liefert zusammen mit der Lipschitz-Bedingung und der Standard-Abschätzung für Integrale

$$\begin{aligned} \|\Phi(x) - \tilde{\Phi}(x)\| &\leq \left\| (\Phi(a) - \tilde{\Phi}(a)) + \int_a^x (\Phi'(\xi) - \tilde{\Phi}'(\xi)) d\xi \right\| \\ &\leq \|\Phi(a) - \tilde{\Phi}(a)\| + L \left| \int_a^x \|\Phi(\xi) - \tilde{\Phi}(\xi)\| d\xi \right|. \end{aligned}$$

Der Rest ist eine unmittelbare Anwendung des Lemmas davor. □

Ist zusätzlich die Abbildung F auf dem Definitionsgebiet G beschränkt durch eine Konstante M , so ist für einen weiteren Punkt $\tilde{a} \in I$

$$\|\tilde{\Phi}(\tilde{a}) - \tilde{\Phi}(a)\| = \left\| \int_a^{\tilde{a}} F(\xi, \tilde{\Phi}(\xi)) d\xi \right\| \leq M |a - \tilde{a}|.$$

Somit erhalten wir die

Folgerung 23.10 *Unter den vorigen Voraussetzungen ist für alle $a, \tilde{a} \in I$:*

$$\|\Phi(x) - \tilde{\Phi}(x)\| \leq (\|\Phi(a) - \tilde{\Phi}(\tilde{a})\| + M |a - \tilde{a}|) e^{L|x-a|}, \quad x \in I.$$

Bemerkung. Einen anderen Beweis der letzten Folgerung findet man z. B. bei GRAUERT–FISCHER [7]. Dort wird sogar die folgende, weit stärkere Aussage über „Fastlösungen“ gezeigt.

Satz 23.11 *Die Abbildung $F : G \rightarrow \mathbb{R}^n$, $G \subset \mathbb{R} \times \mathbb{R}^n$, sei stetig, beschränkt durch eine Konstante M und erfülle eine Lipschitzbedingung mit Konstanten $L > 0$. Es seien $\Phi, \tilde{\Phi} : I \rightarrow \mathbb{R}^n$ differenzierbare Abbildungen, deren Graph in G enthalten sei. Ferner gelte :*

$$\|\Phi'(x) - F(x, \Phi(x))\| \leq \varepsilon \quad \text{und} \quad \|\tilde{\Phi}'(x) - F(x, \tilde{\Phi}(x))\| \leq \tilde{\varepsilon} \quad \text{für alle } x \in I.$$

Dann gilt für alle $x, a, \tilde{a} \in I$ mit $c := \Phi(a)$, $\tilde{c} := \tilde{\Phi}(\tilde{a})$:

$$\|\Phi(x) - \tilde{\Phi}(x)\| \leq \frac{\varepsilon + \tilde{\varepsilon}}{L} (e^{L|x-a|} - 1) + (\|c - \tilde{c}\| + (M + \tilde{\varepsilon}) |a - \tilde{a}|) e^{L|x-a|}. \quad \square$$

Man kann hieraus den Satz über die Stetigkeit der allgemeinen Lösung folgern. Wir wollen aber einem anderen Verfahren⁴⁸ folgen, das den Weissauerschen Fixpunktsatz verwendet und deshalb erst jetzt besprochen werden kann.

Wir müssen dazu noch einmal kurz an unsere Diskussion in Kapitel 20 erinnern und diese auf Differentialgleichungssysteme übertragen. Wir setzen also im folgenden voraus, daß das *stetige Richtungsfeld* $F : G \rightarrow \mathbb{R}^n$, $G \subset \mathbb{R} \times \mathbb{R}^n$ eine offene Menge, lokal einer Lipschitz-Bedingung genügt. Dann gibt es zu beliebiger Anfangsbedingung $(\xi, \eta) \in G$ lokal nahe ξ genau eine Lösung Φ der Differentialgleichung $Y' = F(x, Y)$ mit $\Phi(\xi) = \eta$. Diese kann zu einer Lösung auf einem *maximalen* (automatisch offenen) Intervall fortgesetzt werden, das wir mit $I_{(\xi, \eta)}$ bezeichnen; für die maximale Lösung schreiben wir $\Phi_{(\xi, \eta)}$ und für ihren Wert an der Stelle $x \in I_{(\xi, \eta)}$ auch

$$\Phi_{(\xi, \eta)}(x) =: \Phi(x; \xi, \eta).$$

Diese Funktion ist auf der Menge

$$\tilde{G} := \bigcup_{(\xi, \eta) \in G} I_{(\xi, \eta)} \times \{(\xi, \eta)\} \subset \mathbb{R} \times \mathbb{R} \times \mathbb{R}^n$$

definiert. Sie heißt die *allgemeine Lösung* der Differentialgleichung $Y' = F(x, Y)$ - Der Satz über die Stetigkeit der allgemeinen Lösung ist nun der folgende.

Satz 23.12 *Die Menge $\tilde{G} \subset \mathbb{R}^{n+2}$ ist offen, und die allgemeine Lösung $\Phi(x; \xi, \eta)$ ist stetig auf \tilde{G} .*

⁴⁸Siehe hierzu und zu dem gesamten Anhang: GÜNTHER J. WIRSCHING, *Gewöhnliche Differentialgleichungen*, Manuskript auf der homepage des Autors.

Beweis. Es genügt, die folgende Behauptung zu verifizieren:

Zu jedem Punkt $(a, \xi_0, \eta_0) \in \tilde{G}$ gibt es eine (kompakte) Umgebung $U \subset \mathbb{R}^{n+2}$ und eine stetige Abbildung $\Psi : U \rightarrow \mathbb{R}^n$, so daß für alle (ξ_1, η_1) die Funktion $x \mapsto \Psi(x, \xi_1, \eta_1) =: \psi_1(x)$ eine Lösung des gegebenen Differentialgleichungssystems mit $\psi_1(\xi_1) = \eta_1$ ist.

Denn daraus folgt, daß die allgemeine Lösung $\Phi(x; \xi, \eta)$ auf U mit Ψ übereinstimmt; insbesondere ist $U \subset \tilde{G}$ und damit \tilde{G} offen, und da Ψ stetig ist, muß auch die allgemeine Lösung auf U (und damit überall) stetig sein.

Der Beweis wird in mehreren Schritten durchgeführt.

1. Es sei $(a, \xi_0, \eta_0) \in \tilde{G}$. Nach Definition ist $a \in I_{(\xi_0, \eta_0)}$, und da das maximale Intervall offen ist, gibt es reelle Zahlen α, β mit

$$a, \xi_0 \in (\alpha, \beta) \subset [\alpha, \beta] \subset I_{(\xi_0, \eta_0)}.$$

Es sei weiter $\phi_0(x) := \Phi(x; \xi_0, \eta_0)$ für $\alpha \leq x \leq \beta$.

2. Es gibt eine (kompakte) *Tubenumgebung* des Graphen von ϕ_0 , die ganz in G verläuft, d. h. per definitionem, es gibt eine positive reelle Zahl r , so daß die Menge

$$T_r(\phi_0) := \bigcup_{x \in [\alpha, \beta]} \{x\} \times \overline{B}_r(\phi_0(x)) \subset G \subset \mathbb{R} \times \mathbb{R}^n.$$

Gäbe es nämlich kein solches r , so würde eine Folge $x_j \in [\alpha, \beta]$ und eine Nullfolge $r_j > 0$ existieren, so daß die n -dimensionale „Kugel“ $\{x_j\} \times \overline{B}_{r_j}(\phi_0(x_j))$ nicht in G enthalten ist. Ohne Einschränkung können wir nach Übergang zu einer Teilfolge annehmen, daß die Folge der x_j gegen einen Wert $\xi_\infty \in [\alpha, \beta]$ konvergiert. Dies führt zu einem Widerspruch, da eine offene $(n+1)$ -dimensionale Kugel mit Mittelpunkt $(\xi_\infty, \phi_0(\xi_\infty))$ in der offenen Menge G enthalten ist.

3. Die Tubenumgebung $T_r(\phi_0)$ ist offensichtlich folgenkompakt und damit kompakt. Infolgedessen gibt es, wie man mit der Überdeckungseigenschaft von Heine–Borel sofort verifiziert, eine für ganz $T_r(\phi_0)$ gültige Lipschitz-Konstante L .

4. Wähle jetzt eine positive reelle Zahl ρ so, daß ρ/r kleiner als die Zahlen $e^{-L(\xi_0 - \alpha)}$ und $e^{-L(\beta - \xi_0)}$ ist, und definiere

$$W(F, \phi_0) := \{(x, y) \in \mathbb{R} \times \mathbb{R}^n : \alpha \leq x \leq \beta, \|y - \phi_0(x)\| \leq \rho e^{-L|a-x|}\}.$$

Nach Konstruktion ist dies eine kompakte Umgebung von (a, ξ_0) , die ganz in $T_r(\phi_0)$ enthalten ist. Die gesuchte kompakte Umgebung U definieren wir schließlich durch

$$U = U(\alpha, \beta, F, \phi_0) := [\alpha, \beta] \times W(F, \phi_0) \subset \mathbb{R} \times G.$$

5. Die Menge

$$A := \{\phi \in \mathcal{C}^0(U, \mathbb{R}^n) : \|\phi(x, \xi, \eta) - \phi_0(x)\| \leq \rho e^{-L|a-\xi|} e^{L|x-\xi|}\}$$

ist abgeschlossen im Vektorraum $\mathcal{C}^0(U, \mathbb{R}^n)$, versehen mit der Supremumsnorm. Nach Definition von ρ ist weiter

$$\phi(U) \subset T_r(\phi_0) \quad \text{für alle } \phi \in A$$

und damit insbesondere

$$\sup_U \|\phi - \psi\| \leq 2r \quad \text{für alle } \phi, \psi \in A.$$

6. Als nächstes konstruieren wir einen Operator \mathcal{P} für Funktionen $\phi \in A$ durch

$$(\mathcal{P}\phi)(x, \xi, \eta) := \eta + \int_\xi^x F(t, \phi(t, \xi, \eta)) dt.$$

Um den Weissingerschen Fixpunktsatz 23.7 anwenden zu können, müssen wir zeigen, daß dieser die Menge A in sich abbildet. Wir müssen also einsehen, daß aus $\phi \in A$ folgt: $\mathcal{P}(\phi) \in A$, d. h.

$$\|(\mathcal{P}\phi)(x, \xi, \eta) - \phi_0(x)\| \leq \rho e^{-L|a-\xi|} e^{L|x-\xi|} \quad \text{für alle } (x, \xi, \eta) \in U.$$

Da ϕ_0 das vorgegebene Differentialgleichungssystem löst, ist

$$\int_{\xi}^x F(t, \phi_0(t)) dt = \int_{\xi}^x \phi_0'(t) dt = \phi_0(x) - \phi_0(\xi),$$

was zu der Abschätzung

$$\begin{aligned} \|(\mathcal{P}\phi)(x, \xi, \eta) - \phi_0(x)\| &= \left\| \eta + \int_{\xi}^x F(t, \phi(t, \xi, \eta)) dt - \phi_0(\xi) - \int_{\xi}^x F(t, \phi_0(t)) dt \right\| \\ &\leq \|\eta - \phi_0(\xi)\| + \left\| \int_{\xi}^x (F(t, \phi(t, \xi, \eta)) - F(t, \phi_0(t))) dt \right\| \end{aligned}$$

Anlaß gibt. Wegen $(x, \xi, \eta) \in U$ ist $(\xi, \eta) \in W(F, \phi_0)$ und folglich nach Definition

$$\|\eta - \phi_0(\xi)\| \leq \rho e^{-L|a-\xi|}.$$

Weiter gilt

$$\|F(t, \phi(t, \xi, \eta)) - F(t, \phi_0(t))\| \leq L \|\phi(t, \xi, \eta) - \phi_0(t)\| \leq L \rho e^{-L|a-\xi|} e^{L|t-\xi|}.$$

Damit kann man das noch zu behandelnde Integral wie folgt abschätzen:

$$\begin{aligned} \left\| \int_{\xi}^x (F(t, \phi(t, \xi, \eta)) - F(t, \phi_0(t))) dt \right\| &\leq L \rho e^{-L|a-\xi|} \left| \int_{\xi}^x e^{L|t-\xi|} dt \right| \\ &= \rho e^{-L|a-\xi|} (e^{L|x-\xi|} - 1). \end{aligned}$$

Insgesamt erhält man die gewünschte Ungleichung

$$\|(\mathcal{P}\phi)(x, \xi, \eta) - \phi_0(x)\| \leq \rho e^{-L|a-\xi|} + \rho e^{-L|a-\xi|} (e^{L|x-\xi|} - 1) = \rho e^{-L|a-\xi|} e^{L|x-\xi|}.$$

7. Leider ist aber der soeben konstruierte verallgemeinerte Picard–Operator \mathcal{P} i. A. keine Kontraktion bzgl. der Supremums–Metrik $d_{\infty}(\phi, \psi) := \sup_{(x, \xi, \eta) \in U} \|\phi(x, \xi, \eta) - \psi(x, \xi, \eta)\|$. Man kann nur zeigen, daß man stets

$$d_{\infty}(\mathcal{P}\phi, \mathcal{P}\psi) \leq (\beta - \alpha) L d_{\infty}(\phi, \psi)$$

hat. Dies reicht nur hin, wenn zufälligerweise das Intervall $[\alpha, \beta]$ kurz im Verhältnis zu der Lipschitz–Konstanten L ist.

8. Wir müssen daher das Intervall $[\alpha, \beta]$ in gewissem Sinne „künstlich“ kürzer machen und definieren aus diesem Grunde für eine beliebige reelle Zahl $\lambda > 0$ die Menge

$$U_{\lambda} := \{(x, \xi, \eta) \in U : |x - \xi| \leq \lambda\}.$$

Ist ℓ eine weitere positive reelle Zahl, so gibt es zu $(x, \xi, \eta) \in U_{\lambda+\ell}$ ein x_1 zwischen x und ξ mit $|x - x_1| \leq \ell$ und $|x_1 - \xi| \leq \lambda$. Insbesondere ist dann $(x_1, \xi, \eta) \in U_{\lambda}$, und wegen der offensichtlichen Gleichung

$$(\mathcal{P}\phi)(x, \xi, \eta) - (\mathcal{P}\psi)(x, \xi, \eta) = \phi(x_1, \xi, \eta) - \psi(x_1, \xi, \eta) + \int_{x_1}^x (F(t, \phi(t, \xi, \eta)) - F(t, \psi(t, \xi, \eta))) dt$$

schließt man auf die Ungleichung

$$\sup_{U_{\lambda+\ell}} \|\mathcal{P}\phi - \mathcal{P}\psi\| \leq \sup_{U_\lambda} \|\phi - \psi\| + \ell L \sup_{U_{\lambda+\ell}} \|\phi - \psi\|$$

8. Es sei nun $\lambda > 0$ fest gewählt. Wir behaupten dann, daß für alle $n, k \in \mathbb{N}^*$ die folgende Abschätzung richtig ist:

$$\sup_{U_{k\lambda}} \|\mathcal{P}^n\phi - \mathcal{P}^n\psi\| \leq \sum_{j=0}^k \binom{n}{j} (\lambda L)^{n-j} \cdot \sup_{U_{k\lambda}} \|\phi - \psi\|.$$

Dies beweist man per Induktion nach $n \geq 1$. Für $n = 1$ ist die Aussage offensichtlich richtig, da auf der rechten Seite der Term mit $j = 1$ vorkommt. Sei die Aussage nun schon für ein $n \geq 1$ bewiesen. Dann erhält man mit 7.:

$$\begin{aligned} \sup_{U_{k\lambda}} \|\mathcal{P}^{n+1}\phi - \mathcal{P}^{n+1}\psi\| &\leq \sup_{U_{(k-1)\lambda}} \|\mathcal{P}^n\phi - \mathcal{P}^n\psi\| + (\lambda L) \cdot \sup_{U_{k\lambda}} \|\mathcal{P}^n\phi - \mathcal{P}^n\psi\| \\ &\leq \left(\sum_{j=0}^{k-1} \binom{n}{j} (\lambda L)^{n-j} + \sum_{j=0}^k \binom{n}{j} (\lambda L)^{n-j+1} \right) \cdot \sup_{U_{k\lambda}} \|\mathcal{P}^n\phi - \mathcal{P}^n\psi\| \\ &\leq \left(\sum_{j=1}^k \binom{n}{j-1} (\lambda L)^{n-j+1} + \sum_{j=0}^k \binom{n}{j} (\lambda L)^{n-j+1} \right) \cdot \sup_{U_{k\lambda}} \|\mathcal{P}^n\phi - \mathcal{P}^n\psi\| \\ &\leq \left(\sum_{j=0}^{k+1} \binom{n+1}{j} (\lambda L)^{n-j+1} \right) \cdot \sup_{U_{k\lambda}} \|\mathcal{P}^n\phi - \mathcal{P}^n\psi\|. \end{aligned}$$

9. Wir wählen nun λ so klein, daß $\vartheta := \lambda L < 1$ wird. Sodann wähle man $k \in \mathbb{N}$ so groß, daß $k\lambda > \beta - \alpha$. Dann ist $U_{k\lambda} = U$, und aus dem vorigen Abschnitt folgt für alle $\phi, \psi \in A$ die Abschätzung

$$d_\infty(\mathcal{P}^n\phi, \mathcal{P}^n\psi) \leq \sum_{j=0}^k \binom{n}{j} \vartheta^{n-j} \cdot d_\infty(\phi, \psi).$$

Um den Weissingerschen Fixpunktsatz anwenden zu können, benötigen wir noch die Summierbarkeit der Folge

$$\varepsilon_n := \sum_{j=0}^k \binom{n}{j} \vartheta^{n-j}.$$

Nun zeigt man aber leicht mit dem Wurzelkriterium bei festem j die (absolute) Konvergenz der Reihe

$$\sum_{n=j}^{\infty} \binom{n}{j} \vartheta^{n-j},$$

und hieraus folgt die Behauptung aufgrund des Doppelreihensatzes.

10. Der modifizierte Picard-Operator \mathcal{P} besitzt also auf der Menge $A \subset \mathcal{C}^0(U, \mathbb{R}^n)$ einen eindeutig bestimmten Fixpunkt Ψ . Als solcher ist für feste Anfangswerte $(\xi, \eta) \in W(F, \phi_0)$ die Funktion $x \mapsto \Psi(x, \xi, \eta)$ die Lösung auf dem Intervall $[\alpha, \beta]$ zu den Anfangswerten (ξ, η) . Dies beendet den Beweis. \square

Eine weitere Anwendung des Gronwallschen Lemmas soll diesen Anhang beschließen. Sie behandelt eine wesentliche Verallgemeinerung des Satzes über die Existenz *globaler* Lösungen bei *linearen* Differentialgleichungssystemen (Satz 22.2).

Satz 23.13 *Es sei $I \subset \mathbb{R}$ ein offenes Intervall, und $F : I \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ sei eine stetige Abbildung, die stetig linear beschränkt in der zweiten Variablen sei, d. h. für die mit stetigen Funktionen $\alpha, \beta : I \rightarrow \mathbb{R}$ gilt :*

$$\|F(x, Y)\| \leq \alpha(x) + \beta(x)\|Y\|, \quad x \in I, Y \in \mathbb{R}^n.$$

Dann besitzt jedes Anfangswertproblem der Differentialgleichung $Y' = F(x, Y)$ über ganz I definierte Lösungen.

Beweis. Entsprechend Satz 20.10 kann man nur mit Hilfe des Peanoschen Existenzsatzes, also ohne lokale Lipschitz-Bedingung herleiten, daß es zu stetiger rechter Seite $F : G \rightarrow \mathbb{R}^n$, $G \subset \mathbb{R}^{n+1}$, immer Lösungen $\Phi : J \rightarrow \mathbb{R}^n$ der Differentialgleichung $Y' = F(x, Y)$ zu vorgegebener Anfangsbedingung $(a, c) \in G$ gibt, für die das (dann notwendig offene) Intervall J maximal in dem Sinne ist, daß der Graph der Lösung Φ jedes Kompaktum $K \subset G$ nach beiden Seiten hin verläßt.

Wir nehmen eine solche Lösung $\Phi : J \rightarrow \mathbb{R}^n$, $J \subset I$, in unserer speziellen Situation her und müssen begründen, daß $J = I$ sein muß. Dazu genügt der Nachweis der Existenz einer stetigen Funktion $\mu : I \rightarrow \mathbb{R}$ mit

$$(*) \quad \|\Phi(x)\| \leq \mu(x), \quad x \in J.$$

Denn wäre J (nach einer Seite von a hin) in einem Kompaktum $K \subset I$ enthalten, so müßte auch der Graph von Φ in einer kompakten Teilmenge von $G = I \times \mathbb{R}^n$ enthalten sein im Widerspruch zu dem soeben notierten Sachverhalt.

Zum Nachweis von $(*)$ halten wir $a \in I$ fest und nehmen ohne Beschränkung der Allgemeinheit an, daß die Funktionen α und β positiv sind und links von a monoton fallen bzw. rechts von a monoton steigen. Damit gilt für beliebige $x \in J$:

$$\begin{aligned} \|\Phi(x)\| &= \left\| \Phi(a) + \int_a^x \Phi'(\xi) d\xi \right\| = \left\| \Phi(a) + \int_a^x F(\xi, \Phi(\xi)) d\xi \right\| \\ &\leq \|\Phi(a)\| + \left| \int_a^x (\alpha(\xi) + \beta(\xi)\|\Phi(\xi)\|) d\xi \right| \\ &\leq \|\Phi(a)\| + \left| \int_a^x \alpha(\xi) d\xi \right| + \beta(x) \left| \int_a^x \|\Phi(\xi)\| d\xi \right|. \end{aligned}$$

Mit einer geeigneten Modifizierung des Gronwallschen Lemmas impliziert diese Ungleichung die Abschätzung

$$\|\Phi(x)\| \leq \left(\|\Phi(a)\| + \left| \int_a^x \alpha(\xi) d\xi \right| \right) e^{\beta(x)} =: \mu(x)$$

für alle $x \in J$. Die Funktion μ ist jedoch auf ganz I erklärt und dort stetig. □

