Analysis II

Alexander Grigorian
Universität Bielefeld

WS 2014/15

Contents

Chapter 4

Integralrechnung

In Analysis I haben wir die Operation von Differenzieren (Ableiten) gelernt. Sei I ein Intervall. Für eine Funktion $f : I \to \mathbb{R}$ definiert man die Ableitung $f'(x)$ an eine Stelle $x \in I$ mit

$$f'(x) = \lim_{h \to 0} \frac{f(x+h) - f(x)}{h},$$

vorausgesetzt, dass der Limes existiert. Wir haben auch gelernt, wie man die Ableitung berechnet (die Rechenregeln) und wie die Ableitung für die Untersuchung der Funktion benutzt werden kann.

Jetzt betrachten wir das inverse Problem: gegeben sei eine Funktion f, wie man eine Funktion F mit $F' = f$ bestimmen kann? Das heißt:

wie eine Funktion F durch ihre Ableitung wiederherstellt werden kann?

Diese Frage entsteht in vielen Anwendungen von Mathematik. Zum Beispiel, die Bestimmung der Position $x(t)$ von einem bewegenden Körper durch die gegebene Geschwindigkeit $v(t) = x'(t)$ führt zu diesem Problem. Noch allgemeineres Problem bekommt man aus dem Aktionsprinzip (Zweites Newtonsches Gesetz)

$$ma = F,$$

wobei m die Masse des Körpers ist, $a = a(t)$ die Beschleunigung an der Zeit t und F die bewegende Kraft. Da $a(t) = x''(t)$, so erhalten wir die Gleichung

$$x''(t) = \frac{F}{m}.$$

Ist F als eine Funktion von Zeit t bekannt, so bestimmt man erst x' und danach x. Ist F eine Funktion von x und x' wie häufig der Fall ist, so erhält man eine *Differentialgleichung* – eine Beziehung zwischen x'', x', x die x wiederherstellen lässt.

4.1 Unbestimmtes Integral

Definition. Gilt $F' = f$ auf einem Intervall I, so heißt die Funktion F eine *Stammfunktion* von f auf I.

Nicht alle Funktionen haben Ableitung. Auch nicht alle Funktion haben Stammfunktion. Später in der Vorlesung beweisen wir den folgenden Satz.

Satz. *Jede stetige Funktion auf einem Interval $I \subset \mathbb{R}$ hat eine Stammfunktion auf diesem Intervall.*

Für die Eindeutigkeit von Stammfunktion gilt folgendes.

Satz 4.1 *Ist F eine Stammfunktion von f auf einem Intervall $I \subset \mathbb{R}$, so jede Stammfunktion von f hat die Form $F(x) + C$, wobei C eine beliebige Konstante ist.*

Beweis. Gilt $F' = f$, so gilt auch $(F + C)' = F' = f$. Somit ist $F + C$ auch eine Stammfunktion von f. Umgekehrt, ist G noch eine Stammfunktion von f so gilt auf I die Identität $F' = G' = f$ woraus folgt $(G - F)' = 0$ on I. Nach dem Konstantentest (Satz 3.24 aus Analysis I) ist die Funktion $G - F$ gleich eine Konstante auf I. Bezeichnen wir diese Konstante mit C und erhalten $G(x) = F(x) + C$ für alle $x \in I$, was zu beweisen war. ■

Definition. Die Menge von allen Stammfunktionen von $f(x)$ wird mit

$$\int f(x)\,dx$$

bezeichnet ("Integral von f von x dx"). Dieser Ausdruck heißt auch das *unbestimmte* Integral von f. Nach dem Satz 4.1 ist $\int f(x)\,dx$ eine Funktion bis zur additiven Konstante.

Der Grund für diese Notation wird später geklärt. Im Moment zeigen wir die Beziehung zwischen den Begriffen von Integral und Differential. Das Differential von einer differenzierbaren Funktion F ist der Ausdruck

$$dF = F'(x)\,dx, \tag{4.1}$$

wobei dx is eine unabhängige Variable ist, die das Differential von x heißt. Somit ist dF eine lineare Funktion von dx (für jedes x).

Nach der obigen Definition gilt

$$\left(\int f(x)\,dx\right)' = f(x),$$

was nach (4.1) äquivalent zu

$$d\int f(x)\,dx = f(x)\,dx \tag{4.2}$$

ist. Nach dem Satz 4.1 gilt

$$\int F'(x)\,dx = F(x) + C, \tag{4.3}$$

was sich wie folgt umgeschrieben lässt:

$$\int dF(x) = F(x) + C. \tag{4.4}$$

Der Vergleich von (4.2) und (4.4) zeigt, dass die Symbolen d und $\int$ sich formal wegkürzen lassen (bis zur Konstante), d.h. die Operationen Integral und Differential zueinander invers sind.

Die Operation $f \mapsto \int f dx$ heißt unbestimmte *Integration*. In diesem Kapitel lernen wir die Methoden von unbestimmten Integration. Die einfachste Methode ist die in der Differentialrechnung erhaltenden Identitäten umzukehren. Zum Beispiel, da

$$\left(\frac{x^{n+1}}{n+1}\right)' = x^n, \quad n \neq -1,$$

so ergibt (4.3) die folgende Identität:

$$\boxed{\int x^n dx = \frac{x^{n+1}}{n+1} + C.}$$

Für allgemeines $n \in \mathbb{R} \setminus \{-1\}$ gilt diese Identität auf $(0, +\infty)$, für $n \in \mathbb{N} \cup \{0\}$ gilt sie auf $\mathbb{R}$, und für $n \in (-\mathbb{N})$ – auf $(-\infty, 0)$ und $(0, +\infty)$.

Insbesondere erhalten wir

$$\begin{aligned} \int dx &= x + C, \\ \int x dx &= \frac{x^2}{2} + C, \\ \int \sqrt{x} dx &= \frac{x^{3/2}}{3/2} + .C, \\ \int \frac{dx}{x^2} &= -\frac{1}{x} + C. \end{aligned}$$

Da $(\ln |x|)' = \frac{1}{x}$ auf $(0, +\infty)$ und $(-\infty, 0)$, so erhalten wir

$$\boxed{\int \frac{dx}{x} = \ln |x| + C} \text{ auf } (0, +\infty) \text{ und } (-\infty, 0).$$

Umkehrung von Ableitung von Exponentialfunktion ergibt

$$\boxed{\int \exp(x)\, dx = \exp(x) + C,}$$

und auch

$$\boxed{\int a^x dx = \frac{a^x}{\ln a} + C}$$

für $a > 0, \ a \neq 1$.

Umkehrung von Ableitungen von trigonometrischen Funktionen ergibt:

$$\boxed{\int \sin x dx = -\cos x + C}$$

$$\boxed{\int \cos x dx = \sin x + C},$$

$$\boxed{\int \frac{dx}{\cos^2 x} = \tan x + C},$$

auf jedem Intervall wo $\cos x$ nicht verschwindet,

$$\boxed{\int \frac{dx}{\sin^2 x} = -\cot x + C},$$

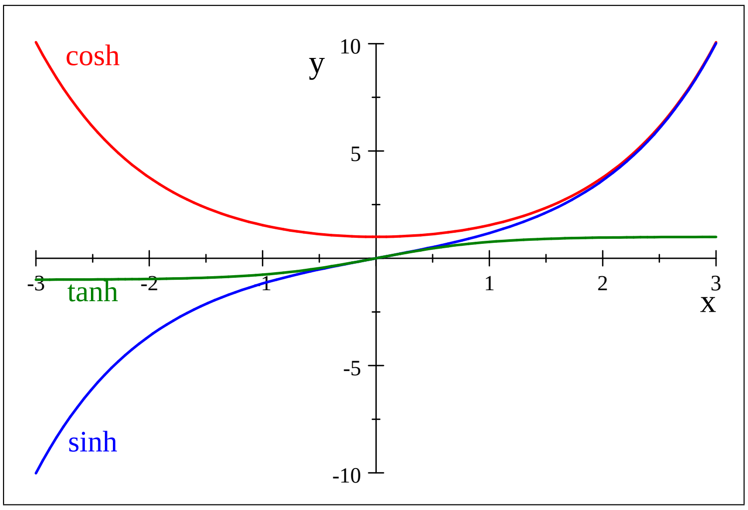
auf jedem Intervall wo $\sin x$ nicht verschwindet. Umkehrung von Ableitungen von inversen trigonometrischen Funktionen ergibt

$$\boxed{\int \frac{dx}{\sqrt{1-x^2}} = \arcsin x + C} \text{ auf } (-1,1)\,,$$

$$\boxed{\int \frac{dx}{1+x^2} = \arctan x + C} \text{ auf } (-\infty,\infty)\,.$$

Betrachten wir die Hyperbelfunktionen

$$\begin{aligned} \cosh x &= \frac{1}{2}\left(e^x + e^{-x}\right), \sinh x = \frac{1}{2}\left(e^x - e^x\right), \\ \tanh x &= \frac{\sinh x}{\cosh x}, \ \coth x = \frac{\cosh x}{\sinh x} = \frac{1}{\tanh x}. \end{aligned}$$



Umkehrung von Ableitungen von den Hyperbelfunktionen ergibt

$$\boxed{\int \sinh x = \cosh x + C}$$

$$\boxed{\int \cosh x dx = \sinh x + C}$$

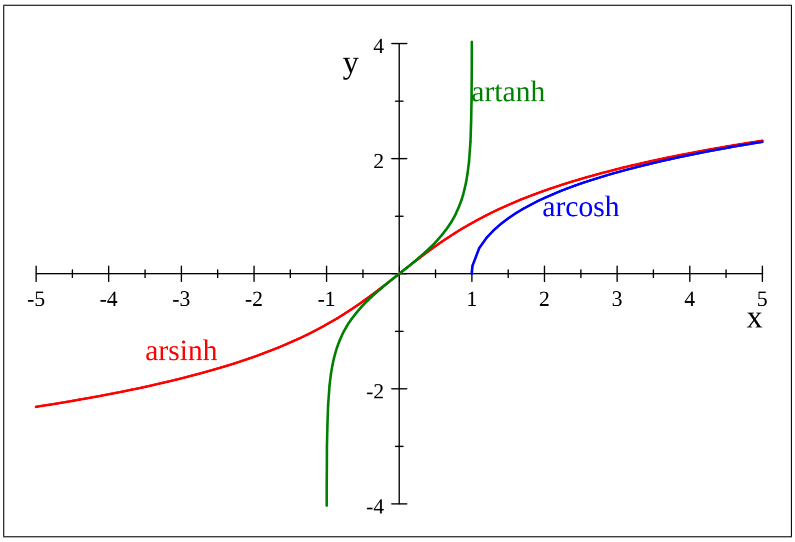
$$\boxed{\int \frac{1}{\cosh^2 x} = \tanh x + C}$$

$$\boxed{\int \frac{1}{\sinh^2 x} = \coth x + C} \text{ auf } (0,\infty) \text{ und } (-\infty,0)$$

Die Hyperbelfunktionen haben die folgenden inversen Funktionen.

Die inverse Funktion von sinh wird mit $\operatorname{arsinh} x$ bezeichnet. Sie hat den Definitionsbereich $(-\infty,\infty)$ und die Ableitung

$$(\operatorname{arsinh} x)' = \frac{1}{\sqrt{x^2+1}}.$$



Die inverse Funktion von cosh wird mit $\operatorname{arcosh} x$ bezeichnet. Sie hat den Definitionsbereich $[1,\infty)$ und die Ableitung auf $(1,\infty)$

$$(\operatorname{arcosh} x)' = \frac{1}{\sqrt{x^2-1}}.$$

Die inverse Funktion von tanh wird mit $\operatorname{artanh} x$ bezeichnet. Sie hat den Definitionsbereich $(-1,1)$ und die Ableitung

$$(\operatorname{artanh} x)' = \frac{1}{1-x^2}.$$

Darüber hinaus gelten die Identitäten:

$$\begin{aligned}
\operatorname{arsinh} x &= \ln\left(x+\sqrt{x^2+1}\right), \quad x \in \mathbb{R} \\
\operatorname{arcosh} x &= \ln\left(x+\sqrt{x^2-1}\right), \quad x \geq 1, \\
\operatorname{artanh} x &= \frac{1}{2}\ln\frac{1+x}{1-x}, \quad -1 < x < 1.
\end{aligned}$$

Somit erhalten wir folgendes:

$$\boxed{\int \frac{dx}{\sqrt{x^2+1}} = \operatorname{arsinh} x + C = \ln\left(x+\sqrt{x^2+1}\right) + C},$$

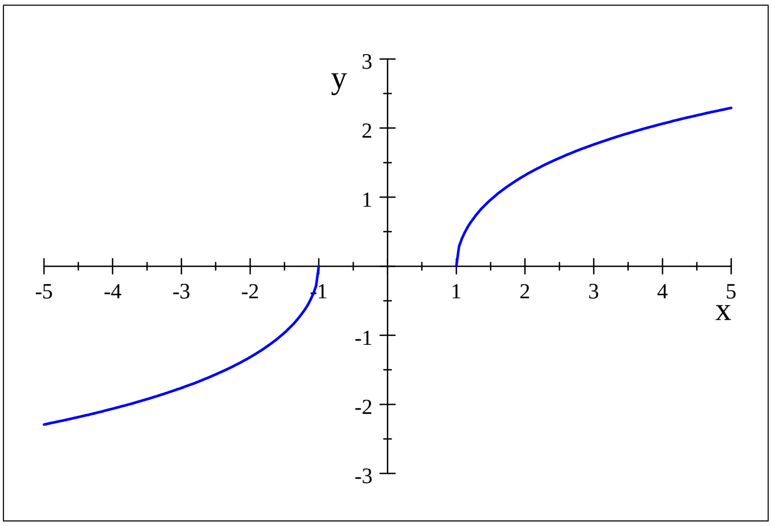
$$\int \frac{dx}{\sqrt{x^2-1}} = \operatorname{arcosh} x + C = \ln\left(x+\sqrt{x^2-1}\right) + C \text{ auf } (1,+\infty)\,, \tag{4.5}$$

$$\int \frac{dx}{1-x^2} = \operatorname{artanh} x + C = \frac{1}{2} \ln \frac{1+x}{1-x} + C \text{ auf } (-1,1)\,. \tag{4.6}$$

Die Identität (4.5) lässt sich wie folgt verallgemeinern:

$$\boxed{\int \frac{dx}{\sqrt{x^2-1}} = \ln\left|x+\sqrt{x^2-1}\right| + C} \text{ auf } (1,+\infty) \text{ und } (-\infty,-1)\,,$$

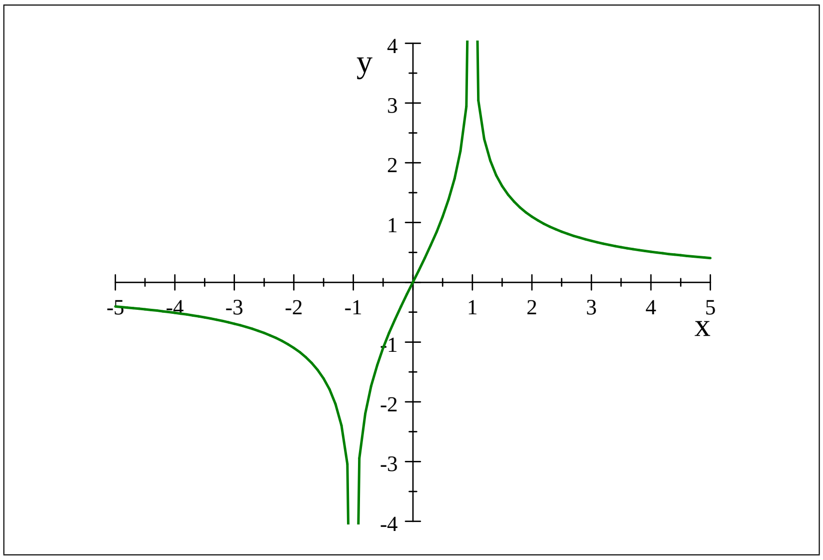
(siehe Aufgaben). Die Funktion $\ln\left|x+\sqrt{x^2-1}\right|$ auf $(1,+\infty)\cup(-\infty,-1)$ heißt der *lange Logarithmus*:



Die Identität (4.6) lässt sich auf die Intervalle $(-\infty,-1)\,,\ (-1,1)\,,\ (1,+\infty)$ wie folgt erweitern:

$$\boxed{\int \frac{dx}{1-x^2} = \frac{1}{2} \ln\left|\frac{1+x}{1-x}\right| + C}$$

(siehe Aufgaben). Die Funktion $\ln\left|\frac{1+x}{1-x}\right|$ heißt der *hohe Logarithmus:*



Die obigen Identitäten in den Rahmen liefern eine einfache Tabelle von Stammfunktionen, die auch *Integraltabelle* heißt. Die Einträge von dieser Tabelle heißen

Grundintegrale. Es gibt lange Tabellen von Stammfunktionen mit tausenden Einträgen. Es gibt auch viele Programme, die Stammfunktion explizit bestimmen können. Diese Programme benutzen die ausführlichen Tabellen von Stammfunktionen und bestimmte Rechenregeln.

Die Operation $f \mapsto \int f(x)\,dx$ heißt Integration oder Integrieren. Die Funktion $f(x)$ heißt der Integrand, die Variable x heißt die Integrationsvariable. Natürlich ist der Wert von Integral unabhängig von der Notation der Integrationsvariable.

Integrieren ist normalerweise viel schwieriger als Differenzieren. Darüber hinaus ist es nicht immer möglich das Integral explizit (durch elementare Funktionen) zu bestimmen; z.B. das ist der Fall für die Integrale $\int e^{x^2}dx$, $\int \frac{\sin x}{x}dx$ und andere.

Unterhalb entwickeln wir die Technik des Integrierens für explizite Berechnung von Integralen (wenn möglich). Diese Technik besteht aus drei allgemeinen Regeln – Linearität, partielle Integration und Substitution, die häufig ein gegebenes Integral auf Grundintegrale zurückzuführen helfen. Da Integrieren eine inverse Operation von Differenzieren ist, so erhält man meist die Rechenregeln von Integrieren als Umkehrung von den Rechenregeln von Differenzieren.

Es gibt auch spezielle Integrationsverfahren für spezielle Klassen von Integranden.

4.2 Linearität des unbestimmten Integrals

Satz 4.2 *Seien f und g zwei Funktion auf einem Intervall I. Existieren die beiden Integrale $\int f dx$ und $\int g dx$ auf I, so für alle $a, b \in \mathbb{R}$ (mit $a \neq 0$ oder $b \neq 0$) gilt*

$$\int (af + bg)\,dx = a\int f dx + b\int g dx. \tag{4.7}$$

Beweis. Wir müssen beweisen, dass die Ableitung der rechten Seite gleich $af + bg$ ist. Die Linearität von Ableitung (Satz 3.16) ergibt

$$\begin{aligned}\left(a\int f dx + b\int g dx\right)' &= a\left(\int f dx\right)' + b\left(\int g dx\right)' \\ &= af + bg\end{aligned}$$

was zu beweisen war. ∎

Beispiel. 1. Bestimmen wir $\int \left(x + \frac{1}{\sqrt{x}}\right)^2 dx$. Nach dem Binomischen Lehrsatz gilt es

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

und somit

$$\begin{aligned}\int \left(x + \frac{1}{\sqrt{x}}\right)^2 dx &= \int x^2 dx + 2\int x^{1/2}dx + \int \frac{1}{x}dx \\ &= \frac{x^3}{3} + \frac{4x^{3/2}}{3} + \ln|x| + C.\end{aligned}$$

2. Bestimmen wir $\int \frac{dx}{\sin^2 x \cos^2 x}$. Bemerken wir, dass

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

woraus folgt

$$\int \frac{dx}{\sin^2 x \cos^2 x} = \int \frac{dx}{\cos^2 x} + \int \frac{dx}{\sin^2 x} = \tan x - \cot x + C.$$

4.3 Partielle Integration

Seien $u(x)$ und $v(x)$ zwei Funktionen auf einem Intervall I. Ist v differenzierbar, so betrachten wir den Ausdruck

$$\int u dv \equiv \int u(x)\, dv(x) := \int u(x)\, v'(x)\, dx.$$

Satz 4.3 *Seien u, v zwei stetig differenzierbare Funktionen auf einem Intervall I. Dann gilt auf diesem Intervall die Identität*

$$\int u dv = uv - \int v du. \tag{4.8}$$

Beweis. Die Identität (4.8) lässt sich ausführlicher wie folgt umschreiben:

$$\int uv' dx = uv - \int vu' dx. \tag{4.9}$$

Da die Funktionen uv' und vu' stetig sind, so die beiden Integrale in (4.9) existieren. Um (4.9) zu beweisen, es reicht zu zeigen, dass die Ableitung der rechten Seite gleich uv' ist. In der Tat gilt es

$$\begin{aligned} \left(uv - \int vu' dx \right)' &= (uv)' - vu' \\ &= (u'v + uv') - vu' \\ &= uv', \end{aligned}$$

was zu beweisen war. ■

Bemerkung. Im Beweis haben wir die Produktregel (Satz 3.16) für die Ableitung $(uv)'$ benutzt. Somit kann die Identität (4.8) als eine Umkehrung von der Produktregel betrachtet werden.

Beispiel. 1. Bestimmen $\int \ln x dx$. Für $u = \ln x$ und $v = x$ haben wir

$$\int \ln x dx = x \ln x - \int x d \ln x = x \ln x - \int x \frac{1}{x} dx = x \ln x - x + C,$$

so dass

$$\boxed{\int \ln x dx = x \ln x - x + C}$$

2. Bestimmen $\int x^2 e^x dx$. Wir benutzen, dass $e^x dx = de^x$. Für $u = x^2$ und $v = e^x$ haben wir

$$\int x^2 e^x dx = \int x^2 de^x = x^2 e^x - \int e^x dx^2 = x^2 e^x - 2\int xe^x dx.$$

Um $\int xe^x dx$ zu bestimmen, wir benutzen den Satz 4.3 wieder, diesmal mit $u = x$ und $v = e^x$:

$$\int xe^x dx = \int xde^x = xe^x - \int e^x dx = xe^x - e^x + C.$$

Somit erhalten wir

$$\int x^2 e^x dx = x^2 e^x - 2xe^x + 2e^x + C.$$

3. Bestimmen $\int \sqrt{1+x^2}dx$. Für $u = \sqrt{1+x^2}$ und $v = x$, erhalten wir

$$\begin{aligned}\int \sqrt{1+x^2}dx &= x\sqrt{1+x^2} - \int \frac{x^2 dx}{\sqrt{1+x^2}} \\ &= x\sqrt{1+x^2} - \int \frac{(1+x^2)\, dx}{\sqrt{1+x^2}} + \int \frac{dx}{\sqrt{1+x^2}} \\ &= x\sqrt{1+x^2} - \int \sqrt{1+x^2}dx + \ln\left(x + \sqrt{x^2+1}\right) + C.\end{aligned}$$

Wir sehen, dass dasselbe Integral in den Beiden Seiten erscheint. Verschieben das Integral nach links und Dividieren durch 2 ergibt

$$\boxed{\int \sqrt{1+x^2}dx = \tfrac{1}{2}x\sqrt{x^2+1} + \tfrac{1}{2}\ln\left(x + \sqrt{x^2+1}\right) + C.}$$

4.4 Integration durch Substitution

Satz 4.4 *Sei u eine differenzierbare Funktion auf einem Intervall I mit $u(I) \subset J$ wobei J noch ein Intervall ist. Sei f eine Funktion auf J mit der Stammfunktion F. Dann gilt*

$$\int f(u(x))\, du(x) = F(u(x)) + C. \tag{4.10}$$

Beweis. Die beiden Funktionen $f(u(x))$ und $F(u(x))$ haben den Definitionsbereich I. Da

$$\int f(u(x))\, du(x) = \int f(u(x))\, u'(x)\, dx,$$

so ist die Identität (4.10) äquivalent zu

$$(F(u(x)))' = f(u(x))\, u'(x).$$

In der Tat erhalten wir nach der Kettenregel (Satz 3.17) und $F' = f$, dass

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

was zu beweisen war. ■

Bemerkung. Da F eine Stammfunktion von f ist, so gilt es

$$\int f(u)\,du = F(u) + C,$$

wobei u hier eine Integrationsvariable ist. Die Identität (4.10) besagt, dass die Integrationsvariable u hier durch eine differenzierbare Funktion $u = u(x)$ ersetzt werden kann. Die Formel (4.10) heißt *die Substitutionsregel* von Integration. Es ist klar aus dem Beweis, dass die Substitutionsregel eine Umkehrung von der Kettenregel ist.

Beispiel. 1. Bestimmen wir

$$\int (ax+b)^n\,dx$$

wobei $a \neq 0$. Da

$$d(ax+b) = adx$$

und somit

$$dx = \frac{1}{a}d(ax+b),$$

so erhalten wir mit der Substitution $u = ax + b$

$$\int (ax+b)^n\,dx = \frac{1}{a}\int (ax+b)^n\,d(ax+b) = \frac{1}{a}\int u^n du.$$

Andererseits gilt es

$$\int u^n du = \begin{cases} \dfrac{u^{n+1}}{n+1} + C, & n \neq -1 \\ \ln|u| + C, & n = -1. \end{cases}$$

woraus folgt

$$\int (ax+b)^n\,dx = \begin{cases} \dfrac{(ax+b)^{n+1}}{a(n+1)} + C, & n \neq 1, \\ \dfrac{1}{a}\ln|ax+b| + C, & n = -1. \end{cases}$$

2. Bestimmen wir

$$\int \frac{dx}{x^2-1}.$$

Das ist ein Grundintegral, aber trotzdem zeigen wir, wie man dieses Integral unabhängig berechnen kann. Da

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

so erhalten wir

$$\begin{aligned} \int \frac{dx}{x^2-1} &= \frac{1}{2}\int \frac{dx}{x-1} - \frac{1}{2}\int \frac{dx}{x+1} \\ &= \frac{1}{2}\int \frac{d(x-1)}{x-1} - \frac{1}{2}\int \frac{d(x+1)}{x+1} \\ &= \frac{1}{2}\ln|x-1| - \frac{1}{2}\ln|x+1| + C \\ &= \frac{1}{2}\ln\left|\frac{x-1}{x+1}\right| + C. \end{aligned}$$

3. Bestimmen wir

$$\int \frac{xdx}{1+x^2}.$$

Da

$$xdx = d\left(\frac{x^2}{2}\right) = \frac{1}{2} d\left(1+x^2\right),$$

so haben wir

$$\int \frac{xdx}{1+x^2} = \frac{1}{2} \int \frac{d\left(1+x^2\right)}{1+x^2}.$$

Substitution $u = 1 + x^2$ ergibt

$$\int \frac{xdx}{1+x^2} = \frac{1}{2} \int \frac{du}{u} = \frac{1}{2} \ln|u| + C = \frac{1}{2} \ln\left(1+x^2\right) + C.$$

4. Bestimmen wir

$$\int \frac{dx}{\sin x}.$$

Mit der Substitution $u = \cos x$ erhalten wir

$$\begin{aligned} \int \frac{dx}{\sin x} &= \int \frac{\sin x dx}{\sin^2 x} = -\int \frac{d\cos x}{\sin^2 x} = \int \frac{d\cos x}{\cos^2 x - 1} = \int \frac{du}{u^2-1} \\ &= \frac{1}{2} \ln\left|\frac{u-1}{u+1}\right| + C = \frac{1}{2} \ln \frac{1-\cos x}{1+\cos x} + C. \end{aligned}$$

Um die Antwort weiter vereinfachen zu können, benutzen wir die trigonometrische Identität

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

was ergibt

$$\boxed{\int \frac{dx}{\sin x} = \ln\left|\tan \tfrac{x}{2}\right| + C.}$$

5. Bestimmen wir

$$\int \arcsin x dx.$$

Wir haben:

$$\begin{aligned} \int \arcsin x dx &= x \arcsin x - \int x d \arcsin x \quad \text{(partielle Integration)} \\ &= x \arcsin x - \int \frac{x}{\sqrt{1-x^2}} dx \\ &= x \arcsin x + \frac{1}{2} \int \frac{d\left(1-x^2\right)}{\sqrt{1-x^2}} \quad \text{(Substitution } u = 1-x^2\text{)} \\ &= x \arcsin x + \frac{1}{2} \int u^{-1/2} du \quad \text{(Grundintegral)} \\ &= x \arcsin x + u^{1/2} + C \\ &= x \arcsin x + \sqrt{1-x^2} + C. \end{aligned}$$

Somit gilt es

$$\boxed{\int \arcsin x dx = x \arcsin x + \sqrt{1-x^2} + C}$$

4.5 Integration von rationalen Funktionen

Ein Polynom $P(x)$ über $\mathbb{R}$ ist eine Funktion der Art

$$P(x) = c_0 + c_1 x + ... + c_N x^N$$

wobei $N \in \mathbb{Z}_+$, $c_0, ..., c_n$ reellwertige Koeffizienten sind und x eine reellwertige Variable. Gilt $c_N \neq 0$ so heißt N der Grad des Polynoms.

Man kann die Polynome auch über $\mathbb{C}$ betrachten, d.h. mit komplexwertigen Koeffizienten und Variable $x \in \mathbb{C}$, aber in diesem Abschnitt brauchen wir nur die Polynome über $\mathbb{R}$.

Eine Funktion f heißt *rational* falls f als Quotient zweier Polynome darstellbar ist, d.h.

$$f(x) = \frac{P(x)}{Q(x)}$$

mit Polynomen $P(x)$ und $Q(x)$.

Integrationsverfahren für die rationalen Integrande basiert auf dem folgenden Satz.

Satz. *(a) Jedes Polynom $Q(x)$ mit reellwertigen Koeffizienten lässt sich wie folgt faktorisieren:*

$$Q(x) = A(x - r_1)^{k_1} ... (x - r_l)^{k_l} \left(x^2 + p_1 x + q_1\right)^{m_1} ... \left(x^2 + p_n x + q_n\right)^{m_n}, \quad (4.11)$$

wobei $n, l \in \mathbb{Z}_+$, $A, r_i, p_j, q_j \in \mathbb{R}$, $k_i, m_j \in \mathbb{N}$, und die Polynome $x^2 + p_j x + q_j$ keine reelle Nullstelle haben. Darüber hinaus alle Zahlen r_i sind verschieden und die Paaren (p_j, q_j) sind auch verschieden.

(b) (Partialbruchzerlegung) Jede rationale Funktion $f(x) = \frac{P(x)}{Q(x)}$ lässt sich wie folgt darstellen:

$$f(x) = R(x) + \sum_{i=1}^{l} \sum_{k=1}^{k_i} \frac{a_{ki}}{(x - r_i)^k} + \sum_{j=1}^{n} \sum_{m=1}^{m_j} \frac{b_{mj} x + c_{mj}}{(x^2 + p_j x + q_j)^m}, \quad (4.12)$$

wobei $R(x)$ ein Polynom ist, r_i, p_j, q_j, k_i, m_j sind wie in (4.11), und $a_{ki}, b_{mj}, c_{mj} \in \mathbb{R}$.

Der Teil (a) wird später in diesem Vorlesungskurs bewiesen und ist eine Folgerung des Fundamentalsatzes der Algebra. Teil (b) wird in Algebra mit Hilfe von Division von Polynomen bewiesen.

Die Identität (4.12) bedeutet, dass jede rationale Funktion sich darstellen lässt als die Summe von einem Polynom und einigen *Partialbrüchen*

$$\frac{a}{(x - r)^k}, \quad \frac{bx + c}{(x^2 + px + q)^m}.$$

Jeder lineare Faktor $(x - r_i)^{k_i}$ in (4.11) ergibt in (4.12) eine Summe

$$\frac{a_1}{(x - r_i)} + \frac{a_2}{(x - r_i)^2} + ... + \frac{a_{k_i}}{(x - r_i)^{k_i}},$$

und jeder quadratische Faktor $(x^2 + p_j x + q_j)^{m_j}$ in (4.11) ergibt in (4.12) eine Summe

$$\frac{b_1 x + c_1}{x^2 + p_j x + q_j} + \frac{b_2 x + c_2}{(x^2 + p_j x + q_j)^2} + ... + \frac{b_2 x + c_2}{(x^2 + p_j x + q_j)^{m_j}}.$$

Zum Beispiel, die rationale Funktion $f(x) = \frac{1}{(x-1)^2(x^2+1)}$ hat die folgende Partialbruchzerlegung:

$$\frac{1}{(x-1)^2(x^2+1)} = \frac{a_1}{x-1} + \frac{a_2}{(x-1)^2} + \frac{bx+c}{x^2+1},$$

wobei die Koeffizienten a_1, a_2, b, c noch bestimmt werden müssen (siehe unterhalb).

Wir benutzen die Partialbruchzerlegung (4.12) um das Integral $\int f(x)\,dx$ zu berechnen. Das Polynom $R(x)$ lässt sich einfach integrieren, da es eine Summe der Glieder der Form αx^k ist. Besprechen wir jetzt, wie integriert man die Partialbrüche.

Der *Partialbruch erster Art* $\frac{a}{(x-r)^k}$ wird wie folgt integriert:

$$\begin{aligned}
\int \frac{a}{(x-r)^k}dx &= a\int \frac{d(x-r)}{(x-r)^k} = (\text{Substitution } u = x - r) \\
&= a\int u^{-k}du = a\begin{cases} \frac{u^{1-k}}{1-k} + C, & k \neq 1, \\ \ln|u| + C, & k = 1, \end{cases} \\
&= a\begin{cases} \frac{(x-r)^{1-k}}{1-k} + C, & k \neq 1, \\ \ln|x-r| + C, & k = 1. \end{cases}
\end{aligned}$$

Integrieren den *Partialbruch zweiter Art* $\frac{bx+c}{(x^2+px+q)^m}$ ist schwieriger. Schreiben wir $x^2 + px + q$ in der Form

$$x^2 + px + q = (x + p/2)^2 + \left(q - p^2/4\right) = u^2 + s^2,$$

mit $u = x + p/2$ und $s = \sqrt{q - p^2/4} > 0$ und erhalten

$$\begin{aligned}
\int \frac{bx+c}{(x^2+px+q)^m}dx &= \int \frac{b'u + c'}{(u^2+s^2)^m}du \\
&= b'\int \frac{udu}{(u^2+s^2)^m} + c'\int \frac{du}{(u^2+s^2)^m}.
\end{aligned}$$

Somit bleibt es die folgenden Integrale zu berechnen:

$$\int \frac{udu}{(u^2+s^2)^m} \quad \text{und} \quad \int \frac{du}{(u^2+s^2)^m}. \tag{4.13}$$

Das erste Integral ist einfach:

$$\begin{aligned}
\int \frac{udu}{(u^2+s^2)^m} &= \frac{1}{2}\int \frac{d(u^2+s^2)}{(u^2+s^2)^m} \quad (\text{Substitution } v = u^2 + s^2) \\
&= \frac{1}{2}\int v^{-m}dv = \frac{1}{2}\begin{cases} \frac{v^{1-m}}{1-m} + C, & m \neq 1, \\ \ln|v| + C, & m = 1, \end{cases} \\
&= \frac{1}{2}\begin{cases} \frac{(u^2+s^2)^{1-m}}{1-m} + C, & m \neq 1, \\ \ln(u^2+s^2) + C, & m = 1. \end{cases}
\end{aligned}$$

Das zweite Integral in (4.13) lässt sich per Induktion nach m berechnen. Dafür setzen wir

$$F_m(u) = \int \frac{du}{(u^2+s^2)^m}$$

und bemerken, dass

$$F_1(u) = \int \frac{du}{u^2+s^2} = \int \frac{du}{s^2\left((u/s)^2+1\right)} = \frac{1}{s}\int \frac{d(u/s)}{(u/s)^2+1} = \frac{1}{s}\arctan\frac{u}{s} + C,$$

d.h.

$$\boxed{\int \frac{du}{u^2+s^2} = \tfrac{1}{s}\arctan\tfrac{u}{s} + C}.$$

Partielle Integration von F_m ergibt

$$\begin{aligned} F_m(u) &= \frac{u}{(u^2+s^2)^m} - \int u d\frac{1}{(u^2+s^2)^m} \\ &= \frac{u}{(u^2+s^2)^m} + 2m\int \frac{u^2 du}{(u^2+s^2)^{m+1}} \\ &= \frac{u}{(u^2+s^2)^m} + 2m\int \frac{u^2+s^2}{(u^2+s^2)^{m+1}}du - 2ms^2\int \frac{du}{(u^2+s^2)^{m+1}} \\ &= \frac{u}{(u^2+s^2)^m} + 2mF_m - 2ms^2F_{m+1}. \end{aligned}$$

Daraus folgt die Relation zwischen F_m und F_{m+1}

$$F_{m+1} = \frac{1}{2ms^2}\left(\frac{u}{(u^2+s^2)^m} + (2m-1)F_m\right),$$

und somit lässt F_m sich per Induktion nach m bestimmen.

Eine Funktion heißt *elementar* falls sie sich durch die Funktionen

$$\text{const}, x^a, e^x, \ln x, \sin x, \cos x, \arcsin x, \arctan x$$

mit Hilfe von Operationen Addition, Multiplikation, Division und Komposition darstellen lässt. Alle Funktionen, die wie bisher benutzt haben, sind elementar. Man sagt, dass eine Funktion f *elementar integrierbar* ist falls $\int f(x)\,dx$ eine elementare Funktion ist. Wir haben oberhalb schon viele Beispiele von elementar integrierbaren Funktionen gesehen. Andererseits, wie es schon erwähnt wurde, es gibt die elementaren Funktionen die nicht elementar integrierbar sind, z.B. e^{x^2} und $\frac{\sin x}{x}$.

Nach dem obigen Argument erhalten wir die folgende Behauptung.

Behauptung. *Jede rationale Funktion $f(x)$ ist elementar integrierbar.*

Beispiel. 1. Bestimmen wir

$$\int \frac{dx}{x^3-x}.$$

Der Nenner lässt sich wie folgt zerlegen:

$$x^3 - x = x(x-1)(x+1).$$

Somit has die Funktion $\frac{1}{x^3-x}$ die folgende Partialbruchzerlegung:

$$\frac{1}{x^3-x} = \frac{1}{x(x-1)(x+1)} = \frac{a}{x} + \frac{b}{x-1} + \frac{c}{x+1}, \tag{4.14}$$

wobei die Koeffizienten a, b, c noch bestimmt werden sollen. Diese Identität soll für alle $x \in \mathbb{R} \setminus \{0, 1, -1\}$ erfüllt werden. Um a zu bestimmen, multiplizieren wir (4.14) mit x:

$$\frac{1}{x^2-1} = a + x\left(\frac{b}{x-1} + \frac{c}{x+1}\right) \tag{4.15}$$

und bemerken, dass diese Identität für alle $x \in \mathbb{R} \setminus \{0, 1, -1\}$ gilt, genau wie (4.14). Aber die beiden Seiten von (4.15) sind auch für $x = 0$ definiert. Nach der Stetigkeit der beiden Seiten von (4.15) gilt die Gleichheit (4.15) auch in $x = 0$. Setzen wir in (4.15) $x = 0$ und erhalten

$$a = -1.$$

Analog ergibt Multiplizieren von (4.14) mit $x - 1$ die folgende Identität

$$\frac{1}{x(x+1)} = b + (x-1)\left(\frac{a}{x} + \frac{c}{x+1}\right),$$

woraus für $x = 1$ folgt

$$b = \frac{1}{2}.$$

Multiplizieren von (4.14) mit $x + 1$ ergibt

$$\frac{1}{x(x-1)} = c + (x+1)\left(\frac{a}{x} + \frac{b}{x-1}\right),$$

woraus für $x = -1$ folgt

$$c = \frac{1}{2}.$$

Somit erhalten wir

$$\frac{1}{x^3-x} = -\frac{1}{x} + \frac{1}{2}\frac{1}{x-1} + \frac{1}{2}\frac{1}{x+1}$$

und

$$\int \frac{dx}{x^3-x} = -\ln|x| + \frac{1}{2}\ln|x-1| + \frac{1}{2}\ln|x+1| + C.$$

2. Bestimmen wir das Integral

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

Die Partialbruchzerlegung der Funktion $f(x) = \frac{1}{(x^2+1)(x-1)^2}$ hat die Form

$$\frac{1}{(x^2+1)(x-1)^2} = \frac{a_1}{(x-1)^2} + \frac{a_2}{x-1} + \frac{bx+c}{x^2+1}, \tag{4.16}$$

wobei die Koeffizienten a_i, b, c noch bestimmt werden sollen. Multiplizieren diese Identität mit $(x-1)^2$ ergibt

$$\frac{1}{x^2+1} = a_1 + a_2(x-1) + (x-1)^2 g(x),$$

wobei $g(x) = \frac{bx+c}{x^2+1}$. Für $x = 1$ erhalten wir

$$a_1 = \frac{1}{2}.$$

Subtrahieren von (4.16) den Glied $\frac{1}{2}\frac{1}{(x-1)^2}$ ergibt

$$\begin{aligned}
\left(\frac{1}{(x^2+1)} - \frac{1}{2}\right)\frac{1}{(x-1)^2} &= \frac{a_2}{x-1} + g(x) \\
-\frac{1}{2}\frac{x^2-1}{(x^2+1)(x-1)^2} &= \frac{a_2}{x-1} + g(x) \\
-\frac{1}{2}\frac{x+1}{(x^2+1)(x-1)} &= \frac{a_2}{x-1} + g(x). \qquad (4.17)
\end{aligned}$$

Multiplizieren mit $x - 1$ ergibt

$$-\frac{1}{2}\frac{x+1}{x^2+1} = a_2 + (x-1)\,g(x),$$

woraus für $x = 1$ folgt

$$a_2 = -\frac{1}{2}.$$

Einsetzen von a_2 in (4.17) lässt uns $g(x)$ bestimmen wie folgt:

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

und somit

$$g(x) = \frac{1}{2}\frac{x^2+1-(x+1)}{(x^2+1)(x-1)} = \frac{1}{2}\frac{x(x-1)}{(x^2+1)(x-1)} = \frac{1}{2}\frac{x}{x^2+1},$$

so dass $g(x)$ wirklich die Form $\frac{bx+c}{x^2+1}$ mit $b = \frac{1}{2}$ und $c = 0$ hat. Somit erhalten wir die folgende Partialbruchzerlegung von f:

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

Jetzt können wir jedes Glied von $f(x)$ integrieren:

$$\begin{aligned}
\int \frac{dx}{(x-1)^2} &= \int \frac{d(x-1)}{(x-1)^2} \quad \text{(Substitute } u = x-1\text{)} \\
&= \int u^{-2}du = \frac{u^{-1}}{-1} + C \\
&= \frac{1}{1-x} + C,
\end{aligned}$$

$$\int \frac{1}{x-1}dx = \int \frac{d(x-1)}{x-1} = \ln|x-1| + C,$$

$$\int \frac{xdx}{x^2+1} = \frac{1}{2}\int \frac{d(x^2+1)}{x^2+1} = \frac{1}{2}\ln\left(x^2+1\right) + C,$$

woraus folgt

$$\int f(x)\,dx = \frac{1}{2}\frac{1}{1-x} - \frac{1}{2}\ln|x-1| + \frac{1}{4}\ln\left(x^2+1\right) + C.$$

3. Bestimmen wir

$$\int \frac{dx}{\left(x^2+2x+5\right)^2}.$$

Das Polynom x^2+2x+5 hat keine Nullstelle, und somit ist die Funktion $\frac{1}{\left(x^2+2x+5\right)^2}$ schon ein Partialbruch zweiter Art. Wir haben

$$x^2+2x+5 = (x+1)^2 + 4,$$

und mit Substitution $u = x+1$ erhalten wir

$$\int \frac{dx}{\left(x^2+2x+5\right)^2} = \int \frac{du}{\left(u^2+4\right)^2}.$$

Zuerst benutzen wir, dass

$$\int \frac{du}{u^2+4} = \frac{1}{2}\int \frac{d\,(u/2)}{(u/2)^2+1} = \frac{1}{2}\arctan\frac{u}{2} + C.$$

Partielle Integration ergibt

$$\begin{aligned}
\int \frac{du}{u^2+4} &= \frac{u}{u^2+4} - \int u d\frac{1}{u^2+4} \\
&= \frac{u}{u^2+4} + \int u \frac{2u}{\left(u^2+4\right)^2} du \\
&= \frac{u}{u^2+4} + 2\int \frac{u^2+4-4}{\left(u^2+4\right)^2} du \\
&= \frac{u}{u^2+4} + 2\int \frac{du}{u^2+4} - 8\int \frac{du}{\left(u^2+4\right)^2}.
\end{aligned}$$

Daraus folgt, dass

$$8\int \frac{1}{\left(u^2+4\right)^2} du = \frac{u}{u^2+4} + \int \frac{1}{u^2+4} du = \frac{u}{u^2+4} + \frac{1}{2}\arctan\frac{u}{2} + C,$$

und somit

$$\int \frac{1}{\left(u^2+4\right)^2} du = \frac{u}{8\left(u^2+4\right)} + \frac{1}{16}\arctan\frac{u}{2} + C,$$

und

$$\int \frac{dx}{\left(x^2+2x+5\right)^2} = \frac{x+1}{8\left(x^2+2x+5\right)} + \frac{1}{16}\arctan\frac{x+1}{2} + C.$$

4.6 Bestimmtes Integral (Riemann-Integral)

Sei $f(x)$ eine Funktion auf einem Intervall $[a,b]$ wobei $a,b \in \mathbb{R}$ und $a < b$. Wir definieren hier den Begriff von *bestimmten Integral*

$$\int_a^b f(x)\,dx. \tag{4.18}$$

Das ist eine reelle Zahl, die man als der Flächeninhalt unter dem Graph der Funktion f betrachten kann. Das Verfahren für Definition des bestimmten Integrals wurde von Riemann eingeführt. Somit heißt das Integral (4.18) auch *Riemann-Integral* oder *Riemannsches Integral*.

Eine *Zerlegung* von dem Intervall $[a,b]$ ist jede endliche streng monoton steigende Folge $\{x_k\}_{k=0}^n$ mit $n \in \mathbb{N}$, $x_0 = a$ und $x_n = b$, d.h.

$$a = x_0 < x_1 < x_2 ... < x_{n-1} < x_n = b.$$

Wir bezeichnen eine Zerlegung mit Z, d.h. Z ist die ganze Folge $\{x_k\}_{k=0}^n$ wie oberhalb.

Gegeben sei eine Zerlegung Z von $[a,b]$, betrachten wir noch eine Folge $\xi = \{\xi_k\}_{k=1}^n$ von Zahlen ξ_k mit $\xi_k \in [x_{k-1}, x_k]$. Dann heißen ξ_k die *Zwischenstellen* von Z.

Definition. Für jede Funktion $f : [a,b] \to \mathbb{R}$ und für jede Zerlegung Z von $[a,b]$ mit den Zwischenstellen ξ definieren wir die *Riemann-Summe* mit

$$S(f,Z,\xi) = \sum_{k=1}^n f(\xi_k)\,\Delta x_k,$$

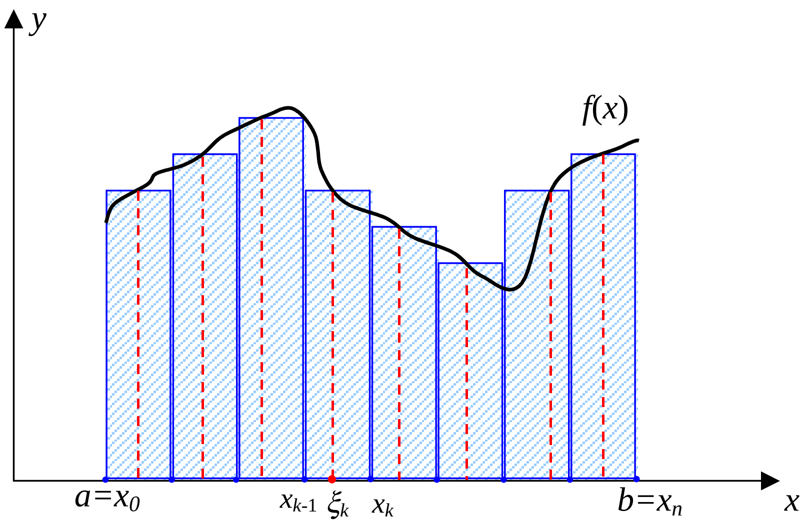
wobei

$$\Delta x_k := x_k - x_{k-1}.$$

Geometrische Bedeutung der Summe $S(f,Z,\xi)$ ist wie folgt. Ist f auf $[a,b]$ nichtnegative, so heißt die folgende Menge

$$\left\{(x,y) \in \mathbb{R}^2 : a \le x \le b,\ 0 \le y \le f(x)\right\}$$

der *Untergraph* von f. Die Riemann-Summe $S(f,Z,\zeta)$ ist gleich die Summe von Flächeninhalten von Rechtecken mit der Basis $[x_{k-1}, x_k]$ und der Höhe $f(\xi_k)$, was eine Annäherung von dem Flächeninhalt des Untergraphes von $f(x)$ ist.



Der Approximationsfehler dieser Annäherung wird fallen wenn die *Feinheit* der Zerlegung gegen 0 geht. Für jede Zerlegung $Z = \{x_k\}_{k=0}^n$ definieren wir die Feinheit von Z mit

$$\mu(Z) = \max_{1\le k\le n} \{\Delta x_k\}.$$

Der Grenzwert von Riemann-Summen $S(f, Z, \xi)$ für $\mu(Z) \to 0$ wird wie folgt definiert.

Definition. Wir schreiben

$$\lim_{\mu(Z)\to 0} S(f, Z, \xi) = A$$

mit einem $A \in \mathbb{R}$, fall für jedes $\varepsilon > 0$ existiert ein $\delta > 0$ so dass für jede Zerlegung Z mit $\mu(Z) < \delta$ und für jede Folge ξ von Zwischenstellen von Z gilt

$$|S(f, Z, \xi) - A| < \varepsilon.$$

Definition. Eine Funktion f auf $[a, b]$ heißt *Riemann-integrierbar* falls der Grenzwert

$$\lim_{\mu(Z)\to 0} S(f, Z, \xi)$$

existiert. Der Wert des Grenzwertes heißt das Riemann-Integral (=bestimmtes Intergal) von f und wird mit

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

bezeichnet.

In anderen Wörter, wir haben nach Definition

$$\int_a^b f(x)\, dx = \lim_{\mu(Z)\to 0} \sum_{k=1}^n f(\xi_k)\, \Delta x_k, \tag{4.19}$$

vorausgesetzt, dass lim existiert. Die Notation $\int_a^b f(x)\,dx$ wurde von by Leibniz vorgeschlagen und wurde so gewählt, um an die Summe $\sum_{k=1}^n f(\xi_k)\,\Delta x_k$ zu erinnern. Für nichtnegative Funktion f heißt der Wert von $\int_a^b f(x)\,dx$ auch der Flächeninhalt des Untergraphes von f.

Die folgenden Fragen sollen unterhalb betrachtet werden.

1. Welche Funktionen sind Riemann-integrierbar?

2. Wie kann man den Wert von Riemann-Integral bestimmen?

3. Wofür benutzt man das Riemann-Integral?

4. Welche Beziehung gibt es zum unbestimmten Integral $\int f(x)\,dx$?

Wir fangen mit zwei Beispielen an.

Beispiel. 1. Sei $f(x) \equiv c$ eine Konstantefunktion. Dann ist f Riemann-integrierbar da für jede Zerlegung Z mit Zwischenstellen ξ gilt

$$S(f,Z,\xi) = \sum_{k=1}^n f(\xi_k)\,\Delta x_k = c\sum_{k=1}^n \Delta x_k = c\,(b-a)$$

und somit der Grenzwert in (4.19) existiert und ist gleich $c\,(b-a)$, d.h.

$$\int_a^b c dx = c\,(b-a)\,. \tag{4.20}$$

2. Sei f die Dirichlet-Funktion

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

Dann ist f auf jedem Intervall $[a,b]$ nicht Riemann-integrierbar. Gegeben sei eine Zerlegung $Z = \{x_k\}_{k=0}^n$, wählen wir alle Zwischenstellen ξ_k irrational. Dann gilt $f(\xi_k) = 0$ und somit

$$S(f,Z,\xi) = 0.$$

Andererseits, für die gleiche Zerlegung wählen wir jetzt die anderen Zwischenstellen ξ_k so dass ξ_k rational sind. Dann gilt $f(\xi_k) = 1$ und somit

$$S(f,Z,\xi) = b-a.$$

Wir sehen, dass $\lim_{\mu(Z)\to 0} S(f,Z,\xi)$ nicht existiert.

4.7 Darboux-Integrierbarkeit

Gegeben sei eine Funktion $f:[a,b]\to\mathbb{R}$ und eine Zerlegung Z von $[a,b]$, definieren wir die *obere Darboux-Summe* von f und Z mit

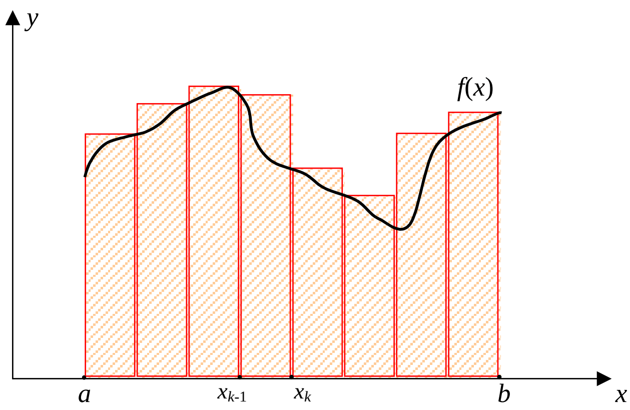
$$S^*(f,Z) = \sum_{k=1}^n (\sup_{[x_{k-1},x_k]} f)\Delta x_k$$

und die *untere Darboux-Summe* mit

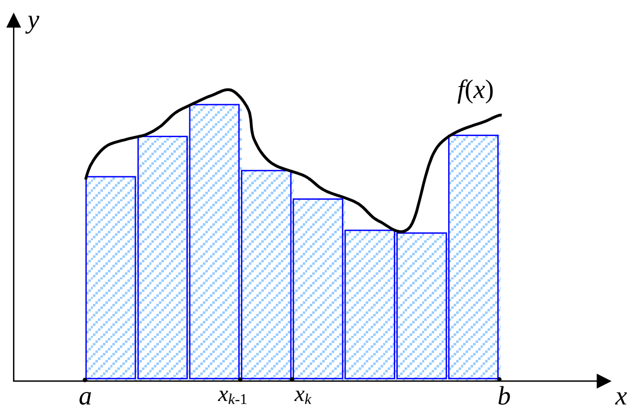
$$S_*(f,Z) = \sum_{k=1}^{n} (\inf_{[x_{k-1},x_k]} f)\Delta x_k.$$

Bemerken wir, dass $S^*(f,Z)$ Element von $(-\infty,+\infty]$ ist und $S_*(f,Z)$ Element von $[-\infty,+\infty)$ ist.

Die Darboux-Summen brauchen keine Zwischenstellen. Die obere Summe $S^*(f,Z)$ ist die Summe von Flächeninhalten von Rechtecken, die den Untergraph von f überdecken:



und die untere Summe $S_*(f,Z)$ ist die Summe von Flächeninhalten von Rechtecken, die im Untergraph enthalten werden:



Da für jede Zwischenstelle $\xi_k \in [x_{k-1},x_k]$ gilt

$$\inf_{[x_{k-1},x_k]} f \le f(\xi_k) \le \sup_{[x_{k-1},x_k]} f,$$

so erhalten wir, dass für jeder Wahl Zwischenstellen ξ von Z gilt

$$S_*(f,Z) \le S(f,Z,\xi) \le S^*(f,Z). \tag{4.21}$$

Definition. Funktion f heißt *Darboux-integrierbar* wenn

$$\lim_{\mu(Z)\to 0} (S^*(f,Z) - S_*(f,Z)) = 0,$$

d.h.

$$\forall \varepsilon > 0 \ \exists \delta > 0 \ \ \forall Z \ \text{ mit } \ \mu(Z) < \delta \text{ gilt } \left| S^*(f, Z) - S_*(f, Z) \right| < \varepsilon.$$

Satz 4.5 *Sei f eine Funktion auf $[a, b]$. Die folgenden drei Eigenschaften sind äquivalent:*

(a) *Funktion f ist Riemann-integrierbar.*

(b) *Funktion f ist Darboux-integrierbar.*

(c) *Die Grenzwerte $\lim_{\mu(Z)\to 0} S_*(f, Z)$ und $\lim_{\mu(Z)\to 0} S^*(f, Z)$ existieren in $\mathbb{R}$ und sind gleich.*

Darüber hinaus unter jeder von Bedingungen $(a), (b), (c)$ gelten die Identitäten:

$$\lim_{\mu(Z)\to 0} S^*(f, Z) = \lim_{\mu(Z)\to 0} S_*(f, Z) = \int_a^b f(x)\, dx = \sup_Z S_*(f, Z) = \inf_Z S^*(f, Z). \tag{4.22}$$

Definition. Die Funktion f heißt integrierbar fall f eine ($\Leftrightarrow$jede) von Bedingungen $(a), (b), (c)$ erfüllt.

Beweis. $(a) \Rightarrow (c)$ Setzen wir

$$A = \int_a^b f(x)\, dx.$$

Nach Definition, $\forall \varepsilon > 0 \ \exists \delta > 0$ so dass für jede Zerlegung Z mit $\mu(Z) < \delta$ und jede Folge von Zwischenstellen ξ gilt

$$|S(f, Z, \xi) - A| < \varepsilon.$$

Wählen wir die Zwischenstellen $\xi_k \in [x_{k-1}, x_k]$ so dass $f(\xi_k)$ sehr nahe zu $\sup_{[x_{k-1}, x_k]} f(x)$ ist. Dann der Wert von $S(f, Z, \xi)$ kann beliebig nahe zu $S^*(f, Z)$ sein, woraus folgt, dass

$$|S^*(f, Z) - A| < \varepsilon$$

und somit

$$\lim_{\mu(Z)\to 0} S^*(f, Z) = A. \tag{4.23}$$

Analog beweist man, dass

$$\lim_{\mu(Z)\to 0} S_*(f, Z) = A$$

und somit (c) gilt. Gleichzeitig haben wir auch die beiden linken Identitäten in (4.22) bewiesen.

$(c) \Rightarrow (b)$ Trivial.

$(b) \Rightarrow (a)$ Für zwei Zerlegungen Z und Z' von $[a, b]$ schreiben wir $Z' \subset Z$ falls Z' als Menge eine Teilmenge von Z ist. Man sagt, dass Z eine *Verfeinerung* von Z' ist.

Behauptung 1. *Im Fall $Z' \subset Z$ gelten die Ungleichungen*

$$S^*(f, Z) \leq S^*(f, Z')$$

und

$$S_*(f, Z) \geq S_*(f, Z').$$

D.h. die obere Summe von Verfeinerung fällt und die untere Summe steigt.

Seien $Z = \{x_i\}_{i=0}^{n}$ und $Z' = \{x'_k\}_{k=0}^{n'}$. Da $Z' \subset Z$, jedes Intervall $\left[x'_{k-1}, x'_k\right]$ von Z' stimmt mit einem Intervall $[x_l, x_m]$ überein so dass

$$x'_{k-1} = x_l < x_{l+1} < ... < x_m = x'_k.$$

Daraus folgt, dass

$$\begin{aligned}
\left(\sup_{[x_{k-1},x_k]} f\right)\left(x'_k - x'_{k-1}\right) &= \sum_{i=l+1}^{m} \left(\sup_{[x_{k-1},x_k]} f\right)(x_i - x_{i-1}) \\
&\geq \sum_{i=l+1}^{m} \left(\sup_{[x_{i-1},x_i]} f\right)(x_i - x_{i-1}).
\end{aligned}$$

Addieren solche Ungleichungen für alle k ergibt $S^*(f, Z') \geq S^*(f, Z)$. Die zweite Ungleichung wir analog bewiesen.

Behauptung 2. *Für zwei beliebige Zerlegungen Z' und Z'' von $[a, b]$ gilt*

$$S_*(f, Z') \leq S^*(f, Z''). \tag{4.24}$$

Die Vereinigung $Z = Z' \cup Z''$ ist auch eine Zerlegung von $[a, b]$. Da $Z' \subset Z$ und $Z'' \subset Z$, so erhalten wir nach Behauptung 1 und (4.21), dass

$$S_*(f, Z') \leq S_*(f, Z) \leq S^*(f, Z) \leq S^*(f, Z''),$$

woraus (4.24) folgt.

Jetzt beweisen wir, dass eine Darboux-integrierbare Funktion f auch Riemann-integrierbar ist. Setzen wir

$$A = \sup_Z S_*(f, Z) \quad \text{und} \quad B = \inf_Z S^*(f, Z). \tag{4.25}$$

Es folgt aus der Behauptung 2, dass

$$A \leq B. \tag{4.26}$$

Nach (4.25) und (4.26) gilt

$$S_*(f, Z) \leq A \leq B \leq S^*(f, Z). \tag{4.27}$$

Beweisen wir, dass

$$\lim_{\mu(Z) \to 0} S(f, Z, \xi) = A. \tag{4.28}$$

Nach (4.21) gilt

$$S_*(f,Z) \le S(f,Z,\xi) \le S^*(f,Z).$$

Vergleich mit (4.27) zeigt, dass die beiden Werte A und $S(f,Z,\xi)$ im gleichen Intervall $[S_*(f,Z), S^*(f,Z)]$ liegen, woraus folgt

$$|S(f,Z,\xi) - A| \le S^*(f,Z) - S_*(f,Z).$$

Da nach der Darboux-Integrabilität die rechte Seite für $\mu(Z) \to 0$ gegen 0 konvergiert, so erhalten wir (4.28) und somit auch die Riemann-Integrabilität von f.

Die dritte Identität von (4.22) folgt aus (4.28), und die vierte wird analog bewiesen, da der obigen Argument auch für B statt A funktioniert. ■

4.8 Integrierbare Funktionen

Korollar 4.6 (Notwendige Bedingung für Integrabilität) *Ist eine Funktion f auf $[a,b]$ integrierbar, so ist f auf $[a,b]$ beschränkt.*

Beweis. Nehmen wir das Gegenteil an, dass $\sup_{[a,b]} f = +\infty$. Für jede Zerlegung $Z = \{x_k\}_{k=0}^n$ gibt es ein Intervall $[x_{k-1}, x_k]$ wo $\sup_{[x_{k-1},x_k]} f = +\infty$, woraus folgt

$$S^*(f,Z) = +\infty.$$

Da immer $S_*(f,Z) < +\infty$, so erhalten wir

$$S^*(f,Z) - S_*(f,Z) = +\infty,$$

und f ist nicht Darboux-integrierbar. Somit muss es gelten $\sup_{[a,b]} f < +\infty$, und analog $\inf_{[a,b]} f > -\infty$, was zu beweisen war. ■

Der folgende Satz gibt uns viele Beispiele von integrierbaren Funktionen.

Satz 4.7 (Hinreichende Bedingungen für Integrierbarkeit)
(a) *Jede stetige Funktion f auf $[a,b]$ ist auf diesem Intervall integrierbar..*
(b) *Jede monotone Funktion f auf $[a,b]$ ist auf diesem Intervall integrierbar.*

Für dem Beweis von (a) brauchen wir den Begriff von gleichmäßiger Stetigkeit.

Definition. Eine Funktion f auf einem Intervall I heißt *gleichmäßig* stetig falls

$$\forall \varepsilon > 0 \;\; \exists \delta > 0 \;\; \forall x, y \in I \text{ mit } |x-y| < \delta \text{ gilt } |f(x) - f(y)| < \varepsilon. \tag{4.29}$$

Erinnern wir uns, dass f stetig auf I ist, falls f stetig an jeder Stelle $x \in I$, d.h. for any $x \in I$ for any $\varepsilon > 0$ there exists $\delta > 0$ such that

$$\forall x \in I \;\; \forall \varepsilon > 0 \; \exists \delta > 0 \;\;\; \forall y \in I \text{ mit } |x-y| < \delta \text{ gilt } |f(x) - f(y)| < \varepsilon. \tag{4.30}$$

In (4.30) hängt δ von x ab, wobei in (4.29) δ gleich für alle x ist, was das Wort "gleichmäßig" erklärt. Offensichtlich ist jede gleichmäßig stetige Funktion auch stetig, aber umgekehrt gilt es nicht immer.

Beispiel. Die Funktion $f(x) = \frac{1}{x}$ ist stetig auf $(0,1)$ aber nicht gleichmäßig stetig. In der Tat, für jedes $\delta > 0$, wählen wir $0 < x < \delta$ und $y = x/2$ so dass $|x - y| < \delta$, während

$$|f(x) - f(y)| = \left|\frac{1}{x} - \frac{1}{y}\right| = \frac{1}{x}$$

beliebige groß werden kann, da x beliebig nahe zum 0 gewählt werden kann.

Lemma 4.8 *Ist $f(x)$ stetig auf einem beschränkten abgeschlossen Intervall I, so ist f auf I gleichmäßig stetig.*

Beweis. Nehmen wir das Gegenteil an, dass f auf I nicht gleichmäßig stetig ist. Die Negation von (4.29) ergibt

$$\exists \varepsilon > 0 \ \ \forall \delta > 0 \ \ \exists x, y \in I \text{ mit } |x - y| < \delta \ \text{ und } \ |f(x) - f(y)| \geq \varepsilon.$$

Wählen wir $\delta = \frac{1}{n}$ für ein $n \in \mathbb{N}$ und somit bekommen $x_n, y_n \in I$ mit

$$|x_n - y_n| < \frac{1}{n} \tag{4.31}$$

und

$$|f(x_n) - f(y_n)| \geq \varepsilon. \tag{4.32}$$

Nach dem Satz von Bolzano-Weierstraß (Satz 2.14) hat die Folge $\{x_n\}$ eine konvergente Teilfolge $\{x_{n_k}\}$. Nach (4.31) erhalten wir, dass auch $\{y_{n_k}\}$ konvergiert und

$$\lim_{k\to\infty} x_{n_k} = \lim_{k\to\infty} y_{n_k}.$$

Nach der Stetigkeit von f daraus folgt

$$\lim_{k\to\infty} f(x_{n_k}) = \lim_{k\to\infty} f(y_{n_k}),$$

was im g zu (4.32) steht. ■

Beweis von Satz 4.7(a)**.** Nach Lemma 4.8 ist die Funktion f gleichmäßig stetig auf $[a,b]$, d.h. sie die Bedingung (4.29) erfüllt. Betrachten wir beliebige Zerlegung $Z = \{x_k\}_{k=0}^n$ von $[a,b]$ mit $\mu(Z) < \delta$. Für jede $x, y \in [x_{k-1}, x_k]$ gilt $|x - y| < \delta$ und somit

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

Daraus folgt, dass

$$\sup_{[x_{k-1},x_k]} f - \inf_{[x_{k-1},x_k]} f \leq \varepsilon$$

und somit

$$\begin{aligned} S^*(f,Z) - S_*(f,Z) &= \sum_{k=1}^{n} \Big(\sup_{[x_{k-1},x_k]} f - \inf_{[x_{k-1},x_k]} f\Big)(x_k - x_{k-1}) \\ &\leq \varepsilon \sum_{k=1}^{n} (x_k - x_{k-1}) \\ &= \varepsilon(b-a). \end{aligned}$$

Da $\varepsilon\,(b-a)$ beliebig klein sein kann, so erhalten

$$\lim_{\mu(Z)\to 0}\left(S^*\left(f,Z\right)-S_*\left(f,Z\right)\right)=0.$$

Somit ist f Darboux-integrierbar und nach dem Satz 4.5 auch integrierbar. ∎

Beweis von Satz 4.7(b)**.** Sei f monoton steigend. Für jede Zerlegung $Z=\{x_k\}_{k=0}^n$ von $[a,b]$ gilt

$$\sup_{[x_{k-1},x_k]} f = f\left(x_k\right) \quad\text{und}\quad \inf_{[x_{k-1},x_k]} f = f\left(x_{k-1}\right)$$

woraus folgt

$$\begin{aligned}
S^*\left(f,Z\right)-S_*\left(f,Z\right) &= \sum_{k=1}^n\left(\sup_{[x_{k-1},x_k]} f-\inf_{[x_{k-1},x_k]} f\right)\left(x_k-x_{k-1}\right)\\
&= \sum_{k=1}^n\left(f\left(x_k\right)-f\left(x_{k-1}\right)\right)\left(x_k-x_{k-1}\right)\\
&\le \mu\left(Z\right)\sum_{k=1}^n\left(f\left(x_k\right)-f\left(x_{k-1}\right)\right)\\
&= \mu\left(Z\right)\left(f\left(b\right)-f\left(a\right)\right).
\end{aligned}$$

Somit erhalten wir

$$\lim_{\mu(Z)\to 0}\left(S^*\left(f,Z\right)-S_*\left(f,Z\right)\right)=0,$$

woraus die Integrabilität von f folgt. ∎

4.9 Fundamentalsatz der Analysis, I

Der nächste Satz etabliert eine Beziehung zwischen bestimmten und unbestimmten Integralen.

Hauptsatz 4.9 (Fundamentalsatz der Analysis Teil 1) *Sei $f\left(x\right)$ eine stetige Funktion auf einem beschränkten abgeschlossenen Intervall $[a,b]$. Sei $F\left(x\right)$ eine Stammfunktion von f auf diesem Intervall. Dann gilt die Identität*

$$\int_a^b f\left(x\right)dx = F\left(b\right)-F\left(a\right). \tag{4.33}$$

Die Identität (4.33) heißt auch *Newton-Leibniz-Formel.* Führen wir die folgende Notation ein:

$$\left[F\right]_a^b := F\left(b\right)-F\left(a\right).$$

Da $F=\int f\left(x\right)dx$, so lässt die Newton-Leibniz-Formel wie folgt umschreiben:

$$\int_a^b f\left(x\right)dx = \left[\int f\left(x\right)dx\right]_a^b.$$

In dieser Form liefert die Newton-Leibniz-Formel eine direkte Beziehung zwischen bestimmten und unbestimmten Integralen.

Man kann (4.33) auch wie folgt umschreiben:

$$\int_a^b F'(x)\,dx = F(b) - F(a) = [F]_a^b .$$

Beweis. Nach dem Satz 4.7 ist f integrierbar. Nach Definition von Riemann-Integral haben wir

$$\int_a^b f(x)\,dx = \lim_{\mu(Z)\to 0} S(f, Z, \xi).$$

Fixieren wir eine Zerlegung $Z = \{x_k\}_{k=0}^n$ von $[a,b]$ und wählen die Zwischenstellen $\xi = \{\xi_k\}_{k=1}^n$ wie folgt. Nach dem Mittelwertsatz (Satz 3.23 aus Analysis I), es gibt $\xi_k \in [x_{k-1}, x_k]$ mit

$$F(x_k) - F(x_{k-1}) = F'(\xi_k)(x_k - x_{k-1}).$$

Für diese ξ_k erhalten wir

$$\begin{aligned} S(f, Z, \xi) &= \sum_{k=1}^n f(\xi_k)(x_k - x_{k-1}) \\ &= \sum_{k=1}^n F'(\xi_k)(x_k - x_{k-1}) \\ &= \sum_{k=1}^n (F(x_k) - F(x_{k-1})) \\ &= F(b) - F(a). \end{aligned}$$

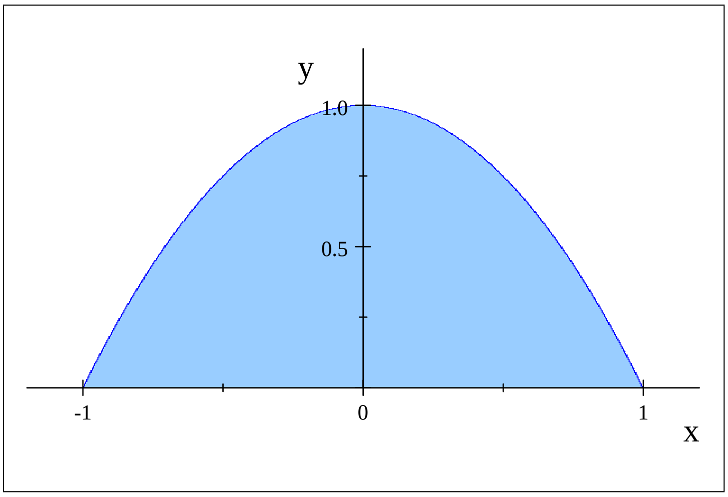
Somit kann der Grenzwert $\lim_{\mu(Z)\to 0} S(f, Z, \xi)$ nur den Wert $F(b) - F(a)$ annehmen, woraus (4.33) folgt. ■

Mit Hilfe von der Newton-Leibniz-Formel kann man die Integrale $\int_a^b f(x)\,dx$ effektiv berechnen. Ist $f \geq 0$, so definieren wir den Flächeninhalt des Untergraphes von f als $\int_a^b f(x)\,dx$.

Beispiel. 1. Für die Funktion $f(x) = 1 - x^2$ erhalten wir

$$\int_{-1}^1 \left(1 - x^2\right) dx = \left[\int \left(1 - x^2\right) dx\right]_{-1}^1 = \left[x - \frac{x^3}{3}\right]_{-1}^1 = \frac{4}{3}.$$

Geometrisch bedeutet dies, dass der Flächeninhalt zwischen der Parabel $y = 1 - x^2$ und der Achse x gleich $4/3$ ist.



Insbesondere beträgt dieser Flächeninhalt genau $2/3$ von den Flächeninhalt von dem kleinsten umschriebenen Rechteck $[-1,1] \times [0,1]$. Diese Regel von $\frac{2}{3}$ wurde erst von Archimedes entdeckt. Er konnte den Flächeninhalt direkt als der Grenzwert von Riemann-Summen berechnen, ohne Newton-Leibniz-Formel zu wissen.

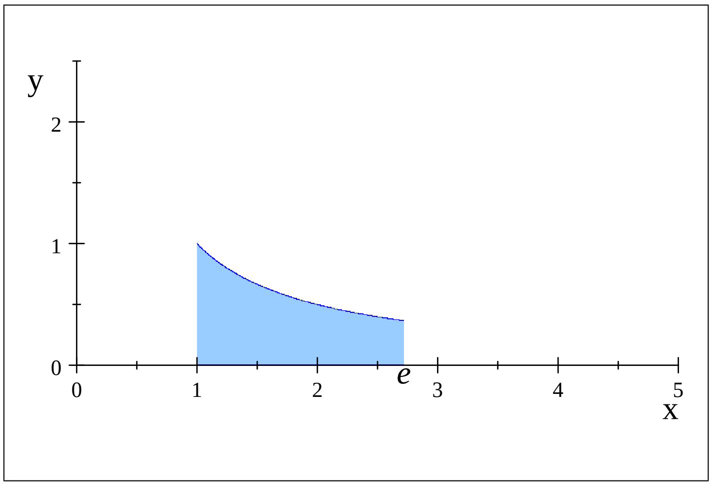
2. Sei $f(x) = \frac{1}{x}$. Für alle $a > 1$ erhalten wir

$$\int_1^a \frac{dx}{x} = \left[\int \frac{dx}{x} \right]_1^a = [\ln x]_1^a = \ln a. \tag{4.34}$$

Die Identität (4.34) lässt sich als eine unabhängige Definition von $\ln a$ benutzen. Danach kann man auch eine neue Definition von exp geben als inverse Funktion von ln . Insbesondere erhält man eine neue Definition von e als die Zahl mit der Eigenschaft

$$\int_1^e \frac{dx}{x} = 1.$$

Der Graph der Funktion $y = \frac{1}{x}$ heißt Hyperbel. Auf dem nächsten Bild ist der Flächeninhalt unter der Hyperbel auf $[1, e]$ gleich 1.



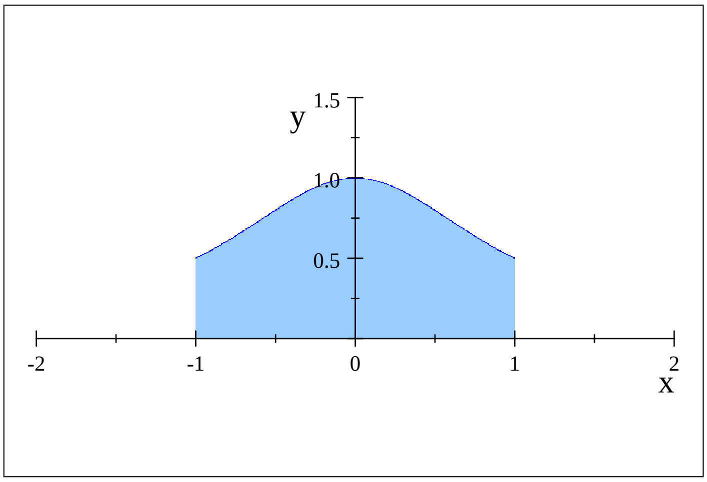
3. Für Funktion $f(x) = \frac{1}{1+x^2}$ und $a > 0$ haben wir

$$\int_0^a \frac{dx}{1+x^2} = \left[\int \frac{dx}{1+x^2} \right]_0^a = [\arctan x]_0^a = \arctan a.$$

Diese Identität lässt sich als unabhängige Definition von arctan benutzen. Dann gibt man an eine neue Definition von π durch

$$\frac{\pi}{4} = \arctan 1 = \int_0^1 \frac{dx}{1+x^2}.$$

Auf dem nächsten Bild ist der Flächeninhalt des Untergraphes von $f(x) = \frac{1}{1+x^2}$ auf $[-1, 1]$ gleich $\frac{\pi}{2}$.



Darüber hinaus kann man eine neue Definition von tan als inverse Funktion von arctan erhalten.

Bemerkung. Erweitern wir die Definition von Integral $\int_a^b f(x)\,dx$ zum Fall $a \geq b$ wie folgt:

$$\int_a^a f(x)\,dx = 0, \tag{4.35}$$

$$\int_a^b f(x)\,dx = -\int_b^a f(x)\,dx \text{ für } a > b. \tag{4.36}$$

Die Newton-Leibniz-Formel gilt dann für beliebige $a, b \in \mathbb{R}$: ist eine Funktion F stetig differenzierbar auf einem Intervall I, so gilt für alle $a, b \in I$

$$\int_a^b F'(x)\,dx = [F]_a^b := F(b) - F(a).$$

Für $a < b$ was dies im Satz 4.9 bewiesen. Für $a = b$ sind die beiden Seiten gleich 0. Für $a > b$ gilt

$$\int_a^b F'(x)\,dx = -\int_b^a F'(x)\,dx = -[F]_b^a = [F]_a^b.$$

Die Zahlen a und b heißen *untere* bzw *obere Grenzen* des Integrals $\int_a^b f(x)\,dx$.

4.10 Weitere Eigenschaften vom bestimmten Integral

Satz 4.10 (Linearität) *Sind die Funktionen f und g integrierbar auf $[a,b]$, so ist auch $\alpha f+\beta g$ integrierbar für alle $\alpha,\beta\in\mathbb{R}$ und*

$$\int_a^b (\alpha f+\beta g)\,dx=\alpha\int_a^b f dx+\beta\int_a^b g dx. \tag{4.37}$$

Beweis. Sei $a<b$. Sei Z eine Zerlegung von $[a,b]$ und ξ eine Folge von Zwischenstellen von Z. Dann

$$\begin{aligned} S(\alpha f+\beta g,Z,\xi) &= \sum_{k=0}^{n}(\alpha f+\beta g)(\xi_k)\,\Delta x_k \\ &= \alpha\sum_{k=0}^{n} f(\xi_k)\,\Delta x_k+\beta\sum_{k=0}^{n} g(\xi_k)\,\Delta x_k \\ &= \alpha S(f,Z,\xi)+\beta S(g,Z,\xi)\,. \end{aligned}$$

Für $\mu(Z)\to 0$ konvergiert die rechte Seite gegen

$$\alpha\int_a^b f(x)\,dx+\beta\int_a^b g(x)\,dx.$$

Somit ist $\alpha f+\beta g$ Riemann-integrierbar und (4.37) gilt. ■

Satz 4.11 (Partielle Integration im bestimmten Integral) *Für stetig differenzierbare Funktionen u,v auf einem Intervall $[a,b]$ gilt*

$$\int_a^b u dv=[uv]_a^b-\int_a^b v du. \tag{4.38}$$

Beweis. Wir haben nach dem Produktregel

$$(uv)'=uv'+vu'.$$

Daraus folgt, dass

$$\begin{aligned} \int_a^b (uv)'\,dx &= \int_a^b uv' dx+\int_a^b vu' dx \\ &= \int_a^b u dv+\int_a^b v du. \end{aligned}$$

Nach dem Satz 4.9 gilt

$$\int_a^b (uv)'\,dx=[ub]_a^b\,,$$

woraus (4.38) folgt. ■

Beispiel. Bestimmen wir $\int_0^\pi e^x \cos x dx$. Da $e^x dx = de^x$, so erhalten wir nach (4.38) mit $u = \cos x$ und $v = e^x$:

$$\begin{aligned}\int_0^\pi e^x \cos x dx &= \left[e^x \cos x\right]_0^\pi - \int_0^\pi e^x d\cos x \\ &= -e^\pi - 1 + \int_0^\pi e^x \sin x dx \\ &= -(e^\pi + 1) + \int_0^\pi \sin x de^x \\ &= -(e^\pi + 1) + \left[e^x \sin x\right]_0^\pi - \int_0^\pi e^x \cos x dx,\end{aligned}$$

woraus folgt

$$\int_0^\pi e^x \cos x dx = -\frac{e^\pi + 1}{2}.$$

Alternativ kann man zunächst das unbestimmte Integral $\int e^x \cos x$ berechnen und erst danach die Newton-Leibniz-Formel anwenden.

Satz 4.12 *Sei $a < b$.*

(a) (Positivität) *Ist f integrierbar auf $[a, b]$ und $f \geq 0$, so gilt*

$$\int_a^b f dx \geq 0.$$

(b) (Monotonie) *Sind f und g integrierbar auf $[a, b]$ und $f \geq g$, so gilt*

$$\int_a^b f dx \geq \int_a^b g dx.$$

Beweis. (a) Für nichtnegative Funktion f sind alle Riemann-Summen auch nichtnegativ, woraus die Aussage folgt.

(b) Nach dem Satz 4.10 und (a) gilt

$$\int_a^b f dx = \int_a^b (f - g)\, dx + \int_a^b g dx \geq \int_a^b g dx.$$

■

Korollar 4.13 *Sei $a < b$. Ist f integrierbar auf $[a, b]$ so gilt*

$$(b - a) \inf_{[a,b]} f \leq \int_a^b f dx \leq (b - a) \sup_{[a,b]} f. \tag{4.39}$$

Beweis. Sei $s = \sup_{[a,b]} f$. Dann $f \leq s$ auf $[a, b]$, und nach dem Satz 4.12 erhalten wir

$$\int_a^b f dx \leq \int_a^b s dx = s\,(b - a)\,,$$

wobei wir auch (4.20) benutzt haben. Die untere Abschätzung wird analog bewiesen.

■

Korollar 4.14 *Sei f eine stetige Funktion auf $[a,b]$. Dann gilt*

$$\left|\int_a^b f dx\right| \le \int_a^b |f|\, dx. \tag{4.40}$$

Beweis. Nach dem Satz 4.7 sind f und $|f|$ integrierbar. Da

$$-|f| \le f \le |f|,$$

so erhalten wir nach dem Satz 4.12

$$-\int_a^b |f|\, dx \le \int_a^b f dx \le \int_a^b |f|\, dx,$$

woraus (4.40) folgt. ■

Bemerkung. Es gilt eine stärkere Aussage: ist f auf $[a,b]$ integrierbar so ist $|f|$ auch integrierbar und (4.40) gilt (siehe Aufgaben).

Satz 4.15 (Mittelwertsatz für Integration) *Ist f stetig auf $[a,b]$, so existiert ein $c \in [a,b]$ mit*

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

Beweis. Der Fall $a = b$ ist trivial. Betrachten wir den Fall $a < b$ (und der Fall $a > b$ ist analog). Nach dem Satz 4.7 ist f integrierbar. Setzen wir

$$m = \inf_{[a,b]} f \quad \text{und} \quad M = \sup_{[a,b]} f.$$

Nach dem Korollar 3.9 aus Analysis I (eine Folgerung aus dem Extremwertsatz und Zwischenwertsatz) ist das Bild von f gleich das Intervall $[m,M]$. Da nach (4.39)

$$\frac{1}{b-a}\int_a^b f dx \in [m,M],$$

so es gibt ein $c \in [a,b]$ mit

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

was zu beweisen war. ■

Satz 4.16 (Additivität) *Sei f eine integrierbare Funktion auf einem Intervall $[a,b]$ mit $a < b$. Dann, für jedes $c \in (a,b)$, ist f auf den Intervallen $[a,c]$ und $[c,b]$ integrierbar und es gilt*

$$\int_a^b f dx = \int_a^c f dx + \int_c^b f dx. \tag{4.41}$$

Beweis. Seien Z' und Z'' die Zerlegungen von $[a,c]$ bzw $[c,b]$. Dann ist $Z = Z' \cup Z''$ eine Zerlegung von $[a,b]$ und es gilt

$$\mu(Z) = \max(\mu(Z'), \mu(Z''))$$

und

$$S^*(f,Z) = S^*(f,Z') + S(f,Z''). \tag{4.42}$$

Die gleiche Identität gilt für S_*. Da f auf $[a,b]$ integrierbar, so haben wir

$$\lim_{\mu(Z)\to 0} (S^*(f,Z) - S_*(f,Z)) = 0. \tag{4.43}$$

Da

$$S^*(f,Z) - S_*(f,Z) = (S^*(f,Z') - S_*(f,Z')) + (S^*(f,Z'') - S_*(f,Z'')),$$

daraus folgt, dass

$$\lim_{\mu(Z')\to 0} (S^*(f,Z') - S_*(f,Z')) = 0 = \lim_{\mu(Z'')\to 0} (S^*(f,Z'') - S_*(f,Z'')).$$

Somit ist f integrierbar auf $[a,c]$ und $[c,b]$. Für $\mu(Z) \to 0$ erhalten wir aus (4.42)

$$\lim_{\mu(Z)\to 0} S^*(f,Z) = \lim_{\mu(Z')\to 0} S^*(f,Z') + \lim_{\mu(Z'')\to 0} S^*(f,Z'').$$

Nach der Identität (4.22) aus dem Satz 4.5 erhalten wir (4.41). ■

Korollar 4.17 *Sei f eine stetige Funktion auf einem Integral I. Dann für alle $a,b,c \in I$ gilt* (4.41).

Diese Aussage gilt auch für integrierbaren Funktionen f.

Beweis. Für $c = a$ oder $c = b$ gilt (4.41) trivial nach (4.35). Für $a = b$ ist (4.41) äquivalent zu

$$0 = \int_a^c f dx + \int_c^a f dx,$$

was nach (4.36) gilt. Seien jetzt a,b,c verschieden. Dann gibt es 6 Fälle wie folgt:

1. $a < c < b$
2. $a < b < c$
3. $b < a < c$
4. $b < c < a$
5. $c < a < b$
6. $c < b < a$

Der Fall $a < c < b$ gilt nach Satz 4.16. Im Fall $a < b < c$ haben wir nach dem Satz 4.16

$$\int_a^c f dx = \int_a^b f dx + \int_b^c f dx,$$

woraus folgt

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

Andere Fälle werden analog betrachtet. ■

4.11 Fundamentalsatz der Analysis, II

Hauptsatz 4.18 (Fundamentalsatz der Analysis Teil 2) *Ist f eine stetige Funktion auf einem Intervall $I \subset \mathbb{R}$, so für jedes $c \in I$ ist die Funktion*

$$F(x) = \int_c^x f(t)\, dt$$

eine Stammfunktion von f. Insbesondere has jede stetige Funktion eine Stammfunktion.

Beweis. Da f stetig ist, so ist das Integral $\int_c^x f(t)\, dt$ nach dem Satz 4.7 wohldefiniert. Wir müssen beweisen, dass $F'(x) = f(x)$ für jedes $x \in I$, d.h.

$$\lim_{\substack{y \in I \setminus \{x\} \\ y \to x}} \frac{F(y) - F(x)}{y - x} = f(x). \tag{4.44}$$

Nach der Additivität des Integrals gilt

$$\begin{aligned} F(y) - F(x) &= \int_c^y f(t)\, dt - \int_c^x f(t)\, dt \\ &= \int_c^y f(t)\, dt + \int_x^c f(t)\, dt \\ &= \int_x^y f(t)\, dt. \end{aligned}$$

Nach dem Satz 4.15 gibt es ein $\xi \in [x, y]$ mit

$$\frac{F(y) - F(x)}{y - x} = f(\xi).$$

Nach der Stetigkeit von f erhalten wir

$$f(\xi) \to f(x) \quad \text{für } y \to x,$$

woraus (4.44) folgt. ■

Die Sätze 4.9 und 4.18 ergeben folgendes.

Korollar 4.19 *Für jede stetige Funktion f auf dem Intervall $[a, b]$ existiert eine Stammfunktion F auf $[a, b]$ und es gilt*

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

Die Existenz der Stammfunktion auf dem ganzen Intervall $[a, b]$ ist wichtig. Hier ist ein Gegenbeispiel, wie man falsches Ergebnis erhält:

$$\int_{-1}^1 \frac{dx}{x} = \left[\int \frac{dx}{x}\right]_{-1}^1 = [\ln|x|]_{-1}^1 = 0.$$

Warum ist diese Berechnung falsch? Die Stammfunktion $\ln|x|$ ist nicht auf dem ganzen Intervall $[-1, 1]$ definiert wie die Newton-Leibniz-Formel anfordert, sondern auf zwei disjunkten Intervallen $(0, +\infty)$ und $(-\infty, 0)$. Somit lässt diese Formel sich für Berechnung von $\int_a^b \frac{dx}{x}$ nur dann anwenden wenn entweder $[a, b] \subset (0, \infty)$ oder $[a, b] \subset (-\infty, 0)$. Darüber hinaus ist die Funktion $\frac{1}{x}$ auf $[-1, 1]$ nicht Riemann-integrierbar, da diese Funktion nicht beschränkt ist.

4.12 Substitutionsregel

Hauptsatz 4.20 (Substitutionsregel im bestimmten Integral) *Seien u eine stetig differenzierbare Funktion auf einem Intervall $I = [a, b]$ und f eine stetige Funktion auf dem Intervall $u(I)$. Dann gilt*

$$\int_a^b f(u(x))\, du(x) = \int_{u(a)}^{u(b)} f(y)\, dy. \tag{4.45}$$

Beweis. Das Bild $u(I)$ ist abgeschlossenes beschränktes Intervall (Theorem 3.8 from Analysis I). Da f auf $u(I)$ stetig ist, so hat f nach dem Satz 4.18 eine Stammfunktion F auf diesem Intervall. Somit gilt nach der Newton-Leibniz-Formel

$$\int_{u(a)}^{u(b)} f(y)\, dy = [F]_{u(a)}^{u(b)} = F(u(b)) - F(u(a)).$$

Nach dem Kettenregel gilt

$$(F \circ u)'(x) = F'(u(x))\, u'(x) = f(u(x))\, u'(x), \tag{4.46}$$

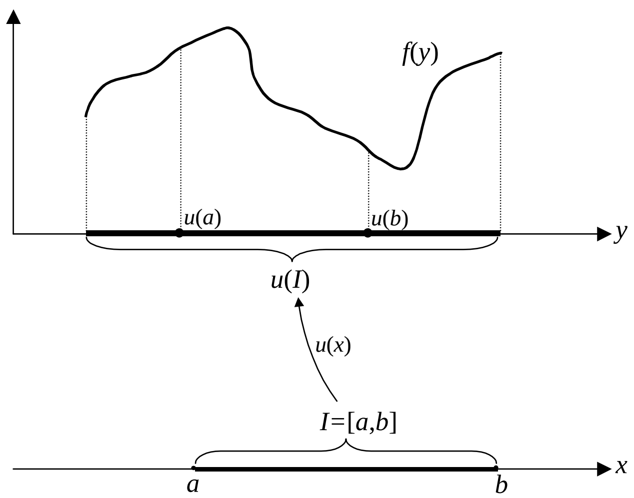
und somit nach der Newton-Leibniz-Formel

$$\int_a^b f(u(x))\, du(x) = \int_a^b f(u(x))\, u'(x)\, dx = [F \circ u]_a^b = F(u(b)) - F(u(a)). \tag{4.47}$$

Vergleichen von (4.46) und (4.47) beweist (4.45). ∎

Der Satz 4.4 gilt auch für integrierbare Funktionen f, aber der Beweis in diesem Fall ist deutlich komplizierter.

Die Identität (4.45) lässt sich als eine Substitution $y = u(x)$ im linken Integral betrachten. Man stellt zuerst den Integrand im gegebenen Integral als $f(u(x))\, u'(x)$ dar. Danach ersetz man $u(x)$ zwei mal durch y und die Grenzen a und b durch $u(a)$ bzw $u(b)$ (die Werte $u(a)$ und $u(b)$ sollen die Grenzen von dem Intervall $u(I)$ nicht unbedingt sein).



Alternativ betrachtet man das rechte Integral

$$\int_A^B f(y)\,dy$$

als gegeben und ersetzt die Variable y zwei mal mit einer Funktion $u(x)$. Man bestimmt a, b aus der Gleichungen

$$A = u(a), \quad B = u(b)$$

und erhält

$$\int_A^B f(y)\,dy = \int_a^b f(u(x))\,du(x).$$

Beispiel. 1. Bestimmen wir $\int_1^2 \frac{dx}{e^x-1}$. Man versucht den Integrand $\frac{dx}{e^x-1}$ in der Form $f(u(x))\,du(x)$ mit einer Funktion $u(x)$ darzustellen, so dass $f(u)$ eine einfachere Funktion ist. Wir haben

$$\int_1^2 \frac{dx}{e^x-1} = \int_1^2 \frac{e^x dx}{e^x(e^x-1)} = \int_1^2 \frac{de^x}{e^x(e^x-1)}.$$

Somit ergibt die Substitution $y = u(x) := e^x$

$$\begin{aligned}
\int_1^2 \frac{dx}{e^x-1} &= \int_{u(1)}^{u(2)} \frac{dy}{y(y-1)} \\
&= \int_e^{e^2} \left(\frac{1}{y-1} - \frac{1}{y}\right) dy \\
&= [\ln|y-1|]_e^{e^2} - [\ln|y|]_e^{e^2} \\
&= \ln\frac{e^2-1}{e-1} - 1 = \ln(e+1) - \ln e = \ln\left(1+e^{-1}\right).
\end{aligned}$$

2. Bestimmen wir $\int_0^1 \sqrt{1-y^2}dy$ mit Hilfe von einer Substitution $y = u(x)$ we folgt. Nehmen wir $y = \sin x$ auf $[0, \pi/2]$ so dass $u(0) = 0$ und $u(\pi/2) = 1$. Dann haben wir

$$dy = \cos x dx$$

und

$$\begin{aligned}
\int_0^1 \sqrt{1-y^2}dy &= \int_{u(0)}^{u(\pi/2)} \sqrt{1-y^2}dy \\
&= \int_0^{\pi/2} \sqrt{1-\sin^2 x}\cos x\,dx \\
&= \int_0^{\pi/2} \cos^2 x\,dx \\
&= \int_0^{\pi/2} \frac{1+\cos 2x}{2}dx \\
&= \frac{\pi}{4} + \frac{1}{4}[\sin 2x]_0^{\pi/2} \\
&= \frac{\pi}{4}.
\end{aligned}$$

Natürlich die gleiche Antwort erhält man mit Hilfe vom unbestimmten Integral

$$\int \sqrt{1-y^2}dy = \frac{1}{2}\arcsin y + \frac{1}{2}y\sqrt{1-y^2} + C$$

und Newton-Leibniz-Formel.

4.13 Länge von Kurve

Wir betrachten hier die Abbildungen mit den Werten in $\mathbb{R}^n$. Jedes $x \in \mathbb{R}^n$ is eine Folge $x = (x_1, ..., x_n)$ von n reellen Zahlen, die die Komponenten von x sind. Für jedes $x \in \mathbb{R}^n$ definieren wir die Norm (oder den Betrag) von x mit

$$\|x\| = \sqrt{x_1^2 + x_2^2 + ... + x_n^2}.$$

Die Norm $\|x\|$ ist eine offensichtliche Verallgemeinerung des Betrages in $\mathbb{R}$ und $\mathbb{C} = \mathbb{R}^2$.

Sei I ein Intervall on $\mathbb{R}$. Jede Abbildung $\varphi : I \to \mathbb{R}^n$ hat n Komponenten $\varphi_1(t), ..., \varphi_n(t)$, so dass jedes φ_i eine reellwertige Funktion von $t \in I$ ist. Wir schreiben

$$\varphi(t) = (\varphi_1(t), ..., \varphi_n(t)).$$

Die Abbildung φ heißt stetig falls alle Komponenten $\varphi_j(t)$ stetig ist, und stetig differenzierbar falls alle Komponenten $\varphi_j(t)$ stetig differenzierbar sind. Im letzten Fall definieren wir die Ableitung φ' von φ mit

$$\varphi' = (\varphi_1', ..., \varphi_n'),$$

so dass φ' auch eine Abbildung von I nach $\mathbb{R}^n$ ist.

Definition. Das Bild $K = \varphi(I)$ von einer stetigen Abbildung $\varphi : I \to \mathbb{R}^n$ heißt *Kurve* und das Paar (I, φ) heißt *Parametrisierung* der Kurve K. Man nennt das Dreifache (K, I, φ) *parametrisierte Kurve.*

Betrachten wir I als ein Zeitintervall und $\varphi(t)$ als die Position von einem bewegenden Körper in $\mathbb{R}^n$ um Zeit t. Die Kurve $\varphi(I)$ ist die Spur des Körpers. Die Parametrisierung (I, φ) lässt sich als der Stundenplan des Körpers betrachten.

Definition. Sei $I = [\alpha, \beta]$ ein beschränktes abgeschlossenes Intervall mit $\alpha < \beta$. Für die parametrisierte Kurve (K, I, φ) mit stetig differenzierbarer Abbildung $\varphi : I \to \mathbb{R}$ definieren wir die *Länge* $L(K, I, \varphi)$ mit

$$L(K, I, \varphi) = \int_\alpha^\beta \|\varphi'(t)\| dt = \int_\alpha^\beta \sqrt{\varphi_1'(t)^2 + ... + \varphi_n'(t)^2} dt. \tag{4.48}$$

Die Ableitung $\varphi'(t)$ ist der Geschwindigkeitsvektor des Körpers und die Norm $\|\varphi'(t)\|$ ist die skalare Geschwindigkeit um Zeitpunkt t. Die Integration von $\|\varphi'(t)\|$ über I ergibt den zurückgelegten Weg des Körpers im Zeitintervall I, was die Länge der Spur $\varphi(I)$ ist.

Beispiel. Betrachten wir die Parametrisierung

$$\begin{aligned} \varphi &: [0,1] \to \mathbb{R}^2 \\ \varphi(t) &= (1-t)\,a + tb \end{aligned}$$

mit zwei Punkten $a, b \in \mathbb{R}^2$. Das Bild $K = \varphi(I)$ ist eine gerade Strecke zwischen a und b. Dann gilt

$$\varphi_1 = (1-t)\,a_1 + tb_1, \quad \varphi_2 = (1-t)\,a_2 + tb_2$$

und

$$\varphi_1' = b_1 - a_1, \quad \varphi_2' = b_2 - a_2$$

und

$$L = \int_0^1 \sqrt{(b_1 - a_1)^2 + (b_2 - a_2)^2} dt = \sqrt{(b_1 - a_1)^2 + (b_2 - a_2)^2}.$$

Beispiel. Betrachten wir die Parametrisierung

$$\begin{aligned} \varphi &: [0,2\pi] \to \mathbb{R}^2 \\ \varphi(t) &= (r\cos t, r\sin t) \end{aligned}$$

mit einem $r > 0$. Das Bild $K = \varphi(I)$ ist der Kreis von Radius r. Da

$$\varphi' = (-r\sin t, r\cos t)$$

und

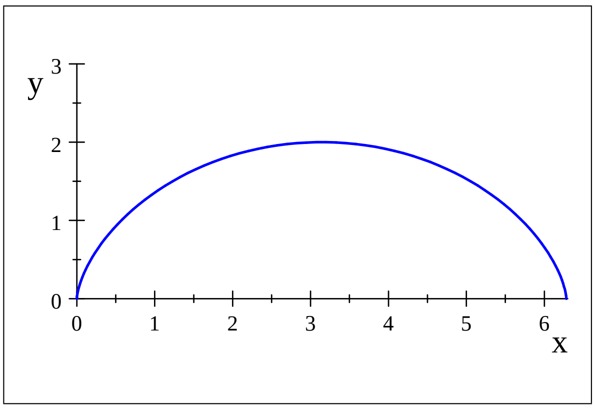
$$\|\varphi'\| = \sqrt{r^2 \sin^2 t + r^2 \cos^2 t} = r,$$

so erhalten wir

$$L = \int_0^{2\pi} r dt = 2\pi r.$$

Beispiel. Die *Zykloide* hat die Parametrisierung

$$\begin{aligned} \varphi &: [0,2\pi] \to \mathbb{R}^2 \\ \varphi(t) &= (t - \sin t, \;\; 1 - \cos t)\,. \end{aligned}$$



Zykloide

Dann gilt

$$\|\varphi'\| = \sqrt{(1-\cos t)^2 + \sin^2 t} = \sqrt{2 - 2\cos t} = 2\sin\frac{t}{2}.$$

Somit ist die Länge der Zykloide gleich

$$L = \int_0^{2\pi} 2\sin\frac{t}{2}dt = 4\int_0^{\pi} \sin s\ ds = 8.$$

Beispiel. Betrachten wir den Graph

$$\Gamma = \{(x, y) : x \in I,\ y = f(x)\}$$

einer Funktion $f : I \to \mathbb{R}$. Der Graph lässt sich betrachten als eine Kurve mit Parametrisierung

$$\begin{aligned} \varphi &: I \to \mathbb{R}^2 \\ \varphi(x) &= (x, f(x)) \end{aligned}$$

Sei $I = [\alpha, \beta]$. Ist f stetig differenzierbar, so erhalten wir

$$\|\varphi'\| = \sqrt{1 + f'(x)^2}$$

und somit

$$L = \int_\alpha^\beta \sqrt{1 + f'(x)^2}dx.$$

Zum Beispiel, für den Graph der Parabel $y = \frac{1}{2}x^2$ auf $[0, 1]$ erhalten wir

$$\begin{aligned} L &= \int_0^1 \sqrt{1+x^2}dx = \left[\frac{1}{2}\ln\left(x + \sqrt{x^2+1}\right) + \frac{1}{2}x\sqrt{x^2+1}\right]_0^1 \\ &= \frac{1}{2}\ln\left(1+\sqrt{2}\right) + \frac{\sqrt{2}}{2}. \end{aligned}$$

Satz 4.21 *Seien I, J zwei abgeschlossene beschränkte Intervalle und $\varphi : I \to \mathbb{R}^n$, $\psi : J \to \mathbb{R}^n$ zwei stetig differenzierbare Abbildungen. Existiert eine monotone stetig differenzierbare Funktion $f : I \to J$ mit $f(I) = J$ und $\varphi = \psi \circ f$, so bestimmen die zwei Parametrisierungen (I, φ), (J, ψ) die gleiche Kurve $K = \varphi(I) = \psi(J)$ und es gilt*

$$L(K, I, \varphi) = L(K, J, \psi).$$

Die Abbildungen φ, ψ und f werden an dem folgenden Diagramm gezeigt:

$$\begin{array}{ccccc} & & \mathbb{R}^n & & \\ & \varphi \nearrow & & \nwarrow \psi & \\ I & & \xrightarrow{f} & & J \end{array}$$

wobei φ die Verkettung von f und ψ ist. Insbesondere gilt für jedes $t \in I$ dass $\varphi(t) = \psi(s)$ mit $s = f(t)$; d.h. zwei bewegende Körper mit Stundenplänen (I, φ) bzw (J, ψ) sich immer am gleichen Punkt befinden, aber um verschiedene Zeitpunkte t und s.

Beweis. Wir haben

$$\varphi(I) = (\varphi \circ f)(I) = \psi(f(I)) = \psi(J),$$

was die erste Aussage ergibt.

Da f monoton ist, so gilt entweder $f'(t) \geq 0$ für alle $t \in I$ (f ist monoton steigend) oder $f'(t) \leq 0$ für alle $t \in I$ (f ist monoton fallend). Betrachten wir den Fall von monoton steigendem f (der Fall von fallendem f ist analog). Für jedes $k = 1, ..., n$ haben wir nach der Kettenregel

$$\varphi_k'(t) = \frac{d}{dt}(\psi_k(f(t))) = \psi_k'(f(t)) f'(t),$$

woraus folgt

$$\|\varphi'(t)\| = \|\psi'(f(t))\| |f'(t)| = \|\psi'(f(t))\| f'(t),$$

wo wir $f' \geq 0$ benutzen.

Seien $I = [\alpha, \beta]$ und $J = [a, b]$ mit $\alpha < \beta$ und $a < b$. Nach (4.48) haben wir

$$L(K, I, \varphi) = \int_\alpha^\beta \|\varphi'(t)\| \, dt = \int_\alpha^\beta \|\psi'(f(t))\| f'(t) \, dt.$$

Da $f(I) = J$ und f monoton steigend ist, so gelten $f(\alpha) = a$ und $f(\beta) = b$. Mit der Substitution $s = f(t)$ erhalten wir

$$L(K, I, \varphi) = \int_\alpha^\beta \|\psi'(f(t))\| \, df(t) = \int_{f(\alpha)}^{f(\beta)} \|\psi'(s)\| \, ds = \int_a^b \|\psi'(s)\| \, ds = L(K, J, \psi),$$

was zu beweisen war. ∎

Beispiel. Betrachten wir den Halbkreis mit zwei Parametrisierungen:

$$\begin{aligned} \varphi &: [0, \pi] \to \mathbb{R}^2 \\ \varphi(t) &= (\cos t, \sin t) \end{aligned}$$

und

$$\begin{aligned} \psi &: [-1, 1] \to \mathbb{R}^2 \\ \psi(s) &= (s, \sqrt{1 - s^2}). \end{aligned}$$

Für die Funktion

$$\begin{aligned} f &: [0, \pi] \to [-1, 1] \\ f(t) &= \cos t \end{aligned}$$

erhalten wir

$$\psi(f(t)) = \left(\cos t, \sqrt{1 - \cos^2 t}\right) = \varphi(t).$$

Somit sind die Längen von den beiden Parametrisierungen gleich. Da

$$L\left(K,\varphi,I\right)=\pi$$

somit erhalten wir auch

$$L\left(K,\psi,J\right)=\pi.$$

Da

$$\|\psi\|=\sqrt{1+\left|\left(\sqrt{1-s^2}\right)'\right|^2}=\sqrt{1+\frac{s^2}{1-s^2}}=\frac{1}{\sqrt{1-s^2}},$$

so erhalten wir

$$\int_{-1}^{1}\frac{ds}{\sqrt{1-s^2}}=\pi.$$

Natürlich lässt dieses Integral sich auch direkt berechnen:

$$\int_{-1}^{1}\frac{ds}{\sqrt{1-s^2}}=\left[\int\frac{ds}{\sqrt{1-s^2}}\right]_{-1}^{1}=\left[\arcsin s\right]_{-1}^{1}=\pi.$$

Chapter 5

Konvergenz von Integralen und Reihen

5.1 Uneigentliches Integral

5.1.1 Definition und Eigenschaften vom uneigentlichen Integral

Das Riemann-Integral $\int_a^b f(x)\,dx$ ist immer für eine Funktion auf einem abgeschlossenen beschränkten Intervall $[a,b]$ definiert ist. In diesem Abschnitt wird diese Definition zu anderen Typen von Integralen erweitert.

Definition. Sei f eine Funktion auf einem Intervall I. Die Funktion f heißt *lokal integrierbar auf* I falls f auf jedem abgeschlossen beschränkten Intervall $J \subset I$ Riemann-integrierbar ist.

Es folgt aus dem Satz 4.7, dass alle stetige Funktionen und alle monotone Funktionen lokal integrierbar sind.

Definition. Sei f eine lokal integrierbare Funktion auf einem Intervall $[a,b)$ mit $-\infty < a < b \le +\infty$. Dann definieren wir das *uneigentliche* Riemann-Integral von f mit

$$\int_a^b f(x)\,dx = \lim_{c\to b-} \int_a^c f(x)\,dx, \tag{5.1}$$

vorausgesetzt, dass der Grenzwert existiert als Element von $\overline{\mathbb{R}}$. Die Notation $c \to b-$ bedeutet, dass $c < b$ und $c \to b$; insbesondere ist das Riemann-Integral $\int_a^c f(x)\,dx$ wohldefiniert.

Ist der Grenzwert in (5.1) endlich, so sagt man, dass das Integral $\int_a^b f(x)\,dx$ an der Grenze b konvergiert.

Ist der Grenzwert unendlich, so sagt man, dass das Integral an der Grenze b bestimmt divergiert.

Existiert der Grenzwert nicht, so sagt man, dass das Integral an b unbestimmt divergiert. Im letzten Fall ist der Wert des Ausdrucks $\int_a^b f(x)\,dx$ nicht definiert.

Die Grenze b für das uneigentliche Integral (5.1) heißt *kritisch.*

Definition. Im Fall wenn f auf dem Intervall $(a, b]$ mit $-\infty \leq a < b < +\infty$ lokal integrierbar ist, wird das uneigentliche Integral analog definiert:

$$\int_a^b f(x)\,dx = \lim_{c \to a+} \int_c^b f(x)\,dx, \tag{5.2}$$

vorausgesetzt, dass der Grenzwert existiert. Die Notation $c \to b+$ bedeutet, dass $c > b$ und $c \to b$. Die Grenze a für das uneigentliche Integral (5.2) heißt kritisch.

Ist f auf einem abgeschlossenen beschränkten Intervall $[a, b]$ definiert und Riemann-integrierbar, so werden jetzt drei Begriffe von Integral für f definiert: eigentliches (normales) Riemann-Integral, das uneigentliche Integral mit kritischer Grenze a und das uneigentliche Integral mit kritischer Grenze b.

Behauptung *Ist f auf $[a, b]$ Riemann-integrierbar, so stimmen drei Werte von $\int_a^b f(x)\,dx$ überein.*

Beweis. Markieren wir die kritische Grenze mit einem Punkt und zeigen, dass

$$\int_a^{\bullet b} f(x)\,dx = \int_a^b f(x)\,dx = \int_{\bullet a}^b f(x)\,dx.$$

Wir benutzen die Stetigkeit der Funktion $F(x) = \int_a^x f(t)\,dt$. Für die linke Identität haben wir

$$\int_a^{\bullet b} f(x)\,dx = \lim_{c \to b-} \int_a^c f(x)\,dx = \lim_{c \to b-} F(c) = F(b) = \int_a^b f(x)\,dx,$$

und die rechte Identität wird analog bewiesen. ■

Beispiel. Betrachten wir die Funktion $f(x) = x^p$ auf $[1, +\infty)$ und das uneigentliche Integral $\int_1^\infty x^p dx$. Im Fall $p \neq -1$ haben wir für jedes $c \in (1, +\infty)$

$$\int_1^c x^p dx = \left[\int x^p dx\right]_1^c = \left[\frac{x^{p+1}}{p+1}\right]_1^c = \frac{c^{p+1}}{p+1} - \frac{1}{p+1}.$$

Im Fall $p > -1$ so divergiert die rechte Seite gegen $+\infty$ für $c \to +\infty$. Somit gilt in diesem Fall

$$\int_1^{+\infty} x^p dx = +\infty \text{ falls } p > -1,$$

d.h. das Integral bestimmt divergent ist. Im Fall $p < -1$ gilt $c^{p+1} \to 0$ für $c \to +\infty$, woraus folgt

$$\int_1^{+\infty} x^p dx = -\frac{1}{p+1} \text{ falls } p < -1,$$

d.h. das Integral konvergent ist.

Im Fall $p = -1$ haben wir

$$\int_1^c \frac{dx}{x} = [\ln x]_1^c = \ln c \to +\infty \text{ für } c \to +\infty,$$

woraus folgt

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

Somit ist das Integral $\int_1^{+\infty} x^p dx$ konvergent für $p < -1$ und bestimmt divergent für $p \geq -1$.

Beispiel. Betrachten wir das uneigentliche Integral $\int_0^\infty \sin x dx$. Für jedes $1 < c < \infty$ gilt

$$\int_0^c \sin x dx = -[\cos x]_0^x = 1 - \cos c.$$

Aber der Grenzwert $\lim_{c\to\infty} \cos c$ existiert nicht. Somit beschließen wir, dass $\int_0^\infty \sin x dx$ unbestimmt divergiert und keinen Wert hat.

Die Eigenschaften von eigentlichem Integral werden für uneigentliches Integral einfach verallgemeinert. Dafür verallgemeinern wir die Notation $[F]_a^b$ zum Fall wenn F auf $[a,b)$ definiert ist wie folgt:

$$[F]_a^b = F(b-) - F(a),$$

wobei

$$F(b-) := \lim_{x\to b-} F(x),$$

vorausgesetzt, dass der Limes existiert, endlich oder unendlich.

Satz 5.1 (Newton-Leibniz-Formel für uneigentliches Integral) *Seien $f(x)$ eine stetige Funktion auf $[a,b)$ mit $-\infty < a < b \leq +\infty$ und F ihre Stammfunktion auf $[a,b)$. Dann gilt*

$$\int_a^b f(x)\, dx = [F]_a^b$$

vorausgesetzt, dass mindestens eine von zwei Seiten wohldefiniert ist.

Beweis. Nach Definition von dem uneigentlichen Integral und Newton-Leibniz-Formel für eigentliches Integral gilt

$$\begin{aligned}\int_a^b f(x)\, dx &= \lim_{c\to b-} \int_a^c f(x)\, dx = \lim_{c\to b-} (F(c) - F(b)) = \lim_{c\to b-} F(c) - F(a) \\ &= F(b-) - F(a) = [F]_a^b.\end{aligned}$$

Insbesondere sehen wir, dass der Grenzwert in der linken Seite genau dann existiert, wenn der Grenzwert in der rechten Seite existiert. ■

Partielle Integration. *Seien u und v stetig differenzierbare Funktion auf $[a,b)$. Dann gilt*

$$\int_a^b u dv = [uv]_a^b - \int_a^b v du \tag{5.3}$$

vorausgesetzt, dass die rechte Seite wohldefiniert ist, d.h. der Wert $(uv)(b-)$ und das uneigentliche Integral $\int_a^b v du$ existieren und deren Differenz wohldefiniert ist.

Beweis ist trivial: da die Funktionen u, v auf $[a,c]$ stetig differenzierbar für jedes $c \in (a,b)$ sind, so gilt es nach dem Satz 4.3

$$\int_a^c u dv = [uv]_a^c - \int_a^c v du.$$

Für $c \to b-$ erhalten wir (5.3).

Substitutionsregel. *Sei u eine stetig differenzierbare Funktion auf einem Intervall $I = [a, b)$ so dass der Wert $u(b-)$ existiert. Sei f eine stetige Funktion auf dem Intervall $u(I)$. Dann gilt*

$$\int_a^b f(u(x))\, du(x) = \int_{u(a)}^{u(b-)} f(y)\, dy, \tag{5.4}$$

vorausgesetzt, dass mindestens eines von zwei uneigentlichen Integralen existiert.

In der Tat gilt für jedes $c \in (a, b)$

$$\int_a^c f(u(x))\, du(x) = \int_{u(a)}^{u(c)} f(y)\, dy.$$

Für $c \to b-$ erhalten wir $u(c) \to u(b-)$ und somit (5.4).

Analog formuliert und beweist man die weiteren Eigenschaften von Integral, wie Linearität, Monotonie, Additivität.

Im Fall wenn die kritische Grenze a ist, d.h. man integriert über $(a, b]$, gelten alle obigen Eigenschaften auch. Zum Beispiel, die Newton-Leibniz-Formel sieht so aus: ist f auf $(a, b]$ stetig, dann für die Stammfunktion F von f gilt

$$\int_a^b f(x)\, dx = [F]_a^b =: F(b) - F(a+),$$

wobei

$$F(a+) = \lim_{x \to a+} F(x).$$

Beispiel. 1. Bestimmen wir $\int_0^1 \frac{dx}{\sqrt{x}}$. Hier ist 0 die kritische Grenze. Nach Newton-Leibniz-Formel erhalten wir

$$\int_0^1 \frac{dx}{\sqrt{x}} = \left[\int x^{-1/2} dx\right]_0^1 = \left[2x^{1/2}\right]_0^1 = 2.$$

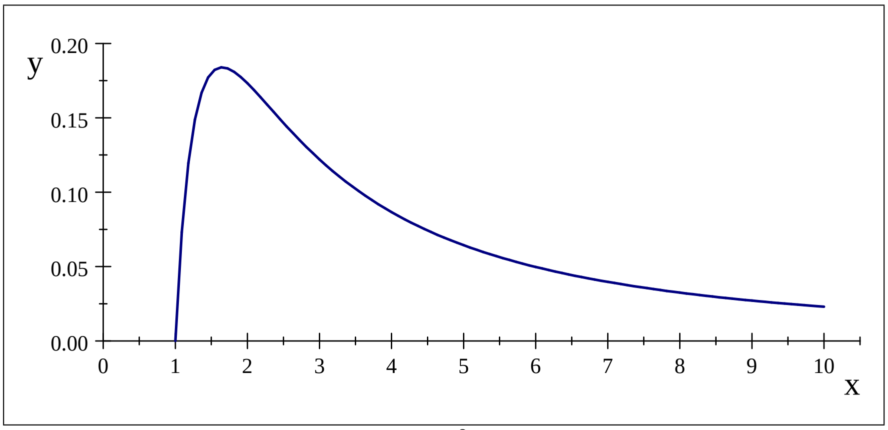
Analog erhalten wir

$$\int_0^1 \frac{dx}{x} = \left[\int \frac{dx}{x}\right]_0^1 = [\ln x]_0^1 = \ln 1 - \ln(0+) = +\infty.$$

2. Bestimmen $\int_1^{+\infty} x^{-2} \ln x dx$, wo die kritische Grenze $+\infty$ ist. Wir erhalten mit Hilfe von der partiellen Integration

$$\begin{aligned}
\int_1^{+\infty} x^{-2} \ln x dx &= -\int_1^{+\infty} \ln x d\frac{1}{x} = -\left[\frac{1}{x} \ln x\right]_1^{+\infty} + \int_1^{+\infty} \frac{1}{x} d \ln x \\
&= \int_1^{+\infty} \frac{1}{x^2} dx = \left[\int x^{-2} dx\right]_1^{+\infty} = -\left[x^{-1}\right]_1^{+\infty} = -(0 - 1) = 1,
\end{aligned}$$

wo wir benutzt haben, das $\frac{\ln x}{x} \to 0$ für $x \to +\infty$.



$y = x^{-2} \ln x$

3. Bestimmen $\int_e^{+\infty} \frac{1}{x \ln^2 x} dx$, wo die kritische Grenze $+\infty$ ist. Mit Hilfe von Substitution $y = u(x) = \ln x$ erhalten wir

$$\begin{aligned} \int_e^{+\infty} \frac{1}{x \ln^2 x} dx &= \int_e^{+\infty} \frac{d \ln x}{\ln^2 x} \\ &= \int_{u(e)}^{u(+\infty)} \frac{dy}{y^2} = \left[\int y^{-2} dy \right]_1^{+\infty} = -\left[y^{-1} \right]_1^{+\infty} = 1. \end{aligned}$$

Betrachten wir jetzt uneigentliches Integral mit beiden kritischen Grenzen.

Definition. Sei f eine lokal integrierbare Funktion auf einem Integral (a, b) mit $-\infty \leq a < b \leq +\infty$. Definieren wir das uneigentliche Integral $\int_a^b f(x)\, dx$ mit beiden kritischen Grenzen a, b mit

$$\int_a^b f(x)\, dx = \int_a^c f(x)\, dx + \int_c^b f(x)\, dx, \tag{5.5}$$

wobei $c \in (a, b)$, vorausgesetzt, dass die beiden Integrale in der rechten Seite existieren als uneigentliche Integrale mit einer kritischen Grenze und deren Summe auch wohldefiniert ist.

Behauptung. *Der Wert von $\int_a^b f(x)\, dx$ in (5.5) ist unabhängig von der Wahl von c.*

Beweis. Sei c' noch ein Punkt in (a, b). Wir haben dann

$$\begin{aligned} \int_a^{c'} f dx + \int_{c'}^b f dx &= \left(\int_a^c f dx + \int_c^{c'} f dx \right) + \left(\int_{c'}^c f dx + \int_c^b f dx \right) \\ &= \int_a^c f dx + \int_c^b f dx + \left(\int_c^{c'} f dx + \int_{c'}^c f dx \right) \\ &= \int_a^c f dx + \int_c^b f dx. \end{aligned}$$

■

Alle Eigenschaften von uneigentlichen Integralen gelten auch für den Fall von zwei kritischen Grenzen. In diesem Fall wird die Notation $[F]_a^b$ noch weiter verallgemeinert wie folgt:

$$[F]_a^b = F(b-) - F(a+),$$

vorausgesetzt, dass die beiden Werte $F(b-)$ und $F(a+)$ existiert und deren Differenz auch wohldefiniert ist. Zum Beispiel, beweisen wir die Newton-Leibniz-Formel in diesem Fall.

Satz 5.2 (Newton-Leibniz-Formel für uneigentliches Integral mit zwei kritischen Grenzen) *Seien $f(x)$ eine stetige Funktion auf (a,b) mit $-\infty \le a < b \le +\infty$ und F eine Stammfunktion von f auch (a,b). Dann gilt*

$$\int_a^b f(x)\,dx = [F]_a^b,$$

vorausgesetzt, dass mindestens eine von zwei Seiten wohldefiniert ist.

Beweis. In der Tat folgt es aus der Definition (5.5) und dem Satz 5.1, dass

$$\begin{aligned}\int_a^b f(x)\,dx &= \int_a^c f(x)\,dx + \int_c^b f(x)\,dx \\ &= [F]_a^c + [F]_c^b \\ &= F(c) - F(a+) + F(b-) - F(c) \\ &= [F]_a^b.\end{aligned}$$

■

Beispiel. 1. Betrachten wir $\int_{-\infty}^{+\infty} x dx$. Die Newton-Leibniz-Formel ergibt

$$\int_{-\infty}^{+\infty} x dx = \left[\frac{x^2}{2}\right]_{-\infty}^{+\infty} = +\infty - (+\infty),$$

was unbestimmter Ausdruck ist. Somit ist das Integral nicht definiert.

2. Bestimmen wir $\int_{-1}^{1} \frac{dx}{\sqrt{1-x^2}}$. Die Funktion $\frac{1}{\sqrt{1-x^2}}$ ist stetig auf $(-1,1)$ aber nicht an ± 1 definiert, so dass die beiden Grenzen ± 1 kritisch sind. Nach de Newton-Leibniz-Formel erhalten wir

$$\int_{-1}^{1} \frac{dx}{\sqrt{1-x^2}} = \left[\int \frac{dx}{\sqrt{1-x^2}}\right]_{-1}^{1} = [\arcsin x]_{-1}^{1} = \frac{\pi}{2} - \left(-\frac{\pi}{2}\right) = \pi.$$

3. Betrachten wir $\int_{-\pi/2}^{\pi/2} \tan x$. Die Funktion $\tan x$ ist stetig auf $(-\pi/2, \pi/2)$, aber an $\pm\pi/2$ ist nicht definiert (unendlich). Wir haben

$$\int_{-\pi/2}^{\pi/2} \tan x dx = \int_{-\pi/2}^{\pi/2} \frac{\sin x}{\cos x} dx = -\int_{-\pi/2}^{\pi/2} \frac{d\cos x}{\cos x} = -\left[\ln|\cos x|\right]_{-\pi/2}^{\pi/2} = -\left(-\infty - (-\infty)\right).$$

Wir haben benutzt, dass $\cos \pi/2 = 0$ and $\ln(0+) = -\infty$. Da die Differenz $-\infty - (-\infty)$ nicht wohldefiniert ist, so existiert das Integral $\int_{-\pi/2}^{\pi/2} \tan x$ nicht.

5.1.2 Konvergenzkriterien von uneigentlichen Integralen

Wir fangen mit der folgenden Beobachtung an.

Satz 5.3 *Sei* $-\infty \leq a < b \leq +\infty$. *Ist* $f : (a,b) \to \mathbb{R}$ *nichtnegativ und lokal integrierbar, so existiert das uneigentliche Integral* $\int_a^b f(x)\,dx$ *mit dem Wert in* $[0,+\infty]$ *an. Insbesondere ist* $\int_a^b f(x)\,dx$ *genau dann konvergent, wenn*

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

Beweis. Für ein $c \in (a,b)$ betrachten wir die Funktion

$$F(x) = \int_c^x f(t)\,dt.$$

Die Funktion $F(x)$ ist monoton steigend, da für $y > x$ gilt

$$F(y) - F(x) = \int_c^y f(t)\,dt - \int_c^x f(t)\,dt = \int_x^y f(t)\,dt \geq 0.$$

Somit existieren die Grenzwerte $F(b-)$ und $F(a+)$. Nach Definition gilt

$$\begin{aligned}\int_a^b f(x)\,dx &= \int_a^c f(x)\,dx + \int_c^b f(x)\,dx \\ &= \lim_{x\to b-}\int_c^x f(t)\,dt + \lim_{x\to a+}\int_x^c f(t)\,dt \\ &= F(b-) - F(a+).\end{aligned}$$

De Differenz $F(b-) - F(a+)$ ist wohldefiniert ist, da die Werte $F(b-)$ und $F(a+)$ nicht gleichzeitig $+\infty$ oder $-\infty$ sein können, weil

$$F(a-) \leq F(c) = 0 \leq F(b+).$$

Daraus auch folgt, dass $F(b+) - F(a-) \in [0,+\infty]$ und somit auch $\int_a^b f(t)\,dt \in [0,+\infty]$. Die zweite Aussage ist offensichtlich. ■

Erinnern wir uns für Vergleichen eine Eigenschaft von den Reihen: für jede nichtnegative Reihe $\sum_{n=1}^{\infty} a_n$ ist ihre Summe immer wohldefiniert als Element von $[0,+\infty]$, und $\sum_{k=1}^{\infty} a_k$ ist genau dann konvergent wenn

$$\sum_{k=1}^{\infty} a_k < +\infty.$$

Der folgende Satz etabliert eine direkte Beziehung zwischen Konvergenz von Reihen und Integralen.

Satz 5.4 (Integralkriterium für Konvergenz von Reihen) *Sei* $f(x)$ *eine nichtnegative monoton fallende Funktion auf* $[1,+\infty)$. *Dann gilt die Äquivalenz*

$$\int_1^{+\infty} f(x)\,dx < \infty \iff \sum_{k=1}^{\infty} f(k) < \infty.$$

Beweis. Da f nichtnegativ und lokal integrierbar ist, so nehmen $\int_1^{+\infty} f(x)\,dx$ und $\sum_{k=1}^{\infty} f(k)$ die Werte in $[0,+\infty]$ an. Wir beweisen, dass die beiden Werte gleichzeitig entweder endlich oder unendlich sind. Fixieren wir ein $n \geq 2$ und betrachten die Zerlegung $Z = \{1, 2, ..., n\}$ von Intervall $[1, n]$. Da $f(x)$ monoton fallend ist, so gilt auf jedem Intervall $[k-1, k] \subset [1, n]$

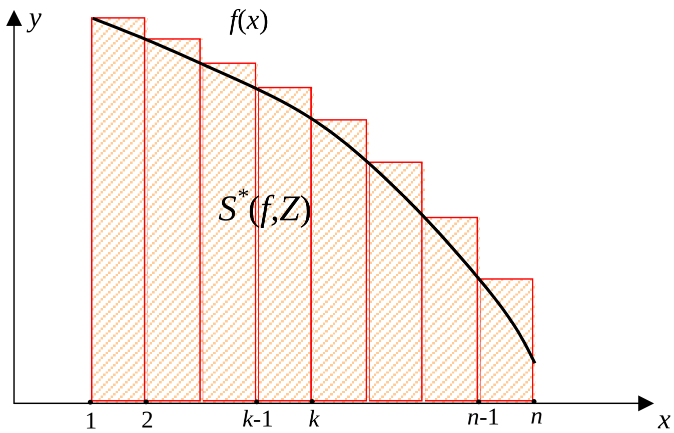
$$\sup_{[k-1,k]} f = f(k-1) \text{ und } \inf_{[k-1,k]} f = f(k).$$

Somit sind die entsprechenden Darboux-Summen gleich

$$S^*(f, Z) = \sum_{k=2}^{n} (\sup_{[k-1,k]} f)(k-(k-1)) = \sum_{k=2}^{n} f(k-1) = f(1) + ... + f(n-1),$$

und

$$S_*(f, Z) = \sum_{k=2}^{n} (\inf_{[k-1,k]} f)(k-(k-1)) = \sum_{k=2}^{n} f(k) = f(2) + ... + f(n)$$



Da immer

$$S_*(f, Z) \leq \int_1^n f(x)\,dx \leq S^*(f, Z),$$

so erhalten wir

$$f(2) + ... + f(n) \leq \int_1^n f(x)\,dx \leq f(1) + ... + f(n-1).$$

Für $n \to \infty$ ergibt diese Ungleichung, dass

$$\sum_{k=1}^{\infty} f(k) - f(1) \leq \int_1^{\infty} f(x)\,dx \leq \sum_{k=1}^{\infty} f(k). \tag{5.6}$$

Somit ist $\int_1^{\infty} f(x)\,dx$ genau dann endlich wenn $\sum_{k=1}^{\infty} f(k) < \infty$. ■

Beispiel. Zeigen wir, dass die Reihe $\sum_{n=1}^{\infty} \frac{1}{n^p}$ genau dann konvergiert, wenn $p > 1$. Im Fall $p \leq 0$ konvergiert die Folge $\left\{\frac{1}{n^p}\right\}$ gegen 0 nicht, und somit ist die Reihe divergent. Im Fall $p > 0$ ist die Funktion $f(x) = \frac{1}{x^p}$ monoton fallend auf $[1, +\infty)$. Nach dem Satz 5.4 konvergiert $\sum_{n=1}^{\infty} \frac{1}{n^p}$ genau dann, wenn

$$\int_1^{+\infty} \frac{dx}{x^p} < \infty, \tag{5.7}$$

und das letzte ist der Fall genau dann wenn $p > 1$.

Definition. Sei f eine lokal integrierbare Funktion auf (a, b). Das Integral $\int_a^b f(x)\, dx$ heißt *absolut konvergent*, falls $\int_a^b |f(x)|\, dx$ konvergent ist, d.h.

$$\int_a^b |f(x)|\, dx < +\infty.$$

Bemerken wir, dass $|f|$ auch lokal integrierbar ist (siehe Aufgaben). Da $|f| \geq 0$, so existiert das Integral $\int_a^b |f(x)|\, dx$ nach dem Satz 5.3.

Satz 5.5 *Ist das Intervall $\int_a^b f(x)\, dx$ absolut konvergent, so ist es auch konvergent und es gilt*

$$\left| \int_a^b f(x)\, dx \right| \leq \int_a^b |f(x)|\, dx.$$

Beweis. Wählen wir ein $c \in (a, b)$ und setzen

$$F(x) = \int_c^x f(t)\, dt \quad \text{und} \quad G(x) = \int_c^x |f(t)|\, dt.$$

Die beiden Grenzwerte $G(a+)$ und $G(b-)$ existieren und sind endlich, da

$$\int_a^b |f(x)|\, dx = \int_a^c |f(x)|\, dx + \int_c^b |f(x)|\, dx = G(b-) - G(a+)$$

(wie im Beweis von Satz 5.3). Beweisen wir dass auch der Grenzwert $F(b-)$ existiert und endlich ist. Es reicht zu zeigen, dass für jede Folge $x_n \to b-$ die Folge $\{F(x_n)\}$ konvergiert, d.h. $\{F(x_n)\}$ eine Cauchy-Folge ist. In der Tat für $x_n > x_m$ erhalten wir

$$\begin{aligned} |F(x_n) - F(x_m)| &= \left| \int_c^{x_n} f(t)\, dt - \int_c^{x_m} f(t)\, dt \right| = \left| \int_{x_m}^{x_n} f(t)\, dt \right| \\ &\leq \int_{x_m}^{x_n} |f(t)|\, dt = G(x_n) - G(x_m). \end{aligned}$$

Nach der Voraussetzung gilt $G(x_n) - G(x_m) \to 0$, woraus auch $|F(x_n) - F(x_m)| \to 0$ folgt. Somit existiert der Grenzwert $F(b-)$ und er ist endlich. Gleiches gilt für $F(a+)$, woraus folgt, dass das Integral

$$\int_a^b f(x)\, dx = \int_a^c f(x)\, dx + \int_c^b f(x)\, dx = -F(a+) + F(b-)$$

konvergiert. ■

Definition. Seien $f(x)$ und $g(x)$ zwei Funktion auf (a,b), die nicht verschwinden. Gilt

$$\lim_{x\to b}\frac{f(x)}{g(x)} = 1,$$

so sagen wir, dass $f(x)$ äquivalent zu $g(x)$ für $x\to b$ ist und schreiben

$$f(x)\sim g(x) \text{ für } x\to b.$$

Behauptung. *Die Relation $f\sim g$ für $x\to b$ ist eine Äquivalenzrelation.*

Beweis. Offensichtlich $f\sim f$ und $f\sim g$ ergibt $g\sim f$. Gelten $f\sim g$ und $g\sim h$ so gilt auch $f\sim h$ da

$$\frac{f}{h} = \frac{f}{g}\frac{g}{h}.$$

■

Behauptung *Gelten $f_1\sim g_1$ und $f_2\sim g_2$ so gelten auch $f_1f_2\sim g_1g_2$ und $\frac{f_1}{f_2}\sim\frac{g_1}{g_2}$.*

Beispiel. 1. $\sin x\sim x$ für $x\to 0$ da $\frac{\sin x}{x}\to 1$ für $x\to 0$.
2. Es gilt

$$x^2+x\sim x^2 \quad \text{für } x\to+\infty$$

da

$$\frac{x^2+x}{x^2} = 1+\frac{1}{x}\to 1 \text{ für } x\to+\infty.$$

Andererseits,

$$x^2+x\sim x \quad \text{für } x\to 0$$

da

$$\frac{x^2+x}{x} = x+1\to 1 \text{ für } x\to 0.$$

Erinnern wir uns an Landau-Symbol o:

$$f(x) = o(g(x)) \text{ für } x\to b \text{ falls } \lim_{x\to b}\frac{f(x)}{g(x)} = 0.$$

Es gilt $f\sim g$ genau dann when $f(x) = g(x)+o(g(x))$ da

$$\frac{f}{g} = 1+\frac{f-g}{g}$$

und $\frac{f}{g}\to 1$ genau dann, wenn $f-g = o(g)$.
Definieren wie das andere Landau-Symbol O.

Definition. Seien f,g zwei Funktionen auf (a,b) und $g(x)>0$. Wir schreiben

$$f(x) = O(g(x)) \text{ für } x\to b$$

und sagen "$f(x)$ ist groß O von $g(x)$ für $x \to b$" falls $|f(x)| \leq Cg(x)$ for alle $x \in (c,b)$ gilt, mit einer Konstante C und einem $c \in (a,b)$. Äquivalente Definition:

$$f(x) = O(g(x)) \text{ für } x \to b \text{ falls } \limsup_{x \to b} \frac{|f(x)|}{g(x)} < +\infty$$

Die Äquivalenz $f \sim g$ ergibt offensichtlich $f = O(g)$.

Satz 5.6 (Vergleichskriterium) *Seien $f(x)$ und $g(x)$ lokal integrierbare Funktionen auf $[a,b)$ und $g(x) > 0$.*

(a) Gelten $f(x) = O(g(x))$ für $x \to b$ und $\int_a^b g(x)\,dx < \infty$, so ist das Integral $\int_a^b f(x)\,dx$ absolut konvergent.

(b) Gilt $f(x) \sim g(x)$ für $x \to b$ so sind die Integrale $\int_a^b f(x)\,dx$ und $\int_a^b g(x)\,dx$ gleichzeitig konvergent bzw divergent.

Beweis. (a) Nach Voraussetzung $f(x) = O(g(x))$ für $x \to b$ existieren $C > 0$ and $c \in (a,b)$ mit

$$|f(x)| \leq Cg(x) \quad \text{für alle } c \leq x < b. \tag{5.8}$$

Schreiben wir

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

Die Funktion f ist auf $[a,c]$ integrierbar, so dass das Integral $\int_a^c |f(x)|\,dx$ eigentlich ist. Nach (5.8) lässt das zweite Integral sich wie folgt abschätzen:

$$\int_c^b |f(x)|\,dx \leq C\int_c^b g(x)\,dx < +\infty.$$

Daraus folgt $\int_a^b |f(x)|\,dx < \infty$ d.h. $\int_a^b f(x)\,dx$ konvergiert absolut.

(b) Die Relation $f(x) \sim g(x)$ ergibt $f(x) = O(g(x))$ und somit nach (a)

$$\int_a^b g(x)\,dx < +\infty \implies \int_a^b f(x)\,dx < +\infty.$$

Die Implikation in die umgekehrte Richtung folgt von $g \sim f$. ■

Beispiel. 1. Untersuchen wir die Konvergenz von $\int_1^{+\infty} \frac{\sqrt{x}dx}{\sqrt{1+x^4}}$. Wir haben

$$f(x) = \frac{\sqrt{x}}{\sqrt{1+x^4}} = \frac{\sqrt{x}}{x^2\sqrt{x^{-4}+1}} \sim \frac{\sqrt{x}}{x^2} = x^{-3/2} \text{ für } x \to +\infty,$$

d.h. $f(x) \sim x^{-3/2}$ für $x \to +\infty$. Da

$$\int_1^{+\infty} x^{-3/2}dx < +\infty,$$

so beschließen wir nach dem Satz 5.6, dass

$$\int_1^{+\infty} \frac{\sqrt{x}dx}{\sqrt{1+x^4}} < +\infty.$$

2. Untersuchen wir die Konvergenz von $\int_1^{+\infty} \frac{\sin x}{x^2} dx$. Da

$$\frac{\sin x}{x^2} = O\left(x^{-2}\right) \text{ für } x \to +\infty,$$

und $\int_1^{\infty} x^{-2} dx < +\infty$, so ist das Integral $\int_1^{\infty} \frac{\sin x}{x^2} dx$ absolut konvergent.

3. Untersuchen wir die Konvergenz von

$$\int_0^a \frac{dx}{\sqrt{\cos x - \cos a}} \tag{5.9}$$

wobei $0 < a < \pi/2$. Die kritische Grenze ist a. Nach Differenzierbarkeit von $\cos x$ folgt es, dass für $x \to a$

$$\cos x - \cos a = -(\sin a)(x-a) + o(x-a) \sim (\sin a)(a-x),$$

woraus folgt

$$\frac{1}{\sqrt{\cos x - \cos a}} \sim \frac{1}{\sqrt{\sin a}\sqrt{a-x}} \text{ für } x \to a$$

Da

$$\int_0^a \frac{dx}{\sqrt{a-x}} = \int_0^a \frac{dy}{\sqrt{y}} = [2\sqrt{y}]_0^a = 2\sqrt{a} < \infty,$$

so beschließen wir, dass das Integral (5.9) absolut konvergent ist.

5.1.3 Bedingte Konvergenz

Definition. Sei f eine lokal integrierbare Funktion auf einem Intervall (a, b). Das uneigentliche Integral $\int_a^b f(x)\, dx$ heißt *bedingt konvergent* falls es konvergent aber nicht absolut konvergent ist.

Der nächste Satz liefert zwei Kriterien für Konvergenz ohne absolute Konvergenz zu benutzen.

Satz 5.7 *Sei $f(x)$ eine stetige Funktion auf $[a, +\infty)$ und $g(x)$ eine stetig differenzierbare monotone Funktion $[a, +\infty)$. Das Integral*

$$\int_a^{\infty} f(x)\, g(x)\, dx$$

konvergiert unter jeder von zwei folgenden Bedingungen:

(a) (Abel-Kriterium) *Integral $\int_a^{\infty} f(x)\, dx$ ist konvergent und $g(x)$ ist beschränkt.*

(b) (Dirichlet-Kriterium) *Die Stammfunktion $F(x) = \int_a^x f(t)\, dt$ ist auf $[a, \infty)$ beschränkt und $g(x) \to 0$ für $x \to +\infty$.*

Beweis. Partielle Integration ergibt

$$\int_a^{\infty} f(x)\, g(x)\, dx = \int_a^{\infty} g(x)\, dF(x) = [Fg]_a^{\infty} - \int_a^{\infty} F(x)\, g'(x)\, dx. \tag{5.10}$$

Wir zeigen, dass die beiden Glieder am rechts endlich sind.

Beweisen wir zunächst, dass das $\int_a^\infty F(x)\, g'(x)\, dx$ absolut konvergent ist. In den beiden Fällen (a) und (b) ist die Funktion F beschränkt. Im Fall (b) ist es Voraussetzung, während im Fall (a) ergibt die Konvergenz von $\int_a^\infty f dx$, dass $\lim_{x\to+\infty} F(x)$ existiert und ist endlich, woraus folgt, dass F beschränkt ist. Sei C eine Konstante mit $|F(x)| \leq C$ für alle $x \in [a, \infty)$. Dann gilt

$$\int_a^\infty |F(x)\, g'(x)|\, dx \leq C \int_a^\infty |g'(x)|\, dx.$$

In den beiden Fällen existiert $\lim_{x\to\infty} g(x)$ und ist endlich. Im Fall (b) ist es Voraussetzung, während im Fall (a) folgt es aus der Monotonie und Beschränktheit von g. Die Monotonie ergibt auch, dass entweder immer $g'(x) \geq 0$ oder immer $g'(x) \leq 0$. Angenommen $g'(x) \geq 0$, so erhalten wir

$$\int_a^\infty |g'(x)|\, dx = \int_a^\infty g'(x)\, dx = [g]_a^{+\infty} = \lim_{x\to+\infty} g(x) - g(a) < +\infty,$$

was die absolute Konvergenz von dem Integral in der rechten Seite von (5.10) beweist .

Um zu beweisen, dass der Ausdruck $[Fg]_a^{+\infty}$ endlich ist, betrachten wir die zwei Fälle separat.

(a) In diesem Fall existiert endlicher $\lim_{x\to+\infty} F(x)$. Da $\lim_{x\to+\infty} g(x)$ auch existiert und endlich ist, so ist $[Fg]_a^{+\infty}$ wohldefiniert und endlich.

(b) Da $F(x)$ beschränkt ist und $g(x) \to 0$ für $x \to +\infty$, so gilt $F(x)\, g(x) \to 0$ für $x \to +\infty$ so dass $[Fg]_a^{+\infty}$ wieder wohldefiniert ist. ■

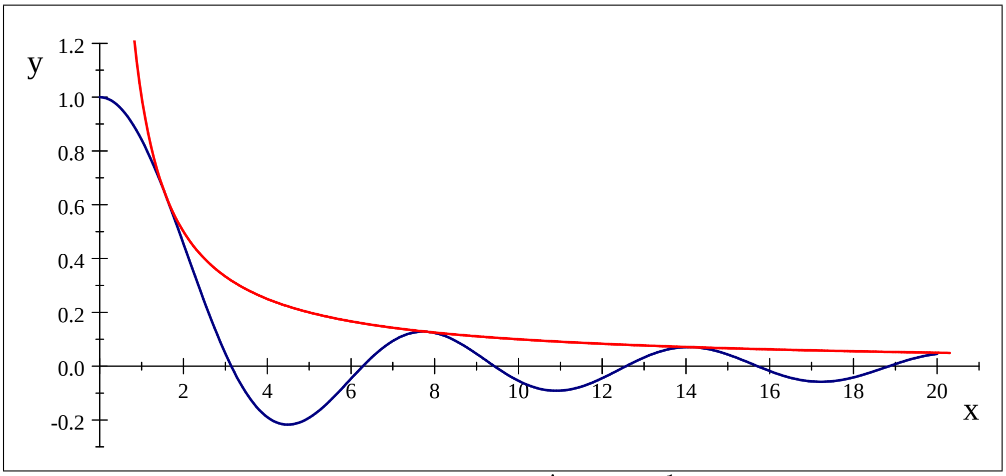
Beispiel. 1. Zeigen wir, dass das Integral $\int_1^\infty \frac{\sin x}{x} dx$ konvergent ist. Dafür betrachten wir die Funktionen $f(x) = \sin x$ und $g(x) = x^{-1}$. Die Stammfunktion

$$F(x) = \int_a^x f(t)\, dt = \int_a^x \sin x dx = -\cos x + \cos a$$

ist offensichtlich beschränkt, während $g(x) \searrow 0$ für $x \to +\infty$. Nach dem Dirichlet-Kriterium ist das Integral $\int_1^\infty \frac{\sin x}{x} dx$ konvergent.

Das Integral $\int_0^\infty \frac{\sin x}{x} dx$ heißt Dirichlet-Integral. Es ist auch konvergent, da die Funktion $\frac{\sin x}{x}$ den Grenzwert 1 für $x \to 0$ hat und somit ist auf $[0, 1]$ integrierbar. Es ist möglich zu beweisen, dass

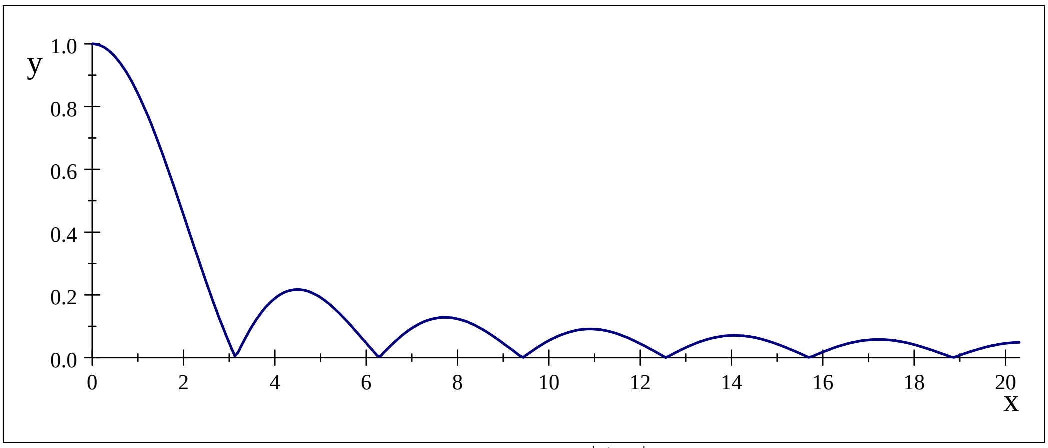
$$\int_0^\infty \frac{\sin x}{x} dx = \frac{1}{2}\pi.$$



Die Funktionen $\frac{\sin x}{x}$ und $\frac{1}{x}$:

Zeigen wir, dass $\int_1^\infty \frac{\sin x}{x}$ bedingt konvergent ist, d.h.

$$\int_1^\infty \frac{|\sin x|}{x} dx = +\infty. \tag{5.11}$$



Die Funktion $\frac{|\sin x|}{x}$

Die Nullstellen von $\sin x$ auf $[1, \infty)$ sind $x_n = \pi n$ für $n \in \mathbb{N}$. Wir haben

$$\begin{aligned}
\int_{x_n}^{x_{n+1}} \frac{|\sin x|}{x} dx &\geq \frac{1}{x_{n+1}} \int_{x_n}^{x_{n+1}} |\sin x| \, dx \\
&\geq \frac{1}{x_{n+1}} \left| \int_{x_n}^{x_{n+1}} \sin x dx \right| \\
&= \frac{1}{x_{n+1}} \left| [\cos x]_{x_n}^{x_{n+1}} \right| \\
&= \frac{2}{\pi (n+1)},
\end{aligned}$$

woraus folgt

$$\int_1^\infty \frac{|\sin x|}{x} dx \geq \sum_{n=1}^\infty \int_{x_n}^{x_{n+1}} \frac{|\sin x|}{x} dx \geq \sum_{n=1}^\infty \frac{2}{\pi (n+1)} = \infty,$$

d.h. (5.11).

2. Zeigen wir, dass $\int_1^\infty \frac{\sin x}{x} \arctan x dx$ konvergiert. Setzen wir $f(x) = \frac{\sin x}{x}$ und $g(x) = \arctan x$. Die Funktion g ist beschränkt und monoton steigend, während $\int_1^{+\infty} f(x)\, dx$ wie oberhalb konvergiert. Nach dem Abel-Kriterium ist das gegebene Integral konvergent.

5.2 Gleichmäßige Konvergenz

5.2.1 Funktionenfolgen

Sei S eine beliebige Menge. Sei $\{f_k\}_{k=1}^\infty$ eine Folge von reellwertigen (oder komplexwertigen) Funktionen auf S.

Definition. Man sagt, dass $\{f_k\}$ gegen eine Funktion f *punktweis auf* S konvergiert falls für jedes $x \in S$ gilt $f_k(x) \to f(x)$ für $k \to \infty$, d.h.

$$\forall x \in S \quad |f_k(x) - f(x)| \to 0 \text{ für } k \to \infty.$$

Die punktweise Konvergenz bezeichnet man mit $f_k \to f$.

Für jede Funktion $f : S \to \mathbb{R}$ definieren wir die sup-Norm von f mit

$$\|f\|_S = \|f\| := \sup_{x \in S} |f(x)|.$$

Definition. Man sagt, dass $\{f_k\}$ gegen f *gleichmässig auf* S konvergiert, falls

$$\|f_k - f\| \to 0 \text{ für } k \to \infty.$$

Die gleichmäßige Konvergenz bezeichnet man mit $f_k \rightrightarrows f$.

Die gleichmäßige Konvergenz ist offensichtlich eine stärkere Bedingung als die punktweise Konvergenz, da sie äquivalent zu

$$\sup_S |f_k - f| \to 0$$

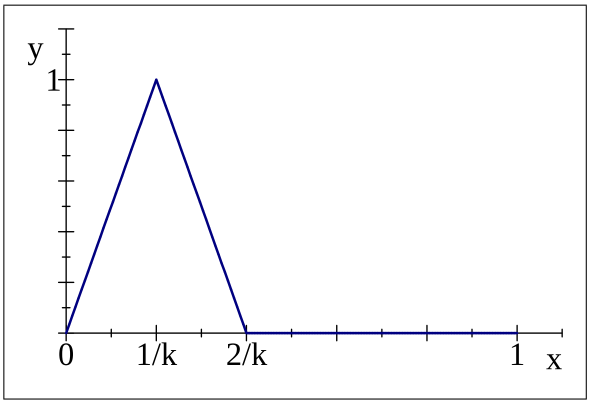
ist, was impliziert $|f_k(x) - f(x)| \to 0$ für jedes $x \in S$.

Die Umkehrung gilt nicht: es gibt punktweis konvergente Folgen die nicht gleichmäßig konvergent sind.

Beispiel. 1. Betrachten wir die Funktionen $f_k(x) = \frac{x}{k}$ auf $S = (0, +\infty)$. Für jedes $x \in S$ gilt offensichtlich $f_k(x) \to 0$ für $k \to \infty$ so dass $f_k \to 0$, aber $f_k \not\rightrightarrows 0$ da $\|f_k\| = \infty$.

2. Ein komplizierteres Beispiel zeigt gleiches auf einem abgeschlossenen beschränkten Intervall. Betrachten wir auf $S = [0, 1]$ die Funktionen

$$f_k(x) = \begin{cases} kx, & 0 \le x \le \frac{1}{k} \\ 2 - kx, & \frac{1}{k} \le x \le \frac{2}{k}, \\ 0, & \frac{2}{k} \le x \le 1. \end{cases}$$



Der Graph von f_k für $k = 5$

Zeigen wir, dass $f_k \to 0$, d.h. für jedes $x \in [0,1]$ gilt $f_k(x) \to 0$. Für $x = 0$ ist es offensichtlich, da $f_k(0) = 0$. Für $x > 0$ gilt $x > 2/k$ für hinreichend große k, woraus $f_k(x) = 0$ folgt. Andererseits, da $\|f_k\| = 1$, so sehen wir, dass $f_k \not\rightrightarrows 0$.

Die gleichmäßige Konvergenz ist wichtig da sie die Stetigkeit von Funktionen bewahrt.

Satz 5.8 *Sei $\{f_k\}$ eine Folge von stetigen Funktionen auf einem Intervall $I \subset \mathbb{R}$. Gilt $f_k \rightrightarrows f$ auf I so ist f auch stetig auf I.*

Beweis. Es reicht zu beweisen, dass f auf jedem beschränkten abgeschlossenen Teilintervall $J \subset I$ stetig ist. Da $\|f_k - f\| \to 0$ für $k \to \infty$, so für jedes $\varepsilon > 0$ existiert $k \in \mathbb{N}$ mit $\|f_k - f\| < \varepsilon/3$, woraus folgt, dass

$$\forall x \in J \quad |f_k(x) - f(x)| < \varepsilon/3.$$

Die Funktion f_k ist auf J gleichmäßig stetig (Lemma 4.8), so dass für ε wie oberhalb

$$\exists \delta > 0 \ \ \forall x, y \in J \text{ mit } |x - y| < \delta \text{ gilt } |f_k(x) - f_k(y)| < \varepsilon/3.$$

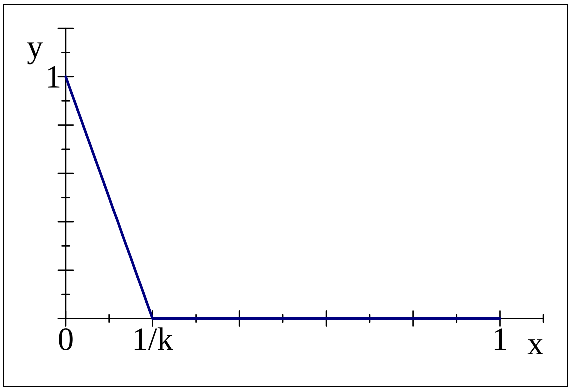
Es folgt, dass für alle $x, y \in J$ mit $|x - y| < \delta$ gilt

$$\begin{aligned} |f(x) - f(y)| &\leq |f(x) - f_k(x)| + |f_k(x) - f_k(y)| + |f_k(y) - f(y)| \\ &< \varepsilon/3 + \varepsilon/3 + \varepsilon/3 = \varepsilon, \end{aligned}$$

was beweist, dass f auf J gleichmäßig stetig ist. ∎

Beispiel. Die punktweise Konvergenz bewahrt die Stetigkeit nicht. Betrachten wir zum Beispiel die Funktionen

$$g_k(x) = \begin{cases} 1 - kx, & 0 \leq x \leq \frac{1}{k}, \\ 0, & \frac{1}{k} \leq x \leq 1. \end{cases}$$



Der Graph von g_k für $k = 5$

Offensichtlich ist g_k stetig auf $[0, 1]$. Für $k \to \infty$ erhalten wir $g_k(x) \to 0$ für $x > 0$ und $g_k(0) \to 1$, d.h. $g_k \to g$ mit

$$g(x) = \begin{cases} 1, & x = 0, \\ 0, & 0 < x \leq 1. \end{cases}$$

Es ist klar, dass g unstetig an 0 ist.

5.2.2 Gleichmäßige Konvergenz von Funktionenreihen

Sei $\{f_k\}_{k=1}^{\infty}$ eine Folge von reellwertigen (oder komplexwertigen) Funktionen auf einer Menge S. Betrachten wir die Funktionenreihe $\sum_{k=1}^{\infty} f_k(x)$ und ihre Partialsummen

$$F_n(x) = \sum_{k=1}^{n} f_k(x).$$

Definition. Die Reihe $\sum_{k=1}^{\infty} f_k(x)$ konvergiert auf S punktweis bzw gleichmäßig falls die Folge $\{F_n\}$ von Partialsummen punktweis bzw gleichmäßig auf S konvergiert.

Die folgende Aussage gibt ein hilfreiches Kriterium für gleichmäßige Konvergenz von Funktionenreihen an.

Satz 5.9 (Weierstraßsches Majorantenkriterium; auch Weierstraßscher *M-Test*) *Sei $\{f_k\}$ eine Funktionenfolge auf S mit*

$$\sum_{k=1}^{\infty} \|f_k\| < \infty.$$

Dann konvergiert die Reihe $\sum_{k=1}^{\infty} f_k(x)$ auf S absolut und gleichmäßig.

Somit folgt die gleichmäßige Konvergenz von einer Funktionenreihe aus der Konvergenz von der numerischen Reihe.

Beweis. Für jedes $x \in S$ gilt

$$\sum_{k=1}^{\infty} |f_k(x)| \leq \sum_{k=1}^{\infty} \|f_k\| < \infty$$

so dass die Reihe $\sum f_k(x)$ absolut konvergent für jedes $x \in S$ ist. Insbesondere ist diese Reihe punktweis konvergent. Setzen wir

$$F(x) = \sum_{k=1}^{\infty} f_k(x)$$

und beweisen, dass $F_n \rightrightarrows F$ für $n \to \infty$. Für jedes $x \in S$ gilt

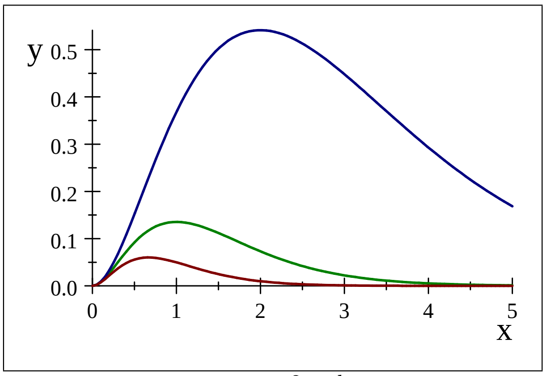
$$|F(x) - F_n(x)| = \left| \sum_{k=n+1}^{\infty} f_k(x) \right| \leq \sum_{k=n+1}^{\infty} |f_k(x)| \leq \sum_{k=n+1}^{\infty} \|f_k\|.$$

Da die rechte Seite unabhängig von $x \in S$ ist, so erhalten wir

$$\|F - F_n\| = \sup_{x \in S} |F(x) - F_n(x)| \leq \sum_{k=n+1}^{\infty} \|f_k\|.$$

Da die rechte Seite für $n \to \infty$ gegen 0 konvergiert, so erhalten wir, dass $\|F - F_n\| \to 0$ und somit $F_n \rightrightarrows F$, was zu beweisen war. ∎

Beispiel. Beweisen wir, dass die Reihe $\sum_{k=1}^{\infty} x^2 e^{-kx}$ gleichmäßig auf $[0, +\infty)$ konvergiert. Die Funktion $f_k(x) = x^2 e^{-kx}$ ist positive für $x > 0$, aber verschwindet at $x = 0$ und konvergiert gegen 0 für $x \to +\infty$. Es folgt, dass f_k eine Maximumstelle auf $[0, \infty)$ hat.



Funktionen $f_k(x) = x^2 e^{-kx}$ für $k = 1, 2, 3$

An der Maximumstelle x von f_k gilt $f_k'(x) = 0$, was äquivalent zu $(\ln f_k)' = 0$, d.h.

$$(2 \ln x - kx)' = 0$$

woraus folgt $x = \frac{2}{k}$. Somit erhalten wir

$$\|f_k\| = \max_{[1,+\infty)} f_k = f_k\left(\frac{2}{k}\right) = \left(\frac{2}{k}\right)^2 e^{-2} = \frac{4e^{-2}}{k^2}.$$

Da die Reihe $\sum \frac{1}{k^2}$ konvergiert, so erhalten wir nach dem Majorantenkriterium dass die Funktionenreihe absolut und gleichmäßig auf $[0, +\infty)$ konvergiert.

Definition. Sei $\{f_k\}_{k=1}^{\infty}$ eine Folge von Funktionen auf einem Intervall I. Die Folge $\{f_k\}$ konvergiert auf I *lokal gleichmäßig*, falls diese Folge auf jedem abgeschlossenen beschränkten Intervall $J \subset I$ gleichmäßig konvergiert. Konvergiert $\{f_k\}$ lokal gleichmäßig gegen f, so schreibt man $f_k \overset{loc}{\rightrightarrows} f$.

Analog konvergiert die Reihe $\sum_{k=1}^{\infty} f_k$ auf I lokal gleichmäßig falls diese Reihe auf jedem abgeschlossenen beschränkten Intervall $J \subset I$ gleichmäßig konvergiert.

Die Beziehung zwischen den Typen von Konvergenz ist wie folgt:

$$f_k \rightrightarrows f \quad \Longrightarrow \quad f_k \overset{loc}{\rightrightarrows} f \quad \Longrightarrow \quad f_k \to f.$$

Beispiel. Die Reihe $\sum_{k=1}^{\infty} x^k$ ist auf $(-1,1)$ lokal gleichmäßig konvergent, da jedes abgeschlossene beschränkte Intervall $J \subset I$ in einem Intervall $[-a,a]$ mit $0 < a < 1$ liegt und somit auf J gilt

$$\sum_{k=1}^{\infty} \left\| x^k \right\|_J \leq \sum_{k=1}^{\infty} a^k < \infty.$$

Andererseits konvergiert die Reihe $\sum_{k=1}^{\infty} x^k$ auf $(-1,1)$ nicht gleichmäßig (siehe Aufgaben).

Satz 5.10 *Sei $\{f_k\}$ eine Folge von stetigen Funktionen auf einem Intervall I. Konvergiert die Folge $\{f_k\}$ lokal gleichmäßig auf I, so ist der Grenzwert $f(x) = \lim f_k(x)$ stetig auf I. Die analoge Eigenschaft gilt auch für die Reihen: die Summe von einer lokal gleichmäßig konvergenten Reihe von stetigen Funktionen ist stetig.*

Beweis. Auf jedem abgeschlossenen beschränkten Intervall $J \subset I$ gilt $f_k \rightrightarrows f$. Nach dem Satz 5.8 ist f stetig auf J und somit ist f stetig auch auf I. Die Aussage für die Reihen folgt, da die Partialsummen stetig und lokal gleichmäßig konvergent sind. ■

5.2.3 Potenzreihen

Betrachten wir eine *Potenzreihe*

$$\sum_{k=0}^{\infty} c_k x^k \tag{5.12}$$

mit Koeffizienten $c_k \in \mathbb{R}$ und bezeichnen wir mit $f(x)$ die Summe der Reihe $\sum_{k=0}^{\infty} c_k x^k$ für alle x wo die Reihe konvergent ist.

Satz 5.11 *Angenommen, dass die Potenzreihe (5.12) für ein $x = x_0 \neq 0$ konvergent ist. Dann konvergiert die Potenzreihe absolut und lokal gleichmäßig auf $(-R,R)$ mit $R = |x_0|$. Insbesondere ist die Summe $f(x) = \sum_{k=0}^{\infty} c_k x^k$ eine stetige Funktion auf $(-R,R)$.*

Beweis. Wählen wir ein $0 < r < R$. Für jedes $x \in [-r,r]$ gilt

$$\left| c_k x^k \right| = \left| c_k x_0^k \left(\frac{x}{x_0} \right)^k \right| \leq \left| c_k x_0^k \right| \left(\frac{r}{R} \right)^k.$$

Die Konvergenz von $\sum_{k=0}^{\infty} c_k x^k$ impliziert, dass $c_k x_0^k \to 0$ für $k \to \infty$. Insbesondere ist die Folge $\left\{ c_k x_0^k \right\}$ beschränkt, zum Beispiel, $\left| c_k x_0^k \right| \leq C$ für alle k und eine Konstante C. Es folgt, dass

$$\left\| c_k x^k \right\|_{[-r,r]} \leq C \left(\frac{r}{R} \right)^k$$

Da $\frac{r}{R} < 1$ und somit die geometrische Reihe $\sum_k \left(\frac{r}{R} \right)^k$ konvergiert, so beschließen wir nach dem M-Test, dass die Potenzreihe $\sum_k c_k x^k$ absolut und gleichmäßig auf $[-r,r]$

konvergiert. Somit ist die Potenzreihe lokal gleichmäßig auf $(-R, R)$ konvergent. Die zweite Aussage folgt aus dem Satz 5.10. ■

Beispiel. Die Potenzreihe $\sum_{k=1}^{\infty} \frac{x^k}{k}$ ist konvergent für $x = -1$, woraus folgt, dass sie auf $(-1, 1)$ konvergent ist und ihre Summe eine stetige Funktion auf $(-1, 1)$ ist.

Liegt x_0 im Definitionsbereich von $f(x) = \sum_{k=0}^{\infty} c_k x^k$ so ist f nach dem Satz 5.11 stetig auf $(-R, R)$ mit $R = |x_0|$. Die Frage entsteht, ob f an der Grenze $x = x_0$ auch stetig ist. Die Antwort wird im folgenden Satz gegeben.

Satz 5.12 (Satz von Abel) *Liegt ein $x_0 \neq 0$ im Definitionsbereich von $f(x) = \sum_{k=0}^{\infty} c_k x^k$, so ist $f(x)$ an $x = x_0$ stetig.*

Beweis. Ohne Beschränkung der Allgemeinheit können wir annehmen, dass $x_0 = 1$ (sonst x durch x/x_0 ersetzen) und $f(x_0) = 0$ (sonst c_0 ändern). Somit gilt

$$\sum_{k=0}^{\infty} c_k = 0,$$

und wir werden beweisen, dass $f(x)$ stetig auf $[0, 1]$ ist, insbesondere an $x = 1$. Bezeichnen wir für jedes $n \in \mathbb{Z}_+$

$$s_n = \sum_{k=0}^{n} c_k$$

und $s_{-1} = 0$, so dass für alle $k \in \mathbb{Z}_+$

$$c_k = s_k - s_{k-1}.$$

Da die Folge $\{s_k\}$ konvergent und somit beschränkt ist, so konvergiert die Reihe

$$\sum_{k=0}^{\infty} s_k x^k$$

für alle $x \in (-1, 1)$ nach dem Majorantenkriterium. Es gilt für alle $x \in (-1, 1)$

$$\begin{aligned} f(x) &= \sum_{k=0}^{\infty} c_k x^k = \sum_{k=0}^{\infty} (s_k - s_{k-1}) x^k \\ &= \sum_{k=0}^{\infty} s_k x^k - \sum_{k=1}^{\infty} s_{k-1} x^k \\ &= \sum_{k=0}^{\infty} s_k x^k - \sum_{k=0}^{\infty} s_k x^{k+1} \\ &= \sum_{k=0}^{\infty} (1 - x) s_k x^k \end{aligned}$$

(Abelsche partielle Summation). Für $x = 1$ gilt diese Identität da $f(1) = 0$. Beweisen wir, dass die obige Reihe gleichmäßig auf $[0, 1]$ konvergiert, woraus die Stetigkeit von f auf $[0, 1]$ folgen wird. Die Partialsumme davon ist

$$f_n(x) = \sum_{k=0}^{n} (1 - x) s_k x^k,$$

und für alle $x \in [0, 1)$ gilt

$$\begin{aligned}
|f(x) - f_n(x)| &= \left|\sum_{k=n+1}^{\infty} (1-x) s_k x^k\right| \leq (1-x) \sup_{k\geq n} |s_k| \sum_{k=n+1}^{\infty} x^k \\
&\leq (1-x) \sup_{k\geq n} |s_k| \frac{1}{1-x} \\
&= \sup_{k\geq n} |s_k| .
\end{aligned}$$

Für $x = 1$ gilt diese Ungleichung auch, da $f(1) = f_n(1) = 0$. Da $s_n \to 0$, so erhalten wir

$$\|f - f_n\|_{[0,1]} \leq \sup_{k\geq n} s_k \to 0 \text{ für } n \to \infty$$

und somit $f_n \rightrightarrows f$ auf $[0, 1]$. Nach dem Satz 5.8 ist f auf $[0, 1]$ stetig. ■

Beispiel. Es gilt die folgende Identität für alle $x \in (-1, 1)$

$$\arctan x = x - \frac{x^3}{3} + \frac{x^5}{5} - \frac{x^7}{7} + ... = \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k+1}}{2k+1} \tag{5.13}$$

(siehe Aufgaben). Die Potenzreihe (5.13) konvergiert auch für $x = 1$ was aus dem Leibniz-Kriterium folgt, da die Glieder monoton fallend und mit abwechselndem Vorzeichen sind. Somit ist die Summe der Reihe stetig an $x = 1$. Da $\arctan x$ auch stetig an $x = 1$ ist, so erhalten wir, dass die Identität (5.13) auch für $x = 1$ gilt (und analog für $x = -1$). Da $\arctan 1 = \pi/4$, so erhalten wir

$$\frac{\pi}{4} = 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + ...$$

5.2.4 Integrations unter gleichmäßiger Konvergenz

Satz 5.13 *Sei $\{f_k\}$ eine Folge von stetigen Funktionen auf einem beschränkten abgeschlossenen Intervall $[a, b]$. Konvergiert f_n gleichmäßig auf $[a, b]$ gegen eine Funktion f, so gilt*

$$\int_a^b f(x)\, dx = \lim_{k\to\infty} \int_a^b f_k(x)\, dx. \tag{5.14}$$

Man kann auch schreiben

$$\int_a^b \lim_{k\to\infty} f_k(x)\, dx = \lim_{k\to\infty} \int_a^b f_k(x)\, dx$$

so dass die Operationen gleichmäßiger lim und $\int_a^b$ vertauschbar sind.

Beweis. Die Funktion f ist stetig nach dem Satz 5.8. Wir haben

$$\left|\int_a^b f_k(x)\, dx - \int_a^b f(x)\, dx\right| = \left|\int_a^b (f_k - f)\, dx\right| \leq \sup_{[a,b]} |f_k - f| (b-a) = \|f_k - f\| (b-a) .$$

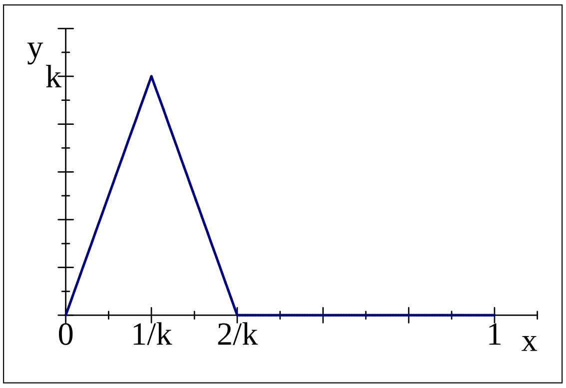
Da $f_k \rightrightarrows f$ und somit $\|f_k - f\| \to 0$, so erhalten wir

$$\left| \int_a^b f_k(x)\, dx - \int_a^b f(x)\, dx \right| \to 0,$$

was zu beweisen war. ■

Beispiel. Die Voraussetzung von der gleichmäßigen Konvergenz ist wichtig für den Satz 5.13. Die punktweise Konvergenz $f_k \to f$ ergibt die Konvergenz von Integralen nicht. Zum Beispiel, betrachten wir die folgenden Funktionen auf $[0,1]$:

$$f_k(x) = \begin{cases} k^2 x, & 0 \le x \le \frac{1}{k}, \\ 2k - k^2 x, & \frac{1}{k} \le x \le \frac{2}{k}, \\ 0, & \frac{2}{k} \le x \le 1. \end{cases}$$



Der Graph von f_k

Es gilt die punktweise Konvergenz $f_k \to 0$ auf $[0,1]$, aber

$$\int_0^1 f_k(x)\, dx = \int_0^{1/k} k^2 x dx + \int_{1/k}^{2/k} \left(2k - k^2 x\right) dx = \frac{1}{2} + 2 - \frac{3}{2} = 1$$

und somit

$$\lim_{k\to\infty} \int_0^1 f_k(x)\, dx = 1 \ne 0 = \int_0^1 \lim_{k\to\infty} f_k dx.$$

Korollar 5.14 *Sei $\{f_k\}$ eine Folge von stetigen auf $[a,b]$ Funktionen. Konvergiert die Reihe $\sum_{k=1}^{\infty} f_k(x)$ gleichmäßig auf $[a,b]$ so gilt*

$$\int_a^b \left(\sum_{k=1}^{\infty} f_k(x) \right) dx = \sum_{k=1}^{\infty} \int_a^b f_k(x)\, dx.$$

In anderen Wörter, die Operationen $\int_a^b$ und $\sum_{k=1}^{\infty}$ über einer Funktionenreihe sind vertauschbar, vorausgesetzt, dass die Reihe gleichmäßig konvergiert.

Beweis. Setzen wir

$$F_n = \sum_{k=1}^{n} f_k \text{ und } F = \sum_{k=1}^{\infty} f_k.$$

Da $F_n \rightrightarrows F$ so erhalten wir nach dem Satz 5.13

$$\int_a^b F_n dx \to \int_a^b F dx,$$

woraus folgt

$$\sum_{k=1}^{\infty} \int_a^b f_k dx = \lim_{n\to\infty} \sum_{k=1}^{n} \int_a^b f_k dx = \lim_{n\to\infty} \int_a^b F_n dx = \int_a^b F dx = \int_a^b \left(\sum_{k=1}^{\infty} f_k \right) dx,$$

was zu beweisen war. ■

Korollar 5.15 *Sei die Potenzreihe $f(x) = \sum_{k=0}^{\infty} c_k x^k$ auf einem Intervall $(-R, R)$ mit $R > 0$ konvergent. Dann gilt für alle $x \in (-R, R)$*

$$\int_0^x f(t)\, dt = \sum_{k=0}^{\infty} \frac{c_k}{k+1} x^{k+1}. \tag{5.15}$$

Beweis. Nach dem Satz 5.11 konvergiert diese Reihe gleichmäßig auf jedem Intervall $[0, x]$ mit $x \in (-R, R)$. Die Integration von 0 bis x ergibt dann

$$\int_0^x f(t)\, dt = \sum_{k=0}^{\infty} \int_0^x c_k t^k dt = \sum_{k=0}^{\infty} \frac{c_k}{k+1} x^{k+1}.$$

■

Beispiel. Betrachten wir die geometrische Potenzreihe

$$\frac{1}{1-x} = \sum_{k=0}^{\infty} x^k,$$

die für $|x| < 1$ konvergiert. Nach (5.15) erhalten wir für alle $x \in (-1, 1)$

$$\int_0^x \frac{dt}{1-t} = \sum_{k=0}^{\infty} \frac{x^{k+1}}{k+1} = x + \frac{x^2}{2} + \frac{x^3}{3} + \frac{x^4}{4} + \ldots$$

Andererseits gilt

$$\int_0^x \frac{dt}{1-t} = -\ln(1-x).$$

Der Vergleich von den beiden Identitäten und der Wechsel von x nach $-x$ ergeben die Identität

$$\ln(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \ldots, \tag{5.16}$$

die für alle $x \in (-1, 1)$ gilt. Die Partialsummen von dieser Reihe sind die Taylor-Polynome der Funktion $\ln(1+x)$.

Bemerken wir, dass die Reihe in (5.16) für $x = 1$ die Leibniz-Reihe ist, die konvergent ist. Nach dem Satz 5.12 ist die Summe der Reihe (5.16) stetig an $x = 1$.

Da $\ln(1+x)$ auch an $x=1$ stetig ist, so sehen wir, dass (5.16) auch an $x=1$ gilt, d.h.

$$\ln 2 = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + ...$$

Beispiel. Betrachten wir $\int_0^x e^{t^2} dt$ (was in elementaren Funktionen nicht darstellbar ist). Es folgt aus der Exponentialreihe, dass

$$e^{x^2} = \sum_{n=0}^{\infty} \frac{x^{2n}}{n!} = 1 + \frac{x^2}{1!} + \frac{x^4}{2!} + ...$$

Da diese Reihe für jedes $x \in \mathbb{R}$ konvergiert, so nach Korollar 5.14 erhalten wir

$$\int_0^x e^{t^2} dt = \sum_{n=0}^{\infty} \int_0^x \frac{t^{2n}}{n!} dt = \sum_{n=0}^{\infty} \frac{x^{2n+1}}{(2n+1)\, n!}.$$

Insbesondere für $x=1$ erhalten wir

$$\int_0^1 e^{t^2} dt = \sum_{n=0}^{\infty} \frac{1}{(2n+1)\,(n!)} = 1 + \frac{1}{3 \cdot 1!} + \frac{1}{5 \cdot 2!} + \frac{1}{7 \cdot 3!} + ...$$

Die Partialsummen dieser Reihe lassen sich als numerische Annäherungen zum Wert von $\int_0^1 e^{t^2} dt$ betrachten. Zum Beispiel, es gelten

$$\sum_{n=0}^{14} \frac{1}{(2n+1)\, n!} \approx 1.462651745\,90716$$

$$\sum_{n=0}^{15} \frac{1}{(2n+1)\, n!} \approx 1.46265174590718$$

$$\sum_{n=0}^{20} \frac{1}{(2n+1)\,(n!)} \approx 1.462\,651745\,90718$$

so dass auch

$$\int_0^1 e^{t^2} dt \approx 1.462\,651745\,90718$$

5.2.5 Ableitung unter gleichmäßiger Konvergenz

Satz 5.16 *Sei $\{f_k\}$ eine Folge von stetig differenzierbaren Funktionen auf einem Intervall $I \subset \mathbb{R}$. Angenommen sei*

- $f_k \to f$ *punktweis auf* I;
- $f_k' \overset{loc}{\rightrightarrows} g$ *auf* I.

Dann gilt $f' = g$. Insbesondere ist f stetig differenzierbar.

Äquivalente Formulierung: konvergiert $\{f_k\}$ punktweis und $\{f_k'\}$ lokal gleichmäßig, so gilt

$$\left(\lim_{k\to\infty} f_k\right)' = \lim_{k\to\infty} f_k',$$

d.h. die Operationen lim und Ableitung vertauschbar sind.

Beweis. Nach der Newton-Leibniz-Formel gilt für alle $x, c \in I$

$$f_k(x) - f_k(c) = \int_c^x f_k'(t)\,dt. \tag{5.17}$$

Nach dem Satz 5.10 ist g stetig auf I. Da $f_k' \rightrightarrows g$ auf $[x, c]$, der Satz 5.13 ergibt

$$\int_c^x f_k'(t)\,dt \to \int_c^x g(t)\,dt \text{ as } k \to \infty.$$

Aus (5.17) erhalten wir für $k \to \infty$, dass

$$f(x) - f(c) = \int_c^x g(t)\,dt.$$

Nach dem Satz 4.18 beschließen wir, dass $f' = g$, was zu beweisen war. ■

Bemerken wir, dass die Behauptung von dem Satz 5.16 über Ableitung ist, während der Beweis auf Integration basiert.

Beispiel. Die Konvergenz $f_k \rightrightarrows f$ allein ergibt $f_k' \to f'$ nicht. Zum Beispiel, die Folge $f_k = \frac{1}{k}\sin kx$ konvergiert gegen 0 gleichmäßig auf $\mathbb{R}$ aber $f_k' = \cos kx$ konvergiert nicht.

Korollar 5.17 *Sei $\{f_k\}$ eine Folge von stetig differenzierbaren Funktionen auf einem Intervall $I \subset \mathbb{R}$. Angenommen sei*

- *Die Reihe $\sum_{k=1}^{\infty} f_k(x)$ ist auf I punktweis konvergent.*
- *Die Reihe $\sum_{k=1}^{\infty} f_k'(x)$ ist auf I lokal gleichmäßig konvergent.*

Dann gilt

$$\left(\sum_{k=1}^{\infty} f_k\right)' = \sum_{k=1}^{\infty} f_k'.$$

Beweis. Anwendung von dem Satz 5.16 für Partialsummen der Reihe ergibt die Behauptung. ■

Satz 5.18 *Sei die Potenzreihe $\sum_{k=0}^{\infty} c_k x^k$ auf einem Intervall $(-R, R)$ mit $R \in (0, +\infty]$ konvergent. Dann ist die Funktion $f(x) = \sum_{k=0}^{\infty} c_k x^k$ unendlich oft differenzierbar auf $(-R, R)$ und es gilt auf $(-R, R)$*

$$f'(x) = \sum_{k=1}^{\infty} c_k k x^{k-1}. \tag{5.18}$$

Beweis. Nach dem Korollar 5.17, es gilt für alle $x \in (-R, R)$

$$f'(x) = \left(\sum_{k=0}^{\infty} c_k x^k\right)' = \sum_{k=0}^{\infty} \left(c_k x^k\right)' = \sum_{k=1}^{\infty} c_k k x^{k-1},$$

vorausgesetzt, dass $\sum_k c_k x^k$ punktweis konvergiert, was gegeben ist, und $\sum_k c_k k x^{k-1}$ lokal gleichmäßig konvergiert. In der Tat zeigen wir, dass die Reihe $\sum_k c_k k x^{k-1}$ gleichmäßig auf jedem Intervall $[-r, r]$ mit $0 < r < R$ konvergiert. Nach dem Majorantenkriterium (Satz 5.9) reicht es zu zeigen, dass

$$\sum_{k=1}^{\infty} \|k c_k x^{k-1}\|_{[r,-r]} < \infty \tag{5.19}$$

Da

$$\left\|c_k k x^{k-1}\right\|_{[-r,r]} = |c_k|\, k r^{k-1},$$

so müssen wir beweisen, dass

$$\sum_{k=1}^{\infty} |c_k|\, k r^{k-1} < \infty.$$

Nach Bernoulli-Ungleichung gilt für alle $\varepsilon > 0$

$$(1+\varepsilon)^k \geq 1 + \varepsilon k > \varepsilon k,$$

woraus folgt

$$k < \frac{(1+\varepsilon)^k}{\varepsilon}$$

und somit

$$\sum_{k=1}^{\infty} |c_k|\, k r^{k-1} < \sum_{k=1}^{\infty} \frac{1}{\varepsilon} |c_k| (1+\varepsilon)^k r^{k-1} = \frac{1}{\varepsilon r} \sum |c_k|\, a^k,$$

wobei $a = (1+\varepsilon) r$. Wählen wir $\varepsilon > 0$ so klein, dass $a < R$. Dann konvergiert die Reihe $\sum c_k a^k$ absolut, woraus (5.19) folgt. Somit ist f stetig differenzierbar und (5.18) gilt.

Da die Potenzreihe (5.18) auf $(-R, R)$ konvergiert, so ergibt die Anwendung von (5.18) zu f' statt f, dass f' stetig differenzierbar ist und

$$f''(x) = \sum_{k=2}^{\infty} c_k k (k-1) x^{k-2}.$$

Per Induktion erhalten wir, dass f unendlich oft auf $(-R, R)$ differenzierbar ist. ■

Beispiel. Die Exponentialreihe

$$\exp(x) = \sum_{k=0}^{\infty} \frac{x^k}{k!},$$

konvergiert auf $(-\infty, +\infty)$. Nach dem Korollar 5.18 erhalten wir

$$\exp(x)' = \sum_{k=0}^{\infty} \left(\frac{x^k}{k!}\right)' = \sum_{k=1}^{\infty} \frac{x^{k-1}}{(k-1)!} = \sum_{n=0}^{\infty} \frac{x^n}{n!} = \exp(x),$$

so dass $(\exp(x))' = \exp(x)$, was wir schon aus Analysis I kennen.

Beispiel. Betrachten wir die Identität

$$\frac{1}{1-x} = \sum_{k=0}^{\infty} x^k$$

auf $(-1, 1)$. Ableiten davon ergibt

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

Wieder Ableiten und Dividieren durch 2 ergeben

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

usw.

Chapter 6

Metrische Räume

In diesem Kapitel entwickeln wir die Basis für Analysis in $\mathbb{R}^n$ und somit auch Analysis von Funktionen mit mehreren Variablen. Wir definieren und untersuchen die Konvergenz von Folgen und Funktionen in $\mathbb{R}^n$ und sogar in allgemeineren *metrischen* Räumen, die mit Hilfe von *Abstandfunktion* definiert werden.

6.1 Abstandfunktion

Definition. Sei X eine beliebige Menge. Eine *Metrik* (=Abstandfunktion) auf X ist eine Funktion $d : X \times X \to \mathbb{R}$ die die folgenden Axiome erfüllt:

1. Positivität: $d(x,y) \geq 0$ für alle $x, y \in X$ und $d(x,y) = 0 \Leftrightarrow x = y$ (somit $d(x,y) > 0$ für alle $x \neq y$).

2. Symmetrie: $d(x,y) = d(y,x)$ für alle $x, y \in X$.

3. Dreiecksungleichung: $d(x,y) \leq d(x,z) + d(y,z)$ für alle $x, y, z \in X$.

Ist d eine Metric auf X, so heißt das Paar (X, d) *metrischer Raum.*

Beispiel. 1. Sei $X = \mathbb{R}$. Dann ist $d(x,y) = |x - y|$ eine Metrik auf $\mathbb{R}$.
2. Für beliebige Menge X ist die folgende Funktion

$$d(x,y) = \begin{cases} 1, & x \neq y, \\ 0, & x = y, \end{cases}$$

eine Metrik. Diese Metrik heißt *diskrete Metrik* auf X.

Unser Hauptbeispiel ist die Menge $\mathbb{R}^n = \underbrace{\mathbb{R} \times \mathbb{R} \times ... \times \mathbb{R}}_{n \text{ times}}$ die aus allen n-Tupeln $(x_1, ..., x_n)$ besteht (wobei alle $x_i \in \mathbb{R}$). Die Elementen von $\mathbb{R}^n$ heißen Vektoren (oder Punkte).

Für $x = (x_1, x_2, ..., x_n)$ heißen alle x_k die Komponenten (oder Koordinaten) von x. Für zwei Vektoren $x, y \in \mathbb{R}^n$ definieren wir die Summe

$$x + y = (x_1 + y_1, x_2 + y_2, ..., x_n + y_n),$$

die auch ein Vektor ist. Für jedes $\lambda \in \mathbb{R}$ und $x \in \mathbb{R}^n$ definieren wir das Produkt λx mit

$$\lambda x := (\lambda x_1, \lambda x_2, ..., \lambda x_n) \in \mathbb{R}^n.$$

Die Menge $\mathbb{R}^n$ mit den obigen Addition und Multiplikation ist ein Vektorraum über $\mathbb{R}$. Erinnern wir uns die Definition von Vektorraum über $\mathbb{R}$ (die reellen Zahlen heißt dann *Skalare*).

Definition. Sei V eine Menge wo die folgenden zwei Operationen definiert werden: Addition

$$x, y \in V \mapsto x + y \in V$$

und Skalarmultiplikation

$$\lambda \in \mathbb{R},\ x \in V \mapsto \lambda x \in V$$

Die Menge V mit diesen Operationen heißt Vektorraum falls die folgenden Axiome erfüllt werden:

1. Nullvektor: es gibt ein $0 \in V$ mit $x + 0 = 0 + x = x$ für alle $x \in V$.
2. Das inverse Element: für jedes $x \in V$ existiert ein $-x \in V$ mit $x + (-x) = (-x) + x = 0$.
3. Assoziativgesetz für Addition: $(x + y) + z = x + (y + z)$.
4. Kommutativgesetz für Addition: $x + y = y + x$.
5. Skalarmultiplikation mit 1: $1x = x$ für alle x.
6. Assoziativgesetz für Skalarmultiplikation: $(\lambda\mu)\, x = \lambda\, (\mu x)$.
7. Distributivgesetz für Addition von Skalaren: $(\lambda + \mu)\, x = \lambda x + \mu x$.
8. Distributivgesetz für Addition von Vektoren: $\lambda\, (x + y) = \lambda x + \lambda y$.

Die Operationen Addition und Skalarmultiplikation heißen zusammen *lineare Operationen.*

Beispiel. 1. Offensichtlich erfüllt $\mathbb{R}^n$ alle Axiome von Vektorraum. Der Nullvektor ist $0 = (0, 0, ..., 0)$ und das inverse Element zu $x \in \mathbb{R}^n$ ist $-x = (-x_1, ..., -x_n)$.

2. Sei S eine beliebige Menge. Bezeichnen wir mit $F(S)$ die Menge von allen reellwertigen Funktionen auf S. Definieren wir Addition von Funktionen $f, g \in F(S)$ mit

$$(f + g)(s) = f(s) + g(s) \quad \forall s \in S$$

und die Skalarmultiplikation mit

$$(\lambda f)(s) = \lambda f(s) \quad \forall s \in S$$

wobei $\lambda \in \mathbb{R}$. Offensichtlich ist $F(S)$ ein Vektorraum wo der Nullvektor 0 ist die Funktion die identisch Null ist, und die inverse Funktion zu f ist $(-f)(s) = -f(s)$.

Insbesondere betrachten wir die Menge $E_n = \{1, 2, ..., n\}$, die aus n Elementen besteht. Dann ist jede Funktion f auf E_n mit der Folge $\{f(1), ..., f(n)\}$ von ihren Werten eindeutig bestimmt. Diese Folge ist ein Element von $\mathbb{R}^n$, so dass f auch als Element von $\mathbb{R}^n$ betrachtet werden kann. Umgekehrt, jedes Element $x \in \mathbb{R}^n$ erzeugt eine Funktion $f \in F(E_n)$ mit $f(k) = x_k$. Somit können wir $F(E_n)$ mit $\mathbb{R}^n$ identifizieren, und diese Identifizierung bewahrt die linearen Operationen.

Definition. Eine Funktion $N : V \to \mathbb{R}$ auf einem Vektorraum V heißt *Norm* falls sie die folgenden Axiome erfüllt:

1. Positivität: $N(x) \geq 0$ und $N(x) = 0 \Leftrightarrow x = 0$ (somit $N(x) > 0$ für alle $x \in V \setminus \{0\}$).

2. Absolute Homogenität: $N(\lambda x) = |\lambda| N(x)\ \forall\ \lambda \in \mathbb{R}$ und $\forall x \in V$.

3. Dreiecksungleichung: $N(x + y) \leq N(x) + N(y)$.

Das Paar (V, N) wobei V ein Vektorraum ist und N eine Norm auf V, heißt *normierter* Vektorraum.

Beispiel. 1. Die Funktion $N(x) = |x|$ auf $\mathbb{R}$ ist eine Norm. Der abstrakte Begriff von Norm ist eine Verallgemeinerung des Begriffes vom Betrag. Um das zu betonen bezeichnet man normalerweise eine beliebige Norm mit $\|x\|$ statt $N(x)$.

2. Sei $V = B(S)$ der Vektorraum von allen *beschränkten* reellwertigen Funktionen auf S. Die Menge $B(S)$ ist offensichtlich ein Unterraum von $F(S)$ und somit auch ein Vektorraum. Die folgende Funktion ist eine Norm in $B(S)$:

$$\|f\| = \sup_{x \in S} |f(x)|$$

(siehe Aufgaben), die heißt die sup-Norm und wird auch mit $\|f\|_{\sup}$ bezeichnet.

3. Sei $V = \mathbb{R}^n$. Die Identifizierung von $\mathbb{R}^n$ mit $B(E_n) = F(E_n)$ wobei $E_n = \{1, 2, ..., n\}$ ergibt die folgende sup-Norm in $\mathbb{R}^n$:

$$\|x\|_{\sup} = \sup_{1 \leq k \leq n} |x_k| = \max_{1 \leq k \leq n} |x_k|.$$

Eine andere Norm in $\mathbb{R}^n$, die heißt die 1-Norm, ist wie folgt definiert:

$$\|x\|_1 = |x_1| + |x_2| + ... + |x_n|.$$

Behauptung. *Ist (V, N) ein normierter Vektorraum, so ist $d(x, y) = N(x - y)$ eine Metrik auf V. Die Metrik d heißt die induzierte Metrik der Norm N.*

Beweis. Beweisen wir alle Axiome von der Metrik, wo wir die Axiome der Norm benutzen.

Positivität: $d(x, y) = N(x - y) \geq 0$ und $d(x, y) = 0 \Leftrightarrow N(x - y) = 0 \Leftrightarrow x = y$.

Symmetrie:

$$d(x, y) = N(y - x) = N((-1)(x - y)) = N(x - y) = d(y, x).$$

Dreiecksungleichung:

$$d(x,y) = N(x-y) = N((x-z)+(z-y)) \leq N(x-z)+N(z-y) = d(x,z)+d(z,y).$$

■

Jeder normierter Vektorraum (V,N) ist somit auch ein metrischer Raum (V,d) mit der induzierten Metrik d.

Beispiel. Für $V = B(S)$ mit der sup-Norm erhalten wir die sup-Metrik:

$$d_{\sup}(f,g) = \|f-g\|_{\sup} = \sup_{x\in S}|f(x)-g(x)|.$$

Insbesondere für $V = \mathbb{R}^n = B(E_n)$ erhalten wir die sup-Metrik

$$d_{\sup}(x,y) = \|x-y\|_{\sup} = \max_{1\leq k\leq n}\{|x_k-y_k|\}.$$

Für $V = \mathbb{R}^n$ mit der 1-Norm erhalten wir die 1-Metrik

$$d_1(x,y) = \|x-y\|_1 = \sum_{k=1}^{n}|x_k-y_k|.$$

Definition. Für jedes $1 \leq p < \infty$ definieren wir die *p-Norm* in $\mathbb{R}^n$ mit

$$\|x\|_p = \left(\sum_{k=1}^{n}|x_k|^p\right)^{1/p}. \tag{6.1}$$

Satz 6.1 *Die p-Norm ist eine Norm in $\mathbb{R}^n$ für alle $p \in [1,\infty)$.*

Beweis. Der Fall $p = 1$ ist offensichtlich, sei jetzt $p > 1$.

Positivität. Es ist klar aus (6.1), dass $\|x\|_p \geq 0$ und

$$\|x\|_p = 0 \Leftrightarrow x_k = 0 \text{ für alle } k = 1,...,n \Leftrightarrow x = 0.$$

Absolute Homogenität:

$$\|\lambda x\|_p = \left(\sum_{k=1}^{n}|\lambda|^p|x_p|^p\right)^{1/p} = |\lambda|\,\|x\|.$$

Die *Dreiecksungleichung* ist deutlich schwieriger und wird erst nach bestimmter Vorbereitung bewiesen.

Behauptung 1. (Hölder-Ungleichung) *Für alle $p,q > 1$ mit*

$$\frac{1}{p}+\frac{1}{q} = 1 \tag{6.2}$$

gilt die folgende Ungleichung

$$|x_1y_1|+|x_2y_2|+...+|x_ny_n| \leq \|x\|_p\|y\|_q \tag{6.3}$$

für alle $x, y \in \mathbb{R}^n$.

Die (6.2) erfüllenden Zahlen heißen die *konjugierten Hölder-Exponenten.*

Gilt $x = 0$ oder $y = 0$ so ist (6.3) trivial. Nehmen wir an, dass $x, y \neq 0$. Die Ungleichung (6.3) ändert sich nicht, wenn x mit einem Skalar $\lambda \neq 0$ multipliziert wird, da die beiden Seiten von (6.3) mit $|\lambda|$ multipliziert werden. Multiplizieren x mit $\lambda = \frac{1}{\|x\|_p}$ und Umbenennen λx zurück nach x lässt uns ohne Beschränkung der Allgemeinheit annehmen, dass $\|x\|_p = 1$. Analog nehmen wir an, dass $\|y\|_q = 1$.

Weiter benutzen wir die Young-Ungleichung

$$\frac{a^p}{p} + \frac{b^q}{q} \geq ab,$$

die unter der Bedingung (6.2) für alle $a, b \geq 0$ gilt. Für $a = |x_k|$ und $b = |y_k|$ erhalten wir

$$\sum_{k=1}^{n} \left(\frac{|x_k|^p}{p} + \frac{|y_k|^q}{q} \right) \geq \sum_{k=1}^{n} |x_k| \, |y_k| \, . \tag{6.4}$$

Mit Hilfe von $\|x\|_p = \|y\|_q = 1$ und (6.2) erhalten wir, dass die linke Seite von (6.4) ist gleich

$$\frac{1}{p} \sum_{k=1}^{n} |x_k|^p + \frac{1}{q} \sum_{k=1}^{n} |y_k|^q = \frac{1}{p} \|x\|_p^p + \frac{1}{q} \|y\|_q^q = \frac{1}{p} + \frac{1}{q} = 1 = \|x\|_p \|y\|_q,$$

woraus (6.3) folgt.

Für zwei Vektoren $x, y \in \mathbb{R}^n$ definieren wir das *Skalarprodukt*

$$x \cdot y := \sum_{k=1}^{n} x_k y_k$$

(nicht mit der Skalarmultiplikation zu vermischen). Bemerken wir, dass $x \cdot y$ eine reelle Zahl ist. Das Skalarprodukt erfüllt offensichtlich die folgenden Eigenschaften:

1. Positivität: $x \cdot x \geq 0$ und $x \cdot x = 0 \Leftrightarrow x = 0$.

2. Symmetrie: $x \cdot y = y \cdot x$

3. Linearität:

   $$(\lambda x) \cdot y = \lambda \, (x \cdot y)$$

   und

   $$(x + y) \cdot z = x \cdot z + y \cdot z$$

   für alle $x, y, z \in \mathbb{R}^n$ und $\lambda \in \mathbb{R}$.

Es folgt aus (6.3) dass

$$x \cdot y \leq \|x\|_p \|y\|_q, \tag{6.5}$$

vorausgesetzt, dass p und q (6.2) erfüllen.

Behauptung 2. *Für konjugierte Hölder-Exponenten* p *und* q *und für alle* $x \in \mathbb{R}^n$ *gilt*

$$\|x\|_p = \sup_{y \in \mathbb{R}^n \setminus \{0\}} \frac{x \cdot y}{\|y\|_q}. \tag{6.6}$$

Nach der Hölder-Ungleichung (6.5) gilt

$$\|x\|_p \geq \frac{x \cdot y}{\|y\|_q}$$

woraus folgt

$$\|x\|_p \geq \sup_{y \neq 0} \frac{x \cdot y}{\|y\|_q}.$$

Jetzt beweisen wir die umgekehrte Ungleichung. Für $x = 0$ sie ist klar. Sonst wählen wir y wie folgt:

$$y_k = x_k \, |x_k|^{p-2}, \quad k = 1, ..., n.$$

Für dieses y haben wir

$$x \cdot y = \sum_{k=1}^{n} |x_k|^p = \|x\|_p^p$$

und mit Hilfe von $q = \frac{p}{p-1}$

$$\|y\|_q = \left(\sum_{k=1}^{n} |y_k|^q \right)^{1/q} = \left(\sum_{k=1}^{n} |x_k|^{(p-1)\frac{p}{p-1}} \right)^{\frac{p-1}{p}} = \|x\|_p^{p-1}.$$

Somit erhalten wir für dieses y, dass

$$\frac{x \cdot y}{\|y\|_q} = \frac{\|x\|_p^p}{\|x\|_p^{p-1}} = \|x\|_p,$$

woraus folgt

$$\sup_{y \neq 0} \frac{x \cdot y}{\|y\|_q} \geq \|x\|_p.$$

As folgt aus diesem Argument, dass sup in (6.6) mit max ersetzt werden kann.

Beweisen wir jetzt die Dreiecksungleichung. Nach Behauptung 2 gilt

$$\begin{aligned} \|x + y\|_p &= \sup_{z \neq 0} \frac{(x + y) \cdot z}{\|z\|^q} = \sup_{z \neq 0} \left(\frac{x \cdot z}{\|z\|^q} + \frac{y \cdot z}{\|z\|^q} \right) \\ &\leq \sup_{z \neq 0} \frac{x \cdot z}{\|z\|^q} + \sup_{z \neq 0} \frac{y \cdot z}{\|z\|^q} \\ &= \|x\|_p + \|y\|_p, \end{aligned}$$

was zu beweisen war. ■

Besonders wichtig ist die 2-Norm:

$$\|x\|_2 = \left(\sum_{k=1}^{n} |x_k|^2 \right)^{1/2},$$

da sie mit dem Skalarprodukt wie folgt verbunden ist: $\|x\|_2 = \sqrt{x \cdot x}$. Da $p = 2$ und $q = 2$ konjugierte Hölder-Exponenten sind, so erhalten wir nach der Hölder-Ungleichung

$$x \cdot y \leq \|x\|_2 \|y\|_2.$$

Dieser spezielle Fall der Hölder-Ungleichung heißt die *Cauchy-Schwarz-Ungleichung.*

Erklären wir jetzt die Beziehung zwischen p-Norm und sup-Norm in $\mathbb{R}^n$.

Behauptung. *Es gilt*

$$\|x\|_p \to \|x\|_{\sup} \text{ für } p \to \infty.$$

(siehe Aufgaben)

Aufgrund dieser Eigenschaft heißt die sup-Norm auch die ∞-Norm und wird mit $\|x\|_\infty$ bezeichnet, d.h.

$$\|x\|_\infty := \|x\|_{\sup} = \lim_{p\to+\infty} \|x\|_p.$$

Bemerken wir, dass $p = \infty$ und $q = 1$ auch konjugierte Hölder-Exponenten sind, da $\frac{1}{\infty} + \frac{1}{1} = 1$. Die Hölder-Ungleichung gilt auch für $p = \infty$ und $q = 1$ wie folgt:

$$x \cdot y \le \|x\|_\infty \|y\|_1$$

(siehe Aufgaben).

In jedem normierten Vektorraum $(V, \|\cdot\|)$ bestimmt die Norm eine Metrik $d(x,y) = \|x-y\|$. Folglich erhalten wir in $\mathbb{R}^n$ die folgende Menge von Metriken: für jedes $p \in [1, +\infty]$

$$d_p(x,y) = \|x-y\|_p = \begin{cases} \left(\sum_{k=1}^n |x_k - y_k|^p\right)^{1/p}, & p < \infty, \\ \max_{1\le k\le n} |x_k - y_k|, & p = \infty. \end{cases}$$

6.2 Metrische Kugel

In einem metrischen Raum (X, d) definieren wir die *metrische Kugeln* wie folgt.

Definition. Für jedes $z \in X$ und $r > 0$ definieren wir die Kugel $B(z,r)$ mit Zentrum z und Radius r wie folgt:

$$B(z,r) = \{x \in X : d(x,z) < r\}.$$

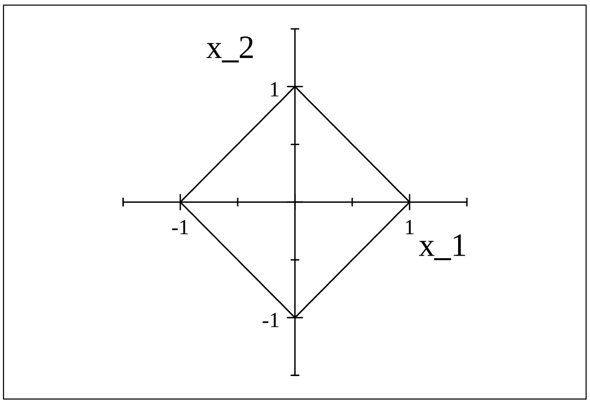
Beispiel. In $\mathbb{R}$ mit der Metrik $d(x,y) = |x-y|$ die Kugel $B(z,r)$ ist das offene Intervall $(z-r, z+r)$.

Beispiel. Betrachten wir $\mathbb{R}^2$ mit der Metrik d_p $(1 \le p \le \infty)$ und beschreiben wir die entsprechende Kugel $B(0,r)$ abhängig von p.

Für $p = 1$ haben wir

$$B(0,r) = \left\{x \in \mathbb{R}^2 : \|x\|_1 < r\right\} = \left\{(x_1,x_2) \in \mathbb{R}^2 : |x_1| + |x_2| < r\right\}$$

Somit ist $B(0,r)$ die Raute (Rhombus) wie auf dem Bild:

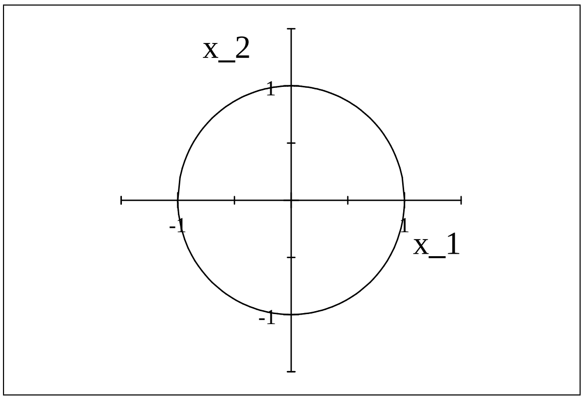


Die metrische Kugel $B(0,1)$ im Fall $p = 1$

Für $p = 2$ haben wir

$$B(0,r) = \left\{x \in \mathbb{R}^2 : \|x\|_2 < r\right\} = \left\{(x_1, x_2) \in \mathbb{R}^2 : x_1^2 + x_2^2 < r^2\right\},$$

und die Kugel ist eine Kreisscheibe

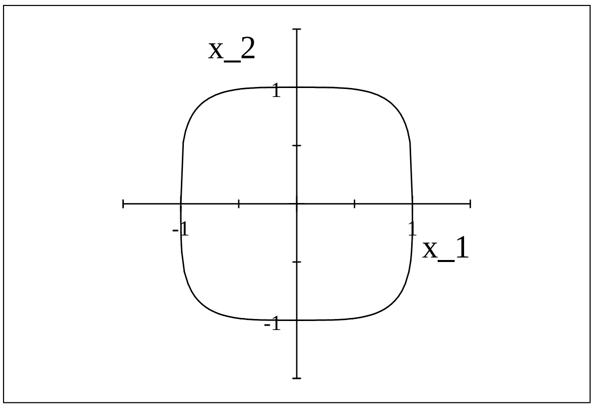


Die metrische Kugel $B(0,r)$ im Fall $p = 2$

Für $p = 4$ haben wir

$$B(0,r) = \left\{x \in \mathbb{R}^2 : \|x\|_4 < r\right\} = \left\{(x_1, x_2) \in \mathbb{R}^2 : x_1^4 + x_2^4 < r^4\right\},$$

was auf dem Bild gezeigt ist:

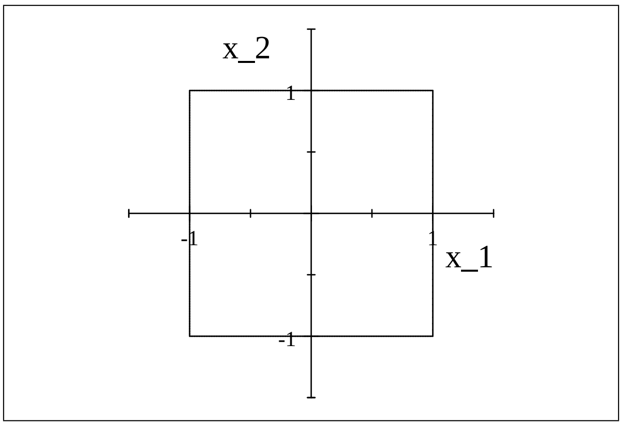


Die metrische Kugel $B(0,r)$ im Fall $p = 4$

Für $p = \infty$ haben wir

$$B(0,r) = \left\{x \in \mathbb{R}^2 : \|x\|_\infty < r\right\} = \left\{(x_1, x_2) \in \mathbb{R}^2 : \max\{|x_1|, |x_2|\} < r\right\},$$

was ein Quadrat ist



Die metrische Kugel $B(0,r)$ im Fall $p = \infty$

Beweisen wir die folgenden Eigenschaften von metrischen Kugeln.

Lemma 6.2 *Seien $B(x_1, r_1)$ und $B(x_2, r_2)$ zwei Kugeln im metrischen Raum (X, d).*
(*a*) *Gilt $d(x_1, x_2) \geq r_1 + r_2$ so sind die Kugeln disjunkt.*
(*b*) *Gilt $d(x_1, x_2) \leq r_1 - r_2$, so haben wir $B(x_2, r_2) \subset B(x_1, r_1)$.*

Beweis. (a) Sei x ein Punkt aus dem Schnitt $B(x_1, r_1) \cap B(x_2, r_2)$. Dann gelten $d(x, x_1) < r_1$ und $d(x, x_2) < r_2$ woraus folgt nach der Dreiecksungleichung

$$d(x_1, x_2) \leq d(x, x_1) + d(x, x_2) < r_1 + r_2,$$

was im g zur Voraussetzung steht.

(b) Für jedes $x \in B(x_2, r_2)$ haben wir

$$d(x, x_1) \leq d(x, x_2) + d(x_1, x_2) < r_2 + d(x_1, x_2) \leq r_1$$

woraus folgt $x \in B(x_1, r_1)$. Somit gilt $B(x_2, r_2) \subset B(x_1, r_1)$. ■

6.3 Konvergenz in metrischen Räumen

Definition. Sei (X, d) ein metrischer Raum. Eine Folge $\{x_n\}_{n=1}^\infty$ von Punkten aus X *konvergiert* gegen ein $a \in X$ falls $d(x_n, a) \to 0$ für $n \to \infty$. Der Punkt a heißt der Grenzwert (=Limes) der Folge $\{x_n\}$ und man schreibt $d\text{-}\lim_{n\to\infty} x_n = a$ oder $x_n \xrightarrow{d} a$.

Äquivalente Definitionen für $x_n \xrightarrow{d} a$:

1. $\forall \varepsilon > 0 \; \exists N \in \mathbb{N} \;\; \forall n \geq N$ gilt $x_n \in B(a, \varepsilon)$ (da $x_n \in B(a, \varepsilon)$ äquivalent zu $d(x_n, a) < \varepsilon$ ist).

2. $\forall \varepsilon > 0$ fast alle x_n liegen in $B(a, \varepsilon)$.

Beispiel. In $\mathbb{R}$ mit Metrik $d(x, y) = |x - y|$ ist die Konvergenz $x_n \xrightarrow{d} a$ äquivalent zu $|x_n - a| \to 0$ und somit zur normalen Konvergenz $x_n \to a$.

Beispiel. Sei S eine Menge. Bezeichnen wir mit $B(S)$ die Menge von allen reellwertigen beschränkten Funktionen auf S. Es ist klar, dass $B(S)$ ein Vektorraum über $\mathbb{R}$ ist. Betrachten auf $B(S)$ die sup-Norm:

$$\|f\|_{\sup} = \sup_{x \in S} |f(x)|.$$

Die sup-Norm ist offensichtlich eine Norm, und we erhalten die entsprechende Metrik

$$d(f, g) = \|f - g\|_{\sup}.$$

Die Konvergenz $f_n \xrightarrow{d} f$ ist äquivalent zu $\|f_n - f\|_{\sup} \to 0$, d.h. zur gleichmäßigen Konvergenz $f_n \rightrightarrows f$ auf S.

Bemerkung. In der Notation $d\text{-}\lim_{n\to\infty} x_n = a$ und $x_n \xrightarrow{d} a$ lässt man häufig "d" ausfallen wenn es klar ist, welche Metrik benutzt wird.

Behauptung. *Jede Folge $\{x_n\}$ in einem metrischen Raum (X, d) hat höchstens einen Grenzwert.*

Beweis. Gelten $x_n \to a$ und $x_n \to b$ so erhalten wir

$$d(a, b) \leq d(a, x_n) + d(x_n, b) \to 0 \text{ für } n \to \infty,$$

woraus $d(a, b) = 0$ und somit $a = b$ folgt. ■

Definition. Seien (X, d_X) und (Y, d_Y) zwei metrische Räume und $f : X \to Y$ eine Funktion (=Abbildung) von X nach Y. Seien $a \in X$ und $b \in Y$. Man schreibt

$$\lim_{x \to a} f(x) = b \quad \text{oder} \quad f(x) \to b \ \text{ für } x \to a \tag{6.7}$$

($f(x)$ konvergiert gegen b für $x \to a$) falls

$$\forall \varepsilon > 0 \ \ \exists \delta > 0 \ \ \forall x \in X \setminus \{a\} \ \text{ mit } \ d_X(x, a) < \delta \text{ gilt } d_Y(f(x), b) < \varepsilon. \tag{6.8}$$

Der Punkt b heißt der Grenzwert (=Limes) von $f(x)$ für $x \to a$.

Bezeichnen wir mit B_X und B_Y die metrischen Kugeln in X bzw Y. Dann ist (6.8) und somit auch (6.7) äquivalent zu

$$\forall \varepsilon > 0 \ \ \exists \delta > 0 \ \ \forall x \in B_X(a, \delta) \setminus \{a\} \ \text{ gilt } f(x) \in B_Y(b, \varepsilon). \tag{6.9}$$

Lemma 6.3 *Die folgenden zwei Bedingungen sind äquivalent.*

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

(*ii*) *Für jede Folge $\{x_n\} \subset X \setminus \{a\}$ mit $x_n \xrightarrow{d_X} a$ gilt $f(x_n) \xrightarrow{d_Y} b$.*

Beweis. $(i) \Rightarrow (ii)$ Nach Voraussetzung gilt (6.9). Wählen wir ein $\varepsilon > 0$ und erhalten nach (6.9) einen Wert $\delta > 0$. Für jede Folge $\{x_n\} \subset X \setminus \{a\}$ mit $x_n \to a$ haben wir nach Definition, dass fast alle x_n in $B_X(a, \delta)$ liegen. Da $x_n \neq a$, so erhalten wir aus (6.9), dass auch fast alle $f(x_n)$ in $B_Y(b, \varepsilon)$ liegen, woraus folgt $f(x_n) \to b$.

$(ii) \Rightarrow (i)$ Angenommen, dass $\lim_{x\to a} f(x) = b$ nicht gilt, so erhalten wir:

$$\exists \varepsilon > 0 \;\; \forall \delta > 0 \; \exists x \in B_X(a, \delta) \setminus \{a\} \;\text{ mit } f(x) \notin B_Y(f(a), \varepsilon). \tag{6.10}$$

Für jedes $k \in \mathbb{N}$ benutzen wir (6.10) mit $\delta = \frac{1}{k}$ und erhalten ein $x_k \in B_X\left(a, \frac{1}{k}\right) \setminus \{a\}$ mit

$$f(x_k) \notin B_Y(f(a), \varepsilon).$$

Für die Folge $\{x_k\}$ gilt $x_k \to a$ aber $f(x_k) \not\to f(a)$, was im Widerspruch zur Voraussetzung (ii) steht. ■

Folglich hat jede Funktion $f : X \to Y$ höchstens einen Grenzwert für $x \to a$.

Definition. Seien X und Y zwei metrische Räume. Eine Funktion $f : X \to Y$ heißt stetig in $a \in X$ falls $\lim_{x\to a} f(x) = f(a)$.

Es folgt aus (6.9), dass die Stetigkeit von f in a äquivalent zur folgenden Bedingung ist

$$\forall \varepsilon > 0 \; \exists \delta > 0 \; \forall x \in B_X(a, \delta) \;\; \text{gilt } f(x) \in B_Y(f(a), \varepsilon). \tag{6.11}$$

Definition. Funktion $f : X \to Y$ heißt stetig falls sie in allen Punkten $a \in X$ stetig ist.

Beispiel. Fixieren wir einen Punkt $x_0 \in X$ und betrachten die Funktion $f : X \to \mathbb{R}$ die mit $f(x) = d(x, x_0)$ definiert ist. Zeigen wir, dass diese Funktion stetig ist, was äquivalent zu $\varepsilon > 0$ there exists $\delta > 0$ such that

$$\forall \varepsilon > 0 \; \exists \delta > 0 \; \forall x \text{ mit } d(x, a) < \delta \;\; \text{gilt } |d(x, x_0) - d(a, x_0)| < \varepsilon.$$

Da nach der Dreiecksungleichung

$$|d(x, x_0) - d(a, x_0)| \leq d(x, a),$$

so reicht es zu nehmen $\delta = \varepsilon$.

6.4 Offene und abgeschlossene Mengen

Sei (X, d) ein metrischer Raum.

Definition. Eine Menge $U \subset X$ heißt *offen* falls für jedes $x \in U$ existiert ein $r > 0$ mit $B(x, r) \subset U$. Eine Menge $F \subset X$ heißt *abgeschlossen* falls das Komplement $X^c := X \setminus F$ offen ist.

Beispiel. 1. Die leere Menge $\emptyset$ und die ganze Menge X sind offen. Somit sind $X = \emptyset^c$ und $\emptyset = X^c$ abgeschlossen.

2. In $\mathbb{R}$ die offenen Intervalle sind offene Mengen und die abgeschlossenen Intervalle sind abgeschlossene Mengen.

3. Zeigen wir, dass offene metrische Kugel $B(z,r) = \{x \in X : d(z,x) < r\}$ eine offene Menge ist. Für jedes $x \in B(z,r)$ gilt $d(z,x) < r$ und somit

$$\varepsilon := r - d(z,x) > 0.$$

Dann $r - \varepsilon = d(z,x)$, und nach Lemma 6.2 erhalten wir, dass $B(x,\varepsilon) \subset B(z,r)$. Somit ist $B(z,r)$ offene Menge.

Satz 6.4 *Alle Mengen hier sind Teilmengen von einem metrischen Raum X.*

(a) *Jede Vereinigung von offenen Mengen ist offen.*

(b) *Der Schnitt endlich vieler offenen Mengen ist offen.*

(c) *Jeder Schnitt von abgeschlossenen Mengen ist abgeschlossen.*

(d) *Die Vereinigung endlich vieler abgeschlossenen Mengen ist abgeschlossen.*

(e) *Eine Menge U ist offen genau dann, wenn U eine Vereinigung von offenen metrischen Kugeln ist.*

(f) *Eine Menge F ist abgeschlossen genau dann, wenn jede konvergente Folge aus F den Grenzwert auch in F hat.*

Beweis. (a) Sei $\{U_\alpha\}_{\alpha\in I}$ ein Mengensystem von offenen Mengen. Zeigen wir, dass die Vereinigung $U = \bigcup_{\alpha\in I} U_\alpha$ offen ist, d.h. für jedes $x \in U$ existiert $r > 0$ mit $B(x,r) \subset U$. In der Tat liegt x in einem U_α. Da U_α offen ist, so existiert $r > 0$ mit $B(x,r) \subset U_\alpha$. Somit gilt auch $B(x,r) \subset U$.

(b) Seien $U_1, U_2, ..., U_n$ offene Mengen. Zeigen wir, dass der Schnitt $U = \bigcap_{k=1}^n U_k$ offen ist. Jedes $x \in U$ liegt in allen U_k. Somit existiert für jedes $k = 1, ..., n$ ein $r_k > 0$ mit $B(x,r_k) \subset U_k$. Dann für

$$r := \min(r_1, r_2, ..., r_n) > 0$$

gilt $B(x,r) \subset U_k$ für alle $k = 1, ..., m$, woraus folgt $B(x,r) \subset U$.

Bemerkung. Der Schnitt unendlich vieler offenen Mengen muss nicht offen sein. Zum Beispiel, der Schnitt von allen offenen Intervallen $\left(-\frac{1}{n}, \frac{1}{n}\right)$ mit $n \in \mathbb{N}$ ist gleich $\{0\}$, was nicht offen ist.

(c) Ist $\{F_\alpha\}_\alpha$ ein Mengensystem von abgeschlossenen Mengen, so ist die Menge

$$\left(\bigcap_{\alpha\in I} F_\alpha\right)^c = \bigcup_{\alpha\in I} F_\alpha^c$$

offen nach (a). Somit ist $\bigcap_{\alpha\in I} F_\alpha$ abgeschlossen.

(d) Sind $F_1, ..., F_n$ abgeschlossene Mengen, so ist die Menge

$$\left(\bigcup_{k=1}^n F_k\right)^c = \bigcap_{k=1}^n F_k^c$$

offen nach (b). Somit ist $\bigcup_{k=1}^n F_k$ abgeschlossen..

(e) Sei U offen. Dann für jedes $x \in U$ existiert ein $r_x > 0$ mit $B(x,r_x) \subset U$. Es ist klar, dass

$$U = \bigcup_{x\in U} B(x,r_x),$$

so dass U eine Vereinigung von offenen Kugeln ist. Umgekehrt, jede Vereinigung von offenen Kugeln ist nach (a) eine offene Menge, da alle Kugeln offen sind.

(f) Sei F abgeschlossen. Sei $\{x_n\}$ eine Folge von Elementen von F die gegen ein $a \in X$ konvergiert. Zeigen wir, dass $a \in F$. Nehmen wir das Gegenteil an, dass $a \notin F$ und somit $a \in F^c$. Da F^c offen ist, so existiert ein $r > 0$ mit $B(a,r) \subset F^c$. Da $x_n \in F$, so erhalten wir, dass $x_n \notin B(a,r)$, woraus folgt, dass $x_n \not\to a$.

Beweisen wir jetzt die Umkehrung: enthält F die Grenzwerte von allen konvergenten Folgen aus F, so ist F abgeschlossen. Dafür zeigen wir, dass F^c offen ist, d.h. für jedes $a \in F^c$ existiert ein $r > 0$ mit $B(a,r) \subset F^c$. Nehmen wir das Gegenteil an: existiert ein $a \in F^c$ so dass keine Kugel $B(a,r)$ mit $r > 0$ in F^c liegt. Insbesondere existiert für jedes $k \in \mathbb{N}$ ein $x_k \in B\left(a, \frac{1}{k}\right)$ mit $x_k \notin F^c$, d.h. $x_k \in F$. Somit erhalten wir eine Folge $\{x_k\}$ von Punkten aus F mit $d(x_k, a) < \frac{1}{k}$, woraus folgt $x_k \to a$. Dann soll auch a in F liegen, was im Widerspruch mit $a \in F^c$ steht. ∎

Bemerkung. Man kann die offenen Menge axiomatisch wie folgt definieren. Betrachten wir eine Menge X und ein Mengensystem $\mathcal{O} \subset 2^X$. Die Elementen von $\mathcal{O}$ heißt offene Mengen, falls $\mathcal{O}$ die folgenden Axiome erfüllt:

1. $\emptyset \in \mathcal{O}$ und $X \in \mathcal{O}$.
2. Jede Vereinigung von Elementen von $\mathcal{O}$ ist wieder ein Element von $\mathcal{O}$.
3. Der Schnitt endlich vieler Elemente von $\mathcal{O}$ ist auch Element von $\mathcal{O}$.

Jedes Mengensystem $\mathcal{O}$ mit diesen Eigenschaften heißt *Topologie* in X, und das Paar $(X, \mathcal{O})$ heißt *topologischer Raum*. Die Topologie in X lässt die Begriffe von Konvergenz von Folgen, stetigen Funktionen usw. definieren.

In diesem Kurs Betrachten wir nur die Topologie von metrischen Raum, die durch eine Metrik definiert ist.

Beispiel. Betrachten wir eine *abgeschlossene* Kugel

$$\overline{B}(z,r) = \{x \in X : d(z,x) \le r\}$$

und beweisen, dass sie abgeschlossene Menge ist. Dafür reicht es zu zeigen, dass das Komplement

$$C := \overline{B}(z,r)^c = \{x \in X : d(z,x) > r\}$$

offen ist. Zeigen wir, dass für jedes $x \in C$ existiert ein $\varepsilon > 0$ mit $B(x,\varepsilon) \subset C$. In der Tat setzen wir

$$\varepsilon := d(z,x) - r > 0$$

so dass $r + \varepsilon = d(z,x)$. Analog zum Lemma 6.2 beweist man, dass die Kugeln $\overline{B}(z,r)$ und $B(x,\varepsilon)$ disjunkt sind. Es folgt, dass $B(x,\varepsilon) \subset C$, was zu beweisen war.

Es folgt, dass jede abgeschlossene Kugel die Grenzwerte von allen konvergenten Folgen aus dieser Kugel enthält.

Satz 6.5 *Seien X, Y zwei metrischen Räumen und $f : X \to Y$ eine Abbildung.*

(a) f ist stetig genau dann, wenn für jede offene Menge $U \subset Y$ das Urbild $f^{-1}(U)$ eine offene Menge in X ist.

(b) f ist stetig genau dann, wenn für jede abgeschlossene Menge $F \subset Y$ das Urbild $f^{-1}(F)$ eine abgeschlossene Menge in X ist.

(c) f ist stetig in $a \in X$ falls $f(x_n) \to f(a)$ für jede Folge $\{x_n\} \subset X$ mit $x_n \to a$.

Beweis. (a) Sei f stetig. Beweisen wir, dass $f^{-1}(U)$ offen für jede offene Menge $U \subset Y$ ist. Für jedes $a \in f^{-1}(U)$ haben wir $f(a) \in U$. Da U offen, so existiert ein $\varepsilon > 0$ mit

$$B_Y(f(a), \varepsilon) \subset U. \tag{6.12}$$

Nach der Stetigkeit von f in a gibt es ein $\delta > 0$ mit

$$x \in B_X(a, \delta) \implies f(x) \in B_Y(f(a), \varepsilon). \tag{6.13}$$

Es folgt aus (6.12) und (6.13) dass

$$f(B_X(a, \delta)) \subset U,$$

woraus folgt

$$B_X(a, \delta) \subset f^{-1}(U),$$

und $f^{-1}(U)$ ist offen.

Sei $f^{-1}(U)$ offen für jede offene Menge $U \subset Y$. Beweisen wir, dass f stetig an jeder Stelle $a \in X$. Wir zeigen: für jedes $\varepsilon > 0$ existiert $\delta > 0$ mit (6.13). Die Kugel $B_Y(f(a), \varepsilon)$ ist eine offene Menge in Y. Somit ist ihres Urbild V eine offene Menge in X. Da $a \in V$, so existiert ein $\delta > 0$ mit

$$B_X(a, \delta) \subset V,$$

woraus folgt

$$f(B_X(a, \delta)) \subset f(V) \subset B_Y(f(a), \varepsilon),$$

und somit (6.13).

(b) Gegeben sei, dass $f^{-1}(F)$ abgeschlossen in X für jede abgeschlossen Menge $F \subset Y$ ist. Setzen wir $U = F^c$ und bemerken, dass

$$f^{-1}(F) = f^{-1}(U^c) = f^{-1}(U)^c$$

und somit

$$\left(f^{-1}(F)\right)^c = f^{-1}(U).$$

Die Menge F ist abgeschlossen genau dass, wenn U offen ist, und $f^{-1}(F)$ ist abgeschlossen genau dann, wenn $f^{-1}(U)$ offen ist. Somit ist die Abgeschlossenheit von $f^{-1}(F)$ für jede abgeschlossene Menge $F \subset Y$ äquivalent zur Offenheit von $f^{-1}(U)$ für jede offene Menge $U \subset Y$, was nach (a) äquivalent zur Stetigkeit von f ist.

(c) Nach Lemma 6.3 $f(x) \to f(a)$ für $x \to a$ gilt genau dann, wenn $f(x_n) \to f(a)$ für jede Folge $\{x_n\} \subset X \setminus \{a\}$ mit $x_n \to a$. Die letzte Bedingung impliziert, dass $f(x_n) \to f(a)$ auch für jede Folge $\{x_n\} \subset X$ mit $x_n \to a$, da für die Glieder x_n der Folge, die a gleich sind, gilt $f(x_n) = f(a)$. ■

Korollar 6.6 *Seien $f : X \to Y$ und $g : Y \to Z$ zwei stetige Abbildungen von metrischen Räumen. Dann ist die Verkettung $g \circ f : X \to Z$ stetig.*

Beweis. Wir haben
$$X \xrightarrow{f} Y \xrightarrow{g} Z.$$
Da $(g \circ f)^{-1} = f^{-1} \circ g^{-1}$, so gilt für jede Menge $U \subset Z$
$$(g \circ f)^{-1}(U) = f^{-1}\left(g^{-1}(U)\right).$$
Ist U offen in Z, so ist $g^{-1}(U)$ offen in Y und somit $f^{-1}(g^{-1}(U))$ ist offen in X. Wir beschließen, dass $(g \circ f)^{-1}(U)$ offen in X ist, und die Abbildung $g \circ f$ stetig nach dem Satz 6.5 ist. ■

Korollar 6.7 *Sind $f, g : X \to \mathbb{R}$ stetige Abbildungen dann auch $f + g$, $fg, f/g$ stetig sind (im Fall f/g vorausgesetzt, dass $g \neq 0$).*

Beweis. Zeigen wir, z.B., dass $f + g$ stetig an jeder Stelle $a \in X$ ist. Für jede Folge $x_n \to a$ gilt nach dem Satz 6.5 $f(x_n) \to f(a)$ und $g(x_n) \to g(a)$. Daraus folgt, dass $f(x_n) + g(x_n) \to f(a) + g(a)$, und nach dem Satz 6.5 beschließen wir, dass $f + g$ an a stetig ist.
■

Betrachten wir die Metriken in einem Vektorraum V, insbesondere in $V = \mathbb{R}^n$.

Definition. Seien N' und N'' zwei Normen in V. Man sagt, dass N' und N'' äquivalent sind, fall es die positiven Konstanten C_1, C_2 gibt mit
$$C_2 N'(x) \leq N''(x) \leq C_1 N'(x) \tag{6.14}$$
für alle $x \in V$.

Es ist einfach zu sehen, dass die Äquivalenz von Normen eine Äquivalenzrelation ist.

Satz 6.8 *Seien N' und N'' zwei äquivalente Normen in V. Seien d' und d'' die von N' bzw N'' induzierten Metriken auf V, d.h.*
$$d'(x, y) = N'(x - y) \quad \textit{und } d''(x, y) = N''(x - y).$$
Dann gelten die folgenden Aussagen.

(a) Die Begriffe der Konvergenz von Folgen in V bezüglich d' und d'' stimmen überein.

(b) Die Begriffe von stetigen Abbildung von V bezüglich d' und d'' stimmen überein.

(c) Die Topologien (=Mengensysteme von offenen Mengen) in V bezüglich d' und d'' stimmen überein.

Beweis. (a) Die Bedingung $x_k \xrightarrow{d'} a$ ergibt $d'(x_k, a) \to 0$, d.h. $N'(x_k - a) \to 0$. Nach (6.14) erhalten wir $N''(x_k - a) \to 0$ woraus folgt $d''(x_k, a) \to 0$ und $x_k \xrightarrow{d''} a$.

(b) Die Stetigkeit lässt sich mit Hilfe von Konvergenz formulieren (siehe Satz 6.5), so dass die Aussage (b) aus (a) folgt.

(c) Die abgeschlossen Mengen lassen sich mit Hilfe von Konvergenz formulieren (siehe Satz 6.4). Somit folgt (c) aus (a). ■

Satz 6.9 *Alle p-Normen in $\mathbb{R}^n$ für $p \in [1, +\infty]$ sind äquivalent.*

Unterhalb wir nehmen immer an, dass die Topologie in $\mathbb{R}^n$ mit Hilfe von einer der p-Normen definiert wird.

Beweis. Es reicht zu zeigen, dass jede p-Norm zu ∞-Norm äquivalent ist. Für jede $p \in [1, \infty)$ gilt

$$\|x\|_p = \left(\sum_{k=1}^n |x_k|^p\right)^{1/p} \geq (\max\{|x_k|\}^p)^{1/p} = \|x\|_\infty$$

und

$$\|x\|_p = \left(\sum_{k=1}^n |x_k|^p\right)^{1/p} \leq (n \max\{|x_k|\}^p)^{1/p} = n^{1/p}\|x\|_\infty$$

woraus folgt

$$\|x\|_\infty \leq \|x\|_p \leq n^{1/p}\|x\|_\infty, \tag{6.15}$$

was zu beweisen war. ■

6.5 Vollständigkeit

Definition. Sei (X, d) ein metrischer Raum. Eine Folge $\{x_n\} \subset X$ heißt Cauchy-Folge, falls

$$d(x_n, x_m) \to 0 \text{ für } n, m \to \infty.$$

Gilt $x_n \to a$, so erhalten wir

$$d(x_n, x_m) \leq d(x_n, a) + d(x_m, a) \to 0 \text{ as } n, m \to \infty.$$

Somit ist jede konvergente Folge auch Cauchy-Folge. Die Umkehrung gilt in $\mathbb{R}$, aber nicht für beliebige metrische Räume.

Beispiel. Betrachten wir den Raum $X = (0, +\infty)$ mit der Metrik $d(x, y) = |x - y|$. Die Folge $x_n = \frac{1}{n}$ ist eine Cauchy-Folge in (X, d), aber diese Folge ist in (X, d) nicht konvergent.

Beispiel. Im Raum $X = [a, b]$ mit $d(x, y) = |x - y|$ ist jede Cauchy-Folge $\{x_n\}$ konvergent ist, da sie einen Grenzwert x in $\mathbb{R}$ hat, und $x_n \in [a, b]$ ergibt $x \in [a, b]$. Gleiches gilt für $X = [a, +\infty)$ und $X = (-\infty, b]$.

Definition. Ein metrischer Raum (X, d) heißt *vollständig* falls jede Cauchy-Schwarz in X konvergent ist. Ein normierter Vektorraum (V, N) heißt vollständig falls der metrische Raum (V, d) mit der induzierten Metrik $d(x, y) = N(x - y)$ vollständig ist. Ein vollständiger normierter Vektorraum heißt *Banachraum*.

Wir haben oberhalb gesehen, dass $(0, \infty)$ nicht vollständig ist, während $[a, b]$, $[a, \infty)$, $(-\infty, b]$ vollständig sind.

Satz 6.10 *Sei S beliebige Menge. Der normierte Vektorraum $B(S)$ von allen reellwertigen beschränkten Funktionen auf S mit der* sup*-Norm ist ein Banachraum.*

Beweis. Bezeichnen wir die sup-Norm mit $\|f\|$, d.h.

$$\|f\| = \sup_S |f| .$$

Sei $\{f_n\} \subset B(S)$ eine Cauchy-Folge von, d.h.

$$\|f_n - f_m\| \to 0 \text{ für } n, m \to \infty. \tag{6.16}$$

Zeigen wir, dass die Folge $\{f_n\}$ gleichmäßig konvergiert. Für jedes $x \in S$ haben wir

$$|f_n(x) - f_m(x)| \le \sup |f_n - f_m| = \|f_n - f_m\| \to 0 \quad \text{für } n, m \to \infty,$$

woraus folgt, dass für jedes $x \in S$ die Folge $\{f_n(x)\}$ von reellen Zahlen eine Cauchy-Folge in $\mathbb{R}$ ist. Dann die Folge $\{f_n(x)\}$ ist konvergent für jedes $x \in S$. Setzen wir

$$f(x) := \lim_{n\to\infty} f_n(x) .$$

Zeigen wir, dass die Funktion f auf S beschränkt ist und dass $f_n \rightrightarrows f$ auf S. Nach (6.16) gilt:

$$\forall \varepsilon > 0 \ \exists N \in \mathbb{N} \ \ \forall n, m \ge N \quad \sup |f_n - f_m| < \varepsilon.$$

Insbesondere für jedes $x \in S$ gilt

$$|f_n(x) - f_m(x)| < \varepsilon.$$

Für $m \to \infty$ erhalten wir, dass für jedes $x \in S$

$$|f_n(x) - f(x)| \le \varepsilon$$

woraus folgt

$$\forall \varepsilon > 0 \ \ \exists N \in \mathbb{N} \ \ \forall n \ge N \quad \|f_n - f\| \le \varepsilon.$$

Somit erhalten wir

$$\|f_n\| \le \|f\| + \varepsilon < \infty$$

d.h. $f \in B(S)$, und $\|f_n - f\| \to 0$ für $n \to \infty$, was zu beweisen war. ∎

Korollar 6.11 *$\mathbb{R}^n$ ist Banachraum bezüglich jeder p-Norm, $p \in [1, \infty]$.*

Beweis. Wir wissen schon, dass $\mathbb{R}^n = B(E_n)$ mit $E_n = \{1, 2, ..., n\}$, da jede Funktion $f : E_n \to \mathbb{R}$ mit der Folge $\{f(1), ..., f(n)\} \in \mathbb{R}^n$ identifiziert wird. Nach dem Satz 6.10 ist $\mathbb{R}^n = B(E_n)$ vollständig bezüglich der sup-Norm, die mit ∞-Norm übereinstimmt. Nach dem Satz 6.9 ist $\mathbb{R}^n$ vollständig auch bezüglich p-Norm. ∎

Korollar 6.12 *Die Menge $C[a, b]$ von allen reellwertigen stetigen Funktionen auf dem Intervall $[a, b]$ ist ein Banachraum bezüglich der* sup*-Norm.*

Beweis. Alle stetige Funktionen auf $[a, b]$ sind beschränkt, so dass $C[a, b] \subset B[a, b]$. Darüber hinaus ist $C[a, b]$ ein Unterraum von $B[a, b]$. Insbesondere ist die sup-Norm eine Norm in $C[a, b]$. Sei $\{f_n\}_{n=1}^{\infty}$ eine Cauchy-Folge in $C[a, b]$. Dann ist $\{f_n\}$ auch eine Cauchy-Folge in $B[a, b]$ und somit konvergiert gleichmäßig gegen einer Funktion $f \in B[a, b]$. Nach dem Satz 5.8 ist f stetig, d.h. $f \in C[a, b]$ und somit ist $C[a, b]$ vollständig. ∎

6.6 Fixpunktsatz von Banach

Definition. Sei (X, d) ein metrischer Raum und $f : X \to X$ eine Selbstabbildung. Ein Punkt $x \in X$ heißt *Fixpunkt* von f falls $f(x) = x$.

In diesem Abschnitt betrachten wir die Bedingungen für Existenz eines Fixpunktes und das Verfahren den Fixpunkt zu bestimmen.

Beispiel. Zeigen wir, dass jede stetige Funktion $f : [0, 1] \to [0, 1]$ einen Fixpunkt hat. In der Tat ist die Funktion $f(x) - x$ nichtnegativ an $x = 0$ und nichtpositiv an $x = 1$. Nach dem Zwischenwertsatz hat die Funktion $f(x) - x$ eine Nullstelle, die ein Fixpunkt von f ist.

Andererseits es gibt viele Abbildungen $f : \mathbb{R} \to \mathbb{R}$ ohne Fixpunkt, z.B. $f(x) = x + 1$.

Beispiel. Sei P eine reellwertige Funktion auf $\mathbb{R}$. Jede Nullstelle x von P erfüllt offensichtlich die Gleichung

$$x = x - cP(x)$$

für beliebige Konstante $c \neq 0$, d.h. die Nullstelle von P stimmt mit dem Fixpunkt der Funktion $f(x) = x - cP(x)$ überein. Berechnen der Nullstelle einer Funktion ist somit äquivalent zur Berechnen des Fixpunktes anderer Funktion.

Definition. Eine Selbstabbildung $f : X \to X$ heißt *Kontraktionsabbildung* falls es eine Konstante $q \in (0, 1)$ gibt mit

$$d(f(x), f(y)) \leq q d(x, y) \quad \forall x, y \in X. \tag{6.17}$$

Hauptsatz 6.13 (Fixpunktsatz von Banach) *Seien (X, d) ein vollständiger metrischer Raum und $f : X \to X$ eine Kontraktionsabbildung. Dann besitzt f genau einen Fixpunkt.*

Beweis. Wählen wir einen beliebigen Punkt $x_0 \in X$ und definieren nach Induktion eine Folge $\{x_n\}_{n=0}^{\infty}$ durch

$$x_{n+1} = f(x_n), \quad n = 0, 1, 2, ...$$

Wir beweisen, dass die Folge $\{x_n\}$ konvergiert und der Grenzwert ein Fixpunkt von f ist.

Bemerken wir, dass

$$d(x_{n+1}, x_n) = d(f(x_n), f(x_{n-1})) \leq q d(x_n, x_{n-1}).$$

Per Induktion erhalten wir, dass

$$d(x_{n+1}, x_n) \leq q^n d(x_1, x_0).$$

Setzen wir $C = d(x_1, x_0)$ und erhalten für alle $n \in \mathbb{N}$

$$d(x_{n+1}, x_n) \leq Cq^n. \tag{6.18}$$

Behauptung. *Es folgt aus (6.18) mit $q < 1$ dass $\{x_n\}$ eine Cauchy-Folge ist.*

In der Tat, für jede $m > n$ erhalten wir nach Dreiecksungleichung und (6.18), dass

$$\begin{aligned} d(x_m, x_n) &\leq d(x_n, x_{n+1}) + d(x_{n+1}, x_{n+2}) + ... + d(x_{m-1}, x_m) \\ &\leq C\left(q^n + q^{n+1} + ... + q^{m-1}\right) \\ &\leq Cq^n \sum_{k=0}^{\infty} q^k \\ &= \frac{Cq^n}{1-q} \to 0 \text{ für } n \to \infty. \end{aligned}$$

Insbesondere gilt $d(x_m, x_n) \to 0$ für $n, m \to \infty$, was bedeutet, dass $\{x_n\}$ eine Cauchy-Folge ist.

Nach der Vollständigkeit von (X, d) konvergiert die Folge $\{x_n\}$ gegen einen Punkt $a \in X$. Daraus folgt $f(x_n) \to f(a)$, da

$$d(f(x_n), f(a)) \leq qd(x_n, a) \to 0.$$

Andererseits $f(x_n) = x_{n+1} \to a$, was ergibt $f(a) = a$. Also, a ist ein Fixpunkt.

Sind a, b zwei Fixpunkte, so gilt es nach (6.17)

$$d(a, b) = d(f(a), f(b)) \leq qd(a, b),$$

was nur dann möglich ist, wenn $d(a, b) = 0$ und somit $a = b$. ∎

Bemerkung. Der Beweis des Fixpunktsatzes ergibt die folgende Methode um den Fixpunkt zu bestimmen bzw anzunähern. Man fängt mit einem beliebigen Punkt x_0 an und bildet induktiv die Folge von *Näherungslösungen* wie folgt:

$$x_{n+1} = f(x_n), \tag{6.19}$$

die gegen Fixpunkt konvergiert. Die Folge $\{x_n\}$ heißt die *Fixpunktiteration.*

Beispiel. Fixieren ein $a > 0$ und betrachten die Funktion $P(x) = x - \frac{a}{x}$ auf $(0, \infty)$, deren Nullstelle ist $x = \sqrt{a}$. Somit ist $\sqrt{a}$ auch ein Fixpunkt der Funktion

$$f(x) = x - \frac{1}{2}P(x) = \frac{1}{2}\left(x + \frac{a}{x}\right).$$

Man kann zeigen, dass f auf $X = [\sqrt{a}, \infty)$ Selbstabbildung und sogar eine Kontraktionsabbildung ist (siehe Aufgaben). Somit konvergiert die Fixpunktiteration $\{x_n\}$ gegen $\sqrt{a}$. Insbesondere lassen sich x_n als Annäherungen von $\sqrt{a}$ betrachten. Nach (6.19) haben wir

$$x_{n+1} = \frac{1}{2}\left(x_n + \frac{a}{x_n}\right),$$

und x_0 kann beliebig in X gewählt werden. Zum Beispiel, sei $a = 2$. Setzen wir $x_0 = 2$ und

$$x_{n+1} = f(x_n) = \frac{1}{2}\left(x_n + \frac{2}{x_n}\right).$$

Dann gilt $x_1 = f(2) = \frac{3}{2}$, $x_2 = f\left(\frac{3}{2}\right) = \frac{17}{12}$, $x_3 = f\left(\frac{17}{12}\right) = \frac{577}{408}$, $x_4 = f\left(\frac{577}{408}\right) = \frac{665\,857}{470\,832}$ und

$$x_5 = f\left(\frac{665\,857}{470\,832}\right) = \frac{886\,731\,088\,897}{627\,013\,566\,048} = 1.41421356237309505...,$$

was schon eine gute Annäherung von $\sqrt{2}$ mit 17 richtigen Nachkommastellen ist.

6.7 Kompakte Mengen

Seien (X, d) ein metrischer Raum space und K eine Teilmenge von X.

Definition. Eine *Überdeckung* von K ist eine Familie $\{U_\alpha\}_{\alpha \in S}$ von Teilmengen von X, die K überdeckt, d.h.

$$K \subset \bigcup_{\alpha \in S} U_\alpha,$$

wobei S eine beliebige Indexmenge ist. Sei T eine Teilmenge von S. Die Familie $\{U_\alpha\}_{\alpha \in T}$ heißt *Teilüberdeckung* von $\{U_\alpha\}_{\alpha \in S}$ falls sie auch K überdeckt. Die Überdeckung $\{U_\alpha\}_{\alpha \in S}$ heißt *offen* falls alle U_α offene Teilmengen von X sind.

In Analysis I haben wir Überdeckung von $[a, b]$ mit offenen Intervallen betrachtet, und haben bewiesen, dass jede solche Überdeckung eine endliche Teilüberdeckung enthält (Überdeckungssatz).

Definition. Eine Menge $K \subset X$ heißt *kompakt* falls jede offene Überdeckung von K eine endliche Teilüberdeckung enthält.

Jede endliche Menge K ist offensichtlich kompakt. Jedes abgeschlossenes beschränktes Intervall $[a, b]$ ist kompakt, was aus dem Überdeckungssatz folgt (später erhalten wir einen unabhängigen Beweis davon). Andererseits, offenes Intervall (a, b) ist nicht kompakt.

Eine von wichtigsten Eigenschaften von Kompaktheit ist die Beziehung zur Stetigkeit.

Satz 6.14 *Seien X und Y zwei metrische Räume spaces und $f : X \to Y$ eine stetige Abbildung. Ist K kompakt in X so ist das Bild $f(K)$ kompakt in Y.*

Somit ist ein stetiges Bild von kompakter Menge wieder kompakt.

Beweis. Sei $\{U_\alpha\}_{\alpha \in S}$ eine offene Überdeckung von $f(K)$, so dass

$$f(K) \subset \bigcup_{\alpha \in S} U_\alpha.$$

Anwendung von Urbildabbildung f^{-1} ergibt

$$K \subset \bigcup_{\alpha \in S} f^{-1}(U_\alpha).$$

Nach dem Satz ist $f^{-1}(U_\alpha)$ eine offene Menge in X und somit ist $\{f^{-1}(U_\alpha)\}_{\alpha \in S}$ eine offene Überdeckung von K in X. Nach der Kompaktheit von K gibt es eine endliche Teilüberdeckung $\{f^{-1}(U_\alpha)\}_{\alpha\in T}$ von K. Dann ist $\{U_\alpha\}_{\alpha\in T}$ eine endliche Teilüberdeckung von $f(K)$, woraus die Kompaktheit von $f(K)$ folgt. ■

Jede Teilmenge $Y \subset X$ vom metrischen Raum X lässt sich als metrischer Raum mit gleicher Metrik d betrachten. Der metrischer Raum (Y,d) heißt *Unterraum* von (X,d). Die Eigenschaften einer Teilmenge $K \subset Y$ können davon abhängen, ob K im metrischen Raum (Y,d) oder (X,d) betrachtet wird. Zum Beispiel, K kann offen in Y sein, aber nicht offen in X (z.B. $K = Y$ ist immer offen in Y, aber nicht unbedingt in X).

Definition. Eine Eigenschaft der Menge $K \subset X$ heißt *intrinsisch* oder *inner* falls sie unabhängig vom enthaltenden metrischen Raum ist.

Wie gesagt, die Eigenschaft von K offen zu sein ist nicht intrinsisch.

Satz 6.15 *Die Kompaktheit von K ist eine intrinsische Eigenschaft.*

Beweis. Sei $K \subset Y \subset X$. Wir müssen beweisen, dass

$$K \text{ ist kompakt in } (Y,d) \Leftrightarrow K \text{ ist kompakt in } (X,d).$$

Sei K kompakt in (Y,d). Sei $\{U_\alpha\}_{\alpha\in S}$ eine offene Überdeckung von K in X. Für jede offene Teilmenge $U \subset X$ ist die Menge $U' = U \cap Y$ offen in Y, da für jedes $x \in U'$ ein $r > 0$ existiert mit $B_X(x,r) \subset U$ und somit mit

$$\begin{aligned} B_Y(x,r) &= \{y \in Y : d(x,y) < r\} \\ &= \{y \in X : d(x,y) < r\} \cap Y \\ &= B_X(x,r) \cap Y \subset U'. \end{aligned}$$

Die Familie $\{U'_\alpha\}_{\alpha\in S}$ ist somit eine offene Überdeckung von K in Y. Nach der Kompaktheit von K in Y gibt es eine endliche Teilüberdeckung $\{U'_\alpha\}_{\alpha\in T}$ von K. Dann ist $\{U_\alpha\}_{a\in T}$ eine endliche Teilüberdeckung von K in X.

Umgekehrt, sei K kompakt in (X,d) und sei $\{V_\alpha\}_{\alpha\in S}$ eine offene Überdeckung von K in (Y,d). Jede offene Menge V in Y ist eine Vereinigung von den Kugeln in Y (Satz 6.4), d.h.

$$V = \bigcup_i B_Y(x_i, r_i)$$

Die Vereinigung von den entsprechenden Kugeln in X ergibt uns eine offene Menge $\widetilde{V}$ in X, d.h.

$$\widetilde{V} = \bigcup_i B_X(x_i, r_i).$$

Es gilt $V = \widetilde{V} \cap Y$, da

$$\begin{aligned} \widetilde{V} \cap Y &= \bigcup_i B_X(x_i, r_i) \cap Y \\ &= \bigcup_i B_Y(x_i, r_i) = V. \end{aligned}$$

Die Familie $\{\widetilde{V}_\alpha\}_{\alpha\in S}$ ist somit eine offene Überdeckung von K in (X,d). Nach der Kompaktheit von K in X gibt es eine endliche Teilüberdeckung $\{\widetilde{V}_\alpha\}_{\alpha\in T}$, woraus folgt, dass $\{V_\alpha\}_{\alpha\in T}$ eine endliche Teilüberdeckung von K in (Y,d) ist. ■

Definition. Eine Teilmenge K vom metrischen Raum (X, d) heißt *folgenkompakt* falls jede Folge in K eine konvergente Teilfolge mit dem Grenzwert in K enthält.

Zum Beispiel, jedes abgeschlossenes beschränktes Intervall $[a, b]$ ist folgenkompakt nach dem Satz von Bolzano-Weierstrass, während offenes Intervall (a, b) nicht folgenkompakt ist.

Definition. Eine Teilmenge K vom metrischen Raum (X, d) heißt *totalbeschränkt* falls es für jedes $\varepsilon > 0$ eine endliche Überdeckung von K mit Kugeln mit Radien ε gibt.

Die letzte Bedingung bedeutet folgendes: für jedes $\varepsilon > 0$ gibt es eine endliche Folge $\{x_1, ..., x_l\}$ von Punkten in X so dass $\{B(x_i, \varepsilon)\}_{i=1}^{l}$ eine Überdeckung von K ist. Jede Folge $\{x_i\}$ mit diesen Eigenschaften heißt *ε-Netz* von K in X. Eine äquivalent Definition von ε-Netz von K: das ist eine endliche Folge $\{x_i\}$ von Elementen von X so dass es für jedes $x \in X$ ein x_i mit $d(x, x_i) < \varepsilon$ gibt. Ein ε-Netz lässt sich als eine endliche Approximation von K mit dem Approximationsfehler $< \varepsilon$ betrachten.

Definition. Eine Teilmenge K von einem metrischen Raum heißt *beschränkt* falls sie in einer metrischen Kugel liegt.

Zum Beispiel, jedes beschränktes Intervall in $\mathbb{R}$ ist eine beschränkte Menge. Bemerken wir die folgenden Eigenschaften von diesen Begriffen.

- Folgenkompaktheit, Totalbeschränktheit und Beschränktheit sind intrinsische Eigenschaften. Für die Folgenkompaktheit ist es offensichtlich nach Definition, für Totalbeschränktheit und Beschränktheit siehe Aufgaben.

- Jede Teilmenge von totalbeschränkter Menge ist auch totalbeschränkt (offensichtlich nach Definition).

- Totalbeschränkte Menge ist beschränkt, aber die Umkehrung davon gilt nicht immer (siehe Aufgabe 81).

Beispiel. Trotzdem jedes beschränktes Intervall I in $\mathbb{R}$ ist totalbeschränkt, da I für jedes $\varepsilon > 0$ eine endliche Menge von Gliedern der Folge $\{\varepsilon k\}_{k \in \mathbb{Z}}$ enthält, die ein ε-Netz ergibt. Daraus folgt, dass jede beschränkte Teilmenge von $\mathbb{R}$ totalbeschränkt ist. Somit in $\mathbb{R}$ sind Beschränktheit und Totalbeschränktheit äquivalent.

Das folgende Lemma ist eine Verallgemeinerung davon für $\mathbb{R}^n$.

Lemma 6.16 *Jede Kugel in $\mathbb{R}^n$ ist totalbeschränkt. Folglich sind in $\mathbb{R}^n$ Beschränktheit und Totalbeschränktheit äquivalent (bezüglich jeder p-Metrik).*

Beweis. Bezeichnen wir die Kugel bezüglich p-Metrik mit $B_p(x, r)$. Es folgt aus (6.15), dass

$$B_p(x, r) \subset B_\infty(x, r),$$

so reicht es zu beweisen, dass $B_\infty(x, r)$ totalbeschränkt ist. Wir beweisen dies per Induktion nach n. Induktionsanfang für $n = 1$ wurde schon im obigen Beispiel gemacht. Da die ∞-Kugel in $\mathbb{R}^n$ ein Produkt von ∞-Kugeln in $\mathbb{R}^{n-1}$ und $\mathbb{R}$ ist, so folgt der Induktionsschritt aus der folgenden Aussage.

Behauptung. *Seien $X \subset \mathbb{R}^m$ und $Y \subset \mathbb{R}^l$ totalbeschränkte Mengen. Dann ist $X \times Y \subset \mathbb{R}^{m+l}$ auch totalbeschränkt.*

Seien $\{x_i\}$ und $\{y_j\}$ die ε-Netze in X bzw Y, d.h. es für jedes $x \in X$ ein x_i mit $d_\infty(x, x_i) < \varepsilon$ gibt und es für jedes $y \in Y$ ein y_j mit $d_\infty(y, y_j) < \varepsilon$ gibt. Dann gilt

$$d_\infty((x, y), (x_i, y_j)) = \max(d_\infty(x, x_i), d_\infty(y, y_j)) < \varepsilon,$$

so dass die Doppelfolge $\{(x_i, y_j)\}$ ein ε-Netz in $X \times Y$ ist. ∎

Hauptsatz 6.17 *Sei (X, d) ein metrischer Raum. Die folgenden drei Bedingungen sind äquivalent.*

(i) X ist kompakt.

(ii) X ist folgenkompakt.

(iii) X ist totalbeschränkt und vollständig.

Beweis. $(i) \Longrightarrow (ii)$ Wir benutzen die folgende Terminologie, wie in Analysis 1. Ein Punkt $x \in X$ heißt *Häufungspunkt* der Folge $\{x_n\}$ falls x der Grenzwert einer Teilfolge von $\{x_n\}$ ist. Ein Punkt $x \in X$ heißt *Verdichtungspunkt* von $\{x_n\}$ falls jede Kugel $B(x, r)$ mit $r > 0$ unendlich viele Glieder der Folge $\{x_n\}$ enthält.

Behauptung. *Ein $x \in X$ ist Häufungspunkt der Folge $\{x_n\}$ genau dann, wenn x Verdichtungspunkt von $\{x_n\}$ ist.*

In der Tat, ist x ein Häufungspunkt, so enthält jeder Kugel $B(x, r)$ fast alle Glieder einer Teilfolge und somit unendlich viele Glieder der Folge $\{x_n\}$. Ist x ein Verdichtungspunkt, so erhält man eine Teilfolge $\{x_{n_k}\}_{k=1}^\infty$ mit $x_{n_k} \in B\left(x, \frac{1}{k}\right)$, die gegen x konvergiert, so dass x ein Häufungspunkt ist.

Sei X kompakt und sei $\{x_n\}$ eine Folge in X. Zeigen wir, dass $\{x_n\}$ eine konvergente Folge enthält, d.h. es einen Verdichtungspunkt von $\{x_n\}$ gibt. Nehmen wir das Gegenteil an, dass jeder Punkt $x \in X$ kein Verdichtungspunkt von $\{x_n\}$ ist. Das bedeutet, dass für jedes $x \in X$ es ein $\varepsilon_x > 0$ gibt so dass die Kugel $B(x, \varepsilon_x)$ nur endlich viele von Glieder der Folge $\{x_n\}$ enthält Die Familie $\{B(x, \varepsilon_x)\}_{x \in X}$ ist eine offene Überdeckung von X, woraus folgt, dass sie eine endliche Teilüberdeckung enthält. Da jede Kugel $B(x, \varepsilon_x)$ nur endlich viele von den Gliedern von der Folge $\{x_n\}$ enthält, so folgt es, dass die Folge $\{x_n\}$ nur endlich viele Glieder enthält, was ein Widerspruch ist.

$(ii) \Longrightarrow (iii)$ Zeigen wir zuerst, dass X vollständig ist. Sei $\{x_n\}$ eine Cauchy-Folge in X. Nach der Folgenkompaktheit von X besitzt $\{x_n\}$ eine konvergente Teilfolge. Daraus folgt, dass auch die ganze Folge $\{x_n\}$ konvergent ist, da die folgende allgemeine Aussage gilt.

Behauptung. *Hat eine Cauchy-Folge $\{x_n\}$ eine konvergente Teilfolge, so ist $\{x_n\}$ auch konvergent.*

Sei $\{x_{n_k}\}_{k=1}^\infty$ konvergente Teilfolge mit dem Grenzwert $a \in X$. Zeigen wir, dass auch $x_n \to a$. Nach Definition von Cauchy-Folge haben wir:

$$\forall \varepsilon > 0 \ \exists N \in \mathbb{N} \ \ \forall n, m \geq N \text{ gilt } d(x_n, x_m) < \varepsilon.$$

Da $x_{n_k} \to a$, so gilt $d(x_{n_k}, a) < \varepsilon$ für fast alle k. Insbesondere gibt es $n_k \geq N$ mit $d(x_{n_k}, a) < \varepsilon$. Für jedes $n \geq N$ haben wir

$$d(x_n, a) \leq d(x_n, x_{n_k}) + d(x_{n_k}, a) \leq 2\varepsilon$$

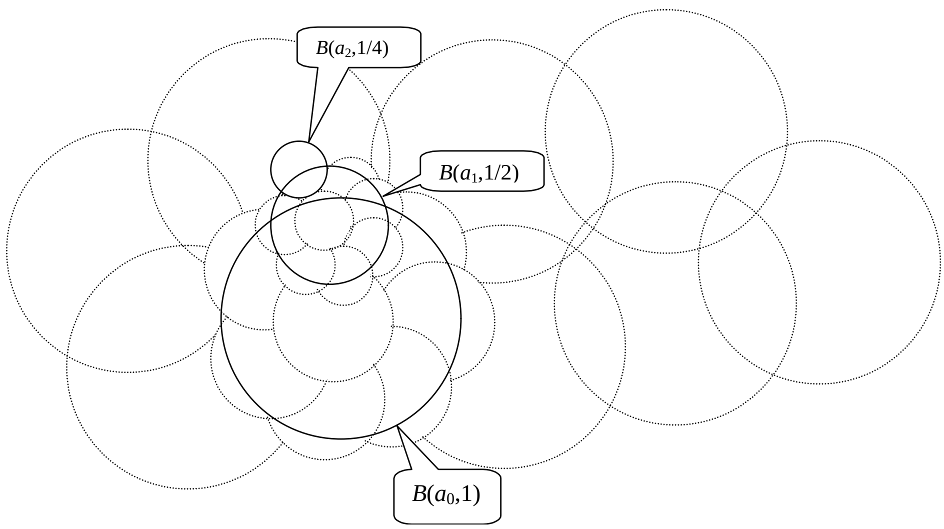
woraus $x_n \to a$ folgt.

Beweisen wir jetzt, dass X totalbeschränkt ist, d.h. es für jedes $\varepsilon > 0$ ein ε-Netz von X gibt. Nehmen wir das Gegenteil an, dass es für ein $\varepsilon > 0$ kein ε-Netz gibt. Definieren wir dann induktiv eine Folge $\{x_n\}_{n=1}^{\infty} \subset X$ wie folgt. Sei $x_1 \in X$ beliebige. Sind $x_1, ..., x_{n-1}$ schon gewählt, so wählen wir x_n wie folgt. Die Kugeln $\{B(x_i, \varepsilon)\}_{i=1}^{n-1}$ überdecken X nicht, da sonst $\{x_i\}_{i=1}^{n-1}$ ein ε-Netz wäre. Somit gibt es einen Punkt in X, der in keiner Kugel $B(x_i, \varepsilon)$, $i = 1, ..., n-1$ liegt, so bezeichnen wir diesen Punkt mit x_n. Nach Konstruktion gilt $d(x_n, x_m) \geq \varepsilon$ für beliebige zwei Indizes $n \neq m$. Daraus folgt, dass $\{x_n\}$ keine Cauchy-Folge ist und darüber hinaus keine Teilfolge von $\{x_n\}$ Cauchy-Folge ist. Somit keine Teilfolge von $\{x_n\}$ konvergiert, was in Widerspruch zur Folgenkompaktheit von X ist.

$(iii) \implies (i)$ Sei $\{U_\alpha\}_{\alpha \in S}$ eine offene Überdeckung von X. Zeigen wir, dass sie eine endliche Teilüberdeckung von X besitzt. Nehmen wir das Gegenteil an: es gibt keine endliche Teilüberdeckung.

Nach Totalbeschränktheit gibt es ein 1-Netz $\{x_1, ..., x_k\}$. Falls jede Kugel $B(x_i, 1)$, $i = 1, ..., k$ eine endliche Teilüberdeckung von $\{U_\alpha\}$ zulässt, so liefert die Vereinigung von allen Teilüberdeckungen von $B(x_i, \varepsilon)$ für $i = 1, ..., k$ eine endliche Teilüberdeckung von X. Somit existiert ein $B(x_i, 1)$ ohne endliche Teilüberdeckung. Bezeichnen wir diese Kugel mit $B(a_0, 1)$ (wobei $a_0 = x_i$).

Es gibt ein $\frac{1}{2}$-Netz $\{y_1, ..., y_l\}$ von X. Betrachten wir nur jene Kugeln $B\left(y_i, \frac{1}{2}\right)$, die mit $B(a_0, 1)$ nicht-leeren Schnitt haben. Wie oberhalb beschließen wir, dass eine von solchen Kugeln $B\left(y_i, \frac{1}{2}\right)$ keine endliche Teilüberdeckung von $\{U_\alpha\}$ zulässt; bezeichnen wir diese Kugel mit $B\left(a_1, \frac{1}{2}\right)$.



Induktiv erhalten wir eine Folge $\{B(a_n, 2^{-n})\}_{n=0}^{\infty}$ von Kugeln mit der Eigenschaften:

(a) jede zwei nacheinander stehende Kugeln haben nicht-leeren Schnitt;

(b) jede Kugel lässt keine endliche Teilüberdeckung von $\{U_\alpha\}$ zu.

Ist $B\left(a_{n-1}, 2^{-(n-1)}\right)$ schon bestimmt, so wählen wir ein ε-Netz $\{z_1, ..., z_m\}$ mit $\varepsilon = 2^{-n}$. Eine von den Kugeln $B(z_i, 2^{-n})$ die mit $B\left(a_{n-1}, 2^{-(n-1)}\right)$ nicht-leeren Schnitt haben, lässt keine endliche Teilüberdeckung von $\{U_\alpha\}$ zu; nennen wir diese Kugel $B(a_n, 2^{-n})$.

Es folgt aus (a) und Lemma 6.2 dass

$$d(a_n, a_{n-1}) < 2^{-(n-1)} + 2^{-n} < 2^{-(n-2)}.$$

Daraus folgt, dass $\{a_n\}$ eine Cauchy-Folge ist (siehe die Behauptung aus dem Beweis von dem Satz 6.13). Nach der Vollständigkeit von X ist die Folge $\{a_n\}$ konvergent, sei $a_n \to a$. Der Punkt a liegt in einer Menge U_α. Da U_α offen ist, so existiert ein $r > 0$ mit $B(a, r) \subset U_\alpha$. Für hinreichend großes n erhalten wir

$$B\left(a_n, 2^{-n}\right) \subset B(a, r), \tag{6.20}$$

da $d(x, a_n) \to 0$ und $2^{-n} \to 0$ für $n \to \infty$ und somit für hinreichend großes n gilt

$$d(a, a_n) + 2^{-n} < r,$$

was (6.20) folgt. Somit erhalten wir $B(a_n, 2^{-n}) \subset U_\alpha$, was im Widerspruch zur Bedingung steht, dass $B(a_n, 2^{-n})$ keine endliche Teilüberdeckung zulässt. ■

Satz 6.18 *Eine Menge $K \in \mathbb{R}^n$ ist kompakt genau dann, wenn K beschränkt und abgeschlossen ist.*

Beweis. Zeigen wir, dass jede kompakte Menge K immer abgeschlossen und beschränkt ist. In der Tat, ist K nicht abgeschlossen, so existiert nach dem Satz 6.4 eine konvergente Folge in K mit dem Grenzwert außer K. Dann hat diese Folge keine konvergente in K Teilfolge, und somit ist K nicht folgenkompakt und nicht kompakt nach dem Satz 6.17. Ist K nicht beschränkt, so ist K nicht totalbeschränkt und somit nicht kompakt nach dem Satz 6.17.

Sei K eine beschränkte abgeschlossene Teilmenge von $\mathbb{R}^n$. Nach Lemma 6.16 ist K totalbeschränkt. Da $\mathbb{R}^n$ nach dem Korollar 6.11 vollständig ist und K abgeschlossen ist, so ist K vollständiger metrischer Raum (Aufgabe 83). Nach dem Satz 6.17 ist K kompakt. ■

Korollar 6.19 (Extremwertsatz) *Seien K eine beschränkte abgeschlossene Teilmenge von $\mathbb{R}^n$ und $f : K \to \mathbb{R}$ eine stetige Funktion. Dann existieren die beiden Werten* $\max_K f := \max f(K)$ *und* $\min_K f := \min f(K)$.

Beweis. Da K nach dem Satz 6.18 kompakt ist, so ist $f(K)$ auch kompakt nach dem Satz 6.14. Somit ist $f(K)$ beschränkte und abgeschlossene Teilmenge von $\mathbb{R}$. Insbesondere sind $\sup_K f$ und $\inf_K f$ endlich. Da $\sup_K f$ und $\inf_K f$ die Grenzwerte von Folgen aus $f(K)$ sind, so liegen sie in $f(K)$ nach der Abgeschlossenheit von $f(K)$. Somit existieren die beiden Werte $\max f(K)$ und $\min f(K)$. ■

Korollar 6.20 *Alle Normen in $\mathbb{R}^n$ sind äquivalent* (siehe Aufgaben).

6.8 Fundamentalsatz der Algebra

Hauptsatz 6.21 (Fundamentalsatz der Algebra) *Jedes Polynom $P(z) = a_0 + a_1 z + ... + a_n z^n$ von Grad $n \geq 1$ mit komplexwertigen Koeffizienten $a_0, a_1, ..., a_n \neq 0$ hat mindestens eine komplexe Nullstelle.*

Beweis. Wir beweisen, dass die reellwertige Funktion $z \mapsto |P(z)|$ eine Minimumstelle in $\mathbb{C}$ besitzt und der Wert des Minimums gleich 0 ist, woraus die Existenz einer Nullstelle folgt. Bemerken wir zunächst, dass

$$|P(z)| \to \infty \text{ für } |z| \to \infty,$$

da

$$P(z) = a_n z^n + a_{n-1} z^{n-1} + ... + a_0$$

ergibt

$$\begin{aligned} |P(z)| &\geq |a_n z^n| - \left|a_{n-1} z^{n-1}\right| - \left|a_{n-2} z^{n-2}\right| - ... - |a_0| \\ &= |z^n| \left(|a_n| - \frac{|a_{n-1}|}{|z|} - \frac{|a_{n-2}|}{|z|^2} - ... \frac{|a_0|}{|z|^n} \right) \\ &\sim |a_n z|^n \text{ für } |z| \to \infty, \end{aligned}$$

und $|a_n z^n| \to \infty$ für $|z| \to \infty$. Wählen wir $R > 0$ so groß, dass $|P(z)| > |a_0|$ für alle $|z| > R$, und betrachten eine abgeschlossene Kugel in $\mathbb{C} = \mathbb{R}^2$:

$$K := \overline{B}(0, R) = \{z \in \mathbb{C} : |z| \leq R\}.$$

Da die Funktion $z \mapsto |P(z)|$ offensichtlich stetig ist, so nimmt die Funktion $|P(z)|$ nach Korollar 6.19 den minimalen Wert in K an einer Stelle $z_0 \in K$ an. Dann gilt

$$|P(z_0)| \leq |P(0)| = |a_0| < |P(z)| \text{ für alle } z \notin K.$$

Somit ist z_0 die Minimumstelle von $|P(z)|$ nicht nur in K sondern auch in $\mathbb{C}$.

Zeigen wir, dass $|P(z_0)| = 0$. Ohne Beschränkung der Allgemeinheit nehmen wir an, dass $z_0 = 0$ (sonst schreiben wir $P(z)$ um als ein Polynom von $(z - z_0)$ und dann $z - z_0$ in z umbenennen). Nehmen wir das Gegenteil an, dass $P(0) \neq 0$, d.h. $a_0 \neq 0$. Ohne Beschränkung der Allgemeinheit nehmen wir an, dass $a_0 = 1$, d.h.

$$P(z) = 1 + a_1 z + ... + a_n z^n.$$

Zeigen wir die Existenz von einem $z \in \mathbb{C}$ mit $|P(z)| < |P(0)| = 1$, was im Widerspruch zur Minimalität von $|P(0)|$ stehen wird. Sei $k \in \{1, ..., n\}$ der minimale Index mit $a_k \neq 0$, so dass

$$P(z) = 1 + a_k z^k + a_{k+1} z^{k+1} + ... + a_n z^n.$$

Jetzt wählen wir $z \in \mathbb{C} \setminus \{0\}$ so dass $a_k z^k$ eine negative reelle Zahl ist, d.h.

$$\arg\left(a_k z^k\right) = \pi. \tag{6.21}$$

Da

$$\arg\left(a_k z^k\right) = \arg a_k + k \arg z,$$

so wird (6.21) erfüllt, vorausgesetzt

$$\arg z = \frac{\pi - \arg a_k}{k}. \tag{6.22}$$

Jetzt wählen wir $|z|$ so klein (aber nicht Null), dass

$$\left|a_k z^k\right| < 1 \quad \text{und} \quad \left|a_{k+1} z^{k+1} + \ldots + a_n z^n\right| < \frac{1}{2}\left|a_k z^k\right|, \tag{6.23}$$

was möglich ist, dass

$$\frac{\left|a_{k+1} z^{k+1} + \ldots + a_n z^n\right|}{\left|a_k z^k\right|} = \frac{\left|a_{k+1} z + \ldots + a_n z^{n-k}\right|}{\left|a_k\right|} \to 0 \text{ für } |z| \to 0.$$

Für z wie in (6.22) und (6.23) gilt

$$\begin{aligned} |P(z)| &\leq \left|1 + a_k z^k\right| + \left|a_{k+1} z^{k+1} + \ldots + a_n z^n\right| \\ &\leq \left|1 + a_k z^k\right| + \frac{1}{2}\left|a_k z^k\right| \qquad (\text{da } a_k z^k \text{ negativ ist}) \\ &= 1 + a_k z^k - \frac{1}{2} a_k z^k \\ &= 1 + \frac{1}{2} a_k z^k \\ &< 1, \end{aligned}$$

was zu beweisen war. ■

Sei P ein Polynom von Grad $n \geq 1$ mit komplexen Koeffizienten und seien $\lambda_1, \ldots, \lambda_l$ are verschiedene komplexe Nullstellen von P. Der Fundamentalsatz der Algebra impliziert die folgende Faktorisierung

$$P(z) = c(z - \lambda_1)^{m_1} \ldots (z - \lambda_l)^{m_l} = c \prod_k (z - \lambda_k)^{m_k}, \tag{6.24}$$

wobei $c \in \mathbb{C} \setminus \{0\}$ und $m_k \in \mathbb{N}$ (siehe Aufgaben). Die Zahl m_k heißt die *Vielfachheit* der Nullstelle λ_k. Es gilt auch

$$m_1 + \ldots + m_l = n.$$

Sei jetzt $P(z)$ ein Polynom mit reellwertigen Koeffizienten. Dann für jede Nullstelle $\lambda \in \mathbb{C} \setminus \mathbb{R}$ mit Vielfachheit m ist auch die komplexe Konjugierte $\overline{\lambda}$ eine Nullstelle von P, und zwar mit gleicher Vielfachheit m. Insbesondere lassen sich die Faktoren $(z - \lambda)^m$ und $\left(z - \overline{\lambda}\right)^m$ in (6.24) gruppieren wie folgt:

$$\left((z - \lambda)\left(z - \overline{\lambda}\right)\right)^m = \left(z^2 + pz + q\right)^m,$$

wobei $p = -\left(\lambda + \overline{\lambda}\right)$ und $q = \lambda\overline{\lambda}$ reel sind. Somit erhalten wir aus (6.24) die folgende reellwertige Faktorisierung

$$P(z) = c \prod_i (z - r_i)^{k_i} \prod_j \left(z^2 + p_j z + q_j\right)^{m_j},$$

wobei r_i alle reelle Nullstellen von P sind und jeder Faktor $z^2 + p_j z + q_j$ zum Paar $\lambda, \overline{\lambda}$ von konjugierten komplexwertigen Nullstellen von P entspricht. Diese Faktorisierung wurde im Abschnitt 4.5 benutzt.

6.9 Zusammenhang

Definition. Eine Teilmenge K von einem metrischen Raum X heißt *zusammenhängend* falls für jede Überdeckung $K \subset U \sqcup V$ von K mit zwei disjunkten offenen Mengen U, V gilt $K \subset U$ oder $K \subset V$.

Die Eigenschaft von Zusammenhang ist eine intrinsische Eigenschaft (siehe Aufgaben).

Satz 6.22 *Seien X und Y zwei metrische Räume und $f : X \to Y$ eine stetige Abbildung. Ist K eine zusammenhängende Teilmenge von X so ist $f(K)$ auch zusammenhängend.*

Beweis. Sei $f(K) \subset U \sqcup V$ eine Überdeckung von $f(K)$ mit disjunkten offenen Mengen. Daraus folgt

$$K \subset f^{-1}(U) \sqcup f^{-1}(V).$$

Da $f^{-1}(U)$ und $f^{-1}(V)$ offen sind, so ergibt der Zusammenhang von K, dass $K \subset f^{-1}(U)$ oder $K \subset f^{-1}(V)$, woraus $f(K) \subset U$ bzw $f(K) \subset V$ folgt. ■

Satz 6.23 *Jedes Intervall $I \subset \mathbb{R}$ ist zusammenhängend. Umgekehrt, jede zusammenhängende Teilmenge von $\mathbb{R}$ ist ein Intervall.*

Beweis. Zeigen wir, dass beliebiges Intervall I zusammenhängend ist. Sei $I \subset U \sqcup V$ eine Überdeckung von I mit offenen Mengen. Definieren wir eine Funktion $f : I \to \mathbb{R}$ wie folgt:

$$f(x) = \begin{cases} 1, & x \in U \cap I, \\ 0, & x \in V \cap I. \end{cases}$$

Beweisen wir, dass f stetig auf I ist. Für jedes $x \in U \cap I$ gibt es ein $\varepsilon > 0$ mit

$$(x - \varepsilon, x + \varepsilon) \subset U,$$

woraus folgt, dass $f \equiv 1$ in $(x - \varepsilon, x + \varepsilon) \cap I$ und somit die Stetigkeit von f an x. Analog ist f stetig an alle Stellen $x \in V \cap I$ und somit auf I.

Sind $U \cap I$ und $V \cap I$ nicht leer, so nimmt f die Werte 0 und 1 an und somit nach dem Zwischenwertsatz soll f auch alle Werte in $[0, 1]$ annehmen, was nicht der Fall ist. Somit soll $U \cap I$ oder $V \cap I$ leer sein, woraus $I \subset V$ oder $I \subset U$ folgt.

Sei K eine zusammenhängende Teilmenge von $\mathbb{R}$. Für $a, b \in K$ soll auch ganzes Intervall $[a, b]$ in K liegen, da sonst es einen Punkt $c \in (a, b) \setminus K$ gibt und wir eine Überdeckung erhalten

$$K \subset (-\infty, c) \sqcup (c, +\infty)$$

mit $K \not\subset (-\infty, c)$ und $K \not\subset (c, \infty)$, was im Widerspruch mit Zusammenhang von K steht. Somit ist K ein Intervall mit den Grenzen $\inf K$ und $\sup K$. ■

Als Folgerung aus den Sätzen 6.22 und 6.23 erhalten wir, dass das Bild der stetigen Abbildung $f : I \to \mathbb{R}$ wieder ein Intervall ist, was in Analysis 1 schon bewiesen wurde.

Für jede zwei Punkte $x, y \in \mathbb{R}^n$ bezeichnen wir mit $[x, y]$ die Menge

$$[x, y] := \{(1-\lambda)\, x + \lambda y : \lambda \in [0, 1]\}\,.$$

Die Menge $[x, y]$ ist die gerade Strecke zwischen x und y.

Definition. Eine Menge $K \subset \mathbb{R}^n$ heißt *Sterngebiet*, falls es einen Punkt $a \in K$ gibt mit

$$x \in K \Rightarrow [a, x] \in K.$$

Der Punkt a heißt ein *Sternzentrum.*

Beispiel. Zeigen wir, dass jede offene oder abgeschlossene Kugel in $\mathbb{R}^n$ bezüglich einer Norm immer ein Sterngebiet ist. Sei K eine (offene oder abgeschlossene) Kugel mit Zentrum a und Radius r, d.h.

$$K = B\,(a, r) = \{x \in \mathbb{R}^n : \|x - a\| < r\} \;\text{ oder }\; K = \overline{B}\,(a, r) = \{x \in \mathbb{R}^n : \|x - a\| \le r\}\,.$$

Zeigen wir, dass a ein Sternzentrum von K ist, d.h. für jedes $x \in K$ gilt $[a, x] \subset K$. Jeder Punkt $y \in [a, x]$ lässt sich wie folgt darstellen:

$$y = (1-\lambda)\, a + \lambda x \;\text{ für ein } \lambda \in [0, 1]\,,$$

woraus folgt, dass

$$\|y - a\| = \|\lambda x - \lambda a\| = |\lambda|\, \|x - a\| \le \|x - a\|$$

und somit $y \in K$ und $[a, x] \subset K$.

Satz 6.24 *Jedes Sterngebiet K in $\mathbb{R}^n$ ist zusammenhängend.*

Beweis. Sei a ein Sternzentrum von K, so dass für jedes $x \in K$ gilt $[a, x] \subset K$. Sei $K \subset U \sqcup V$ eine Überdeckung von K mit offenen Mengen. Nehmen wir an, dass $a \in U$. Es folgt, dass U, V auch eine Überdeckung von $[a, x]$ für jedes $x \in K$ ist, d.h.

$$[a, x] \subset U \sqcup V.$$

Die Strecke $[a, x]$ ist das Bild von $[0, 1]$ unter der Abbildung

$$\begin{aligned} \varphi &: [0, 1] \to \mathbb{R}^n \\ \varphi(\lambda) &= (1-\lambda)\, a + \lambda x. \end{aligned}$$

Da φ offensichtlich stetig ist, so ist $[a, x] = \varphi([0, 1])$ zusammenhängend nach dem Satz 6.22. Somit soll $[a, x]$ in einer von U, V enthalten. Da $a \in U$, es folgt, dass $[a, x] \subset U$, woraus folgt $x \in U$ und $K \subset U$. ■

Beispiel. Bestimmen wir das Bild der Funktion

$$f(x, y, z) = \frac{x^2 + y^2 + z^2}{(x + y + z)^2}$$

auf der Menge

$$K = \left\{(x, y, z) \in \mathbb{R}^3 : x, y, z > 0\right\}.$$

Die Funktion $f : K \to \mathbb{R}$ ist offensichtlich stetig. Da K Sterngebiet ist und somit zusammenhängend, so ist das Bild $f(K)$ zusammenhängende Teilmenge von $\mathbb{R}$ und somit ein Intervall. Bestimmen wir die Grenzen des Intervalls, d.h. $\sup f$ und $\inf f$, und ob diese dem Intervall gehören. Wir haben

$$x^2 + y^2 + z^2 < x^2 + y^2 + z^2 + 2xy + 2xz + 2yz = (x + y + z)^2$$

so dass $f(x, y, z) < 1$ und $\sup f \leq 1$. In der Tat gilt $\sup f = 1$ da für $y \to 0$ und $z \to 0$ erhalten wir $f(x, y, z) \to 1$. Insbesondere liegt $\sup f$ nicht in $f(K)$.

Andererseits, nach der Ungleichung zwischen arithmetischen und quadratischen Mittelwerten gilt[1]

$$x^2 + y^2 + z^2 \geq \frac{1}{3}(x + y + z)^2, \tag{6.25}$$

woraus folgt

$$f(x, y, z) \geq \frac{1}{3}$$

und somit $\inf f \geq \frac{1}{3}$. Da $f(1, 1, 1) = \frac{1}{3}$, so erhalten wir $\inf f = \frac{1}{3} \in f(K)$. Es folgt, dass

$$f(K) = [\frac{1}{3}, 1).$$

[1] Es gilt die Identität

$$3\left(x^2 + y^2 + z^3\right) = (x + y + z)^2 + (x - z)^2 + (y - z)^2 + (x - y)^2,$$

woraus (6.25) folgt.

Chapter 7

Differentialrechnung in $\mathbb{R}^n$

7.1 Partielle und totale Differenzierbarkeit

Seien Ω eine offene Teilmenge von $\mathbb{R}^n$ und $f : \Omega \to \mathbb{R}^m$ eine Funktion auf Ω mit den Werten in $\mathbb{R}^m$. Wir benutzen die folgende Notation für die Komponenten von x und $f(x)$

$$x = (x_1, ..., x_n)$$

und

$$f(x) = (f_1(x), ..., f_m(x)),$$

wobei jedes $f_k(x)$ eine reellwertige Funktion auf Ω ist. Man schreibt $f_k(x)$ auch in der Form $f_k(x_1, ..., x_n)$ so dass f_k sich als eine reellwertige Funktion von n reellen Variablen betrachten lässt.

Fixieren wir ein $j = 1, ..., n$ und ein $k = 1, ..., m$ und betrachten die Funktion

$$x_j \mapsto f_k(x_1, ..., x_j, ...x_n),$$

wobei alle x_i mit $i \neq j$ als Konstanten betrachtet werden. Ist diese Funktion differenzierbar, so betrachten wir ihre Ableitung.

Definition. Die Ableitung von f_k bezüglich x_j heißt *partielle Ableitung* 1-er Ordnung von f und wird mit $\frac{\partial f_k}{\partial x_j}$ bezeichnet, d.h.

$$\frac{\partial f_k}{\partial x_j}(x) := \lim_{h\to 0} \frac{f_k(x_1, ...x_j + h, ...x_n) - f(x_1, ..., x_j, ..., x_n)}{h}.$$

In dieser Notation wird ein rundes ∂ statt eines geraden d benutzt. Die Ableitung heißt partiell da es nur eine Variable x_j von n Variablen benutzt wird.

Es gibt noch andere Notation für partielle Ableitungen wie folgt:

$$\frac{\partial f_k}{\partial x_j} = \partial_j f_k = (f_k)_{x_j}.$$

Definition. Existieren die partielle Ableitungen $\partial_j f_k(x)$ für alle k und j, so heißt die Funktion f *partiell differenzierbar* in x. In diesem Fall lässt sich die Menge von allen partiellen Ableitung von f in einer $m \times n$ Matrix anordnen wie folgt:

$$J_f := \left(\frac{\partial f_k}{\partial x_j}\right) = (\partial_j f_k) = \begin{pmatrix} \partial_1 f_1 & \partial_2 f_1 & \dots & \partial_n f_1 \\ \partial_1 f_2 & \partial_2 f_2 & \dots & \partial_n f_2 \\ \dots & \dots & \dots & \dots \\ \partial_1 f_m & \partial_2 f_m & \dots & \partial_n f_m \end{pmatrix}, \tag{7.1}$$

wobei $k = 1, ..., m$ ein Zeilenindex ist und $j = 1, ..., n$ ein Spaltenindex. Die Matrix $J_f = J_f(x)$ heißt die *Jacobi-Matrix von* f.

Beispiel. Betrachten wir eine Funktion $f(x) : \mathbb{R}^2 \to \mathbb{R}^1$. In diesem Fall ist J_f eine 1×2 Matrix, d.h. die Zeile

$$J_f = (\partial_1 f, \partial_2 f).$$

Setzen wir

$$f(x, y) = \begin{cases} \frac{xy}{x^2+y^2}, & x \neq 0, \\ 0, & x = 0. \end{cases} \tag{7.2}$$

In jedem Punkt $(x, y) \neq (0, 0)$ ist diese Funktion differenzierbar bezüglich x und y, und es gilt

$$\partial_1 f = \frac{\partial f}{\partial x} = \frac{y(x^2+y^2) - xy2x}{(x^2+y^2)^2} = \frac{y^3 - yx^2}{(x^2+y^2)^2}$$

und analog

$$\partial_2 f = \frac{\partial f}{\partial y} = \frac{x^3 - xy^2}{(x^2+y^2)^2}.$$

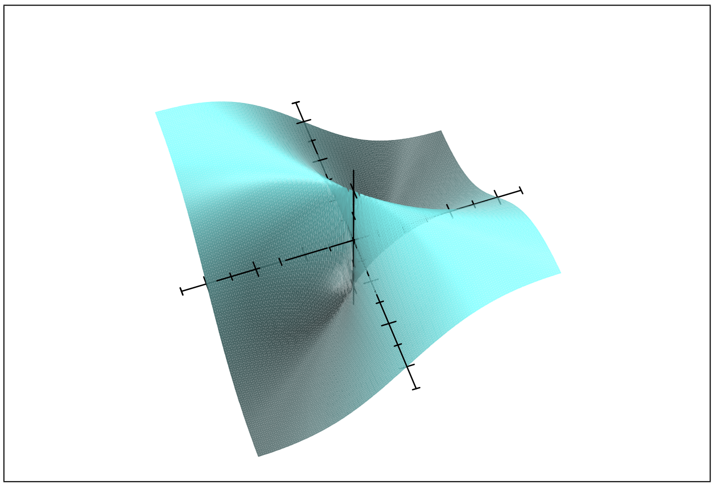
Zeigen wir, dass f auch in $(0, 0)$ partiell differenzierbar. In der Tat haben wir nach Definition

$$\frac{\partial f}{\partial x}(0, 0) = \lim_{h \to 0} \frac{f(h, 0) - f(0, 0)}{h} = 0$$

da $f(h, 0) = 0$, und analog $\frac{\partial f}{\partial y}(0, 0) = 0$. Somit ist f partiell differenzierbar in allen Punkten von $\mathbb{R}^2$.

Allerdings ist die Funktion f nicht stetig im Punkt $(0, 0)$, da für jedes $t \neq 0$ gilt $f(t, t) = \frac{1}{2}$ und somit

$$\lim_{t \to 0} f(t, t) = \frac{1}{2} \neq 0 = f(0, 0).$$



Funktion $f(x, y) = \frac{xy}{x^2+y^2}$

Erinnern wir uns daran, dass die Differenzierbarkeit von Funktion von einer Variable die Stetigkeit ergibt. Diese Eigenschaft ist sehr gewünscht, und um sie zu bekommen soll der Begriff von partieller Differenzierbarkeit verstärkt werden.

Definition. Eine Funktion $f : \Omega \to \mathbb{R}^m$ heißt *total differenzierbar* in $x \in \Omega$ falls es eine lineare Abbildung $A : \mathbb{R}^n \to \mathbb{R}^m$ gibt mit

$$f(x+h) - f(x) = Ah + o(h) \text{ für } h \to 0. \tag{7.3}$$

Fixieren wir eine Norm $\|\cdot\|$ in $\mathbb{R}^n$ und bezeichnen mit $B(x,r)$ die metrische Kugel bezüglich der induzierten Metrik $\|x-y\|$. Da die Menge Ω offen ist, so gibt es eine Kugel $B(x,r)$ mit $r>0$ die in Ω liegt. Somit für alle $h \in \mathbb{R}^n$ mit $\|h\| < r$ liegt $x+h$ in Ω und $f(x+h)$ wohldefiniert ist.

Die Abbildung $A : \mathbb{R}^n \to \mathbb{R}^m$ heißt linear falls sie mit den linearen Operationen in $\mathbb{R}^n$ vertauschbar ist, d.h.

$$\begin{aligned} A(u+v) &= A(u) + A(v) \\ A(\lambda u) &= \lambda A(u) \end{aligned}$$

für alle $u, v \in \mathbb{R}^n$ und $\lambda \in \mathbb{R}$. Für lineare Abbildungen ist es üblich zu schreiben Au statt $A(u)$, was wir in (7.3) benutzt haben.

Das Landau-Symbol $o(h)$ bezeichnet in (7.3) eine Funktion $\varphi(h)$ mit den Werten in $\mathbb{R}^m$ und mit

$$\frac{\|\varphi(h)\|}{\|h\|} \to 0 \text{ für } h \to 0, \tag{7.4}$$

wobei im Zähler eine Norm in $\mathbb{R}^m$ steht. Da alle Normen in $\mathbb{R}^n$ (und in $\mathbb{R}^m$) äquivalent sind, so ist die Richtigkeit von (7.4) (und (7.3)) unabhängig von der Wahl von den Normen in $\mathbb{R}^n$ bzw $\mathbb{R}^m$.

Die Identität (7.3) bedeutet, dass für eine lineare Abbildung A gilt

$$\frac{\|f(x+h) - f(x) - Ah\|}{\|h\|} \to 0 \text{ für } h \to 0.$$

Man kann auch sagen, dass die Differenz $f(x+h) - f(x)$ sich als die Summe von zwei Gliedern darstellen lässt: der führende lineare Glied Ah und der Rest $o(h)$, der als Approximationsfehler für die Approximation von $f(x+h) - f(x)$ mit Ah betrachtet werden kann.

Definition. Die Variable h in (7.3) heißt das *Differential* von x und wird auch mit dx bezeichnet (so dass $dx \in \mathbb{R}^n$ eine unabhängige Variable ist). Die Function $h \mapsto Ah$ heißt das *Differential* der Funktion f in x und wird auch mit $df(x)$ bezeichnet, so that $df = Adx$. Die lineare Abbildung A heißt die *totale Ableitung* von f in x und wird mit $\frac{df}{dx}(x)$ oder $f'(x)$ bezeichnet.

Mit dieser Notation lässt (7.3) sich wie folgt umschreiben:

$$f(x+h) - f(x) = f'(x)h + o(h),$$

was mit der entsprechenden Eigenschaft von Ableitung in Analysis 1 übereinstimmt.

Beispiel. Sei $f : \mathbb{R}^n \to \mathbb{R}^m$ eine lineare Abbildung, d.h. $f(x) = Ax$. Dann gilt

$$f(x+h) - f(x) = Ah,$$

woraus folgt, dass f in jedem Punkt x differenzierbar ist und $f'(x) = A$.

Es ist aus Linearer Algebra bekannt, dass jede lineare Abbildung $A : \mathbb{R}^n \to \mathbb{R}^m$ sich als eine $m \times n$ Matrix $\mathcal{A}$ darstellen lässt, nämlich $Ah = \mathcal{A}h$ wobei der Vektor $h \in \mathbb{R}^n$ als ein Spaltenvektor betrachtet wird. Somit ist $\mathcal{A}h$ das Produkt von $m \times n$ Matrix $\mathcal{A}$ und $n \times 1$ Matrix h, was eine $m \times 1$ Matrix ergibt, d.h. einen Spaltenvektor in $\mathbb{R}^m$, wie erwartet. Normalerweise identifizieren wir die Abbildung A mit der Matrix $\mathcal{A}$ und bezeichnen mit A sowohl die Abbildung als auch ihre Matrix.

Insbesondere lässt sich die totale Ableitung $f'(x)$ als eine $m \times n$ Matrix betrachten.

Satz 7.1 *Ist die Funktion $f : \Omega \to \mathbb{R}^m$ total differenzierbar im Punkt $x \in \Omega$, so gilt folgendes.*
(a) f ist stetig in x.
(b) f ist partiell differenzierbar in x und es gilt

$$f'(x) = J_f(x). \tag{7.5}$$

Bemerken wir, dass $f'(x)$ und $J_f(x)$ die $m \times n$ Matrizen sind. Es folgt aus (7.5), dass die totale Ableitung $f'(x)$ eindeutig bestimmt ist, falls sie existiert.

Wir betonen, dass die totale Differenzierbarkeit aus der partiellen Differenzierbarkeit nicht folgt, was das Beispiel der Funktion (7.2) zeigt.

Beweis. (a) Wir benutzen den Begriff von *Operatornorm* der linearen Abbildung $A : \mathbb{R}^n \to \mathbb{R}^m$:

$$\|A\| := \sup_{h \in \mathbb{R}^n \setminus \{0\}} \frac{\|Ah\|}{\|h\|}. \tag{7.6}$$

Man kann zeigen, dass immer $\|A\| < \infty$ (siehe Aufgaben). Es folgt aus (7.6), dass

$$\|Ah\| \leq \|A\| \, \|h\|.$$

Da

$$f(x+h) - f(x) = Ah + o(h) \text{ für } h \to 0$$

so erhalten wir

$$\|f(x+h) - f(x)\| \leq \|Ah\| + \|o(h)\| \leq \|A\| \, \|h\| + \|o(h)\| \to 0$$

für $h \to 0$, woraus die Stetigkeit von f in x folgt.

(b) Sei $f'(x) = A = (a_{kj})$ wobei k der Zeilenindex ist und j der Spaltenindex. In der Identität

$$f(x+h) - f(x) = Ah + o(h) \tag{7.7}$$

wählen wir $h = (0, ..., h_j, ...0)$ d.h. alle Komponenten von h außer h_j gleich 0 sind. Dann gilt

$$Ah = \begin{pmatrix} a_{11} & ... & a_{1j} & ... & a_{1n} \\ ... & ... & ... & ... & ... \\ a_{k1} & ... & a_{kj} & ... & a_{kn} \\ ... & ... & ... & ... & ... \\ a_{m1} & ... & a_{mj} & ... & a_{mn} \end{pmatrix} \begin{pmatrix} 0 \\ ... \\ h_j \\ ... \\ 0 \end{pmatrix} = \begin{pmatrix} a_{1j}h_j \\ ... \\ a_{kj}h_j \\ ... \\ a_{mj}h_j \end{pmatrix}.$$

Die Identität von k-en Komponenten in (7.7) ergibt

$$f_k(x+h) - f_k(x) = a_{kj}h_j + o(h_j),$$

d.h.

$$f_k(x_1, ..., x_j + h_j, ..., x_n) - f_k(x_1, ..., x_n) = a_{kj}h_j + o(h_j)$$

woraus folgt, dass

$$\partial_j f_k(x) = a_{kj}.$$

Da die Einträge von $J_f(x)$ gleich $\partial_j f_k(x)$ sind, so erhalten wir die Identität (7.5). ■

Unter einer zusätzlichen Bedingung folgt die totale Differenzierbarkeit aus der partiellen Differenzierbarkeit wie im nächsten Satz.

Satz 7.2 *Sei $f : \Omega \to \mathbb{R}^m$ partiell differenzierbar in allen Punkten von Ω. Sind alle Ableitungen $\partial_j f_k$ stetig in einem Punkt $x \in \Omega$ so ist f total differenzierbar in x.*

Beweis. Betrachten wir erst den Fall $m = 1$, so dass f reellwertig ist. Ohne Beschränkung der Allgemeinheit nehmen wir an, dass $x = 0$.

Die Jacobi-Matrix J_f ist eine $1 \times n$ Matrix, d.h. ein Zeilenvektor

$$J_f = (\partial_1 f, ..., \partial_n f).$$

Die totale Differenzierbarkeit von f in 0 wird folgen, falls wir bewiesen, dass

$$f(h) - f(0) = J_f(0)\,h + o(h) \text{ für } h \to 0.$$

Wählen wir die Norm in $\mathbb{R}^n$ als die ∞-Norm. Sei $r > 0$ so klein, dass $B(0,r) \subset \Omega$. Dann nehmen wir an, dass $\|h\| < r$. Gegeben sei ein $h \in \mathbb{R}^n$ mit $\|h\| < r$, betrachten wir eine Folge $\{a_k\}_{k=0}^n$ von Punkten in $\mathbb{R}^n$ wie folgt:

$$a_k = (h_1, h_2, ..., h_k, 0, ..., 0).$$

Insbesondere haben wir

$$\begin{aligned} a_0 &= 0 \\ a_1 &= (h_1, 0, ..., 0) \\ &\;\;... \\ a_n &= (h_1, ..., h_n) = h. \end{aligned}$$

Es gilt $\|a_k\| \leq \|h\| < r$ so dass alle Punkte a_k in Ω liegen. Dann haben wir

$$f(h) - f(0) = f(a_n) - f(a_0) = \sum_{k=1}^{n} (f(a_k) - f(a_{k-1})). \tag{7.8}$$

Um die Differenz $f(a_k) - f(a_{k-1})$ abzuschätzen, betrachten wir die Funktion

$$g(t) = f(h_1, ..., h_{k-1}, t, 0, ..., 0),$$

so dass

$$g(0) = f(a_{k-1}) \quad \text{und} \quad g(h_k) = f(a_k).$$

Die Funktion $g(t)$ ist für alle $t \in [0, h_k]$ wohldefiniert und ist in t differenzierbar, da

$$g'(t) = \partial_k f(h_1, ..., h_{k-1}, t, 0, ..., 0).$$

Der Mittelwertsatz ergibt für ein $\xi \in [0, h_k]$, dass

$$f(a_k) - f(a_{k-1}) = g(h_k) - g(0) = g'(\xi) h_k = \partial_k f(b_k) h_k \tag{7.9}$$

wobei

$$b_k = (h_1, ..., h_{k-1}, \xi, 0, ..., 0).$$

Insbesondere gilt $\|b_k\| \le \|h\|$ und somit $b_k \to 0$ für $h \to 0$. Es folgt aus (7.8) und (7.9), dass

$$\begin{aligned} f(h) - f(0) &= \sum_{k=1}^n \partial_k f(b_k) h_k \\ &= \sum_{k=1}^n \partial_k f(0) h_k + \sum_{k=1}^n (\partial_k f(b_k) - \partial_k f(0)) h_k \\ &= J_f(0) h + o(h), \end{aligned}$$

da

$$\frac{\left\|\sum_{k=1}^n (\partial_k f(b_k) - \partial_k f(0)) h_k\right\|}{\|h\|} \le \sum_{k=1}^n |\partial_k f(b_k) - \partial_k f(0)| \to 0$$

für $h \to 0$ weil $b_k \to 0$ und $\partial_k f$ stetig in 0 ist.

Betrachten wir jetzt den Fall von beliebigen m. Da jede Komponente f_k eine Funktion von Ω nach $\mathbb{R}$ ist, so ergibt der vorige Fall die totale Differenzierbarkeit von f_k in x, d.h.

$$f_k(x+h) - f_k(x) = A_k h + o(h) \text{ für } h \to 0, \tag{7.10}$$

wobei A_k eine $1 \times n$ Matrix ist, d.h. ein Zeilenvektor. Sei A die $m \times n$ Matrix mit k-ter Zeile A_k für $k = 1, ..., m$. Es folgt aus (7.10), dass

$$f(x+h) - f(x) = Ah + o(h) \quad \text{für } h \to 0,$$

woraus die totale Differenzierbarkeit von f in x folgt. ■

Definition. Die Funktion $f : \Omega \to \mathbb{R}^m$ heißt *stetig partiell differenzierbar in* Ω, falls sie in allen Punkten $x \in \Omega$ partiell differenzierbar und alle partielle Ableitungen $\partial_j f_k$ stetig in Ω sind.

Korollar 7.3 *Jede stetig partiell differenzierbare in Ω Funktion ist total differenzierbar in Ω.*

Beweis. Folgt aus dem Satz 4.7. ■

Bemerkung. Häufig benutzt man die folgende einfachere Terminologie:

$$\begin{aligned} \text{Differenzierbarkeit} &\equiv \text{totale Differenzierbarkeit} \\ \text{stetige Differenzierbarkeit} &\equiv \text{stetige partielle Differenzierbarkeit.} \end{aligned}$$

Die stetige Differenzierbarkeit ergibt somit die Differenzierbarkeit.

7.2 Rechenregeln für totale Ableitung

7.2.1 Linearität

Satz 7.4 (Linearität) *Sind die Funktionen $f, g : \Omega \to \mathbb{R}^n$ total differenzierbar in einem Punkt $x \in \Omega$, so ist auch ihre lineare Kombination $af + bg$ mit $a, b \in \mathbb{R}$ total differenzierbar in x und es gilt*

$$(af + bg)'(x) = af'(x) + bg'(x).$$

Beweis. Nach Definition gilt für $h \to 0$

$$f(x+h) - f(x) = f'(x)h + o(h)$$

und

$$g(x+h) - g(x) = g'(x)h + o(h).$$

Es folgt, dass die Funktion $F = af + bg$ erfüllt

$$F(x+h) - F(x) = (af'(x) + bg'(x))h + o(h)$$

woraus $F'(x) = af'(x) + bg'(x)$ folgt. ■

7.2.2 Kettenregel

Satz 7.5 (Kettenregel) *Seien $U \subset \mathbb{R}^n$ und $V \subset \mathbb{R}^m$ offene Mengen. Sei eine Funktion $g : U \to V$ differenzierbar in einem Punkt $x \in U$ und $f : V \to \mathbb{R}^l$ differenzierbar im Punkt $g(x) \in V$. Dann ist die Komposition $f \circ g : U \to \mathbb{R}^l$ differenzierbar in x und es gilt*

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

Bemerken wir, dass

$$\mathbb{R}^n \overset{g'(x)}{\longrightarrow} \mathbb{R}^m \overset{f'(g(x))}{\longrightarrow} \mathbb{R}^l.$$

Somit ist das Produkt (=Komposition) $f'(g(x))g'(x)$ von den linearen Abbildung wohldefiniert und ist eine lineare Abbildung von $\mathbb{R}^n$ nach $\mathbb{R}^l$, genauso, wie die totale Ableitung $(f \circ g)'(x)$.

Beweis. Wir haben

$$g(x+a) - g(x) = g'(x)a + \varphi(a),$$

wobei $\varphi(a) = o(a)$ für $a \to 0$. Setzen wir $v = g(x)$ so dass

$$f(v+b) - f(v) = f'(v)b + \psi(b),$$

wobei $\psi(b) = o(b)$ für $b \to 0$. Wählen wir b so dass $v + b = g(x+a)$, d.h.

$$b = g(x+a) - g(x) = g'(x)a + \varphi(a). \tag{7.11}$$

Es folgt, dass

$$\begin{aligned} f \circ g\left(x+a\right) - f \circ g\left(x\right) &= f\left(v+b\right) - f\left(v\right) \\ &= f'\left(v\right) b + \psi\left(b\right) \\ &= f'\left(v\right) g'\left(x\right) a + f'\left(v\right) \varphi\left(a\right) + \psi\left(b\right). \end{aligned} \tag{7.12}$$

Es bleibt zu zeigen, dass

$$f'\left(v\right) \varphi\left(a\right) + \psi\left(b\right) = o\left(a\right) \text{ für } a \to 0. \tag{7.13}$$

Es gilt

$$\left\| f'\left(v\right) \varphi\left(a\right) \right\| \leq \left\| f'\left(v\right) \right\| \left\| \varphi\left(a\right) \right\| = o\left(\|a\|\right)$$

so dass

$$f'\left(v\right) \varphi\left(a\right) = o\left(a\right) \text{ für } a \to 0.$$

Es folgt aus (7.11), dass

$$\|b\| \leq \left\| g'\left(x\right) a \right\| + \left\| \varphi\left(a\right) \right\| \leq \left\| g'\left(x\right) \right\| \|a\| + \left\| \varphi\left(a\right) \right\| = O\left(\|a\|\right),$$

und somit

$$\left\| \psi\left(b\right) \right\| = o\left(\|b\|\right) = o\left(O\left\|a\right\|\right) = o\left(\|a\|\right)$$

woraus $\psi\left(b\right) = o\left(a\right)$ folgt. ■

Korollar 7.6 *Unter der Bedingungen des Satzes* 7.5 *gilt*

$$\frac{\partial \left(f \circ g\right)_k}{\partial x_j}\left(x\right) = \sum_{i=1}^{m} \frac{\partial f_k}{\partial v_i}\left(g\left(x\right)\right) \frac{\partial g_i}{\partial x_j}\left(x\right), \tag{7.14}$$

wobei v das Argument von f bezeichnet.

Man schreibt auch

$$\partial_j \left(f \circ g\right)_k \left(x\right) = \sum_{i=1}^{m} \partial_i f_k \left(g\left(x\right)\right) \partial_j g_i \left(x\right).$$

Beweis. Die Ableitung $\frac{\partial(f\circ g)_k}{\partial x_j}$ ist der (k,j)-Eintrag von der Jacobi-Matrix und somit auch von $\left(f \circ g\right)'$ (nach dem Satz 7.1). Da

$$\left(f \circ g\right)'\left(x\right) = f'\left(v\right) g'\left(x\right)$$

wobei $v = g\left(x\right)$, so gilt für den (k,j)-Eintrag nach der Regel von Matrizenmultiplikation

$$\begin{aligned} \frac{\partial \left(f \circ g\right)_k}{\partial x_j} &= \left(f'\left(v\right) g'\left(x\right)\right)_{kj} = \sum_{i=1}^{m} f'\left(v\right)_{ki} g'\left(x\right)_{ij} \\ &= \sum_{i=1}^{m} \frac{\partial f_k}{\partial v_i}\left(v\right) \frac{\partial g_i}{\partial x_j}\left(x\right), \end{aligned}$$

was zu beweisen war. ■

Beispiel. Betrachten wir den Fall $n = m = 2$ und $l = 1$. Bezeichnen wir die Koordinaten in $U \subset \mathbb{R}^2$ mit (x, y) und die Koordinaten in $V \subset \mathbb{R}^2$ mit (u, v) so dass $g = g(x, y)$ und $f = f(u, v)$. Bezeichnen wir mit u, v auch die Komponenten von g, d.h.

$$g(x, y) = (u(x, y), v(x, y)).$$

Dann ergibt (7.14)

$$\partial_x (f(u(x, y), v(x, y))) = (\partial_u f)(u(x, y), v(x, y)) \partial_x u(x, y) + (\partial_v f)(u(x, y), v(x, y)) \partial_x v(x, y)$$

und die ähnliche Identität für $\partial_y f$. Kurz schreibt man:

$$\partial_x f = \partial_u f \ \partial_x u + \partial_v f \ \partial_x v.$$

So, die Regel ist wie folgt: um die Komposition $f(u(x, y), v(x, y))$ in x abzuleiten, leitet man f bezüglich u bzw v ab, multipliziert die Ableitung mit $\partial_x u$ bzw $\partial_x v$, und danach addiert die zwei Produkte.

Z.B. sei $f(u, v) = u^v$ und $g(x, y) = (x^2 + y, xy^2)$, so dass $u = x^2 + y$ und $v = xy^2$ und

$$f(u, v) = \left(x^2 + y\right)^{xy^2}.$$

Wir erhalten

$$\begin{aligned} \partial_x f(u, v) &= \partial_u f \ \partial_x u + \partial_v f \ \partial_x v \\ &= v u^{v-1} 2x + \ln u \ u^v y^2 \\ &= \left(x^2 + y\right)^{xy^2 - 1} 2x^2 y^2 + \ln\left(x^2 + y\right) \left(x^2 + y\right)^{xy^2} y^2 \end{aligned}$$

und analog

$$\begin{aligned} \partial_y f(u, v) &= \partial_u f \ \partial_y u + \partial_v f \ \partial_y v \\ &= \left(x^2 + y\right)^{xy^2 - 1} xy^2 + \ln\left(x^2 + y\right) \left(x^2 + y\right)^{xy^2} 2xy \end{aligned}$$

Korollar 7.7 (Ableitung der inversen Funktion) *Seien $U \subset \mathbb{R}^n$ und $V \subset \mathbb{R}^n$ offene Mengen. Sei $g : U \to V$ eine bijektive Funktion die in einem Punkt $x \in U$ differenzierbar ist. Sei die inverse Funktion $f = g^{-1}$ im Punkt $v = g(x)$ differenzierbar. Dann es gilt*

$$f'(v) = g'(x)^{-1}. \tag{7.15}$$

Die beiden linearen Abbildung $f'(v)$ und $g'(x)$ sind von $\mathbb{R}^n$ nach $\mathbb{R}^n$ und somit werden mit $n \times n$ Matrizen dargestellt. Dann ist $f'(v)$ die inverse Matrix von $g'(x)$.

Beweis. Die Komposition $f \circ g$ ist die identische Funktion $I : U \to U$, d.h.$I(x) = x$. Dann gilt $I'(x) = \mathrm{Id}$ wobei $\mathrm{Id} : \mathbb{R}^n \to \mathbb{R}^n$ die identische Abbildung ist. Somit haben wir nach der Kettenregel

$$\mathrm{Id} = I'(x) = (f \circ g)'(x) = f'(v) g'(x),$$

woraus (7.15) folgt. ■

Beispiel. Betrachten wir die kartesische (x,y) Koordinaten in $\mathbb{R}^2 \setminus \{0\}$ als Funktionen von den Polarkoordinaten (r,θ), d.h.

$$(x,y) = (r\cos\theta, r\sin\theta) =: g(r,\theta).$$

Die totale Ableitung von F existiert und stimmt mit der Jacobi-Matrix überein

$$g' = J_g = \begin{pmatrix} \partial_r x & \partial_\theta x \\ \partial_r y & \partial_\theta y \end{pmatrix} = \begin{pmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{pmatrix}, \tag{7.16}$$

da J_g stetig bezüglich (r,θ) ist. Sei h die inverse Abbildung von g, d.h. h ergibt die Polarkoordinaten durch die kartesischen Koordinaten,

$$(r,\theta) = h(x,y).$$

Dann wir haben nach (7.15)

$$\begin{aligned} h' &= (g')^{-1} = \begin{pmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{pmatrix}^{-1} \\ &= \begin{pmatrix} \cos\theta & \sin\theta \\ -\frac{1}{r}\sin\theta & \frac{1}{r}\cos\theta \end{pmatrix} = \begin{pmatrix} x/r & y/r \\ -y/r^2 & x/r^2 \end{pmatrix}. \end{aligned}$$

Andererseits gilt

$$h' = \begin{pmatrix} \partial_x r & \partial_y r \\ \partial_x \theta & \partial_y \theta \end{pmatrix}$$

woraus folgt

$$\begin{aligned} \partial_x r &= \frac{x}{r}, \quad \partial_y r = \frac{y}{r} \\ \partial_x \theta &= -\frac{y}{r^2}, \quad \partial_y \theta = \frac{x}{r^2}. \end{aligned} \tag{7.17}$$

Natürlich erhält man diese Identitäten auch direct aus $r = \sqrt{x^2+y^2}$ and $\tan\theta = y/x$.

Sei f eine total differenzierbare Funktion von (x,y) einer offenen Teilmenge von $\mathbb{R}^2 \setminus \{0\}$. Einsetzen x und y als Funktionen von r,θ ergibt uns f als Funktion von r,θ. Mit Hilfe von (7.16) erhalten wir

$$\partial_r f = \partial_r (f \circ g) = \partial_x f \partial_r x + \partial_y f \partial_r y = \partial_x f \cos\theta + \partial_y f \sin\theta$$

und

$$\partial_\theta f = \partial_x f \partial_\theta x + \partial_y f \partial_\theta y = r(-\partial_x f \sin\theta + \partial_y f \cos\theta).$$

Umgekehrt, ist f eine total differenzierbare Funktion von (r,θ), so erhalten wir mit Hilfe von (7.17)

$$\partial_x f = \partial_r f \partial_x r + \partial_\theta f \partial_x \theta = \frac{x}{r}\partial_r f - \frac{y}{r^2}\partial_\theta f$$

und

$$\partial_y f = \partial_r f \partial_y r + \partial_\theta f \partial_y \theta = \frac{y}{r}\partial_r f + \frac{x}{r^2}\partial_\theta f.$$

Zum Beispiel, für $f = r^a \sin b\theta$ erhalten wir

$$\partial_x f = ar^{a-1} \sin b\theta \frac{x}{r} - r^a b \cos b\theta \frac{y}{r^2} = r^{a-1} \left(a \sin b\theta \cos\theta - b \cos b\theta \sin\theta \right)$$

und

$$\partial_y f = ar^{a-1} \sin b\theta \frac{y}{r} + r^a b \cos b\theta \frac{x}{r^2} = r^{a-1} \left(a \sin b\theta \sin\theta + b \cos b\theta \cos\theta \right)$$

7.3 Richtungsableitung und Mittelwertsatz

Definition. Sei f eine reellwertige Funktion auf einer offenen Menge $\Omega \subset \mathbb{R}^n$. Für jedes $x \in \Omega$ und für jeden Vektor $v \in \mathbb{R}^n$ definieren wir die Richtungsableitung

$$\partial_v f(x) = \frac{\partial f}{\partial v}(x) := \lim_{t \to 0} \frac{f(x+tv) - f(v)}{t} = \left. \frac{df(x+tv)}{dt} \right|_{t=0},$$

wobei t eine reelle Variable ist.

Die partielle Ableitung $\partial_j f$ ist ein spezieller Fall der Richtungsableitung. In der Tat betrachten wir den Basisvektor

$$e_j = (0, ..., 0, \overset{j}{\widehat{1}}, 0, .., 0)$$

wobei die Eins an der Position j steht. Dann gilt

$$\partial_j f = \partial_{e_j} f,$$

da

$$\begin{aligned} \partial_{e_j} f &= \lim_{t \to 0} \frac{f(x + te_j) - f(x)}{t} = \\ &= \lim_{t \to 0} \frac{f(x_1, ..., x_j + t, ..., x_n) - f(x_1, ..., x_n)}{t} = \partial_j f. \end{aligned}$$

Satz 7.8 *Ist f total differenzierbar in x dann für jedes $v \in \mathbb{R}^n$ existiert die Richtungsableitung $\partial_v f(x)$ und es gilt*

$$\partial_v f(x) = f'(x) v = \sum_{j=1}^{n} \partial_j f(x) v_j.$$

Beweis. Es folgt aus der Definition von totaler Differenzierbarkeit, dass

$$f(x+tv) - f(x) = f'(x)(tv) + o(t) \quad \text{für } t \to 0$$

Dividieren durch t ergibt

$$\partial_v f = \lim_{t \to 0} \frac{f(x+tv) - f(x)}{t} = \lim_{t \to 0} \left(f'(x) v + \frac{o(t)}{t} \right) = f'(x) v.$$

■

Insbesondere sehen wir, dass die Abbildung $v \mapsto \partial_v f(x)$ linear ist, was aus der Definition nicht offensichtlich ist.

Im nächsten Satz benutzen wir die Strecke $[x, y]$ zwischen zwei Punkten $x, y \in \mathbb{R}^n$, d.h.

$$[x, y] = \{(1-t)x + ty : 0 \leq t \leq 1\}.$$

Satz 7.9 (Mittelwertsatz) *Sei f eine reellwertige total differenzierbare Funktion auf einer offenen Menge $\Omega \subset \mathbb{R}^n$. Seien x, y zwei Punkte in Ω mit $[x, y] \subset \Omega$. Dann es gibt einen Punkt $\xi \in [x, y]$ mit*

$$f(y) - f(x) = f'(\xi)(y - x).$$

Beweis. Setzen wir $v = y - x$ und betrachten die Funktion

$$g(t) = f(x + tv) \quad \text{für} \quad t \in [0, 1],$$

so dass $g(0) = f(x)$ und $g(1) = f(y)$. Die Funktion g ist in $[0, 1]$ differenzierbar als Komposition von den Funktionen $t \mapsto x + tv$ und f. Nach dem Mittelwertsatz aus Analysis I gibt es ein $s \in [0, 1]$ mit

$$g(1) - g(0) = g'(s).$$

Setzen wir

$$\xi := x + sv = x + s(y - x) = (1 - s)x + sy,$$

so dass $\xi \in [x, y]$, und bemerken, dass

$$\begin{aligned} \partial_v f(\xi) &= \left.\frac{df(\xi + tv)}{dt}\right|_{t=0} = \left.\frac{df(x + (t + s)v)}{dt}\right|_{t=0} \\ &= \left.\frac{df(x + \tau v)}{d\tau}\right|_{\tau=s} = g'(s). \end{aligned}$$

Nach dem Satz 7.8 erhalten wir

$$f(y) - f(x) = g(1) - g(0) = g'(s) = \partial_v f(\xi) = f'(x)v = f'(x)(y - x),$$

was zu beweisen war. ■

7.4 Partielle Ableitungen höherer Ordnung und Satz von Schwarz

Seien $\Omega \subset \mathbb{R}^n$ eine offene Menge und und f eine reellwertige Funktion auf Ω. Existiert die partielle Ableitung $\partial_j f$ in Ω, so man kann diese Funktion weiter ableiten und die partielle Ableitung 2-ter Ordnung betrachten:

$$\partial_i(\partial_j f).$$

Existiert diese Ableitung, so bezeichnet man sie mit $\partial_{ij} f$ oder $\frac{\partial^2 f}{\partial x_i \partial x_j}$. Für $i = j$ schreibt man $\partial_{ii} f = \frac{\partial^2 f}{\partial x_i^2}$. If Fall $i \neq j$ heißt $\frac{\partial^2 f}{\partial x_i \partial x_j}$ die gemischte Ableitung.

Analog kann man die partiellen Ableitungen höherer Ordnung betrachten wie folgt

$$\partial_{i_1} \partial_{i_2} ... \partial_{i_k} f = \partial_{i_1 i_2 ... . i_k} f = \frac{\partial^k f}{\partial x_{i_1} \partial x_{i_2} ... \partial x_{i_k}}.$$

Die Zahl k hier heißt die Ordnung der Ableitung. Die Ableitung der Ordnung 0 ist die Funktion f selbst.

Satz 7.10 (Satz von Hermann Schwarz) *Angenommen, dass die Funktion $f : \Omega \to \mathbb{R}$ in Ω die beiden partiellen Ableitungen $\partial_{ij} f$ und $\partial_{ji} f$ hat. Sind $\partial_{ij} f$ und $\partial_{ji} f$ in einem Punkt $x \in \Omega$ stetig, so gilt $\partial_{ij} f(x) = \partial_{ji} f(x)$.*

Beweis. Ohne Beschränkung der Allgemeinheit nehmen wir an, dass $x = 0$ und $i = 1$, $j = 2$. Im Beweis werden die Variablen $x_3, ..., x_n$ konstant sein. Somit können wir die Funktion f als eine Funktion nur von x_1, x_2 betrachten, d.h. we können auch annehmen, dass $n = 2$. Bezeichnen wir die Koordinaten in $\mathbb{R}^2$ mit (x, y).

Da die Funktion $f = f(x, y)$ in der Nähe von $(0, 0)$ definiert ist, so gibt es ein $\varepsilon > 0$ so dass f im Quadrat

$$Q = \left\{(x, y) \in \mathbb{R}^2 : |x| < \varepsilon, |y| < \varepsilon\right\}$$

definiert ist. Betrachten wir im Q die Funktion

$$F(x, y) = f(x, y) - f(x, 0) - f(0, y) + f(0, 0)$$

und beweisen die folgende Aussage: für jedes $(x, y) \in Q$ existieren $s \in [0, x]$ und $t \in [0, y]$ mit

$$F(x, y) = \partial_{21} f(s, t)\, xy. \tag{7.18}$$

Dafür fixieren wir $(x, y) \in Q$ und betrachten die Funktion

$$\varphi(s) = f(s, y) - f(s, 0) \text{ für } |s| < \varepsilon,$$

so dass

$$\begin{aligned} F(x, y) &= f(x, y) - f(x, 0) - (f(0, y) - f(0, 0)) \\ &= \varphi(x) - \varphi(0). \end{aligned}$$

Nach Voraussetzung ist φ differenzierbar. Nach dem Mittelwertsatz existiert ein $s \in [0, x]$ mit

$$\varphi(x) - \varphi(0) = \varphi'(s)\, x = (\partial_1 f(s, y) - \partial_1 f(s, 0))\, x. \tag{7.19}$$

Die Funktion

$$\psi(t) = \partial_1 f(s, t) \quad \text{für } |t| < \varepsilon$$

ist differenzierbar und somit existiert ein $t \in [0, y]$ mit

$$\psi(y) - \psi(0) = \psi'(t)\, y = \partial_2 \partial_1 f(s, t)$$

d.h.

$$\partial_1 f(s,y) - \partial_1 f(s,0) = \partial_{21} f(s,t)\, y. \tag{7.20}$$

Einsetzen in (7.19) ergibt

$$\varphi(x) - \varphi(0) = \partial_{21} f(s,t)\, xy,$$

woraus (7.18) folgt.

Analog beweist man die Existenz von $\widetilde{s} \in [0,x]$ und $\widetilde{t} \in [0,y]$ mit

$$F(x,y) = \partial_{12} f\left(\widetilde{s},\widetilde{t}\right) xy. \tag{7.21}$$

Dafür stellen wir F wie folgt dar:

$$\begin{aligned} F(x,y) &= (f(x,y) - f(0,y)) - (f(x,0) - f(0,0)) \\ &= \widetilde{\psi}(y) - \widetilde{\psi}(0), \end{aligned}$$

wobei

$$\widetilde{\psi}(t) = f(x,t) - f(0,t).$$

Dann existiert ein $\widetilde{t} \in [0,y]$ mit

$$\widetilde{\psi}(y) - \widetilde{\psi}(0) = \widetilde{\psi}'\left(\widetilde{t}\right) y = \left(\partial_2 f\left(x,\widetilde{t}\right) - \partial_2 f\left(0,\widetilde{t}\right)\right) y = \left(\widetilde{\varphi}(x) - \widetilde{\varphi}(0)\right) y,$$

wobei

$$\widetilde{\varphi}(s) = \partial_2 f\left(s,\widetilde{t}\right).$$

Nach dem Mittelwertsatz erhalten wir ein $\widetilde{s} \in [0,x]$ mit

$$\widetilde{\varphi}(x) - \widetilde{\varphi}(0) = \widetilde{\varphi}'(\widetilde{s})\, x = \partial_1 \partial_2 f\left(\widetilde{s},\widetilde{t}\right) x,$$

woraus (7.21) folgt.

Beim Vergleichen von (7.18) und (7.21) erhalten wir für $x, y \neq 0$, dass

$$\partial_{21} f(s,t) = \partial_{12} f\left(\widetilde{s},\widetilde{t}\right). \tag{7.22}$$

Bemerken wir, dass die Variablen $s, t, \widetilde{s}, \widetilde{t}$ die Funktionen von x, y sind. Nach Konstruktion konvergieren $s, t, \widetilde{s}, \widetilde{t}$ gegen 0 für $x, y \to 0$. Da $\partial_{21} f$ und $\partial_{12} f$ in $(0,0)$ stetig sind, so erhalten wir aus (7.22) für $x, y \to 0$, dass

$$\partial_{21} f(0,0) = \partial_{12} f(0,0),$$

was zu beweisen war. ∎

Ohne die Voraussetzung von Stetigkeit können Sie Ableitungen $\partial_{ij} f$ und $\partial_{ji} f$ verschieden sein, wie im nächsten Beispiel.

Beispiel. Betrachten wir die folgende Funktion in $\mathbb{R}^2$

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

und berechnen $\partial_{12} f(0)$ und $\partial_{21} f(0)$. Nach Definition gilt

$$\partial_{12} \partial f(0,0,) = \partial_1 \partial_2 f(0,0) = \lim_{x \to 0} \frac{\partial_2 f(x,0) - \partial_2 f(0,0)}{x}.$$

So, berechnen wir zuerst $\partial_2 f(x,0)$:

$$\partial_2 f(x,0) = \lim_{y \to 0} \frac{f(x,y) - f(x,0)}{y} = \lim_{y \to 0} x \frac{x^2 - y^2}{x^2 + y^2} = x.$$

Da auch $\partial_2 f(0,0) = 0$, so erhalten wir

$$\partial_1 \partial_2 f(0,0) = \lim_{x \to 0} \frac{x - 0}{x} = 1.$$

Analog haben wir

$$\partial_1 f(0,y) = \lim_{x \to 0} \frac{f(x,y) - f(0,y)}{x} = \lim_{x \to 0} y \frac{x^2 - y^2}{x^2 + y^2} = -y$$

und

$$\partial_{21} f(0,0) = \lim_{y \to 0} \frac{\partial_1 f(0,y) - \partial_1 f(0,0)}{y} = \lim_{y \to 0} \frac{-y - 0}{y} = -1.$$

Somit gilt $\partial_{12} f(0) \neq \partial_{21} f(0)$.

Definition. Eine Funktion $f : \Omega \to \mathbb{R}$ heißt *k-fach stetig differenzierbar* falls alle partielle Ableitungen von f der Ordnung $\leq k$ existieren und stetig in Ω sind. Die Menge von allen k-fach stetig differenzierbaren Funktionen auf Ω wird mit $C^k(\Omega)$ bezeichnet. Insbesondere wird mit $C(\Omega) = C^0(\Omega)$ die Menge von allen stetigen Funktionen auf Ω bezeichnet.

Es ist klar aus der Definition, dass $C^k(\Omega) \subset C^{k-1}(\Omega)$.

Korollar 7.11 *Für jede Funktion $f \in C^k(\Omega)$ ist der Wert von jeder partiellen Ableitung der Ordnung $\leq k$ unabhängig von der Reihenfolge von Ableiten. D.h., für jede Folge $i_1, ..., i_m$ von $m \leq k$ Indizes und für jede Permutation $j_1, ..., j_m$ von $i_1, ..., i_m$ gilt $\partial_{i_1 \dots i_m} f = \partial_{j_1 \dots j_m} f$.*

Beweis. Nach dem Satz 7.10 gilt folgendes: jede zwei nacheinander stehende Indizes in der Folge $i_1, ... i_m$, z.B. i_l und i_{l+1}, lassen sich vertauschen ohne den Wert von $\partial_{i_1 \dots i_m} f$ zu ändern, da

$$\begin{aligned} \partial_{i_1 \dots i_l i_{l+1} \dots i_m} f &= \partial_{i_1 \dots i_{l-1}} \left(\partial_{i_l} \partial_{i_{l+1}} \right) \partial_{i_{l+2} \dots i_m} f \\ &= \partial_{i_1 \dots i_{l-1}} \left(\partial_{i_{l+1}} \partial_{i_l} \right) \partial_{i_{l+2} \dots i_m} f \\ &= \partial_{i_1 \dots i_{l+1} i_l \dots i_m} f. \end{aligned}$$

Da jede Permutation $j_1, ..., j_m$ von $i_1, ..., i_m$ sich aus $i_1, ..., i_m$ mit Hilfe von einer Reihe von Vertauschen von nacheinander stehenden Indizes erhalten lässt, so gilt $\partial_{i_1 \dots i_m} f = \partial_{j_1 \dots j_m} f$ ■

Für Funktionen $f \in C^k(\Omega)$ benutzen wir die folgende Notation für die partiellen Ableitungen m-er Ordnung mit $m \leq k$:

$$\frac{\partial^m f}{\partial x_1^{\alpha_1} \partial x_2^{\alpha_2} ... \partial x_n^{\alpha_n}}, \tag{7.23}$$

wobei $\alpha_1 + ... + \alpha_n = m$. Diese Notation bedeutet folgendes: man leitet f α_1 mal in x_1, α_2 mal in x_3 usw. ab, d.h.

$$\frac{\partial^m f}{\partial x_1^{\alpha_1} \partial x_2^{\alpha_2} ... \partial x_n^{\alpha_n}} := \partial_{\underbrace{1...1}_{\alpha_1} \underbrace{2...2}_{\alpha_2} \cdots \underbrace{n...n}_{\alpha_n}} f \ .$$

Nach Korollar 7.11 lässt jede partielle Ableitung von f der Ordnung $\leq k$ sich in der Form (7.23) darstellen.

Definition. Jede Folge $\alpha = (\alpha_1, ..., \alpha_n)$ von nichtnegativen ganzen Zahlen α_k heißt *Multiindex* von Dimension n. Die Menge von allen Multiindizes von Dimension n wird mit $\mathbb{I}^n$ bezeichnet. Für jeden Multiindex $\alpha \in \mathbb{I}^n$ definieren wir die Ordnung (Betrag) von α mit

$$|\alpha| = \alpha_1 + ... + \alpha_n.$$

Für jede Funktion $f \in C^k(\Omega)$ und für jeden Multiindex $\alpha \in \mathbb{I}^n$ mit $|\alpha| \leq k$ definieren wir die α-Ableitung von f mit

$$D^\alpha f = \frac{\partial^{|\alpha|} f}{\partial x_1^{\alpha_1} \partial x_2^{\alpha_2} ... \partial x_n^{\alpha_n}}.$$

Insbesondere $D^0 f = f$.

Es folgt aus dem Korollar 7.11, dass für jede Funktion $f \in C^k(\Omega)$ und für beliebige Multiindizes α, β mit $|\alpha| + |\beta| \leq k$ gilt

$$D^{\alpha+\beta} f = D^\alpha \left(D^\beta f\right). \tag{7.24}$$

7.5 Taylorformel

Im nächsten Satz benutzen wir die folgende Notation: für jeden Multiindex $\alpha = (\alpha_1, ..., \alpha_n) \in \mathbb{I}^n$ setzen wir

$$\alpha! := \alpha_1!...\alpha_n!,$$

und für jeden Vektor $h = (h_1, ..., h_n) \in \mathbb{R}^n$

$$h^\alpha := h_1^{\alpha_1}...h_n^{\alpha_n}$$

(wobei $h_i^{\alpha_i} = 1$ für $\alpha_i = 0$).

Hauptsatz 7.12 (Taylorformel mit der Restgliedform nach Peano) *Sei Ω eine offene Teilmenge von $\mathbb{R}^n$. Für jede Funktion $f \in C^k(\Omega)$ mit $k \geq 0$ und für jedes $x \in \Omega$ gilt*

$$f(x+h) = \sum_{\{\alpha \in \mathbb{I}^n : |\alpha| \leq k\}} \frac{D^\alpha f(x)}{\alpha!} h^\alpha + o\left(\|h\|^k\right) \ \text{für } h \to 0. \tag{7.25}$$

Umgekehrt, gilt für reelle Koeffizienten c_α

$$f(x+h) = \sum_{\{\alpha\in\mathbb{I}^n:|\alpha|\le k\}} c_\alpha h^\alpha + o\left(\|h\|^k\right) \text{ für } h\to 0, \tag{7.26}$$

so haben wir $c_\alpha = \frac{D^\alpha f(x)}{\alpha!}$.

Nach der Offenheit von Ω liegt $x+h$ für hinreichend klein $\|h\|$ in Ω, so dass $f(x+h)$ wohldefiniert ist.

Definition. Die Funktion

$$T_k(h) := \sum_{\{\alpha\in\mathbb{I}^n:|\alpha|\le k\}} \frac{D^\alpha f(x)}{\alpha!}h^\alpha = \sum_{\{\alpha\in\mathbb{I}^n:|\alpha|\le k\}} \frac{1}{\alpha_1!...\alpha_n!}\frac{\partial^{|\alpha|} f(x)}{\partial x_1^{\alpha_1}...\partial x_n^{\alpha_n}} h_1^{\alpha_1}...h_n^{\alpha_n} \tag{7.27}$$

heißt *Taylor-Polynom* der Ordnung k von der Funktion f im Punkt x. Die ausführliche Notation für das Taylor-Polynom ist $T_{k,f}(h;x)$.

Es ist klar, dass $T_k(h)$ wirklich ein Polynom bezüglich der Variablen $h_1,...,h_n$ ist und $T_k(0) = f(x)$. Die Taylorformel lässt sich wie folgt umschreiben:

$$f(x+h) = T_k(h) + o(\|h\|^k) \text{ für } h\to 0. \tag{7.28}$$

Somit ist $T_k(h)$ eine Approximation von $f(x+h)$ für kleine Werte von h mit dem Approximationsfehler $o(\|h\|^k)$.

Die zweite Aussage des Satzes 7.12 bedeutet folgendes. Gilt

$$f(x+h) = P(h) + o(\|h\|^k) \text{ für } h\to 0 \tag{7.29}$$

für ein Polynom $P(h) = \sum c_\alpha h^\alpha$ von h des Grades $\le k$, so ist P identisch gleich T_k. Somit ist T_k das einzige Polynom des Grades $\le k$, das (7.29) erfüllt.

Seien $n=1$ und Ω ein offenes Intervall. Dann ergibt (7.27)

$$T_k(h) = \sum_{\alpha=0}^{k} \frac{f^{(\alpha)}(x)}{\alpha!}h^\alpha$$

und wir erhalten die Taylorformel aus Analysis I:

$$f(x+h) = \sum_{\alpha=0}^{k} \frac{f^{(\alpha)}(x)}{\alpha!}h^\alpha + o\left(|h|^k\right) \quad \text{für } h\to 0.$$

Sei n beliebig. Berechnen wir explizit $T_k(h)$ für $k=0,1,2,3$. Für $|\alpha|=0$ gilt $\alpha=(0,...,0)$ und somit

$$\frac{D^\alpha f(x)}{\alpha!}h^\alpha = f(x)$$

und somit

$$T_0(h) = f(x) = \text{const}.$$

Für $|\alpha| = 1$ gilt

$$\alpha = (0,...,\overset{i}{\widehat{1}},...,0)$$

mit einziger 1 an einer Position i und mit allen anderen Komponenten gleich 0. Somit haben wir

$$\frac{D^\alpha f(x)}{\alpha!} h^\alpha = \partial_i f(x)\, h_i$$

und

$$\sum_{\{\alpha \in \mathbb{I}^n : |\alpha|=1\}} \frac{D^\alpha f(x)}{\alpha!} h^\alpha = \sum_{i=1}^n \partial_i f(x)\, h_i = \partial_1 f(x)\, h_1 + ... + \partial_n f(x)\, h_n.$$

Es folgt, dass

$$T_1(h) = f(x) + \partial_1 f(x)\, h_1 + ... + \partial_n f(x)\, h_n = f(x) + f'(x)\, h.$$

Für $|\alpha| = 2$ gibt es zwei Möglichkeiten: entweder

$$\alpha = (0, ..., \overset{i}{\widehat{2}}, ..., 0)$$

mit 2 an einer Position i oder

$$\alpha = (0, ..., \overset{i}{\widehat{1}}, ..., \overset{j}{\widehat{1}}, ..., 0),$$

wobei 1 zweimal an den Positionen $i < j$ steht. Im ersten Fall haben wir

$$\frac{D^\alpha f(x)}{\alpha!} h^\alpha = \frac{\partial_{ii} f(x)}{2} h_i^2,$$

und im zweiten Fall

$$\frac{D^\alpha f(x)}{\alpha!} h^\alpha = \partial_{ij} f(x)\, h_i h_j,$$

woraus folgt

$$\sum_{|\alpha|=2} \frac{D^\alpha f(x)}{\alpha!} h^\alpha = \sum_i \frac{\partial_{ii} f(x)}{2} h_i^2 + \sum_{i<j} \partial_{ij} f(x)\, h_i h_j$$

Somit erhalten wir

$$T_2(h) = f(x) + \sum_i \partial_i f(x)\, h_i + \sum_i \frac{\partial_{ii} f(x)}{2} h_i^2 + \sum_{i<j} \partial_{ij} f(x)\, h_i h_j. \tag{7.30}$$

Analog bestimmt man T_3:

$$T_3(h) = T_2(h) + \sum_i \frac{\partial_{iii} f(x)}{6} h_i^3 + \sum_{i \neq j} \frac{\partial_{iij} f(x)}{2} h_i^2 h_j + \sum_{i<j<l} \partial_{ijl} f(x)\, h_i h_j h_l. \tag{7.31}$$

Beispiel. Betrachten wir den Fall $n = 2$ und $x = 0$. Dann bezeichnen wir die Koordinaten in $\mathbb{R}^2$ mit (x, y) statt (x_1, x_2) und setzen $h = (x, y)$. Es folgt aus (7.30), dass das Taylor-Polynom T_2 in 0 ist wie folgt:

$$T_2(x, y) = f(0) + \partial_x f(0)\, x + \partial_y f(0)\, y + \frac{1}{2}\partial_{xx} f(0)\, x^2 + \frac{1}{2}\partial_{yy} f(0)\, y^2 + \partial_{xy} f(0)\, xy, \tag{7.32}$$

und aus (7.31)

$$T_3(x,y) = T_2(x,y) + \frac{1}{6}\partial_{xxx} f(0)\, x^3 + \frac{1}{6}\partial_{yyy} f(0)\, y^3 + \frac{1}{2}\partial_{xxy} f(0)\, x^2 y + \frac{1}{2}\partial_{xyy} f(0)\, x y^2.$$

Zum Beispiel, berechnen wir T_2 und T_3 in 0 für die Funktion $f(x,y) = \cos(x-y)$. Dafür brauchen wir alle partielle Ableitungen $\partial_x f$, $\partial_y f$, $\partial_{xx} f$, $\partial_{xy} f$, $\partial_{yy} f$ in 0. Wir haben

$$\partial_x f = -\sin(x-y), \quad \partial_y f = \sin(x-y),$$

und

$$\partial_{xx} f = -\cos(x-y), \quad \partial_{xy} = \cos(x-y), \quad \partial_{yy} f = -\cos(x-y)$$

woraus folgt

$$\partial_x f(0,0) = \partial_y f(0,0) = 0$$

und

$$\partial_{xx} f(0,0) = \partial_{yy} f(0,0) = -1 \text{ und } \partial_{xy} f(0,0) = 1.$$

Somit erhalten wir aus (7.32)

$$T_2(x,y) = 1 - \frac{1}{2}x^2 - \frac{1}{2}y^2 + xy. \tag{7.33}$$

Da die dritten Ableitungen $\partial_{xxx} f$, $\partial_{yyy} f$, $\partial_{xxy} f$, $\partial_{xyy} f$ gleich $\pm\sin(x-y)$ sind und somit in $(0,0)$ verschwinden, so erhalten wir die gleiche Formel auch für T_3:

$$T_3(x,y) = 1 - \frac{1}{2}x^2 - \frac{1}{2}y^2 + xy. \tag{7.34}$$

Alternativ lassen sich T_2 und T_3 wie folgt bestimmen. Mit Hilfe von Taylorformel für $\cos x$ und $\sin x$ aus Analysis 1 erhalten wir

$$\begin{aligned}
\cos(x-y) &= \cos x \cos y + \sin x \sin y \\
&= \left(1 - \frac{x^2}{2} + o\left(x^3\right)\right)\left(1 - \frac{y^2}{2} + o\left(y^3\right)\right) + \left(x + o\left(x^2\right)\right)\left(y + o\left(y^2\right)\right) \\
&= 1 - \frac{x^2}{2} - \frac{y^2}{2} + xy + o\left(\|h\|_\infty^3\right),
\end{aligned}$$

woraus (7.33) und 7.34) nach der Eindeutigkeit des Taylor-Polynoms folgen.

Beweis von dem Satz 7.12. Fixieren wir ein $x \in \Omega$, setzen

$$R_k(h) := f(x+h) - T_k(h) \tag{7.35}$$

und beweisen per Induktion nach k, dass

$$R_k(h) = o\left(\|h\|^k\right) \text{ für } h \to 0 \tag{7.36}$$

Induktionsanfang. Für $k = 0$ wird (7.36)

$$f(x+h) - f(x) = o(1) \text{ für } h \to 0,$$

was nach der Stetigkeit von f gilt.

Induktionsschritt von $k-1$ nach k. Fixieren wir einen Index $i=1,...,n$ und bemerken, dass $\partial_i f \in C^{k-1}(\Omega)$.

Behauptung. *Das Taylor-Polynom der Ordnung $k-1$ von $\partial_i f$ ist gleich $\partial_i T_k$, d.h.*

$$T_{k-1,\partial_i f}(h) = \partial_i T_{k,f}(h). \tag{7.37}$$

Nach (7.27) haben wir

$$\partial_i T_k(h) = \sum_{\{\alpha\in\mathbb{I}^n:|\alpha|\le k\}} \frac{D^\alpha f(x)}{\alpha!}\partial_i h^\alpha. \tag{7.38}$$

Im Fall $\alpha_i = 0$ hängt h^α von h_i nicht ab, woraus folgt $\partial_i h^\alpha = 0$. Somit können wir annehmen, dass in der obigen Summe nur die Werte von α mit $\alpha_i \ge 1$ benutzt werden. Bezeichnen wir

$$\beta = (0,...,\overset{i}{\widehat{1}},...,0). \tag{7.39}$$

Dann ist $\alpha-\beta$ ein Multiindex der Ordnung $|\alpha|-1$ und es gelten die Identitäten:

$$\partial_i h^\alpha = \partial_i\left(h_1^{\alpha_1}...h_i^{\alpha_i}...h_n^{\alpha_n}\right) = \alpha_i(h_1^{\alpha_1}...h_i^{\alpha_i-1}...h_n^{\alpha_n}) = \alpha_i h^{\alpha-\beta},$$

$$\alpha! = \alpha_1!...\alpha_i!...\alpha_n! = \alpha_i\left(\alpha_1!...\left(\alpha_i-1\right)!...\alpha_n!\right) = \alpha_i\left(\alpha-\beta\right)!,$$

$$D^\alpha f = D^{\alpha-\beta}D^\beta f = D^{\alpha-\beta}\partial_i f.$$

Einsetzen in (7.38) und Wechsel $\gamma = \alpha-\beta$ ergeben:

$$\begin{aligned}\partial_i T_k(h) &= \sum_{\{\alpha\in\mathbb{I}^n:|\alpha|\le k,\alpha_i\ge 1\}} \frac{D^{\alpha-\beta}\partial_i f(x)}{\alpha_i(\alpha-\beta)!}\alpha_i h^{\alpha-\beta} \\ &= \sum_{\{\gamma\in\mathbb{I}^n:|\gamma|\le k-1\}} \frac{D^\gamma(\partial_i f)(x)}{\gamma!}h^\gamma = T_{k-1,\partial_i f}(h),\end{aligned} \tag{7.40}$$

was (7.37) beweist.

Nach (7.37) und nach der Induktionsvoraussetzung für die Funktion $\partial_i f \in C^{k-1}(\Omega)$ erhalten wir

$$(\partial_i f)(x+h) = \partial_i T_k(h) + o\left(\|h\|^{k-1}\right) \text{ für } h\to 0,$$

was äquivalent zu

$$\partial_i R_k(h) = o\left(\|h\|^{k-1}\right) \text{ für } h\to 0 \tag{7.41}$$

ist (bemerken wir, dass $(\partial_{x_i} f)(x+h) = \partial_{h_i}(f(x+h))$ nach der Kettenregel).

Die Funktion $R_k(h)$ ist in einer Kugel $B(0,\varepsilon)$ wohldefiniert und gehört zur Klasse $C^k(B(0,\varepsilon))$. Da $k\ge 1$, so ist R_k differenzierbar in dieser Kugel. Nach dem Mittelwertsatz 7.9 erhalten wir, dass für ein Punkt $\xi\in[0,h]$

$$R_k(h) = R_k(h) - R_k(0) = R_k'(\xi)h = \sum_{i=1}^n \partial_i R_k(\xi)h_i.$$

Da $\|\xi\| \leq \|h\|$, so erhalten wir aus (7.41)

$$|R_k(h)| \leq \sum_{i=1}^{n} |\partial_i R_k(\xi)| \, \|h\|_\infty = o\left(\|\xi\|^{k-1}\right) \|h\|_\infty = o\left(\|h\|^k\right),$$

woraus (7.36) folgt.

Jetzt beweisen wir die Eindeutigkeit des Taylor-Polynoms. Gilt (7.26), so setzen wir

$$b_\alpha = \frac{D^\alpha f(x)}{\alpha!} - c_\alpha$$

und

$$Q(h) = \sum_{\{\alpha \in \mathbb{I}^n : |a| \leq k\}} b_\alpha h^\alpha,$$

so dass $Q(h)$ ein Polynom von h des Grades $\leq k$ ist. Es folgt aus (7.25) and (7.26), dass

$$Q(h) = o\left(\|h\|^k\right) \quad \text{für } k \to \infty. \tag{7.42}$$

Wir müssen beweisen, dass $b_\alpha = 0$ für alle α, was wir aus (7.42) gewinnen.

Zuerst zeigen wir, dass $Q(h) \equiv 0$ für alle $h \in \mathbb{R}^n$. Setzen wir $h = tv$ für $v \in \mathbb{R}^n$ und $t \in \mathbb{R}$ und stellen $Q(h)$ wie folgt dar:

$$Q(h) = \sum_{\{\alpha \in \mathbb{I}^n : |\alpha| \leq k\}} b_\alpha (tv)^\alpha = \sum_{\{\alpha \in \mathbb{I}^n : |\alpha| \leq k\}} b_\alpha t^{|\alpha|} v^\alpha = \sum_{j=0}^{k} Q_j(v) \, t^j,$$

wobei $Q_j(v)$ die Polynome von v sind. Für festes v und für $t \to 0$ erhalten wir aus (7.42)

$$\sum_{j=0}^{k} Q_j(v) \, t^j = o\left(t^k\right) \text{ für } t \to 0.$$

Da die linke Seite hier ein Polynom von t des Grades $\leq k$ ist, so erhalten wir nach der Taylor-Formel aus Analysis 1, dass alle Koeffizienten von diesem Polynom verschwinden, d.h.

$$Q_j(v) = 0 \text{ für alle } j = 0, ..., k \text{ und für alle } v \in \mathbb{R}^n,$$

woraus folgt

$$Q(h) = 0 \text{ für alle } h \in \mathbb{R}^n. \tag{7.43}$$

Beweisen wir jetzt per Induktion nach k, dass unter der Bedingung (7.43) alle Koeffizienten b_α von Q verschwinden. Für $k = 0$ ist das offensichtlich, da $Q(h) \equiv b_0 = \text{const}$. Für Induktionsschritt bemerken wir, dass analog zu (7.40) und mit β aus (7.39)

$$\partial_i Q(h) = \sum_{\{\alpha \in \mathbb{I}^n : |\alpha| \leq k, \alpha_i \geq 1\}} \alpha_i b_\alpha h^{\alpha - \beta} = \sum_{\{\gamma \in \mathbb{I}^n : |\gamma| \leq k-1\}} b'_\gamma h^\gamma,$$

wobei $\gamma = \alpha - \beta$ und $b'_\gamma = \alpha_i b_\alpha$. Da $\partial_i Q(h) \equiv 0$ und $\partial_i Q$ ein Polynom des Grades $\leq k - 1$ ist, so erhalten wir nach der Induktionsvoraussetzung, dass $b'_\gamma = 0$ für alle γ, woraus folgt, dass $b_\alpha = 0$ für alle α mit $\alpha_i \geq 1$. Da i beliebig ist, so erhalten wir $b_\alpha = 0$ für alle $\alpha \neq 0$. Für $\alpha = 0$ gilt $b_0 = Q(0) = 0$ auch. Somit sind alle b_α gleich 0, was zu beweisen war. ■

7.6 Lokale Extrema

Wir benutzen hier die Taylorformel um lokale Extrema von Funktionen in $\mathbb{R}^n$ zu bestimmen. Sei Ω eine offene Teilmenge von $\mathbb{R}^n$.

Definition. Eine Funktion $f : \Omega \to \mathbb{R}$ hat in einem Punkt $x \in \Omega$ *lokales Maximum* falls es eine Kugel $B(x,r) \subset \Omega$ gibt wo $f(x)$ der maximale Wert von f ist, d.h.

$$f(x) \geq f(y) \text{ für alle } y \in B(x,r).$$

Der Punkt x heißt die *lokale Maximumstelle* von f. Analog definiert man *lokale Minimumstelle*. Der Punkt x heißt *lokale Extremumstelle* von f, falls x lokale Maximum- oder Minimumstelle ist.

In diesem Abschnitt besprechen wir die notwendigen und hinreichenden Bedingungen für lokale Extrema. Wir fangen mit der Verallgemeinerung des Satzes von Fermat.

Satz 7.13 *Sei $x \in M$ eine lokale Extremumstelle von f in Ω. Ist f in x differenzierbar, so gilt $f'(x) = 0$.*

Die Bedingung $f'(x) = 0$ ist äquivalent zu $\partial_1 f(x) = \partial_2 f(x) = ... = \partial_n f(x) = 0$.

Beweis. Fixieren wir ein $h \in \mathbb{R}^n$ und betrachten die folgende Funktion $\varphi(t)$ von reeller Variable t:

$$\varphi(t) = f(x + th).$$

Diese Funktion ist in einem Intervall um 0 definiert und hat in 0 eine lokale Extremumstelle. Da φ in 0 differenzierbar, so erhalten wir nach dem Satz von Fermat, dass $\varphi'(0) = 0$. Da $\varphi'(0) = f'(x)h$, so folgt $f'(x)h = 0$ für alle $h \in \mathbb{R}^n$, woraus folgt $f'(x) = 0$. ■

Sei f in Ω differenzierbar. Die Punkte $x \in \Omega$ wo $f'(x) = 0$ heißen die *kritischen Punkten* von f. Um die lokalen Extremumstellen von f zu bestimmen, soll man zuerst alle kritischen Punkten finden und danach jeden kritischen Punkt weiter zu untersuchen.

Beispiel. Betrachten wir die Funktion $f(x,y) = \sin x \cos y$ in $\Omega = (-\pi,\pi) \times (-\pi,\pi)$. Da

$$\partial_x f = \cos x \cos y \quad \text{und} \quad \partial_y f = -\sin x \sin y,$$

so erfüllen die kritischen Punkte das System

$$\begin{cases} \cos x \cos y = 0 \\ \sin x \sin y = 0, \end{cases}$$

d.h. entweder $\cos x = 0$ und $\sin y = 0$ oder $\cos y = 0$ und $\sin x = 0$. Somit erhalten wir die folgenden kritischen Punkte:

$$\left(\frac{\pi}{2}, 0\right), \left(-\frac{\pi}{2}, 0\right), \left(0, \frac{\pi}{2}\right), \left(0, -\frac{\pi}{2}\right).$$

Wie wir später sehen, zwei davon sind wirklich die lokalen Extremumstellen und die anderen zwei sind nicht.

In Analysis 1 wird die hinreichende Bedingung für lokale Extrema mit Hilfe von zweiter Ableitung gegeben.

Definition. Für jede Funktion $f \in C^2(\Omega)$ definieren wir die totale zweite Ableitung $f''(x)$ als die folgende $n \times n$ Matrix:

$$f''(x) = (\partial_{ij} f(x))_{i,j=1}^n = \begin{pmatrix} \partial_{11} f(x) & \partial_{12} f(x) & ... & \partial_{1n} f(x) \\ \partial_{21} f(x) & \partial_{22} f(x) & ... & \partial_{2n} f(x) \\ ... & ... & ... & ... \\ \partial_{n1} f(x) & \partial_{n2} f(x) & ... & \partial_{nn} f(x) \end{pmatrix}.$$

Diese Matrix heißt auch die *Hesse-Matrix* von f.

Nach dem Satz 7.10 gilt $\partial_{ij} f = \partial_{ji} f$ so dass die Hesse-Matrix symmetrisch ist.

Jede symmetrische $n \times n$ Matrix $A = (a_{ij})$ mit reellen Einträgen bestimmt eine Bilinearform

$$Q(u,v) = \sum_{i,j=1}^n a_{ij} u_i v_j$$

wobei $u, v \in \mathbb{R}^n$, und die entsprechende *quadratische Form*

$$Q(u) := Q(u,u) = \sum_{i,j=1}^n a_{ij} u_i u_j$$

als Funktion von $u \in \mathbb{R}^n$.

Im Fall $A = f''(x)$ heißt die Bilinearform Q von $f''(x)$ das *zweite Differential* von f und wird mit $d^2 f(x)$ bezeichnet, d.h.

$$d^2 f(x)(u,v) = \sum_{i,j=1}^n \partial_{ij} f(x) u_i v_j.$$

Nach Aufgabe 109 gilt die folgende Identität

$$d^2 f(x)(u,v) = \partial_v (\partial_u f)(x)$$

für alle $u, v \in \mathbb{R}^n$, was als eine äquivalente Definition von $d^2 f(x)$ betrachtet werden kann.

Definition. Eine symmetrische $n \times n$ Matrix A (und ihre quadratische Form Q) heißt

- *positiv semidefinit* falls $Q(u) \geq 0$ für alle $u \in \mathbb{R}^n$; in diesem Fall schreiben wir $A \geq 0$;
- *positive definit* falls $Q(u) > 0$ für alle $u \neq 0$; Schreibweise $A > 0$;
- *negativ semidefinit* falls $Q(u) \leq 0$ für alle $u \in \mathbb{R}^n$; Schreibweise $A \leq 0$;
- *negativ definit* falls $Q(u) < 0$ für alle $u \neq 0$; Schreibweise $A < 0$;
- *indefinit* falls $Q(u)$ positive und negative Werte annimmt.

Bemerken wir, dass $Q(0) = 0$. Somit ist 0 eine Minimumstelle von Q genau dann, wenn $Q(u) \geq 0$ für alle $u \geq 0$, d.h. wenn $A \geq 0$. Analog ist 0 eine Maximumstelle von Q genau dann, wenn $A \leq 0$.

Beispiel. Die identische Matrix $A = \mathrm{id}$ erzeugt die quadratische Form $Q(u) = u_1^2 + ... + u_n^2$, die offensichtlich positiv definit ist, so dass $\mathrm{id} > 0$. Offensichtlich ist 0 eine Minimumstelle von Q.

Im Fall $n = 2$ betrachten wir die Matrix $A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ deren quadratische Form ist

$$Q(u) = a_{11}u_1^2 + a_{12}u_1u_2 + a_{21}u_2u_1 + a_{22}u_2^2 = 2u_1u_2.$$

Da $Q(u)$ positive und negative Werte annimmt, so ist A in diesem Fall indefinit, und 0 ist keine Extremumstelle.

Satz 7.14 *Seien $\Omega \subset \mathbb{R}^n$ eine offene Menge und f eine Funktion von $C^2(\Omega)$. Sei x ein kritischer Punkt von f, d.h. $f'(x) = 0$.*

(a) (Notwendige Bedingung für lokales Extremum) *Ist x eine lokale Maximumstelle von f, so gilt $f''(x) \leq 0$. Ist x eine lokale Minimumstelle von f so gilt $f''(x) \geq 0$.*

(b) (Hinreichende Bedingung für lokales Extremum) *Gilt $f''(x) < 0$ so ist x eine lokale Maximumstelle von x. Gilt $f''(x) > 0$ so ist x eine lokale Minimumstelle von f.*

Als eine Folgerung von (a) erhalten wir folgendes: ist $f''(x)$ indefinit so ist x keine Extremumstelle.

Beweis. Nach (7.30) haben wir für das Taylor-Polynom von f in x

$$\begin{aligned} T_2(h) &= f(x) + \sum_{i=1}^n \partial_i f(x) h_i + \sum_{i=1}^n \frac{\partial_{ii} f(x)}{2} h_i^2 + \sum_{1 \leq i < j \leq n} \partial_{ij} f(x) h_i h_j \\ &= f(x) + \sum_{i=1}^n \partial_i f(x) h_i + \frac{1}{2} \sum_{i,j=1}^n \partial_{ij} f(x) h_i h_j \\ &= f(x) + f'(x) h + \frac{1}{2} Q(h), \end{aligned}$$

wobei Q die quadratische Form der Hesse-Matrix $f''(x)$ ist. Da $f'(x) = 0$, so erhalten wir nach der Taylorformel

$$f(x+h) - f(x) = \frac{1}{2} Q(h) + o\left(\|h\|^2\right) \text{ für } h \to 0. \tag{7.44}$$

(a) Sei x eine lokale Minimumstelle. Beweisen wir, dass $f''(x) \geq 0$ d.h. $Q(h) \geq 0$ für alle $h \in \mathbb{R}^n$. Ersetzen wir in (7.44) h mit th für $t \in \mathbb{R}$ und erhalten

$$f(x+th) - f(x) = \frac{1}{2} Q(h) t^2 + o\left(t^2\right) \quad \text{für } t \to 0.$$

Da x eine lokale Minimumstelle ist, so gilt $f(x+th) \geq f(x)$ für hinreichend kleinen Werte von t, woraus folgt

$$\frac{1}{2} Q(h) t^2 + o\left(t^2\right) \geq 0.$$

Dividieren mit t^2 ergibt für $t \to 0$, dass $Q(h) \geq 0$, was zu beweisen war.

Analog beweist man $f''(x) \leq 0$ im Fall wenn x eine lokale Maximumstelle ist.

(b) Sei $f''(x) > 0$, d.h. $Q(h) > 0$ für alle $h \in \mathbb{R}^n \setminus \{0\}$. Beweisen, wir, dass

$$f(x+h) > f(x)$$

für alle $h \neq 0$, vorausgesetzt, dass $\|h\|$ hinreichend klein ist. Nach (7.44) reicht es zu beweisen, dass für solche Werte von h

$$\frac{1}{2}Q(h) + o\left(\|h\|^2\right) > 0.$$

Die Funktion $Q(h)$ ist offensichtlich stetig auf $\mathbb{R}^n$, da Q eine lineare Kombination von stetigen Funktionen $h_i h_j$ ist. Betrachten wir die Menge

$$S = \{h \in \mathbb{R}^n : \|h\| = 1\}$$

(die der Rand der Kugel $B(0,1)$ ist). Die Menge S ist offensichtlich beschränkt (da $S \subset B(0,2)$) und abgeschlossen, da S das Urbild von $\{1\}$ unter der stetigen Abbildung $h \mapsto \|h\|$ ist. Nach dem Extremwertsatz (Korollar 6.19) besitzt die Funktion $Q(h)$ ein Minimum auf S. Sei $m = \min_S Q$.

Da $Q|_S > 0$, so gilt $m > 0$. Für jedes $h \neq 0$ gilt $\frac{h}{\|h\|} \in S$, woraus folgt

$$Q\left(\frac{h}{\|h\|}\right) \geq m,$$

und somit

$$Q(h) \geq m\|h\|^2.$$

Andererseits für jedes $\varepsilon > 0$ gilt

$$\left|o\left(\|h\|^2\right)\right| < \varepsilon\|h\|^2,$$

vorausgesetzt, dass $\|h\|$ hinreichend klein ist. Wählen wir $\varepsilon < \frac{1}{2}m$, z.B. $\varepsilon = \frac{1}{3}m$, und erhalten, dass für alle $h \neq 0$ mit hinreichend kleiner Norm $\|h\|$ gilt

$$\frac{1}{2}Q(h) + o\left(\|h\|^2\right) \geq \left(\frac{1}{2}m - \varepsilon\right)\|h\|^2 > 0,$$

was zu beweisen war. Analog beweist man, dass im Fall $f''(x) < 0$ der Punkt x eine Maximumstelle ist. ∎

Beispiel. Betrachten wir wieder die Funktion $f(x,y) = \sin x \cos y$ in $\Omega = (-\pi,\pi) \times (-\pi,\pi)$. Wir wissen schon, dass diese Funktion vier kritische Punkte hat:

$$\left(\frac{\pi}{2},0\right), \left(-\frac{\pi}{2},0\right), \left(0,\frac{\pi}{2}\right), \left(0,-\frac{\pi}{2}\right).$$

Bereichen wir die Hesse-Matrix von f:

$$f''(x) = \begin{pmatrix} \partial_{xx}f & \partial_{xy}f \\ \partial_{yx}f & \partial_{yy}f \end{pmatrix} = \begin{pmatrix} -\sin x\cos y & -\cos x\sin y \\ -\cos x\sin y & -\sin x\cos y \end{pmatrix}.$$

Daraus folgt:

$$f''(\frac{\pi}{2},0)=\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}<0,$$

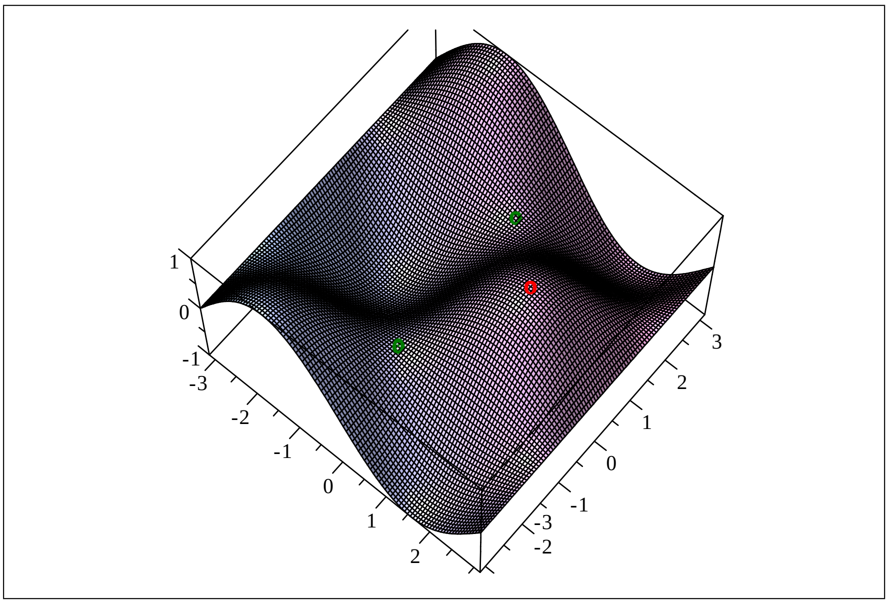
so dass $\left(\frac{\pi}{2},0\right)$ eine lokale Maximumstelle ist;

$$f''(-\frac{\pi}{2},0)=\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}>0,$$

so dass $\left(-\frac{\pi}{2},0\right)$ eine lokale Minimumstelle ist;

$$f''(0,\frac{\pi}{2})=\begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix} \text{ und } f''(0,-\frac{\pi}{2})=\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

sind indefinit, so dass weder $\left(0,\frac{\pi}{2}\right)$ noch $\left(0,-\frac{\pi}{2}\right)$ lokale Extremumstelle ist.



Die Funktion $f(x,y)=\sin x\cos y$

Bemerkung. Die Definitheit von einer symmetrischen Matrix $A=(a_{ij})_{i,j=1}^n$ lässt sich mit Hilfe von Eigenwerten wie folgt bestimmen. Da A symmetrisch ist, so sind alle Eigenwerte $\lambda_1,...,\lambda_n$ von A reell. Die quadratische Form $Q(u)$ von A lässt sich mit Hilfe von einer linearen Transformation $v=v(u)$ zur Diagonalform führen:

$$Q(u)=\lambda_1 v_1^2+...+\lambda_n v_n^2.$$

Somit erhalten wir die äquivalenten Bedingungen:

1. $A>0 \Leftrightarrow$ alle $\lambda_i>0$
2. $A\geq 0 \;\Leftrightarrow$ alle $\lambda_i\geq 0$
3. $A<0 \;\Leftrightarrow$ alle $\lambda_i<0$
4. $A\leq 0 \;\Leftrightarrow$ alle $\lambda_i\leq 0$

5. A ist indefinite $\Leftrightarrow$ es gibt i, j mit $\lambda_i > 0$ und $\lambda_j < 0$.

Im Fall $n = 2$ die Vorzeichen von λ_1 und λ_2 lassen sich leicht bestimmen ohne die Werte von λ_1 und λ_2 zu berechnen. Es ist bekannt, dass

$$\begin{aligned}\lambda_1 + \lambda_2 &= \operatorname{Spur} A = a_{11} + a_{22} \\ \lambda_1 \lambda_2 &= \det A = a_{11}a_{22} - a_{12}^2.\end{aligned}$$

Im Fall $\det A < 0$ folgt es, dass die Eigenwerte verschiedene Vorzeichen haben und somit ist die Matrix A indefinit. Im Fall $\det A > 0$ und $\operatorname{Spur} A > 0$ sind die Eigenwerte positiv und somit $A > 0$. Im Fall $\det A > 0$ und $\operatorname{Spur} A < 0$ sind die Eigenwerte negativ und somit $A < 0$.

Es gibt auch andere Methoden um die Definitheit von A zu bestimmen. Zum Beispiel, das Sylvester-Kriterium besagt folgendes: $A > 0$ genau dann, wenn alle führenden Hauptminoren von A positiv sind, d.h. für alle $1 \leq k \leq n$,

$$\det (a_{ij})_{i,j=1}^k > 0.$$

7.7 Satz von der impliziten Funktion

Betrachten wir das folgende Problem: gegeben sei eine Funktion $F(x, y)$ von zwei reellen Variablen, man bestimme y als Funktion von x aus der Gleichung $F(x, y) = 0$. Gibt es eine Funktion $f(x)$ so dass

$$F(x, y) = 0 \Leftrightarrow y = f(x), \tag{7.45}$$

so sagt man, dass die Funktion $f(x)$ durch die Gleichung $F(x, y) = 0$ *implizit* definiert wird. Häufig wird der Begriff "implizite Funktion" benutzt, was bedeutet nicht anderes als "implizit definierte Funktion".

Betrachten wir die Menge

$$M = \left\{(x, y) \in \mathbb{R}^2 : F(x, y) = 0\right\}.$$

Die Existenz von der impliziten Funktion f ist dann äquivalent zur Bedingung, dass M der Graph einer Funktion ist.

Beispiel. Die Menge von Punkten $(x, y) \in \mathbb{R}^2$ die die Gleichung

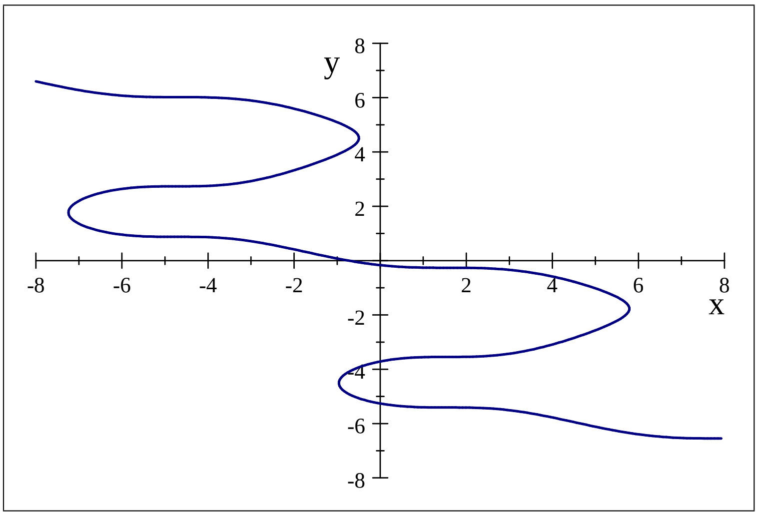
$$x^2 + y^2 = 1 \tag{7.46}$$

erfüllen, ist ein Kreis. Der Kreis ist kein Graph, aber besteht aus zwei Graphen von den Funktionen $y = \pm\sqrt{1 - x^2}$ auf $x \in [-1, 1]$. Diese Funktionen werden implizit von der Gleichung (7.46) im Bereichen $\{y \geq 0\}$ bzw $\{y \leq 0\}$ definiert.

Betrachten wir noch eine andere Gleichung:

$$x + \cos x + y + 5 \sin y = 0, \tag{7.47}$$

die sich nicht explizit lösen lässt.



Die Menge von Punkten (x, y), die (7.47) erfüllen

Die Menge von Punkten $(x, y) \in \mathbb{R}^2$ die (7.47) erfüllen, ist eine Kurve, die sich in mehreren Graphen teilen lässt, und zwar zwischen den Wendepunkten. Somit gibt es mehrere impliziten Funktionen, die von (7.47) definiert werden, abhängig von dem Definitionsbereich von (7.47).

Betrachten wir jetzt eine allgemeinere Situation when x ein Punkt in $\mathbb{R}^n$ ist und y ein Punkt in $\mathbb{R}^m$. Wir betrachten das Paar (x, y) als Element von $\mathbb{R}^{n+m}$ mit den Komponenten

$$(x_1, ..., x_n, y_1, ..., y_m).$$

Sei Ω eine offene Teilmenge von $\mathbb{R}^{n+m}$ wo eine Funktion $F : \Omega \to \mathbb{R}^m$ definiert ist. Wir möchten die Gleichung $F(x, y) = 0$ bezüglich y lösen und somit eine Funktion $y = f(x)$ erhalten. Diese Gleichung sieht ausführlich so aus:

$$\begin{cases} F_1(x_1, ..., x_n, y_1, ..., y_m) = 0 \\ F_2(x_1, ..., x_n, y_1, ..., y_m) = 0 \\ ... \\ F_m(x_1, ..., x_n, y_1, ..., y_m) = 0 \end{cases}$$

wobei $F_1, ..., F_m$ die Komponenten von F sind. Das ist ein System von m skalaren Gleichungen mit m Unbekannten $y_1, ..., y_m$ und mit n Parametern $x_1, ..., x_n$. Solches System kann sehr kompliziert sein, aber wir formulieren unterhalb die einfachen Bedingungen, die mindestens lokale Lösbarkeit garantieren.

Beispiel. Betrachten wir ein Beispiel, wo F eine lineare Abbildung bezüglich y ist, d.h.

$$F(x, y) = A(x)\, y + B(x),$$

wobei $A(x)$ eine $m \times m$ von x abhängige Matrix ist und $B(x) \in \mathbb{R}^m$. Die Gleichung $F(x, y) = 0$ ist ein lineares System

$$A(x)\, y + B(x) = 0,$$

das genau dann lösbar ist, wenn die Matrix $A(x)$ invertierbar ist. In diesem Fall erhalten wir

$$y = -A^{-1}(x)\, B(x).$$

Sei die Funktion $F : \Omega \to \mathbb{R}^m$ differenzierbar. Die Jacobi-Matrix von F lässt sich wie folgt darstellen

$$J_F = \begin{pmatrix} \partial_{x_1} F_1 & ... & \partial_{x_n} F_1 & \partial_{y_1} F_1 & ... & \partial_{y_m} F_1 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \partial_{x_1} F_m & ... & \partial_{x_n} F_m & \partial_{y_1} F_m & ... & \partial_{y_m} F_m \end{pmatrix} = \left(\partial_x F \mid \partial_y F\right), \qquad (7.48)$$

wobei $\partial_x F$ die Jacobi-Matrix von F bezüglich x ist und $\partial_y F$ die Jacobi-Matrix von F bezüglich y ist. Es ist klar, dass $\partial_x F$ eine $m \times n$ Matrix ist und $\partial_y F$ eine $m \times m$ Matrix. Insbesondere ist $\partial_y F$ eine quadratische Matrix.

Zum Beispiel, im Fall $F = A(x)\, y + B(x)$ haben wir $\partial_y F(x, y) = A(x)$. Somit ist in diesem Fall die Lösbarkeit der Gleichung $F(x, y) = 0$ äquivalent zur Invertierbarkeit der Matrix $\partial_y F$.

Definition. Let Ω eine offene Teilmenge von $\mathbb{R}^k$. Eine Funktion $f : \Omega \to \mathbb{R}^m$ heißt *l-fach stetig differenzierbar* falls alle partielle Ableitungen $D^\alpha f_j$ der Ordnung $|\alpha| \leq l$ von allen Komponenten f_j existieren und stetig in Ω sind. Die Menge von allen l-fach stetig differenzierbaren Funktionen $f : \Omega \to \mathbb{R}^m$ wird mit $C^l(\Omega, \mathbb{R}^m)$ bezeichnet.

Häufig schreibt man kurz $f \in C^l$ statt $f \in C^l(\Omega, \mathbb{R}^m)$ und sagt, dass f eine Funktion der Klasse C^l ist.

Hauptsatz 7.15 (Der Satz von der impliziten Funktion) *Seien Ω eine offene Menge in $\mathbb{R}^{n+m}$ und F eine Funktion der Klasse $C^l(\Omega, \mathbb{R}^m)$ mit $l \geq 1$. Gelten für einen Punkt $(a, b) \in \Omega$ die Bedingungen*

$$F(a, b) = 0 \;\; \textit{und} \;\; \partial_y F(a, b) \;\; \textit{ist invertierbar,} \qquad (7.49)$$

so existieren offene Mengen $U \subset \mathbb{R}^n$ und $V \subset \mathbb{R}^m$ mit $a \in U$, $b \in V$, $U \times V \subset \Omega$, und eine Funktion $f : U \to V$ mit

$$F(x, y) = 0 \;\Leftrightarrow\; y = f(x) \quad \textit{sofern } x \in U, y \in V. \qquad (7.50)$$

Darüber hinaus ist f von der Klasse $C^l(U, \mathbb{R}^m)$ und es gilt für alle $x \in U$ die Identität

$$f'(x) = -\left(\partial_y F\right)^{-1} \partial_x F(x, f(x)) \qquad (7.51)$$

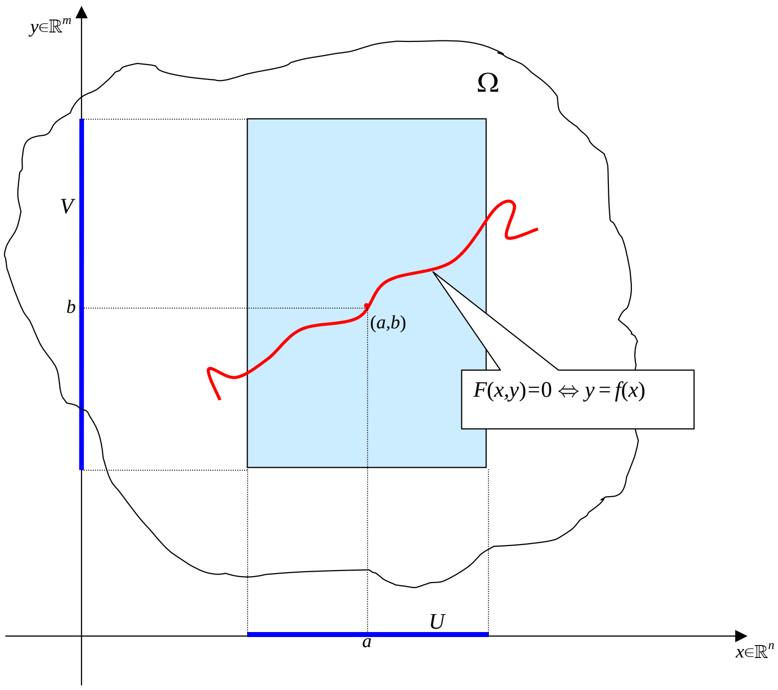


Bild zum Satz 7.15

Betrachten wir die Mengen

$$M = \{(x,y) \in \Omega : F(x,y) = 0\}$$

und

$$N = \{(x,y) \in \Omega : \det \partial_y F(x,y) = 0\}.$$

Die Bedingung (7.49) bedeutet, dass $(a,b) \in M \setminus N$. Die Existenz von f mit (7.50) bedeutet, dass die Menge M in der Nähe von (a,b) (nämlich in $U \times V$) der Graph einer Funktion ist. Somit lässt sich der Satz 7.15 wie folgt kurz umformulieren: in der Nähe von jedem Punkt $(a,b) \in M \setminus N$ ist die Menge M ein Graph. Oder noch kürzer: die Menge $M \setminus N$ ist lokal ein Graph. Das Wort "lokal" bedeutet genau "in der Nähe von jedem Punkt", und "ein Graph" bedeutet "der Graph einer Funktion von x".

Beispiel. Betrachten wir wieder die Funktion (7.47), d.h.

$$F(x,y) = x + \cos x + y + 5\sin y$$

mit $x, y \in \mathbb{R}$. Wir haben

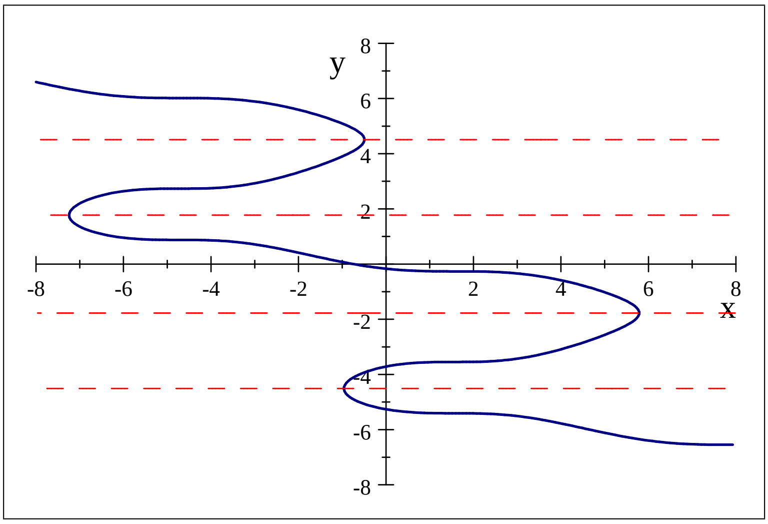
$$\partial_y F = 1 + 5\cos y.$$

Betrachten wir die Mengen M und N wie oberhalb, d.h.

$$M = \left\{(x,y) \in \mathbb{R}^2 : x + \cos x + y + 5\sin y = 0\right\}$$

und

$$\begin{aligned} N &= \left\{(x,y) \in \mathbb{R}^2 : \cos y = -\frac{1}{5}\right\} \\ &= \left\{(x,y) \in \mathbb{R}^2 : y = \pm\arccos\left(-\frac{1}{5}\right) + 2\pi k, \quad k \in \mathbb{Z}\right\}. \end{aligned}$$



Die Mengen M und N

Am obigen Bild M ist die Kurve und N ist die Vereinigung von waagerechten Geraden. Der Schnitt $M \cap N$ besteht aus allen Wendepunkten der Kurve M. Die Menge M außerhalb des Schnittes ist lokal der Graph einer Funktion $y = f(x)$.

Sei $y = f(x)$ eine implizit von $F(x, y) = 0$ gegebene Funktion. Es folgt aus (7.51), dass

$$f'(x) = -\frac{\partial_x F}{\partial_y F} = -\frac{1 - \sin x}{1 + 5\cos y} = -\frac{1 - \sin x}{1 + 5\cos f(x)}.$$

Obwohl die Funktion $f(x)$ nicht explizit bekannt ist, die Formel (7.51) ergibt die Ableitung $f'(x)$ explizit durch x und $f(x)$.

Bemerkung. Die Formel (7.51) lässt sich leicht gewinnen, vorausgesetzt, dass die implizite Funktion $y = f(x)$ existiert und differenzierbar ist. Dann erfüllen sie für alle $x \in U$ die Gleichung

$$F(x, f(x)) \equiv 0.$$

Ableiten von der Funktion $g(x) = F(x, f(x))$ ergibt nach der Kettenregel

$$g'(x) = F'(x, y) \begin{pmatrix} \mathrm{Id} \\ f'(x) \end{pmatrix} = (\partial_x F \mid \partial_y F) \begin{pmatrix} \mathrm{Id} \\ f'(x) \end{pmatrix} = \partial_x F + (\partial_y F) f'(x),$$

woraus die folgende Identität in U folgt

$$\partial_x F + (\partial_y F) f'(x) = 0,$$

und somit auch (7.51).

Dieser Argument gilt als Beweis von (7.51) nur dann, wenn die Differenzierbarkeit von f schon bekannt ist. Allerdings im Beweis des Satzes 7.15 werden wir die Differenzierbarkeit von f zusammen mit (7.51) erhalten, aber nicht zuvor.

7.8 Satz von der inversen Funktion

Hauptsatz 7.16 (Satz von der inversen Funktion) *Seien W eine offene Teilmenge von $\mathbb{R}^n$ und f eine Funktion der Klasse $C^l(W, \mathbb{R}^n)$ mit $l \geq 1$. Ist $f'(p)$ in einem*

Punkt $p \in W$ invertierbar, so existieren offene Teilmengen U und V von $\mathbb{R}^n$, so dass $p \in U \subset W$, $f(p) \in V$, und $f|_U$ eine Bijektion von U nach V ist; insbesondere ist die inverse Funktion $f^{-1} : V \to U$ wohldefiniert. Darüber hinaus liegt f^{-1} in der Klasse $C^l(V, \mathbb{R}^n)$ und es gilt

$$\left(f^{-1}\right)'(y) = \left(f'(x)\right)^{-1} \tag{7.52}$$

für alle $y \in V$ und $x = f^{-1}(y)$.

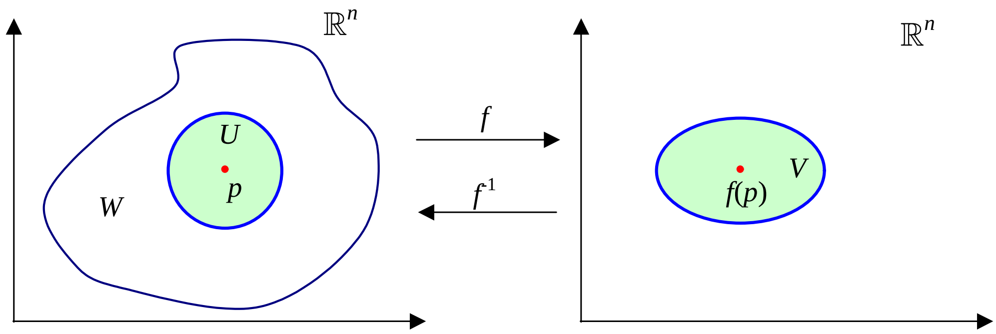


Bild zum Satz 7.16

Eine kurze Umformulierung des Satzes 7.16: in der Nähe von jedem Punkt p, wo die Matrix $f'(p)$ invertierbar ist, ist auch die Funktion f invertierbar. Betrachten wir die Menge

$$W_0 = \{p \in W : \det f'(p) = 0\}.$$

Dann in $W \setminus W_0$ ist f lokal invertierbar.

Bemerkung. Sei $f : I \to \mathbb{R}$ eine Funktion der Klasse C^1 auf einem Intervall $I \subset \mathbb{R}$ mit $f'(x) \neq 0$ für alle $x \in \mathbb{R}$. Zeigen wir, dass die inverse Funktion f^{-1} existiert auf ganzem Bild $J = f(I)$. Da f' stetig ist, so gilt nach dem Zwischenwertsatz, dass entweder $f'(x) > 0$ für alle $x \in I$ oder $f'(x) < 0$ für alle $x \in I$. Dann erhalten wir nach dem Satz von den inversen Funktion aus Analysis 1 (Satz 3.27): f^{-1} existiert auf J, ist differenzierbar, und

$$\left(f^{-1}\right)'(y) = \frac{1}{f'(x)}$$

für alle $y \in J$ und $x = f^{-1}(y)$. Der Satz 7.16 ergibt auch die Existenz von f^{-1} aber nur lokal, in einer Umgebung von einem Punkt p. Für globale Existenz von f^{-1} in Dimension $n \geq 2$ benötigt man zusätzliche Bedingung, die wir hier nicht angeben.

Beispiel. Betrachten wir die Funktion $f : \mathbb{R}^2 \to \mathbb{R}^2$ wie folgt:

$$f(x, y) = \left(x^2 - y^2, 2xy\right). \tag{7.53}$$

Die totale Ableitung

$$f'(x, y) = J_f(x, y) = \begin{pmatrix} \partial_x f_1 & \partial_y f_1 \\ \partial_x f_2 & \partial_y f_2 \end{pmatrix} = \begin{pmatrix} 2x & -2y \\ 2y & 2x \end{pmatrix}$$

ist genau dann invertierbar, wenn $\det f'(x,y) = 4(x^2+y^2) \neq 0$ d.h. für $(x,y) \neq 0$. Setzen wir $W = \mathbb{R}^2 \setminus \{0\}$ und erhalten nach dem Satz 7.16, dass f in W lokal invertierbar.

Mit Hilfe von der komplexen Variable $z = x + iy$ können wir die Funktion (7.53) wie folgt darstellen: $f(z) = z^2$. Jede komplexe Zahl $w \in W$ hat genau zwei Werte von $\sqrt{w}$, d.h. es gibt zwei Werte von z mit $f(z) = w$. Es folgt, dass das Bild $f(W)$ gleich W, aber die Funktion $f : W \to W$ nicht injektiv ist und somit nicht (global) invertierbar ist.

Beweis. Setzen wir

$$F(x,y) = y - f(x),$$

so dass die Gleichung $y = f(x)$ äquivalent zu

$$F(x,y) = 0$$

ist. Die Funktion $F(x,y)$ ist für alle $(x,y) \in \Omega := W \times \mathbb{R}^n$ definiert und nimmt die Werte in $\mathbb{R}^n$ an. Offensichtlich haben wir $F \in C^l(\Omega, \mathbb{R}^n)$.

Mit Hilfe von dem Satz von der impliziten Funktion (Satz 7.15) lösen wir die Gleichung $F(x,y) = 0$ bezüglich x und somit erhalten x als Funktion von y. Dafür brauchen wir die Invertierbarkeit von $\partial_x F$ in einem Punkt. Offensichtlich haben wir

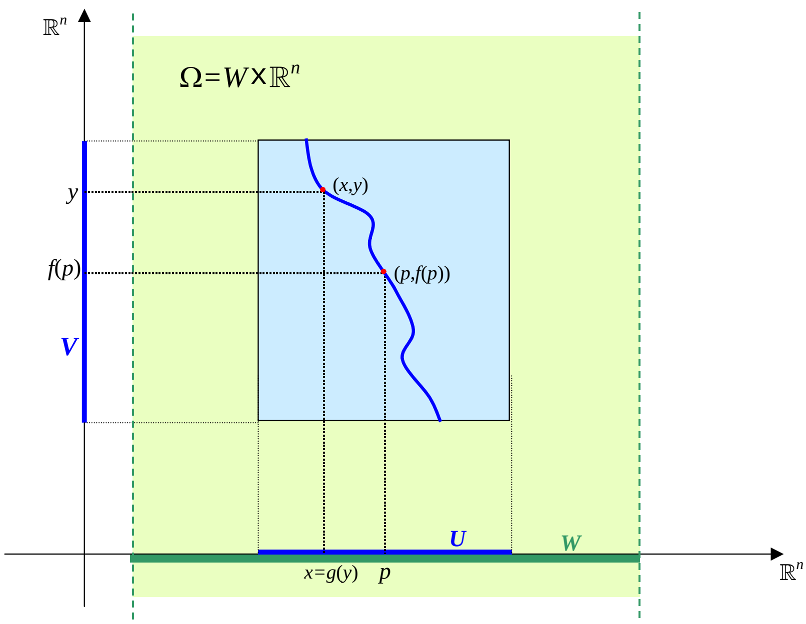
$$\partial_x F(x,y) = -f'(x).$$

Da $f'(p)$ invertierbar, so erhalten wir, dass $\partial_x F(p, f(p))$ invertierbar. Da auch $F(p, f(p)) = 0$, so erhalten wir nach dem Satz 7.15 folgendes: es gibt die Umgebungen $U \subset W$ von p und $V \subset \mathbb{R}^n$ von $f(p)$ und eine Funktion $g : V \to U$ von der Klasse C^l mit

$$F(x,y) = 0 \Leftrightarrow x = g(y) \text{ für alle } x \in U \text{ und } y \in V,$$

d.h.

$$y = f(x) \Leftrightarrow x = g(y) \text{ für alle } x \in U \text{ und } y \in V. \tag{7.54}$$



Funktion $g : V \to U$

Weiter bestimmen wir $f \circ g$ und $g \circ f$. Für jedes $y \in V$ setzen wir $x = g(y)$. Da $x \in U$, so erhalten aus (7.54) $f(g(y)) = f(x) = y$, d.h.

$$f \circ g = \mathrm{Id}_V\,. \tag{7.55}$$

Allerdings $g \circ f$ ist nicht unbedingt gleich Id_U . Für jedes $x \in U$ setzen wir $y = f(x)$. Es ist uns nicht gegeben, dass $f(x) \in V$. Im Fall $f(x) \in V$ gilt $y \in V$ und wir erhalten aus (7.54), dass

$$g(f(x)) = x \ \text{ für alle } x \in U \text{ mit } \ f(x) \in V.$$

Die zusätzliche Bedingung $f(x) \in V$ ist äquivalent zu $x \in f^{-1}(V)$ wobei f^{-1} hier die Urbildabbildung ist. Somit erhalten wir

$$g(f(x)) = x \ \text{ für alle } x \in U \cap f^{-1}(V) =: U_0. \tag{7.56}$$

Da f stetig ist, so ist die Menge $f^{-1}(V)$ offen und somit ist U_0 auch offen. Wir haben auch $p \in U_0$ da $p \in U$ und $f(p) \in V$.

Bemerken wir, dass das Bild von g in U_0 liegt da nach (7.55) $f(g(V)) = V$, woraus folgt $g(V) \subset f^{-1}(V)$ und somit $g(V) \subset f^{-1}(V) \cap U = U_0$. Es folgt, dass die Bedingung (7.54) auch für U_0 anstatt U gilt. Dann erhalten wir aus (7.56)

$$g \circ f = \mathrm{Id}_{U_0},$$

was zusammen mit (7.55) ergibt, dass $g : V \to U_0$ die inverse Funktion von $f : U_0 \to V$ ist. Es bleibt nur U_0 in U umbenennen.

Um (7.52) zu beweisen, leiten wir die Identität $f \circ g = \mathrm{Id}_V$ ab. Nach der Kettenregel gilt für jedes $y \in V$

$$\mathrm{id} = (f \circ g)' = f'(x)\, g'(y)\,,$$

wobei $x = g(y) = f^{-1}(y)$, was $g'(y) = f'(x)^{-1}$ ergibt. ■

7.9 Beweis von dem Satz von der impliziten Funktion

Beweis von dem Satz 7.15. Ohne Beschränkung der Allgemeinheit nehmen wir an, dass $a = 0$ in $\mathbb{R}^n$ und $b = 0$ in $\mathbb{R}^m$. In den Vektorräumen $\mathbb{R}^n$, $\mathbb{R}^m$, $\mathbb{R}^{n+m}$ wählen wir die ∞-Norm. Da F in 0 differenzierbar ist und $F(0) = 0$, so haben wir

$$F(h) = F'(0)\, h + \varphi(h)\,, \tag{7.57}$$

wobei die Funktion $\varphi : \Omega \to \mathbb{R}^m$ erfüllt

$$\varphi(h) = o(\|h\|) \ \text{ für } h \to 0.$$

We behaupten, dass

$$\varphi \in C^l(\Omega, \mathbb{R}^m)\,, \quad \varphi(0) = 0, \quad \varphi'(0) = 0. \tag{7.58}$$

In der Tat, es folgt aus (7.57), dass

$$\varphi(h) = F(h) - F'(0)\,h,$$

woraus die ersten zwei Eigenschaften in (7.58) folgen. Da

$$\varphi'(h) = F'(h) - F'(0),$$

so erhalten wir auch $\varphi'(0) = 0$.

Da $h \in \mathbb{R}^{n+m}$, so bezeichnen wir $h = (x, y)$ mit $x \in \mathbb{R}^n$ und $y \in \mathbb{R}^m$. Setzen wir

$$A = \partial_x F(0) \quad \text{und} \quad B = \partial_y F(0).$$

Dann gilt

$$F'(0) = (A \mid B)$$

und

$$F'(0)\,h = (A \mid B)\begin{pmatrix} x \\ y \end{pmatrix} = Ax + By.$$

Somit erhalten wir aus (7.57)

$$F(x, y) = Ax + By + \varphi(x, y). \tag{7.59}$$

Die Gleichung $F(x, y) = 0$ lässt sich wie folgt umschreiben:

$$Ax + By + \varphi(x, y) = 0.$$

Da die Matrix B invertierbar ist, so ist diese Gleichung äquivalent zu

$$y = -B^{-1}(Ax + \varphi(x, y)).$$

Setzen wir

$$G(x, y) = -B^{-1}(Ax + \varphi(x, y)) \tag{7.60}$$

und erhalten die folgende Äquivalenz:

$$F(x, y) = 0 \Leftrightarrow y = G(x, y).$$

Die weitere Idee von Beweis ist, dass die Abbildung $y \mapsto G(x, y)$ für jedes x in der Nähe von 0 in $\mathbb{R}^n$ eine Kontraktionsabbildung in der Nähe von 0 in $\mathbb{R}^m$ ist. Nach dem Fixpunktsatz von Banach können wir daraus beschließen, dass diese Abbildung einen Fixpunkt hat, dass heißt, die Gleichung $y = G(x, y)$ bezüglich y lösbar ist, was y als eine Funktion von x liefert. Um diese Idee rigoros zu machen, wir müssen den Definitionsbereich der Abbildung $y \mapsto G(x, y)$ bestimmen, wo diese Abbildung eine Selbstabbildung und auch eine Kontraktion ist. Darüber hinaus soll der Definitionsbereich ein vollständiger metrischer Raum sein.

Die Funktion $G(x, y)$ erfüllt die folgenden Eigenschaften:

$$G \in C^l(\Omega, \mathbb{R}^m), \quad G(0) = 0, \quad \partial_y G(0) = 0,$$

die trivial aus (7.58) und (7.60) folgen. Wählen wir hinreichend kleine positive Konstanten ε und δ aus den folgenden Bedingungen (i)-(iii).

(i) Da alle partielle Ableitungen $\partial_{y_i} G_j$ stetig sind und in 0 verschwinden, so gilt für hinreichend kleines $\varepsilon > 0$ die Eigenschaft:

$$\forall x \in \mathbb{R}^n, \forall y \in \mathbb{R}^m \text{ mit } \|x\| \leq \varepsilon,\ \|y\| \leq \varepsilon \text{ gilt } (x, y) \in \Omega \text{ und } |\partial_{y_i} G_j(x, y)| \leq \frac{1}{2m} \tag{7.61}$$

für alle i, j.

(ii) Da $\det \partial_y F$ stetig ist und $\det \partial_y F(0) \neq 0$ nach der Invertierbarkeit von $\partial_y F(0)$, so gilt für hinreichend kleines ε auch die folgende Eigenschaft:

$$\forall x \in \mathbb{R}^n\ \forall y \in \mathbb{R}^m \text{ mit } \|x\| \leq \varepsilon,\ \|y\| \leq \varepsilon \text{ ist } \partial_y F(x, y) \text{ invertierbar.} \tag{7.62}$$

(iii) Da die Funktion $x \mapsto G(x, 0)$ stetig ist und $G(0, 0) = 0$, so existiert für jedes $\varepsilon > 0$ (insbesondere für ε wie in (7.61) und (7.62)) ein $\delta \in (0, \varepsilon]$ mit

$$\forall x \in \mathbb{R}^m \text{ mit } \|x\| \leq \delta \text{ gilt } \|G(x, 0)\| < \frac{1}{2}\varepsilon. \tag{7.63}$$

Da $\delta \leq \varepsilon$ so liegt der Punkt $(x, 0)$ in Ω und somit ist $G(x, 0)$ wohldefiniert.

Bezeichnen wir mit U und V die folgenden Kugeln

$$U = \{x \in \mathbb{R}^n : \|x\| < \delta\},\ V = \{y \in \mathbb{R}^m : \|y\| < \varepsilon\}$$

und betrachten auch die abgeschlossenen Kugeln

$$\overline{U} = \{x \in \mathbb{R}^n : \|x\| \leq \delta\},\quad \overline{V} = \{y \in \mathbb{R}^m : \|y\| \leq \varepsilon\}.$$

Nach der Wahl von ε und δ gilt $\overline{U} \times \overline{V} \subset \Omega$. Der weitere Beweis wird schrittweise durchgeführt.

Schritt 1. *Zeigen wir, dass für alle $x \in \overline{U}$ und $y, y' \in \overline{V}$ gilt*

$$\|G(x, y) - G(x, y')\| \leq \frac{1}{2}\|y - y'\|. \tag{7.64}$$

Nach dem Mittelwertsatz erhalten wir für jede Komponente G_j of G und für ein $\xi \in [y, y']$, dass

$$\begin{aligned} |G_j(x, y) - G_j(x, y')| &= |\partial_y G_j(x, \xi)(y - y')| \\ &= \left| \sum_{i=1}^m \partial_{y_i} G_j(x, \xi)(y_i - y_i') \right| \leq \frac{1}{2m} m \|y - y'\|_\infty, \end{aligned}$$

woraus (7.64) folgt. Hier haben wir (7.61) benutzt, was die Konstante $\frac{1}{2m}$ in (7.61) erklärt.

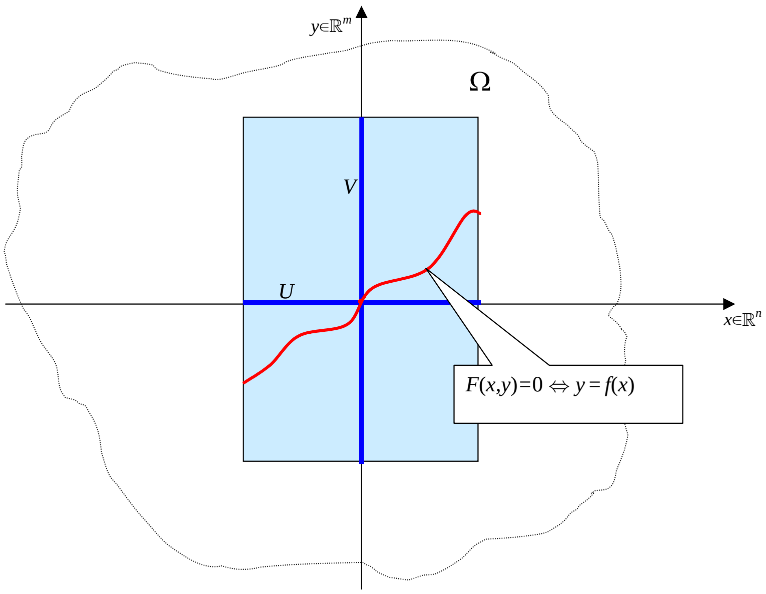
Schritt 2. *Zeigen wir, dass für alle $x \in \overline{U}$ und $y \in \overline{V}$ gilt $G(x, y) \in V$.*

We erhalten mit Hilfe von (7.64) und (7.63)

$$\|G(x, y)\| \leq \|G(x, y) - G(x, 0)\| + \|G(x, 0)\| < \frac{1}{2}\|y\| + \frac{1}{2}\varepsilon \leq \varepsilon.$$

Schritt 3. *Beweisen wir, dass es eine Funktion $f : \overline{U} \to \overline{V}$ gibt, so dass*

$$y = f(x) \Leftrightarrow F(x, y) = 0 \quad \textit{für alle } x \in \overline{U} \textit{ und } y \in \overline{V}. \tag{7.65}$$



Funktion f

Für jedes $x \in \overline{U}$ betrachten wir die Selbstabbildung von $\overline{V}$

$$\overline{V} \ni y \mapsto G(x,y) \in \overline{V},$$

die nach Schritt 2 wohldefiniert ist. Die Menge $\overline{V}$ ist eine abgeschlossene Teilmenge von $\mathbb{R}^m$ und somit ist ein vollständiger metrischer Raum (siehe Aufgaben). Nach (7.64) ist diese Abbildung eine Kontraktion. Nach dem Fixpunktsatz von Banach (Satz 6.13) gibt es genau einen Fixpunkt dieser Abbildung, den wir mit $f(x)$ bezeichnen. Diese Funktion ist für jedes $x \in \overline{U}$ definiert und nimmt die Werte in $\overline{V}$ an. Nach Definition von f erhalten wir, dass für $x \in \overline{U}$ und $y \in \overline{V}$ die Gleichung $G(x,y) = y$ äquivalent zu $y = f(x)$ ist. Da $G(x,y) = y$ äquivalent zu $F(x,y) = 0$ ist, so erhalten wir (7.65).

Schritt 4. *Beweisen wir: es gibt eine Konstante C mit*

$$\|f(x) - f(x')\| \leq C\|x - x'\| \text{ für alle } x, x' \in \overline{U}. \tag{7.66}$$

Die Funktionen, die (7.66) erfüllen, heißen *Lipschitz-stetig*. Es ist klar, dass eine Lipschitz-stetige Funktion auch stetig ist. Insbesondere ist die Funktion f stetig in $\overline{U}$.

Es folgt aus der Identität $f(x) = G(x, f(x))$, dass

$$\begin{aligned} \|f(x) - f(x')\| &= \|G(x, f(x)) - G(x', f(x'))\| \\ &\leq \|G(x, f(x)) - G(x', f(x))\| \\ &\quad + \|G(x', f(x)) - G(x', f(x'))\|. \end{aligned} \tag{7.67}$$

Nach (7.64) gilt

$$\|G(x', f(x)) - G(x', f(x'))\| \leq \frac{1}{2}\|f(x) - f(x')\|.$$

Um den ersten Glied in (7.67) abzuschätzen, benutzen wir für jede Komponente G_j von G den Mittelwertsatz: es gibt ein $\xi \in [x, x']$ mit

$$G_j(x, f(x)) - G_j(x', f(x)) = \partial_x G_j(\xi, f(x))(x - x') = \sum_{i=1}^{n} \partial_{x_i} G_j\left(\xi_j, f(x)\right)(x_i - x_i').$$

Da die Funktion $\partial_{x_i} G_j(x, y)$ stetig ist, so ist sie nach dem Extremwertsatz beschränkt auf $\overline{U} \times \overline{V}$, da diese Menge beschränkt und abgeschlossen ist. Also, es existiert eine Konstante M mit $|\partial_{x_i} G_j(x, y)| \leq M$ für alle $x \in \overline{U}$, $y \in \overline{V}$ alle i, j, woraus folgt

$$|G_j(x, f(x)) - G_j(x', f(x))| \leq Mn\|x - x'\|$$

und somit

$$\|G(x', f(x)) - G(x', f(x'))\| \leq Mn\|x - x'\|.$$

Es folgt aus (7.67) dass

$$\|f(x) - f(x')\| \leq Mn\|x - x'\| + \frac{1}{2}\|f(x) - f(x')\|,$$

was (7.66) mit $C = 2Mn$ ergibt.

Schritt 5. *Beweisen wir: die Funktion $f : U \to \mathbb{R}^m$ ist in U differenzierbar und*

$$f'(x) = -(\partial_y F)^{-1} \partial_x F(x, f(x)). \tag{7.68}$$

Fixieren wir ein $x_0 \in U$ und beweisen, dass f differenzierbar in x_0 ist. Setzen wir $y_0 = f(x_0)$. Nach der Differenzierbarkeit von F in (x_0, y_0) gilt

$$F(x, y) - F(x_0, y_0) = A(x - x_0) + B(y - y_0) + \varphi(x, y) \tag{7.69}$$

mit

$$A = \partial_x F(x_0, y_0), \quad B = \partial_y F(x_0, y_0)$$

und

$$\varphi(x, y) = o(\|x - x_0\| + \|y - y_0\|) \text{ für } x \to x_0 \text{ und } y \to y_0. \tag{7.70}$$

Wir benutzen (7.69) mit $x \in U$ und $y = f(x)$. Nach Definition von f haben wir

$$F(x, y) = F(x_0, y_0) = 0. \tag{7.71}$$

Nach (7.62) ist die Matrix B invertierbar. Es folgt aus (7.69) und (7.71), dass

$$f(x) - f(x_0) = y - y_0 = -B^{-1}A(x - x_0) - B^{-1}\varphi(x, f(x)). \tag{7.72}$$

Um die Differenzierbarkeit von f in x_0 daraus zu gewinnen, reicht es zu zeigen, dass

$$B^{-1}\varphi(x, f(x)) = o(\|x - x_0\|) \text{ für } x \to x_0.$$

Da die Norm der Matrix B^{-1} endlich ist, so reicht es zu beweisen, dass

$$\varphi(x, f(x)) = o(\|x - x_0\|) \text{ für } x \to x_0. \tag{7.73}$$

Nach (7.70) haben wir

$$\varphi(x, f(x)) = o(\|x - x_0\| + \|f(x) - f(x_0)\|) \text{ für } x \to x_0,$$

woraus (7.73) folgt, da nach (7.66)

$$\|f(x) - f(x_0)\| \leq C\|x - x_0\|.$$

Es folgt aus (7.72), dass

$$f'(x_0) = -B^{-1}A = -\partial_y F(x_0, f(x_0))^{-1} \partial_x F(x_0, f(x_0)).$$

Schritt 6. *Beweisen wir, dass* $f \in C^l(U, \mathbb{R}^m)$.

Induktionsanfang für $l = 1$. Da die Funktion f und alle partiellen Ableitungen von F stetig sind, so es folgt aus (7.68), dass auch $f'(x)$ stetig ist, woraus $f \in C^1$ folgt.

Induktionsschritt von $l - 1$ nach l. Beweisen wir, dass $F \in C^l$ ergibt $f \in C^l$. Nach der Induktionsvoraussetzung gilt $f \in C^{l-1}$. Da $\partial_x F$ und $\partial_y F$ von der Klasse C^{l-1} sind, so folgt es aus (7.68), dass $f' \in C^{l-1}$ (siehe Aufgaben), woraus $f \in C^l$ folgt. ∎

7.10 Flächen in $\mathbb{R}^n$

7.10.1 Linearer Unterraum

Es gibt die folgenden zwei Wege um einen Unterraum von $\mathbb{R}^n$ zu bestimmen.

Für jede lineare Abbildung $A : \mathbb{R}^m \to \mathbb{R}^n$ ist das Bild

$$\operatorname{im} A = \{Au : u \in \mathbb{R}^m\}$$

ein Unterraum von $\mathbb{R}^n$. Man sagt, dass der Unterraum $S = \operatorname{im} A$ durch die *parametrische* Gleichung $v = Au$ gegeben wird, was bedeutet, dass jeder Punkt $x \in S$ sich als $x = Au$ mit $u \in \mathbb{R}^m$ darstellen lässt. Der Punkt u heißt *Parameter.* Wir haben

$$\dim S = \dim \operatorname{im} A = \operatorname{rg} A,$$

da die linear unabhängigen Spalten von A eine Basis in $\operatorname{im} A$ liefern und der Rang $\operatorname{rg} A$ gleich die maximale Anzahl von linear unabhängigen Spalten (oder Zeilen) der Matrix A ist.

Für jede lineare Abbildung $B : \mathbb{R}^n \to \mathbb{R}^k$ ist der Kern

$$\ker B = \{x \in \mathbb{R}^n : Bx = 0\}$$

ein Unterraum von $\mathbb{R}^n$. Man sagt, dass der Unterraum $S = \ker B$ durch die Gleichung $Bx = 0$ gegeben wird. Nach dem Rangsatz gilt

$$\dim S = \dim \ker B = n - \operatorname{rg} B.$$

Es ist leicht zu sehen, dass jeder Unterraum von $\mathbb{R}^n$ sich in den beiden Formen darstellen lässt: sowohl parametrisch wie als Kern.

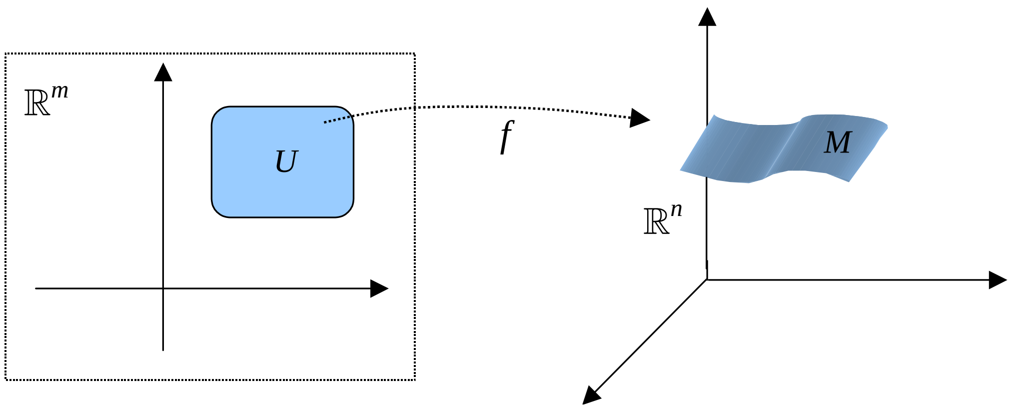
7.10.2 Parametrische Gleichung einer Fläche

Seien M eine Teilmenge von $\mathbb{R}^n$ und m eine ganze Zahl zwischen 1 und n.

Definition. Die Menge M heißt *m-dimensionale Fläche* falls es eine offene Menge $U \subset \mathbb{R}^m$ und eine Abbildung $f : U \to \mathbb{R}^n$ gibt, so dass gilt:

1. $M = f(U)$;
2. f ist injektiv;
3. f ist stetig differenzierbar;
4. f' ist nichtsingulär, d.h. $\operatorname{rg} f'(u) = m$ für alle $u \in U$.

Das Paar (U, f) heißt *Parametrisierung* von M. Das Dreifache (M, U, f) heißt *parametrisierte Fläche*. Die parametrisierte Fläche gehört zur Klasse C^l falls $f \in C^l$.



m-dimensionale parametrisierte Fläche (M, U, f)

Wir bezeichnen die Punkte in U mit u und nennen u Parameter. Nach Definition gilt für jeden Punkt $x \in M$ die eindeutige Darstellung $x = f(u)$ mit $u \in U$. Man sagt auch, dass M mit der *parametrischen Gleichung* $x = f(u)$ gegeben wird. Die Komponenten $u_1, ..., u_m$ von dem Parameter u heißen die *lokalen Koordinaten* von x. In ausführlicher Form sieht die parametrische Gleichung wie folgt aus:

$$\begin{cases} x_1 = f_1(u_1, .., u_m) \\ \dots \\ x_n = f_n(u_1, ..., u_m). \end{cases}$$

Die Ableitung $f' = \left(\frac{\partial f_i}{\partial u_j}\right)$ ist eine $n \times m$ Matrix (die Jacobi-Matrix) und die Bedingung $\operatorname{rg} f'(x) = m$ bedeutet, dass der Rang von $f'(u)$ an jeder Stelle $u \in U$ maximal ist.

Im Fall $m = 1$ nehmen wir an, dass U ein offenes Intervall ist. Die 1-dimensionale parametrisierte Fläche (M, U, f) ist offensichtlich eine parametrisierte Kurve.

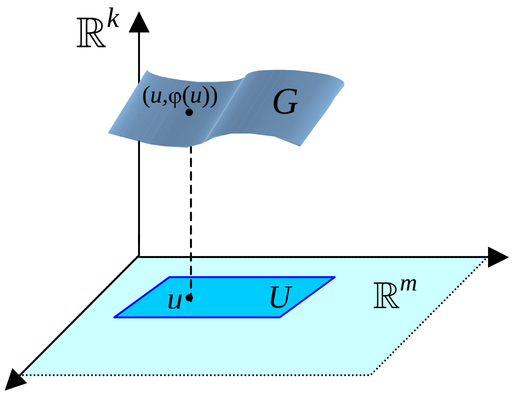
Beispiel. (*m-dimensionale Ebene*) Sei $A : \mathbb{R}^m \to \mathbb{R}^n$ eine lineare Abbildung oder, was äquivalent ist, eine $n \times m$ Matrix. Angenommen, dass $m \leq n$ und $\operatorname{rg} A = m$, wir erhalten, dass A eine Parametrisierung von $\operatorname{im} A$ ist. Somit ist der Unterraum $\operatorname{im} A$ von $\mathbb{R}^n$ eine m-dimensionale Fläche.

Betrachten wir jetzt eine *affine* Abbildung $f : \mathbb{R}^m \to \mathbb{R}^n$ d.h. $f(u) = Au + b$ wobei A eine Matrix wie oberhalb ist und $b \in \mathbb{R}^n$. Da $f' = A$ und somit $\operatorname{rg} f' = m$, so ist die Menge $M = f(\mathbb{R}^m)$ eine m-dimensionale Fläche. Offensichtlich ist M das Bild des Unterraums $\operatorname{im} A$ unter Translation $x \mapsto x + b$, d.h. M eine m-dimensionale *Ebene* ist.

Beispiel. *(Graphen)* Seien U eine offene Teilmenge von $\mathbb{R}^m$ und $\varphi : U \to \mathbb{R}^k$ eine stetig differenzierbare Funktion. Betrachten wir den Graph von φ

$$G = \{(u, \varphi(u)) \in \mathbb{R}^n : u \in U\}, \tag{7.74}$$

wobei das Paar $(u, \varphi(u))$ ein Element von $\mathbb{R}^n$ mit $n = m + k$ ist.



Der Graph G der Funktion φ

Lemma 7.17 *Der Graph G ist eine m-dimensionale Fläche.*

Beweis. Es folgt aus (7.74), dass $G = f(U)$ wobei die Abbildung $f : U \to \mathbb{R}^n$ wie folgt definiert wird:

$$f(u) = (u, \varphi(u)).$$

Die Abbildung f ist offensichtlich injektiv und stetig differenzierbar. Es gilt

$$f'(u) = \left(\begin{array}{|c|} \hline \mathrm{id} \\ \hline \varphi'(u) \\ \hline \end{array} \right), \tag{7.75}$$

wobei id die identische $m \times m$ Matrix ist und $\varphi'(u)$ eine $k \times m$ Matrix. Da die ersten m Zeilen der Matrix $f'(u)$ linear unabhängig sind, so erhalten wir $\operatorname{rg} f'(u) = m$. Somit ist (G, U, f) eine Parametrisierung und G ist eine m-dimensionale Fläche. ∎

Der nächste Satz wird ohne Beweis angegeben.

Satz. *Jede m-dimensionale Fläche ist lokal der Graph einer stetig differenzierbaren Funktion.*

Im Beweis dieses Satzes wird der Satz von der impliziten Funktion benutzt und die Bedingung, dass die Parametrisierung der Fläche nichtsingulär sein soll, spielt eine wichtige Role. Dafür betrachten wir ein Beispiel.

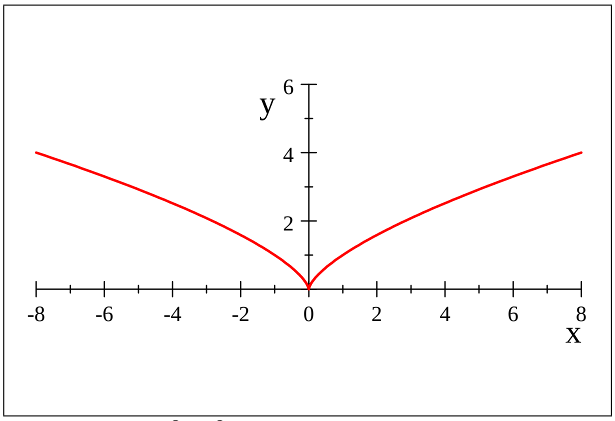
Beispiel. Die Abbildung

$$\begin{aligned} f &: \mathbb{R} \to \mathbb{R}^2 \\ f(u) &= \left(u^3, u^2\right) \end{aligned}$$

erfüllt die Bedingungen 1-3 von der Definition der Parametrisierung, aber die Bedingung 4 gilt nicht im Punkt $u = 0$ da

$$f'(0) = \left(3u^2, 2u\right)|_{u=0} = (0,0)$$

und somit $\operatorname{rg} f'(0) = 0 < 1$. Das Bild von f ist eine Kurve und sogar der Graph der Funktion $y = x^{2/3}$, aber diese Funktion ist in 0 nicht differenzierbar.



Due Kurve $f(u) = (u^3, u^2)$ ist der Graph der Funktion $y = x^{2/3}$

7.10.3 Tangentialebene

Definition. Seien U eine offene Teilmenge von $\mathbb{R}^m$ und $f : U \to \mathbb{R}^n$ eine differenzierbare Abbildung. Die *Tangentialabbildung* von f im Punkt $u \in U$ ist die folgende affine Abbildung

$$\begin{gathered} \tau : \mathbb{R}^m \to \mathbb{R}^n \\ \tau(h) = f(u) + f'(u)\,h. \end{gathered} \tag{7.76}$$

In anderer Wörter ist $\tau(h)$ das Taylor-Polynom der Ordnung 1 der Funktion f in u. Nach der Differenzierbarkeit von f gilt

$$f(u+h) = f(u) + f'(u)\,h + o(\|h\|) = \tau(h) + o(\|h\|)$$

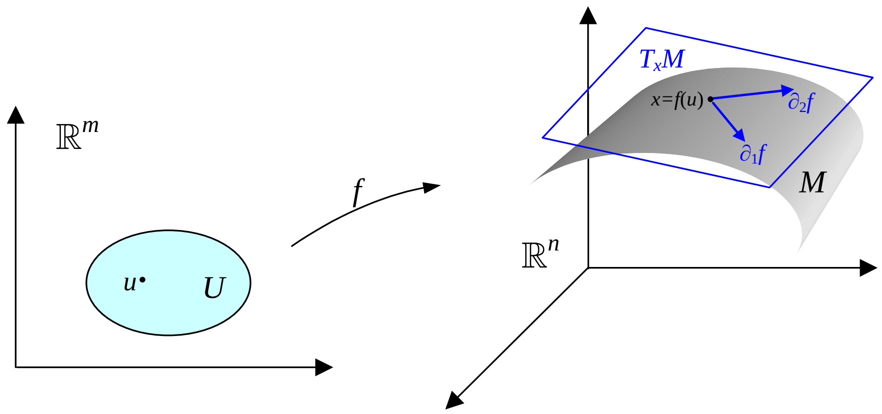
für $h \to 0$, so dass $\tau(h)$ eine affine Approximation von $f(u+h)$ für die kleinen Werte von $\|h\|$ ist.

Definition. Sei (M, U, f) eine m-dimensionale parametrisierte Fläche in $\mathbb{R}^n$. Die *Tangentialebene* T_xM an M im Punkt $x = f(u) \in M$ ist das Bild der Tangentialabbildung τ von f im Punkt u.

Es folgt aus (7.76), dass

$$T_xM = \operatorname{im}\tau = x + \operatorname{im} f'(u)\,.$$

Da $\operatorname{rg} f'(u) = m$, so ist T_xM eine m-dimensionale Ebene und $(\mathbb{R}^n, \tau)$ ist die Parametrisierung von T_xM. Die Ebene T_xM ist eine “gute” Approximation der Fläche M in der Nähe von x.



Tangentialebene T_xM

In Anwendungen ist es bequem die parametrische Gleichung (7.76) wie folgt darstellen:

$$\tau(h) = f(u) + h_1 \partial_1 f(u) + h_2 \partial_2 f(u) + ... + h_m \partial_m f(u),$$

wobei $h_1, ..., h_m$ die Komponenten von h sind und $\partial_1 f, ..., \partial_m f$ die Spalten von der Matrix f' sind, die eine Basis in $\operatorname{im} f'(u)$ liefern.

Beispiel. Betrachten wir die Parametrisierung des Kreises

$$\begin{aligned} f &: (0, 2\pi) \to \mathbb{R}^2 \\ f(u) &= (\cos u, \sin u) \end{aligned}$$

Da $f'(u) = (-\sin u, \cos u)$, so ist die Tangente (=1-dimensionale Tangentialebene)

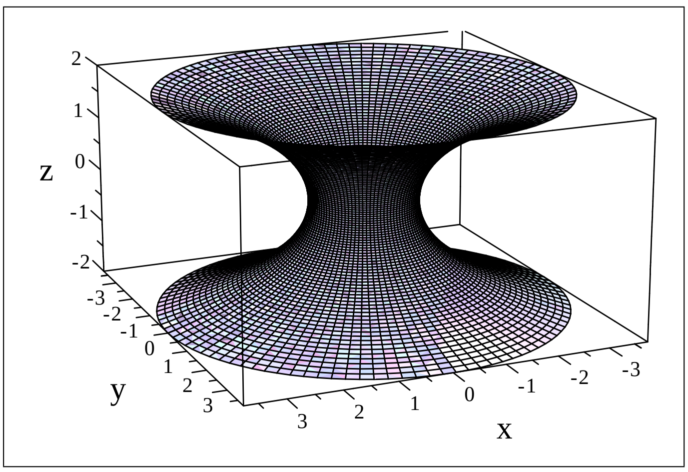
$$\begin{aligned} \tau(h) &= (\cos u, \sin u) + h(-\sin u, \cos u) \\ &= (\cos u - h \sin u, \sin u + h \cos u). \end{aligned}$$

Die Richtung $(-\sin u, \cos u)$ der Tangente ist offensichtlich orthogonal zum Radiusvektor $(\cos u, \sin u)$.

Beispiel. Betrachten wir die Parametrisierung

$$\begin{aligned} f &: U \to \mathbb{R}^3 \\ f(u, v) &= (\cosh u \cos v, \cosh u \sin v, u) \end{aligned}$$

wobei $U = \mathbb{R} \times (0, 2\pi)$. Die Fläche $M = \operatorname{im} f$ heißt *Katenoid.*



Katenoid

Berechnen wir die Ableitung:

$$f' = (\partial_u f, \partial_v f) = \begin{pmatrix} \sinh u \cos v & -\cosh u \sin v \\ \sinh u \sin v & \cosh u \cos v \\ 1 & 0 \end{pmatrix}$$

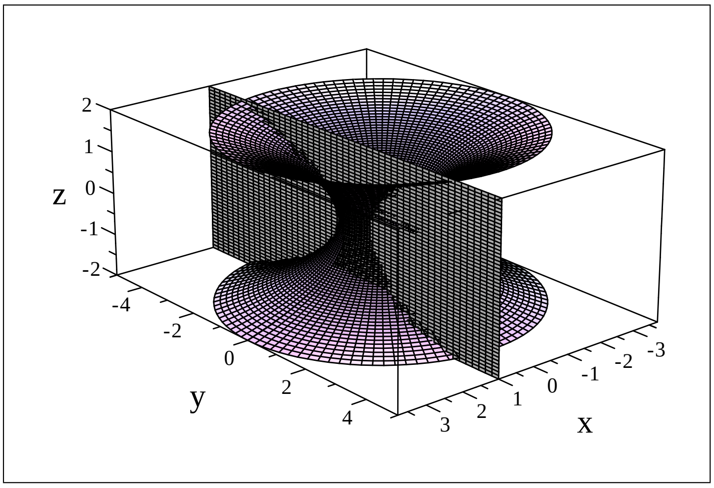
woraus folgt $\operatorname{rg} f' = 2$. Die Tangentialabbildung von f im Punkt (u, v) ist

$$\begin{aligned} \tau(h) &= f(u,v) + h_1 \partial_u f + h_2 \partial_v f \\ &= f(u,v) + h_1 (\sinh u \cos v, \sinh u \sin v, 1) + h_2 (-\cosh u \sin v, \cosh u \cos v, 0) . \end{aligned}$$

Zum Beispiel für $u = v = 0$ haben wir $x = f(0,0) = (1,0,0)$ und

$$\tau(h) = (1,0,0) + h_1 (0,0,1) + h_2 (0,1,0) .$$

Die Tangentialebene $T_x M = \operatorname{im} \tau$ geht durch $(1,0,0)$ und wird von den Vektoren $(0,0,1)$ und $(0,1,0)$ erzeugt.



Katenoid und Tangentialebene

7.10.4 Implizite Flächen

Satz 7.18 *Seien Ω eine offene Teilmenge von $\mathbb{R}^n$ und $F : \Omega \to \mathbb{R}^k$, $k < n$, eine stetig differenzierbare Funktion. Gilt $\operatorname{rg} F'(p) = k$ in einem Punkt $p \in \Omega$, so existiert eine offene Menge W mit $p \in W \subset \Omega$ so dass die Null-Niveaumenge*

$$M = \{x \in W : F(x) = 0\}$$

eine $(n-k)$-dimensionale Fläche ist.

Darüber hinaus gilt für jedes $x \in M$

$$T_x M = x + \ker F'(x) . \tag{7.77}$$

Die von der Gleichung $F(x) = 0$ gegebene Menge heißt *implizite Fläche*. Der Satz 7.18 besagt, dass jede implizite Fläche in der Nähe von jedem Punkt p mit nichtsingulärem $F'(p)$ eine Fläche ist.

Es folgt aus (7.77), dass $X \in T_x M$ äquivalent zu $X - x \in \ker F'(x)$ ist, d.h. zu

$$F'(x)(X - x) = 0.$$

Das ist die Gleichung der Tangentialebene $T_x M$.

Beweis. Setzen wir $m = n - k$. Die Matrix $F'(x)$ hat k Zeilen und n Spalten. Da $\operatorname{rg} F'(p) = k$, so existieren k linear unabhängige Spalten von $F'(p)$. Ohne Beschränkung der Allgemeinheit nehmen wir an, dass die letzten k Spalten, d.h. die Spalten $m+1, ..., n$ linear unabhängig sind. Setzen wir $x = (u, v)$ wobei

$$u = (x_1, ..., x_m) \in \mathbb{R}^m \quad \text{and} \quad v = (x_{m+1}, ..., x_n) \in \mathbb{R}^k$$

und schreiben die Gleichung $F(x) = 0$ in der Form $F(u, v) = 0$. Da

$$F'(p) = (\partial_u F(p) \quad | \quad \partial_v F(p))$$

und die Matrix $\partial_v F(p)$ aus den letzten k Spalten von der Matrix $F'(p)$ besteht, so erhalten wir nach der Voraussetzung, dass $\partial_v F(p)$ invertierbar ist.

Nach dem Satz von der impliziten Funktion existieren offene Mengen $U \subset \mathbb{R}^m$, $V \subset \mathbb{R}^k$ mit $p \in U \times V \subset \Omega$ und eine C^1-Funktion $\varphi : U \to V$ mit

$$F(u, v) = 0 \;\Leftrightarrow v = \varphi(u) \;\text{ für alle } u \in U \text{ und } v \in V. \tag{7.78}$$

Setzen wir $W = U \times V$ und bemerken, dass die Menge

$$\begin{aligned} M &= \{x \in W : F(x) = 0\} \\ &= \{(u, v) \in U \times V : F(u, v) = 0\} \\ &= \{(u, v) \in U \times V : v = \varphi(u)\} \end{aligned}$$

der Graph der Funktion $v = \varphi(u)$ in U ist. Nach Lemma 7.17 ist M eine m-dimensionale Fläche.

Diese Fläche hat die Parametrisierung $x = f(u)$, wobei $f(u) = (u, \varphi(u))$. Die Tangentialebene an M im Punkt $x = f(u)$ ist

$$T_x M = x + \operatorname{im} f'(u).$$

Es bleibt zu beweisen, dass

$$\operatorname{im} f'(u) = \ker F'(x).$$

Nach (7.78) gilt $F(u, \varphi(u)) = 0$ für alle $u \in U$, was äquivalent zu $F \circ f = 0$. Ableiten von dieser Identität ergibt

$$F'(x) f'(u) = 0,$$

woraus folgt

$$\operatorname{im} f'(u) \subset \ker F'(x). \tag{7.79}$$

Da $\operatorname{rg} F'(x) = k$ in der Nähe von p, so erhalten wir nach dem Rangsatz

$$\dim \ker F'(x) = n - \operatorname{rg} F'(x) = n - k = m.$$

Andererseits nach dem Beweis von Lemma 7.17 gilt

$$\dim \operatorname{im} f'(u) = \operatorname{rg} f'(u) = m.$$

Somit haben die Unterräume $\ker F'(x)$ und $\operatorname{im} f'(u)$ die gleiche Dimensionen, woraus folgt, dass sie übereinstimmen. ■

Beispiel. Betrachten wir die Gleichung $x_1^2 + ... + x_n^2 = 1$, die die Einheitssphäre bezüglich der 2-Norm bestimmt. Die Sphäre ist die Null-Niveaumenge der Funktion $F(x) = x_1^2 + ... + x_n^2 - 1$. Wir haben

$$F'(x) = (\partial_1 F, ..., \partial_n F) = 2(x_1, ..., x_n)$$

woraus die Gleichung der Tangentialebene folgt:

$$\sum_{i=1}^{n} x_i (X_i - x_i) = 0.$$

Insbesondere sehen wir, dass $X - x$ orthogonal zu x ist.

7.11 Gewöhnliche Differentialgleichungen

Eine gewöhnliche Differentialgleichung (abgekürzt mit *DGL*) hat die Form

$$F\left(x, y, y', ..., y^{(n)}\right) = 0, \tag{7.80}$$

wobei x eine unabhängige reelle Variable ist, $y = y(x)$ eine gesuchte Funktion (und $y^{(k)}$ die k-te Ableitung von y), und F eine gegebene Funktion von $n+2$ Variablen. Die Zahl n, die die maximale Ordnung der Ableitung $y^{(k)}$ in (7.80) ist, heißt die *Ordnung* von der DGL. Man sagt auch, dass (7.80) eine DGL n-ter Ordnung ist.

Die Gleichung (7.80) heißt "*differential*", weil sie die Ableitungen der gesuchten Funktion enthält. Eigentlich stellt die Gleichung (7.80) eine Beziehung zwischen verschiedenen Ableitungen von $y(x)$ dar. Die Differentialgleihung (7.80) heißt "*gewöhnlich*", weil die Ableitungen $y^{(k)}$ gewöhnlich sind, im Gegensatz zu partiellen Ableitungen. Es gibt auch die partiellen Differentialgleichungen, wo die gesuchte Funktion von mehreren Variablen abhängt und deshalb die partiellen Ableitungen benutzt werden müssen, aber in diesem Kurs betrachten wir nur gewöhnliche DGLen.

Gewöhnliche DGLen entstehen in verschiedenen Gebieten von Mathematik, als auch in Wissenschaften und Technik, da viele Naturgesetze mittels Differentialgleichungen formuliert werden können. In meisten Anwendungen braucht man eine Lösung $y(x)$ von (7.80) (mit gegebenen Randbedingungen) analytisch oder numerisch zu ermitteln. Es gibt bestimmte spezielle Typen von DGLen, die sich explizit analytisch lösen lassen. Andererseits, für ziemlich generellen Typen von GDLen kann man verschiedene Eigenschaften von Lösungen beweisen ohne sie explizit zu berechnen, z.B. die Existenz und Eindeutigkeit von Lösungen, Differenzierbarkeit, usw.

Betrachten wie die DGLen 1. Ordnung: $F(x, y, y') = 0$. Häufig kann diese Gleichung bezüglich y' gelöst werden, und man erhält die DGL in der *expliziten* Form:

$$y' = f(x, y), \tag{7.81}$$

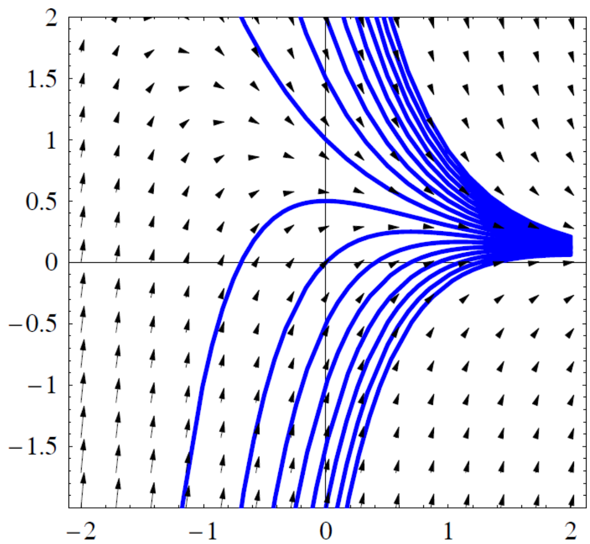
wobei $y = y(x)$ eine gesuchte reelle Funktion einer reellen Variablen x, und $f(x, y)$ eine gegebene Funktion von zwei reellen Variablen. Wir betrachten das Paar (x, y) als ein Punkt in $\mathbb{R}^2$. Der Definitionsbereich von f ist dann eine Teilmenge Ω von $\mathbb{R}^2$. Die Menge Ω heißt auch der Definitionsbereich von DGL (7.81).

Definition. Sei $y(x)$ eine reelle Funktion, die auf einem Intervall $I \subset \mathbb{R}$ definiert ist. Die Funktion $y(x)$ heißt (*spezielle*) *Lösung* von (7.81) genau dann, wenn

1. für jedes $x \in I$, der Punkt $(x, y(x))$ ein Element von Ω ist;
2. $y(x)$ an jeder Stelle $x \in I$ differenzierbar ist;
3. für jedes $x \in I$, die Gleichung $y'(x) = f(x, y(x))$ erfüllt ist.

Die Gesamtheit von allen speziellen Lösungen von (7.81) heißt die *allgemeine* Lösung.

Die Lösingen von (7.81) lassen sich graphisch wie folgt darstellen. Der Graph einer speziellen Lösung heißt *Integralkurve* der Gleichung. Offensichtlich ist jede Integral-Kurve im Definitionsbereich Ω enthalten. Dass die Lösing $y(x)$ die Gleichung $y' = f(x, y)$ erfüllt bedeutet, dass die Tangente zur Integralkurve an jeder Stelle (x, y) die Steigung $f(x, y)$ hat. Offensichtlich kann man die Steigung an jeder Stelle $(x, y) \in \Omega$ bestimmen ohne die DGL zu lösen. Jeder Stelle $(x, y) \in \Omega$ entspricht eine *Richtung*: eine Gerade durch (x, y) mit der Steigung $f(x, y)$. Die Gesamtheit von allen Richtungen heißt das *Richtungsfeld* der DGL. Es ist klar, dass die Tangente zu jeder Integralkurve an jeder Stelle ein Element des Richtungsfeldes ist. Lösen von (7.81) hat die folgende graphische Bedeutung: man verbindet die Elemente des Richtungsfeldes durch eine Integralkurve.



Integralkurven eines Richtungsfeldes

In der Regel lassen sich die allgemeinen DGLen nicht explizit analytisch lösen. Wir zeigen hier ein Beispiel wo man die allgemeine Lösung mit Hilfe von Integration bestimmen kann.

Beispiel. Hängt f von y nicht ab, so haben wir die Gleichung $y' = f(x)$. In diesem Fall ist y die Stammfunktion von f. Angenommen, dass f stetig ist, erhalten wir die allgemeine Lösung

$$y = \int f(x)\,dx = F(x) + C,$$

wobei $F(x)$ eine Stammfunktion von $f(x)$ ist und C eine beliebige Konstante ist.

Beispiel. Betrachten wir eine DGL

$$y' = y$$

und ermitteln erst alle positive Lösungen. Angenommen $y(x) > 0$ auf einem Intervall I, können wir durch y dividieren. Da

$$\frac{y'}{y} = (\ln y)',$$

erhalten wir die äquivalente Gleichung

$$(\ln y)' = 1.$$

Daraus folgt, dass

$$\ln y = \int dx = x + C,$$

also

$$y = e^C e^x = C_1 e^x,$$

wobei $C_1 = e^C$. Da $C \in \mathbb{R}$ beliebig reell ist, ist $C_1 = e^C$ beliebig positive. Daher sind alle positiven Lösungen $y(x)$ auf I wie folgt:

$$y = C_1 e^x, \quad C_1 > 0.$$

Angenommen, $y(x) < 0$ für alle $x \in I$, erhalten wir ebenso

$$\frac{y'}{y} = (\ln(-y))'$$

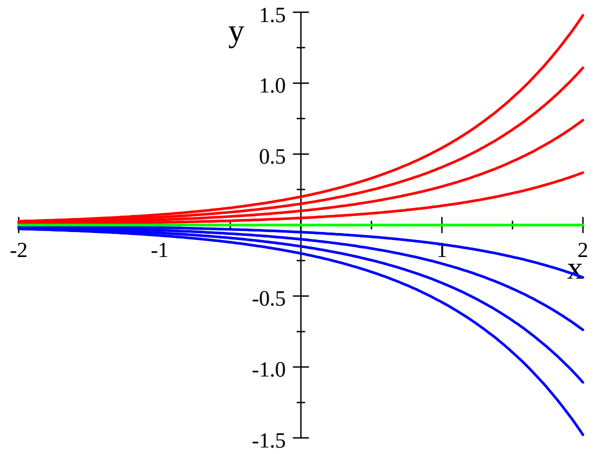
und

$$y = -C_1 e^x,$$

wobei $C_1 > 0$. So, jede Lösing $y(x)$, die immer entweder positive oder negative auf I bleibt, hat die Form

$$y(x) = Ce^x,$$

wobei $C > 0$ oder $C < 0$. Es ist klar, dass $C = 0$ auch eine Lösung $y \equiv 0$ ergibt. Die Integralkurven der Lösungen $y = Ce^x$ sind auf Fig. 7.11 gezeichnet worden. Bemerken wir, dass für jeden Punkt $(x_0, y_0) \in \mathbb{R}^2$ es eine Integralkurve gibt, die durch diesen Punkt geht. Man kann zeigen, dass die Gesamtheit von Lösungen $y = Ce^x$ mit $C \in \mathbb{R}$ die allgemeine Lösung von $y' = y$ ist.



Die Integralkurven von $y' = y$

Sei Ω eine offene Teilmenge von $\mathbb{R}^2$ und $f : \Omega \to \mathbb{R}$ eine gegebene Funktion. Betrachten wir jetzt das *Anfangswertproblem*

$$\begin{cases} y' = f(x, y), \\ y(x_0) = y_0, \end{cases} \tag{7.82}$$

wobei (x_0, y_0) ein gegebener Punkt in Ω ist. Eine Funktion $y : I \to \mathbb{R}^n$ heißt die Lösung von (7.82) falls y eine Lösung von $y' = f(x, y)$ auf einem Interval I ist, $x_0 \in I$ und $y(x_0) = y_0$. Der Graph der Lösung $y(x)$ ist eine Integralkurve, die durch den Punkt (x_0, y_0) geht.

Jetzt können wir den Hauptsatz von DGLen formulieren (ohne Beweis).

Hauptsatz 7.19 (Satz von Picard-Lindelöf) *Seien die Funktion f und die partielle Ableitung $\partial_y f$ stetig in Ω. Dann hat das Anfangswertproblem* (7.82) *eine Lösung für jedes $(x_0, y_0) \in \Omega$. Sind $y_1(x)$ und $y_2(x)$ zwei Lösungen von* (7.82), *dann gilt $y_1(x) = y_2(x)$ im gemeinsamen Definitionsbereich von y_1 und y_2.*

Der Beweis basiert auf dem Fixpunktsatz von Banach. Die Bedingung von Stetigkeit von $\partial_y f$ lässt sich durch eine schwachere Bedeutung von lokale Lipschitz-Stetigkeit von f in y ersetzt werden. Unter der einzigen Bedingung von der Stetigkeit von f gilt die Eindeutigkeit von Lösungen nicht.

Beispiel. Für die Differentialgleichung $y' = y$ erfüllt die Funktion $f(x, y) = y$ die Voraussetzungen des Satzes 7.19. Somit hat das Anfangswertproblem eine eindeutige Lösung für beliebige Anfangsbedingung. Da es für jeden Punkt $(x_0, y_0) \in \mathbb{R}^2$ eine Lösung $y = Ce^x$ gibt, die durch diesen Punkt geht, so ist $y = Ce^x$ die allgemeine Lösung von $y' = y$.

Beispiel. Betrachten wir die DGL

$$y' = \sqrt{|y|},$$

im Bereich $\mathbb{R} \times \mathbb{R}$. Diese DGL hat offensichtlich eine Lösung $y = 0$. In den Bereichen $y > 0$ und $y < 0$ lösen wir die DGL mit Hilfe von Trennung der Variablen. Im Bereich $y > 0$ erhalten wir

$$\int \frac{dy}{\sqrt{y}} = \int dx,$$

$$2\sqrt{y} = x + C,$$

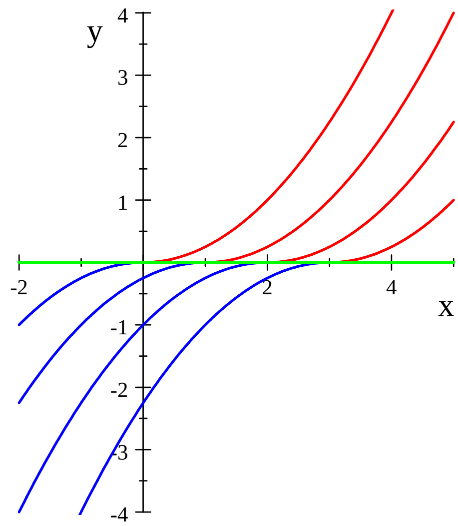
und

$$y = \frac{1}{4}(x+C)^2, \ \ x > -C, \tag{7.83}$$

wobei die Beschränkung $x > -C$ aus der vorherigen Gleichung kommt. Ebenso, im Bereich $y < 0$ erhalten wir

$$y = -\frac{1}{4}(x+C)^2, \ x < -C. \tag{7.84}$$

Die Integralkurven der Lösungen (7.83) und (7.84) werden unterhalb gezeichnet.



Die Integralkurven von $y' = \sqrt{|y|}$

Wir sehen, dass die Integralkurven aus den Bereichen $y > 0$ and $y < 0$ schneiden die Linie $y = 0$, die auch eine Lösung ist. Das ermöglicht Erstellung von mehreren Lösungen wie folgt: für jedes Paar von reellen Zahlen $a < b$, betrachten wir die Funktion

$$y(x) = \begin{cases} -\frac{1}{4}(x-a)^2, & x < a, \\ 0, & a \le x \le b, \\ \frac{1}{4}(x-b)^2, & x > b, \end{cases} \tag{7.85}$$

die durch Verkleben von drei anderen Lösungen gewonnen wird und offensichtlich eine Lösung für alle $x \in \mathbb{R}$ ist. Erlauben wir a zu sein $-\infty$ oder b zu sein $+\infty$, mit der offensichtlichen Bedeutung von (7.85) in diesen Fällen, stellt (7.85) die allgemeine Lösung von $y' = \sqrt{|y|}$ dar. Es ist jetzt klar, dass durch jeden Punkt $(x_0, y_0) \in \mathbb{R}^2$ unendliche viele Integralkurven der DGL gehen, und die Eindeutigkeit im Anfangswertproblem gilt nicht. Bemerken wir, dass die Funktion $f(x,y) = \sqrt{|y|}$ keine Ableitung $\partial_y f$ im Punkt $y = 0$ hat.