

Mathematik für Physiker III/Analysis III

Ausarbeitung einer Vorlesung
vom Wintersemester 2006/07

JOACHIM WEIDMANN

Fachbereich Informatik und Mathematik
der Universität Frankfurt

Stand 9. Februar 2007

Teil I

Gewöhnliche Differenzialgleichungen

1 Einführung

1.1 Grundbegriffe

Eine *gewöhnliche Differenzialgleichung der Ordnung n* ist eine Gleichung der Form

$$F(x, y, y', \dots, y^{(n)}) = 0, \quad \text{genauer } F(x, y(x), y'(x), \dots, y^{(n)}(x)) = 0$$

wobei x eine reelle Variable ist (häufig wird dies auch t für die Zeit sein) und $y = y(x)$ eine gesuchte Funktion mit Werten in \mathbb{R} oder \mathbb{C} . (Natürlich wird man nur dann sagen, die Gleichung ist von der Ordnung n , wenn F tatsächlich von $y^{(n)}$ abhängt. Auch Funktionen mit Werten in \mathbb{R}^m oder \mathbb{C}^m sind möglich, $y = (y_1, \dots, y_m)$; man schreibt dann ausführlicher

$$F(x, y_1, \dots, y_m, y'_1, \dots, y'_m, \dots, y_1^{(n)}, \dots, y_m^{(n)}) = 0,$$

wobei dann i.allg. auch F vektorwertig sein muß, d. h. es handelt sich dann um ein System von Gleichungen. Die Bezeichnung „gewöhnliche Differenzialgleichung“ bezieht sich auf die Tatsache, dass nur „gewöhnliche“ Ableitungen nach der einen Variablen x vorkommen (im Gegensatz zu „partiellen“ Ableitungen in „partiellen Differenzialgleichungen“).

Beispiel 1.1 Eine grundlegende Beziehung in der Physik ist:

$$\begin{aligned} \text{Kraft} - \text{Masse} \times \text{Beschleunigung} &= 0 \\ k(t) - m\ddot{x} &= 0. \end{aligned}$$

Also ein Spezialfall einer Differenzialgleichung der Ordnung 2. □

Die Funktion F ist auf einer Teilmenge

$$D \subset \mathbb{R}^{n+2} \quad \text{oder} \quad \mathbb{R} \times \mathbb{C}^{n+1} \quad \text{bzw.} \quad D \subset \mathbb{R}^{1+m(n+1)} \quad \text{oder} \quad \mathbb{R} \times \mathbb{C}^{m(n+1)}$$

erklärt. Eine *Lösung* der Differenzialgleichung (in dem hier zur Diskussion stehenden elementaren Sinne) ist eine n mal stetig differenzierbare Funktion $y : I \rightarrow \mathbb{R}(\mathbb{C}, \mathbb{R}^m)$ oder \mathbb{C}^m auf einem Intervall I mit

$$(x, \eta(x), \dots, \eta^{(n)}(x)) \in D \quad \text{und} \quad F(x, \eta(x), \dots, \eta^{(n)}(x)) = 0 \quad \text{für alle } x \in I.$$

Bemerkung 1.2 *Nicht jede Differenzialgleichung ist lösbar. So besagt z. B. die Gleichung $|y'| = -|y - x|$ einerseits, dass $y' \equiv 0$ sein muß und andererseits $y(x) = x$, ein offensichtlicher Widerspruch.*

Aus rein theoretischer Sicht genügt es, Differenzialgleichungen der Ordnung 1 zu betrachten, d. h. Gleichungen (bzw. Systeme) der Form

$$F(x, y, y') = 0 \quad \text{bzw.} \quad F_j(x, y_1, \dots, y_m, y'_1, \dots, y'_m) = 0 \quad \text{für } j = 1, \dots, k.$$

Eine Gleichung (bzw. ein System) der Ordnung n kann nämlich ganz einfach in ein System von Gleichungen der Ordnung 1 übergeführt werden, indem man definiert:

$$Y = (y_1, \dots, y_n) := (y, y', \dots, y^{(n-1)})$$

(wobei hier i. allg. $y_j \in \mathbb{R}^m$ oder \mathbb{C}^m ist), und für diese vektorwertige Funktion die Gleichung

$$\tilde{F}(x, Y, Y') = 0$$

betrachtet mit

$$\tilde{F}(x, Y, Y') = \begin{pmatrix} y'_1 - y_2 \\ y'_2 - y_3 \\ \vdots \\ y'_{n-1} - y_n \\ F(x, y_1, \dots, y_n, y'_1, \dots, y'_n) \end{pmatrix}.$$

Man sieht sofort, dass $y = y_1$ genau dann die ursprüngliche Gleichung $F(x, y, y', \dots, y^{(n)}) = 0$ löst, wenn Y die Gleichung $\tilde{F}(x, Y, Y') = 0$ löst.

Soweit es theoretische Überlegungen betrifft, beschränken wir uns deshalb auf Differenzialgleichungen erster Ordnung.

Gleichungen der bisher betrachteten allgemeinen Form heißen *implizite Differenzialgleichungen*. Sie sind nur selten in dieser Form lösbar. Wir gehen deshalb in der Regel davon aus, dass die Gleichung $F(x, y, y') = 0$ nach y' aufgelöst werden kann, also in der Form

$$y' = f(x, y) \quad \text{explizite Differenzialgleichung}$$

geschrieben werden kann (bzw. $y'_j = f_j(x, y_1, \dots, y_n)$ für $j = 1, \dots, n$ im Fall eines Systems).

Gleichungen dieser Art können (im reellen Fall und für $m = 1$) leicht geometrisch gedeutet werden: Jedem Punkt $(x, y) \in D \subset \mathbb{R}^2$ ist durch $y' = f(x, y)$ eine Steigung, also eine Richtung in der (x, y) -Ebene, zugeordnet. Die Gesamtheit dieser Richtungen wird als *Richtungsfeld* bezeichnet.

Die Lösungen der Differenzialgleichung sind genau die Funktionen $y = y(x)$, für die in jedem Punkt des Graphen die Tangentialrichtung mit der durch $y' = f(x, y)$ vorgegebenen Richtung übereinstimmt. Häufig kann man so leicht qualitative Aussagen über die Lösungen finden. (Im Fall der impliziten Differenzialgleichung gibt es dagegen u. U. in einem Punkt mehrere „zulässige“ Richtungen: so ist z. B. für die Gleichung $|y'| - |y| = 0$ sowohl $y' = y$ wie auch $y' = -y$ möglich.)

1.2 Einfachste Differenzialgleichungen und Problemstellung

Beispiel 1.3 Der einfachste Fall einer Differenzialgleichung (üblicherweise gar nicht als solche bezeichnet) ist

$$y' = f(x) \quad (\text{die rechte Seite ist von } y \text{ unabhängig}).$$

Für ein stetiges f bedeutet dies gerade, dass y eine Stammfunktion von f ist; Offenbar ist für jedes x_0 aus dem Definitionsbereich von f und jedes $y_0 \in \mathbb{R}$ oder \mathbb{C} durch

$$y(x) := y_0 + \int_{x_0}^x f(t) dt$$

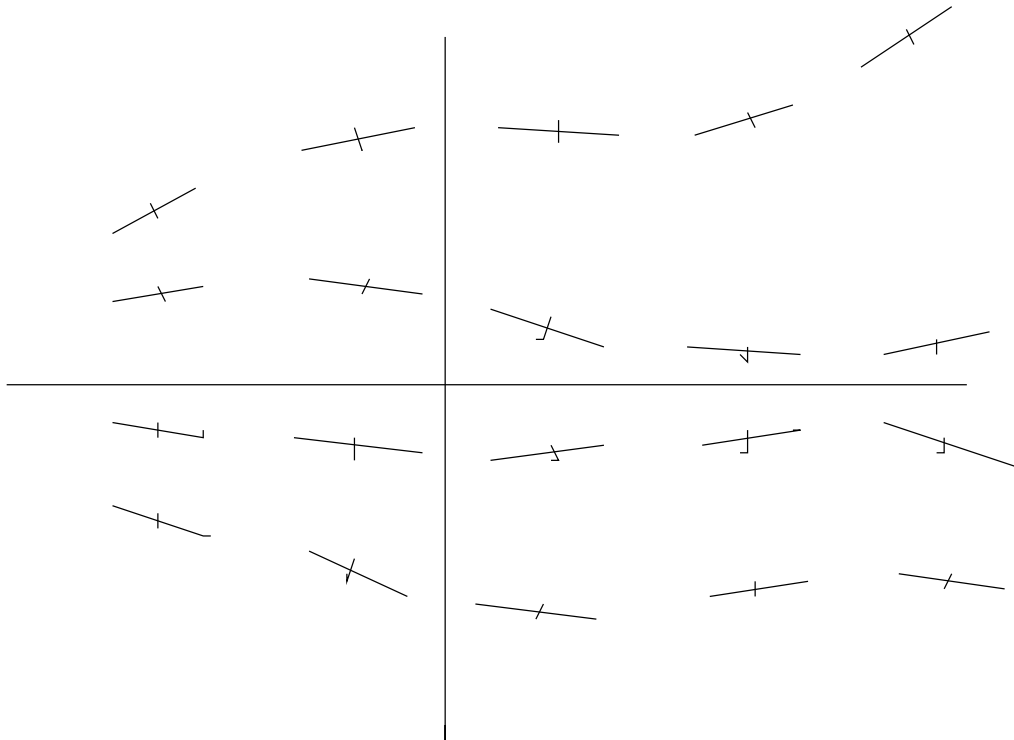


Abbildung 1: Richtungsfeld

eine Lösung mit dem *Anfangswert* $y(x_0) = y_0$ am *Anfangspunkt* x_0 gegeben. Diese Lösung ist durch den Anfangswert eindeutig bestimmt, denn für die Differenz zweier Lösungen der Gleichung zum gleichen Anfangswert gilt $w'(x) = 0$ für alle x und $w(x_0) = 0$, also $w \equiv 0$. (Dies ist im wesentlichen die Methode, wie wir auch später immer wieder die Eindeutigkeit zeigen werden.)

Betrachten wir die Lösungen y_{x_0, y_0} zum Anfangswert y_0 beim Anfangspunkt x_0 und y_{x_1, y_1} zum Anfangswert y_1 beim Anfangspunkt x_1 , so gilt offenbar

$$\begin{aligned} |y_{x_0, y_0}(x) - y_{x_1, y_1}(x)| &= \left| y_0 - y_1 + \int_{x_0}^x f(t) dt - \int_{x_1}^x f(t) dt \right| \\ &\leq |y_0 - y_1| + \left| \int_{x_0}^{x_1} f(t) dt \right|. \end{aligned}$$

Werden Anfangspunkt und Anfangswert wenig geändert, so ändert sich also die Lösung (global) wenig; sie *hängt global stetig vom Anfangspunkt und vom Anfangswert ab*.

„Lokal“ erhält man sogar eine *stetige Abhängigkeit von der rechten Seite* der Gleichung: Für verschiedene rechte Seiten f_0, f_1 gilt nämlich bei Anfangspunkten x_0, x_1 und Anfangswerten y_0, y_1 , wenn wir die entsprechenden Lösungen mit $y_0(\cdot)$ bzw. $y_1(\cdot)$ bezeichnen:

$$|y_0(x) - y_1(x)| \leq |y_0 - y_1| + \left| \int_{x_0}^{x_1} f_0(t) dt \right| + \left| \int_{x_1}^x |f_0(t) - f_1(t)| dt \right|.$$

(In dieser Abschätzung kann man die Rollen von $\{x_0, y_0, f_0\}$ und $\{x_1, y_1, f_1\}$ vertauschen, was u. U. die Abschätzung verbessern kann.) Solche Eigenschaften gelten auch für allgemeinere Gleichungen und sind für die Anwendungen von allergrößter Wichtigkeit. \square

1:17.10.06

Beispiel 1.4 Eine gleichmäßig beschleunigte Bewegung (z. B. der freie Fall) wird beschrieben durch die Differenzialgleichung (wobei hier die abhängige Variable x statt y und die Zeit t statt der Variablen x gewählt wird)

$$x'' = b \quad (\text{in der Physik üblicherweise : } \ddot{x} = b)$$

wobei b die (konstante) Beschleunigung ist (z. B. ist die Erdbeschleunigung nahe der Erdoberfläche $b = g \sim -9,81 \frac{m}{sec^2}$). Mit $x_1 = x, x_2 = x'$ geht diese Gleichung über in das System

$$\begin{aligned} x_1' &= x_2 & (x_1(t) &= \text{Ort zur Zeit } t), \\ x_2' &= b & (x_2(t) &= \text{Geschwindigkeit zur Zeit } t). \end{aligned}$$

Aus der zweiten Gleichung folgt (vgl. Beispiel 1.3) $x_2(t) = v_0 + bt$, wobei v_0 die Geschwindigkeit zur Zeit 0 ist (Anfangsgeschwindigkeit). Einsetzen in die erste Gleichung ergibt $x_1' = v_0 + bt$, also (wiederum nach Beispiel 1.3)

$$x_1(t) = x_0 + v_0 t + \frac{1}{2} b t^2,$$

wobei x_0 der Ort zur Zeit 0 ist. □

(Natürlich kann man auch direkt, d. h. ohne Übergang zu einem System, die Gleichung integrieren: Aus $x'' = b$ folgt $x'(t) = v_0 + bt$, und damit $x(t) = x_0 + v_0 t + \frac{1}{2} b t^2$.)

Man sieht dieser Lösung sofort an, dass für jedes kompakte Intervall $[t_1, t_2] \subset \mathbb{R}$ gilt:

Aus $x_{0,n} \rightarrow x_0, v_{0,n} \rightarrow v_0, b_n \rightarrow b$ folgt für die entsprechenden Lösungen $x_n(t) \rightarrow x(t)$ gleichmäßig in $[t_1, t_2]$.

Diese gleichmäßige Konvergenz gilt aber nicht auf ganz \mathbb{R} ; auch wenn x_0 und x_1, v_0 und v_1 bzw. b_1 und b nahe beisammen liegen, weichen die zugehörigen Lösungen für große $|t|$ i. a. weit voneinander ab. Man sagt, die Differenzialgleichung ist **nicht stabil**. □

Aus den bisherigen Überlegungen lassen sich leicht die Wünsche ablesen, die man an eine Theorie der Differenzialgleichungen haben wird. Wir beschränken uns dabei wie oben schon angedeutet auf Gleichungen erster Ordnung:

Ziele einer Theorie gewöhnlicher Differenzialgleichungen

1. *Überblick* über alle Lösungen einer Gleichung, – wichtiger ist aber:
2. Für beliebige Anfangswerte die *Existenz* und *Eindeutigkeit*:
 - (α) existiert mindestens eine Lösung (*Existenz*),
 - (β) existiert höchstens eine Lösung (*Eindeutigkeit*); zusammen mit (α) existiert dann also genau eine Lösung.
3. Die Lösung sollte *stetig vom Anfangswert* und eventuell auch *vom Anfangspunkt* und den in der Gleichung enthaltenen Funktionen (bzw. Koeffizienten) *abhängen*. Allgemeiner: *Stetige Abhängigkeit von den Daten*. – Dies wird i. allg. nur lokal (d. h. in der Nähe von x_0) gelten, während weit von x_0 entfernt die Abweichung sehr groß wird, auch wenn die Anfangswerte nahe beisammen liegen (in jedem Fall sollte die Abweichung klein werden, wenn die Anfangswerte hinreichend nahe zusammen liegen).

4. Für wichtige Typen von Gleichungen möchte man Verfahren zur *expliziten Bestimmung der Lösungen* haben, eventuell wenigstens zur näherungsweise Bestimmung (numerisch).
5. In Fällen, wo eine explizite Lösung nicht möglich oder zu schwierig ist, möchte man wenigstens möglichst gute *qualitative* und/oder *quantitative Aussagen* über die Lösung haben, z. B. ihr asymptotisches Verhalten für große x -Werte.

1.3 Übungsaufgaben

1. a) Man skizziere die Kurvenschar

$$A = \{y(x) = Ce^x : C \in \mathbb{R}\}.$$

- b) Man bestimme eine Differenzialgleichung für die Schar B der Kurven, die in jedem Punkt (x, y) die durch den Punkt verlaufende Kurve der Schar A senkrecht schneidet.
- c) Man skizziere die Schar B .
- d) Man bestimme die Schar B als Gesamtheit der Lösungen der in b) gefundenen Differenzialgleichung.

2. Die Differenzialgleichung

$$(DG) \quad y' + f(x)y = g(x)$$

habe die Lösungen $y_1(x) = x^2$ und $y_2(x) = x^{-2}$.

- a) Ist z_1 Lösung der Differenzialgleichung (DG) , so ist z_2 genau dann eine Lösung, wenn $y := z_1 - z_2$ die Gleichung

$$y' + f(x)y = 0$$

löst.

- b) Man bestimme f und g aus (DG) .
- c) Man bestimme alle Lösungen der Differenzialgleichung (DG) .

3. a) Für welche $a \in \mathbb{R}$ hat die Gleichung

$$y'' - 2xy' + 2ay = 0$$

ein Polynom vom Grad k ($k \in \mathbb{N}_0$) als Lösung?

- b) Für $k = 3$ bestimme man das Polynom mit höchstem Koeffizienten = 1, das diese Gleichung löst.

2 Elementar lösbare Differenzialgleichungen

In diesem Kapitel werden für einige Klassen gewöhnlicher Differenzialgleichungen erster Ordnung Verfahren zur expliziten Lösung angegeben. Existenz- und Eindeutigkeitsfragen für die entsprechenden Anfangswertprobleme werden dabei in der Regel nebenbei mitbeantwortet.

2.1 Differenzialgleichungen mit getrennten Variablen

Eine Differenzialgleichung mit *getrennten Variablen* hat die Gestalt

$$y' = f(x)g(y).$$

Sie heißt so, weil sie für $g \neq 0$ in der Form $y'/g(y) = f(x)$ geschrieben werden kann; hier stehen die Variablen x und y auf verschiedenen Seiten der Gleichung. (Den trivialen Fall $g \equiv 1$ haben wir bereits in Beispiel 1.3 behandelt.)

Satz 2.1 *Seien $f : I_1 \rightarrow \mathbb{R}, g : I_2 \rightarrow \mathbb{R}$ stetig, $x_0 \in I_1, y_0 \in I_2, g(y_0) \neq 0$. Dann gibt es eine Umgebung von x_0 , in der das Anfangswertproblem*

$$\text{AWP } y' = f(x)g(y), \quad y(x_0) = y_0$$

genau eine Lösung besitzt. Wählen wir Stammfunktionen

$$G \text{ von } \frac{1}{g} \text{ und } F \text{ von } f \text{ mit } G(y_0) = F(x_0)$$

so erhalten wir die Lösung durch Auflösen der Gleichung

$$G(y) = F(x)$$

nach y . Diese Auflösung ist stets möglich. (Das entsprechende gilt auch, wenn f komplexwertig ist, G eine Stammfunktion von $1/g$ im Sinne der Funktionentheorie.)

Beweis. Existenz: Wegen $g(y_0) \neq 0$ ist $g(y) \neq 0$ für y nahe y_0 , d. h. G ist streng monoton nahe y_0 und somit ist G dort umkehrbar. Wegen $F(x_0) = G(y_0)$ existiert also eine Funktion $y(\cdot)$, die auf einer Umgebung von x_0 definiert ist mit

$$y(x) = G^{-1}(F(x)) \quad \text{und} \quad y(x_0) = y_0.$$

Nach der Kettenregel und der Formel für die Ableitung einer Umkehrfunktion gilt also

$$\begin{aligned} y'(x) &= \frac{d}{dx} G^{-1}(F(x)) = \frac{1}{G'(G^{-1}(F(x)))} F'(x) = \frac{1}{G'(y(x))} F'(x) \\ &= g(y(x))f(x), \end{aligned}$$

d.h. $y(\cdot)$ ist Lösung des Anfangswertproblems.

Eindeutigkeit: Sei z irgendeine Lösung des AWP. Wegen $g(z(x_0)) = g(y_0) \neq 0$ ist dann $g(z(x)) \neq 0$ für x nahe x_0 , also auf Grund der Differenzialgleichung

$$\frac{z'(x)}{g(z(x))} = f(x) \quad \text{für } x \text{ nahe } x_0.$$

Integration von x_0 bis x liefert (mit Hilfe der Substitution $s = z(t)$)

$$\begin{aligned} F(x) &= F(x_0) + \int_{x_0}^x f(t) dt = G(y_0) + \int_{x_0}^x \frac{z'(t)}{g(z(t))} dt = G(y_0) + \int_{y_0}^{z(x)} \frac{1}{g(s)} ds \\ &= G(z(x)) \end{aligned}$$

d. h. z ist die Lösung der Gleichung $F(x) = G(z)$, wobei F und G Stammfunktionen von f bzw. $\frac{1}{g}$ sind mit $F(x_0) = G(y_0)$, also $z(x) = G^{-1}(F(x)) = y(x)$. ■

Explizit führt man die Lösung einer Gleichung mit getrennten Variablen meist am bequemsten wie folgt durch:

- (i) Man bestimmt eine Stammfunktion G von $1/g$,
- (ii) man bestimmt eine Stammfunktion F von f ,
- (iii) man löst $G(y) = F(x) + C$ nach y auf; das liefert (u.U. für jedes C mehrere) Lösungen y_C ,
- (iv) man bestimmt C so, dass der vorgeschriebene Anfangswert angenommen wird (ein solches C gibt es immer; man braucht nämlich nur mit den in (i) und (ii) gefundenen F und G zu wählen $C = G(y_0) - F(x_0)$.)

Beispiel 2.2 Die Differenzialgleichung $y' = xy^2$ hat zunächst die triviale Lösung $y(x) \equiv 0$. – Im Sinne der obigen Überlegungen ist hier $f(x) = x$, $g(y) = y^2$ (also insbesondere $g(y_0) \neq 0$ für alle $y_0 \neq 0$). Stammfunktionen von f und $1/g$ sind

$$F(x) = \frac{1}{2}x^2, \quad G(y) = -\frac{1}{y}.$$

Aus $G(y) = F(x) + C$ folgt also durch Auflösen nach y

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

Der Anfangswert $y(x_0) = y_0 \neq 0$ liefert schließlich C :

$$y_0 = -\left(\frac{1}{2}x_0^2 + C\right)^{-1} \implies C = -\frac{1}{y_0} - \frac{1}{2}x_0^2.$$

Diese Lösung hat bei x mit $x^2 = -2C$ eine Singularität (falls $C \leq 0$ ist), d. h. sie ist *nicht* über $x = \pm(-2C)^{1/2}$ hinaus fortsetzbar. Keine dieser Lösungen mündet allerdings in die bereits oben erwähnte Null-Lösung ein. □

Im *Spezialfall* $f(x) \equiv 1$ erhalten wir die *autonome Differenzialgleichung*

$$y' = g(y).$$

Mit $F(x) := x - x_0$ und $G(y) := \int_{y_0}^y \frac{1}{g(s)} ds$ wird dann

$$y(x) = G^{-1}(x - x_0).$$

Wenn G irgend eine Stammfunktion von $1/g$ ist, kann die Lösung angegeben werden in der Form

$$y(x) = G^{-1}(x + C) \quad \text{wobei } C \text{ so zu wählen ist, dass } y(x_0) = y_0 \text{ gilt.}$$

Beispiel 2.3 Ein sehr spezieller Fall ist die *lineare Differenzialgleichung* erster Ordnung

$$y' = -ay.$$

Mit

$$G(y) = - \int_{y_0}^y \frac{1}{as} ds = \frac{1}{a} (\ln |y_0| - \ln |y|)$$

(wenn y_0 und y gleiches Vorzeichen haben) folgt, was man natürlich auch hätte erraten können,

$$\begin{aligned} y(x) &= \pm \exp \left\{ -a(x - x_0) + \ln |y_0| \right\} \quad \text{für } y_0 \gtrless 0, \\ &= y_0 \exp \left\{ -a(x - x_0) \right\} \quad \text{für } y_0 \neq 0, \\ y(x) &\equiv 0 \quad \text{für } y_0 = 0. \end{aligned}$$

□

Beispiel 2.4 Wir betrachten die Gleichung

$$y' = |y|^{1/2}.$$

Für $y_0 > 0$ und $y > 0$ gilt

$$G(y) = \int_{y_0}^y s^{-1/2} ds = 2 \left(y^{1/2} - y_0^{1/2} \right)$$

und somit

$$\begin{aligned} y(x) &= \left\{ \frac{1}{2}(x - x_0) + y_0^{1/2} \right\}^2 = \frac{1}{4} \left(x - x_0 + 2y_0^{1/2} \right)^2 \\ &= \frac{1}{4}(x - c)^2 \quad \text{mit } c := x_0 - 2y_0^{1/2}. \end{aligned}$$

Wegen $y' = |y|^{1/2} > 0$ ist dies nur für $x > c$ eine Lösung der Differenzialgleichung.

Entsprechend erhält man für $y_0 < 0$ und $y < 0$

$$y(x) = -\frac{1}{4}(x - d)^2 \quad \text{für } x < d := x_0 + 2(-y_0)^{1/2}.$$

Außerdem ist natürlich

$$y(x) = 0 \quad \text{für alle } x \in \mathbb{R}$$

eine Lösung. Die Gesamtheit aller Lösungen ist somit:

- (i) $y(x) = 0$ für alle $x \in \mathbb{R}$,
- (ii) $y(x) = \begin{cases} 0 & \text{für } x \leq c, \\ \frac{1}{4}(x-c)^2 & \text{für } x \geq c, \end{cases}$
- (iii) $y(x) = \begin{cases} -\frac{1}{4}(x-d)^2 & \text{für } x \leq d, \\ 0 & \text{für } x \geq d, \end{cases}$
- (iv) $y(x) = \begin{cases} -\frac{1}{4}(x-d)^2 & \text{für } x \leq d, \\ 0 & \text{für } d \leq x \leq c, \\ \frac{1}{4}(x-c)^2 & \text{für } x \geq c. \end{cases}$

Läßt man $d = -\infty$ und $c = +\infty$ zu, so sind allein in (iv) alle Lösungen enthalten. \square

In diesem Fall ist also das Anfangswertproblem i. allg. *nicht eindeutig* lösbar; wir werden später sehen unter welcher zusätzlichen Bedingung die Eindeutigkeit garantiert ist.

2: 20. Oktober 2006

Beispiel 2.5 Ganz ähnliche Verhältnisse treten auf bei der Differentialgleichung

$$y' = 2 \operatorname{sign}(y)|y|^{1/2}.$$

Für $y > 0$ lautet die Gleichung $y' = 2y^{1/2}$ mit der Lösung $y(x) = (x-c)^2$; da für $y = (x-c)^2$ auf Grund der Differentialgleichung $y' \geq 0$ gelten müsste, ist dies jedoch nur für $x \geq c$ eine Lösung.

Entsprechend erhält man für $y < 0$ die Lösung $y(x) = -(x-d)^2$ im Bereich $x \geq d$.

Schließlich ist natürlich $y(x) \equiv 0$ eine Lösung, womit man insgesamt die folgenden Lösungen erhält:

- (i) $y(x) = 0$ für alle $x \in \mathbb{R}$,
- (ii) $y(x) = \begin{cases} 0 & \text{für } x \leq c, \\ (x-c)^2 & \text{für } x \geq c, \end{cases}$
- (iii) $y(x) = \begin{cases} 0 & \text{für } x \leq d, \\ -(x-d)^2 & \text{für } x \geq d. \end{cases}$

Für $c \rightarrow \infty$ bzw. $d \rightarrow \infty$ erhält man die Lösung (i) aus (ii) bzw. (iii). \square

Für $y_0 = 0$ ist in beiden Beispielen die obige Voraussetzung $g(y) \neq 0$ nicht erfüllt. Tatsächlich geht dadurch in diesen Beispielen die Eindeutigkeit verloren, während dies in Beispiel 2.3 (trotz $g(0) = 0$) nicht der Fall war. Der folgende Satz macht deutlich, wann trotz $g(y_0) = 0$ der Anfangswert y_0 zu einer lokal eindeutigen Lösung führt.

Satz 2.6 In der Differentialgleichung $y' = f(x)g(y)$ sei f stetig, $g(y_0) = 0$ und $g(y) \neq 0$ für y nahe y_0 und $y \neq y_0$. Sind die bei y_0 uneigentlichen Integrale

$$\int_{y_0}^{y_0+\delta} \frac{1}{g(s)} ds \quad \text{und} \quad \int_{y_0-\delta}^{y_0} \frac{1}{g(s)} ds$$

divergent für geeignetes $\delta > 0$, so ist $y(x) \equiv y_0$ die einzige Lösung des Anfangswertproblems

$$y' = f(x)g(y), \quad y(x_0) = y_0.$$

Mit anderen Worten: es gibt keine andere Lösung, die (wie in obigem Beispiel) in diese Lösung einmündet.

Beweis. Nehmen wir an, dass das AWP eine Lösung $y(x) \neq y_0$ besitzt. Ist $y(x) \neq y_0$ rechts von x_0 , so gibt es x_1 und x_2 mit

$$y_0 = y(x_1) \neq y(x) \quad \text{und} \quad g(y(x)) \neq 0 \quad \text{für} \quad x_0 \leq x_1 < x < x_2.$$

Dann gilt Substitution $s = y(t)$

$$\int_{y(x_1+\eta)}^{y(x)} \frac{1}{g(s)} ds = \int_{x_1+\eta}^x \frac{1}{g(y(t))} y'(t) dt = \int_{x_1+\eta}^x f(t) dt.$$

Für $\eta \rightarrow 0+$ divergiert nach Voraussetzung die linke Seite, während die rechte offensichtlich konvergiert. Das ist ein Widerspruch. Entsprechend verfährt man, wenn $y(x) \neq y_0$ links von x_0 gilt. ■

2.2 Differenzialgleichungen der Form $y' = f(ax + by + c)$

Im Fall $b = 0$ ist die Gleichung trivial; sei also o.E. $b \neq 0$. Für die neue Funktion

$$u(x) := ax + by(x) + c$$

erhält man eine (autonome) Differenzialgleichung mit getrennten Variablen

$$\begin{aligned} u' &= a + by' = a + bf(ax + by(x) + c) \\ &= a + bf(u). \end{aligned}$$

Ist u Lösung dieser Gleichung, so ist

$$y(x) = \frac{1}{b}(u(x) - ax - c)$$

Lösung der ursprünglichen Gleichung, denn es gilt

$$y' = \frac{1}{b}(u' - a) = \frac{1}{b}(a + bf(u) - a) = f(u) = f(ax + by + c).$$

Offensichtlich kann jede Lösung auf diese Weise gewonnen werden; wenn nämlich y eine Lösung ist, dann ist $u = ax + by + c$ Lösung der autonomen Differenzialgleichung.

Um das entsprechende AWP mit $y(x_0) = y_0$ zu lösen, muß man für $u(x)$ die Lösung der autonomen Differenzialgleichung mit $u(x_0) = by_0 + ax_0 + c$ wählen.

Beispiel 2.7 $y' = (x + y)^2$ (also $a = 1$, $b = 1 \neq 0$ und $c = 0$).

Die Differenzialgleichung für $u = x + y$ lautet $u' = 1 + u^2$ und hat die Lösung $u = \tan(x + C)$ (denn es ist $G(u) = \arctan u$ im Sinne von Abschnitt 2.1); hier ist C so zu wählen, dass für $y(x) = u(x) - x$ die Anfangsbedingung erfüllt ist. □

2.3 Die homogene Differenzialgleichung $y' = f\left(\frac{y}{x}\right)$

Wir betrachten homogene Differenzialgleichungen der Form $y' = f\left(\frac{y}{x}\right)$. Mit $u(x) := \frac{y(x)}{x}$ erhält man

$$u' = \frac{y'x - y}{x^2} = x^{-1} \left(y' - \frac{y}{x} \right) = x^{-1} (f(u) - u),$$

also eine Differenzialgleichung mit getrennten Variablen. Ist u Lösung dieser Gleichung, so löst

$$y(x) := xu(x)$$

die ursprüngliche Gleichung. Andererseits kann jede Lösung so gewonnen werden (die Division durch 0 für $x = 0$ stört nicht, da laut Differenzialgleichung für $x = 0$ auch y verschwinden muß).

Beispiel 2.8 $y' = \frac{y}{x} - \frac{x^2}{y^2}$, $y(1) = 1$. Hier ist also $f(t) = t - t^{-2}$ und die Differenzialgleichung für u lautet

$$u' = \frac{1}{x}(u - u^{-2} - u) = -\frac{1}{x} \frac{1}{u^2}, \quad u(1) = 1.$$

Mit $f(x) = -\frac{1}{x}$ und $g(u) = u^{-2}$ (im Sinne von Abschnitt 2.1) erhalten wir also

$$\begin{aligned} F(x) &= -\ln x, & G(u) &= \frac{1}{3}u^3, \\ u &= G^{-1}(F(x) + C) = \left\{ 3(-\ln x + C)^{1/3} \right\}. \end{aligned}$$

Mit der Anfangsbedingung $u(1) = 1$ folgt $C = 1/3$, also

$$y(x) = xu(x) = x\{1 - 3 \ln x\}^{1/3}.$$

□

2.4 Differenzialgleichungen der Form $y' = f\left(\frac{ax + by + c}{\alpha x + \beta y + \gamma}\right)$

Wir unterscheiden zwei Fälle:

(a) $\text{Det} \begin{pmatrix} a & b \\ \alpha & \beta \end{pmatrix} = 0$, z. B. $(a, b) = \lambda(\alpha, \beta)$. Die Differenzialgleichung hat also die Form

$$y' = f\left(\frac{\lambda(\alpha x + \beta y) + c}{(\alpha x + \beta y) + \gamma}\right) = g(\alpha x + \beta y)$$

mit $g(s) := f\left(\frac{\lambda s + c}{s + \gamma}\right)$. Dies ist eine Gleichung des Typs wie sie in Abschnitt 2.2 behandelt wurde.

(b) $\text{Det} \begin{pmatrix} a & b \\ \alpha & \beta \end{pmatrix} \neq 0$. Dann hat das Gleichungssystem

$$\begin{aligned} ax + by + c &= 0 \\ \alpha x + \beta y + \gamma &= 0 \end{aligned}$$

eine eindeutig bestimmte Lösung (x, y) (ξ, η) . Für die neuen Variablen

$$\bar{x} := x - \xi, \quad \bar{y} := y - \eta$$

gilt dann

$$\begin{aligned} \bar{y}(\bar{x}) &= y(x) - \eta = y(\bar{x} + \xi) - \eta, \\ \bar{y}'(\bar{x}) &= \frac{\partial}{\partial x} \{y(\bar{x} + \xi) - \eta\} = y'(\bar{x} + \xi) = f \left(\frac{a(\bar{x} + \xi) + by(\bar{x} + \xi) + c}{\alpha(\bar{x} + \xi) + \beta y(\bar{x} + \xi) + \gamma} \right) \\ &= f \left(\frac{a(\bar{x} + \xi) + b(\bar{y}(\bar{x}) + \eta) + c}{\alpha(\bar{x} + \xi) + \beta(\bar{y}(\bar{x}) + \eta) + \gamma} \right) \\ &\quad (\text{nach Konstruktion von } (\xi, \eta)) \\ &= f \left(\frac{a\bar{x} + b\bar{y}(\bar{x})}{\alpha\bar{x} + \beta\bar{y}(\bar{x})} \right) = f \left(\frac{a + b\bar{y}/\bar{x}}{\alpha + \beta\bar{y}/\bar{x}} \right) \\ &= g \left(\frac{\bar{y}}{\bar{x}} \right) \quad \text{mit} \quad g(s) := f \left(\frac{a + bs}{\alpha + \beta s} \right). \end{aligned}$$

Somit ist die Lösung dieser Differentialgleichung auf die Lösung einer homogenen Differentialgleichung zurückgeführt. Die Lösung der ursprünglichen Gleichung ist

$$y(x) = \bar{y}(\bar{x}) + \eta = \bar{y}(x - \xi) + \eta.$$

2.5 Die lineare Differentialgleichung erster Ordnung

Allgemein heißt eine Differentialgleichung $F(x, y, y', \dots, y^{(n)}) = 0$ *linear*, wenn F eine affine Funktion von $y, y', \dots, y^{(n)}$ ist, d. h.

$$F(x, y, y', \dots, y^{(n)}) = g(x) + \sum_{j=0}^n f_j(x)y^{(j)}.$$

Speziell hat also die *explizite lineare Differentialgleichung erster Ordnung* die Form

$$y' + f(x)y = g(x).$$

Die Gleichung heißt

$$\begin{aligned} \textit{homogen}, & \quad \text{falls } g \equiv 0 \text{ gilt:} & y' + f(x)y = 0, \\ \textit{inhomogen}, & \quad \text{falls } g \not\equiv 0 \text{ gilt:} & y' + f(x)y = g(x). \end{aligned}$$

$g(\cdot)$ heißt die *rechte Seite* der Differentialgleichung.

Die *homogene Gleichung* $y' = -f(x)y$ ist eine Gleichung mit getrennten Variablen, die wir sofort lösen können:

$$\begin{aligned} g(y) &= -y, \quad G(y) = -\ln |y|, \quad F(x) = \int_{x_0}^x f(t) dt, \\ y(x) &= \pm \exp \left\{ - \int_{x_0}^x f(t) dt + C \right\}. \end{aligned}$$

Die Anfangsbedingung $y(x_0) = y_0$ liefert also

$$y(x) = y_0 \exp \left\{ - \int_{x_0}^x f(t) dt \right\}.$$

Die Lösung existiert auf dem gesamten Intervall, auf dem f (stetig) definiert ist.

3: 24. Oktober 2006

Die *inhomogene Gleichung* wird mit einem (auch für Systeme und damit für lineare Gleichungen höherer Ordnung, vgl. Abschnitt 4.3 und 6.1) wichtigen *Trick* gelöst: In Anlehnung an die allgemeine Lösung der homogenen Gleichung $C \exp \left\{ - \int_{x_0}^x f(t) dt \right\} =: C\tilde{y}(x)$ setzt man für die Lösung der inhomogenen Gleichung an

$$y(x) = C(x) \exp \left\{ - \int_{x_0}^x f(t) dt \right\} =: C(x)\tilde{y}(x);$$

man ersetzt also die Konstante C durch eine Funktion $C(\cdot)$: *Variation der Konstanten*. (Da $\tilde{y}(x) \neq 0$ für alle x gilt, läßt sich $y(\cdot)$ jedenfalls so schreiben; die Frage ist nur, ob man $C(x)$ auch berechnen kann. Wenn überhaupt eine stetig differenzierbare Lösung $y(\cdot)$ existiert, ist jedenfalls $C(\cdot) = y(\cdot)/\tilde{y}(\cdot)$ stetig differenzierbar) – Einsetzen in die (inhomogene) Gleichung liefert

$$C'(x)\tilde{y}(x) + C(x)\tilde{y}'(x) + C(x)f(x)\tilde{y}(x) = g(x).$$

Wegen $\tilde{y}' + f(x)\tilde{y} = 0$ heben sich die Terme 2 und 3 heraus. Es bleibt

$$C'(x)\tilde{y}(x) = g(x), \quad C'(x) = \frac{g(x)}{\tilde{y}(x)} = g(x) \exp \left\{ \int_{x_0}^x f(t) dt \right\}$$

und somit

$$C(x) = \int_{x_0}^x g(s) \exp \left\{ \int_{x_0}^s f(t) dt \right\} ds + C_0 \quad \text{mit} \quad C_0 = y_0.$$

Satz 2.9 Sei $f : I \rightarrow \mathbb{C}$ stetig. Dann hat das Anfangswertproblem

$$\text{AWP} \quad y' + f(x)y = g(x), \quad y(x_0) = y_0$$

die auf ganz I definierte eindeutig bestimmte Lösung

$$y(x) = \exp \left\{ - \int_{x_0}^x f(t) dt \right\} \left[y_0 + \int_{x_0}^x g(s) \exp \left\{ \int_{x_0}^s f(t) dt \right\} ds \right].$$

Die Lösung des inhomogenen Anfangswertproblems läßt sich also schreiben als Summe einer speziellen Lösung der inhomogenen Gleichung (hier mit Anfangswert 0), und einer Lösung der homogenen Gleichung (die den geforderten Anfangswert annimmt).

Beweis. Offenbar löst dieses y das Anfangswertproblem (s. oben). Sind y und z Lösungen, so ist $w := y - z$ Lösung des homogenen Anfangswertproblems

$$w' + f(x)w = 0, \quad w(x_0) = 0.$$

Dann ist aber $w \equiv 0$. Wäre nämlich $w(x_1) \neq 0$, so wäre (vgl. obige Formel für die Lösung der homogenen Gleichung, die durch ihren Anfangswert eindeutig bestimmt ist) $w(x) \neq 0$ für alle x ; das ist ein Widerspruch zu $w(x_0) = 0$. ■

2.6 Die exakte Differenzialgleichung

Eine Differenzialgleichung

$$y'h(x, y) + g(x, y) = 0 \quad \text{mit} \quad h, g : D \rightarrow \mathbb{R}, D \subset \mathbb{R}^2$$

heißt *exakt*, wenn eine stetig differenzierbare Funktion $F : D \rightarrow \mathbb{R}$, existiert mit

$$h(x, y) = F_y(x, y), \quad g(x, y) = F_x(x, y).$$

Ein solches F heißt eine *Stammfunktion* der Gleichung. Die Differenzialgleichung liefert

$$\begin{aligned} \frac{d}{dx}F(x, y(x)) &= D_1F(x, y(x)) + D_2F(x, y(x))y'(x) \\ &= g(x, y(x)) + h(x, y(x))y'(x) = 0, \end{aligned}$$

d. h. auf einer Lösungskurve der Differenzialgleichung ist F konstant; die Lösungskurven sind die Niveaulinien von F .

So sind z. B. die Kreise mit Mittelpunkt 0

$$y_c(x) = \pm \sqrt{c^2 - x^2}$$

die Niveaulinien von $F(x, y) = x^2 + y^2$, also die Lösungen der Differenzialgleichung

$$2y'y + 2x = 0 \quad \text{bzw.} \quad y'y + x = 0.$$

Satz 2.10 Seien $h, g : D \rightarrow \mathbb{R}$ ($D \subset \mathbb{R}^2$) stetig mit $h(x, y) \neq 0$. Ist die Differenzialgleichung

$$y'h(x, y) + g(x, y) = 0 \quad \text{exakt mit Stammfunktion} \quad F \in C^1(D),$$

so erhält man durch Auflösen der Gleichungen $F(x, y) = C$ ($C \in \mathbb{R}$) alle Lösungen der Differenzialgleichung. (Die Bedingung $h = D_2F \neq 0$ garantiert, dass die Gleichungen $F_f(x, y) = C$ tatsächlich nach y auflösbar sind: Für jedes x gibt es nur ein y mit $F(x, y) = C$.)

Wie stellt man nun fest, ob die Differenzialgleichung exakt ist?

Satz 2.11 Ist $D \subset \mathbb{R}^2$ einfach zusammenhängend und sind $g, h : D \rightarrow \mathbb{R}$ stetig differenzierbar, so ist die Differentialgleichung $y'h(x, y) + g(x, y) = 0$ genau dann exakt, wenn gilt

$$g_y(x, y) = h_x(x, y) \quad \text{für alle } (x, y) \in D.$$

Für einen beliebigen Punkt $(x_0, y_0) \in D$ ist eine Stammfunktion F gegeben durch

$$F(x, y) = \int_{(x_0, y_0)}^{(x, y)} g(s, t) ds + h(s, t) dt.$$

(Verschiedene Stammfunktionen – für verschiedene (x_0, y_0) unterscheiden sich nur durch eine additive Konstante. Falls D nicht einfach zusammenhängend ist, gilt diese Aussage in jeder einfach zusammenhängenden Teilmenge von D , also jedenfalls lokal; da aber auch die Lösungen der Differentialgleichung i.allg. nur lokal existieren, stört dies nicht weiter.)

Beweis. Mit dem Vektorfeld $k(x, y) = (g(x, y), h(x, y))$ und einem beliebigen $(x_0, y_0) \in D$ definieren wir (als Kurvenintegral über eine beliebige Kurve von (x_0, y_0) nach (x, y))

$$F(x, y) = \int_{(x_0, y_0)}^{(x, y)} k_1(s, t) ds + k_2(s, t) dt = \int_{(x_0, y_0)}^{(x, y)} g(s, t) ds + h(s, t) dt.$$

Wegen $\text{rot}k = h_x - g_y = 0$ ist die Integrabilitätsbedingung erfüllt, d. h. das Kurvenintegral ist wegunabhängig; es ist deshalb gleichgültig, welcher Weg von (x_0, y_0) nach (x, y) gewählt wird. Weiter gilt für dieses F offenbar $F_x = g, F_y = h$, d. h. F ist ein Potential von k . ■

Beispiel 2.12 In der Differentialgleichung

$$xe^{xy}y' + 2x + ye^{xy} = 0$$

haben wir

$$\begin{aligned} h(x, y) &= xe^{xy}, & h_x(x, y) &= e^{xy} + xye^{xy}, \\ g(x, y) &= 2x + ye^{xy}, & g_y(x, y) &= e^{xy} + xye^{xy}, \end{aligned}$$

d. h. die Differentialgleichung ist exakt. Für F erhalten wir mit obigem Satz ((x_0, y_0) beliebig in \mathbb{R}^2)

$$\begin{aligned} F(x, y) &= \int_{(x_0, y_0)}^{(x, y)} g(s, t) ds + h(s, t) dt = \int_{(x_0, y_0)}^{(x, y)} (2s + te^{st}) ds + se^{st} dt \\ &= \int_{(x_0, y_0)}^{(x, y)} (2s + y_0 e^{sy_0}) ds + \int_{(x_0, y_0)}^{(x, y)} xe^{xt} dt = (s^2 + e^{sy_0}) \Big|_{x_0}^x + e^{xt} \Big|_{y_0}^y \\ &= x^2 + e^{xy_0} - x_0^2 - e^{x_0y_0} + e^{xy} - e^{xy_0} \\ &= x^2 + e^{xy} - (x_0^2 + e^{x_0y_0}) = x^2 + e^{xy} - C(x_0, y_0). \end{aligned}$$

(Mit etwas Phantasie hätte man F natürlich auch erraten können.) Damit erhalten wir nun die Lösungen

$$y = \frac{1}{x} \ln(C - x^2) \quad (C \in \mathbb{R}).$$

□

Es ist natürlich ein seltener Zufall, wenn eine Differenzialgleichung exakt ist. Unter Umständen kann sie durch Multiplikation mit einer Funktion $M(x, y) \neq 0$ exakt gemacht werden; eine solche Funktion wird als *integrierender Faktor* oder *Eulerscher Multiplikator* bezeichnet. Die Menge der Lösungen wird durch eine solche Multiplikation natürlich nicht verändert.

Beispiel 2.13 Die Gleichung

$$2y'x + y = 0 \quad (h = 2x, g = y, h_x = 2, g_y = 1)$$

ist nicht exakt. Durch Multiplikation mit $x^{-1/2}$ wird sie exakt:

$$2y'x^{1/2} + yx^{-1/2} = 0, \quad h = 2x^{1/2}, \quad g = yx^{-1/2}, \quad F(x, y) = 2y\sqrt{x}.$$

Ebenso könnte man mit y multiplizieren

$$2y'xy + y^2 = 0, \quad h = 2xy, \quad g = y^2, \quad F(x, y) = xy^2.$$

(In beiden Fällen kann man die Stammfunktion leicht erraten; wenn dies nicht gelingt, muß man wie im obigen Beispiel vorgehen.) Die Lösungen der Differenzialgleichung sind also

$$y = Cx^{-1/2}.$$

□

Nach obigem Satz über die Exaktheit einer Differenzialgleichung ist offenbar M genau dann ein integrierender Faktor, wenn gilt: $(Mg)_y = (Mh)_x$, d. h.

$$M_y g + M g_y = M_x h + M h_x.$$

In dieser Form ist das Verfahren schwer anwendbar. Häufig gelingt es aber, einen integrierenden Faktor zu finden, der nur von x oder nur von y abhängt:

Satz 2.14 (i) $M(x)$ ist genau dann ein integrierender Faktor, wenn gilt

$$\frac{M'(x)}{M(x)} = \frac{g_y(x, y) - h_x(x, y)}{h(x, y)}.$$

Ein solcher existiert, wenn hier die rechte Seite nur von x abhängt.

(ii) $M(y)$ ist genau dann ein integrierender Faktor, wenn gilt

$$\frac{M'(y)}{M(y)} = \frac{h_x(x, y) - g_y(x, y)}{g(x, y)}.$$

Ein solcher existiert, wenn hier die rechte Seite nur von y abhängt.

Hiermit sieht man z. B., dass es in Beispiel und Beispiel 2.13 integrierende Faktoren gibt, die nur von x oder y abhängen.

4: 27. Oktober 2006

Beispiel 2.15 Die Differentialgleichung

$$xyy' + \frac{1}{2}(xy^2 + y^2) = 0$$

ist nicht exakt:

$$h(x, y) = xy, \quad h_x(x, y) = y, \quad g(x, y) = \frac{1}{2}(xy^2 + y^2), \quad g_y(x, y) = (x + 1)y,$$

also

$$h_x(x, y) \neq g_y(x, y).$$

Es gilt aber

$$\frac{1}{h}(g_y - h_x) = 1 \quad (\text{hängt also nur von } x \text{ ab}).$$

Aus $M'/M = 1$ erhält man also den integrierenden Faktor

$$M = e^x$$

und somit die exakte Differentialgleichung

$$xye^x y' + \frac{1}{2}(x + 1)y^2 e^x = 0.$$

Als Stammfunktion erhält man (erraten oder berechnen wie in Beispiel 2.12)

$$F(x, y) = \frac{1}{2}y^2 x e^x,$$

und somit die Lösungen

$$y = C \left\{ 2 \frac{1}{x} e^{-x} \right\}^{1/2} = C' x^{-1/2} e^{-x/2}.$$

□

2.7 Übungsaufgaben

1. Man bestimme alle Lösungen der Differentialgleichung

$$y' = -3|y|^{2/3}$$

und bestimme die Menge der Anfangsdaten, für die das AWP *nicht lokal* eindeutig lösbar ist.

2. a) Man beschreibe ein Verfahren zur Lösung der Differentialgleichung

$$y'' = f(y),$$

indem man die mit y' multiplizierte Gleichung integriert.

- b) Auf diese Weise löse man die Differentialgleichung

$$y'' = -y.$$

3. a) Man löse das Anfangswertproblem

$$(1 + e^x)yy' = e^x, \quad y(1) = 1.$$

- b) Man bestimme alle Lösungen von

$$y' = (x - y)^2 + 1.$$

4. Man bestimme alle Lösungen der Differentialgleichung

$$xy' - y(1 + xy) = 0.$$

Anleitung: Es gibt einen integrierenden Faktor, der nur von y abhängt.

5. Sei $g : \mathbb{R} \rightarrow \mathbb{R}$ stetig mit $\lim_{x \rightarrow \infty} g(x) = b$, $a \in (0, \infty)$. Man zeige mit Hilfe der Lösungsformel aus der Vorlesung, dass für jede Lösung der Differentialgleichung $y' + ay = g(x)$ gilt

$$\lim_{x \rightarrow \infty} y(x) = \frac{b}{a}.$$

6. Man bestimme alle Lösungen der Differentialgleichung.

$$y' = \frac{x - y}{x + 2y}.$$

Anleitung: Umschreiben als homogene Differentialgleichung.

7. a) Man zeige: Die Tangenten der Lösungskurven einer linearen Differentialgleichung erster Ordnung in den Punkten mit gleicher Abszisse schneiden sich in einem Punkt.
 b) Die Gesamtheit dieser Schnittpunkte bildet eine Kurve, die *Leitkurve*. Man gebe die Leitkurve in Parameterdarstellung.
 c) Man bestimme die Leitkurve der Differentialgleichung $y' - e^{-x}y = 1 - xe^{-x}$ in kartesischen Koordinaten.

8. Man löse die Differentialgleichung

$$xy' + y = y^2 \ln x \quad \text{für } x > 0$$

mit Hilfe der Substitution $z = xy$.

9. Ein Punkt P_1 bewege sich auf der x_1 -Achse der (x_1, x_2) -Ebene mit Geschwindigkeit 1 in positiver Richtung. Ein weiterer Punkt P_2 bewege sich in \mathbb{R}^2 so, dass
- (i) der Abstand der Punkte P_1 und P_2 konstant gleich 1 ist,
 - (ii) der Punkt P_2 sich immer in Richtung P_1 bewegt.

Man stelle ein Differenzialgleichungssystem

$$x'_j = f_j(t, x_1, x_2), j = 1, 2$$

für die Bewegung von P_2 auf.

Anleitung: Jede der Bedingungen (i) und (ii) liefert eine Gleichung; aus diesen beiden Gleichungen gewinnt man das System.

Ergebnis: $x'_1 = (x_1 - t)^2$, $x'_2 = x_2(x_1 - t)$.

10. Man löse das System aus Aufgabe 9.

Hilfen :

$$\frac{d}{dx} \tanh x = 1 - (\tanh x)^2$$
$$\frac{d}{dx} \ln \cosh x = \tanh x.$$

3 Existenz, Eindeutigkeit und stetige Abhängigkeit

In zahlreichen konkreten Fällen haben wir gesehen, dass das Anfangswertproblem

$$\text{AWP : } y' = f(x, y), \quad y(x_0) = y_0$$

(zumindest in einer Umgebung von x_0 , d. h. lokal) genau eine Lösung besitzt. Wir haben aber auch gesehen, dass diese Lösung in manchen Fällen nicht eindeutig bestimmt ist. Welche Voraussetzungen an das AWP sind nun eigentlich ausreichend, um die Existenz und die Eindeutigkeit zu garantieren?

3.1 Die Lipschitzbedingung

Wir betrachten also Anfangswertprobleme der obigen Form mit

$$f : G \rightarrow \mathbb{R}^m \quad (\text{bzw. } \mathbb{C}^m), \quad G \subset \mathbb{R}^{m+1} \quad (\text{bzw. } \mathbb{R} \times \mathbb{C}^m) \quad \text{offen}$$

(im folgenden legen wir o. E. in der Regel den allgemeineren Fall komplexwertiger Funktionen y bzw. y_j zu Grunde). Ausführlich geschrieben handelt es sich also um das System

$$\begin{aligned} y_1'(x) &= f_1(x, y_1(x), \dots, y_m(x)), \\ &\vdots \\ y_m'(x) &= f_m(x, y_1(x), \dots, y_m(x)). \end{aligned}$$

Man sagt f erfüllt (bezüglich y) eine (globale) *Lipschitzbedingung* (bezüglich y) mit *Lipschitzkonstante* L , wenn gilt

$$|f(x, y) - f(x, \tilde{y})| \leq L|y - \tilde{y}| \quad \text{für } (x, y), (x, \tilde{y}) \in G;$$

f erfüllt *lokal* eine *Lipschitzbedingung* (bezüglich y), wenn es zu jedem Punkt $(x_0, y_0) \in G$ eine Umgebung U in G gibt so, dass $f|_U$ eine Lipschitzbedingung mit einer Lipschitzkonstanten $L(U)$ erfüllt.

Beispiel 3.1 a) $f : \mathbb{R} \times \mathbb{C} \rightarrow \mathbb{C}$, $f(x, y) = xy^2$ genügt lokal einer Lipschitzbedingung:

$$|f(x, y) - f(x, \tilde{y})| = |x(y + \tilde{y})(y - \tilde{y})| \leq |x(y + \tilde{y})||y - \tilde{y}|.$$

Die Funktion genügt aber keiner globalen Lipschitzbedingung: für $y = \tilde{y} + 1$ und $x \neq 0$ gilt nämlich $|f(x, y) - f(x, \tilde{y})| = |x||2y + 1| = L|y - \tilde{y}|$ mit $L \rightarrow \infty$ für $y \rightarrow \infty$.

b) $f : \mathbb{R} \times \mathbb{C} \rightarrow \mathbb{C}$, $f(x, y) = |y|^{1/2}$ genügt (bei 0) keiner Lipschitzbedingung. Es wird sich zeigen, dass das der Grund dafür ist, dass das AWP für $y' = |y|^{1/2}$ mit Anfangswert $y_0 = y(x_0) = 0$ nicht eindeutig lösbar ist.

□

Für den Nachweis der Gültigkeit einer Lipschitzbedingung ist folgender Satz sehr nützlich.

Satz 3.2 Sei $G \subset \mathbb{R} \times \mathbb{C}^m$ offen, $f : G \rightarrow \mathbb{C}^m$ stetig und bezüglich y stetig partiell differenzierbar (d. h. f und $D_y f$ sind stetig in G). Dann gilt:

- a) f genügt lokal einer Lipschitzbedingung.
 b) Ist G konvex und sind die partiellen Ableitungen von f nach y in G beschränkt, so genügt f einer Lipschitzbedingung.

Beweis. a) Sei $(x_0, y_0) \in G$. Dann existiert ein $r > 0$ so, dass

$$U := \left\{ (x, y) \in \mathbb{R} \times \mathbb{C}^m : |x - x_0| \leq r, |y - y_0| \leq r \right\} \subset G$$

gilt, und in U alle partiellen Ableitungen von f nach y beschränkt sind. Nach dem m -dimensionalen Mittelwertsatz gilt dann

$$|f(x, y) - f(x, \tilde{y})| \leq L|y - \tilde{y}| \quad \text{für } (x, y), (x, \tilde{y}) \in U$$

mit

$$L = \sup \left\{ |D_y f(x, y)| : (x, y) \in U \right\},$$

wobei mit $|D_y f(x, y)|$ die Norm der linearen Abbildung $D_y f(x, y) : \mathbb{C}^m \rightarrow \mathbb{C}^m$ gemeint ist.

Eine solche Abschätzung läßt sich aber auch elementar, nur mit Hilfe des 1-dimensionalen Mittelwertsatzes, zeigen, denn es gilt

$$\begin{aligned} |f(x, y) - f(x, \tilde{y})| &\leq \left| f(x, y) - f(x, \tilde{y}_1, y_2, \dots, y_m) \right| \\ &\quad + \left| f(x, \tilde{y}_1, y_2, \dots, y_m) - f(x, \tilde{y}_1, \tilde{y}_2, y_3, \dots, y_m) \right| \\ &\quad \vdots \\ &\quad + \left| f(x, \tilde{y}_1, \dots, \tilde{y}_{m-1}, y_m) - f(x, \tilde{y}) \right| \\ &\leq \tilde{L} \left\{ |y_1 - \tilde{y}_1| + \dots + |y_m - \tilde{y}_m| \right\} = \tilde{L} \sqrt{m} |y - \tilde{y}| \\ &= \hat{L} |y - \tilde{y}| \end{aligned}$$

mit $\tilde{L} := \sup \left\{ |D_{y_j} f(x, y)| : (x, y) \in U, j = 1, \dots, m \right\}$.

- b) In diesem Fall kann $U = G$ gewählt werden. ■

3.2 Existenz und Eindeutigkeit

Satz 3.3 (Eindeutigkeitssatz) Sei $G \subset \mathbb{R} \times \mathbb{C}^m$ offen, $f : G \rightarrow \mathbb{C}^m$ stetig und erfülle lokal eine Lipschitzbedingung bezüglich y . Sind y und $z : I \rightarrow \mathbb{C}^m$ Lösungen des Anfangswertproblems

$$y' = f(x, y) \quad \text{mit } y(x_0) = z(x_0) \quad \text{für ein } x_0 \text{ aus } I,$$

so gilt $y(x) = z(x)$ für alle $x \in I$. (Kurz: Eine lokale Lipschitzbedingung impliziert globale Eindeutigkeit.)

Beweis. Würde die Behauptung nicht gelten, so gäbe es

- (i) ein $\tilde{x} > x_0, \tilde{x} \in I$ mit $y(\tilde{x}) \neq z(\tilde{x})$, oder
- (ii) ein $\tilde{x} < x_0, \tilde{x} \in I$ mit $y(\tilde{x}) \neq z(\tilde{x})$.

O.E. nehmen wir an, dass der erste Fall vorliegt. Sei

$$x_1 := \sup \left\{ x \in I : y(s) = z(s) \quad \text{für} \quad x_0 \leq s \leq x \right\} < \tilde{x}.$$

Da y und z stetig sind, gilt

$$y(x_1) = z(x_1).$$

Sei r so, dass

$$U := \left\{ (x, y) \in \mathbb{R} \times \mathbb{C}^m : |x - x_1| \leq r, |y - y(x_1)| \leq r \right\} \subset G$$

gilt und f in U eine Lipschitzbedingung mit Lipschitzkonstante $L = L(U)$ erfüllt. Da y und z stetig sind, gibt es ein $\delta > 0$ mit

$$|y(x) - y(x_1)| \leq r, \quad |z(x) - y(x_1)| \leq r \quad \text{für} \quad |x - x_1| < \delta;$$

dabei kann o. E. $\delta \leq r$ gewählt werden. Für $x \in I$ mit $0 \leq x - x_1 < \delta$ gilt also

$$|y(x) - z(x)| = \left| \int_{x_1}^x \left\{ f(t, y(t)) - f(t, z(t)) \right\} dt \right| \leq \int_{x_1}^x L |y(t) - z(t)| dt.$$

Mit

$$M(x) := \sup \left\{ |y(t) - z(t)| : x_1 \leq t \leq x \right\} \quad \text{für} \quad x \geq x_1$$

gilt also für $0 \leq s - x_1 \leq x - x_1 \leq \delta$

$$|y(s) - z(s)| \leq L(s - x_1)M(s) \leq L(x - x_1)M(x).$$

Bildet man links das Supremum über alle $s \in (x_1, x)$, so erhält man

$$M(x) \leq L(x - x_1)M(x).$$

was für x nahe bei x_1 nur möglich ist, wenn $M(x) = 0$ gilt. Das ist aber im Widerspruch zur Konstruktion von x_1 . ■

Aus Beispiel 2.4, $y' = |y|^{1/2}$, wissen wir bereits, dass die Eindeutigkeit der Lösung eines Anfangswertproblems i. allg. nicht gegeben ist, wenn $f(\cdot, \cdot)$ keine Lipschitzbedingung bezüglich y erfüllt. Im folgenden werden wir sehen, dass andererseits die Lipschitzbedingung ausreicht, um lokal auch die Existenz einer Lösung zu garantieren (die dann zwangsläufig eindeutig bestimmt ist).

5: 31. Oktober 2006

Satz 3.4 (Existenzsatz von Picard–Lindelöf) a) Ist $G \subset \mathbb{R} \times \mathbb{C}^m$ offen und $f : G \rightarrow \mathbb{C}^m$ stetig und erfüllt lokal eine Lipschitzbedingung bezüglich y , so gibt es zu beliebigem $(x_0, y_0) \in G$ ein $\varepsilon > 0$ so, dass das Anfangswertproblem

$$\text{AWP} \quad y' = f(x, y), \quad y(x_0) = y_0$$

in $[x_0 - \varepsilon, x_0 + \varepsilon]$ lösbar ist (und nach Satz 3.4) eindeutig.

- b) Ist $[x_0, \hat{x})$ das größte rechts von x_0 gelegene rechtsoffene Intervall, auf dem die Lösung existiert, und ist K eine kompakte Teilmenge von G , so gibt es ein $\tilde{x} \in (x_0, \hat{x})$ mit $(x, y(x)) \notin K$ für $x > \tilde{x}$, d. h. die Lösung nähert sich für $x \rightarrow \hat{x}$ dem Rand von G bzw. wird unendlich. (Entsprechendes gilt, wenn $(\hat{x}, x_0]$ das größte links von x_0 liegende linksoffene Intervall ist, auf dem die Lösung existiert.) Die Lösung kann also immer soweit fortgesetzt werden, bis sie an den Rand von G stößt.

Beweis. a) Da G offen ist, existiert ein $r > 0$ mit

$$U := \left\{ (x, y) \in \mathbb{R} \times \mathbb{C}^m : |x - x_0| \leq r, |y - y_0| \leq r \right\} \subset G.$$

In U erfüllt f eine Lipschitzbedingung bezüglich y mit einer Lipschitzkonstanten $L = L(U)$. Da f stetig ist, existiert ein $M \geq 0$ mit

$$|f(x, y)| \leq M \quad \text{für} \quad (x, y) \in U.$$

Sei

$$\varepsilon := \min\left\{r, \frac{r}{M}\right\}, \quad I := [x_0 - \varepsilon, x_0 + \varepsilon].$$

(Diese Wahl von ε läßt sich leicht begründen: Der Betrag (die Norm) der Ableitung einer Lösung y in U ist höchstens M und das Intervall I ist so klein gewählt, dass eine Kurve, die in (x_0, y_0) beginnt und immer maximale Ableitung hat, den Rand von U frühestens in den Randpunkten von I erreicht.)

Eine Funktion $y : I \rightarrow \mathbb{C}^m$ löst offenbar genau dann das AWP, wenn die *Integralgleichung*

$$y(x) = y(x_0) + \int_{x_0}^x f(t, y(t)) \, dt, \quad x \in I$$

gilt. Der Vorteil dieser Integralgleichung gegenüber der Differenzialgleichung mit dem Anfangswert liegt u. a. darin, dass sie sowohl die Differenzialgleichung als auch die Anfangsbedingung enthält. Zur Lösung dieser Integralgleichung benutzt man das *Iterationsverfahren von Picard*, ein ganz zentrales Verfahren in der Theorie der Differenzialgleichungen:

$$\begin{aligned} y_0(x) &= y_0 \quad \text{für alle} \quad x \in I, \\ y_{k+1}(x) &= y_0 + \int_{x_0}^x f(t, y_k(t)) \, dt \quad \text{für} \quad k = 0, 1, 2, \dots, \quad x \in I. \end{aligned}$$

Wir zeigen: Die Folge (y_k) ist für alle $k \in \mathbb{N}$ auf ganz I definiert und konvergiert auf I gleichmäßig gegen eine Lösung der Integralgleichung (und damit des AWP).

(i) Alle y_k sind auf ganz I definiert: dafür genügt es, induktiv zu zeigen:

$$|y_k(x) - y_0| \leq r \quad \text{für alle } x \in I, k \in \mathbb{N}_0.$$

k = 0: Dies ist klar, da $y_0(x) = y_0$ gilt.

k \implies k + 1: Sei $|y_k(x) - y_0| \leq r$ für alle $x \in I$. Dann folgt aus der Rekursionsformel

$$|y_{k+1}(x) - y_0| \leq \left| \int_{x_0}^x f(t, y_k(t)) dt \right| \leq \left| \int_{x_0}^x |f(t, y_k(t))| dt \right| \leq M\varepsilon \leq r.$$

(ii) Es gilt: $|y_{k+1}(x) - y_k(x)| \leq ML^k \frac{|x - x_0|^{k+1}}{(k+1)!}$ für alle $x \in I, k \in \mathbb{N}_0$.

$$\mathbf{k = 0:} \quad |y_1(x) - y_0| = \left| \int_{x_0}^x f(t, y_0) dt \right| \leq M|x - x_0| = ML^0 \frac{|x - x_0|^{0+1}}{(0+1)!}.$$

$$\begin{aligned} \mathbf{k - 1 \implies k:} \quad |y_{k+1}(x) - y_k(x)| &\leq \left| \int_{x_0}^x |f(t, y_k(t)) - f(t, y_{k-1}(t))| dt \right| \\ &\leq \left| \int_{x_0}^x L|y_k(t) - y_{k-1}(t)| dt \right| \leq L \frac{ML^{k-1}}{k!} \left| \int_{x_0}^x |t - x_0|^k dt \right| \\ &= ML^k \frac{|x - x_0|^{k+1}}{(k+1)!}. \end{aligned}$$

(iii) Gleichmäßige Konvergenz der Folge (y_k) auf I : Aus (ii) folgt

$$|y_{k+1}(x) - y_k(x)| \leq ML^k \frac{\varepsilon^{k+1}}{(k+1)!} \quad \text{für alle } x \in I, k \in \mathbb{N}_0.$$

Also ist die Reihe

$$\sum_{k=0}^{\infty} |y_{k+1}(x) - y_k(x)| \quad \text{in } I \text{ gleichmäßig konvergent.}$$

Somit existiert im Sinne der gleichmäßigen Konvergenz

$$\begin{aligned} y(x) &:= \lim_{k \rightarrow \infty} y_{k+1}(x) = \lim_{k \rightarrow \infty} \sum_{\ell=0}^k (y_{\ell+1}(x) - y_{\ell}(x)) + y_0 \\ &= y_0 + \sum_{k=0}^{\infty} (y_{k+1}(x) - y_k(x)), \end{aligned}$$

und $y(\cdot)$ ist stetig in I .

(iv) $y(\cdot)$ ist Lösung der Integralgleichung: Wegen

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

konvergiert $f(x, y_k(x))$ gleichmäßig gegen $f(x, y(x))$ und somit gilt

$$\begin{aligned} y(x) &= \lim_{k \rightarrow \infty} y_k(x) = \lim_{k \rightarrow \infty} \left\{ y_0 + \int_{x_0}^x f(t, y_{k-1}(t)) dt \right\} \\ &= y_0 + \int_{x_0}^x f(t, y(t)) dt. \end{aligned}$$

Eine genaue Durchsicht des Beweises zeigt, dass die Lipschitzkonstante L nur bei den Konvergenzuntersuchungen gebraucht wird. Die Abschätzung M für $|f(\cdot, \cdot)|$ geht in die Abschätzung von y_k bzw. y und in die Konvergenz ein.

b) Nehmen wir an, dass die Behauptung nicht gilt. Dann existiert ein Kompaktum $K \subset G$ und eine Folge $x_n \in [x_0, \hat{x}]$ mit $x_n \rightarrow \hat{x}$ und $(x_n, y(x_n)) \in K$ für alle n . Also gibt es eine Teilfolge (\tilde{x}_n) von (x_n) mit $(\tilde{x}_n, y(\tilde{x}_n)) \rightarrow (\hat{x}, \hat{y}) \in K$. Ist $\varepsilon = \varepsilon(\hat{x}, \hat{y}) > 0$ für (\hat{x}, \hat{y}) so definiert, wie in Teil a) $\varepsilon(x_0, y_0)$ für (x_0, y_0) , so ist für hinreichend große n das

$$\text{AWP } y' = f(x, y) \quad \text{mit } y_n = y(\tilde{x}_n) \quad \text{von oben}$$

mindestens in $[\tilde{x}_n, \tilde{x}_n + \varepsilon/2]$ lösbar und somit ist y über \hat{x} hinaus fortsetzbar. ■

Besonders einfache Differenzialgleichungen können mit Hilfe des Picard'schen Iterationsverfahrens tatsächlich explizit gelöst werden, z. B.:

Beispiel 3.5 Für das Anfangswertproblem

$$\text{AWP } y' = 2xy, \quad y(0) = y_0$$

liefert die Iteration

$$\begin{aligned} y_0(x) &= y_0 \\ y_1(x) &= y_0 + 2 \int_0^x ty_0 dt = y_0 + y_0 x^2 = y_0(1 + x^2), \\ y_2(x) &= y_0 + 2 \int_0^x ty_0(1 + t^2) dt = y_0(1 + x^2 + \frac{1}{2}x^4), \end{aligned}$$

und weiter durch Induktion (Beweis!)

$$y_k(x) = y_0 \left(1 + x^2 + \frac{1}{2!}x^4 + \frac{1}{3!}x^6 + \dots + \frac{1}{k!}x^{2k} \right).$$

Diese Folge konvergiert auf ganz \mathbb{R} (lokal gleichmäßig) gegen die Lösung des AWP

$$y(x) = y_0 e^{x^2},$$

die man natürlich schneller durch Trennung der Variablen erhalten hätte. □

Der folgende Satz liefert uns mehr Information über den (evtl. maximalen) Bereich, auf dem das

$$\text{AWP } y' = f(x, y), \quad y(x_0) = y_0$$

lösbar ist.

6: 3. November 2006

Satz 3.6 a) Sei $G = \{(x, y) \in \mathbb{R} \times \mathbb{C}^m : x_0 \leq x \leq x_0 + a, |y - y_0| \leq b\}$, $f : G \rightarrow \mathbb{C}^m$ stetig, und erfülle eine (bezüglich G lokale) Lipschitzbedingung bezüglich y . Ist

$$M := \max \{|f(x, y)| : (x, y) \in G\} \quad \text{und} \quad \alpha := \min\{a, b/M\},$$

so ist das AWP auf $[x_0, x_0 + \alpha]$ (eindeutig) lösbar. Entsprechendes gilt für ein Intervall links von x_0 .

b) Ist $G = \{(x, y) \in \mathbb{R} \times \mathbb{C}^m : x_0 \leq x \leq x_0 + a\}$ ein Streifen, $f : G \rightarrow \mathbb{C}^m$ stetig, und erfüllt f eine globale Lipschitzbedingung bezüglich y , so ist das AWP auf ganz $[x_0, x_0 + a]$ eindeutig lösbar. Entsprechendes gilt für einen Streifen links von x_0 . (Eigentlich wird nur eine lokale Lipschitz-Bedingung und $|f(x, y)| \leq M_0 + L|y|$ benutzt.)

c) Ist $G = \mathbb{R} \times \mathbb{C}^m$, f stetig, und erfüllt f eine globale Lipschitzbedingung bezüglich y , so ist das AWP auf ganz \mathbb{R} eindeutig lösbar. Entsprechendes gilt für den Halbraum $\mathbb{R} \times [x_0, \infty)$ bzw. $\mathbb{R} \times (-\infty, x_0]$

Beweis. Die Eindeutigkeit ergibt sich in jedem Fall unmittelbar aus Satz 3.3. Es bleibt also die jeweilige Existenzaussage zu beweisen.

a) Wegen $|y'(x)| \leq M$ stößt die Lösung frühestens bei $x_0 + a$ oder $x_0 + b/M$ an den Rand von G , je nachdem, welches der kleinere Wert ist. Mit Satz 3.4 b) folgt damit die Behauptung.

b) Mit $M_0 := \max \{|f(x, 0)| : x_0 \leq x \leq x_0 + a\}$ gilt offenbar auf Grund der Lipschitzbedingung

$$|f(x, y)| = M_0 + L|y|$$

und somit für $g(x) := (1 + |y(x)|^2)^{1/2}$

$$\begin{aligned} g'(x) &= \frac{1}{g(x)} \operatorname{Re} \langle y(x), y'(x) \rangle \leq \frac{1}{g(x)} |y(x)| |f(x, y(x))| \\ &\leq |f(x, y(x))| \leq M_0 + L|y(x)| \leq Kg(x). \end{aligned}$$

Daraus folgt

$$|y(x)| \leq g(x) \leq K' e^{Kx}$$

in dem Bereich, in dem y existiert. Also kann die Lösung den Rand des Streifens erst bei $x_0 + a$ erreichen.

c) Dies folgt aus Teil b), da die Lösung auf jedem Streifen rechts von x_0 existiert. ■

3.3 Gleichungen n-ter Ordnung

Was liefern die obigen Resultate für Gleichungen erster Ordnung nun für *Gleichungen n-ter Ordnung*? Wir betrachten dazu Gleichungen

$$y^{(n)} = f(x, y, y', \dots, y^{n-1})$$

mit $f : G \rightarrow \mathbb{C}^m$, $G \subset \mathbb{R} \times \mathbb{C}^{nm}$ offen. Man sagt in diesem Fall, f erfüllt (ggf. lokal) eine *Lipschitzbedingung* bezüglich y , wenn für

$$f(x, Y) := f(x, y_0, y_1, \dots, y_{n-1}), \quad Y = (y_0, y_1, \dots, y_{n-1})$$

gilt

$$|f(x, Y) - f(x, \tilde{Y})| \leq L|Y - \tilde{Y}|.$$

Satz 3.7 Sei $G \subset \mathbb{R} \times \mathbb{C}^{mn}$ offen, $f : G \rightarrow \mathbb{C}^m$ stetig und erfülle lokal eine Lipschitzbedingung bezüglich y .

a) Eindeutigkeit: Sind $y, z : I \rightarrow \mathbb{C}^m$ Lösungen von

$$y^{(n)} = f(x, y, y', \dots, y^{n-1}) \quad \text{mit} \quad y(x_0) = z(x_0), \dots, y^{(n-1)}(x_0) = z^{(n-1)}(x_0),$$

so gilt $y(x) = z(x)$ für alle $x \in I$.

b) Existenz: Ist $(x_0, y_{0,0}, y_{0,1}, \dots, y_{0,n-1}) \in G$, so existiert ein $\varepsilon > 0$ so, dass in $[x_0 - \varepsilon, x_0 + \varepsilon]$ genau eine Lösung des Anfangswertproblems

$$\text{AWP} \quad \begin{aligned} y^{(n)} &= f(x, y, y', \dots, y^{(n-1)}), \\ y^{(j)}(x_0) &= y_{0,j} \quad \text{für} \quad j = 0, 1, \dots, n-1 \quad (y^{(0)}(\cdot) = y(\cdot)) \end{aligned}$$

existiert.

(Wie für Gleichungen erster Ordnung lassen sich auch hier genauere Aussagen darüber machen, wie groß der Bereich ist, in dem die Lösung existiert, vgl. Satz 3.6.)

Beweis. Wir wissen, dass dieses Anfangswertproblem äquivalent ist zu folgendem Anfangswertproblem erster Ordnung für $Y = (y_0, \dots, y_{n-1})$

$$\text{AWP} \quad Y' = \tilde{f}(x, Y), \quad Y(x_0) = (y_{0,0}, y_{0,1}, \dots, y_{0,n-1}),$$

wobei \tilde{f} definiert ist durch

$$\tilde{f}(x, Y) = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_{n-1} \\ f(x, y_0, \dots, y_{n-1}) \end{pmatrix}.$$

Wegen

$$\begin{aligned} & \left| f(x, y_0, \dots, y_{n-1}) - f(x, \tilde{y}_0, \dots, \tilde{y}_{n-1}) \right| \leq |\tilde{f}(x, Y) - \tilde{f}(x, \tilde{Y})| \\ & \leq |y_1 - \tilde{y}_1| + \dots + |y_{n-1} - \tilde{y}_{n-1}| + \left| f(x, y_0, \dots, y_{n-1}) - f(x, \tilde{y}_0, \dots, \tilde{y}_{n-1}) \right| \\ & \leq \sqrt{n-1}|Y - \tilde{Y}| + \left| f(x, y_0, \dots, y_{n-1}) - f(x, \tilde{y}_0, \dots, \tilde{y}_{n-1}) \right| \end{aligned}$$

erfüllt offenbar \tilde{f} genau dann (lokal) eine Lipschitzbedingung bezüglich Y , wenn f (lokal) eine Lipschitzbedingung bezüglich $(y, y', \dots, y^{(n-1)})$ erfüllt. Damit ergibt sich die Behauptung aus den obigen Resultaten über Gleichungen erster Ordnung. ■

3.4 Zur stetigen Abhängigkeit von den Daten

In diesem kurzen Abschnitt soll ein einfaches Resultat über die stetige Abhängigkeit der Lösung einer Differenzialgleichung (erster Ordnung) $y' = f(x, y)$ von den „Daten“ bewiesen werden. Unter Daten wollen wir hier verstehen:

- die „rechte Seite“, d. h. die Funktion $f(\cdot, \cdot)$,
- den Anfangswert y_0 an der vorgegebenen (festen) Anfangsstelle x_0 .

(Natürlich könnte man auch die Anfangsstelle x_0 zu den Daten rechnen und die stetige Abhängigkeit von x_0 untersuchen. Ist die Anfangsstelle \tilde{x}_0 statt x_0 , so bedeutet dies nichts anderes als ein Ersetzen der rechten Seite durch $\tilde{f}(x, y) = f(x + (\tilde{x}_0 - x_0), y)$; wenn f z. B. gleichmäßig stetig bezüglich x ist, ist damit die stetige Abhängigkeit von x_0 im folgenden Resultat enthalten.)

Der folgende Satz sagt im wesentlichen aus: Wenn die rechte Seite $f(\cdot, \cdot)$ und der Anfangswert nur „wenig“ verändert werden, ändert sich auch die Lösung des Anfangswertproblems (zumindest in einem „kleinen“ Intervall um die Anfangsstelle x_0) nur „wenig“.

Wir betrachten die Anfangswertprobleme

$$\begin{aligned} y' &= f(x, y), & y(x_0) &= y_0, \\ z' &= g(x, z), & z(x_0) &= z_0. \end{aligned}$$

Die Voraussetzungen des Satzes sind so, dass das erste Anfangswertproblem auf dem zur Diskussion stehenden Intervall existiert und eindeutig bestimmt ist (Lipschitzbedingung für f). Falls das zweite Anfangswertproblem nicht eindeutig lösbar sein sollte, ist der Satz so zu verstehen, dass er für jede Lösung z gilt. Fordert man für g eine lokale Lipschitzbedingung, so ist auch das zweite Anfangswertproblem auf dem fraglichen Intervall eindeutig lösbar. Im übrigen sind die Voraussetzungen an f und g völlig symmetrisch, die Rollen von f und g können also ausgetauscht werden.

Satz 3.8 Sei $J \subset \mathbb{R}$ ein Intervall, $D \subset \mathbb{C}^m$ (oder \mathbb{R}^m) offen, $f, g : J \times D \rightarrow \mathbb{C}^m$ (oder \mathbb{R}^m), $x_0 \in J, a \in D, r, \varrho > 0$ mit $[x_0, x_0+r] \subset J$ und $K(a, 2\varrho) \subset D$. Für $x \in [x_0, x_0+r]$ und $y, z \in K(a, 2\varrho)$ gelte mit geeigneten $\beta, \gamma, \varepsilon > 0$

$$\begin{aligned} |f(x, y)| &\leq \beta, & |g(x, y)| &\leq \beta, \\ |f(x, y) - f(x, z)| &\leq \gamma|y - z|, \\ |f(x, y) - g(x, y)| &\leq \varepsilon. \end{aligned}$$

Dann gilt für y_0, z_0, x mit $y_0, z_0 \in K(a, \varrho)$ und $x_0 \leq x \leq x_0 + r_0$ mit $r_0 := \min\{r, \varrho/\beta\}$

$$|y(x) - z(x)| \leq |y_0 - z_0|e^{\gamma|x-x_0|} + \frac{\varepsilon}{\gamma} \left(e^{\gamma(x-x_0)} - 1 \right).$$

(Es sei angemerkt, dass der erste Term verschwindet, wenn die Anfangswerte gleich sind, der zweite dann, wenn die rechten Seiten gleich sind.) Die entsprechende Aussage gilt für ein Intervall links von x_0 .

Beweis. Das erste Anfangswertproblem ist auf Grund der Voraussetzungen an f auf dem Intervall $[x_0, x_0 + r_0]$ eindeutig lösbar mit $y(x) \in K(a, 2\varrho)$ für x aus diesem Intervall. Sofern das zweite Anfangswertproblem lösbar ist, gilt

$$|z(x) - z_0| \leq \left| \int_{x_0}^x g(t, z(t)) dt \right| \leq \beta|x - x_0|$$

und somit auch $z(x) \in K(a, 2\varrho)$ für $x \in [x_0, x_0 + r_0]$. Deshalb können alle Voraussetzungen an f und g benutzt werden. Wir erhalten zunächst

$$\begin{aligned} |y'(x) - z'(x)| &= \left| f(x, y(x)) - g(x, z(x)) \right| \\ &\leq \left| f(x, y(x)) - f(x, z(x)) \right| + \left| f(x, z(x)) - g(x, z(x)) \right| \\ &\leq \gamma|y(x) - z(x)| + \varepsilon, \end{aligned}$$

$$\begin{aligned} \frac{d}{dx}|y(x) - z(x)|^2 &= 2 \operatorname{Re} \left\langle y(x) - z(x), y'(x) - z'(x) \right\rangle \\ &\leq 2|y(x) - z(x)||y'(x) - z'(x)| \\ &\leq 2\left\{ \gamma|y(x) - z(x)| + \varepsilon \right\}|y(x) - z(x)| \\ &= 2\gamma|y(x) - z(x)|^2 + 2\varepsilon|y(x) - z(x)|, \end{aligned}$$

$$\frac{d}{dx}|y(x) - z(x)|^2 - 2\gamma|y(x) - z(x)|^2 \leq 2\varepsilon|y(x) - z(x)|.$$

Setzen wir nun

$$\psi(x) := e^{-\gamma(x-x_0)}|y(x) - z(x)|,$$

so erhalten wir

$$\begin{aligned} \psi(x)\psi'(x) &= \frac{1}{2} \frac{d}{dx} \psi(x)^2 = \frac{1}{2} \frac{d}{dx} \left\langle e^{-2\gamma(x-x_0)}|y(x) - z(x)|^2 \right\rangle \\ &= \frac{1}{2} e^{-2\gamma(x-x_0)} \left\{ \frac{d}{dx} |y(x) - z(x)|^2 - 2\gamma|y(x) - z(x)|^2 \right\} \\ &\leq e^{-2\gamma(x-x_0)} \varepsilon |y(x) - z(x)| \\ &= \varepsilon e^{-\gamma(x-x_0)} \psi(x). \end{aligned}$$

Für den Rest des Beweises müssen wir zwei Fälle unterscheiden:

Ist $\psi(x) \equiv 0$, so ist die Behauptung offensichtlich richtig.

Sei also $\psi(x) \neq 0$. Wir definieren

$$t_0 = t_0(x) := \inf \left\{ t \in (x_0, x) : \psi(s) > 0 \text{ für } s \in (t, x) \right\},$$

das ist das kleinste t_0 , für das ψ auf (t_0, x) positiv ist; wegen der Stetigkeit von ψ ist also $\psi(t_0) = 0$. In (t_0, x) kann die zuletzt bewiesene Ungleichung durch $\psi(x)$ dividiert werden, d. h. es gilt

$$\psi'(t) \leq \varepsilon e^{-\gamma(t-x_0)} \text{ für } t \in (t_0, x),$$

also nach Integration

$$\psi(x) = \psi(t_0) + \frac{\varepsilon}{\gamma} \left\{ e^{-\gamma(t_0-x_0)} - e^{-\gamma(x-x_0)} \right\}.$$

Multiplikation mit $e^{\gamma(x-x_0)}$ liefert

$$\begin{aligned} |y(x) - z(x)| &\leq e^{\gamma(x-x_0)} \psi(t_0) + \frac{\varepsilon}{\gamma} \left\{ e^{\gamma(x-t_0)} - 1 \right\} \\ &= \begin{cases} 0 + \frac{\varepsilon}{\gamma} \left\{ e^{\gamma(x-t_0)} - 1 \right\} & \text{falls } t_0 > x_0, \\ \left| y(x_0) - z(x_0) \right| e^{\gamma(x-x_0)} + \frac{\varepsilon}{\gamma} \left\{ e^{\gamma(x-x_0)} - 1 \right\} & \text{falls } t_0 = x_0 \end{cases} \\ &\leq \left| y_0 - z_0 \right| e^{\gamma(x-x_0)} + \frac{\varepsilon}{\gamma} \left\{ e^{\gamma(x-x_0)} - 1 \right\}, \end{aligned}$$

weil im 1. Fall natürlich $y_0 = z_0$ gelten muß. Das ist die behauptete Abschätzung. ■

7: 7. November 2006

Bemerkung 3.9 *Wie gut oder schlecht sind diese Abschätzungen?*

Explizite Berechnung der Beispiele

– $f(x, y) = \gamma y$ und $g(x, z) = \gamma z$ mit $y_0 \neq z_0$, also

$$\begin{aligned} y(x) &= y_0 e^{\gamma(x-x_0)}, & z(x) &= z_0 e^{\gamma(x-x_0)}, \\ |y(x) - z(x)| &= |y_0 - z_0| e^{\gamma(x-x_0)}. \end{aligned}$$

– $f(x, y) = \gamma y$ und $g(x, z) = \gamma z + \varepsilon$ mit $y_0 = z_0$

$$\begin{aligned} y(x) &= y_0 e^{\gamma(x-x_0)}, & z(x) &= y_0 e^{\gamma(x-x_0)} + \frac{\varepsilon}{\gamma} \left(e^{\gamma(x-x_0)} - 1 \right) \\ |y(x) - z(x)| &= \frac{\varepsilon}{\gamma} \left(e^{\gamma(x-x_0)} - 1 \right) \end{aligned}$$

zeigt, dass die im Satz angegebenen Abschätzungen, obwohl sie sehr grob erscheinen, optimal sind.

3.5 Übungsaufgaben

1. Man zeige (ohne die Differentialgleichung explizit zu lösen): Das Anfangswertproblem

$$y' = \left\{ (x-1)^2 + (y-3)^2 \right\}^{-1}, \quad y(0) = 0$$

ist auf ganz \mathbb{R} eindeutig lösbar.

2. Wir betrachten das Anfangswertproblem

$$y' = (1 - y^4)y = y - y^5, \quad y(0) = y_0.$$

- a) Für alle $y_0 \in \mathbb{R}$ ist das AWP eindeutig lösbar (womit noch nichts über den Existenzbereich ausgesagt ist).
 b) Man bestimme die Lösungen für $y_0 = -1, 0, 1$.
 c) Für $0 < |y_0| < 1$ ist das AWP auf ganz \mathbb{R} lösbar.
 d) Man skizziere die Lösungen aus den Teilen b und c.
 e) Das AWP ist für jedes $|y_0| > 1$ **nicht** auf ganz \mathbb{R} lösbar. Man skizziere auch diese Lösungen.
Anregung: Zum Vergleich $y' = -cy^5$ mit geeignetem $c > 0$.

3. Man zeige, dass das AWP

$$y' = (y-1)|y|^{1/2}, \quad y(0) = y_0 > 1$$

nicht für alle $x > 0$ lösbar ist.

Anleitung: Man vergleiche die Lösung mit der des AWP: $z' = (z-1)^{3/2}, z(0) = y_0$.

4. Ein Massepunkt mit Masse m und Koordinaten $x(t) = (\xi_1(t), \xi_2(t), \xi_3(t))$ bewegt sich in einem Kraftfeld, das ein zweimal stetig differenzierbares Potential $V: \mathbb{R}^3 \rightarrow \mathbb{R}$ besitzt:

$$mx'' = -\text{grad } V(x).$$

- a) Man transformiere dieses System in ein System erster Ordnung und zeige, dass dieses lokal eine Lipschitzbedingung erfüllt.
 b) Ist $x: J \rightarrow \mathbb{R}^3$ eine Lösung des Anfangswertproblems mit $x(t_0) = x_0, x'(t_0) = x_1 (t_0 \in J, x_0, x_1 \in \mathbb{R}^3)$, so gilt

$$V(x(t)) \leq \frac{m}{2}|x_1|^2 + V(x_0) \quad \text{für alle } t \in J.$$

Anleitung: Man zeige, dass die *Energie* $E(t) = \frac{m}{2}|x'(t)|^2 + V(x(t))$ konstant ist.

- c) Gilt $V(x) \rightarrow \infty$ für $x \rightarrow \infty$, so sind $x(t)$ und $x'(t)$ beschränkt.
 d) Ist $V(x) \geq c > -\infty$ für alle $x \in \mathbb{R}^3$, so ist jedes Anfangswertproblem auf ganz \mathbb{R} lösbar.

5. Man bestimme die Lösungsmatrix von

$$y' = \begin{pmatrix} 1 & 1 \\ 0 & -1 \end{pmatrix} y \quad \text{für } x_0 = 0$$

(mit Hilfe des Picard'schen Iterationsverfahrens).

4 Systeme linearer Differenzialgleichungen, allgemeiner Fall

In diesem Abschnitt betrachten wir Systeme linearer Differenzialgleichungen erster Ordnung

$$y' = A(x)y + b(x) \quad (x \in J \subset \mathbb{R})$$

mit einem beliebigen Intervall $J \subset \mathbb{R}$ und stetigen Funktionen

$$A : J \longrightarrow \mathbb{C}^{m \times m} \quad (= m \times m\text{-Matrizen}), \quad b : J \longrightarrow \mathbb{C}^m.$$

Gesucht sind also stetig differenzierbare Lösungen

$$y = (\eta_1, \dots, \eta_m)^t : J \longrightarrow \mathbb{C}^m.$$

Ausführlich geschrieben lautet das Differenzialgleichungssystem also

$$\begin{aligned} \eta_1'(x) &= \sum_{k=1}^m a_{1k}(x)\eta_k(x) + b_1(x), \\ &\vdots \\ \eta_m'(x) &= \sum_{k=1}^m a_{mk}(x)\eta_k(x) + b_m(x), \end{aligned}$$

bzw. etwas abgekürzt

$$\eta_j'(x) = \sum_{k=1}^m a_{jk}(x)\eta_k(x) + b_j(x) \quad (j = 1, \dots, m).$$

Wir betrachten zunächst den allgemeinen Fall, wo $A(\cdot)$ tatsächlich von x abhängt (variable Koeffizienten) und untersuchen dann im folgenden Abschnitt speziell den Fall, wo $A(\cdot)$ konstant ist (konstante Koeffizienten).

4.1 Die Lösungsmatrix

Die Existenz und Eindeutigkeit der Lösungen von Anfangswertproblemen ist in diesem Fall unter sehr allgemeinen Bedingungen beweisbar:

Satz 4.1 Sind $A : J \longrightarrow \mathbb{C}^{m \times m}$ und $b : J \longrightarrow \mathbb{C}^m$ stetig (d. h. alle $a_{jk}(\cdot)$ und $b_j(\cdot)$ sind stetig), $x_0 \in J, y_0 \in \mathbb{C}^m$, so ist das

$$\text{AWP} \quad y' = A(x)y + b(x), \quad y(x_0) = y_0$$

in ganz J eindeutig lösbar.

Beweis. Wir zeigen: Ist J_0 ein kompaktes Teilintervall von J und

$$G_0 := \{(x, y) \in \mathbb{R} \times \mathbb{C}^m : x \in J_0\} \subset G := \{(x, y) \in \mathbb{R} \times \mathbb{C}^m : x \in J\},$$

so erfüllt die rechte Seite

$$f(x, y) := A(x)y + b(x)$$

in G_0 eine Lipschitzbedingung (und somit erfüllt f lokal eine Lipschitzbedingung in G):

$$\begin{aligned} |f(x, y) - f(x, \tilde{y})| &= |A(x)(y - \tilde{y})| \\ &\leq \left\{ \sum_{j=1}^m \left| \sum_{k=1}^m a_{jk}(x)(\eta_k - \tilde{\eta}_k) \right|^2 \right\}^{1/2} \\ &\quad (\text{mit } L' := \max \{ |a_{jk}(x)| : 1 \leq j, k \leq m, x \in J_0 \}) \\ &\leq L' \left\{ \sum_{j=1}^m \left| \sum_{k=1}^m |\eta_k - \tilde{\eta}_k| \right|^2 \right\}^{1/2} \\ &\leq L' \left\{ \sum_{j=1}^m \left(\sum_{k=1}^m 1 \right) \sum_{k=1}^m |\eta_k - \tilde{\eta}_k|^2 \right\}^{1/2} \\ &= mL'|y - \tilde{y}| = L|y - \tilde{y}| \quad (\text{mit } L := mL'). \end{aligned}$$

Mit Satz 3.6 b) folgt, dass das AWP auf ganz J_0 eindeutig lösbar ist. Da J_0 ein beliebiges kompaktes Teilintervall von J ist, folgt damit die Behauptung. \blacksquare

Im folgenden untersuchen wir die Struktur der Lösungsmenge eines linearen Systems genauer:

Satz 4.2 *Das Matrixwertige Anfangswertproblem*

$$\text{AWP} \quad Y' = A(x)Y, \quad Y(x_0) = E = \begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{pmatrix}.$$

besitzt für jedes $x_0 \in J$ genau eine Lösung $Y : J \rightarrow \mathbb{C}^{m \times m}$. Für jedes $y_0 \in \mathbb{R}^m$ ist $y(x) := Y(x)y_0$ die (eindeutig bestimmte) Lösung des

$$\text{AWP} \quad y' = A(x)y, \quad y(x_0) = y_0.$$

Die Matrixfunktion $Y(\cdot)$ heißt die Lösungsmatrix des Systems (zum Anfangspunkt x_0); für jedes $x \in J$ ist $Y(x)$ invertierbar.

Beweis. Für jedes $j \in \{1, \dots, m\}$ sei $y_j(\cdot)$ die (eindeutig bestimmte) Lösung des

$$\text{AWP} \quad y_j' = A(x)y_j, \quad y_j(x_0) = e_j = (\delta_{1j}, \dots, \delta_{mj})^t.$$

$Y(x)$ sei die Matrix mit den Spalten $y_j(x)$,

$$Y(x) := \begin{pmatrix} y_1(x) & \vdots & y_2(x) & \vdots & \dots & \vdots & y_m(x) \end{pmatrix}.$$

¹Dies folgt natürlich (weniger explizit) auch aus dem früheren Satz 3.2, weil $y \mapsto A(x)y$ stetig nach y differenzierbar ist.

Offensichtlich ist $Y(\cdot)$ stetig differenzierbar, und es gilt

$$\begin{aligned} Y(x_0) &= \begin{pmatrix} e_1 \vdots e_2 \vdots \dots \vdots e_m \end{pmatrix} = E, \\ Y'(x) &= \begin{pmatrix} y'_1(x) \vdots y'_2(x) \vdots \dots \vdots y'_m(x) \end{pmatrix} \\ &= \begin{pmatrix} A(x)y_1(x) \vdots A(x)y_2(x) \vdots \dots \vdots A(x)y_m(x) \end{pmatrix} \\ &= A(x)Y(x), \end{aligned}$$

d. h. $Y(\cdot)$ ist die gesuchte matrixwertige Funktion; die Eindeutigkeit folgt daraus, dass die j -te Spalte von $Y(\cdot)$ notwendig die Lösung des AWP $y'_j = A(x)y_j$, $y_j(x_0) = e_j$ ist.

Weiter gilt

$$\begin{aligned} Y(x_0)y_0 &= Ey_0 = y_0, \\ (Y(x)y_0)' &= Y'(x)y_0 = (A(x)Y(x))y_0 = A(x)(Y(x)y_0), \end{aligned}$$

d. h. $y(x) := Y(x)y_0$ löst das AWP $y' = A(x)y$, $y(x_0) = y_0$.

Es bleibt die Invertierbarkeit von $Y(x)$ zu beweisen. Nehmen wir an, dass $Y(x_1)$ nicht invertierbar ist. Dann gibt es ein $y_0 \in \mathbb{R}^m \setminus \{0\}$ mit $Y(x_1)y_0 = 0$. Die Lösung des AWP $y' = A(x)y$, $y(x_0) = y_0$ erfüllt also bei x_1 die Anfangsbedingung $y(x_1) = 0$ und ist somit (Eindeutigkeitssatz) identisch gleich 0, im Widerspruch zu $y(x_0) = y_0 \neq 0$. ■

Um die Inverse der Lösungsmatrix differenzieren zu können, benötigen wir den folgenden

Hilfssatz 4.3 Sei $Y : J \rightarrow \mathbb{C}^{m \times m}$ stetig differenzierbar und $Y(x)$ invertierbar für alle $x \in J$. Dann ist $Z(\cdot) = Y(\cdot)^{-1}$ in J stetig differenzierbar und es gilt

$$Z'(x) = -Z(x)Y'(x)Z(x) = -Y(x)^{-1}Y'(x)Y(x)^{-1}.$$

(Das Resultat ist insbesondere auf Lösungsmatrizen anwendbar. Man vergleiche die obige Formel mit der Formel für die Ableitung der Inversen einer reellwertigen Funktion: $(1/y)'(x) = -y'(x)y(x)^{-2}$. Man beachte, dass $Y'(x)$ und $Y(x)^{-1}$ i. allg. nicht vertauschbar sind.)

Beweis. Wir beweisen zunächst die Stetigkeit von $Z : J \rightarrow \mathbb{C}^{m \times m}$. Nach der Cramerschen Regel hat jedes Matrixelement von $Z(\cdot)$ die Form

$$(\text{Det } Y(\cdot))^{-1} \times \{\text{Polynom in den Matrixelementen von } Y(\cdot)\}.$$

Also ist jedes Matrixelement von $Z(\cdot)$, und somit $Z(\cdot)$ selbst, stetig.

Damit folgt nun auch leicht die Differenzierbarkeit und die Formel für die Ableitung:

$$\begin{aligned} \lim_{h \rightarrow 0} \frac{1}{h} \left(Z(x+h) - Z(x) \right) &= \lim_{h \rightarrow 0} Z(x+h) \frac{1}{h} \left(Y(x) - Y(x+h) \right) Z(x) \\ &= -Z(x)Y'(x)Z(x). \end{aligned}$$

■

Korollar 4.4 Ist $Y(\cdot)$ die Lösungsmatrix von $y' = A(x)y$ (zum Anfangspunkt x_0), so ist $(Y(\cdot)^{-1})^t$ die Lösungsmatrix des transponierten Systems $Z' = -A(x)^t Z$. (Dabei ist A^t die zu A transponierte Matrix; es gilt offenbar $(AB)^t = B^t A^t$, $A'(x)^t = (A(x)^t)'$.)

Beweis. Sei wieder $Z(\cdot) = Y(\cdot)^{-1}$. Dann gilt

$$\begin{aligned} (Z(x)^t)' &= Z'(x)^t = -\left(Z(x)Y'(x)Z(x)\right)^t \\ &= -\left(Z(x)A(x)Y(x)Z(x)\right)^t = -\left(Z(x)A(x)\right)^t \\ &= -A(x)^t Z(x)^t, \\ Z(x_0)^t &= \left(Y(x_0)^{-1}\right)^t = E^t = E. \end{aligned}$$

■

4.2 Fundamentalsystem und Lösungsmatrix

Wir betrachten weiterhin das homogene System 1. Ordnung

$$y' = A(x)y, \quad x \in J, \quad y: J \longrightarrow \mathbb{C}^m.$$

Ein System von m Funktionen $\{y_1, \dots, y_m\}$ heißt ein *Fundamentalsystem* dieser Differentialgleichung, wenn gilt

- (a) Die Funktionen y_1, \dots, y_m sind Lösungen der Differentialgleichung,
- (b) die Funktionen y_1, \dots, y_m sind linear unabhängig, d. h. aus $\sum_{j=1}^m c_j y_j(x) = 0$ für alle $x \in J$ folgt $c_j = 0$ für $j = 1, \dots, m$.

Satz 4.5 Die Lösungen y_1, \dots, y_m des homogenen Systems bilden genau dann ein Fundamentalsystem, wenn eine der folgenden Bedingungen erfüllbar ist.

- (b') Für jedes $x \in J$ sind die Vektoren $y_1(x), \dots, y_m(x)$ linear unabhängig in \mathbb{C}^m , oder
- (b'') es gibt ein $x^* \in J$ so, dass die Vektoren $y_1(x^*), \dots, y_m(x^*)$ in \mathbb{C}^m linear unabhängig sind.

8: 10. November 2006

Beweis. Es ist die Äquivalenz $(b) \Leftrightarrow (b'')$ zu beweisen.

$(b) \implies (b'')$: Nehmen wir an, dass (b') nicht gilt, d. h. es gibt ein $x_0 \in J$ und $(c_1, \dots, c_m) \neq (0, \dots, 0)$ mit $\sum c_j y_j(x_0) = 0$. Dann ist $y(\cdot) = \sum c_j y_j(\cdot)$ die Lösung des AWP $y' = A(x)y$, $y(x_0) = \sum c_j y_j(x_0) = 0$, also $y(x) = 0$ für alle $x \in J$, d. h. (b) gilt nicht. — Die Implikationen $(b') \implies (b'')$ und $(b'') \implies (b)$ sind offensichtlich. ■

Die oben benutzten Lösungen $y_j(\cdot)$ mit $y_j(x_0) = e_j$ ($j = 1, \dots, m$) bilden offenbar ein spezielles Fundamentalsystem.

Die Lösungsmatrix läßt sich aber auch mit Hilfe eines beliebigen Fundamentalsystems erzeugen, wie der folgende Satz zeigt:

Satz 4.6 Ist $\{z_1, \dots, z_m\}$ ein (beliebiges) Fundamentalsystem von $y' = A(x)y$ und $Z(x)$ die Matrix mit den Spalten $z_1(x), \dots, z_m(x)$,

$$Z(x) := \begin{pmatrix} z_1(x) & \vdots & z_2(x) & \vdots & \dots & \vdots & z_m(x) \end{pmatrix},$$

so ist $Z(x)$ für jedes $x \in J$ invertierbar und es gilt

$$\begin{aligned} Z(x) &= Y(x)Z(x_0), \\ Y(x) &= Z(x)Z(x_0)^{-1}, \end{aligned}$$

wobei $Y(\cdot)$ die Lösungsmatrix zum Anfangspunkt x_0 ist.

Beweis. Auf Grund des obigen Resultats gilt offenbar

$$z_j(x) = Y(x)z_j(x_0).$$

Damit folgt sofort die erste Gleichung

$$Z(x) = Y(x)Z(x_0).$$

Da $\{z_1, \dots, z_m\}$ ein Fundamentalsystem ist, ist die Matrix

$$Z(x) = \begin{pmatrix} z_1(x) & \vdots & z_2(x) & \vdots & \dots & \vdots & z_m(x) \end{pmatrix}$$

invertierbar für jedes $x \in J$. Multiplikation der eben bewiesenen Identität mit $Z(x_0)^{-1}$ von rechts ergibt

$$Z(x)Z(x_0)^{-1} = Y(x).$$

■

Korollar 4.7 Sei $\{z_1, \dots, z_m\}$ ein (beliebiges) Fundamentalsystem. Dann ist jede Lösung der homogenen Gleichung $y' = A(x)y$ eine Linearkombination der Funktionen z_1, \dots, z_m , genauer:

$$\begin{aligned} y(x) &= Y(x)y_0 = Z(x)\left(Z(x_0)^{-1}y_0\right) = Z(x)(t_1, \dots, t_m)^t \\ &= \sum_{j=1}^m t_j z_j(x) \quad (t_j = j\text{-te Komponente von } Z(x_0)^{-1}y_0). \end{aligned}$$

Die Menge der Lösungen der homogenen Differentialgleichung $y' = A(x)y$ bildet also einen m -dimensionalen Vektorraum über \mathbb{C} . Jedes Fundamentalsystem ist eine Basis dieses Raumes.

4.3 Das inhomogene System

Nachdem wir einen vollständigen Überblick über die Lösungen des homogenen Systems $y' = A(x)y$ haben, wollen wir nun daran gehen, das inhomogene System

$$y' = A(x)y + b(x) \quad (b \neq 0)$$

zu untersuchen.

Satz 4.8 Sei $\{z_1, \dots, z_m\}$ ein (beliebiges) Fundamentalsystem der homogenen Gleichung $y' = A(x)y$, \hat{y} eine beliebige Lösung des inhomogenen Systems $y' = A(x)y + b(x)$. Dann hat die Lösung des inhomogenen Anfangswertproblems

$$\text{AWP} \quad y' = A(x)y + b(x), \quad y(x_0) = y_0$$

die Gestalt

$$y(x) = \hat{y}(x) + \sum_{j=1}^m c_j z_j(x) \quad (\text{mit geeigneten } c_j);$$

genauer gilt für das AWP $y(x_0) = y_0$

$$y(x) = \hat{y}(x) + Z(x)Z(x_0)^{-1}(y_0 - \hat{y}(x_0)).$$

Beweis. Da $Z(x)Z(x_0)^{-1}$ die Lösungsmatrix des homogenen Systems ist, ist klar, dass das zuletzt angegebene $y(\cdot)$ die inhomogene Gleichung löst. Weiter gilt

$$y(x_0) = \hat{y}(x_0) + Z(x_0)Z(x_0)^{-1}(y_0 - \hat{y}(x_0)) = y_0,$$

d. h. y hat die richtigen Anfangswerte. Wegen der Eindeutigkeit muß also die Lösung diese Gestalt haben. Da $Z(x)\left(Z(x_0)^{-1}(y_0 - \hat{y}(x_0))\right)$ eine Linearkombination der z_1, \dots, z_m ist, ist damit alles bewiesen. ■

Eine explizite Lösungsformel, die nicht die Kenntnis einer speziellen Lösung der inhomogenen Gleichung erfordert, gibt der folgende Satz (die im Abschnitt 2.5 gefundene Formel für die Lösung **einer** inhomogenen Differenzialgleichung erster Ordnung ($m = 1$) ist ein Spezialfall hiervon).

Satz 4.9 Die eindeutig bestimmte Lösung y des

$$\text{AWP} \quad y' = A(x)y + b(x), \quad y(x_0) = y_0$$

ist gegeben durch

$$y(x) = Y(x)y_0 + Y(x) \int_{x_0}^x Y(t)^{-1}b(t) dt$$

(wobei $Y(\cdot)$ die Lösungsmatrix des homogenen Systems zum Anfangspunkt x_0 ist).

Beweis. Es gilt offenbar

$$\begin{aligned}
 y(x_0) &= Y(x_0)y_0 + Y(x_0) \int_{x_0}^{x_0} Y(t)^{-1}b(t) dt \\
 &= Ey_0 + E0 = y_0, \\
 y'(x) &= Y'(x)y_0 + Y'(x) \int_{x_0}^x Y(t)^{-1}b(t) dt + Y(x)Y(x)^{-1}b(x) \\
 &= A(x) \left\{ Y(x)y_0 + Y(x) \int_{x_0}^x Y(t)^{-1}b(t) dt \right\} + b(x) \\
 &= A(x)y(x) + b(x).
 \end{aligned}$$

■

Der Beweis des letzten Satzes ist zwar sehr einfach, verrät aber nicht, wie man diese Lösung findet. Deshalb sei hier noch ein anderes Vorgehen beschrieben: Die Spaltenvektoren $y_j(x)$ der Lösungsmatrix $Y(x) = (y_1(x) \dotscolumn{m}{\dotscolumn{m}} y_m(x))$ bilden für jedes $x \in J$ eine Basis in \mathbb{C}^m (vgl. obige Charakterisierung eines Fundamentalsystems). Es gibt also für die Lösung $y(\cdot)$ des inhomogenen Systems und für jedes $x \in J$ einen Vektor

$$y_0(x) = \left(\eta_{0,1}(x), \dots, \eta_{0,m}(x) \right)^t \in \mathbb{R}^m$$

mit

$$\begin{aligned}
 y(x) &= \sum_{j=1}^m \eta_{0,j}(x) y_j(x) = Y(x)y_0(x), \\
 y_0(x) &= Y(x)^{-1}y(x).
 \end{aligned}$$

Existiert eine stetig differenzierbare Lösung $y(\cdot)$, so ist also wegen $y_0(x) = Y(x)^{-1}y(x)$ auch $y_0(\cdot)$ stetig differenzierbar und wir erhalten durch Einsetzen von $y(x) = Y(x)y_0(x)$ in die inhomogene Differenzialgleichung

$$\begin{aligned}
 A(x)Y(x)y_0(x) + b(x) &= A(x)y(x) + b(x) = y'(x) = Y'(x)y_0(x) + Y(x)y_0'(x) \\
 &= A(x)Y(x)y_0(x) + Y(x)y_0'(x).
 \end{aligned}$$

Hieraus ergibt sich $Y(x)y_0'(x) = b(x)$ und somit die Gleichung für $y_0(\cdot)$

$$y_0'(x) = Y(x)^{-1}b(x),$$

aus der durch Integration folgt:

$$y_0(x) = y_0 + \int_{x_0}^x Y(t)^{-1}b(t) dt.$$

Einsetzen in $y(x) = Y(x)y_0(x)$ liefert die bereits oben bewiesene Lösungsformel.

Formal haben wir oben in der allgemeinen Lösung der homogenen Gleichung $Y(x)y_0$ den konstanten Vektor y_0 durch einen variablen Vektor $y_0(x)$ ersetzt, und es stellte sich heraus, dass $y_0(\cdot)$ tatsächlich berechnet werden kann. Man nennt deshalb diese Methode die *Variation der Konstanten*.

4.4 Übungsaufgaben

1. a) Man bestimme die Lösungsmatrix von

$$y' = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} y \quad \text{für } x_0 = 0.$$

- b) Man löse das Anfangswertproblem

$$y' = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} y + \begin{pmatrix} x-1 \\ 2x \end{pmatrix}, \quad x(0) = \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

Hilfe: $\begin{pmatrix} -x \\ 0 \end{pmatrix}$ ist eine Lösung des inhomogenen Systems.

5 Systeme linearer Differenzialgleichungen mit konstanten Koeffizienten

In diesem Abschnitt betrachten wir das homogene lineare Differenzialgleichungssystem $y' = Ay$, wobei jetzt A eine konstante $m \times m$ -Matrix ist.

5.1 Die Lösungsmatrix $\exp(\mathbf{x}A)$

Um die Lösungsmatrix des Systems zu bestimmen, lösen wir das matrixwertige Anfangswertproblem

$$\text{AWP} \quad Y' = AY, \quad Y(x_0) = E$$

mit Hilfe der Picard'schen Iteration:

$$Y_0(x) = E, \\ Y_{k+1}(x) = E + \int_{x_0}^x AY_k(t) dt \quad \text{für } k = 0, 1, 2, \dots$$

Damit erhalten wir explizit

$$Y_1(x) = E + \int_{x_0}^x AE dt = E + (x - x_0)A, \\ Y_2(x) = E + \int_{x_0}^x A(E + (t - x_0)A) dt = E + (x - x_0)A + \frac{(x - x_0)^2}{2}A^2,$$

und weiter durch Induktion (mit $A^0 = E$)

$$Y_{k+1}(x) = E + \int_{x_0}^x A \sum_{j=0}^k \frac{(t - x_0)^j}{j!} A^j dt = \sum_{j=0}^{k+1} \frac{(x - x_0)^j}{j!} A^j.$$

Wie man aus dem Beweis des Existenzsatzes von Picard–Lindelöf (oder aber auch mit dem gleichen Beweis, mit dem man die Konvergenz der Exponentialreihe beweist) sieht, ist diese Folge auf jedem kompakten Teilintervall von \mathbb{R} gleichmäßig konvergent gegen eine stetige matrixwertige Funktion

$$Y(x) := \sum_{j=0}^{\infty} \frac{(x - x_0)^j}{j!} A^j,$$

die das obige Anfangswertproblem löst, d. h. $Y(\cdot)$ ist die gesuchte Lösungsmatrix.

Die formale Ähnlichkeit mit der Exponentialreihe legt die folgende Schreibweise nahe

$$e^{xA} = \exp(xA) := \sum_{j=0}^{\infty} \frac{x^j}{j!} A^j.$$

Mit dieser Bezeichnung haben wir

$$Y(x) = \exp \left\{ (x - x_0)A \right\}.$$

Da Y die homogene Differenzialgleichung $Y' = AY$ erfüllt, folgt

$$\begin{aligned} \frac{d}{dx} \exp \{ (x - x_0)A \} &= Y'(x) = AY(x) \\ &= A \exp \{ (x - x_0)A \}, \end{aligned}$$

erneut eine Eigenschaft, die von einer Exponentialfunktion zu erwarten ist.

Beispiel 5.1 für eine Diagonalmatrix $A = \begin{pmatrix} \lambda_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \lambda_m \end{pmatrix}$ ist offenbar

$$e^{xA} = \begin{pmatrix} e^{\lambda_1 x} & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & e^{\lambda_m x} \end{pmatrix}$$

□

Beispiel 5.2 Wir wollen für $A = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}$ (einen 2×2 -Jordan-Kasten) e^{xA} berechnen. Zunächst sieht man leicht

$$A^2 = \begin{pmatrix} \lambda^2 & 2\lambda \\ 0 & \lambda^2 \end{pmatrix}, \quad A^3 = \begin{pmatrix} \lambda^3 & 3\lambda^2 \\ 0 & \lambda^3 \end{pmatrix},$$

und durch Induktion allgemein

$$A^j = \begin{pmatrix} \lambda^j & j\lambda^{j-1} \\ 0 & \lambda^j \end{pmatrix} \quad \text{für } j \in \mathbb{N}_0.$$

Somit erhalten wir

$$\begin{aligned} e^{xA} &= \sum_{j=0}^{\infty} \frac{x^j}{j!} A^j = \sum_{j=0}^{\infty} \frac{x^j}{j!} \begin{pmatrix} \lambda^j & j\lambda^{j-1} \\ 0 & \lambda^j \end{pmatrix} \\ &= \sum_{j=0}^{\infty} \frac{(x\lambda)^j}{j!} E + \sum_{j=1}^{\infty} \frac{(x\lambda)^{j-1}}{(j-1)!} \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \\ &= \begin{pmatrix} e^{x\lambda} & xe^{x\lambda} \\ 0 & e^{x\lambda} \end{pmatrix}. \end{aligned}$$

□

9: 14. November 2006

5.2 $\exp(xA)$ für diagonalisierbares A

Besonders einfach werden die Verhältnisse, wenn A diagonalisierbar ist, d. h. wenn es eine invertierbare Matrix M gibt so, dass $M^{-1}AM$ Diagonalgestalt hat,

$$M^{-1}AM = \text{Diag}(\lambda_1, \dots, \lambda_m) = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_m \end{pmatrix},$$

wobei dann natürlich $\lambda_1, \dots, \lambda_m$ die Eigenwerte von A sind (mit Vielfachheit gezählt). Es sei hier daran erinnert, dass A genau dann diagonalisierbar ist, wenn A m linear unabhängige Eigenvektoren besitzt; $\text{Diag}(\lambda_1, \dots, \lambda_m)$ ist dann gerade die Darstellung von A bezüglich der Basis, die durch diese Eigenvektoren gebildet wird. A hat sicher dann m linear unabhängige Eigenvektoren, wenn A m verschiedene Eigenwerte hat.

Für diagonalisierbares A gilt offenbar

$$A^j = (MDM^{-1})^j = MD^jM^{-1}$$

und somit

$$\begin{aligned} e^{xA} &= \sum_{j=0}^{\infty} \frac{x^j}{j!} A^j = M \left\{ \sum_{j=0}^{\infty} \frac{x^j}{j!} D^j \right\} M^{-1} \\ &= M \begin{pmatrix} e^{x\lambda_1} & & 0 \\ & \ddots & \\ 0 & & e^{x\lambda_m} \end{pmatrix} M^{-1}. \end{aligned}$$

M und M^{-1} können wie folgt explizit angegeben werden: Ist $\{a_1, \dots, a_m\}$ eine Basis aus Eigenvektoren und ist $\{a'_1, \dots, a'_m\}$ eine dazu *duale Basis*, d. h. es gilt

$$\langle a'_j, a_k \rangle = \delta_{jk} \quad \text{für } j, k = 1, \dots, m,$$

wobei $\langle \cdot, \cdot \rangle$ das Skalarprodukt in \mathbb{C}^m ist, d. h. $\langle x, y \rangle = \sum_{j=1}^m \bar{x}_j y_j$, so ist

$$M = (a_1 \mid a_2 \mid \dots \mid a_m), \quad M^{-1} = \begin{pmatrix} \overline{a'_1} \\ \overline{a'_2} \\ \vdots \\ \overline{a'_m} \end{pmatrix},$$

d. h. M ist die Matrix mit den Spalten a_1, \dots, a_m und M^{-1} ist die Matrix mit den Zeilenvektoren $\overline{a'_1}, \dots, \overline{a'_m}$, den zu a_1, \dots, a_m dualen Vektoren.

Für die Lösungsmatrix erhält man damit

$$Y(x) = e^{(x-x_0)A} = M \begin{pmatrix} e^{(x-x_0)\lambda_1} & & 0 \\ & \ddots & \\ 0 & & e^{(x-x_0)\lambda_m} \end{pmatrix} M^{-1}.$$

Das Anfangswertproblem

$$\text{AWP} \quad y' = Ay, \quad y(x_0) = y_0,$$

hat also die Lösung

$$\begin{aligned} y(x) &= Y(x)y_0 = M \begin{pmatrix} e^{(x-x_0)\lambda_1} & & 0 \\ & \ddots & \\ 0 & & e^{(x-x_0)\lambda_m} \end{pmatrix} M^{-1}y_0 \\ &= M \begin{pmatrix} e^{(x-x_0)\lambda_1} & & 0 \\ & \ddots & \\ 0 & & e^{(x-x_0)\lambda_m} \end{pmatrix} \begin{pmatrix} \langle a'_1, y_0 \rangle \\ \langle a'_2, y_0 \rangle \\ \vdots \\ \langle a'_m, y_0 \rangle \end{pmatrix} \end{aligned}$$

$$= M \begin{pmatrix} e^{(x-x_0)\lambda_1} \langle a'_1, y_0 \rangle \\ e^{(x-x_0)\lambda_2} \langle a'_2, y_0 \rangle \\ \vdots \\ e^{(x-x_0)\lambda_m} \langle a'_m, y_0 \rangle \end{pmatrix} = \sum_{j=1}^m \langle a'_j, y_0 \rangle e^{(x-x_0)\lambda_j} a_j.$$

Ist speziell A hermitesch, so hat es in jedem Fall m linear unabhängige Eigenvektoren a_1, \dots, a_m , die sogar o. E. als Orthonormalbasis von \mathbb{C}^m gewählt werden können. Dann ist offenbar $\{a'_1, \dots, a'_m\} = \{a_1, \dots, a_m\}$ und somit

$$M^{-1} = \overline{M^t} = M^*, \quad \text{d. h. } M \text{ ist unitär.}$$

Für $y(\cdot)$ erhalten wir dann

$$y(x) = \sum_{j=1}^m \langle a_j, y_0 \rangle e^{(x-x_0)\lambda_j} a_j.$$

Insgesamt haben wir bewiesen:

Satz 5.3 a) Sei A eine $m \times m$ -Matrix mit m linear unabhängigen Eigenvektoren a_1, \dots, a_m und zugehörigen Eigenwerten $\lambda_1, \dots, \lambda_m$, außerdem sei $\{a'_1, \dots, a'_m\}$ eine zu $\{a_1, \dots, a_m\}$ duale Basis. Dann ist die Lösung des

$$\text{AWP} \quad y' = Ay, \quad y(x_0) = y_0$$

gegeben durch

$$y(x) = \sum_{j=1}^m \langle a'_j, y_0 \rangle e^{(x-x_0)\lambda_j} a_j.$$

b) Ist speziell A hermitesch und a_1, \dots, a_m eine Orthonormalbasis von Eigenelementen, so gilt

$$y(x) = \sum_{j=1}^{\infty} \langle a_j, y_0 \rangle e^{(x-x_0)\lambda_j} a_j.$$

Dies sind offenbar Linearkombinationen der Lösungen der Anfangswertprobleme mit $y(x_0) = a_j$ ($j = 1, \dots, m$), wobei die Entwicklungskoeffizienten durch $\langle a'_j, y_0 \rangle$ bzw. $\langle a_j, y_0 \rangle$ gegeben sind. Für $x = x_0$ hat man gerade die Entwicklung von y_0 nach der Basis $\{a_1, \dots, a_m\}$.

Beispiel 5.4 Wir betrachten das System

$$y' = \begin{pmatrix} 1 & -1 & 2 \\ -1 & 1 & 2 \\ 1 & 1 & 0 \end{pmatrix} y, \quad \text{also } y' = Ay \quad \text{mit } A = \begin{pmatrix} 1 & -1 & 2 \\ -1 & 1 & 2 \\ 1 & 1 & 0 \end{pmatrix}.$$

Aus (Entwicklung nach der 3. Zeile)

$$\begin{aligned} \det(A - \lambda E) &= \det \begin{pmatrix} 1 - \lambda & -1 & 2 \\ -1 & 1 - \lambda & 2 \\ 1 & 1 & -\lambda \end{pmatrix} = -\lambda^3 + 2\lambda^2 + 4\lambda - 8 \\ &= -(\lambda - 2)^2(\lambda + 2) \end{aligned}$$

folgt, dass $\lambda = 2$ und $\lambda = -2$ Eigenwerte sind, $\lambda = -2$ einfach, $\lambda = 2$ mit (zunächst nur) algebraischer Vielfachheit 2. Den Eigenvektor zu $\lambda = -2$ erhalten wir aus

$$(A + 2E) \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} 3 & -1 & 2 \\ -1 & 3 & 2 \\ 1 & 1 & 2 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = 0, \quad \begin{array}{l} 3a - b + 2c = 0, \\ -a + 3b + 2c = 0, \\ a + b + 2c = 0. \end{array}$$

Da sich die dritte Gleichung aus den ersten beiden ergibt, bleibt das System zu lösen, das aus den ersten 2 Gleichungen besteht. Es hat (z. B.) die Lösung

$$\begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}.$$

Dies ist (bis auf einen Faktor) **der** Eigenvektor zum Eigenwert -2 .

Die Eigenvektoren zu $\lambda = 2$ erhalten wir aus

$$(A - 2E) \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} -1 & -1 & 2 \\ -1 & -1 & 2 \\ 1 & 1 & -2 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = 0, \quad \begin{array}{l} -a - b + 2c = 0, \\ -a - b + 2c = 0, \\ a + b - 2c = 0. \end{array}$$

Dieses System ist äquivalent zu der einzigen Gleichung

$$a + b - 2c = 0$$

mit (z. B.) den Lösungen

$$\begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} \quad \text{und} \quad \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.$$

Dies sind 2 Eigenvektoren zum Eigenwert 2. Insgesamt hat also A eine Basis von Eigenvektoren a_1, a_2, a_3 zu den Eigenwerten $\lambda_1 = -2, \lambda_{2,3} = 2$:

$$a_1 = \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}, \quad a_2 = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \quad a_3 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

(da diese nicht orthogonal sind, würde Normierung keinerlei Vorteil bringen). Damit haben wir ein Fundamentalsystem des Differenzialgleichungssystems:

$$\begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix} e^{-2x}, \quad \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} e^{2x}, \quad \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} e^{2x}.$$

Die zu den Eigenvektoren a_1, a_2, a_3 gehörige duale Basis ist

$$a'_1 = \frac{1}{4} \begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix}, \quad a'_2 = \frac{1}{2} \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \quad a'_3 = \frac{1}{4} \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix}.$$

Es ist also

$$M = \begin{pmatrix} 1 & 1 & 1 \\ 1 & -1 & 1 \\ -1 & 0 & 1 \end{pmatrix}, \quad M^{-1} = \begin{pmatrix} 1/4 & 1/4 & -1/2 \\ 1/2 & -1/2 & 0 \\ 1/4 & 1/4 & 1/2 \end{pmatrix}$$

und somit die Lösungsmatrix

$$\begin{aligned} Y(x) = e^{xA} &= M \begin{pmatrix} e^{2x} & 0 & 0 \\ 0 & e^{2x} & 0 \\ 0 & 0 & e^{-2x} \end{pmatrix} M^{-1} \\ &= \begin{pmatrix} \frac{1}{2}(e^{2x} - e^{-2x}) & 0 & \frac{1}{2}(e^{2x} + e^{-2x}) \\ -\frac{1}{2}e^{-2x} & e^{2x} & \frac{1}{2}e^{-2x} \\ -\frac{1}{4}e^{2x} - \frac{1}{2}e^{-2x} & -\frac{1}{2}e^{2x} & -\frac{1}{4}e^{2x} + \frac{1}{2}e^{-2x} \end{pmatrix}, \end{aligned}$$

zusammen mit

$$\begin{aligned} Y(x)^{-1} &= M \begin{pmatrix} e^{-2x} & & 0 \\ & e^{-2x} & \\ 0 & & e^{2x} \end{pmatrix} M^{-1} \\ &= \begin{pmatrix} \frac{1}{2}(e^{-2x} - e^{2x}) & 0 & \frac{1}{2}(e^{-2x} + e^{2x}) \\ -\frac{1}{2}e^{2x} & e^{-2x} & \frac{1}{2}e^{2x} \\ -\frac{1}{4}e^{-2x} - \frac{1}{2}e^{2x} & -\frac{1}{2}e^{-2x} & -\frac{1}{4}e^{-2x} + \frac{1}{2}e^{2x} \end{pmatrix} \end{aligned}$$

ist dann jedes Anfangswertproblem für dieses System lösbar. \square

5.3 Ein Hamiltonsches System; Stabilität, Instabilität

Wir betrachten ein mechanisches System mit m Freiheitsgraden mit

$$\begin{aligned} \text{Koordinaten} & \quad x_1, \dots, x_m, \\ \text{Kinetischer Energie} & \quad T = \frac{1}{2} \sum_{j=1}^m (\dot{x}_j)^2 = \frac{1}{2} |\dot{x}|^2, \\ \text{Potentieller Energie} & \quad U = \frac{1}{2} \langle x, Bx \rangle \end{aligned}$$

mit einer symmetrischen $m \times m$ -Matrix B . Mit den

$$\text{Impulskoordinaten} \quad p_j = \dot{x}_j \quad (j = 1, \dots, m)$$

ist dann die *Hamiltonsche Funktion*

$$H(p, x) = T + U = \frac{1}{2} \{ |p|^2 + \langle x, Bx \rangle \}.$$

Damit erhalten wir die *Hamiltonschen Differenzialgleichungen*

$$\begin{aligned} \dot{x}_j &= \frac{\partial}{\partial p_j} H(p, x) = p_j, \\ \dot{p}_j &= -\frac{\partial}{\partial x_j} H(p, x) = -(Bx)_j \end{aligned} \quad (j = 1, \dots, m),$$

(wobei $(Bx)_j$ die j -te Komponente von Bx ist) oder, kompakter geschrieben,

$$\dot{x} = p, \quad \dot{p} = -Bx.$$

Mit der neuen Funktion $y : \mathbb{R} \rightarrow \mathbb{R}^{2m}$

$$y(t) = \left(x_1(t), \dots, x_m(t), p_1(t), \dots, p_m(t) \right)$$

erhalten wir das Differenzialgleichungssystem

$$y' = Ay \quad \text{mit der Matrix} \quad A = \begin{pmatrix} 0 & I \\ -B & 0 \end{pmatrix}.$$

Aus

$$A^1 = \begin{pmatrix} 0 & I \\ -B & 0 \end{pmatrix} \quad \text{und} \quad A^2 = \begin{pmatrix} -B & 0 \\ 0 & -B \end{pmatrix}$$

folgt durch Induktion für alle $k \in \mathbb{N}_0$

$$A^{2k} = \begin{pmatrix} (-B)^k & 0 \\ 0 & (-B)^k \end{pmatrix}, \quad A^{2k+1} = \begin{pmatrix} 0 & (-B)^k \\ (-B)^{k+1} & 0 \end{pmatrix},$$

und somit

$$e^{tA} = \begin{pmatrix} \sum_{n=0}^{\infty} \frac{(-1)^n t^{2n}}{(2n)!} B^n & \sum_{n=0}^{\infty} \frac{(-1)^n t^{2n+1}}{(2n+1)!} B^n \\ \sum_{n=0}^{\infty} \frac{(-1)^{n+1} t^{2n+1}}{(2n+1)!} B^{n+1} & \sum_{n=0}^{\infty} \frac{(-1)^n t^{2n}}{(2n)!} B^n \end{pmatrix}.$$

Da B symmetrisch ist, gilt mit einer Orthonormalbasis $\{b_1, \dots, b_m\}$ von Eigenelementen,

$$M = (b_1 \mid b_2 \mid \dots \mid b_m), \quad M^{-1} = \begin{pmatrix} \overline{b_1} \\ \vdots \\ \overline{b_m} \end{pmatrix},$$

und, mit zugehörigen Eigenwerten $\lambda_1, \dots, \lambda_m$,

$$B = M \text{Diag}(\lambda_1, \dots, \lambda_m) M^{-1}.$$

Im folgenden betrachten wir zwei spezielle Fälle im Detail.

Stabiler Fall: Es sei $U(x) > 0$ für $x \neq 0$, d. h. A ist positiv definit, $\lambda_j > 0$ für $j = 1, \dots, m$. Sei $\nu_j > 0$ so, dass $\lambda_j = \nu_j^2$ gilt. Dann ist

$$\begin{aligned} M^{-1} \left\{ \sum_{n=0}^{\infty} \frac{(-1)^n t^{2n}}{(2n)!} B^n \right\} M &= \sum_{n=0}^{\infty} \frac{(-1)^n t^{2n}}{(2n)!} \begin{pmatrix} \nu_1^{2n} & & 0 \\ & \ddots & \\ 0 & & \nu_m^{2n} \end{pmatrix} \\ &= \begin{pmatrix} \sum_{n=0}^{\infty} \frac{(-1)^n t^{2n}}{(2n)!} \nu_1^{2n} & & 0 \\ & \ddots & \\ 0 & & \sum_{n=0}^{\infty} \frac{(-1)^n t^{2n}}{(2n)!} \nu_m^{2n} \end{pmatrix} \\ &= \begin{pmatrix} \cos \nu_1 t & & 0 \\ & \ddots & \\ 0 & & \cos \nu_m t \end{pmatrix}, \end{aligned}$$

also

$$\sum_{n=0}^{\infty} \frac{(-1)^n t^{2n}}{(2n)!} B^n = M \begin{pmatrix} \cos \nu_1 t & & 0 \\ & \ddots & \\ 0 & & \cos \nu_m t \end{pmatrix} M^{-1}.$$

Entsprechend folgt

$$\begin{aligned} \sum_{n=0}^{\infty} \frac{(-1)^n t^{2n+1}}{(2n+1)!} B^n &= M \begin{pmatrix} \frac{1}{\nu_1} \sin \nu_1 t & & 0 \\ & \ddots & \\ 0 & & \frac{1}{\nu_m} \sin \nu_m t \end{pmatrix} M^{-1}, \\ \sum_{n=0}^{\infty} \frac{(-1)^{n+1} t^{2n+1}}{(2n+1)!} B^{n+1} &= M \begin{pmatrix} -\nu_1 \sin \nu_1 t & & 0 \\ & \ddots & \\ 0 & & -\nu_m \sin \nu_m t \end{pmatrix} M^{-1}, \end{aligned}$$

und damit ist auch e^{tA} explizit gegeben.

10: 17. November 2006

$$e^{ta} = \begin{pmatrix} M \begin{pmatrix} \cos \nu_1 t & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \cos \nu_m t \end{pmatrix} M^{-1} & M \begin{pmatrix} \frac{1}{\nu_1} \sin \nu_1 t & \dots & 0 \\ \vdots & \dots & \vdots \\ 0 & \dots & \frac{1}{\nu_m} \sin \nu_m t \end{pmatrix} M^{-1} \\ M \begin{pmatrix} -\nu_1 \sin \nu_1 t & \dots & 0 \\ \vdots & \dots & \vdots \\ 0 & \dots & -\nu_1 \sin \nu_1 t \end{pmatrix} M^{-1} & M \begin{pmatrix} \cos \nu_1 t & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \cos \nu_m t \end{pmatrix} M^{-1} \end{pmatrix}$$

Die Lösung des

$$\text{AWP} \quad y' = Ay, \quad y(0) = \begin{pmatrix} x_0 \\ p_0 \end{pmatrix} \in \mathbb{R}^{2m}$$

ist gegeben durch

$$y(t) = e^{tA} \begin{pmatrix} x_0 \\ p_0 \end{pmatrix}.$$

Dabei berechnen wir zunächst nur die ersten m (Orts-) Komponenten $x(t) = (y_1(t), \dots, y_m(t))$,

$$\begin{aligned} x(t) &= \sum_{n=0}^{\infty} \frac{(-1)^n t^{2n}}{(2n)!} B^n x_0 + \sum_{n=0}^{\infty} \frac{(-1)^n t^{2n+1}}{(2n+1)!} B^n p_0 \\ &= M \begin{pmatrix} \cos \nu_1 t & & 0 \\ & \ddots & \\ 0 & & \cos \nu_m t \end{pmatrix} M^{-1} x_0 + M \begin{pmatrix} \frac{1}{\nu_1} \sin \nu_1 t & & 0 \\ & \ddots & \\ 0 & & \frac{1}{\nu_m} \sin \nu_m t \end{pmatrix} M^{-1} p_0. \end{aligned}$$

Mit $M = (b_1 \vdots b_2 \vdots \dots \vdots b_m)$ und $M^{-1} = M^t$ folgt hieraus

$$x(t) = \sum_{j=1}^m \left\{ \langle b_j, x_0 \rangle \cos \nu_j t + \langle b_j, p_0 \rangle \frac{1}{\nu_j} \sin \nu_j t \right\} b_j.$$

Die letzten m (Geschwindigkeits-) Komponenten $p(t) = (y_{m+1}(t), \dots, y_{2m}(t))$ erhält man entsprechend, oder einfacher durch Differenzieren dieses Ausdrucks:

$$\dot{x}(t) = \sum_{j=1}^m \left\{ -\langle b_j, x_0 \rangle \nu_j \sin \nu_j t + \langle b_j, p_0 \rangle \cos \nu_j t \right\} b_j.$$

Die Lösung ist *stabil* weil $y(t)$ (bzw. $x(t)$) **und** $p(t)$ beschränkt bleiben für alle t .

Instabiler Fall: Mindestens ein Eigenwert von B ist negativ. Wir nehmen hier speziell an:

$$\lambda_1 < 0 < \lambda_j \quad \text{für } j = 2, \dots, m.$$

Es ist dann

$$\begin{aligned} \text{Diag}(\lambda_1, \dots, \lambda_m) &= \begin{pmatrix} -\nu_1^2 & & & 0 \\ & \nu_2^2 & & \\ & & \ddots & \\ 0 & & & \nu_m^2 \end{pmatrix}, \\ \text{Diag}(\lambda_1, \dots, \lambda_m)^n &= \begin{pmatrix} (-1)^n \nu_1^{2n} & & & 0 \\ & \nu_2^{2n} & & \\ & & \ddots & \\ 0 & & & \nu_m^{2n} \end{pmatrix}, \end{aligned}$$

und somit

$$\begin{aligned} \sum_{n=0}^{\infty} \frac{(-1)^n t^{2n}}{(2n)!} B^n x_0 &= M \sum_{k=0}^{\infty} \frac{(-1)^k t^{2k}}{(2k)!} \text{Diag}((-1)^k \nu_1^{2k}, \nu_2^{2k}, \dots, \nu_m^{2k}) M^{-1} \\ &= M \begin{pmatrix} \cosh \nu_1 t & & & 0 \\ & \cos \nu_2 t & & \\ & & \ddots & \\ 0 & & & \cos \nu_m t \end{pmatrix} M^{-1}, \\ \sum_{n=0}^{\infty} \frac{(-1)^n t^{2n+1}}{(2n+1)!} B^n &= M \sum_{n=0}^{\infty} \frac{(-1)^n t^{2n+1}}{(2n+1)!} \text{Diag}((-1)^n \nu_1^{2n}, \nu_2^{2n}, \dots, \nu_m^{2n}) M^{-1} \\ &= M \begin{pmatrix} \frac{1}{\nu_1} \sinh \nu_1 t & & & 0 \\ & \frac{1}{\nu_2} \sin \nu_2 t & & \\ & & \ddots & \\ 0 & & & \frac{1}{\nu_m} \sin \nu_m t \end{pmatrix} M^{-1}, \\ \sum_{n=0}^{\infty} \frac{(-1)^{n+1} t^{2n+1}}{(2n+1)!} B^{n+1} &= M \sum_{n=0}^{\infty} \frac{(-1)^{n+1} t^{2n+1}}{(2n+1)!} \text{Diag}(-\nu_1^2, \nu_2^2, \dots, \nu_m^2)^{n+1} M^{-1} \\ &= M \begin{pmatrix} -\nu_1 \sinh \nu_1 t & & & 0 \\ & -\nu_2 \sin \nu_2 t & & \\ & & \ddots & \\ 0 & & & -\nu_m \sin \nu_m t \end{pmatrix} M^{-1}, \end{aligned}$$

also

$$e^{tA} = \begin{pmatrix} M \begin{pmatrix} \cosh \nu_1 t & & & 0 \\ & \cos \nu_2 t & & \\ & & \ddots & \\ 0 & & & \cos \nu_m t \end{pmatrix} M^{-1} & M \begin{pmatrix} \frac{1}{\nu_1} \sinh \nu_1 t & & & 0 \\ & \frac{1}{\nu_2} \sin \nu_2 t & & \\ & & \ddots & \\ 0 & & & \frac{1}{\nu_m} \sin \nu_m t \end{pmatrix} M^{-1} \\ M \begin{pmatrix} -\nu_1 \sinh \nu_1 t & & & 0 \\ & -\nu_2 \sin \nu_2 t & & \\ & & \ddots & \\ 0 & & & -\nu_m \sin \nu_m t \end{pmatrix} M^{-1} & M \begin{pmatrix} \frac{1}{\nu_1} \sinh \nu_1 t & & & 0 \\ & \frac{1}{\nu_2} \sin \nu_2 t & & \\ & & \ddots & \\ 0 & & & \frac{1}{\nu_m} \sin \nu_m t \end{pmatrix} M^{-1} \end{pmatrix}$$

Damit erhalten wir in diesem Fall

$$\begin{aligned} x(t) &= \left\{ \langle b_1, x_0 \rangle \cosh \nu_1 t + \langle b_1, p_0 \rangle \frac{1}{\nu_1} \sinh \nu_1 t \right\} b_1 \\ &\quad + \sum_{j=2}^m \left\{ \langle b_j, x_0 \rangle \cos \nu_j t + \langle b_j, p_0 \rangle \frac{1}{\nu_j} \sin \nu_j t \right\} b_j, \\ y(t) &= \left\{ \langle b_1, x_0 \rangle \nu_1 \sinh \nu_1 t + \langle b_1, p_0 \rangle \cosh \nu_1 t \right\} b_1 \\ &\quad + \sum_{j=2}^m \left\{ -\langle b_j, x_0 \rangle \nu_j \sin \nu_j t + \langle b_j, p_0 \rangle \cos \nu_j t \right\} b_j. \end{aligned}$$

Diese Lösung heißt *instabil*, weil sie i. allg. nicht beschränkt ist; sie ist genau dann beschränkt, wenn $\langle b_1, x_0 \rangle = \langle b_1, p_0 \rangle = 0$ gilt.

Ist z. B. $\lambda_1 = 0$ und $\lambda_j = \nu_j^2 > 0$ für $j = 2, \dots, m$, so erhält man offenbar

$$\begin{aligned} x(t) &= \left\{ \langle b_1, x_0 \rangle + \langle b_1, p_0 \rangle t \right\} b_1, \\ &\quad + \sum_{j=2}^m \left\{ \langle b_j, x_0 \rangle \cos \nu_j t + \langle b_j, p_0 \rangle \frac{1}{\nu_j} \sin \nu_j t \right\} b_j, \\ y(t) &= \langle b_1, p_0 \rangle b_1 + \sum_{j=2}^m \left\{ -\langle b_j, x_0 \rangle \nu_j \sin \nu_j t + \langle b_j, p_0 \rangle \cos \nu_j t \right\} b_j. \end{aligned}$$

Auch diese Lösung ist unbeschränkt, wenn $\langle b_1, p_0 \rangle \neq 0$ gilt.

Mit diesen Informationen ist es nun leicht, alle möglichen Situationen zu studieren.

5.4 Nicht diagonalisierbares A

Ist A nicht diagonalisierbar, weil zu wenige ($< m$) Eigenvektoren existieren, so liegt es nahe, $\exp(xA)$ mit Hilfe der Jordan'schen Normalform zu berechnen (vgl. Beispiel 5.1). Jedes A kann mit Hilfe einer invertierbaren Matrix V auf die Jordan'sche Normalform transformiert werden:

$$A = V \begin{pmatrix} A_1 & & 0 \\ & \ddots & \\ 0 & & A_n \end{pmatrix} V^{-1}, \quad V^{-1}AV = \begin{pmatrix} A_1 & & 0 \\ & \ddots & \\ 0 & & A_n \end{pmatrix},$$

mit $\ell_j \times \ell_j$ -Matrizen.

$$A_j = \begin{pmatrix} \lambda_j & 1 & & 0 \\ 0 & \lambda_j & & \\ & & \ddots & \\ & & & \lambda_j & 1 \\ 0 & & & 0 & \lambda_j \end{pmatrix} \quad \text{Jordan-Kasten;}$$

dabei sind die λ_j die Eigenwerte von A . Wie im Beispiel 5.1 kann man berechnen (vgl. auch Übungsaufgabe 1)

$$\exp(xA_j) = \begin{pmatrix} e^{x\lambda_j} & xe^{x\lambda_j} & \frac{x^2}{2!}e^{x\lambda_j} & \frac{x^3}{3!}e^{x\lambda_j} & \dots \\ & e^{x\lambda_j} & xe^{x\lambda_j} & \frac{x^2}{2!}e^{x\lambda_j} & \dots \\ & & e^{x\lambda_j} & xe^{x\lambda_j} & \dots \\ 0 & & & e^{x\lambda_j} & \dots \end{pmatrix}.$$

Die Matrix $\exp(xV^{-1}AV) = V^{-1}\exp(xA)V$ setzt sich offenbar aus diesen Kästen zusammen. Die Spalten dieser (großen) Matrix bilden bekanntlich ein Fundamentalsystem des Systems $z' = (V^{-1}AV)z$. Ein Fundamentalsystem von $y' = Ay$ erhält man hieraus durch Transformation mit V :

$$\text{für } y = Vz \text{ gilt } y' = (Vz)' = Vz' = VV^{-1}AVz = Ay.$$

Unter Berücksichtigung der Rechenregeln für Matrizen erkennt man daran, dass ein Fundamentalsystem existiert, das wie folgt geschrieben werden kann:

$$\begin{pmatrix} p_{j,\ell,1}(x) \\ p_{j,\ell,2}(x) \\ \vdots \\ p_{j,\ell,m}(x) \end{pmatrix} e^{x\lambda_j} \quad \text{mit} \quad j \equiv 1, \dots, \ell_j = \dim A_j,$$

wobei die $p_{j,\ell,k}$ Polynome vom Grad $\leq \ell - 1$ sind. Dies reicht in der Regel aus, um das Differenzialgleichungssystem $y' = Ay$ mit Hilfe geeigneter „Ansätze“ zu lösen:

Beispiel 5.5 $y' = \begin{pmatrix} 1 & -1 \\ 4 & -3 \end{pmatrix} y$. Aus

$$\det(A - \lambda E) = \det \begin{pmatrix} 1 - \lambda & -1 \\ 4 & -3 - \lambda \end{pmatrix} = \lambda^2 + 2\lambda + 1 = (\lambda + 1)^2 = 0$$

folgt, dass $\lambda = -1$ Eigenwert mit (algebraischer) Vielfachheit 2 ist. Es gibt aber nur einen Eigenvektor

$$(A + 1E) \begin{pmatrix} a \\ b \end{pmatrix} = 0, \quad \begin{matrix} 2a = b \\ 4a = 0 \end{matrix}, \quad 2a = b, \quad \text{z. B.} \quad \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}.$$

Es ist also hier $n = 1$, $\dim A_1 = 2$. Eine Lösung ist gegeben durch

$$\begin{pmatrix} 1 \\ 2 \end{pmatrix} e^{-x}.$$

Eine zweite Lösung hat die Form

$$\begin{pmatrix} a \\ b \end{pmatrix} e^{-x} + \begin{pmatrix} c \\ d \end{pmatrix} x e^{-x}.$$

Einsetzen in die Differenzialgleichung liefert (nach Multiplikation mit e^x)

$$\begin{pmatrix} -a + c \\ -b + d \end{pmatrix} - \begin{pmatrix} c \\ d \end{pmatrix} x = \begin{pmatrix} a - b \\ 4a - 3b \end{pmatrix} + \begin{pmatrix} c - d \\ 4c - 3d \end{pmatrix} x$$

und hieraus durch Vergleich der Koeffizienten bei x^1 und x^0 :

$$\begin{aligned} -\begin{pmatrix} c \\ d \end{pmatrix} &= \begin{pmatrix} c - d \\ 4c - 3d \end{pmatrix} \implies (\text{Rechnung wie oben}) \quad \begin{pmatrix} c \\ d \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \\ \begin{pmatrix} -a + c \\ -b + d \end{pmatrix} &= \begin{pmatrix} a - b \\ 4a - 3b \end{pmatrix} \implies \mathbf{2a = 2b = 1} \implies \text{z. B. } \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 0 \\ -1 \end{pmatrix}. \end{aligned}$$

Wir haben damit die weitere Lösung

$$\begin{pmatrix} x \\ -1 + 2x \end{pmatrix} e^{-x},$$

und somit das Fundamentalsystem

$$\begin{pmatrix} 1 \\ 2 \end{pmatrix} e^{-x}, \quad \begin{pmatrix} x \\ -1 + 2x \end{pmatrix} e^{-x}.$$

Im Sinne von Satz 4.6 ist also

$$Z(x) = \begin{pmatrix} e^{-x} & x e^{-x} \\ 2e^{-x} & (-1 + 2x)e^{-x} \end{pmatrix},$$

womit man die Lösungsmatrix (z. B. für $x_0 = 0$) erhält:

$$\begin{aligned} Y(x) &= Z(x)Z(0)^{-1} = e^{-x} \begin{pmatrix} 1 & x \\ 2 & -1 + 2x \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 2 & -1 \end{pmatrix}^{-1} \\ &= e^{-x} \begin{pmatrix} 1 & x \\ 2 & -1 + 2x \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 2 & -1 \end{pmatrix} = e^{-x} \begin{pmatrix} 1 + 2x & -x \\ 4x & 1 - 2x \end{pmatrix}. \end{aligned}$$

Dieses $Y(\cdot)$ könnte man natürlich auch leicht finden, indem man für die beiden Spaltenvektoren die Linearkombinationen der obigen Lösungen wählt, die für $x = 0$ die Anfangsbedingungen $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ bzw. $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ erfüllen.

Hieraus kann man auch noch leicht berechnen

$$Y(x)^{-1} = e^x \begin{pmatrix} 1 - 2x & x \\ -4x & 1 + 2x \end{pmatrix},$$

so dass dann jedes inhomogene Anfangswertproblem im Prinzip explizit lösbar wird. \square

5.5 Übungsaufgaben

5.1 a) Man löse die Anfangswertprobleme

$$y' = \begin{pmatrix} \lambda & 1 & & 0 \\ & \lambda & 1 & \\ & & \ddots & \ddots \\ 0 & & & \lambda & 1 \\ & & & & \lambda \end{pmatrix} y, \quad y(0) = e_j \quad (j = 1, \dots, m),$$

indem man das System von unten her auflöst.

b) Man gebe die Lösungsmatrix dieses Systems für $x_0 = 0$ an,

5.2 Ein Punkt A , der durch ein „Gummiband“ mit dem Ursprung verbunden ist, bewege sich geradlinig mit konstanter Geschwindigkeit u vom Ursprung weg. Auf dem Gummiband bewegt sich ein zweiter Punkt B mit konstanter Geschwindigkeit v (relativ zum Gummiband). Für welche v holt der Punkt B den Punkt A in endlicher Zeit ein?

5.3 Man bestimme ein Fundamentalsystem des Differenzialgleichungssystems

$$y' = \begin{pmatrix} 5 & 3 & \sqrt{2} \\ 3 & 5 & -\sqrt{2} \\ 0 & 0 & 2 \end{pmatrix} y.$$

5.4 Mit Hilfe der Theorie der Differenzialgleichungen zeige man für jede $m \times m$ -Matrix A

$$\exp(sA) \exp(tA) = \exp((s+t)A) \quad \text{für alle } s, t \in \mathbb{R}.$$

5.5 Man bestimme zwei stetig differenzierbare Funktionen $C, S : \mathbb{R} \rightarrow \mathbb{C}$ mit

$$\begin{aligned} C'(x) &= -S(x), & S'(x) &= C(x), \\ C(0) &= 1, & S(0) &= 0 \end{aligned}$$

und zeige, dass sie durch diese Forderungen eindeutig bestimmt sind.

6 Lineare Differenzialgleichungen der Ordnung n

Eine lineare Differenzialgleichung der Ordnung n in expliziter Form hat die Gestalt

$$\begin{aligned} y^{(n)} &= a_0(x)y + a_1(x)y' + \dots + a_{n-1}(x)y^{(n-1)} + b(x) \\ &= \sum_{j=0}^{n-1} a_j(x)y^{(j)} + b(x), \end{aligned}$$

wobei wir stets voraussetzen werden, dass die Koeffizientenfunktionen $a_j : J \rightarrow \mathbb{C}$ und die Funktion $b : J \rightarrow \mathbb{C}$ stetig auf einem Intervall $J \subset \mathbb{R}$ sind. Im ersten Unterabschnitt werden wir die allgemeine Lösungstheorie behandeln, während wir im zweiten Unterabschnitt den Spezialfall konstanter Koeffizienten a_j untersuchen; in letzterem Fall läßt sich leicht ein explizites Lösungsverfahren angeben.

6.1 Allgemeine Theorie

Mit der neuen vektorwertigen Funktion (vgl. Abschnitt 1.1)

$$\underline{y}(x) = \begin{pmatrix} y_0(x) \\ y_1(x) \\ \vdots \\ y_{n-1}(x) \end{pmatrix} := \begin{pmatrix} y(x) \\ y'(x) \\ \vdots \\ y^{(n-1)}(x) \end{pmatrix}, \quad x \in J$$

geht die obige Differenzialgleichung über in das System

$$\begin{aligned} y'_0 &= y_1 \\ y'_1 &= y_2 \\ &\vdots \\ y'_{n-2} &= y_{n-1} \\ y'_{n-1} &= a_0(x)y_0 + \dots + a_{n-1}(x)y_{n-1} + b(x), \end{aligned}$$

oder kürzer, in Matrix-Schreibweise,

$$\underline{y}' = A(x)\underline{y} + \underline{b}(x)$$

mit

$$A(x) = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \\ a_0(x) & a_1(x) & a_2(x) & \dots & a_{n-1}(x) \end{pmatrix}, \quad \underline{b}(x) = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ b(x) \end{pmatrix}.$$

Nach unseren früheren Resultaten (Ý 4) gibt es zu diesem System und jedem $x_0 \in J$ eine eindeutig bestimmte *Lösungsmatrix* $Y(\cdot)$, die Lösung des Matrixanfangswertproblems

$$\text{AWP} \quad Y' = A(x)Y, \quad Y(x_0) = E.$$

Mit dieser Lösungsmatrix $Y(\cdot)$ ist dann für jedes Anfangswertproblem

$$\text{AWP} \quad \underline{y}' = A(x)\underline{y} + \underline{b}(x), \quad \underline{y}(x_0) = \underline{y}_0$$

die eindeutig bestimmte Lösung gegeben durch

$$\underline{y}(x) = Y(x)\underline{y}_0 + Y(x) \int_{x_0}^x Y(t)^{-1} \underline{b}(t) dt.$$

Eigentlich interessiert uns jedoch nur $y(\cdot)$, die erste Komponente von $\underline{y}(\cdot)$. Damit erhalten wir zunächst:

Satz 6.1 Seien $a_j, b : J \rightarrow \mathbb{C}$ stetig, $x_0 \in J$. Dann hat das Anfangswertproblem

$$\text{AWP} \quad \begin{aligned} y^{(n)} &= a_0(x)y + a_1(x)y' + \dots + a_{n-1}(x)y^{(n-1)} + b(x), \\ y(x_0) &= y_0, \quad y'(x_0) = y_1, \dots, \quad y^{(n-1)}(x_0) = y_{n-1} \end{aligned}$$

genau eine Lösung.

Unser Ziel ist es, aus obiger Formel für $\underline{y}(x)$ eine Formel für die erste Komponente $y(x)$ von $\underline{y}(x)$ zu gewinnen.

Wir nennen n Funktionen $z_1, \dots, z_n : J \rightarrow \mathbb{C}$ ein *Fundamentalsystem* der homogenen Differentialgleichung

$$y^{(n)} = a_0(x)y + a_1(x)y' + \dots + a_{n-1}(x)y^{(n-1)},$$

wenn sie

- (i) die Differentialgleichung lösen, und
- (ii) linear unabhängig sind (in $C(J)$, d.h. aus $\sum_{j=1}^n c_j z_j(x) \equiv 0$ folgt $c_1 = \dots = c_n = 0$).

Es gilt: $\{z_1, \dots, z_n\}$ ist genau dann ein Fundamentalsystem dieser Gleichung, wenn $\{\underline{z}_1, \dots, \underline{z}_n\}$ mit

$$\underline{z}_j(x) = \left(z_j(x), z_j'(x), \dots, z_j^{(n-1)}(x) \right)^t$$

ein Fundamentalsystem des homogenen Systems $\underline{y}' = A(x)\underline{y}$ ist. Es gilt nämlich $\sum_{j=1}^n c_j z_j(x) = 0$ für alle $x \in J$ genau dann, wenn $\sum_{j=1}^n c_j z_j^{(k)}(x) = 0$ für alle $x \in J$ und $k \in \{0, \dots, n-1\}$ gilt.

11: 21. November 2006

Da die Spaltenvektoren der Lösungsmatrix $Y(\cdot)$ ein Fundamentalsystem von $\underline{y}' = A(x)\underline{y}$ bilden, bilden also die Funktionen der ersten Zeile von $Y(\cdot)$ ein Fundamentalsystem der Differentialgleichung n -ter Ordnung. Und da jede Lösung eines homogenen Systems erster Ordnung Linearkombination der Lösungen eines Fundamentalsystems ist, folgt:

Satz 6.2 Ist $\{z_1, \dots, z_n\}$ ein Fundamentalsystem der homogenen Differenzialgleichung

$$y^{(n)} = a_0(x)y + \dots + a_{n-1}(x)y^{(n-1)},$$

so ist jede Lösung dieser Gleichung darstellbar als Linearkombination von z_1, \dots, z_n . — Die Menge der Lösungen der homogenen linearen Differenzialgleichung n -ter Ordnung bilde also einen n -dimensionalen Vektorraum.

Für $n-1$ mal stetig differenzierbare Funktionen $z_1, \dots, z_n : J \rightarrow \mathbb{C}$ ist die Wronskideterminante definiert durch

$$W(x) = W(z_1, \dots, z_n; x) := \det \begin{pmatrix} z_1(x) & z_2(x) & \dots & z_n(x) \\ z_1'(x) & z_2'(x) & \dots & z_n'(x) \\ \vdots & \vdots & \dots & \vdots \\ z_1^{(n-1)}(x) & z_2^{(n-1)}(x) & \dots & z_n^{(n-1)}(x) \end{pmatrix}.$$

Für diese Wronskideterminante gilt:

Satz 6.3 Seien z_1, \dots, z_n Lösungen der homogenen Differenzialgleichung n -ter Ordnung

$$y^{(n)} = a_0(x)y + \dots + a_{n-1}(x)y^{(n-1)} \quad (x \in J),$$

$W(\cdot)$ die Wronskideterminante von z_1, \dots, z_n . Dann gilt für alle $x_0, x \in J$

$$W(x) = W(x_0) \exp \left\{ \int_{x_0}^x a_{n-1}(s) ds \right\},$$

Die folgenden Aussagen sind äquivalent:

- (i) $\{z_1, \dots, z_n\}$ ist ein Fundamentalsystem,
- (ii) $W(x) \neq 0$ für ein $x \in J$,
- (iii) $W(x) \neq 0$ für alle $x \in J$.

Beweis. Aus der (gewöhnlichen) Produktregel ergibt sich, dass die Ableitung einer Determinante gleich der Summe der Determinanten ist, die man erhält, indem man jeweils eine Zeile der Matrix differenziert. Das liefert in diesem Fall

$$W'(x) = \det \begin{pmatrix} z_1(x) & z_2(x) & \dots & z_n(x) \\ z_1'(x) & z_2'(x) & \dots & z_n'(x) \\ \vdots & \vdots & \dots & \vdots \\ z_1^{(n-2)}(x) & z_2^{(n-2)}(x) & \dots & z_n^{(n-2)}(x) \\ z_1^{(n)}(x) & z_2^{(n)}(x) & \dots & z_n^{(n)}(x) \end{pmatrix};$$

alle anderen Summanden verschwinden, da die entsprechenden Matrizen jeweils zwei identische Zeilen enthalten. In der letzten Zeile kann man nun auf Grund der Differenzialgleichung ersetzen

$$z_j^{(n)}(x) = a_0(x)z_j(x) + a_1(x)z_j'(x) + \dots + a_{n-1}(x)z_j^{(n-1)}(x).$$

Auf Grund der Linearität (bezüglich der Zeilen) der Determinante ist also die obige Determinante gleich der Summe von n Determinanten, in denen jeweils in der letzten Zeile $z_j^{(n)}(x)$ ersetzt wird durch $a_k(x)z_j^{(k)}(x)$ für $k = 0, \dots, n-1$; nur die letzte dieser Determinanten ist nicht Null. Damit folgt

$$W'(x) = \det \begin{pmatrix} z_1(x) & \dots & z_n(x) \\ z_1'(x) & \dots & z_n'(x) \\ \vdots & & \vdots \\ z_1^{(n-2)}(x) & \dots & z_n^{(n-2)}(x) \\ a_{n-1}(x)z_1^{(n-1)}(x) & \dots & a_{n-1}(x)z_n^{(n-1)}(x) \end{pmatrix} = a_{n-1}(x)W(x).$$

Daraus folgt die behauptete Formel durch Integration. Diese Formel liefert auch die Äquivalenz von (ii) und (iii).

(i) \Rightarrow (ii): Wenn (ii) nicht gilt, gibt es ein $x_0 \in J$ und $c_1, \dots, c_n \in \mathbb{C}$ mit $(c_1, \dots, c_n) \neq (0, \dots, 0)$ und $y(x) := \sum_{j=1}^n c_j z_j^{(k)}(x_0) = 0$ für $k = 0, \dots, n-1$. Dann ist aber $y(\cdot)$ die Lösung des homogenen AWP mit Anfangspunkt x_0 und Anfangswerten $y(x_0) = y'(x_0) = \dots = y^{(n-1)}(x_0) = 0$, also $y(x) \equiv 0$. Das heißt aber, dass $\{z_1, \dots, z_n\}$ linear abhängig ist, im Widerspruch zu (i).

(ii) \Rightarrow (i): Wenn (i) nicht gilt, gibt es $c_1, \dots, c_n \in \mathbb{C}$ mit $(c_1, \dots, c_n) \neq (0, \dots, 0)$ und $\sum_{j=1}^n c_j z_j(x) \equiv 0$. Dann ist aber auch $\sum_{j=1}^n c_j z_j^{(k)}(x) \equiv 0$ für alle $k = 0, \dots, n-1$ und somit $W(x) = 0$ im Widerspruch zu (ii).

Die Äquivalenz von (i) und (ii) folgt aber auch aus der Tatsache, dass $\{z_1(\cdot), \dots, z_n(\cdot)\}$ genau dann ein Fundamentalsystem der Gleichung n -ter Ordnung ist, wenn die Spalten obiger Matrix ein Fundamentalsystem des entsprechenden Systems ist. ■

Damit haben wir alle Vorbereitungen, um auch die inhomogene Gleichung zu lösen:

Satz 6.4 Sei $\{z_1, \dots, z_n\}$ ein Fundamentalsystem der homogenen Differenzialgleichung

$$y^{(n)} = a_0(x)y + \dots + a_{n-1}(x)y^{(n-1)},$$

$W(x) = W(z_1, \dots, z_n; x)$ die Wronskideterminante,

$$K(x, t) := \frac{1}{W(t)} \det \begin{pmatrix} z_1(t) & \dots & z_n(t) \\ \vdots & & \vdots \\ z_1^{(n-2)}(t) & \dots & z_n^{(n-2)}(t) \\ z_1(x) & \dots & z_n(x) \end{pmatrix}$$

Dann ist

$$y(x) = \int_{x_0}^x K(x, t)b(t) dt$$

die (eindeutige bestimmte) Lösung der inhomogenen Gleichung

$$y^{(n)} = a_0(x)y + \dots + a_{n-1}(x)y^{(n-1)} + b(x)$$

mit den Anfangswerten

$$y(x_0) = y'(x_0) = \dots = y^{(n-1)}(x_0) = 0.$$

Die Lösung zu beliebigen Anfangswerten

$$y(x_0) = y_0, \quad \dots, \quad y^{(n-1)}(x_0) = y_{n-1}$$

erhält man durch Addition der Lösung der homogenen Gleichung (Linearkombination der z_j), die diese Anfangsbedingungen erfüllt.

Beweis. Die gesuchte Lösung $y(\cdot)$ ist die erste Komponente der Lösung $\underline{y}(\cdot)$ des zugehörigen Systems:

$$\underline{y}(x) = \left(\underline{y}(x) \right)_1 = \left(Y(x) \int_{x_0}^x Y(t)^{-1} \underline{b}(t) dt \right)_1;$$

diese ist also zu berechnen.

Sei $Z(x)$ die Matrix mit den Spalten

$$\underline{z}_j(x) = \left(z_j(x), z'_j(x), \dots, z_j^{(n-1)}(x) \right)^t.$$

Dann gilt für die Lösungsmatrix $Y(\cdot)$ nach Satz 4.6

$$Y(x) = Z(x)Z(x_0)^{-1}, \quad Y(x)^{-1} = Z(x_0)Z(x)^{-1},$$

und somit

$$\begin{aligned} \underline{y}(x) &= Z(x)Z(x_0)^{-1} \int_{x_0}^x Z(x_0)Z(t)^{-1} \underline{b}(t) dt \\ &= Z(x) \int_{x_0}^x Z(t)^{-1} \underline{b}(t) dt = \int_{x_0}^x Z(x)Z(t)^{-1} \underline{b}(t) dt. \end{aligned}$$

Für den Vektor $Z(t)^{-1} \underline{b}(t)$ gilt nach der Cramer'schen Regel

$$\begin{aligned} &\left(Z(t)^{-1} \underline{b}(t) \right)_j = \\ &= \frac{1}{W(t)} \det \begin{pmatrix} z_1(t) & \dots & z_{j-1}(t) & 0 & z_{j+1}(t) & \dots & z_n(t) \\ z'_1(t) & \dots & z'_{j-1}(t) & 0 & z'_{j+1}(t) & \dots & z'_n(t) \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ z_1^{(n-2)}(t) & \dots & z_{j-1}^{(n-2)}(t) & 0 & z_{j+1}^{(n-2)}(t) & \dots & z_n^{(n-2)}(t) \\ z_1^{(n-1)}(t) & \dots & z_{j-1}^{(n-1)}(t) & b(t) & z_{j+1}^{(n-1)}(t) & \dots & z_n^{(n-1)}(t) \end{pmatrix} \\ &\quad \text{(mit dem Determinanten-Entwicklungssatz, } j\text{-te Spalte)} \\ &= (-1)^{n+j} \frac{b(t)}{W(t)} \det \begin{pmatrix} z_1(t) & \dots & z_{j-1}(t) & z_{j+1}(t) & \dots & z_n(t) \\ z'_1(t) & \dots & z'_{j-1}(t) & z'_{j+1}(t) & \dots & z'_n(t) \\ \vdots & & \vdots & \vdots & & \vdots \\ z_1^{(n-2)}(t) & \dots & z_{j-1}^{(n-2)}(t) & z_{j+1}^{(n-2)}(t) & \dots & z_n^{(n-2)}(t) \end{pmatrix}. \end{aligned}$$

Damit folgt, speziell für die erste Komponente

$$\begin{aligned} \left(Z(x)Z(t)^{-1}\underline{b}(t) \right)_1 &= \sum_{j=1}^n z_j(x) \left(Z(t)^{-1}\underline{b}(t) \right)_j \\ &= \sum_{j=1}^n z_j(x) (-1)^{n+j} \frac{b(t)}{W(t)} \det \begin{pmatrix} z_1(t) & \dots & z_{j-1}(t) & z_{j+1}(t) & \dots & z_n(t) \\ z'_1(t) & \dots & z'_{j-1}(t) & z'_{j+1}(t) & \dots & z'_n(t) \\ \vdots & & \vdots & \vdots & & \vdots \\ z_1^{(n-2)}(t) & \dots & z_{j-1}^{(n-2)}(t) & z_{j+1}^{(n-2)}(t) & \dots & z_n^{(n-2)}(t) \end{pmatrix} \\ &\quad (\text{und mit erneuter Anwendung des Entwicklungssatzes}) \\ &= K(x, t)b(t). \end{aligned}$$

Integration von x_0 bis x liefert die Behauptung. ■

6.2 Lineare Differenzialgleichungen der Ordnung n mit konstanten Koeffizienten

Im Fall konstanter Koeffizienten ist es relativ leicht, ein Fundamentalsystem der homogenen Gleichung explizit anzugeben. Diese Ergebnisse lassen sich natürlich aus den entsprechenden Sätzen für Systeme erster Ordnung mit konstanten Koeffizienten ableiten. Es ist aber auch nicht schwer, sie auf direktem Weg zu erhalten:

Mit L bezeichnen wir im folgenden den *Differenzialausdruck*, der definiert ist durch

$$Ly = y^{(n)} - a_{n-1}y^{(n-1)} - \dots - a_1y' - a_0y$$

mit Konstanten $a_0, \dots, a_{n-1} \in \mathbb{C}$. Wie man einerseits leicht erraten kann, wie andererseits aber aus den Resultaten für Systeme klar ist, sollte der Lösungsansatz

$$y(x) = e^{\lambda x} \quad (\text{Exponentialansatz})$$

zum Ziel führen. Einsetzen dieses Ansatzes in die Differenzialgleichung liefert:

$$0 = L\{e^{\lambda x}\} = p(\lambda)e^{\lambda x}$$

mit dem *charakteristischen Polynom*

$$p(\lambda) = \lambda^n - a_{n-1}\lambda^{n-1} - \dots - a_1\lambda - a_0.$$

Wie man leicht nachrechnet, gilt auch

$$p(\lambda) = (-1)^{n+1} \det \begin{pmatrix} -\lambda & 1 & & & & \\ 0 & -\lambda & 1 & & & 0 \\ \vdots & \vdots & \vdots & \ddots & & \\ 0 & 0 & 0 & \dots & 1 & \\ 0 & 0 & 0 & \dots & -\lambda & 1 \\ a_0 & a_1 & a_2 & \dots & a_{n-1} & -\lambda \end{pmatrix}$$

Diese Form von $p(\cdot)$ erinnert sehr deutlich an das zur Differenzialgleichung n -ter Ordnung äquivalente System erster Ordnung.

Offenbar ist also $y = e^{\lambda x}$ genau dann eine Lösung der Differentialgleichung, wenn λ eine Nullstelle von $p(\cdot)$ ist. Abgesehen davon, dass wir die lineare Unabhängigkeit dieser Lösungen noch nicht bewiesen haben, haben wir damit in dem Fall, in dem $p(\cdot)$ nicht n verschiedene Nullstellen hat, noch nicht genügend viele Lösungen für ein Fundamentalsystem. Die noch fehlenden Lösungen liefert der folgende Satz, in dem wir dann auch die lineare Unabhängigkeit des gesamten Systems beweisen.

Satz 6.5 Seien $\lambda_1, \dots, \lambda_k \in \mathbb{C}$ die verschiedenen Nullstellen von $p(\cdot)$ mit Vielfachheiten ν_1, \dots, ν_k $\sum_{j=1}^k \nu_j = n$ (vgl. Fundamentalsatz der Algebra). Dann ist ein Fundamentalsystem der Differentialgleichung $Ly = 0$ gegeben durch

$$e^{\lambda_j x}, \quad x e^{\lambda_j x}, \dots, \quad x^{\nu_j - 1} e^{\lambda_j x} \quad \text{für } j = 1, \dots, k.$$

Beweis. a) Wir zeigen zunächst: Wenn λ eine ν -fache Nullstelle von $p(\cdot)$ ist, dann sind

$$e^{\lambda x}, \quad x e^{\lambda x}, \dots, \quad x^{\nu - 1} e^{\lambda x}$$

Lösungen der Differentialgleichung $Ly = 0$. Aus der Gleichung $L(e^{\lambda x}) = p(\lambda)e^{\lambda x}$ folgt durch j -malige Differenziation nach λ (Vertauschbarkeit der Differenziationsreihenfolge)

$$\begin{aligned} L(x^j e^{\lambda x}) &= L\left(\frac{\partial^j}{\partial \lambda^j} e^{\lambda x}\right) = \frac{\partial^j}{\partial \lambda^j} L(e^{\lambda x}) = \frac{\partial^j}{\partial \lambda^j} \{p(\lambda)e^{\lambda x}\} \\ &\quad \left(\text{mit } (fg)^{(j)} = \sum_{\ell=0}^j \binom{j}{\ell} f^{(\ell)} g^{(j-\ell)}\right) \\ &= \sum_{\ell=0}^j \binom{j}{\ell} p^{(\ell)}(\lambda) x^{j-\ell} e^{\lambda x}. \end{aligned}$$

Ist λ_0 eine ν -fache Nullstelle von p , also $p(\lambda) = (\lambda - \lambda_0)^\nu p_0(\lambda)$, so ist

$$p^{(\ell)}(\lambda_0) = 0 \quad \text{für } \ell = 0, \dots, \nu - 1,$$

und somit

$$L\left(x^j e^{\lambda_0 x}\right) = 0 \quad \text{für } j = 0, \dots, \nu - 1.$$

b) Es bleibt die lineare Unabhängigkeit der im Satz angegebenen Lösungen zu beweisen. Dazu nehmen wir an, dass gilt

$$\sum_{j=1}^k \sum_{\ell=1}^{\nu_j} c_{j\ell} x^{\ell-1} e^{\lambda_j x} = 0 \quad \text{für alle } x.$$

Es ist zu zeigen, dass dann alle $c_{j\ell}$ verschwinden.

Wir führen diesen Beweis durch Induktion nach k .

k = 1: Auf Grund der Annahme gilt dann

$$\sum_{\ell=1}^{\nu_1} c_{\ell} x^{\ell-1} e^{\lambda_1 x} = 0 \quad \text{für alle } x,$$

also

$$\sum_{\ell=1}^{\nu_1} c_\ell x^{\ell-1} = 0 \quad \text{für alle } x.$$

Das ist aber nur möglich, wenn alle c_ℓ verschwinden (da nur das 0-Polynom für alle x verschwindet).

k \implies k + 1: Die Induktionsannahme – Gültigkeit der Behauptung für k verschiedene Nullstellen – kann wie folgt formuliert werden: Sind $q_j(\cdot)$ Polynome vom Grad $< \nu_j$ mit $\sum_{j=1}^k q_j(x) e^{\lambda_j x} = 0$ für alle x , so gilt $q_j(x) \equiv 0$ für $j = 1, \dots, k$. Auf Grund der obigen Annahme gilt mit geeigneten Polynomen $p_j(\cdot)$ vom Grad $< \nu_j$

$$\sum_{j=1}^{k+1} p_j(x) e^{\lambda_j x} = 0 \quad \text{für alle } x.$$

Multiplikation mit $e^{-\lambda_{k+1}x}$ liefert

12: 24. November 2006

$$\sum_{j=1}^k p_j(x) e^{(\lambda_j - \lambda_{k+1})x} = -p_{k+1}(x) \quad \text{für alle } x.$$

Differenziert man diese Gleichung $\nu_{k+1} + 1$ mal, so wird die rechte Seite Null, und mit Polynomen $P_j(\cdot)$ vom gleichen Grad wie $p_j(\cdot)$ (wobei das Null-Polynom vereinbarungsgemäß den Grad -1 hat) gilt dann

$$\sum_{j=1}^k P_j(x) e^{(\lambda_j - \lambda_{k+1})x} = 0 \quad \text{für alle } x.$$

Die Induktionsannahme liefert dann $P_j(\cdot) = 0$ (d. h. $P_j = -1$) für $j = 1, \dots, k$. Da die p_j den gleichen Grad wie P_j haben, folgt daraus $p_j(\cdot) = 0$ für $j = 1, \dots, k$ und mit obiger Gleichung auch $p_{k+1}(\cdot) = 0$. \square

Mit Hilfe dieses Fundamentalsystems kann nun also das in 6.1 entwickelte Verfahren zur Lösung eines beliebigen (inhomogenen) Anfangswertproblems eingesetzt werden.

Bemerkung 6.6 Ist die Differentialgleichung reell (d. h. alle Koeffizienten a_0, \dots, a_{n-1} sind reell), so hat $p(\cdot)$ mit jeder nicht-reellen Nullstelle $\lambda = \sigma + i\tau$ auch die konjugiert komplexe Nullstelle $\bar{\lambda} = \sigma - i\tau$. Die zu diesen Nullstellen gehörigen Lösungen kann man dann auch ersetzen durch Real- und Imaginärteil der oben angegebenen Lösungen:

$$\begin{aligned} \operatorname{Re} x^j e^{(\sigma \pm i\tau)x} &= x^j e^{\sigma x} \cos \tau x, \\ \operatorname{Im} x^j e^{(\sigma + i\tau)x} &= -\operatorname{Im} x^j e^{(\sigma - i\tau)x} = x^j e^{\sigma x} \sin \tau x. \end{aligned}$$

Für die Lösung gewisser inhomogener Gleichungen ist der folgende Satz nützlich:

Satz 6.7 Ist b (genau) k -fache Nullstelle von $p(\lambda) = \sum_{j=0}^n a_j \lambda^j$, so hat die inhomogene Gleichung

$$y^{(n)} - \sum_{j=0}^{n-1} a_j y^{(j)} = ce^{bx}$$

eine Lösung der Form $y(x) = dx^k e^{bx}$; dabei ist $d = \frac{c}{p^{(k)}(\lambda)}$. – Ist speziell $p(b) \neq 0$ (d. h. λ ist 0-fache Nullstelle von p), so ist $y(x) = \frac{c}{p(\lambda)} e^{bx}$ eine Lösung.

Beweis. Einsetzen von $y(x) = dx^k e^{\lambda x}$ in die Differenzialgleichung und Verwenden der Formel für $L(x^k e^{\lambda x})$ aus dem Beweis des vorhergehenden Satzes liefert $L(x^k e^{\lambda x}) = p^{(k)}(\lambda) e^{\lambda x}$, also $dp^{(k)}(\lambda) = c$, $d = c/p^{(k)}(\lambda)$. ■

Beispiel 6.8 Die (homogene) Schwingungsgleichung

$$u'' + 2hu' + w_0^2 u = 0.$$

Dabei ist $u(t)$ die Auslenkung vom Nullpunkt, $h > 0$ ein Reibungskoeffizient und w_0^2 die Federkonstante.

Die charakteristische Gleichung ist

$$p(\lambda) = \lambda^2 + 2h\lambda + w_0^2 = 0,$$

also

$$\lambda_{1,2} = -h \pm \sqrt{h^2 - w_0^2}.$$

Drei Fälle sind zu unterscheiden:

$h^2 > w_0^2$: Also $\lambda_1 \neq \lambda_2$ reell,

$$u_{1,2}(t) = \exp \left\{ \left(-h \pm \sqrt{h^2 - w_0^2} \right) t \right\}.$$

Wegen $h > \sqrt{h^2 - w_0^2}$ gehen beide Lösungen für $t \rightarrow \infty$ gegen 0 ohne zu schwingen (*aperiodischer Fall*).

$h^2 = w_0^2$: Also $\lambda_1 = \lambda_2$,

$$u_1(t) = e^{-ht}, \quad u_2(t) = te^{-ht}.$$

Auch hier gehen noch beide Lösungen für $t \rightarrow \infty$ gegen 0 ohne zu schwingen (*aperiodischer Grenzfall*).

$h^2 < w_0^2$: Also $\lambda_{1,2} = -h \pm i\sqrt{w_0^2 - h^2}$,

$$u_1(t) = e^{-at} \cos \sqrt{w_0^2 - h^2} t, \quad u_2(t) = e^{-at} \sin \sqrt{w_0^2 - h^2} t.$$

Beide Lösungen schwingen mit Frequenz $\frac{1}{2\pi} w_0$, gehen aber für $t \rightarrow \infty$ gegen 0 (*gedämpfte Schwingung*).

Ist speziell $h = 0$, so erhält man

$$u_1(t) = \cos(w_0 t), \quad u_2(t) = \sin(w_0 t),$$

eine ungedämpfte Schwingung mit Frequenz $\frac{1}{2\pi}w_0$ (deshalb wurde die Konstante w_0 gewählt).

Unter den zugehörigen inhomogenen Gleichungen interessiert insbesondere die Gleichung

$$u'' + 2hu' + w_0^2 u = c \cos wt.$$

Man hat also eine periodische „Anregung“ mit Frequenz $\frac{1}{2\pi}w$.

Man kann diese Gleichung mit den oben behandelten Methoden lösen. Eine Lösung der inhomogenen Gleichung findet man aber auch auf folgendem Weg: Man betrachtet die Gleichung mit der rechten Seite ce^{iwt} (deren Realteil gerade die richtige rechte Seite ist). Kann man diese lösen, so ist der Realteil eine Lösung der ursprünglichen Gleichung. Die komplexe Gleichung hat nach obigem Satz 6.7 im Fall $p(iw) \neq 0$ eine Lösung

$$\tilde{u}(t) = \tilde{d}e^{iwt}$$

mit

$$\tilde{d} = \frac{c}{p(iw)} = \frac{c}{w_0^2 - w^2 + 2ihw} = c \frac{w_0^2 - w^2 - 2ihw}{(w_0^2 - w^2)^2 + 4h^2w^2},$$

also die komplexe Lösung

$$\tilde{u}(t) = c \frac{w_0^2 - w^2 - 2ihw}{(w_0^2 - w^2)^2 + 4h^2w^2} e^{iwt}.$$

Der Realteil hiervon ist

$$\begin{aligned} u(t) &= c \frac{w_0^2 - w^2}{(w_0^2 - w^2)^2 + 4h^2w^2} \cos(wt) + c \frac{2hw}{(w_0^2 - w^2)^2 + 4h^2w^2} \sin(wt) \\ &= d \cos(wt + \varphi), \end{aligned}$$

die *erzwungene Schwingung*; dies klappt, wenn $w \neq w_0$ oder $hw \neq 0$ ist.

Ist $h = 0$ und $w = w_0$, also $p(iw) = 0$, so hat die komplexe Gleichung die Lösung

$$\tilde{u}(t) = \tilde{d}te^{iwt} \quad \text{mit} \quad \tilde{d} = \frac{c}{p'(iw)}.$$

Dies liefert eine reelle Lösung

$$u(t) = dt \cos(wt + \varphi);$$

das ist der *Resonanzfall*. (Dazu kommt natürlich immer noch eine geeignete Lösung der homogenen Gleichung, so gewählt, dass die Anfangsbedingung erfüllt ist.) \square

6.3 Übungsaufgaben

- 6.1 a) Man bestimme alle Lösungen der Differentialgleichung

$$xy'' - y' = x^2.$$

Anleitung: Für die homogene Gleichung mache man den Ansatz $y(x) = x^r$.

- b) Die Wronskideterminante des Fundamentalsystems verschwindet für $x = 0$. Warum ist dies kein Widerspruch zu dem Satz aus der Vorlesung, nach dem $W(x)$ entweder für alle oder für kein x verschwindet?

- 6.2 a) Man zeige: Die inhomogene Differentialgleichung

$$u'' + \omega_0^2 u = c \sin \omega t \quad (\omega_0, \omega \in \mathbb{R}, \omega_0 \neq 0)$$

hat eine Lösung der Form

- (i) $d \sin \omega t$, falls $\omega \neq \omega_0$ ist,
- (ii) $dt \cos \omega t$, falls $\omega = \omega_0$ ist.

b) Man löse für alle $\omega \in \mathbb{R}$ das Anfangswertproblem

$$u'' + \omega_0^2 u = c \sin \omega t, \quad u(0) = 0, \quad u'(0) = 1.$$

6.3 Sei $q : [0, 1] \rightarrow \mathbb{R}$ stetig. Ist für ein $\lambda \in \mathbb{R}$ das Randwertproblem (RWP)

$$-u'' + qu = \lambda u, \quad u(0) = u(1) = 0 \tag{*}$$

nichttrivial lösbar, so heißt λ *Eigenwert* und u *Eigenfunktion* des RWP.

a) Eigenfunktionen u, v zu verschiedenen Eigenwerten λ, μ sind orthogonal, d. h.

$$\int_0^1 \overline{u(x)} v(x) dx = 0.$$

b) Für $q = 0$ bestimme man alle Eigenwerte und Eigenfunktionen.

6.4 Für das RWP (*) aus Aufgabe 3 zeige man, dass die Eigenwerte keinen endlichen Häufungspunkt haben.

Anleitung: Aus $\lambda_n \rightarrow \lambda_0$, u_n zugehörige Eigenfunktionen (z. B. mit $u_n'(0) = 1$) folgt $u_n \rightarrow u_0$ gleichmäßig in $[0, 1]$.

7 Der Existenzsatz von Peano

In Abschnitt 3 haben wir gesehen, dass für ein Differenzialgleichungssystem $y' = f(x, y)$, in dem f eine lokale Lipschitzbedingung bezüglich y erfüllt, jedes Anfangswertproblem in einer Umgebung des Anfangspunktes eindeutig lösbar ist. Hier beweisen wir nun, dass für ein solches System mit lediglich stetigem f jedes Anfangswertproblem in einer Umgebung des Anfangspunktes lösbar ist; wie wir schon an Beispielen gesehen haben (vgl. Beispiele 2.4 und 2.5), kann unter diesen Umständen die Eindeutigkeit nicht garantiert werden. Der folgende Abschnitt dient der Vorbereitung.

7.1 Der Satz von Arzela–Ascoli

Der Satz von Arzela–Ascoli, den wir hier beweisen wollen, dient uns als Hilfsmittel beim Beweis des Existenzsatzes von Peano. Er ist darüberhinaus einer der grundlegenden Sätze der Analysis. In der Funktionalanalysis ist er die Basis zahlreicher Kompaktheitskriterien in Funktionsräumen (auch die Grundlage des folgenden Beweises des Satzes von Peano ist in Wirklichkeit ein Kompaktheitsargument).

Sei $\Omega \subset \mathbb{R}^m$ (oder \mathbb{C}^m), $M \subset C(\Omega)$. Die Funktionen $f \in M$ (bzw. die Menge M) heißen **gleichgradig stetig**, wenn gilt

$$\forall \varepsilon > 0 \quad \exists \delta = \delta(\varepsilon) > 0 \text{ so, dass } \forall x, x' \in \Omega \text{ mit } |x - x'| < \delta \\ \text{und } \forall f \in M \text{ gilt } |f(x) - f(x')| < \varepsilon.$$

(Insbesondere ist dann jedes einzelne $f \in M$ **gleichmäßig** stetig.) Wir sagen M ist **beschränkt**, wenn ein $C \geq 0$ existiert mit $|f(x)| \leq C$ für alle $x \in \Omega$ und $f \in M$.

Satz 7.1 (Satz von Arzela–Ascoli) Sei $\Omega \subset \mathbb{R}^m$ (oder \mathbb{C}^m) kompakt, (f_n) eine Folge aus $C(\Omega)$ so, dass $M = \{f_n : n \in \mathbb{N}\}$ gleichgradig stetig und beschränkt ist. Dann gibt es eine Teilfolge von (f_n) , die gleichmäßig konvergiert.

Beweis. Sei $\{x_i : i \in \mathbb{N}\}$ eine abzählbare Menge in Ω , die in Ω dicht ist (d. h. zu jedem $x \in \Omega$ und jedem $\varepsilon > 0$ gibt es ein $i = i(x, \varepsilon)$ mit $|x - x_i| < \varepsilon$; eine solche Menge existiert immer; ist $\Omega \subset \overline{\Omega}_0$, so kann z. B. die Menge der „rationalen“ Punkte aus Ω gewählt werden). Der Beweis besteht aus zwei Schritten:

- (i) Es gibt eine Teilfolge (g_k) von (f_n) so, dass für jedes $i \in \mathbb{N}$ die Folge $(g_k(x_i))_{k \in \mathbb{N}}$ konvergent ist.

Zum Beweis verwenden wir das **Diagonalfolgenverfahren**: Die Folge $(f_n(x_1))$ ist in \mathbb{C} beschränkt. Es gibt also eine Teilfolge $(f_{1,n})$ von (f_n) , für die $(f_{1,n}(x_1))$ konvergiert. Die Folge $(f_{1,n}(x_2))$ ist wieder beschränkt in \mathbb{C} . Es gibt also eine Teilfolge $(f_{2,n})$ von $(f_{1,n})$, für die $(f_{2,n}(x_2))$ konvergiert, usw. So erhält man sukzessive Teilfolgen $(f_{k,n})_{n \in \mathbb{N}}$, für die $(f_{k,n}(x_k))_{n \in \mathbb{N}}$ konvergiert, wobei $(f_{k+1,n})_{n \in \mathbb{N}}$ jeweils eine Teilfolge von $(f_{k,n})_{n \in \mathbb{N}}$ ist. Natürlich ist $(f_{k,n}(x_i))_{n \in \mathbb{N}}$ konvergent für alle $i \leq k$.

Die **Diagonalfolge** $(g_k) = (f_{k,k})$ ist für jedes $\ell \in \mathbb{N}$ bis auf endlich viele Anfangsglieder eine Teilfolge von $(f_{\ell,n})_{n \in \mathbb{N}}$, d. h. $(g_k(x_i))_{k \in \mathbb{N}}$ ist für jedes i konvergent. Damit ist (i) bewiesen.

13: 28. November 2006

- (ii) Die in (i) konstruierte Teilfolge (g_k) ist gleichmäßig konvergent.

Hierzu ist zu beweisen:

$$\forall \varepsilon > 0 \quad \exists N \in \mathbb{N} \text{ so, dass } \forall k, \ell \geq N \text{ und } \forall x \in \Omega \text{ gilt } |g_k(x) - g_\ell(x)| < \varepsilon.$$

Zu vorgegebenen $\varepsilon > 0$ sei $\delta = \delta(\varepsilon/3)$ gemäß der Definition der gleichgradigen Stetigkeit gewählt. Die kompakte Menge Ω läßt sich durch endlich viele Kugeln $K_p (p = 1, \dots, N_0)$ mit Durchmesser $< \delta$ überdecken. Für jedes $p \in \{1, \dots, N_0\}$ gibt es ein i_p mit $x_{i_p} \in K_p$. Da die Folgen $(g_k(x_{i_p}))_{k \in \mathbb{N}}$ für alle $p = 1, \dots, N_0$ konvergieren, gibt es ein $N \in \mathbb{N}$ so, dass gilt

$$\forall k, \ell \geq N \text{ und } \forall p \in \{1, \dots, N_0\} \text{ gilt } |g_k(x_{i_p}) - g_\ell(x_{i_p})| < \frac{\varepsilon}{3}.$$

Ist $x \in \Omega$ beliebig, so gibt es ein $p = p(x) \in \{1, \dots, N_0\}$ mit $x \in K_p$; also gilt (auf Grund der oben getroffenen Wahl von δ) für $k, \ell \geq N$

$$\begin{aligned} |g_k(x) - g_\ell(x)| &\leq |g_k(x) - g_k(x_{i_p})| + |g_k(x_{i_p}) - g_\ell(x_{i_p})| + |g_\ell(x_{i_p}) - g_\ell(x)| \\ &\leq \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon. \end{aligned}$$

Das ist die gleichmäßige Konvergenz der Folge (g_k) . □

7.2 Der Existenzsatz von Peano

Satz 7.2 (Existenzsatz von Peano) Sei $G \subset \mathbb{R} \times \mathbb{C}^m$ offen, $f : G \rightarrow \mathbb{C}^m$ stetig, $(x_0, y_0) \in G$. Dann existiert ein $\alpha > 0$ so, dass das Anfangswertproblem

$$\text{AWP} \quad y' = f(x, y), \quad y(x_0) = y_0$$

in $J = [x_0 - \alpha, x_0 + \alpha]$ lösbar ist. Eine Abschätzung für die Größe von α ergibt sich aus dem Beweis. (Über die Eindeutigkeit kann im allgemeinen nichts gesagt werden; das zeigt z. B. das früher behandelte AWP $y' = |y|^{1/2}$ mit $y(x_0) = 0$ – Es kann gezeigt werden, dass auch in diesem Fall die Lösung bis zum Rand von G fortgesetzt werden kann; vgl. z. B. W. WALTER §I.7.)

Beweis. Wir wählen ein „Rechteck“ R in $G \subset \mathbb{R} \times \mathbb{C}^m$,

$$R := \left\{ (x, y) \in \mathbb{R} \times \mathbb{C}^m : |x - x_0| < \alpha_1, |y - y_0| \leq \alpha_2 \right\} \subset G,$$

und setzen f auf den „Streifen“

$$S := \left\{ (x, y) \in \mathbb{R} \times \mathbb{C}^m : |x - x_0| < \alpha_1 \right\}$$

stetig und beschränkt fort, z.B. durch

$$\hat{f}(x, y) := \begin{cases} f(x, y) & \text{für } (x, y) \in R, \\ f(x, y_0 + \alpha_2 e) & \text{für } (x, y) = (x, y_0 + ce) \text{ mit } (x, y_0) \in R \text{ und } c > \alpha_2, \end{cases}$$

wobei e ein beliebiger Einheitsvektor in \mathbb{C}^m ist; d.h. $\hat{f}(x, \cdot)$ ist außerhalb $|y - y_0| \leq \alpha_2$ konstant auf Strahlen die von y_0 ausgehen. Offenbar ist

\hat{f} stetig, und es gilt: $|\hat{f}(x, y)| \leq C := \max \{ |f(x, y)| : (x, y) \in R \}$ für alle $(x, y) \in S$.

Für jedes $\beta > 0$ definieren wir $y_\beta : (-\infty, x_0 + \alpha_1] \longrightarrow \mathbb{C}^m$ durch

$$y_\beta(x) := \begin{cases} y_0 & \text{für } x \leq x_0, \\ y_0 + \int_{x_0}^x \hat{f}(t, y_\beta(t - \beta)) dt & \text{für } x_0 < x \leq x_0 + \alpha_1. \end{cases}$$

Diese Funktion ist tatsächlich induktiv wohldefiniert, denn: um y_β auf $(x_0, x_0 + \beta]$ zu berechnen, werden nur die Werte von y_β für $t \leq x_0$ benötigt; diese sind vorgegeben. Um y_β auf $(x_0 + k\beta, x_0 + (k+1)\beta]$ zu berechnen, werden nur die Werte von y_β für $t \leq x_0 + k\beta$ benötigt; diese sind im vorherigen Schritt berechnet worden.

Die Funktionenmenge $\{y_\beta : \beta > 0\}$ ist beschränkt,

$$|y_\beta(x)| \leq |y_0| + \alpha_1 C,$$

und gleichgradig stetig, denn es gilt

$$|y'_\beta(x)| \leq C, \quad \text{also} \quad |y_\beta(x) - y_\beta(x')| \leq C|x - x'|,$$

d. h. in der Definition der gleichgradigen Stetigkeit kann $\delta(\varepsilon) := \frac{\varepsilon}{C}$ gewählt werden. Nach dem Satz von Arzela–Ascoli enthält also die Folge $(y_{1/n})$ (als Funktionen auf $[x_0, x_0 + \alpha_1]$) eine gleichmäßig konvergente Teilfolge (y_{β_n}) , d. h. es existiert eine stetige Funktion $y : [x_0, x_0 + \alpha_1] \longrightarrow \mathbb{C}$ mit

$$y_{\beta_n}(x) \longrightarrow y(x) \quad \text{gleichmäßig für } n \longrightarrow \infty.$$

Wegen

$$\begin{aligned} & \left| y_{\beta_n}(x - \beta_n) - y(x) \right| \\ & \leq \left| y_{\beta_n}(x - \beta_n) - y_{\beta_n}(x) \right| + \left| y_{\beta_n}(x) - y(x) \right| \\ & \leq \beta_n C + \left| y_{\beta_n}(x) - y(x) \right| \end{aligned}$$

gilt dann auch

$$y_{\beta_n}(x - \beta_n) \longrightarrow y(x) \quad \text{gleichmäßig auf } [x_0, x_0 + \alpha_1].$$

In der Definition von y_{β_n} ,

$$y_{\beta_n}(x) = \begin{cases} y_0 & \text{für } x \leq x_0, \\ y_0 + \int_{x_0}^x \hat{f}(t, y_{\beta_n}(t - \beta_n)) dt & \text{für } x_0 < x \leq x_0 + \alpha_1 \end{cases}$$

kann also der Grenzübergang $n \longrightarrow \infty$ ausgeführt werden:

$$y(x) = y_0 + \int_{x_0}^x \hat{f}(t, y(t)) dt \quad \text{für } x_0 \leq x \leq x_0 + \alpha_1,$$

d. h. y ist Lösung des

$$\text{AWP} \quad y' = \hat{f}(x, y), \quad y(x_0) = y_0 \quad \text{im Intervall} \quad [x_0, x_0 + \alpha_1].$$

Für $|x - x_0| \leq \alpha := \min \left\{ \alpha_1, \frac{\alpha_2}{C} \right\}$ gilt

$$\left| y(x) - y_0 \right| \leq \left| \int_{x_0}^x \hat{f}(t, y(t)) dt \right| \leq C|x - x_0| \leq C\alpha \leq \alpha_2,$$

d. h. es gilt

$$\hat{f}(x, y(x)) = f(x, y(x)) \quad \text{für} \quad |x - x_0| \leq \alpha.$$

Somit löst y das vorgegebene Anfangswertproblem in $[x_0, x_0 + \alpha]$. Entsprechend findet man ein Stück der Lösung links von x_0 . \square

Bemerkung 7.3 Für das früher behandelte Anfangswertproblem

$$\text{AWP} \quad y' = |y|^{1/2}, \quad y(x_0) = 0,$$

das bekanntlich nicht eindeutig lösbar ist, liefert das Verfahren des obigen Beweises offenbar

$$y_\beta(x) = 0 \quad \text{für alle} \quad x \geq x_0, \beta > 0,$$

und somit die Lösung $y(x) \equiv 0$.

Bemerkung 7.4 Man könnte auch versucht sein, zur Konstruktion einer Lösung im Fall des Satzes von Peano wieder die Picard'sche Iteration zu verwenden. Tatsächlich erhielte man (zumindest in einem kleinen Intervall) eine beschränkte und gleichgradig stetige Folge (y_n) , die also eine gleichmäßig konvergente Teilfolge (y_{n_k}) enthält. Aber in

$$y_{n_k+1}(x) = y_0 + \int_{x_0}^x f(t, y_{n_k}(t)) dt$$

kann man nicht zur Grenze übergehen.

7.3 Übungsaufgaben

- 7.1 Sei $\Omega \subset \mathbb{R}^m$, $f_n : \Omega \rightarrow \mathbb{C}$ eine Folge von Funktionen, deren Einschränkungen auf jede kompakte Teilmenge von Ω gleichgradig stetig sind. Dann existiert eine Teilfolge von (f_n) , die auf jeder kompakten Teilmenge von Ω gleichmäßig konvergiert. Anleitung: Diagonalfolgenverfahren.

8 Einführung in die Lebesguesche Integrationstheorie

Das Ziel dieses Abschnitts ist es, das Riemann–Integral bzw. das Integral für Regelfunktionen zu verallgemeinern, so dass mehr Funktionen integrierbar sind (für Riemann–integrierbare bzw. Regelfunktionen soll das Integral natürlich mit dem früheren Integral übereinstimmen). Insbesondere wird das neue Integral bessere Eigenschaften bezüglich Vertauschung von Grenzübergang und Integration haben.

Es gibt im Wesentlichen zwei Möglichkeiten vorzugehen:

- man entwickelt eine vollständige Maßtheorie (was recht technisch und mühsam ist) und kann dann relativ schnell das Integral bezüglich dieses Maßes erklären;
- man entwickelt die Maßtheorie nur ein kleines Stück weit (Prämaß), kann dann das Integral definieren und dessen Eigenschaften beweisen. Anschließend ist es einfach, das Maß zu erklären und zu untersuchen.

Das erste Verfahren wird von den Wahrscheinlichkeitstheoretikern vorgezogen. Das zweite entspricht eher den Bedürfnissen der Analysis und der Funktionalanalysis, weshalb wir so vorgehen wollen.

8.1 Prämaße und Nullmengen

Sei X eine Menge. Eine nichtleere Familie \mathcal{R} von Teilmengen von X heißt ein *Mengenring*, wenn aus $A, B \in \mathcal{R}$ folgt $A \cup B \in \mathcal{R}$ und $A \setminus B = A \cap (X \setminus B) \in \mathcal{R}$; dann gilt auch $A \cap B = A \setminus (A \setminus B) \in \mathcal{R}$ und, da stets ein $A \in \mathcal{R}$ existiert, $\emptyset = A \setminus A \in \mathcal{R}$ (dagegen gehört X selbst i. allg. nicht zu \mathcal{R}).

Zum Beispiel bildet in \mathbb{R}^m die Familie der Mengen, die als Vereinigungen von endlich vielen beschränkten m –dimensionalen (und ggf. niedrigerdimensionalen, ausgearteten) Intervallen darstellbar sind, einen Mengenring, die *Figuren* in \mathbb{R}^m . — Auf einer Teilmenge von \mathbb{R}^m , insbesondere auf einer niedrigerdimensionalen Mannigfaltigkeit S in \mathbb{R}^m bilden die Durchschnitte von S mit den Figuren in \mathbb{R}^m einen Mengenring.

Sei nun \mathcal{R} ein Mengenring in X . Eine Funktion $\mu : \mathcal{R} \rightarrow [0, \infty)$ heißt *Prämaß* auf (X, \mathcal{R}) , wenn sie σ –*additiv* ist, d. h., wenn gilt: Aus $A_n \in \mathcal{R}$ ($n \in \mathbb{N}$), $A_n \cap A_m = \emptyset$ für $n \neq m$ und $\bigcup_{n=1}^{\infty} A_n \in \mathcal{R}$ folgt $\mu(\bigcup_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} \mu(A_n)$. — Insbesondere ergibt sich $\mu(\emptyset) = 0$ und $\mu(A \cup B) = \mu(A) + \mu(B)$ für $A, B \in \mathcal{R}$ mit $A \cap B = \emptyset$.

Beispiel 8.1 In \mathbb{R}^m sei $\lambda(A)$ das m –dimensionale Volumen für Figuren A in \mathbb{R}^m : Wenn A als Vereinigung disjunkter Intervalle dargestellt ist, dann ist $\lambda(A)$ die Summe der Volumina dieser Intervalle; ausgeartete Intervalle haben dabei das Volumen 0. Dieses Prämaß nennen wir das *Lebesguesche Prämaß* auf \mathbb{R}^m . □

14: 1. Dezember 2006

Beispiel 8.2 Für eine rechtsstetige, nicht fallende Funktion $\phi : \mathbb{R} \rightarrow \mathbb{R}$ definieren wir ²

$$\mu((a, b)) = \phi(b-) - \phi(a), \quad \mu(\{a\}) = \phi(a) - \phi(a-).$$

²Wenn ϕ nicht rechtsstetig vorausgesetzt wird, kann man definieren $\mu((a, b)) = \phi(b-) - \phi(a+)$ und $\mu(\{a\}) = \phi(a+) - \phi(a-)$. Der Nachteil ist, dass in diesem Falle ϕ durch μ nicht eindeutig bestimmt ist. — Die Forderung der Rechtsstetigkeit hat zur Folge, dass ϕ durch μ eindeutig bestimmt wird.

Jede Vereinigung von endlich vielen Intervallen kann dargestellt werden als eine disjunkte Vereinigung von Intervallen der Form (a, b) und $\{a\}$. Aus diesem Grund kann $\mu(A)$ für jede Teilmenge A von \mathbb{R} definiert werden, die als Vereinigung endlich vieler Intervalle darstellbar ist. Dieses μ nennt man das durch ϕ erzeugte *Lebesgue–Stieltjessche Prämaß* ϕ . \square

In beiden Fällen haben wir die σ -Additivität von μ nicht überprüft. Aufgabe 8.1 gibt Hinweise zum Beweis der σ -Additivität für den Fall des Beispiels 8.2.

Sei μ ein Prämaß auf (X, \mathcal{R}) . Eine Teilmenge N von X heißt μ -Nullmenge, wenn für jedes $\varepsilon > 0$ eine Folge (A_n) in \mathcal{R} existiert, mit

$$N \subset \bigcup_n A_n \quad \text{und} \quad \sum_n \mu(A_n) < \varepsilon.$$

Offensichtlich gilt: Jede Teilmenge einer Nullmenge ist eine Nullmenge. In \mathbb{R}^m ist jede endliche und jede abzählbare Menge eine Nullmenge bezüglich des Lebesgueschen Prämaßes λ .³

Satz 8.3 Die Vereinigung von abzählbar vielen μ -Nullmengen N_k ist wieder eine μ -Nullmenge.

Beweis. Für jedes $\varepsilon > 0$ und jedes $k \in \mathbb{N}$ existiert eine Folge $(A_{k,n})_{n \in \mathbb{N}}$ aus \mathcal{R} so, dass

$$N_k \subset \bigcup_n A_{k,n} \quad \text{und} \quad \sum_n \mu(A_{k,n}) < 2^{-k} \varepsilon.$$

Daraus folgt

$$\bigcup_k N_k \subset \bigcup_{k,n} A_{k,n} \quad \text{und} \quad \sum_{k,n} \mu(A_{k,n}) < \varepsilon,$$

d. h. die Vereinigung der N_k ist eine μ -Nullmenge. \blacksquare

Man sagt, eine Aussage gilt μ -fast überall (μ -f.ü. in X), oder für μ -fast alle x in X (für μ -f.a. $x \in X$), wenn eine μ -Nullmenge N existiert so, dass die Aussage für alle $x \in X \setminus N$ gilt, z. B.: $f(x) = g(x)$, oder $f(x) > g(x)$, oder $f_n(x) \rightarrow f(x)$ für $n \rightarrow \infty$, oder ... f.ü.⁴

8.2 Das Integral für Elementarfunktionen

Sei \mathcal{R} ein Mengenring in X . Eine Funktion $f : X \rightarrow \mathbb{C}$ heißt eine *Elementarfunktion* (genauer, eine \mathcal{R} -Elementarfunktion), wenn endlich viele disjunkte Mengen $A_1, A_2, \dots, A_n \in \mathcal{R}$ und $c_1, c_2, \dots, c_n \in \mathbb{C}$ existieren so, dass

$$f(x) = \sum_{j=1}^n c_j \chi_{A_j}(x) = \begin{cases} c_j & \text{für } x \in A_j (j = 1, \dots, n), \\ 0 & \text{für } x \in X \setminus \bigcup_{j=1}^n A_j. \end{cases}$$

³Im Sinne des JORDANSchen Inhalts hat eine Menge A in \mathbb{R}^m den Inhalt 0, wenn es endlich viele Intervalle gibt, deren Vereinigung die Menge A enthält und deren Inhalt $< \varepsilon$ ist.

⁴In der Mathematik wird auch die Ausdrucksweise „fast alle“ (ohne Zusatz) benutzt. Sei besagt „für alle bis auf endlich viele“.

Diese Menge von Funktionen bildet einen komplexen Vektorraum $E(X, \mathcal{R})$. In dem vorhergehenden Beispiel 8.1 ($X = \mathbb{R}^m$, \mathcal{R} die Familie von Figuren in \mathbb{R}^m) heißen diese Funktionen auch *Treppenfunktionen*.

Sei μ ein Prämaß auf (X, \mathcal{R}) . Für \mathcal{R} -Elementarfunktionen $f = \sum_{j=1}^n c_j \chi_{A_j}$ ist das μ -Integral definiert durch

$$\int f(x) \, d\mu(x) = \int f \, d\mu := \sum_{j=1}^n c_j \mu(A_j).$$

Dieser Ausdruck ist wohldefiniert obwohl die Mengen A_j durch f i. allg. nicht eindeutig bestimmt sind. Die Abbildung $E(X, \mathcal{R}) \rightarrow \mathbb{C}$, $f \mapsto \int f \, d\mu$ ist offensichtlich

- *linear*: $\int (af + bg) \, d\mu = a \int f \, d\mu + b \int g \, d\mu$, und
- *positiv*: aus $f \geq 0$ folgt $\int f \, d\mu \geq 0$.

Es gilt außerdem

$$\begin{aligned} \left| \int f \, d\mu \right| &= \left| \sum_k c_j \mu(A_j) \right| \leq \sum_k |c_j| \mu(A_j) = \int |f| \, d\mu, \\ \int |f + g| \, d\mu &\leq \int (|f| + |g|) \, d\mu = \int |f| \, d\mu + \int |g| \, d\mu. \end{aligned}$$

Also definiert

$$\|f\|_1 := \int |f(x)| \, d\mu(x)$$

eine *Halbnorm* auf $E(X, \mathcal{R})$.

Wir sagen, dass eine Folge von Funktionen $f_n : X \rightarrow \mathbb{C}$ μ -fast gleichmäßig gegen eine Funktion $f : X \rightarrow \mathbb{C}$ konvergiert, wenn für jedes $\varepsilon > 0$ eine Folge $(A_k^\varepsilon)_{k \in \mathbb{N}}$ in \mathcal{R} existiert so, dass gilt

$$\sum_k \mu(A_k^\varepsilon) < \varepsilon \quad \text{und} \quad f_n(x) \rightarrow f(x) \quad \text{gleichmäßig in} \quad X \setminus \bigcup_k A_k^\varepsilon.$$

Aus μ -fast gleichmäßiger Konvergenz folgt offenbar Konvergenz μ -fast überall. Zum Beispiel konvergiert die Folge $f_n(x) = x^n$ in $[0, 1]$ oder $[0, 1)$ λ -fast gleichmäßig (wenn λ das Lebesguesche Prämaß ist); dieses Beispiel zeigt aber auch, dass es im allgemeinen keine μ -Nullmenge gibt, außerhalb der die Folge gleichmäßig konvergiert.

Satz 8.4 Sei (f_n) eine $\|\cdot\|_1$ -Cauchyfolge in $E(X, \mathcal{R})$. Dann gilt:

- a) Es existiert eine Funktion $f : X \rightarrow \mathbb{C}$ und eine Teilfolge (f_{n_k}) von (f_n) so, dass

$$f_{n_k}(x) \rightarrow f(x) \quad \mu\text{-fast gleichmäßig.}$$

- b) Wenn (g_ℓ) und (h_ℓ) Teilfolgen von (f_n) sind, die μ -fast überall konvergieren, dann gilt

$$g_\ell(x) - h_\ell(x) \rightarrow 0 \quad \mu\text{-fast überall.}$$

Beweis. a) Die Teilfolge (f_{n_k}) sei so gewählt, dass gilt

$$\|f_{n_k} - f_{n_{k+1}}\|_1 \leq 3^{-k-1} \quad \text{für alle } k \in \mathbb{N}.$$

Mit

$$M_k := \left\{ x \in X : \left| f_{n_k}(x) - f_{n_{k+1}}(x) \right| \geq 2^{-k-1} \right\} \quad \text{und} \quad N_\ell := \bigcup_{k \geq \ell} M_k$$

konvergiert die Reihe $\sum_k (f_{n_k}(x) - f_{n_{k+1}}(x))$, und somit die Folge $(f_{n_k}(x))$, gleichmäßig in $X \setminus N_\ell$ für jedes ℓ . Wegen

$$\mu(M_k)2^{-k-1} \leq \int_{M_k} |f_{n_k} - f_{n_{k+1}}| d\mu \leq \|f_{n_k} - f_{n_{k+1}}\|_1 \leq 3^{-k-1}$$

gilt außerdem

$$\begin{aligned} \mu(M_k) &\leq 2^{k+1}3^{-k-1} = \left(\frac{2}{3}\right)^{k+1}, \\ \sum_{k \geq \ell} \mu(M_k) &\leq \sum_{k \geq \ell} \left(\frac{2}{3}\right)^{k+1} \rightarrow 0 \quad \text{für } \ell \rightarrow \infty, \end{aligned}$$

und daraus folgt, dass (f_{n_k}) μ -fast gleichmäßig konvergiert.

b) Es gelte $g_\ell(x) \rightarrow g(x)$ und $h_\ell(x) \rightarrow h(x)$ μ -f. ü. Dann ist zu zeigen, dass gilt $g(x) = h(x)$ μ -f. ü. Offensichtlich ist $(p_n) = (g_1, h_1, g_2, h_2, \dots)$ eine $\|\cdot\|_1$ -Cauchyfolge in $E(X, \mathcal{R})$. Eine Teilfolge (p_{n_k}) gemäß Teil a des Satzes kann so ausgewählt werden, dass die Funktionen mit ungeraden Indizes aus (g_ℓ) entnommen sind und die mit geraden Indizes aus (h_ℓ)

$$(p_{m_k}) = (g_{m_1}, h_{k_1}, g_{m_2}, h_{k_2}, \dots).$$

Es gibt dann eine Funktion $f : X \rightarrow \mathbb{C}$ und eine μ -Nullmenge $N \subset X$ so, dass $p_{n_k}(x) \rightarrow f(x)$ für $x \in X \setminus N$ gilt. Selbstverständlich gilt das auch für die Teilfolgen mit ungeraden/geraden Indizes; andererseits existieren μ -Nullmengen N_1 und N_2 so dass diese Folgen in $X \setminus N_1$ gegen g und in $X \setminus N_2$ gegen h konvergieren. Dies impliziert, dass $g(x) = f(x) = h(x)$ für alle x im Komplement der μ -Nullmenge $N \cup N_1 \cup N_2$ gilt. ■

15: 5. Dezember 2006

Satz 8.5 a) Für $A, A_j \in \mathcal{R}$ ($j \in \mathbb{N}$) mit $A \subset \cup_j A_j$ gilt $\mu(A) \leq \sum_j \mu(A_j)$.

b) Für eine Folge (f_n) aus $E(X, \mathcal{R})$ mit $f_n(x) \searrow 0$ μ -f. ü. gilt $\|f_n\|_1 \rightarrow 0$.

Beweis. a) Mit $B_1 := A_1$, $B_{n+1} := A_{n+1} \setminus (\cup_{j=1}^n B_j)$ für $n \in \mathbb{N}$ gilt $A = \cup_j (A \cap B_j)$, wobei die Mengen $A \cap B_j$ disjunkt sind. Die σ -Additivität von μ impliziert

$$\mu(A) = \sum_j \mu(A \cap B_j) \leq \sum_j \mu(B_j) \leq \sum_j \mu(A_j).$$

b) Aus $0 \leq \int f_{n+1} d\mu \leq \int f_n d\mu \leq \int f_1 d\mu$ folgt, dass (f_n) eine $\|\cdot\|_1$ -Cauchyfolge ist. Aus Satz 8.4 und der Monotonie folgt deshalb, dass (f_n) μ -fast gleichmäßig gegen 0 konvergiert. Für jedes

$\varepsilon > 0$ existiert eine Folge (A_j) in \mathcal{R} so, dass $\sum_j \mu(A_j) < \varepsilon$ und $f_n \rightarrow 0$ gleichmäßig in $X \setminus (\cup_j A_j)$. Wählt man n_0 so, dass

$$f_n(x) < \varepsilon \text{ für } n \geq n_0 \text{ und } x \in X \setminus \left(\bigcup_j A_j \right)$$

gilt, so folgt zusammen mit Teil a des Satzes

$$\begin{aligned} \|f_n\|_1 &= \int f_n \, d\mu \\ &\leq \mu\left(\left\{x \in X : f_n(x) > \varepsilon\right\}\right) \max f_1 + \varepsilon \mu\left(\left\{x \in X : f_1(x) \neq 0\right\}\right) \\ &\leq \sum_j \mu(A_j) \max f_1 + \varepsilon \mu\left(\left\{x \in X : f_1(x) \neq 0\right\}\right) \\ &\leq \varepsilon \left\{ \max f_1 + \mu\left(\left\{x \in X : f_1(x) \neq 0\right\}\right) \right\} \end{aligned}$$

für $n \geq n_0$, und deshalb $\|f_n\|_1 \rightarrow 0$. ■

Satz 8.6 Sei (f_n) eine $\|\cdot\|_1$ -Cauchyfolge in $E(X, \mathcal{R})$, die μ -f.ü. konvergiert. Dann sind die folgenden Aussagen äquivalent:

$$(\alpha) \quad f_n(x) \rightarrow 0 \text{ } \mu\text{-f. ü.}, \quad (\beta) \quad \|f_n\|_1 \rightarrow 0.$$

Beweis. $(\alpha) \Rightarrow (\beta)$: Annahme: $\|f_n\|_1 \not\rightarrow 0$. Ohne Einschränkung können wir annehmen, dass $f_n \geq 0$ gilt (sonst ist $|f_n|$ an Stelle von f_n zu betrachten) und $\|f_n\|_1 \rightarrow c > 0$. Dann existiert eine Teilfolge (n_j) von \mathbb{N} so, dass

$$\|f_{n_1}\|_1 \geq \frac{2}{3}c, \quad \|f_{n_j} - f_{n_{j+1}}\|_1 < \frac{1}{3}c \left(\frac{1}{2}\right)^{j+1}.$$

Die Folge (\tilde{f}_j) mit

$$\tilde{f}_j(x) := \min \left\{ f_{n_1}(x), \dots, f_{n_j}(x) \right\}$$

ist nicht-wachsend mit

$$\begin{aligned} \tilde{f}_j &\geq f_{n_1} - |f_{n_1} - f_{n_2}| - \dots - |f_{n_{j-1}} - f_{n_j}|, \\ \|\tilde{f}_j\|_1 &= \int \tilde{f}_j \, d\mu \geq \|f_{n_1}\|_1 - \|f_{n_1} - f_{n_2}\|_1 - \dots - \|f_{n_{j-1}} - f_{n_j}\|_1 \\ &> \frac{2}{3}c - \frac{1}{3}c = \frac{1}{3}c. \end{aligned}$$

Also gilt $\|\tilde{f}_j\|_1 \not\rightarrow 0$, und mit Satz 8.5 b folgt, dass $\tilde{f}_j(x)$ nicht μ -f. ü. gegen 0 konvergiert. Zusammen mit $f_{n_j} > \tilde{f}_j$ ergibt sich ein Widerspruch zu (α) .

$(\beta) \Rightarrow (\alpha)$: Offensichtlich ist die Folge $(f_1, 0, f_2, 0, \dots)$ auch eine $\|\cdot\|_1$ -Cauchyfolge. Die Teilfolgen (f_n) und $(0, 0, \dots)$ sind μ -f. ü. konvergent gegen $f(x) := \lim_{n \rightarrow \infty} f_n(x)$ bzw. 0. Mit Satz 8.5 b folgt daraus $f(x) = 0$ μ -f. ü. ■

Korollar 8.7 a) (f_n) und (g_n) seien $\|\cdot\|_1$ -Cauchyfolgen in $E(X, \mathcal{R})$ mit $f_n(x) - g_n(x) \rightarrow 0$ μ -f. ü. Dann gilt

$$\left| \int (f_n - g_n) d\mu \right| \leq \|f_n - g_n\|_1 \rightarrow 0, \quad \|f_n\|_1 - \|g_n\|_1 \rightarrow 0.$$

b) Ist (f_n) eine $\|\cdot\|_1$ -Cauchyfolge in $E(X, \mathcal{R})$ mit $\lim_{n \rightarrow \infty} f_n(x) \geq 0$ μ -f. ü., so folgt

$$\lim_{n \rightarrow \infty} \int f_n d\mu \geq 0.$$

Beweis. a) Aus der Voraussetzung folgt, dass auch $(f_n - g_n)$ eine $\|\cdot\|_1$ -Cauchy Folge ist, und deshalb gilt nach Satz 8.6 $\|f_n - g_n\|_1 \rightarrow 0$. Der zweite Teil folgt aus der Dreiecksungleichung für die Halbnorm $\|\cdot\|_1$.

b) Offensichtlich ist auch $g_n(x) := \max\{0, f_n(x)\}$ eine $\|\cdot\|_1$ -Cauchyfolge und erfüllt $f_n(x) - g_n(x) \rightarrow 0$ μ -f. ü. Wegen $g_n \geq 0$ folgt

$$\lim_{n \rightarrow \infty} \int f_n d\mu = \lim_{n \rightarrow \infty} \int g_n d\mu \geq 0$$

aus Teil a. ■

8.3 Integrierbare Funktionen

Eine Funktion $f : X \rightarrow \mathbb{C}$ heißt μ -integrierbar, wenn eine $\|\cdot\|_1$ -Cauchyfolge (f_n) in $E(X, \mathcal{R})$ mit $f_n(x) \rightarrow f(x)$ μ -f. ü. existiert. In diesem Falle ist das μ -Integral von f definiert durch

$$\int f(x) d\mu(x) = \int f d\mu := \lim_{n \rightarrow \infty} \int f_n d\mu.$$

Die *Eindeutigkeit* der Definition und die *Positivität* des μ -Integrals ergeben sich aus Korollar 8.7. Die *Linearität* ergibt sich beim Grenzübergang aus der Linearität des Integrals für Elementarfunktionen.

Wenn f μ -integrierbar ist, dann ist auch $|f|$ μ -integrierbar (weil mit (f_n) auch $(|f_n|)$ eine $\|\cdot\|_1$ -Cauchy Folge ist mit $|f_n(x)| \rightarrow |f(x)|$ μ -f. ü.) und zusammen mit der Positivität ergibt sich die *Dreiecksungleichung*

$$\int |f + g| d\mu \leq \int (|f| + |g|) d\mu = \int |f| d\mu + \int |g| d\mu.$$

Bemerkung 8.8 a) Die *Dirichletfunktion* $f(x) = 1$ für rationale x und $f(x) = 0$ für irrationale x ist Lebesgue-integrierbar mit Integral 0, da sie fast überall mit der Nullfunktion übereinstimmt.

b) Jede Riemann-integrierbare Funktion ist Lebesgue-integrierbar: Seien (ϕ_n) und (ψ_n) Folgen von Treppenfunktionen mit $\phi_n \leq f \leq \psi_n$ und $\int (\psi_n - \phi_n) dx \rightarrow 0$; o. E. können (ϕ_n) nichtfallend und (ψ_n) nichtwachsend gewählt werden. Dann sind (ϕ_n) und (ψ_n) $\|\cdot\|_1$ -Cauchyfolgen mit $\phi_n(x) \rightarrow f(x)$ und $\psi_n(x) \rightarrow f(x)$ λ -f. ü.

Die Menge aller μ -integrierbaren Funktionen auf X bezeichnen wir mit $\mathcal{L}^1(X, \mu)$. Offensichtlich ist $\mathcal{L}^1(X, \mu)$ ein komplexer (bzw. reeller, falls nur reelle Funktionen betrachtet werden) Vektorraum und $\|f\|_1 := \int |f| d\mu$ ist eine *Halbnorm* auf $\mathcal{L}^1(X, \mu)$ (man beachte, dass dies im allgemeinen keine Norm ist, da für alle f mit $f(x) = 0$ μ -f. ü. gilt $\|f\|_1 = \int |f| d\mu = 0$).

16: 8. Dezember 2006

Satz 8.9 Sei (f_n) eine $\|\cdot\|_1$ -Cauchyfolge in $\mathcal{L}^1(X, \mu)$. Dann gibt es ein $f \in \mathcal{L}^1(X, \mu)$ mit

$$\int f d\mu = \lim_{n \rightarrow \infty} \int f_n d\mu \quad \text{und} \quad \int |f - f_n| d\mu \rightarrow 0 \quad \text{für} \quad n \rightarrow \infty;$$

es existiert eine Teilfolge (f_{n_k}) mit $f_{n_k}(x) \rightarrow f(x)$ μ -fast gleichmäßig, insbesondere also μ -fast überall.

Beweis. Aus der Definition der μ -Integrierbarkeit (von f_n) und Satz 8.4 folgt: Für jedes $n \in \mathbb{N}$ existiert ein $h_n \in E(X, \mathcal{R})$ und eine Folge $(A_\ell^n)_\ell$ in \mathcal{R} so, dass gilt

$$\|f_n - h_n\|_1 \leq 2^{-n}, \quad \sum_{\ell} \mu(A_\ell^n) < 2^{-n}$$

und

$$|f_n(x) - h_n(x)| \leq 2^{-n} \quad \text{in} \quad X \setminus \bigcup_{\ell} A_\ell^n.$$

Also ist (h_n) eine $\|\cdot\|_1$ -Cauchyfolge, und nach Satz 8.4 existiert ein $f : X \rightarrow \mathbb{C}$ ($f \in \mathcal{L}^1(X, \mu)$) und eine Teilfolge (h_{n_k}) so, dass für jedes $\varepsilon > 0$ eine Folge (A_q) in \mathcal{R} existiert, mit

$$\sum_q \mu(A_q) < \varepsilon \quad \text{und} \quad h_{n_k}(x) \rightarrow f(x) \quad \text{gleichmäßig in} \quad X \setminus \bigcup_q A_q.$$

Dies impliziert

$$f_{n_k}(x) \rightarrow f(x) \quad \text{gleichmäßig in} \quad X \setminus \left\{ \left(\bigcup_q A_q \right) \cup \left(\bigcup_{\ell} \bigcup_{n \geq N} A_\ell^n \right) \right\}.$$

Da dies für jedes $\varepsilon > 0$ und $N \in \mathbb{N}$ gilt, erhält man

$$f_{n_k}(x) \rightarrow f(x) \quad \mu\text{-fast gleichmäßig,}$$

und deshalb

$$\lim_{n \rightarrow \infty} \int f_n d\mu = \lim_{k \rightarrow \infty} \int f_{n_k} d\mu = \lim_{k \rightarrow \infty} \int h_{n_k} d\mu = \int f d\mu.$$

Da auch $(|f - f_n|)$ eine $\|\cdot\|_1$ -Cauchyfolge ist mit $|f(x) - f_n(x)| \rightarrow 0$ μ -f. ü., folgt $\int |f - f_n| d\mu \rightarrow \int 0 d\mu = 0$. ■

8.4 Grenzwertsätze

Die bemerkenswertesten Eigenschaften des Lebesgueschen Integrals zeigen sich im Zusammenhang mit den folgenden äußerst nützlichen Grenzwertsätzen.

Satz 8.10 (B. Levi) *Ist (f_n) eine nicht-fallende Folge in $\mathcal{L}^1(X, \mu)$ mit $\int f_n d\mu \leq C$ für alle $n \in \mathbb{N}$, dann gibt es ein $f \in \mathcal{L}^1(X, \mu)$ mit*

$$f_n(x) \rightarrow f(x) \quad \mu\text{-f. ü.} \quad \text{und} \quad \int f_n d\mu \rightarrow \int f d\mu.$$

Gilt analog für fallende Folgen (f_n) mit $\int f_n d\mu \geq C$.

Beweis. Aus der Monotonie folgt, $\int |f_n - f_m| d\mu = \int f_n d\mu - \int f_m d\mu$ für $n > m$. Also ist (f_n) eine $\|\cdot\|_1$ -Cauchyfolge, und mit Satz 8.9 impliziert dies das gewünschte Ergebnis. Wegen der Monotonie gilt die μ -f. ü.-Konvergenz nicht nur für eine Teilfolge. ■

Satz 8.11 *Für $f, g \in \mathcal{L}^1(X, \mu)$ sind auch die Funktionen $\min\{f, g\}$, $\max\{f, g\}$, f_+ , f_- , $\operatorname{Re} f$ und $\operatorname{Im} f$ in $\mathcal{L}^1(X, \mu)$.*

Beweis. Seien (f_n) und (g_n) die $\|\cdot\|_1$ -Cauchyfolgen in $E(X, \mathcal{R})$ welche nach Definition der μ -Integrierbarkeit von f und g existieren. Dann ist $(\min\{f_n, g_n\})$ eine $\|\cdot\|_1$ -Cauchyfolge in $E(X, \mathcal{R})$ mit $\min\{f_n, g_n\} \rightarrow \min\{f, g\}$ μ -f. ü.; dies impliziert $\min\{f, g\} \in \mathcal{L}^\infty(X, \mu)$. Auf die gleiche Art und Weise finden wir $\max\{f, g\}$, $f_+ = \max\{f, 0\}$, $f_- = -\min\{f, 0\} \in \mathcal{L}^1(X, \mu)$. Entsprechend folgt, dass $(\operatorname{Re} f_n)$ und $(\operatorname{Im} f_n)$ $\|\cdot\|_1$ -Cauchyfolgen sind, die μ -f. ü. gegen $\operatorname{Re} f$ bzw. $\operatorname{Im} f$ konvergieren; das impliziert, dass $\operatorname{Re} f$ und $\operatorname{Im} f$ in $\mathcal{L}^1(X, \mu)$ liegen. ■

Satz 8.12 (Lebesgue; dominierte bzw. majorisierte Konvergenz) *Seien $f_n (n \in \mathbb{N})$ und g in $\mathcal{L}^1(X, \mu)$ mit $|f_n(x)| \leq g(x)$ μ -f. ü. für alle $n \in \mathbb{N}$ und $f_n(x) \rightarrow f(x)$ μ -f. ü., dann gilt*

$$f \in \mathcal{L}^1(X, \mu) \quad \text{und} \quad \int f d\mu = \lim_{n \rightarrow \infty} \int f_n d\mu \quad (g \text{ wird als integrierbare Majorante bezeichnet}).$$

Beweis. Offenbar können wir o. E. annehmen, dass die f_n reellwertig sind (Beweis!). Aus Satz 8.11 folgt, dass $\max\{f_n, \dots, f_{n+k}\}$ in $\mathcal{L}^1(X, \mu)$ liegt, und somit folgt aus dem Satz von B. Levi 8.10

$$g_n := \sup\{f_k : k \geq n\} = \lim_{k \rightarrow \infty} \max\{f_n, \dots, f_{n+k}\} \in \mathcal{L}^1(X, \mu).$$

Weil $\int g_n d\mu \geq -\int g d\mu$ ist, kann der Satz von B. Levi (8.10) auch auf die fallende Folge (g_n) angewandt werden; dies impliziert

$$g_n(x) \searrow f(x) \quad \mu\text{-f. ü.}, \quad f \in \mathcal{L}^1(X, \mu) \quad \text{und} \quad \int g_n d\mu \rightarrow \int f d\mu.$$

Auf dem gleichen Weg erhalten wir für $h_n := \inf\{f_k : k \geq n\}$

$$h_n \nearrow f(x) \quad \mu\text{-f. ü.} \quad \text{und} \quad \int h_n d\mu \rightarrow \int f d\mu.$$

Zusammen mit $h_n \leq f_n \leq g_n$ folgt $\int f_n d\mu \rightarrow \int f d\mu$. ■

Satz 8.13 (Lemma von Fatou) Wenn (f_n) eine Folge von nichtnegativen Funktionen in $\mathcal{L}^1(X, \mu)$ ist mit $\int f_n d\mu \leq C$ für alle $n \in \mathbb{N}$ und $f_n \rightarrow f$ μ -f.ü., dann gilt

$$f \in \mathcal{L}^1(X, \mu) \quad \text{und} \quad \int f d\mu \leq \liminf_{n \rightarrow \infty} \int f_n d\mu.$$

Beweis. Für $h_n := \inf\{f_n, f_{n+1}, \dots\}$ gilt $h_n(x) \nearrow f(x)$ μ -f.ü. Aus $h_n \leq f_{n+k}$ für alle $n, k \in \mathbb{N}$ folgt,

$$\int h_n d\mu \leq \liminf_{k \rightarrow \infty} \int f_{n+k} d\mu = \liminf_{k \rightarrow \infty} \int f_k d\mu.$$

Zusammen mit dem Satz von B. Levi 8.10 ergibt dies

$$f \in \mathcal{L}^1(X, \mu) \quad \text{und} \quad \int f d\mu = \lim_{n \rightarrow \infty} \int h_n d\mu \leq \liminf_{k \rightarrow \infty} \int f_k d\mu,$$

also die Behauptung. ■

Bemerkung 8.14 Im Satz von Lebesgue kann nicht auf die integrierbare Majorante verzichtet werden, auch nicht, wenn die Folge $(\int |f_n| d\mu)$ beschränkt ist: Dazu sei $X = \mathbb{R}, \mu = \lambda$ und $f_n := \chi_{[n, n+1]}$ (charakteristische Funktion des Intervalls $[n, n+1]$). Offenbar gilt $f_n(x) \rightarrow f(x) = 0$ für alle x und $\int f_n d\mu = 1 \rightarrow 1 \neq 0 = \int f d\mu$.

Im Lemma von Fatou gilt i. allg. nicht die Gleichheit, wie das gleiche Beispiel zeigt. I. allg. kann auch nicht \liminf durch \lim ersetzt werden, weil dieser i. allg. nicht existiert, wie man z. B. an der Folge $(f_1, g_1, f_2, g_2, f_3, \dots)$ mit $g_n = 0$ sieht.

Bemerkung 8.15 (Zusammenfassung der Grenzwertsätze)

In jedem Falls sei (f_n) eine Folge aus \mathcal{L}^1 .

B. Levi: $(f_n) \nearrow, \int f_n \leq C \Rightarrow \exists f \in \mathcal{L}^1, f_n \rightarrow f$ μ -f.ü. $\int f_n \rightarrow \int f$.

Lebesgue: $g \in \mathcal{L}^1, |f| \leq g, f_n \rightarrow f$ μ -f.ü. $\Rightarrow f \in \mathcal{L}^1, \int f_n \rightarrow \int f$.

Fatou: $f_n \geq 0, \int f_n \leq C, f_n \rightarrow f$ μ -f.ü. $\Rightarrow f \in \mathcal{L}^1, \int f \leq \liminf \int f_n$.

8.5 Messbare Mengen und Funktionen, Maße

Der Einfachheit halber nehmen wir an, dass (X, μ) σ -endlich ist, d. h. es existiert eine Folge (X_n) in \mathcal{R} mit $X = \cup_n X_n$.

Eine Funktion $f : X \rightarrow \mathbb{C}$ heißt μ -messbar, wenn eine Folge (f_n) in $E(X, \mathcal{R})$ existiert mit $f_n(x) \rightarrow f(x)$ μ -f.ü. (Im nicht σ -endlichen Fall müßten wir definieren: f ist μ -messbar, wenn in jedem $A \in \mathcal{R}$ die Funktion f der μ -f.ü.-Limes einer Folge von Elementarfunktionen ist; wir überlassen es dem Leser, sich ggf. die im folgenden notwendigen Modifikationen zu überlegen.)

Die Dirichletfunktion ($= 1$ für rationale, $= 0$ für irrationale x) ist offenbar Lebesgue-messbar, da sie λ_n fast überall gleich 0 ist. Auf Grund der Definition der Integrierbarkeit ist klar, dass jede μ -integrierbare Funktion μ -messbar ist. Der folgende Satz kehrt dies Aussage teilweise um:

Satz 8.16 Wenn f μ -messbar ist, $g \in \mathcal{L}^1(X, \mu)$ und $|f(x)| \leq g(x)$ μ -f.ü., dann gilt $f \in \mathcal{L}^1(X, \mu)$.

Beweis. Sei (f_n) eine Folge aus $E(X, \mathcal{R})$ mit $f_n(x) \rightarrow f(x)$ μ -f. ü. Dann gilt nach Satz 8.11

$$\tilde{f}_n := \max \left\{ -g, \min\{f_n, g\} \right\} \in \mathcal{L}^1(X, \mu),$$

und $\tilde{f}_n(x) \rightarrow f(x)$ μ -f. ü., $|\tilde{f}_n| \leq g$. Die Anwendung des Satzes von Lebesgue 8.12 impliziert $f \in \mathcal{L}^1(X, \mu)$. ■

17: 12. Dezember 2006

Satz 8.17 a) Wenn f und g μ -messbar sind, dann sind auch $f+g, fg, f/g, |f|, \dots$ μ -messbar.

b) Wenn f_n μ -messbar ist für jedes n und $f_n(x) \rightarrow f(x)$ μ -f. ü. gilt, dann ist auch f μ -messbar.

Beweis. a) Dies ist offensichtlich nach der Definition der Messbarkeit.

b) Seien $X_n \in \mathcal{R}$ ($n \in \mathbb{N}$) disjunkt mit $X = \cup_n X_n$. Wenn $h : X \rightarrow \mathbb{R}$ definiert ist durch $h(x) = n^{-2} \mu(X_n)^{-1}$ für $x \in X_n$, dann ist $h \in \mathcal{L}^1(X, \mu)$ und $h(x) > 0$ für alle $x \in X$. Somit ist

$$g_n(x) := \frac{h(x)f_n(x)}{h(x) + |f_n(x)|} \quad \mu\text{-messbar mit } |g_n(x)| \leq h(x),$$

also $g_n \in \mathcal{L}^1(X, \mu)$ nach Satz 8.16. Für $n \rightarrow \infty$ gilt

$$g_n(x) \rightarrow g(x) := \frac{h(x)f(x)}{h(x) + |f(x)|} \quad \text{und} \quad |g(x)| \leq h(x).$$

Daraus folgt $g \in \mathcal{L}^1(X, \mu)$, woraus die μ -Messbarkeit von $f = hg/(h - |g|)$ folgt (um diese letzte Identität zu beweisen unterscheidet man die Fälle $f(x) \geq 0$ und $f(x) < 0$ und damit $g(x) \geq 0$ bzw. $g(x) < 0$). ■

Eine Teilmenge $A \subset X$ heißt μ -messbar wenn die charakteristische Funktion χ_A ($= 1$ in A , $= 0$ in $X \setminus A$) μ -messbar ist. Wenn A μ -messbar ist, dann ist das Komplement $X \setminus A$ ebenfalls messbar (Beweis!).

Das Prämaß μ wird nun, ohne dass wir die Bezeichnung ändern, fortgesetzt auf die Familie der μ -messbaren Mengen, indem wir definieren:

$$\mu(A) := \begin{cases} \int \chi_A d\mu & \text{falls } \chi_A \in \mathcal{L}^1(X, \mu), \\ \infty & \text{falls } \chi_A \notin \mathcal{L}^1(X, \mu). \end{cases}$$

Satz 8.18 a) Eine Teilmenge N von X ist genau dann eine μ -Nullmenge, wenn N μ -messbar ist mit $\mu(N) = 0$.

b) Wenn A_n ($n \in \mathbb{N}$) disjunkte μ -messbare Mengen sind, dann ist $\cup_n A_n$ μ -messbar und $\mu(\cup_n A_n) = \sum_n \mu(A_n)$.

c) Wenn (A_n) eine fallende Folge von μ -messbaren Mengen ist mit $\mu(A_n) < \infty$ für ein n , dann ist $\cap_n A_n$ μ -messbar und $\mu(\cap_n A_n) = \lim_{n \rightarrow \infty} \mu(A_n)$.

Beweis. a) \Rightarrow : Wenn N eine μ -Nullmenge ist, dann ist $\chi_N = 0$ μ -f.ü. und deshalb $\mu(N) = \int \chi_N d\mu = 0$.

\Leftarrow : Ist N μ -messbar mit $\mu(N) = 0$, so ist χ_N μ -messbar mit $\int \chi_N d\mu = 0$; dann kann der Satz von B. Levi 8.10 auf die Folge $(n \cdot \chi_N)_{n \in \mathbb{N}}$, angewandt werden; dies liefert die Existenz eines $f \in \mathcal{L}^1(X, \mu)$ mit $n \cdot \chi_N \rightarrow f$ μ -f.ü. für $n \rightarrow \infty$, also $\chi_N = 0$ μ -f.ü. Das bedeutet, dass N eine μ -Nullmenge ist.

b) Man beachte, dass

$$\chi_{\bigcup_{n=1}^m A_n} = \sum_{n=1}^m \chi_{A_n} \nearrow \sum_{n=1}^{\infty} \chi_{A_n} = \chi_{\bigcup_{n=1}^{\infty} A_n}.$$

Wenn $\sum_n \mu(A_n) < \infty$ ist, impliziert Satz von B. Levi 8.10 das Ergebnis,

$$\mu\left(\bigcup_{n=1}^{\infty} A_n\right) = \int \chi_{\bigcup_{n=1}^{\infty} A_n} d\mu = \lim_{n \rightarrow \infty} \int \sum_{n=1}^n \chi_{A_n} d\mu = \sum_{n=1}^{\infty} \mu(A_n).$$

Wenn $\sum_n \mu(A_n) = \infty$, dann ist $\chi_{\bigcup_n A_n}$ nicht integrierbar (Beweis!); dann sind beide Seiten der Gleichung unendlich.

c) In diesem Fall haben wir

$$\chi_{A_n} \searrow \chi_{\bigcap_n A_n} = \lim_{n \rightarrow \infty} \chi_{A_n},$$

und somit nach dem Satz von Lebesgue 8.12

$$\mu\left(\bigcap_{n=1}^{\infty} A_n\right) = \int \chi_{\bigcap_n A_n} d\mu = \lim_{n \rightarrow \infty} \int \chi_{A_n} d\mu = \lim_{n \rightarrow \infty} \mu(A_n),$$

womit die Behauptung bewiesen ist. ■

Eine Familie \mathcal{A} von Teilmengen von X heißt eine σ -Algebra in X , wenn gilt: $\emptyset \in \mathcal{A}, A \in \mathcal{A}$ impliziert $X \setminus A \in \mathcal{A}$, und $A_n \in \mathcal{A}$ für $n \in \mathbb{N}$ impliziert $\bigcup_n A_n \in \mathcal{A}$; insbesondere ist auch $X = X \setminus \emptyset$ in \mathcal{A} . Eine Abbildung $\mu : \mathcal{A} \rightarrow [0, \infty]$ heißt ein Maß auf (X, \mathcal{A}) , wenn gilt: Aus $A_n \in \mathcal{A}$ ($\setminus \in \mathbb{N}$) mit $A_n \cap A_m = \emptyset$ für $n \neq m$ folgt $\mu(\bigcup_n A_n) = \sum_n \mu(A_n)$; man sagt, μ ist σ -additiv. Das Tripel (X, \mathcal{A}, μ) — oder gelegentlich auch kurz (X, μ) — heißt ein Maßraum.

Der obige Satz 8.18 besagt, dass $(X, \mathcal{A}_\mu, \mu) = (\mathcal{X}, \mu)$ ein Maßraum ist, wenn \mathcal{A}_μ die Familie der μ -messbaren Mengen ist.

Auf Grund der Definition einer μ -Nullmenge gilt offenbar: Jede Teilmenge einer μ -Nullmenge ist wiederum eine μ -Nullmenge; ein Maß mit dieser Eigenschaft nennt man *vollständig*.

Ist $A \subset X$ eine μ -messbare Teilmenge, so definieren wir

$$\int_A f d\mu := \int \chi_A f d\mu \quad \text{falls } \chi_A f \in \mathcal{L}^1(X, \mu).$$

Für $A \cap B = \emptyset$ und $\chi_{A \cup B} f \in \mathcal{L}^1(X, \mu)$ gilt dann offenbar

$$\int_{A \cup B} f d\mu = \int_A f d\mu + \int_B f d\mu.$$

Satz 8.19 Eine Funktion $f : X \rightarrow \mathbb{R}$ ist genau dann μ -messbar, wenn $M_c := \{x \in X : f(x) \geq c\}$ für jedes $c \in \mathbb{R}$ μ -messbar ist (das gleiche gilt für die Teilmengen, auf denen $f(x) > c, \leq c, < c, c \leq f(x) \leq d, c < f(x) < d, c \leq f(x) < d, c < f(x) \leq d$ gilt).

Beweis. \Leftarrow : Wenn M_c für jedes $c \in \mathbb{R}$ μ -messbar ist, dann sind die Mengen

$$X \setminus M_t = \{x \in X : f(x) < t\} \quad \text{und} \quad M(s, t) := \{x \in X : s \leq f(x) < t\}$$

ebenfalls μ -messbar. Also sind die Funktionen

$$f_n(x) = \begin{cases} \frac{k}{n} & \text{für } x \in M\left(\frac{k-1}{n}, \frac{k}{n}\right) \\ 0 & \text{sonst} \end{cases} \quad \text{für } k = -n^2 + 1, \dots, n^2,$$

μ -messbar mit $f_n(x) \rightarrow f(x)$ für alle $x \in X$; daraus folgt die μ -Messbarkeit von f .

\Rightarrow : Ohne Einschränkung können wir annehmen, dass $c = 1$ gilt (andernfalls betrachtet man $f - c + 1$ statt f). Die μ -Messbarkeit von f impliziert die μ -Messbarkeit von $g := \min\{1, \max\{0, f\}\}$. Wegen $g^n(x) \rightarrow \chi_{M_c}(x)$ folgt hieraus die μ -Messbarkeit von M_c .

Der Beweis der entsprechenden Aussagen für die anderen Teilmengen kann dem Leser überlassen werden (insbesondere mit geeignetem Komplement und Durchschnittsbildungen). \blacksquare

18: 15. Dezember 2006

Auf Grund der bisher bewiesenen Eigenschaften ist es offensichtlich, dass sehr viele Teilmengen von \mathbb{R} bzw. Funktionen $f : \mathbb{R} \rightarrow \mathbb{C}$ Lebesgue-messbar sind; es scheint deshalb gar nicht selbstverständlich, dass es überhaupt nicht-Lebesgue-messbare Teilmengen von \mathbb{R} bzw. Funktionen $f : \mathbb{R} \rightarrow \mathbb{C}$ gibt. Deshalb stellen wir hier die Konstruktion einer nicht-Lebesgue-messbaren Teilmenge von \mathbb{R} vor (deren charakteristische Funktion ist dann nicht Lebesgue-messbar):

Beispiel 8.20 (Nicht Lebesgue-messbare Menge/Funktion) Sei $X = [0, 1)$. Wir nennen zwei Punkte $x, y \in [0, 1)$ äquivalent, $x \sim y$, wenn $x - y$ rational ist. Dies ist offensichtlich eine Äquivalenzrelation, und wir können X in Äquivalenzklassen zerlegen (wobei x und y genau dann zu verschiedenen Klassen gehören, wenn $x - y$ irrational ist).

Sei nun $M \subset X$ so gewählt, dass es genau ein Element aus jeder Äquivalenzklasse enthält (hier benutzen wir offenbar das Auswahlaxiom). Für jedes $r \in \mathbb{Q}_0 := \mathbb{Q} \cap [0, 1)$ (\mathbb{Q} = rationale Zahlen) sei

$$M_r := \{r + x \bmod 1 : x \in M\} \quad \text{insbesondere ist also } M_0 = M.$$

Dann gilt $M_{r_1} \cap M_{r_2} = \emptyset$ für $r_1, r_2 \in \mathbb{Q}_0, r_1 \neq r_2$, und

$$X = \bigcup_{r \in \mathbb{Q}_0} M_r.$$

Man sieht nun leicht, dass M nicht Lebesgue-messbar ist. Wäre nämlich M Lebesgue-messbar, so müßte gelten

$$(\alpha) \quad \lambda(M) = 0 \quad \text{oder} \quad (\beta) \quad \lambda(M) > 0.$$

Wir benutzen nun die offensichtliche Tatsache, dass das Lebesguemaß translationsinvariant ist, d. h. es gilt $\lambda(A) = \lambda(A + x \bmod 1)$ für jede Lebesgue-messbare Menge $A \subset [0, 1)$ und jedes $x \in \mathbb{R}$ (vgl. Aufgabe 8.2). Daraus folgt $\lambda(M_r) = \lambda(M)$ für jedes $r \in \mathbb{Q}_0$ und somit mit der σ -Additivität

$$1 = \lambda([0, 1)) = \sum_{r \in \mathbb{Q}_0} \lambda(M_r) = \begin{cases} 0 & \text{im Fall } (\alpha), \\ \infty & \text{im Fall } (\beta). \end{cases}$$

In beiden Fällen ist dies ein Widerspruch. Also ist die Menge M und somit die Funktion χ_M nicht Lebesgue-messbar. \square

Anmerkung 8.21 Ausgehend von einem Prämaß μ auf einem Mengerring \mathcal{R} in X haben wir ein Maß μ auf einer σ -Algebra in X gewonnen. Zur Erinnerung: Eine σ -Algebra \mathcal{A} in X ist eine Familie von Teilmengen von X mit

- $\emptyset \in \mathcal{A}$
- $A \in \mathcal{A} \Rightarrow X \setminus A \in \mathcal{A}$ (und damit auch $X = X \setminus \emptyset \in \mathcal{A}$),
- $A_n \in \mathcal{A}$ für $n \in \mathbb{N} \Rightarrow \bigcup_{n \in \mathbb{N}} A_n \in \mathcal{A}$.

Ist \mathcal{A}_0 eine beliebige Familie von Teilmengen von X , so gibt es immer eine kleinste σ -Algebra, die \mathcal{A}_0 enthält, die von \mathcal{A}_0 erzeugte σ -Algebra \mathcal{A} . (Dies ist der Durchschnitt aller σ -Algebren, die \mathcal{A}_0 enthalten. Es gibt mindestens eine σ -Algebra, die \mathcal{A}_0 enthält, nämlich die Potenzmenge von X .)

In \mathbb{R}^m (insbesondere auch \mathbb{R} , und allgemeiner in jedem topologischen Raum) wird die σ -Algebra, die durch die Familie der offenen (oder abgeschlossenen, oder offenen und abgeschlossenen) Teilmengen erzeugt wird als *Borelsche σ -Algebra* bezeichnet. Da die σ -Algebra der Lebesgue-messbaren Teilmengen alle offenen Mengen enthält, sind also alle Borelschen Mengen Lebesgue-messbar; dies gilt für alle durch Prämaße auf den Figuren in \mathbb{R}^m erzeugten Maße. Diese Maße werden deshalb auch als *Borelmaße* bezeichnet. Die oben konstruierte nicht Lebesgue-messbare Menge ist also sicher keine Borelmenge.

8.6 Produktmaße; der Satz von Fubini-Tonelli

Seien \mathcal{R}_X und \mathcal{R}_Y Mengerringe, μ_X und μ_Y Prämaße auf X bzw. Y . Die Mengen der Form $A \times B \subset X \times Y$ mit $A \in \mathcal{R}_X$ und $B \in \mathcal{R}_Y$ erzeugen einen Mengerring $\mathcal{R}_X \times \mathcal{R}_Y$ in $X \times Y$, und durch

$$\mu\left(\bigcup_j (A_j \times B_j)\right) := \sum_j \mu_X(A_j) \mu_Y(B_j)$$

für $A_j \in \mathcal{R}_X$, $B_j \in \mathcal{R}_Y$ mit $A_j \times B_j \cap A_k \times B_k = \emptyset$ wird ein Prämaß $\mu := \mu_X \times \mu_Y$ auf $(X \times Y, \mathcal{R}_X \times \mathcal{R}_Y)$ definiert. Mit der oben durchgeführten Konstruktion definiert dies ein Maß $\mu := \mu_X \times \mu_Y$ auf $X \times Y$, das als *Produktmaß* von μ_X und μ_Y bezeichnet wird. Das folgende Lemma ist für die Untersuchung von Produktmaßen wesentlich:

Lemma 8.22 Wenn $N \subset X \times Y$ eine μ -Nullmenge ist, und

$$N(x) = \{y \in Y : (x, y) \in N\},$$

dann ist

$$N_X := \left\{ x \in X : N(x) \text{ ist nicht } \mu_Y\text{-Nullmenge} \right\}$$

eine μ_X -Nullmenge, d. h. für μ_X -f. a. $x \in X$ ist $N(x)$ eine μ_Y -Nullmenge. (Für die Umkehrung vergleiche man Aufgabe 8.12.)

Beweis. Da N eine μ -Nullmenge ist, gibt es $C_k = A_k \times B_k$ mit $A_k \in \mathcal{R}_X$ und $B_k \in \mathcal{R}_Y$ so, dass $N \subset \bigcup_{k=1}^{\infty} C_k$, $\sum_{k=1}^{\infty} \mu(C_k) < \infty$ und, dass jedes $z \in N$ in unendlich vielen C_k enthalten ist (Beweis?). Aus

$$\sum_{k=1}^{\infty} \int \chi_{A_k} d\mu_X \int \chi_{B_k} d\mu_Y = \sum_{k=1}^{\infty} \mu_X(A_k) \mu_Y(B_k) = \sum_{k=1}^{\infty} \mu(C_k)$$

ergibt sich, dass die Folge

$$\left(\sum_{k=1}^n \chi_{A_k}(x) \int \chi_{B_k} d\mu_Y \right)_{n \in \mathbb{N}} \quad \text{aus} \quad E(X; \mathcal{R}_X)$$

nichtfallend ist mit beschränkter Integralfolge. Nach dem Satz von B. Levi 8.10 existiert eine μ_X -Nullmenge F_X mit

$$\sum_{k=1}^{\infty} \chi_{A_k}(x) \int \chi_{B_k} d\mu_Y < \infty \quad \text{in } X \setminus F_X.$$

Dies impliziert, wiederum mit dem Satz von B. Levi 8.10, dass für jedes $x \in X \setminus F_X$ die Folge $\left(\sum_{k=1}^n \chi_{A_k}(x) \chi_{B_k}(y) \right)_{n \in \mathbb{N}}$ für μ_Y -f. a. $y \in Y$ beschränkt ist. Aus der Konstruktion der Folge (C_k) ergibt sich für $x \notin F_X$, dass für μ_Y -f. a. $y \in Y$ der Punkt (x, y) nicht in N liegt, d. h. $x \notin N_X$. Also gilt $N_X \subset F_X$, d. h. N_X ist eine μ_X -Nullmenge. ■

Satz 8.23 (Satz von Fubini–Tonelli) Eine Funktion $f : X \times Y \rightarrow \mathbb{C}$ liegt genau dann in $\mathcal{L}^1(X \times Y, \mu)$, wenn gilt: f ist μ -messbar, $|f(x, \cdot)| \in \mathcal{L}^1(Y, \mu_Y)$ für μ_X -f. a. $x \in X$ und

$$F(x) := \begin{cases} \int_Y |f(x, y)| d\mu_Y(y), & \text{falls } |f(x, \cdot)| \in \mathcal{L}^1(Y, \mu_Y), \\ 0, & \text{sonst} \end{cases}$$

ist μ_X -integrierbar (\Rightarrow : Satz von Fubini, \Leftarrow : Satz von Tonelli). In diesem Fall liegt die Funktion $x \mapsto \int_Y f(x, y) d\mu_Y(y)$ in $\mathcal{L}^1(X, \mu_X)$ und es gilt

$$\int_{X \times Y} f(x, y) d\mu(x, y) = \int_X \left\{ \int_Y f(x, y) d\mu_Y(y) \right\} d\mu_X(x).$$

Entsprechendes gilt, wenn die Rollen von x und y vertauscht werden; für $f \in \mathcal{L}^1(X \times Y, \mu)$ kann die Integrationsreihenfolge vertauscht werden. – Ist $f \geq 0$ μ -messbar, so gilt $f \in L^1(X, \mu)$ genau dann, wenn eines der iterierten Integrale endlich ist; es stimmen dann die beiden iterierten Integrale mit $\|f\|_1$ überein.

Beweis. \Rightarrow (**Fubini**): Aus $f \in \mathcal{L}^1(X \times Y, \mu)$ folgt, dass eine $\|\cdot\|_1$ -Cauchyfolge (f_n) aus $E(X \times Y, \mathcal{R}_X \times \mathcal{R}_Y)$ existiert mit $f_n(x, y) \rightarrow f(x, y)$ für μ -f. a. $(x, y) \in X \times Y$. Ohne Einschränkung können wir annehmen, dass gilt (ggf. Auswahl einer Teilfolge von (f_n))

$$f_n(x, y) = \sum_{j=1}^n h_j(x, y) \quad \text{mit} \quad \sum_{j=1}^{\infty} \|h_j\|_1 < \infty.$$

Aus Lemma 8.22 folgt, dass es eine μ_X -Nullmenge N_1 gibt so, dass gilt

$$f_n(x, y) \rightarrow f(x, y) \quad \text{für } \mu_Y\text{-f. a. } y \in Y \text{ und alle } x \in X \setminus N_1. \quad (1)$$

Da die Behauptung offensichtlich für Elementarfunktionen gilt, ist

$$\left(\sum_{j=1}^n \int_Y |h_j(x, y)| \, d\mu_Y(y) \right)_{n \in \mathbb{N}} \quad \text{eine } \|\cdot\|_1\text{-Cauchyfolge in } \mathcal{L}^1(X, \mu_X), \quad (2)$$

und somit ist auch

$$\left(\int_Y f_n(x, y) \, d\mu_Y(y) \right)_{n \in \mathbb{N}} \quad \text{eine } \|\cdot\|_1\text{-Cauchyfolge in } \mathcal{L}^1(X, \mu_X). \quad (3)$$

Es gibt eine μ_X -Nullmenge N_2 so, dass die Folge in (3) für $x \in X \setminus N_2$ konvergiert; also ist

$$\left(f_n(x, \cdot) \right) \quad \text{eine } \|\cdot\|_1\text{-Cauchyfolge in } \mathcal{L}^1(Y, \mu_Y) \text{ für } x \in X \setminus N_2. \quad (4)$$

Aus (1) und (4) folgt, dass $f(x, \cdot) \in \mathcal{L}^1(Y, \mu_Y)$ gilt (das ist die erste Aussage), und

$$\int_Y f_n(x, y) \, d\mu_Y(y) \rightarrow \int_Y f(x, y) \, d\mu_Y(y) \quad \text{für } x \in X \setminus (N_1 \cup N_2). \quad (5)$$

Aus (3) und (5) folgt die behauptete Gleichung:

$$\begin{aligned} \int \left\{ \int f(x, y) \, d\mu_Y(y) \right\} \, d\mu_X(x) &= \lim_{n \rightarrow \infty} \int \left\{ \int f_n(x, y) \, d\mu_Y(y) \right\} \, d\mu_X(x) \\ &= \lim_{n \rightarrow \infty} \int f_n \, d(\mu_X \times \mu_Y) = \int f \, d(\mu_X \times \mu_Y) = \int f(x, y) \, d\mu(x, y). \end{aligned}$$

Wenn wir dieses Resultat auf $|f|$ anwenden, erhalten wir die Behauptung für F .

\Leftarrow (**Tonelli**): Sei $A_n \in \mathcal{R}_X, B_n \in \mathcal{R}_Y$ mit $A_n \subset A_{n+1}, B_n \subset B_{n+1}, X = \cup_n A_n, Y = \cup_n B_n$,

$$f_n(x, y) = \begin{cases} f(x, y) & \text{falls } |f(x, y)| \leq n \text{ und } (x, y) \in A_n \times B_n, \\ 0 & \text{sonst.} \end{cases}$$

Dann ist $f_n \in \mathcal{L}^1(X \times Y, \mu_X \times \mu_Y)$, und die Folge $(|f_n|)$ ist nichtfallend mit $f_n(x, y) \rightarrow f(x, y)$. Aus der \Rightarrow -Richtung dieses Satzes folgt

$$\int |f_n| \, d(\mu_X \times \mu_Y) = \int \left\{ \int |f_n| \, d\mu_Y \right\} \, d\mu_X \leq \int \left\{ \int |f| \, d\mu_Y \right\} \, d\mu_X < \infty$$

und somit mit dem Satz von B. Levi 8.10

$$|f| \in \mathcal{L}^1(X \times Y, \mu_X \times \mu_Y).$$

Da f μ -messbar ist, folgt $f \in \mathcal{L}^1(X \times Y, \mu_X \times \mu_Y)$. ■

19: 19. Dezember 2006

Im Fall des Lebesguemaßes auf \mathbb{R}^m schreiben wir das Integral in der Form

$$\int f(x) dx \quad \text{bzw.} \\ \int f(x_1, \dots, x_n) d(x_1, \dots, x_n) = \int \dots \int f(x_1, \dots, x_n) dx_1 \dots dx_n.$$

8.7 Der Satz von Radon–Nikodym

Sind μ und ν Maße auf X (mit entsprechenden σ -Algebren \mathcal{A}_μ und \mathcal{A}_ν von μ - bzw. ν -messbaren Mengen), so heißt ν *absolut stetig* bezüglich μ (oder: *μ -absolut stetig*; kurz $\nu \ll \mu$), falls jede μ -messbare Menge auch ν -messbar ist (d. h. $\mathcal{A}_\mu \subset \mathcal{A}_\nu$) und jede μ -Nullmenge auch eine ν -Nullmenge ist.

Ist $h \in \mathcal{L}^1(X, \mu)$, $h \geq 0$, so wird durch

$$\nu(A) := \int \chi_A(x) h(x) d\mu(x) \quad \text{für } \mu\text{-messbares } A$$

offenbar ein Maß auf X definiert (auf der σ -Algebra \mathcal{A}_μ), das absolut stetig bezüglich μ ist: Ist A eine μ -Nullmenge, so ist $\chi_A(x)h(x) = 0$ μ -f.ü., also $\nu(A) = 0$. Der folgende Satz liefert die wichtige Umkehrung dieser Aussage.

Satz 8.24 (Radon–Nikodym) *Seien μ und ν Maße auf X , (X, μ) sei σ -endlich. Ist ν absolut stetig bezüglich μ und $\nu(X) < \infty$, so existiert ein $h \in L^1(X, \mu)$ mit*

$$\nu(A) = \int \chi_A h d\mu \quad \text{für jede } \mu\text{-messbare Menge } A, \quad (1)$$

$$\int f d\nu = \int fh d\mu \quad \text{für jedes } f \in L^1(X, \nu); \quad (2)$$

h heißt die μ -Dichte von ν , oder die Radon–Nikodym–Ableitung von ν bezüglich μ .

Beweis. (Der Beweis dieses Satzes wird in der Vorlesung größtenteils übergangen, da er wesentlich ein Hilfsmittel aus der Hilbertraumtheorie benutzt.) Es genügt, den Fall $\mu(X) < \infty$ zu betrachten; der allgemeine Fall folgt leicht mit Hilfe der σ -Endlichkeit von μ durch Zusammenstückeln.

Sei $\tau := \mu + \nu$ mit $\mathcal{A}_\tau := \mathcal{A}_\mu$. Dann ist die Abbildung

$$L^2(X, \tau) \rightarrow \mathbb{K}, \quad f \mapsto \int f d\nu$$

ein stetiges lineares Funktional, denn aus der Schwarzischen Ungleichung folgt für alle $f \in L^2(X, \tau)$

$$\left| \int f d\nu \right| \leq \int |f| d\tau = \int 1|f| d\tau \leq \tau(X)^{1/2} \|f\|_2.$$

Nach dem Rieszschen Darstellungssatz (Satz ??) gibt es also ein $g \in L^2(X, \tau)$ mit

$$\int f d\nu = \int gf d\tau \quad \text{für jedes } f \in L^2(X, \tau). \quad (1)$$

Wegen $\chi_A \in L^2(X, \tau)$ für jede μ -messbare Teilmenge A von X folgt hieraus

$$\nu(A) = \int \chi_A \, d\nu = \int g\chi_A \, d\tau = \int_A g \, d\tau.$$

Aus $0 \leq \nu(A) \leq \tau(A)$ folgt $0 \leq \int_A g \, d\tau \leq \tau(A)$ für jede μ -messbare Teilmenge A von X , also

$$0 \leq g(x) \leq 1 \quad \text{für } \tau\text{-f. a. } x \in X$$

(wäre $g(x) < 0$ oder > 1 auf einer Teilmenge M mit $\tau(M) > 0$, dann würde $\int_M g \, d\tau < 0$ bzw. $\int_M g \, d\tau > \tau(M)$ folgen).

Aus (1) folgt

$$\begin{aligned} \int (1-g)f \, d\nu &= \int gf \, d\tau - \int gf \, d\nu \\ &= \int gf \, d\mu \quad \text{für jedes } f \in L^2(X, \tau). \end{aligned} \tag{2}$$

Sei nun

$$N := \{x \in X : g(x) = 1\}, \quad L := X \setminus N.$$

Aus (2) folgt dann

$$\mu(N) = \int \chi_N \, d\mu = \int g\chi_N \, d\mu = \int (1-g)\chi_N \, d\nu = 0$$

also $\nu(N) = 0$, da ν absolut stetig bezüglich μ ist. Wählen wir $f = (1 + g + g^2 + \dots + g^n)\chi_A$ in (2), so folgt

$$\begin{aligned} \int_A (1 - g^{n+1}) \, d\nu &= \int_A (1-g)(1 + g + g^2 + \dots + g^n) \, d\nu \\ &= \int_A g(1 + g + g^2 + \dots + g^n) \, d\mu. \end{aligned}$$

Die Funktionenfolgen unter den Integralen sind nicht-fallend, die Folgen der Integrale sind beschränkt durch $\int_A 1 \, d\nu = \nu(A)$. Die Folge unter dem linken Integral konvergiert punktweise gegen χ_L ; nach dem Satz von B. Levi 8.10 konvergiert also das Integral auf der rechten Seite gegen $\nu(L \cap A)$. Wiederum aus dem Satz von B. Levi 8.10 folgt also die Existenz einer Funktion $h \in L^1(X, \mu)$ mit

$$g(1 + g + g^2 + \dots + g^n) \rightarrow h \quad \mu\text{-f. ü.},$$

und

$$\int_A g(1 + g + g^2 + \dots + g^n) \, d\mu \rightarrow \int_A h \, d\mu$$

für jede μ -messbare Teilmenge $A \subset X$. Da N eine ν -Nullmenge ist, folgt daraus

$$\nu(A) = \nu(L \cap A) + \nu(N \cap A) = \nu(L \cap A) = \int_A h \, d\mu.$$

Die zweite Behauptung ist offensichtlich für Elementarfunktionen und folgt für den allgemeinen Fall durch Grenzübergang. ■

Bemerkung 8.25 Die Voraussetzung der σ -Endlichkeit von μ im Satz von Radon–Nikodym 8.24 ist wesentlich: Sei $X = [0, 1]$ und μ das Zählmaß eingeschränkt auf die σ -Algebra der Lebesgue-messbaren Teilmengen von $[0, 1]$. Da nur die leere Menge eine μ -Nullmenge ist, ist es offensichtlich, dass das Lebesgue-Maß λ absolut stetig bezüglich μ ist. Wir zeigen, dass λ keine Dichte bezüglich μ hat: Für eine solche Dichte h müßte gelten, entweder

(a) $h \equiv 0$, also $\lambda(A) = 0$ für jede Lebesgue-messbare Menge $A \subset [0, 1]$, oder

(b) es gibt ein $x_0 \in [0, 1]$ mit $h(x_0) \neq 0$, also $\lambda(\{x_0\}) = h(x_0) \neq 0$.

Beide Möglichkeiten führen also zu einem Widerspruch.

8.8 L^p -Räume

Sei (X, μ) ein Maßraum, $\mathcal{L}^1(X, \mu)$ haben wir bereits definiert (der Raum der (absolut) integrierbaren Funktionen). Wir definieren jetzt allgemeiner für $1 \leq p < \infty$:

$$\begin{aligned} \mathcal{L}^p(X, \mu) &= \left\{ f : X \rightarrow \mathbb{C}, f \text{ } \mu\text{-messbar, } |f|^p \in \mathcal{L}^1(X, \mu) \right\} \\ \|f\|_p &:= \left\{ \int_X |f(x)|^p d\mu(x) \right\}^{1/p}. \end{aligned}$$

$\|\cdot\|_p$ ist eine Halbnorm auf $\mathcal{L}^p(X, \mu)$. Nur die Dreiecksungleichung macht etwas Mühe (man findet den Beweis in jeder Einführung in die Funktionalanalysis). Andererseits ist offensichtlich, dass $\|\cdot\|_p$ keine Norm ist. Definiert man

$$\mathcal{N}_p(X, \mu) = \left\{ f \in \mathcal{L}^p(X, \mu) : \|f\|_p = 0 \right\} = \left\{ f \in \mathcal{L}^p(X, \mu) : f(x) = 0 \text{ } \mu\text{-f.ü.} \right\},$$

so ist dies offensichtlich ein Untervektorraum von $\mathcal{L}^p(X, \mu)$. Definiert man

$$L^p(X, \mu) = \mathcal{L}^p(X, \mu) / \mathcal{N}_p(X, \mu),$$

d. h. man identifiziert Elemente aus $\mathcal{L}^p(X, \mu)$, die μ -f.ü. übereinstimmen, so ist $\|f\|_p$ gleich für alle Elemente einer Äquivalenzklasse, und $\|f\|_p = 0$ gilt nur für die Elemente aus $\mathcal{L}^p(X, \mu)$, die äquivalent zur Nullfunktion sind. Also ist

$$L^p(X, \mu) \text{ mit } \|\cdot\|_p \quad (\text{auch } (\mathcal{L}^p(X, \mu), \|\cdot\|_p))$$

ein normierter Raum (die Norm eines Elements dieses Raumes erhält man, indem man $\|f\|_p$ für einen beliebigen Repräsentanten ausrechnet).

Für $X \subset \mathbb{R}^m$ mit dem Lebesgue-Maß λ schreiben wir meist $L^p(x)$ statt $L^p(X, \lambda)$.

Für die Anwendungen ist wichtig:

Satz 8.26 Für jedes $p \in [1, \infty)$ ist $L^p(X, \mu)$ mit $\|\cdot\|_p$ vollständig (d. h. jede $\|\cdot\|_p$ -Cauchyfolge ist konvergent). Weiter gilt: Jede Cauchyfolge enthält eine μ -f.ü. konvergente Teilfolge.

Beweis. Sei (f_n) eine Cauchyfolge (wir denken uns für jedes f_n einen beliebigen – aber fest gehaltenen – Repräsentanten, den wir wieder als f_n bezeichnen). Zu jedem $j \in \mathbb{N}$ existiert ein $n_j \in \mathbb{N}$ mit

$$\|f_n - f_m\|_p \leq \frac{1}{2^j} \text{ für } n, m \geq n_j,$$

also

$$\|f_{n_{j+1}} - f_{n_j}\|_p \leq \frac{1}{2^j} \text{ für alle } j.$$

Sei

$$g_k(x) := \sum_{j=1}^k |f_{n_{j+1}}(x) - f_{n_j}(x)|.$$

Dann ist die Folge $(g_k(\cdot)^p)$ nicht-fallend, und es gilt (Dreiecksungleichung)

$$\int g_k(x)^p d\mu(x) = \|g_k\|_p^p \leq \left\{ \sum_{j=1}^k \frac{1}{2^j} \right\}^p < 1.$$

Aus dem Satz von B. Levi folgt, dass $(g_k(\cdot)^p)$, also auch $(g_k(\cdot))$, μ -f.ü. konvergiert. Also ist auch

$$f_{n_k}(x) = f_{n_1}(x) + \sum_{j=1}^{k-1} (f_{n_{j+1}}(x) - f_{n_j}(x))$$

μ -f.ü. konvergent gegen eine μ -messbare Funktion f .

Es bleibt zu zeigen, dass $f \in \mathcal{L}^p(X, \mu)$ ist und $\|f_n - f\|_p \rightarrow 0$ gilt: Für jedes $\varepsilon > 0$ seien $n(\varepsilon)$ und $j(\varepsilon)$ so gewählt, dass gilt

$$\int |f_{n_j}(x) - f_n(x)|^p d\mu(x) = \|f_{n_j} - f_n\|_p^p < \varepsilon \text{ für } j > j(\varepsilon) \text{ und } n > n(\varepsilon).$$

Die Folge $(|f_{n_j}(x) - f_n(x)|^p)_j$ ist also nicht-negativ mit beschränkter Integralfolge und

$$|f_{n_j}(x) - f_n(x)|^p \rightarrow |f(x) - f_n(x)|^p \text{ für } j \rightarrow \infty, \mu \text{-f.a. } x \in X.$$

Mit dem Lemma von Fatou folgt

$$|f - f_n|^p \in \mathcal{L}^1(X, \mu), \quad \int |f - f_n|^p d\mu < \varepsilon \text{ für } n \geq n(\varepsilon).$$

Also ist $f \in \mathcal{L}^p(X, \mu)$ und es gilt $f_n \rightarrow f$. Mit der durch f repräsentierten Äquivalenzklasse ist damit der Satz bewiesen. ■

Die Tatsache, dass $L^p(X, \mu)$ vollständig ist, wogegen z. B. $C[a, b]$ mit der Norm $\|\cdot\|_p$ nicht vollständig ist, ist einer der Hauptgründe, weshalb das Lebesgueintegral eingeführt wird.

Der Raum $\mathcal{L}^\infty(X, \mu)$ ⁵ ist der Raum der μ -messbaren, *wesentlich beschränkten* Funktionen auf X , d. h. der μ -messbaren Funktionen, für die ein $C = C(f)$ existiert mit

$$|f(x)| \leq C(f) \quad \mu\text{-f.ü.}$$

Das kleinste $C(f)$ dieser Art wird als $\|f\|_\infty$ bezeichnet (das *wesentliche Supremum* von $|f|$). Identifiziert man auch hier Funktionen, die μ -f.ü. übereinstimmen, so erhält man den Raum $L^\infty(X, \mu)$. Auf diesem Raum ist $\|\cdot\|_\infty$ eine Norm, und es gilt

Satz 8.27 $L^\infty(X, \mu)$ ist vollständig bezüglich der Norm $\|\cdot\|_\infty$.

Der Beweis ist wesentlich einfacher als der des vorhergehenden Satzes.

20: 22. Dezember 2006

⁵Es sei an die ∞ -Norm auf $C[a, b]$ erinnert:

$$\|f\|_\infty := \max\{x \in [a, b]\}.$$

Der Raum $C[a, b]$ ist bezüglich der Norm $\|\cdot\|_\infty$ offenbar vollständig (jede Cauchy-Folge konvergiert gleichmäßig).

8.9 Übungen

8.1 Sei μ das gemäß Beispiel 8.2 durch eine nichtfallende rechtsstetige Funktion ϕ erzeugte Lebesgue–Stieltjessche Prämaß auf \mathbb{R} .

- Zu jedem Intervall I und jedem $\varepsilon > 0$ gibt es ein offenes Intervall J mit $I \subset J$ und $\mu(J) < \mu(I) + \varepsilon$.
- Ist I ein abgeschlossenes Intervall und (I_n) eine disjunkte Folge von Intervallen mit $I = \cup_n I_n$, so gilt $\mu(I) = \sum_n \mu(I_n)$.
- μ ist σ –additiv.

8.2 a) Man beweise die Substitutionsregel für das Lebesgueintegral mit Hilfe der Substitutionsregel für das Riemannintegral.

- Mit Hilfe von Teil a beweise man die Translationsinvarianz des Lebesgue–Maßes ($\lambda(A) = \lambda(A + x)$ für jede Lebesgue–messbare Menge $A \subset \mathbb{R}$ und jedes $x \in \mathbb{R}$).

8.3 Sei \mathcal{R} der Mengenring in \mathbb{R} , der aus den endlichen Vereinigungen von Intervallen besteht, $\varrho : \mathcal{R} \rightarrow [0, \infty)$ sei definiert durch

$$\varrho(A) := \begin{cases} 1 & \text{falls ein } \varepsilon > 0 \text{ existiert mit } (0, \varepsilon) \subset A, \\ 0 & \text{sonst.} \end{cases}$$

Dann gilt: aus $A_1, A_2, \dots, A_m \in \mathcal{R}$ und $A_n \cap A_\ell = \emptyset$ für $n \neq \ell$ folgt

$$\varrho\left(\bigcup_{n=1}^m A_n\right) = \sum_{n=1}^m \varrho(A_n) \quad (\text{endlich additiv}).$$

ϱ ist aber nicht σ –additiv.

8.4 a) Für ein Kompaktum $K \subset \mathbb{R}^m$ sei $f : K \rightarrow \mathbb{R}$ Riemann–integrierbar. Dann ist f Lebesgue–integrierbar.

b) Für eine Funktion $f : \mathbb{R}^m \rightarrow \mathbb{R}$ seien f und $|f|$ uneigentlich Riemann–integrierbar. Dann ist f Lebesgue–integrierbar.

8.5 Die *Cantor–Menge* C entsteht aus dem Intervall $[0, 1]$, indem zunächst das offene mittlere Drittel herausgenommen wird, aus den zwei verbleibenden Intervallen wieder jeweils das offene mittlere Drittel usw., also

$$C = [0, 1] \setminus \left\{ \left(\frac{1}{3}, \frac{2}{3}\right) \cup \left(\frac{1}{9}, \frac{2}{9}\right) \cup \left(\frac{7}{9}, \frac{8}{9}\right) \cup \dots \right\}.$$

- C ist eine Lebesgue–Nullmenge.
- C besteht genau aus den Punkten $a \in \mathbb{R}$ mit $a = \sum_{j=1}^{\infty} a_j 3^{-j}$ mit $a_j \in \{0, 2\}$.
- C hat die Mächtigkeit von $[0, 1]$ (ist also insbesondere nicht abzählbar).

8.6 Eine verallgemeinerte Cantor-Menge, C' erhält man, indem man zunächst aus $[0, 1]$ ein offenes Intervall herausnimmt, aus den zwei verbleibenden abgeschlossenen Intervallen jeweils wieder ein offenes Intervall etc.

- a) $D := (0, 1) \setminus C'$ ist offen und C' der Rand von D .
- b) Sei $\varepsilon > 0$. Bei geeigneter Wahl der offenen Intervalle, die man herausnimmt, erhält man eine verallgemeinerte Cantor-Menge C' mit Lebesguemaß $> 1 - \varepsilon$.

8.7 Sei $f : \mathbb{R}^m \rightarrow \mathbb{R}$ beschränkt, $f(x) = 0$ für $|x| \geq R$,

$$\begin{aligned}\bar{f}(x) &:= \inf_Q \left\{ \sup\{f(y) : y \in Q\} : x \in \overset{\circ}{Q} \right\}, \\ \underline{f}(x) &:= \sup_Q \left\{ \inf\{f(y) : y \in Q\} : x \in \overset{\circ}{Q} \right\},\end{aligned}$$

wobei Q Quader in \mathbb{R}^m sind und $\overset{\circ}{A}$ das Innere der Menge A bezeichnet.

- a) Es gilt $\underline{f}(x) \leq \bar{f}(x)$ für alle x , und $\bar{f}(x) = \underline{f}(x) = f(x)$ genau dann, wenn f im Punkt x stetig ist.
- b) Für jedes $n \in \mathbb{N}$ sei $\{Q_{n,j} : j \in \mathbb{N}\}$ eine disjunkte Überdeckung von \mathbb{R}^m mit Quadern, deren maximale Kantenlänge für $n \rightarrow \infty$ gegen 0 konvergiert,

$$\begin{aligned}\underline{f}_n(x) &= \inf\{f(y) : y \in Q_{n,j}\} \\ \bar{f}_n(x) &= \sup\{f(y) : y \in Q_{n,j}\}\end{aligned} \quad \text{für } x \in Q_{n,j}.$$

Dann gilt für alle x , die nicht auf einer Kante der $Q_{n,j}$ liegen

$$\underline{f}_n(x) \rightarrow \underline{f}(x), \quad \bar{f}_n(x) \rightarrow \bar{f}(x).$$

- c) \bar{f} und \underline{f} sind Lebesgue-integrierbar und es gilt

$$\overline{\int} f(x) dx = \int \bar{f}(x) dx, \quad \underline{\int} f(x) dx = \int \underline{f}(x) dx,$$

wobei $\overline{\int}$ bzw. $\underline{\int}$ das Ober- bzw. Unterintegral, \int das Lebesgueintegral bedeuten.

- d) f ist genau dann Riemann-integrierbar, wenn f fast überall (bzgl. des Lebesgue-Maßes) stetig ist.

8.8 Jedes $f \in \mathcal{L}_1(X, \mu)$ läßt sich schreiben in der Form $f = f_1 - f_2 + i(f_3 - f_4)$, wobei für jedes f_j eine nicht fallende $\|\cdot\|_1$ -Cauchyfolge $(g_{j,n})_n$ aus $E(X, \mathcal{R}_X)$ existiert mit $g_{j,n}(x) \rightarrow f_j(x)$ μ -f. ü.

8.9 Sei $f : [a, b] \mapsto \mathbb{R}$ Lebesgue-integrierbar. Ist $\int_I f dx = 0$ für jedes Intervall $I \subset [a, b]$, so gilt $f(x) = 0$ fast überall. (Es genügt auch, offene oder abgeschlossene Intervalle zu benutzen).
Anleitung: Man benutze z. B. die folgenden Beweisschritte:

- a) $\int g f dx = 0$ für jede Treppenfunktion g ,
- b) $\int_A f dx = 0$ für jede Lebesgue-messbare Teilmenge A von $[a, b]$,

- c) ist $h : [a, b] \rightarrow \mathbb{R}$ Lebesgue-integrierbar mit $h(x) \geq 0$ f. ü. und $\int_{[a,b]} h d\mu = 0$, so gilt $h(x) = 0$ f. ü.

8.10 a) Zu jeder Familie von Teilmengen einer Menge X gibt es eine kleinste σ -Algebra, die diese Mengen enthält. Die auf diese Weise aus den offenen Mengen (oder den abgeschlossenen Mengen) des \mathbb{R}^m erzeugte Familie ist die *Borelsche σ -Algebra* \mathcal{B} , ihre Elemente heißen *Borel-Mengen*.

- b) Ist ϱ ein Maß auf \mathbb{R}^m , das im Sinne dieses Kapitels aus einem Prämaß auf den Figuren im \mathbb{R}^m erzeugt wird (z. B. das Lebesgue-Maß), so ist jede Borel-Menge eine ϱ -messbare Menge.
- c) Eine Funktion auf \mathbb{R}^m heißt *Borel-messbar*, wenn das Urbild jeder offenen Menge eine Borel-Menge ist. Ist ϱ wie in Teil b, so gilt: Jede Borel-messbare Funktion ist ϱ -messbar.

8.11 Sei (X, μ) ein Maßraum. Eine Funktion $f : X \rightarrow \mathbb{R}$ ist genau dann messbar, wenn für jede Borel-Menge $A \subset \mathbb{R}$ das Urbild $\{x \in X : f(x) \in A\}$ μ -messbar ist.

8.12 Es gilt auch die Umkehrung von Lemma 8.22: Ist N eine $\mu_X \times \mu_Y$ -messbare Teilmenge von $X \times Y$ mit der Eigenschaft, dass für μ_X -f. a. $x \in X$ die Menge $N(x) = \{y \in Y : (x, y) \in N\}$ eine μ_Y -Nullmenge ist, so ist N eine $\mu_X \times \mu_Y$ -Nullmenge. *Anleitung:* Satz von Tonelli 8.23.

8.13 a) Für $-\infty < a < b < \infty$ und $1 \leq p < q \leq \infty$ gilt $\mathcal{L}^q(a, b) \subset \mathcal{L}^p(a, b)$.

- b) Für $f \in \mathcal{L}^\infty(a, b)$ gilt $\|f\|_\infty = \lim_{p \rightarrow \infty} \|f\|_p$.

Hinweis: Es gilt $|f(x)| \leq \|f\|_\infty$ f. ü., und $\forall \varepsilon > 0 \exists A_\varepsilon \subset (a, b)$ mit $\lambda(A_\varepsilon) > 0$ und $|f(x)| > \|f\|_\infty - \varepsilon$ für $x \in A_\varepsilon$.

- c) Für $1 \leq p, q \leq \infty$ und $p \neq q$ gilt keine der Inklusionen $\mathcal{L}^p(\mathbb{R}) \subset \mathcal{L}^q(\mathbb{R})$ und $\mathcal{L}^q(\mathbb{R}) \subset \mathcal{L}^p(\mathbb{R})$.

8.14 a) Zu jedem $f \in \mathcal{L}^p(\mathbb{R})$ ($1 \leq p < \infty$) und jedem $\varepsilon > 0$ gibt es ein $C = C(\varepsilon, f)$ und ein $g \in \mathcal{L}^\infty(\mathbb{R})$ mit $|g(x)| \leq C$, $g(x) = 0$ für $|x| \geq C$ und $\|f - g\|_p < \varepsilon$.

- b) Zu jedem $f \in \mathcal{L}^p(\mathbb{R})$ ($1 \leq p < \infty$) und jedem $\varepsilon > 0$ gibt es eine Treppenfunktion g mit $\|f - g\|_p < \varepsilon$.

- c) Zu jedem $f \in \mathcal{L}^p(\mathbb{R})$ ($1 \leq p < \infty$) und jedem $\varepsilon > 0$ gibt es eine Treppenfunktion g mit rationalen Werten, deren Konstanzintervalle rationale Randpunkte haben, und für die gilt $\|f - g\|_p < \varepsilon$.

- d) In $\mathcal{L}^p(\mathbb{R})$ gibt es eine abzählbare dichte Teilmenge (d. h. $\mathcal{L}^p(\mathbb{R})$) ist *separabel*).

8.15 Eine Funktion $f : [a, b] \rightarrow \mathbb{C}$ heißt *absolut stetig*, wenn es ein $g \in \mathcal{L}^1(a, b)$ gibt mit

$$f(x) = \int_a^x g(t) dt.$$

Die Funktion g heißt *Lebesgue-Ableitung* von f .

- a) Jede absolut stetige Funktion $[a, b] \rightarrow \mathbb{C}$ ist stetig.

b) Es gibt stetige Funktionen, die nicht absolut stetig sind.

Hinweis: z. B. $f : [0, 1] \rightarrow \mathbb{R}$, $f(x) = x \sin \frac{1}{x^2}$ für $x \neq 0$, $f(0) = 0$.

c) Ist $f : [a, b] \rightarrow \mathbb{C}$ absolut stetig, so gibt es zu jedem $\varepsilon > 0$ ein $\delta > 0$ mit: Für $a \leq x_1 < y_1 \leq x_2 < y_2 \leq \dots \leq x_n < y_n \leq b$ mit

$$\sum_{k=1}^n |y_k - x_k| < \delta \quad \text{gilt} \quad \sum_{k=1}^n |f(y_k) - f(x_k)| < \varepsilon.$$

9 Kurven, Kurvenlänge und Kurvenintegrale

9.1 Parameterdarstellung

Eine Kurve Γ in \mathbb{R}^m denken wir uns in der Regel gegeben durch eine Parameterdarstellung

$$\gamma : [a, b] \rightarrow \mathbb{R}^m, \quad t \mapsto \gamma(t),$$

wobei z. B. $\gamma(t)$ den Ort eines Teilchens (hier also in $\mathbb{R}^1, \mathbb{R}^2$ oder \mathbb{R}^3) zur Zeit t beschreibt. Meist werden wir annehmen, dass γ zumindest stückweise stetig differenzierbar ist.

$\gamma'(t)$ (in der Physik meist $\dot{\gamma}(t)$) ist dann die Geschwindigkeit, mit der der Punkt $\gamma(t)$ durchlaufen wird (bzw. die Geschwindigkeit zur Zeit t).

Einfache Beispiele sind:

Beispiel 9.1 a) *Strecke* von x nach y :

$$\gamma : [0, 1] \rightarrow \mathbb{R}^m, \quad \gamma(t) := x + t(y - x).$$

Es kann aber auch eine andere Parameterdarstellung interessant sein:

$$\begin{aligned} \tilde{\gamma} : \left[-\frac{\pi}{2}, \frac{\pi}{2}\right] &\rightarrow \mathbb{R}^m, & \tilde{\gamma}(t) &:= \frac{1}{2}(x + y + (\sin t)(y - x)), \\ \hat{\gamma} : [-1, 1] &\rightarrow \mathbb{R}^m, & \hat{\gamma}(t) &:= \frac{x + y}{2} + t \frac{y - x}{2}. \end{aligned}$$

b) *Kreislinie* mit Radius r um den Nullpunkt in \mathbb{R}^2 (im mathematisch positiven Sinn, d. h. entgegen dem Uhrzeigersinn, durchlaufen):

$$\gamma : [0, 2\pi] \rightarrow \mathbb{R}^2, \quad \gamma(t) := (r \cos t, r \sin t),$$

oder

$$\tilde{\gamma} : [0, 1] \rightarrow \mathbb{R}^2, \quad \tilde{\gamma}(t) := (r \cos 2\pi t, r \sin 2\pi t),$$

oder, mehrfach durchlaufen (k -fach):

$$\gamma_k : [0, 2\pi] \rightarrow \mathbb{R}^2, \quad \gamma_k(t) := (r \cos kt, r \sin kt).$$

c) *Schraubenlinie* mit *Radius* r und *Ganghöhe* c :

$$\gamma : \mathbb{R} \rightarrow \mathbb{R}^3, \quad \gamma(t) := (r \cos t, r \sin t, \frac{c}{2\pi}t),$$

oder, nur ein *Umlauf*:

$$\gamma_1 : [0, 1] \rightarrow \mathbb{R}^3, \quad \gamma_1(t) := (r \cos 2\pi t, r \sin 2\pi t, ct).$$

d) *NEILSche Parabel*: $\gamma : \mathbb{R} \rightarrow \mathbb{R}^2, \gamma(t) = (t^2, t^3)$.

e) Die *Zykloide* ist die Bahn eines Punktes auf dem Rand des Einheitskreises, der auf der x -Achse abrollt:

$$\gamma : \mathbb{R} \rightarrow \mathbb{R}^2, \quad \gamma(t) = (t - \sin t, 1 - \cos t),$$

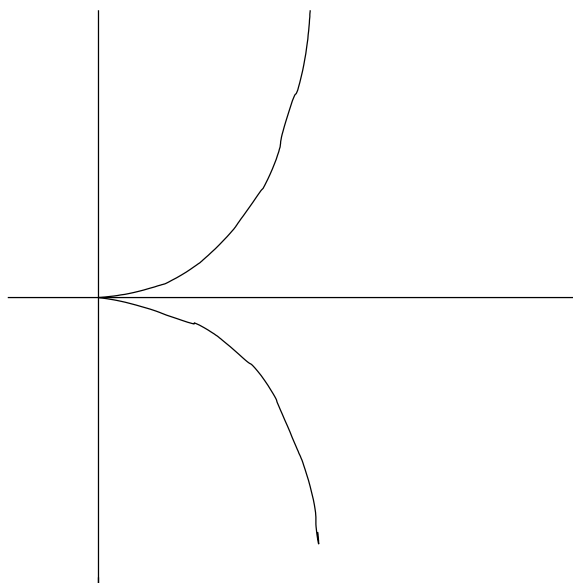


Abbildung 2: Neilsche Parabel

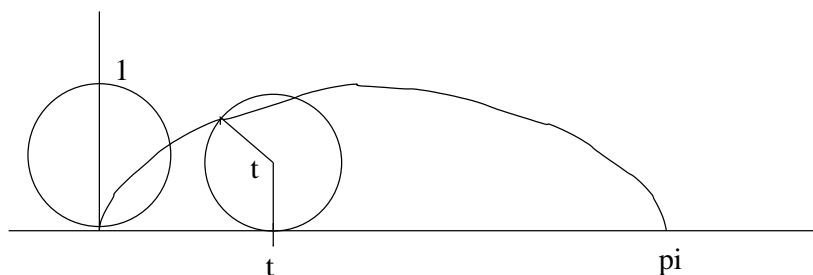


Abbildung 3: Zykloide

f) *Logarithmische Spirale:*

$$r : \mathbb{R} \rightarrow (0, \infty), \quad r(\varphi) = ae^\varphi \quad \text{bzw.} \quad \varphi(r) = \ln\left(\frac{r}{a}\right).$$

Die zweite Darstellung gibt der Kurve ihren Namen. Eine Parameterdarstellung ist

$$\gamma(\varphi) = (ae^\varphi \cos \varphi, ae^\varphi \sin \varphi).$$

□

Bei der Neilschen Parabel und der Zykloide fällt auf, dass die Kurve, obwohl die Parameterdarstellung beliebig differenzierbar ist, in einzelnen Punkten nicht „glatt“ ist. Woran liegt das? Zunächst fällt auf, dass dort die „Geschwindigkeit“ $\gamma'(t)$ verschwindet. Die Kurve kann deshalb „plötzlich“ die Richtung ändern.

Die meisten Kurven lassen sich, zumindest stückweise, als Graphen darstellen: $\gamma(t) = (t, f(t))$.

Kreis: $\gamma : [-r, r] \rightarrow \mathbb{R}^2$, $\gamma(t) = (t, \pm\sqrt{r^2 - t^2})$.

Neilsche Parabel: Offenbar ist $y(x) = \pm x^{3/2}$ ($x > 0$) oder $x(y) = |y|^{2/3}$. Dementsprechend erhält man

$$\gamma : [0, \infty) \rightarrow \mathbb{R}^2, \quad \gamma(t) := (t, \pm t^{3/2})$$

oder

$$\tilde{\gamma} : \mathbb{R} \rightarrow \mathbb{R}^2, \quad \tilde{\gamma}(t) := (|t|^{2/3}, t).$$

Eine Parameterdarstellung γ einer Kurve Γ heißt *regulär*, bzw. *regulär im Punkt* $\gamma(t)$, wenn γ' nicht verschwindet, bzw. $\gamma'(t) \neq 0$ ist. Eine Kurve Γ heißt *regulär*, wenn es eine reguläre Parameterdarstellung von Γ gibt.

Die Neilsche Parabel ist nicht regulär im Nullpunkt. Das könnte zunächst an der gewählten Parameterdarstellung liegen. Tatsächlich kann aber die Neilsche Parabel keine überall reguläre Parameterdarstellung haben, da sie im Nullpunkt keine wohldefinierte Richtung hat.

Andererseits ist die Strecke von x nach y (vgl. obiges Beispiel a) natürlich regulär, wobei die Darstellung γ regulär ist, wogegen die Darstellung $\tilde{\gamma}$ in den beiden Randpunkten nicht regulär ist. Die Darstellung

$$\hat{\gamma} : [-1, 1] \rightarrow \mathbb{R}^m, \quad \hat{\gamma}(t) := \frac{x+y}{2} - t^3 \frac{y-x}{2}$$

ist nicht regulär im Mittelpunkt.

Wenn sich zwei Kurven (Γ_1, γ_1) und (Γ_2, γ_2) in einem Punkt $\gamma_1(t_1) = \gamma_2(t_2)$ schneiden und beide Kurven in diesem Punkt regulär sind, gilt

$$\langle \gamma_1'(t_1), \gamma_2'(t_2) \rangle = \cos \text{Winkel}(\gamma_1'(t_1), \gamma_2'(t_2)) \|\gamma_1'(t_1)\| \|\gamma_2'(t_2)\|,$$

d. h. der *Schnittwinkel* ist gegeben durch

$$\text{Winkel}(\gamma_1'(t_1), \gamma_2'(t_2)) = \arccos \frac{\langle \gamma_1'(t_1), \gamma_2'(t_2) \rangle}{\|\gamma_1'(t_1)\| \|\gamma_2'(t_2)\|}.$$

9.2 Rektifizierbare Kurven, Kurvenlänge

Eine Kurve (Γ, γ) heißt *rektifizierbar*, wenn ihr auf folgendem Weg eine Länge zugeschrieben werden kann: Ist $\gamma : [a, b] \rightarrow \mathbb{R}^m$ eine Parameterdarstellung von Γ und gibt es ein C so, dass für alle Zerlegungen $\alpha = t_0 < t_1 < \dots < t_n = b$ des Intervalls $[a, b]$ gilt

$$L_\gamma(t_0, t_1, \dots, t_n) := \sum_{i=1}^n |\gamma(t_i) - \gamma(t_{i-1})| \leq C,$$

so existiert das Supremum aller $L_\gamma(t_0, \dots, t_n)$ ($n \in \mathbb{N}$). Dieses wird mit $L_\Gamma = \text{Länge der Kurve } \Gamma$ bezeichnet.

Nun ist zwar ziemlich klar, dass z. B. die Schraubenlinie oder die Neilsche Parabel keine endliche Länge haben. Ist aber das Parameterintervall beschränkt (und abgeschlossen) und γ stetig, so wird man das schon eher erwarten. Tatsächlich gilt das nicht:

Beispiel 9.2 Die Kurve Γ mit $\gamma : [0, 1] \rightarrow \mathbb{R}^2, \mapsto \left(t, t \sin \frac{1}{t}\right)$ ($\gamma(0) = 0$) ist nicht rektifizierbar. Das sieht man, wenn man für die Zerlegung von $[0, 1]$ immer mehr Punkte nimmt in denen $\sin \frac{1}{t} = \pm 1$ gilt. \square

Eine wichtige Klasse rektifizierbarer Kurven wird durch folgenden Satz beschrieben.

Satz 9.3 Eine Kurve Γ mit stetig differenzierbarer Parameterdarstellung $\gamma : [a, b] \rightarrow \mathbb{R}^m$ ist rektifizierbar und hat die Länge

$$L_\gamma = \int_a^b |\gamma'(t)| dt = \int_a^b \left(\gamma_1'(t)^2 + \gamma_2'^2 + \dots + \gamma_m'(t)^2 \right)^{1/2} dt.$$

Beweis. (Anschaulich ist das klar: Wenn man $|\gamma'(t)|$ als Geschwindigkeit betrachtet, mit der ein Punkt die Kurve durchläuft, ist der zurückgelegte Weg – also die Kurvenlänge – gleich dem Integral über $|\gamma'(\cdot)|$.) Nun etwas genauer, wie es obiger Definition entspricht:

$$|\gamma(t_i) - \gamma(t_{i-1})| = \left\{ \sum_{k=1}^m |\gamma_k(t_i) - \gamma_k(t_{i-1})|^2 \right\}^{1/2} = (t_i - t_{i-1}) \left\{ \sum_{k=1}^m \gamma_k'(\tau_{i,k})^2 \right\}^{1/2}$$

mit $\tau_{i,k}$ zwischen t_{i-1} und t_i . Also unterscheidet sich $\sum_{i=1}^n |\gamma(t_i) - \gamma(t_{i-1})|$ bei zunehmender Verfeinerung immer weniger von der Riemann-Summe⁶

$$\sum_{i=1}^n \left\{ \sum_{k=1}^m \gamma_k'(t_i)^2 \right\}^{1/2} (t_i - t_{i-1}) = \sum_{i=1}^n |\gamma'(t_i)| (t_i - t_{i-1})$$

für das Integral $\int_a^b |\gamma'(t)| dt$. ■

Bis auf die Neilsche Parabel sind die Längen der obigen Kurve leicht zu berechnen (und elementargeometrisch bekannt).

Hier wollen wir die Länge der Zykloide berechnen: Nach Beispiel 9.1 e gilt

$$\begin{aligned} \gamma'(t) &= (1 - \cos t, \sin t), \\ |\gamma'(t)|^2 &= (1 - \cos t)^2 + \sin^2 t = 1 - 2 \cos t + \cos^2 t + \sin^2 t \\ &= 2 - 2 \cos t = 2 \left(\sin^2 \frac{t}{2} + \cos^2 \frac{t}{2} \right) - 2 \cos^2 \frac{t}{2} + 2 \sin^2 \frac{t}{2} = 4 \sin^2 \frac{t}{2}, \\ |\gamma'(t)| &= 2 \left| \sin \frac{t}{2} \right| \end{aligned}$$

und damit die Länge (einer Periode) der Zykloide:

$$L = \int_0^{2\pi} 2 \left| \sin \frac{t}{2} \right| dt = 2 \int_0^{2\pi} \sin \frac{t}{2} dt = -4 \cos \frac{t}{2} \Big|_0^{2\pi} = 8.$$

(Es ist überraschend, dass hier kein Ausdruck entsteht, der π enthält, sondern eine ganze Zahl.)

Es ist noch interessant zu sehen, wie die Kurvenlänge in zwei speziellen Formen der Parameterdarstellung berechnet werden kann:

⁶Ist man $\{|t_i - t_{i-1}| : i = 1, \dots, n\} < \delta$ mit $|\gamma_k'(t) - \gamma_k'(s)| < \varepsilon$ für $|t - s| < \delta$ und $k = 1, \dots, n$, so ist

$$\left| |\gamma'(t_i)| - \left\{ \sum_{k=1}^m \gamma_k'(\tau_{i,k})^2 \right\}^{1/2} \right| \leq \left\{ \sum_{k=1}^m \left(\gamma_k'(t_i) - \gamma_k'(\tau_{i,k}) \right)^2 \right\}^{1/2} \leq \sqrt{m} \varepsilon.$$

Also unterscheiden sich die beiden Summen für eine solche Zerlegung höchstens um $(b - a)\sqrt{m}\varepsilon$.

Satz 9.4 a) Ist die Kurve als Graph der stetig differenzierbaren Funktion $g : [a, b] \rightarrow \mathbb{R}^{m-1}$ gegeben, so ist die Kurvenlänge

$$L = \int_a^b \left(1 + |g'(t)|^2\right)^{1/2} dt.$$

b) Ist eine ebene Kurve in Polarkoordinaten $r : [a, b] \rightarrow \mathbb{R}^2$, $\varphi \mapsto r(\varphi)$ mit stetig differenzierbarem $r(\cdot)$ gegeben, so ist die Kurvenlänge

$$L = \int_a^b \left(r'(\varphi)^2 + r(\varphi)^2\right)^{1/2} d\varphi.$$

Beweis. a) Hier ist $\gamma(t) = (t, g(t))$, also $|\gamma'(t)| = (1 + |g'(t)|^2)^{1/2}$, woraus die Behauptung folgt.

b) Hier ist $\gamma(\varphi) = (r(\varphi) \cos \varphi, r(\varphi) \sin \varphi)$, also

$$\begin{aligned} |\gamma'(\varphi)|^2 &= (r'(\varphi) \cos \varphi - r(\varphi) \sin \varphi)^2 + (r'(\varphi) \sin \varphi + r(\varphi) \cos \varphi)^2 \\ &= r'(\varphi)^2 + r(\varphi)^2, \end{aligned}$$

woraus auch in diesem Fall die Behauptung folgt. ■

9.3 Kurvenintegrale

Die bei einer Bewegung eines Körpers im Kraftfeld $k(\cdot)$ geleistete Arbeit ist (bei geradlinigier Bewegung) gleich der Weglänge mal der Kraftkomponente in Richtung des Weges. Durchläuft der Körper eine Kurve (Γ, γ) , so zerlegt man zunächst das Parameterintervall

$$a = t_0 < t_1 < \dots < t_n = b.$$

Die im (Zeit-)Intervall $[t_{k-1}, t_k]$ geleistete Arbeit ist näherungsweise

$$\left\langle k(\gamma(t_{k-1})), \gamma(t_k) - \gamma(t_{k-1}) \right\rangle \sim \left\langle k(\gamma(t_{k-1})), \gamma'(t_{k-1}) \right\rangle (t_k - t_{k-1}).$$

Die über das ganze Zeitintervall $[a, b]$ geleistete Arbeit ist also

$$\sum_{k=1}^n \left\langle k(\gamma(t_{k-1})), \gamma'(t_{k-1}) \right\rangle (t_k - t_{k-1}).$$

Dies ist die/eine Riemannsumme für das Integral

$$\int_{\Gamma} k(x) dx := \int_a^b \left\langle k(\gamma(t)), \gamma'(t) \right\rangle dt.$$

Aus der Vorüberlegung zu dieser Definition ist offensichtlich, dass dieses Integral nicht von der Wahl der Parameterdarstellung abhängt.

Man kann auch anders zu dieser Definition des Integrals kommen: Die *Leistung* zur Zeit t – also die Ableitung von $A(t)$ der bis zur Zeit t geleisteten Arbeit – ist gleich dem Produkt aus

der Geschwindigkeit und der Komponente des Kraftfeldes in Richtung der Geschwindigkeit, also $\langle k(\gamma(t)), \gamma'(t) \rangle$. Integration liefert wieder das gleiche Integral.

Haben wir nun allgemein in einer offenen Menge $A \subset \mathbb{R}^m$ ein stetiges Vektorfeld $k : A \rightarrow \mathbb{R}^m$ und eine Kurve Γ in A mit stetig differenzierbarer Parameterdarstellung $\gamma : [a, b] \rightarrow \mathbb{R}^m$, so definieren wir das *Integral des Vektorfeldes k längs der Kurve Γ* durch

$$\int_{\Gamma} k(x) dx := \int_a^b \langle k(\gamma(t)), \gamma'(t) \rangle dt.$$

Gibt es eine stetig differenzierbare Funktion $V : A \rightarrow \mathbb{R}$ mit $k(x) = \text{grad } V(x)$, so heißt k ein *Gradientenfeld* und V ein *Potential* von k . Der folgende Satz charakterisiert die Gradientenfelder durch ihre Kurvenintegrale:

Satz 9.5 *Sei $A \subset \mathbb{R}^m$ offen, $k : A \rightarrow \mathbb{R}^m$ ein stetiges Vektorfeld auf A . Dann ist k genau dann ein Gradientenfeld, wenn die Kurvenintegrale über k wegunabhängig sind (d. h. der Wert des Integrals hängt nur von Anfangs- und Endpunkt der Kurve ab). Es gilt dann*

$$\int_{\Gamma} k(x) dx = \int_a^b \langle k(\gamma(t)), \gamma'(t) \rangle dt = V(\gamma(b)) - V(\gamma(a)).$$

Beweis. \Rightarrow : Sei $k = \text{grad } V$. Dann gilt

$$\int_a^b \langle k(\gamma(t)), \gamma'(t) \rangle dt = \int_a^b \sum_{j=1}^m D_j V(\gamma(t)) \gamma'_j(t) dt = \int_a^b \frac{d}{dt} V(\gamma(t)) dt = V(\gamma(b)) - V(\gamma(a)).$$

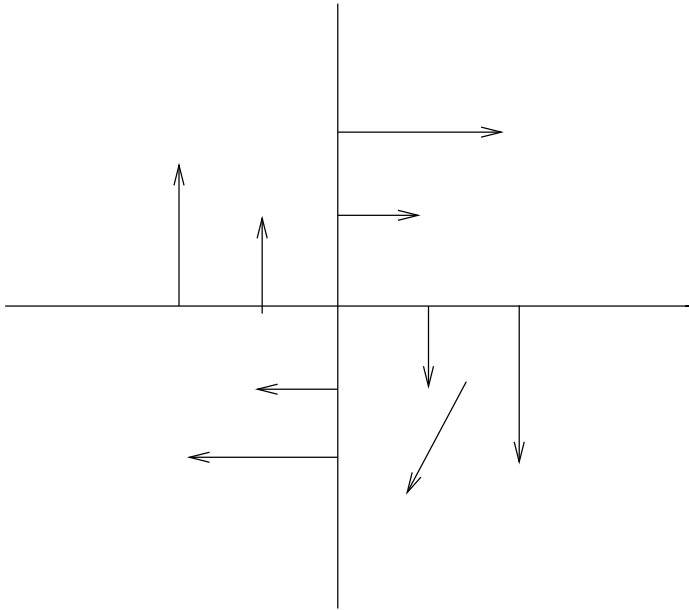
\Leftarrow : Sei das Kurvenintegral wegunabhängig. Wir konstruieren ein Potential von k . Dazu sei $x_0 \in A$ fest gewählt, Γ_x ein beliebiger Weg von x_0 nach x in A ,

$$V(x) := \int_{\Gamma_x} k(z) dz$$

(damit ist $V(\cdot)$ wohldefiniert, da das Integral wegunabhängig ist). Dann gilt

$$\begin{aligned} D_j V(x) &= \lim_{h \rightarrow 0} \frac{1}{h} (V(x + h e^j) - V(x)) = \lim_{h \rightarrow 0} \frac{1}{h} \left\{ \int_{x_0}^{x + h e^j} \dots - \int_{x_0}^x k(z) dz \right\} \\ &= \lim_{h \rightarrow 0} \frac{1}{h} \int_x^{x + h e^j} k(z) dz = \lim_{h \rightarrow 0} \frac{1}{h} \int_0^h \langle k(x + t e^j), e^j \rangle dt \\ &= \lim_{h \rightarrow 0} \frac{1}{h} \int_0^h k_j(x + t e_j) dt = k_j(x), \end{aligned}$$

d. h. V ist ein Potential von k . ■

Abbildung 4: Vektorfeld $k(x) = (\xi_2, -\xi_1)$ **Beispiel 9.6** Das Vektorfeld

$$k : \mathbb{R}^2 \rightarrow \mathbb{R}^2, \quad k(x) = (\xi_2, -\xi_1)$$

hat kein Potential. Auf Grund der Skizze

sieht man sofort ein, dass das Kurvenintegral über Kreise um Null nicht verschwindet.

Außerdem: Wäre $k = \text{grad } V = (D_1V, D_2V)$, so würde folgen (beachte, dass dieses k beliebig oft stetig differenzierbar ist)

$$\text{rot } k(x) := D_1k_2 - D_2k_1 = D_1D_2V - D_2D_1V = 0.$$

Tatsächlich gilt aber $\text{rot } k = -2$. (Die Bedeutung der *Rotation* rot werden wir später kennen lernen.)

□

Eine Schwäche des obigen Satzes liegt darin, dass man einem Vektorfeld zu nächst nicht einfach ansieht, ob es ein Gradientenfeld ist, wenn man nicht schon ein Potential kennt. Ist k ein stetig differenzierbares Gradientenfeld $k = \text{grad } V$, so gilt (da dann V zwei mal stetig differenzierbar ist) für alle $i, j \in \{1, \dots, m\}$

$$D_ik_j - D_jk_i = D_iD_jV - D_jD_iV = 0.$$

Man kann sich also fragen, ob das ausreicht um zu garantieren, dass k ein Gradientenfeld ist. Im allgemeinen Nein!

Beispiel 9.7 Sei $k : \mathbb{R}^2 \setminus \{0\}$ das Vektorfeld

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

Dann gilt zwar $D_1k_2 = D_2k_1 = 0$, aber die Integrale über Kreise um 0 verschwinden nicht, d. h. k ist kein Gradientenfeld. Der folgende Satz macht deutlich, woran das liegt. □

Satz 9.8 (Integrabilitätsbedingung) Sei $G \in \mathbb{R}^m$ offen und einfach zusammenhängend (d. h. jede geschlossene Kurve in G läßt sich in G [d. h. ohne das man dabei G verläßt] stetig auf einen Punkt zusammenziehen⁷), $k : G \rightarrow \mathbb{R}^m$ ein stetig differenzierbares Vektorfeld mit $D_i k_j - D_j k_i = 0$ für alle $i, j \in \{1, \dots, m\}$. Dann ist k ein Gradientenfeld.

Beweis. Der Einfachheit halber betrachten wir nur sternförmige Gebiete: Ein Gebiet $G \subset \mathbb{R}^m$ heißt *sternförmig* bezüglich x_0 , wenn mit jedem $x \in G$ auch die Verbindungsstrecke mit x_0 in G liegt.

Ohne Einschränkung sei G sternförmig bezüglich 0. Wir definieren $(\gamma : [0, 1] \rightarrow \mathbb{R}^m, \gamma(t) = tx)$

$$V(x) := \int_0^x k(z) dz = \int_0^1 \langle k(tx), x \rangle dt.$$

V ist damit wohldefiniert und offensichtlich (Beweis) stetig in G . Weiter gilt

$$\begin{aligned} \frac{\partial}{\partial x_j} V(x) &= \int_0^1 \frac{\partial}{\partial x_j} \langle k(tx), x \rangle dt = \int_0^1 \frac{\partial}{\partial x_j} \sum_{\ell=1}^m k_\ell(tx) x_\ell dt \\ &= \int_0^1 \left\{ k_j(tx) + \sum_{\ell=1}^m x_\ell \frac{\partial}{\partial x_j} k_\ell(tx) \right\} dt \\ &\quad \text{(hier wird die Voraussetzung benutzt)} \\ &= \int_0^1 \left\{ k_j(tx) + \sum_{\ell=1}^m x_\ell \frac{\partial}{\partial x_\ell} k_j(tx) \right\} dt = \int_0^1 \left\{ k_j(tx) + t \langle \text{grad } k_j(tx), x \rangle \right\} dt \\ &\quad \left(\text{Produktregel für } t \cdot k_j \text{ und Kettenregel für } \frac{\partial}{\partial t} k_j(tx) \right) \\ &= \int_0^1 \frac{\partial}{\partial t} (tk_j(tx)) dt = 1 \cdot k_j(1x) - 0 \cdot k_j(0x) = k_j(x) \end{aligned}$$

Also ist V ein Potential von k . ■

10 Umkehrfunktion und implizit definierte Funktion

10.1 Umkehrfunktion

Aus dem 1-dimensionalen Fall wissen wir: Ist $f : I \rightarrow \mathbb{R}$ stetig und streng monoton, so ist auch das Bild $f(I)$ ein Intervall und f besitzt eine *Umkehrfunktion* (*Inverse*) $f^{-1} : f(I) \rightarrow I$. Ist zusätzlich f differenzierbar mit $f'(x_0) \neq 0$, so ist f^{-1} an der Stelle $y_0 = f(x_0)$ differenzierbar mit

$$(f^{-1})(f(x_0)) = \frac{1}{f'(x_0)} \quad \text{bzw.} \quad (f^{-1})'(y_0) = \frac{1}{f'(f^{-1}(y_0))}.$$

⁷In \mathbb{R}^2 bedeutet das, dass im Innern der Kurve keine Punkte liegen, die nicht zu G gehören; $\mathbb{R}^2 \setminus \{0\}$ ist nicht einfach zusammenhängend; dagegen ist $\mathbb{R}^m \setminus \{0\}$ für $m \geq 3$ einfach zusammenhängend

Ist f stetig differenzierbar mit $f'(x) \neq 0$, so ist auch f^{-1} stetig differenzierbar (die strenge Monotonie folgt dann bereits aus der Forderung $f'(x) \neq 0$).

Hier soll das Resultat auf Funktionen mehrerer Variablen verallgemeinert werden. Die Monotonie läßt sich nicht übertragen. Dagegen erweist es sich als sinnvoll, die Forderung $f'(x) \neq 0$ durch $\det Df(x) \neq 0$ zu ersetzen.

Wir sehen dies zunächst an folgendem einfachen Beispiel (wobei man beachte, dass jede differenzierbare Funktion lokal durch eine affine Funktion approximiert werden kann):

Beispiel 10.1 Sei L eine $m \times m$ -Matrix, $a \in \mathbb{R}^m$, f die affine Funktion

$$f : \mathbb{R}^m \rightarrow \mathbb{R}^m, \quad f(x) = a + Lx.$$

Offensichtlich ist f in jedem $x \in \mathbb{R}^m$ differenzierbar mit Ableitung $Df(x) = L$ für alle $x \in \mathbb{R}^m$, denn es gilt

$$f(z) = a + Lz = a + Lx + L(z - x) = f(x) + L(z - x) + \varphi_x(z) \quad \text{mit } \varphi_x(z) \equiv 0.$$

Da $y = a + Lx$ genau dann für alle y eindeutig nach x auflösbar ist, wenn L invertierbar ist, $x = L^{-1}(y - a)$, ist f genau dann invertierbar, wenn L invertierbar ist, und es gilt

$$f^{-1} : \mathbb{R}^m \rightarrow \mathbb{R}^m, \quad f^{-1}(y) = L^{-1}(y - a) = b + L^{-1}y \quad \text{mit } b = -L^{-1}a.$$

Insbesondere ist auch f^{-1} differenzierbar mit Ableitung $(Df^{-1})(y) = L^{-1} = (Df(y))^{-1}$ für alle $y \in \mathbb{R}^m$. \square

Da eine differenzierbare Funktion $f : A \rightarrow \mathbb{R}^m$ (mit $A \subset \mathbb{R}^m$ offen) in der Nähe von x_0 durch die affine Funktion

$$F_{x_0} : \mathbb{R}^m \rightarrow \mathbb{R}^m, \quad F_{x_0}(x) = f(x_0) + Df(x_0)(x - x_0)$$

sehr gut approximiert wird, liegt es nahe, dass f in der Nähe von x_0 umkehrbar ist, wenn die Matrix $Df(x_0)$ invertierbar ist. Genauer: annähernd sollte für y nahe $f(x_0)$ gelten

$$f^{-1}(y) \sim F_{x_0}^{-1}(y) = Df(x_0)^{-1}(y - f(x_0)) + x_0 = f^{-1}(f(x_0)) + Df(x_0)^{-1}(y - f(x_0)).$$

Die Ableitung von f^{-1} bei $f(x_0)$ sollte also gerade $(Df(x_0))^{-1}$ sein. Da die Approximation nur lokal gut ist, ist aber zu erwarten, dass auch diese Aussage nur lokal gilt (vgl. dazu den folgenden Satz und das darauffolgende Beispiel).

Im folgenden Satz benutzen wir eine neue Sprechweise: ist $x_0 \in A \subset \mathbb{R}^m$, so heißt U eine offene Umgebung von x_0 in A , wenn U offen ist mit $x_0 \in U \subset A$. Im Fall $A = \mathbb{R}^m$ nennen wir U einfach eine offene Umgebung von x_0 .

Satz 10.2 (Umkehrfunktion) Sei $A \subset \mathbb{R}^m$ offen, $f : A \rightarrow \mathbb{R}^m$ stetig differenzierbar, $x_0 \in A$, $y_0 = f(x_0)$, $Df(x_0)$ invertierbar. Dann gibt es eine offene Umgebung U von x_0 in A und eine offene Umgebung V von y_0 so, dass

$$f : U \rightarrow V \quad \text{bijektiv}$$

ist und $Df(x)$ für alle $x \in U$ invertierbar ist. Die Abbildung

$$g := (f|_U)^{-1} : V \rightarrow U$$

ist stetig differenzierbar mit

$$Dg(f(x)) = (Df(x))^{-1} \quad \text{für } x \in U \quad \text{bzw.} \quad Dg(y) = (Df(g(y)))^{-1} \quad \text{für } y \in V.$$

Beispiel 10.3 Wie oben schon vermutet, gilt die Aussage des Satzes (insbesondere die Invertierbarkeit) nur global. Dies wird durch folgendes Beispiel belegt:

$$A = \mathbb{R}^2 \setminus \{0\} \subset \mathbb{R}^2, \quad f(\xi_1, \xi_2) = (\xi_1^2 - \xi_2^2, 2\xi_1, \xi_2).$$

(Setzt man $z = \xi_1 + i\xi_2$, so ist offenbar $f(\xi_1, \xi_2) = (\operatorname{Re} z^2, \operatorname{Im} z^2)$, d. h. bei f handelt es sich um die Darstellung der komplexen Quadrierung in \mathbb{R}^2 .)

Es gilt

$$Df(\xi_1, \xi_2) = \begin{pmatrix} 2\xi_1 & -2\xi_2 \\ 2\xi_2 & 2\xi_1 \end{pmatrix}, \quad \text{also } \det Df(\xi_1, \xi_2) = 4(\xi_1^2 + \xi_2^2) \neq 0 \text{ f\u00fcr alle } x \in A.$$

Trotzdem ist f nicht injektiv, denn es gilt f\u00fcr jedes $x = (\xi_1, \xi_2) \in A$

$$f(-\xi_1, -\xi_2) = f(\xi_1, \xi_2);$$

das ist nichts anderes als die Tatsache, dass jede komplexe Zahl $\neq 0$ zwei Quadratwurzeln hat. \square

Beweis von Satz 10.2. Da A offen ist, und f stetig differenzierbar, existiert ein $\delta > 0$ mit $U := K_\delta(x_0) \subset A$ und ⁸

$$|Df(x) - Df(x_0)| \leq \frac{1}{2} \left| Df(x_0)^{-1} \right|^{-1} =: \lambda \text{ f\u00fcr } x \in A \text{ mit } |x - x_0| < \delta.$$

Wir werden zeigen: Mit diesem U und $V := f(U)$ gilt die Behauptung des Satzes.

1. Schritt: $Df(x)$ ist invertierbar f\u00fcr $x \in U$: F\u00fcr alle $x \in U$ und $y \in \mathbb{R}^m \setminus \{0\}$ gilt (dabei benutzen wir, dass f\u00fcr eine invertierbare Matrix L gilt $|y| \leq |L^{-1}| |Ly|$ bzw. $|Ly| \geq |L^{-1}|^{-1} |y|$)

$$\begin{aligned} |Df(x)y| &\geq |Df(x_0)y| - |Df(x)y - Df(x_0)y| \\ &\geq |Df(x_0)^{-1}|^{-1} |y| - \frac{1}{2} |Df(x_0)^{-1}|^{-1} |y| \\ &= \frac{1}{2} |Df(x_0)^{-1}|^{-1} |y| > 0, \end{aligned}$$

d. h. f\u00fcr $x \in U$ ist $Df(x)$ invertierbar.

2. Schritt: f ist auf U injektiv: F\u00fcr $x \in U$ schreiben wir f in der Form

$$f(x) = Df(x_0)x + F(x) \text{ mit } F(x) := f(x) - Df(x_0)x.$$

Dann gilt f\u00fcr alle $x, x' \in U$ mit $x \neq x'$ (wieder benutzen wir $|Ly| \geq |L^{-1}|^{-1} |y|$)

$$\begin{aligned} |f(x) - f(x')| &= |Df(x_0)(x - x') + F(x) - F(x')| \\ &\geq |Df(x_0)(x - x')| - |F(x) - F(x')| \\ &\geq |Df(x_0)^{-1}|^{-1} |x - x'| - |DF(\xi)| |x - x'| \\ &= 2\lambda |x - x'| - |Df(\xi) - Df(x_0)| |x - x'| \\ &\geq 2\lambda |x - x'| - \lambda |x - x'| = \lambda |x - x'| > 0. \end{aligned}$$

⁸dabei ist f\u00fcr eine $m \times m$ -Matrix L mit $|L|$ die „Operatornorm“ von L gemeint, $|L| := \sup\{|Lx| : x \in \mathbb{R}^m, |x| \leq 1\}$. Es gilt z. B. $|L| \leq \sum_{i,j=1}^m |\ell_{ij}|$ und $|L| \leq \left\{ \sum_{i,j=1}^m |\ell_{ij}|^2 \right\}^{1/2}$; insbesondere ist $|L|$ klein, wenn alle a_{ij} klein sind.

Also ist $f|_U$ injektiv.

Damit ist jedenfalls $f : U \rightarrow V$ bijektiv.

3. Schritt: V ist offen: Dazu zeigen wir: Ist

$$y' = f(x') \in V = f(U), \quad \delta' := \delta - |x' - x_0|, \quad \varepsilon = \varepsilon(x') := \frac{\lambda}{2}\delta',$$

so ist $K_\varepsilon(y') \in V$, d. h.: zu jedem $z \in K_\varepsilon(y')$ gibt es ein $x \in U = K_\delta(x_0)$ mit $f(x) = z$.

Wegen $|z - y'| < \varepsilon = \frac{\lambda}{2}\delta'$ gibt es ein $r < \delta'$ mit $|z - y'| < \frac{\lambda}{2}r$.

Wir betrachten die Hilfsfunktion

$$\psi : \overline{K}_r(x') \rightarrow \mathbb{R}_+, \quad x \mapsto |f(x) - z|^2.$$

Gesucht ist also ein $x \in K_r(x') \subset U$ mit $f(x) = z$, d. h. mit $\psi(x) = 0$; da ψ nicht negativ ist, ist dies ein Minimum von ψ . – Da ψ auf $\overline{K}_r(x')$ stetig ist, nimmt es dort sein Minimum an; es ist zu zeigen, dass dieser minimale Wert 0 ist. Für x auf dem Rand von $\overline{K}_r(x')$, d. h. $|x - x'| = r$, gilt

$$\begin{aligned} \psi(x) &= |f(x) - z|^2 \geq \left\{ |f(x) - f(x')| - |y' - z| \right\}^2 \\ &\geq \left\{ \lambda|x - x'| - \frac{\lambda}{2}r \right\}^2 = \left\{ \lambda r - \frac{\lambda}{2}r \right\}^2 = \frac{1}{4}\lambda^2 r^2. \end{aligned}$$

Andererseits gilt im Mittelpunkt x' von $\overline{K}_r(x')$

$$\psi(x') = |f(x') - z|^2 = |y' - z|^2 < \left(\frac{1}{2}\lambda r\right)^2 = \frac{1}{4}\lambda^2 r^2.$$

Also nimmt ψ sein Minimum im Innern der Kugel $K_r(x')$ an, etwa in $\tilde{x} \in K_r(x')$. Dann muß aber gelten

$$\begin{aligned} 0 &= \text{grad } \psi(\tilde{x}) = \left(\dots, D_j \sum_{k=1}^m |f_k(\tilde{x}) - z_k|^2, \dots \right) \\ &= \left(\dots, 2 \sum_{k=1}^m D_j f_k(\tilde{x})(f_k(\tilde{x}) - z_k), \dots \right) = 2(Df(\tilde{x}))^t(f(\tilde{x}) - z). \end{aligned}$$

Da $Df(\tilde{x})$ invertierbar ist, gilt dies auch für $(Df(\tilde{x}))^t$, und somit ist $f(\tilde{x}) = z$, d. h. \tilde{x} ist das gesuchte x mit $z = f(x)$. Damit ist bewiesen, dass V offen ist.

4. Schritt: Es bleibt schließlich nur noch zu zeigen $g := (f|_U)^{-1}$ ist in $y_0 = f(x_0)$ differenzierbar mit $Dg(y_0) = Df(x_0)^{-1}$: (O. E. können wir uns auf x_0 beschränken, da jedes $x \in A$ als Ausgangspunkt gewählt werden kann.) Für $y \in \mathbb{R}^m$ mit $|y| < \varepsilon$ (ε wie oben) gilt

$$g(y_0 + y) = x_0 + x \in U = K_\delta(x_0).$$

Wegen

$$y = f(x_0 + x) - y_0 = f(x_0 + x) - f(x_0)$$

folgt aus den obigen Überlegungen

$$|y| \geq \lambda|x| \quad \text{bzw.} \quad |x| \leq \frac{1}{\lambda}|y|.$$

Außerdem gilt (da f in x_0 die Ableitung $Df(x_0)$ hat)

$$y - Df(x_0)x = f(x_0 + x) - f(x_0) - Df(x_0)x = \varphi_{f,x_0}(x_0 + x),$$

$$Df(x_0)^{-1}y - x = Df(x_0)^{-1}\varphi_{f,x_0}(x_0 + x)$$

mit

$$\frac{1}{|x|}\varphi_{f,x_0}(x_0 + x) \rightarrow 0 \text{ für } x \rightarrow 0.$$

Damit erhalten wir

$$\begin{aligned} |g(y_0 + y) - g(y_0) - Df(x_0)^{-1}y| &= |(x_0 + x) - x_0 - Df(x_0)^{-1}y| \\ &= |x - Df(x_0)^{-1}y| = |Df(x_0)^{-1}\varphi_{f,x_0}(x_0 + x)| \\ &\leq |Df(x_0)^{-1}|\varphi_{f,x_0}(x_0 + x)| \end{aligned}$$

mit

$$\frac{1}{|y|}|Df(x_0)^{-1}|\varphi_{f,x_0}(x_0 + x)| \leq \frac{1}{\lambda}|Df(x_0)^{-1}|\frac{|\varphi_{f,x_0}(x_0 + x)|}{|x|} \rightarrow 0$$

für $y \rightarrow 0$ (was mit $x \rightarrow 0$ gleichbedeutend ist). Also gilt

$$g(y_0 + y) = g(y_0) + Df(x_0)^{-1}y + \varphi_{g,y_0}(y_0 + y)$$

mit

$$\frac{1}{|y|}\varphi_{g,y_0}(y_0 + y) \rightarrow 0 \text{ für } y \rightarrow 0,$$

d. h. g ist in $y_0 = f(x_0)$ differenzierbar mit $Dg(y_0) = Df(x_0)^{-1}$. ■

22: 12. Januar 2007

Beispiel 10.4 $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$, $f(x, y) = \begin{pmatrix} x^3 + x + y \\ x^3 + x - y \end{pmatrix}$ ist überall differenzierbar mit

$$Df(x, y) = \begin{pmatrix} 3x^2 + 1 & 1 \\ 3x^2 + 1 & -1 \end{pmatrix}, \text{ speziell } \text{Det}Df(0, 0) = \text{Det} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} = -2 \neq 0.$$

Also ist f in einer Umgebung von $(0, 0)$ ($f(0, 0) = (0, 0)$) umkehrbar mit

$$Df^{-1}(0, 0) = \frac{1}{2} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}.$$

□

10.2 Implizit definierte Funktionen

Wir betrachten erst die wesentlich überschaubarere Situation in \mathbb{R}^2 . Ist $g : \mathbb{R}^2 \rightarrow \mathbb{R}$ eine „vernünftige“ (was das genau heißt sehen wir später; z. B. stetig differenzierbare) Funktion, so wird durch die Nullstellenmenge

$$M := \left\{ (x, y) \in \mathbb{R}^2 : g(x, y) = 0 \right\}$$

in vielen Fällen eine Kurve in \mathbb{R}^2 beschrieben. So ist z. B. für $g(x, y) = x^2 + y^2 - 1$ die Menge M gerade die Einheitskreislinie. Da sich Kurven in der Regel (zumindest stückweise) als Graph darstellen lassen, kommen wir zu der

Frage: Gibt es eine Menge $U \subset \mathbb{R}$ (oder vielleicht mehrere derartige Mengen) und ein $f : U \rightarrow \mathbb{R}$ so, dass der Graph von f ein Stück der Nullstellenmenge M von g beschreibt. In anderen Worten, kann die Gleichung $g(x, y) = 0$ (wenigstens lokal) nach y aufgelöst werden, $y = f(x)$ mit $g(x, f(x)) = 0$? Eine solche Funktion f heißt durch g *implizit definiert*.

Offensichtlich ist M sicher dann ein Graph einer (allerdings i. Allg. nicht notwendig stetigen und überall definierten) Funktion f , wenn die Funktion

$$g_x : y \mapsto g(x, y)$$

für jedes x streng monoton ist, denn dann gibt es für jedes x höchstens ein y mit $g_x(y) = g(x, y) = 0$.

Dies gilt aber schon für so einfache Funktionen wie die oben genannte

$$g(x, y) = x^2 + y^2 - 1$$

nicht global. Allerdings läßt sich die Einheitskreislinie stückweise als Graph von stetig differenzierbaren Kurven darstellen, nämlich

$$y = \pm\sqrt{1-x^2} \quad \text{für } -1 < x < 1 \quad \text{bzw.} \quad x = \pm\sqrt{1-y^2} \quad \text{für } -1 < y < 1.$$

Der folgende Satz ergibt hierfür z. B.

$$y' = -\frac{D_1(x^2 + y(x)^2 - 1)}{D_2(x^2 + y(x)^2 - 1)} = -\frac{2x}{2y(x)} = -\frac{x}{\pm\sqrt{1-x^2}} = \mp \frac{x}{\sqrt{1-x^2}}.$$

Bevor wir den allgemeinen Fall betrachten, formulieren und beweisen wir den entsprechenden Satz für den bisher betrachteten einfachen Fall, in dem die Situation noch geometrisch einfach erfaßbar ist.

Satz 10.5 Sei $A \subset \mathbb{R}^2$ offen, $g \in C^1(A) = C^1(A, \mathbb{R})$,

$$(x_0, y_0) \in M := \left\{ (x, y) \in A : g(x, y) = 0 \right\} \quad \text{mit } D_2g(x_0, y_0) \neq 0.$$

Dann existieren offene Intervalle $U, V \subset \mathbb{R}$ mit

$$x_0 \in U, \quad y_0 \in V, \quad U \times V \subset A$$

und eine stetige Funktion

$$f : U \rightarrow V \quad \text{mit } M \cap (U \times V) = G_f := \{(x, f(x)) : x \in U\}$$

und

$$Df(x) = f'(x) = -\frac{D_1g(x, f(x))}{D_2g(x, f(x))} \quad \text{für } x \in U.$$

(Insbesondere kann die Ableitung von f angegeben werden, ohne dass man die Funktion f genau kennt.)

Beweis. Wegen $D_2g(x_0, y_0) \neq 0$ und der Stetigkeit von D_2g existiert ein Rechteck $U' \times V$ um (x_0, y_0) mit

$$D_2g(x, y) \neq 0 \quad \text{für } (x, y) \in U' \times V,$$

d. h. $g(x, \cdot)$ ist streng monoton in V für jedes $x \in U'$. Also gibt es zu jedem $x \in U'$ höchstens ein $y \in V$ mit $g(x, y) = 0$; es gibt also ein Teilintervall $U \subset U'$ mit $x_0 \in U$ so, dass für jedes $x \in U$ genau ein $y =: f(x) \in V$ existiert mit $g(x, y) = 0$.

Wenn wir hier die Differenzierbarkeit von f als gegeben betrachten (im Beweis des folgenden Satzes ergibt sich diese für den allgemeinen Fall), dann folgt aus $h(x) := g(x, f(x)) \equiv 0$ mit $k(x) := (x, f(x))$

$$\begin{aligned} 0 &= h'(x) = Dh(x) = Dg(k(x))Dk(x) \\ &= \left\langle \left(D_1g(x, f(x)), D_2g(x, f(x)) \right), (1, f'(x)) \right\rangle \\ &= D_1g(x, f(x)) + f'(x)D_2g(x, f(x)). \end{aligned}$$

woraus die behauptete Formel für f' folgt. ■

Beispiel 10.6 Für $g(x, y) = x - y^2$ ist das Nullstellengebilde

$$\left\{ (x, y) \in \mathbb{R}^2 : x = y^2 \right\}, \quad \left\{ (x, y) \in \mathbb{R}^2 : x > 0, y = \pm\sqrt{x} \right\}.$$

Die durch $g(x, y) = 0$ implizit definierte Funktion ist also

$$(0, \infty) \rightarrow \mathbb{R}, \quad x \mapsto \sqrt{x} \quad \text{oder} \quad x \mapsto -\sqrt{x}.$$

Im Nullpunkt ist der Satz nicht anwendbar, da dort $D_2g(0, 0) = 0$ ist (dort ist aber auch die Wurzel nicht differenzierbar).

Für andere Punkte gilt

$$\frac{d}{dx} (\pm\sqrt{x}) = -\frac{D_1g(x, y(x))}{D_2g(x, y(x))} = -\frac{1}{-2y(x)} = \frac{1}{2(\pm\sqrt{x})}.$$

Wir haben damit einen anderen Weg gefunden, die Ableitung der Wurzelfunktion zu bestimmen. Dies kann natürlich entsprechend auf andere Umkehrfunktionen angewandt werden. □

Korollar 10.7 Sei $A \subset \mathbb{R}^2$ offen, $g : A \rightarrow \mathbb{R}$ stetig differenzierbar mit

$$\text{grad } g(x, y) \neq 0 \quad \text{für alle } (x, y) \in M := \{(x, y) \in A : g(x, y) = 0\}.$$

Dann ist M endliche Vereinigung von Graphen stetig differenzierbarer Funktionen f_j (zum Teil y als Funktion von x , zum Teil x als Funktion von y).

Beweis. In jedem Punkt $(x_0, y_0) \in M$ ist $D_1g(x_0, y_0) \neq 0$ oder $D_2g(x_0, y_0) \neq 0$. Deshalb gibt es ein ganzes Rechteck $U \times V$ um (x_0, y_0) mit $D_1g(x, y) \neq 0$ für alle $(x, y) \in U \times V$ oder $D_2g(x, y) \neq 0$ für alle $(x, y) \in U \times V$, also eine Funktion

$$f : U \rightarrow V \quad \text{oder} \quad \tilde{f} : V \rightarrow U$$

mit

$$M \cap (U \times V) = G_f \quad \text{oder} \quad M \cap (U \times V) = G_{\tilde{f}}^{-1},$$

(wobei G_h^{-1} aus G_h durch Vertauschen der Komponenten hervorgeht). In konkreten Fällen ist klar, dass sich M durch endlich viele solcher Rechtecke überdecken läßt; im allgemeinen Fall wird hier der Satz von Heine–Borel benötigt (vgl. z. B. H. Heuser, Lehrbuch der Analysis 2, §157). ■

Satz 10.8 (Implizite Funktionen) Sei $m, n \in \mathbb{N}$, $A \subset \mathbb{R}^{m+n}$ offen, $g \in C^1(A, \mathbb{R}^n)$. Die Punkte aus \mathbb{R}^{m+n} schreiben wir in der Form

$$(x, y) = (\xi_1, \dots, \xi_m, \eta_1, \dots, \eta_n).$$

Es sei

$$L_1(x, y) := D_x g(x, y) = (D_i g_j(x, y))_{\substack{i=1, \dots, m \\ j=1, \dots, n}} = \begin{pmatrix} D_1 g_1(x, y) & \dots & D_m g_1(x, y) \\ \vdots & \ddots & \vdots \\ D_1 g_n(x, y) & \dots & D_m g_n(x, y) \end{pmatrix},$$

$$L_2(x, y) := D_y g(x, y) = (D_{m+j} g_i(x, y))_{i,j=1, \dots, n} = \begin{pmatrix} D_{m+1} g_1(x, y) & \dots & D_{m+n} g_1(x, y) \\ \vdots & \ddots & \vdots \\ D_{m+1} g_n(x, y) & \dots & D_{m+n} g_n(x, y) \end{pmatrix}$$

und somit

$$Dg(x, y) = (L_1(x, y) : L_2(x, y)).$$

Für ein

$$(x_0, y_0) \in M := \{(x, y) \in A : g(x, y) = 0\}$$

sei die Matrix $L_2(x_0, y_0)$ invertierbar.

Dann existieren offene Mengen $U \in \mathbb{R}^m$ und $V \subset \mathbb{R}^n$ mit

$$(x_0, y_0) \in U \times V \subset A$$

und eine stetig differenzierbare Funktion

$$f : U \rightarrow V \quad \text{mit} \quad M \cap (U \times V) = G_f,$$

$$Df(x) = -L_2(x, f(x))^{-1} L_1(x, f(x)).$$

Beweisidee: In der Nähe von (x_0, y_0) gilt mit „großer Genauigkeit“

$$g(x, y) \sim g(x_0, y_0) + L_1(x_0, y_0)(x - x_0) + L_2(x_0, y_0)(y - y_0).$$

Auflösung der Gleichung $g(x, y) = 0$ nach y in der Nähe von (x_0, y_0) mit $g(x_0, y_0) = 0$ sollte also möglich sein, wenn die Gleichung

$$L_1(x_0, y_0)(x - x_0) + L_2(x_0, y_0)(y - y_0) = 0$$

nach y auflösbar ist. Das ist genau dann möglich, wenn $L_2(x_0, y_0)$ invertierbar ist. Die Lösung ist

$$f(x) = y = y_0 + L_2(x_0, y_0)^{-1} L_1(x_0, y_0)(x - x_0).$$

Als Ableitung erhalten wir, insbesondere im Punkt $x = x_0$

$$Df(x_0) = L_2(x_0, y_0)^{-1} L_1(x_0, y_0).$$

Beweis. Wir definieren die folgenden Hilfsfunktionen

$$\begin{aligned} G &: A \rightarrow \mathbb{R}^{m+n} & (x, y) &\mapsto (x, g(x, y)), \\ j_1 &: \mathbb{R}^m \rightarrow \mathbb{R}^{m+n} & x &\mapsto (x, 0) \quad (\text{Injektion}), \\ p_2 &: \mathbb{R}^{m+n} \rightarrow \mathbb{R}^n, & (x, y) &\mapsto y \quad (\text{Projektion}). \end{aligned}$$

Im folgenden sei E_m die m -dimensionale Einheitsmatrix, $0_{m,n}$ die m -zeilige und n -spaltige Nullmatrix. Damit gilt offensichtlich

$$DG(x_0, y_0) = \begin{pmatrix} E_m & 0_{mn} \\ L_1(x_0, y_0) & L_2(x_0, y_0) \end{pmatrix},$$

$$\det DG(x_0, y_0) = \det L_2(x_0, y_0) \neq 0$$

da $L_2(x_0, y_0)$ invertierbar ist. Also gibt es eine Umgebung von (x_0, y_0) in A , ohne Einschränkung in der Form $U_1 \times V$ wählbar mit offenen Umgebungen U_1 und V von x_0 bzw. y_0 , so, dass $G : U_1 \times V \rightarrow \mathbb{R}^{m+n}$ eine stetig differenzierbare Inverse $(G|_{U_1 \times V})^{-1}$ hat; wegen $G(x_0, y_0) = (x_0, g(x_0, y_0)) = (x_0, 0)$ ist $(G|_{U \times V})^{-1}$ jedenfalls auf $U \times \{0\}$ definiert für eine geeignete Umgebung $U \subset U_1$ von x_0 in \mathbb{R}^m .

Wir zeigen, dass

$$f := p_2 \circ G^{-1} \circ j_1 : U \rightarrow V$$

die gewünschte Funktion ist:

- j_1 bildet U auf $U \times \{0\}$ ab, $x \mapsto (x, 0)$,
- G^{-1} bildet $U \times \{0\}$ in $U \times V$ ab, $(x, 0) \mapsto (x, y)$ mit $g(x, y) = 0$,
- p_2 bildet $U \times V$ nach V ab, $(x, y) \mapsto y$, wobei für die hier in Frage stehenden (x, y) das y die Eigenschaft hat, dass $g(x, y) = 0$ ist.

Insgesamt bildet also f jedes $x \in U$ auf das $y \in V$ ab, für das $g(x, y) = 0$ gilt; wir haben also die gesuchte Funktion von f auf einer Umgebung U von x_0 „explizit“ dargestellt.

Nun bleibt nur noch die Ableitung zu berechnen:

$$Df(x) = Dp_2(G^{-1}(j_1(x))) DG^{-1}(j_1(x)) Dj_1(x).$$

Der erste und der letzte Faktor sind einfach zu berechnen,

$$Dp_2(\cdot) = \begin{pmatrix} 0_{n,m} & E_n \end{pmatrix}, \quad Dj_1 = \begin{pmatrix} E_m \\ 0_{n,m} \end{pmatrix}.$$

Den mittleren Faktor erhalten wir aus dem Satz über die Umkehrfunktion (Satz 10.2). $DG(x, y)$ haben wir bereits oben angegeben (wobei für (x_0, y_0) jetzt (x, y) zu setzen ist). Man sieht leicht, dass gilt

$$(DG(x, y))^{-1} = \begin{pmatrix} E_m & 0_{mn} \\ L_1(x, y) & L_2(x, y) \end{pmatrix}^{-1} = \begin{pmatrix} E_m & 0_{m,n} \\ -L_2^{-1}L_1 & L_2^{-1} \end{pmatrix} \quad \text{mit} \quad L_j := L_j(x, y),$$

also, mit der Formel für die Ableitung der Umkehrfunktion,

$$\begin{aligned} DG^{-1}(j_1(x)) &= DG^{-1}(x, 0) = \left(DG(G^{-1}(x, 0)) \right)^{-1} \\ &= \left(DG(x, f(x)) \right)^{-1} = \begin{pmatrix} E_m & 0_{m,n} \\ -M_2^{-1}M_1 & M_2^{-1} \end{pmatrix} \quad \text{mit} \quad M_j := L_j(x, f(x)), \end{aligned}$$

und damit erhalten wir

$$\begin{aligned} Df(x) &= \begin{pmatrix} 0_{n,m} & E_n \end{pmatrix} \begin{pmatrix} E_m & 0_{m,n} \\ -M_2^{-1}M_1 & M_2^{-1} \end{pmatrix} \begin{pmatrix} E_m \\ 0_{n,m} \end{pmatrix} \\ &= \begin{pmatrix} 0_{n,m} & E_n \end{pmatrix} \begin{pmatrix} E_m \\ -M_2^{-1}M_1 \end{pmatrix} = -M_2^{-1}M_1 \\ &= -L_2(x, f(x))^{-1}L_1(x, f(x)). \end{aligned}$$

Das ist die behauptete Formel. ■

Völlig analog zu Korollar ?? gilt:

Korollar 10.9 Sei $A \in \mathbb{R}^m$ offen, $g : A \rightarrow \mathbb{R}$ stetig differenzierbar mit

$$\text{grad } g(x) \neq 0 \quad \text{für alle } x \in M := \{x \in A : g(x) = 0\}.$$

Dann ist M endliche Vereinigung von Graphen stetiger Funktionen f_j (jeweils eine Koordinate als Funktion der anderen $m-1$).

10.3 Bedingte Extrema

Häufig will man Extrema von Funktionen auf Flächen oder Kurven bestimmen, *bedingte Extrema*. Dies ist auch wichtig, wenn man Extrema auf Teilgebieten von \mathbb{R}^m einschließlich ihres Randes bestimmen will: Mit der früher beschriebenen Methode findet man lokale Extrema im Innern. Dies muß man dann noch vergleichen mit den lokalen Extrema auf dem Rand.

Satz 10.10 (Methode der Lagrange-Multiplikatoren) Sei $A \subset \mathbb{R}^m$ offen,

$$\begin{aligned} g_j : A \rightarrow \mathbb{R} \text{ stetig differenzierbar, } M_j := \{x \in A : g_j(x) = 0\} \text{ für } j = 1, \dots, k, \\ M := \bigcap_{j=1}^k M_j, \quad x_0 \in M \text{ und } \{\text{grad } g_j(x_0) : j = 1, \dots, k\} \text{ linear unabhängig.} \end{aligned}$$

Ist $f : A \rightarrow \mathbb{R}$ stetig differenzierbar und hat f in x_0 ein bedingtes lokales Extremum auf M , so gilt

$$\text{grad } f(x_0) \in L\{\text{grad } g_j(x_0) : j = 1, \dots, k\},$$

d. h. es gibt reelle Zahlen $\lambda_1, \dots, \lambda_k$ (Lagrange-Multiplikatoren) mit

$$\text{grad } f(x_0) = \sum_{j=1}^k \lambda_j \text{grad } g_j(x_0).$$

23: 16. Januar 2007

Beweis. Ist $\gamma : [a, b] \rightarrow M$ eine (in x_0) differenzierbare Kurve in M durch x_0 , $\gamma(c) = x_0$ für ein $c \in (a, b)$, so hat die Funktion

$$h : [a, b] \rightarrow \mathbb{R}, \quad h(t) = f(\gamma(t))$$

in c ein lokales Extremum. Da h in c differenzierbar ist, gilt

$$0 = h'(c) = \langle \text{grad } f(\gamma(c)), \gamma'(c) \rangle,$$

d. h. $\text{grad } f(x_0)$ steht in x_0 senkrecht auf jeder Kurve in M durch x_0 .

Da $\gamma'(c)$ jede Richtung haben kann, die orthogonal zu $\{\text{grad } g_j(x_0) : j = 1, \dots, k\}$ ist (hier wird die lineare Unabhängigkeit von $\text{grad } g_j(x_0) : j = 1, \dots, k\}$ benutzt; dann hat der Tangentialraum von M in x_0 die Dimension $m - k = \dim\{\text{grad } g_j(x_0) : j = 1, \dots, k\}^\perp$), folgt aus

$$\bigcap_{j=1, \dots, k} \{\text{grad } g_j(x_0)\}^\perp = \left(L\{\text{grad } g_j(x_0) : j = 1, \dots, k\}\right)^\perp,$$

dass $\text{grad } f(x_0)$ eine Linearkombination der $\text{grad } g_j(x_0)$ sein muß. ■

Bemerkung 10.11 Die Anwendung dieses Satzes geht so: Man löst das Gleichungssystem

$$\begin{aligned} g_j(x) &= 0 \quad (j = 1, \dots, k) \\ \text{grad } f(x) - \sum_{j=1}^k \lambda_j \text{grad } g_j(x) &= 0 \end{aligned}$$

$k + m$ Gleichungen für die $k + m$ Unbekannten $\xi_1, \dots, \xi_m, \lambda_1, \dots, \lambda_k$ auf, wobei die λ_j eigentlich nicht interessieren. Dann überprüft man anhand der Funktionswerte in den gefundenen Punkten x , wo das Maximum bzw. Minimum liegt.

Will man das Maximum bzw. Minimum auf einem berandeten Gebiet bestimmen, so sucht man mit der früheren Methode die Extrema im Inneren, mit dieser Methode die Extrema auf dem Rand, und vergleicht diese miteinander.

Beispiel 10.12 $A = \mathbb{R}^2$, $g(x) = \xi_1^2 + \xi_2^2 - 1$, also

$$M := \{x \in \mathbb{R}^2 : \xi_1^2 + \xi_2^2 = 1\} \quad \text{Einheitskreislinie.}$$

Es sollen die Extrema der Funktion

$$f(x) = \xi_1 + 2\xi_2$$

auf M bestimmt werden. Man hat also das Gleichungssystem

$$\xi_1^2 + \xi_2^2 - 1 = 0 \tag{1}$$

$$(1, 2) - \lambda(2\xi_1, 2\xi_2) = 0 (= \text{grad } f(x) + \lambda \text{grad } g(x)) \tag{2}$$

zu lösen. Aus (2) folgt $\lambda \neq 0$ und

$$\xi_1 = \frac{1}{2\lambda}, \quad \xi_2 = \frac{1}{\lambda} \tag{3}$$

Einsetzen in (1) liefert

$$\frac{1}{4\lambda^2} + \frac{1}{\lambda^2} - 1 = 0, \quad \lambda^2 = \frac{5}{4}, \quad \lambda = \pm \frac{1}{2}\sqrt{5}.$$

Einsetzen in (3) liefert

$$\xi_1 = \pm \frac{1}{\sqrt{5}}, \quad \xi_2 = \pm \frac{2}{\sqrt{5}}.$$

Die Funktionswerte in diesen Punkten sind

$$f\left(\pm \frac{1}{\sqrt{5}}, \pm \frac{2}{\sqrt{5}}\right) = \pm\sqrt{5} \sim \pm 2, 236 \dots$$

Also liegt in $\left(\frac{1}{\sqrt{5}}, \frac{2}{\sqrt{5}}\right)$ das Maximum, in $\left(-\frac{1}{\sqrt{5}}, -\frac{2}{\sqrt{5}}\right)$ das Minimum auf M vor.

Es gibt mindestens zwei weitere Möglichkeiten, dieses Problem zu lösen:

- Man stellt den Kreis in Parameterdarstellung dar und führt damit das Problem auf ein 1-dimensionales Problem zurück.
- Man stellt fest, dass f auf den Geraden $\xi_2 = -\frac{1}{2}\xi_1 + c$ konstant ist (Wert = c) und bestimmt die c , für die diese Geraden Tangenten sind.

□

11 Berechnung m -dimensionaler Integrale

11.1 Elementare Regeln

Ist I ein „Intervall“ (Quader) in \mathbb{R}^m ,

$$I = \{x \in \mathbb{R}^m : a_j < \xi_j < b_j\} \quad \text{mit} \quad -\infty \leq a_j < b_j \leq \infty$$

und ist $f : I \rightarrow \mathbb{R}$ (oder \mathbb{C}) Lebesgue-integrierbar, so kann nach Fubini das Integral über I iterativ berechnet werden:

$$\int_I f(x) \, dx = \int_{a_m}^{b_m} \dots \int_{a_2}^{b_2} \left\{ \int_{a_1}^{b_1} f(\xi_1, \dots, \xi_m) \, d\xi_1 \right\} d\xi_2 \dots d\xi_m,$$

wobei die Reihenfolge der Integrale beliebig vertauscht werden kann. (Man beachte, dass Lebesgue-Integrierbarkeit stets eine absolute Integrierbarkeit ist, d. h. auch $|f|$ ist integrierbar. Existieren nur die uneigentlichen Integrale

$$\int_I f(x) \, dx := \lim_{n \rightarrow \infty} \int_{I_n} f(x) \, dx \quad (\text{mit } I_n \nearrow I),$$

so gilt für „ $\int \dots$ “ die entsprechende Formel im Allgemeinen nicht.)

Satz 11.1 Ist $A \subset \mathbb{R}^{p+q}$ Lebesgue-messbar und $f : A \rightarrow \mathbb{R}$ Lebesgue-integrierbar, so gilt

$$\int_A f(x) \, dx = \int_{A'} \left\{ \int_{B_y} f(y, z) \, dz \right\} dy$$

mit

$$A' := \{y \in \mathbb{R}^p : \exists z \in \mathbb{R}^q \text{ mit } (y, z) \in A\},$$

$$B_y := \{z \in \mathbb{R}^q : (y, z) \in A \text{ für } y \in A'\}.$$

Beweis. Nach dem Satz von Fubini gilt

$$\begin{aligned} \int_A f(x) \, dx &= \int \chi_A(x) f(x) \, dx = \int \left\{ \int \chi_A(y, z) f(y, z) \, dz \right\} dy \\ &= \int \left\{ \int \chi_{B_y}(z) f(y, z) \, dz \right\} \chi_{A'}(y) \, dy = \int_A \left\{ \int_{B_y} f(y, z) \, dz \right\} dy. \end{aligned}$$

■

Speziell in \mathbb{R}^2 kann man A häufig in der Form

$$A = \left\{ x = (\xi_1, \xi_2) \in \mathbb{R}^2 : a_1 \leq \xi_1 \leq b_1, \varphi(\xi_1) \leq \xi_2 \leq \psi(\xi_1) \right\}$$

darstellen, oder als Vereinigung solcher Mengen (wobei eventuell die Rollen von ξ_1 und ξ_2 vertauscht sind). Dann gilt im gleichen Sinn wie im Satz formuliert

$$\int_A f(x) \, dx = \int_{a_1}^{b_1} \int_{\varphi(\xi_1)}^{\psi(\xi_1)} f(\xi_1, \xi_2) \, d\xi_2 \, d\xi_1.$$

Entsprechende Überlegungen sind in \mathbb{R}^m möglich, wenn A z. B. die Form

$$A = \left\{ x \in \mathbb{R}^m : (\xi_1, \dots, \xi_{m-1}) \in A_{m-1} \subset \mathbb{R}^{m-1}, \right. \\ \left. \varphi_m(\xi_1, \dots, \xi_{m-1}) \leq \xi_m \leq \psi_m(\xi_1, \dots, \xi_{m-1}) \right\}$$

hat, wobei u. U. A_{m-1} entsprechend dargestellt werden kann.

Beispiel 11.2 Sei A die rechte Hälfte des Einheitskreises, $f(x) = \xi_1$. Wir wollen $\int_A f(x) dx$ berechnen. Offensichtlich ist

$$A = \left\{ x \in \mathbb{R}^2 : 0 \leq \xi_1 \leq 1, -\sqrt{1-\xi_1^2} \leq \xi_2 \leq \sqrt{1-\xi_1^2} \right\}$$

also

$$\int_A f(x) dx = \int_0^1 \int_{-\sqrt{1-\xi_1^2}}^{\sqrt{1-\xi_1^2}} \xi_1 d\xi_2 d\xi_1 = \int_0^1 \xi_1 \int_{-\sqrt{1-\xi_1^2}}^{\sqrt{1-\xi_1^2}} d\xi_2 d\xi_1 \\ = \int_0^1 2\xi_1 \sqrt{1-\xi_1^2} d\xi_1 = -\frac{2}{3} (1-\xi_1^2)^{3/2} \Big|_0^1 = \frac{2}{3}.$$

Da offenbar $\int_A \xi_2 dx = 0$ gilt, ist also der Schwerpunkt dieses Halbkreises bei $\xi_2 = 0$ und $\xi_1 = 2/3$ dividiert durch die Fläche (=Gesamtmasse) $\pi/2$, also $\xi_2 = 0$ und $\xi_1 = 4/3\pi \approx 0.42$. \square

Beispiel 11.3 Wir betrachten in \mathbb{R}^m die Kugel um 0 mit Radius r :

$$K_{m,r} = \left\{ x \in \mathbb{R}^m : |x| \leq r \right\} = \left\{ x \in \mathbb{R}^m : \xi_1^2 + \dots + \xi_{m-1}^2 + \xi_m^2 \leq r^2 \right\} \\ = \left\{ x \in \mathbb{R}^m : \xi_1^2 + \dots + \xi_{m-1}^2 \leq r^2, \right. \\ \left. -(r^2 - \xi_1^2 - \dots - \xi_{m-1}^2)^{1/2} \leq \xi_m \leq (r^2 - \xi_1^2 - \dots - \xi_{m-1}^2)^{1/2} \right\} \\ \vdots \\ = \left\{ x \in \mathbb{R}^m : -r \leq \xi_1 \leq r, -(r^2 - \xi_1^2)^{1/2} \leq \xi_2 \leq (r^2 - \xi_1^2)^{1/2}, \dots, \right. \\ \left. -(r^2 - \xi_1^2 - \dots - \xi_{m-1}^2)^{1/2} \leq \xi_m \leq (r^2 - \xi_1^2 - \dots - \xi_{m-1}^2)^{1/2} \right\}.$$

Damit kann also z. B. das Integral über die Einheitskugel $K_{3,r}$ in \mathbb{R}^3 geschrieben werden in der Form

$$\int_{K_3} f(x) dx = \int_{-r}^r \int_{-\sqrt{r^2-\xi_1^2}}^{\sqrt{r^2-\xi_1^2}} \int_{-\sqrt{r^2-\xi_1^2-\xi_2^2}}^{\sqrt{r^2-\xi_1^2-\xi_2^2}} f(\xi_1, \xi_2, \xi_3) d\xi_3 d\xi_2 d\xi_1.$$

\square

11.2 Transformationsformel (Substitutionsregel)

Für die explizite Berechnung von Integralen ist (analog zur 1-dimensionalen Substitutionsregel) häufig die folgende *Transformationsformel (Substitutionsregel)* nützlich:

Satz 11.4 (Transformationsformel) Seien $A, B \subset \mathbb{R}^m$ Lebesgue-messbar, $\varphi : A \rightarrow B$ stetig differenzierbar und surjektiv (evtl. bis auf eine Lebesgue-Nullmenge), es gebe eine Lebesgue-Nullmenge $N \subset A$ so, dass φ in $A' = A \setminus N$ eine stetig differenzierbare Inverse hat. Eine Funktion $f : B \rightarrow \mathbb{R}$ ist genau dann Lebesgue-integrierbar, wenn $f \circ \varphi : A \rightarrow \mathbb{R}$ Lebesgue-integrierbar ist, und es gilt

$$\int_B f(x) dx = \int_A f(\varphi(y)) \left| \det D\varphi(y) \right| dy.$$

Vorbemerkung zum Beweis. Da φ stetig differenzierbar ist, ist $\varphi(N)$ wieder eine Lebesgue-Nullmenge (Beweis!). Wir können also in A auf die Menge N und in B auf die Menge $\varphi(N)$ verzichten, d. h. es genügt, die obige Gleichung mit A' statt A und $B' = B \setminus \varphi(N)$ statt B zu beweisen; es ist dann φ stetig differenzierbar mit stetig differenzierbarer Inversen. — Zum Beweis benötigen wir die folgenden Hilfssätze.

Hilfssatz 11.5 Ist L eine invertierbare reelle $m \times m$ -Matrix, so gibt es orthogonale $m \times m$ -Matrizen U_1 und U_2 und eine Diagonalmatrix $D = \text{diag}(d_1, \dots, d_m)$ mit $d_j > 0$ für $j = 1, \dots, m$ so, dass gilt $L = U_1 D U_2$.

Beweis. $L^* L$ ist symmetrisch und invertierbar mit positiven Eigenwerten d_1^2, \dots, d_m^2 . Also gibt es eine orthogonale Matrix U so, dass gilt:

$$L^* L = U^{-1} D^2 U, \quad U L^* L U^{-1} = D^2 \quad \text{mit } D = \text{diag}(d_1, \dots, d_m).$$

Mit $U_1 := L U^{-1} D^{-1}$ und $U_2 := U$ gilt dann

$$U_1^* U_1 = (D^{-1} U L^*) (L U^{-1} D^{-1}) = D^{-1} D^2 D^{-1} = I,$$

d. h. U_1 und U_2 sind orthogonal, und mit

$$U_1 D U_2 = (L U^{-1} D^{-1}) D U = L (U^{-1} D^{-1} D U) = L$$

folgt die Behauptung. ⁹ ■

24: 19. Januar 2007

Hilfssatz 11.6 Für jede Lebesgue-messbare Menge $M \subset \mathbb{R}^m$ gilt:

- a) Für jedes $a \in \mathbb{R}^m$ ist auch $M + a = \{x + a : x \in M\}$ Lebesgue-messbar mit $|M + a| = |M|$; das Lebesgue-Maß ist translationsinvariant.
- b) Ist $D = \text{diag}(d_1, \dots, d_m)$, so ist auch $DM = \{Dx : x \in M\}$ Lebesgue-messbar mit $|DM| = |M| \prod_{j=1}^m |d_j| = |M| \det(D)$.

⁹Alternativ zu diesem Beweis kann man $|L| := U^{-1} D U$ definieren. Es gilt dann $|L|^2 = L^* L$ und $\left| |L|x \right| = \langle |L|x, |L|x \rangle = \langle |L|^2 x, x \rangle = \langle L^* L x, x \rangle = |Lx|^2$. Also wird durch $V : |L|x \mapsto Lx$ eine unitäre Abbildung V definiert, mit der gilt $L = V|L| = V U^{-1} D U$.

- c) Ist U eine orthogonale Matrix, so ist auch $UM = \{Ux : x \in M\}$ Lebesgue-messbar mit $|UM| = |M|$.
- d) Ist φ eine invertierbare affine Abbildung, $\varphi(x) = a + Lx$, mit einer (invertierbaren) Matrix L , so ist auch $\varphi(M) = \{\varphi(x) : x \in M\}$ Lebesgue-messbar mit $|\varphi(M)| = |M| |\det L| = |M| |\det D\varphi|$. (Das gilt natürlich auch, wenn L nicht invertierbar ist: Dann ist $\det L = 0$ und da LM in einem echten Teilraum von \mathbb{R}^m liegt, ist auch $|LM| = 0$.)

Beweis. a) Offensichtlich ist $|Q+a| = |Q|$ für jeden Quader Q . Damit folgt die Translationsinvarianz des Lebesgue-Integrals; mit χ_M ist also auch χ_{M+a} integrierbar mit gleichem Integral, also $|M+a| = |M|$.

b) Die Aussage gilt für Quader, also für Integrale charakteristischer Funktionen von Quadern und überträgt sich damit auf Integrale und somit auf das Maß beliebiger Lebesgue-messbarer Mengen.

c) Das orthogonale Bild jedes Quaders ist offenbar Lebesgue-messbar; es handelt sich um einen verdrehten Quader. Sei Q_0 irgendein fester Quader (z. B. der Einheitswürfel) und es gelte $|UQ_0| = \alpha|Q_0|$. Wegen Teil a und Teil b gilt dann für jeden Quader in \mathbb{R}^m und somit für jede Lebesgue-messbare Menge M (insbesondere auch für Kugeln): $|UM| = \alpha|M|$. Es bleibt zu zeigen, dass $\alpha = 1$ gilt. dies erkennt man aber sofort bei Anwendung auf eine Kugel um 0, die unter U auf sich abgebildet wird.

d) Dies ergibt sich durch Zusammensetzen der Teile a, b und c und mit obigem Hilfssatz. ■

Hilfssatz 11.7 Sei $A \subset \mathbb{R}^m$ eine offene Menge, $Q \subset A$ ein Quader in \mathbb{R}^m , und $\varphi : A \rightarrow \mathbb{R}^m$ stetig differenzierbar mit stetig differenzierbarer Inverser. Dann ist $\varphi(Q)$ Lebesgue-messbar mit

$$|\varphi(Q)| = \int_Q |\det D\varphi(x)| dx.$$

Beweis. $\varphi(Q)$ ist Lebesgue-messbar, denn der Rand von $\varphi(Q)$ besteht aus Hyperflächenstücken, d. h. die charakteristische Funktion von $\varphi(Q)$ ist f.ü. stetig..

Nehmen wir an, dass $|\varphi(Q)| > \int_Q |\det D\varphi(x)| dx$ gilt. Dann gibt es ein $\varepsilon > 0$ mit

$$|\varphi(Q)| \geq \int_Q |\det D\varphi(x)| dx + \varepsilon|Q|.$$

Durch Halbieren der Kantenlängen zerlegen wir Q in 2^m Quader. Unter diesen muß einer sein, wir bezeichnen ihn mit Q^1 , für den gilt

$$|\varphi(Q^1)| \geq \int_{Q^1} |\det D\varphi(x)| dx + \varepsilon|Q^1|.$$

Iterieren wir dieses Verfahren, so erhalten wir eine Folge (Q^k) von ineinander geschachtelten Quadern, deren Kantenlänge das 2^{-k} -fache der Kantenlänge von Q sind, mit

$$\frac{|\varphi(Q^k)|}{|Q^k|} \geq \frac{1}{|Q^k|} \int_{Q^k} |\det D\varphi(x)| dx + \varepsilon$$

für $k \in \mathbb{N}$. Da die Q^k einen gemeinsamen Punkt x_0 enthalten, o. E. können wir annehmen, dass dieser gleich 0 ist, gilt mit $L := D\varphi(0)$

$$\liminf_{k \rightarrow \infty} \frac{|\varphi(Q^k)|}{|Q^k|} \geq \lim_{k \rightarrow \infty} \frac{1}{|Q^k|} \int_{Q^k} |\det D\varphi(x)| dx + \varepsilon = |\det L| + \varepsilon. \quad (*)$$

Andererseits folgt aus der Differenzierbarkeit von φ

$$|\varphi(x) - \varphi(0) - Lx| \leq |x|\delta(|x|) \quad \text{mit } \delta(r) \rightarrow 0 \text{ für } r \rightarrow 0 \quad (**)$$

(o. E. können wir annehmen, dass $\delta(\cdot)$ wachsend ist, also $\delta(r) \searrow 0$ für $r \searrow 0$).

Sei Q_0 der Quader mit Zentrum 0, der durch entsprechende Verschiebung von Q entsteht, $\tilde{Q}_0 := LQ_0$, also nach Hilfssatz 11.6

$$|\tilde{Q}_0| = |\det L| |Q_0| = |\det L| |Q|, \quad |Q| = \frac{|\tilde{Q}_0|}{|\det L|}.$$

Dann gibt es eine Folge (x_k) mit $x_k \in 2^{-k}Q_0$ so, dass

$$Q^k = x_k + 2^{-k}Q_0$$

gilt. Wegen $0 \in Q^k$ gilt, mit $d =$ Durchmesser von Q ,

$$|x| \leq 2^{-k}d, \quad \text{also } |x|\delta(x) \leq 2^{-k}d\delta(2^{-k}d) \quad \text{für alle } x \in Q^k.$$

Mit $A_\varepsilon := \{x \in \mathbb{R}^m : d(x, A) \leq \varepsilon\}$ folgt also aus (**)

$$\begin{aligned} \varphi(Q^k) &= \varphi(x_k + 2^{-k}Q_0) \subset \varphi(0) + Lx_k + (2^{-k}\tilde{Q}_0)_{2^{-k}d\delta(2^{-k}d)} \\ &= \varphi(0) + Lx_k + 2^{-k}(\tilde{Q}_0)_{d\delta(2^{-k}d)}, \end{aligned}$$

also

$$|\varphi(Q^k)| \leq 2^{-km}|(\tilde{Q}_0)_{d\delta(2^{-k}d)}|$$

und, wegen $\delta(r) \rightarrow 0$ für $r \rightarrow 0$,

$$\frac{|\varphi(Q^k)|}{|Q^k|} \leq \frac{2^{-km}|(\tilde{Q}_0)_{d\delta(2^{-k}d)}|}{2^{-km}|Q|} = |\det L| \frac{|(\tilde{Q}_0)_{d\delta(2^{-k}d)}|}{|\tilde{Q}_0|} \rightarrow |\det L| \quad \text{für } k \rightarrow \infty.$$

Das ist ein Widerspruch zu (*).

Entsprechend zeigt man, dass $|\varphi(Q)| < \int_Q |\det D\varphi(x)| dx$ nicht gelten kann. Hier wird benutzt

$$\varphi(Q^k) \supset \varphi(0) + Lx_k + (2^{-k}\tilde{Q}_0)_{-\varepsilon_k}, \quad |\varphi(Q^k)| \geq 2^{-km}|(\tilde{Q}_0)_{-d\delta(2^{-k}d)}|$$

mit $M_{-\varepsilon} := \{x \in \mathbb{R}^m : d(x, \mathbb{R}^m \setminus M) \geq \varepsilon\}$. ■

Beweis von Satz 11.4. Aus Hilfssatz 11.7 folgt die Behauptung für Treppenfunktionen auf B , also $f = \sum_k f_k \chi_{Q_k}$ mit Quader Q_k in B :

$$\begin{aligned} \int_B f(s) dx &= \sum_k f_k |Q_k| = \sum_k f_k \int_{\varphi^{-1}(Q_k)} |\det D\varphi(x)| dx \\ &= \sum_k \int_{\varphi^{-1}(Q_k)} f(\varphi(x)) |\det D\varphi(x)| dx = \int_{\varphi^{-1}(B)} f(\varphi(x)) |\det D\varphi(x)| dx. \end{aligned}$$

Daraus ergibt sich die Behauptung für beliebige Lebesgue-integrierbare Funktionen. Insbesondere folgt die Integrierbarkeit von $f(\varphi(\cdot))|\det D\varphi(\cdot)|$ über A aus der Integrierbarkeit von f über B .

Aus der Integrierbarkeit von $f \circ \varphi|\det D\varphi|$ über A folgt umgekehrt die Integrierbarkeit von f über B wegen $f(x) = f(\varphi(\varphi^{-1}(x)))|\det D\varphi(\varphi^{-1}(x))||\det D\varphi^{-1}(x)|$. ■

11.3 Polarkoordinaten

Wir wenden dieses Resultat nun auf Polarkoordinaten in \mathbb{R}^2 und \mathbb{R}^3 an:

Beispiel 11.8 *Polarkoordinaten* in \mathbb{R}^2 : Die Abbildung

$$\varphi : [0, \infty) \times [0, 2\pi) \rightarrow \mathbb{R}^2, \varphi(r, \phi) = (r \cos \phi, r \sin \phi)$$

ist stetig differenzierbar mit

$$\begin{aligned} D\varphi(r, \phi) &= \begin{pmatrix} \cos \phi & -r \sin \phi \\ \sin \phi & r \cos \phi \end{pmatrix}, \\ \det D\varphi(r, \phi) &= r \cos^2 \phi + r \sin^2 \phi = r. \end{aligned}$$

φ ist offenbar surjektiv und mit der Nullmenge

$$N = \left\{ (0, \phi) : 0 \leq \phi < 2\pi \right\}$$

ist φ auf $(0, \infty) \times [0, 2\pi) = [0, \infty) \times [0, 2\pi) \setminus N$ injektiv.

Zum Beispiel bildet φ die Menge

$$A = \left\{ (r, \phi) : 0 \leq r \leq R, 0 \leq \phi < 2\pi \right\}$$

auf die Kreisscheibe mit Radius R

$$B = K_{2,R} = \left\{ x \in \mathbb{R}^2 : |x| \leq R \right\}$$

ab. Also ist

$$\begin{aligned} \int_{K_{2,R}} f(x) \, dx &= \int_A f(r \cos \phi, r \sin \phi) r \, d(r, \phi) \\ &= \int_0^R r \int_0^{2\pi} f(r \cos \phi, r \sin \phi) \, d\phi \, dr. \end{aligned}$$

Ist f nur von r abhängig, $f(x) = \tilde{f}(|x|)$ (rotationssymmetrisch), so folgt

$$= \int_0^R r 2\pi \tilde{f}(r) \, dr = 2\pi \int_0^R r \tilde{f}(r) \, dr.$$

Speziell ergibt sich damit

$$\int_{K_{2,R}} 1 \, dx = 2\pi \int_0^R r \, dr = R^2\pi,$$

die Fläche der Kreisscheibe mit Radius r

□

Beispiel 11.9 *Zylinderkoordinaten* in \mathbb{R}^3 : Die Abbildung φ

$$\varphi : [0, \infty) \times [0, 2\pi) \times \mathbb{R} \rightarrow \mathbb{R}^3, \quad \varphi(r, \phi, z) = (r \cos \phi, r \sin \phi, z)$$

ist stetig differenzierbar mit

$$D\varphi(r, \phi, z) = \begin{pmatrix} \cos \phi & -r \sin \phi & 0 \\ \sin \phi & r \cos \phi & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$\text{Det } d\varphi(r, \phi, z) = r \cos^2 \phi + r \sin^2 \phi = r.$$

Das Integral über Zylinder $Z_{h,R}$ von 0 bis h um die z -Achse mit Radius R hat damit die Form

$$\int_{Z_{h,R}} f(x) \, dx = \int_0^h \int_0^R \int_0^{2\pi} r f(r \cos \phi, r \sin \phi, z) \, d\phi \, dr \, dz.$$

Diese Koordinaten sind offenbar dann besonders geeignet, wenn f nur von r und z abhängt. \square

Beispiel 11.10 *Polarkoordinaten* in \mathbb{R}^3 : Wir betrachten die Abbildung

$$\varphi : [0, \infty) \times [0, 2\pi) \times [0, \pi] \rightarrow \mathbb{R}^3,$$

$$\varphi(r, \phi, \vartheta) = (r \cos \phi \sin \vartheta, r \sin \phi \sin \vartheta, r \cos \vartheta).$$

Offenbar ist φ stetig differenzierbar mit

$$D\varphi(r, \phi, \vartheta) = \begin{pmatrix} \cos \phi \sin \vartheta & -r \sin \phi \sin \vartheta & r \cos \phi \cos \vartheta \\ \sin \phi \sin \vartheta & r \cos \phi \sin \vartheta & r \sin \phi \cos \vartheta \\ \cos \vartheta & 0 & -r \sin \vartheta \end{pmatrix},$$

und durch Entwicklung nach der letzten Zeile

$$\begin{aligned} \det D\varphi(r, \phi, \vartheta) &= \cos \vartheta (-r^2 \sin^2 \phi \sin \vartheta \cos \vartheta - r^2 \cos^2 \phi \sin \vartheta \cos \vartheta) \\ &\quad - r \sin \vartheta (r \cos^2 \phi \sin^2 \vartheta + r \sin^2 \phi \sin^2 \vartheta) \\ &= -r^2 \sin \vartheta \cos^2 \vartheta - r^2 \sin \vartheta \sin^2 \vartheta = -r^2 \sin \vartheta, \\ |\det d\varphi(r, \phi, \vartheta)| &= r^2 \sin \vartheta. \end{aligned}$$

φ ist offenbar surjektiv, und mit der Nullmenge

$$\begin{aligned} N &= \left\{ (0, \phi, \vartheta) : 0 \leq \phi < 2\pi, 0 \leq \vartheta \leq \pi \right\} \\ &\quad \cup \left\{ (r, \phi, 0) : 0 < r < \infty, 0 \leq \phi < 2\pi \right\} \\ &\quad \cup \left\{ (r, \phi, \pi) : 0 < r < \infty, 0 \leq \phi < 2\pi \right\} \end{aligned}$$

ist φ auf $[0, \infty) \times [0, 2\pi) \times [0, \pi] \setminus N$ injektiv.

Zum Beispiel bildet ϕ die Menge

$$A = \left\{ (r, \phi, \vartheta) : 0 \leq r \leq R, 0 \leq \phi < 2\pi, 0 \leq \vartheta \leq \pi \right\}$$

auf die Kugel

$$B = K_{3,R} = \left\{ x \in \mathbb{R}^3 : |x| \leq R \right\}$$

ab. Es ist also

$$\begin{aligned} & \int_{K_{3,R}} f(x) \, dx \\ &= \int_A f(r \cos \phi \sin \vartheta, r \sin \phi \sin \vartheta, r \cos \vartheta) r^2 \sin \vartheta \, d(r, \phi, \vartheta) \\ &= \int_0^R r^2 \int_0^{2\pi} \int_0^\pi f(r \cos \phi \sin \vartheta, r \sin \phi \sin \vartheta, r \cos \vartheta) \sin \vartheta \, d\vartheta \, d\phi \, dr. \end{aligned}$$

Ist f nur von r abhängig, $f(x) = \tilde{f}(|x|)$ (sphärisch symmetrisch), so folgt

$$= \int_0^R r^2 \tilde{f}(r) 2\pi \int_0^\pi \sin \vartheta \, d\vartheta \, dr = 4\pi \int_0^R r^2 \tilde{f}(r) \, dr,$$

Speziell ergibt sich

$$\int_{K_{3,R}} 1 \, dx = 4\pi \int_0^R r^2 \, dr = \frac{4}{3}\pi R^3,$$

das Volumen der Kugel mit Radius R . □

Beispiel 11.11 Wir wollen das Integral über einen Kegel $K_{r,h}$ der Höhe h über einem Kreis mit Radius r berechnen:

$$\begin{aligned} K_{r,h} &= \left\{ x \in \mathbb{R}^3 : 0 \leq \xi_3 \leq h, \sqrt{\xi_1^2 + \xi_2^2} \leq r - \xi_3 \frac{r}{h} \right\} \\ &= \left\{ x \in \mathbb{R}^3 : 0 \leq \xi_3 \leq h, 0 \leq |\xi_2| \leq r \left(1 - \frac{\xi_3}{h}\right), \right. \\ &\quad \left. 0 \leq |\xi_1| \leq \sqrt{(1 - \xi_3/h)^2 r^2 - \xi_2^2} =: \varrho(\xi_2, \xi_3) \right\}. \end{aligned}$$

Es ist also

$$\begin{aligned} \int_{K_{r,h}} f(x) \, dx &= \int_0^h \int_{-r(1-\xi_3/h)}^{r(1-\xi_3/h)} \int_{-\varrho(\xi_2, \xi_3)}^{\varrho(\xi_2, \xi_3)} f(\xi_1, \xi_2, \xi_3) \, d\xi_1 \, d\xi_2 \, d\xi_3 \\ &= \int_0^h \int_{\{|\xi_1, \xi_2| \leq r(1-\xi_3/h)\}} f(\xi_1, \xi_2, \xi_3) \, d(\xi_1, \xi_2) \, d\xi_3. \end{aligned}$$

Ist f nur von $\varrho = |(\xi_1, \xi_2)|$ abhängig, $f(x) = \tilde{f}(|(\xi_1, \xi_2)|)$, so folgt (vgl. Beispiel 11.8)

$$= \int_0^h \int_{K_{2,r(1-\xi_3/h)}} \tilde{f}(|(\xi_1, \xi_2)|) \, d(\xi_1, \xi_2) \, d\xi_3 = \int_0^h 2\pi \int_0^{r(1-\xi_3/h)} \tilde{f}(\varrho) \, d\varrho \, d\xi_3.$$

Ist f nur von ξ_3 abhängig, $f(x) = \hat{f}(\xi_3)$, so folgt

$$= \int_0^h \hat{f}(\xi_3) \int_{K_{2,r(1-\xi_3/h)}} d(\xi_1, \xi_2) \, d\xi_3 = \int_0^h \hat{f}(\xi_3) \pi r^2 \left(1 - \frac{\xi_3}{h}\right)^2 \, d\xi_3.$$

Speziell erhalten wir für das Volumen des Kegels $K_{r,h}$

$$\int_{K_{r,h}} 1 \, dx = \int_0^h \pi r^2 \left(1 - \frac{\xi_3}{h}\right)^2 d\xi_3 = -\frac{1}{3} \pi r^2 h \left(1 - \frac{\xi_3}{h}\right)^3 \Big|_0^h = \frac{1}{3} \pi r^2 h.$$

Wie bei den vorhergehenden Beispielen können wir auch eine Parameterdarstellung des Kegels verwenden:

$$\begin{aligned} \varphi &: [0, 2\pi) \times [0, r] \times [0, h] \rightarrow \mathbb{R}^3, \\ \varphi(\phi, \varrho, \xi) &= \left(\varrho\left(1 - \frac{\xi}{h}\right) \cos \phi, \varrho\left(1 - \frac{\xi}{h}\right) \sin \phi, \xi\right). \end{aligned}$$

Es gilt

$$D\varphi(\phi, \varrho\xi) = \begin{pmatrix} -\varrho\left(1 - \frac{\xi}{h}\right) \sin \phi & \left(1 - \frac{\xi}{h}\right) \cos \phi & 0 \\ \varrho\left(1 - \frac{\xi}{h}\right) \cos \phi & \left(1 - \frac{\xi}{h}\right) \sin \phi & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

$$\det D\varphi(\phi, \varrho\xi) = -\varrho\left(1 - \frac{\xi}{h}\right)^2 \sin^2 \phi - \varrho\left(1 - \frac{\xi}{h}\right)^2 \cos^2 \phi = -\varrho\left(1 - \frac{\xi}{h}\right)^2,$$

$$\int_{K_{2,r}} f(x) \, dx = \int_0^{2\pi} \int_0^r \int_0^h f\left(\varrho\left(1 - \frac{\xi}{h}\right) \cos \phi, \varrho\left(1 - \frac{\xi}{h}\right) \sin \phi, \xi\right) \varrho\left(1 - \frac{\xi}{h}\right)^2 d(\phi, \varrho, \xi).$$

Ist speziell f nur von ξ abhängig, $f(x) = \hat{f}(\xi)$, so gilt

$$\begin{aligned} &= \int_0^h \hat{f}(\xi) \int_0^{2\pi} \int_0^r \varrho\left(1 - \frac{\xi}{h}\right) d\varrho d\phi d\xi \\ &= \pi r^2 \int_0^h \hat{f}(\xi) \left(1 - \frac{\xi}{h}\right)^2 d\xi, \\ \text{Volumen} &= -\frac{1}{3} \pi r^2 h \left(1 - \frac{\xi}{h}\right)^3 \Big|_0^h = \frac{1}{3} \pi r^2 h, \end{aligned}$$

wie wir dies bereits oben erhalten hatten. Die Berechnung des Integrals über eine Funktion, die nur von $\varrho = |(\xi_1, \xi_2)|$ abhängt, ist in dieser Darstellung nicht so einfach. \square

25: 23. Januar

11.4 Zwei wichtige Integrale

Es ist häufig nützlich über die Existenz oder Nicht-Existenz der Integrale der folgenden beiden Beispiele Bescheid zu wissen (im Rahmen des Riemannintegrals sind dies *uneigentliche* Integrale).

Beispiel 11.12 In \mathbb{R}^m ($m > 1$) untersuchen wir das Integral

$$\int_{K_R} \frac{1}{|x|^\alpha} \, dx \quad (\alpha \in \mathbb{R}).$$

Mit $A_n := K_R \setminus K_{1/n}$ für $n \rightarrow \infty$ oder allgemein $A_\varepsilon = K_R \setminus K_\varepsilon$ für $\varepsilon \rightarrow 0$ gilt in jedem Fall (gleichgültig ob das Integral endlich oder unendlich ist)

$$\int_{K_R} \frac{1}{|x|^\alpha} \, dx = \lim_{\varepsilon \rightarrow 0} \int_{A_\varepsilon} \frac{1}{|x|^\alpha} \, dx,$$

also

$$\begin{aligned} \int_{K_R} \frac{1}{|x|^\alpha} dx &= \lim_{\varepsilon \rightarrow 0} \int_{K_R \setminus K_\varepsilon} \frac{1}{|x|^\alpha} dx = \lim_{\varepsilon \rightarrow 0} F_m \int_\varepsilon^R r^{m-1-\alpha} dr \\ &= \lim_{\varepsilon \rightarrow 0} F_m \begin{cases} \frac{r^{m-\alpha}}{m-\alpha} \Big|_\varepsilon^R & \text{für } m-\alpha \neq 0, \\ \ln r \Big|_\varepsilon^R & \text{für } m-\alpha = 0, \end{cases} \end{aligned}$$

wobei F_m die Oberfläche der Sphäre mit Radius 1 in \mathbb{R}^m ist. Der Limes existiert genau dann, wenn $m-\alpha > 0$ ist; dann gilt

$$\int_{K_R} \frac{1}{|x|^\alpha} dx = \frac{F_m}{m-\alpha} R^{m-\alpha} \quad \text{für } \alpha < m.$$

Das Integral existiert genau dann (bzw. ist genau dann endlich), wenn $\alpha < m$ ist. \square

Beispiel 11.13 In \mathbb{R}^m ($m > 1$) untersuchen wir das Integral

$$\int_{\mathbb{R}^m \setminus K_R} \frac{1}{|x|^\alpha} dx \quad (\alpha \in \mathbb{R}).$$

Für A_n wählen wir $K_n \setminus K_R$ oder allgemeiner $A_s = K_s \setminus K_R$ für $s \rightarrow \infty$, d. h. wir haben zu untersuchen

$$\begin{aligned} \lim_{s \rightarrow \infty} \int_{K_s \setminus K_R} \frac{1}{|x|^\alpha} dx &= \lim_{s \rightarrow \infty} F_m \int_R^s r^{m-1-\alpha} ds \\ &= \lim_{s \rightarrow \infty} F_m \begin{cases} \frac{r^{m-\alpha}}{m-\alpha} \Big|_R^s & \text{für } m-\alpha \neq 0, \\ \ln r \Big|_R^s & \text{für } m-\alpha = 0. \end{cases} \end{aligned}$$

Dieser Limes existiert genau dann, wenn $m-\alpha < 0$ ist; dann gilt

$$\int_{\mathbb{R}^m \setminus K_R} \frac{1}{|x|^\alpha} dx = \frac{F_m}{\alpha-m} R^{m-\alpha} \quad \text{für } \alpha > m.$$

Das Integral existiert genau dann (bzw. ist genau dann endlich), wenn $\alpha > m$ ist. \square

Insbesondere folgt aus den beiden Beispielen, dass $\int_{\mathbb{R}^m} \frac{1}{|x|^\alpha} dx$ für kein $\alpha \in \mathbb{R}$ existiert.

11.5 Parameterabhängige Integrale und Zerlegung der Eins

Wir diskutieren die Frage der Differenzierbarkeit bzw. der Differentiation parameterabhängiger Integrale nach dem Parameter. Der Einfachheit halber werden wir nur Integrale bezüglich einer Variablen betrachten.

Satz 11.14 (Differentiation parameterabhängiger Integrale) Sei $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ stetig und besitze eine stetige partielle Ableitung nach der zweiten Variablen, $g, h : \mathbb{R} \rightarrow \mathbb{R}$ seien stetig differenzierbar. Dann gilt:

$$(a) \quad \frac{d}{dy} \int_a^b f(x, y) dx = \int_a^b D_2 f(x, y) dx,$$

$$(b) \quad \frac{d}{dy} \int_a^{h(y)} f(x, y) dx = h'(y)f(h(y), y) + \int_a^{h(y)} D_2 f(x, y) dx,$$

$$(c) \quad \frac{d}{dy} \int_{g(y)}^b f(x, y) dx = -g'(y)f(g(y), y) + \int_{g(y)}^b D_2 f(x, y) dx,$$

$$(d) \quad \frac{d}{dy} \int_{g(y)}^{h(y)} f(x, y) dx = h'(y)f(h(y), y) - g'(y)f(g(y), y) + \int_{g(y)}^{h(y)} D_2 f(x, y) dx.$$

Beweis. Zunächst gilt offensichtlich nach dem Hauptsatz der Differenzial- und Integralrechnung

$$\frac{d}{dz} \int_a^z f(x, y) dx = f(z, y) \tag{i}$$

und, durch explizite Berechnung der Differenzialquotienten

$$\begin{aligned} \frac{d}{dy} \int_a^z f(x, y) dx &= \lim_{h \rightarrow 0} \frac{1}{h} \int_a^z (f(x, y+h) - f(x, y)) dx \\ &= \lim_{h \rightarrow 0} \frac{1}{h} \int_a^z \int_y^{y+h} D_2 f(x, t) dt dx \\ &= \lim_{h \rightarrow 0} \frac{1}{h} \int_y^{y+h} \int_a^z D_2 f(x, t) dx dt = \int_a^z D_2 f(x, y) dx. \end{aligned} \tag{ii}$$

(a) Dies folgt aus (ii) mit $z = b$.

(b) Mit $k(y) := (h(y), y)$, $\ell(z, w) := \int_a^z f(x, w) dx$ folgt

$$\begin{aligned} \frac{d}{dy} \int_a^{h(y)} f(x, y) dx &= \frac{d}{dy} \ell(h(y), y) = \frac{d}{dy} \ell(k(y)) = D\ell(k(y)) Dk(y) \\ &= \left\langle \left(D_1 \ell(h(y), y), D_2 \ell(h(y), y) \right), \left(k_1'(y), k_2'(y) \right) \right\rangle \\ &\quad \text{(mit (i) für } D_1 \ell \text{ und (ii) für } D_2 \ell) \\ &= \left\langle \left(f(h(y), y), \int_a^{h(y)} D_2 f(x, y) dx \right), \left(h'(y), 1 \right) \right\rangle \\ &= h'(y)f(h(y), y) + \int_a^{h(y)} D_2 f(x, y) dx. \end{aligned}$$

(c) Wird analog zu (b) bewiesen.

(d) Mit der Zerlegung

$$\int_{g(y)}^{h(y)} f(x, y) \, dx = \int_{g(y)}^c f(x, y) \, dx + \int_c^{h(y)} f(x, y) \, dx$$

ergibt sich die Aussage (d) aus (b) und (c). ■

Satz 11.15 (Zerlegung der Eins) Sei $K \subset \mathbb{R}^m$ kompakt, $\{O_1, \dots, O_n\}$ eine Überdeckung von K aus (endlich vielen) offenen Mengen. Dann gibt es beliebig oft stetig differenzierbare Funktionen $\varphi_1, \dots, \varphi_n (\in C_0^\infty(\mathbb{R}^m))$ mit

$$\sum_{j=1}^n \varphi_j(x) = 1 \quad \text{für } x \in K, \quad \varphi_j(x) = 0 \quad \text{für } x \notin O_j.$$

(Das entsprechende gilt für eine unendliche, jedoch lokal endliche, Überdeckung.)

Beweis. Für jedes $\delta > 0$ und $j = 1, \dots, n$ sei

$$O_j^\delta := \left\{ x \in O_j : |x - y| > \delta \quad \text{für alle } y \in \mathbb{R}^m \setminus O_j \right\}$$

(das ist die Menge der Punkte von O_j , die vom Rand von O_j einen Abstand $> \delta$ haben). Offenbar ist O_j^δ offen.

Es gibt ein $\varepsilon > 0$ so, dass $K \subset O^\varepsilon := \bigcup_{j=1}^n O_j^{2\varepsilon}$ gilt. Dies folgt leicht daraus, dass $\{O^\varepsilon : \varepsilon > 0\}$ eine offene Überdeckung von K ist; es gibt also eine endliche Teilüberdeckung $\{O^{\varepsilon_1}, \dots, O^{\varepsilon_k}\}$; mit $\varepsilon := \min\{\varepsilon_j : j = 1, \dots, k\}$ folgt die Behauptung.

Sei nun für jedes $\varepsilon > 0$

$$\varphi_\varepsilon(x) := \begin{cases} c_\varepsilon \exp\left(-\frac{1}{\varepsilon^2 - x^2}\right) & \text{für } |x| < \varepsilon, \\ 0 & \text{für } |x| \geq \varepsilon \end{cases}$$

und c_ε so, dass $\int \varphi_\varepsilon(x) \, dx = 1$ gilt. Für

$$\begin{aligned} \chi_j &= \text{charakteristische Funktion von } O_j^\varepsilon \quad (j = 1, \dots, n), \\ \chi_0 &= \text{charakteristische Funktion von } \mathbb{R}^m \setminus \bigcup_{j=1}^n O_j^\varepsilon \end{aligned}$$

definieren wir

$$\tilde{\varphi}_j(x) := \int \varphi_\varepsilon(x - y) \chi_j(y) \, dy \quad (j = 0, 1, \dots, n)$$

(die *Faltung* von φ_ε mit χ_j). Dann gilt (vgl. Satz 11.14 a)

$$\begin{aligned} \tilde{\varphi}_j \in C_0^\infty(\mathbb{R}^m) \quad \text{mit} \quad \tilde{\varphi}_j(x) &= \begin{cases} 1 & \text{in } O_j^{2\varepsilon}, \\ 0 & \text{außerhalb } O_j \end{cases} \quad \text{für } j = 1, \dots, n, \\ \tilde{\varphi}_0 \in C^\infty(\mathbb{R}^m) \quad \text{mit} \quad \tilde{\varphi}_0(x) &= \begin{cases} 1 & \text{in } \mathbb{R}^m \setminus \bigcup_{j=1}^n O_j, \\ 0 & \text{in } O^\varepsilon = \bigcup_{j=1}^n O_j^{2\varepsilon}. \end{cases} \end{aligned}$$

Also ist $\sum_{k=0}^n \tilde{\varphi}_k(x) \neq 0$ für alle $x \in \mathbb{R}^m$. Definieren wir

$$\varphi_j(x) := \left\{ \sum_{k=0}^n \tilde{\varphi}_k(x) \right\}^{-1} \tilde{\varphi}_j(x) \quad \text{für } j = 1, \dots, n,$$

so folgt die Behauptung des Satzes. ■

11.6 Übungsaufgaben

11.1 Sei P ein reelles Polynom vom Grad $r > 0$ in m Variablen. Für jede kompakte Teilmenge $K \subset \mathbb{R}^m$ und jedes $\lambda \in \mathbb{R}$ ist

$$M := \{x \in K : P(x) = \lambda\}$$

eine Lebesgue-Nullmenge.

Anleitung: O. E. $\lambda = 0$, Induktion nach r , Induktionsverankerung $r = 1$. Beim Induktionsschritt ist (z. B.) zu zeigen:

- a) $N_0 = \{x \in K : P(x) = 0, \text{grad } P(x) = 0\}$ ist eine Jordan-Nullmenge.
- b) Für jedes $\varepsilon > 0$ gibt es ein $\delta > 0$ so, dass

$$N_\varepsilon^- := \{x \in K : P(x) = 0, |\text{grad } P(x)| \leq \delta\}$$

in der Vereinigung K_ε endlich vieler Intervalle mit Gesamtvolumen $\leq \varepsilon$ enthalten ist.

- c) $N_\varepsilon^+ := M \cap (K \setminus K_\varepsilon)$ ist eine Vereinigung endlich vieler Hyperflächen.

11.2 Der Schwerpunkt eines homogenen Körpers B in \mathbb{R}^3 ist gegeben durch

$$\frac{1}{V_B} \left(\int_B x \, d(x, y, z), \int_B y \, d(x, y, z), \int_B z \, d(x, y, z) \right),$$

wobei V_B das Volumen des Körpers B ist. Man berechne den Schwerpunkt der oberen Hälfte der Einheitskugel.

11.3 Man bestimme den Schwerpunkt des Kegels über dem Halbkreis $\{(\xi_1, \xi_2) : \xi_2 \geq 0, \xi_1^2 + \xi_2^2 \leq r\}$ mit der Spitze in $(0, 0, h)$.

11.4 Man berechne das Volumen und den Schwerpunkt der oberen Hälfte des Ellipsoids

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1 \quad (a, b, c > 0)$$

mit Hilfe der Transformation

$$x = ar \sin \vartheta \cos \varphi, \quad y = br \sin \vartheta \sin \varphi, \quad z = cr \cos \vartheta.$$

11.5 Man berechne

$$(a) \quad A = \int_{\mathbb{R}^2} \exp(-|x|^2) dx, \quad (b) \quad B = \int_{-\infty}^{\infty} \exp(-t^2) dt.$$

Anleitung: Man zeige $A = B^2$.

11.6 Man berechne

$$\int_0^{\pi} \left\{ \int_x^{\pi} \frac{\sin y}{y} dy \right\} dx.$$

11.7 1. Man skizziere die Archimedische Spirale

$$r(\varphi) = \varphi \quad \text{für } 0 \leq \varphi \leq 2\pi$$

(und mache insbesondere das Verhalten der Kurve im Nullpunkt deutlich).

2. Man berechne die Fläche, die durch die Archimedische Spirale und das Intervall $[0, 2\pi]$ auf der x -Achse begrenzt wird. *Anleitung:* Man kann das Flächenstück als Bild des Dreiecks $((0, 0), (2\pi, 0), (2\pi, 2\pi))$ der (φ, r) -Ebene darstellen.

11.8 Für welche $\alpha, \beta \in \mathbb{R}$, $\beta \leq 1$, existiert das Integral

$$\int_A |x|^{-\alpha} dx \quad \text{mit } A = \left\{ x \in \mathbb{R}^m : \xi_1 \geq 1, \xi_2^2 + \dots + \xi_m^2 \leq \xi_1^{2\beta} \right\}?$$

11.9 Ist $f : \mathbb{R}^m \rightarrow \mathbb{R}$ über jede kompakte Teilmenge von $\mathbb{R}^m \setminus \{0\}$ Lebesgue-integrierbar, und gilt

a) $|f(x)| \leq |x|^{-\alpha}$ mit $\alpha < m$, so ist f integrierbar über K_R .

b) $|f| \leq |x|^{-\alpha}$ mit $\alpha > m$, so ist f über $\mathbb{R}^m \setminus K_R$ integrierbar.

11.10 Man berechne den Schwerpunkt des Kegels über dem Halbkreis $\{(\xi_1, \xi_2) \in \mathbb{R}^2 : \xi_2 > 0, \xi_1^2 + \xi_2^2 \leq r\}$ mit Spitze in $(0, 0, h)$.

11.11 Für $m \geq 3$ sei

$$\varphi_m : [0, \infty) \times [0, 2\pi) \times \underbrace{[0, \pi] \times \dots \times [0, \pi]}_{m-2 \text{ mal}} \rightarrow \mathbb{R}$$

$$\varphi_m(r, \varphi, \vartheta_1, \dots, \vartheta_{m-2}) = (r \cos \varphi \sin \vartheta_1 \dots \sin \vartheta_{m-2}, r \sin \varphi \sin \vartheta_1 \dots \sin \vartheta_{m-2}, r \cos \vartheta_1 \sin \vartheta_2 \dots \sin \vartheta_{m-2}, \dots, r \cos \vartheta_{m-3} \sin \vartheta_{m-2}, r \cos \vartheta_{m-2})$$

(m -dimensionale Polarkoordinaten).

- a) φ_m ist surjektiv; die Einschränkung auf $(0, \infty) \times [0, 2\pi) \times (0, \pi) \times \dots \times (0, \pi)$ ist injektiv.
- b) Für $m = 4$ berechne man die Funktionaldeterminante $\det D\varphi_4(r, \varphi, \vartheta_1, \vartheta_2)$.
- c) Mit dieser Parameterdarstellung berechne man das Volumen der Kugel mit Radius R in \mathbb{R}^4 .

11.12 Das *Trägheitsmoment* eines Körpers K mit Dichte $\varrho : K \rightarrow \mathbb{R}$ in bezug auf eine Drehachse g ist $T_g := \int_K \varrho(x) d(x, g) dx$, wobei $d(x, g)$ der Abstand des Punktes x von der Achse g ist.

Man berechne das Trägheitsmoment einer homogenen Kugel mit Dichte ϱ und Radius r um eine Achse durch den Mittelpunkt.

11.13 Zwei gerade Kreiszylinder mit gleichem Radius R liegen so, dass ihre Achsen sich senkrecht schneiden. Man bestimme das Volumen der innerhalb beider Zylinder liegenden Menge.

12 Oberflächenintegrale; der allgemeine Satz von Stokes

Im folgenden definieren wir Integrale über Hyperflächen (insbesondere Kurven in \mathbb{R}^2 und Flächen in \mathbb{R}^3) und beweisen den grundlegenden Satz über den Zusammenhang zwischen Volumen- und Oberflächenintegral (im Spezialfall \mathbb{R}^1 ist dies der Hauptsatz der Differenzial- und Integralrechnung, der eine Beziehung zwischen dem Integral der Funktion f über ein Intervall und den Randwerten einer Stammfunktion F von f herstellt, bzw. zwischen den Randwerten einer Funktion f und dem Integral der Ableitung f' über das Intervall).

12.1 Oberflächenintegrale

Wir nehmen im folgenden zunächst an, dass eine *Hyperfläche* M als Graph einer stetig differenzierbaren Funktion h dargestellt wird:

$$M = \left\{ x = (y, z) : y \in A, z = h(y) \right\} \quad \text{mit } A \subset \mathbb{R}^{m-1}, h : A \rightarrow \mathbb{R},$$

wobei es natürlich keine Rolle spielt, dass hier $z(\cdot)$ für die letzte Komponente steht.

Das Integral einer (zunächst stetigen) Funktion $f : M \rightarrow \mathbb{R}$ über die Fläche M denken wir uns (zur Motivation für eine richtige Definition) definiert als Grenzwert von Summen über Terme der Form

$$\begin{aligned} & (\text{Fläche eines kleinen Flächenstücks in } M) \times \\ & (\text{Funktionswert von } f \text{ in einem Punkt dieses Flächenstücks}). \end{aligned}$$

Es dürfte bequemer sein, dieses „Integral“ als ein Integral über die „ebene“ Fläche A , die Projektion von M auf \mathbb{R}^{m-1} , zu definieren; dabei sind natürlich die entsprechenden Flächenstücke mit einem Faktor, der die Neigung des Flächenstücks auf M gegenüber A berücksichtigt, zu multiplizieren: dieser Faktor ist nach Pythagoras $(1 + |\text{grad } h(y)|^2)^{1/2}$. Damit kommen wir zu der Definition

$$\int_M f(x) \, d\sigma(x) := \int_A f(y, h(y)) \left(1 + |\text{grad } h(y)|^2\right)^{1/2} dy;$$

dabei soll „ $d\sigma(x)$ “ andeuten, dass es sich um ein Oberflächenintegral handelt.

Läßt sich M nicht als ganzes als Graph darstellen, so gehen wir zumindest davon aus, dass man M in kleinere Teile zerlegen kann, die sich als Graphen darstellen lassen, und definieren das Integral stückweise. (Nach dem Satz über implizite Funktionen (Satz ??), in Verbindung mit dem Satz von Heine–Borel, ist dies z. B. dann der Fall, wenn M mit Hilfe einer stetig differenzierbaren Funktion $g : \mathbb{R}^m \rightarrow \mathbb{R}$ beschrieben wird durch

$$M := \left\{ x \in \mathbb{R}^m : g(x) = 0 \right\} \quad \text{mit } \text{grad } g(x) \neq 0 \text{ für } x \in M.$$

Wie kann man nun das Oberflächenintegral berechnen, wenn M in einer beliebigen Parameterdarstellung gegeben ist? Dabei wird sich dann auch ergeben, dass das Integral nicht von der Darstellung von M abhängt.

Wir betrachten zunächst die durch die Darstellung als Graph erzeugte spezielle Parameterdarstellung von M

$$\varphi : \mathbb{R}^{m-1} \supset A \rightarrow \mathbb{R}^m, \quad \varphi(y) = (\varphi_1(y), \dots, \varphi_m(y)) := (y, h(y)) = (\eta_1, \dots, \eta_{m-1}, h(y)).$$

Mit diesem φ und mit

$$\varphi^{(j)} := (\varphi_1, \dots, \hat{\varphi}_j, \dots, \varphi_m) \quad (\text{das ist } (\varphi_1, \dots, \varphi_m) \text{ ohne } j\text{-te Komponente}),$$

gilt (einfache Rechnung!)

$$D\varphi(y) = \begin{pmatrix} D_1\varphi_1(y) & \dots & D_{m-1}\varphi_1(y) \\ \vdots & & \vdots \\ D_1\varphi_m(y) & \dots & D_{m-1}\varphi_m(y) \end{pmatrix} = \begin{pmatrix} 1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & 1 \\ D_1h(y) & \dots & D_{m-1}h(y) \end{pmatrix},$$

$$1 + |\text{grad } h(y)|^2 = \sum_{j=1}^m |\det D\varphi^{(j)}(y)|^2 =: G_\varphi(y),$$

also

$$\int_M f(x) \, d\sigma(x) = \int_A f(\varphi(y)) G_\varphi(y)^{1/2} \, dy;$$

26.: 26. Januar 2007

$G_\varphi(\cdot)$ heißt die *Gramsche Determinante* von $D\varphi$. Diese Bezeichnung geht darauf zurück, dass gilt:

$$G_\varphi(y) = \text{Det} \left(\left(\langle D_i\varphi(y), D_j\varphi(y) \rangle \right)_{i,j=1,\dots,m} \right);$$

das ist die Gramsche Determinante der m Vektoren $\{D_1\varphi(y), \dots, D_m\varphi(y)\}$ vgl. B.Walter: Analysis II, § 8.8. Die folgenden Überlegungen könnten entsprechend mit dieser Darstellung von $G_\varphi(\cdot)$ durchgeführt werden. (Man nennt übrigens die Parameterdarstellung φ *regulär* in y , wenn $G_\varphi(y) \neq 0$ ist; dies wird zunächst keine Rolle spielen.)

Wir verwenden die obige Definition von G_φ und dem Integral für beliebige stetig differenzierbare Parameterdarstellungen und zeigen nun die Unabhängigkeit des von der Wahl der Parameterdarstellung. Da jede Darstellung als Graph als eine spezielle Parameterdarstellung betrachtet werden kann, ergibt sich damit gleichzeitig die Unabhängigkeit der Integrale von der Wahl der Darstellung als Graph.

Satz 12.1 *Sei M wie oben gegeben. Ist $\psi : B \rightarrow M$ ($B \subset \mathbb{R}^{m-1}$) eine beliebige stetig differenzierbare Parameterdarstellung von M , so gilt*

$$\int_M f(x) \, d\sigma(x) = \int_B f(\psi(z)) G_\psi(z)^{1/2} \, dz.$$

Die angegebenen Formeln für das Oberflächenintegral liefern also für alle Darstellungen als Graph bzw. für alle Parameterdarstellungen den gleichen Wert.

Beweis. Wie eben gezeigt wurde, gilt (mit φ aus der obigen Definition von \int_M)

$$\int_M f(x) \, d\sigma(x) = \int_A f(\varphi(y)) G_\varphi(y)^{1/2} \, dy$$

(Substitution $k = \varphi^{-1} \circ \psi : B \rightarrow A$; auf Grund der speziellen Struktur von φ gilt offenbar $k(y) = (\psi_1(y), \dots, \psi_{m-1}(y))$, d. h. k ist eine stetig

$$\begin{aligned}
& \text{differenzierbare Variablentransformation)} \\
&= \int_B f(\varphi(k(z))) G_\varphi(k(z))^{1/2} |\det Dk(z)| dz \\
&= \int_B f(\psi(z)) \left\{ \sum_{j=1}^m \left(\det D\varphi^{(j)}(k(z)) \right)^2 \right\}^{1/2} |\det Dk(z)| dz \\
&= \int_B f(\psi(z)) \left\{ \sum_{j=1}^m \left(\det D\varphi^{(j)}(k(z)) \det Dk(z) \right)^2 \right\}^{1/2} dz \\
&\quad \left(\det D\varphi^{(j)}(k(z)) \det Dk(z) = \det \left\{ D\varphi^{(j)}(k(z)) Dk(z) \right\} \right. \\
&\quad = \det D(\varphi^{(j)} \circ k)(z) = \det D(\varphi^{(j)} \circ \varphi^{-1} \circ \psi)(z) \\
&\quad = \det D(\varphi \circ \varphi^{-1} \circ \psi(z) \text{ ohne } j\text{-te Komponente}) \\
&\quad = \det D(\psi(z) \text{ ohne } j\text{-te Komponente}) = \det D\psi^{(j)}(z) \left. \right) \\
&= \int_B f(\psi(z)) \left\{ \sum_{j=1}^m (\det D\psi^{(j)}(z))^2 \right\}^{1/2} dz \\
&= \int_B f(\psi(z)) G_\psi(z)^{1/2} dz.
\end{aligned}$$

womit die Behauptung bewiesen ist. ■

Die Formel zur Berechnung von Oberflächenintegralen, die eine Darstellung von M als Graph zu Grunde legt, ist in der Regel besonders einfach auszuwerten. Wir werden deshalb im folgenden meist diese Darstellung benutzen.

Wir betrachten einige Beispiele. Zunächst in \mathbb{R}^2 ; Hyperflächen in \mathbb{R}^2 sind natürlich Kurven. Ist eine Kurve Γ in \mathbb{R}^2 als *Graph* dargestellt, $z = h(y)$ für $a \leq y \leq b$, so erhalten wir mit der ersten Definition

$$\int_\Gamma f(x) d\sigma(x) = \int_\Gamma f(x) ds(x) = \int_a^b f(y, h(y)) \sqrt{1 + h'(y)^2} dy,$$

während wir mit einer beliebigen stetig differenzierbaren *Parameterdarstellung* $\varphi : [a, b] \rightarrow \mathbb{R}^2$ aus der zweiten Darstellung erhalten

$$\int_\Gamma f(x) d\sigma(x) = \int_a^b f(\varphi(t)) \sqrt{\varphi_1'(t)^2 + \varphi_2'(t)^2} dt.$$

Beides sind Formeln, die wir bereits kennen.

Beispiel 12.2 Für *Kreislinie* $S_{1,r}$ mit Radius r um 0

$$h : [-r, r] \rightarrow \mathbb{R}, \quad h(y) = \pm \sqrt{r^2 - y^2}, \quad \text{grad } h(y) = h'(y) = \frac{y}{\mp \sqrt{r^2 - y^2}},$$

$$\begin{aligned}
\int_{S_{1,r}} f(x) d\sigma(x) &= \int_{-r}^r f\left(y, \sqrt{r^2 - y^2}\right) \sqrt{1 + \frac{y^2}{r^2 - y^2}} dy \\
&\quad + \int_{-r}^r f\left(y, -\sqrt{r^2 - y^2}\right) \sqrt{1 + \frac{y^2}{r^2 - y^2}} dy \\
&= \int_{-r}^r \left\{ f\left(y, \sqrt{r^2 - y^2}\right) + f\left(y, -\sqrt{r^2 - y^2}\right) \right\} \frac{r}{\sqrt{r^2 - y^2}} dy.
\end{aligned}$$

Integration über $f(\cdot) \equiv 1$ liefert dies den Kreisumfang $2 \int_{-r}^r r(r^2 - y^2)^{-1/2} dy = ???$ (Integrationsübung); es sei dem Leser überlassen, die Integration auszuführen, auf dem folgenden Weg geht es allerdings noch etwas einfacher.

Andererseits erhalten wir aus der Parameterdarstellung

$$\begin{aligned} \varphi : [0, 2\pi] &\rightarrow \mathbb{R}^2, & \varphi(t) &= (r \cos t, r \sin t), = \varphi, \\ G_\varphi(t) &= (\det D\varphi^{(1)}(t))^2 + (\det D\varphi^{(2)}(t))^2 = \varphi_1'(t)^2 + \varphi_2'(t)^2 = r^2, \end{aligned}$$

von $S_{1,r}$ für das Integral die Formel

$$\int_{S_{1,r}} f(x) \, d\sigma(x) = \int_0^{2\pi} f(r \cos t, r \sin t) r \, dt.$$

für den Kreisumfang $\int_0^{2\pi} r \, dt = 2\pi r$. □

Ist eine Fläche M in \mathbb{R}^3 als *Graph* dargestellt, $z = h(y)$ für $y \in A \subset \mathbb{R}^2$, so erhalten wir mit der ersten Definition

$$\int_M f(x) \, d\sigma(x) := \int_A f(y, h(y)) \sqrt{1 + D_1 h(y)^2 + D_2 h(y)^2} \, dy.$$

Ist die Fläche M in einer beliebigen stetig differenzierbaren *Parameterdarstellung* $\varphi : \mathbb{R}^2 \supset A \rightarrow \mathbb{R}^3$ gegeben, so erhalten wir aus der zweiten Definition

$$\begin{aligned} \int_M f(x) \, d\sigma(x) &= \int_A f(\varphi(x)) G_\varphi(y)^{1/2} \, dy \\ &= \int_A f(\varphi(y)) \left\{ \left(D_1 \varphi_1(y) D_2 \varphi_2(y) - D_2 \varphi_1(y) D_1 \varphi_2(y) \right)^2 \right. \\ &\quad + \left(D_1 \varphi_2(y) D_2 \varphi_3(y) - D_2 \varphi_2(y) D_1 \varphi_3(y) \right)^2 \\ &\quad \left. + \left(D_1 \varphi_3(y) D_2 \varphi_1(y) - D_2 \varphi_3(y) D_1 \varphi_1(y) \right)^2 \right\}^{1/2} \, dy. \end{aligned}$$

Beispiel 12.3 Für das Integral über die *Kugeloberfläche* $S_{2,r}$ mit Radius r um 0

$$h : K_r \rightarrow \mathbb{R}^3, \quad h(y) = \pm \sqrt{r^2 - |y|^2}, \quad \text{grad } h(y) = \mp \frac{1}{\sqrt{1 + |y|^2}} (\eta_1, \eta_2),$$

$$\begin{aligned} \int_{S_{2,r}} f(x) \, d\sigma(x) &= \int_{K_r} f\left(y, \sqrt{r^2 - |y|^2}\right) \sqrt{1 + \frac{|y|^2}{r^2 - |y|^2}} \, dy \\ &\quad + \int_{K_r} f\left(y, -\sqrt{r^2 - |y|^2}\right) \sqrt{1 + \frac{|y|^2}{r^2 - |y|^2}} \, dy \\ &= \int_{K_r} \left\{ f\left(y, \sqrt{r^2 - |y|^2}\right) + f\left(y, -\sqrt{r^2 - |y|^2}\right) \right\} \frac{r}{\sqrt{r^2 - |y|^2}} \, dy. \end{aligned}$$

Integration über $f(\cdot) \equiv 1$ liefert Kugeloberfläche $2 \int_{K_r} r(r^2 - |y|^2)^{-1/2} dy = \dots$ (Übung).

Andererseits hat $S_{2,r}$ die Parameterdarstellung

$$\varphi : [0, 2\pi] \times [0, \pi] \rightarrow \mathbb{R}^3, \quad \varphi(\phi, \vartheta) = \left(r \cos \phi \sin \vartheta, r \sin \phi \sin \vartheta, r \cos \vartheta \right)$$

mit der Gramschen Determinante

$$\begin{aligned}
 G_\varphi(\phi, \vartheta) &= (\det D\varphi^{(1)}(\phi, \vartheta))^2 + (\det D\varphi^{(2)}(\phi, \vartheta))^2 + (\det D\varphi^{(3)}(\phi, \vartheta))^2 \\
 &\quad + \left\{ \det \begin{pmatrix} r \cos \phi \sin \vartheta & r \sin \phi \cos \vartheta \\ 0 & -r \sin \vartheta \end{pmatrix} \right\}^2 \\
 &\quad + \left\{ \det \begin{pmatrix} 0 & -r \sin \vartheta \\ -r \sin \phi \sin \vartheta & r \cos \phi \cos \vartheta \end{pmatrix} \right\}^2 \\
 &= \left\{ \det \begin{pmatrix} -r \sin \phi \sin \vartheta & r \cos \phi \cos \vartheta \\ r \cos \phi \sin \vartheta & r \sin \phi \cos \vartheta \end{pmatrix} \right\}^2 \\
 &= r^4 \left\{ (-\sin^2 \phi \sin \vartheta \cos \vartheta - \cos^2 \phi \sin \vartheta \cos \vartheta)^2 \right. \\
 &\quad \left. + (-\cos \phi \sin^2 \vartheta)^2 + (-\sin \phi \sin^2 \vartheta)^2 \right\} \\
 &= r^4 \left\{ \sin^2 \vartheta \cos^2 \vartheta + \cos^2 \phi \sin^4 \vartheta + \sin^2 \phi \sin^4 \vartheta \right\} \\
 &= r^4 \left\{ \sin^2 \vartheta \cos^2 \vartheta + \sin^2 \vartheta \sin^2 \vartheta \right\} = r^4 \sin^2 \vartheta
 \end{aligned}$$

(das ist vielleicht nicht ganz überraschend, vgl. Beispiel 11.10). Daraus ergibt sich für das Integral über $S_{2,r}$

$$\int_{S_{2,r}} f(x) \, d\sigma(x) = \int_0^\pi \int_0^{2\pi} f(r \cos \phi \sin \vartheta, r \sin \phi \sin \vartheta, r \cos \vartheta) r^2 \sin \vartheta \, d\phi \, d\vartheta.$$

Für die Kugeloberfläche folgt also $\int_0^\pi \int_0^{2\pi} r^2 \sin \vartheta \, d\phi \, d\vartheta = 4\pi r^2$. □

Beispiel 12.4 Als weiteres Beispiel betrachten wir den *Kegelmantel* $M_{r,h}$ des Kegels der Höhe h über dem Kreis mit Radius r : Als Graph dargestellt haben wir (wobei, um eine Verwechslung mit der Höhe h zu vermeiden jetzt $g(\cdot)$ statt $h(\cdot)$ benutzt wird)

$$\begin{aligned}
 g(y) &= h - \frac{h}{r}|y| \text{ für } y \in \mathbb{R}^2 \text{ mit } |y| \leq r, \quad \text{grad } h(y) = \frac{h}{r|y|}(\eta_1, \eta_2) \\
 \int_{M_{r,h}} f(x) \, d\sigma(x) &= \int_{|y| \leq r} f\left(y, h - \frac{h}{r}|y|\right) \left(1 + \frac{h^2}{r^2}\right)^{1/2} \, dy \\
 &= \frac{\sqrt{r^2 + h^2}}{r} \int_{|y| \leq r} f\left(y, h - \frac{h}{r}|y|\right) \, dy.
 \end{aligned}$$

Speziell erhalten wir für die Oberfläche dieses Kegelmantels

$$\begin{aligned}
 F(M_{r,h}) &= \int_{M_{r,h}} 1 \, d\sigma(x) = \frac{\sqrt{r^2 + h^2}}{r} \int_{|y| \leq r} 1 \, dy \\
 &= \frac{\sqrt{r^2 + h^2}}{r} \pi r^2 = \sqrt{r^2 + h^2} \pi r = \pi r L;
 \end{aligned}$$

man beachte, dass hier $L := \sqrt{r^2 + h^2}$ die Länge der Mantellinie ist während $2\pi r$ der Kreisumfang ist, d. h. es gilt Mantelfläche = $\frac{1}{2}$ (Länge der Mantellinie) \times (Kreisumfang), wie man dies aus der Elementargeometrie kennt.

Betrachtet man statt dessen die Parameterdarstellung

$$\varphi : [0, r] \times [0, 2\pi] \rightarrow \mathbb{R}^3, \quad \varphi(\varrho, \phi) = (\varrho \cos \phi, \varrho \sin \phi, h(1 - \varrho/r))$$

mit

$$\begin{aligned} G_\varphi(\varrho, \phi) &= (\det D\varphi^{(1)}(\varrho, \phi))^2 + (\det D\varphi^{(2)}(\varrho, \phi))^2 + (\det D\varphi^{(3)}(\varrho, \phi))^2 \\ &= \det \begin{pmatrix} \sin \phi & \varrho \cos \phi \\ -h/r & 0 \end{pmatrix}^2 + \det \begin{pmatrix} \cos \phi & -\varrho \sin \phi \\ -h/r & 0 \end{pmatrix}^2 + \det \begin{pmatrix} \cos \phi & -\varrho \sin \phi \\ \sin \phi & \varrho \cos \phi \end{pmatrix}^2 \\ &= \left(\frac{\varrho h}{r} \cos \phi\right)^2 + \left(\frac{h}{r} \sin \phi\right)^2 + \varrho^2(\cos^2 \phi + \sin^2 \phi) \\ &= \varrho^2 h^2 / r^2 + \varrho^2 = L^2 \varrho^2 / r^2. \end{aligned}$$

Also gilt

$$\begin{aligned} \int_{M_{r,h}} f(x) \, d\sigma(x) &= \int_0^r \int_0^{2\pi} f(\varphi(\varrho, \phi)) G_\varphi(\varrho, \phi)^{1/2} \, d\phi \, d\varrho \\ &= \int_0^r \int_0^{2\pi} f(\varrho \cos \phi, \varrho \sin \phi, h(1 - \varrho/r)) \frac{L\varrho}{r} \, d\phi \, d\varrho. \end{aligned}$$

Speziell für $f = 1$ erhält man

$$F(M_{r,h}) = \int_0^r \int_0^{2\pi} \frac{L\varrho}{r} \, d\phi \, d\varrho = \pi r L,$$

wieder die Oberfläche des Kegelmantels. □

27: 30. Januar 2007

12.2 Der allgemeine Satz von Stokes

Als Vorstufe für den allgemeinen Satz von Stokes beweisen wir den

Satz 12.5 Sei Q ein Intervall in \mathbb{R}^m ,

$$Q = \left\{ x \in \mathbb{R}^m : a_j \leq \xi_j \leq b_j \text{ für } j = 1, \dots, m \right\}.$$

Wir schreiben Q in der Form $Q = A \times [a_m, b_m]$ mit

$$A = \left\{ y \in \mathbb{R}^{m-1} : a_j \leq \eta_j \leq b_j \text{ für } j = 1, \dots, m-1 \right\}.$$

Mit einer stetig differenzierbaren Funktion $h : A \rightarrow \mathbb{R}$ sei

$$M := \left\{ (y, h(y)) : y \in A \right\} \quad \text{und} \quad \Omega := \left\{ (y, \xi_m) \in Q : \xi_m \approx h(y) \right\}.$$

a) Für jedes $y \in A$, d. h. $(y, h(y)) \in M$, ist

$$n(y, h(y)) = \pm \left\{ 1 + |\text{grad } h(y)|^2 \right\}^{-1/2} \left(D_1 h(y), \dots, D_{m-1} h(y), -1 \right)$$

die bezüglich Ω äußere Normale auf M im Punkt $(y, h(y))$ (hier gilt das $+$ -Zeichen, wenn in der Definition von Ω das $>$ -Zeichen gilt; entsprechend für $-$ und $<$).

- b) Ist $f : Q \rightarrow \mathbb{R}$ stetig differenzierbar mit $f(x) \equiv 0$ in der Nähe des Randes von Q , so gilt, wobei $n_j(\cdot)$ die j -te Komponente von $n(\cdot)$ ist,

$$\int_{\Omega} D_j f(x) dx = \int_M f(x) n_j(x) do(x) \text{ für } j = 1, \dots, m.$$

(Das entsprechende Resultat gilt, wenn in der Darstellung von M die k -te Komponente ($1 \leq k < m$) als Funktion der übrigen $m - 1$ Komponenten aufgefaßt wird.)

Beweis. O.E. betrachten wir nur den Fall, wo in der Definition von Ω das $>$ -Zeichen und in $n(y, h(y))$ das $+$ -Zeichen stehen; der Leser führe selbst die Modifikationen für den anderen Fall durch.

- a) Wir können M auch schreiben in der Form

$$M = \left\{ (y, \xi_m) \in \mathbb{R}^m : g(y, \xi_m) = 0 \right\} \quad \text{mit} \quad g(y, \xi_m) := h(y) - \xi_m.$$

Dann wissen wir, dass

$$\text{grad } g(y, h(y)) = \left(D_1 h(y), \dots, D_{m-1} h(y), -1 \right)$$

die (nichtnormierte) Normale auf M im Punkt $(y, h(y))$ ist und in die Richtung zeigt, wo $g > 0$ ist, also $\xi_m < h(y)$; d. h. bezüglich Ω nach außen. Damit folgt die Aussage über $n(y, h(y))$ durch Normierung.

- b) Es bleibt die Integralformel zu beweisen. Wir unterscheiden zwei Fälle:

$j = m$:

$$\begin{aligned} \int_{\Omega} D_m f(x) dx &= \int_A \int_{h(y)}^{b_m} D_m f(y, t) dt dy = \int_A \left\{ f(y, b_m) - f(y, h(y)) \right\} dy \\ &\quad (\text{da } f(y, t) \equiv 0 \text{ für } t \text{ nahe } b_m) \\ &= - \int_A f(y, h(y)) dy \\ &= - \int_A f(y, h(y)) \underbrace{\left\{ 1 + |\text{grad } h(y)|^2 \right\}^{-1/2}}_{=-n_m(y, h(y))} \left\{ 1 + |\text{grad } h(y)|^2 \right\}^{1/2} dy \\ &\quad (\text{nach Definition des Oberflächenintegrals}) \\ &= \int_M f(x) n_m(x) do(x). \end{aligned}$$

$j < m$:

$$\begin{aligned} \int_{\Omega} D_j f(x) dx &= \int_A \int_{h(y)}^{b_m} D_j f(y, t) dt dy \\ &\quad (\text{Differentiation parameterabhängiger Integrale, Satz 11.14}) \\ &= \int_A \left\{ D_j \int_{h(y)}^{b_m} f(y, t) dt + f(y, h(y)) D_j h(y) \right\} dy \\ &\quad (\text{wegen } A = A' \times [a_j, b_j], y = (y', \eta_j)) \\ &= \int_{A'} \int_{a_j}^{b_j} D_j \int_{h(y', \eta_j)}^{b_m} f(y', \eta_j, t) dt d\eta_j dy' + \int_A f(y, h(y)) D_j h(y) dy \end{aligned}$$

$$\begin{aligned}
&= \int_{A'} \left\{ \int_{h(y', b_j)}^{b_m} f(y', b_j, t) dt - \int_{h(y', a_j)}^{b_m} f(y', a_j, t) dt \right\} dy' \\
&\quad + \int_A f(y, h(y)) D_j h(y) dy \\
&\quad \text{(ins erste Integral gehen nur Werte von } f \text{ auf dem Rand von } \Omega \text{ ein)} \\
&= 0 + \int_A f(y, h(y)) \underbrace{D_j h(y) \left\{ 1 + |\text{grad } h(y)|^2 \right\}^{-1/2}}_{=n_j(y, h(y))} \left\{ 1 + |\text{grad } h(y)|^2 \right\}^{1/2} dy \\
&= \int_M f(x) n_j(x) do(x).
\end{aligned}$$

Damit ist die Formel für alle j bewiesen. ■

Mit Hilfe des eben bewiesenen Satzes und des Satzes von Heine–Borel sind wir nun in der Lage, den allgemeinen Satz von Stokes zu beweisen.

Satz 12.6 (Allgemeiner Stokesscher Satz) *Sei Ω eine beschränkte offene Menge in \mathbb{R}^m mit stetig differenzierbarem Rand $M = \partial\Omega$. Ist $n(x)$ für jedes $x \in M$ die bezüglich Ω äußere Normale auf M im Punkt x und $f : \Omega \cup M \rightarrow \mathbb{R}$ stetig differenzierbar, so gilt für $j = 1, \dots, m$*

$$\int_{\Omega} D_j f(x) dx = \int_M f(x) n_j(x) do(x).$$

Beweis. Für jedes $x \in \Omega \cup M$ sei Q_x ein offenes Intervall so, dass

entweder $Q_x \cap M$ als Graph darstellbar ist,
oder $Q_x \cap M = \emptyset$.

Dann ist $\{Q_x : x \in \Omega \cup M\}$ eine offene Überdeckung der kompakten Menge $\Omega \cup M$. Also genügen endlich viele $Q_1 := Q_{x_1}, \dots, Q_n := Q_{x_n}$ um Ω zu überdecken.

$\{\varphi_1, \dots, \varphi_n\}$ sei eine zur Überdeckung $\{Q_1, \dots, Q_n\}$ gehörige Zerlegung der Eins (vgl. Satz 11.15). Dann gilt (wegen $\sum_{k=1}^n \varphi_k(x) = 1$ für alle $x \in \Omega \cup M$)

$$\begin{aligned}
\int_{\Omega} D_j f(x) dx &= \int_{\Omega} D_j \left(\sum_{k=1}^n \varphi_k(x) f(x) \right) dx = \sum_{k=1}^n \int_{\Omega} D_j \left(\varphi_k(x) f(x) \right) dx \\
&\quad \text{(da } \varphi_k(x) f(x) = 0 \text{ außerhalb von } Q_k) \\
&= \sum_{k=1}^n \int_{\Omega \cap Q_k} D_j \left(\varphi_k(x) f(x) \right) dx.
\end{aligned}$$

Wir untersuchen die einzelnen Integrale und unterscheiden dazu zwei Fälle:

$Q_k \subset \Omega$ (d. h. $Q_k \cap M = \emptyset$):

$$\begin{aligned}
&\int_{\Omega \cap Q_k} D_j \left(\varphi_k(x) f(x) \right) dx = \int_{Q_k} D_j \left(\varphi_k(x) f(x) \right) dx \\
&= \int_{Q'_k} \int_{a_j}^{b_j} D_j \left(\varphi_k(x) f(x) \right) d\xi dx' \\
&= \int_{Q'_k} \left\{ \varphi_k(x', b_j) f(x', b_j) - \varphi_k(x', a_j) f(x', a_j) \right\} dx' \\
&= 0 \quad \text{(da } \varphi_k f(x) = 0 \text{ am Rand von } Q_k) \\
&= \int_M \varphi_k(x) f(x) n_j(x) do(x) \quad \text{(da } \varphi_k(x) = 0 \text{ auf } M).
\end{aligned}$$

$Q_k \cap M$ ist als Graph darstellbar: ein ξ_l als Funktion h der übrigen $m-1$ Koordinaten; $\Omega \cap Q_k$ ist dann die Menge der Punkte aus Q_k mit $\xi_l >$ oder $\xi_l < h(\xi_1, \dots, \xi_{l-1}, \xi_{l+1}, \dots, \xi_m)$: Mit Satz 12.5 folgt also

$$\begin{aligned} \int_{\Omega \cap Q_k} D_j(\varphi_k(x)f(x)) \, dx &= \int_{Q_k \cap M} \varphi_k(x)f(x)n_j(x) \, d\sigma(x) \\ &\quad (\text{da } \varphi_k(x)f(x) = 0 \text{ außerhalb von } Q_k) \\ &= \int_M \varphi_k(x)f(x)n_j(x) \, d\sigma(x). \end{aligned}$$

Insgesamt erhalten wir also

$$\begin{aligned} \int_{\Omega} D_j f(x) \, dx &= \sum_{k=1}^n \int_{\Omega \cap Q_k} D_j(\varphi_k(x)f(x)) \, dx = \sum_{k=1}^n \int_M \varphi_k(x)f(x)n_j(x) \, d\sigma(x) \\ &= \int_M f(x)n_j(x) \, d\sigma(x), \end{aligned}$$

womit der Satz bewiesen ist. ■

Eine genaue Durchsicht der obigen Beweise zeigt, dass die Aussage des allgemeinen Stokesschen Satzes auch dann gilt, wenn der Rand von Ω gewisse Ecken und/oder Kanten hat. Bei Bedarf überlegt man sich das am besten für den Einzelfall, da es schwierig und auch nicht lohnend ist, eine möglichst allgemeine Form des Satzes anzugeben.

Als Anwendung betrachten wir den *Auftrieb* eines in eine inkompressible Flüssigkeit eingetauchten Körpers. Die physikalische Aussage, die wir dabei benutzen ist: in einer Tiefe t unter der Oberfläche der Flüssigkeit herrscht ein Druck pro Flächeneinheit von $t\rho$, wobei ρ die Dichte der Flüssigkeit ist (und zwar ist dieser Druck unabhängig von der Lage des Flächenstücks). Dazu kommt natürlich noch der evtl. an der Oberfläche herrschende Druck, der aber, wie man leicht sieht, keinen Beitrag liefert; ebenso spielt es keine Rolle, wie tief der Körper unter der Oberfläche ist, sofern er nur vollständig eingetaucht ist.

Satz 12.7 *Der Auftrieb eines (glatten) Körpers K in einem flüssigen oder gasförmigen Medium ist (unabhängig von der Form des Körpers und der Eintauchtiefe) gleich dem Gewicht $G(K)$ der durch den Körper verdrängten Menge des Mediums; die Kraft wirkt senkrecht nach oben (hat also keine horizontale Komponente).*

Beweis. Wir gehen von der physikalischen Annahme aus, dass die spezifische Dichte des Mediums nur von ξ_3 (der senkrechten Komponente) abhängt, $\rho(x) = \tilde{\rho}(\xi_3)$ (bei einer inkompressiblen Flüssigkeit hat man konstantes ρ , was den Beweis kaum vereinfacht).

Der Druck in einem Punkt mit 3-ter Komponente ξ_3 ist also

$$p(x) = \tilde{p}(\xi_3) = \tilde{p}(0) + \int_{\xi_3}^0 \tilde{\rho}(t) \, dt.$$

In einem Punkt x der Oberfläche M des Körpers mit äußerer Normale $n(x)$ wirkt auf einer Flächeneinheit der Oberfläche also die Kraft $-p(x)n(x)$. Damit folgt für die j -te Komponente der insgesamt

auf den Körper wirkenden Kraft k :

$$\begin{aligned} k_j &= \int_M \{-p(x)n_j(x)\} \, d\sigma(x) = - \int_M \tilde{p}(\xi_3)n_j(x) \, d\sigma(x) \\ &\quad \text{(allg. Stokesscher Satz mit } f(x) = p(x)) \\ &= \begin{cases} - \int_K D_j \tilde{p}(\xi_3) \, dx = 0 & \text{für } j = 1, 2, \\ - \int_K D_3 \tilde{p}(\xi_3) \, dx = \int_K \tilde{q}(\xi_3) \, dx = G(K) & \text{für } j = 3 \end{cases} \end{aligned}$$

Damit sind alle Aussagen des Satzes bewiesen. ■

28: 2. Februar 2007

12.3 Integrale auf allgemeineren Mannigfaltigkeiten

Sei $A \subset \mathbb{R}^k$ mit $k < m$ eine offene Teilmenge, $\varphi : A \rightarrow \mathbb{R}^m$ stetig differenzierbar,

$$\varphi^{(j_1, \dots, j_k)}(y) = \left(\varphi_{j_1}(y), \dots, \varphi_{j_k}(y) \right) \quad 1 \leq j_1 < j_2 < \dots < j_k \leq m.$$

Dann ist die *Gramsche Determinante* von φ definiert durch

$$G_\varphi(y) = \sum_{1 \leq j_1 < j_2 < \dots < j_k \leq m} \left(\det D\varphi^{(j_1, \dots, j_k)}(y) \right)^2;$$

es folgt wieder aus der Linearen Algebra, dass gilt (vgl. G. Fischer: Lineare Algebra; Determinanten–Multiplikationstheorem)

$$G_\varphi(y) = \det \left(\left(\langle D_i \varphi(y), D_j \varphi(y) \rangle \right)_{i,j=1, \dots, k} \right).$$

Ist $G_\varphi(y) \neq 0$ für alle $y \in A$, d. h. φ ist *regulär*, so heißt das Bild $M = \varphi(A)$ eine differenzierbare Mannigfaltigkeit (allgemeiner ist eine differenzierbare Mannigfaltigkeit ein Gebilde, das stückweise in dieser Form dargestellt werden kann; für die Parameterdarstellungen φ, ψ von zwei Teilstücken muß dann $\varphi^{-1} \circ \psi$, dort wo es definiert ist [d. h. auf dem ϕ -Urbild des Schnittes der Wertebereiche von φ und ψ] stetig differenzierbar sein mit $\det D(\varphi^{-1} \circ \psi)(y) \neq 0$, *Verträglichkeit der „lokalen Karten“*.) Das Integral einer (z. B. stetigen) Funktion auf M wird definiert durch

$$\int_M f(x) \, d\sigma(x) := \int_A f(\varphi(y)) G_\varphi(y)^{1/2} \, dy.$$

Der Faktor $G_\varphi(\cdot)^{1/2}$ steht hier, weil er gleich dem Volumen des durch die Vektoren $D_1 \varphi(\cdot), \dots, D_k \varphi(\cdot)$ aufgespannten Parallelepipedes ist, dem differentiellen Bild des Einheitswürfels im Punkt $y \in A \subset \mathbb{R}^k$.¹⁰ Wie in Abschnitt 26.1 kann die Unabhängigkeit dieses Integrals von der Parameterdarstellung bewiesen werden.

Ein einfacher Spezialfall ist:

¹⁰Sind x_1, \dots, x_k linear unabhängige Vektoren in \mathbb{R}^m , $\{e_1, \dots, e_k\}$ eine ONB in $L\{x_1, \dots, x_k\}$, $x_j = \sum_{l=1}^k c_{jl} e_l$ für $j = 1, \dots, k$, so gilt für das Volumen V des durch x_1, \dots, x_k aufgespannten Parallelepipedes

$$\begin{aligned} V^2 &= \left(\det \begin{pmatrix} c_{11} & \cdots & c_{1k} \\ \vdots & & \vdots \\ c_{k1} & \cdots & c_{kk} \end{pmatrix} \right)^2 = \det \begin{pmatrix} c_{11} & \cdots & c_{1k} \\ \vdots & & \vdots \\ c_{k1} & \cdots & c_{kk} \end{pmatrix} \det \begin{pmatrix} c_{11} & \cdots & c_{k1} \\ \vdots & & \vdots \\ c_{1k} & \cdots & c_{kk} \end{pmatrix} \\ &= \det \begin{pmatrix} \sum c_{1l} c_{1l} & \cdots & \sum c_{1l} c_{kl} \\ \vdots & & \vdots \\ \sum c_{kl} c_{1l} & \cdots & \sum c_{kl} c_{kl} \end{pmatrix} = \det \begin{pmatrix} \langle x_1, x_1 \rangle & \cdots & \langle x_1, x_k \rangle \\ \vdots & & \vdots \\ \langle x_k, x_1 \rangle & \cdots & \langle x_k, x_k \rangle \end{pmatrix}. \end{aligned}$$

Beispiel 12.8 *Kurve in \mathbb{R}^m* : In diesem Fall ist A ein Intervall $[a, b]$ und

$$G_\varphi(y) = \sum_{1 \leq j \leq m} \varphi'_j(y)^2,$$
$$\int_M f(x) \, do(x) = \int_a^b f(\varphi(y)) \left\{ \sum_{j=1}^m \varphi'_j(y)^2 \right\}^{1/2} dy.$$

Dies ist das bereits bekannte Integral über eine Kurve, vgl. S. 138. Für beliebige m hatten wir das bisher nicht betrachtet, außer im Fall $f \equiv 1$, wo wir die bereits bekannte Formel für die Kurvenlänge erhalten (vgl. §22). \square

12.4 Übungsaufgaben

12.1 Man berechne die Gramsche Determinante für die Polarkoordinatendarstellung der Fläche

$$\left\{ (x, y, z) \in \mathbb{R}^3 : \frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1 \right\}.$$

12.2 Man berechne die Oberfläche der Einheitssphäre in \mathbb{R}^4 nach dem Vorbild von Beispiel 12.2 und 12.3.

13 Die klassischen Integralsätze von Gauß, Green und Stokes

13.1 Die Sätze von Gauß und Green

Die beiden Sätze dieses Abschnitts sind einfache Folgerungen aus dem allgemeinen Satz von Stokes. Es sei hier Ω eine offene Teilmenge von \mathbb{R}^m mit einem stetig differenzierbaren Rand $M = \partial\Omega$ wie dies z. B. in Satz 12.6 vorausgesetzt wurde.

Für ein stetig differenzierbares Vektorfeld $f : \Omega \rightarrow \mathbb{R}^m$ ist die *Divergenz* definiert durch

$$\operatorname{div} f(x) := \sum_{j=1}^m D_j f_j(x).$$

Der folgende Satz liefert eine anschauliche physikalische Deutung der Divergenz.

Satz 13.1 (Gaußscher Divergenzsatz) Sei $\Omega \subset \mathbb{R}^m$ offen mit stetig differenzierbarem Rand M ; für jedes $x \in M$ sei $n(x)$ die normierte bezüglich Ω äußere Normale auf M im Punkt x . Dann gilt

$$\int_{\Omega} \operatorname{div} f(x) \, dx = \int_M \langle f(x), n(x) \rangle \, d\sigma(x).$$

Bemerkung 13.2 a) Dieser Satz stellt offenbar die direkte Verallgemeinerung des Hauptsatzes der Differenzial- und Integralrechnung in \mathbb{R}^1 auf den \mathbb{R}^m dar (für ein Intervall $[a, b]$ zeigt die äußere Normale $n(x)$ in a nach links, in b nach rechts, ist also in a negativ, in b positiv).

b) Faßt man das Vektorfeld f als *Geschwindigkeitsfeld* einer Strömung in \mathbb{R}^m auf, so gibt offenbar $\langle f(x), n(x) \rangle$ den Fluß pro Flächeneinheit durch ein „kleines“ Flächenstück von M um den Punkt x von Ω nach $\mathbb{R}^m \setminus \Omega$ an. Das Integral auf der rechten Seite gibt also den gesamten Fluß von Ω durch M nach $\mathbb{R}^m \setminus \Omega$ an. Da dies entsprechend für Teilgebiete von Ω gilt, kann (nach Grenzübergang für immer kleinere Volumina) $\operatorname{div} f(x)$ als *Quellendichte* des Vektorfeldes f aufgefaßt werden; das Integral auf der linken Seite stellt dann die gesamte *Quellenstärke* des Vektorfeldes f in Ω dar.

Beweis von Satz 13.1. Durch Anwendung von Satz 12.6 auf f_j statt f erhalten wir

$$\begin{aligned} \int_{\Omega} \operatorname{div} f(x) \, dx &= \sum_{j=1}^m \int_{\Omega} D_j f_j(x) \, dx = \sum_{j=1}^m \int_M f_j(x) n_j(x) \, d\sigma(x) \\ &= \int_M \sum_{j=1}^m f_j(x) n_j(x) \, d\sigma(x) = \int_M \langle f(x), n(x) \rangle \, d\sigma(x), \end{aligned}$$

was zu beweisen war. ■

Beispiel 13.3 In \mathbb{R}^m betrachten wir das Vektorfeld

$$f : \mathbb{R}^m \rightarrow \mathbb{R}^m, \quad f(x) = x.$$

Dann gilt

$$\operatorname{div} f(x) = \sum_{j=1}^m D_j x_j = \sum_{j=1}^m 1 = m.$$

Wählen wir z.B.

$$\begin{aligned}\Omega &= K_{m,r} \text{ Kugel mit Radius } r \text{ um } 0, \\ M &= S_{m-1,r} \text{ Sphäre mit Radius } r \text{ um } 0,\end{aligned}$$

so folgt für $V(K_{m,r}) = \text{Volumen von } K_{m,r}$ und $O(S_{m-1,r}) = \text{Oberfläche von } S_{m-1,r}$ aus obigem Satz mit $f(x) = x$

$$\begin{aligned}mV(K_{m,r}) &= \int_{K_{m,r}} m \, dx = \int_{K_{m,r}} \operatorname{div} f(x) \, dx = \int_{S_{m-1,r}} \langle f(x), n(x) \rangle \, do(x) \\ &= \int_{S_{m-1,r}} r \, do(x) = rO(S_{m-1,r}).\end{aligned}$$

Dies bestätigt insbesondere die folgenden bekannten Formeln:

$$\mathbf{m} = 2: \quad V(K_{2,r}) = \pi r^2, \quad O(S_{1,r}) = 2\pi r.$$

$$\mathbf{m} = 3: \quad V(K_{3,r}) = \frac{4}{3}\pi r^3, \quad O(S_{2,r}) = 4\pi r^2.$$

Als weiteres Beispiel betrachten wir

$\mathbf{m} = 4$: Zunächst gilt (iterative Integration)

$$\begin{aligned}V(K_{4,r}) &= \int_{K_{4,r}} 1 \, dx = \int_{-r}^r \int_{K_{3,(r^2-t^2)^{1/2}}} 1 \, dx' \, dt = \int_{-r}^r V(K_{3,(r^2-t^2)^{1/2}}) \, dt \\ &= \frac{4}{3}\pi \int_{-r}^r (r^2 - t^2)^{3/2} \, dt.\end{aligned}$$

Mit der Substitution

$$t = \varphi(\tau) \text{ mit } \varphi: \left[-\frac{\pi}{2}, \frac{\pi}{2}\right] \rightarrow \mathbb{R}, \quad \varphi(\tau) = r \sin \tau, \quad \varphi'(\tau) = r \cos \tau$$

erhalten wir also ¹¹

$$V(K_{4,r}) = \frac{4}{3}\pi r^4 \int_{-\pi/2}^{\pi/2} \cos^4 \tau \, d\tau$$

¹¹Dabei benutzen wir

$$\begin{aligned}\int \cos^4 \tau \, d\tau &= \int \cos^3 \cos \tau \, d\tau = \cos^3 \tau \sin \tau + 3 \int \cos^2 \tau \sin^2 \tau \, d\tau \\ &= \cos^3 \tau \sin \tau + 3 \int \cos^2 \tau \, d\tau - 3 \int \cos^4 \tau \, d\tau \\ &= \frac{1}{4} \cos^3 \tau \sin \tau + \frac{3}{4} \int \cos^2 \tau \, d\tau \\ &= \frac{1}{4} \cos^3 \tau \sin \tau + \frac{3}{4} \cos \tau \sin \tau + \frac{3}{4} \int \sin^2 \tau \, d\tau \\ &= \frac{1}{4} \cos^3 \tau \sin \tau + \frac{3}{4} \cos \tau \sin \tau + \frac{3}{4} \int 1 \, d\tau - \frac{3}{4} \int \cos^2 \tau \, d\tau \\ &= \frac{1}{4} \cos^3 \tau \sin \tau + \frac{3}{8} \cos \tau \sin \tau + \frac{3}{8} \tau.\end{aligned}$$

$$\begin{aligned}
&= \frac{4}{3}\pi r^4 \left\{ \left(\frac{1}{4} \cos^3 \tau + \frac{3}{8} \cos \tau \right) \sin \tau + \frac{3}{8} \tau \right\} \Big|_{-\pi/2}^{\pi/2} \\
&= \frac{4}{3}\pi r^4 \left\{ \frac{3}{8} \frac{\pi}{2} + \frac{3}{8} \frac{\pi}{2} \right\} = \frac{1}{2}\pi^2 r^4.
\end{aligned}$$

Mit obiger Formel ergibt sich daraus (vgl. auch Aufgabe 12.2)

$$O(S_{3,r}) = \frac{4}{r}V(K_{4,r}) = 2\pi^2 r^3.$$

Formeln für Volumen der Kugeln und Oberflächen der Sphären beliebiger Dimension findet man z. B. in O.Forster: Analysis 3, Beispiel 5.7. \square

Der folgende Satz ist besonders nützlich beim Umgang mit dem *Laplaceschen Differenzialausdruck* $\Delta = \sum_{j=1}^m D_j^2$.

Satz 13.4 (Greensche Formel) Seien $\Omega \subset \mathbb{R}^m$ offen mit stetig differenzierbarem Rand M und die Funktionen $f, g : \Omega \cup M \rightarrow \mathbb{R}$ zweimal stetig differenzierbar. Dann gilt

$$\begin{aligned}
&\int_{\Omega} \left\{ f(x)\Delta g(x) - (\Delta f(x))g(x) \right\} dx \\
&= \int_M \left\{ f(x) \langle \text{grad } g(x), n(x) \rangle - g(x) \langle \text{grad } f(x), n(x) \rangle \right\} d(x).
\end{aligned}$$

(Dabei ist $\langle \text{grad } h(x), n(x) \rangle =: „\partial h(x)/\partial n(x)“$ für $x \in M$ die Richtungsableitung von h in Richtung der äußeren Normalen auf M , kurz die Normalableitung von h auf M .)

Bemerkung 13.5 Dieses Resultat entspricht der 1-dimensionalen Formel

$$\begin{aligned}
&\int_a^b \left\{ f(x)g''(x) - f''(x)g(x) \right\} dx \\
&= \left\{ f(b)g'(b) - f'(b)g(b) \right\} - \left\{ f(a)g'(a) - f'(a)g(a) \right\}.
\end{aligned}$$

Beweis von Satz 13.4. Mit Hilfe des Divergenzsatzes (Satz 13.1) folgt

$$\begin{aligned}
&\int_{\Omega} \left\{ f(x)\Delta g(x) - (\Delta f(x))g(x) \right\} dx \\
&= \int_{\Omega} \text{div} \left\{ f(x) \text{grad } g(x) - g(x) \text{grad } f(x) \right\} dx \\
&= \int_M \left\{ \langle f(x) \text{grad } g(x), n(x) \rangle - \langle g(x) \text{grad } f(x), n(x) \rangle \right\} d\sigma(x).
\end{aligned}$$

Das ist die gewünschte Gleichung. \blacksquare

Besonders wichtig ist der folgende Spezialfall der Greenschen Formel:

Korollar 13.6 Sind f und g wie im Satz 13.4 und gilt außerdem $f(x) = g(x) = 0$ auf M (oder $f(x) = 0$ und $\text{grad } f(x) = 0$, oder $g(x) = 0$ und $\text{grad } g(x) = 0$), so folgt

$$\int_{\Omega} \left\{ f(x)\Delta g(x) - g(x)\Delta f(x) \right\} dx = 0.$$

13.2 Der klassische Satz von Stokes in der Ebene

Wir beweisen hier zunächst den klassischen Stokesschen Satz im Spezialfall der Ebene \mathbb{R}^2 . Den entsprechenden Satz im Raum \mathbb{R}^3 werden wir im folgenden Abschnitt beweisen.

Für ein stetig differenzierbares Vektorfeld $f : \Omega \rightarrow \mathbb{R}^2$ ($\Omega \subset \mathbb{R}^2$) definieren wir die *Rotation*

$$\operatorname{rot} f(x) := D_1 f_2(x) - D_2 f_1(x).$$

(Würde man f als Vektorfeld in \mathbb{R}^3 betrachten, das nicht von ξ_3 abhängt mit $f_3(x) = 0$, so wäre dies gerade die dritte Komponente der Rotation von f , vgl. Abschnitt 13.3; allerdings wären in diesem Fall die anderen beiden Komponenten gleich Null.)

Satz 13.7 (Stokesscher Satz in der Ebene) *Sei Ω ein Gebiet in \mathbb{R}^2 mit stetig differenzierbarer Randkurve $M = \Gamma$, die so orientiert sei, dass Ω links von Γ liegt (Umlauf um Ω entgegen dem Uhrzeigersinn; mathematisch positiv) $\gamma : [a, b] \rightarrow \mathbb{R}^2$, $\gamma(a) = \gamma(b)$.*

Ist $f : \Omega \cup \Gamma \rightarrow \mathbb{R}^2$ ein stetig differenzierbares Vektorfeld, so gilt

$$\int_{\Omega} \operatorname{rot} f(x) \, dx = \int_{\Gamma} f_1(x) \, d\xi_1 + f_2(x) \, d\xi_2 = \int_a^b \langle f(\gamma(t)), \gamma'(t) \rangle \, dt.$$

Beweis. Mit dem Gaußschen Satz (angewandt auf f_2 mit $j = 1$ und auf f_1 mit $j = 2$) erhalten wir

$$\begin{aligned} \int_{\Omega} \operatorname{rot} f(x) \, dx &= \int_{\Omega} (D_1 f_2(x) - D_2 f_1(x)) \, dx \\ &= \int_M \{f_2(x)n_1(x) - f_1(x)n_2(x)\} \, d\sigma(x) \\ &= \int_M \langle (f_1(x), f_2(x)), (-n_2(x), n_1(x)) \rangle \, d\sigma(x) \\ &= \int_a^b \langle (f_1(\gamma(t)), f_2(\gamma(t))), (-n_2(\gamma(t)), n_1(\gamma(t))) \rangle |\gamma'(t)| \, dt. \end{aligned}$$

Der Vektor $(-n_2(x), n_1(x))$ entsteht aus dem äußeren Normalenvektor $n(x) = (n_1(x), n_2(x))$ durch Drehung um $90^\circ = \pi/2$ entgegen dem Uhrzeigersinn, nämlich durch Anwenden der Matrix

$$\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} \cos \frac{\pi}{2} & -\sin \frac{\pi}{2} \\ \sin \frac{\pi}{2} & \cos \frac{\pi}{2} \end{pmatrix}.$$

Also ist $(-n_2(x), n_1(x))$ der Tangentialvektor im Punkt x an die Kurve Γ , und zwar in Richtung der Orientierung von Γ . Folglich können wir wie folgt weiterrechnen:

$$= \int_a^b \langle f(\gamma(t)), \frac{\gamma'(t)}{|\gamma'(t)|} \rangle |\gamma'(t)| \, dt = \int_a^b \langle f(\gamma(t)), \gamma'(t) \rangle \, dt.$$

Das ist die behauptete Gleichung. ■

Betrachtet man speziell ein Vektorfeld f mit $\operatorname{rot} f(x) \equiv 1$, z. B. $f(x) = (-\xi_2, 0)$ oder $(0, \xi_1)$ oder $\frac{1}{2}(-\xi_2, \xi_1)$, so liefert der obige Satz eine Formel für die Fläche, ausgedrückt durch das Kurvenintegral von f über die Randkurve (vgl. Aufgabe 13.3).

13.3 Der klassische Satz von Stokes im \mathbb{R}^3

Wir denken uns jetzt eine („krumme“) Fläche F in \mathbb{R}^3 dargestellt als Graph über einem ebenen Flächenstück \tilde{F} im \mathbb{R}^2 . Den allgemeinsten interessierenden Fall kann man sich aus solchen Flächenstücken zusammengesetzt denken (doppelt auftretende Kurvenstücke heben sich dabei im Integral weg, da sie in unterschiedlichen Richtungen durchlaufen werden). Sei also

$$F = \left\{ \left(\xi_1, \xi_2, h(\xi_1, \xi_2) \right) \in \mathbb{R}^3 : (\xi_1, \xi_2) \in \tilde{F} \right\}$$

mit einer stetig differenzierbaren Funktion $h : \tilde{F} \rightarrow \mathbb{R}$. $\tilde{\Gamma}$ sei die Randkurve von \tilde{F} , $\tilde{\gamma} : [a, b] \rightarrow \mathbb{R}^2$, wobei \tilde{F} wieder links von $\tilde{\Gamma}$ liege. Dann ist die Randkurve Γ von F gegeben durch

$$\gamma : [a, b] \rightarrow \mathbb{R}^3, \quad \gamma(t) = \left(\tilde{\gamma}(t), h(\tilde{\gamma}(t)) \right).$$

29: 06. Februar 2007

Weiter sei $n(x)$ die bezüglich ξ_3 nach oben zeigende Normale auf F im Punkt $x \in F$. (Damit kann die „Orientierung“ der Fläche F und der Randkurve Γ auch ohne Bezug auf \tilde{F} und $\tilde{\Gamma}$ wie folgt beschrieben werden. Schreitet man aufrecht im Sinne der Normalen $n(x)$ in Richtung der Kurve Γ voran, so liegt F links. So verstanden darf $n(x)$ auch nach unten zeigen; dann wäre auch die Orientierung der Randkurve zu ändern.)

Für ein stetig differenzierbares Vektorfeld $f : A \rightarrow \mathbb{R}^3$ sei die *Rotation* von f definiert durch

$$\text{rot } f(x) := \left(D_2 f_3(x) - D_3 f_2(x), D_3 f_1(x) - D_1 f_3(x), D_1 f_2(x) - D_2 f_1(x) \right).$$

Wie schon oben angemerkt, entspricht die dritte Komponente gerade der Rotation in \mathbb{R}^2 . Mit diesen Vereinbarungen gilt

Satz 13.8 (Stokesscher Integralsatz in \mathbb{R}^3) Seien F und Γ wie oben, $f : F \cup \Gamma \rightarrow \mathbb{R}^3$ stetig differenzierbar (d. h. f ist auf einer offenen \mathbb{R}^3 -Umgebung von $F \cup \Gamma$ stetig differenzierbar). Dann gilt

$$\begin{aligned} \int_F \langle \text{rot } f(x), n(x) \rangle \, d\sigma(x) &= \int_{\tilde{F}} f_1(x) \, d\xi_1 + f_2(x) \, d\xi_2 + f_3(x) \, d\xi_3 \\ &= \int_a^b \langle f(\gamma(t)), \gamma'(t) \rangle \, dt. \end{aligned}$$

(Im Beweis tauchen an einer Stelle zweite Ableitungen von h auf, d. h. es wird eigentlich benutzt, dass h zweimal stetig differenzierbar ist. Durch geeignete Approximation folgt die Aussage aber auch für stetig differenzierbares h .)

Beweis. Mit den obigen Bezeichnungen gilt

$$\begin{aligned} &\int_{\Gamma} \langle f(\gamma(t)), \gamma'(t) \rangle \, dt \\ &= \int_a^b \left\langle f\left(\tilde{\gamma}(t), h(\tilde{\gamma}(t))\right), \left(\tilde{\gamma}'_1(t), \tilde{\gamma}'_2(t), D_1 h(\tilde{\gamma}(t))\tilde{\gamma}'_1(t) + D_2 h(\tilde{\gamma}(t))\tilde{\gamma}'_2(t)\right) \right\rangle \, dt \end{aligned}$$

$$\begin{aligned}
&= \int_a^b \left\{ f_1(\tilde{\gamma}(t), h(\tilde{\gamma}(t))) \tilde{\gamma}'_1(t) + f_2(\tilde{\gamma}(t), h(\tilde{\gamma}(t))) \tilde{\gamma}'_2(t) \right. \\
&\quad \left. + f_3(\tilde{\gamma}(t), h(\tilde{\gamma}(t))) \left(D_1 h(\tilde{\gamma}(t)) \tilde{\gamma}'_1(t) + D_2 h(\tilde{\gamma}(t)) \tilde{\gamma}'_2(t) \right) \right\} dt \\
&= \int_a^b \left\{ \underbrace{\left(f_1(\tilde{\gamma}(t), h(\tilde{\gamma}(t))) + f_3(\tilde{\gamma}(t), h(\tilde{\gamma}(t))) D_1 h(\tilde{\gamma}(t)) \right)}_{=:k_1(\tilde{\gamma}(t))} \tilde{\gamma}'_1(t) \right. \\
&\quad \left. + \underbrace{\left(f_2(\tilde{\gamma}(t), h(\tilde{\gamma}(t))) + f_3(\tilde{\gamma}(t), h(\tilde{\gamma}(t))) D_2 h(\tilde{\gamma}(t)) \right)}_{=:k_2(\tilde{\gamma}(t))} \tilde{\gamma}'_2(t) \right\} dt \\
&= \int_a^b \langle k(\tilde{\gamma}(t)), \tilde{\gamma}'(t) \rangle dt \\
&\quad \text{(Satz von Stokes in der Ebene)} \\
&= \int_{\tilde{F}} \operatorname{rot} k(\xi_1, \xi_2) d(\xi_1, \xi_2) = \int_{\tilde{F}} \left\{ \frac{\partial}{\partial \xi_1} k_2(\xi_1, \xi_2) - \frac{\partial}{\partial \xi_2} k_1(\xi_1, \xi_2) \right\} d(\xi_1, \xi_2) \\
&= \int_{\tilde{F}} \left\{ \frac{\partial}{\partial \xi_1} \left(f_2(\xi_1, \xi_2, h(\xi_1, \xi_2)) + f_3(\xi_1, \xi_2, h(\xi_1, \xi_2)) D_2 h(\xi_1, \xi_2) \right) \right. \\
&\quad \left. - \frac{\partial}{\partial \xi_2} \left(f_1(\xi_1, \xi_2, h(\xi_1, \xi_2)) + f_3(\xi_1, \xi_2, h(\xi_1, \xi_2)) D_1 h(\xi_1, \xi_2) \right) \right\} d(\xi_1, \xi_2) \\
&= \int_{\tilde{F}} \left\{ D_1 f_2(\xi_1, \xi_2, h(\xi_1, \xi_2)) + D_3 f_2(\xi_1, \xi_2, h(\xi_1, \xi_2)) D_1 h(\xi_1, \xi_2) \right. \\
&\quad + D_1 f_3(\xi_1, \xi_2, h(\xi_1, \xi_2)) D_2 h(\xi_1, \xi_2) \\
&\quad + D_3 f_3(\xi_1, \xi_2, h(\xi_1, \xi_2)) D_1 h(\xi_1, \xi_2) D_2 h(\xi_1, \xi_2) \quad (*) \\
&\quad + f_3(\xi_1, \xi_2, h(\xi_1, \xi_2)) D_1 D_2 h(\xi_1, \xi_2) \quad (\dagger) \\
&\quad - D_2 f_1(\xi_1, \xi_2, h(\xi_1, \xi_2)) - D_3 f_1(\xi_1, \xi_2, h(\xi_1, \xi_2)) D_2 h(\xi_1, \xi_2) \\
&\quad - D_2 f_3(\xi_1, \xi_2, h(\xi_1, \xi_2)) D_1 h(\xi_1, \xi_2) \\
&\quad - D_3 f_3(\xi_1, \xi_2, h(\xi_1, \xi_2)) D_2 h(\xi_1, \xi_2) D_1 h(\xi_1, \xi_2) \quad (*) \\
&\quad \left. - f_3(\xi_1, \xi_2, h(\xi_1, \xi_2)) D_2 D_1 h(\xi_1, \xi_2) \right\} d(\xi_1, \xi_2) \quad (\dagger) \\
&\quad \text{(die Zeilen mit (*) bzw. (\dagger) heben sich weg)} \\
&= \int_{\tilde{F}} \left\{ \left(D_1 f_2(\xi_1, \xi_2, h(\xi_1, \xi_2)) - D_2 f_1(\xi_1, \xi_2, h(\xi_1, \xi_2)) \right) \cdot 1 \right. \\
&\quad - \left(D_2 f_3(\xi_1, \xi_2, h(\xi_1, \xi_2)) - D_3 f_2(\xi_1, \xi_2, h(\xi_1, \xi_2)) \right) D_1 h(\xi_1, \xi_2) \\
&\quad \left. - \left(D_3 f_1(\xi_1, \xi_2, h(\xi_1, \xi_2)) - D_1 f_3(\xi_1, \xi_2, h(\xi_1, \xi_2)) \right) D_2 h(\xi_1, \xi_2) \right\} d(\xi_1, \xi_2) \\
&= \int_{\tilde{F}} \left\langle \operatorname{rot} f(\xi_1, \xi_2, h(\xi_1, \xi_2)), \left(-D_1 h(\xi_1, \xi_2), -D_2 h(\xi_1, \xi_2), 1 \right) \right\rangle d(\xi_1, \xi_2) \\
&\quad \text{(wegen } M = \{x \in \mathbb{R}^3 : g(x) := \xi_3 - h(\xi_1, \xi_2) = 0\} \text{ ist)}
\end{aligned}$$

$$\begin{aligned}
& (-D_1 h(\xi_1, \xi_2), -D_2 h(\xi_1, \xi_2), 1) = \sqrt{1 + |\operatorname{grad} h(\xi_1, \xi_2)|^2} n(\xi_1, \xi_2, h(\xi_1, \xi_2)) \\
& = \int_{\tilde{F}} \left\langle \operatorname{rot} f(\xi_1, \xi_2, h(\xi_1, \xi_2)), n(\xi_1, \xi_2, h(\xi_1, \xi_2)) \right\rangle \sqrt{1 + |\operatorname{grad} h(\xi_1, \xi_2)|^2} d(\xi_1, \xi_2) \\
& = \int_F \left\langle \operatorname{rot} f(x), n(x) \right\rangle do(x),
\end{aligned}$$

womit der Satz bewiesen ist. ■

Bemerkung 13.9 Das Kurvenintegral über die (geschlossene) Randkurve Γ von F liefert die Arbeit die geleistet wird, wenn sich ein Körper unter Einwirkung des Kraftfeldes f längs Γ bewegt; man nennt dieses Integral die *Zirkulation* des Feldes f längs der geschlossenen Kurve Γ . Der Stokessche Satz besagt, dass die Zirkulation von f längs Γ gleich dem Fluß des Feldes $\operatorname{rot} f$ durch die Fläche F ist. Da der Stokessche Satz auch für alle Teilstücke von F entsprechend gilt, kann man $\operatorname{rot} f(\cdot)$ als die *Wirbeldichte* (Zirkulationsdichte) von f bezeichnen. Aus dem Satz ergibt sich insbesondere für $m = 3$ das bereits in Abschnitt 9 bewiesene Kriterium für die Wegunabhängigkeit des Kurvenintegrals.

Es fällt auf, dass das Integral über F nicht wirklich von der Fläche abhängt, sondern nur von der Kurve Γ in die sie „eingespannt“ ist. Das ist allerdings nicht weiter überraschend, denn wegen

$$\begin{aligned}
\operatorname{div} \operatorname{rot} f &= D_1(D_2 f_3 - D_3 f_2) + D_2(D_3 f_1 - D_1 f_3) + D_3(D_1 f_2 - D_2 f_1) \\
&= (D_1 D_2 f_3 - D_2 D_1 f_3) + (D_2 D_3 f_1 - D_3 D_2 f_1) + (D_3 D_1 f_2 - D_1 D_3 f_2) \\
&= 0
\end{aligned}$$

ist $\operatorname{rot} f$ divergenzfrei für jedes zweimal stetig differenzierbare Vektorfeld f .

Beispiel 13.10 Die *Induktionsgleichung* besagt (bis auf physikalische Konstanten)

$$\int_{\Gamma} E_1(x, t) d\xi_1 + E_2(x, t) d\xi_2 + E_3(x, t) d\xi_3 = -\frac{d}{dt} \int_F \langle H(x, t), n(x) \rangle do(x)$$

mit E = elektrische, H = magnetische Feldstärke.

Mit dem Satz von Stokes folgt daraus die *zweite Hauptgleichung* des elektromagnetischen Feldes

$$\operatorname{rot} E(x, t) = -\frac{d}{dt} H(x, t).$$

□

13.4 Übungsaufgaben

13.1 Sei S die Oberfläche der Kugel mit Radius R um den Nullpunkt in \mathbb{R}^3 , $n(x)$ die äußere Normale von S in x . Mit Hilfe des Divergenzsatzes berechne man

$$\int_S \left\langle \left(\xi_1^3, \xi_2^3, \xi_3^3 \right), n(x) \right\rangle do(x),$$

Ergebnis: $\frac{12}{5} \pi R^5$.

13.2 Mit Hilfe des Stokes'schen Satzes berechne man das Kurvenintegral von

$$k(x) = \left(2\xi_2, -2\xi_1, 2\xi_1\xi_3^2 \right)$$

über die Kurve $\gamma : [0, 2\pi] \rightarrow \mathbb{R}^3, t \mapsto (\cos t, \sin t, 5)$.

Ergebnis: -8π .

13.3 1. Sei Ω ein Gebiet in der Ebene, γ die Randkurve von Ω (so, dass Ω links von γ liegt), $\gamma : [a, b] \rightarrow \mathbb{R}^2$. Dann ist die Fläche $F(\Omega)$ des Gebietes Ω gleich der Hälfte des Kurvenintegrals des Vektorfeldes $k(x, y) = (-y, x)$ über γ , d. h.

$$F(\Omega) = \frac{1}{2} \int_{\gamma} (-\xi_2, \xi_1) dx = \frac{1}{2} \int_a^b \langle \gamma(t), \gamma'(t) \rangle dt.$$

2. Man berechne die Fläche unter einer Periode der Zykloide (vgl. § ?? der Vorlesung).

Ergebnis: 3π .